

# Top 5 Best Practices for Tick Data Analytics

The trading landscape has changed disruptively over the last few decades, and today, we are in a new normal. Volatility and the ability to extract profits from the current markets are elusive as of late. Even with some of the latest political situations challenging the world and causing uncertainties, the volatility has not returned consistently to the markets. This is causing ripple effects and business model changes at firms that have relied upon intraday trading for decades. Changes from an intraday trading model to one where positions are carried overnight have multiple impacts. First, the increase in capital requirements associated with holding positions overnight on most exchanges increases the capital outlays of the firm. Second, algorithms that will compute trading strategies will need to process a much larger volume of data at increased speeds. In this note, we discuss how the trading firms can leverage a modern infrastructure for tick data analytics to solve these challenges.

## 1 | Leverage Single Copy Architecture

A few years ago, it was common for single day market data streams to be sub 1 terabyte. Today, it is not uncommon for those streams to exceed 5, 10, and 20 terabytes per day. There are architectures where the customer can have a copy of that data for every compute node they support. There is also a centralized shared storage model where a single copy of all that market data can be centrally managed. Managing one single copy of market data is far less onerous than managing a copy for every compute node. 20 compute nodes equal 20 market data copies. In the shared storage configuration, hundreds of thousands to millions of simultaneous algorithmic jobs have been run on a cluster that holds multiple decades of market data. Scale Out Object like Dell EMC ECS and Scale Out File like Dell EMC PowerScale, offer choices of platforms of the protocol, but both offer a single copy architecture in central storage.

## 2 | Run Tick Data Analytics and AI/DL Workloads Simultaneously with Data in Place

It is increasingly common for market data to be used for both quantitative analysis and artificial intelligence (AI) / Deep Learning (DL) simultaneously. The co-location of data that supports both types of jobs & compute is very common in these environments. No data movement or specialized configurations are required to run both tick data analytics and deep learning from the same storage location with PowerScale or ECS.

## 3 | Take Advantage of Multi-Cloud

Hybrid cloud has garnered an increasing amount of attention in the past year. The tick data workloads are being adapted to leverage cloud and on-premises resources. Dell Technologies Cloud Storage for Multi-cloud allows a single copy of market data to reside in a hybrid cloud environment. This single copy of data can be utilized by the 3 major cloud service providers at the same time, without any need of moving or copying the data. All the while, data sovereignty and access control are maintained by the individual organization. Data access is handed over to the cloud service providers from the storage systems, and not managed by a cloud service provider. It is managed on the cluster by the user, which allows market data to be treated as the valuable data element that it is.

## 4 | Benefit from High-Performance Object Technology

ECS has come forward with a VERY high-performance object. Object technology is fairly new in the tick data space, not all technology in the tick data ecosystem supports object today. We are seeing a growing interest in object storage technology for tick data though, from long term retention to high performance utilization. The newest high-performance object is a game changer for those who want to leverage object technology in their tick data analytics or artificial intelligence.

## 5 | Leverage Massive Scalability

Tick data can require very large data volumes and the ability to store and manage hundreds of millions of files. PowerScale offers to small and mid-sized organizations entry-level high-performance All-Flash storage, with flexibility and longevity for future expansions. PowerScale's scale-out NAS allows you to non-disruptively scale capacity from terabytes to over 50 petabytes all in a single file system. ECS's scale out Object technology allows you to non-disruptively scale capacity from terabytes to over 500 Pb all in a single namespace. Imagine the simplicity of storing all your tick data under a single file system and namespace for all your applications to use. This provides less operational overhead to manage this scalability.