




# The HP Neoview data warehousing platform for business intelligence

Ronald Wulff  
EMEA, BI Solution Architect  
HP Software - Neoview




© 2006 Hewlett-Packard Development Company, L.P.  
The information contained herein is subject to change without notice




## Topics

- Neoview
  - Hardware technology and platform offerings
  - Software technology and ecosystem
  - Keys to excellence



2



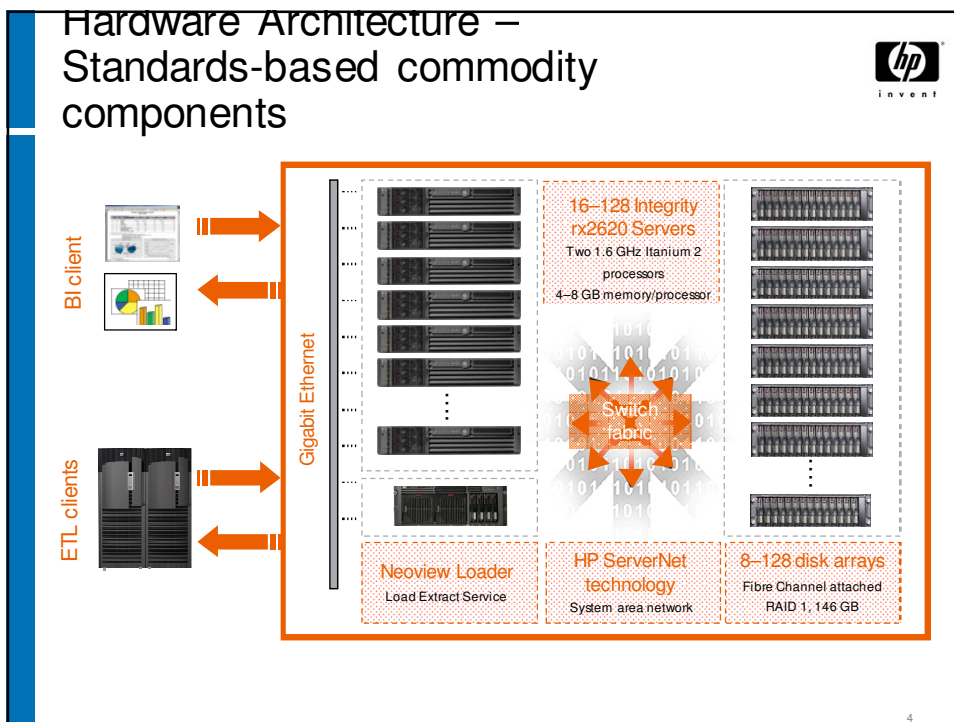
## The HP Neoview platform

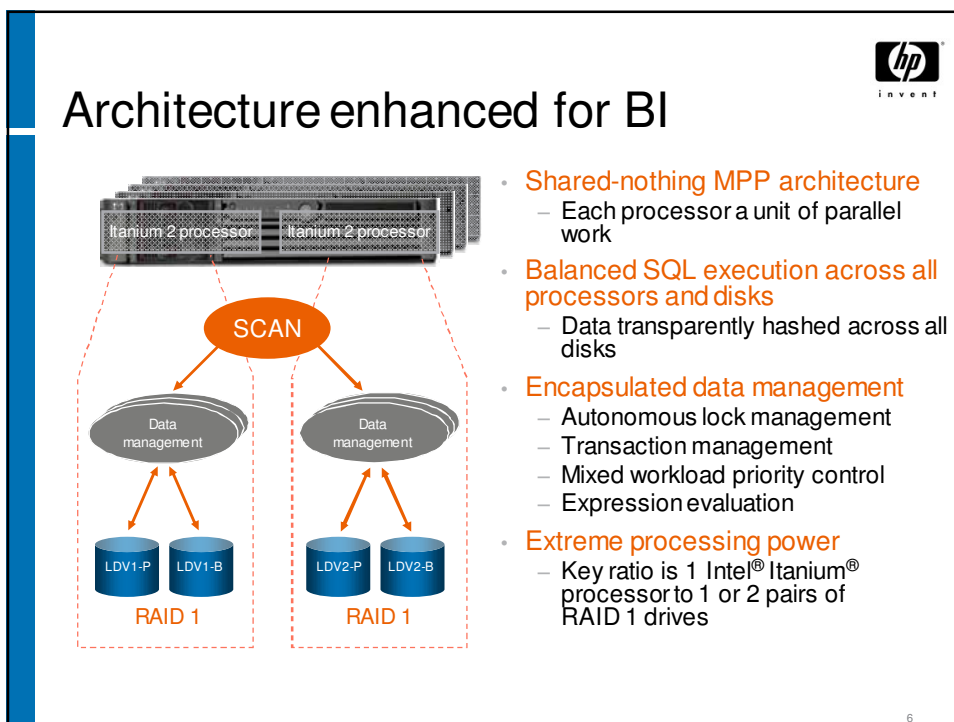
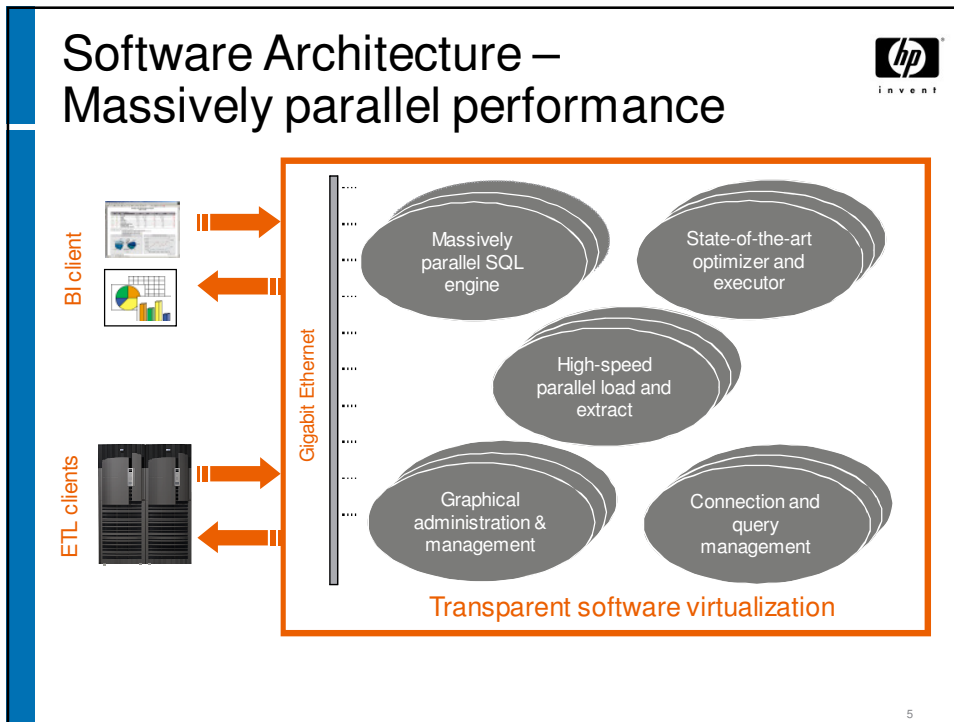
An innovative enterprise data warehouse

Enterprise class	Appliance-like simplicity	Affordable
<ul style="list-style-type: none"> <li>Highly available 24x7</li> <li>Readily scales 100s of processors</li> <li>High-performance massively parallel SQL database</li> <li>Handles complex queries and multiple users</li> </ul>	<ul style="list-style-type: none"> <li>Preconfigured, integrated, and tested (hardware and software)</li> <li>Little maintenance or database tuning</li> <li>Easy to incorporate into existing IT environments</li> <li>Remotely monitored and managed by HP</li> </ul>	<ul style="list-style-type: none"> <li>Up to 40% less than the cost of traditional offerings*</li> <li>Standard-based commodity components</li> <li>Reduced operating costs</li> <li>Reduced DBA costs</li> </ul> <p style="font-size: small; margin-top: 10px;">* Based on publicly available pricing data versus equivalent Teradata and IBM configurations.</p>


Surrounded by world-class HP Services

3





## Parallel UOW drives MPP performance




- **Measured scan rate**
  - 130+ MB/sec./LDV
  - Greater than 1 TB per minute\*
- **Measured active load rate**
  - 1+ MB/sec./processor
  - Greater than 20 TB per day\*
- **Measured extract rate**
  - 2.5+ MB/sec./processor
  - Greater than 50 TB per day\*

\* For largest configuration.

7

## More than simply parallel SQL




- State-of-the-art optimizer and executor
- Innovative SQL technology
  - Indexed data clustering
  - Mixed workload priority support
  - Fault-tolerant process pairing
  - Accelerated aggregations
  - Materialized views
  - Accelerated joins
  - Multidimensional access method
  - Logical RID secondary indexing
  - Co-located secondary indexing
  - Query analytic functions
  - Performance/usability extensions

**“Smarter” SQL parallelism**

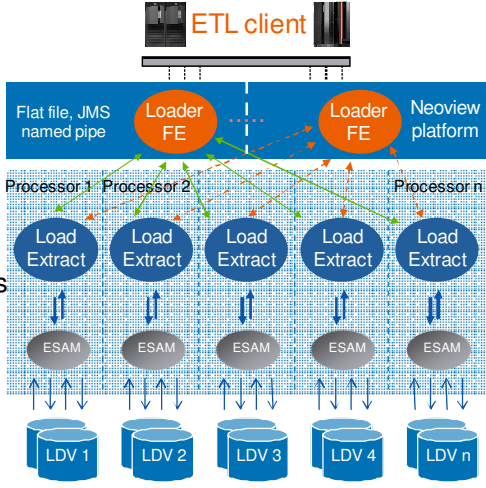
- Reduced sort, access, and scan I/O
- Reduced index dependencies
- Reduced aggregate dependencies

8




## High-speed load and extraction

- MPP facility for loading and extracting data
- Fully online and real time
  - Accelerated row inserts
- Supports inserts, updates, *upserts*, and *delserts*
- Multiple transport options
- Supports leading ETL products or client-built application
- Extreme performance
  - Hash partition awareness
  - Rowset efficiency
  - Non-atomic error recovery
  - Multiple loader support




The diagram illustrates the high-speed load and extraction architecture. At the top, an ETL client connects to a Neoview platform containing Loader Front Ends (FE). These loaders interface with a multi-processor (MPP) environment consisting of Processor 1, Processor 2, and Processor n. Each processor contains a Load Extract component and an ESAM (Extract, Sort, Aggregate, Merge) component. The ESAM components are connected to a distributed set of Load Data Volumes (LDV 1, LDV 2, LDV 3, LDV 4, LDV n). Data flow is shown as bidirectional between the loaders and processors, and between the processors and the LDVs.

9



## Easy to incorporate into existing environments

- A complete, preconfigured solution that can be rapidly deployed, easily managed, and is compatible with existing BI applications
- Preserves existing investments in tools, people, and processes



The diagram shows the Neoview platform acting as a central hub for data integration and query/reporting/analysis. On the left, under 'Data integration', are logos for Ascential Software, Business Objects, Informatica, and SAS. On the right, under 'Query, reporting, and analysis', are logos for Business Objects, Cognos, SAS, MicroStrategy, and Hyperion.

10

## hp invent

# Prioritized mixed workload support

### Prioritized SQL I/O

- User-/role-based assignment
- Application/query priority assigned to SQL I/O requests
- Prioritizes I/O requests for ESAM and processor execution
- Cluster-wide prioritization control

### Benefits

- Superior mixed workload support
- Service level agreement (SLA) fulfillment
- Allows concurrent maintenance and query (strategic and tactical) processing

The diagram illustrates a three-node system (LDV 1, LDV 2, LDV n). Each node has a queue, an ESAM (Enterprise Storage Access Method) component, a cache, and a Primary & RAID 1 pair of LDVs. A vertical dashed line separates LDV 2 from LDV n, labeled 'Processor or segment boundary'. Arrows show data flow from Operation P110, Query P150, and Query P180 to their respective queues. The queues are labeled with values: LDV 1 (110, 150, 180), LDV 2 (110), and LDV n (110, 150, 180). The ESAM and cache components are shown with bidirectional arrows, and the LDVs are shown with bidirectional arrows.

11


## hp invent

# Innovative cluster-aware optimization

- State-of-the-art optimizer (Cascades)
  - Top-down, branch, and bound
  - Cost-based and rules driven
  - Multipass optimization
  - "Equal-height" histogram statistics
  - Pushes predicates down as far as possible
  - Eliminates sorts when feasible, syntactically and semantically
- Parallel and nonparallel plan consideration
- Optimization for both complex queries and real-time workloads


The diagram shows a central 'Compiler & optimizer' box. It is connected to several optimization methods: Join methods, Aggregation methods, Access methods, Statistics (histograms), ESAM/ESP processing, and Parallel options. Below the optimizer is a 'Cost Vector' box containing: Processor cycles, Disk I/O, I/O transfers, Memory utilization, Message transfers, and Temporal storage. To the right of the Cost Vector is a 'Cost optimized access plan' box.

12



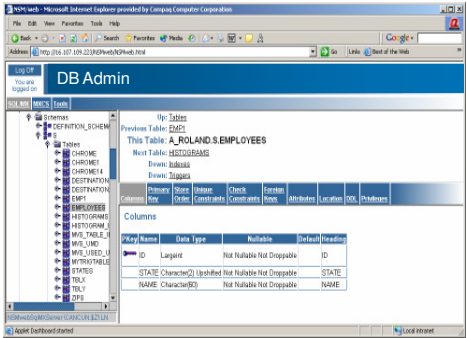
## Simplicity

Deployment	<ul style="list-style-type: none"> <li>• Delivered preconfigured, integrated, and tested</li> <li>• Compatible with existing BI infrastructure</li> <li>• Data model agnostic</li> <li>• No new learning or staffing requirements</li> </ul>
Administration	<ul style="list-style-type: none"> <li>• No table spaces to define and manage, no partition boundaries to define and tune</li> <li>• Indexes and materialized views are optional</li> <li>• Simple GUI tools, scripting interface, and automated administration services</li> </ul>
Management	<ul style="list-style-type: none"> <li>• Performance and workload dashboard</li> <li>• Self-managing subsystems</li> <li>• Secure remote monitoring, support, and incident analysis</li> </ul>



## Simplified administration

- Neoview DB Admin
  - SQL objects
  - Users and roles
  - ODBC/JDBC connections
  - Materialized views
  - Scheduling data services
- Neoview Script
  - Scripting and command line interface
- Neoview Dashboard
  - Performance and workload monitoring

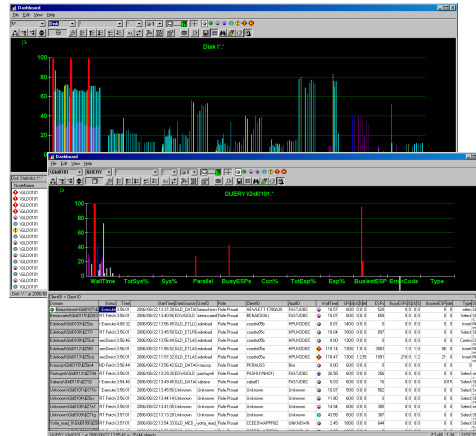


The screenshot shows the Neoview DB Admin web interface. On the left is a tree view of database schemas. The main area displays the definition for a table named 'EMPLOYEES' in the 'A\_ROLAND.S' schema. The table has three columns: 'ID' (Largeint), 'STATE' (Character(2)), and 'NAME' (Character(50)). Each column has a 'Not Nullable' and 'Not Droppable' attribute.

PK	Column	Data Type	Nullable	Default	PrimaryKey
	ID	Largeint	Not Nullable	Not Droppable	ID
	STATE	Character(2)	Not Nullable	Not Droppable	STATE
	NAME	Character(50)	Not Nullable	Not Droppable	NAME

## Powerful yet user-friendly performance monitoring and workload management

- Graphical display of alerts, reports, and dashboards
- View compile and query execution statistics
  - Query ID, text, explain plan, and estimated costs
  - Detailed query execution statistics after query completes
  - Gross process-level execution statistics for a running query
- Threshold-based notification
- Preemptive and runaway query control



15

## Secure remote support


- Instant Service Enterprise Edition (ISEE) is a standard HP remote support product
- ISEE uses a multilevel, layered security structure
  - Encryption
  - Authentication
  - Industry-standard security protocols
  - Best practices
- Single, secure access point

Example of ISEE security technologies	
<b>Privacy of data</b> <ul style="list-style-type: none"> <li>SSL with RSA and RC4 encryption ciphers for Advanced Configuration</li> <li>RC4 encryption ciphers for Standard Configuration</li> <li>IPsec VPN or SSH with 3DES encryption cipher</li> </ul>	<b>Authenticity of data</b> <ul style="list-style-type: none"> <li>SSL with X.509v3 certificate standard</li> <li>MD5 message digest with X.509v3 digital certificate standard</li> </ul>
<b>Integrity of data</b> <ul style="list-style-type: none"> <li>MD5 message digest with X.509v3 digital certificate standard</li> </ul>	<b>Authenticity of users</b> <ul style="list-style-type: none"> <li>PKI authentication (X.509v3 digital certificate standard)</li> <li>Windows NT domain authentication</li> </ul>

**Standard HP technology**


16






## Unrivaled availability

- Reliable and failure resilient hardware
- Patented fault-tolerant software “process-pair” technology
- Elimination of offline administration windows
- Platform is continuously available, in spite of any single point hardware or software failure
- Leverages 30 years of proven HP NonStop system engineering



17



## Keys to HP excellence

- Massively parallel performance and scalability
  - Shared-nothing combined with driver-level SQL efficiency
  - Scalable infrastructure virtualization
  - Balanced data distributions and processor utilization
- Fully “active” load and extract facility for extreme workloads
- Robust native connectivity for industry-standard BI products

plus

- Innovative SQL technology drives breakthrough performance for mixed and concurrent workloads
- Engineered appliance-like simplicity
- “Always on” hardware and software technology

18