High performance analytical solutions Solutions

Cisco tests quasardb database on UCS servers

The agility of companies depends on their capacity to analyse and anticipate. An appropriate and rapid analysis of large amounts of data is a real challenge to which analytical solutions must respond. In-memory type solutions are an asset for achieving the necessary level of performance in this area. In order to extend their field of use, quasardb offers a distributed database capable of feeding these in-memory systems very quickly. Analytical solutions, therefore, dispose of all the data required to make their queries in real time. It is in this context that quarsardb and Cisco tested the behaviour of a quasardb cluster with a configuration of Cisco UCS servers.

Cisco

Cisco addresses the needs of Big Data and analytics by offering:
- Network infrastructures (Nexus, ACI) and servers (UCS)
- Data integration software (Cisco Data Virtualization)
- A massively parallel processing system (Parstream)
- A range of Cisco Connected analytical solutions
- Partnerships with software producers

Quasardb

Quasardb is a company that offers a distributed key-value database.

- **Strengths:**
  Performance and expansion capability

- **Use cases:**
  Trading, risk calculation, cyber security, dynamic pricing, churn management, fraud detection, etc.

The Big Data challenge: Analyse more and more data

More and more fields use are demanding analytical solutions to process large volumes of historical data and recent data in real time. Such requirements justify the use of 'in-memory" databases. As the database resides in the server memory, one is freed from any performance problems caused by data transfers. The reality is more restrictive. In many cases, the volume of data of a large company exceeds the technological limit of an in-memory database. In order for analytical solutions to be able to exploit the power of in-memory technology without limiting their scope of use, it is necessary to be able to quickly load the data relevant to the analyses into the server memory as needed in a timely fashion.

Extend the power of in-memory technology to all the company data

Quasardb is a new generation of distributed database designed to maintain high performance in extreme conditions of use in terms data volume and the number of queries. The Quasardb design allows the data required by analysis tools to be quickly loaded into the memory of a server.

All software can interface to Quasardb thanks to an API designed for this purpose and the use of data in its native format eliminates any need for conversion.

Quarsardb identified several principles in order to achieve these results:

- **Reduce the transfer to only the data useful to the analytical tools.**
  A tag function refines the selection of data, then Quarsardb limits itself to the transfer of the registrations concerned.
- **Ensure performance linearity and availability.**
  The distributed architecture of Quarsardb uses the Chord algorithm to optimise performance. Consistent performance is maintained when adding a node to a cluster or adding another cluster.
- **Ensure the frugality of operations.** The database code works as closely as possible with the hardware in order to limit it to essential use only.

Quasardb tests its performance on Cisco UCS servers in a 10-Gigabit environment.

High performance can only achieved if there is a perfect match between the software architecture and the hardware infrastructure.

1- The tests were designed to ensure that:

- Each Quasardb cluster node uses the bandwidth (10 GB per second in the test environment) in an optimal manner and that load distribution is done in a homogeneous manner on all nodes.
- Performance increases linearly with the addition of nodes.

2- Configuration of the test:

- **Analytical simulation server:**
  Cisco UCS C240M4 **-** 2 CPU Intel E5-2697 v3 - 256GB – 2 x 40 GB VIC 1385
- **Switch**:
  Cisco Nexus 9372 PQ (2 40GB connections to the simulation server – 8 10GB connections to two fabric interconnects)
- **Fabric interconnect**:
  2 x Cisco UCS 6248 equipped with the UCS manager software
- **1 cluster of 7 quasardb nodes:**
  Each node consists of a Cisco UCS C240M4 server - 2 Intel E5-2697 v3 CPUs - 256GB - 2 x 10GB VIC - 24 1TB DD. The servers run on Linux.

The linearity of performance verified by the tests

The tests were carried out successively on 1 node, 2 nodes, 4 nodes and 7 nodes by making 120,000 transactions per node. The simulations reveal the following:

- **An optimal use of the bandwidth** with a transfer of 8GB per second for one node due to the optimisation of protocols by quarsardb. The test did not make use of compression mechanisms. Their use during production will further optimise the performances obtained. One could even increase the global bandwidth by using, for example, the latest Cisco fabric interconnect (UCS 6332) in order to dispose of a full 40GB architecture.

- **Linearity in load increase**. A progressive increase in the number of nodes in a cluster did not result in any degradation of performance. The throughput increased proportionately to the number of nodes to achieve 56GB per second with 7 nodes. The quarsardb distributed database automatically redistributes the data to all nodes to ensure consistent access to each node. This avoids a disparity in performance between nodes. With the 7-node cluster, 1TB of data can be loaded in 125 seconds.
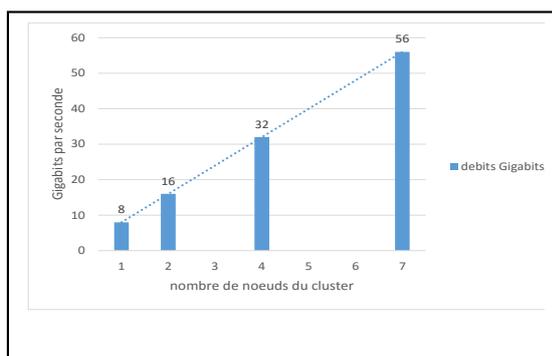
- **Performance maintained in the event of a node failure.** The database redistributes the data evenly to all active nodes. Another effect of this approach, the persistence of data, is ensured by the Quasardb database and avoids having to manage complex replication mechanisms at the in-memory level.

Quasardb and Cisco complementarity for going beyond limits

In a context in which it is necessary to take the continuing growth of the volume of data into account, it is important to ensure that the architecture of the database and the network and server infrastructure are well aligned in order to support future developments with the same level of performance, and to do so without having to make massive new investments. The objective is to maintain a native hardware infrastructure and software development perspective.

- **The quarsardb architecture** provides unlimited scalability of capacity and does so by adding nodes or clusters without performance degradation.

- **The Cisco infrastructure** adapts itself natively to the throughput requirements of the Quarsardb database by adding network links or by increasing their speed and moving from 10Gbits to 40Gbits or to 100Gbits. The non-blocking infrastructure and the low latency ensure the optimisation of the network links.

Powerful and expanding

Companies facing the challenges of performance and expansion now have the opportunity to meet them today. To this end, each component of the solution must be designed with this in mind and none of them must penalise the functioning of the whole. This is what the test has demonstrated.