# Accelerate SAS® Analytics Workloads by 5X with WekaFS™ and Destiny Corporation

**"Using this reference design, organizations will be able to consolidate more users and workloads with the same SAS licensing in place, delivering exceptional economic value.**

**Not only will this improve time to market and business outcomes, but it will eliminate racks of server infrastructure for a significantly improved total cost of ownership.**

*Shailesh Manjrekar, Head of AI, WekaIO*

SAS analytics software is one of the most widely deployed business intelligence tools, with 94% of the Fortune Global 1000 customers using SAS applications to power their businesses. Through innovative software and services, SAS empowers and inspires customers around the world to transform data into intelligence. Data-driven applications in artificial intelligence (AI) and machine learning (ML) using SAS analytics are increasing in volume and complexity as organizations look to enable predictive and cognitive analytics for financial services industry (FSI), risk management, healthcare, and IoT use cases.

While SAS Grid Manager and now SAS Viya® are architected to meet the demands of modern business analytics and to take optimal advantage of modern multi-core CPUs, their performance and scale is ultimately constrained by access to volumes of data and storage I/O responsiveness. Even today's leading on-premises and cloud storage solutions cannot keep up with the capabilities of modern CPUs and servers. As a result, SAS Grid Manager and SAS Viya application servers can remain vastly underutilized as they wait for data to be served from the storage system when operating across large data sets.

To enable predictive and cognitive analytics for these heavy random and sequential read/write mixed workloads, the underlying infrastructure needs to accelerate job completion times and data pipelines by delivering ultra-low latency with massive ingest bandwidths.

## OVERCOME CHALLENGES WHEN MODERNIZING SAS, R, AND PYTHON APPLICATIONS

IT organizations face challenges when they modernize their operations because they need to maintain legacy applications and deal with undocumented, inefficient application logic and performance limitations due to long running jobs, processes, and pipelines. Digital transformation is driving new use cases, business models, and application architectures to cater to AI/ML and IoT. SAS users are also increasingly using open source front ends to SAS, like R and Python.

Chief Data Officers (CDOs) and Chief Administration Officers (CAOs) are faced with big problems, including improved SAS user productivity for long running jobs that are not meeting Service Level Agreements (SLAs). There must be better utilization, scalability, and stability of application servers that crash because of resource issues, and the IT team must keep up with the storage capacity demands from increased data volumes and multiple copies of the same data. Furthermore, lowering SAS licensing costs and the overall costs for SAS

**"As a trusted advisor to our clients, we always look for new methods and technologies to help our clients do their analytics work.**

**Destiny chose to partner with WekaIO because we like the disruptive capabilities of this technology that elegantly solve our clients' problems with no change needed to SAS code.**

*Dana Rafiee, Managing Director, Destiny Corporation – Business and Technology Consulting, specializing in SAS analytics*

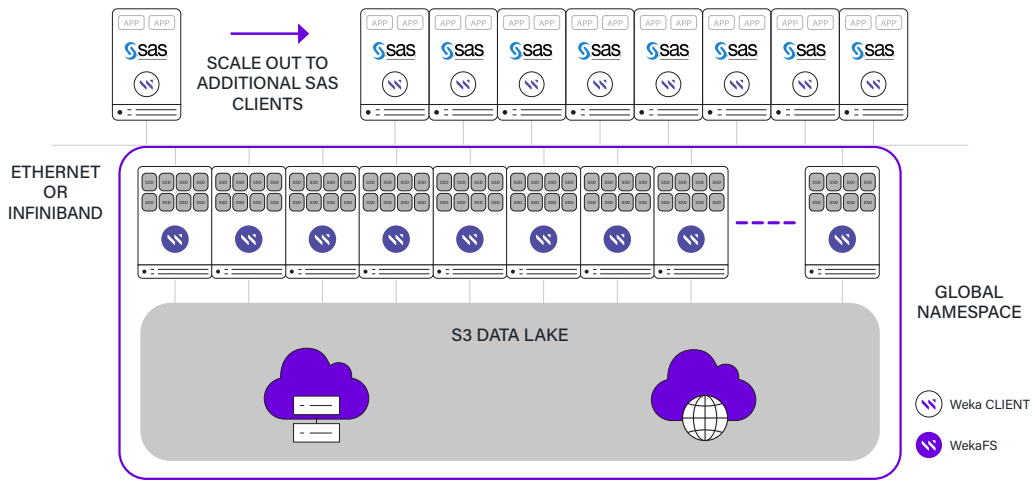


*Figure 1 – WekaFS and SAS Analytics in a Production Environment*

**ACCELERATE TIME TO MARKET**
- Business outcomes, time to market, long running jobs and data loads for SAS analytics at scale

**OPTIMIZE YOUR INVESTMENT**
- Run 10X more workloads with the same SAS licensing
- Improve TCO, eliminating racks of server infrastructure
- Maximize storage costs with a hybrid storage data lake

**SIMPLIFY MANAGEMENT**
- No code changes, continue to use SAS Enterprise Guide / Data Miner

**AGILE RESPONSIVENESS**
- Spin up and down SAS servers in the public cloud to meet your workload requirements

**ENTERPRISE USE CASES**

**Artificial Intelligence**
- Predictive to Cognitive Analytics

**Machine Learning**
- Classification, Regression, Clustering

**Natural Language Processing**
- Speech Recognition
- Contextual Marketing

**ENTERPRISE VERTICAL SOLUTIONS**
- Banking
- Government
- Healthcare and Life sciences
- Manufacturing and Energy
- Telco
- Retail

**COMPLIMENTARY SAS HEALTH CHECK AVAILABLE FOR QUALIFIED CLIENTS! GET STARTED AND GO TO: HTTPS://BIT.LY/30X02GC**

infrastructure, including servers, storage, backup, disaster recovery, and maintenance, is another important challenge for CDOs and CAOs.

## ACCELERATE SAS FOUNDATION, SAS GRID MANAGER, AND SAS VIYA WORKLOADS

WekaIO™ (Weka), working closely with Destiny Corporation, a SAS consulting company, has devised a reference design for significant runtime and wall clock time improvements for long-running SAS jobs, all while providing the best economics. The design comprises of SAS application servers configured either with SAS Foundation, SAS Grid Manager, or SAS Viya software. These servers connect over InfiniBand or 100Gbit Ethernet to a storage system running WekaFS™, the world's fastest file system. The shared storage system starts at 8 nodes with NVMe SSDs and can scale to hundreds of nodes, depending on the throughput requirements. In addition, WekaFS seamlessly integrates object storage for the best economics on multi-petabyte data sets. Figure 1 provides an overview of a typical implementation of WekaFS with SAS.

The reference design is tuned to ensure over 300MB/sec of bandwidth per SAS core while running an I/O intensive SAS test suite. This was devised by Destiny Corporation based on its 30 years of experience with SAS applications. The test suite involves reading and writing over 10 million SAS tables with data loads and multi-threaded SAS base procedures. The suite represents commonly encountered SAS functions – summary calculations, statistics, summary reports and data transformation and small and large file random and sequential access, and a mix of 60% Write and 40% Read I/O patterns.

On a reference design, scaling from 1 SAS client running 32 concurrent sessions to 8 clients running 256 concurrent sessions, WekaFS demonstrated linear scaling for this challenging test suite. The solution measures up to 106GB/sec of mixed Read/Write bandwidth while only taking 320 SAS cores to achieve this performance. In addition, the CPU-time to real-time ratio is maintained at close to 1, indicating that the cores are fully utilized, without cycles wasted on data I/O. This reference design can be deployed either on-premises, in AWS, or in a hybrid model to get similar results. WekaFS matches the parallelism of the SAS Grid Manager compute environment to enable the highest data loads and ETL (extract, transform and load) processing while providing enterprise data management capabilities to SAS workloads.

| PERFORMANCE | 1 CLIENT | 2 CLIENTS | 4 CLIENTS | 8 CLIENTS |
|---|---|---|---|---|
| Scale No. of Sessions | 32 | 64 | 128 | 256 |
| R/W Ratios | 40 Read :60 Write | 40 Read :60 Write | 40 Read :60 Write | 40 Read :60 Write |
| Scale Throughput | 13.3 = 5.5 + 7.8GB/sec | 26.6 = 11 + 15.6GB/sec | 53.2 = 22 + 31.2GB/sec | 106 GB/s = 44 + 62GB/sec |
| No. of cores | 40 | 80 | 160 | 320 |
| Throughput per core | 332MB/sec | 332MB/sec | 332MB/sec | 332MB/sec |
| Weka Storage Nodes | 8 nodes | 8 nodes | 10 nodes | 16 nodes |

*Table 1 – Performance Scaling of WekaFS*

## ENTERPRISE READINESS WITH ACCELERATED DATAOPS

In addition to accelerating your SAS workloads, the reference design leverages object storage data lakes to extend the file namespace to provide the best economics by leveraging both flash and HDD media. The solution also provides checkpointing for SAS Grid Manager and SAS Viya deployments along with data lifecycle management using file namespace snapshots for backup and recovery.

* SPEC® SFS2014 Results: https://bit.ly/35UHLDk
WekaFS™ Data Sheet: https://bit.ly/36BRJdc

WEKA

910 E Hamilton Avenue, Suite 430, Campbell, CA 95008 T: 408.335.0085 E: info@weka.io www.weka.io