

ActiveScale™ Family Configuration Guide

Introduction

The world is awash in a sea of data. Unstructured data from our mobile devices, emails, social media, clickstreams, log files, sensor data, video, audio, and more is clogging our data centers. Finding a scalable and economic way to store and protect this data is a real challenge. When this data is correctly stored and accessed, it can provide deeper insights into business operations, market conditions, and customer behaviors. Insights based on data can provide a competitive advantage. Historical data can provide the rich history needed by advanced analytics or to take advantage of new technology by analyzing old data. Traditional storage architectures don't meet these needs and looking for a way to implement a "data forever" strategy can seem daunting. This data is invaluable, and needs to be protected with a robust system with high data durability and availability.

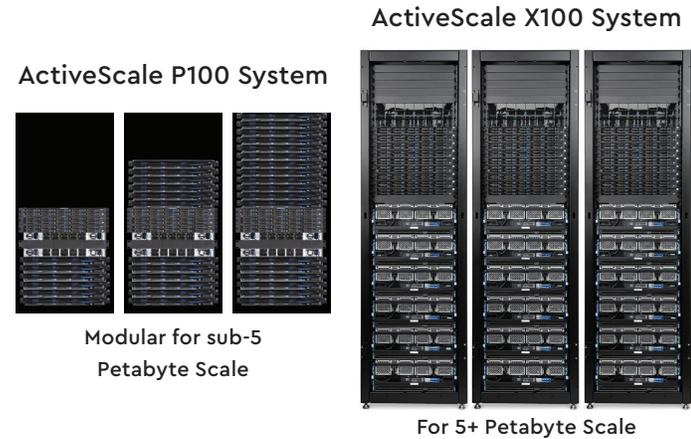


Figure 1: ActiveScale Family of Object Storage

The ActiveScale family is designed to meet your unstructured data needs by providing massive scalability, robust fault tolerant capabilities, with high data durability and system availability. The ActiveScale system is available on two different platforms: P-series modular solution for flexible configurations and X-series for a highly efficient integrated solution. ActiveScale Cloud Manager (CM) provides cloud-based system analytics for Active Archive or ActiveScale systems with historical trending, multi-system monitoring, and SLA friendly metrics to better manage your business.

Object storage is the right platform for unstructured data for several reasons. Traditional storage is based on RAID data protection. As disk drives grow beyond 6TB, disk rebuild times become so long that the opportunity for a second failure and data loss becomes a real concern. Additionally, NAS systems tend to bog down beyond 7-8 racks which limit their scale. Finally, the performance/price/capacity design points of object storage are aimed at economic, high-scale unstructured data, whereas block and file systems are more focused on high performance and are less sensitive to cost.

Not all object storage systems are created equal, and ActiveScale is differentiated through patented technologies such as BitSpread® and BitDynamics™. BitSpread provides dynamic data placement of objects, creating a system that tolerates hardware problems and easily incorporates new capacity without rebalancing. BitDynamics performs background data integrity checks and self-healing for outstanding data integrity. These two technologies define the answer to customers who need long-term data storage that is economic and highly scalable with extreme data durability.

When deciding which solution meets your needs, consider the ActiveScale P100 System modular object storage for needs under 5PB and where cost of entry is important. For larger needs, or very rapidly growing data, use the ActiveScale X100 System offers our lowest cost per capacity and greatest scalability. Both P and X series use the same operating system, so the features on either system will be the same. Both systems support scale-up and scale-out growth, but due to the hardware configurations, they grow differently.

The ActiveScale operating system is based on advanced erasure coding object storage that can provide up to 19 nines of data durability for your demanding environments and long retention data. Our BitDynamics technology provides background data integrity verification and self-healing, also an important characteristic for long-duration data. ActiveScale also features our patented BitSpread feature which provides dynamic data placement for non-disruptive upgrades and fault handling. Other capabilities include end-to-end encryption for sensitive data, object versioning for rapid recovery from accidental deletions and ransomware attacks, and strong consistency for always-fresh data.



Figure 2: ActiveScale P100 Storage Node

ActiveScale P100 System Scale-up Configurations

Scale-up is traditionally a way to add more capacity behind existing system nodes. The P100 scales-up in increments of 720 or 864TB based on the choice of 10 or 12TB HelioSeal® HDDs and that are packaged as six 1RU storage nodes. The 10TB HDD-based systems allow you to start with a minimum of 720TB and scale-up to 2.1PB or start with 12TB HDDs at 864TB and scale-up to 2.5PB which is managed as a single namespace. If you start with a 10TB HDD-based system you can scale up with 12TB HDDs, for extra flexibility.

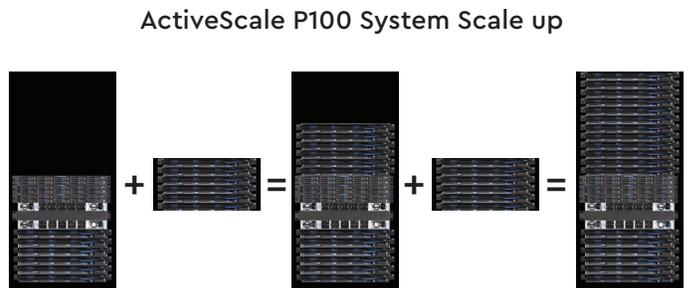


Figure 3: ActiveScale P100 Scale-up Configurations

Scale-out Configurations

Scale-out is defined as adding additional system nodes to support new capacity. The ActiveScale P100 scales-out to three units in a single location with up to 7PB of raw capacity per site managed as a single namespace. If the units are in the same location, the object chunks will all be local, and all three can comprise a single namespace for ease of management. If a multiple location configuration is desired, the object chunks are spread across three locations in such a way that if one of the locations becomes unavailable, all the objects are still available and strongly consistent. Maximum raw capacity for a multiple location configuration is 23PB. Similarly to scale-up capability, you can start with a 10TB HDD-based system and scale-out with 12TB HDDs for flexibility.

There are different erasure encoding options for single and multiple location configurations, and the encoding needs to be configured before the first system is installed. It is not possible to take an existing single location P100 and make it a multiple location configuration, so if the ultimate intent is to have a dispersed configuration, work with your ActiveScale team to make sure it is configured correctly.

ActiveScale P100 One Location System Scale-out

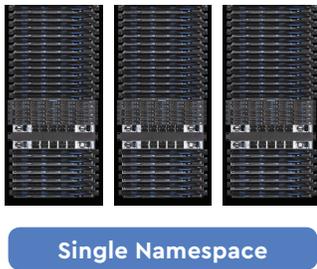


Figure 4: ActiveScale P100 Scale-out Configuration: One Location

ActiveScale P100 Three Location System Scale-out

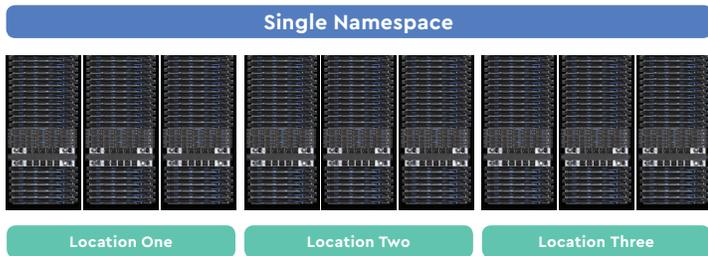


Figure 5: ActiveScale P100 Scale-out Configuration: Three Locations

Performance Configurations

One of the highlights of a modular architecture is the ability to optimize for performance or capacity. If you have a need for more performance, scaling-out to include more storage nodes with a minimum of capacity (720TB) for maximum throughput.



Figure 6: ActiveScale P100 Performance Configuration

Capacity Configurations

If you want a capacity-optimized solution, then you'll want to scale-up the P100 before adding more system nodes to keep the cost/TB low. If you're looking for the lowest cost/TB, the ActiveScale X100 might be the right solution if you need 5PB or more of capacity.



Figure 7: ActiveScale P100 Capacity Configuration

ActiveScale X100 System Scale-up Configurations

Growing the X100 is a little different since it is a rack-integrated system, unlike the P100's modular architecture. Optionally, the X100 may be ordered as the elements without a rack, if desired. The X100 storage has 6 storage shelves with 7 sleds in each shelf. Each sled has 14 of either 10TB or 12TB Western Digital Heliioseal HDDs. ActiveScale X100 can scale from 840TB to 5.88PB in a single instance or 1008TB to 7PB (raw capacities) with the 12TB HDDs. To scale-up you replace unpopulated sleds with populated sleds. Each shelf must have the same sled population in each of the positions. For instance, at sled position C shown below, each shelf's C position sled must be the same configuration.



Figure 8: ActiveScale X100 Storage Sled in Shelf



Figure 9: ActiveScale X100 HDD in Storage Sled

Scale-out Configurations

Scale-out with the X100 means adding more racks. A scale-out of up to nine racks in the same location means the object chunks will all be local, and they can be managed as a single namespace. If a multiple location configuration is desired, the object chunks are spread across the three locations with up to 9 racks in such a way that if one of the locations becomes unavailable, all the objects are still available in a strongly consistent manner. Maximum scale-out capacity is 63PB raw capacity.

There are different erasure encoding options for single and multiple location configurations, and the encoding needs to be configured before the first system is installed. It is not possible to take an existing one location X100 and make it multiple location system, so if the ultimate intent is to have a multiple location configuration, work with your ActiveScale team to make sure it is configured correctly.

ActiveScale X100 One Location System Scale-out

Single Namespace



One Location

Figure 10: ActiveScale X100 Scale-out Configuration: One Location

ActiveScale X100 Three Location System Scale-out

Single Namespace



Location One

Location Two

Location Three

Figure 11: ActiveScale X100 Scale-out Configuration: Three Locations

System Availability

ActiveScale takes great care with data durability to assure that data read is data as written with technology that performs background integrity checks and self-healing. System availability, however, is different from data durability. If there is an earthquake, storm, or widespread power outage, an entire data center site could become unavailable. System availability speaks to the ability to tolerate this level of outage and continue operations from the remaining sites. There are two system-level approaches to tolerating this kind of outage: replication and geo-spreading.

Geo-Spreading

Geo-spreading with ActiveScale is an object storage capability that allows an object to be spread across three different locations in such a way that if one location is unavailable, the object is still available at the two other locations. This is unlike traditional architectures where 2-site replication means the data would be copied in both locations to respond to system unavailability. The system in a geo-spread configuration is comprised of multiple racks (instances) managed as a single namespace or failure domain. The system takes an object and breaks it into chunks of data and parity and spreads those across the three locations with enough parity to rebuild the objects even if one location (and the third of the chunks on it) are unavailable. In ActiveScale, this is done with strong consistency, so there is no chance of getting stale data.

In a 3-geo configuration, if one site becomes unavailable the data will still be available in the two available sites and new objects will be written to the two available sites. When the formerly unavailable third site becomes available, data will again be spread across all three sites. Management of a 3-geo configuration is very efficient since it is a single namespace and is managed as a single system with multiple instances from a single pane of glass.

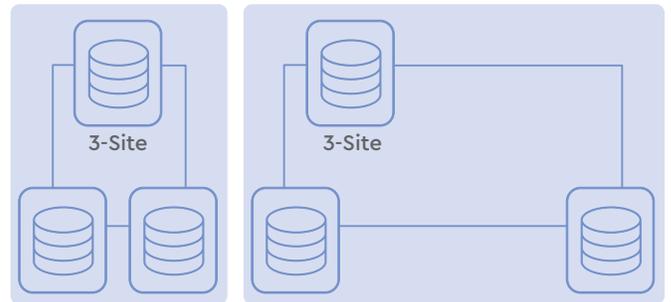


Figure 12: ActiveScale Geo-Spread Configurations Form a Single Namespace

Performance of the multi-site system will be affected by distance, since it is a strongly consistent system. As distances increase, latency in the system will increase. The link among systems will be the customer's WAN, so understanding the available bandwidth is an important part of system performance. In a well-implemented system you might estimate the latency to be about 1-2 ms per 100km of distance.

Because ActiveScale geo-spreading has strong consistency, it is synchronous and you will not get stale data and there is no RPO (recovery point objective) impact. By contrast, 2-site asynchronous replication provides the alternative system availability solution for long distances that might make geo-scaling inappropriate.

Replication to ActiveScale

ActiveScale also provides the option of 2-site asynchronous replication to provide a different way to preserve system availability. Two-site replication operates at the bucket or instance level for customers where a geo-spread configuration is not appropriate, perhaps due to long distances that would impact system performance. Asynchronous replication does not need to complete both primary and secondary site writes before acknowledging to the application, so performance is not decremented by long distance. Because replication is asynchronous, it also means there is an RPO impact since data in flight may be lost when the primary site becomes unavailable. This is also true with block and file storage asynchronous replication, so most data centers understand how to manage an asynchronous replication environment.

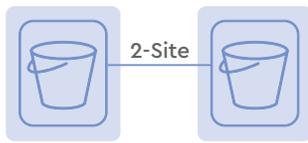


Figure 13: 2-Site Asynchronous Replication is Comprised of Two Namespaces

Additionally, asynchronous replication is comprised of two namespaces. The logic for this is that long distances would slow down the object chunk data and parity handling which creates additional WAN traffic and during an object rebuilding chunks may be required from the other location, significantly impacting performance. This does mean that the two locations are managed as two separate systems.

ActiveScale replication also has a two-way replication option. Replication is configured when the bucket is set up. To enable two-way replication buckets in both sites will need to set up with replication. This ensures that changes made in one site will be reflected in the other. The traffic is two-way between the sites, but not full duplex. Again, review the WAN capacity for sufficient bandwidth when enabling 2-site asynchronous replication. Control commands are sent out of band to reduce the overhead. Finally, 2-way asynchronous replication and 3-geo scale out are not mutually exclusive, you can have them work together if you have buckets you want to share between two 3-geo namespaces. You will need to understand the WAN implications as always, but we suggest you work with your solution architect to assure the solution configuration will meet the application needs.

Replication to Amazon® AWS™

In addition to 2-site asynchronous replication to other ActiveScale systems, replication to a bucket in Amazon AWS is also supported. This provides yet another way to provide high system availability by having data available on-site and off-site. Replication to AWS is one-way replication so that changes to the objects in the ActiveScale bucket will automatically update the AWS bucket. Changes to objects in the AWS bucket will need to be manually replicated back to the ActiveScale, if desired.



Figure 14: Hybrid Cloud Replication to AWS

In addition to using hybrid cloud replication as another disaster recovery tool, another interesting use case is analytics on AWS. Data replication to AWS for analytic processing allows you to take advantage of the compute power and analytic toolsets in AWS. Replicate the data to AWS, perform the analytics, send the results back to ActiveScale, and delete the data in AWS to reduce the cost of storing the data in AWS. You don't need to export the data back to ActiveScale, it is already there, and you save the export expense.

The same WAN considerations as in 2-site asynchronous replication apply to hybrid cloud replication, and you should work with your solution architect to assure the solution configuration will meet the application needs.

Management

Managing petabyte-scale storage can be a challenge. In order to make the storage administrator as efficient as possible, the right tools are essential. ActiveScale has developed a two-tier strategy. The first tier is operational where ActiveScale System Manager (SM) handles the day-to-day operations such as provisioning new storage and users, identifying hardware failures and performance problems. ActiveScale SM is a graphical tool that allows the storage administrator to manage the entire namespace, even if it includes multiple instances of the system. This means that from a single pane of glass the administrator can identify problems anywhere in the namespace, and then drill down to find individual component problems that need resolution.

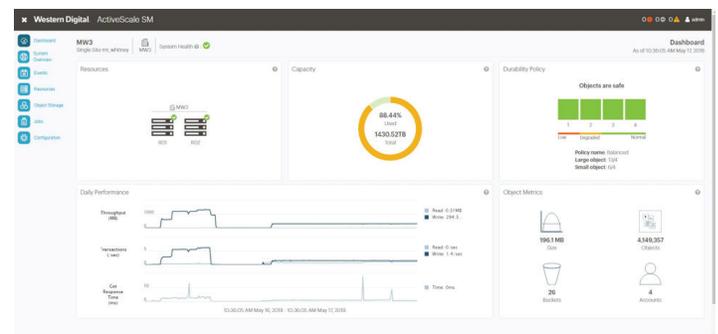


Figure 15: ActiveScale System Management

The second tier of management is ActiveScale Cloud Management (CM). ActiveScale CM is the more strategic tool, and can see all the Active Archive and ActiveScale systems that an enterprise has installed. It is a cloud-based system analytics tools that can be accessed from anywhere. Data is continually gathered and is uploaded daily from ActiveScale and Active Archive systems to the ActiveScale CM cloud. ActiveScale CM allows the administrator to better manage their service

level agreements by seeing who the active users are, which are the most active buckets, and by using the historical data, compare July of last year to July of this year to see if there is a onetime problem, or a trend that needs to be addressed differently. Additionally, there is a predictive element that will forecast based on historical capacity growth, and a report shows when additional storage may need to be ordered.

There are no additional charges for ActiveScale CM or SM, they are included with the ActiveScale P100 and X100 systems.

Conclusion

ActiveScale facilitates your development of a "data forever" storage architecture. Scalability is available as both scale-up and scale-out features, as well as a multiple location configuration for high system availability. ActiveScale boasts a flexible architecture due to our patented dynamic data placement that adapts to many hardware problems and easy adoption of new capacity, unlike static data placement found in other object storage systems. Finally, ActiveScale's BitDynamics background data integrity checks with self-healing provide peace of mind for long-duration data storage. For more information go to www.wdc.com/dc-systems

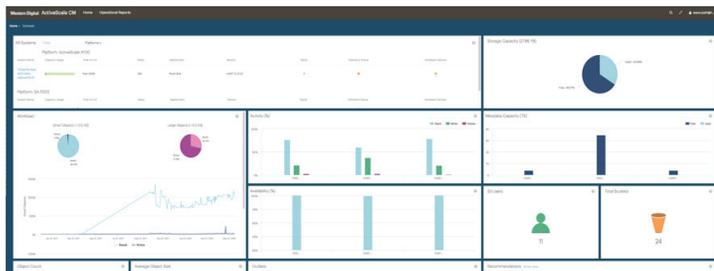


Figure 16: ActiveScale Cloud Management

Western Digital.

5601 Great Oaks Parkway
 San Jose, CA 95119, USA
US (Toll-Free): 800.801.4618
International: 408.717.6000

www.westerndigital.com

© 2018 Western Digital Corporation or its affiliates. All rights reserved. Western Digital, the Western Digital logo, ActiveScale, BitSpread, BitDynamics, and Helioseal are registered trademarks or trademarks of Western Digital Corporation or its affiliates. Amazon, AWS, and Amazon S3 are trademarks of Amazon.com, Inc. or its affiliates in the United States and/or other countries. All other marks are the property of their respective owners.