WEKA

# CryoSPARC Workflow Integrations with WEKA

REFERENCE ARCHITECTURE

# Contents

# Executive
## Summary

Cryo-electron microscopy (Cryo-EM) is a technology that allows scientists to use electron microscopes to look at and determine the 3D structure of various molecules, protein assemblies, viruses and more. Cryo-EM is the primary technique currently being used to understand protein shapes in the areas of drug discovery, vaccination research, and other life sciences areas. During the use of a Cryo-EM system, massive quantities of images from the microscope (micrographs) are generated.

CryoSPARC is a popular software package that integrates with the Cryo-EM machines and manages the entire workflow process from scanning and data collection through the various steps in the computational process to get to the final 3D result.

The workflow associated with Cryo-EM micrograph processing consists of preparing the molecular / biological sample to be scanned, scanning the sample and creating a data set of micrographs, processing the micrographs through a workflow of computational procedures that isolate and refine the images to make them useful, then combine the results to form a 3D view of the sample that was scanned as well as all the associated information about how the sample structure is constructed.

The raw data per sample scanning process can be in the 5-10TB size, and the process to get to the 3D output may involve multiple iterations of processing in the computation pipeline to resolve the 3D image appropriately. In many cases, the entire workflow including scanning a new sample will need to be repeated in order to achieve the appropriate results. This results in a massive amount of data that needs to be stored and retrieved at high speed in order to make the pipeline work efficiently. WEKA's unique architecture and extensive feature set provides the necessary storage capabilities to accelerate this pipeline.

Many organizations that use CryoSPARC are taking advantage of the ability to rent GPUs and scale the compute functions in the cloud in an elastic manner for processing the micrograph images. Being truly software defined WEKA is able to operate in these environments and provide the same benefits in the cloud as on-prem.

# 1. Introduction

## Audience

This paper is for those who design and manage CryoSPARC systems and have storage and data access challenges within their CryoSPARC environment.

We have organized this document to address critical items for enabling successful design, implementation, and transition to operation.

## Purpose

This document covers the following subject areas:

- Overview of WEKA.
- Overview of CryoSPARC and WEKA integration.
- The benefits of CryoSPARC on WEKA.
- Recommendations for architecting a complete CryoSPARC solution in the cloud using the WEKA Data Platform, including design and configuration considerations.
- Benchmarks for performance on WEKA.

## Document Version History

| Version Number | Published | Notes |
|---|---|---|
| 1.0 | July 2023 | Author: Joel Kaufman. |

# 2. CryoSPARC

## Overview of CryoSPARC

CryoSPARC is a popular software package that integrates with the Cryo-EM machines and provides the algorithms and computational functions to transform the output of the Cryo-EM systems into meaningful 3d images. In addition, CrySPARC manages the entire workflow process from scanning and data collection through the various steps in the computational process to get to the final 3D result.

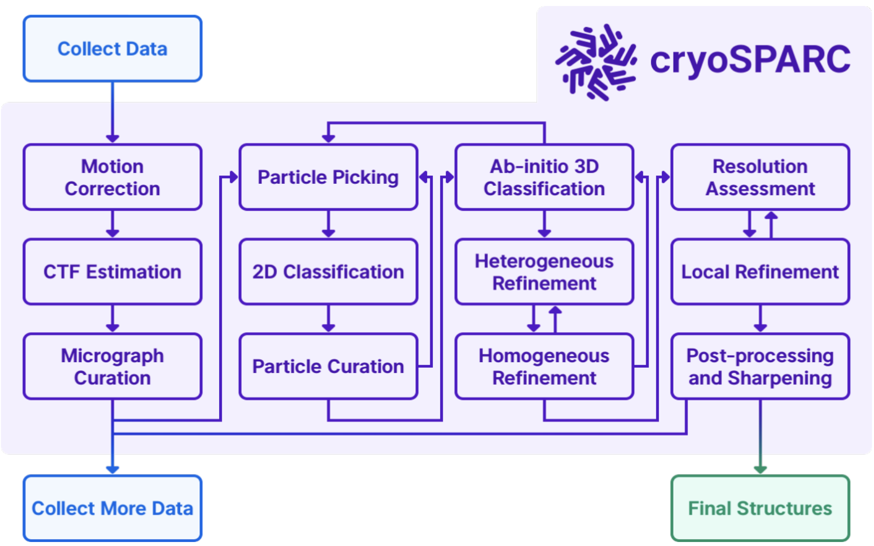


**FIG. 1** CryoSPARC process steps

This process consists of 12 discrete steps that each have unique requirements for IO to process the data rapidly, which can be a challenge for most storage solutions. When processing the 12 computational steps, CryoSPARC utilizes a master node/worker node model for the computational cluster. The master node is essentially a job scheduler, catalog and witness for the work being done by the worker nodes.

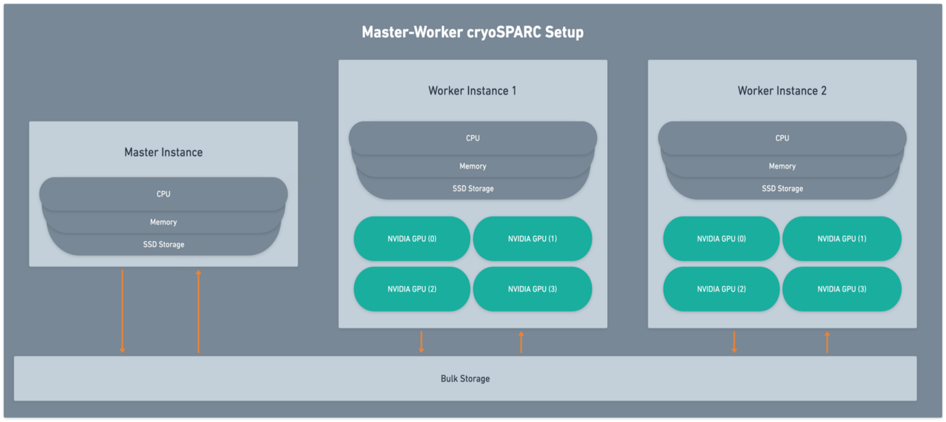


**FIG. 2** CryoSPARC master/worker nodes

In order to use CryoSPARC, there are a number of requirements for the deployment such as the use of NVIDIA CUDA Toolkits to access the required NVIDIA GPUs for processing, The GPUs themselves, and crucially, a shared high performance filesystem for the underlying storage for all stages of the process. The shared filesystem is required as each project created by a user is associated with a single project directory in the filesystem that all CryoSPARC nodes must be able to read from and write to during the various phases of computation.

# 3. WEKA

WekaFS is a software defined data platform that can run natively on common X86 servers/cloud instances/VMs either on-premises or in major hyperscalars such as AWS, GCP, Azure and OCI. WekaFS, the filesystem that powers the data platform, uniquely integrates with NVMe drives and high-performance networking in these servers to create a multi-protocol capable storage cluster to deliver data at unbeatable speeds and scale to clients.

## Performance at Scale

WekaFS is the world's fastest and most scalable POSIX-compliant file system. It is designed to transcend the limitations of legacy file systems that leverage local storage, NFS, or block storage, making it ideal for data-intensive AI and HPC workloads. WekaFS is a clean sheet design integrating NVMe-based flash storage, object storage for tiering and data protection, and ultra-low latency interconnect fabrics such as 200GbE or InfiniBand, creating a highly performant scale-out storage system with the breakthrough ability to handle streaming throughput, random I/O and metadata intensive workloads on the same filesystem at the same time. WekaFS performance scales linearly as more servers are added to the storage cluster allowing the infrastructure to scale with the increasing demands of the business. WEKA can scale both on-prem or in the cloud.

## Expandable Global Namespace over S3 Object Store

WekaFS delivers best-of-breed performance from the NVMe flash tier, and the namespace can expand to any S3 compatible object store, on-premises or in the cloud. This optional hybrid storage model with the ability to extend the global namespace to lower-cost media in an object store delivers a cost-effective data lake without compromising performance. This integrated tiering to multiple S3 compatible targets enables cost-effective data lifecycle management for older or less used data.

## Advanced Durability and Security

Large and agile datasets frequently require a data versioning capability. This is achieved using WEKA's instant and space-efficient snapshot capability. The Snap-To-Object (S2O) feature captures a point-in-time copy of the entire file system (flash and object store) namespace that can then be presented as another file namespace instance to any other WEKA system whether on-prem or in a private or public cloud. With additional management integration, WEKA's integrated snapshots and end-to-end encryption features ensure that data is always backed up and secure throughout its lifecycle. WekaFS also provides immutability and data mobility for these datasets with instant recovery. WekaFS can seamlessly back up to multiple cloud targets providing backup, DR, and data governance capability.

## Cloud Bursting and Data Mobility

In addition to providing versioning, Weka's snap-to-object feature offers additional benefits beyond backup and DR to the public cloud; it enables secure data portability from on-premises to the public cloud for organizations that require access to on-demand GPU and CPU resources in the public cloud.

## Multiprotocol and Container Support

Along with POSIX access, WekaFS supports multiple file access protocols, including NFS, SMB, S3, and Nvidia GDS for maximum compatibility and interoperability. Different protocols can access the same data sets and files concurrently, maximizing efficiency by preventing the need to copy data to different locations to support different protocols. Hadoop and Spark environments also benefit from the performance of a shared file system through a fully integrated connector that allows WekaFS to overcome the limitations of HDFS and function as a single, easy-to manage data lake for all forms of analytics.

Organizations are also increasingly adopting containers deployed on Kubernetes (K8s) for a variety of workloads. Using the WekaFS CSI plug-in, organizations now have flexibility in how and where they deploy containerized applications. It provides easy data mobility from on-premises to the cloud and back while delivering the best storage performance and latency. Figure 3 provides an overview of WekaFS in a typical production deployment.

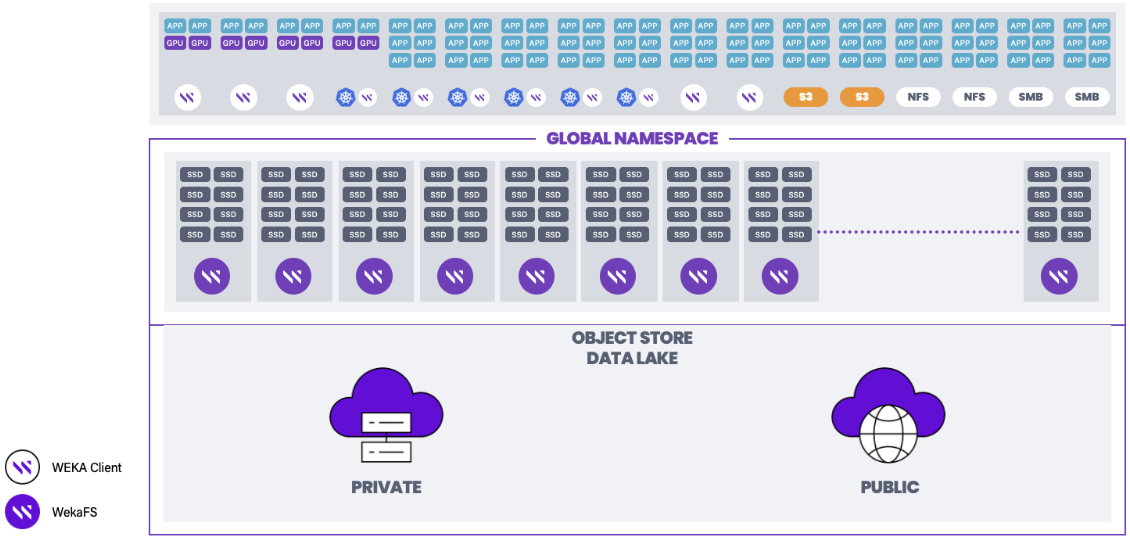


**FIG. 3** WekaFS

# 4. Using CryoSPARC with WEKA

For WEKA to work with CryoSPARC, it needs to be connected to the master and worker nodes. For highest performance, WEKA recommends the use of our POSIX driver (aka WEKA Client). In Figure 4, we show an example of a deployed environment in AWS.

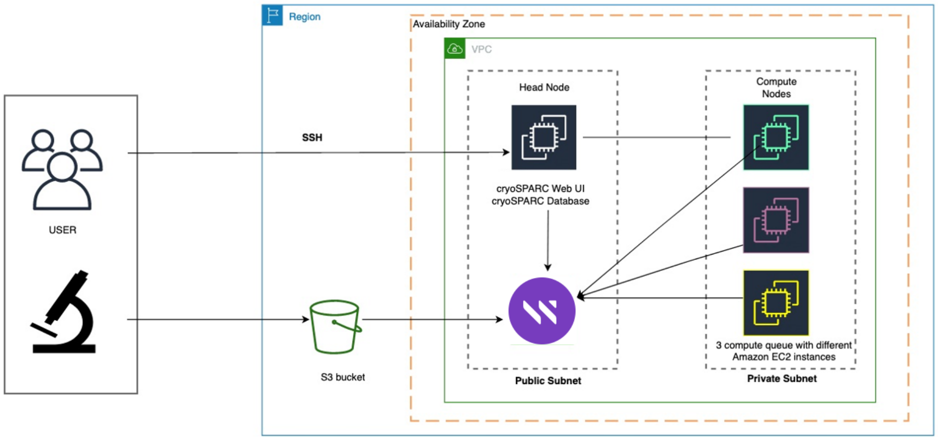


**FIG. 4** WEKA connecting to CryoSPARC nodes in AWS

### Best Practices

Use the WEKA Client whenever possible. WEKA's continued innovation and support of Magnum IO GPUDirect Storage technology provides low-latency, direct access between GPU memory and storage. This frees CPU cycles from servicing the I/O operations and delivers higher performance for the CryoSPARC workload.

## Validation and Benchmarking

Working with a third party consultancy, Clovertex (see appendix for details), WEKA has benchmarked CryoSPARC in a cloud based hyperscalar environment and produced a meaningful comparison of WEKA vs. AWS FSx Lustre when backing CryoSPARC environments.

## 5. Environment Overview

The environment tested was set up similar to Figure 4 above. The only difference is that there are no external users, and a pre-built dataset was used. The tests were fully automated with no manual process of moving data through the CryoSPARC pipeline.

### Test Environment Configuration

| | |
|---|---|
| **User pipeline simulation:** | 10 simultaneous users, 1 CryoSPARC live session and 5 SLURM lanes/queues of concurrent data sent to multiple GPU instances. |
| **Dataset:** | EMPIAR-10288 cannibinoid receptor protein complex data consisting of 2756 TIFF images comprising 476GB of data. Each users output dataset is stored in an individual directory, as shown in Figure 5. |
| **Cryosparc clients:** | Comparisons of 5 different GPU based cloud instances were used, ranging from 1-8 GPUs. G4dn.4xlarge, G5.4xlarge, P3.8xlarge, G5.24xlarge, and G5.48xlarge |
| **WEKA Storage cluster:** | 8x i3en.3xlarge AWS instances with up to 25Gbit networking and 1x 7.5TB SSDs. A single 35TB filesystem was presented for all the data directories to reside in and was mounted on the clients via the WEKA POSIX agent |
| **Amazon FSx Lustre :** | A single uncompressed 35TiB filesystem at 500MB/S/TB throughput was provisioned and presented for all the data directories to reside in and was mounted on the clients via the AWS recommended Lustre mount command. |

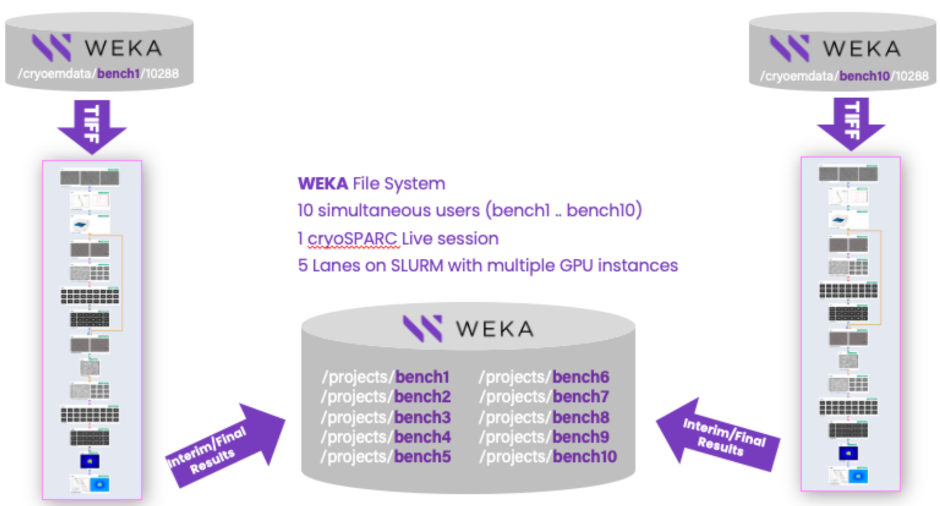


**FIG. 5** Benchmark Workflow

For the workflow, each user had in an individual copy of the Cryo-EM source data set in a dedicated directory. And after the data was processed by CryoSPARC, the final data was written to a dedicated project directory. While the data was identical in each source directory, in real life each one of these directories could have different Cryo-EM image sets in them. CryoSPARC is designed to handle multiple datasets running in all different stages of the pipeline concurrently. This benchmark is a good example of concurrency in a real-world environment.

# 6. Validation and Benchmark comparison

A full simulation of the entire dataset workflow running through different GPUs using either a FSx Lustre and a WEKA storage system produced the following results:

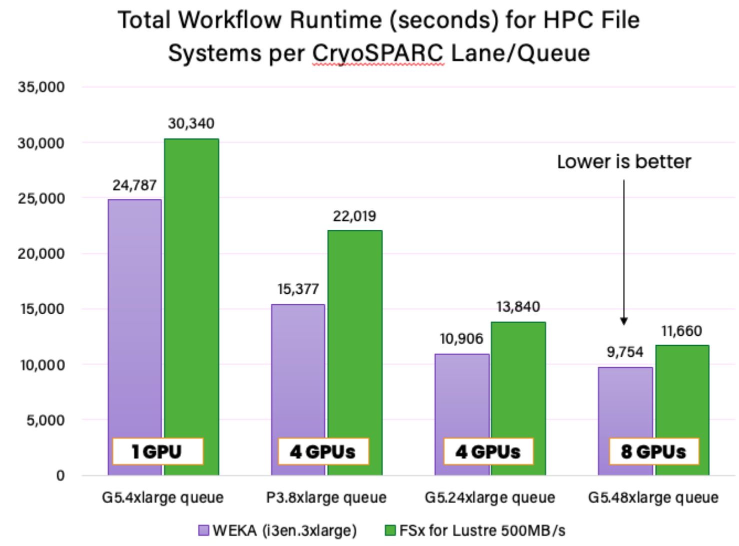


**FIG. 6** Runtime comparison for CryoSPARC lanes

## Interpreting the Results

The results are broken down based on the different GPU instances used. The different GPU instances have different characteristics. The G5 series use the Nvidia A10G GPU, While the P3 series use the Tesla V100 GPU. The instances also have different CPUs, memory allocations, networking, and more. The differences in runtime between WEKA and FSx Lustre reflect the ability of the filesystem to keep the CryoSPARC queue as saturated as possible across different GPUs. In these cases, as shown in Table 1, the combined time for a full queue run resulted in a WEKA advantage across the board.

| GPU Instance Type | WEKA Time | FSx Lustre Time | WEKA Advantage |
|---|---|---|---|
| G5.4xlarge | 24,787 | 30,340 | 19% faster |
| P3.8xlarge | 15,377 | 22,019 | 30% faster |
| G5.24xlarge | 10,906 | 13,840 | 21% faster |
| G5.48xlarge | 9,754 | 11,660 | 17% faster |

**TABLE 1** Runtime comparison for CryoSPARC lanes

While a specific time advantage is key in understanding how a filesystem can accelerate the use of a CryoSPARC pipeline, this performance can also allow for cost optimizations on the compute side as well. The formula is simple: cost of the GPU instance expressed in $/hr multiplied by the length of time for the queue to run gives you the cost per queue on each compute instance. Because the GPU instance cost is a constant, the calculation shows that any performance advantage becomes a direct cost advantage at the same percentage. Table 2 and figure 7 show the comparison in cost per queue between WEKA and FSx Lustre.

| GPU Instance Type | Linux on demand $/hr | Linux 1yr reserved $/hr | WEKA Time | FSx Lustre Time | WEKA Advantage |
|---|---|---|---|---|---|
| G5.4xlarge | $1.624 | $1.023 | 24,787 | 30,340 | 19% less cost |
| P3.8xlarge | $12.24 | $8.354 | 15,377 | 22,019 | 30% less cost |
| G5.24xlarge | $8.144 | $5.130 | 10,906 | 13,840 | 21% less cost |
| G5.48xlarge | $16.288 | $10.261 | 9,754 | 11,660 | 17% less cost |

**TABLE 2** Cost advantage of WEKA per compute queue

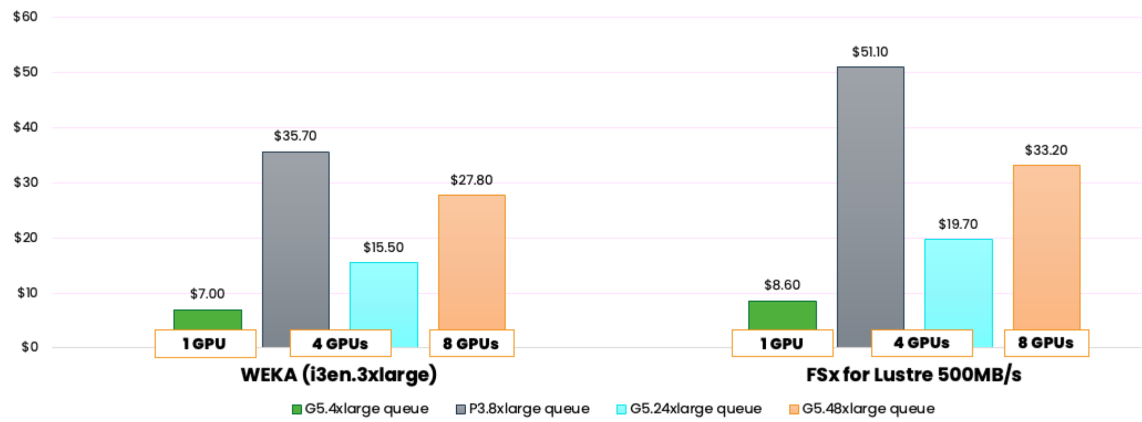


**FIG. 7** Figure 7: cost per queue of various GPU instances (1yr reserve cost)

Figures 6 and 7 in particular highlight an interesting point: While WEKA delivers a speed advantage by accelerating how fast each GPU can process the pipeline, There is also the question of cost of the GPU instances themselves compared to the performance they deliver. For example, a G5.48xlarge instance can complete a queue 2.5x faster than a G5.4xlarge instance, but the G5.48xlarge is 3.9x the cost of the G5.4xlarge. The question then becomes: does your organization value speed over cost, or is there a balance to be made? Depending on operational workflow, how often the analysis needs to get done, number of concurrent users, utilization rates of the Cryo-EM machine and more will determine whether it is better to have a few faster instances or broaden out to more of the slower ones. WEKA in all cases helps make it faster and more cost effective compared to FSx Lustre.

## Storage costs

Accelleration of the compute layer can realize significant cost savings, but there is also the ongoing cost of the data storage layer as well. In this case, WEKA does a fantastic job of matching up to FSx Lustre for a comparable capacity point.

| Storage | Infrastructure cost per month | Software License Cost per month | Total Monthly Cost | WEKA Advantage |
|---|---|---|---|---|
| FSx-Luster, 500MB/sec/TB persistent, 35 TB uncompressed | $0.34 per GB * 35000GB= $11,900 per month | N/A. FSx licensing is based on capacity and performance only. | $11,900 | |
| WEKA 8x i3en.3xlarge, 35 TB filesystem, all flash, uncompressed | 1 year reserved I3en.3xlarge=$738 per month *8 instances = $5,904 per month | WEKA XCL flash capacity license: $2500/month | $8,404 | 30% less cost |

**TABLE 3** Cost advantage of all-flash WEKA

This cost comparison only applies to the environment that was used for the benchmark. As you grow to higher capacities, it may be advantageous to take less frequently accessed data and transparently tier it off to a lower cost object store while retaining a high performance flash environment. Table 4 shows what a 500TB environment might look like with both all flash and a tiered environment of 100TB of flash and 400TB of object store capacity. For FSx Lustre we've reduced the performance to 250MB/s/TB as the additional capacity will provide significant performance compared to the smaller capacity 500MB/s environment above. The 250MB/s/TB environment is less expensive as well. All pricing comes from the AWS online pricing calculators.

| Storage | Infrastructure cost per month | Software License Cost per month | Total Monthly Cost | WEKA Advantage |
|---|---|---|---|---|
| FSx-Luster, 250MB/sec/TB persistent, 500 TB uncompressed | $0.21 per GB * 500,000GB= $105,000 per month | N/A. FSx licensing is based on capacity and performance only. | $105,000 | |
| WEKA 77x i3en.3xlarge, 500 TB filesystem, all flash, uncompressed | 1 year reserved I3en.3xlarge=$738 per month *77 instances = $56,826 per month | WEKA XCL flash capacity license: $17,637/month | $74,463 | 30% less cost |
| WEKA 16x i3en.3xlarge, 500 TB filesystem, 100TB flash, 400TB object, uncompressed | 1 year reserved I3en.3xlarge=$738 per month *16 instances = $11,808 per month. 400TB AWS S3 standard object storage=$8,400 per month | WEKA XCL flash capacity license: $6333/month, WEKA DTO object capacity license: $1250/month | $27,791 | 75% less cost |

**TABLE 4** Cost advantage of Tiered WEKA

# 7. Conclusion

As modern genomics and life sciences work advances, the need for high resolution views into micro-level structures in proteins continues to rise. Analytic and visualization technologies such as CryoSPARC continue to evolve to meet the demands of these life sciences environments but can be significantly enhanced by the choice of the underlying filesystem and storage architecture.

Traditional and even some modern storage infrastructures are unsuitable for demanding life sciences applications due to an inability to handle multiple types of IO across the entire data pipeline that these applications drive. This slows down the time to results from the application, drives up complexity, increases cost, and limits scale. WEKA has architected a scalable and robust filesystem leveraging modern technologies that pushes the boundaries of innovation and performance.

The results show performance scalability across the entire CryoSPARC environment when deployed in AWS, resulting in significant savings in wall-clock time for processing multiple pipelines. It demonstrates that by using WEKA in the cloud to scale the data infrastructure, we allow the GPUs to take advantage and accelerate time to insights, while reducing costs.

# 8. Appendix

## References

| Product & Technology | |
|---|---|
| **Weka File System** | |
| WEKA overview | https://www.weka.io/how-it-works/ |
| WekaFS Architecture White Paper | https://www.weka.io/resources/white-papers/wekaio-architectural-whitepaper/ |
| **CryoSPARC** | |
| CryoSPARC overview | https://cryosparc.com/ |
| CryoSPARC user guide | https://guide.cryosparc.com/ |
| **EMPIAR Datasets** | |
| Cryo-EM images of cannibinoid receptor protein complex | https://www.ebi.ac.uk/empiar/EMPIAR-10288/ |
| **Partners** | |
| Clovertex | Clovertex manages petabytes of data and hundreds of environments in AWS to support drug discovery. With large scale HPC workloads running in production, Clovertex specializes in building scientific applications that utilize AWS services for life-sciences, from discovery research to clinical trials to manufacturing and supply chain. https://www.clovertex.com |

# 9. About WEKA

Weka offers a modern subscription software-based data platform delivering 10x+ performance and scale demanded by today's cloud and AI workloads. With the simplicity of NAS, the performance of SAN or DAS, and the scale of object storage, no more compromises between Simplicity, Speed, or Scale. Learn more at www.weka.io or follow us on Twitter @WekaIO.





WKA-355-01 07/2023