

Prepare Your Network for the Cloud and Beyond

Managing QoS for your cloud-based applications



Product Disclaimer: LiveAction has renamed their software solution, formerly known as “LiveAction” to “LiveNX.” From 2016 and on, LiveNX will remain the official name for the software solution.

TABLE OF CONTENTS

Introduction	3
Cloud Services Overview	4
QoS Best Practices for Cloud Traffic.....	5
Discussion and Planning	5
Investigation and Design	6
Classifying Cloud Traffic	7
Marking and Queuing for Cloud Traffic	8
Additional WAN Network Considerations.....	9
QoS Implementation Summary.....	10
Tools for a Successful Cloud Services Implementation	11
Network Traffic Baselining and Understanding	12
QoS Design and Provisioning	14
QoS Monitoring and Analysis	17
Traffic Generation and Analysis	19
Conclusion	22
More Information.....	22
About LiveAction.....	22

INTRODUCTION

While there has been an extreme amount of hype surrounding cloud computing, there is a good chance your company is using or will use an important cloud-computing business service in the near future. The operational efficiencies, ease of implementation, and flexible purchasing commitment provided by cloud computing are compelling enough for companies to use in multiple facets of their business.

Since cloud computing services (cloud services) are heavily focused on applications, oftentimes little thought goes into the impact a shift to cloud services has on an organization’s network infrastructure. Switching to a cloud consumption model represents a significant shift in the characteristics of the traffic a network will need to deal with, but the primary concern should always be how these changes will impact the quality of user experience. Answering this question is especially important since most cloud services use standard Internet protocols and applications as the interface for accessing their services. This means, any Quality of Service (QoS) policies you have in place, will by default, treat these applications as casual web traffic and assign them to a Best Effort class.

This white paper will clarify the different types of cloud services, explain implications cloud services have for WAN and Internet architectures, and provide tips on preparing your network with QoS policies to support cloud services. If you’re a system administrator involved in deploying cloud services, this white paper will help you understand the issues faced by the network engineers in your company and how to better communicate with them and line of business (LOB) stakeholders when project planning and implementation occur.

BEFORE CLOUD COMPUTING

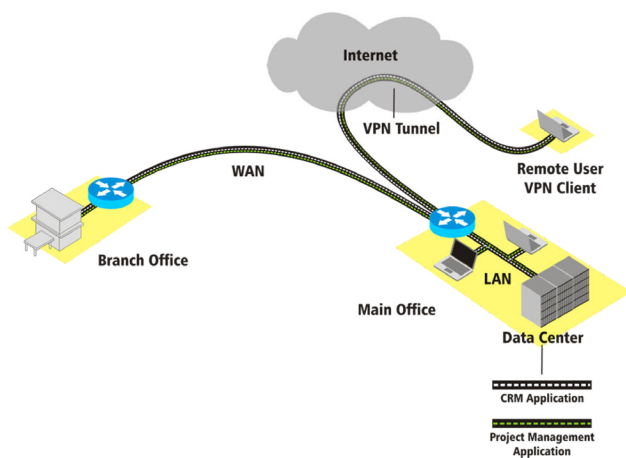


Figure 1. Enterprise applications hosted in main office

AFTER CLOUD COMPUTING

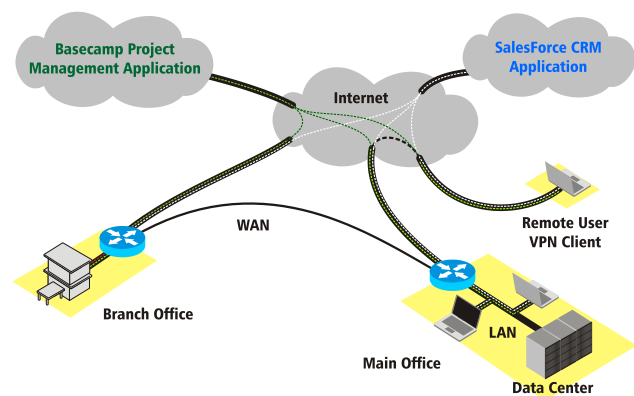


Figure 2. Cloud-based applications

CLLOUD SERVICES OVERVIEW

In the simplest terms, cloud computing is the system that provides utility computing in a scalable and easily consumable form via the Internet. While the concept of utility computing is not new, only the recent maturity of technologies such as high-powered multi-core CPUs, virtualization, and fast and ubiquitous Internet access have transformed the adoption of cloud computing into services that provide real business value.

While cloud services are founded on the concept of utility computing, they can be packaged in many different ways, and in some forms, they do not work with a utility computing pricing model. There is still some flux in cloud service models and offerings, but presently the industry has settled on three primary cloud services models: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). The following table summarizes the different cloud services:

Service	Description	Examples	Notes
SaaS	A software application hosted on an Internet accessible infrastructure that eliminates the need to install, run and store data on the user's own computer. The service is typically paid for on a monthly basis.	SalesForce CRM Google Apps 37signals Basecamp Microsoft Office 365	Completely eliminates the need for end-users to purchase and manage computing infrastructure and software.
PaaS	A computing infrastructure and runtime environment accessible via the Internet whereby users can upload and run application code designed for that environment. The infrastructure is typically paid for on a usage basis (amount of computing resources consumed).	Microsoft Azure Google App Engine Force.com	Eliminates the need to purchase and manage computing infrastructure and license. PaaS also manages extra platform services like database and billing. In addition, PaaS can lock end-users into a specific platform environment.
IaaS	A computing infrastructure accessible via the Internet whereby users can upload and run a virtual machine running their application code. The infrastructure is typically paid for on a usage basis (amount of computing resources consumed).	Amazon Web Services (uses EC2) Rackspace.com	Eliminates the need to purchase and manage computing infrastructure and provides the flexibility to move to other platforms.

QOS BEST PRACTICES FOR CLOUD TRAFFIC

Just like any other IT project, before making any changes to your network and application infrastructure, it is important to plan ahead. Here are the major steps you need to follow when planning for network QoS for cloud services:

1. **Communicate with the other IT teams:** Especially in large companies there will be multiple teams involved in rolling out cloud services. If you're putting unified communications in the cloud, you may be working with a different group than if you were implementing Google Apps for Business. In either case, hopefully those teams are reaching out to you. If they aren't, don't hesitate to interject yourself in the conversation—mention that you'd like to discuss what needs to be done before the rollout rather than try to fix things afterwards.
2. **Understand the traffic:** Cloud services will cause a shift in traffic and usage patterns by adding traffic to and from the cloud service provider. If you're replacing an internal application with a cloud application, that legacy traffic will disappear, changing traffic patterns to and from your data centers. You should develop a good understanding of your current conditions by baselining your network traffic. You can learn more about your cloud service traffic by understanding the services involved (see *Classifying Cloud Traffic*) and by tracking live traffic during a pilot rollout.
3. **Develop initial design:** Put together a rough design of how you want to implement QoS for cloud services into your current architecture. Cloud service traffic fits best in QoS models with 4 or more classes. If baselining indicates that you need to add bandwidth or change the network architecture, incorporate this into the plan.
4. **Review your plan:** Present your plan to other IT teams and key stakeholders in your company. Business and technology leaders will want to know what you are changing and how it might impact services already in place. If the current network architecture needs updating (upgraded WAN services, new routers, or even a different network design), you'll also need to make that justification.
5. **Design the details:** Most cloud service traffic will use existing Internet protocols, so you should know the IP addresses of the services you will be connecting to, in order to distinguish them from casual Internet traffic.
6. **Implement and test:** Try your policies in the lab or as a pilot deployment. This is where you can make sure everything is running correctly and fine tune bandwidth allocations or reassign protocols to different classes if necessary.
7. **Production roll-out:** Implement the policies in a production environment and monitor and fine tune as you go. This step is similar to the pilot project or lab testing, but it is occurring on the live network and on a larger scale.

Discussion and Planning

Before cloud services are deployed, some discussion between the application teams and network teams should occur. This will help proactively address network performance problems that may impact the end-user experience with cloud and other services already running on the network.

Most cloud services are delivered via the Internet using standard Internet client applications. This is appealing for companies since Internet access gives users the flexibility to use the service from virtually anywhere. The downside is that without adjusting your QoS policies, standard Internet protocols such as HTTP and HTTPS will be treated the same as casual Web browsing. Whether or not the anticipated load on the network will have an impact on the end user is best determined in the context of your current network design and the services running across it.

Some specialized services may use dedicated connections directly to the service provider if the bandwidth and quality requirements dictate it. Examples of specialized services are healthcare applications, unified communications, and offsite backup hosted in the cloud. In these cases, you should consider your WAN architecture and the impact cloud services used by your branch offices will have on the network.

Here's a list of issues and questions that should be addressed and considered during the discussion phase:

- **What type of service is being used?** – SaaS, PaaS, IaaS
- **What are the typical activities associated with the service?** – Web GUI interaction, file uploads/downloads, console session interaction
- **What is the nature of the cloud service traffic?** – Typical bandwidth requirements, burstiness of the traffic, real-time vs. non-real-time, business importance of the traffic
- **What are the usage patterns?** – Number of users, frequency of usage, usage vs. time of day, locations of use (branch/HQ/remote)
- **How is Internet access provided for your branch offices?** – Through the HQ vs. dedicated Internet access at each branch
- **Is the cloud service replacing a service that is currently hosted in-house?** – Internal CRM vs. Salesforce CRM

Investigation and Design

Before proceeding to the design stage, baseline what's happening today on your network. You can use a flow monitoring and reporting tool. If you have QoS policies in place, use a tool that allows you to monitor QoS class performance. Using this information you will want to determine the typical utilization of your current network applications and find and fix any trouble spots you encounter. The data may even indicate that further network infrastructure changes are needed to support your new cloud services.

Once you understand where things are today, you'll want to know a little more about the type of protocols and traffic characteristics used by the cloud services.

- SaaS services like Microsoft Office 365, Google Docs and 37signals Basecamp will typically have a lot of light HTTP and HTTPS traffic with occasional large file transfers if documents are uploaded or downloaded.
- PaaS services will likely have minimal impact during application development and post deployment. There may be windows of high network use when applications are uploaded during testing phase or revised after deployment.
- IaaS traffic will vary by the business usage and can be characterized by large file transfers using protocols like SFTP. For example, if storage services like Amazon S3 are used for data back-up or for financial or scientific data storage, this will likely result in regular transfers of large amounts of data.

With this information you can put together a rough design, share it with other IT teams and seek approval from your directors and executives. With the plan approved, it's time to work out the details of your QoS policies.

The primary QoS mechanisms that will be put into use are:

- Identification and classification
- Marking and queuing
- Traffic shaping for WAN links

Classifying Cloud Traffic

Identifying and classifying your cloud traffic is the first step to creating policies. You can use this step during your pilot and production roll-outs to make sure you are properly classifying the traffic before taking any other QoS actions. Also, by using a QoS monitoring tool with a monitoring policy, you will get a good idea of the traffic levels you can expect from your cloud services.

For identifying and classifying cloud services, a couple of approaches can be taken. You can directly use the IP address and/or port numbers used to access the service and create an access control list (ACL). If you need to get the IP addresses for your cloud applications, you can talk to your cloud service provider, or use a flow-monitoring tool to discover them during the course of operation.

An easier method is also recommended. You can take advantage of Cisco's advanced NBAR2 classification engine that can recognize cloud-based applications using deep packet inspection. Using this method you would only need to select the cloud-based application name and then manage the application with marking or queuing policies. Cisco updates their NBAR2 classifications regularly—you can verify which cloud-based applications are supported [here](#).

Marking and Queuing for Cloud Traffic

Once traffic is classified, the appropriate marking action can be taken. The marking action will depend on the class model you are using, as well as the QoS treatment you intend to give the traffic. As an example, let’s consider three of Cisco’s class models:

4-Class Model	8-Class Model	12-Class Model
Real-Time/Streaming	Voice	Voice
	Interactive Video	Real-Time Interactive
		Multimedia Conferencing
Streaming Video	Broadcast Video	
	Multimedia Streaming	
Signaling/Control	Call Signaling	Call Signaling
	Network Control	Network Control
		Network Management
Critical Data	Critical Data	Transactional Data
		Bulk Data
Best Effort	Best Effort	Best Effort
	Scavenger	Scavenger

In the 4- and 8-class models, you will likely map cloud service traffic to the Critical Data or possibly Best Effort class. In the more granular models such as the 11- or 12-class, you have the choice of using the Transactional Data, Bulk Data or Best Effort classes. Transactional Data will include the services that involve user interactions where timely responses are expected. These are typically the CRM, ERP and similar business applications. Bulk Data will be used for high-throughput traffic involving no user interaction such as backups and uploads to cloud storage services (i.e., SFTP and FTP).

When finalizing marking and queuing for your QoS policies, keep these Cisco best practices in mind:

- Mark packets as close to the source of the traffic as possible (i.e. the access switch connected to the device). If this is not practical, packets can be marked at other points in the network where congestion is likely to occur.
- The majority of traffic will be classified as default, so enough bandwidth should be provided to support this type of traffic.
- Real-time traffic should use priority queues and be assigned adequate bandwidth. However, you should limit the overall priority queue to 33% of the available bandwidth to prevent the starvation of other application traffic.

- The total bandwidth allocation for classes other than default should not exceed 75% of a link’s capacity to allow for Layer 2 overhead and Best Effort traffic.
- Recreational or Scavenger traffic should be policed as close to the source as possible to prevent unnecessary bandwidth usage if it exceeds a certain threshold.

Marking close to the source will also be needed on the Internet traffic coming into your WAN edge from the cloud service provider (downstream cloud service Internet traffic).

Additional WAN Network Considerations

At the branch office router and HQ WAN aggregator, there will likely be a couple of additional factors you need to take into consideration:

- Speed mismatches between the LAN and WAN/Internet connections
- QoS treatment for traffic traversing a service provider network

Almost without exception, there will be a speed mismatch when traffic is traveling from an interface on the internal side of your aggregator router to one connected to the WAN or Internet. One method of handling this is to use a hierarchical policy that nests a queuing policy within a shaping policy. This prioritizes traffic within a sub-line rate shaping policy to match the contracted WAN or Internet service capacity. Remember, your capacity is determined by the terms of the services you have contracted from your provider. Just because you have a 100 Mbs link to the provider does not mean that you have a 100 Mbs contracted rate.

HOW TO MATCH MY LINK SPEED TO MY SERVICE PROVIDER’S RATE

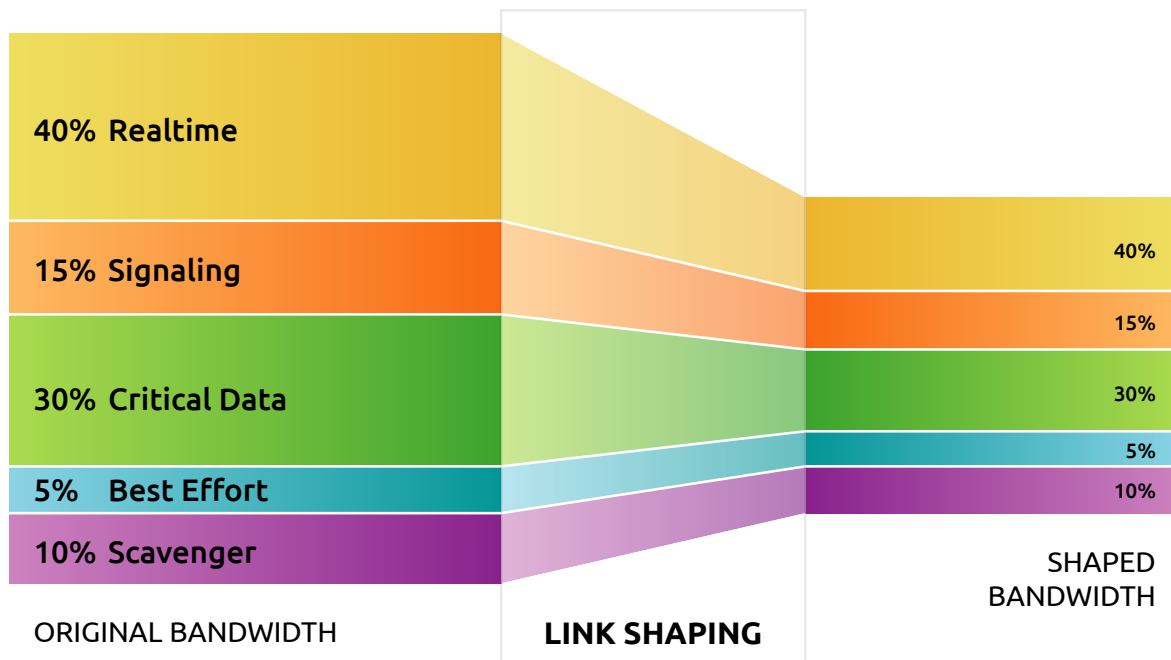


Figure 3. Using hierarchical policies and link shaping to accommodate speed mismatches

The second consideration is if you have an MPLS service with multiple Class of Service (CoS) levels. These typically come in flavors such as 3- or 4-class models. The table below outlines Cisco best practices for remarking and mapping a 12-class model to a 4-class MPLS service. As the table below shows, Transactional Data is mapped to SP-Critical 1 and Bulk Data is lumped into SP-Best Effort.

Application	DSCP Remarking	4-Class SP Model
Voice	EF	SP Real-Time (RTP/UDP) 30%
Real-Time Interactive	CS4 CS5	
Network Control	CS6	SP Critical 1 (TCP) 20%
Transactional Data	AF2 AF3	
Call Signaling	CS3	
Broadcast Video	CS5 CS2	SP Critical 2 (UDP) 20%
Multimedia Conferencing	AF4 AF2	
Multimedia Streaming	AF3 AF2	
Network Management	CS2	
Bulk Data	AF1 DF	SP – Best Effort 30%
Best Effort	CS1 DF	
Scavenger	DF	

QOS IMPLEMENTATION SUMMARY

The guidelines in this section may or may not be applicable given the needs of your business, the architecture of your LAN and WAN, and the traffic types running over it. Make sure to consider all the factors that are unique to your business when designing your QoS policies.

Tools for a Successful Cloud Services Implementation

When deploying QoS for cloud services, or even investigating and troubleshooting performance issues with cloud services, having the right tools will help accelerate the implementation and troubleshooting of cloud services and provide a better end-user experience.

LiveNX, LiveAction’s application-aware network performance management software with QoS control, is designed to simplify network management. With innovative and unique QoS capabilities, it combines several functions for implementing QoS for cloud services into a cohesive interface. The following table summarizes tasks required for implementing cloud services and the LiveNX capabilities that will help:

Function	LiveNX Capabilities
Network traffic baselining and understanding	<ul style="list-style-type: none"> • End-to-end traffic flow visualization • Powerful NetFlow reports for historical traffic analysis • Traffic filtering by IP address, application, DSCP, Ports, etc. • NBAR2 application recognition
QoS monitoring and analysis	<ul style="list-style-type: none"> • QoS audit report for understanding and validating existing policies • QoS baselining of existing policies with SNMP and NBAR2 QoS Reports • Immediate validation of QoS policy changes
QoS design and provisioning	<ul style="list-style-type: none"> • Wizards and templates for easy policy creation based on Cisco best practices • Rules checking and references to ensure error-free changes • Graphical policy editor for fine-tuning policies • NBAR2 for simple traffic classification by application name • Graphical ACL editor for traffic classification by IP address • Single-click policy rollout or updates to multiple devices • Policy snapshot and rollback
Synthetic traffic generation and analysis using Cisco IP SLA	<ul style="list-style-type: none"> • Traffic generation and analysis for measuring the impact of QoS policies in a controlled manner • Create and manage synthetic HTTPS, UDP and video traffic • Monitor results of synthetic traffic tests in summary and detailed reports

Network Traffic Baseline and Understanding

LiveNX has several analysis tools available for the information gathering that you will need in your cloud service implementation.

LiveNX's end-to-end flow visualization helps with identifying current traffic patterns and gives you the power to troubleshoot any potential issues that need to be resolved.

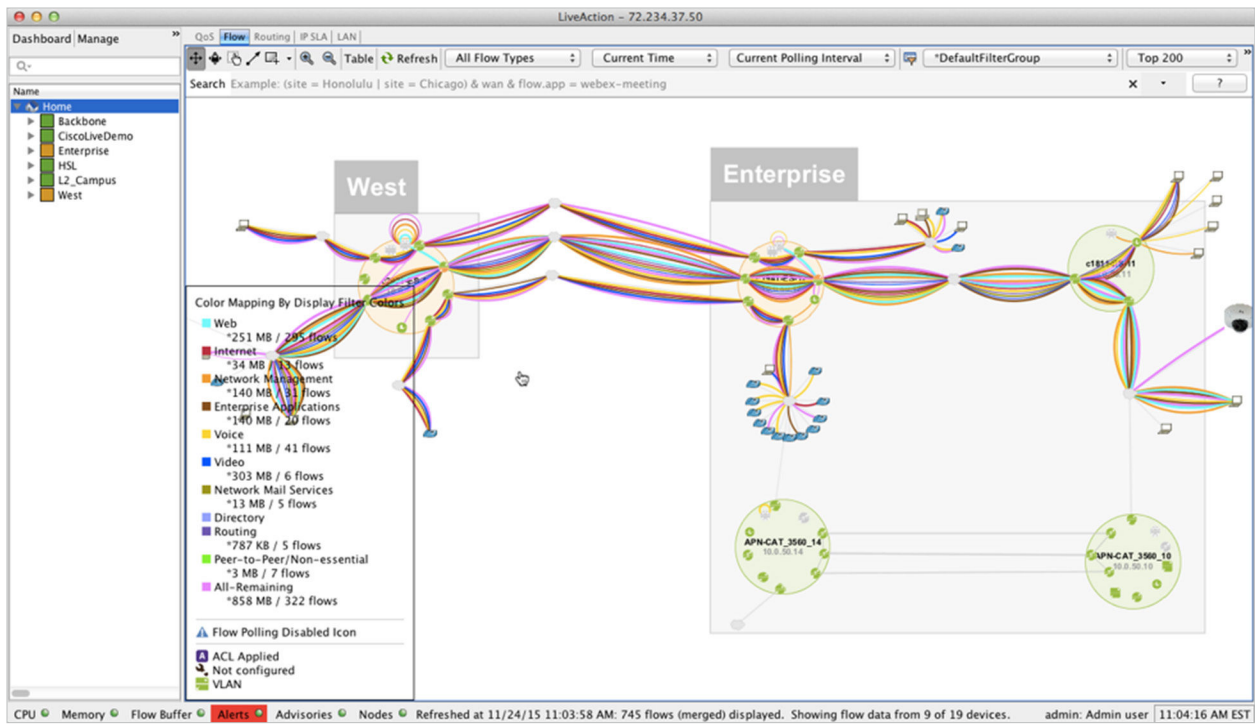


Figure 4. Viewing network-wide traffic flows using LiveNX's topology screen

LiveNX's playback allows a DVR like approach to identify which traffic may need to be addressed with your QoS policy implementation.

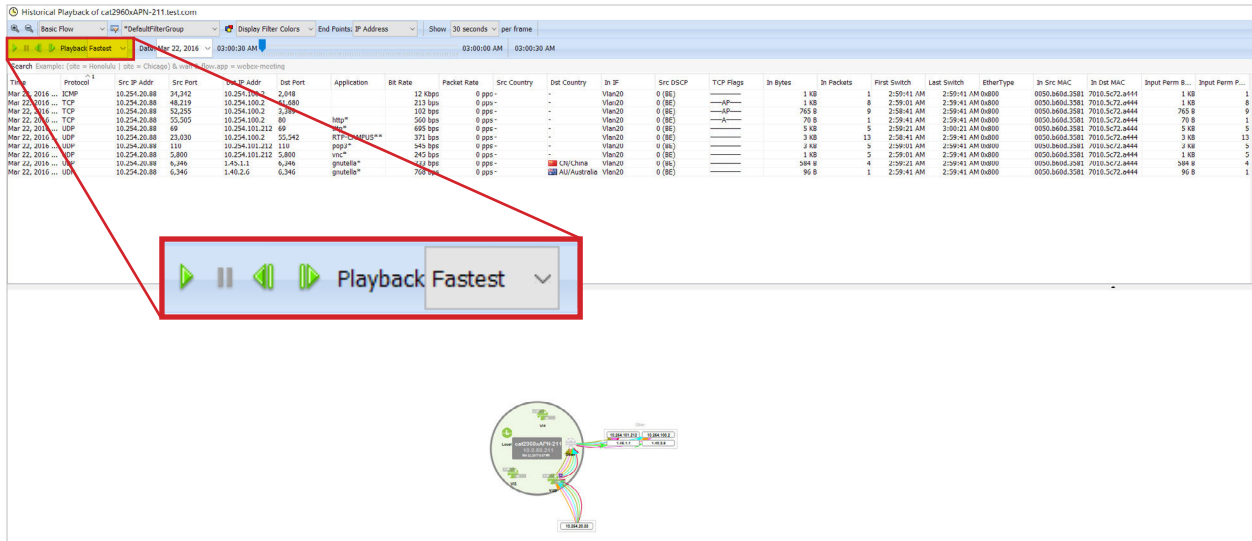


Figure 5. Playback LiveNX's historical flow data

Rich reports will give current utilization for the entire network, or down to a per-device and per-interface for selected time periods for Applications, Top-Talkers, DSCP markings, etc.

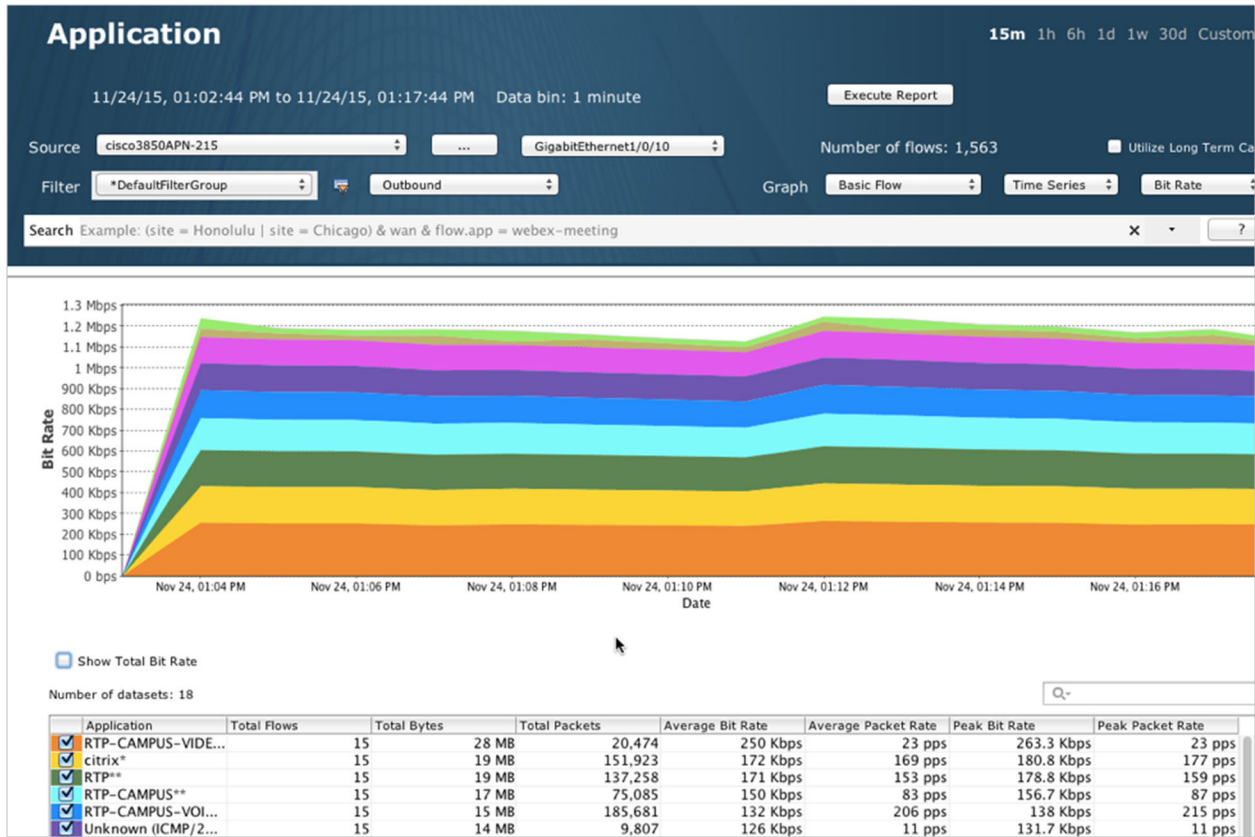


Figure 6. Using LiveNX’s historical flow data to verify BW utilization by application

QoS Design and Provisioning

LiveNX greatly simplifies QoS policy creation with expert capability in the form of GUI wizards, templates, and interactive screens for designing, deploying, and adjusting QoS policies.

Policies can be created using built-in wizards based on Cisco design recommendations.

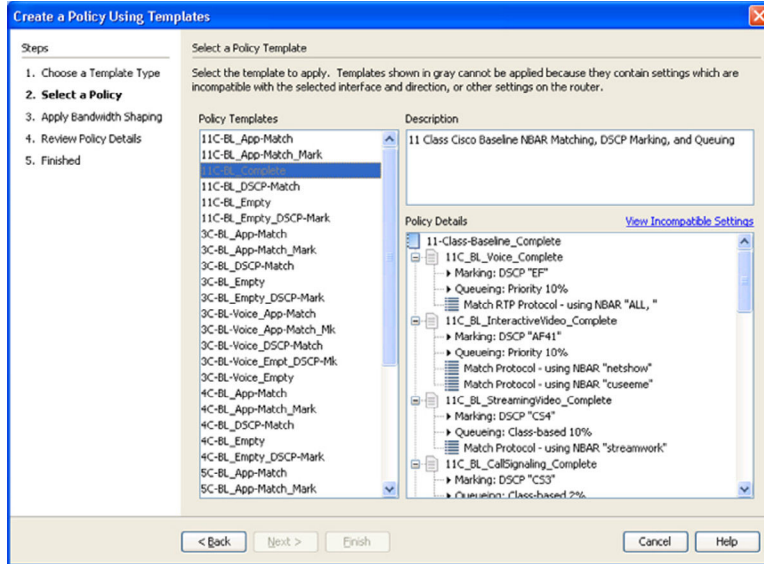


Figure 7. Creating a QoS policy using LiveNX's template wizard

Policies can also be customized and managed with a comprehensive QoS policy editor/manager.

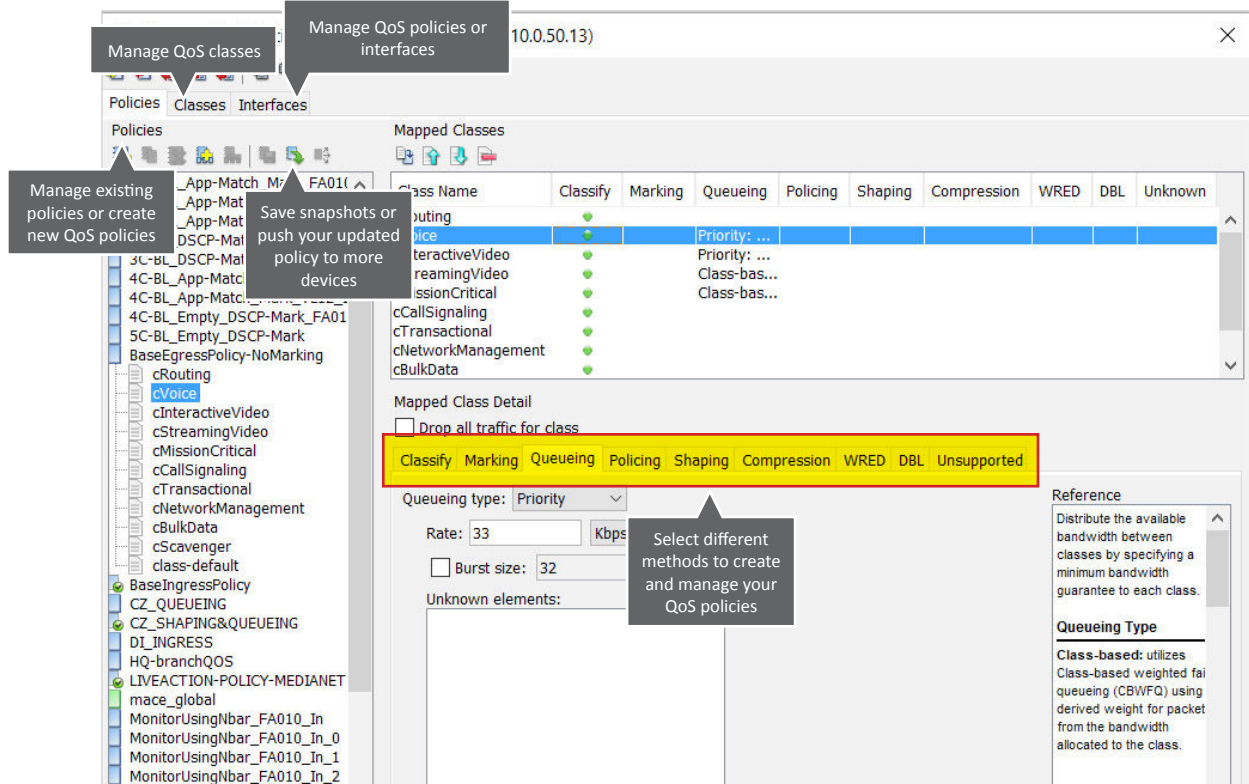


Figure 8. Using LiveNX's QoS editor and manager to adjust policies

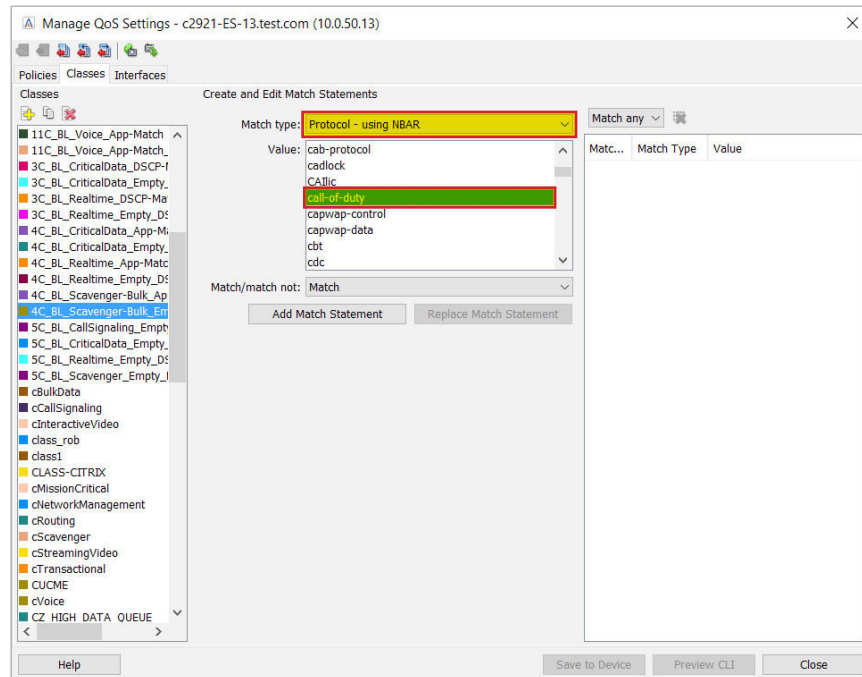


Figure 9. Using LiveNX’s QoS policy editor and NBAR2 to manage cloud-based applications based on advanced classification techniques

ACLs can be created and edited using LiveNX’s built in ACL editor:

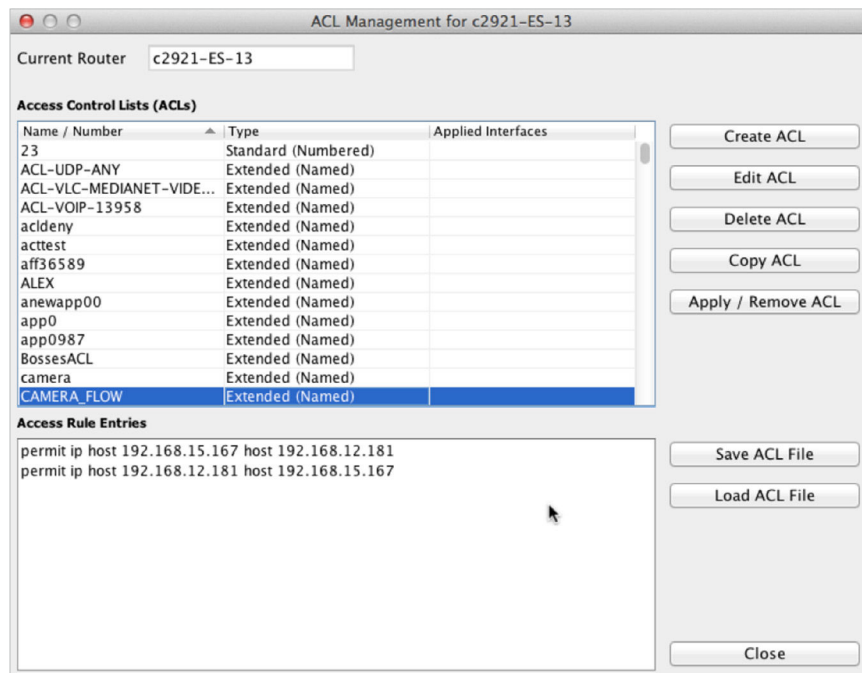


Figure 10. Create, edit and apply ACLs

Or ACLs can be instantly created using NetFlow data:

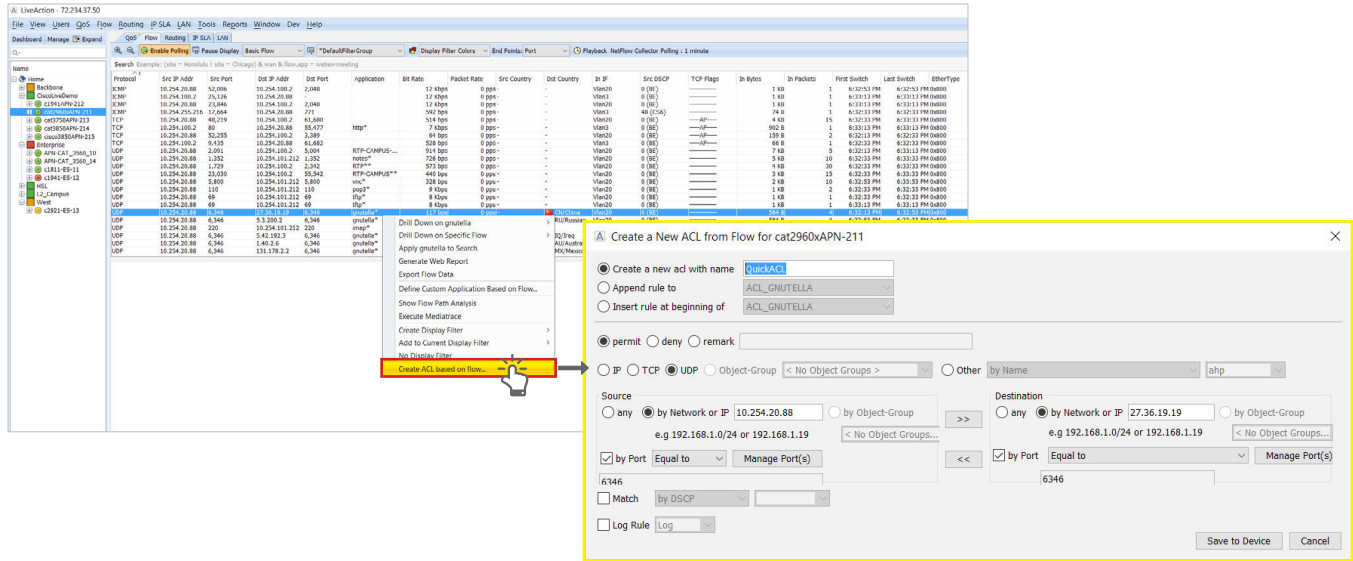


Figure 11. Use NetFlow to create ACLs

QoS Monitoring and Analysis

LiveNX's QoS monitoring tools will give you access to information that will help you better understand the performance of existing policies, and immediately validate changes made along the way.

If policies are already in use you can run a QoS Audit Report to quickly identify and understand any performance or configuration issues.

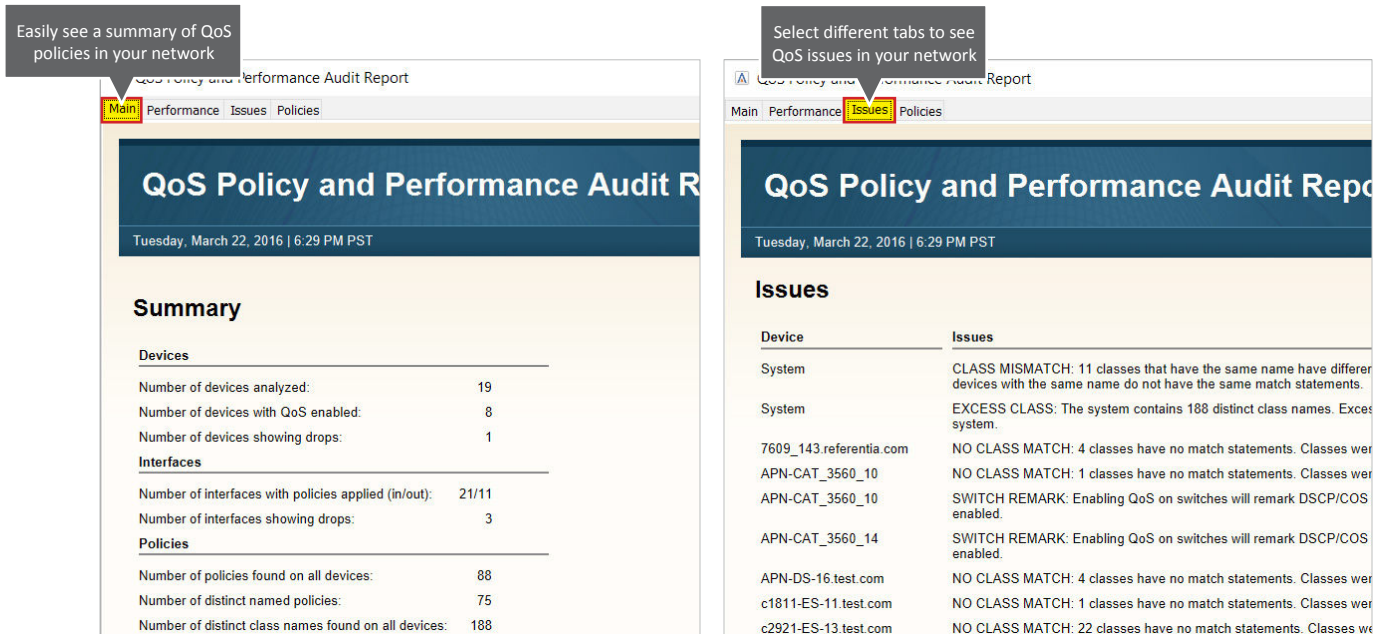


Figure 12. QoS Audit Report automatically summarizes policy issues

You can view the performance of a QoS Policy on the Input or Output of an interface on a per-class basis and even compare pre- and post-policy performance and NBAR2 application throughput. You can view this information in real time, or on a historical basis using LiveNX QoS Reports.

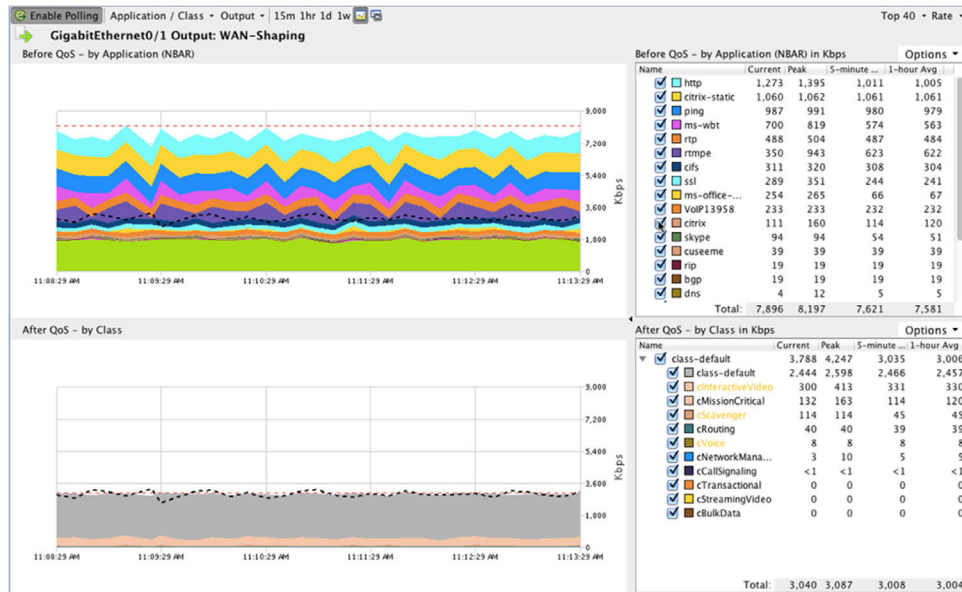


Figure 13. Viewing real-time results of applied QoS policies

A simplified editor is available for quick adjustment of queue settings and reserve bandwidth adjustments for policies that are actively in use.

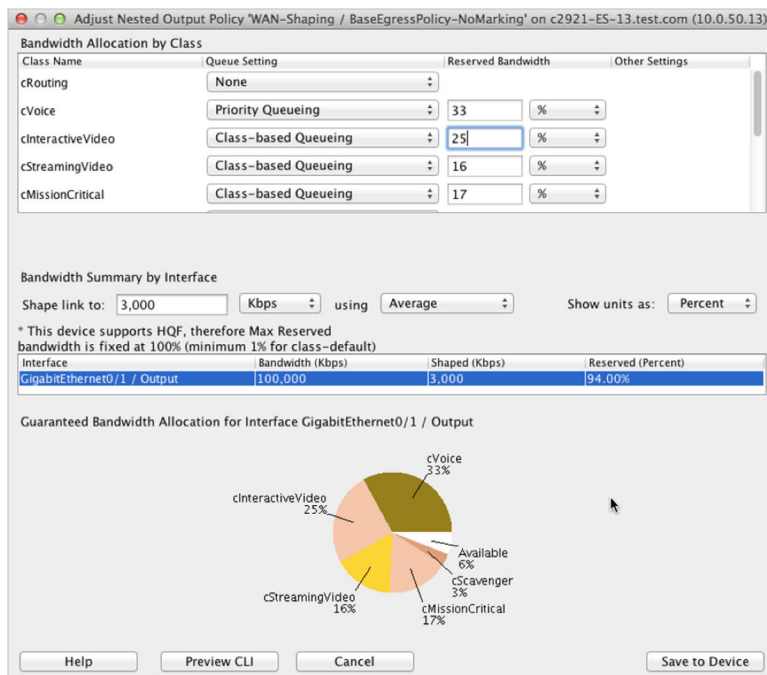


Figure 14. Easily view and adjust QoS policies

In addition, troubleshooting and implementation will be much easier with LiveNX's real-time, system-wide flow visualization. This is critical for quickly gaining end-to-end awareness of the traffic flowing across the network and identifying points of congestion, incorrectly configured services, and invalid traffic.

Using LiveNX Flow Path Analysis you can track down a single user's conversation and view the performance of that conversation hop-by-hop through the network.

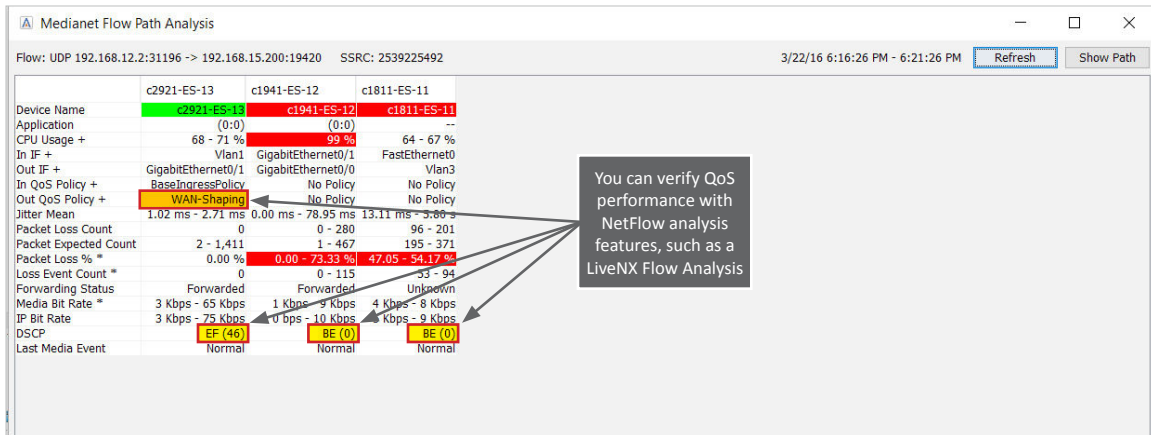


Figure 15. Flow Path Analysis

Traffic Generation and Analysis

Whether experimenting in the lab or verifying performance in an operational network, LiveNX helps you unlock Cisco's IP SLA capabilities for traffic generation and analysis. LiveNX helps you quickly set up tests using HTTP, UDP, Jitter, Voice and other traffic types, visualize where those tests are active, and analyze the results.

You can easily create an IP SLA test with a right-click on a managed device.

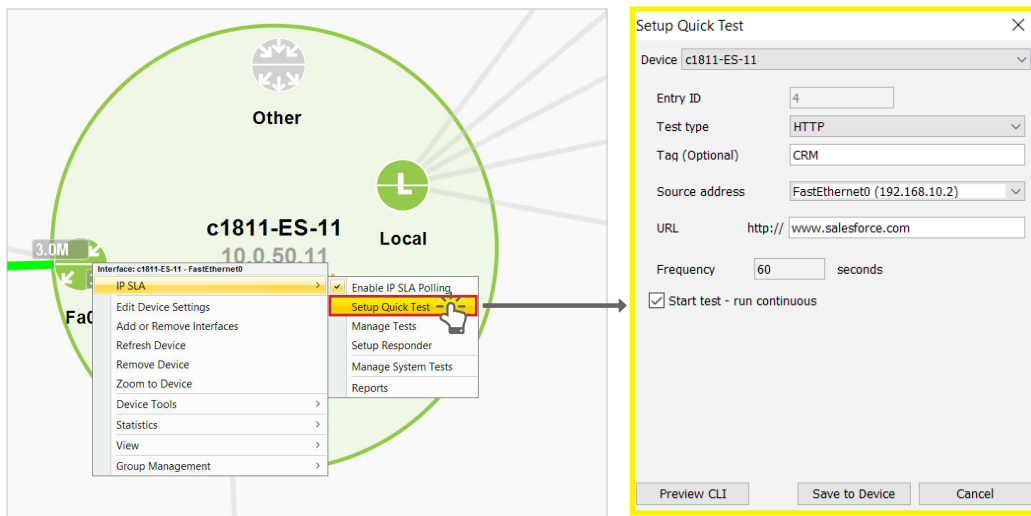


Figure 16. Create an IP SLA test

You can then easily manage and adjust IP SLA test settings to meet your network specifications.

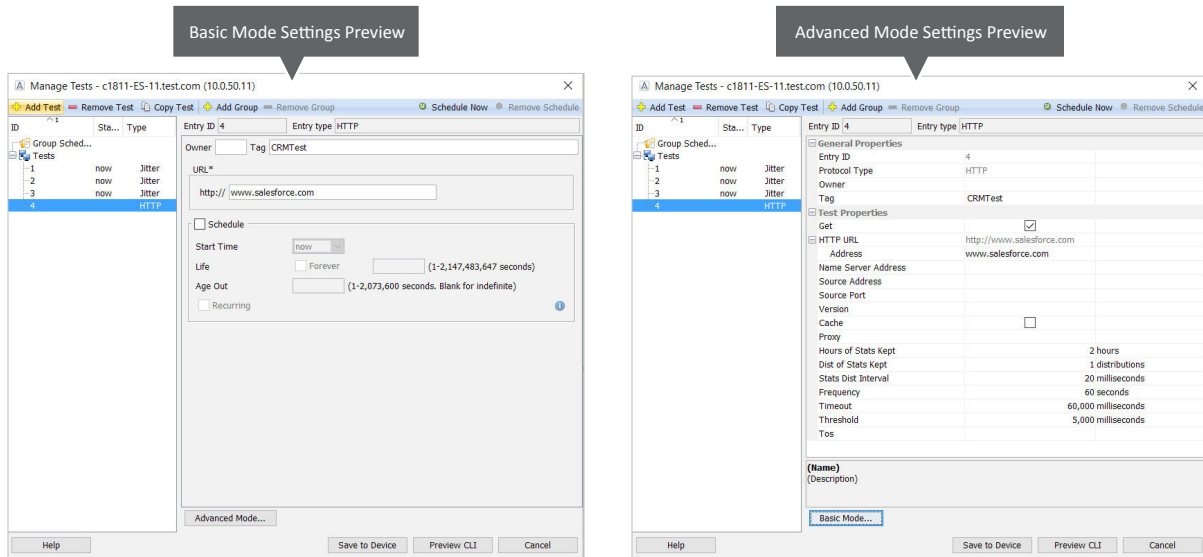


Figure 17. Manage tests using basic and advanced test configuration views

With LiveNX’s visual topology you can easily see the status of an IP SLA test.

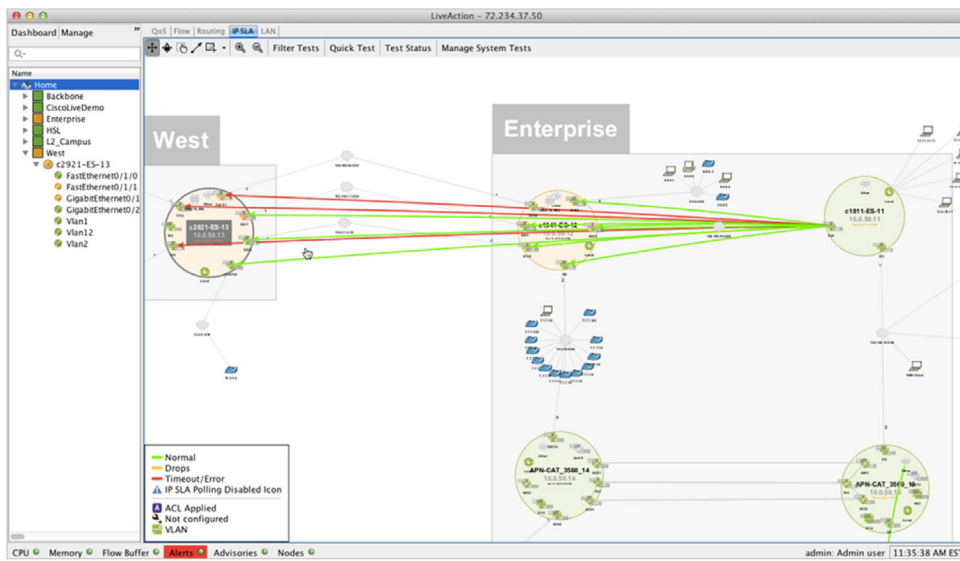
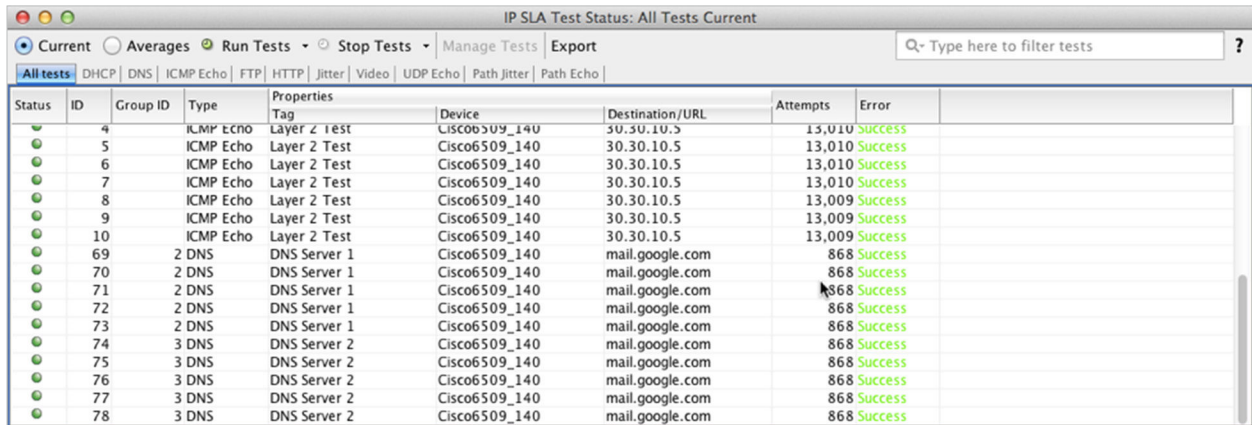


Figure 18. View of running traffic tests through network topology

You can bring up the test status for all running IP SLA tests on the topology.



Status	ID	Group ID	Type	Properties Tag	Device	Destination/URL	Attempts	Error
●	4		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,010	Success
●	5		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,010	Success
●	6		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,010	Success
●	7		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,010	Success
●	8		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,009	Success
●	9		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,009	Success
●	10		ICMP Echo	Layer 2 Test	Cisco6509_140	30.30.10.5	13,009	Success
●	69	2	DNS	DNS Server 1	Cisco6509_140	mail.google.com	868	Success
●	70	2	DNS	DNS Server 1	Cisco6509_140	mail.google.com	868	Success
●	71	2	DNS	DNS Server 1	Cisco6509_140	mail.google.com	868	Success
●	72	2	DNS	DNS Server 1	Cisco6509_140	mail.google.com	868	Success
●	73	2	DNS	DNS Server 1	Cisco6509_140	mail.google.com	868	Success
●	74	3	DNS	DNS Server 2	Cisco6509_140	mail.google.com	868	Success
●	75	3	DNS	DNS Server 2	Cisco6509_140	mail.google.com	868	Success
●	76	3	DNS	DNS Server 2	Cisco6509_140	mail.google.com	868	Success
●	77	3	DNS	DNS Server 2	Cisco6509_140	mail.google.com	868	Success
●	78	3	DNS	DNS Server 2	Cisco6509_140	mail.google.com	868	Success

Figure 19. Test traffic status summary table

You can drill down on the data in the summary table to LiveNX IP SLA reports to view the performance of a single test over a selectable period of time.



Figure 20. Analyze IP SLA performance with LiveNX Reports

CONCLUSION

Cloud services can provide immense operational efficiencies for applications such as office productivity, offsite data storage and analysis, and application development and hosting. Whether your company is migrating existing desktop applications to the cloud or introducing new services hosted in the cloud, application changes will drastically affect your network traffic patterns. By planning for these changes and properly implementing QoS for your new cloud services, you can avoid future headaches for both end-users and IT teams. Ultimately, your enterprise will be able to optimize operational efficiencies and user experience to the fullest.

MORE INFORMATION

User Experience Monitoring

Find out why—and how—User Experience Monitoring can accelerate problem resolution and simplify your application performance monitoring challenges.

Upcoming Webinars

Check out our updated webinar schedule—gain insights from our special presenters about topics like QoS, Hybrid WAN Management, Capacity Planning and more.

Additional Resources

Case studies, white papers, eBooks and more are available for your learning on the LiveAction resources page.

LiveNX and LiveUX Downloads

Free downloads of [LiveNX](#) and [LiveUX](#) are available now. Visit our webpage to discover more details and benefits of LiveNX and LiveUX.

ABOUT LIVEACTION

LiveAction provides comprehensive and robust solutions for Network Performance Management. Key capabilities include Cisco Intelligent WAN visualization and service assurance, best-practice QoS policy management, and application-aware network performance management. LiveAction software's rich GUI and visualization provide IT teams with a deep understanding of the network while simplifying and accelerating management and troubleshooting tasks.

©2016 LiveAction, Inc. All rights reserved. LiveAction, the LiveAction logo and LiveNX Software are trademarks of LiveAction. Other company and product names are the trademarks of their respective companies.

LiveAction, Inc. · 3500 West Bayshore Road · Palo Alto, CA 94303 · USA · +1 (888) 881-1116