Table of contents

Executive summary .............................................................................................................................................................. 2
Storage tiers: Opportunity and challenge .......................................................................................................................... 3
Brief overview of volume mapping ..................................................................................................................................... 4
  Physical disks ..................................................................................................................................................................... 4
  Logical disks ....................................................................................................................................................................... 4
  Common provisioning groups ......................................................................................................................................... 5
  CPGs and workloads ......................................................................................................................................................... 5
  Virtual volumes ................................................................................................................................................................. 5
Requirements ........................................................................................................................................................................ 6
HP 3PAR Adaptive Optimization Software ......................................................................................................................... 6
  CPG as tiers in AO configuration ...................................................................................................................................... 6
  Data locality ....................................................................................................................................................................... 6
  Tiering analysis algorithm ................................................................................................................................................ 7
  Average tier access rate densities .................................................................................................................................. 7
  Design tradeoff: Granularity of data movement ........................................................................................................... 8
HP 3PAR Adaptive Flash Cache ....................................................................................................................................... 9
HP 3PAR Thin Deduplication ............................................................................................................................................ 9
Implementing Adaptive Optimization ................................................................................................................................. 9
  Sizing for Adaptive Optimization .................................................................................................................................... 9
  Configuring AO using region I/O density reports ........................................................................................................... 10
  Using AO configuration with VVsets ............................................................................................................................ 16
  Adaptive Optimization with Remote Copy .................................................................................................................... 17
Use case .............................................................................................................................................................................. 18
  Accelerating workloads by adding SSDs ........................................................................................................................ 18
  Lowering cost per GB by configuring a three-tier configuration with SSD, FC, and NL ......................................... 20
Summary ............................................................................................................................................................................. 21
Executive summary

Now more than ever, IT managers are under pressure to deliver the service levels necessary for a wide variety of mission-critical applications at the lowest possible cost. The introduction of solid-state drive (SSD) technology has created enormous demand for an optimization solution capable of leveraging this new drive class to improve service levels without raising costs. Traditional approaches to service-level optimization have not been successful in meeting this need due to limitations such as:

- High cost of placing entire volumes onto SSDs
- Inability to scale sub-volume optimization to accommodate large, mission-critical data centers
- Insufficient administrative controls and override mechanisms
- Inability to move data without impacting service levels
- Unproven technologies that introduce unnecessary risk

New opportunities exist to optimize the cost and performance of storage arrays, thanks to the availability of a wide range of storage media such as SSDs, high-performance hard disk drives (HDDs), and high-capacity HDDs. But these opportunities come with the challenge of doing it effectively and without increasing administrative burdens, because the tradeoffs for storage arrays are different from CPU memory hierarchies.

HP 3PAR Adaptive Optimization (AO) enables the HP 3PAR StoreServ to act as a hybrid storage array where the StoreServ can blend flash-based SSDs and HDDs to provide high performance at an affordable price. HP 3PAR Adaptive Optimization Software delivers the next generation in autonomic storage tiering by taking a fine-grained, highly automated approach to service-level improvement. Using the massively parallel, widely striped, and highly granular HP 3PAR architecture as a foundation, Adaptive Optimization leverages the proven sub-volume data movement engine built into the HP 3PAR Operating System Software. The result is highly reliable, non-disruptive, cost-efficient sub-volume storage tiering that delivers the right quality of service (QoS) at the lowest transactional cost.

Adaptive Optimization enables enterprise and cloud data centers to improve service levels effortlessly, on a large scale and for a lower total cost than other solutions. This approach enables data centers using HP 3PAR systems to meet enterprise and cloud performance demands within a smaller footprint and lower storage equipment cost than by using only Fibre Channel (FC) storage. This level of cost savings is made possible in part by application-specific thresholds and comprehensive support for thin and fat volumes as well as volume copies. Adaptive Optimization enables IT managers to react swiftly to changing business needs while delivering service-level improvement over the entire application lifecycle—autonomically and non-disruptively.

With this highly autonomic technology, IT managers can now achieve non-disruptive, cost- and performance-efficient storage within even the largest and most demanding enterprise and cloud environments.

This white paper explains some of the tradeoffs, describes the technology that adaptively optimizes storage on HP 3PAR StoreServ Storage, and illustrates its effectiveness with performance results.
Storage tiers: Opportunity and challenge

Modern storage arrays support multiple tiers of storage media with a wide range of performance, cost, and capacity characteristics—ranging from inexpensive Serial ATA (SATA) HDDs that can sustain only about 75 IOPS to more expensive flash memory-based SSDs that can sustain thousands of IOPS. Volume RAID and layout choices enable additional performance, cost, and capacity options. This wide range of cost, capacity, and performance characteristics is both an opportunity and a challenge.

Figure 1. Autonomic tiering with HP 3PAR StoreServ

The opportunity is that the performance and cost of the system can be improved by correctly placing the data on different tiers; move the most active data to the fastest (and most expensive) tier and move the idle data to the slowest (and least expensive) tier. The challenge, of course, is to do this in a way that minimizes the burden on storage administrators while also providing them with appropriate controls.

Currently, data placement on different tiers is a task usually performed by storage administrators—and their decisions are often based not on application demands but on the price paid by the users. If they don’t use careful analysis, they may allocate storage based on available space rather than on performance requirements. At times, HDDs with the largest capacity may also have the highest number of accesses. However, the largest HDDs are often the slowest HDDs. This can create significant performance bottlenecks.

There is an obvious analogy with CPU memory hierarchies. Although the basic idea is the same (use the smallest, fastest, most expensive resource for the busiest data), the implementation tradeoffs are different for storage arrays. While deep CPU memory hierarchies (first-, second-, and third-level caches; main memory; and finally paging store) are ubiquitous and have mature design and implementation techniques, storage arrays typically have only a single cache level (the cache on disk drives usually acts more like a buffer than a cache). Automatic tiering in storage arrays is a recent development, and not commonplace at all. The industry still has much to learn about it.
Brief overview of volume mapping

Before you can understand HP 3PAR Adaptive Optimization, it is important to understand volume mapping on HP 3PAR StoreServ Storage as illustrated in figure 2.

**Figure 2. HP 3PAR volume management**

The HP 3PAR Operating System has a logical volume manager that provides the virtual volume abstraction. It comprises several layers, with each layer being created from elements of the layer below.

**Physical disks**

Every physical disk (PD) that is admitted into the system is divided into 1 GB chunklets. A chunklet is the most basic element of data storage of the HP 3PAR OS. These chunklets form the basis of the RAID sets; depending on the sparing algorithm and system configuration, some chunklets are allocated as spares.

**Logical disks**

The logical disk (LD) layer is where the RAID functionality occurs. Multiple chunklet RAID sets, from different PDs, are striped together to form an LD. All chunklets belonging to a given LD will be from the same drive type. LDs can consist of all nearline (NL), FC, or SSD drive type chunklets.

There are three types of logical disks:

1. User (USR) logical disks provide user storage space to fully provisioned virtual volumes.
2. Shared data (SD) logical disks provide the storage space for snapshots, virtual copies, thinly provisioned virtual volumes (TPVVs), or thinly deduped virtual volumes (TDVVs).
3. Shared administration (SA) logical disks provide the storage space for snapshot and TPVV administration. They contain the bitmaps pointing to which pages of which SD LD are in use.

The LDs are divided into regions, which are contiguous 128 MB blocks. The space for the virtual volumes is allocated across these regions.
Common provisioning groups

The next layer is the common provisioning group (CPG), which defines the LD creation characteristics, such as RAID type, set size, disk type for chunklet selection, plus total space warning and limit points. A CPG is a virtual pool of LDs that allows volumes to share resources and to allocate space on demand. A thin provisioned volume created from a CPG will automatically allocate space on demand by mapping new regions from the LDs associated with the CPG. New LDs of requested size will be created for fully provisioned volumes and regions are mapped from these LDs.

CPGs and workloads

HP 3PAR StoreServ performs efficiently for any type of workload, and different workloads can be mixed on the same array. These different workloads may need different types of service levels to store their data. For example, for high-performance mission-critical workloads, it may be best to create volumes with RAID 5 protection on SSD or RAID 1 protection on fast class [FC or serial-attached SCSI (SAS) performance HDDs]. For less-demanding projects, RAID 5 on FC drives or RAID 6 on NL drives may suffice. For each of these workloads, you can create a CPG to serve as the template for creating virtual volumes (VVs) for the workload. VVs can be moved between CPGs with the HP 3PAR Dynamic Optimization (DO) software command `tunevv`, thereby changing their underlying physical disk layout and hence their service level.

Virtual volumes

The top layer is the VV, which is the only data layer visible to hosts. VVs draw their resources from CPGs and the volumes are exported as virtual logical unit numbers (VLUNs) to hosts.

A VV is classified by its type of provisioning, which can be one of the following:

- Full: Fully provisioned VV, either with no snapshot space or with statically allocated snapshot space
- TPVV: Thin provisioned VV, with space for the base volume allocated from the associated user CPG and snapshot space allocated from the associated snapshot CPG (if any)
- TDVV: Thin deduped VV, with space for the base volume allocated from the associated user CPG (SSD tier only) and snapshot space allocated from the associated snapshot CPG (if any)
- CPVV: Commonly provisioned VV. The space for this VV is fully provisioned from the associated user CPG and the snapshot space allocated from the associated snapshot CPG

TPVVs and TDVVs associated with the same CPG share the same LDs and draw space from that pool as needed, allocating space on demand in small increments for each controller node. As the volumes that draw space from the CPG require additional storage, the HP 3PAR OS automatically increases the logical disk storage until either all available storage is consumed or, if specified, the CPG reaches the user-defined growth limit, which restricts the CPG’s maximum size. The size limit for an individual VV is 16 TB.

Figure 2 illustrates the volume mapping for both non-tiered as well as tiered (adaptively optimized) volumes. For non-tiered VVs, each space (user, snap, or admin) is mapped to LD regions within a single CPG, therefore, is in a single tier. For tiered VVs, each space can be mapped to regions from different CPGs.

Finally, remember that although this mapping from VVs to VV spaces to LDs to chunklets is complex, the user is not exposed to this complexity because the system software automatically creates the mappings.

The remainder of this white paper describes how Adaptive Optimization tiering is implemented and the benefits that can be expected.
Requirements

HP 3PAR Adaptive Optimization is a licensed feature of HP 3PAR OS and is supported on all HP 3PAR StoreServ Storage systems. For more information on HP 3PAR Adaptive Optimization licensing, contact your HP representative or authorized HP partner. Creating and managing AO requires HP 3PAR StoreServ Management Console (SSMC) 4.5 or later, or SSMC version 2.1 or later. To use the command line, you must install HP 3PAR Command Line Interface 3.1.2 MU2 or later. Certain features and reports of HP 3PAR Adaptive Optimization described in this paper are only available from HP 3PAR SSMC 4.5 or later and HP 3PAR OS 3.1.3 MU1 and later. The ability to specify minimum and maximum sizes of each tier is supported in HP 3PAR OS 3.2.2 and later.

HP 3PAR Adaptive Optimization Software

CPG as tiers in AO configuration

Before creating an AO configuration, it is required to create CPGs. CPGs are used as tiers within AO. An AO configuration can have a maximum of three tiers, and a minimum of two tiers are required to define a new AO configuration. A CPG can be part of only one AO configuration. So, every AO configuration will need a different set of CPGs. Virtual volumes that are not part of an AO configuration have all their regions mapped to LDs belonging to a single CPG. However, in the case of AO, VVs will have regions mapped to LDs from different CPGs (tiers). Data placement for a particular region is decided based on statistics collected and analyzed by AO for each region.

HP 3PAR Adaptive Optimization leverages data collected by HP 3PAR System Reporter (SR) on node. The SR on node periodically collects detailed performance and space data that is used by AO for the following:

- Analyze the data to determine the volume regions that should be moved between tiers
- Instruct the array to move the regions from one CPG (tier) to another
- Provide the user with reports that show the impact of Adaptive Optimization

Refer to the HP 3PAR System Reporter white paper for more details.

Data locality

Sub-LUN tiering solutions such as AO provide high value when there is a lot of locality of data. Locality of data means there is a limited area of the data address space used by a server or application that receives a large percentage of the I/O requests compared to the rest of the data space. This is the result of how a server’s applications use the data on the LUNs.

A relatively small portion of the data receives a lot of I/O and is hot, while a larger portion of the data space receives very few or no I/O and is considered cold. AO is not responsible for causing locality of data. Sub-LUN tiering solutions move small pieces of a LUN up and down between tiers based on how hot or cold the data pieces are. In the case of AO, the pieces are 128 MB regions.

Typically, over 80 percent of IOPS is served by less than 20 percent of addressable capacity. If such a LUN is part of a single tier, then that tier should be capable of handling the maximum IOPS requirement even when most of the capacity will not be accessed as frequently. With Adaptive Optimization, the LUN will be spread over multiple tiers; a small amount of capacity that is accessed the most will be moved to the fastest tier and any capacity that is least accessed will be moved to the slowest and cheapest tier. Due to high data locality, the fastest tier can be small as compared to the other tiers and still provide the required IOPS. Figure 3 gives an example of how I/O will be distributed on a LUN that has a single tier and of a LUN that is spread across multiple tiers.

An application that does not have high data locality will have IOPS spread throughout the LUN and a small fast tier will not be a good fit; capacity allocated from the slower tiers will also have a similar access rate as the fastest tier. Such an application will not perform optimally with a tiered LUN. It should be deployed on a LUN created using a single tier.

The section about region I/O density reports will explain in detail how to find out if a particular group of applications is suitable for sub-LUN tiering.
Tiering analysis algorithm

The tiering analysis algorithm that selects regions to move from one tier to another considers several things described in the following sections.

**Space available in the tiers**

If the space in a tier exceeds the tier size (or the CPG warning limit), then the algorithm will first try to move regions out of that tier into any other tier with available space, in an attempt to lower the tier’s size below the limit. If no other tier has space, then the algorithm logs a warning and does nothing (Note: If the warning limit for any CPG is exceeded, the array will generate a warning alert). If space is available in a faster tier, it chooses the busiest regions to move to that tier. Similarly, if space is available in a slower tier, it chooses the idlest regions to move to that tier. The average tier service times and average tier access rates are ignored when data is being moved because the size limits of a tier have been exceeded.

**Average tier service times**

Normally, HP 3PAR Adaptive Optimization tries to move busier regions in a slow tier into a higher performance tier. However, if a higher performance tier gets overloaded (too busy), performance for regions in that tier may actually be lower than regions in a slower tier. In order to prevent this, the algorithm does not move any regions from a slower to a faster tier if the faster tier’s average service time is not lower than the slower tier’s average service time by a certain factor (a parameter called `svctFactor`). There is an important exception to this rule because service times are only significant if there is sufficient IOPS load on the tier. If the IOPS load on the destination tier is below another value (a parameter called `minDstIops`), then we do not compare the destination tier’s average service time with the source tier’s average service time. Instead, we use an absolute threshold (a parameter called `maxSvctms`).

**Average tier access rate densities**

When not limited, as described earlier, by lack of space in tiers or by high average tier service times, Adaptive Optimization computes the average tier access rate densities (a measure of how busy the regions in a tier are on average, calculated with units of IOPS per gigabyte per minute). It also compares them with the access rate densities of individual regions in each tier. Then, it decides whether to move the region to a faster or slower tier.

We first consider the algorithm for selecting regions to move from a slower to a faster tier. For a region to be considered busy enough to move from a slower to a faster tier, its average access rate density or `acrr(region)` must satisfy these two conditions:

First, the region must be sufficiently busy compared to other regions in the source tier:

\[ \text{acrr}(\text{region}) > \text{srcAvgFactorUp}(\text{Mode}) \times \text{acrr}(\text{srcTier}) \]

Where `acrr(srcTier)` is the average access rate density of the source (slower) tier and `srcAvgFactorUp(Mode)` is a tuning parameter that depends on the mode configuration parameter. Note that by selecting different values of `srcAvgFactorUp` for performance, balanced, or cost mode values, HP 3PAR Adaptive Optimization can control how aggressive the algorithm is in moving regions up to faster tiers.

Second, the region must meet one of two conditions: it must be sufficiently busy compared with other regions in the destination tier, or it must be exceptionally busy compared with the source tier regions. This second condition is added to
cover the case in which a very small number of extremely busy regions are moved to the fast tier, but then the average access rate density of the fast tier creates too high a barrier for other busy regions to move to the fast tier:

\[
\text{accr}(\text{region}) > \min\left((\text{dstAvgFactorUp}(\text{Mode}) \times \text{accr}(\text{dstTier})), (\text{dstAvgMaxUp}(\text{Mode}) \times \text{accr}(\text{srcTier}))\right)
\]

The algorithm for moving idle regions down from faster to slower tiers is similar in spirit—but instead of checking for access rate densities greater than some value, the algorithm checks for access rate densities less than some value:

\[
\text{accr}(\text{region}) < \text{srcAvgFactorDown}(\text{Mode}) \times \text{accr}(\text{srcTier})
\]

\[
\text{accr}(\text{region}) < \max\left((\text{dstAvgFactorDown}(\text{Mode}) \times \text{accr}(\text{dstTier})), (\text{dstAvgMinDown}(\text{Mode}) \times \text{accr}(\text{srcTier}))\right)
\]

AO makes a special case for regions that are completely idle (\(\text{accr}(\text{region}) = 0\)). These regions are moved directly to the lowest tier, even when performance mode is selected.

**Minimum and maximum space within a tier**

HP 3PAR InForm OS version 3.2.2 introduces the ability to manage the space AO consumes directly in each tier. The previous discussion illustrates how AO monitors the system and moves data between tiers in response to available space and data locality of access. This can result in little or no data in a given tier. In cases where tier 0 is an SSD tier and all SSD space is dedicated to AO, it is desirable to maximize the SSD space utilization. HP 3PAR InForm OS version 3.2.2 introduces new options to the "createaocfg," "setaocfg," and "startao" commands to support a user-specified minimum or maximum size for each tier.

The new command options are -t0min, -t1min, and -t2min to set a minimum size for a tier and -t0max, -t1max, and -t2max to set a maximum size for a tier. Specifying a minimum size target for tier 0 of 400 GB (setaocfg -t0min 400G), for example, will instruct AO to utilize a minimum of 400 GB in tier 0.

The space can be specified in MB (default), GB (g or G), or TB (t or T). Specifying a size of 0 instructs AO not to enforce a minimum or maximum size target.

**Example**

A new array is purchased that includes eight 480 GB SSDs. The intention is to use all the space from all eight SSDs as tier 0 in an AO configuration. AO considers space constraints (e.g., is tier 1 out of space?) and region density (what regions in tier 1 are receiving more IOPS than other regions?) to determine which regions should move to tier 0. This could result in little or no SSD space being used by AO.

Using the new options to the createaocfg, setaocfg, or startao commands to set a minimum size will make effective use of the SSD space. The command setaocfg -t0min 1 TB <aocfg name> would result in a minimum of 1 TB of the SSD space being utilized by AO.

Consult the 3.2.2 CLI Reference Manual or CLI help output for more details on command syntax.

**Design tradeoff: Granularity of data movement**

The volume space to LD mapping has a granularity of either 128 MB (user and snapshot data) or 32 MB (admin metadata) and that is naturally the granularity at which the data is moved between tiers. Is that the optimal granularity? On the one hand, having fine-grain data movement is better since we can move a smaller region of busy data to high-performance storage without being forced to bring along additional idle data adjacent to it. On the other hand, having a fine-grain mapping imposes a larger overhead because HP 3PAR Adaptive Optimization needs to track performance of a larger number of regions, maintain larger numbers of mappings, and perform more data movement operations. Larger regions also take more advantage of spatial locality (the blocks near a busy block are more likely to be busy in the near future than a distant block).
**HP 3PAR Adaptive Flash Cache**

HP 3PAR OS 3.2.1 introduces a new feature called HP 3PAR Adaptive Flash Cache (AFC), which provides read cache extensions by leveraging HP 3PAR first-in-class virtualization technologies. This functionality allows dedicating a portion of SSD capacity as an augmentation of the HP 3PAR StoreServ data cache, reducing application response time for read-intensive I/O workloads. This feature can coexist with Adaptive Optimization.

In order to understand how much of the existing SSD capacity should be allocated to AFC, refer to the [HP 3PAR Adaptive Flash Cache technical white paper](#). If customers already have Adaptive Optimization configured and no available space on the SSD tier, they may set a warning limit on the CPG SSD tier to free up some space to then allocate to AFC. If the array is running 3.2.2 or later, space can be freed using the AO tier CFG limit options (e.g., –t0max) to the createaocfg, setaocfg, and startao commands.

**HP 3PAR Thin Deduplication**

HP 3PAR OS 3.2.1 MU1 introduces a new feature called HP 3PAR Thin Deduplication, which allows provisioning of TDVVs to an SSD tier. While the feature can coexist on the same system where Adaptive Optimization is running and configured, TDVVs cannot be provisioned on a CPG that is part of an AO configuration, and the system will prevent creating an AO configuration if a CPG has TDVVs associated with it. A system can be configured with a shared pool of SSDs that may be used for sub-tiering (AO), cache augmentation (AFC), and provisioning of TDVVs, TPVVs, or full VVs.

**Implementing Adaptive Optimization**

**Sizing for Adaptive Optimization**

It’s important to size the array correctly when setting up an HP 3PAR array with Adaptive Optimization. Table 1 gives recommendations on how to size the array if you do not have the region density reports. Details on how to interpret region density reports are described later in this paper.

<table>
<thead>
<tr>
<th>Configuration type</th>
<th>SSD tier</th>
<th>SAS 10k/15k tier</th>
<th>NL tier</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 tiers SSD-FC/SAS</td>
<td>At least 5% of the capacity*</td>
<td>95% of the capacity</td>
<td>N/A</td>
</tr>
<tr>
<td>2 tiers FC/SAS-NL</td>
<td>N/A</td>
<td>Minimum 60% of the capacity</td>
<td>Maximum 40% of the capacity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100% of the IOPS target</td>
<td>0% of the IOPS target</td>
</tr>
<tr>
<td>3 tiers SSD-FC/SAS-NL</td>
<td>At least 5% of the capacity*</td>
<td>Minimum 55% of the capacity</td>
<td>Maximum 40%</td>
</tr>
</tbody>
</table>

* Or minimum disk requirement for SSD

When new data (new VVs or new user space for a thin volume) is created, it will be created in the default CPG defined at volume creation time (best practice is to use the FC tier). Adaptive Optimization will not migrate regions of data to other tiers until the next time the AO configuration is executed. It is, therefore, important that the FC disks have enough performance and capacity to accommodate the performance or capacity requirements of new applications that are provisioned to the system. It is a best practice to size the solutions assuming the NL tier will contribute zero percent of the IOPS required from the solution.
Sizing for SSD tier

It is relatively straightforward to size the FC tier based on the IOPS requirement, but the IOPS delivered by the SSD tier will be based on the number of SSD drives in the array with respect to the overall capacity as well as the application data locality. The following example walks through how much IOPS to expect per SSD in different scenarios for an array sized for 50 TB.

- If SSD tier is less than 5 percent of the total capacity, then it will occupy between 1.25 TB and 2.5 TB of the total capacity. This requirement can be met with 480 GB SSD drives and each drive may only average 550 IOPS. Even though each SSD drive can handle a lot more IOPS, less amount of capacity used by SSD tier will translate to less number of IOPS per drive. This is because there are very few applications that will do very heavy I/O to such a small percentage of their data space.
- If SSD tier is between 5 percent and 10 percent, then it will occupy between 2.5 TB and 5 TB of the total capacity. This requirement can be met with 480 GB or 1.92 TB SSD drives and each drive will get around 1,000 to 1,100 IOPS. The number of IOPS per drive increase as we add more capacity by doubling the drive size.
- If SSD tier is more than 10 percent, then the customer will choose the 1.92 TB or 3.84 TB SSD drives. Each of these drives will provide on an average 2,150 IOPS. This is possible because the SSD capacity has increased by doubling the size of each drive.

Configuring AO using region I/O density reports

Region I/O density reports are used to identify applications and volumes that are suitable for AO as well as to check that AO is moving the regions as expected after the initial implementation. Starting from HP 3PAR OS 3.1.3 MU1, System Reporter license is not needed to produce region I/O density reports. The cumulative region I/O density is most useful when setting up AO and the regular bar chart-based region I/O density is useful to get insights into how AO has performed over time. These reports can be run via the command line or from the HP 3PAR SSMC.

Cumulative region I/O density reports

These reports give a good indication about the locality of data for a particular CPG. These reports can be generated against CPGs and AO configurations; when setting up a configuration for the first time, reports should be generated against single CPGs, as they will help identify CPGs that are a good candidate for an AO configuration. There are two types of cumulative reports: percentage-based and numbers-based. The percentage type report has percent capacity on X-axis and % access rate on Y-axis. Whereas for the numbers-type report the total capacity is on X-axis and total access rate is on Y-axis. Figure 4 gives an example of the percentage-type cumulative region I/O density report.

![Cumulative region I/O density report—percentage](image)
A CPG is a possible candidate for AO if the curve for that CPG is in the left top corner. Such a CPG will serve most of its I/O from a small amount of addressable capacity. The report in figure 4 has two visible curves both in the left top corner. So both are possible candidates for AO. The red colored curve tells that almost 90 percent of the I/O for that CPG is serviced by 5 percent of the capacity. Similarly, the green curve tells that almost 90 percent of the I/O for that CPG is serviced by 1 percent of the capacity.

From this report, it seems that it will help if we add the two CPGs to an AO configuration and let AO move this hot capacity to SSD. But, this report is a normalized report based on percentage, so you do not have the actual IOPS or capacity numbers. We first need to find out how much capacity is hot and how much total I/O this small capacity is serving. This information is given by the numbers-type cumulate region I/O density report.

Figure 5 gives an example of the numbers-type report. Here capacity in GiB is on X-axis and access rate I/O/min is on Y-axis. In this report, we see that it will be useful to move the FastClass.cpg into an AO configuration that uses SSDs for tier 0. However, the OS_Images, which was on the left top corner in figure 4, has very small I/O and so it will not be beneficial to move this CPG to an AO configuration.

**Figure 5.** Cumulative region I/O density report—numbers

![Cumulative region I/O density](image)

**Region I/O density reports**

Region I/O density is an indication of how busy a region of data is. A region I/O density report is a set of histograms with I/O rate buckets. The space GiB histogram shows the capacity in GiB for all regions in the I/O rate buckets. The I/O/min histogram shows the total IO's/min for the regions in the I/O rate buckets. The example results in figure 6 is describe region I/O density after HP 3PAR Adaptive Optimization has run for a while.

Both charts are histograms, with the X-axis showing the I/O rate density buckets; the busiest regions are to the right and the idiest regions are to the left. The chart on the left shows on the Y-axis the capacity for all the regions in each bucket, while the chart on the right shows on the Y-axis the total IO's/min for the regions in each bucket. As shown in the charts, the FC tier (tier 1) occupies very little space but absorbs most of the I/O accesses, whereas the nearline tier (tier 2) occupies most of the space but absorbs almost no accesses at all. This environment is a good fit for Adaptive Optimization.
There is also a cumulative histogram that adds up all the bucket values from left to right. Figure 7 shows the cumulative region I/O density report for the same AO configuration as shown in figure 6.

Using these charts together, we can get a view into how densely I/O values are grouped across the data space and determine how large different tiers of storage should be. These reports are most useful when run against an AO configuration as they display the distribution of space and I/O across all tiers in the AO configuration.

From HP 3PAR OS version 3.1.3 onwards, you can display region density report for each VV in an AO configuration or CPG. This report has two use cases: to find which VVs are best suited for AO and to find if certain VVs need a different AO policy. Certain VVs could have a different I/O profile than the rest of the VVs in the Adaptive Optimization configuration (AOCFG). You might find out that some VVs are not well suited for Adaptive Optimization, as they do not have enough locality of data. Using the per VV region density report, you can now find such VVs and move them to a CPG outside the AO configuration using Dynamic Optimization. Figure 8 shows the region density reports for some of the VVs in an AO configuration with two tiers. As shown in the chart, the volumes using the FC tier have more I/O but consume very little space. Whereas the volumes using the NL tier have very little I/O but consume most of the space.
Creating and managing AO

Simple administration is an important design goal, which makes it tempting to automate Adaptive Optimization completely. The administrator need not configure anything. However, analysis indicates that some controls are, in fact, desirable for administration simplicity. Since HP 3PAR StoreServ Storage is typically used for multiple applications—often for multiple customers—HP allows administrators to create multiple Adaptive Optimization configurations so that they can use different configurations for different applications or customers. Figure 9 shows the configuration settings for an Adaptive Optimization configuration.

An AO configuration is made up of up to three tiers. Each tier corresponds to a CPG. In figure 9, SSD CPG is used for tier 0, RAID 1 FC CPG is used for tier 1, and RAID 5 FC CPG is used for tier 2. The warning and limits displayed in figure 9 have to be configured for each CPG individually by editing the CPG. More details on how to edit the warning and limit for the CPG is available in the HP 3PAR SSMC user guide.

Make sure to define tier 0 to be on a higher performance level than tier 1, which in turn should be higher performance than tier 2. For example, you may choose RAID 1 with SSDs for tier 0, RAID 5 with FC drives for tier 1, and RAID 6 with NL or SATA drives for tier 2.
Best practices encourage you to begin your Adaptive Optimization configurations with your application CPG starting with tier 1. For example, tier 1 could be CPG using your FC or SAS physical disks. This allows you to add both higher and lower tier capabilities at a later date. If you don’t have a higher or lower tier, you can add either or both at a later date by using a new CPG, such as tier 0 using SSDs or tier 2 using NL. Or, you could have CPG tiers with RAID 1 or RAID 5 and RAID 6. The main point is that you should begin with middle CPG tier 1 when configuring Adaptive Optimization with your application.

It is also important to specify the schedule when a configuration will move data across tiers along with the measurement duration preceding the execution time. This allows the administrator to schedule data movement at times when the additional overhead of that data movement is acceptable (for example, non-peak hours). You can also set the schedule as to when Adaptive Optimization should stop working before the next measurement period.

The following data movement modes are available:

1. **Performance mode**—biases the tiering algorithm to move more data into faster tiers.
2. **Cost mode**—biases the tiering algorithm to move more data into the slower tiers.
3. **Balanced mode**—is a balance between performance and cost.

The mode configuration parameter does not change the basic flow of the tiering analysis algorithm, but rather it changes certain tuning parameters that the algorithm uses.

If the SSD tier is used only for AO, then it is recommended to disable the raw space alert for the SSD tier. AO manages the amount of space used on each tier and it can fill up over 90 percent of a small SSD tier. But, this will generate alerts about raw space availability. These alerts can be ignored if the SSD tier is used only by AO. These alerts can be disabled using the following command:

```bash
setsys –param RawSpaceAlertSSD 0
```

### Scheduling AO policies

After creating the AO configuration, the configuration has to be scheduled to run on a regular interval. Figure 10 shows how to create the scheduled task for an AO configuration. The max run time specifies how long AO should move the regions once it is scheduled. The measurement interval specifies the duration for which the AO configuration should be analyzed for data movement. Schedule specifies when to run the AO task; the task can recur daily, once, multiple times daily, or advanced. Figure 11 shows details of the scheduled tasks.

**Figure 10. Schedule AO configuration**

![Schedule AO Configuration](image)
Figure 11. Start AO task summary

<table>
<thead>
<tr>
<th>Name</th>
<th>Resource</th>
<th>Date</th>
<th>State</th>
<th>User</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move Regions (AO: AO, CPG from T1 to T2)</td>
<td>not set</td>
<td>Jul 2, 2015 12:10:23 PM PDT</td>
<td>Completed 3m24s</td>
<td>3parsys</td>
</tr>
</tbody>
</table>

The AO Space Moved Report shown in figure 12 provides details about the capacity migrated across tiers. This report can be used to verify the effectiveness of the AO configuration and should be checked regularly.

Figure 12. AO Space Moved Report

Freeing unused space from CPGs using **compact** operation

The `startao` command has a `-compact` auto option that runs `compactcpg` only if one or more of the following conditions are met (otherwise, `compactcpg` is not run):

1. There is unused space in the CPG and the current allocated space is above the warning limit.
2. The unused space in the CPG is more than a certain fraction (25 percent) of the CPG space. This is the total unused space across user, snapshot, and admin space.
3. The space available to grow the CPG (i.e., free chunklets for the CPG) is less than four times the CPG growth increment. This can be examined by comparing the `LDFree` output of `showspace -cpg` with `showcpg -sdg`.

If desired, `compactcpg` can be scheduled separately from `startao` to always run `compactcpg`. 
Using \textit{min\_iops} option

From HP 3PAR OS version 3.1.3 onwards, a new option \textit{-min\_iops} was added to the startao command. If this option is used, AO will not execute the region moves if the average LD IOPS for the AOCFG during the measurement interval is less than \textit{-min\_iops} value specified with this option. If \textit{-min\_iops} is not specified, then the default value is 50. The \textit{-min\_iops} option can be used to prevent movement of normally busy regions to slower tiers if the application associated with an AOCFG is down or inactive (i.e., total LD IOPS for the AOCFG is less than \textit{-min\_iops}) during the measurement interval. This ensures that when the application resumes normal operation, its busy regions are still in the faster tiers.

Removing a tier from an AO configuration

A CPG will not be tracked by AO after it is removed from the AO configuration and any existing data in that CPG will not be moved to other CPGs. Hence, it’s important to move all data out of a tier before it is removed from the AO configuration. The easiest way to drain the tier is by setting the CPG warning level to 1. This will hint AO to not move any new data to this CPG and to move existing data from this tier to the other CPGs.

Configuring the SSD pool to be used by multiple CPGs

The SSD tier may be used by CPGs that are not part of an AO configuration. In this case it is a best practice to set warning limits on the SSD AO configurations or tier limits when running 3.2.2. This helps prevent the system from allocating all available free space to the AO tier 0 CPG.

For example, if a system has 10 TB of usable SSD capacity, the user can create an \textit{AO\_tier0\_SSDr5} and \textit{Dedupe\_SSDr5} and set a warning limit on the AO CPGs to 5 TB so no more than 5 TB of usable SSD will be allocated for that AO configuration. For the other \textit{Dedupe\_SSDr5} CPG, the users can decide to set a warning or limits on the CPGs or VVs depending on how they want to manage the shared pool of space and when they want to be alerted that the pool is running short of space.

Using AO configuration with VVsets

Starting from HP 3PAR OS 3.2.1, a customer can configure AO region moves at a more granular level than all the volumes within a CPG. This can be done by scheduling region moves at the virtual volume set (VVset) that serves as a proxy for an application.

This option is useful if a customer wants to have multiple volumes within the same AO Configuration CPG and schedule data movements at different times with different schedules and modes.

VVsets is an autonomic group object that is a collection of virtual volumes. VVsets help to simplify the administration of VVs and reduce human error. An operation like exporting a VVset to a host will export all member volumes of the VVset. Adding a volume to an existing VVset will export the new volume automatically to the host or the host set the VVset is exported to.

VVsets have a number of use cases beyond reducing the administration of their volume members, such as enabling simultaneous point-in-time snapshots of a group of volumes with a single command. There is no requirement that volumes in a VVset have to be exported to hosts for an AO configuration to work at the VVset level.

An environment may have workloads with two distinct IO profiles; profile\_1 business hours and profile\_2 non\_business\_hours.

For profile\_2 only a subset of volumes are used and these applications can be logically assigned to a VVset called \textit{AO\_Profile\_2}.

In order to make the volumes layout ready for out of business hours activity, one can schedule an Adaptive Optimization run that will target only the volumes that belong to VVset \textit{AO\_Profile\_2}, without impacting other volumes that are part of the AOCFG:

- 8:00 PM—Run AO against \textit{AO\_Profile\_2} measurement from last night, (12:00 AM to 5 AM). Maximum run for 4 hours.
- 5:00 AM—Run AO on entire AO config to prepare for day activity, (9:00 AM to 7:00 PM). Maximum run for 4 hours.

Alternatively, instead of running the AO run against all volumes, it’s possible to create VVset \textit{AO\_Profile\_1} and run the AO only against that VVset. Without this feature, we would have needed two different AO configurations and hence different set of CPGs.

A new option \textit{-vv} that takes a comma-separated list of VV names or set patterns has been added to three commands \textit{startao}, \textit{srgiodensity}, and \textit{sraomoves}. 
If -vv option is used with startao, the command then runs AO on the specified AOCFG, but only to matching VVs. This allows a user to run AO on separate applications using the same AOCFG.

For srgiodensity, the -vv option can be used to filter the volumes in the report when used along with the -withvv option.

And, the -vv option for sracomoves will allow the user to see the region moves associated with individual VVs. The min_iops option described in the previous sections can be used at a more granular level when implementing AO with VVsets.

This option is configurable via command line only; for more information refer to the HP 3PAR Command Line Interface reference guide.

Following example schedules AO moves every day at 23:00; this command will only move regions belonging to VVs in VVset AO_Profile_2.

```
ccreatesched "startao -btsecs -1h -vv set:AO_Profile_2 AOCFG" "0 23 * * *"
AO_out_of_business_hours
```

**Adaptive Optimization with Remote Copy**

Although Adaptive Optimization coexists and works with replicated volumes, it’s important that users take into account the following considerations when using AO on destination volumes for HP 3PAR Remote Copy. The I/O pattern on the Remote Copy source volumes is different from the I/O pattern on the Remote Copy target volumes; hence, AO will act differently when moving regions on the primary volume when compared to the data movement on the destination volume.

- With synchronous replication and asynchronous streaming replication, the target volumes will receive all the write I/O requests; however, the target volume will see none of the read requests on the source volume. AO will only see the write I/O on the target volumes and will move the regions accordingly. If the application is read intensive, then the hot regions will be moved to tier 0 on the source array but will not be moved to tier 0 on the target array. Upon failover to the remote array, the performance of the application may be impacted as the region that was hot and in tier 0 earlier (on the source array) may not be in tier 0. This scenario also applies to HP 3PAR Peer Persistence.

- With periodic asynchronous replication mode, write I/O operations are batched and sent to the target volumes periodically; hence, the I/O pattern on target volumes is very different from that on source volumes. If AO is configured on these volumes, the resulting sub-tiering will be different from the sub-tiering done on the source volumes.

Considering the above scenarios, data layout on the target volume may be different on the secondary system. And in case of a failover, performance levels may differ. In these cases, a good way to rebalance the hottest data to the upper tiers is to manually run the AO policy 30 minutes after failover (so the system has enough I/O statistics for the new I/O profile to identify the right regions to move). When AO and RC are used, it is recommended to configure a portion of the SSD capacity for AFC and to size the target volumes on a single tier that is capable of handling all the I/O requirements of the application after failover.
Use case

Accelerating workloads by adding SSDs

In a two-tier configuration with SSD and fast-class drives, AO provides performance acceleration by lowering the average svctimes. In the provided example, a customer had an array with 120 FC drives with a backend IOPS of over 30,250. Each drive was getting around 250 IOPS and had a service time of over 20 ms. Figure 13 shows the IOPS and service times reported by the statport command. The customer engaged HP to help them resolve their performance problem.

Figure 13. Back-end IOPS and service time before implementing AO

From the HP 3PAR StoreServ CLI command statport, it was clear that the array was getting a lot more I/O than what it was sized for. And, they had not yet completed the migration activity and expected a lot more load on the array. The customer had two options: either add more FC drives or analyze and check if adding SSD drives and enabling AO will help. Figure 14 shows the region I/O density report for the FC CPG.

Figure 14. Region I/O density report

Looking closer at the charts, we see that about 4 TB (left side) of capacity is driving the bulk of I/O (right side) = 500,000 I/O/min.

The customer decided to add 24 SSD drives and use AO. Figure 15 shows the statpd report after enabling AO. The SSD drives are now serving around 1,000 IOPS each and the FC drives are now serving around 200 IOPS each. And, figure 16 shows how service times reduced from over 20 ms to less than 10 ms.
Figure 15. PD IOPS after implementing AO

![Figure showing PD IOPS after implementing AO]

Each SSD drive servicing around 1000 IOPS after enabling AO

IOPS per FC drive is reduced to around 200

So, by enabling AO, the customer was able to reduce the IOPS on FC drives, thereby improving the response time. The I/O profile also had good locality of data that helped AO to put the hot regions in SSD tier.

Figure 16. Service time reduction as result of AO

![Figure showing service time reduction as result of AO]

Service time

Read Service Time
Write Service Time
Total Service Time

Time

2013-10-30 17:00:22
2013-10-30 05:00:10
2013-10-30 03:00:45
2013-10-30 01:00:30
2013-10-29 23:00:15
2013-10-29 21:00:00
2013-10-29 19:00:45
2013-10-29 17:00:30
2013-10-29 15:00:15
2013-10-29 13:00:00
2013-10-29 11:00:45
2013-10-29 09:00:30
2013-10-29 07:00:15
2013-10-29 05:00:00
2013-10-29 03:00:45
2013-10-29 01:00:30
2013-10-28 23:00:15
2013-10-28 21:00:00
2013-10-28 19:00:45
2013-10-28 17:00:30
2013-10-28 15:00:15
2013-10-28 13:00:00
2013-10-28 11:00:45
2013-10-28 09:00:30
2013-10-28 07:00:15
2013-10-28 05:00:00
2013-10-28 03:00:45
2013-10-28 01:00:30
2013-10-27 23:00:15
2013-10-27 21:00:00
2013-10-27 19:00:45
2013-10-27 17:00:30
2013-10-27 15:00:15
2013-10-27 13:00:00
2013-10-27 11:00:45
2013-10-27 09:00:30
2013-10-27 07:00:15
2013-10-27 05:00:00
2013-10-27 03:00:45
2013-10-27 01:00:30
2013-10-26 23:00:15
2013-10-26 21:00:00
2013-10-26 19:00:45
2013-10-26 17:00:30
2013-10-26 15:00:15
2013-10-26 13:00:00
2013-10-26 11:00:45
2013-10-26 09:00:30
2013-10-26 07:00:15
2013-10-26 05:00:00
2013-10-26 03:00:45
2013-10-26 01:00:30
2013-10-25 23:00:15
2013-10-25 21:00:00
2013-10-25 19:00:45
2013-10-25 17:00:30
2013-10-25 15:00:15
2013-10-25 13:00:00
2013-10-25 11:00:45
2013-10-25 09:00:30
2013-10-25 07:00:15
2013-10-25 05:00:00
2013-10-25 03:00:45
2013-10-25 01:00:30
Lowering cost per GB by configuring a three-tier configuration with SSD, FC, and NL

This section describes the real benefits that a customer has from using HP 3PAR Adaptive Optimization. The customer had a system with 96 300 GB 15k rpm FC drives and 48 1 TB 7.2k rpm NL drives. The customer had 52 physical servers connected and running VMware® with more than 250 virtual machines (VMs).

The workload was mixed (development and QA, databases, file servers, and more) and they needed more space to accommodate many more VMs that were scheduled to be moved onto the array. However, they faced a performance issue: they had difficulty managing their two tiers (FC and NL) in a way that kept the busier workloads on their FC disks. Even though the NL disks had substantially less performance capability (because there were fewer NL disks and they were much slower), they had larger overall capacity.

As a result, more workloads were allocated to them and they tended to be busier while incurring long latencies. The customer considered two options: either they would purchase additional 96 FC drives, or they would purchase additional 48 NL drives and 16 SSD drives and use HP 3PAR Adaptive Optimization to migrate busy regions onto the SSD drives. They chose the latter and were pleased with the results (illustrated in figure 17).

Figure 17. Improved performance after Adaptive Optimization

Before HP 3PAR Adaptive Optimization, as described in the charts—and even though there are fewer NL drives—they incur greater IOPS load than the FC drives in aggregate and consequently have very poor latency (~40 ms) compared with the FC drives (~10 ms). After HP 3PAR Adaptive Optimization was executed for a little while, as shown in figure 17, the IOPS load for the NL drives dropped substantially and the load was transferred mostly to the SSD drives.

HP 3PAR Adaptive Optimization moved approximately 33 percent of the IOPS workload to the SSD drives even though that involved moving only one percent of the space. Back-end performance improved in two ways: the 33 percent of the IOPS that were serviced by SSD drives got very good latencies (~2 ms), and the latencies for the NL drives improved (from ~40 ms to ~15 ms). The front-end performance also improved significantly as most of the frequently accessed data was serviced by SSD tier. Moreover, the investment in the 16 SSD drives permitted them to add even more NL drives in the future, because the SSD drives have both space and performance headroom remaining.
Summary

HP 3PAR Adaptive Optimization is a powerful tool for identifying how to configure multiple tiers of storage devices for high performance. Its management features can deliver results with reduced effort. As in all matters concerning performance, your results may vary but proper focus and use of HP 3PAR Adaptive Optimization can deliver significant improvements in device utilization and total throughput.

Learn more at hp.com/go/3PARStoreServ