

**Department für Informatik**  
Abteilung Informationssysteme

# Projektdokument

Business Intelligence in the Cloud for Energy



**BICE**

Projekt:	BICE
Auftraggeber:	Marco Haas, Oliver Norkus
Auftragnehmer:	Cloud-BI
Autoren:	Projektgruppe Cloud-BI
Datum:	4. Oktober 2015

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Projektorganisation</b>	<b>2</b>
2.1	Projektmanagement . . . . .	2
2.2	Rollen . . . . .	2
2.3	Vorgehensweise . . . . .	5
2.4	Meta-Kommunikation . . . . .	6
2.4.1	Kommunikations-Radar . . . . .	6
2.4.2	Einsatz . . . . .	6
2.4.3	Auswertung . . . . .	8
2.4.4	Konsequenzen . . . . .	8
<b>3</b>	<b>Anforderungen</b>	<b>9</b>
3.1	Zielsetzung . . . . .	9
3.2	Anforderungserhebung . . . . .	9
3.2.1	Produkteinsatz . . . . .	9
3.2.2	Deckungsbeitragsrechnung . . . . .	9
3.2.3	Gesamtkostenrechnung . . . . .	12
3.2.4	Kennzahlen . . . . .	13
3.2.5	Use Cases . . . . .	13
3.3	Funktionale Anforderungen . . . . .	18
3.4	Nichtfunktionale Anforderungen . . . . .	20
3.5	System-Anforderungen . . . . .	22
3.5.1	Musskriterien . . . . .	22
3.5.2	Wunschkriterien . . . . .	23
3.5.3	Abgrenzungskriterien . . . . .	24
<b>4</b>	<b>Entwurf</b>	<b>26</b>
4.1	Überblick . . . . .	26
4.2	Komponenten . . . . .	27
4.2.1	Local Client . . . . .	27
4.2.2	User Management . . . . .	27
4.2.3	Load Management & Balancing . . . . .	28
4.2.4	Data Retrieval und Accounting . . . . .	28
4.2.5	Datenbanksystem und ETL . . . . .	29
4.3	Kennzahlen . . . . .	31
4.3.1	Gesamtkosten . . . . .	32

4.3.2	Deckungsbeitrag . . . . .	34
4.4	Betriebsbedingungen . . . . .	34
<b>5</b>	<b>Implementation</b>	<b>36</b>
5.1	Infrastrukturebene . . . . .	36
5.1.1	Load Balancing & Autoscaling . . . . .	36
5.2	Plattformebene . . . . .	38
5.3	Softwareebene . . . . .	39
5.3.1	Local Client . . . . .	39
5.3.2	Dispatcher & Session Management . . . . .	42
5.3.3	Authentication Management . . . . .	43
5.3.4	Data Retrieval . . . . .	44
5.3.5	Accounting . . . . .	45
<b>6</b>	<b>Evaluation</b>	<b>46</b>
6.1	Leistungsspektrum . . . . .	46
6.2	Qualität der Software . . . . .	46
6.2.1	Quelltext-Ebene . . . . .	46
6.2.2	Komponenten-Ebene . . . . .	47
6.2.3	System-Ebene . . . . .	52
6.3	Rahmenbedingungen . . . . .	52
<b>7</b>	<b>Zusammenfassung und Ausblick</b>	<b>55</b>
7.1	Zusammenfassung . . . . .	55
7.2	Ausblick . . . . .	55
7.2.1	Customer Relationship Management . . . . .	56
7.2.2	ene't Datenbanken . . . . .	57
7.2.3	Energiedaten-Management . . . . .	58
7.2.4	Kundendaten-Datenbank . . . . .	58
<b>8</b>	<b>Glossar</b>	<b>60</b>
<b>9</b>	<b>Literatur</b>	<b>63</b>
<b>A</b>	<b>Benutzerhandbuch</b>	<b>64</b>
A.1	Login . . . . .	64
A.2	Calculation-Page . . . . .	64
A.3	Accounting . . . . .	65
A.4	Logout . . . . .	66
<b>B</b>	<b>Seminararbeiten</b>	<b>68</b>
<b>C</b>	<b>Protokolle</b>	<b>152</b>

# Abbildungsverzeichnis

2.1	Organigramm . . . . .	3
2.2	Kommunikations-Radar . . . . .	7
3.1	Bezeichnung der eEPK-Symbole . . . . .	10
3.2	eEPK Deckungsbeitragsrechnung . . . . .	11
3.3	eEPK Gesamtkostenrechnung . . . . .	12
3.4	Übersicht über die Use Cases . . . . .	14
4.1	Übersicht über das BICE System . . . . .	26
4.2	Local Client Komponentendiagramm . . . . .	27
4.3	User- und Load Management und Balancing Komponentendiagramm . . . . .	29
4.4	Struktur der Daten . . . . .	30
4.5	Data Retrieval, Accounting, Database System und ETL Komponentendiagramm . . . . .	31
4.6	MER-Diagramm Gesamtkosten . . . . .	33
4.7	MER-Diagramm Deckungsbeitrag . . . . .	35
6.1	Graphik Performanztest . . . . .	50
6.2	Graphik Referenzwerte . . . . .	51
A.1	Login-Page . . . . .	64
A.2	Calculation-Seite . . . . .	64
A.3	Customer-Dropdown . . . . .	65
A.4	Customer Site Dropdown . . . . .	65
A.5	Datepicker . . . . .	66
A.6	Graph . . . . .	67
A.7	Tabelle . . . . .	67
A.8	Accounting-Seite . . . . .	67

# Tabellenverzeichnis

2.1	Rollenübersicht . . . . .	3
2.2	Auswertung Kommunikations-Radar . . . . .	8
3.1	Beschreibung Deckungsbeitragsrechnung . . . . .	13
3.2	Kurzbeschreibung aller Use Cases . . . . .	15
3.3	Funktionale Anforderungen . . . . .	18
3.4	Funktionale Anforderungen (Fortsetzung) . . . . .	19
3.5	Nichtfunktionale Anforderungen . . . . .	20
3.6	Nichtfunktionale Anforderungen (Fortsetzung) . . . . .	21
3.7	Musskriterien funktionale Anforderungen Musskriterien . . . . .	22
3.8	Nichtfunktionale Anforderungen der Musskriterien . . . . .	23
3.9	Funktionale Anforderungen der Wunschkriterien . . . . .	23
3.10	Nichtfunktionale Anforderungen der Wunschkriterien . . . . .	24
4.1	Kennzahlensteckbrief Gesamtkosten . . . . .	32
4.2	Kennzahlensteckbrief Deckungsbeitrag . . . . .	34
6.1	Ergebnisse eines Blackboxtests . . . . .	47
6.2	Ergebnisse eines Blackboxtests . . . . .	48
6.3	Empirischer Mittelwert und empirische Varianz . . . . .	49
6.4	Ergebnisse der Anpassungstest . . . . .	50

# 1 Einleitung

Im Kontext der Globalisierung und der letzten Finanzkrisen zeichnet sich ab, dass Unternehmen sich an eine Welt anpassen müssen, die sich rapide verändert und von unsicheren und unvollständigen Informationen geprägt ist, um weiter im Wettbewerb zu bestehen. Sie müssen schnell auf Veränderungen und Trends reagieren können und dementsprechend agil und flexibel müssen auch die Entscheidungsprozesse gestaltet sein. Aus diesem Grund steigen auch die Anforderungen an die verwendeten IT-Werkzeuge. Es werden proaktive und integrierte Systeme zur Entscheidungsunterstützung gebraucht, die hochverfügbar und erreichbar sind.[NS14]

*Business Intelligence in the Cloud for Energy (BICE)* soll eine Realisierung eines solchen Systems werden, das einen konkreten Anwendungsfall aus der Energiewirtschaft bedient. Dabei handelt es sich um die Deckungsbeitragsrechnung, die die Beschaffung, Aggregation und Analyse relevanter Daten umfasst. Zu diesen Daten gehören unter anderem Leistungs-, Arbeits- und Netznutzungskosten sowie einige Kundendaten. Mit den aktuellen Mitteln dauert die Deckungsbeitragsrechnung jedoch zu lange, um die notwendige Agilität und Flexibilität in der Entscheidungsfindung zu gewährleisten. Das System BICE soll die Komplexität der Deckungsbeitragsrechnung verringern. Derzeit sind für die Deckungsbeitragsrechnung mehrere Prozesse notwendig und die Daten müssen aus verschiedenen Quellen geladen werden. Die Deckungsbeitragsrechnung soll als ein Prozess und als Service verfügbar gemacht werden.

Dieses Dokument stellt die finale Projektdokumentation des BICE-Projekts dar. Dabei wird detailliert auf die allgemeine Projektorganisation, die Anforderungserhebung, den konzeptionellen Entwurf, die technische Realisierung (Implementation), das Testing sowie die anschließende Evaluation des BICE-Systems eingegangen. Weiter enthält dieses Dokument im Anhang die Seminararbeiten der Projektgruppenmitglieder sowie sämtliche verfasste Protokolle während der Projektphase.

## 2 Projektorganisation

In diesem Kapitel werden die Projektmanagement-Methoden sowie die sich daraus ergebenden Rollen vorgestellt, die von den Mitgliedern der Projektgruppe umgesetzt, bzw. bekleidet wurden. Daraus lassen sich Erkenntnisse schließen, wie ein solches Projekt in der Wirtschaft umgesetzt werden könnte.

### 2.1 Projektmanagement

Für die Umsetzung des Projektes wurde die Projektmanagement-Methode Scrum verwendet. Scrum ist ein iteratives Projektmanagementverfahren, welches eine enge Bindung zwischen dem Entwicklerteam und dem Kunden schaffen soll. Der iterative Charakter fördert darüber hinaus die Kommunikation innerhalb des Entwicklerteams sowie die Möglichkeit, schnell auf auftretende Probleme zu reagieren. Eine detaillierte Beschreibung dieser Projektmanagementmethode kann in der Seminararbeit über Scrum nachgelesen werden, die sich im Anhang des Projektdokuments befindet.

### 2.2 Rollen

Für die Bearbeitung des Projektes wurden einzelne Rollen identifiziert, die von den Teilnehmern der Projektgruppe zur Realisierung des Projektes eingenommen wurden. Scrum sieht die Rollen Kunde, Entwicklungsteam und Scrum Master vor. Die Rolle des Kunden übernehmen im Fall des BICE-Projekts der Betreuer der Projektgruppe Oliver Norkus und der Sparingspartner aus der Energiewirtschaft Marco Haas von Ceyoniq Consulting GmbH. Die studentischen Mitglieder der Projektgruppe bilden das Entwicklungsteam. Der Scrum Master stammt ebenfalls aus dem Entwicklungsteam.

Neben den Hauptrollen, welche durch Scrum vorgegeben werden, wurden darüber hinaus einzelne Rollen innerhalb des Entwicklerteams identifiziert. Zu diesen Rollen zählen der Projektleiter, der stellv. Projektleiter, der Architektur Manager, der Entwickler, der Server Administrator, der Qualitätsmanager, der Homepage Beauftragte und der Dokumentenbeauftragte. Das zugehörige Diagramm ist in der Abbildung 2.1 zu sehen. Die Aufteilung der Rollen ist in Tabelle 2.1 zu sehen. Im Folgenden werden die einzelnen Rollen kurz beschrieben.

**Projektleiter u. Scrum Master:** Die Rollen des Projektleiters und des Scrum Masters werden im Falle der Projektgruppe von einem einzelnen Studenten übernommen, da sowohl der Projektleiter, als auch der Scrum Master ähnliche Aufgabengebiete vereinen. Dem Projektleiter obliegt die operative Planung und Steuerung des Projekts. Der

## 2 Projektorganisation

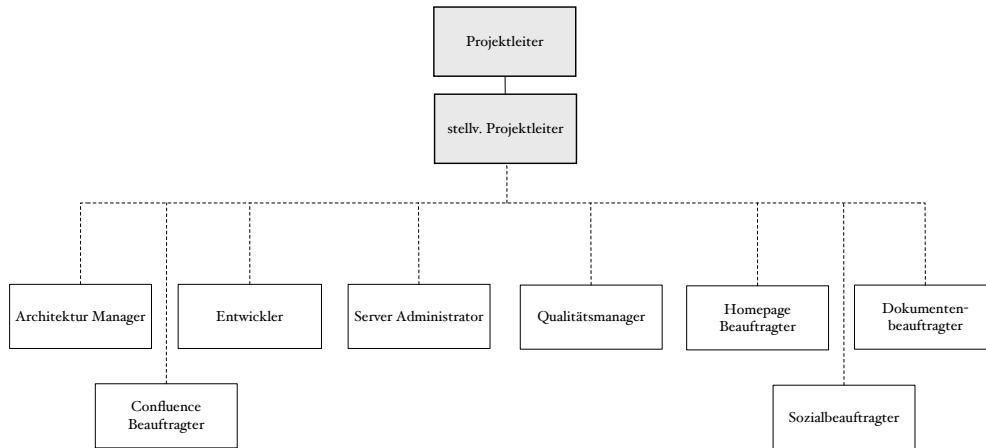


Abbildung 2.1: Das Organigramm zur Rollenverteilung in der Projektgruppe.

Person	Rolle
Clark, Brian	Entwickler
Friedrich Björn	Entwickler, Homepage Beauftragter
Izadpanah, Babak	Entwickler, Architektur Manager, Sozialbeauftragter
Merkel, Florian	Entwickler, Server Administrator
Schweer, Ilias	Entwickler, stellv. Projektleiter, Qualitätsmanager
Zimak, Alexander	Entwickler, Projektleiter, Confluence Beauftragter

Tabelle 2.1: Übersicht über die Teammitglieder und denen von ihnen übernommenen Rollen

Scrum Master ist für den reibungslosen Ablauf des Scrum verantwortlich. Er soll sich, wie der Projektleiter auch, um persönliche Konflikte im Team, die Zusammenarbeit und die Kommunikation zwischen den Teammitgliedern kümmern. Außerdem ist er auch, wie der Projektleiter, für Projektorganisation zuständig. Im Normalfall hat der Projektleiter weiterreichende Befugnisse als der Scrum Master. Er ist dem Team gegenüber weisungsbefugt und ist verantwortlich für die Ressourcenplanung. Der Scrum Master ist nicht weisungsbefugt und die Ressourcenplanung wird vom Team selbst durchgeführt.

**stellv. Projektleiter:** Der stellv. Projektleiter übernimmt die Aufgaben, des Projektleiters, falls dieser verhindert ist. Sollte dies der Fall sein, hat der stellv. Projektleiter dieselben Befugnisse und Entscheidungsspielräume, wie der Projektleiter.

**Entwickler:** Jedes Teammitglied hat die Rolle des Entwicklers inne. Der Entwickler arbeitet an der Erstellung der Bestandteile zur Realisierung des Produkts. Dies umfasst alle zugehörigen Aufgaben wie Dokumentation, Programmierung, Wartung, Administration ect. Hinzu kommen die Aufgaben, die das Team im Scrum Prozess übernehmen soll. Dazu zählen Aufwandsabschätzungen, Product Backlog Refinement und die Sprint



## 2 Projektorganisation

Planung.

**Dokumentenbeauftragter:** Damit die Dokumente eine einheitliche Form haben, erstellt der Dokumentenbeauftragte Vorlagen. Insbesondere bei häufig wiederkehrenden Dokumenten, wie bspw. Protokolle, ist eine einheitliche Form wichtig. Die umfassenderen Dokumente, wie das Anforderungsdokument, das Pflichtenheft und der Abschlussbericht müssen strukturiert werden. Die Anfertigung des Grundgerüsts übernimmt der Dokumentenbeauftragte.

**Architektur Manager:** Der Architektur Manager hat eine administrative Aufgabe. Die Entwicklung und Dokumentation der Architektur wird vom gesamten Team gemeinsam übernommen. Der Architektur Manager sorgt für den strukturierten Aufbau der Architektur und ist erster Ansprechpartner bei Fragen bezüglich der Architektur. Darüber hinaus ist er hauptverantwortlich für alle konzeptionellen und praktischen Arbeiten an der Architektur.

**Confluence Beauftragter:** Für die Kollaboration haben wird während des Projektes Confluence eingesetzt. Damit die gemeinschaftlich gesammelten Erfahrungen und das gemeinschaftlich gesammelte Wissen abgerufen werden können, muss es strukturiert dokumentiert werden. Der Confluence Beauftragte sorgt dafür, dass alle Dokumente katalogisiert werden und im Confluence zum Abruf bereitstehen. Weiterhin organisiert er die Struktur im Confluence und sorgt dafür, dass sie eingehalten wird.

**Homepage Beauftragter:** Der Homepage Beauftragte kümmert sich um die Homepage der Projektgruppe ([www.cloud-bi.informatik.uni-oldenburg.de](http://www.cloud-bi.informatik.uni-oldenburg.de)). Seine Aufgabe ist es, die Homepage zu erstellen, zu hosten und zu pflegen. Außerdem obliegt ihm die Aufgabe sich um alle anfallenden administrativen Aufgaben bzgl. der Homepage zu kümmern.

**Qualitätsmanager:** Für alle Dokumente wurde eine Stufe für die Qualitätssicherung eingebaut. Der Qualitätsmanager soll für einen Standard bei den Dokumenten sorgen. Die Stufe der Qualitätssicherung wird in allen Prozessen, die Dokumente beinhalten, berücksichtigt.

**Server Administrator:** Für die Kollaboration werden verschiedene Tools (JIRA, Confluence, SVN) eingesetzt. Diese müssen auf Servern installiert und administriert werden. Diese Aufgabe übernimmt der Server Administrator. Weiterhin pflegt er die Systeme und steht den Mitgliedern des Teams bei Fragen und Problemen bzgl. der Systeme zur Seite.

**Sozialbeauftragter:** Neben der fachlichen Kompetenz im Team ist die Zusammenarbeit und die Stimmung von essentieller Bedeutung. Der Sozialbeauftragte organisiert Team-Events, um die Kommunikation im Team zu fördern und den Zusammenhalt zu stärken.

## 2.3 Vorgehensweise

Es wurde beschlossen, dass während der Bearbeitung des Projektes eine bestimmte Abfolge von Arbeitsschritten einzuhalten ist. Da Scrum als Projektmanagement- bzw. Softwareentwicklungsmethode verwendet wird, wurden die einzelnen Prozesse und Phasen für die Bedürfnisse der Projektgruppe angepasst. Dabei wurden die Rollen nicht in dem Ausmaß verfolgt, wie es Scrum eigentlich vorsieht. So war der Produkt Owner zum Teil stark in die Entwicklung des Systems involviert, gleiches gilt für den Scrum Master. Für ein genaues Verständnis der Scrum Methodik steht die Seminararbeit über Scrum zur Verfügung, welche im Anhang zu finden ist.

- Product Backlog: Wurde von der Projektgruppe, in Teamarbeit, regelmäßig aktualisiert und erweitert.
- Sprint Planning: Wurde jede Woche von der Projektgruppe durchgeführt.
- Sprint Backlog: Wurde während des Sprint Planning erstellt.
- Daily Scrum: Aufgrund des Arbeitsumfeldes einer studentischen Projektgruppe wurde auf das Daily Scrum verzichtet. Der Scrum Master konnte bei Konflikten jedoch jeder Zeit kontaktiert werden.
- Sprint Execution: Die Dauer eines Sprints wurde auf eine Woche festgelegt.
- Sprint Review: Der Sprint Review wurde einmal wöchentlich, parallel mit dem Sprint Planning und der Sprint Retrospective, im Beisein des Product Owners durchgeführt.
- Sprint Retrospective: Die Retrospective wurde einmal wöchentlich durchgeführt.

Neben dem Vorgehen bezüglich der Scrum Methodik hat sich die Projektgruppe auch bezüglich der Anforderungserhebung mittels Interviews auf ein Vorgehensmodell geeinigt. Das Vorgehen wird in den folgenden Schritten iterativ durchgeführt.

1. Fragenkatalog für das Interview erstellen.
2. Interview mit dem Kunden aus der Energiewirtschaft führen, aufzeichnen und ein Protokoll anfertigen.
3. Erkenntnisse aus dem Interview auswerten und Ermitteln von Anforderungen aus diesen.
4. Präsentation der Ergebnisse.

## 2.4 Meta-Kommunikation

Neben der fachlichen Arbeit gehört zur Teamarbeit eine gute Zusammenarbeit. Dafür ist die Kommunikation im Team ein zentraler Bestandteil. Wird im Team nicht richtig kommuniziert, kann das zu Konflikten führen. Gibt es Konflikte im Team, wird in die Bereinigung dieser viel Energie investiert, die besser in die Arbeit hätte fließen können. Die Projektgruppe hat einige allgemeine Regeln für Kommunikation und Feedback festgelegt. Zur Analyse und Förderung der Kommunikation im Team wurde ein Kommunikations-Radar nach Peter Schlötter verwendet. Neben dem Kommunikations-Radar gibt es den Kommunikations-Tacho. Dieser ist jedoch nicht so detailliert wie das Kommunikations-Radar und trifft eine allgemeine Aussage über die Stimmung im Team. Zur Förderung der Kommunikation im Team gibt es auch Events. Das Team hat bspw. am Boule-Turnier der Projektgruppen teilgenommen und hat einen Abend gemeinsam gebowled. Bei der Verwendung ist zu beachten, dass es sich um subjektive Momentaufnahmen handelt und nicht um eine objektive Messung.

### 2.4.1 Kommunikations-Radar

Das Kommunikations-Radar soll das Bewusstsein des Teams auf die Kommunikation lenken. Durch das Ausfüllen tritt die Person einen Schritt zurück und betrachtet die Kommunikation von der Position eines Beobachters aus. Die Zahlenwerte sorgen dafür, dass die Person intensiver über den Aspekt nachdenkt und sorgfältig abwägt, wie viele Punkte sie geben soll. In der Abbildung 2.2 ist das Kommunikations-Radar, das verwendet wurde, zu sehen.

Für die Skalierung haben wir Einser Schritte gewählt und uns entschlossen, auch negative Werte zuzulassen. Durch die negativen Werte ist intuitiv klar, welche bei Missfallen gewählt werden müssen. Die Skala reicht von -3 bis +3, wobei -3 die größte Unzufriedenheit repräsentiert und +3 die maximale Zufriedenheit.

### 2.4.2 Einsatz

Die Bedeutung der einzelnen Punkte und warum sie gewählt wurden, soll im Folgenden erklärt werden.

**Kooperation:** Zur Kooperation gehört die Ausrichtung der Strategie des Teams. Dieser Punkt soll dazu anregen über diese Ausrichtung nachzudenken und zu überprüfen, ob die Prioritäten der Teammitglieder gleich gesetzt sind und ob Zielkonflikte vorliegen. Es kann natürlich vorkommen, dass es widersprüchliche Prioritäten gibt. Manchmal ist dieser Konflikt jedoch hilfreich, bspw. bei Qualität und Kosten. Neben dieser strategischen Kooperation gibt es noch die formale Kooperation.

Die formale Kooperation beschreibt die Praxis der Hierarchie und der Organisation. In der Projektgruppe sind alle gleichberechtigt, trotzdem wird ein Projektleiter gewählt, der eine Autorität darstellt. Gerade durch diese zwiespältige Situation kann es zu Problemen kommen. Geht man bei Problemen direkt zum Betreuer oder zum Projektleiter,

## 2 Projektorganisation

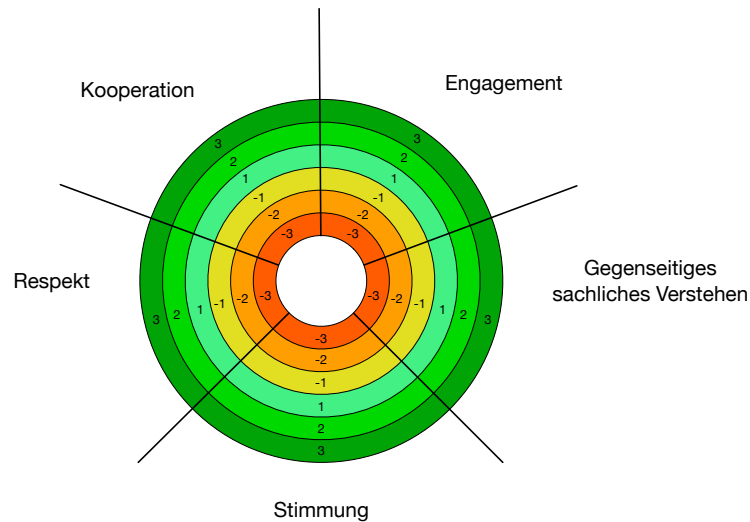


Abbildung 2.2: Das von der Projektgruppe eingesetzte Kommunikations-Radar mit den verschiedenen Ausprägungen und der Skalierung.

können sich Konflikte entwickeln, die den Anlass der Konflikte übersteigen. Meistens wäre eine interne Klärung im Team völlig ausreichend gewesen.

Um diesen Konflikten vorzubeugen, ist das Kriterium ins Kommunikations-Radar aufgenommen worden.

**Engagement:** Engagement soll die wahrgenommene Aktivität des Teams abbilden. Die Einschätzung dieses Aspekts ist schwierig. Häufig nimmt man die Arbeit der anderen Teammitglieder nicht wahr, weil sie alleine stattfindet. Außerdem lässt sich an den Ergebnissen nicht immer ablesen, wie viel Arbeit und Zeit investiert wurden. Ein weiterer Punkt, der zu dem Aspekt Engagement gehört, ist die Bereitschaft Aufgaben und Verantwortung zu übernehmen.

**Gegenseitiges sachliches Verstehen:** Dieser Punkt ist wegen der Zusammensetzung der Projektgruppe besonders wichtig. Neben Vertretern aus dem Bereich Wirtschaftsinformatik und Informatik gab es zwei Mitglieder, bei denen Deutsch nicht die Muttersprache ist. Neben der Sprache kann es auch durch die Denkweisen und den Bildungsstand zu Problemen kommen. Die genannten Ursachen lassen sich zwar nicht einfach beseitigen, aber das Wissen hilft, um diese Konflikte vorzubeugen.

**Stimmung:** Dieser Punkt ist einer der allgemeinsten Aspekte. Allerdings auch einer der wichtigsten. Die Stimmung trägt dazu bei, dass das Team in einer produktiven Atmosphäre arbeiten kann und zuversichtlich ist, gesteckte Ziele auch zu realisieren.

## 2 Projektorganisation

**Respekt:** Respekt ist ein Begriff der häufig sehr unterschiedlich verstanden wird. Als dieser Punkt in das Kommunikations-Radar aufgenommen wurde, hat sich das Team auf die folgende Definition verständigt:

*Respekt ist die auf Anerkennung, Bewunderung beruhende Achtung einer anderen Person.*

Weiterhin bedeutet für uns Respekt auch, dass man über kleinere Fehler und Eigenheiten der anderen Teammitglieder hinwegsieht.

### 2.4.3 Auswertung

Die Auswertung erfolgt wie das Ausfüllen anonym. Die ausgefüllten Kommunikations-Radare wurden an den Betreuer der Projektgruppe gegeben. Dieser hat eine E-Mail mit dem Ergebnis verfasst. In der Tabelle 2.2 ist die Auswertung für das Kommunikations-Radar vom 02.07.2015 zu sehen. Für jeden Aspekt wurde die Gesamtpunktzahl sowie der maximale und der minimale Wert erfasst. Dieses Vorgehen sorgt in der Gruppe dafür, dass das Zugehörigkeitsgefühl gestärkt wird. Das Ergebnis eines jeden Mitglieds zählt zu dem Gesamtergebnis dazu und ist dennoch anonym. Weiterhin wird die Diskussion von den vergebenen Punkten abgelenkt und es wird über die Sachverhalte an sich und mögliche Lösungswege gesprochen. Der Einzelne tritt somit in den Hintergrund.

Aspekt/Werte	Minimum	Maximum	Durchschnitt
Kooperation	1	2	1,83
Engagement	1	2	1,16
gegenseitiges sachl. Verständnis	-1	2	1,33
Stimmung	1	3	2,33
Respekt	1	3	2,16

Tabelle 2.2: Die Auswertung des Kommunikations-Radars vom 02.07.15.

### 2.4.4 Konsequenzen

Während des SCRUM-Review wurden auch die Ergebnisse des Kommunikations-Radars besprochen und über Maßnahmen diskutiert. Diese Maßnahmen wurden dann direkt umgesetzt.

Die Auswertung, ob die Maßnahmen erfolgreich waren, geschieht erst zwei Wochen später. In dieser Zeit sollten die Maßnahmen erste Verbesserungen zeigen.

# 3 Anforderungen

Dieses Kapitel beschäftigt sich mit der Anforderungserhebung des BICE-Projektes.

## 3.1 Zielsetzung

Das System BICE soll die Komplexität der Deckungsbeitragsrechnung verringern. Derzeit sind für die Deckungsbeitragsrechnung mehrere Prozesse notwendig und die Daten müssen aus verschiedenen (heterogenen) Quellen extrahiert und transformiert werden. Der Prozess der Deckungsbeitragsrechnung soll daher als Service verfügbar gemacht werden. Idealerweise soll der Service via Internet bereitgestellt werden, um ihn so für sämtliche mobile Endgeräte mit einer Internetverbindung verfügbar zu machen. Aus dieser Vision eines Systems zur Deckungsbeitragsrechnung, welche als Service angeboten werden kann, ergeben sich entsprechende Anforderungen für die Entwicklung. Der folgende Abschnitt beschäftigt sich mit den fachlichen und technischen Anforderungen, die sich an ein solches System stellen.

## 3.2 Anforderungserhebung

In diesem Abschnitt wird näher auf die Anforderungserhebung und den Business Case des BICE Systems eingegangen. Als Grundlage der Anforderungserhebung wurden Experten-Interviews mit einem Mitarbeiter der Ceyoniq Consulting GmbH durchgeführt. Aus diesen konnten die nötigen Erkenntnisse bezüglich des zu realisierenden Business Cases ermittelt werden.

### 3.2.1 Produkteinsatz

Das BICE System befasst sich mit zwei verschiedenen Business Cases: Zum einen die Berechnung des Deckungsbeitrags und zum anderen die Berechnung der Gesamtkosten. Beide Business Cases werden jeweils in den folgenden Abschnitten beschrieben und sind dazu unterstützend als erweiterte Ereignisgesteuerte Prozesskette (eEPK) dargestellt (Abbildung 3.2 und Abbildung 3.3).

### 3.2.2 Deckungsbeitragsrechnung

Die Vorbedingung für diesen Business Case lautet, dass ein Kunde bereits mit Energie versorgt wird und nun, in Anbetracht eines auslaufenden Vertragsverhältnisses, die Erstellung eines neuen Angebots vom Energielieferanten geprüft werden soll. Der Business Case Deckungsbeitrag stellt folglich einen Bestandskundenprozess dar. Der Lieferant

### 3 Anforderungen

berechnet mithilfe des Cloud-BI-Systems den Deckungsbeitrag. Die dafür notwendigen Kundendaten sind bereits im System hinterlegt. Als Output dieser Funktion entsteht der Deckungsbeitrag. Dieser geht als Input in die nächste Funktion *Deckungsbeitrag analysieren* hinein, die ebenfalls mithilfe des Cloud-BI-Systems umgesetzt wird. Sollte der Deckungsbeitrag auch für einen theoretischen neuen Vertrag genügen, kann der ursprüngliche Vertrag verlängert werden. Die andere Option lautet, dass Bedarf für ein neues Angebot besteht. Hierbei wird nun vom Lieferanten mithilfe des Cloud-BI-Systems und des ursprünglichen Deckungsbeitrags als Input ein neues Angebot (Output) erstellt. Danach wird das Angebot dem Kunden zur Verfügung gestellt und dieser entscheidet darüber. Die im Anschluss daran möglichen Ereignisse lauten: Der Kunde nimmt das neue Angebot an, der Kunde lehnt das Angebot endgültig ab oder der Kunde fordert ein neues/besseres Angebot. Letzteres führt wieder dazu, dass vom Lieferanten ein Angebot erstellt werden muss. Siehe hierzu Abbildung 3.2 auf der nächsten Seite. Die dabei verwendeten Symbole für die eEPK-Diagramme werden in Abbildung 3.1 vorgestellt.



Abbildung 3.1: Bezeichnung der eEPK-Symbole

### 3 Anforderungen

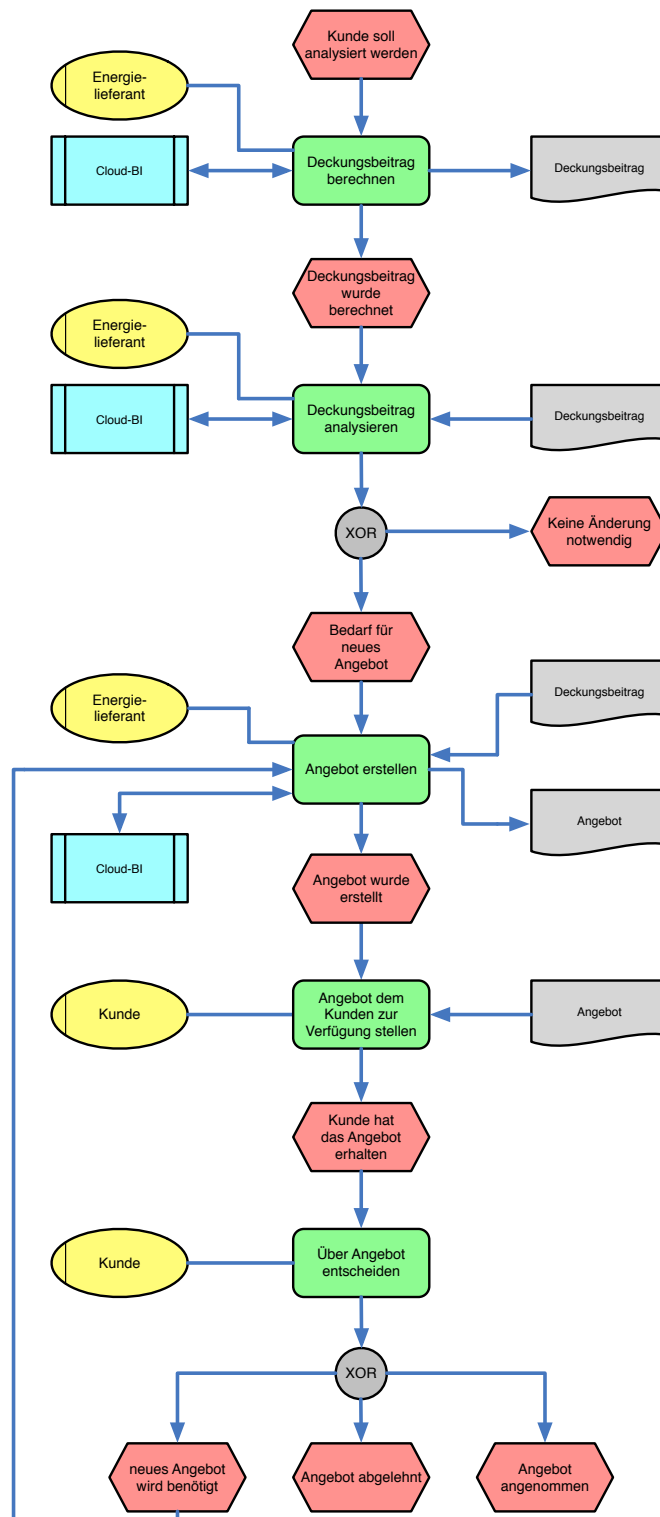


Abbildung 3.2: eEPK Deckungsbeitragsrechnung



### 3.2.3 Gesamtkostenrechnung

Dieser Business Case stellt einen Neukundenprozess dar. Dem potenziellen Neukunden soll ein Angebot unterbreitet werden. Erforderliche Kundendaten müssen zunächst in das System eingepflegt werden. Weiter lässt der Energielieferant mithilfe des Cloud-BI-Systems die Gesamtkosten berechnen. Die Gesamtkosten setzen sich zusammen aus dem Arbeitspreis, Leistungspreis und den Netznutzungskosten (Inputwerte) und stellen damit den Mindestpreis dar, der erzielt werden muss, um keine Verluste zu verzeichnen (in Ct/kWh). Mit den Gesamtkosten zzgl. der gewünschten Marge als Input wird dann vom Lieferanten mithilfe des Cloud-BI-Systems ein Angebot erstellt. Der restliche Prozess ist analog zum ersten Business Case.

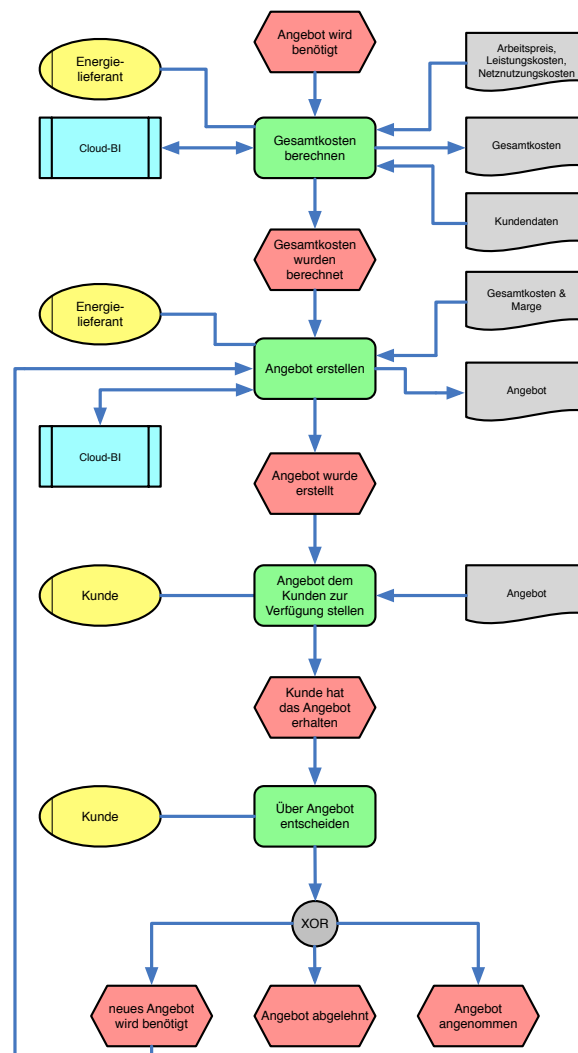


Abbildung 3.3: eEPK Gesamtkostenrechnung

### 3 Anforderungen

Deckungsbeitrag	Beschreibung
Deckungsbeitrag I	Differenz aus dem Erlös und den Arbeitskosten
Deckungsbeitrag II	Differenz aus Deckungsbeitrag I und den Leistungskosten
Deckungsbeitrag III	Differenz aus Deckungsbeitrag II und den Netznutzungskosten. Dieser Deckungsbeitrag entspricht dem Gewinn.

Tabelle 3.1: Beschreibung der Deckungsbeiträge

#### 3.2.4 Kennzahlen

Aus den beiden zuvor genannten Use Cases ergeben sich zwei Key Performance Indicator, die im BICE System berechnet und untersucht werden sollen. Diese sind die zuvor genannten *Gesamtkosten* pro Viertelstunde für die Versorgung eines potenziellen Kunden durch den Energieversorger und die verschiedenen Deckungsbeiträge pro Viertelstunde für die Versorgung eines existierenden Kunden des Energieversorgers. Die Gesamtkosten berechnen sich aus der Summe der Arbeits-, Leistungs- und Netznutzungskosten, die bei der Stromerzeugung und -Lieferung anfallen. Während die ersten beiden Anteile der Gesamtkosten von den für die Stromproduktion verwendeten Kraftwerken abhängen, sind die Netznutzungskosten von der Netzregion abhängig, in denen sich die Einspeisepunkte der Kraftwerke bzw. die Entnahmestellen der Kunden befinden. Diese Kennzahl wird dazu verwendet, um für den Kunden ein Lieferangebot zu erstellen. Bei der Deckungsbeitragsrechnung werden die vier Deckungsbeiträge berücksichtigt, die in Tabelle 3.1 beschrieben sind. Die Kennzahlen sollen über verschiedene Dimensionen und in unterschiedlicher Granularität aggregiert werden können. Die Dimensionen sind *Zeit*, *Kunde*, *Kraftwerk* und *Netzregion*. Zu diesem Zweck werden die Kennzahlen in einer Datenbank hinterlegt.

#### 3.2.5 Use Cases

Aus den beiden Use Cases Deckungsbeitragsrechnung und Gesamtkostenrechnung ergeben sich einige kleinere Use Cases, welche in diesem Abschnitt visualisiert und beschrieben werden. Eine Übersicht über alle Use Cases ist in Abbildung 3.4 zu sehen. Des Weiteren sind alle Use Cases in Tabelle 3.10 aufgelistet. Abschnitt 3.2.5 enthält die detaillierten Beschreibungen der einzelnen Use Cases.

##### Beschreibungen der Use Cases

In allen Use Cases wird der *Energielieferant* als Akteur genannt. Dieser ist Mitarbeiter oder Beauftragter eines Energieversorgungsunternehmens, welcher mit dem BICE-System über dessen Benutzeroberfläche interagiert.

**Use Case UC1 Angebot erstellen** Der Energielieferant erstellt mit Hilfe des BICE Systems ein Lieferangebot für seinen Kunden.

### 3 Anforderungen

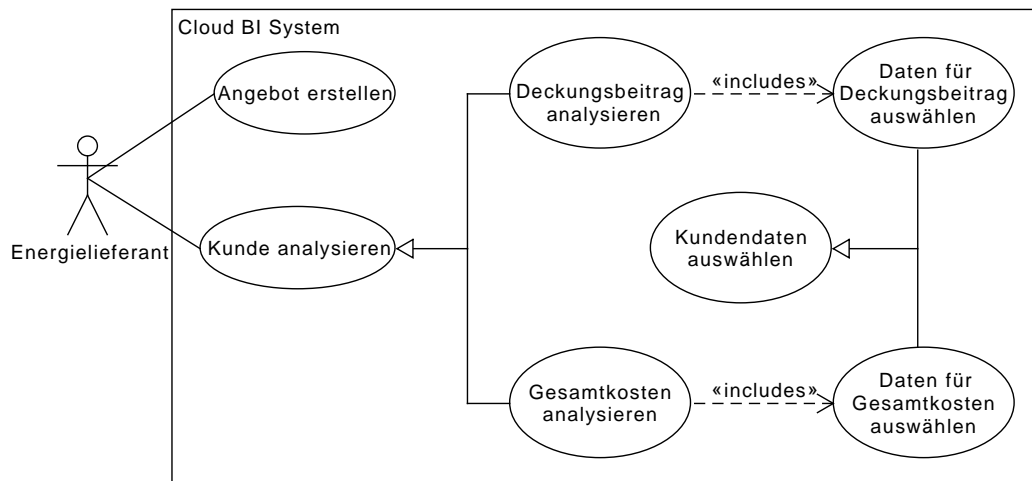


Abbildung 3.4: Übersicht über die Use Cases

#### Akteure:

- Energienlieferant

#### Vorbedingungen:

- Der angestrebte Profit liegt im System vor.
- Der Energienlieferant hat eine Gesamtkosten- oder Deckungsbeitragsanalyse mit dem System durchgeführt.

**Ablauf** Der Energienlieferant bekommt die im System hinterlegte angestrebte Handelsspanne angezeigt. Mit ihr und den aus der Analyse gewonnenen Informationen erstellt er ein Angebot für den Kunden.

#### Nachbedingungen:

- Der Energienlieferant hat ein Angebot erstellt.

#### Use Case UC2 Kunde analysieren

Der Energienlieferant analysiert die Daten eines Kunden.

#### Akteure:

- Energienlieferant

### 3 Anforderungen

ID	Titel	Kurzbeschreibung
UC1	Angebot erstellen	Der Energielieferant erstellt mit Hilfe des BICE Systems ein Lieferangebot für seinen Kunden.
UC2	Kunde analysieren	Der Energielieferant analysiert die Daten eines Kunden.
UC2a	Deckungsbeitrag analysieren	Der Energielieferant analysiert die Deckungsbeiträge für einen Kunden.
UC2b	Gesamtkosten analysieren	Der Energielieferant analysiert die Gesamtkosten, die die Versorgung eines potenziellen Kunden verursacht.
UC3	Kundendaten auswählen	Der Energielieferant wählt aus, welche Daten je Kunde analysiert werden sollen.
UC3a	Daten für Deckungsbeitrag auswählen	Der Energielieferant wählt aus über welchen Kunden er Daten zum Deckungsbeitrag erhalten möchte.
UC3b	Daten für Gesamtkosten auswählen	Der Energielieferant wählt aus über welchen Kunden er Daten zu den Gesamtkosten erhalten möchte.

Tabelle 3.2: Kurzbeschreibung aller Use Cases

#### Vorbedingungen:

- Die benötigten Daten liegen im System vor.

**Beschreibung des Ablaufs:** Der Energielieferant wählt aus, auf welche Art die Daten des Kunden ausgewertet werden sollen und erhält die gewünschten Informationen.

#### Nachbedingungen:

- Der Energielieferant hat die gewünschten Informationen erhalten.

**Anmerkungen:** Welche Informationen dem Energielieferanten zur Verfügung gestellt werden können ist in den spezialisierten Varianten dieses Anwendungsfalls (UC3a und UC3b) beschrieben.

#### Use Case UC2a Deckungsbeitrag analysieren

Der Energielieferant analysiert die Deckungsbeiträge für einen Kunden.

#### Akteure:

- Energielieferant

### 3 Anforderungen

**Beschreibung des Ablaufs:** Der Energielieferant wählt die Deckungsbeitrags-analysefunktion am Frontend des BICE Systems aus. Er bekommt Informationen darüber, zu welchen Anteilen Arbeits-, Leistungs- und Netznutzungskosten im Preis für eine kWh Strom für einen bestimmten Kunden enthalten sind. Außerdem sieht er, wie viel Profit er bei diesem Preis erzielt. Die angezeigten Informationen können über verschiedene Dimensionen aggregiert angezeigt werden. Diese sind: Zeit, Ort, Kraftwerk und Kunde.

**Anmerkungen:** Dieser Anwendungsfall ist ein Spezialfall von UC2 und erbt dessen Vor- und Nachbedingungen.

#### Use Case UC2b Gesamtkosten analysieren

Der Energielieferant analysiert die Gesamtkosten, die die Versorgung eines potenziellen Kunden verursacht.

#### Akteure:

- Energielieferant

**Beschreibung des Ablaufs:** Der Energielieferant wählt die Gesamtkostenanalysefunktion am Frontend des BICE Systems aus. Er erhält Informationen darüber, wie viel ein Kunde mindestens pro kWh Strom bezahlen muss, damit der Energielieferant keinen Verlust macht. Der Energielieferant kann sich die Kosten aggregiert über ein Jahr, monatlich, täglich oder viertelstündlich anzeigen lassen. Zusätzlich kann er die Anteile von Arbeits-, Leistungs- und Netznutzungskosten an den Gesamtkosten sehen.

**Anmerkungen:** Dieser Anwendungsfall ist ein Spezialfall von UC2 und erbt dessen Vor- und Nachbedingungen.

#### Use Case UC3 Kundendaten auswählen

Der Energielieferant wählt aus welche Daten er über welchen Kunden erhalten möchte.

#### Akteure:

- Energielieferant

#### Vorbedingungen:

- Die Daten über den Kunden sind im BICE System hinterlegt.

**Beschreibung des Ablaufs:** Der Energielieferant wählt am Frontend des BICE Systems den Kunden aus, über den er Daten erhalten möchte. Danach wählt er aus, welche Daten das System ihm liefern soll.

**Nachbedingungen:**

- Der Energielieferant hat die gewünschten Daten erhalten.

**Anmerkungen:** Welche Daten dem Energielieferanten durch das System bereitgestellt werden können ist in den spezialisierten Varianten dieses Anwendungsfalls (UC3a und UC3b) beschrieben.

**Use Case UC3a Daten für Deckungsbeitrag auswählen**

Der Energielieferant wählt aus über welchen Kunden er Daten zum Deckungsbeitrag erhalten möchte.

**Akteure:**

- Energielieferant

**Beschreibung des Ablaufs:** Der Energielieferant wählt am Frontend des BICE Systems den Kunden aus, über den er Daten erhalten möchte. Danach wählt er aus, dass er Informationen zu den Deckungsbeiträgen des gewählten Kunden betrachten möchte.

**Anmerkungen:** Dieser Anwendungsfall ist ein Spezialfall von UC3 und erbt dessen Vor- und Nachbedingungen.

**Use Case UC3b Daten für Gesamtkosten auswählen**

Der Energielieferant wählt aus über welchen Kunden er Daten zu den Gesamtkosten erhalten möchte.

**Akteure:**

- Energielieferant

**Beschreibung des Ablaufs:** Der Energielieferant wählt am Frontend des BICE Systems den Kunden aus, über den er Daten erhalten möchte. Danach wählt der Energielieferant die Zählpunkte des Kunden aus, an denen eine Versorgung stattfinden soll. Die gewünschten Informationen werden dann für den Energielieferanten sichtbar angezeigt.

**Anmerkungen:** Dieser Anwendungsfall ist ein Spezialfall von UC3 und erbt dessen Vor- und Nachbedingungen.

### 3.3 Funktionale Anforderungen

In diesem Abschnitt werden die funktionalen Anforderungen in Form von Tabellen aufgeführt. Jede funktionale Anforderung hat eine eindeutige Identifikationsbezeichnung (ID), einen Namen, eine Beschreibung sowie den Grund für diese Anforderung (Ursache). Die funktionalen sowie die nicht funktionalen Anforderungen ergeben sich aus den zuvor ermittelten Use Cases. Aufgrund der Anzahl der funktionalen Anforderungen verteilen sich diese auf zwei Tabellen auf den nächsten beiden Seiten.

ID	Name	Beschreibung	Ursache
001F	Deckungsbeitrag berechnen	Berechnung des Deckungsbeitrags im System, welcher sich aus <i>Erlös - Gesamtkosten</i> ergibt.	Manuelle Berechnung des Deckungsbeitrags erfordert hohen Arbeits- und Zeitaufwand.
002F	Gesamtkosten berechnen	Berechnen der Gesamtkosten durch das System. Gesamtkosten bestehen aus Leistungspreis, Arbeitspreis und Netznutzungskosten.	Manuelle Berechnung der Gesamtkosten erfordert hohen Arbeits- und Zeitaufwand.
003F	Kennzahlen in Ct/kWh	Die berechneten Kennzahlen sollen in Ct/kWh ausgegeben werden.	Ct/kWh ist die übliche Einheit in dieser Branche.
004F	Kennzahlen zeitlich aggregieren	Kennzahlen sollen je 15 Minuten, je Stunde, je Tag, je Monat und je Jahr berechnet werden können.	Keine zeitliche Aggregation der Daten schränkt die Analysemöglichkeiten des Nutzers ein.

Tabelle 3.3: Funktionale Anforderungen

### 3 Anforderungen

ID	Name	Beschreibung	Ursache
005F	Ergebnis mathematisch visualisieren	Eine mathematische Darstellung des Ergebnisses, aufgeschlüsselt nach Ergebnis und den einzelnen Einflussfaktoren, muss gewährleistet werden.	Wird nur der das reine Ergebnis visualisiert, ist die Zusammensetzung des Ergebnisses nicht erkenntlich.
006F	Ergebnis der Berechnung als Diagramm visualisieren.	Die berechneten Daten müssen in Diagramm-Form visualisiert werden.	Reine Zahlenwerte beinhalten kaum Informationen bezüglich einzelner Dimensionen (Zeit, Region).
007F	Datenkommunikation zwischen den Systemen	Schnittstellen für Daten Im- und Export definieren.	Daten müssen in das System im- und exportiert werden können.
008F	Betrag aggregieren	Der berechnete Betrag soll über verschiedene Dimensionen aggregiert werden können (Zeit, Region, Menge etc.).	Kennzahlen enthalten mehr Informationen, wenn sie in verschiedene Dimensionen aggregiert werden können.
009F	Benutzer-spezifische Arbeitsumgebung	Benutzer kann gewünschten Kunden auswählen, die gewünschte Kennzahl und Visualisierungsart selektieren.	Statische Anwendungen sind nur für einzelne Kunden sinnvoll.

Tabelle 3.4: Funktionale Anforderungen (Fortsetzung)



### 3.4 Nichtfunktionale Anforderungen

In diesem Abschnitt werden die nichtfunktionalen Anforderungen in Form von Tabellen aufgeführt. Diese Tabellen sind analog zu den Tabellen der funktionalen Anforderungen aufgebaut. Aufgrund der Anzahl der nichtfunktionalen Anforderungen verteilen sich diese auf zwei Tabellen auf den nächsten beiden Seiten.

ID	Name	Beschreibung	Ursache
001N	Erweiterbarkeit gewährleisten	Das System muss hinsichtlich neuer Komponenten und Quellen erweiterbar sein.	Nicht erweiterbare Systeme verlieren schnell an Nutzen.
002N	Schnelle Antwortzeiten	Echtzeitberechnung soll gewährleistet werden.	Lange Bearbeitungszeiten wirken sich negativ auf die Benutzerzufriedenheit aus.
003N	Verfügbarkeit für Endgeräte	Das Programm soll auf vielfältigen Endgeräten betrieben werden können (Desktop, Laptop, Tablet, Smartphone etc.).	Die Bereitstellung von Software für nur ein Endgerät ist auf dem heutigen Softwaremarkt nicht mehr zeitgemäß.
004N	Vergleichbarkeit	Die Ergebnisse der Berechnung müssen vergleichbar sein.	Ergebnisse, welche keine Vergleichbarkeit aufweisen, haben nur geringen Nutzen für Kunden.
005N	Branchenweit einsetzbar	Sämtliche Energielieferanten, bzw. mehrere Energielieferanten eines Konzerns (deutschlandweit).	Auf nur einen Kunden zugeschnittene Software schmälert das Marktpotenzial.

Tabelle 3.5: Nichtfunktionale Anforderungen

### 3 Anforderungen

ID	Name	Beschreibung	Ursache
006N	Berechtigungskonzept	Verschiedene Benutzerrollen (Trennung je Lieferant).	Werden betriebsinterne Daten für Unbefugte einsehbar, könnte sich das negativ auf die Wirtschaftlichkeit auswirken.
007N	Benutzer-spezifische Konfiguration	Nutzer können benötigte IT-Ressourcen jederzeit selbstständig nutzen, ordnen und einrichten.	Anforderungen an Flexibilität und Customizing müssen erfüllt werden.
008N	Einheitlichen Zugang bereitstellen (Internet)	Standardbasierter Netzzugriff. Nutzer können unabhängig von Ort und Zeit über das Internet auf den Service zugreifen.	Zugriff auf den Service von nur definierten Orten behindert die Flexibilität der Endanwender.
009N	Simultane Nutzung	Mehrere Benutzer/Kunden müssen gleichzeitig mit dem Service arbeiten können (Mehrmandantenfähigkeit).	Der Service muss für mehrere Kunden zeitgleich unabhängig voneinander betrieben werden können.
010N	Nutzungsbedarf bestimmen	Ressourcen müssen je nach Bedarf skalierbar sein.	Chaotischer und unvorhersehbarer Ressourcenbedarf muss lösbar sein.
011N	Nutzungsbasierte Abrechnung	Abrechnung je nach Verbrauch (Pay-per-Use).	eigene Kapitalbindung kann zu Problemen führen.
012N	Korrektheit	Der Dienst muss korrekt implementiert und dokumentiert sein.	Fehler innerhalb der Software und ungenaue Dokumentation führen zu vermeidbaren Problemen.

Tabelle 3.6: Nichtfunktionale Anforderungen (Fortsetzung)

### 3.5 System-Anforderungen

In diesem Kapitel werden die zuvor ermittelten Anforderungen unter Berücksichtigung des BICE-Systems in drei Kategorien unterteilt: Die Muss-, die Wunsch- und die Abgrenzungskriterien. Die Musskriterien sind die Anforderungen, die realisiert werden müssen, um die Funktionalität des Prototypen zu gewährleisten. Die Wunschkriterien sind für die Funktionalität nicht zwingend erforderlich, allerdings ist eine Umsetzung wünschenswert. Die Abgrenzungskriterien grenzen die Funktionalität des Systems gegenüber weiteren Einsatzbereichen in der Energiewirtschaft ab. Die Auswahl der Kriterien erfolgt im Hinblick auf den Prototypen und nicht auf einen realen Einsatz in der Wirtschaft. Die Bezeichnungen der einzelnen Anforderungen entsprechen denen aus dem Kapitel der fachlichen Anforderungen.

#### 3.5.1 Musskriterien

ID	Name	Beschreibung
001F	Deckungsbeitrag berechnen	Berechnung des Deckungsbeitrags im System, welcher sich aus <i>Erlös - Gesamtkosten</i> ergibt.
002F	Gesamtkosten berechnen	Berechnen der Gesamtkosten durch das System. Gesamtkosten bestehen aus Leistungspreis, Arbeitspreis und Netznutzungskosten.
003F	Kennzahlen in Ct/kWh	Die berechneten Kennzahlen sollen in Ct/kWh ausgegeben werden.
004F	Kennzahlen zeitlich aggregieren	Kennzahlen sollen je 15 Minuten, je Stunde, je Tag, je Monat und je Jahr berechnet werden können.
005F	Ergebnis mathematisch visualisieren	Eine mathematische Darstellung des Ergebnisses, aufgeschlüsselt nach Ergebnis und den einzelnen Einflussfaktoren, muss gewährleistet werden.

Tabelle 3.7: Funktionale Anforderungen der Musskriterien

### 3 Anforderungen

ID	Name	Beschreibung
008N	Einheitlichen Zugang bereitstellen (Internet)	Standardbasierter Netzzugriff. Nutzer können unabhängig von Ort und Zeit über das Internet auf den Service zugreifen.
012N	Korrektheit	Der Dienst muss korrekt implementiert und dokumentiert sein.

Tabelle 3.8: Nichtfunktionale Anforderungen der Musskriterien

#### 3.5.2 Wunschkriterien

ID	Name	Beschreibung
006F	Ergebnis der Berechnung als Diagramm visualisieren.	Die berechneten Daten müssen in Diagramm-Form visualisiert werden.
007F	Daten-kommunikation zwischen den Systemen	Schnittstellen für Daten Im- und Export definieren.
008F	Betrag aggregieren	Der berechnete Betrag soll über verschiedene Dimensionen aggregiert werden können (Zeit, Region, Menge etc.).
009F	Benutzerspezifische Arbeitsumgebung	Benutzer kann gewünschten Kunden auswählen, die gewünschte Kennzahl und Visualisierungsart selektieren.

Tabelle 3.9: Funktionale Anforderungen der Wunschkriterien

### 3 Anforderungen

ID	Name	Beschreibung
001N	Erweiterbarkeit gewährleisten	Das System muss hinsichtlich neuer Komponenten und Quellen erweiterbar sein.
002N	Schnelle Antwortzeiten	Echtzeitberechnung soll gewährleistet werden.
003N	Verfügbarkeit für Endgeräte	Das Programm soll auf vielfältigen Endgeräten betrieben werden können (Desktop, Laptop, Tablet, Smartphone etc.).
004N	Vergleichbarkeit	Die Ergebnisse der Berechnung müssen vergleichbar sein.
005N	Branchenweit einsetzbar	Sämtliche Energielieferanten, bzw. mehrere Energielieferanten eines Konzerns (deutschlandweit).
006N	Berechtigungskonzept	Verschiedene Benutzerrollen (Trennung je Lieferant).
007N	Benutzerspezifische Konfiguration	Nutzer können benötigte IT-Ressourcen jederzeit selbstständig nutzen, ordnen und einrichten.
009N	Simultane Nutzung	Mehrere Benutzer/Kunden müssen gleichzeitig mit dem Service arbeiten können (Mehrmandantenfähigkeit).
010N	Nutzungsbedarf bestimmen	Ressourcen müssen je nach Bedarf skalierbar sein.
011N	Nutzungsbasierte Abrechnung	Abrechnung je nach Verbrauch (Pay-per-Use).

Tabelle 3.10: Nichtfunktionale Anforderungen der Wunschkriterien

#### 3.5.3 Abgrenzungskriterien

Das System soll nur die Vertriebler von Energieversorgern unterstützen. Mit BICE sollen sich Angebote für potenzielle Kunden erstellen lassen und Deckungsbeiträge und Gesamtkosten vorhandener Kunden analysieren lassen.

BICE kann nicht für die Einsatzplanung von Kraftwerken übernehmen oder unterstützen. Die Einsatzplanung umfasst die Analyse von Lastprofilen und der Auslastung von Kraftwerken. Eine Prognose des Energieertrags aus den Erneuerbaren Energien auf Basis von Wetterdaten ist in diesem Zusammenhang ebenfalls ausgeschlossen. Aufgrund der Analysen wird die Auslastung der Kraftwerke so geplant, dass im Optimalfall genau so viel Strom produziert wird wie aktuell verbraucht wird.

Das System unterstützt keine Direktvermarktung. Direktvermarktung ist der Verkauf des erzeugten Stroms direkt an den Großkunden und nicht an den lokalen Netzbetreiber. Direktvermarktung wird für Anlagen, die ab 2014 in Betrieb genommen werden und eine installierte Leistung von 500 KW oder mehr haben bzw. für Anlagen die ab 2016 in Betrieb genommen werden und eine installierte Leistung von mindestens 100 KW haben,

### 3 Anforderungen

vom EEG (Erneuerbare Energien Gesetz) vorgeschrieben.

Der Stromhandel wird in diesem Projekt nicht berücksichtigt. Der Strompreis wird als gegeben angenommen. Alle Prozesse und Faktoren, die zur Berechnung des Strompreises und dessen Höhe beeinflussen werden nicht berücksichtigt. Zu nennen sind hier die Preise für die Energieträger und die Preisentwicklung an der EEX (European Energy Exchange). Gerade die Preise an der Strombörse machen sich in Echtzeit beim Strompreis bemerkbar.

Die Bundesnetzagentur hat die "Geschäftsprozesse zur Kundenbelieferung mit Elektrizität" (GPKE) vereinheitlicht. Das umfasst die Prozesse an sich und die verwendeten Dateiformate. Diese zum Lieferantenwechselprozess vorgelagerten und nachgelagerten Prozesse unterstützt BICE nicht. Die GPKE umfassen die folgenden Prozesse:

- Kündigung
- Lieferende
- Lieferbeginn
- Ersatzversorgung
- Zählerstand- und Zählwertübermittlung/Messwertübermittlung
- Netznutzungsabrechnung
- Stammdatenänderung
- Geschäftsdatenanfrage
- Bestandsliste

BICE wird mit Blick auf End-To-End Prozesse entwickelt. Da die Komplexität im Rahmen der Projektgruppe beherrschbar bleiben muss, werden die oben genannten Aspekte nicht berücksichtigt und sind nur bei einer Weiterentwicklung denkbar.

## 4 Entwurf

Dieses Kapitel beschäftigt sich mit dem konzeptionellen Entwurf des BICE-Systems, der die zuvor durchgeführte Anforderungserhebung auf ein technisches Grundgerüst überführt.

### 4.1 Überblick

Das BICE System besteht aus einem Cloud-basierten Backend, auf das mit einem Endgerät über das Internet zugegriffen werden kann. Als Grundlage dient eine Hardware-Infrastruktur, auf der als Plattform eine beliebige Anzahl an Virtual Machines (VM) bereitstehen, die die einzelnen BICE-Softwarekomponenten bereitstellen. Wichtig hierbei ist, dass die Hardwareressourcen (Infrastrukturebene) sowie die darauf liegende Plattformebene entsprechende Eigenschaften, wie Skalierbarkeit, Monitoring der Virtual Machines, Orchestration und Load Balancing erfüllen, um die Cloud-Charakteristika zu gewährleisten. Eine Übersicht über das System ist in Abbildung 4.1 zu sehen. Ebenso dient ein In-Memory Datenbanksystem als Quelle der Kennzahlen und Standortdaten für das Berechnen der Key Performance Indicators (KPI's). Eine In-Memory-Lösung ist insofern wichtig, da wesentlich höhere Zugriffsgeschwindigkeiten erzielt werden können. In den folgenden Unterkapiteln werden die einzelnen BICE-Softwarekomponenten, das Datenbanksystem sowie die Kennzahlen konzeptionell aufgeführt und beschrieben.

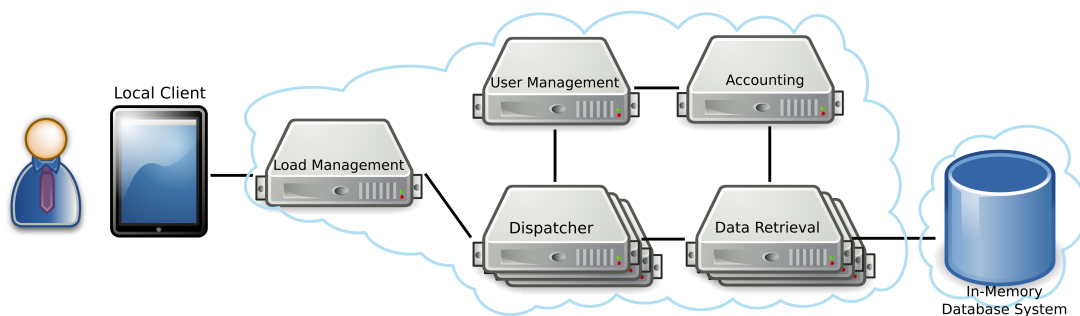


Abbildung 4.1: Übersicht über das BICE System

## 4.2 Komponenten

In diesem Abschnitt werden die einzelnen BICE-Softwarekomponenten auf konzeptionelle Weise beschrieben.

### 4.2.1 Local Client

Die **Local Client** Komponente wird lokal bei dem Nutzer, zum Beispiel im Browser, ausgeführt. Diese Komponente enthält folgende drei Unterkomponenten:

- **Report Display:** Komponente, die für die Darstellung der Benutzungsoberfläche und Anzeigen der Reports verantwortlich ist. Über Data Binding ist diese Komponente mit der Client Business Logic Komponente verbunden.
- **Report Configurator:** Diese Komponente dient zur Konfiguration der Benutzungsoberfläche. Sie ist dafür verantwortlich, dass Ein- und Ausblenden von Oberflächenelementen zu signalisieren und komplexere Aufrufe als Reaktion auf Benutzerinteraktion an die Business Logic weiterzuleiten.
- **Client Business Logic:** Dies ist die Kernkomponente des Clients. Sie enthält den Großteil der Funktionalität des Clients und kommuniziert mit dem Server.

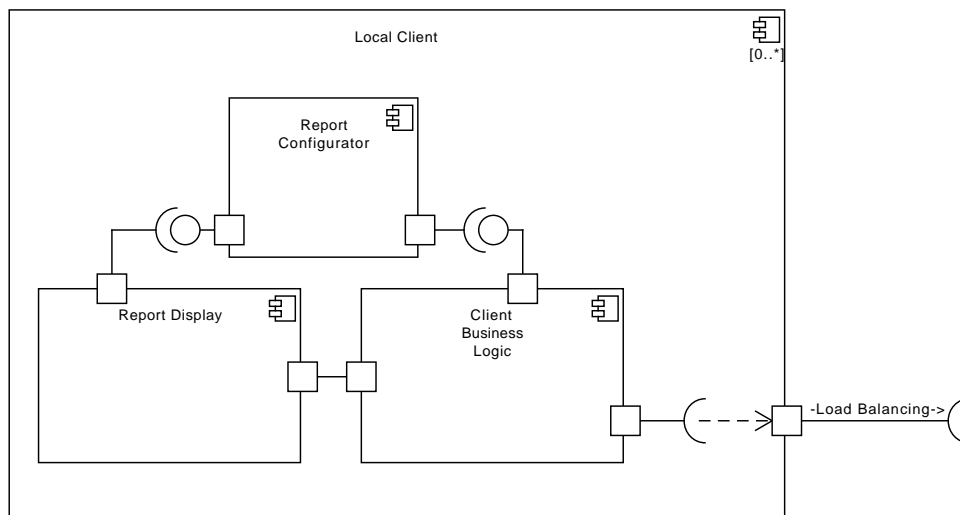


Abbildung 4.2: Local Client Komponentendiagramm

### 4.2.2 User Management

Das **User Management** ist für das allgemeine Verwalten der Benutzer des Systems zuständig und besteht aus folgenden Komponenten:



- **User Manager:** Zuständig für das Erstellen, Löschen, etc. von Benutzern.
- **User Database:** Verwaltet eine Liste aller Benutzer und der zugehörigen Energielieferanten und stellt diese für A&S Management und Accounting zur Verfügung.
- **Authentication & Session Management:** Diese Komponente stellt die Login-Funktionalität bereit und verwaltet aktive Benutzersessions.
- **Authorization & Rights Management:** Diese Komponente verwaltet die Benutzerrechte und ist für die Authorisierung von Anfragen der Clients zuständig.

### 4.2.3 Load Management & Balancing

Das **Load Management & Balancing** ist für die Skalierbarkeit des Systems verantwortlich. Es besteht aus folgenden Komponenten:

- **Resource Monitoring:** Misst die Ressourcennutzung der physischen und virtuellen Serversysteme.
- **Resource Management:** Kann bei Bedarf weitere Ressourcen (=virtuelle Server) zur Verfügung stellen und nicht weiter benötigte Ressourcen freigeben, bzw. herunterfahren.
- **Load Balancing:** Zuständig für die Verteilung der Client-Anfragen auf die Serverkomponenten gemäß der Auslastung der einzelnen Serverkomponenten. Die Weiterleitung erfolgt an die Dispatcher Komponente.
- **Dispatcher:** Diese Komponente nimmt die Anfragen des Clients (die über das Load Balancing weitergeleitet wurden) entgegen. Je nachdem, um welche Art von Anfrage es sich dabei handelt, verteilt der Dispatcher diese an die entsprechende Serverkomponente. Mögliche Serverkomponenten sind: User Manager, Data Retrieval und ETL-Manager. Zuvor wird mit Hilfe des Authentication & Session Management überprüft, ob die Anfrage zu einer gültigen Benutzersession gehört.

### 4.2.4 Data Retrieval und Accounting

Mit Hilfe der **Accounting** Komponente wird aufgezeichnet, wie viele Ressourcen ein Benutzer verbraucht. Die Benutzungsdaten erhält es von der Data Retrieval Komponente. Diese Komponente kann mit Hilfe des User Management die Nutzungsdaten einzelner Nutzer den Energielieferanten (Mandanten) zuordnen. Die **Data Retrieval** Komponente sorgt für die Kommunikation zwischen Client und Datenbank. Anfragen der Clients werden in entsprechende Datenbankabfragen umgewandelt und an die Datenbank gesendet. Diese Komponente enthält folgende Unterkomponente:

- **Query Generator:** Zuständig für das Umwandeln von Anfragen der Clients in Datenbankabfragen. Diese Anfragen werden an die Datenbank (Database System) weitergeleitet und die Antwort entsprechend weitergeleitet.

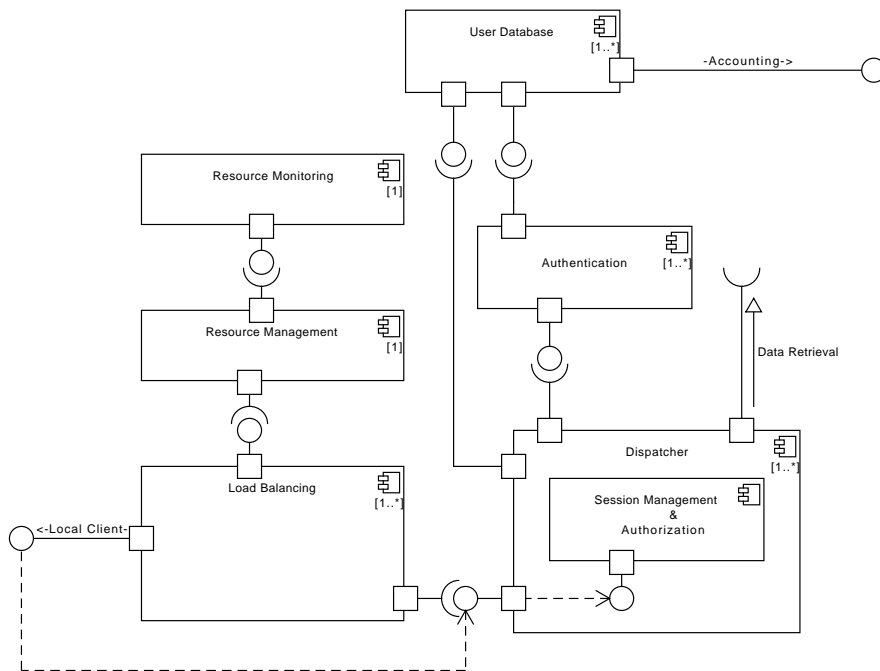


Abbildung 4.3: User- und Load Management und Balancing Komponentendiagramm

#### 4.2.5 Datenbanksystem und ETL

Das In-Memory Datenbanksystem (**Database System**) hält den gesamten Datenbestand in für die Anfragen sinnvoller Form (siehe weiter unten). Über den ETL-Prozess - der in der Entwicklung des BICE-Systems aufgrund geringerer Priorität nicht umgesetzt wird - werden Daten in das Datenbanksystem geladen. Dieser könnte ggf. durch eine Benutzerinteraktion angestoßen werden. Der **ETL-Manager** könnte neue ETL-Processor erstellen, vorhandene entfernen und den ETL-Prozess starten. Der **ETL-Processor** wäre zuständig für das Laden und Transformieren von Daten aus den Quellsystemen in die Datenbank. Die Komponente könnte durch den Source Monitor oder den ETL-Manager zu dieser Aktion aufgefordert werden. Der **Source Monitor** würde die Quell-Systeme überwachen und bei Änderungen den ETL-Prozess starten. Die Daten zur Berechnung der KPI's werden in einer Datenbank gespeichert. Da diese nur zu Analysezwecken verwendet werden soll, liegt es zunächst nahe, sie in einem Star- oder Snowflakeschema anzuordnen. Damit sind jedoch immer Redundanzen vorhanden, die möglicherweise Konsistenzprobleme verursachen. Die redundante Datenspeicherung ist im Hinblick auf die Verwendung eines In-Memory Datenbanksystems eher ungünstig, da der verwendbare Speicherplatz weniger kostengünstig skaliert als bei traditionellen Datenbanksystemen. Darüber hinaus ermöglicht die Datenhaltung im Primärspeicher einen schnelleren Zu-

#### 4 Entwurf

griff, sodass an dieser Stelle ein stärker normalisiertes Datenbankschema verwendet werden kann. Vor diesem Hintergrund ist die in Abbildung 4.4 abgebildete Datenstruktur entwickelt worden.

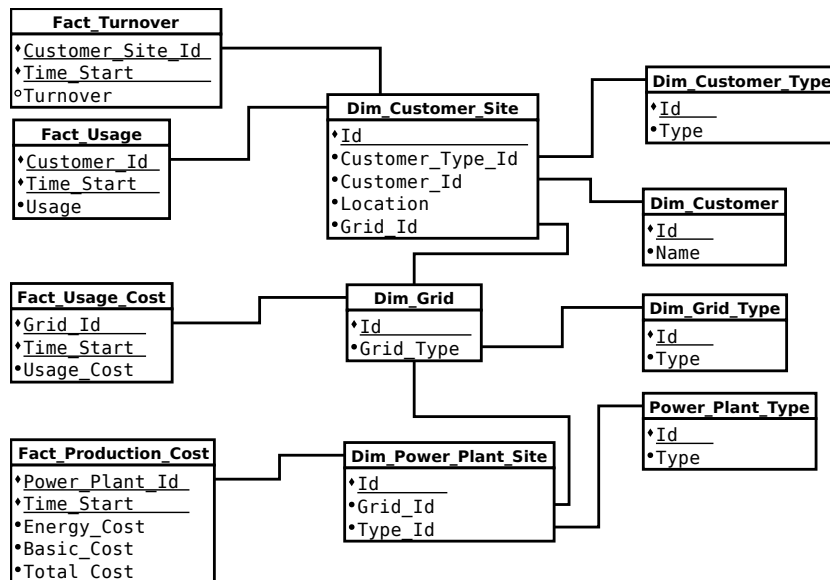


Abbildung 4.4: Struktur der Daten

Die Faktentabellen dienen zur Speicherung der eigentlichen Informationen. Alle Datenfakten besitzen einen partiellen Primärschlüssel **Time\_Start**. Dieser bezeichnet den Startzeitpunkt eines Viertelstunden-Intervalls, für das die entsprechenden Informationen gültig sind. Auf diese Weise ist die Zeitdimension in der Datenbank abgebildet.

In der **Fact\_Turnover** Tabelle ist der Gesamtumsatz bezogen auf einen Kundenstandort hinterlegt. Für zukünftige Werte kann dieser auch NULL betragen, was durch den nicht ausgefüllten Kreis dargestellt ist. Eine weitere vom Kundenstandort abhängige Größe ist der Verbrauch, der in der **Fact\_Usage** Tabelle gespeichert ist. Die beiden anderen Faktentabellen besitzen außer der Zeit keine gemeinsame Dimension.

Kraftwerk- und Kundenstandorte (**Dim\_Customer\_Site** bzw. **Dim\_Power\_Plant\_Site**) sind jeweils in einem (ggf. dem selben) Stromnetz angesiedelt. Da verschiedene Standorte an unterschiedlichen Netzen angeschlossen sein können, sich also möglicherweise auf unterschiedlichen Ebenen der Netzhierarchie befinden, kann der Netztyp über die

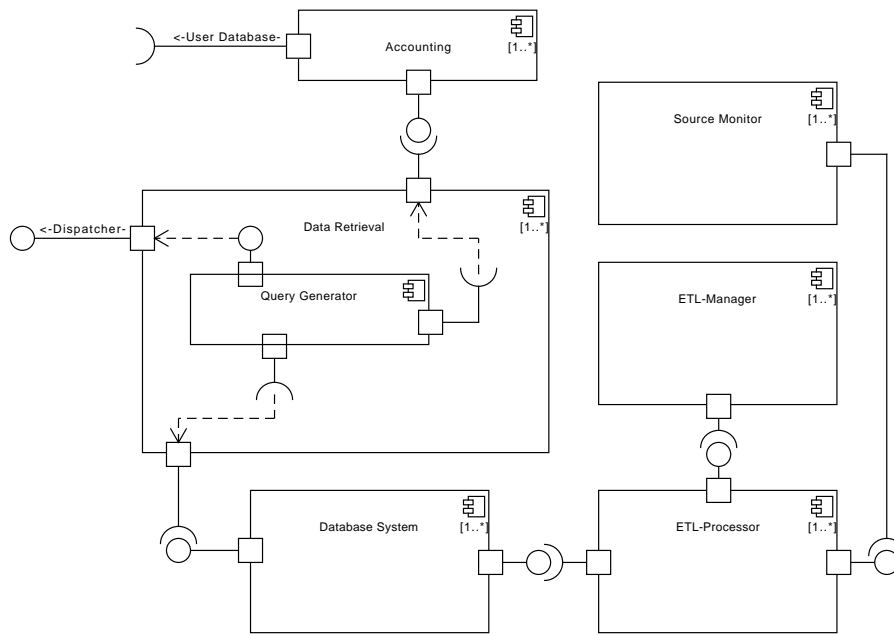


Abbildung 4.5: Data Retrieval, Accounting, Database System und ETL Komponentendiagramm

`Dim_Grid_Type` Tabelle bestimmt werden. Die Kundenstandorte sind einem Kunden (`Dim_Customer`) zugeordnet, der den jeweiligen Standort besitzt. Außerdem haben sie einen Standorttyp (`Dim_Customer_Type`), der die Art des Verbrauchers identifiziert. Dieser identifiziert Standorte beispielsweise als Lagerhäuser, Kaufhäuser oder industrielle Fertigungsanlagen. Auch der Typ eines Kraftwerks (`Dim_Power_Plant_Type`) wird in der Datenbank gespeichert. Verschiedene Kraftwerkstypen unterscheiden sich u.a. in den Produktionskosten pro Kilowattstunde und der Zeit, die zum Hochfahren benötigt wird. Deshalb ist für die Einsatzplanung der Kraftwerkstyp von besonderer Bedeutung.

### 4.3 Kennzahlen

In diesem Abschnitt werden die beiden Kennzahlen Gesamtkosten und Deckungsbeitrag behandelt. Dies geschieht mithilfe von Multidimensional Entity Relationship Models (MER-Diagramme) sowie Kennzahlensteckbriefen.

## 4.3.1 Gesamtkosten

<b>ID</b>	GK
<b>Bezeichnung</b>	Gesamtkosten
<b>Kennzahlensystem</b>	Energiewirtschaftliche Kennzahlen
<b>Beschreibung</b>	Gesamtkosten sind sämtliche Kosten, die aufgebracht werden müssen, um Strom über einen gegebenen Zeitraum an einen bestimmten Endkunden zu liefern. Sie setzen sich aus Arbeitskosten, Leistungskosten und Netznutzungskosten zusammen.
<b>Dimensionen</b>	Zeit, Kunde, Produktion, Netzregion
<b>Exempl. Analysefragen</b>	Was würde es einen Energielieferanten kosten, IKEA Oldenburg für ein Jahr zu versorgen?
<b>Datenquellen</b>	Die Daten zur Berechnung stammen aus angeschlossenen Quellsystemen, wie SAP IS-U oder dem Energy Data Management (EDM).
<b>Berechnung</b>	$Arbeitskosten + Leistungskosten + Netznutzungskosten = Gesamtkosten$
<b>Aktualität</b>	Die Kennzahl wird bei jedem Aufruf oder jeder Aktualisierung neu berechnet.
<b>Eskalationsregeln</b>	Sollten die Gesamtkosten einen kritischen Wert (vorher festgelegt) überschreiten, wird dies mittels einer Meldung angezeigt.
<b>Reports &amp; Dashboards</b>	Die Gesamtkosten werden sowohl mathematisch als auch in einem Diagramm dargestellt.
<b>Self-Service-BI</b>	Systembenutzer können Analysen direkt auf dem BI-CE Frontend laufen lassen.
<b>Verantwortlichkeiten</b>	Projektgruppe Cloud-BI

Tabelle 4.1: Kennzahlensteckbrief Gesamtkosten

## 4 Entwurf

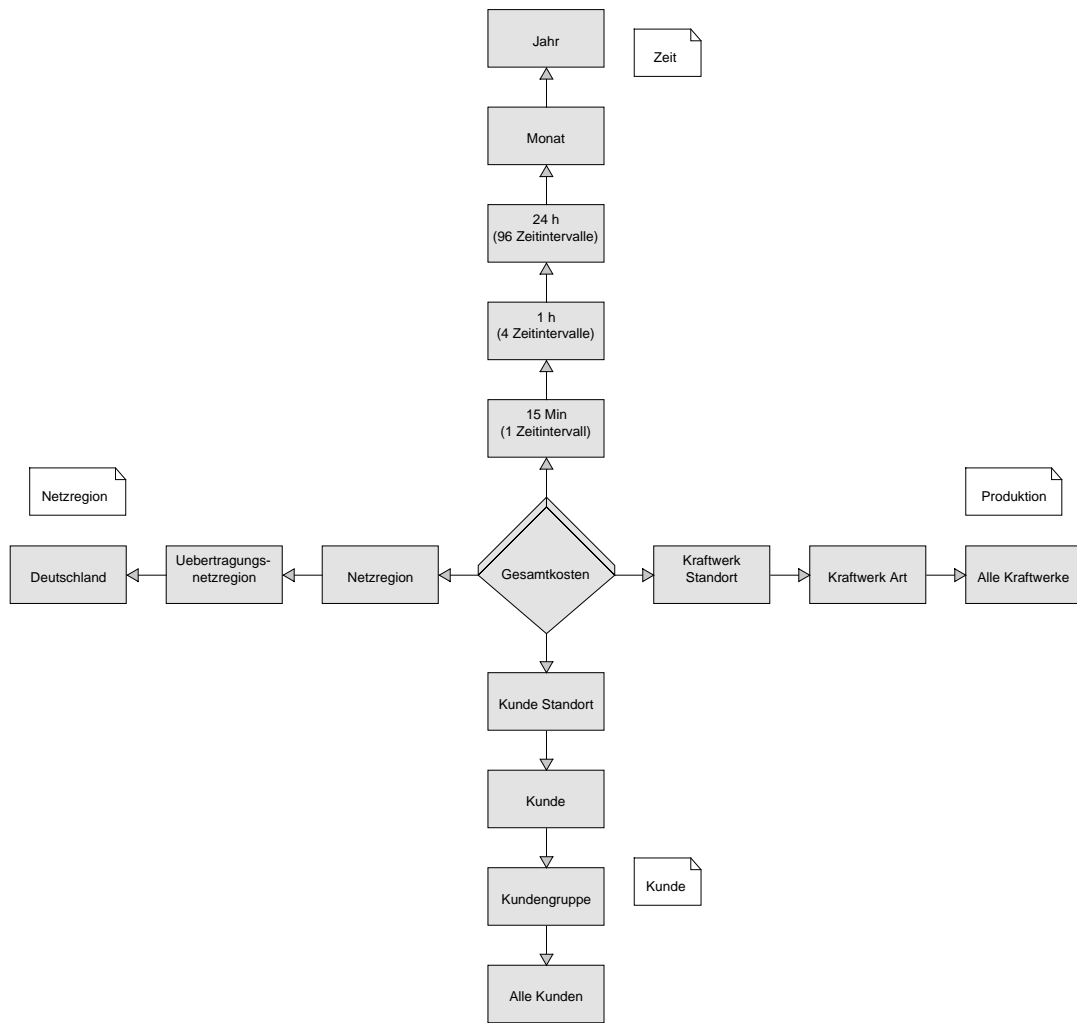


Abbildung 4.6: MER-Diagramm Gesamtkosten

### 4.3.2 Deckungsbeitrag

<b>ID</b>	DB
<b>Bezeichnung</b>	Deckungsbeitrag
<b>Kennzahlensystem</b>	Energiewirtschaftliche Kennzahlen
<b>Beschreibung</b>	Der Deckungsbeitrag beschreibt den Gewinn über einen gegebenen Zeitraum, welcher sich aus dem Erlös des gelieferten Stroms, gemindert um die Gesamtkosten ergibt.
<b>Dimensionen</b>	Zeit, Kunde, Produktion, Netzregion
<b>Exempl. Analysefragen</b>	Wie viel Gewinn wird am Standort Ikea Oldenburg über einen Zeitraum von 6 Monaten erwirtschaftet?
<b>Datenquellen</b>	Die Daten zur Berechnung stammen aus angeschlossenen Quellsystemen, wie SAP IS-U oder dem Energy Data Management (EDM).
<b>Berechnung</b>	$Erlös - Gesamtkosten = Deckungsbeitrag$
<b>Aktualität</b>	Die Kennzahl wird bei jedem Aufruf oder jeder Aktualisierung neu berechnet.
<b>Eskalationsregeln</b>	Sollte der Deckungsbeitrag negativ sein, wird dies mittels einer Meldung angezeigt.
<b>Reports &amp; Dashboards</b>	Der Deckungsbeitrag wird sowohl mathematisch als auch in einem Diagramm dargestellt.
<b>Verantwortlichkeiten</b>	Projektgruppe Cloud-BI

Tabelle 4.2: Kennzahlensteckbrief Deckungsbeitrag

## 4.4 Betriebsbedingungen

Die Betriebsbedingungen beschreiben die technischen Voraussetzungen für die Nutzung des Systems. Um die Funktion des BICE-Systems vollständig nutzen zu können, ist ein Endgerät (Personal Computer, Smartphone oder Tablet) mit Zugang zum Internet erforderlich. Für bestimmte Anwendungen können auch technische Einrichtungen wie z.B. ein Energie Daten Management oder ein Customer Relationship Management notwendig sein. Ebenfalls müssen entsprechend notwendige Quellsysteme/-datenbanken auf Seite des Kunden eingebunden werden (Siehe dazu auch Kapitel Ausblick).

## 4 Entwurf

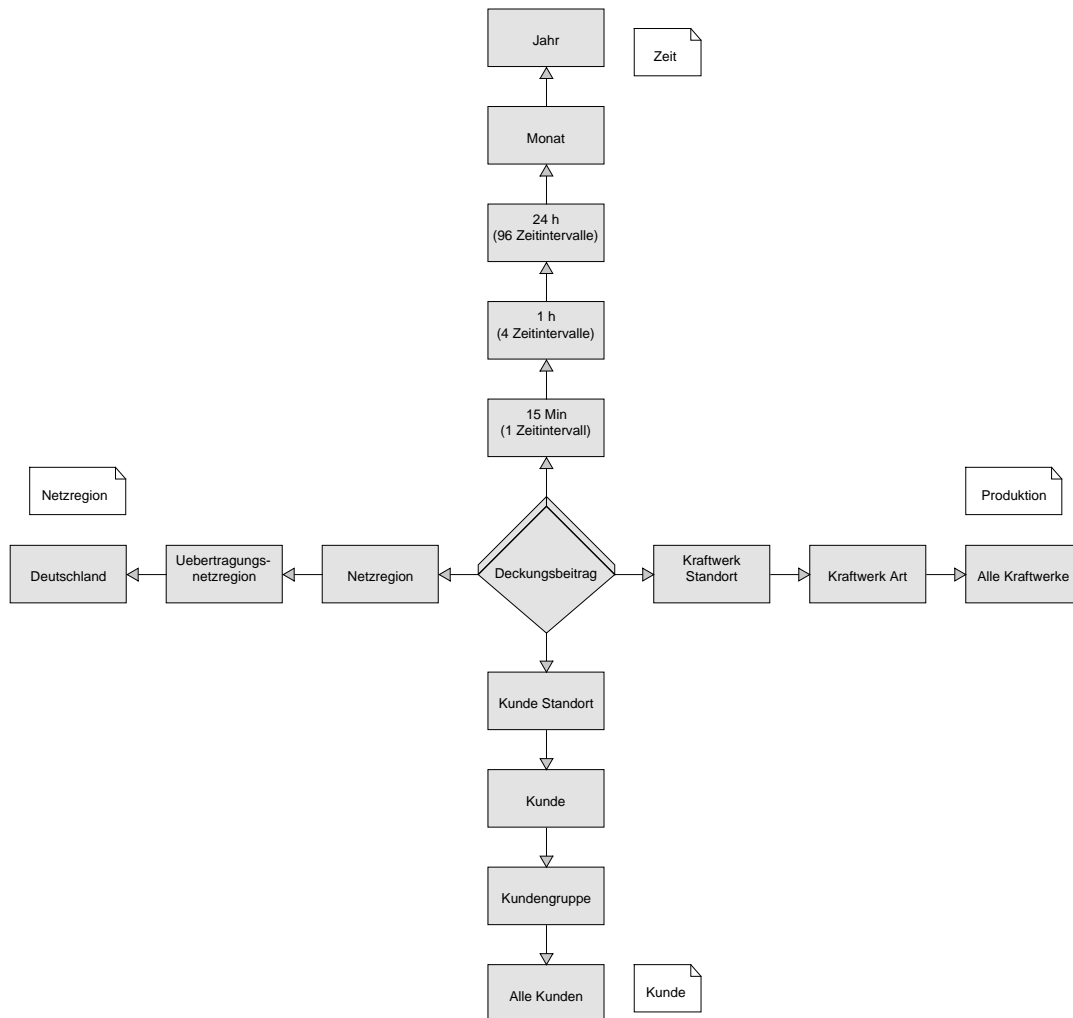


Abbildung 4.7: MER-Diagramm Deckungsbeitrag



# 5 Implementation

Dieses Kapitel beschäftigt sich mit der technischen Realisierung der Komponenten aus dem vorherigen Kapitel. Es werden alle Komponenten der Infrastruktur-, Plattform- und Softwareebene im Einzelnen vorgestellt und eine Installationsanleitung beigelegt.

## 5.1 Infrastrukturebene

Als Infrastruktur für das BICE-System dient eine virtuelle Maschine auf einer HP Converged Cloud Infrastruktur. Ihr stehen 10GB Festplattenspeicher und 64GB Arbeitsspeicher sowie 16 logische CPU-Kerne zur Verfügung. Außerdem existiert eine In-Memory Datenbank, die von den Komponenten der Softwareebene verwendet wird, um die Datenanalysefunktionen des Systems bereitzustellen. Hierbei handelt es sich um eine SAP HANA Instanz. Auf der VM wird Ubuntu 14.04 als Betriebssystem verwendet, auf dem das OpenStack Framework aufgesetzt wurde. Diese Infrastruktur wird vom Hasso-Plattner-Institut für den Zeitraum der Projektgruppe zur Verfügung gestellt. Durch das OpenStack Framework können je nach Last automatisch virtuelle Maschinen gestartet und gestoppt werden. Diese bilden die Grundlage für die Plattformebene des Systems.

### 5.1.1 Load Balancing & Autoscaling

Devstack ist ein Script, um eine OpenStack-Installation zu verwalten. Dabei werden nur bestimmte Linux Distributionen unterstützt. Im Falle des BICE-Projekts wurde Ubuntu Server 14.04 genutzt, da es bereits vorinstalliert war.

#### Installation

Der Installationsprozess kann im Quick Start Guide von Devstack nachgelesen werden. <http://docs.openstack.org/developer/devstack/>

- Devstack aktualisieren

Die aktuelle Version des Devstack Scripts muss heruntergeladen werden. Das kann mittels dem Programm "Git" durchgeführt werden.

```
git clone https://git.openstack.org/openstack-dev/devstack
```

- Konfiguration `local.conf`

In der aktuellen Version der Datei "`devstack/local.conf`" müssen die gewünschten Dienste mit den gewünschten Passwörtern und Tokens konfiguriert werden.

## 5 Implementation

- Run `stack.sh`

Das Script `stack.sh` muss ausgeführt werden. Dies kann einige Minuten in Anspruch nehmen. Wurde das Script erfolgreich durchlaufen, wird die URL für das OpenStack Subsystem 'Horizon' angezeigt. Unter dieser URL ist eine Web-Console zur Steuerung der Horizon-Komponente zu finden. Zusätzlich kann ein Terminal Programm genutzt werden. Zur Nutzung des Terminal Programms sollten die System-Variablen definiert werden. Dies kann mittels des Scripts "devstack/openrc" mit den Argumenten für die Benutzer und Gruppen durchgeführt werden.

z.B. `source devstack/openrc admin admin`

- Installationsschlüssel

In den Beispielen wurde der SSH Schlüssel "bice.key"/"bice.key.pub" genannt.

Um den SSH Schlüssel zu generieren, kann das Programme 'ssh-keygen' benutzt werden.

```
ssh-keygen -t rsa -f bice.key
```

Der öffentliche Schlüssel kann mit folgendem Befehl auf OpenStack geladen werden:

```
nova keypair-add --pub-key bice.key.pub bice
```

OpenStack nutzt den Schlüssel auf den VM-Instanzen.

- Stack Scripts Laden

- `stack.yaml`: Definiert den Großteil der Stack-Ressourcen.
- `server.yaml`: Definiert den Web Server.

Der Stack kann mit folgendem Befehl erreicht werden:

```
heat stack-create teststack -f stack.yaml
```

### Parts

Der sogenannte "Stack," ist eine Gruppe von Ressourcen, um einen gemeinsamen Zweck zu erfüllen. Für das BICE System ist es eine Kombination von VM Instanzen, die die Hauptarbeit übernehmen, um die Infrastruktur der VM's zu unterstützen.

**OS Image** Wie bereits erwähnt haben wir das Ubuntu Cloud Server Betriebssystemimage in der Version 14.04 verwendet.

**VM Definition** OpenStack benutzt den Begriff „Flavor“ für die Definition von virtuellen Maschinen. Im BICE-System wird der Flavor „m1.heat“ verwendet, was eine virtuelle Maschine mit einer CPU, 512MB RAM und keine Festplatte definiert. Die gesamte Software wird beim Start der VM heruntergeladen und installiert. Diese Variante wurde gewählt, da nur 10GB Festplattenspeicher auf dem Server in der HP Converged Cloud zur Verfügung gestanden hat.

**VM Startup Script** Das Startskript ist so konzipiert, dass es alle Komponenten installiert, die notwendig sind um es von einer Basisinstallation des Betriebssystemimages in den Betriebszustand zu versetzen.

Zunächst werden die SSH-Schlüssel installiert, die für das Herunterladen der Komponenten der Softwareebene aus einem SVN-Repository. Danach muss der SVN-Server gegenüber SSH als bekannter Host deklariert werden, da sonst kein automatisches Herunterladen möglich ist. Als nächstes werden die notwendigen Softwarepakete aus dem Ubuntu-Repository installiert. Schließlich werden die WAR-Dateien, als die Komponenten der Softwareebene vorliegen, aus dem SVN heruntergeladen. Nach dem Abschluss dieser Operation ist die VM Einsatzbereit.

**Load Balancer** Ziel war es einen Load Balancer nach dem „Round Robin“-Verfahren arbeiten. Aufgrund eines Problems mit der Konfiguration des OpenStack Netzwerkmoduls (Neutron) war es uns nicht möglich dies umzusetzen, da Neutron ein zwingender Bestandteil des Loadbalancingmechanismus von OpenStack ist.

**Installationsbeschreibung** Das OpenStack Framework wird mit Hilfe eines Installationscripts auf dem System installiert. Daraufhin muss die Konfiguration der einzelnen OpenStack Komponenten erfolgen.

## 5.2 Plattformebene

Auf virtuellen Maschinen wird die Plattform für die Softwarekomponenten bereitgestellt. Als Betriebssystem kommen Ubuntu Cloud Server 14.04. Es handelt sich dabei um Linux Distributionen, die speziell auf den Einsatz in Cloud-basierten Anwendungen abgestimmt wurden. Dieses Grundsystem beherbergt einen Apache Tomcat Application Server, der also Container für die Komponenten der Softwareebene fungiert. Als solcher ist er dafür verantwortlich, dass Aufrufe bestimmter URLs an die entsprechenden Servlets weitergeleitet werden. Darüber hinaus bietet der Servlet Container eine Sessionverwaltung an. Tomcat ermöglicht es außerdem, Servlets verteilt in einem Cluster auszuführen und die Last auf die einzelnen Knoten des Clusters zu verteilen. Die Sessionverwaltung kann entsprechend konfiguriert werden, dass die sessionrelevanten Daten beim Ausfall eines Knotens an einem anderen Knoten zur Verfügung stehen. In der aktuellen Implementation wird Tomcat in der Version 1.7 eingesetzt.

**Installationsbeschreibung** Das Aufsetzen und Starten der VMs wird durch das OpenStack Framework übernommen. Durch entsprechende Konfiguration wird auf dem Betriebssystem-Basis-Image automatisch die notwendige Software installiert und gestartet. Danach werden die Softwarekomponenten aus einem SVN-Repository geladen und durch den Application Server ausgeführt.

## 5.3 Softwareebene

Auf der Software ebene sind die Komponenten angesiedelt welche im vorherigen Entwurfskapitel genauer beschreiben worden sind. Neben der technischen Umsetzung, wird auch eine Installationsbeschreibung mit angehängt.

### 5.3.1 Local Client

Der Local Client wurde auf HTML5-Basis umgesetzt, um den plattformunabhängigen Betrieb der Software zu gewährleisten. Welche Technologien im Allgemeinen verwendet worden sind, wird im folgenden Abschnitt genauer beschrieben.

#### HTML5 Technologien

Nachdem der Local Client anfänglich auf SAP's OpenUI5 entwickelt worden ist, wurde festgestellt, dass der Aufwand bezüglich der Modifizierung, des Debuggings und der Wartung die Erwartungen überschreiten wird. Die folgenden Untersuchungen von OpenUI5 ergaben, dass es zwar gut unterstützt wird, jedoch nicht alle aktuellen Ansätze JavaScript-basierter Entwicklung unterstützt.[Fav] Als Alternative zu SAP OpenUI5 wurde eine Lösung auf Basis von AngularJS und Bootstrap verwendet, da es bezüglich dieser Bibliotheken ein großes Spektrum an Dokumentationen, Handbüchern und Foren gibt. Darüber hinaus gibt der Modulare Charakter die Möglichkeit, viele Support-Bibliotheken von Drittanbietern zu nutzen und so die Integration der Module in das Projekt zu vereinfachen.

#### Libraries

Der "Bower Package Manager" wurde benutzt, um die genutzten Bibliotheken zu verwalten. Bower hat den Vorteil, dass abhängige Bibliotheken automatisch mit installiert werden. Es gibt auch die Möglichkeit, Versionen zu verwalten und Updates zu installieren. Alle Bibliotheken wurden im Ordner "bower\_components" installiert. Die folgenden Bibliotheken wurden benutzt:

**Angular Route (ngRoute)** Angular Route erlaubt die Verbindung von Pfad-Elementen in den URL-Ankern mit HTML-Templates, welche im DOM der übergeordneten HTML Seite speziell definiert sind. Das bedeutet, dass der aktuelle Zustand des Systems als Teil der URL angezeigt wird.

**angular-chart.js** Angular-chart.js liefert ein AngularJS Modul, welches Chart.js-Objekte mit AngularJS Objekten verbindet.

**AngularJS** AngularJS ist ein JavaScript Framework, welches dabei hilft MVC-basierte (oder ähnliche) Muster zu implementieren. So können "Controller in JavaScript und einfache Attributdefinitionen in HTML definiert werden. AngularJS verbindet dann die

## 5 Implementation

Attributdefinitionen und die JavaScript-Definitionen miteinander, sodass komplexe Verbindungen auf einem einfachen und direkten Weg umgesetzt werden können. Darüber hinaus stellt Angular ein modulares Definitions-System zur Verfügung, welches eine einfache Erweiterung des Frameworks ermöglicht. Es ist eine extrem populäre Bibliothek mit über 1500 aktiven Entwicklern auf GitHub verfügbar. <https://github.com/angular/angular>

**Bootstrap** Bootstrap ist eine CSS Bibliothek, welche viele CSS Definitionen bereitstellt, um eine moderne Webseite zu erstellen. Mittels der selektiven Nutzung von Klassen kann das UI durch JavaScript-Routinen, welche einfache CSS Klassen in das DOM integrieren, modifiziert werden.

**Chart.js** Chart.js ist eine erweiterbare Diagramm Bibliothek mit einem einfachen und übersichtlichen Design.

**UI Bootstrap** Diese Bibliothek stellt Abkürzungen und die Einbindung von Bootstrap und AngularJS im Projekt bereit.

**jquery** JQuery ist eine weit verbreitete Bibliothek mit vielen Funktionen zur Steuerung von Webseiten.

### Umsetzung

Nachdem die Grundlagen der genutzten Technologien vorgestellt worden sind, wird im Folgenden auf die einzelnen Bestandteile des Local Clients eingegangen.

**index.html** Der Eingangspunkt des Local Clients ist "index.html". Dieses HTML Dokument lädt alle JavaScript Bibliotheken und CSS Definitionen. Die index.html definiert darüber hinaus die Navigationsleiste im Kopfbereich der HTML-Seite und einen Platzhalter für die Login-Seite und die Inhalts-Seiten. Sollte der Nutzer nicht eingeloggt sein, wird der Login Screen angezeigt. Ansonsten wird mittels Angular Route das jeweilige HTML-Fragment geladen, welches in der URL spezifiziert ist.

**app.js** In der Datei "app.js" sind die Business Logic und die globalen Variablen des Controllers definiert. Daher enthält diese Datei den Großteil des Java Script Codes.

**BiceController** Dieser Controller wird genutzt, um globale Variablen auf Applikations-Ebene zu verwalten, wie die Bestimmung des zu nutzenden Servers, wenn der Nutzer eingeloggt ist.

**BiceDatepickerController** Die Logik hinter den Datepickern (für den Start- und Endzeitpunkt) werden durch diesen Controller realisiert.

## 5 Implementation

**BiceAccountingController** Mittels dieses Controllers werden dem Nutzer alle Informationen bezüglich der genutzten Ressourcen bereitgestellt. Darüber hinaus werden alle Elemente der Accounting-Seite durch den Controller gesteuert.

**CompanyController** Der Company Controller erhält die Liste der Firmen von der BICE API und zeigt diese mittels AngularJS auf der Calculation-Page an. Darüber hinaus beinhaltet der CompanyController die Logik für das Customer-Dropdown Menü.

**LocationController** Der Location Controller erhält die Liste der Standorte zu einem bestimmten Kunden durch das BICE API und zeigt diese mittels AngularJS auf der Calculation-Page an. Darüber hinaus beinhaltet der LocationController die Logik für das Customersite-Dropdown Menü.

**LoginController** Die Steuerung der Elemente der Login-Page, werden durch diesen Controller realisiert. (Der aktuelle Login-Status wird im Bice Controller gehalten, da dieser auch von den anderen Controllern genutzt werden kann)

**GraphController** Der GraphController führt die Überwachung der angegebenen Zeitspanne, das Laden der angefragten Daten vom Server und die Verarbeitung der Daten durch und bindet die Daten in den Graphen ein. Sowohl die Graphen-, als auch die Tabellenanzeige werden durch diesen Controller gesteuert. controller.

**TestController** Wird genutzt um Routinen für Testfälle auszuführen.

**Chart.StackedLine.js** Der Graph wird in einem HTML5 Canvas angezeigt, welcher ein angepasstes Objekt aus dem Line Chart aus der Chart.js nutzt. Für unseren Anwendungsfall ist es wichtig einen Graphen zu visualisieren, welcher aus gestapelten Bereichen besteht. Daher wurde ein Graph mit gestapelten Linien entwickelt, welcher den nachgelagerten Bereich auf den vorherigen legt und bei Mouseinteraktion eine Legende der einzelnen Bereiche mit den zugehörigen Werten anzeigt. Darüber hinaus beinhaltet die Datei eine Logik, welche die farbliche Füllung des Bereichs nur bis zum darunter liegenden Bereich steuert und nicht den kompletten Bereich bis zur X-Achse einfärbt.

### Installationsbeschreibung

Analog zu den anderen BICE-Modulen ist der Local Client als Web-Archiv auf dem Host-Server (z.B. Tomcat) eingerichtet. Aufgrund der Tatsache, dass Tomcat automatisch statische Daten der WAR-Dateien nutzt, wurde eine Datei zur Nutzung der Web-Archiv Definitionen zu "WEB-INF/web.xml" hinzugefügt. Diese Datei erlaubt es den Local Client Ordner archiviert mittels Java's "jar" Programm zu nutzen. Die War-Datei "biceng.war", welche das Front-End enthält, wurde wie folgt installiert:

`jar cf biceng.war` Nach diesem Schritt ist der Local Client bereit, um auf dem Tomcat Server deployed zu werden. Zum Deployen kann die Manageroberfläche des Tomcat-Servers genutzt werden.

### 5.3.2 Dispatcher & Session Management

Wie die anderen Serverkomponenten ist auch der Dispatcher als Java Servlet umgesetzt. Unter Verwendung von *Jersey*, einer Implementierung der *javax-rs* Spezifikation, stellt die Komponente eine Representational State Transfer-Schnittstelle für die anderen Komponenten bereit. Die Implementierung als *javax-rs* konformes Servlet ermöglicht es, das Servlet vom Application Server unabhängig zu entwickeln. Das Servlet wird von einem Tomcat Server ausgeführt und verwendet JSON als Format für die Datenübertragung. Der Einsatz als Servlet ermöglicht es die Hypertext Transport Protocol-Session für das Speichern von sessionspezifischen Daten. Konkret handelt es sich dabei um Informationen über den Benutzer und seinen Anmeldestatus. In der aktuellen Implementation wird der In-Memory Session Store des Servlet-Containers verwendet, was bedeutet, dass bei einem Ausfall des Servers die Daten verloren gehen.

**Detaillierte Beschreibung** Der Dispatcher ist als *javax-rs* Servlet implementiert. Kern des Dispatchers ist die Java-Klasse `DispatcherServlet`, in der die Handler für die HTTP-Anfragen definiert sind. Für jede Art Anfrage, die über die REST-Schnittstelle zur Verfügung gestellt wird, existiert eine entsprechende Methode in der Klasse, die mit dem jeweiligen HTTP-Request Typ (wie z.B. GET, POST, etc.) und optional mit einem Pfad und dem **HTTP Content-Type** annotiert wird. Auf diese Weise kann der Application Server erkennen, welche Methoden aufgerufen werden sollen. Wenn der Dispatcher eine Anfrage erhält, überprüft er zunächst den Anmeldestatus des Local Client, welcher die Anfrage gesendet hat. Falls sich der Benutzer bereits angemeldet hat wird die Anfrage ggf. umgeformt und an die Zielkomponente weitergeleitet. Die Antwort wird an den Local Client gesandt, sobald die Serverkomponente antwortet. Da der Dispatcher die Verbindung offen hält, bis die Zielkomponente geantwortet hat, sollte der Local Client diese Anfragen asynchron ausführen.

**Schnittstelle** Die Schnittstelle des Dispatchers definiert unterschiedliche Anfragen, die jeweils durch den **URL-Pfad**, den **HTTP-Content-Type** und **HTTP-Anfrage-Methode** identifiziert sind. Alle Handler im Dispatcher reagieren auf GET-Anfragen mit `application/json` als Content-Type. Das Präfix aller Pfade lautet `service` und ist in einer Konfigurationsdatei festgelegt. Relativ dazu wurden weitere Pfade definiert. Damit der *Local Client* Anfragen stellen kann, die dazu führen, dass Daten abgerufen werden, muss er sich zunächst gegenüber dem System authentifizieren. Dies geschieht durch das Übermitteln von Benutzername und Passwort als Query-Parameter an den Pfad `service/login`. Im Dispatcher wird daraufhin mit Hilfe eines Session Stores der Login-Status des Clients festgehalten. Mit allen weiteren Anfragen, die der *Local Client* stellt, können Daten aus der Datenbank abgerufen werden, bis dieser die Session durch einen Aufruf des Pfades `service/logout` beendet oder 30 Minuten lang keine Aktivität stattfindet.

Unter dem Pfad `service/getData` ist die Funktionalität zum Abrufen von KPI-Werten entsprechend der ausgewählten Filter erreichbar. Die Filtereinstellungen werden als Query-Parameter übergeben. Die zu übergebenden Parameter sind:

## 5 Implementation

- `id`: identifiziert einen Standort eines Kunden des Energieversorgers
- `timeStart`: gibt den Startzeitpunkt des Intervalls für die KPI-Werte an
- `timeEnd`: gibt den Endzeitpunkt des Intervalls für die KPI-Werte an

Eine typische Abfrage könnte `http://host/service/getData?id=1&timeStart='2015-04-20 00:00:00'&timeEnd='2015-04-27 00:00:00'` sein. Eine entsprechende Anfrage wird an die Data Retrieval Komponente gesendet, worauf der Dispatcher ein JSON-Objekt erhält, welches direkt an den Local Client weitergeleitet wird.

Der Pfad `service/getCompanies` ermöglicht das Abrufen einer Liste von Kunden des Energieversorgers, welche ebenfalls als JavaScript Object Notation-Objekt an den Local Client gesendet wird. Diese Abfrage sieht keine Parameter vor.

Des Weiteren ist noch der Pfad `service/getLocations` vorhanden, welcher als Parameter eine `customerId` erwartet und die Standorte des entsprechenden Kunden als JSON-Objekt liefert.

Neben den Daten zu den Kunden des Energieversorgers kann dieser auch Informationen zur Verwendung des BICE Cloud-Dienstes und den Kosten seiner Anfragen abrufen. Dies geschieht mit durch das Senden einer Anfrage an `service/accountingInfo`. Es sind dabei folgende Parameter erforderlich:

- `startDate`: Startdatum des Zeitraums, für den Abrechnungsinformationen angefragt werden sollen.
- `endDate`: Enddatum des Zeitraums, für den Abrechnungsinformationen angefragt werden sollen.
- `granularity`: Granularität der Informationen.

Das Format der Parameter `startDate` und `endDate` muss dem Schema `YYYY-mm-DD` entsprechen. Es müssen also eine vierstellige Jahreszahl, ein zweistelliger Monat und ein zweistelliger Tag angegeben werden. `granularity` kann die Werte `day`, `month` oder `year` annehmen, um die entsprechende Granularität auszuwählen.

### 5.3.3 Authentication Management

Auch das AuthenticationManagement ist als Servlet implementiert. Es kommuniziert über Java Database Connectivity mit der Benutzerdatenbank um zu überprüfen, ob eine gegebene Kombination von Benutzername und Passwort existiert.

#### Detaillierte Beschreibung

**Interfaces** Das Authentication Management bietet nur eine Abfrage für anderen Komponenten an, die über `/login` zu erreichen ist. Erwartet werden die Parameter `userName` und `password`, die die zu überprüfende Kombination von Benutzername und Passwort enthalten.



### 5.3.4 Data Retrieval

Das Data Retrieval ist für die Übermittlung der Anfrage des Dispatcher an die Datenbank und die Rückgabe der Daten zum Dispatcher verantwortlich und ist ebenfalls als Servlet implementiert. Die Anfragen an die Datenbank werden vor dem Absenden dynamisch generiert. Weiterhin übermittelt das Data Retrieval die notwendigen Abrechnungsdaten an das Accounting und stellt diese auch in geeigneter Form mit Hilfe des Accounting bereit. Im Folgenden werden die eingesetzten Technologien beschrieben. Bei Technologien, die auch in anderen Komponenten verwendet werden, wird auf das entsprechende Kapitel verwiesen, um Wiederholungen zu vermeiden.

**Detaillierte Beschreibung** In der `DataRetrieval` Klasse ist die Funktionalität der Data Retrieval Komponente implementiert. Diese stellt sie über eine HTTP-Schnittstelle für den Dispatcher bereit. Die Daten der Kunden des Energieversorgers lädt die Komponenten mit Hilfe der JDBC-API und den entsprechenden Treibern für die verwendete Datenbank. Im Falle des BICE-Systems handelt es sich dabei um eine SAP HANA Datenbank. Außerdem verwendet das Data Retrieval die Accounting Komponente, um Abrechnungsdaten zu speichern und zur Verfügung zu stellen. Diese ist eine Java-Komponente und wird durch entsprechende Methodenaufrufe angesprochen.

**Schnittstelle** Wie auch beim Dispatcher wird die Funktionalität der Data Retrieval Komponente über eine HTTP-Schnittstelle zur Verfügung gestellt, die GET-Anfragen mit `application/json` Content-Type akzeptiert. Der Wurzelfad aller Anfragen lautet auch hier `service`.

Unter dem Wurzelfad ist die Funktionalität zur Abfrage von Daten erreichbar. Analog zur Schnittstelle des Dispatchers sind auch hier die Parameter `id`, `startDate` und `endDate` vorhanden. Darüber hinaus existiert noch der Parameter `userId`, der den Systembenutzer identifiziert. Während die anderen drei Parameterwerte vom Local Client stammen und durch den Dispatcher an das Data Retrieval weitergeleitet werden, wird die der `userId` Parameter durch den Dispatcher selbst festgelegt. Mit Hilfe dieses Parameters kann einerseits sichergestellt werden, dass nur die Daten des angemeldeten Benutzers abgefragt werden können, andererseits wird er auch an das Accounting übermittelt, sodass die Ressourcennutzung dem entsprechenden Benutzer zugeordnet werden kann.

- `id`: identifiziert einen Standort eines Kunden des Energieversorgers
- `startDate`: gibt den Startzeitpunkt des Intervalls für die KPI-Werte an
- `endDate`: gibt den Endzeitpunkt des Intervalls für die KPI-Werte an
- `userId`: die Datenbank ID des angemeldeten Benutzers

Neben der Möglichkeit KPI-Werte abzufragen ermöglicht die Schnittstelle auch die Abfrage von vorhandenen Kunden (`service/getCompanies`) und den dazugehörigen Kundenstandorten (`service/getLocations`). Die Parameter sind bei diesen Abfragen die

gleichen wie im Dispatcher. Eine Anfrage an `service/accountingInfo` ermöglicht das Abfragen von Abrechnungsdaten unter Verwendung der Accounting-Komponente. Die Parameter dieser Abfrage sind die gleichen wie die dem Dispatcher übergebenen Parameter unter dem gleichen Pfad.

### 5.3.5 Accounting

Die Accounting Komponente verwendet JDBC um abrechnungsrelevante Informationen in einer Datenbank zu speichern bzw. diese bereitzustellen. Sie ist als Java-Klasse implementiert und wird aus dem Data Retrieval heraus aufgerufen.

**Interfaces** Folgende Methoden werden von der Accounting Komponente zur Verfügung gestellt.

- `JSONObject getUsage(String startDate, String endDate, long userId, String granularity, Connection con)`: Diese Methode liefert für ein Start- und Enddatum (`startDate` und `endDate`), einer `userId` und einer Granularität ein JSON-Objekt, in dem die verwendeten Ressourcen sowie den Preis, den das Energieversorgungsunternehmen für die Ressourcenverwendung zahlen muss. Das Format für `startDate` und `endDate` entspricht dem im Dispatcher bzw. Data Retrieval verwendeten Format. `con` ist das Objekt, das die Verbindung zur Datenbank verwaltet und wird vom Data Retrieval bereitgestellt.
- `public void insertUsage(long userId, String md5sum, Connection con)`: Diese Methode wird nach dem Ausführen einer Datenbankabfrage aufgerufen. Sie erwartet die id des Benutzers, der die Abfrage ausgelöst hat sowie den MD5-Hash des Abfrage-Strings. Der Hash ist notwendig, um die Ressourcenauslastung, die durch die Abfrage entstanden ist, vom Datenbanksystem zu erhalten. Auch hier ist die Verbindung zur Datenbank durch den `con` Parameter gewährleistet.

## 6 Evaluation

Die Problemstellung, der sich die Projektgruppe stellen musste, war die Entwicklung eines BI-Cloud-Dienstes für die Deckungsbeitragsrechnung in der Energiewirtschaft. In diesem Kapitel wird untersucht, ob diese Problemstellung von der Projektgruppe erfolgreich umgesetzt worden ist. Als Ergebnis der Entwicklung ist das BICE System entstanden, welches in seinen Bestandteilen bereits in den vorherigen Kapiteln vorgestellt worden ist.

### 6.1 Leistungsspektrum

Das BICE System ist ein Cloud-basiertes BI-Auswertungstool, welches plattformunabhängig als Dienst bereitgestellt und genutzt werden kann. Es bietet eine serverseitige Datenverarbeitung und Aggregation in nahezu Echtzeit, somit kann der Dienst Endgeräteunabhängig mit einer hohen Performance genutzt werden. Zur Nutzung des Dienstes wird dem Nutzer eine Applikation im Web-Browser bereitgestellt, welche sich je nach genutztem Endgerät an die Displaygröße anpasst, um eine übersichtliche Visualisierung zu gewährleisten. Die Kosten für die Nutzung des Dienstes werden je nach genutzten Ressourcen berechnet und sind so je nach Größe und Volumen der zu analysierenden Daten Kundenindividuell abrechenbar. Zur Nutzung des Dienstes ist ein Kundenkonto erforderlich, welches bei den Serveradministratoren angefordert werden kann. Hat sich der Kunde eingeloggt, kann der Nutzer sich den Deckungsbeitrag für einen bestimmten Standort einer Kunden über den gewünschten Zeitraum sowohl grafisch als auch tabellarisch anzeigen lassen. Darüber hinaus werden dem Kunden die durchschnittlichen Kosten für die Stromproduktion und der Durchschnittliche Profit für den ausgewählten Zeitraum angezeigt.

### 6.2 Qualität der Software

Um zu gewährleisten, dass das BICE-System gemäß den Anforderungen und Kundenerwartungen arbeitet, ist ein Testkonzept entwickelt und umgesetzt worden. Darin sind unterschiedliche Abstraktionsebenen berücksichtigt und entsprechende Tests ausgewählt worden.

#### 6.2.1 Quelltext-Ebene

Auf dieser Ebene sind Unit-Tests in Form von Black- und Whitebox-Tests zum Einsatz gekommen. Dadurch ist sichergestellt worden, dass die getesteten Ausführungspfade bei

definierten Eingaben die erwarteten Ausgaben liefern. Da das Testen aller Ausführungspfade sehr aufwendig ist, sind nur die wichtigsten Teile getestet worden. Bei Whitebox-Tests werden Einblicke in den Quellcode vorausgesetzt. Bei Blackbox-Tests dagegen wird der Quellcode nicht betrachtet. In 6.1 und 6.2 sind die Ergebnisse solcher Blackbox-Tests aufgeführt. Dabei wurden bestimmte Funktionen, wie der Login-Vorgang, mit verschiedenen Grundbedingungen durchgeführt.

Testfall	Eingabe	Ausgabe	
		Schlüssel	Wert
Test1	Richtige Benutzername Richtige Kennwort	success	true
Test2	Falsche Benutzername Richtige Kennwort	error	"Query returned no results"
Test3	Falsche Benutzername Falsche Kennwort	error	"Query returned no results"
Test4	Richtige Benutzername Falsche Kennwort	success	false

Tabelle 6.1: Ergebnisse eines Blackboxtests

### 6.2.2 Komponenten-Ebene

Das Zusammenspiel der einzelnen Komponenten wird auf dieser Ebene überprüft. Während der Entwicklung sind die Komponenten auf einem Entwicklungsserver ausgeführt worden, wodurch ein kontinuierliches Testen der *Integration* ermöglicht worden ist. Ebenfalls wurde die Datenbank-Komponente anhand eines Performanztests ausgiebig geprüft. Der Performanztest soll die Abfragezeiten von BICE testen. Es werden die Werte für die gesamte Abfragedauer und die Verarbeitungszeit auf der Datenbank erfasst. Das Testprogramm ist in Java implementiert und steht als \*.jar Datei zur Verfügung. Java wurde gewählt, damit das Programm auf jedem Betriebssystem in der JVM in gleicher Weise ausgeführt werden kann.

Zur statistischen Analyse wird das Programm R verwendet. Es werden der Kolmogorov-Smirnov und der  $\chi^2$ -Anpassungstest angewandt, um auf Normalverteilung zu testen. Liegt eine Normalverteilung vor, können Aussagen zur Wahrscheinlichkeit der Abfragedauer gemacht werden. Außerdem sollen die Werte Aufschluss über die Skalierbarkeit des Systems geben.

#### Rahmenbedingungen

Es nehmen vier Clients an dem Performanztest teil. Alle Clients müssen über das Internet und einen VPN mit den HPI FSOC Ressourcen verbunden sein. Allen Clients wird eine ausführbare \*.jar Datei mit dem Programm für den Performanztest zur Verfügung gestellt.

Es wird die Zeit gemessen, die vom Absetzen der Anfrage bis zum vollständigen Erhalt der

## 6 Evaluation

Testfall	Eingabe	Ausgabe
Test1	Richtige URL Richtiger Benutzername Richtiges Kennwort	<u>Connection-Objekt</u>
Test2	Falsche URL Richtiges Benutzername Richtiges Kennwort	<u>Exception</u>
Test3	Richtige URL Falscher Benutzername Richtiges Kennwort	<u>Exception</u>
Test4	Richtige URL Richtiger Benutzername Falsches Kennwort	<u>Exception</u>
Test5	Falsche URL Falscher Benutzername Richtiges Kennwort	<u>Exception</u>
Test6	Richtige URL Falscher Benutzername Falsches Kennwort	<u>Exception</u>
Test7	Falsche URL Richtiger Benutzername Falsches Kennwort	<u>Exception</u>
Test8	Falsche URL Falsches Benutzername Falsches Kennwort	<u>Exception</u>

Tabelle 6.2: Ergebnisse eines Blackboxtests

Antwort vergeht. Alle anderen Prozeduren des Programms werden nicht berücksichtigt. Dadurch lässt sich die Performanz des gesamten BICE-Systems testen. Das Programm verhält sich exakt wie ein normaler Client. Es loggt sich zu Beginn ein und stellt alle Anfragen an die Dispatcher Komponente. Die Anfragen werden zufällig generiert, um ein möglichst realistisches Szenario nachzustellen. Es werden nur Abfragen generiert, zu denen Werte in der Datenbank vorhanden sind. Außerdem ist die Zeiterfassung somit weitestgehend von der Hardware des Clients unabhängig. Der Plan-Cache der HANA Datenbank wurde vor der Ausführung geleert. Es sind also keine Abfragepläne vorhanden, auf die zurückgegriffen werden kann.

Die Ausführungszeit auf der SAP HANA Datenbank wird über das Accounting abgefragt. Das Accounting erfasst die durchschnittliche Rechenzeit, die für einen Query benötigt wird. Die Zeit liegt in Mikrosekunden vor und wird durch Multiplikation mit dem entsprechenden Faktor in Millisekunden umgerechnet.

**Statistische Analyse**

Zur statistischen Auswertung werden der einseitige Kolmogorov–Smirnov–Anpassungstest und der  $\chi^2$ –Anpassungstest herangezogen. Mit beiden Tests wird geprüft, ob die Grundgesamtheit, aus der die Stichprobe stammt, normalverteilt ist. Ist sie normalverteilt, kann eine Aussage über die Wahrscheinlichkeit getroffen werden, mit der eine Abfrage eine bestimmte Zeit dauert. Die Parameter  $\mu$  und  $s^2$  werden mit den folgenden Formeln berechnet:

$$\mu = \frac{1}{n} \sum_{i=1}^n i$$

bzw.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x - \mu)^2$$

$n$  ist der Stichprobenumfang. Die Werte in Tabelle 6.3 wurden für  $\mu$  und  $s^2$  ermittelt.

empirischer Mittelwert ( $\mu$ )	77,85
empirische Varianz ( $s^2$ )	625,03

Tabelle 6.3: Empirischer Mittelwert und empirische Varianz

Bei beiden Anpassungstests handelt es sich um Hypothesentests. Es werden die folgenden Hypothesen  $H_0$  und  $H_1$  aufgestellt:

$H_0$ : Die Grundgesamtheit ist  $\mathcal{N}(77,85; 625,03)$  verteilt.

$H_1$ : Die Grundgesamtheit ist nicht  $\mathcal{N}(77,85; 625,03)$  verteilt.

Sind die Testgrößen  $D$  bzw.  $\chi^2$  kleiner als bestimmte kritische Werte, dann wird  $H_0$  angenommen, andernfalls abgelehnt. Alternativ lässt sich das Ergebnis an dem  $p$ –Wert ablesen. Der  $p$ –Wert gibt an, mit welcher Wahrscheinlichkeit eine vorliegende Stichprobe bei einer entsprechenden gezogen wird. Der kritische Wert  $d_\alpha$  für den Kolmogorov–Smirnov–Anpassungstest berechnet sich aufgrund des großen Stichprobenumfangs wie folgt:

$$d_\alpha = \frac{\sqrt{-0.5 \ln \left( \frac{\alpha}{2} \right)}}{\sqrt{n}}$$

Der kritische Wert für den  $\chi^2$ –Anpassungstest ist ein Quantil der Verteilung und kann aus der entsprechenden Tabelle abgelesen werden.

Für ein Signifikanzniveau von  $\alpha = 0,001$  ergibt sich für den kritischen Wert:

$$d_{0,001} = \frac{\sqrt{-0.5 \ln \left( \frac{0,001}{2} \right)}}{\sqrt{400}} = 0,0974737$$

Für das Quantil zum Signifikanzniveau  $\alpha = 0,001$  und 399 Freiheitsgraden ergibt sich:

$$\chi_{399; 0,001} = 476,6$$

## 6 Evaluation

Die Prüfgrößen sind jeweils deutlich größer als die kritischen Werte ( $0,77 > 0,097$  und  $3204,9 > 476,6$  siehe Tabelle 6.4). Damit entstammt die Stichprobe keiner normalverteilten Grundgesamtheit und es können keine weiteren Aussagen in diesem Zusammenhang getroffen werden.

Test	Werte	Ergebnis
Kolmogorov-Smirnov	$D = 0,77; p = 0,5952$	nicht normalverteilt
$\chi^2$	$\chi^2 = 3204,9; p < 2,2 \cdot 10^{-16}$	nicht normalverteilt

Tabelle 6.4: Ergebnisse der Anpassungstest

### Auswertung

Die statistischen Analysen gezeigt haben, dass die Stichprobe nicht aus einem normalverteilten Wahrscheinlichkeitsraum stammt. Die Werte geben dennoch Aufschluss über die Performanz und die Skalierbarkeit von BICE.

Die erhobenen Werte für den Performanztest sind in Abbildung 6.1 als Graph dargestellt. Die Werte der verschiedenen Clients sind farblich unterschiedlich markiert. Die

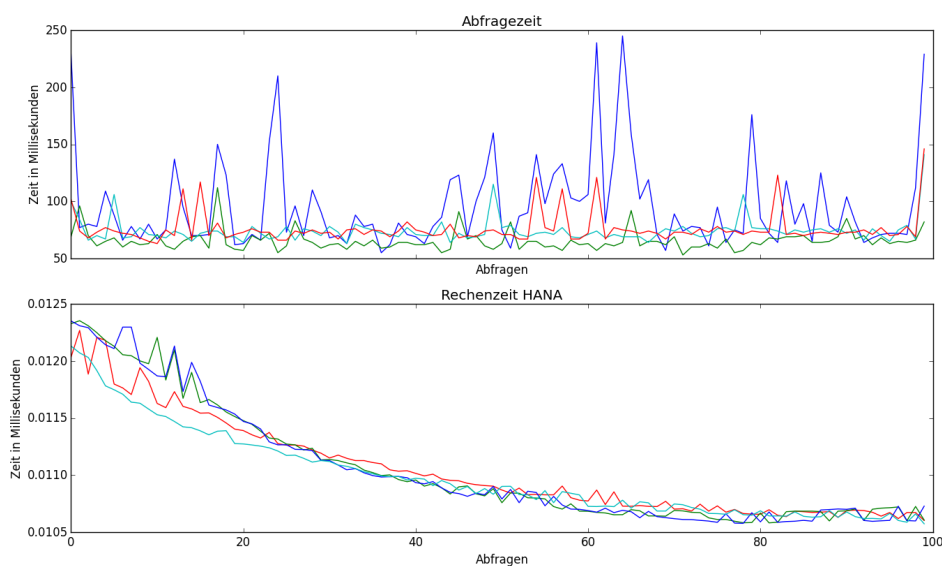


Abbildung 6.1: Der obere Graph zeigt die Gesamtzeit der Abfrage und der untere die zugehörige Rechenzeit, die auf der SAP HANA Datenbank benötigt wurde.

Farben in beiden Graphiken korrespondieren. Die blaue Kurve im unteren Graphen ist die zugehörige Rechenzeit auf der Datenbank zur blauen Kurve in der oberen Graphik usw.

## 6 Evaluation

Die Abfragen wurden zufällig gewählt. Die Rechenzeit auf der Datenbank befindet sich im Mikrosekundenbereich und die Schwankungen der Abfragezeit im Millisekundenbereich. Deshalb können die Schwankungen auf die Internetverbindung zurückgeführt werden und nicht auf BICE.

Die stark abfallenden Kurven der Rechenzeit auf der Datenbank lassen sich mit dem Plan Cache erklären. Bevor der Test gestartet wurde ist der Plan Cache geleert worden. Es lagen also keine Abfragepläne auf der Datenbank vor. Für die ersten Anfragen wurde ein neuer Plan berechnet. Im weiteren Verlauf konnte auf vorhandene Pläne zurückgegriffen werden. Ca. ab der 70. Anfrage mussten nahezu keine neuen Pläne berechnet werden. Deshalb stabilisiert sich die Rechenzeit ab dem Zeitpunkt.

Die Werte zeigen, dass die SAP HANA Datenbank für das vorliegende System gut geeignet ist.

Um die Skalierbarkeit zu testen wurden Referenzanfragen separat ausgeführt. Während der Bearbeitung der Referenzanfragen wurde BICE nicht anderweitig belastet. Die erfassten Werte sind in Abbildung 6.2 dargestellt. Die Abfragezeit der Referenzabfragen

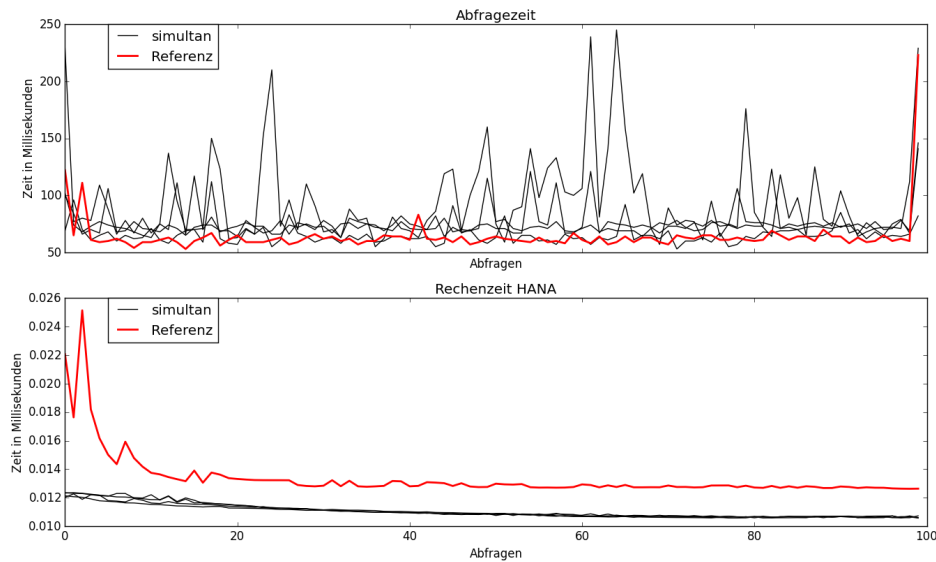


Abbildung 6.2: Der rote Graph stellt die Werte der Referenzabfrage dar. Die schwarzen Linien entsprechen den Werten aus Abbildung 6.1.

liegt unter bzw. mit der Abfragezeit der simultanen Abfragen gleich auf. Die Peaks lassen sich wieder auf die Internetverbindung zurückführen. Die Ressourcen und die gewählten Methoden zur Skalierbarkeit sind angemessen und funktionieren.

Das Verhalten der Rechenzeit auf der Datenbank lässt sich durch Caches erklären. Die Rechenzeit bei den Referenzanfragen liegt über der, der simultanen Anfragen. Bei den simultanen Anfragen werden Abfragepläne berechnet und im Plan Cache vorgehalten. Die Wahrscheinlichkeit, dass zwei Clients eine Anfrage stellen, die mit demselben Plan



ausgeführt werden kann, ist relativ hoch und wird mit der Anzahl der Abfragen größer. Bei der einzelnen Abfrage ist die Wahrscheinlichkeit, dass ein Plan mehrfach genutzt werden kann geringer. Deshalb muss für jede Anfrage ein neuer Plan erstellt werden. Asymptotisch wird sich die Rechenzeit für die Referenzanfragen der Rechenzeit, der simultanen Abfragen annähern.

Da die simultanen Abfragen alle Zeitgleich bearbeitet wurden kann eine Aussage über die Güte der Skalierbarkeit getroffen werden. Wäre die Skalierbarkeit nicht gewährleistet, dann wäre die Abfragezeit bei den simultanen Abfragen auf einem deutlich höheren Niveau als die, der Referenzanfragen. Die Rechenzeit auf der Datenbank ist zu kurz, um einen deutlichen Einfluss auf die Abfragezeit zu haben. Demnach ist die Skalierbarkeit erfolgreich umgesetzt worden. Der Flaschenhals ist die Übertragung über das Internet. Wobei mit maximalen Werten von ca. 250 Millisekunden immer noch eine ausreichend hohe Performanz gewährleistet ist.

### 6.2.3 System-Ebene

Die Funktionalität des Systems als Ganzes ist in *Systemtests* ebenfalls kontinuierlich mit Hilfe des Entwicklungsservers geschehen. Der Abnahmetest des Systems ist durch die Einbeziehung des Kunden in den Entwicklungsprozess iterativ gestaltet worden, sodass auf geänderte Anforderungen dynamisch reagiert werden konnte. Dies ist in Form von drei Präsentationen geschehen, in denen der Stand des Projekts vorgestellt wurde. Dem Kunden ist dabei die Gelegenheit für Anmerkungen und Fragen gegeben worden, sodass dessen Vorstellungen und Einwände bei der Entwicklung berücksichtigt werden konnten. Zusätzlich wurde auch zwischen den Präsentationen bei Unklarheiten Rücksprache mit dem Kunden gehalten.

## 6.3 Rahmenbedingungen

Im Anforderungskapitel wurden Musskriterien vorgestellt, welche für einen erfolgreichen Abschluss des Projektes erfüllt werden müssen. Ob und in welchem Umfang dies der Fall ist, wird im Folgenden untersucht.

- Anforderung 001F: Deckungsbeitrag berechnen: Die Berechnung des Deckungsbeitrages wurde komplett umgesetzt. Der Nutzer kann sich unter Angabe der gewünschten Filter den Deckungsbeitrag berechnen lassen, unabhängig von Größe und Herkunft der Daten.
- Anforderung 002F: Gesamtkosten berechnen: Die Berechnung der Gesamtkosten wurde mit der Berechnung des Deckungsbeitrags zusammengelegt. Die Anforderung gilt damit als erfüllt.
- Anforderung 003F: Kennzahlen in Ct/kWh: Die Kennzahlen wurden in Ct./kWh visualisiert.

## 6 Evaluation

- Anforderung 004F: Kennzahlen zeitlich aggregieren: Diese Anforderung wurde unter Absprache mit dem Kunden vernachlässigt, da lediglich die 15 Minuten Intervalle eine optimale Auswertung der Daten ermöglicht.
- Anforderung 005F: Kennzahlen mathematisch visualisieren: Eine mathematische Visualisierung ist gewährleistet, da in der tabellarischen Ansicht die einzelnen Kostenpunkte, die Gesamtkosten sowie der Profit gut erkennbar sind.
- Anforderung 008N: Einheitlichen Zugang bereitstellen: Das BICE System ist Endgeräte-unabhängig über das Internet erreichbar und somit wird ein einheitlicher Zugang bereitgestellt.
- Anforderung 008N: Korrektheit: Die Dokumentation des Systems wurde in den vorherigen Kapiteln und Abschnitten abgedeckt, darüber hinaus wurden die einzelnen Komponenten durch die zuvor erwähnten Tests geprüft und validiert.

Somit wurden alle Musskriterien innerhalb der Projektlaufzeit umgesetzt und das Projekt kann als erfolgreich deklariert werden. Über die Musskriterien hinaus wurden auch noch einige Wunschkriterien umgesetzt, welche im Folgenden analog zu den Musskriterien vorgestellt werden.

- Anforderung 006F: Ergebnis der Berechnung als Diagramm visualisieren: Diese Anforderung wurde zur Zufriedenheit des Kunden umgesetzt.
- Anforderung 007F: Datenkommunikation zwischen den Systemen: Es wurden im Projektzeitraum keine Schnittstellen zu externen Systemen entwickelt.
- Anforderung 008F: Betrag aggregieren: Diese Anforderung wurde verworfen.
- Anforderung 009F: Benutzerspezifische Arbeitsumgebung: Der Benutzer kann auf der Weboberfläche seine eigenen Kunden auswählen und die gewünschte Visualisierungsform wählen.
- Anforderung 002N: Schnelle Antwortzeiten: Mittels der verwendeten In-Memory Datenbank können die angefragten Daten in nahezu Echtzeit berechnet und aggregiert werden. Auch die Weiterleitung der Daten an den Local Client besitzt eine guten Performance. Die Anforderung wurde darüber hinaus durch den Performance Test aus dem Abschnitt Qualität der Software kontrolliert und validiert.
- Anforderung 004N: Vergleichbarkeit: Die Vergleichbarkeit der Ergebnisdaten wird aufgrund der gleichen Art und Weise der Berechnung des Deckungsbeitrages gewährleistet. Die tabellarische Darstellung der angefragten Daten eignet sich am Besten, um Vergleiche durchzuführen.
- Anforderung 005N: Branchenweit einsetzbar: Die Software ist an keinen Nutzer gebunden und kann von mehreren unterschiedlichen Energielieferanten simultan genutzt werden.

## 6 Evaluation

- Anforderung 006N: Berechtigungskonzept: Ein Berechtigungskonzept wurde aufgrund der knappen Terminierung des Projektes nicht umgesetzt.
- Anforderung 007N: Benutzerspezifische Konfiguration: Der Nutzer kann den Dienst zu jederzeit erreichen und nutzen .
- Anforderung 009N: Simultane Nutzung: Eine simultane Nutzung des Systems ist möglich und wurde durch Tests geprüft und validiert.
- Anforderung 010N: Nutzungsbedarf bestimmen: Wurde nicht erfolgreich implementiert.
- Anforderung 011N: Nutzungsbasierte Abrechnung: Wird durch das Accounting gewährleistet.

Die Auswertung der gestellten Anforderungen zeigt, dass neben den Musskriterien auch eine Reihe von Wunschkriterien umgesetzt wurden. Somit kann das Projekt BICE als erfolgreich abgeschlossen angesehen werden.

# 7 Zusammenfassung und Ausblick

Dieses Kapitel schließt das Projektdokument mit einer Zusammenfassung und einem Ausblick ab.

## 7.1 Zusammenfassung

Das BICE-Projekt begann mit der Aneignung von Grundlagen in Form von Seminararbeiten, dessen Themen dem Gebiet Business Intelligence in der Cloud und Energieversorgung zuzuordnen sind und von jedem Projektteilnehmer zu verfassen waren. Bevor mit der Realisierung des BICE-Systems begonnen werden konnte, wurde die grundlegende Projektorganisation definiert und entsprechende Funktionen und Rollen den Projektmitgliedern zugeordnet. Danach wurde die Anforderungserhebung durchgeführt. Dazu zählten Experten-Interviews, um die Anforderungen für das BICE-System zu sammeln und nach Soll-, Muss- und Wunschkriterien zu sortieren. Nun konnte die Entwurfsphase beginnen. Hierbei wurden im Speziellen UML-Diagramme angefertigt und damit eine konzeptionelle Architektur entworfen. Anhand des Entwurfs konnte im Anschluss daran die eigentliche Implementation beginnen. Dazu zählten die Installation der erforderlichen Ressourcen und das Umsetzen der einzelnen BICE-Komponenten. Abschließend wurde das Projekt im Allgemeinen sowie das BICE-System selbst evaluiert. Dabei standen vor allem verschiedene Arten von Testszenerarien im Vordergrund, um das System zu validieren. Parallel zu sämtlichen Projektschritten wurden entsprechende Dokumentationen, wie ein Anforderungsdokument, das Pflichtenheft sowie das allgemeine Projektdokument, angefertigt.

## 7.2 Ausblick

Das realisierte BICE-System verwendet beispielhafte realitätsnahe Testdaten, um die Kernfunktionen durchführen zu können. Weiterhin sollte bei einer Weiterentwicklung des Systems die Möglichkeit eines Failovers beim Ausfall einzelner virtueller Maschinen realisiert werden, was im Rahmen der Projektgruppe nicht mehr möglich war. Zu diesem Zweck muss zunächst die Möglichkeit geschaffen werden, den Session Store einzelner Clients außerhalb des eigentlichen Containers zu verwalten. Dafür gibt es mehrere Möglichkeiten.

**Tomcat Clustering** Die Clustering-Funktionalität von Tomcat kann dazu verwendet werden, die sessionrelevanten Daten auf alle andere Knoten im Cluster zu replizieren.

Für große Cluster ist dies allerdings nicht empfohlen, da dann Performanceprobleme auftreten können. Daher bietet sich die Replizierung auf einen *Backupknoten* an.

**Externe Bibliotheken** Für die Verwendung eines Tomcat Clusters ist allerdings eine spezielle Netzwerkkonfiguration notwendig, was unter Umständen unerwünscht ist. Stattdessen können auch externe Frameworks verwendet werden, die eine Sessionverwaltung anbieten. Apache Shiro ist ein solches Framework, das beliebige Speichermöglichkeiten für Sessions durch die Implementierung entsprechender Interfaces anbietet und bereits Komponenten zum Anbinden vieler Speicher-Backends mitbringt. Da es nicht gelungen ist, das Loadbalancing innerhalb des Zeitrahmens umzusetzen muss dieses bei einer Weiterentwicklung des BICE-Systems auf jeden Fall thematisiert werden. Im Folgenden werden weitere mögliche Quellsysteme und entsprechende Schnittstellen aufgeführt und beschrieben. Neben dieser grundlegenden Funktionalität kann das BICE-System noch um das Einbinden externer Quellen erweitert werden. Einige solcher Quellsysteme und die möglichen Schnittstellen werden in den folgenden Abschnitten beschrieben.

### 7.2.1 Customer Relationship Management

Für die Schnittstelle zum Customer Relationship Management (CRM) wurde ein XML-Schema definiert. Dieses könnte BICE Kunden zur Verfügung gestellt werden. Extensible Markup Language ist im Gegensatz zur Dokumententypdefinition (DTD) flexibler und XML-Dokumente können mit der Schemadefinition validiert werden.

Das XML-Dokument muss genau ein Element „customerdata“ enthalten. Das Element kann beliebig viele Elemente vom Typ „product“, maximal ein Element „contactaddress“ und genau ein Element „company“ enthalten. Weiterhin gibt es zwei Attribute. Das Attribut „customervalue“ vom Typ „positiveInteger“ beschreibt den Wert des Kunden für das Unternehmen. Das Attribut „date“ vom Typ „date“ enthält das aktuelle Datum. Das Element „product“ muss die folgenden Werte in der genannten Reihenfolge enthalten. Nach dem Semikolon ist die Anzahl in eckigen Klammern notiert.

1. Name des Produkts (name) vom Typ string; [1]
2. Wert des Produkts (value) vom Typ float; [1]
3. Beginn der Vertragslaufzeit (contractstart), falls vorhanden vom Typ date; [0,1]
4. Ende der Vertragslaufzeit (contractend), falls vorhanden vom Typ date; [0,1]

Die Kontaktdaten „contactaddress“ müssen die folgenden Informationen in der genannten Reihenfolge enthalten. Nach dem Semikolon ist die Anzahl in eckigen Klammern notiert.

1. Titel (title) vom Typ string ist optional; [0,1]
2. Nachname (lastname) vom Typ string; [1]
3. Vorname (firstname) vom Typ string; [1]

4. Straße (street) vom Typ string; [1]
5. Hausnummer ( housenumber) vom Typ string; [1]
6. Postleitzahl (zipcode) vom Typ string; [1]
7. Stadt (city) vom Typ string; [1]
8. Bundesland (state) vom Typ string; [1]
9. Staat (country) vom Typ string; [1]
10. Telefonnummer vom Typ unsigned Integer; [0,1]
11. Mobilnummer vom Typ unsigned Integer; [0,1]
12. E-Mail Adresse vom Typ string; [0,1]

Das Element „company“ enthält die Daten der Firma. Die Daten müssen in der genannten Reihenfolge genau ein Mal aufgeführt werden.

1. Name der Firma (company) vom Typ string
2. Straße (street) vom Typ string
3. Hausnummer ( housenumber) vom Typ string
4. Postleitzahl (zipcode) vom Typ string
5. Stadt (city) vom Typ string
6. Bundesland (state) vom Typ string
7. Staat (country) vom Typ string

### 7.2.2 ene't Datenbanken

Es gibt Unternehmen, die Schnittstellen zu ene't Datenbanken entwickelt haben. Eine Übersicht findet sich unter <http://www.enet.eu/importmoeglichkeiten>. ene't selbst bietet ebenfalls Analysetools zur Kalkulation der Netznutzungskosten, ect. an. Die vom System benötigten Daten liegen in den Datenbanken „Netznutzung Strom“ und „Endkumentarife Strom“. Die Datenbanken stehen als Download bei ene't zur Verfügung.

#### Netznutzung Strom

Die Datenbank kann in den folgenden Formaten abgerufen werden:

- Access2000
- MS-Structured Query Language
- CSV
- ORACLE

### Endkundertarife Strom

Diese Datenbank wird als relationale Datenbank realisiert. Sie kann in den folgenden Formaten abgerufen werden:

- Access2000
- MS-SQL
- CSV
- ORACLE

### 7.2.3 Energiedaten-Management

Auf dem Energiemarkt ist UN/EDIFACT ein gängiger Standard. Deshalb werden die Schnittstellen zu dem Quellsystem Energiedaten-Management (EDM) auf diesem Standard basieren.

Aus dem EDM stammen die Lastgänge des Kunden. Die Informationen liegen im XML Format vor. Es wird sich an der neuesten Version des UN/EDIFACT Standards orientiert (Version D.14B).

### 7.2.4 Kundendaten-Datenbank

Über die Schnittstelle zur Kundendaten-Datenbank werden die Kundendaten eingelesen. Diese werden benötigt, um den Lastprofilen die Kunden und Adressen zu zuordnen. Die Schnittstelle wird mit Hilfe eines XML-Schemas definiert. Zur Validierung wird das XML-Schema den BICE Kunden zur Verfügung gestellt. Im Folgenden wird das XML-Schema detailliert erläutert.

Das XML-Dokument muss genau ein Element „customer“ enthalten. Dieses Element kann bel. viele Adressen von Kundenstandorten enthalten, muss aber min. eine enthalten. Die Kontaktdaten sind optional. Falls ein Update der Daten erfolgt, muss die Kontaktadresse nicht neu eingelesen werden, wenn sie sich nicht verändert hat. Das Datum dient zur Überprüfung des Alters der Daten. Das Datum wird im Standardformat „date“ innerhalb der XML-Datei gespeichert. Eine „siteaddress“ muss die folgenden Elemente in der entsprechenden Reihenfolge genau ein Mal enthalten:

1. Name der Firma (company) vom Typ string
2. Straße (street) vom Typ string
3. Hausnummer (hounumber) vom Typ string
4. Postleitzahl (zipcode) vom Typ string
5. Stadt (city) vom Typ string
6. Bundesland (state) vom Typ string

### 7. Staat (country) vom Typ string

Die Kontaktdaten „contactaddress“ müssen die folgenden Informationen in der genannten Reihenfolge enthalten. Nach dem Semikolon ist die Anzahl in eckigen Klammern notiert.

1. Titel (title) vom Typ string ist optional; [0,1]
2. Nachname (lastname) vom Typ string; [1]
3. Vorname (firstname) vom Typ string; [1]
4. Straße (street) vom Typ string; [1]
5. Hausnummer (housenumber) vom Typ string; [1]
6. Postleitzahl (zipcode) vom Typ string; [1]
7. Stadt (city) vom Typ string; [1]
8. Bundesland (state) vom Typ string; [1]
9. Staat (country) vom Typ string; [1]
10. Telefonnummer vom Typ unsigned Integer; [0,1]
11. Mobilnummer vom Typ unsigned Integer; [0,1]
12. E-Mail Adresse vom Typ string; [0,1]

Die Anbindung dieser Quellsysteme ist eine Voraussetzung um weitere Prozesse der Energiewirtschaft als Cloud-Dienst anzubieten.



## 8 Glossar

### **Arbeitspreis**

Unter den Arbeitspreis fallen Kosten, die entstehen, um den Strom zu produzieren, also z.B. Rohstoffe, Transport, usw.

### **Business Intelligence in the Cloud for Energy**

BICE ist das Acronym für Business Intelligence in the Cloud for Energy. BICE ist der Name des Systems.

### **Confluence**

Confluence ist ein Wiki und zugleich ein Kollaborationswerkzeug. Im Wiki werden das von der Projektgruppe gemeinsam erlangte Wissen dokumentiert.

### **Customer Relationship Management**

Customer Relationship Management beschreibt die Verwaltung aller Interaktionen mit den Kunden. Meistens wird CRM mit Hilfe von IT-Systemen realisiert.

### **Deckungsbeitrag**

Der Deckungsbeitrag ist die Differenz zwischen den erzielten Erlösen (Umsatz) und den variablen Kosten. Es handelt sich also um den Betrag, der zur Deckung der Fixkosten zur Verfügung steht. Der Deckungsbeitrag kann sowohl auf die Gesamtmenge eines Produktes bezogen sein, als auch auf eine Mengeneinheit.

### **Energieversorgungsunternehmen**

Energieversorgungsunternehmen beliefern Endkunden mit Energie. Dazu zählen neben Strom auch Gas, Öl und Fernwärme.

### **erweiterte Ereignisgesteuerte Prozesskette**

Mit Hilfe von eEPK Diagrammen lassen sich Geschäftsprozesse grafisch darstellen.

### **Extensible Markup Language**

XML definiert eine Reihe von Regeln zur Definition von Dateien. Diese Regeln sorgen dafür, dass die Dateien sowohl von Maschinen, als auch von Menschen lesbar sind. Die Regeln wurden vom World Wide Web Consortium aufgestellt und sind ein offener Standard.

### **Gesamtkosten**

Gesamtkosten sind sämtliche Kosten, die aufgebracht werden müssen, um Strom über einen gegebenen Zeitraum an einen bestimmten Endkunden zu liefern. Sie setzen sich aus Arbeitskosten, Leistungskosten und Netznutzungskosten zusammen.

### **Hypertext Transport Protocol**

HTTP ist ein zustandsloses Protokoll, um Daten in einem Netzwerk zu übertragen. Im Internet wird es hauptsächlich zur Übertragung von Webseiten benutzt.

### **In-Memory Datenbanksystem**

Ein In-Memory Datenbanksystem nutzt zur Speicherung der Daten den Arbeitsspeicher eines Computers anstelle traditioneller Festplattenspeicher. Dadurch werden signifikant höhere Zugriffsgeschwindigkeiten erzielt. [PZ11]

### **Java Database Connectivity**

JDBC ist eine einheitliche Schnittstelle zu Datenbanken der Java Platform. Für JDBC stehen Treiber für Datenbanken verschiedener Hersteller bereit.

### **JavaScript Object Notation**

JSON ist ein Dateiformat, das zum Austausch von Daten zwischen Anwendungen verwendet wird. Jedes JSON-Objekt ist gültiger JavaScript Code.

### **JIRA**

JIRA ist ein Ticketmanagementsystem und Kollaborationswerkzeug. JIRA ist für das Scrum Entwicklungsverfahren optimiert.

### **Key Performance Indicator**

KPI's sind Kennzahlen, anhand derer der Fortschritt oder der Erfüllungsgrad hinsichtlich wichtiger Zielsetzungen oder kritischer Erfolgsfaktoren innerhalb einer Organisation gemessen und/oder ermittelt werden kann.

### **Kilowattstunde**

Wattstunde ist eine Einheit, die nicht im internationalen Einheitensystem (SI) definiert ist, aber zugelassen wurde. Deshalb ist sie eine rechtsgültige Einheit. Sie beschreibt die Energie bzw. Arbeit. Wird das Präfix Kilo verwendet, dann wird die Kilowattstunde mit 1000 multipliziert.  $1\text{kWh} \equiv 1000\text{Wh}$

### **Lastprofil**

Ein Lastprofil oder auch Lastgang oder Lastkurve ist ein Diagramm, das auf der Ordinate die abgenommene Leistung gegenüber der Zeit, auf der Abszisse aufträgt. Lastprofile werden in der Energiewirtschaft u.a. zur Prognose von Energiebedarf verwendet.

### **Leistungspreis**

Der Leistungspreis (auch als Grundpreis bezeichnet) ist der Preis für die vom Versorger bezogenen Leistungen in Cent je kWh. Je nach Preismodell wird die höchste gemessene Leistung des Jahres (Jahresleistungspreis) oder die eines Monats (Monatsleistungspreis) berechnet. Z.B. Personal, Unterhaltungskosten.

### **Netznutzungskosten**

Netznutzungskosten sind Gebühren, die von Netzbetreibern für die Nutzung des jeweiligen Stromnetzes erhoben werden. Diese Gebühren sind in den Preisen, die der Endverbraucher an seinen Anbieter bezahlt, enthalten. Netznutzungskosten sind je nach Auslastung dynamisch.

### **Representational State Transfer**

REST ist ein Programmierparadigma für verteilte Systeme.

### **Structured Query Language**

Bei SQL handelt es sich um eine Sprache zur Verwendung von Datenbanken. Neben dem Einfügen, Löschen oder Verändern von Daten ist auch die Definition des Schemas möglich. SQL ist ein Standard in diesem Bereich und wird von einem Großteil der Datenbanken unterstützt.

### **SVN**

Apache Subversion ist eine zentrale Versionsverwaltung für Dateien und Verzeichnisse.

### **Virtual Machines**

Als virtuelle Maschine (kurz VM) wird die Nachbildung eines Rechnersystems bezeichnet. Die virtuelle Maschine bildet die Rechnerarchitektur eines real in Hardware existierenden oder hypothetischen Rechners nach. [Cra06]

## 9 Literatur

- [Cra06] Iain D. Craig: *Virtual Machines*. Springer, 2006.
- [Fav] Max Favilli: *Why new SAP UI will be a fiasco*. URL: <https://maxfavilli.com/sapui5> (besucht am 04.10.2015).
- [Ham09] Oliver Hammerstein: "SCRUM". In: *projektMANAGEMENT aktuell* 4 (2009), S. 28–32.
- [NS14] Oliver Norkus; Jürgen Sauer: "BI in der Cloud - Einordnung, Nutzen, Potenziale". In: *ERP Management* (2014).
- [PZ11] Hasso Plattner; Alexander Zeier: *In-Memory Data Management: An Inflection Point for Enterprise Applications*. 1. Aufl. Springer, 2011.

# A Benutzerhandbuch

## A.1 Login

Auf der Startseite des BICE-Systems wird ein Login-Fenster angezeigt. Hier müssen der Username sowie das zugehörige Passwort in die dafür vorgesehenen Input-Felder eingetragen werden und die Eingabe mit Betätigen des Login-Buttons bestätigt werden. War die Eingabe erfolgreich, wird man auf die Calculation-Seite weitergeleitet. War die Eingabe fehlerhaft, wird eine Fehlermeldung ausgegeben und die Daten können erneut eingegeben werden.

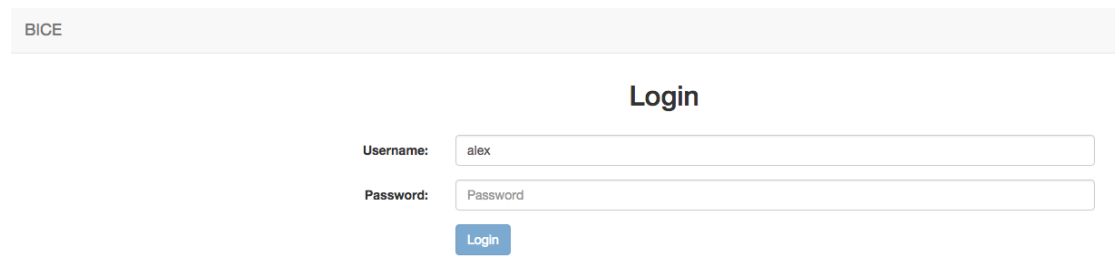


Abbildung A.1: Login-Page

## A.2 Calculation-Page

Auf der Calculation-Seite gibt es mehrere Eingabefelder, welche für die Datenaggregation ausgefüllt werden müssen. Diese sind wie folgt auszufüllen:

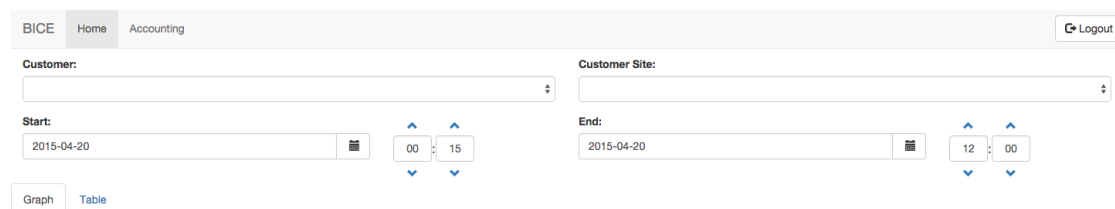


Abbildung A.2: Calculation-Seite

1. Kundenauswahl: Der Nutzer wählt zuerst einen Kunden aus dem Dropdown-Menü Customer aus.

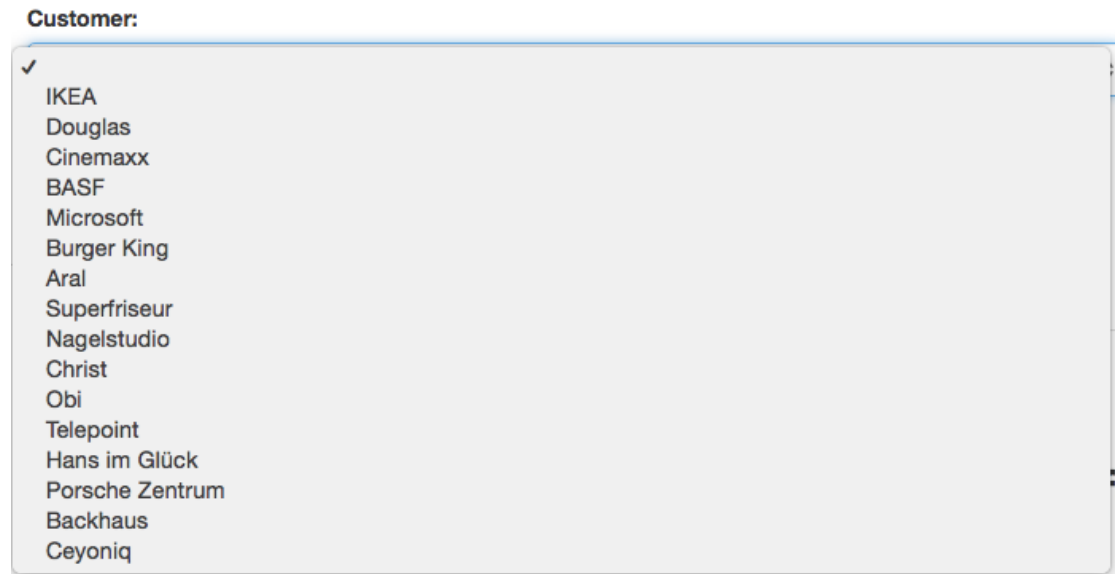


Abbildung A.3: Customer-Dropdown

2. Kundenstandortauswahl: Nach der Auswahl des Kunden muss der Benutzerstandort aus dem Dropdown-Menü Customer Site gewählt werden.

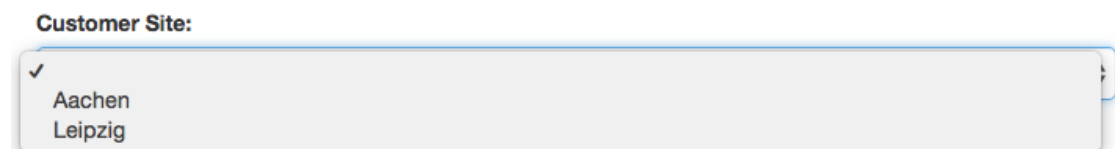


Abbildung A.4: Customer Site Dropdown

3. Zeitraumauswahl: Das Start-Datum sowie -Uhrzeit und das End-Datum sowie -Uhrzeit können hier festgelegt werden
4. Datumsauswahl: Der Nutzer kann mittels der Reiter unterhalb der Auswahlfelder zwischen grafischer und tabellarischer Darstellung wählen. Auf dem Graphen können die Werte eines Zeitpunkts via Mouseover detailliert angezeigt werden.

### A.3 Accounting

Bestätigt man im Headermenu den Button Accounting, wird man auf die Accounting-Seite weitergeleitet. Hier kann sich der Nutzer die in Anspruch genommenen Ressourcen

## A Benutzerhandbuch

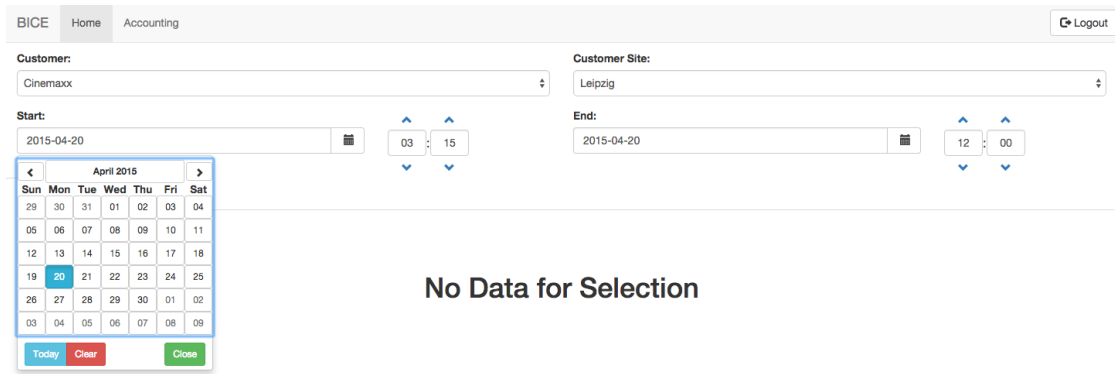


Abbildung A.5: Datepicker

über einen bestimmten Zeitraum anzeigen lassen. Angezeigt werden sowohl die Einheiten, welche in Mikrosekunden angezeigt werden, als auch die anfallenden Kosten in €. Nach Bedarf kann mittels des Dropdown-Menüs auch die Granularität der angezeigten Daten verändert werden (Tag, Monat, Jahr).

### A.4 Logout

Mit Betätigen des Logout-Buttons wird man sicher vom System ausgeloggt und muss sich für weitere Berechnungen neu anmelden.

# A Benutzerhandbuch

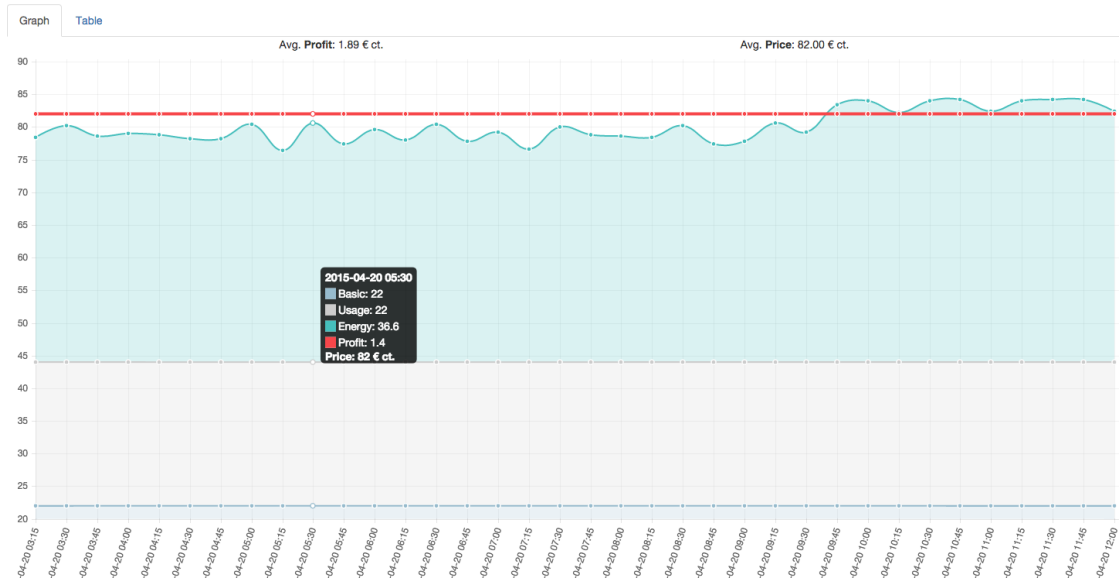


Abbildung A.6: Graph

Timestamp	Basic Cost	Energy Cost	Usage Cost	Total Cost	Profit
2015-04-20 03:15	22 € ct.(28.06%)	34.4 € ct.(43.88%)	22 € ct.(28.06%)	78.4 € ct.	3.6 € ct.
2015-04-20 03:30	22 € ct.(27.43%)	36.2 € ct.(45.14%)	22 € ct.(27.43%)	80.2 € ct.	1.8 € ct.
2015-04-20 03:45	22 € ct.(27.99%)	34.6 € ct.(44.02%)	22 € ct.(27.99%)	78.6 € ct.	3.4 € ct.
2015-04-20 04:00	22 € ct.(27.85%)	35 € ct.(44.30%)	22 € ct.(27.85%)	79 € ct.	3 € ct.
2015-04-20 04:15	22 € ct.(27.92%)	34.8 € ct.(44.16%)	22 € ct.(27.92%)	78.8 € ct.	3.2 € ct.
2015-04-20 04:30	22 € ct.(28.13%)	34.2 € ct.(43.73%)	22 € ct.(28.13%)	78.2 € ct.	3.8 € ct.

Abbildung A.7: Tabelle

BICE Home Accounting Logout

User: alex  
Mandator: Lieferant1

Start: 2015-09-06 End: 2015-10-04

Granularity: Day

Day	Units	Price
2015-10-3	52898 µs	5.28 €
2015-10-4	71599 µs	7.15 €
<b>Total</b>	<b>124497 µs</b>	<b>12.44€</b>

Abbildung A.8: Accounting-Seite



## B Seminararbeiten

Auf den folgenden Seiten befinden sich jeweils die vollständigen Seminararbeiten der Mitglieder der Projektgruppe. Folgende Themen wurden behandelt:

- Big Data - Grundlagen, Einsatzgebiete und Technologien
- In Memory Computing
- Energiewirtschaft und Energiewende
- Systeme in der Energiewirtschaft
- Grundlagen: Business Intelligence
- Scrum - Agiles Projektmanagement

# Seminararbeit

## „Big Data - Grundlagen, Einsatzgebiete und Technologien“

Ilias Schweer

Department für Informatik  
Abteilung Informationssysteme  
ilias.schweer@uni-oldenburg.de  
Matrikelnummer: 1457587  
Betreuer: Ralf Krause

**Abstract:** Big Data ist ein abstrakter Oberbegriff, der Daten beschreibt, die bezüglich Volumen, Varietät, Schnelligkeit (des Erzeugens und der Verarbeitung) und der Richtigkeit, klassische Technologien und Methoden an ihre Grenzen stoßen lassen. Ziel von Big Data ist es, eine große Menge von größtenteils unstrukturierten Daten zu sammeln und möglichst in Echtzeit zugänglich zu machen, zu verarbeiten und daraus nützliche Informationen zu gewinnen. Dieser Prozess beinhaltet die Phasen Erwerb und Aufzeichnung, die Extraktion und Bereinigung, die Aggregation und Integration, die Analyse und Modellierung und abschließend die Interpretation der Daten. Dabei sind einige Herausforderungen, die die finanziellen Aspekte der Realisierung, Know-How und allgemein technologische Aspekte der Umsetzung von Big Data betreffen, zu berücksichtigen. Big Data findet heutzutage in vielen Bereichen, wie der Wirtschaft, im Gesundheitswesen und in der Forschung große Anwendung bei Echtzeitanalysen, Entscheidungsprozessen sowie Personalisierungen. Die technologische Umsetzung findet dabei in der heutigen Zeit vermehrt mit dem Apache Hadoop Framework statt, welcher mithilfe des MapReduce Algorithmus große Datenmengen strukturieren und somit viele der Anforderungen von Big Data erfüllen kann.

## 1 Einleitung

In der heutigen Zeit zeichnet sich eine deutliche Entwicklung zu einer Vernetzung der Bevölkerung der Welt durch mit dem Internet verbundenen Endgeräten ab. Ein Großteil der Menschen auf diesem Planeten verfügen über solche mobile Endgeräte und kommunizieren ausgeprägt in dieser Form. In der Forschung werden heutzutage Satellitendaten ausgewertet, umfangreiche Experimente mithilfe von technologischer Ausrüstung durchgeführt sowie im Gesundheitswesen Millionen von Patientendaten gesammelt. Diese aufgeführten Entwicklungen sorgen primär durch die sehr hohe Anzahl an Teilnehmern der Verbindungen für ein enormes Datenaufkommen sowohl von Nutzdaten, als auch von Profildaten und statistischen Daten. So sind 90% der gesamten Daten der Weltgeschichte im Zeitraum von 2010 bis 2012 entstanden. Allein das Wachstum des Datenvolums im Jahre 2012 auf das Vorjahr bezogen betrug über 40%. In all diesen erzeugten Daten befindet sich ein erhebliches Potential unter anderem für Unternehmen, einen Mehrwert und somit

Wettbewerbsvorteile zu generieren. In der Rohform allerdings eher nutzlos, können durch analytische Prozesse, Modellierungstechniken und geeigneter Technologien diese Daten in nützliche Informationen transformiert werden. In diesem Bereich spielt der Begriff Big Data eine wesentliche Rolle. Genau wie zum Beispiel Cloud Computing stellt Big Data heutzutage einen der größten Technologie-Antreiber dar. [Kin14, KTG13] Ziel dieser Arbeit ist es, dem Leser einen allgemeinen Eindruck und wesentliche Kenntnisse zu Big Data und den umliegenden Aspekten zu vermitteln. Zunächst werden die theoretischen Grundlagen behandelt. Danach wird auf die Phasen und technologischen Herausforderungen von Big Data eingegangen. Im Anschluss daran erfolgt eine praktische Sichtweise auf Big Data mit Einsatzgebieten und Potentialen. Hiernach werden Technologien bezüglich Big Data behandelt und als Abschluss dieser Arbeit erfolgt ein Fazit mit Ausblick.

## 2 allgemeine Begriffsbildung

Der Begriff Big Data wurde erstmals 1970 genannt. In der damaligen Zeit wurde der Begriff Big Data von Autoren zumeist im Sinne von großen Datensätzen im Bereich der Computermodellierung und Software- oder Hardwareentwicklung gebraucht. [Kin14] In der heutigen Zeit behandelt der Begriff Big Data einen viel weitreichenderen Bereich. Big Data stellt dabei ein „kulturelles, technologisches und wissenschaftliches Phänomen [dar], das sich aus kommunikationswissenschaftlichen, informationstechnologischen und betriebswirtschaftlichen Komponenten zusammensetzt.“ [Kin14] DeRoos definiert Big Data wie folgt: „extracting insight from an immense volume, variety and velocity of data, in context, beyond what was previously possible“. [DeR12] Im Allgemeinen werden durch Big Data Daten beschrieben, die im Bezug auf Ausmaß (Menge), Geschwindigkeit, Struktur und Varietät bisherige (bzw. konventionelle) Datenbanksysteme überfordern. [Kin14] Eine weitere Definition von Big Data der *Federal Big Data Commission* der *TechAmerica Foundation* lautet: „Big Data is a term that describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information.“ [Com12]

Vier Charakteristika zeichnen Big Data aus: [Kin14, Win11, Com12, BBAa11]

1. Umfang (**Volume**): Eine sehr hohe Menge an Daten, die aufgenommen und analysiert werden muss, um daraus nützliche und wertvolle Informationen zu generieren, um daraus wiederum Entscheidungen treffen zu können. Dazu werden entsprechend hohe Hardwareressourcen benötigt. Dabei steigt der Umfang mit der Anzahl an Quellen. Treiber hierfür sind Steigerungen der Anzahl der Datenquellen sowie der Datentiefe und Auflösung.
2. Varietät (**Variety**): Die Daten sind nicht einheitlich strukturiert. Sie stammen aus verschiedenen und neuen Quellen innerhalb sowie ausserhalb der Organisation und die Struktur der Daten kann stark variieren. Ebenfalls ist es möglich, dass die Strukturierungsform vorher unbekannt ist. Die Varietät bezieht sich auf die Darstellungsart, die Heterogenität des Datentyps und der semantischen Interpretation. Genauer

können Daten in drei Strukturierungsvarianten unterteilt werden (siehe auch Abb. 1):

- **strukturierte** Daten können in einer herkömmlichen relationalen Datenbankstruktur (Reihen und Spalten) gespeichert werden. Beispiele für strukturierte Daten sind Transaktionsdaten, Daten aus dem Enterprise Resource Planning und Customer Relationship Management.
  - **semi-strukturierte** Daten sind zum Teil strukturiert, enthalten jedoch auch unstrukturierte Elemente. Standardisierte Tools können diese unstrukturierten Elemente nicht verarbeiten, weshalb komplexere Verfahren benötigt werden. Beispiele hierfür sind Weblogs, Umfragen, E-Mails und Leistungsbewertungen.
  - **unstrukturierte** Daten benötigen zur Verarbeitung, Zählung, Evaluierung sowie Filtration neue Technologien. Zu dieser Variante kann man zum Beispiel Internetdaten wie Kundenrezensionen, Blogs, Daten aus sozialen Netzwerken und Wetterdaten zählen.
3. Schnelllebigkeit (**Velocity**): Die Daten kommen in einer hohen Geschwindigkeit in das Datenmanagementsystem und müssen dort schnell analysiert, verstanden und weiterverarbeitet werden, um Entscheidungen treffen zu können. Treiber hierfür sind ebenfalls das Ansteigen der Datenquellen, verbesserte Rechnerleistung von datengenerierenden Geräten und die stärkere und ausgeprägtere Konnektivität von Geräten.
4. Richtigkeit (**Veracity**): bezeichnet die Qualität und Quelle der aufgenommenen Daten. Treiber hierfür ist, dass datenbasierte Entscheidungen Nachvollziehbarkeit, Begründbarkeit und Rückverfolgbarkeit erfordern und somit gerechtfertigt sein sollten.

Im Rahmen von Big Data ist es nun Ziel, sämtliche Daten zu sammeln, ob strukturiert oder nicht, zusammenzufassen und zu analysieren, um aus diesen nützliche und wertvolle Informationen zu gewinnen, um damit wiederum Entscheidungen treffen zu können. [Bar13] Die dabei entstehenden unstrukturierten Datenzusammenführungen können anhand drei Kategorien klassifiziert werden: [KTGH13]

- **Mensch-zu-Mensch** Daten, die aus der Kommunikation zwischen Menschen entstehen. Dazu zählen zum Beispiel Daten aus sozialen Netzwerken wie Facebook.
- **Mensch-zu-Maschine** Daten, die aus der Kommunikation zwischen Menschen und Maschinen oder Diensten entstehen. Beispiele hierfür sind Daten aus der Nutzung von Geldautomaten oder Daten aus der Nutzung von E-Commerce Anwendungen (wie Amazon).
- **Maschine-zu-Maschine** Daten, die aus der Kommunikation zwischen Diensten und Maschinen entstehen. Hierunter sind zum Beispiel Sensordaten oder GPS Positionsinformationen zu nennen.

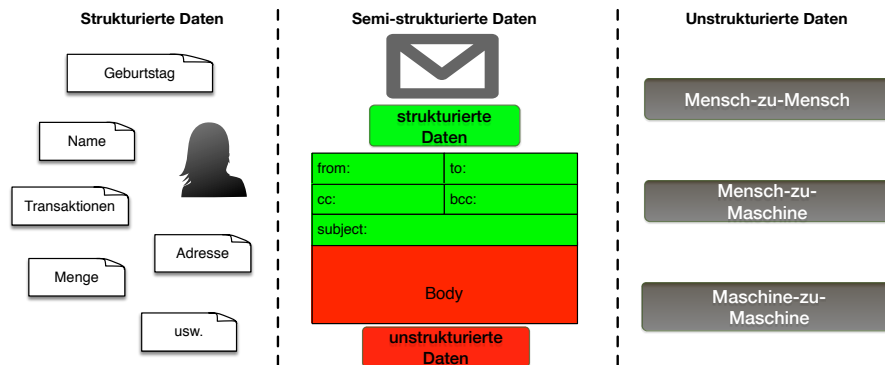


Abbildung 1: Datenvielfalt in Anlehnung an [KTGH13]

Wie bereits erwähnt haben die gesammelten Rohdaten im Rahmen von Big Data keinen signifikanten Nutzen. Erst durch analytische Prozesse und Modellierungstechniken können aus den Rohdaten Schlüsse gezogen werden, die einen Mehrwert generieren können.

[Kin14] Der *Vice President of Engineering* von Facebook, Jay Parikh, äußerte sich 2012 dazu wie folgt: „Big data really is about having insights and making an impact on your business. If you aren't taking advantage of the data you're collecting, then you just have a pile of data, you don't have big data.“ [Constine 2012] Zusammenfassend handelt es sich bei Big Data um ein Vorgehen, bei dem aus vielen und unterschiedlichen Datenquellen große Mengen Daten, die unterschiedliche Strukturierungsformen aufweisen können, mit einer hohen Geschwindigkeit gesammelt und weiterverarbeitet werden, um daraus mithilfe analytischer Prozesse nützliche Informationen zu extrahieren.

## 2.1 Abgrenzung zu Business Intelligence

Business Intelligence (BI) und Big Data verfolgen teils ähnliche Ziele und tauchen entsprechend oft zusammen auf, weshalb in diesem Abschnitt eine kurze Abgrenzung dieser beiden Begriffe aufgeführt wird. Business Intelligence wird als „[...] betriebliche Entscheidungsunterstützung durch einen integrierten [und] auf das Unternehmen bezogenen IT-basierten Gesamtansatz“ [GT09] definiert. Die im Rahmen von BI verwendeten Daten werden aus verschiedenen Abteilungen einer Organisation extrahiert, transformiert und in einem Data Ware House (DWH) abgelegt. Wie auch bei Big Data können nun mithilfe von Modellierungstechniken und analytischen Prozessen, wie zum Beispiel Data Mining, aus den im DWH gelagerten Daten nützliche und zur unternehmerischen Entscheidungsfindung relevanter Informationen generiert werden. Der Unterschied zwischen BI und Big Data liegt auf der Ausrichtung der verwendeten Daten. BI-Lösungen setzen strukturierte, konsistente und beständige Daten voraus. Big Data Systeme dagegen sind speziell auf die

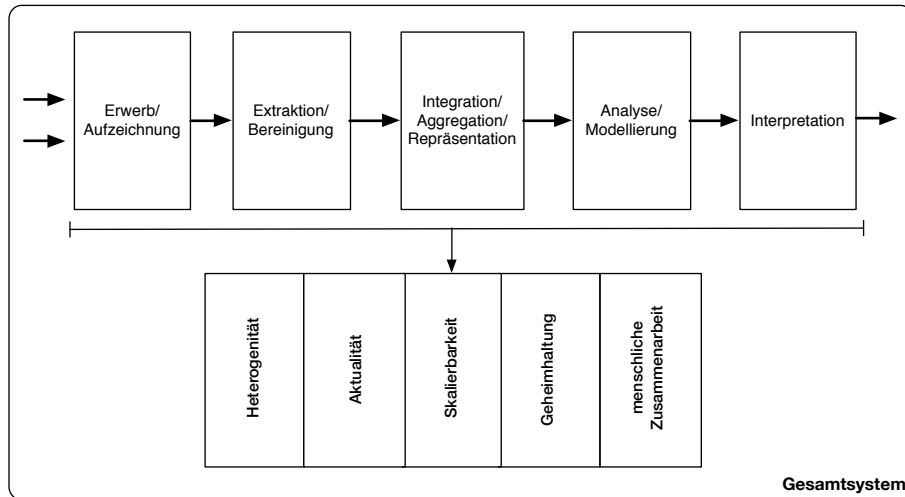


Abbildung 2: Big Data Prozess und Herausforderungen in Anlehnung an [BBAa11]

zuvor in dieser Arbeit behandelten Arten von Daten ausgelegt. Daraus ergeben sich Unterschiede in den von Big Data und BI verwendeten Technologien und Techniken. Bei sehr großen und unstrukturierten Datenmengen reichen die von BI verwendeten konventionellen relationalen Datenbanksysteme nicht mehr aus, sodass neuere Konzepte notwendig sind (siehe dazu fünftes Kapitel). [KTGH13]

### 3 Phasen des Big Data Prozesses und technologische Herausforderungen

Im folgenden Abschnitt wird der genaue Ablauf des Big Data Prozesses behandelt und die einzelnen Phasen erläutert. Abbildung 2 visualisiert den Ablauf des Big Data Prozesses und die dabei zu beachtenden Herausforderungen.

1. **Erwerb und Aufzeichnung von Daten:** Die Rohdaten, die in ein Big Data System aufgenommen werden, entstehen nicht aus dem Nichts. Die Daten werden in der heutigen Welt aus einer enormen Anzahl von Datenquellen aufgezeichnet und aufgenommen. Dazu zählen beispielsweise Teleskope, die Millionen von Terabyte Daten pro Tag generieren oder aber wissenschaftliche Experimente und Simulationen, die ebenfalls Petabytes von Daten produzieren können. Dabei ist es möglich, dass ein Großteil dieser Daten keine weitere Relevanz hat. Hier gilt es diese Rohdaten intelligent zu filtern und in der Größe zu komprimieren, ohne dabei nützliche und wertvolle Daten unbewusst auszufiltern. Ebenfalls spielt die automatische Generierung von Metadaten für die weiteren Phasen eine wichtige Rolle. Diese Meta-

Herausforderung	Anteil in %
Bewältigung der Datenmenge	45
Hoher Datenverkehr (Netze & Router)	38
Datenverwaltung (Archivierung & Retrieval)	34
Speicherung und Zugriff strukturierter Daten (Storage)	34
Master-Data-Management (Datenkonsolidierung)	27
Speicherung und Zugriff unstrukturierter Daten (Storage)	25

Tabelle 1: Herausforderungen bei Big Data in Anlehnung an [Kin14]

daten sollten Auskunft darüber geben, welche Daten wie und wo aufgenommen und gemessen wurden. [BBAA11]

2. **Extraktion und Bereinigung:** Die nun gewonnenen Daten befinden sich häufig in einem für weitere Analysen ungeeigneten Format. Diese Eigenschaft der Daten erschwert eine effiziente Analyse. Aus diesem Grund ist es notwendig mithilfe von Extraktionsprozessen die benötigten Informationen herauszuziehen und in einer strukturierten Form für weitere Analysen auszudrücken. [BBAA11]
3. **Integration, Aggregation und Repräsentation:** Der nächste Schritt stellt die Lagerung der Daten dar. Dadurch, dass die Daten heterogene Eigenschaften aufweisen, ist es nicht möglich die enormen Mengen von Daten in einem konventionellen Datenbanksystem zu laden. [BBAA11] Geeignete Technologien werden im fünften Kapitel dieser Arbeit näher betrachtet.
4. **Analyse und Modellierung:** Als nächstes folgt die Analyse der Daten, um wertvolle Informationen zu generieren. Dies geschieht mithilfe von Analysen- und Modellierungstechniken (wie mathematische Algorithmen). [BBAA11, Com12]
5. **Interpretation:** Abschließend müssen die gewonnenen Informationen verstanden werden, um sie dann interpretieren zu können und dadurch Entscheidungen auf Grundlage dieser Informationen und Modelle zu treffen. [BBAA11]

Auf diese Phasen des Big Data Prozesses wirken verschiedene Herausforderungen, die die Anforderungen und die Durchführung von Big Data betreffen können. Im Rahmen dieser Arbeit werden ausschließlich technologische und technische Herausforderungen aufgeführt. Herausforderungen bezüglich kultureller und gesellschaftlicher, ethischer, organisatorischer sowie der Rechtslage betreffender Aspekte im Rahmen von Big Data werden aus Gründen des Umfangs dieser Arbeit nicht mit einbezogen.

Die Herausforderungen, die bei Big Data auftreten können, hängen davon ab, welche Ziele die Organisation mithilfe von Big Data Systemen lösen möchte. Die Wahl der richtigen Technologien und Ressourcen stellt dabei bereits vor der Nutzung von Big Data ein erhebliches Problem für viele Unternehmen dar. Das resultiert daraus, dass es bisher keine Einheitslösung für Big Data Projekte gibt und deshalb eine individuelle Konfiguration

und Inbetriebnahme notwendig ist. Ebenfalls erfordern Big Data Systeme einen erheblichen finanziellen Aufwand sowie ein hohes Know-How der IT-Abteilung, welches viele Organisationen nicht aufbringen können. Tabelle 1 zeigt auf, welche Herausforderungen bei den befragten deutschen Unternehmen laut einer Studie der International Data Corporation (IDC) im Jahre 2012 bei Big Data am ehesten auftreten (Mehrfachnennung war möglich). Unternehmen empfinden die enormen Datenmengen und allgemein die hohe Komplexität von Big Data als hinderlich. Dies betrifft zum einen, wie bereits aufgezeigt, die Realisierung, also die Inbetriebnahme von Big Data Systemen, als auch die Nutzung eines schon eingerichteten Systems. So können viele Unternehmen keinen Nutzen aus Konsumentendaten generieren und haben Probleme bei der Speicherung der Daten. Ebenfalls sehen viele Unternehmen den Schutz der Informationen als problematisch an. Gerade in dem Zeitraum kurz nach Realisierung eines Big Data Systems sind sogenannte Zero-Day-Attacken kritisch, da in diesem Zeitraum mögliche Sicherheitslücken unbekannt sind und lange unentdeckt bleiben können. Die Auswertung der Daten, also Analyseprozesse, stellen dabei eine weitere Herausforderung dar. Diese sind teils nicht voll ausgereift und erfordern Verbesserungen hinsichtlich Aggregationslevel, der Erkennung von Mustern und der Vorhersage von Verhalten. [Kin14]

## **4 Big Data in der Praxis**

Dieses Kapitel behandelt den praktischen Aspekt von Big Data und geht dabei auf Potentiale und Einsatzgebiete von Big Data in der heutigen Unternehmenswelt ein.

Um eine etwas genauere Vorstellung zu erhalten, um welche enormen Datenmengen es sich im Kontext von Big Data handelt, beginnt das Kapitel mit einigen statistischen Informationen bezüglich Big Data.

Im Jahre 2011 betrug der gesamte Umfang von Big Data 1,8 Trillion Gigabyte (entspricht 500 Milliarden Dokumenten) und laut Schätzungen verdoppelt sich diese Menge alle zwei Jahre. Rein auf Unternehmensdaten bezogen nimmt in den nächsten Jahren der Umfang dieser um 650% zu, wobei davon 80% unstrukturierte Daten ausmachen werden. Weiterhin wird im Laufe des nächsten Jahrzehnts die zehnfache Anzahl an Servern zur Speicherung dieser Daten notwendig sein. [Kin14, KTGH13]

Allgemein bietet Big Data den Vorteil, Produktivitäts- und Umsatzsteigerungen bei Unternehmen und Organisationen zu erzielen. Laut einer Studie der International Business Machines Corporation (IBM) generierten Unternehmen, welche Big Data erfasst, analysiert und dadurch Vorhersagen treffen konnten, eine Umsatzsteigerung von 10,8% und konnten ebenfalls eine Aufwertung des Aktienkurses um 7,4% verzeichnen. Die Ergebnisse einer Studie des MIT Sloan Management Review in Kooperation mit dem IBM Institute for Business Value implizieren ebenso wertvolle Potentiale beim Einsatz Big Data: Unter 3000 Führungskräften und Analysten aus Organisationen in 108 Staaten aus 30 verschiedenen Branchen wurde erhoben, welche Bedeutung Datenanalyse bezüglich Wettbewerbsfähigkeit und Unternehmenserfolg haben. So nutzen erfolgreiche Unternehmen doppelt so häufig Datenanalytik als ihre schwächeren Konkurrenten in derselben Branche



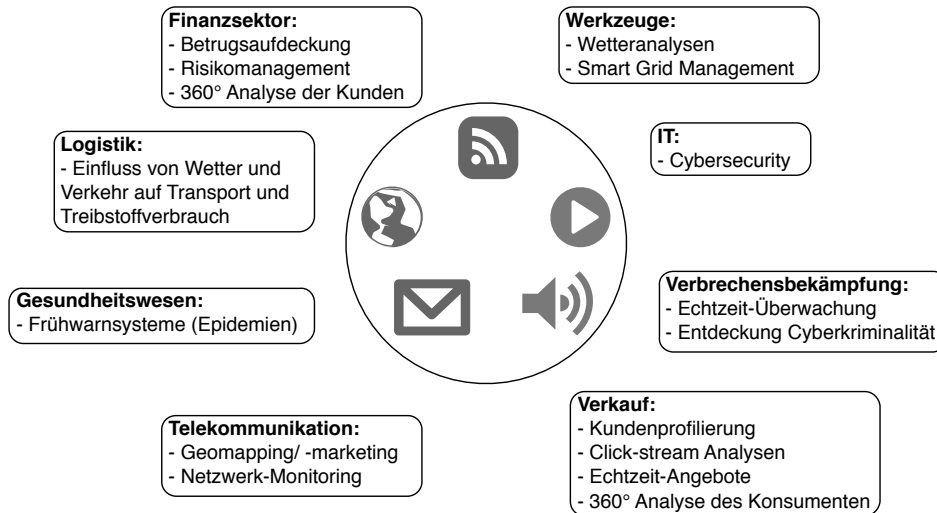


Abbildung 3: Auswahl an Einsatzgebieten von Big Data in Anlehnung an [DeR12]

und haben dadurch erhebliche Vorteile beim Wachstum, bei der Leistungsfähigkeit und der Wettbewerbsfähigkeit. Allgemein entscheidend ist dabei das benötigte Know-How. Unternehmen mit Datenanalysten gelingt es häufiger Umsatz durch Big Data zu erzielen, als Unternehmen ohne Datenanalysten. [Kin14]

Abbildung 3 zeigt die enorme Fülle der Einsatzgebiete von Big Data. Diese reichen vom Finanzsektor über das Gesundheitswesen bis zur Logistik. Diese Auswahl nennt allerdings nur einige Beispiele innerhalb der Einsatzgebiete. Im Folgenden werden vier allgemeinere Einsatzgebiete von Big Data näher erläutert: [Kin14, Win11, Com12]

- **Echtzeitanalysen und -experimente:** Mithilfe von Big Data können Organisationen Echtzeitanalysen und -experimente durchführen, um zum Beispiel die Auswirkungen von Veränderungen bei Produkten und Services hinsichtlich Durchlaufzeit, Leistung und Kosten zu erkennen. Unternehmen wie Amazon und Google nutzen solche Analysen dazu, in Echtzeit Personalisierungen ihrer Webseiten anhand bestimmter Nutzerkriterien, wie Geschlecht und Nationalität, zu ermöglichen. Im Telekommunikationssektor werden Echtzeitanalysen genutzt, um Probleme der Konnektivität oder der Bandbreite frühstmöglich zu erkennen und zu verhindern.
- **Personalisierungen:** Durch Big Data können Personen, Objekte und Orte analysiert werden. Hierbei werden zum Beispiel persönliche Attribute, Beziehungen, Gefühle und Produktinteressen zusammengeführt, um eine 360-Grad-Sicht zu gewinnen (360-Grad-Analyse). Quellen für solche Erkenntnisse können Click-Streams, E-Mails, Aufzeichnungen von Call-Center-Anrufen, Überwachungsvideos aus Kaufhäusern

und Social Media sein. Primäre Ziele, die durch diese 360-Grad-Analysen erreicht werden sollen, sind die Entwicklung von gezielten Marketingstrategien, die Optimierung von Cross- und Up-Selling-Möglichkeiten sowie die Verhinderung von Abwanderungen. Diese Profilerstellung wird auch von Versicherungen genutzt, um Versicherungskosten zu prognostizieren und individuelle Angebote zu erstellen. Ein weiteres Beispiel für Individualisierung des Angebots stammt aus dem Gesundheitswesen. Durch Analyse von Patientendaten kann ermittelt werden, welche Behandlung bei spezifischen Krankheiten am wirksamsten ist. Hierbei werden die Ergebnisse der angewandten Behandlungsmethoden verglichen. Dazu kann der gesamte Pool an Patientendaten Ärztinnen und Ärzten in Echtzeit zugänglich gemacht werden, um die Behandlung zu verbessern und die Kostenstruktur effizienter zu gestalten.

- **Innovationen:** Organisationen, wie Telekommunikationsanbieter, Banken und Einzelhändler, können mithilfe von Big Data Daten aus Social Media Beiträgen auswerten und dazu nutzen, neue Markttrends zu identifizieren und damit neue, darauf zugeschnittene Services und Produkte zu entwickeln und anzubieten. Hierbei kann das Feedback der Konsumenten, welches passiv oder aktiv preisgegeben wird, direkt in die Entwicklung neuer Innovationen mit einfließen.
- **Planungs- und Entscheidungsprozesse:** Ein sehr häufig erwähntes Einsatzgebiet stellt die durch Big Data optimierte Unterstützung bei der Entscheidungsfindung für Entscheidungsbefugte in Organisationen dar. Steuerbehörden aus den USA nutzen automatisierte Risikoalgorithmen, um Anomalien in den Steuerdaten zu erkennen und die betroffene Person zur weiteren Prüfung zu markieren. Ebenfalls können durch Big Data Planungsprozesse unterstützt und verbessert werden. Dies findet zum Beispiel bei gemeinnützigen Organisationen Einsatz, die durch Big Data den Ressourceneinsatz optimieren können. Dabei können detaillierte Momentaufnahmen der Bevölkerung gewonnen werden, um den aktuellen Status der Bevölkerung zu erkennen. Damit können Krisenausbrüche prognostiziert und besser verstanden werden und die Hilfsgüterlogistik verbessert werden.

## 5 Technologien

Dieses Kapitel behandelt den technologischen Aspekt von Big Data. Wie in den vorherigen Kapiteln bereits erläutert, reichen für die Anforderungen, die Big Data stellt, traditionelle und klassische SQL-Datenbanksysteme nicht aus. Zu groß ist das Volumen der Daten, zu schnell kommen diese Daten an und müssen weiterverarbeitet werden. Um diese Anforderungen zu befriedigen, mussten die bestehenden technologischen Konzepte überdacht und weiterentwickelt werden. Wie auch bei klassischen Datenbanksystemen werden im Rahmen von Big Data riesige Server-Cluster verwendet, um die Aufgaben verteilt und gemeinsam zu lösen. Jeder teilnehmende Server übernimmt einen Teil der Daten, sodass der Prozess parallel und schnell durchgeführt werden kann. Weiter müssen die Server des Clusters kompakt und preiswert sein, da bis zu hunderttausend Server benötigt werden. Die Kosten für solch einen Server-Cluster wären ansonsten beträchtlich hoch. Einen weiteren

Aspekt stellt die auf den Server-Clustern operierende Anwendung dar. Diese muss mit der verteilten Server-Infrastruktur und den enormen Datenmengen arbeiten können.[Win11] Zusammenfassend werden die Schlüsselideen hinter Big Data nun aufgelistet, dabei steht I/O für In- and Output: [WSV14]

- Die Daten werden auf mehrere Server innerhalb des Clusters verteilt (Netzwerk I/O Geschwindigkeit  $\ll$  Lokaler Datenträger I/O Geschwindigkeit).
- Die Anwendungen werden in Richtung der Daten verteilt und nicht umgekehrt (die Daten haben ein wesentlich größeres Volumen als die Anwendungen).
- Soweit wie möglich finden die Datenprozesse lokal (auf den Servern im Cluster) statt und laufen nicht über das Netzwerk (Netzwerk I/O Geschwindigkeit  $\ll$  Lokaler Datenträger I/O Geschwindigkeit).
- Zufälliger Datenzugriff wird durch sequentiellen Datenzugriff ersetzt (Übertragungsrates  $\ll$  Datenträgerzugriffszeit).

Allgemein betrachtet werden die bisher genannten Anforderungen und benötigten Eigenschaften unter anderen von sogenannten *Not Only SQL* (NOSQL) Datenbanksystemen erfüllt und bereitgestellt. NOSQL Datenbanksysteme sind für Vorhaben konzipiert, bei denen relationale Datenbanken nicht mehr alleine ausreichen. Sie haben ihre Stärke bei verteilten Systemen mit großen Datenmengen, bei denen kaum Änderungen der Datenbestände auftreten, sondern meist ausschließlich stetig neue hinzugefügt werden (zum Beispiel Twitter Posts oder Internet Server Protokolldateien). NOSQL ist dabei ein Oberbegriff für Datenbanksysteme und kann in drei Varianten gegliedert werden: [KTGH13]

- **Dokumenten-orientierte Datenbanken** sind speziell für die Speicherung von Dokumenten konzipiert, also für semi-strukturierte Daten. Sie erlauben das Suchen innerhalb der Dokumente. Beispiele für solche Datenbanken lauten MongoDB, BaseX und Apache CouchDB.
- **Graphen-orientierte Datenbanken** sind dahingehend optimiert, stark vernetzte Informationen zu speichern und die Beziehung zwischen Entitäten darzustellen. Beispiele hierfür sind Neo4j und ArangoDB.
- **Key-Value-Datenbanken** lassen sich in In-Memory Varianten und On-Disk Varianten unterteilen. Allgemein werden bei Key-Value Datenbanken unter bestimmten Schlüsseln Werte gespeichert. Bei In-Memory ist jeder Server des Clusters mit genügend Arbeitsspeicher ausgestattet und die eintreffenden Daten werden dort hineingeladen. Dadurch lassen sich kurze Zugriffszeiten erzielen. Bei der On-Disk Variante werden die Daten auf der Festplatte gespeichert und eignen sich aus diesem Grund eher als Datenspeicher. Beispiele sind Apache Cassandra und Aerospike. [KTGH13]

Diese Aufzählung von verschiedenen Datenbanksystemen stellt dabei nur einen kurzen und oberflächlichen Einblick von Möglichkeiten bezüglich Big Data dar. Am häufigsten

verbreitet im Rahmen von Big Data und damit essentiell für diese Arbeit ist der **Apache Hadoop Framework**. Hadoop ist ein freies, auf Java basierendes Framework, um effizient und hochskalierbar verteilte Berechnungen von Aufgaben zu übernehmen. Die Entwicklung begann im Jahre 2005. Zunächst war es dafür konzipiert, das Projekt Nutch<sup>1</sup> zu unterstützen, kapselte sich danach aber davon ab und wurde ein eigenständiges Projekt der Apache Foundation. Hadoop wird als MapReduce System bezeichnet, da Hadoop den MapReduce Algorithmus nutzt. MapReduce Systeme lassen sich allgemein wie folgt charakterisieren: [WSV14]

- Es werden skalierbare und als Service nutzbare Computerressourcen verwendet.
- Die Daten werden auf die einzelnen Knoten des Clusters verteilt, ohne, dass vordefinierte Kriterien beachtet werden müssen.
- Der Nutzer muss nur zwei voneinander getrennte Prozesse definieren: Map und Reduce.

Die zuvor genannten Schlüsseleigenschaften von Technologien hinter Big Data setzt Hadoop (durch den MapReduce Algorithmus) folgendermaßen um: [WSV14]

- Die Daten werden in große Blöcke gesplittet und auf dem Hadoop Distributed File System (HDFS) gespeichert. Da das HDFS ein verteiltes Dateisystem ist, werden die Daten auf alle Knoten des Clusters redundant verteilt (standardmäßig liegt bei Hadoop der Replikationsfaktor bei 3).
- Die Anwendungsbibliotheken sowie der Map- und Reduce-Anwendungscode werden auf alle mit Aufgaben versehenen Knoten verteilt.
- Jeder Knoten des Computerclusters liest lokal die Daten und führt den Map-Prozess aus.
- Die Daten werden sequentiell und für je eine Aufgabe auf einen Block zur selben Zeit gelesen (Ein Block hat typischerweise eine Größe von 64MB bis 128MB).

Der MapReduce Algorithmus besteht aus den beiden Schritten Map und Reduce. Diese werden, wie bereits erwähnt, vom Nutzer individuell für eine Anwendung definiert. Abbildung 4 zeigt vereinfacht visualisiert den MapReduce-Prozess. Die Map-Phase stellt dabei einen Aufnahme- und Transformationsschritt dar, bei der einzelne Eingabedatensätze parallel verarbeitet werden können. Die Reduce-Phase ist ein Aggregations- und Verdichtungsschritt, bei der alle zugehörigen Datensätze von einer einzelnen Entität zusammengeführt werden. Ein entscheidender Nachteil des MapReduce Konzepts ist, dass es nicht für iterative Algorithmen geeignet ist. Dies ist insofern von Bedeutung, da ein Großteil datenwissenschaftlicher Algorithmen von Natur aus iterativ sind. Der MapReduce Algorithmus berechnet jede Iteration als einen eigenständigen MapReduce Prozess. Da jedoch jede Iteration die produzierten Daten der vorherigen Iterationen benötigen könnte, verursacht der MapReduce Prozess unnötigen I/O und der allgemeine Durchlauf wird negativ beeinflusst. [WSV14]

---

<sup>1</sup>ein open-source web-search-engine Projekt

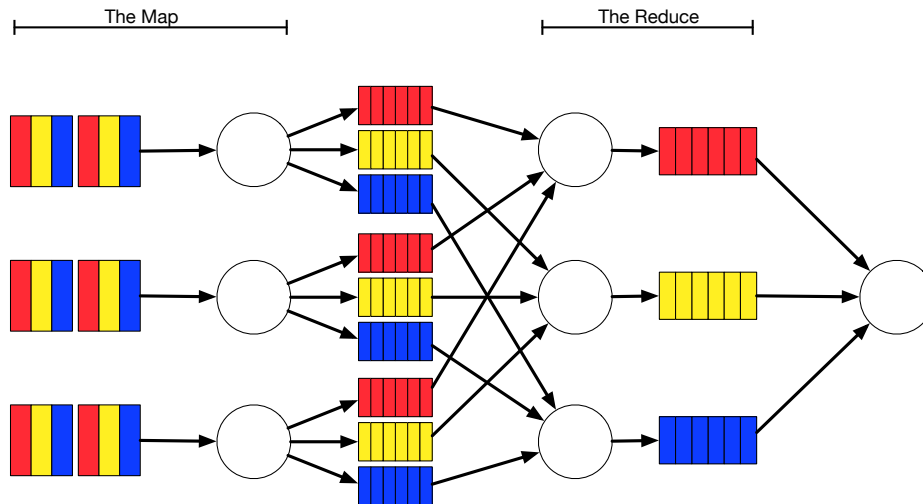


Abbildung 4: MapReduce Illustration in Anlehnung an [WSV14]

## 6 Fazit und Ausblick

Schlussendlich kann festgehalten werden, dass für viele Organisationen Big Data eine essentielle Rolle spielt, bestimmte Anforderungen und Ziele zu erreichen, beziehungsweise diese Ziele auf eine effiziente und vorteilhafte Weise zu erreichen. Das Volumen der Daten in sämtlichen Bereichen, wie Forschung, Gesundheit, Justiz und Wirtschaft, wird zukünftig stets durch neue Datenquellen und steigender Bevölkerungsanzahl weiter zunehmen und die Relevanz und Notwendigkeit von Big Data wird nicht geringer sein als in der heutigen Zeit. Dabei gilt es die Konzepte, Methoden und Technologien den immer wachsenden Anforderungen und Herausforderungen anzupassen. Hier sind als Beispiel neue Speicherarchitekturen im Cloud-Bereich und effizientere Algorithmen zur Auswertung der Daten zu nennen. Zusätzlich zu den technologischen Aspekten sind im Umfeld von Big Data kulturelle und gesellschaftliche, ethische, organisatorische sowie der Rechtslage betreffende Aspekte nicht zu vernachlässigen. Dabei sind besonders Datenschutz- und Sicherheit sowie Privatsphäre hervorzuheben. Big Data hat in der heutigen Zeit durch relativ gut umsetzbarer Technologien, wie der Hadoop Framework, eine durchaus praktische Phase erreicht, wird aber, wie im IT-Bereich häufig, einer stetigen Entwicklung ausgesetzt sein.

## Literatur

- [Bar13] Pavlo Baron. *Big Data für IT-Entscheider - Riesige Datenmengen und moderne Technologien gewinnbringend nutzen*. Carl Hanser Verlag, München, 2013.

- [BBAa11] Philip Bernstein, Elisa Bertino, Divyakant Agrawal und andere. Challenges and Opportunities with Big Data 2011-1. *Cyber Center Technical Reports*, (Paper 1), 2011.
- [Com12] Federal Big Data Commission. Demystifying Big Data: A Practical Guide To Transforming The Business of Government. TechAmerica Foundation, 2012.
- [DeR12] Dirk DeRoos. What is Big Data and how does it fit into an Information Integration Strategy. *Information Integration & Governance Forum*, 2012.
- [GT09] Markus Grünwald und Dirk Traubner. Business Intelligence. *Informatik Spektrum*, 5(32), 2009.
- [Kin14] Stefanie King. *Big Data - Potential und Barrieren der Nutzung im Unternehmenskontext*. Springer VS, München u. a., 2014.
- [KTGH13] Dominik Klein, Phuoc Tran-Gia und Matthias Hartmann. Big Data. *Informatik Spektrum*, 3(36):319–323, 2013.
- [Win11] Richard Winter. Big Data: Business Opportunities, Requirements and Oracle's Approach. Bericht, Winter Corperation, Cambridge, MA USA, 2011.
- [WSV14] Sameer Wadkar, Madhu Siddalingaiah und Jason Venner. *Pro Apache Hadoop*. Apress, Berkeley, CA, 2. Auflage, 2014.

# Seminararbeit „In Memory Computing“

Brian David Clark

Department für Informatik  
Abteilung Informationssysteme  
brian.clark@uni-oldenburg.de  
Matrikelnummer: 2414587  
Betreuer: Cornelius Ludmann

**Abstract:** Decreasing RAM prices allow for corporate databases to be economically loaded entirely into main memory. This not only allows for design patterns based around disk based databases to be discarded, but processing and analyzing methodology to shift towards new and more efficient designs.

## 1 Introduction: Disruptive Technologies

“Disruptive technologies bring to a market a very different value proposition than had been available previously. ... Products based on disruptive technologies are typically cheaper, simpler, smaller, and, frequently, more convenient to use.” [Chr03, p. xviii]

In his book *The Innovator’s Dilemma*, Clayton M. Christensen describes a tendency for entrenched companies to focus further R&D efforts into developing their existing technology, in an attempt to make regular progress to sustain their current product line. These entrenched companies neglect focusing on developing technologies that are (at the time) clearly inferior to the existing technology. However, if the rate of technological progress outstrips growth in demand, eventually the low end of the market falls to the new technology, due to the new technology being “good enough” for an application, while the new technology is simultaneously superior in some other attribute, (i.e. price). With time, the new technology becomes sufficient for the mainstream market, causing a shift in from existing established, entrenched technologies for the alternatives. It is critical for for a business’s future success to understand when (what appears to initially be) an inferior technology has the potential become “good enough” to supplant what was previously accepted as standard.

Since the development of SQL, disk based relational databases have become the accepted way to store and inspect data. However, the technical limitations of being able to contain and process the data simultaneously are subsiding, allowing for new opportunities that were previously infeasible.

## **1.1 Cheaper RAM**

The price of RAM per megabyte have been consistently falling over time. [PZ12, p. 10] As a result, it continuously becomes more economical to implement functionality which requires greater amounts of RAM. While multi-terabyte drives have been within the US Government's capability since 2004 [Mel04], the price of memory is entering the range where it is feasible for corporations.

Plattner and Zeier state: "... a medium-sized enterprise system today contains 100 GB of transactional data and 1 TB of read-only analytical data. Values for a large company are more than 35 TB of transactional data and 40 TB of read-only analytical data." [PZ12, p. 26]

Plattner and Zeier also suggest that as old data is never written and rarely read, it can be maintained in a slower but cheaper hardware rather than in the main memory. [PZ12, p. 104] This further reduces the demand for memory.

Ultimately, the ability to store the database entirely in memory is possible for a wider variety of organisations, as the economic feasibility makes it good enough.

## **1.2 More CPUs (Multi-core, Bus Access)**

Since 2002 increases to CPU frequency over time have waned, but rather the number of CPU cores per chip have increased in compensation. [PZ12, p. 11] A side effect of this is that modern systems need a different mechanism to access memory. While PCs originally used a shared Front Side Bus, Intel has been developing a NUMA (Non-Uniform Memory Architecture) setup called Quick Path Interconnect (QPI) to divide RAM into segments, with each core having direct access to one segment, (and an underlying bus to use remote segments). [Pla13, p. 26] So long as the thread is running on a core which is associated with a RAM segment, threads are allowed greatly increased simultaneous memory access.

## **1.3 System Parallelization and Clustering**

Multiple systems can be combined to work as a cluster. For example Google has been using large clusters of commodity hardware [Tay03] instead of large singular systems. A significant benefit is that as they are less invested in any single machine [Tay03], and by they can operate at a bigger scale through the development of distributed processes (e.g. Map Reduce [Lam10, p. 3]).



## **2 Database Functions**

System R and SQL were designed to implement a simple, consistent interface. At the same time, it is able to isolate the raw data structure from the end users. The use of an optimizer allows the system to choose how to best evaluate the query by estimating disk I/O and CPU usage. Users could simultaneously manipulate data, with guards to maintain consistency. [CAB<sup>+</sup>81]

### **2.1 Recovery (Durability)**

RAM requires power to retain its contents. As such, in-memory databases must rely a form of persistent storage to store data between reboots. Additionally as committed transactions must be durable (i.e. not be lost), a record of the transaction must be written to persistent storage.

### **2.2 Write Optimized vs. Read Optimized**

As row based databases maintain a record in a continuous memory location, it is quick to read or write an individual record.

For OLTP operations, which frequently need to repeatedly insert a single record quickly, row based layouts are optimal, as the new record can be added in a single batch write [PZ12, p. 79].

### **2.3 Column Based Databases**

A column based database groups and continuously stores the records of each attribute together, rather than by individual records. [PZ12, p. 72]

The primary advantage column stores is that aggregate operations can quickly scan all records' attributes.

## **3 The In-Memory Database**

The RAM memory market has advanced to a point where it is economically feasible to store entire databases in memory. This allows for a new type of DBMS, which instead of being built around disk access is instead built around memory access. As such, the traditional disk-based mechanisms for data storage and processing need to be reevaluated to determine if alternative methods provide a better possible solution.

Processes based on working around the Disk I/O limitations can be eliminated, and other processes can be simplified. The entire system around the database can be simplified, extended, or otherwise updated for the new possibilities.

### **3.1 Eliminate disk access**

A memory access is approximately over 100 times faster than that of disk access. [PZ12, p. 9] Read operations can be completed significantly faster. CPU cycles waiting for Disk I/O to finish are wasted. By keeping the database store in memory, a scan across the database can be completed significantly faster than from disk.

Removing disk access has the potential to greatly increase query performance. For example, SAP HANA was able to execute a query over 33 million records in 3 seconds, when it had previously required 45 minutes. [Pla11]

### **3.2 Utilize optimal memory access patterns**

While faster than disk, memory is still slower than the CPU, and thus becomes the new bottleneck for the system. [Pla11] Therefore it is critical to implement memory access in a way that takes advantage of the hardware. As storing data redundantly would result in additional memory use, this would increase memory use. However, the goal is to maintain a small lean memory footprint, allowing the entire database to be read quickly.

### **3.3 Integrate OLTP and OLAP into a single system (Complexity/Cost)**

With disk-based database systems, OLTP and OLAP have been separate instances, in part due to the need to optimize the database for either read or write operations, sacrificing the performance of one for the other. [PZ12, p. 9] With an in memory database operations occur fast enough that a single, unified database can be used. This reduces system complexity, as well as the cost of operating two database systems and the associated overhead.

### **3.4 Utilize multiple cores (Parallelization)**

Large queries can be broken down into individual operations, which can be assigned out to other threads, and the results collected and merged. Relational operations allow the database management system to devise how this breakdown occurs, thus the application programmer does not need to know the internal layout of the database. To quickly on a large set of data, the data corpus can be divided up into partitions, which are assigned between multiple threads running on their own processors, allowing the data to be operated

on in parallel. [PZ12, p. 58-59]

Further parallelization can occur through the use of SIMD (Single Instruction Multiple Data) instructions. This allows multiple operations to occur with per instruction. [PZ12, p. 60]

### **3.4.1 Data Partitioning**

Another strategy is to store data in partitions according to a common attribute. Such partitioning can help to reduce individual table size, allowing for faster queries. [PZ12, p. 103]

## **4 SanssouciDB Implementation**

SanssouciDB is a prototype database for enterprise applications. [PZ12, p. 30] Its goal is to demonstrate the integration of OLTP and OLAP is possible through the performance enhancements of an in-memory database.

### **4.1 Parallelization**

SanssouciDB has implemented several algorithms for use in parallel.

A *scan* operation is used to evaluate a predicate on a column. [PZ12, p. 61]

An *aggregation* operation computes values for a column. [PZ12, p. 62]

A *join* is used to map value combination in tables to rows in the table. [PZ12, p. 62]

It should be noted that the hash tables used in aggregation and join operations are capable of concurrent insertions without locking, allowing for multiple threads to work simultaneously.

### **4.2 Aging data**

While companies may retain data for years, they may only frequently work on recent data, (i.e. the current and previous years). Additionally, the old data is no longer being changed, and is effectively read-only. [PZ12, p. 103-104] An application can define the age of the data, and then SanssouciDB can automatically manage the aging of data during maintenance. [PZ12, p. 105]

An additional advantage is that if the data is not regularly used, it can be stored on slower media to reduce costs. [PZ12, p.104]

### 4.3 Main Store and Differential Buffer

To maintain efficient operation, the majority of the data needs to be stored contiguously. As such, it is difficult to shift data when a write operation would change the internal alignment of data. [PZ12, p. 114]

To effectively implement write operations, SanssouciDB uses a *differential buffer* to quickly take the written data. To take notice of the writes, read operations need to first use the (optimized) main store, followed by checking the differential store to see if any updates occurred. [PZ12, p. 114]

As the differential buffer is not read optimized, it needs to be merged into the main store once a threshold is reached. During this process, a new differential store is created for incoming writes, read operations consult the old main store and the two differential buffers, while a new main store is built from combining the previous main store and differential buffer. Once the table is free from read locks, the old main store and differential store are discarded, leaving the new main store and differential buffer. [PZ12, p. 116-117]

### 4.4 Hybrid Column Store

As only certain attributes are likely candidates for aggregate operations, keeping the other attributes in row based form accelerates accessing an individual record. When properly aligned to queries, this can provide an efficient implementation. [PZ12, p.73-75]

It should be noted that if the queries to be executed are changed, the layout of the hybrid database needs to be reconstructed. [PZ12, p. 72-73]

### 4.5 Lightweight Compression

As memory bandwidth is a limiting factor, methods to minimize the memory footprint enhance performance. [PZ12, p. 64-65] Heavier compression routines could be applied, but for frequently access data would add significant computational overhead. [PZ12, p. 70]

*Dictionary encoding* uses smaller bitstrings to represent longer bitstrings, with the conversion being stored in a lookup table. [PZ12, p. 65]

*Sparse encoding* is useful when the majority of records have the same value. The individual record can then use a single bit to indicate if the attribute is an exception to the common value, and a separate store can be used to contain the exception values. [PZ12, p. 66-67]

*Run-length encoding* is applicable when similar values are stored consecutively for long runs. The value is stored, along with the range for which it runs. [PZ12, p. 66]

*Cluster encoding* divides the data into blocks. Blocks with a single value are compressed as just that value. [PZ12, p. 67]

*Indirect encoding* also operates on blocks of data. While a global dictionary exists, a local

dictionary for each block is used, which references the global dictionary. [PZ12, p. 67-68]

#### 4.6 Insert-Only

*Insert-only* is a methodology where new values are appended to the database, and old values are retained but invalidated. This allows the history of the database to be recorded, perhaps to comply with legal requirements or to allow old values to be retrieved for a detailed analysis at a later time. [PZ12, p. 123]

Insert-only also allows for a simpler MVCC (Multiversion Concurrency Control) implementation, in that read operations can end at the point where the read transaction started, while write transactions continue to append to the list. [PZ12, p. 126]

While insert-only databases have the disadvantage that they consume more memory, they do allow supporting tables (e.g. aggregate data, change history, etc.) to be eliminated. [PZ12, p. 131-132]

### 5 Alternatives to SQL and Relational Databases

Large quantities of data exist as unstructured (or semistructured) data, (i.e. text). By nature this is not suited for SQL. [Lam10, p.7] Many companies need to be able to process such data quickly and efficiently, which has led to the use of several alternatives.

#### 5.1 MapReduce

*MapReduce* consists of *mapping* phase and a *reducing* phase. In the mapping phase, the input data on the node is processed (e.g. filtered, sorted, transformed, etc.) by a defined *map* function into Key/Value pairs. All common keys are grouped together. A *reduce* function is used to further process the collective data associated by a key from *map*. [Lam10, p. 12-13]

MapReduce can handle raw processing tasks (e.g. image processing) that SQL cannot, because SQL is limited to structured data with a clear relational schema. [Lam10, p. 7]

#### 5.2 NoSQL

*NoSQL* is a name used for a collection of databases which don't follow the traditional SQL and relational based approach. Each of these databases has their own approach, and there is no clear or strong definition for them. Some of the common themes of these databases are that they do not use a defined schema and are designed to run on large clusters. [Fow12]

*Key/Value* systems with a simple schema, with a key mapped to a value, (which is often another data structure). [EFHB11, p. 7] It should be noted that MapReduce works Key/Value pairs. [Lam10, p. 7]

*Document* based databases store document. This can be further parsed into specific document formats, (e.g. JSON), at which point they are similar to a Key/Value store.[EFHB11, p. 8]

*Graph* databases store data in graphs defined by triples. [EFHB11, p. 8] This allows various graph theory applications to be applied as analysis, (e.g. distance in a social network). [EFHB11, p. 208] This can even be integrated with RDF and other semantic web technologies. [EFHB11, p. 227]

## 6 Conclusion

The computer market has changed dramatically since System R and SQL where first developed. During that time, SQL and disk based relational databases have proven to be a reliable and useful technology.

However, the demands for data processing and analysis are shifting, and alternate technologies are becoming viable.

While NoSQL databases provide interesting solutions that are invaluable in some cases, there will always be cases where a relational database with a well defined schema will be more appropriate. Here, the biggest change is the economic feasibility to store databases entirely in RAM. It is now reasonable to store small to mid sized databases in main memory, and if the RAM price per megabyte continues to fall over time, large databases will eventually be viable for more organizations as well.

The speed difference that in-memory databases have is considerable, giving these kinds of databases the opportunity to disrupt other aspects of the field. It appears that some of the primary benefits of in-memory databases will include the ability to quickly perform analysis in real time, synchronized over a wide range of sources. As previous methodologies have been designed for years with slower disk based access in mind, it will take time for the new opportunities provided to be fully realized.

It will take time for in-memory databases to fully mature, but as they develop the implications will be worked out. The direct implication of in-memory databases is that queries can be ran over significant data sets without waiting for disk I/O. Indirectly, this enables applications and functionality that were previously considered infeasible.

## Literatur

[CAB<sup>+</sup>81] Donald D. Chamberlin, Morton M. Astrahan, Michael W. Blasgen, James N. Gray, W. Frank King, Bruce G. Lindsay, Raymond Lorie, James W. Mehl, Thomas G. Price, Franco Putzolu, Patricia Griffiths Selinger, Mario Schkolnick, Donald R. Slutz, Irving L.

- Traiger, Bradford W. Wade und Robert A. Yost. A History and Evaluation of System R. *Commun. ACM*, 24(10):632–646, Oktober 1981.
- [Chr03] Clayton M. Christensen. *The Innovator's Dilemma: The Revolutionary Book that Will Change the Way You Do Business (Collins Business Essentials)*. Collins Business, 2003.
- [EFHB11] S. Edlich, A. Friedland, J. Hampe und B. Brauer. *NoSQL: Einstieg in die Welt nichtrelationaler Web 2.0 Datenbanken*. Hanser, 2011.
- [Fow12] Martin Fowler. NosqlDefinition. <http://martinfowler.com/bliki/NosqlDefinition.html>, Januar 2012. Accessed: 2014-11-11.
- [Lam10] Chuck Lam. *Hadoop in Action*. Manning Publications Co., Greenwich, CT, USA, 1st. Auflage, 2010.
- [Mel04] Chris Mellor. US govt buys world's biggest RAM disk. <http://news.techworld.com/storage/3578680/us-govt-buys-worlds-biggest-ram-disk/>, Marz 2004. Accessed: 2014-11-11.
- [Pla11] Hasso Plattner. SanssouciDB: An In-Memory Database for Processing Enterprise Workloads. In Theo Härder, Wolfgang Lehner, Bernhard Mitschang, Harald Schning und Holger Schwarz, Hrsg., *BTW*, Jgg. 180 of *LNI*, Seiten 2–21. GI, 2011.
- [Pla13] H. Plattner. *A Course in In-Memory Data Management: The Inner Mechanics of In-Memory Databases*. Springer Berlin Heidelberg, 2013.
- [PZ12] H. Plattner und A. Zeier. *In-Memory Data Management: Technology and Applications*. SpringerLink : Bücher. Springer Berlin Heidelberg, 2012.
- [Tay03] Allison Taylor. Google's Secret: 'Cheap and Fast' Hardware. <http://www.pcworld.com/article/112891/article.html>, Oktober 2003. Accessed: 2014-11-15.

# **Seminararbeit**

## **„Energiewirtschaft und Energiewende“**

Florian Merkel

Department für Informatik  
Abteilung Informationssysteme  
florian.merkel@uni-oldenburg.de  
Matrikelnummer: 1110837  
Betreuer: Oliver Norkus, Marco Haas

### **1 Einleitung und Motivation**

Ein Leben in der heutigen Gesellschaft einer Industrienation wäre ohne die ständige Verfügbarkeit von Energie in Form von Elektrizität und Wärme kaum vorstellbar. Umso wichtiger ist die Frage, wie die Energieversorgung der Zukunft aussehen kann. Zur Zeit basiert diese hauptsächlich auf fossilen Brennstoffen: Öl, Kohle, Gas und Uran. Diese sind jedoch nur begrenzt vorhanden und ihre Verwendung, ausgenommen Uran, trägt zu einem erhöhten CO<sub>2</sub>-Ausstoß und damit zum Klimawandel bei. Aus diesem Grund wird der Wechsel zu einer Versorgung mit erneuerbaren Energien immer wichtiger. Diese Veränderung bringt jedoch auch Schwierigkeiten mit sich.

In dieser Arbeit wird ein Einblick in den heutigen und vergangenen Energiemarkt in Deutschland bzw. der Europäischen Union gewährt. Darüber hinaus wird die Energiewende in Deutschland und die daraus entstehenden Herausforderungen betrachtet. Der Fokus der Arbeit liegt dabei auf der elektrischen Energie, da sich die Projektgruppe ebenfalls darauf beschränkt. Durch das bessere Verständnis des Strommarkts soll die Erarbeitung von informationstechnischen Lösungen zur effizienteren Vermarktung von Energie erleichtert werden.

Zunächst werden Entwicklung, Struktur und Funktionsweise des Stromnetzes sowie des Strommarkts erörtert. Daraufhin wird der Einfluss der Energiewende auf beide beschrieben.

### **2 Grundlagen**

Um die heutige Struktur sowohl des Energiemarkts als auch des Netzes zu Verteilung von Energie in Form von Strom nachvollziehen zu können muss zunächst deren Geschichte erklärt werden, da diese einen bleibenden Einfluss hinterlassen hat.



## 2.1 Entwicklung

Zu Beginn der Elektrifizierung sind zunächst kleine *Inselnetze* errichtet worden. Ein Inselnetz bezeichnet ein Stromnetz, das nur von wenigen bzw. einem Kraftwerk mit Strom versorgt wird und nicht mit anderen Netzen verbunden ist. Die Inselnetze haben Industrieanlagen, Bürogebäude, Hotels und andere ähnlich große Gebäude mit Elektrizität versorgt, die in erster Linie für die Beleuchtung verwendet worden ist. Schon vor dem Jahr 1900 hat es in Deutschland etwa 500 solcher kleiner Kraftwerke gegeben. Da auch die Nachfrage an elektrischer Beleuchtung in Privathaushalten, öffentlichen Plätzen und Straßen sowie die Versorgung der öffentlichen Verkehrsmittel mit Elektrizität gewachsen ist, haben sich Investoren gefunden, die städtische Kraftwerke betrieben haben. Der erste, der ein solches Kraftwerk hat errichten lassen, ist Thomas Alva Edison mit der *Pearl Street Station* im Jahr 1882 in New York gewesen. Drei Jahre später ist das erste deutsche Kraftwerk zur öffentlichen Versorgung durch die *Aktiengesellschaft Städtische Elektrizitätswerke* in Berlin eröffnet worden. Auch in ländlichen Regionen sind solche Kraftwerke entstanden, die oft von den jeweiligen Bundesländern teilweise oder ganz finanziert und geführt worden sind [Sch06]. Die unterschiedlichen Netze der Betreiber sind mit der Zeit miteinander verbunden worden, sodass ein Verbundnetz entstanden ist [Gmb, Far04]. In ein solches wird an mehreren Stellen Strom ins Netz eingespeist, was zu einem höheren Regelungsaufwand führt, jedoch bei geeigneter Netzinfrastruktur eine höhere Versorgungssicherheit gewährleistet. Das europäische Verbundnetz etwa ist vermascht, mit dem Ziel es *n-1-sicher* zu machen. Das bedeutet, dass das Netz auch dann noch betrieben werden soll, wenn eine beliebige Verbindung ausfällt. Auf Grund des hohen Rechenaufwands wird die n-1-Sicherheit nur für möglichst viele, aber nicht alle möglichen Lastverteilungen simulativ überprüft. Wenn jedoch Teile des Netzes aus Sicherheitsgründen abgetrennt werden müssen, können diese im Inselbetrieb arbeiten [Kon07, Sch06].

Wie diese Stromnetze strukturiert sind und welche Auswirkungen die Strukturen auf den Energiemarkt in Deutschland haben folgt in den nächsten Abschnitten.

## 2.2 Struktur des Stromnetzes

Das Stromnetz ist in mehrere Ebenen unterteilt, die unterschiedliche Spannungen führen und für unterschiedliche Zwecke verwendet werden. Es folgt eine Übersicht über die existierenden Netzebenen mit Erklärungen, welche Netzebene wofür verwendet wird.

**Transportnetze** Diese Netze dienen der Übertragung von Strom über weite Strecken und werden auch *Höchstspannungsnetze* genannt. Sie werden in Deutschland mit 220 oder 380 kV Nennspannung betrieben. Nur "große" Kraftwerke von über 300 MW speisen in die Transportnetze Strom ein. Die Transportnetze sind vermascht, wodurch der Strom je nach Bedarf und Bezugsabsprachen in unterschiedliche und wechselnde Richtungen durch das Netz fließt. Dies geschieht sowohl in den durch den jeweiligen Netzbetreiber verwalteten Regelzonen als auch darüber hinaus. Hauptsächlich werden aber die an das

Transportnetz angeschlossenen Übertragungsnetze versorgt [Sch06, GV].

**Übertragungsnetze** In den Übertragungsnetzen bzw. *Hochspannungsnetzen* wird eine Betriebsspannung von 110 kV verwendet. Daran angeschlossen sind die regionalen Verteilnetze und Sonderkunden wie industrielle Großabnehmer. Außerdem können auch Kraftwerke von bis zu 300 MW Leistung an Übertragungsnetzen angeschlossen sein. Im Gegensatz zu den Transportnetzen sind diese meist strahlenförmig mit ihrem jeweiligen Verteilknoten verbunden. Zum Teil sind die Übertragungsnetze jedoch auch vermascht und werden wie die Transportnetze verwendet [Sch06, GV].

**Verteilnetze** Die Verteilnetze oder *Mittelspannungsnetze* erhalten ihre Energie aus den Übertragungsnetzen und verteilen diese lokal. Durch die Verteilnetze werden Ortsnetze, Schwerpunkstationen in Ballungsgebieten sowie industrielle Großabnehmer beliefert. Bei letzterem spricht man von *Primärverteilung*. Die Netzspannung in Verteilnetzen liegt bei 10 oder 20 kV [Sch06].

**Niederspannungsnetze** Für die Übertragung der Energie zu den Endkunden sind die Niederspannungsnetze verantwortlich. Dazu zählen Ortsnetze, Netze zur Versorgung von Industrieanlagen und Ballungsgebieten. Genau wie die Verteil- und Übertragungsnetze sind auch die Niederspannungsnetze sternförmig [Sch06, GV].

Die angegebenen Spannungen sind die *Nennspannungen* der jeweiligen Netze, d.h. die Betriebsspannungen im Normalbetrieb. Außer in den Niederspannungsnetzen ist eine Abweichung von  $\pm 15\%$  im Dauerbetrieb zulässig [Kon07].

Neben der Nennspannung in den jeweiligen Netzen ist noch eine weitere Größe wichtig: die *Netzfrequenz*. Sie beträgt im europäischen Verbundnetz 50 Hz. Die Netzfrequenz ergibt sich direkt aus der Technik zur Stromerzeugung, bei der Dreiphasengeneratoren zum Einsatz kommen. Bei diesen sind drei Leiterspulen um einen Magneten angeordnet. Durch kinetische Energie, z.B. aus einer Dampfturbine, wird dieser Magnet zum Rotieren gebracht. Da sich dadurch das Magnetfeld in den Spulen ständig ändert, wird elektrischer Strom induziert. Die Netzfrequenz ist somit direkt von der Drehfrequenz des Magneten im Kraftwerk abhängig. Diese Art des Netzbetriebs ermöglicht es, an jedem beliebigen Punkt im Netz bestimmen zu können, ob die benötigte Leistung verfügbar ist. Das ist deshalb möglich, weil eine Entnahme von Energie aus dem Stromnetz zu einem Abfall der Netzfrequenz führt und gleichermaßen das Einspeisen einen Anstieg der Netzfrequenz bewirkt. Des Weiteren lässt sich der so erzeugte *Dreiphasenwechselstrom* leichter höhere Spannungen transformieren als es mit Gleichstrom der Fall wäre, was wiederum eine verlustärmere Übertragung über weite Strecken ermöglicht. Da die mechanischen Komponenten im Kraftwerk nur in einem gewissen Rahmen funktionieren, ohne Schaden zu nehmen, müssen Abweichungen ausgeglichen werden. Wie das passiert wird in Unterabschnitt 2.5 erörtert. Da im Stromnetz keine Energie gespeichert werden kann, muss immer so viel Strom erzeugt werden, wie gerade benötigt wird [Kon07].

Seit Beginn der Energiewende verändert sich die Struktur der Stromnetze zunehmend.

Dies betrifft insbesondere die Nieder- und Mittelspannungsnetze. Die Art dieser Veränderung, ihre Hintergründe und ihre Auswirkungen werden in Abschnitt 3 erörtert.

### 2.3 Entwicklung des Energiemarkts

Aufgrund der erheblichen Kosten die der Ausbau und Betrieb eines Stromnetzes mit sich bringt und des eingeschränkten Platzes der für die Infrastruktur zur Verfügung steht haben die Energieversorgungsunternehmen (EVU) räumliche Monopolstellungen innegehabt. Diese sind 1920 durch Demarkationsverträge abgesichert worden, in denen die EVU festlegt haben wer in welchem geographischen Bereich Kunden mit Energie hat beliefern dürfen (sog. *Regelzonen*), wodurch jeglicher noch vorhandener Wettbewerb unter den EVU zum Erliegen gekommen ist. Die einzige Konkurrenz hat die Eigenstromerzeugung in Industrieanlagen dargestellt. Das 1935 erlassene *Energiewirtschaftsgesetz* und die zugehörigen Verordnungen haben die Monopolstellung der EVU gefestigt, indem sie zum einen die gängige Praxis der Vergabe von Konzessionsverträgen durch die Kommunen an die lokal ansässigen EVU festgeschrieben haben und zum anderen die Errichtung von Anlagen zur Eigenstromerzeugung eingeschränkt haben. Im Gegenzug für die Garantie der alleinigen Versorgung durch ein EVU hat man sich niedrige Strompreise und eine sichere Stromversorgung erhofft [Far04].

Die nächste wesentliche Veränderung der beschriebenen Verhältnisse ist erst 1998 mit der Umsetzung der *EU Richtlinie 96/92/EG* durch das *Energiewirtschaftsgesetz - EnWG* eingetreten. Es hat erste Liberalisierungsbemühungen angestoßen, welche jedoch teilweise nur langsam vorangegangen sind. Deshalb ist 2005 die *EU-Richtlinie 2003/54/EG* mit dem *Zweiten Gesetz zur Neuregelung des Energiewirtschaftsrechts* in nationales Recht umgesetzt worden, welches das EnWG ersetzt hat und dessen Eckdaten 2011 und 2012 angepasst worden sind. Eine wichtige Neuerung besteht in der *Entflechtung (Unbundling)* von EVU mit mehr als 100.000 Kunden.[Sch06] Das bedeutet die Trennung der Bereiche *Erzeugung*, *Netzbetrieb* und *Vertrieb* sowohl gesellschaftsrechtlich als auch buchhalterisch, operationell und informationell. Für EVU mit weniger Kunden gelten jedoch einige Ausnahmen. Darüber hinaus ist gesetzlich festgelegt worden, dass jeder Stromkunde seinen Versorger frei wählen kann. Auf diese Weise haben Endkunden die Möglichkeit erhalten zwischen mehreren Stromversorgern und diesen wiederum zwischen mehreren Produzenten zu wählen, wodurch ein marktregulierender Wettbewerb gewährleistet werden soll. Ein weiterer Aspekt ist die Regulierung der Netznutzungsentgelte durch ein Regulierendes Organ, die Bundesnetzagentur. Diese muss dafür sorgen, dass Übertragungs- und Verteilnetzbetreiber den Zugang zu ihren Netzen diskriminierungsfrei für alle Netznutzer gestaltet. Insbesondere müssen die gleichen Netznutzungsentgelte für Kunden und Nichtkunden der Netzbetreiber gelten. Für Großkunden, die mindestens 10 GWh im Jahr verbrauchen und auf mindestens 7.500 Benutzungsstunden im Jahr kommen dürfen die Netzbetreiber die Entgelte jedoch bis zu 50% niedriger festsetzen. Eine Regulierung dieser Art ist notwendig, da die Netzbetreiber nach wie vor ein natürliches Monopol besitzen [Kon07].

## 2.4 Akteure am Energiemarkt

Vor dem Unbundling haben die EVU Produktion, Transport und Vertrieb des von ihnen gehandelten Stroms übernommen. Üblicherweise ist das in ihren jeweiligen Regelzonen geschehen. Falls es in anderen Teilen des Verbundnetzes zu Ausfällen gekommen ist, ist jedoch auch Strom über die Verbundnetzleitungen in andere Regelzonen gespeist worden. Am Stromfluss selbst hat das Unbundling nichts geändert, da die Netztopologien gleich geblieben sind. Die Entflechtung hat jedoch bewirkt, dass die zuvor integrierten Teile der EVU jetzt als eigenständige Akteure am Energiemarkt agieren und weitere Unternehmen, die nicht ursprünglich Teil eines EVU waren, ebenfalls die Möglichkeit haben am Energiehandel teilzunehmen [Kon07].

**Kunden** Die Kunden sind Abnehmer elektrischer Energie. Sie werden in *leistungsbe-messene Kunden* und *Grundversorgungskunden* unterteilt. Erstere sind in der Regel Betreiber von Industrieanlagen, die oftmals auch Eigenstromerzeugung betreiben. Zweitere sind üblicherweise private Haushalte bzw. Kunden mit einem Jahresverbrauch von weniger als 100 kWh. Sie unterscheiden sich in der Art wie der Strom abgerechnet wird [Kon07].

**(Energie-)Versorgungsunternehmen** Ein Unternehmen, das ein Energieversorgungsnetz betreibt oder Energie an andere liefert wird nach §3 Abs. 18 ein Energieversorgungsunternehmen genannt [Kon07].

**Grundversorger** Als Grundversorger wird das Versorgungsunternehmen bezeichnet, welches die meisten Haushalts- bzw. Grundversorgungskunden in einem Netzgebiet beliefert. Für Grundversorger besteht die Verpflichtung, alle Grundversorgungskunden in ihrer Regelzone ans Netz anzuschließen [Kon07].

**Netzbetreiber** Die Netzbetreiber werden in zwei Gruppen eingeteilt. Die *Übertragungs-netzbetreiber* verwalten die 220 kV und 380 kV Netze und bilden, über Kuppelleitungen verbunden, das deutsche Verbundnetz. Ihre Aufgabe besteht in der Spannungs- und Frequenzhaltung und im Ausgleich von Fahrplanabweichungen in ihrem Netzgebiet. Was das konkret bedeutet wird in Unterabschnitt 2.5 erklärt. Die *Verteilnetzbetreiber* sind für die Netze von bis zu 110 kV zuständig und verteilen den Strom aus den Transport- und Übertragungsnetzen an die Verteilnetze. Bei einem Vergleich mit Unterabschnitt 2.2 fällt auf, dass die Benennung der Netzbetreiber nicht ganz konsistent zur Benennung der Netzebenen ist. Die hier verwendete Benennung spiegelt allerdings die in der Literatur vorkommenden Bezeichnungen wider [Kon07].

**Erzeuger** Die Stromproduktion wird von den Erzeugern übernommen. Dazu zählen Kleinproduzenten, die in der Regel unabhängig agieren, aber auch Verbundunternehmen mit Kraftwerkparcs [Kon07].

**Lieferanten** Zu den Lieferanten gehören sowohl Kraftwerksbetreiber als auch Stromhändler. Während die ersten nur den selbst erzeugten Strom am Markt anbieten sind die Händler sowohl am Kauf als auch Verkauf von Strom beteiligt. Das *EnWG* verbietet es allerdings sowohl Händler als auch Produzent zu sein. Lieferanten müssen die Fahrpläne für ihre Entnahme- und Einspeisestellen viertelstündlich an die Übertragungsnetzbetreiber übermitteln [Kon07].

**Energiebörse** Wie mit Aktien und anderen Gütern kann auch mit Strom gehandelt werden. Aufgabe der Energiebörse ist es eine Plattform bereitzustellen, die finanziell, rechtlich und technisch sicher ist. An der Energiebörse wird neben Strom auch mit Erdgas und anderen Primärenergieträgern sowie Emissionszertifikaten gehandelt. Die *European Energy Exchange (EX)* in Leipzig ist die in Deutschland und Kontinentaleuropa führende Energiehandelsplattform.[AG]

## 2.5 Regelung

Wie bereits in Unterabschnitt 2.2 erwähnt, muss immer gerade so viel Energie in das Stromnetz eingespeist werden wie benötigt wird, da diese im Netz nicht gespeichert werden kann. Aus diesem Grund müssen die Übertragungsnetzbetreiber, die für die Frequenzhaltung verantwortlich sind, den erwarteten Stromverbrauch modellieren. Dies geschieht mit sog. *Fahrplänen*, die Angeben zu welcher Tageszeit wie viel Strom benötigt wird. Zur Erstellung der Fahrpläne wird eine Regelzone in sog. *Bilanzkreise* unterteilt. Einem Bilanzkreis gehört mindestens ein Netznutzer der jeweiligen Regelzone an und jeder Netznutzer gehört zu genau einem Bilanzkreis. Die Netznutzer in einem Bilanzkreis benennen einen Bilanzkreisverantwortlichen, der einen Tag im Voraus Fahrpläne mit viertelstündlicher Auflösung für die Einspeise- und Entnahmestellen in seinem Bilanzkreis erstellt und an den Übertragungsnetzbetreiber der Regelzone, in der sich der Bilanzkreis befindet, übermittelt. Der entsprechende Übertragungsnetzbetreiber wird Bilanzkoordinator genannt. Mit Hilfe dieser Fahrpläne können die Übertragungsnetzbetreiber planen wie viel Leistung zu welcher Zeit im Netz verfügbar sein muss. Der tatsächliche Verbrauch weicht jedoch in der Regel vom erwarteten Verbrauch ab und das sekundlich und nicht viertelstündlich [Kon07].

Zum Ausgleichen dieser Abweichungen werden die Kraftwerke im gesamten Verbundnetz im Normalbetrieb etwa 3-5% unter der Leistungsgrenze betrieben. Bei Bedarf kann dann die Leistung der Kraftwerke noch erhöht werden (*Primärregelung*). Dies geschieht automatisch innerhalb von 15 bis 30 Sekunden und kann mindestens 15 Minuten aufrecht erhalten werden. Die vorgehaltene Leistung der Primärregelung, die *Primärregelreserve*, soll möglichst schnell wieder einsatzfähig sein, nachdem sie hat aktiviert werden müssen. Dazu wird spätestens 30 Sekunden nach einer Inanspruchnahme automatisch die *Sekundärregelung* aktiv. Sie wird von Kraftwerken in einer Regelzone, die im Teillastbetrieb arbeiten und bei Bedarf hochgefahren werden können, übernommen. Die Kraftwerke müssen nach 15 Minuten die benötigte Leistung liefern können und dazu in der Lage sein sie mindestens 1 Stunde lang zu erbringen. Auch diese Reserve wird von der nächsten

Stufe, der Tertiärregelung, abgelöst. Das geschieht nach spätestens 15 Minuten. Die Tertiärregelung wird im Gegensatz zu den vorherigen beiden Stufen manuell aktiviert. Die Primär-, Sekundär- und Tertiärregelung werden vom Übertragungsnetzbetreiber durchgeführt. Die nächste Regelstufe jedoch, die *Stundenreserve*, liegt in der Verantwortung des Bilanzkreisverantwortlichen und wird in Anspruch genommen, falls im Bilanzkreis nicht genug Leistung zur Verfügung steht. Falls jedoch nach allen Regelversuchen nicht genügend Leistung im Netz verfügbar ist, tritt ein Notfallplan in Kraft. Dieser sieht die schrittweise Trennung von Netzteilen vor. Sollte dies keine ausreichende Wirkung zeigen, werden Kraftwerke vom Netz genommen, um Schaden an der Technik zu vermeiden und der Strom fällt aus. Der beschriebene Ablauf tritt bei einer Unterversorgung des Netzes ein. Bei einer Überversorgung wird die Leistung von Kraftwerken entsprechend heruntergefahren. Dies geschieht allerdings nur unterhalb der Höchstspannungsebene [Kon07].

## 2.6 Stromhandel

Der Handel mit elektrischer Energie findet einerseits an der Energiebörse statt, andererseits wickeln Akteure am Energiemarkt auch Geschäfte außerhalb der Börse ab. An der Energiebörse EEX sind zwei Arten von Börsengeschäften möglich. Zum einen ist dies der Handel über den Spotmarkt, zum anderen über den Terminmarkt

Am Spotmarkt werden standardisierte Produkte für die physische Lieferung von Strom angeboten. Es können Verträge über die Bereitstellung einer bestimmten konstanten Leistung zu bestimmten Stunden (*Stundenkontrakte* bzw. Stundenblöcken (*Blockkontrakte*)) geschlossen werden. Je nach Art des Vertrags wird die Lieferung am selben oder am Folgetag ausgeführt. Dies geschieht durch Fahrplananmeldung bei den Übertragungsnetzbetreibern. Über den Spotmarkt werden Regelleistungen (bei auten Abweichungen vom Fahrplan) für den aktuellen Bedarf und Leistungen für den nächsten Tag einkaufen.

Der Terminmarkt ist für längerfristige Geschäfte gedacht. Er ermöglicht das Handeln von *Futureskontrakten* und *Optionskontrakten*, die zur Absicherung von Preisrisiken dienen [Kon07].

## 3 Energiewende

Neben der Liberalisierung des Energiemarkts beeinflusst die *Energiewende* wie Strom erzeugt und transportiert wird und wie der Handel mit Strom in Zukunft aussehen wird. Mit dem Begriff *Energiewende* ist die Verwirklichung einer nachhaltigen Versorgung mit Strom, Wärme und Mobilität gemeint [Mau13]. Hinsichtlich der schwindenden Vorräte an rentabel erschließbaren fossilen Brennstoffen und des durch den steigenden Ausstoß an Treibhausgasen wie Methan und CO<sub>2</sub> verursachten Klimawandel beinhaltet dies auch ein Verringern der Abhängigkeit von den fossilen Energieträgern Gas, Kohle und Öl. In Deutschland kommt noch dazu, dass auch die klimaneutrale Erzeugung von Strom aus Kernenergie auf Grund der mit ihr Verbundenen Risiken beendet werden soll [HM12].

Neben den ökologischen Aspekten der Energiewende ist auch der energiepolitische Betrachtungswinkel nicht außer Acht zu lassen. Eine Energieversorgung, die größtenteils oder vollständig durch erneuerbare Energien abgedeckt ist reduziert die Abhängigkeit von Öl- und Gasimporten aus politisch weniger stabilen Ländern. Historische Ereignisse wie die erste Ölkrise in den 1970er Jahren verdeutlichen, dass eine energiewirtschaftliche Abhängigkeit schwerwiegende Folgen für die Wirtschaft eines Landes haben kann [Kon07].

Die gesetzlichen Vorgaben zu Bewältigung der Energiewende in Deutschland legt das *Erneuerbare Energien Gesetz (EEG)* fest. Näheres dazu ist in Unterabschnitt 3.1 zu finden.

### **3.1 Gesetzgebung**

Die Anfänge der Legislatur zur Energiewende sind im 1990 verabschiedeten *Strompreisgesetz* zu sehen. Dieses wurde 2000 durch das *Gesetz für den Vorrang Erneuerbarer Energien (EEG)* ersetzt, welches bis heute mit einer Vielzahl an Veränderungen, die jüngste im Jahr 2014, besteht. Ziel des EEG ist unter anderem die Erhöhung des Anteils regenerativer Energieerzeugung. Als erneuerbare Energien zur Stromerzeugung sind im EEG Solar- und Windkraft sowie die Biogasverstromung und die Geothermie berücksichtigt. Bis zum Jahr 2020 soll dieser Anteil mindestens 35%, bis 2030 mindestens 50% und bis 2040 mindestens 65% betragen [BfWuE]. Das EEG in seiner heutigen Form gibt unter anderem vor, dass Strom aus erneuerbaren Energien (EE) von den Verteilnetzbetreibern abgenommen, übertragen und verteilt werden muss. Darüber hinaus sind die Verteilnetzbetreiber verpflichtet, Anlagen zur Erzeugung von EE-Strom an das Netz anzuschließen und 20 Jahre lang, zumindest bis zur Gesetzesnovelle 2014, für den eingespeisten Strom eine Mindestvergütung zu bezahlen. Der von den Verteilnetzbetreibern abgenommene Strom muss wiederum durch die Übertragungsnetzbetreiber abgenommen und vergütet werden. Umgekehrt gilt dies allerdings nicht: Die Übertragungsnetzbetreiber müssen den Strom aus EE an der Börse vertreiben [Kon07]. Die Höhe der Vergütungen für unterschiedliche Formen der EE Kraftwerke wurden seit der erstmaligen Einführung immer wieder - und in der Regel nach unten - angepasst. Für größere Neuanlagen wird es ab 1. Januar 2015 allerdings keine Vergütung mehr geben. Stattdessen muss der Strom an der Energiebörse vermarktet werden. Die Differenz zwischen den bisher gültigen Vergütungen und dem durchschnittlichen Marktpreis wird als sogenannte *Marktprämie* erstattet. Bei kleineren Kraftwerken und Bestandsanlagen besteht die Wahl zwischen Vergütung und Marktprämie [BfWuE]. Die Mindestvergütung bzw. die Marktprämie muss nicht von den Versorgungsunternehmen gezahlt werden. Stattdessen wird sie als *EEG-Umlage* auf die Endkunden abgewälzt, sofern diese nicht davon befreit sind [Kon07].

### **3.2 Technische und wirtschaftliche Implikationen**

Obwohl bei der Energiegewinnung aus erneuerbaren Energien keine Brennstoff- und kaum Arbeitskosten anfallen, sind die Kapital- bzw. Instandhaltungskosten pro kW Leistung we-

sentlich höher. Dies hängt mit dem geringeren Nutzungsgrad zusammen. So tragen beispielsweise Windkraftanlagen trotz höherer vorhandener Gesamtleistung weniger als halb so viel zur erzeugten Strommenge bei wie die vorhandenen Kohlekraftwerke, da der Wind sehr unregelmäßig weht. Aus diesem Grund muss nicht nur der Ausbau sondern auch der Betrieb von Windkraftwerken und Solaranlagen subventioniert werden, was über die bereits erwähnte EEG-Umlage bzw. die Marktprämie geschieht. Darüber hinaus entsteht durch die Unregelmäßigkeit und Unvorhersehbarkeit der Einspeisung aus den genannten Quellen ein erhöhter Regelungsaufwand. Da sich nicht zuverlässig vorhersehen lässt, wie stark der Wind weht und wie die Abschattung von Photovoltaikanlagen ausfällt, ist die Fahrplanerstellung unter Berücksichtigung dieser Energiequellen kaum möglich. Ein Netzbetreiber muss also zur Deckung seines Energiebedarfs in Kauf nehmen, dass er möglicherweise Strom bezahlt, den er nicht braucht und zusätzlich noch den Wind- und Solarstrom ausregeln muss, was zu höherem Verschleiß und einem möglicherweise unrentablen Betrieb von Kraftwerken führt. Erwartet ein Netzbetreiber stattdessen eine größere Deckung seines Energiebedarfs aus erneuerbaren Quellen, die dann allerdings ausbleibt, muss er zusätzlichen Strom zu einem höheren Preis einkaufen. Ist die Einspeisung sogar so groß, dass er sie nicht in seinem Netz aufnehmen kann muss er sie trotzdem vergüten und zusätzlich dafür sorgen, dass die Kapazitäten in Zukunft aufgenommen werden können. Dies führt letztendlich zu höheren Preisen für die Endverbraucher [Kon07].

Da es nicht Möglich ist, die Einspeisung durch Wind- und Solarkraftwerke zu kontrollieren, werden stattdessen Ansätze verfolgt, die Last im Stromnetz zu steuern. Mit *Smart Metern* sollen verbrauchsintensive Geräte wie Geschirrspüler, Waschmaschinen und Wäschetrockner genau dann eingeschaltete werden, wenn gerade viel Strom erzeugt wird. Auch entsprechend gut isolierte Kühl- und Gefrierschränke könnten dann ihre Arbeit aufnehmen. Eine weitere Anwendung tut sich auf Grund der zunehmenden Elektromobilität auf, denn auch das Laden von E-Bikes und anderen Elektrofahrzeugen kann von der *Smart Meter* Technologie profitieren. Für den Netzbetreiber hat die Verwendung von intelligenten Stromzählern (*Smart Metern*) den Vorteil, dass er so Lastspitzen vermeiden und eine gleichmäßigere Lastkurve erreichen kann, mit der die Planung effizienter durchgeführt werden kann.

Ein weiterer Ansatz zur Lösung der unvorhersehbaren Einspeisung mancher erneuerbarer Energien besteht in *virtuellen Kraftwerken*. Diese bezeichnen den Zusammenschluss unterschiedlicher und dezentraler Kleinkraftwerke, die von einer Stelle aus kontrolliert werden. Durch die Diversifizierung soll die Wirtschaftlichkeit der Einzelanlagen erhöht und Nachteile der einzelnen Erzeugungsarten ausgeglichen werden [Sch06]. Beispielsweise könnte durch den Zusammenschluss eines Windparks, eines Photovoltaikkraftwerks und eines Pumpspeicherkraftwerks bei überschüssiger Einspeisung, wenn die Sonne scheint und der Wind weht, das Pumpspeicherkraftwerk zur Speicherung verwendet werden. Bei Ausfall je einer der erneuerbaren Quellen könnte die andere den Verlust ausgleichen und beim Ausfall von beiden könnte das Pumpspeicherkraftwerk gespeicherten Strom einspeisen [DFBK<sup>+</sup>09].

Bei einer Kombination der beiden Ansätze spricht man von einem *Smart Grid* [Sch06].

Eine zunehmende Ausstattung mit intelligenten Stromzählern erfordert jedoch auch einen höheren Kommunikationsaufwand. Mehrere technische Möglichkeiten zur Kommunika-



tion stehen zur Verfügung: Die Signalübertragung auf Niederspannungsleitungen bis zur nächsten Trafostation, die Vernetzung über das Internet und die Verwendung von Mobilfunknetzen. Pilotprojekte zur Erforschung der Technik existieren bereits [Kon07].

## 4 Zusammenfassung

Die Arbeit hat einen Überblick über den deutschen bzw. europäischen Energiemarkt gewährt. Es wurden auf die Energiewende in Deutschland und deren Umsetzung durch das Erneuerbare-Energien-Gesetz eingegangen. Dabei wurden die Probleme, die aus der Kombination der Gesetzgebung und der Unvorhersehbarkeit von Wind- und Solarkraft eingegangen. Schließlich wurden Lösungsansätze für die identifizierten Problem vorgestellt, die sich zur Zeit in Entwicklung befinden. Mit den behandelten Grundlagen und in Kombination mit den anderen im Rahmen der Projektgruppe *Business Intelligence as a Service* angefertigten Seminararbeiten wird die Erstellung von informationstechnischen Lösungen für eine effizientere Vermarktung von Strom im europäischen Markt erleichtert.

## Literatur

- [AG] European Energy Exchange AG. At the Centre of European Energy Trading. <http://www.eex.com/blob/68250/bc1150c78328d90dc18a14f9dd6cbacf/e-eex-unternehmen-februar-2014-pdf-data.pdf>, Abgerufen am 13.11.14, 12:47 Uhr.
- [BfWuE] Referat Öffentlichkeitsarbeit Bundesministerium für Wirtschaft und Energie. BMWi - Erneuerbare Energien - Das Erneuerbare-Energien-Gesetz. <http://bmwi.de/DE/Themen/Energie/Erneuerbare-Energien/erneuerbare-energien-auf-einen-blick.html>, Abgerufen am 17.11.14, 19:04.
- [DFBK<sup>+</sup>09] Bert Droste-Franke, Holger Berg, Annette Kötter, Jörg Krüger, Karsten Mause, Johann-Christian Pielow und I Romey T Zieseemer. Brennstoffzellen und virtuelle Kraftwerke. *Ethics of science and technology assessment*, 36:43–131, 2009.
- [Far04] Alexander Faridi. Der regulierende Eingriff des Energiewirtschaftsgesetzes in den Wettbewerb zwischen öffentlicher und industrieller Stromerzeugung in den 30er Jahren. *Zeitschrift für Unternehmensgeschichte/Journal of Business History*, 49(2):173–197, 2004.
- [Gmb] 50Hertz Transmission GmbH. Historischer Hintergrund - Stromnetz - 50Hertz Transmission GmbH. <http://www.50hertz.com/de/50Hertz/Historischer-Hintergrund>, Abgerufen am 12.11.14, 15:51 Uhr.
- [GV] José M. G. González Vázquez. *Ein Referenzmodellkatalog für die Energiewirtschaft - Management von Informationsmodellen für Softwareproduktmanager bei der Anforderungsanalyse*. Dissertation.
- [HM12] Harald Heinrichs und Gerd Michelsen. *Nachhaltigkeitswissenschaften*. Springer-Verlag, 2012.
- [Kon07] Panos Konstantin. *Praxisbuch Energiewirtschaft*. Springer, 2007.

- [Mau13] Klaus-Dieter Maubach. Fundamente der Energiewende (1980–1998). In *Energiewende*, Seiten 41–58. Springer, 2013.
- [Sch06] Adolf Schwab. *ELEKTROENERGIESYSTEME: ERZEUGUNG, TRANSPORT, AOBEBE*. Springer, 2006.

**Seminararbeit**  
**„Systeme in der Energiewirtschaft“**

Björn Friedrich

Department für Informatik  
Abteilung Informationssysteme  
bjoern.friedrich@uni-oldenburg.de  
Matrikelnummer: 2023167  
Betreuer: Oliver Norkus und Marco Haas

**Abstract:**

Die vorliegende Arbeit soll einen Überblick über die Systeme in der Energiewirtschaft geben, die für ein Business Intelligence Reporting System von Interesse sind. Zuerst wird die Motivation des Vorhabens beschrieben und dann auf die einzelnen Systeme eingegangen. Es werden Customer Relationship Management, Energiedaten Management, ene't und SAP-Industry Solution Utilities vorgestellt. Für jedes System wird auf die Eingangsdaten, Ausgangsdaten und die Funktion des Systems eingegangen. Zum Abschluss gibt es eine kurze Zusammenfassung und die Bedeutung dieser Systeme für die Projektgruppe.

## **Inhaltsverzeichnis**

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Projektgruppe . . . . .	1
1.2	Business Intelligence und Cloud Computing . . . . .	1
1.3	Energiewirtschaft . . . . .	2
<b>2</b>	<b>Systeme</b>	<b>3</b>
2.1	Customer Relationship Management . . . . .	3
2.2	Energiedaten Management . . . . .	4
2.3	ene't . . . . .	4
2.4	Industry Solution Utilities . . . . .	6
<b>3</b>	<b>Schlussbemerkung</b>	<b>9</b>

# 1 Einleitung

## 1.1 Projektgruppe

Die Projektgruppe „Business Intelligence as a Service“ ist bestrebt, die beiden Gebiete Cloud Computing und Business Intelligence zu verbinden. Das übergordnete Ziel ist ein Business Intelligence System in der Cloud (Abb. 1). Der Prototyp soll mit einer Fallstudie aus der Energiewirtschaft getestet werden. Einem Kundenberater soll es möglich sein, einem Kunden vor Ort ein Angebot zu machen. Dabei soll der Preis pro Kilowattstunde vom System berechnet werden. Der Preis soll unter Einbezug aktueller Marktpreise, Leistungskosten und Netznutzungskosten berechnet werden.

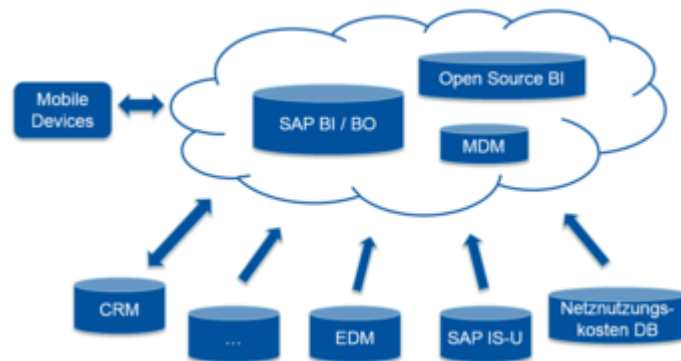


Abbildung 1: Eine bildliche Skizze des Systems, das die Projektgruppe entwickeln soll. In der Cloud sind die Systeme abgebildet, die für ein Business Intelligence System notwendig sind. Es stehen mehrere zur Auswahl. Unter der Cloud sind die Quellsysteme zu finden. Links sieht man die Anbindung der mobilen Geräte. [Nor]

## 1.2 Business Intelligence und Cloud Computing

Business Intelligence beschreibt die systematische Analyse von Daten. Die Analyse umfasst Daten des eigenen Unternehmens, Daten über die Mitbewerber und die Marktentwicklung. Ziel ist es, aus aktuellen Daten Wissen zu erhalten auf dessen Grundlage geschäftliche Entscheidungen getroffen werden können. Schnell auf Veränderungen am Markt reagieren und Trends frühzeitig erkennen zu können, stellt einen Vorteil im Wettbewerb dar. Informationen über diese Ereignisse müssen aktuell und zu jeder Zeit verfügbar sein. Insbesondere für mobile Endgeräte ist dies eine Herausforderung. Die Analyse der Daten erfordert ein hohes Maß an Rechenleistung und die Masse der Daten eine enorme Speicherkapazität. Die Leistung, die für die Analyse und Aufbereitung der Daten insgesamt vonnöten ist, kann nur durch Server bereitgestellt werden.

Mobile Endgeräte übernehmen die Visualisierung der Daten. Die Analyse und Datenspeicherung übernimmt ein System in der Cloud.

Das National Institute of Standards and Technology definiert Cloud Computing wie folgt: „Cloud Computing ist ein Modell, das es erlaubt bei Bedarf jederzeit, und überall bequem über ein Netz auf einen geteilten Pool von konfigurierbaren Rechnerressourcen (z. B. Netze, Server, Speichersysteme, Anwendungen und Dienste) zuzugreifen, die schnell und mit minimalem Managementaufwand oder geringer Serviceprovider-Interaktion zur Verfügung gestellt werden können.“ [MG]. Diese Definition wird ebenfalls von der European Network and Information Security Agency und dem Bundesamt für Sicherheit in der Informationstechnik verwendet und kann deshalb als gängige Definition angesehen werden. [fSidI]

Mit einem entsprechenden System ist es möglich Business Intelligence als eine Dienstleistung bzw. als Cloud-Service für Unternehmen anzubieten. Die Projektgruppe entwickelt den Prototypen so eines Systems. Dieser wird mit einem Anwendungsfall aus der Energiewirtschaft erprobt.

### **1.3 Energiewirtschaft**

Die Energiewirtschaft befindet sich im Umbruch. Mit den Erneuerbaren Energien wandelt sich die Struktur der Energieproduktion von einer zentralisierten in eine dezentralisierte Struktur. Dieser Wandel stellt die Energiewirtschaft vor große Herausforderungen. Durch die Erneuerbaren Energien, die daraus resultierende Dezentralisierung und das Smart-Metering haben sich die Anforderungen an Business Intelligence grundlegend verändert. Hinzu kommt der European Energy Exchange (im Folgenden EEX). Am EEX wird mit Strom gehandelt. Neben Termingeschäften sind auch Over-The-Counter Geschäfte möglich. Bei Termingeschäften wird ausgehandelt wie viel Strom, zu welchem Preis an einem Zeitpunkt in der Zukunft geliefert werden soll. Solche Geschäfte lassen sich von den Versorgern gut kalkulieren, weil weit im Vorfeld bekannt ist wie viel Strom gebraucht wird. Damit wird die Versorgung langfristig abgesichert. Bei Over-The-Counter Geschäften werden Geschäfte kurzfristig getätigt. Es kann Strom gekauft werden, der zeitnah produziert wird. Damit wird vom Versorger unvorhersehbarer Strombedarf abgedeckt. Der Preis wird von Angebot und Nachfrage bestimmt und variiert dementsprechend. [TK09]

Die Stromgewinnung aus den Erneuerbaren Energien ist zu einem großen Teil wetterabhängig. Niemand kann genau vorhersagen welche Wetterbedingungen zu einem bestimmten zukünftigen Zeitpunkt herrschen. Gerade bei Wetterdaten ist deshalb eine schnelle Analyse erforderlich. Weiterhin werden durch das Smart-Metering deutlich mehr Daten erfasst. Das Smart-Metering ermöglicht es in deutlich kürzeren Abständen Messwerte für einzelne Kunden zu erfassen. Während früher immer nur ein Messwert pro Jahr für jeden Kunden und seine Produkte vorlagen, kann mit Smart-Metering alle 15 min. ein Wert erhoben werden. Diese Veränderung haben eine Flexibilisierung zur Folge. Die Unternehmen in der Energiewirtschaft müssen heutzutage noch schneller auf einen noch flexibleren Markt reagieren. Eine schnelle Verfügbarkeit dieser Informationen und der daraus resultierenden Ergebnisse ist für ein Unternehmen wettbewerbsrelevant. Es gibt in der Ener-

gewirtschaft einige Quellsysteme, die für die Projektgruppe von Bedeutung sind. Deshalb werden im weiteren Verlauf diese Quellsysteme vorgestellt. Zu jedem System werden die Funktion, die Eingabedaten und die Informationen, die das System bereitstellt beschrieben. Abschließend wird kurz auf die Bedeutung des Systems und seiner Informationen für die Projektgruppe eingegangen. Als Schlussbemerkung wird ein Blick auf die Ausgabedaten der Systeme geworfen und welche Informationen das System der Projektgruppe bereitstellen wird.

## **2 Systeme**

In der Energiewirtschaft gibt es eine Vielzahl an Systemen, die Daten erfassen und bereitstellen. Nicht alle sind für die Projektgruppe von Bedeutung. Im Folgenden werden das Customer Relationship Management, das Energiedaten-Management, die Datenbanken der Firma ené't und die Industry Solution Utilities von SAP vorgestellt.

### **2.1 Customer Relationship Management**

Customer Relationship Management bezeichnet den strategischen Ansatz, der zur vollständigen Planung, Steuerung und Durchführung aller interaktiven Prozesse mit den Kunden genutzt wird. Customer Relationship Management hilft einem Unternehmen neue Kunden zu akquirieren und Bestandskunden enger an sich zu binden. Eine langfristige Kundenbindung dient primär dazu den Gewinn zu erhöhen und das Unternehmenswachstum zu fördern. Das kann z.B. durch eine geringere Preissensitivität der gebundenen Kunden erreicht werden. Durch die Analyse von Kundendaten wie gekaufte Produkte, Präferenzen, in Anspruch genommene Dienstleistungen etc. können Kunden personalisierte Vorschläge zu Produkten, Dienstleistungen oder besonderen Angeboten gemacht werden. Im Fokus steht einem Kunden möglichst viele Produkte zu verkaufen und nicht mehr die Produkte möglichst vielen Kunden zu verkaufen, um Marktanteile zu maximieren. Hier findet sich auch das Data-Mining. Es werden Kundengruppen analysiert und aufgrund der erhaltenen Daten Kunden, die einer ähnlichen Gruppe angehören, entsprechende Produkt- und Dienstleistungsvorschläge oder Angebote gemacht. Customer Relationship Management wird häufig mit Unterstützung von Informationssystemen realisiert. [Hol]

Durch den Einsatz von Informationssystemen lassen sich die Kundendaten effizient aus jedem Bereich des Unternehmens zusammenführen. Für eine optimale Kundenbearbeitung muss eine ganzheitliche Abbildung des Kunden im System vorliegen. Es werden Daten über den gesamten Kundenlebenszyklus gesammelt. Dadurch entsteht eine große Menge von Daten und der Aufwand einer Analyse ist beträchtlich. Ohne Unterstützung durch ein Informationssystem ist eine differenzierte Kundbetreuung schwer möglich. Ein besonderes Augenmerk wird auf profitable Kunden gelegt.

Die Informationen, die durch ein Customer Relationship Management System bereitgestellt werden, sind für die Projektgruppe von großer Wichtigkeit. Der Anwendungsfall beinhaltet den direkten Kundenkontakt und das Produktangebot. Die Informationen des



Customer Relationship Management Systems werden benutzt, um dem Kunden ein personalisiertes Angebot zu unterbreiten.

Neben einem personalisierten Angebot muss das Unternehmen auch die Ware liefern können. Das Energiedaten Management hilft Unternehmen dabei, diese Kalkulationen vorzunehmen.

## 2.2 Energiedaten Management

Die Liberalisierung des Strom- und Gasmarktes hat die Grundlage für einen besseren Wettbewerb geschaffen. Die Europäische Kommission hat 1997 mit ihrer Richtlinie 96/92/EC die Grundlage für die Liberalisierung in Europa geschaffen. Durch die Öffnung der Strommärkte sollen die Energiepreise für den Endverbraucher sinken. [Kem] Die Monopolstrukturen sollten aufgebrochen werden und allen Anbietern den freien Zugang zum Markt ermöglichen. Die bestehende Netzinfrastruktur kann als natürliches Monopol betrachtet werden. Der Bau eines eigenen Netzes ist nicht nur kostenintensiv, vorallem stehen umwelt- und landschaftsplanerische Aspekte dem Bau weiterer Leitungen entgegen. Deshalb muss allen Marktteilnehmern der freie Netzzugang ermöglicht werden. In Deutschland ist die Netzinfrastruktur Teil vertikal-integrierter Versorgungsunternehmen. Um zu verhindern, dass die Netze zu unternehmerischen Zwecken missbraucht werden, müssen Erzeugungs-, Übertragungs-, Verteilungs- und nicht-elektrizitätswirtschaftliche Tätigkeiten als rechnerisch getrennte Unternehmensteile geführt werden. [Rid03] Dieser Wandel führt zu erhöhten Anforderungen im Bereich Energiedaten Management.

Das Energiedaten Management ist wichtig für die Energieplanung, die Portfoliooptimierung, die Einsatzplanung, den Energiehandel, das Risikomanagement und die Marktanalyse. Durch die Nutzung der Netzinfrastruktur durch viele Stromanbieter ist es wichtig zu überprüfen welcher Anbieter wie viel Strom in das Netz eingespeist und an Kunden geliefert hat. Sofern keine Smart-Meter verfügbar sind wird der Stromverbrauch von Großkunden im Viertelstundenraster und der von Kleinkunden einmal pro Jahr erfasst. Die Verbrauchsmengen werden anhand von statistischen Daten geschätzt und in Lastprofile umgesetzt. Für jede Kundengruppe gibt es ein entsprechendes Lastprofil (Abb. 2).

Für die Projektgruppe ist insbesondere das Lastprofil von Bedeutung. Wenn ein Unternehmen weitere Kunden aquirieren will, dann muss bekannt sein wie viel Strom diese voraussichtlich verbrauchen werden. Das hat nicht nur Einfluss auf den Preis. Das Unternehmen kann kalkulieren, ob es den benötigten Strom überhaupt produzieren oder liefern kann.

## 2.3 ene't

Ein wichtiger Punkt bei der Kalkulation eines Angebots ist die Marktsituation. Um die Marktsituation zu erfassen muss bekannt sein, welche weiteren Anbieter es bei dem Kunden gibt, welche Produkte sie anbieten und zu welchen Konditionen. Weiterhin sind die Konzessionsabgaben, die Netznutzungsentgelte, die Kraft-Wärme-Kopplungs-Aufschläge

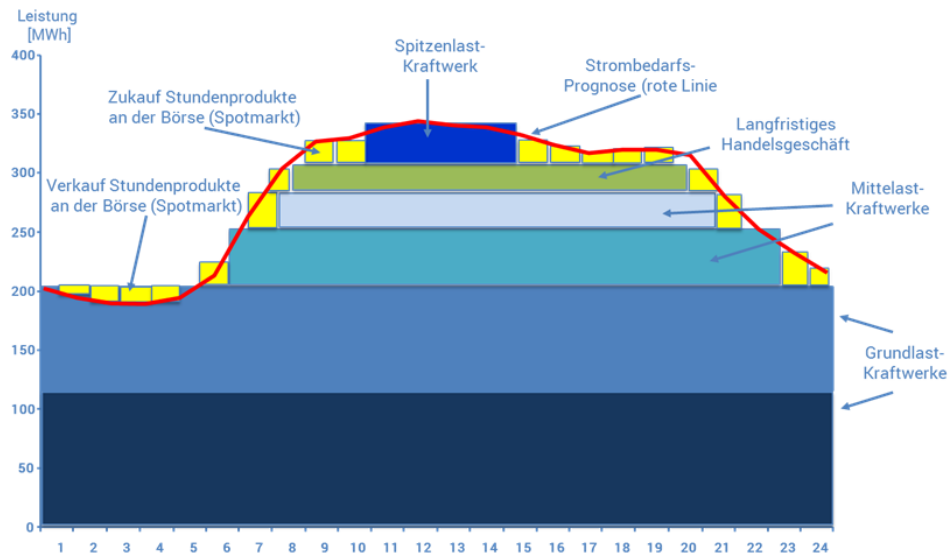


Abbildung 2: Das Diagramm zeigt ein beispielhaftes Lastprofil. Es zeichnet den Leistungsbedarf in mWh in Abhängigkeit von der Tageszeit in Stunden auf. Die eingefärbten Blöcke illustrieren die verschiedenen Lastarten. [Haa]

(im Folgenden KWK-Aufschläge), die Netzinfrastruktur und die Voraussetzungen, unter denen ein Produkt geordert werden kann wichtige Faktoren, die ein Angebot beeinflussen. Die Firma e'net stellt Datenbanken mit entsprechenden Informationen für den Strommarkt und den Gasmarkt zur Verfügung. Diese Datenbanken werden täglich aktualisiert. [eGc] Für die Fallstudie der Projektgruppe sind nur die Datenbanken für den Strommarkt interessant. Zwei Datenbanken werden detailliert betrachtet. Als erstes die Datenbank „Netznutzung Strom“. Als zweites die Datenbank „Endkumentarife“. Die Datenbank „Netznutzung Strom“ enthält Informationen welcher Anbieter, welche Region beliefert. [eGa] Die Datenbank „Endkumentarife“ stellt Informationen über alle Produkte, die an einem Standort verfügbar sind bereit. [eGb] Die Datenbank „Netznutzung Strom“ enthält die folgenden Daten: [eGb]

1. Netzbetreiberstammdaten
2. Netzatlas
3. Preisregelungen
4. Gemeinden und Gemeindegrößen
5. KWK-Aufschläge

Die Netzbetreiberstammdaten umfassen die Adressdaten aller aktuellen und historischen Netzbetreiber. Zudem sind zu jedem Netzbetreiber die Informationen zu den jeweiligen Ansprechpartnern vorhanden.

Der Netzatlas gibt einen Überblick über alle Betreiber des Hoch-, Mittel- und Niederspannungsnetzes. Für jeden Betreiber sind die Lieferorte verzeichnet.

Die Kategorie Preisregelung umfasst alle aktuellen und historischen Preisregelungen.

Die Informationen zu den Gemeinden und deren Größe sind zur Berechnung der Konzessionsabgaben notwendig.

Weiterhin sind alle Kosten die durch das KWK-Gesetz entstehen verzeichnet. Diese Mittel werden zur Förderung und Modernisierung der Kraft-Wärme-Kopplungsanlagen aufgewendet.

Die Datenbank „Endkundentarife“ enthält die folgenden Daten: [eGa]

1. Aktuelle Strompreise aller deutschen Netzanbieter
2. Übersicht über alle angebotenen Produkte
3. Konditionen abhängig vom Standort
4. Informationen unter welchen Voraussetzungen ein Produkt geordert werden kann
5. Zusatzprodukte zu vorhandenen Produkten

Die Daten aus beiden Datenbanken sind für das Erstellen eines Angebots wichtig. In der ersten Datenbank finden sich Informationen zu den Auflagen, die Einfluss auf den Preis haben. In der zweiten Datenbank finden sich alle Informationen zu den Mitbewerbern und ihren Produkten. Auf Grundlage dieser Informationen kann das Angebot optimal für den Kunden und das Unternehmen kalkuliert werden.

Wenn Angebote für Kunden erstellt werden müssen diese auch verwaltet werden. Die Industry Solution Utilities von SAP sind eine Software-Lösung, die die Verwaltung informationstechnisch unterstützt.

## **2.4 Industry Solution Utilities**

Die Industry Solution Utilities (im Folgenden IS-U) sind eine Software-Lösung, die von SAP speziell für Versorgungs- und Dienstleistungsunternehmen angeboten wird. IS-U dient zur Verwaltung und Abrechnung von Kunden. Das umfasst alle Arten von Kunden, Tarfkunden, Sonderkunden und Dienstleistungskunden. Außerdem können auch Interessenten mit der Software verwaltet werden. Zur Prozessoptimierung unterstützt IS-U automatisierte Abläufe wie Neuanlage eines Hausanschlusses oder dessen Sperrung. Die SAP Standardkomponenten Instanthalter, Kundenservice, Vertrieb, Vertragskontokorrent und SAP Business Workflow sind in IS-U integriert.

IS-U hat einen großen Funktionsumfang. Mit den Grundfunktionen lassen sich Adressdaten verwalten und Termine steuern. Die Stammdaten enthalten alle Daten, die für längere

Zeit unverändert bleiben. Es wird zwischen kaufmännischen und technischen Stammdaten unterschieden. Zu den kaufmännischen Stammdaten zählen die Geschäftspartner, deren Vertragskonten und der Vertrag abrechenbarer Service. Hinzu kommen die Serviceanbieter und deren nicht abrechenbarer Service.

Zu den technischen Stammdaten gehören Zählpunkte, deren technische Installation, Anschluss, Anschlussobjekt/TP, die Verbrauchsstelle, der Geräteplatz, das Gerät und die Anlage. Dazu kommen noch Daten über das Netz.

Mit der Komponente „Geräteverwaltung“ lassen sich die Installation, Ablesung und Beglaubigung aller Geräte des Versorgungsunternehmens verwalten. Um diese Funktionalität bereitzustellen, werden andere Komponenten benötigt. Die Daten zu dem Gerät stammen aus der Komponente „Instanthalter“. Der Gerätetyp und die Gerätenummer werden in der Komponente „Grunddaten Logistik“ gespeichert.

Die „Abrechnung“ ermöglicht es die Versorgungs- und Dienstleistungen des Versorgungsunternehmens abzurechnen. Es werden neben Tarifkunden auch Sonderkunden unterstützt. Es können Jahresverbrauchsabrechnungen, aber auch Verbrauchsabrechnungen mit kürzeren Zyklen getätigt werden. Die Komponente „Abrechnung“ bezieht Ableseergebnisse aus der Komponente „Geräteablesung“ und gibt ihre Ergebnisse an die „Fakturierung“ weiter.

Die „Fakturierung“ erstellt Belege für die IS-U Verbrauchsabrechnung. Außerdem verwaltet sie Stornierungsprozesse, Abschlagspläne und unterstützt die Ermittlung von Steuern, Gebühren und Abgaben. Weiterhin werden die Zusatzfunktionen wie Verzinsung, Mahnen und Sperren bereitgestellt.

Mit der Komponente „Kundenservice“ lassen sich Serviceleistung und Geschäftsvorfälle abwickeln. Es werden Daten zu den Konten, Ableseergebnissen und Abschlägen bereitgestellt. Dadurch können Kunden schnell Auskünfte erteilt werden. Die Daten werden über das Customer-Interaction-Center oder das Front-Office präsentiert. Beides beschreibt weitere Komponenten, die in IS-U bei Bedarf integriert werden können. Das Customer-Interaction-Center unterstützt den Kundendienst und Call-Center bei der Kommunikation mit den Kunden. Es bietet eine Telefonsteuerung mit Zugriff auf Geschäftstransaktionen. Das System führt eine umfassende Anrufprotokollierung durch und ermöglicht es das Gespräch aufzuzeichnen. Das Front-Office bietet ähnliche Funktionalitäten an. Allerdings wird hier auf die Verwaltung von Telefongesprächen verzichtet. Darüber hinaus arbeitet die Komponente „Kundenservice“ mit der Komponente „Internet Self Services“ zusammen. „Internet Self Services“ ermöglicht es dem Kunden seine Kundendaten über das Internet selbst zu verwalten. Er kann Einzugsermächtigungen erteilen, Rechnungszahlungen veranlassen, Rechnungen anzeigen lassen, Kontenauskünfte einholen, Einzüge melden, Rechnungsanschriften ändern, Zählerstände erfassen, Versorgerwechsel melden, sich registrieren und Passwörter ändern, Verbrauchshistorien einsehen, Rückrufe initiieren und ein neues Passwort anfordern.

IS-U bietet auch ein Energiedaten Management. Diese Komponente leistet das was im Abschnitt „Energiedaten Management“ (2.2) beschrieben wurde. Es bietet eine zentrale Datenbank für Energiedaten wie Energieverbräuche, Preise etc. und ermöglicht Bilanzierung und Fahrplan-Management. Darüber hinaus lassen sich Daten einzelner Intervalle abrechnen (Real-Time-Pricing-Abrechnung). Ein wichtiges Feature ist der Zugriff auf die Profile. IS-U Energiedaten Management ermöglicht es über das Internet auf Profile zuzugreifen. Mit Hilfe des „Work-Management“ lassen sich Arbeitsaufträge planen, kalkulieren und

ausführen. Neben Arbeitsaufträgen mit Kundenbezug wie Erstellung und Verstärkung eines Hausanschlusses, Energieberatung, außerplanmäßigen Zählerablesungen, Sperrung und Freischaltung ist auch die Verwaltung von Arbeitsaufträgen ohne Kundenbezug möglich. Letztere umfassen Wartungsmaßnahmen, Reparaturen, Neuinstallationen, Erweiterungen oder Demontage von technischen Einrichtungen. Zur Prozessoptimierung gibt es Vorlagen, die durch Eingabe von Parametern angepasst werden können. Das System erstellt dann automatisch den passenden Auftrag.

Die betriebswirtschaftlichen Abläufe eines Entsorgungsunternehmens können mit der Komponente „SAP Waste and Recycling“ unterstützt werden. Die Komponente umfasst Logistik-, Abrechnungs-, Service- und Kundenbetreuungssysteme.

Eine wichtige Komponente ist der „Unternehmensübergreifende Datenaustausch“. Sie verwaltet Deregulierungsdaten wie Zählpunkte, Netze, Serviceanbieter etc., hilft bei der Abwicklung von Datenaustauschprozessen und der Prozessbearbeitung. Durch die Liberalisierung des Energiemarktes müssen bei einem Versorger- oder Anbieterwechsel deutlich mehr Unternehmen miteinander kommunizieren. Deshalb ist ein effizienter Datenaustausch zwischen den Unternehmen wichtig. [SAP]

Für die Projektgruppe sind nicht alle Funktionen und Daten, die IS-U bietet relevant. Speziell die Komponenten, die mit Kundenservice, Abrechnung und Abfall und Recycling in Zusammenhang stehen sind für die Projektgruppe nicht von großer Bedeutung.

### **3 Schlussbemerkung**

Die vorgestellten Systeme stellen alle in der einen oder anderen Form Daten bereit, die für die Projektgruppe wichtig sind. Das Customer-Relationship-Management beinhaltet ein umfassendes Bild des Kunden im Unternehmen. Um ein speziell auf den Kunden zugeschnittenes Angebot erstellen zu können, sind diese Informationen äußerst wichtig.

Das Energiedaten Management ist nicht nur für die Kalkulation der Angebote wichtig, sondern auch für die Produktion des gesamten Unternehmens. Mit Hilfe der Lastprofile kann in einem gewissen Rahmen vorausgesagt werden, wie viel Energie zu welcher Zeit benötigt wird. Diese Profile werden Angeboten zugrunde gelegt, um zu ermitteln, ob und wie der Kunde beliefert werden kann. Dabei sind auch die Informationen aus den Datenbanken von ene't nützlich. Es werden Informationen zur Marktanalyse und Netzinfrastruktur bereitgestellt. Für die Marktanalyse werden die Strompreise aller deutschen Netzanbieter, deren Produkte, Zusatzprodukte und Konditionen bereitgestellt. Ein Netzatlas, die Preisregelungen, die Gemeindegroßen und die KWK-Aufschläge sind die Informationen, die über die Netzinfrastruktur zur Verfügung gestellt werden. Diese Informationen gibt es neben dem Strommarkt auch für den Gasmarkt. Letzterer ist für die Projektgruppe nicht relevant.

SAP stellt mit seinen Industry Solution Utilities eine umfangreiche Software-Lösung für Versorgungs- und Dienstleistungsunternehmen bereit. IS-U bietet eine Reihe von Daten und Funktionalitäten, die für die Projektgruppe relevant sind. Dazu gehören die Grundfunktionen, um Adressen, Regionalstrukturen und Termine für Ablesungen, Abschläge und Abrechnungen zu verwalten. Es werden die Stammdaten wie Geschäftspartner, Verträge, Vertragskonten, Anschlussobjekte, Verbrauchsstellen, Anlagen und Geräteplätze bereitgestellt. Außerdem lässt sich mit IS-U Energiedaten Management betreiben. Naturgemäß kann die Projektgruppe kein marktreifes System entwickeln, sondern nur einen Prototypen. Deshalb gibt es einige Daten und Funktionalitäten der IS-U, die für die Projektgruppe nicht notwendig sind. Insbesondere sind die Daten und Funktionen im Bereich Kundenservice und Abfallbeseitigung und Recycling.

## Literatur

- [eGa] ene't GmbH. Datenbanken Endkumentarife. <https://www.enet.eu/datenbanken-endkumentarife/datenbank.html>. (28.09.2014).
- [eGb] ene't GmbH. Datenbanken Netznutzung. <https://www.enet.eu/datenbanken-netznutzung/datenbanken.html>. (28.09.2014).
- [eGc] ene't GmbH. ene't Datenbanken. <https://www.enet.eu/datenbanken.html>. (28.09.2014).
- [fSidI] Bundesamt für Sicherheit in der Informationstechnik. Cloud Computing Grundlagen. [https://www.bsi.bund.de/DE/Themen/CloudComputing/Grundlagen/Grundlagen\\_node.html](https://www.bsi.bund.de/DE/Themen/CloudComputing/Grundlagen/Grundlagen_node.html). (30.10.2014).
- [Haa] M. Haas. IT Veränderungen durch EEG, Direktvermarktung und virtuelle Kraftwerke. PowerPoint Präsentation von Marco Haas.
- [Hol] H. Holland. Customer Relationship Management. <http://wirtschaftslexikon.gabler.de/Definition/customer-relationship-management-crm.html>. (01.11.2014).
- [Kem] C. Kemfert. Märkte unter Strom: Die Folgen der Strommarktliberalisierung. <http://www.presse.uni-oldenburg.de/einblicke/38/3kemfert.pdf>. (25.10.2014).
- [MG] P. Mell und T. Grance. The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology. <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>. (30.10.2014).
- [Nor] O. Norkus. Projektgruppe Business Intelligence as a Service. <http://www.uni-oldenburg.de/informatik/is/lehre/projektgruppen/business-intelligence-as-a-service/>. (14.11.2014).
- [Rid03] N. Ridder. *Öffentliche Energieversorgungsunternehmen im Wandel: Wettbewerbsstrategien im liberalisierten deutschen Strommarkt*. Tectum Verlag, Marburg, 2003.
- [SAP] SAP. SAP Utilities. [http://help.sap.de/saphelp\\_utilities472/helpdata/de/c6/4dce68eafc11d18a030000e829fbbd/frameset.htm](http://help.sap.de/saphelp_utilities472/helpdata/de/c6/4dce68eafc11d18a030000e829fbbd/frameset.htm). (25.10.2014).
- [TK09] A. Kießling T. Kästner. *Energie in 60 Minuten - Ein Reiseführer durch die Stromwirtschaft*. VS Verlag für Sozialwissenschaften, Wiesbaden, 2009.

# **Seminararbeit**

## **„Grundlagen: Business Intelligence“**

Babak Izadpanah

Department für Informatik  
Abteilung Informationssysteme  
babak.izadpanah@uni-oldenburg.de  
Matrikelnummer: 2054817

Betreuer: Dr. Marco Grawunder und Oliver Norkus, M. Sc.

**Abstract:** In dieser Seminararbeit habe ich mich mit den Grundbausteinen von Business Intelligence beschäftigt.

Durch die Darstellung und Verarbeitung von Daten mit Hilfe von Business Intelligence kann ein System zur Datenverwertung vorgestellt werden. Es zeigt sich, dass ein komplexes System nötig ist, um in der heutigen Zeit schnell auf Informationen zurückgreifen zu können. Durch die Vorstellungen der unterschiedlichen Möglichkeiten der Datenverarbeitung ist zu erkennen, dass viele verschiedene individuell zugeschnittene Lösungsansätze nötig sind, damit wir in der Zukunft auf Daten noch schneller und gezielter zugreifen und so zukünftig Entscheidungen besser treffen zu können.

## **1 Motivation**

Für die meisten Organisationen ist es heute eine Tatsache, dass allgemeine Daten ein Schatz für sie sind und speziell Umfragedaten die "Energiequelle" der digitalen Wirtschaft sind. Die richtige Information zur richtigen Zeit abrufen zu können, ist für ein Unternehmen von entscheidender Bedeutung.

Um wirklich effektiv mit den vielen verschiedenen Daten und Formaten zu arbeiten, ist ein Business Intelligence (BI) System nötig, also ein System, das aus digitalen Daten relevante und einfach abrufbare Informationen machen kann. Technisch gesagt ist dies die unternehmensweite, maßgeschneiderte Optimierung von Datenspeicherung, Aufbereitung und Bereitstellung. Mit BI liefert die Informationstechnologie dem Management eine kaufmännisch-ausgerichtete Hilfe und ermöglicht das Aufstellen von realistischen Prognosen.

## **2 Definition: Business Intelligence**

Business Intelligence klingt für viele Menschen wie ein Marketing-Modewort. Jedoch kennt jeder Unternehmer, vielleicht nur unter einem anderen Namen BI. Intelligenz bedeu-



tet, dass mit Wissen, Verständnis und Bedeutung komplexe Probleme gelöst und dadurch bessere Entscheidungen getroffen werden können. Menschen treffen zu jeder Zeit Entscheidung, aber sie müssen sicher sein, dass die Entscheidung richtig und gut ist. Für eine gute Entscheidung braucht man Informationen (Data) und Intelligenz. Die Informationen bieten Entscheidungshilfen. Deshalb ist es sehr wichtig, dass man zur richtigen Zeit die richtigen Informationen zur Hand hat. Der gesamte Prozess ist eigentlich Business Intelligence und mit einfacheren Worten: Business Intelligence bedeutet der richtigen Person, die richtigen Information zur richtigen Zeit zur Verfügung zu stellen, damit die richtigen Entscheidungen getroffen werden können.

Im Internet gibt es viele verschiedene Definitionen von Business Intelligence z.B.:

Man versteht unter Business Intelligence die Integration von Strategien, Prozessen und Techniken, um aus verteilten und inhomogenen Unternehmens-, Markt- und Wettbewerberdaten erfolgskritisches Wissen über Status, Potentiale und Perspektiven zu erzeugen [SO09].

und Business Intelligence kann man außerdem wie folgt definieren:

”Business Intelligence is the process of transforming data into information and, through discovery into knowledge [Sch06]”. Mit einfachen Worten ist „Business Intelligence“, ein zusammenfassender Prozess (Tools, Wissen, Technologie, Risiken und Anwendung), der die Vergangenheit analysiert und so die zukünftigen Entscheidungen verbessern kann. Durch BI werden die richtigen Informationen in die richtigen Hände gegeben und dies geschieht in einem Format, das die Daten sehr schnell verständlich macht.

Man kann nicht sagen das BI nur Tools wie Oracle oder SQL beinhaltet oder lediglich Data-Warehouse und Data-Marts erfordert. BI hilft dem Management zu verstehen, wo die richtigen Änderungen vorzunehmen sind und welche Veränderungen positive, negative oder die stärksten Auswirkungen haben.

Der grundlegende Bestandteil von BI ist ein zentraler Datenbestand, üblicherweise ein Data-Warehouse. Zentral ist hier im Sinne analogischer Datensicht gemeint und bedeutet das BI Prozesse zusammenfasst, die einer Organisation helfen ihre Leistungsfähigkeit zu steigern.

Allerdings gibt es einige Herausforderungen, z.B. liegen verteilte Daten in verschiedenen Systemen und Entscheidungen werden auf Basis unvollständiger Daten getroffen. Die Auswertung großer Datenmengen ist ohne die Unterstützung durch IT-Systeme schwierig. Mit BI können diese Missstände behoben werden und nützliche, umfassende und zuverlässige Daten können zu einer erhöhten Qualität der Entscheidungen beitragen.

### **3 OLAP vs. OLTP**

OLAP ist die Abkürzung für „Online Analytical Processing“. OLAP ermöglicht es Analyseprozesse auf Unternehmensdaten interaktiv (online) durchzuführen. OLAP-Systeme werden im Dialogbetrieb benutzt. Dabei muss die Antwortzeit möglichst kurz gehalten werden. Deshalb sind OLAP von OLTP logisch und physikalisch getrennt.

Die Daten werden bei OLAP subjektbezogen und verdichtet unter spezieller Berücksichtigung historischer Verlaufsdaten vorgehalten. Aktualisierung von Daten ist durch Abzüge

(„Snapshots“) operativer Systeme möglich. Nur das Lesen ist von verfügbaren Daten ist schnell. Bei der Historisierung muss mit einer mittleren bis langen Antwortzeit gerechnet werden, weil lange Lesetransaktionen auftreten.

Der operative OLTP-Ansatz oder „Online Transaction Processing“ unterscheidet sich vom OLAP-Ansatz. Beim OLTP-Ansatz wiederholen sich die Datenbankprozesse häufig, sind strukturiert und bestehen aus isolierten, atomaren Transaktionen. Diese arbeiten mit aktuellsten Daten und greifen lesend sowie schreibend meist nur auf wenige Datensätze über Primärschlüssel zu. Bei OLAP hingegen steht die historisch aggregierte Information im Vordergrund. Die Analysen sind komplexer Abfragen und der Zugriff erfolgt meist nur lesend. Die Tabelle fasst die wichtigsten Unterschiede stichwortartig zusammen.

Beim OLTP-Ansatz gibt es viele Benutzer und kurze, einfache Transaktionen. Hinzu kommt die Echtzeitanforderung und das kurze Gedächtnis, bspw. Online-Shopsysteme oder Bankautomaten. Dagegen gibt es beim OLAP-Ansatz weniger Benutzer und komplexere Queries mit langer Laufzeit. Aber es treten nur lesende Transaktionen und die Analyse historischer Daten mit Data Mining Verfahren und Controlling auf. Um die Daten nutzen zu können, müssen vorher einige Probleme gelöst werden:

- verteilte Daten liegen auf verschiedenen Systemen
- Heterogenitätsproblem
- oftmals verborgene Änderungen in der Semantik der Daten
- erfordert Transport großer Datenmengen über das Netzwerk
- bei wachsenden Datensätzen kann man leicht den Überblick verlieren
- historische Sicht-Datenmengen wachsen immer weiter
- operative Systeme brauchen die historischen Daten nicht
- Manager brauchen viele der operativen Daten nicht
- Komplexität der Transformationen

Die Lösung ist ein DWH. Dieses System wird im folgenden vorgestellt. Hierbei werden die technischen Grundlagen und Funktionsweisen eines DWH Systems erläutert.

## **4 Data Warehouse-Architektur**

Der Vater des DWH (William Inmon) hat DWH wie folgt definiert: „A DWH is a subject-oriented, integrated, non-volatile, and time-variant collection of data in support of management’s decision“ [Inm05].

Ein DWH ist subject-oriented, weil die Daten für Verkäufer und Personen, Produkte, etc. sind. Die Daten stammen aus verschiedenen Quellen (integriert) und sie werden ohne Änderung über die Zeit (nichtflüchtig) gespeichert. Der Vergleich von Daten über die Zeit

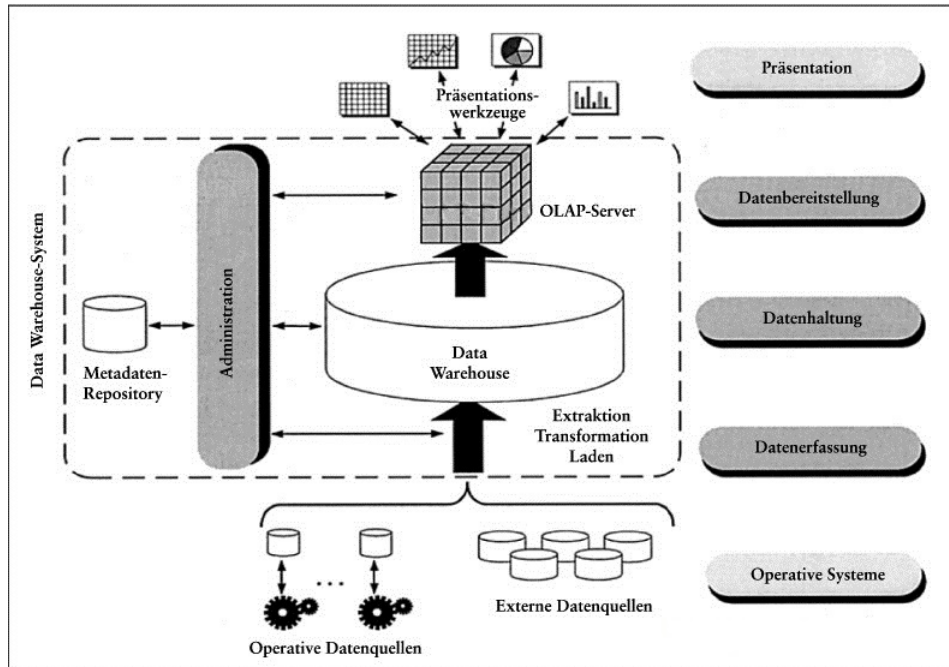


Abbildung 1: Architektur eines Data Warehouse-Systems [Bau]

ist ebenfalls möglich (zeitvariant).

Im Internet gibt es viele verschiedene Definitionen zu DWH z.B. Data Warehousing ist kein Produkt, sondern der Prozess der Zusammenführung und des Managements von Daten aus verschiedenen Quellen mit dem Zweck, eine einheitliche, detaillierte Sicht auf den einzelnen Geschäftsbereich oder das gesamte Unternehmen zu erhalten.

Jeder Produktionsfaktor eines Unternehmens basiert auf vier Produktionsfaktoren, Arbeit, Lage, Kapital und Wissen. Letzterer avanciert beim richtigen Einsatz zum entscheidenden Wettbewerbsfaktor. Informationen sind hierbei Bestandteil von Wissen. Die Unternehmen erhalten ihre Informationen aus internen oder externen Quellen. Die Verarbeitung dieser Informationen wird schwerer, wenn die Daten in heterogener Form, in unterschiedlicher Qualität, in verschiedenen Datenformaten oder in unterschiedlichen Datenquellen vorliegen. Mit einem DWH können die heterogenen Daten vereint und eine analytische Verwendung ermöglicht werden.

”Data Warehouse is a subject oriented, integrated, non-volatile, and time variant collection of data in support of management’s decision-making process [Inm96].” In der Praxis reicht die Definition von Inmon nicht aus. Ein Data Warehouse ist eine physische Datenbank, die eine integrierte Sicht auf beliebige Daten zu Analysezwecken ermöglicht. Das DWH ist dabei ein Teil von BI. BI führt die Integrationsmöglichkeiten von Strategien, Prozessen, Anwendungen und Technologien mit dem Wissensmanagement zusammen.

#### 4.1 ETL-Prozess (Extraktion, Transformation und Laden)

Unter ETL versteht man den Prozess der Datenintegration in ein DWH. Dabei werden die Daten aus unterschiedlichen operativen Quellsystemen extrahiert, unter Einhaltung der Datenqualität umgewandelt, bereinigt, angeglichen, transformiert und abschließend in ein Zielsystem integriert und geladen.

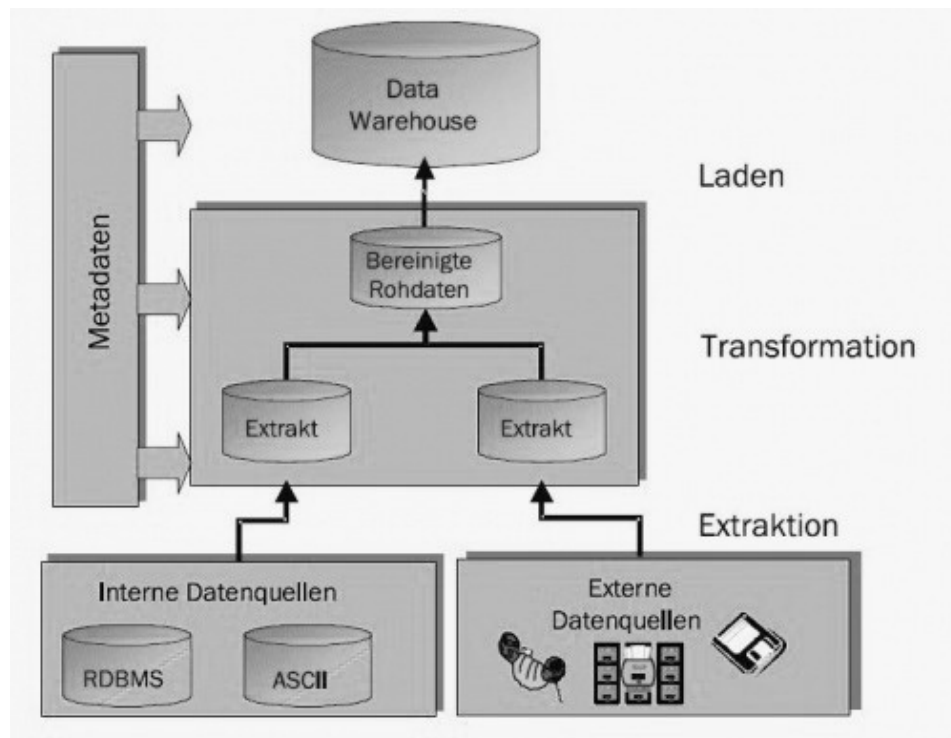


Abbildung 2: Extraktion, Transformation und Laden (ETL) [Manb]

##### 4.1.1 Anforderungen von DWH-Data

Um einen sicheren und effizienten ETL-Prozess und DWH-Umgebung zu haben und die genannten Probleme (3) zu vermeiden, gibt es Vorgaben, die eingehalten werden müssen. Die Vorgaben wurden von W.H. Inmon für ein DWH und den damit verbundenen ETL-Prozess wie folgt festgelegt:

- Subjektorientierung:  
Informationen sollten sich nicht an den operativen Geschäftsprozessen orientieren, sondern an den Subjekten des Unternehmens. Subjekte sind z.B. Produktgruppen oder der Kundenstamm.

- **Integration:**  
Informationen sind oftmals nur in verschiedenen Formaten und Strukturen vorhanden. Daher sollte eine Anpassung auf die von dem Unternehmen gewünschte Formatvorlage erfolgen, um Analysemöglichkeiten von einer einheitlichen Datenbasis zu gewährleisten.
- **Zeitraumbezug:**  
Auf der kleinsten Datensatzebene sollte, sofern nicht vorhanden, ein einheitlicher Zeitraumbezug verwendet werden. Dies ermöglicht Zeitreihenanalysen und Erstellung von Trends.
- **Nicht-Volatilität:**  
Volatilität bedeutet zu einem speziellen Zeitpunkt eine Veränderung an den vorhandenen Daten. Daten sollten bei einem DWH niemals verändert werden, sondern nur archiviert und erneut geladen.

#### **4.1.2 Daten Extraktion**

Am Anfang werden die relevanten Daten und Datenstrukturen aus den verschiedenen Quellsystemen in den temporären Arbeitsbereich (staging area) übernommen. Um die Daten in einem DWH auf einen aktuellen Stand zu halten, bietet es sich an die Extraktion in einem regelmäßigen zeitlichen Ablauf zu initiieren. Dies kann entweder synchron oder asynchron gesteuert werden, wobei bei der asynchronen Extraktion zwischen periodischer, ereignisgesteuerter oder anfragegesteuerter Extraktion unterschieden wird.

#### **4.1.3 Daten Transformation**

Die Datentransformation im ETL-Prozess ist die zentrale Komponente. Die Daten müssen an die Schemata und Datenqualität der Vorgaben, der Anwender und Zielstrukturen angepasst werden. Die Transformationen können in zwei Bereiche eingeteilt werden: Datenintegration und Datenbereinigung.

Datenintegration beschäftigt sich mit der Beseitigung von semantischen Mängeln und überprüft die Daten auf inhaltliche Aspekte.

Datenbereinigung umfasst die Korrektur syntaktischer Mängel. Dabei wird versucht bei den zu integrierenden Daten festgestellte Qualitätsmängel zu beseitigen. Qualitätsmängel treten oft im Zusammenhang mit mehreren Datenquellen auf.

#### **4.1.4 Daten Laden**

Am Ende des ETL-Prozesses werden die bereinigten und transformierten Daten aus dem temporären Arbeitsbereich physisch in die Datenbank des DWH geladen. Um die Datenintegrität nicht zu gefährden, ist der Zugriff für die Endanwender, zu diesem Zeitpunkt auf sämtliche abhängige Analysetools und OLAP-Dienste gesperrt. Es wichtig, sicherzustellen, dass der Prozess höchst effizient durchgeführt wird, um die Wahrscheinlichkeit eines Ausfalls zu minimieren. In Abhängigkeit von den, in der Extraktion geladenen Daten

muss zwischen einem erstmalig initiierten Fall und einer regelmäßigen Aktualisierung (inkrementelles Laden) unterschieden werden. Beim erstmaligen Laden werden alle wichtigen Daten aus den Quellen in die DWH-Datenbank übernommen, während beim inkrementellen Laden nur Änderungsdaten eingespielt werden.

## 4.2 DWH-Cube (Fakten und Dimensionen)

”Das multidimensionale Datenmodell wird oft kurz mit dem Begriff Datenwürfel, engl. data cube, charakterisiert.” Dieser Datenwürfel ist eine mehrdimensionale Darstellung der Analysedaten (Kennzahlen). Die Achsen eines dreidimensionalen Datenwürfels werden auch als Dimensionen bezeichnet und sind in der Regel hierarchisch gegliedert. So werden als Beispiel für Hierarchie einzelne Filialen in einer Stadt zusammengefasst und mehrere Städte einem Land zugeordnet. Dies ermöglicht innerhalb von Dimensionen eine Betrachtung mit unterschiedlichen Verdichtungsstufen.

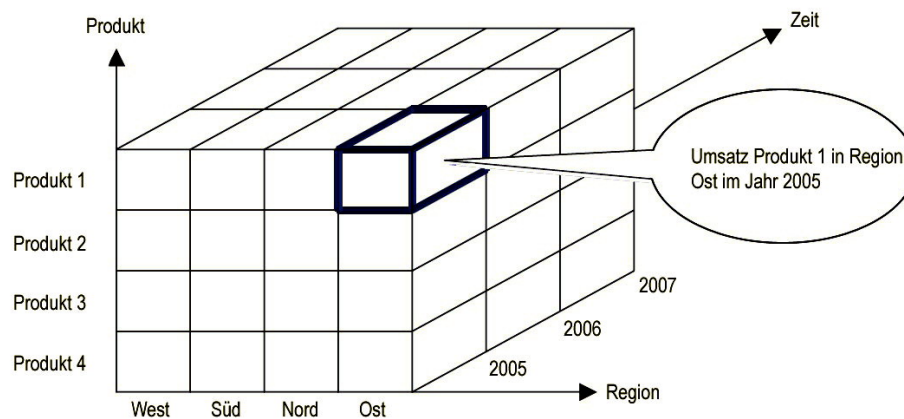


Abbildung 3: Daten als Würfel mit drei typischen Dimensionen [Mana]

## 4.3 Slicing- Dicing- Roll-Up- Drill-Down

Mit Hilfe des Würfels, seinen Dimensionen und verschiedenen Operationen kann eine Auswahl von bestimmten Daten (z.B alle Produkte) getroffen und über Daten aggregiert (z.B alle Produkte von West) werden. Als Beispiel könnte man anstatt über ein Jahr zu aggregieren, einzelne Monate (Drill-Down) betrachten. Ein entgegengesetztes Vorgehen ist ebenfalls möglich, statt den Monat zu betrachten, können die Jahre (Roll-Up) betrachten.

#### **4.4 Data Marts**

Wenn auf Daten die verschiedenen Operationen (Drill-Down, Roll-Up, Slice, Dice, Pivot) angewandt werden, dann sind die Ergebnisse Data Marts. Data Marts enthalten die Daten spezieller Form. Sie stellen eine analyseorientierte Sicht auf Teile des DWH-Datenbestandes dar.

#### **4.5 Datenanalyse (OLAP)**

Analysen können mit festen Reportings oder mit flexiblen Werkzeugen geplant, regelmäßig, kontinuierlich, ad-hoc oder mit Data Mining Verfahren durchgeführt werden.

#### **4.6 Metadaten-Repository**

Wesentlicher Erfolgsfaktor eines DWH sind die Metadaten. Metadaten dienen zur technischen Beschreibung des DWH, zur Vermeidung von Fehlinterpretationen und Nachvollziehbarkeit der Prozesse.

Die Metadaten enthalten die Schemata, die Dokumentation, Quellbeschreibungen, Datentypen, die Prozessbeschreibungen des Konfigurationsmanagements und der Verarbeitung.

### **5 Data Warehouse-System (DWS) -Referenzarchitektur**

Ein DWH ist eine Datenbank, die die Aufgabe hat integrierte Daten für Analysezwecke bereitzustellen. Die Architektur muss den Anspruch der unabhängigen Nutzung einzelner Komponenten - hinsichtlich der Auslastung - erfüllen. Weiterhin muss die längerfristige Bereitstellung und die Mehrfachverwendbarkeit integrierter Daten gewährleistet werden. Die DWH-Architektur muss die Möglichkeit der Erweiterbarkeit bieten, um ggf. neue Datenquellen zu integrieren. Ein weiteres Kriterium ist die automatisierte Verarbeitung innerhalb der einzelnen Prozessschritte.

Nachfolgend wird das Referenzmodell für die Architektur von DWH-Systemen (nach Bauer u. Günzel 2009) sowie die zugrundeliegenden Bausteine der Referenzarchitektur vorgestellt. Das Referenzmodell stellt einen Bauplan bereit, welcher die oben beschriebenen Anforderungen an eine DWH-Architektur erfüllen soll.

#### **5.1 Referenzarchitektur**

Eine Referenzarchitektur ist ein Modellmuster, also ein idealtypisches Model. Dieses Modell soll die gestellten Anforderungen erfüllen und muss gegenüber Änderungen robust

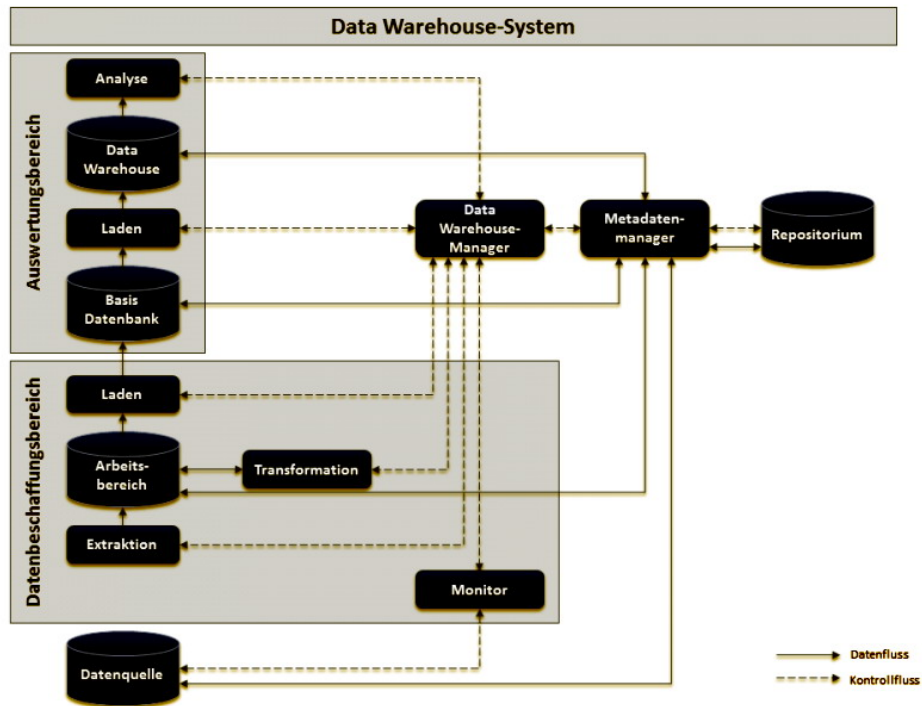


Abbildung 4: DWS-Referenzarchitektur [BGVZ09]

sein. Auch eine gewisse Ästhetik soll die Referenzarchitektur aufweisen.

**Komponentenübersicht** Anhand der Referenzarchitektur gibt es die nachfolgenden Einzelkomponenten in einem DWH-System.

**Data Warehouse Manager** Die zentrale Steuerungskomponente des DWH-Systems ist der Data Warehouse Manager (DWHM). Das Hauptaugenmerk liegt auf der Initiierung und Steuerung des Datenbeschaffungsprozesses, sowie der Regulierung der Analyseanfragen. Der Datenbeschaffungsprozess wird in regelmäßigen Zeitintervallen, durch Datenänderung oder durch explizites Verlangen eines Benutzers angestoßen. Während der Datenverarbeitung steuert der DWHM die korrekte Reihenfolge der Verarbeitung. Verarbeitungsfehler werden vom DWHM zentral dokumentiert und kommuniziert mit dem Metadatenmanager.

**Datenquellen** Datenquellen sind Systeme, die Daten bereitstellen. Sie enthalten die zu integrierenden, heterogenen, realen Daten. Die Daten sind nicht Bestandteil des DWH-Systems. Allerdings sind die Quellen der Ausgangspunkt des Datenflusses innerhalb des



DWH-Systems. Da die Datenquellen zu Analysen genutzt werden, ist die Auswahl geeigneter Datenquellen und somit die Datenqualität ein wichtiger Erfolgsfaktor bei der Integration eines DWH-Systems.

**Monitor** Der Monitor überwacht die Datenquellen auf Änderungen. Innerhalb der Datenquellen erkennt der DWH-Monitor, wenn Daten manipuliert wurden und kümmert sich um die inkrementelle Alternierung der Daten in der Basisdatenbank und des DWH. Ein vollständiges Löschen und anschließendes Laden ist aufgrund von Performanceüberlegungen nicht praktikabel. Wegen der Heterogenität der Quellen existiert im Allgemeinen pro Datenquellen ein Monitor. Jeder Monitor verfolgt eine eigene Monitoring-Strategie, um Änderungen zu identifizieren.

**Arbeitsbereich** Der Arbeitsbereich ist der temporäre Speicherort des ETL-Prozesses. Notwendige Transformationen werden direkt in diesem Zwischenspeicher durchgeführt ohne, dass der laufende Betrieb, der Datenquellen und der Basisdatenbank beeinflusst wird. Erst nach Abschluss aller Verarbeitungsschritten werden die Daten in die Basisdatenbank geladen und anschließend im Arbeitsbereich gelöscht.

**Extraktionskomponente** Bei der Extraktion werden die relevanten Daten in den temporären Arbeitsbereich kopiert. Die Extraktionskomponente ist für die Übertragung der Daten aus den Datenquellen in den Arbeitsbereich zuständig. Abhängig von der Monitoring-Strategie unterscheidet sich die eingesetzte Technik zur Extraktion der Daten. Eine weitere Aufgabe der Komponente ist es, die Auswahl von Quellen bzw. den Ausschnitt von Quellen zu steuern. Hierzu ist es erforderlich festzulegen, wie die Daten im Data Warehouse ausgewertet werden sollen.

**Transformationskomponente** Durch Transformation werden die Daten im Arbeitsbereich transformiert, bereinigt und integriert. Die Transformationskomponente verändert die Struktur und den Inhalt der Daten mit dem Ziel, die heterogenen Daten in ein einheitliches internes Format zu überführen. Man bezeichnet diesen Prozess der Datenstandardisierung auch als Datenmigration. Des Weiteren werden unter Verwendung von Plausibilitätsprüfungen Datenverunreinigungen erkannt und korrigiert. Dieser Verarbeitungsschritt wird als Datenbereinigung bezeichnet. Sind alle Transformationsschritte durchgeführt worden, erhält man die integrierten Daten.

**Ladekomponente** Beim Laden werden die Daten als erstes in die integrierte Basisdatenbank kopiert und anschließend in das DWH geladen. Sie dienen als Grundlage für verschiedene Analysen. Mit anderen Worten: die Ladekomponente ist für die Übertragung der Daten aus dem Arbeitsbereich in die Basisdatenbank bzw. in das DWH zuständig. Dieser Prozess erfolgt direkt im Anschluss an die Transformationsschritte. In der Referenzarchitektur gibt es zwei Ladekomponenten. Die Erste lädt die analyseunabhängigen Daten in die Basisdatenbank und die Zweite überträgt die analysespezifischen Daten in das DWH. Existiert keine Basisdatenbank gibt es nur eine Ladekomponente.

**Basisdatenbank** Die Basisdatenbank im DWH-Referenzmodell ist zwischen dem Arbeitsbereich und dem DWH angesiedelt. Sie sammelt und integriert umfassend in Zeit und Granularität den operativen Datenbestand (Integrationsfunktion).

**Data Warehouse** Das DWH ist eine Datenbank, die im Hinblick auf Analysenanfragen aufgebaut wird und sich damit grundlegend von der Basisdatenbank unterscheidet. Die dort gespeicherten Daten sind standardisiert bzw. integriert und werden dauerhaft vorgehalten.

**Analyse** Unter Analyse versteht man die eigentlichen Operationen auf Daten des DWH. Die Analyse umfasst alle Operationen, die mit den Daten des DWH durchgeführt werden können. Neben der Abfrage und Darstellungen von Daten, werden Analysefunktionen genutzt, um neue Informationen für Folgesysteme zu generieren. Dies kann durch einfache arithmetische Operationen (z.B. Aggregation) bis hin zu komplexen statistischen Untersuchungen (z.B. Data Mining) erfolgen.

**Repository** Im Repository werden die Metadaten des DWH-Systems abgelegt. Die Metadaten enthalten Informationen, die den Aufbau, die Wartung und die Administration des DWH-Systems vereinfachen und die Informationsgewinnung mit Hilfe des DWH-Systems ermöglichen.

**Metadatenmanager** Der Metadatenmanager unterstützt Prozesse der Versionierung und Integration, des Konfigurationsmanagement, für den Zugriff, Anfragen und Navigation auf den Metadaten. Diverse Komponenten des DWH-System greifen auf die Metadaten für die Entwicklung, Administration und Analyse sowie die Navigation und das Editieren zu. Hierfür stellt der Metadatenmanager entsprechende Lese- und Schreibschnittstellen bereit.

**Data Mart** Das Konzept des DWH soll eine zentrale Lösung für analytische Verarbeitung in einem Unternehmen anbieten. Ein Data Mart (DM) ist eine Datenstruktur, die spezialisierte Daten zugeschnitten an die analytischen Anforderungen eines Unternehmensbereichs, vorhält, bspw. Daten für Finanz-, Rechnungs-, Personalwesen etc. Es gibt zwei Typen von DMs. Der Erste ist ein abhängiger DM und er beinhaltet eine Teilmenge eines DWH, die nach Organisationsbereichen gegliedert ist. Der Andere, ist ein unabhängiger DM. Das Konzept der unabhängigen DMs besteht darin, dass einzelne Organisationsbereiche (z.B. Abteilung) sich eigenständige DMs aufbauen. Die Integration der Daten erfolgt nicht über ein DWH, sondern wird direkt aus den Daten der operativen Systeme vorgenommen.

## **6 Risiken von DWH**

### **6.1 Fehlende Planung**

Die DWH-Einführung sollte gut geplant werden. Dies beinhaltet das Aufstellen eines realistischen Budgets und Zeitplans. Zudem sollten die dazu benötigten Ressourcen genau bestimmt werden.

Das Beitreiben eines DWH benötigt Pflege und kostet somit Zeit und Geld. Wird keine vorzeitige und vollumfängliche Planung durchgeführt, kann es zu unvorhergesehene Schwierigkeiten kommen. Insbesondere sind hier finanzielle Probleme zu nennen.

### **6.2 Falsche Daten**

Die richtige Auswahl von Datenquellen ist von zentraler Bedeutung. Analysen sollten immer auf qualitativ hochwertigen Daten basieren, um eine möglichst Entscheidungsgrundlage zu bilden. Zudem ist eine fortlaufende Zuführung von aktuellen Daten notwendig. Werden die falschen oder veraltete Daten erfasst, kann es zu schwerwiegenden Fehlentscheidungen kommen.

### **6.3 Falsche Interpretation**

Auch DWH-Systeme können nicht alle relevanten Daten erfassen und richtig analysieren. Daher kann es auch zu Fehlinterpretationen der Daten kommen. Es kann also geschehen, dass der Kunde für ihn irrelevante Informationen erhält.

### **6.4 Einführung der Benutzer**

Die Endbenutzer sollten mittels Training mit dem System vertraut gemacht werden. So soll sichergestellt werden, dass diese den richtigen Umgang erlernen und die Möglichkeiten des DWH kennenlernen. Fehlende Kenntnisse führen zu Demotivation der Mitarbeiter und zu falschen Entscheidungen.

## **7 Analyse-OLAP**

Unter Analyse versteht man den Prozess der Untersuchung und Präsentation der Daten mit unterschiedlichen Methoden. Das DWH verwaltet die, für die Analyse relevanten Daten und bietet verschiedene Funktionen für deren Auswertung. Dabei dienen die Daten als

Grundlage für operative und strategische Entscheidungen und werden von den Anwendern in unterschiedlicher Weise ausgewertet.

Die Abkürzung OLAP steht für die Bezeichnung "Online Analytical Processing" und ist ein Analyseverfahren, bei dem nicht mehr einzelne Werte aus dem Datenbestand abzurufen, sondern eine multidimensionale Sicht auf die Datenbasis zu erhalten. Dabei spielt das multidimensionale Datenmodell eine zentrale Rolle. Die multidimensionale Sicht ist erforderlich, um Entscheidungen zu treffen, schnell und auf vielschichtige Weise auf Informationen zugreifen zu können. Dazu reicht eine zweidimensionale Sicht auf Informationen, wie bspw. in einer Tabelle nicht aus. In einer Tabelle können nur die Zusammenhänge zweier Variablen dargestellt werden. Um nun weitere Variablen in die Betrachtung einbeziehen zu können, bietet die OLAP-Technologie die Darstellung der Werte in Form eines multidimensionalen Würfels.

## 7.1 Funktionen und Operationen

Das Kernstück der Datenanalyse mit OLAP ist die Betrachtung eines multidimensionalen Datenmodells, welches häufig in Form eines Würfels dargestellt wird. Die folgenden Operationen und Funktionen können auf diesen Würfel angewandt werden:

**Pivotierung (Rotation)** Bei dieser Operation kann der Anwender den Würfel durch Vertauschen der Dimensionen um seine eigene Achse drehen und die Daten aus beliebigen Perspektiven betrachten.

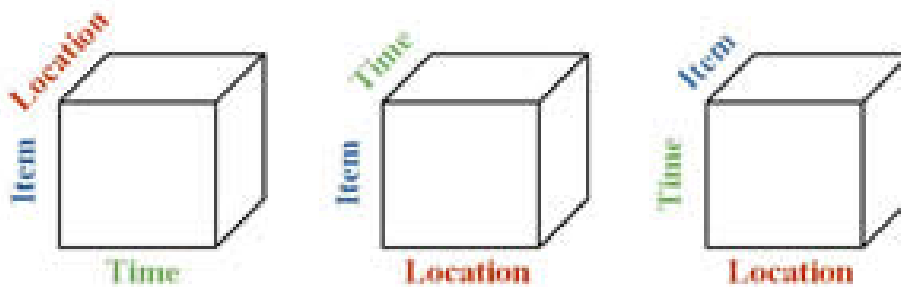


Abbildung 5: Pivotierung [Chi]

**Roll-Up** Bei dieser Operation wird die detaillierte Sicht der Daten wieder verdichtet.

**Drill-Down** Beim Drill-Down ist es möglich aus dem Würfel der verdichteten (aggregierten) Daten, die detaillierte Sicht der Daten zu erstellen.

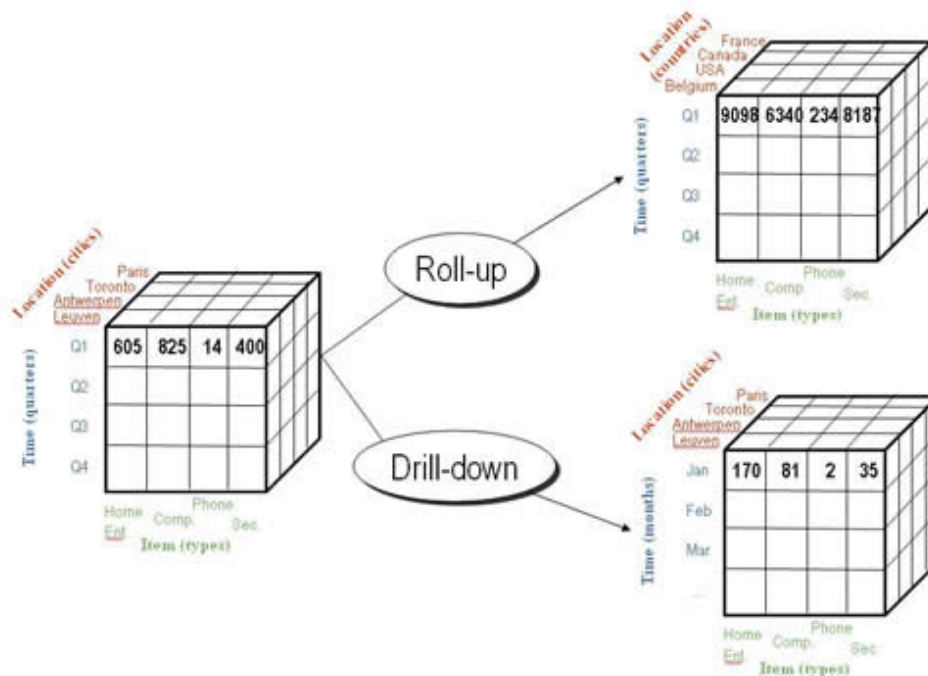


Abbildung 6: Roll-Up und Drill-Down [Chi]

**Drill-Across** Hierbei wird die Sicht von einem Würfel auf einen anderen Würfel gerichtet.

**Slice** ist das Herausschneiden einer Scheibe aus dem Würfel, um damit eine bestimmte Sicht auf einen Teil der Daten zu erhalten. Bildlich lässt sich dieses wie in Abbildung 7 darstellen.

**Dice** lässt sich in zwei Formen beschreiben: Zum Einen kann der Würfel gedreht werden, um so einen Überblick aus verschiedenen Blickwinkel über die Daten zu erhalten. Zum Anderen wird über das Dice ein Teilwürfel des gesamten Würfel betrachtet. Dies entspricht Ad-hoc-Anfragen.

## 7.2 OLAP Modelle

In der OLAP Technologie sind verschiedene Datenmodelle zu unterscheiden, die die Daten, die häufig in einer relationalen Datenbank vorliegen, aus dem DWH aus einer Tabellen in die Zellenstruktur transformiert.

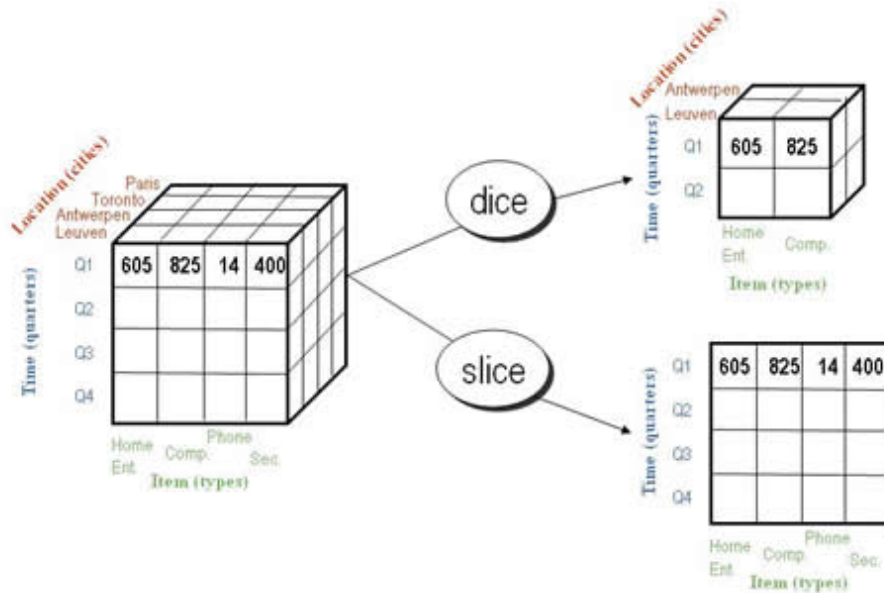


Abbildung 7: Dice und Slice [Chi]

### 7.2.1 ROLAP (Relational OLAP)

Bei der Analyse der Daten werden diese direkt aus der Datenbank geladen und virtuell zur Weiterverarbeitung zur Verfügung gestellt. Die Speicherung erfolgt im relationalen System. Die generierten Berichte und deren Basisdaten werden in einer relationalen Datenbank gespeichert.

### 7.2.2 MOLAP (Multidimensional OLAP)

Die Daten werden multidimensional gespeichert. Die erzeugten Berichte und Basisdaten werden in einer multidimensionalen Speicherstruktur abgelegt.

### 7.2.3 DOLAP (Desktop OLAP)

Bei dieser Variante wird die Verarbeitung direkt auf dem Client vorgenommen. Es handelt sich um eine dezentrale Verarbeitung, der zu analysierenden Daten.

### 7.2.4 HOLAP (Hybrid OLAP)

HOLAP bietet die Vorteile von ROLAP und MOLAP. In der Praxis findet oft eine Mischform des OLAP statt. Die einzelnen Modelle haben individuelle Vor- und Nachteile, die

im Praxiseinsatz gegeneinander abzuwägen sind. Hierbei werden die Eigenschaften des ROLAP und MOLAP genutzt, um die Vorteile der relationalen und multidimensionalen Datenbanksysteme zusammenzuführen.

## 8 Data Mining

Das Data Warehouse stellt die relevanten Unternehmensdaten mit neutral angelegten Datensätzen, auf relativ niedriger Verdichtungsstufe zur Verfügung und bildet so die Grundlage für leistungsfähige Instrumente der Datenanalyse, wie zum Beispiel das Data Mining. Aufgabe des Data Mining ist die Filterung der Daten, um sie für eine weitergehende Analyse vorzubereiten.

Unter Data Mining versteht man eine Menge von Datenanalysemethoden. Umstritten bleibt jedoch welche konkreten Verfahren dem Data Mining zuzuordnen sind. Eine allgemein anerkannte Definition beschreibt Data Mining als nicht triviale Entdeckung gültiger, neuer, potentiell nützlicher und verständlicher Muster in großen Datenbeständen.

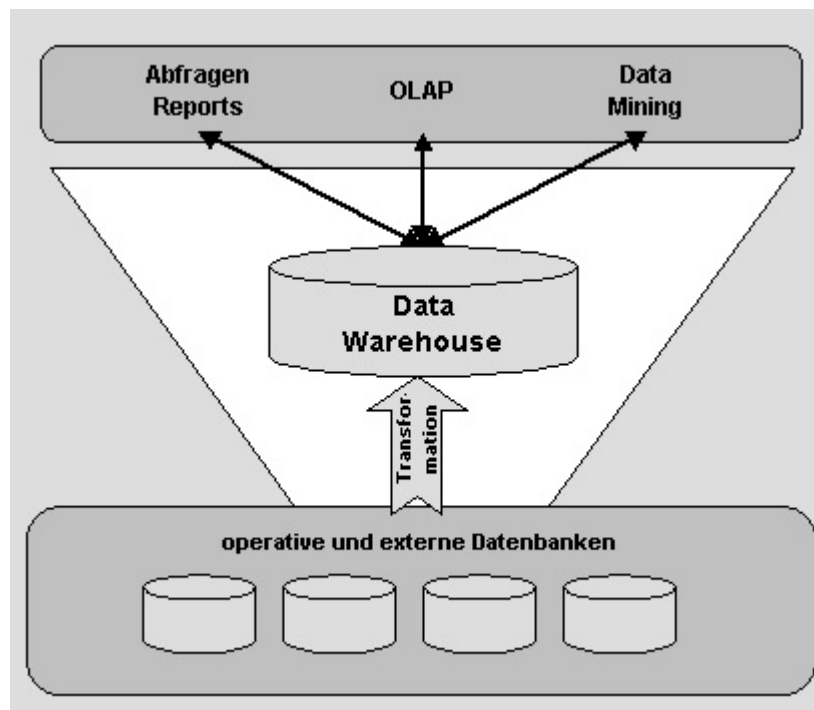


Abbildung 8: Die Architektur eines Data Warehouse [Kon]

Data Mining bezeichnet die Ermittlung "bisher unbekannter Zusammenhänge, Muster und Trends und ist ein Analyseansatz, der darauf abzielt, Beziehungsmuster, wie z.B. Regelmäßigkeiten und Auffälligkeiten, in den zugrunde liegenden Daten zu ermitteln und

durch logische oder funktionale Beziehungszusammenhänge abzubilden". Die ermittelten Muster und Werte müssen dann in ansprechender Form aufbereitet werden, damit diese weiter verarbeitet werden können. Dazu gibt es eine Reihe von Visualisierungsinstrumenten, fallbasierte Systeme, Clusterverfahren, Entscheidungsbaumverfahren und konnektionistische Systeme. Welches Verfahren verwendet wird, hängt von der Zielsetzung des Data Mining, von den Eigenschaften der zu analysierenden Daten und von der Darstellungsform der zu ermittelten Beziehungsmuster ab.



## Literatur

- [Bau] H. Bauer, A./Günzel. Data Warehouse-Systems. <http://www.dasWirtschaftslexikon.com>. Abgerufen am:20.10.2014.
- [BGVZ09] Andreas Bauer, Holger Günzel, Anca Vaduva und Thomas Zeh. Aspekte einer Referenzarchitektur. *Data-Warehouse-Systeme. Architektur Entwicklung Anwendung*, 3:33–38, 2009.
- [Chi] Yoo Myung Chio. OLAP Tutorial. <http://www.cis.drexel.edu>. Abgerufen am:24.10.2014.
- [Inm96] W Inmon. Building the Data Warehouse, New York, Chichester, Brisbane, Toronto, Singapur, 1996.
- [Inm05] W Inmon. H.(2005): Building the Data Warehouse, 2005.
- [Kon] M. Konetzny. Data Warehouse. <http://www.mkonetzny.de>. Abgerufen am:3.11.2014.
- [Mana] Dr Klaus Manhart. BI-Analysemethoden OLAP and Data Mining. <http://www.techannel.de>. Abgerufen am:24.10.2014.
- [Manb] Dr Klaus Manhart. Datenaufbereitung durch den ETL-Prozess. <http://www.techannel.de>. Abgerufen am:24.10.2014.
- [Sch06] Holger Schrödl. *Business Intelligence mit Microsoft SQL Server 2005*. BoD-Books on Demand, 2006.
- [SO09] Andreas Seufert und Karsten Oehler. *Grundlagen Business Intelligence*, Jgg. 1. BoD-Books on Demand, 2009.

# Scrum - Agiles Projektmanagement

Alexander Zimak

Department für Informatik  
Abteilung Informationssysteme  
alexander.zimak@uni-oldenburg.de  
Matrikelnummer: 1451857  
Betreuer: Thomas Vogelgesang

**Abstract:** Scrum ist ein Prozessmodell, welches als Werkzeug für agiles Projektmanagement eingesetzt werden kann. Mithilfe von Scrum können Projekte, Prozesse oder ganze Organisationen optimiert werden. Diese Arbeit soll dabei helfen, die Vorgehensweise innerhalb von Scrum zu verstehen und den Charakter von Scrum zu erkennen. Dazu wird, neben der Erklärung des Scrum-Prozesses, detailliert auf die einzelnen Komponenten von Scrum eingegangen. Um ein Verständnis dafür zu erlangen, wie Scrum auf verschiedene Arbeitsumgebungen angewandt werden kann, wird Scrum anhand eines Anwendungsszenarios angepasst und vorgestellt. Abschließend wird auf die Vorteile und Risiken von Scrum eingegangen. Scrum ist eine Möglichkeit, die hilft, sich selbst und die Organisation zu optimieren und kann daher für ambitionierte und innovationsgetriebene Firmen, Projektleiter und Mitarbeiter nachhaltigen Erfolg ermöglichen.

## 1 Einleitung

Viele Unternehmen sind bei der Produktentwicklung in selbst auferlegten statischen und veralteten Abläufen gefangen und versuchen Methoden genau nach Plan und Schritt für Schritt in die Realität umzusetzen. Dabei werden Aufträge soweit gegliedert und unterteilt, dass die ausführenden Mitarbeiter keinerlei Spielraum mehr bei der Umsetzung ihrer Aufgaben besitzen. Ihnen werden strikte Vorschriften gegeben, wie sie zu arbeiten haben und wie Aufgaben umgesetzt werden müssen. Das führt dazu, dass der Spaß an der Arbeit verloren geht und dass selbst kreative Arbeitsbereiche, wie die Softwareentwicklung, zu Fließbandarbeit verkommen. Um sich von diesem strikten methodischen Denken zu lösen, ist ein Übergang von einer vorausplanenden zu einer lernenden Arbeitsweise notwendig. Eine mögliche lernende Arbeitsweise stellt Scrum dar. In Scrum geht es nach Nonaka und Takeuchi um die Zusammenarbeit mehrerer Fachdisziplinen innerhalb der Produktentwicklung. Scrum stellt eine Grundüberzeugung, Philosophie und Arbeitsweise mit klar definierten Rollen und einem sehr einfachen Prozessmodell dar. Dabei hilft Scrum, den Mitarbeitern innerhalb eines Teams wieder die nötigen Freiheiten zu geben, die diese benötigen, um ihre Kreativität auszuschöpfen und Verantwortung für die geleistete Arbeit zu übernehmen. Das bedeutet einen Übergang vom prototypischen Arbeiter, der sich nur auf seine Arbeit konzentriert und nicht nach links und rechts schaut, hin zum teleologisch arbeitenden Menschen, der sich seiner Fähigkeiten und Potenziale bewusst ist und diese richtig vermarkten kann.

Scrum stellt somit keine Methode dar, sondern lediglich ein Rahmen, welcher je nach den gegebenen Umständen angepasst werden muss, um Verbesserungen herbeizuführen. Dabei ist es nicht, wie ursprünglich konzipiert, nur auf die Produktentwicklung anwendbar, sondern kann auch zum Steuern großer Projekte oder ganzer Organisationen genutzt werden. Es ist aber darauf zu achten, dass Scrum nicht zeigt wie etwas besser gemacht wird, sondern lediglich was besser gemacht werden muss bzw. kann. Der entscheidende Vorteil von Scrum gegenüber lineareren Prozessmodellen ist hierbei, dass Hindernisse und Dysfunktionen sichtbar werden und so systematisch behoben werden können. Durch den iterativen Charakter ist Scrum in der Lage, eine stetige Optimierung der Arbeitsweisen zu ermöglichen.[Glo11]

## **1.1 Aufbau der Arbeit**

In dieser Arbeit geht es darum, Scrum strukturiert und übersichtlich vorzustellen. Dafür wird im ersten Abschnitt Scrum anhand des Prozesses erläutert. Nachdem der Prozess vermittelt wurde, wird auf die einzelnen Bestandteile und Komponenten von Scrum eingegangen, bevor die Anpassbarkeit von Scrum anhand eines Anwendungsszenarios näher erklärt wird. Abschließend geht diese Arbeit noch auf die Vorteile und Risiken ein, die Scrum mit sich bringt.

## **2 Scrum**

Um Scrum verständlich zu vermitteln, reicht es nicht aus, lediglich alle Bestandteile eines Scrum-Projektes im Einzelnen vorzustellen. Aufgrund der engen Beziehungen der einzelnen Komponenten von Scrum, macht es mehr Sinn, diese innerhalb des Prozesses vorzustellen, da der Prozess die komplette Funktionalität von Scrum veranschaulicht. Im folgenden Abschnitt wird daher der Prozess vorgestellt und erläutert. Auf die genauen Bedeutungen und Spezifikationen der einzelnen Begrifflichkeiten wird im Anschluss detailliert eingegangen. In einem Scrum-Projekt gibt es vier verschiedene Begrifflichkeiten, die essenziell für das Verständnis des Scrum-Prozesses sind. Diese vier Begriffe lauten Rolle, Artefakt, Aktivität und Methode. Dabei stellen Rollen verschiedene Personen oder Personengruppen innerhalb von Scrum dar. Artefakte sind Dokumente, die innerhalb von Scrum vorkommen. Unter Aktivitäten verbergen sich alle Vorgänge, die von den Personen unter Zuhilfenahme der Artefakte innerhalb von Scrum durchgeführt werden müssen. Der Begriff Methoden beschreibt Verfahren, die bei der Durchführung verschiedener Aktivitäten oder bei der Erstellung von Artefakten angewandt werden können.

### 3 Der Prozess

Im folgenden Abschnitt wird der Scrum-Prozess vorgestellt und graphisch veranschaulicht (Abbildung 1). Um die Komplexität des Prozesses auf ein anschauliches Niveau herunter zu brechen, wurde der Prozess in drei Phasen eingeteilt: Die *Planungsphase*, die *Entwicklungsphase* und die *Nachbetrachtungsphase*. Darüber hinaus werden die auftretenden, für Scrum spezifischen Begrifflichkeiten kurz umschrieben. Die genauen Spezifikationen und Eigenschaften der einzelnen Begrifflichkeiten werden im Anschluss erläutert. Neben der Durchführung der drei Phasen sollte im gesamten Projektverlauf darauf geachtet werden, dass das Projekt eine gewisse Transparenz besitzt, da die Transparenz des Projekts eine genauere und sichere Planung ermöglicht und darüber hinaus auch das Produktbewusstsein der Stakeholder fördert.

#### 3.1 Planungsphase:

In diesem Abschnitt wird die Planungsphase behandelt. Es werden alle wichtigen Schritte, die in der Planungsphase durchgeführt werden, kurz beschrieben. Vor Beginn eines Scrum-Projektes müssen bereits einige Voraussetzungen erfüllt werden. Es gibt drei unterschiedliche Rollen innerhalb von Scrum, die im Vorfeld sinnvoll vergeben werden müssen. Diese Rollen sind die des *Product Owners*, die des *Scrum Masters* und die des *Entwicklungsteam*. Dabei ist der Product Owner derjenige, der die Verantwortung für den wirtschaftlichen Erfolg des Projektes trägt, der Scrum Master derjenige, der den Ablauf des Projekts kontrolliert und steuert und das Entwicklungsteam, das aus den Mitarbeitern besteht, die Entwicklungsarbeiten innerhalb des Projektes verrichten. Darüber hinaus gibt es Rollen außerhalb von Scrum, die zur Durchführung eines erfolgreichen Projektes beitragen können. Neben dem *Kunden* zählen auch die *Nutzer* und das *Management* dazu. Der Product Owner sollte jemand sein, der in engem Kontakt zum Kunden steht und dessen Interessen und Anforderungen optimal vertreten kann. Der Product Owner erstellt zu Beginn das so genannte *Product Backlog*. Dieses enthält alle Anforderungen und Wünsche des Kunden in abstrahierter, sortierter und priorisierter Form. Nachdem das Product Backlog vom Product Owner erstellt worden ist, beginnt das *Sprint Planning*. Im ersten Sprint Planning wird das Vorgehen für den ersten *Sprint* (Arbeitsabschnitt) festgelegt. Darüber hinaus wird die sogenannte *Definition of Done*, oder kurz DoD festgelegt. Der letzte Schritt innerhalb der Planung ist das Erstellen des *Sprint Backlog*, welcher alle Backlog-Einträge enthält, die im nächsten Sprint bearbeitet werden sollen. [Glo11][Wir11][SAP13]

#### 3.2 Entwicklungsphase:

Die Entwicklungsphase stellt die zweite Phase innerhalb des Scrumprozesses dar. In dieser geht es um die Schritte und Aktionen, die innerhalb der eigentlichen Entwicklung des Produktes ausgeführt werden müssen. Wurden im Sprint Planning alle relevanten The-

men bearbeitet, folgt der nächste Schritt, der Sprint. Der Sprint ist der Schritt in Scrum, in dem die eigentliche Produktentwicklung stattfindet. Im Sprint bearbeitet das Entwicklungsteam alle im Sprint Backlog enthaltene Einträge. Der Sprint hat eine vorher definierte Länge und kann weder verkürzt noch verlängert werden. Während des Sprints arbeitet das Entwicklungsteam selbstständig. Im *Daily Scrum* erhalten der Scrum Master und der Product Owner einen Einblick in die Fortschritte des Entwicklungsteams innerhalb des Sprints. Das Daily Scrum ist ein tägliches Meeting, welches dem Informationsaustausch dient. Mögliche Hindernisse oder Unstimmigkeiten, die während des Daily Scrum aufgedeckt wurden, werden an den Scrum Master gemeldet. Dieser beschäftigt sich anschließend mit der Lösung der Hindernisse, um einen reibungslosen Ablauf von Scrum zu gewährleisten. Das Entwicklungsteam muss während der Sprintphase neben der eigentlichen Produktentwicklung auch einige Artefakte bearbeiten. Zu diesen Aufgaben gehört unter anderem das Pflegen des *Product Increments*. Im Product Increment werden alle bearbeiteten Einträge aus dem Sprint Backlog aufgeführt, die der Definition of Done entsprechen. Neben den aufeinander folgenden Aktivitäten gibt es auch Aktivitäten die durchgehend durchgeführt werden. So zum Beispiel das Product Backlog Refinement, in welchem alle Änderungen bezüglich der Einträge im Product Backlog behandelt werden, wie beispielsweise das Entfernen, Hinzufügen oder Detaillieren von Einträgen. Ist der Sprintzeitraum beendet, folgt die Nachbetrachtung. [Glo11][Wir11][SAP13]

### 3.3 Nachbetrachtungsphase:

In der dritten Phase des Scrum-Prozesses werden die Ergebnisse des vorangegangenen Sprints ausgewertet und es wird der Ablauf des Sprints kontrolliert und bewertet. Im Anschluss an den Sprint folgt das *Sprint Review*. In dieser Aktivität wird das Product Increment überprüft, welches während des Sprints vom Entwicklungsteam gefertigt worden ist. Anhand des Increments wird ersichtlich, ob ein Sprint erfolgreich war oder ob es zu Problemen bei der Bearbeitung der Einträge kam. Unter Verwendung dieser Erkenntnisse kann der Product Owner den Sprint als vollständig abgeschlossen deklarieren und das Product Backlog kann entsprechend angepasst werden. Des Weiteren präsentiert das Entwicklungsteam die Ergebnisse des Sprints. Bei der Präsentation können neben dem Product Owner und dem Scrum Master, auch die Stakeholder (Kunde, Nutzer etc.) anwesend sein. Das Ziel des Sprint Review ist es, einen Dialog aller Beteiligten zu fördern, um so eine mögliche Ideenfindung für zukünftige Sprints zu begünstigen und konstruktives Feedback bezüglich der bisherigen Entwicklungen zu sammeln. Nachdem alle Beteiligten über den Fortschritt des Projektes informiert worden sind, folgt nun die *Sprint Retrospektive*. Sie bildet den letzten Schritt innerhalb des Sprintprozesses ab. Die Sprint Retrospektive wird meist in einem geschützten Raum abgehalten, um so das Wohlbefinden der Beteiligten zu stärken. Es wird die Arbeitsweise innerhalb der vorangegangenen Planungs- und Entwicklungsphase untersucht und angepasst, um zukünftig effizienter und effektiver arbeiten zu können.

Wurden alle drei Prozessphasen durchlaufen, werden diese aufgrund der iterativen Struktur von Scrum mit den angepassten Artefakten und optimierten Arbeitsweisen erneut durchgeführt. Dieser Prozess wiederholt sich bis das Projekt in seiner Gänze abgeschlossen ist. Die

iterative Struktur von Scrum wird in der Abbildung 1 anschaulich dargestellt. [Glo11][Wir11][SAP13]

## **4 Rollen**

In diesem Abschnitt werden die drei Hauptrollen, die in Scrum eingenommen werden können, mit den zugeordneten Eigenschaften und Aufgaben vorgestellt. Bestimmte Aktivitäten, Artefakte und Methoden, die in diesem Zusammenhang vorgestellt werden, können in den dazugehörigen Abschnitten nachgelesen werden.

### **4.1 Product Owner:**

Der Product Owner ist die Person, die für den wirtschaftlichen Erfolg des entwickelten Produktes verantwortlich ist. Der Product Owner ist im Normalfall eine Einzelperson, die als direktes Kommunikationsglied zwischen dem Kunden und dem Scrum-Team steht und während des Projektes nicht ausgetauscht werden sollte. Um den Erfolg des Projektes zu gewährleisten, ist der Product Owner für die Konzeption und Mitteilung einer klaren Produktvision zuständig. Aus dieser ermittelt und extrahiert der Product Owner einzelne Produkteigenschaften und Aufgaben die zur Erstellung des Produktes erfüllt werden müssen. Diese Aufgaben und Eigenschaften notiert der Product Owner im Product Backlog. Der Product Owner besitzt darüber hinaus die alleinige Entscheidungsgewalt über das Produkt, die Eigenschaften und die Reihenfolge der Implementierung. Während der Durchführung des Projektes mit Scrum ist der Product Owner für die Kontrolle und Aktualisierung des Product Backlogs zuständig. Darüber hinaus ist er für die Erstellung von Anforderungen für das Sprint Backlog und die Abnahme der Ergebnisse am Ende eines jeden Sprints zuständig.[Glo11][SAP13]

### **4.2 Scrum Master:**

Der Scrum Master kann als Projektleiter innerhalb des Scrum-Projekts verstanden werden. Er ist für den reibungslosen Ablauf des Projekts verantwortlich. Genau wie der Product Owner stellt auch der Scrum Master eine Einzelperson dar, die jedoch während des Projektes wechseln kann. Der Scrum Master stellt eine Führungsperson für das Entwicklungsteam dar und arbeitet eng mit diesem zusammen. Er geht jedoch meist nicht zum Entwicklungsteam. Als Führungsperson müssen einige Aufgaben erfüllt werden, um den reibungslosen Ablauf der Entwicklungen zu gewährleisten. Zu diesen zählen die Koordination der Abläufe, das Einführen von Scrum-Regeln sowie die Kontrolle, dass diese eingehalten werden, die Leitung der Meetings und das Beheben von Strungen und Hindernissen innerhalb des Projekts. Auch wenn der Scrum Master eine Führungsperson für das Entwicklungsteam darstellt, gibt er einzelnen Teammitgliedern weder Anweisungen, Beurteilungen noch verhängt er Disziplinarstrafen. Er behandelt das Entwicklungsteam als eine Einheit. Zu Beginn

der Entwicklung des Produktes stellt die Stelle des Scrum Masters eine Vollzeitstelle dar, da das Umstellen der gewohnten Ablufe, das Zusammenwachsen des Teams und das Einlernen der Rollen sehr zeitaufwendig ist. Ist Scrum etabliert, beschftigt sich der Scrum Master mit den Aufgaben eines Change Managers. Das bedeutet, er beschftigt sich neben der Untersttzung des Scrum-Teams mit der Verbreitung von Scrum im Unternehmen und der Sensibilisierung der Mitarbeiter.[Glo11][SAP13]

### **4.3 Entwicklungsteam:**

Die dritte Rolle innerhalb von Scrum stellt das Entwicklungsteam dar. Wie der Name vermuten lsst, beschftigt sich das Entwicklungsteam mit der eigentlichen Entwicklung des Produktes. Das Entwicklungsteam ist ein selbstorganisiertes Team, welches nur bedingt vom Product Owner oder Scrum Master in der Arbeitsweise angeleitet werden sollte. Um einen reibungslosen Ablauf innerhalb der Entwicklungen zu gewhrleisten, ist eine interdisziplinre Besetzung des Scrum-Teams essenziell. Das bedeutet, dass alle fr das Projekt benötigten Rollen von den Mitgliedern abgedeckt werden mssen. Zu diesen zhlen zum Beispiel Entwickler oder Dokumentationsexperten. Von Vorteil wre es darber hinaus, wenn alle Teammitglieder mehrere Rollen einnehmen knnen. So knnen unvorhergesehenen Ausfalle (z.B. durch Krankheit) kompensiert werden und die Entwicklungen reibungslos fortgefhhrt werden. Ein weiterer Faktor auf den bei der Zusammenstellung des Entwicklungsteams geachtet werden sollte, ist die Teamgre. Jedes Entwicklungsteam sollte alle benötigten Rollen abdecken knnen, jedoch dabei einen berschaubaren Koordinationsaufwand beibehalten. Daher wird eine Teamgre von 3 - 9 Personen als optimal angesehen. Dabei sollte das Team sowohl nach innen und nach auen immer als Einheit auftreten. Es gibt also keine hierarchische Struktur innerhalb des Teams. Wurde ein geeignetes Team zusammengestellt, bekommt dieses bestimmte Aufgaben, fr die es whrend der Entwicklung verantwortlich ist. Zu diesen zhlt, neben der Schtzung des Umfangs der einzelnen Eintrge im Product Backlog, das Herunterbrechen dieser in sogenannte *Tasks*. Die Hauptaufgabe besteht in der Erfllung der Tasks innerhalb des vorgegebenen Sprintzeitraums. [Glo11][SAP13]

### **4.4 Rollen auerhalb von Scrum:**

Neben den Rollen innerhalb von Scrum gibt es auch Rollen auerhalb der Scrum-Umgebung, die untersttzend bei der Entwicklung eines Produkts wirken knnen. Zu diesen gehren der Kunde, die Nutzer und das Management. Der Kunde ist der Auftraggeber und steht whrend des Entwicklungsprozesses in Kontakt mit dem Product Owner, um einen Einblick in den Fortschritt des Projekts zu bekommen und die von ihm gestellten Anforderungen zu aktualisieren. Auch die Nutzer sollten in die Entwicklung miteinbezogen werden, da sie Feedback ber die Benutzerfreundlichkeit oder Sinnhaftigkeit von bestimmten Produkteigenschaften liefern knnen. Das Management sorgt dafr, dass die ntigen Rahmenbedingungen fr die Durchfhhung von Scrum vorhanden sind. Die Aufgabe des Managements ist dafr zu sorgen, dass gengend materielle Ressourcen zugnglich gemacht werden und dass Rum-

lichkeiten und Arbeitsmittel gestellt werden.[Glo11][SAP13]

## **5 Artefakte**

Artefakte stellen Dokumente dar, die innerhalb des Serumprozesses benötigt werden, um gewisse Arbeitsschritte ordnungsgemäß auszuführen. Im Folgenden werden die einzelnen Artefakte in ihrer Gesamtheit vorgestellt und erläutert.

### **5.1 Product Backlog:**

Das Product Backlog stellt den individuellen Anforderungskatalog des Product Owners dar. Dieser enthält alle Anforderungen, die der Kunde an das geordnete Produkt stellt. Die gestellten Anforderungen weisen häufig eine hohe Komplexität auf und werden daher vom Product Owner abstrahiert und auf einzelne Aufgaben heruntergebrochen. Das Product Backlog enthält keine expliziten Anforderungen, sondern lediglich abstrahierte Anforderungen, die einen gewissen Interpretationsspielraum zulassen. So entstehen einzelne Einträge, die vom Entwicklungsteam umgesetzt werden können. Wurde ein erstes Product Backlog erstellt, wird dieses während des Projektes immer wieder angepasst und aktualisiert, um so auf Änderungen und Abweichungen der Produkthanforderungen zu reagieren und diese miteinzubeziehen.[Glo11][SAP13]

### **5.2 Product Increment:**

Das Product Increment, enthält alle im Sprint vollständig bearbeiteten Tasks. Alle im Increment enthaltene Tasks und Einträge, müssen der festgelegten Definition of Done entsprechen. Anhand des Product Increments lässt sich feststellen, ob am Ende des Sprints alle zu erledigenden Aufgaben erledigt worden sind.[Glo11][SAP13]

### **5.3 Sprint Backlog:**

Der Sprint Backlog besteht aus den Product Backlog-Einträgen, die für den Sprint ausgewählt wurden. Er beinhaltet neben dem Product Increment eine Prognose über die möglichen Funktionalitäten des nächsten Increments und den dafür benötigten Arbeitsaufwand. Das bedeutet, er enthält alle zu erledigenden Aufgaben innerhalb des Sprints sowie die Reihenfolge der Umsetzung. Eine häufig verwendete Visualisierungsform des Sprint Backlogs ist das Taskboard, das im Abschnitt über Methoden genauer beschrieben wird. [Glo11][SAP13]



#### **5.4 Definition of Done:**

Die Definition of Done, beschreibt das Verständnis des Scrum Teams unter welchen Bedingungen eine Aufgabe als erledigt gilt. Die Definition of Done wird im ersten Sprint Planning festgelegt und während des Entwicklungsprozesses aktualisiert und angepasst. Die DoD enthält Qualitätskriterien, Einschränkungen und allgemeine, nicht funktionale Anforderungen, die erfüllt werden müssen um eine Aufgabe als abgeschlossen zu markieren. Das kann bei der Entscheidung, wie viele Tasks innerhalb eines Sprints erledigt werden sollen, behilflich sein. [Glo11]

#### **5.5 Task:**

Ein Task ist ein durch das Entwicklungsteam definierter Arbeitsschritt. Dieser hat eine maximale Bearbeitungsdauer von einem Tag und dient der Aufwandsschätzung innerhalb des Sprint Plannings.[Glo11]

### **6 Aktivitten**

Unter Aktivitten werden alle Arbeitsabschnitte aufgeführt und detailliert erklärt, welche während des Scrum-Prozesses auftreten und bearbeitet werden müssen.

#### **6.1 Sprint:**

Der Sprint repräsentiert einen Arbeitsabschnitt innerhalb des Entwicklungsprozesses von Scrum. Sprints sind immer gleichlange Arbeitsabschnitte, in denen sich das Entwicklungsteam mit der Abarbeitung der geforderten Anforderungen befasst. Das heißt, es soll immer eine vorher definierte Anzahl an zu erfüllenden Aufgaben erledigt werden. Es sollte kein Einfluss auf das Entwicklungsteam während des Sprints genommen werden, da das Entwicklungsteam im Sprint selbstorganisiert die geforderten Anforderungen abarbeitet. Abgegrenzt wird der Sprint vom Sprint Planning am Anfang und vom Sprint Review und der Sprint Retrospektive am Ende des Sprints. Die Dauer eines Sprints sollte 1 - 4 Wochen nicht überschreiten, da sonst aufkommende Probleme innerhalb der Entwicklung zu lange aufgeschoben werden könnten. Sollte es zu erheblichen Komplikationen innerhalb des Sprints kommen, kann dieser auch abgebrochen und mit einer Sprint Retrospektive beendet werden. [Glo11][SAP13]

## **6.2 Sprint Planning:**

Im Sprint Planning wird festgelegt, was im Sprint entwickelt werden kann und wie bestimmte Aufgaben im Sprint zu bearbeiten sind. Um festzulegen, was im Sprint entwickelt werden kann, stellt der Product Owner die zu erfüllenden Anforderungen aus dem Product Backlog in priorisierter Reihenfolge vor. Um ein gemeinsames Verständnis der zu erfüllenden Aufgaben zu erlangen, werden vom Product Owner und dem Entwicklungsteam verschiedene Einsatz- und Akzeptanzkriterien vereinbart. Diese helfen bei der Entscheidung, wann etwas fertig gestellt werden soll und helfen dem Entwicklungsteam einzuschätzen, wie viele Backlog-Einträge sie im Sprint bearbeiten können. Das bedeutet, dass die Entscheidung über die Anzahl der zu erfüllenden Backlog-Einträge beim Entwicklungsteam und die Reihenfolge der Abarbeitung beim Product Owner liegt. Gemeinsam wird aus diesen Entscheidungen das Sprintziel ermittelt. Im Gegensatz dazu, was entwickelt wird, liegt die Entscheidung darüber, wie entwickelt wird, einzig beim Entwicklungsteam. Das Entwicklungsteam einigt sich darüber, wie der Ablauf der Bearbeitung aussieht und welche Tasks erfüllt werden müssen, um das Sprintziel zu erreichen. Hierfür können auch kleine Gruppen innerhalb des Entwicklungsteams gebildet werden, um spezielle Fragestellungen, wie die der Architektur oder der Dokumentation, zu klären. Das Ergebnis, das sich am Ende des Sprint Plannings ergibt, ist der sogenannte Sprint Backlog.[Glo11][SAP13]

## **6.3 Product Backlog Refinement:**

Das Product Refinement ist ein fortlaufender Prozess, welcher vom Product Owner und dem Entwicklungsteam entwickelt wird. Innerhalb des Product Backlog Refinements werden Details zu Einträgen hinzugefügt, Einträge gelöscht, geschätzt oder geordnet. Dabei können Informationen von Stakeholdern hilfreich für die Planung und Schätzung von Einträgen sein. Im Umfang sollte das Product Backlog Refinement nicht mehr als zehn Prozent der Arbeitszeit ausmachen.[Glo11][SAP13]

## **6.4 Sprint Review:**

Das Sprint Review erfolgt immer am Ende eines jeden Sprints. Das Sprint Review dient der Überprüfung des Product Increments, um notwendige Änderungen am Product Backlog vorzunehmen. Zum Beispiel muss, wenn ein Eintrag aus dem Sprint Backlog nicht erfüllt worden ist, dieser wieder in das Product Backlog eingefügt und neu priorisiert werden. Sind alle Einträge aus dem Sprint Backlog abgearbeitet, erfolgt die Abnahme durch den Product Owner. Neben der Überprüfung der geleisteten Arbeit, dient das Sprint Review auch der Präsentation der Ergebnisse. Hierfür nehmen neben dem Scrum Team auch die Stakeholder an dem Review teil. Die Präsentation der bisherigen Ergebnisse fördert den Dialog und die Ideenfindung. Stakeholder, die sonst wenig Einblick in die Entwicklung innerhalb des Scrum-Teams erhalten, können ihre Meinung in Form von konstruktivem Feedback kundtun und so die

zukünftige Entwicklung unterstützen. Zum Beispiel könnte ein Anwender bemerken, dass etwas richtig implementiert worden ist, jedoch die Funktionalität nie genutzt werden wird. Mithilfe dieser Informationen kann der Product Owner nun das Product Backlog anpassen. Die Dauer eines solchen Reviews sollte nicht mehr als eine Stunde pro Sprintwoche betragen. [Glo11][SAP13]

## **6.5 Sprint Retrospektive:**

Im Gegensatz zum Sprint Review wird die Sprint Retrospektive unter Ausschluss der Stakeholder abgehalten. Sie dient dem Scrum-Team zur Überprüfung der Arbeitsweisen innerhalb des Entwicklungsprozesses. Um eine gewohnte und offene Atmosphäre zu schaffen, wird die Retrospektive in einem geschützten und vertrautem Umfeld abgehalten. Sie dient allein der Förderung der Arbeitsweisen und der Kommunikation innerhalb des Scrum-Teams. Der Scrum Master nimmt dabei eine unterstützende Rolle ein und hilft bei der Suche nach Praktiken und Techniken zur Verbesserung der Arbeitsweisen. Im Allgemeinen hat sich die 5 Phasen Struktur bei der Durchführung der Retrospektive bewährt. Die 5 Phasen sind folgende:

### **1. offene Atmosphäre schaffen:**

- Schaffen einer offenen und ehrlichen Atmosphäre, um den Mitarbeitern Sicherheit zu vermitteln.

### **2. Informationen sammeln:**

- Sammeln von Informationen, um einen Ist-Zustand zu ermitteln und so mögliche Abweichungen vom Soll-Zustand zu erkennen.

### **3. Erkenntnisse entwickeln:**

- Warum treten Probleme auf und wie können diese behoben werden, bzw. was läuft ordnungsgemäß.

### **4. Vereinbaren von wichtigen Schritten:**

- Aus den Erkenntnissen werden konkrete, sinnvolle und realistische Schritte ermittelt, um die zukünftige Arbeitsweise zu optimieren.

### **5. Abschluss:**

- Zusammenfassen der Erkenntnisse und Dokumentieren der vereinbarten Schritte

Für die Dokumentation der Schritte gibt es unterschiedliche Methoden. Eine Methode wäre es, die ermittelten Hindernisse und Verbesserungsmaßnahmen in das Product Backlog aufzunehmen und diese zu einem geeigneten Zeitpunkt innerhalb eines Sprints zu bearbeiten.[Foe14]

## 6.6 Daily Scrum:

Das Daily Scrum ist, wie der Name vermuten lässt, ein tägliches Meeting innerhalb des Sprints. In einem Daily Scrum sind alle Personen innerhalb des Scrum-Teams anwesend, jedoch sind Product Owner und Scrum Master inaktiv und lediglich Zuhörer. Die Mitglieder des Entwicklungsteams beantworten die Fragen, *was habe ich gemacht? und was mache ich als nächstes? und sind Hindernisse aufgetreten?* Sollte sich zum Beispiel herausstellen, dass ein Task länger als einen Tag dauern wird, muss dieser aufgesplittet werden, um so die zulässige maximale Dauer nicht zu überschreiten. Das Daily Scrum dient allein dem Informationsaustausch. Für Diskussionen und angeregte Gespräche ist hier kein Platz. Daher beträgt die Dauer eines Daily Scrum lediglich 15 Minuten. Sollten Probleme und Hindernisse auftreten, die nicht gelöst werden können, werden diese nicht innerhalb des Daily Scrum besprochen, sondern an den Scrum Master weitergegeben.[Glo11][SAP13]

## 7 Methoden

### 7.1 Stories schreiben:

Stories dienen der Erhebung von Anforderungen für das Product Backlog. Sie stellen sogenannte Nutzer- oder Anwendungsszenarien dar. Zum Beispiel benötigt der Kunde eines Webshops einen Login, an dem er sich mit seinen Nutzerdaten anmelden kann. Aus solchen Szenarien können Anforderungen erhoben werden, ein Login muss implementiert werden oder ein System zur Speicherung von Nutzerdaten muss entwickelt werden. Diese Szenarien folgen immer einem allgemeinen Muster: Als *Nutzer*, will ich *Ziel/Wunsch*, damit *Nutzen*. Das heißt: Wer will was womit tun? Stories müssen immer den Akzeptanzkriterien entsprechen, die im Product Refinement Backlog festgehalten worden sind. Darüber hinaus müssen die Stories immer an den Umfang der Sprints angepasst werden. Wurden Anforderungen aus den Stories erhoben, können diese vom Product Owner in den Product Backlog übernommen werden. Anschließend werden diese nach Nutzen, Risiko und Kosten sortiert. Stories, die Fragen nach dem *„wie“* beinhalten, gehören nicht in den Product Backlog, sondern werden im Sprint Planning behandelt.[Pic09]

### 7.2 Burndown Chart:

Ein Burndown Chart (Abbildung 2) ist ein Graph, der die geleistete sowie die noch zu verrichtenden Aufgaben visualisiert darstellt. Innerhalb von Scrum gibt es zwei gebräuchliche Varianten: Zum einen den Sprint Burndown und zum anderen den Release Burndown. Der Aufbau bei beiden ist identisch, beide bestehen aus einem Koordinatensystem mit X- und Y-Achse. Lediglich bei der Skalierung der Achsen und dem abgebildeten Inhalt unterscheiden sich die beiden Burndown-Varianten. Beim Sprint Burndown (Abbildung 2) bildet die X-Achse den Zeitverlauf und die Y-Achse die Anzahl der zu erledigenden Tasks

in Stunden oder Anzahl ab. Durch die geplanten Dauern der zu erledigenden Tasks kann so eine Ideallinie (blaue Linie) berechnet werden. Diese fällt konstant ab und schneidet am Sprintende die X-Achse. Während des Sprints wird die reale Linie (rote Linie) ermittelt, die im Idealfall am Ende des Sprints auch die X-Achse schneidet. So können Abweichungen innerhalb der Planung erkannt werden. Beim Release Burndown wird auf der X-Achse der Zeitverlauf in Sprints abgebildet und auf der Y-Achse die Anzahl des Umfangs, also aller noch zu erledigenden Aufgaben, aufgeteilt in gleich große Teile. Das hilft bei der Schätzung des Gesamtumfangs und der Festlegung des Liefertermins. [Glo11]

### **7.3 Taskboard:**

Das Taskboard (Abbildung 3) ist eine Technik, die zur Visualisierung des Sprint Backlogs genutzt werden kann. Auf dem Taskboard lassen sich alle Einträge aus dem Product Backlog erkennen, die für einen Sprint ausgewählt worden sind. Darüber hinaus sind auch alle dazugehörigen Aufgaben und Bearbeitungszustände aufgeführt. Das Taskboard besteht, wie in der Abbildung 3 zu sehen, aus einer Tabelle mit 4 Spalten. Die erste Spalte (Story) enthält die Einträge aus dem Sprint Backlog, die anderen drei Spalten enthalten Aufgaben oder Tasks im jeweiligen Bearbeitungszustand. Die zweite Spalte (To Do) enthält dabei alle noch zu erledigenden Aufgaben oder Tasks, Spalte drei (In Process) alle in Bearbeitung und Spalte vier (To Verify) alle fertigen Tasks und Aufgaben. Die Tasks und Aufgaben in Spalte vier müssen anschließend nur noch vom Product Owner abgenommen werden. Das Taskboard wird häufig im Daily Scrum verwendet, um den Fortschritt innerhalb eines Sprints darzustellen. Sollte ein Task die maximale Länge von einem Tag überschreiten oder auf eine andere Art und Weise hinderlich sein, wird dieser im Taskboard markiert (roter Punkt). Zur Behebung der Fehlerquelle sollten Gegenmaßnahmen in Form neuer Tasks in das Taskboard aufgenommen werden. [Glo11]

## **8 Anwendungsszenario studentische Projektgruppe**

Da Scrum lediglich als eine Art Schablone dient, die dabei helfen kann, eine verbesserte und übersichtlichere Projektdurchführung zu erreichen, muss diese an die gegebenen Umstände des jeweiligen Arbeitsumfeldes angepasst werden. Nicht jedes Arbeitsumfeld ist gleich und weist die gleichen Ressourcen, Arbeitsmittel und kulturellen Hintergründe auf. In diesem Abschnitt wird ein möglicher Lösungsansatz vorgestellt. Dieser sieht vor, Scrum innerhalb einer Gruppe von Studenten als Projektmanagement-Werkzeug anzuwenden. Die Projektgruppe (PG) Business Intelligence as a Service (BIAaaS) hat die Aufgabe bekommen, sich im Laufe einer einjährigen Projektlaufzeit damit zu beschäftigen, ein analytisches Informationssystem in der Cloud zu implementieren. Das entstandene Produkt sollte im Stande sein, die Anforderungen des Kunden, in diesem Fall ein Kunde aus der Energiebranche, zu erfüllen.

## 8.1 Plaungsphase:

Um Scrum innerhalb der PG umzusetzen, bedarf es einiger Vorberlegungen, genau wie es im allgemeinen Scrum-Umfeld der Fall ist. Zuerst mssen die bentigten Rollen identifiziert und verteilt werden. Im PG-Umfeld wrde sich der Betreuer als Product Owner eignen, da dieser das grte Know-how auf dem Gebiet der Projektentwicklung besitzt und bereits in Kontakt mit dem Kunden steht. Des Weiteren knnte dieser innerhalb des Sprint Review entscheiden, ob die erbrachten Leistungen den gestellten Anforderungen entsprechen und so die Lieferung eines Produktes sichern. Neben dem Product Owner mssen noch die Rollen des Scrum Masters und die des Entwicklungsteams vergeben werden. Diese knnen bzw. sollten von den Studenten besetzt werden. Die Position des Scrum Masters sollte von jemandem bekleidet werden, der bereits ein gewisses Know-how im Umgang mit Scrum besitzt und alle Komponenten kennt, um so den reibungslosen Ablauf des Scrum-Prozesses zu sichern. Dieser sollte aber ber die Aufgaben des Scrum Masters hinweg auch in das Entwicklungsteam eingebunden werden, da die Aufgaben des Scrum Masters in diesem Anwendungsfall keine Vollzeitstelle erfordern. Im Gegensatz zu einem Scrum Master in einem greren Unternehmen ist es in der PG nicht die Aufgabe des Scrum Master, Scrum zu verbreiten und andere Mitarbeiter fr Scrum zu sensibilisieren. Es gibt auch keine eingefahrenen Arbeitsablupe, die umgestellt werden mssten. Wurde der Scrum Master ausgewählt, bekleiden alle brigen Mitglieder der Projektgruppe das Entwicklungsteam. Welche Aufgabenbereiche die jeweilige Mitarbeiter innerhalb des Entwicklungsteam bearbeiten, entscheidet sich nach der Erhebung der Anforderungen im Sprint Planning. Rollen, die mit groer Wahrscheinlichkeit vergeben werden mssen, sind die des Entwicklers, des Dokumentationsbeauftragten und die des Software-Architekten. Ein sinnvolles Szenario wre es, wenn innerhalb der Projektgruppe jedes Mitglied alle Rollen abdecken knnte, um so die Ausfle zu kompensieren und wenn ntig, in kleinen Teams bestimmte Tasks zu bearbeiten. Dabei ist hervorzuheben, dass die wichtige Interdisziplinaritt im PG-Umfeld nicht gegeben ist, da alle Studenten einen hnlichen Wissensstand besitzen und nur geringfgige Spezialisierungen besitzen. Wenn aber alle PG-Mitglieder smtliche Rollen innehaben knnen, knnen Vorteile wie zum Beispiel die des Pair Programming genutzt werden, um die Entwicklung des Products voran zu treiben und Fehler zu vermeiden.

Nachdem die grundlegenden Rollen verteilt worden sind, kann mit der eigentlichen Planungsphase des Projektes begonnen werden. Der Verlauf des Sprintdurchlaufs wird untersttzend in (Abbildung 4) visualisiert. Im ersten Schritt sollte der Product Backlog, im Gegensatz zum allgemeinen Vorgehen, von den Studenten (Entwicklungsteam) in Eigenarbeit entwickelt werden. Der Product Owner sollte hierbei untersttzend zur Seite stehen und die Studenten auf fehlende oder ungenaue Anforderungen aufmerksam machen. Um die ntigen Anforderungen zu ermitteln, wrde sich die Methode der Stories eignen. So knnen relativ einfache Anwendungsfile in speziell fr das Produkt notwendigen Anforderungen aufgespalten werden. Die Einbeziehung der Studenten macht hier insofern Sinn, da diese durch die Erstellung der Anforderungen bereits einen Einblick in die fachlichen Anforderungen des Produkts bekommen. Auerdem hilft es ihnen bei der genaueren Schtzung der zu bewltigenden Tasks. Die Priorisierung der Eintrge im Backlog, kann sowohl vom gesamten Scrum-Team als auch vom Product Owner alleine vorgenommen werden. Um den Lerneffekt der PG zu steigern, wre die Schtzung und Priorisierung der Eintrge durch

die Studenten sinnvoll.

Nachdem das anfängliche Product Backlog erstellt worden ist, geht das Scrum-Team zum ersten Sprint Planning über. Auch beim Sprint Planning sollten alle Mitglieder des Scrum-Teams teilnehmen. Im Sprint Planning sollten sowohl die Länge der Sprints als auch die Definition of Done, die Art und Weise des Arbeitens und der Inhalt des Sprints geklärt werden. Für die Länge der Sprints sollte ein angemessener Zeitraum gewählt werden, der ausreicht, um einen Fortschritt in der Entwicklung des Produkts erkennbar zu machen. Um das Projekt am Laufen zu halten und einen kontinuierlichen Arbeitsvorgang zu gewährleisten, würde sich ein Sprintzeitraum von einer Woche anbieten. Die Studenten sollten mit dem Product Owner den Umfang der im Sprint zu bearbeitenden Backlog-Einträge festlegen. Hat das Entwicklungsteam die Anzahl der zu erledigenden Einträge aus dem Product Backlog festgelegt, werden diese in das Sprint Backlog übernommen. Neben der Anzahl der zu bearbeitenden Einträge, sollte sich das Entwicklungsteam auch über die Art und Weise der Bearbeitung Gedanken machen und bestimmte Richtlinien festlegen. Hier sollten Dinge geklärt werden, wie die Rollenverteilung, die Möglichkeit in kleinen Teams zu arbeiten oder die Art und Weise der Kommunikation. Wurden diese Punkte festgehalten, kann der erste Sprint beginnen.

## **8.2 Entwicklungsphase:**

Innerhalb des Sprints bearbeiten die Studenten die Aufträge aus dem Sprint Backlog in der vorher definierten Reihenfolge. Das Daily Scrum, welches innerhalb des Sprints täglich abgehalten werden sollte, wird aufgrund der dezentralen Bearbeitung der Tasks in ein Weekly Scrum geändert und innerhalb eines wöchentlichen Treffens abgehalten. Lediglich das Product Increment und das Sprint Backlog werden während der Bearbeitungszeit gepflegt, um so einen Überblick über den Projektfortschritt zu erhalten und zu gewährleisten, dass nicht zwei Mitglieder unabgesprochen am selben Task arbeiten. Sollten Schwierigkeiten oder Hindernisse während der Bearbeitung der Tasks auftreten, sollten sich die Mitglieder des Entwicklungsteams direkt an den Scrum Master wenden. Der Scrum Master kann dann die auftretenden Hindernisse beseitigen. Das Sprint Backlog kann in Form eines digitalen Taskboards visualisiert werden, um es von überall zugänglich zu machen und um nicht nur erledigte Einträge aus dem Product Increment anzuzeigen, sondern auch solche, die noch in Bearbeitung sind.

## **8.3 Nachbetrachtungsphase:**

Ist die festgelegte Zeit für den Sprint abgelaufen, findet sich das komplette Scrum-Team in einem kombinierten Treffen aus Sprint Review und Sprint Retrospektive, wie in (Abbildung 4) zu erkennen, zusammen. Begonnen wird hier wie im Allgemeinen mit dem Sprint Review. Hierbei sollten je nach Fortschritt die Kunden im Rahmen einer kurzen Präsentation der Ergebnisse eingeladen werden. Nach der Präsentation sollte dem Kunden und dem Product Owner die Möglichkeit gegeben werden, ihre Meinung zu äußern, um so

Stärken und Schwächen des Produkts aufzudecken. Aufgrund der geringen Bearbeitungszeit ist es jedoch nicht notwendig, dass der Kunde an jedem Review teilnimmt. Darüber hinaus entscheidet der Product Owner, ob alle zu erfüllenden Anforderungen erfolgreich abgeschlossen wurden und das Product Increment somit vollständig ist.

Nach dem Sprint Review folgt dann die Sprint Retrospektive. Diese sollte durchgeführt werden, um die zukünftige Arbeitsweise effektiver und effizienter zu gestalten. Hierbei sollte der Scrum Master im Dialog mit dem Entwicklungsteam unterschiedliche Techniken und Praktiken diskutieren, die die Zusammenarbeit fördern können. Darüber hinaus sollten Burndown Charts erstellt werden, um das Projekt transparent zu gestalten und die zukünftige Schätzung zu unterstützen. An dieser Stelle wird der erste Sprintdurchlauf abgeschlossen und der Kreislauf wieder von vorne beginnen.

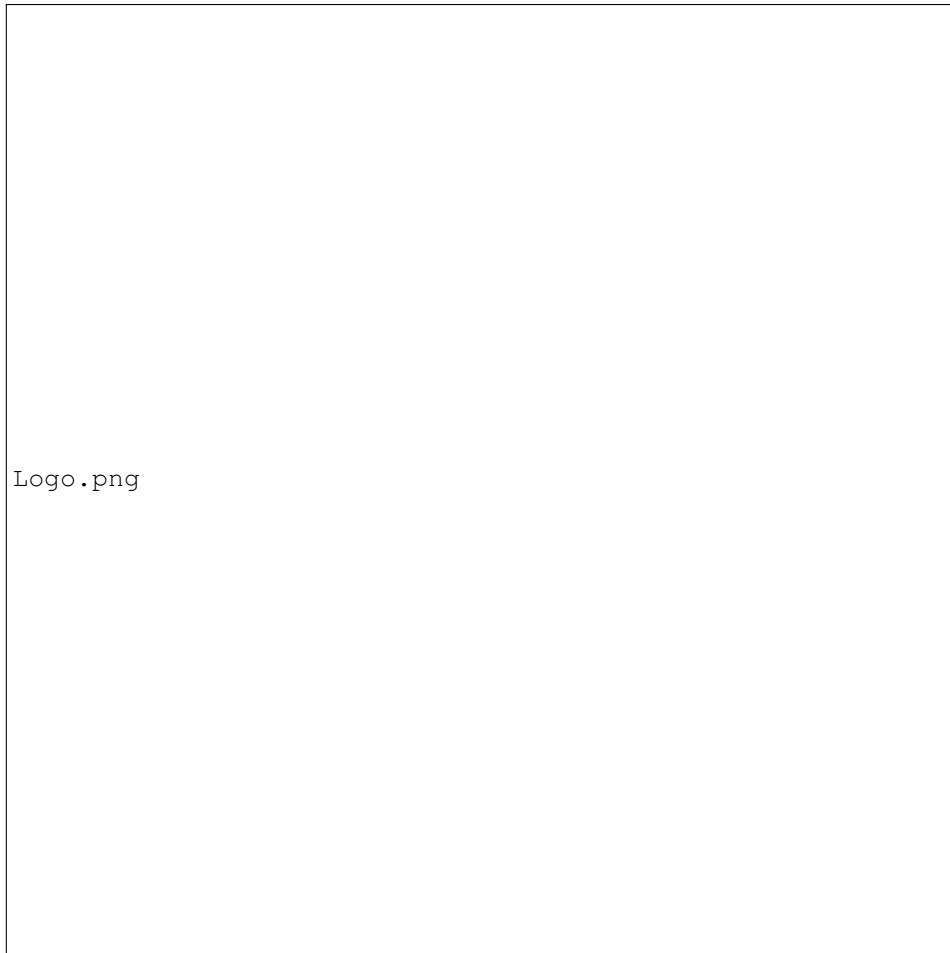
## 9 Schlussbetrachtung

Abschließend ist zu sagen, dass Scrum eine Möglichkeit für ambitionierte Projektleiter und Mitarbeiter ist, um Prozesse, Projekte und gar die ganze Organisation neu zu gestalten und aus den alten statischen Abläufen heraus zu führen. Scrum besitzt großes Potenzial und kann bei Mitarbeitern eine versteckte Innovationskraft wecken, welche produktiv für die Realisierung der gesteckten Ziele und Aufgaben genutzt werden kann. Es geht darum, den Spaß an der Arbeit wieder zu entdecken, sich selbst und die eigene Leistung zu hinterfragen und sich stetig weiter zu entwickeln. Das kann zu einem viel harmonischeren Arbeitsumfeld führen, bringt dabei gleichzeitig aber auch Vorteile bei der Produktivität und der Qualität der Arbeit. Da Führungskräfte ständig auf der Suche nach Verbesserung und Optimierung sind, bietet Scrum eine Philosophie, die genau diese Anforderungen erfüllt. Wird Scrum richtig genutzt, wird der Optimierungsprozess nie beendet, denn nach jedem Durchlauf stellt sich erneut die Frage, was besser gemacht werden kann und wie es besser gemacht werden kann. Damit unterstützt Scrum, mithilfe eines einfachen Prozessmodells und klar definierten Rollen, den existierenden Wunsch nach Perfektion. Neben den großen Potenzialen von Scrum darf aber nicht vergessen werden, dass Scrum auch Risiken birgt. Nicht jeder ist in der Lage, oder bereit, Scrum effektiv und effizient umzusetzen. Kopiert man Scrum von anderen Firmen, wird man scheitern. Es geht viel mehr darum seine eigene Arbeitsweise zu finden und zu entwickeln, denn jedes Unternehmen und jede Arbeitsgemeinschaft ist anders und weist andere Eigenschaften auf. Scrum darf nicht als festes Vorgehensmodell angesehen werden. Scrum gibt lediglich einen empfohlenen Rahmen vor, der genutzt werden kann, um die Arbeitsabläufe zu optimieren und die Mitarbeiter zu motivieren. Damit das funktioniert, muss es treibende Kräfte geben, die den Willen haben, Innovationen herbei zu führen und vor den anfallenden Aufgaben nicht zurückschrecken. Es muss Mitarbeiter geben, die sich mit Scrum auseinandersetzen und es zu ihrer Philosophie machen, um es optimal an die gegebenen Umstände anzupassen. Es ist nicht einfach, Scrum richtig einzuführen, da die Einführung mit viel Aufwand verbunden ist. Ist Scrum jedoch einmal etabliert, kann es die weitere Entwicklung stark positiv beeinflussen. Dabei sollte man sich immer vor Augen führen "Doing is a way of thinking". [Glo11]



## **Literatur**

- [Foe14] Malte Foegen. *Der Ultimative Scrum Guide 2.0*. wibas GmbH, 2014.
- [Glo11] Boris Gloger. *Scrum*, Jgg. 3. Hanser, 2011.
- [Pic09] Roman Pichler. *Scrum – Agiles Projektmanagement erfolgreich einsetzen*. d.punkt Verlag, 2009.
- [SAP13] SAP. *Agile Software Engineering*. SAP, 2013.
- [Wir11] Ralf Wirdemann. *Scrum Mit User Stories*, Jgg. 2. Hanser, 2011.



Logo.png

Abbildung 1: Scrum-Prozess in Anlehnung an it-agile.de

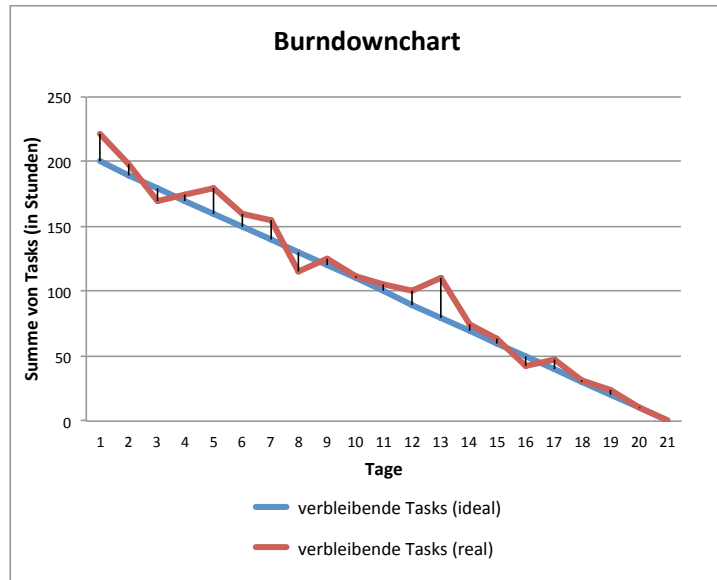


Abbildung 2: Burndown Chart in Anlehnung an [mountaingoatsoftware.com](http://mountaingoatsoftware.com)

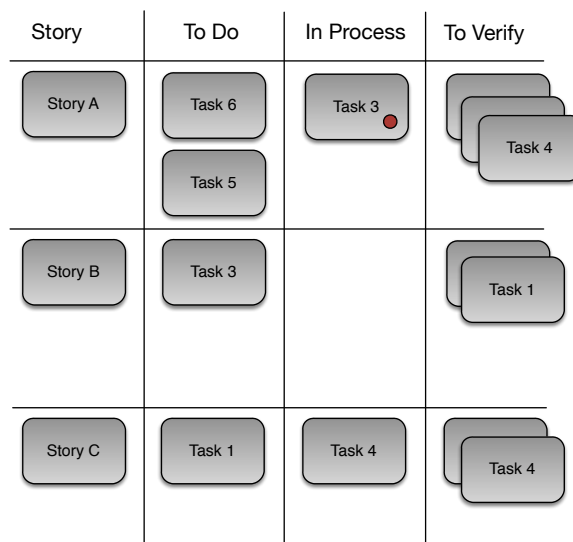


Abbildung 3: Taskboard in Anlehnung an [mountaingoatsoftware.com](http://mountaingoatsoftware.com)

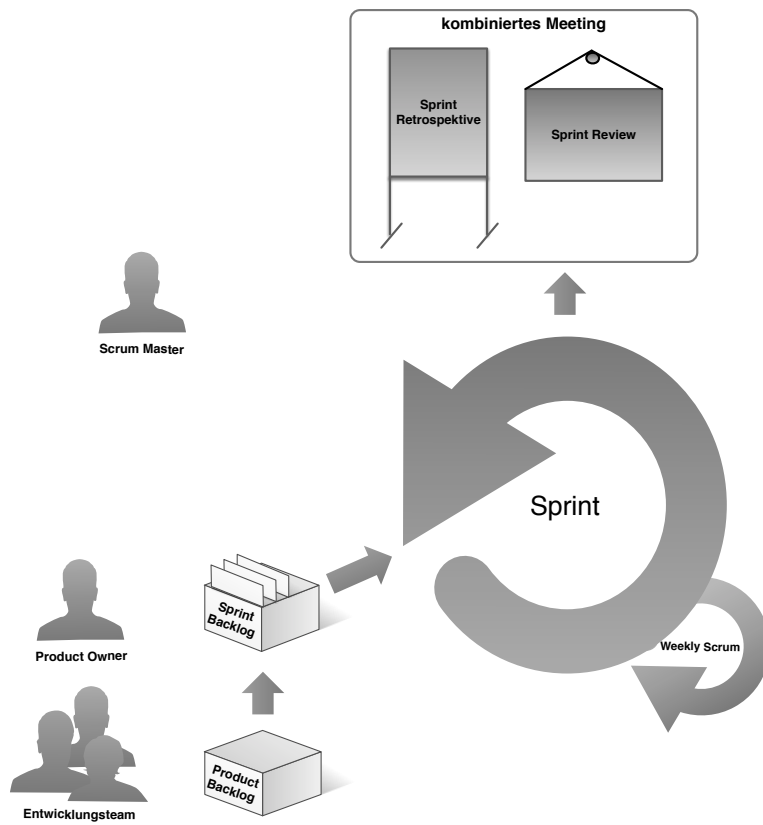


Abbildung 4: Scrum Prozess in PG eigene Darstellung

## **C Protokolle**

Auf den folgenden Seiten befinden sich die Protokolle sämtlicher Sitzungen der Projektgruppe während des Projektzeitraumes.

# Protokoll

## Projektgruppe BIaaS

Sitzungsdatum	16. Oktober 2014
Uhrzeit	10:00 - 11:30 Uhr
Protokollant	Hussein Khalil Hussein
Moderator	Oliver Norkus
Anwesend	Ilias Schweer, Brian Clark, Björn Friedrich, Oliver Norkus, Hussein Khalil Hussein, Babak Izadpanah, Alexander Zimak, Haji Shuker Qasim.
Abwesend	Florian Merkel (Arbeit)
Nächster Protokollant	Ilias Schweer

## Tagesordnung

1. Datenmanagementsystem und Zeiterfassungssystem
2. Rollenverteilung
3. Einheitlichkeit der Textverarbeitungssoftware
4. Vorlage der Seminararbeit
5. Empfehlenswerte Termine
6. Projektgruppen Termine

## 1 Datenmanagementsystem und Zeiterfassungssystem

Die Entscheidung der Projektgruppe für ein geeignetes Datenmanagementsystem fiel auf das Produkt *Confluence*. Confluence ist ein geeignetes Datenmanagementsystem, das auch gleichzeitig als Wiki-Software genutzt werden kann. Darüber hinaus haben wir uns für die bessere Koordination des Projekts für das Projektmanagementtool Jira entschieden. Beide Softwaresysteme wurden von Atlassian entwickelt und können miteinander verknüpft werden.

## 2 Rollenverteilung

Im nächsten Treffen werden erste Rollen für die Projektmitglieder verteilt.

### **3 Einheitlichkeit der Textverarbeitungssoftware**

Um während der Projektarbeit einheitliche Abgaben von Dokumentationen und Texten zu gewährleisten, müssen sämtliche Dokumente mit LateX verfasst werden. Alle Textdokumente müssen als PDF abgegeben werden.

### **4 Vorlage der Seminararbeit**

Die zur Verfügung gestellte LateX-Vorlage soll mit folgenden zusätzlichen Informationen im Kopf der Arbeit ergänzt werden:

- a.) Die Art der Arbeit (Seminararbeit).
- b.) Die Matrikelnummer des Studenten.
- c.) Der jeweilige Betreuer der Seminararbeit.

### **5 Empfehlenswerte Termine**

Am 6. November 2014 um 10.00 Uhr findet die Verteidigung der Bachelorarbeit von den beiden Gruppenmitgliedern Ilias Schweer und Alexander Zimak statt. Das Thema der Bachelorarbeit ist ein interessantes Thema aus dem Business Intelligence Umfeld: „Konzeption eines Modells zur Beschreibung von BI-Cloud-Angeboten“. Des Weiteren findet die Zwischenpräsentation der Projektgruppe LSOG am 21.10.2104 um 14 Uhr statt.

### **6 Projektgruppen Termine**

Im nächsten Treffen wird ein regelmäßiges wöchentliches Treffen vereinbart. Die Abgabe der Seminararbeiten wird am 17. November 2014 sein. Die Seminararbeit wird als Pdf Format in Confluence hochgeladen. Am 20. November 2014 finden die Vorträge der einzelnen Seminararbeiten statt. Jeder hat 25 Minuten Zeit, um seine Seminararbeit zu präsentieren. Anschließend werden über 5 Minuten Fragen gestellt.

# Protokoll

## Projektgruppe BIaaS

Sitzungsdatum	30. Oktober 2014
Uhrzeit	08:15-09:25 Uhr
Ort	Offis Raum O-100
Protokollant	Ilias Schweer
Moderator	Brian Clark
Anwesend	Ilias Schweer, Brian Clark, Björn Friedrich, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak, Haji Shuker Qasim
Abwesend	Hussein Khalir Hussein (zurückgetreten am 29.10.14)
Nächster Protokollant	Brian Clark

## Tagesordnung

1. Jira
2. Confluence
3. Rollen
4. Marketing
5. Versionsverwaltung
6. Organisatorisches

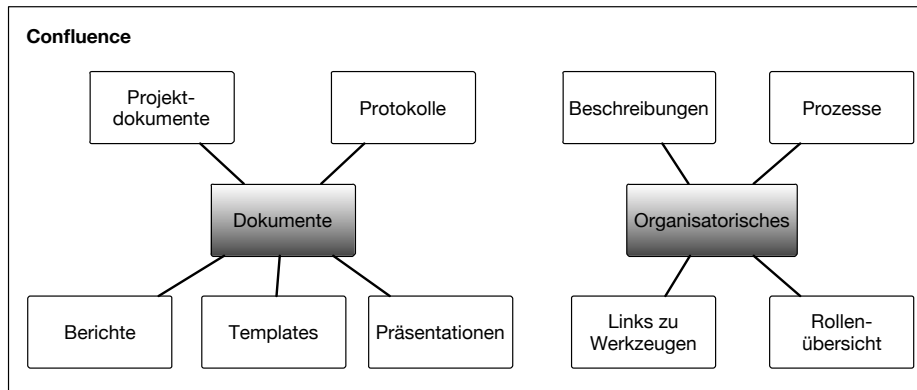
### 1 Jira

Jira wurde von Florian auf dem gegebenen UNIX-Server aufgesetzt. Dazu wurde von Florian eine Oracle Datenbank aufgesetzt. Ein Schema auf dieser Datenbank für Jira/Confluence ist zu dieser Zeit noch nicht erstellt worden. Die Benutzerkonten für die Gruppenmitglieder werden von Florian eingerichtet. Wünsche für den Benutzernamen können die Teilnehmer bis Sonntag, 02.11.14 an Florian richten, ansonsten wird der Benutzername nach dem Schema „erster Buchstabe des Vornamens plus Nachname“ (Bsp.: Ilias Schweer wird zu ischweer) erstellt. Jedes Konto erhält zudem ein generiertes Standardpasswort, welches von den Mitgliedern im Anschluss manuell geändert werden sollte. Die Serveradresse für Jira lautet **134.106.56.34** und der Port lautet **8080**. Als Zeiterfassungsplugin wird ICTime verwendet.



## 2 Confluence

Confluence ist ebenfalls von Florian aufgesetzt worden und ist in Jira integriert. Es wird dementsprechend kein separates Benutzerkonto benötigt. Der Port für Confluence lautet **8090**. In Confluence ist eine Wiki-Funktion enthalten. Diese wird von der Projektgruppe genutzt. Confluence dient der Lagerung von Dokumenten jeglicher Art. Folgende Struktur der Kategorien von Dokumenten wurde erarbeitet:



## 3 Rollen

Folgende Rollen wurden an Gruppenmitglieder verteilt:

Rolle	Name
Projektleiter	Alexander Zimak
Serverbeauftragte	Florian Merkel, Brian Clark
Dokumentenbeauftragter	Ilias Schweer
Webpräsenz	Björn Friedrich
Sozialbeauftragter	Haji Shuker Qasim

## 4 Marketing

Es soll eine Webpräsenz für dieses Projekt eingerichtet werden. Die zuständige Person ist Björn.

## 5 Versionsverwaltung

Als Tool zur Versionsverwaltung wird zukünftig Apache Subversion (SVN) genutzt. Die Relevanz für eine Versionsverwaltung und damit auch die Einrichtung dieser erfolgt erst nach den Seminararbeiten. Als Standardzeichencodierung für Textdokumente auf dem SVN-Server wurde UTF-8 festgelegt.

## 6 Organisatorisches

- Aus der Gruppe zurückgetretene Personen werden aus dem E-Mail-Verteiler entfernt.

- Antwort-Mails sollen zukünftig an den Verteiler gerichtet sein.
- Das Protokoll des vorherigen Treffens wird von Ilias überarbeitet.
- Der jeweils nächste Moderator schickt einige Tage vor dem Meeting die von ihm erarbeitete Tagesordnung an die Teilnehmer.
- Das Protokoll ist bis spätestens vier Tage nach dem Meeting online zu stellen.
- Die Treffen beginnen von nun an immer c.t. (wenn nicht anders kommuniziert)

### **6.1 Nächster Termin**

Das nächste Treffen findet am Dienstag, 11.11.14 um 08 Uhr c.t. in Raum O-100 im Offis statt.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	11. November 2014
Uhrzeit	08:19-08:42 Uhr
Ort	Offis Raum O-100
Protokollant	Brian Clark
Moderator	Ilias Schweer
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Oliver Norkus, Haji Shuker Qasim, Ilias Schweer, Alexander Zimak
Abwesend	—
Nächster Protokollant	Björn Friedrich

## Tagesordnung

1. Letzte Protokolle
2. Jira / Confluence
3. Weiteres Vorgehen
4. Organisatorisches

### 1 Letzte Protokolle

Die beiden letzten Protokolle (vom 16.10.2014 und 30.10.2014) wurden akzeptiert.

### 2 Jira / Confluence

Während Confluence (auf Port 8090) über das Internet erreichbar ist, ist Jira (auf Port 8080) noch nicht erreichbar. Florian wird Auskunft über die Fehler einholen.

### 3 Weiteres Vorgehen

- Auf den Servern SVN installieren.
- Benutzerkonten für die Anwendungen auf dem Server anlegen.

- Die Website für die Projektgruppe könnte auf unserer VM oder einem Uni Server betrieben werden.
- Es muss ein passender Name für die PG-Webseite gefunden werden. Die vorgeschlagenen Namen sind "CBI" und "CloudBI".

## **4 Organisatorisches**

- Zur Teambildung wollen wir an einem Mittwoch im November oder Dezember auf den Weihnachtsmarkt gehen.

### **4.1 Nächster Termin**

Das nächste Treffen findet am Dienstag, 25.11.14 um 08 Uhr c.t. im OFFIS U-64 statt.

# Protokoll

Projektgruppe  
BIaaS

Datum	25.11.2014
Uhrzeit	08:20h -09:55h
Protokollant	Björn Friedrich
Moderator	Brian D. Clark
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Haji Shuker Qasim, Alexander Zimak
Abwesend	-
Nächster Protokollant	Babak Izadpanah

## Tagesordnung

1. Vortrag Oliver Norkus: Grundlagen Modellierung und Architektur
2. Planung der Anforderungserhebung
3. JIRA/Confluence
  - a) Beide über die Ports erreichbar?
  - b) Hat jeder sein Benutzerkonto getestet?
  - c) Lizenz für Confluence
4. Name der Projektgruppe
5. Workflow Protokoll
6. Organisatorisches

## 0 letztes Protokoll

Das Protokoll ist soweit in Ordnung. Alexander überarbeitet den Satzbau und die Formulierungen. Dann wird das endgültige Protokoll per E-Mail an alle verteilt.

# 1 Modellierung und Architektur (Vortrag Oliver Norkus)

Die präsentierten Folien sind im Confluence verfügbar. Es gibt drei Ergebnistypen nach March und Smith (1995). In der ersten Spalte sind die Begriffe nach March und Smith zu sehen und in der zweiten die entsprechenden Ergebnisse der Projektgruppe.

1. Konstrukte: Fach-Terminologie aus der Energiewirtschaft, BI, Cloud ect. und evtl. ein Glossar
2. Modelle: User Stories
3. Methoden: Scrum als Entwicklungsprozess
4. Instanzen: Der Prototyp

Auf Folie 12 ist eine Abbildung zu sehen, die entsprechend abgeändert in der Abschlussdokumentation auftauchen könnte. Die Graphik illustriert das Phasenmodell des Design Science Research. Die einzelnen Phasen bedeuten für die Projektgruppe folgendes:

## Phase 1

Die Problemidentifikation und die Motivation der Projektgruppe steht für die Projektgruppenmitglieder nicht im Vordergrund. Das Thema hat Relevanz und die Carl von Ossietzky Universität ist in diesem Themenbereich engagiert.

## Phase 2

In der Phase 2 wird die Projektgruppe die Anforderungsanalyse durchführen.

## Phase 3

In dieser Phase wird das System modelliert und implementiert. Danach steht ein Prototyp zur Verfügung, der in Phase 4 eingesetzt werden kann.

## Phase 4

In dieser Phase wird der Prototyp mit den Daten für den Anwendungsfall getestet.

## Phase 5

In der vorletzten Phase wird der Prototyp und seine Funktionalität dem Kunden präsentiert. In diesem Fall einem speziellen Kunden, nämlich Marco Haas. Ist der Kunde mit dem System zufrieden, dann hat das System die Evaluation erfolgreich durchlaufen.

## Phase 6

Die letzte Phase ist für die Projektgruppe nur bedingt von Bedeutung. Ist das System so weit fertiggestellt, dass es in der Öffentlichkeit präsentiert werden kann, dann wird das in dieser Phase vorbereitet.

## 2 Anforderungserhebung

Es wird zwischen fünf und sechs Termine geben. Ein Termin soll ca. 1h dauern und es werden jeweils Kunden anwesend sein. Bei diesen Terminen muss nicht die gesamte Projektgruppe anwesend sein. Eine Delegation von drei bis vier Leuten reicht aus. Für die Gespräche werden spezielle Rollen an die Projektgruppenmitglieder verteilt. Im Gespräch sollen die Anforderungen erhoben werden. Ziel ist den Anforderungskatalog bis zum 23.12.2014 fertigzustellen. Dazu gibt es den folgenden Meilensteinplan:

### Meilensteinplan

1. 02.12.2014: Entscheidungen: Tool, Vorgehen, Erhebungskonzept, Rollenverteilung und evtl. Vorlage für das Anforderungsdokument und Use Cases.
2. 12.12.2014: Alle Termine wahrgenommen und die nötigen Artefakte liegen vor.
3. 16.12.2014: Die erhobenen Anforderungen sind sauber aufgeschrieben und aufgearbeitet.
4. 21.12.2014: Die Anforderungsdokumente werden dem QS-Manager vorgelegt.
5. 23.12.2014: Der QS-Manager übersendet Oliver Norkus, Marco Haas und etwaigen anderen Personen die Dokumente.

## 3 JIRA/Confluence

Es wurde beschlossen, dass Marco Haas ein Benutzerkonto für Confluence erhalten soll. Weiterhin sollen die Seminararbeiten und die Präsentationen im PDF Format im Confluence bereitgestellt werden. Es werden entsprechende Bereiche im Confluence angelegt. Die Dateinamen sind nach dem folgenden Schema zu generieren:

Thema, Seminararbeit bzw. Präsentation, Name, Vorname

### a) Beide über die Ports erreichbar?

JIRA und Confluence sind teilweise erreichbar. Jedes Projektgruppenmitglied testet den Zugang von zu Hause aus und schreibt Florian bis morgen eine E-Mail mit dem Ergebnis.

### b) Hat jeder sein Benutzerkonto getestet?

siehe a

### c) Lizenz für Confluence

Florian spricht mit Ralf Krause bzgl. der Lizenz. Die Projektgruppe wird eine Lizenz erhalten. Notfalls kauft die Carl von Ossietzky Universität neue.

## 4 Name der Projektgruppe

Die Projektgruppe wurde auf den Namen Cloud-BI getauft.

## 5 Workflow Protokoll

Bis zum Treffen am 02.12.2014 soll eine graphische Darstellung des Prozesses präsentiert werden. Der Prozess dauert drei Tage. Phase 4 und 5 finden am selben Tag statt. Der folgende Ablauf ist für die Anfertigung und Verifikation eines Protokolls einzuhalten:

### Phase 1

Die Rohfassung wird vom Protokollanten am Tag des Treffens fertiggestellt und an die Mitglieder der Projektgruppe übersandt. (Zeitpunkt:  $t_0$ )

### Phase 2

Alle Mitglieder der Projektgruppe haben einen Tag Zeit, um Verbesserungsvorschläge zu machen. (Zeitpunkt:  $t_0 + 1$ )

### Phase 3

Der Protokollant hat einen Tag Zeit, das Protokoll gemäß den Rückmeldungen anzupassen. (Zeitpunkt:  $t_0 + 2$ )

### Phase 4

Der QS Manager erhält das fertige Protokoll zur Validierung. Er hat einen Tag Zeit, bis das Protokoll an Oliver geschickt werden muss. (Zeitpunkt  $t_0 + 3$ )

### Phase 5

Der QS Manager schickt das Protokoll an Oliver Norkus und lädt es im Confluence hoch. (Zeitpunkt  $t_0 + 3$ )

## 6 Organisatorisches

Der nächste Termin ist am Donnerstag den 27.12.2014 um 08:00h c.t. in der ARBI. Der Besuch des Weihnachtsmarktes wurde auf den 08.12.14 um 18 Uhr terminiert.



# Protokoll

Projektgruppe  
BlaaS

Datum	27.11.2014
Uhrzeit	08:15h - 10:00h
Protokollant	Björn Friedrich
Moderator	-
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Ilias Schweer, Haji Shuker Qasim, Alexander Zimak
Abwesend	-
Nächster Protokollant	Babak Izadpanah

## 1 Tagesordnung

1. Tagesordnung für das Treffen am 02.12.2014
2. Tools
3. Planung Termine
4. Aufgaben

## 2 Tagesordnung 02.12.2014

1. Protokoll
2. Workflow Protokoll
  - a) Workflow Diagramm vorstellen und von Oliver Norkus verifizieren lassen
3. Präsentation des Meilensteinplans
4. Organisation Termine
  - a) Terminvorschläge
  - b) Erhebungskonzept
  - c) Fragenkatalog
  - d) Verteilung des Fragenkatalogs vorab?
5. Tools und evtl. Rollenverteilung
6. JIRA/Confluence

- a) Verlinkung
- b) Hochladen

7. Weihnachtsmarkt am 03.12.2014

### 3 Tools

Das Erstellen von Diagrammen soll mit dem Tool UMLet erfolgen. Das Tool ist in Java implementiert und steht somit für alle Betriebssysteme zur Verfügung. Bei Bedarf steht auch ein Plugin für Eclipse bereit.

### 4 Planung Termine

Folgende Termine werden vorgeschlagen. In der linken Spalte sind die Termine zu sehen und in der rechten die Personen, die an den Terminen teilnehmen können.

Mittwoch 14:00h - 16:00h:	Björn (nur am 03.12.), Florain, Brian, Babak, Alexander, Ilias
Donnerstag 08:00h - 10:00:	Brian, Florian, Babak, Ilias, Alexander, Haji
Donnerstag 14:00h - 16:00h:	Brian, Florian, Babak, Ilias, Alexander
Freitag 08:00h - 12:00h:	Brian (nur von 08:00h - 10:00h), Florian, Babak, Ilias, Haji, Alexander

### 5 Aufgaben

Björn: Fragenkatalog bis Sonntag  
Ilias: Diagramm Workflow Protokoll bis Sonntag  
Florian: Diagramm Meilensteinplan bis Sonntag

# Protokoll

Projektgruppe  
BlaaS

Sitzungsdatum	4. Dezember 2014
Uhrzeit	08:15-09:00 Uhr
Ort	Offis Raum E06
Protokollant	Banak Izadpanah
Moderator	Alexander Zimak
Anwesend	Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak, Haji Shuker Qasim
Abwesend	Björn Friedrich (musste arbeiten)
Nächster Protokollant	Haji Shuker Qasim

## Tagesordnung

1. Letztes Protokoll
2. Workflow Protokoll
3. Interview
4. Tools
5. Jira / Confluence
6. Organisatorisches

### 1 Letztes Protokoll

Das vorherige Protokoll wurde abgenommen.

### 2 Workflow Protokoll

Die Visualisierung des Workflows zum Verfassen von Protokollen wurde vorgestellt. Sie ist prinzipiell gut, aber die Notation entspricht keinem Standard und muss noch einmal überarbeitet werden.

### 3 Interview

- Termine:
  - Donnerstag, 04.12.14, 14-16 Uhr (OFFIS / O100)
  - Donnerstag, 11.12.14, 08-10 Uhr (OFFIS / U64)
  - Donnerstag, 18.12.14, 14-16 Uhr (in den Räumlichkeiten der Ceyoniq)
- erster Fragenkatalog:
  - Wie sieht der Business Case aus?
  - Wie und welche Kennzahl soll dargestellt werden?
  - Soll es verschiedene Benutzerrollen geben?
- Rollen während der Interviews:
  - Moderatoren (einer bis zwei), die Fragen stellen.
  - Protokollanten (mind. drei), die das Interview schriftlich festhalten.
  - Technische Hilfsmittel wie ein Diktiergerät, um nachträgliche Verbesserung vorzunehmen.

### 4 Tools

Für die UML-Modellierungsphase wird UMLet verwendet. Für die Zeitverfassung arbeiten wir mit Jira. Da Jira aktuell noch mit einer Testlizenz läuft, erfasst zunächst jeder Projektteilnehmer manuell seine eigene Projektarbeitszeit.

### 5 Jira / Confluence

Jira ist nun auf Port 18080 erreichbar.

### 6 Organisatorisches

- Der letzte Termin am 02.12.2014 musste ausfallen (Oliver fiel krankheitsbedingt aus).
- Am Montag, den 08.12.2014 treffen wir uns um 18:00 Uhr vor Galeria Kaufhof, um gemeinsam den Weihnachtsmarkt zu besuchen.
- Das nächste Treffen findet am Dienstag, 09.12.2014 um 08:15 Uhr c.t. in Raum U-64 im Offis statt.

# Protokoll: Interview vom 04.12.2014

## Projektgruppe BIaaS

Sitzungsdatum	4. Dezember 2014
Uhrzeit	08:15-09:00 Uhr
Ort	Offis Raum O-100
Protokollant	Brian Clark, Babak Izadpanah, Florian Merkel, Ilias Schweer
Moderator	Alexander Zimak
Anwesend	Brian Clark, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak
Befragter	Marco Haas

Der Business Case sieht vor, dass das System als Cloud System für Business Intelligence Analysen realisiert wird. Für den Anwendungsfall Energiewirtschaft bedeutet das, dass ein Keyaccounter zu jeder Zeit und an jedem Ort einem Kunden ein Angebot machen kann. Dieses Angebot wird on Demand berechnet werden. Es werden nur Großkunden betrachtet. Ein Kunde gilt als Großkunde, wenn er mehr als 100.000 kWh pro Jahr verbraucht. Der Deckungsbeitrag muss für 15 Minuten-Intervalle berechnet werden. Das System braucht alle relevanten Daten des Kunden. Das ist u.a. der Standort, an dem die Energie verbraucht wird. Diese Daten liegen im IS-U oder CRM System. Bei neuen Kunden muss eine Stammdatenanfrage beim bisherigen Versorger durchgeführt werden. Der Kunde benötigt einen Preis, den er mit anderen Anbietern vergleichen kann. Der Lieferant muss dem Kunden einen Preis für die kWh machen, der für eine deutschlandweite Versorgung gilt. Die Daten über alle Standorte in Deutschland müssen aggregiert werden. Die Netznutzungskosten an den verschiedenen Standorten können sich unterscheiden. Der Keyaccounter muss einen Preis berechnet bekommen anhand dessen er entscheiden kann ab welchem Preis pro kWh sich die Versorgung des Kunden rechnet.

Weiterhin muss für die Berechnung ein Fahrplan erstellt werden. Dieser muss den unterschiedlichen Energiebedarf zu verschiedenen Zeiten berücksichtigen. Beispielsweise wird am Wochenende, an Feiertagen usw. eine andere Menge an Energie benötigt als an Werktagen.

Das System muss neben der Berechnung des Preises pro kWh noch drei Deckungsbetreibungsrechnungen durchführen. Dem Keyaccounter müssen die Netznutzungskosten, die Arbeitskosten und die Leistungskosten bereitgestellt werden. Der Preis pro kWh muss für viertel Stunden Intervalle sichtbar und für jeden Standort einzeln und in der Summe sein. Die Darstellung kann als Tabelle oder Diagramm erfolgen.

Das System muss ein User-Management ermöglichen. Es soll drei verschiedene Rollen geben: der Kunde, der Netzbetreiber und der Lieferant bzw. der Händler.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	9. Dezember 2014
Uhrzeit	08:15-09:15 Uhr
Ort	Offis Raum E06
Protokollant	Florian Merkel
Moderator	Babak Izadpanah
Anwesend	Björn Friedrich, Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak, Haji Shuker Qasim
Abwesend	-
Nächster Protokollant	Haji Shuker Qasim

## Tagesordnung

1. Letztes Protokoll
2. Workflow
3. Auswertung des Interviews
4. Zweiter Fragenkatalog und Anforderungen
5. JIRA / Confluence
6. Sonstiges

### 1 Letztes Protokoll

Das Protokoll vom letzten mal wurde abgenommen. Der PG Name *Cloud BI* muss noch im vorletzten Protokoll vermerkt werden.

### 2 Workflow

Die neue Visualisierung des Workflows zur Verfassung von Protokollen wird vorgestellt. Die Notation (ECPK) ist so in Ordnung. Die Grafik wird im Confluence hochgeladen.

### 3 Auswertung des Interviews

Das EEPK Diagramm zur Visualisierung des Business Case wird vorgestellt. Dies sind die gewünschten Veränderungen:

- Der Begriff Energiehändler bzw. Energielieferant soll statt Händler verwendet werden.
- Es muss klarer werden, was mit Preis gemeint ist (→Gesamtkosten)

Das EEPK-Diagramm beschreibt nur die Errechnung der Gesamtkosten. Es muss noch ein weiteres erstellt werden, das die Deckungsbeitragsrechnung zeigt.

Die Gesamtkosten sind der Betrag, der die Basis für ein Angebot an einen neuen Kunden bildet. Ein Deckungsbeitrag ist ein Anteil am Erlös. Es können verschiedene Deckungsbeiträge berechnet werden (Arbeitskosten, Leistungskosten, Netznutzungskosten, Profit, ...). Die beiden Werte unterscheiden sich im wesentlichen in der „Richtung“ der Berechnung. Bei den Gesamtkosten werden die Teilkosten addiert. Für die Deckungsbeiträge wird der Erlös in seine Bestandteile aufgeteilt.

Bis zum Interview am 11.12.14 muss das zweite EEPK-Diagramm und ein Usecase-Diagramm erstellt werden. Bis zur nächsten Sitzung soll ein MER-Diagramm erstellt werden. MER-Diagramme werden kurz erklärt.

### 4 Zweiter Fragenkatalog und Anforderungen

Der Fragenkatalog muss noch bis zum 11.12.14 erstellt werden. Die ersten Anforderungen werden bis zum 11.12.14 konkretisiert und festgehalten.

### 5 JIRA / Confluence

JIRA und Confluence brauchen noch Lizenzen. Diesbezüglich wird von Florian nach der Sitzung eine E-Mail an Atlassian geschickt.

### 6 Sonstiges

Es soll darauf geachtet werden, dass interne E-Mails nur an den entsprechenden Verteiler geschickt werden. Die Verteiler sollen so konfiguriert werden, dass dem Betreff von E-Mails automatisch ein verteilerspezifisches Tag zugeordnet werden. Die Seminararbeiten sollen bis zur Abgabe des Abschlussberichts überarbeitet werden. Dabei darf der Inhalt nicht verändert werden. Rechtschreibfehler und die Platzierung von Quellenangaben sollen beispielsweise korrigiert werden. Für die Raumreservierung steht für kurzfristige Anfragen ggf. das Raumbüro der Universität zur Verfügung. Zur längerfristigen Blockung des Raums im OFFIS wird eine Doodle Umfrage zur Terminfindung erstellt. Es wird eine Rufnummernliste mit den Handynummern der PG-Teilnehmer angelegt.

## **7 Nächster Termin**

Das nächste Treffen findet am Dienstag, 16.12.2014 um 08:15 Uhr c.t. in Raum U-64 im Offis statt.



# Protokoll: Interview vom 09.12.2014

## Projektgruppe BIaaS

Sitzungsdatum	9. Dezember 2014
Uhrzeit	08:25-10:00 Uhr
Ort	Offis Raum E-02
Protokollant	Brian Clark, Babak Izadpanah, Florian Merkel, Haji Shuker Qasim, Ilias Schweer
Moderator	Alexander Zimak
Anwesend	Brian Clark, Babak Izadpanah, Florian Merkel, Oliver Norkus, Haji Shuker Qasim, Ilias Schweer, Alexander Zimak
Befragter	Marco Haas

Zu Beginn des Interviews wurden zwei eEPK Diagramme, die Use Cases und die Anforderungstabelle vorgestellt. Das eEPK für den Deckungsbeitrag ist in Ordnung. Eventuell könnte man die Quellsysteme ebenfalls in das Diagramm einfügen. Dem eEPK Diagramm für die Gesamtkosten muss eine Begriffserklärung hinzugefügt werden. Weiterhin muss die Anforderungstabelle überarbeitet werden. Die Anforderung 003F muss so modifiziert werden, dass der Energielieferant detailliert abgebildet wird und der Kunde nur aggregiert. Möglicherweise muss die Anforderung in zwei Anforderungen aufgeteilt werden.

Bei der Anforderung 004F fehlt die Kommunikation zwischen den Systemen. Weiterhin müssen die nichtfunktionalen Anforderung zur Analysegeschwindigkeit und zu den Charakteristika von Cloud-Diensten hinzugefügt werden. (Für mich sind das eher funktionale Anforderungen) Es soll ein Service angeboten werden, der branchen- und bundesweit von mehreren Kunden gleichzeitig genutzt werden kann.

Es wurde die Frage diskutiert, wie sich der Leistungspreis und der Arbeitspreis zusammensetzen und voneinander abgrenzen. Der Leistungspreis beinhaltet die Unterhaltskosten für die Kraftwerke. Der Arbeitspreis die Kosten, die anfallen, wenn Strom erzeugt wird. Das ist z.B. der Preis für Steinkohle zzgl. Transport usw. Beide Preise müssen viertelstündlich berechnet werden, denn sie sind abhängig von Zeit und dem aktuellen Marktpreis. Die Nutzungskosten sind nicht linear. Deshalb müssen sie auch nahezu in Echtzeit berechnet werden. Die lastabhängigen Produktionskosten werden sind nicht linear, werden aber der Einfachheit halber als linear angenommen. Die Netznutzungskosten ebenfalls zeitabhängig und müssen deshalb auch für ein viertelstündliches Intervall berechnet werden. Großkunden erhalten im Gegensatz zu Privatkunden zwei Rechnungen. Auf der einen werden die Netznutzungskosten aufgeführt und

auf der anderen der Leistungsverbrauch. Bei Privatkunden werden die Netznutzungskosten vom Lieferanten verwaltet.

Weiterhin wurde erörtert über welche Dimensionen aggregiert werden soll. Es wurden die folgenden sechs Dimensionen festgelegt:

- Kunde (z.B. IKEA)
- Geschäftsstelle (z.B. Oldenburg)
- Zählpunktnummer (“MAC-Adresse des Entnahmepunktes“)
- Netz (z.B. EWE-Netz)
- Viertelstunde
- Leistungskosten und Arbeitskosten für die gesamten Standorte

Die Angebote werden für Verträge erstellt, die eine Laufzeit von einem Jahr haben. Hinzu kommt, dass jeder Benutzer nur mit seinen Daten arbeiten kann. Die Daten anderer Benutzer dürfen nicht sichtbar sein.

Abschließend wurden die benötigten OLAP-Funktionen diskutiert. Die Daten müssen über die Zeit auswertbar sein. Jeweils in den Zeiträumen von 15 min. über 1h und 24h bis hin zu einem Monat. Dann sollen die Daten über den Kunden und eine Kundengruppe oder über Kraftwerke mit den Attributen Produkt, Art des Kraftwerks und die Kosten des Kraftwerks ausgewertet werden können. Weiterhin sollen die Daten auch nach Standort und Übertragungsnetzregion sortierbar sein. Die Profile müssen in einem Gesamtprofil über alle Regionen und Zählpunkte zusammengefasst werden können. Weiterhin muss der Leistungspreis und der Arbeitspreis über alle Regionen und Zählpunkte aggregiert werden können.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	16. Dezember 2014
Uhrzeit	08:20-10:04 Uhr
Ort	Offis Raum U64
Protokollant	Haji Shuker Qasim
Moderator	Florian Merkel
Anwesend	Oliver Norkus, Ilias Schweer, Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Haji Shuker Qasim, Alexander Zimak
Abwesend	-
Nächster Protokollant	Alexander Zimak

## Tagesordnung

1. Letztes Protokoll
2. Lastenheft
3. Website
4. Arbeitszeiten
5. Etikette
6. Termine
7. Sonstiges
8. Präsentation

### 1 Letztes Protokoll

Die Erklärung des Deckungsbeitrags ist nicht ganz korrekt und muss überarbeitet werden. Ebenfalls soll in der finalen Version ein Glossar enthalten sein. Dazu müssen noch einige Tippfehler korrigiert werden.

## 2 Lastenheft

Sämtliche Anmerkungen von Oliver bzgl. des Lastenhefts stehen in der entsprechenden PDF, darunter:

- Der Projekttitel muss angepasst werden. Der Begriff Realtime ist nicht angebracht.
- Kapitel fünf (Lieferumfang), sechs (Phasen- und Meilensteinplanung) und acht (Abnahmekriterien und Qualitätsanforderung) werden für das Dokument nicht benötigt und können entfernt werden.
- Das Use Case Diagramm muss überarbeitet werden: Kundendaten auswählen als <include> von Daten analysieren. Bei den Akteuren sollen keine Vorbedingungen stehen.
- Im MockUp soll im Frontend ein Footer pro Dimension oder mehrere Dimensionen für z.B. Kunde oder Region der Energielieferanten enthalten sein. Dazu soll es eine Visualisierung in Form eines Diagramms geben.
- Die Anforderungstabelle muss hinsichtlich einiger Punkte überarbeitet werden (nicht einfach die klassischen Cloud-Charakteristika als Beschreibung und Titel übernehmen).
- Im MER-Diagramm soll der Standort in die Kundendimension verschoben werden.

## 3 Website

Björn soll einen Vorschlag für die PG-Website bis zum nächsten Treffen am 6.1.2015 entwickeln. Die Website soll eine Projektbeschreibung und ein Logo enthalten. Jedes PG-Mitglied kann dazu entsprechende Texte für die Website in Confluence bereitstellen. Eine lokale Speicherung der ersten Website-Version sollte bis Anfang Januar verfügbar sein.

## 4 Arbeitszeiten

Die Zeiterfassung in Jira soll bis zum 19.12.2014 von jedem Mitglied erstellt worden sein. Dazu kann jeder seine Zeiten in Form *von-bis* oder benötigte Zeit angeben. Oliver kann jemanden vom BTC bitten, eine kleine Vorstellung über das Zeitmanagement in Jira zu halten.

## 5 Etikette

Folgende Punkte sollen im Rahmen von Interviews in der Wirtschaft aufgrund geltender Etikette (Verhaltensregeln) von der Gruppe eingehalten werden:

- Außer dem offiziellen Protokollant und Moderator soll niemand seinen Laptop oder Tablet benutzen (offen stehen haben).
- Es sollen keine privaten Gegenstände, wie Taschen oder Essen auf dem Tisch frei rumliegen.
- Jeder soll sich während des Interviews auf das Thema konzentrieren und damit einen guten Einfluss auf den Befragten ausüben.

## 6 Termine

- > Das nächste Interview findet am Freitag, den 19.12.14 um 13:00 Uhr im Ceyoniq statt (Ammerländer Heerstraße 70).
- > Abgabe des Lastenhefts am Dienstag, den 23.12.14.
- > Das nächste Treffen findet am Dienstag, 6.1.2015 um 08:15 Uhr c.t. in Raum U-64 im Offis statt.

## 7 Sonstiges

- Björn hat den ersten Entwurf des Sequenzdiagramms vorgestellt.
- Alex hat kurz das Vorgehen bei der Zeiterfassung in Jira erklärt.
- Florian soll die Email von Atlassian beantworten.
- Oliver hat vorgeschlagen einen Artikel über das Projekt in einem wissenschaftlichen Magazin mit uns zu veröffentlichen. Jeder der möchte, kann bei der Erstellung mitwirken und ist damit Autor des Artikels. Frist für die Abgabe des Manuskripts ist der 26.1.2015.

## 8 Präsentation

- Die Folien sollen etwas konkreter sein.
- Welche Herausforderungen sind genau gemeint?
- Auf Folie 3 soll das Wort Höhere zu Höhe geändert werden.
- Auf Folie 4 soll eine Zusammenfassung für nichtfunktionale Anforderungen stehen.
- Auf Folie 5 soll der Punkt Verfügbarkeit konkretisiert werden.
- Folien über die Timeline und dem groben Überblick fehlen noch.

# Protokoll: Interview vom 19.12.2014

## Projektgruppe BIaaS

Sitzungsdatum	19. Dezember 2014
Uhrzeit	13:00-15:00 Uhr
Ort	CEYONIQ Ammerländer Heerstraße 70, 26129 Oldenburg
Protokollant	Brian Clark, Babak Izadpanah, Florian Merkel, Ilias Schweer, Haji Shuker Qasim
Moderator	Alexander Zimak
Anwesend	Brian Clark, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak, Haji Shuker Qasim
Befragter	Marco Haas

Während des Interviews wurden vier Fragen erörtert. Begonnen wurde mit der Frage nach der manuellen Anbindung von Quellsystemen oder Dateien. Im Anschluss wurden die Eingabewerte diskutiert und in wie weit eine Prognose des Verbrauchs stattfinden soll. Zum Schluss wurde die Frage nach der Auslastungsvisualisierung der Kraftwerke beantwortet.

Die Daten werden aus festen Quellen bezogen. Also wird ein manueller Austausch von Quellen nicht möglich sein. Es werden Schnittstellen definiert, über die die Systeme angesprochen werden.

Man muss über verschiedene Dimensionen aggregieren und filtern können. Beides kann mit OLAP-Operationen realisiert werden. Der betrachtete Zeitraum ist großer Bedeutung. Es soll z. B. ein Jahr betrachtet werden können, um zu ermitteln, ob der Kunde sich rechnet oder ob der Preis neu kalkuliert werden muss.

Prognosedaten sind für die Projektgruppe wichtig. Mit den prognostizierten Daten kann entschieden werden, ob es günstiger ist den Strom an der Börse zu kaufen oder ihn im eigenen Kraftwerk zu produzieren. Die Prognosen müssen nicht selbst erstellt werden. Sie liegen in einem EDM-System und können von dort abgerufen werden. In der Visualisierung sollen die prognostizierten Werte und die normalen Werte in verschiedenen Farben dargestellt werden. Fehlt in den prognostizierten Daten eine Variable, dann wird diese gleich null gesetzt.

Die Fahrpläne der Kraftwerke müssen in der Auswertung sichtbar sein. Es muss ersichtlich werden, ob die Kraftwerke zu stark ausgelastet sind oder ob der Strom für den Kunden produziert werden kann. Sind keine Kapazitäten vorhanden, kann der Strom von der Börse gekauft werden.

Die Projektgruppe hat in der kommenden Zeit zwei wichtige Termine. Der erste Termin findet Mitte März im Rahmen einer Expertenkommission (EPK) am Niedersächsischen Ministerium für Wissenschaft und Kultur in Hannover statt. Bis zu diesem Zeitpunkt sollte ein präsentabler Prototyp fertiggestellt sein. Dieser sollte mit Beispieldaten vorgeführt werden können. Der zweite Termin findet zwischen dem 15. und 20. Januar mit Herrn Prof. Dr. Dr. - h.c. Hans-Jürgen Appelrath und Herrn Prof. Dr. Jürgen Sauer statt. Es werden die erhobenen Anforderungen präsentiert. Die heute verwendete Präsentation soll wiederverwendet werden. Dafür sind einige Anpassungen erforderlich. Die erste Seite muss das Datum enthalten und auf der Vorlage basieren. Auf der 3. Seite fehlen die Quellen, der Unterschied zwischen Analyseleistung und Abfrageleistung muss herausgestellt werden und die Verfügbarkeit und Erreichbarkeit sind dasselbe. Auf der Seite 4 muss genau spezifiziert werden, was große Unternehmen sind. In der Graphik Cloud-BI muss eine Verbindung zwischen "Angebot erstellen" und "Kunde analysieren" eingefügt werden. Weiterhin ist der Abstand zwischen den Zeilen nicht notwendig.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	6. Januar 2015
Uhrzeit	08:15-10:00 Uhr
Ort	Offis Raum U64
Protokollant	Alexander Zimak
Moderator	Haji Shuker Qasim
Anwesend	Oliver Norkus, Ilias Schweer, Brian Clark, Björn Friedrich, Babak Izadpanah, Haji Shuker Qasim, Alexander Zimak
Abwesend	Florian Merkel (Erkrankung)
Nächster Protokollant	Brian Clark

## 1 Tagesordnung

1. Letztes Protokoll
2. Seminararbeiten PG Cuberunner
3. Auswertung Anforderungskatalog
4. Architekturdesign und Technologien
5. Außendarstellung (Homepage, Artikel)
6. Jira Zeiterfassung
7. Sonstiges

## 2 Letztes Protokoll

Der Firmenname der CEYONIQ muss berichtigt werden. Das Protokoll wurde von Oliver Norkus und der Gruppe abgenommen.



### **3 Seminararbeiten PG Cuberunner**

Folgende Seminararbeiten wurden vorgestellt:

- Ilias Schweer
  - Stammdatenmanagement und Konsolidierung
- Alexander Zimak
  - Anforderungsanalyse und Konzeptarbeit in BI Projekten
- Björn Friedrich
  - Data Mining and Knowledge Discovery in Databases
  - Vergleich Microsoft & SAP Business Intelligence
- Babak Izadpanah
  - SAP NW BW & SAP BO 4.0
- Brian Clark
  - Self-Service Business Intelligence mit Microsoft Excel 2010, PowerPivot und Power View

Die noch fehlenden Seminararbeiten werden am 13.01.2015 vorgestellt.

### **4 Auswertung des Anforderungskatalog**

Der Anforderungskatalog wurde als zufriedenstellend befunden. Um Änderungen am Dokument vorzunehmen, muss das Feedback von Marco Haas abgewartet werden.

### **5 Architekturdesign und Technologien**

Nachdem das Lastenheft erstellt worden ist, erfolgt der Übergang zum Pflichtenheft. Diesbezüglich sollen erste Fragen geklärt werden. Zu diesen zählen das Ermitteln der nötigen UML-Diagramme und der Übergang vom MER-Diagramm zum Data Fact Table. Darüber hinaus soll eine erste Überlegung stattfinden, welche Bestandteile der BI Referenzarchitektur von Bauer und Günzel für unsere Zwecke in die Cloud ausgelagert werden können bzw. sollen. Im Zusammenhang des Architekturdesigns und der Technologien, soll geprüft werden ob das Angestrebte Produkt auch unter Einsatz von Open Source Software realisierbar wäre.

### **6 Außendarstellung (Homepage, Artikel)**

Oliver Norkus hat der PG vorgeschlagen einen Artikel für eine Fachzeitschrift zu verfassen. Es gibt unterschiedliche Fachzeitschriften in denen eine Veröffentlichung möglich wäre. Zu diesen zählen das ERP-Management Magazin sowie BI-Spektrum. Der geforderte Inhalt der Artikel unterscheidet sich je nach Fachzeitschrift. Beim Verfassen des Artikels soll im Allgemeinen darauf geachtet werden, dass der Use Case der PG im Mittelpunkt steht.

Bezüglich der Website gibt es noch keine neuen Erkenntnisse.

## **7 Jira Zeiterfassung**

Oliver Norkus ist mit der Zeiterfassung zufrieden. Zukünftig werden alle Aufwände wöchentlich gebucht, als Stichtag gilt immer der Sonntag der aktuellen Woche.

## **8 Sonstiges**

Unter Sonstiges ist festzuhalten, dass die Tonbandmitschnitte der Kundeninterviews in das Confluence geladen werden müssen und die Auswertungen dieser Interviews fertig gestellt werden sollen. Darüber hinaus sollen Ilias Schweer und Alexander Zimak eine Übersicht der für das Projekt geeigneten Produkte von Microsoft und SAP erstellen und vorstellen.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	13. Januar 2015
Uhrzeit	08:17-10:16 Uhr
Ort	Offis Raum U64
Protokollant	Brian Clark
Moderator	Haji Shuker Qasim
Anwesend	Oliver Norkus, Ilias Schweer, Brian Clark, Björn Friedrich, Babak Izadpanah, Haji Shuker Qasim, Alexander Zimak
Abwesend	-
Nächster Protokollant	Björn Friedrich

## 1 Tagesordnung

1. Letztes Protokoll
2. Seminararbeiten PG Cuberunner
3. Präsentation BI in der Cloud von Microsoft & SAP
4. Homepage
5. UML
6. Architektur
7. Data-Fact-Table
8. Sonstiges

## 2 Letztes Protokoll

Das letzte Protokoll muss dahingehend überarbeitet werden, dass der Grund für die Erstellung der Zusammenfassungen der Seminararbeiten der PG Cuberunner ersichtlich sein soll. Ebenfalls muss ein kleiner Fehler überarbeitet werden.

### **3 Seminararbeiten PG Cuberunner**

Folgende Seminararbeiten der PG Cuberunner wurden vorgestellt:

- Alexander Zimak
  - BI Projektmanagement
- Brian Clark
  - Mobile BI
- Florian Merkel
  - Testen
  - In-Memory

Die noch fehlenden Seminararbeiten werden am 20.01.2015 vorgestellt.

### **4 Präsentation BI in der Cloud von Microsoft & SAP**

Ilias Schweer und Alexander Zimak halten eine kurze Präsentation über die BI-Cloud Angebote von Microsoft und SAP. Die Gruppe soll einen Termin finden, an dem die Lösungen von SAP evaluiert und erprobt werden sollen. Eine Person soll diesbezüglich Vorbereitungen dafür treffen.

### **5 Homepage**

Es kann ein Content Management System für die Webseite benutzt werden. Ein Workflow Prozess soll ebenfalls vorhanden sein. Bis nächste Woche soll eine URL für die Website verfügbar sein. Die Homepage soll bei Ende Januar online sein. Björn Friedrich wird die Inhaltsstruktur per Email schicken.

### **6 UML**

Babak hält eine Präsentation über das weitere Vorgehen der UML-Diagramm-Phase. Als nächste Schritte stehen Sequenz- und Objektdiagramme an und der Data-Fact-Table muss weiter angepasst werden.

### **7 Architektur**

Für das Projekt ist die Rolle eines Architekturmanagers von nun an nötig. Diese Rolle bekleidet Babak. Weiter soll betrachtet werden, ob im Projekt VaaS gleich SaaS, MaaS gleich PaaS und DaaS gleich IaaS ist. Bis zum nächsten Treffen soll ein erster Plan für die Architektur stehen.

### **8 Data-Fact-Table**

Florian stellt seinen ersten Entwurf des Data-Fact-Tables vor. Eine Anmerkung ist, dass eine Hierarchie eingefügt werden soll.

## 9 Sonstiges

Ein internes Treffen ist für Donnerstag geplant. Die Präsentation über Open-Source Lösungen für unser Projekt von Brian Clark wurde aufgrund von Zeitmangel auf nächste Woche verschoben.

# Protokoll

## Projektgruppe BIaaS

Datum	20.01.2015
Uhrzeit	08:15h - 09:45h
Protokollant	Björn Friedrich
Moderator	Brian D. Clark
Anwesend	Brian Clark, Björn Friedrich, Marco Haas, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak
Abwesend	-
Nächster Protokollant	Babak Izadpanah

## Tagesordnung

0. Letztes Protokoll
1. Besprechung Lastenheft
2. Open Source
3. Architektur
4. Website/Homepage
5. Data-Fact-Table

## 0 Letztes Protokoll

Das Protokoll ist soweit in Ordnung. Nur der Moderator muss noch von Haji Shuker Qasim in Alexander Zimak geändert werden, weil Alexander spontan eingesprungen ist.

## 1 Besprechung Lastenheft

Oliver und Marco haben ihre Anmerkungen zum Lastenheft vorgestellt. Die Dokumente mit ihren Notizen wurden der Projektgruppe ausgehändigt. Anmerkungen siehe Dokumente.

## **2 Open Source**

Brian hat eine Auswahl an Open Source Software vorgestellt, die für die PG von Nutzen sein könnten. In der nächsten Woche wird noch einmal genau geklärt, inwieweit einzelne Produkte zusammenarbeiten können und auf welcher Ebene sie angesiedelt sind. Die Folien sind im Confluence zu finden.

## **3 Architektur**

Dieser Punkt wurde auf Donnerstag den 22.01.15 verschoben. Die PG wird sich während des Testtages ca. 2h nehmen, um sich um die Architektur zu kümmern.

## **4 Website/Homepage**

WebPress wird von Jörg eingerichtet. Allerdings dauert das noch ca. eine Woche. Die Verfügbarkeit der URL [www.cloud-bi.informatik.uni-oldenburg.de](http://www.cloud-bi.informatik.uni-oldenburg.de) wird derzeit geprüft. Eine Antwort von Prof. Dr. Jürgen Sauer steht jedoch noch aus.

Marco wird auf der Homepage als Sparringspartner aus der Energiewirtschaft aufgeführt werden. Auch die Verlinkung der zwei Ceyoniq Seiten (<http://www.ceyoniq.com> und <http://www.ceyoniq-consulting.com/sap-bihana>) sind von ihm genehmigt worden.

## **5 Data-Fact-Table**

Dieser Punkt wurde auf Donnerstag den 22.01.15 verschoben.

# Protokoll

## Projektgruppe BIaaS

Sitzungsdatum	27. Januar 2015
Uhrzeit	08:20-09:45 Uhr
Ort	Offis Raum U63
Protokollant	Babak Izadpanah
Moderator	Björn Friedrich
Anwesend	Ilias Schweer, Brian Clark, Oliver Nor- kus, Björn Friedrich, Babak Izadpanah, Florian Merkel
Abwesend	Alexander Zimak (Wegen Erkeltung)
Nächster Protokollant	Florian Merkel

## Tagesordnung

1. Letztes Protokoll
2. Vortrag Open Source
3. Zusammenfassungen Seminararbeiten (SQL, CRM und Nachhaltigkeit in BI)
4. Meilensteinplan Architekturkonzeption
5. Artikel ERP
6. Struktur und Inhalt Homepage
7. Sonstiges

## 1 Letztes Protokoll

Das Protokoll ist soweit in Ordnung. Einzige Anmerkung ist, dass Marco nicht als Kunde auf Homepage und sonstigen Texten oder Artikeln aufgeführt werden soll, sondern als Sparingspartner.

## 2 Vortrag Open Source

Brian hielt eine Präsentation über mögliche Open Source Lösungen für unser Projekt. Dabei fokussierte er sich auf Grundlagen und Beispielen der Open Source Möglichkeiten. Bis zum nächsten Mal soll Brian grafisch eine mögliche Architektur des Systems mit Open Source Lösungen vorstellen.



### **3 Zusammenfassungen von PG-CubeRunner (SQL, CRM und Nachhaltigkeit von BI)**

Zunächst hielt Babak zwei Referate, eins über Microsoft SQL Server 2012 Microsoft Business Intelligence und das andere über Analytical CRM. Im Anschluss hielt Ilias eine kurze Präsentation über Nachhaltigkeit in BI.

### **4 Meilensteinplan Architekturkonzeption**

- Die Gruppe kann selbst einen Vorschlag bzgl. der Architektur machen.
- Bis März soll die Architektur stehen, sodass dann die Implementierungsphase anfangen kann.
- Für die Präsentation für Herrn Prof. Sauer und Herrn Prof. Appelrath muss bis zum 28.1.2015 um 9 Uhr ein Terminvorschlag stehen.
- Das Lastenheft muss bis Samstag 12 Uhr überarbeitet an Oliver geschickt sein.
- Das nächste Treffen für die Architekturkonzeption findet am Donnerstag, 29.01.2015 statt. Die erste schriftliche Version des Architekturdiagramms soll am 3.02.15 vorgestellt werden.

### **5 Artikel ERP**

Das Manuskript des Artikels wurde von Oliver fristgerecht eingereicht. Nun muss auf eine Antwort gewartet werden.

### **6 Struktur und Inhalt Homepage**

- Nicht ausschließlich Text, sondern ebenso einige Bilder sollten auf der Homepage zu finden sein.
- Der Inhalt soll sich aus vorhandenen Präsentationen (Olivers und Florians) und ebenso aus dem Artikel zusammensetzen.
- Jeder Teilnehmer soll sein Profil in Confluence ausfüllen, damit Björn diese Informationen nutzen kann.
- Bis Ende des Monats ist die erste Version online.
- Das Logo sollen wir verfeinern und anpassen.

### **7 Sonstiges**

Das Komponentendiagramm wird beim nächsten Treffen weiter bearbeitet.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	3. Februar 2015
Uhrzeit	08:15-09:45 Uhr
Ort	Offis Raum U63
Protokollant	Florian Merkel
Moderator	Babak Izadpanah
Anwesend	Björn Friedrich, Ilias Schweer, Brian Clark (verspätet, entschuldigt), Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak, Marco Haas
Abwesend	-
Nächster Protokollant	Ilias Schweer

## Tagesordnung

1. Letztes Protokoll
2. Data Fact Table
3. MVC-Architektur
4. Open Source
5. Website
6. Zwischenstandspräsentation
7. Sonstiges

### 1 Letzes Protokoll

Zu Punkt 4: Anstatt „Samstag“ sollte das Datum des gemeinten Tages (31.01.15) im Protokoll stehen. Wenn die Änderung durchgeführt worden ist, gilt das Protokoll als abgenommen.

### 2 Data Fact Table

Das Modell der Data Fact Table wird vorgestellt. Es wurden folgende Anmerkungen bzw. Änderungsvorschläge geäußert:

- Die Zeit einzelner Einträge muss nicht in Dimensionstabellen festgehalten werden.

- Die „Meter Point“ Tabelle fällt weg. Sie ist für den Anwendungsfall nicht wichtig, da die relevanten Daten schon in der Verbindung mit der Netzregion liegen und die zyklische Abhängigkeit zu inkonsistenten Daten führen kann.
- Der „Meter Point“ taucht als Spalte bei der „Customer Site“ auf und wird POD bzw. Point of Delivery genannt.

### 3 MVC-Architektur

Der erste Entwurf der Architektur wird vorgestellt. Er soll noch einmal generischer und in UML-Notation umgesetzt werden. Dann soll in zwei Gruppen jeweils eine softwarespezifische Architektur mit SAP HANA und Open Source Software erstellt werden. Zu diesem Zweck wird am Donnerstag, den 05.02.2015 ein Architektur-Workshop durchgeführt.

### 4 Open Source

VoltDB wird als mögliche Alternative zu HANA betrachtet. Ohne weitere Maßnahmen ist es jedoch ungeeignet, da die Open Source Version keine Durability sicherstellt. HANA bietet diese Eigenschaft, die Trial-Version ist jedoch auf 1GB Speicher beschränkt. Es wird die Möglichkeit einer kombinierten Datenhaltung mit einem traditionellen MySQL+HANA System genannt. Auch eine vorübergehende parallele Entwicklung von 2 Prototypen ist denkbar, bei dem einer auf HANA und einer auf MySQLs In-Memory Technologie basiert. Es wird die Ausführung eines Scripts auf der HANA Cloudplattform demonstriert, mit dem Daten aus der Datenbank abgefragt werden und sowohl textuell als auch graphisch mit einem Diagramm dargestellt werden können. Die graphische Oberfläche des Clients kann z.B. mit den Javascript Bibliotheken SAP UI5 oder dojo erstellt werden. Bei der Verwendung von HANA mit UI5 ist zu beachten, dass die HTML-Seite über die gleiche Domain erreichbar ist, wie die Datenbankinstanz, die Javascript Bibliotheken dürfen auf einem anderen Server liegen. Die Frage, welche Sprache für die Controller-Logik eingesetzt werden soll, kommt auf. Java, ggf. mit Hibernate, wird als Möglichkeit genannt.

### 5 Website

Die Website ist über die Domain [www.cloud-bi.informatik.uni-oldenburg.de](http://www.cloud-bi.informatik.uni-oldenburg.de) erreichbar. Beim Aufrufen ändert sich die URL in [ems.informatik.uni-oldenburg.de](http://ems.informatik.uni-oldenburg.de). Dies kann ggf. noch geändert werden. Die Farbe des Designs muss noch auf das der Universität Oldenburg abgestimmt und eventuell das Uni-Logo eingefügt werden. Außerdem muss eventuell ein Logo für die PG erstellt werden. Dazu werden einige Entwürfe gezeigt und diskutiert.

### 6 Zwischenstandpräsentation

Die Zwischenstandpräsentation findet am 25.02.2015 um 14 Uhr statt. Der Raum steht noch nicht fest. Die erste Version der Folien soll zum 10.02.2015 fertig vorgestellt werden. Für die Präsentation müssen ein oder mehr Redner benannt werden.

## 7 Sonstiges

Am 17.03.2015 und am 24.03.2015 kann Oliver nicht am PG-Treffen teilnehmen. Das Treffen am 17.03.2015 soll deshalb auf den 16.03.2015 oder den 24.03.2015 verlegt werden. Marco Haas ist am 24.03.2015 eventuell anwesend. Bis zum Vormittag des 09.02.2015 soll eine erste Version des Pflichtenhefts fertig sein. Dieses soll die Ergebnisse des Architektur-Workshops vom 05.02.2015 berücksichtigen.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	10. Februar 2015
Uhrzeit	08:17-10:05 Uhr
Ort	Offis Raum U64
Protokollant	Ilias Schweer
Moderator	Florian Merkel
Anwesend	Björn Friedrich, Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	-
Nächster Protokollant	Alexander Zimak

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. Pflichtenheft
3. Architektur
4. Projektstand und weiteres Vorgehen (Meilensteinplan)
5. Zwischenstandspräsentation
6. Terminfindung für Bowling
7. Sonstiges

## 0 Weekly Scrum

Der Scrum-Master präsentiert die Übersicht der Aufgaben des Weekly Scrums:

- Letzte Woche:
  - MVC Konzept
  - Data Fact Table
  - Homepage
  - Open Source Lösungen

- Aktueller Stand:
  - Pflichtenheft 1.0
  - Zwei Architekturteams
  - Zwischenstandspräsentation (Agenda)
- Geplanter Stand:
  - Pflichtenheft 1.1
  - Zwischenstandspräsentation 1.0

## 1 Letztes Protokoll

Der angegebene Raum im letzten Protokoll ist nicht korrekt (U63 → U64). Ebenso ist der Name *Dojo* klein geschrieben, allerdings sollte er groß geschrieben sein.

## 2 Pflichtenheft

Die Gliederung des Pflichtenhefts wurde durchgegangen und diskutiert. Folgende Änderungen sollten durchgeführt werden:

- Die Bezeichnung Grenzkriterien im ersten Kapitel ist unklar und sollte zu Musskriterien geändert werden.
- Die Bezeichnung des vierten Kapitels sollte von Systemüberblick zu Systemkonzeption geändert werden.
- Der Unterpunkt Systemarchitektur sollte auf die Position 4.4 rücken und die Unterpunkte 4.4 und 4.3 entsprechend eine Position nach oben (dann 4.2 und 4.3).
- Kapitel sechs kann entfernt werden (Produkt-Leistungen).
- Unter Kapitel acht (Qualitätsanforderungen) sollten die Nichtfunktionalen Anforderungen aufgeführt werden.
- Kapitel neun sollte Testkonzept heißen.
- Das zehnte Kapitel kann entfernt werden.
- Das Glossar kann entfernt werden

Prinzipiell brauchen im Pflichtenheft (wie auch im Lastenheft) keine Kapitel bezüglich der Rollenverteilung, Meilenstein- sowie Terminpläne enthalten sein, da diese Inhalte letztendlich in einem allgemeinen Projektdokument zu finden sein sollen, welches parallel angefertigt werden muss.

### **3 Architektur**

Florian stellte der Gruppe das MVC-Architektur-Diagramm vor. Es wurde besprochen, dass die Verbindung von Model (Client) zum Controller führen soll und ebenso eine Verbindung von Modell (Client) zu Controller (Client) führen soll. Ebenso wurde beschlossen, dass die Produktionsanlagen weiterhin in der Konzeption eine Rolle spielen. Danach stellte Florian den Data Fact Table vor. Hierbei wurde besprochen, dass die bisherige sinnvollste Lösung für den zeitlichen Bezug ein einheitliches/gleichmäßiges Intervall von 15 Minuten darstellt. Beginn und Ende eines Zeitraums werden stets an dieses Intervall angeglichen. Für längere Zeiträume werden diese 15 Minuten Intervalle dann hochmultipliziert. Als Lösungsansatz wurden Data-Marts genannt. Die Gruppe soll sich diesbezüglich (insbesondere aus Performance-Gründen) weiter informieren und recherchieren. Dazu sollte sich ein Gruppenmitglied mit der Dissertation von Martin Tröschel mit dem Thema Aktive Einsatzplanung in virtuellen Kraftwerken beschäftigen. Weiter hält Oliver eine kurze Präsentation mit dem Thema Abfrageperformance. Die entsprechenden Folien lädt Oliver in Kürze ins Confluence. Dazu soll ein Gruppenmitglied Recherche bzg. Performance tätigen.

### **4 Projektstand und weiteres Vorgehen**

Alexander stellte eine Terminübersichtstabelle mit vergangenen und zukünftigen Iterationen vor. Ebenso präsentierte er einen Meilensteinplan als Gantt-Chart. Diesbezüglich wurde besprochen, dass die Testphase länger ausgelegt werden sollte und parallel zur Implementation laufen sollte. Allgemein sollte bei der Terminplanung das Scrum-Prinzip stärker einbezogen werden, also die einzelnen Phasen sollten mehr parallelisiert werden. Im Anschluss stellte Oliver den Meilensteinplan aus seiner damaligen Projektgruppe vor. An diesen können wir uns orientieren.

### **5 Zwischenstandspräsentation**

Die Gliederung der Zwischenstandspräsentation wurde diskutiert. Hierbei wurde beschlossen, dass keine Grundlagenfolien zum Thema Cloud und BI nötig sind. Das Wort Produkt sollte nicht verwendet werden. Ansonsten ist Oliver mit der Gliederung einverstanden. Die bisher ausgewählten Redner lauten Alexander, Björn und Ilias.

### **6 Terminfindung für Bowling**

Das soziale Event Bowling wurde auf den 18.02.2015 um 19 Uhr terminiert. Es findet statt im Bowling Center Oldenburg (BCO). Oliver wird diesbezüglich eine Bestätigung via E-Mail versenden.

### **7 Sonstiges**

Anstatt festgelegte Termine und Fristen immer im Protokoll suchen zu müssen, sollte in Confluence eine Terminübersicht, bzw. ein PG-Kalender angelegt werden. Ebenso sollte demnächst eine genau Terminabsprache für die festen Sitzungen im nächsten Semester stattfinden, damit möglichst die vollständige Gruppe stets anwesend sein kann.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	17. Februar 2015
Uhrzeit	08:15-09:45 Uhr
Ort	Offis Raum U63
Protokollant	Alexander Zimak
Moderator	Ilias Schweer
Anwesend	Björn Friedrich, Ilias Schweer, Brian Clark , Oliver Norkus, Babak Izadpanah, Alex- ander Zimak
Abwesend	Florian Merkel (entschuldigt)
Nächster Protokollant	Brian Clark

## Tagesordnung

1. Weekly Scrum
2. Letztes Protokoll
3. Pflichtenheft
4. Zwischenpräsentation
5. Architektur
6. Kurze Vorträge zu Abfrageperformance und Einsatzplanung
7. Meilensteinplan und Terminübersicht
8. Sonstiges



# 1 Weekly Scrum

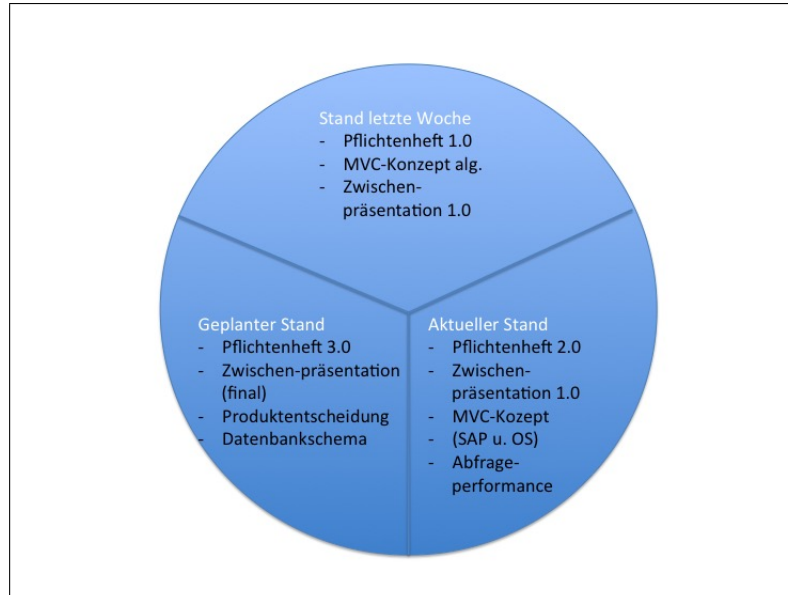


Abbildung 1: Weekly Scrum

## 2 Letzes Protokoll

Tagespunkt 4 (Projektstand) soll sprachlich noch einmal überarbeitet werden.

## 3 Pflichtenheft

Keine Anmerkungen seitens Oliver Norkus, da es noch keine bewertbare Version des Pflichtenheftes gibt.

## 4 Zwischenpräsentation

Der zum Termin aktuelle Stand der Zwischenpräsentation wurde besprochen und angepasst. Neben der Änderung der Gliederung, wurde eine höhere Fokussierung auf das eigene Projekt gefordert. Die Präsentation soll einem roten Pfaden folgen und die Übergänge sollen sinnvoll und in sich schlüssig gestaltet werden. Des Weiteren wurden inhaltliche Fehler sowie falsch gewählte Begrifflichkeiten identifiziert. Die neue Gliederung der Zwischenpräsentation sieht wie folgt aus:

1. Begrüßung
2. Problemstellung
3. Grundlagen
4. Projektidee

5. Projektanforderungen

6. Projektstand

## **5 Architektur**

Bei der Betrachtung der Systemarchitektur des Modells wurde festgestellt, dass die Komponente Controller um einen Query Optimizer erweitert werden sollte. Darüber hinaus sollten in zukünftigen Sprints die Kardinalitäten der einzelnen Komponenten zueinander bestimmt werden und eine detaillierte Beschreibung der einzelnen Komponenten erstellt werden. Zur Stärkung der Grundkenntnisse in der Softwarearchitektur soll bis zum 24.02.2015 das 3. Kapitel im Buch "Handbuch der Software-Architektur" gelesen werden.

## **6 Kurze Vorträge zu Abfrageperformance und Einsatzplanung**

Björn Friedrich hat einen Kurzvortrag über das Thema Anfrageperformance gehalten und dabei herausgestellt, dass nicht alle Parameter der Anfrageperformance für unseren Anwendungsfall notwendig sind. Der Vortrag von Florian Merkel bezüglich der Einsatzplanung wurde krankheitsbedingt auf den 24.02.2015 verschoben.

## **7 Meilensteinplan und Terminübersicht**

Alexander Zimak hat eine überarbeitete Version des Meilensteinplans vorgestellt, dieser muss noch hinsichtlich der Farbgebung und Benennung der einzelnen Phasen und Meilensteine überarbeitet werden.

## **8 Sonstiges**

- Confluence: Alle Dokumente und Modelle sollen zukünftig in das Confluence geladen werden unabhängig vom Grad des Zustandes.
- PG Kalender: Es liegt in der Verantwortlichkeit des Projektleiters alle relevanten Termine sowie die Urlaube der einzelnen Gruppenmitglieder im Kalender zu pflegen.
- Product Backlog: Die Aufwände aller offenen Aufgaben aus dem Product Backlog sollen geschätzt werden.
- Homepage: Brian Clark muss sein Profil im Confluence bis zum 24.02.2015 gepflegt haben.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	24. Februar 2015
Uhrzeit	08:18 - 9:28 Uhr
Ort	Offis Raum U64
Protokollant	Brian Clark
Moderator	Björn Friedrich
Anwesend	Björn Friedrich, Brian Clark, Oliver Nor- kus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	Ilias Schweer (verhindert)
Nächster Protokollant	Babak Izadpanah

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. Architektur
3. Data Fact Table
4. Zwischenpräsentation
5. Sonstiges

## 0 Weekly Scrum

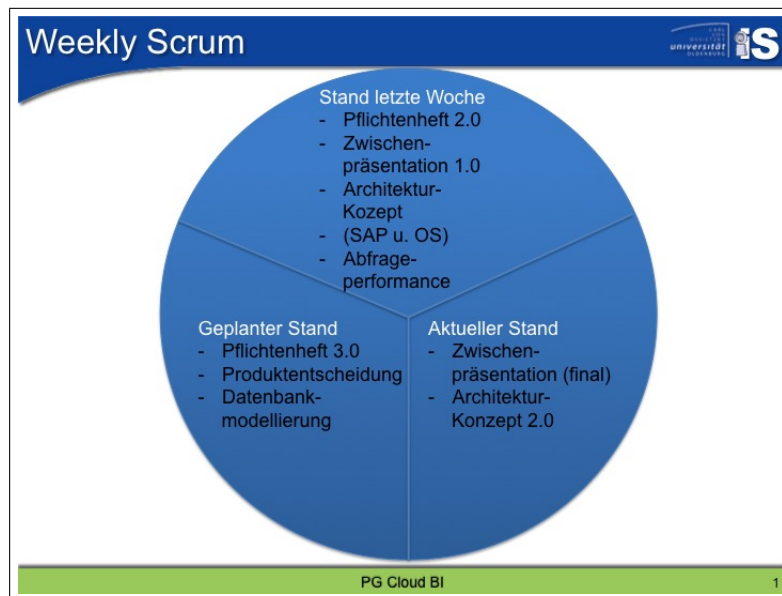


Abbildung 1: Weekly Scrum

## 1 Letztes Protokoll

Es gibt ein paar Rechtschreibfehler im letzten Protokoll.

## 2 Architektur

Babak hat den aktuellen Zustand des UML-Komponenten-Diagramms vorgestellt. Dazu gibt es folgende Empfehlungen:

- Eine *Accounting & Billing* Komponente mit Nutzungsmessungen fehlt.
- Die *Load Balancer* Komponente braucht mehr Verbindungen.
- Die Kardinalitäten für die Komponenten werden benötigt.

Bezüglich der *Accounting* Module muss entschieden werden, ob die Berechnung pro Zeit, Report, Speicher, oder CPU-Benutzung sein soll. Zusätzlich ist möglicherweise eine Cache-Funktion notwendig, allerdings kann dies später entschieden werden.

## 3 Data Fact Table

Florian hat den aktuellen Zustand des Data Fact Table gezeigt.

## 4 Zwischenpräsentation

Alex hat den aktuellen Zustand der Zwischenpräsentation gezeigt.

- Die Veränderungen des Probelaufs waren integriert.
- Die neue Data Fact Table muss noch integriert werden.

## 5 Sonstiges

- Es ist angebracht, ein Raum zu finden, der für die Semesterferien regelmäßig genutzt werden kann.
- Wir sollen einen Antrag für das Hasso Plattner Institute Future SOC Lab bis zum 27. März 2015 stellen.
- Auf der Website fehlt noch Brians Profil.
- Wir brauchen noch ein Sequenzdiagramm für den Architekturplan.

# Protokoll

Projektgruppe  
BlaaS

Sitzungsdatum	05. März 2015
Uhrzeit	10:05-11:30 Uhr
Ort	Offis Raum U63
Protokollant	Florian Merkel
Moderator	Alexander Zimak
Anwesend	Ilias Schweer, Oliver Norkus, Florian Merkel, Alexander Zimak
Abwesend	Babak Izadpanah (2 Min. Verspätung), Björn Friedrich (10 Min. Verspätung), Brian Clark (20 Min. Verspätung, entschuldigt)
Nächster Protokollant	Babak Izadpanah

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. Pflichtenheft
3. Architektur
4. Product Backlog
5. HPI Future SOC Lab
6. Sonstiges

## 0 Weekly Scrum

Der Scrum-Master präsentiert die Übersicht der Aufgaben des Weekly Scrums:

- Letzte Woche:
  - Zwischenpräsentation
  - Architektur
- Aktueller Stand:

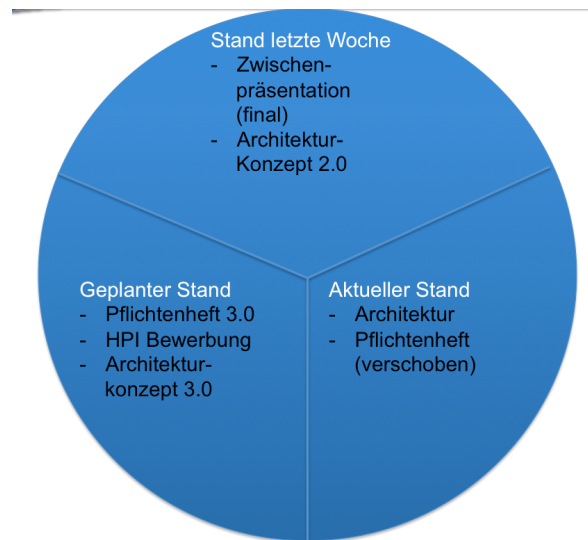


Abbildung 1: Weekly Scrum

- Pflichtenheft ist nicht auf dem geplanten Stand
- Architektur wurde und wird weiterentwickelt
- Geplanter Stand:
  - HPI Bewerbung
  - Architekturkonzept 3.0

## 1 Letztes Protokoll

- Für das weekly Scrum soll nicht die Folie sondern die Abbildung auf der Folie ins Protokoll.
- Der Punkt zu „Load Balancer“ bei Abschnitt 2 (Architektur) ist unklar und muss überarbeitet werden.
- Bei Abschnitt 5 (Sonstiges) muss der korrekte Name des Hasso Plattner Institute und die korrekte Präposition beim Datum (bis) verwendet werden.

## 2 Pflichtenheft

Das Pflichtenheft befindet sich nicht in einem präsentierbaren Zustand und bedarf noch viel Arbeit. Bis zum 20.03.2015 soll es diesen Stand erreichen.

## 3 Architektur

Inhaltlich hat die statische Sicht auf die Architektur einen guten Stand erreicht. Die Kardinalitäten der einzelnen Komponenten werden besprochen, wobei folgende Anmerkungen gemacht werden.

- Das User Management muss ggf. skalierbar sein, sollte also [1..\*] mal vorkommen können.
- Es sollte möglich sein, mehrere Datenbanksysteme (DBS) im System einzubinden, also [1..\*], weil manche Daten eventuell nur auf bestimmten DBS liegen dürfen.
- Es wird ein Source Monitor pro Quelle benötigt, also können [1..\*] im System existieren.
- Ebenso werden mehrere ETL Prozessoren vorkommen: [1..\*].

Zur Frage der möglichen Umsetzung gibt es folgende Anmerkungen.

- Das User Management muss ggf. nicht implementiert, aber vorgesehen werden. Die Verwendung von OpenID ist ebenfalls denkbar.
- Die Hardwarevirtualisierung beim Load Balancing kann ignoriert werden, da eine eigene Implementierung keinen Mehrwert im Vergleich zu bestehenden Systemen bietet.
- Das DBS inklusive ETL-Prozessor und Source Monitor kann als Cloud Dienst (HANA-Instanz) in Anspruch genommen werden.
- Das Hauptaugenmerk bei der Implementierung liegt also auf:
  - Local Client
  - Data Retrieval
  - Datenbanklogik
  - Load Balancing für die Verteilung von Clients auf Serverinstanzen

Während das Komponentendiagramm größtenteils fertig ist, muss die Interaktion (Sequenzdiagramme), die Einteilung in Pakete (Paketdiagramm) und die Verwendung von Hardwarekomponenten bzw. Systemumgebungen (Deploymentdiagramm) noch beschrieben werden. Das Paketdiagramm muss nicht unbedingt UML sein. Darüber hinaus muss noch einmal überdacht werden, welches Pay-per-Use Modell zum Einsatz kommen soll und wie die Umsetzung aussehen kann. Die Architektur soll am 10.05.15 geschehen in der Sitzung präsentiert werden. In der Präsentation soll die Richtung der Schnittstellen im Komponentendiagramm noch einmal ausführlich erklärt werden.

## 4 Product Backlog

Das Product Backlog wird vorgestellt. Es fällt auf, dass es noch viele überfällige Aufgaben gibt. Weitere Aufgaben werden während der Sitzung ins Backlog eingetragen. Diese umfassen Aufgaben, die die Architektur betreffen, und Aufgaben, die die Bewerbung beim HPI Future SOC Lab betreffen.

## 5 HPI Future SOC Lab

Bis zur Sitzung am 10.03.15 wird Brian eine Schätzung für den Ressourcenbedarf fertigstellen. Zu diesem Zweck wird er sich mit Marco Haas in Verbindung setzen.



## 6 Sonstiges

- Falls etwas nicht fertig wird muss **frühzeitig** Bescheid gegeben werden.
- Der Beginn von Terminen ist immer s.t.
- Das Wiki im Confluence sollte gepflegt werden. Alexander übernimmt diese Aufgabe als Wiki-Beauftragter. Dokumente werden auf einer Seite hochgeladen und dann von Alexander an die richtige Stelle verschoben.
- Die Homepage sollte noch weiter überarbeitet werden.
  - Das Layout ist suboptimal (links zu breit, Text zu klein)
  - Anstatt „Meilenstein“: „Zwischenpräsentation
  - Es sollten mehr Informationen über das Projekt vorhanden sein
  - Eine Roadmap / ein Projektplan sollten vorhanden sein.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	10. März 2015
Uhrzeit	8:15-9:45 Uhr
Ort	Offis Raum U63
Protokollant	Babak Izadpanah
Moderator	Florian Merkel
Anwesend	Ilias Schweer, Brian Clark, Björn Friedrich, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Nächster Protokollant	Brian Clark

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. Architektur
3. Product Backlog
4. HPI Future SOC Lab
5. Sonstiges

## 0 Weekly Scrum

In diesem Abschnitt des Protokoll soll nicht mehr das Tortendiagramm gezeigt werden, sondern die Tabelle von Alexander. Die Tabelle ist detaillierter und übersichtlicher als das Diagramm. In der Tabelle steht, an welchen Tasks letzte Woche gearbeitet wurde und welche Tasks für die kommende Woche geplant werden.

## 1 Letztes Protokoll

Das letzte Protokoll ist noch nicht im Confluence hochgeladen worden. Deshalb gibt es seitens Oliver noch keine Anmerkungen dazu. Das Hochladen muss so schnell wie möglich nachgeholt werden.

## 2 Architektur

Es sind die folgenden Diagramme vorgestellt worden:

1. Paketdiagramm
2. Komponentendiagramm

Es lagen noch keine Sequenzdiagramme vor. Deshalb wurde besprochen, welche Abläufe als Sequenzdiagramme abzubilden sind. Zu den einzelnen Diagrammen sind folgende Bemerkungen gemacht worden:

### **Paketdiagramm:**

Die einzelnen Komponenten wurden in Pakete eingeteilt. Dabei wurde überlegt, welche Komponenten semantisch zusammengehören. Es muss festgelegt werden welche Komponenten in der Cloud und welche lokal auf Endgeräten.

### **Komponentendiagramm:**

- Dem Komponentendiagramm sind neue Komponenten hinzugefügt worden und die wurden Kardinalitäten angepasst.
- Das Load Balancer Interface wurde umgedreht. Der Load Balancer kann auch als Proxy benutzt werden.
- Dispatcher: Leitet Anfragen an die entsprechenden Komponenten weiter
- User Manager: Bietet die Funktionen neuen Benutzer anlegen, Kennwort ändern und Benutzer löschen.
- ETL Manager: Benutzer können neue ETL Prozesse erstellen
- Die User Manager Komponenten wurden auseinander gezogen, weil alle Unterkomponenten verschiedene Kardinalitäten haben.
- Die User Manager Komponente soll auch in der Cloud liegen.

**Sequenzdiagramm:** Es liegen noch keine Diagramme vor. Aber es wurde eine Liste von benötigten Sequenzdiagrammen vorbereitet. Damit kann entschieden werden, welche Funktionen für das Projekt wichtig sind und auf welche sich die Projektgruppe konzentriert.

- Anlegen eines neuen Benutzers
- Hinzufügen einer neuen Quelle
- ETL-Prozess wird durch einen Benutzer gestartet
- ETL-Prozess wird durch den Source Monitor gestartet
- Benutzer erstellt eine OLAP Anfrage ( Filter, Daten aktualisieren, Drill-Down, Rotate ect.)
- Ein neuer Backend Server wird gestartet

### 3 Product Backlog

Sprint Backlog					
Task	Priorität	Fälligkeiten	Status	Verantwortlicher	
Sequenzdiagramme anpassen	A	17.	Märactive	Björn	
Komponentendiagramm überarbeiten	A	17.	Märactive	Babak	
HPI Bewerbung	A	17.	Märactive	Brian	
In Scope/Out of Scope Definition	A	17.	Märactive	Florian	
Architekturübersichtsdiagramm	A	17.	Märactive	Florian	
Pflichtenheft Kapitel überarbeiten	A	20.	Märactive	Ilias	

Abbildung 1: Backlog

Das Product Backlog ist überarbeitet worden. Alle neuen Aufgaben wurden eingetragen und Mitgliedern der PG zugewiesen.

### 4 HPI Future SOC Lab

Für die Bewerbung hat Brian eine Kalkulation der benötigten Hardwareressourcen durchgeführt. Am nächsten Donnerstag (12.03.2015) wird er eine überarbeitete Version der Tabelle präsentieren.

### 5 Sonstiges

- Björn plant ein Auslandssemester in Japan zu machen. Deswegen kann er bei der Abschlusspräsentation nicht dabei sein. Während des Auslandssemester wird er fleißig remote arbeiten und seine Aufgaben weiterhin wahrnehmen. In der Implementierungs- und Dokumentationsphase lässt sich das bewerkstelligen. Für Gespräche steht u.a. auch Skype zur Verfügung.
- Ilias wird Alexander als Projektleiter während dessen Urlaub vertreten.
- Am nächste Dienstag (17.03.2015) kann Oliver nicht bei der PG-Sitzung sein dabei. Der Termin kann auf Freitag (20.03.2015) verschoben werden oder ohne Oliver stattfinden.
- Die Sitzung für die Überarbeitung der Architektur findet am Donnerstag von 13 bis 16 Uhr in U61 statt.

# Protokoll Architekturtreffen

Projektgruppe  
BIaaS

Sitzungsdatum	12. März 2015
Uhrzeit	13:05 - 15:45 Uhr
Ort	Offis Raum U61
Protokollant	Ilias Schwerer
Moderator	-
Anwesend	Björn Friedrich, Brian Clark, Oliver Nor- kus, Babak Izadpanah, Florian Merkel, Ili- as Schwerer
Abwesend	Alexander Zimak (beurlaubt)
Nächster Protokollant	-

## Tagesordnung

1. Komponentendiagramm (mit Cloudmarkierungen)
2. In Scope und Out of Scope
3. Swim Lanes
4. UML für das Pflichtenheft
5. HPI Bewerbung

## 1 Komponentendiagramm (mit Cloudmarkierungen)

Florian präsentiert das Komponentendiagramm mit den jeweiligen Cloudmarkierungen. Der mittlere Teil soll noch mit in die Cloud gefügt werden. HANA Datenbank könnte komplett alleine für sich stehen. ETL-Prozessor und Manager zusammen in einer Cloud. Apache JCloud bietet eine geeignete Schnittstelle zwischen Datenbank und ETL-Prozessor. Das Komponenten-diagramm mit den Cloudmarkierungen sollte etwas allgemeiner/abstrahierter sein. Gegliedert in Software/Platform/Infrastructure.

## 2 In Scope und Out of Scope

In Scope:

- User-Datenbank
- Accounting-Komponente (Monitoring der Nutzung)

- Data Retrieval (Query Optimizer und Generator)
- Dispatcher
- Load Balancing
- Resource Management
- Resource Monitoring
- Local Client Komponente

Out of Scope:

- User Management Komponente (siehe unten)
- ETL-Prozessor und Manage (siehe unten)

Es sollte eher eine statische Nutzerkomponente genutzt werden, ohne Registrieren, Passwort vergessen, usw.. Wir gehen davon aus, dass Daten im DWH liegen, der ETL-Prozess wird erstmal außer Acht gelassen. Pay-per-Use Art sollte zunächst nur nach Anzahl an Anfragen definiert sein, um dann später optional Verfeinerungen (Zeit, Ressourcen, etc.) vorzunehmen.

### 3 Swim Lanes

Komponente	Iteration 1	Iteration 2	Iteration 3	Iteration 4
User Management	statische Logins, ohne Registrierung	Protokollierung des Aufrufs, Acc.-DB	Rechnung für einen Zeitraum, Zugriff * Kosten	-
Data Retrieval	Dummy QO, statische Anfragen an DB	Daten für Graph und Report übertragen	Dyn. Query	-
Local Client	Click Prototyp, Interface definieren	Daten von DR anzeigen, Login anbinden	Mehr Daten anzeigen, Design anpassen	Dyn. Query
ETL	Testdaten in DB laden, Schema anlegen, DB aufsetzen	-	-	-
Disp. L.M.	Webserver konfigurieren, (SVN-Script für www-Root)	Einfache Anfrageverarbeitung	-	-

### 4 UML für das Pflichtenheft

- Komponentendiagramm —> Ilias
- Wolkendiagramm —> Florian
- Packagediagramm —> Florian
- Deploymentdiagramm —> Florian

- Sequenzdiagramme (vom Standardprozess, wie in den EPKs) —> Björn
- Kommunikationsdiagramm —> Björn

## **5 HPI Bewerbung**

Zunächst soll der von der Gruppe bearbeitete Antrag an Oliver geschickt werden. Danach sollen Olivers Anmerkungen umgesetzt werden und etwaige Fragen/Probleme telefonisch mit dem HPI geklärt werden.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	31. März 2015
Uhrzeit	08:21 - 10:15 Uhr
Ort	Offis Raum U64
Protokollant	Ilias Schweer
Moderator	Babak Izadpanah
Anwesend	Björn Friedrich, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak, Ilias Schweer
Abwesend	Brian Clark (Urlaub)
Nächster Protokollant	Alexander Zimak

## Tagesordnung

0. Product Backlog
1. Letztes Protokoll
2. Pflichtenheft
3. HPI Bewerbung
4. Entwicklungsstand und -planung (Swim Lanes)
5. Auslandssemester von Björn
6. Sonstiges



## 0 Product Backlog

Sprint Backlog				
Task	Priorität	Fälligkeiten	Status	Verantwortlicher
Datenbankschema und Testdaten anlegen	A	31. März	active	Ilias
Dummy QO einrichten	A	31. März	active	Björn
Webserver Konfigurieren	A	31. März	active	Florian
statische Logins ohne Registrierung	A	31. März	active	Babak
klickbarer Prototyp, Interface definiton	A	31. März	active	Alex

Abbildung 1: Product Backlog

## 1 Letztes Protokoll

In den Protokollen und auch allgemein sollten keine Marken-, bzw. Produktnamen verwendet werden. Ebenfalls gilt es ein paar Rechtschreib- sowie Inhaltsfehler zu korrigieren (falsches Datum, ein Wort zu viel und eine fehlerhafte Abbildung).

## 2 Pflichtenheft

Oliver hat folgende Anmerkungen zur ersten Version des Pflichtenhefts:

- In Tabelle 1.1 muss bei 004F die Wortwahl (berechnen/berichten) überdacht werden.
- 008F ist in der Formulierung überflüssig und sollte ein Muss-Kriterium sein.
- 009F sollte ebenfalls ein Muss-Kriterium sein.
- Bei 009N muss die Formulierung überarbeitet werden (Mehrmandantenfähigkeit nicht unbedingt durch simultane Nutzung, eher Skalierbarkeit).
- 010N sowie 011N sind eher Muss-Kriterien.
- In Kapitel zwei sollte der Anwendungsbereich detaillierter beschrieben sein.
- Das vierte Kapitel soll allgemein präziser und umfangreicher formuliert sein. In-Scope und Out-of-Scope sollten genauer erklärt sein. Hierbei liegt der Fokus auf Skalierbarkeit und Flexibilität. Die Reihenfolge der Abschnitte sollte auch geändert werden, sodass erst der allgemeine Überblick aufgeführt ist und dann das Architekturdiagramm folgt. Ebenso sollten die Komponenten genauer beschrieben werden.
- Allgemein gibt es einige sprachliche und inhaltliche Mängel, jedoch ist der Gesamteindruck positiv. Das von Oliver mit Anmerkungen versehene Exemplar des Pflichtenhefts wird noch eingescannt und veröffentlicht.

## 3 HPI Bewerbung

Die von der Gruppe angefertigte Bewerbung für das HPI Future SOC Lab wurde fristgerecht eingereicht.

## 4 Entwicklungsstand und -planung (Swim Lanes)

- Babak hat sich mit dem User Management beschäftigt und stellte diesbezüglich eine kurze Präsentation seiner Einschätzung der Möglichkeiten vor. Hierbei wurden allerdings einige Aspekte, wie das Session Management sowie der Zusammenhang zwischen User und Rollen (Rechte) vernachlässigt.
- Florian hat sich mit der Dispatcher Komponente sowie dem Load Management beschäftigt und stellte ebenfalls seine Einschätzung der Möglichkeiten in Form einer kurzen Präsentation vor.
- Alexander hat einen klickbaren Mockup angefertigt und stellte diesen vor. Hierbei ergab sich eine Frage an Marco: Wie sollen die einzelnen Preise von verschiedenen Netzregionen visualisiert werden?
- Björn hat weiter an der Homepage gearbeitet.
- Ilias hat sich mit der Datenbank und dem ETL-Prozess beschäftigt.

## 5 Auslandssemester Björn

Björn beginnt sein Auslandssemester am 01. August 2015 und wird seinen Urlaub in den letzten vier Wochen der Projektphase nehmen.

## 6 Sonstiges

- Die Gruppe nimmt kollektiv vom 06. April bis einschließlich 12. April Urlaub (Björn jedoch nicht).
- Das nächste Treffen findet am 14.04 statt.
- Alexander wird eine Doodle-Umfrage bezüglich der Terminfindung für das nun kommende Semester erstellen.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	16. April 2015
Uhrzeit	12:15-13:50 Uhr
Ort	Offis Raum U64
Protokollant	Alexander Zimak
Moderator	Björn Friedrich
Anwesend	Björn Friedrich, Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	-
Nächster Protokollant	Brian Clark

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. Aktueller Stand der Swim Lanes
3. Aktueller Stand des Pflichtenheftes
4. Product Backlog
5. Sonstiges

## 0 Weekly Scrum

Der Scrum-Master präsentiert die Übersicht der Aufgaben des Weekly Scrums (siehe Abbildung 1).

## 1 Letztes Protokoll

Es müssen einige Änderungen am alten Protokoll vorgenommen werden. In Punkt 2 (Pflichtenheft) soll im 2. Unterpunkt der Satz von „ist die Formulierung überflüssig“ in „sollte umformuliert werden“ umformuliert werden. Im Unterpunkt 4 sollte der Satz „Mehrmandantenfähigkeit nicht unbedingt durch simultane Nutzung, eher Skalierbarkeit“ umformuliert werden, zum Beispiel „Der Begriff Skalierbarkeit findet hier eher Anwendung, da Mehrmandantenfähigkeit nicht die simultane Nutzung des Systems beschreibt“ . In Unterpunkt 6 soll die Bezeichnung „Kapitel zwei“ durch den genauen Namen des Kapitels ersetzt werden.

Sprint Backlog				
Task	Priorität	Fälligkeiten	Status	Verantwortlicher
Datenbankschema und Testdaten anlegen	A	14. Apr	active	Ilias
Dummy QO einrichten	A	14. Apr	active	Björn
Webserver Konfigurieren	A	14. Apr	active	Florian
statische Logins ohne Registrierung	A	14. Apr	done	Babak
klickbarer Prototyp, Interface definiton	A	14. Apr	active	Alex
Bericht Swim Lane ETL	A	14. Apr	done	Ilias
Bericht Swim Lane Local Client	A	14. Apr	done	Alex
Bericht Swim Lane Data Retrieval	A	14. Apr	done	Björn
Bericht Swim Lane Load Manager/Dispatcher	A	14. Apr	done	Florian
Bericht Swim Lane User Mgt.	A	14. Apr	done	Babak

Abbildung 1: Weekly Scrum

## 2 aktueller Stand der Swim Lanes

- **User Management:** Es wurden zwei auf unterschiedlichen Programmiersprachen basierende Prototypen vorgestellt. Der erste Prototyp wurde mit Java implementiert und besitzt noch kein Session-Tracking. Die zweite Prototyp wurde mit php Implementiert und besitzt bereits ein Session-Tracking. Der Prototyp auf Basis von php wird von Babak Izadpanah präferiert, da Babak mehr Kenntnisse im php-Umfeld besitzt. Im nächsten Sprint wird sich Babak mit dem Tool OpenID beschäftigen und es auf seine Tauglichkeit prüfen.
- **Dispatcher und Load Manager:** Erster Prototyp auf Basis von Ruby on Rails wurde vorgestellt. Im nächsten Sprint werden weitere Technologien getestet und mit dem Local Client-Verantwortlichen eine geeignete Technologie ermittelt.
- **Data Retrieval:** Björn Friedrich hat sich mit dem Data Retrieval beschäftigt und als mögliche Programmiersprache für die Implementation Objective-C vorgeschlagen. Darüber hinaus hat sich Björn mit benötigten Schnittstellen beschäftigt, diese wären Schnittstellen zum Dispatcher, zum Accounting, zum A&R Management und zur Data Base. Für weitere Schritte muss hier noch auf eine Technologieentscheidung gewartet werden.
- **ETL:** Ilias Schwer, welcher die Swim Lane ETL bearbeitet, hat die bereits erstellten Data Fact Table in Excel Tabellen überführt und so eine erste Datenbankstruktur erstellt. Im nächsten Schritt werden diese Tabellen mit Testdaten gefüllt, um so bereits erste Dummy Daten für das BICE System zu speichern. Auch hier muss noch auf die Technologieentscheidung gewartet werden.
- **Local Client:** Es gibt bereits einen klickbaren Prototypen, welcher im kommenden Schritt in einen HTML5 Prototypen überführt wird. Darüber hinaus wurden alle nötigen Benutzer, Funktionen und Inhalte der Benutzer GUI dokumentiert.

## 3 Stand des Pflichtenheftes

Björn Friedrich hat sich während des PG-Urlaubs mit der Überarbeitung des Pflichtenheftes beschäftigt. Änderung im Vergleich zum alten Exemplar sind die genaue Beschreibung der

Direktvermarktung und das Überarbeiten des Schnittstellen und Produktfunktionen Kapitels. Bezüglich des Testkonzepts hat sich Björn bereits erste Gedanken über mögliche Testszenarien gemacht, zu diesen zählen die Performance und Usabilitytests. Darüber hinaus hat Björn sich bei Marco Haas über den Stromhandel und die sich daraus für uns ergebenden Anforderungen erkundigt.

## 4 Product Backlog

Product Backlog						
Aufgaben	Unteraufgaben	Tasks	Prioritäten	Fälligkeiten	Status	
Pflichtenheft	Zielbestimmung	Kapitel überarbeiten	B	03. März	open	
	Produkteinsatz	Produktel überarbeiten	B	03. März	open	
		Produktkonfiguration	Hard und Softwarekomponenten festlegen	B	24. Feb	open
	Systemüberblick	Kapitel überarbeiten	Kapitel verfassen	B	27. Feb	open
			Kapitel überarbeiten	B	03. März	open
		Kapitel verfassen	A	17. Feb	active	
	Produktfunktionen	Kapitel überarbeiten	B	03. März	open	
		Kapitel verfassen	A	17. Feb	active	
	Benutzerschnittstellen	Kapitel überarbeiten	B	03. März	open	
		Qualitäts-Zielbestimmung	Kapitel überarbeiten	B	03. März	open
		Testkonzept	Kapitel überarbeiten	B	03. März	open
	Implementation	Infrastruktur		C		open
		Plattform		C		open
Software			C		open	
Testphase	Testfallimplementation		C		open	
	Prozess für wöchentliche Tests		C		open	
Dokumentation	Glossar bearbeiten		C		open	
Administration						
Swim Lanes	ETL	Datenbankschema und Testdaten anlegen	A	23. Apr	active	
	Data Retrieval	Dummy QO einrichten	A	23. Apr	active	
	Load Manager, Dispatcher	Webserver Konfigurieren	A	23. Apr	active	
	User Management	statische Logins ohne Registrierung	A	23. Apr	active	
	Local Client	klickbarer Prototyp, Interface definiton	A	23. Apr	active	
	Local Client	Html 5 GUI Prototyp	A	23. Apr	active	
	User Management	Open ID testen	A	23. Apr	active	
	User Management	Session Managemnet in Java	A	23. Apr	active	
	Load Manager, Dispatcher	Framework Tests	A	23. Apr	active	
	ETL	Test Daten erstellen	A	23. Apr	active	
	ETL	MySql Datenbank aufsetzen	A	23. Apr	active	
	Data Retrieval	Schnittstellen definition	A	23. Apr	active	
		Jetty Server aufsetzen	A	23. Apr	active	

Abbildung 2: Product Backlog

## 5 Sonstiges

- Oliver hat der PG mitgeteilt, dass der verfasste Artikel für das ERP-Magazin angenommen wurde und in der nächsten Ausgabe gedruckt werden wird. Diesbezüglich soll Björn den Artikel mit in den Inhalt der Homepage aufnehmen.
- Oliver hat darüber hinaus vorgeschlagen, einen weiteren Artikel zu verfassen. An diesem werden Brian, Florian, Björn und Oliver gemeinsam arbeiten.
- Als Team Event wurde ein gemeinsames Grillen vorgeschlagen, für welches eine Doodle Umfrage erstellt werden soll.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	23. April 2015
Uhrzeit	12:15-13:50 Uhr
Ort	Offis Raum U64
Protokollant	Brian Clark
Moderator	Alexander Zimak
Anwesend	Björn Friedrich, Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	-
Nächster Protokollant	Björn Friedrich

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. Aktueller Stand der Swim Lanes
3. HPI Future SOC Lab
4. Artikel
5. Aktueller Stand des Pflichtenheftes
6. Sonstiges

## 0 Weekly Scrum

Mitglied	aktueller Sprint	letzter Sprint
Babak	Java Session implementieren (Probe) und dafür PHP benutzen	XAMPP aufsetzen und damit OpenID testen mit PHP und Java
Björn	Artikel für den Workshop	Dummy QO und erste Arbeiten am Data Retrieval
Alexander	Prototypen in HTML5 überführen	Prototypen in HTML5 überführen
Florian	Am Artikel arbeiten, Pflichtenheft bearbeiten	Web application frameworks evaluieren
Brian	Java Server aufsetzen, bei GUI helfen	GUI Prototype, OpenID, Java Server erstellen
Ilias	DB-Tabellen mit Testdaten füllen, Protokollvorlage anpassen, Pflichtenheftkapitel überarbeiten, QS	MySQL Server aufgesetzt, Client konfiguriert, Tabellen erstellt, Fremdschlüssel (Constraints) eingefügt

## 1 Letztes Protokoll

Es sind einige Rechtschreibfehler enthalten. Die Qualitätssicherung muss präziser und konsequenter durchgeführt werden.

## 2 Aktueller Stand der Swim Lanes

Für den aktuellen Stand der Swim Lanes siehe Weekly Scrum Tabelle.

## 3 HPI Future SOC Lab

Eine Liste der Gruppenmitglieder (also der Benutzer) muss für die HPI Future SOC Lab erstellt und zugestellt werden, um den Zugang zu der HANA-Plattform zu erhalten. Am Ende des Projekts benötigt HPI einen Report mit unseren Ergebnissen und zusätzlich ein übersichtliches Plakat, das unsere Projektergebnisse zusammenfassend visualisiert.

## 4 Artikel

Es wurde beschlossen, einen weiteren Artikel für die Konferenz von Prof. Dr. Lehnhoff zu verfassen. Der Artikel kann danach, sollte er gut genug sein, für das Magazin ERICS vorgesehen werden. Die Abgabefrist ist der 4.5.2015.

## 5 Stand des Pflichtenheftes

Das Pflichtenheft wurde teilweise von einigen Mitgliedern weiter bearbeitet. Als nächstes werden konkret Olivers Anmerkungen zum Pflichtenheft in der Version 1.0 umgesetzt.

## 6 Sonstiges

- Das Protokoll soll bis spätestens Montags um 9 Uhr in Confluence sein.



# Protokoll

Projektgruppe  
BlaaS

Datum	7.05.2015
Uhrzeit	12:15h - 13:30h
Protokollant	Babak Izadpanah
Moderator	Björn Friedrich
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak
Abwesend	-
Nächster Protokollant	Florian Merkel

## Tagesordnung

0. Weekly Scrum
1. Letztes Protokoll
2. HPI
3. Scrum
4. Product Backlog
5. Sonstige

## 0 Weekly Scrum

Mitglied	aktueller Sprint	letzter Sprint
Babak	Java Session implementieren (Probe) und dafür PHP benutzen	XAMPP aufsetzen und damit OpenID testen mit PHP und Java
Björn	Artikel für den Workshop	Dummy QO und erste Arbeiten am Data Retrieval
Alexander	Prototypen in HTML5 überführen	Prototypen in HTML5 überführen
Florian	Am Artikel arbeiten, Pflichtenheft bearbeiten	Web application frameworks evaluieren
Brian	Java Server aufsetzen, bei GUI helfen	GUI Prototype, OpenID, Java Server erstellen
Ilias	DB-Tabellen mit Testdaten füllen, Protokollvorlage anpassen, Pflichtenheftkapitel überarbeiten, QS	MySQL Server aufgesetzt, Client konfiguriert, Tabellen erstellt, Fremdschlüssel (Constraints) eingefügt

## 1 Letztes Protokoll

Bei dem Protokoll für 30.04.2015 müssen noch Rechtschreibfehler korrigiert werden. Außerdem müssen bei den Swim Lanes die Aktivitäten im Protokoll detaillierter beschrieben werden.

## 2 HPI

Die E-Mail mit den Zugangsdaten für die HANA Instanz ist nur an Oliver gegangen. Er hat sie an alle Mitglieder der PG weitergeleitet.

## 3 Scrum

Björn hat den Scrum-Vortrag von Jan Christian Halfbrodt zusammengefasst und kurz erläutert, was für die PG noch sinnvoll wäre. Scrum eignet sich bei Projekten, die länger als ein halbes Jahr dauern. Denn bis sich das ganze Team an den Prozess gewöhnt hat und die Prozesse etabliert sind, ist einige Zeit notwendig. Für die Teammitglieder bedeutet Scrum ein entspannteres Arbeiten während der Projekte und führt zu einer höheren Qualität der Produkte. Scrum ist nur die Retrospektive und das Sprint Review ein selbsteilender Prozess. Die Projektgruppe sollte unbedingt die Retrospektive einführen. Die Retrospektive findet ohne den Product Owner statt und kann deshalb bei den internen PG-Treffen stattfinden. In der Retrospektive werden nicht nur fachliche Dinge besprochen, sondern auch Aspekte bzgl. der Meta-Kommunikation.

Weiterhin könnte das Schätzen des Aufwands für Aufgaben ausgeprägter sein. Allerdings ist noch nicht klar, wie das umgesetzt werden könnte.

## 4 Product Backlog

### Babak-Ilias:

Babak und Ilias arbeiten zusammen, um Ilias' Testdaten in unsere HANA Instanz zu laden

und eine rudimentäre Login-Funktion zu programmieren.

**Brian:**

Brian hilft bei der Implementierung der GUI.

**Alex:**

Alex arbeitet weiter an der GUI.

**Björn:**

Björn bereitet den Performancetest vor.

**Florian:**

Florian kümmert sich weiterhin um Serverangelegenheiten und dem Client.

## 5 Sonstige

- Am 14.05.2015 haben wir keine Sitzung. Die nächste Sitzung ist am 21.05.15
- Am 27.05. um 17 Uhr planen wir Grillen und Volleyball spielen. In Wechloy spielen wir Volleyball und grillen danach bei Alex und Ilias. Wir schreiben ein Liste der Sachen, die mitgebracht werden sollen.
- Bouleturnier für die Projekgruppen findet wahrscheinlich Mitte Juni statt. Sobald genauere Informationen verfügbar und die Anmeldung möglich ist, kümmert sich Alex darum.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	21. Mai 2015
Uhrzeit	10:05-11:30 Uhr
Ort	Offis Raum U64
Protokollant	Florian Merkel
Moderator	Alexander Zimak
Anwesend	Ilias Schweer, Oliver Norkus, Florian Merkel, Alexander Zimak, Björn Friedrich, Brian Clark, Babak Izadpanah
Abwesend	-
Nächster Protokollant	Babak Izadpanah

## Tagesordnung

0. Letztes Protokoll
1. Weekly Scrum
2. Aktueller Stand der Swimlanes
3. Artikel
4. Sonstiges

## 0 Letztes Protokoll

Das letzte Protokoll kann nicht eingesehen werden, da Confluence nicht erreichbar ist. Das Protokoll wird nächste Woche besprochen.

## 1 Aktueller Stand der Swimlanes

Folgende Tätigkeiten wurden ausgeführt:

### Data Retrieval (Ilias & Björn)

- Das Datenbankschema zur Speicherung der Energiedaten wurde aufgesetzt (lokal & HANA).
- Eine Java-Anwendung zur Durchführung von Performancetests auf der lokalen Datenbank und der HANA Datenbank wurde entwickelt.

Ergebnis der Performancetests: mit den zur Zeit vorhandenen Datenmengen ist die Netzwerklatenz die einzige bestimmende Größe.

### **Client (Alexander & Brian)**

- Die Benutzeroberfläche wurde erweitert.
- Verschiedene Diagrammdarstellungen können ausgewählt werden.
- Filter sollen über einen Popup Dialog einstellbar sein.
- Es existieren Tabs zum Auswählen der Funktionen „Deckungsbeitragsrechnung“ und „Gesamtkostenberechnung“

Die Datenbankanbindung muss noch umgesetzt werden.

### **User Management (Babak)**

- Eine Tabelle zur Speicherung der Benutzerdaten wurde in der HANA Datenbank angelegt.
- Ein beispielhafter Loginvorgang mit der Benutzerdatenbank über ein Webinterface wurde umgesetzt.

Die Benutzertabelle muss noch mindestens um eine Spalte ergänzt werden, über die ein Benutzer einem Mandanten zugeordnet wird.

### **Dispatcher (Florian)**

- Mit Hilfe von node.js wurden über einen Webservice json Daten bereitgestellt.
- Diese können mit einem parametrisierten HTTP GET abgerufen werden.
- Die Daten wurden in einem von UI5s Diagrammtypen dargestellt.

Bei den dargestellten Daten handelt es sich um Dummy-Daten. Im späteren System werden diese vom Data Retrieval abgefragt. Das Verwalten von Sessions wurde angesprochen. Zu diesem Zweck muss eine Datenbank vorhanden sein, in der der Session Store abgelegt wird. Die Funktionalität kann durch express.js und einem entsprechenden Datenspeicher-Plugin bereitgestellt werden.

### **Load Management**

- Brian hat OpenStack auf Ubuntu installiert.

## **2 Sonstiges**

Das interne Treffen am Mi., 27.05.2015 wurde geplant. Folgende Entscheidungen wurden getroffen:

- Oliver bringt Bier mit.

- Florian bringt Nudelsalat oder Kartoffelsalat mit.
- Babak bringt Salat mit.
- Alexander und Ilias kümmern sich um allgemeines wie Brot und Kohle.
- Die Kosten dafür werden geteilt.

Außerdem wurden folgende Punkte angesprochen:

- Nächste Woche finden Einzelgespräche mit Oliver statt.
- Björn schlägt eine Methode zur Bewertung der Zusammenarbeit in der Gruppe vor.
- Die Alpha-Präsentation muss geplant werden.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	28. Mai 2015
Uhrzeit	12:16-12:48 Uhr
Ort	Offis Raum U64
Protokollant	Ilias Schweer
Moderator	Florian Merkel
Anwesend	Ilias Schweer, Oliver Norkus, Florian Merkel, Alexander Zimak, Björn Friedrich, Brian Clark, Babak Izadpanah
Abwesend	-
Nächster Protokollant	Alexander Zimak

## Tagesordnung

0. Letzte Protokolle
1. Aktueller Stand der Swimlanes
2. Alpha-Präsentation
3. Jira, Confluence, SVN
4. Sonstiges

## 0 Letztes Protokoll

Allgemein soll stärker auf inhaltliche Formulierung geachtet werden. Einige Aussagen aus den Treffen müssen nicht unbedingt in das Protokoll mit aufgenommen werden, wie z.B. Überlegungen zur Planung von Sozialveranstaltungen. Dazu sollen die Berichte zu den Swimlanes, also die schon erledigten sowie die geplanten Aufgaben jedes Gruppenmitgliedes, detaillierter in den Protokollen beschrieben werden.

## 1 Aktueller Stand der Swimlanes

In dem aktuellen Sprint wurden im Wesentlichen theoretische Überlegungen und Planungen bezüglich Schnittstellen und Zusammenführen der einzelnen Systemkomponenten sowie den Ablauf der Datenweitergabe betrieben, um für die Alpha-Version einen ersten Durchschnitt des Systems präsentieren zu können (also Anfrage von Kennzahlen auf Clientseite -> Abfragen der Daten aus Datenbank -> Visualisieren der Informationen).

## **2 Alpha-Präsentation**

Oliver kümmert sich um einen geeigneten Termin für die Alpha-Präsentation. Dieser wird frühestens beim nächsten Meeting (04.06.15) stehen und die Präsentation wird ungefähr weitere zwei Wochen später stattfinden. Für den Inhalt der Präsentation müssen noch einige Fragen geklärt werden, wie z.B. das Ressourcenmanagement oder die Skalierbarkeit realisiert wird (ggf. Skalierungsalgorithmus zeigen). Bis zum 11. Juni soll ein Konzept für die Präsentation ausgearbeitet werden (Vortragende, Agenda, usw.).

## **3 JIRA, Confluence, SVN**

Der für JIRA, Confluence und SVN genutzte Server ist seit dem 17.05.15 nicht mehr erreichbar. Ralf wurde unmittelbar informiert, konnte jedoch bisher das Problem nicht lösen und ist bis zum 05.07.15 nicht im OFFIS anwesend. Diesbezüglich wurden einige Alternativen besprochen, wie das Nutzen von ARBI-Ressourcen. Ebenfalls wird sich Oliver darum kümmern, dass der Server wieder erreichbar wird.

## **4 Sonstiges**

Da das Sozialevent Volleyball aufgrund von ungünstigem Wetter und dem Ausfall von Oliver nicht realisiert wurde, wird dieses im Sommer nachgeholt. Ausserdem wurde vereinbart, als weiteres nächstes Sozialevent Kart fahren zu gehen.



# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	04. Juni 2015
Uhrzeit	12:15 - 13:25 Uhr
Ort	Offis Raum U64
Protokollant	Alexander Zimak
Moderator	Ilias Schweer
Anwesend	Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	Björn Friedrich (entschuldigt)
Nächster Protokollant	Björn Friedrich

## Tagesordnung

0. Letztes Protokoll
1. aktueller Stand der Swim Lanes
2. SVN, Jira/Confluence, Alternativen
3. Alpha Präsentation
4. Sonstiges

## 0 Letztes Protokoll

In den zukünftigen Protokollen sollen sowohl die internen Gruppenentscheidungen, als auch die Aufgaben der einzelnen PG-Teilnehmer mit einbezogen werden. Die Aufgaben der PG-Teilnehmer sollen nach dem folgenden Muster ins Protokoll eingefügt werden:

**Name:**

- **letzte Woche:** Hier sind die Aufgaben zu finden, die im Protokoll, welches eine Woche zuvor verfasst wurde, unter dem Punkt 'nächste Woche' aufgeführt wurden. Der PG-Teilnehmer präsentiert kurz seine Ergebnisse und begründet, warum manche Aspekte nicht umgesetzt worden sind (sofern dies der Fall ist).
- **nächste Woche:** Unter diesem Punkt sind die geplanten Aufgaben für den nächsten Sprint zu finden.

# 1 Aktueller Stand der Swimlanes

- **Babak Izadpanah:**

- **letzte Woche:** Babak hat sich im vergangenen Sprint mit der Realisierung und Integration des User Managements unter SAP Hana beschäftigt. Hierfür hat er die User Tabelle erneuert sowie eine Tabelle für Mandanten auf der Hana Datenbank angelegt. Darüber hinaus hat er sich mit der Nutzung von Servlets für das Session Management auseinandergesetzt. Hier gibt es aber noch keine lauffähige Version. Weiter hat sich Babak mit Florian und Brian auf eine gemeinsame Technologie zur Integration der verschiedenen Softwareteile geeinigt (Apache tomCat Server).
- **nächste Woche:** In der nächsten Woche geht es für Babak um die Integration des User Managements mit den anderen Softwarekomponenten des BI-Cloud Systems.

- **Björn:**

- **letzte Woche:** Björn hat in der vergangenen Woche eine RESTful Schnittstelle zwischen dem Data Retrieval und dem Dispatcher entwickelt. Diese ist im jetzigen Stadium in der Lage Daten vom Data Retrieval als JSON-Datei weiterzugeben und diese auf einer Webseite zu visualisieren. Das Data Retrieval ist im jetzigen Stand in der Lage mit statischen Queries Daten von der Datenbank zu lesen. Darüber hinaus hat sich Björn mit Alexander über einen möglichen Ablauf der Alpha-Präsentation verständigt.
- **nächste Woche:** In der nächsten Woche wird sich Björn mit der Integration des Data Retrievals mit den anderen Softwarekomponenten beschäftigen. Er wird auch eine kurze Zusammenfassung seiner Swim Lane, sowie einen Zeitplan für seine verbleibende Zeit bei der PG vorstellen (Arbeitszeit, Urlaub etc.).

- **Ilias:**

- **letzte Woche:** Ilias hat sich wie in jedem Sprint mit der Qualitätssicherung der anfallenden Dokumente beschäftigt. Er hat darüber hinaus Björn in der Entwicklung der RESTful Schnittstelle unterstützt und sich weiterhin mit dem Thema der Date Extraction aus der Hana Datenbank beschäftigt.
- **nächste Woche** Ilias beschäftigt sich in der nächsten Woche mit der Integration des Data Retrievals auf dem Server. Darüber hinaus wird er weiter daran arbeiten, dass Queries dynamisch erstellt und abgefragt werden können.

- **Alexander:**

- **letzte Woche:** Alexander hat sich zusammen mit Brian mit der Anbindung eigener Datenquellen an die Grafiken in der GUI beschäftigt, ist jedoch noch zu keiner Lösung gekommen, da es kaum möglich ist im Browser zu debuggen. Im Layout der Calculation View hat Alexander ein paar visuelle Anpassungen vorgenommen. Darüber hinaus hat sich Alexander mit der Konzeption einer Agenda für die Alpha-Präsentation auseinandergesetzt. Organisatorische Aufgaben waren, wie jede Woche Teil der Aufgaben von Alexander.
- **nächste Woche:** In der nächsten Woche wird sich Alexander mit der Integration des Clients auf dem Server befassen, als auch die Datenanbindung an die Grafik realisieren.

- **Brian:**

- **letzte Woche:** Brian hat bei den internen Treffen im Bereich der GUI, des User Managements und des Dispatchers unterstützend gewirkt. Darüber hinaus hat er sich mit der Anbindung eigener Daten an die GUI beschäftigt.
- **nächste Woche:** Im nächsten Schritt wird sich Brian mit der Installation von OpenStack auf dem Blade beschäftigen und sich einen Überblick über TomCat verschaffen.

- **Florian:**

- **letzte Woche:** Florian hat in der letzten Woche die Probleme auf der VM bereinigt, sodass wir diese wieder gemeinsam nutzen können. Er hat außerdem in der ARBI ein zweites SVN eingerichtet, das wir als Backup nutzen können, falls die VM erneut ausfallen sollte. Florian hat sich auch bezüglich Redmine als Jira/Confluence Alternative erkundigt. Dies wurde aber aufgrund der wieder funktionierenden VM verworfen. Weiter hat sich Florian mit dem Session Management auf TomCat Servern beschäftigt. Hierbei hat Florian schon einige Erkenntnisse bezüglich des Arbeitens mit Sessions auf einzelnen VM's gesammelt. Bei multiplen VM's sind jedoch noch weitere Kenntnisse erforderlich.
- **nächste Woche** Im nächsten Schritt wird sich Florian mit der Integration der Softwarekomponenten auf dem Server beschäftigen sowie die Datenanbindung der einzelnen Komponenten unterstützen.

## 2 JIRA, Confluence, SVN

Aufgrund des Ausfalls der PG-internen VM, auf der SVN, Jira und Confluence laufen, wurde PG-intern über einige Alternativen gesprochen. Die PG hat sich darauf geeinigt ein SVN auf einem Server der ARBI aufzusetzen und als Jira/Confluence-Ersatz die Software Redmine einzusetzen. Die Installation von Redmine wurde wieder verworfen, da die PG-interne VM wieder gestartet worden ist. Das SVN wurde eingerichtet und wird von der PG als Backup-System genutzt.

## 3 Alpha-Präsentation

Der Termin der Alpha-Präsentation wird bis spätestens nächste Woche festgelegt. Die Agenda sieht zum jetzigen Zeitpunkt wie folgt aus:

1. **Motivation**
2. **Anwendungsfall**
3. **Architektur**
4. **Prototyp**
5. **Ausblick**

## 4 Sonstiges

- Oliver hat angemerkt, dass wir alle drei Dokumente (Projektdateien, Organigram, Projektdokument) sowohl ausgedruckt, als auch digital auf CD/USB Stick bereitstellen müssen.
- Alexander erkundigt sich bezüglich der Preise für ein Kart-Team-Event.
- Die PG wird sich im nächsten internen Treffen mit dem Kommunikationsradar als mögliches neues Projektmanagement-Tool beschäftigen.

# Protokoll

Projektgruppe  
BlaaS

Sitzungsdatum	11. Juni 2015
Uhrzeit	12:15 - 13:18 Uhr
Ort	Offis Raum U64
Protokollant	Björn Friedrich
Moderator	Alexander Zimak
Anwesend	Ilias Schweer, Brian Clark, Oliver Norkus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	Brian (Urlaub)
Nächster Protokollant	Babak Izadpanah

## Tagesordnung

0. Letztes Protokoll
1. aktueller Stand der Swim Lanes
2. Alpha Präsentation
3. Sonstiges

## 0 Letztes Protokoll

Bis jetzt gibt es keine Anmerkungen zum Protokoll.

## 1 Aktueller Stand der Swimlanes

- **Babak Izadpanah:**
  - **letzte Woche:** Babak hat weiter an dem Session-Management gearbeitet. Die Session-ID soll via JSP zu den anderen Komponenten geschickt werden. Weiterhin hat er an einer Methode gearbeitet die Session-ID und die User-ID gemeinsam zu übergeben.
  - **nächste Woche:** Der nächste Schritt ist die Integration der Login-Funktion in BICE und die Installation auf dem Server. Weiterhin muss das Servlet mit der Funktion Klasse verbunden werden, die die Logik enthält.

- **Björn:**
  - **letzte Woche:** Björn hat mit Ilias weiter am Data Retrieval gearbeitet. Der REST Service ist voll funktionsfähig. Es können zwei Anfragen gestellt werden. Die erste gibt alle Deckungsbeiträge zurück und der zweite die Gesamtkosten. Die aktuelle Version wurde auf dem Server installiert und ist bei vorhandener VPN Verbindung erreichbar.
  - **nächste Woche:** Die nächsten Schritte sind die Anpassung des TimeStamp-Formats und die Ausgabe der Daten als JSON Objekt. Bei den JSON Objekten muss insbesondere darauf geachtet werden, dass sie den Vorgaben des Local Client entsprechen.
- **Ilias:**
  - **letzte Woche:** Ilias hat sich wie in jedem Sprint mit der Qualitätssicherung der anfallenden Dokumente beschäftigt. Er hat darüber hinaus Björn in der Entwicklung der RESTful Schnittstelle unterstützt und sich weiterhin mit dem Thema der Date Extraction aus der Hana Datenbank beschäftigt. Weiterhin hat er alle notwendigen Queries erstellt und bei der Integration des Data Retrieval auf dem Server mitgewirkt.
  - **nächste Woche:** Die nächsten Schritte sind die Berechnung der Intervalle und die Änderung des TimeStamp. Die Aggregation der über die Intervalle soll auf der Datenbank geschehen. Dafür müssen entsprechende Funktionen auf der Datenbank implementiert werden.
- **Alexander:**
  - **letzte Woche:** Alexander hat sich zusammen mit Florian an der RESTful Schnittstelle vom Local Client zum Dispatcher gearbeitet. Aufgrund unvorhergesehener Schwierigkeiten ist die Schnittstelle noch nicht funktionstüchtig. Auf dem Local Client können nun Filter gesetzt werden. Es kann der Firmenname und der Start- bzw. Endzeitpunkt ausgewählt werden.
  - **nächste Woche:** Weitere Arbeiten am Local Client. Die Datenanbindung an die Datenbank muss noch fertiggestellt werden und die Installation auf unserem Blade muss noch durchgeführt werden. Weiterhin wird die GUI modifiziert und ins Englische übersetzt.
- **Brian:**
  - **letzte Woche:** Brian ist bei diesem Treffen nicht anwesend. Er hat jedoch die Arbeiten am Local Client, Dispatcher und Data Retrieval unterstützt.
  - **nächste Woche:** Brian unterstützt hauptsächlich das Local Client Team, weil dort die meisten Arbeiten noch zu erledigen sind. Weiterhin dient er als Ansprechpartner für die anderen Teams.
- **Florian:**
  - **letzte Woche:** Florian hat mit Alex an der Verbindung zwischen dem Local Client und dem Dispatcher gearbeitet.
  - **nächste Woche:** Die nächste Aufgabe ist die Integration des Dispatchers und die Verbindung von Dispatcher und Local Client bzw. dem Data Retrieval.

## 2 $\alpha$ -Präsentation

Die Präsentation soll maximal ein Stunde dauern. Davon sind 45 min. für den Vortrag vorgesehen und 15 min. für Fragen. Die Redner werden Babak, Björn, Brian und Florian sein. Die Muttersprachler unterstützen Babak und Brian, indem Texte korrigiert und Formulierungen überprüft werden. Die Präsentation gliedert sich wie folgt:

1. **Einleitung** Dauer: 5 min.; Redner: Björn
2. **Anwendungsfall** Dauer: 5 min.; Redner: Björn
3. **Architektur** Dauer: 10 min.; Redner: Florian  
Hier soll das Komponentendiagramm gezeigt werden. Die einzelnen Bereiche werden farblich markiert. Die Farben grenzen die Kategorien ab.
4. **Technologien und Entwurfsentscheidungen** Dauer: 10 min.; Redner: Brian
5. **Prototyp** Dauer: 5 min.; Redner: Babak  
An dieser Stelle wird der Prototyp präsentiert.
6. **Ausblick** Dauer 5 min.; Redner: Babak  
Der Ausblick umfasst alle Funktionen, die bis zur  $\beta$  noch implementiert werden.

Jeder Redner legt den Inhalt seiner Folien selbst fest. Abschließend wird die Präsentation von einer Person erstellt, um ein einheitliches Design zu gewährleisten. Am nächsten Donnerstag (19.06.2015) wird die Präsentation probenhalber gehalten.

## 3 Sonstiges

- Björn hat seinen Plan für die Remote Arbeit vorgestellt. Ist soweit akzeptiert worden. Kritik und Vorschläge sind willkommen.
- Alex' Informationen bzgl. des Kartfahrens in Rastede und Bremen eingeholt. Rastede ist flexibler, weil es schneller zu erreichen ist als Bremen. Ob und wann wir das machen wird wann anders geklärt.
- Babak hat einige Entwürfe für ein Logo vorgestellt. Drei sind in die engere Auswahl gekommen. Eine endgültige Entscheidung wird zu einem späteren Zeitpunkt getroffen.

# Protokoll

Projektgruppe  
BlaaS

Datum	18.06.2015
Uhrzeit	12:15h - 13:35h
Ort	Offis Raum U61
Protokollant	Babak Izadpanah
Moderator	Björn Friedrich
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak
Abwesend	-
Nächster Protokollant	Florian Merkel

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. Präsentation
3. Sonstiges

## 0 Letztes Protokoll

Wenn es Seitens Oliver nach zwei Wochen keine Anmerkungen gibt, dann gilt das Protokoll als abgenommen. Bis jetzt gibt es keine Anmerkungen, weder von der PG, noch Oliver zum Protokoll.

## 1 Aktueller Stand der Swim Lanes

Es wurde zuerst die aktuelle Version der Benutzeroberfläche und ihre Fähigkeiten gezeigt, weil sie bei der Alpha-Präsentation am 19.06.2015 live vorgeführt werden soll. Seitens Oliver gab es einige Verbesserungsvorschläge bzgl. der Darstellung. Oliver hat seine Vorstellung an die Tafel gezeichnet. Die Anmerkungen werden bis zur Präsentation eingearbeitet. Bei den Swim Lanes kann gesagt werden, dass jede Komponente, wie User Manager, Dispatcher, Data Retrieval und Local Client funktionsfähig sind. Es wurde überlegt, wie der Load-Balancer umgesetzt werden kann und welche Skalierungsstrategie verfolgt werden kann. Jede Person hat ihre Aufgabe wie folgt erfüllt:



- Babak:

**letzte Woche:** Der nächste Schritt ist die Integration der Login-Funktion in BICE und die Installation auf dem Server. Weiterhin muss das Servlet mit der Funktionsklasse verbunden werden, die die Logik enthält.

**nächste Woche:** Babak kümmert sich, wie jedes Gruppenmitglied, fokussiert um Dokumentation.

- Alex:

**letzte Woche:** Alexander hat sich zusammen mit Florian an der RESTful-Schnittstelle vom Local Client zum Dispatcher gearbeitet. Aufgrund unvorhergesehener Schwierigkeiten ist die Schnittstelle noch nicht funktionstüchtig. Auf dem Local Client können nun Filter gesetzt werden. Es kann der Firmenname und der Start- bzw. Endzeitpunkt ausgewählt werden.

**nächste Woche:** Das Bearbeiten und Erstellen von Dokumenten für die Projektdokumentation stellt die nächste Aufgabe von Alexander dar.

- Björn:

**letzte Woche:** Die letzten Arbeiten am Data Retrieval müssen abgeschlossen werden, damit sie bei der Präsentation live vorgeführt werden kann. Es sind noch einige Anpassungen bzgl. der JSON Objekte, die zum Local Client geschickt werden, notwendig. Weiterhin soll eine weitere Kennzahl mit einem der Queries abgefragt werden. Dazu muss der Query entsprechend verändert werden. Außerdem gibt es Rundungsfehler bei der Berechnung. Es muss getestet werden, wie sich die einzelnen Rundungsmethoden von SAP (ROUND\_UP, ROUND\_DOWN, ROUND usw.) verhalten.

**nächste Woche:** Es stehen keine größeren Arbeiten am Data Retrieval an. Als nächstes wird der Fokus auf die Dokumentation gelegt.

- Florian:

**letzte Woche:** Die nächste Aufgabe ist die Integration des Dispatchers und die Verbindung von Dispatcher und Local Client bzw. dem Data Retrieval.

**nächste Woche:** Florian arbeitet an der Dokumentation von Dispatcher und User Management. Außerdem recherchiert er wie Load Balancing mit OpenStack funktioniert.

- Ilias:

**letzte Woche:** Die letzten Arbeiten am Data Retrieval müssen abgeschlossen werden, damit sie bei der Präsentation live vorgeführt werden kann. Es sind noch einige Anpassungen bzgl. der JSON Objekte, die zum Local Client geschickt werden, notwendig. Weiterhin soll eine weitere Kennzahl mit einem der Queries abgefragt werden. Dazu muss der Query entsprechend verändert werden. Außerdem gibt es Rundungsfehler bei der Berechnung. Es muss getestet werden, wie sich die einzelnen Rundungsmethoden von SAP (ROUND\_UP, ROUND\_DOWN, ROUND usw.) verhalten.

**nächste Woche:** Es stehen keine größeren Arbeiten am Data Retrieval an. Als nächstes wird der Fokus auf die Dokumentation gelegt.

- Brian:

**letzte Woche:** Die nächste Aufgabe ist die Integration des Dispatchers und die Verbindung von Dispatcher und Local Client bzw. dem Data Retrieval.

**nächste Woche:** Installieren von OpenStack entsprechende Konfiguration sowie Dokumentation.

## 2 Alpha-Präsentation

Es wurde von der Gruppe die aktuelle Version der Präsentation gezeigt. Währenddessen wurden gemeinsam die Folien und die Agenda überarbeitet. Verbesserungen und Optimierungen sind noch bei der Benutzungsoberfläche notwendig. Jeder Redner hat den Inhalt seiner Folien festgelegt. Die Präsentation gliedert sich wie folgt:

- Björn: Einleitung und Anwendungsfall
- Florian: Architektur
- Brian: Technologien und Entwurfsentscheidungen
- Babak: Prototyp und Ausblick

Alle Redner schicken einen groben Entwurf mit ihren Stichpunkten bzw. Bilder an Ilias. Dieser wird alles zu einer Präsentation zusammenfassen und darauf achten, dass sie ein einheitliches Format hat.

Morgen (Freitag, 19.06.2015) um 09:30h findet die Präsentation statt. Alle PG-Mitglieder sollen sich bis 09:00h im Offis einfinden.

## 3 Sonstiges

- Der Artikel muss nochmal von den Autoren Korrektur gelesen werden. Nach der Präsentation liest Brian den Artikel, um sprachliche Optimierungen vorzunehmen. Die finale Version muss am Freitag (19.06.2015) eingereicht werden. Alexander sendet Oliver bis Freitag einen aktuellen Screenshot mit einer englischen Version der Benutzungsoberfläche zu.
- Das letzte Logo in der ersten Reihe soll weiter verfeinert werden. Die Quadrate sollen nicht mehr gekippt sein. Von diesem Logo soll es zwei Versionen geben. Bei der ersten soll die Schrift bleiben wo sie ist. Bei der anderen soll der BICE Schriftzug unter dem Logo stehen.
- Da nicht alle PG-Mitglieder Kart fahren wollen, wird die Aktivität ausgetauscht. Stattdessen wird nach einer alternativen Aktivität gesucht.
- Das Kommunikationsradar wurde ausgefüllt und an Oliver übergeben. Er wertet es bis morgen (19.06.2015) aus und lässt uns das Ergebnis per E-Mail zukommen.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	25. Juni 2015
Uhrzeit	12:15 - 13:30 Uhr
Ort	Offis Raum U64
Protokollant	Florian Merkel
Moderator	Babak Izadpanah
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Ilias Schweer, Oliver Norkus, Florian Merkel (verspätet), Alexander Zimak
Abwesend	-
Nächster Protokollant	Ilias Schweer

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swimlanes
2. Feedback zur Alpha-Präsentation
3. Aufbau/Struktur der Dokumentation
4. Auswertung des Kommunikationsradars
5. Homepage

## 0 Letztes Protokoll

Das letzte Protokoll wurde angenommen.

## 1 Aktueller Stand der Swimlanes

In der letzten Woche wurde hauptsächlich an der Dokumentation des Projekts gearbeitet. Jeder hat den ihm im Montagstreffen zugewiesenen Teil der Dokumentation bearbeitet. Es folgen daher nur die Arbeiten, die darüber hinaus gehen.

- **Babak**
  - **letzte Woche:** Die Alpha-Präsentation wurde vorbereitet und gehalten.

- **diese Woche:** Die Umsetzung von einer tomcat-Realms-basierten Rechte- und Sessionverwaltung im Hauptspeicher bzw. einer Datenbank wurden untersucht und beispielhaft implementiert.
- **Alexander**
  - **letzte Woche:** Im Rahmen der Arbeiten an der Dokumentation wurde nach einer Vorlage für das Projektdokument gesucht. Da eine solche in allgemeingültiger Form jedoch nicht existiert wird eine Vorlage nach eigenem Ermessen erarbeitet.
  - **diese Woche:** Fehler in der Oberfläche des Local Client werden beseitigt. Darüber hinaus wird der Login über das Usermanagement integriert.
- **Björn**
  - **letzte Woche:** Neben dem Bearbeiten der Dokumentation wurde die Alpha-Präsentation bearbeitet und gehalten.
  - **diese Woche:** Die Dokumentation des Data Retrieval wird weiter bearbeitet. Darüber hinaus werden Änderungen zur Verbesserung der Abfrageperformance durchgeführt. Außerdem müssen die Anfragen an die sich ggf. ändernden Bedürfnisse des Local Client angepasst werden.
- **Florian**
  - **letzte Woche:** Neben dem Bearbeiten der Dokumentation wurde die Alpha-Präsentation vorbereitet und gehalten.
  - **diese Woche:** Die Dokumentation des Dispatchers wird weiter bearbeitet. Außerdem müssen die vom Dispatcher bereitgestellten Schnittstellen mit dem Local Client bzw. Data Retrieval abgestimmt werden.
- **Ilias**
  - **letzte Woche:** Arbeiten am Data Retrieval wurden durchgeführt. Außerdem wurde die Vorlage für die Projektdokumentation angepasst.
  - **diese Woche:** Alternativen zur on-the-fly Berechnung der Kennzahlen im SELECT-Query wurden recherchiert. Unter anderem wurden Trigger-basierte Ansätze zur Vorberechnung der gewünschter Kennzahlen untersucht.
- **Brian**
  - **letzte Woche:** Die Installation von OpenStack wurde begonnen, jedoch bereitete die Rechteverwaltung auf dem Blade Schwierigkeiten. Außerdem wurde die Alpha-Präsentation vorbereitet und gehalten.
  - **diese Woche:** Die OpenStack Installation wird abgeschlossen. Im nächsten Schritt sollen Instanzen durch eine API erstellt/entsorgt werden.

## 2 Feedback zur Alpha-Präsentation

Das Feedback zur Alpha-Präsentation wurde besprochen. Es wurden die Anmerkungen und Fragen nach der Präsentation sowie weitere Anmerkungen von Oliver Norkus thematisiert. Aus dem Feedback konnten Erkenntnisse für zukünftige Präsentationen der Projektgruppe

gewonnen werden.

Zur Veranschaulichung des Use Case und des Nutzens des BICE Systems sollte der aktuelle Zustand (Deckungsbeitragsrechnung mit Tabellenkalkulation) anhand eines Beispiels erklärt werden. Auf diesem Beispiel sollte dann die weitere Präsentation aufbauen um eine kontinuierliche “Geschichte” zu präsentieren.

Die Begrifflichkeiten sollten an die Zielgruppe der Präsentation angepasst bzw. der Zielgruppe der Präsentation entsprechend erklärt werden. Der in der Präsentation verwendete Begriff “Load Balancing” ist im Kontext der Energiewirtschaft anders zu verstehen als in der IT. Außerdem sollte die Komponente “Query Optimizer” umbenannt werden, da die Funktionalität der Komponente nicht der von ebenso benannten Komponenten in Informations- bzw. Datenbanksystemen entspricht.

Darüber hinaus ist aufgefallen, dass die Darstellung der Kennzahlen wie sie im aktuellen Prototypen implementiert ist, nicht sinnvoll ist und einer Überarbeitung bedarf.

Ein Protokoll der Präsentation wird von der Projektgruppe erstellt.

### **3 Aufbau/Struktur der Dokumentation**

Die Struktur der Projektdokumentation wurde besprochen. Es wurde festgelegt, dass die Dokumentation die Inhalte des Lastenhefts unter der Überschrift *Anforderungen* eingepflegt werden sollen. Die Inhalte des Pflichtenhefts sollen unter der Überschrift *Entwurf* zu finden sein. Die Seminararbeiten und die Protokolle werden der Projektdokumentation als Anhänge hinzugefügt und im Text an der entsprechenden Stelle darauf verwiesen. Für die Projektdokumentation ist folgende Gliederung vorgesehen.

1. Einleitung
2. Projektorganisation / Vorgehen
3. Anforderungen
4. Entwurf
5. Implementierung
6. Evaluation
7. Fazit
8. Literatur
9. Anhang

Neben der Projektdokumentation müssen auch ein Artikel für das Hasso-Plattner-Institut geschrieben und ein zugehöriges Poster erstellt werden, die das Projekt, dessen Durchführung und seine Ergebnisse beschreiben.

## **4 Kommunikationsradar**

Die Ergebnisse des Kommunikationsradars wurden ausgewertet. Obwohl im Durchschnitt alle Bereiche außer dem gegenseitigen Sachverständnis mit 5 oder 6 bewertet wurden ist die niedrigste Bewertung in allen Bereichen 3 Punkte. Ohne die einzelnen Punkte konkret anzusprechen ist es schwierig Maßnahmen zur Verbesserung zu beschließen, wodurch jedoch die Anonymität des Verfahrens verloren ginge. Aus diesem Grund sollen Maßnahmen zur allgemeinen Verbesserung in den jeweiligen Bereichen erarbeitet werden.

## **5 Homepage**

Die Homepage soll um die Inhalte der Alpha-Präsentation erweitert werden. Außerdem soll der Artikel der INFORMATIK 2015 in die Liste der Publikationen mit aufgenommen werden.

# Protokoll der Alpha-Präsentation von BICE

## Projektgruppe BIaaS

Datum	19.06.2015
Uhrzeit	09:30h - 10:30h
Ort	Offis Raum E02
Protokollant	Ilias Schweer
Vortragenden	Florian Merkel, Björn Friedrich, Babak Izadpanah, Brian Clark
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak, Jürgen Sauer, Marco Grawunder, Marco Haas
Abwesend	-
Nächster Protokollant	-

## Tagesordnung

1. Präsentation
2. Diskussion/Feedback

## 1 Präsentation

Die Projektgruppenmitglieder Björn Friedrich, Florian Merkel, Babak Izadpanah und Brian Clark tragen die Präsentation der Alpha-Version des BICE-Systems vor.

## 2 Diskussion/Feedback

Das Feedback zur Alpha-Präsentation wurde besprochen. Es wurden die Anmerkungen und Fragen nach der Präsentation sowie weitere Anmerkungen von Oliver Norkus thematisiert. Aus dem Feedback konnten Erkenntnisse für zukünftige Präsentationen der Projektgruppe gewonnen werden. Folgende Anmerkungen wurden erwähnt:

- Zur Veranschaulichung des Use Cases und des Nutzens des BICE-Systems sollte der aktuelle Zustand (Deckungsbeitragsrechnung mit Tabellenkalkulation) anhand eines Beispiels erklärt werden. Auf diesem Beispiel sollte dann die weitere Präsentation aufbauen, um eine kontinuierliche "Geschichte" zu präsentieren.

- Die Begrifflichkeiten sollten an die Zielgruppe der Präsentation angepasst bzw. der Zielgruppe der Präsentation entsprechend erklärt werden. Der in der Präsentation verwendete Begriff “Load Balancing” ist im Kontext der Energiewirtschaft anders zu verstehen, als in der IT. Außerdem sollte die Komponente “Query Optimizer” umbenannt werden, da die Funktionalität der Komponente nicht der von ebenso benannten Komponenten in Informations- bzw. Datenbanksystemen entspricht.
- Darüber hinaus ist aufgefallen, dass die Darstellung der Kennzahlen, wie sie im aktuellen Prototypen implementiert ist, nicht sinnvoll ist und einer Überarbeitung bedarf.



# Protokoll

Projektgruppe  
BIaaS

Datum	02.07.2015
Uhrzeit	12:18h - 13:15h
Ort	Offis Raum U61
Protokollant	Ilias Schweer
Moderator	Florian Merkel
Anwesend	Brian Clark, Björn Friedrich, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer, Alexander Zimak
Abwesend	-
Nächster Protokollant	Alexander Zimak

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. Beta Backlog
3. Testdaten
4. Sonstiges

## 0 Letztes Protokoll

Im letzten Protokoll fehlt der Hinweis, dass Oliver am 09.07 verhindert ist. Aus diesem Grund muss nach einem Alternativtermin gesucht werden.

## 1 Aktueller Stand der Swim Lanes

- Babak:

**letzte Woche:** Die Umsetzung von einer tomcat-Realms-basierten Rechte- und Sessionverwaltung im Hauptspeicher bzw. einer Datenbank wurden untersucht und beispielhaft implementiert.

**nächste Woche:** Integration von Usermanagement in die Dispatcher-Komponente. Bearbeitung entsprechender Kapitel im Pflichtenheft.

- Alex:
  - letzte Woche:** Fehler in der Oberfläche des Local Client werden beseitigt. Darüber hinaus wird der Login über das Usermanagement integriert.
  - nächste Woche:** Bearbeitung/Erstellung der Dokumentation des Local Client Kapitels sowie Bearbeitung der GUI.
  
- Björn:
  - letzte Woche:** Die Dokumentation des Data Retrieval wird weiter bearbeitet. Darüber hinaus werden Änderungen zur Verbesserung der Abfrageperformance durchgeführt. Außerdem müssen die Anfragen an die sich ggf. ändernden Bedürfnisse des Local Client angepasst werden.
  - nächste Woche:** Arbeiten an der Dokumentation (DEA für das Data Retrieval erstellt, mit der Dokumentation für REST begonnen sowie Korrekturen an der Dokumentation für das Data Retrieval)
  
- Florian:
  - letzte Woche:** Die Dokumentation des Dispatchers wird weiter bearbeitet. Außerdem müssen die vom Dispatcher bereitgestellten Schnittstellen mit dem Local Client bzw. Data Retrieval abgestimmt werden.
  - nächste Woche:** Dokumentation bearbeiten (Dispatcher + User Management), dazu Implementierung von Accounting planen und ggf. anfangen.
  
- Ilias:
  - letzte Woche:** Alternativen zur on-the-fly Berechnung der Kennzahlen im SELECT-Query wurden recherchiert. Unter anderem wurden Trigger-basierte Ansätze zur Vorberechnung der gewünschter Kennzahlen untersucht.
  - nächste Woche:** Erstellen/Bearbeiten der Dokumentation des Data Retrieval Kapitels (Beschreibung der getLocation und getCompanies Methoden, allgemeine Beschreibung des Data Retrievals)
  
- Brian:
  - letzte Woche:** Die OpenStack Installation wird abgeschlossen. Im nächsten Schritt sollen Instanzen durch eine API erstellt/entsorgt werden.
  - nächste Woche:** Realisierung des Zugriffs von Open Stack über eine Konsole. Ebenfalls Realisierung einer VM von Tomcat auf Open Stack. Dazu Recherche über Dynamic Scaling/Load Balancing.

## 2 Beta Backlog

Beta Backlog						
Aufgaben	Unteraufgabe	Tasks	Priorität	Fälligkeit	Status	Verantwortlicher
Local Client						
	Graphik anpassen	Graphen Visualisierung ändern		09. Juli	active	Alex
	Hierarchische Tabelle			10. Aug	open	
	User Management			10. Aug	open	
User management						
	User Management	Schnittstelle zum Dispatcher entwickeln		09. Juli	active	Babak/Florian
	Session Management implementieren			10. Aug	open	
	Accounting Komponenten entwerfen			10. Aug	open	
	Accounting Report erstellen			10. Aug	open	
Load Balancing						
	Recherche			10. Aug	active	Brian
	geeignete Methode wählen			10. Aug	open	
Data Retrieval						
	REST-Schnittstelle anpassen			10. Aug	open	
	Datenbankzugriff anpassen			10. Aug	open	
Database						
	Testtabelle für Deckungsbeiträge erstellen			10. Aug	open	
	Testdaten erzeugen und einpflegen			10. Aug	open	
Dokumentation						
	Pflichtenheft	Swim Lane Dokumentation		10. Aug	active	Alle
	Gesamtdokument			10. Aug	active	Ilias

Abbildung 1: Beta Backlog

## 3 Testdaten

Wir werden im Laufe der nächsten Woche Testdaten für das BICE-System erhalten. Dazu sendet Ilias die bestehenden Testdaten als .csv Datei an Oliver. In diesem Schema werden dann die neuen Testdaten zur Verfügung gestellt.

## 4 Sonstiges

- Die PG ist bei dem diesjährigen Boule-Turnier offiziell angemeldet. Dieses findet am Mittwoch, 08.07.15 ab 14 Uhr statt.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	08. Juli 2015
Uhrzeit	10:14 - 11:00 Uhr
Ort	Offis Raum U64
Protokollant	Alexander Zimak
Moderator	Ilias Schweer
Anwesend	Ilias Schweer, Björn Friedrich, Oliver Nor- kus, Babak Izadpanah, Florian Merkel, Alexander Zimak
Abwesend	Brian Clark (entschuldigt)
Nächster Protokollant	Brian Clark

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. Stand der Dokumentation
3. Beta Backlog
4. Sonstiges

## 0 Letztes Protokoll

Es gab keine Anmerkungen zum letzten Protokoll.

## 1 Aktueller Stand der Swimlanes

- **Babak Izadpanah:**
  - **letzte Woche:** Integration vom User-Management in die Dispatcher-Komponente. Bearbeitung entsprechender Kapitel im Pflichtenheft.
  - **nächste Woche:** Weiterentwicklung der User-Managementkomponente und Anbindung an die Dispatcher Komponente sowie Anbindung an die Datenbank. Des Weiteren wurde das User-Managementkapitel im Pflichtenheft weiter bearbeitet.

- **Björn:**
  - **letzte Woche:** Arbeiten an der Dokumentation (DEA für das Data Retrieval erstellt, mit der Dokumentation für REST begonnen sowie Korrekturen an der Dokumentation für das Data Retrieval).
  - **nächste Woche:** Das Format für den Time Stamp wurde für die Rückgabewerte des Data Retrievals angepasst, vorläufig noch mittels eines Workarounds. Es wurde jedoch schon eine geeignetere Lösung gefunden, welche in Zukunft umgesetzt wird. In der Dokumentation wurden Anpassungen für die Rückgabewerte an den Local Client vorgenommen. Darüber hinaus wurden Vorbereitungen für einen Performanctest getroffen, eine entsprechende Testklasse wurde implementiert.
- **Ilias:**
  - **letzte Woche:** Erstellen/Bearbeiten der Dokumentation des Data Retrieval Kapitels (Beschreibung der getLocation und getCompanies Methoden, allgemeine Beschreibung des Data Retrievals)
  - **nächste Woche:** Ilias hat sich mit Änderungen am Data Retrieval, sowie dem Bearbeiten des Pflichtenhefts beschäftigt.
- **Alexander:**
  - **letzte Woche:** Bearbeitung/Erstellung der Dokumentation des Local Client Kapitels sowie Bearbeitung der GUI.
  - **nächste Woche:** Alexander hat sich mit der Anpassung der Tabellenstruktur auf der GUI beschäftigt und eine neue hierarchische Tabelle eingefügt, darüber hinaus wurden mit Björn und Ilias die Anzeigefehler der Alpha-Präsentation behoben.
- **Brian:**
  - **letzte Woche:** Realisierung des Zugriffs von Open Stack über eine Konsole. Ebenfalls Realisierung einer VM von Tomcat auf Open Stack. Dazu Recherche über Dynamic Scaling/Load Balancing.
  - **nächste Woche:** Brian hat versucht ein Ubuntu-Image auf OpenStack zum Laufen zu bekommen, dies ist jedoch noch nicht möglich. Darüber hinaus hat sich Brian damit beschäftigt eine Tomcat Instanz auf dem Ubuntu Image zu installieren. Brian hat sich außerdem mit der Recherche der Apache Stratos Dokumentation beschäftigt.
- **Florian:**
  - **letzte Woche:** Dokumentation bearbeiten (Dispatcher + User Management), dazu Implementierung von Accounting planen und ggf. anfangen.
  - **nächste Woche:** Anpassung des User-Management Kapitels im Pflichtenheft. Weiterentwicklung der Dispatcher Komponente, sowie der Accounting Komponente.

## 2 Stand der Dokumentation

Die Dokumentationsteile Dispatcher, Local Client, Data Retrieval und User Management wurden zur Prüfung an Oliver Norkus übergeben. Oliver wird wenn nötig Feedback liefern.

### **3 Beta Backlog**

Anpassungen des Beta Backlogs werden im nächsten internen Treffen vorgenommen.

### **4 Sonstiges**

- Die Anmeldung zur Modulprüfung wurde freigeschaltet und jedes PG-Mitglied hat bis zum 31.07 Zeit sich für die Modulprüfung anzumelden. Sollte dies versäumt werden, kann das PG-Mitglied nicht nachgetragen werden und somit die PG nicht erfolgreich abschließen.
- Es wurden organisatorische Angelegenheiten und das Regelwerk für das Boule-Turnier besprochen.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	23. Juli 2015
Uhrzeit	10:15 - 12:17 Uhr
Ort	Offis Raum U64
Protokollant	Alexander Zimak
Moderator	Björn Friedrich
Anwesend	Ilias Schweer, Björn Friedrich, Oliver Nor- kus, Alexander Zimak, Brian Clark
Abwesend	Florian Merkel (entschuldigt), Babak Izadpanah (entschuldigt)
Nächster Protokollant	Brian Clark

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. Stand der Dokumentation
3. HPI
4. Wahl des Logos
5. Sonstiges

## 0 Letztes Protokoll

Es gab keine Anmerkungen zum letzten Protokoll.

## 1 Aktueller Stand der Swimlanes

- **Babak Izadpanah:**
  - **letzte Woche:** Weiterentwicklung der User-Managementkomponente und Anbin-  
dung an die Dispatcher Komponente sowie Anbindung an die Datenbank. Des Wei-  
teren wurde das User-Managementkapitel im Pflichtenheft weiter bearbeitet.
  - **nächste Woche:** Babak hat Urlaub und wird danach am User-Management wei-  
terarbeiten.

- **Björn:**
  - **letzte Woche:** Das Format für den Time Stamp wurde für die Rückgabewerte des Data Retrievals angepasst, vorläufig noch mittels eines Workarounds. Es wurde jedoch schon eine geeignetere Lösung gefunden, welche in Zukunft umgesetzt wird. In der Dokumentation wurden Anpassungen für die Rückgabewerte an den Local Client vorgenommen. Darüber hinaus wurden Vorbereitungen für einen Performance-test getroffen, eine entsprechende Testklasse wurde implementiert.
  - **nächste Woche:** Änderung des Timestamp-Formates in das Format (YYYY-MM-DD hh:mm). Sowie die Recherche bezüglich des SAP Plan Cache.
- **Ilias:**
  - **letzte Woche:** Ilias hat sich mit Änderungen am Data Retrieval, sowie dem Bearbeiten des Pflichtenhefts beschäftigt.
  - **nächste Woche:** Ilias beschäftigt sich mit der Weiterführung der QS der Dokumentation. Darüber hinaus wird er sich mit Björn mit der Recherche bezüglich des SAP Plan Cache beschäftigen.
- **Alexander:**
  - **letzte Woche:** Alexander hat sich mit der Anpassung der Tabellenstruktur auf der GUI beschäftigt und eine neue hierarchische Tabelle eingefügt, darüber hinaus wurden mit Björn und Ilias die Anzeigefehler der Alpha-Präsentation behoben.
  - **nächste Woche:** Wird sich weiter mit der Hierarchischen-Tabelle beschäftigen. Darüber hinaus wird er sich mit Marco in Verbindung setzen um eine geeignete graphische Visualisierung zu wählen.
- **Brian:**
  - **letzte Woche:** Brian hat versucht ein Ubuntu-Image auf OpenStack zum Laufen zu bekommen, dies ist jedoch noch nicht möglich. Darüber hinaus hat sich Brian damit beschäftigt eine Tomcat Instanz auf dem Ubuntu Image zu installieren. Brian hat sich außerdem mit der Recherche der Apache Stratos Dokumentation beschäftigt.
  - **nächste Woche:** Brian wird sich mit dem Load Balancing beschäftigen, dafür wird er OpenStack Heat genauer unter die Lupe nehmen.
- **Florian:**
  - **letzte Woche:** Anpassung des User-Management Kapitels im Pflichtenheft. Weiterentwicklung der Dispatcher Komponente, sowie der Accounting Komponente.
  - **nächste Woche:** Weiterentwicklung der Dispatcher Komponente, sowie der Accounting Komponente.

## 2 Stand der Dokumentation

Die angefertigten Dokumentationsteile müssen vertieft und in das Pflichtenheft eingebunden werden.



### **3 HPI**

Alexander wurde als Plakatverantwortlicher festgelegt. Es wurde weiterhin beschlossen, dass es sinnvoll wäre, als geschlossene PG zur Vorstellung beim HPI zu fahren. Alexander wird sich erkundigen, ob die Uni Oldenburg für diesen Anlass Fördergelder bereitstellen würde.

### **4 Wahl des Logos**

Die PG hat sich für das angehängte Logo entschieden.

### **5 Sonstiges**

- Bis zur Alpha sollte es möglich sein mittels eines Proxy-Servers geräte-unabhängig auf das BICE System zugreifen zu können.
- Oliver wird vom 10.08 bis 23.08 im Urlaub sein.
- Ilias und Alex fehlen beim Treffen am 30.07 aufgrund einer Klausur.



**BICE**

Abbildung 1: BICE

# Protokoll

Projektgruppe  
BIaaS

Datum	30.07.2015
Uhrzeit	12:15h - 12:30h
Ort	Offis Raum U61
Protokollant	Babak Izadpanah
Moderator	Florian Merkel
Anwesend	Brian Clark, Babak Izadpanah, Florian Merkel und Oliver Norkus
Abwesend	Björn Friedrich (in Japan), Ilias Schweer und Alexander Zimak (Klausurtermin)
Nächster Protokollant	Ilias Schweer

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. HPI
3. Sonstige

## 0 Letztes Protokoll

Das letzte Protokoll war zur Zeit der Sitzung leider noch nicht verfügbar.

## 1 Aktueller Stand der Swim Lanes

- **Babak Izadpanah:**
  - **letzte Woche:** Babak hatte Urlaub und wird danach am User-Management weiterarbeiten.
  - **nächste Woche:** Die Schnittstelle zwischen Usermanagement und Dispatcher wurde realisiert.

- **Björn:**
  - **letzte Woche:** Änderung des Timestamp-Formates in das Format (YYYY-MM-DD hh:mm). Sowie die Recherche bezüglich des SAP Plan Cache.
  - **nächste Woche:** Verfassen des Kapitels zum Kommunikationsradar für die Projektdokumentation sowie Arbeiten an den Tests.
- **Ilias:**
  - **letzte Woche:** Ilias beschäftigt sich mit der Weiterführung der QS der Dokumentation. Darüber hinaus hat er sich mit Björn mit der Recherche bezüglich des SAP Plan Cache beschäftigt.
  - **nächste Woche:** Planung der Testfälle mit Björn. Allgemeine Dokumentation und Recherche bezüglich Timestamp und Plan-Cache.
- **Alexander:**
  - **letzte Woche:** Alexander hat sich weiterhin mit der Hierarchischen-Tabelle beschäftigt. Darüber hinaus hat er sich mit Marco in Verbindung gesetzt, um eine geeignete graphische Visualisierung zu wählen.
  - **nächste Woche:** Arbeiten am Local Client (grafische Oberfläche, Schnittstellen, usw.).
- **Brian:**
  - **letzte Woche:** Brian hat sich mit dem Load Balancing beschäftigt. Dafür hat er sich mit der Konfiguration von OpenStack Heat beschäftigt.
  - **nächste Woche:** Brian beschäftigt sich mit Templates für Load-Balancing und Auto-scaling.
- **Florian:**
  - **letzte Woche:** Beschäftigen mit der Accounting-Komponente.
  - **nächste Woche:** Weiterentwicklung der Dispatcher Komponente sowie der Accounting Komponente.

## 2 HPI

Alex sammelt weiterhin Informationen bezüglich einer möglichen Finanzierung der Reise zu den HPI-Vorträgen. Die Vorlage für das Plakat zu unserem Projekt ist im Confluence verfügbar und muss weiter bearbeitet werden.

## 3 Sonstiges

- Am 13.08.2015 und 20.08.2015 findet keine Sitzung statt.
- Am 24.08.2015 findet die Präsentation unserer Beta statt.

# Protokoll

Projektgruppe  
BIaaS

Datum	06.08.2015
Uhrzeit	12:15h - 12:45h
Ort	Offis Raum U61
Protokollant	Ilias Schweer
Moderator	Babak Izadpanah
Anwesend	Brian Clark, Babak Izadpanah, Oliver Norkus, Ilias Schweer, Alexander Zimak
Abwesend	Björn Friedrich, Florian Merkel (beide entschuldigt)
Nächster Protokollant	Alexander Zimak

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. HPI
3. Beta Präsentation
4. Sonstiges

## 0 Letztes Protokoll

Es gibt keine Anmerkungen zu dem letzten Protokoll.

## 1 Aktueller Stand der Swim Lanes

- Babak:

**letzte Woche:** Die Schnittstelle zwischen Usermanagement und Dispatcher wurde realisiert.

**nächste Woche:** Arbeiten an der User Management Dokumentation und allgemeine Fehlerbehebungen/Optimierungen an der User Management Komponente.

- Alex:

**letzte Woche:** Arbeiten am Local Client (grafische Oberfläche, Schnittstellen, usw.).

**nächste Woche:** Alexander hatte Urlaub und hat deshalb keine Aufgaben wahrgenommen.

- Björn:

**letzte Woche:** Verfassen des Kapitels zum Kommunikationsradar für die Projektdokumentation sowie Arbeiten an den Tests.

**nächste Woche:** Arbeiten am Pflichtenheft und Projektdokument (Glossar, Accounting, Rollen, Kommunikationsradar).

- Florian:

**letzte Woche:** Weiterentwicklung der Dispatcher Komponente sowie der Accounting Komponente.

**nächste Woche:** Florian hat in dieser Woche keine Aufgaben wahrgenommen, da er Urlaub hat.

- Ilias:

**letzte Woche:** Planung der Testfälle mit Björn. Allgemeine Dokumentation und Recherche bezüglich Timestamp und Plan-Cache.

**nächste Woche:** Ilias hatte Urlaub und hat deshalb keine Aufgaben wahrgenommen.

- Brian:

**letzte Woche:** Brian beschäftigt sich mit Templates für Load-Balancing und Auto-scaling.

**nächste Woche:** Arbeiten an verschiedenen Komponenten und Schnittstellen.

## 2 HPI

- Auf der Homepage des HPI gibt es für uns wichtige Vorlagen und Bestimmungen für das Plakat sowie das Paper, welche im Rahmen des Projektes umgesetzt werden sollen.
- Die entsprechenden Vorträge finden am 04.11.15 statt. Die Abgabefrist für das Plakat/Paper ist am 04.10.15.
- Bis zur Beta muss von der Gruppe eine erste Version des Plakats sowie des Papers vorbereitet und an Oliver zur Überprüfung geschickt werden.
- Alexander hat eine erste Konversation mit der verantwortlichen Person bezüglich Fördergelder für die Reise zu den HPI-Vorträgen gehalten. Leider ergab dieses Gespräch keine Fortschritte, sodass Alexander einen weiteren Versuch neue Erkenntnisse zu gewinnen startet.

## 3 Beta Präsentation

- Für die Beta Präsentation ist es ausreichend, die Load Balancing Komponente auf konzeptionelle Weise vorzustellen.
- Das User Management muss vollständig implementiert und funktionsbereit vorgestellt werden.
- Ebenso gilt dies für die Accounting Komponente.
- Es wurde beschlossen, dass die Beta Präsentation für ein fachkundiges Publikum gedacht ist, das bereits die vorherigen Präsentationen mitverfolgt hat.

## 4 Sonstiges

- Für eine mögliche nächste Sozialaktivität wurde Volleyballspielen in Erwägung gezogen. Ein möglicher Zeitraum ist Ende August (nach der Beta Präsentation).
- Die Abschlusspräsentation wird nach hinten verschoben.
- Die nächsten beiden Meetings werden aufgrund von Urlaub nicht stattfinden.

# Protokoll

Projektgruppe  
BIaaS

Datum	27.08.2015
Uhrzeit	12:15h - 13:00h
Ort	Offis Raum U64
Protokollant	Babak Izadpanah
Moderator	Ilias Schweer
Anwesend	Brian Clark, Babak Izadpanah, Florian Merkel, Oliver Norkus, Ilias Schweer und Alexander Zimak
Abwesend	Björn Friedrich (entschuldigt)
Nächster Protokollant	Brian Clark

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes bzw. Nachbesprechung der Beta (-Präsentation)
2. Weiteres Vorgehen
3. HPI Paper und Poster
4. Sonstiges

## 0 Letztes Protokoll

Es gab keine Anmerkungen zum letzten Protokoll.

## 1 Aktueller Stand der Swim Lanes

- Der Stand des Projekts zur Beta Präsentation kam beim Publikum gut an und war zufriedenstellend.
- Die GUI, das Accounting sowie das User-Management waren die Kernfeatures der Präsentation und sind implementiert.
- Der BICE-Dienst ist bisher ausschließlich über einen lokalen Proxy erreichbar. Für die Abschlusspräsentation könnte ein über das Internet erreichbarer Proxy verwendet werden.



## 2 Weiteres Vorgehen

- Erstellung des Product Backlog mit den Aufgaben bis zum Projektabschluss. Dieser wird beim nächsten Meeting besprochen.

## 3 HPI Paper und Poster

- Der Titel des Papers sollte nicht den identischen Namen des vorherigen veröffentlichten Paper haben.
- Das Poster soll auf Englisch verfasst sein.
- Das Design und die allgemeine Formatierung des Posters sind gelungen. Allerdings muss stärker auf die Architektur und die verwendeten Technologien und Ressourcen eingegangen werden.
- Die Schrift und die Abbildungen sollen etwas kleiner sein.
- Für den HPI Future SOC Day muss die Projektgruppe Anfahrt und Unterbringung selbstständig organisieren.
- Bisher gibt es seitens der Uni Oldenburg keine Möglichkeit sich die Exkursion finanzieren zu lassen. Oliver wird aber nochmals im OFFIS und bei Marco nachfragen. Ausserdem wartet Babak noch auf Antworten von ISO, HGAS und ASTA.

## 4 Sonstiges

- Bis nächsten Donnerstag kümmert sich Oliver um Testdaten.
- Als Sozialevent wurde Grillen vorgeschlagen.
- Am 17.09.2015 haben wir keine Sitzung. Stattdessen treffen wir uns am 18.09.2015 um 14 Uhr.
- Das Treffen am 01.10.2015 verschieben wir auf den 05.10.2015 um 14 Uhr, wo die endgültige Abgabe erfolgt.
- Beim nächsten Treffen ist Herr Sauer anwesend.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	10. September 2015
Uhrzeit	12:15-13:15 Uhr
Ort	Offis Raum U64
Protokollant	Alexander Zimak
Moderator	Babak Izadpanah
Anwesend	Ilias Schweer, Oliver Norkus, Babak Izadpanah, Alexander Zimak, Jürgen Sauer, Marco Haas
Abwesend	Brian Clark (entschuldigt), Florian Merkel (entschuldigt), Björn Friedrich (entschuldigt)
Nächster Protokollant	Brian Clark

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swim Lanes
2. Backlog
3. HPI Paper and Poster
4. Sonstiges

## 0 Letztes Protokoll

Es gab keine Anmerkungen zum letzten Protokoll.

## 1 aktueller Stand der Swim Lanes

- **Babak Izadpanah:**
  - **letzte Woche:** Einarbeitung in Angular/Bootstrap.
  - **nächste Woche:** Entwicklung der Unit-Test-Umgebung und der ersten Tests.
- **Björn:**
  - **letzte Woche:** Bearbeitung des HPI Papers sowie des Projektdokuments.

- **nächste Woche:** Bearbeitung des HPI Papers sowie des Projektdokuments.
- **Ilias:**
  - **letzte Woche:** Bearbeiten und Einpflegen der Testdaten in die HANA Datenbank.
  - **nächste Woche:** Bearbeitung des Projektdokuments und des HPI Posters (einschließlich Abbildung).
- **Alexander:**
  - **letzte Woche:** Bearbeitung des Product Backlogs sowie Bearbeitung des Testkonzepts und Recherche bezüglich passender Tools.
  - **nächste Woche:** Bearbeitung des Pflichtenhefts und des Projektdokuments.
- **Brian:**
  - **letzte Woche:** Brian hatte letzte Woche Urlaub.
  - **nächste Woche:** Brian implementiert das Skript fürs Load Balancing. Anbindung zwischen Dispatcher und Client wiederherstellen.
- **Florian:**
  - **letzte Woche:** Bearbeitung des Session-Managements auf OpenStack.
  - **nächste Woche:** Anpassung des Dispatchers, aufgrund der Änderung am Data Retrieval.

## 2 Backlog

Es gab keine Anmerkungen bezüglich des Backlogs.

## 3 HPI Paper and Poster

- **HPI Paper:** Der Abschnitt bezüglich der Anforderungen soll gekürzt werden, dafür eine Grafik der Architektur eingefügt werden. Darüber hinaus muss das Kapitel über die Evaluation noch geschrieben werden. Weitere Anpassungen sind die Verallgemeinerung der Mail Adressen nach dem Muster [Vorname].[Nachname]@uni-oldenburg.de und die Anpassung des Titels zu "Business Intelligence in the Cloud for Energy (BICE)".
- **Poster:** Die Anordnung der Abbildungen und des Textes sollten verändert werden. Hier bietet sich eine zentrale Abbildung mit Textspalten links und rechts neben der Abbildung an. Darüber hinaus soll die Abbildung noch einmal überarbeitet und genau gegliedert werden.

## 4 Sonstiges

- Die Gliederung des Projektdokuments muss noch angepasst werden. Das Pflichten- und Lastenheft soll unter Anforderung zusammengefasst werden und die Kapitel Entwurf, Implementation und Test hinzugefügt werden. Die Seminararbeiten und Protokolle sollen im Anhang aufgeführt werden.

- Marco benötigt die Teilnehmerliste für den Future SOC Lab Day. Diese wird Alexander ihm zukommen lassen.
- Das nächste PG Treffen wurde auf den 19.09 verlegt.

# Protokoll

Projektgruppe  
BIaaS

Sitzungsdatum	24. September 2015
Uhrzeit	12:15 - 13:10 Uhr
Ort	Offis Raum U64
Protokollant	Florian Merkel
Moderator	Alexander Zimak
Anwesend	Brian Clark (verspätet), Babak Izadpanah, Ilias Schweer, Oliver Norkus, Florian Merkel, Alexander Zimak
Abwesend	Björn Friedrich (entschuldigt)
Nächster Protokollant	Ilias Schweer

## Tagesordnung

0. Letztes Protokoll
1. Aktueller Stand der Swimlanes
2. HPI Poster und Paper
3. Sonstiges

## 0 Letztes Protokoll

Zum Punkt *HPI Paper and Poster* gab es folgende Anmerkungen:

- Die E-Mail Adressen sollen in geschweiften, nicht eckigen, Klammern stehen.
- Im Titel des Papers sollen keine Abkürzungen eingeführt werden.

## 1 Aktueller Stand der Swimlanes

### Babak

- **letzte Woche:** Babak hat sich in den Einsatz von JUnit eingearbeitet und begonnen die Unit-Tests für den Quelltext zu planen. Dabei soll nicht alles getestet werden.
- **nächste Woche:** Das Testkonzept soll fertiggestellt und umgesetzt werden. Bei der Dokumentation muss beachtet werden, dass nicht nur Unit-Tests Erwähnung finden. Beispielsweise wurden auch inkrementelle Systemtests mit dem Kunden durchgeführt. (→ Zwischenpräsentationen mit Marco Haas)

## Björn

- **letzte Woche:** Björn hat an der Umsetzung der Performance-Tests und am HPI-Paper gearbeitet.
- **nächste Woche:** Björn wird weiterhin an den Performance-Tests arbeiten. Arbeiten am Paper werden auf die Zeit nach der Abgabe verschoben.

## Brian

- **letzte Woche:** Brian hat am Load Balancing mit OpenStack gearbeitet. Es werden virtuelle Maschinen gestartet, doch die Verbindung zwischen diesen und dem Load Balancer funktioniert noch nicht.
- **nächste Woche:** Brian wird sich weiterhin mit der Konfiguration des OpenStack-Frameworks beschäftigen.

## Ilias

- **letzte Woche:** Ilias hat die Testdaten in die Datenbank eingepflegt und das DataRetrieval entsprechend geändert. Außerdem wurde am Projektdokument gearbeitet.
- **nächste Woche:** Ilias wird primär am Projektdokument arbeiten.

## Alexander

- **letzte Woche:** Alexander hat am Projektdokument gearbeitet.
- **nächste Woche:** Alexander wird primär am Projektdokument arbeiten. Im Anforderungskapitel werden die Inhalte des Lastenhefts und zum Teil des Pflichtenhefts eingearbeitet. Verwendete Technologien werden im Umsetzungskapitel beschrieben.

## Florian

- **letzte Woche:** Florian hat an der Oberfläche des Accounting gearbeitet und mit Ilias zusammen das DataRetrieval für die neuen Daten abgeändert. Außerdem wurden Apache Shiro als mögliches Framework für ein Session-Management-System identifiziert.
- **nächste Woche:** Florian wird sich mit der Implementierung des Session-Managements mit Shiro befassen und das Accounting überarbeiten.

## 2 Backlog

Das Backlog hat sich seit dem letzten Mal dahingehend verändert, dass weitere Tasks hinzugekommen sind, was die Dokumentation betrifft.

### 3 UI

Die aktuelle Version der Nutzungsoberfläche wurde vorgestellt. Es wurde angemerkt, dass zur Zeit die Datenwerte, die für ein 15-Minuten-Intervall gelten, als einzelne Punkte angezeigt werden, auch wenn Daten für einen größeren Zeitraum angezeigt werden, sodass eine Aggregation möglicherweise sinnvoller ist. *Nachtrag: Nach Rücksprache mit Marco Haas wurde entschieden, das aktuelle Verhalten beizubehalten.* Des Weiteren sorgt die aktuelle Darstellung von Erlös und Profit für Verwirrung. Die Erlöslinie im Graphen wird rot angezeigt, in der Legende ist die rote Farbe dem Profit zugeordnet. In der Tabelle liegt auch eine andere Verwendung der Farben vor. Die verwendeten Farben müssen daher sowohl im Graphen als auch in der Tabelle überarbeitet werden. In der Darstellung des Accounting fehlt die Anzeige des Benutzers und des Mandanten. Zur Zeit werden außerdem keine Werte angezeigt, da diese in der Datenbank einer falschen Benutzer ID zugeordnet werden. Möglicherweise sollen alle Werte eines Datenbank-Mandanten angezeigt werden können.

### 4 HPI Poster und Paper

Das Poster in der aktuellen Version wurde vorgestellt. Es gab folgende Anmerkungen:

- Browser- und Betriebssystemicons: Hier soll durch Punkte vermittelt werden, dass noch mehr als die durch die Icons symbolisierten Browser und Betriebssysteme unterstützt werden.
- Die weißen Ecken beim Bild des iPads sollen verschwinden.
- Durch eine entsprechende Beschriftung soll verdeutlicht werden, dass die blauen Kästen virtuelle Maschinen darstellen.
- Durch Überlagern von Komponenten kann Platz gespart werden.
- Die Breite der Abbildung muss ggf. angepasst werden um die Textplatzierung zu optimieren.
- Die Namen der Uni und des Departments sollen auf Englisch übersetzt werden.
- Jürgen Sauer muss auf dem Plakat erwähnt werden.

Zum Paper gab es folgende Anmerkungen:

- Einige Anmerkungen wurden noch nicht eingepflegt.
- Das Paper für das HPI sollte erst nach der Abgabe des Projekts am 05.10.25 weiter bearbeitet werden.
- Im Paper soll herausgestellt werden, welche Erkenntnisse durch das Projekt gewonnen wurden  
→ Die Realisierung eines Cloud-Dienstes ist mit den eingesetzten Technologien und der entwickelten Architektur möglich.

## 5 Sonstiges

Da Brian nicht mit zum HPI fahren kann, um das Projekt vorzustellen, muss Rücksprache mit Marco Haas bezüglich der Zimmer-Reservierung gehalten werden. Vor der Vorstellung am HPI in Potsdam soll noch ein Social Event stattfinden.