VLDB Challenges in Very Large Enterprises
Panelists

Chair
• Dr. Michael L. Brodie, Chief Scientist, Verizon

Problem Owners
• Dr. Hans-Peter Steiart, Research & Technology, DaimlerChrysler AG

Solution Owners
• Adam Bosworth, VP, Engineering, BEA Systems Inc
• James Hamilton, Architect, Microsoft SQL Server, Microsoft
• Pat Selinger, IBM Fellow and VP, Data Management Architecture and Technology, IBM
## Very Large Enterprises

**Problem Drivers**
- Data Growth
- Data Life Cycle

<table>
<thead>
<tr>
<th>Very Large Enterprise</th>
<th>Global 500</th>
<th>Fortune 500</th>
<th>Revenues ($ B US)</th>
<th>Employees (1,000s)</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daimler Chrysler</td>
<td>7</td>
<td>N/A</td>
<td>$ 136</td>
<td>373</td>
<td></td>
</tr>
<tr>
<td>Verizon</td>
<td>26</td>
<td>11</td>
<td>$ 67</td>
<td>248</td>
<td>Petabyte</td>
</tr>
</tbody>
</table>
OLTP Workload Growth

OLTP/sec Triples 2001-2004

- Dec-04
- Dec-02
- Jan-01

Projected average workload

OLTP Doubles by 12/04

- Four Years 1/01 - 12/04
- Two Years 1/01 - 12/02

Projected Workload Growth Rate (average)
DSS Workload Growth

Inflight Queries Double by 04

Projected workload (average)
(concurrent inflight queries)

DSS Workload Triples by 04

Projected workload growth rate
(average)
Database Growth

Size: OLTP Doubles; DSS Triples by 04

On 12 2004

On 12 2002

On Jan 2001

Growth: DSS & OLTP Double by 04

Jan 01 - Dec 04

Jan 01 - Dec 02

Respondents' Projected Database Size (average) (TB)

Respondents' Projected Database Growth Rate (average)
VLE Storage Growth

VLE X Storage Triples 02-06

Calendar Year

Terabytes

2001 2002 2003 2004 2005 2006

0 500 1000 1500 2000 2500 3000 3500

IP SAN/NAS
OS Vertical
OS Shared
FC SAN
System390
Data Life Cycle

Life Cycle Actions

• Create
• Store
• Replicate
• Protect
• Update
• Archive
• Exchange, exchange, exchange, …

Factors

• History: 40+ years of Mergers & Acquisitions
• Growth: Automation & Partnering
• Protection: security, confidentiality, …
• The Grand Challenge: semantics of data
Grand Challenge: $1 Trillion/year

Integration Cost Estimates

- 13% of IT spend: $752 B / year US (Giga estimate; May 2002)
- 25-40% of all IT projects (various)
- 6% of US IT spending: $610B / year US (IDC, May 2002)
- 7% of IT spending: $1.3T / year worldwide (IDC, May 2002)
- 28+% of worldwide consulting: $ 160 B/year (Gartner, March 2002)
- 43% of e-business worldwide consulting: $53 B / year (Gartner)
- 1.75% to annual IT budget on EAI and B2Bi (Forrester, Dec 2001)
- 10-30% of IT budgets (David Sink, IBM, InformationWeek, May 27, 2002)

Data Quality Cost Estimates

- $600 B / year US (Data Warehouse Institute, 2002)
VLE VLDB Challenges

Data Management

- **Global Data Management**
  - Significant improvement in dealing automatically with semantics

- **Database Engineering**
  - Automated DBA

- **Comprehensive Data Management Architecture**
  - Data architecture: Web Services, mid-tier, distributed data

Storage Management

- **Data Protection**
  - Integrated products: DBMS, replication, …

- **Data Utilization**
  - Automated DBA, Storage Virtualization, Hierarchical Storage Management for distributed system