

# Service Level Monitoring with Cisco IOS Service Assurance Agent

**Cisco IOS Service Assurance Agent (SAA) is an embedded performance-monitoring agent in Cisco IOS software. It provides a scalable, effective, and low-cost solution for service level monitoring. This document provides an overview on SAA as well as some of the major benefits.**

## Service Level Management

A few years ago, the acronym “www” Internet was jokingly referred to as the World Wide Wait, as the download of one page often took several minutes. At the time, most business users were reluctant to rely on network services for any critical tasks. Over the years, both bandwidth and routing intelligence have increased dramatically, resulting in tremendous service quality improvements. At the same time, technologies advancements have powered networks to a higher level. IP convergence makes it possible to carry data, voice, and video in the same network, providing new service capabilities such as virtual showrooms, web-based conferencing, and Video on Demand (VoD). More and more businesses have adopted the Internet as an important productivity tool, and it is now truly an integral part of our work and life.

As large numbers of users adopted the Internet, there came the demand for better user experiences. Users need clear IP-based voice services, faster responses from financial transactions, and smooth video conferencing. This demand is profoundly important to business users, because any degradation of services directly impacts

profits. Network Service Providers (NSPs) who cannot deliver these expectations risk losing customers to competitors. More NSPs are scrambling to offer service level agreements (SLAs) to their customers. SLAs provide customers a degree of predictability regarding the services they are receiving.

Enterprises are also setting service level objectives for their networks. While enterprise Information Technology (IT) teams do not necessarily need SLAs, the requirements are often similar. Service level objectives are a vital part of an IT team’s main responsibilities. An IT team is accountable to the enterprise in much the same way that a NSP is accountable to its customers, so defined levels of service are a requirement for both.

SLAs vary, but typically include the following provisions:

- **Availability:** the percentage of time the service is available; for example, 98% availability
- **Predictability:** network delay, jitter, and packet loss for the specific class of services; for example, average 100 ms round-trip delay between branches, with less than 3 percent packet loss for data VPN services.
- **Reparation:** mean time to repair.



- Service: customer service hour and process, for example, response to network problems within 2 hours.
- Credit: reimbursement schedule for violating the guaranteed service level, for example, 15 percent credit toward the next month's service fee.

A successful SLA offering requires network performance monitoring for the following activities:

- Capacity Planning  
In order to guarantee a certain service level, it is extremely important to assess the capacity of the network, the traffic mix, and the traffic pattern. Only by monitoring the network over time can one provide the optimal service level guarantees.
- Service Validation  
Any guarantee requires validation. A network provider has to continuously monitor the service level to make real-time adjustments and to plan for future services. Also, customers can ensure that their expectations are met by reviewing the measurements. Offering real-time measurement to customers is also a valuable marketing tool for promoting service step-ups.
- Real-time Troubleshooting  
Just as the traffic in the network is dynamic and changes constantly, so does the service level of the traffic. Real-time monitoring ensures that the network provider identifies and resolves any problem in a timely manner. Real-time monitoring also gives the users more confidence in their expectations.

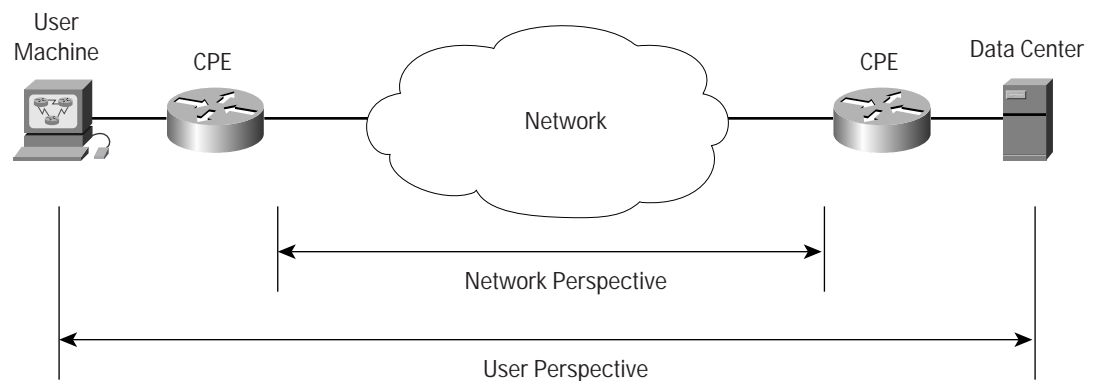
## Performance Measurement Tools

There are a number of network performance measurement tools in the market. These tools differ in the following dimensions: scope of measurement, collection method, and measurement method.

### Scope of Measurement

Scope of measurement refers to what the measurement is about. In general, network performance measurement tools can monitor performance either from the user's perspective or from the network's perspective.

Figure 1  
Measurement Scope





## User Perspective

Some measurement tools examine the performance between a user console and a data center server to give the service level from the user's perspective. Typically, a software agent is installed in both the user's machine and the server. By capturing the timestamps of the user's request and response, the agent can determine the actual application response time. This approach has the benefit of determining the end-user's performance experience.

There are four major disadvantages to this method:

- The measurement must be reported to a network management system (NMS) to be useful. This approach would not be feasible if the NMS had to gather measurement from a large number of end-user stations.
- Polling information out of the end-user stations is subject to intrusion of privacy.
- The response time measurement depends on the load of the user's machine.
- The applications running on the user's machine usually fall out of the administrator's responsible domain. An SLA based on these measurements will likely fail.

## Network Perspective

The other approach is to examine the performance between the customer routers, commonly known as customer premises equipment (CPE) devices. These measurements show the network's impact on the overall performance experience for the user. Although the measurements do not reflect the true "end-to-end" performance, they are within the responsible domain of the network provider. Further, since the number of CPEs is substantially fewer than the number of end-user stations, this approach fares much better in scaling than the first approach. Not surprisingly, most network providers prefer this approach. SAA is a monitoring tool for measuring performance from the network perspective.

## Collection Method

There are two main collection methods used by network monitoring tools: external probe and embedded agent.

### External Probe

Network performance can be measured by placing specialized hardware probes in the network. These probes can either actively send out synthetic packets to other probes, or passively observe packets flowing across the probes.

The major benefit of the external probe method is that it does not interfere with the network traffic, and generally provides accurate reading of the network performance. However, adding hardware probes to the network means additional costs and maintenance. In order to measure performance levels in the overall network, it is necessary to deploy these probes in many locations. A typical SLA requires performance measurement between CPEs to the core network; therefore, it is necessary to deploy these probes at each CPE location. It is expensive and difficult to deploy these specialized probes. Usually, network providers use this method to monitor selective links of the network, such as the link between the point-of-presence (POP) sites.

### Embedded Agent

Another collection method is to embed a software agent in the routers, which serves as a software "probe". This method eliminates the need to deploy and manage hardware probes and is therefore more cost effective. Since the agents reside in the routers, they have precise knowledge of the routers' delays and can therefore provide very accurate performance measurements. SAA is a performance measurement agent embedded in routers powered by Cisco IOS Software.



## Measurement Method

There are two methods available for calculation of the actual measurements: passive (observed) or active (synthetic).

### Passive

The passive method observes the existing packets across the network. By correlating the sources, destinations, and timestamps of the packets, a passive monitoring tool can provide the measurements of the actual traffic. Since it only observes the traffic, it does not interfere with the actual operation. However, processing all packet headers is extremely resource-intensive and can only be done in a specialized external probe. Additionally, because the passive method only looks at the actual traffic, it cannot be used for future service planning.

### Active

In contrast to the passive method, the active method (i.e., Cisco IOS SAA) generates synthetic packets. This is analogous to testing a traffic delay on the highway by driving across it. Although it does inject additional traffic into the network, the added traffic is typically negligible and completely predictable.

Additionally, synthetic traffic allows a network provider to plan new services by generating the new type of traffic into the network and assessing the performance.

## Cisco IOS Service Assurance Agent

SAA is an embedded, synthetic performance-monitoring tool to measure service performance from the network perspective. Because SAA is embedded in Cisco IOS Software, it is platform-independent for all Cisco routers and switches running Cisco IOS Software.

## Major Benefits

### Extensive Traffic Coverage

SAA supports a comprehensive list of protocols and operation types, including Internet Control Message Protocol (ICMP), Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Hypertext Transfer Protocol (HTTP), Dynamic Host Configuration Protocol (DHCP), Domain Name System (DNS), File Transfer Protocol (FTP), Asynchronous Transfer Mode (ATM), Frame Relay, Systems Network Architecture (SNA), email, and SAP applications.

Each operation provides a metric of performance measurement relevant to a particular purpose, making the tool both flexible and powerful. For example, the metric for the UDP jitter operation can provide information for round-trip delay, one-way delay, delay variance (also called jitter), as well as packet loss. The collection of different operations provides a convenient way to monitor different applications such as website access and voice over IP simultaneously.

SAA can also be used to monitor a Multiprotocol Label Switching (MPLS) Virtual Private Network (VPN) backbone. An SAA operation can be attached to a Virtual Route Forwarding (VRF) name. SAA can use this feature to measure the performance level specific to any particular VPN tunnel.



### Real-time, Accurate Measurements

For network providers, the usual service boundary is between the source and the destination CPEs. The processing delays in the end CPEs are usually out of the range of responsibility. A Cisco end device can function as an SAA responder in order to factor out such processing delays. The responder has the intimate knowledge of the router and thus can allow the source to subtract the processing delay from the measurement calculations.

SAA can also provide hop-by-hop performance information. These metrics can potentially pinpoint the exact location of traffic congestion and allow timely reparation.

Further, each operation in SAA can be configured to monitor per-class traffic. This functionality makes SAA an effective tool for Quality of Service (QoS) deployment.

### Flexible Scheduling

SAA can be configured to run at any given time, or continuously over any time interval. This functionality is important for continuous SLA validation. If performance issues do occur, SAA can be run at any time for troubleshooting purposes.

Additionally, an SAA operation can be activated when a pre-set threshold is crossed. For example, a network provider can set the threshold of the round-trip delay to that listed in the SLA. If the real-time round-trip measurement exceeds this threshold, a separate SAA operation can be activated dynamically to collect hop-by-hop response time. The dynamically activated operation can collect more real-time, detailed statistics that have a strong correlation to the actual problem. The administrator can promptly isolate and address the problem, dramatically improve SLA fulfillment, and thus increase customer satisfactions.

### Flexible Reporting Mechanism

Obtaining the measurements is another important aspect of performance monitoring. SAA provides data access through the command line interface (CLI) or Simple Network Management Protocol (SNMP) via RTTMON MIB. The CLI is a convenient way to configure and run SAA operations, as well as to view statistics. For a larger scale monitoring, external applications can programmatically access the SAA measurements using the RTTMON MIB.

SAA stores some amount of historical data. In addition to providing the latest measurement, it can also store this data in an aggregated format. With the new enhanced history feature, a user or an application can configure SAA to store aggregated measurements in "buckets." For example, SAA can be configured to store forty-eight buckets, and each bucket maintains fifteen minutes of the aggregated measurements. With this configuration, it can store twelve hours of performance information.

SAA can also report information asynchronously. A user or an application can configure an SAA operation with a set of thresholds. An SNMP trap will be generated if any threshold is crossed. A potential use of such trigger is to link it with a paging and email system. When a threshold is crossed, the responsible administrator can be notified in real time. Further, a threshold crossing can trigger a new SAA operation to complete more detailed investigations such as hop-by-hop measurements. These features can significantly reduce the mean time to repair, a key component in most SLAs.



## Supported in All Cisco IOS Software-Based Platforms

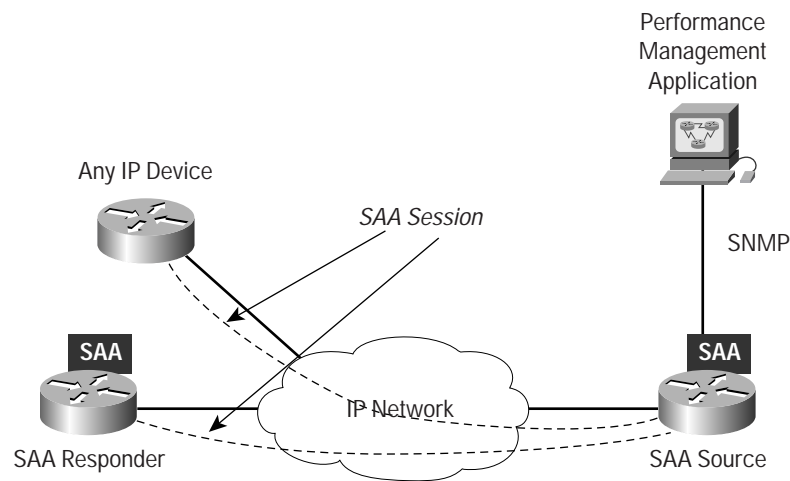
Because SAA is a software agent embedded in Cisco IOS Software, it is included in all routers that run Cisco IOS Software. A network provider can monitor the performance between anywhere in the network: core, edge, and the CPEs. Monitoring can be done anytime, anywhere without deploying any physical probe. This unique advantage makes SLA activation easy and fast.

### How SAA Works

SAA uses synthetic traffic to measure network performance between two routers. As depicted by Figure 2, it starts when the SAA source sends a synthetic packet to the destination. Once the destination router receives the packet, depending on the type of operation, it will respond with time-stamp information for the source to make the calculation on performance metrics.

There are three classes of operations, each differing from one another in terms of the listening port number, number of packets sent, and network requirements.

Figure 2  
SAA Operations



### ICMP Operations

With ICMP operations, the source SAA sends several ICMP packets to the destination. The destination router, which can be any IP device, echoes with replies. The source SAA uses the sent and received timestamps to calculate the response time.

The ICMP Echo is the simplest operation and resembles the traditional extended ping utility. It measures only the response time between the source and the destination.

The ICMP Path Echo and Path Jitter operations use the traceroute utility to identify the whole path. By subsequently sending ICMP packets to the path nodes and correlate the measurements, SAA can provide hop-by-hop round-trip delay as well as jitter information. The Path Echo and Path Jitter operations are particularly helpful in troubleshooting, because they detect bottlenecks.



Lastly, a user or an application can use the Loose Source Routing (LSR) options to specify a path. This feature can specify the exact hop-by-hop path of the ICMP packets.

### Special Purpose Non-Responder-Based Operations

These operations are used to monitor specific traffic, such as HTTP, FTP and DHCP. The destination can be any IP device that supports the protocol. These protocols usually have well known port numbers (i.e, port 80 for HTTP).

The measurement metrics vary from protocol to protocol. Typically, the most important measurement is the server response time, as the destination acts as the server of such protocol.

If a session is open during the measurement time, it will be closed at the completion of operation. For example, in DHCP measurement, the SAA source requests an IP address from a DHCP server. After the server response time is calculated, the SAA source releases the acquired IP address.

### Responder-Based Operations

A responder-based operation requires the SAA responder to be enabled in the destination router. An SAA responder is an SAA feature that allows the router to participate in the performance measurement.

The SAA responder provides the following functionality:

- SAA source can communicate with the responder using the SAA control protocol. The control protocol allows the source to instruct the responder to listen to any given port number. This not only avoids port number conflicts, but also provides some measure of protection against Denial of Service (DOS) attacks. The MD5 authentication method can be used in the control protocol to increase security.
- Some performance metrics require the SAA responder to participate in the measurement. For example, the calculation of one-way delay requires the timestamp when the packet arrives at the destination router. The calculation of packet-loss requires the responder to monitor out-of-sequence packets in order to assess how the packets are lost.
- The SAA responder can also dramatically increase the measurement accuracy by providing an assessment to the source of processing delay in the destination router. The processing delay of the routers is a measurement error from a network performance monitoring perspective. This error can be substantial when the router's CPU utilization is high.

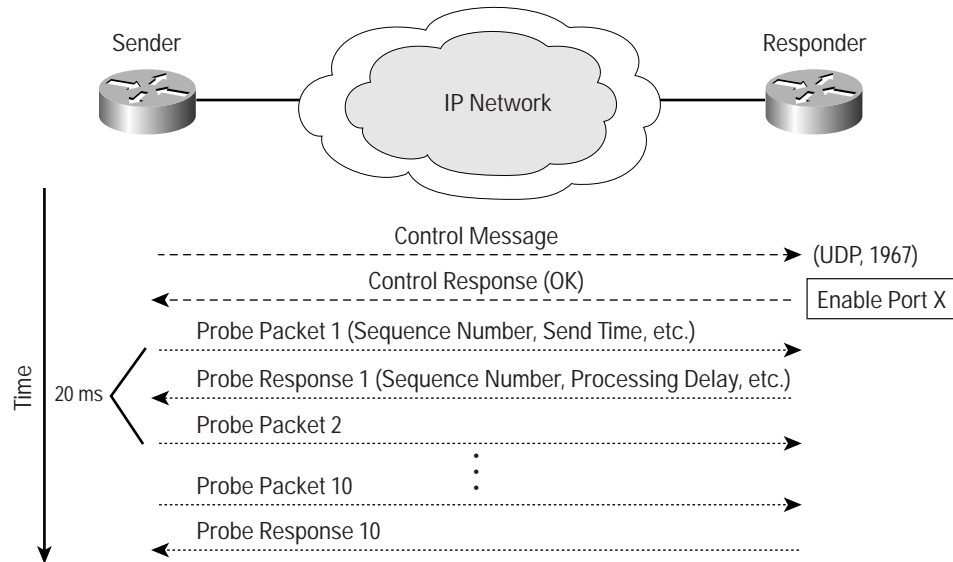
The SAA responder is embedded in the router, so it has intimate knowledge of the router and can provide information, such as the processing delays of the SAA packets. In its replies to the SAA source, the SAA responder includes information such as processing delays. The SAA source can then remove the delays in its final calculations.

The UDP Jitter and UDP Echo probes are examples of responder-based operations. The UDP Jitter operation requires the destination device to be enabled as a responder, but this is optional for the UDP Echo operation.

Figure 3 shows the packet flows for a UDP Jitter operation. For this example, the operation is configured to send 10 consecutive packets with 20 ms time spacing, which is the default setting for UDP Jitter.



Figure 3  
UDP Jitter Operation



### SAA Applications

Table 1 shows the different SAA operation types as well as their key applications.

Table 1 SAA Operation Types and Key Applications

Operation	Measurement Capability	Key Applications
<b>UDP Echo</b>	Round-trip delay	Accurate measurement of response time of UDP traffic
<b>UDP Jitter</b>	Round-trip delay, one-way delay, jitter, packet loss Note: One-way delay requires time synchronization between the SAA source and target routers.	Most common operations for networks that carry UDP traffic, such as voice.
<b>TCP Connect</b>	Connection time	Website performance monitoring
<b>DNS</b>	DNS lookup time	Website performance monitoring, trouble-shooting
<b>DHCP</b>	Round-trip time to get an IP address	Website performance monitoring, trouble-shooting
<b>FTP</b>	Round-trip time to transfer a file	Website performance monitoring
<b>HTTP</b>	Round-trip time to get a web page	Website performance monitoring
<b>ICMP Echo</b>	Round-trip delay	Troubleshooting and availability measurement
<b>ICMP Path Echo</b>	Round-trip delay for the full path Note: The path can be determined using LSR.	Troubleshooting



Table 1 SAA Operation Types and Key Applications

Operation	Measurement Capability	Key Applications
ICMP Path Jitter	Round-trip delay, jitter and packet loss for the full path	Troubleshooting
DLSw+	Peer tunnel performance	Data Link Switching (DLSw) peer tunnel Performance Monitoring
Frame Relay	Circuit availability, round-trip delay, frame delivery ratio	WAN SLA monitoring
ATM	Availability, round-trip delay and delivery ratio. Supported through Visual Network UpTime.	WAN SLA monitoring

### Performance Management Applications Supporting SAA

Because SAA is an intelligent agent embedded in the router, it provides only raw statistics. The programmatic interface allows external applications to configure and retrieve performance metrics. The external applications can then use the metrics to analyze and present the measurements graphically, as well as to generate service-level monitoring reports.

SAA is both powerful and versatile; many industry-leading applications have integrated the feature into their products. The following is a partial list of the applications that use SAA to collect performance metrics:

- Cisco Systems
  - CiscoWorks Internetworking Performance Monitor (IPM)  
The CiscoWorks IPM is a performance management application that can obtain real-time performance statistics from SAA and other MIBs in the router. It provides a graphical front-end, and is thus a lightweight yet complete performance monitoring application.
  - Cisco Network Service Performance Engine (CNS PE)  
The CNS PE is mediation software that provides aggregated performance statistics to high-level monitoring applications. It allows rapid integration with network management applications to take advantage of not only SAA but also other performance management tools in the router including NetFlow, Radius Call Detail Record (CDR), various MIBs, and text files. An application can use this information to build a correlated, complete picture of the network's health.
  - VPN Solution Center (VPN SC)  
The VPN SC is a comprehensive management application for managed VPN services. It provides different aspects of VPN management: provisioning, performance monitoring, and billing. Since one of the most important aspects of managed VPN services is performance monitoring, it uses SAA as the main tool to provide real-time measurements.
- Agilent Technology
  - Firehunter
- Brix Networks
  - BrixWox

- Concord
  - eHealth
- InfoVista
  - VistaView
  - PowerView
- Visual Networks
  - IPInsight
  - UpTime

## Additional Information

Further details about SAA, including links to usage documentation can be found at <http://www.cisco.com/go/saa/>



Corporate Headquarters  
Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, CA 95134-1706  
USA  
[www.cisco.com](http://www.cisco.com)  
Tel: 408 526-4000  
800 553-NETS (6387)  
Fax: 408 526-4100

European Headquarters  
Cisco Systems International BV  
Haarlerbergpark  
Haarlerbergweg 13-19  
1101 CH Amsterdam  
The Netherlands  
[www-europe.cisco.com](http://www-europe.cisco.com)  
Tel: 31 0 20 357 1000  
Fax: 31 0 20 357 1100

Americas Headquarters  
Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, CA 95134-1706  
USA  
[www.cisco.com](http://www.cisco.com)  
Tel: 408 526-7660  
Fax: 408 527-0883

Asia Pacific Headquarters  
Cisco Systems, Inc.  
Capital Tower  
168 Robinson Road  
#22-01 to #29-01  
Singapore 068912  
[www.cisco.com](http://www.cisco.com)  
Tel: +65 317 7777  
Fax: +65 317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the **Cisco Web site at [www.cisco.com/go/offices](http://www.cisco.com/go/offices)**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia  
Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland  
Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland  
Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden  
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

All contents are Copyright © 1992–2002, Cisco Systems, Inc. All rights reserved. Cisco, Cisco IOS, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company.  
(0206R) 202822.C/ETMG 9/02