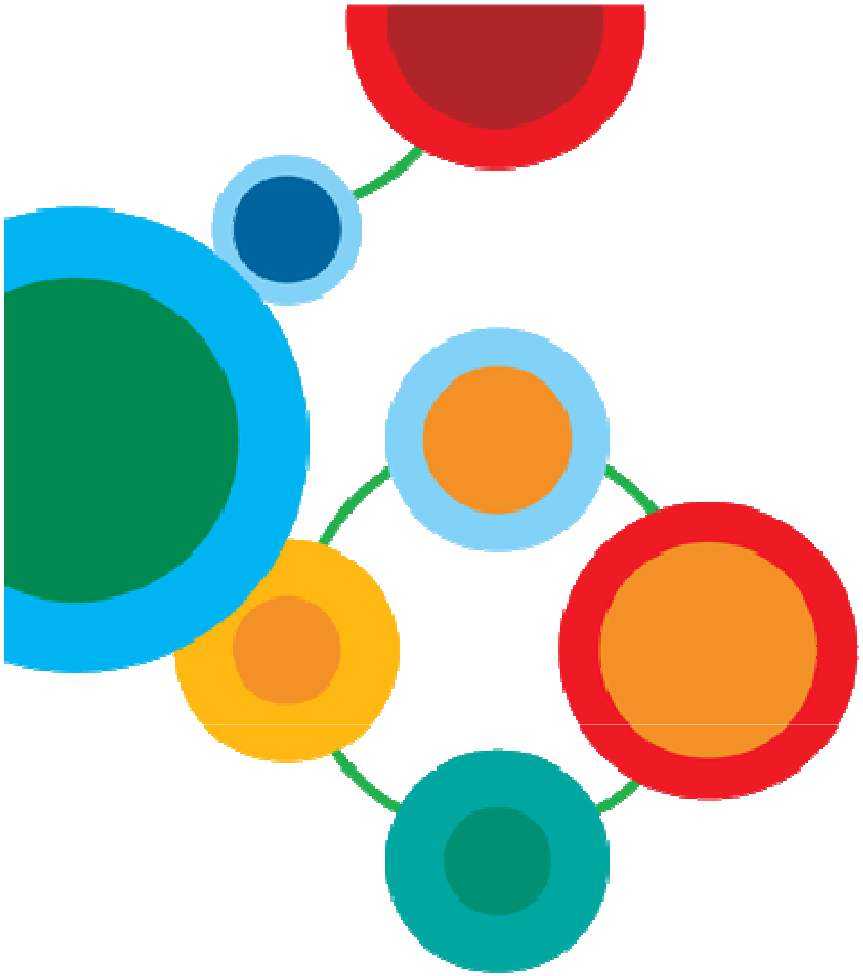


DB2 for z/OS
Trends and Directions

John Campbell
IBM DB2 for z/OS Development

campbelj@uk.ibm.com

Disclaimer:

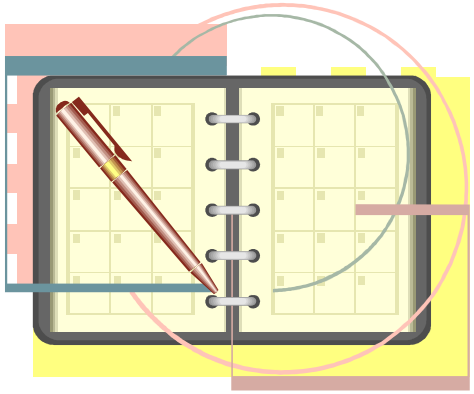
Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The Information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance Disclaimer:

This document contains performance information based on measurements done in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the numbers stated here.

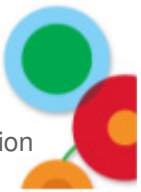
Agenda

- DB2 for z/OS technical strategy and roadmap
- DB2 10 quick review
- DB2 11 preview



DB2 for z/OS Technical Strategy

- Availability, scalability and performance are core values
- Through tight integration with IBM System z stack
  - Industry leading DBMS for mission critical applications
  - Delivering excellent performance, 24*7 availability, near limitless scalability, tight security
  - Platform of choice for workload consolidation
- Reduce cost of ownership
  - Grow workload without growing staff
  - Leveraging zIIPs and zAAPs
  - Autonomics to improve DBA productivity
- Advanced features for new and enhanced applications
  - Integration with larger enterprise
  - Growing with technological advances
  - Rich growing set of SQL and XML functionality
- Grow analytics capabilities
  - Do not move the data - simplification, less complexity, cost savings
  - Provide a hybrid architecture for OLTP and analytics in the same database
  - Exploit first class native DB2 z/OS query performance capabilities



DB2 Deep Synergy with IBM System z drives high value

- Hardware-based Coupling Facility for efficient DB2 data sharing (availability and scale out)
- zIIP engines to offload certain DB2 workloads
- Hardware instructions for Unicode conversion
- Hardware data compression & encryption
- Cross-memory, memory protection keys
- Sorting
- z/OS Workload Manager (WLM)
- z/OS Security Server (RACF)
- 1 MB page size
- Volume-level and object-level FlashCopy
- Solid state disks (SSDs)
- System z High Performance FICON (zHPF) with multi-track data transfer and FICON Express 8

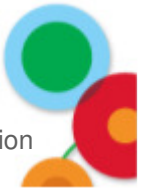
- Mission-critical data server of choice:
  - High availability, scalability, performance
  - Ease of operations
  - Low TCO: running mixed workloads at high utilization on fewer images

DB2 and IBM zIIP Add Value to Database Work

Portions of the following DB2 9 & DB2 10 workloads may benefit from zIIP or zAAP

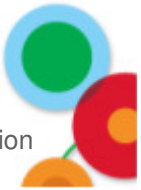
DB2 9 in blue DB2 10 in green

- DRDA over TCP/IP connections
  - DB2 9 for z/OS remote native SQL procedures
  - DB2 9 XML parsing, DB2 schema validation
  - Increased portion of DRDA redirected to zIIPs – up to 60%
    - Improved performance via reduced processor switching
- Requests that use parallel queries
  - Higher percentage of parallel queries in DB9 are zIIP eligible
  - More DB2 10 queries are eligible enabling more parallelism
- DB2 utilities LOAD, REORG and REBUILD functions used to maintain index structures and sort
  - DB2 10 RUNSTATS – options other than column group, inline
- DB2 10 buffer pool prefetch and deferred write



Strategy: Continuous Availability and RAS Leadership

- Remove all causes for planned outages
- Eliminate outages due to a failure of any single hardware or software component
- Streamline backup/recovery process by moving to volume-based COPY/RECOVER
- Leadership in disaster recovery solutions for both RPO (avoidance of data loss) and RTO (fast recovery time)
- Lead in overall product quality
- Lead in customer service



Aspects of Availability

High Availability
Fault-tolerant, failure-resistant infrastructure supporting continuous application processing

Continuous Operations
Non-disruptive backups and system maintenance coupled with continuous availability of applications

Disaster Recovery
Protection against unplanned outages such as disasters through reliable, predictable recovery

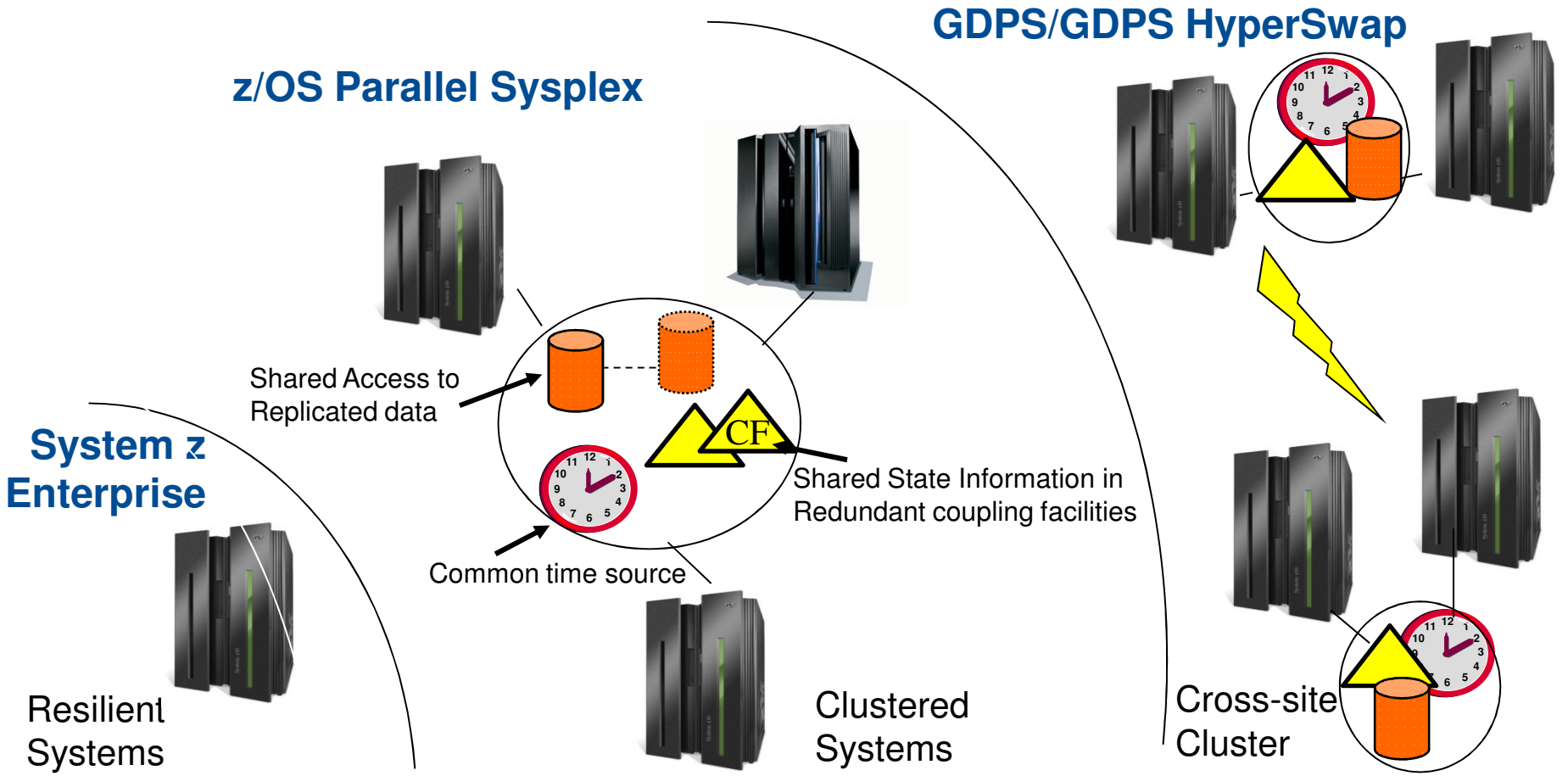
Protection of critical business data

Operations continue after a disaster

Recovery is predictable and reliable

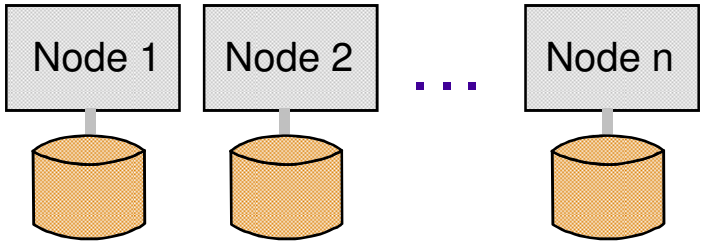
Costs are predictable and manageable

System z Availability Building Blocks



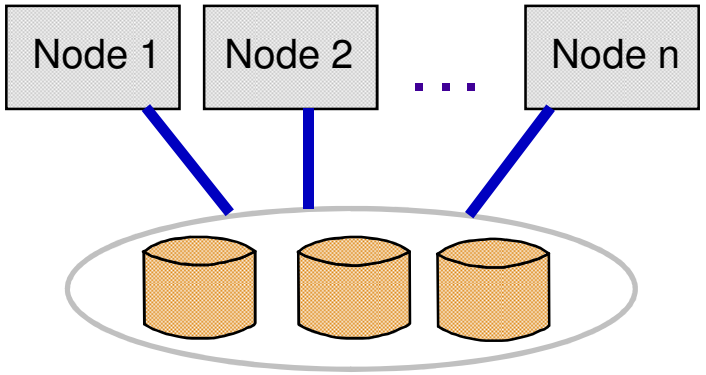
High Availability … Continuous Operations … Continuous Availability

Alternative Parallel DBMS Architectures



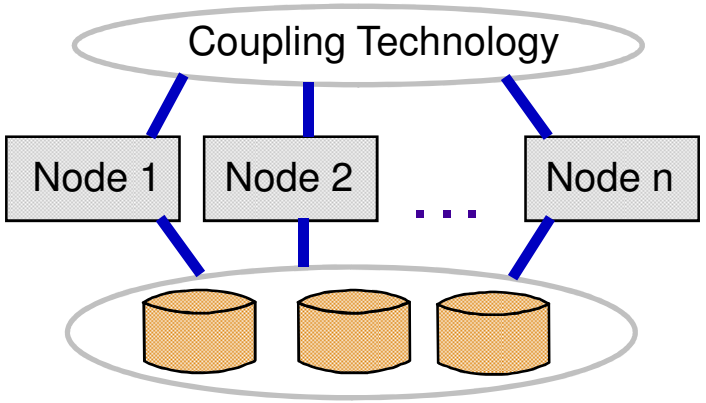
Shared Nothing

- Database is partitioned
- No disks are shared amongst the nodes
- Data repartioning necessary as nodes are added
- Susceptible to skewed access patterns
- Strong architecture for DW/BI (or “sharding”)



Shared Disks

- No database partition necessary
  - But partitioning can give better performance
- Strong fail-over characteristics
- Dynamic load balancing
- Inter-node concurrency and coherency control mechanisms are needed
  - Messaging overhead limits performance, scalability

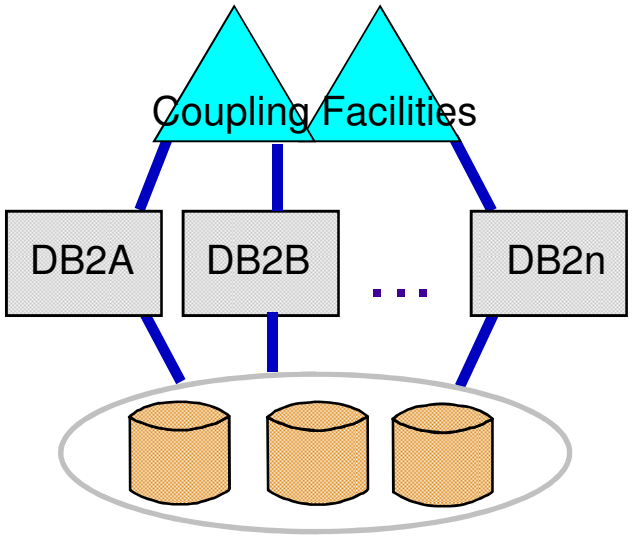


Shared Data <- DB2 for z/OS Data Sharing

- Best performance, availability
- Coupling facility is used as hardware assist for efficient concurrency and coherency control
- Continuous availability model (not fail-over)
- Flexible growth
- Messaging overhead minimized, excellent scalability

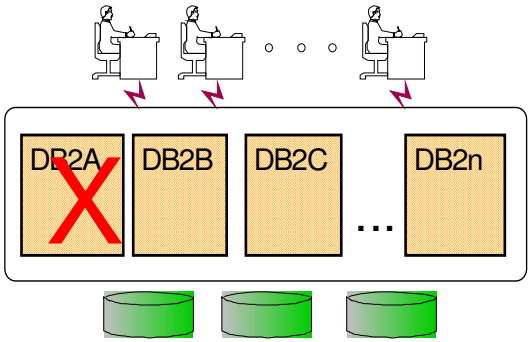
Data Sharing Design for Continuous Availability

- Goal: Continuous availability across planned or unplanned outage of any single hardware or software element
- Single points of failure eliminated:
  - DB2 subsystem or z/OS system
  - CPC (or CEC)
  - I/O path
  - Hiperswap for disk controller high availability
- Strategy:
  - Remove all causes for planned outages
  - Build on legacy of robust, fault tolerant z/OS components
  - On a failure:
    - Isolate failure to lowest granularity possible
    - Automate recovery and recover fast



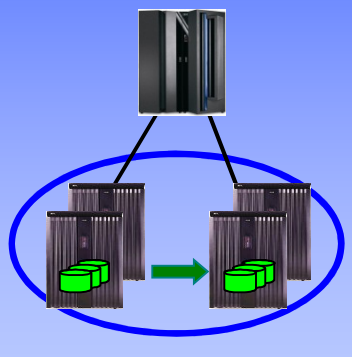
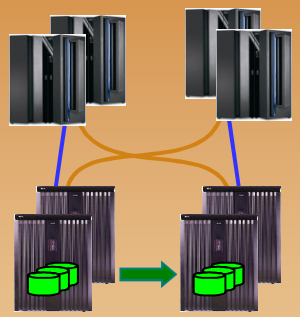
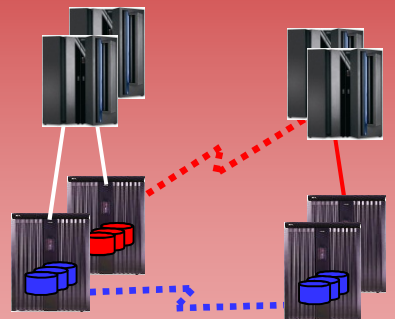
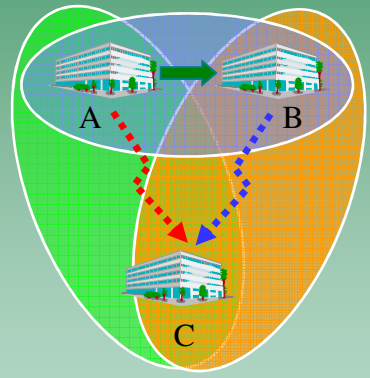
Data Sharing Design for Continuous Availability …

- Goal: Continuous availability across any planned or unplanned outage of any single hardware or software element
- IBM System z reliability means that system failures are rare
- But if they happen, ensure that you have tuned to fully utilize data sharing's capabilities to mask the failure
- The failed member holds "retained locks" in the CF to protect in-flight data
  - Should be fine granularity
    - Pages/rows



  - Data with retained locks not available until recovery completes
    - Can still be accessed by UR readers
    - And probably CS readers – if lock avoidance works
  - All other data fully available from surviving members
- Recommendation: Applications must be Sysplex-enabled and running on more than one DB2 member!

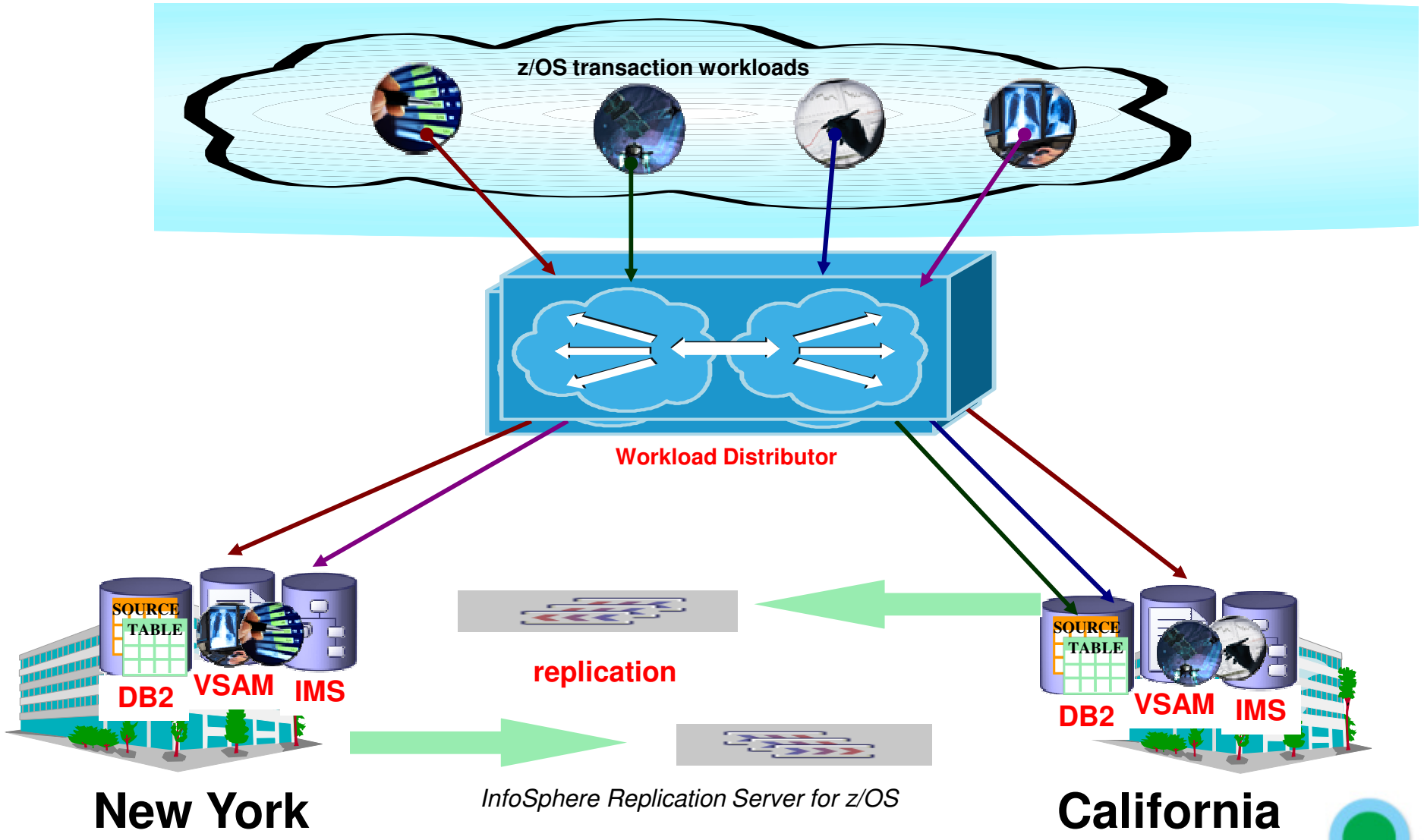
GDPS: What are customers doing today ?

Continuous Availability of Data within a Data Center	Continuous Availability / Disaster Recovery within a Metropolitan Region	Disaster Recovery at Extended Distance	Continuous Availability Regionally and Disaster Recovery Extended Distance
Single Data Center Applications remain active Continuous access to data in the event of a storage subsystem outage	**Two Data Centers** Systems remain active Multi-site workloads can withstand site and/or storage failures	**Two Data Centers** Rapid Systems Disaster Recovery within "seconds" of Data Loss Disaster recovery for out of region interruptions	**Three Data Centers** High availability for site disasters Disaster recovery for regional disasters
			
GDPS/HyperSwap Mgr RPO=0 & RTO=0	GDPS/PPRC RPO=0 & RTO<1 hr	GDPS/GM & GDPS/XRC RPO secs & RTO <1 hr	GDPS/MGM & GDPS/MzGM

RPO – Recovery Point Objective

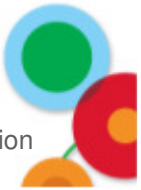
RTO – Recovery Time Objective

Active/Active Sites concept



Strategy: Performance and Scalability

- Scale on both large SMPs (scale up) and Parallel Sysplex (scale out)
- Efficient resource utilization to lower costs
  - Compression, specialty engines, MIPS reduction features
- Move to 64-bit exclusive
  - Done as of DB2 10
- Exploit large memories for improved performance and easier management
- Tight integration with latest z hardware capabilities for leading performance

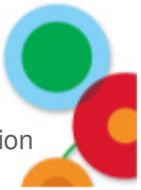


DB2 and zEnterprise EC12

- Faster CPU – 1.25x compared to z196
  - 20-28% CPU reduction measured with DB2 OLTP workloads
  - 25% reduction measured with DB2 query and utilities workloads
  - Less compression overhead with DB2 data (1-15%)
- 50% More System Capacity to help consolidation
  - Excellent synergy with DB2 10 scalability
- New Features DB2 plans to exploit
  - 2GB real storage frame support
    - Additional CPU savings, especially for very large memory
  - DB2 code backed by large frames for CPU reductions
- Flash memory improves SVC dump and z/OS paging performance

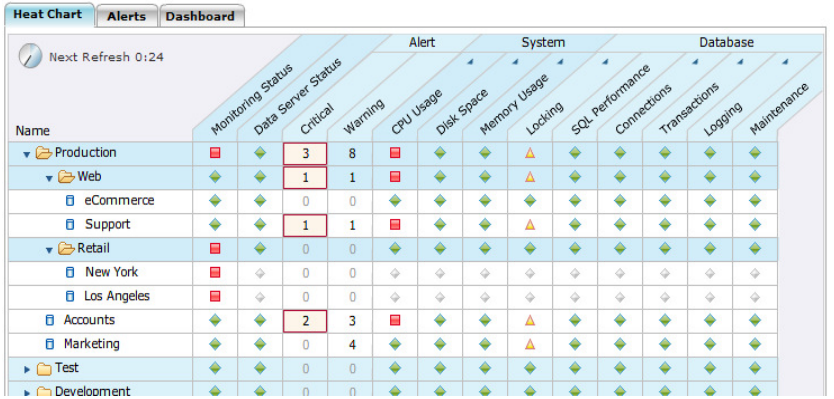
Strategy: Autonomics and Simplification

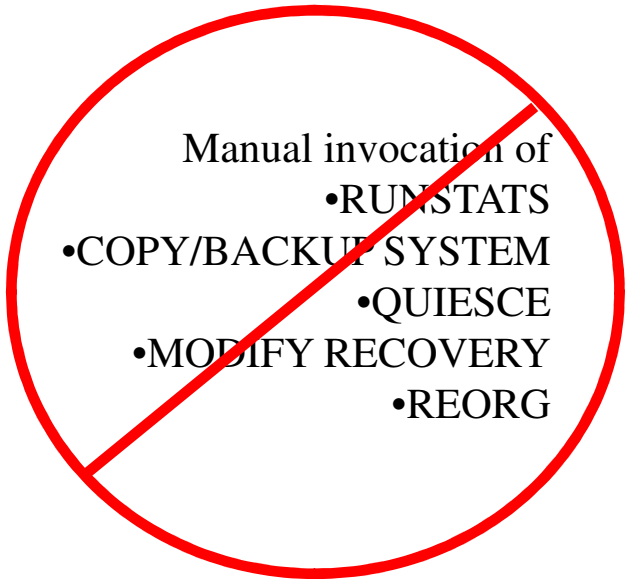
- Make DB2 more self-managing for the DBAs and system programmers
  - Reduce number of "knobs“
  - Improve system robustness by removing the need for manual intervention
  - Avoid the need for REORGs
- Provide best of breed utilities



DB2 10: Productivity – Doing More with Less!

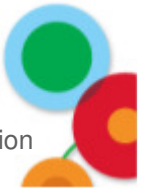
- Easier scaling, simpler memory management
- Reduce contention, more online processing
- Reduced need for REORG
  - Build compression dictionary on the fly
  - Index list prefetch enhancements
  - Row-level sequential detection
- Configure IBM UDFs and stored procedures
- Statement level monitoring
- New DSNTIJXZ job to update migration input datasets with current zparm values
- Query performance management improvements
- DDF thread management enhancements



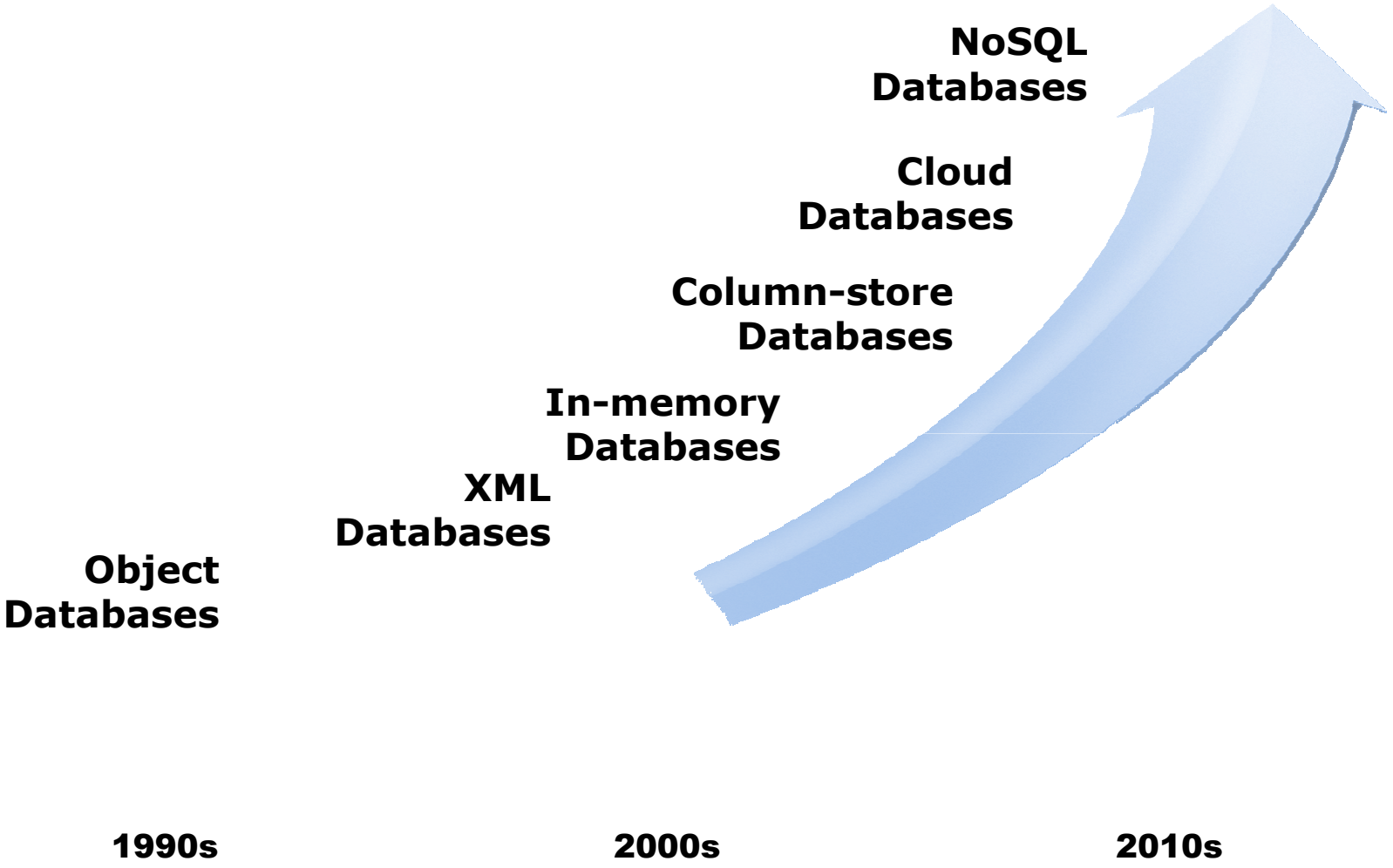


Strategy: Advanced Application Features

- Provide full DB2 Family compatible SQL for all key OLTP SQL requirements
- Provide new SQL features to simplify applications and improve performance
- Support new SQL functions required by key ISV applications
- Extend DB2 SQL language to simplify porting of apps to DB2 z/OS
- Support emerging mobile and NoSQL applications

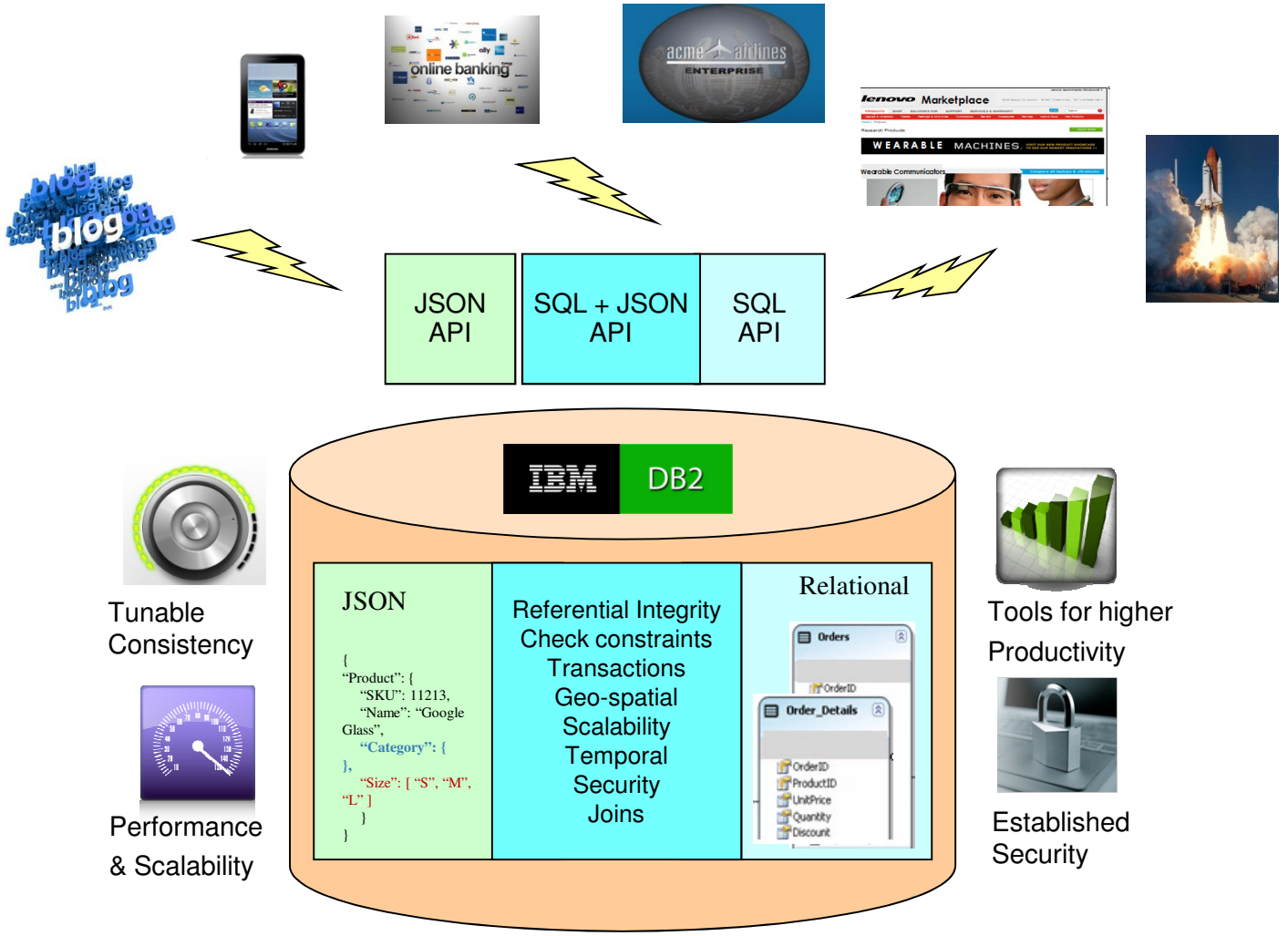


New Technology Emerges



JSON Database Technology Preview

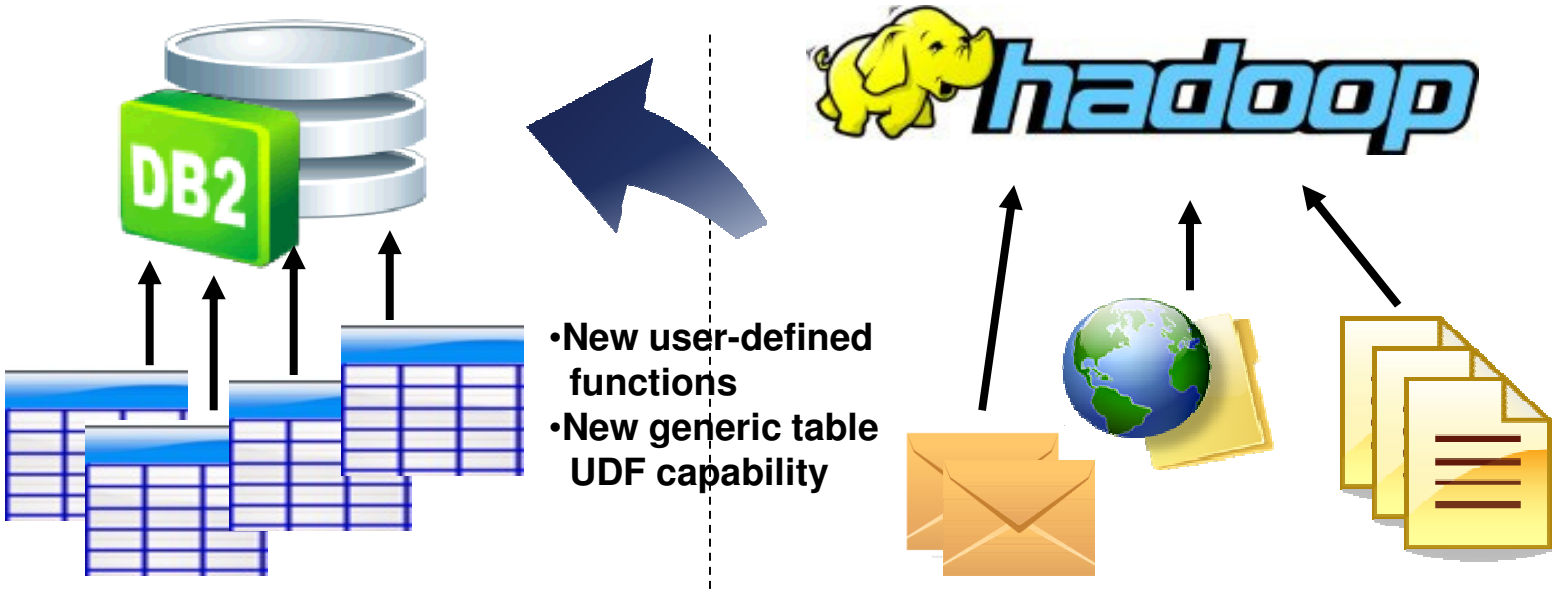
Providing the best of both worlds



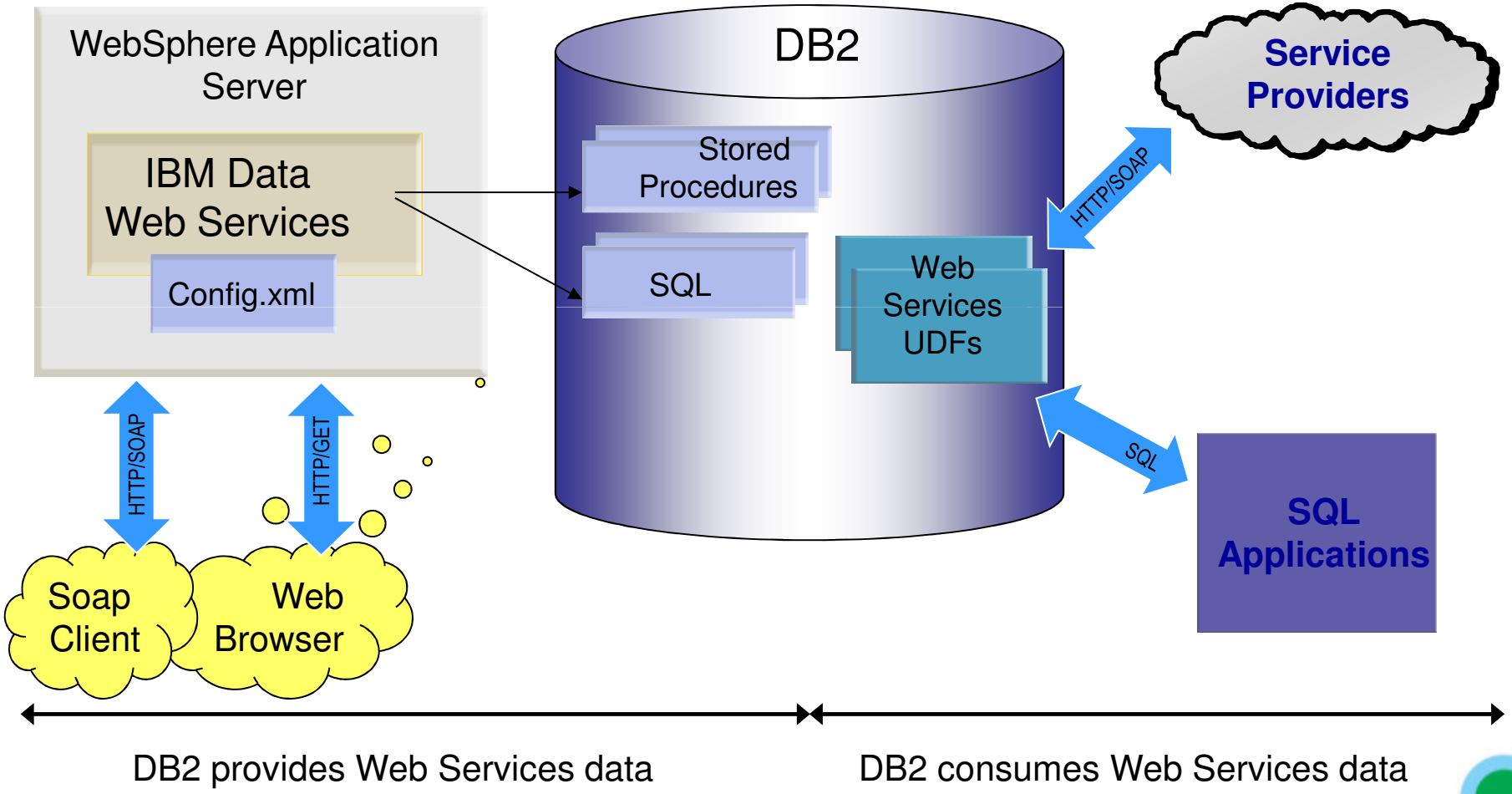
DB2 11: Big Data Integration

- Much of the world's operational data resides on z/OS
- Unstructured data sources are growing fast
- There is a need to integrate this data so that insights from BigData sources can drive business actions
- DB2 is providing the connectors and the database capability to allow DB2 applications to easily and efficiently Hadoop data in IBM BigInsights

Use of unstructured raw data is GROWING

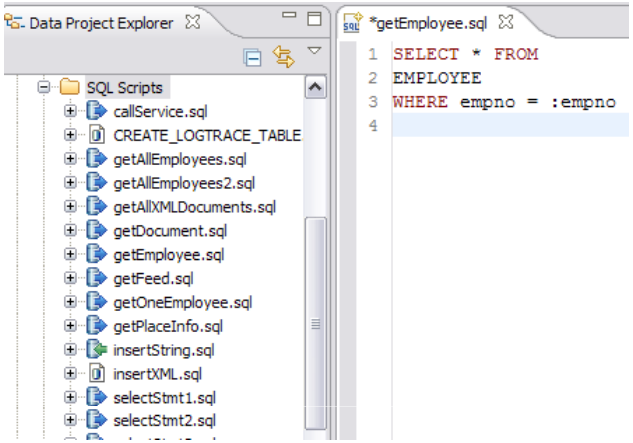


DB2 Web Services Overview (producer + consumer)

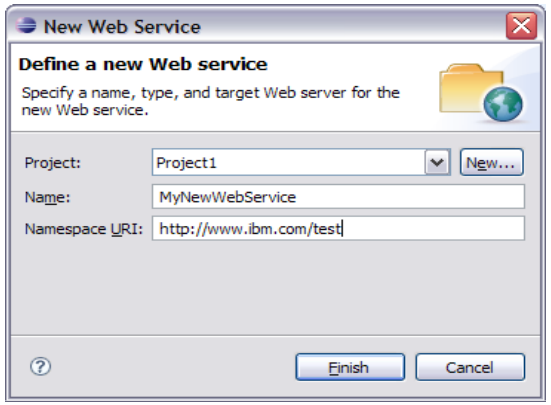


Optim Development Studio Tooling for Data Web Services

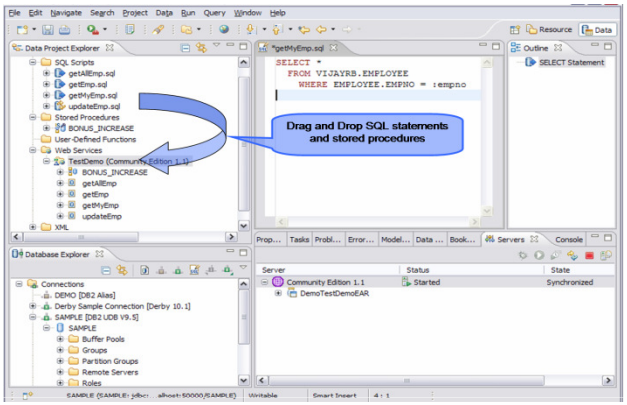
1. Develop Statements



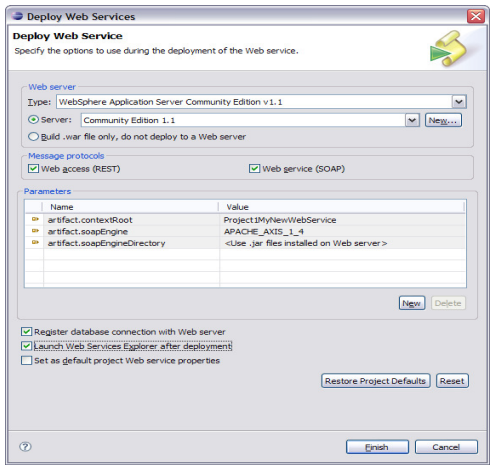
2. Create Service



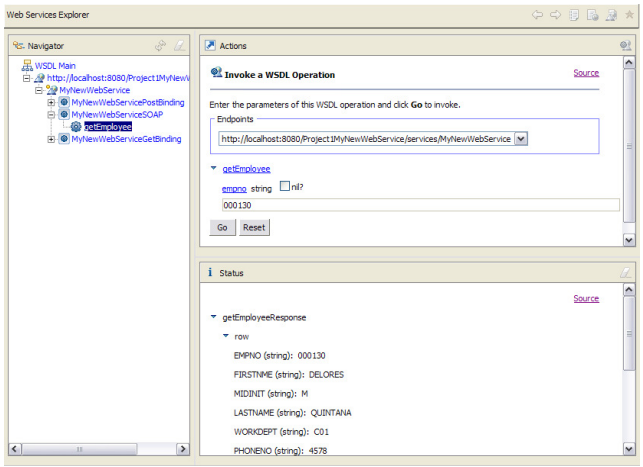
3. Drag 'n drop



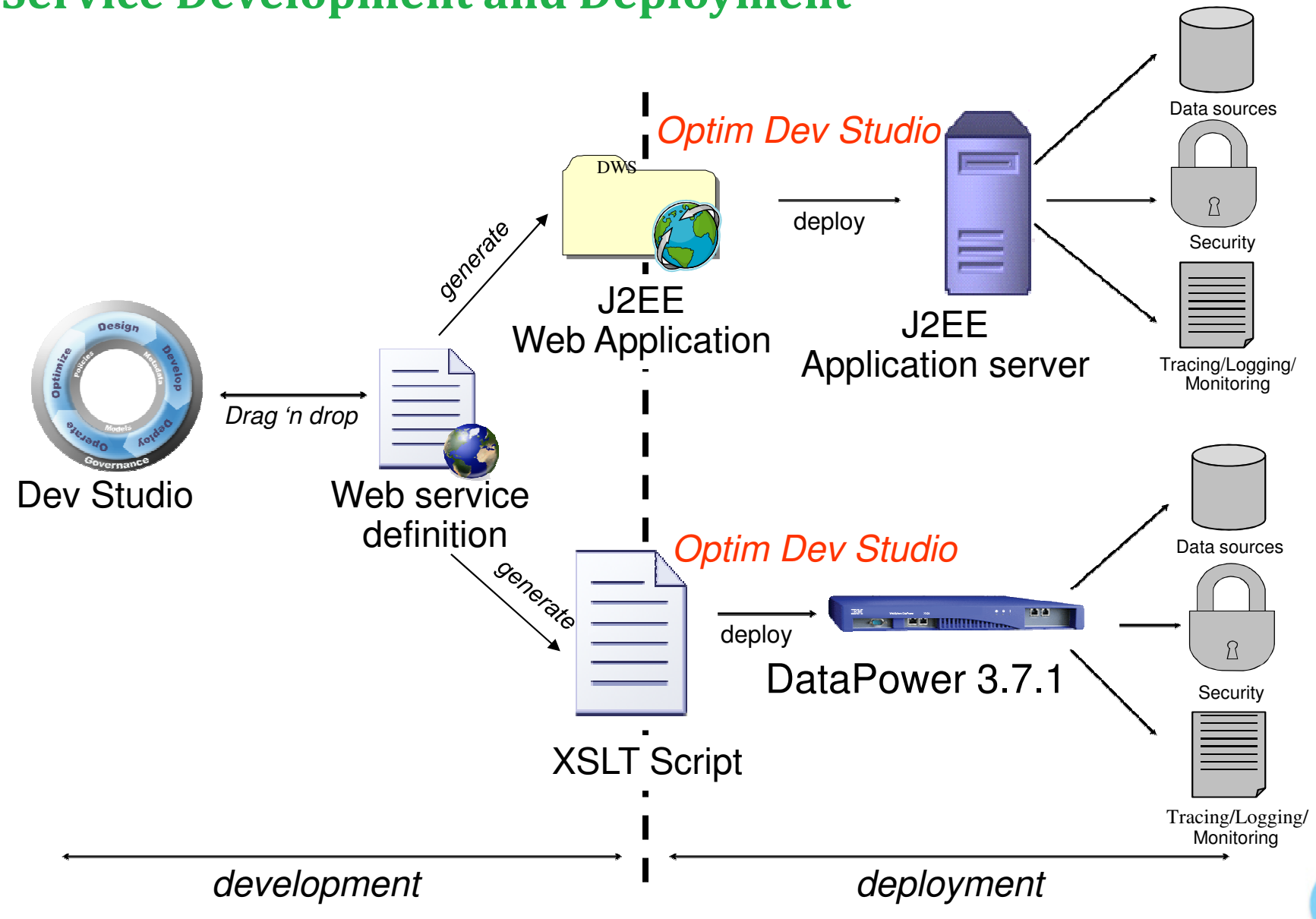
4. Deploy Service



5. Test Service

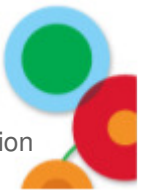


Service Development and Deployment



Strategy: Grow Analytics Capabilities

- Do not move the data!
  - Move analytics to data instead of moving data to analytics
  - Reduced copies of data
    - Simplification, less complexity
    - Cost savings
  - Run analytics against current data instead of older copies
  - Improved security by keeping data on z platform
- Provide a hybrid architecture for OLTP and analytics in the same database
- Exploit first class native DB2 z/OS query performance capabilities



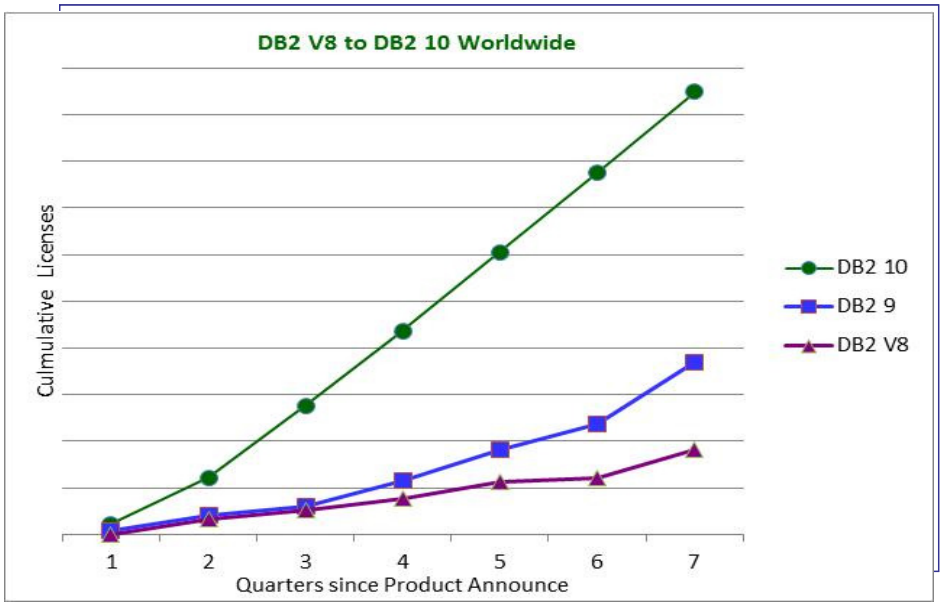
DB2 10 for z/OS

- CPU reductions for transactions, queries, and batch
  - Out-of-the-box CPU reductions of 5-10% for traditional workloads
  - Up to additional 10% CPU savings using new functions or avoiding constraints
  - Out-of-the box CPU reductions of up to 20% for new workloads
- Scales with less complexity and cost
  - 5-10x more concurrent users – up to 20,000 per subsystem
  - Significant scale-up capabilities in addition to existing scale-out support
  - Consolidate to fewer subsystems and LPARs
- Improved operational efficiencies and lower administration cost
  - Automatic diagnostics, tuning, and compression
- Even better performance
  - Elapsed time improvement for small LOBS and Complex Queries

- ***64 bit Evolution Virtual Storage Relief***
- ***Temporal Data***
- ***Integrated XML Support***
- ***Query Processing Enhancements***
- ***Business Security & Compliance***
- ***Better Productivity***

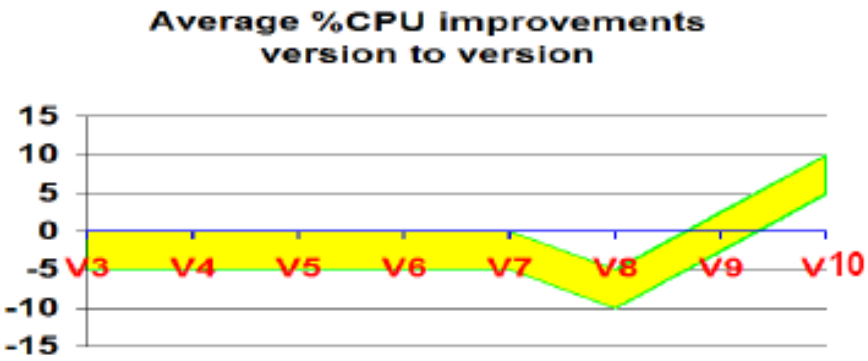
DB2 10 for z/OS Snapshot

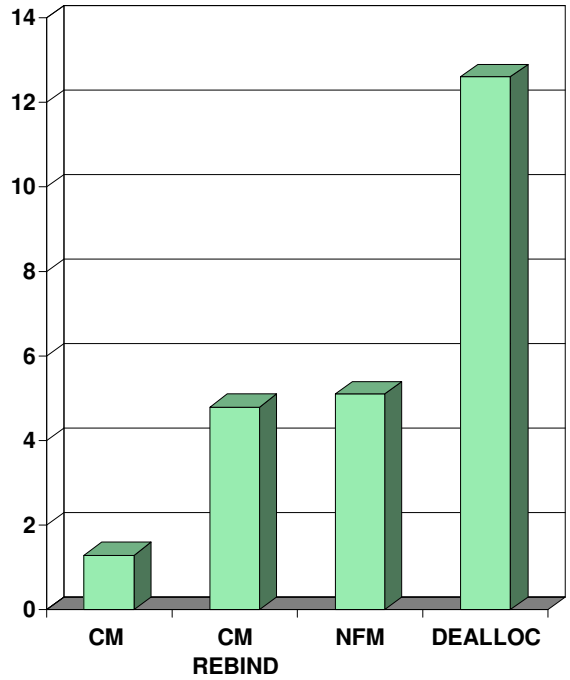
- Fastest uptake
  - +2x customers
  - +2x licenses
  - 25% coming from DB2 V8
- Adoption Driven by:
  - Performance improvements without application changes
  - DBM1 ASID 31-bit virtual storage constraint relief for more threads
  - Security, RAS improvements
  - Bitemporal data



DB2 10 Performance

- Most customers can see a 5% - 10% out-of-the-box CPU reduction (transactions and batch) after rebind
- Synergistic operation with latest System z hardware
- Measurements of IBM Relational Warehouse Workload (IRWW) with Data Sharing
  - Base: DB2 9 NFM REBIND with PLANMGMT EXTENDED
  - DB2 9 NFM → DB2 10 CM without REBIND showed 1.3% CPU reduction
  - DB2 10 CM REBIND with same access path showed 4.8% CPU reduction
  - DB2 10 NFM brought 5.1% CPU reduction
  - DB2 10 CM or NFM with RELEASE DEALLOCATE 12.6% CPU reduction from DB2 9

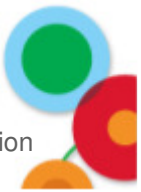




CPU Percent reduced from DB2 9

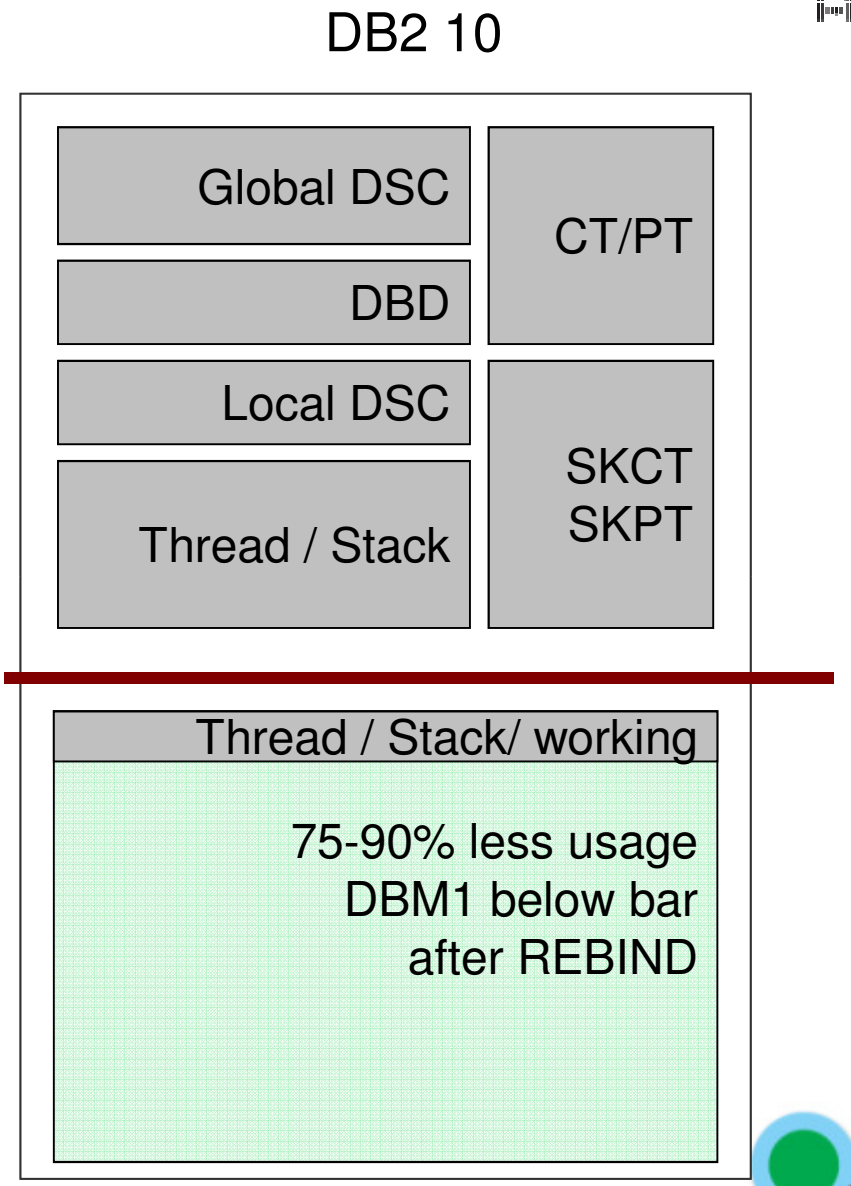
DB2 10: Some of the Interesting Performance Features

- CM:
  - High performance DBATs
  - Parallel index I/O at insert
  - Index list prefetch
  - SQLPL performance improvements
  - SQL pagination performance enhancement
  - 1M size real storage page frames for buffer pools
- NFM:
  - Hash access to data
  - Unique index include columns
  - Inline LOBs

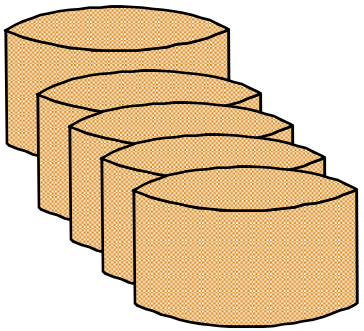


DB2 10 Virtual Storage Constraint Relief

- DBM1 below 2GB
  - 75-90% less usage in DB2 10 compared to DB2 9
  - Some of working storage (stack, xproc storage) stays below 2GB
- Larger number of threads
  - Possible data sharing member consolidation
- Reduce CPU with more real storage
  - More use of RELEASE(DEALLOCATE)
  - Larger MAXKEEPD values for KEEPDYNAMIC=YES



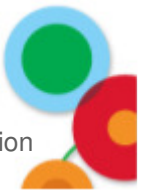
Changes in DB2 10 Catalog & Directory



- Improve availability and productivity
- Increase maximum size substantially
- Reduce contention: BIND, Prepare, Utilities
  - DDL concurrency also improved from removal of DBD01 hash anchor locks
- Catalog changes: Remove links, hashes
  - Many more table spaces, partition by growth
  - Row level locking, reordered row format
  - CLOB and BLOB columns for long strings
    - Inline for performance
  - Online reorganization and check
  - More automatic: DB2-managed SMS-controlled
  - Allow query of SYSLGRNX
  - Allow SQL statements in catalog to be queried with normal SQL

DB2 10 Continuous Availability: More Online Schema Changes

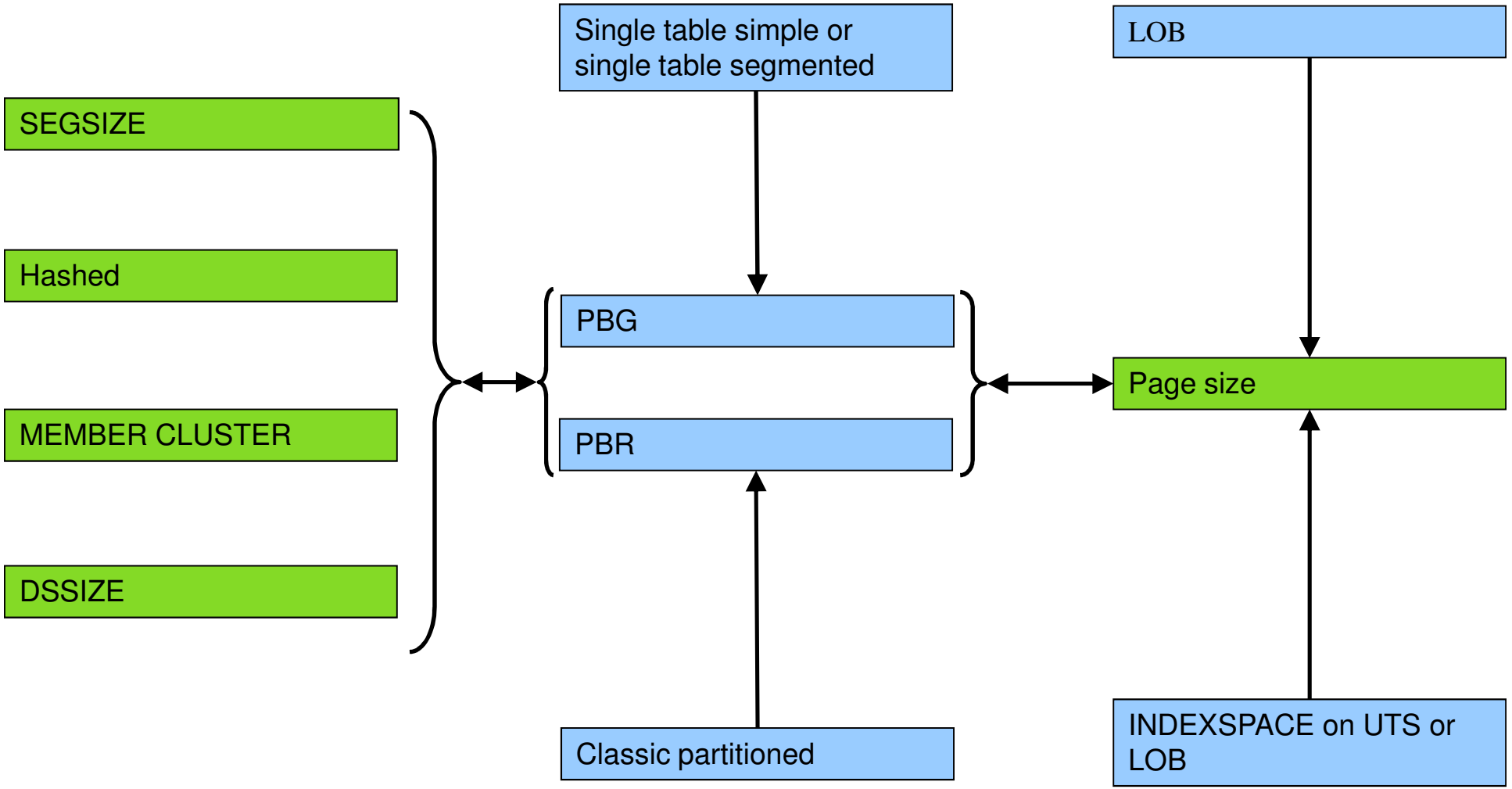
- Universal Table Space (UTS) type was introduced in DB2 9
  - Two types: Partition By Growth (PBG) and Partition By Range (PBR)
  - Combined advantages of the old segmented and classic partitioned table spaces
  - UTS is the strategic table space type going forward, many new functions are enabled for UTS only
  - But DB2 9 gives no easy way to convert existing TS's to UTS
  - DB2 10 allows for ALTER + Online REORG – easy, non disruptive
  - DB2 10 allows MEMBER CLUSTER for UTS
- DB2 10: Convert to Universal Table Space (UTS) with no outage
  - Single table simple -> UTS/PBG
  - Single table segmented -> UTS/PBG
  - Classic partitioned -> UTS/PBR



DB2 10 Continuous Availability: More Online Schema Changes …

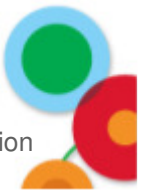
- DB2 10: Online ALTER of other UTS attributes
  - DSSIZE (also for XML or LOB table spaces)
  - Page size (also for LOB table spaces)
  - SEGSIZE (also for segmented or XML table spaces)
- DB2 10: Alter index page size for indexes on UTS's
- DB2 10: Convert PBG/PBR -> Hash
- DB2 10: Other schema change enhancements
  - Table space no longer needs to be stopped to alter MAXROWS
  - Object no longer needs to be stopped to alter BPOOL in data sharing

Online Schema – DB2 10

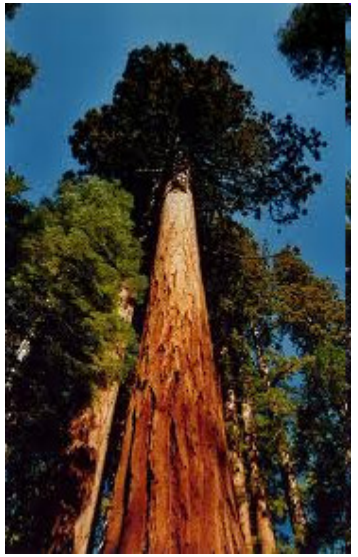


DB2 10 Other Availability Improvements

- Access currently committed data
- Change DDF location alias names online
  - New MODIFY DDF ALIAS command
  - New “dynamic alias” concept allows you to dynamically switch connections to an alias to different members
- Online DDF CDB changes
  - LOCATIONS, IPNAMES, IPLIST
- Dynamic add of active logs
  - New –SET LOG NEWLOG option
- Pre-emptable backout



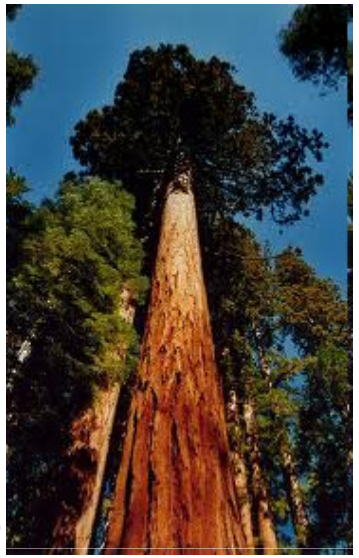
DB2 11 Major Themes



- **Performance Improvements**
  - Improving efficiency, reducing costs, no application changes
  - 0-5% for OLTP, 5-15% for update intensive batch
  - 5-20% for query workloads
  - Less overhead for data de-compression
  - Exploitation of new zEC12 hardware features
- **Continuous Availability Features**
  - Improved autonomics which reduces costs and improves availability
  - Making online changes without affecting applications
  - Online REORG improvements, less disruption
  - DROP COLUMN, online change of partition limit keys
  - Extended log record addressing capacity (1 yottabyte)
  - BIND/REBIND, DDL break into persistent threads

DB2 11 Major Themes …

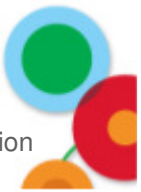
- **Enhanced business analytics**
  - Faster, more efficient performance for query workloads
  - Temporal and SQLPL enhancements
  - Transparent archiving
  - SQL improvements and IDAA enhancements
- **Simpler, faster DB2 version upgrades**
  - Decouple application changes required from DB2 release upgrade
  - Access path stability improvements
  - Product stability: support pre GA customer production



DB2 for z/OS 11
Early Support Program

- Started March 8
- 21 Customers in pre-production systems
  - 4 Customers running DB2 11 in New Function Mode
  - 5 Customers successfully tested DB2 11 migration with fallback to DB2 10
  - 12 successful migrations for DB2 11 compatibility mode
  - 1 customer successfully running DB2 10 and DB 11 in coexistence

A key design goal of DB2 11 is easier, faster upgrades

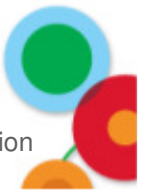


DB2 11 Some Planned RAS Improvements

- BIND / DDL concurrency with persistent threads
  - Use of persistent threads likely will increase in V10 with virtual storage constraint relief
- More online schema changes
  - Alter partitioning limit keys
  - DROP column
  - Point in time recovery support for some deferred schema changes
- REORG avoidance
  - Automatic cleanup of index pseudo deleted entries
- Online REORG improvements – if a REORG is needed, then the goal is non-disruptive
  - SWITCH phase performance improvements
  - Drain improvements to improve concurrency of log apply and switch phases
  - REORG REBALANCE SHRLEVEL(CHANGE)

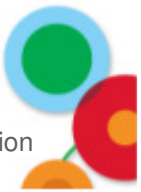
DB2 11 Some Planned DB2 11 RAS Improvements …

- Open data set limit raised to 200K
- Cancel DDF Threads – new FORCE option
- Easier query performance management
  - Optimizer input to RUNSTATS
- DB2/RACF authorization control enhancements
- Buffer pool management improvements
  - New FRAMESIZE BP attribute for direct control of z/OS large page frame usage
  - Max/min size for WLM system-managed BPs



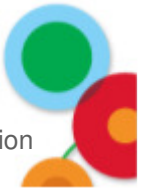
DB2 11 – Other Planned Availability Enhancements

- BACKUP SYSTEM & RESTORE SYSTEM improvements
- Data availability improvements in REORG
- Faster recovery from inline image copies
- Faster recovery of catalog & directory
- Improved recoverability for long-running REORGs
- Remove some recoverability restrictions associated with online schema
- Improved data integrity checking on pageset open



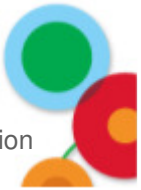
Extended RBA Problem Statement

- DB2's Relative Byte Address (RBA) for logging is 6 bytes
  - Gives 256TB of log record addressing capacity per DB2 subsystem/member
- With heavy sustained logging rates, DB2 can exhaust the 6-byte RBA
  - DSNJ032I and DSNJ033E warning messages
    - Alert-level = 'WARNING' when RBA reaches x'F00000000000'
    - Alert-level = 'CRITCIAL' when RBA reaches x'FFFF00000000'
  - Manual recovery actions are needed
    - Data Sharing: shut down the affected member and start a new member in its place
    - Non Data Sharing: reset all PGLOGRBA values back to zero (extended outage)
    - Documented in the DB2 Administration Guide
  - If alert-level reaches 'CRITICAL' then DB2 terminates to protect data integrity and force recovery actions
    - Reason code 00D10251
    - ACCESS(MAINT) restart allowed to prepare for recovery actions



Extended LRSN Problem Statement

- The data sharing Log Record Sequence Number (LRSN) is derived from the 8-byte time-of-day clock which hits end of range in 2042
- However, some data sharing groups have a non-zero LRSN "delta" which gets added to the TOD clock
  - If a non-zero "delta" exists, then the LRSN will hit end of range prior to 2042
  - Use DSNJU004 to determine if you have a non-zero LRSN delta value
  - A "delta" value could be set when data sharing is enabled or re-enabled
    - Whenever the end-of-log RBA of the enabling member is past the TOD clock
- Some non data sharing customers have enabled data sharing to circumvent RBA nearing end-of-range
  - This would cause a non-zero LRSN delta, so LRSN hits end of range before 2042
- 6-byte LRSN value has precision to only 16 microseconds
  - Can cause LRSN 'spinning' which burns extra CPU and aggravated log latch contention
  - V9 NFM addresses most LRSN spin situations, and V10 NFM enhances further. But some spins still exist due to the 16 usec granularity (log latch not held, page latches are)

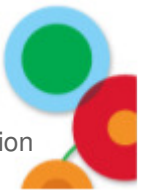


DB2 11 Planned Solution for Extended RBA/LRSN

- Expand the RBA and LRSN to 10 bytes
  - RBA addressing capacity of 1 yottabyte (2**80)
  - LRSN extended on left by 1 byte, on the right by 3 bytes
    - >30,000 years and 16Mx more precision
  - 8 bytes is not sufficient to solve LRSN issues and may not give sufficient capacity for the longer term
- NFM only (6 byte RBA/LRSN continues to be used in CM)
- Once in NFM, DB2 continues to use 6-byte values until you take action to convert
- Two conversion tasks – total flexibility in order of implementation
  - Convert BSDSes to new format to enable logging with larger RBAs/LRSNs
  - Convert pagesets to new page format via REORG
- These tasks are optional
  - If you don't care about larger RBAs/LRSNs then you don't have to convert
  - But performance will be better if you convert BSDSes (avoid internal conversion overhead on log write and LRSN spin)
- BSDSes can be converted without converting pagesets
- Pagesets can be converted in a piecemeal fashion
  - Expectation is that most customers will roll the conversion over a period of weeks to months

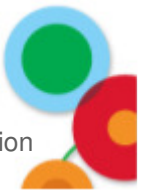
Online ALTER Partition Limit Keys

- Currently: Affected partitions are set to REORP
  - These partitions cannot be accessed
  - REORG is run to redistribute the data and remove the status
- DB2 11: Alter limit key is treated as a pending alter (NFM)
  - The affected partitions are set to AREOR
  - Online REORG **must** be run to materialise the pending changes
  - Supported table spaces types are:
    - UTS – partitioned by range (PBR)
    - Classic partitioned table spaces (table controlled partitioning)
  - The new limit keys are materialized in SYSTABLEPART in the SWITCH phase



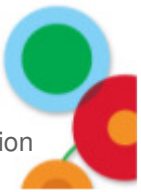
DB2 11 Some Performance Improvements

- Suppress-null indexes
  - Index entries not created when all values for indexed columns are NULL
  - Reduced index size, improved insert/update/delete performance, Oracle compatibility
  - Improved utility CREATE INDEX performance
- DDF performance improvements
  - Reduced SRB scheduling on tcp/ip receive using new CommServer capabilities
  - Improved autocommit OLTP performance
  - DRDA package based continuous block fetch
- xProcs above the bar
  - 31-bit virtual storage constraint relief enabled by RMODE 64 support in z/OS 1.13 and above
  - Enables other internal performance improvements



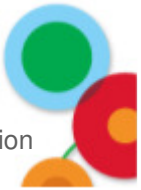
DB2 11 Some Performance Improvements …

- DSCF/IFC enhancements
  - Move WWFR, CCB above the bar
- zIIP enablement for all SRB-mode DB2 system agents that are not very response time critical
- Avoid cross-memory overhead for writing log records
- INSERT performance
  - Latch contention reduction for classes 6, 14, 19
  - CPU reduction for Insert column processing and log record creation
  - Page fix/free avoidance in GBP write

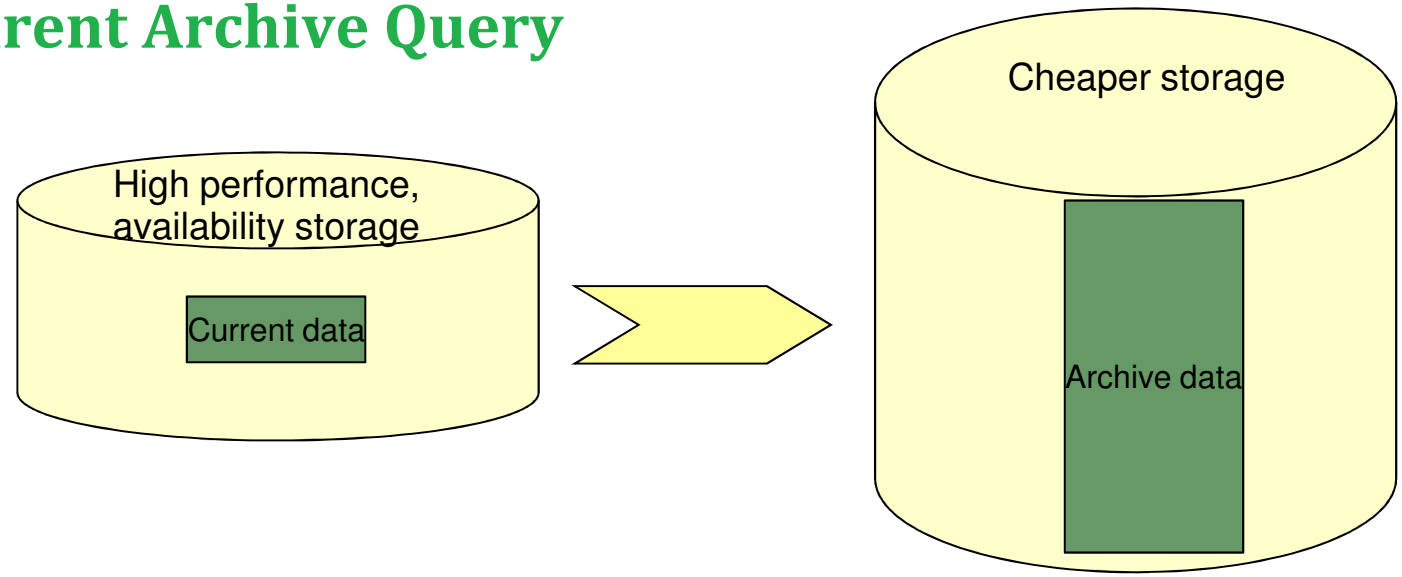


DB2 11 Some Planned New Application Features

- Global variables
- SQLPL improvements (performance, manageability, function)
  - Autonomous transactions
  - Array data type support
- Alias/synonym support for Sequence Objects
- Grouping sets
- DPSI performance improvements
- DGTT performance improvements
- Temporal data enhancements
  - Support for views
  - Special register support
- Transparent archive query
  - New DDL to relate the current table to the archive
  - Applications can query current + archive with no SQL changes
- Enhancements for DB2 and Big Data integration



Transparent Archive Query



- Applications can query current + archive with no SQL changes
  - By default, data is retrieved from base table only, as usual
  - Set a new global variable when archive data is desired
  - DB2 automatically converts SQL to UNION ALL via dynamic plan switching technique (high performance)
- Archiving process is user-controlled
- Move_To_Archive global variable allows DELETEs to be automatically archived
- Leverages DB2 10 temporal constructs for archiving use cases
- Future potential for more IDAA synergy

DB2 11 Planning

- Dual mode migration (CM, ENFM, NFM)
- Migration from DB2 10 NFM only (no skip)
- z/OS 1.13 or above
- z10 or above
- No pre-V9 bound packages
- Sysplex query parallelism support is removed