Data Warehouse and Business Intelligence on System z

Fritz Oosterbeek, IT-Specialist
Data Warehouse on System z
Agenda

• Market Trends: From Business Automation to Business Optimization

• Operational BI: What is it, Examples

• Why key markets trends are playing off for System z: Positioning System z strengths for Business Intelligence and Data Warehousing

• What is and will be available on System z

• Summary
New Initiatives Increasingly Focus on Optimization
Organizations Striving for Competitive Advantage

Business Optimization Growth is **2 Times Faster** than Business Automation

*Business Optimization Growth*

- **11.1% CGR**
  - $117B
  - Information Agenda

- **5.1% CGR**
  - $594B
  - Application Agenda

**Competitive Advantage**

- Dynamic Demand Planning
- Financial Risk Insight
- Supply Chain Management
- ERP & Financials
- Call Center Operations
- Customer Profitability

**Faster Processing, Reduced Costs**

New Initiatives Increasingly Focus on Optimization

Organizations Striving for Competitive Advantage

Business Optimization Growth is **2 Times Faster** than Business Automation

*Business Optimization Growth*

- **11.1% CGR**
  - $117B
  - Information Agenda

- **5.1% CGR**
  - $594B
  - Application Agenda

**Competitive Advantage**

- Dynamic Demand Planning
- Financial Risk Insight
- Supply Chain Management
- ERP & Financials
- Call Center Operations
- Customer Profitability

**Faster Processing, Reduced Costs**
Leveraging Information to Create Business Value

Insightful, Relevant Information When and Where it’s Needed

Information On Demand
- Optimize Each Transaction
- Call Centers, Field Ops

Dynamic Warehousing / Operational BI

OLAP & Data Mining
- Merchandising, Inventory, Operations

Traditional Data Warehousing

Query & Reporting
- Financials, Sales
## Examples of Dynamic Warehousing in Action

**Enabling Information On Demand for Business Advantage**

<table>
<thead>
<tr>
<th>Traditional warehousing</th>
<th>Dynamic warehousing</th>
</tr>
</thead>
</table>
| **Insurance fraud analysis and reporting** | Identifying potentially fraudulent claims prior to approval and payment  
*Transforms healthcare* |
| **Reporting on customer issues** | Identifying possible related issues, churn risk and cross-sell opportunities while engaged with the customer  
*Transforms customer service* |
| **Historical sales analysis and reporting** | Understanding relevant customer info to identify cross-sell opportunities & improve negotiating position at point of sale |
| **Crime statistics and reporting** | Identifying related incidents and potential suspects prior to arriving at the crime scene  
*Transforms crime fighting* |
Getting Information to Decision Makers: Operational BI

What is it?
“Operational BI assembles data from the business as it happens, analyzes it and makes it available to drive decisions across the business. It is the linking of the historical analysis systems with the operational business process systems.”

Why do I need it?
- **Business Optimization**
  - More value if information is in the hands of decision makers at time of decision
- **Business Agility**
  - Enables fact based decisions to be made at point of decision
- **Compliance**
  - Provides executives with immediate Information to drive the business while enabling regulatory information distribution

“Operational BI starts by asking what outcome company officials want, rather than seeing what data is lying around and then generating a report. It truly works when managers forget it exists and the technology becomes a seamless and invisible part of the business process.”

Source: Operational BI Comes of Age, TDWI
Getting the Information to Where it Provides the Most Value

Who is using business intelligence tools today?

Who can get most value out of business insights?

Typically only 18% of potential BI users actually make use of this vital information
The Evolving Landscape of Business Intelligence

BI Capability

1985

Transactions (Requests)

1990

Occasional

1995

Hundreds

2005

Thousands

2008

Millions

“C” Level

Board Room

Executive

Managerial

 Analyst Level (Analytics)

Customer Facing Personnel
(e.g., Service Center)

Customers

KPI applications

Customer Service

Self Service

Users

<50

<500

<1,000

K’s

M’s
Aggregation Level and Decision Making

**Data Aggregation Level**

- Massive Data Aggregation

- Light Data Aggregation

**“C” Level**
- Executive
- Managerial
- Analyst Level
- Sales and Customer Service
  - (customer, account, transaction)

**Decision Making**
- Strategic
- Strategic but Focused
- Strategic but More Narrow
- Customer based
- Transactional
# The Three Levels of Business Intelligence

(Imhoff)

<table>
<thead>
<tr>
<th></th>
<th>Strategic BI</th>
<th>Tactical BI</th>
<th>Operational BI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Business focus</strong></td>
<td>Achieve long-term business goals</td>
<td>Manage tactical initiatives to achieve strategic goals</td>
<td>Manage and optimize daily business operations</td>
</tr>
<tr>
<td><strong>Primary users</strong></td>
<td>Executives &amp; business analysts</td>
<td>Executives, analysts &amp; LOB managers</td>
<td>Analysts, LOB managers and users, and operational processes</td>
</tr>
<tr>
<td><strong>Time-frame</strong></td>
<td>Months to years</td>
<td>Days to weeks to months</td>
<td>Intra-day</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>Historical data</td>
<td>Historical data</td>
<td>Real-time, low-latency &amp; historical data</td>
</tr>
</tbody>
</table>

*System z Sweet Spot*
Where you put your Data Matters…
Confidence in System z, z/OS and DB2 for z/OS

• 25 of the top 25 WW banks (1)
• 23 of the top 25 US retailers (2)
• 9 of the top 10 global life/health insurance providers (3)
• The world’s largest OLTP System is Land Registry for England and Wales. 23.1 TB system (4)
• UPS' shipping system achieved a peak workload of 1.1 Billion SQL statements per hour

(2) US Retailers from National Retail Federation July 2005: www.stores.org/pdf/TOP100printwithad.pdf
**Why** key markets trends are playing off for System z: Positioning System z strengths for Business Intelligence and Data Serving
Mission Critical Workloads Require Highest QoS

More than 90% of Global 2000 companies plan to incorporate analytics into multiple operational applications that access the data warehouse by 2010, but fewer than 15% of data warehouses have been designed to provide high availability, failover, disaster recovery and the remaining components of mission-critical systems.

By the end of 2009, 90% of Global 2000 companies will have implemented some type of mission-critical dependency between the warehouse and at least one revenue supporting or cost-controlling operational application — up from less than 25% in 2007.

Fewer than 15% of data warehouses in 2007 have been designed to provide high availability, failover, disaster recovery and the remaining components of mission-critical systems.

Business Impact:
The majority of companies using data from a data warehouse for in-line, operational analytics have reported that a data warehouse production failure caused operational systems to cease daily operations, resulting in lost revenue or increased costs.

1 Operational Analytics and the Emerging Mission-Critical Data Warehouse, 14 May 2007
Key questions for your DW Platform Design

1. Which DBMS to use for Datawarehousing
2. Where to place your Datawarehouse
3. How to solve the biggest challenge for dynamic DW: the mixed workload performance challenge
4. Plan for growth, the requirement for Scalability
5. TCO for your DW
1. Which DBMS to use for Datawarehousing

“The transactional DBMSs have an edge that challenges the DW DBMSs (such as Teradata)”

Gartner Data Warehouse Magic Quadrant, 2006
Which DBMS to use for Datawarehousing

“The transactional DBMSs have an edge that challenges the DW DBMSs (such as Teradata)”

Gartner Data Warehouse Magic Quadrant, 2006
2. Where to place your Datawarehouse

The larger the amount of data, the more likely it is to deploy operational BI centrally.

A centralized approach can incorporate Information management to consolidate
- data integration
- data transformation
- metadata management
rather than have these information processes handled differently by each department or line of business.

Simplification
- IBM Mainframes already host most company’s “true” real-time data
- Distributed alternatives requires additional systems, complex and costly data movement
- Increased requirements for compliance across end-to-end data integration and analytic components
2. Where to place your Datawarehouse: DW on z Solution Architecture

Within a data sharing environment, the data warehouse resides in the same group as the transactional data.
3. How to solve the biggest challenge for dynamic DW: The mixed workload performance challenge

The emergence of issues based on a mix of 4 DW workloads

- Continuous (near-real-time) data loading – similar to an OLTP workload
- Large numbers of standard reports
- An increasing number of true ad hoc query users
- An increasing level of analytics and BI-oriented functionality in OLTP

The **mixed workload performance** will become the single most important performance issue in DW

Gartner Data Warehouse Magic Quadrant, 2006
Traditional workload management approach

- Screen queries before they start execution
- Time consuming for DBAs.
- Some large queries slip through the crack.
- Running these queries degrade system performance.
- Cancellation of the queries wastes CPU cycles.
Workload Management

The ideal workload manager policy for data warehousing:

**Consistent favoring of shorter running work**........
through WLM period aging

**with select favoring of critical business users**
through WLM explicit prioritization of critical users
Workload Management

<table>
<thead>
<tr>
<th>Period</th>
<th>Duration</th>
<th>Velocity</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5000</td>
<td>80</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>50000</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>100000</td>
<td>40</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1000000</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>discretionary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Period aging can be used to lower the priorities of a query execution after it exceeds a specific amount of CPU time. This way, short running queries run with higher priority while long running queries are not blocking important resources.
Workload Management assigned to Service classes

Some queries (the “CEO” queries) still need to return in a minimum amount of time. Their priority can override the period aging through their service class assignments.

1 hour to complete

- Service class A: velocity 50, importance 3

5 minutes to complete

- Service class B: velocity 100, importance 1
4. Plan for growth, the requirement for **Scalability**: Scalability with DB2: **Single Query Parallelism**

- CPU Intensive Queries

**Time in seconds**

- 10 CPU SMP: 1874s
- 20 CPU Sysplex: 1020s
- 20 CPU: Perfect Scalability 937s

92% measured scalability
4. Plan for growth, the requirement for **Scalability:**

Scalability with DB2: **Multiuser Throughput**

- **Query Throughput**
  - 293 completions @ 3 Hours (30 users)
  - 551 completions @ 3 Hours (60 users)

- **Scalability**
  - 94%
5. **TCO** for your DW

When system z is the lowest cost platform

... just some pricing aspects for DWing
Mainframe Cost Per Unit of Work Goes Down as Workload Increases

Most TCO benchmarks compare single applications

Most businesses operate here

Cost per unit of work

Mainframe

Distributed scale out

Data Center Workload
Mainframe Cost Per Unit of Work Goes Down as Workload Increases

Where is the Cross Over Point?

It depends on your environment …

Cost per unit of work

Most TCO benchmarks compare single applications

Mainframe

Most businesses operate here

Distributed scale out

Typical large business runs 2000-4000 applications

Data Center Workload

~ 100 distributed servers according to Robert Francis Group

© 2009 IBM Corporation
Mainframe Cost Per Unit of Work Goes Down as Workload Increases

Data Center Staffing Levels for System z Have Not Increased Despite Large Increase in MIPS

\[
\begin{align*}
\text{2001} & : 5,994 \text{ MIPS} & \text{2004} & : 13,272 \text{ MIPS} \\
\text{289 Ops Staff} & & \text{260 Ops Staff} \\
\end{align*}
\]

\[
\frac{5,994}{289} = 21 \text{ MIPS/HC} \quad \frac{13,272}{260} = 51 \text{ MIPS/HC}
\]

Source: Gartner
Worldwide IT spending trend is playing off for system z

Power and cooling spend may eventually exceed new server spending

2000 – Raw processing “horsepower” is the primary goal, while the infrastructure to support it is assumed ready

2006 – Raw processing “horsepower” is a given, but the infrastructure to support deployment is a limiting factor

Management/Administration costs are already exceeding new server spending

Source: IDC, Virtualization 2.0: The Next Phase in Customer Adoption, Doc #204904, Dec 2006
Building on a strong track record of technology innovation with specialty engines, IBM is introducing the System z9 Integrated Information Processor.

Centralized data sharing across mainframes

- Internal Coupling Facility (ICF) 1997
- Integrated Facility for Linux (IFL) 2001
- System z9 Application Assist Processor (zAAP) 2004
- IBM System z9 Integrated Information Processor (IBM zIIP) 2006

Support for new workloads and open standards
- Support for new workloads and open standards
- Designed to help improve resource optimization for eligible data workloads within the enterprise
- Incorporation of JAVA into existing mainframe solutions
DB2 V8 (and later) exploitation of IBM zIIP value add

- Portions of the following DB2 for z/OS V8 workloads may benefit from zIIP*:
  1 - ERP, CRM, Business Intelligence or other enterprise applications
    - Via DRDA over a TCP/IP connection (enclave SRBs, not stored procedures or UDFs)
  2 - Data warehousing applications*
    - Requests that utilize parallel queries
  3 - DB2 for z/OS V8 utilities LOAD, REORG & REBUILD*
    - DB2 utility functions used to maintain index maintenance structures
Measured zIIP offload for a DW workload

TPC-H like queries

% offload to zIIP
... last not least
Hardware-assisted data compression

Compression Ratios Achieved

Effects of Compression on Elapsed Time

In V9 Indexes can also be compressed
TCA vs. TCO for the Mainframe

• Do not compare Total Cost of Acquisition with distributed systems

• Always look at Total Cost of Ownership
  • Cost of Hardware
  • Cost of Software
  • Environmentals
  • Cost of Labor
  • Financial terms

• Don’t be shy to promote Mainframe strengths
Summary of why DW on z

DWH z Solution

zLinux

z/OS

z/OS

Data Warehouse on z
Key questions for your DW Platform Design - Summary

1. Which DBMS to use for Datawarehousing
2. Where to place your Datawarehouse
3. How to solve the biggest challenge for dynamic DW: the mixed workload performance challenge
4. Plan for growth, the requirement for Scalability
5. TCO for your DW
Top 10 Reasons for DW and BI on System z

10. DB2 9 for z/OS
9. Consolidation through Virtual Partitioning
8. Mixed Workload Support
7. Disaster Recovery
6. Where the data resides - Collocation
5. Environment Savings - power, cooling, foot print
4. Security
3. Availability
2. Administration Costs
1. Total Cost of Ownership
What is available on System z & How to implement
System z and Information On Demand
Unlocking the Business Value of Information for Competitive Advantage

Better Business Outcomes

Business Optimization

End-to-end Capabilities

Optimization

Automation

Customer & Product Profitability
Financial Risk Insight
Workforce Optimization
Dynamic Supply Chain
Multi-channel Marketing

Flexible Architecture for Leveraging Existing Investments

Business Intelligence & Performance Management

Information Integration, Warehousing & Management

Enterprise Content Management, Enterprise Data Management

Data Warehousing on System z
Information Server for System z
IMS Integration with Info2.0
InfoSphere MDM Server for System z

Cognos 8 BI for System z

DB2 9 for z/OS
IMS 10
Content Manager v8.4
Content Manager OnDemand v8.4

Other Information & Application Sources
Data Warehouse & BI Architecture for System z

Core Offering for Enterprise Data Warehouse and BI:

- InfoSphere Information Server for Linux on System z
  - Complete, Consistent, Cost Effective
  - Same products, functionality, and leverages IFLs and Hipersockets
- DB2 Value Unit Edition
  - An alternate value point for new DB2 for z/OS workloads
- Cognos 8 BI for Linux on System z
  - A comprehensive System z offering for Enterprise BI

Watch for 2009 Additions!
IBM Information Server for System z

- COMPLETE All products
- CONSISTENT Same functionality in the Linux for System z products
- COST EFFECTIVE Can leverage lower-cost IFL MIPS with native z/OS data

IBM Information Server for System z

**Understand**
- Information Analyzer for Linux on System z
- Business Glossary for Linux on System z
- Data Architect

**Cleanse**
- QualityStage for Linux on System z

**Transform**
- DataStage for Linux on System z
- DataStage for z/OS
- DataStage MVS

**Deliver**
- Federation Server & Classic Federation
- Data Event Publisher & Classic Data EP
- Replication Server & Classic Replication

Platform Services

- Parallel Processing Services
- Connectivity Services
- Metadata Services
- Administration Services
- Deployment Services

IBM Information Server for System z
IBM Cognos 8 BI for Linux on System z

- Adhoc query, reporting and analysis (Query Studio, Report Studio & Analysis Studio)
- Dashboards and charting (Cognos Connection & Report Viewer)
- Event management (Event Studio)
- Integration with Microsoft Office (Go! Office and CAFE)
- Cube building (Transformer)
- Go! Mobile
- Go! Dashboards

Introduction to IBM Cognos 8 BI for Linux on System z
“10 TB Study”
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101437
InfoSphere Warehouse on System z

- Lower cost way to design, populate and optimize a DB2 for z/OS data warehouse
- Leverages Linux on System z to target DB2 for z/OS
InfoSphere Warehouse on System z

Client Layer
- Design and admin client
- BI / Reporting tools and Apps

Application Server

Data Warehouse Server

Source Systems

Design Studio
- Eclipse
- IE/Firefox

Admin Console
- Windows / Linux

WebSphere App Server

Cubing Services Engine
- SQW Runtime

Cognos 8 BI for System z

Third Parties / BPs
- Excel

IBM MQ Advise

Cube Metadata

DB2 for z/OS

IMS

VSAM

RDBMS

JDBC/DB2 Connect

MDX

SQL

Windows / Linux

Linux on System z Partition / IFL
Why Linux on System z for the solution?

- **Unify the infrastructure – Get it all “in the box”**
  - Manageability and Environmental benefits

- **Significant cost savings**
  - MIPs charged at IFL rate … NOT z/OS rate
    - All processing is on Linux for System z, except the z/OS data access
    - Minimizes impact on other z/OS software costs
    - DB2 access qualifies for zIIP specialty engine

- **Keep your data access and information integration processes close to your data**
  - Eliminate "wire" connectivity – data can flow over hipersockets
  - Simplify, less parts
Information On Demand
Acquisitions and Continued Innovation Drive Increased Value

Delivering End-to-End Capabilities
Data Management
Enterprise Content Management
Information Integration, Warehousing & Management
Business Intelligence & Performance Management
Information On Demand
Investment, Innovation and Deliveries Drive New Capabilities for System z

Unlocking the Business Value of Information for Competitive Advantage

Cognos 8 BI for System z
Data Studio pureQuery for z/OS
InfoSphere MDM Server for System z
Optim/Data Governance for System z
Information Server for System z

2008…
and more…

Ascential, & DWL Acquisitions 2005
FileNet Acquisition 2006
DB2 9 for z/OS 2007
IMS 10
Content Manager 8.4
Princeton Acquisition

2007

Business Optimization
Unlocking the Business Value of Information
for Competitive Advantage
Data Warehouse on System z – Trends and Directions

- Business Intelligence Performance Mgmt Reporting
- Analysis & Discovery
- Design & Management
- Data Retention
- Information Integration
- Master Data Mgmt.
- Extreme Performance
- DataServer

- Alphablox
- DataQuant
- Cognos on z/Linux
- Design Studio
- Data Archiving
- Federation, Event Publishing
- Information Server
- Embedded Data Movement
- MDM Server
- OLAP
- InfoSphere Warehouse on System z

- Share everything
- Unlimited scalability
- Workload Management
- Data Compression
- DB2 9

- Reporting
- ETL
- BI
- MDM
- Olap

© 2009 IBM Corporation
Recent papers

- Data Warehousing and Business Intelligence for IBM System z

- Operational business intelligence: The renaissance of the mainframe
  [http://www.ibm.com/software/data/businessintelligence/systemz/]

- Data Warehousing with the DB2 Family for System z Customers
  Leveraging Platform Strengths
Some key Redbooks

• Enterprise Data Warehousing with DB2 9 for z/OS
  • http://www.redbooks.ibm.com/abstracts/sg247637.html

• 50 TB Data Warehouse Benchmark on IBM System z
  • This is the draft

• DB2 for z/OS: Data Sharing in a Nutshell
  • http://www.redbooks.ibm.com/abstracts/sg247322.html

• System Programmer’s Guide To: Workload Manager
  • http://www.redbooks.ibm.com/abstracts/sg246472.html

• Workload Management for DB2 Data Warehouse, REDP-3927
  • http://www.redbooks.ibm.com/abstracts/redp3927.html
Other recent articles in the press

- Enhanced Query Parallelism with zIIP processors
  - February 2008

- Operational BI and System z
  - March 2008

- Business Intelligence’s New Look: IBM extends its BI portfolio with Cognos 8 BI for Linux on System z
  - July / August 08

- Take the Reins - An Information On Demand Strategy helps deliver a competitive edge for today’s businesses
  - July / August 08

- Three part series: Myths of Doing BI on the mainframe

- Data Warehousing With DB2 for z/OS … Again!!!
  - June/July 2008
Contact

- Feel free to call us any time!
- Via your Sales contact, Technical Field Sales,
- or call me directly:

  Fritz Oosterbeek  
  IT-Specialist  
  Information Management  
  Data Warehouse on System z

  Gustav-Heinemann-Ufer 120-122  
  D-50968 Köln  
  Germany  
  phone: +49-221-304-2017  
  mobile: +49-151-1429 3714  
  email: oosterf@de.ibm.com
Thank You