White Paper

Affordable, Scalable, Reliable OLTP in a Cloud and Big Data World: IBM Db2 pureScale

Sponsored by: IBM

Carl W. Olofson
August 2017

IN THIS WHITE PAPER

This white paper discusses the concept of shared data scale-out clusters, as well as how they deliver continuous availability and why they are important for delivering scalable transaction processing support. It also contrasts this approach, taken in a relational database context, with clustering approaches employed by NoSQL databases and Hadoop applications, showing the importance of the relational model and a persistent governing schema in large volume transaction processing.

It goes on to discuss the specific advantages offered by IBM's Db2 pureScale, which is designed to deliver the power of server scale-out architecture, enabling enterprises to affordably develop and manage transactional databases that can meet the requirements of a cloud-based world with rapid transaction data growth, in an affordable manner.

SITUATION OVERVIEW

Data Growth Versus IT Budget Constraints

IT managers are being told constantly that they must do more with less. They must handle more databases, more data, and growing volumes of transactions and analytic queries with the same or smaller budgets. The advent of Big Data, fueled by the need to retain and analyze streaming data, machine-generated data, social media data, and data from the Internet of things (IoT), has added to this demand.

With growing data volumes has come an increased demand for throughput and for support by the same database of more, different, and more complex database operations. In addition, demand for increased intelligence in the execution of transactions has led to the need to mix update transactions with complex queries. All this requires more processing power.

Enabling databases to grow sufficiently to handle this demand for processing power involves using scaling technologies in the relational database management systems (RDBMSs). If a database requires large amounts of shared memory to handle multiple complex queries with a substantial compute component to the workload, a "scale up" approach may be employed, seating a database on a server with a large number of sockets and large amounts of memory. If the objective is the flexible distribution of transaction operations or simpler queries where large numbers of concurrent users are involved, a "scale out" approach, involving clustering, is generally preferred; the kind of clustering depends on the workload involved.

Scale-Out Database Clustering

Clustering is an approach to database deployment that involves using multiple servers (called "nodes") that operate as a single database server. Adding more nodes increases the compute power of the cluster overall. The database server may be supporting one or more databases, but in either case, the nodes cooperate as a single system.

Database Clustering Versus Sharding

RDBMS-managed clustering is distinct from "sharding," where a database is broken up into pieces (called "shards") and managed on multiple servers. With sharding, each server, whether running as part of an OS-managed cluster or as multiple servers on a network, acts as a single database server; these servers coordinate some of their operations but cannot, for instance, ensure referential integrity across servers or implement ACID (atomicity, consistency, isolation, and durability) for the collection of shard servers as a unit. Also, because each shard operates as a database server, the application must be deployed in a way that accounts for the organization of the shards.

Clustering Approaches

There are two basic kinds of clustering, and each addresses a distinct workload. These are commonly called "shared nothing" and "shared storage" (we used to call this latter approach "shared disk," but since storage systems today can manage persistent storage in memory, in solid state memory, in solid state drives, or on spinning disks, the term "shared storage" is now more precise).

The "shared nothing" approach involves dividing the database data into partitions, and assigning each partition to a node in the cluster. Each such node manages the data for its partition exclusively. Unlike sharding, clustering involves such close cooperation between the nodes that they behave externally as a single database system, distributing queries across the nodes, executing queries in parallel, gathering up the results, and returning a single result set to the application. Because the shared nothing approach facilitates parallel query processing (it is also called massively parallel processing, or MPP), it is considered the better approach for analytic databases, such as data warehouses, that have few concurrent update threads and many complex, often long-running query threads. (Note that an example of this, in the case of IBM, would be the Database Partition Feature [DPF] in the Db2 Advanced Server Edition and the Advanced Workgroup Edition. This configuration of Db2 is also available in an integrated system format called PureData System for Operational Analytics.)

"Shared storage" involves distributing the compute work among cluster nodes that share the data in a common storage resource (usually deployed as network-attached storage or a storage area network), with the work being allocated across the nodes using a load balancing scheme. This approach is considered better when dealing with databases that support many update threads and where update activity is more or less constant and requests are fairly compact, such as transactional databases. As a result, technology that involves concurrent update control (including lock management) and shared buffers is critical to the smooth performance and scalability of the database.

Clustering for High Availability

Clustering is commonly employed for high availability. This is often done by pairing an "active" database server with a "passive" or "standby" server that takes over if the active database server goes down. Such an approach is unnecessary when deploying with a shared nothing or shared storage cluster because it is generally considered a fundamental requirement that the RDBMS, in addition to supporting scalability, must implement each node as a failover node for another so that such clustering approaches have high availability built in. Nonetheless, it is important for anyone considering a clustered database approach to examine how the RDBMS vendor handles failover within the cluster.

Requirements of a Shared Storage Clustering RDBMS

In summary, the following are requirements that the user should expect for any shared storage clustering RDBMS:

- The database appears as a single unit to the application, requiring no special application configuration or specialized SQL.
- Failover is built in so that the database administrator (DBA) may select the failover data replication strategy, and the RDBMS will do the rest.
- Adding or removing nodes in the cluster should be a simple operation with little administrative overhead and should be transparent to the application.
- The RDBMS should have features that enable largely autonomous operation, including automated node workload balancing, tuning, failover, and repair.

Db2 pureScale

IBM's RDBMS, Db2, offers a configuration called pureScale, which can operate on both x86 and Power systems. pureScale enables Db2 to operate as a shared storage cluster, interacting with the application as a single database system. pureScale leverages architectural concepts from its mainframe cousin, along with technologies from IBM's Systems and Technology Group (STG) and Tivoli, to provide a clustered database environment that is largely autonomic and easy to administer. With Version 11 of Db2, all editions now include the ability to deploy pureScale, not just the advanced editions.

How Db2 pureScale Clustering Works

The nodes in the pureScale cluster act cooperatively to optimize their execution of SQL transactions on the database. Their database buffers and locks are synchronized by a master cache manager, called the Cluster Caching Facility, using advanced technology to ensure continuously consistent database images on all the nodes. This can be done using either a purely software approach over conventional 10 Gigabit Ethernet or (on systems that support it, such as Power and x86 systems) a hardware-assisted approach using RDMA (remote direct memory access) that works over InfiniBand or 10 Gigabit Ethernet networks. With Version 11, Db2 pureScale now also supports 40Gb RDMA over Converged Ethernet (RoCE) or InfiniBand networks. A pureScale cluster also provides transparent failover in the event that one node crashes. Even the cache manager has a standby to which it can fail over if necessary. The system uses parallel logs to prevent a bottleneck on log I/O. The result is a collection of database server nodes that act as a unit, with smooth, continuous operation that is transparent to the application and thus requires no specialized SQL.

How Db2 pureScale Delivers Scale-Out Benefits

pureScale offers Db2 users a pay-as-you-grow approach to systems management. Rather than allocating system resources now for the anticipated load of a few years from now, the user may provision the database with the appropriate processing power for today's workload and then, as that workload demand expands, simply add more nodes to the cluster. Nodes may be added without interrupting database operation. pureScale features automatic workload balancing across the nodes and, according to customers, near linear scalability.

Because Db2 pureScale is simply a configuration of Db2, existing users can adopt it without converting data or reorganizing the database in any way. Existing Db2 applications will run without alteration and require no changes in order to take advantage of the scalability of pureScale. There are no partitions to define and no special indexes to create.

Db2 delivers scale-out benefits in the following ways:

- Transparent, cooperative operation requiring no special code or tuning changes, fully compatible with any standard configuration of Db2
- Automatic failover, again without requiring special application code (This capability also means that the nodes can be maintained, in most cases, through rolling database system maintenance, eliminating the need for scheduled maintenance downtime.)
- Simple node insertion procedure for scaling up the cluster, requiring a minimum of administrative involvement and no cluster outage
- Self-management features that ensure self-tuning, load balancing, and even operation without special tuning or monitoring

With Version 11, Db2 pureScale now offers more flexible licensing terms. A new configuration that is available involves the use of an "admin node" in the cluster. An admin node is a Db2 node that is functionally limited to administrative functions and incurs a much lower license fee as a result. In addition, Version 11 features Db2 Direct licensing, which is a monthly subscription model that can be applied to Db2 in the datacenter or in the cloud.

Deploying a shared storage database cluster can be a daunting administrative task, and maintenance can be tricky. With Version 11, IBM has reduced the complexity at both levels so that a reasonably competent DBA should have no trouble setting up and maintaining a fairly large and robust Db2 pureScale cluster.

Cloud Computing and Continuous Availability

Whether in the datacenter or in cloud service environments, enterprises are moving to cloud computing as the foundational architecture for IT. This means making server and storage resources fungible in order to achieve the elastic scalability that the cloud demands. Where the public cloud is concerned, it often also means ensuring continuous availability. Scale-out configurations like Db2 pureScale are logical approaches to satisfying these requirements from a relational database management perspective, and Version 11 provides direct support for cloud deployment and cloud-based licensing. Specifically, pureScale supports IBM's SoftLayer cloud platform.

Real-Time Decisioning and Columnar IMDB

Increasingly, enterprises must depend on intelligence in the application, and on data that is immediately available to the user, to ensure the kind of smart transaction processing that produces ongoing competitiveness. Blending transaction processing with complex analytics performed on live transactional data is a key element of real-time decisioning. A number of RDBMS vendors, including IBM, are developing increasingly sophisticated ways of doing this without compromising performance. Currently, Db2 offers the means of leveraging scalable OLTP and memory-optimized columnar complex analytic query processing with BLU Acceleration to address this requirement. In the latest release of Db2, BLU Shadow Tables, which are BLU Acceleration tables that mirror the row-wise transactional tables in the database, are maintained for high-speed analytics close to the transactions.

FUTURE OUTLOOK

Big Data Integration

Increasingly, enterprises are learning to integrate Big Data into their operations. This means not only ingesting data from streaming sources, such as machine-generated data from sensors or feeds from data services, and using that data for analytics but also applying that data to transactional workloads. It also results in more complex transactions and puts greater strain on the database server to process such transactions while delivering uniform performance that meets the service-level agreements that are established for each workload. Flexible scalability such as that delivered by shared storage clustering is key to providing such even performance.

While Db2 has native support for JSON documents, fully integrating data from sources like Hadoop into analytical-transaction workloads can still be a challenge. IBM is addressing this issue with the federation capabilities built into Db2, which can federate a range of data sources, including, in addition to relational databases, Hadoop. Db2 federation works with Hadoop either through Apache Hive or IBM's Big SQL.

CHALLENGES/OPPORTUNITIES

RDBMS faces a number of challenges, including meeting the demands of an integrated cloud environment, more efficiently dealing with unstructured data and data from Big Data sources, such as Hadoop and NoSQL databases, and managing mixed workload environments. All vendors in this space are evolving their technologies rapidly, and the pressure is on IBM to keep pace with or outpace the pack.

CONCLUSION

The combination of online processing pressures, new data sources, and more complex applications has posed key problems for fixed-resource RDBMSs, particularly where transaction processing is involved. The response needs to include a scale-out architecture that can allow DBAs to expand (or shrink) the processing capacity of the database server to meet the needs of the applications it serves. Such a scale-out architecture is best expressed using a shared storage cluster that operates, from the point of view of the application, as a single database server.

The need for such a capability is made all the more acute in relation to deployment in the cloud, where virtualized resources may be readily assigned. Advances of Db2 in support of a cloud environment are examples of flexible resource usage that is sorely needed in the cloud.

In reviewing such requirements, and options for addressing them, the enterprise should consider the following:

- Any vendor's solution should offer near linear scalability with little administrative effort.
- The clustering approach should appear transparent to applications and users.
- Failover and recovery operations should be automatic, requiring no special user intervention.
- The cluster should deliver scalable performance without special conditioning of the database or the SQL that accesses it.

- The cluster should be deployable in the datacenter or in the cloud, and when in the cloud, it should fully leverage the multitenancy, virtualization, and elastic scalability aspects of the public cloud.
- For existing Db2 users, the solution should work without requiring data conversion or application changes.

Taking all the above into account, the enterprise should compare IBM Db2 pureScale with any other offering being considered for implementing a clustered, scalable database configuration.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com