EMC VMAX3 FAMILY FEATURE OVERVIEW — A DETAILED REVIEW FOR OPEN AND MAINFRAME SYSTEM ENVIRONMENTS

ABSTRACT

This white paper describes the features that are available for the EMC® VMAX3™ Family storage systems. These features are available when running HYPERMAX OS 5977 and Solutions Enabler V8. Throughout this document, the term EMC VMAX Family is applicable to all VMAX 100K, VMAX 200K, and VMAX 400K.

March, 2016
## Table of Contents

**Executive Summary** ................................................................................................................................. 5

**Audience** .................................................................................................................................................. 5

**Introduction** .............................................................................................................................................. 6
- VMAX3 hardware models ............................................................................................................................... 6
- HyperMAX OS ............................................................................................................................................. 7
- Data at Rest Encryption (D@RE) ................................................................................................................ 7

**Management Software** ............................................................................................................................ 8
- Solutions Enabler ....................................................................................................................................... 8
- Unisphere® for VMAX .................................................................................................................................. 8
- Unisphere 360™ .......................................................................................................................................... 11
- Unisphere Planning Section .......................................................................................................................... 11
- VMAX REST API ........................................................................................................................................ 13

**Open Systems New Features** .................................................................................................................. 14
- Virtual Provisioning in VMAX3, Pre-configured arrays, Storage Resource Pools .................................. 14
- 128 KB track size ....................................................................................................................................... 14
- Device Expansion ....................................................................................................................................... 15
- VMAX3 FAST.............................................................................................................................................. 15
- Service Level Provisioning ............................................................................................................................ 15
- FAST External - FAST.x™ .......................................................................................................................... 19
- FAST.X™ Incorporation ............................................................................................................................. 21
- Database Storage Analyser (DSA) and FAST Hinting ............................................................................... 21
- TimeFinder SnapVX .................................................................................................................................... 24
- Enhanced SRDF ......................................................................................................................................... 25
- SRDF/Metro.................................................................................................................................................. 26
- Integrated Data Protection with ProtectPoint .............................................................................................. 27
- Embedded NAS (eNAS) .............................................................................................................................. 28
- Virtual Data Mover synchronous replication ............................................................................................. 29
- The front-end port map and recommended practices ............................................................................ 29
- Open Systems Scalability ............................................................................................................................ 29
Enhancement to group management features ...................................................... 30
Host I/O limits per storage group ................................................................. 31
VMware VAAI support ........................................................................... 31
VMware VASA support .......................................................................... 32

MAINFRAME SUPPORT ............................................................................ 32
Mainframe Enabler ............................................................................... 32
GDDR .................................................................................................... 32
IBM z Systems Compatibility Enhancements ........................................... 32

HARDWARE ENHANCEMENTS .................................................................. 33
VMAX3 engines ................................................................................... 33
Multi-Core emulation ....................................................................... 33
Dual MMCS .......................................................................................... 34
Dynamic Virtual Matrix/Infiniband Fabric .............................................. 34
New director emulations ................................................................. 35
Vault to FLASH .................................................................................. 35
6 Gb/s SAS back-end/drive infrastructure ........................................... 36
16 Gb/s FC front-end support ............................................................ 36
16 Gb/s FICON and zHPF Support ....................................................... 36
Ultra dense Disk Array Enclosure support and mixing ......................... 36
Local RAID ......................................................................................... 37
Dense configurations ......................................................................... 37
Bay dispersion .................................................................................... 38
Third-Party racking ........................................................................... 39

CONCLUSION .......................................................................................... 39
REFERENCES ........................................................................................... 40
EXECUTIVE SUMMARY

Organizations around the globe need IT infrastructures that can deliver instant access, at all times, to the massively increasing volumes of data associated with traditional transaction processing and big data use cases such as Data warehousing, OLAP, and so on. At the same time, these organizations are trying to understand how the new generation of applications, built around the world of social, mobile, cloud, and big data (known as 3rd Platform) can be leveraged to provide economies and new business opportunities. With the Q1 2016 HYPERMAX OS release VMAX3 extends its support to the 1st platform Mainframe technologies allowing EMC to extend the core values first launched to Open Systems users: simplicity of ordering and service level performance management while bringing VMAX’s established data services such as SRDF and SnapVX to the mainframe space. EMC is at the heart of this discussion and is redefining data center cloud platforms to build the bridge between these two worlds to modernize and deliver the next generation Hybrid Cloud.

The best cloud based storage architectures require:

- Massive capacity scaling.
- Massive performance scaling.
- Flexibility to handle highly fluctuating workloads yet maintain consistent service levels, all the time.
- A data protection infrastructure that can scale with a cloud scale deployment.
- Reduced costs by converging infrastructure.
- A usage model that is automated and almost totally hands off.

The VMAX3 100K, 200K, and 400K arrays are designed to meet and exceed these requirements through:

- Scaling to 5760 drives, mixing Flash SAS and NL-SAS.
- Leveraging a scale out architecture up to 8 engines, 384 cores, 16TB cache and 256 physical host connections.
- Service Level Objectives (SLOs) and an architecture that dynamically adjusts system resources to where they are needed most.
- The most advanced local and remote replication solutions in the industry.
- A converged storage platform that runs powerful storage and application workloads on VMAX3.
- Mainframe Support, z/OS, z/TF, z/VM, Linux on System z all supported (3380 and 3390)
- An ease of use model that is unique in the high end storage arena.

This paper outlines in detail how the VMAX3 delivers these capabilities.

AUDIENCE

This white paper is intended for EMC customers and EMC personnel.
INTRODUCTION

EMC VMAX3 is incredibly well positioned to solve the CIO challenge of embracing a modernized flash-centric data center and hybrid cloud while simultaneously trying to simplify, automate and consolidate IT operations. VMAX3 isn't just bigger, better and faster – which it is – VMAX3 was designed as a data services platform that specifically addresses the new requirements of the modern data center while continuing to deliver the reliability and availability our customers have relied on for years.

With VMAX3, the industry’s leading tier 1 array has evolved into a thin hardware platform with complete set of rich software data services servicing internal and now external block storage. VMAX data services are delivered by a highly resilient, agile hardware platform that offers global cache, CPU (processing) flexibility, performance and HA at scale able to meet the most demanding storage infrastructures.

VMAX3 also radically simplifies management at scale though service level objectives (SLOs). SLOs change the customer conversation from "how many disks of which type to allocate" to "what performance does an application need. Automation within VMAX3 assigns the needed resources to meet the performance target, and continually adjusts to maintain it. Tier 1 storage management can now be done in a matter of minutes, and doesn't require extensively trained IT storage gurus. By delivering these capabilities, VMAX3 improves overall staff productivity allowing them to focus on the needs of the business, rather than the management of technology.

EMC VMAX3 arrays continue the legacy of all Symmetrix / DMX and VMAX arrays that have come before it. It allows IT departments to automate, modernize and converge their data center infrastructure while delivering mission-critical storage with the scale, performance, availability, security and agility they've come to rely on for years.

VMAX3 hardware models

The VMAX3 Family with HYPERMAX OS 5977 encompasses three new array models, VMAX 100K, VMAX 200K and VMAX 400K (referred to as VMAX3 arrays). All three models can be configured as all flash arrays or Hybrid arrays, and all share the same software and hardware features. The key differentiator between models is the number of CPU cores per engine and how big they can scale. The key scalability numbers are shown in Figure 1.

The VMAX3 arrays provide unprecedented performance and scale. Ranging from the single or dual-engine 100K up to the eight-engine VMAX 400K, these new arrays offer dramatic increases in floor and tile density with engines and high capacity disk enclosures for both 2.5” and 3.5” drives consolidated in the same system bay.
In addition, the VMAX 100K, 200K, and 400K arrays support the following.

- Hybrid or all-flash configurations.
- System bay dispersion of up to 82 feet (25 meters) from the first system bay.
- Optional third-party racking.

The VMAX3 arrays support the use of native 6Gb/s SAS 2.5” drives, 3.5” drives, or a mix of both drive types. Individual system bays can house either one or two engines and up to four or six high density disk array enclosures (DAEs).

In previous versions of the VMAX Family, the operating system was called Enginuity. Starting with VMAX3, the operating system is called HYPERMAX OS. All VMAX3 arrays are 100% Virtual Provisioned and pre-configured in the factory. The arrays are built for management simplicity, extreme performance, and massive scalability in a small footprint. With VMAX3, storage can be rapidly provisioned with a desired Service Level Objective to meet the needs of the most demanding workloads at scale.

**HYPERMAX OS**

The VMAX3 arrays introduce the industry’s first open storage and hypervisor converged operating system, HYPERMAX OS. HYPERMAX OS combines industry-leading high availability, I/O management, quality of service, data integrity validation, storage tiering, and data security with an open application platform. It features the first real-time, non-disruptive storage hypervisor that manages and protects embedded services by extending VMAX high availability to services that traditionally would have run external to the array. It also provides direct access to hardware resources to maximize performance. The hypervisor can be non-disruptively upgraded.

HYPERMAX OS runs on top of the Dynamic Virtual Matrix leveraging its scale-out flexibility of cores, cache, and host interfaces. The VMAX3 hypervisor reduces external hardware and networking requirements, delivers higher levels of availability, and dramatically lowers latency.

HYPERMAX OS provides the following services:

- Manages system resources to intelligently optimize performance across a wide range of I/O requirements and ensures system availability through advanced fault monitoring, detection, and correction capabilities.
- Provides concurrent maintenance and serviceability features.
- Interrupts and prioritizes tasks from microprocessors such as ensuring fencing off failed areas takes precedence over other operations.
- Offers the foundation for specific software features available through EMC’s disaster recovery, business continuity, and storage management software.
- Provides functional services for VMAX3 arrays and for a large suite of EMC storage application software.
- Defines the priority of each task, including basic system maintenance, I/O processing, and application processing.

**Data at Rest Encryption (D@RE)**

Data at Rest Encryption D@RE provides hardware-based, on-array, back-end encryption for VMAX3 family systems running HYPERMAX OS Q1 2015 SR or above. Back-end encryption protects your information from unauthorized access when drives are removed from the system. D@RE provides encryption on the back end using SAS I/O modules that incorporate XTS-AES 256-bit data-at-rest encryption. These modules encrypt and decrypt data as it is being written to or read from a drive. All configured drives are encrypted, including data drives, and spares. In addition, all array data is encrypted, including Symmetrix File System and Vault contents.

D@RE incorporates RSA® Embedded Key Manager for key management. D@RE keys are self-managed, and there is no need to replicate keys across volume snapshots or remote sites. RSA Embedded Key Manager provides a separate, unique DEK for all drives in the array including spare drives.
By securing data on enterprise storage, D@RE ensures that the potential exposure of sensitive data on discarded, misplaced, or stolen media is reduced or eliminated. As long as the key used to encrypt the data is secured, encrypted data cannot be read. In addition to protecting against threats related to physical removal of media, this also means that media can readily be repurposed by destroying the encryption key used for securing the data previously stored on that media.

D@RE is compatible with all VMAX3 system features, allows for encryption of any supported logical drive types or volume emulations and delivers powerful encryption without performance degradation or disruption to existing applications or infrastructure.

**MANAGEMENT SOFTWARE**

**Solutions Enabler**

Solutions Enabler provides your host with the Symmetrix Command Line Interface (SYMCLI). The SYMCLI is a comprehensive command set for managing your environment. SYMCLI commands can be invoked on the command line or within scripts. These commands can be used to monitor device configuration and status and perform control operations on devices and data objects within your EMC VMAX storage environment.

Solutions Enabler 8.0 or above is required for discovering VMAX3 arrays running HYPERMAX OS 5977.

**Unisphere® for VMAX**

Unisphere for VMAX is an advanced Graphical User Interface (GUI) that provides a common EMC user experience across storage platforms. Unisphere for VMAX enables customers to easily provision, manage, and monitor VMAX environments. Unisphere 8.1 has been enhanced to support the new capabilities of the VMAX3 family.

With the release of HYPERMAX OS Q3 2015 SR it is possible to run Unisphere for VMAX as an Guest Operating system directly on the VMAX3 controllers within the VMAX3 native Hypervisor. This option removes the need to an external management host to control and manage the VMAX3 array. Embedded management must be specified when ordering the VMAX3 system as CPU and memory requirements must be sized appropriately. Please see the [Unisphere 8.1 Documentation](https://support.emc.com) set available on https://support.emc.com for more information.

Unisphere for VMAX offers big-button navigation and streamlined operations to simplify and reduce the time required to manage a data center; it also simplifies storage management under a common framework.

Unisphere for VMAX 8.1 contains a number of task-oriented dashboards to make monitoring and configuring VMAX systems intuitive and easy.

The Storage Group Dashboard displays information about application storage groups and whether or not they are meeting their SLO requirements. Administrators can quickly navigate from this dashboard to gather more in-depth performance statistics.

The Protection Dashboard displays information about storage groups that are protected by local and remote replication. From here administrators can right-click on the storage group to protect their applications with Snapshots/SRDF® or configure the constructs for ProtectPoint™. The protection dashboard now includes a backup category that contains information about storage groups containing devices configured for ProtectPoint making it easy to monitor and gauge how applications are being protected and what recovery options exist.

**Unisphere Performance Event Correlation**

Event correlation within performance views makes it easier for users to identify root cause for any performance issues. Intuitive displays with contextual drilldowns direct the attention of the storage administrator to the likely cause of issues. Figure 3 shows the event correlation dashboard. In this instance storage group zp_157 is not meeting its SLO response time, the likely causes are highlighted to be either front end or disk. Clicking on the storage group will bring the user to a detailed view showing the utilization of each of the components from the storage group perspective enabling faster root cause analysis to be done and the storage admin to determine the corrective action needed.
**Figure 2. Storage Groups Dashboard - Event Correlation**

Unisphere Performance Host views

Host view of Performance enables administrators to explore performance metrics from the host (Initiator Group) or cluster (Host group). These views enable the storage admin to quickly identify if performance problems are internal or external to the array and determine the cause. Figure 4 shows a sample of a host view report built into Unisphere.

**Figure 3. Unisphere Performance All Hosts view**

The All Host view shows an overview of all hosts in the system. Administrators can drill down to get a more detailed view of each host as shown in Figure 5. From these views individual host initiator statistics and Key Performance indicators can be investigated at a host level helping to determine workloads on the host or cluster level. It is also possible to create custom charts with additional metrics.
Unisphere for VMAX 8.1 also includes new troubleshooting capabilities to provide an easy way for the admin to type any object (Host WWN, Storage Group, Host) and view any matching object across all managed arrays. This cuts the time spent searching multiple managed systems - one of the most time-consuming issues for a Storage administrator. Figure 6 shows the Unisphere search facility and how to access.

**Figure 4. Unisphere Individual host view**

**Figure 5. Accessing the Unisphere Search Facility**
**Unisphere 360™**

Unisphere 360 is a new management software option for any customer running more than one VMAX in a single data center. It provides the ability to aggregate and monitor up to 200 VMAX systems in a data center through the use of embedded management (eManagement) available in VMAX3. Unisphere 360 removes the need to run Solutions Enabler or Unisphere on an external host for every VMAX3, it provides a consolidated view into the entire VMAX (VMAX & VMAX3) infrastructure in a single data center.

Unisphere 360 for VMAX facilitates better insights across an entire data centre by providing storage administrators the ability to view site level health reports for every VMAX or coordinate compliance to code levels and other infrastructure maintenance requirements.

Figure 7 Shows the Unisphere 360 Dashboard.

---

**Figure 6. Unisphere 360 Dashboard**

![Unisphere 360 Dashboard](image)

---

**Unisphere Planning Section**

Unisphere for VMAX 8.0.3 and higher provide VMAX administrators with a planning view of their storage landscape with the planning dashboard. In order to utilize the planning utilities of Unisphere for VMAX the system needs at least 10 days of performance data. Unisphere for VMAX has the ability to project future utilization based on usage trends observed over time. Using this trending information administrators are able to project and plan for future growth. The planning section has 2 projection dashboards:

1. Array Projection Dashboard
2. Capacity Projection Dashboard

---

**Array Projection Dashboard**

The array protection dashboard is accessed by clicking on the Performance Icon and selecting Plan. This opens a heatmap view of the array component over the selected time range by default, when initially opened, the Utilization Projection dashboard displays heatmap data for the next month. The timeframe is customizable to the next week, next month, next 3 months or the next 6 months. The heatmap is updated to show the administrator at a glance if any component utilization is expected to increase over the time period selected.

The administrator can configure the time ranges for the dashboard by selecting an appropriate time range from the drop-down menu as shown in Figure 8.
If any component shows hot (Orange/Red) over the timeframe selected, the administrator can click the component element in the array to display a chart for the selected component showing the current utilization values and the projection for the selected time frame.

Capacity Projection Dashboard

The capacity projection dashboard planning view is accessed by changing the category from Array to Thin Pool as shown in Figure 9, from this view administrators are able to view how quickly their array capacity is being utilized and determine when to add additional capacity. The days to full metric gives an indication based on current growth rates how quickly the underlying thin pools on each SRP are being consumed and when they are likely to reach capacity.

Administrators can schedule both Array and Capacity planning reports to be emailed on a regular basis to ensure that they are aware of growth requirements.
VMAX REST API

The EMC Unisphere for VMAX REST (Representational State Transfer) API allows for accessing diagnostic performance data, accessing configuration data, and performing storage provisioning operations for VMAX hardware through robust APIs.

The idea is to provide a simple API that provides a higher level of abstraction for a simpler operational model. These APIs can be used in any of the programming environments that support standard REST clients such as web browsers and programming platforms that can issue HTTP requests. The Unisphere for VMAX REST API supports both XML and JavaScript Object Notation (JSON) MIME types.

The VMAX REST API enables service providers and customers to easily integrate and manage VMAX storage into existing orchestration providing the ability to automate day-to-day storage tasks can be automated to allow for easier management of resources at scale.
**OPEN SYSTEMS NEW FEATURES**

**Virtual Provisioning in VMAX3, Pre-configured arrays, Storage Resource Pools**

All VMAX3 arrays arrive pre-configured from the factory with Virtual Provisioning Pools ready for use. A VMAX3 array pools all the drives in the array into Storage Resource Pools (SRP) which provide physical storage for thin devices that are presented to hosts through masking views. Storage Resource Pools are managed by Fully Automated Storage Tiering (FAST®) and require no initial setup by the storage administrator, reducing the time to I/O and radically simplifying the management of VMAX3 storage. With SRP, capacity is monitored at the SRP level, and RAID levels and binding are no longer considerations for the storage administrator as all devices are ready for use upon creation and RAID is implemented under the covers of the SRP preconfigured at the EMC factory.

Figure 10 shows the SRP components and the relationship to the storage group (SG) used for masking thin devices to the host applications. Note there is a 1:1 relationship between disk groups and data pools. Each disk groups specifies a RAID protection, disk size, technology, and rotational speed, forming the basis for each of the preconfigured thin pools. Every VMAX3 array comes from the factory with the bin file (configuration) already created. This means best practices for deployment - TDAT sizes, RAID protection, and data pools - will already be in place.

Figure 9.  **Storage Resource Pool components**

With the new preconfigured SRP model VMAX3 provides all the benefits of Thin Provisioning with none of the complexity.

For more details on managing, monitoring and modifying Storage Resource Pools refer to the **Solutions Enabler Array Management CLI Guide** Part of the **Solutions Enabler Documentation set** available at [https://support.emc.com/products/2071_Solutions-Enabler](https://support.emc.com/products/2071_Solutions-Enabler)

**128 KB track size**

VMAX3 arrays support 128 KB track size. This is the allocation unit for storage from SRPs for thin provisioned devices. When thin devices are created in a VMAX3 array they consume no space from the SRP until they are written to. With previous generations, a VMAX host write to a track would result in 12 tracks being allocated. If only one of these 12 tracks had been written to following the allocation, the unwritten tracks in the allocated space could not be reclaimed to the system. With HYPERMAX OS 5977 this is no longer the case; a host write will only cause 1 track to be allocated.
Device Expansion

VMAX3 arrays no longer require meta devices to support large device sizes. The maximum device size in VMAX3 can now scale to 64TB and devices can be resized non-disruptively. Unisphere for VMAX and Solutions Enabler provide mechanisms to expand a device or expand a group of devices seamlessly to the host. This feature requires HYPERMAX OS Q3 2015 SR and Unisphere for VMAX 8.1 with Solutions Enabler 8.1.

Figure 10. Expanding a device with Unisphere 8.1

VMAX3 FAST

FAST, automates the identification of active or inactive application data for the purposes of reallocating that data across different performance/capacity pools within a VMAX3 array. FAST proactively monitors workloads at both the LUN and sub-LUN level to identify busy data that would benefit from being moved to higher-performing drives, while also identifying less-busy data that could be moved to cost-effective drives, without affecting existing performance.

The VMAX3 FAST engine:

- Enables you to set performance levels by Service Level Objective.
- Actively manages and delivers the specified performance levels.
- Provides high-availability capacity to the FAST process.
- Delivers defined storage services based on a mixed drive configuration.

Service Level Provisioning

VMAX3 FAST provides the ability to deliver variable performance levels through Service Level Objectives (SLO). Thin devices can be added to storage groups, and storage groups can be assigned a specific Service Level Objective to set performance expectations. The service level objective defines the response time target for the storage group. FAST continuously monitors and adapts to the workload in order to maintain (or meet) the response time target.

There are five available Service Level Objectives, varying in expected average response time targets. There is an additional Optimized Service Level Objective that has no explicit response time target associated with it. To ensure consistency across the platform, Service Level Objectives are fixed and may not be modified; however a storage groups SLO may be changed by the user to match changing performance goals of the application. Table 1 lists available service level objectives:
Table 1  Service Level Objectives and Expected Average Response Time

<table>
<thead>
<tr>
<th>Service Level Objective</th>
<th>Behavior</th>
<th>Expected Average Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diamond</td>
<td>FLASH performance</td>
<td>0.8 ms</td>
</tr>
<tr>
<td>Platinum</td>
<td>Emulates performance between FLASH and 15K RPM drive</td>
<td>3.0 ms</td>
</tr>
<tr>
<td>Gold</td>
<td>Emulates 15K RPM performance</td>
<td>5.0 ms</td>
</tr>
<tr>
<td>Silver</td>
<td>Emulates 10K RPM performance</td>
<td>8.0 ms</td>
</tr>
<tr>
<td>Bronze</td>
<td>Emulates 7.2K RPM performance</td>
<td>14.0 ms</td>
</tr>
<tr>
<td>Optimized (default)</td>
<td>Achieves optimal performance by placing most active data on higher performing storage and least active data on most cost-effective storage</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The actual response time of an application associated with each Service Level Objective will vary based on the actual workload seen on the application, and will depend on average I/O size, read/write ratio, and the use of local or remote application.

Table 2 lists the workload types that may be added to the chosen Service Level Objective, (with the exception of Optimized), to further refine response time expectations.

Table 2  Service level workloads

<table>
<thead>
<tr>
<th>Workload</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLTP</td>
<td>Small block I/O workload</td>
</tr>
<tr>
<td>OLTP with Replication</td>
<td>Small block I/O workload with local or remote replication</td>
</tr>
<tr>
<td>DSS</td>
<td>Large block I/O workload</td>
</tr>
<tr>
<td>DSS with Replication</td>
<td>Large block I/O workload with local or remote replication</td>
</tr>
</tbody>
</table>

It is possible to change SGs assignment from one Service Level Objective to another non-disruptively. Thus, if a Service Level Objective does not provide sufficient performance, the user can change it to a better Service Level Objective online. Or conversely, the user can switch to a lower performance SLO. Figure 12 shows the breakdown of the preconfigured SLO Diamond with OLTP workload including the read, write, random and sequential mix that is used in the workload planning for this workload.
Once an application is in compliance with its associated Service Level Objective, promotions to higher performing drives will stop. Subsequent movements for the application will look to maintain the response time of the application below the upper threshold of the compliance range.

With Solutions Enabler 8.0.2 and Unisphere for VMAX 8.0.2 or higher, the pre-defined Service Level Objectives can be renamed. For example, Platinum could be renamed to Mission Critical to more accurately describe its use within the business.

Unisphere 8.0.2 and higher also provides functionality to create reference workloads from existing storage groups. Reference workloads can be used to more accurately describe the workload profile based on existing workloads on the system and set performance expectations based on actual workloads running on the VMAX3 system.

When workloads are being provisioned alerting can be configured to ensure that storage admins are notified if a storage group fails to meet its Service Level Objective. Unisphere for VMAX 8.0.2 also has a Service Level Report detailing how each storage group is performing. The dashboard indicates how much of the time each storage group is meeting its SLO and the percentage of time it is not in compliance. This dashboard can be set to run as a report and scheduled to be distributed by email on a user defined schedule.

Note: When assessing the performance for a storage group, Workload Planner calculates its weighted response time for the past 4 hours and for the past 2 weeks, and then compares the two values to the maximum response time associated with its given SLO. If both calculated values fall within (under) the SLO defined response time band, the compliance state is STABLE. If one of them is in compliance and the other is out of compliance, then the compliance state is MARGINAL. If both are out of compliance, then the compliance state is CRITICAL.
Refer to Tech Note *EMC VMAX3 Service Level Provisioning with Fully Automated Storage Tiering (FAST)*

Unisphere for VMAX 8.1 now provides a view of response time buckets for each individual storage group enabling the administrators giving storage administrators a more in depth view into the response time for their application IO. A sample user dashboard showing the response time buckets is shown in Figure 14. The grey band shows the response time band for the SLO associated with the storage group. In the example shown the majority of the IO for this storage group over the observed time period falls within or under the response time for the SLO while a small number of IOs are outliers, overall the storage group is meeting its average response time as a result.

**Figure 14. Response Time Buckets**
**FAST External - FAST.x™**

FAST.x™ enables non-VMAX arrays to be abstracted behind a VMAX3 platform to provide additional cost effective storage tiers, and enable VMAX3 features on these arrays (SRDF, SLOs, SnapVX, ProtectPoint etc.). FAST.X supports volumes from XtremIO, CloudArray, non EMC-Arrays, and ViPR objects to be exposed within a Storage Resource Pool (SRP). The FAST technology used by Service Level Objective is augmented to understand the performance profile of XtremIO, CloudArray (and any non-EMC storage by loading and profiling the arrays over a discovery period).

Figure 15 shows the additional benefit of an XtremIO with FAST.X on VMAX3. When used with XtremIO, FAST.X provides a very high performance tier of storage that is also gives very effective data reduction. When compared to a Diamond SLO, if the data reduction is significant, it could represent a more cost effective option. In addition, FAST.X also provides VMAX3 enterprise level protection services (SRDF and SnapVX for example) as well as to improve write performance due to VMAX3 large write caches).

**Figure 14. FAST.X™ With XtremIO**

Figure 16 shows the basic architecture for FAST.X with XtremIO. External arrays are connected via redundant pairs of Fiber Channel through Storage Area Network (SAN) via ports mapped to external director (DX) emulation in the VMAX3. Each external director DX has 2 ports configured. Each port is connected and zoned to a port on each storage controller of the XtremIO array. Storage volumes are presented from the EMC XtremIO Array to VMAX3 DX ports and consumed as additional capacity either as a Service Level in an existing SRP or as a separate SRP.
Storage administrators can increase available capacity from external storage using Unisphere for VMAX or Solutions Enabler via the add device wizard or the symconfigure command set. It is also possible to remove capacity from the external array using the drain options. Figure 17 shows how to access the management options in Unisphere for the external storage. Note: Capacity must be made visible from XtremIO through the host provisioning wizards of element manager to the VMAX3 array to be ingested for use by FAST.X.

If XtremIO devices are added into an existing SRP, FAST will include XtremIO devices in its algorithms scoring the XtremIO devices for data placement as an additional tier of Flash storage. The storage admin simply provisions storage from the VMAX3 in the same fashion as any other device and will either change the source SRP on the provisioning wizard or select a Service Level Objective that includes XtremIO.

FAST.X with cloud array enables VMAX to provide a very cost effective integrated cold data offload capability. Cloud Array should be used as a separate SRP or in a well understood workload as a tier in a primary SRP. FAST.X with CloudArray and non-EMC storage is supported on VMAX3 family systems running HYPERMAX OS Q3 2015 SR or above. FAST.X with XtremIO is supported with HYPERMAX OS 5977.596.583 with a required HYPERMAX OS epack.
FAST.X™ Incorporation

With Solutions Enabler 8.2 external provisioning through FAST.X provides the ability to use external storage as tiered storage within the VMAX3 array. Encapsulation allows the importation of LUNS from external arrays by preserving their existing data set and can be detached from FAST.X with integral data set presented to other hosts. Encapsulation mode for FAST.X combines both of these features by INCORPORATING existing data on external LUNS and immediately handing control to the FAST engine after virtualisation. This means the data is retained and FAST engine can move the data according to its algorithm and all data services (SnapVX /SRDF etc.) are supported without restriction. While the data is under FAST control and is free to be moved between tiers depending on the SLO assigned the incorporated storage from the external array is also available to the FAST engine for provisioning. The FAST engine will determine the response time of the new storage and provision data to it should it meet the needs of a particular storage groups SLO.

Database Storage Analyser (DSA) and FAST Hinting

Unisphere for VMAX V8.0.1 introduced Database Storage Analyzer (DSA). DSA is an application that is part of the Unisphere base install. It provides a database to storage performance troubleshooting solution for databases such as Oracle and Microsoft SQL, running on EMC Symmetrix and VMAX storage systems.

DSA is a Unisphere for VMAX feature and requires no additional license to run. It supports database to storage correlation by providing a shared view of how performance issues correlate to database and storage level activities. This view is accessible by database administrators (DBAs) and storage administrators (SAs). The view presents I/O metrics such as IOPS, throughput and response time from both the database and the storage system, which helps to quickly identify gaps between the database I/O performance and the storage I/O performance DSA supports Symmetrix arrays running Enginuity 5671-5876 and VMAX3 arrays running HYPERMAX OS 5977 or higher. DSA is fully compatible with VMware environments so even if the database is running on virtual environment, the DB/Storage correlation works.

Figure 18 shows the DSA Dashboard screen where DBAs and SAs can see at a glance their monitored databases response times and IOPs. From the initial dashboard, administrators can dive deeper into any that have a problem to identify if it is storage related or application related. If a read database response time is higher than the expected response time (configurable) then the response time status bar would show either yellow or red. The bubble chart represents the ratio between the DB response time (X Axis) and the storage response time (Y Axis); ideally everything should be on the left lower corner of the chart. We can see in the chart shown in Figure 18 the database named Ora11 has a very high DB response time but a low storage response time so it is seen as an outlier on the right of the chart. This can be investigated further using the analytics views of DSA shown in Figure 19.
Figure 17. DSA Dashboard

Figure 18. DB/Storage correlation
With FAST SLO management SAs can define performance targets for their applications and allow the array to manage resources automatically at a storage group level. There are instances where additional information might be available outside of the storage array metrics that could be provided to fine tune the behaviour of applications storage in advance of anticipated events or expected activities. Unisphere 8.2 and DSA provide new hinting capabilities where SAs/DBAs can provide hints to FAST in relation to specific Oracle/SQL databases and their objects and expected increases in activities due to known upcoming events. Hints are passed from DSA directly to the FAST engine. Hints should only include a portion of the database rather than the entire database.

There are 3 types of hints:

- **Top**
  - Main use case: periodical process such as end-of-month process where the assumption is that most of the data is inactive during the month.
  - This hint would cause the array to promote all data for the hinted database object to FLASH.

- **Very high**
  - Main use case: Ensure that a given process gets better response time than other processes in the SG even if the data was inactive.
  - In this case there is no immediate promotion to FLASH; however, it marks all extents including the inactive ones as active with a target SLO as Platinum OLTP.

- **High**
  - Main use case: Ensures a given database object gets better response time than other object extents in the SG. It is less powerful as it only promotes active extents.

For example, the accounts payable table will be busy on the accounting database for the 28th-30th September 9am-5pm 2015. During this time the workload on the storage for this table should be given the highest priority Diamond SLO promoting all the relevant data to FLASH prior to the start of the activity. Once the hint window has passed the regular SLO for the storage group will once again be applied.

**Figure 19. DSA Hinting**

It is important to note that FAST hinting will have no effect on workloads that are associated with Bronze or Diamond SLO. Diamond workloads already utilize the fastest storage resources and will not benefit from any additional movement. Workloads assigned Bronze SLO typically require lower response times importance and therefore hints are not supported for data on those workloads.
Note that the DSA DB/Storage correlation works for older generation Symmetrix and VMAX3 however the DB hinting feature only supports VMAX3 with HYPERMAX OS code 5977 Q3 2015 SR and above. Hinting is only supported on Oracle Databases, support for SQL Server is planned for a future version.

**Viewing Databases in Unisphere Performance Dashboards**

Once DSA starts to collect database statistics, it creates two database entities in Unisphere that can be viewed under the performance of Unisphere as a new category. One represents the database datafiles devices and the other represents the redo log devices. This gives the storage administrator storage dashboards from the context of a given database. For example, the Utilization dashboard would only display the hardware components that are associated with the database IO path. Figure 8 shows high front end utilization as well as high queue depths that impact the performance of database ora11.

![Figure 20. Database - Utilization dashboard](image)

**TimeFinder SnapVX**

EMC TimeFinder® software delivers point-in-time copies of volumes that can be used for backups, decision support, data warehouse refreshes, or any other process that requires parallel access to production data.

TimeFinder in HYPERMAX OS 5977 for VMAX3 introduces TimeFinder SnapVX, which combines the best aspects of the previous TimeFinder offerings, adding new ease-of-use features and increased scalability.

In arrays running HYPERMAX OS, TimeFinder SnapVX lets you non-disruptively create point-in-time copies (snapshots) of critical data. SnapVX creates snapshots by storing changed tracks (deltas) directly in the SRP of the source device. With SnapVX, you do not need to specify a target device and source/target pairs when you create a snapshot. If there is ever a need for the application to use the point-in-time data, you can create links from the snapshot to one or more target devices. If there are multiple snapshots and the application needs to find a particular point-in-time copy for host access, you can link and relink until the correct snapshot is located.

SnapVX supports TimeFinder/Clone, TimeFinder VP Snap, and TimeFinder/Mirror via emulations that transparently convert commands to SnapVX commands. You can still run scripts that use TimeFinder/Clone, TimeFinder VP Snap, and TimeFinder/Mirror commands, but the underlying mechanism will be SnapVX.
In HYPERMAX OS arrays, SnapVX supports up to 256 snapshots per source device (including any emulation mode snapshots). The legacy session limits (16 copies) still apply to the emulations.

**Figure 21. Example of SnapVX Snapshot**

You can set snapshots to automatically terminate after a specified number of days or at a specified date and time, Figure 22 shows multiple snapshots of a production volume with a Time to Live (TTL) of 1 Day. HYPERMAX OS will only terminate the snapshot if it does not have any links. If it does have links, HYPERMAX OS will terminate the snapshot when the last link has been unlinked. Writes to a linked target device will only be applied to the linked target and will not change the point in time of the snapshot itself. Snaps can be deleted out of order without affecting the chain of snaps.

For more information refer to [Tech Note VMAX3 Local Replication Suite TimeFinder SnapVX and TimeFinder Emulation](#)

**Enhanced SRDF**

The Symmetrix Remote Data Facility (SRDF) family of software is the gold standard for remote replication in mission critical environments. Built for the industry-leading high-end VMAX hardware architecture, the SRDF family of solutions is trusted for disaster recovery and business continuity. SRDF enables remote data services to provide 6 nines data availability 24x7xForever operations with synchronous SRDF. The SRDF family offers unmatched deployment flexibility and massive scalability to deliver a wide range of distance replication capabilities.

With VMAX 3 and HYPERMAX OS 5977, SRDF has been enhanced to provide increased performance of devices involved in remote replication, removing the need to configure meta-devices or increase device counts. This allows you to achieve desired performance goals with less complex implementation even in the most demanding environments. Since meta devices are no longer required with VMAX3 and are supported with previous generation SRDF, pairs are supported between existing VMAX meta devices (Enginuity 5876) and VMAX3 non-meta devices. SRDF on VMAX3 by default enables Geometry Compatibility Mode (GCM) to handle odd size source devices from a previous generation system; this addresses the conversion without the need for user intervention.

With VMAX3, all directors are capable of supporting a variable number of ports for SRDF up to 16 ports for every director configured with emulation for SRDF. The number of SRDF groups supported on an individual VMAX3 SRDF director (RA) has increased from 64 to 250, which is also the total number of SRDF groups per VMAX3 array. All VMAX3 array models are capable of supporting enhanced hardware compression for bandwidth optimization on both IP and fibre links.

Another key change in the architecture is the Delta Set Extension feature which is designed to increase availability in SRDF/A environments. There is no longer a need to configure a separate pool in the array and there is no need for a DSE pool to spill over in the remote (R2) array. Instead the SRP will have a maximum DSE capacity associated with it (specified in GBs). DSE capacity is set when the VMAX3 array is configured, resulting in a less complex configuration for the storage administrator to manage.
SRDF/A has also been improved to provide shorter and more predictable Recovery Point Objectives (RPO). This is done through a new multi-cycle mode which allows more than two delta sets to exist on the source array. SRDF/A Multi-cycle Mode (MCM) has the benefit of always cycle switching based on the user set cycle time, so the cycles are now predictable and much smaller when applied to the secondary side. In most cases, this eliminates the need for DSE on the secondary side.

The internal workings of these features and changes are discussed in more detail refer to the Solutions Enabler SRDF CLI Guide Part of the Solutions Enabler Documentation set available at https://support.emc.com/products/2071_Solutions-Enabler.

**SRDF/Metro**

SRDF/Metro significantly changes the traditional behavior with respect to the R2 device availability to better support host applications in high-availability environments. With SRDF/Metro, the SRDF R2 device is also read/write accessible to the host and will take on the federated personality of the primary R1 device (geometry, device World-Wide Name (WWN), and so on). By providing this federated personality on the R2 device, both R1 and R2 devices can appear as a single virtual device across the two SRDF paired arrays for host presentation. With both the R1 and R2 devices being accessible, the host or hosts (in the case of a cluster) can read and write to both R1 and R2 devices. SRDF/Metro ensures that each copy remains current and consistent, and addresses any write conflicts that might occur between the paired SRDF devices. A witness array can be Additionally SRDF/Metro can avail of external storage via FAST.X including XtremIO. Figure 23 shows the basic architecture of SRDF/Metro. For details of supported configurations refer to the Elab interoperability Navigator at https://elabnavigator.emc.com/eln/elnhome.

**Figure 22. SRDF/Metro Witness Configuration**

SRDF/Metro is fully compatible with FAST/SLO provisioning. When configuring SRDF/Metro, each site’s storage groups should have the same Service Level Objective set to ensure predictable performance. The internal FAST engine on both arrays incorporates FAST statistics from both sides to make sure each side represents the workload as a whole. This ensures consistent placement of data and performance from either side. Site Bias can be configured to control which site will be primary in the event of link separation or site outage.

SRDF Metro is supported on VMAX3 family systems running HYPERMAX OS Q3 2015 SR or above. For More information on SRDF/Metro refer to VMAX 3 SRDF Family CLI User Guide available as part of the Solutions Enabler 8.1 Documentation Set Available on https://support.emc.com/products/2071_Solutions-Enabler/ from Sept 2015.
Integrated Data Protection with ProtectPoint

EMC ProtectPoint provides faster, more efficient backup while eliminating the backup impact on application servers by integrating primary storage on HYPERMAX OS arrays with protection storage for backups on EMC Data Domain systems. ProtectPoint reduces cost and complexity by eliminating traditional backup applications while still providing the benefits of native backups. The VMAX3 primary storage keeps track of and sends changed blocks for application directly to Data Domain appliance, which uses these changed blocks to create full backups in native format. Figure 24 shows the basic architecture of EMC ProtectPoint and the VMAX3 software features that enable its operation.

Figure 23. EMC ProtectPoint Architecture

EMC ProtectPoint encapsulates Data Domain devices copying data to/from these devices for backup and restore operations. With this configuration, backups are archived in permanent storage within the Data Domain appliance using one of the following methods:

- Initiate directly from VMAX3 storage using the ProtectPoint agent
- Initiate automatically via scripts to send backups direct to DataDomain

Backups from Protect Point eliminate backup impact on the application by removing the server from the data path of the backup and minimizing the backup window needed.

ProtectPoint gives application owners and database administrators complete control of their own backup, recovery, and replication directly from their native application utilities. This empowers application owners by providing them the control they desire without additional cost and complexity. It eliminates backup impact on the local area network (LAN) and minimizes storage area network (SAN) bandwidth requirements by sending only unique data from primary storage to protection storage.

Since ProtectPoint backs up data to a Data Domain system, it’s protected with the Data Domain Data Invulnerability Architecture that provides the industry’s best defence against data integrity issues. Inline write-and-read verification, continuous fault detection, and self-healing ensure that backup data is accurately stored, retained, and recoverable throughout its lifecycle on a Data Domain system.

For detailed information on ProtectPoint and supported database environments refer to the ProtectPoint Solutions Guide available on https://support.emc.com
Embedded NAS (eNAS)

VMAX3 incorporates a light-weight hypervisor with its HYPERMAX OS operating system, which enables embedded data services directly on the storage array and delivers new levels of efficiency to enterprise workloads. Embedded NAS (eNAS) is the first data service to be offered in this way. Virtual instances called Software Data Movers and Control Stations run within the HYPERMAX OS architecture on VMAX3 in the FA emulation of the VMAX3 directors.

The eNAS implementation is fully integrated into the VMAX3 platform and is configurable through Unisphere for VMAX. eNAS provides flexible and secure multi-protocol file sharing (NFS, CIFS/SMB 3.0) as well as multiple file server identities (CIFS and NFS servers), enabling file server consolidation/multi-tenancy, built-in async file level remote replication (File Replicator), built-in Network Data Management Protocol (NDMP), and Anti-Virus.

Unisphere for VMAX has been enhanced to include a file management dashboard and interface to create and automatically provision storage to the ENAS Data Movers; this facilitates File I/O for end customers. Detailed NAS active management (Share management, replication, Quotas, etc.), is performed within standard EMC Unisphere for VNX File. The link between Unisphere for VMAX and Unisphere for VNX File is based on context sensitive Link and launch. Figure 25 shows the Unisphere for File dashboard which provides the interface for interacting with the embedded NAS components and adding storage. Figure 26 Highlights the Scalability for eNAS across the VMX3 family for HyperMAX OS Q1 2016 Code release.

Figure 24. Unisphere for File Dashboard

![Unisphere for File Dashboard](image)

Figure 25. eNAS Scalability

<table>
<thead>
<tr>
<th>Category</th>
<th>EMC VMAX 100K</th>
<th>EMC VMAX 200K</th>
<th>EMC VMAX 400K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum eNAS CPU Cores</td>
<td>10</td>
<td>22</td>
<td>66</td>
</tr>
<tr>
<td>Software Data Movers</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>eNAS Capacity (usable)</td>
<td>256TB</td>
<td>1.5 PB</td>
<td>3.5 PB</td>
</tr>
<tr>
<td>Maximum Ethernet Ports</td>
<td>24</td>
<td>48</td>
<td>96</td>
</tr>
</tbody>
</table>

For detailed information on eNAS, refer to the EMC VMAX3 Unified Embedded NAS Technical Notes available on [http://www.emc.com](http://www.emc.com).
Virtual Data Mover synchronous replication

Virtual Data Mover (VDM) synchronous replication (also called file auto recovery) provides the ability to use the embedded NAS (eNAS) for File CLI to manually fail over or move a VDM from a source eNAS system to a destination eNAS system. The failover or move uses block-level SRDF synchronous replication (SRDF/S), providing zero-data-loss protection for an unplanned failover.

The synchronous replication feature consolidates VDMs, file systems, file-system checkpoint schedules, CIFS servers, NFS file shares, networking, and VDM configurations into their own separate pools. This feature works for a true DR where the source is unavailable. For DR support in the event of an unplanned failover, it provides an option to recover and clean up the source system to prepare it as a future destination.

You can also manually initiate failover and reverse operations using the EMC File Auto Recover Manager (FARM). The FARM option provides the ability to automatically fail over a selected sync-replicated VDM on a source eNAS system to a destination eNAS system. Use FARM to set up monitoring of configured synch-replication VDMs and to trigger automatic failover due to VDM, file system, control station, or IP network unavailability that would cause the NAS client to lose access to data.

The front-end port map and recommended practices

Each VMAX3 engine has two directors supporting up to 16 front Fibre Channel (FC) ports on each director. Figure 27 shows the possible port numbering for physical ports in each director if it is fully populated with 4-port FC IO Modules. Front-end ports are available in Slots 2 and 3, or Slots 8 and 9. With VMAX3 arrays running HYPERMAX OS 5977 there is only a single emulation instance of a specific type (FA (Front End), DA (Back end), DX (ProtectPoint), RF(SRDF), EF (FICON), etc.) available per director board. VMAX3 front-end ports can be assigned to the emulation by the storage administrator.

Figure 26. VMAX3 FC Front End slots

It is recommended that when connecting a host or cluster to VMAX3:

- 2-4 Front-end paths are configured in the port group for masking and zoned to the host (single initiator zoning is recommended).
- For Cabling options one approach is to connect all even numbered ports to fabric A and odd numbered ports to fabric B.
  - In single engine systems with this approach select 2 I/O ports spanning both SAN fabrics on each director, for example, Port 28 & 29 on both directors 1 and 2.
  - In a multi-engine system if you choose an odd and even port on at least 2 different directors spanning different engines this will spread load for performance and ensure fabric redundancy; for example, Port 28 & 29 on directors 1 and 3.

For more information on port to director emulation support refer to Solutions Enabler Array Management CLI guide available on https://support.emc.com

Open Systems Scalability

With the new VMAX3 architecture some scalability limits have changed from previous generations of VMAX. The following limits affect VMAX 100K, 200K and 400K systems:

- Devices can be up to 64 TB in size.
The maximum number of host addressable devices is 64K per VMAX3 array.
The maximum number of storage groups, port groups, and masking views is 64K per VMAX3 array.
The maximum number of devices addressable through each port has been increased to 4K. There are no meta devices in the VMAX3 array so this number is harder to reach than with previous systems.

Enhancement to group management features

Cascaded storage groups provide storage administrators with a great deal of flexibility: administrators can nest multiple storage groups and associated Service Level Objectives within a single parent storage group used for both FAST and for the masking view. Unisphere for VMAX has been updated to make this implementation easier. The user interface has been simplified so that administrators can provision storage for multiple service level objectives from a single screen, as shown in Figure 28. In this example, a parent storage group named Finance is created with three child groups for Payroll, Accounting, and Reporting. Each child group has its own Service Level Objective and drive capacity. Expanding the storage group is as easy as creating it, the wizard brings you to the same screen and you can increase the volume count for the group you require.

Figure 27. Provisioning cascaded storage groups in Unisphere

Until now, if a system was implemented without cascaded storage groups this could not be easily changed. Often it was necessary to maintain two sets of storage groups: one for the host access in the masking view and a second set for FAST Policy association. Unisphere and Solutions Enabler now also provide capability to convert a standalone storage group to a cascaded model, enabling more flexible configurations. These changes also mean that a single parent storage group with a number of children can be implemented after the fact, making the transition to multiple services offering simpler.

With previous releases, storage implementations that had been designed without FAST in mind were complex to manage and required careful planning for changes. Solutions Enabler 8.0 now allows for the following:

- Moving devices between child storage groups of a parent storage group when the masking view uses the parent group.
- Moving devices between storage groups when a view is on each storage group and both the initiator group (IG) and the port group (PG) elements are common to the views (initiators and ports from the source group must be present in the target).
• Moving devices from a storage group with no masking view or service level to one in a masking view. This is useful as you can now create devices and automatically add to storage group from CLI, so a staging group may exist. Figure 29 shows a device being moved between storage groups to improve response times. Figure 29 also shows the common components and the possible directions devices may move between storage groups.

**Figure 28. Non-disruptive movement of devices between storage groups.**

### Host I/O limits per storage group

This feature allows every storage group (up to the maximum number of storage groups per array) to be associated with a host I/O limit. Quota limits are evenly distributed across the available directors within the associated port group.

Setting a Host I/O limit on a storage group ensures that applications cannot exceed their set limit, therefore reducing the potential of impacting other applications. Host I/O limits also provide greater levels of control on performance allocation in multi-tenant or Cloud environments, providing the predictability required to service more customers on the same array.

This feature manages expectations of application administrators with regard to performance, and provides incentives for users to upgrade their performance service levels.

### VMware VAAI support

EMC’s implementation of the vSphere Storage APIs for Array Integration, or VAAI, is done natively within the HYPERMAX OS software and therefore does not require additional software to enable support for block devices. However if eNAS is implemented, VAAI plugin for NAS is required. It can be downloaded from EMC support and installed on the ESXi hosts to enable VAAI on NAS. VMAX3 arrays with HYPERMAX OS 5977 supports all 4 VAAI primitives: Atomic Test and Set, T10 Unmap (with ESXi 5.0 U1+), Block Zero, and Full Copy. ESXi version 5.0 or higher is required for connectivity to VMAX3 arrays.
For more information and best practices, refer to the white paper Implementing VMware VSPHERE Storage API for Storage Awareness with Symmetrix Storage Arrays, available on EMC.com. Also refer to EMC ELAB navigator for the latest supportability information at https://elabnavigator.emc.com/.

**VMware VASA support**

vSphere Storage APIs for Storage Awareness (VASA) are vCenter server-based APIs that enable storage arrays to inform the vCenter server about their configurations, capabilities, and storage health and events. They make an administrator aware of the physical storage topology, capability and state. EMC supports vSphere Storage APIs for Storage Awareness through the VASA provider, which is downloadable from https://support.emc.com.

VMAX3 VASA integration requires the 2.x VASA provider.

**MAINFRAME SUPPORT**

With the HYPERMAX OS Q1 2016 release, the functionality of VMAX3 to becomes available to mainframe customers. Along with leveraging VMAX3’s massive array of data services (RDF, SnapVX etc.) it also offers automated tiering by Service levels in three varieties: Diamond, Bronze and Optimized.

**Mainframe Enabler**

Mainframe Enabler (MFE) is a suite of z/OS-based products for managing the VMAX3 in a z/OS environment. MFE commands can be used to monitor device configurations and statuses and used to perform control operations on devices and data objects within the storage environment. MFE 8.0 is released as part of the HYPERMAX OS Q1 2016 release but can also be used to manage older VMAX and Symmetrix models.

**GDDR**

GDDR version 5.0 is supported as part of the HYPERMAX OS Q1 2016 release with VMAX3. GDDR is a business continuity automation solution for both unplanned and data centre site switch operations and restart operations following site disaster, ranging from the loss of compute capacity and/or disk array access through total loss of a single data centre or regional disaster including loss of dual data centres.

**IBM z Systems Compatibility Enhancements**

With the inclusion of mainframe support for VMAX3 the following IBM 2107 support and copy features are included:

- Query Host Access (QHA) CCW support
- PPRC Soft Fence
- Non-Disruptive State Save (NDSS)
- zHyperwrite
HARDWARE ENHANCEMENTS

VMAX3 engines

The VMAX3 hardware infrastructure for the 100K, 200K, and 400K arrays is designed around new engines. A VMAX3 engine is a pair of physically separate director boards housed in the same enclosure. Each director board is physically independent each having its own power and redundant hardware components. The director board supports 24 (100K), 32 (200K), or 48 (400K) 1vy Bridge cores, and Intel hyper-threading technology. The only physical differences between the engines in 100K, 200K, and 400K arrays are the dual inline memory module (DIMM) populations and CPUs (in both core frequency and number of cores). Figure 30 below shows the common hardware components of the VMAX3 engines and the Management Module Control Station (MMCS), which is only present in the first system bay of each VMAX3 array. The VMAX3 engine is designed to be modular and easily serviced. Directors can be removed at the front for service upgrades without the need to disconnect any cabling from front-end or back-end I/O modules.

Figure 29. Fully configured VMAX3 Engine rear view

Multi-Core emulation

Multi-Core emulation provides additional CPU and physical port utilization capabilities to HYPERMAX OS emulations, extending the existing code architecture and improving overall performance. One key feature of Multi-Core emulation is pre-defined core mappings that allow specification of performance characteristics based on expected I/O profiles and usage. VMAX3 arrays can be configured to be front-end centric (allocating more CPU cores to handle host I/O), back-end centric (allocating more CPU cores to handle disk I/O), or the default baseline configuration where CPU cores are evenly distributed between front and back end.
Figure 30. VMAX3 Multi-Core Emulation

Figure 31 shows the default Multi-Core emulation in VMAX3 arrays. Cores are pooled for front end, back end, and for HYPERMAX OS functions. All of the CPU cores on the director will work on I/O from all of the ports. This helps ensure VMAX3 directors’ ports are always balanced.

Dual MMCS

All VMAX3 arrays contain two Management Module Control Systems (MMCS) in system bay 1. This helps to increase system availability as there are multiple access points to the system for remote access. If there is a failure in either MMCS, the system is able to dial home from the remaining MMCS for remote recovery or diagnose if hardware replacement is required. The MMCS replaces the Service Processor that was present in earlier VMAX models.

Dynamic Virtual Matrix/Infiniband Fabric

The Dynamic Virtual Matrix provides the Global Memory interface between directors with more than one engine. The Dynamic Virtual Matrix is composed of multiple elements, including Infiniband Host Channel Adapter (HCA) endpoints, Infiniband Interconnects (switches), and high-speed passive, active copper, and optical serial cables to provide a Virtual Matrix interconnect.

A fabric Application Specific Integrated Circuit (ASIC) switch resides within a special Management Interface Board Enclosure (MIBE), which is responsible for Virtual Matrix initialization and management. Each fabric port connects back to an Infiniband switch housed in the first system bay cabinet. The Infiniband switches are only present in multi-engine systems and are added with the addition of a second engine. Infiniband switches are installed in pairs and each director has a path to Fabric switch A and B. Fabric switches are supported by standby power supplies for vault activities to ensure all cache data gets vaulted.
The VMAX 100K and 200K arrays support 8 Interconnect ports; the VMAX 400K array supports 16 Interconnect ports. These are shown below in Figure 32 and Figure 33.

**Figure 31. 12 Port Infiniband fabric switch VMAX100K/200K**

![12 Port Infiniband fabric switch VMAX100K/200K](image)

**Figure 32. 18 port Infiniband fabric switch VMAX400K**

![18 port Infiniband fabric switch VMAX400K](image)

The cabling on the Infiniband switches is simpler than previous VMAX models, enabling faster setup time by EMC field personnel.

**New director emulations**

Two new director emulations have been introduced with HYPERMAX OS 5977 and VMAX3: Infrastructure Manager (IM) and HYPERMAX OS Data Services emulation\(^1\). The IM emulation is an aggregation of common infrastructure tasks previously distributed across all director types. This consolidation is intended to allow other directors to devote their CPU resources to I/O specific work only, without interference from the demands of the infrastructure tasks. The HYPERMAX OS Data Services emulation also provides a consolidation of various functionalities, with the main goals being to both reduce I/O path latency and introduce better scalability for various HYPERMAX OS applications.

**Table 3  Emulation Types in VMAX3**

<table>
<thead>
<tr>
<th>Emulation</th>
<th>Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>Back</td>
<td>Disk Services (Da for SAS Disks)</td>
</tr>
<tr>
<td>DX</td>
<td>Back</td>
<td>Director External (Used for FAST.X and ProtectPoint)</td>
</tr>
<tr>
<td>IM</td>
<td>Middle</td>
<td>Infrastructure Management</td>
</tr>
<tr>
<td>EDS</td>
<td>Middle</td>
<td>Enginuity Data Services</td>
</tr>
<tr>
<td>FA</td>
<td>Front</td>
<td>Front End (Fiber Channel)</td>
</tr>
<tr>
<td>FE</td>
<td>Front</td>
<td>Front End (FCoE Fiber Channel over Ethernet 10GbE)</td>
</tr>
<tr>
<td>SE</td>
<td>Front</td>
<td>Front End (iSCSI Ethernet 10GbE)</td>
</tr>
<tr>
<td>RA</td>
<td>Front</td>
<td>Remote Adapter (Fiber Channel)</td>
</tr>
<tr>
<td>RE</td>
<td>Front</td>
<td>Remote Adapter (Ethernet 10GbE/Gige)</td>
</tr>
</tbody>
</table>

**Vault to FLASH**

Vaulting is the process of saving Global Memory data to a reserved space during an offline event. Vault to FLASH provides vaulting of Global Memory data to an internal flash I/O module. The feature provides the following advantages:

- Better performance by enabling larger Global Memory per director that is capable of being saved within 5 minutes.
- The system is lighter and requires fewer batteries.
- The system is easier to configure as there is no longer a requirement to reserve capacity on back-end drives for vault space.

---

\(^1\) May also be referred to as EDS or ED.
No minimum drive count/engine.

6 Gb/s SAS back-end/drive infrastructure

All VMAX3 models utilize 6 Gb/s SAS (Serial Attached SCSI) drives with back-end configuration that provides improved performance. SAS is a high-speed reliable protocol that uses the same low-level technology as Fiber Channel encoding. SAS topology is different from Fiber Channel as SAS uses a connectionless tree structure with unique paths to individual devices. Routing tables store these paths and help to route I/O to the required locations. Figure 30 shows the VMAX3 Engine with 2x2 Port 6Gb/s SAS I/O modules per director each providing 8 lanes of SAS for connectivity to back-end Disk Array Enclosures (DAE).

16 Gb/s FC front-end support

All VMAX3 models support 16 Gb/s FC front-end support which is the current fastest available Fibre channel speed at time of publishing and keeps VMAX a leader in the FC market. 4 port 16 Gb/s FC I/O modules are available and can auto negotiate to support 4/8/16 Gb/s line speeds. It is possible to configure up to 32 16 Gb/s FC front end ports in a VMAX3 engine.

16 Gb/s FICON and zHPF Support

With the release of HYPERMAX OS Q1 2016 release support for 16Gb/s FICON channel adapter cards. Each card consists of a 4 Port 16 Gb/s I/O module based on the Qlogic ship set. Each port auto-negotiates to 4,8 and 16 Gb speeds. 
The VMAX3 will also support zHPF which includes list prefetch and Di-directional support, Format Write Commands and exploitation of xHPF by BSAM, QSAM and BPAM access methods.

In addition VMAX3 will support the following FICON enhancements announced by IBM z13:

- Forward Error Correction (FEC) support of 16Gb FICON
- FICON Dynamic Routing (FIDR)
- Read Diagnostic parameters to enable SAN management products to display diagnostic for 16Gb/s links.

Ultra dense Disk Array Enclosure support and mixing

VMAX3 systems provide two types of Disk Array Enclosure (DAE): ultra-high density (120 2.5” drives) DAEs and standard density (60 3.5” drives). Figure 34 shows the physical hardware for each of the DAE units.

Every DAE is designed to have 4 power zones and can thus tolerate the loss of power to any one zone (this would require loss of two separate power supplies). DAEs can be added to systems in single increments if using RAID 1, RAID5 (3+1), and R6 (6+2). However, if your system contains RAID5 (7+1) or RAID6 (14+2), adding DAEs may only be possible in pairs. If required, it is possible to configure RAID6 (14+2) and RAID5 (7+1) across 4 DAE so that only one member resides in any power zone.

For maximum flexibility, DAEs can be mixed behind engines to accommodate 2.5” and 3.5” form factor drives in the same array. A VMAX3 engine is able to support up to 6 DAEs (720 x 2.5” drives, 360 x 3.5”) drives, or a mixed DAE (combination of the two).

When the system is configured at the factory, drives are distributed across engines in balanced configurations to provide the optimal performance in the array. When drives are added it is expected that they will also be added in a balanced manner.
Local RAID

VMAX 100K, 200K and 400K arrays implement local RAID which requires all members of a RAID group to be located behind the same engine. This provides local access and control over I/O for all RAID members and reduces the number of messages and Global Memory operations that need to be carried out for RAID operations, lowering I/O overhead and improving RAID performance. Local RAID also eliminates the need for cross-bay cabling in direct/daisy chain DAEs. This allows for the dispersion of systems at engine/bay level (position around any obstacles or across an aisle), making the VMAX3 systems the most flexible storage system in the industry.

Dense configurations

All VMAX3 arrays can be configured with a single engine per cabinet and up to 6 DAE. Alternatively a system can be configured to have 2 engines per system bay with 4 DAEs (up to 480 2.5") drives to provide a much denser storage configuration. With drives up to 1.2TB (10k), the dual-engine systems can contain over 500TB (raw) per rack with 64 host ports and up to 4TB of cache in a single standard floor tile. Figure 35 shows the single engine and dense configuration systems for VMAX3 arrays.
Bay dispersion

VMAX 100K, 200K and 400K system racks can be physically separated by up to 25 meters to avoid columns and other obstacles in the data center without a need to reserve empty floor tiles for future array growth. Any VMAX3 system bay can be placed anywhere in your data center as long as it is within 82 feet (25 meters) of the first system bay which houses the Infiniband Dynamic Virtual Matrix switches. Figure 36 shows a possible dispersion option for 8-engine VMAX 400K with 2 adjacent system bays and 6 system bays dispersed at a distance of 25M each from system bay 1.
Third-Party racking

All VMAX 100K, 200K and 400K arrays support industry standard 19-inch racks and optional third-party racking to conform to your data center infrastructure. Third Party racks must meet the dimensions set out in the EMC VMAX Family VMAX 100K, 200K, 400K Planning Guide available on https://support.emc.com

CONCLUSION

This paper was intended provide an introduction to the VMAX3 family of arrays and features yet there's much more to learn about VMAX3 capabilities which cannot possibly be listed in this short whitepaper. With VMAX3, the industry’s leading tier 1 array has evolved into a thin hardware platform with complete set of rich software data services servicing internal and now external block storage. VMAX data services are delivered by a highly resilient, agile hardware platform that offers global cache, CPU flexibility, performance and HA at scale able to meet the most demanding storage infrastructures.

VMAX3 arrays are designed and built for management simplicity, extreme performance, hyperconsolidation and massive scalability in a dense footprint. With the VMAX3 Family, storage can be rapidly provisioned with the Service Level Objective deployment model.

The VMAX3 Family introduced the storage industry’s first Dynamic Virtual Matrix that brings data services closer to the storage they access reducing the sprawl of hardware appliances in the datacentre. The VMAX3 Data Services Platform allows flexible storage infrastructure decisions that are not bound by what is capable within a “frame”. A workload’s performance, response time and data services data can all be presented via the VMAX data services platform though the actual storage device may come from an external device with FAST.X (on or off premise and CloudArray) reducing cost/GB and cost to serve with a platform that is steeped in heritage established on tried and tested Reliability, Availability and Serviceability (RAS). This approach provides hyper consolidation, excellent Total Cost of Ownership, simple and agile management and is able to meet customers’ current and future needs for Platform 2, 2.5 requirements. VMAX3 allows customers to Automate, Modernize & Converge their entire IT infrastructure to deliver IT as a service.
REFERENCES

A full list of VMAX3 Whitepapers and Technical Notes can be obtained here

EMC VMAX Family VMAX 100K, 200K, 400K Planning Guide

Solutions Enabler 8.1 Documentation Set