OPTIMIZING VM STORAGE PERFORMANCE & CAPACITY
Tintri Customers Leverage New Predictive Analytics SaaS

OCTOBER 2016

THE NET NET Today we are seeing big impacts on storage from the huge increase in the scale of an organization's important data (e.g. Big Data, Internet Of Things) and the growing size of virtualization clusters (e.g. never-ending VM's, VDI, cloud-building). In addition, virtualization adoption tends to increase the generalization of IT admins. In particular, IT groups are focusing more on servicing users and applications and no longer want to be just managing infrastructure for infrastructure's sake. Everything that IT does is becoming interpreted, analyzed, and managed in application/business terms, including storage to optimize the return on their total IT investment. To move forward, an organization's storage infrastructure not only needs to grow internally smarter, it also needs to become both VM and application aware.

While server virtualization made a lot of things better for the over-taxed IT shop, delivering quality storage services in hypervisor infrastructures with traditional storage created difficult challenges. In response Tintri pioneered per-VM storage infrastructure. The Tintri VMstore has eliminated multiple points of storage friction and pain. In fact, it's now becoming a mandatory checkbox across the storage market for all arrays to claim some kind of VM-centricity. Unfortunately, traditional arrays are mainly focused on checking off rudimentary support for external hypervisor APIs that only serve to re-package the same old storage. The best fit to today's (and tomorrow's) virtual storage requirements will only come from fully engineered VM-centric storage and application-aware approaches as Tintri has done.

However, it's not enough to simply drop in storage that automatically drives best practice policies and handles today's needs. We all know change is constant, and key to preparing for both growth and change is having a detailed, properly focused view of today's large scale environments, along with smart planning tools that help IT both optimize current resources and make the best IT investment decisions going forward. To meet those larger needs, Tintri has rolled out a Tintri Analytics SaaS-based offering that applies big data analytical power to the large scale of their customer's VMstore VM-aware metrics.

In this report we will look briefly at Tintri's overall "per-VM" storage approach and then take a deeper look at their new Tintri Analytics offering. The new Tintri Analytics management service further optimizes their app-aware VM storage with advanced VM-centric performance and capacity management. With this new service, Tintri is helping their customers receive greater visibility, insight and analysis over large, cloud-scale virtual operations. We'll see how "big data" enhanced intelligence provides significant value and differentiation, and get a glimpse of the payback that a predictive approach provides both the virtual admin and application owners.

CRITICAL STORAGE CHALLENGES IN VIRTUALIZED ENVIRONMENTS

Optimal storage should aim to provide the lowest operating cost, the most reliable data protection and the best application IO and storage services. This simply isn't feasible using traditional storage in

today's virtualized environments for a number of reasons. Server virtualization made a lot of things better for the over-taxed IT shop and helped bring budgets back in control. However delivering quality storage services in a hypervisor infrastructure created severe challenges including:

- Increased complexity and cost in maintaining effective VM to storage mapping
- Poor performance (e.g. from the virtual IO blender effect when disparate VM IO is blindly aggregated inside the hypervisor)
- Difficult if not impossible per-VM storage insight for troubleshooting or planning
- Optimizing and balancing dynamic VM workloads and QoS needs with storage services

This last problem is key – without per-VM insight, IT will struggle with every storage task, but especially when it comes to providing business visible services including optimizing capacity utilization (e.g. cost), assuring availability and recovery (e.g. realistic DR), and guaranteeing application QoS (e.g. protection and performance).

VM-centric Storage Analytics

Effective storage management requires visibility from VM to storage. Tintri nailed that need fundamentally with the VMstore architecture. This revolutionary approach delivered storage resources as required per-VM – fully managed at the VM level without LUNs, volumes, striping, or other traditional storage array complexities. Tintri offers additional VM-centric storage capabilities like per-VM storage provisioning, VM storage rebalancing, per-VM snapshots and space-efficient clones, and the DevOps focused copy data management capability of SyncVM to manage per-VM data recovery and database synchronization. VMstore also fully supports both VMware VAAI and Microsoft Hyper-V ODX API's for full hypervisor service integration.

Now, Tintri has taken their per-VM management approach to a new level with a SaaS-based Analytics offering. The new Tintri Analytics augments existing VMstore storage management with a per-VM performance/capacity visualization and predictive analysis service.

Tintri Analytics collects and aggregates up to 3 years of detailed per-VM storage performance metrics (latency, throughput, etc. as forwarded by VMstores daily), and then dynamically serves browser-based visualization and analysis query and reporting. Tintri leverages solutions like Apache Spark, Elastic Search, etc. as a big data analytics platform that can support up to 160,000 VMs worth of time-series data.

Before considering the actual analytics, at a high level we'd like to point out some significant benefits that this "Management as a Service" (MaaS) approach delivers to the harried IT admin:

- The Tintri VMstore takes on all the service and data management tasks which frees up IT staff (and data center resources) to focus on business strategy, planning and execution related activities

- As a scalable cloud service, Tintri Analytics provides sub-second query responses to even the largest client queries across hundreds of thousands of VM's worth of data without having to spend any time maintaining on-premise storage management tools or databases

Ultimately, IT must start working at the application level to offer more relevant capabilities to their business clients. Tintri recognized that a per-VM analysis by itself isn't enough when working at large scale, or is always useful when working with end users or application owners. For example, many applications consist of multiple VMs working in concert – the business user may only perceive the health and performance of the whole set as the application. Therefore Tintri Analytics helps characterize groups of related VMs into larger workload constructs that can map to business application and/or organizational groupings for reporting, analysis and predictive planning. This

directly enables IT to better communicate and work with their business clients in more cloud-like service provider role.

Tintri Analytics SaaS/MaaS

At the first look at Tintri Analytics, we immediately note that they have distilled out and summarized key information on the front "page" default dashboard. There are the three key metrics that are intended to intelligently summarize the current storage situation across the environment:
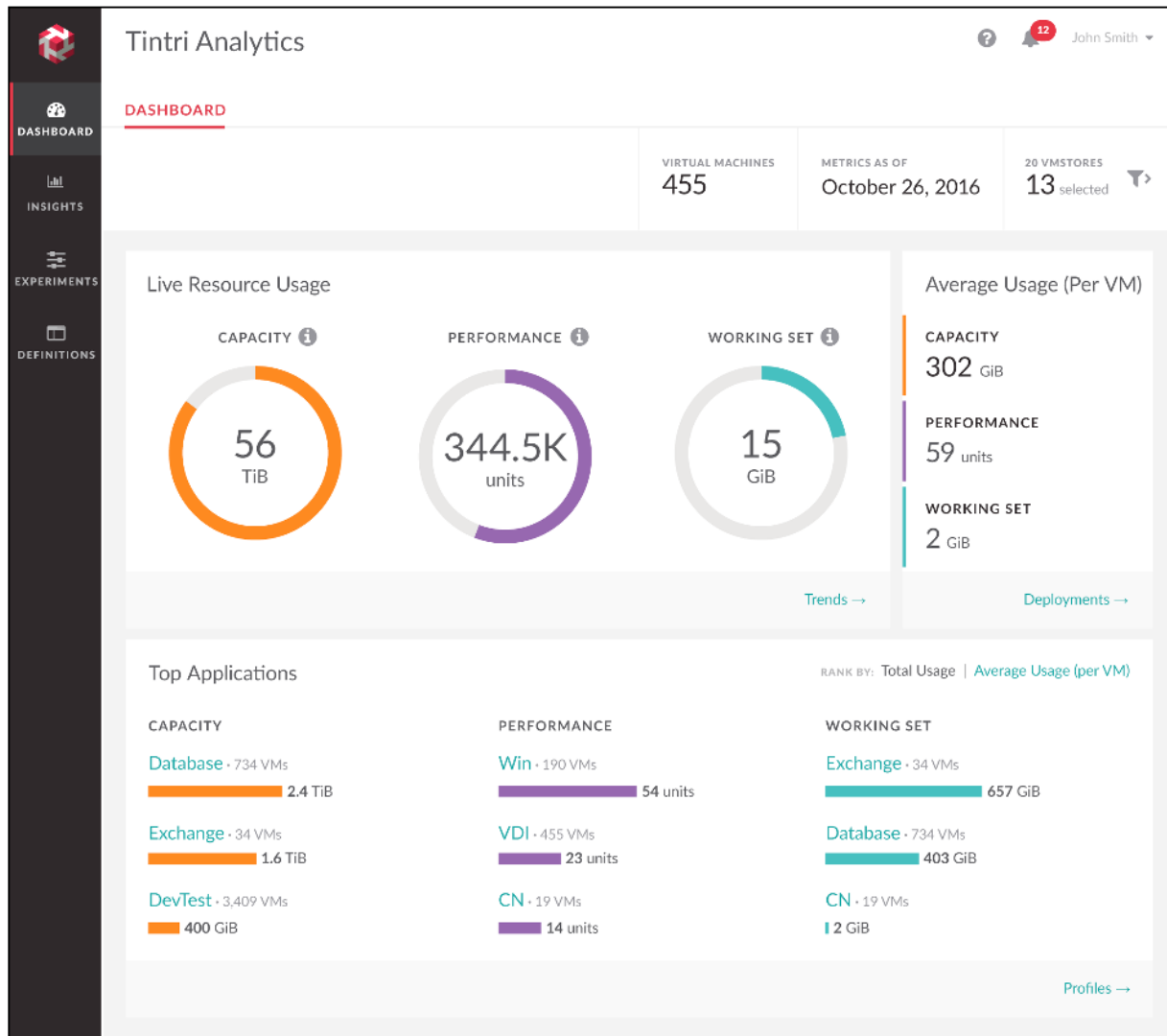

Figure 1 - Tintri Analytics dashboard

1. **Capacity** – Storage capacity is relatively straightforward, but we note with Tintri that since no storage is wasted, lost, isolated in LUN's or file store constructs, that the reported used and remaining capacities are truer and directly mapped to VMs. This enables direct and accurate "thinking and planning" tasks without the unreliable "fudge factors" required when managing traditional storage in virtual environments.

2. **Performance** – Understanding storage performance can be complex with traditional arrays, but here it is reported in terms of specially aligned "units" that intelligently aggregate lower-level IOPs and throughput metrics to provide a single VMstore key performance indicator.

This simple performance metric directly highlights the remaining performance capacity and identifies any impending "shortfalls" in performance requiring attention.

3. **Working Set** – Interestingly, we think this is a great calculated metric reflecting the effective active data set size, similar to a process working set size metric for memory usage. This data working set can be used as an indication of how much flash is needed in order to serve 99% of all read/write requests from flash.

There are also secondary metrics that show the current "live" and total VM allocations, the top "applications" (as grouped above), and a profile of the average VM requirements. Needless to say, the total information presented in just this first dashboard page used to take hours if not days of professional services to distill out of low-level storage management utilities.

As expected from a modern analysis solution, one can launch off into drill downs at any point to explore subsets of the environment as desired – by data center, cluster, network, application grouping, and of course per-VM.

In addition there are some interesting visual/analytical capabilities to specially note. When combined with the ability to drill down, these analytics can quickly solve or address almost any storage environment management need for supporting detail or planning recommendation. We'll next look closer at three of them:

- **Projections** – Looking forward in time
- **Histograms** – Understanding the behavior of groups of VMs
- **Experiments** – Evaluating "what-if?" growth and infrastructure change scenarios

PROJECTIONS – ORGANIC GROWTH

The time-series charting views can show up to three years of history, and can drill down into application groups of VMs. When a sufficient history of data has accumulated, automatically built seasonally adjusted projections up to six months forward can be overlaid to show the future expected utilizations assuming continuing organic growth. Intelligently and uniquely, these projections are cast over those key calculated metrics outlined above (capacity, performance, and working set) to directly identify future potential issues in those terms (and on a drilldown application or other grouping if desired).
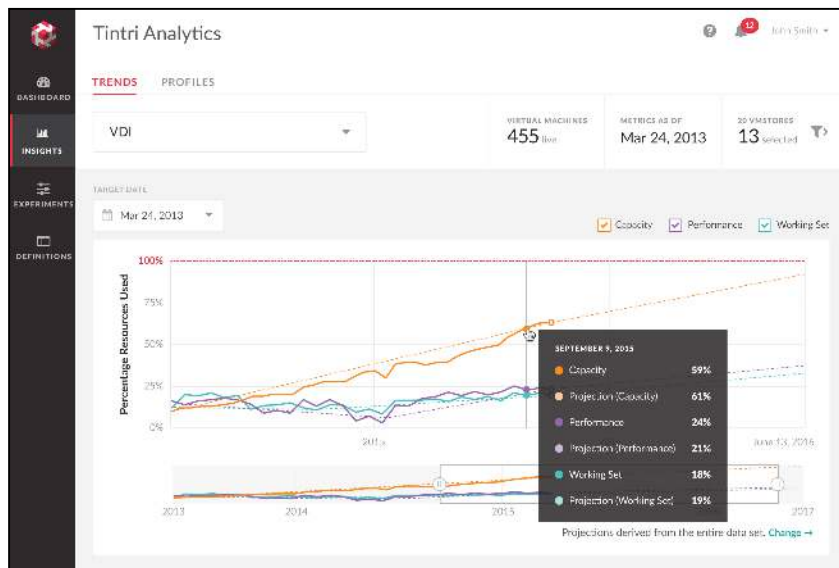

Figure 2 - Tintri Analytics Trends (with projections)

Simple exploration here will readily show which applications (or departments, etc.) are growing fastest in each important dimension, and if any are going to hit important limits anytime soon.

HISTOGRAMS - DISTRIBUTIONS AND PROFILES

Any performance consultant will tell you that its not enough to look at a time-series perspective, sometimes you want to examine some distribution of a data set. Here Tintri Analytics offers easy-to-use histogram-based views that assist with profiling tasks.

For example, if there are a large number of VM's in a cluster (or large application), the histogram view enables one to see the natural clustering of those VMs (e.g. by performance "units"). This perspective can help the analyst identify the "average" profile, or help group or isolate important or exceptional VMs for further analysis or in making planning decisions.

EXPERIMENTS – WHAT-IF?

Perhaps the most advanced capability is found in the experiments views. Here the analyst can run "what-if?" models to answer questions about the impact of future changes in usage or VMs deployed (and where), and then further evaluate potential upgrades to storage. A common modeling scenario might start by modeling "how many more VMs worth of this application will fit…".
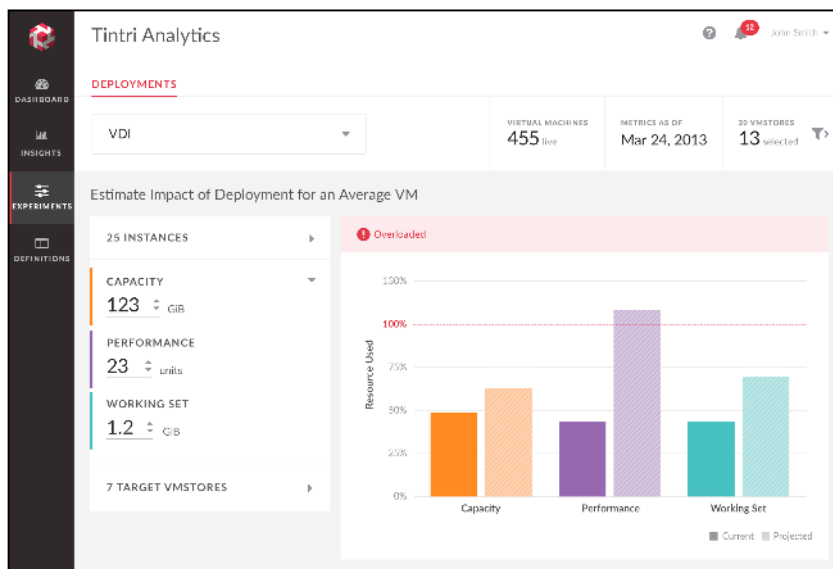

The underlying prediction engine accounts for and predicts along all of those same key dimensions introduced above (capacity, performance, working set), so the admin/analyst is always working in the same mindset looking at current status or making future plans.

Next we'll see how these fundamental analytical capabilities might be used in common virtual storage management activities.

Figure 3 - Tintri Analytics - Experiments

Using Tintri Analytics

To be honest, much of the Tintri solution obviates the need to spend as much admin time dealing with storage management issues. Tintri's per-VM QoS and the dynamic Tintri Scale-out balancing take out a lot of traditional storage administration tasks. Still, in the virtualized data center, some common storage performance tasks best accomplished with detailed insight might include the following:

1. **Performance Troubleshooting** – With Tintri's per-VM storage each VM's usage is isolated from another (and throttled when necessary) and can be balanced dynamically across a cluster of VMstores (with Tintri Scale-out), but there still can be rogue VM's and unexpected application usage that need to be identified and dealt with.

    The Analytics dashboard will first show if the problem is an environment-wide shortfall, or instead may be more isolated in nature. One can readily drill down into subsets of the environment to "box" the issue. Then projections and profiles at finer grained subsets of the environment can be immediately explored to compare expected application usage with current usage patterns, VMs within an app with each other, or otherwise identify unexpected utilization patterns.

2. **Performance Management** – Managing performance starts with understanding and profiling workloads (aka workload characterization). Tintri Analytics directly supports this activity with natural histogram views on the key metrics that enable quick profiling. Identified applications or other relevant subsets of VMs can be readily grouped for future views and reports, including for the purpose of keeping tabs on organic growth projections by application.

   We should mention that working and reporting at the application level with smart metrics helps IT work more closely with business-side clients. By working in terms and in units that better reflect their IT client perspective, IT can provide better for business enablement.

3. **Capacity Planning** – Certainly the automatic projections of organic growth provide great input to the capacity planning process, but it's the experiments capability that really shines here. Because predictions are based on a model that accounts for capacity, performance and working set size at the same time, the capacity planner doesn't have to separately worry about each dimension.

   While this model doesn't enable playing with every conceivable "knob" that a virtual admin can possibly turn (e.g. all the VM level settings in the hypervisor), here predictions can be used to evaluate scenarios concerning how much storage will be required to support a growing or changing application. In addition, scenarios can evaluate the profile/size of VMs as deployed hypothetically to targeted VMstores.

4. **Optimizing Operations** – When it comes to actually adding a VM to the infrastructure, or adding a new VMstore to the storage pool, Tintri's existing management automatically optimizes placement and overall re-balancing. After deployment, Tintri Analytics can uniquely provide the application level insight and validation that application-level operations are proceeding as planned.

TANEJA GROUP OPINION

It's rare to find storage management solutions that offer any kind of actionable predictive capabilities, much less application-level insight. Tintri is definitely setting a new standard for the wider storage industry to catch up to by including these capabilities in their increasingly intelligent solution. Tintri is closing the feedback loops that eventually will lead to fully self-operating and self-optimizing application storage services.

We think Management as a Service (MaaS) is not simply a hot trend, but rather the inevitable future for systems management overall. In general, call home support is fast evolving into more real-time, automated, intelligent, big data driven and optimized operational services. Tintri Analytics is no exception here, but rather an early leader in this space, extending Tintri's call home support services to provide differentiating and directly accessible customer analysis and insight capabilities. We have no doubt that as the intelligence inside grows, more operational benefits will emerge – and there is plenty of room to learn from the growing body of big data that Tintri is accumulating from across its user base.

We also want to note that customers adopting Tintri feed into the trend of reducing silo IT operations in favor of more virtual, generalized IT. We think this trend also lets IT folks move and work closer to the application owners and users – usually over on the business side of the fence. It's clear that solutions that help IT work at the application level are going to have the biggest impact (and future role). While "per-VM" is definitely the right and necessary approach in the virtual environment, the application awareness that Tintri Analytics now brings to storage takes their solution to a whole new level of value. As IT in general adopts a service provider mentality, management solutions with application awareness will help the IT "service provider" do more for their business clients.

Which leads to perhaps the best part of this new offering – the cost. Tintri is currently making it available "free" to all of Tintri clients with active support agreements. Tintri seems to love its customers!

NOTICE: The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors that may appear in this document.