SAP HANA:
In-Memory Data Management for Enterprise Applications

Dr. Alexander Zeier
Massachusetts Institute of Technology (MIT)
Visiting Professor

March 23rd 2012
SAP Academic Conference Americas, San Antonio
Transactional (OLTP) and analytical (OLAP) data processing has to be on one system again.

Enterprise applications have to reflect latest developments in:
- Hardware, such as:
  - Multi-core processors
  - Huge Main Memory
- Data management, such as:
  - Column-oriented storage
  - Light-weight compression
In-Memory Technology Enables Combining OLTP and OLAP in Real-Time

- Data-centric architecture: In-Memory database serves as **single source** of truth for ERP data
- Architecture based on 4 distinct pillars
  - Multi-Core computing
  - In-Memory
  - Column and Row Store
  - Insert-Only
- Enables informed management decisions based on up-to-the-moment data through real-time combination of
  - Transactional applications
  - Analytical applications
**In-Memory Data Management**

**Advances in Hardware**

- Multi-Core Architecture (8 x 10core CPU per blade)
- Parallel scaling across blades
- One blade ~$50,000 = 1 Enterprise Class Server
- 64bit address space – 2TB in current servers
- 100GB/s data throughput
- Dramatic decline in price/performance

**Advances in Software**

- Row and Column Store
- Compression
- Partitioning
- No Aggregate Tables
- Insert Only
- On-the-fly extensibility
Combined column and row store

Insert only for time travel

Active/passive data store

Dynamic multi-threading within nodes

No aggregate tables

On-the-fly extensibility

Map reduce

Minimal projections

Bulk load

Partitioning

Analytics on historical data

Single and multi-tenancy

Object to relational mapping

Group Key

Any attribute as index

Multi-core/parallelization

Lightweight Compression

SQL interface on columns & rows

Reduction of layers

Text Retrieval and EXploration

No disk

In-Memory/HANA Enterprise Data Management | SAP UA Conference | March 23rd 2012 | Dr. Alexander Zeier, MIT
Two Different Principles of Physical Data Storage: Row vs. Column Store

Row Store

<table>
<thead>
<tr>
<th>Document Number</th>
<th>Document Date</th>
<th>Sold-To Party</th>
<th>Order Value</th>
<th>Status</th>
<th>Sales Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>95769214</td>
<td>2009-10-01</td>
<td>584</td>
<td>10.24</td>
<td>CLOSED</td>
<td>Germany Frankfurt</td>
</tr>
<tr>
<td>95769215</td>
<td>2009-10-01</td>
<td>1215</td>
<td>124.35</td>
<td>CLOSED</td>
<td>Germany Berlin</td>
</tr>
<tr>
<td>95779216</td>
<td>2009-10-21</td>
<td>584</td>
<td>47.11</td>
<td>OPEN</td>
<td>Germany Berlin</td>
</tr>
<tr>
<td>95779217</td>
<td>2009-10-21</td>
<td>454</td>
<td>21.20</td>
<td>OPEN</td>
<td>Germany Frankfurt</td>
</tr>
</tbody>
</table>

Column Store

<table>
<thead>
<tr>
<th>Doc Num</th>
<th>Doc Date</th>
<th>Sold-To</th>
<th>Value</th>
<th>Status</th>
<th>Sales Org</th>
</tr>
</thead>
</table>

Row

Row 1
Row 2
Row 3
Row 4
SELECT * 
FROM Sales Orders 
WHERE Document Number = '95779216'  
(OLTP-style query)

SELECT SUM (Order Value) 
FROM Sales Orders 
WHERE Document Date > 2009-01-20  
(OLAP-style query)
Dictionary Compression

- Reduces I/O operations to main memory (bottleneck)
- Operations directly on compressed data

Typical compression factor for enterprise software 10
In financial applications up to 50
Table Characteristics

Row Store
- Small tables
- Frequent updates
- Materialized aggregates

Column Store
- Large tables
- Rare updates
- Dynamic aggregates

Text
- Crawler
- Join structured & unstructured data

Transactional Data
- Direct access to tuples
- Blade-local transactions
- Status updates
- Active / passive

Historical Data
- Sequential access
- No updates
Innovative In-Memory / HANA Applications
Nowadays Financials

Base Tables

Accounting Document Header

Accounting Document Items

Change History

Materialized Aggregates

General Ledger

Accounts Payable

Accounts Receivable

Material Ledger

Sales Ledger

Tax Ledger

Fixed Asset

Cash Ledger

Materialized Views

General Ledger Items

Accounts Payable Items

Accounts Receivable Items

Material Ledger Items

Dunning

Materialized Views

Sales Ledger Items

Tax Ledger Items

Fixed Asset Items

Cash Ledger Items

Payments

Reporting Cubes

Indices
Simplified Financials (Target)

Only base tables, algorithms, and some indices
Customer Study 1: Dunning Run in < 1s?

- Dunning run determines all open and due invoices
- Customer defined queries on 250M records
- Current system: 20 min
- New logic: 1.5 sec
  - In-memory column store
  - Parallelized stored procedures
  - Simplified Financials
Dunning Application
Tracing pharmaceutical packages in Europe

15 bn packages / 35 bn read events per year

Prototype with 12 billions records with response time: 23 ms
Medical doctors have all patient data at hand to apply personalized medicine.

Medical researchers perform real-time analysis to define cohorts for clinical studies.

International research initiative for exchanging relevant tumor data started at World Health Summit 2011 in Berlin.

In-Memory Technology as key-enabler for real-time analysis.

Provider for information at your fingertips (iPad).
HANA Oncolyzer - combining Structured and Unstructured Data

<table>
<thead>
<tr>
<th>Events</th>
<th>Description</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery</td>
<td>04 Mai 1992</td>
<td>Last Contact 04 Mai 1992</td>
</tr>
<tr>
<td>Surgery</td>
<td>04 Mai 1992</td>
<td>Tumor Death Tod tumorbedingt, keine nähere Angabe</td>
</tr>
<tr>
<td>Time Skip</td>
<td></td>
<td>Used Medicine Vincristin</td>
</tr>
<tr>
<td>System Therapy</td>
<td>System Therapy</td>
<td></td>
</tr>
<tr>
<td>01 Februar 1994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td>01 Februar 1994</td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td>01 Juni 1994</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>15 Februar 1994</td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td>01 Juni 1994</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>61 years</td>
<td></td>
</tr>
<tr>
<td>Start Date</td>
<td>01 Februar 1994</td>
<td></td>
</tr>
<tr>
<td>End Date</td>
<td>01 März 1994</td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>29 days</td>
<td></td>
</tr>
<tr>
<td>Intention</td>
<td>Palliativ bzw. symptomatisch</td>
<td></td>
</tr>
<tr>
<td>Ratio</td>
<td>♂ 59%, ♂ 41%</td>
<td></td>
</tr>
</tbody>
</table>
HANA Oncolyzer was presented on CeBIT 2012 to Germany’s Chancellor Angela Merkel as SAP’s Innovation 2012
First Results of Customers using SAP HANA

- **1,000x** Faster: Many (Dunning, Aging, ...)
- **10,000x** Faster: NongFu Spring, Essar Group, SAP IT, Cornell, Charmer Sunbelt
- **100,000x** Faster: YodoBashi, MKI

**OR**

- 24+ Hours to **3.8S**: Food and Beverage / Distribution - Logistics
- 15+ Hours To **4.8S**: Project Management / Services, Profitability, Performance
- 30 Days to **28S**: Manufacturing - Order to Cash
- 3 Days to **2s**: Retail / Insurance - Incentives
All Findings are Summarized in the Book “In-Memory Data Management”

This book is the culmination of five years worth of in-memory research

- **PART I – An Inflection Point for Enterprise Applications**
  - Overview of our vision of how in-memory technology will change enterprise applications

- **PART II – A Single Source of Truth through In-Memory**
  - Technical foundations of in-memory data management
  - In-depth description of how we intend to realize our vision

- **PART III – How In-Memory Changes the Game**
  - Resulting implications on the development and capabilities of enterprise applications

-> Book launched at Cebit 2011, SAP Product HANA is available since June 2011.

-> **New extended Book** Edition “In-Memory Data Management - Technology and Applications” focusing on Application Development will be available for Sapphire May 2012.
In-Memory/HANA Drives Worldwide Innovation

Book Launch at CeBIT 2011 with Vice-President of the European Commission Neelie Kroes
SAP and HPI win the German Innovation Award 2012 for SAP HANA!

This year’s winners were announced am March 16, 2012 in Munich, Germany.

Please feel free to contact me:

Dr. Alexander Zeier
Massachusetts Institute of Technology (MIT)
Visiting Professor
Executive Director MIT Forum for SC Innovation
Email: zeier@mit.edu
Website with list of over 150 Publications: http://zeier.mit.edu