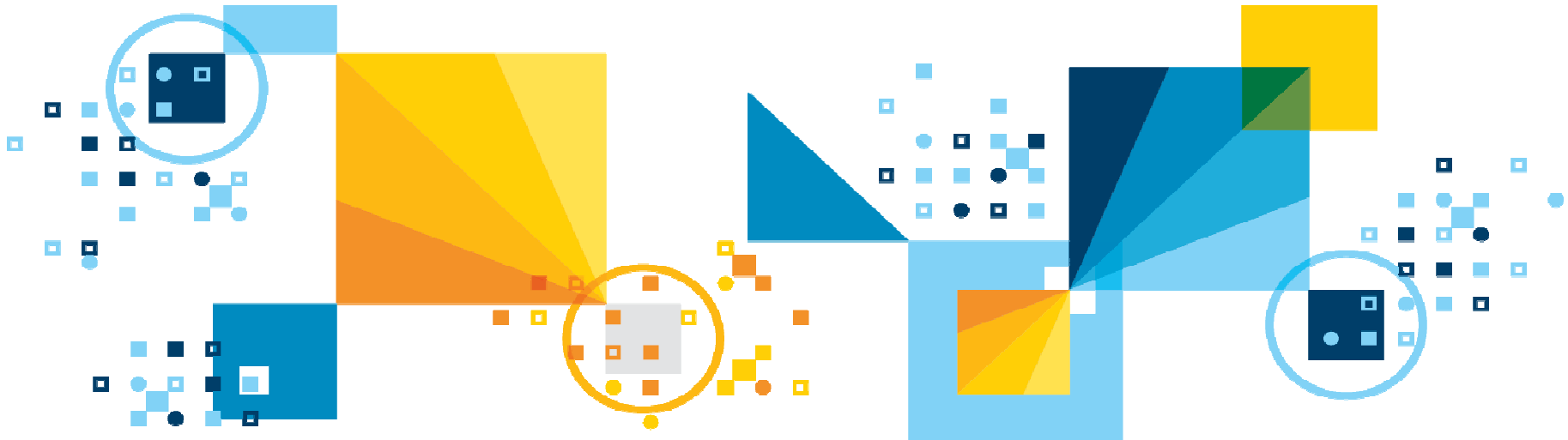


# DB2 LUW Version 11.1

## A Technical Overview

**Matt Huras, IBM Distinguished Engineer and DB2 Chief Architect ([huras@ca.ibm.com](mailto:huras@ca.ibm.com))**  
**Berni Schiefer, IBM Fellow ([schiefer@ca.ibm.com](mailto:schiefer@ca.ibm.com))**



# Safe Harbor Statement

*Copyright © IBM Corporation 2016. All rights reserved.*

*U.S. Government Users Restricted Rights - Use, duplication, or disclosure restricted by GSA ADP Schedule Contract with IBM Corporation*

**THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON CURRENT THINKING REGARDING TRENDS AND DIRECTIONS, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. FUNCTION DESCRIBED HEREIN MY NEVER BE DELIVERED BY I BM. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS AND/OR SOFTWARE.**

IBM, the IBM logo, ibm.com and DB2 are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)

# DB2 Version 11.1 Highlights ( GA: June, 2016)

## Core Mission Critical Workloads : Extending DB2 Leadership

### Comprehensive Enterprise Security



#### Enterprise Encryption

- Centralized Key Managers (KMIP)

### Availability 2<sup>nd</sup> only to DB2 for zOS



#### Simple Fast Deployment

- Up and running in hours

#### Even Greater Availability

- Zero data loss DR with HADR
- More online management

#### More Platforms Supported

- Power Linux (LE)
- Virtualization for RDMA (x86)

### Significant Core Database Advances



#### Very Large Database Performance

- Higher user throughput

#### Simpler, Faster, More Online Upgrades

- Faster, no need for offline backup
- Streamlined HADR upgrade
- DB2 Version 9.7 direct to 11.1

## Warehousing Workloads : Most Consumable, Most Scalable In-Memory Warehousing Platform

### Massive Scale Warehousing at In-Memory Performance



#### MPP BLU Scalability

- PB scale in-memory warehousing

### Next Gen In-Memory Performance, Function & Workloads



- **Faster ELT/ETL performance**
- **More Query Workloads Optimised**
- **More Function supported**
  - Generated Columns
  - RCAC
  - OLAP + BLU Perf

### Enhanced Compatibility



#### Multi-Lingual SQL Advances

- Postgres SQL

#### Support for European Languages

- Codepage 819

# DB2 Version 11.1 Highlights ( GA: June, 2016)

## Core Mission Critical Workloads : Extending DB2 Leadership

### Comprehensive Enterprise Security



#### Enterprise Encryption

- Centralized Key Managers (KMIP)

### Availability 2<sup>nd</sup> only to DB2 for zOS



#### Simple Fast Deployment

- Up and running in hours

#### Even Greater Availability

- Zero data loss DR with HADR
- More online management

#### More Platforms Supported

- Power Linux (LE)
- Virtualization for RDMA (x86)

### Significant Core Database Advances



#### Very Large Database Performance

- Higher user throughput

#### Simpler, Faster, More Online Upgrades

- Faster, no need for offline backup
- Streamlined HADR upgrade
- DB2 Version 9.7 direct to 11.1

## Warehousing Workloads : Most Consumable, Most Scalable In-Memory Warehousing Platform

### Massive Scale Warehousing at In-Memory Performance



#### MPP BLU Scalability

- PB scale in-memory warehousing

### Next Gen In-Memory Performance, Function & Workloads



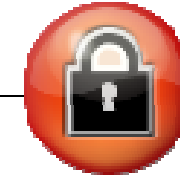
- Faster ELT/ETL performance
- More Query Workloads Optimised
- More Function supported
  - Generated Columns
  - RCAC
  - OLAP + BLU Perf

### Enhanced Compatibility



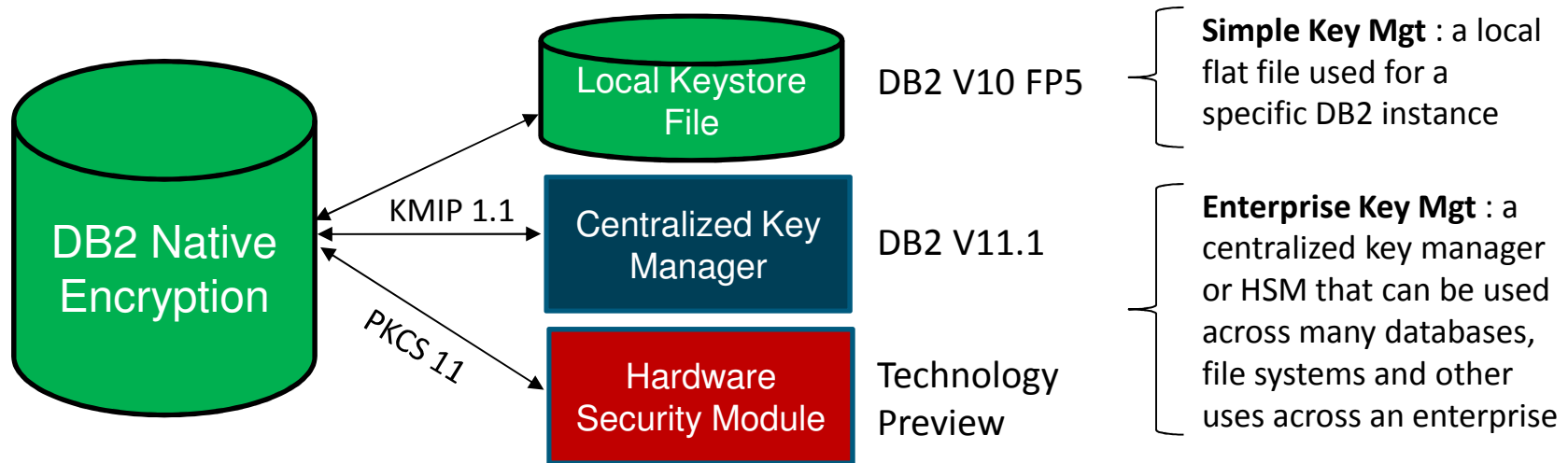
#### Multi-Lingual SQL Advances

- Postgres SQL
- Support for European Languages
- Codepage 819



# Encryption and Enterprise Key Management

- V11.1 adds support for KMIP 1.1 compliant centralized key managers**
  - Initially validated on IBM Security Key Lifecycle Manager (ISKLM)



## Example configuration changes

### Local Keystore

```
update dbm cfg using keystore_type pkcs12
update dbm cfg using keystore_location /home/thomas/keystores/localkeystore.p12
```

### Centralized KMIP Key Manager

```
update dbm cfg using keystore_type kmip
update dbm cfg using keystore_location /home/thomas/keystores/isklm.cfg
```

*Indicates KMIP protocol will be used to interact with key manager*

*Location of configuration file containing host, port and other details of centralized key manager*

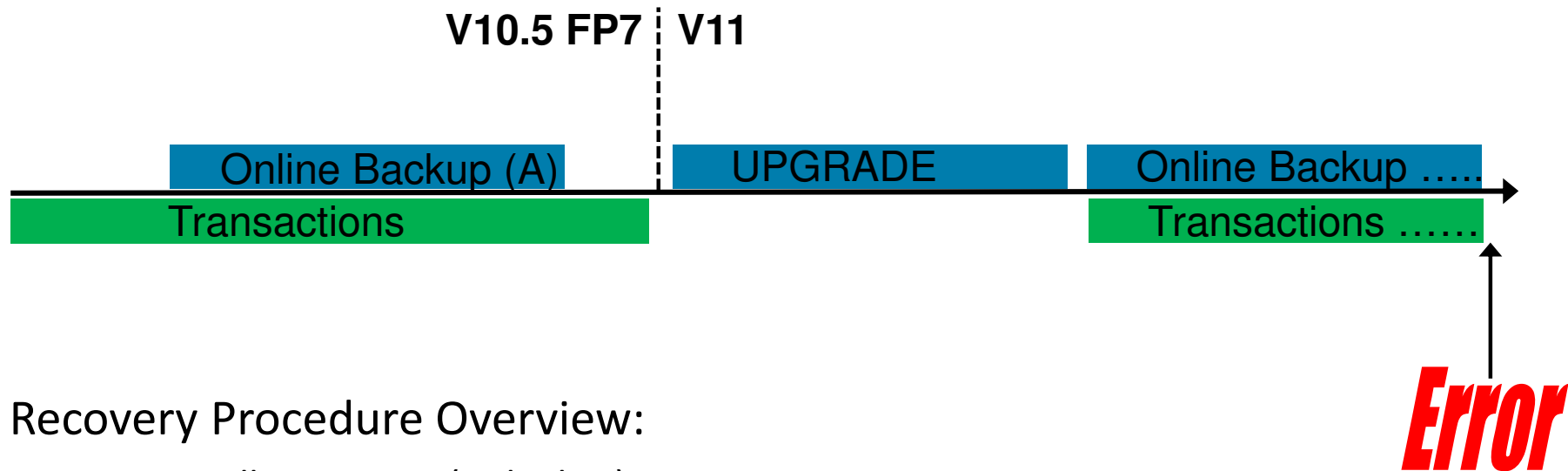
## Streamlined Upgrade Process

- **Upgrade directly from Version 9.7, 10.1 and 10.5 (3 releases back)**
- **Ability to roll-forward through database version upgrades**
  - Users are no longer required to perform an offline backup of existing databases before or after they upgrade
  - A recovery procedure involving roll-forward through database upgrade now exists
  - Applies to all editions and configurations except Database Partitioning Feature (DPF)
  - Pre-req: must start from DB2 Version 10.5 Fix Pack 7, or later
- **HADR environments can now be upgraded without the need to re-initialize the standby database after performing an upgrade on the primary database**
  - Applies to all editions except pureScale
  - Pre-req: must start from DB2 Version 10.5 Fix Pack 7, or later



# UPGRADE without Offline Backup

- No more need to take an offline backup to ensure recoverability across upgrade !
- Example scenario :

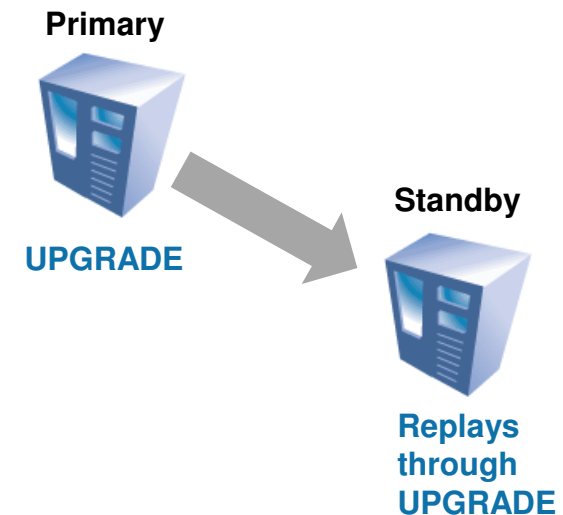


- Recovery Procedure Overview:
  1. Re-install V10.5 FP7 (or higher)
  2. Restore online backup (A)
  3. Rollforward to a desired point-in-time just before the Error
    - Receive SQL2463N or SQL2464N indicating the start of upgrade
  4. Re-install V11
  5. Continue Rollforward

# Streamlined HADR Version Upgrade

- No more need to re-initialize the standby database during version upgrades
  - Saves significant time & effort especially with large and numerous databases
- Procedure overview :

1. **PRIMARY :**  
DEACTIVATE database; DB2STOP instance  
Upgrade the instance using db2iupgrade
2. **STANDBY :**  
DEACTIVATE database; DB2STOP instance  
Upgrade the instance using db2iupgrade
3. **STANDBY : UPGRADE the database**
  - Returns successful – this indicates that the standby is now waiting for log data from a subsequent UPGRADE issued on the primary
4. **PRIMARY : UPGRADE the database**
  - Will ship log data to standby; Standby replays these log records
5. **PRIMARY : ACTIVATE the database**
  - Now primary and standby can resume normal operations

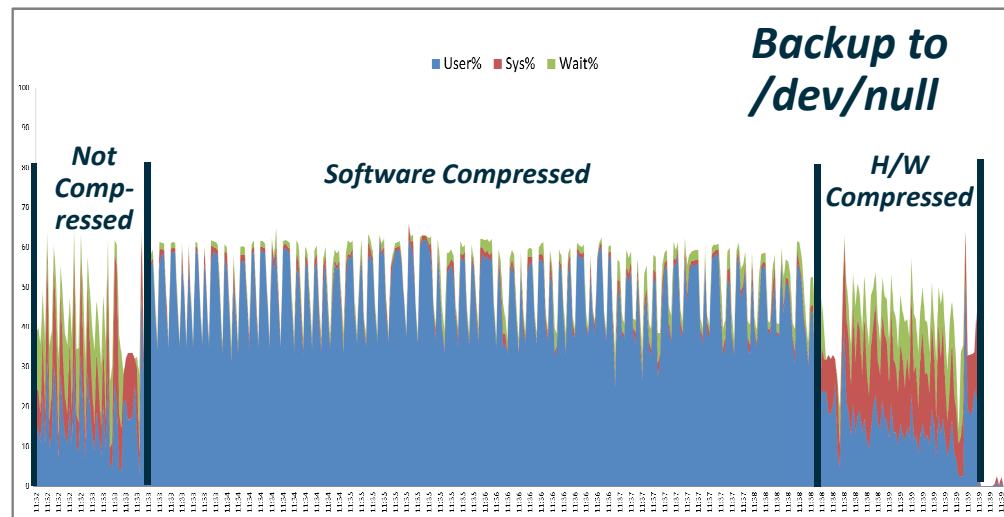
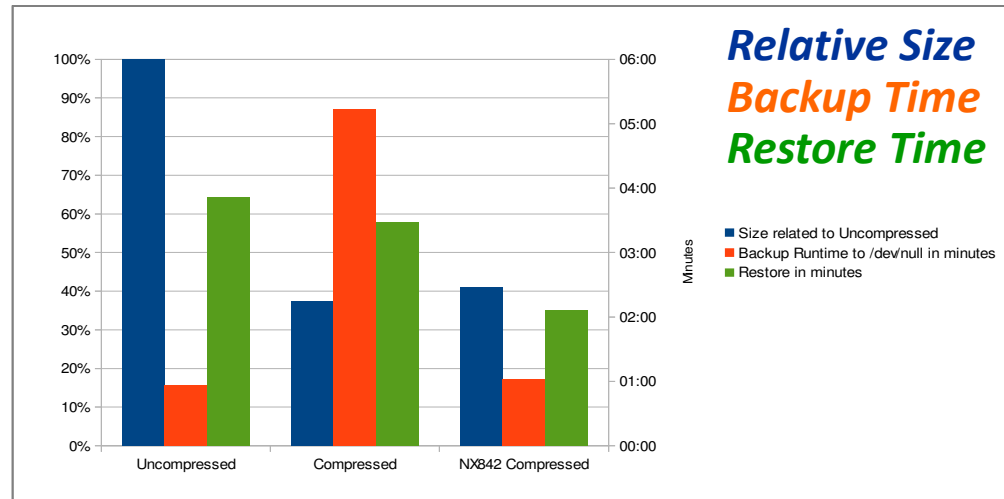




# Backup and Log Compression Acceleration

## Hardware Acceleration on POWER 7+, Power 8

- Concept : offload compression to ‘nx842’ hardware compression accelerator**
  - Accelerator present on Power 7+, Power 8
  - Used for AIX “Active Memory Expansion” compression
- Benefits :**
  - Very significant reduction in CPU consumption
  - Very significant reduction in elapsed time
  - Maintains almost all of the compression storage benefits



# Backup and Log Compression Acceleration

## *Details*

- **Backup/Restore**

- Invocation

```
BACKUP DATABASE mydbname COMPRESS COMPRLIB libdb2nx842.a
```

```
- or -
```

```
db2set DB2_BCKP_COMPRESSION=NX842
```

```
BACKUP DATABASE mydbname COMPRESS
```

- Restore automatically decompresses
- Automatic reversion to software compression if accelerator not present

- **Log Archival**

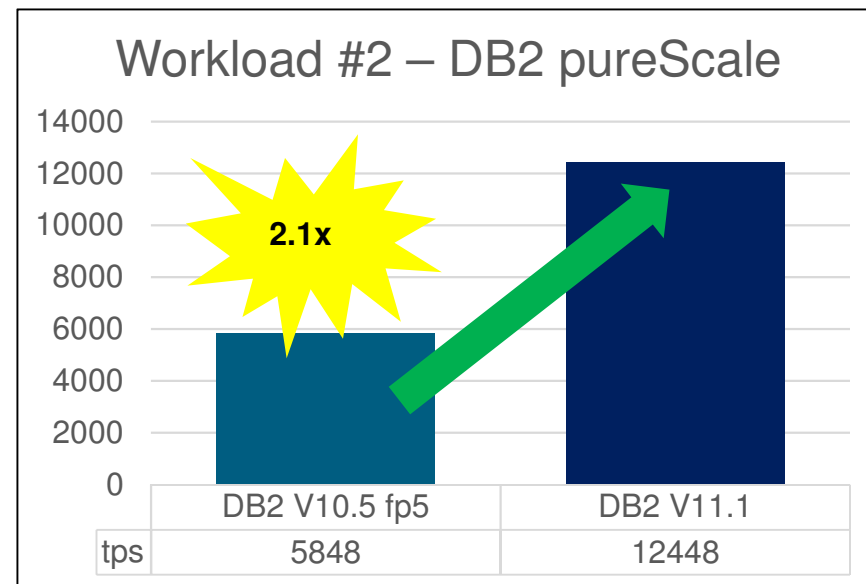
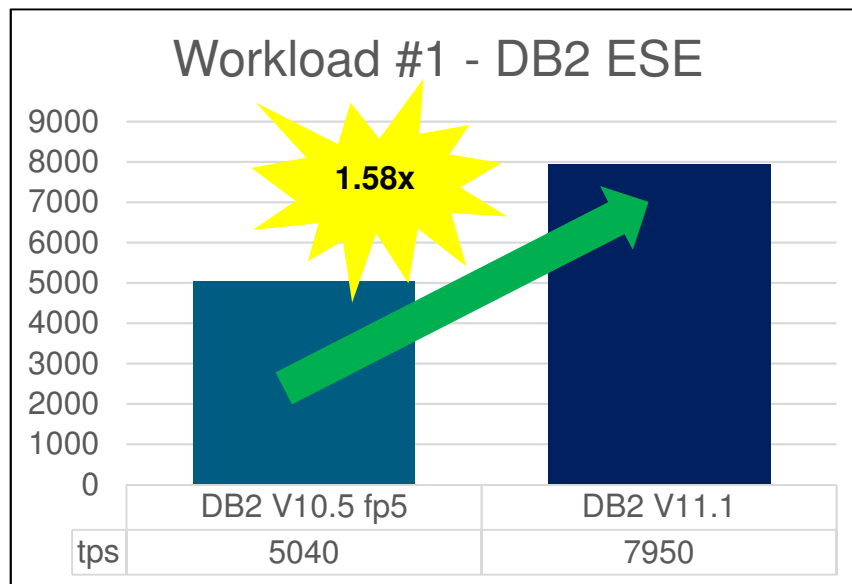
- Invocation

```
UPDATE DB CFG FOR mydbname USING LOGARCHCOMPR1 nx842
```

- Retrieve automatically decompresses
- Automatic reversion to software compression if accelerator not present

# Improved Performance for Highly Concurrent Workloads

- **V11 revamps DB2’s internal bufferpool latching protocol**
  - Significantly reduces contention
  - Benefits most pronounced on high concurrency transactional workloads



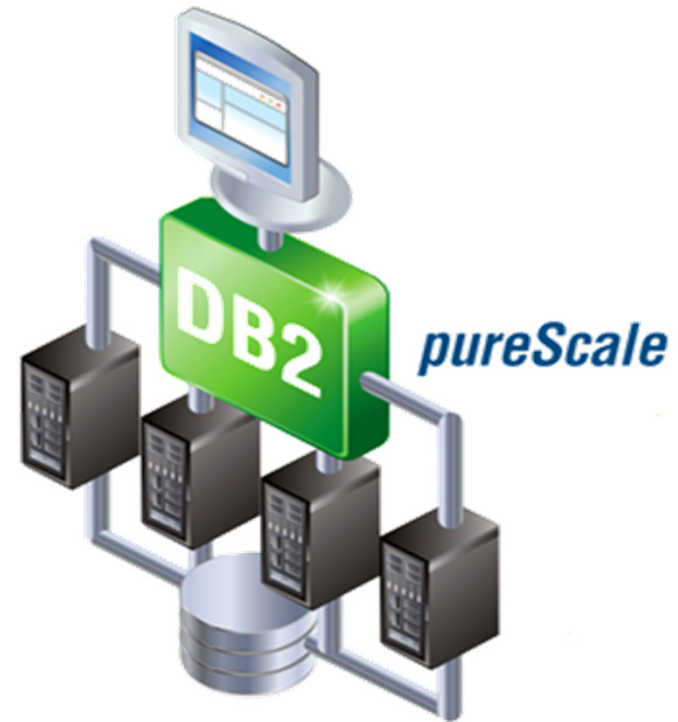
- Workload 1 based on an industry benchmark standard
- POWER7 32c, 512 GB

- Workload 2 implements a warehouse-based transactional order system
- 4 members, 2 CFs with 16c, 256 GB

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

## DB2 V11 pureScale Feature Enhancements

- Easier Installation and 'Up and Running'
- Power Linux Little-Endian (LE) support
- Linux Virtualization Enhancements
- HADR and GDPC Enhancements
- Performance Enhancements
- Increased Workload Balancing Flexibility
- Manageability Improvements
- Text Search support



# DB2 pureScale: Simplified Install and Deployment

## ■ Fast Up and Running

- Up and running in hours
- Increased advantage relative to competitive cluster databases

## ■ Install re-engineering includes:

- “**Push-Button**” install for pureScale clusters
  - Install complexity **reduced** by at least **40%** (sockets) or **25%** (RDMA)
  - **Smarter defaults**, intuitive options, parallel & quick pre-deployment validation across hosts
- 30-step native GPFS replication setup reduced to **simple 4-step process**
  - Also simplified conversion to GPFS replication post install via db2cluster
- **Increased Resiliency** for aborted/partial installations, **clean rollback for clean re-execution**

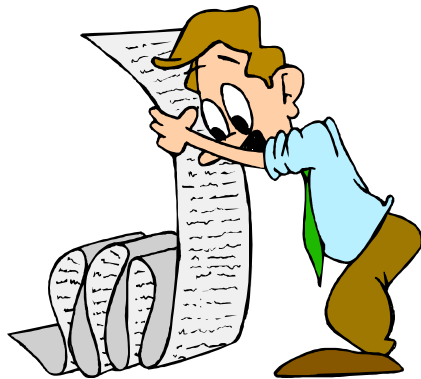
## ■ Additional assistance via:

- **Simplified** documentation
- **Enhanced** pre-checking of storage, elimination of manual verification of firmware libraries
- **Increased intuitive** and **user-friendly** errors & warnings



# Simplified pureScale Storage Replication Deployment

**Today**



**V11**



1. Takes ~8 native GPFS commands to create a replicated file system with the standard three redundancy groups.
2. Takes ~24 native GPFS commands to convert a non-replicated FS to a replicated FS.
3. Takes ~8 native GPFS commands to add a new disk to an existing replicated FS
4. Takes ~7 native GPFS commands to remove a disk from existing replicated FS

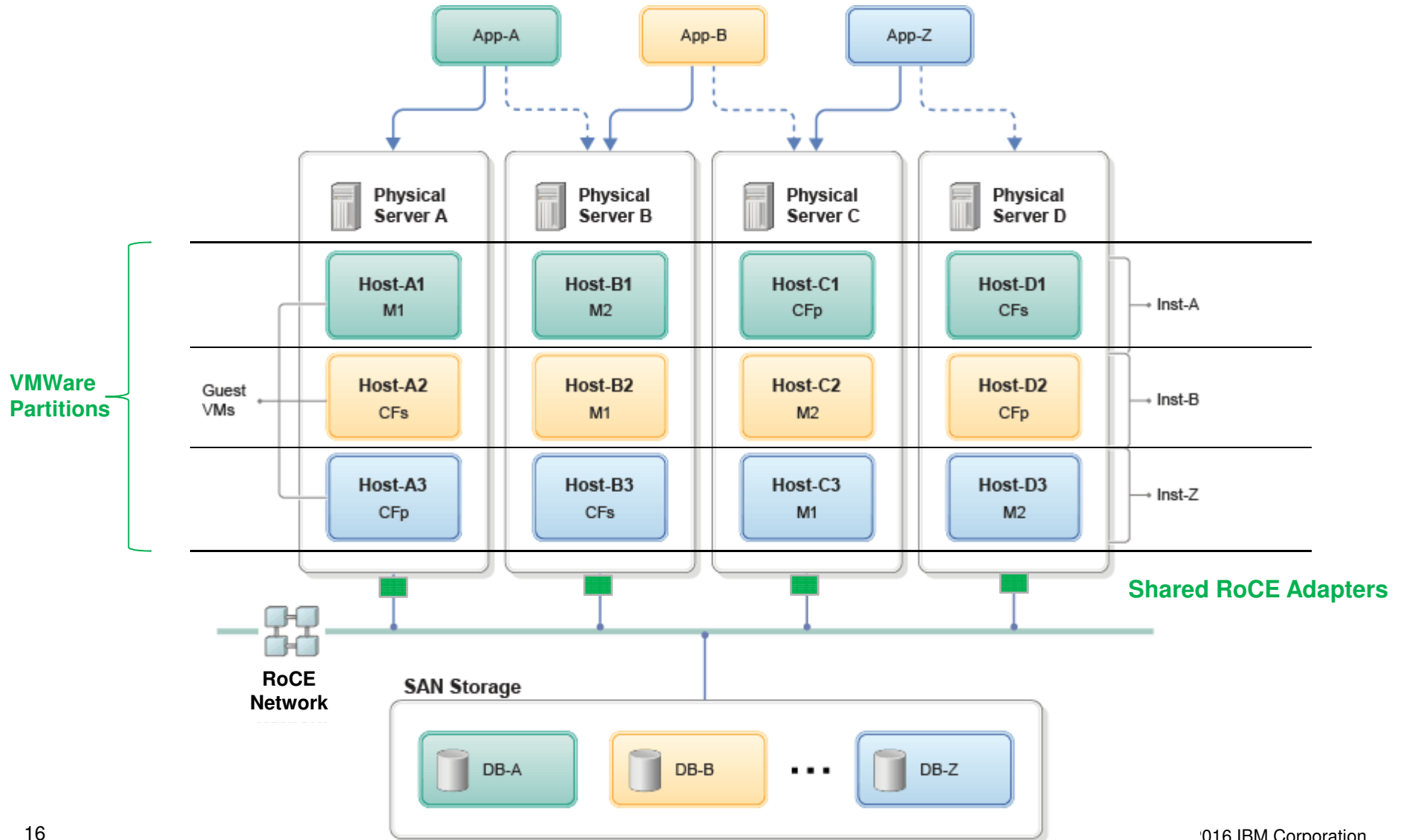


1. 1 db2cluster command
2. 2 db2cluster commands (one for conversion and one for adding storage)
3. 1 db2cluster command
4. 1 db2cluster command

## Linux Virtualization Enhancements

- **Virtualized access to RDMA over Converged Ethernet (RoCE) adaptors**
  - Via RoCE Single-Root I/O Virtualization (SR-IOV) support
  - Allows a single adapter to be shared across multiple VMWare partitions
  
- **What is “Single-Root I/O Virtualization” (SR-IOV) ?**
  - Standard that enables one PCI Express (PCIe) adapter to be presented as multiple separate logical devices to virtual machines
  - Allow the virtual machines to run native RoCE and achieve near wire speed performance
  - Can be enabled on Mellanox ConnectX-3/ConnectX-3 Pro/Connect X-3 VPI adapters for Ethernet

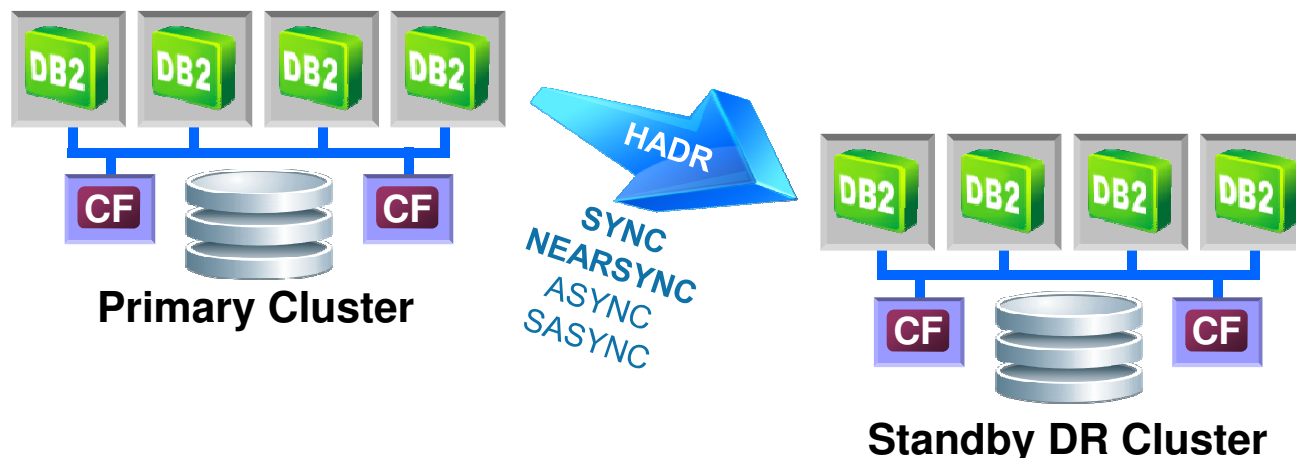
# VMWare RoCE Adapter Sharing Example



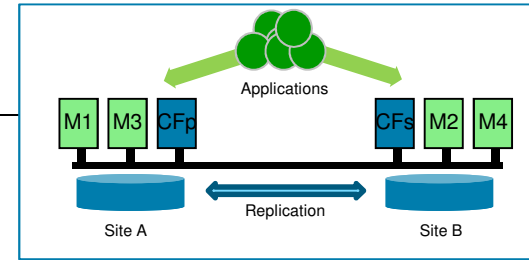


## HADR Support for SYNC and NEARSYNC Mode

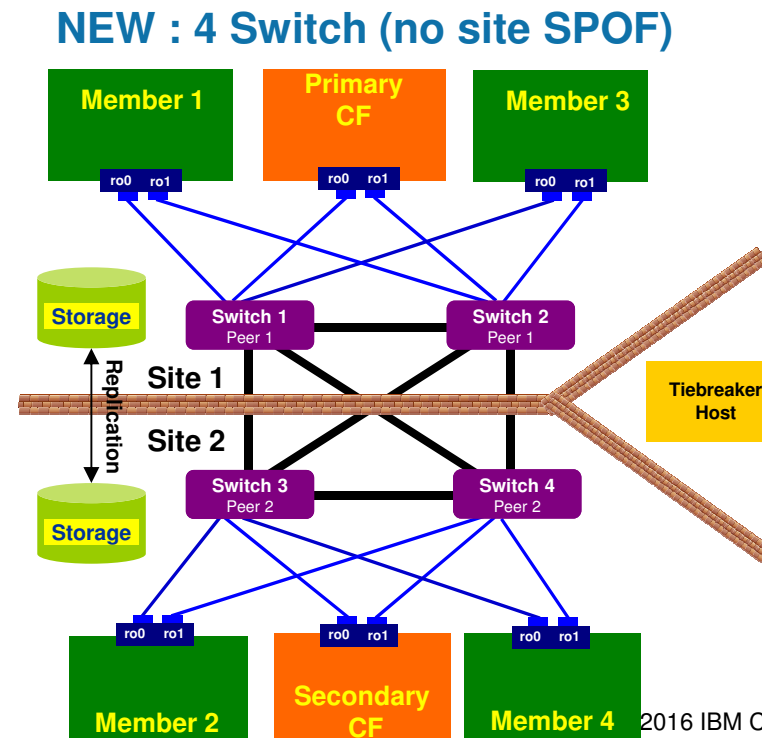
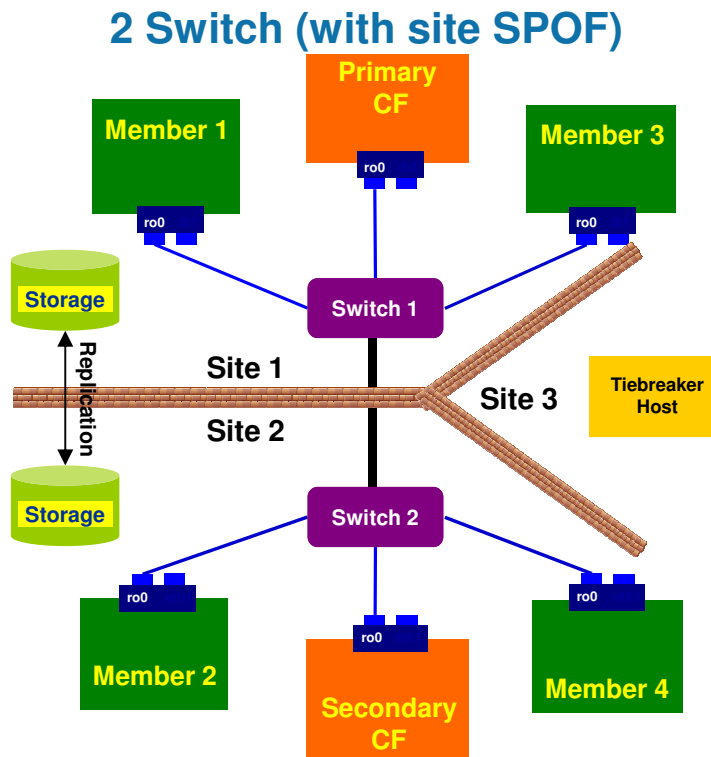
- Combines pureScale and HADR to provide a near continuously available system with robust RPO=0 disaster recovery
- Other capabilities & enhancements include
  - HADR rolling update supported, eg.
    1. **On STANDBY CLUSTER** : Perform pureScale rolling update and commit
    2. **Issue TAKEOVER** (New primary cannot form HADR connection with (now downlevel) new standby)
    3. **On NEW STANDBY (OLD PRIMARY) CLUSTER** : Update and Commit, then Activate
  - In V11.1, HADR log send and replay can occur **during crash recovery**
    - Allows logs written during crash recovery to be replayed while crash recovery is occurring
      - Previously, log send and replay was disabled during crash recovery
    - Allows more rapid attainment of PEER state
    - Especially important in pureScale during **online** member crash recovery
    - Support added for both pureScale and non-pureScale



# GDPC Support Enhancements

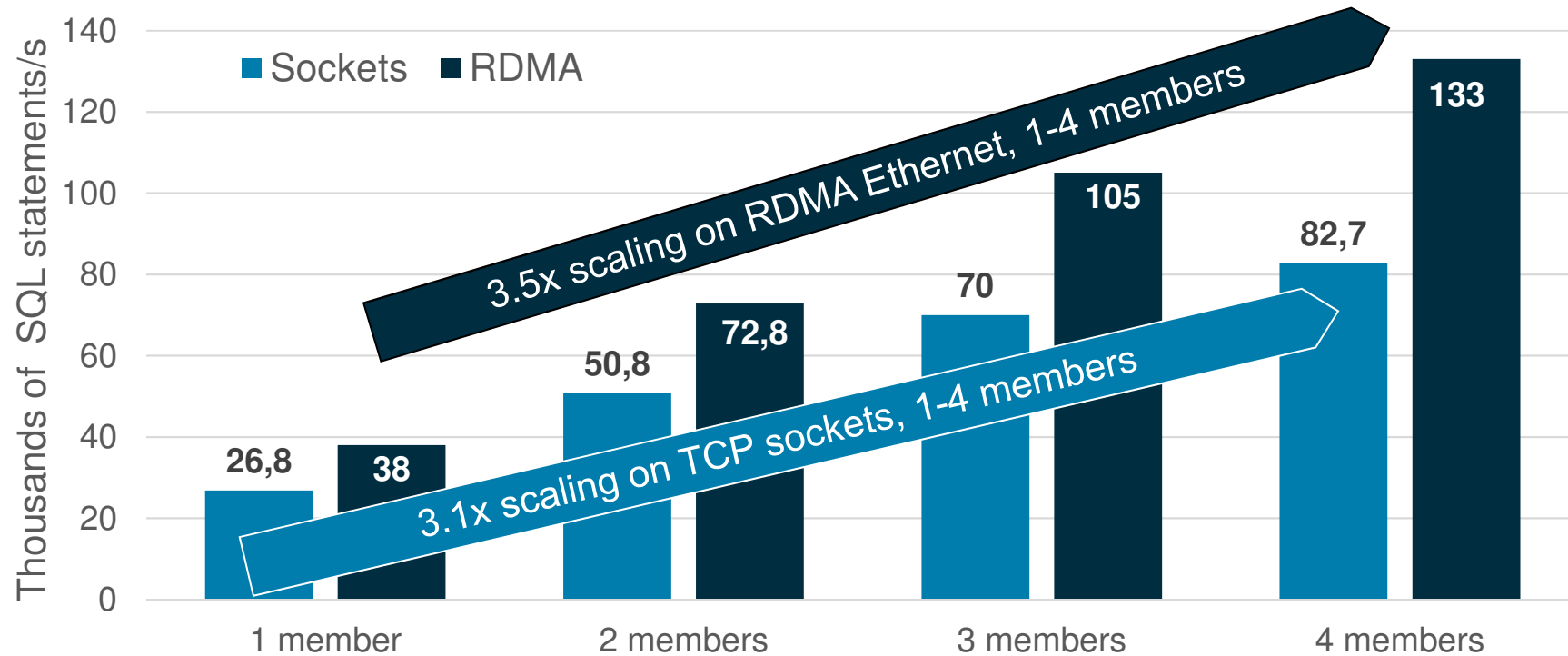


- **DB2 V11 adds improved high availability for *Geographically Dispersed DB2 pureScale Clusters (GDPC)* for both RoCE & TCP/IP**
  - **Multiple adapter ports** per member and CF to support higher bandwidth and improved redundancy at the adapter level
  - **Dual switches** can be configured at each site to eliminate the switch as a site-specific single point of failure (i.e. 4-switch configuration)



# Horizontal Scaling with DB2 pureScale on POWER Linux

Scale-out Throughput – DB2 pureScale on LE POWER Linux



- 80% read / 20% write OLTP workload
- POWER8 4c/32t, 160 GB LBP
- 10 Gb RoCE RDMA Ethernet / 10 Gb TCP sockets

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

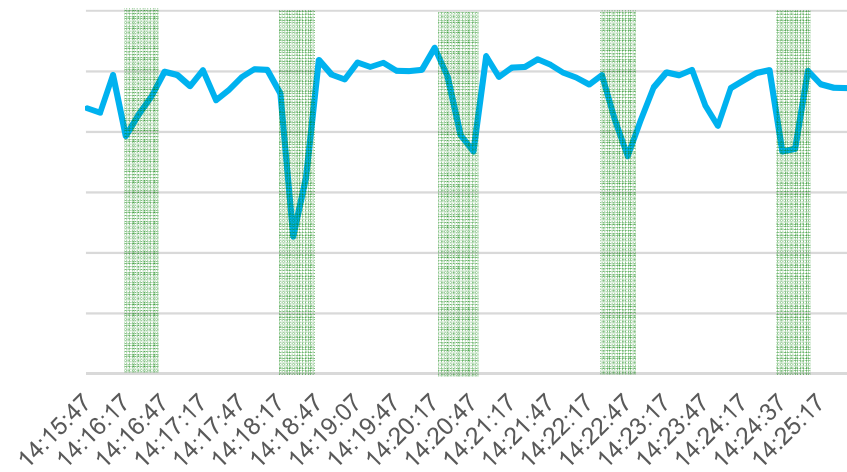
## Improved Table TRUNCATE Performance in pureScale

- **More efficient processing of Group Bufferpool (GBP) pages**
  - Speeds up truncate of permanent tables especially with large GBP sizes
  - Helps DROP TABLE and LOAD / IMPORT / INGEST with REPLACE option
  - Enables improved batch processing with these operations
- **Example**
  - Workload with INGEST (blue) and TRUNCATE (green) of an unrelated table
  - DB2 v11.1 has much smaller impact on OLTP workload than DB2 10.5 fp5

Application throughput - DB2 v10.5 fp5



Application throughput – DB2 v11.1



## Increased Workload Balancing Flexibility with pureScale

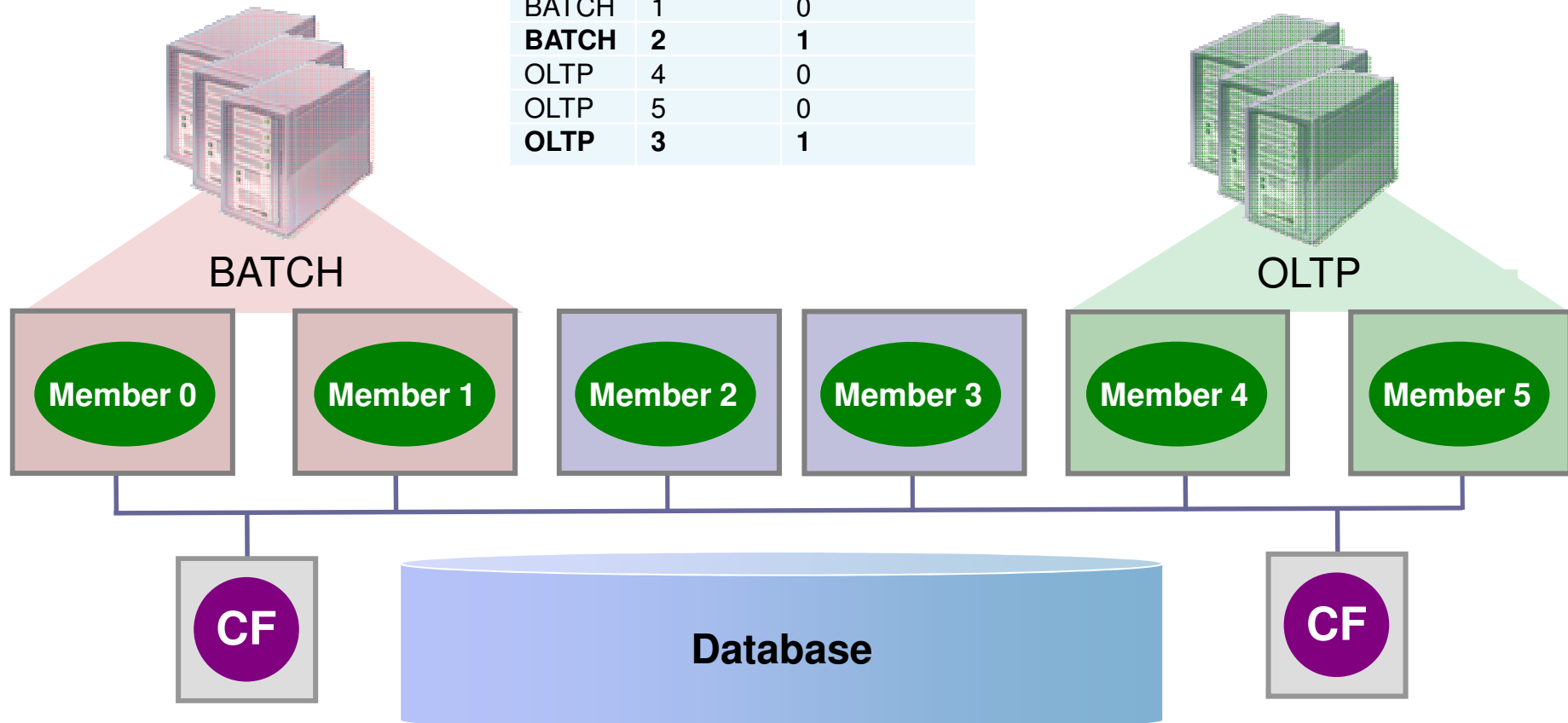
- **Version 11.1 adds member subset configuration to allow explicit definition of alternate members**
  - A new **Failover Priority** allows you to explicitly define the members to be used as alternates in the event a primary member of the subset fails
    - Members with a failover priority of 0 (the default priority) are considered primary members
    - Members with failover priority 1-254 are considered alternative members
  - If a primary member fails, an alternate member is automatically selected to service the member subset workload, from a lower failover priority
    - Selection starts at failover priority 1, then failover priority, etc
  
- **Allows users currently using client affinity to move their configuration to member subset, allowing them to benefit from:**
  - Centralized server based configuration (as opposed to client based configuration in db2dsdriver.cfg file)
  - Dynamic server side reconfiguration

# V11: More Flexibility with New FAILOVER\_PRIORITY

```
CALL SYSPROC.WLM_ALTER_MEMBER_SUBSET( 'BATCH', NULL, '(ADD 2 FAILOVER_PRIORITY 1)');
CALL SYSPROC.WLM_ALTER_MEMBER_SUBSET( 'OLTP', NULL, '(ADD 3 FAILOVER_PRIORITY 1)');
```

SUBSET	MEMBER	FAILOVER_PRIORITY
BATCH	0	0
BATCH	1	0
<b>BATCH</b>	<b>2</b>	<b>1</b>
OLTP	4	0
OLTP	5	0
<b>OLTP</b>	<b>3</b>	<b>1</b>

This information is available from  
 SYSCAT.MEMBERSUBSETMEMBERS and  
 db2pd -membersubsetstatus -detail



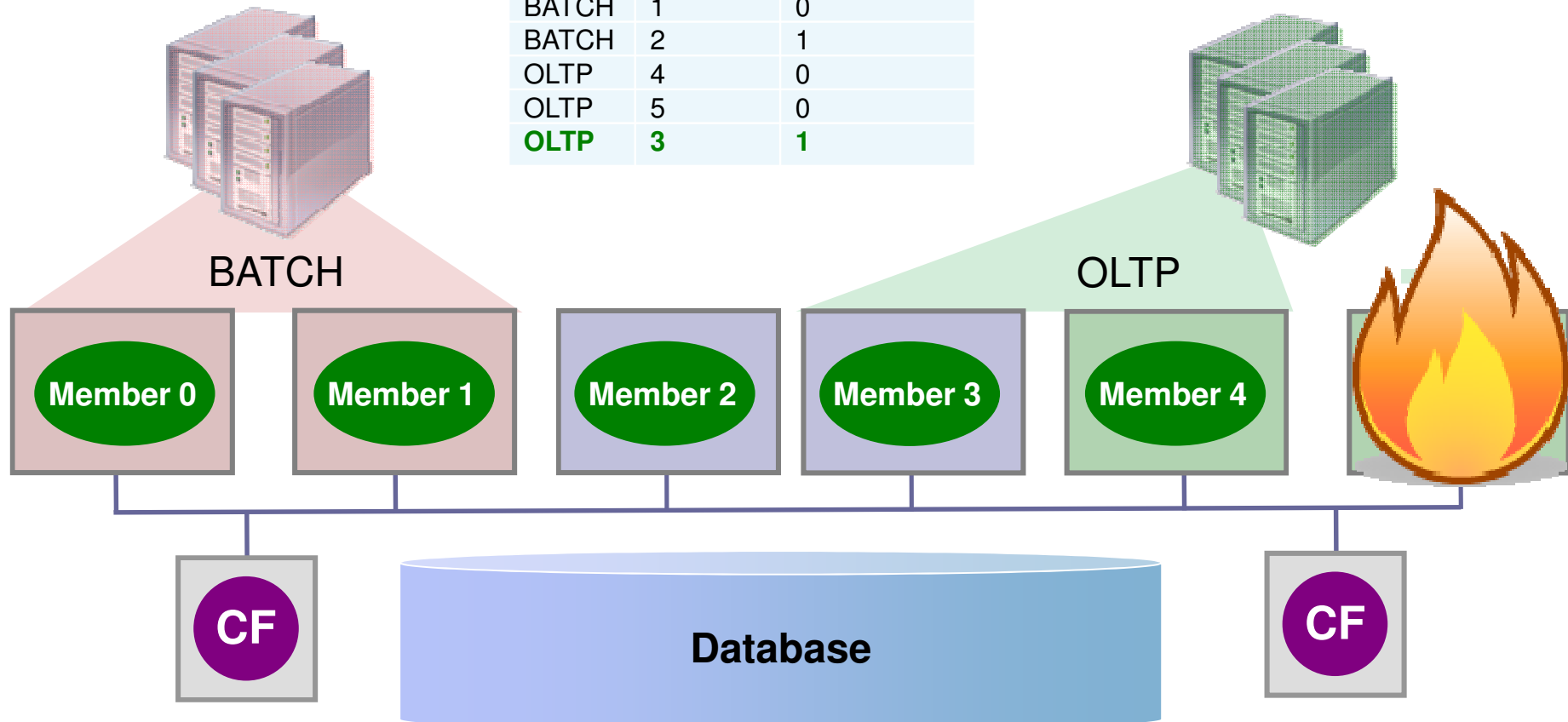
# Member Subsets : FAILOVER PRIORITY

```
CALL SYSPROC.WLM_ALTER_MEMBER_SUBSET( 'BATCH', NULL, '(ADD 2 FAILOVER_PRIORITY 1)');
```

```
CALL SYSPROC.WLM_ALTER_MEMBER_SUBSET( 'OLTP', NULL, '(ADD 3 FAILOVER_PRIORITY 1)');
```

SUBSET	MEMBER	FAILOVER_PRIORITY
BATCH	0	0
BATCH	1	0
BATCH	2	1
OLTP	4	0
OLTP	5	0
<b>OLTP</b>	<b>3</b>	<b>1</b>

This information is available from  
 SYSCAT.MEMBERSUBSETMEMBERS and  
 db2pd -membersubsetstatus -detail



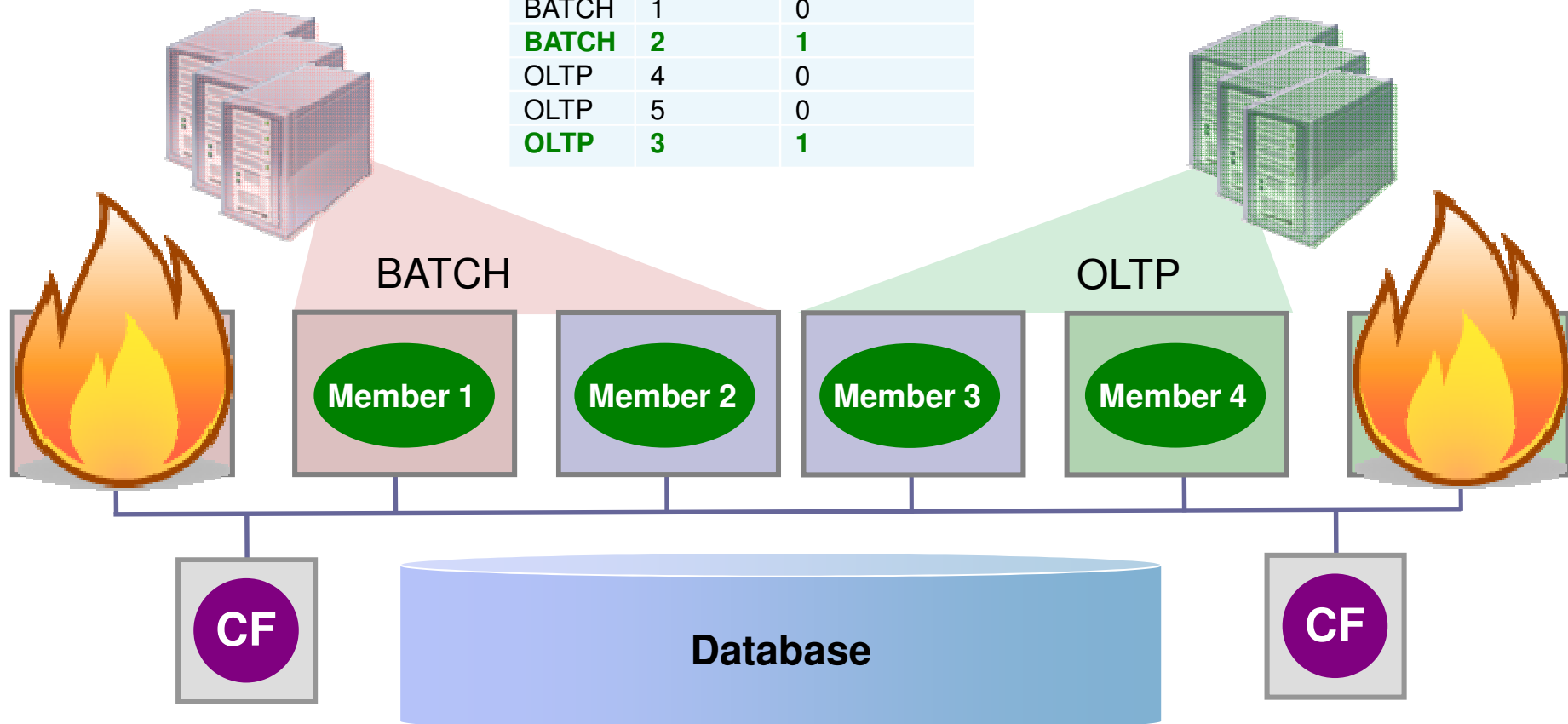
# Member Subsets : FAILOVER PRIORITY

```
CALL SYSPROC.WLM_ALTER_MEMBER_SUBSET( 'BATCH', NULL, '(ADD 2 FAILOVER_PRIORITY 1)');
```

```
CALL SYSPROC.WLM_ALTER_MEMBER_SUBSET( 'OLTP', NULL, '(ADD 3 FAILOVER_PRIORITY 1)');
```

SUBSET	MEMBER	FAILOVER_PRIORITY
BATCH	0	0
BATCH	1	0
<b>BATCH</b>	<b>2</b>	<b>1</b>
OLTP	4	0
OLTP	5	0
<b>OLTP</b>	<b>3</b>	<b>1</b>

This information is available from SYSCAT.MEMBERSUBSETMEMBERS and db2pd -membersubsetstatus detail





# DB2 Version 11.1 Highlights ( GA: June, 2016)

## Core Mission Critical Workloads : Extending DB2 Leadership

### Comprehensive Enterprise Security



#### Enterprise Encryption

- Centralized Key Managers (KMIP)

### Availability 2<sup>nd</sup> only to DB2 for zOS



#### Simple Fast Deployment

- Up and running in hours

#### Even Greater Availability

- Zero data loss DR with HADR
- More online management

#### More Platforms Supported

- Power Linux (LE)
- Virtualization for RDMA (x86)

### Significant Core Database Advances



#### Very Large Database Performance

- Higher user throughput

#### Simpler, Faster, More Online Upgrades

- Faster, no need for offline backup
- Streamlined HADR upgrade
- DB2 Version 9.7 direct to 11.1

## Warehousing Workloads : Most Consumable, Most Scalable In-Memory Warehousing Platform

### Massive Scale Warehousing at In-Memory Performance



#### MPP BLU Scalability

- PB scale in-memory warehousing

### Next Gen In-Memory Performance, Function & Workloads



- **Faster ELT/ETL performance**
- **More Query Workloads Optimised**
- **More Function supported**

- Generated Columns
- RCAC
- OLAP + BLU Perf

### Enhanced Compatibility



#### Multi-Lingual SQL Advances

- Postgres SQL

#### Support for European Languages

- Codepage 819

## **New SQL and DB2 BLU Enhancements**

- **Extended SQL Support**
- **Additional Netezza, Oracle & Postgres Compatibility Support**
  
- **Massively Parallel Processing (MPP)**
- **Performance Improvements**
- **Security Enhancements (RCAC)**
- **SQL functions optimized for BLU**

## New CREATE FUNCTION statement for aggregate UDFs

- **The CREATE FUNCTION (aggregate interface) statement allows you to create your own aggregate functions**
  - An aggregate function returns a single value that is the result of an evaluation of a set of like values, such as those in a column within a set of rows
  - Use your choice of programming language (C++ recommended)
  - Four sections within the function are defined for each stage of the aggregation process (via UDFs themselves)
    - INITIATE, ACCUMULATE, MERGE, FINALIZE
  
- **Hyperloglog user-defined aggregate function**

```
CREATE OR REPLACE FUNCTION hll(VARCHAR(10)) RETURNS INT
SPECIFIC hll AGGREGATE WITH (registers VARCHAR(16) FOR BIT DATA)
USING INITIALIZE PROCEDURE hll_initialize
    ACCUMULATE PROCEDURE hll_accumulate
    MERGE PROCEDURE hll_merge
    FINALIZE FUNCTION hll_finalize;
```

## **BINARY and VARBINARY data**

### ▪ **BINARY and VARBINARY data types**

- Allow binary string data to be stored and manipulated without the overhead of using a BLOB type
- A binary string is a sequence of bytes that are used to store as pictures, sound, or mixed media
- BINARY and VARBINARY data types are compatible with each other and are compatible with the BLOB data type
- Binary string data types are not compatible with character string data types, except those character strings that are defined as FOR BIT DATA
- Support for BINARY and VARBINARY data types enhances compatibility with other relational database management systems
- BINARY can contain 1-254 bytes
- VARBINARY can be up to 32672 bytes

## Compatibility features for Netezza Performance Server

- **Use the SQL\_COMPAT session variable to activate the following optional NPS compatibility features**
  - SET SQL\_COMPAT='NPS'
- **Controlled Compatibility Features**
  - Double-dot notation
    - You can use double-dot notation to specify a database object
  - TRANSLATE parameter syntax
    - TRANSLATE (*char-string-exp*, *from-string-exp* ,*to-string-exp*)
  - Grouping by SELECT clause columns
    - You can specify the ordinal position or exposed name of a SELECT clause column when grouping the results of a query
  - Routines written in NZPLSQL
    - The NZPLSQL language can be used in addition to the SQL PL language.
- **On by default Compatibility Features**
  - JOIN USING, UPDATE FROM, CREATE TEMP syntax alternative to DGTT, LIMIT OFFSET, NULL ORDERING

## Many New SQL Functions

- **Date and Time Functions (31)**

- e.g. OVERLAPS(), DAYOFMONTH(), AGE(), ...

- **Statistical Functions (9)**

- e.g. COVARIANCE\_SAMP(), MEDIAN(), (), WIDTH\_BUCKET(), ...

- **Regular Expression Functions (7)**

- e.g. REGEXP\_LIKE(), REGEXP\_SUBSTR(), REGEXP\_COUNT(), ...

- **Bitwise Functions (12)**

- e.g. INT4AND(), INT4OR(), INT4XOR(), INT4NOT(), ...

- **Misc. Functions (7)**

- e.g. BTRIM(), RAWTOHEX(), ...

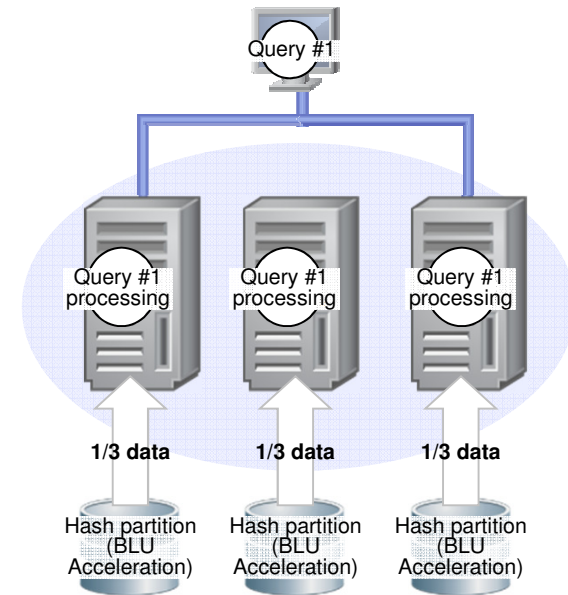
# BLU Acceleration: MPP Scale Out

## Technology

- Pervasive SMP & MPP Query Parallelism
- Inter-partition query parallelism simultaneous with intra-partition- parallelized, memory-optimized, columnar, SIMD-enabled, BLU processing

## Value Proposition

- Improve Response Time
  - All servers contribute to the processing of a query
- Massively Scale Data
  - Significantly beyond current practical limits
- Streamline BLU Adoption
  - Add BLU Acceleration to existing data warehouses

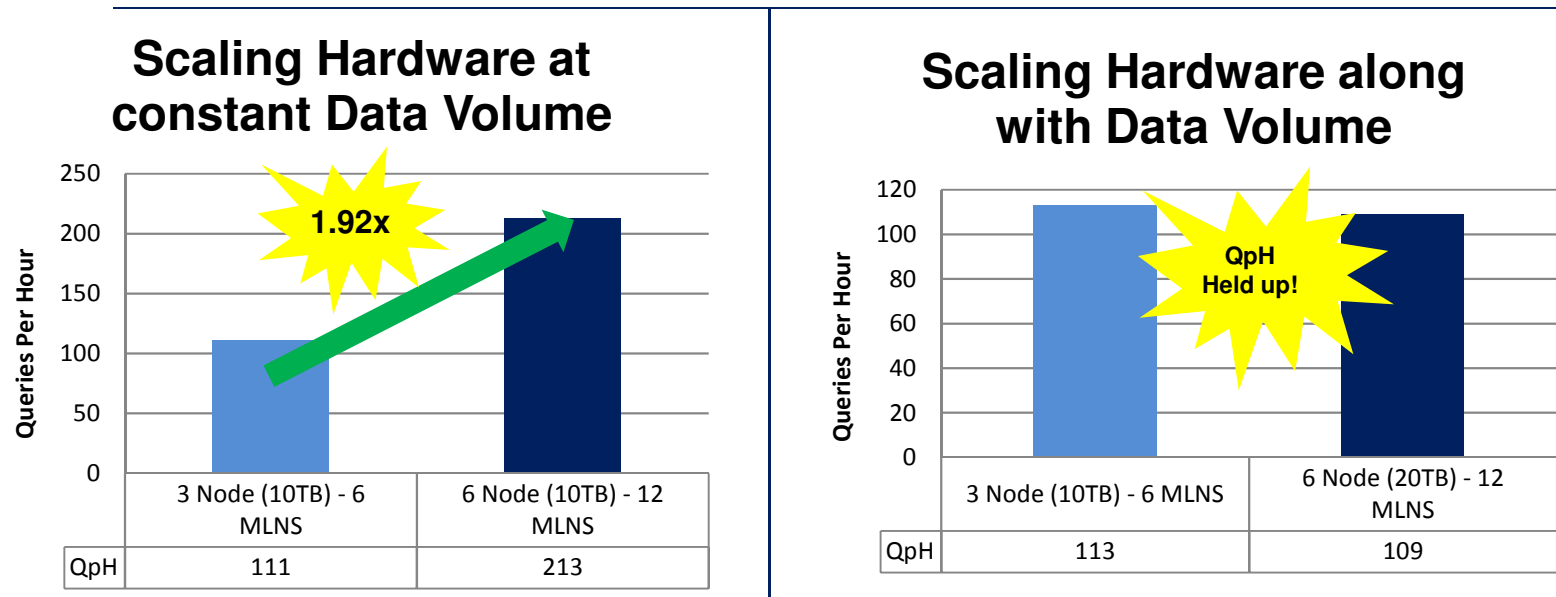


**MPP = Massively Parallel Processing**

DB2 10.5 BLU Capacity	DB2 V11.1 BLU Capacity
10s of TB	1000s of TB
100s of Cores	1000s of Cores

# Demonstrating BLU MPP Linear Scaling

- **DB2 Version 11.1 on an IBM Power Systems E850 Cluster**



- **Scaling was measured in two different ways**

- Doubling the hardware but keeping the database constant
- Doubling the hardware and doubling the database size
- Both tests used the BD Insights Heavy Analytics Internal Workload

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

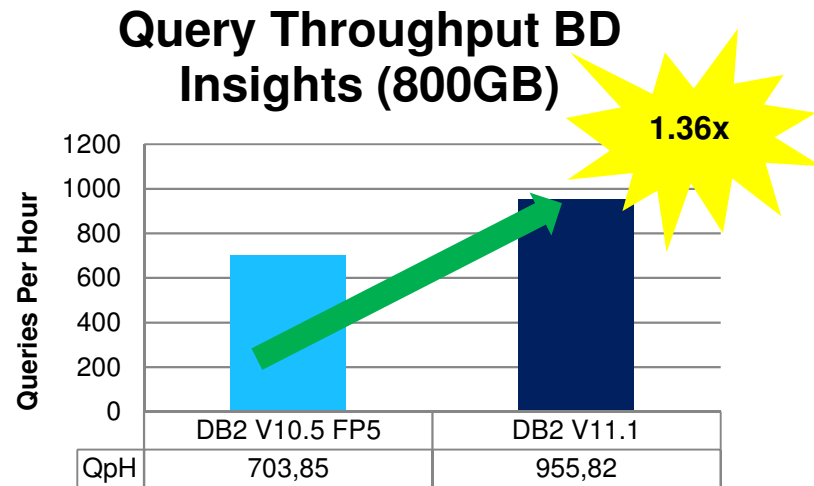


## Performance Improvements

- **BLU Acceleration includes a number of performance improvements in the columnar engine**
  - Faster SQL MERGE processing
  - Nested Loop Join Support
  - Industry Leading Parallel Sort
  - Improved SORTHEAP utilization
  - Query Rewrite Improvements
  - Push-down of a number of SQL & OLAP functions into the BLU engine
  - Optimized SQL support for BLU

# Demonstrating BLU Single Instance Improvement

## DB2 Version 11.1 on Intel Haswell EP



### Reasons for Improvement

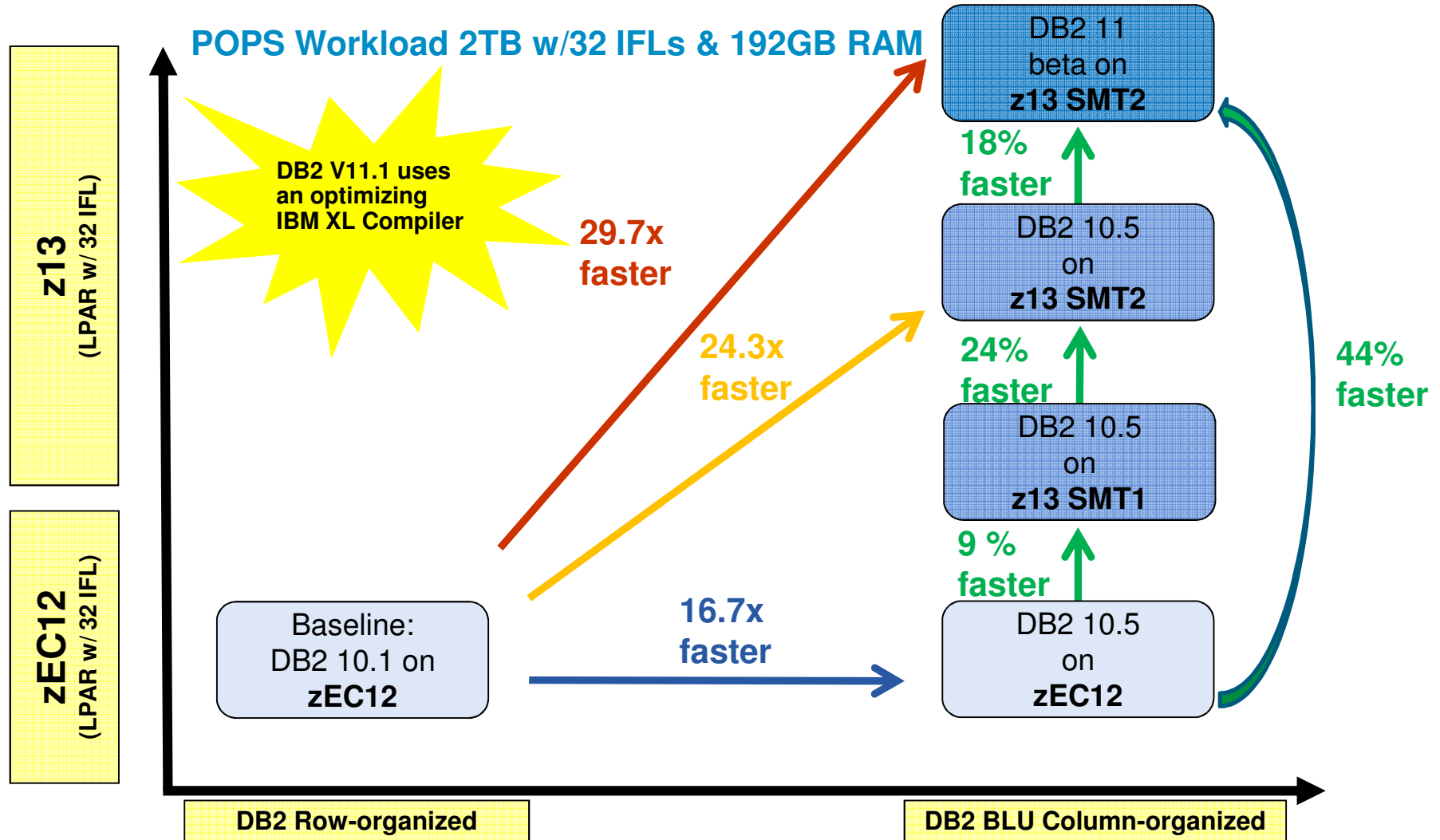
- Native BLU Evaluation**
  - Native Sort
  - Native OLAP (usually combined with sort)
  - Enables query plans to remain as much as possible within the columnar engine
- Query Rewrite Improvements**
  - Find areas to improve degree determination and improve parallel use
- Improved SORTHEAP Utilization**
  - SORTHEAP used for building hash tables for JOINS, GROUP BYs, and other runtime work
  - Efficient use allows for more concurrent intra-query and inter-query operations to co-exist.

## Configuration Details

- 2 socket, 36 core Intel Xeon E5-2699 v3 @ 2.3GHz
- 192GB RAM
- BD Insights Internal Multiuser Workload 800GB

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

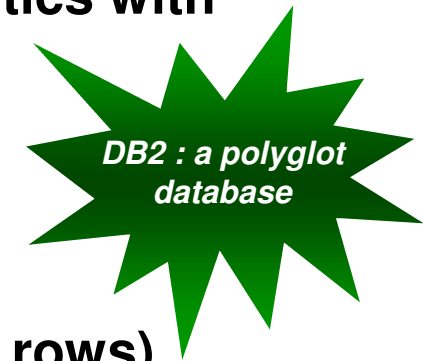
# Tuned for Linux on z13 Mainframe Systems



Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

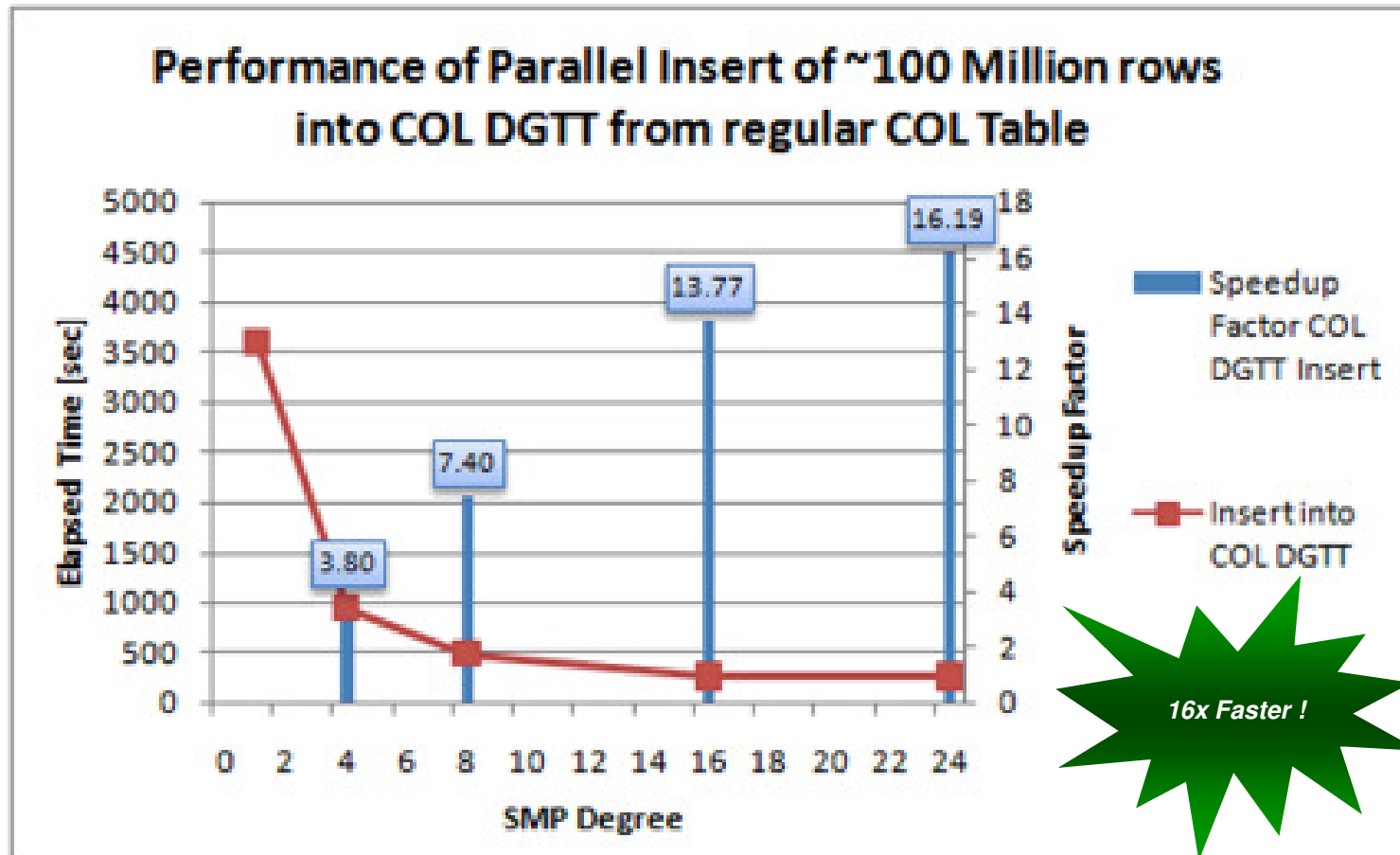
## Optimized SQL Support for Columnar Tables

- **SQL OLAP improvements for deeper in-database analytics with column-organized tables**
- **Additional Oracle Compatibility Support**
  - Wide rows
  - Logical character support (CODEUNITS32)
- **DGTT support (except not logged on rollback preserve rows)**
  - Parallel insert into not-logged DGTT from BLU source
- **IDENTITY and EXPRESSION generated columns**
- **European Language support (Codepage 819)**
- **NOT LOGGED INITIALLY support**
- **Row and Column Access Control (RCAC)**
- **ROWID Support**
- **Faster SQL MERGE processing**
- **Nested Loop Join Support**



# BLU Acceleration: Massive Gains for ELT & ISV Apps

*BLU Declared Global Temporary Table (not-logged DGTT) Parallelism*



Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

# SQL Functions Optimized for Columnar Mode

## ▪ String Functions

- LPAD, RPAD
- TO\_CHAR
- INITCAP

## ▪ Numeric Functions

- POWER, EXP, LOG10, LN
- TO\_NUMBER
- MOD
- SIN, COS, TAN, COT, ASIN, ACOS, ATAN
- TRUNCATE

## ▪ Date and Time Functions

- TO\_DATE
- MONTHNAME, DAYNAME

## ▪ Miscellaneous

- COLLATION\_KEY

## OLAP Functions Support by BLU

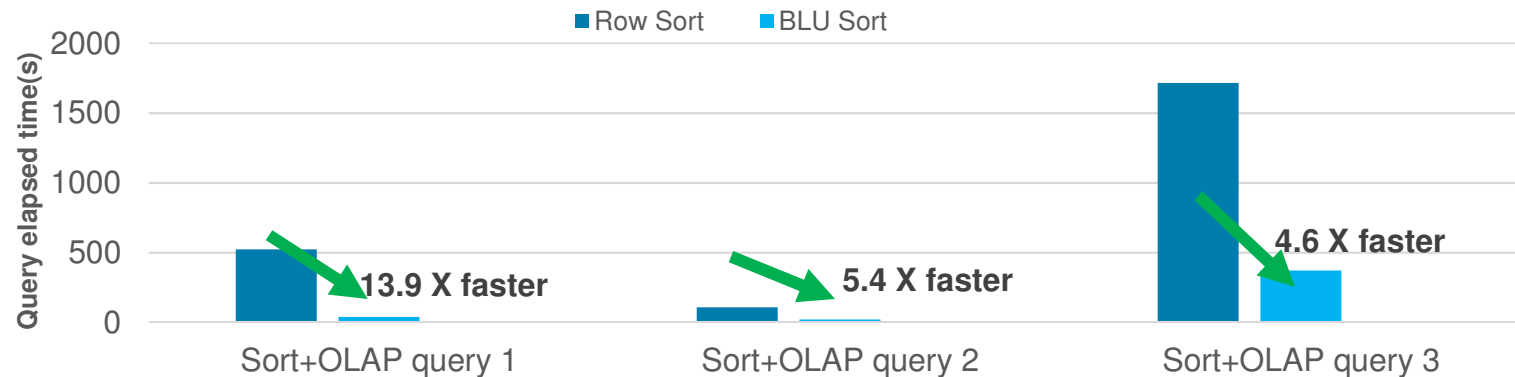
- **OLAP functions supported by BLU:**
  - RANK, DENSE\_RANK, ROW\_NUMBER
- **OLAP column functions supported by BLU:**
  - AVG
  - COUNT, COUNT\_BIG
  - MIN, MAX
  - SUM
  - FIRST\_VALUE
  - RATIO\_TO\_REPORT
- **Note: Window aggregation group clause is limited to:**
  - ROWS/RANGE BETWEEN UNBOUNDED PRECEDING AND UNBOUNDED FOLLOWING
  - ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW
  - ROWS BETWEEN CURRENT ROW AND CURRENT ROW (not supported for FIRST\_VALUE)





# Industry Leading Parallel Sort

- **Leverages the latest sort innovations from IBM TJ Watson Research and DB2 Development**
  - Enhancements can increase BLU Acceleration performance by as much as 13.9X
- **BLU Sort+OLAP on SMP Environment**



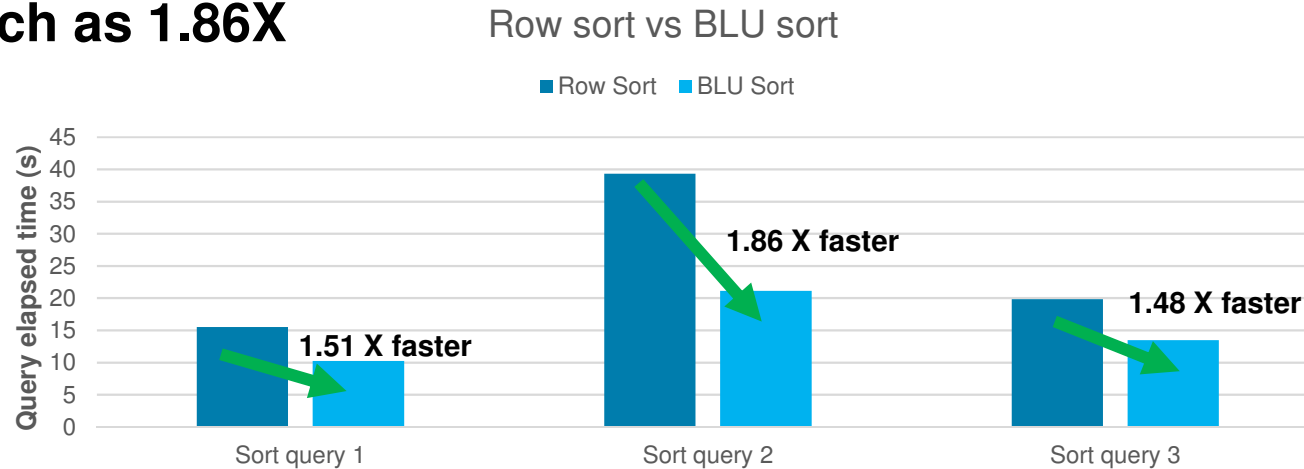
## ▪ Configuration Details

- On 4-socket Intel Xeon platform with 72 Cores and 742G RAM
- 1 TB TPC-DS database
- Query scenarios involving multiple sort and OLAP operations

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

## BLU Sort on MPP Environment

- Enhancements can increase BLU Acceleration performance by as much as 1.86X



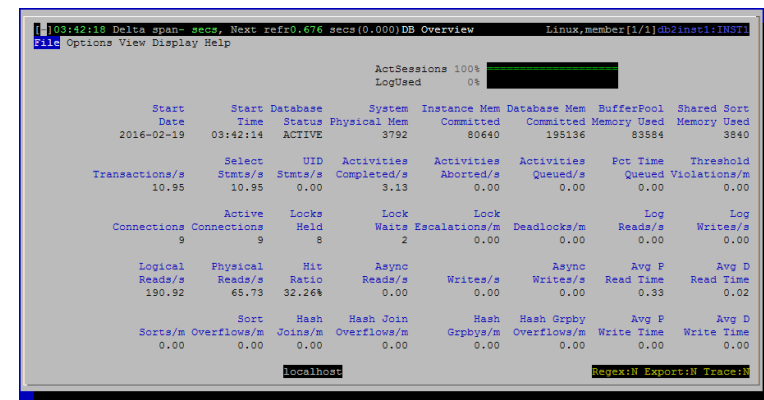
- Configuration Details**

- BLU MPP with 1 physical node/4-MLNs and 8 cores per MLN
- 4-socket, 40 Cores Intel Xeon platform and 1 TB RAM
- Used Sort Target Workload 1 TB (an internal workload based on the TPC-DS schema), with 10-Million rows to sort for each scenarios
- Query scenarios involving single ORDER BY concentrating on sort operation

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

# dsmtop

- A new easy-to-use text-based User Interface (UI) for managing DB2
- Similar to db2top, but uses DB2's lightweight in-memory monitoring functions
  - i.e. the 'MON\_GET\_\*' interfaces (not the older snapshot interfaces)
- Additional features include
  - Metrics about newer DB2 features, eg.
    - BLU (column store)
    - pureScale
    - Workload management
  - Increased coverage of memory usage & reorgs
  - Windows platform support !
  - Menus for not-so-experienced DBAs
  - Monitoring remote databases



The screenshot shows the dsmtop utility interface. At the top, it displays the time '10:42:18', the instance name 'Delta span- secs, Next ref:0.676 secs(0.000)DB Overview', and the platform 'Linux,member[1/1]db2inst1:INST1'. Below this, there are several sections of data:

Start Date	Start Time	Database Status	System Physical Mem	Instance Mem Committed	Database Mem Committed	BufferPool Memory Used	Shared Sort Memory Used
2016-02-19	03:42:14	ACTIVE	3792	80640	195136	83584	3840

Transactions/s	Select Stmts/s	UID Stmts/s	Activities Completed/s	Activities Aborted/s	Activities Queued/s	Pct Time	Threshold
10.95	10.95	0.00	3.13	0.00	0.00	0.00	0.00

Active Connections	Locks Held	Lock Waits	Lock Escalations/m	Lock Deadlocks/m	Log Reads/s	Log Writes/s
9	9	8	2	0.00	0.00	0.00

Logical Reads/s	Physical Reads/s	Hit Ratio	Async Reads/s	Async Writes/s	Avg P Read Time	Avg D Read Time
190.92	65.73	32.26%	0.00	0.00	0.33	0.02

Sorts/m	Sort Overflows/m	Hash Joins/m	Hash Join Overflows/m	Hash Grpby/m	Hash Grpby Overflows/m	Avg P Write Time	Avg D Write Time
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

At the bottom, it shows 'localhost' and 'Regex:N Export:N Trace:N'.

*dsmtop is available on developerworks now, and included in DB2 Version 11.1*

## Summary

- **DB2 Version 11.1 is a major new version**
- **This talk is focused on traditional software deployment in enterprise data centers**
  - Many of the same capabilities are (or already have been) available in the cloud
    - DB2 for Cloud
    - dashDB
    - IBM dashDB™ Transactional
      - <http://www-01.ibm.com/support/docview.wss?uid=swg21961758>
- **Significant enhancements for all DB2 topologies**
  1. Single server scale-up SMP
  2. Continuously available DB2 pureScale
  3. MPP scale-out
- **Substantial performance, scalability and SQL function improvements**
- **Please return for future DB2Night shows**

Additional DB2 Version 11.1 deep dives to cover features outlined here in more detail and others we didn't even have time to mention

# Questions?



**Torsten Roeber**

*Technical Support  
Competence Center  
Software Support  
Datamanagement*

*IBM Deutschland GmbH  
Am Weiher 24  
64541 Kelsterbach, Germany  
Office: +49 (0)7034-274-1726  
Mobil: +49 (0)1607411349  
[torsten.roeber@de.ibm.com](mailto:torsten.roeber@de.ibm.com)*

# Thank You

**IBM®**

## Legal Disclaimer

- © IBM Corporation 2016. All Rights Reserved.
- The information contained in this publication is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this publication, it is provided AS IS without warranty of any kind, express or implied. In addition, this information is based on IBM's current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this publication or any other materials. Nothing contained in this publication is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.
- References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. Product release dates and/or capabilities referenced in this presentation may change at any time at IBM's sole discretion based on market opportunities or other factors, and are not intended to be a commitment to future product or feature availability in any way. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.
- If the text contains performance statistics or references to benchmarks, insert the following language; otherwise delete:  
Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.
- If the text includes any customer examples, please confirm we have prior written approval from such customer and insert the following language; otherwise delete:  
All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.
- Please review text for proper trademark attribution of IBM products. At first use, each product name must be the full name and include appropriate trademark symbols (e.g., IBM Lotus® Sametime® Unyte™). Subsequent references can drop "IBM" but should include the proper branding (e.g., Lotus Sametime Gateway, or WebSphere Application Server). Please refer to <http://www.ibm.com/legal/copytrade.shtml> for guidance on which trademarks require the ® or ™ symbol. Do not use abbreviations for IBM product names in your presentation. All product names must be used as adjectives rather than nouns. Please list all of the trademarks that you use in your presentation as follows; delete any not included in your presentation. IBM, the IBM logo, Lotus, Lotus Notes, Notes, Domino, Quickr, Sametime, WebSphere, UC2, PartnerWorld and Lotusphere are trademarks of International Business Machines Corporation in the United States, other countries, or both. Unyte is a trademark of WebDialogs, Inc., in the United States, other countries, or both.
- If you reference Adobe® in the text, please mark the first use and include the following; otherwise delete:  
Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
- If you reference Java™ in the text, please mark the first use and include the following; otherwise delete:  
Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.
- If you reference Microsoft® and/or Windows® in the text, please mark the first use and include the following, as applicable; otherwise delete:  
Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.
- If you reference Intel® and/or any of the following Intel products in the text, please mark the first use and include those that you use as follows; otherwise delete:  
Intel, Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
- If you reference UNIX® in the text, please mark the first use and include the following; otherwise delete:  
UNIX is a registered trademark of The Open Group in the United States and other countries.
- If you reference Linux® in your presentation, please mark the first use and include the following; otherwise delete:  
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both. Other company, product, or service names may be trademarks or service marks of others.
- If the text/graphics include screenshots, no actual IBM employee names may be used (even your own), if your screenshots include fictitious company names (e.g., Renovations, Zeta Bank, Acme) please update and insert the following; otherwise delete: All references to [insert fictitious company name] refer to a fictitious company and are used for illustration purposes only.