



Technical FAQ

Advanced Drive Partitioning Technical FAQ

James Wright, NetApp

May 2016

Overview

This document addresses the most frequently asked question about ONTAP's Advanced Drive Partitioning functionality which includes Entry internal HDD partitioning, SSD Flash Pool partitioning, and AFF root-data sharing Use Case.

TABLE OF CONTENTS

1 Terminology	3
2 Advanced Drive Partitioning Overview	3
Why Advanced Drive Partitioning?	3
What is Advanced Drive Partitioning?	4
3 Root-data Partitioning for HDDs	4
What is Root-data partitioning for Entry with Internal HDD and RAID DP Use Case in ONTAP 8.3.x?	4
What is RAID-TEC?	5
What is Root-data partitioning for Entry with Internal HDD and RAID-TEC in ONTAP 9.0?	6
What are the design goals of root-data (R-D) partitioning?	6
How do we configure an Active/Passive configuration on a FAS2520 with 8.3.x style root-data (R-D) sharing partitioning?	7
How do we configure an Active/Active configuration on a FAS2520 with root-data (R-D) sharing?	8
Can we add an external shelf to a system that is running Data ONTAP 8.3 root-data sharing?	9
What is the usable capacity profile for various drives using root-data sharing?	9
How do I upgrade my 8.2.x or 8.3.x Entry system with internal HDD whole drives to 8.3.x or ONTAP 9 to enable Advanced Drive Partitioning with root-data sharing?	9
4 Flash Pool SSD Partitioning	10
What is Flash Pool SSD Partitioning?	10
What is an SSD Storage Pool?	11
What is an Allocation Unit (AU)?	11
Is there a concept of ownership in an SSD Storage Pool?	12
Can I use system manager to provision cache using Flash Pool SSD Partitioning?	12
How do I provision SSD cache in a Flash Pool aggregate?	12
5 Root-data partitioning for SSDs (AFF)	13
What is 8.3.x style root-data (R-D) sharing for SSD?	13
What is 9.0.x style root-data-data (R-D2) sharing for SSDs and why is it better than 8.3.x (R-D) sharing with 1-shelf AFF 24-drive Active/Active configurations?	14
How is ONTAP 9 style root-data-data (R-D2) SSDs sharing different in a 2-shelf AFF 24-drive Active/Active configuration than 1-shelf?	15
How many SSDs can be partitioned with 8.3.x style root-data (R-D) sharing & ONTAP 9 root-data-data (R-D2) sharing in an AFF system; what is the least number of drives that are required for SSD root-data partitioning in 8.3.x and ONTAP 9.	16
6 Supported Configurations	17
What platforms are supported with Advanced Drive Partitioning?	17
What drives are supported with Advanced Drive Partitioning?	17
7 Additional Readings	18

1 Terminology

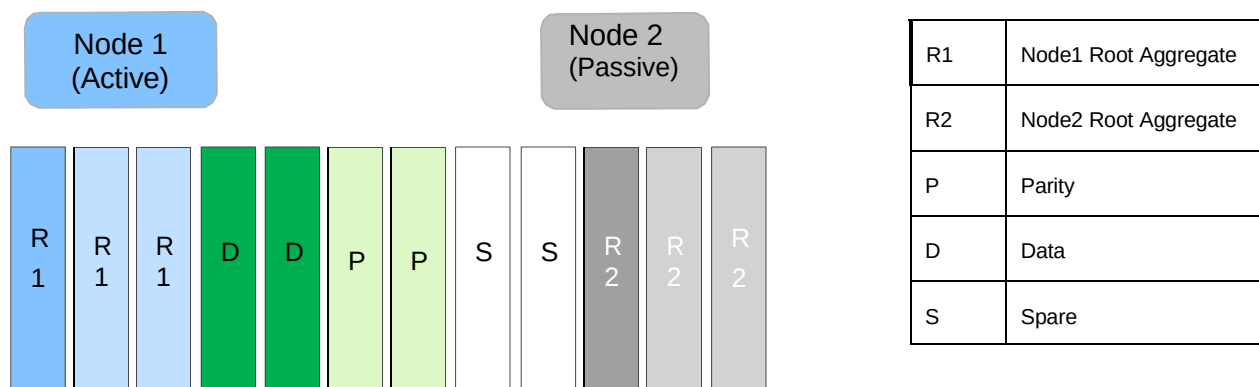
- **Physical disk:** Refers to a whole disk (non-partitioned) as we know it today.
- **Shared disk:** Refers to a partitioned disk shared between aggregates and a HA pair.
- **Disk Partition:** Disk Partition is a slice of a physical HDD or an SSD.
- **Root Partition:** Refers to a smaller partition used by root aggregate
- **Data Partition:** Refers to a larger partition used by data aggregate
- **Storage Pool:** Storage Pool is a collection of partitioned SSDs.
- **MUA:** Minimum Unit of Allocation that can be added to a Flash Pool aggregate.
- **Allocation Unit:** A Minimum Unit of Allocation (MUA) that can be added to a Flashpool aggregate.
- **RAID-TEC:** RAID Triple Erasure Coding or RAID Triple Parity is the use of 3 drives for RAID parity to improve storage resiliency.
- **AFF:** All-FAS-Flash Optimized 8xxx High platform with only SSD storage
- **FAS with SSD:** Non-Optimized Entry, Mid, High platform with only SSD storage

2 Advanced Drive Partitioning Overview

Why Advanced Drive Partitioning?

8.2.x and earlier cDOT HA configurations require dedicated root aggregate without user data for each controller. Even if root aggregates are configured as RAID-4 (which is not recommended), this results in significant storage efficiency challenges. Not only does this serve as a competitive disadvantage, but it was also a barrier to cDOT adoption.

FAS2520 HA pair, Active/Passive Configuration running cDOT (without drive partitioning)



Following is the configuration:

- 1) Two dedicated cDOT Root Aggregates:
 - a. Root Aggregate 1: 1D + 2P
 - b. Root Aggregate 2: 1D + 2P
- 2) One Data Aggregate:
 - a. Data Aggregate: 2D + 2P

The above configuration demonstrates the efficiency challenges faced in the entry space especially with

cDOT deployments. With two root aggregates requiring Raid DP and two drives for hot spare we're left just four drives to create the user data aggregate. Comparably, storage efficiency on 7-mode Active-Passive configuration is much better since a dedicated root aggregate is not required – although it is recommended. Still, there's storage squandering with dedicated root aggregate on the passive node, which is another hurdle in the extremely cost and capacity sensitive entry segment.

What is Advanced Drive Partitioning?

Advanced Drive Partitioning is a feature initially released in ONTAP 8.3 and further enhanced in ONTAP 9.0 that addresses storage efficiency challenges in the entry space, offers increased cache allocation provisioning flexibility, and reduces overall \$/GB across AFF and FAS with SSD platforms through drive slicing. It is divided into three use cases:

1. Root-data Partitioning for HDDs
2. Flash Pool SSD Partitioning
3. Root-data Partitioning for SSDs (AFF and FAS with SSD)

3 Root-data Partitioning for HDDs

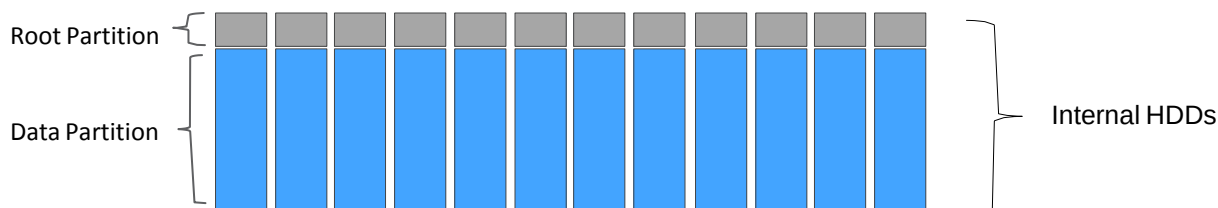
What is Root-data partitioning for Entry with Internal HDD and RAID DP Use Case in ONTAP 8.3.x?

Root-data partitioning or root-data (R-D) sharing for the Entry internal HDD use case in ONTAP 8.3.x addresses usable capacity challenges by significantly reducing root aggregate overhead. Sharing HDDs enables the root aggregate to use less space, leaving more room to provision data aggregates. Correspondingly storage utilization is improved for your system.

Each internal HDD is shared by two partitions:

- 1) **Root Partition:** Root partition is a smaller partition and it is used to create root aggregates.
- 2) **Data Partition:** Data partition is a larger partition and it is used to create data aggregates.

Figure 2: Root-data partition:



Root-data sharing also allows you to create Active-Active configuration in a 12 drive entry system, this was previously not possible with Clustered DATA ONTAP (in pre-8.3 versions) due to limited spindles and dedicated root aggregate requirement.

The primary reason for using a dedicated root aggregate is hosting user data volumes on the same aggregate as the root volume can increase the risk of the aggregate becoming inconsistent –more information on this topic can be found in [Storage Subsystem Technical FAQ](#)

Sharing HDDs into root and data partitions lets you manage spare capacity efficiently. Users can assign partitions for spare capacity instead of dedicating the whole drive. Especially in Active/Passive configurations where the passive node is not serving data and is only hosting the root aggregate. Since root aggregates are provisioned using the smaller root-partitions, you can assign smaller root partitions as spare partitions instead of needlessly using the whole drive - thus reducing spare capacity overhead.

Each partitioned HDD can be shared with up to two nodes and two aggregates (a data aggregate and a root aggregate) within a HA pair. There are some tradeoffs to sharing an HDD such as, lack of fault and performance isolation. For example, an internal drive failure can affect the root aggregate as well as the data aggregate. Therefore, RAID DP is enforced by design to increase resiliency and enable best practices. Moreover, you can have RAID DP on both the root aggregates as well data aggregates.

R-D sharing for Entry with internal HDDs in RAID DP is a compelling feature which makes NetApp Entry systems extremely competitive. However 8.3.x's method of provisioning was just the beginning, as less resilient SATA HDD (FSAS or MSATA) 6TB or greater require additional protections that a two parity drive scheme could not provide. Hence the need in ONTAP 9.0 to develop an improved Data Protection methodology, that could offer greater fault tolerance. RAID-TEC or RAID Triple Erasure Coding with triple parity implemented on R-D Entry with internal HDD systems would fill that void.

What is RAID-TEC?

RAID Triple Erasure Coding or RAID-TEC is NetApp software RAID technology that leverages triple parity to protect against data loss from the failure of up to three drives (or 2 drives and 1 media bit error, 1 drive and 2 media errors or 3 media bit errors) within a single RAID group (RG). RAID-TEC is an extension of RAID DP which uses 2 parity disks (dual failure protection) and its predecessor RAID-4 that has a single parity (1 drive protection) for drive reconstruction. Upgrades to RAID-TEC and downgrades back to RAID DP for Data aggregates can be performed non-disruptively. And with the resilience RAID-TEC offers one has the flexibility to increase RAID group sizes as needed which is an added benefit of upgrading since triple parity is not supported in ONTAP releases prior to 9.0. The Use Case targeted for RAID-TEC is large capacity SATA drives 6TB or greater however all current shipping drives are supported.

The diagram in Figure 3 below provides an overview of RAID-TEC features. Most notably, RAID-TEC is the default for SATA $\geq$ 6TB and a requirement for SATA $\geq$ 10TB.

Figure 3: RAID-TEC

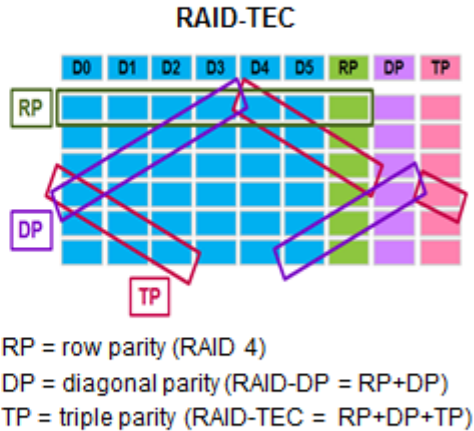
RAID-TEC Triple Parity Protection

ONTAP 9.0

- Reduced risk of data loss with very large NL-SAS HDDs requires more RAID protection
 - Smaller RAID-DP RG for 6 & 8T HDD

Support bigger RAID groups by adding a 3rd parity drive

- RAID-TEC RG NL-SAS HDDs ≥6T up to 1.5–2x bigger than RAID-DP
- RAID-TEC default for 6TB & 8TB NL-SAS HDDs
 - Default RG size = 21, max RG size = 29
 - Required for 10TB and larger NL-SAS
 - Also supported w/10k rpm SAS HDDs and SSDs



What is Root-data partitioning for Entry with Internal HDD and RAID-TEC in ONTAP 9.0?

ONTAP 9.0 Entry with internal HDD root-data sharing follows the same logic as 8.3.x style but the difference is instead of provisioning partitions with RAID-DP, Data aggregates are configured as RAID-TEC or RAID Triple Erasure Coding with triple parity. All drive types (SAS, SATA, SSD) are eligible to utilize triple parity for the Data aggregate however Root aggregates will be populated as RAID-TEC as well if the HDD are SATA (NL-SAS or MSATA) greater or equal to 6TB and there are 16 or more drives available in the HA.

Employing triple parity enables greater data protection and improves system resiliency which is desirable in light of the concerns with large capacity SATA drive viability and risk of data loss due to three drives failing. In fact RAID-TEC provides 100x more defenses from data loss events than RAID DP.

RAID-TEC is exclusive to ONTAP 9.0. For the Entry with Internal HDD root-data sharing use case it is only supported with systems that were shipped new from factory, upgraded from 8.3.x with partitioned root/data aggregate or reinitialized as 9.0 with partitioning.

See [RAID-TEC FAQ](#) for additional information on RAID Triple Erasure Coding with triple parity.

What are the design goals of root-data (R-D) partitioning?

Simplicity, traditional aggregate provisioning model and auto-partitioning are the goals of root-data sharing.

As we introduced root-data partitioning or root-data sharing capabilities, we wanted to ensure no added complexities were incorporated into the system's administration. Simplicity was a required design goals baked-in from the foundation.

This prevents System admins having to consider any of the following:

- Learn new commands to provision aggregates.
- Make choices about partition sizes

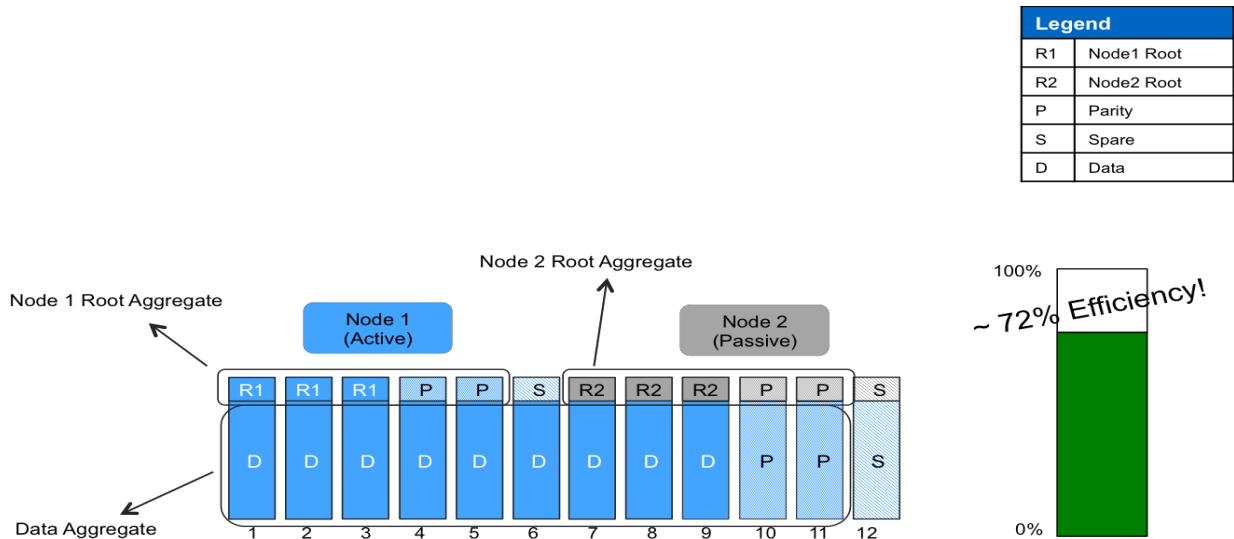
- Decide which disks get partitioned or any other details about partitioning.

Hence, we are preserving the traditional aggregate provisioning model.

One item to note here is that partitioning happens automatically during system initialization time. ONTAP automatically creates root-data partitioning on each internal HDD when system is zeroing the drives.

How do we configure an Active/Passive configuration on a FAS2520 with 8.3.x style root-data (R-D) sharing partitioning?

Figure 4: FAS2520 Active/Passive configuration



- Two dedicated cDOT root aggregates:
 - R1: 3D + 2P + 1S
 - R2: 3D + 2P + 1S
- One data aggregate:
 - D1: 9D + 2P + 1S

In the above diagram, ONTAP created root aggregates using smaller root-partitions and provides 12 data partitions to create a data aggregate. Note that this configuration is using RAID DP on both the root aggregates and data aggregate. At the same time spare capacity is allocated based on each aggregate.

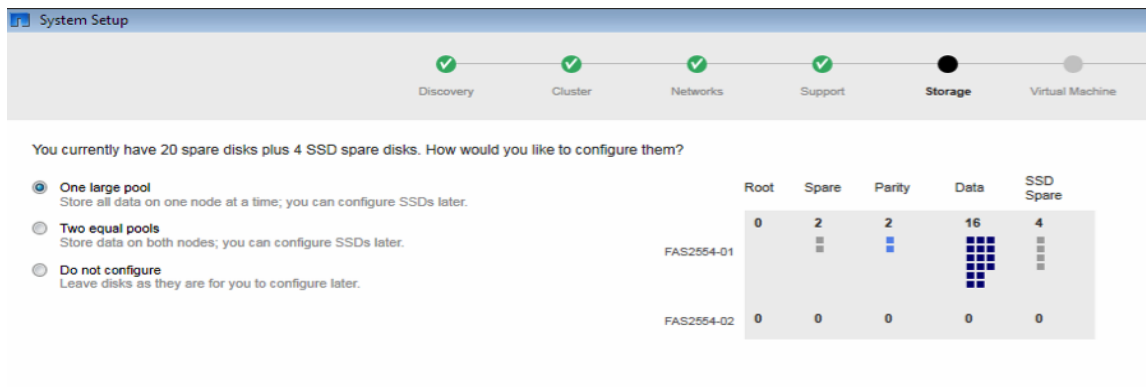
System Setup allows you to create Active/Passive and Active/Active configuration during system setup. It automates the process of assigning drives to an active system by letting you select from the following three options:

1. Create one large pool (Active/Passive)
2. Create two equal pools (Active/Active)
3. Do not configure.

Below is Figure 5, the GUI of a System Setup creating an Active/Passive configuration. You can also create an Active/Passive configuration from command line using the following CLIs:

- storage disk removeowner -data true <disk>
- storage disk assign -data true -disk <disk> -owner <node_name>

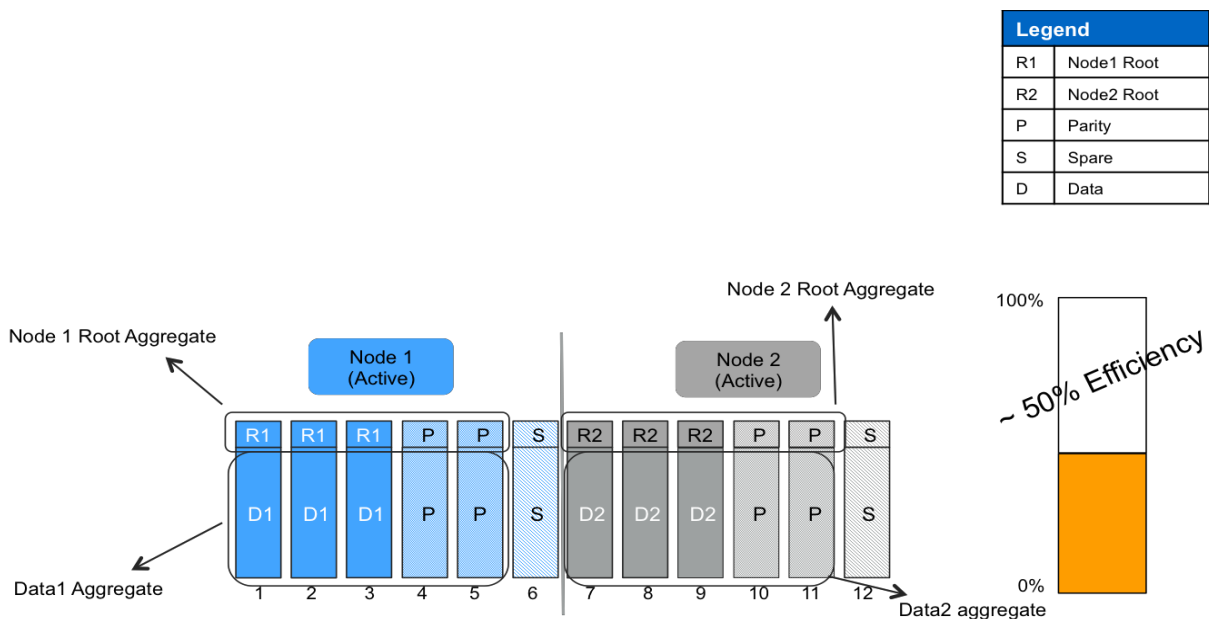
Figure 5: System Setup Active/Passive configuration



How do we configure an Active/Active configuration on a FAS2520 with root-data (R-D) sharing?

The diagram below shows how root-data partitioning is configured using Active/Active configuration on a FAS2520 system. This configuration was previously not possible in the pre-8.3 releases. One can take advantage of workload isolation and multi-tenancy using Active/Active configuration from 12 drive entry system. Overall, storage efficiency is slightly low in this configuration because of two the data aggregates which adds to the parity and spare capacity overhead.

Figure 6: FAS2520 Active/Active configuration



“Figure 5: System Setup Active/Passive configuration” depicts a system setup creating an Active/Active configuration. You can also create an Active/Active configuration from command line using the following CLIs:

- storage disk removeowner –data true <disk>
- storage disk assign –data true –disk <disk> -owner <node_name>

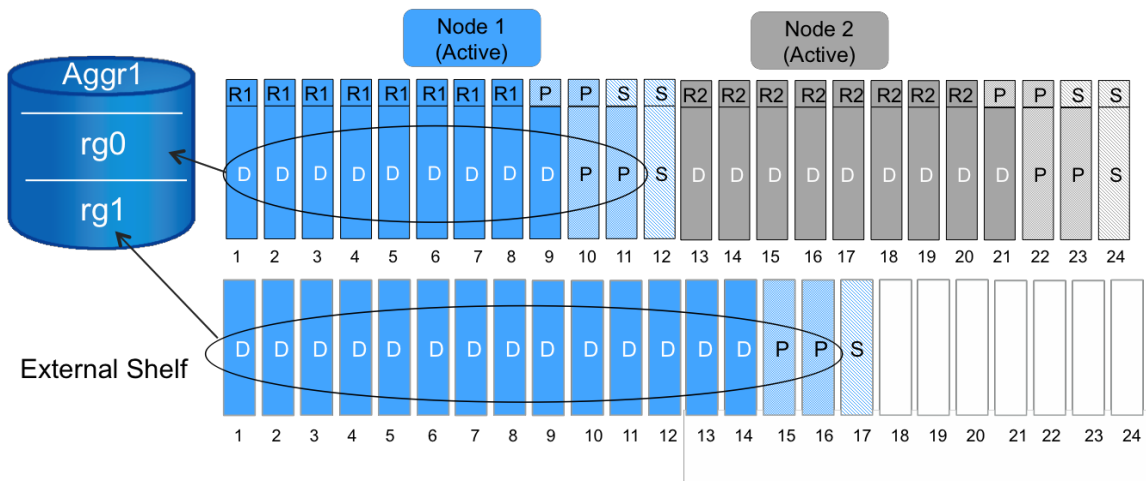
System Setup allows you to create Active/Passive and Active/Active configuration during system setup. It automates the process of assigning drives to an active system by letting you select from the following three options:

- Create one large pool (Active/Passive)
- Create two equal pools (Active/Active)
- Do not configure.

Can we add an external shelf to a system that is running Data ONTAP 8.3 root-data sharing?

Yes, you can attach an external shelf to an entry system with root-data (R-D) sharing. If you are looking to grow your existing aggregates with internal partitioned drives, you can attach an external shelf and grow the existing aggregate by creating a separate raid group using external drives. Following is an example of mixing partitioned drives with non-partitioned drives in the same aggregate.

Figure 7: Attaching an external shelf with root-data (R-D) partitioning



In this example, Aggr1 consists of two raid groups, rg0 and rg1; rg0 consists of internal partitioned drives and rg1 consists of external non-partitioned drives. However, if you decide to add the external non-partitioned drives on to the same raid group, say, rg0 ONTAP will then right size (or partition) the external drives.

What is the usable capacity profile for various drives using root-data sharing?

Please see Synergy for usable capacity values for all the different drive types and configuration.

How do I upgrade my 8.2.x or 8.3.x Entry system with internal HDD whole drives to 8.3.x or ONTAP 9 to enable Advanced Drive Partitioning with root-data sharing?

Upgrading existing whole drive (non-partitioned) 8.2.x or 8.3.x Entry with internal HDD systems to ONTAP 8.3.x or 9.0 with root-data sharing **requires** reinitializing the root aggregates. We are studying approaches recently submitted by Professional Services teams for disruptive and non-disruptive upgrades that migrate data of existing 8.2.x and 8.3.x deployments containing whole root layouts and restore the data back to root-data sharing 8.3.x or 9.0 environments. Thus far no procedure has achieved Engineering sanction however we will continue to monitor the situation as circumstances change.

When upgrading without reinitializing, a system will maintain its original root format. This is consistent in whole drive and partitioned configurations. If a configuration originates as an 8.2.x or 8.3.x with whole drives for root and is updated without reinitialization to 8.3.x or 9.0, its root and data drives will remain intact as non-partitioned whole drives. 8.3.x style partitioned setups will maintain their root style after upgrading to 9.0 as well.

The good news is all 8.3.x and 9.0 systems are eligible for Flashpool cache partitioning if SSD are present regardless of whole or partitioned root-data status.

Can we utilize root-data HDD partitioning and SSD Flashpool partitioning on one system simultaneously?

Yes. On an entry system root-data partitioning will be pre-enabled by ONTAP 8.3 or later during system initialization. If you have SSDs on your system, you can create storage pool and take advantage of Flash Pool SSD partitioning.

4 Flash Pool SSD Partitioning

What is Flash Pool SSD Partitioning?

In this use case each SSD is divided into four equal partitions. SSDs are provisioned as a cache in a flash pool aggregate. This use case is very different from root-data (R-D) sharing partitioning. Root-data partitioning allows root aggregates to be stored on smaller root-partitions, which helps in reducing root aggregate space utilization and allows for increased usable capacity. Flash Pool SSD Partitioning on the other hand is used just for caching and it enables you to share each SSD with up to four flash pool aggregates. Sharing SSDs as a caching tier in your flash pool aggregate reduces parity and spare capacity overhead because you can now allocate a single SSD for parity/spare and share it among multiple flash pool aggregates.

Without partitioning, the capacity and performance of a SSD is only dedicated to a single aggregate. With partitioning, you can share SSDs, which are capable of generating significant IOPs and throughput among multiple Flashpool aggregate. Therefore, it improves SSD storage utilization and helps you achieve higher performance.

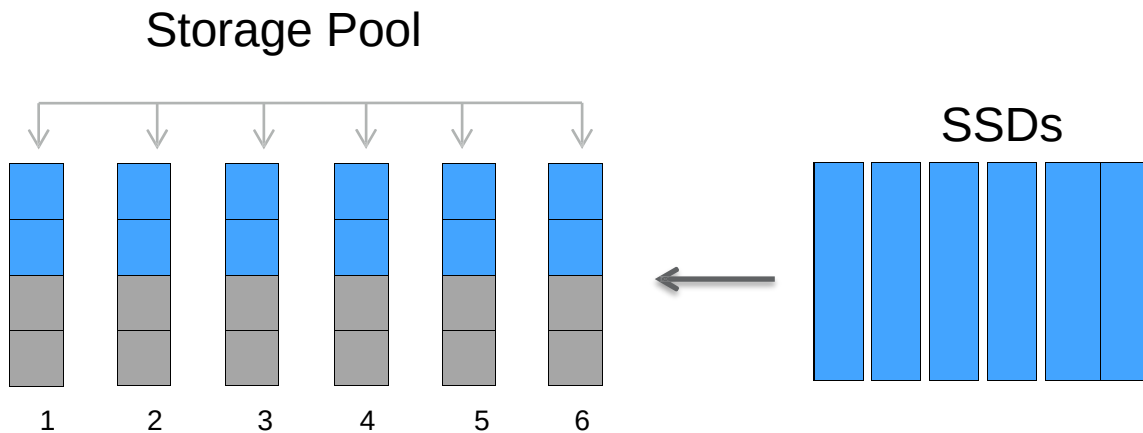
Unlike root-data partitioning, SSDs in this use case are not partitioned automatically because SSDs here are used just for data provisioning (caching). Thus, it is designed to provide more control over to the users. Therefore, users have to manually enable partitioning by creating an SSD storage pool.

What is an SSD Storage Pool?

Storage Pool is a new concept which is introduced as part of the flash pool SSD partitioning functionality. It increases cache-allocation flexibility by allowing you to provision cache based on capacity rather than the number of drives. It is a new capacity-based provisioning model, which allows system admins to provision cache based on capacity, instead of traditional drive-based provisioning. NetApp's intelligent I/O scheduling module allows SSDs to be shared by both nodes in an HA pair and with up to four flash pool aggregates supported.

Furthermore, storage pool can be created using *storage pool create command*, SSDs are partitioned when SSD storage pool is created. Following is an example of an SSD storage pool with 6 SSD drives:

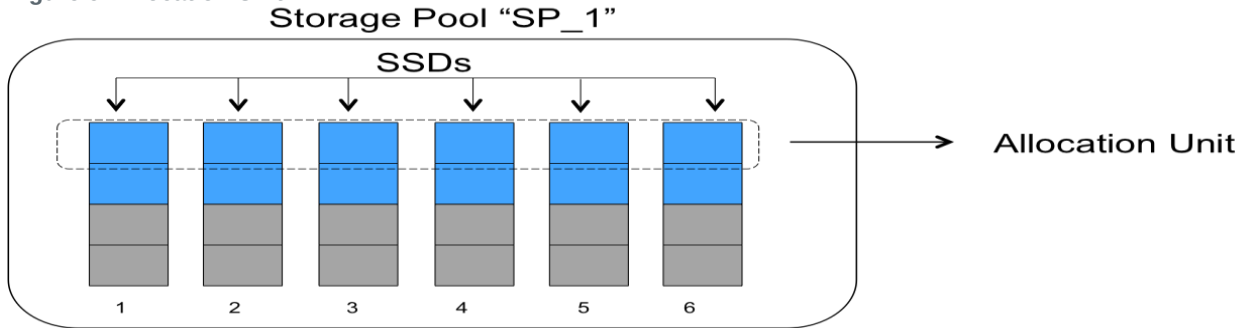
Figure 8: Storage Pool



What is an Allocation Unit (AU)?

Storage pool is divided into four allocation units. Each allocation unit contains one partition from each SSD in an SSD storage pool i.e. one allocation unit represents 25% of the cache from the storage pool. Allocation unit is also a minimum unit of allocation (MUA) that can be added to a flash pool aggregate as a single raid group. Therefore, you can also think of an AU as a single raid group.

Figure 9: Allocation Unit



Is there a concept of ownership in an SSD Storage Pool?

Yes. When you create a SSD storage pool, by default two allocations are assigned to each node. There is an option to change the ownership of allocation units (AUs) by running *storage pool reassign* command.

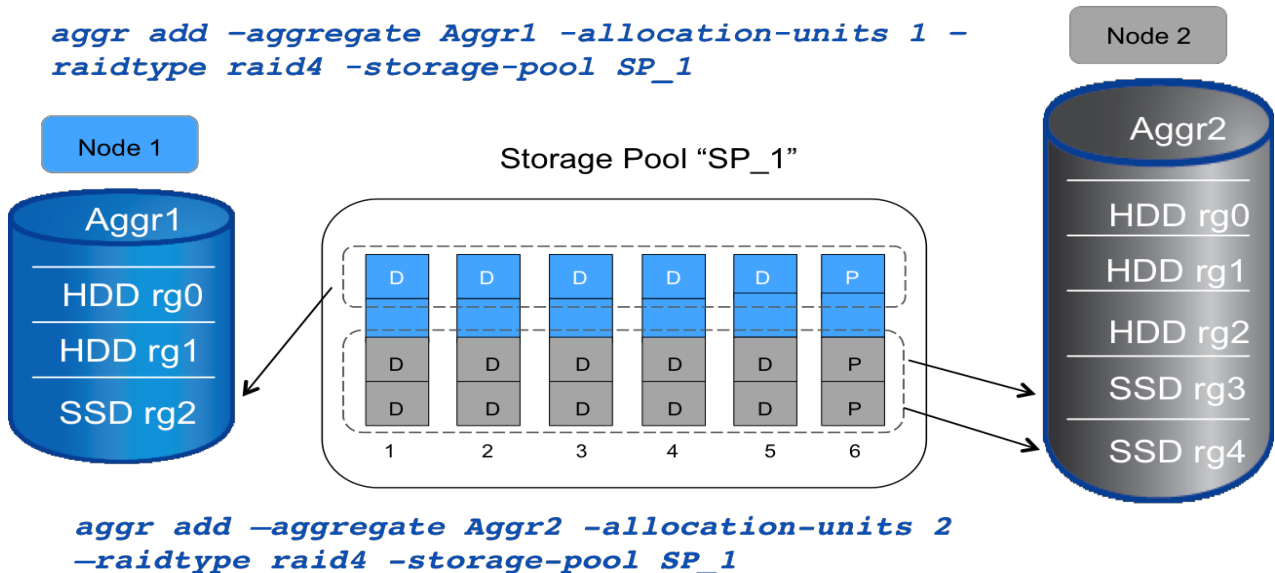
Can I use system manager to provision cache using Flash Pool SSD Partitioning?

Yes. System Manager is fully supported with Flash Pool SSD partitioning. It allows you to create SSD storage pool and assign allocation units to your flash pool aggregates. System Manager's user friendly GUI allows you to easily administer your cache without having to remember CLIs. The tool also recommends best practices for managing spare drives.

How do I provision SSD cache in a Flash Pool aggregate?

As mentioned earlier, in order to provision cache in a flash pool aggregate, create an SSD storage pool and then assign allocation units based on your needs. For example the following diagram shows how to provision a flash pool aggregate using partitioned SSDs.

Figure 10: Flash SSD Partitioning using storage pool



5 Root-data partitioning for SSDs (AFF)

What is 8.3.x style root-data (R-D) sharing for SSD?

8.3.x style root-data (R-D) sharing for SSD is very similar to the first use case i.e. root-data (R-D) sharing partitioning for HDD. The same logic as Entry with internal HDD is applied to AFF and FAS with SSD to increase usable capacity by significantly reducing the root aggregate space allocation. Only SSD can be provisioned in a root-data sharing model in non-Entry platforms, HDD are not eligible.

Root-data partitions are created during system-initialization and ONTAP divides each SSD into two partitions:

- 1) **Root Partition:** Root partition is a smaller partition and it is used to create root aggregate.
- 2) **Data Partition:** Data partition is a larger partition and it is used to create data aggregate.

Root-data sharing for SSD efficiency manages root aggregate allocation which opens up more space to provision data aggregates. This enables reduced \$/GB across the board on our Entry, Mid and High-end AFF and FAS with SSD systems.

Sharing SSDs into root and data partitions lets you manage spare capacity efficiency. Users can assign partitions for spare capacity instead of dedicating the whole drive. This technique reduces parity and spare overhead.

Up to 48 SSD can be partitioned with 8.3.x style of root-data (R-D) sharing using RAID DP for the root and data aggregate. From a partition ownership perspective each node owns the data and root constituents of a drive. You'll note later in some ONTAP 9 root data partitioning scenarios that may not be the case. Ownership of a single drive's data and root elements could be owned by a single node or both of the HA partners in a 9.0 environment. We'll touch on how ONTAP 9 SSD root data partitioning uses something called Root-Data-Data (R-D2) sharing in a later section.

The following is the configuration of an 8.3.x style root-data (R-D) sharing 24-SSD HA Active/Active AFF:

- 1) Two dedicated cDOT Root Aggregates:
 Root Aggregate 1: 8 Data + 2 Parity + 2 Spare
 Root Aggregate 2: 8 Data + 2 Parity + 2 Spare
- 2) Two Data Aggregates:
 Data Aggregate 1: 9 Data + 2 Parity + 1 Spare
 Data Aggregate 2: 9 Data + 2 Parity + 1 Spare

From an efficiency standpoint a total of 18 partitions are allocated for Data, 4 for Parity and 2 for spares - yields:

- Data-to-Parity ratio of 18:4 or 9:2.
- Overhead tax of parity & spares of 6 partitions
- Data-to-Overhead ratio of 18:6 or 3:1

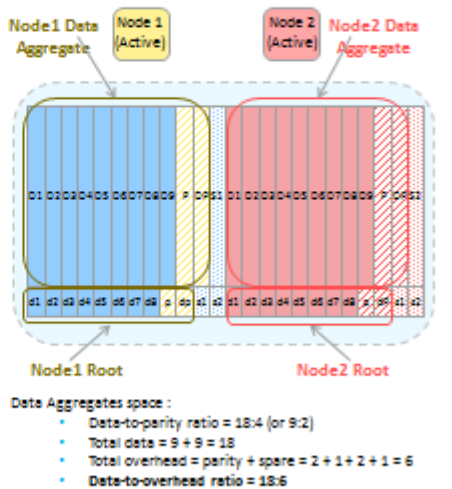
Not too shabby for single shelf 24-SSD AFF HA Active/Active setup, with 8.3.x R-D sharing however there is room for improvement and we'll see how the same layout in ONTAP 9.0 fares in comparison.

Figure 11: 24-drive 1-shelf HA Active/Active AFF with 8.3.x style root-data (R-D) SSD sharing

ADP Root-Data (R-D) with SSD Partitioning

Data ONTAP 8.3

- Provisions root aggregates using small partitions on many drives
 - Up to 48 SSDs on AFF and FAS with SSD systems
- Increases usable data capacity and storage efficiency
 - Root aggregate uses only the space needed
 - Eliminates the need for dedicated parity drives
 - Recapture lost capacity for User Data



NetApp

What is 9.0.x style root-data-data (R-D2) sharing for SSDs and why is it better than 8.3.x (R-D) sharing with 1-shelf AFF 24-drive Active/Active configurations?

8.3.x pioneered root-data (R-D) SSD sharing functionality for AFF and FAS with SSD environments. In that Use Case up to 48 SSD could be subdivided into 2 partitions - 1 large partition for the data aggregate and 1 smaller partition for the root aggregate. Each drive's data and root slice is owned by the same controller in the HA pair. ONTAP 9.0 takes 8.3.x (R-D) provisioning scheme a step further and introduces root-data-data (R-D2) SSD sharing in which partitions are broken into 3 slices instead of 2. The smaller root partition remains the same as 8.3.x style partitioning, but for data the constituents are sliced in half. In 1-shelf 24-drive Active/Active AFF or FAS with SSD R-D2 Active/Active layout this yields a nearly 17% increase in storage efficiency as the data partition ownership is split between the nodes, where one node is apportioned one data partition and the partner owns the other equally sized partition. This arrangement saves space by eliminating 2 parity drives and 1 spare in comparison to 8.3.x's R-D sharing layout without compromising performance which about the same for the two schemes. Due to core dump requirements there's a minimum SSD size of 400G or greater that can be partitioned in 9.0 R-D2 SSD sharing scheme; SSD smaller than 400G will be provisioned with 8.3.x R-D SSD sharing model.

The following is an ONTAP 9 style root-data-data (R-D2) SSD sharing 1-shelf configuration with 24-SSD HA AFF:

1) Two dedicated cDOT Root Aggregates:

Root Aggregate 1: 8 Data + 2 Parity + 2 Spare

Root Aggregate 2: 8 Data + 2 Parity + 2 Spare

2) Two Data Aggregates:

Data Aggregate 1: 21 Data + 2 Parity + 1 Spare

Data Aggregate 2: 21 Data + 2 Parity + 1 Spare

From an efficiency standpoint let's examine how ONTAP 9 root-data-data (R-D2) sharing in a 1-shelf AFF 24-drive Active/Active HA stacks up.

A total of 42 partitions are allocated for Data, 4 for Parity and 2 for spares – yields:

- Data-to-Parity ratio of 42:4 or 21:2
- Overhead of tax of parity & spares of 6 partitions
- Data-to-Overhead ratio of 21:3 or 7:1

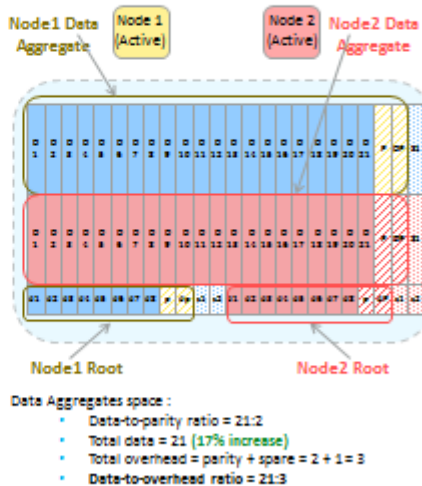
This represents a ~17% improvement over 8.3.x style root-data SSD sharing AFF with a single shelf – a substantial increase.

Figure 12: 24-drive 1-shelf HA Active/Active AFF with ONTAP 9 style root-data-data (R-D2) SSD sharing

Root-Data-Data (R-D2) with SSD Partitioning

ONTAP 9.0

- Divides SSDs into 2 large data partitions and 1 small root partition
 - Root partition sizes are same as with 8.3.x R-D
 - Data partitions are half the size
- 1 data partition per SSD is assigned to each node in the HA pair
 - Only 2 SSDs total are needed for RAID-DP parity for both aggregates (vs 2 for each aggr)
 - Only 1 SSD is needed for hot spare partitions
 - Result is 3 more SSDs for data and 17% more usable capacity (with a 24-drive configuration)
 - Requires SSD ≥ 400G for R-D2, otherwise 8.3.x method used
- Performance is comparable to 8.3.x R-D partitioning



NetApp

How is ONTAP 9 style root-data-data (R-D2) SSDs sharing different in a 2-shelf AFF 24-drive Active/Active configuration than 1-shelf?

For AFF or FAS with SSD Active/Active configuration in 9.0 root-data-data (R-D2) SSD sharing scheme of 2-shelves or 48-drives (SSD ≥ 25 up to 48 drives) the partition ownership arrangement is somewhat different than a 1-shelf layout. If you recall, the 1-shelf AFF ONTAP 9 (R-D2) SSD sharing model had 3 slices per drive consisting of 1 small root partition and 2 equally sized data partitions, where each node owned 1 of 2 data slices. In a 2-shelf (R-D2) SSD sharing paradigm the partition ownership is similar to a root-data (R-D) sharing model in that one node owns the drive's partitions for both root and data. In creating the new arrangement 2 RAID groups (rg0 and rg1) are formed per Data aggregate for each controller. Space efficiency wise ONTAP 9.0 AFF HA 2-shelf 48-drive Active/Active with root-data-data (R-D2) sharing and 8.3.x root-data (R-D) sharing scheme with 48-drive, 2-shelves are essentially the same but the additional RAID group present in R-D2 gives you greater flexibility than before.

The following is an ONTAP 9 style root-data-data (R-D2) SSD sharing 2-shelf configuration with 24-SSD HA Active/Active AFF:

- 1) Two dedicated cDOT Root Aggregates:

Root Aggregate 1: 20 Data + 2 Parity + 2 Spare

Root Aggregate 2: 20 Data + 2 Parity + 2 Spare

2) Two Data Aggregates with 2 RAID group each:

Data Aggregate 1: 21 Data + 2 Parity + 1 Spare (rg0)

21 Data + 2 Parity + 1 Spare (rg1)

Data Aggregate 2: 21 Data + 2 Parity + 1 Spare (rg0)

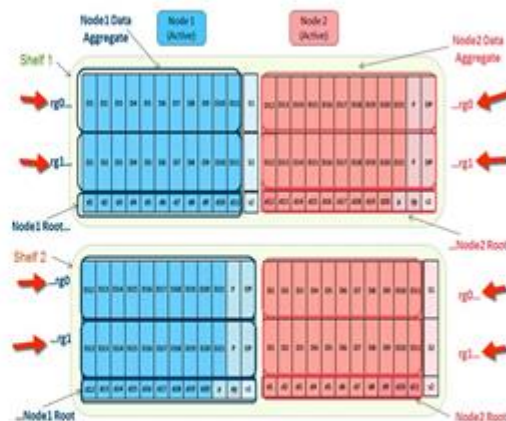
21 Data + 2 Parity + 1 Spare (rg1)

Figure 13: 48-drive 2-shelf HA Active/Active AFF with ONTAP 9 style root-data-data (R-D2) SSD sharing

Root-Data-Data (R-D2) SSD partitioning – 2 shelves

ONTAP 9.0

- R-D2 SSD sharing with 2-shelves provisions (3) partitions but slice ownership similar to Root-Data method
 - Both Data partitions per SSD owned by (1) node
 - (2) RAID-Groups are created for each node
 - For AFF & FAS all SSD Use Cases with > 24 SSD, up to 48 Drives



NetApp

How many SSDs can be partitioned with 8.3.x style root-data (R-D) sharing & ONTAP 9 root-data-data (R-D2) sharing in an AFF system; what is the least number of drives that are required for SSD root-data partitioning in 8.3.x and ONTAP 9.

A maximum of 48 SSDs can be partitioned per HA pair in 8.3.x root-data (R-D) SSD sharing and ONTAP 9.0 style root-data-data (R-D2) SSD sharing models. Additional SSD beyond the 48-drive threshold will remain as whole spares. And on the opposite end of the spectrum, at least 8 SSDs are required to enable root-data (8.3.x) or root-data-data (9.0) sharing functionality on an AFF or FAS with SSD HA pair. This minimum requirement is driven by the size of the root aggregate needed by ONTAP.

6 Supported Configurations

What platforms are supported with Advanced Drive Partitioning?

Entry = Root-data partitioning for HDDs (use case #1)

Flash Pool = Flash pool SSD partitioning (use case #2)

AFF = Root-data partitioning for SSDs (use case #3)

Figure 14: Supported ONTAP and Platform Configurations

Supported Configurations

Data ONTAP	ONTAP 8.3.x & ONTAP 9.0	✓
Platforms	FAS25xx – Entry with HDD, Flash Pool and Entry FAS with SSD	✓
	FAS22xx – Entry with HDD, Flash Pool and Entry FAS with SSD	✓
	FAS32xx – Flash Pool and FAS with SSD	✓
	FAS62xx – Flash Pool and FAS with SSD	✓
	FAS80xx – Flash Pool, AFF and FAS with SSD	✓

What drives are supported with Advanced Drive Partitioning?

Figure 15: Supported Drives

Supported Configurations

Drive Support		
Drive Category	Drive Type	Capacity
Capacity	SATA/NL-SAS – 7.2K RPM	2TB, 3TB, 4TB, 4TB NSE, 6TB, 6TB NSE & 8TB
Performance	SAS – 10K RPM	600GB, 900GB, 900GB NSE, 1.2TB, 1.8TB & 1.8TB NSE
Ultra Performance	SSD	200GB, 400GB, 800GB, 800GB NSE, 1.6TB, 1.6TB NSE, 3.8TB & 15.3TB
Comments	- Root aggregate is configured by cDOT automatically - User to provision Flash Pool SSD Partitions - User to provision data aggregates	

7 Additional Readings

AFF 15TB SSD Quick Info Guide:

<https://fieldportal.netapp.com/content/378300?assetComponentId=378929>

AFF Tech FAQ: <https://fieldportal.netapp.com/content/296217?assetComponentId=296749>

AFF TR: <http://www.netapp.com/us/media/tr-4505.pdf>

RAID-TEC FAQ:

<https://fieldportal.netapp.com/content/409728?assetComponentId=410359>

Storage Subsystem FAQ: <https://fieldportal.netapp.com/content/195603?assetComponentId=195701>

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

Go further, faster®



www.netapp.com

Partitioning FAQ

NetApp Confidential – Limited Use Only