Marc Staimer, President & CDS Dragon Slayer Consulting

W H I T E P A P E R

Eliminating Application SDS Performance And Capacity Management Contortionism

With Predictive Analytics

Eliminating Application SDS Performance & Capacity Management Contortionism

Marc Staimer, President & CDS of Dragon Slayer Consulting

Introduction To A Nightmare

It's 2:20 AM when the IT director is suddenly woken by their mobile phone buzzing urgently with an alarm from the data center. Several mission-critical applications are running horrendously slow. They've dropped below performance thresholds. Adrenaline racing, the IT director hastily gets out of bed, jumps online to the web portal, and verifies there is indeed a catastrophic mission-critical application performance issue occurring and getting steadily worse. Throwing on some clothes, the IT director quickly jumps in the car, drives rapidly into the office, rushes to the data center to see what and where the problem is, to figure out the cause, and determine how to fix it. The IT director upon arrival finds that the graveyard shift has not yet isolated the problem. In fact, they are highly frustrated and have no idea what's causing it. As the entire team digs in to try to find the problem cause it suddenly goes away. Nothing changed but the mission-critical application seems to have recovered and performance is back within normal operating parameters. What the heck happened? A painstaking post mortem fails to reveal the answer and the problem can't be solved unless it occurs again (and it will at the most inconvenient time possible according to Murphy) and can be trapped, traced, and isolated. That requires sophisticated hardware, software, and extensive expertise. This may sound like an outlier nightmare scenario, but regrettably, it occurs far too frequently in today's 7 by 24 by 365 data center.

The betting odds are that a performance problem will most likely occur somewhere in the storage. IT pros have a love-hate relationship with their storage. Storage is a basic necessity today for all operating systems, file systems, hypervisors, and applications. That may change sometime down the road as the barrier between memory and storage is breached. But for the foreseeable future, storage is an absolute necessity.

The hate part of the relationship comes from the fact that storage systems have traditionally required a high level of specialized expertise and manually labor-intensive tasks. These manual, labor-intensive tasks such as system set up, LUN or file system provisioning, data protection and frequency per LUN or file system, and of course ongoing management. Storage systems evolved over the years to become simpler and more automated reducing the required expertise and increasing amounts of manual requirements. However, storage also tends to be the root cause for numerous application performance problems. Performance problems such as inconsistent response times, application timeouts unacceptably slow response times, and even application crashes.

Over the years, external shared storage systems have evolved to have some historical analytical capabilities that help pinpoint where system performance problems are occurring. The problem could be in a specific physical volume or LUN, virtual equivalents, storage controller, front-end controller, back-end IO, caching tier, etc. The rapid market trend to lower cost software defined storage (SDS) has made analytics orders of magnitude more difficult. Storage can reside in different server nodes, different external storage systems, across different networks, and different infrastructure. Isolating, let alone anticipating, application performance problems in a SDS ecosystem can be an overwhelming challenge.

The love part of the relationship comes from the fact that although storage or SDS is the cause of their performance problems, it also tends to be the fix. Sometimes it's as easy as increasing storage capacity, or adding more storage drives, or adding faster drives such as solid-state flash drives (SSD), or adding more memory, etc. In too many cases these fixes do not solve the performance ills. They merely shift the performance bottleneck from one place in the storage architecture to another. Figuring out the actual application performance problem storage root cause is a time consuming, labor-intensive exercise.

The other IT pro nightmare scenario that no storage administrator likes to talk about is when an application runs out of storage. That makes the previous nightmare scenario seem like a walk in the park. No IT pro ever wants their applications, especially mission-critical applications, to run out of storage. Very bad things transpire when that happens. The applications become unstable and/or crash and/or their data becomes corrupted. IT director phones light up like the northern lights during a solar flare and spend hours, days, even weeks repairing the damage. And yet, as virtualization has permeated most data centers, the probability of that occurring has progressively increased. The most common way to avoid this is to over-provision the amount of storage allocated to the applications. Over-provisioning storage capacity is a "guestimate" approach based on project storage consumption rates for that application. If the guess is too high, storage capacity is wasted or orphaned. If it's too low, it only deferred the problem.

Thin provisioning can alleviate orphaned storage from excessive capacity over-provisioning. Although, that too has been known to be problematic when other thin provisioned LUNs or volumes consume the physical capacity pool more rapidly than expected. And while there might be threshold alerts in the storage system or third-party storage resource management software (SRM), the lack of predictive analytics and velocity tracking can lead to dire situations.

These are the types of problems FalconStor set out to solve with predictive analytics as part of its FreeStor SDS.

Table of Contents

Software Defined Storage (SDS) Application Performance and Capacity Consumption Problems

Software-defined storage (SDS) has exploded into the storage market and with good reason. As the amount of data requiring storage keeps growing at an exponential rate, storing that data the traditional way on legacy arrays or file systems has become operationally and financially unsustainable. SDS holds the promise of significantly reducing storage systems, storage infrastructure, storage networks, storage management, storage laborious tasks, and storage costs (CapEx and OpEx) far in excess of 50%. There are numerous definitions for SDS. The general definition is storage software that is separate from the storage hardware it controls and manages to deliver data storage services. That's a pretty broad definition allowing many implementations. The most common SDS implementation is full storage services suite software that executes as a virtual storage appliance (VSA) or physical storage appliance (PSA).

VSA

The VSA is the storage software stack and services executing on a hypervisor virtual machine (VM) in a physical host. The VSA has become quite popular in hypervisor clusters and hyper-converged infrastructure in that it enables lower cost server side embedded storage hardware media (both HDDs and SSDs) to be shared among all nodes in the cluster. VSAs also virtualize, pool, share, deduplicate, compress, snapshot, replicate, as well as repurpose DAS, SAN, or JBOF storage for physical and virtual machine workloads. The storage interfaces are primarily iSCSI, NFS, and/or SMB.

PSA

The PSA is identical to a VSA except that it executes in a dedicated x86 physical host as an application running on Linux or Windows or even in a Docker container. Everything else is the same as the VSA.

VSAs and PSAs are ordinarily deployed in ALUA (asynchronous LUN access) pairs. This is also known as active/active pairs. Some implementations can be scaled out with multiple clustered pairs or stretched across multiple geographic locations.

Problem Solving Frustration

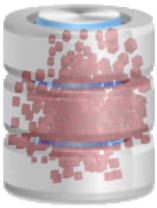
A key issue for SDS VSAs or PSAs is the underlying infrastructure complexity. SDS masks that complex infrastructure and makes it appear simple via virtualization. Masking complexity is a good thing until something breaks and that mask becomes a major obstacle. When applications have storage performance issues, isolating and figuring out what's causing those issues is a time-consuming and frequently frustrating exercise. Consider the data path for a SDS assuming the target SDS is in a different physical host:

- *The application data exits the VM through the hypervisor file system;*
- *Through the physical host memory;*
- *Onto the memory channel;*
- *Across the physical host PCIe controller;*
- *Onto the PCIe fabric;*
- *Through a physical storage adapter (FC, iSCSI over Ethernet, TCP/IP over Ethernet);*
- *Past a transceiver;*
- *Through a cable;*
- *Onto the storage network;*
- *Past another transceiver;*
- *Through a minimum of one standard Ethernet or Fibre Channel switch;*
- *Back out through yet another transceiver;*
- *Back onto the storage network;*
- *Across a different cable;*
- *Through yet another transceiver;*
- *Into another storage adapter in the VSA or PSA;*
- *Across the VSA/PSA PCIe fabric;*
- *Through the VSA/PSA PCIe controller;*
- *Onto the memory channel;*
- *Thru the memory;*
- *Across the hypervisor file system (if VSA);*
- *Thro the OS or VM;*
- *Processed by the VSA or PSA;*
- *Then back across the same path and piece parts in the SDS;*
- *To the embedded physical storage media or DAS to that physical host;*
- *If the physical storage is external to the SDS host;*

- *Back onto the storage network;*
- *Thru yet another transceiver,*
- *Past the physical storage adapter on the storage system;*
- *Across the storage system PCIe fabric;*
- *Thru the storage system PCIe controller;*
- *Onto the memory channel;*
- *Thru the memory;*
- *Processed in the Storage CPU;*
- *Back out thru the memory, PCIe controller, PCIe fabric, internal storage fabric, storage controller, and storage media;*
- *And back through the same tortuous path to the originating application;*
- *Multiple round trips for every transaction.*

If the target SDS is on the same physical host as a VSA, several of the steps are virtualized, but unfortunately still exist.

All of this is made transparent to the application. Obviously, there are many touch points in the SDS data path any one of which can cause application performance problems. Masking the complexity generally simplifies application storage management but makes application performance troubleshooting a considerable time sink.



What makes this situation pricklier is that the vast majority of SDS systems cannot prognosticate when an application performance problem is going to occur. Some can tell you an application performance problem has occurred when a performance threshold is breached, but not that it will occur based on predictive analytics. Nor will it tell you where the problem lies within the storage system. And third party application performance monitoring will tell you when there is an application performance problem, but not what's causing it.

Difficulty Figuring Out Actionable Insight And Information

Those SDS systems that do provide performance information tend to provide either too much data detail or too little. Too much data detail makes it exceedingly laborious to decipher actionable information. Too little data makes it nearly impossible. The dubious underlying premise is that the user not the system must have significant expertise. That has been historically true but trends are showing decreasing market levels of storage expertise. User expertise is in operating not creating. Script writing is becoming less prevalent. Doing more with less for years upon years has generally forced IT into being more generalists than specialists. Storage administration is not commonly seen today as a "glamorous" job.

Reactive Capacity Consumption

As mentioned in the introduction, capacity management is always a storage issue and is frequently exacerbated by SDS implementations. Several SDS implementations can provide alerts when capacity utilization thresholds have been crossed, but there is no anticipation as to when it will happen. It is a binary alert. Nor do they measure the velocity of capacity consumption. This means that a high consumption velocity can overtake the entire assigned capacity pool by the time an alert has occurred causing catastrophic application failure. Something no IT pro ever wants to deal with.

The common work-around in preventing that kind of catastrophic application failure is to overprovision. Overprovisioning capacity makes sure that the application never runs out of capacity even though it's a very high-cost safeguard. Storage capacity can be orphaned. Reducing that cost requires effective thin provisioning. But even thin provisioning can get caught up in a high-velocity consumption situation.

These are the problems FalconStor Software set out to solve with their innovative SDS FreeStor®.

FreeStor And Its Intelligent Predictive Analytics

FreeStor is a pure SDS software system architected from the ground up to be SDS as a PSA or VSA. It's a highly efficient SDS based on more than 16 years of storage virtualization market experience. FreeStor is not re-purposed or re-designated software. FreeStor is purpose built to solve many storage problems aggravating users, Enterprises, Cloud Service Providers, and Hyperscalers. It neatly separates and abstracts the storage image from the hardware it manages while delivering the complete range of storage functionality and services required today:

- *Virtualized storage pooling*
- *Inline deduplication.*
- *Snapshot copy management services.*
- *WAN-optimized replication.*
- *Automated HA.*
- *Common data service across storage systems, storage media, storage infrastructures, platforms, types, and vendors.*
- *Intuitive storage implementation, provisioning, orchestration, capacity management, as well as all data services.*
- *Instinctive HTML-5 GUI and full RESTful API.*

What sets FreeStor apart from the market is that it solves those infuriating application storage performance and capacity consumption problems. FreeStor does this with its Intelligent Predictive Analytics.

Prognostic Application Performance Management With Simplified Troubleshooting

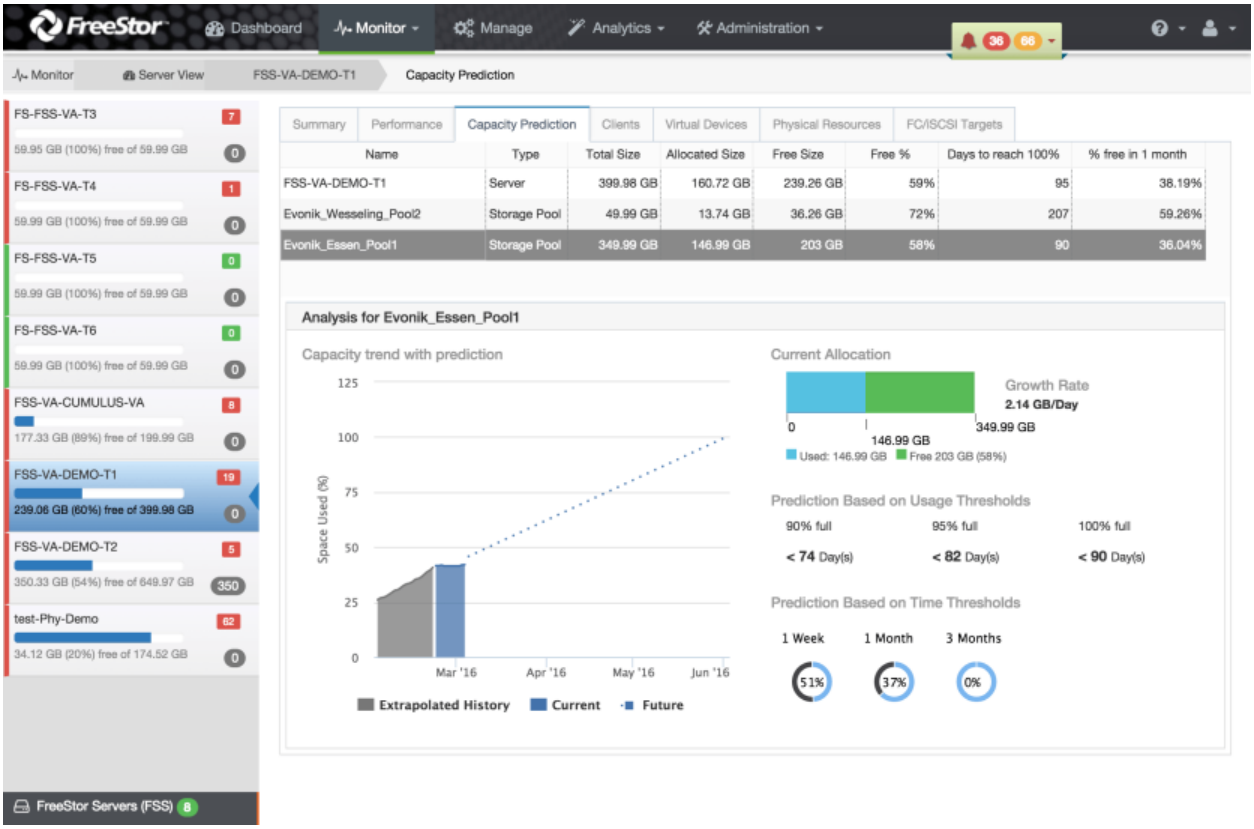
FreeStor Intelligent Predictive Analytics helps identify and pinpoint application storage performance bottlenecks and their storage causes. It does this by providing real-time performance, health and inventory information from the entire storage stack, including the physical drives (flash SSDs and HDDs), storage pools, virtual LUNs and connected the connected physical hosts. The analysis of this data in real-time is provided through the use of smart rules and alerts to automatically detect storage infrastructure changes in addition to changes affecting SLAs. General rules are used to provide alerts on configuration policies and data-to-day provisioning issues such as snapshot consumption is trending high.

These analytics convey uniform information for all of the underlying storage systems and/or host embedded storage even when they are different system types or vendors. FreeStor Intelligent Predictive Analytics comes with extensive reporting framework of more than 15 custom reports that can be filtered exported to PDF, email, or SMS. The monitoring, analytics, reporting, configuring, orchestrating, provisioning, and more scales to 100s of storage servers managed from a single pane of glass with the historical and real-time trends plus the ability to troubleshoot performance in the same GUI view. FreeStor's instinctive GUI enables the vast majority of storage tasks, provisioning, reporting, and troubleshooting to be completed with a few simple clicks measured in seconds.

Since no two storage administrators are alike, the dashboards are highly customizable so that storage administrators can individually configure them for personal preferences. There are also native IOS and Android storage monitoring apps available recognizing that storage administrators are unlikely to be chained to their desks.

Capacity Consumption, Capacity Consumption Projection, And Capacity Planning Management

In addition to the performance capabilities, FreeStor Intelligent Predictive Analytics can prognosticate storage pool and server capacity consumption based on current and historical trending. It captures storage pool and server storage capacity consumption in as it's occurring and the speed or velocity of that consumption. It then extrapolates that consumption on both a time-based view (what will the space consumption and availability be in 3 months) and percentage utilization view (when does it hit 90% capacity utilization). It also calculates the daily average consumption on various levels. See screen shot below.



What This Means...

No more sleepless nights with wake up calls at "0 dark thirty" because an application performance just is critically poor.

No more worrying if an application, especially a mission-critical application will run out of storage prematurely.

No more aggravation painstakingly attempting to deduce and isolate exactly what aspect of the storage is causing application performance to decline below its service level agreement, especially intermittent application performance problems.

No more wasted hours training on multiple storage systems and infrastructure.

No more over-buying and over-provisioning storage capacity.

No more manually entering storage performance or capacity consumption data into a database or a spreadsheet in an attempt to visualize actionable insight and information.

No more application storage performance and capacity management contortionism.

In a nutshell, FreeStor SDS with Intelligent Predictive Analytics delivers, easier, simpler storage, storage infrastructure, and brain dead simple implementations, provisioning, orchestration, management, troubleshooting, and operations.

Summary

SDS enables server-side storage (HDDs, SSDs, and storage systems) to be shared with multiple hosts and virtual machines. Application storage performance management is far too frequently a non-trivial operation. It doesn't have to be. FreeStor SDS with Intelligent Predictive Analytics makes that non-trivial…trivial.

For More Information

Contact FalconStor Software at: www.falsonstor.com or via email salesinfo@falconstor.com

Paper sponsored by FalconStor Software, Inc. About the author: *Marc Staimer, as President of the 18-year-old Dragon Slayer Consulting in Beaverton, OR, is well known for his in-depth and keen understanding of user problems, especially with storage, networking, applications, and virtualization. Marc has published thousands of technology articles and tips from the user perspective for internationally renowned online trades including SearchStorage.com, SearchCloudStorage.com, SearchSolidStateStorage.com, SearchSMBStorage.com, SearchVirtualStorage.com, SearchStorageChannel.com, SearchModernInfrastructure.com, SearchVMware.com, SearchDataBackup.com, SearchDisasterRecovery.com, SearchDataCenter.com, SearchServerVirtualization.com, SearchVirtualDesktop.com, SearchNetworking.com, and Network Computing. Marc has additionally delivered hundreds of white papers, webinars, and seminars to many well-known industry giants such as: Brocade, Cisco, DELL, EMC, Emulex (Avago), HDS, HP, LSI (Avago), Mellanox, NEC, NetApp, Oracle, QLogic, SanDisk; as well as smaller, less well-known vendors/startups including: Asigra, Caringo, Cleversafe, Cloudian, Clustrix, Condusiv, DH2i, Diablo, FalconStor, Gridstore, Nexenta, Neuxpower, NetEx, NoviFlow, Permabit, Qumulo, Scality, Tegile, and many more. His speaking engagements are always well attended, often standing room only, because of the pragmatic, immediately useful information provided.* *Marc can be reached at marcstaimer@me.com, (503)-579-3763, in Beaverton OR, 97007.*