Planning for Performance Assurance

No Longer an Afterthought

By Jim Metzler
Jim@Kubernan.Com

Introduction

A recent Kubernan Brief [1] entitled "The Performance Management Mandate" pointed out that unlike the situation following the dot-com implosion when IT was regarded primarily as a cost to be minimized, the IT function is now regarded as a key player in the ongoing roll-out of important business initiatives. In a growing number of instances, performance management is a critical component of IT's ability to support these initiatives.

One class of initiatives that is impacted by the IT organization's performance management capabilities is data center initiatives. Market research conducted by Kubernan indicates that half of IT organizations either recently have, or soon will, consolidate servers out of branch offices and into centralized data centers. In addition, one third of IT organizations either already have, or soon will, reduce the number of data centers that they support on a worldwide basis. These consolidation initiatives have multiple business benefits. These initiatives do, however, increase the distance between the end user and the applications that they need to access. With this added distance comes additional latency, jitter and packet loss, which drives the need for IT organizations to also implement performance management upgrades.

Another class of initiatives that is impacted by the IT organization's performance management capabilities is the deployment of Unified Communications (UC) functionality. Market research recently conducted by Kubernan[2] indicates that companies perceive a number of benefits from UC, most notably higher employee productivity, which was the top driver for 75% of the respondents to a recent survey. These survey respondents quantified the benefits from that productivity improvement and it included 29% higher customer satisfaction, 27% higher employee satisfaction, and 18% more sales. Given these business benefits, companies are beginning to make significant deployments of UC. IT organizations, however, need to ensure that they have implemented the requisite performance management functionality that will ensure that delay-sensitive UC applications exhibit acceptable performance so that the business benefits are actually achieved.

"The Performance Management Mandate" also highlighted the fact that just as IT is changing to become more business relevant, network management also needs to change and be able to demonstrate more business value. To put this into context, most business managers and CXOs don't care very much about network connectivity. They consider this to be a utility that IT is expected to deliver as easily and seamlessly as electric power. They do, however, care about the applications that support the company's key business processes. In response, over the last few years many network managers' jobs have expanded to include ensuring adequate application performance levels.

Unfortunately many IT organizations have not developed effective performance management capabilities and often regard it as something that can be put off until an IT initiative is either well into the implementation stage or is already in place. In many cases, the lack of effective performance management functionality limits the ability of

[1] The Performance Management Mandate, http://www.webtorials.com/abstracts/KubernanBrief-1-1.htm
[2] 2007-2008 Mobile Unified Communications Market Report, http://www.webtorials.com/main/resource/papers/kubernan/sotm07-4/paper/2007-Mobile-UC.pdf

the IT organization to be successful with the implementation, fine-tuning and ongoing management of these initiatives.

The goal of this brief is to suggest ways that IT organizations can successfully evolve their approach to managing the IT infrastructure to include a focus on performance. To achieve that goal, the brief will focus on processes as well as some key technologies that IT organizations can use to gather the management data that IT organizations must have in order to prepare for a key business initiative, and to support the rollout, and management of the new solution once it's up and running.

The Importance of Performance Assurance

As pointed out in "The Performance Management Mandate", in the majority of instances in which the performance of an application is degrading, the degradation is noticed first by the end user and not by the IT organization. I recently asked a couple of CIOs about the impact of this fact. One CIO stated that it causes the entire company to question the competence of IT. Another CIO was more direct and stated that it caused the IT organization "to look like a bunch of bumbling idiots." In order to be regarded as successful, IT organizations need to eliminate the image of being "bumbling idiots".

A few years ago the industry began to evolve from where network management focused almost exclusively on the availability of network devices to where it's current focus often includes the performance of networks and applications. Quality of Experience (QoE), which is what end-users notice and judge IT's performance by, is not just based on device availability. In fact it's a common experience for network managers to have end-users complaining about poor application performance while the NOC lights—based on SNMP (Simple Network Management Protocol) information about device status—are all green.

Effective performance assurance requires application-level awareness. Dynamic dependency mapping between resources and application is critical. A proactive monitoring approach that reduces the risk of QoE degradation is also essential. IT organizations must be able to correlate infrastructure performance, application response time metrics and end-to-end network performance in order to understand the end-user perspective. The key is to leverage any and all embedded instrumentation, capture rich performance metric data, consolidate, correlate and analyze this data to provide a robust view of end user experience. This rich view of the end user experience allows IT organization to meet service level requirements whether these requirements are explicitly described in a formal document, or are implicit in the expectation of the company's senior management.

The Performance Assurance Process

Performance assurance can be modeled as a four-part process as illustrated in Figure 1. These primary components of performance assurance are:

Baselining: Baselining provides a reference point from which service quality and application delivery effectiveness can be measured. It does this by quantifying the key characteristics (e.g., response time, utilization, delay) of applications and various IT resources including servers, WAN links and routers. Baselining allows an IT organization to understand the normal behavior of those applications and other IT resources.

Deployment Planning and Management: All IT organizations strive to avoid any disruption of business operations caused by the deployment of a new initiative. In order to minimize disruption, it's important to be able to use baseline data to accurately estimate the impact of the rollout on other systems and applications and to eliminate any potential issues in advance. It is also important to have the capability to be able to quickly identify and resolve to any issues that do arise.

Fine-Tuning: Few IT initiatives experience a perfect implementation. In particular, the first iteration of a complex initiative typically will not perform in an optimal manner, no matter how carefully the IT organization baselined and planned the rollout. There are almost always improvements that can be made that will deliver increased user satisfaction or productivity, and improve the ROI of the project.

Ongoing Management: Once everything is up and running and the fine tuning has been completed, the IT organization has to keep the network and the applications running, and, as much as possible, detect and eliminate problems before they impact end-users. The process does not end here as ongoing management leads back to baselining that enables the IT organization to successfully support the current functionality as well as the next initiative.

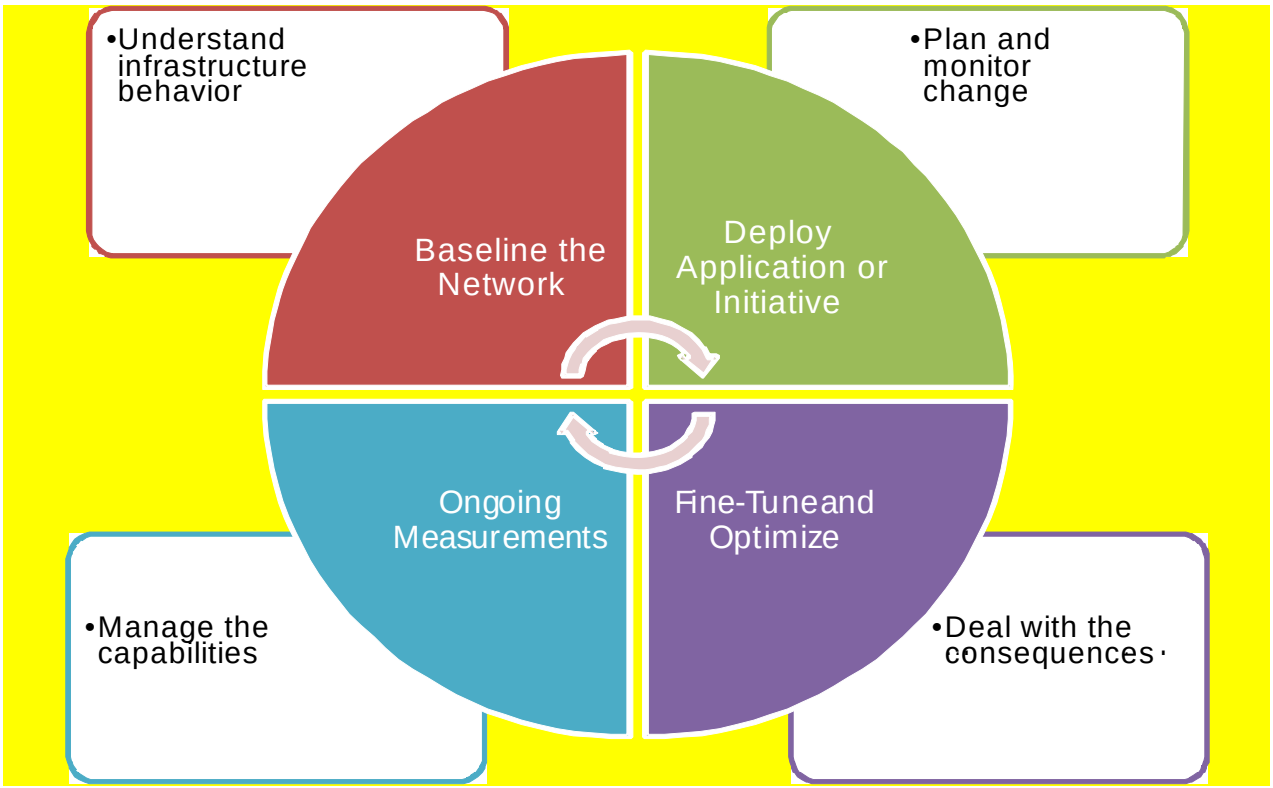

Figure 1: Performance Assurance

Sources of Management Data

IT organizations can typically rely on having access to management data from SNMP MIBs (Management Information Bases) on network devices such as switches and routers. Indeed, this is the foundation of today's device-centric approach to management that was mentioned in the introduction. SNMP from network devices provides data link layer visibility across the entire enterprise network and captures parameters such as the number of packets sent and received, the number of packets that are discarded, and the overall link utilization.

However, SNMP data tells the IT organization nothing about the applications involved, the servers the data is coming from, or the user to which the data is being delivered - all of which is important information for understanding the impact of network misbehavior. In addition, SNMP data does not provide any insight into such things as class of service, which is important for QoS management.

To get more granular information, many organizations turn to NetFlow, a Cisco IOS software feature and also the name of a Cisco protocol. Within NetFlow, a network flow is defined as a unidirectional sequence of packets between a given source and destination. For example, the branch office router outputs a flow record after it determines that the flow is finished. This record contains information that gives the IT organization answers to questions such as:

- Source and destination addresses: Where does the traffic originate? Who's affected?
- Ports: What application is involved? Is it one of the critical applications that business managers care about?
- Class of service: What's the desired priority of the traffic?
- Packets and bytes: How much traffic has been transmitted?
- Flow timestamps: How many packets or bytes per second?

Because it can provide answers to these questions, NetFlow represents a more advanced source of management data than SNMP MIBs. A new Internet standard, IP Flow Information Export (IPFIX), is being developed based on NetFlow Version 9. The IETF is in the final stages [3] of approving IPFIX (RFC 3917). IPFIX specifies how to log IP packets as they flow through a router, switch or other networking device and report that information to network management and accounting systems.

Although NetFlows supplies a great deal of information about application usage, it lacks data about application performance. For this, even more granular data, such as is supplied by Cisco's IP Service Level Agreements (IP SLAs), is required. IP SLA, a feature of IOS, is an active traffic monitoring capability, based on synthetic traffic, that collects real-time information about response time, one-way latency, jitter, packet loss, voice quality, and other network statistics.

One common use of IP SLA is to measure performance by sending one or more packets to a Cisco router, using the timestamp information on the packet to calculate actual performance statistics. These measurements can be one-way, or, if the destination router is configured as an IP SLA responder, two-way. IP SLA operation can be scheduled for a particular time, or operated continuously over a time interval. By setting the Diff-Serv Code Point (DSCP) bits, IP SLA can also monitor per class-traffic as an aid to QoS management. Devices configured for IP SLA operation can trigger SNMP alerts if measurements exceed or fall below a threshold. Thus, IP SLA and similar technologies provide key management data that can be used for ongoing performance monitoring, troubleshooting, and baselining.

NetFlow and IP SLA are both significant components of an overall network management strategy. However, there is a new approach to network management that is beginning to provide significant value. This approach is based on not only embedding instrumentation, but is also based on embedding analysis. Examples of this approach include Cisco's Network Analysis Modules (NAM) as well as Cisco's Bandwidth Quality Manager (BQM). The NAM Traffic Analyzer Software is an embedded, Web-enabled management station that monitors, analyzes and troubleshoots traffic. BQM offers end-to-end network service quality monitoring and provides visibility and analysis of traffic, bandwidth and service quality on IP access networks. A future brief will discuss NAM and BQM in detail.

Summary and Call to Action

As recently as a few years ago, there was little discussion in the industry about performance assurance. Now driven in part by the need to support initiatives such as data center consolidation and UC deployment, performance assurance is a top of mind issue for the vast majority of IT organizations.

In order to implement effective performance assurance, IT organizations must implement processes that correspond to the four stages that are depicted in Figure 1. In order for these processes to be effective, however, IT organizations also need to incorporate more sophisticated sources of management data than is provided by SNMP MIBs. The best time to incorporate this information is not when a key initiative is being deployed, as this approach limits the ability of the IT organization to ensure the successful implementation, fine-tuning and ongoing

[3] Active Working Group (http://mirror.switch.ch/ftp/doc/ietf/ipfix/ipfix-charter.txt)

management of these initiatives. Rather, IT organizations should begin now to integrate new sources of management data, potentially using a network refresh as the trigger to justify incorporating these new data sources. This approach has a higher probability of ensuring the success of a given initiative, and the careers of IT professionals.

A Word from the Sponsor – Cisco Systems

Cisco offers the Network Application Performance Analysis (NAPA) Solution suite to address the customer needs by leading the way in providing a holistic and comprehensive solution for "Performance Management."

Cisco's combined products and services offering, the NAPA solution leverages integrated and advanced rich instrumentation from the "Intelligent Information Network," provides for 24x7 real time monitoring and analysis and offers advanced application and network design and performance troubleshooting capabilities.

The Cisco NAPA solution is an industry-leading, business-class, secure and proven end-to-end performance management solution suite designed, packaged, configured and priced specifically for medium to large business.

For more information on Cisco's NAPA solution, go to http://www.cisco.com/go/napas.

About Kubernan™

Kubernan™, a joint venture of industry veterans Steven Taylor and Jim Metzler, is devoted to performing in-depth analysis and research in focused areas such as Metro Ethernet and MPLS, as well as in areas that cross the traditional functional boundaries of IT, such as Unified Communications and Application Delivery. Kubernan's focus is on providing actionable insight through custom research with a forward looking viewpoint. Through reports that examine industry dynamics from both a demand and a supply perspective, the firm educates the marketplace both on emerging trends and the role that IT products, services and processes play in responding to those trends.

Kubernan is the Greek root word for *helmsman* as well as the phrases to guide and to steer. As such, the name Kubernan reflects our mission of guiding the innovative development and usage of IT products and services.

For more information and for additional IT Impact Briefs, please visit http://www.Kubernan.com.

Kubernan Briefs
Vol.1, Number 5

Published by Kubernan
www.Kubernan.com

Cofounders:
Jim Metzler
jim@kubernan.com
Steven Taylor
taylor@kubernan.com