InfiniDB®
Hardware Sizing Guide

Release: 4.6
Document Version: 4.6-1

InfiniDB Hardware Sizing Guide
July 2014

InfiniDB may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from InfiniDB, the furnishing of this document does not give you any license to these patents, trademarks copyrights, or other intellectual property. The information in this document is subject to change without notice. InfiniDB shall not be liable for any damages resulting from technical errors or omissions which may be present in this document, or from use of this document

Contents

1 Overview....4
2 Data Size and Data Volumes 5
3 Workload and Active Data Set 6
4 Memory Requirements 6
5 Linear Scale Expectations....7

1 Overview

InfiniDB delivers tremendous I/O efficiency versus traditional row-based DBMS through a number of techniques. This I/O efficiency allows for queries to scale linearly with additional processing power without encountering traditional bottlenecks. Adding cores can be accomplished either by scaling up and leveraging today's multi-core systems, or by scaling out and leveraging commodity hardware.

Sizing discussions for InfiniDB involve understanding both the business requirements related to the size of the data, as well as the target workload to be processed.

Data size requirements will dictate the drive capacity requirements and workload processing requirements will dictate the CPU and memory requirements.

Note: *For details on the InfiniDB commands necessary to complete some of the analysis below make sure to consult the InfiniDB Administrator's Guide. You should also consult the InfiniDB Performance Tuning Guide during this analysis. Also note that some concepts discussed here apply only to InfiniDB v4.5 and later.*

2 Data Size and Data Volumes

Configuration decisions based on data size are used to determine the capacity of the storage system, and do not necessarily determine the required processing power, or the number of servers. InfiniDB supports multiple data volumes, and the number of data volumes can be scaled independently of the number of servers.

Data Volumes should be configured to have between 4 and 8 disks configured in a RAID 10 configuration, delivering 350 to 450 MB/s read throughput per volume. Data volumes can be configured to support between 1 TB and 3 TB of storage each. RAID configurations that achieve redundancy via parity (RAID 5, 6, etc.) are also suitable and can be used if the load rate requirements are minimal and the load operations don't overlap with significant query workload. A good starting point is to allocate 2 data volumes per I/O channel.

Care should be taken to avoid having to place import datasets on the same disks and/or I/O channels as the database volumes. This can significantly affect import performance. Configurations using a separate UM and the standard import process normally do not have to worry about this since there are no database volumes on the UM in such a configuration.

For example, assuming a 5:1 compression ratio and 2TB data volumes, storage for 40TB of raw data would require 5-6 data volumes to allow for some future growth. Adding additional storage to InfiniDB is quite easy, but it is not an online operation, so to minimize the number of maintenance windows required some capacity planning is required.

Configuration on Amazon Web Services should maximize the number of Elastic Block Store volumes to maximize overall I/O throughput.

Careful consideration of the filesystem format is required. There are benefits and drawbacks to each:

- ext2: a very reliable, very efficient format. Volume size is limited to 4TB.
- ext3: also very reliable. The journal, while providing some benefit to recovery, incurs an automatic double-write penalty. Limited to 4TB. There is little reason to use ext3 if ext4 is available.
- ext4: very reliable, still has the journal double-write penalty, but is virtually unlimited in its volume size.

Regardless of the filesystem format chosen, InfiniDB creates relatively few, very large files. Thus the number of inodes allocated can be significantly reduced from the default by as much as a factor of 256.

3 Workload and Active Data Set

Workload requirements for a database system are a function of Query Size, Query Concurrency, and the Active Data Set.

- Query Size: The typical or average number of blocks touched to support the query. This can be measured directly using calgetstats(),and through analysis of query syntax, schema, and table cardinalities.
- Query Concurrency: The number of active, concurrent queries typically executing. Note that this may be different from the number of active users.
- Active Data Set: If the vast majority of queries are focused on the same subset of data, then the Active Data Set is smaller than the Data Size. For example, if the overall data set covers 36 months of data, but 95% of the queries are focused on the most recent month for Analysis, then the Active Data Set would be 1/36th the size of the total data.

The output of the Workload calculations is a metric in terms of blocks touched per second. Typical processing capabilities for modern CPU can be between 10,000 and 50,000 block operations per second per core. These metrics can be used to calculate a projected ability to process workload. Newer CPUs will deliver higher processing capabilities, and processing data from cache will also result in higher block processing rates.

4 Memory Requirements

Memory is critical to both query flexibility and performance with InfiniDB. Memory is used to avoid I/O from storage as well as providing space for memory based join and aggregation operations. Minimum specification for evaluation of InfiniDB is 32GB, but the typical sizing for full-scale testing/production use is higher. A significant number of customers go into production with between 96GB and 512GB of memory.

Analysis with dimension tables under 1 million rows and minimal need for self-join of fact tables can be handled with 32GB of memory. For more complex analysis, large dimension tables, and self-join of fact tables a larger memory configuration is recommended. The largest configuration in production is approximately 200GB allocated for `TotalUmMemory`.

The recommended amount of memory for the data buffer cache to avoid I/O from storage is 20% of the Active Data Set. For example, with 36 months of history and 36 TB of raw data size, the Active Data Set would be approximately 1 TB, and a data buffer cache of approximately 200GB would be appropriate.

5 Linear Scale Expectations

InfiniDB offers scalable disk resources and scalable processing that allows for linear scaling of queries. In general, doubling the processing power and I/O throughput will deliver twice the query performance, cutting query time in half.

Scaling expectations:

- Doubling the processing and I/O while doubling the Active Data Set will deliver consistent query performance.
- Doubling the processing and I/O with static data will deliver near linear scalability. Some better than linear scaling can occur when additional memory allows for caching more blocks and reducing a disk bottleneck.
- Query times will increase in direct proportion as the number of concurrent queries.

Note that doubling the total data size, while the Active Data Set remains the same will not generally impact performance. For example if the Active Data Set is 1 month, but the historical data grows from 36 months of history to 72 months of history the query performance will remain consistent.