

**ACTIVE DATA WAREHOUSING:
WHITE PAPER – THOMAS WEILER**



ACTIVE DATA WAREHOUSING

Ein modernes Informationssystem deckt inzwischen weit mehr Funktionalitäten als die Rolle des „klassischen Datenlieferanten“ ab. Lag zu Beginn der Fokus mit einfachen Reporting-Systemen einzig auf der Datenbereitstellung, sind heutige Informationssysteme zur Drehscheibe des Daten- und Informationsflusses gewachsen. Begünstigt durch parallele Entwicklungen der Technologie, neuen Prozessmodellen wie EAI und der Einbezug in operative Abläufe positionieren sich derartige Systeme mittlerweile als vollständiger Teil der Informationsinfrastruktur, im operativen wie auch im dispositiven Sinne. Klassische Data Warehouse Systeme erfüllten diese Eigenschaft noch nicht, die stetige Weiterentwicklung der Data Warehouse Technologie und Methodik hat sie in den vergangenen Monaten jedoch in ihrer Evolution eine Stufe höher geführt und eine Alternative zu allfälligen Problemen jeder gewachsenen Informationsinfrastruktur (Dezentralisierung, Realtime-Prozesse, Prozesskommunikation, Datenqualität, etc.) geschaffen. Um diesem Entwicklungssprung Rechnung zu tragen, hat sich für derartige neue Informationssysteme der Begriff „Active Data Warehousing“ (im Folgenden kurz ADW) durchgesetzt.



Was konkret unter einem ADW-System zu verstehen ist und wie sich ADW vom klassischen Warehousing unterscheidet, verdeutlicht dieses Dokument.

1 MOTIVATION

Data Warehousing (DW) ist der Sammelbegriff für Informationssysteme in den unterschiedlichsten Ausprägungen und mittlerweile ein etabliertes, unternehmerisches Instrument. Wir finden einfache Reportingsysteme auf lokaler PC-Basis ebenso wie hoch performante, mehrschichtige Client-Server Architekturen im Netzverbund auf heterogenen Plattformen mit einem vielschichtigen Anwendungsportfolio. Komplexe DW-Systeme bieten mittlerweile neben der Standardanwendung „Reporting“ Funktionalitäten in den Bereichen „Datenmanagement“, „Anwendungsintegration“, „(End)Benutzerkommunikation“, „Metadatenmanagement“ oder „Data Mining“ und sind zu einem tragenden Baustein der betrieblichen Infrastruktur gewachsen. Technisch stehen Systeme auf vielen Plattformen für jedwellige Anforderungen zur Verfügung, applikatorisch sind unzählige Komponenten für DW-Systeme (DWS) verfügbar und aus Sicht der Geschäftsprozesse bieten DW-Systeme weit mehr Potential als die Rolle des puren Datenlieferanten. Das Datum an sich ist nach wie vor die zentrale Komponente eines DWS. Es muss korrekt, aktuell, vollständig und auf die Anforderung zugeschnitten sein, damit der Geschäftsprozess effizient durchführbar ist. Ein Geschäftsprozess funktioniert in seinen Grundzügen wie jeder andere Prozess, indem Eingabeinformationen durch Verarbeitung Ausgabeinformationen erzeugen und weitere Prozesse anstoßen. Klassische DWS sind dabei auf der Datenebene ausgereift, was ihnen häufig aus Sicht der Zukunftsorientierung fehlt, ist die optimale – das heißt, für jeden zu unterstützenden Geschäftsprozess den Anforderungen entsprechende - Einbettung der Prozesse in die gesamte Geschäftsprozesslogik.

Ein Web-Shop bspw. stellt ein geeignetes Instrument zur Umsetzung automatisierter Geschäftsprozesslogik dar. Informationen über Kunden und Produkte sind während der Kundenaktivitäten identifizier- und bereits elektronisch verfügbar. Betreiben sie allerdings zwischen ihrem CRM- bzw. Bestandssystem und



dem Web-Shop eine Batch-Schnittstelle (d.h. der Datenaustausch findet periodisch über Datenextrakte statt), verlieren Sie bspw. kostbare Zeit, um unmittelbar nach der Kundenaktion reagieren zu können oder Präventivmassnahmen einzuleiten. Die fehlende Komponente im System ist hierbei das aktive Handeln bei Eintreten bestimmter Kundenereignisse. Dieses Beispiel zeigt bereits eine Kernproblematik gewachsener IT-Strukturen: Die Anwendungsintegration.

Ein weiteres Beispiel ist die Missbrauchserkennung und damit mögliche –vorbeugung im Telekommunikationsbereich. Aus den Gesprächsdaten lassen sich bestimmte Verhaltensmuster der Kunden ableiten, wozu dedizierte Anwendungen und Methoden existieren. Aufgrund des Vergleichs der aktuellen Gesprächsdaten mit dem Verhaltensmuster ist mit hoher Wahrscheinlichkeit ein Missbrauch erkennbar und es können entsprechende Maßnahmen ergriffen werden. Die Qualität eines solchen „Fraudmanagement“ ist dabei sehr stark von der Aktualität der Informationen abhängig, d.h. Informationen müssen „near realtime“ verfügbar sein.

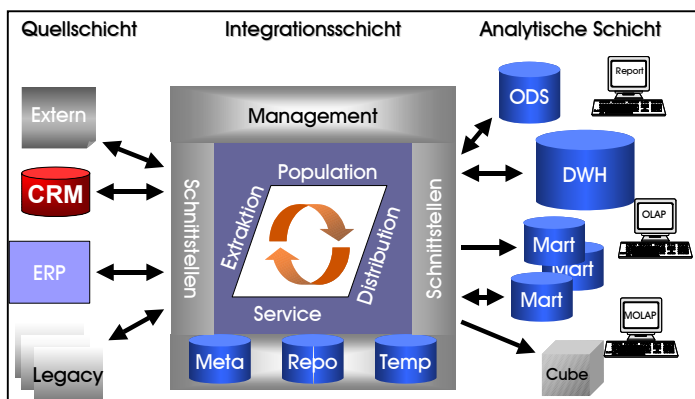
Als drittes Beispiel sei die Integration umfassender Kundendaten in Call Center Systeme genannt. Hier kann die Daten- und Prozessunterstützung nur dann ihr volles Potential ausschöpfen, wenn sie vollständige und höchst aktuelle Informationen für die Kundeninteraktion zu Verfügung stellt. Vielfach pflegen Kunden ihre Stammdaten wie Adressen, Kontenverbindungen etc., um in einem



nachgelagerten Gespräch mit dem Call Center diese Informationen erneut zu aktualisieren. Auch dieser relativ einfache Vorgang unterstreicht das Problem der Daten- und Anwendungsintegration. Hierbei wird sehr deutlich, wo die Problematik –und damit auch das Entwicklungspotential bzgl. ADW- liegt. Es reicht nicht, konsolidierte Daten aufzubereiten, sie müssen auch zur richtigen Zeit der richtigen Anwendung zur Verfügung gestellt werden, d.h. Daten- und Anwendungsintegration; oder um es mit verbreiteten Schlagwörtern zu beschreiben: „enterprise wide actionalbe data integration“.

Klassische Warehouse Systeme können zwar die Dateninhalte in den drei Beispielen bereitstellen, das Problem ist jedoch i.d.R. die Anwendungsintegration dieser aktuellen, maßgeschneiderten, zentralen und anforderungsgerecht aufbereiteten Daten, d.h. Informationen.

Das zentrale Thema hierbei - die Anwendungsintegration von Informationen - wurde in der Vergangenheit mit unterschiedlichen Ansätzen angegangen. Ältere Anwendungsintegrationen basieren auf unternehmenseigenen Datenaustauschverfahren, moderne Ansätze nutzen sog. Processware oder EAI-Tools. Allen gemeinsam ist hierbei, dass sie dem Gedanken der Dezentralisierung folgen. Informationen liegen verteilt bei ihren Master-Anwendungen und werden per Definition, bei Bedarf, bei Anforderung oder einem bestimmten Ereignis an weitere Anwendungen übermittelt. Das Ergebnis sind entweder hohe Datenredundanzen inklusive Qualitätseinbussen oder aufwendige Kommunikations-Infrastrukturen; in jedem Falle aber ein komplexes Design der Anwendungs-, Schnittstellen- und Kommunikationslogik.



Wendet man sich vom Gedanken der Dezentralisierung ab, ist die erste Idee ein zentraler Datentopf mit einem Grundgerüst an Schnittstellen, das jede potentielle Anwendung bedienen und nutzen kann (mainstream). Dies ist quasi die einfache Kurzbeschreibung eines Data Warehouses (DWH), bei dem passive Bestandsinformationen im Vordergrund stehen. Für die Anwendungsintegration sind diese Daten jedoch zeitnah und transaktionsorientiert abzubilden, was dem Charakter eines operational data store (ODS) entspricht. Und zu guter Letzt müssen diese Informationen auch noch Trigger für weitere Systeme darstellen, d.h. sie müssen Auslöser für Aktionen sein und zwar genau dann, wenn Geschäftsprozesse dies verlangen (event-getriebene und –treibende Datenhaltung). Sub-

summiert man diese Eigenschaften und Anforderungen, so beschreibt es ein System, welches man als „Active Data Warehousing“ (ADW) bezeichnet.

Hierbei bedarf der Begriff „near realtime“ besonderer Beachtung. Die klassische Interpretation des Begriffs basiert auf rein zeitlichen Überlegungen des Datentransfers und deutet ihn als Verarbeitung in Echtzeit bzw. mit geringen Latenzzeiten. Für das ADW hieße das, die Aktualität der ADW-Daten entspricht der Aktualität der Quelldaten mit Abweichungen im Sekundenbereich, begründet durch die Verarbeitungszeit der Datenüberführung in das ADW. Dieser Ansatz ist eine rein technische Sicht und für das ADW unzureichend. Im ADW ist der zentrale Ausgangspunkt stets der Geschäftsprozess, welcher die fachlichen Anforderungen an das ADW definiert. Und hierbei ist near realtime bedarfsorientiert zu sehen, d.h. unabhängig von der Kontinuität oder Periodizität der Datenüberführung ist near realtime die Bereitstellung zum richtigen Zeitpunkt, also genau dann, wenn der Prozess die Daten benötigt. Die Aktualität kann dabei von Sekunden bis Monate betragen. Eine bankfachliche Deckungsprüfung benötigt bspw. höchst aktuelle Werte, eine Monatsabschlussverarbeitung sicherlich nicht. Near realtime sollte demnach nicht als Echtzeitverarbeitung verstanden werden, sondern als Bereitstellung zum richtigen Zeitpunkt. In der Praxis hat sich ein common sense von einer Stunde Aktualität durchgesetzt, welches man als near realtime bezeichnet.

Das übergeordnete Ziel des ADW ist somit, konsolidierte und höchst aktuelle Daten den Anwendungen zur Verfügung zu stellen, die an einem Geschäftsprozess beteiligt sind, sowohl bei Anforderung als auch aus dem ADW-System selbst heraus getriggert. Etwas vereinfacht ausgedrückt, aber damit sehr gut vorstellbar ist ADW die Erweiterung klassischer DWH-Systeme um die Anwendungsintegration (die parallel im Bereich EAI ebenfalls das zentrale Thema ist) und Integration operativer- und event-getriebener Prozesse.

Zusammengefasst führen diese Herausforderungen zu einer weiteren Evolutionsstufe von Data Warehouses, dem ADW. Die Weiterentwicklung richtet sich wesentlich an folgenden Kriterien aus:

- zentrale, redundanzfreie Datenhaltung auf verschiedenen Aggregations-, Veredelungs- und Aktualitäts-Ebenen,
- realtime Datenaktualität,
- Einbindung in operative Prozesse,
- eventgetriebene Aktionssteuerung von nachgelagerten Geschäftsprozessen und
- realtime Datenbezug von Informationen im Request/Reply-Modus jederzeit möglich.

2 DEFINITION

Active Data Warehousing ist somit ein Prozess, der auf den klassischen Ansätzen des Data Warehouse Prozesses basiert, diese Architekturen jedoch in kritischen Bereichen aktueller Geschäftsprozesse erweitert und aktiv in diese Prozesse integriert wird. Es unterliegt dabei keiner starren Architektur (bspw. stets die gleichen Datenlayer), sondern ist architektonisch ähnlich einem Baukastensystem situationsgerecht zu formen. Beispiele dieses kreativen Designs finden sich im nächsten Kapitel „Architekturen“.

Für die Definition sind an dieser Stelle die möglichen Komponenten des Baukastens zu beschreiben. Um den oben genannten Evolutionsprung zu verdeutlichen, erfolgt die Komponentenbeschreibung in zwei Schritten, zunächst die Komponenten des klassischen DWH und anschließend exemplarische Änderungen/Erweiterungen für Active Data Warehousing“.



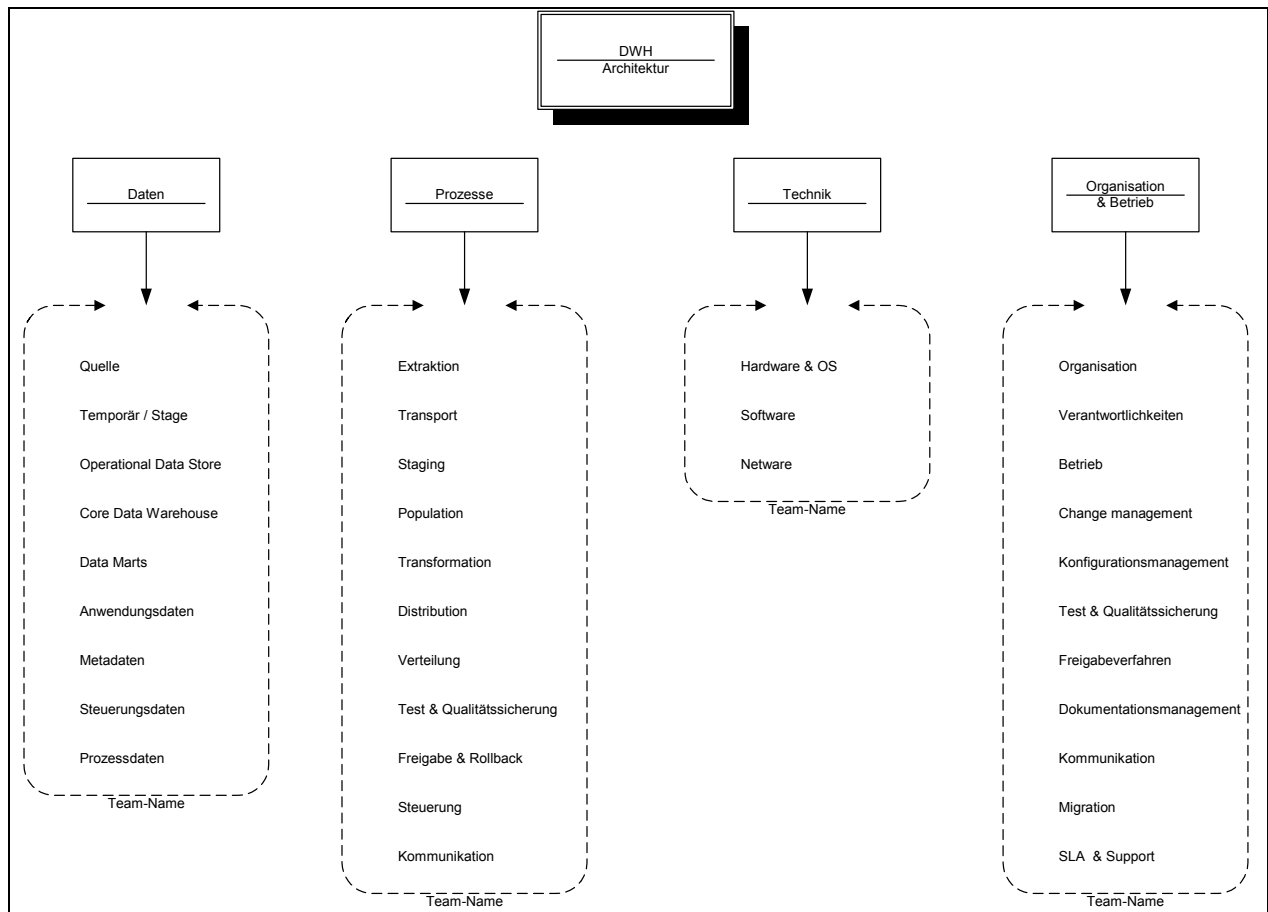
2.1 ÜBERSICHT - KOMPONENTEN DES DATA WAREHOUSING

Die Architektur eines klassischen Warehouses basiert auf vier Säulen,

- der Datenarchitektur zur Definition sämtlicher Datenschichten,
- der Prozessarchitektur, welche die Vorgänge definiert,

- der technischen Architektur, die sämtliche technischen Komponenten beschreibt und
- der Organisation / des Betriebs, welche organisatorische Aspekte adressiert und den betrieblichen Ablauf definiert.

Jede Säule adressiert dabei mehrere Teilkomponenten, wie in folgender Grafik dargestellt.



2.2 DATENARCHITEKTUR

In der Datenarchitektur geht es zentral um die Frage, welche Datenschichten aufgebaut werden müssen. Logisch wird zwischen neun Datenschichten unterschieden (Quelle, Stage, ..., Prozessdaten siehe Grafik untenstehend). Physikalisch können diese Schichten übergreifend modelliert werden.

Daten								
Quelle	Temporär Stage	Operational Data Store	Core Data Warehouse	Data Marts	Anwendungsdaten	Metadaten	Steuerungsdaten	Prozessdaten
Art (oprativ/dispositiv)	Lokalität	Modellstruktur	Modellstruktur	Aggregations-ebene	Datenart (extern?)	Art	Parametrisierung	Art
Datenqualität	Persistenz	Aktualität	Konsolidierung	Shared dimensions	Lokalität	Umfang	Automatische Fortschreibung	Umfang
Abhängigkeit zu anderen Systemen	Funktionalität	Füllungs-mechanismus	Versionierung	Conformed dimensions	Metadaten	Aktive / passive	Releasefähig	Reaktive Daten
Batchfenster	Kontrolle	Funktionalität	Historie	Rückwirkende Änderungen	Pflege	Nutzung	(De)zentral	Metadatenbestandteil
Zugriffsart		Reporting	Kundenzentrik	Modellierung	Nutzung	Schnittstellen	Schnittstellen	Lokalität
Delta-verarbeitung		Historie	Transformations-komplexität	Dimensions-fakten	Integration	(De)zentral	Aktualisierung	Struktur
Modellierung		Versionierung	Backup/Rollback	Security	Releasefähig	Fortschreibung	Lokalität	Nutzung
Redundanzen		Transformations-komplexität	Security	Historie	Backup	Aktualität	Toolgesteuert	Historie
Aktualität		Normalisierung	Datensicherheit	Tollspezifika		Automatisiert	Überwachung	
Fachlichkeit		Quality-Check	Analytik	Fachlichkeit		Technisch	Reporting	
Nachlieferungen			Operative Referenzen	Multifact Marts		Fachlich	bidirektional (aktiv / reaktiv)	
Parametrisierung			RI / Qualität	Logical/Physical		Explorable	Historie	
Releaseplanung			Redundanzen	Mart-Kombinationen		Content sensitiv		
Testvorgaben			Granularität	Datenschutz und -sicherheit		Historie		
				Zugriff				

Speziell für ADW sind hierbei neue Punkte zu beachten, die im folgenden kurz adressiert werden. Quellen müssen den Aktualisierungsanforderungen entsprechend angebunden werden. Dies bedeutet i.d.R. einen replikationsähnlichen Datenabzug von Änderungsdaten (changed data capture). Wird dieser Mechanismus nicht vom Prozess der Extraktion abgedeckt, müssen die Quellsysteme dahingehend erweitert werden.

Besondere Bedeutung erlangen operational data stores (ods), deren Hauptfunktion innerhalb des ADW die operativen Aspekte darstellen. Gerade die real-time Verarbeitung und operativen services sind nur über eine weitere Stufe zwischen den Quellsystemen und dem core data warehouse umsetzbar. Insofern erlebt der ODS eine Renaissance.

Ein ODS hält somit subjekt-orientierte, integrierte, aktuelle und detaillierte Daten vor und bildet die wichtigste Vorstufe (Quelle) für das core data warehouse. In Abhängigkeit von der Aktualisierungsfrequenz klassifiziert man eine ODS in die Klasse I, wenn Daten real time aktualisiert werden, Klasse II bei (i.d.R. stündlicher) Aktualisierung über Standardmethoden wie z.B. Replikation, Klasse III bei (täglicher) Batch-Aktualisierung. Häufig wird noch eine weitere Klasse IV definiert, in der ein ODS nicht nur von den operativen Systemen versorgt wird, sondern ebenfalls mit komplexen Berechnungsergebnissen des core warehouses oder nachgelagerten Marts versorgt wird. Hier entstehen Informationen, die ebenfalls im operativen Betrieb wertvoll sind. Als Beispiel seien kundenbezogene Daten wie der „Life-time value“ oder Kundenwertigkeit genannt. Diese Maßgrößen, häufig im Warehouse/Marts recht komplex berechnet, stellen für die CRM-Applikation und den Call Center Basisdaten dar, werden aber normalerweise nur periodisch aufbereitet.

2.3 PROZESSARCHITEKTUR

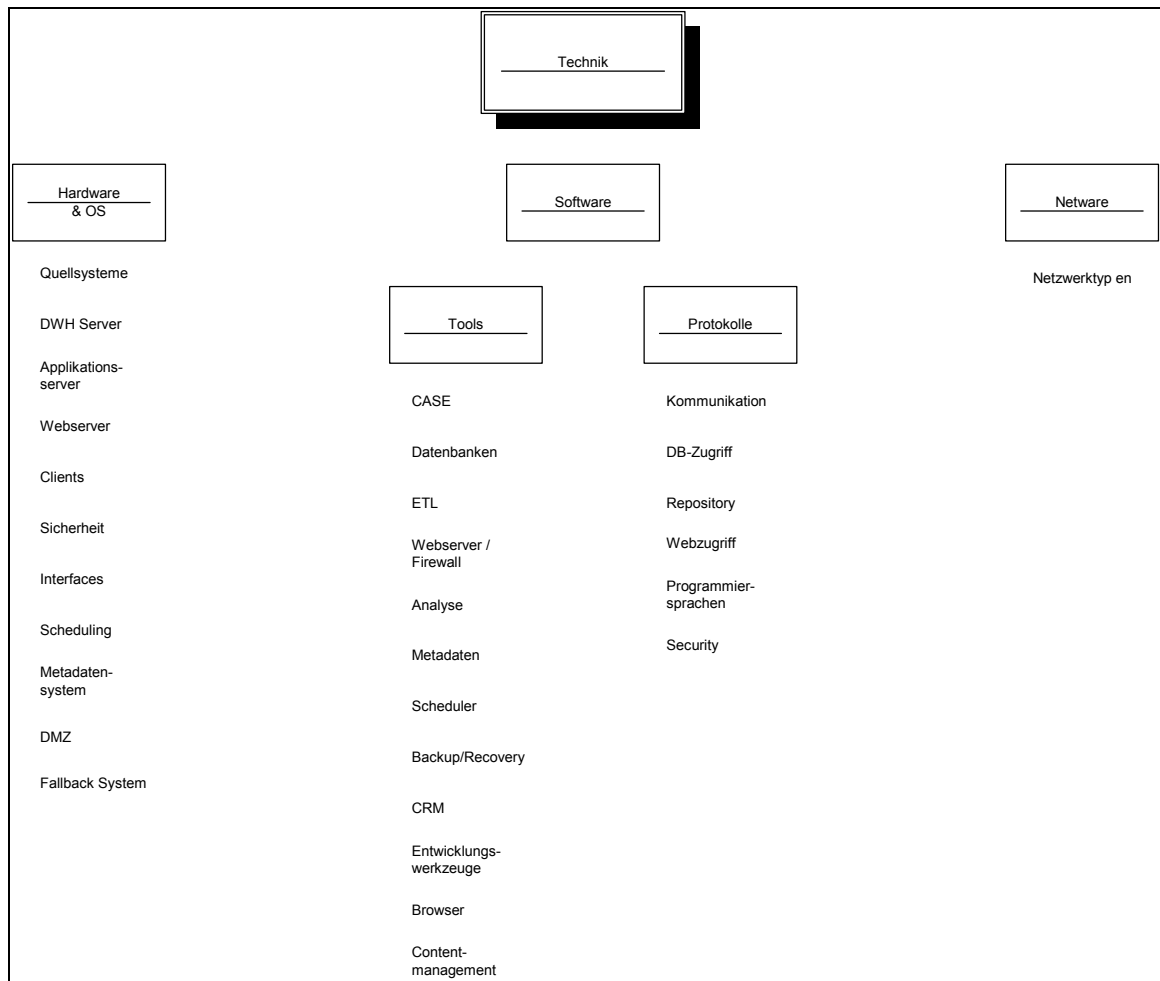
Prozesse								
Extraktion	Transport Staging	Population	Transformation	Distribution	Verteilung	Test & Qualitätssicherung	Freigabe & Rollback	Steuerung & Kommunikation
Zugriffsart	Transportweg	Transformationen	Fachlichkeiten	Logik	Mechanismus	Routinen	Prozessumsetzung	Prozessumsetzung
Delta	Kommunikation	Komplexität	Herausforderungen	Aggregation	Security	Lokalität	Integration	Parametrisierung
Parametrisiert	Transaktion	Versionierung	Kernprobleme	Security	Nutzerkonzept	QS-Massnahmen	Automatisierung	Logging
Kontrollfunktion	Persistenz	Historisierung	Kritische Prozesse	Modellumsetzung	Steuerung	Integration		Transaktionsmechanismus
Qualitychecks	RI / Checks	Key-Generierung	Performance	Load-mechanismen	Toolgesteuert	Metadaten		Restartfähigkeit
Sonderlieferungen	Laderoutinen	Anonymisierung	Qualität	Toolunterstützung	Zusatzfunktionalitäten	Steuerungsdaten		Backup/Recovery
Onlinetestmöglichkeit	Modellierung	Konsolidierung	Tool-erweiterungen			Fachliche Prüfungen		
Automatische Ansteuerung	Sicherung	Key-Mapping	Parallelisierung			Lernend		
Transaktionsfähig	Push / Pullprinzip	Rekursion						
Metadaten	Generationenunterstützung	Tooleinsatz						
	Funktionalitäten	Parametrisierbarkeit						
	Löschmechanismus	Restartmechanismen Sonderverarbeitungen (Nachlieferungen)						
		Protokollierung						
		Transaktionsmechanismen						

Analog zur Datenarchitektur stellen sich auch für die Prozessarchitektur neue Aufgaben. Wie schon erwähnt, muss der Extraktionsprozess sowohl Änderungsdatenabzüge als auch vollständige Initial Loads unterstützen. Die Datenüberführung muss zeitnah erfolgen, des weiteren dürfen nur kurze Zeitfenster belegt werden. Aggregationen sind inkrementell aufzubauen und in kurzen Perioden zu berechnen. Population und Transformationen bedürfen Qualitätsprozessen „on the fly“, Referenzierungen müssen eindeutig umgesetzt werden, Backup/Recovery muss im Betrieb möglich sein, Ausfallsicherheit und Verfügbarkeit ist bei allen Prozessen zu beachten.

Generell wachsen die Batch-orientierten Prozesse im klassischen Warehousing zu komplexen online-Prozessen im ADW. Sie müssen schneller mehr aktuelle Daten und Informationen verarbeiten. Dies sind im Übrigen Eigenschaften, welche die Softwarehersteller in den letzten Versionen ihrer Werkzeuge besonders adressiert haben.

Besonders seien hier die ETL-Tools genannt, die gerade im Bereich Datenzugriff- und Transport neue Möglichkeiten integriert haben (vielfach aus dem EAI-Bereich). Um mehr Daten schneller verarbeiten zu können, ist unbedingte Voraussetzung, das Datenvolumen zu minimieren. Dies wird im sog. Changed data capture adressiert. Der Transport ist über vielfältige Middleware-Techniken möglich, teilweise bieten die Hersteller eigene Middlewareslösungen an. Der Trend ist somit eindeutig, Funktionalitäten der EAI-Tools in Hinblick auf die Anwendungsintegration abzudecken. Aus dieser Konkurrenzsituation entsteht für die klassischen Data Warehouse Infrastrukturen der Vorteil, dass sie mit bestehenden Komponenten bereits für den Schritt zum ADW vorbereitet sind.

2.4 TECHNISCHE ARCHITEKTUR



In der technischen Architektur betreffen die Erweiterungen überwiegend Tools und Protokolle. Insbesondere neue Versionen der Datenbanken und ETL-Tools bieten vielfach eine realtime-Verarbeitung an und können damit die Nutzung spezieller Middleware wie etwa EAI-Tools überflüssig machen. Web- und Applikationsserver werden zumeist eingesetzt, um einen service-orientierten Dienst (request-reply) zu implementieren. Hinsichtlich der Analysetools tragen datenbankbasierte Werkzeuge mit Datenaufbereitung bei Anforderung der Datenaktualitätsanforderung Rechnung. Scheduler haben weiterhin große Bedeutung, da die Prozesse in Richtung event-getrieben implementiert werden. Backup/Recovery wird zu einem zentralen Thema (häufig bei DWH Systemen vernachlässigt), da die Einbindung in das operative Geschäft Hochverfügbarkeit verlangt. Security muss über alle Datenlayer berücksichtigt werden, da der Datenzugriff potentiell wahlfrei erfolgen kann.

Zentrale Themen der technischen Architektur sind für ein ADW System die Kommunikation und die Datenbank. Während die Kommunikation auf standardisierte Mechanismen basieren muss (Informationsbus) und für diese Schnittstellen zu allen Werkzeugen zur Datenüberführung bereitstellen muss, ist für die eingesetzten Datenbanksysteme ein performantes Verhalten für den mixed-workload entscheidend. Diese Anforderung resultiert aus dem parallelen Aktualisieren (schreibender Zugriff) und Lesen der Informationen in einem ADW-System. Damit sind die Datenbankmodelle nicht mehr nur einem Ziel folgend (Analyse oder normalisiert) zu entwickeln, sondern performant für beide Nutzungsarten zu entwickeln.

Besondere Bedeutung hat weiterhin die Skalierbarkeit des Systems, da ADW größere Datenmenge im Vergleich zum klassischen DWH bedeutet, sowohl für die technischen Komponenten, als auch für die Daten- und Prozessarchitektur.

2.5 ORGANISATION / BETRIEB

Organisation & Betrieb						
Organisation & Verantwortlichkeiten	Betrieb	Change management	Konfigurationsmanagement	Test-, Qualitäts- & Freigabeverfahren	Kommunikation & Dokumentation	SLA & Support
Aufbauorga	Organisation	Prozessdefinition	Existent	Verantwortlichkeiten	Kanäle	Definition umfanglich
Ablauforga	Ressourcen	Verantwortliche Personen	Releaseplanung	Automatisierung	Informationsgehalt & -umfang	Abgestimmt und akzeptiert
DWH Competence Center	Entwicklung	Akzeptanz	DWH- oder IT getrieben	Eskalationsroutinen	Anforderungsgerecht	Inhalt
Rollendefinition	Problemzonen	Reaktionszeiten	Vorgehensmodell	Akzeptanz	Zugriff SLA konform	Grenzwert oder mit Puffer
Aufgabenverteilung	Automatisierungsgrad	Kommunikation	Treibend oder getrieben	Einbindung Fachabteilung	Benutzerakzeptanz	Hotline verfügbar
Mitarbeiter-backup	SLA	Budgetierung	Strategische Planung berücksichtigend	Systemübergreifender Ansatz	Strukturkonform	Besetzung Hotline & Support
Interne/Externe Ressourcen	Support	Status	Performance	Etabliert	Metadaten-gestützt	Etabliert
Auslastung		Planung		Prüfproben-mechanismus	Toolkonformität	
Know how Transfer					Bi-direktional	
Tooleinsatz						
Unterstützungs-notwendigkeit						
Budgetierung						

Bezüglich der Organisation / des Betriebes steht im wesentlichen die Sicherstellung der SLA's mit hoher Anforderung im Vordergrund. Ausfälle sind ähnlich den operativen Systemen nicht vertretbar, Releases und Änderungen bedürfen einem definierten Management und das Monitoring des Systems hat hohe Bedeutung ebenso wie das Konfigurationsmanagement. Die teilweise unscharfen Betriebsaspekte im klassischen Warehousing müssen in einem ADW System fest definierten Prozessen weichen.

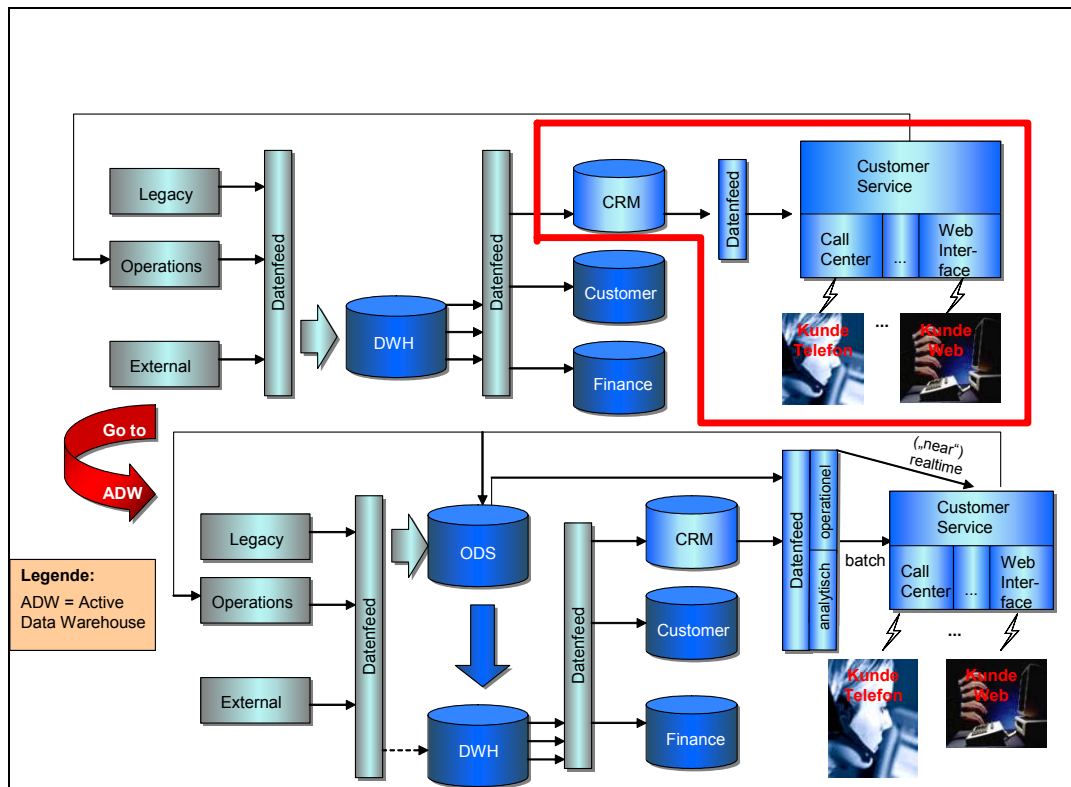
3 ARCHITEKTUREN

In diesem Kapitel werden zwei typische –in ihrer Komplexität stark unterschiedliche- ADW-Architekturen beschrieben. Das erste Beispiel zeigt eine moderate Erweiterung des klassischen Warehousing hin zu einem relativ einfachen ADW System, während das zweite ein neu entwickeltes und hoch komplexes ADW System darstellt.

3.1 ERWEITERUNG EINES KLASSISCHEN WAREHOUSE SYSTEMS IN RICHTUNG ACTIVE DATA WAREHOUSING

Das bereits erwähnte Beispiel der Call Center Anbindung an ein DWH System, bei dem konsolidierte Daten in der Kundeninteraktion genutzt werden, ist heutzutage vielfach mittels Batch-Schnittstellen realisiert. Im Sinne von ADW entspricht die Erweiterung der Unterstützung von operativen CRM-Daten mit hoher Aktualität, welche aus einem ODS abgeleitet werden.

Folgende Grafik verdeutlicht diesen Schritt:



3.2 AUFBAU EINES KOMPLEXEN ADW SYSTEMS

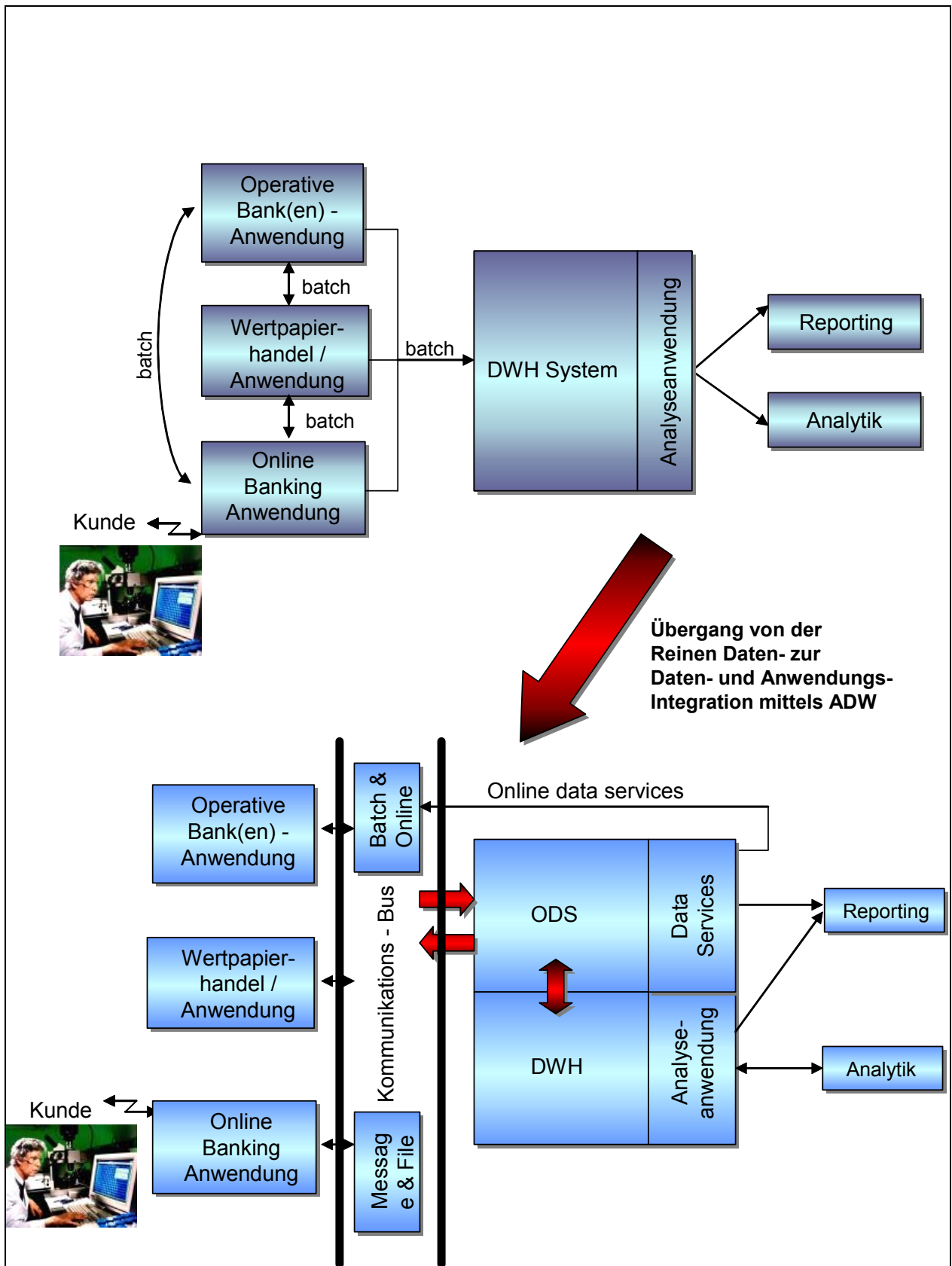
In einem komplexen ADW System geht die Anwendungsintegration noch einen entscheidenden Schritt weiter, indem der ODS (und teilweise damit auch das DWH) stärker in die operativen Geschäftsabläufe eingreift. Z.B. sind hier über sog. data services online-Abfragen operativer Anwendungen an den ODS möglich (konkrete Beispiele sind aktuelle spatenübergreifende Vertragsverhältnisse/Polices eines Kunden bei Versicherungsunternehmen, Vermögensausweise oder Deckungsprüfungen bei Bankinstituten und im aktiven Bereich Transaktionsauffälligkeiten oder ungewöhnliche Telekommunikations-Gesprächsmuster, Buchungsverhalten im Sinne von „alerting“).

Die entscheidende Stärke der ODS/DWH-Architektur innerhalb des ADW liegt darin, dass nur sie über aktuelle und konsolidierte (hiermit verfügt man nicht nur über den sog single point of truth wie es im klassischen DWH der Fall ist, sondern zusätzlich auch noch über den aktuellen) Daten verfügen.

Damit sind prädestiniert, jegliche –auch operative- Anwendungen mit zeitkritischen Datenanforderungen zu versorgen, was ganz speziell durch die e-Anwendungen größte Bedeutung gewonnen hat (seien es Web-Shops, CRM-Anwendungen, Online-Anwendungen bei Banken und Versicherungen oder Telco-Anwendungen).

Dabei zeichnen sich bestehende Systeme vor allem durch Batch-orientierten Datenaustausch zwischen den verschiedenen Anwendungen sowie einer DWH-Einbahnstraßen-Struktur aus. ADW setzt hier mit einer zentralen Konsolidierungs- und Kommunikationsplattform an, wie folgende Grafik zeigt.

Folgende Grafik zeigt einen Ausschnitt eines ADW Systems im Finanzbereich:



4 FAZIT

Active Data Warehousing stellt eine Alternative zur Informations-Konsolidierung und –Einbettung in die Informationsinfrastruktur dar. Der Evolutionsschritt vom klassischen Warehousing zum ADW ist ein Meilenstein, ADW ein hoch komplexes System. Es vereint die Informations- und Anwendungsintegration auf dispositiver und operativer Ebene und stellt damit eine Daten- und Informationsdrehscheibe für Unternehmen dar. Damit kann es als wirkungsvolles und effizientes Instrument zur Modernisierung gewachsener und ineffizienter Informations-Infrastrukturen eingesetzt werden.

SARACUS – EIN KOMPETENTER PARTNER

saracus hat sich seit mehr als zehn Jahren auf Data Warehousing konzentriert und alle Entwicklungsstufen mit vorangetrieben. Auch im aktuellen Bereich Active Data Warehousing betreut und realisiert die saracus große Projekte.

Speziell die Erfahrungen im Hinblick auf unterschiedliche Architekturen und die herstellerneutrale Architekturauswahl haben die saracus consulting zu einem wertvollen Partner gemacht.

SARACUS CONSULTING GMBH
HAFENWEG 46
D-48155 MÜNSTER
TEL. +49 (0251) 98721-0
FAX +41 (0251) 98721-26
INTERNET: WWW.SARACUS.COM
EMAIL: SARACUS@SARACUS.COM

