IBM®

Technical **Forum** & 17 al 21 Octubre 2011

Executive Briefing

Imagine **PODER** Imagine **CAPACIDAD**

DB2 pureScale on POWER
Workload Optimized Systems

Workloads Are Different

Business Analytics	OLTP

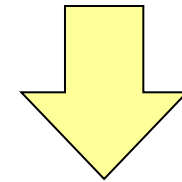**Business Analytics**
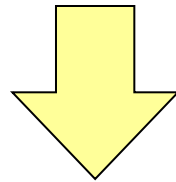
- Complex queries against a data warehouse
- Read only
- Star schema design often used to speed up queries
- Different modes of operation

OLTP

- Transactions against operational data
- Reads and writes
- Normalized schema design to eliminate redundancy
- Multiple user throughput operation

Modes Of Operation Have Different Performance Objectives

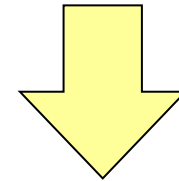| Business Analytics		OLTP
Complex analytic queries	Operational reports with high concurrency	Multiple users running transactions concurrently
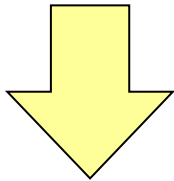
Complete queries in the fastest
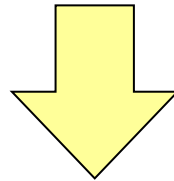
Achieve maximum throughput

Achieve maximum throughput

Competitor's One-Size-Fits-All Database Machine

versus

IBM Workload Optimized Systems

IBM workload optimized systems provide reliable transaction integrity, leading performance, and system flexibility.

Optimized for Workloads that …

Ensure every business transaction completes successfully and valuable data is protected

Provide top speed and 24x365 availability to serve the globally integrated enterprise

Allow configurations that meet today's needs, with rapid and easy growth to keep pace with business

…and built for the way your business really works

Our systems – optimized for transactions – have a proven track record

1950s…1960s

TPF: Airline Reservation System

S/360: 1ST Binary Compatible computer family
IMS: Transaction & Database System

1990s

IMS, CICS, and DB2 Parallel Sysplex:
    High-scale Application and Data Serving

1970s...1980s

System 38 and AS/400:
    Integrated Application and Data Serving
S/370: Online Transaction Processing
DB2: Relational Database System

2000s

DB2 pureScale on PowerHA: High-Scale
    Database Management
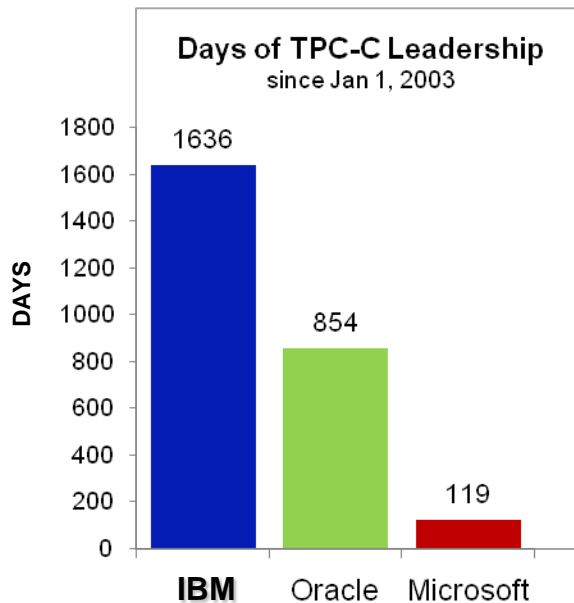**WebSphere Edge Server**: High-scale
    Web Application Serving
DB2 9: Industry's 1st Hybrid Relational &
    XML Database System
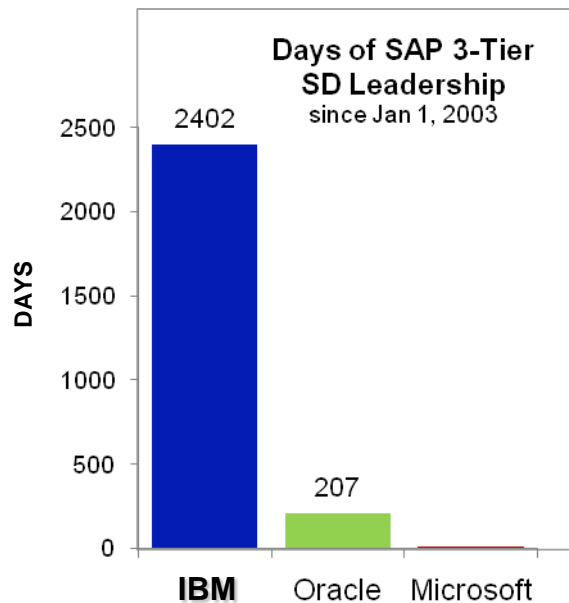**Datapower:** XML & Web services
    appliance
Power 7

Trusted performance is measured by the speed to handle the world's most demanding business transactions.

TPC-C performance leadership

Days of TPC-C Leadership
since Jan 1, 2003

IBM: 1636 | Oracle: 854 | Microsoft: 119

SAP SD performance leadership

Days of SAP 3-Tier SD Leadership
since Jan 1, 2003

IBM: 2402 | Oracle: 207 | Microsoft: (none)

SPECjEnterprise2010 Performance leadership

Days of SPEC Enterprise 2010 Leadership
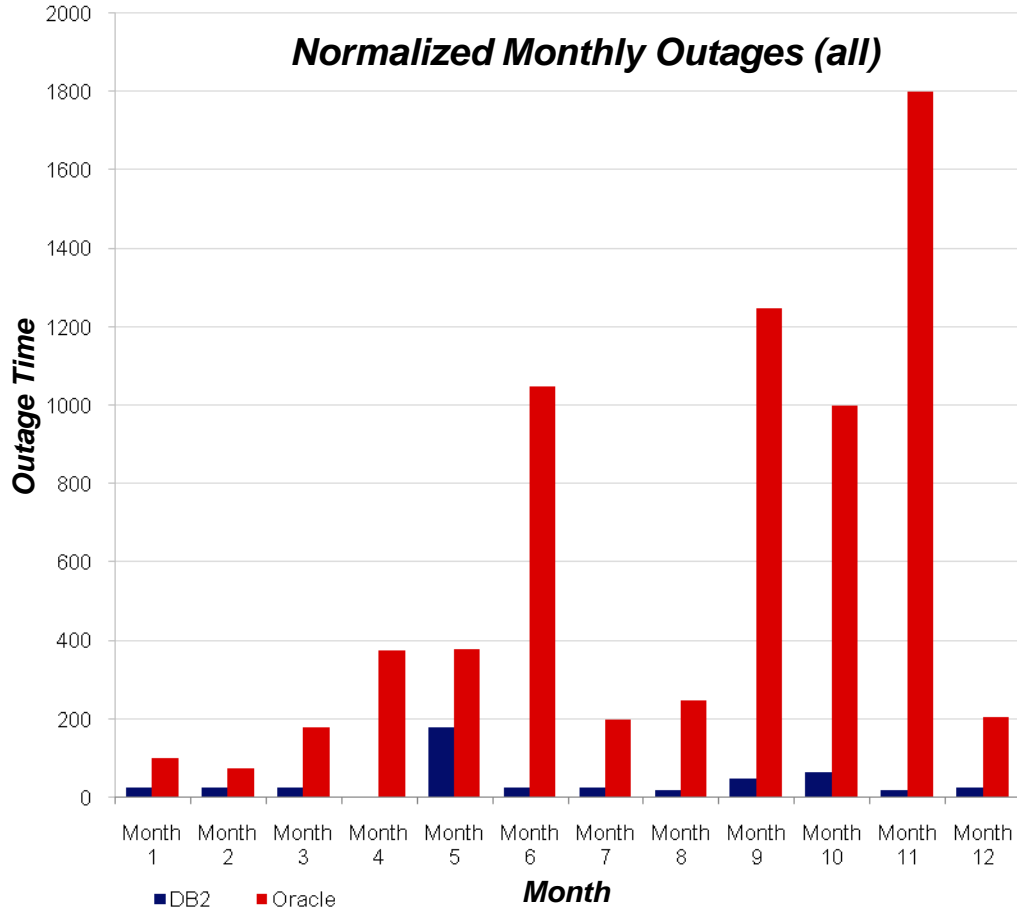
IBM: 60 | Oracle: 0
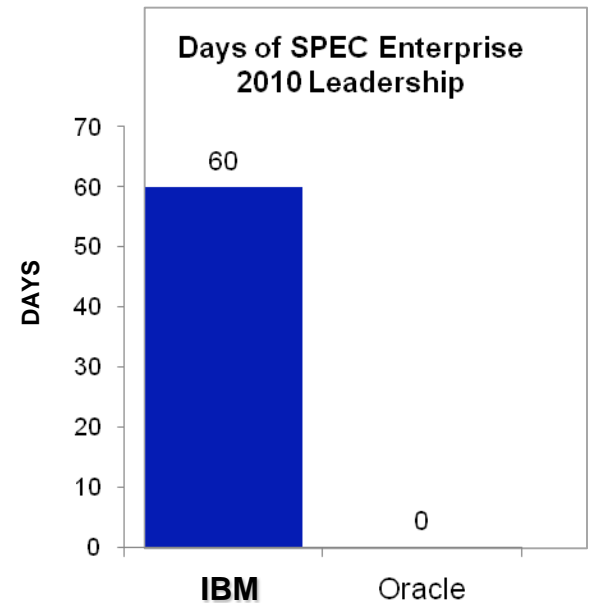
Source: IBM-maintained records of performance benchmark leadership. TPC-C and SAP 3-Tier SD leadership days are up to and including 22 Feb 2010. SPECjEnterprise2010 up to 05 Mar 2010.

IBM offers the highest levels of reliability for critical business systems.

Database Reliability*

Normalized Monthly Outages (all)

Chart showing Outage Time (y-axis, 0 to 2000) vs Month (x-axis, Month 1 through Month 12), comparing DB2 (dark blue) and Oracle (red).

- Month 1: DB2 ~25, Oracle ~100
- Month 2: DB2 ~25, Oracle ~75
- Month 3: DB2 ~25, Oracle ~180
- Month 4: Oracle ~375
- Month 5: DB2 ~180, Oracle ~380
- Month 6: DB2 ~25, Oracle ~1050
- Month 7: DB2 ~25, Oracle ~200
- Month 8: DB2 ~20, Oracle ~250
- Month 9: DB2 ~50, Oracle ~1250
- Month 10: DB2 ~65, Oracle ~1000
- Month 11: DB2 ~20, Oracle ~1800
- Month 12: DB2 ~25, Oracle ~205

■ DB2 ■ Oracle

Data Availability**

Line chart showing % of Data Available vs Time (~seconds), with Node Failure indicated. Comparing DB2 pureScale (green) and Oracle RAC (red).

— DB2 pureScale
— Oracle RAC

Node Failure

* Source: Solitaire whitepaper: DB2 Performance on IBM System p® and System x®11g
** Source: Internal IBM testing for DB2 pureScale availability combined with several public sources for Oracle RAC availability.

7

Do You Know -

What the Cost of a One Hour Outage is?

?

Do You Know -

What the Cost of a One Hour Outage is?

Industry	Average Cost Downtime Per Hour
Brokerage Services	$6.5 million
Energy	$2.8 million
Credit Card	$2.6 million
Telecomm	$2.0 million
Manufacturing	$1.6 million
Financial	$1.5 million
Retail	$1.1 million
Pharmaceutical	$1.0 million
Industry Average	$2.4 million

Source: Network Computing, The Meta Group, and Contingency Planning Research

DB2 pureScale
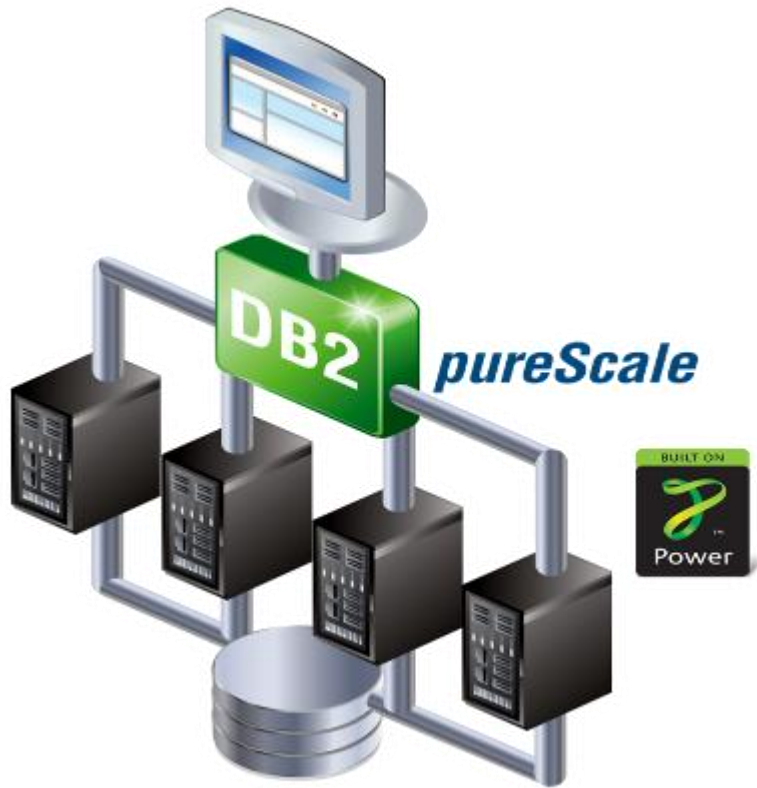
Unlimited Capacity

- Buy only what you need, add more without service interruption

Application Transparency

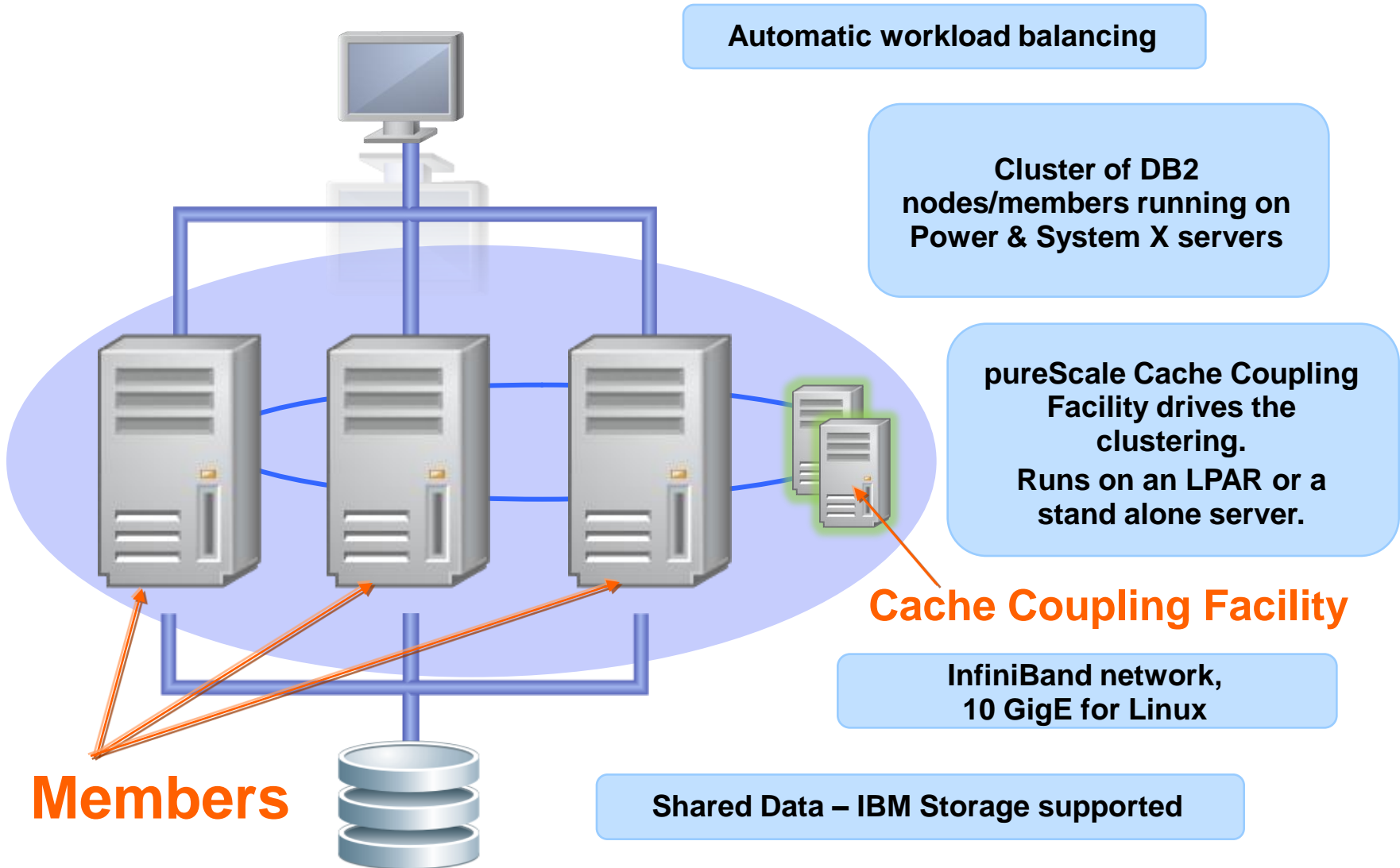- Avoid the risk and cost of application changes

Continuous Availability

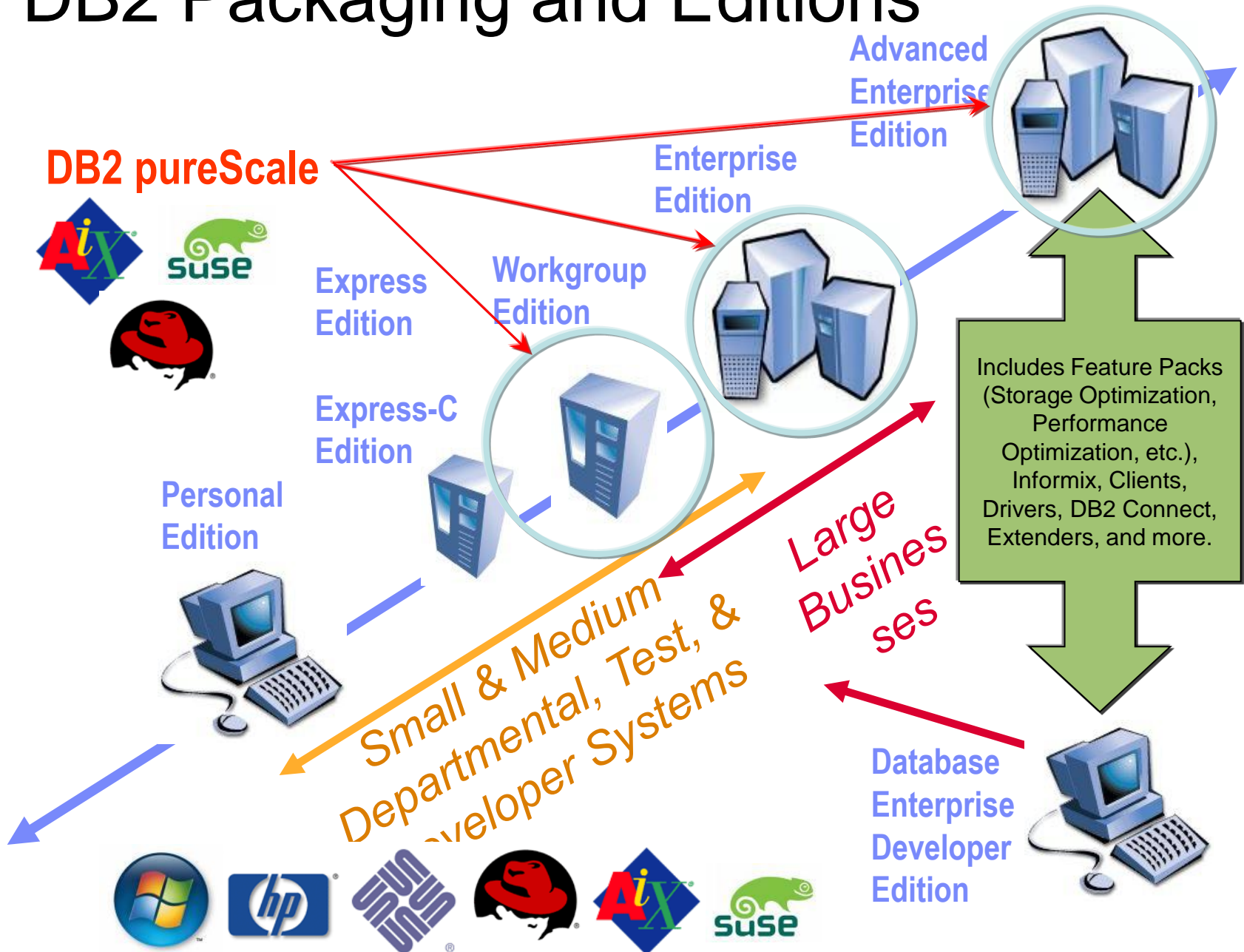- Deliver uninterrupted access to your data with consistent performance

DB2's new OLTP continuous availability feature provides unlimited capacity, transparent to applications

Leverages the architecture of z/OS: the Gold Standard of reliability and scalability

DB2 pureScale Architecture

Automatic workload balancing

Cluster of DB2 nodes/members running on Power & System X servers

pureScale Cache Coupling Facility drives the clustering.
Runs on an LPAR or a stand alone server.

Cache Coupling Facility

InfiniBand network, 10 GigE for Linux

Members

Shared Data – IBM Storage supported

DB2 Packaging and Editions

Advanced Enterprise Edition

DB2 pureScale

Enterprise Edition

Express Edition

Workgroup Edition

Express-C Edition

Personal Edition

Includes Feature Packs (Storage Optimization, Performance Optimization, etc.), Informix, Clients, Drivers, DB2 Connect, Extenders, and more.

Small & Medium

Departmental, Test, & Developer Systems

Large Businesses

Database Enterprise Developer Edition

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
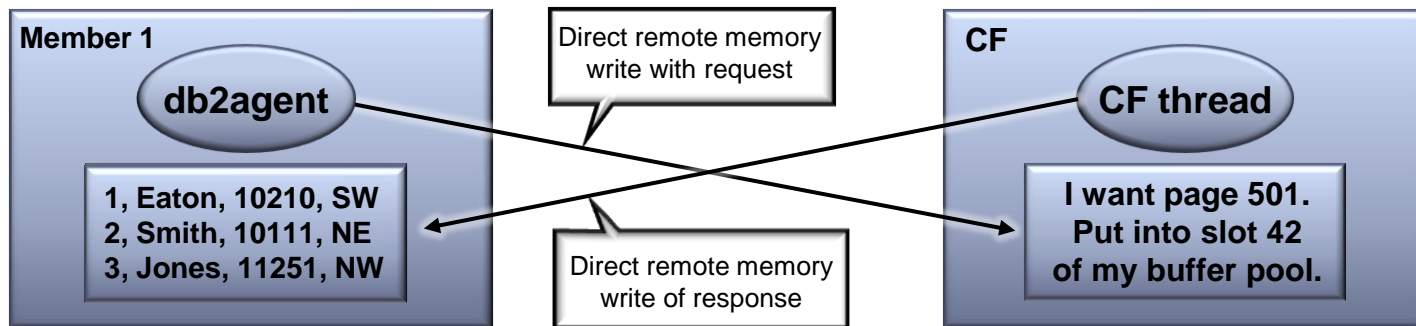   - No IP socket calls, no interrupts, no context switching

- 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

Member 1	Member 1	Member 1

CF

Group Buffer Pool

Group Lock Manager

PowerHA pureScale
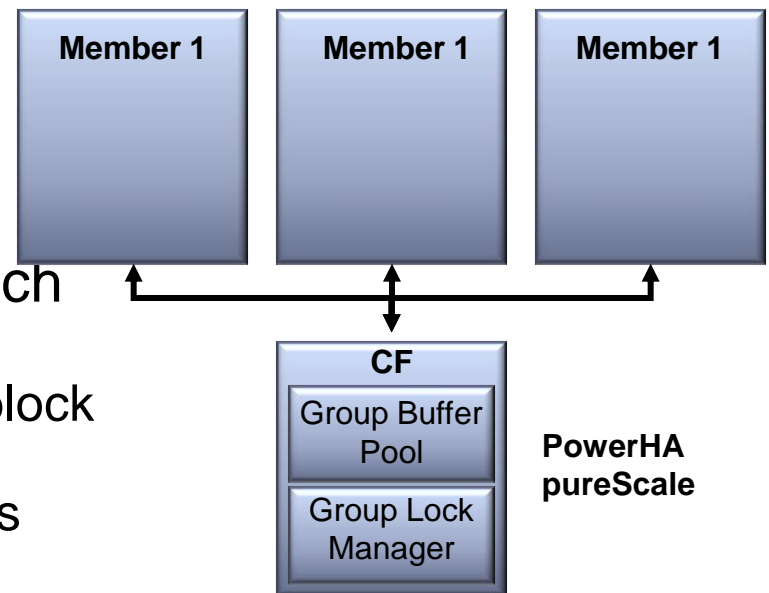
The Advantage of DB2 Read and Register with RDMA

1. DB2 agent on Member 1 writes directly into CF memory with:
   – Page number it wants to read
   – Buffer pool slot that it wants the page to go into
2. CF either responds by writing directly into memory on Member 1:
   – That it does not have the page   **or**
   – With the requested page of data
- Total end to end time for RAR is measured in microseconds
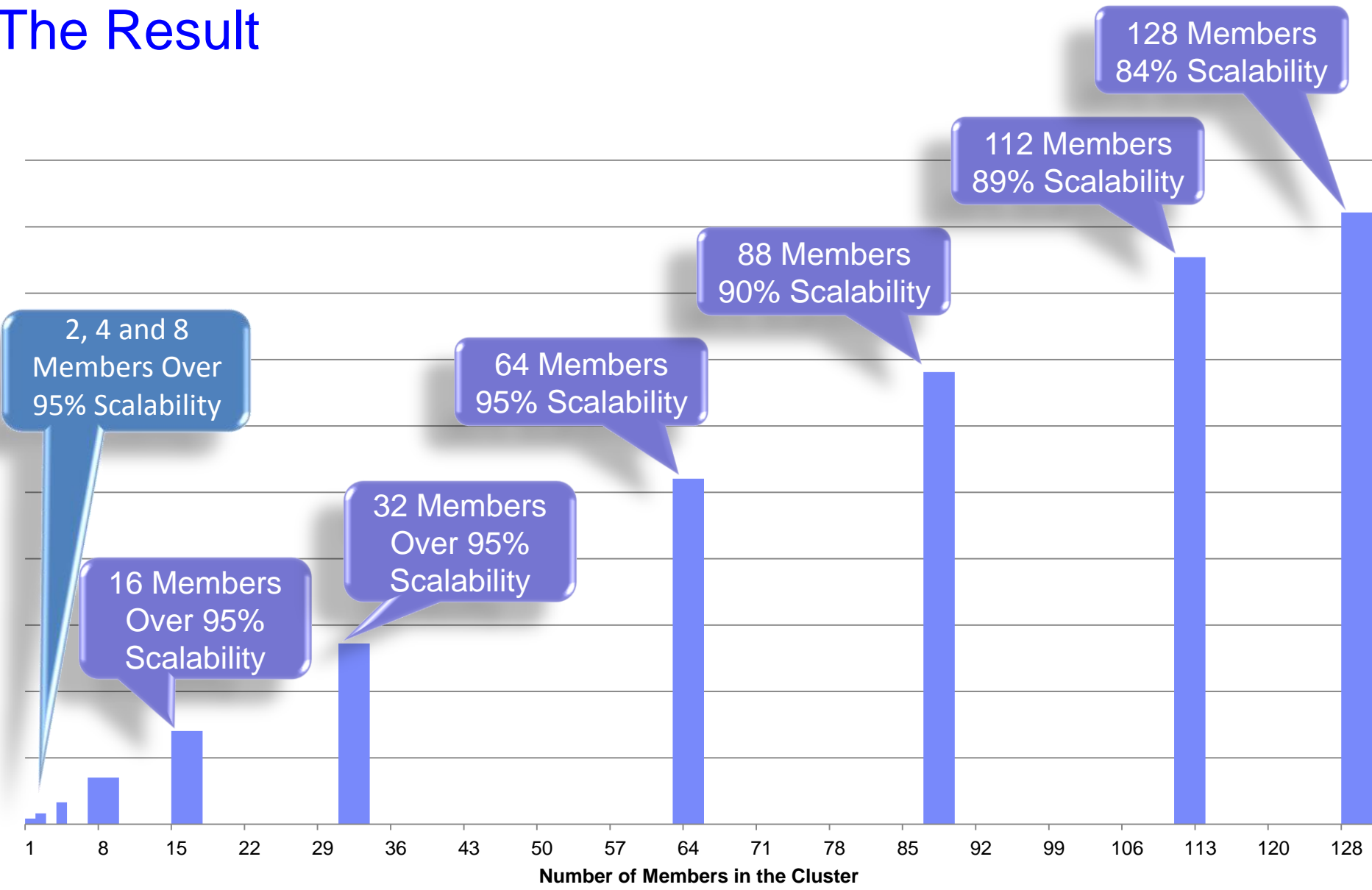- Calls are very fast, the agent may even stay on the CPU for the response

Member 1

db2agent

1, Eaton, 10210, SW
2, Smith, 10111, NE
3, Jones, 11251, NW

Direct remote memory write with request

Direct remote memory write of response

CF

CF thread

I want page 501.
Put into slot 42
of my buffer pool.

Much more scalable, does not require locality of data

Proof of DB2 pureScale Architecture Scalability

- How far will it scale?

- Take a web commerce type workload
  - Read mostly but **not read only**

- Don't make the application cluster aware
  - **No routing of transactions to members**
  - Demonstrate transparent application scaling

- Scale out to the 128 member limit and measure scalability

The Result

2, 4 and 8 Members Over 95% Scalability

16 Members Over 95% Scalability

32 Members Over 95% Scalability

64 Members 95% Scalability

88 Members 90% Scalability

112 Members 89% Scalability

128 Members 84% Scalability

Number of Members in the Cluster

1   8   15   22   29   36   43   50   57   64   71   78   85   92   99   106   113   120   128

Unlimited Capacity

- DB2 pureScale has been designed to grow to whatever capacity your business requires

- Flexible licensing designed for minimizing costs of peak times

- Only pay for additional capacity when you use it even if for only a single day

Solution:
Use DB2 pureScale and add another server for those two days, and only pay sw license fees for the days you use it.
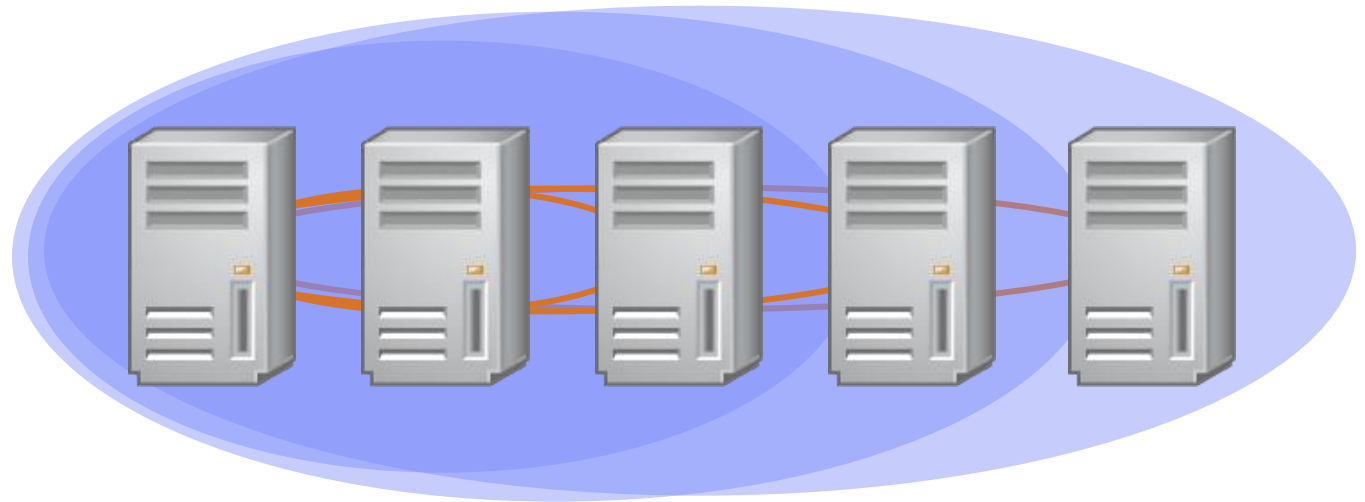
Over 100+ node architecture validation has been run by IBM

Application Transparency

Take advantage of extra capacity instantly

- No need to modify your application code
- No need to tune your database infrastructure
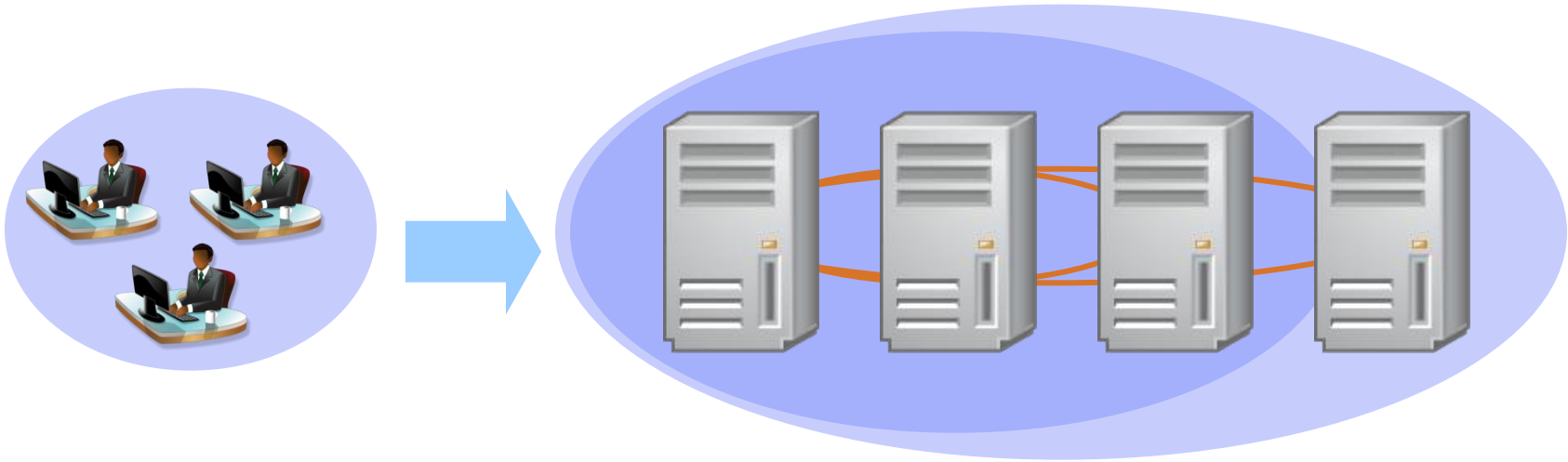
Your DBAs can add capacity without re-tuning or re-testing

Your developers don't even need to know more nodes are being added

Continuous Availability

- Protect from infrastructure outages
  - Architected for no single point of failure

Automatic workload balancing

Duplexed secondary global lock and memory manager

Tivoli System Automation automatically handles all component failures

DB2 pureScale stays up even with multiple node failures

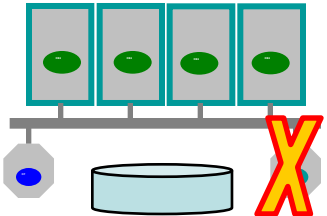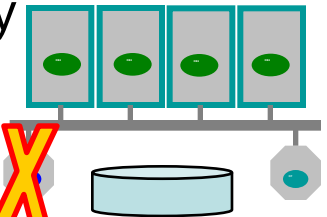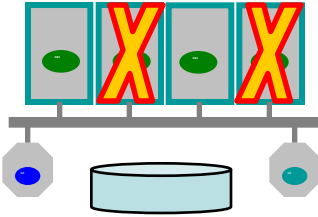Shared disk failure handled using disk replication technology
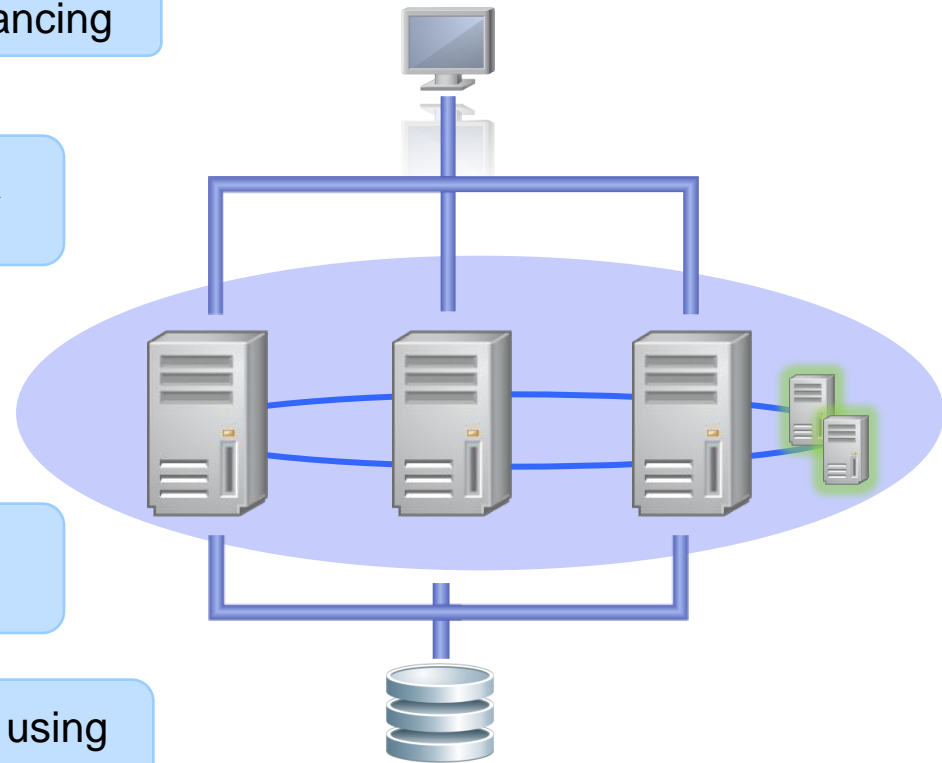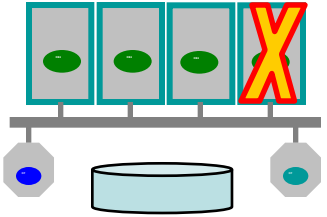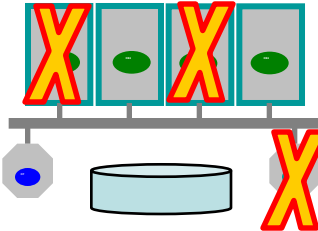
Recover Instantaneously From Node Failure
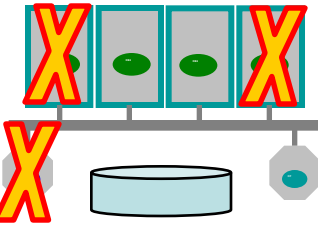
- **Protect from infrastructure related outages**
  - Redistribute workload to surviving nodes immediately
  - Completely redundant architecture
  - Recover in-flight transactions on failing node in as little as 15 seconds including detection of the problem

Application Servers and DB2 Clients
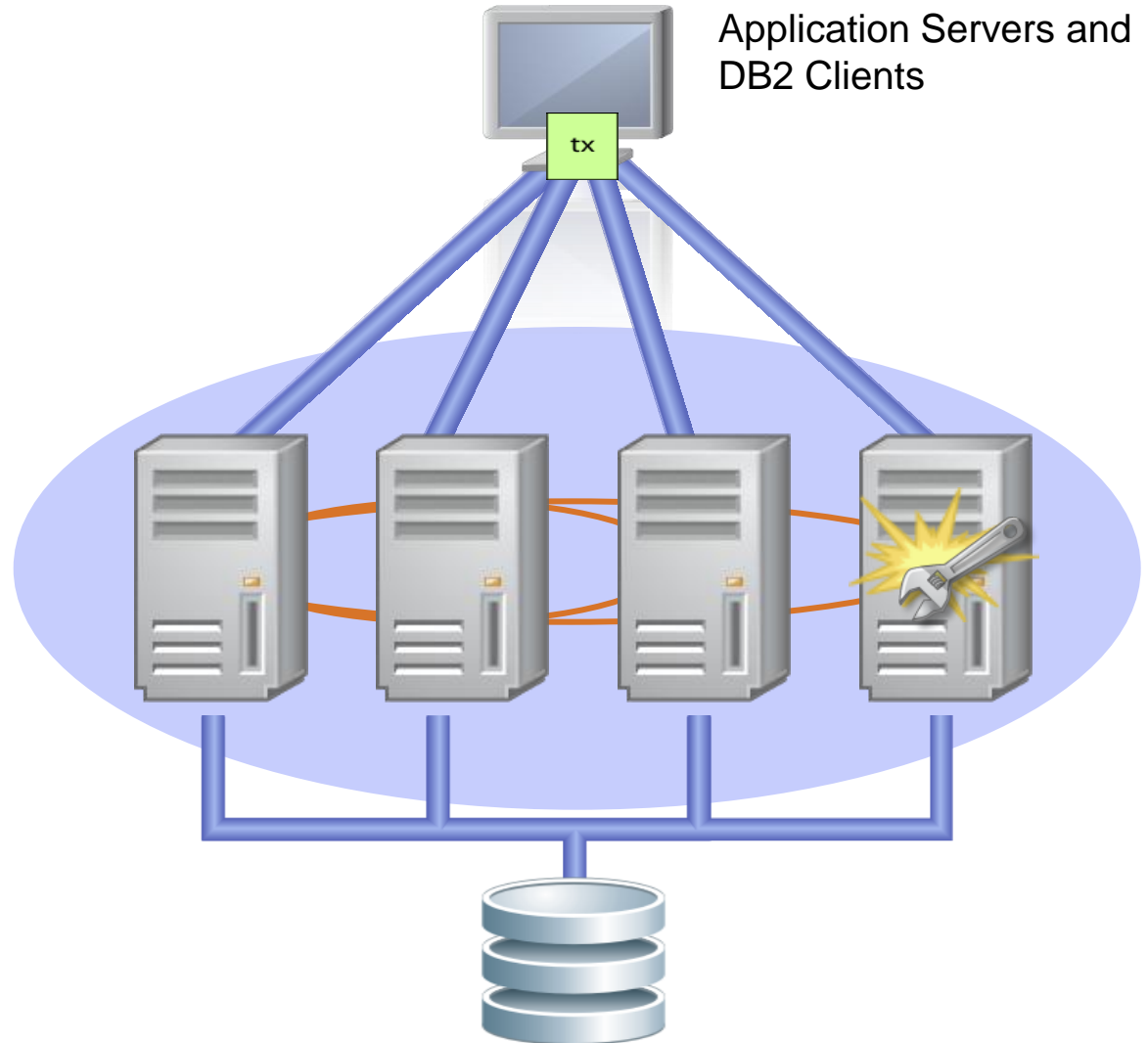
Summary (Single Failures)

Failure Mode	Other Members Remain Online ?	Automatic & Transparent ?	Comments
Member	✔	✔ Connections to failed member transparently move to another member	Only data that was inflight on failed member remains locked temporarily.
Primary *PowerHA pureScale*	✔	✔	Momentary "blip" in CCF *pureScale* service. Transparent to members (In-flight *CCF pureScale* requests just take a few more seconds before completing normally.)
Secondary *PowerHA pureScale*	✔	✔	Momentary "blip" in *PowerHA pureScale* service. Transparent to members (In-flight *PowerHA pureScale* requests just take a few more seconds before completing normally.)

Simultaneous Failures

Failure Mode	Other Members Remain Online ?	Automatic & Transparent ?	Comments
	✔	✔ Connections to failed member transparently move to another member	Only data that was inflight on failed members remains locked temporarily. Recoveries done in parallel.
	✔	✔ Connections to failed member transparently move to another member	Same as member failure. Momentary, transparent, "blip" in *PowerHA pureScale* service.
	✔	✔ Connections to failed member transparently move to another member	Same as member failure. Momentary, transparent, "blip" in *PowerHA pureScale* service.

Moving Your Applications to DB2 is Easy

Break Free
From High Database Costs

Proven Results

- Easily move your applications from Oracle database to DB2

- Leverage existing skills and people without re-training

- Applications moved to DB2 run quickly with full native execution

- Customers are no longer locked into Oracle RAC

- Integrated, cross-platform tools supporting Oracle database as well

- Customers and partners have moved in only days

Summary – What can DB2 pureScale Do For You?

- Deliver higher levels of scalability and superior availability

- Better concurrency during regular operations

- Better concurrency during member failure

- Result in less application design and rework for scalability

- Improved SLA attainment

- Lower overall costs for applications that require high transactional performance and ultra high availability

Questions & Answers

Thank you!

Ibm.com/software/data/infosphere/smart-analytics-system

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

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.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, 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.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput 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 improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.