

ESG Economic Validation

Quantifying the Value of Pure FlashStack Converged Infrastructure Solution versus Cloud-based IaaS

By Aviv Kaufmann, Senior IT Validation Analyst

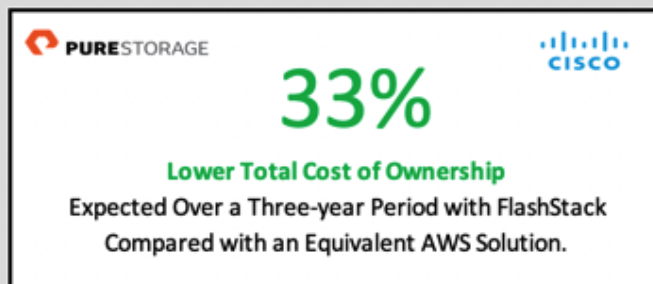
March 2019

Executive Summary

Converged infrastructure solutions have proven to greatly simplify on-premises infrastructure. Choosing the right converged solution can provide your organization with cloud-like agility at a cost savings, while allowing you to control your own operations and freeing you from dependence on the public cloud.

ESG was engaged by Pure Storage and Cisco to develop a quantitative model examining how an investment in FlashStack Converged Infrastructure Solutions could deliver efficiency and cost savings over a three-year period when compared with public cloud-based IaaS. The analysis is designed to help IT organizations determine the fully burdened costs and benefits of deploying a FlashStack solution based on Pure Storage FlashArray//X-series all-flash storage arrays versus satisfying the same requirements with cloud-based compute instances and storage. This analysis is built upon ESG's evaluation of FlashStack, ESG qualitative and quantitative market research with IT decision makers, and ESG's general familiarity with on-premises and public cloud IaaS offerings. This analysis is designed to provide prospective customers with a comprehensive picture of the potential direct and indirect cost and benefit drivers they should contemplate when considering an investment in FlashStack converged solutions.

ESG leveraged this model to compare the expected costs of deploying 700 consolidated virtual servers in the AWS cloud versus satisfying the same technical requirements with a FlashStack solution over a three-year period. The model predicts that, through consolidation and efficiency savings, the FlashStack on-premises solution will result in an expected total cost of ownership that is 33% lower than the expected total costs of the AWS solution.



Introduction

This ESG Economic Validation focuses on the quantitative and qualitative benefits organizations can expect by deploying FlashStack on-premises compared with leveraging AWS EC2 compute and EBS storage to handle the same workloads.

Challenges

It is the goal of every IT organization to quickly, safely, and cost-effectively deliver services that are agile enough to support the ever-changing needs of the business without noticeable disruption. The task often proves quite challenging to deliver on-premises due to the planning, testing, and troubleshooting required to bring application, compute, network, and storage services from leading vendors together into a high-performance, highly available, scalable, and secure solution. It is no surprise that 66% of IT organizations believe that their IT environment is more complex than it was just two years ago.¹

As a result, many organizations have turned to the flexibility and agility offered by public cloud-based infrastructure-as-a-service (IaaS) solutions like Amazon Web Services (AWS). IT organizations see advantages in these public cloud offerings provided by their fast spin-up and provisioning, OpEx-based consumption, and minimal resources needed for management. However, a recent ESG study revealed that quite a few organizations have decided to repatriate workloads that they had moved to cloud-based services back to on-premises infrastructure. These organizations most cited improved data security, improved/more predictable performance, improved availability/reliability, improved data recovery capabilities, and more predictable costs as the top benefits that their organization has achieved or expects to achieve from moving public cloud-based applications/workloads back to its own data centers (see Figure 1).²

Figure 1. Benefits of/Aspirations for Cloud Repatriation



Source: Enterprise Strategy Group

The challenge of many organizations is to deliver cloud-like agility and flexibility on-premises while limiting the complexity and operational cost of maintaining point solutions. For this reason, many have turned to converged infrastructure solutions, especially ones that have cloud-like advantages in their technology and business models. Pretested, pre-integrated converged infrastructure solutions provide a faster time to value and improved operational efficiency when

¹ Source: ESG Research Report, [2019 Technology Spending Intentions Survey](#), February 2019.

² Source: ESG Master Survey Results, [Tipping Point: Striking the Hybrid Cloud Balance](#), October 2018.

compared with do-it-yourself (DIY) systems. And offerings that further reduce the ongoing management and upgrade complexity offer IT organizations even more advantages. The right solution can offer improved performance, security, efficiency, and even cost savings when compared with public cloud offerings.

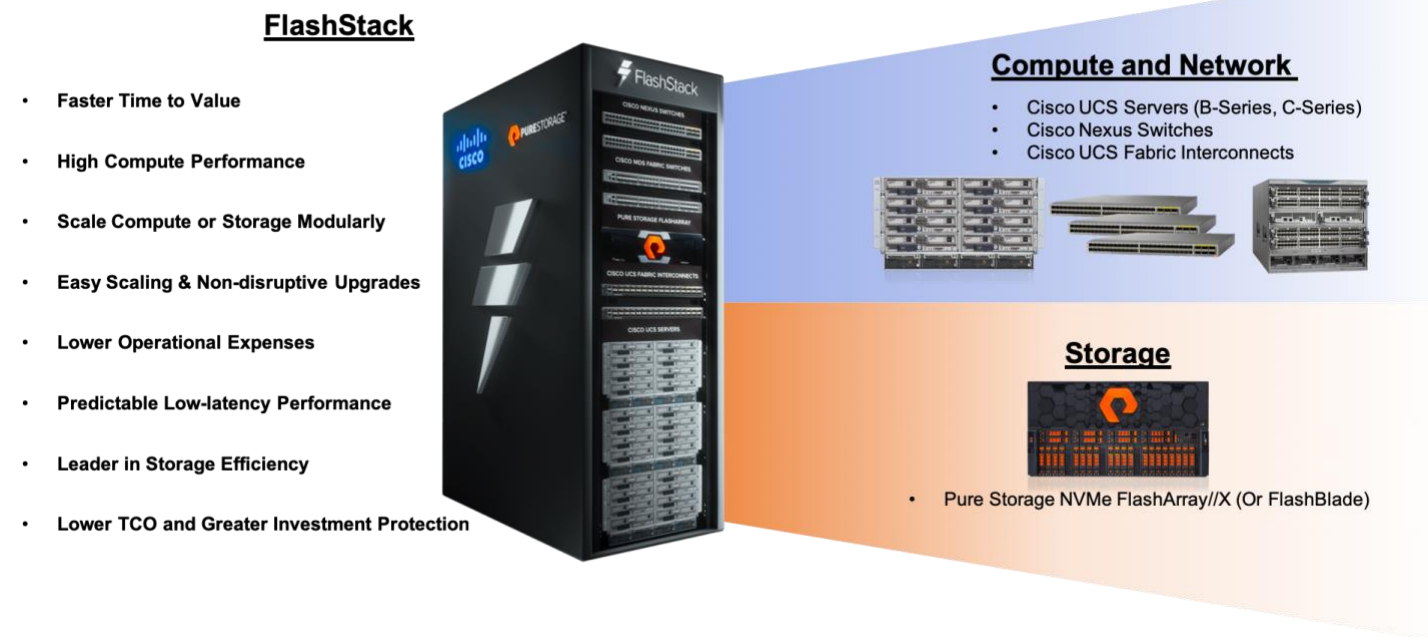
The Solution: FlashStack Converged Infrastructure Solutions Powered by Pure Storage

FlashStack is a high-performance, highly available converged infrastructure platform that combines best-of-breed technologies from Pure Storage and Cisco into a powerful, enterprise-grade, converged platform for business-critical applications, DevOps, and modern analytics. FlashStack is VM-aware, hybrid cloud ready, and comprised of compute (database, virtualization, application, and management servers based on Cisco UCS platform), network (three-layer network and SAN technologies from Cisco), and NVMe storage (Pure FlashArray//X or FlashBlade) (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 80 Gbps of throughput connectivity and can be provisioned with up to two SSDs or HDDs for localized storage.

Pure FlashArray//X is a 100% NVMe solution that provides FlashStack with Pure's data-centric architecture to deliver scalable, predictable latency and storage density for enterprise-level mixed workload consolidation. FlashArray//X includes:

- 100% NVMe performance: As low as 250 µs latency; Up to 2x faster than previous gen AFAs; NVMe & NVMe-oF.
- Efficient & affordable: Average of 5:1 data reduction, 10:1 total efficiency; all array software included.
- Ultra-dense consolidation: 3PBs effective in 6U; proven 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.
- Non-disruptive software updates and non-disruptive controller upgrades (no forklift upgrades).
- Evergreen Storage business model, combining Pure's non-disruptive upgrades with investment protection.

Figure 2. FlashStack Modern Converged Infrastructure 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, and 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, VMware, and SAP HANA among many others.

FlashStack Economic Validation Overview

ESG's model predicts that deploying FlashStack on-premises can provide cost savings and opportunities for economic benefit when compared with satisfying the same requirements using a public cloud infrastructure-as-a-service (IaaS). ESG found that FlashStack provides economic savings in three primary areas:

1. Compute Density and Consolidation Advantages
2. Storage Efficiency Savings
3. Performance and Cost Predictability

ESG reviewed Pure's economic value propositions by evaluating a series of existing customer case studies, comparing that information with knowledge of cloud-based solutions, and comparing the results of technical and economic models.



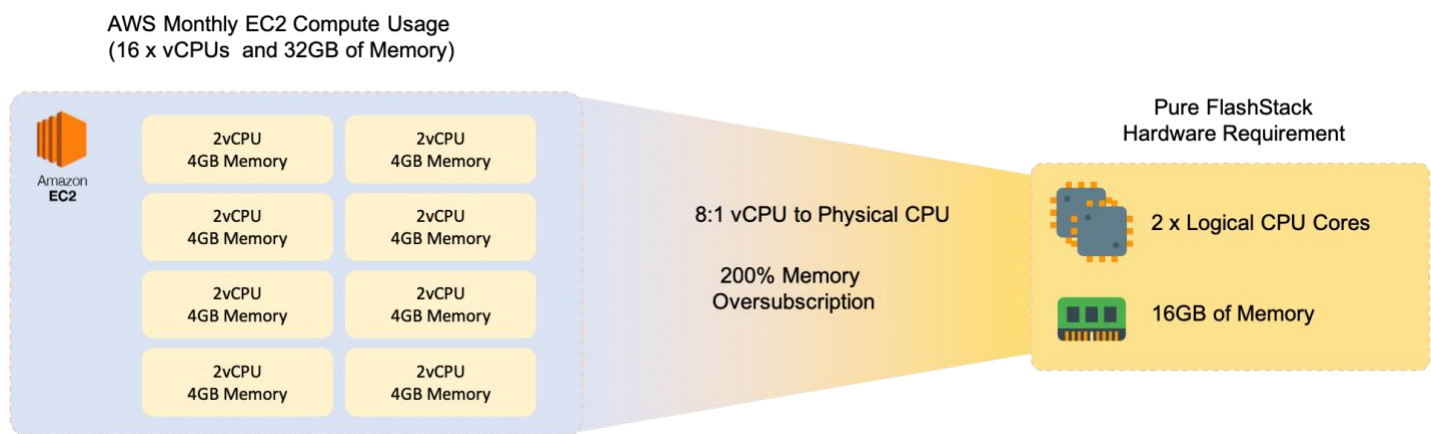
Compute Density and Consolidation Advantages

ESG's model (based on B-Series UCS blade servers) demonstrates how organizations can expect to reap the economic benefits of compute and memory density for their virtualized workloads. Virtual servers and virtual desktops can be deployed on shared compute and memory resources, allowing workloads to share physical resources and extending the value of the hardware investment. This is known as oversubscription, which is a common practice in virtualized environments. For instance, multiple VMs can be deployed using a single physical or logical core if the deployed virtualized machines require only a fraction of its cycles (VMware recommends a ratio of 6:1 to 8:1 for most environments). Similarly,

since not all memory will be typically consumed at once, and not all VMs will be powered on at the same time, memory can be oversubscribed in a similar manner. Memory also benefits from a form of compression and cross-OS deduplication known as transparent page sharing (TPS) and sharing of RAM through the use of a memory balloon driver.

The compute instances running on physical hardware in the AWS cloud benefit from these same density savings, but the savings are realized by the IaaS provider (AWS) and not passed on to the consumer, who effectively pays monthly for every vCPU and GB of memory that they have deployed regardless of the physical resources that are actually required to support the instances. By deploying a FlashStack on-premises, the organization realizes the economic benefits of compute and memory density. While the actual savings are dependent on such factors as processor technology, vSphere version, and workload requirements, the net result is that an on-premises FlashStack deployment can use fewer physical cores and less memory to satisfy the technical requirements of virtual machines and consolidate AWS workloads.

Figure 3. Consolidation of AWS Instances on FlashStack



Source: Enterprise Strategy Group



Storage Efficiency Savings

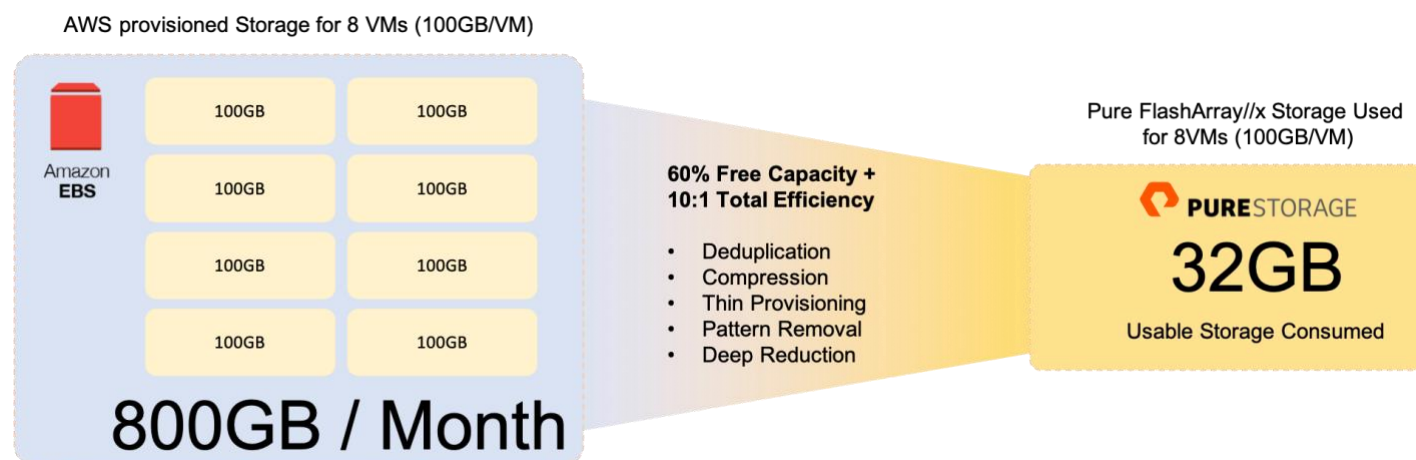
The latest FlashStack offerings are built with Pure Storage FlashArray//X all-NVMe flash accelerated storage arrays to consolidate workloads and handle the high-capacity and high-performance storage requirements that are demanded from a converged infrastructure. FlashArray//X scales seamlessly to provide storage capacity and predictable latency for all dedicated and shared resources that are deployed on FlashStack.

Pure Storage arrays provide some of the highest performing data reduction technologies available. Like the compute and density advantages described above, FlashStack passes the savings provided by the disparity between capacity required (effective capacity for workloads) and capacity stored on FlashArray (used capacity written to the array) on to the organization, allowing for far less capacity to be deployed than is required. As is the case with compute and memory, you must pay for every GB of storage that you provision on AWS (whether it has been used or not) and any advantages of data reduction are realized by AWS, not the consumer.

Pure Storage provides “always-on” data reduction for all workloads that includes the highest efficiency global, inline deduplication and compression, pattern detection and removal (including zeros), copy reduction, thin provisioning, and post-process compression. The overall data reduction rate and total efficiency (as seen across existing deployments) are reported in real time on the Pure website, with an impressive 5:1 average global data reduction ratio and 10:1 total

efficiency ratio (including thin provisioning). In addition, Pure provides a Right-Size Guarantee (RSG) for overall efficiency for new deployments, which can be extended with any capacity increase. The Pure FlashArray//X can deliver 3PB effective in just 6U of rack space, further providing savings by lowering power, cooling, and floorspace costs per GB, and lowering administrative costs by allowing administrators to consolidate more applications into a single array.

Figure 4. Example: Storage Capacity Savings to Provide Each of Eight VMs with a 100GB Volume



Source: Enterprise Strategy Group



Performance and Cost Predictability

Pure Storage FlashArray//X was purpose-built with end-to-end shared and accelerated NVMe storage that is up to 2X faster than previous generation AFAs at zero cost premium (\$/GB) versus SAS-based SSDs. FlashArray//X can scale to as high as 3PB of effective capacity. Array components, including controllers and flash modules, can be non-disruptively upgraded to improve capacity and performance when needed, and are designed to be backward-compatible, further simplifying upgrades and increasing array agility. In addition, Pure's Evergreen storage subscriptions allow for years of non-disruptive upgrades of hardware and capacity with no wasted investment. All of this means that organizations are able to consolidate and grow workloads within FlashStack based on their needs, while delivering predictable, low-latency performance, agility, and simple capacity management. While the cloud is built for scale, performance can be unpredictable at times, making guaranteeing SLAs for workloads difficult (with general-purpose storage) or costly (using provisioned IOPS storage).

Ever-changing workload and storage requirements can make it difficult to predict cloud costs, which can accumulate quickly. An investment in FlashStack provides cost predictability while still delivering cloud-like agility to handle the unpredictable needs of the business. While some capacity and performance overhead must be purchased upfront to provide this agility, the costs are known. In contrast, keeping track of cloud spend across all departments within an organization can require substantial operational overhead.

Economic Modeling

ESG's Pure FlashStack Economic Value Model

ESG leveraged the information collected through vendor-provided material, public and industry knowledge of economics and technologies, and the results of customer studies to create a TCO/ROI model that compares the costs and benefits of deploying, managing, and supporting FlashStack on-premises versus deploying the equivalent compute and storage

requirements with AWS EC2 instances and EBS storage over a three-year period. Because no two organizations are the same, the model was designed to allow for great flexibility in characterizing the organization's environment and its workload requirements. The validated assumptions and logic contained in this model are also used as the basis for ESG's *FlashStack TCO/ROI Calculator* in use by Pure Storage field sales teams.

Please note that cloud costs change often and vary by region. The default AWS costs used in this modeled scenario were based on those published for the US East N. Virginia region at the time of publication.

Example: Using the EVM to Predict Savings for a Modeled Organization

To illustrate the economic advantage that an organization can achieve through a FlashStack deployment, ESG ran a set of assumptions through the ESG Economic Value Model. The scenario modeled a large organization with requirements to consolidate its 700 virtualized servers. The virtualized servers consisted of bronze, silver, gold, and platinum level offerings. The requirements for the virtualized servers are shown in Table 1.

Table 1. Virtualized Server Requirements

	Qty.	vCPU / VM	Mem / VM	GB / VM	IOPS/VM	CPU Class	Storage Class
Bronze	200	1	4	30	100	Standard	Standard
Silver	200	2	4	60	250	Standard	Standard
Gold	200	4	8	120	500	Standard	Standard
Platinum	100	8	16	240	1,000	High Freq.	Low Latency

Source: Enterprise Strategy Group

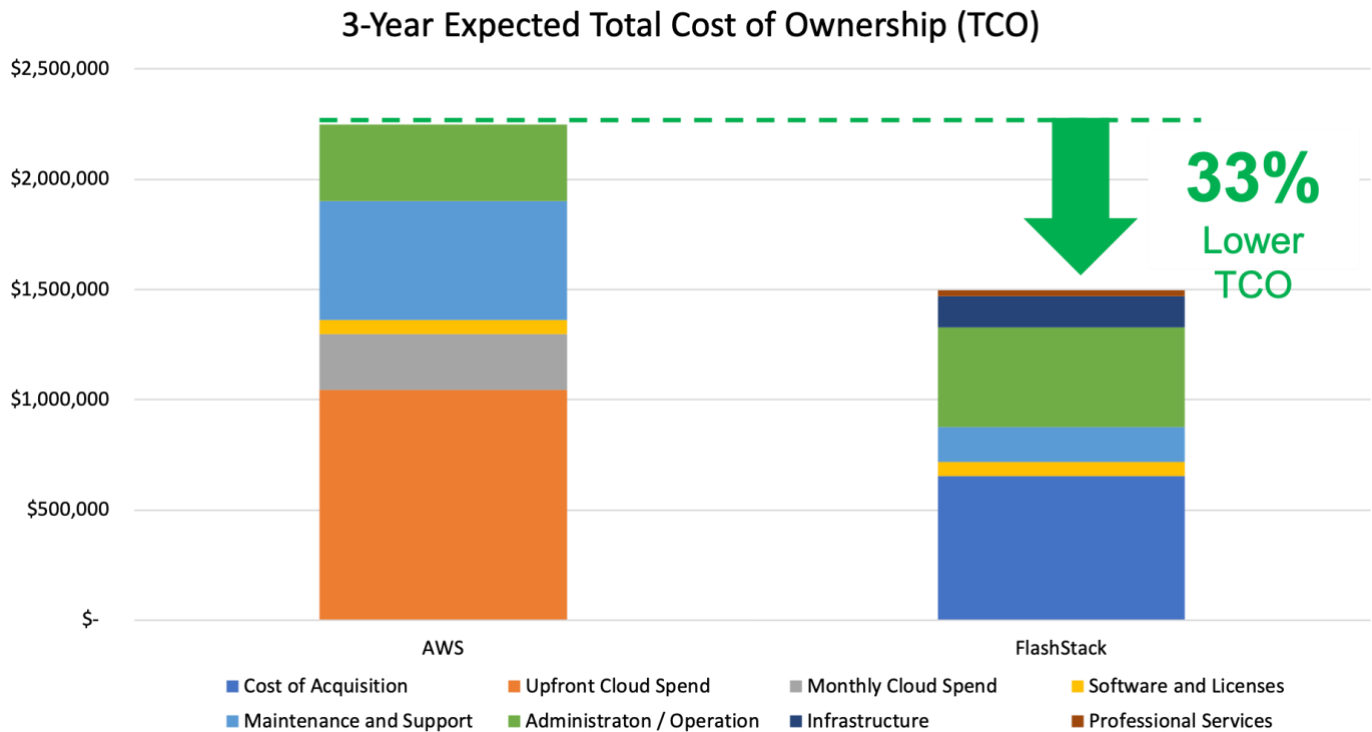
To meet the requirements on AWS, ESG assumed that the bronze and silver leveraged c5.large instances, the gold leveraged c5.xlarge instances, and the platinum required c5.2xlarge instances with io1 provisioned storage to guarantee a similar level of predictable, low latency that would be delivered by the Pure Storage system. All other virtualized servers deployed the cost-effective gp2 storage.

ESG assumed typical discounts on the cost of acquisition for the FlashStack solution. ESG assumed conservative data reduction efficiency (4.2:1) and core overcommit ratios of 10:1 for bronze, 6:1 for silver, 4:1 for gold, and 1:1 for platinum. Memory overcommit was assumed to be 4:1, 2:1, 2:1, and 1:1 for the four levels described above. ESG assumed that the capacity allocated to each VM initially contained 60% unwritten capacity and a 20% free storage requirement on the Pure Storage array to allow for annual growth in capacity.

ESG conservatively assumed that the compute instances for the AWS solution would be paid using a three-year all upfront option that results in the greatest savings (as high as 75% savings for some instance types) to the organization at the expense of limiting the flexibility of the solution. Since software requirements can vary drastically, ESG assumed that any additional software licenses required for software running on the VMs in the cloud or on-premises would be equivalent and conservatively estimated licenses as 10% of hardware spend.

With this information entered, the ESG EVM calculated that by deploying Pure FlashStack on-premises, the modeled organization could expect total savings of roughly \$752K over a three-year period instead of deploying on AWS. As expected, the costs to operate the Pure FlashStack solution were higher, but these additional costs were much lower than the savings provided in other key areas. The model predicted additional spending of \$144K for power/cooling/floorspace and Internet connectivity and \$26K for professional services to install the solution. Administration costs for the FlashStack solution were roughly 30% higher and considered hardware, cloud instance, and virtualization/application administration. The total savings resulted in an expected ROI of 33%. A summary of the modeled results is shown in Figure 5.

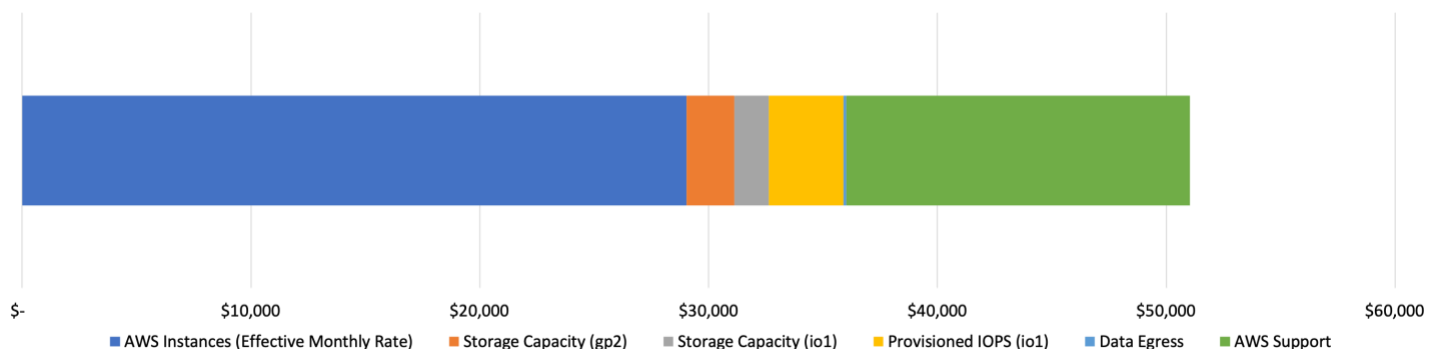
Figure 5. Summary of Savings Expected with Pure FlashStack for a Medium-sized Organization



Source: Enterprise Strategy Group

The majority of the savings were a result of the cost of the FlashStack solution being lower than the combined cost of the three-year upfront reserved instance pricing and cumulative monthly cloud spend for storage capacity, provisioned IOPS, and expected network egress traffic (minimal). In addition, the AWS solution required much more expensive support costs, with the enterprise support starting at a minimum of \$15K per month and growing from there as a percentage of monthly cloud spend. For this scenario, ESG’s model predicted a support and maintenance cost that was 71% lower, and upfront cost savings (cost of acquisition versus three-year upfront reserved instances) of \$393K. The reserved instances would all be paid upfront; however, Figure 6 summarizes the effective monthly cloud costs of the AWS solution calculated by dividing this upfront cost over 36 months to better understand the relative cost differences for the AWS solution.

Figure 6. Expected “Effective” Monthly Cloud Spend for the AWS Modeled Scenario



Source: Enterprise Strategy Group

The Bigger Truth

Today's businesses requirements progress and transform at a rate faster than ever before. The burden placed on the IT organization to provide an infrastructure that is scalable, high-performance, and cost-effective has led many to look to cloud service providers rather than deploy, manage, and maintain infrastructure in-house. However, like any service provided to others, over time, the advantages of not having to do work in-house comes at a cost premium, and as workload and capacity requirements grow, the cost savings predicted by the organization can quickly fade away and scale up to a complex-to-calculate cost impediment.

ESG evaluated the benefits of the Pure FlashStack converged infrastructure solution with a focus on its ability to consolidate cloud-based workloads and how the solution can pass the economic savings provided by compute, memory, and storage density and efficiency advantages to the organization, rather than to the cloud service provider. In addition, the predictable, low latency provided by the Pure Storage//X array allows organizations to consolidate workloads with costs that are highly predictable as workload requirements shift. While there is still some manpower involved to manage FlashStack, Pure and Cisco have streamlined the planning, deployment, management, and maintenance of the solution to help reduce the expected operational cost of the infrastructure.

ESG leveraged knowledge of the industry and solutions to create an economic pricing model to analyze the expected costs of deploying 700 mixed-requirement virtual servers leveraging AWS EC2 and EBS storage versus deploying a FlashStack solution to meet the same set of technical requirements on-premises. While some operational costs were higher as expected (i.e., administration, deployment, power/cooling/floorspace, and professional services), the FlashStack solution provided a three-year total cost of ownership that was 33% lower than that of the AWS solution. The ability to overcommit compute and memory and take advantage of Pure's storage efficiency savings passed much of the savings generally realized by the cloud service provider on to the organization.

The real-world scenario ESG looked at provides a glimpse into potential savings and benefits that Pure FlashStack can deliver by helping organizations to cost-effectively consolidate workloads while delivering predictable low-latency performance and cloud-like agility and flexibility. Because every organization's requirements are different, ESG recommends that organizations consider contacting Pure Storage to analyze the potential savings that can be provided based on their own business requirements. If your organization is ready to take back control of your infrastructure, Pure FlashStack may be the perfect converged solution for you.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.



Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides market intelligence and actionable insight to the global IT community.

© 2019 by The Enterprise Strategy Group, Inc. All Rights Reserved.

