2007 ASUG Annual Conference

April 22-25

Georgia World Congress Center

Atlanta, Georgia, USA

Managing Data Growth with SAP Netweaver 2004s NLS

Dr. Michael Hahne

Session 109

EDUCATING | NETWORKING | INFLUENCING

Learning Points

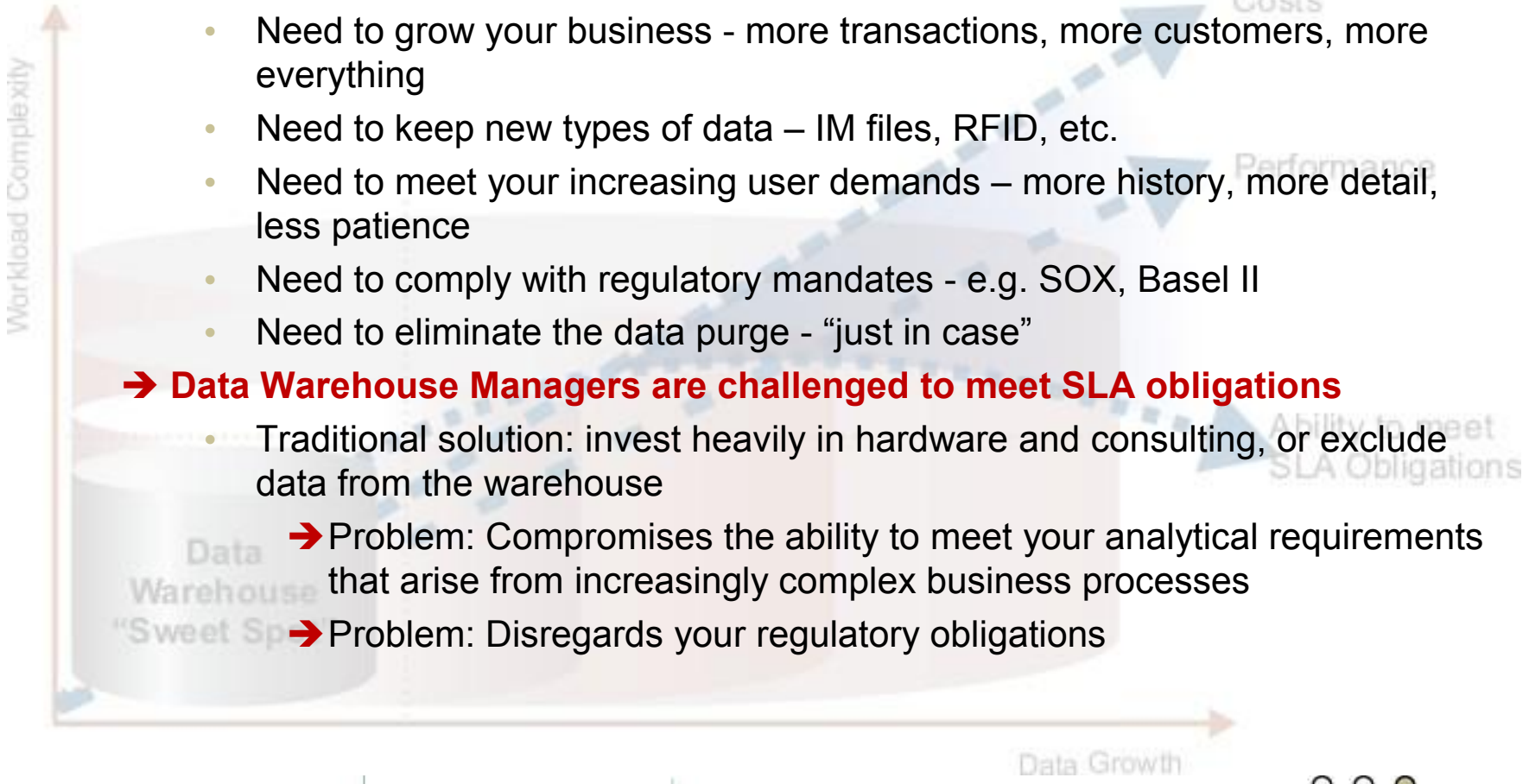
- How you can reduce terabytes of data to a tiny footprint.
- How you can realize the benefits of adding a central data management strategy based on Nearline Storage as part of your overall Information Lifecycle Management (ILM) approach.
- How you can lower your TCO and improve operational efficiency — with rapid ROI.
- How you can satisfy legal and audit needs for retaining records, and meet Service Level Agreements.

What Companies Are Facing Today…

- Unprecedented explosion in data volumes driven by:
  - Need to grow your business - more transactions, more customers, more everything
  - Need to keep new types of data – IM files, RFID, etc.
  - Need to meet your increasing user demands – more history, more detail, less patience
  - Need to comply with regulatory mandates - e.g. SOX, Basel II
  - Need to eliminate the data purge - “just in case”

➔ **Data Warehouse Managers are challenged to meet SLA obligations**

  - Traditional solution: invest heavily in hardware and consulting, or exclude data from the warehouse
    - ➔ Problem: Compromises the ability to meet your analytical requirements that arise from increasingly complex business processes
    - ➔ Problem: Disregards your regulatory obligations



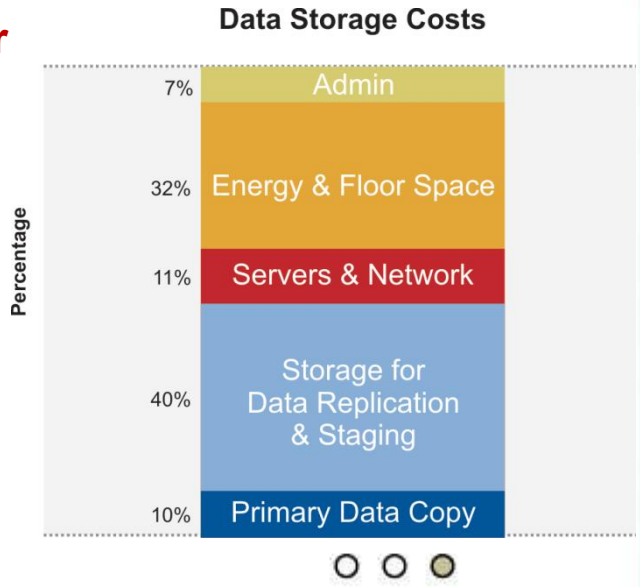
Why Not Just Add More Storage ?

- Data volumes are growing faster than the price/performance ratios of disk storage technology.
- Fast disks are still expensive
- Data stored in production environments requires failover and backup technology
- For every $1 a company spends on data storage devices, an estimated additional $5 to $10 is required to manage those devices over the lifetime of the equipment

➔ **Total costs > $150,000 to $200,000 per TB per year**

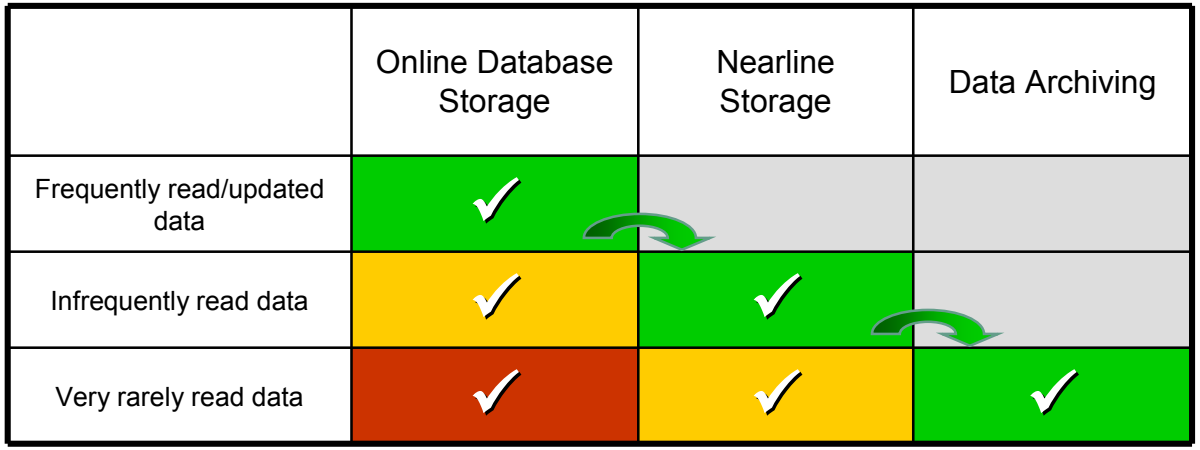
➔ More importantly, large volumes of data have adverse effects on your system's responsiveness:

- Performance of data loading
- Performance of change runs, rollups, and so on
- Backup and recovery times
- Migration and upgrade times



The Solution: SAP's ILM Strategy

Split the data according to age or frequency of access:



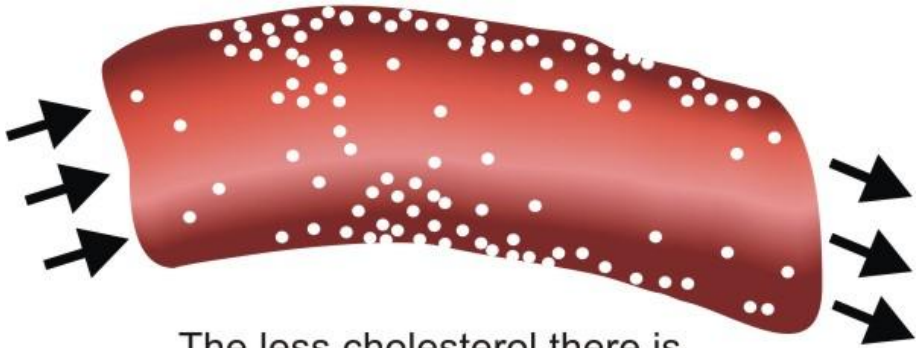
SAP has introduced an Information Life Cycle (ILM) architecture that enables SAP BI Data Warehouse Managers to:

1) Keep a "skinny", responsive relational database within SAP BI
2) Keep *all* your data accessible and usable over time
3) Keep compliant with your analytic and legal requirements
4) Keep your system availability levels according to SLA obligations
5) Keep control of your budget

Bill Inmon's Opinion about Performance Issues Surrounding NLS

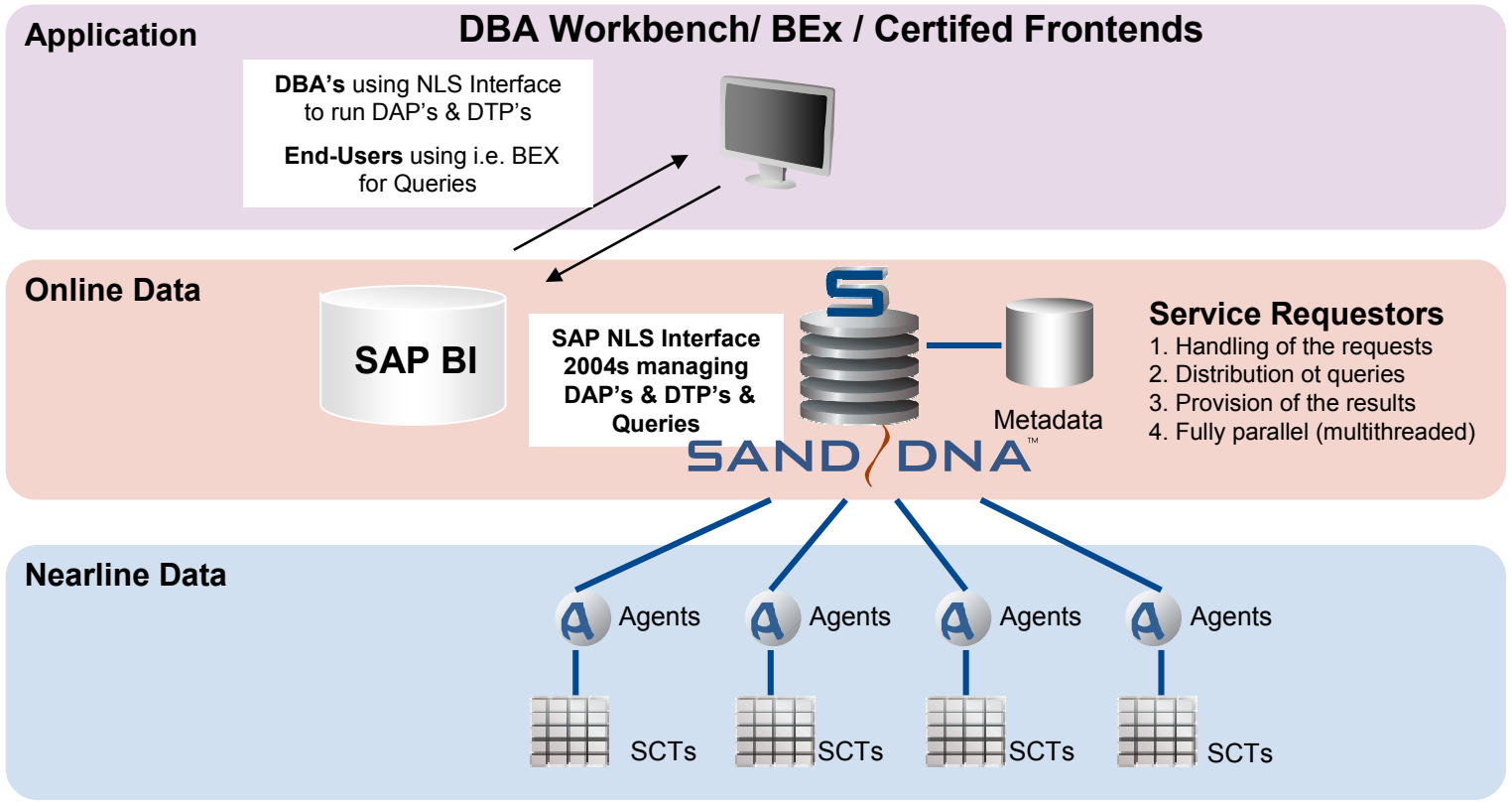
"Indeed, leaving infrequently accessed data on disk storage greatly **HURTS** performance […] **because mixing infrequently used data with actively used data is like adding lots of cholesterol into the blood stream.**"

Information Lifecycle Management for Data Warehousing:
Matching Technology to Reality
An Introduction to SAND/DNA Access
By W.H. Inmon (2005)



The less cholesterol there is,
the more efficient the flow of blood

SAP NW 2004s BI & SAND/DNA architecture



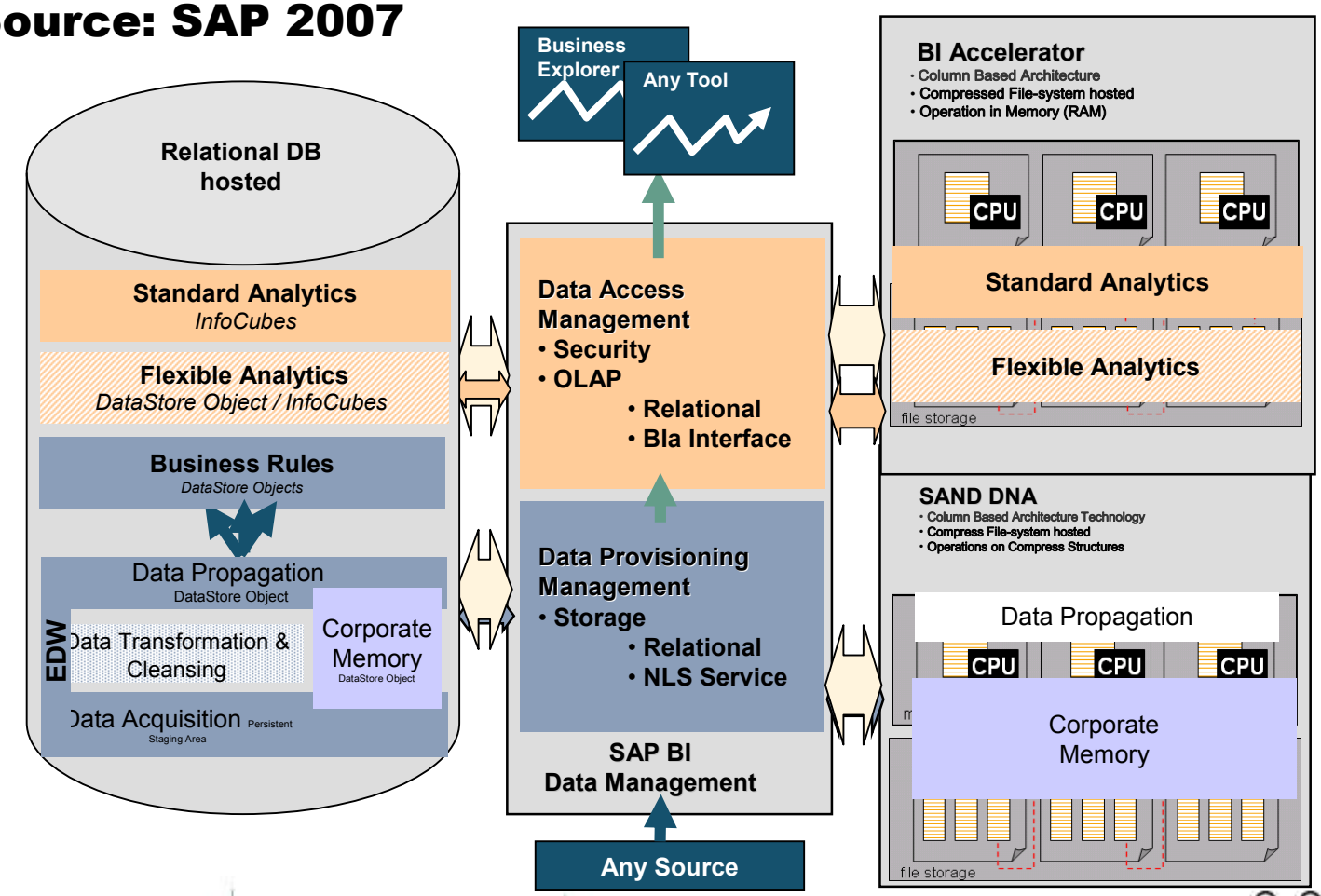
SAND/DNA for SAP BI

- Powered by NetWeaver for SAP BW 3.1/3.5, and certified and integrated into SAP NetWeaver BI 2004s infrastructure
- Data compression of at least 85% (frequently as high as 95% depending on data)
- Does not require index building, but still allows any data field to be accessed within Data Transfer Processes (DTP‘s), or via BEx or any SAP BI-certified front-end tool
- Can facilitate migration from SAP BW 3.1/3.5 to SAP NetWeaver BI 2004s and enable establishment of “Corporate Memory”
- Can run on the same server as SAP BI, or on a different server
- Software solution runs on most popular operating systems (Tru64, Solaris, AIX, HP-UX, Linux, Windows), database independent - no special hardware required
- Integrated with major archiving solutions like Opentext LiveLink and Network Appliance SnapLock to enable full Information Life Cycle Management in accordance with SAP's recommended approach

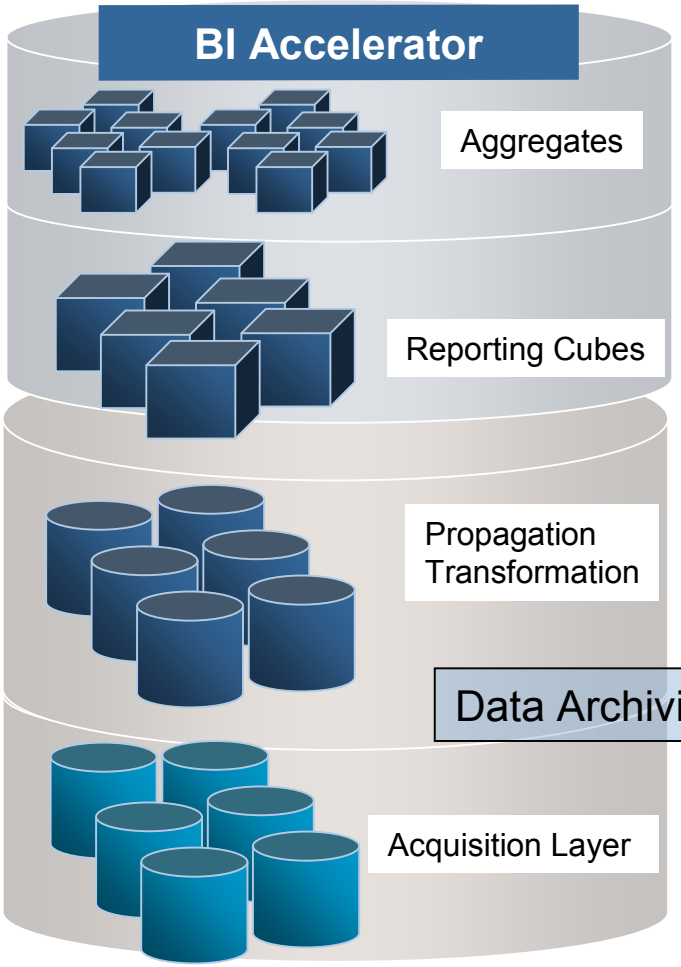
SAP NetWeaver BI 2004s Nearline & SAND/DNA Functionality

- Ability to use nearline DataStore Objects for Data Transfer Processes (DTP) that derive new InfoCubes or DataStore Objects
- Integration with “Nearline LookUp API” and Direct access to DataStore Objects and InfoCubes in nearline storage via BEx or any other certified Business Intelligence tool
- Transparent access to combined data from SAP NetWeaver BI 2004s and SAND/DNA and direct drill-down from InfoCubes into detailed data in DataStore Objects stored nearline
- Administration of all operations via Data Archiving Processes (DAP) and Process Chains
- Ability to restore nearline DataStore Objects and InfoCubes back into the SAP BI
- *Reduced administration requirements:*
  - *Automatic storage of designated tables*
  - *No need to create indexes*
  - *Any table/column can be queried*
  - *Aggregations “on the fly”*
  - *Automatic creation of SAND/DNA data structures*
  - *Automatic adaptation to structural changes in SAP BI*

Managing and Hosting Data in BI
Enabling a SAP BI Data Hub
Source: SAP 2007



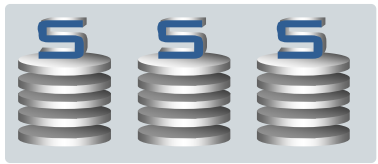
Lesson learned : Nearline concentrates on Detailed Data



- **Relieving SAP BI from detailed data**
- **Compressed by more than 85%**
- **Used as a "Corporate Memory"**

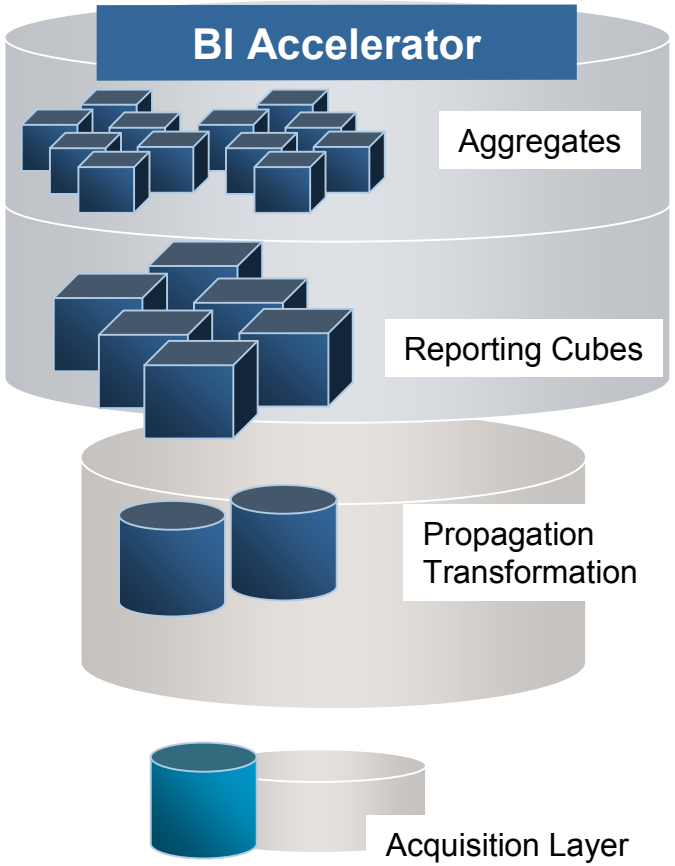
  - Details in its "pure" form
  - Infrequently used detailed data
  - "Just-in-Case" data
  - Aged and historical data
  - Legacy data

Data Archiving Process (DAP)



Efficient Corporate Memory

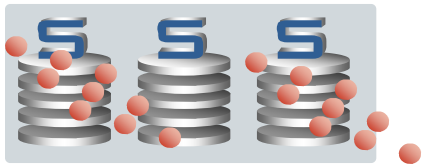
Usage of “Corporate Memory”



Greater Flexibility in Responding to New Analytical Requirements

- deriving new InfoCubes or DSO‘s
- building new KPI‘s based on historical data

Data Transfer Process (DTP) & Look Up API



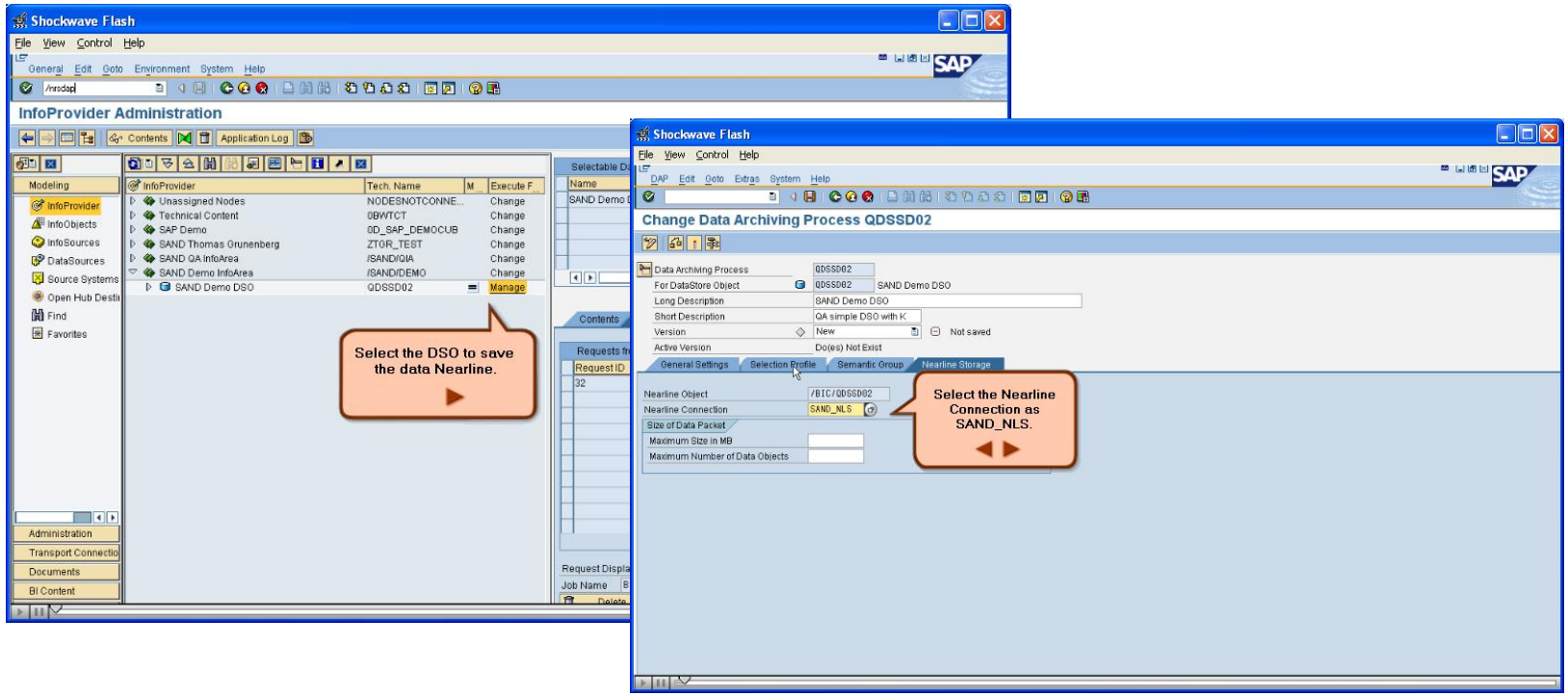
Efficient Corporate Memory

Configuring SAND/DNA With SAP NetWeaver 2004s

- How to nearline data in BI capabilities of SAP NetWeaver 2004s:
  - SAP NetWeaver 2004s supports NLS for InfoCubes and DataStore objects
  - Transparent access to online and nearlined data for queries
  - SAP NetWeaver 2004s supports an open interface for certified partners
  - SAND/DNA integrates with SAP NetWeaver 2004s via the SAP nearline adapter

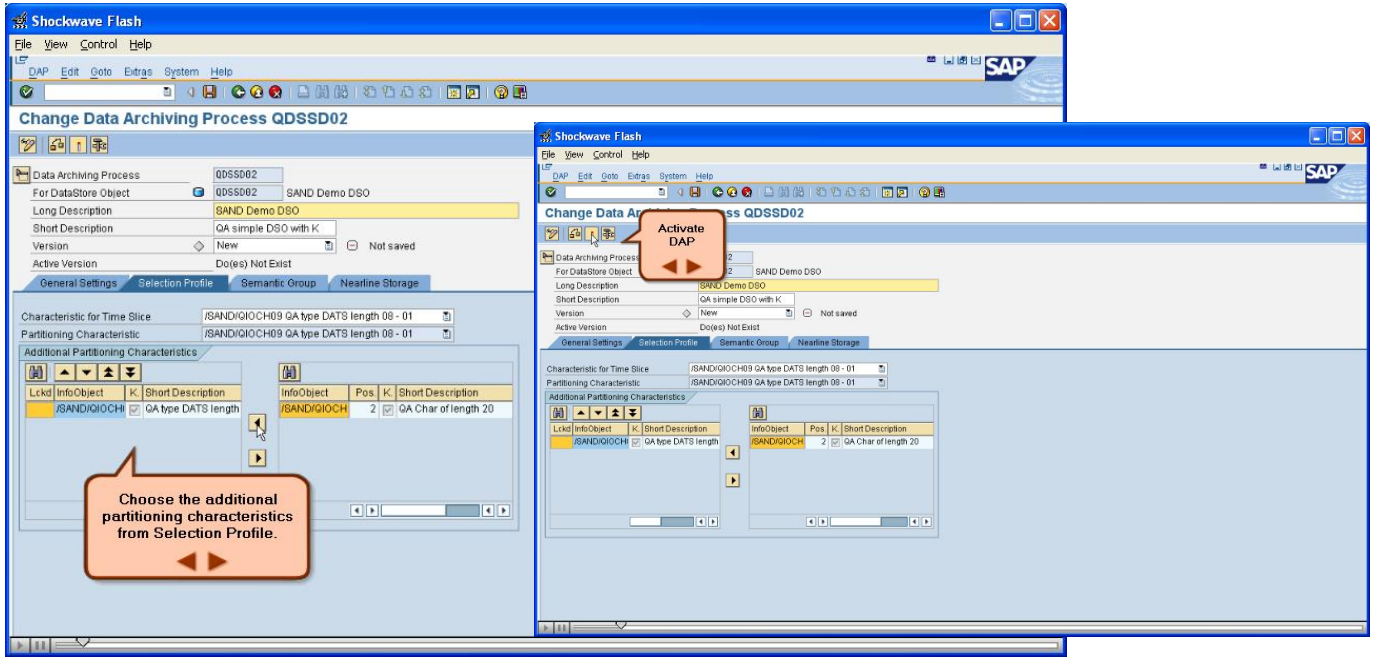
How to Nearline Data via SAND/DNA:

- From the Data Warehousing Workbench, select the DataStore Object you wish to nearline
- In the Nearline Storage tab, Select Nearline Connection to use



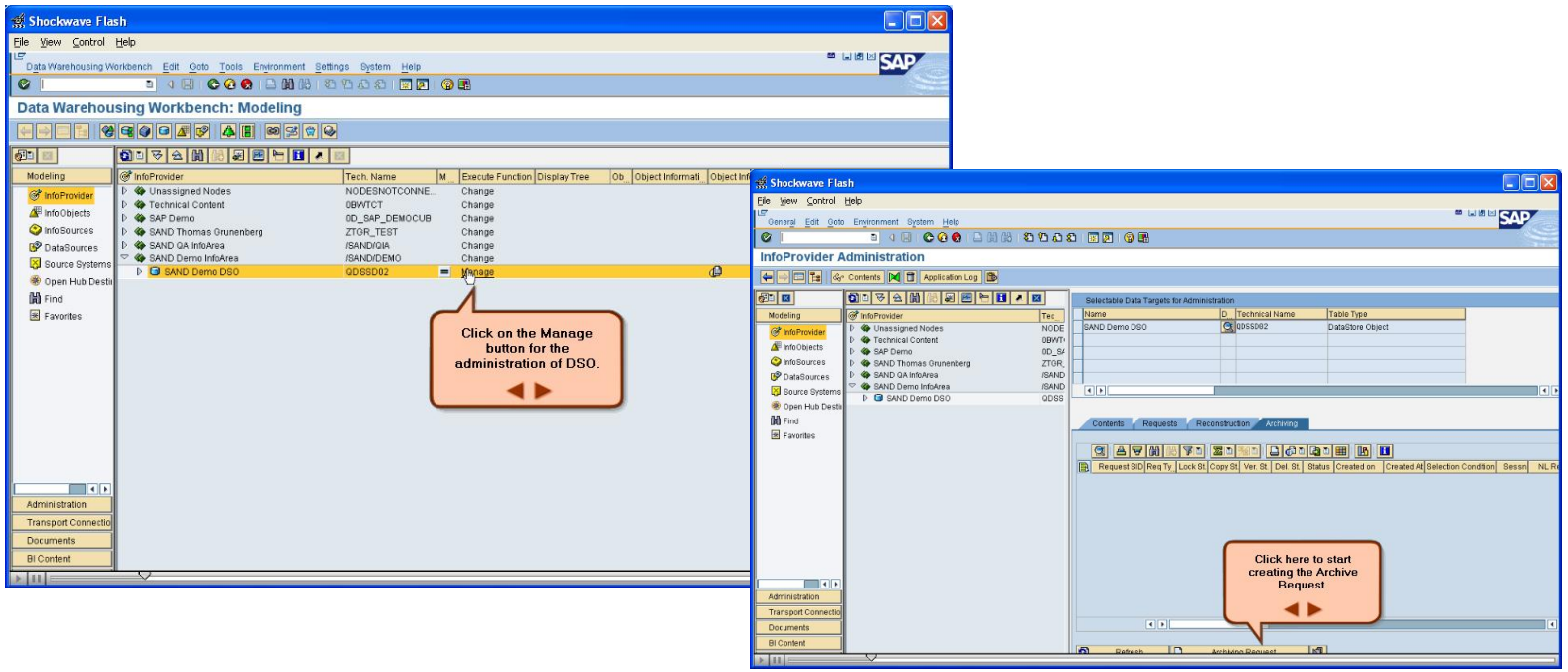
How to Nearline Data via SAND/DNA: (cont.)

- From the Selection Profile tab, choose the additional partitioning characteristics
- In this case we are using a date InfoObject
- Activate the Data Archiving Process (DAP)



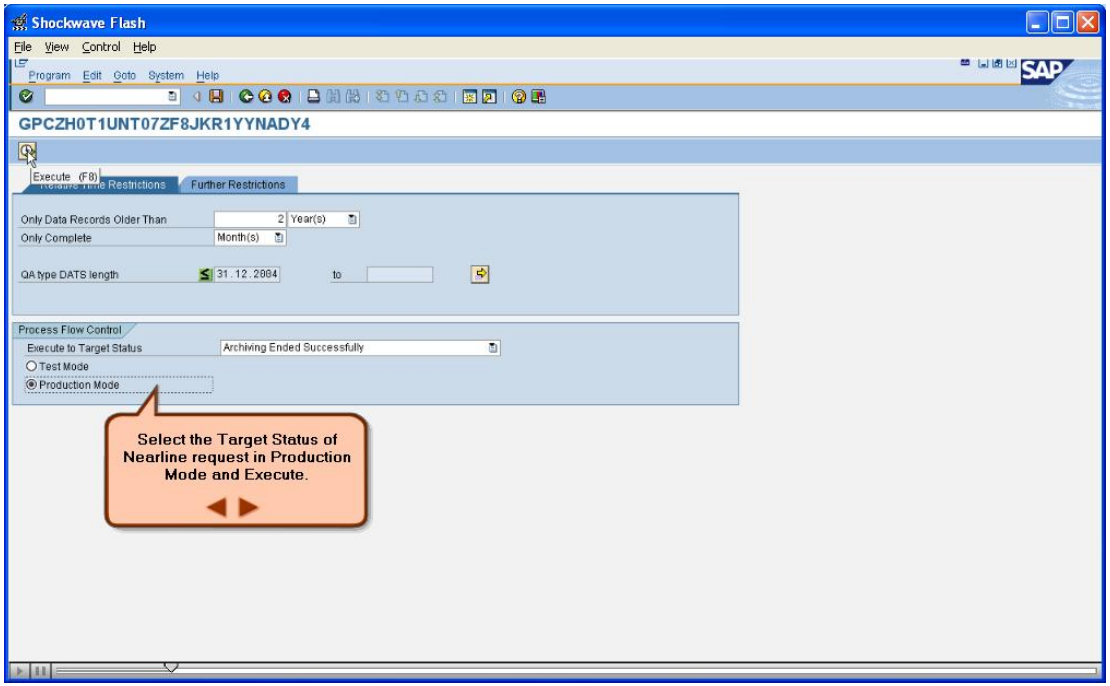
How to Nearline Data via SAND/DNA: (cont.)

- Click “Manage” to access the InfoProvider Administration screen
- Click “Archiving Request” from the Archiving tab



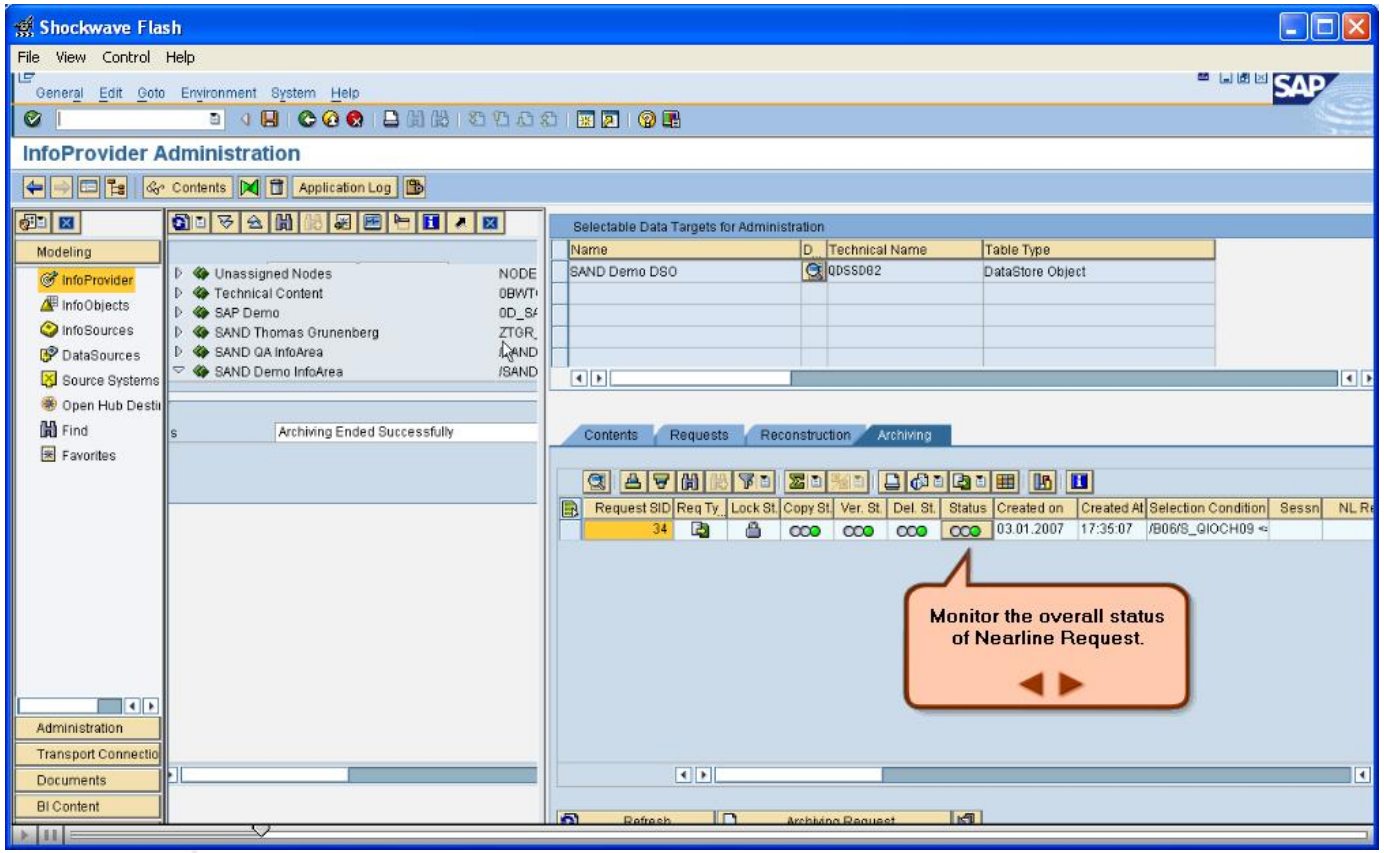
How to Nearline Data via SAND/DNA: (cont.)

- Select the nearline criteria
- Select the Target Status and execute



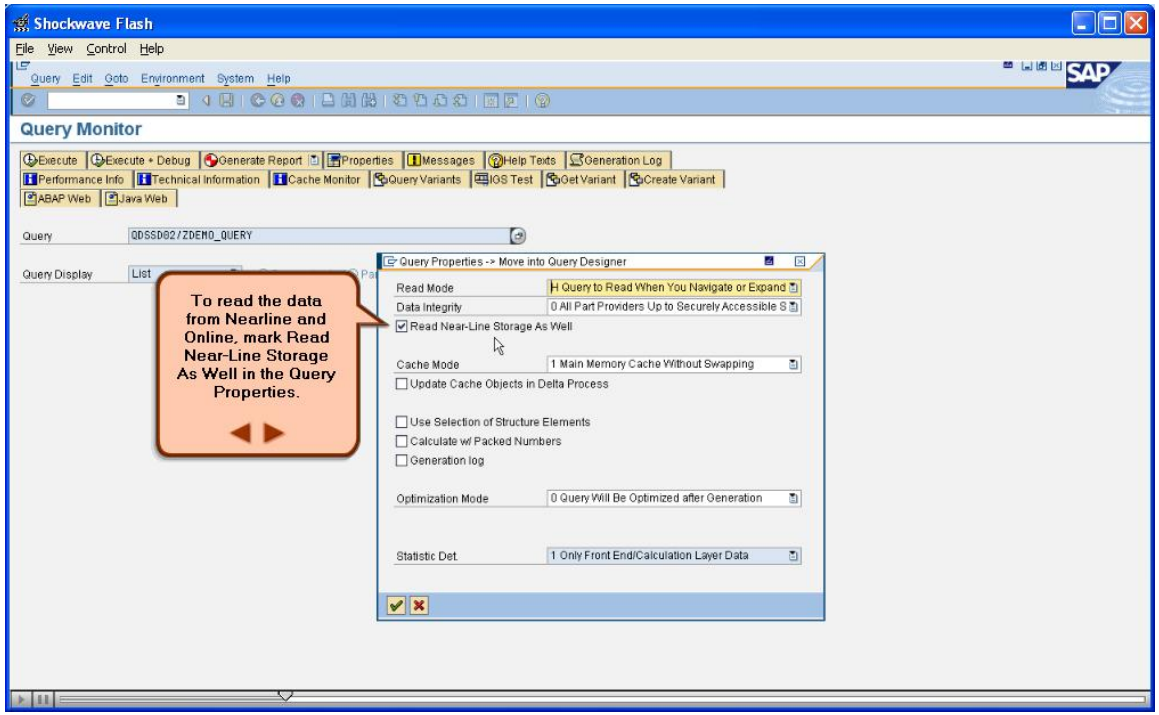
How to Monitor Nearline Data in SAND/DNA:

- After executing the DAP, you may monitor the progress of the nearline request

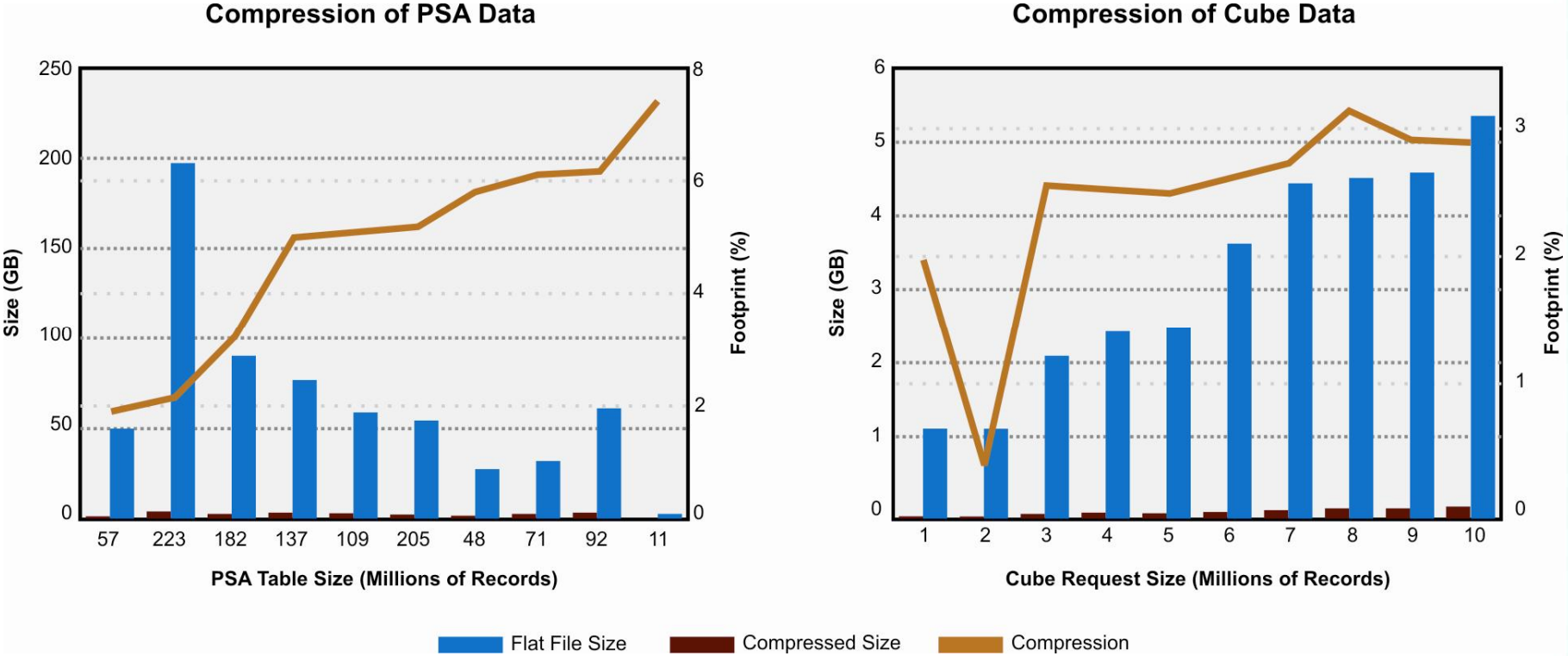


How to Set a Query to Read from Nearline and Online Storage:

- Execute the Query Monitor
- Check the “Read Near-Line Storage As Well” field

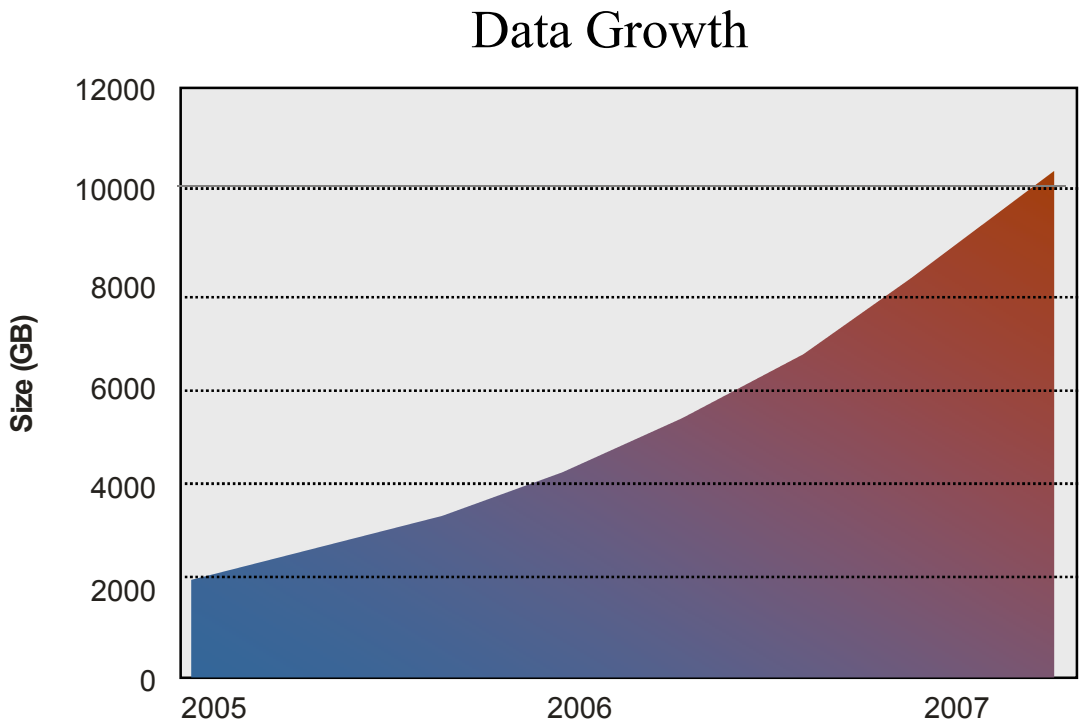


Typical Compression Ratios of SAND/DNA



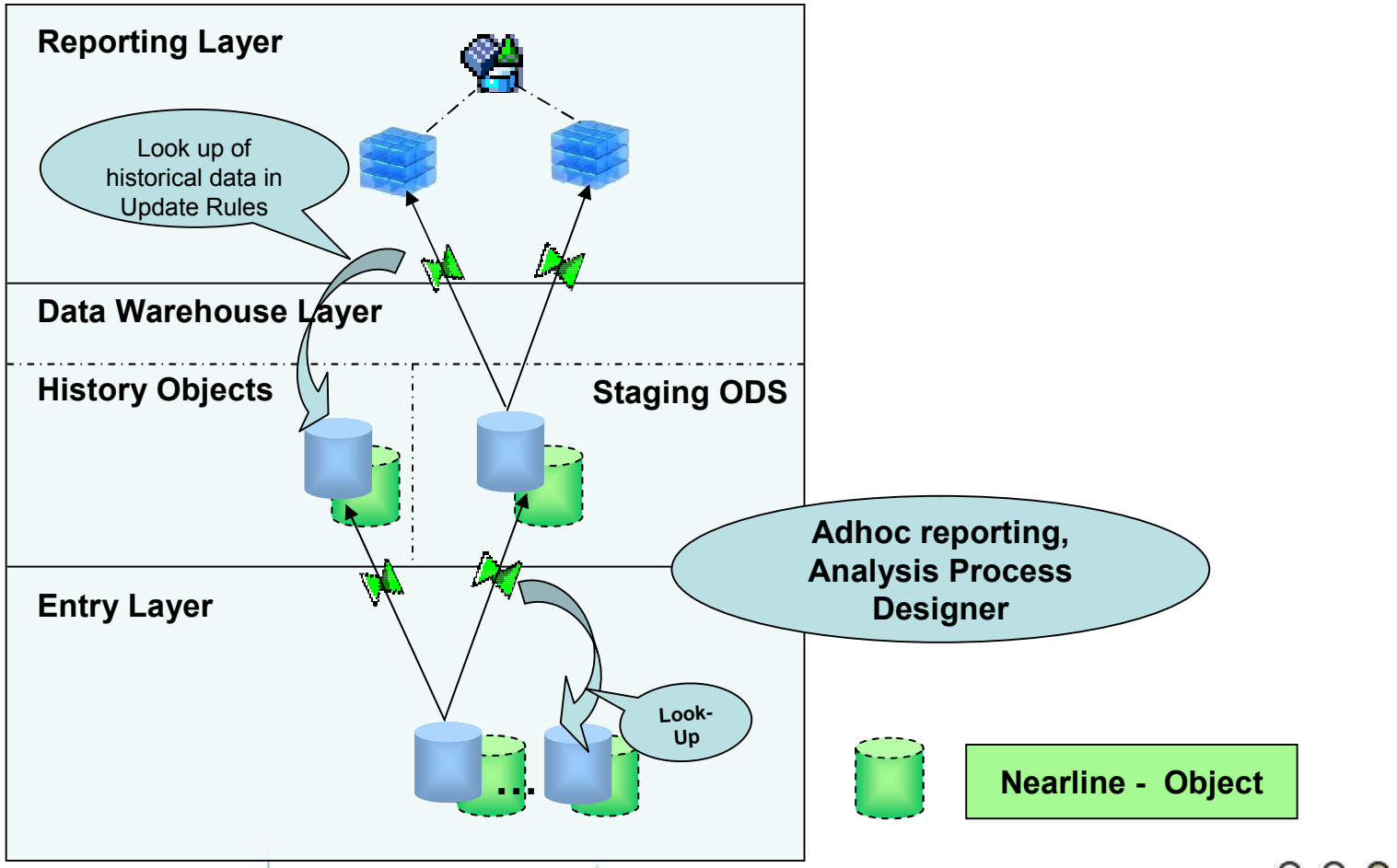
Volkswagen Bank Case Study

"1 TB of data in our SAP BI production environment generates 5 TB for failover and backup processes." *- Adrian Bourcevet, Volkswagen Bank GmbH Germany*

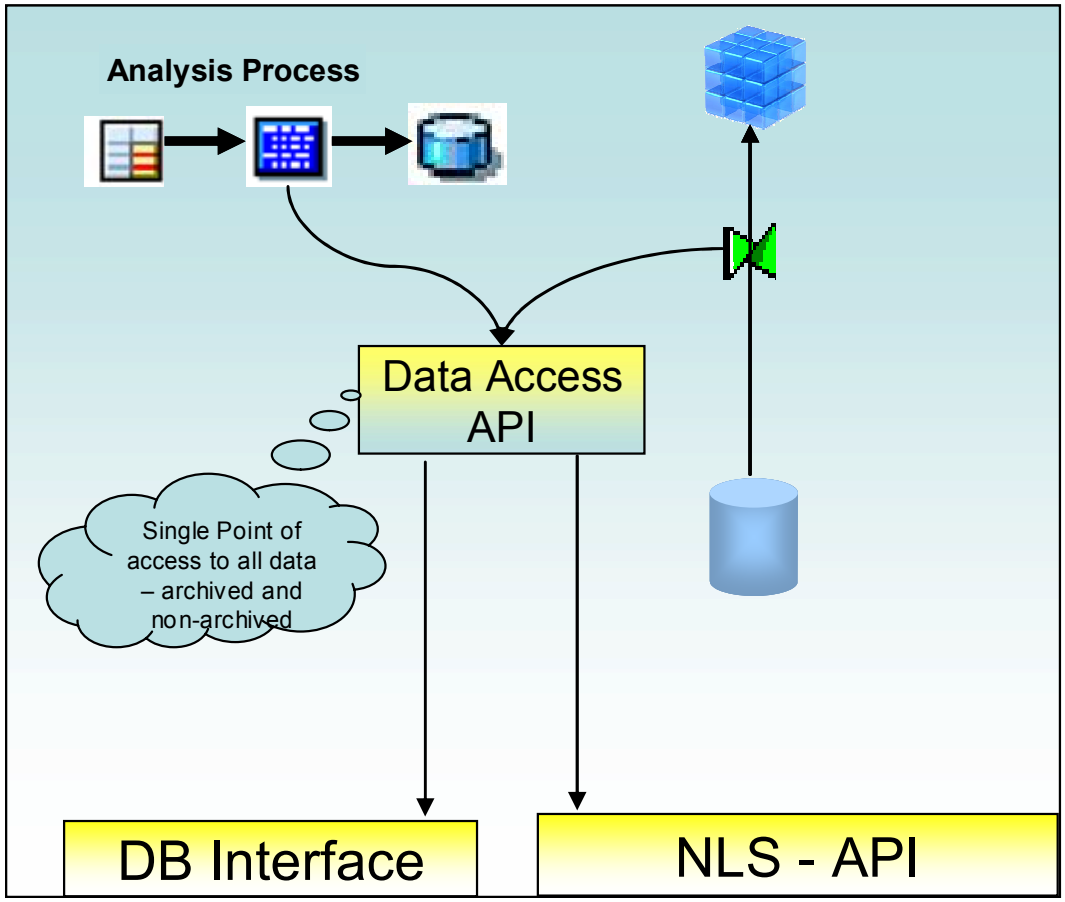


Volkswagen Bank

Typical Data Flow Requirements at Volkswagen Bank



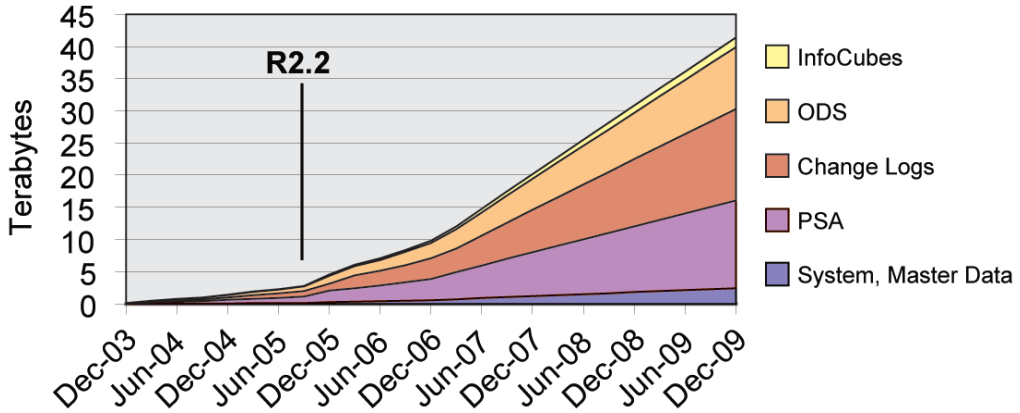
Requirement: “Look up” Functionality



US Government Case Study

- SAP BW database was growing at an unsustainable rate
- Limited funding for disk resources
- Performance risk

➔ *Data management strategy urgently required*

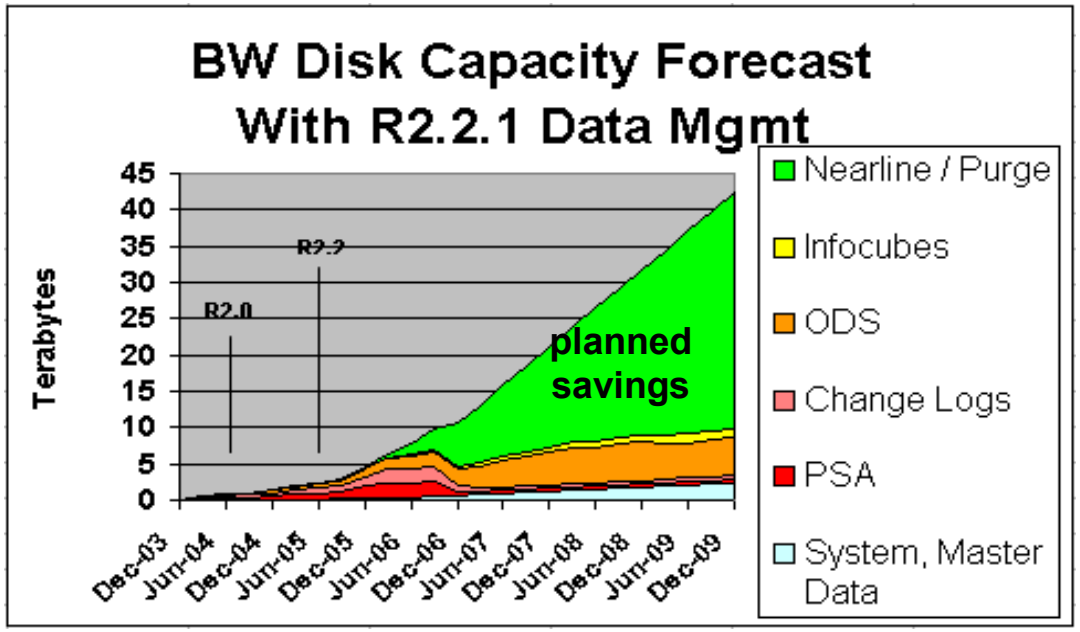


Database Growth:

- SAP BW database currently 5 TB (used space)
- Approximate growth rate at 400 GB/month
- Expected database size 10 TB by Dec 2006

US Government Case Study (cont.)

SAP BW Forecast with Data Management Strategy



Return on Investment

- Volkswagen Bank:
  - 90% data compression (on average), still available for use in reporting or as the basis for new DataStore objects or InfoCubes
  - Low total cost of ownership, due to the need for far less administrative support as compared with standard archiving solutions

- US Government:
  - ROI in less than 6 months (immediately after production go-live)
  - Savings of over 50% on related storage infrastructure thereafter
  - About 95% compression
  - Reduced data footprint eases replication/bandwidth issues

Best Practices

- Move your infrequently used data to nearline
- Have an SAP NetWeaver nearlining/ILM strategy for BI before you experience performance or maintenance issues
- Have a nearline strategy in place so you can react quickly to unplanned events such as audit, M&A, new business directions - and avoid penalties, lost revenue, and customer dissatisfaction.

Key Learnings

- You can lower your TCO and improve operational efficiencies with Nearline
- You can keep more data at your fingertips to respond to changing business needs, trend analysis, and regulatory compliance
- You can change/update data more easily when it's kept nearline (vs. archive)
- You can stop throwing away your data or choosing what data to keep as you upgrade - keep it all!

SAND/DNA supports and is certified with both SAP NetWeaver 3.1/3.5 and SAP NetWeaver 2004s

Thank you for participating.
Please remember to complete and return your evaluation form following this session.

For ongoing education on this area of focus, visit the Year-Round Community page at **www.asug.com/yrc**

ASUG®
AMERICAS' SAP USERS' GROUP

Session Code:
109

EDUCATING | NETWORKING | INFLUENCING