

Seagate and Platina unlock the value of data by reimagining active archive solutions

Solution Summary

Customers can compose and manage active archives in a cost-effective manner using open source software stacks, such as Platina-supported community Ceph and Kubernetes, and low-cost industry-standard servers, Seagate storage enclosures, and Platina Cluster Heads, all managed from a centralized software controller, Platina Command Center (PCC). Ceph storage clusters can be tuned to meet very stringent requirements for scalability and data robustness.

Benefits Summary

- Dramatically simpler deployment, management, and monitoring of Ceph clusters, starting from bare metal.
- Performance with minimal effort for media applications with fast and easy retrieval of large blocks of video
- Archival data migration to a Ceph-based object storage platform, duplication and migration costs can be eliminated.
- All static data available “online” and accessible within milliseconds
- Cloud storage reserved for aged out data and disaster recovery efforts, minimizing costs

Using Seagate’s high-capacity storage enclosures and Platina’s composable infrastructure platform for Bare Metal, Kubernetes and Ceph object storage, enterprises can quickly unlock the value in their data archives by migrating to a Distributed Active Data Archive.

Archives have historically been tape-based solutions because of their size (10s of PB) and cost, but the recent trend is to store archival data sets on HDD-based platforms, particularly in the public cloud. The continued growth of such archives brings several distinct challenges for customers.

Tape-based archives are fragile and need periodic technology updates, which take several years and only provide static, or offline, storage. Archives residing in public clouds can lead to punitive data egress charges that can be 10x the data storage costs. And on-premise OEM storage solutions have untenable costs at multi-petabyte scale.

Introduction

A Distributed Active Data Archive, as envisioned by Platina and Seagate, is a private cloud that can be spread across multiple physical locations to

provide cost-effective, low-latency and high-capacity data storage services to a variety of applications.

The Challenge

Archive content can reside on a variety of storage media and platforms, starting with tape libraries at customer premises or in tape repositories at companies like Iron Mountain, but that is not the only location for stranded content. Content that has already been digitized may have been moved to a public cloud or placed on expensive customer-owned OEM storage. All these alternatives present their own business challenges.

There is great fragility in the tape ecosystem today. For example, all LTO tape is now manufactured by just two companies, Sony, and Fuji. In 2019, Sony and Fuji were embroiled in a patent lawsuit that resulted in a blockage of import of all LTO-8 tape into the US from both manufacturers. Fortunately, the lawsuit has been settled, but until this

resolution occurred, it created a shortage of tape and forced users to utilize the older and lower-capacity LTO-7 generation. Likewise, on the equipment side, with the exit of Oracle from the market, there is only one remaining major manufacturer of tape library platforms. Taking these developments into consideration, the future of tape seems uncertain.

As regards public cloud options, although companies have investigated them for active archives, most have learned that it is unacceptable due to storage costs and punitive egress bandwidth charges.

Finally, the cost structure of commercially available Ceph from vendors makes it cost-prohibitive for large archives. And branded storage from OEMs, such as NetApp, Dell EMC, and HPE, is also untenable for similar reasons.

Solution Approach

Tape storage, so long depended upon for archives, needs to be replaced with a cost-effective, simple-to-deploy solution that can offer a dependable method to store large amounts of data for the long term. Importantly, that data needs to be instantly accessible for active use by consumers and users—something that is unachievable using tape systems.

Customers are now evaluating private cloud architectures where they can

store and control their own data and do so at a much lower cost. Furthermore, many of these organizations have a mission critical need for a highly available active archive that must maintain long term data integrity over tens of petabytes (PB) of data. Because of this requirement, a Ceph-based software defined storage platform is an ideal solution to replace a tape archive or other stranded media content.

Seagate Enabling Technology

A key component of this software-defined storage solution is the Seagate® Exos® E 4U106 JBOD (just a bunch of disks) high density disk array, which supports a variety of Seagate Exos enterprise SAS HDDs, including the industry-leading Seagate Exos 16TB nearline HDD,

as well as Seagate Nytro® SAS flash SSDs. The Seagate Exos E 4U106 combines the highest-density packing of storage in a 4U form factor with the highest-capacity HDD devices, providing the leading storage capacity of any disk array on the market.



Exos E 4U106 is the data center's largest storage building block. The JBOD is suited for data storage requirements that are expected to grow rapidly. It delivers industry-first capacity and density without sacrificing data access speed. A vast amount of data fits in a 4U rack by leveraging up to 106 high-capacity Seagate

nearline enterprise SAS HDDs in a single enclosure, which holds up to an unprecedented 1.7PB¹ of raw capacity. As both the designer and manufacturer of the HDDs and the storage platform, Seagate can optimize the system design to allow full performance of all 106 disks utilizing patented cooling and system technology.

Partner Solution

Platina's key value proposition is a dramatic simplification of the provisioning, management, and monitoring of Distributed Active Data Archives. Platina technology has two key elements:

- An innovative platform the company has created—the Cluster Head—that combines the functionality of a top-of-rack switch with that of a traditional server and can be used to compose scalable compute and storage clusters; and
- The Platina Command Center (PCC) software controller, which can be used to manage Distributed Active Data Archives.

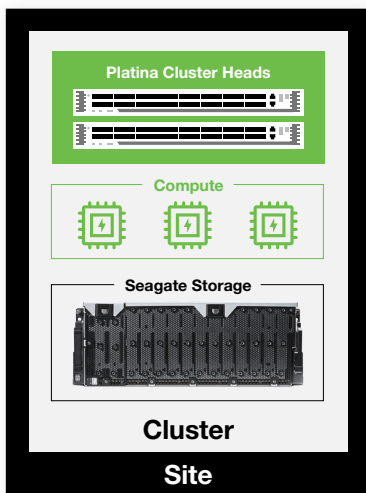
When operating in the public cloud, the management of all infrastructure and services is taken care of by the cloud operator. Platina's solution is purpose-designed to eliminate the repetitive, time-consuming, and expensive DIY procedures customers must endure to run their archives in non-public cloud environments.

Platina's Cluster Heads and PCC software controller are an integral part of a Distributed Active Data Archive built using

Ceph. PCC dramatically simplifies the deployment, management, and monitoring of Ceph clusters, starting from bare metal. There is extensive use of policies to guide repeatable actions, prune repetitive and error prone manual steps, and simplify the user's experience.

The Platina Seagate solution provides for automated bare metal deployment, which saves substantial time and effort in bringing up clusters, modifying them, or iterating on deployment configurations. This includes installation of the operating system, and management and monitoring of the hardware when operational. It also provisions core networking, storage, and system software services with point-and-click actions.

The complex software stacks of Ceph have also been collapsed by PCC into a small number of clicks in the GUI. Many defaults for common configurations are computed behind the scenes and presented as pre-populated values in the GUI rather than having to be entered manually. The result is a rapid and repeatable Ceph deployment experience with all data entry limited to one GUI screen.



¹ Using Exos 16TB HDDs

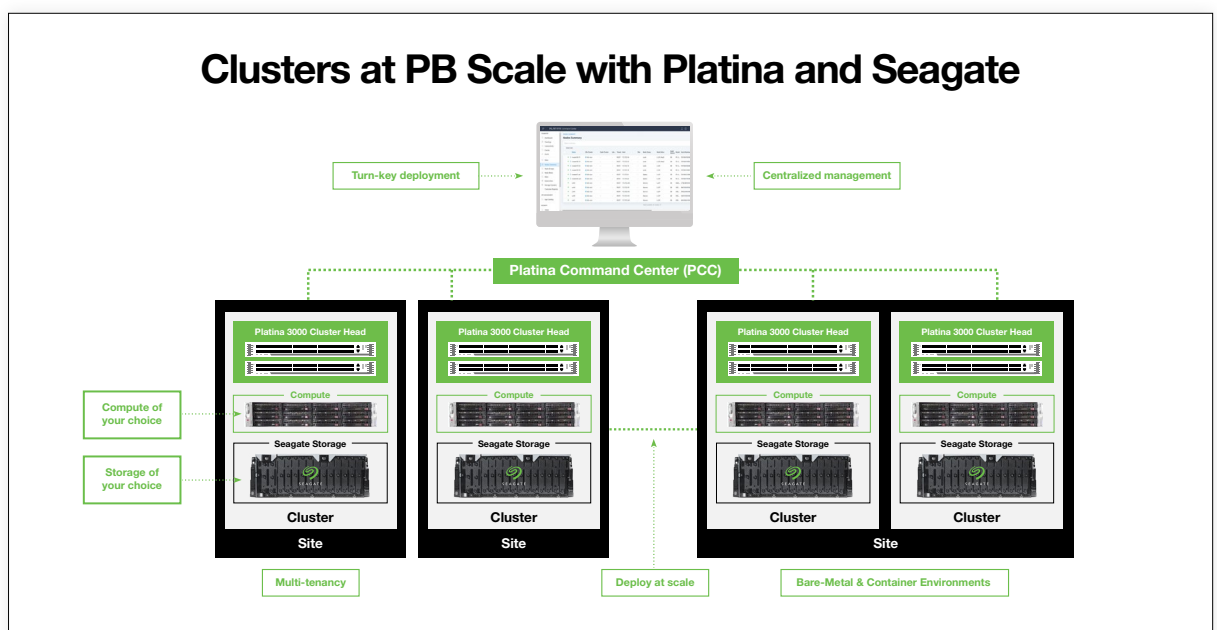


Benefits

- PCC dramatically simplifies the deployment, management, and monitoring of Ceph clusters, starting from bare metal. There is extensive use of policies to guide repeatable actions and simplify the user's experience.
- Performance with minimal effort for media applications—When using an active archive system for media content, the retrieval of large blocks of video becomes both easy and fast. This means that edited videos are available in a much faster time frame and production costs are lowered due to efficiency in editing.
- By migrating archival data to a Ceph-based object storage platform, the overhead of repeatedly duplicating and migrating tape content can be eliminated.
- All static data becomes available “online” and can be accessed with millisecond latencies versus tens of minutes from a tape library, or an hour or more when the data is held on tape offsite.
- Data can continue to be aged out to cloud storage strictly as an insurance policy for extraordinary disaster recovery, but otherwise not used, thus minimizing costs.

Features

- Automatically discover and inventory the equipment used in a cluster.
- Create and validate all cluster-wide network configurations and topologies.
- Provision all server and storage hardware and software components.
- Configure the servers, storage, and networking gear with system software and applications, including, but not limited to:
 - Operating systems of user's choice,
 - Core services (name resolution, address management, network routing, drive naming and grouping, block storage, file storage, object storage, security, etc.),
 - Application deployment environments (e.g., containers via Kubernetes, virtual machines, related tools, etc.),
 - Customer applications and application specific components.
 - Add-on packages
- Manage and monitor all deployed infrastructure, services, and applications.



In Conclusion

Although tapes are very cost-effective for fully passive archives, the tape ecosystem is now extremely fragile, and an active archive is required for online data use. Public cloud suffers from slow retrieval times and egress bandwidth charges that are layered on top of the storage costs. The cost structure of commercially available Ceph from vendors makes it cost-prohibitive for large archives. And branded storage from OEMs, such as NetApp, Dell EMC, and HPE, etc., is also untenable for similar reasons.

Platina and Seagate's Distributed Active Data Archive solution enables customers to migrate from a tape-based or public cloud media archive to an online Ceph active archive in a very cost-effective manner. By providing sophisticated networking and infrastructure capabilities, core software services, and application provisioning, management, and monitoring capabilities, it can be implemented quickly and administered by a small operations team, making your data actively available.

Take your passive archive to the next generation with Platina and Seagate's Distributed Active Data Archive solution and make your data readily available, ensuring that you have a scalable long-term solution today.

For more information on Seagate, visit www.seagate.com

For more information on Platina, visit www.platinasystems.com

seagate.com

© 2020 Seagate Technology LLC. All rights reserved. Seagate, Seagate Technology, and the Spiral logo are registered trademarks of Seagate Technology LLC in the United States and/or other countries. Exos, the Exos logo, Nytro, and the Nytro logo are either trademarks or registered trademarks of Seagate Technology LLC or one of its affiliated companies in the United States and/or other countries. All other trademarks or registered trademarks are the property of their respective owners. When referring to drive capacity, one gigabyte, or GB, equals one billion bytes and one terabyte, or TB, equals one trillion bytes. Your computer's operating system may use a different standard of measurement and report a lower capacity. In addition, some of the listed capacity is used for formatting and other functions, and thus will not be available for data storage. Seagate reserves the right to change, without notice, product offerings or specifications. SC739.1-2006US June 2020

