NSS Enriching Knowledge for Information and Communication Technology Curriculum Series - (2) ICT Development and Applications in Hong Kong: Database Security & ICT Trend

10 Jan 2008

IT Trend Briefing :

Virtual, Scalable, Distributed Data Provisioning (Grid Computing)

ORACLE

Benefits of Grid Computing

- Better information faster, easier
  - Perform more work with fewer resources
  - Spread work across resources
  - Access to resources on demand
- Faster response to changing business priorities
  - Instantly and dynamically realign IT resources as business needs change
- Reduced IT costs
  - Improve utilization of existing resources
- All these benefits are realized when resources are
  - Virtual
  - Scalable
  - Distributed
  - Cost-effective

Virtual Computing Resources

- The Grid enables virtualization of resources across servers, data centers, and enterprises
- Types of virtual resources
  - Network
  - Storage
  - CPU
  - Data
- The Grid's complexity should be hidden from the applications
- To the end-users, they just "plug-in"

Grid Computing Resources must be Virtual, Scalable, Distributed

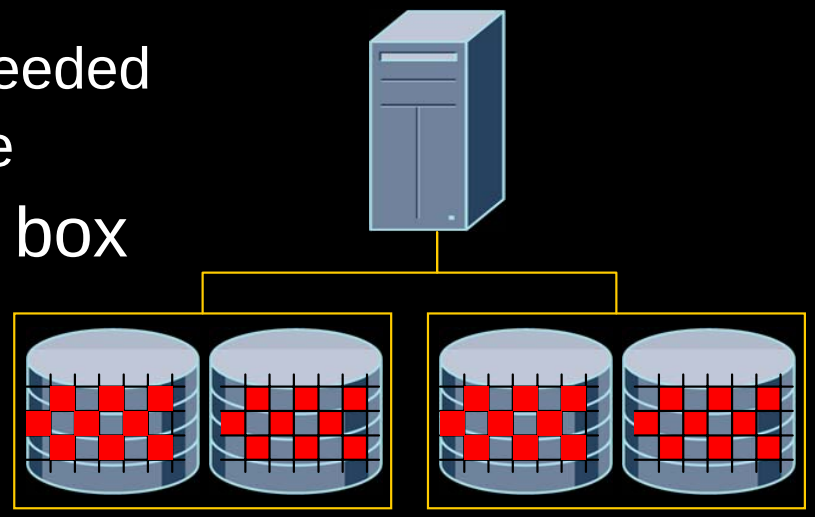
- Storage
- Database Servers
- Application Servers
- Provisioning and Management Tools

Align Storage with Business

- Islands of data
  - “My storage is 30% utilized and growing 50% a year”
- Disk farms of industry standard disks
  - Consolidate into SAN or NAS
  - Provision as needed

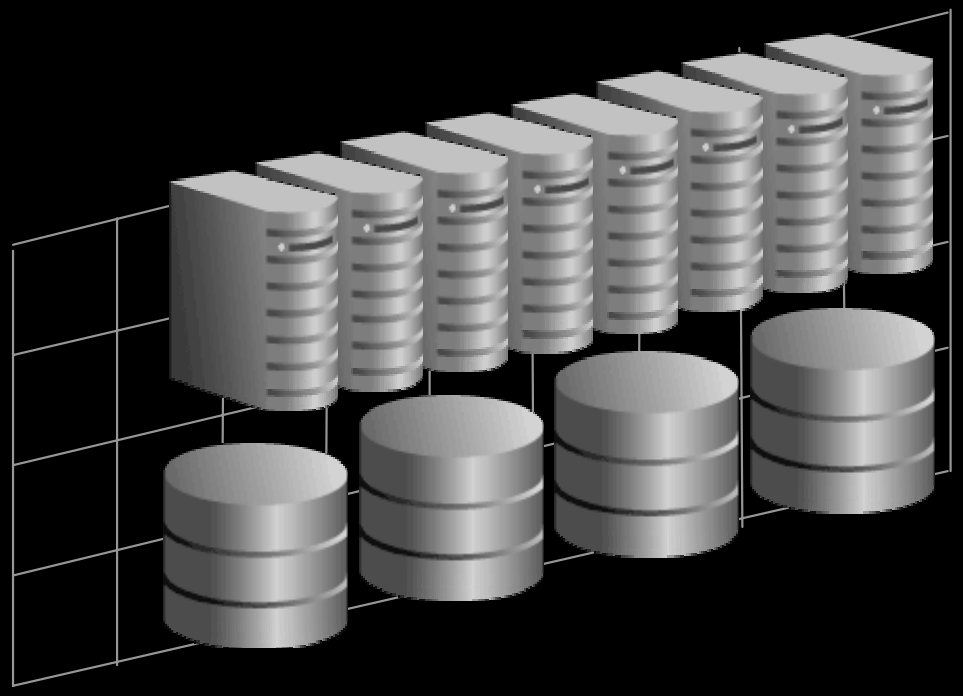
Oracle ASM (Automatic Storage Manager)

- Use groups of industry standard storage
- Automatically mirror and stripe data, immune to disk failure
- Manage all storage needs for the databases
  - Provisions storage capacity automatically to Oracle 10*g* as needed
  - no volume manager or file system needed
  - efficiently add/remove storage online
- Highest performance I/O out of the box
- Eliminate constant I/O tuning by load balancing to avoid hot spots



Align Processing with the Business

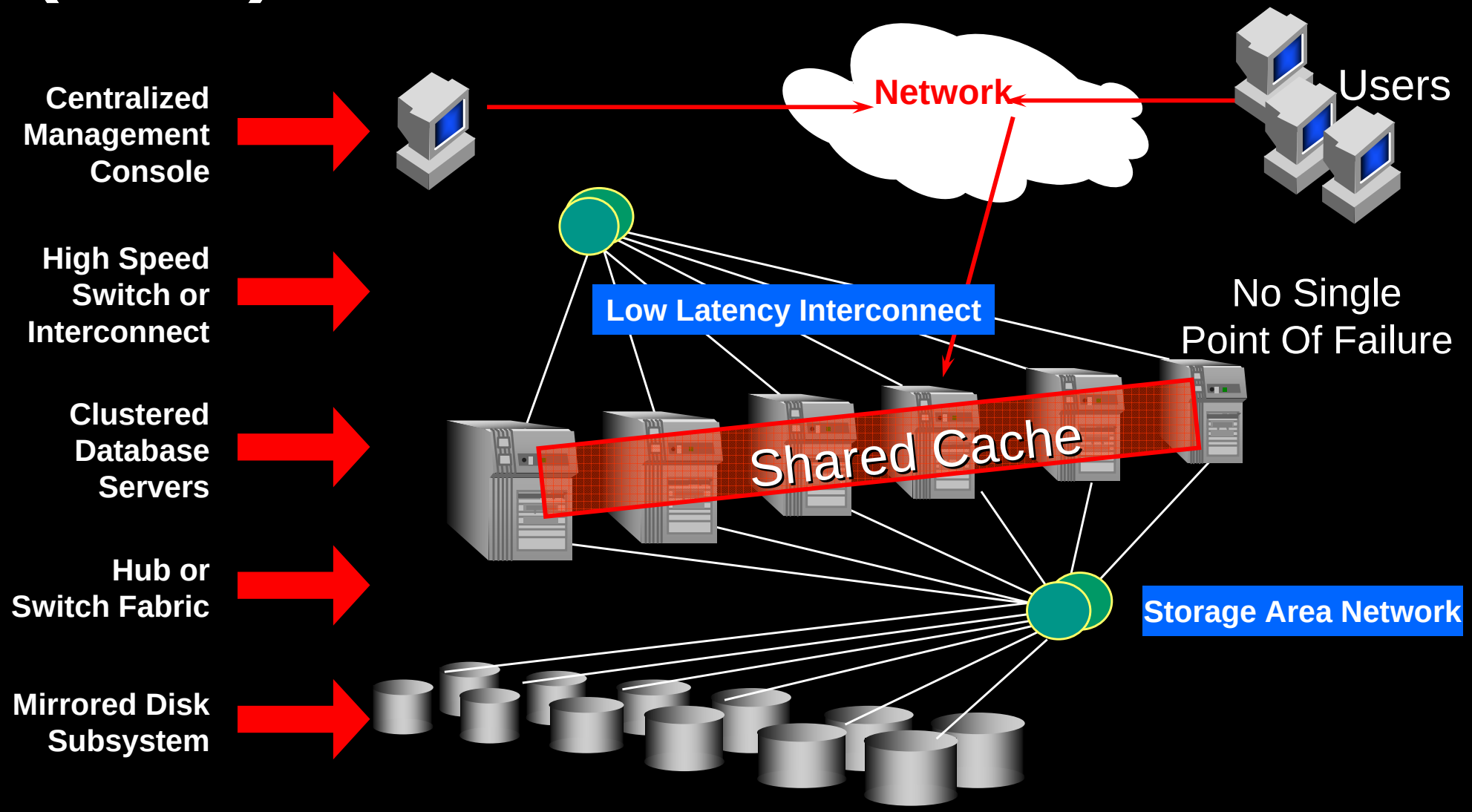
- Islands of computation
  - "15% utilization of CPU is exceptional"
- Standardize resources
  - Blades provide lowest cost, highest performance
  - Not Self-healing, Disposable
- Share virtual resources
- Provision resources as required
- Scale out



Databases on the Grid

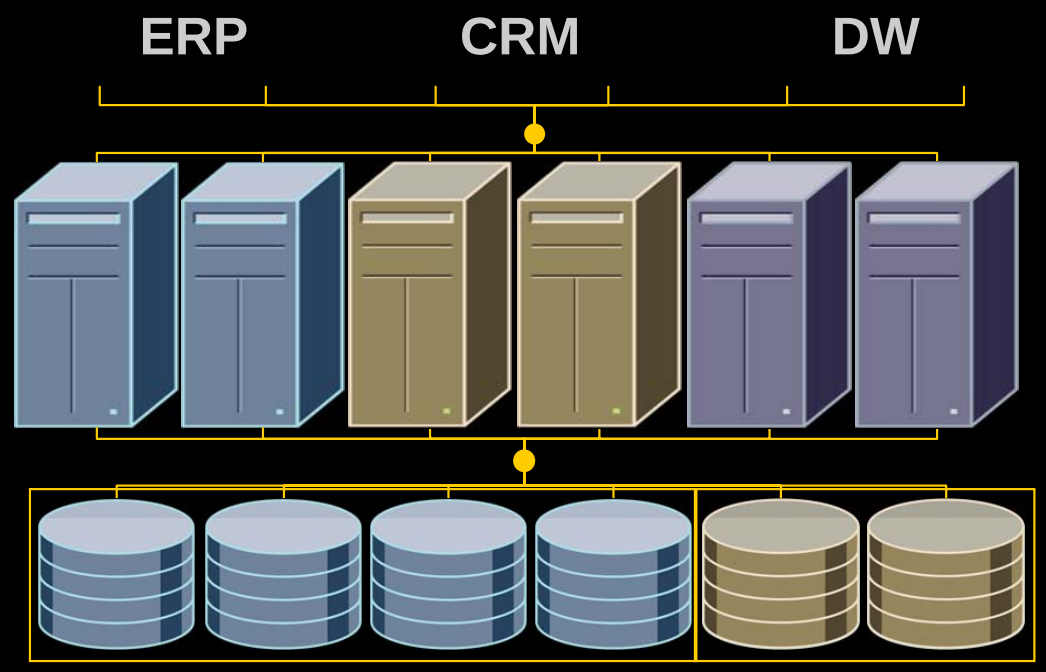
- Database clustering with shared disk
  - Low cost
  - highest quality of service
  - Scalability AND availability
- Add/drop servers as needs change
- Automatically balance load across servers
  - Automated routing of service requests to least loaded instance
  - If a server fails, re-allocate to surviving servers
  - If capacity is added, re-allocate to new servers

Real Application Clusters (RAC)



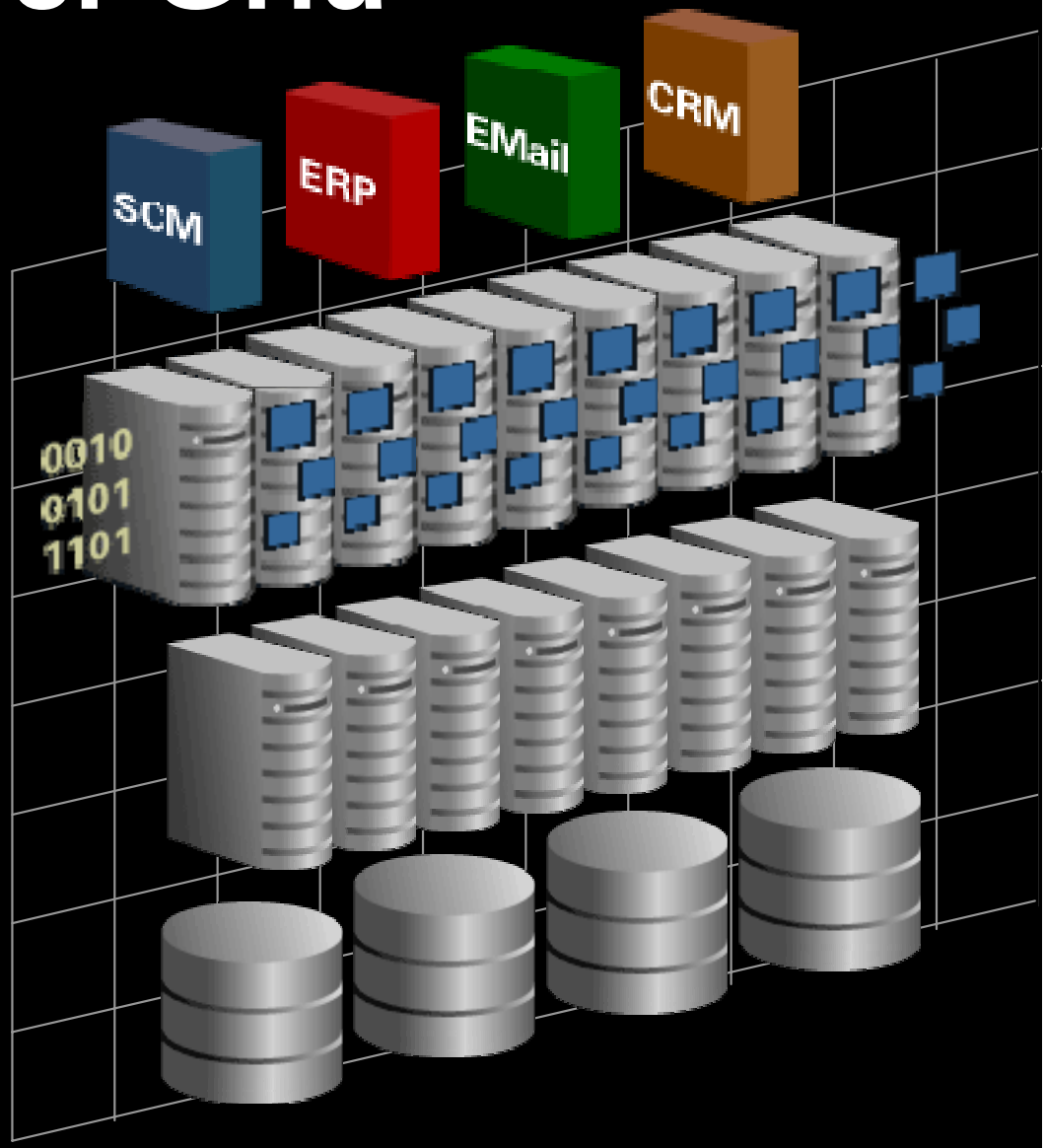
Oracle Database 10*g* – Real Application Clusters

- Improved Robustness
  - Cluster Verification Utility
  - Improved OCR, Voting Disk resilience
- Clusterware API for integrated application availability
- Connection Pool load balancing
- Certified to 100 nodes
- Improved performance



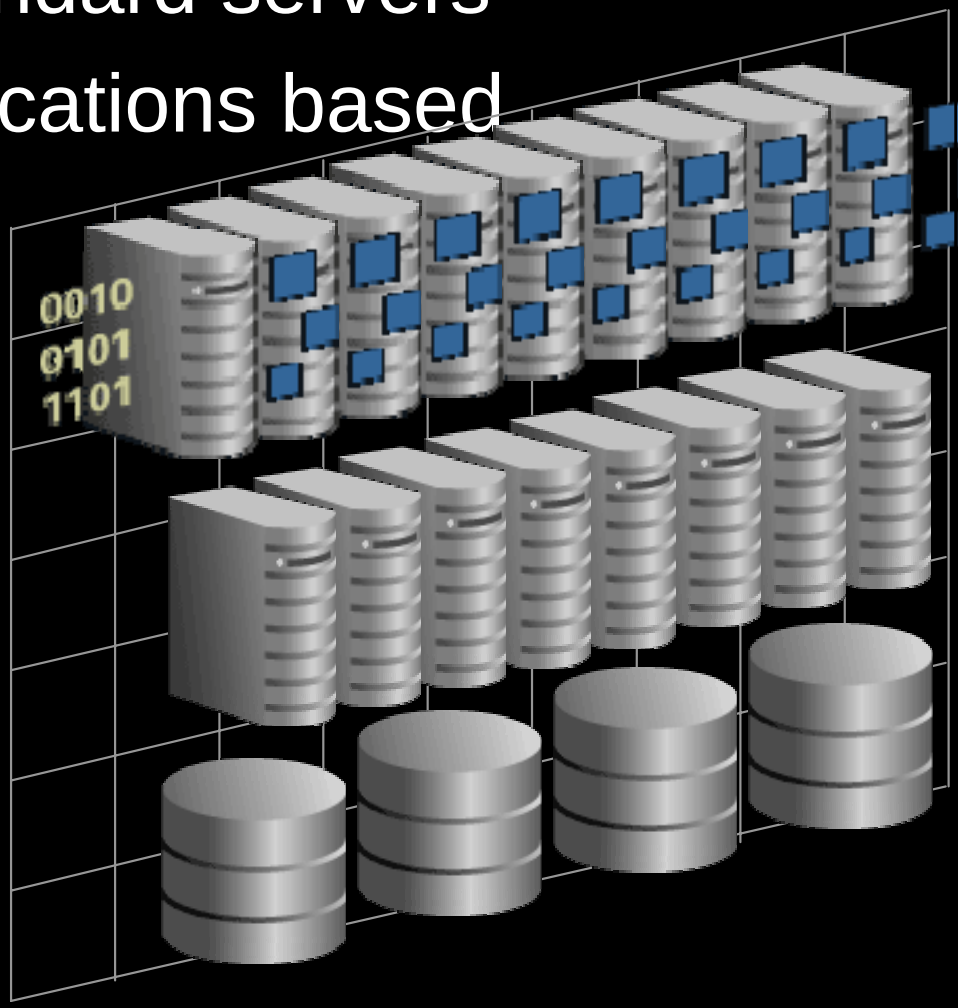
Application Server Grid

- Complete, integrated application server clusters
- End-to-end transparent application fail-over
  - Fast fault recovery in seconds
- Application-specific load balancing policies
  - Schedules
  - Runtime metrics



Oracle Fusion Middleware Application Servers

- Use groups of industry standard servers
- Allocate resources to applications based on business needs
- Get highest availability
- Dynamically add / remove servers



Virtual Computing Resources

- The Grid enables virtualization of resources across servers, data centers, and enterprises
- Types of virtual resources
  - Network
  - Storage
  - CPU
  - Data
- “Information integration” key to unlocking the **Virtual, Distributed** data resources on the Grid
- Efficient communication facility helps **Scaling** the moving of large amounts of data and network traffic

Information Integration

- Provides a unified view of all information regardless of
  - Data type
  - Data store
  - Physical location
- Provides access to data when and where you need it
  - Optimizes information access regardless of physical location of data
  - Enables virtualization of data in Grids

Methods of Information Integration

- Consolidate Information
- Federate Information
- Share Information

Consolidate Information

- Consolidate heterogeneous data into a central database
  - Any data (audio, video, XML, email, messages, etc…)
  - Any platform (Linux, Windows, Solaris, HPUX, AIX, Tru64, OpenVMS, OS/390)
  - Proven scalability
    - Hundreds of Very Large Database customers
    - Industry leading benchmarks

Benefits of Consolidation

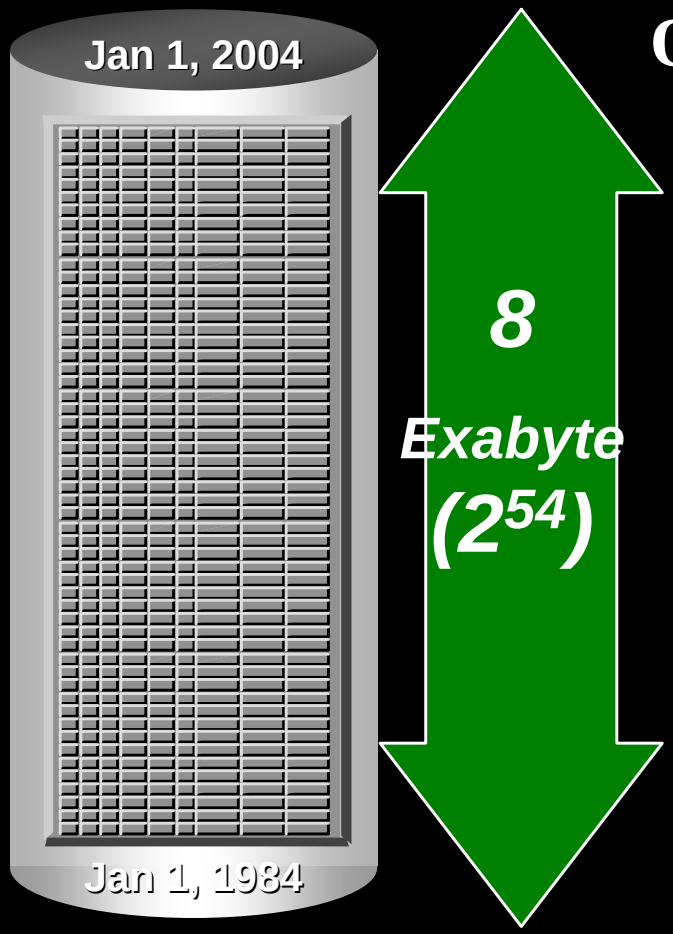
- Simplest form of Information Integration
- Lower ongoing administration costs
  - Fewer servers and databases to manage
- Faster application deployment
  - Deploy application once, at central server
- Faster access to global data
  - Queries and DML run faster against centralized data

Massive database support

- OLTP
  - CATalog database, CUSTomer database, and ACB database (order, shipment, and payment information) – 4 node clusters
- DW
  - Query
    - 16 nodes, 23 TB DB (68 TB disk)
    - 50,000 queries/week (mostly complex)
  - Historic Clickstream:
    - 4 nodes, 17 TB DB (42 TB disk)
    - 3-15+ month old clickstream data stored
  - ETL:
    - 8 nodes, 7 TB DB (36 TB disk)
- Migration from SMP to HP/Linux/RAC saved $6M ($2.5M vs $8.5M) in capital costs alone

Tackle VLDB data volume — Oracle *10g* Partitioning

Large Data Volume



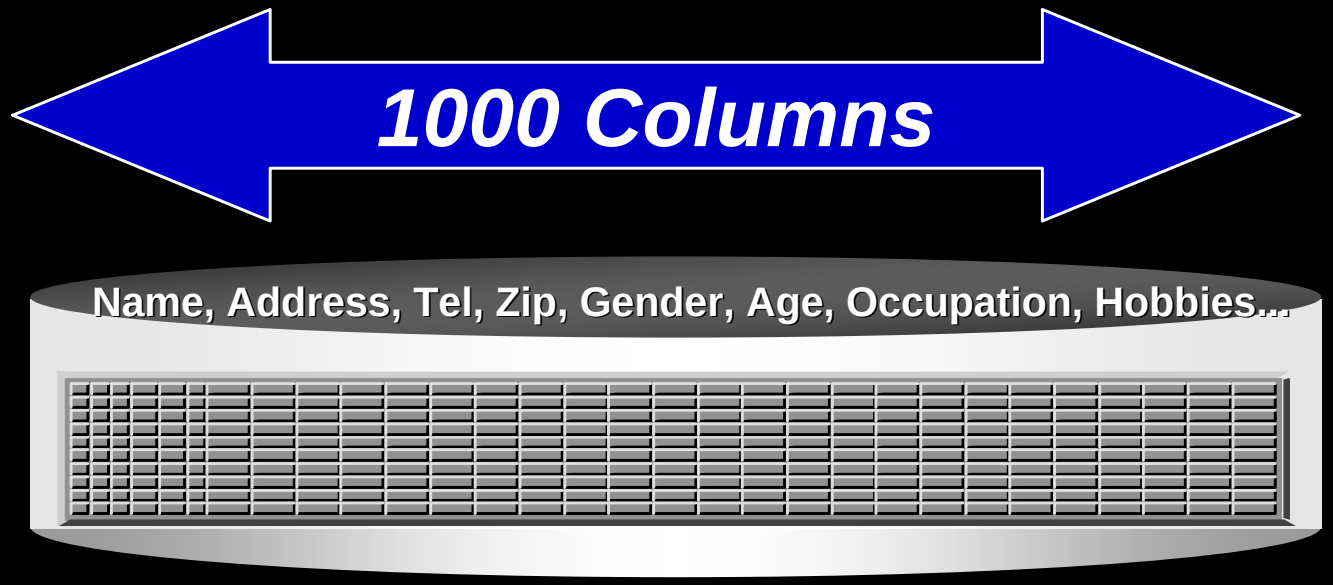
Oracle 10*g* 8E

Oracle 9*i* 512P

1 E = 1024 P
1 P = 1024 T
1 T = 1024 G
1 G = 1024 M

Lots of Conc. Users

High Complexity

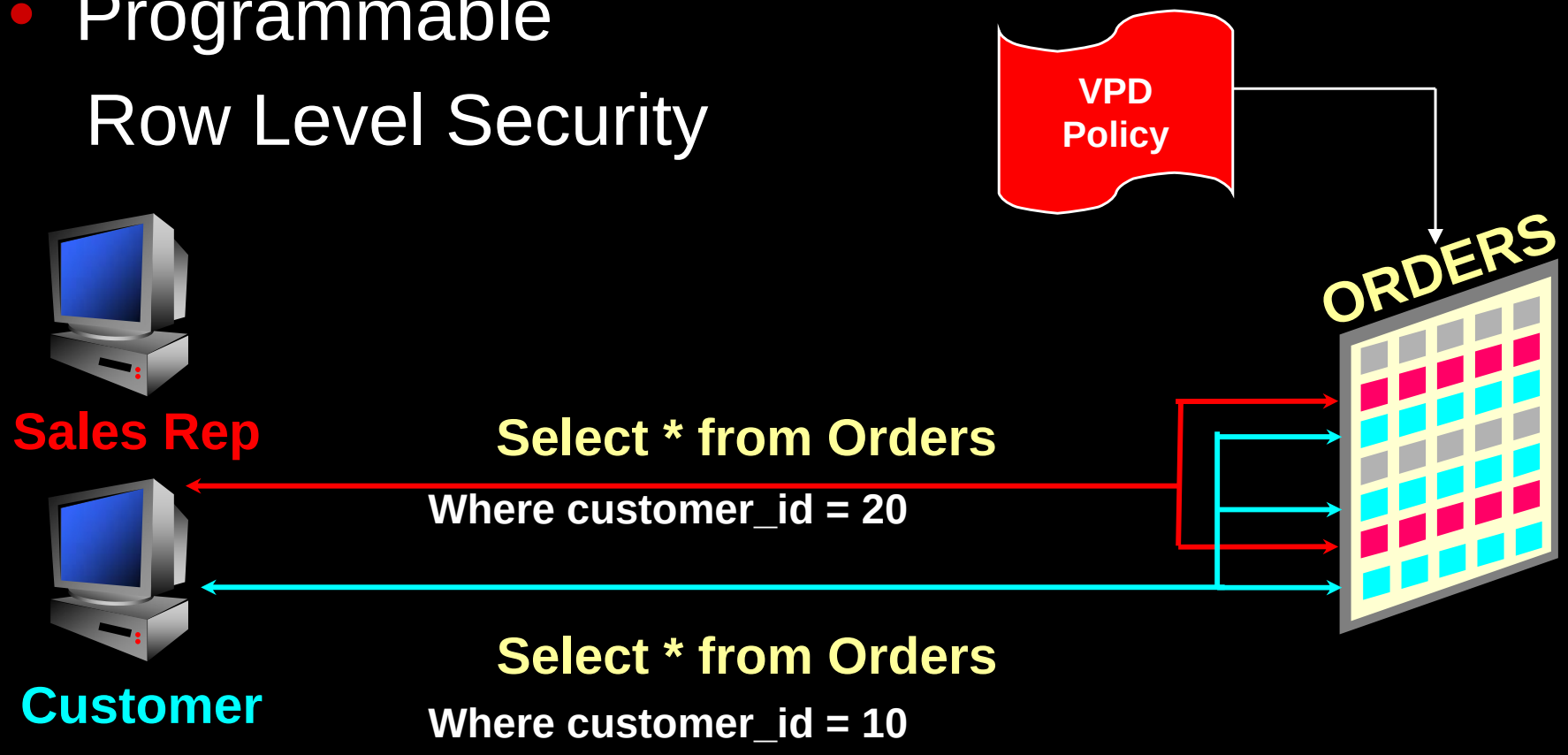


Virtual Private Database

Oracle Database 10g Virtual Private Database

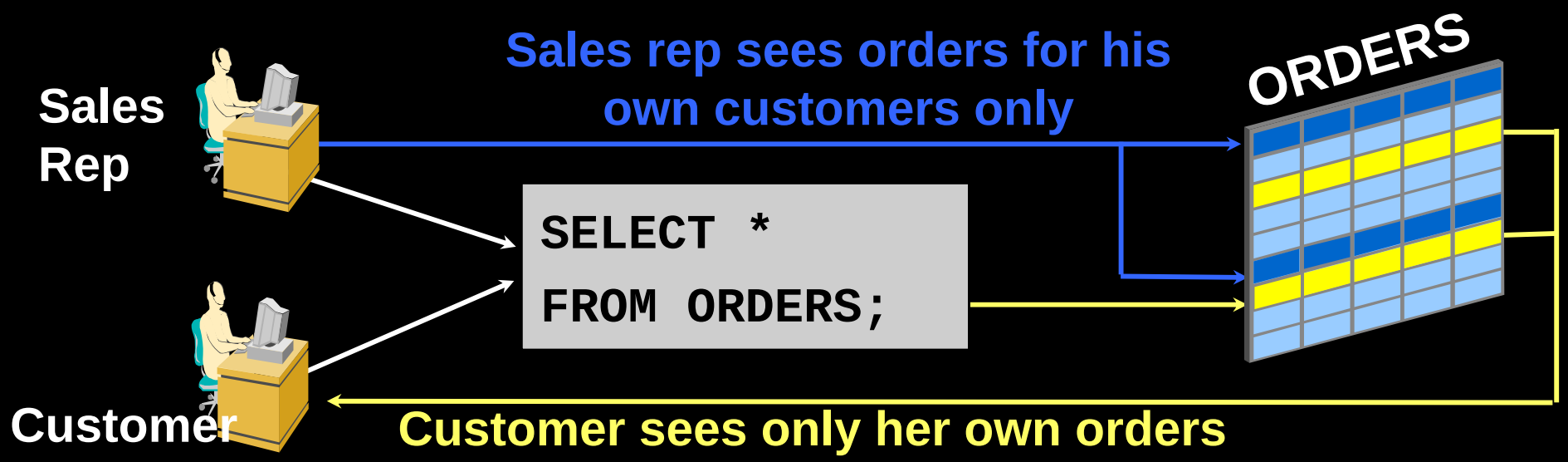
- Database enforced
- Programmable

Row Level Security



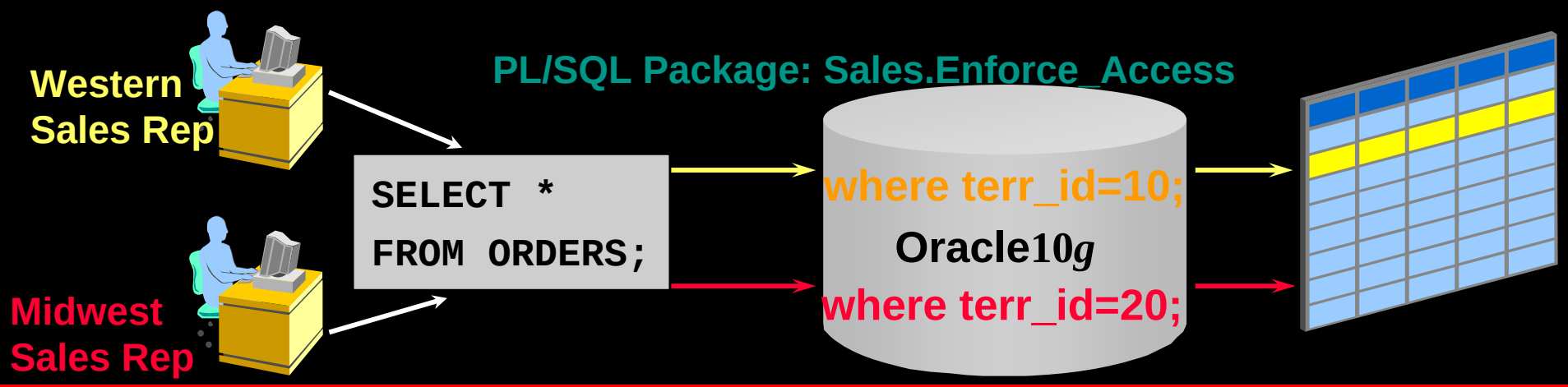
Virtual Private Database

- Data access is managed at the database level
  - Fine-Grained Access Control: Enforced at server
  - Application Context: Determines access control condition
- Users only see data that they have access to
  - Conditions can differ by users

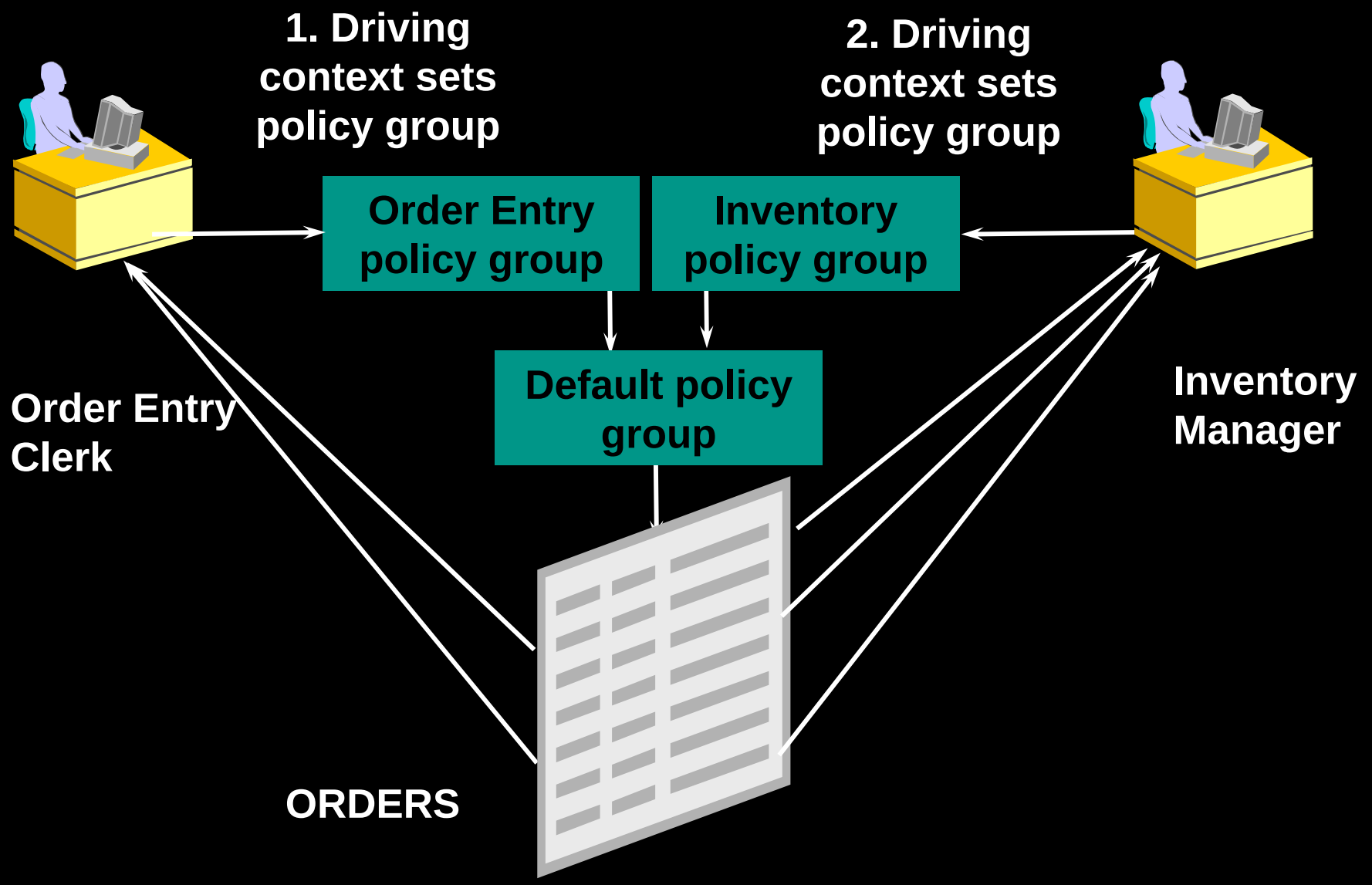


Virtual Private Database

- Add policy to any table, view, or (now) synonym
  - Application administrator binds PL/SQL package with table using Oracle supplied API (dbms_rls)
- Dynamically rewrite SQL
  - Query modification based on PL/SQL package assigned to a table, view or synonym
  - “Where” clause appended to SQL statement



Partitioned Fine-grained Access Control



Oracle Database 10g Virtual Private Database

- Column Relevant Policies
  - Policy enforced only if specific columns are referenced
  - Increases row level security granularity

Select store_id, revenue…
(enforce)

Store ID	Revenue	Department	
AX703	10200.34	Finance	✗
B789C	8020.34	Engineering	OK
JFS845	12341.34	Legal	✗
SF78SD	13243.34	HR	✗

Oracle Database 10g Virtual Private Database

- Column Filtering
  - Optional VPD configuration to return all rows but filter out column values in rows which don't meet criteria

Select revenue…..(enforce)

Store ID	Revenue	Department	
AX703		Finance	OK
B789C	8020.34	Engineering	OK
JFS845		Legal	OK
SF78SD		HR	OK

Oracle Database 10g Virtual Private Database

- Performance enhancements
  - Static Polices
  - Dynamic Policies
- Static Polices
  - Good choice for hosting when policy doesn't change
- Dynamic Policies
  - Good choice when policy returns different predicates
    - Time of day enforcement

Methods of Information Integration

- Consolidate Information
- Federate Information
- Share Information

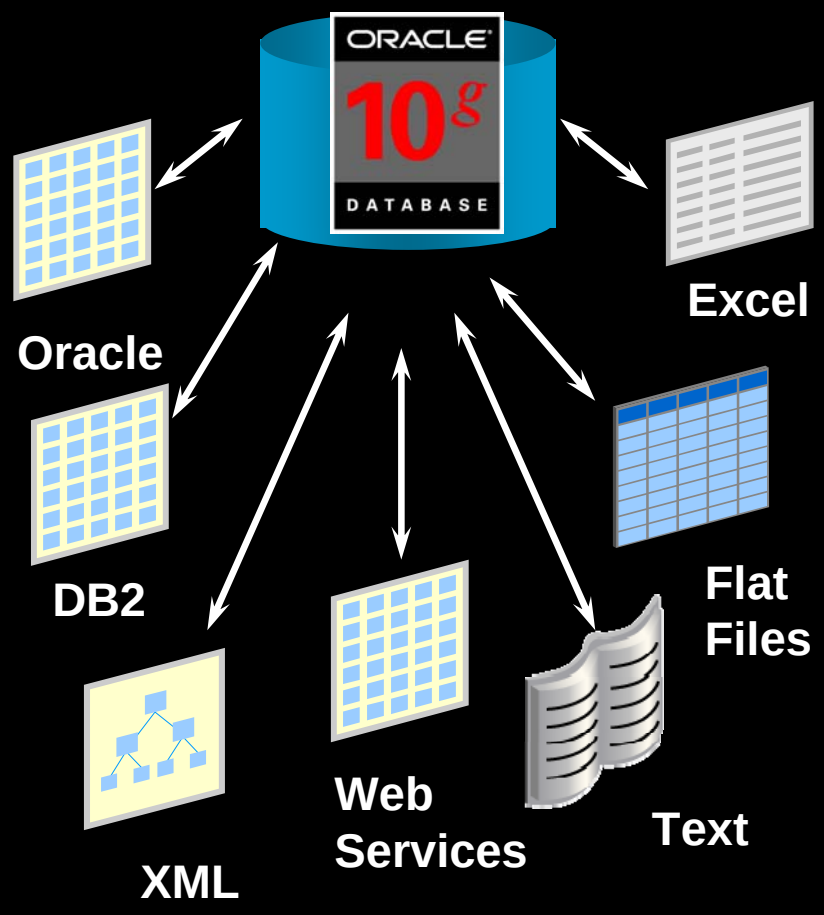
Federate Information

- Federate data in multiple data stores into a single virtual database
  - Access both Oracle and non-Oracle data
  - Access both structured and unstructured data
  - Hide physical location of data from applications

Benefits of Federation

- Fast integration
  - Simple changes yield immediate results
- Supports integration of data that cannot be consolidated
  - Legacy applications
  - Data requiring local ownership
- Ad Hoc Integration to infrequently accessed data

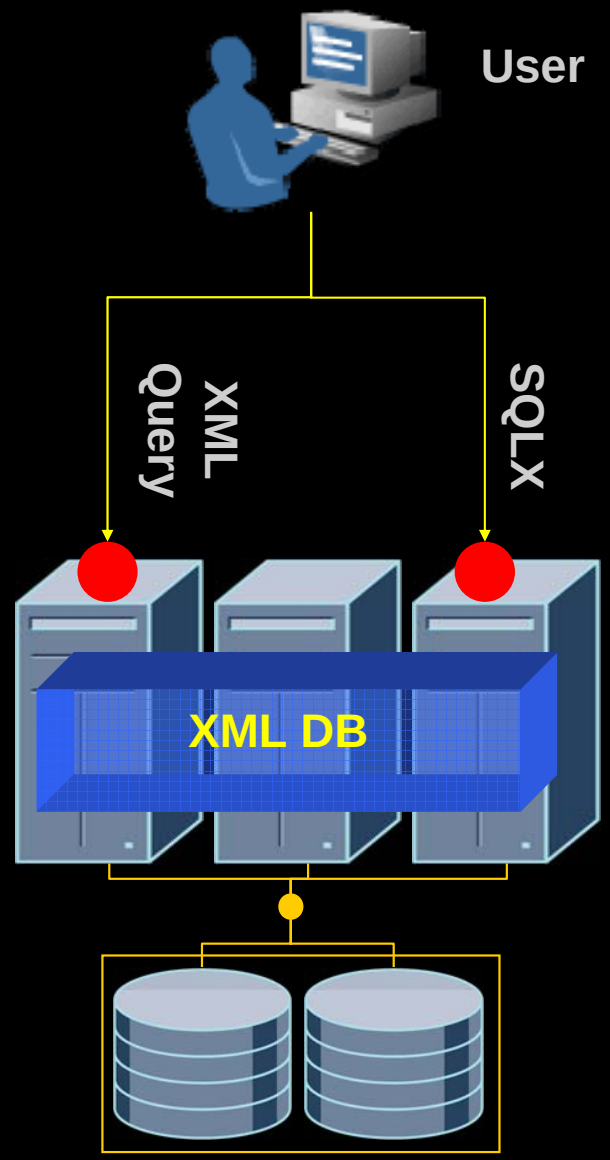
Features



- Multiple access methods
  - Distributed SQL
  - XQuery
- Any data source
  - Oracle and non-Oracle data stores
  - External files
  - Web services
- Advanced indexing technology

Access Data with XML Query

- Native XML Query support in the Database
- XML Query is emerging as the consensus query language for XML
- First mainstream commercial database release to support XML Query
  - Native = Fast, Efficient
- Can mix-and-match XML Query and SQL in same query



Access Oracle and non-Oracle Data Stores

- Heterogeneous data integration built into the Oracle database
  - SQL Translation
  - Data Dictionary Translation
  - Procedure execution on non-Oracle system
- Transparent Gateways
  - Available for most major non-Oracle data stores (e.g., DB2, SQL Server, Sybase, etc.)
  - Optimized, certified, end-to-end support
- Generic Connectivity
  - Gateway to any 3rd-party ODBC or OLEDB driver

Access External Files

- Access native XML Files
  - Query via XQuery or SQL
- Access proprietary file types via filters
  - Transform into XML files
  - Over 100 file types supported by supplied filters
- Access external tables
  - Map data in external files into tables

Methods of Information Integration

- Consolidate Information
- Federate Information
- Share Information

Share Information

- Share information between users, applications, and databases
  - Move or copy information as needed
  - Traditionally implemented as replication or message queuing
  - Has evolved to include warehouse loading, event notification, workflow, and EAI

Benefits of Sharing

- Greater flexibility
  - Locate data where you want it
  - Control access
  - Share information with business partners
- Faster access to local data
  - Queries and DML run faster against smaller, local data store
- Higher Availability
  - Applications not dependent on availability of remote data stores

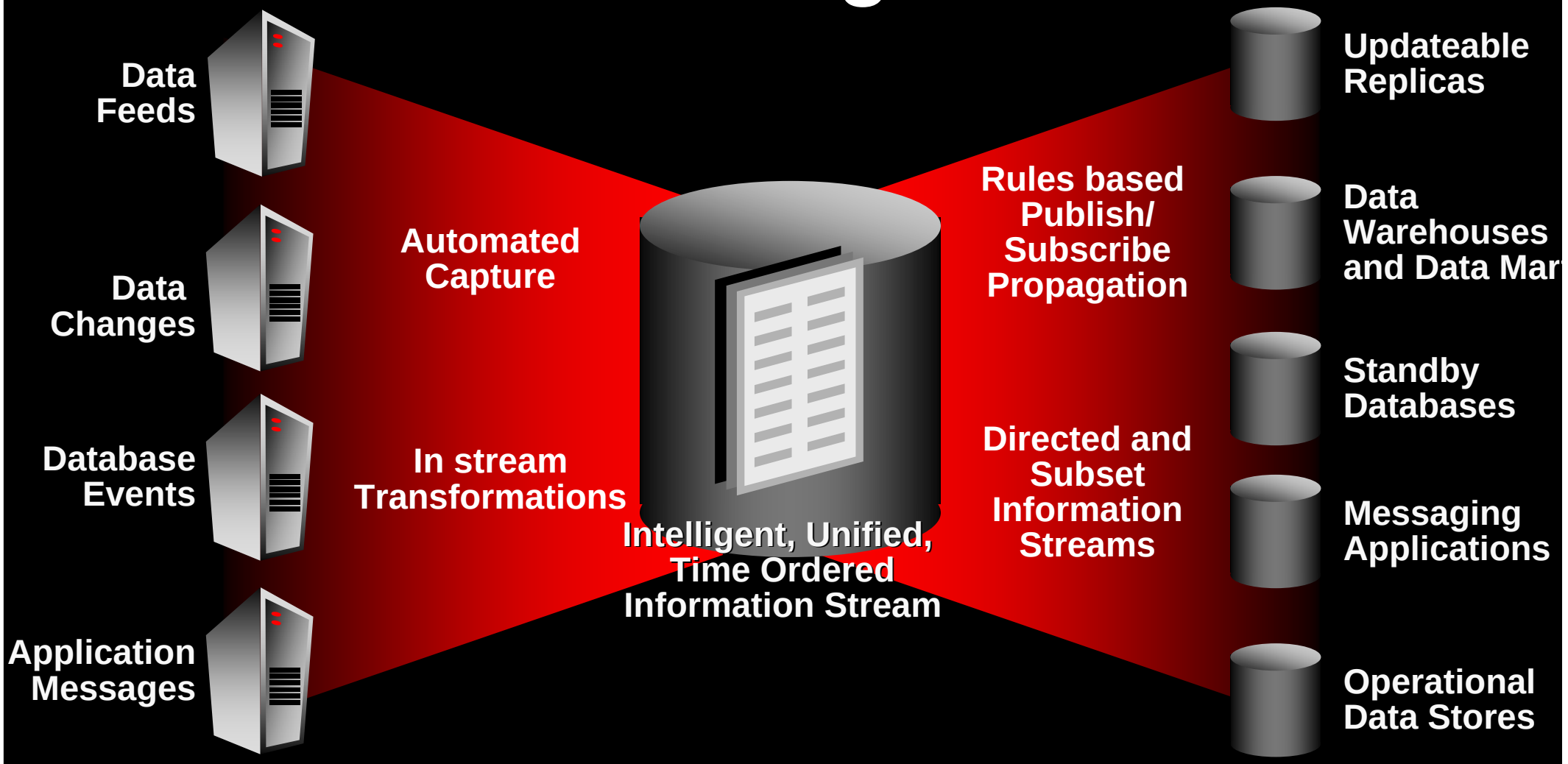
Information Sharing Features

- Streams
  - Message Queuing
  - Replication
  - Workflow
  - Much more
- Materialized Views
- Transportable Tablespaces
- Data Pump
- File Group and Tablespace Repositories

Eliminate Complexity with Oracle Streams

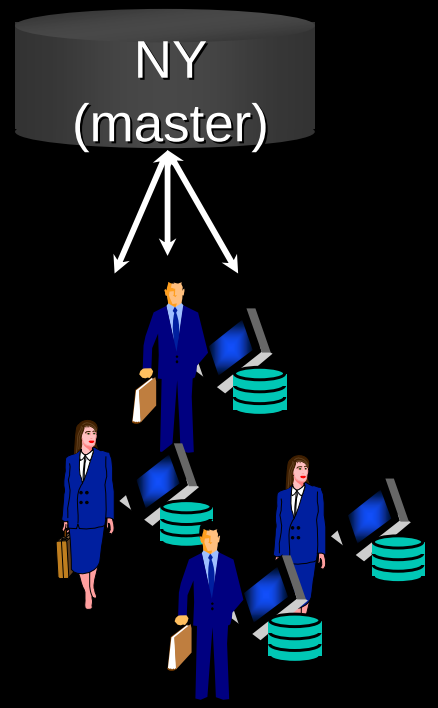
- Unifies all enterprise information into a single Stream
  - Unifies database, messaging, replication, publish/subscribe APIs and capabilities
- Avoids inherent problems with multiple technologies:
  - Multiple development models
  - Multiple operational models
  - Builders do the integration
- Not just a unified branding of different features

Unified Infrastructure for Information Sharing



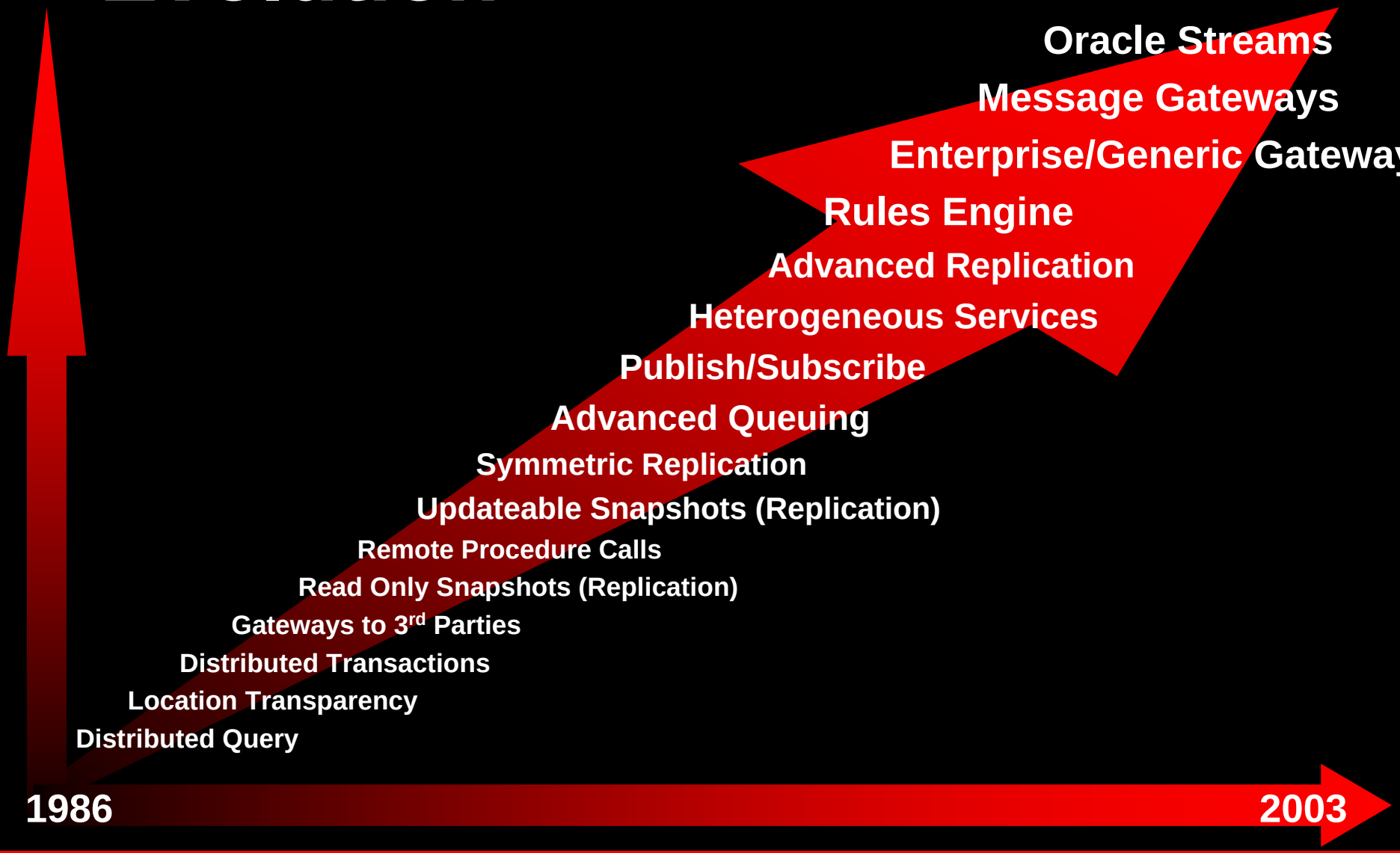
Share Data with Materialized Views

- Data warehousing
  - Compute and store aggregated data, such as sums or averages, for faster queries
- Mobile computing
  - Easy mass deployment
  - Disconnected computing
- Replication
  - Incrementally maintain local subset of data
  - Read-only or updateable with conflict resolution



Updateable Materialized Views

Oracle Information Integration Evolution



Summary

- Information Integration methods :
  - Consolidate Information—Scalable access to large amounts of information in single physical database
  - Federate Information—Data can be located where needed but accessed as if it is a single database
  - Share Information—Multiple users and applications can share data/information kept separately by different technologies more effectively
- Information Integration & Grid-ready software infrastructure critical to Grid Computing success