UNDERSTANDING STORAGE RESOURCE POOLS ON VMAX3 AND VMAX ALL FLASH STORAGE ARRAYS

Configuration recommendations and best practices for single or multiple SRPs

ABSTRACT

This white paper explains the concept of Storage Resource Pools (SRPs) and how they are implemented in Dell EMC VMAX3 and VMAX All Flash storage arrays. It provides the information required to make decisions that follow best practices regarding the configuration of SRPs.

September, 2016

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Organizations around the globe need IT infrastructures that can deliver instant access to the huge volumes of data intrinsic to traditional transaction processing/data warehousing and to a new generation of applications built around the world of social, mobile, cloud, and big data. Dell EMC is redefining Data Center Cloud Platforms to build the bridge between these two worlds to form the next generation Hybrid Cloud.

Essential to this is the ability to quickly and efficiently allocate storage resources. VMAX3 and VMAX All Flash arrays are pre-configured in the factory with Virtual Provisioning Pools from which thin devices can be quickly and easily assigned to hosts and applications. All physical drives in the array are placed in Storage Resource Pools (SRPs), which provide the physical storage for thin devices that are presented to hosts using masking views.

SRPs are managed by Fully Automated Storage Tiering (FAST) and require no configuration operations to be performed by the storage administrator. This simplifies the initial configuration of new VMAX3 and VMAX All Flash arrays significantly and greatly reduces the time to I/O. Storage capacity is monitored at the SRP level and RAID considerations and thin device binding are no longer issues of concern for the storage administrator when creating and assigning devices. This is because all devices are available as soon as they are created and RAID protection is a function of the SRP itself and not a property of an individual device. This new array design and method of configuring and allocating storage greatly reduces the amount of time and effort required to manage and monitor the VMAX3 and VMAX All Flash array.

AUDIENCE

This white paper is intended for anyone who needs to understand the concept of SRPs and how they relate to array configuration and storage provisioning in the VMAX3 and VMAX All Flash. This document specifically targets Dell EMC customers, sales, and field technical staff who are designing and implementing a VMAX3 or VMAX All Flash storage array or are considering a future implementation.

UNDERSTANDING VMAX3 SRPS AND RELATED ELEMENTS

SRPs are comprised of one or more data pools, which contain the pre-configured data (or TDAT) devices that provide storage for the thin devices (TDEVS) that are created and presented to hosts or applications. Physical storage for the TDAT devices is provided by disk groups, which contain physical drives. In order to understand SRPs and the role they play in the configuration and management of the VMAX3, it is important to understand these elements, which are the underlying entities that comprise SRPs.

Data Pools

A data pool, also known as a thin pool, is a collection of data devices of the same emulation and RAID protection type. All data devices configured in a single disk group are contained in a single data pool. As such, all the data devices are configured on drives of the same technology type, capacity, and, if applicable, rotational speed.

The VMAX3 storage array supports up to 510 data pools. Data pools are preconfigured within the storage array and their configuration cannot be modified using management software.

Disk Groups

A disk group is a collection of physical drives sharing the same physical and performance characteristics. Drives are grouped based on technology, rotational speed, capacity, and desired RAID protection type.

Each disk group is automatically configured with data devices (TDATs) upon creation. A data device is an internal logical device dedicated to providing physical storage to be used by thin devices. All data devices in the disk group are of a single RAID protection type and all are the same size. Because of this, each drive in the group has the same number of hyper volumes (hypers) created on them, with each hyper being the same size. There are 16 hypers configured on each drive.

The VMAX3 storage array supports up to 510 internal disk groups. Disk groups are preconfigured within the storage array and their configuration cannot be modified using management software. Dell EMC Customer Service may add physical drives to a disk group, but drives cannot be removed.

Storage Resource Pools

A Storage Resource Pool (SRP) is a collection of disk groups configured into thin data pools constituting a FAST domain whose performance and reliability is tightly coupled. This means that data movement performed by FAST is done within the boundaries of the SRP. Application data belonging to thin devices can be distributed across all data pools within the SRP to which it is associated. TimeFinder snapshot data and SRDF/A DSE (delta set extension) data is also written to pools within an SRP.

By default, a VMAX3 storage array has a single SRP containing all the configured data pools. This single SRP configuration is appropriate for the vast majority of production environments.

There is no restriction on the combination of drive technology types and RAID protection within an SRP. When moving data between data pools, FAST will differentiate the performance capabilities of the pools based on both rotational speed (if applicable) and RAID protection.

While an SRP may contain multiple data pools, individual data pools can only be a part of one storage resource pool.

The VMAX3 All Flash storage array supports up to five storage resource pools, but requires approval from Dell EMC for configurations containing more than a single SRP. When multiple storage resource pools are configured, one of the storage resource pools must be marked as being the default storage resource pool.

Storage resource pools are preconfigured within the storage array. Drives may be added to existing SRPs and additional SRPs may be configured later, if required, however it is not possible to split an existing SRP into multiple SRPs.

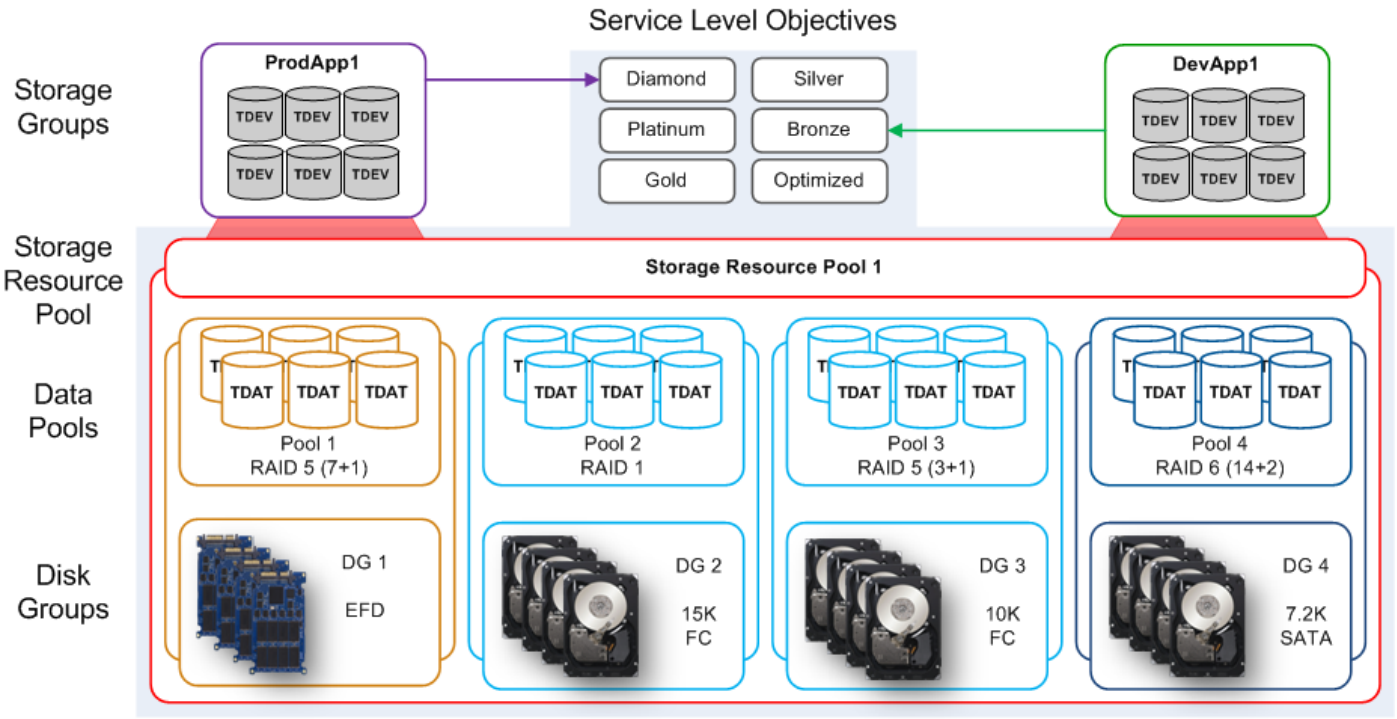

Figure 1. VMAX3 SRP and FAST related elements

UNDERSTANDING VMAX ALL FLASH SRPS AND RELATED ELEMENTS

The concepts regarding SRPs on a VMAX3 also apply to the VMAX All Flash, but there are differences due to the nature of the all EFD array. The All Flash is comprised of one data pools which contains the pre-configured data (or TDAT) devices that provide storage for the thin devices (TDEVS) that are created and presented to hosts or applications. The physical storage for the TDAT devices is provided by a disk group which contains physical flash (or EFD) drives. In order to understand SRPs and the role they play in the configuration and management of the VMAX All Flash, it is important to understand these elements, which are the underlying entities that comprise SRPs.

Data Pools

A data pool, also known as a thin pool, is a collection of data devices of the same emulation and RAID protection type. All data devices configured in a single disk group are contained in a single data pool. As such, all the data devices are configured on drives of the same technology type and capacity.

The VMAX All Flash storage array supports up to 510 data pools. Data pools are preconfigured within the storage array and their configuration cannot be modified using management software.

Disk Groups

A disk group is a collection of physical drives sharing the same physical and performance characteristics. Flash drives are grouped based on technology, capacity, and desired RAID protection type.

Each disk group is automatically configured with data devices (TDATs) upon creation. A data device is an internal logical device dedicated to providing physical storage to be used by thin devices. All data devices in the disk group are of a single RAID protection type and all are the same size. Because of this, each drive in the group has the same number of hyper volumes (hypers) created on them, with each hyper being the same size. There are 16 hypers configured on each drive unless compression is configured. In that case, there will be 64 hypers per physical drive.

The VMAX All Flash storage array supports up to 510 internal disk groups. Disk groups are preconfigured within the storage array and their configuration cannot be modified using management software. Dell EMC Customer Service may add physical drives to a disk group, but drives cannot be removed.

Storage Resource Pools

A Storage Resource Pool (SRP) is a collection of disk groups configured into thin data pools constituting a FAST domain whose performance and reliability is tightly coupled. This means that data movement performed by FAST is done within the boundaries of the SRP. Application data belonging to thin devices can be distributed across all data pools within the SRP to which it is associated. TimeFinder snapshot data and SRDF/A DSE (delta set extension) data is also written to pools within an SRP.

By default, a VMAX All Flash storage array has a single SRP containing all the configured data pools. This single SRP configuration is appropriate for the vast majority of production environments.

There is no restriction on the combination of RAID protection within an SRP. When moving data between data pools, FAST will differentiate the performance capabilities of the pools based on RAID protection.

While an SRP may contain multiple data pools, individual data pools can only be a part of one storage resource pool.

The VMAX All Flash storage array supports up to five storage resource pools, but requires approval from Dell EMC for configurations containing more than a single SRP. When multiple storage resource pools are configured, one of the storage resource pools must be marked as being the default storage resource pool.

Storage resource pools are preconfigured within the storage array. Drives may be added to existing SRPs and additional SRPs may be configured later, if required, however it is not possible to split an existing SRP into multiple SRPs.

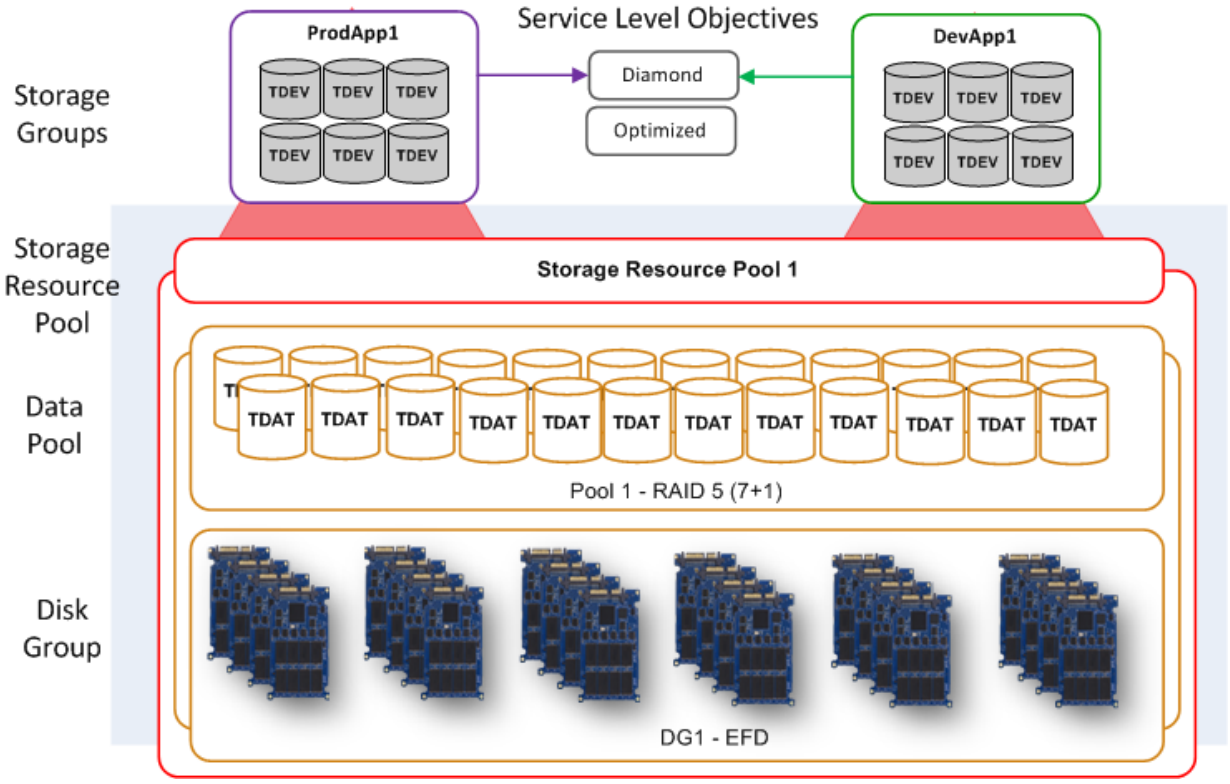

Figure 1. VMAX All Flash SRP and FAST related elements

FAST AND ADDITIONAL RELATED ELEMENTS

It is also necessary to understand other elements related to Fully Automated Storage Tiering (FAST) in order to understand how SRPs relate to storage allocation and performance on both VMAX3 and VMAX All Flash arrays.

Service Level Provisioning

VMAX3 and VMAX All Flash radically simplifies storage provisioning by eliminating the need to manually assign physical storage resources to hosts and applications. Instead, the storage performance required for an application is specified during the provisioning process by associating a pre-defined service level objective to the application through the storage group containing its thin devices. Application data is then dynamically allocated by FAST across storage resources of differing performance characteristics to achieve the overall performance required by the application.

This ability to provision to service levels is inherently available on all VMAX3 and VMAX All Flash storage arrays because all arrays are virtually provisioned with FAST permanently enabled.

Virtual Provisioning

Virtual Provisioning allows an increase in capacity utilization by enabling more storage to be presented to a host than is physically consumed and by allocating storage only as needed from a shared virtual pool. Virtual Provisioning also simplifies storage management by making data layout easier through automated wide striping and by reducing the steps required to accommodate application growth.

Virtual Provisioning uses a type of host-accessible device called a virtually provisioned device, also known as a thin device (TDEV), which does not need to have physical storage allocated at the time the devices are created and presented to a host. All thin devices are associated with the default SRP upon creation. The physical storage that is used to supply storage capacity to thin devices comes from data (TDAT) devices within an SRP. These data devices are dedicated to the purpose of providing the actual physical storage used by virtually provisioned devices.

When data is written to a portion of the virtually provisioned device, the VMAX3 and VMAX All Flash array allocates physical storage from the pool and maps that storage to a region of the virtually provisioned device including the area targeted by the write. These allocation operations are performed in small units of storage called virtually provisioned device extents, which are 1 track (128 KB) in size. These extents are also referred to as chunks.

When data is read from a virtually provisioned device, the data being read is retrieved from the appropriate data device in the storage resource pool where the data was written. When more storage is required to service existing or future virtually provisioned devices, data devices can be added to existing data pools within the SRP.

Storage groups

A storage group is a logical collection of VMAX thin devices that are to be managed together, typically constituting a single application. Storage groups can be associated with a storage resource pool, a service level objective, or both. Associating a storage group with an SRP defines the physical storage to which data in the storage group can be allocated. The association of a service level objective defines the response time objective for that data.

By default, storage groups will be associated with the default storage resource pool and managed under the Optimized SLO. A storage group is considered 'FAST managed' when it has an explicit SLO or SRP assigned to it. When a storage group is a parent storage group with an associated child group, the SLO or SRP are associated with the child group. Parent storage groups cannot have SLO or SRPs associated with them.

Devices may be included in more than one storage group, but may only be included in one storage group that is 'FAST managed'. This ensures that a single device cannot be managed by more than one service level objective or have data allocated in more than one storage resource pool. Individual thin devices cannot have a SLO or SRP assigned to them.

The VMAX3 and VMAX All Flash storage array supports up to 16,384 storage groups, each of which may contain up to 4,096 devices. Each storage group name may be comprised of up to 64 alphanumeric characters, hyphens (-), and underscores (_). Storage group names are not case sensitive.

VMAX3 Service Level Objectives

A service level objective (SLO) defines an expected average response time target for a storage group. By associating a service level objective to a storage group that contains devices from an application, FAST automatically monitors the performance of the application and adjusts the distribution of extent allocations within a storage resource pool in order to maintain or meet the response time target.

There are six available service level objectives, five of which, Diamond, Platinum, Gold, Silver, and Bronze, have defined expected average response time targets. There is also an additional SLO called Optimized, which has no explicit response time target associated with it.

By default, all data in the VMAX3 and VMAX All Flash storage array is managed by the Optimized SLO. However, an explicit response time target may be set for an application by associating it with one of the other five service level objectives. The actual response time of an application associated with each service level objective will vary based on the observed workload and will depend on average IO size, read/write ratio, the use of local or remote replication, along with the availability of other resources within the array.

A detailed description of the available service level objectives is available in FAST and VMAX3 and VMAX All Flash documentation available at support.emc.com.

VMAX All Flash Service Level Objectives

There are two available service level objectives. One of them, Diamond, has defined expected average response time targets. There is also an additional SLO called Optimized, which has no explicit response time target associated with it.

By default, all data in the VMAX All Flash storage array is managed by the Optimized SLO. However, because the VMAX3 contains all flash, or EFD devices, bit the Optimized setting and the Diamond setting will be, in effect, the same setting. The actual response time of an application associated with each service level objective will vary based on the observed workload and will depend on average IO size, read/write ratio, the use of local or remote replication, along with the availability of other resources within the array.

A detailed description of the available service level objectives is available in FAST and VMAX documentation available at support.emc.com.

CONFIGURATIONS WITH A SINGLE SRP

The default VMAX3 and VMAX All Flash system configuration contains a single storage resource pool. For the majority of environments, a single SRP system will be the best configuration for both performance and ease of management.

Advantages of a single SRP system

The advantage of a single SRP system is the simplicity with which storage creation, allocation, and management can be performed. This ease of use inherent in the VMAX3 and VMAX All Flash, which was one of the main goals in the design of the arrays, is most easily recognized and experienced with a single SRP configuration. With a single SRP and devices under FAST control, the storage administrator can simply create the required devices and add them to a storage group with the appropriate Service Level Objective. Once that is done, the physical location of the data is determined by FAST, requiring no further management by the storage administrator to ensure optimal availability and performance. Both mainframe, and open systems can be configured in a single SRP, either sharing physical disk groups or with isolated disk groups for each emulation type.

Because of the ease of management inherent in single SRP systems, Dell EMC recommends that a single SRP be used in all cases that don't have specific requirements for multiple SRPs.

Disadvantages of a single SRP system

A single SRP system does not allow for physical isolation of drives. This may be required for regulatory, performance or fault tolerance reasons.

Specific performance requirements may be another reason that multiple SRPs may be considered. Performance requirements can generally be addressed with a single SRP by using FAST and choosing an appropriate SLO, however, specific agreements between the array owner and tenants, company best practices, or other requirements may lead to a multiple SRP configuration.

A potential disadvantage with single SRP systems is the risk of multiple drive failures affecting the entire system. This risk is extremely small and is almost completely mitigated by the VMAX3 and VMAX All Flash sparing mechanism, which begins the process of replacing a failing physical drive with a permanent spare as soon as the HYPERMAX OS determines that it is likely to fail. In systems handling an unusually high I/O load, spare rebuild times are increased. This increases the chance, however small, that a second drive failure in the same SRP will cause an outage. Using RAID6 as the local RAID protection scheme when appropriate, using large drives, and configuring SRDF remote replication will mitigate these concerns, however, if SRDF is not being used and the disk group is large, multiple SRPs may be warranted.

Single SRP systems may not be acceptable for reasons that are not technical. For example, multiple, isolated pools of physical storage may be a requirement stated in an RFP that all competing storage vendors must be able to deliver on in order to bid.

CONFIGURATIONS WITH MULTIPLE SRPs

While the vast majority of environments will benefit from a single SRP configuration, there are certain user, regulatory, or business requirements that can only be met with multiple SRPs.

Advantages of a multiple SRP system

Despite not being the default configuration, multiple SRP systems offer some benefits over single SRP systems for specific use cases. Multiple SRP systems may be considered in multi-tenant situations where isolation of workload or dedicated physical drives is required. This segregation may be desired to prevent a tenant, who shares a single SRP with other tenants, from assigning high performing SLOs for multiple applications thereby potentially causing the performance to decline for others who share the SRP. Multiple SRPs will allow the physical disks to be isolated, but complete physical segregation of a specific SRP is not possible because VMAX3 and VMAX All Flash hardware, aside from the physical drives, will always be shared across the array regardless of the configuration of the SRPs.

If a configuration is large enough that a single SRP will exceed the maximum recommended disk group size, multiple SRPs may be needed. Configurations requiring SRPs with an unusually large amount of capacity may simply be the result of a large production environment, or they may be related to other things such as particular local replication requirements. For example, physical separation between clone source devices and clone targets may be required in certain circumstances, such as when the space needed by clone targets is large enough that the number of devices required in a single SRP would violate the maximum recommended SRP size. It will also protect against certain user errors, such as an administrator accidentally oversubscribing the source SRP, leaving the target pool without the required space to create clone targets.

The need to segregate drives or data to adhere to legal requirements is a common and valid reason why multiple SRP array configurations may be adopted. Though things like DAEs, power, and engines will always be shared within the array, physical drives can be segregated to meet government or industry mandated physical data separation.

Spindle isolation may also be required for performance reasons. Depending on the particular configuration, extreme performance requirements may require separate SRPs. For example, a VMAX3 configuration may be designed using a small number of flash drives with the remaining physical drives being 10k or 15k RPM with RAID1 protection in order to satisfy extreme performance requirements. A multiple SRP environment may also be warranted with certain operating systems because of similar high performance needs. For example, SRPs for use with IBM i (formerly AS/400) may be designed in this way to isolate disk resources from what is being used by other operating systems attached to the array.

Disadvantages of a multiple SRP system

Two of the causes of this increased management complexity can be viewed as disadvantages over a single SRP system. Firstly, application data cannot span multiple SRPs, which forces the storage administrator to be concerned with choosing an appropriate SRP for each application. Performance planning must be done on each individual SRP in a multiple SRP system instead of on only a single SRP that encompasses the entire array. This means that the administrator must plan ahead of time for any possible I/O bursts and for the maximum required performance for each application based on the SRP that it will be assigned to. This is much more time-consuming for the administrator than it would be to simply assign an SLO to a storage group in a single SRP system and let FAST handle any required moves to relocate busy extents onto a higher performing storage tier. Secondly, with more than one SRP, FAST optimization is limited because FAST can only make performance-based extent moves within an SRP, not between them. This means that if storage tiering is going to be used within an array, each SRP must have multiple storage tiers if the data within each SRP is to be managed by FAST. This is not the most efficient or cost-effective way to manage a VMAX3 and VMAX All Flash array.

Smaller SRPs can also be an issue in and of themselves. This is because SRPs containing lower spindle counts can potentially lead to reduced performance unless they are large enough for the data to be spread widely enough across the physical drives that comprise the disk group.

Another disadvantage for the storage administrator is the need to monitor and manage available capacity in multiple SRPs. Having a multiple SRP system with the same capacity as a single SRP system increases the possibility of running out of space in a given SRP because each SRP will contain a smaller amount of capacity leaving a greater chance that an SRP may not have enough free space to satisfy extent allocation for new or existing volumes.

Multiple SRP configurations will also require more physical drives than similarly sized configurations using a single SRP. This is because adequate spares will be required in each SRP to properly protect against physical drive failure. This holds true for all drive types in each SRP including flash drives. This can be a significant additional expense depending on the configuration of the SRPs and how many disk groups exist in each.

One of the major advantages of the Symmetrix array throughout the history of the product has been the ability of the array to monitor itself and keep itself healthy. This has included things like proactive dial-homes, cache scrubbing, sparing, and many other functions that allow the Symmetrix to operate at maximum reliability and efficiency. The VMAX3 and VMAX All Flash keeps itself healthy in the same way earlier generations of Symmetrix arrays did. One of these ways that is unique to the VMAX3 and VMAX All Flash is the ability of the HYPERMAX OS to promote data outside of the rules of a specific SLO, if necessary. If moving data outside of the chosen SLO is required to keep the array healthy, the data will be moved, however, because data cannot be moved between SRPs, this function is limited to movement only within the SRP in which the data resides. This limits where the data can be placed in a multiple SRP system and could prevent the array from placing the data in the optimal location.

Regardless of the reason for configuring them, in addition to the above disadvantages, multiple SRP systems result in additional management complexity over single SRP systems, at least in the area of capacity allocation, capacity management, and performance.

Note: Multiple SRP configurations are not supported without approval from VMAX engineering. Dell EMC recommends discussing any proposed multiple SRP configuration with local Dell EMC field representatives.