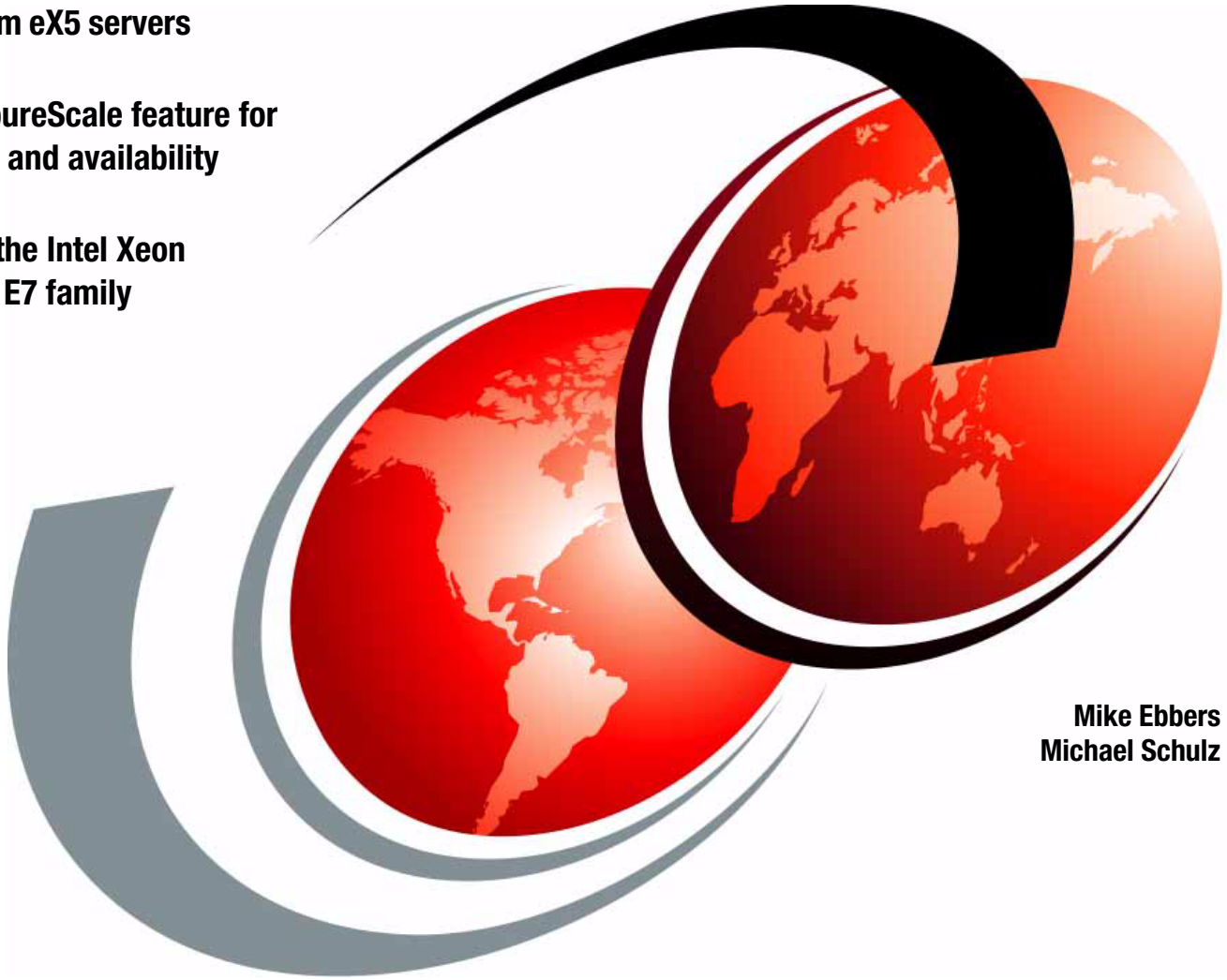
IBM® Information Management Software

Highly Available and Scalable Systems with IBM eX5 and DB2 pureScale

- Introduction to the new IBM System eX5 servers
- IBM DB2 pureScale feature for scalability and availability
- Based on the Intel Xeon processor E7 family



Mike Ebbers
Michael Schulz

Redpaper

ibm.com/redbooks

International Technical Support Organization

Highly Available and Scalable Systems with IBM eX5 and DB2 pureScale

April 2011

REDP-4742-00

Note: Before using this information and the product it supports, read the information in “Notices” on page v.

First Edition (April 2011)

This edition applies to the IBM System x eX5 family of servers and IBM DB2 pureScale feature for Enterprise Server Edition 9.8.0 5724-Y69.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of 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.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US 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 http://www.ibm.com/legal/copytrade.shtml

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

DB2®
DS8000®
Dynamic Infrastructure®
GPFS™
IBM Systems Director Active Energy Manager™
IBM®
OnForever™
pureScale™
Redbooks®
Redpaper™
Redbooks (logo)  ®
Smarter Planet™
System x®
X-Architecture®
z/OS®

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

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo 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 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.

Preface

The pressures related to managing transactional databases are increasing rapidly. Business growth, the drive to consolidate databases, and the need to deploy new, data-intensive technologies are fostering the massive expansion of data volumes and application workloads. At the same time, the move toward real-time computing requires faster and more reliable data access, especially when databases are used to drive client-facing applications. Businesses need simpler and more cost-effective strategies for expanding their database environment.

IBM® and Intel® provide an answer to this challenge with the IBM DB2® pureScale™ feature and the latest generation of IBM System x® eX5 servers based on the Intel Xeon® E7 processor family. The combined solution enables clients to scale mission-critical, performance-sensitive databases simply, using affordable, industry-standard servers.

This IBM Redpaper™ publication describes an IBM solution containing high availability and scalability for mission-critical databases on System x. The audience includes executives and other decision-makers, consultants, and architects.

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Mike Ebbers is an IBM Consulting IT Specialist and Project Leader at the International Technical Support Organization (ITSO), Poughkeepsie Center. He has worked with IBM mainframe hardware and software products since 1974 in the field, in education, and in the ITSO.

Michael Schulz is a Competitive Consultant working on the Information Management Technology Ecosystems team in IBM Houston. He works with clients and IBM Business Partners to move their applications to the IBM Information Management portfolio.

Thanks to the following people for their contributions to this project:

David Watts
International Technical Support Organization, Raleigh Center

Paul Field
IBM Raleigh

Yvonne Chan
Fraser McArthur
IBM Toronto

Greg Cnossen
Dennis B. Fallis
Intel

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks® publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks
- Send your comments in an email to:

  redbooks@us.ibm.com
- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks
- Follow us on Twitter:

  http://twitter.com/ibmredbooks
- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806
- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm
- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

The need for high-performing clusters

The pressures related to managing transactional databases are increasing rapidly. Business growth, the drive to consolidate databases, and the need to deploy new, data-intensive technologies are fostering the massive expansion of data volumes and application workloads. At the same time, the move toward real-time computing requires faster and more reliable data access, especially when databases are used to drive client-facing applications. Businesses need simpler and more cost-effective strategies for expanding their database environment. IBM and Intel provide an answer to this challenge with the IBM DB2 pureScale feature and the latest generation of IBM System x eX5 servers based on the Intel Xeon E7 processor family. The combined solution enables clients to scale mission-critical, performance-sensitive databases simply, using affordable, industry-standard servers.

With the IBM DB2 pureScale feature[1], multiple IBM System x servers can be combined into a highly performing cluster that provides virtually unlimited capacity with continuous availability and complete application transparency. A new server can be integrated into the cluster with a few commands. No application changes are required, and workloads are automatically rebalanced to optimize performance. This is a breakthrough capability that makes it far easier to handle today's exploding data volumes. IT organizations can now take a building block approach to database expansion by scaling large, mission-critical databases incrementally and affordably.

- IBM integrates hardware, software, services and support into comprehensive, highly optimized solutions that deliver better and more reliable value out of the box.
- IBM DB2 software enhancements such as autonomic management, deep compression (up to 80%), and native XML support provide higher business value at substantially lower total costs.
- Modular server designs offer pay-as-you-grow scalability, affordable starting points, and easy expansion up to eight sockets (80 cores, 160 threads) and 6 TB of memory per two-node server.

This IBM Redpaper publication presents a business case for the solution and discusses its components.

[1] http://www-01.ibm.com/software/data/db2/linux-unix-windows/editions-features-purescale.html

2

Dealing with business growth

We live in a time of rapid data growth. Business growth can accelerate this increase, as can mergers, acquisitions, new application deployments, and application upgrades that increase data processing workloads. Database solutions must be able to absorb increasing data volumes and expanding workloads, while sustaining performance and avoiding disruptive capacity upgrades.

2.1 The effect of growth on transactional databases

Transactional databases are the lifeblood of an enterprise. Critical data must be available quickly and without interruption to keep the business running at peak efficiency. If a key database goes down or slows down, dependent applications suffer the same fate, often costing the business millions of dollars in lost revenue and creating dissatisfied clients. Unplanned downtime is not the only risk: in today's global, 24x7 business environment, taking down a key database for upgrades or maintenance can be very costly.

In almost every industry you can name, whether retail, manufacturing, telecom, government, or finance, transactions are the fuel that drives the business. They can be credit card payments at the mall, sell orders on the stock exchange, hotel bookings, or mobile phone call connections.

More transactions mean more money coming in, more orders going out, more communication with clients and suppliers, more activity, and more business. In an increasingly instrumented, interconnected, and smarter world, transaction volumes will only grow; estimates of the amount of data created each year have been rising exponentially, and that trend shows no sign of abating in the years ahead.

2.1.1 The challenges of scaling mission-critical databases

Organizations that master online transaction processing (OLTP) and develop the ability to easily handle ever-larger transaction volumes have the opportunity to differentiate themselves from the competition and stay a step ahead of the market. For businesses ready to take this next step, their choice of an OLTP platform is critical. It must not only provide outstanding performance; it must also deliver that performance every minute of every day, regardless of fluctuating loads or maintenance needs. It must also stand ready to expand and adapt smoothly as the organization it supports grows and changes.

Traditionally, scaling capacity for enterprise databases has been particularly challenging because of the need to host the database on a single server. Even a large multiprocessor system can take only so many processors, memory chips, and I/O cards before reaching its full capacity. The next step would be to replace the entire system, which is an inherently expensive proposition.

2.2 The solution of clustering

What is needed is a cost-effective system that is incrementally scalable, allowing it to grow to handle increases in the workload.

Clustered database solutions offer an alternative approach to database scaling, one that has been available for several years. However, expanding a cluster has been a resource-intensive operation, requiring substantial time and expertise to repartition data, rebalance workloads, and optimize performance.

2.2.1 IBM eX5 servers and IBM DB2 pureScale to the rescue

IBM and Intel have been working for years to create systems that change the economics of x86 computing. One result of that collaboration is an OLTP platform for clustered, distributed systems that delivers unprecedented flexibility and scalability. IBM System x eX5 servers based on the Intel Xeon processor E7 product family provide an ideal platform for

data-intensive workloads in mission-critical environments, such as the x3850 x5 shown in Figure 2-1.

Figure 2-1 IBM System x3850 X5

The DB2 pureScale feature, an optional feature of DB2 for Linux®, UNIX® and Microsoft® Windows® (DB2 for LUW), enables these servers to be clustered together to support truly massive workloads with excellent performance, linear scalability, and high availability. Together, IBM eX5 servers and DB2 pureScale are fundamentally changing the economics of scaling large, mission-critical databases.

With the release of the eX5 servers with more processing power and memory scalability than ever before in an x86 platform, this paper shows you how you can get the most out of these high-performance servers. The IBM eX5 servers allow you to scale to new performance highs, process transactions quickly, improve reliability, and get query results even faster than on previous generation servers.

Capacity and availability with DB2 pureScale

DB2 pureScale provides near-limitless capacity and continuous availability running on System x servers powered by Intel Xeon processors. With this new solution from IBM, organizations can build responsive, highly available OLTP systems that quickly scale in any direction needed, while remaining completely transparent to enterprise applications.

DB2 pureScale takes proven DB2 technology and combines it with the latest distributed system techniques. In a DB2 pureScale cluster, one server, which can be duplexed to avoid a single point of failure, provides centralized lock management services, a global cache for data pages, and other services. This eliminates the challenges and limitations of hosting databases on distributed systems. Members of the pureScale cluster can simultaneously access a shared database for read and write operations. They are also connected directly to the central server through remote direct memory access (RDMA) and DB2 pureScale technology, running over either high-performance InfiniBand or 10 Gb network links.

The result is a cluster of DB2 systems that, to OLTP applications, looks and acts like a single fast server. DB2 pureScale technology raises the bar for availability and scalability in OLTP systems. Downtime can be virtually eliminated because if a cluster node fails, the system automatically fails over to an active node. DB2 provides full access to all pages of data that do not need to be recovered, and interrupted transactions are recovered or rolled back within seconds.

During the recovery, other members of the cluster can continue to process transactions, so the systems and clients that depend on the cluster can continue business as usual. Likewise, planned outages become a thing of the past because IT departments can perform rolling maintenance one server at a time, without bringing down the entire system. DB2 pureScale clusters can be easily and quickly scaled up; adding new servers simply requires a few commands. You can take a building block approach to database expansion, beginning with two partially populated servers and adding capacity incrementally and without limit as requirements grow.

IBM has also built flexibility into DB2 pureScale licensing: clients can pay for the DB2 and DB2 pureScale they need, when they need it, and only for the period of time in use. That is welcome news for IT managers who are looking to cut costs. The great news for developers is that no application changes are needed to benefit from DB2 pureScale clusters. Applications do not need to be cluster-aware, so there is no need to alter or replace existing code. And, because DB2 offers compatibility technologies such as native support for commonly used syntax and the PL/SQL procedural language, organizations can run applications written for other database platforms with few or no code alterations.

System x servers grow with your business

The IBM System x3850 X5 server allows freedom of choice with extremely flexible configurations plus memory expansion capabilities. A modular building block design lets you customize your system for current needs while providing the ability to react to changing workloads. Expand from a two-processor system up to four processors. Add a second system to create an eight-processor system. Start with two memory DIMMs and expand up to 192 with a dual-node system and two MAX5 memory expansion options that provide FlexNode capabilities for dynamic partitioning. Reallocate resources as your environment changes. The x3850 X5 meets your needs today, while providing an easy, cost-effective upgrade path to change your environment when you are ready.

Do more with your investments

Six-core, eight-core, and ten-core processor options and expanded memory capacity let you manage memory-intensive applications with higher utilization and greater throughput and bandwidth. Consolidate older servers onto fewer industry-leading IBM X-Architecture® eX5 systems for a lower total cost. Optimize your investments with more and larger virtual machines per server and lower licensing costs, while achieving greater business productivity. The MAX5 memory expansion drawer with large-scale memory capacity enables organizations to place an entire database in memory and run queries in real time, speeding the path from data to decision.

Consolidate with confidence

A long-time leader in reliability, many of the features IBM pioneered with earlier generations of X-Architecture have now become industry standards. The x3850 X5 continues to build on this leadership, enabling you to consolidate workloads without sacrificing availability or your peace of mind. OnForever™ features such as redundant power and cooling help minimize the potential for catastrophic business disruption. Other features include enhanced IBM Memory ProteXion for greater data integrity, memory mirroring, and Predictive Failure Analysis for increased system uptime.

Own and operate with ease

Consolidated single points of management and an energy-smart design help reduce the time and money needed to run your system. Advanced light path diagnostics and around-the-clock remote access management facilitate ease of use and proactive problem-solving. In addition, IBM Systems Director Active Energy Manager™ helps monitor, measure, and manage power consumption to lower the energy draw and costs.

3

A client case study

IBM DB2 pureScale feature on System x3690 X5 servers powered by Intel Xeon processors helps companies provide uninterrupted data access for mission-critical applications. Here is how one company is applying this solution.

3.1 Overview of case study company

This case study involved a global distributor of communication and security products, electrical and electronic wire and cable, fasteners, and other small parts, which has more than 100,000 clients in 52 countries resulting in $5 billion in sales.

The company combines superior product offerings with unparalleled technological expertise and innovative supply chain services to provide customized solutions designed to meet each client's specific needs. The company's innovative inventory management programs, 218 warehouses (more than 7 million square feet), and highly specialized sales forces are instrumental in helping its clients find best-fit solutions, improve efficiency, and reduce total costs.

This company maintains a complex inventory of more than 425,000 products. To keep revenue flowing for this Fortune 500 company, sales teams around the world need fast and highly reliable access to product information and to the company's inventory and ordering systems. Even brief interruptions can be costly and disruptive, not only for this vendor itself, but also for its clients, many of whom rely on the company to provide time-critical products and materials.

When business teams wanted to deploy mission-critical applications that were not supported in the existing mainframe environment, but yet were essential to the sales, ordering, and invoicing processes, a solution was needed to deliver comparable levels of service for the underlying database. That solution was found in IBM DB2 pureScale feature running on IBM System x3690 X5 servers powered by the Intel Xeon processor E7 family. This simple, powerful database clustering solution not only meets the company's performance and high availability requirements, but also provides unique flexibility and nearly unlimited scalability for future growth. Table 3-1 summarizes the case study for this company.

Table 3-1 Case study summary for a Fortune 500 company

Challenge	Ensure uninterrupted data access for mission-critical applications supporting users in more than 50 countries.
Solution	A simple, powerful, highly-available database cluster - IBM DB2 pureScale running on three IBM System x3690 X5 servers based on the Intel Xeon processor E7 family.
Benefits	Exceptional performance and high availability - Little or no cluster latency and recovery from a failed node in just milliseconds - The leading scalability and reliability of IBM x3690 X5 servers - The exceptional scalable performance and the advanced reliability, availability, and serviceability (RAS) features of the x3690 and Intel Xeon processor E7 family Easy testing, deployment and management - No application changes required - No need to tune the database or balance workloads Flexibility and scalability for the future - Adding servers to the cluster with just two commands - Applications can be added or removed with ease

The company was looking to upgrade two key business applications and add another.

- A new third-party vendor application that calculates taxes across the company's diverse global markets and integrates with its ordering system.
- A custom application, based on third-party code, originates all product and parts information. It also populates the company's online and print catalogs, brochures, and other key sales and marketing publications.

- The third custom application, also based on third-party software, publishes invoices for reprints and research. Unlike the other two, it operates in batch mode, imaging and electronically filing all invoices at the end of each day.

According to the director of Information Technology, “Performance and scalability are important for all three applications. High availability is even more critical. Downtime would be disruptive to our sales, ordering, and invoicing processes. It could also impact our customers in negative ways, so we absolutely need to keep these applications up and running.”

To address these requirements, the company is deploying DB2 pureScale running on System x3690 X5 servers based on the Intel Xeon processor E7 family.

3.1.1 Verifying performance and high availability in the company’s environment

The company was intrigued by the potential value of simple, high-performance database clustering. According to the director of Information Technology, “We have a long-standing and very satisfactory working relationship with IBM and Intel. Our experience with their products and technologies has been very positive, and I was excited to set up a cluster and see what it could do.”

To test the solution they set up a database cluster using a pair of four-socket x3850 X5 servers based on the Intel Xeon processor E7 family. Because the DB2 pureScale clustering solution is completely transparent to applications, there was no need to modify the applications in any way, which not only simplified setup but significantly reduced testing requirements.

The company tested both performance and failover for the clustered solution, and was more than happy with the results. “Performance is very impressive, and so is the resilience of the cluster. If a server goes down, DB2 pureScale recovers in milliseconds, and we have continuous access to the data through the remaining servers. In most cases, this solution will completely shield our customers from a partial outage in our data center.” The solution will also help reduce planned downtime, because it enables rolling maintenance of the servers without bringing down the cluster.

3.1.2 Moving into production

Following the successful tests, the company is moving to deploy IBM DB2 pureScale on three x3690 X5 servers running Linux in its production environment. Each of these two-socket servers is configured with the Intel Xeon processor E7 family. The IBM DB2 and the IBM DB2 pureScale feature components reside on all servers in the cluster to provide three-way redundancy. The company is currently using 12 of the available 16 cores per server, with nine cores per system dedicated to DB2 and three dedicated to the DB2 pureScale Cluster Caching Facility (CF), which manages the cache and locks.

Every server in the cluster will have access to all data and be able to perform any transaction. If a server goes down, no information will be lost or unavailable. The cluster will continue running at two-thirds capacity and workloads will be automatically rebalanced to optimize performance across the remaining two servers. Based on current workloads, the two-server configuration will be sufficient to meet service level agreements. After the failed server is reconnected, workloads will be automatically rebalanced again and the restored cluster will continue to deliver optimized performance.

3.1.3 A foundation for flexible and unlimited growth

According to the director of IT, he chose the IBM clustered database solution first and foremost to meet the stringent high-availability requirements of his mission-critical applications. The scalability and flexibility of the solution are added bonuses. “Our database isn't all that big right now, but the scalability and transparency of the cluster could ultimately be huge for us. It means we can add capacity and integrate new applications at any time, with no coding and without having to retune the database or manually rebalance workloads. If we spin off a new business, or even if we just continue to grow at our historical rate, this gives us a lot of flexibility going forward.”

With the IBM DB2 pureScale feature running on IBM System x3690 X5 servers based on Intel Xeon processors, the company has found a powerful, flexible, and cost-effective database solution for supporting mission-critical applications for its global users. It has also established a foundation for supporting unpredictable growth and change in the future. With today's growing dependence on real-time, data-intensive business processes, that combination could provide substantial value for many other large and growing companies.

4

Why run IBM DB2 pureScale on eX5 servers

With System x and Intel Xeon processors, a powerful and scalable architecture is available to build cost-effective, high-performing, and highly scalable systems. If you add DB2 and the IBM DB2 pureScale feature on top of that architecture, you have a combination of the best industry standard servers and database technology that allows for building systems that provide a high level of performance and availability.

4.1 High-performance clustering for near-linear scalability

A fully configured IBM eX5 server provides a powerful and resilient database engine. Add the IBM DB2 pureScale feature and these servers become building blocks that can be combined without limits to support truly massive data volumes and application workloads with exceptional performance and high availability. More important, adding a server to the cluster is a simple process. Unlike other database clustering solutions, there is no need to repartition the database, manually rebalance workloads, or adapt applications.

The technology that enables this near-linear scalability is not altogether new. It has been at work in IBM mainframes for many years, helping those systems support enormous workloads with the very highest levels of reliability. Continued advancements in industry-standard processor, server, and networking technologies have enabled IBM to bring this technology to its Intel Xeon processor-based System x servers, which is something that was not feasible until these advances.

The shared disk architecture used in DB2 pureScale allows each member of the cluster to access the complete database. There is no partitioning of data used that would limit this. That architecture enables any server in the cluster to process any request, which allows workload balancing across members. It also helps ensure that if one or more servers fail, the remaining servers are not cut off from any portion of the data, thereby enabling normal operations to continue without disruption.

DB2 pureScale and System x servers enable organizations to achieve near-linear scalability, high availability, and complete application transparency on affordable, industry-standard servers.

Every server in a DB2 pureScale cluster also has direct, point-to-point access to a centralized data caching and locking facility through a high-speed network. This centralized cache is the “secret sauce” IBM has ported from its mainframe architecture and is the key to enabling low-latency database performance in a clustered server environment. When an individual server accesses data to perform a transaction, the data page is stored not only in the server’s local memory and cache subsystems but also in the centralized cache. From this repository, data and locking information can be accessed almost instantly by any other server in the cluster.

To make this architecture work, data transfers among the clustered servers and the centralized cache must be very fast, that is, fast enough so that minimal latency is introduced into database transactions. To meet this requirement, DB2 pureScale takes advantage of Remote Direct Memory Access (RDMA). This lightweight memory access protocol allows each member in the cluster to directly access the memory of the centralized cache and vice versa. It helps dramatically reduce the communications overhead required to request and transfer data, and it enables each server to quickly access and deliver data and other critical information to the centralized cache. As a result, the centralized cache reflects the activities of each server in the cluster almost instantaneously.

Every server has high-speed access not only to all data in the shared storage system but also to all cached data and locking information. The high-speed, centralized caching and locking facility virtually eliminates the traditional performance challenges of distributed database solutions. In a traditional cluster, when a server receives a request, it has to search for cached data among other servers in the cluster. Older clustering solutions use distributed caching and locking mechanisms, leading to time-consuming searches and introducing latencies that slow performance. As the cluster grows, the searches become more complex and the latencies increase. By eliminating these complex and time-consuming operations, DB2 pureScale enables fast performance and linear scalability even for very large clusters.

Organizations can now use affordable IBM System x servers to support data volumes and application workloads of any size. As data volumes and application workloads grow, they can simply add additional servers to the cluster to scale capacity. Organizations no longer need to replace entire servers or disrupt operations. In addition, they can build their cluster using the servers that best match their workloads and service-level requirements. Organizations gain the flexibility to optimize capability versus total cost.

4.2 Application transparency helps eliminate hidden costs

DB2 pureScale requires no application changes, database partitioning, or manual workload balancing to maintain optimized performance. An application running on a single, large server can be moved onto a DB2 pureScale cluster quickly and easily. New servers can be added to the cluster at any time, all with no code changes and very little administrative overhead.

This is not the case for a traditional database cluster that relies on distributed caching and locking mechanisms. With these traditional clusters, each server in the cluster must be as close as possible to the data it needs to minimize data access latencies. Satisfying this requirement introduces a great deal of administrative overhead. The database must be partitioned, workloads must be balanced, and applications must be made “cluster-aware” so that they can interface efficiently with this complex environment. These are time-consuming processes that require a significant level of expertise. They are also ongoing requirements because the demands of data partitioning, workload balancing, and application awareness all change as the data, applications, and hardware infrastructure evolve. Changes to applications are particularly resource-intensive, because they typically require extensive testing and validation before they can be implemented in mission-critical production systems.

DB2 pureScale eliminates these requirements. With the software’s centralized caching and locking facility, every server is equally close to all data, so there is no need to partition the database. Any server can perform any transaction. This capability makes it relatively easy to balance workloads across the cluster, and DB2 pureScale performs this function automatically.

Just as important, the centralized caching and locking facility enables DB2 pureScale to be completely transparent to applications. No changes are required to run an application on a DB2 pureScale cluster, and no special tuning is required to optimize performance as the cluster grows. Scaling performance and capacity is quick and simple. A new server can be integrated with just a few commands, and workloads are automatically rebalanced to help optimize performance. Adding and removing servers from a cluster is easy enough that IT can even scale clusters up or down in response to seasonal or monthly workload fluctuations. IBM offers flexible licensing policies that can help businesses take advantage of this flexibility to achieve significant cost savings.

4.3 High availability for uninterrupted operations

DB2 pureScale adds yet another layer of high availability to the silicon- and system-level RAS features that are built into eX5 servers. Because every member of a DB2 pureScale cluster is active and has access to all data, whether cached or stored, a failed server affects only transactions that are currently in process on that particular server. The remaining servers continue operating as usual, and they can instantly take on any new requests that would have gone to the failed system. DB2 pureScale seamlessly maintains normal operations following a server failure by leveraging advanced RAS features. A graphical representation of this concept is illustrated in Figure 4-1.

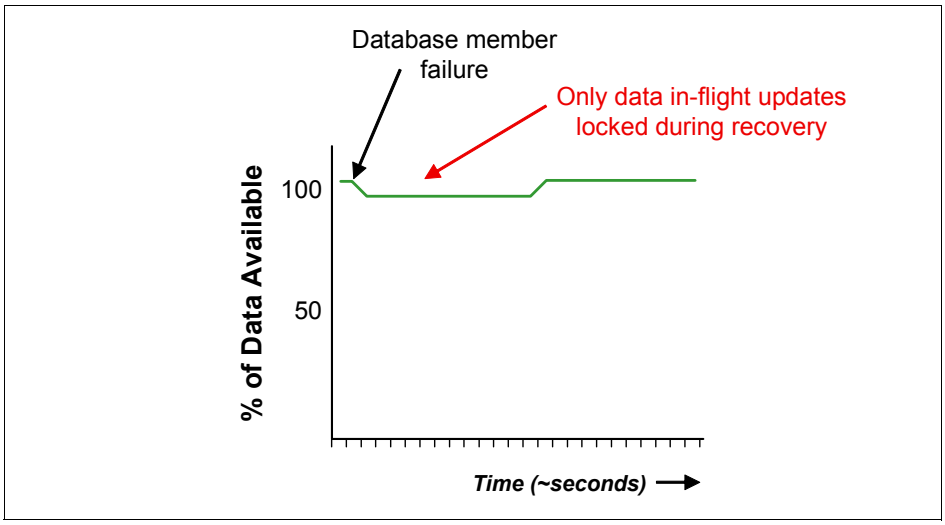


Figure 4-1 Normal operations are maintained during an outage

As in any transactional database, transactions that are interrupted when a server fails must be either recovered or rolled back. A partially completed transaction would not only corrupt the transaction itself but also compromise the integrity of the data. This is another area where the DB2 pureScale centralized caching and locking technology delivers tremendous value. With a centralized information repository, DB2 pureScale software is always aware of the status of every transaction. It also knows which data pages are “dirty,” meaning which pages contain data that has been changed by a transaction but not yet recorded to disk.

Following a server failure, therefore, DB2 pureScale software always has all the information needed to perform these tasks:

- Maintain normal operations without interruption

  A server failure affects only interrupted transactions and dirty data pages directly associated with those transactions. With the distributed caching and locking architecture of a traditional cluster, there is a delay while individual servers share information to determine which data pages are dirty. Until those issues are fully resolved, no data page can be safely used because of the risk of using old data in new transactions. DB2 pureScale software eliminates this delay.

- Recover or roll back interrupted transactions quickly

  Because of its centralized caching and locking facility, DB2 pureScale software knows which transactions need to be recovered or rolled back without accessing any additional information from any other servers in the cluster. Each server also maintains recovery processes that are sitting idle but are ready in the event of a failure. As a result, there is no need for the operating system to create a process and allocate memory to it. Because no information has to be shared and recovery processes can be instantly activated, complete recovery of all interrupted transactions is exceptionally fast. Full recovery typically takes less than 20 seconds, including the time needed to detect the failure. That detection time is not always included in the database recovery specifications of other vendors.

DB2 pureScale not only maintains normal operations without interruption, but also continues to deliver optimized performance by automatically rebalancing workloads across the remaining servers in the cluster. As soon as the failed server is restored, workloads are again automatically rebalanced.

To help ensure there are no single points of failure in a cluster, the DB2 pureScale feature components are always hosted on at least two physical servers. If one server fails, the other

takes over instantly to maintain normal operations.[1] With this approach, a DB2 pureScale cluster provides exceptionally high availability. Traditional single-server database implementations, in contrast, typically require an expensive, high-end hot standby server to provide comparable service-level assurances.

In addition, because DB2 pureScale software automatically rebalances workloads as servers are added or removed from the cluster, taking a server down for physical maintenance is both simple and nondisruptive. Hardware and software remain fully operational, and performance is dynamically optimized based on available hardware resources.

4.4 Automatic workload balancing

One of the great features of DB2 pureScale is the integrated workload balancing capability, which ensures that workloads get evenly distributed among cluster members. All members within the cluster keep load information for the entire cluster so that either the next connection or next transaction can be routed to the least loaded member. The routing occurs automatically and is transparent to the application. Figure 4-2 shows workloads before and after balancing is utilized.

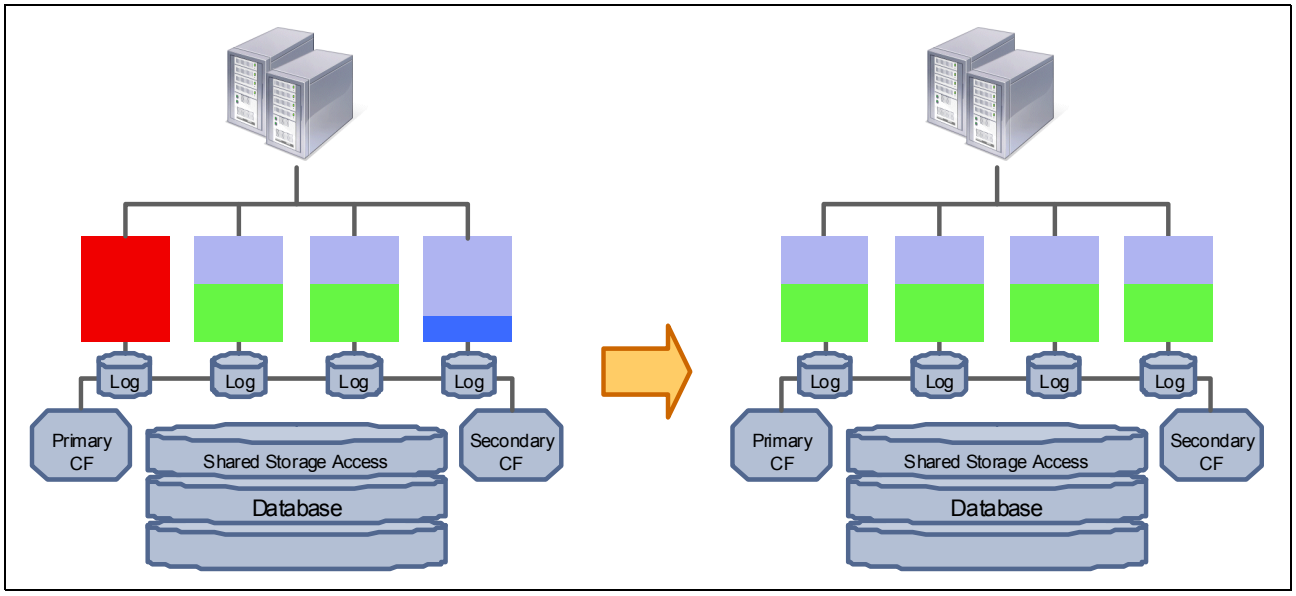


Figure 4-2 Workload balancing

In a failover scenario, the load of the failed member is then evenly distributed to the remaining members. After the failed member comes back into the cluster, requests are then again sent to all cluster members.

[1] A DB2 pureScale cluster can be implemented on as few as two physical servers, in which case both servers are configured such that they are able to perform both cluster management and transaction processing functions as needed. For larger clusters, the DB2 pureScale components can be hosted on dedicated systems.

5

Technology overview

IBM and Intel have been working together for many years to deliver optimized computing solutions that help businesses address these critical challenges more effectively. Increasingly, we have been introducing enterprise-class technologies into industry-standard systems and solutions, so enterprise clients can support heavier workloads and more critical business processes on affordable, Intel Xeon processor-based servers. Today's IBM System x eX5 servers featuring Intel Xeon processors from the E7 family deliver breakthrough performance, scalability, flexibility, and value for enterprise computing. These servers are designed with advanced memory subsystems, proactive management, and enhanced RAS (reliability, availability, serviceability) features which make them the ideal Intel processor-based server platforms for enterprising computing. IBM adds additional value by engineering and tuning these servers in combination with select software stacks to deliver highly optimized platforms for specific workloads.

In combination with IBM DB2 software, IBM eX5 servers based on Intel Xeon processors deliver next-generation data management capabilities on scalable and highly reliable two-socket, four-socket, and eight-socket systems. A single system can have up to eight sockets, 80 cores, 160 threads, and 6 TB of memory per two-node server. Clients can host large, mission-critical databases on these flexible and affordable servers. They can also benefit from IBM software enhancements such as autonomic features, deep compression, and integrated XML to obtain better value from their data at lower total cost.

With IBM DB2 pureScale, an optional feature of IBM DB2, clients can start small and scale incrementally, without disruption, using eight-socket, four-socket, or even two-socket servers as their workloads and data volumes grow.

5.1 DB2 pureScale

Designed for organizations that run online transaction processing (OLTP) applications on distributed systems, IBM DB2 pureScale offers clustering technology that helps deliver high availability and exceptional scalability transparent to applications. DB2 pureScale is based on technology from IBM z/OS®. It is available as an option on IBM DB2 Workgroup Edition, Enterprise Server Edition, and Advanced Enterprise Edition. DB2 pureScale reduces the risk and cost of meeting ever-changing business demands by providing unlimited capacity, continuous availability, and application transparency.

5.1.1 Transparent scalability

Transparent scalability in DB2 pureScale allows for more flexibility when deploying the cluster. Instead of provisioning servers up front for workloads that may occur down the road and binding financial capital, DB2 pureScale lets you instantly add servers when you need them and remove servers if business needs decrease. Along with flexibility, DB2 pureScale offers breakthrough efficiency. Lab tests have shown that the pureScale architecture can deliver 84% scalability for a web commerce workload.

With DB2 pureScale, you do not need to change your application code or perform additional tuning to your database to efficiently run on multiple members. Thanks to a proven, scalable architecture, you can grow your application to meet the most demanding business requirements. You can also run applications written for other database software with few or no changes. DB2 offers native support for commonly used syntax and PL/SQL code, making it easier to move from other databases to DB2.

5.1.2 Continuous availability for your business

DB2 pureScale, through the use of highly reliable clustering technology which includes redundant servers, network, and storage systems, provides continuous availability and eliminates any single point of failure. It provides unlimited access to data with consistent performance. In the event of an unplanned outage, the ability of DB2 pureScale to instantly recover keeps your database running. Maintenance does not require any downtime or interruption to the environment. The system is tolerant to system outages and the user is unaffected by system maintenance.

5.1.3 The DB2 pureScale architecture

Leveraging the experience and know-how IBM gathered from its sysplex technology, DB2 pureScale allows for easy expansion and continuous availability. Through the use of DB2 integrated load balancing and re-routing mechanisms, the cluster appears as one database system to clients, as shown in Figure 5-1. Developers do not need to make application adjustments to run on this clustered environment. This allows for an easier and faster transition to the new architecture.

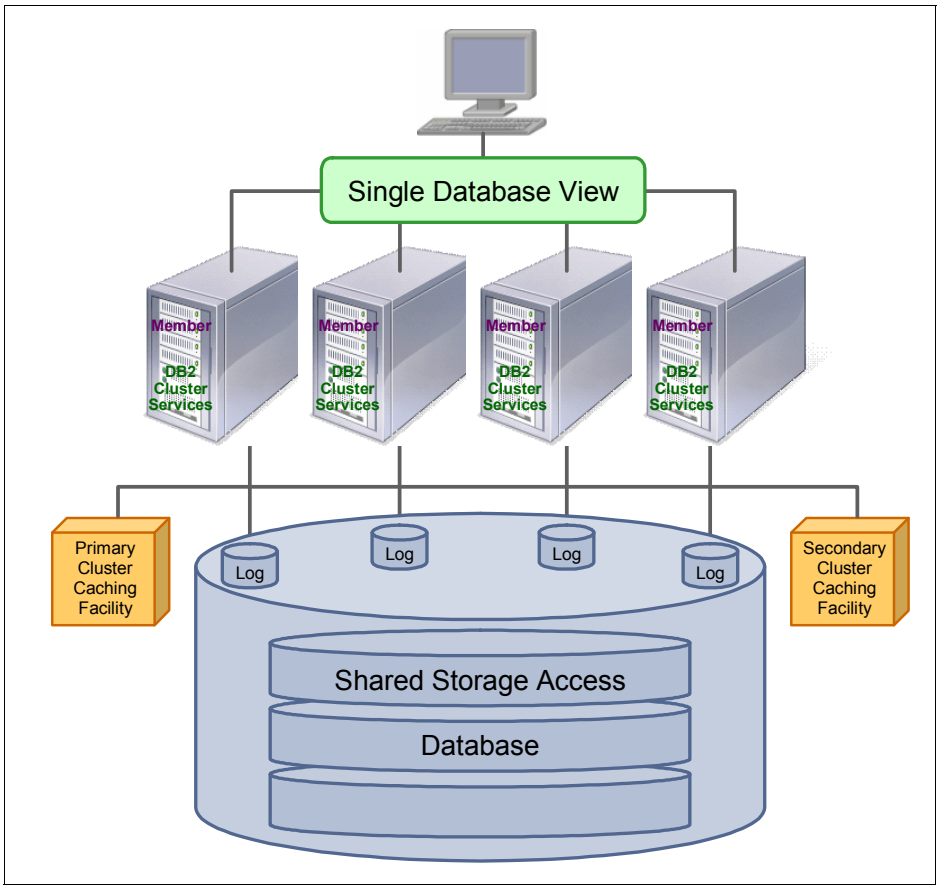


Figure 5-1 DB2 pureScale architecture overview

The DB2 engine runs on each individual member server of the cluster to enable cooperation between each other and provide coherent access to the database from any member. DB2 pureScale provides integrated cluster services for failure detection and recovery automation, and the IBM General Parallel File System (GPFS™) cluster file system for a shared-disk architecture. The cluster caching facility (CF) provides the central instance that holds the Group Buffer Pool (GBP) and Global Lock Manager (GLM). These components ensure a direct way for cluster members to communicate across the cluster and hold locks and buffers to increase database performance.

Communication between cluster members and the CF utilizes either InfiniBand or 10 Gb Ethernet as a low latency, high speed connection. This allows for ultra-fast RDMA access to update information directly in the member servers' memory without CPU utilization.

5.1.4 Handling a system failure

A sudden software or hardware failure can be highly disruptive, even in a system that employs redundant components. DB2 pureScale incorporates several design features to deliver fault tolerance that not only can keep the database instance available, but also minimizes the effect of component failures on the rest of the database system, as illustrated in Figure 5-2.

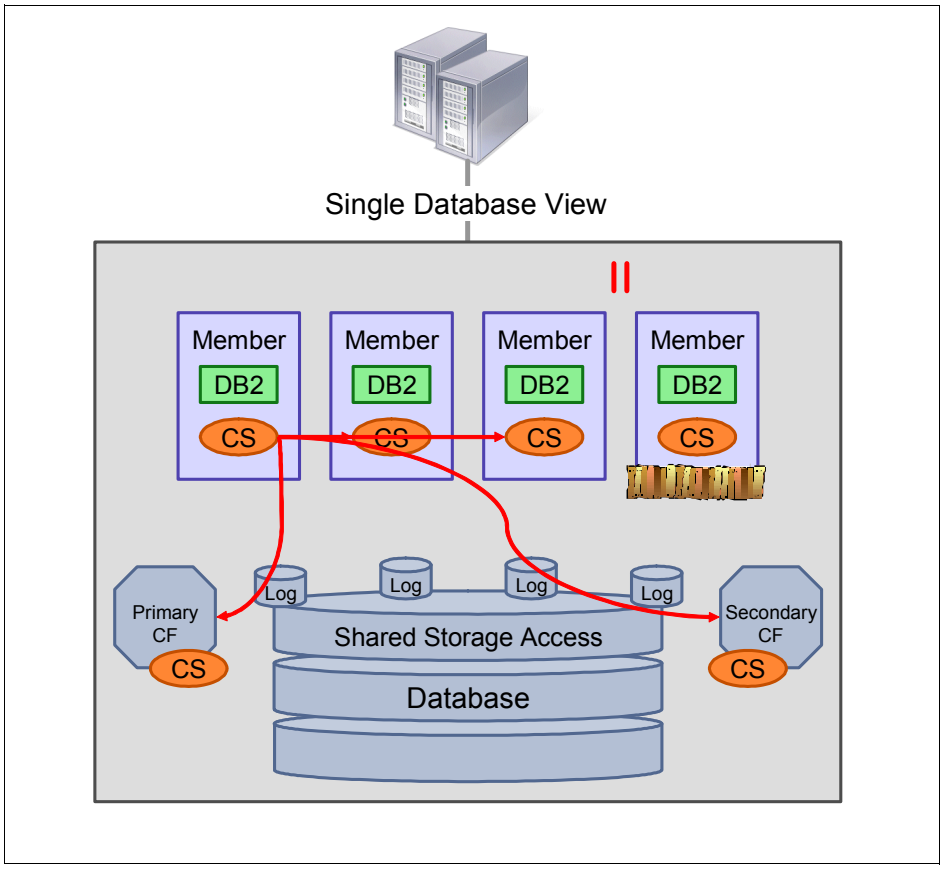


Figure 5-2 DB2 pureScale handling a member failure

Robust heartbeat detection ensures that failed components are identified and isolated rapidly. Recovery from component failures is fully automatic and requires no intervention.

If a member fails while processing database requests, it is immediately fenced off from the rest of the system. During the failure, most of the data on the shared disk storage remains available to active members processing database requests. Only the data that was in-flight on the failed member is temporarily held by a retained lock until DB2 pureScale completes the automated member crash recovery.

After a software failure, the member is restarted on its home host and recovery is performed. The member resumes transaction processing as soon as recovery is complete. After a hardware failure, the member restarts on another host (a process known as “restart light”) so that the data can be recovered. As soon as its home host is available again, the member fails back to that host, restarts, and resumes processing.

After a software or hardware failure on the primary cluster caching facility, a secondary, duplexed cluster caching facility automatically takes over the primary role. This takeover is transparent to applications and causes only minimal delay because of the continuous duplexing of locking and caching information between cluster caching facilities. The instance remains available.

5.1.5 Performing maintenance without downtime

In every IT infrastructure, one thing is certain: the need for updates and fixes. On a single system, this would mean taking down the service that the system is providing, performing the

maintenance task (for example, applying an operating system level patch), and then bringing the system back up and hoping that the patch applied without problems. This can be a daunting task and it usually requires off-hours work. In a cluster this task can be achieved much more easily, as illustrated in Figure 5-3.

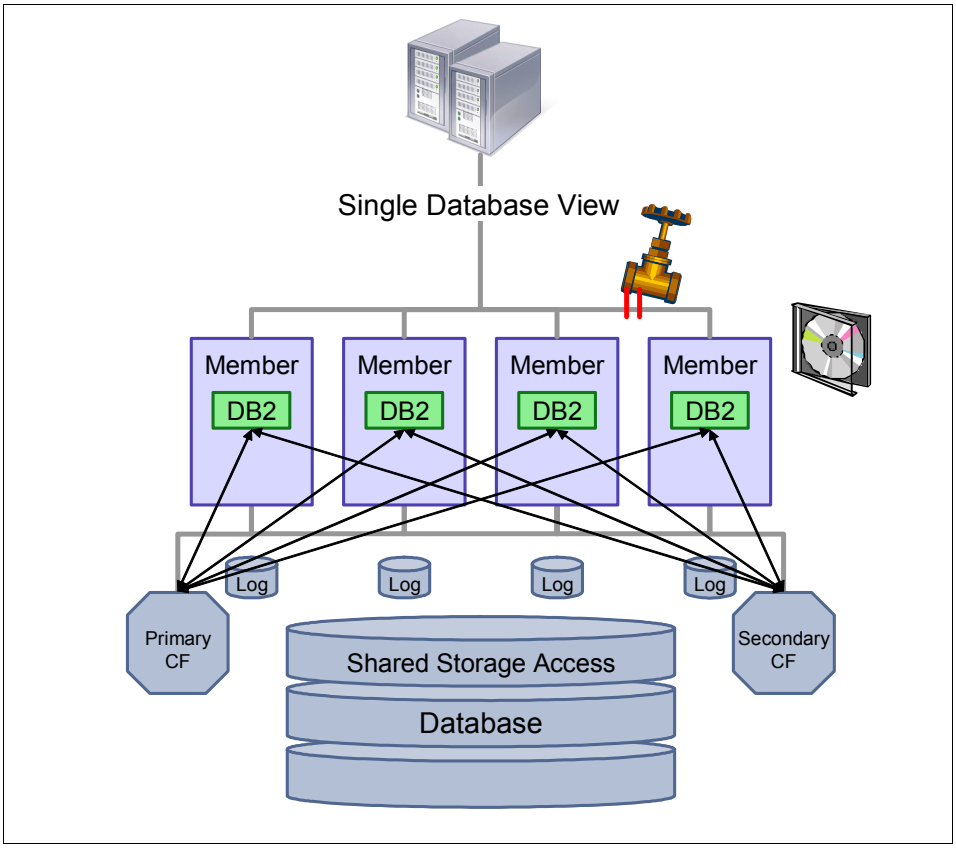


Figure 5-3 Stealth maintenance in a DB2 pureScale cluster

DB2 pureScale allows the system administrator to simply take one system at a time out of the cluster, perform the necessary maintenance tasks, and then rejoin the member to the cluster without any service interruption for the users. The proper way to do this is to “quiesce” the cluster member in question and drain it from all client requests. After this is done and no new requests are handled by this system, the system administrator can take the system offline, perform the maintenance, and after successfully completing the task, bring the machine back up and into the cluster. Users will not notice this maintenance at all because their service was kept available by DB2 pureScale.

5.2 A new generation of powerful, enterprise-class servers

IBM eX5 servers based on the Intel Xeon processor E7 product family represent a major leap forward in supporting mission-critical workloads on industry-standard servers. With on-demand scalability, flexible partitioning, exceptional memory capacity, and advanced reliability features, the eX5 servers provide powerful data center platforms that can be deployed affordably and then scaled easily to meet changing requirements.

The Intel Xeon processor E7 product family plays an essential role in delivering the performance, scalability, and reliability to support demanding applications on large, enterprise-class systems. Servers based on these processors can help IT organizations

scale performance as workloads grow and maintain high availability, while preserving the compelling cost advantages of Intel Xeon processor-based solutions.

5.2.1 The IBM eX5 family

The IBM eX5 product portfolio represents the fifth generation of servers built upon Enterprise X-Architecture. Enterprise X-Architecture is the culmination of bringing IBM technology, often derived from our experience in high-end enterprise servers, to the x86 server market. Now with eX5, IBM scalable systems technology for Intel processor-based servers has been delivered to mid-sized and high-end x86 server systems. These servers can be expanded on demand, and configured by using a building block approach that optimizes system design for your workload requirements.

As a part of the IBM Smarter Planet™ initiative, our Dynamic Infrastructure® charter guides us to provide servers that improve service, reduce cost, and manage risk. These servers scale to more CPU cores, memory, and I/O than previous systems, thereby enabling them to handle greater workloads than the systems they supersede. Power efficiency and machine density are optimized, making them affordable to own and operate.

The ability to modify memory capacity independently of the processors, and the high-speed local storage options such as eXFlash, mean that these system can be highly utilized, yielding the best return from your application investment. eXFlash is a dedicated SSD-based storage array for eX5 systems deployed for high performance database applications and workloads. eXFlash solutions incorporate a unique IBM controller that can support up to 1.6 TB of capacity, and leverage enhanced company I/O technologies to eliminate system bottlenecks. These systems allow your enterprise to grow in processing, I/O, and memory dimensions, so you can provision what you need now, and expand the system to meet future requirements. System redundancy and availability technologies are more advanced than previously available in the x86 systems.

The systems in the eX5 family that support DB2 pureScale are the IBM System x3850 X5 and the x3690 X5. Each system can scale with additional memory by adding an IBM MAX5 memory expansion unit to the server. MAX5 is short for memory expansion for eX5. The x3850 X5 can also be scaled by connecting two together to form a two-node system.

The x3850 X5 is a 4U highly rack-optimized server. The x3690 X5 is a 2U rack-optimized unit that brings additional features and performance to the mid tier. The x3850 X5 and the workload-optimized x3950 X5 are the new flagship servers of the IBM x86 server family. These systems are designed for maximum utilization, reliability, and performance for compute-intensive and memory-intensive workloads.

5.2.2 x3850 X5

The IBM System x3850 X5 server is a unique and compelling high-end modular server with up to four Intel Xeon processors. This server is a modular building-block server. The x3850 X5 can be configured with one or two chassis to build a scale-up server with up to 160 logical threads and 3 TB memory. The x3850 X5 can be configured with a 1U MAX5 memory drawer to build a 4-socket, 1.5 TB memory configuration.

The x3850 X5 server is the fifth generation of the Enterprise X-Architecture, delivering innovation with enhanced reliability and availability features to enable optimal performance for databases, enterprise applications, and virtualized environments. The x3850 X5 is shown in Figure 5-4.

Figure 5-4 IBM System x3850 X5

Features of the x3850 X5 include:

- Increased performance with more processing power and memory than ever before on x86
- Intel Xeon processor E7 family with up to 10 cores and HyperThreading
- Up to 64 DIMM slots per chassis delivering up to 2 TB of high-speed PC3-10600 double data rate (DDR3) memory
- Up to eight sockets and 128 DIMMs with QPI scaling for larger databases, enterprise, and mission-critical workloads
- EXA scaling with dual node systems with MAX5 for performance scaling up to 8 sockets with FlexNode dynamic partitioning and node failover
- Memory reliability and availability with memory ProteXion with Chipkill, memory mirroring, memory sparing, Intel SMI Lane Failover, SMI packet retry, SMI Clock failover
- Low-power cost-effective memory with Advanced Buffer eXecution chip
- High-performing databases and fast time to value for database workloads
- eXFlash for high performance solid state drive storage
- Advanced networking capabilities with Emulex 10 Gb Ethernet Virtual Fabric Adapter
- Integrated Management Module (IMM) for enhanced systems management capabilities
- Serial Attach SCSI (SAS) and SATA HDD storage options, ServerRAID and Host Bus Adapter options, and SSD storage options provide capacity, performance and data protection options for internal storage

5.2.3 x3690 X5

The x3690 X5 delivers innovation with enhanced reliability and availability features to enable optimal performance for databases, enterprise applications, and virtualized environments. It is a 2U, two-socket, scalable system that offers up to four times the memory capacity of

current two-socket servers. It is part of the Intel Xeon processor E7 family, available in sizes of 6-core, 8-core, and 10-core. Most models of the x3690 X5 servers can be combined with the IBM MAX5 memory expansion for up to 2 TB of memory. The x3690 X5 is shown in Figure 5-5.

Figure 5-5 IBM System x3690 X5

The x3690 X5 has the following features:

- Up to two sockets for Intel Xeon processors. Depending on the processor model, processors have six, eight, or ten cores.
- Memory that is implemented using high-speed PC3-10600 and PC3-8500 DDR3 memory technology at up to 1066 MHz bus speed.
- Up to 32 DIMMs in the base system (16 on the system planar, 16 on an optional memory mezzanine), plus an additional 32 DIMMs with optional 1U MAX5 memory expansion unit: a total of 64 DIMM slots.
- Uses Intel QuickPath Interconnect (QPI) technology for processor-to-processor connectivity and Intel Scalable Memory Interconnect (SMI) processor-to-memory connectivity, as follows:
  - Intel QPI link topology at up to 6.4 Gbps with four QPI links per CPU.
  - Intel SMI link topology at up to 6.4 Gbps with four SMI links per CPU.
- Ethernet networking capabilities with a Broadcom 5709 dual Gb Ethernet adapter that is standard on all models.
- Emulex 10 Gb dual-port Ethernet Virtual Fabric adapter that is standard on certain models and optional on all others.
- Power management savings.
- Memory ProteXion with Chipkill, memory mirroring, memory sparing, Intel SMI Lane Failover, SMI packet retry, SMI Clock failover.
- SAS-based internal storage with RAID 0, 1, or 10 to maximize throughput and ease installation; optional RAID-5 and 6.
- Up to 16 hot-swap 2.5-inch SAS or SATA HDDs, up to 8 TB of maximum internal storage. The system includes (as standard) one HDD backplane that can hold four drives; a second and third backplane are optional for an additional 12 drives.
- Up to 24 hot-swap 1.8 inch drives with eXFlash high-IOPS solid-state storage technology for larger, faster databases.
- Five PCIe 2.0 slots.
- Integrated Management Module (IMM) for enhanced systems management capabilities.

- 2U rack-optimized, tool-free chassis.
- Rear access hot-swap redundant power supplies for easy access.

5.3 Other key hardware components

In a highly available and highly scalable environment, the servers that are used are important. Equally important is the storage technology used to store the data and make it accessible in a high-performing and highly reliable fashion to those servers in the cluster, as well as the networking infrastructure to connect cluster members and clients.

5.3.1 Storage infrastructure

Because all cluster members in a pureScale cluster share the same data, technology that allows a shared storage architecture is required. Best practice is to use a SAN-based environment. There is a wide range of storage providers and storage topologies. Usually the storage division within a business mandates certain storage types and topologies. The DB2 pureScale feature is designed to work with a wide range of Fibre Channel attached storage. DB2 pureScale uses GPFS and, in general, all storage supported by GPFS is supported by DB2 pureScale. It will also run on almost any disk subsystem similar to existing, non-DB2 pureScale Feature deployments.

How many disks you need for a DB2 pureScale solution depends largely on the type of workload and on storage requirements. In general, the disk requirements for a given workload are similar to those of the DB2 Enterprise Server Edition. At least two separate disks are required for the cluster: the quorum disk, and the disk that holds the database-related files.

5.3.2 Networking

There are a number of considerations to keep in mind when looking at the communication infrastructure with regard to a high availability cluster. On one side is the internal communication between cluster members. In the case of pureScale we strongly suggest the use of InfiniBand, as explained in the following section.

The other side is the communication of clients with the cluster itself. This is usually an Ethernet-based TCP/IP network, which allows clients to communicate with the cluster as a whole.

InfiniBand

At the core of the DB2 pureScale feature on System x hardware is the InfiniBand (IB) communications infrastructure, which consists of IB Quad Data Rate (QDR) adapters, IB cables, and IB switches. InfiniBand is used as a high speed, low latency interconnect between members that allows for ultra-fast RDMA access to update data directly in memory of the respecting members without wasting CPU cycles. This helps to increase performance of the overall system, because the CPUs of the individual cluster members can be purely used to process requests. As a lower-cost alternative, it is also possible to use 10 Gb Ethernet interconnect within the cluster to connect the members. Although not as fast or low latency as InfiniBand it can be sufficient, depending on the usage scenario.

5.3.3 The Intel Xeon processor E7 family

The x3850 X5 is powered by the highly scalable Intel Xeon processor E7 family. Built to handle the most demanding applications, the Intel Xeon processor E7 family delivers a quantum leap in enterprise computing performance. The Intel Xeon processor E7 family features Intel Advanced Reliability Technology that provides automatic detection and correction of errors, dynamic reassignment of workloads across processors, interconnect error detection and recovery, and individual virtual machine recovery in virtualized environments. It is designed to provide greater processing throughput and more energy efficiency from the following features:

- Greater throughput:
  - 16 DIMMs per processor. Using 32 GB DIMMs, a four-processor x3850 X5 with 64 DIMMs supports 2 TB RAM.
  - Up to 2.4 GHz.
  - 10 cores / 20 threads.
  - Up to 30 MB of last level cache.
- More efficiency:
  - More performance within the same maximum CPU thermal design profile (TDP) as prior versions.
  - Lower partial active and idle power using Intel Intelligent Power Technology.
  - Support for low voltage DIMMs.
  - Reduced power memory buffers.

The enhancements in performance, scalability, energy efficiency, and reliability from the features are summarized in Figure 5-6 and discussed in the following sections.

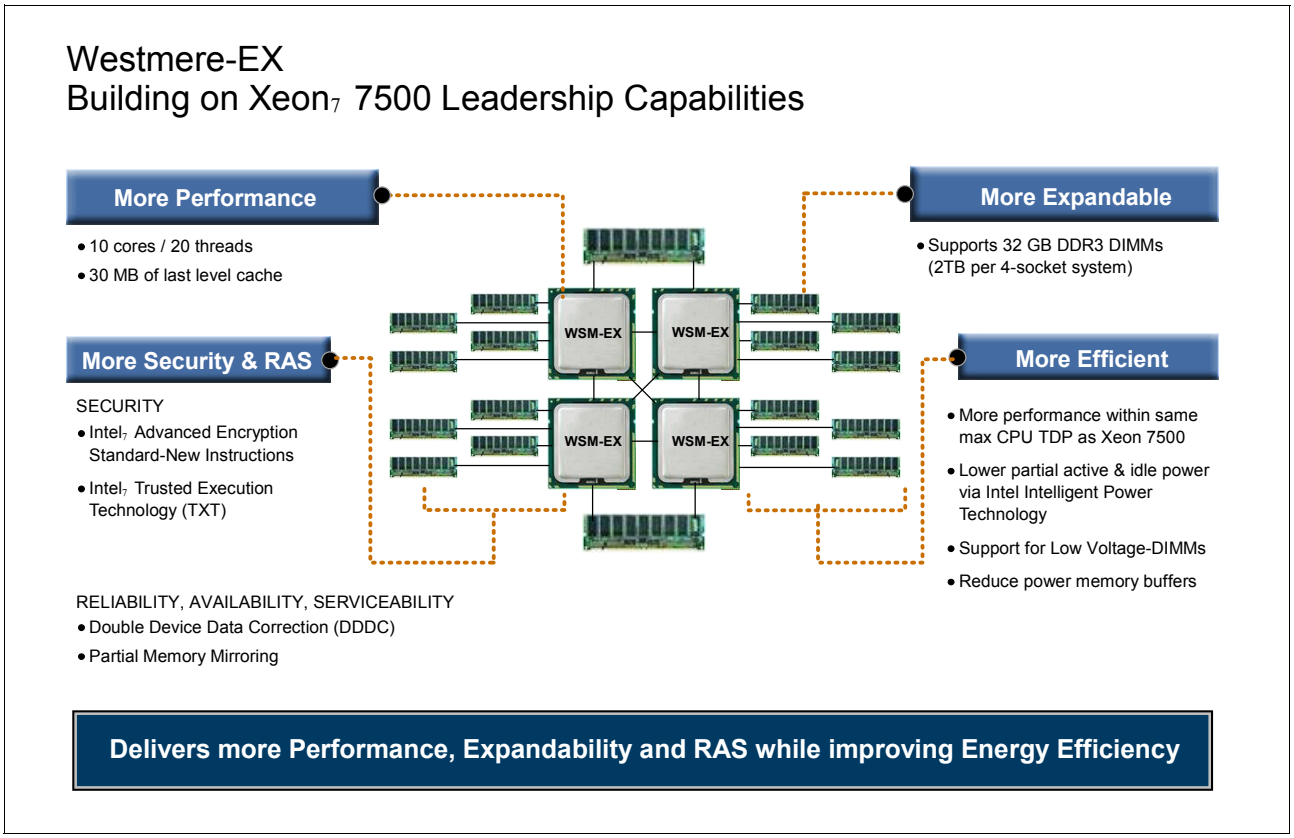


Figure 5-6 The Intel Xeon processor E7 product family

5.3.4 Performance

The Intel Xeon processor E7 family offers outstanding performance for heavy, data-intensive workloads. Built on a 32nm microarchitecture, this product family can provide up to 10 cores, 20 threads, and 30 MB of shared cache per processor. Four advanced, high-bandwidth interconnect links allow multiple processors to be directly connected to each other to increase performance and reduce latency.

These capabilities deliver tangible performance results. Using the Intel Xeon E7 product family, organizations can accelerate database transactions by up to 40% compared with previous-generation four-socket processors.[1] As a result, organizations can generate results faster or process larger data volumes without having to increase the server footprint.

5.3.5 Scalability

The Intel Xeon processor E7 family provides the scalability required for adding or expanding workloads to accommodate new business opportunities. Intel QuickPath Interconnects allow scaling of processors from two to eight sockets. The Quad-Channel Integrated Memory

[1] This claim is based on performance on three industry-standard, common enterprise benchmarks (TPC Benchmark* C, SPECint*_rate_base2006, and SPECfp*_rate_base2006) comparing the best published/submitted results on a 4-socket server equipped with the Intel Xeon E7-4870 processor compared with a 4-socket server using the Intel Xeon processor X7460 and 4-socket server using the Intel Xeon processor X7560, as of March 26, 2010. Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing.

Controller supports up to 16 memory slots per processor socket, enabling administrators to deploy up to 2 TB of memory in a 4-socket system to handle peak demands and leave headroom for database growth. The IBM MAX5 memory expansion drawer adds 32 DIMM slots for up to a total of 3 TB of memory in a 4-socket system. This large-scale memory capacity enables organizations to place an entire database in memory and run queries in real time, speeding the path from data to decision.

We also have node scalability. One x3850 X5 node can scale to two, with the option of EXA scaling using MAX5 for FlexNode dynamic partitioning and failover capability. This gives these systems the ability to dynamically change to dual nodes and back to single node.

5.3.6 Energy efficiency

The x3850 has a power-optimized, energy-smart design for enhanced performance per watt. IBM exclusive eXFlash SSD technology is designed to improve application and database performance and to reduce storage and energy costs.

- New lower power memory DIMMs at 1.35V (versus 1.5V) provide additional energy efficiencies.
- Integrated power gates, which allow one or more cores to be operating while idle cores have power completely shut off.
- Memory power.

Active Energy Manager

IBM Systems Director Active Energy Manager measures, monitors, and manages the energy components of the eX5 systems. Monitoring functions include power trending, thermal trending, PDU support, and support facility providers. Management functions include power capping and power savings mode.

This application helps clients monitor energy consumption to allow better utilization of available energy resources. The application software enables clients to trend actual energy consumption and corresponding thermal loading of IBM Systems running in their environment with their applications. IBM Active Energy Manager is illustrated in Figure 5-7.

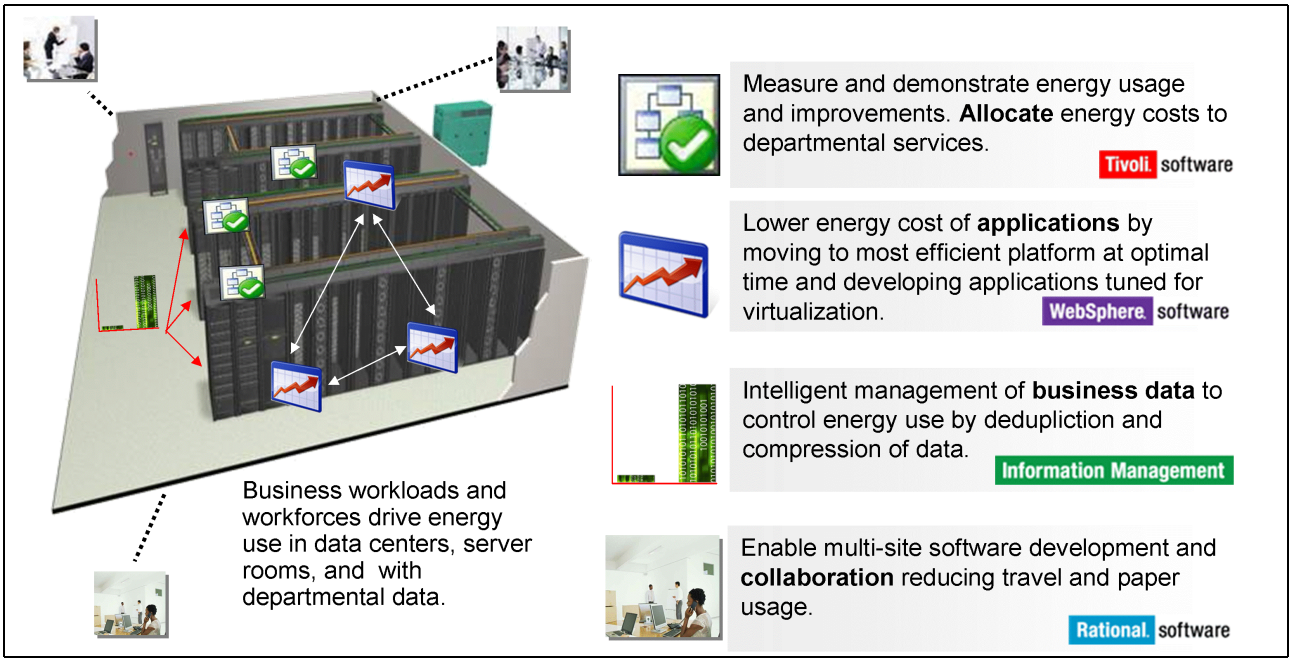


Figure 5-7 IBM Active Energy Manager

Active Energy Manager can help with the following tasks:

- Allocate less power and cooling infrastructure to IBM servers
- Lower the power usage on select IBM servers
- Plan for the future by viewing trends of power usage over time
- Determine power usage for all components of a rack
- Retrieve temperature and power information through wireless sensors
- Collect alerts and events from facility providers related to power and cooling equipment
- Better understand energy usage across your data center by performing the following tasks:
  - Identify energy usage
  - Measure cooling costs accurately
  - Monitor IT

Reliability

The Intel Xeon processor E7 family includes more than 20 features designed to deliver the reliability, availability, and security required for demanding, mission-critical workloads, as explained here:

- Self-healing capabilities enable continued operation even in the event of component failures. Machine Check Architecture-recovery (MCA-r) technology works with the operating system to recover from memory errors without crashes.
- The Intel Xeon processor E7 family adds support for Double Data Device Correction (DDDC), which facilitates recovery from two DRAM device failures.
- Support for Advanced Encryption Standard-New Instructions (AES-NI) enhances security by significantly reducing the performance penalties usually experienced with pervasive encryption.

The Intel Xeon processor E7 family capitalizes on an Intel microarchitecture designed to provide the high-end scalability and availability needed to support large, mission-critical databases. Figure 5-8 illustrates the RAS features of the Intel Xeon processor.

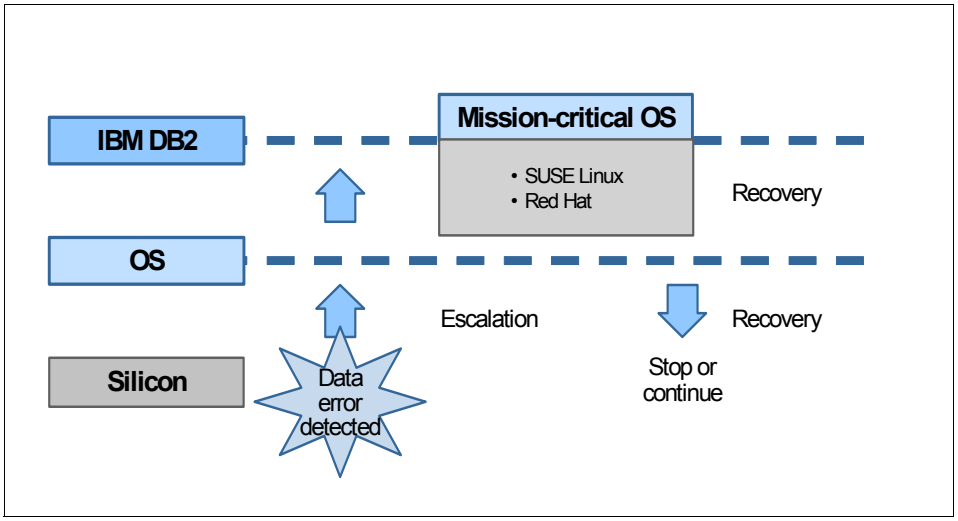


Figure 5-8 RAS features of the Intel Xeon processor

The Intel Xeon processor E7 family includes more than 20 reliability, availability, and serviceability (RAS) features including Machine Check Architecture Recovery, which enables synchronized error handling and automated recovery from many types of previously uncorrectable errors.

IBM eX5 servers offer more than just scalable performance, however. They also deliver major enhancements to improve data integrity and system resilience. IBM supports and extends the silicon-based RAS features of the Intel Xeon processor E7 product family with features such as advanced error management, dynamic node partitioning, and automatic failover to improve utilization and uptime for mission-critical workloads.

Servers based on this architecture also include features such as integrated management modules and advanced light path diagnostics to provide sophisticated monitoring and proactive problem resolution. Redundant power supplies help eliminate a key source of potential downtime. And if issues should arise, easily accessible components help speed time to repair.

These and other features help organizations protect data and meet service level agreements, while refocusing resources on innovation.

The latest version of IBM DB2 software has been tuned to take advantage of Intel Xeon processor enhancements. It also provides valuable software features such as new autonomic management that help to improve performance and free up resources to reduce total cost. New compression capabilities in IBM DB2 software are especially valuable for large database implementations; this feature can reduce storage requirements by as much as 80% to help lower storage costs and improve performance. Because more data can be kept in system memory, there is less need for performance-robbing disk I/O operations. Processors have faster access to data, which improves utilization to help reduce transaction latencies and increase throughput.

For detailed information about IBM eX5 servers, download the IBM eX5 Portfolio Overview at `http://www.redbooks.ibm.com/abstracts/redp4650.html` or visit the IBM website at `http://www.ibm.com/systems/info/x86servers/ex5`.

5.3.7 IBM M3 servers: granular scaling for smaller databases

IBM M3 servers are smaller, two-socket systems based on the Intel Xeon processor E7-2800 product family, providing up to 6 cores and 12 threads per socket. The key advantage of these smaller servers is that they enable more granular addition of resources to a DB2 pureScale cluster.

IBM 3650 M3 servers can be configured with up to two processors from the Intel Xeon processor E7-2800 product family and 192 GB of memory. Although these are powerful and reliable servers, they provide fewer resources than a comparable two-socket IBM eX5 server Additionally, the silicon- and system-level RAS features on M3 servers are not as advanced as those of eX5 servers. IT organizations should take these factors into account when designing their solutions, because more servers may be required per cluster to deliver comparable performance and uptime assurances.

For more information about IBM 3650 M3 servers, visit the IBM website at http://www-03.ibm.com/systems/x/hardware/rack/x3650m3/.

5.4 Configurations to match specific needs

Currently IBM DB2 pureScale software can be used on IBM System x servers, including two-socket, four-socket, or eight-socket eX5 servers based on the Intel Xeon processor E7 product family, as well as two-socket M3 servers. The choice will depend on workloads, expected growth rates, and service level requirements.

5.4.1 Smaller database workloads

In many cases a database may not be very large or have a very high load, but it still needs to be available during business hours. An example of this could be the back-end of a supermarket providing data for daily business operations such as cash registers and inventory tracking. Even for small stores with only a few registers, availability is as important as in a larger store. DB2 pureScale allows you to build systems that meet a need based on the real world applicability of the solution, rather than on a stringent matrix. With as few as two servers and a shared storage device, you can build such a configuration. Figure 5-9 illustrates two IBM System x 3690 eX5 servers connected to an IBM DS3500 storage system.

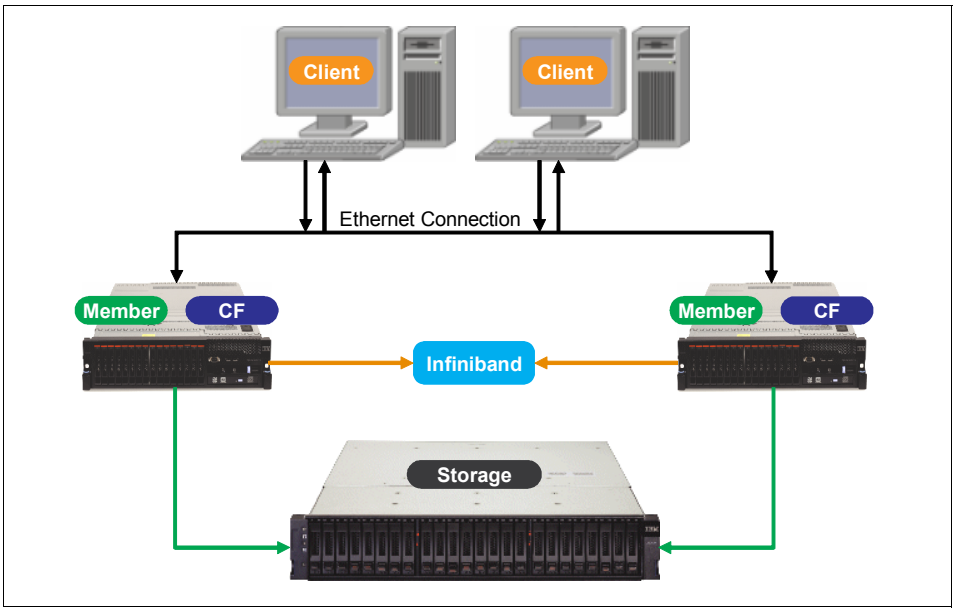


Figure 5-9 Small-size sample DB2 pureScale configuration

5.4.2 Medium-size workloads

In many organizations databases come in various sizes. If a small configuration does not suffice in terms of performance or level of protection against failures, a midsized configuration can be achieved by separating the CF and the member services onto different machines. The machines for CF and member services do not need to be the same type or level of server. In the smaller configuration, only one machine can fail before database availability is compromised.

By separating services, in this case onto four machines, a higher level of resiliency is achieved. The configuration illustrated in Figure 5-10 uses IBM System x 3850 eX5 servers as members and CF machines connected to an IBM DS5020 storage system.

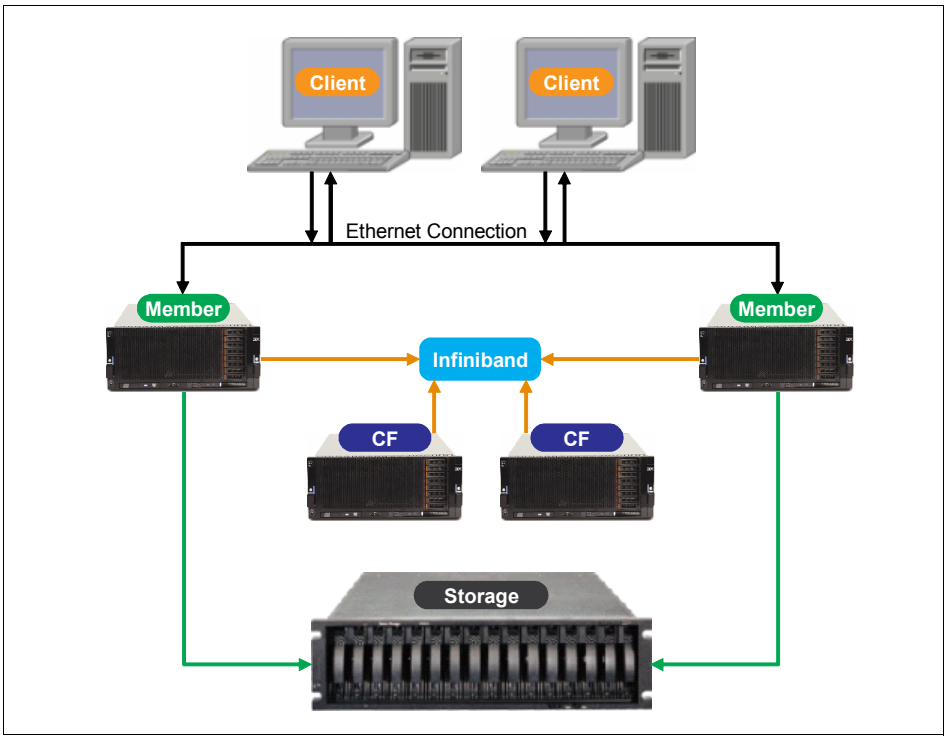


Figure 5-10 Midsize sample DB2 pureScale configuration

5.4.3 Larger workloads and configuration

For large databases and higher workloads, DB2 pureScale allows users to scale out based on their needs. Figure 5-11 shows a five-member cluster. (This is not, however, the largest configuration that can be built with DB2 pureScale.) The advantage is that DB2 pureScale allows users to grow the cluster as they see the need.

This is an example of a higher-end configuration with a higher-end storage infrastructure ready to grow even bigger. As in the medium-sized configuration, IBM System x 3850 eX5 servers are used as members and CFs. It is also possible to mix server models for different functions. The storage system illustrated in Figure 5-11 is the IBM DS8000®.

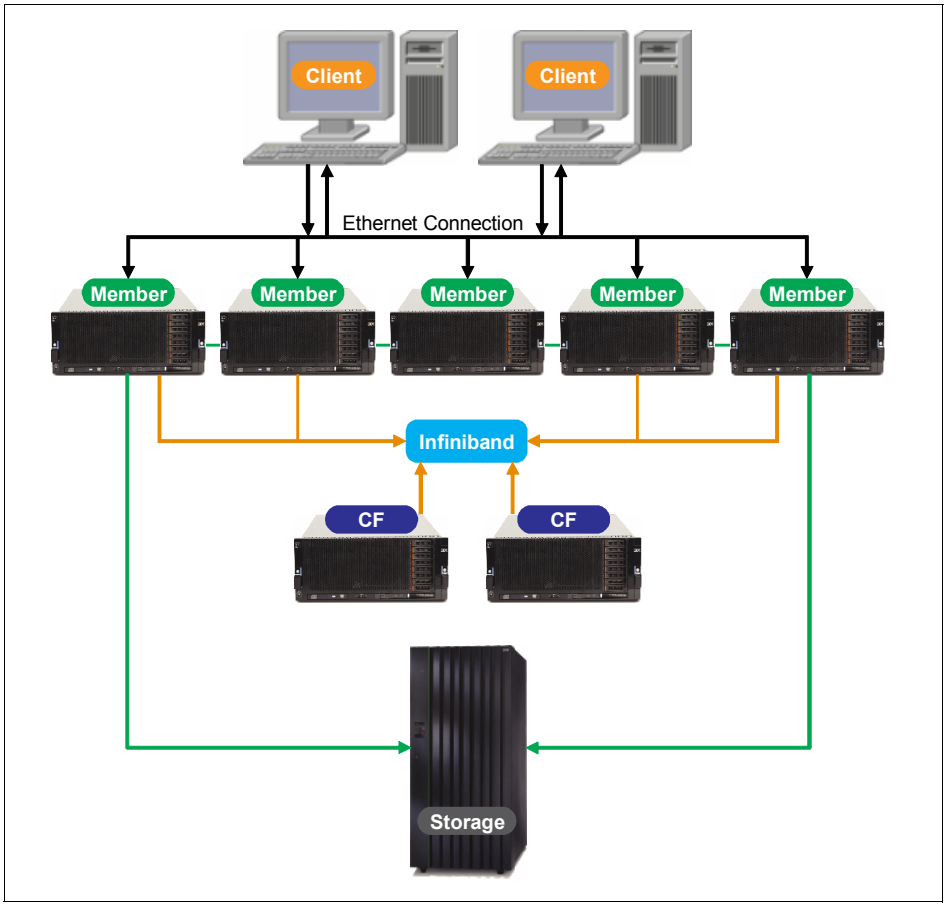


Figure 5-11 Large sample DB2 pureScale configuration

6

The value of IBM DB2 pureScale on IBM eX5 servers

IBM eX5 servers based on the Intel Xeon processor E7 product family deliver a major leap forward in performance, scalability, and mission-critical capability for data-intensive enterprise workloads. These servers offer a powerful, high-value hardware infrastructure for IBM DB2 implementations, with affordable starting points, incremental scalability, and an array of mainframe-inspired features to support demanding requirements for high availability and data integrity.

With the IBM DB2 pureScale feature, clients can combine a large number of these powerful servers into a high-performance cluster that delivers virtually unlimited scalability with high availability and complete application transparency.

By eliminating the challenges that have plagued database clustering solutions from other vendors and supporting affordable Intel Xeon processor-based servers, DB2 pureScale changes the paradigm for deploying and growing large, mission-critical databases. As increasing volumes of data flow into the enterprise, businesses can now scale their infrastructure on demand and without limit, enabling them to store and use their data far more effectively and at a much lower cost.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, 3690 X5, and BladeCenter HX5,* REDP-4650

  http://www.redbooks.ibm.com/abstracts/redp4650.html

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Online resources

These websites are also relevant as further information sources:

- IBM website for DB2 pureScale on LUW

  http://www-01.ibm.com/software/data/db2/linux-unix-windows/editions-features-purescale.html
- IBM DB2 pureScale Feature for Enterprise Server Edition

  http://publib.boulder.ibm.com/infocenter/db2luw/v9r8/index.jsp?topic=/com.ibm.db2.luw.container.doc/doc/c0056942.html
- IBM website for System x eX5

  http://www-03.ibm.com/systems/info/x86servers/ex5/

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Highly Available and Scalable Systems with IBM eX5 and DB2 pureScale

Redpaper™

Introduction to the new IBM System eX5 servers

IBM DB2 pureScale feature for scalability and availability

Based on the Intel Xeon processor E7 family

The pressures related to managing transactional databases are increasing rapidly. Business growth, the drive to consolidate databases, and the need to deploy new, data-intensive technologies are fostering the massive expansion of data volumes and application workloads. At the same time, the move toward real-time computing requires faster and more reliable data access, especially when databases are used to drive client-facing applications. Businesses need simpler and more cost-effective strategies for expanding their database environment.

IBM and Intel provide an answer to this challenge with the IBM DB2 pureScale feature and the latest generation of IBM System x eX5 servers based on the Intel Xeon E7 processor family. The combined solution enables clients to scale mission-critical, performance-sensitive databases simply, using affordable, industry-standard servers.

This IBM Redpaper publication describes an IBM solution containing high availability and scalability for mission-critical databases on System x. The audience includes executives and other decision-makers, consultants, and architects.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

REDP-4742-00