EIM 119

High-Performance Analytic Appliance

The Road to Analyzing Business as it Happens

Brian Wood, Product Strategist

October 2010

SAP TECHED 10

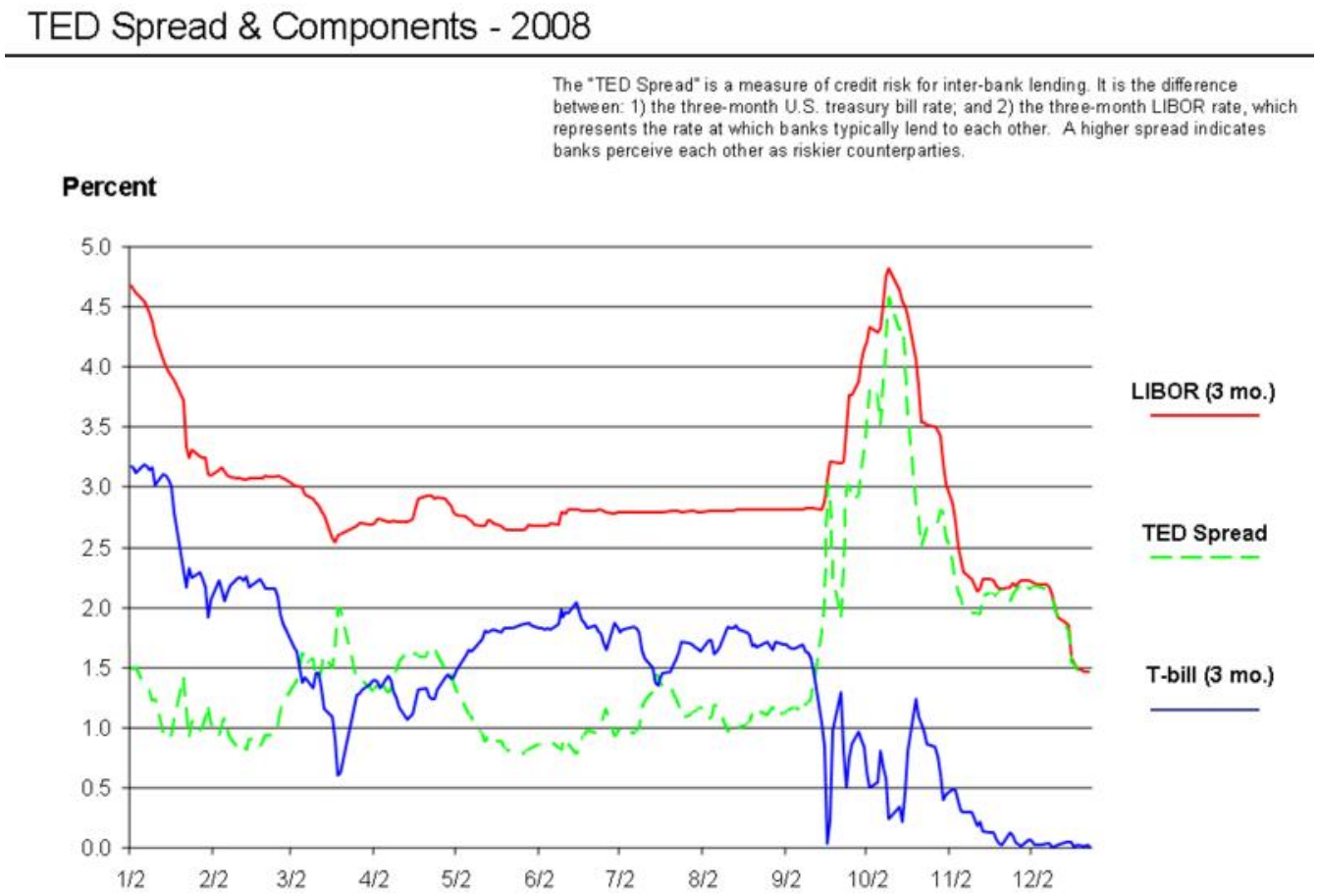
Disclaimer

This presentation outlines our general product direction and should not be relied on in making a purchase decision. This presentation is not subject to your license agreement or any other agreement with SAP. SAP has no obligation to pursue any course of business outlined in this presentation or to develop or release any functionality mentioned in this presentation. This presentation and SAP's strategy and possible future developments are subject to change and may be changed by SAP at any time for any reason without notice. This document is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. SAP assumes no responsibility for errors or omissions in this document, except if such damages were caused by SAP intentionally or grossly negligent.

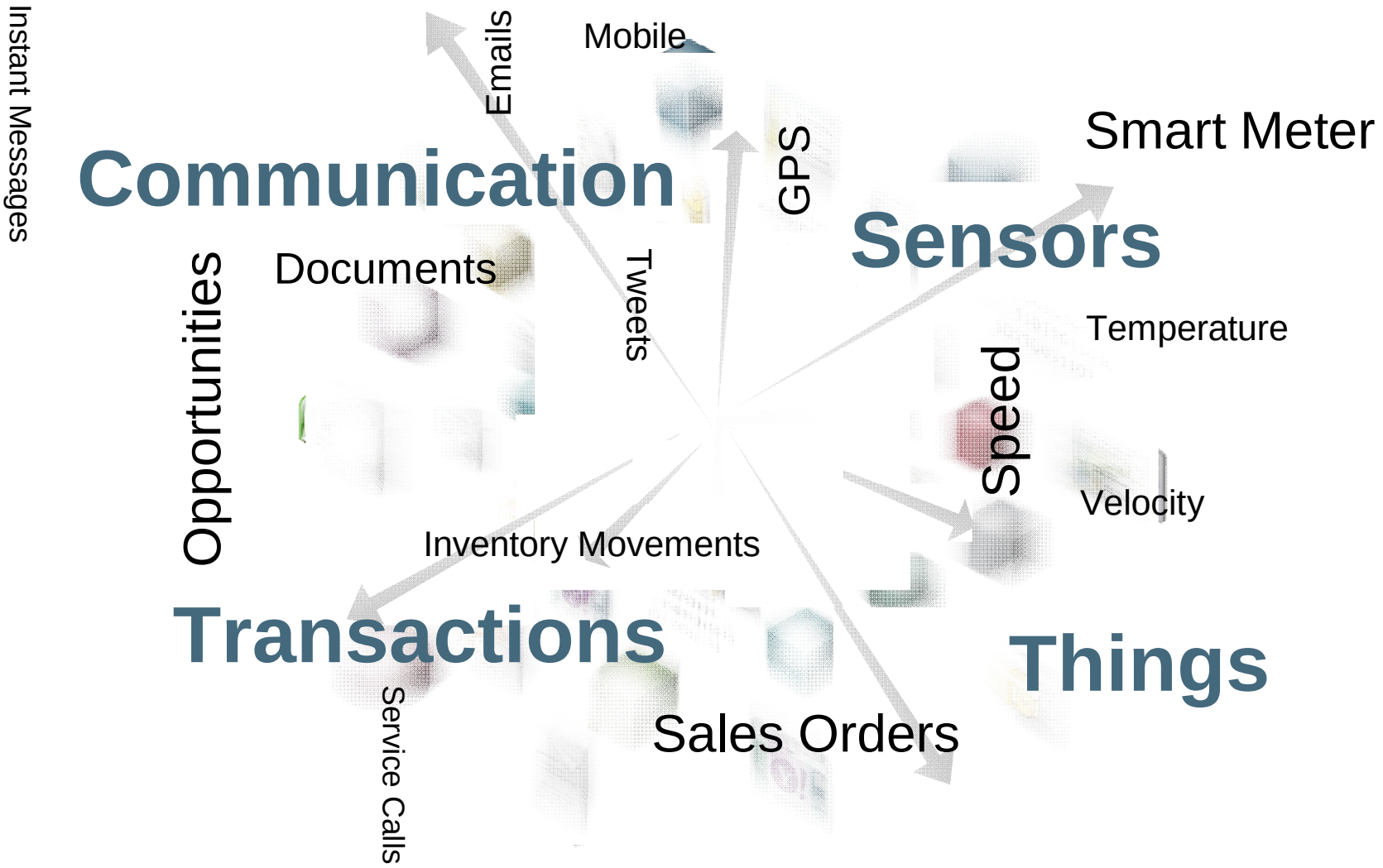
AGENDA

1. The Challenge
2. SAP In Memory Computing
3. HANA

Speed of Change



Information Explosion



Mobile Access

Challenge

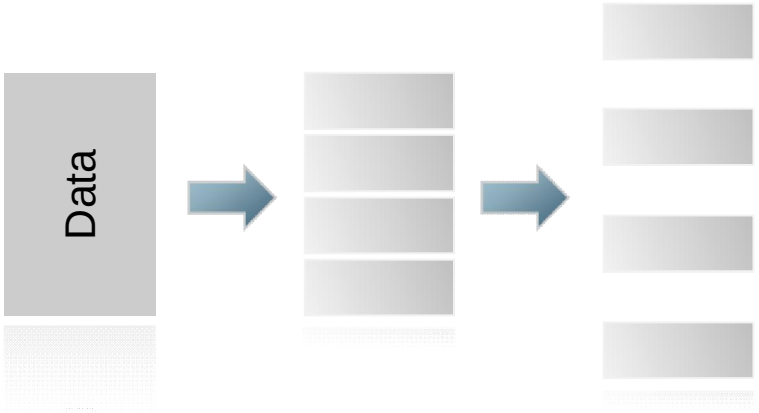
Immediate results

Massivly growing volumes of data

High Flexibility

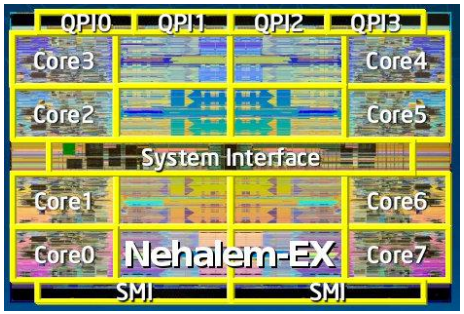
Scale – Concept

Partitioning



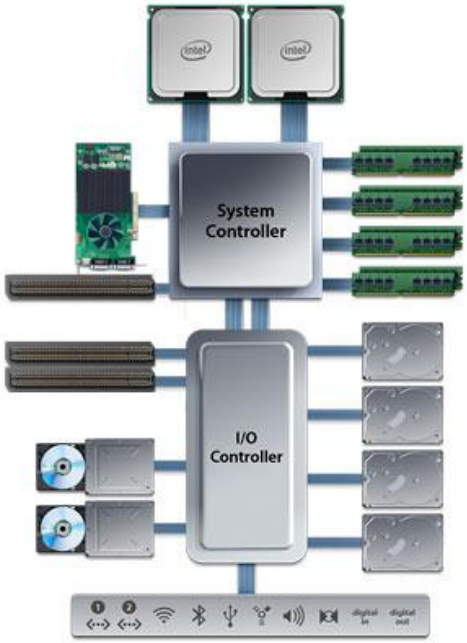
If you happen to have a BIG problem,
you tend to divide and conquer.

Scale – Hardware/Cores

2002	2006	2010
		
1 Core	4 Cores	**8 Cores**
120nm process	65nm process	45nm process
1.8 GHz	1.6-3GHz	2.26GHz
32 bit	64 bit	64 bit

Scale – Hardware/CPUs

2007

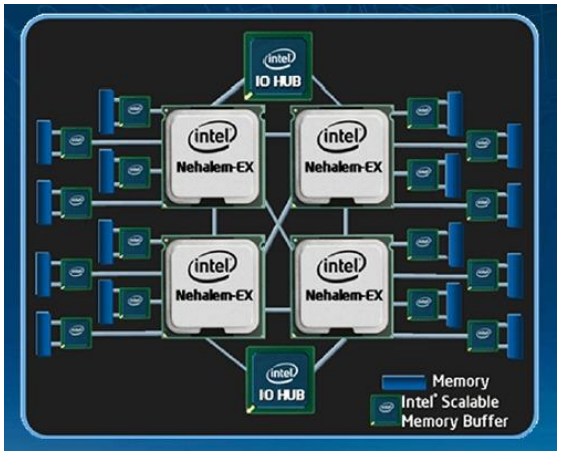


2 CPUs per Server

Memory via Controller

CPUs via Controller

2010



4 CPUs per Server

Memory Controller on Chip

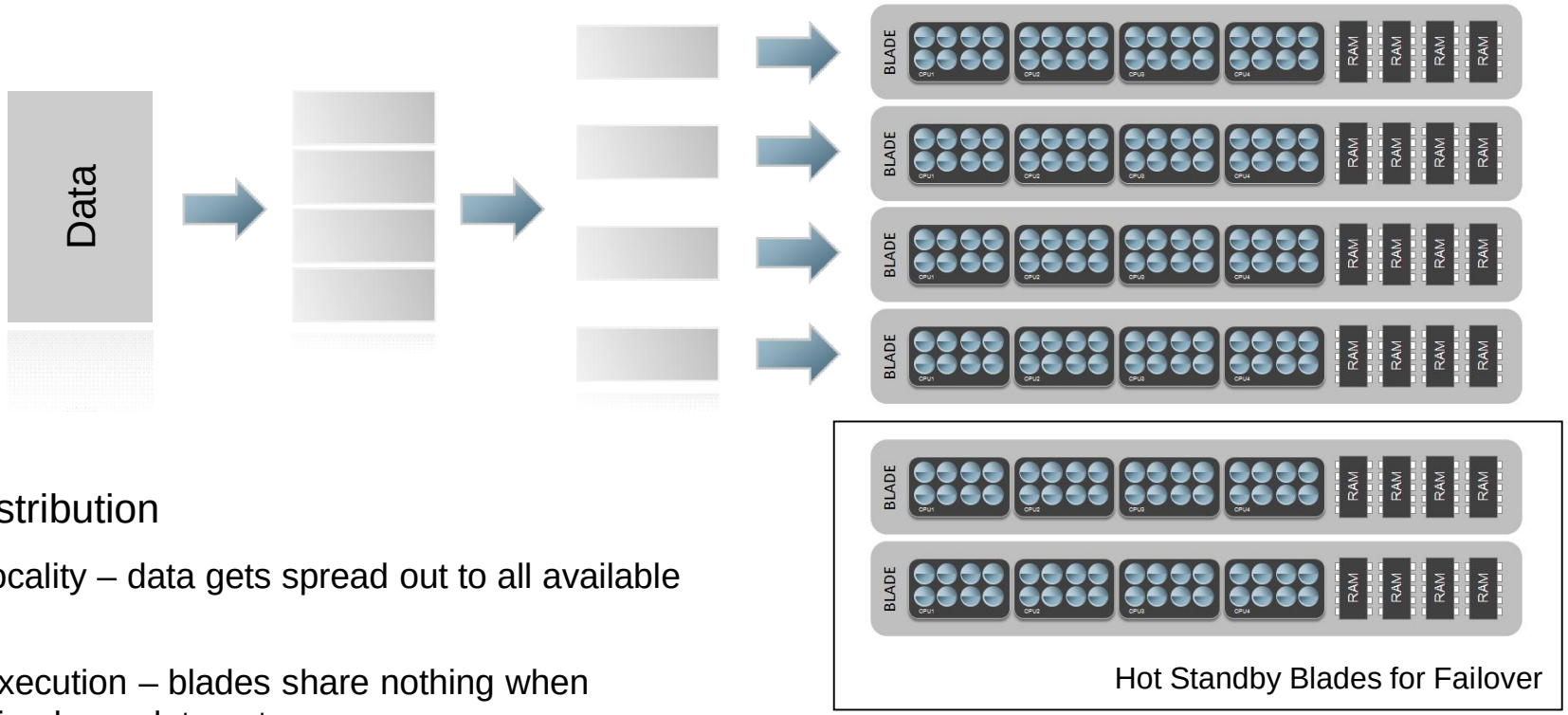
QPI

Scale – Hardware

1. Dimension – Mulicore CPUs	8 Cores per CPU
2. Dimension – Multi-CPU Boards	4 CPUs per Blade
3. Dimension – Multiple Blades	4 Blades

= 128 cores!

Scale – SW side distribute across cores



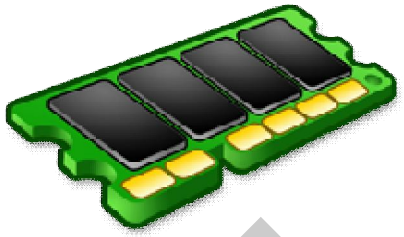
Data Distribution

- RAM locality – data gets spread out to all available cores
- MPP execution – blades share nothing when crunching large data sets
- Failover - Individual blades may fail without causing problems

Fast - Concept

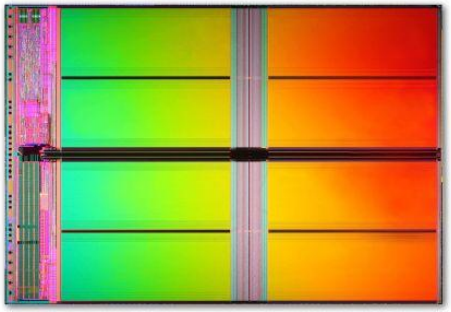
Stick the data in memory ...

DATA



... to keep Cores happy

Fast – Memory Volume



32 bit Systems	2^32 = 4,294,967,296 4GB limit per CPU
64 bit Systems	2^64 = 18,446,744,073,709,551,616 Only constraint by physics (64/128 modules)

Capacity per module grows with new production processes (8GB – 16GB – 32GB)

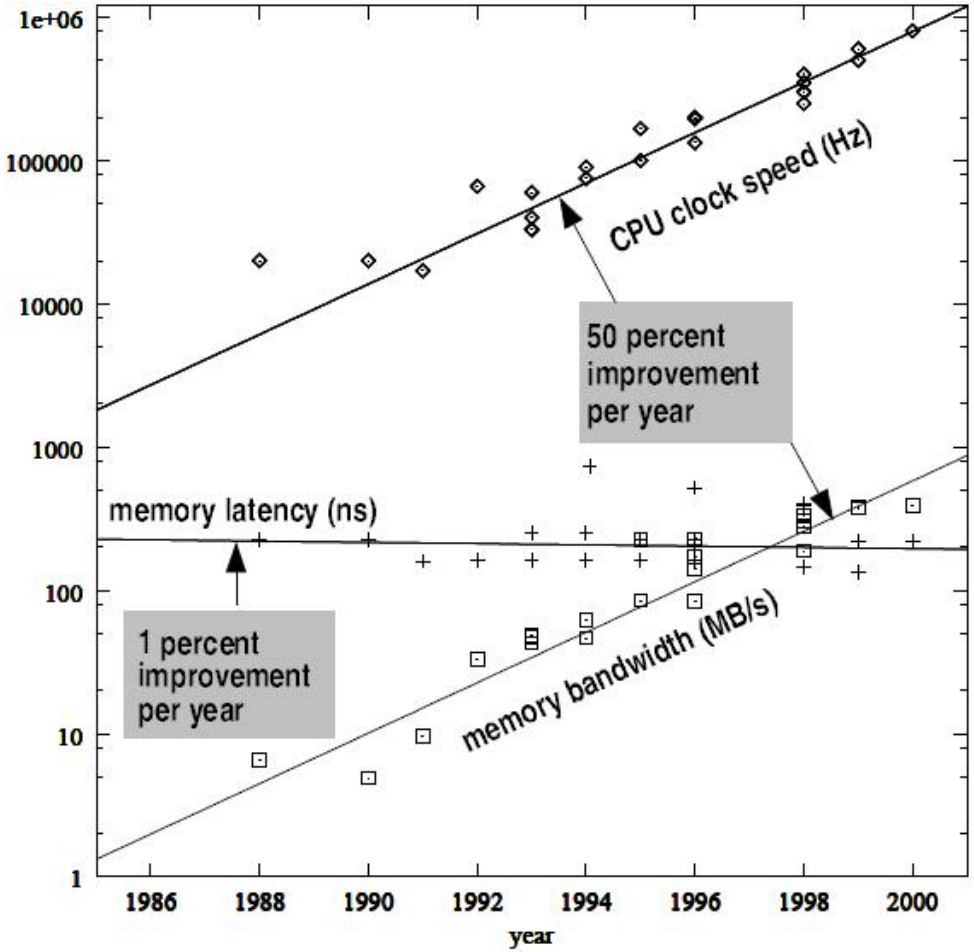
Price per module shrinks with new production processes (-30% with 32nm)

1 blade with 64 modules can hold up to **1TB** (16GB)

Fast – Memory Latency

Memory bandwidth increasing at 50% per year but latency remains pretty much constant!

Software needs to be smart about data access, it cannot just assume that data in memory equals fast!



Fast – SW side optimization for memory

conceptual view

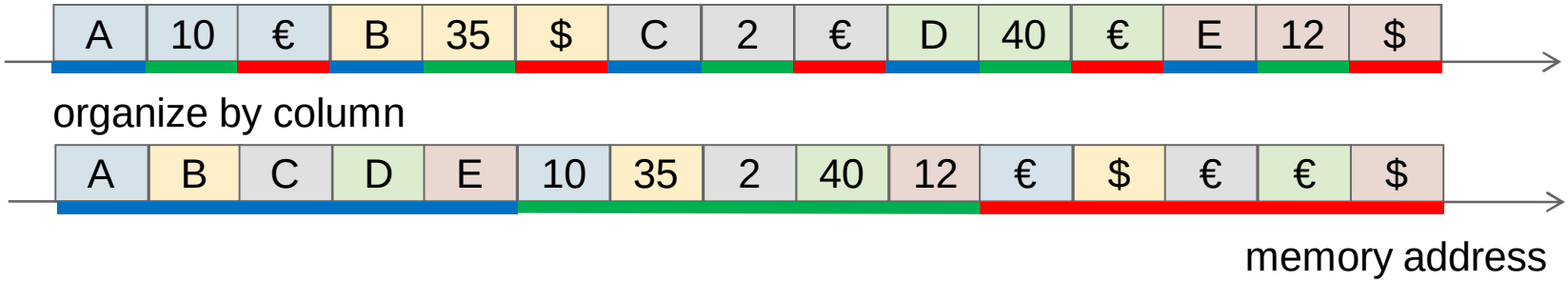
A	10	€
B	35	$
C	2	€
D	40	€
E	12	$

mapping to memory

Conventional databases store records in rows

Storing data in columns enables faster in-memory processing of operations such as aggregates

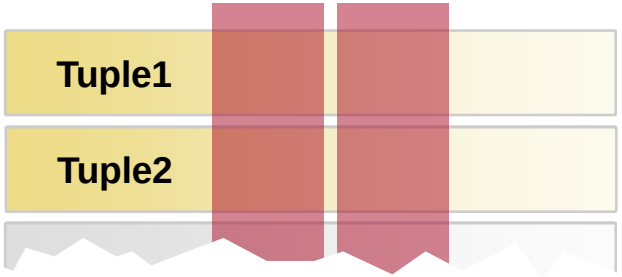
- Columnar layout supports sequential memory access
- A simple aggregate can be processed in one linear scan



Fast – Further SW side optimization for speed

Row wise

stores tables by row



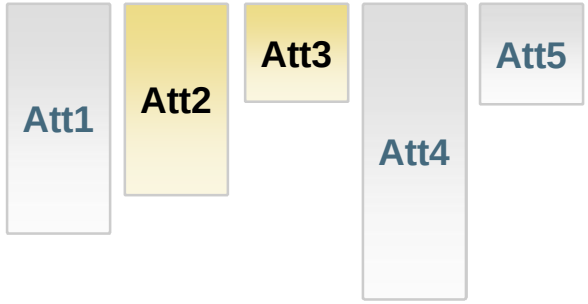
Data records are available as complete tuples in one read.

Compression is limited.

Accessing only few attributes for each tuple is an expensive operation.

Column wise

stores tables by column



Columns can be accessed in one read.

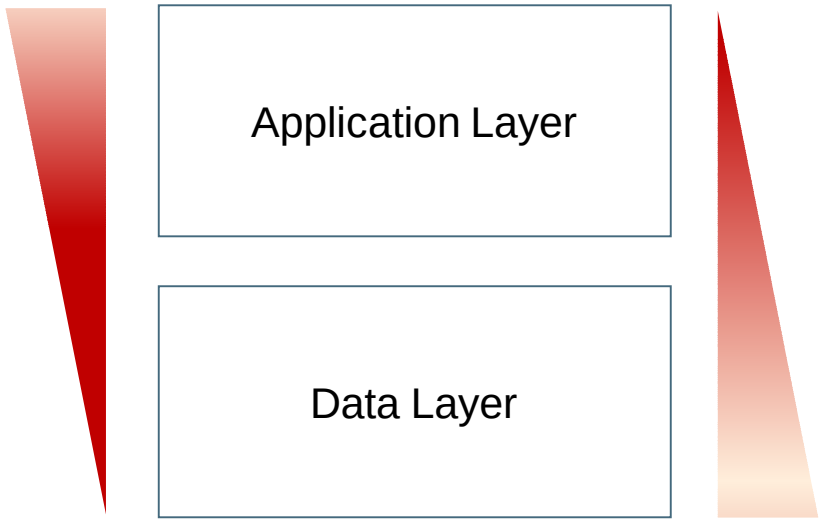
Columns contain only one data type, enabling very high compression (10x)

Accessing all attributes for one tuple (record) is an expensive operation.

Fast – Further SW side optimization for speed

Delegation of data intense operations to the in-memory computing

Today‘s applications execute many data intense operations in the application layer

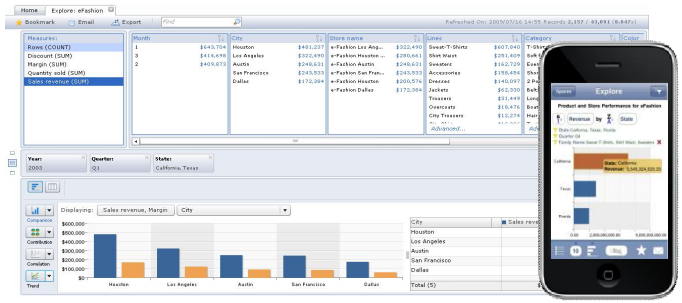


High performant apps delegate data intense operations to the in-memory computing

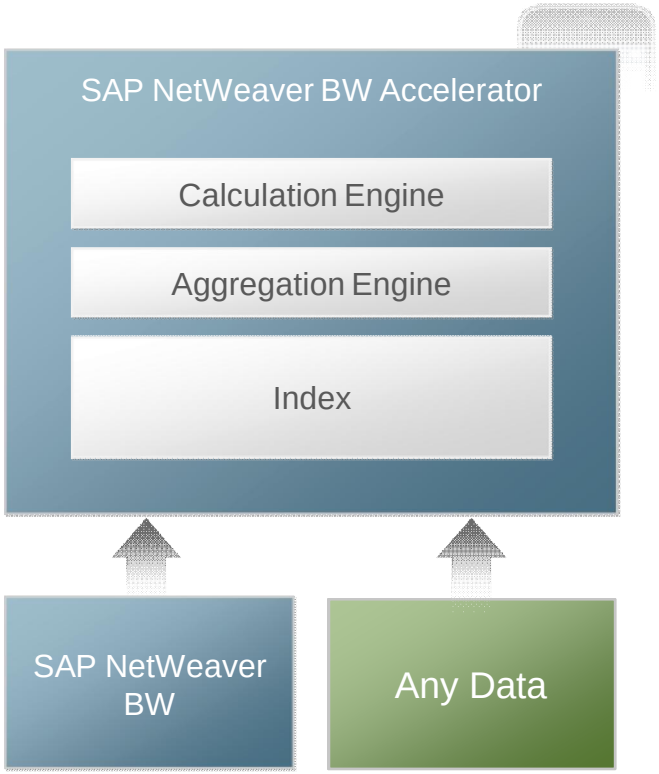
In-Memory Computing Imperative - Avoid movement of detailed data – calculate first, then move results

Technology Proof

SAP Business Objects Explorer Accelerated
for intuitive enterprise data exploration



- Intuitive user experience for casual business users
- Search and explore large volumes of enterprise data to discover relationships and uncover root cause
- Business users gain immediate 'insight at the speed of thought' without needing assistance from a business analyst or IT
- 1000+ productive installations



Technology Proof

SAP High Volume Customer Segmentation

For Marketing Users

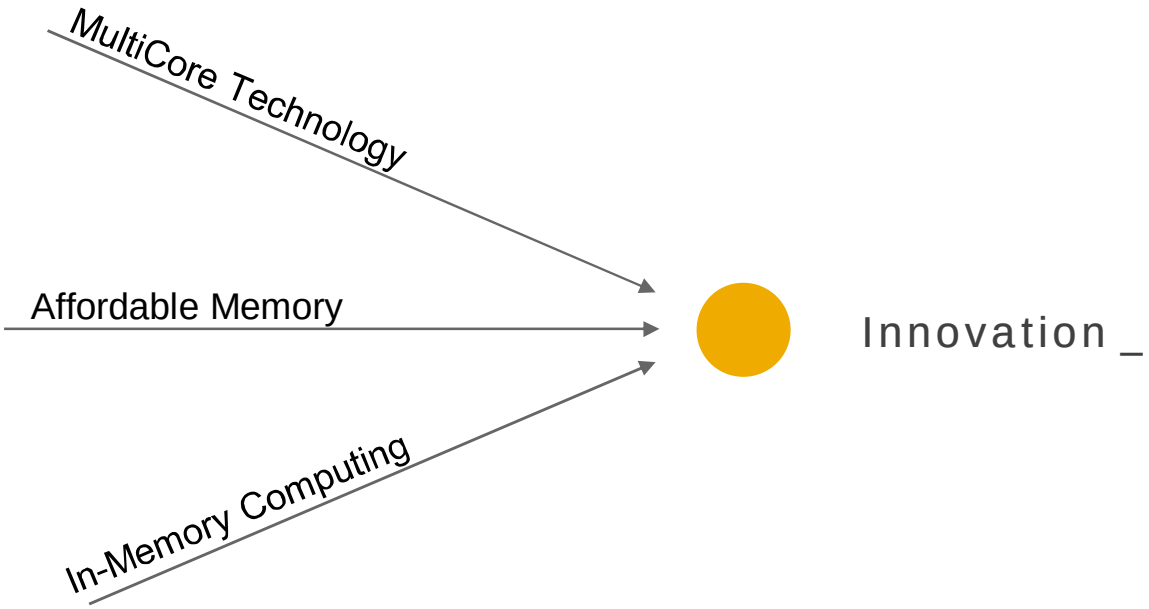
- Immediately select specific customer segment based on unique attributes from entire SAP CRM customer database
- Drag and drop attribute selection for real time segmenting
- Associate segmented customers with campaign in SAP CRM

SAP NetWeaver Enterprise Search

For Analysts

- Search for data across different sources simultaneously
  - structured data (e.g. ERP applications & business intelligence data)
  - unstructured data (e.g. PDFs, Microsoft Office formats, HTML)
- Single entry point for your work

Pivotal moment



- Disk-based row-stores were and are optimized for the disk I/O bottleneck
- The advances in hardware of the last two decades allow to fundamentally rethink this basic design goal

AGENDA

1. The Challenge
2. SAP In Memory Computing
3. HANA

SAP In-Memory Computing

In-Memory Computing

Technology that allows the processing of
massive quantities of real time data
in the main memory of the server
to provide immediate results from
analyses and transactions

More than an In-Memory DB

Only SAP...

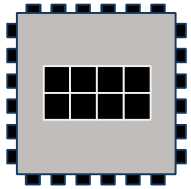
- ...combines transactional and analytical applications enabled by in-memory computing
- ...can conduct analytics, performance management, operations in a single system – real time link between insight, foresight and action
- ...can enhance existing investments with in-memory capabilities: SAP customers implement in-memory computing without disruption – no change to IT roadmaps.
- ...collaborates with industry leading partners to deliver in-memory computing solutions

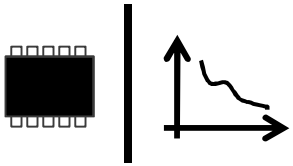
SAP In-Memory Computing Engine

Multi-Core Architecture

Massive parallel scaling with many blades

Low price for performance





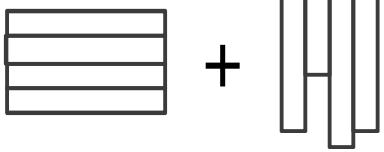
64bit, up to 2TB (today)

100GB/s data throughput

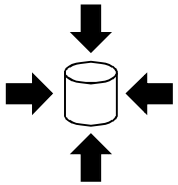
Low price forperformance

HW Technology Innovations

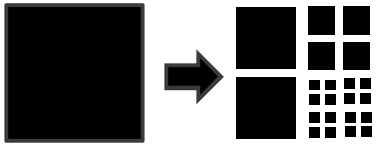
SAP SW Technology Innovations



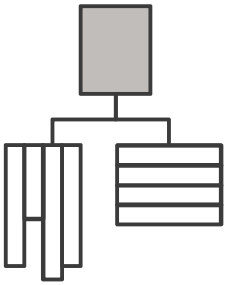
Row and Column Store



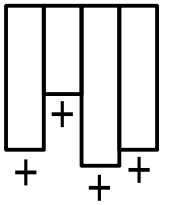
Compression



Partitioning



No Aggregate Tables



Persistency

AGENDA

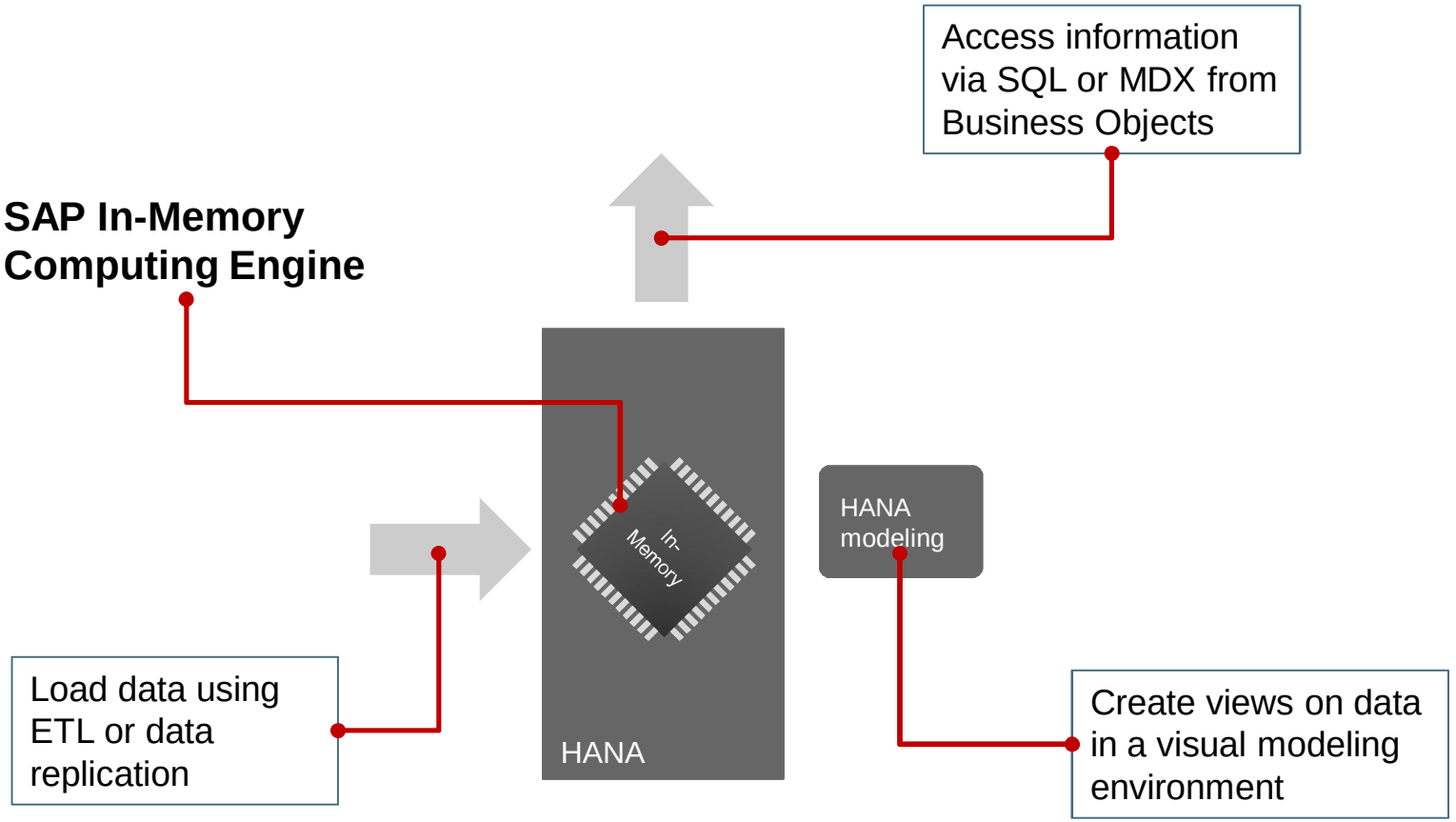
1. The Challenge
2. SAP In Memory Computing
3. HANA

SAP High Performance Analytic Appliance (SAP HANA) – current ideas

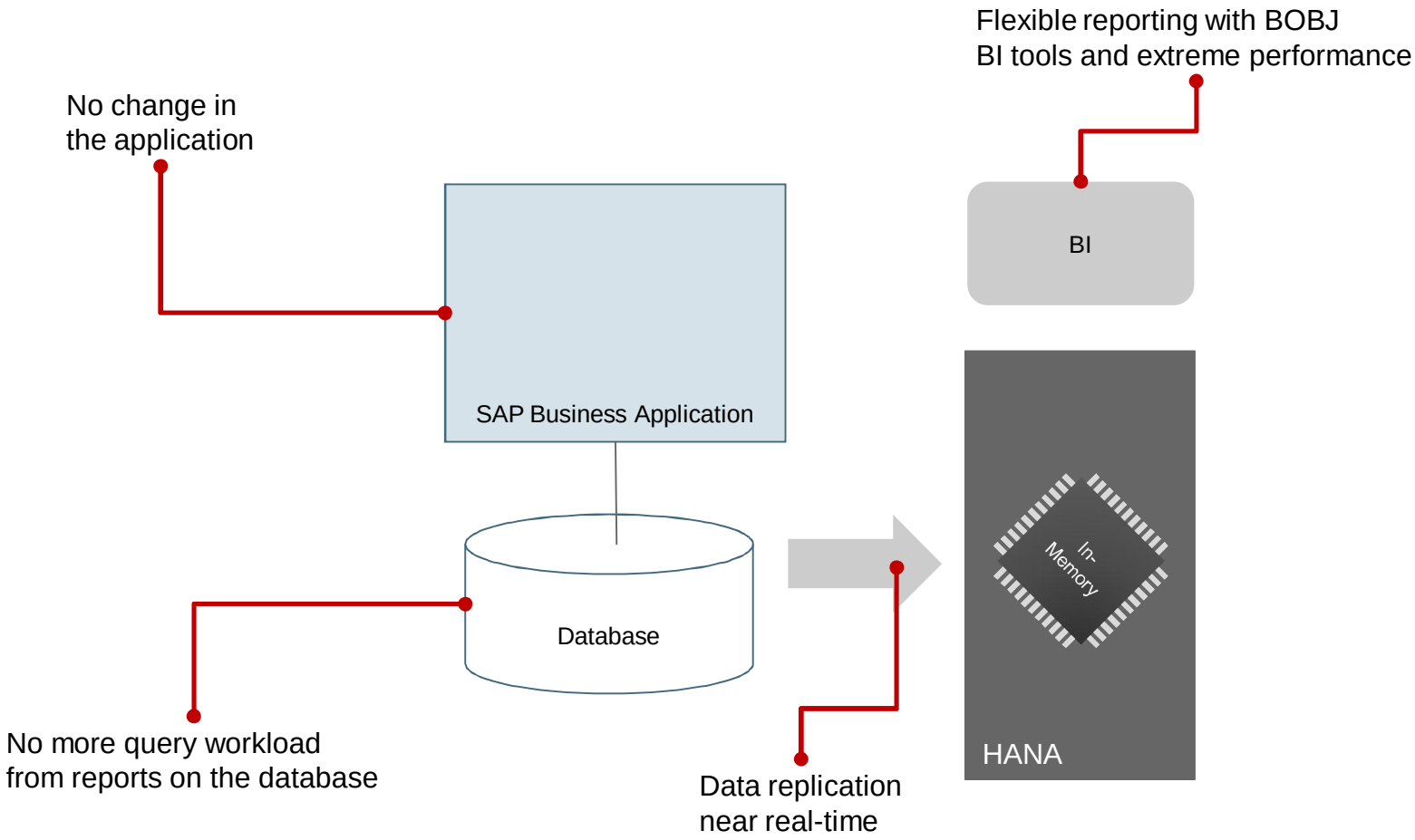
SAP HANA is planned to be ...

... the first product which incorporates the SAP In-Memory Computing Engine

... shipped as an appliance with our hardware partners



Potential use case with operational reporting



Use case

Challenge

SAP GFO was looking for a solution to implement sales forecast reports.

The sales forecast report should reflect sales activities in real time, in order to provide immediate insight into business for management.

Management should be able to navigate data withour constraints to get a proper understanding for all opportunities and sales activities.

Conditions

No changes to the CRM system.

No performance impact on the database running underneath the CRM system.

No ABAP programming.

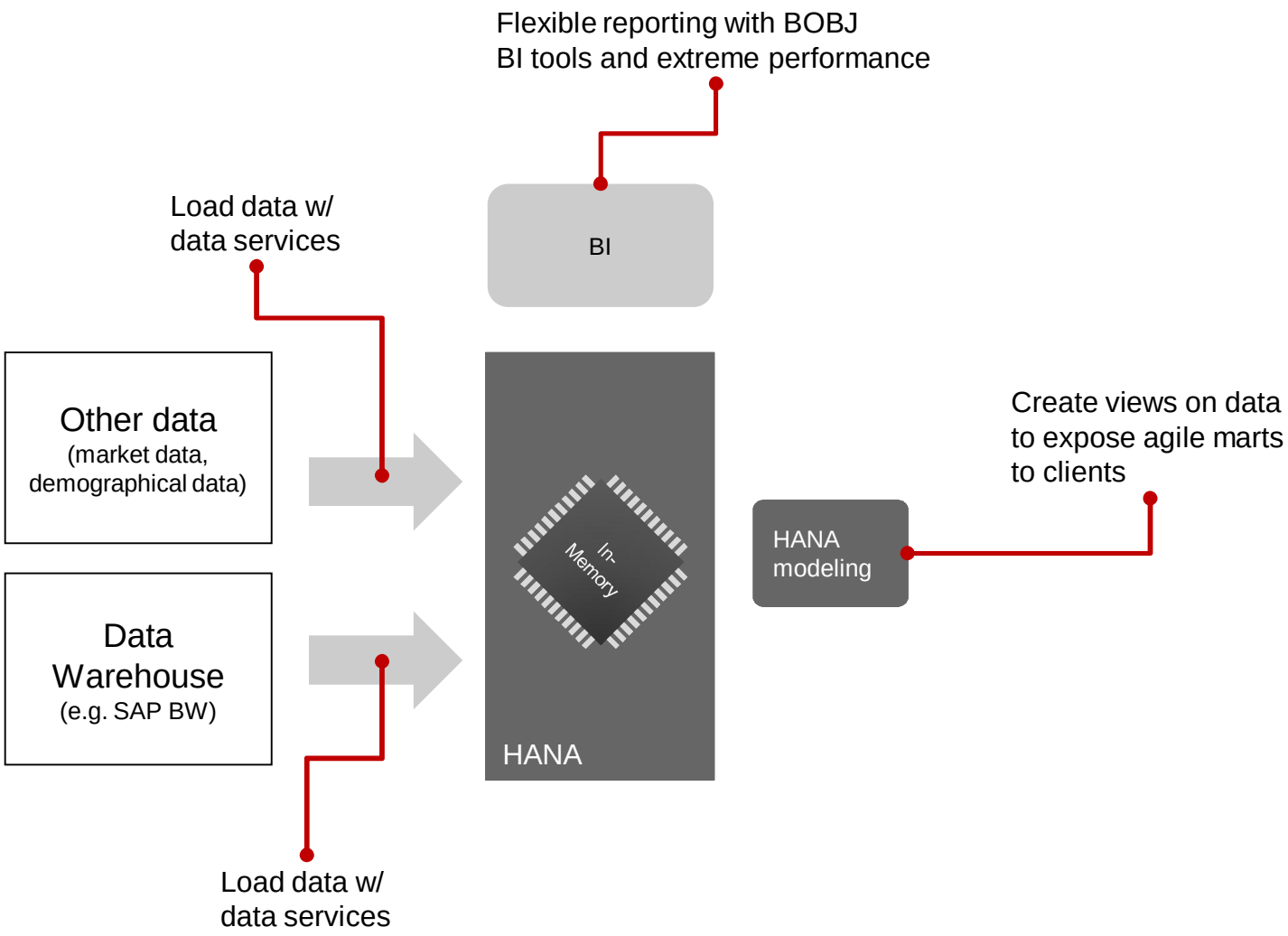
Solution

Replicate CRM data into SAP HANA.

Create analytic views for consumption in BI tools.

Consume analytic views with SBOP Explorer on iPads.

Potential use case for agile data marts



Use case

Challenge

Customer has got cost data for produced goods in SAP BW.

Business departments would like to simulate the imact of changes in BOM before entering these changes in the business system.

Conditions

No changes to BW models in production.

Results needed in 3 business days.

Cost data is not allowed to leave managed systems

Solution

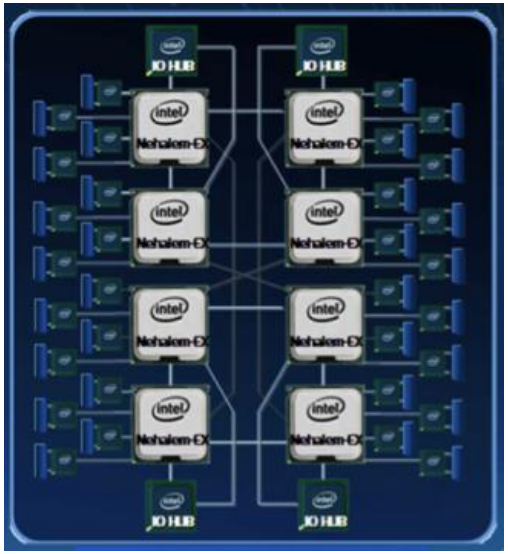
Load the cost model from BW into HANA.

Read BOM data into Excel, change it and write it into HANA.

Create an analytic view combining the cost data with the new BOM data.

Consume analytic view wiht SBOP Advanced Analysis

Some early numbers



Volume tests ran with 460bn records.

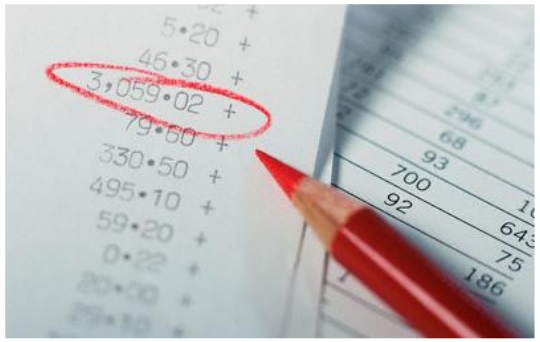
Scan speed – 1million records / ms / core

Aggregation speed – 10m records / sec / core

200x price performance improvement

SAP In-Memory Computing

To Empower Your Organization...



...plan smarter

...run faster

...perform better

- Next wave of technology innovation
- Combined in-memory analytics & transactional applications
- Our strategy is to provide the continuous real-time link between insight, foresight and action
- SAP HANA is planned to go into ramp-up by end-of-year

Feedback

Please complete your session evaluation.

Be courteous — deposit your trash,
and do not take the handouts for the following session.

THANK YOU !

© 2010 SAP AG. All Rights Reserved