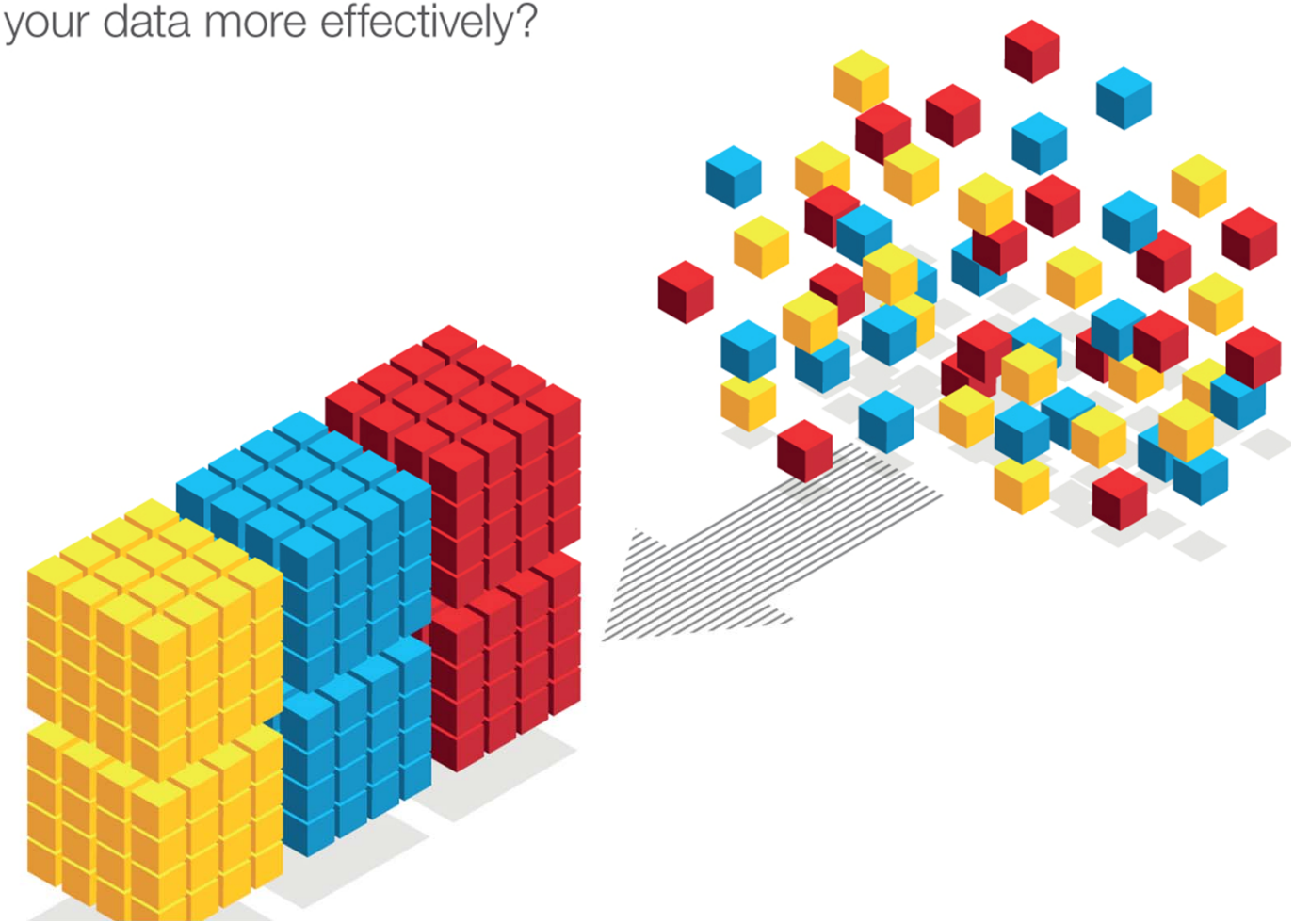
Turn the data you have into the information you need

How will you manage your data more effectively?



IBM DB2 pureScale

Unlimited capacity

Application transparency

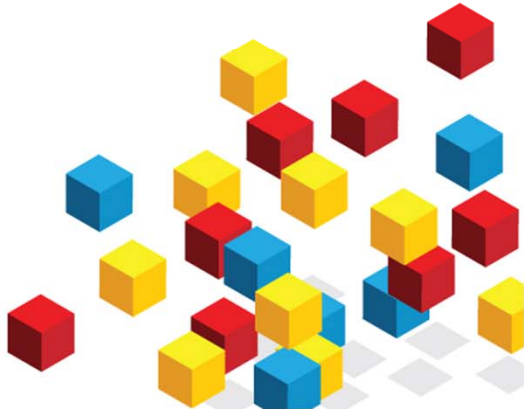
Continuous availability

Vinnie Cardoso

Data Management Technical Pre-sales

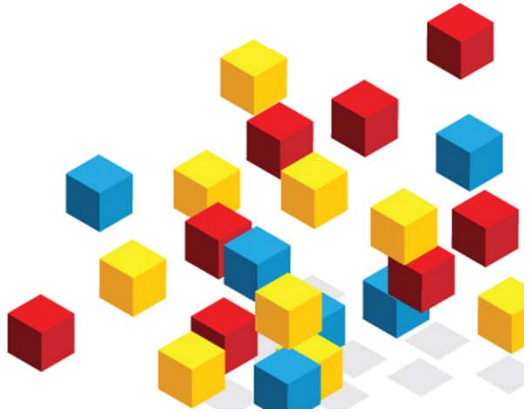
IBM SWG – A/NZ

13/07/2010



Agenda

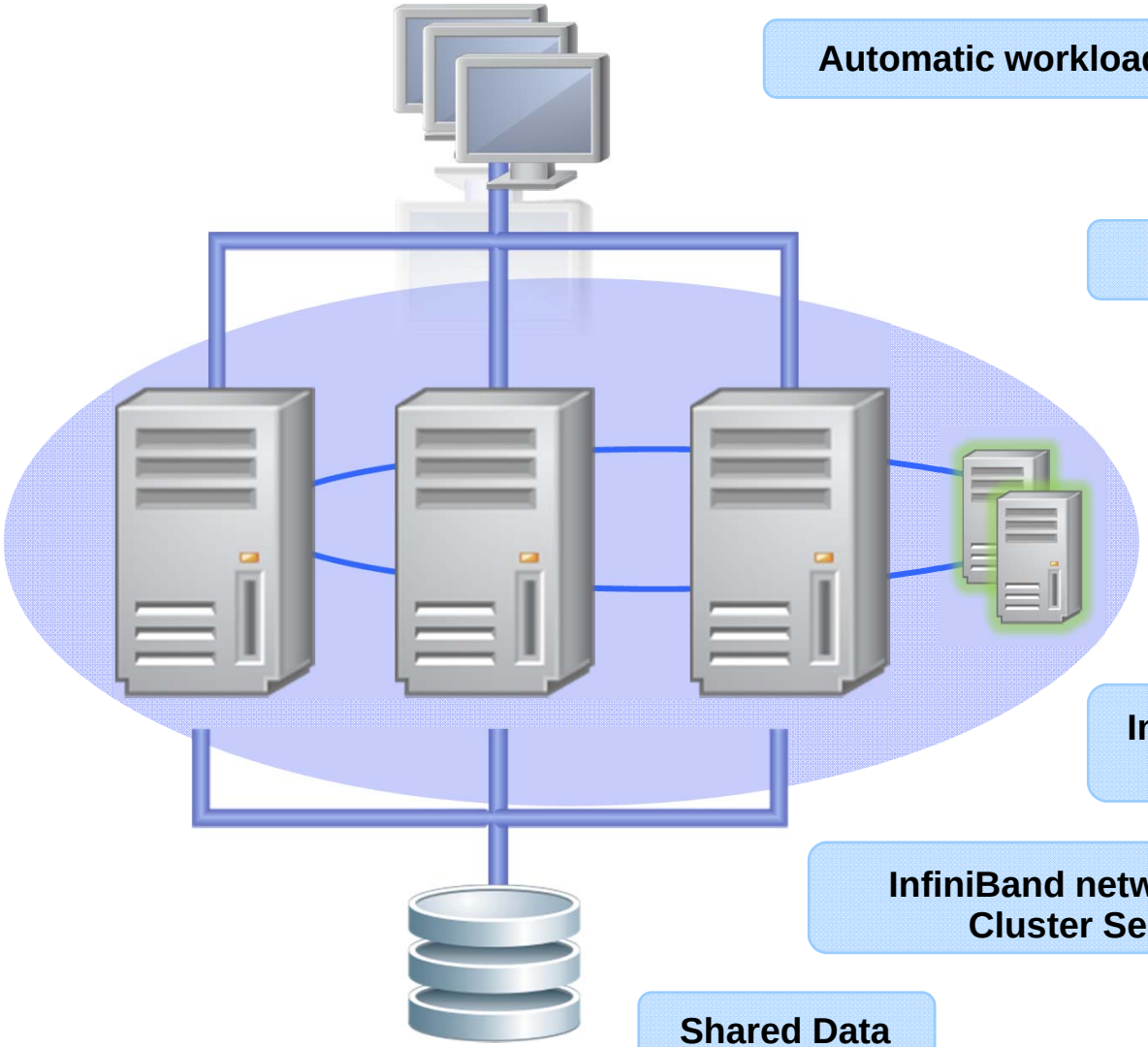
- Technology overview & architecture
- Supported environments
- DB2 pureScale demo introduction & scope
- Demo



Introducing DB2 pureScale

- Virtually unlimited capacity
  - Buy only what you need, add capacity as your needs grow
- Application transparency
  - Avoid the risk and cost of application changes
- Continuous availability
  - Deliver uninterrupted access to your data with consistent performance

DB2 pureScale Architecture



Automatic workload balancing

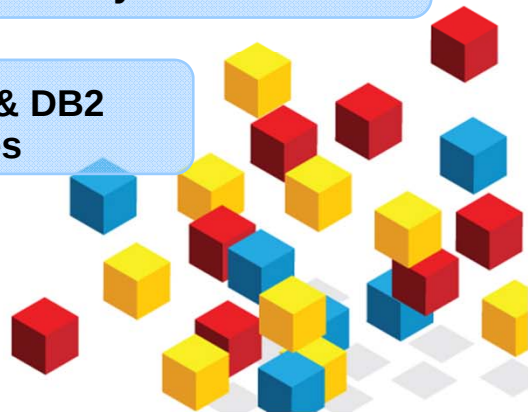
Cluster of DB2 nodes

Leverages global lock and buffering technology derived from z/OS

Integrated Failure Detection & Recovery Automation

InfiniBand network & DB2 Cluster Services

Shared Data



DB2 for z/OS Data Sharing is the Gold Standard

- Everyone recognizes DB2 for z/OS as the “Gold” standard for scalability and high availability

- Even Oracle agrees:

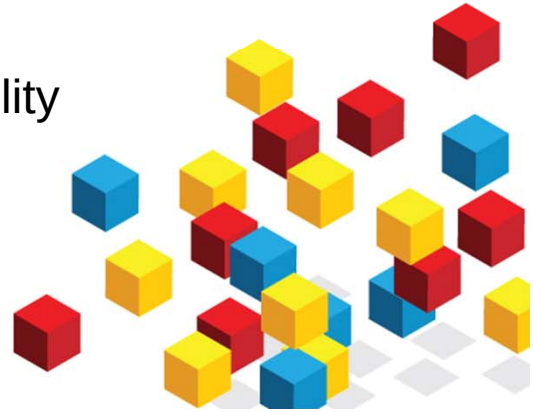
eWEEK.COM

Database

In Larrys Own Words By: Matthew Symonds

I make fun of a lot of other databases- all other databases, in fact, except the mainframe version of DB2. Its a first-rate piece of technology.

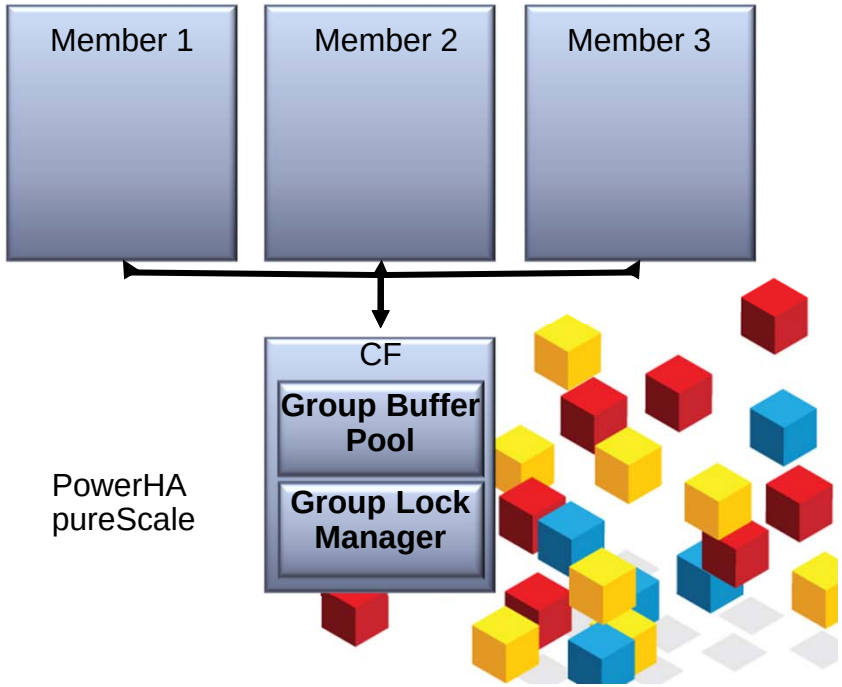
- Why?
  - The **Coupling Facility**!!
    - **Centralized locking**, **centralized buffer pool** deliver superior scalability and superior availability
  - The entire environment on z/OS uses the Coupling Facility



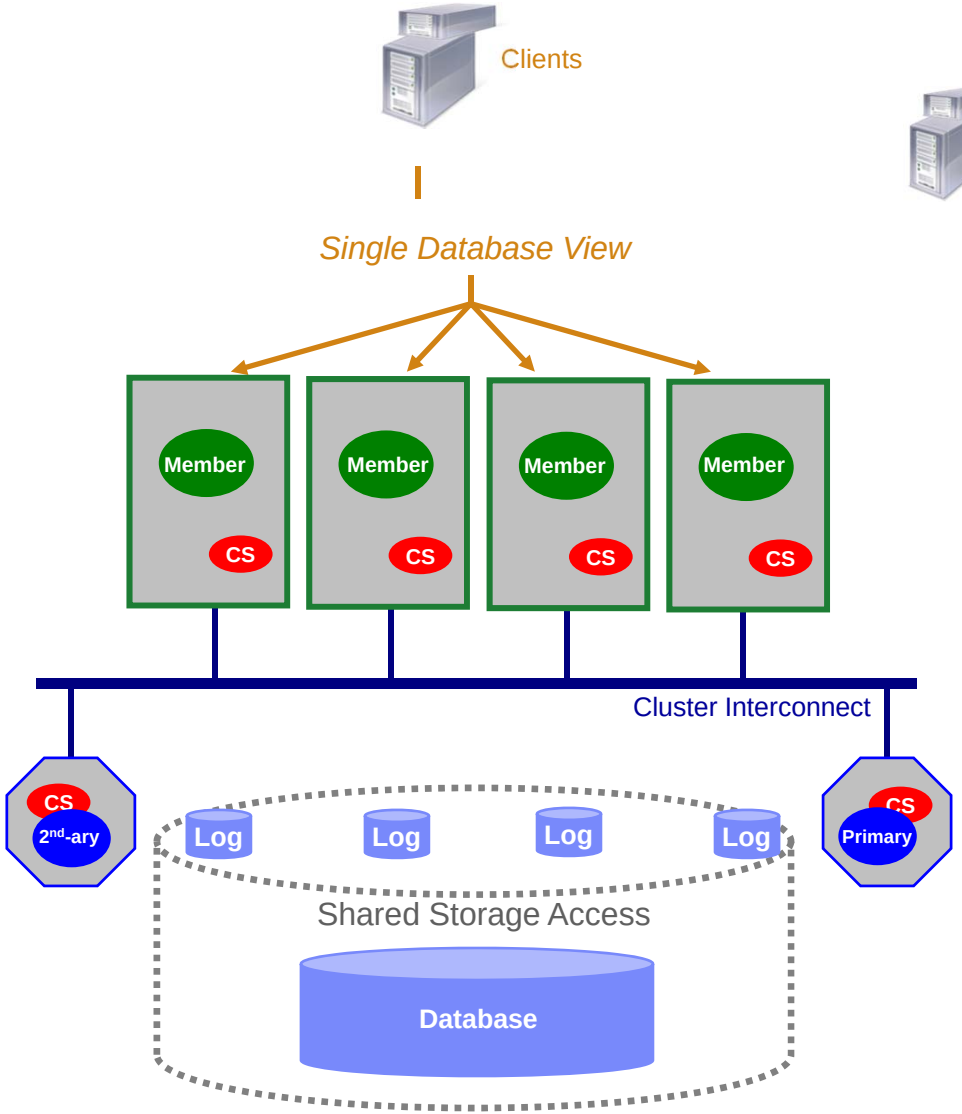
The Key to Scalability and High Availability

- Efficient Centralized Locking and Caching
  - As the cluster grows, DB2 maintains one place to go for locking information and shared pages
  - Optimized for very high speed access
    - DB2 pureScale uses Remote Direct Memory Access (RDMA) to communicate with the powerHA pureScale server
    - No IP socket calls, no interrupts, no context switching

- Results
  - Near Linear Scalability to large numbers of servers
  - Constant awareness of what each member is doing
    - If one member fails, no need to block I/O from other members
    - Recovery runs at memory speeds



Technology Overview



Clients connect anywhere,… … see single database

- Clients connect into any member
- Automatic load balancing and client reroute may change underlying physical member to which client is connected

DB2 engine runs on several host computers

- Co-operate with each other to provide coherent access to the database from any member

Integrated cluster services

- Failure detection, recovery automation, cluster file system
- In partnership with STG (GPFS,RSCT) and Tivoli (SA MP)

Low latency, high speed interconnect

- Special optimizations provide significant advantages on RDMA-capable interconnects (eg. Infiniband)

Cluster caching facility (CF) from STG

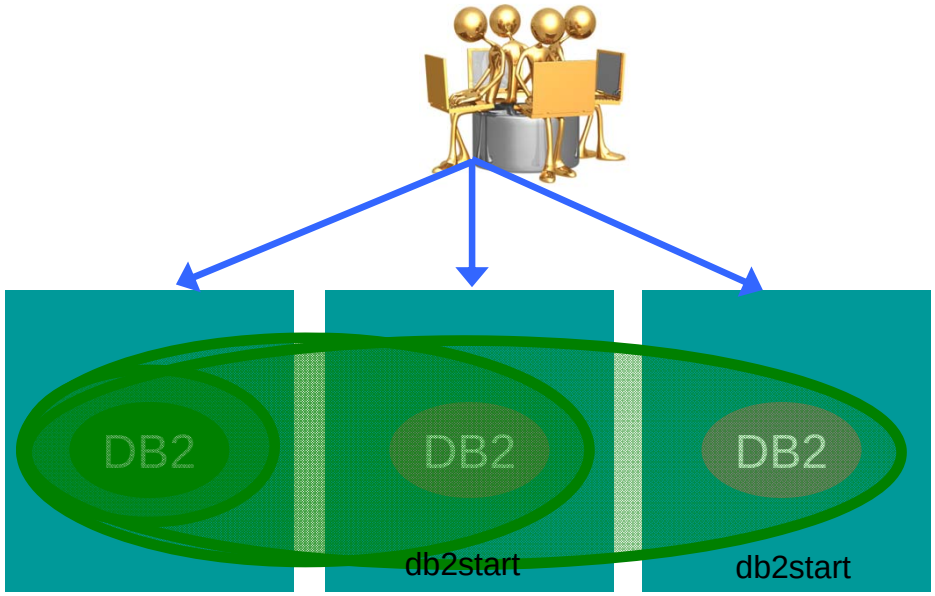
- Efficient global locking and buffer management
- Synchronous duplexing to secondary ensures availability

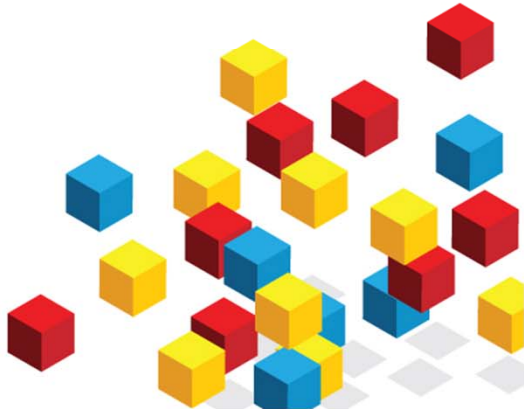
Data sharing architecture

- Shared access to database
- Members write to their own logs
- Logs accessible from another host (used during recovery)

Virtually Unlimited Capacity: Simple Growth

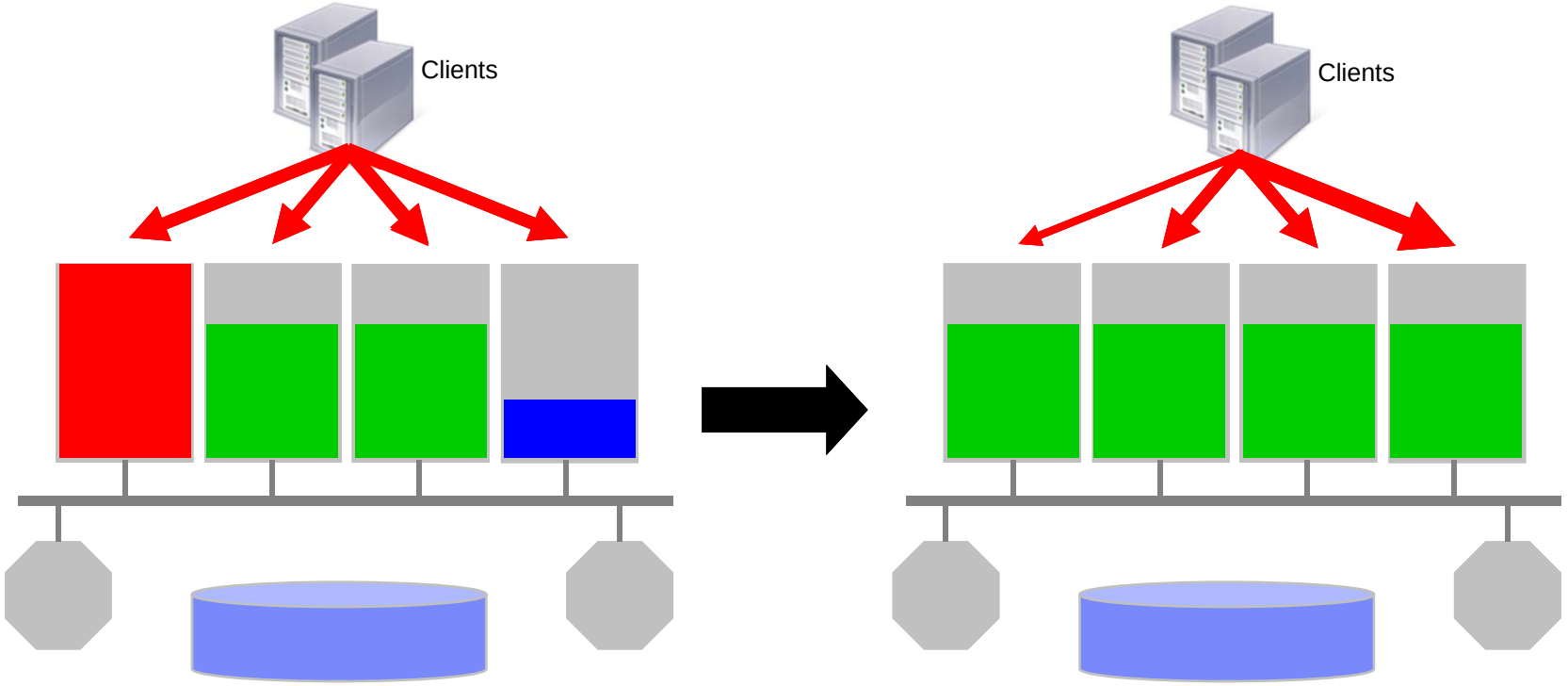
- By enable a member
- Easily add a new member to the cluster
- Just a ‘db2start’ of the member and it comes online and workload rebalances to take advantage of extra capacity
- demonstrates ease of increasing the size of cluster to meet increased demands





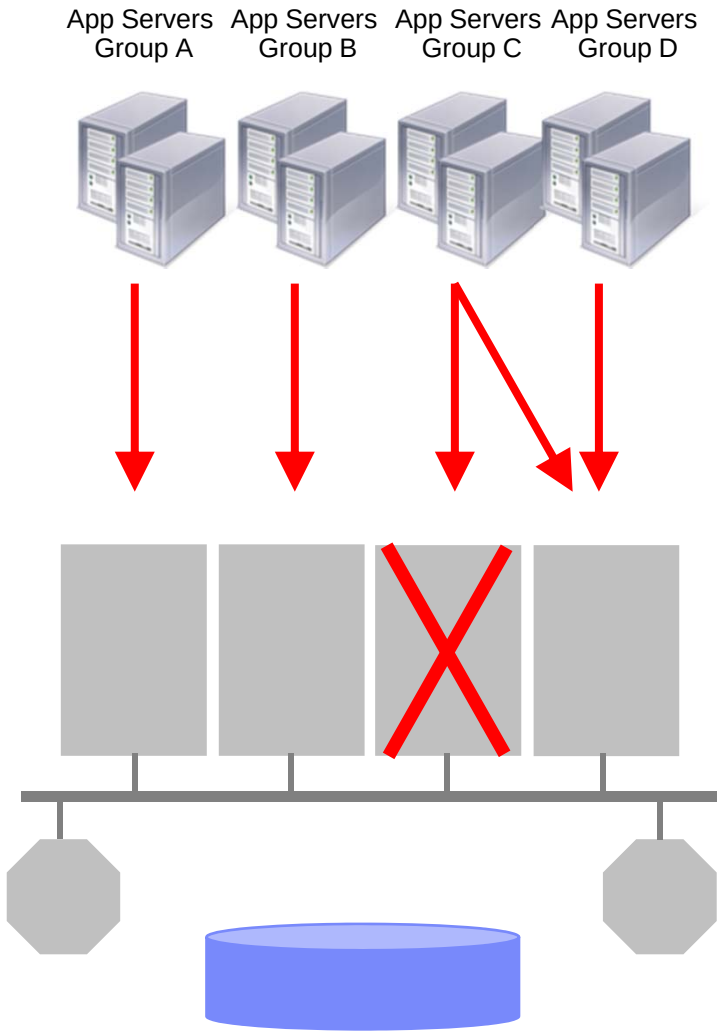
Workload Balancing

- Run-time load information used to automatically balance load across the cluster
  - Shares design with system z Sysplex
  - Load information of all members kept on each member
  - Piggy-backed to clients regularly
  - Used to route next connection (or optionally next transaction) to least loaded member
  - Routing occurs automatically (transparent to application)
- Failover
  - Load of failed member evenly distributed to surviving members automatically
- Fallback
  - Once the failed member is back online, fallback does the reverse



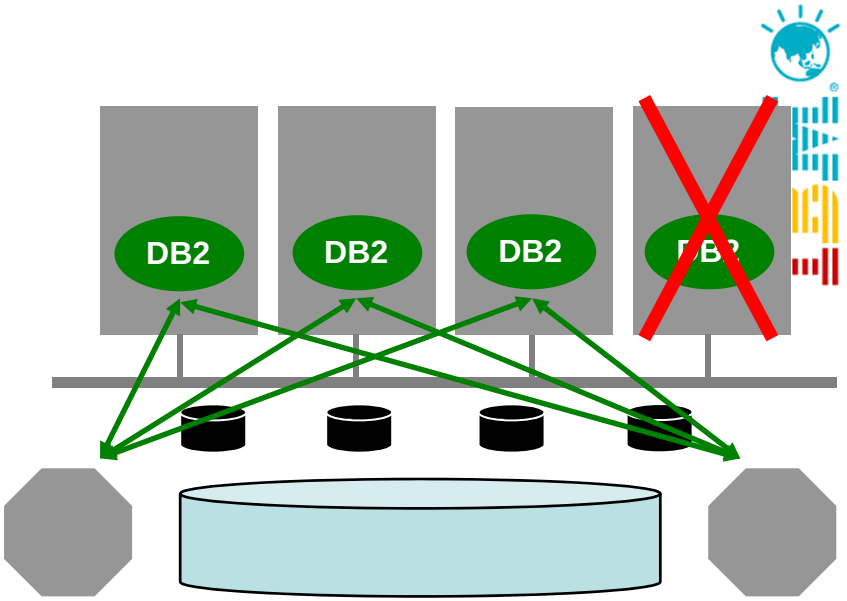
Optional Affinity-Based Routing

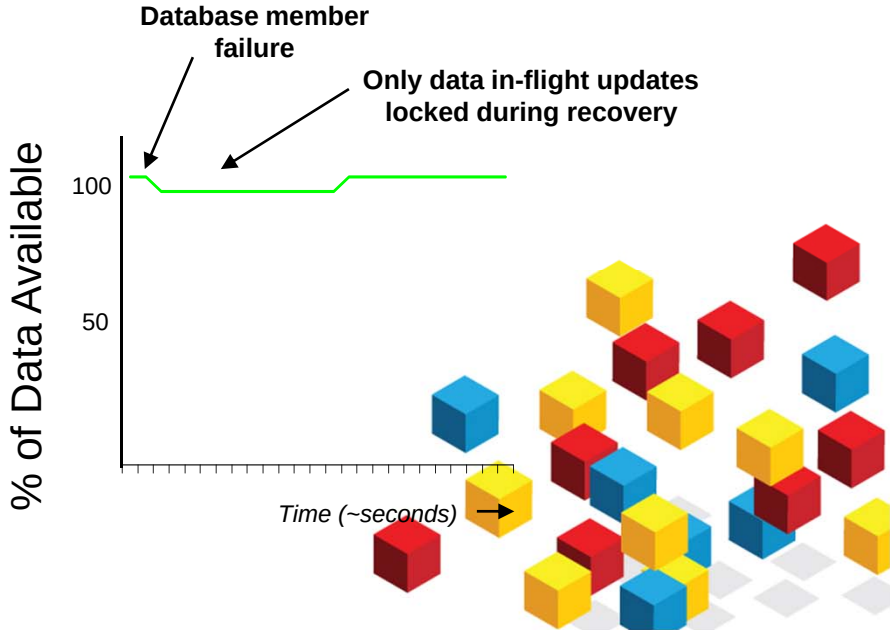
- Allows you to target different groups of clients or workloads to different members in the cluster
  - Maintained after failover …
    … and fallback
- Example use cases
  - Consolidate separate workloads/applications on same database infrastructure
  - Minimize total resource requirements for disjoint workloads
- Easily configured through client configuration
  - db2dsdriver.cfg file



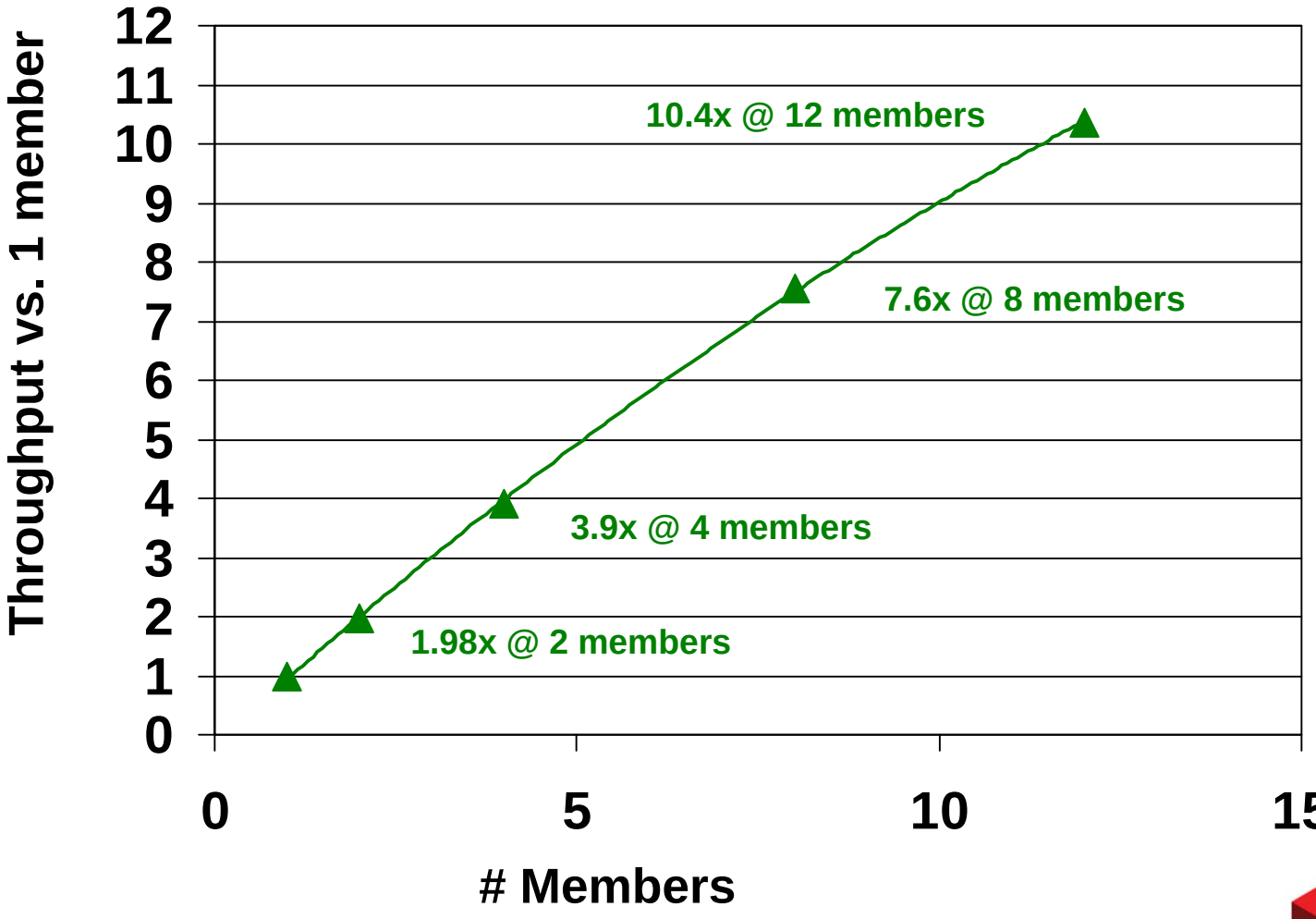
Online Recovery

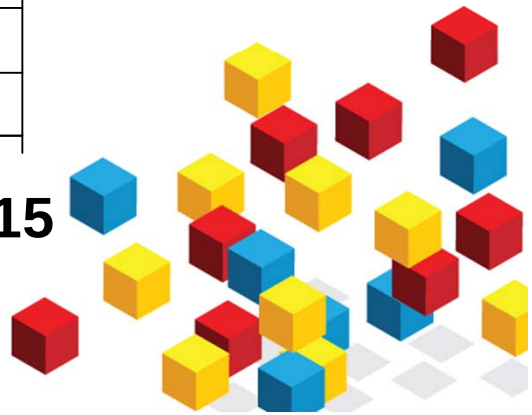
- A key DB2 pureScale design point is to maximize availability during failure recovery processing
- When a database member fails, only data in-flight on the failed member remains locked during the automated recovery
  - In-flight = data being updated on the member at the time it failed





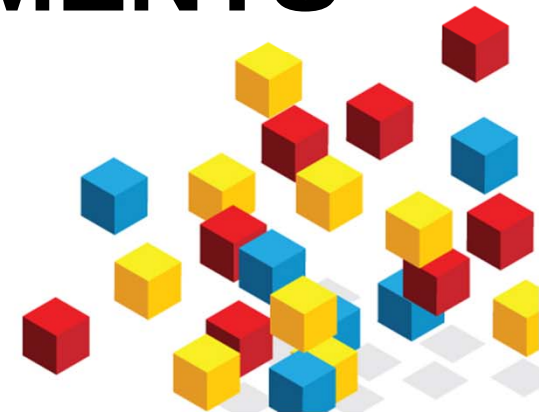
Scalability example





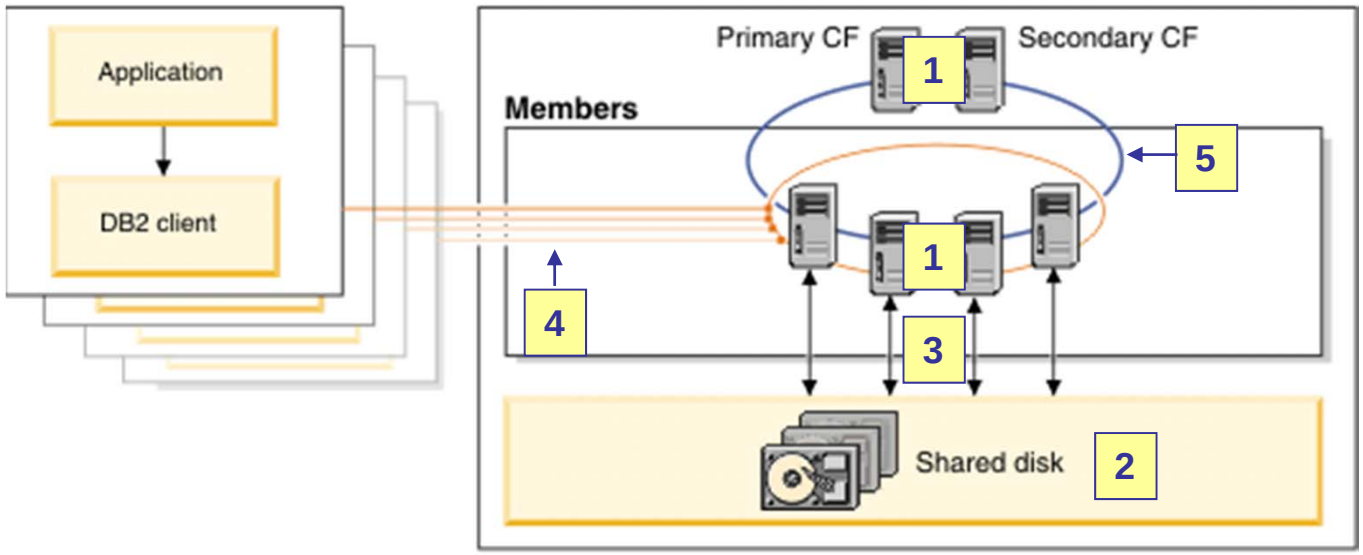
DB2 pureScale

SUPPORTED ENVIRONMENTS

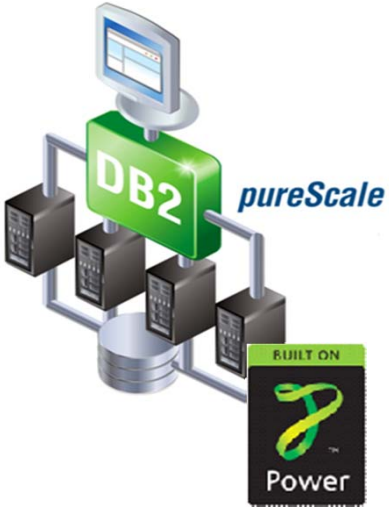


DB2 pureScale: Hardware requirements (POWER)

AIX 6.1 / 7.1



1. POWER6 or POWER7 servers
   - Including P6 550, 595; P7 710, 720, 730, 750, 755, 770, 780, 795
2. Disk storage supported by IBM General Parallel File System (GPFS)
3. Direct SAN connection
4. Ethernet network
5. InfiniBand network
   - IB adapters (varies depending on server type)
   - IB switch (e.g. 7874-024 24-port 4x DDR InfiniBand Edge Switch)



DB2 pureScale POWER6 Clusters – Typical Configuration

DB2 pureScale components

System configuration

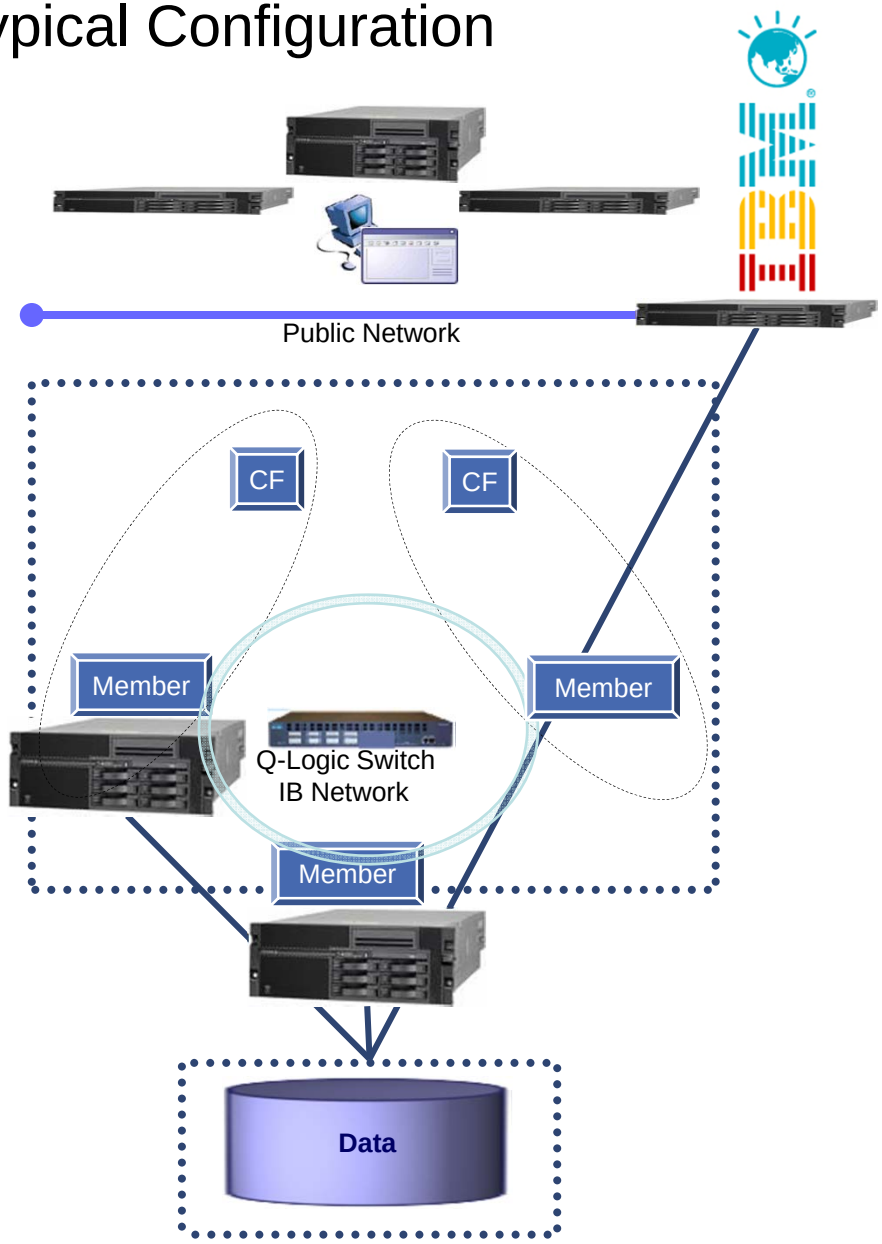
- DB2 pureScale servers & application workload server:
  - 3 x P6 550 servers
    - 8 x 8192MB RDIMMs
    - 4 x 2-core 5.0 GHz cpu
    - Each P550 server split into 2 LPARs

Storage Information

- Internal storage for software installations
  - 6x300GB 15K RPM SAS Disk Drive
- External storage
  - DS5100 with 5 x 16-Pak 146.8 GB/15K DDM
  - Defined as 8 devices of ~500GB each

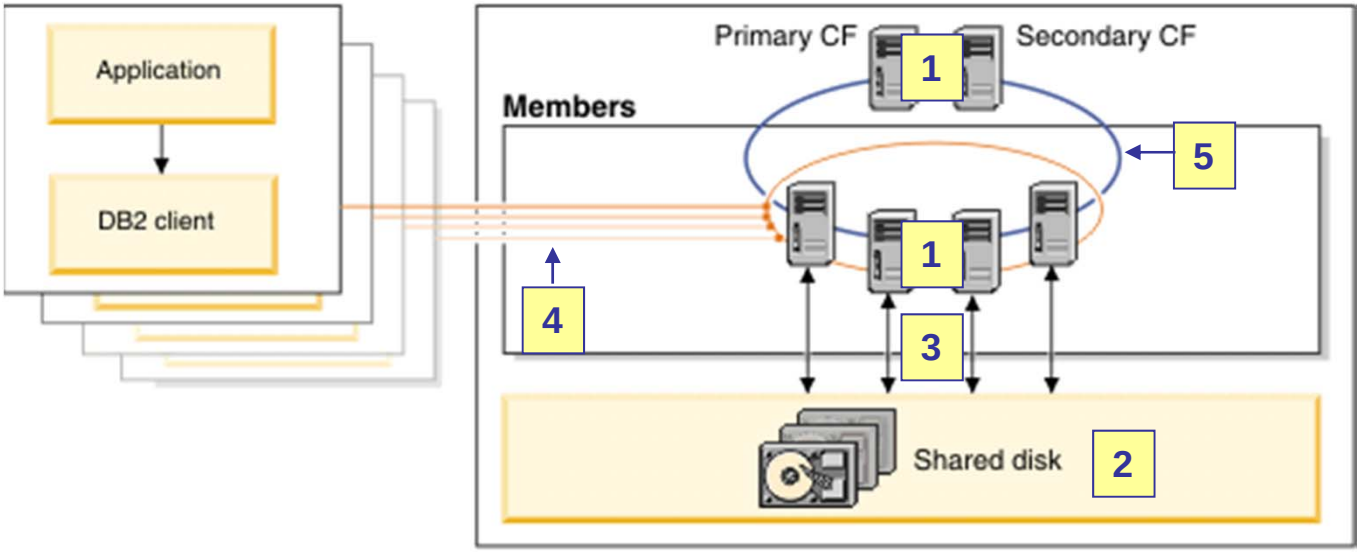
Network Information

- Infiniband switch (7874-024)
- Infiniband (FC 5609) NIC
- Public Ethernet 1GB NIC2



DB2 pureScale: Hardware requirements (Linux)

Suse Linux Enterprise Server 10



1. System x servers
   - Including 3650 M3, 3690 X5, 3850 X5
2. Disk storage supported by IBM General Parallel File System (GPFS)
3. Direct SAN connection
4. Ethernet network
5. InfiniBand network
   - IB adapters (Mellanox ConnectX-2 VPI adapter card)
   - IB switch (Mellanox InfiniScale IV QDR InfiniBand Switch)

- Or 10 Gigabit Ethernet (10GE) network

DB2 pureScale System x Clusters – Typical Configuration

DB2 pureScale components

System configuration

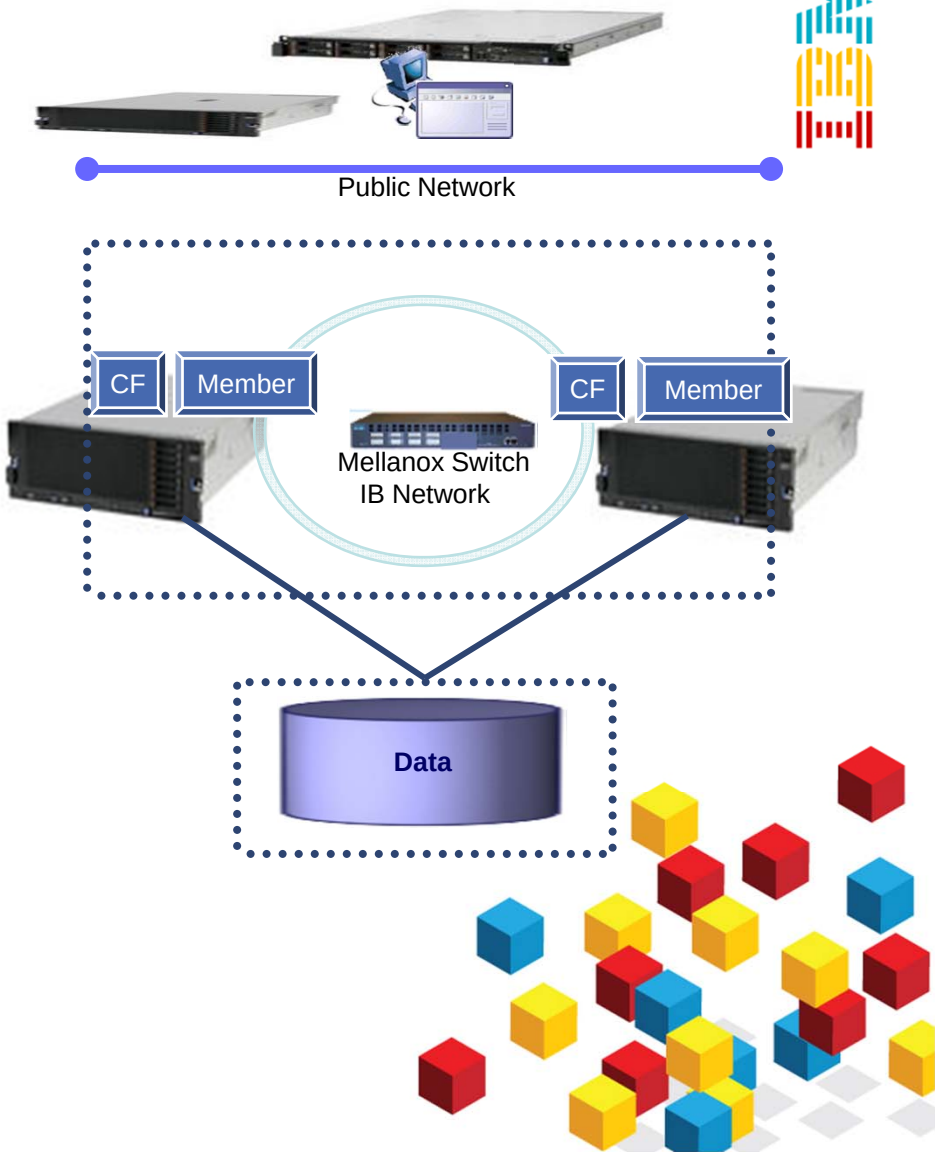
- DB2 pureScale servers & application workload server:
  - 2 x x3850 X5 servers
    - 4 x 4GB DIMMs (Quad-Rank x8)
    - 2 x 4-core 2.27 GHz Xeon 7560 CPU
  - 1 x x3550 M3 server
    - 3 x 8192MB DIMMS
    - 2 x 6-core 2.93 GHz Xeon 5670 CPU

Storage Information

- External storage
  - DS5020 with 5 x 16-Pak 146.8 GB/15K DDM
  - Defined as 3 devices of ~1.4TB, and 1 of ~300MB

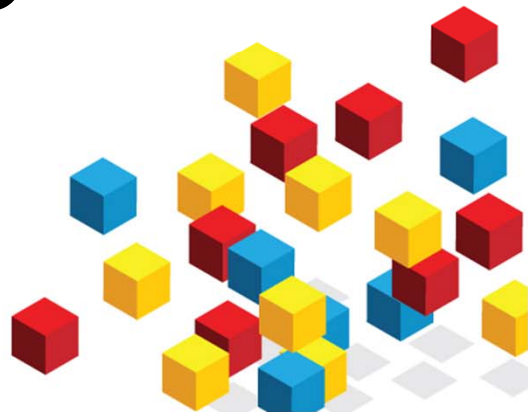
Network Information

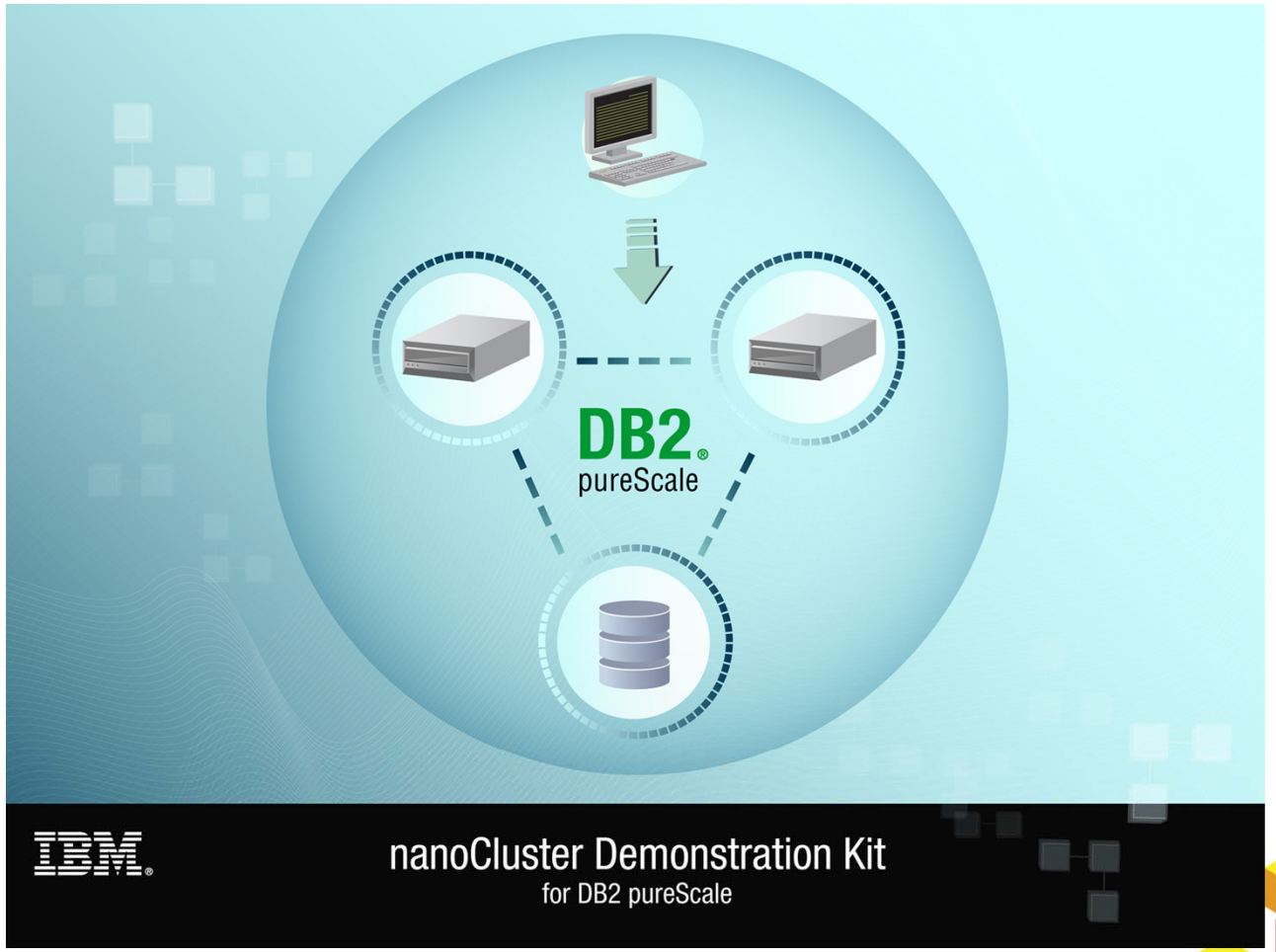
- Mellanox Infiniband switch (MIS5030Q-1SFC)
- Mellanox Infiniband (MHQH29B-XTR) NIC
- Public Ethernet 1GB NIC2



Using the IBM pureScale NanoCluster demokit

DB2 PURESCALE DEMO




DB2
pureScale
IBM
nanoCluster Demonstration Kit
for DB2 pureScale

Software Stack

2 Data Nodes

OS (Novell):
Novell SUSE Linux
Enterprise Server 11 SP1

Software Components:
- DB2 Enterprise 9.8 FP2
- IBM GPFS
- IBM TSAMP

Storage and
Application Node

OS (Novell):
Novell SUSE Linux
Enterprise Server 11 SP1

Software Components:
Tech Explorer for DB2 9.8
DB2 Express-C 9.7.3
WebSphere Application
server 7
Apache
PHP/Perl

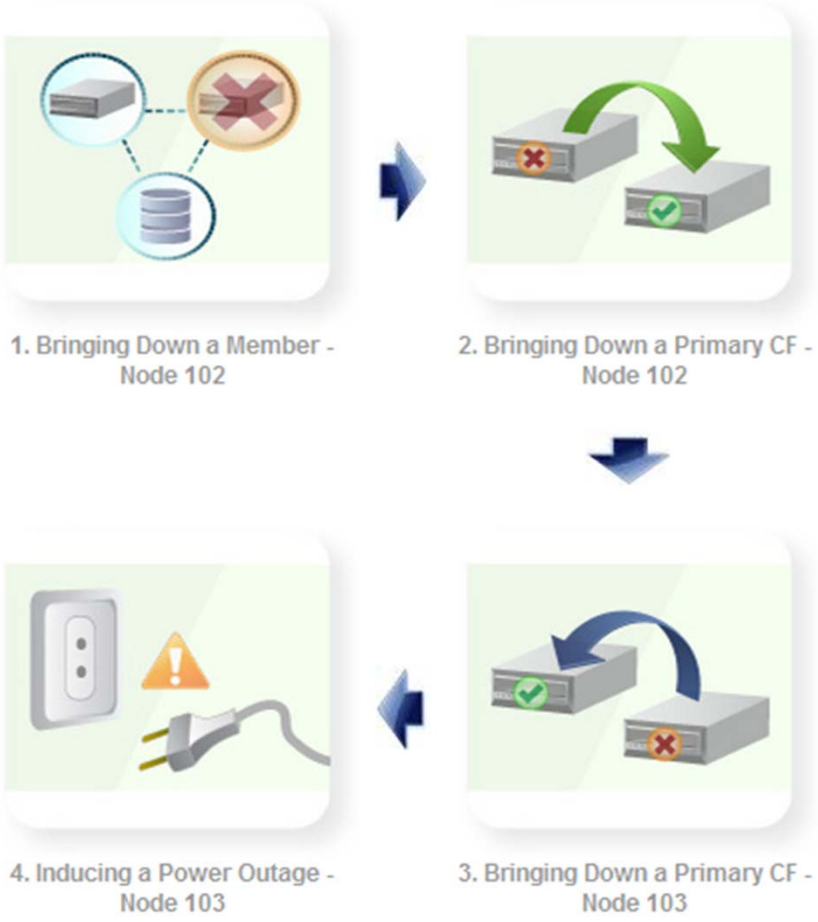
DB2 Discovery Kit
Demo Webpages

Major Components

- **Hardware:**
  - 3x light weight and inexpensive machines based on Intel Atom dual core CPUs, 4GB of RAM, 160GB HDDs
  - 1x Gigabit Ethernet switch
  - 1x 16GB USB flash drive

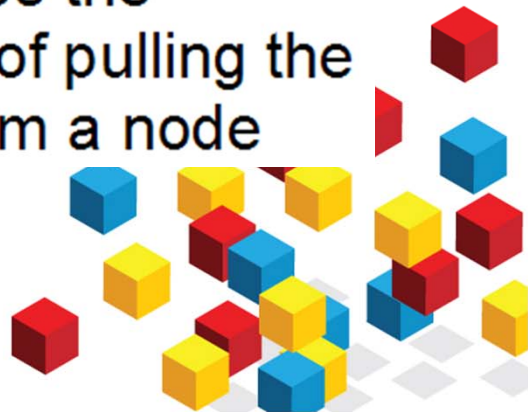
- **Software:**
  - SUSE Linux Enterprise Server 11 SP1
  - DB2 9.8 FP2
- **Demo:**
  - Technology Explorer for DB2 with SDTW workload demo
    - Demonstrate active-active high availability of DB2 pureScale

Test cases



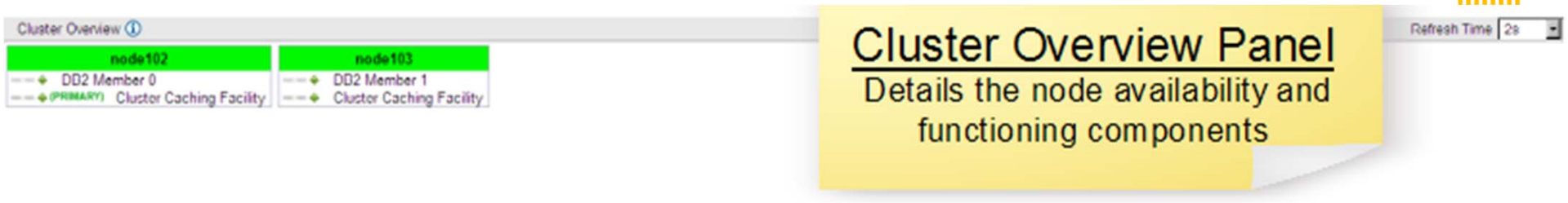
Walkthrough 1 – 3 showcases the recovery process to of bringing down the Member and the CF

Walkthrough 4 showcases the recovery of pulling the power from a node

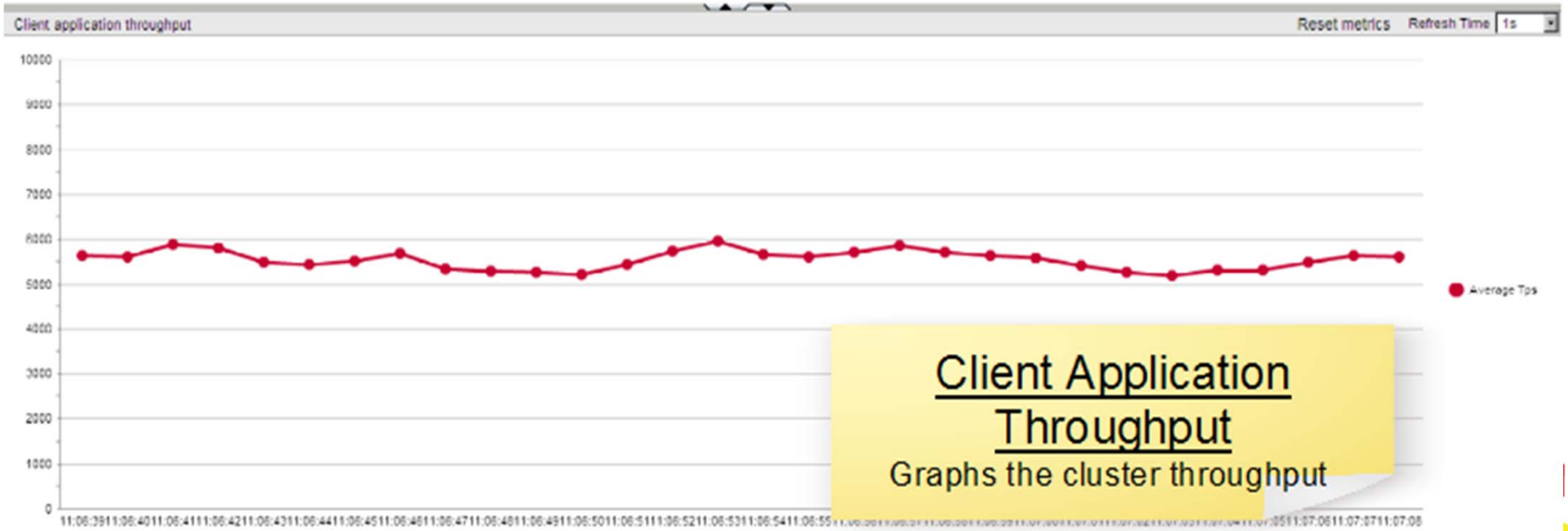


Demonstrations – Technology Explorer for DB2 (TE)

- View current pureScale information



Cluster Overview Panel
Details the node availability and functioning components



Client Application Throughput
Graphs the cluster throughput

DB2 pureScale NanoCluster

DEMO