Technical White Paper

Minimizing Production Downtime with FlashStack Enabled by Cisco UCS and Pure Storage FlashArray//X

By Alex Arcilla, Validation Analyst

May 2020

This ESG Technical White Paper was commissioned by Pure Storage and is distributed under license from ESG.

Contents

ESG Technical White Papers

The goal of ESG Technical White Papers is to educate IT professionals about information technology solutions for companies of all types and sizes. ESG Technical White Papers are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objectives are to explore some of the more valuable features and functions of IT solutions, show how they can be used to solve real customer problems, and identify any areas needing improvement. The ESG Validation Team's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments.

Introduction

This ESG Technical White Paper explores how organizations can minimize downtime when migrating a FlashStack environment equipped with Cisco UCS 6200 Series Fabric Interconnects and Pure Storage FlashArray//M to Cisco UCS 6400 Series Fabric Interconnects and Pure Storage FlashArray//X.

Background

Downtime in production environments remains a huge concern for organizations. According to ESG research, organizations most often cited losing customer confidence (51%), revenue loss (46%), and missed business opportunities (41%) as possible business impacts should end-users experience application downtime or lose data (see Figure 1).[1] It does not help that any efforts to prevent these impacts are hindered by complexity within their IT environments, as a separate ESG research study uncovered that 95% of organizations consider their IT environments equally or more complex than they were two years ago.[2]

Figure 1. Top Eight Business Impacts of Application Downtime or Lost Data

Source: Enterprise Strategy Group

In light of complex IT environments, organizations must always consider how easily and quickly they can upgrade both hardware and software that will accommodate evolving workload performance and scalability requirements. It is critical for organizations operating in an ever-changing business environment to make the upgrade process non-disruptive so that production downtime is minimized.

The Solution: FlashStack Modern Converged Infrastructure from Cisco and Pure

FlashStack is a Reference Architecture Plus designed to deliver high performance, high availability, and enterprise-grade features and functionality. It combines best-of-breed technologies from Cisco and Pure Storage to support business-critical

[1] Source: ESG Master Survey Results, *Real-world SLAs and Availability Requirements*, May 2018.
[2] Source: ESG Master Survey Results, *2020 Technology Spending Intentions Survey*, January 2020.

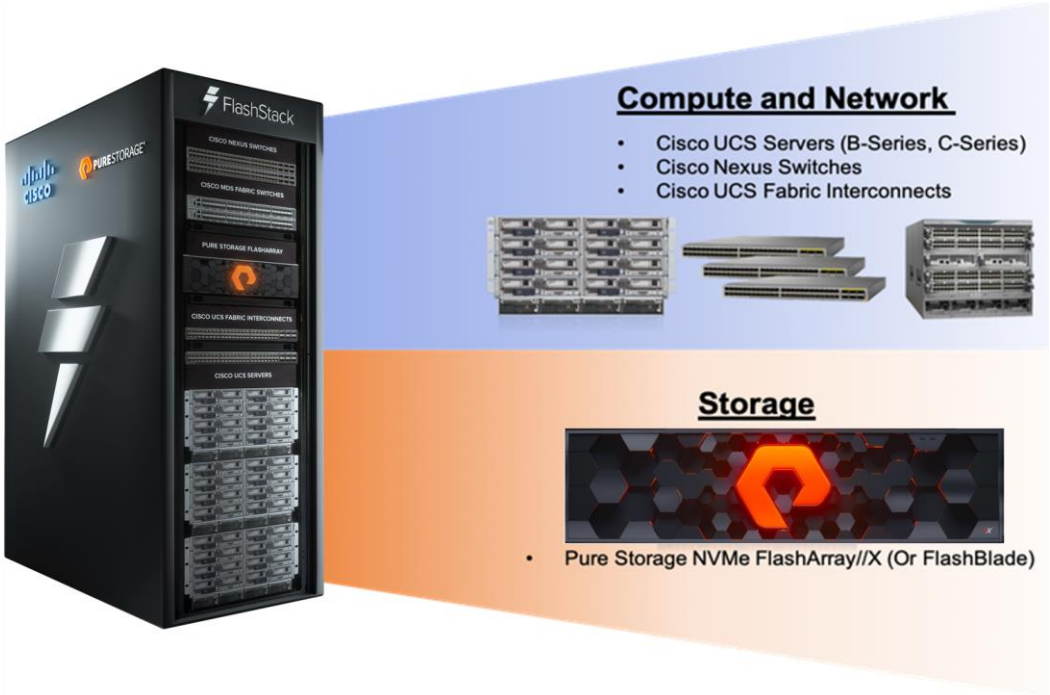
applications, DevOps, and modern analytics. FlashStack is VM-aware, hybrid-cloud-ready, and composed of compute (database, virtualization, application, and management servers based on Cisco UCS platform), network (three-layer network and SAN technologies from Cisco), and non-volatile memory express (NVMe) storage (Pure Storage FlashArray//X) (see Figure 2).

FlashStack can be built using B-Series or C-Series UCS servers. Cisco UCS B-Series blade servers provide FlashStack with an economical high-performance, high-density compute solution. The UCS B200 M5 servers leverage the latest generation Intel Xeon scalable processors and can scale up to 3,072 GB of RAM. A single UCS M5 blade server can provide up to 160 Gbps of throughput connectivity. Currently, a pair of Cisco UCS Fabric Interconnects supports up to 160 Cisco UCS servers in a single Cisco UCS domain, supporting the high scalability of FlashStack.

Pure Storage FlashArray//X is a 100% NVMe solution that enables FlashStack to deliver Pure Storage's modern data experience, providing organizations with scalable, predictable latency and storage density for enterprise-level mixed workload consolidation. FlashArray//X includes:

- 100% NVMe performance: As low as 150 microsecond latency.
- DirectFlash Fabric support: NVMe over fabric (NVMe-oF) via RDMA over converged Ethernet (RoCE).
- Storage and cost efficiency: Average of 5:1 data reduction, 10:1 total efficiency; all array software included.
- Ultra-dense consolidation: 3PBs effective in 6U, designed for 99.9999% availability, plus always-on QoS.
- Simplicity: AI-driven management and predictive support; REST API for cloud orchestration.
- Always-on data encryption and rapid data-locking capability.
- Array and cloud-based snapshots, and ActiveCluster synchronous active/active or asynchronous replication.
- Evergreen Storage business model, combining Pure Storage's non-disruptive upgrades with investment protection.
- Non-disruptive software updates and non-disruptive controller upgrades (no forklift upgrades).

Figure 2. FlashStack Reference Architecture Plus from Pure and Cisco



Source: Enterprise Strategy Group

FlashStack can be deployed in a variety of flexible options to suit the needs of the business, as it is always pretested and validated by both Cisco and Pure. In addition, validated solutions exist that provide faster time to value for many popular

workloads and applications such as Oracle, SAP HANA, VMware (virtual server infrastructure (VSI) and virtual desktop infrastructure (VDI)), Microsoft (VSI), and Citrix (VDI).

How Downtime Is Minimized

ESG analyzed the FlashStack platform to determine how it can help organizations to minimize downtime when the platform is live in a production network. We consider the case in which an organization is upgrading storage controllers to migrate from FlashArray//M to FlashArray//X. The organization also wants to migrate from the Cisco 6200 Series to the Cisco UCS 6400 Series Fabric Interconnect to achieve higher network speeds.[3]

Hardware is constantly changing; new innovations emerge, existing ones become more affordable, and component failures can happen. Therefore, infrastructure solutions that closely tie the software functionality and the configuration information directly to the specific hardware components place a tremendous cost burden on the organization. In order to upgrade or even replace failed hardware components, IT admins must often reconfigure the surrounding environment, wasting personnel cycles and introducing unnecessary risk to the environment. With stateless architecture, upgrading to the latest technology or replacing a failed controller does not require the burden of reconfiguring the environment.

Ease of Pure Storage Controller Upgrades

Pure Storage's FlashArray//X relies on a stateless storage controller architecture to enable non-disruptive controller upgrades. Host I/O write data is not stored on the FlashArray//X controllers, but on decoupled NV-RAM devices. Integrity of the data currently stored in the FlashArray//X is preserved.

Because data integrity is preserved, controller upgrades can be completed in a live production network with minimal performance impact. As the front-ends of the dual controller architecture are active/active, paths from each controller are cross-connected to redundant back-end storage shelves. Both controllers continuously communicate with each other.

The primary controller passes traffic to a designated primary back-end. Both the primary and secondary back-ends have been built and sized to handle all workloads supported by FlashArray. In the event that one controller fails, the other controller can handle the entire storage system load.

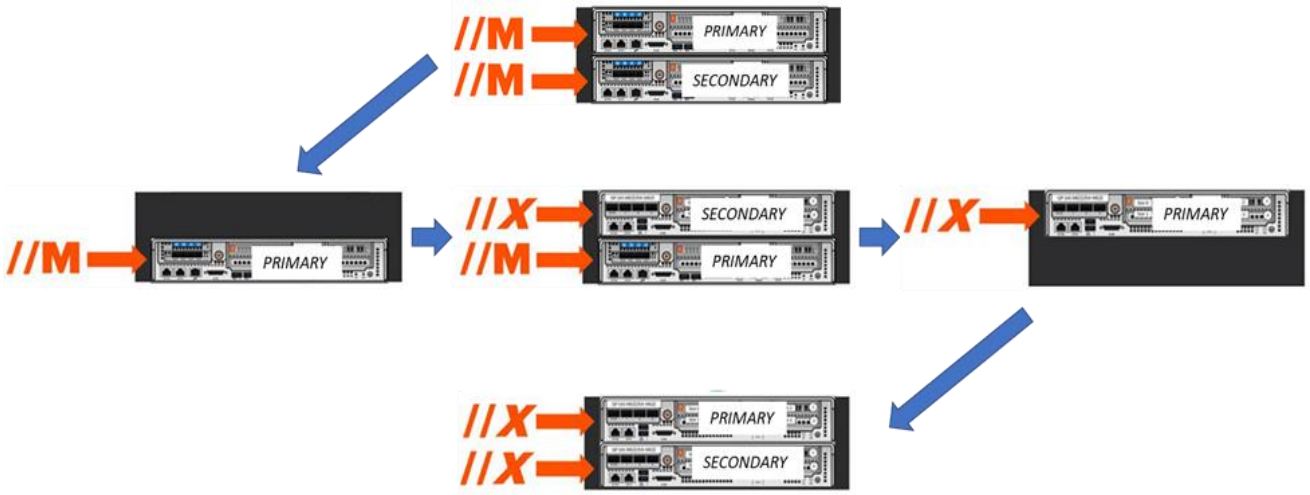
Organizations can non-disruptively upgrade the controllers, one at a time, by failing over traffic to one controller while replacing the other. Since the controllers are cross-connected with each back-end, FlashArray will remain in service. ESG's past evaluation of controller failovers in previous versions of the FlashArray have shown that the failover time was minimal and did not adversely affect I/O.[4] With non-disruptive upgrades, IT no longer needs to designate maintenance windows that interrupt normal IT operations. This is especially important as more enterprises consolidate and mix workloads that are "mission-critical" to the business on a single FlashArray//X, making downtime unacceptable.

In Figure 3, we illustrate the simplicity of upgrading the controllers when migrating from FlashArray//M (a previously supported model within FlashStack) to the FlashArray//X.

[3] The Cisco 6454 supports 48 10/25 GE and Fibre Channel over Ethernet (FCoE) ports, six 40/100 GE and FCoE ports, and eight ports supporting both 10/25 GE and 8/16/32 Gbps Fibre Channel.

[4] Source: ESG Lab Validation, *Pure Storage Purity ActiveCluster: Synchronous Replication with Automatic Failover*, March 2018.

Figure 3. Upgrading Controllers When Migrating from FlashArray//M to FlashArray//X



Source: Enterprise Strategy Group

Using Pure Storage's management CLI, the secondary controller is designated the primary controller, prompting the array to direct all storage traffic through the secondary controller. The FlashArray//M's primary controller is physically removed and replaced with a FlashArray//X controller. Once installed and updated with the latest software release, the FlashArray//X controller is designated the secondary controller.

To remove the remaining FlashArray//M controller, a storage administrator must first designate the newly installed FlashArray//X controller as the primary. When the FlashArray//M controller is shut down, all storage traffic flows through the primary FlashArray//X controller. The older controller is removed and replaced. Once updated with the latest software release, the redundant FlashArray//X controller becomes "secondary." Both FlashArray//X controllers are now active.

Upgrading controllers in this fashion also helps to minimize overall costs, as no additional systems or tools are necessary to migrate data off the array when performing the controller upgrade. As the FlashArray has been sized to keep all virtualized workloads running regardless of whether a controller goes down, data can remain in place.

The Impact of Pure Storage's Evergreen Storage

To further simplify upgrades, Pure Storage designed the Evergreen Storage program to remove the burden of the typical three- to five-year upgrade cycles. It combines Pure Storage's non-disruptive upgrade capability with other benefits that enable subscribers to reduce the overall maintenance and cost of Pure Storage arrays over their lifecycle (see Figure 4).

With Evergreen, enterprises can upgrade any storage array component, depending on their specific business requirements. At a minimum, Evergreen subscribers receive a free controller every three years ("Free Every Three"). If subscribers wish to upgrade the controllers within the three-year period (e.g., take advantage of newer controllers released during that time), they can also trade in older controllers for next-generation releases, or even skip generations via Upgrade Flex.

Figure 4. Pure Evergreen Storage

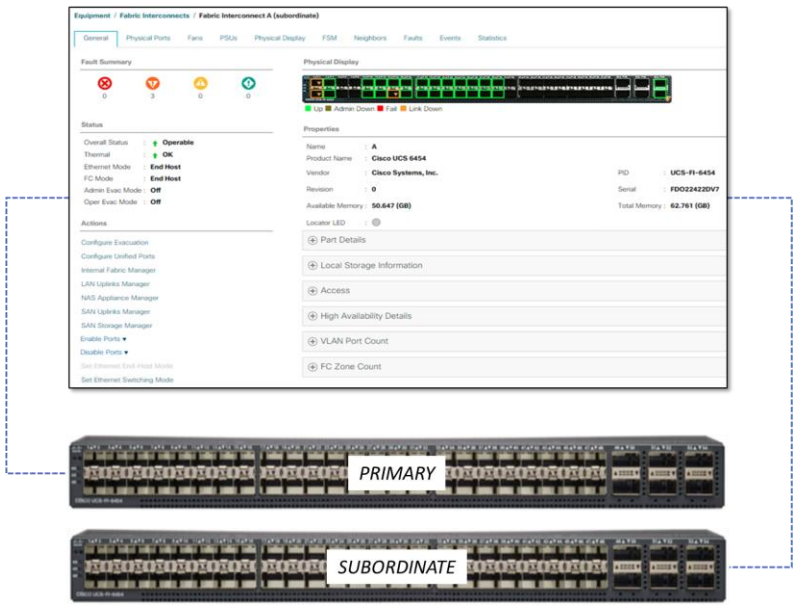
Source: Enterprise Strategy Group

Pure Storage's approach to non-disruptive upgrades is not limited to the controllers. Subscribers can upgrade to newer flash storage media without the need to migrate data off the current FlashStack array. For those migrating to FlashArray//X, customers can leverage Capacity Consolidation to receive trade-in credit for SSDs toward the cost of the newer NVMe drives. The Evergreen storage program also covers non-disruptive software upgrades, as all subscribers are entitled to future software updates and releases.

Non-disruptive controller upgrades are only one facet of the Pure Evergreen Storage program that helps organizations to decrease downtime. Evergreen comprehensively helps customers to decrease business downtime for the life of the FlashStack platform.

Ease of Cisco UCS Fabric Interconnect Upgrades

To further ease upgrades, Cisco UCS also employs a stateless architecture. Enterprises leverage this architecture via service profiles. Rather than configuring hardware components individually to support a given workload, a service profile defines, in software, the server, storage, and network resources required for a specific workload.



Service profiles are created in software via the Cisco UCS Manager, the management GUI, which is hosted on the Cisco UCS Fabric Interconnects. Once created, the service profiles are stored within the fabric interconnects and deployed via the Cisco UCS Manager. Essentially, each service profile treats FlashStack as a hardware resource pool.

When a service profile is deployed, the Cisco UCS Manager automatically configures the server, adapters, fabric extenders, and fabric interconnects to match the configuration specified in the profile. This automation of device configuration reduces the number of manual configuration steps, further minimizing potential downtime.

Also, since the hardware resources have already been configured via service profiles, workloads will

continue to operate. The upgrade process does not require any reboot of the application or virtual machine running already on FlashStack.

The process for upgrading the fabric interconnects is similar to the process for upgrading the FlashArray storage controllers. Because the service profiles are maintained by both fabric interconnects, service will not be disrupted when shutting down either one. Service remains uninterrupted.

When upgrading the fabric interconnects, the same firmware must be running on both the Cisco UCS 6200 Series and 6400 Series before hardware is replaced**.** Once that is verified, an administrator evacuates all traffic and shuts down the active uplink and server ports of the subordinate fabric interconnect before uncabling and removing it.

After cabling the new subordinate Cisco 6454 Fabric Interconnect and reconfiguring the uplink and server ports, it will automatically synchronize with the configuration and database and state information from the primary Cisco UCS 6200 Fabric Interconnect. Once data paths are verified between the Cisco 6454 Fabric Interconnect and the servers, the 6454 is promoted to "primary" status. The process is repeated to remove and replace the remaining Cisco 6200 Fabric Interconnect. During this time, no disruption in traffic has occurred. Downtime is minimized, and mission-critical workloads continue to operate.

Why This Matters

Organizations using a reference architecture will need to upgrade hardware components to meet evolving business requirements. Typically, enterprises need to impose maintenance windows or shut down the production network. For mission-critical applications that organizations must operate around the clock, this is simply not acceptable.

The FlashStack Modern Reference Architecture Plus from Cisco and Pure Storage relies on a stateless architecture to perform in-service, non-disruptive upgrades of both the FlashArray controllers and the Cisco UCS Fabric Interconnects. By leveraging the active/active redundancy and the stateless architecture of both the FlashArray controller and the Cisco UCS Fabric Interconnects, organizations can upgrade both while workloads continue to run with minimal performance impact, thus decreasing application downtime and data loss.

The Bigger Truth

For any downtime incurred in a live production IT environment, organizations risk losing customer confidence, revenue, and potential business opportunities.[5] While they cannot afford these losses, upgrading critical IT resources in the field can still present a significant challenge in today's complex IT environments. To mitigate any risks, enterprises should use solutions architected for non-disruptive upgrades so as to not incur application downtime or lose data.

FlashStack is a Reference Architecture Plus, combining Pure Storage FlashArray and Cisco UCS, that can help organizations decrease overall downtime in their business environments. By leveraging a stateless architecture, organizations can upgrade both the FlashArray controllers and the Cisco UCS Fabric Interconnects without disrupting ongoing business operations. IT can upgrade these components without shutting down the entire system and migrating workloads to another system to continue service. Instead, upgrades are completed while the system remains online, minimizing the need for extensive planned downtime or maintenance windows. Organizations can then continue to run their mission-critical workloads with little risk to the business while minimizing operational expenses related to maintenance upgrades.

Pure Storage also increases uptime via Evergreen Storage, a program designed to extend the life of customers' investments in Pure Storage solutions. With the "Free Every Three" and "Upgrade Flex" program components, subscribers can upgrade controllers to take advantage of newer technology and features without performing a forklift upgrade of the entire system.

[5] Source: ESG Master Survey Results, *Real-world SLAs and Availability Requirements*, May 2018.

The program also aids in minimizing capital expenses, as well as the time typically spent on planning and purchasing newer IT resources.

The evaluation presented in this document is based on ESG's analysis of the FlashStack architecture. Due to the many variables in each production environment, it is important to perform planning and testing in your own environment to validate the viability and efficacy of how FlashStack can indeed minimize downtime typically related to upgrades.

Organizations, especially those running mission-critical applications, cannot afford to experience downtime, as the consequences are dire to the business. Performing upgrades in your IT infrastructure is a given. To decrease complexity, downtime, and cost, in-service upgrades are ideal. ESG believes that FlashStack enables an upgrade process that will reduce IT complexity and minimize overall downtime. If application downtime and lost data are tangible risks to still mitigate, ESG suggests taking a closer at the FlashStack Reference Architecture Plus.