RED HAT STORAGE

GAIN A COMPETITIVE ADVANTAGE WITH FINANCIAL ANALYTICS

Access useful insights and cost-effective tick data storage with Red Hat and Penguin Computing

TECHNOLOGY OVERVIEW

INTRODUCTION

A significant portion of financial data analytics is performed by sifting through large volumes of tick data. However, enterprises struggle to retain as much data as they would like due to the cost and capacity constraints of their underlying storage systems. Penguin Computing and Red Hat Storage have built a cost-effective and scalable solution to enable financial institutions to make more insightful decisions based on tick data.

THE PROBLEM

Financial institutions, traders, and researchers closely track a number of asset class types on an ongoing basis. These may include securities, equities, futures, interest rates, foreign exchange, exotic instruments, cash indices, or full order book data.

Any upward or downward movement – represented by a tick – over a period of time is extremely valuable to run analytics, build historical simulations, develop trading and market-making strategies, and build transaction-cost models.

Tick data could contain price, volume and many other dimensions – such as bid or ask prices, bid or ask sizes, quote time, trade time, or exchange information – for each point of granularity.

In an uptick trade, the current transaction occurs at a price higher than the previous transaction. In a downtick trade, the current transaction occurs at a lower price than the previous transaction. Consequently, a zero tick refers to a trade where the transaction price has not changed.

Historical tick data may be used with a variety of analytical applications to improve profitability, including:

- Low-frequency calculations.
- Risk modeling and profit and loss reporting.
- Mark-to-market, stress tests, and Monte Carlo simulations.

At the end of each trading day, all the tick data collected through a number of data sources is stored as distinct Historical Database Partitioned Format (hdpf) files for that day. Each file is typically written as a large sequential stream of blocks.

The granularity of data stored within each hdpf file depends on how closely the enterprise would like to track the security and what kind of analytics must be run on the data. The higher the resolution of tick data collected, the larger the dataset size and, hence, the amount of storage capacity required.

Storage capacity for tick data could very quickly grow from hundreds of gigabytes to terabytes or even petabytes. Efficient tick data storage is, therefore, a critical business challenge for most financial institutions.

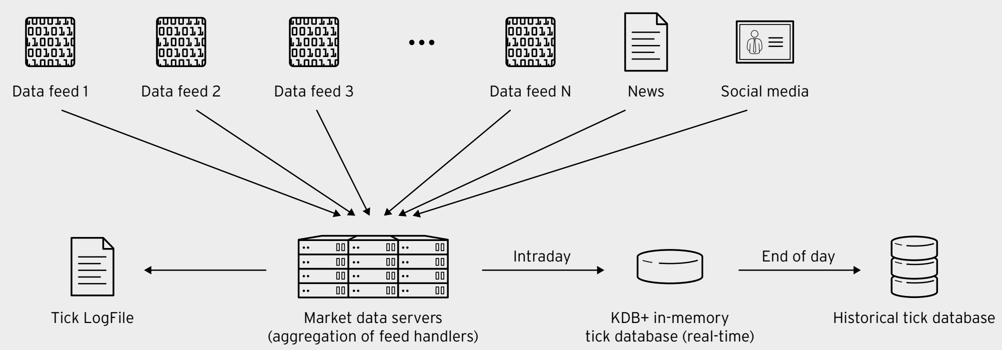

Figure 1. Solution Architecture

WITH TICK DATA, CUSTOMERS CAN EXPERIENCE ISSUES SUCH AS:

- **Quality of financial analytics.** The limitations of traditional storage restrict financial institutions' ability to retain as much data with as much granularity as they would like.
- **Compliance with standards.** Regulatory standards in the financial industry mandate specific data retention durations and policies.
- **Cost of storage.** The exorbitant cost of storage appliances and rigidity in capacity expansion further stress financial enterprises on tick data storage.
- **Performance.** Below-par performance in reading or writing tick data can lead to unacceptable latencies in making buy/sell decisions based on the analysis of data.

A PARTNERED APPROACH

The joint solution between Red Hat Gluster Storage and Penguin Computing is built on the core premise that software-defined storage deployed on industry-standard hardware can deliver the performance, scale, and flexibility needed by financial institutions struggling with tick data storage challenges.

The production system is built using Red Hat Gluster Storage across dedicated Penguin servers, with redundancies built in. The off-site standby is a lightweight version of the on-site production unit and offers additional redundancy for failover and backup.

WITH THE COMBINED RED HAT AND PENGUIN SOLUTION, CUSTOMERS CAN BENEFIT FROM:

- **Deeper analytics.** The ability to look further back in time by storing more data with greater granularity helps enterprises find more useful and actionable insights.
- **Lower cost.** Customers switching from storage appliances like EMC Isilon to Red Hat Gluster Storage for tick data storage report cost savings on solution acquisition and ongoing maintenance.
- **Better performance.** Customers switching from monolithic, traditional storage solutions to Red Hat Gluster Storage find a significant gain in read/write performance.

- **Lower risk**. The off-site standby is asynchronously geo-replicated with the production site, allowing for greater resiliency and high availability, in addition to offering compliance with regulatory standards.
- **Tried and tested configurations**. Experts at Penguin Computing and Red Hat Storage have identified the optimal combination of software and hardware based on the specific access patterns of the use case.

PERFORMANCE RESULTS

By profiling a similar workload on an EMC Isilon system we find that Red Hat Gluster Storage can sustain the rigorous workload pattern of a large number of directory creations as tick data is stored in distributed time series hdpf files across many directories. In fact, Red Hat Gluster Storage outperforms a 16-node EMC Isilon cluster with only four nodes of Red Hat Gluster Storage, which directly translates into cost savings for the customer.

The following graphs featuring Red Hat benchmark tests offer additional details on the workload and comparative performance.

Throughput ranges from 1.3 GB/s (maximum) to 0.2 GB/s (minimum).

The average sustained throughput is ~0.6 GB/s.

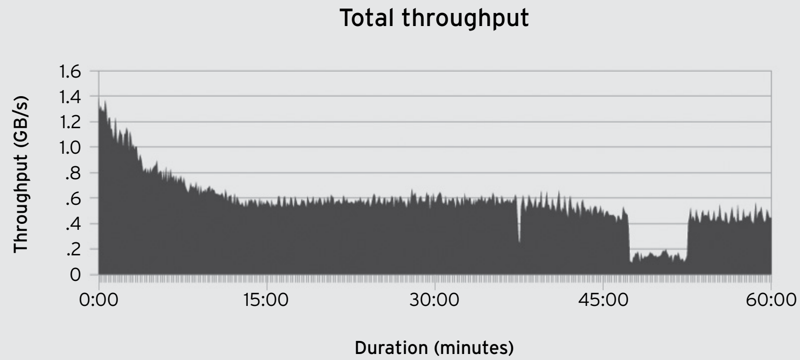

Figure 2. Total cluster throughput as measured over a period of 1 hour (6-second intervals)

The input/output (I/O) workload is predominantly read-oriented.

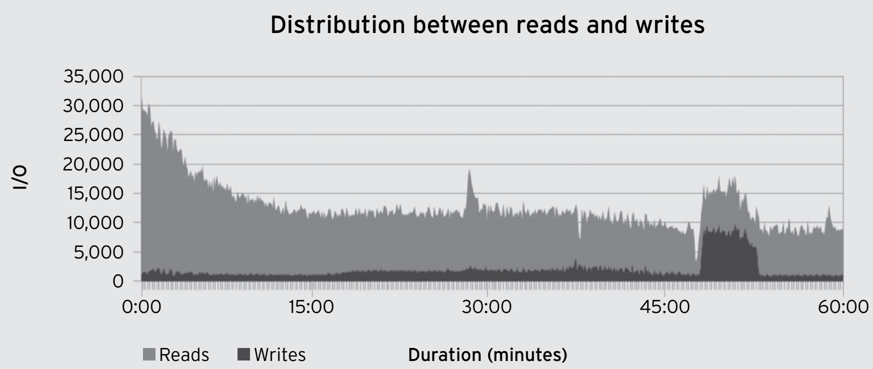

Figure 3. Distribution between reads and writes

Writes comprise ~15% of the workload.

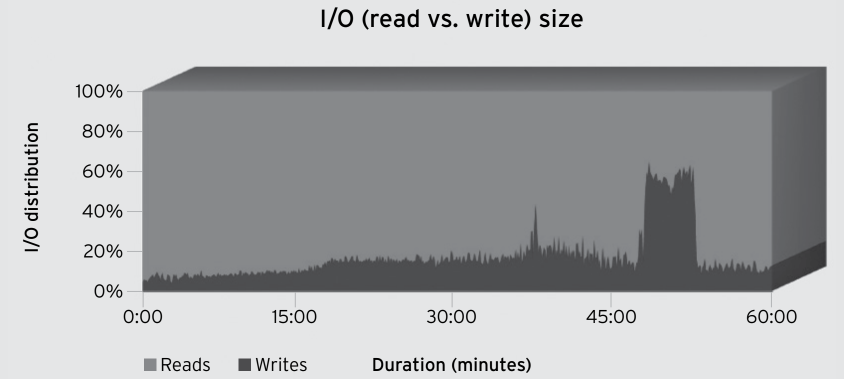

Figure 4. I/O (read vs. write) size

The average read I/O size is ~51 KB.

The average write I/O size is ~18 KB.

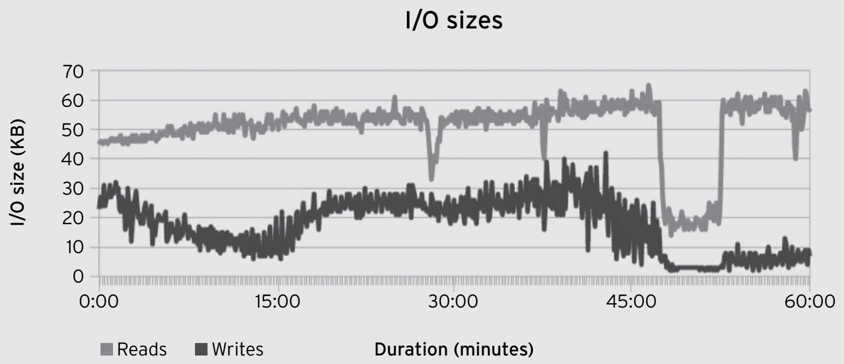

Figure 5. I/O (read vs. write) size

Most write I/O sizes are between 20-29 KB (24 KB average).

Most read sizes are between 50-59 KB (54 KB average).

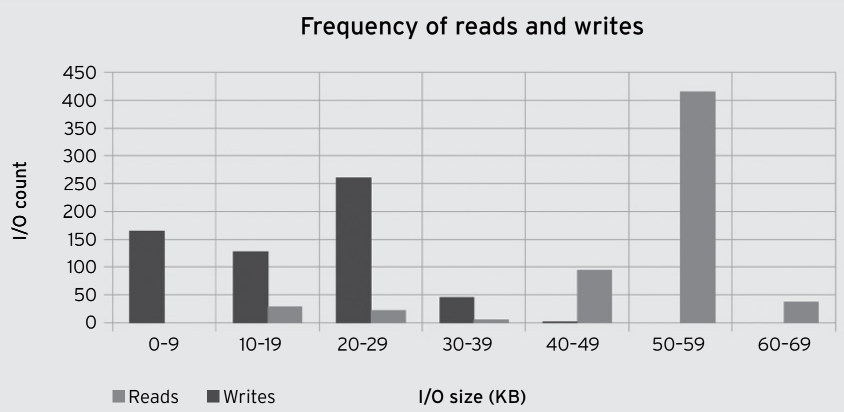

Figure 6. Frequency of reads and writes

TECHNICAL DETAILS

COMMON HARDWARE ACROSS PRODUCTION AND OFF-SITE

- Seagate 512e sector size 6TB, 7200RPM, HDD spinning drives
- Dual Intel Xeon E5-2650 V3 CPU (with 10 Cores, 2.3 GHz), per Red Hat's test system
- 256GB memory (16 x 16GB DIMMs), per test system performance and memory consumption results during stress testing

PRODUCTION SITE

For on-site deployment, we have a two-way replicated Red Hat Gluster Storage deployment, with a minimum of 500TB usable storage after replication.

- **Redundant array of independent disks (RAID) level:** Each RAID volume will be built with 10+2 RAID 6 volumes. This setup provides sustainability for up to two drive failures per RAID volume.
- **Gluster replication**: The Gluster distributed storage with two-way replication to sustain node failures.
- **Component failure resiliency**: Dual SAS expander backplane and 2 x RAID cards to sustain expander and RAID card failures. This also provides efficient resource distribution – for example, a rebuild on a RAID volume would not degrade performance on another RAID volume.
- **Rack-level power failure resiliency**: Distributing data replicas across multiple racks ensures failure resiliency at the rack level.

Configuration details

- **Server**: 10 x 4924 (4U 24 Bay) with Dual SAS expander backplane
- **Processor**: Dual Xeon E5-2650V3 (10 cores, 2.3 Ghz)
- **RAID cards**: 2 x LSI 9361 RAID card with BBU and cache vault
- **Memory**: 256GB (16 x 16GB DIMMs)
- **Drives**: 24 x 512e Seagate 6TB 3.5" 7200RPM
- **OS drives**: 2 x 512GB Micron M600 SSDs
- **Storage**: 1440TB raw / 1200TB usable / 600TB usable after replication

Design choice

This is a well-balanced solution, as each RAID volume is built on a dedicated RAID card with cache vault and a dedicated SAS expander. This does not take down the whole node on a RAID card or SAS expander failure. Also, it sustains a total of four drive failures per node (two drive failures per RAID volume).

Expansion plans

Storage can be expanded by adding 2 x 4924 nodes (one per rack) as an expansion unit. Each expansion unit provides 288TB raw / 240TB usable / 120TB replicated usable storage.

OFF-SITE

The off-site cluster provides a single replica of the data stored in on-site storage. The off-site cluster has also been built with less component redundancy.

- **RAID level**: Each RAID volume will be built with 10+2 RAID 6 Volumes, which provide sustainability for up to two failures per drive
- **Component failure resiliency**: We chose to use single SAS expander backplane with one RAID card per server, which provides no tolerance on either of the component failures

Configuration details

- **Server**: 5 x 4924 (4U 24 Bay) with single SAS expander backplane
- **Processor**: Dual Xeon E5-2650V3 (10 cores, 2.3 Ghz)
- **RAID cards**: 1 x LSI 9361-8i RAID card with BBU and cache vault
- **Memory**: 256GB (8 x 16GB DIMMs)
- **Drives**: 24 x 512e Seagate 6TB 3.5" 7200RPM
- **OS drives**: 2 x 512GB Micron M600 SSDs
- **Storage**: 720TB raw / 600TB usable

Design choice

This is a slimmed-down version of the on-site storage unit but is a linearly scaled version of the solution, as each server holds 2 x RAID volumes. This takes down the whole node on a RAID card or SAS expander in the event of a failure. It also sustains a total of four drive failures per node (or two drive failures per RAID volume).

Expansion plans

Storage may be expanded by adding 1 x 4924 nodes as an expansion unit. Each expansion unit provides 144TB raw / 120TB usable storage.

LEARN MORE

To find out more about this joint solution from Red Hat and Penguin Computing, visit redhat.com/storage.

facebook.com/redhatinc
@redhatnews
linkedin.com/company/red-hat

NORTH AMERICA
1 888 REDHAT1

EUROPE, MIDDLE EAST, AND AFRICA
00800 7334 2835
europe@redhat.com

ASIA PACIFIC
+65 6490 4200
apac@redhat.com

LATIN AMERICA
+54 11 4329 7300
info-latam@redhat.com

redhat.com
F5809_0217