



IBM Software Group

## Power of zSeries and DB2 Together



John J. Campbell  
Distinguished Engineer  
DB2 for z/OS Development  
eMail: [CampbelJ@uk.ibm.com](mailto:CampbelJ@uk.ibm.com)

© 2005 IBM Corporation

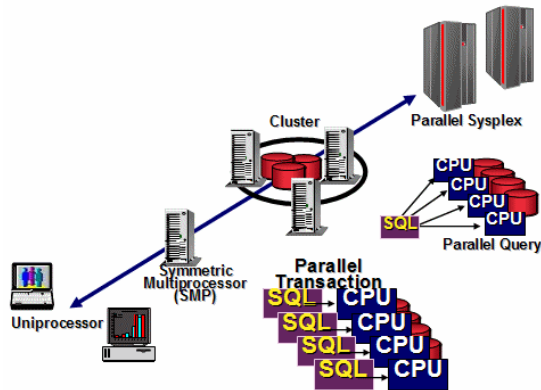
## Vision

Maintain and extend DB2 for z/OS leadership as the industry's premier OLTP database engine offering the industry's best scalability, performance, availability, and reliability across all classes of OLTP and mixed workloads:

- ▶ Traditional z/OS CICS and IMS applications
- ▶ ERP, CRM, SCM (SAP, PeopleSoft, Siebel, etc)
- ▶ Java servlets and EJBs (WebSphere, WebLogic, etc)
- ▶ Database-centric workloads using stored procedures
- ▶ Database server for n-tier architectures using transaction monitors



# DB2 for z/OS Areas of Investment



## Scalability:

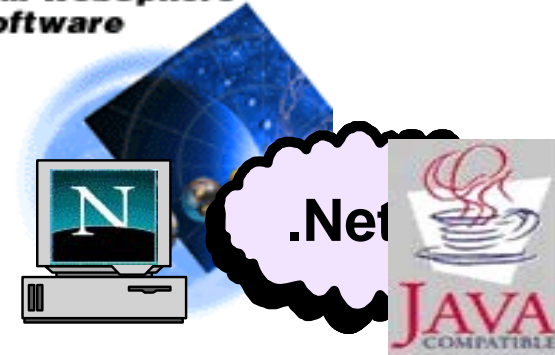
Users, Data, Queries, Transactions



## Availability:

Planned (reorg) & Unplanned (restart), Utilities & Administration

IBM WebSphere Software



## Access: Web & DRDA

Java, DB2 Connect, .Net, WebSphere, Information Integration, Replication



## Extensibility:

New data & function for e-business



## Management:

Java Control Center, Visual Tools, DB2 Tools



## Applications:

Vendors & Development, SQL & Object Portability



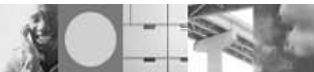
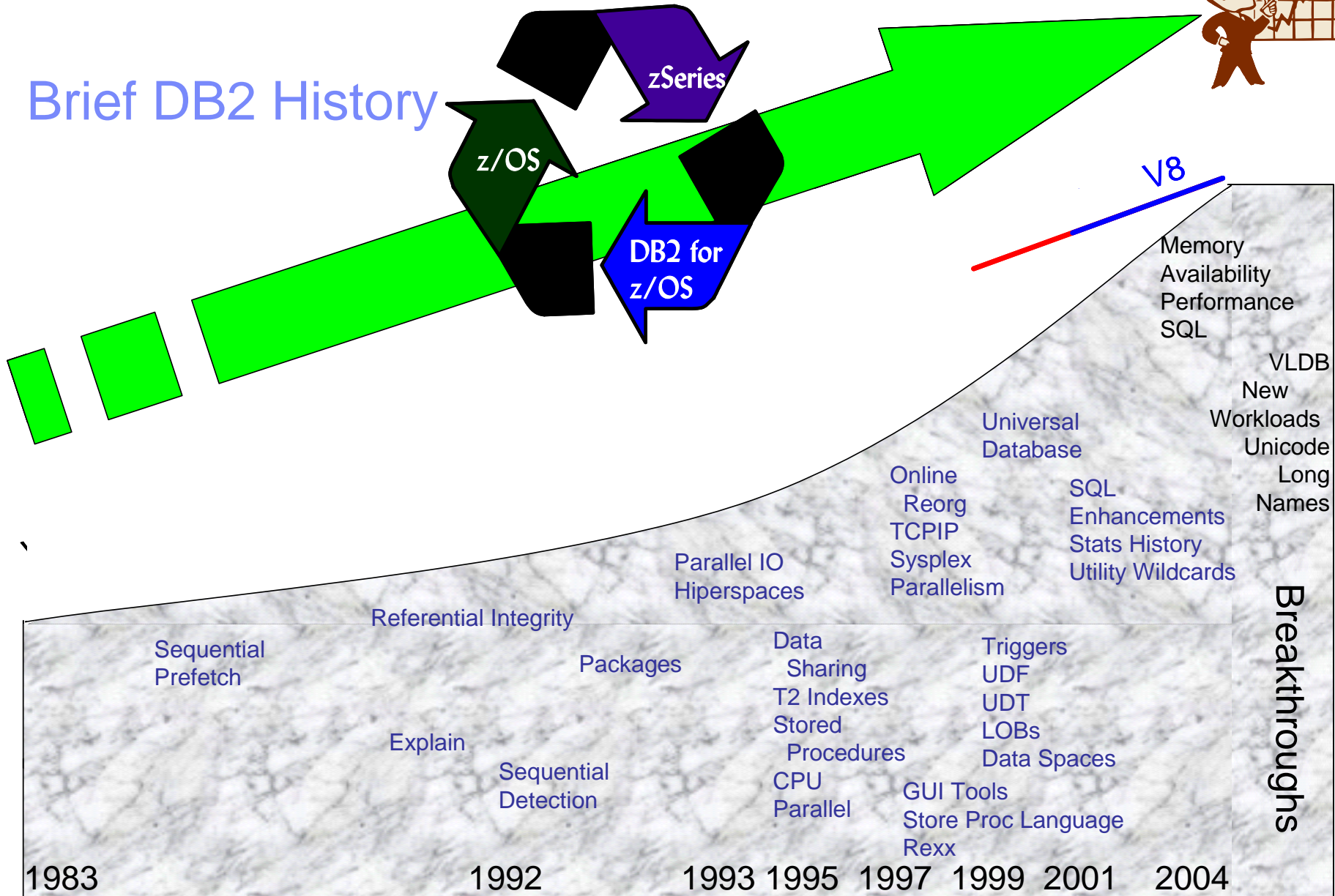
## DB2 for z/OS Strategic Direction

- Continue to Evolve as “The” Enterprise Server for Network Computing Solutions
  - ▶ Centralized Processing of Massive Quantities of Data
  - ▶ DB2 Family Compatibility
  - ▶ OLTP, ERP, Batch, e-business
- Premier Use of Parallel Technology and z/OS - zSeries Integration
  - ▶ Data Sharing
  - ▶ Continuous Availability
  - ▶ Incremental Growth
  - ▶ Work Load Balancing
- Improved Application Enablement Through Extensive SQL Functionality
  - ▶ Object-Oriented Extensions
  - ▶ Rules Driven Integrity
  - ▶ Vendor Enabling and Portability
- Continual Performance Improvement
  - ▶ Optimizer Technology
  - ▶ Transaction, Query, Batch, Utility Concurrency
  - ▶ Utilities and Tools Focus





# Brief DB2 History



# Synergy with zSeries and z/OS

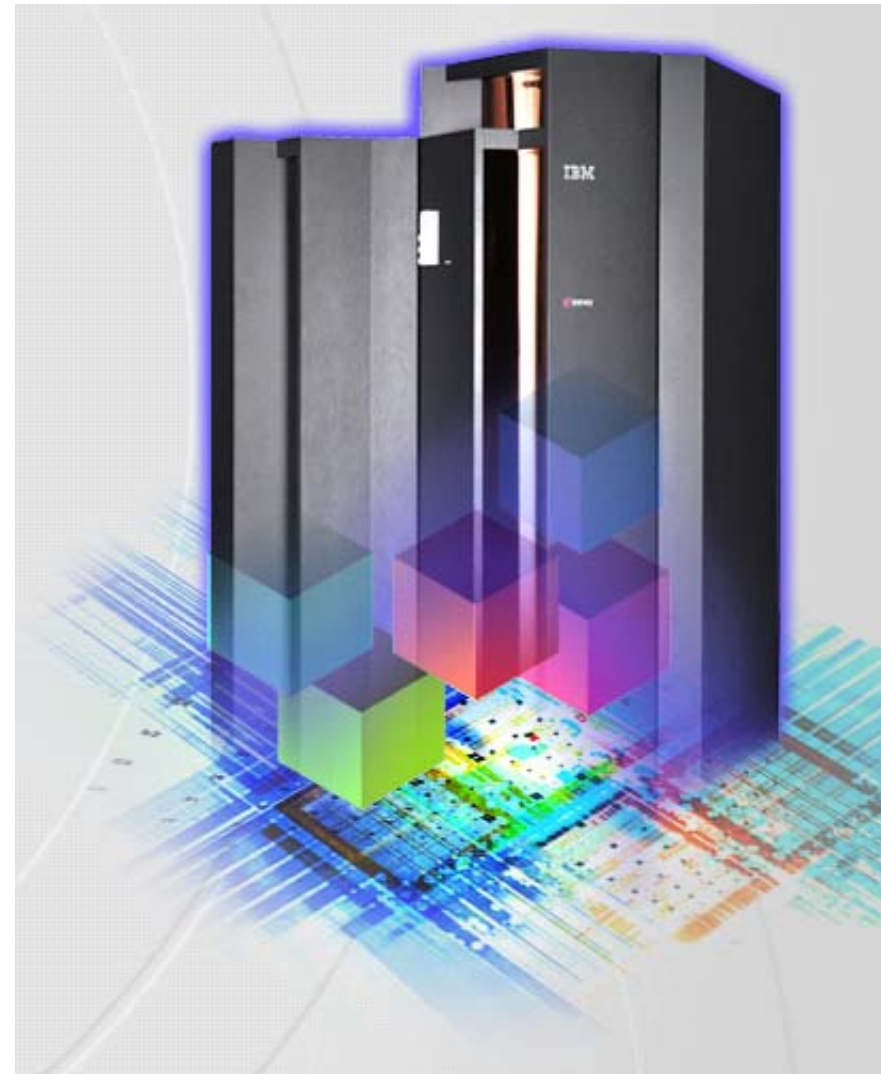
- Continuous Availability Features
  - ▶ z/OS recovery services for non-stop operations
  - ▶ Manage mixed workloads to achieve business goals
  - ▶ Disaster recovery, business continuance
- Large memory
  - ▶ 64 bit addressing
  - ▶ Cross memory and multiple address spaces
- Performance and Scaling Features
  - ▶ Special hardware instructions
    - Compression, sort, Unicode translation, encryption
  - ▶ Hipersockets
  - ▶ Extended addressing for large tables
- Parallel Sysplex Clusters
  - ▶ High performance data sharing using all three CF structure models
  - ▶ Robust failure recovery: >50% of the engineering effort
  - ▶ Single system image using z/OS coupling services
  - ▶ 100's of person years of DB2 research and development





## zSeries Server Innovations for On Demand Readiness

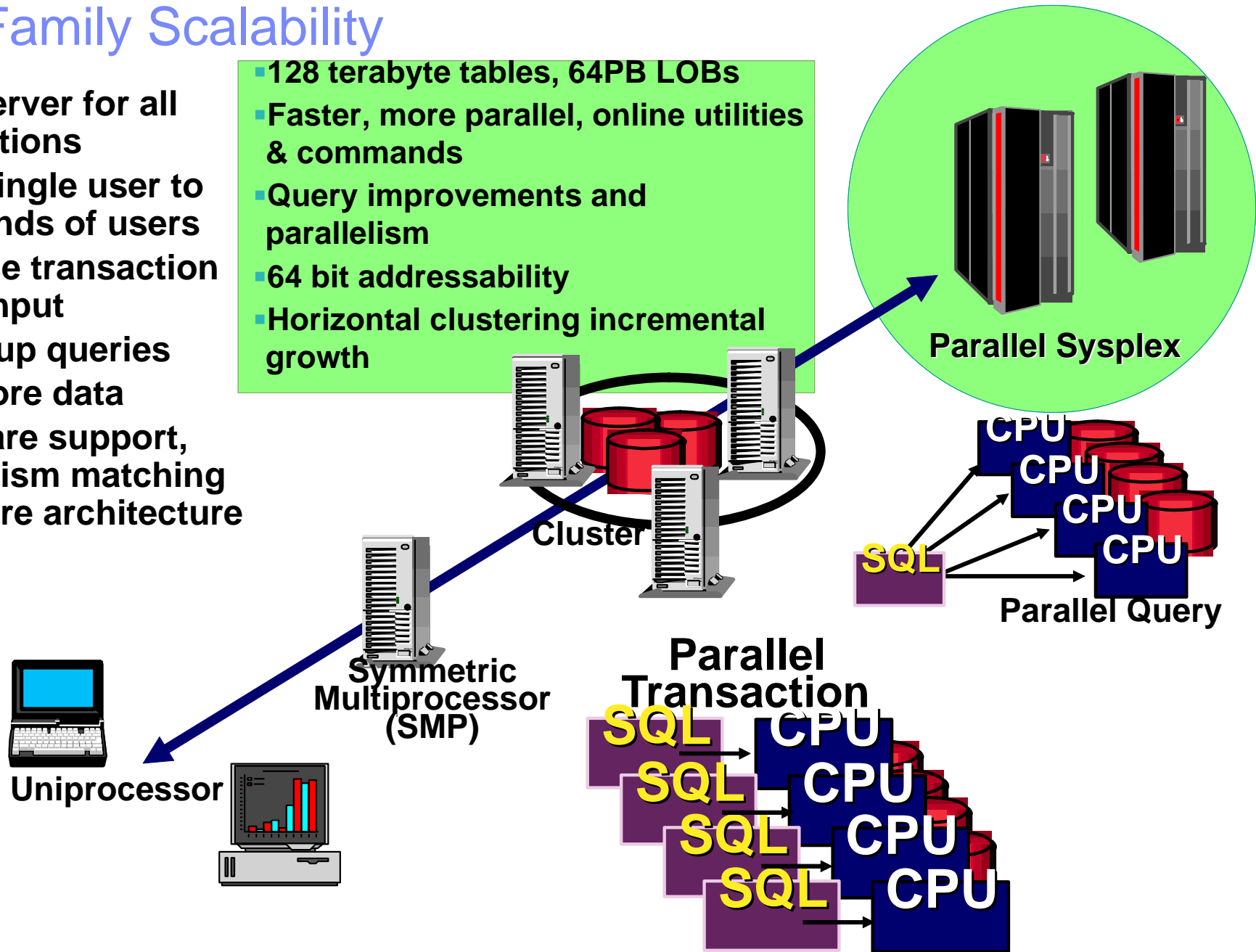
- Resource Virtualization
- Business Resiliency and Security
- Intelligent Workload Management
- Business Integration
- 64 bit z/Architecture
- Special Designated workload processors
  - ▶ Coupling (ICF)
  - ▶ Linux® (IFL)
  - ▶ Java™ (zAAP) workloads



# DB2 Family Scalability

- Data Server for all applications
- From single user to thousands of users
- Increase transaction throughput
- Speed up queries
- Add more data
- Hardware support, parallelism matching hardware architecture

- 128 terabyte tables, 64PB LOBs
- Faster, more parallel, online utilities & commands
- Query improvements and parallelism
- 64 bit addressability
- Horizontal clustering incremental growth





## DB2 Scalability ...

- 64 bit addressability - 1 TB buffer pools
- Long Names
- 128 TB tables
- Multiple encoding schemes (EBCDIC, Unicode, ASCII)
- Unicode catalog, SQL parsing
- 4096 partitions per table
- Up to 100,000 open datasets per DB2 with z/OS V1R5
- Up to 150,000 connections per DB2 member
- Up to 2000 active threads at a time
- 2MB SQL Statements
- Enhanced optimization tolerant of differing data types
  - ▶ Java and C/C++ workload benefit
  - ▶ Joining unlike character or numeric data types
- 225 way joins

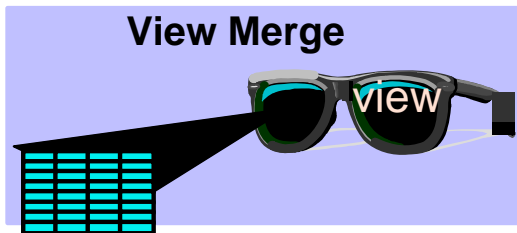


# Mature Optimizer

Static SQL  
Dynamic SQL  
Dynamic w/ Static Security

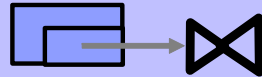
## SQL Optimizer

Query Rewrite

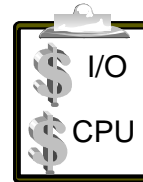


Predicate translation

Transformation  
 OR → IN  
 Subquery to Join

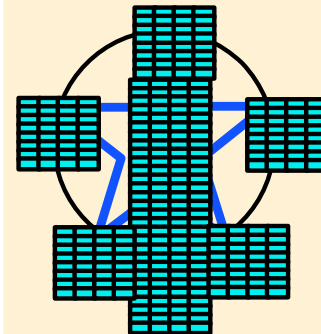


Multiple Locking Levels

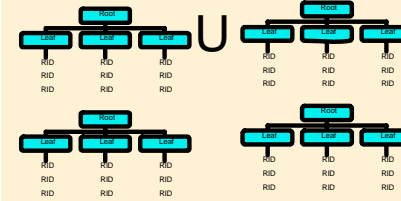


Indexing Techniques

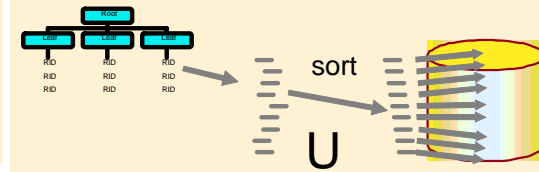
Star-schema



Index and/oring

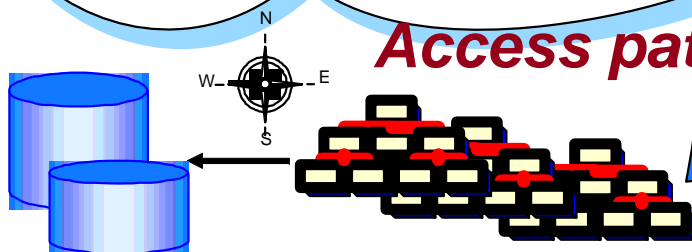


List Prefetch



Index-only access

*Access path*



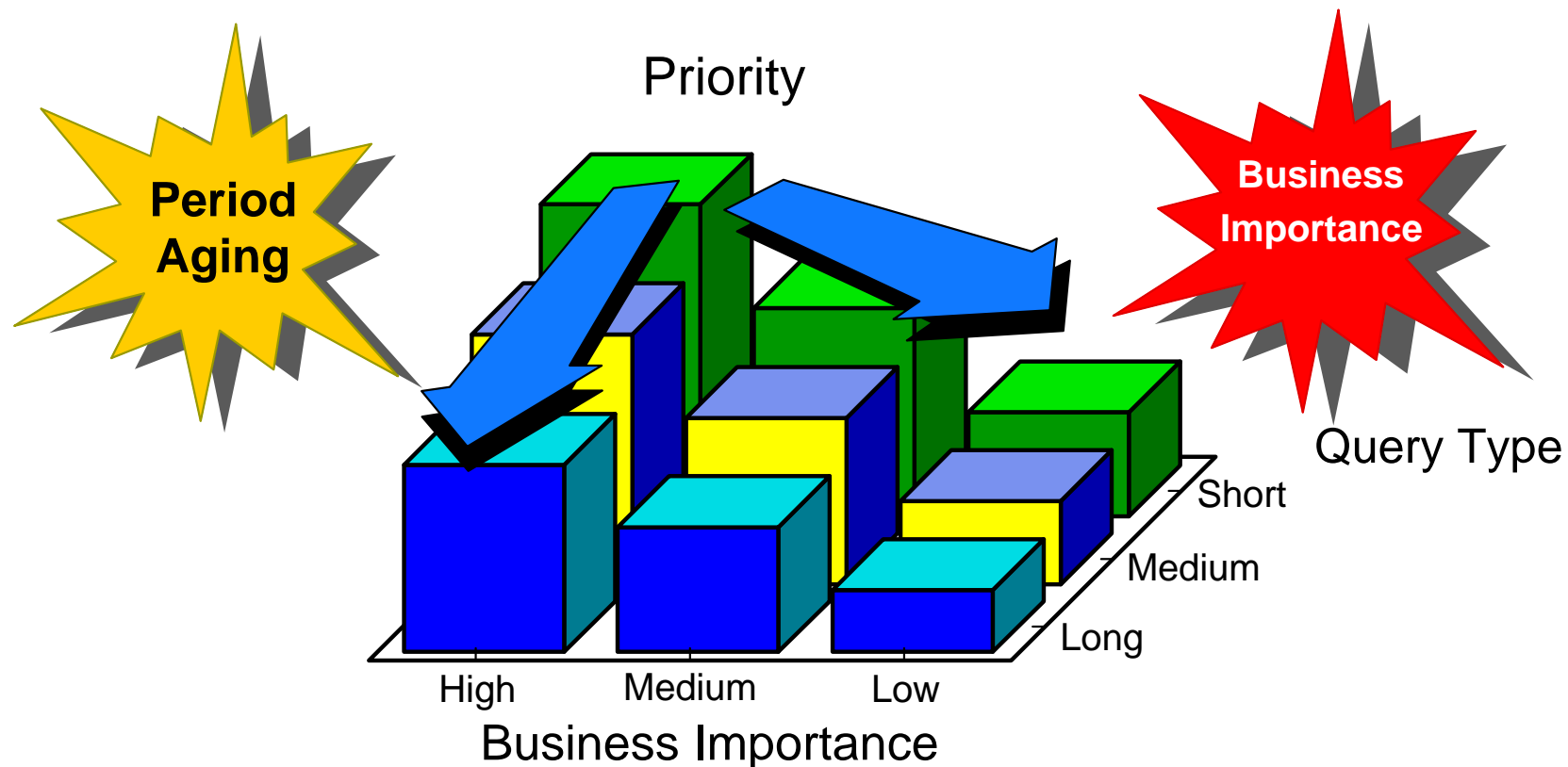
Heuristics

Sequential Detection  
List Prefetch

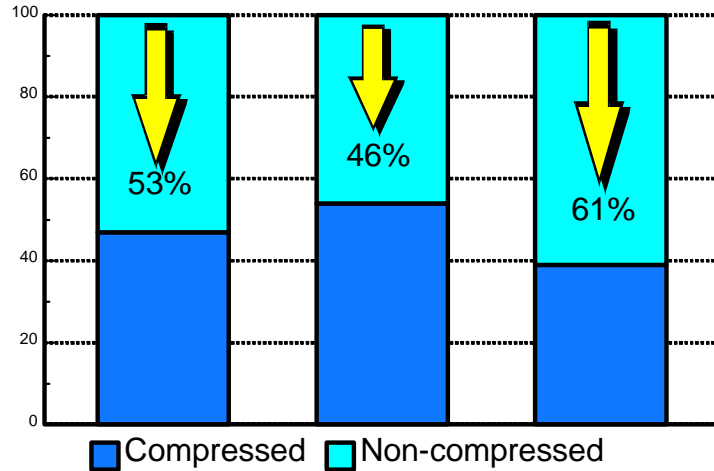


## z/OS Workload Manager

- Business importance of tasks honored even for mixed workloads
- Long-running work will not impact short running work



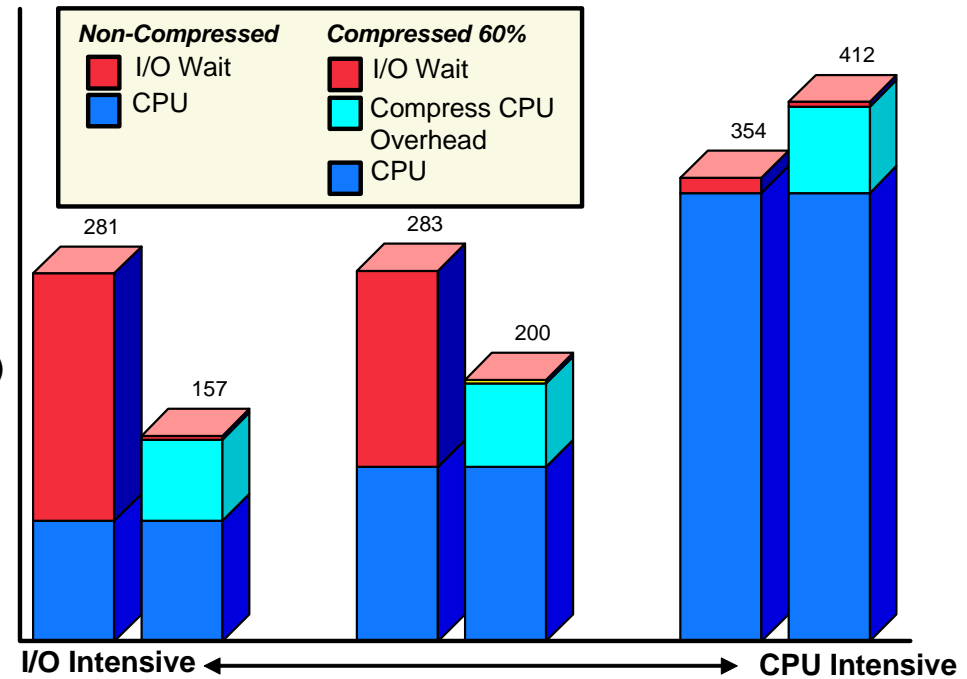
# Hardware Assisted DB2 Compression



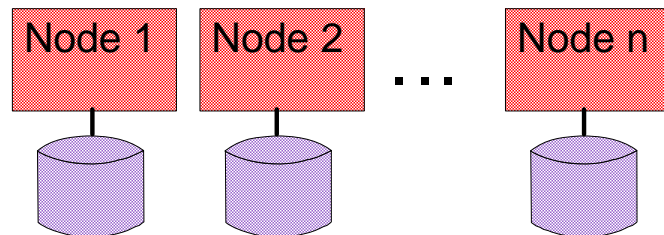
← Sample Compression Ratios Achieved

→ Effects of Compression on Elapsed Time

Elapsed time (sec)

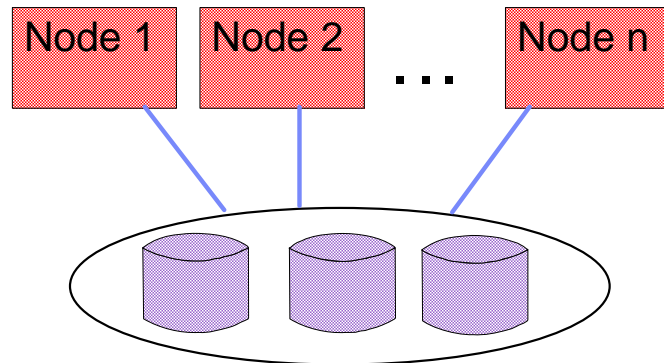


# Different Forms of Clustering



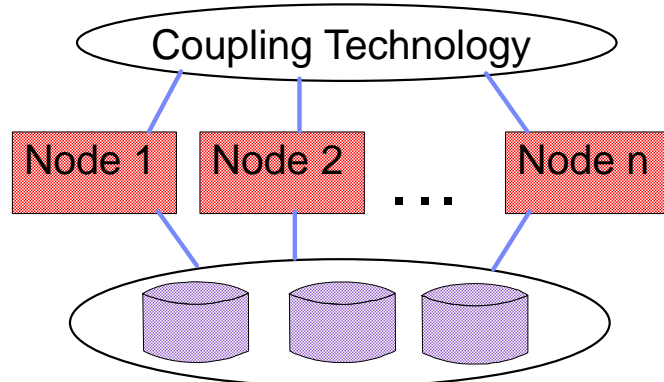
## Shared Nothing (SN) - DB2 for LUW

- Database partitioned, no disks shared amongst nodes
- Excellent scaling for query or partitioned OLTP
- Distributed commit may be necessary



## Shared Disks (SDi) - Oracle RAC

- No database partitioning necessary
- But partitioning needed for good performance
- OLTP requires inter-node concurrency and buffer coherency control mechanisms
- Extensive messaging and disk I/O overhead
- Dynamic load balancing not possible
- Architecture not scalable



## Shared Data (SDa) - DB2 for z/OS

- Aggressive use of z/OS Parallel SysPlex coupling facility hardware avoids heavy SDi overheads
- Excellent scalability for OLTP, ERP and mixed workloads
- Data partitioning not necessary for good performance
- Strong fail-over and dynamic load balancing
- Flexible non-disruptive growth



# Continuous Availability

## 1. Unplanned/system

- ▶ H/W failure (CPU, memory, server, coupling facility, etc)
- ▶ S/W failure (OS, DBMS, clusters)
  - If the 'failure masking' aspects fail, then DBMS crash recovery performance becomes critical for a highly resilient system
- ▶ Reliability also plays in here
  - Frequency of failure is at least as important as how the failure is handled

## 2. Unplanned/data

- ▶ H/W
  - DASD (RAID, DBMS partial write detection)
  - DASD controller (DBMS log mirroring, z/OS HyperSwap, etc)
  - If the 'failure masking' aspects fail, then data recovery performance/throughput becomes critical for a highly resilient system
- ▶ Application data corruption (point-in-time recovery)
- ▶ Data health checking
- ▶ Site failure (disaster recovery)





# Continuous Availability ...

## 3. Planned/system

- ▶ Online reconfiguration:
  - Buffer pools
  - System parameters
  - etc.
- ▶ Adding CPU, memory capacity
- ▶ Rolling S/W upgrades for OS or DBMS
  - SW patches
  - Version upgrades

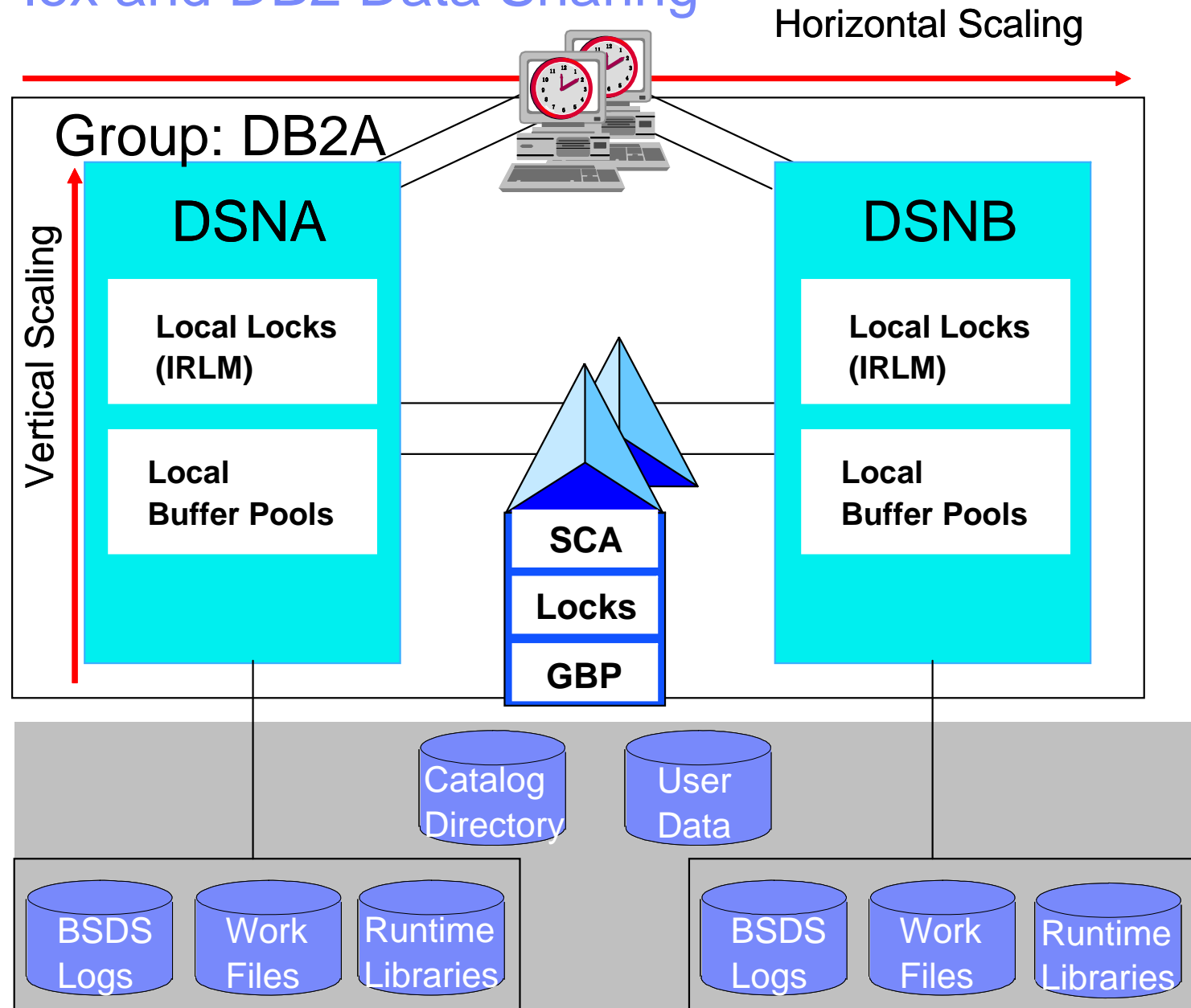
## 4. Planned/data

- ▶ Online schema change
- ▶ Data/index reorg for defrag or reclaiming space
- ▶ Online create index
- ▶ Online load
- ▶ Data partitioning



# z/OS Parallel SysPlex and DB2 Data Sharing

- Sysplex Timers
- Coupling Facilities
  - Internal (ICF)
  - External
  - Lock structures
  - Group Buffer Pools
  - Communications
- Coupling Links
- Up to 32 DB2 Members
- CF use based on Inter-DB2 interests
- Shared Catalog/Directory
- Separate BSDS, Logs
- Shared disk
- Dynamic VIPA
- Hiperswap



# Continuous Availability ...

- Avoiding **Planned** Outages
  - ▶ H/W & S/W (OS & DB2)
    - Rolling maintenance
    - Rolling version upgrades
  - ▶ Data Sharing is not needed for:
    - Online parameter changes
    - Online schema evolution
    - Online Reorg, Load, Check, etc.
    - Hot & incremental backups
    - Online statistics collection
    - Tooling for log based backup creation
    - Cross Loading of DB2 & non-DB2 data



## Continuous Availability ...

- Avoiding **Unplanned** Outages
  - ▶ Redundant Timers, CFs, Links
  - ▶ Member failure maintains access to unlocked data (immediate failover)
  - ▶ Restart Light protocol used to release retained locks making all data available
  - ▶ CF structure failures reestablished by:
    - Failover to duplexed copy
    - Dynamic structure rebuild on the fly
  - ▶ Data Sharing is not needed for:
    - Duplexed Logs, Archives
    - Fast Log Apply processing
      - Used automatically for DB2 restart
      - Optional for object and LPL recovery
  - ▶ Reliable OS & Hardware



## Key difference with DB2 for z/OS

Keep the data available

vs.

Try to recover fast



# Online Reorganization

## *Availability of Data During REORG*

Allows Read and Write Access During REORG

- Very small window when readers are drained
- Slightly larger window when writers are drained

REORG Index Only, or Data and Indexes at same time

Image Copy Taken During Process

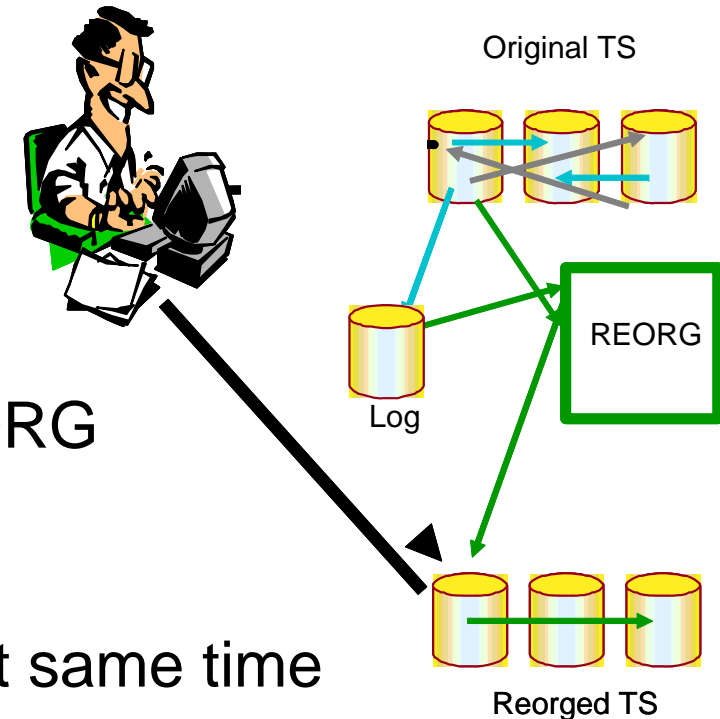
Requires Additional DASD during REORG

FASTSWITCH option for shadow datasets

BUILD2 phase repairs RIDs in Secondary Indexes

- Parallel BUILD2 Phase
- Avoided if DPSI's are used
- 

*Introduced in V5 & continually improved*



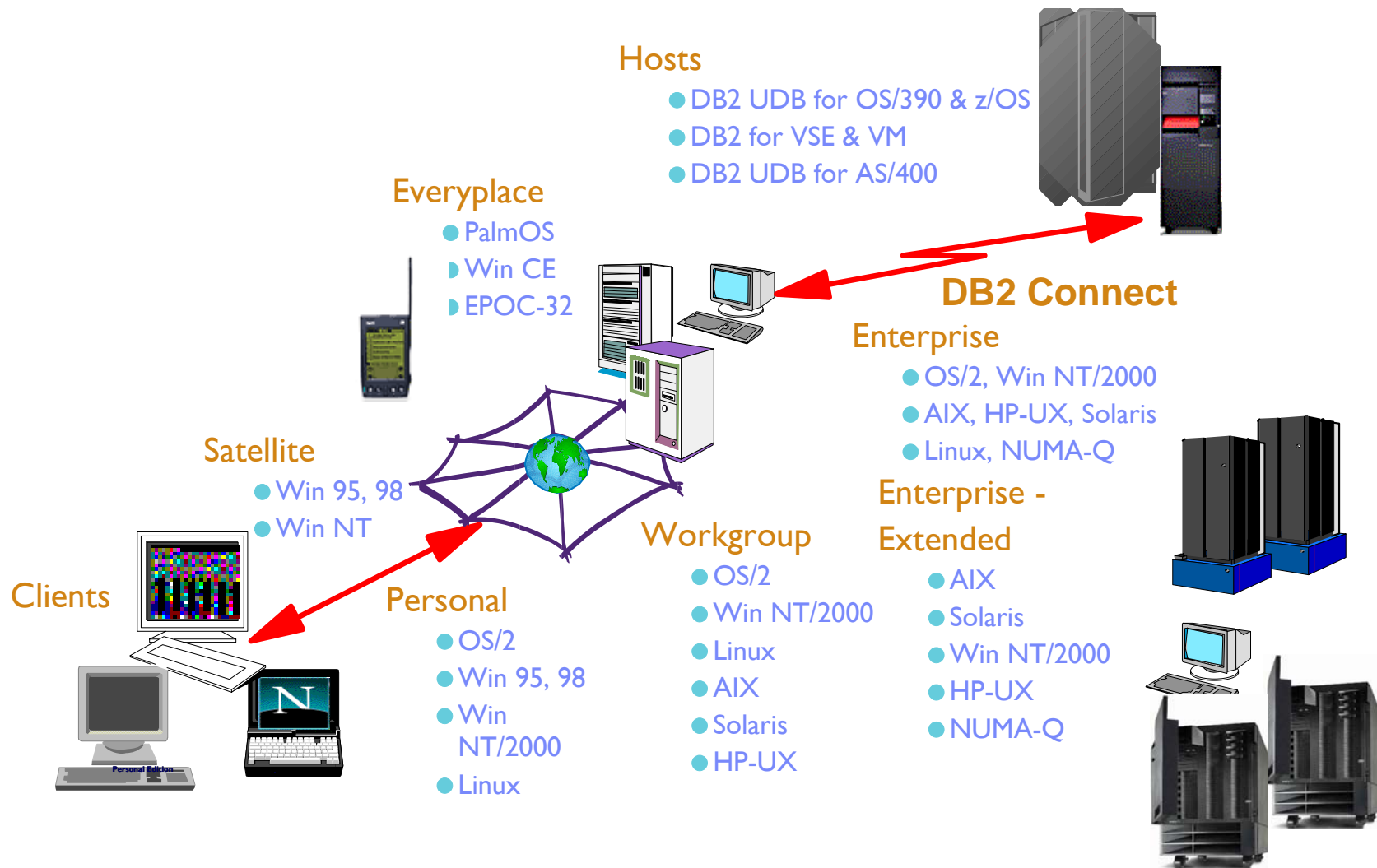


# Advanced Backup and Recovery

- Data recovery to a specific copy, point-in-time, or to current point of failure
- Data & Indexes backed up & recovered at the partition level
- System wide backup & recovery option
- Cancel thread & deferred restart options to avoid backouts of long units of work
  - ▶ Enables access to data faster
  - ▶ Tracking option for long running units of work
- DB2 Tracker site
- GDPS with Extended Remote Copy or Peer-to-peer remote copy
- Online changeable options to adjust checkpoint frequency and tune crash restart times
- Log usage is tracked @ the object level (SYSLOGRNX)
  - ▶ Logs with no activity for an object are skipped during a recovery
- Logging is spread across members in a data sharing group
- The log also supports operational performance
  - ▶ Deferred write / force at commit protocol
  - ▶ Buffer pool / Group buffer pool thresholds to moderate write activity



# zSeries and DB2 for z/OS Connectivity



# DB2 for z/OS for Network Computing

- **DB2 Connect**
  - ▶ **Member routing**
    - Balance SysPlex workload
    - Fault tolerance/failover
    - 7x24 Availability
    - Single System Image
  - ▶ **Performance/scalability**
    - Connection pooling
    - Thread (DBAT) reuse
    - XA support for TCP/IP
  - ▶ **Monitoring**
    - Current user of thread
    - End user workstation
    - Application name



## DB2 for z/OS for Network Computing ...

- Open standard interfaces for database access
  - ▶ SQL, ODBC, JDBC, SQLJ
  - ▶ DRDA over TCP/IP or SNA
  
- Stored Procedures
  - ▶ Reduces network traffic in a client server environment
  - ▶ Provides highly reusable code
  - ▶ Can be written in 3GL, Java, SQL Procedure Language (PSM)
  - ▶ Can be in a different language than that of the calling program
  - ▶ Support for MQSeries, access to VSAM files and IMS databases, issue DB2 commands
  - ▶ Can return multiple result sets
  
- Performance and Scalability
  - ▶ WLM workload prioritization
  - ▶ Thread (DBAT) pooling with up to 150K connections per server
  - ▶ Up to 2000 active threads/DBATs

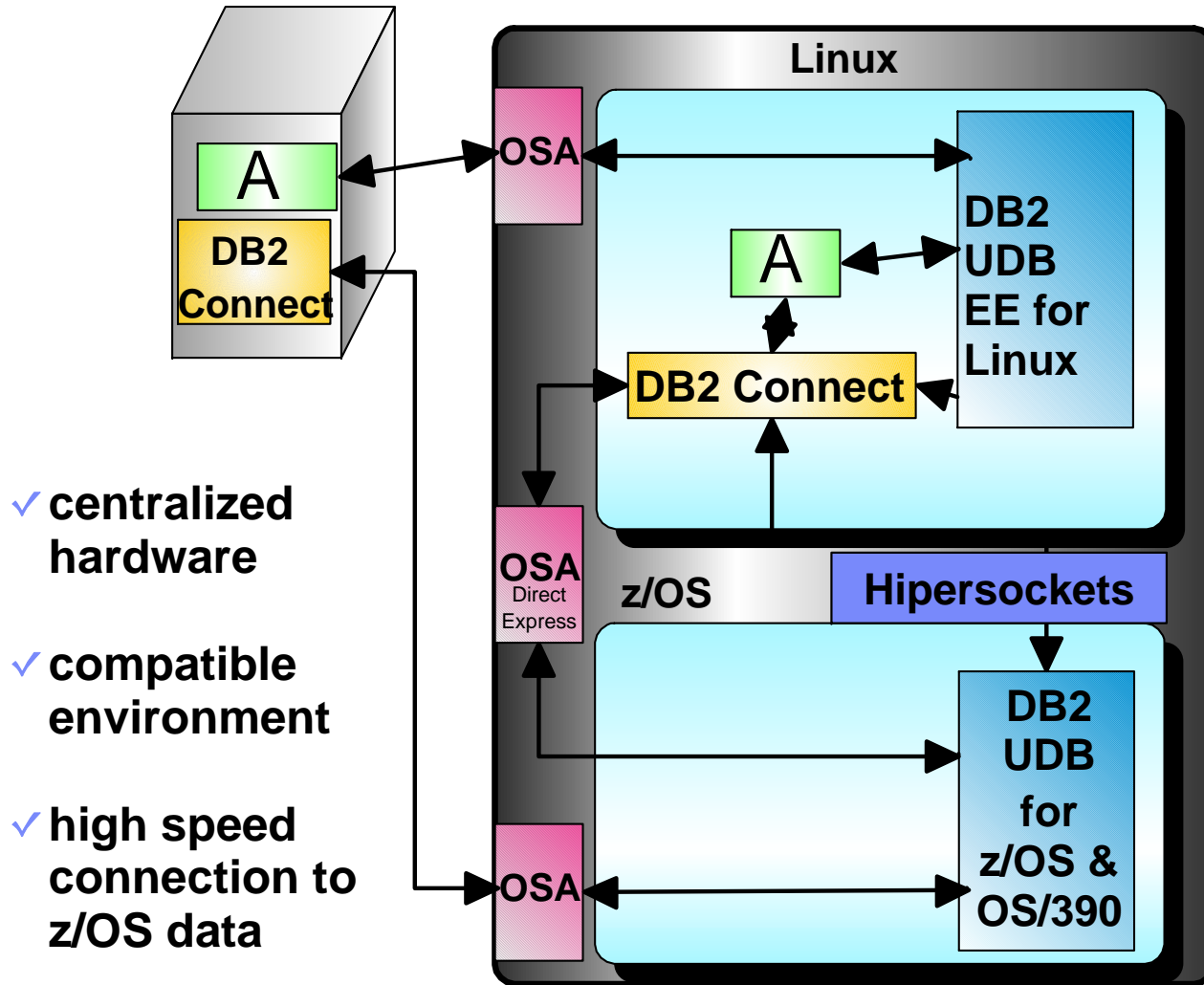


## DB2 for z/OS for Network Computing ...

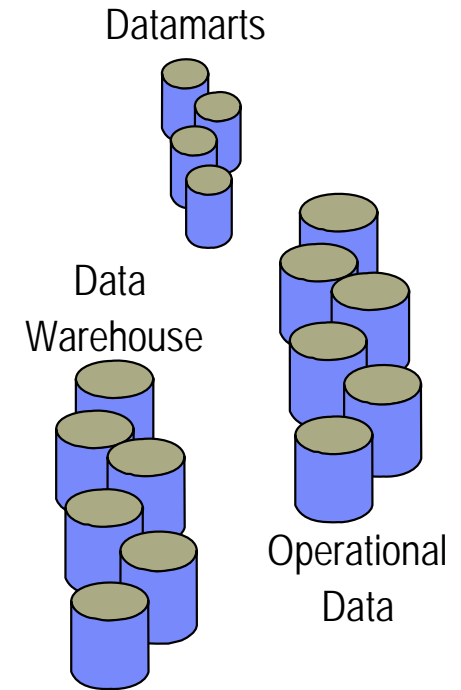
- Enhanced Stored Procedures
  - ▶ Easy to use Stored Procedure Language PSM (ANSI)
    - Familiar to Oracle, Informix, Sybase, SQL Server application developers
    - Supported by all DB2 platforms
  - ▶ Stored Procedure Builder which can be integrated with popular interactive development environments
    - Single tool supporting the entire DB2 Family with one-step build to any DB2 server platform
    - SQL coding assistants
    - Test execution of installed SPs
    - Source level remote debug of installed SPs
  - ▶ Develop on one platform and deploy on other server platforms
    - e.g. Develop and test on DB2 for LUW, use in production on DB2 for z/OS



# zSeries Connectivity: z/OS & Linux

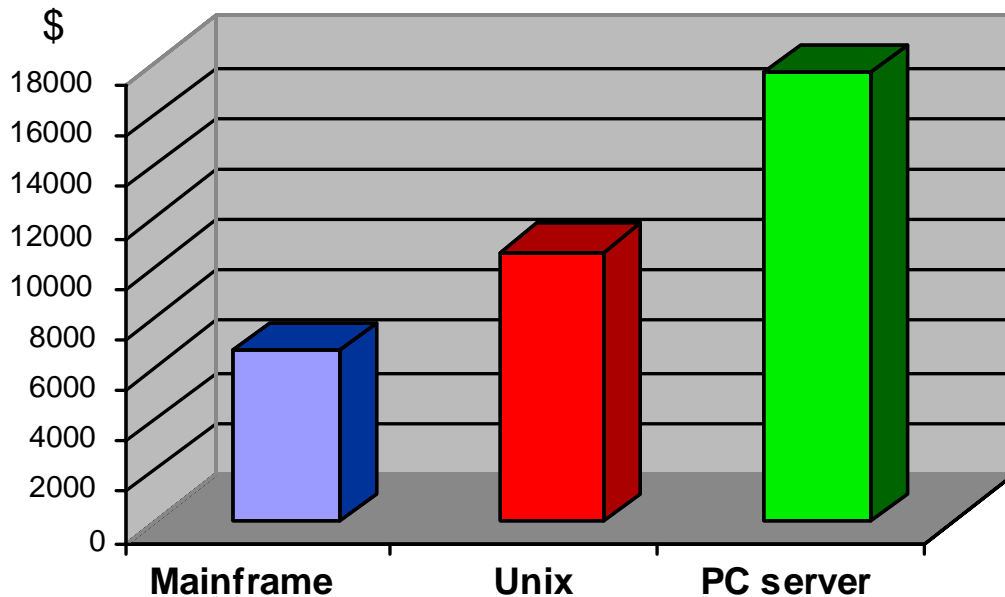


- ✓ centralized hardware
- ✓ compatible environment
- ✓ high speed connection to z/OS data





## zSeries TCO Advantage



**Total annual cost per user over a five-year period**

Source: *The Dinosaur Myth*, www.arcati.com, 2004



“We believe that simple Cost of Ownership comparisons between the mainframe and distributed platforms are often misleading - dangerously so. Many distributed costs are surprisingly well hidden within the enterprise; but with the mainframe data center, what you pay is what you get.

And what you get is unparalleled scalability, very high utilization levels, and mature centralized management. **This in turn significantly reduces the need for technical support, simplifies change management, and allows a more flexible approach to business continuity. “**

Mark Lillycrop, Arcati Research

## Benefits of zSeries, z/OS and DB2 Together

- Robust, proven, industrial strength database used by the largest WW enterprises
- Open standard interfaces (application, connectivity, web enablement)
- Very large database support
- Online utilities and automated recovery
- 7 x 24 availability (99.999%)
- Balanced performance and workload management
  - ▶ Across spectrum of application types
  - ▶ Mixed workloads
  - ▶ Many concurrent users
- Cross platform, scalable solution to meet changing business needs
- High performance host connectivity
- Performance monitoring and tuning
- Serviceability



## zSeries Rated Top Again

Table 2  
OLTP/ERP Serving Summary: Weighted Data

	IBM z900	HP Non-Stop 86000	Sun Sun Fire 15K (Starcats)	HP Super-dome	IBM p690 (Regatta)	IBM i890 (Regatta)	Fujitsu Prime-power 2500	Linux Cluster	HP Pro-Liant	IBM xSeries x440	Dell Power-Edge	Unisys ES7000 Orion
Technology	205	201	137	159	150	169	150	84	97	97	91	121
Manageability	220	178	146	169	138	190	156	83	87	82	75	116
Market Momentum	193	127	250	245	258	186	226	168	280	274	265	189
Business Practice	194	171	154	183	199	196	189	124	172	191	180	183
User-Definable Criteria	0	0	0	0	0	0	0	0	0	0	0	0
Total	812	677	687	756	745	741	721	459	636	644	611	609
Rank	1	7	6	2	3	4	5	12	9	8	10	11

Source: Gartner Research (August 2003)



# Why Customers Pick DB2 for z/OS

- Unsurpassed Experience and Leadership
  - ▶ Processor
    - RAS and SMP clustering
  - ▶ Operating System
    - Dynamic and automatic features
  - ▶ Data Base
    - DSS <---> OLTP workloads concurrently
  - ▶ Space Management
    - Policy based data placement/movement
- Handle diverse and unpredictable workloads for mission critical applications, while maximizing resource utilization
- Integration of products - not platform agnostic
  - ▶ Processor, Operating System, Data Base
  - ▶ DASD Space Management, ESS

