

Architektur und Implementierung eines
kommerziellen Grid Resource Brokers auf Basis
der Globus Infrastruktur

Michael Gerber

23. Mai 2004

Aufgabensteller Prof. Dr.-Ing. Wilhelm G. Spruth

Institut für Informatik, Universität Leipzig

Wilhelm-Schickard-Institut für Informatik, Eberhard-Karls-Universität Tübingen

Betreuer Joachim Franz

IBM Business Consulting Service, Stuttgart

Abgabe-Datum 31. Mai 2004

Erklärung der Selbständigkeit

Ich versichere, dass ich die vorliegende Arbeit selbständig angefertigt und mich fremder Hilfe nicht bedient habe. Alle Stellen, die wörtlich oder sinngemäß veröffentlichtem oder unveröffentlichtem Schrifttum entnommen sind, habe ich als solche kenntlich gemacht.

Danksagung

Inhaltsverzeichnis

1	Einleitung	14
1.1	Die Transaktionsmaschine	14
1.1.1	Einleitung	14
1.1.2	Processing Domain	15
1.1.3	Channel Controller und Profiler	16
1.1.4	Masterflow Controller	16
1.1.5	Carrier/Subworkflow Controller	16
1.1.6	Service Request Work Unit	17
1.2	Band Distribution Algorithm for Processing Domains	17
1.3	Aufgabenstellung und Ziele der Diplomarbeit	18
2	Der kommerzielle Grid Resource Broker	19
2.1	Einleitung	19
2.2	Aufbau	19
2.2.1	Central Broker Part	23
2.2.1.1	Profiler	24
2.2.1.2	Bulletin Board	26
2.2.1.3	Collection Service	27
2.2.1.4	On-Demand Information Service	28
2.2.1.5	Band Mapping Unit	30
2.2.1.6	Distributor	32
2.2.1.7	Receiver	32
2.2.1.8	Mediator	33
2.2.2	Distributed Broker Part	34
2.2.2.1	On-Demand Information Service	35

2.2.2.2	Auction Unit	36
2.2.2.3	Band Transformer	38
2.2.2.4	Bulletin Board	40
2.2.2.5	Data Collection Unit	42
2.2.2.6	Dispatcher	43
2.2.3	SRWU-Broker Segment	44
2.2.3.1	SRWU-Id	44
2.2.3.2	State	45
2.2.3.3	Customer Service Element	45
2.2.3.4	Auction Settings	46
2.2.3.5	WLM-Settings	47
2.2.4	Bid	47
2.2.5	PDP	48
2.2.6	PSE	48
2.3	Funktionsbeschreibung	49
2.3.1	Annahme und Verifikation einer SRWU	51
2.3.1.1	Annahme einer gültigen SRWU	51
2.3.1.2	Ablehnung einer ungültigen SRWU	53
2.3.1.3	Vermittlung einer nicht verarbeitbaren SRWU	54
2.3.2	Bestimmung einer geeigneten PD	55
2.3.2.1	Bestimmung einer PD durch eine erfolgreiche Auktion	56
2.3.2.2	Wiederholung nach erfolgloser Auktion	57
2.3.2.3	Wiederholung nach erfolgloser Reservierung	59
2.3.2.4	SRWU-Abgabe nach erfolgloser Auktion	60
2.3.2.5	SRWU-Abgabe nach erfolgloser Reservierung	61
2.3.2.6	Band-Management	62
2.3.2.7	SRWU-Scheduling	63
2.3.3	Übergabe eines Auftrags an eine PD	64
2.3.3.1	Übergabe einer gültigen SRWU	64
2.3.3.2	Ablehnung einer ungültigen SRWU	65
2.3.3.3	Starten einer SRWU durch den Dispatcher	66
2.3.4	Übergabe eines Response von einer PD	67
2.3.5	Data Collection	68
2.3.5.1	Collection Service	68

2.3.5.2	Aktualisieren des GIS durch den Distributed Broker Part	69
2.3.5.3	Sammeln lokaler PD-Daten durch die Data Collection Unit	70
2.3.5.4	Aktualisierung der globalen PD-Daten erzwingen	71
2.3.6	Zusammenfassung	72
3	Quality of Service und Service Level Agreements	73
3.1	QoS des kommerziellen Grid Resource Brokers	74
3.1.1	Deadline	74
3.1.2	Timeline	74
3.1.3	Budget	74
3.1.4	Velocity	74
3.1.5	Failure	74
4	Mikroökonomische Ansätze und Auktionsverfahren	75
4.1	Einleitung	75
4.2	Modell von Ferguson	76
4.3	HOP1	76
4.4	Auktionsverfahren	77
4.4.1	Sealed Bid Auction	77
4.4.2	Dutch Auction	77
4.4.3	English Auction	77
4.4.4	Hybrid Auction	77
4.4.5	Continuous Double Auction	78
4.5	Ökonomischer Ansatz beim kommerziellen Grid Resource Broker	78
4.5.1	Auktion beim kommerziellen Grid Resource Broker	79
4.5.1.1	Erstellung von Gebotsanfragen	79
4.5.1.2	Erstellung von Geboten	80
4.5.1.3	Auswertung von Geboten	81
4.5.1.4	Reservierungsablauf	82
4.5.2	Spezielles Auktionsverfahren	83
5	Implementierung und Simulation	84
5.1	GridSim und SimJava	84

<i>INHALTSVERZEICHNIS</i>	10
5.1.1 SimJava	84
5.1.2 GridSim	84
5.1.2.1 Netzwerksimulation	85
5.1.2.2 CPU-Simulation	85
5.1.2.3 Grid-Ressourcen	85
5.1.2.4 GIS	85
5.1.2.5 Grid Resource Broker	85
5.2 Implementierung des kommerziellen Grid Resource Brokers	86
5.2.1 Einleitung	86
5.2.2 GridSim-Modifikationen	87
5.2.2.1 Modifikation der Grid-Ressource	87
5.2.2.2 Modifikation der Netzwerkresource	88
5.2.3 Experiment	88
5.2.4 Statistik	88
6 Globus	89
6.1 Einleitung	89
6.2 Komponenten	90
6.3 Kommerzieller Grid Resource Broker und Globus	90
7 Performance Messung	91
7.1 Überblick	91
7.2 Referenzmessungen	91
7.3 Optimierung des Auktionsverfahrens	91
7.4 On-Demand Value	91
7.5 Optimierung durch die TimeTable	91
7.6 Anpassung der Bänder	91

Abbildungsverzeichnis

1.1	Transaktionsmaschine	15
2.1	Kommerzieller Grid Ressource Broker	20
2.2	Übersicht Grid-Komponenten	21
2.3	Versenden einer SRWU	21
2.4	Versenden eines Message Containers	22
2.5	Übermittlung von PD Informationen	22
2.6	Central Broker Part	23
2.7	Profiler	24
2.8	Profiler-Authentifizierung	25
2.9	Benutzerprofil	25
2.10	Bulletin Board	26
2.11	Collection Service	27
2.12	On-Demand Information Service	28
2.13	Message Container	29
2.14	Band Mapping Unit	30
2.15	Übersicht Auktion	31
2.16	Distributor, Receiver, Mediator	32
2.17	Distributed Broker Part	34
2.18	On-Demand Information Service	35
2.19	Auction Unit	36
2.20	Auktionsablauf auf dem Central Broker Part	37
2.21	Band Transformer	38
2.22	Bulletin Board	40
2.23	Time Table	41

2.24	Data Collection Unit	42
2.25	Dispatcher	43
2.26	SRWU Übersicht	44
2.27	SRWU-Broker Segment	44
2.28	SRWU-Id	44
2.29	State	45
2.30	Customer Service Element	45
2.31	Auction Settings	46
2.32	WLM-Settings	47
2.33	Bid	47
2.34	Use Case #01: Übersicht Kommerzieller Grid Resource Broker	49
2.35	Use Case #02: Verarbeitung einer SRWU	50
2.36	Use Case #03: Annahme einer gültigen SRWU	51
2.37	Use Case #04: Ablehnung einer ungültigen SRWU	53
2.38	Use Case #05: Vermittlung einer nicht verarbeitbaren SRWU	54
2.39	Use Case #06: Bestimmung einer PD durch eine erfolgreiche Auktion	56
2.40	Use Case #07: Wiederholung nach erfolgloser Auktion	57
2.41	Use Case #08: Wiederholung nach erfolgloser Reservierung	59
2.42	Use Case #09: SRWU-Abgabe nach erfolgloser Auktion	60
2.43	Use Case #10: SRWU-Abgabe nach erfolgloser Reservierung	61
2.44	Use Case #11: Band-Management	62
2.45	Use Case #12: SRWU-Scheduling	63
2.46	Use Case #13: Übergabe einer gültigen SRWU	64
2.47	Use Case #14: Ablehnung einer ungültigen SRWU	65
2.48	Use Case #15: Starten einer SRWU durch den Dispatcher	66
2.49	Use Case #16: SRWU-Scheduling	67
2.50	Use Case #17: Collection Service	68
2.51	Use Case #18: Aktualisierung des GIS durch den Distributed Broker Part	69
2.52	Use Case #19: Sammeln lokaler PD-Daten durch die Data Collection Unit	70
2.53	Use Case #20: Aktualisierung der globalen PD-Daten erzwingen	71
5.1	PD-Informationen	87

Tabellenverzeichnis

2.2	Übersicht Kommerzieller Grid Resource Broker	49
2.4	Verarbeitung einer SRWU	50
2.6	Annahme einer gültigen SRWU	52
2.8	Ablehnung einer ungültigen SRWU	54
2.10	Vermittlung einer ungültigen SRWU	55
2.12	Bestimmung einer PD durch eine erfolgreiche Auktion	57
2.14	Wiederholung nach erfolgloser Auktion	58
2.16	Wiederholung nach erfolgloser Reservierung	60
2.18	SRWU-Abgabe nach erfolgloser Auktion	61
2.20	SRWU-Abgabe nach erfolgloser Reservierung	62
2.22	Band-Management	63
2.24	SRWU-Scheduling	64
2.26	Übergabe einer gültigen SRWU	65
2.28	Ablehnung einer ungültigen SRWU	66
2.30	Starten einer SRWU durch den Dispatcher	67
2.32	Übergabe eines Response von einer PD	68
2.34	Collection Service	69
2.36	Aktualisierung des GIS durch den Distributed Broker Part	70
2.38	Sammeln lokaler PD-Daten durch die Data Collection Unit	71
2.40	Aktualisierung der globalen PD-Daten erzwingen	72
4.1	Erstellung von Gebotsanfragen	80
4.2	Erstellung von Geboten	81
4.3	Auswertung von Geboten	82
4.4	Reservierungsablauf	82

Kapitel 1

Einleitung

1.1 Die Transaktionsmaschine

1.1.1 Einleitung

Das Konzept der Transaktionsmaschine (TE) basiert auf Arbeiten von W. G. Spruth und J. Franz und wurde konzeptionell in [11] beschrieben. Die Transaktionsmaschine bildet die architekturelle Basis für ein Hochleistungs-Transaktionssystem. Dieses ermöglicht Geschäftsaufträge (z.B. Wertpapierorders, Zahlungen, etc.) auf der Basis einer serviceorientierten Verarbeitung mit möglichst hoher Leistung abzuarbeiten. Siehe hierzu auch die Arbeiten [13, 2, 22]. Die nachfolgende Abbildung zeigt die wesentlichen Architekturkomponenten der Transaktionsmaschine. Diese sind

- Die Processing Domain
- Der Channel Controller und Profiler
- Der Masterflow Controller
- Der Carrier/Subworkflow Controller
- Die Service Request Work Unit

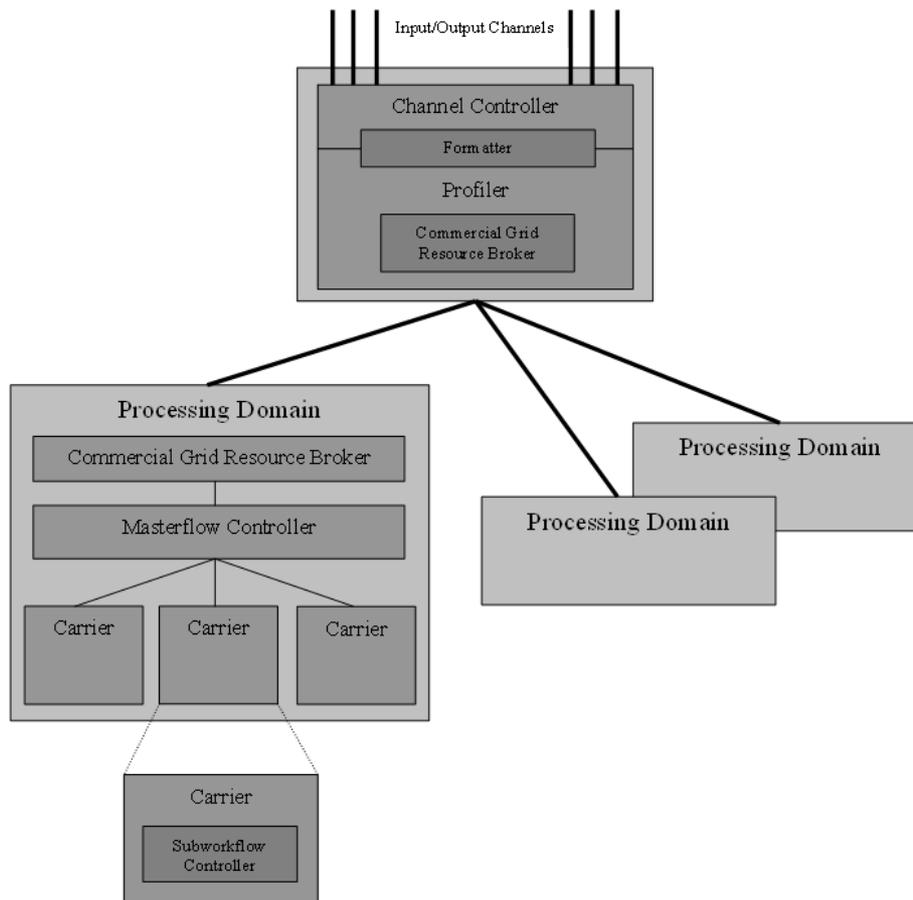


Abbildung 1.1: Transaktionsmaschine

In den folgenden Abschnitten werden die Komponenten der Transaktionsmaschine kurz beschrieben. Ein für die Verarbeitungssteuerung der TE zuständiger Algorithmus BDAfPD (Band Distribution Algorithm for Processing Domains) wird in Kapitel 1.2 auf Seite 17 erläutert.

1.1.2 Processing Domain

Die Transaktionsmaschine benutzt das Konzept der Processing Domain(s) (PD), um heterogene Computerressourcen oder Systeme in Form einer einheitlichen Computerleistung zu bündeln (clustern). Eine PD kann aus einer beliebigen Struktur, wie Einzelrechner, homogene Cluster oder heterogene Computersysteme bestehen, die sich als ein Verarbeitungszentrum darstellt und für einen Leistungsabnehmer eine bestimmte Verarbeitungsleistung (Service) anbietet [11]. Die interne Verwaltung erfolgt völlig autonom nach systemspezifischen Strategien. Die Processing Domain stellt dabei keine Computerressourcen im klassischen Sinn dar. Sie übernimmt stattdessen die Aufgabe eines Leistungserbringers (Service Providers). Dieser bietet einen Verarbeitungsservice an, der in Service Units (SU) gemessen wird. Für die eigentliche Verarbeitung sind spe-

zielle Komponenten (Carrier), wie Workload Manager oder Workflow Controller auf der PD verantwortlich. Der durch eine Service Unit definierte Service setzt sich aus einer Reihe von elementaren Basiseinheiten zusammen. Zu diesen können CPU-, I/O-, Hauptspeicher- und Multiprogramming-Einheiten gehören. Für die Bestimmung des gesamten Service erhält jede Basiseinheit ein individuelles Servicegewicht. Dieses ist von der Leistungsfähigkeit der jeweiligen PD abhängig. Ebenso spielt der Preis für die Basiseinheit eine Rolle. Das Angleichen der Basiseinheiten auf einer PD und bezüglich einer Referenz PD bildet eine normierte Abrechnungsgröße für den verbrauchten Service in einer definierten Zeiteinheit (SU). Die Dauer eines Zeitintervalls (Period) wird als Service Intervall (Service Period) bezeichnet und ist für das gesamte System eindeutig. Eine PD bietet ihre Serviceleistungen in einem Processing Domain Profile (PDP) an. Neben der Anzahl an SUs, die verarbeitet werden können, werden auch Angaben über Kosten, Sicherheitsstufen, unterstützte Produkte, Verfügbarkeit, etc. der PD im PDP abgebildet. Auf dieser Basis lassen sich Vereinbarungen über die Verarbeitungsqualität (Quality of Service) treffen. Die Aufträge werden in Form von Service Request Work Units (SRWU) übergeben.

1.1.3 Channel Controller und Profiler

Der Channel Controller/Profiler (CC/P) ist eine Kommunikations- und Vermittlungseinheit zwischen Leistungsabnehmer (externen Systemen, Auftragsgeber) und den Verarbeitungskomponenten der Transaktionsmaschine. Der Channel Controller überwacht unterschiedliche Ein- und Ausgabekanäle (z.B. Internet, Intranet, FTP usw.) und nimmt Aufträge in verschiedenen Formaten entgegen [11]. Diese werden von einem Formatter in das neutrale Format einer Service Request Work Unit (SRWU) umgewandelt. Dadurch wird eine schnelle und einheitliche Verarbeitung ermöglicht. Für eine Antwort (SRWU Response) wird ein entsprechendes Ausgabeformat erstellt und durch den Channel Controller an den Leistungsabnehmer übertragen. Die Vermittlung der SRWUs an die PDs erfolgt über einen Meta-Scheduling Algorithmus den Band Distribution Algorithm for Processing Domains (BDAfPD), der von dem nachfolgend beschriebenen kommerziellen Grid Resource Broker umgesetzt wird. Dazu ermittelt der Profiler die genauen Anforderungen der Leistungsabnehmer (Service Level Agreements-SLA) einer SRWU, die durch den BDAfPD eingehalten werden müssen. Der Broker sammelt Informationen über die Eigenschaften einer PD und gleicht diese mit den SLAs ab (matching).

1.1.4 Masterflow Controller

Der Masterflow Controller (MFC) ist für die Verwaltung und Ausführung einer SRWU innerhalb einer PD zuständig. Er generiert Subworkflow Prozesse und startet diese innerhalb einer geeigneten Verarbeitungsumgebung (Carrier). Dabei werden Eigenschaften wie Parallelisierung und Fehlerbehandlung unterstützt. Für eine optimierte Verarbeitung konfiguriert der MFC den lokalen Workload Manager entsprechend den Anforderungen der SRWU [11].

1.1.5 Carrier/Subworkflow Controller

Ein Carrier bildet eine gesicherte Transaktionsumgebung für die Ausführung von Subworkflow Prozessen, die vom MFC generiert werden. Der Subworkflow Controller

(SWFC) nutzt diese Umgebung um technische Transaktionen auf Workflow-Ebene durchzuführen. Dabei müssen die ACID-Eigenschaften sichergestellt werden [13].

1.1.6 Service Request Work Unit

Die Service Request Work Unit (SRWU) repräsentiert ein universelles Format zur Definition von Aufträgen, die verarbeitet werden sollen. Sie enthält für jede Komponente der Transaktionsmaschine ein Segment mit Informationen (z.B. Master- und Subworkflow Definitionen), über welche die Abarbeitung gesteuert wird. Das Customer Service Element (CSE) definiert einen Bereich, der die Business-Anforderungen der SRWU aufnimmt. Dazu gehören neben einer Auftraggeberkennung, ein SLA-Profil, der Typ der SRWU, die Anzahl geschätzter Service Units und der Typ des Bandes, in welchem die SRWU verarbeitet werden soll.

1.2 Band Distribution Algorithm for Processing Domains

Der Band Distribution Algorithm for Processing Domains (BDAfPD) ist ein Meta-Scheduling Algorithmus für die Verteilung von SRWUs in einem Netzwerk von PDs unter Einhaltung eines QoS. Der BDAfPD erlaubt eine serviceorientierte Sicht auf die Ressourcen, unabhängig von deren Architektur. Ermöglicht wird dieser Ansatz durch die Definition von Produktklassen, die als Bänder bezeichnet werden. Diese sind in der Lage jeweils einen bestimmten QoS zu erfüllen. Anhand ihres Processing Domain Profiles können für jede Processing Domain die möglichen Bänder ermittelt werden. Geschäftsprodukte bzw. deren Geschäftsprozesse (SRWUs) mit ähnlichen SLA-Definitionen werden in einem Band zusammengefasst, das für die Verarbeitung am besten geeignet ist. Eine mögliche Einteilung könnte wie folgt aussehen. Spezialprodukte, mit hohen Qualitätsanforderungen und höheren Preisen, Standardprodukte mit geringeren Qualitätsanforderungen und moderaten Preisen sowie Lightprodukte mit sehr geringen Qualitätsanforderungen und niedrigen Preisen. Der BDAfPD gestattet den PDs sich optimal auf die definierten Bänder einzustellen. Eine Möglichkeit besteht darin nur solche Bänder zu unterstützen, für welche die PD optimal geeignet ist, zum Beispiel ein-/ausgabeintensive Produkte auf PDs mit hoher I/O-Leistung. Eine andere Strategie könnte die Spezialisierung einer auf die Produkte mit dem höchsten PD-Gewinn sein. Standardprodukte können oft sehr schnell verarbeitet werden und bringen durch ihre hohe Anzahl meist den gewünschten Gewinn. Ebenso kann aber auch ein Mix aus verschiedenen Bändern angeboten werden. Der BDAfPD bietet die Möglichkeit sich dynamisch an die jeweiligen Bedingungen anzupassen. Sollte in einem Band Kapazität ungenutzt bleiben, kann diese in andere Bänder verschoben werden. Auftretende Lastspitzen werden so abgefangen. Ein weiteres Konzept des BDAfPD ist die Vermeidung von Strafgeldern. Durch die Angabe der Kapazität der Bänder in Service Units werden nur so viele SRWUs angenommen, wie verarbeitet werden können. Die Bestimmung eines Performance Index (PI) ermöglicht eine noch bessere Einhaltung des allgemeinen QoS.

1.3 Aufgabenstellung und Ziele der Diplomarbeit

Die Aufgabe dieser Diplomarbeit ist der Entwurf eines kommerziellen Grid Resource Brokers auf Basis der Globus Infrastruktur. Dieser soll als Teil der Transaktionsmaschine arbeiten und einen kommerziellen Meta-Scheduling Algorithmus (Band Distribution Algorithm for Processing Domains) umsetzen. Dafür ist der Entwurf eines aktiven Collection Service notwendig, der in Verbindung mit einem Grid Information Service arbeitet.

Eine teilweise Umsetzung des kommerziellen Grid Resource Brokers erfolgt in Java auf Basis von Simjava 1.2 und GridSim 2.1.

Diese Diplomarbeit hat das Ziel, einen kommerziellen Grid Resource Broker auf Basis der Globus Infrastruktur zu entwerfen. Dieser soll als Teil der Transaktionsmaschine im Channel Controller/Profiler arbeiten und dabei das

Kapitel 2

Der kommerzielle Grid Resource Broker

2.1 Einleitung

Bei einem Grid handelt es sich um eine heterogene Umgebung von Computerressourcen mit unterschiedlichen Hardwarearchitekturen, Betriebssysteme, Softwareangebote und Verwaltungsstrategien, die über ein Netzwerk miteinander verbunden sind. Ein wesentliches Merkmal ist die Globalität der Ressourcen, d.h. sie müssen weder im selben administrativen Bereich liegen noch lokal erreichbar sein. In einem serviceorientierten Grid können Provider Leistungen anbieten, die von Benutzern genutzt werden. Als Vermittler zwischen beiden Parteien kommen Grid Resource Broker zum Einsatz. Ihre Aufgabe besteht darin, freie Ressourcen im Grid zu finden und die Jobs oder Aufträge der Benutzer zuzuweisen.

Der kommerzielle Grid Resource Broker (kGRB) unter Benutzung des BDAfPD verwendet Grid-Technologien, um ein Netzwerk von Processing Domains zu verwalten. Durch das Einsetzen unterschiedlicher Bänder entsteht ein heterogenes Umfeld, das durch den BDAfPD optimiert wird.

2.2 Aufbau

Der kommerzielle Grid Resource Broker setzt sich aus mehreren Brokerinstanzen in einer verteilten Brokerumgebung zusammen. Die Hauptkomponenten sind der Central Broker Part (CBP) und der Distributed Broker Part (DBP). Beide haben fest definierte Aufgaben. Der DBP ist für die lokale Optimierung einer PD zuständig. Er sorgt für die Einhaltung und Anpassung der Bänder, die Durchführung von Auktionen so-

wie die Planung der Ausführung von SRWUs unter Berücksichtigung des QoS. Auf jeder PD muss sich eine Instanz des DBP befinden. Der Central Broker Part optimiert das gesamte System. Die Hauptaufgabe liegt in der Unterstützung der Distributed Broker Part-Instanzen durch Kontrolle und Verteilung ankommender SRWUs und durch die Leitung von Auktionen. Nach außen verbirgt der Central Broker Part die innere Struktur des Systems und präsentiert sich als ein Leistungsanbieter (Service-Provider) oder Vermittler (Mediator). Der systemweit angebotene Service wird in einem Service Provider Element (PSE) veröffentlicht. Die Einträge im PSE repräsentieren eine Zusammenfassung aus den statischen und dynamischen Profilen der PDs.

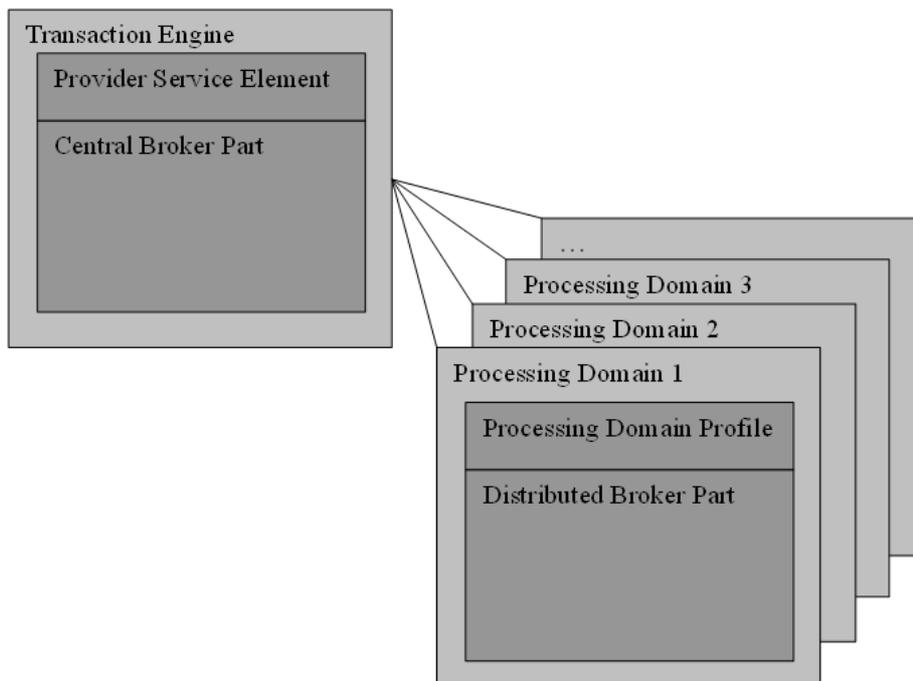


Abbildung 2.1: Kommerzieller Grid Ressource Broker

Der kommerzielle Grid Resource Broker setzt auf einer Grid-Infrastruktur auf. Die dort angebotenen Dienste werden für den Daten- und Nachrichtenverkehr genutzt. Der Broker setzt eine verallgemeinerte Grid-Umgebung mit folgenden Komponenten voraus

- Job Transfer Service (JTS)
- Grid Information Service (GIS)

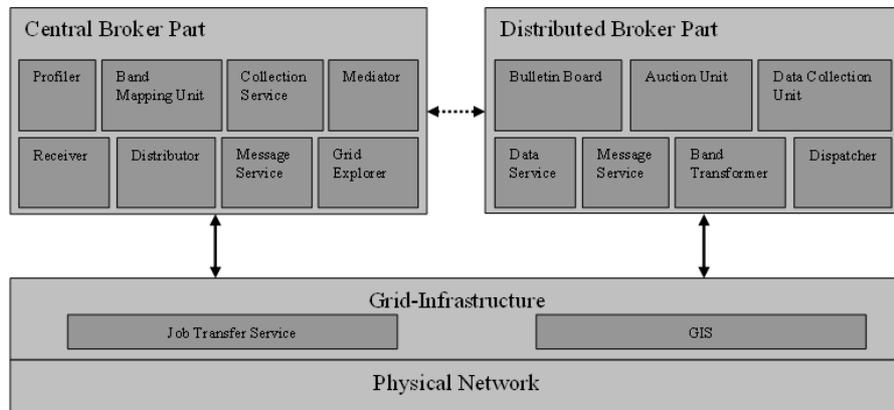


Abbildung 2.2: Übersicht Grid-Komponenten

Der Job Transfer Service gestattet den Transport beliebiger Aufträge zwischen Central Broker Part und Distributed Broker Part. Versendet werden ausschließlich Service Request Work Units und Nachrichten in Form von Message Container (siehe Kapitel 2.2.1.4 auf Seite 28).

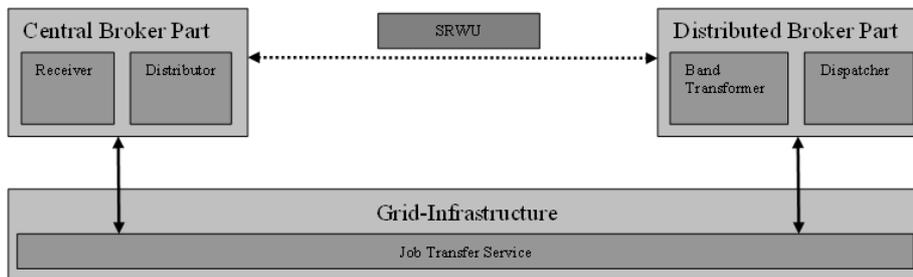


Abbildung 2.3: Versenden einer SRWU

Die Message Container dienen hauptsächlich der Übermittlung von Auktionsinformationen (z.B. Gebote einer PD, Reservierungen, etc.) zwischen der Band Mapping Unit des Central Broker Parts und der Auction Unit des Distributed Broker Parts. Aufgrund ihres generischen Formats kommen die Message Container auch bei der Nachrichtenübermittlung anderer Broker-Komponenten zum Einsatz (siehe Kapitel 2.2.1.3 auf Seite 27).

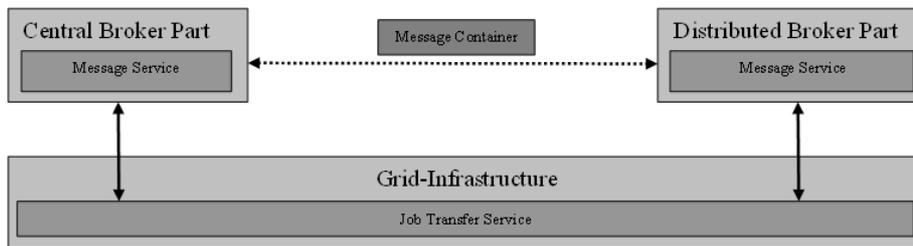


Abbildung 2.4: Versenden eines Message Containers

Der Grid Information Service bildet einen zentralen Sammelpunkt für Informationen über PDs (z.B. unterstützte Bänder, Preise, Auslastung, etc.). Der Distributed Broker Part versorgt den Dienst mit den entsprechenden Daten, so dass der Central Broker Part auf diese zugreifen kann.

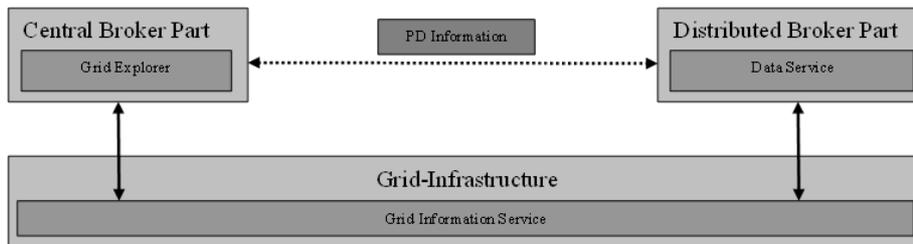


Abbildung 2.5: Übermittlung von PD Informationen

Aktuelle Grid-Lösungen wie Globus-Toolkit oder Legion bieten Schnittstellen an, die die benötigten Dienste bereits zur Verfügung stellen bzw. deren Umsetzung ermöglichen [9, 12, 18, 19, 24]. Dienste wie JMS erlauben eine standardisierte Umsetzung des Message Services.

2.2.1 Central Broker Part

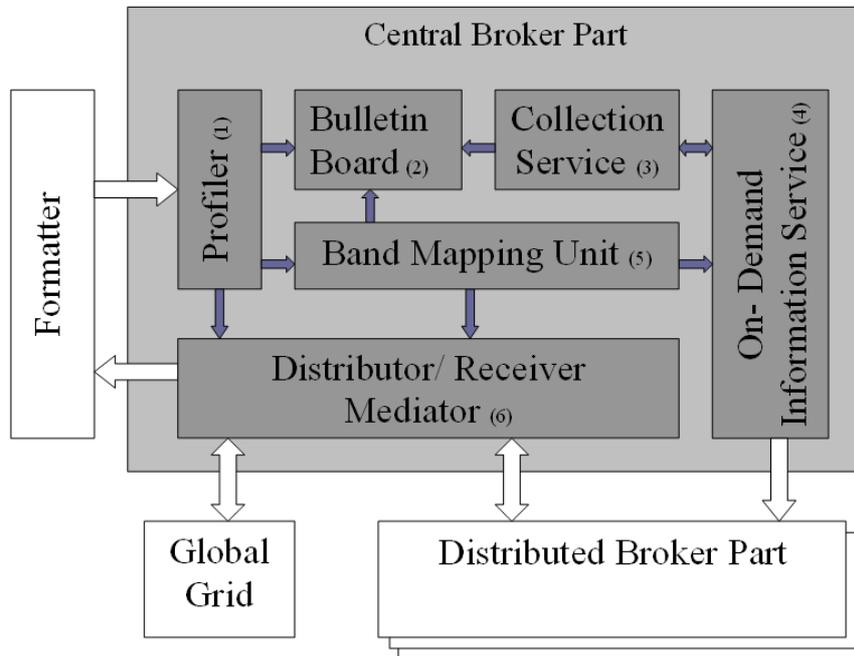


Abbildung 2.6: Central Broker Part

Der Central Broker Part besteht aus mehreren Komponenten, die in unterschiedlichen Threads laufen können und über ein periodisches Signal synchronisiert werden. Dieses wird vom Central Broker Part im Abstand eines Service Intervalls erzeugt und an die Broker-Komponenten übergeben. Die Dauer eines Service Intervalls ist systemweit eindeutig und wird durch eine Service Unit definiert.

2.2.1.1 Profiler

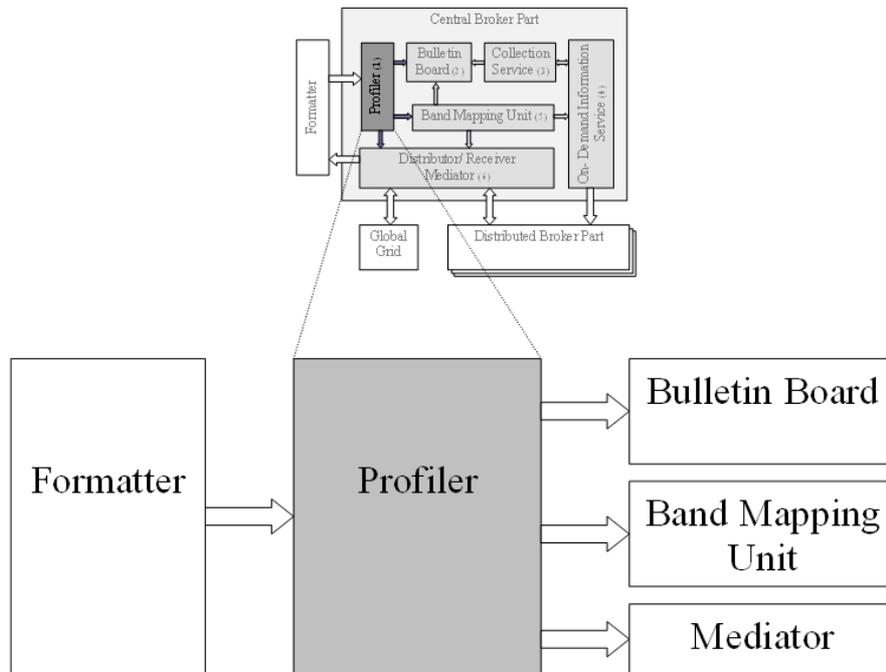


Abbildung 2.7: Profiler

Die Transaktionsmaschine ist in der Lage Aufträge in unterschiedlichen Formaten entgegen zu nehmen. Ein Formatter erstellt daraus Service Request Work Units. Diese werden anschließend dem Profiler übergeben. Zu dessen Aufgaben gehört neben der Authentifizierung auch die Verwaltung der Benutzerprofile von Leistungsabnehmern. Zwischen authentifizierten Leistungsabnehmern und dem kommerziellen Grid Resource Broker bestehen Verträge in Form von Service Level Agreements (SLA). Diese sind in den Benutzerprofilen abgelegt. Für jeden Leistungsabnehmer ist genau definiert, welche Auftragsstypen (SRWU-Typen) akzeptiert werden, die Qualität der Verarbeitung sowie Abrechnungs- und Preisinformationen (siehe Kapitel 2.2.3.3 auf Seite 46).

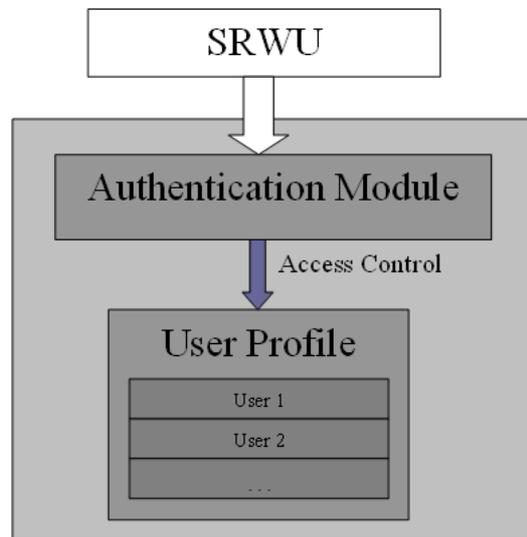


Abbildung 2.8: Profiler-Authentifizierung

Jeder SRWU eines Typs wird ein Standardqualitätsprofil zugeordnet. Zusätzlich können noch Varianten vereinbart werden. Dabei handelt es sich um Abweichungen vom Standardqualitätsprofil unter Berücksichtigung des Preises. Der Zugriff wird durch den Authentifizierungsmechanismus kontrolliert. Wenn eine Berechtigung vorliegt, wird der SRWU das gewünschte SLA-Profil zugeordnet.

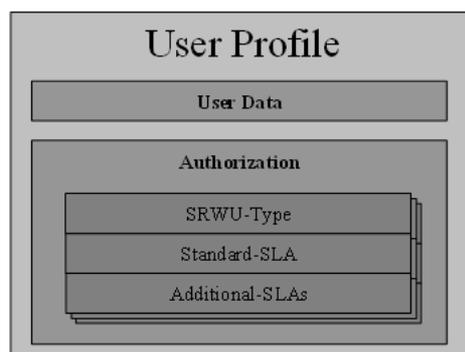


Abbildung 2.9: Benutzerprofil

Die benötigte Verarbeitungsleistung einer SRWU wird in Service Units angegeben. Der Profiler bietet einen Mechanismus an, welcher aus historischen Messdaten die voraussichtliche Anzahl an SUs abschätzt, die für eine Verarbeitung nötig sind. Die notwen-

digen Informationen werden von einem Collection Service gesammelt und im Bulletin Board abgelegt.

2.2.1.2 Bulletin Board

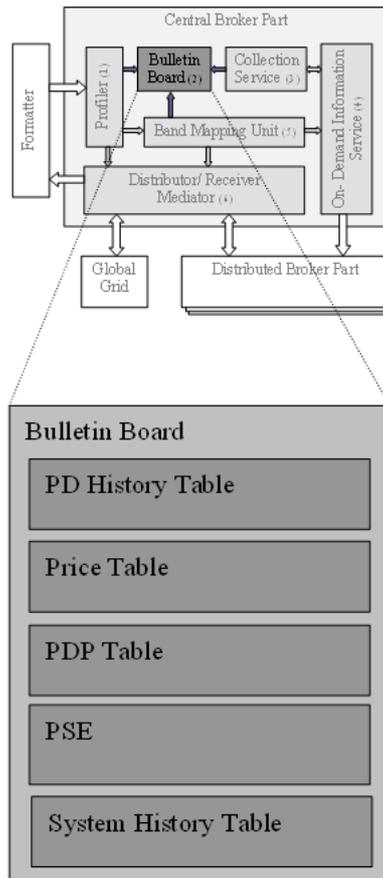


Abbildung 2.10: Bulletin Board

Das Bulletin Board bietet Strukturen an, um Informationen zu speichern und anzubieten, die allen Komponenten des Central Broker Parts zugänglich gemacht werden sollen. Wichtig sind vor allem Daten über die Auslastung der PDs und deren Einhaltung des QoS. PDs können aber auch über den On-Demand Information Service ihr PDP veröffentlichen oder aktuelle Preisinformationen abgeben. Der Central Broker Part bestimmt auf dieser Grundlage sein Serviceangebot und kann Optimierungen durchführen. Dynamische Daten, wie Auslastungsinformationen, werden in einer History Table abgelegt. Diese Struktur ermöglicht das Sammeln und Ablegen von Informationen über einen längeren Zeitraum. Innerhalb einer Datenreihe ist somit auch der

Zugriff auf ältere (historische) Werte möglich, die bereits durch neuere ersetzt wurden. Als Speichermedium dient eine Liste mit einer einstellbaren festen Größe. Neue Werte werden an die erste Stelle der Liste geschrieben und die älteren Einträge um eine Stelle nach hinten verschoben. Wenn die Kapazität nicht mehr ausreicht, wird nach dem FIFO-Prinzip der älteste Wert entfernt. Der Mechanismus orientiert sich an dem Verfahren, das bei dem Workload Manager in z/OS-Systemen zum Einsatz kommt [1, 21].

2.2.1.3 Collection Service

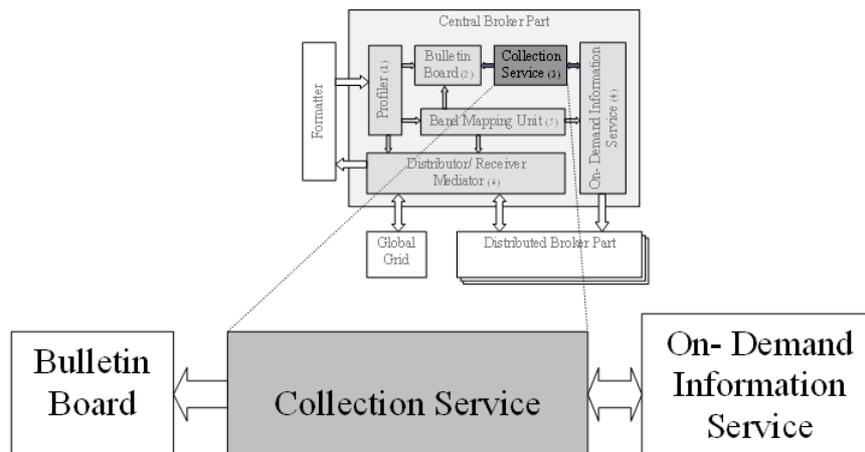


Abbildung 2.11: Collection Service

Der Collection Service ist ein Dienst, der von den Komponenten des Central Broker Parts benutzt wird, um aktiv Daten über die PDs im System zu sammeln. Diese werden anschließend im Bulletin Board abgelegt. Der Zugriff auf die PD-Daten erfolgt über den On-Demand Information Service. Bei den abgerufenen Informationen handelt es sich hauptsächlich um Daten, die statisch sind oder eine geringe Dynamik aufweisen. Dazu gehört zum Beispiel das Serviceangebot einer PD, Qualitätsangaben (QoS), unterstützte Bänder, Preisübersichten sowie Mittelwerte über die Auslastung der PD und Verarbeitungsgeschwindigkeit der SRWUs. Der Aufruf des Collection Service erfolgt periodisch oder auf Anfrage einer Broker-Komponente. Die Dauer des Intervalls ist sehr stark von der Architektur und Größe des gesamten Systems abhängig. Der zusätzliche Nachrichtenverkehr muss unbedingt beachtet werden, da hier wichtige Netzwerkre-sourcen verloren gehen. Aufgrund der statischen Natur der Daten empfiehlt sich die Verwendung eines relativ großen Intervalls von mehreren Sekunden oder Minuten. Sollten größere Änderungen auf einer PD auftreten hat diese die Möglichkeit selbständig den Collection Service zu informieren.

2.2.1.4 On-Demand Information Service

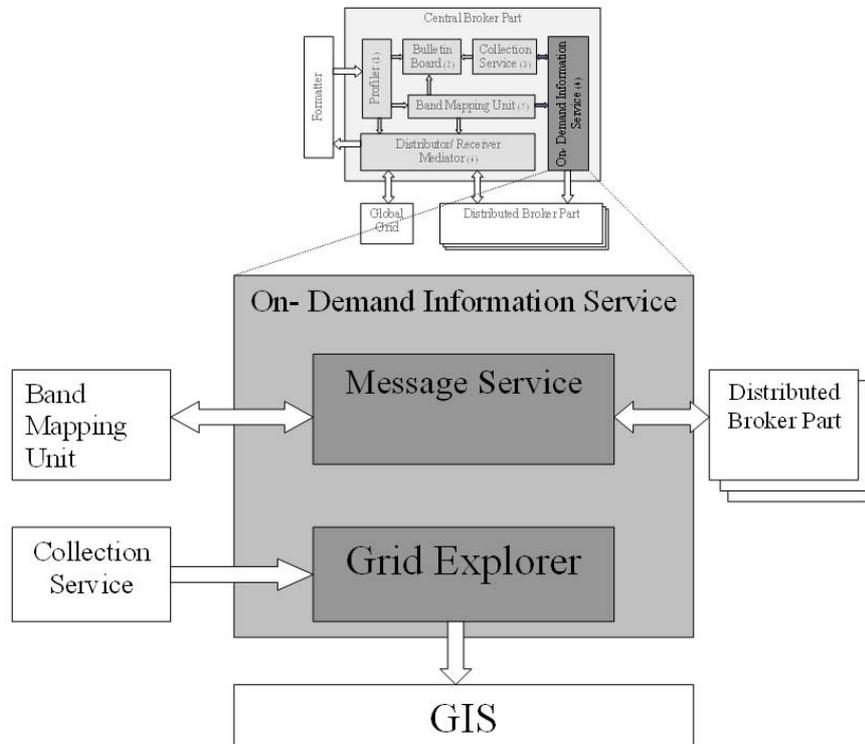


Abbildung 2.12: On-Demand Information Service

Der On-Demand Information Service ist der Kommunikationsdienst des Central Broker Parts. Der gesamte Nachrichtenverkehr mit den PDs erfolgt über diese Komponente. Besonders wichtig ist eine schnelle und dennoch flexible Arbeitsweise, die durch den Einsatz spezieller Message Container erreicht wird. Dieses Datenformat ermöglicht durch seine homogene Struktur ein universelles Verarbeiten unterschiedlicher Nachrichten, unabhängig vom Sender bzw. Empfänger. Die eigentliche Nachricht ist in einem Datenteil untergebracht.

Der Header des Message Containers enthält alle Informationen, um ausgehende Nachrichten an die richtige PD zu senden bzw. eingehende Nachrichten an die entsprechende Broker-Komponente zu übermitteln. Für den eigentlichen Transport in der Grid-Umgebung kann der Message Service auf Standardverfahren zurückgreifen (z.B. JMS). Ein weiterer wichtiger Bestandteil des On-Demand Information Service ist der Grid Explorer. Dieser stellt einen erweiterten Grid Information Service (GIS) dar, der es ermöglicht Informationen über die PDs des Systems zu sammeln. Auf diese Weise

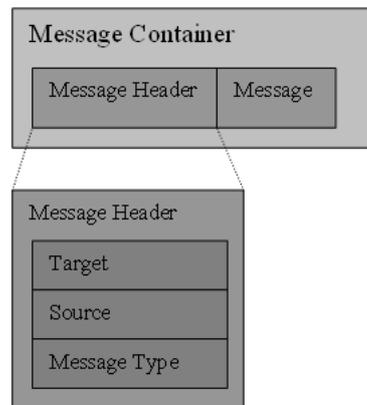


Abbildung 2.13: Message Container

kann der Collection Service mit Daten versorgt werden. Der Grid Explorer wird aber auch benutzt um die Verfügbarkeit von PDs zu prüfen. Dies ist besonders dann wichtig, wenn eine starke Dynamik im System vorliegt und PDs beliebig aus dem System entfernt werden können.

2.2.1.5 Band Mapping Unit

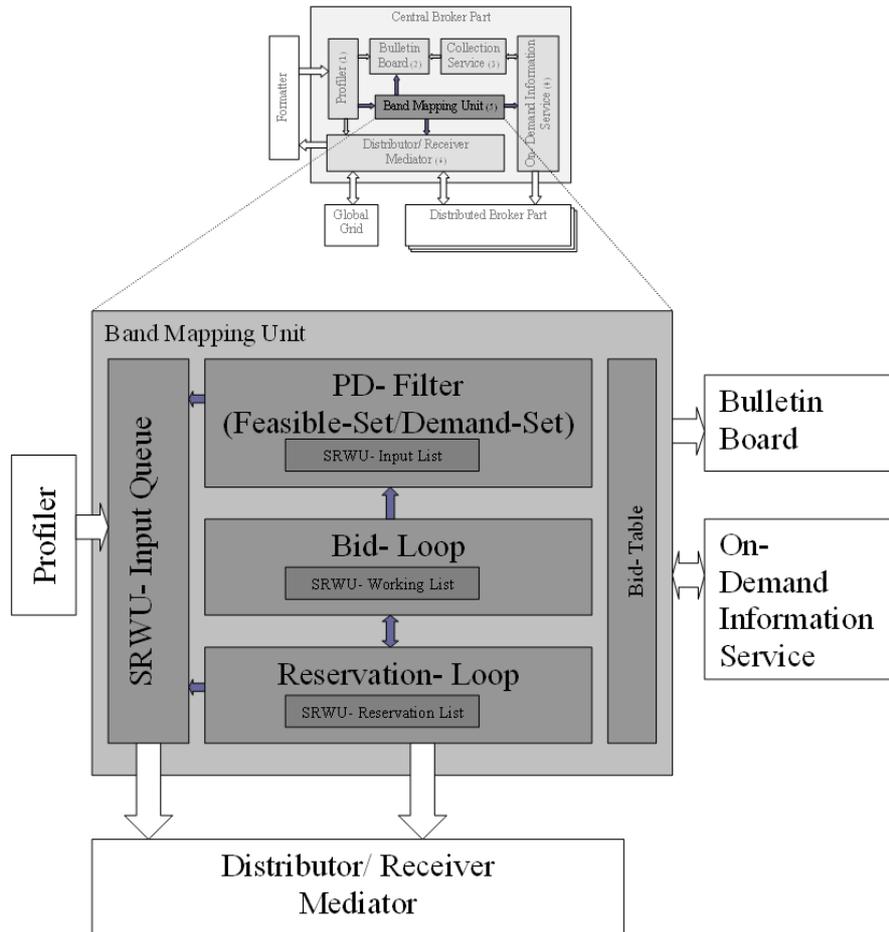


Abbildung 2.14: Band Mapping Unit

Die Band Mapping Unit bildet den Kern des Central Broker Parts. Im Mittelpunkt steht ein Algorithmus, der unter ökonomischen Gesichtspunkten SRWUs auf die PDs des Systems verteilt (BDAfPD). Die Hauptaufgabe ist die Maximierung des Gewinns durch Umsetzung des BDAfPD. Dieser überprüft die Eigenschaften der PDs anhand ihres PD Profiles und bildet diese auf die Bänder ab. Somit kann die Einhaltung des QoS der Bänder gewährleistet werden. Neben den statischen Merkmalen aus den PDPs spielen aber auch noch andere Faktoren, bei der Einhaltung des Quality of Service, eine wichtige Rolle. Einen großen Einfluss hat die Auslastung der PD. Ein Load Balancing Algorithmus versucht daher den Workload möglichst gleichmäßig zu verteilen. Verwendet wird eine Kombination aus einem klassischen zentralisierten Lastverteilungsverfahren und einem adaptierten Auktionsmechanismus.

Die Band Mapping Unit arbeitet in einer getakteten Schleife, d.h. dass die Ausführung

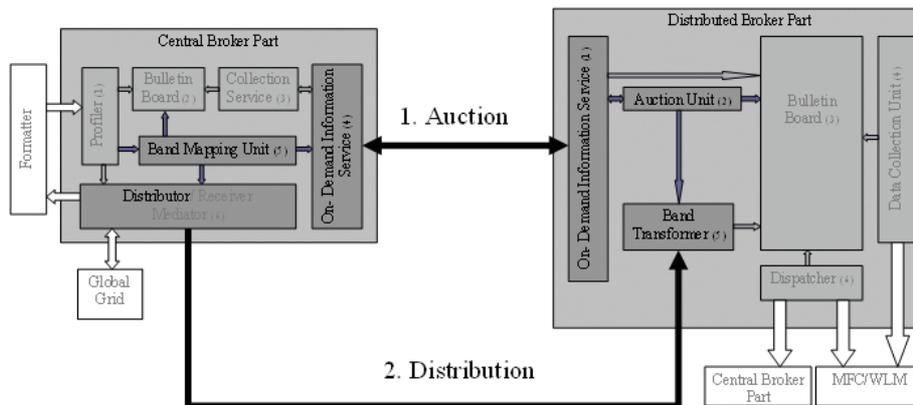


Abbildung 2.15: Übersicht Auktion

periodisch gestartet wird. Der Zeitraum entspricht einem Service Intervall. SRWUs, die nach dem Start eintreffen, werden in einer Input-Queue zwischengespeichert und im nächsten Durchlauf berücksichtigt. Dieser Sammelmechanismus erlaubt das Verarbeiten ganzer Gruppen von Service Request Work Units. Dadurch wird zum einen der Kommunikationsaufwand reduziert und zum anderen sind sehr schnell Rückschlüsse auf die Ankunftsdaten im Service Intervall möglich. Der Auktionsmechanismus erwartet die Bildung eines Feasible-Sets und eines Demand-Sets. Dabei handelt es sich um eine Vorauswahl von PDs, die an der Versteigerung teilnehmen dürfen. Bei der Bestimmung des Demand-Sets wird ein Teil des BDAfPD umgesetzt. Dieser liest Informationen über die PDs aus dem Bulletin Board, die vom Collection Service gesammelt wurden. Nur PDs mit entsprechender Band-Unterstützung werden im Demand-Set aufgenommen. Die dynamische Anpassung der Bandkapazitäten und die Umsetzung der SLAs unter Einhaltung des QoS erfolgen nicht in der Band Mapping Unit. Diese Teile des BDAfPD liegen in der Verantwortung der PDs. Die Umsetzung erfolgt in den Komponenten des Distributed Broker Parts (siehe Kapitel 2.2.2.3 auf Seite 38). Der Demand-Set bildet somit einen Filter für die Bänder der PDs. Da diese durch den Distributed Broker Part dynamisch an den Eingangsdatenstrom angepasst werden, beeinflusst der Demand-Set die Arbeitsweise der PDs. Für eine erhöhte Flexibilität wird die Funktion zur Bestimmung des Demand-Sets über ein Plugin realisiert. Der sich anschließende Auktionsmechanismus kann ebenfalls individuell festgelegt werden. Durch die Verwendung eines Plugins können beliebige Auktionsvarianten implementiert werden.

2.2.1.6 Distributor

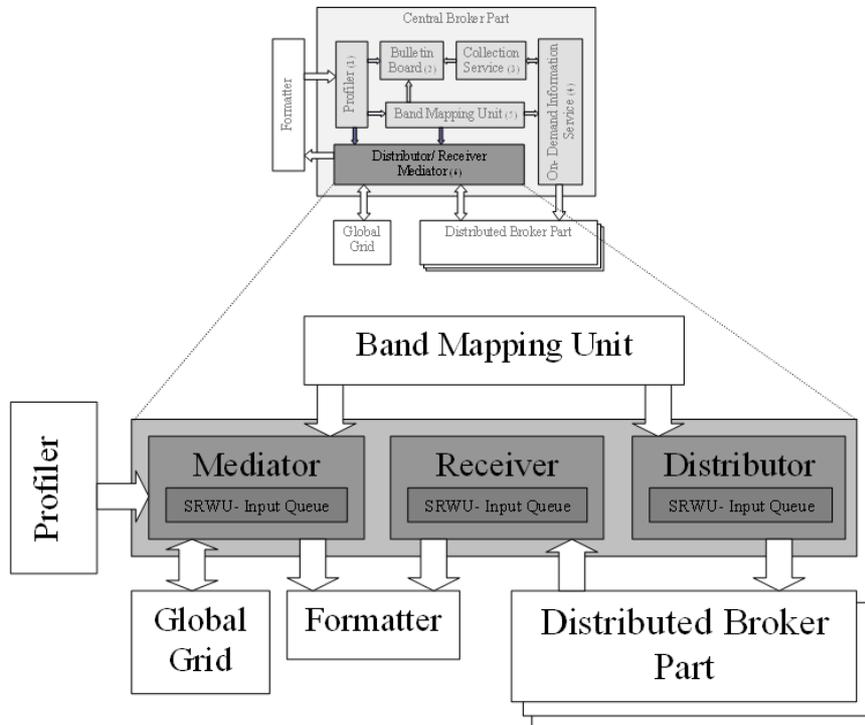


Abbildung 2.16: Distributor, Receiver, Mediator

Der Distributor empfängt SRWUs von der Band Mapping Unit und übermittelt diese an eine festgelegte PD. Die Komponente muss daher die Schnittstelle des Transportmediums kennen und die entsprechenden Funktionen umsetzen. Eine weitere wichtige Aufgabe ist die Datenflusskontrolle. Langsame Transportmedien, wie sie in Grid-Umgebungen vorkommen können, oder langsame Empfänger sollen keinen Einfluss auf die Performance der Band Mapping Unit haben. Der Distributor stellt daher eine Input-Queue als Puffer zur Verfügung. Diese wird von der Band Mapping Unit verwendet, um SRWUs zu übergeben. Der Distributor kann anschließend die nötigen Schritte für das Versenden einleiten, während die Band Mapping Unit ihre Arbeit fortsetzt.

2.2.1.7 Receiver

Der Receiver ermöglicht dem Central Broker Part das Empfangen einer SRWU, die von einer PD gesendet wurde. Die Komponente kennt die Schnittstelle des Transportmediums und bietet über diese einen entsprechenden Empfängerdienst an. Die Hauptaufgabe des Receivers ist das Empfangen von Response-SRWUs. Diese enthalten Informationen über verarbeitete Aufträge einer PD. Der Receiver übergibt die Response-SRWUs an den Formatter, der eine formatierte Ausgabe für den Leistungsabnehmer erstellt. Eine andere Aufgabe des Receivers ist das Empfangen abgelehnter SRWUs

von den PDs. Das sind Aufträge, die aufgrund von Verzögerungen (z.B. durch hohe Netzauslastung) nicht den vereinbarten Empfangszeitpunkt eingehalten haben. Der Receiver übergibt diese SRWUs der Band Mapping Unit, um die Möglichkeit einer Neuverteilung zu untersuchen.

2.2.1.8 Mediator

Der Mediator ist eine Vermittlungseinheit zu anderen Service Providern außerhalb des Verwaltungsbereichs des kommerziellen Grid Resource Brokers (z.B. in ein globales Grid). Diese Aufgabe wird auf Provisionsbasis durchgeführt, d.h. für eine erfolgreiche Vermittlung muss eine Gebühr von dem Leitungsabnehmer bezahlt werden. Langfristig wird in den meisten Fällen eine lokale Verarbeitung der SRWUs einen höheren Gewinn einbringen, da Leitungsabnehmer über einen längeren Zeitraum auf den Einsatz eines Vermittlers verzichten werden und einen direkten Zugang wählen. Der Mediator sollte daher nur zum Abfangen von Lastspitzen zum Einsatz kommen bzw. für Produkte deren Verarbeitung nicht den erwünschten Gewinn einbringt. Vor allem langläufige Prozesse mit geringen Gewinnerwartungen sollten hier beachtet werden, wenn in der nächsten Zeit profitablere Aufträge erwartet werden. In einer weltweiten Grid-Umgebung könnten somit Ressourcen belegt werden, die aufgrund von Zeitverschiebungen keine hohe Auslastung aufweisen. Der Einsatz des Mediators als freier Vermittler ist derzeit für reale Umgebungen eher unrealistisch. Das größte Problem stellt heute die Abrechnung von erbrachten Leistungen dar. Derzeitige Entwicklungen in Grid Infrastrukturen und Grid Software erweitern aber ständig die Möglichkeiten. Erste praktische Umsetzungen einer marktorientierten globalen Grid-Architektur werden im Rahmen des Gridbus-Projekts durchgeführt. Über einen einfachen Grid Resource Broker (Nimrod/G) mit grafischer Benutzeroberfläche können weltweit verteilte Ressourcen allokiert werden. Dabei werden einfache Optimierungen der Kosten bzw. der Verarbeitungsgeschwindigkeit (Budget/Deadline-Constraints) durchgeführt. Nimrod/G wurde bereits erfolgreich in einer realen Grid-Umgebung getestet [3, 4, 6]. Weiter Projekte waren z.B. Spawn, Mariposa und POPCORN. Um Serviceleistungen in einer globalen Grid-Umgebungen anbieten bzw. nutzen zu können sind Standardisierungen nötig [15]. Die Open Grid Service Architecture (OGSA) ist bemüht, auf der Basis von Web Services, eine standardisierte Grid Architektur zu beschreiben [16, 10]. Dabei kommen Standards wie Simple Object Access Protocol (SOAP), Web Service Description Language (WSDL) und Web Service Inspection Language (WSIL) zum Einsatz. SOAP ist für die Unterstützung eines Messaging Service zuständig. Auf der Basis von XML erlaubt SOAP die Kommunikation zwischen einem Leistungsanbieter und einem Leistungsabnehmer über eine Vielzahl von Transportprotokollen wie HTTP, FTP, Java Messaging Service (JMS), etc. WSDL und WSIL basieren ebenfalls auf XML. Während WSDL für die Schnittstellenbeschreibung eines Leistungsanbieters zuständig ist, können über WSIL in dezentralisierter Form die Serviceleitungen eines Leistungsanbieters abgefragt werden. Für die Erstellung einer robusten Grid-Umgebung sieht die OGSA eine Vielzahl an Schnittstellen vor, die unterschiedliche Aufgaben erfüllen [9].

GirdService-portType Erlaubt das Auffinden eines Service und das Management seiner Gültigkeitsdauer (für den Fall, dass sich der Service ändert).

Notification-Source-portType Erlaubt das Senden von Benachrichtigungen (z.B. Statusänderungen zwischen verschiedenen Services).

Notification-Sink-portType Erlaubt das Empfangen von Benachrichtigungen.

Notification-Subscription-portType Erlaubt das Festlegen von Benachrichtigungen, die für eine bestimmte Dauer durch den Notification-Source-portType gesendet werden dürfen.

Registration-portType Erlaubt es einer Service-Instanz, sich für einen öffentlichen Zugriff zu registrieren bzw. zu sperren.

Factory-portType Erlaubt das dynamische Erstellen von Service-Instanzen, über die die gesamte Serviceaktivität abgehandelt wird.

HandleResolver-portType Ermittelt über einen Service-Handle die Referenz auf eine Service-Instanz.

2.2.2 Distributed Broker Part

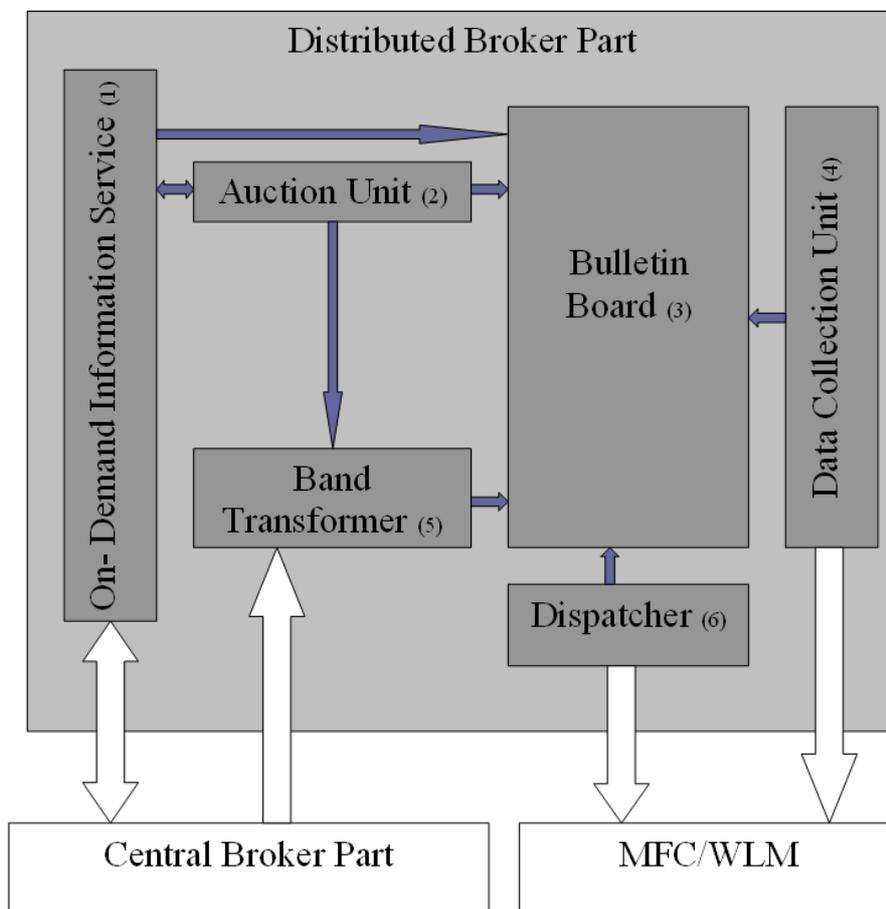


Abbildung 2.17: Distributed Broker Part

Der Distributed Broker Part ist Bestandteil jeder PD. Er besteht aus mehreren Komponenten, die in unterschiedlichen Threads laufen können und über ein periodisches

Signal miteinander synchronisiert werden. Dieses wird vom Distributed Broker Part im Abstand eines Service Intervalls erzeugt und an die Broker-Komponenten übergeben.

2.2.2.1 On-Demand Information Service

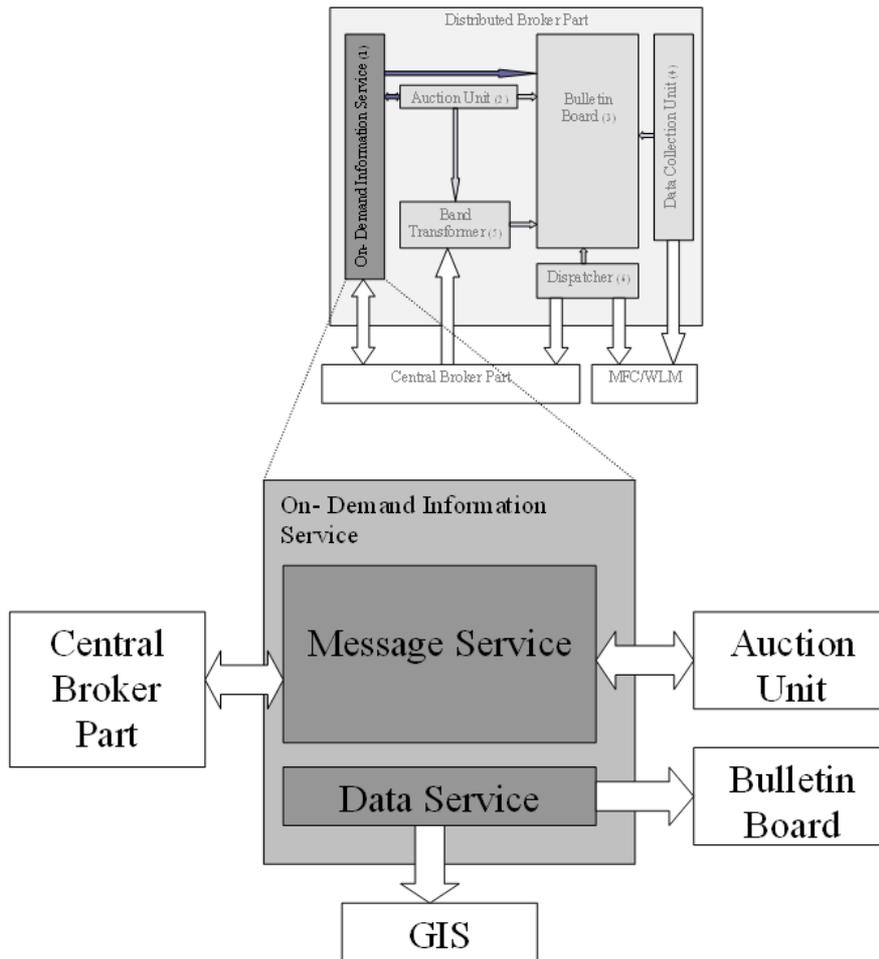


Abbildung 2.18: On-Demand Information Service

Der On-Demand Information Service des Distributed Broker Parts bildet das Gegenstück zum On-Demand Information Service des Central Broker Parts. Er stellt einen Kommunikationsdienst dar, der einen Nachrichtenverkehr mit dem Central Broker Part ermöglicht. Für eine flexible und schnelle Arbeitsweise wird auch hier der Message Container verwendet. Die Verteilung eingehender Nachrichten auf die Broker-Komponenten erfolgt anhand der Informationen im Nachrichten-Header. Der Data Service stellt eine weitere wichtige Funktion des On-Demand Information Service zur Verfügung. Auf

Anfrage erstellt dieser eine Struktur mit wesentlichen dynamischen und statischen Informationen über die PD. Dazu gehören Daten über die aktuelle Auslastung, genutzter und freier Service, Preise und Gewinn sowie die unterstützten Bänder. Die Informationen werden dem Central Broker Part über das GIS übermittelt, wo sie vom Collection Service verarbeitet werden. Diese Komponente ist auch für den Aufruf des Data Service verantwortlich, um aktiv Daten über die PDs zu sammeln. Bei wichtigen Änderungen ist der Distributed Broker Part auch in der Lage eigenmächtig den Servicedienst zu starten.

2.2.2.2 Auction Unit

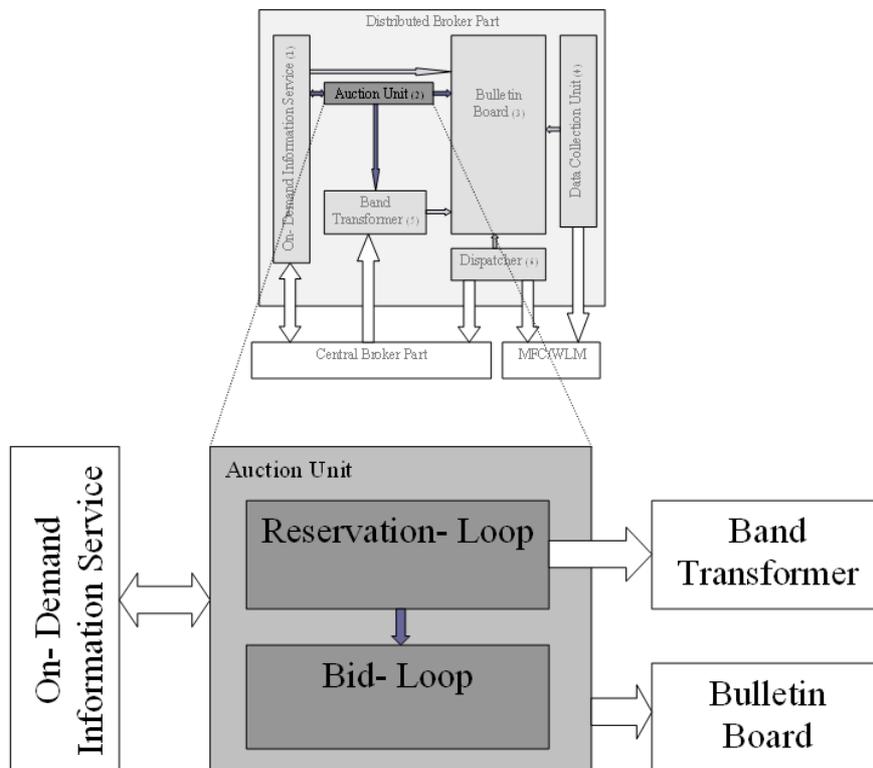


Abbildung 2.19: Auction Unit

Die Verteilung von SRWUs auf die PDs erfolgt durch einen Auktionsmechanismus. Jede PD muss daher eine Schnittstelle zur Verfügung stellen, welche die Durchführung einer Auktion ermöglicht. Beim Distributed Broker Part wird diese Aufgabe von der Auction Unit übernommen. Dabei muss sie eine Reihe wichtiger Punkte beachten. Die Auktion wird von der Band Mapping Unit des Central Broker Parts geleitet. Wichtigste Voraussetzung ist daher die Kompatibilität der Auktionsmodelle. Mögliche Verfahren sind zum Beispiel die Sealed Bid Auction, die English Auction oder die Hybrid Auction. Die Verwendung eines Plugin-Systems ermöglicht eine individuelle Anpassung.

Der prinzipielle Ablauf der Auktion verläuft stets nach dem gleichen Muster. Im ersten Schritt findet die Verarbeitung von Reservierungen statt. Darunter versteht man gewonnene Auktionen, für die eine Bestätigung durch die Auction Unit erstellt werden muss. Anschließend können neue Gebote generiert und abgegeben werden. Für die Kommunikation kommt dabei der Message Service des On-Demand Information Service zum Einsatz. Die Auktionsinformationen werden in Form von Message Container zwischen dem Central Broker Part und dem Distributed Broker Part ausgetauscht.



Abbildung 2.20: Auktionsablauf auf dem Central Broker Part

Die Umsetzung des BDAfPD ist die Hauptaufgabe des kommerziellen Grid Resource Brokers. Den größten Anteil übernimmt dabei der Distributed Broker Part. Für die Verwaltung der Bänder ist der Band Transformer verantwortlich. Die Auction Unit steuert den zeitlichen Ablauf, d.h. den Start des Band Transformers, und belegt die freie Servicekapazität der PD unter Einhaltung der Bandgrenzen. Der Aufruf des Band Transformers erfolgt zwischen der Verarbeitung von Reservierungen und der Generierung neuer Gebote. Dieser Zeitpunkt garantiert, dass nur freie Bandkapazitäten verschoben werden können. Reservierungen stellen gewonnene Auktionen dar. Die entsprechende Anzahl an Service Units muss auf der PD zur Verfügung gestellt werden. Dies wird bei der Reservierungsverarbeitung sichergestellt. Anschließend kann der Band Transformer die verbleibende freie Servicekapazität gefahrlos auf die Bänder verteilen. Jedem Band wird dabei eine bestimmte Anzahl freier Service Units zugeordnet. Die Auction Unit darf nur diese für die Erstellung neuer Gebote verwenden. Eine Überbelegung des Bandes würde die Einhaltung des QoS gefährden und unter Umständen zu hohen Gewinnverlusten führen.

Die benötigten Informationen, für die Durchführung einer Auktion unter Einhaltung des QoS, werden dem Bulletin Board entnommen. Das wichtigste Arbeitsmedium ist dabei die Time Table. Sie bietet Strukturen an, über die ein beliebiger Zeitraum überwacht bzw. verwaltet werden kann (siehe Kapitel 2.2.2.4 auf Seite 40). Die Auction Unit nutzt diese Fähigkeit, um die Ausführung von SRWUs unter Berücksichtigung ihrer Service Level Agreements zu planen. Dadurch ist es möglich den Start- und Endzeitpunkt bzw. die Dauer der Verarbeitung einer SRWU genau festzulegen. Somit kann bereits bei der Erstellung eines Gebots überprüft werden, ob die entsprechenden SLAs eingehalten werden können. Bei der Reservierung wird dann der endgültige Verarbeitungsplan in der Time Table festgelegt.

2.2.2.3 Band Transformer

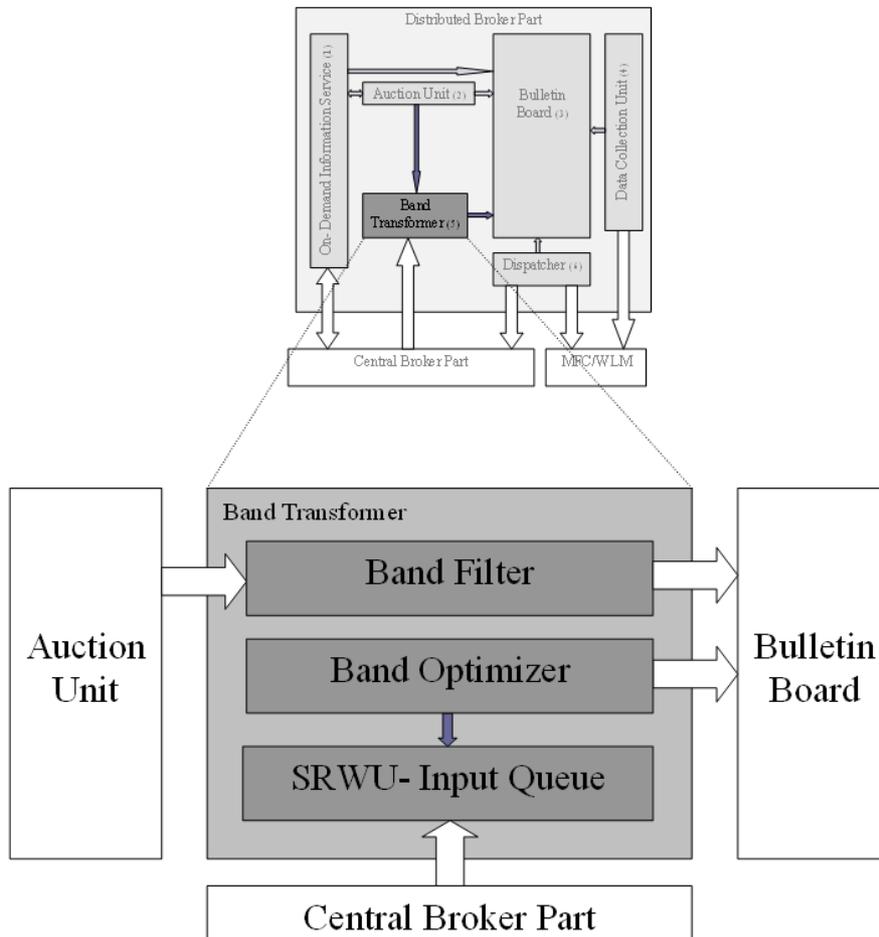


Abbildung 2.21: Band Transformer

Der Band Transformer ist für die Verwaltung der Bänder auf der PD zuständig. Dabei wird der größte Teil des BDAfPD umgesetzt. Während sich die Auction Unit um die Belegung der Bänder kümmert, d.h. Service Units zu einem möglichst hohen Preis versteigert, ist der Band Transformer für die Einstellung der Bandkapazitäten verantwortlich. Da dabei sehr unterschiedliche Strategien zum Einsatz kommen können, ist eine Anpassung durch die Verwendung eines Filter-Plugins möglich. Der Grund für diese Vielfalt liegt in den unterschiedlichen Zielen, Aufgaben und Fähigkeiten der PDs. So können zum Beispiel PDs vollständig auf dynamische Bandkapazitäten verzichten. Andere sind möglicherweise in der Lage, nur genau ein Band zu unterstützen, da andernfalls der QoS nicht eingehalten werden kann. Für die Systemoptimierung sind wiederum volldynamische Bandkapazitäten angebracht, die sich entsprechend der SRWU-Ankunftsrate einstellen. Auf diese Weise kann auch der Central Broker Part Einfluss auf die PDs nehmen, indem er die Anzahl der SRWUs pro Band für eine PD

durch den Demand-Set begrenzt. Die Verschiebung der Bandgrenzen kann nur unter gewissen Bedingungen ablaufen. Die Gesamtkapazität der PD darf nicht überschritten werden, d.h. die Summe der Kapazitäten der Bänder darf die Kapazität der PD nicht überschreiten. Es dürfen nur freie Kapazitäten verschoben werden, um den QoS der bereits zugewiesenen Service Units einhalten zu können. Reservierungen haben Vorrang vor einer Bandverschiebung. Mit der Abgabe eines Gebots geht die PD die vertragliche Bindung ein, den damit angebotenen Service zur Verfügung zu stellen. Nur die Ablehnung des Gebots durch den Central Broker Part befreit die PD von dieser Verpflichtung. Die Informationen über eine Bandverschiebung werden in der Time Table im Bulletin Board eingetragen. Somit sind auch Verschiebungen über einen beliebigen Zeitraum möglich.

Betriebssysteme und WLM bieten häufig die Möglichkeit die Ausführung von Prozessen durch Angabe von Parametern zu beeinflussen. Selbst einfache Systeme erlauben zum Beispiel die Angabe von Prioritäten. Komplexe Workload Manager, wie sie in z/OS-Systemen eingesetzt werden, erweitern diese Möglichkeiten noch erheblich. So können hier sogenannte Goals angegeben werden, die es erlauben Prozesse zum Beispiel auf ihre Antwortzeit zu optimieren. Der Band Transformer nutzt diese Möglichkeiten, indem er den SRWUs bestimmte Parameter zuweist. Diese sind so gewählt, dass der geforderte QoS des entsprechenden Bandes optimal eingehalten wird. Für diese Aufgabe nimmt der Band Transformer die SRWUs vom Central Broker Part entgegen. Anhand des Bandes werden entsprechende Parameter zugewiesen. Anschließend wird die SRWU für die Verarbeitung in der Time Table abgelegt.

2.2.2.4 Bulletin Board

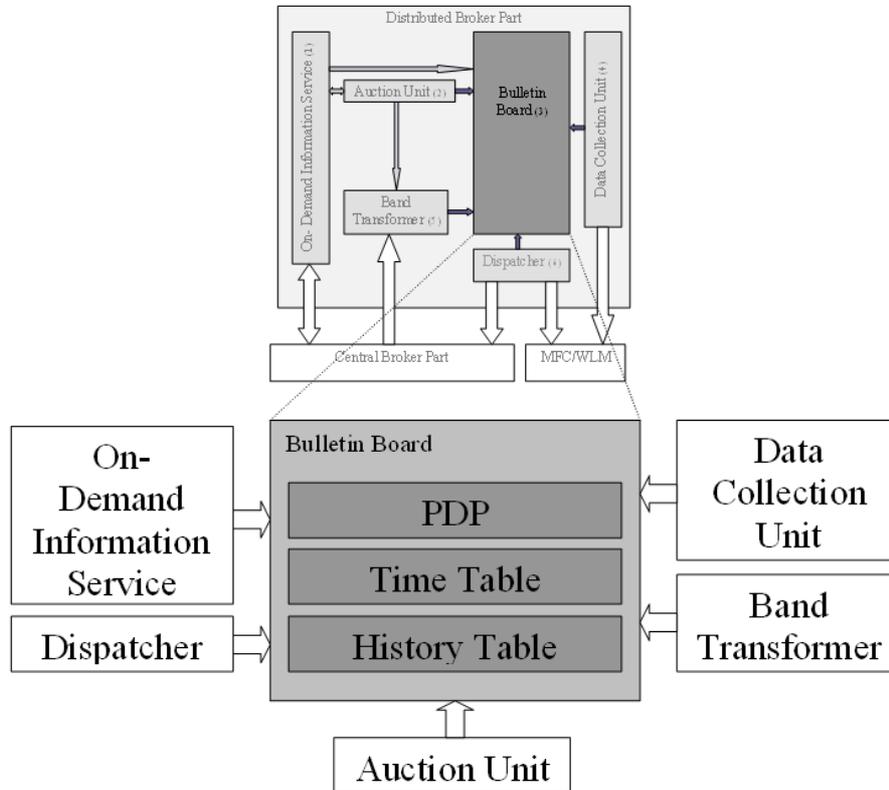


Abbildung 2.22: Bulletin Board

Das Bulletin Board des Distributed Broker Parts ist für die Speicherung von Informationen zuständig, die den Broker-Komponenten zugänglich gemacht werden sollen. Die wichtigste Speicherstruktur ist die Time Table. Diese erlaubt das Verwalten von Bändern, SRWUs und Auktionsinformationen über einen beliebigen Zeitraum. Dieser wird in Perioden mit einer Periodendauer von einem Service Intervall zerlegt. Dadurch ist es möglich, alle Angaben über die Kapazität einer PD in der Time Table in Service Units vorzunehmen. In jeder Periode werden zunächst die Bänder definiert. Die entsprechende Kapazität wird durch den Band Transformer festgelegt. Ein Band besteht aus einer bestimmten Anzahl von Slots. Ein Slot ist die kleinste Verwaltungseinheit der Time Table. Er kann eine beliebige Anzahl an SRWUs oder Auktionsdaten aufnehmen. Zu den Auktionsdaten gehören Reservierungen und Gebote, die von der Auction Unit erstellt werden. Die Kapazität eines Slots ist pro Band konstant. Nur die Anzahl kann bei dynamischen Bandgrenzen variieren. Die Aufgabe der Slots besteht in der Gruppierung von SRWUs oder Auktionsdaten mit gleichen Eigenschaften, wie WLM-Parameter oder Priorität. Diese Einteilung erlaubt eine Klassifizierung der SRWUs. Der Workload Manager kann sich somit optimal auf die Verarbeitung einer ganzen SRWU-Gruppe einstellen, ohne die Strategie (Goal) während der Verarbeitung ändern zu müssen. Weiterhin erhöht sich die Zugriffsgeschwindigkeit auf bestimmte SRWUs

(z.B. SRWUs mit der gleichen Priorität), da diese zusammengefasst sind. Die Arbeit des Dispatchers, der die SRWUs aus der Time Table für die Verarbeitung entnimmt, und des Workload Managers kann somit erheblich vereinfacht werden.

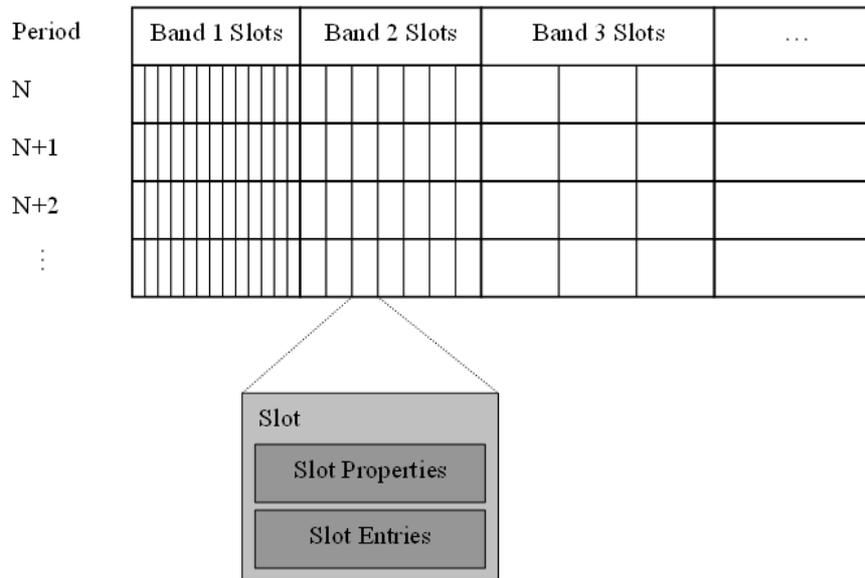


Abbildung 2.23: Time Table

Eine andere Verwaltungseinheit im Bulletin Board ist die History Table (siehe Kapitel 2.2.1.2 auf Seite 26). In ihr können die Broker-Komponenten Messwerte ablegen und abrufen, die öffentlich zugänglich gemacht werden sollen. Auf dem Distributed Broker Part dient die History Table vor allem dem Ablegen von Auslastungsinformationen der PD. Diese werden von dem On-Demand Information Service benutzt, um den Grid Information Service mit Daten über die PD zu versorgen. Aber auch die anderen Komponenten nutzen diese Informationen zur Einhaltung des QoS.

2.2.2.5 Data Collection Unit

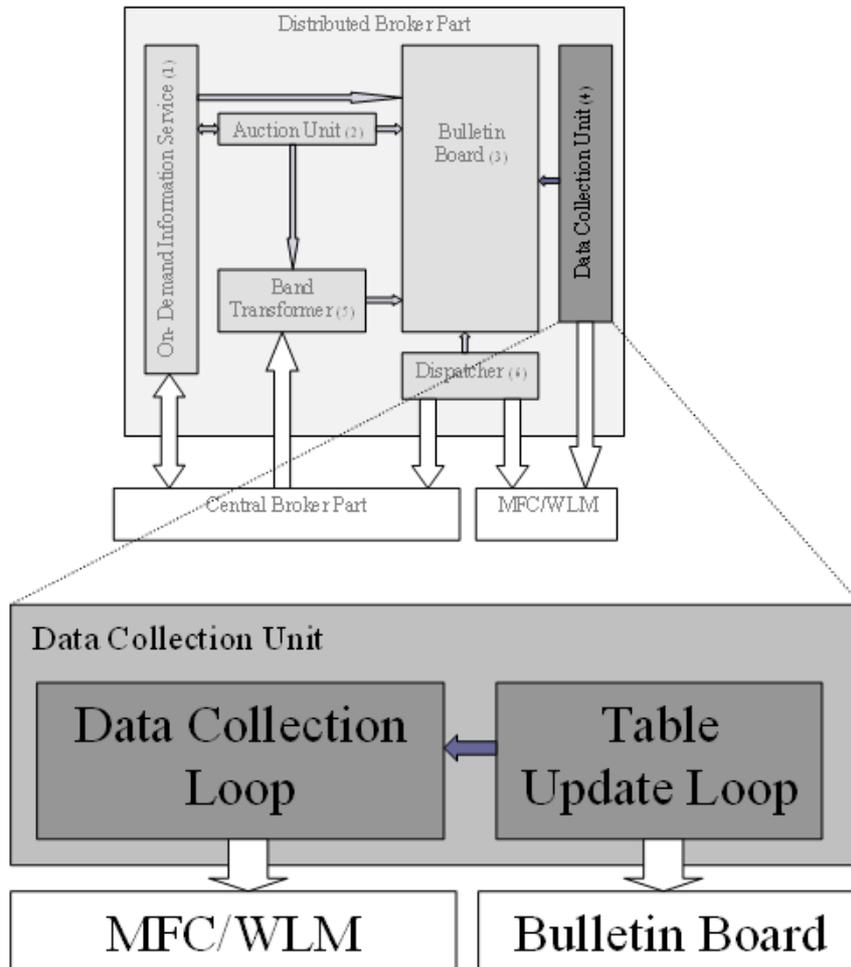


Abbildung 2.24: Data Collection Unit

Der kommerzielle Grid Resource Broker setzt für seine Arbeit sehr genaue Werte über den aktuellen Zustand der PDs voraus. Die Data Collection Unit ist für das Sammeln von Daten über die aktuelle Auslastung, den genutzten Service, Antwortzeiten und Performance Index verantwortlich. Der Workload Manager oder Masterflow Controller liefert die gewünschten Informationen von der PD. Die Data Collection Unit sammelt die Daten über den Zeitraum eines Service Intervalls. Diese werden statistisch zusammengefasst (Mittelwertbildung, Summenbildung, etc.) und über das Bulletin Board veröffentlicht.

2.2.2.6 Dispatcher

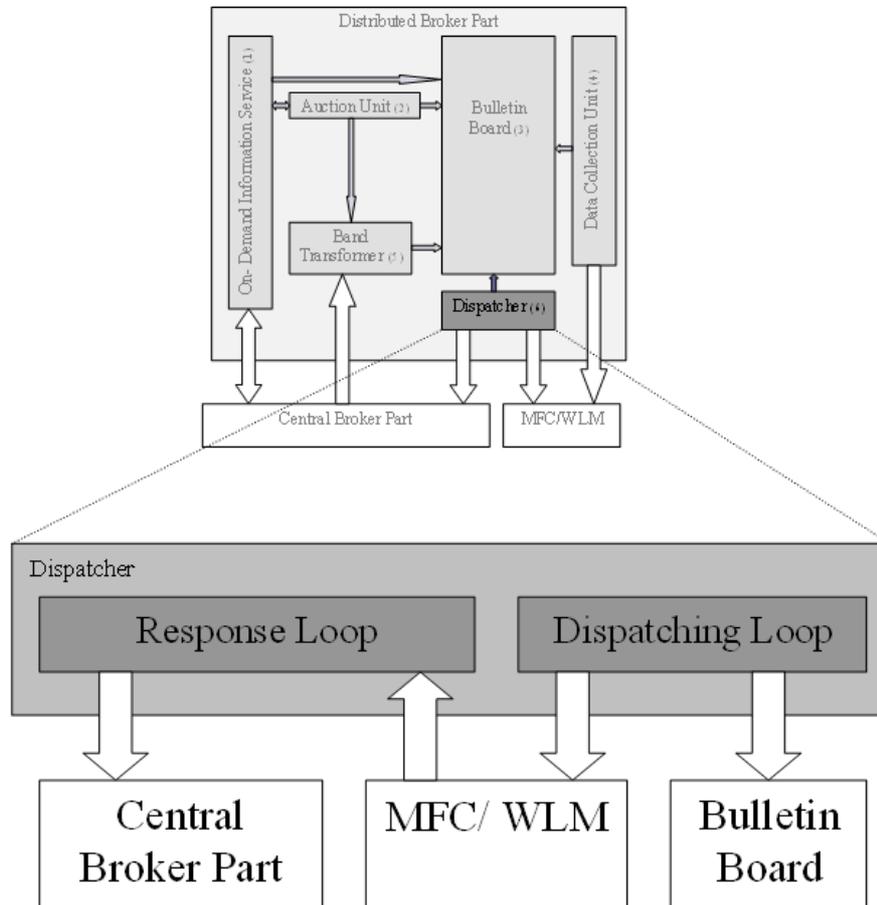


Abbildung 2.25: Dispatcher

Die Auction Unit erstellt während der Auktion einen Eintrag in der Time Table, der den genauen Startzeitpunkt einer SRWU angibt und die Anzahl an Service Units festlegt, die pro Periode zugewiesen werden sollen. Der Band Transformer belegt diese Verteilung mit einer SRWU, nachdem er ihr einige WLM-Parameter (Priorität, Enklave, etc.) zugewiesen hat. Der Dispatcher hat die Aufgabe die Time Table zu überwachen. Wenn für eine SRWU der Startzeitpunkt erreicht wurde, wird sie der Time Table entnommen der PD zur Verarbeitung übergeben. Der Dispatcher kann dabei auch nur Teile der SRWU entnehmen, wenn dies durch die Time Table festgelegt ist. Aufgrund seiner direkten Schnittstelle zur PD, übernimmt der Dispatcher auch die Aufgabe Response-SRWUs von der PD entgegenzunehmen. Diese werden anschließend an den Receiver des Central Broker Parts übermittelt.

2.2.3 SRWU-Broker Segment

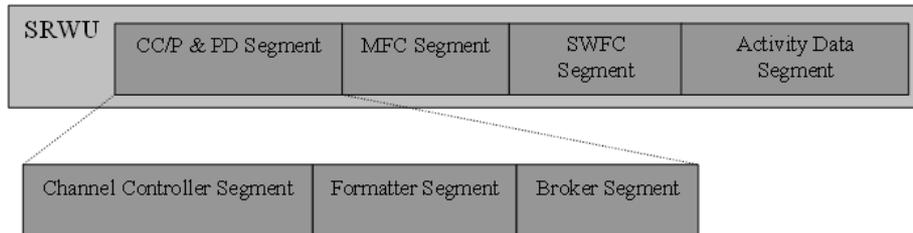


Abbildung 2.26: SRWU Übersicht

Die Service Request Work Unit ist ein universelles Datenformat zur Definition von Benutzeraufträgen. Sie ist dabei in unterschiedliche Segmente aufgeteilt. Für jede Komponente der Transaktionsmaschine existiert ein eigener Abschnitt, mit verschiedenen Steuer- und Kontrollparametern.

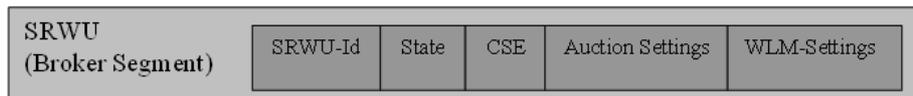


Abbildung 2.27: SRWU-Broker Segment

Das Segment des kommerziellen Grid Resource Brokers besteht aus folgenden Teilen

- SRWU-Id
- State
- Customer Service Element
- Auction Settings
- WLM-Settings

2.2.3.1 SRWU-Id



Abbildung 2.28: SRWU-Id

Die SRWU-Id dient der Identifizierung der SRWU innerhalb der Transaktionsmaschine. Sie ist Systemweit eindeutig und wird bereits vom Formatter zugewiesen.

2.2.3.2 State



Abbildung 2.29: State

Der Parameter State kennzeichnet den Status der SRWU-Verarbeitung. Folgende Werte sind möglich

New Die SRWU ist neu im System und nimmt noch an keiner Auktion teil.

Bid Die SRWU nimmt an einer Auktion teil.

Reservation Die SRWU besitzt eine gültige Reservierung auf einer PD und ist damit für eine Ausführung berechtigt

Rejection Die SRWU wurde von einer Broker-Komponente abgelehnt.

Ready Die SRWU befindet sich in der Time Table und ist für eine Ausführung bereit.

Run Die SRWU oder Teile der SRWU befinden sich in der Verarbeitung.

Finished Die SRWU wurde vollständig und fehlerfrei verarbeitet. Dieser Wert ist nur bei einer Response-SRWU gültig.

Failed Die Verarbeitung der SRWU ist fehlgeschlagen. Dieser Wert ist nur bei einer Response-SRWU gültig.

2.2.3.3 Customer Service Element

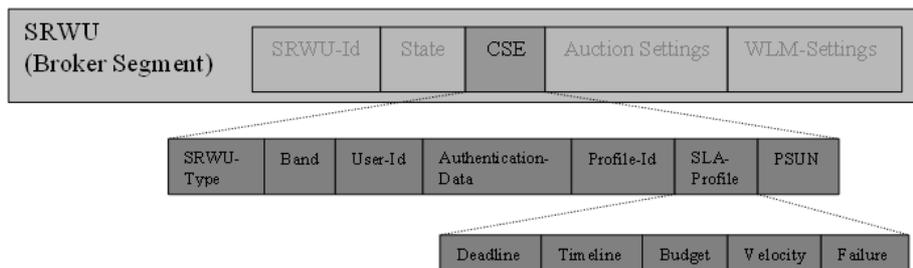


Abbildung 2.30: Customer Service Element

Das Customer Service Element enthält die Benutzervorgaben für die Verarbeitung der SRWU. Der Leistungsabnehmer legt nur einen Teil der Parameter fest. Ein großer Teil wird durch den Profiler erzeugt.

SRWU-Type Legt den Typ der SRWU fest, durch den das Band ermittelt wird. Dieser Parameter wird durch den Formatter bestimmt.

Band Kennzeichnet das Band, in dem die SRWU verarbeitet werden soll. Der Profiler legt diesen Parameter über den SRWU-Typ und dem SLA-Profil fest.

User-Id Enthält die Benutzer-Identifikation, mit der sich ein Auftraggeber vorstellt. Die Id verweist auf das Benutzerprofil im Profiler.

Authentication-Data Enthält Daten, die die Echtheit des Auftraggebers bestätigen (z.B. Passwort)

Profile-Id Kann durch den Auftraggeber wahlweise angegeben werden. Enthält einen Verweis auf ein spezielles SLA-Profil, das vom Standardprofil abweicht. Wenn dieser Parameter nicht angegeben wird, füllt der Profiler das Feld mit der Id des Standardprofils.

SLA-Profile Wird vom Profiler ausgefüllt und enthält Angaben über Deadline, Timeline, Failure, Velocity und Budget und wie diese eingehalten werden sollen (Service Level Agreements).

PSUN Enthält die geschätzte Anzahl an Service Units, die für die Verarbeitung benötigt werden. Dieser Parameter wird vom Profiler ausgefüllt.

2.2.3.4 Auction Settings

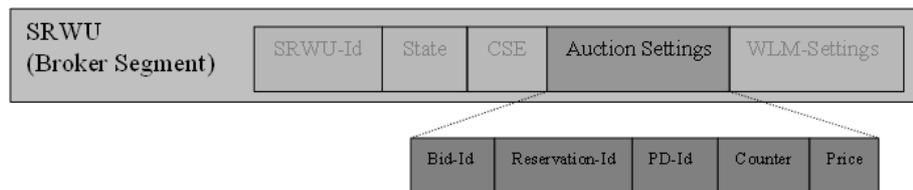


Abbildung 2.31: Auction Settings

Die Auction Settings sind Parameter, die vor oder während der Auktion festgelegt werden und den Auktionsablauf regeln.

Bid-Id Enthält die Id des Gebots, das den Zuschlag bei der Auktion erhält.

Reservation-Id Enthält die Id der Reservierungsbestätigung, die durch eine PD erstellt wird.

PD-Id Enthält die Id der PD, die die Auktion für die SRWU gewonnen hat.

Counter Ein Zähler, der durch den Profiler initialisiert wird und festlegt, wie oft die SRWU erfolglos an einer Auktion teilnehmen darf.

Price Der Preis, der bei der Auktion ausgehandelt wurde.

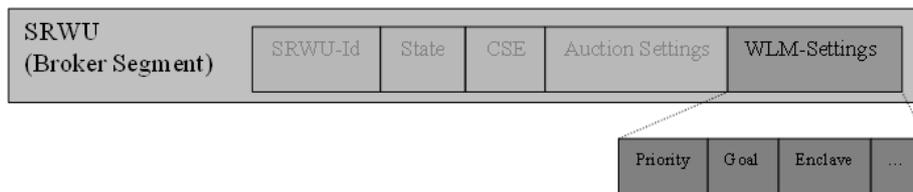


Abbildung 2.32: WLM-Settings

2.2.3.5 WLM-Settings

Die WLM-Settings werden durch den Band-Transformer festgelegt und sind abhängig vom Band und von den Eigenschaften der PD. Viele Systeme erlauben mindestens das Festlegen der Priorität (Priority). Weitere Parameter können sein

- Goal
- Enclave
- usw.

2.2.4 Bid

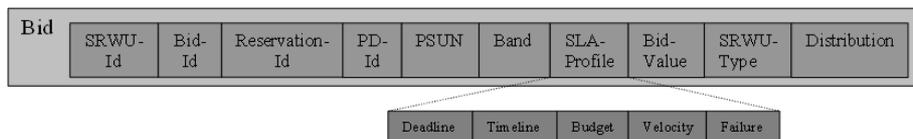


Abbildung 2.33: Bid

Die Datenstruktur Bid kommt bei der Durchführung einer Auktion zum Einsatz. Sie wird verwendet um Gebotsanfragen, Gebote, Ablehnungen, Reservierungsanfragen, Reservierungen und Reservierungsablehnungen zu erstellen. Folgende Felder können belegt werden

SRWU-Id Kennzeichnet die SRWU, für die das Bid zuständig ist.

Bid-Id Kennzeichnet eindeutig ein Gebot von einer PD.

Reservation-Id Kennzeichnet eindeutig eine Reservierung von einer PD.

PD-Id Kennzeichnet eindeutig die PD, die ein Gebot oder eine Reservierung abgegeben hat.

PSUN Anzahl der Service Units, die benötigt werden.

Band Kennzeichnet das Band, in dem die Service Units versteigert werden.

SLA-Profile Gibt die Service Level Agreements an, die für die versteigerten Service Units eingehalten werden müssen. Diese bewegen sich nur im Toleranzbereich des Bandes.

Bid-Value Gibt den Wert des Gebots an, d.h. den Preis, den die PD angegeben hat.

SRWU-Type Legt den Typ der SRWU fest, für die die Auktion stattfindet.

Distribution Erlaubt das Festlegen und Versteigern von Service Units über einen Zeitraum, d.h. für einen oder mehrere beliebige Zeitpunkte in der Zukunft kann eine PD Service Units versteigern (siehe Kapitel 4.5.2 auf Seite 83).

State Gibt die Art der Bid-Datenstruktur an, d.h. in welchem Status befindet sich die Auktion.

2.2.5 PDP

2.2.6 PSE

2.3 Funktionsbeschreibung

Der kommerzielle Grid Resource Broker arbeitet als Teil der Transaktionsmaschine in der Channel Controller/Profiler Komponente. Durch einen autonomen Loadbalancing- und Scheduling-Algorithmus verbirgt er die innere Struktur der verwalteten PDs und präsentiert sich als virtueller Service Provider.

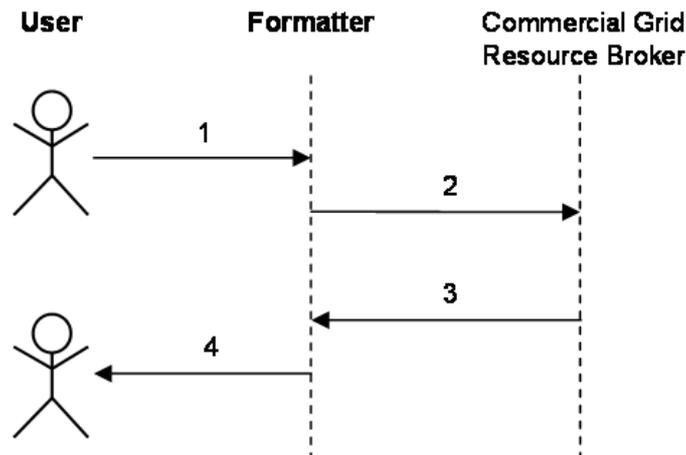


Abbildung 2.34: Use Case #01: Übersicht Kommerzieller Grid Resource Broker

Use Case #01	Übersicht Kommerzieller Grid Resource Broker
Use Case Beschreibung	
Die Transaktionsmaschine wartet über den Formatter auf eingehende Serviceanforderungen von Leistungsabnehmer (Benutzer, Kunden oder andere externe Systeme). Diese werden in das interne Format einer SRWU umgewandelt (1). Der Formatter übergibt die SRWU an den kommerziellen Grid Resource Broker, der für die Verarbeitung verantwortlich ist (2). Als Reaktion sendet dieser eine Response-SRWU (3). Diese enthält Informationen über evtl. abgelehnte Aufträge oder dem Ergebnis bzw. Status der Verarbeitung. Der Formatter erstellt aus der Response-SRWU eine benutzerspezifische Ausgabe und übermittelt sie an den Leistungsabnehmer (User-Response) (4).	
Eingabedaten	Individuelle Serviceanforderung (Online-Formular, EC-Karten Buchung, etc.), die vom Formatter in eine SRWU verwandelt werden kann.
Ausgabedaten	Individueller User-Response (Auftragbestätigung, Auftragsablehnung, etc.), der vom Formatter aus einer Response-SRWU erstellt wurde.

Tabelle 2.2: Übersicht Kommerzieller Grid Resource Broker

Die Trennung in Central Broker Part und Distributed Broker Part erlaubt eine Aufteilung der Funktionen des kommerziellen Grid Resource Brokers unter Berücksichtigung einer verteilten Verarbeitung (Distributed Processing).

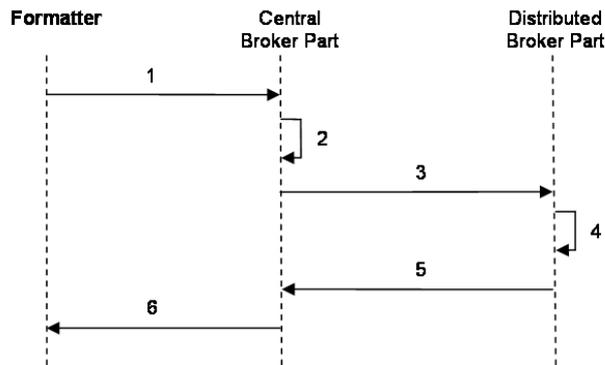


Abbildung 2.35: Use Case #02: Verarbeitung einer SRWU

Use Case #02	Verarbeitung einer SRWU
Use Case Beschreibung	
<p>Der Formatter übergibt eine erstellte SRWU an die zentrale Brokerinstanz der Transaktionsmaschine den Central Broker Part (1). Neben einer Authentifizierung werden hier Ausführungsprofile (SLA, PSUN, etc.) zugewiesen und eine zentrale Optimierung des Eingangsdatenstroms durchgeführt (2). Die SRWU wird durch ein Auktionsverfahren an eine PD bzw. an deren Distributed Broker Part übergeben. Dabei kommen sowohl Mechanismen zur Lastverteilung, als auch zur Gewinnoptimierung zum Einsatz (3). Der Distributed Broker Part veranlasst die lokale Verarbeitung der SRWU auf der PD. Neben der Einhaltung und Anpassung der Bandgrenzen wird die Ausführung des Auftrags über einen Scheduling-Algorithmus geplant (4). Nach der vollständigen Verarbeitung der SRWU auf einer PD sendet der Distributed Broker Part eine Response-SRWU an den Central Broker Part (5), der diese an den Formatter übergibt (6).</p>	
Eingabedaten	Durch den Formatter erstellte SRWU.
Ausgabedaten	Response-SRWU vom kommerziellen Grid Resource Broker.

Tabelle 2.4: Verarbeitung einer SRWU

Die Funktionsweise des kommerziellen Grid Resource Brokers gliedert sich in die Bereiche

1. Annahme und Verifikation einer SRWU
2. Bestimmung einer geeigneten PD
3. Übergabe eines Auftrags
4. Übergabe eines Response
5. Data Collection

2.3.1 Annahme und Verifikation einer SRWU

Für jede SRWU liegen Verarbeitungsinformationen in den Benutzerprofilen des Profilers vor. Der Zugriff wird durch einen Authentifizierungsmechanismus geschützt, um einen Missbrauch zu verhindern. Nach erfolgreicher Überprüfung der Echtheit eines Auftrags können die Verarbeitungsinformationen (SLA, Band, PSUN) zugewiesen werden. Beim Auftreten eines Fehlers erfolgt die Ablehnung der SRWU bzw. die Vermittlung durch den Mediator.

2.3.1.1 Annahme einer gültigen SRWU

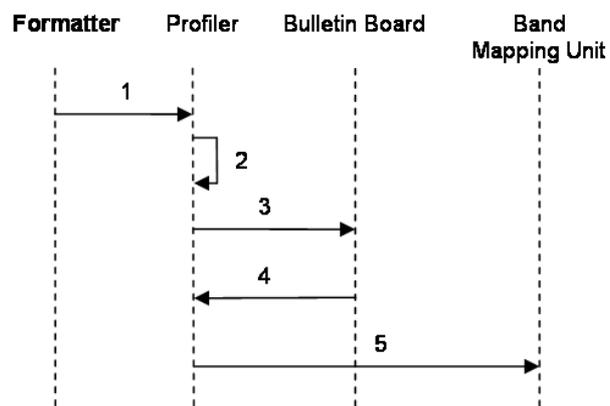


Abbildung 2.36: Use Case #03: Annahme einer gültigen SRWU

Use Case #03	Annahme einer gültigen SRWU
Use Case Beschreibung	
<p>Der Profiler wartet auf eine neue SRWU, die ihm vom Formatter übergeben wird (1). Die SRWU enthält neben einem SRWU-Typ, Authentifizierungsdaten und eine Benutzer-Id, über die ein Authentifizierungsmechanismus die Berechtigung auf ein Benutzerprofil prüft. Diese wird benutzt um der SRWU ein Verarbeitungsprofil (Band, SLA) zuzuweisen. Wenn keine Profilkennung übermittelt wurde, wird ein Standardprofil verwendet. (2). Das Bulletin Board liefert die voraussichtliche Anzahl an SU (PSUN), welche für eine Verarbeitung benötigt werden (3 bis 4). Die initialisierte SRWU wird anschließend von der Band Mapping Unit entgegengenommen (5).</p>	
Eingabedaten	<p>Von dem Formatter erstellte SRWU. Enthält</p> <ul style="list-style-type: none"> • Benutzer-Id • Authentifizierungsdaten (Passwort) • SRWU-Typ • Profilkennung (optional)
Ausgabedaten	<p>Initialisierte SRWU. Enthält zusätzlich</p> <ul style="list-style-type: none"> • SLA-Profil • Band-Typ • PSUN

Tabelle 2.6: Annahme einer gültigen SRWU

2.3.1.2 Ablehnung einer ungültigen SRWU

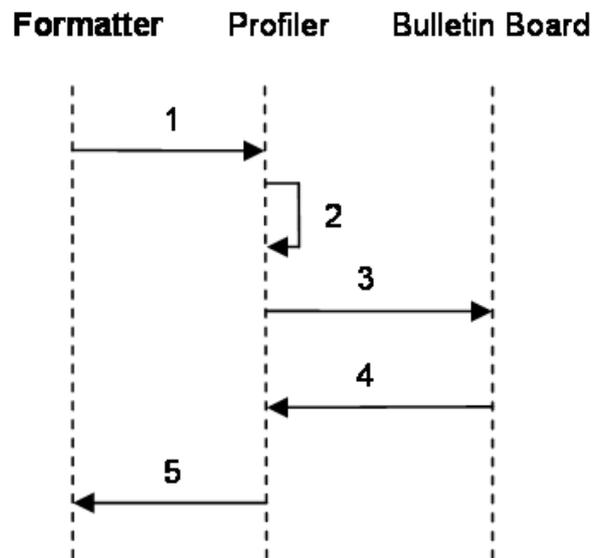


Abbildung 2.37: Use Case #04: Ablehnung einer ungültigen SRWU

Use Case #04	Ablehnung einer ungültigen SRWU
Use Case Beschreibung	
<p>Der Profiler wartet auf eine neue SRWU, die ihm vom Formatter übergeben wird (1). Die SRWU enthält neben einem SRWU-Typ, Authentifizierungsdaten und eine Benutzer-Id, über die ein Authentifizierungsmechanismus die Berechtigung auf ein Benutzerprofil prüft. Diese wird benutzt um der SRWU ein Verarbeitungsprofil (Band, SLA) zuzuweisen. Wenn keine Profilkennung übermittelt wurde, wird ein Standardprofil verwendet. (2). Das Bulletin Board liefert die voraussichtliche Anzahl an SU (PSUN), welche für eine Verarbeitung benötigt werden (3 bis 4). Der Profiler erkennt eine fehlgeschlagene Authentifizierung oder den Zugriff auf ein unberechtigtes Verarbeitungsprofil. Die SRWU wird als fehlerhaft markiert und an den Formatter übergeben (5).</p>	
Eingabedaten	Von dem Formatter erstellte SRWU. Enthält <ul style="list-style-type: none"> • Benutzer-Id • Authentifizierungsdaten (Passwort) • SRWU-Typ • Profilkennung (optional)
Ausgabedaten	Nicht initialisierte SRWU. Enthält zusätzlich <ul style="list-style-type: none"> • Fehlermeldung

Tabelle 2.8: Ablehnung einer ungültigen SRWU

2.3.1.3 Vermittlung einer nicht verarbeitbaren SRWU

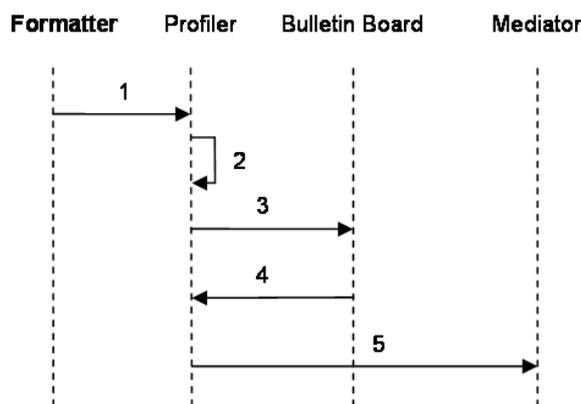


Abbildung 2.38: Use Case #05: Vermittlung einer nicht verarbeitbaren SRWU

Use Case #05	Vermittlung einer nicht verarbeitbaren SRWU
Use Case Beschreibung	
<p>Der Profiler wartet auf eine neue SRWU, die ihm vom Formatter übergeben wird (1). Die SRWU enthält neben einem SRWU-Typ, Authentifizierungsdaten und eine Benutzer-Id, über die ein Authentifizierungsmechanismus die Berechtigung auf ein Benutzerprofil prüft. Diese wird benutzt um der SRWU ein Verarbeitungsprofil (Band, SLA) zuzuweisen. Wenn keine Profilkennung übermittelt wurde, wird ein Standardprofil verwendet. (2). Das Bulletin Board liefert die voraussichtliche Anzahl an SU (PSUN), welche für eine Verarbeitung benötigt werden (3 bis 4). Der Profiler erkennt den Zugriff auf ein unberechtigtes Verarbeitungsprofil nach erfolgreicher Authentifizierung. Die Benutzerdaten erlauben dem Profiler die Vermittlung der SRWU in einen anderen administrativen Bereich, so dass die SRWU an den Mediator übergeben wird (5).</p>	
Eingabedaten	Von dem Formatter erstellte SRWU. Enthält <ul style="list-style-type: none"> • Benutzer-Id • Authentifizierungsdaten (Passwort) • SRWU-Typ • Profilkennung
Ausgabedaten	Nicht initialisierte SRWU.

Tabelle 2.10: Vermittlung einer ungültigen SRWU

2.3.2 Bestimmung einer geeigneten PD

Die Bestimmung einer geeigneten PD zur Ausführung einer SRWU erfolgt in Zusammenarbeit von Central Broker Part und Distributed Broker Part. Über einen Auktionsmechanismus wird ein ökonomischer Preis für die Verarbeitung bestimmt, der Angebot, Nachfrage und Kosten berücksichtigt. Dieser Preis bildet die Grundlage für die Auswahl der richtigen PD.

Die detaillierte Beschreibung des Auktionsmechanismus befindet sich im Kapitel 4.5 auf Seite 78.

2.3.2.1 Bestimmung einer PD durch eine erfolgreiche Auktion

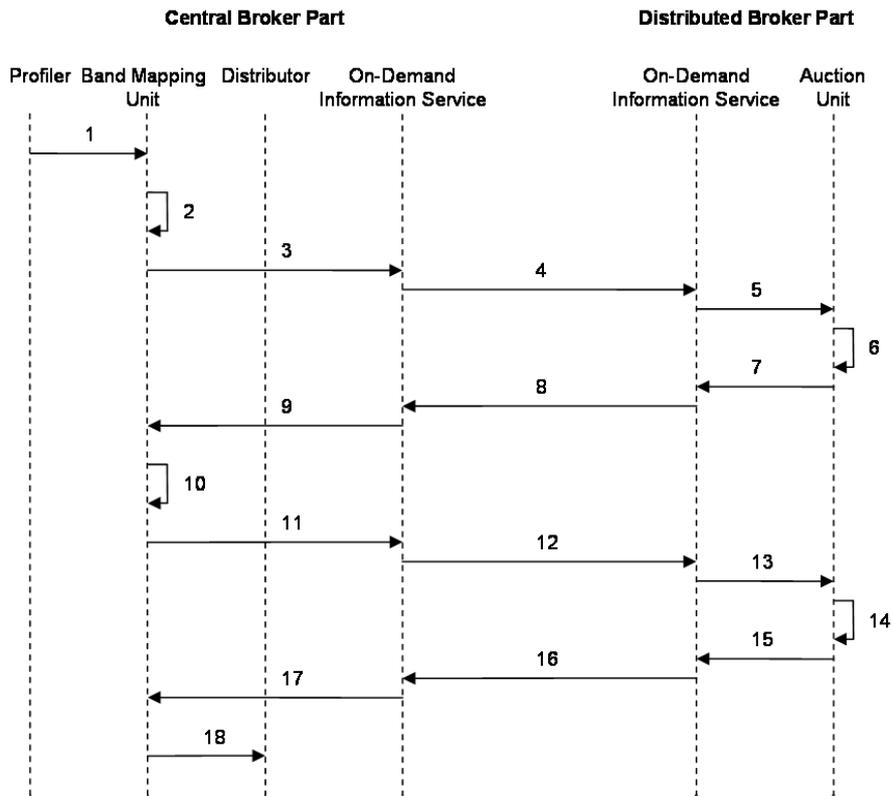


Abbildung 2.39: Use Case #06: Bestimmung einer PD durch eine erfolgreiche Auktion

Use Case #06	Bestimmung einer PD durch eine erfolgreiche Auktion
Use Case Beschreibung	
<p>Die Band Mapping Unit wartet auf neue SRWUs, die in einer SRWU-Input Queue gesammelt werden (1). Eine Liste von SRWUs wird der SRWU-Input Queue entnommen und für jede SRWU wird der Feasible-Set/Demand-Set bestimmt und eine Gebotsanfrage erstellt (2). Über den On-Demand Information Service wird die Gebotsanfrage an die Auction Unit der Ziel-PDs gesendet (3 bis 5). Diese passen ihre Bandkapazitäten an die Anforderungen an und erstellen für jede Gebotsanfrage ein Gebot (6). Dieses wird an die Band Mapping Unit zurückgesendet (7 bis 9). Jedes Gebot wird überprüft und das Beste (bester Preis) der SRWU zugewiesen (10). Über den On-Demand Information Service versendet die Band Mapping Unit eine Reservierungsanfrage für die gewonnene Auktion an die Auction Unit der entsprechenden Sieger-PD (11 bis 13). Diese überprüft den Zuschlag und plant die Ausführung der SRWU (14) ein. Die Bestätigung der Reservierung wird über den On-Demand Information Service an die Band Mapping Unit gesendet (15 bis 17). Diese übergibt die SRWU an den Distributor, um sie an die PD zu verschicken (18).</p>	
Eingabedaten	Initialisierte SRWU an die Band Mapping Unit.
Ausgabedaten	SRWU mit einer gültigen Reservierung für die PD.

Tabelle 2.12: Bestimmung einer PD durch eine erfolgreiche Auktion

2.3.2.2 Wiederholung nach erfolgloser Auktion

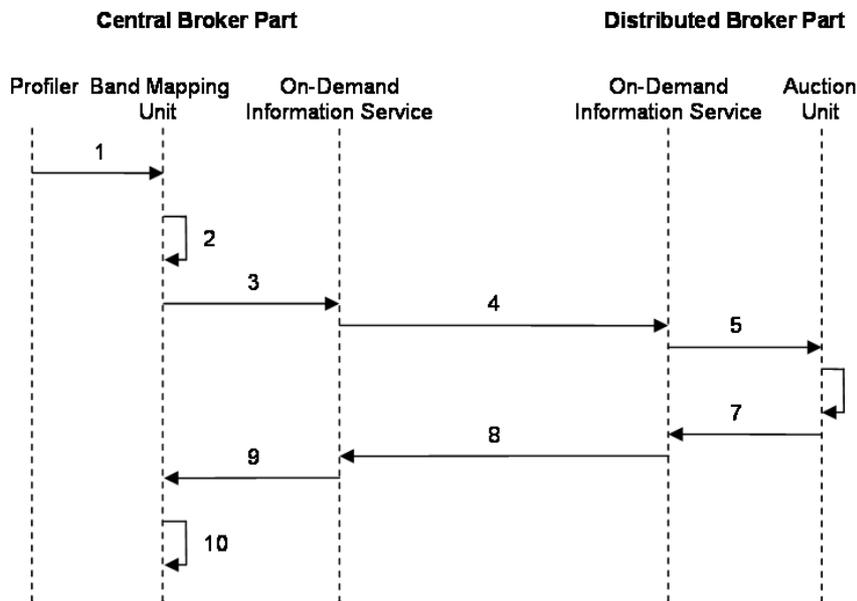


Abbildung 2.40: Use Case #07: Wiederholung nach erfolgloser Auktion

Use Case #07	Wiederholung nach erfolgloser Auktion
Use Case Beschreibung	
<p>Die Band Mapping Unit wartet auf neue SRWUs, die in einer SRWU-Input Queue gesammelt werden (1). Eine Liste von SRWUs wird der SRWU-Input Queue entnommen und für jede SRWU wird der Feasible-Set/Demand-Set bestimmt und eine Gebotsanfrage erstellt (2). Über den On-Demand Information Service wird die Gebotsanfrage an die Auction Unit der Ziel-PDs gesendet (3 bis 5). Diese passen ihre Bandkapazitäten an die Anforderungen an und erstellen für jede Gebotsanfrage ein Gebot (6). Dieses wird an die Band Mapping Unit zurückgesendet (7 bis 9). Jedes Gebot wird überprüft und das Beste (bester Preis) der SRWU zugewiesen. Da kein geeignetes Gebot gefunden werden kann, wird die SRWU für eine Wiederholung der Auktion erneut in die SRWU-Input Queue gestellt (10). Dabei verringert sich die Anzahl der möglichen Auktionsteilnahmen der SRWU (Counter in den Auction Settings). Dieser Wert begrenzt die Anzahl der Wiederholungen.</p>	
Eingabedaten	Initialisierte SRWU an die Band Mapping Unit.
Ausgabedaten	keine

Tabelle 2.14: Wiederholung nach erfolgloser Auktion

2.3.2.3 Wiederholung nach erfolgloser Reservierung

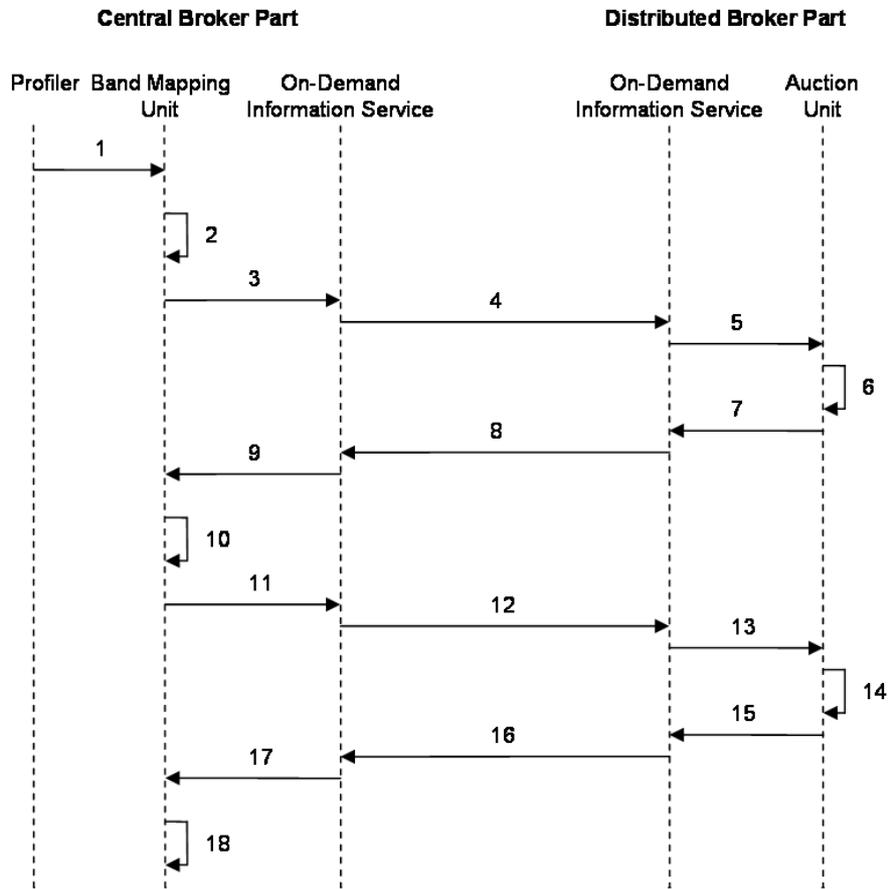


Abbildung 2.41: Use Case #08: Wiederholung nach erfolgloser Reservierung

Use Case #08	Wiederholung nach erfolgreicher Reservierung
Use Case Beschreibung	
<p>Die Band Mapping Unit wartet auf neue SRWUs, die in einer SRWU-Input Queue gesammelt werden (1). Eine Liste von SRWUs wird der SRWU-Input Queue entnommen und für jede SRWU wird der Feasible-Set/Demand-Set bestimmt und eine Gebotsanfrage erstellt (2). Über den On-Demand Information Service wird die Gebotsanfrage an die Auction Unit der Ziel-PDs gesendet (3 bis 5). Diese passen ihre Bandkapazitäten an die Anforderungen an und erstellen für jede Gebotsanfrage ein Gebot (6). Dieses wird an die Band Mapping Unit zurückgesendet (7 bis 9). Jedes Gebot wird überprüft und das Beste (bester Preis) der SRWU zugewiesen (10). Über den On-Demand Information Service versendet die Band Mapping Unit eine Reservierungsanfrage für die gewonnene Auktion an die Auction Unit der entsprechenden Sieger-PD (11 bis 13). Diese überprüft den Zuschlag und stellt einen Fehler fest (14). Die Ablehnung der Reservierung wird über den On-Demand Information Service an die Band Mapping Unit gesendet (15 bis 17). Diese übergibt die SRWU wieder an die SRWU-Input Queue für eine Wiederholung der Auktion (18). Dabei verringert sich die Anzahl der möglichen Auktionsteilnahmen der SRWU (Counter in den Auction Settings). Dieser Wert begrenzt die Anzahl der Wiederholungen.</p>	
Eingabedaten	Initialisierte SRWU an die Band Mapping Unit.
Ausgabedaten	keine

Tabelle 2.16: Wiederholung nach erfolgreicher Reservierung

2.3.2.4 SRWU-Abgabe nach erfolgreicher Auktion

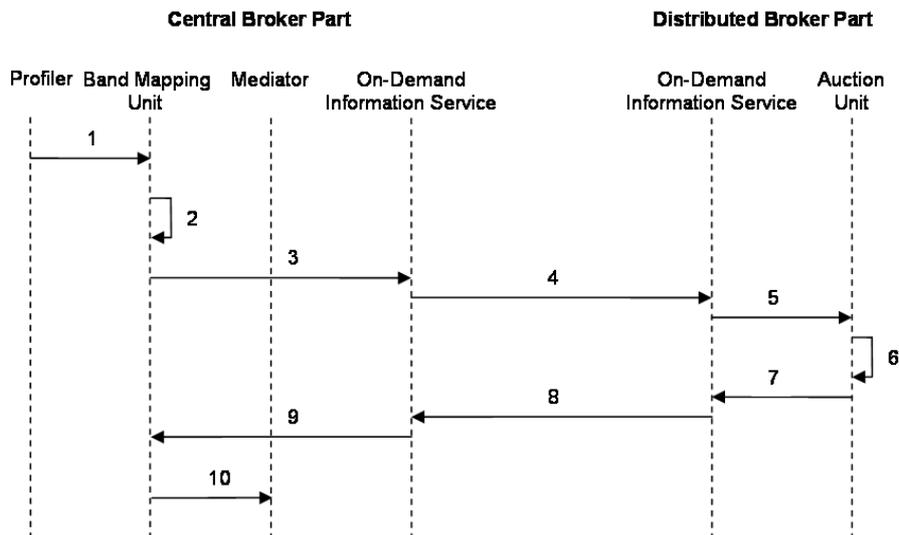


Abbildung 2.42: Use Case #09: SRWU-Abgabe nach erfolgreicher Auktion

Use Case #09	SRWU-Abgabe nach erfolgreicher Auktion
Use Case Beschreibung	
Die Band Mapping Unit wartet auf neue SRWUs, die in einer SRWU-Input Queue gesammelt werden (1). Eine Liste von SRWUs wird der SRWU-Input Queue entnommen und für jede SRWU wird der Feasible-Set/Demand-Set bestimmt und eine Gebotsanfrage erstellt (2). Über den On-Demand Information Service wird die Gebotsanfrage an die Auction Unit der Ziel-PDs gesendet (3 bis 5). Diese passen ihre Bandkapazitäten an die Anforderungen an und erstellen für jede Gebotsanfrage ein Gebot (6). Dieses wird an die Band Mapping Unit zurückgesendet (7 bis 9). Jedes Gebot wird überprüft und das Beste (bester Preis) der SRWU zugewiesen. Da kein geeignetes Gebot gefunden werden kann, wird die SRWU für eine Wiederholung der Auktion vorgesehen. Da die maximale Anzahl an Auktionsdurchgängen erreicht ist (Counter in den Auction Settings), wird die SRWU dem Mediator übergeben (10).	
Eingabedaten	Initialisierte SRWU an die Band Mapping Unit.
Ausgabedaten	Abgelehnte SRWU an den Mediator.

Tabelle 2.18: SRWU-Abgabe nach erfolgreicher Auktion

2.3.2.5 SRWU-Abgabe nach erfolgreicher Reservierung

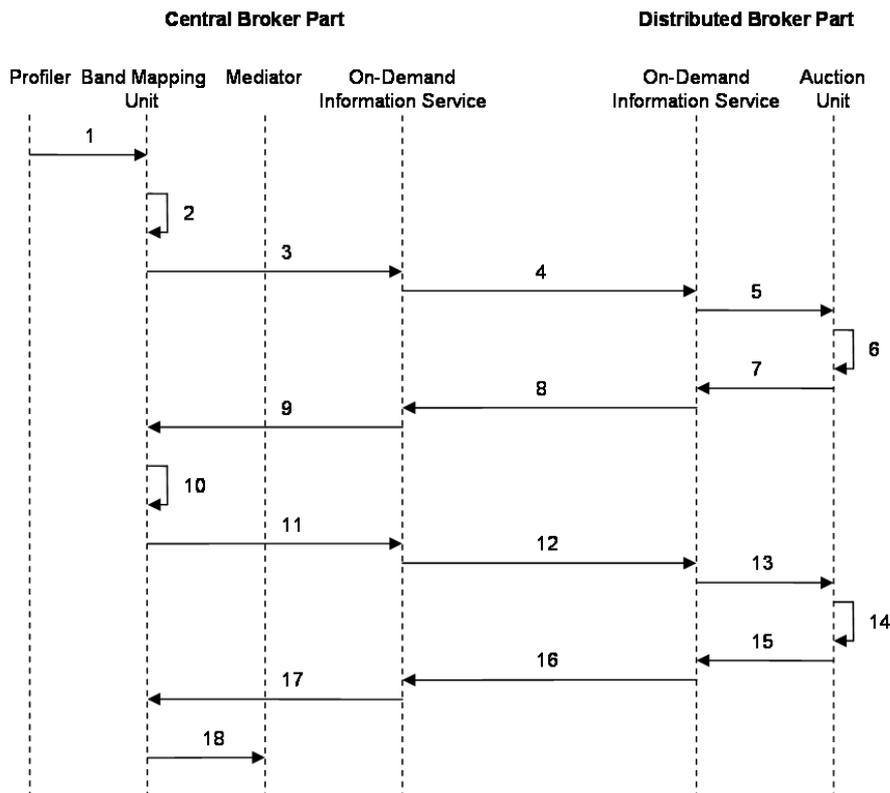


Abbildung 2.43: Use Case #10: SRWU-Abgabe nach erfolgreicher Reservierung

Use Case #10	SRWU-Abgabe nach erfolgreicher Reservierung
Use Case Beschreibung	
<p>Die Band Mapping Unit wartet auf neue SRWUs, die in einer SRWU-Input Queue gesammelt werden (1). Eine Liste von SRWUs wird der SRWU-Input Queue entnommen und für jede SRWU wird der Feasible-Set/Demand-Set bestimmt und eine Gebotsanfrage erstellt (2). Über den On-Demand Information Service wird die Gebotsanfrage an die Auction Unit der Ziel-PDs gesendet (3 bis 5). Diese passen ihre Bandkapazitäten an die Anforderungen an und erstellen für jede Gebotsanfrage ein Gebot (6). Dieses wird an die Band Mapping Unit zurückgesendet (7 bis 9). Jedes Gebot wird überprüft und das Beste (bester Preis) der SRWU zugewiesen (10). Über den On-Demand Information Service versendet die Band Mapping Unit eine Reservierungsanfrage für die gewonnene Auktion an die Auction Unit der entsprechenden Sieger-PD (11 bis 13). Diese überprüft den Zuschlag und stellt einen Fehler fest (14). Die Ablehnung der Reservierung wird über den On-Demand Information Service an die Band Mapping Unit gesendet (15 bis 17). Diese stellt fest, dass die SRWU an keiner weiteren Auktion mehr teilnehmen darf, da die maximale Anzahl an Auktionsdurchgängen erreicht ist (Counter in den Auction Settings). Die SRWU wird anschließend dem Mediator übergeben (18).</p>	
Eingabedaten	Initialisierte SRWU an die Band Mapping Unit.
Ausgabedaten	Abgelehnte SRWU an den Mediator.

Tabelle 2.20: SRWU-Abgabe nach erfolgreicher Reservierung

2.3.2.6 Band-Management

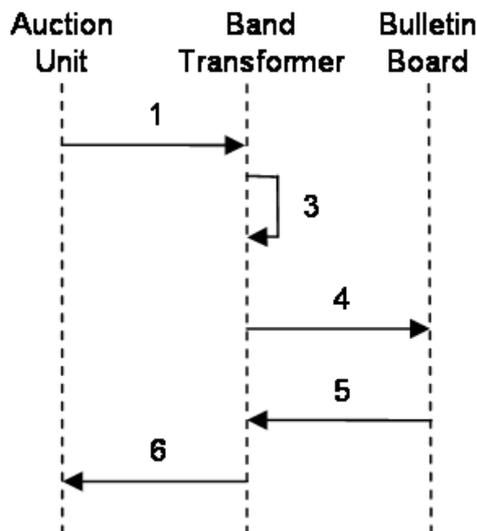


Abbildung 2.44: Use Case #11: Band-Management

Use Case #11	Band-Management
Use Case Beschreibung	
Der Band Transformer empfängt von der Auction Unit sämtliche Informationen über eingetroffene Gebotsanfragen (1). Diese werden analysiert und auf dieser Basis die neuen Bandkapazitäten bestimmt. Wenn nötig, können Gebotsanfragen bereits abgelehnt werden (z.B. Gebotsanfragen für SRWUs, die ein geändertes SLA-Profil aufweisen und aufgrund einer hohen Bandauslastung der QoS nicht garantiert werden kann) (2). Das Bulletin Board schreibt die neuen Bandgrenzen in die Time Table (3 bis 4). Die Gebotsanfragen bzw. erstellte Ablehnungen werden anschließend wieder der Auction Unit übergeben (5).	
Eingabedaten	Alle eingetroffenen Gebotsanfragen.
Ausgabedaten	Gebotsanfragen und evtl. erstellte Ablehnungen.

Tabelle 2.22: Band-Management

2.3.2.7 SRWU-Scheduling

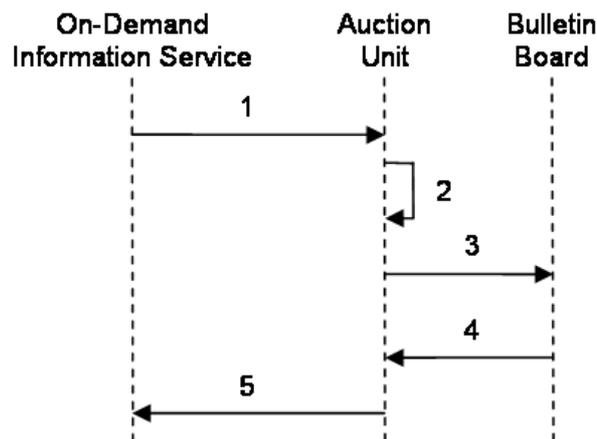


Abbildung 2.45: Use Case #12: SRWU-Scheduling

Use Case #12	SRWU-Scheduling
Use Case Beschreibung	
Der On-Demand Information Service übergibt der Auction Unit die eingetroffenen Reservierungsanfragen, die eine gewonnene Auktion kennzeichnen (1). Für jede Reservierungsanfrage wird ein Ausführungsplan erstellt, der die SLAs der entsprechenden SRWU berücksichtigt (2). Das Bulletin Board schreibt den Ausführungsplan fest in die Time Table (3 bis 4). Als Antwort wird dem On-Demand Information Service eine Reservierungsbestätigung übergeben (5).	
Eingabedaten	Reservierungsanfragen.
Ausgabedaten	Antwort für jede Reservierungsanfrage.

Tabelle 2.24: SRWU-Scheduling

2.3.3 Übergabe eines Auftrags an eine PD

2.3.3.1 Übergabe einer gültigen SRWU

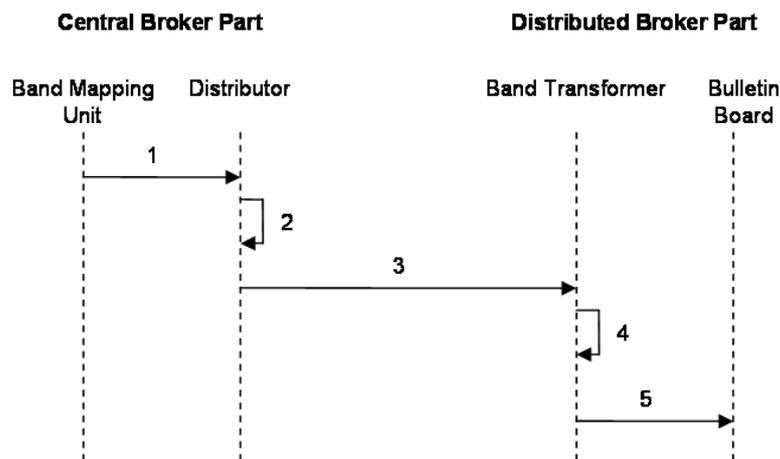


Abbildung 2.46: Use Case #13: Übergabe einer gültigen SRWU

Use Case #13	Übergabe einer gültigen SRWU
Use Case Beschreibung	
Die Band Mapping Unit übergibt dem Distributor eine SRWU, zur Übermittlung an eine PD (1). Die Id der Ziel-PD wird der SRWU entnommen und in eine Grid-Adresse umgewandelt (z.B. in eine IP-Adresse) (2). Anschließend erfolgt das Versenden an den Band Transformer der Ziel-PD (3). Anhand des Typs werden der SRWU WLM-Parameter zugeordnet (Priorität, Enklave usw.), die für eine optimale Verarbeitung sorgen (4). Zum Abschluss schreibt das Bulletin Board die SRWU in die Time Table. Dabei wird ihr der Verarbeitungsplan aus der Reservierungsphase zugeordnet (5). Dabei wird ihr der Verarbeitungsplan aus der Reservierungsphase zugeordnet (5).	
Eingabedaten	SRWU mit gültiger Reservierung.
Ausgabedaten	keine

Tabelle 2.26: Übergabe einer gültigen SRWU

2.3.3.2 Ablehnung einer ungültigen SRWU

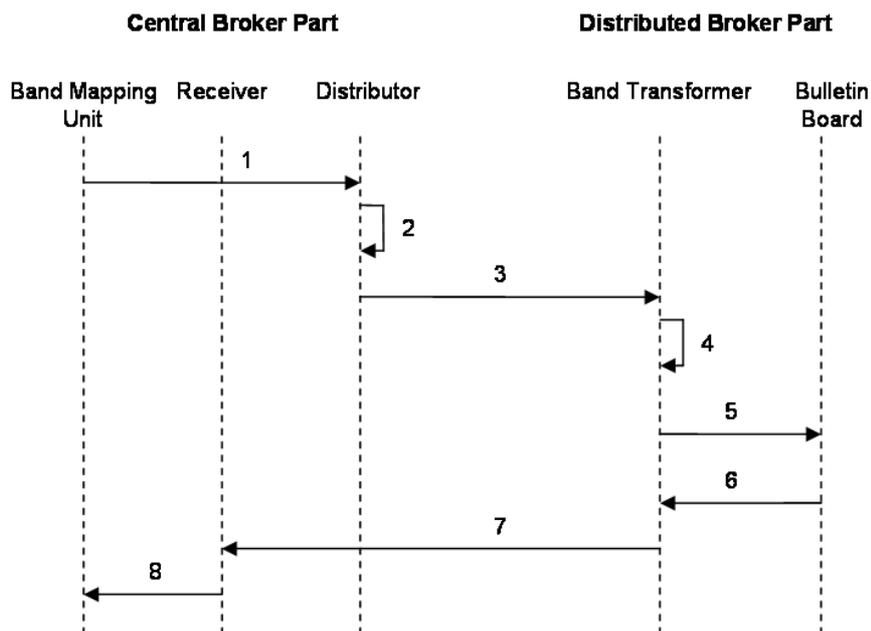


Abbildung 2.47: Use Case #14: Ablehnung einer ungültigen SRWU

Use Case #14	Ablehnung einer ungültigen SRWU
Use Case Beschreibung	
<p>Die Band Mapping Unit übergibt dem Distributor eine SRWU, zur Übermittlung an eine PD (1). Die Id der Ziel-PD wird der SRWU entnommen und in eine Grid-Adresse umgewandelt (z.B. in eine IP-Adresse) (2). Anschließend erfolgt das Versenden an den Band Transformer der Ziel-PD (3). Anhand des Typs werden der SRWU WLM-Parameter zugeordnet (Priorität, Enklave usw.), die für eine optimale Verarbeitung sorgen (4). Zum Abschluss schreibt das Bulletin Board die SRWU in die Time Table. Dabei wird ihr der Verarbeitungsplan aus der Reservierungsphase zugeordnet (5). Das Bulletin Board stellt fest, dass der Verarbeitungsplan nicht mehr gültig ist, was z.B. auf eine Verzögerungen bei der Übertragung zurückgeführt werden kann ist. Die SRWU wird dem Band Transformer als fehlerhaft zurückgegeben und von dort an den Receiver des Central Broker Parts gesendet (6 bis 7). Der Receiver erkennt, dass es sich dabei um eine abgewiesene SRWU handelt und übergibt sie sofort der Band Mapping Unit für eine neue Auktion (8).</p>	
Eingabedaten	SRWU mit gültiger Reservierung.
Ausgabedaten	Abgelehnte SRWU

Tabelle 2.28: Ablehnung einer ungültigen SRWU

2.3.3.3 Starten einer SRWU durch den Dispatcher

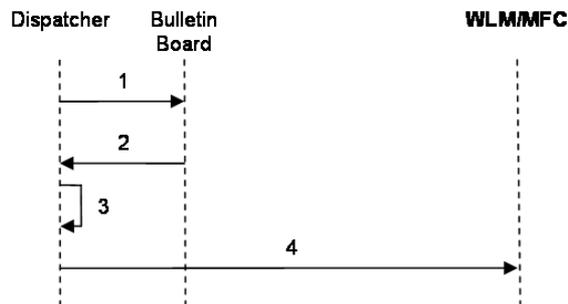


Abbildung 2.48: Use Case #15: Starten einer SRWU durch den Dispatcher

Use Case #15	Starten einer SRWU durch den Dispatcher
Use Case Beschreibung	
Der Dispatcher fordert periodisch das Bulletin Board auf, alle zur Ausführung bereiten SRWUs zu liefern (1). Das Bulletin Board überprüft die Time Table und gibt die gewünschten Informationen zurück (2). Der Dispatcher sucht sich die SRWUs heraus, die für die Ausführung vorgesehen sind. Dabei können auch nur Teile einer SRWU betroffen sein (3). Über die PD-Schnittstelle übergibt der Dispatcher die Aufträge an den Masterflow Controller bzw. dem Workload Manager der PD (4).	
Eingabedaten	keine
Ausgabedaten	SRWU oder Teile einer SRWU.

Tabelle 2.30: Starten einer SRWU durch den Dispatcher

2.3.4 Übergabe eines Response von einer PD

Die Verarbeitung einer SRWU findet auf einer PD statt. Bei interaktiven Aufträgen wird die Rückgabe einer Antwort oder eines Ergebnis erwartet. Die PD erstellt in solchen Fällen Response-SRWUs, die die gewünschten Informationen enthalten.

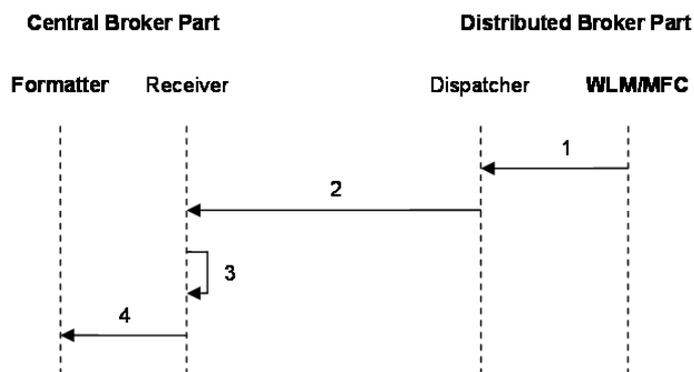


Abbildung 2.49: Use Case #16: SRWU-Scheduling

Use Case #16	Übergabe eines Response von einer PD
Use Case Beschreibung	
Nach vollständiger Verarbeitung einer SRWU durch den Masterflow Controller bzw. Workload Manager wird dem Dispatcher eine Response-SRWU übergeben (1). Diese wird sofort an den Receiver des Central Broker Parts gesendet (2). Der Receiver erkennt, dass es sich um eine Response-SRWU handelt und wertet statistische Daten für die SRWU aus, z.B. Anzahl ausgeführter SRWUs eines Typs pro Service Intervall, die im Bulletin Board veröffentlicht werden (3). Anschließend wird die Response-SRWU dem Formatter übergeben, der einen formatierte Ausgabe erzeugen kann (4).	
Eingabedaten	Response-SRWU.
Ausgabedaten	Response-SRWU.

Tabelle 2.32: Übergabe eines Response von einer PD

2.3.5 Data Collection

Das Sammeln von Informationen über die PDs erfolgt lokal durch den Distributed Broker Part und global durch den Central Broker Part. Unterstützt wird der Mechanismus durch ein Grid Information Service (GIS), der die gewünschten Daten in formatierter Form global zur Verfügung stellt.

2.3.5.1 Collection Service

Der Collection Service ist ein Dienst zur Aktualisierung der PD-Informationen im globalen Bulletin Board. Für die Datensammlung wird auf das GIS zugegriffen.

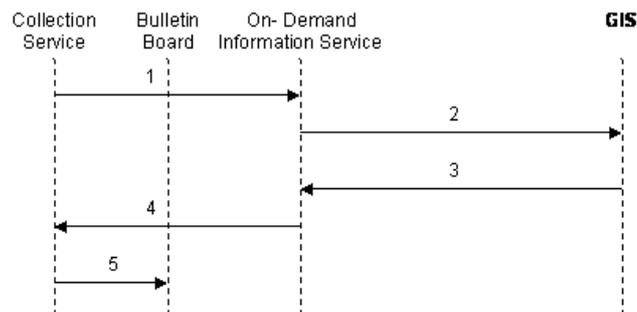


Abbildung 2.50: Use Case #17: Collection Service

Use Case #17	Collection Service
Use Case Beschreibung	
Der Collection Service startet periodisch die Aktualisierung der globalen Daten des Central Broker Parts. Dazu fordert er den On-Demand Information Service auf die Informationen über alle PDs vom GIS abzurufen (1). Der On-Demand Information Service ermittelt alle verfügbaren PDs und stellt für diese eine Anfrage an den Grid Information Service (2). Der GIS sendet die Rohdaten an den On-Demand Information Service, der sie wiederum an den Collection Service übergibt (3 bis 4). Der Collection Service fasst die Informationen zusammen und übergibt sie in formatierter Form dem Bulletin Board (5).	
Eingabedaten	keine
Ausgabedaten	Daten aller PDs <ul style="list-style-type: none"> • Auslastung • QoS • Durchsatz • Kapazität • etc.

Tabelle 2.34: Collection Service

2.3.5.2 Aktualisieren des GIS durch den Distributed Broker Part

Der Grid Information Service aktualisiert periodisch seine gespeicherten Informationen. Der Parameter TTL (time-to-live) gibt das Aktualisierungsintervall an.

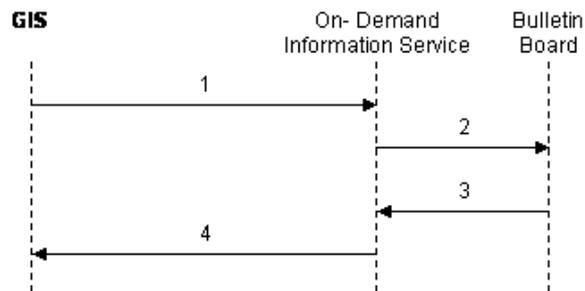


Abbildung 2.51: Use Case #18: Aktualisierung des GIS durch den Distributed Broker Part

Use Case #18	Aktualisierung des GIS durch den Distributed Broker Part
Use Case Beschreibung	
Über das API der Grid-Infrastruktur (z.B. das API des Globus-Toolkit) wartet der On-Demand Information Service auf Anfragen. Der GIS sendet eine Aufforderung an den On-Demand Information Service alle Informationen über die registrierten Messreihen zu liefern (1). Der On-Demand Information Service benachrichtigt das Bulletin Board, das die gewünschten Daten zurückgibt (2 bis 3). Diese werden über das API an den GIS übermittelt (4).	
Eingabedaten	Aktualisierungsanfrage
Ausgabedaten	Daten einer PD <ul style="list-style-type: none"> • Auslastung • QoS • Durchsatz • Kapazität • etc.

Tabelle 2.36: Aktualisierung des GIS durch den Distributed Broker Part

2.3.5.3 Sammeln lokaler PD-Daten durch die Data Collection Unit

Die Informationen über eine PD werden lokal durch den Distributed Broker Part gesammelt. Dazu greift er auf den Workload Manager oder Masterflow Controller zu, die entsprechende Daten über aktuelle Auslastung, Antwortzeit, freie Kapazität usw. zur Verfügung stellen.

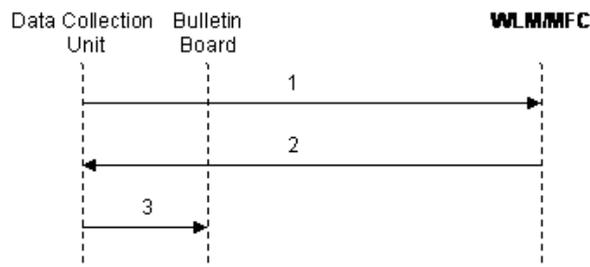


Abbildung 2.52: Use Case #19: Sammeln lokaler PD-Daten durch die Data Collection Unit

Use Case #19	Sammeln lokaler PD-Daten durch die Data Collection Unit
Use Case Beschreibung	
Die Data Collection Unit sendet in einem Service Intervall mehrmals Anfragen an den WLM/MFC, der Informationen über die aktuelle Auslastung, den Durchsatz, etc. liefert (1 bis 2). Diese Daten werden über den Zeitraum eines Service Intervalls gesammelt und anschließend an das Bulletin Board übergeben (3).	
Eingabedaten	keine
Ausgabedaten	Daten einer PD, über ein Service Intervall gesammelt <ul style="list-style-type: none"> • Auslastung • Durchsatz • Kapazität • etc.

Tabelle 2.38: Sammeln lokaler PD-Daten durch die Data Collection Unit

2.3.5.4 Aktualisierung der globalen PD-Daten erzwingen

Eine PD kann bei Bedarf den Collection Service veranlassen, ihre Daten im globalen Bulletin Board zu aktualisieren. Dadurch können größere Änderungen auf einer PD schnell veröffentlicht werden.

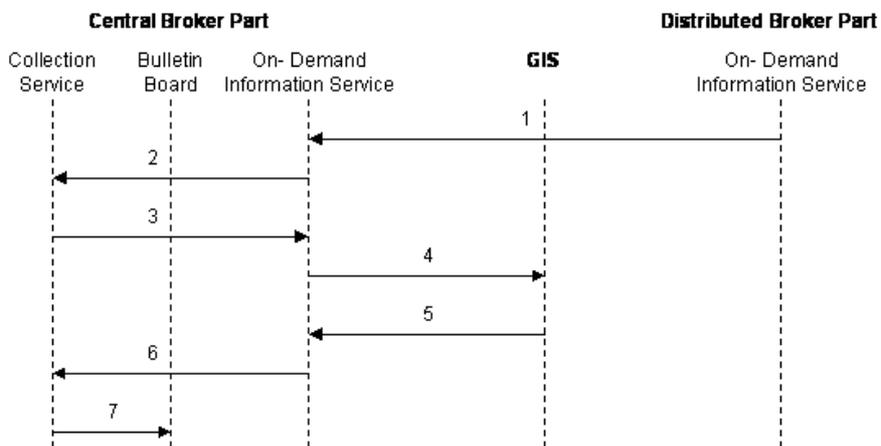


Abbildung 2.53: Use Case #20: Aktualisierung der globalen PD-Daten erzwingen

Use Case #20	Aktualisierung der globalen PD-Daten erzwingen
Use Case Beschreibung	
<p>Der Distributed Broker Part sendet über den On-Demand Information Service einen Message Container an den Collection Service des Central Broker Parts. Dieser enthält eine Aufforderung an den Collection Service die globalen Daten der entsprechenden PD zu aktualisieren (1 bis 2). Der Collection Service startet den Prozess, indem er den globalen On-Demand Information Service auffordert die Informationen über die spezielle PD vom GIS abzurufen (3). Der On-Demand Information Service stellt für die PD eine Anfrage an den Grid Information Service (4). Der GIS sendet die Rohdaten an den On-Demand Information Service, der sie wiederum an den Collection Service übergibt (5 bis 6). Der Collection Service fasst die Informationen zusammen und übergibt sie in formatierter Form dem Bulletin Board (7).</p>	
Eingabedaten	Message Container mit Aktualisierungsanfrage einer PD.
Ausgabedaten	<p>Daten einer PD</p> <ul style="list-style-type: none"> • Auslastung • QoS • Durchsatz • Kapazität • etc.

Tabelle 2.40: Aktualisierung der globalen PD-Daten erzwingen

2.3.6 Zusammenfassung

Die vorgestellten Use Cases wurden im Rahmen des Grobentwurfs des kommerziellen Grid Ressource Brokers erstellt. Sie verdeutlichen das wesentliche Verhalten des kGRB bei

- Der Interaktion mit einem Leistungsabnehmer
- Der Annahme und Verifikation einer SRWU in Verbindung mit einer Authentifizierung und Profilzuweisung
- Der Bestimmung einer geeigneten PD durch den Einsatz von Auktionsverfahren, der dynamischen Anpassung der Bänder und einem TimeTable-basierten Scheduling-Verfahren
- Dem Transport von SRWUs bzw. Response-SRWUs in einer Grid-Umgebung, sowie dem Starten einer SRWU auf einer PD
- Dem aktiven Sammeln und Aufbereiten von PD-Informationen in Verbindung mit einem Grid Information Service

Auf der Basis dieser Use Cases wurde das Feindesign und Implementierung des kommerziellen Grid Resource Brokers durchgeführt (siehe Kapitel 5.2 auf Seite 86).

Kapitel 3

Quality of Service und Service Level Agreements

Dieses Kapitel beschreibt die Grundlagen für den Einsatz von QoS- und SLA-Definitionen. Nach einer kurzen Einführung wird auf die speziellen Anforderungen des kommerziellen Grid Resource Brokers eingegangen.

Heutige Computersysteme bieten kaum eine Möglichkeit die Qualität von Serviceleistungen, wie Verarbeitungsgeschwindigkeit oder Antwortzeit, anzugeben und zu garantieren. Das Ziel ist vielmehr die Optimierung des Systems in einem Bereich, um die Leistungsfähigkeit zu erhöhen, z.B. Maximierung des Durchsatzes. Vor allem im kommerziellen Umfeld, wie Banken oder Versicherungen, entsteht mehr und mehr die Nachfrage nach Qualitätsgarantien. Für diese sind die Teilnehmer auch bereit entsprechend zu bezahlen [20]. Unter dem Quality of Service versteht man Angaben, die die minimalen Zusicherungen für ein System definieren. Nach [6] gibt es zwei Arten von QoS. Die Einteilung erfolgt in Hard QoS und Soft QoS. Die Angabe eines Hard QoS garantiert die Einhaltung konkreter Werte für einen abzuarbeitenden Auftrag, z.B. die Zusicherung einer konkreten Antwortzeit. Der Soft QoS beruht auf statistischen Messungen, die in einem festgelegten Zeitraum durchgeführt werden. Er gibt den prozentualen Anteil an Aufträgen an, für die ein konkreter Wert, z.B. eine bestimmte Antwortzeit, eingehalten werden muss. Das Anbieten eines Quality of Service stellt unterschiedliche Anforderungen an das System. Für einen Soft QoS müssen genaue Messungen über den angebotenen Service und die Auslastung des Systems (Performance Index) vorliegen. Unter der Voraussetzung, dass die Servicequalität mit steigendem Performance Index abnimmt, kann ein Mechanismus die Annahme von Aufträgen regulieren. Ein Hard QoS stellt wesentlich höhere Ansprüche an das System. Neben sehr genaue Angaben über eine Vielzahl von Systemparameter ist auch ein Schedulingmechanismus notwendig, der die Ausführung von Aufträgen exakt planen und garantieren kann. Echtzeitfähigkeit ist vermutlich eine Grundvoraussetzung für Hard QoS, da dadurch die Ausführung einer Serviceleistung erzwungen werden kann.

Während der QoS die Fähigkeiten des Systems ausdrückt, geben die Service Level Agreements die Forderungen eines Auftrags wieder. SLAs geben also die minimalen Qualitätszusicherungen an, die vom System erfüllt werden müssen. Diese können nicht willkürlich festgelegt werden. SLAs sind vertragliche Vereinbarungen, die beide Parteien gemeinschaftlich aushandeln. Ebenso wird ein Preisniveau für die erbrachten

Leistungen bzw. Strafgeelder für Verstöße festgelegt.

3.1 QoS des kommerziellen Grid Resource Brokers

Die unterschiedlichen Arten von Qualität, die ein System bieten kann, sind sehr vielfältig. Für den kommerziellen Grid Resource Broker wurden folgende Qualitätsmerkmale definiert [11].

- Deadline
- Timeline
- Budget
- Velocity
- Failure

Unterstützt wird nur ein Soft QoS. Für die Einhaltung ist der BDAfPD verantwortlich. Im Rahmen der Simulation ist auch die Einhaltung eines Hard QoS möglich, sofern die Plugins der Ressourcen Manager entsprechend ausgelegt sind.

3.1.1 Deadline

Unter der Deadline versteht man einen Zeitpunkt, bei dem ein Auftrag fertiggestellt sein muss. Der entsprechende QoS-Parameter gibt an, wie gut der Zeitpunkt eingehalten wird.

3.1.2 Timeline

Unter der Timeline versteht man einen Zeitpunkt, bei dem ein Auftrag gestartet werden muss. Der entsprechende QoS-Parameter gibt an, wie gut der Zeitpunkt eingehalten wird.

3.1.3 Budget

Das Budget gibt einen maximalen Preis an, für den ein Auftrag abgearbeitet wird. Der QoS-Parameter gibt an, wie gut der obere Grenze eingehalten wird.

3.1.4 Velocity

Velocity ist ein Wert für die Geschwindigkeit, mit der ein Auftrag abgearbeitet wird. Er repräsentiert somit die erwartete Antwortzeit. Der QoS-Parameter gibt an, wie gut die Antwortzeit eingehalten wird.

3.1.5 Failure

Der Wert Failure repräsentiert die Fehlerrate bei der Ausführung eines Auftrags. Der QoS-Parameter gibt eine Garantie, über die Anzahl der bearbeiteten Aufträge.

Kapitel 4

Mikroökonomische Ansätze und Auktionsverfahren

In diesem Kapitel wird das Konzept der Mikroökonomie im Bereich der Lastverteilung in Computernetzen beschrieben. Die Grundlage bildet eine rein marktwirtschaftliche Sichtweise auf ein System. Als Beispiel wird anschließend das Modell von Ferguson beschrieben, das diesen Ansatz vollständig umsetzt [8]. Als Repräsentant eines klassischen Lastverteilungsverfahrens wird der HOP1-Algorithmus kurz vorgestellt [8]. Der dritte Abschnitt widmet sich dem kommerziellen Grid Resource Broker. Es wird geklärt, welche Teile des mikroökonomischen Ansatz hier Verwendung finden. Den Abschluss bildet eine Übersicht über die gängigsten Auktionsverfahren.

4.1 Einleitung

Der mikroökonomische Ansatz stellt den Versuch dar, für eine verteilte Computerumgebung (Cluster), einen Lastverteilungsmechanismus zu entwerfen [8, 7, 17]. Dieser weicht von den klassischen Methoden ab, indem kein Wissen über die Auslastung der Ressourcen vorhanden ist. Das Konzept orientiert sich an einer marktwirtschaftlichen Umgebung, wo Preise für Produkte festgelegt werden und diese sich durch Angebot und Nachfrage selbst regulieren. Die Mikroökonomie funktioniert auf ähnliche Weise. Hier bieten Computerressourcen ihre Kapazität für einen Preis an. Ein Job, der auf einer Ressource ausgeführt werden soll, muss diesen Betrag von einem festen Budget bezahlen. Die Preise sind variabel und werden über ein Auktionsverfahren ermittelt. Dadurch stehen die Jobs in einer permanenten Konkurrenz um die Ressourcen.

Die Lastverteilung ergibt sich aus folgender Überlegung. Der Preis einer Ressource wird bei einer hohen Auslastung entsprechend hoch sein, da vermutlich sehr viele Jobs dafür geboten haben. Ressourcen, die eine geringere Auslastung aufweisen, werden dementsprechend niedrigere Preise verlangen. Da die Jobs möglichst wenig von ihrem Budget bezahlen wollen, werden sie die günstigere Ressource wählen wodurch sich die Last verteilt.

4.2 Modell von Ferguson

Ferguson hat ein Modell zur Lastverteilung in verteilten Umgebungen entwickelt, das unter mikroökonomischen Gesichtspunkten arbeitet. In seinem Vorschlag unterscheidet er zwischen CPU- und Netzwerkressourcen. Dabei verbindet das Netzwerk die CPUs miteinander. Jede Ressource besitzt ein individuelles Gewicht, als Leistungsfaktor und einen individuellen Preis. Die CPUs verwalten Bulletin Boards, wo die Daten aller benachbarter CPUs eingetragen werden. Die Informationen werden nach Bedarf aktualisiert. Jobs stellen die Verbraucher dar, die die Kapazität der Ressourcen nutzen wollen. Dabei nehmen sie CPU- und Netzwerkkapazität in Anspruch. Die Anforderungen an die Ressourcen sind am Anfang für die Jobs bekannt. Jede CPU und jeder Job erhält einen Agenten, der für die Umsetzung einer Geschäftsstrategie verantwortlich ist. Für eine CPU steht die Maximierung des Gewinns im Vordergrund. Der Job möchte so preiswert wie möglich verarbeitet werden. Er kann dabei eine Wahl zwischen Kosten und Geschwindigkeit treffen.

Die marktwirtschaftlichen Prozesse werden durch ein Auktionsverfahren umgesetzt. Angebot und Nachfrage sollen die Preise bestimmen. Jobs besitzen ein Budget, von dem sie nur so viel bezahlen wollen, wie gerade nötig ist. Dabei haben sie die Möglichkeit auf andere CPUs zu wechseln. Beachtet werden müssen die Kosten für den Transport, die die Netzwerkressourcen erheben. Durch die Bulletin Boards können sich die Jobs über die Netzwerk- und CPU-Preise informieren. Bei einer CPU handelt es sich bei den Kosten nur um einen Richtwert. Der endgültige Preis wird durch eine Auktion bestimmt. Über die Kosten und den Geschwindigkeitsfaktor wählt eine Job eine CPU aus und nimmt an der Versteigerung ihrer Kapazität teil. Der Höchstbietende erhält den Zuschlag und wird anschließend auf der CPU verarbeitet. Die Jobs verwenden zu Beginn nicht ihr gesamtes Budget für ein Gebot. Das Ziel ist, so günstig wie möglich verarbeitet zu werden. Wenn Jobs eine Auktion verlieren, erhöhen sie den maximalen Wert für ein Gebot und nehmen an einer weiteren Auktion teil. Die CPUs legen ihre veröffentlichten Preise entsprechend dem Höchstgebot fest. Je höher dieser Wert wird, desto weniger Jobs nehmen an der Auktion teil. Die Nachfrage sinkt und demzufolge auch das Höchstgebot. Nehmen keine Jobs an einer Versteigerung teil, wird der veröffentlichte Preis gesenkt. Eine ausführliche Beschreibung des Modells ist unter [8] zu finden.

4.3 HOP1

Der HOP1 ist ein klassischer Algorithmus zur Lastverteilung in einer P2P-Umgebung. Im Gegensatz zum mikroökonomischen Verfahren wird hier versucht auf direktem Weg den Workload optimal zu verteilen. Die Umgebung besteht aus einer Anzahl Prozessoren, die miteinander verbunden sind. Das Netzwerk ermöglicht eine Kommunikation bzw. den Transport von Jobs. Jeder Prozessor führt eine Tabelle mit den Auslastungsinformationen der anderen Prozessoren. Die Aktualisierung der Daten wird bei jeder Workloadänderung durchgeführt. Auf einem Prozessor befindet sich eine Warteschlange, in der neue Jobs abgelegt werden. Wird der Prozessor frei, erhält er den kleinsten Job aus der Warteschlange für eine Verarbeitung (Shortest Job First, SJF).

Die Lastverteilung wird von jedem Prozessor durchgeführt. Dabei wird nach dem Prinzip vorgegangen die Leistung eines anderen Prozessors nie weiter als die eigene zu verschlechtern. In den meisten Fällen wird sich die Leistung durch Erhöhung der Auslas-

tung (PI) verschlechtern. Jeder Prozessor besitzt für den PI eine obere Schranke. Beim Eintreffen eines neuen Jobs gelangt dieser in die Warteschlange, die entsprechend der Jobgröße sortiert wird. Anschließend findet eine neue Berechnung der Auslastung statt. Liegt diese höher als die obere Schranke, wird versucht eine Lastverteilung durchzuführen. Der Algorithmus sucht den Prozessor mit dem kleinsten PI. Die lokale Auslastungstabelle ermöglicht dabei einen schnellen Zugriff auf die benötigten Informationen. Anschließend wird versucht einen Job an diesen Prozessor abzugeben. Hierbei können zwei Strategien zum Einsatz kommen, best-fit und first-fit. Die best-fit Variante sucht den Job aus der eigenen Warteschlange, der unter der Vorgabe, den fremden PI nicht weiter als den eigenen zu verschlechtern, am besten geeignet ist. Die first-fit Methode versucht den ersten Job abzugeben. Der Vorgang wird solange wiederholt, wie eine Jobabgabe möglich ist, d.h. der PI nicht verschlechtert wird.

4.4 Auktionsverfahren

4.4.1 Sealed Bid Auction

Die Sealed Bid Auction ist ein Auktionsverfahren, das in einem Durchlauf abgeschlossen wird. Jeder Bieter gibt für das zu ersteigende Objekt ein geheimes und verbindliches Gebot ab. Vor dem Abschluss der Auktion werden diese Werte nicht veröffentlicht. Gewinner ist der Bieter, der das höchste Gebot abgegeben hat. Das ersteigerte Objekt wird für den Wert des höchsten Gebots verkauft.

4.4.2 Dutch Auction

Die Dutch Auction ist ein Auktionsverfahren, bei dem genau ein Gebot abgegeben wird. Für das zu ersteigende Objekt wird ein initialer Startwert festgelegt und veröffentlicht. Dieser wird solange stufenweise verringert, bis ein Gebot abgegeben wird. Der Bieter erhält den Zuschlag für das Objekt. Der Preis entspricht dem verringerten Startwert bei Auktionsende.

4.4.3 English Auction

Die English Auction ist ein Auktionsverfahren, bei dem mehrere Durchgänge möglich sind. Ein Bieter gibt für das zu ersteigende Objekt ein Gebot ab. Dieses wird veröffentlicht und ist jedem zugänglich. Andere Bieter dürfen neue Gebote abgeben, die über dem alten liegen müssen. Eine andere Variante dieses Verfahrens erhöht bei Abgabe eines neuen Gebots das alte um einen konstanten Wert. Das Ende der Auktion bestimmt ein festgelegter Zeitpunkt oder ein Zeitraum, in dem nicht mehr geboten wird. Gewinner ist der Bieter, mit dem höchsten Gebot. Das ersteigerte Objekt wird für den Wert des höchsten Gebots verkauft.

4.4.4 Hybrid Auction

Die Hybrid Auction ist ein Auktionsverfahren, das durch einen Annäherungsprozess versucht einen optimalen Preis für ein Objekt zu bestimmen. Zu Beginn wird ein Startwert (Preis) festgelegt. Gehen für das zu ersteigende Objekt mehrere Gebote ein,

wird der Startwert erhöht. Findet sich kein Bieter, so wird der Startwert verringert. Dieser Vorgang wiederholt sich so lange, bis genau ein Gebot abgegeben wird. Der Bieter erhält den Zuschlag für das Objekt, das für den aktuellen Preis verkauft werden muss.

4.4.5 Continuous Double Auction

Bei der Continuous Double Auction handelt es sich um ein Auktionsverfahren, bei dem es zwei bietende Gruppen gibt. Die erste Gruppe bietet Objekte über eine Art Sealed Bid Auction an, d.h. die abgegebenen Gebote sind Preise für die Objekte. Diese werden in einer sortierten Liste verwaltet. Das niedrigste Gebot steht an der ersten und das höchste an der letzten Stelle. Die zweite Gruppe versucht die Objekte ebenfalls über eine Sealed Bid Auction zu ersteigern. Die abgegebenen Gebote kennzeichnen die Preise, die für ein Objekt gezahlt werden würde. Die Werte werden in einer sortierten Liste verwaltet, wobei der erste Eintrag der größte und der letzte der niedrigste Wert ist. In der nächsten Stufe werden beide Listen miteinander verglichen. Sollte das niedrigste Gebot der ersten Gruppe kleiner sein als das höchste Gebot der zweiten Gruppe, so erhalten beide den Zuschlag für einander. Beide Listen werden der Reihe nach abgearbeitet, bis keine Zuordnungen mehr möglich sind.

Das Verfahren funktioniert nur dann vernünftig, wenn es um Objekte der gleichen Art geht. Für ein heterogenes Angebot müssen unterschiedliche Listen geführt werden.

4.5 Ökonomischer Ansatz beim kommerziellen Grid Resource Broker

Der kommerzielle Grid Resource Broker verwendet mikroökonomische Ansätze zur Optimierung der Lastverteilung und zur Ermittlung eines Preises für die Serviceleistung einer PD (On-Demand Value). Dieser Marktwert berücksichtigt das aktuelle Verhältnis von Angebot (PD) und Nachfrage (SRWU), um den Broker langfristig in einem kommerziellen Umfeld konkurrenzfähig zu halten. Trotz dieser Vorgabe soll der Gewinn maximiert werden.

In dem administrativen Bereich des kommerziellen Grid Resource Brokers werden die Preise für eine Serviceleistung von den PDs in eigener Verantwortung festgelegt. Dabei kommt eine On-Demand Value Funktion zum Einsatz, die sehr unterschiedliche Parameter berücksichtigen kann. Zu den wichtigsten gehören

- Auslastung der PD (PI)
- Produktklasse (Band) und Service Level Agreements der SRWU
- QoS-Profil der PD unter Berücksichtigung der aktuellen Auslastung
- Produktionskosten der PD und minimaler Preis
- Nachfragevolumen der SRWUs
- Rabatte
- Preisentwicklung bei langfristigen Reservierungen

- u.v.m.

Der Einsatz komplexer Funktionen aus dem Börsenbereich sind hier vorgesehen. Eine PD berechnet ihren Preis auf Anfrage. Diese Anfrage ist Teil eines Auktionsverfahrens, das zwischen Central Broker Part und Distributed Broker Part stattfindet. Bei der Auktion stehen die PDs miteinander in einem Konkurrenzkampf um den Erhalt von SRWUs. Sie werden demnach versuchen ihren Service so günstig wie möglich, unter Beachtung ihrer On-Demand Funktion, anzubieten.

4.5.1 Auktion beim kommerziellen Grid Resource Broker

Der Einsatz eines Auktionsverfahrens ermöglicht das Aushandeln von Preisen unter einem marktwirtschaftlichen Gesichtspunkt. SRWUs repräsentieren Käufer von Serviceleistungen, mit einem festgelegten Budget. Die PDs bilden die Anbieter oder Verkäufer des Service, mit einem bestimmten Preis. Die Zuweisung der SRWUs an die PDs wird von dem Auktionsmechanismus übernommen. Der Central Broker Part übernimmt die Aufgabe der Auktionsleitung. Die Auction Unit des Distributed Broker Part vertritt dabei das Interessen einer PD. Der allgemeine Ablauf gliedert sich in die folgenden Abschnitte.

1. Erstellung von Gebotsanfragen
2. Erstellung von Geboten
3. Auswertung von Geboten
4. Reservierungsablauf

Unter Beachtung dieser Einteilung kann jede Auktionsform eingesetzt werden.

4.5.1.1 Erstellung von Gebotsanfragen

Dieser Schritt erstellt für jede vorhandene SRWU eine Gebotsanfrage, mit Informationen über Anzahl der Service Units, Band, Preise und Service Level Agreements. Bereits hier wird eine Vorauswahl an PDs getroffen, die für eine Auktion in Frage kommen.

1	Erstelle eine Gebotsanfrage	Am Anfang der Auktion liegt dem Central Broker Part eine Liste mit SRWUs vor, die verteilt werden sollen. Für jeden Eintrag wird eine Gebotsanfrage mit den Informationen einer SRWU erstellt.
2	Bestimme den Feasible-Set	Der Feasible-Set für eine Gebotsanfrage enthält alle PDs, die aufgrund ihrer veröffentlichten Preise für eine Auktion in Frage kommen.
3	Bestimme den Demand-Set	Der Demand-Set wird aus dem Feasible-Set gebildet. Er enthält alle PDs, die spezielle Anforderungen der Gebotsanfrage erfüllen. Da der BDAfPD zum Einsatz kommt, muss Servicekapazität im Band der Gebotsanfrage angeboten werden.
4	Sende Gebotsanfrage an die PDs im Demand-Set	Die PDs, die für eine Auktion vorgesehen sind, müssen über ihrer Teilnahme unterrichtet werden. Jedem Eintrag im Demand-Set wird daher eine Gebotsanfrage zugesendet. Sollten mehrere SRWUs an der selben PD interessiert sein, werden die entsprechenden Gebotsanfragen in einer Nachricht übermittelt.

Tabelle 4.1: Erstellung von Gebotsanfragen

4.5.1.2 Erstellung von Geboten

Die Gebote werden stets durch den Distributed Broker Part auf den PDs erstellt. Als Grundlage dienen die Gebotsanfragen, die vom Central Broker Part übermittelt wurden. Die On-Demand Value Funktion liefert den Preis, d.h. die Höhe des Gebots.

1	Starte Verarbeitung der Gebotsanfragen	Das Erstellen von Geboten geschieht On-Demand, d.h. wenn der Central Broker Part eine entsprechende Anfrage stellt. Das Eintreffen einer Gebotsanfrage startet den Mechanismus.
2	Führe Reservierungsverarbeitung durch	Ein wichtiger Punkt bei einer Auktion ist die Verarbeitung von Reservierungen. Diese werden vom Central Broker Part benutzt um den Zuschlag für eine gewonnene Auktion zu bestätigen. Eine Reservierung sichert der PD eine SRWU zu. Die entsprechende Servicemenge (SU) muss jetzt endgültig in der Time Table reserviert werden. Diese Kapazität steht für weitere Auktionen nicht mehr zur Verfügung.
3	Erstelle Gebote	Beim Erstellen von Geboten werden Preise über die On-Demand Value Funktion generiert. Angegeben werden stets die Kosten für eine Service Unit. Die Grundlage für die Berechnung bilden die Gebotsanfragen. Hier können Angaben über Serviceanforderungen, Qualität oder Band entnommen werden.
4	Führe Gebotsabgaben durch	Anhand der Gebotsanfragen wurden Gebote erstellt und mit Preisen versehen. In diesem Schritt wird überprüft, welche Gebote tatsächlich abgegeben und welche verworfen oder abgelehnt werden. Zu dem wichtigsten Kriterium gehört die vorhanden Kapazität an Service Units aber auch der erwartete Gewinn. Wenn ein Gebot abgegeben wird, muss es in die Time Table eingetragen werden. Bereits an dieser Stelle erfolgt eine zeitliche Planung, d.h. die Verteilung der Service Units auf die Perioden in der Time Table wird durchgeführt.
5	Übermittle Gebote an den Central Broker Part	Die Auswahl an abgegebenen Geboten wird an den Central Broker Part gesendet. Damit sind diese für die PD bindend.

Tabelle 4.2: Erstellung von Geboten

4.5.1.3 Auswertung von Geboten

Die Auswertung der Gebote findet im Central Broker Part statt. Hier wird jeder SRWU das optimale Gebot zugewiesen, wie es durch die Auktion ausgehandelt wird. Es sind aber auch Strategien möglich, die im Widerspruch zur Mikroökonomie stehen und unter Umständen einer konkurrenzfähigen Preisentwicklung entgegenwirken. So können zum Beispiel Gebote spezieller PDs bevorzugt werden, um deren Auslastung zu erhöhen.

1	Gebote sammeln	Vor der Auswertung müssen die Gebote aller PDs gesammelt werden, für die eine Gebotsanfrage abgegeben wurde. Die Beachtung eines Timeouts ist besonders wichtig, um die Verzögerung gering zu halten.
2	Gebote ordnen	Die eingetroffenen Gebote der PDs erhalten eine Reihenfolge, die ihre Wichtigkeit widerspiegelt. In dem ökonomischen Umfeld der Auktion spielt der Preis die übergeordnete Rolle.
3	Gebot wählen	Das beste Gebot wird als Gewinner der Auktion ausgewählt.
4	Reservierung durchführen	Für das Gebot, das die Auktion gewonnen hat, muss eine Überprüfung auf der PD erfolgen. Diese stellt sicher, dass die Versteigerung zulässig ist. Der Vorgang wird durch den Reservierungsmechanismus realisiert.

Tabelle 4.3: Auswertung von Geboten

4.5.1.4 Reservierungsablauf

Der Vorgang der Reservierung ist ein Verwaltungsprozess zum Abschluss einer Auktion. Ein Gebot erhält durch den Central Broker Part den Zuschlag. Die Reservierung sorgt für die Bestätigung des Zuschlags durch den Distributed Broker Part. Diese berechtigt eine SRWU auf einer PD verarbeitet zu werden. Bei der Überprüfung wird getestet, ob dem Zuschlag ein reguläres Gebot zugrunde liegt.

1	Reservierung erstellen	Eine Reservierung enthält wichtige Informationen über die Auktion. Die wichtigste ist die Id des Gebots. Im einfachsten Fall reicht diese Information aus, um eine Reservierung bestätigen zu können. Bei komplexen Auktionsformen können aber noch weitere Daten definiert werden.
2	Reservierung versenden	Die PD, die das Sieger-Gebot abgegeben hat, erhält die Reservierung. Das Versenden erfolgt zusammen mit den Geboten. Dadurch lässt sich der Nachrichtenverkehr reduzieren.
3	Reservierung prüfen	Bevor die Auction Unit neue Gebote erstellt werden die Reservierungen verarbeitet. Die Reihenfolge ist wichtig, um die entsprechende Kapazität vor einer neuen Versteigerung zu schützen. Jedes Gebot muss in der Time Table eingetragen sein. Nur wenn ein Eintrag vorliegt ist die Reservierung zulässig.
4	Reservierung eintragen	Wenn eine Reservierung zulässig ist, wird diese in der Time Table eingetragen und somit die Kapazität vor weiteren Auktionen geschützt. Das Gebot wird anschließend aus der Tabelle entfernt.
5	Reservierung bestätigen	Dem Central Broker Part wird eine Bestätigung der Reservierung geschickt. Damit ist die Auktion abgeschlossen und die Kapazität der SRWU auf einer PD fest eingeplant.
	SRWU versenden	Besitzt eine SRWU eine gültige Reservierung, wird sie an die entsprechende PD verschickt.

Tabelle 4.4: Reservierungsablauf

4.5.2 Spezielles Auktionsverfahren

Die Ablaufsteuerung des Auktionsmechanismus unterstützt den Einsatz verschiedener Auktionsverfahren. Für den Einsatz im Hochlastbereich unter Verwendung des kommerziellen Grid Resource Brokers werden Optimierungen empfohlen. Dabei steht hauptsächlich die Reduzierung des Nachrichtenverkehrs im Vordergrund, der bei Auktionen auftritt. Auf diese Weise können wichtige Netzwerkressourcen eingespart werden und der Verarbeitungsaufwand in den Brokerkomponenten wird verringert. Einen weiteren wichtigen Faktor stellt die Wartezeit der SRWUs dar. Iterative Auktionsverfahren (Englich Auction, Dutch Auction) nähern sich über mehrere Auktionsdurchläufe dem Ergebnis an. Je länger diese Auktion dauert, desto geringer ist die Zeit, die für die eigentliche Verarbeitung der SRWU zur Verfügung steht.

Die Sealed Bid Auction ermöglicht den Abschluss einer Auktion in einem Durchgang. Das optimierte Auktionsverfahren basiert daher auf dieser Methode. Die Optimierung erfolgt durch den Einsatz spezieller Gebote, die durch den Central Broker Part flexibel verarbeitet werden können. Bei der klassischen Sealed Bid Auction wird für jede Gebotsanfrage ein eigenes Gebot erstellt. Dabei können nur so viele Gebote erstellt werden, wie freie Kapazität vorhanden ist. Die restlichen Gebotsanfragen mussten bereits abgelehnt werden. Bieten mehrere PDs für die gleiche Gebotsanfrage, erhält höchstens eine PD den Zuschlag. Die Kapazität der übrigen Gebote geht bei dieser Auktion verloren. Erst durch weitere Auktionsdurchgänge kann diese Kapazität schrittweise belegt werden.

Die optimierte Variante der Sealed Bid Auction erlaubt den PDs eine Kapazität zu versteigern, ohne diese an eine bestimmte Gebotsanfrage zu binden. Angegeben wird nur das Band und die erlaubte Servicemenge. Diese kann auch über mehrere Perioden (Service Intervalle) angeboten werden. Der Central Broker Part darf die angebotene Servicemenge in eigener Verantwortung mit SRWUs belegen. Dabei werden ökonomische und leistungsorientierte Gesichtspunkten beachtet, d.h. möglichst preiswert und mit guter Lastverteilung und Beachtung der SLAs. In der Reservierungsphase prüfen die PDs die Belegungsvorschläge des Central Broker Parts und führen unter Umständen noch eigene Optimierungen in der Belegung durch (z.B. zur besseren Einhaltung der SLAs). Je besser die Kenntnisse des Central Broker Part über die Belegungsstrategie (Scheduling-Strategien) der PDs sind, desto optimaler kann das gesamte System arbeiten. Durch die zentrale Zuweisung der SRWUs an die Gebote kann eine wesentlich größere Kapazität in einem Auktionsdurchgang belegt werden. Der Nachrichtenverkehr reduziert sich und die SRWUs gelangen schneller auf die Ziel-PD. Dieses Verfahren wird hier als Capacity Sealed Bid Auction bezeichnet.

Kapitel 5

Implementierung und Simulation

5.1 GridSim und SimJava

Das GridSim ToolKit 2.1 ist eine Sammlung von Java-Klassen, über die der Aufbau und Betrieb einer Grid-Umgebung oder eines Clusters simuliert werden kann. Es ist Teil des GridBus Project und wurde von Manzur Murshed und Rajkumar Buyya entwickelt und betreut. Als Open Source Projekt kann GridSim beliebig erweitert oder verändert werden [5].

5.1.1 SimJava

Die Basis für die GridSim-Simulation bildet SimJava-1.2. Dabei handelt es sich um eine ereignisorientierte diskrete Simulationsumgebung für Java. Für die Visualisierung und Animation des Ablaufs steht eine entsprechende Schnittstelle zur Verfügung [14].

SimJava basiert auf der Verwaltung von Entitäten (Entity). Dabei handelt es sich um Java-Threads, die auf Ereignisse (Event) reagieren können. Ein Event ist ein Java-Objekt, das zu einem definierten Zeitpunkt an eine Entität übergeben wird. Die Entität führt eine beliebige Verarbeitung durch. Dabei können neue Events erzeugt werden, die an SimJava übergeben werden. Die Dauer der Verarbeitung hat keinen Einfluss auf die interne Uhr von SimJava. Die Verarbeitung verläuft für die Simulation zeitlos. Über die SimJava-Uhr wird der Zeitpunkt ermittelt, an dem ein Event verteilt wird. Erst wenn alle Events eines Zeitpunktes verarbeitet wurden, wird zu dem nächsten Zeitpunkt übergegangen, an dem ein Event vorliegt. Sollte sich keines mehr finden lassen, wird die Simulation beendet. Bei der Erstellung eines Events können nicht nur die Ziel-Entität und der Ausführungszeitpunkt festgelegt werden. Es steht auch ein Feld zur Verfügung, durch das Daten übermittelt werden können.

5.1.2 GridSim

Das GridSim Toolkit 2.1 basiert auf der SimJava-Simulationsumgebung. Alle Komponenten der Grid-Umgebung werden als SimJava-Entitäten definiert. Die Simulation des

Netzwerkverkehrs und der verbrauchten Rechenzeit erfolgt über SimJava-Events. Die Hauptbestandteile einer GridSim-Umgebung sind

- Netzwerkressourcen
- Grid-Ressourcen
- GIS
- Grid Resource Broker

5.1.2.1 Netzwerksimulation

Die Simulation des Netzwerks in GridSim erfolgt durch das Versenden von Events. Der Datenteil dient zum Aufnehmen und Transportieren der Nachricht bzw. der Daten. Durch die Angabe einer Verzögerung beim Eintreffen des Events, können unterschiedliche Netzwerkgeschwindigkeiten simuliert werden.

5.1.2.2 CPU-Simulation

Die Verarbeitung von Aufträgen auf einer CPU erfolgt über das Versenden von Events. Die benötigte Rechenzeit wird durch die Angabe einer Verzögerung simuliert. Das Eintreffen des Events markiert das Ende der Verarbeitung. Die Simulation mehrerer Aufträge auf einer CPU vergrößert die Verzögerung und verschiebt den Endzeitpunkt entsprechend weit in die Zukunft.

5.1.2.3 Grid-Ressourcen

Ein Grid-Ressource besitzt sowohl eine Netzwerkschnittstelle, als auch eine beliebige Anzahl an CPUs. Für die Verarbeitung wird ein Preis festgelegt, den die Aufträge bezahlen müssen. Die CPU-Ressourcen können im Space-Shared Modus (nur ein Auftrag pro CPU) oder in einem Time-Shared Modus (beliebig viele Aufträge auf einer CPU) arbeiten.

5.1.2.4 GIS

Der GIS simuliert einen Grid Information Service für die Grid-Umgebung. Dabei handelt es sich jedoch nur um ein Ressourcen-Verzeichnis, das die Adressen der verfügbaren Grid-Ressourcen enthält. Die gewünschten statischen und dynamischen Informationen (Architektur, Preis, Auslastung, usw.) müssen direkt von der Grid-Ressource abgefragt werden.

5.1.2.5 Grid Resource Broker

Der Grid Resource Broker kann beliebig viele Jobs von mehreren Benutzern entgegennehmen und verteilt diese für die Verarbeitung auf die Grid-Ressourcen. Der Vorgang gliedert sich in mehrere Phasen. In der ersten Phase werden über den GIS die verfügbaren Grid-Ressourcen ermittelt. Von diesen ruft der Grid Resource Broker Informationen über Preis, Auslastung, Leistungsfähigkeit, Aufbau usw. ab. In der zweiten Phase wird für jeden Job eine passende Grid-Ressource ermittelt. Der Grid Resource Broker unterstützt 3 Strategien

Time-Optimisation Für jede Grid-Ressource wird der Zeitpunkt berechnet, an dem ein zugewiesener Job beendet wird. Die Grid-Ressourcen werden nach aufsteigender Endezeit sortiert. Entsprechend dieser Reihenfolge wird der Grid-Ressource der Job zugewiesen, die als erstes unter Berücksichtigung eines gesetzten Budgets für die Ausführung geeignet ist.

Cost-Optimisation Die Grid-Ressourcen werden nach aufsteigenden Kosten sortiert. Entsprechend dieser Reihenfolge werden sie solange mit Jobs belegt, wie die gestzten Endetermine (Deadlines) eingehalten werden.

Cost-Time-Optimisation Diese Strategie kombiniert Time- und Cost-Optimisation. Dazu werden Gruppen erstellt (Resource Groups-RGs), die Grid-Ressourcen mit den den gleichen Kosten enthalten. Diese Gruppen werden nach aufsteigenden Kosten sortiert (Cost-Optimisation) und entsprechend dieser Reihenfolge behandelt. Innerhalb einer Gruppe werden Jobs per Time-Optimisation zugewiesen.

In der dritten Phase werden die Jobs an die Grid-Ressourcen übergeben. Dort werden sie anschließend verarbeitet. In der letzten Phase werden die beendeten Jobs wieder von den Grid-Ressourcen entgegengenommen.

5.2 Implementierung des kommerziellen Grid Resource Brokers

5.2.1 Einleitung

Im Rahmen dieser Diplomarbeit wurde der kommerziellen Grid Resource Broker auf Basis des GridSim Toolkits 2.1 umgesetzt. Die Implementation umfasst die Teile

- Bestimmung einer geeigneten PD
- Übergabe eines Auftrags an eine PD
- Übergabe eines Response von einer PD
- Data Collection

Aufgrund der unzureichenden GIS-Unterstützung durch GridSim, arbeitet der Teil Data Collection nur in einem passiven Modus, d.h. die PDs sind für die Aktualisierung ihre Daten im globalen Bulletin Board selbst verantwortlich. Diese senden periodisch Message Container an den Collection Service, die die entsprechenden Informationen enthalten.

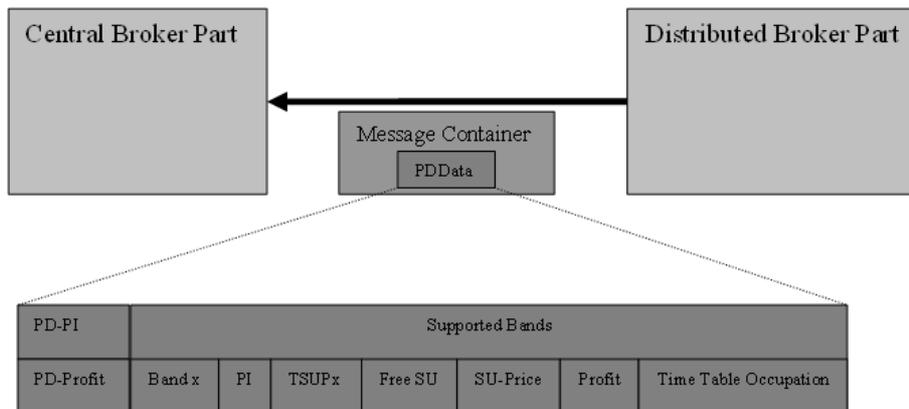


Abbildung 5.1: PD-Informationen

Für die Simulation nicht implementiert ist der Teil

- Annahme und Verifikation einer SRWU

Die entsprechenden Aufgaben werden bereits bei der Generierung der SRWUs durch das Experiment durchgeführt (siehe Kapitel).

Die Konfiguration und Auswertung der Simulation erfolgt vollständig über Excel-Tabellen. Diese erlauben ein schnelles und intuitives Erstellen und Ändern von Simulations-Szenarien. Der Umgang mit diesen Tabellen wird in dem Installationsdokument der Simulationsumgebung beschrieben.

5.2.2 GridSim-Modifikationen

Aufgrund der hohen Funktionsvielfalt und Komplexität des kommerziellen Grid Resource Brokers waren grundlegende Modifikationen des GridSim Toolkits nötig. Die größten Änderungen betreffen

- Die Grid-Ressourcen und die damit verbundene Ressourcen-Definition
- Die Netzwerkressourcen

Weiterhin wurden die Schnittstellen der GridSim-Komponenten an die Anforderungen des kommerziellen Grid Resource Brokers angepasst (siehe Javadoc). Der GridSim-Grid Resource Broker wurde durch den kommerziellen Grid Resource Broker ersetzt.

5.2.2.1 Modifikation der Grid-Ressource

Der Einsatz des Distributed Broker Part befreit die Grid-Ressource von finanziellen und marktorientierten Aufgaben, wie das Festlegen von Preisen. Der Schwerpunkt liegt stattdessen auf der reinen Verarbeitung von Aufträgen, die als SRWU übergeben werden. Die Grid-Ressource bildet somit eine vereinfachte Processing Domain. Diese wurde um die Unterstützung mehrerer Ressourcen Manager (RM) erweitert. Die

Verwendung von Plugins ermöglicht die Implementierung beliebig komplexer RM-Verfahren. Jede Grid-Ressource besitzt eine Instanz des Distributed Broker Parts. Dieser kann über das Interface Resource auf die Grid-Resource zugreifen (siehe Javadoc). Übergeben werden SRWUs, die um einen Datenteil erweitert sind. Dieser enthält ein Ressourcen-Gridlet, das das interne Format auf der Grid-Ressource bildet. Für die Verarbeitung stehen 3 Ressourcen Manager zur Verfügung. Diese erhalten über ein Round Robin-Verfahren Systemzeit von der Grid-Ressource. Die Gesamtzeit der 3 Ressourcen Manager ist auf der Grid-Ressource konstant und für die Simulation auf ein Service Intervall festgelegt. Im Rahmen dieses Intervalls können die Zeitscheiben für jeden Ressourcen Manager individuell festgelegt werden. Der Distributed Broker Part besitzt dazu die Möglichkeit über das Interface Resource die Zeiten zu verändern. Auf diese Weise können z.B. dynamische Bandkapazitäten auf die Ressourcen Manager abgebildet werden. Nach der Übergabe der SRWU entnimmt die Grid-Ressource das Ressourcen-Gridlet und weist es einem RM zu. Über die WLM-Parameter der SRWU kann der kommerzielle Grid Resource Broker die Wahl des Ressourcen Managers steuern.

5.2.2.2 Modifikation der Netzwerkresource

Die Netzwerkresource der GridSim-Umgebung gestattet keine interaktive Kommunikation zwischen den Komponenten der Grid-Umgebung. Die Modifikation erlaubt das parallele Senden und Empfangen unterschiedlicher Daten und Nachrichten. Ermöglicht wird dies durch das Interface Network, das von den GridSim-Komponenten implementiert werden muss (siehe Javadoc) und die neue Kommunikationsschnittstelle bildet.

5.2.3 Experiment

5.2.4 Statistik

Kapitel 6

Globus

6.1 Einleitung

Globus ist die aktuelle Bemühung verschiedener Unternehmen und Einrichtungen (IBM, Microsoft, Cisco Systems, Argonne National Laboratory, University of Chicago, University of Southern California Information Sciences Institute, University of Edinburgh, Swedish Royal Institute of Technology, High Performance Computing Laboratory Northern Illinois University, Max Planck Institute, Imperial College, Monash University, DARPA, NSF, NASA, etc.) ein Software-Paket zu entwerfen und zu implementieren, das die Planung und den Betrieb einer Grid-Umgebung erlaubt. Die Grundlage bildet die Open Grid Service Architecture (OGSA), über die ein standardisierter Aufbau und Zugriff ermöglicht wird (siehe Kapitel 2.2.1.8 auf Seite 33 und [9, 15, 16]). Die aktuelle Globus-Version (3.x) setzt bereits eine Vielzahl der Vorschläge der OGSA um. Unter www.globus.org sind die entsprechenden Pakete frei erhältlich. Die Hauptkomponenten werden in Kapitel 6.2 auf der nächsten Seite beschrieben. Für die Entwicklung von Grid-Software, die auf Globus basiert, steht ein entsprechendes API zur Verfügung, das vollständig in C implementiert ist. Das Globus Developer's Kit (Commodity Grid-CoG) erlaubt aber auch den Einsatz folgender Technologien

- Java
- Python
- Web Services
- CORBA
- Java Server Pages
- Perl
- Matlab

Die entsprechenden ausführlichen Beschreibungen sind unter www.globus.org zu finden.

Literaturverzeichnis

- [1] AMAN, J. ; EILERT, C. K. ; EMMES, D. ; YOCOM, P. ; DILLENBERGER, D. : *Adaptive algorithms for managing a distributed data processing workload*. 1997
- [2] BEYERLE, M. : *Architekturanalyse der Java Virtual Machine unter z/OS und Linux*. Fakultät für Informations- und Kognitionswissenschaften, Universität Tübingen, 2003
- [3] BUYYA, R. ; ABRAMSON, D. ; GIDDY, J. : *Nimrod/G: An Architecture for a Resource Management and Scheduling System in Global Computational Grid*. IEEE Computer Society Press, 2000
- [4] BUYYA, R. ; ABRAMSON, D. ; GIDDY, J. ; STOCKINGER, H. : *Economic models for resource management and scheduling in Grid computing*. Grid Computing and Distributed Systems Laboratory, 2002
- [5] BUYYA, R. ; MURSHED, M. : *GridSim: A Toolkit for the Modeling and Simulation of Distributed Resource Management and Scheduling for Grid Computing*
- [6] CHELIOTIS, G. ; KENYON, C. : *Grid Economics*. 2002
- [7] FERGUSON, D. ; NIKOLAOU, C. ; YEMINI, Y. : *An Economy for Managing Replicated Data in Autonomous Decentralized Systems*. 1993
- [8] FERGUSON, D. ; YEMINI, Y. ; NIKOLAOU, C. : *Microeconomic Algorithms for Load Balancing in Distributed Computer Systems*. 1988
- [9] FERREIA, L. ; BERSTIS, V. ; ARMSTRONG, J. ; KENDZIERSKI, M. ; NEUKOETTER, A. ; TAKAGI, M. ; BING-WO, R. ; AMIR, A. ; MURAKAWA, R. ; HERNANDEZ, O. ; MAGOWAN, J. ; BIEBERSTEIN, N. : *Introduction to Grid Computing with Globus*. 2002
- [10] FOSTER, I. : *WS-Resource Framework: Globus Alliance Perspectives*
- [11] FRANZ, J. ; SPRUTH, W. G.: *The Transaction Engine: High transaction rate processing for Core Business Applications*. 2003
- [12] GROUP, L. : *Legion 1.7 Reference Manual*. 2001
- [13] HORNUNG, T. : *Entwurf und Implementierung einer transaktionalen Subwork-flowsteuerung*. 2003
- [14] HOWELL, F. ; MCNAB, R. : *simjava: A discrete event simulation package for Java*

- [15] III, J. H. ; BROWN, J. S.: *Service Grid: The Missing Link in Web Services*. 2002
- [16] KISHIMOTO, H. ; SAVVA, A. ; SNELLING, D. : *OGSA Fundamental Services: Requirements for Commercial GRID Systems*. 2002
- [17] LALIS, S. ; NIKOLAOU, C. ; PAPADAKIS, D. ; MARAZAKIS, M. : *Market-driven Service Allocation in a QoS-capable Environment*
- [18] LUTHER, A. ; BUYYA, R. ; RANJAN, R. ; VENUGOPAL, S. : *Alchemi: A .NET-based Grid Computing Framework and its Integration into Global Grids*. Grid Computing and Distributed systems Laboratory
- [19] NATRAJAN, A. ; NGUYEN-TUONG, A. ; HUMPHREY, M. A. ; HERRICK, M. ; CLARKE, B. P. ; GRIMSHAW, A. S.: *The Legion Grid Portal*. 2001
- [20] PLUM, C. : *QoS-Management mit mobilen Agenten*. 2003
- [21] ROONEY, W. ; KUBALA, J. P. ; MAERGNER, J. ; YOCOM, P. B.: *Intelligent Resorce Director*. 2002
- [22] SCHLODDER, M. : *Architektur und Implementierung eines Datenzugriffsdienstes für eine Transaktionsmaschine*. Fakultät für Informations- und Kognitionswissenschaften, Universität Tübingen, 2004
- [23] SPRUTH, W. G. ; HERRMANN, P. ; KEBSCHULL, U. : *Einführung in z/Os und OS/309*. 2003
- [24] YU, J. ; VENUGOPAL, S. ; BUYYA, R. : *A Market-Oriented Grid Directory Service for Publication and Discovery of Grid Service Providers and their Services*. Grid Computing and Distributed Systems Laboratory

Akronymverzeichnis

SU	Service Unit
PD	Processing Domain
BDAfPD	Band Distribution Algorithm for Processing Domain
SRWU	Service Request Work Unit
TSUP	Total Service Units Projected
PSE	Provider Service Element
CSE	Customer Service Element
PDP	Processing Domain Profile
PSUN	Projected Number of Service Units Needed
CSE	Customer Service Element
SLA	Service Level Agreements
QoS	Quality of Service
MFC	Masterflow Controller
WLM	Workload Manager
CBP	Central Broker Part
DBP	Distributed Broker Part
RAS	Reliability, Availability, Serviceability
SWFC	Carrier and Subworkflow Controller
CC/P	Channel Controller, Profiler
ACID	Atomicity, Consistency, Isolation, Durability
GIS	Grid Information Service
PI	Performance Index
kGRB	kommerzieller Grid Resource Broker

P2P	Peer-To-Peer
SJF	Shortest Job First
TE	Transaction Engine
OGSA	Open Grid Service Architecture
GSI	Grid Security Infrastructure
RM	Ressourcen Manager