INFINIDAT

Infinidat InfiniBox® and SAS® Analytics

SAS is one of the leading analytic platforms for AI, machine learning and business intelligence. The resources required to support the SAS platform, such as CPU, memory and I/O, are very predictable and isolated in such a way that each can be optimized to improve performance, reliability, and availability for all phases of SAS workloads.

Optimizing SAS Performance

Of all resources that affect performance, the most influential in any SAS environment is I/O. You can add memory or CPUs to a server farm, but if the I/O sub-system is not capable of keeping up with the workflow, the other resources could be under-utilized waiting for I/O to complete. It is important when choosing an enterprise storage platform to look for a provider that will perform well supporting every phase of the SAS processing cycle.

ACCELERATE
The Analytics Life Cycle

DATA

DEPLOYMENT   DISCOVERY

The SAS Processing Cycle

Data Access: Getting the data from a source to the SAS environment requires large block writes to the SAS servers.

ETL: In this phase, data is then mapped, transformed and loaded into SAS data structures.

Usage: This is where the user community begins to analyze, synthesize and distill the data to provide the answers SAS was designed to deliver.

Each of these phases have characteristics that have vastly different I/O requirements, and each utilizes the storage system in different ways.

I/O Testing Tool: To give the end user a sense for what a particular I/O system can support, SAS provides an I/O testing tool that simulates these three phases. The SAS support page describes the tool, how to use it, and how to interpret the results.

Throughput Rates: Performance of your SAS Analytics deployment is determined by the throughput rates for each of the 2 SAS file systems: **/saswork** and **/sasdata**. This document presents test results for running these two file systems on an Infinidat enterprise storage platform. At a minimum, SAS recommends an average throughput rate of 125MB/s per physical core in the system.

Test 1: Results for Single Server Running SAS

This test was conducted using a single, large bare-metal Linux system with 2 volumes configured: /sasdata and /saswork. The server had 20 cores and 128GB of RAM. There were 4 x 16GB Fiber Channel HBA's connecting to the InfiniBox F63xx. The rhel_iotest.sh script was run against each file system.

SAS I/O Testing Tool Results:

Read throughput rate: 308.12 megabytes/second per physical core

Write throughput rate: 256.91 megabytes/second per physical core

InfiniMetrics®, the Infinidat performance monitoring tool, shows the I/O performance through the entire SAS Processing cycle.

Performance Notes:

▶ The green line shows peak read throughput rate is over 6GB/s, with very low latency

▶ The red line shows peak write throughput rate is 5.3GB/s

The graph shows 3 phases of processing: Data copy, ETL, and read data. Data copy and ETL are write intensive, with data copy using 1MB write block size. The Reads are small block sequential reads.

INFINIDAT

Test 2: Results For 3 Servers Running SAS

The second test was to show how much SAS workload can be placed on a single InfiniBox F63xx. Each server is a large bare-metal Linux system, as described in Test 1, used to drive the SAS workload, run simultaneously on all 3 servers. The results are as follows:

Server 1: SAS I/O Testing Tool Results

READ THROUGHPUT RATE: 262.23 megabytes/second per physical core

WRITE THROUGHPUT RATE: 225.86 megabytes/second per physical core

Server 2: SAS I/O Testing Tool Results

READ THROUGHPUT RATE: 293.67 megabytes/second per physical core

WRITE THROUGHPUT RATE: 230.52 megabytes/second per physical core

Server 3: SAS I/O Testing Tool Results:

READ THROUGHPUT RATE: 252.74 megabytes/second per physical core

WRITE THROUGHPUT RATE: 254.76 megabytes/second per physical core

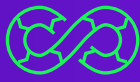InfiniMetrics shows the I/O performance through the entire SAS Processing cycle.

Performance Notes:

▶ Read throughput shows peak read throughput rate is over 16GB/s

▶ Write throughput exceeded 14.8GB/s
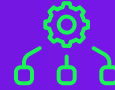
INFINIDAT

**High
Performance**   **High Availability
and Reliability**   **Multi-Petabyte
Scale**   **Simple and Powerful
Management**   **Low TCO and
High ROI**

InfiniBox Advantages for SAS

SAS requires high-performance, high-resiliency storage at scale. That system should also be easy to manage so that administrators can focus on the needs of the business instead of the needs of the infrastructure. InfiniBox excels at meeting these requirements while also delivering a low total cost of ownership (TCO) with a rapid return on investment (ROI). InfiniBox SSA, Infinidat's 100% solid-state enterprise storage system takes performance to another level for workloads that require consistent, microsecond latency for every I/O.

InfiniBox is built on 5 fundamental principles that are a perfect match for SAS solutions:

1. High Performance

Groundbreaking, real-world application performance mainly contributed by architecture innovations, neural cache design, and real-time analytics. A large telecommunication company saw query times go from 40 minutes to 15 seconds after they implemented InfiniBox, over using direct attached storage.

2. High Availability and Reliability

SLA-based outcomes, guarantees 100% availability, performance, and cyber storage recoverability.

3. Multi-Petabyte Scale

Storage densities, capacities, and efficiency that maximize floor space and storage platform consolidation up to 17PB of effective capacity in a 42U rack, to lower TCO and create a greener data center.

4. Simple and Powerful Management

InfiniOps enables ease-of-use and ease of automation that significantly increases the petabyte (PB) per full-time-employee (FTE) management ratio.

5. Low Total Cost of Ownership (TCO) and High Return on Investment (ROI)

A global investment management company stated using InfiniBox was 50% less expensive than their previous enterprise array vendor. Infinidat offers a TCO of 30-50% less on a price per TB basis compared to the competition. According to the IDC Business Value of Infinidat InfiniBox Snapshot, InfiniBox delivers an average ROI of 165% over 5 years with a 48% reduction of operating costs and a payback period of 11 months.*

* Business Value Snapshot, sponsored by Infinidat, Business Value of Infinidat InfiniBox, December 2022

INFINIDAT

Infinidat InfiniBox for SAS

The InfiniBox enterprise storage platform delivers groundbreaking, real-world application performance, 100% availability guaranteed, and capacity density at multi-petabyte scale for a multitude of mixed workloads, synchronous and asynchronous replication, data-at-rest encryption, and InfiniSafe® cyber storage recovery capabilities, assure maximum data security and reliability.

With InfiniBox, SAS implementations are simpler to implement, operate, and protect while reducing costs and delivering the results that help companies discover, innovate, and drive their success.

Key Features

▶ Groundbreaking real-world application performance mainly contributed by architecture innovations, neural cache design, and real-time analytics

▶ SLA-based outcomes, guarantees 100% availability, performance, and cyber storage recoverability

▶ Storage densities, capacities, and efficiency that maximize floor space and storage platform consolidation to lower TCO and create a greener data center

▶ InfiniOps enables ease-of-use that significantly reduces the petabyte (PB) per full-time-employee (FTE) management ratio

▶ Economy of scale that dramatically lowers TCO and increases ROI for large scale, data intensive IT operations resulting in best time-to-value

InfiniBox for SAS in Action

Configuring Infinidat storage for SAS is extremely simple. SAS requires 2 main file systems, /saswork and /sasdata. Deploying on InfiniBox requires a single volume for each. There is no reason to create multiple volumes for each which complicates the configuration. When a single volume is created on InfiniBox, using InfiniRaid, the data objects are spread across every spindle in the frame.

InfiniBox F63xx   **Multiple SAS Nodes Scale-out configuration**

Mount points for each

▶ /saswork

▶ /sasdata

Conclusion

Analytics platforms such as SAS require rapid access to critical datasets with predictable performance regardless of the workload. Since the data is vital to your business, it should also be highly reliable, easy to deploy, and scale with your business growth. The testing with SAS publicly available test tools shows that a single InfiniBox can deliver performance advantages whether running in a single SAS node or in a scale-out SAS model. SAS running on InfiniBox enables customers to focus on the needs of their business instead of worrying about the cost and complexity of the storage infrastructure by providing:

▶ **Faster Ingest**–increase the ingest speed and overall amount of data ingested WHILE supporting multiple user report.

▶ **Highest Availability**–capture all critical data with ultra-low latency and near-instant write acknowledgement. 100% availability guaranteed means your Splunk/InfiniBox solution is always open for business.

▶ **Speed/Capacity/Lowest TCO**–with groundbreaking speed AND petabyte-scale capacity, InfiniBox addresses the speed and capacity in a single solution, at the lowest total cost of ownership (TCO), providing best time-to-value.

INFINIDAT