

WekaIO Improves Performance and Stability for Geospatial HPC Workloads in the AWS Cloud



“In the market of parallel filesystems, WekaIO Matrix is distinguished from the other players for its modernism. You can feel this modern approach everywhere: in the product and technology, in Support and in the cloud-oriented options.

Alessandro Menegaz, Cloud IT Manager at TRE ALTAMIRA

USE CASE

Geospatial HPC Workload in the AWS Cloud

COMPANY

TRE ALTAMIRA

BUSINESS CHALLENGES

- Production capacity limits
- Time-to-market dependency on storage infrastructure
- Constrained innovation a major disaster
- Compatibility with Ubuntu Linux Distribution

SOLUTION

WekaIO Matrix™ on AWS

RESULTS

- No limit on production capacity
- Improved time-to-market by 7x
- Reduced costs by 3x
- Improved innovation has led to new products and services

TRE ALTAMIRA, headquartered in Milan, Italy with offices worldwide, is a leader in measuring ground and structural movement from satellite radar data for oil and gas, mining, civil engineering and geohazard industries. The company uses the images it acquires to analyze the movements on areas that interest its customers. Its services include measuring surface deformation, earth fractures, reservoir model calibration and caprock integrity surveillance. The company can monitor gas fields to ensure reservoir pressure and surface uplift remains within safe operational limits. For instance, acquired images are used by a major oil company to monitor its oil fields, and input data is provided by many space exploration agencies such as NASA and the European Space Agency.

The data acquired in TRE ALTAMIRA analyses consists of large datasets that each could exceed 4TB in size. Generated by internally developed software, this data is ingested into a Linux high performance compute (HPC) cluster for processing in chains of more than 100 steps. Some chains include thousands of parallel jobs that need to access the ingested data and provide partial results of previous steps at the same time, requiring significant processing power and use of high-performance, low-latency storage.

THE CHALLENGE

The processing power and the performance of the storage have a strong impact on TRE ALTAMIRA's business because they determine the company's production capacity and time-to-market. Additionally, these two factors strictly impact the company's innovation capabilities—they couldn't consider new products or services that would require too many processing and/or storage resources.

“Until last year all our results were produced on our on-premises data-center so production capacity, time to market and innovation capabilities were limited by the number and the power of our machines and by the performance and the capacity of our storage,” says Alessandro Menegaz, Cloud IT Manager for TRE ALTAMIRA.

Menegaz continues. “The production capacity was limited to 30 analyses per week, so we simply could not accept more orders than the ones compatible with this number or we had to extend the delivery time. Sometimes it took more than three weeks to produce a result simply because all our resources were already in use. These limitations did not encourage innovation, and sometimes in the past, new promising algorithms were discarded due to the lack of resources.”

TRE ALTAMIRA had three criteria any new system needed to meet:

- It needed to be compatible with their Ubuntu Linux distribution
- It needed to be faster than their on-premises storage solution which processed data at 5GBps per 100TB and, at a minimum, Lustre's performance on the cloud
- It needed to be stable

THE SOLUTION: WekaIO MATRIX ON AWS

TRE ALTAMIRA had been using the public cloud to run its complex workflows and initially turned to Lustre for its production HPC software. But, Lustre didn't support its installed Ubuntu Linux distribution and required a risky source code recompile that failed on several occasions, thus affecting business. Based

on the recommendation from an Amazon Web Services solutions architect, they chose WekaIO Matrix, a parallel filesystem. Matrix software is installed on each node of a four macro-block HPC cluster. TRE ALTAMIRA software transfers the input data for analysis on the filesystem. The processed data is then archived on Amazon S3 (Simple Cloud Storage Service) and the entire HPC cluster is shut down.

"In the last 20 years, we tested and used in production several different software solutions: GFS, GPFS, PANFS, and Lustre. The performance and ease of installation of Matrix is something we had never experienced before, and the Support is outstanding. We received answers in minutes over a dedicated Slack channel, and this helped a lot during the adoption phase," explained Menegaz.

BENEFITS ACHIEVED AND RETURN ON INVESTMENT

The benefits and return on investment TRE ALTAMIRA achieved by using WekaIO Matrix include:

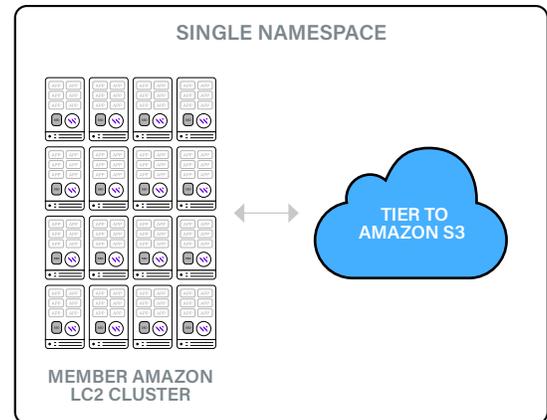
- Increased production capacity by a factor of 7x. Previous on-premises analyses would have exceeded two weeks and consumed all available resources. By running Matrix in AWS, analyses times were reduced from two weeks to less than two days and only required a portion of available resources.
- Costs scale linearly with the workload. If there are no running analyses, there is no cost. By shutting down the filesystem after completing the processing, analysis costs were reduced from approximately \$60,000 per analysis to less than \$20,000, a 3x reduction.
- There is no longer a dependency between time-to-market and the number of concurrent analyses, making the sales process faster and simplifying the product development process.
- It encouraged innovation. Products and services that had been abandoned in the past due to the lack of infrastructure resources are now feasible with AWS and WekaIO Matrix.
- Installation, setup, and upgrade of the product is quick and easy. TRE ALTAMIRA upgraded the software multiple times to access new features and each time the provisioning of a new custom AMI required less than 30 minutes and was non-disruptive.
- World-class support. Using modern collaboration tools such as Slack and automated monitoring provides rapid, proactive response to support scenarios.

Menegaz summarizes, "In the market of parallel filesystems, WekaIO Matrix is distinguished from the other players for its modernism. You can feel this modern approach everywhere: in the product and technology, in the Support, and in the cloud-oriented options."

For a free trial of Matrix on AWS, go to <https://www.weka.io/get-started>.

About AWS

For 10 years, Amazon Web Services has been the world's most comprehensive and broadly adopted cloud platform. AWS offers more than 90 fully-featured services for compute, storage, databases, analytics, mobile, Internet of Things (IoT) and enterprise applications from 42 Availability Zones (AZs) across 16 geographic regions in the U.S., Australia, Brazil, Canada, China, Germany, India, Ireland, Japan, Korea, Singapore and the UK. AWS services are trusted by millions of active customers around the world monthly — including the fastest growing startups, largest enterprises, and leading government agencies — to power its infrastructure, make it more agile and lower costs. Learn more about AWS, visit aws.amazon.com.



HPC Storage on AWS is possible with WekaIO Matrix.

