White Paper

Cloud for Financial Modeling

Use Cloud Resources to Run More Financial Models—Economically, Easily, and Right Now

AVERE

By Jeff Tabor & Greg Ulepic

November 2015

Use Cloud Resources to Run More Financial Models Economically, Easily, and Right Now

Financial services workloads require massive compute and storage resources. Satisfying that demand has traditionally meant investing equally substantial capital dollars in IT infrastructure. Physical ownership of 5,000 cores, for example, can require an outlay in excess of $3 million. With such prohibitive costs, technology buyers struggle to keep pace with unabating demand for compute and storage capacity.

The compromise many firms make is to place limits on the workloads run on in-house infrastructure. To produce results in available timeframes, analysts may limit testing frequency, the scope of modeling (ruling out testing at scale, for example), the number of asset classes used as input, and other such variables. Unfortunately, these restrictions can contribute to excessive risk, missed opportunity, and competitive disadvantage, particularly for smaller institutions with limited budget allocation for IT infrastructure.

The cloud, boasting nearly infinite compute and storage capacity, holds the potential to shatter traditional testing limits, enabling firms to apply tens of thousands of compute cores to modeling workloads. As a result, analysts can run more simulations in less time, perform testing at scale, leverage more inputs, and return results faster—all without capital investment. Pay-as-you-consume pricing eliminates capital expenditure, and per-core pricing that starts as low as pennies per hour offers economy no corporate-owned, on-premises data center can match.

So what prevents financial institutions from fully embracing cloud solutions? Concerns related to security, data migration, protocol incompatibility, scale, and manageability keep many businesses from moving forward with cloud strategies. Avere Systems technology removes these roadblocks to make the cloud a practical solution—one that can immediately deliver cost-effective, on-demand compute for financial analysis workloads like risk modeling and backtesting and nearly infinite storage capacity for market-data collection.

The following sections describe just a few of the opportunities Avere-enabled cloud compute and storage solutions can offer financial services organizations.

Use the Compute Cloud for Economy and On-demand Cores

Economical Compute

Major cloud compute offerings such as Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure each provide access to the latest processor technology at unprecedented scale and low cost. In mid-2015, for example, the Amazon Web Services Elastic Compute Cloud (EC2) service priced its on-demand c4.xlarge instance at just $0.22 per hour. (See https://aws.amazon.com/ec2/pricing/ for current Amazon EC2 pricing, https://cloud.google.com/compute/pricing for the latest Google Compute Engine pricing that includes a model for per-minute billing after the first 10 minutes of usage, and https://azure.microsoft.com/en-us/pricing/ for Microsoft Azure rates).

Cloud vendors also continue to innovate in pricing models, making spare compute capacity in their data centers available at extremely low cost for simulations with less-stringent completion requirements. In most cases, firms can adapt modeling queues to take advantage of periods when low-cost compute is available. The savings can be substantial—for example, Google Cloud Engine preemptible virtual machines prices start at $0.01 per core hour for a 70-percent savings over standard instances. AWS offers its own pricing innovation with spot instances, available via a continuous auction of spare compute capacity. AWS reports that clients can save as much as 90 percent by bidding on spare EC2 instances. Such aggressive pricing models make cloud compute attractive for applications where large numbers of compute cores are used.

On-demand Compute

Cloud service providers offer access to effectively infinite resources, making it possible for analysts to apply as much compute as desired to simulation and

modeling workloads. Instantly spin up thousands of cloud compute cores, run the simulation, then equally rapidly shut down the cores—and attendant rental fees. Because renting 1,000 cores for 10 hours, for example, costs the same as renting 10,000 cores for one hour, analysts have the flexibility to use as many cores as needed to complete models in less time, run more simulations, or run more complex models with more asset classes. The availability of on-demand compute capacity for any specified duration empowers the business without requiring costly ownership of continually aging compute assets.

Eliminate Roadblocks with the Avere-enabled Cloud

Minimize Latency and Data Movement

Effective use of the cloud, however, does require getting past several roadblocks. For example, the complex and time-consuming job of moving and synchronizing datasets has kept many firms from moving risk models and backtesting to the cloud. High latency between cloud compute and on-premises datasets can make cloud compute essentially unusable. While cloud-based block storage is available, it is an expensive place to store large volumes of data and migrating and synchronizing the precise subset of data needed in the cloud is a management challenge.

The Avere Systems solution eliminates manual data copying and management, providing low-latency, transparent access to modeling data wherever it resides. Avere virtual and physical FXT Edge filers ensure high-performance access to data whether the data is local or remote, NAS- or object-based, on premises (including in data centers or at remote sites), or on a public cloud like the AWS platform or the GCP. For the compute cloud, Avere vFXT Edge filers deliver low-latency access by automatically caching active data, storing it on SSDs provisioned alongside the compute instances in the cloud and ensuring low-latency access to cached data. The Avere solution allows firms to maintain the bulk of their modeling data on-premises, caching (and encrypting) only active data in the cloud.

Handling all read, write, and metadata operations near the compute, the Edge filer allows simulations to run at high performance. Avere technologies make it

practical to leverage compute-cloud resources to replace provisioned IT infrastructure or as bursting compute, supplementing on-premises compute-farm capacity for fastest response and allowing analysts, for example, to quickly run new models to assimilate data from major market events.

Secure Data and Simplify

Avere FXT Edge filers provide high degrees of security, encrypting all cached and in-transit data with AES-256 encryption and customer-supplied and customer-managed encryption keys. Avere filers leverage the OpenSSL Federal Information Processing Standard (FIPS) Object Module for data and network encryption and are FIPS 140-2 Level 1 compliant to ensure the highest standards of data security.

Avere solutions also eliminate issues of interface incompatibility to simplify access to cloud resources. FXT Edge filers use object APIs to connect to public and private clouds. Transparent to both applications and users, the Avere solution automatically translates between those APIs and file-system protocols (NFS/SMB/CIFS) to enable seamless access to any data, anywhere. Without having to make changes to applications or access methods, firms can make nearly immediate use of cloud services.

Maximize Performance and Scale

Avere solutions scale from three to as many 50 FXT nodes to deliver millions of operations per second and throughput beyond 100 gigabytes per second. Each Edge filer added to the cluster contributes more CPUs and DRAM for performance, as well as more SSD capacity to support larger working sets and even higher cache hit rates.

Avere also makes it easier to take advantage of more cloud cores (including across multiple availability zones) without requiring changes to existing workflow processes and without the complex configuration and setup associated with specialized cloud file structures. Each Avere FXT node can connect to 10,000 NFS clients, enabling a 50-node total of 500,000 clients (IP addresses).

Unleash the Compute Cloud on Risk Models and Backtesting

Over the past several years, economic events and increasing regulatory requirements have been driving financial firms to measure risk even more comprehensively. As a result, risk models, for example, have increased in complexity to include more portfolio assets, more randomly generated paths, and more granular timesteps.

Such increased complexity can dramatically impact the load on compute farms. For example, while calculating the Greeks using Monte Carlo simulations for five asset classes might take an hour to run on several hundred compute cores, increasing the calculation to include ten asset classes can drive requirements to 1,500 cores for results in 24 hours.

To fully appreciate the economic and time-to-results benefits of the cloud, consider the impact of running an exceptionally large workload on cloud compute cores. For example, one set of test results suggests that firms could feasibly run as many as one million large-workload jobs (such as risk models with up to 10 assets, some 100,000 paths, and more than 1,000 timesteps), using approximately 1,500 instances (>50,000 vCPUs) to complete the entire queue in a single day at a cost well under $100,000.

But to take full advantage of cloud scale, firms must effectively manage data across thousands or tens of thousands of cores—and not all cluster file systems work seamlessly with large-scale, parallelized jobs run in the cloud. Hedge fund firms are gaining this value from Avere today. One firm in particular recently abandoned its use of a popular cloud file system because of the complexity associated with scaling compute nodes beyond 10,000 cores. Now an Avere customer, the firm uses Avere technology to utilize many times more than the previous 10,000 core limit to run analyses on model data stored on a NAS array at the firm's on-premises data center. When risk analysis jobs run in the compute cloud, the firm's Avere vFXT cluster automatically populates its cache with the latest quant model data from the NAS system.

Avere technology provides the ideal solution for firms that want to utilize the cloud to run parallel workloads requiring access to hundreds or thousands of small files. The Avere vFXT solution also enables writing out log files at scale. Avere enables the scale required for CPU-intensive backtesting workloads, providing seamless access to the multiple terabytes of on-premises historical data required to run these simulations. Avere makes it possible to utilize the cloud to do backtesting more often and with results returned dramatically faster for improved investing strategies and risk management.

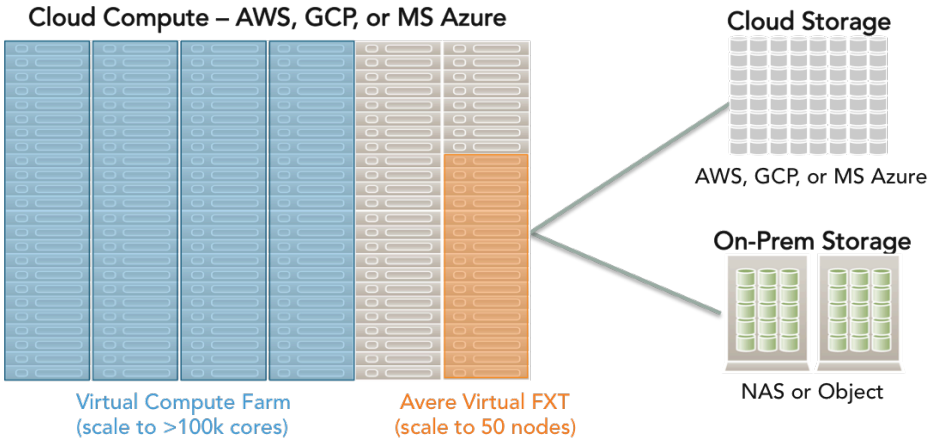

Figure 1: Avere vFXT clusters enables scaling cloud compute beyond 30,000 cores for risk analysis and backtesting

Use the Storage Cloud for Economy and Infinite Capacity

Along with compute cores, cloud services such as AWS, GCP, and Microsoft Azure offer infinitely scalable, low-cost storage. Companies such as IBM/Cleversafe and HGST/Amplidata have similar solutions for installation in private data centers. These solutions are based on the same underlying object storage technology that allows for the lowest cost per terabyte of storage available today. For example, Amazon Simple Storage Service (S3) costs as low as $0.0275 per GB per month, and Google Cloud Storage costs $0.026 per GB per month for Standard Storage and $0.01 per GB per month for Nearline

Storage. An IBM/Cleversafe private object solution can be up to 72-percent less than a traditional RAID-based system.[1]

Unleash the Storage Cloud for Fast Access to More Market Data

Taking advantage of the storage cloud lets financial services firms more cost-effectively retain market data. With many institutions collecting more than 200TB of new market data each month, the cost of maintaining on-premises capacity is becoming increasingly prohibitive. Using Avere, firms can store the most active data on premises, then as the data ages, seamlessly move it off to a lower-cost storage tier such as Google Cloud Storage. Google Cloud Platform offers three storage classes—Standard Storage, Durable Reduced Availability (DRA) Storage, or Nearline Storage—to meet varying access and cost parameters.

As in the case of the compute cloud, the storage cloud comes with the challenges of incompatible (and unfamiliar) access protocols and high latency caused by distance and/or bandwidth restrictions—all issues that Avere solutions effectively resolve, as described previously in this paper. Avere provides seamless data migration from on-premises storage to the cloud (via Avere FlashMove® software) and makes it easy for businesses to actively read, modify, and write data located in Amazon, Google, and Microsoft object-based cloud storage—without latency and without disruptions to users.

Benefits: Find More Opportunities, Faster

The Avere-enabled cloud provides financial services institutions a smooth on-ramp to cloud scale and economy. On-demand access to both compute cores and storage capacity gives firms new flexibility and agility to align infrastructure resources with business needs. With the ability to more quickly run more analyses using more input data, analysts can be better equipped to

[1] https://www.cleversafe.com/documents/public/Forrester-TEI.pdf

manage market volatility and to identify and move on successful investing strategies.

Avere enables **massive scale**—scale seamlessly across thousands or tens of thousands of compute cores, including across zones within a single cloud service provider and across compute clouds offered by multiple service providers. Businesses are already using Avere solutions to sustain production runs on more than 40,000 cloud cores.

Avere delivers cloud **economy**—instead of spending tens of millions of dollars on in-house compute assets, financial institutions can run simulations at five cents per core hour or as low as a penny per hour for non-time-critical workloads that can take advantage reduced-rate spare cloud capacity.

Avere technology provides **user-friendly access**—use cloud resources the same as data center infrastructure, taking advantage of the cloud without moving data, without changing applications and workflows, and without changing how analysts work.

For more information about using Avere FXT Edge filers to harness the power of the cloud for financial services applications, http://www.averesystems.com/solutions/industries/financial-analysis.

More Cores, More Inputs, More Speed, More Economy

Financial services institutions take advantage of the Avere-enabled cloud to run more complex simulations faster and without the high costs of procuring, installing, and maintaining data center infrastructure. Avere solutions are helping firms:

- Get faster answers—spend the same amount to rent 10,000 cores for one day or 1,000 cores for ten days
- Run more simulations and what-if scenarios with more input variables
- Easily set up compute cores in just minutes without touching any hardware

- Use cloud bursting for short-team peak workloads and avoid capital investment in aging technology assets
- As market data ages, easily move it to economical cloud storage and avoid capital investment in additional on-premises disk or tape capacity
- Shrink on-going data center operational expenses

About Avere Systems

Avere is radically changing the economics of data storage. Avere's hybrid cloud solutions give companies—for the first time—the ability to end the rising cost and complexity of data storage and compute via the freedom to store and access files anywhere in the cloud or on premises, without sacrificing the performance, availability, or security of enterprise data. Based in Pittsburgh, Avere is led by veterans and thought leaders in the data storage industry and is backed by investors Lightspeed Venture Partners, Menlo Ventures, Norwest Venture Partners, Tenaya Capital, and Western Digital Capital. For more information, visit www.averesystems.com

Jeff Tabor is the Sr. Director of Product Management & Marketing at Avere Systems

Greg Ulepic is the Central Area Sales Director at Avere Systems