Disks and aggregates

ONTAP 9

NetApp
July 15, 2022

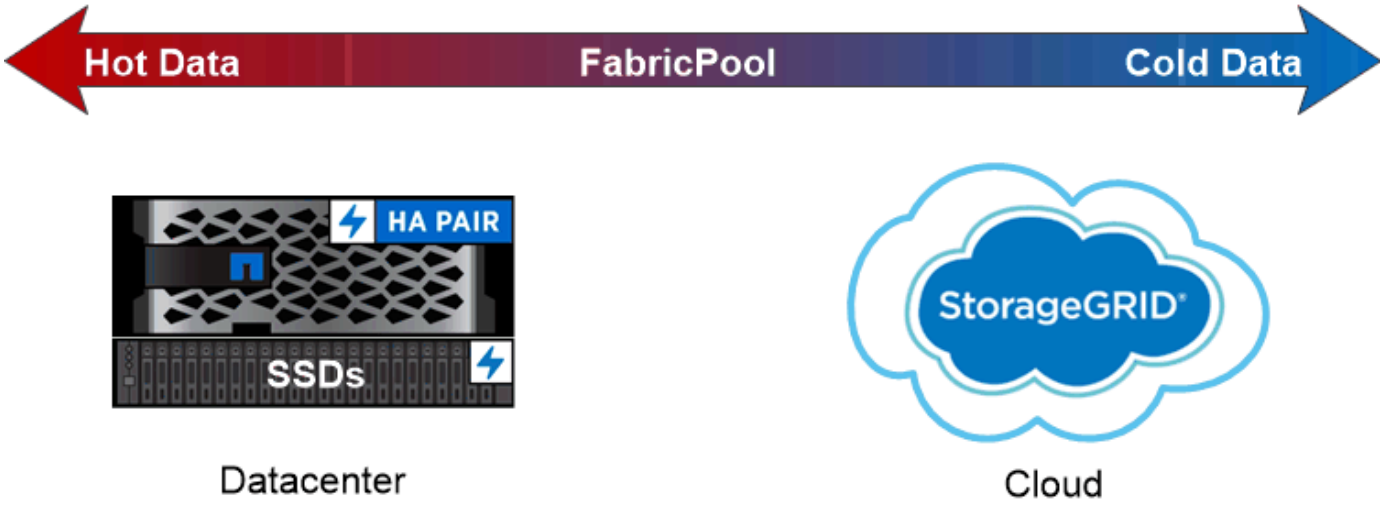
Table of Contents

Disks and aggregates

Disks and aggregates overview

Aggregates are containers for the disks managed by a node. You can use aggregates to isolate workloads with different performance demands, to tier data with different access patterns, or to segregate data for regulatory purposes.

- For business-critical applications that need the lowest possible latency and the highest possible performance, you might create an aggregate consisting entirely of SSDs.
- To tier data with different access patterns, you can create a *hybrid aggregate,* deploying flash as high-performance cache for a working data set, while using lower-cost HDDs or object storage for less frequently accessed data. A *FlashPool* consists of both SSDs and HDDs. A *FabricPool* consists of an all-SSD aggregate with an attached object store.
- If you need to segregate archived data from active data for regulatory purposes, you can use an aggregate consisting of capacity HDDs, or a combination of performance and capacity HDDs.



You can use a FabricPool to tier data with different access patterns, deploying SSDs for frequently accessed "hot" data and object storage for rarely accessed "cold" data.

Aggregates and RAID groups

Modern RAID technologies protect against disk failure by rebuilding a failed disk's data on a spare disk. The system compares index information on a "parity disk" with data on the remaining healthy disks to reconstruct the missing data, all without downtime or a significant performance cost.

An aggregate consists of one or more *RAID groups.* The *RAID type* of the aggregate determines the number of parity disks in the RAID group and the number of simultaneous disk failures the RAID configuration protects against.

The default RAID type, RAID-DP (RAID-double parity), requires two parity disks per RAID group and protects against data loss in the event of two disks failing at the same time. For RAID-DP, the recommended RAID group size is between 12 and 20 HDDs and between 20 and 28 SSDs.

You can spread out the overhead cost of parity disks by creating RAID groups at the higher end of the sizing recommendation. This is especially the case for SSDs, which are much more reliable than capacity drives. For HDD aggregates, you should balance the need to maximize disk storage against countervailing factors like the longer rebuild time required for larger RAID groups.

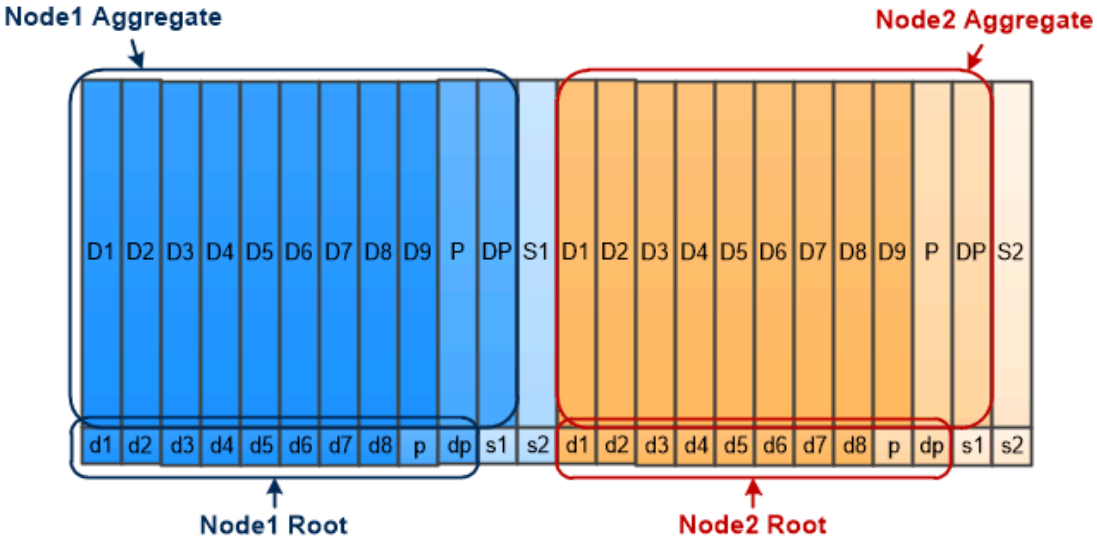
Root-data partitioning

Every node must have a root aggregate for storage system configuration files. The root aggregate has the RAID type of the data aggregate.

System Manager does not support root-data or root-data-data partitioning.

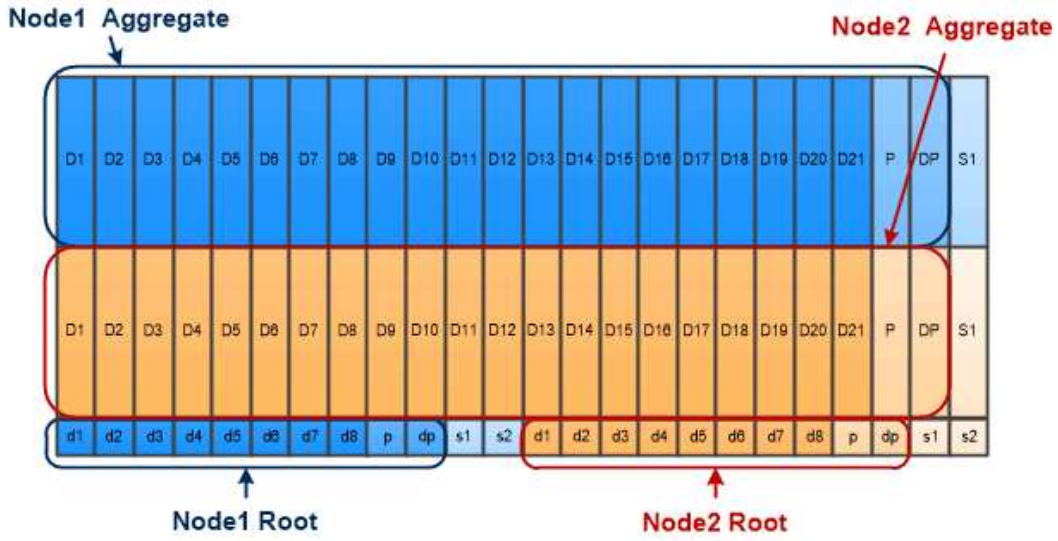
A root aggregate of type RAID-DP typically consists of one data disk and two parity disks. That's a significant "parity tax" to pay for storage system files, when the system is already reserving two disks as parity disks for each RAID group in the aggregate.

Root-data partitioning reduces the parity tax by apportioning the root aggregate across disk partitions, reserving one small partition on each disk as the root partition and one large partition for data.



Root-data partitioning creates one small partition on each disk as the root partition and one large partition on each disk for data.

As the illustration suggests, the more disks used to store the root aggregate, the smaller the root partition. That's also the case for a form of root-data partitioning called *root-data-data partitioning*, which creates one small partition as the root partition and two larger, equally sized partitions for data.



Root-data-data partitioning creates one small partition as the root partition and two larger, equally sized partitions for data.

Both types of root-data partitioning are part of the ONTAP *Advanced Drive Partitioning (ADP)* feature. Both are configured at the factory: root-data partitioning for entry-level FAS2xxx, FAS9000, FAS8200, FAS80xx and AFF systems, root-data-data partitioning for AFF systems only.

Drives partitioned and used for the root aggregate

The drives that are partitioned for use in the root aggregate depend on the system configuration. Knowing how many drives are used for the root aggregate helps you to determine how much of the drives' capacity is reserved for the root partition, and how much is available for use in a data aggregate.

The root-data partitioning capability is supported for entry-level platforms, All Flash FAS platforms, and FAS platforms with only SSDs attached.

For entry-level platforms, only the internal drives are partitioned.

For All Flash FAS platforms and FAS platforms with only SSDs attached, all drives that are attached to the controller when the system is initialized are partitioned, up to a limit of 24 per node. Drives that are added after system configuration are not partitioned.