ActualTech Media

Solving the Top 5 Virtualized Application and Infrastructure Problems

By David Davis, vExpert

Co-Founder, ActualTech Media

May, 2015

uila

Table of Contents

ActualTech Media

Introduction

Today's companies no longer hold back when it comes consolidating servers using server virtualization. The servers (and their applications) dubbed "the low hanging fruit" of virtualization have long since been virtualized. Now the only fruit that's left are the business-critical applications and they're being virtualized at a faster rate than ever before. They aren't being virtualized because it's easy to do so, but rather because the underlying infrastructure hardware and software are mature and robust.

Once virtualized those business-critical applications often initially seem to work with the already-virtualized server, storage, and hypervisor. However, after time, the great performance that you and your end users have come to expect from those business-critical applications, day after day, stops. Why? Is it an application or infrastructure problem? What's the root cause of the problem? How do you solve it? Where do you even go to find these answers?

Inevitably this is exactly where too many companies find themselves today — lost and with more questions than answers.

To learn how to solve the top 5 virtualized application and infrastructure problems, read on.

Key Factors

From talking with hundreds of companies, it's clear that the most common virtualized application and infrastructure problems all relate back to the following key factors:

- **Root cause** —Elusive to most data center applications, data center admins struggle to identify the root cause of a problem fast enough to prevent the impact of negative application performance on their users
- **Complete visibility** —It's rare to find an organization that doesn't use a half-dozen or more applications to manage, monitor, and troubleshoot their data center. These fragmented tools lead finger-pointing, missed opportunities to prevent downtime, and delays in problem resolution.
- **Over and under provisioning** —We all want to maximize the return on infrastructure but too many companies seesaw between over and under provisioning their infrastructure, thus causing either wasted resources or, the much more damaging, performance / capacity bottlenecks across the 4 critical data center resources: compute, memory, storage, and network.

- **Lack of application-oriented metrics** —Management and monitoring tools provide hundreds of metrics but rarely do they relate to anything meaningful to end users or application owners. The real measure of the infrastructure is the response time of the business-critical applications as viewed by the end users who use them daily.
- **Infrastructure and application silos** —At most companies, every infrastructure stack and application group are broken down into management silos. This siloed management approach causes communication barriers, troubleshooting delays, and finger-pointing, further adding to the stress and downtime during problem scenarios. This exasperates the issue of each management silo having its own management tool and limited insight into other management silos.

With this high-level view of the key factors effecting the application and infrastructure, now let's learn what we have determined are the top 5 application and infrastructure problems and how to solve them.

#1 Isolating Root Cause in Single and Multi-Tier Applications

Virtual infrastructures are comprised of many different pieces of hardware and software making root cause identification challenging for even simple applications. However, since companies today need to virtualizing complex, multi-tiered, business-critical applications, the challenge of root cause identification had become greater than ever.

Let's look at an example of a 4-tier business application made up of a load balancer, front end server, business application server, and database each running on separate virtual machines. The virtual machines are all reliant on the underlying virtual infrastructure that provides the CPU, memory, storage I/O, and network I/O resources. The first challenge for any performance and capacity monitoring tool is gain visibility into all the resources that could cause trouble for an application. Typically those resources are made up of:

- CPU
- Memory
- Storage (I/O or capacity)
- Network I/O
- Application services

Many tools will try to do this by installing guest OS agents across the infrastructure, but this is inefficient and just increases the management burden. Instead, you should look for tools that can gain the level of insight needed without using guest OS agents.

Even with insight into these resources, rarely can management tools automatically learn the dependencies required for an application to work. Examples of dependencies might be DNS, a database server, file share access, and more. Also, just because an application has all the infrastructure resources available that it needs, an application can't work without its application dependencies.

Finally, even with its application dependencies in place, how are those dependencies performing? Are DNS lookups experiencing high latency and that's what is causing a financial application to run slow?

The **solution** to this challenge is to select a management tool that has both:

1) Complete visibility into all the infrastructure resources
2) Complete visibility into application dependencies
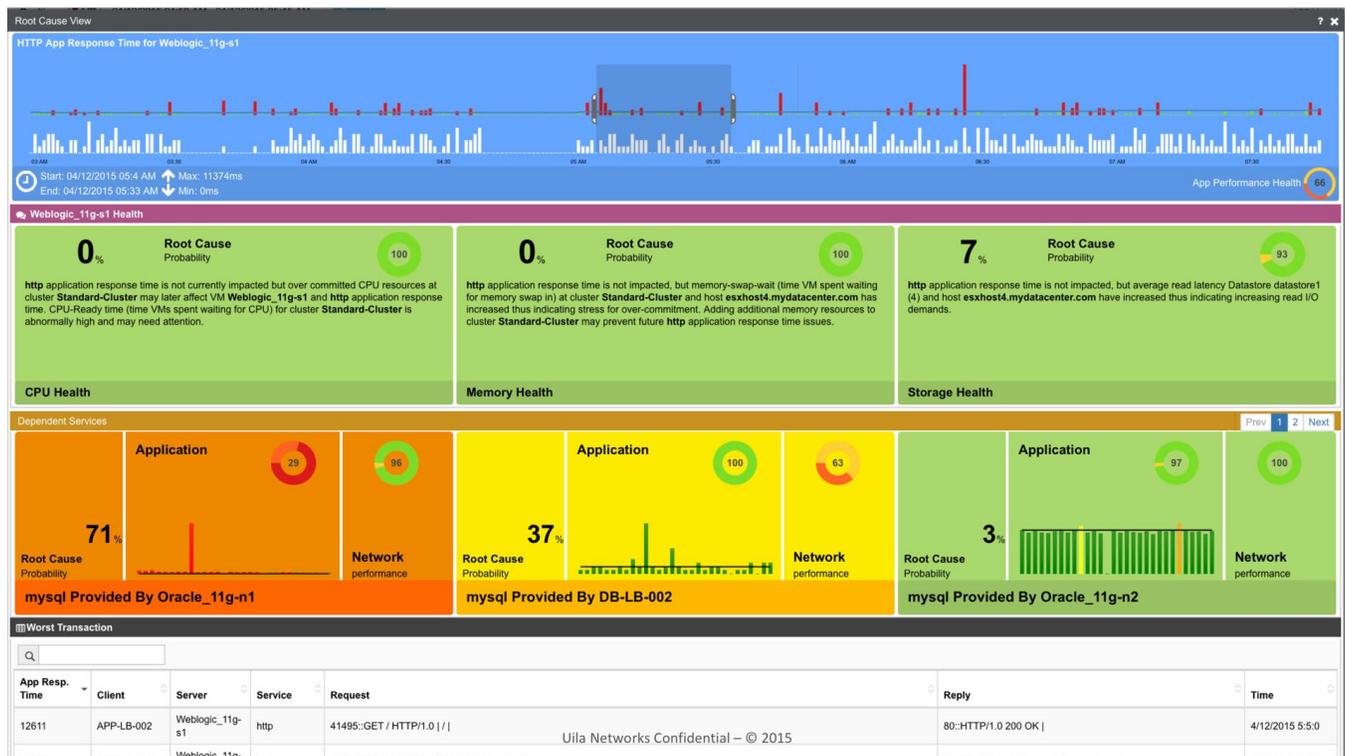
Figure 1

There is a new company in the world of virtualization monitoring called Uila. Their new virtualization monitoring tool unique in many ways. Throughout this tech brief, we'll highlight how Uila is able to help solve these common virtualized applications and infrastructure problems.

When it comes to isolating root cause in single and multi-tiered virtualization applications, Uila – creator of a leading application and infrastructure performance monitoring tool – is able to correlate application response time to infrastructure health across all the tiers of the multi-tiered application, as shown in Figure 1. At the same time, it's able to provide real-time application transaction latency so you can understand the end user experience.

#2 Preventing Network Latency for Critical Applications

Many virtual infrastructure management tools have insight into CPU, memory, and storage, but it's the network by far that is the most overlooked resource critical to application performance. After all, without reliable high-speed networking, applications can't access their dependencies or that access is slow (causing the application to be slow).

Most companies either decide that slow application performance must be the fault of the network (the "guilty until proven otherwise" theory) OR they spend a tremendous amount of time and resources troubleshooting a slow application only to find out later that it was a networking issue. Virtualization management and monitoring tools should provide network visibility so that they can monitor network utilization along with all the other infrastructure resources when performing root cause analysis.

Of all the network statistics that one could monitor, network latency is the most critical. Network latency is the time that a packet takes to traverse a network connection. It's high network latency that can quickly kill application performance (especially for multi-tiered applications).
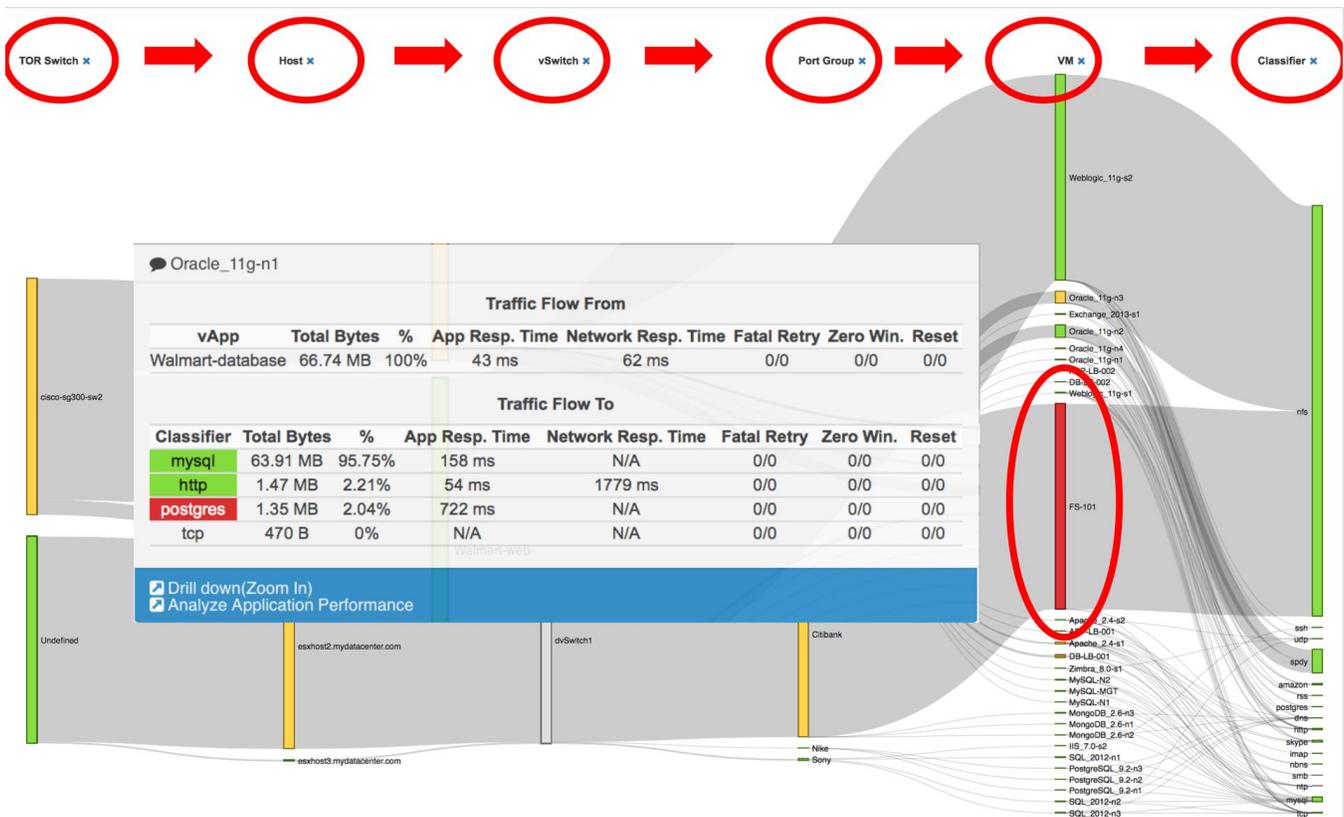
Figure 2

In order to monitor and prevent network latency, the **solution** is to select a management tool that provides comprehensive monitoring of application latency on the network. For example, Figure 2 shows how Uila monitors the network latency of a multi-tiered application and is able to quickly identify which of the tiers is experiencing high network latency, resulting in application performance degradation for end users.

#3 Identifying vCPU Over and Under-Provisioning

We all know that a virtual machine that doesn't have enough vCPU to provide resources for CPU hungry applications will result in poor application performance. Certainly you need a monitoring tool to identify that shortage. However, significant over-provisioning of vCPU across many virtual machines can also result in a performance penalty.

No matter the resource (vCPU, virtual memory, storage I/O, or network I/O), ensure that the virtualization monitoring tool that you select can recognize under provisioning as well as over provisioning.
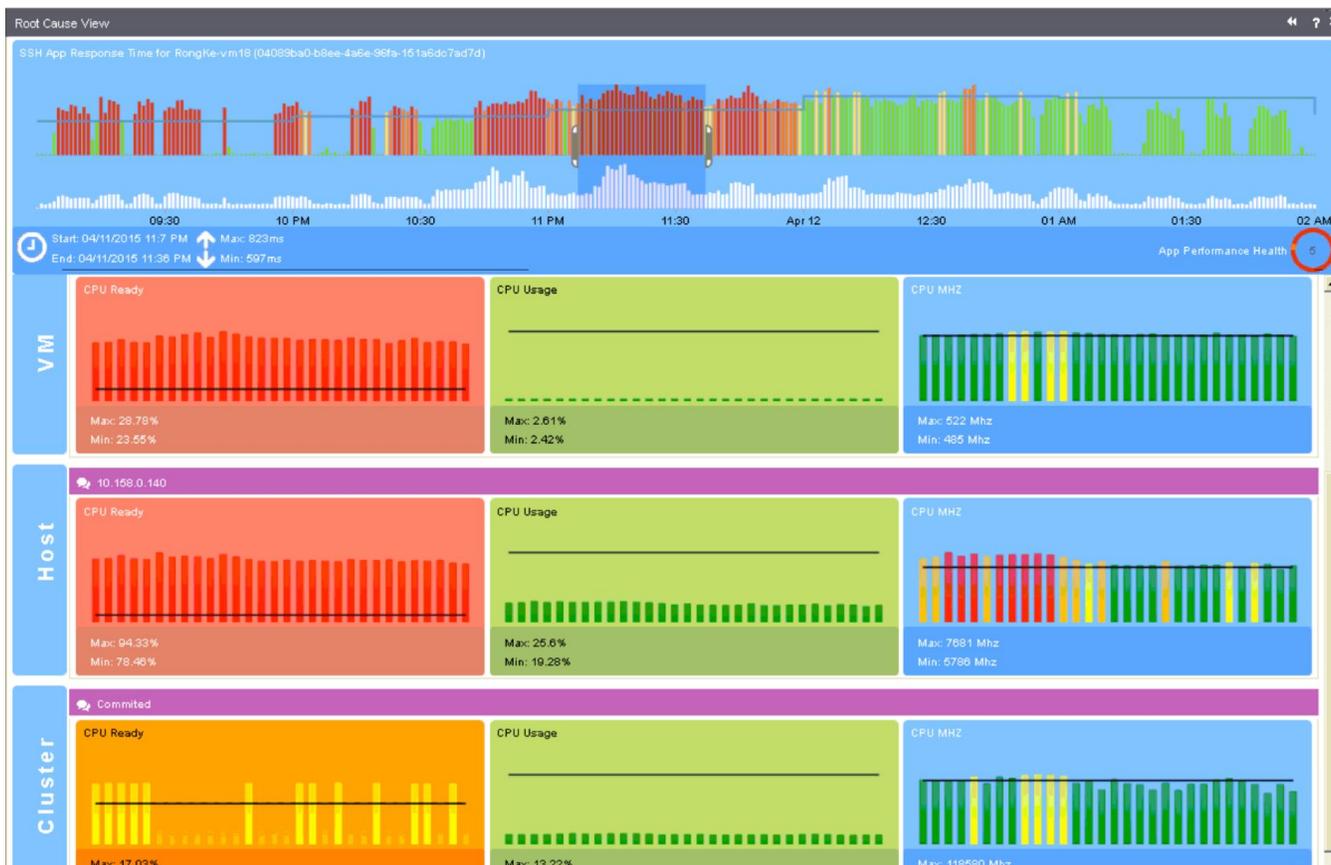
Figure 3

As you can see in Figure 3, Uila is able to identify over- and under-vCPU provisioning across clusters, hosts, and virtual machines. Not only will under and over provisioned VMs be identified, but the application performance impact will be identified all the way down to the specific application and time period in which the end user applications were negatively effected.

#4 Troubleshooting Storage I/O Contention and I/O Latency for Tier-1 Applications

As the utilization of the virtual machines in the virtual infrastructure increases and as the number of virtual machines virtualized increases, so does storage I/O utilization. In fact, VMware did an analysis of all the calls that they received in their support group and found that the vast majority of all problems could be tied back to storage issues. It's common for companies to be caught off guard by storage I/O issues for two different reasons:

1) When new applications are virtualized it can be difficult to size their storage I/O requirements. Enterprise applications tend to use more storage I/O than expected, and the result can be an unexpected storage I/O bottleneck.

2) Storage I/O is very "bursty" and is very dependent on the nature of the application. The result is that storage I/O contention issues can come and go quickly, making them difficult to identify. In many cases, storage I/O contention can happen so quickly that common monitoring tools don't notice it or label it as a bottleneck; however, end users will likely still notice when they experience delayed application response.

Like the network latency issue we discussed previously, storage latency in multi-tiered business-critical enterprise applications can be devastating when it happens at just one of the tiers and requires extensive troubleshooting to identify.

Since storage I/O throughput is a very expensive resource (and today's price of flash storage makes it cost prohibitive for all applications), most organizations can't over-provisioning storage I/O much more than they need to keep their applications running with acceptable performance. So what's the solution?

In order to be able to provision just the right amount of storage I/O for the various tiers of a multi-tiered application, the **solution** is to know the application dependencies and know the storage I/O average utilization and maximum utilization between all the tiers. Only with that information can you properly provision (but not over provision) the storage I/O between the multiple tiers.

Figure 4 shows how Uila is able to report both the application dependencies as well as the storage I/O utilization between each.
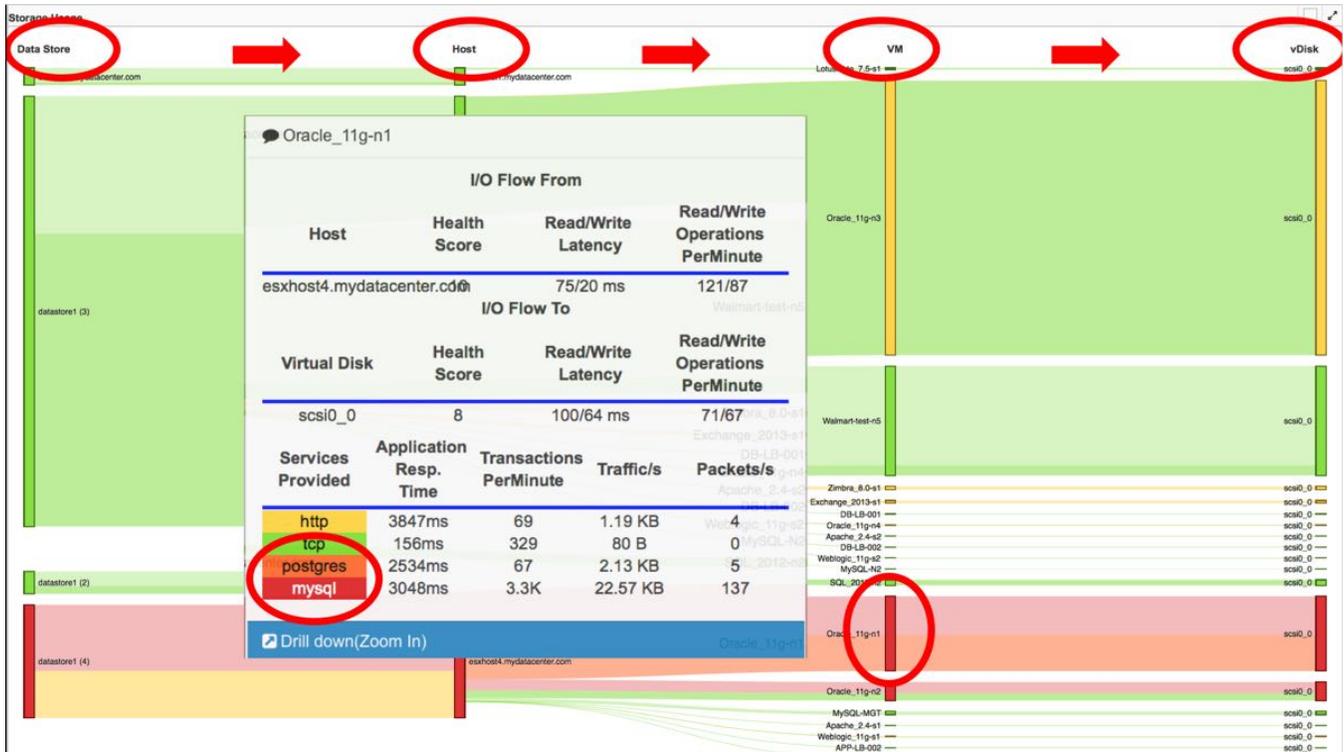
Figure 4

Knowledge is power and with this kind of knowledge, you (or your tool) now have the power to identify storage I/O bottlenecks between any of the tiers of a tier 1 application. This allows you to prevent finger-pointing between the different infrastructure silos and prevent or minimize downtime for the users of the application (shown in Figure 5).
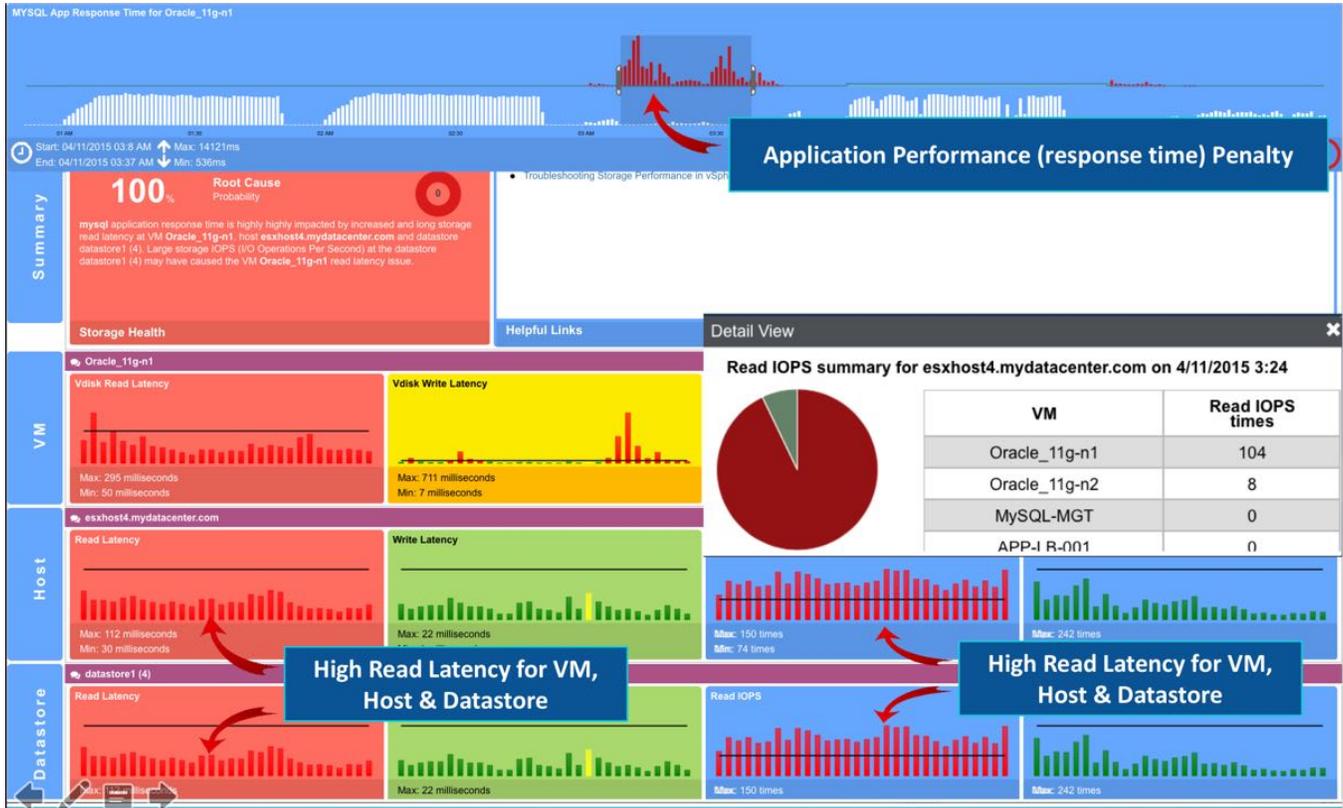
Figure 5

#5 Measuring and Preventing Virtual Application Downtime

When a problem does occur, finding the root cause and being able to solve it quickly is vital. Even better than that would be to be able to monitor capacity over time and prevent problems across the entire infrastructure and all the applications — before they happen.

The ultimate **solution** for measuring and preventing application and infrastructure downtime, for complex multi-tiered business-critical applications is to have a comprehensive solution for infrastructure AND application monitoring. As you can see in Figure 6, Uila provides such a solution by showing network, CPU, storage, and memory health, across all hosts and clusters and predicts future application disruption.
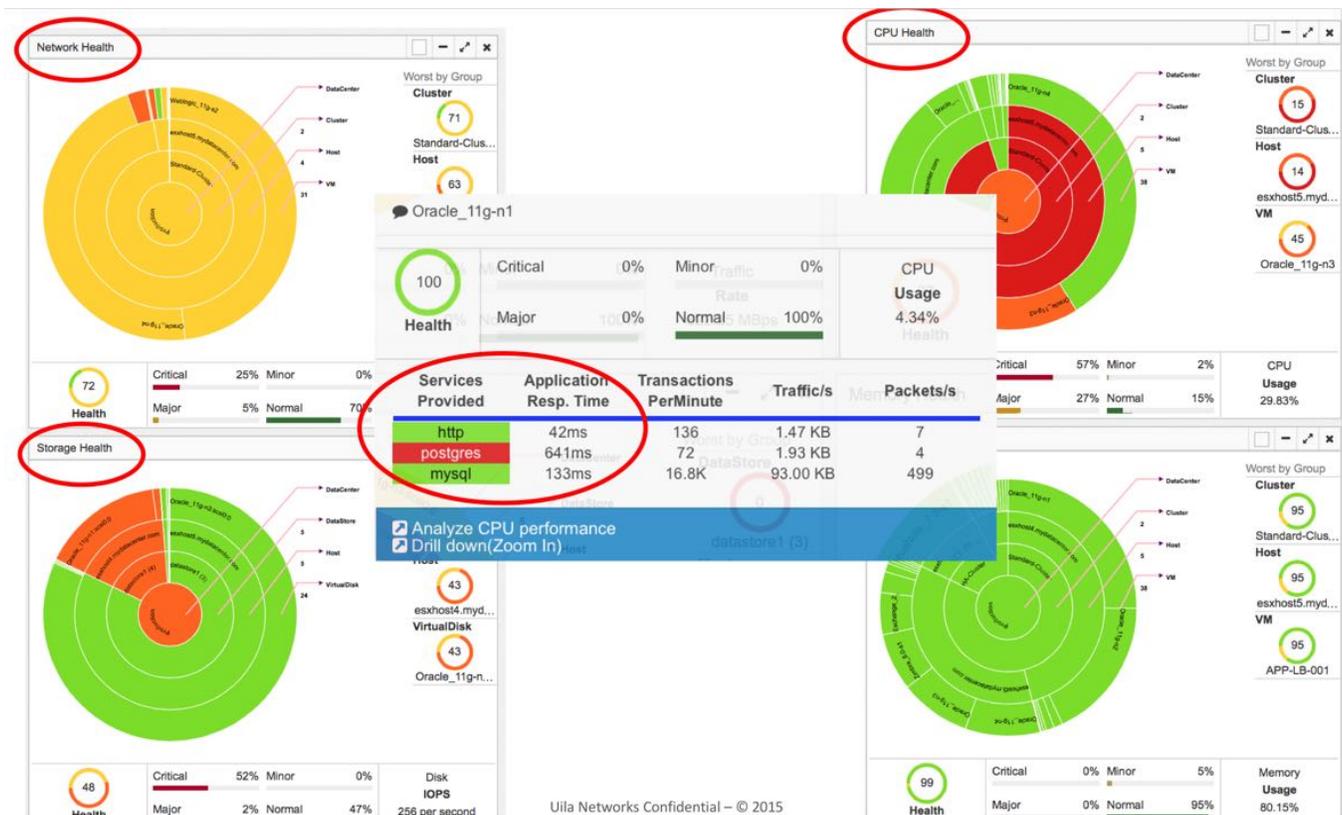
Figure 6

Additionally, Uila is able to trend application response time (Figure 7) over time, and correlate that with infrastructure resource utilization to quickly identify what caused application performance degradation in the past so that it can be prevented in the future.
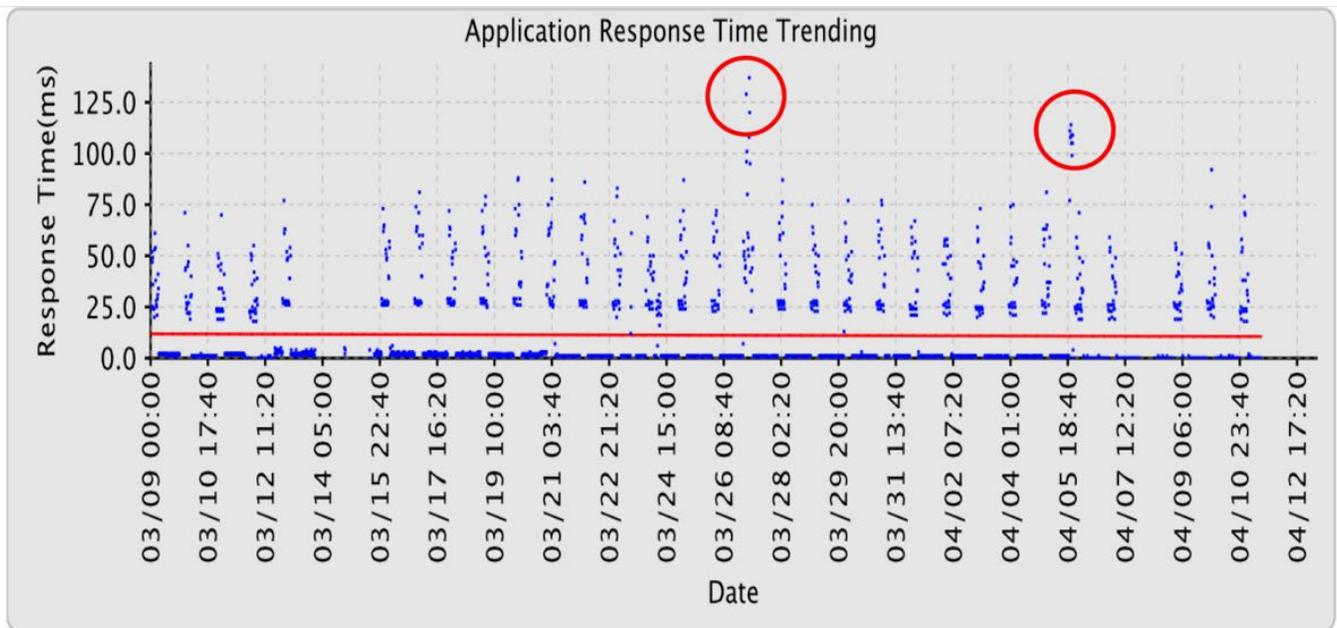
Figure 7

Selecting the Best Virtualization and Application Monitoring Tool

You started the virtual infrastructure to maximize return on investment and to ease administrative burden. In order for any virtual infrastructure to be successful by pushing capacity utilization, you must be able to predict and identify capacity bottlenecks to maintain optimal end user application performance. Thus, it's inevitable that any company running a virtual infrastructure will need a performance and capacity monitoring tool.

When you go to select your first tool or your next tool, look for the latest features and capabilities which include:

- Performance / capacity problem prediction (and super-fast troubleshooting)
- Root cause problem identification
- Complete visibility across the infrastructure (including the virtual network) and applications
- Application-awareness and application performance monitoring over time, all without having to install agents in each virtual machine

To register for a free trial of Uila's virtualization performance and capacity monitoring tool (which provides all of these capabilities), visit

www.Uila.com

About the Author

David Davis is a well-known virtualization and cloud computing expert, author, speaker, and evangelist. David's library of popular virtualization video training courses can be found at Pluralsight.com. He holds several certifications including VCP5, VCAP, CCIE, and has been awarded the VMware vExpert award 6 years running. His website, covering the latest in virtualization is VirtualizationSoftware.com.

About Uila

Uila offers a fresh approach to virtual infrastructure management by integrating the monitoring of both IT infrastructure and applications. Their innovative product is a game-changer in the data center and allows companies to measure application performance over time (without agents), identify root cause of problems, remediate efficiency, prevent future performance issues, and reduce IT total cost of ownership. Find out more about Uila at www.Uila.com.