



*The Essential Guide to*  
**VoIP Implementation and Management**



*Written by:*  
*John Q. Walker and Jeffrey T. Hicks*

# The Essential Guide to VoIP Implementation and Management

by John Q. Walker and Jeffrey T. Hicks  
*NetIQ Corporation*

## Chapter Six: Establishing VoIP SLAs

<b>Determining What to Measure in a VoIP SLA</b>	<b>3</b>
Availability.....	3
Call Setup Performance.....	4
Call Quality.....	5
Incident Tracking.....	6
<b>Implementing VoIP SLAs</b>	<b>7</b>
Define Responsibilities.....	7
Identify Service Levels for VoIP and Other Applications.....	8
Negotiate the Service-Level Agreement.....	9
Deploy Tools to Measure SLAs.....	12
Manage SLA Compliance and Enforcement.....	14
<b>Conclusion</b>	<b>16</b>
<b>References</b>	<b>16</b>
About the Authors.....	17

# Chapter Six: Establishing VoIP SLAs

Chapter Six of our book, “*The Essential Guide to VoIP Implementation and Management*,” by John Q. Walker and Jeffrey T. Hicks of NetIQ Corporation, explains the SLAs that are needed for a VoIP implementation.

We’re serializing this book, releasing a chapter a month for seven months. A revised, bound edition is to be published in autumn 2002 by McGraw-Hill. We appreciate your comments on how to improve it. We’ll acknowledge the feedback we use in our printed book.

If you’re new to VoIP, we recommend downloading the previous chapters and reading them first.

- Our Introductory Chapter introduces the terminology and acronyms that recur frequently in the telephony and data networking communities – a set of ideas essential to a good understanding of VoIP.
- Chapter 1 explains how to build a business case for VoIP – the first step in the planning process of a VoIP deployment.
- Chapter 2 details a step-by-step process for deploying VoIP successfully using classic IT project management: plan thoroughly, evaluate and experiment with the designs and devices you plan to use, and then roll out the deployment incrementally, building on your successes and learning as you proceed.
- Chapter 3 explains how to decide whether to outsource VoIP implementation and management or do it yourself.
- Chapter 4 discusses quality of service (QoS) and tuning techniques that can help improve the network characteristics important for call quality.
- Chapter 5 discusses ongoing VoIP management, a crucial part of ensuring the continued success of a VoIP deployment.

This chapter discusses service-level agreements (SLAs) for VoIP. SLAs have become a common way of defining how network performance will be measured and guaranteed by service providers. We introduced the concept of SLA management in the last chapter as part of network performance management. In this chapter, we’ll go into more detail and discuss some typical SLAs for VoIP, the metrics these SLAs are based on, and how to implement a VoIP SLA – within your organization or with a service provider.

An SLA is exactly what the name implies – an agreement between two parties about the delivery of a certain level of service. Why should you consider implementing an SLA? Because you’re depending on a service provider, such as an ISP, to deliver a service, such as WAN access, that meets certain quality expectations. And when the delivered service doesn’t meet those expectations, it costs you – time and money. An SLA can help you recover some of the cost incurred during a period of degraded or unavailable service. Equally, an SLA can represent a promise you make to VoIP phone users for a guaranteed level of quality and availability. But, in either case, to create an effective SLA, you need to consider what you want to measure when evaluating the service and what expectations you and your network’s users have for the network service provided.

# Determining What to Measure in a VoIP SLA

In a VoIP deployment, you want your phone users to be as happy with the level of service being delivered as they were with the PSTN. Service-level agreements provide a target for the actual performance your VoIP system delivers. At its root, an SLA is quite a simple matter: you define performance and availability goals and then monitor the system to see how you're doing. But, just as a VoIP SLA can include a huge variety of performance metrics to be monitored, so can it be defined in many different ways.

What kind of SLA is best for you? It depends. First of all, do you need an internal or external SLA, or both? Let's start by defining internal and external SLAs:

**Internal SLA:** An SLA within your enterprise. It typically describes the expectations between users of the system and another internal organization, such as your IT group.

**External SLA:** An SLA between your company or organization and a third-party service provider. The service provider may be an ISP, carrier, or other VoIP outsourcer.

The type of SLA you need is determined by business needs and user expectations. You may end up with both types of SLAs, or you may only develop one. However, a key component of both types of SLAs is the metrics that will be measured. The metrics you emphasize should be based on your overall business goals.

We've categorized the key SLA metrics into four main groups – Availability, Call setup Performance, Call Quality, and Incident Tracking. Each of these categories contains several sub-metrics. Let's discuss each group of metrics in more detail.

## Availability

As discussed in previous chapters, today's PSTN users expect to hear a dial tone 99.999% of the time when they pick up a phone. This expectation makes availability important in any VoIP SLA. But, what exactly does the term availability include? To begin with, you want to hear a dial tone when you pick up the phone; to state the case simply, **no dial tone** equates to being unavailable. Next, when you dial a phone number, you expect the call to go through; **no ringing** or **no busy signal** equates to being unavailable. Finally, while you're talking, you expect to be able to complete the call without being disconnected; **abnormal call termination** equates to being unavailable.

In addition to directly measuring downtime, here are some of the underlying sub-metrics to monitor and include in an SLA that tracks availability:

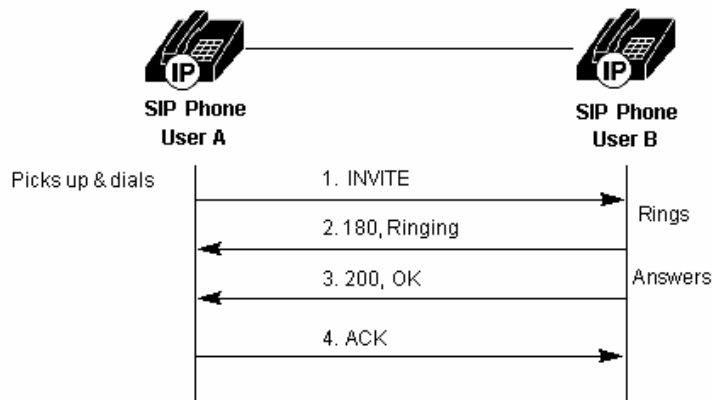
- **IP PBX availability.** Is the IP PBX (or equivalent server) active and functioning properly? If not, how long was it down due to hardware, software, or network problems? Was the IP PBX unavailable while patches or updates were applied? If multiple IP PBXs are involved, this metric captures the availability statistics of all of them.
- **Network availability.** VoIP calls depend on the IP network. Is network connectivity available? If not, how long has the outage persisted? Was the network unavailable during router updates or configuration changes? You probably already have some SLAs in place for network availability. Consider updating them for VoIP by making them more stringent.

- **Network service availability.** Many IP phones and VoIP servers rely on critical network services to perform call routing. Are the DNS and DHCP servers available? Are these network services running? If DNS and DHCP are unavailable, users may not be able to make any calls.
- **Call completion percentage.** This metric is sometimes referred to as the “answer seizure rate” in the telephony community. It represents the percentage of attempted calls that were successfully completed. A low call completion percentage generally points to declining availability.
- **Abnormal disconnections.** These refer to calls that were not ended by one of the talkers. How many calls that were in progress were abnormally terminated? Where did these disconnections occur? A high number of dropped calls may point to declining availability.
- **Line busy.** How many times did a caller receive a busy signal? Was the line really busy, or was the busy signal an indicator of oversubscribed gateway ports? The key metric here is the number of busy call attempts caused by oversubscription of resources.

Because they’re so basic to the health of your entire system – and to users’ satisfaction levels – availability metrics are the first group of statistics to put in a VoIP SLA. Next, you need to think about what happens when availability is good, but performance problems plague call setup.

## Call Setup Performance

When you make a VoIP call, a complex series of events has to occur in sequence, and without errors. The first set of events, the call setup phase, takes care of getting a dial tone, dialing the phone number, and getting a result, either ringing at the desired location or a busy signal. Several different protocols are used for call setup in various VoIP implementations, and all could experience poor performance. Call setup protocols such as H.323, SIP, MGCP, and Megaco operate principally using the TCP protocol, sending a large number of different flows between the IP phones and VoIP server to establish a call between two parties. Figure 1 shows an example from SIP.



**Figure 1.** Call setup can involve many network flows. This is the simplest example of call setup using SIP between two IP phones. [1]

Poor call setup performance can affect the user's initial perception of the call. If the normal sequence of tones and responses is not provided quickly enough, users may get impatient and hang up. Call setup can be divided into two subphases:

- **Dial tone response time.** How much time elapses from the moment you pick up the phone until you hear a dial tone? This time is the average dial tone response time. If the delay is long enough, users may think the system is unavailable. Today, most IP phones will generate a dial tone almost instantaneously, so this metric may not be a major issue for your systems. However, it's still vitally important. A good upper bound to use is 2 to 3 seconds. Any additional delay would not be acceptable to most users.
- **Call setup response time.** Sometimes referred to as "post dial delay," this metric applies to the amount of delay between the time you dial the phone number and the time you hear ringing or get a busy signal. This is the average call setup response time. How much delay is too much for this metric? It depends on the users and their expectations. But if the delay becomes too long, some users may get frustrated and hang up, thinking the call cannot go through. A good upper bound for call setup response time is 2.5 seconds [2]. This matches the average call setup response time in the PSTN.

You should also pay close attention to call setup response times for calls between the VoIP system and the PSTN. There's extra work involved in routing IP calls over analog lines, and gateway signaling protocols, such as MGCP and Megaco, are added to the mix. Translations between signaling protocols at the gateway may add additional delays to the call setup time.

Call setup metrics must be considered in any VoIP SLA. Next, you need to take into account what happens when availability is good and call setup completes quickly, but the call quality is poor.

## Call Quality

Users have well-established expectations for a VoIP system: it ought to sound as good as the PSTN. As a result, call quality is a key component of a VoIP SLA.

The Mean Opinion Score (MOS) is the standard metric for user perception of call quality. The SLA for all calls should be drafted in terms of the MOS scale, from 1.0 to 5.0. A MOS of 4.0 or higher is considered toll quality or equivalent to the PSTN. We consider that a MOS of 4.0 is good, 3.6 and above is acceptable, and anything below 3.6 should not be considered acceptable for business-quality calls.

USER SATISFACTION	MOS
Very Satisfied	4.4
Satisfied	4.3
Some Users Dissatisfied	4.0
Many Users Dissatisfied	3.6
Nearly All Users Dissatisfied	3.1
Not Recommended	2.6
	1.0

**Figure 2. User satisfaction shown on a MOS scale.**

When you incorporate the MOS into your SLA, you also need tools that can monitor the performance of calls on your network and calculate a MOS from what is measured. In our second chapter, “Planning for VoIP,” we discussed the network performance metrics that make up the MOS:

- **Delay.** VoIP traffic is intolerant of excessive delay. Long delays can make phone calls sound like walkie-talkie conversations. SLAs for delay are usually specified as a maximum allowable in milliseconds, for packet delivery. 150 ms is a widely accepted limit for end-to-end delay; if your delay exceeds 150 ms, it is likely that the MOS will decline.
- **Jitter.** Variations in packet arrival times can cause packets to be discarded and VoIP call quality to suffer. SLAs for jitter are usually specified in maximum allowable milliseconds of variability in delay among packets transmitted from the same source. 50 ms is a good number for the maximum allowable jitter.
- **Lost packets.** If lost, VoIP packets are not retransmitted. Lost packets thus result in clipped syllables or even missing words in a call. SLAs for lost packets are usually specified as a maximum allowable percentage loss of all packets sent. A packet loss percentage of less than 0.50% is recommended for good-quality VoIP.

A fourth component of the MOS, the codec, is usually a fixed parameter per call – so it isn’t something that is measured in real time. Chapter 2 includes a table that shows standard quality impairments for each codec.

There’s no need to spell out separate SLA metrics for each of these measurements; the call-quality portion of your SLA should specify the MOS. However, you may choose to set up monitoring thresholds for these metrics, to trigger early warnings that quality is declining. For example, you may set a threshold that is triggered whenever end-to-end delay rises above 150 ms. Although the measured MOS may not yet have declined, a delay above 150 ms signals a potential reduction in call quality.

## Incident Tracking

If any of the availability, call setup, or call quality SLA metrics decline, you’ll want to determine why. Whenever an SLA metric deteriorates and crosses a threshold, the incident-tracking metrics come into play. Incidents often occur because of an outage or severe degradation that needs to be repaired. Scheduled changes and ongoing maintenance are also

tracked as incidents. Several metrics are usually part of an SLA to deal solely with incident tracking:

- **Mean Time to Repair (MTTR).** When an SLA value is violated, how long does it take for the provider to fix a problem, make an upgrade, or perform required maintenance? The time it takes to resolve each incident is averaged to get the MTTR. The MTTR is probably the most common SLA metric that deals with incident tracking. For this metric, lower numbers are better.
- **Mean Time Between Failures (MTBF).** Do failures occur frequently or only rarely? This metric defines the average time between failure incidents. The MTBF gives an indication of how often failures are occurring and can help identify potential availability problems. Proper monitoring of the MTBF can also inform you if failures are occurring more often. For this metric, higher numbers are better.

When tracking downtime or low call quality, set thresholds proportional to the SLA metrics and configure events or alerts to be sent when the thresholds are crossed. Tie the events into your fault management and event response systems. Early warnings may prevent SLA violations for the crucial VoIP network metrics and help you avoid triggering violations of incident tracking metrics.

A good VoIP SLA should include incident-tracking metrics to give you an expectation for how rapidly your service provider will respond when service levels are not being met. In turn, incident-tracking metrics help to guarantee the high availability, call setup performance, and call quality already specified in the SLA.

## Implementing VoIP SLAs

Armed with a good understanding of the SLA metrics important for VoIP, you can begin the process of actually implementing the right SLAs for your enterprise. Like many topics we've covered in this book, we view this as a staged process. Working through the five stages discussed below will bring you the results you're seeking – a trouble-free VoIP system, transparent to your end users.

The first stage is defining who's responsible for each role in the SLA implementation process. Next is identifying the right VoIP service levels for your enterprise. The service-level agreement itself then needs to be negotiated. After that, measurements can begin, to determine when the SLA is being met and when it isn't. You'd like this stage to be automatic, so you'll choose among the tools available to make this painless. The last stage is managing and enforcing the SLA.

### Define Responsibilities

The implementation of a service-level agreement should be segmented into close to a dozen roles. Most of the following apply more readily to external SLAs – those agreed to with a third-party provider. However, each implicit task must also be completed by someone when an internal SLA is being developed, and must also remain someone's responsibility once the SLA is in force.

- Who defines the SLA? Who decides which metrics are important for the organization?

- Who writes the contract and guides it through the negotiations? Who determines the penalties?
- Who manages the network, its equipment, and the related computer hardware and software? These are the items against which the service covered by the SLA is being measured. Who's responsible for maintaining them?
- Who takes the measurements for the metrics specified in the SLA? Who assures the quality of the measurements as they're taken?
- Who manages the SLA? Who specifies the thresholds for the metrics and gets notified when they've been crossed?
- Who responds to and resolves incidents when a managed threshold or an SLA metric is crossed?
- Who does the accounting for the SLA? Who measures the percentage of compliance, and presents it to the provider or service recipient at the end of each week or month?
- Who enforces compliance with the SLA? This involves determining penalties and collecting them.
- Who decides when it's time to get a new provider, when compliance with the SLA becomes an issue?

In a small organization, many of these roles may be handled by the same person. But, in larger organizations, these responsibilities are probably divided among several people who need to communicate well with one another.

## Identify Service Levels for VoIP and Other Applications

Before beginning your SLA contract negotiations, you need to determine what you're going to measure and what the target SLA values should be. Although the VoIP SLA metrics discussed in "Determining What to Measure in a VoIP SLA" on page 3 are all important in varying degrees, your own SLA should comprise the metrics that are most meaningful to your business. A good piece of advice to take to heart is that you must "[c]reate SLAs that are based on the quality of end-user application experience, rather than IT metrics which customers cannot or do not want to understand." [3] VoIP is an example of a network application whose value is driven nearly 100% by the perceptions – positive or negative – of ordinary users.

For VoIP, the relevant metrics for an SLA are the metrics discussed above: availability, call quality, call setup performance, and incident resolution time. The first three are end-user, end-to-end measurements – they're what counts. You could, alternatively, measure their lower-layer constituents only, such as delay and packet loss, but it's best to consider them principally for diagnostic purposes, to reduce the "find and fix" time. You should also include measurements related to problem repair, like mean time to repair and incident management: how problems are identified, submitted, and passed among the team.

Now here's another consideration. Having added VoIP to the network, do you now need to add SLAs for your other business-critical applications, which you may not have been monitoring before? These include e-mail, groupware, e-commerce, and industry-specific business programs. Maybe you should have had a response time SLA for your SAP applications, but you weren't aware of much dissatisfaction before you deployed VoIP. Our advice is this: adding VoIP traffic to a system near its capacity may significantly increase the response time

of other business-critical applications. Application performance is something you now really need to pay attention to.

To establish the SLA targets for your most important applications, begin by establishing performance baselines. These let you know what's actually possible and provide a starting point. You know that at the time you take the baseline, here is where MOS, response time, or throughput stands. You obviously shouldn't write an SLA to support 100 VoIP calls with a minimum MOS of 3.9, if there's clearly not enough bandwidth or other resources to support 100 calls with that quality.

Start your VoIP call quality baseline with the results provided by the VoIP-readiness assessment we discussed in chapter 2. Obviously you need to work from the "last" assessment – the one where the MOS met your standards and was conducted after you'd done all the necessary upgrades and eliminated all the bottlenecks or other problems that were identified. It's similarly straightforward to get baselines for the network performance of other applications, such as response time, throughput, or packet loss. Don't go overboard; be sure to measure what's important to end users for each application. For example, for a database application, focus on the response time for queries or updates, rather than on throughput.

Similarly, create your availability baseline using the best numbers you have to describe your current availability statistics. Without making any significant changes to the network or the users, this gives you a place to start, a place where users know what to expect.

If you want to monitor application response time, avoid insisting that all requests be met with a response time of 1 second or less. That's unrealistically strict. Instead, it would be better to state that 95 percent of requests must have no more than a 1-second response time and 5 percent may have a response time of between 1 second and 5 seconds.

Starting from these baselines – the expected and observable behavior – create some SLA targets. SLA targets are values representing performance that's so bad, it's no longer acceptable. For example, if your VoIP baseline MOS is 4.15 today, you might create an SLA target that reads something like this:

MOS of 4.0 or above 85% of the time; 3.9 or above 95% of the time;  
and 3.8 or above 100% of the time; measured on 10 concurrent calls  
with the G.711 codec, between Raleigh and Houston.

## Negotiate the Service-Level Agreement

The intention of an SLA is to spell out which services are to be provided, how the services will operate, and what should happen if their performance doesn't meet the expected service levels. But, a certain amount of negotiation, compromise, and perhaps even controversy will undoubtedly enter into any SLA you implement. A recent article noted that "SLAs are like insurance policies. Just like life insurance doesn't guarantee life, SLAs don't guarantee levels of service. They provide you with compensation in case something goes wrong." [4]

Because we anticipate some give-and-take in the relationships affected by a VoIP SLA, we've drawn up a top ten list, topics to be addressed in your VoIP SLA negotiations:

### 1. Specify the SLA metrics and their target values.

The section "Identify Service Levels for VoIP and Other Applications" on page 8 describes this topic in detail. The measurements that affect the end-user experience are important to your organization and should be included in your VoIP SLA: availability,

call quality, and call setup performance. You may also want to include metrics for other applications so that their performance doesn't degrade because of the addition of VoIP, as well as a metric for incident rates and their rates of resolution.

## **2. Describe how the SLA metrics are measured and who measures them.**

The section "Define Responsibilities" on page 7 identifies the wide range of roles and responsibilities associated with creating and enforcing a VoIP SLA. Do you take SLA compliance measurements in your organization, or are they taken by a third party or the service provider? If the provider does the measurements, how do you, the customer, verify them? Tools for taking measurements are described in the next section, "Deploy Tools to Measure SLAs" on page 12.

The SLA should describe in detail how the measurements are to be taken. It should specify the locations to be monitored. And, the SLA should spell out how measurements and compliance should be handled if an end-to-end metric involves multiple ISPs.

It should also explicitly describe what time periods are covered. We noticed the following quotation in an ISP offering brochure: "A high-end VoIP carrier will offer 99.99% availability, which does not include scheduled maintenance windows where the carrier may take down the network to upgrade equipment; clean or switch fibres [sic] *or perform any other work that could lead to network downtime.*" [5, our italics] Wow! A lot of time may elapse in these periods that are not included in the availability agreement; what time periods are covered in your SLAs and how are they measured?

## **3. Describe the SLA reports and their schedule.**

Your service provider should demonstrate its compliance with the SLA by sending you monthly or even weekly reports. In your negotiations, make sure the contract specifies what metrics and what parts of the network will be included in the reports. It should also say how often you'll get reports.

## **4. Allow requests to review SLA compliance information on demand.**

The SLA should establish a procedure for requesting SLA compliance information on demand. This can be helpful for troubleshooting. For example, if you're experiencing a delay problem, information from your ISP may help you narrow the problem down to a WAN link that you don't control.

## **5. Specify the turnaround times for change requests, by severity.**

You'll make changes as you gain more experience with your VoIP system. For example, as you add new locations and new users, you may want to add more locations to be monitored as part of the SLA. This may require a change request to your service provider. What's the expected turnaround time for the change request? It's also reasonable to include a prioritization scheme in the SLA's timetable for such requests. A slow or over-burdened link may be one of your highest-severity items and should be expedited accordingly.

## **6. Specific support-staff and help-desk levels.**

The last thing you want is to be placed on hold indefinitely when there's an SLA-related fire to put out. Get it in writing: How many people are available to support your VoIP system when incidents occur? What is their skill level? What hours do they work?

## **7. Schedule periodic reviews and adjustments to contract provisions.**

Your initial VoIP deployment will no doubt change over time: new users, new locations, new applications, new hardware, more bandwidth, mergers, and so on. These may cause your SLA requirements to change. Don't let them get too far out of date. Schedule regular reviews with your service provider.

## **8. Describe the rewards for great compliance and penalties for non-compliance.**

What penalties should you build into your SLA? And if an unsatisfactory situation drags on, how long do the penalties build up before you call it quits?

CommWeb.com points out that "It's easy for service providers to promise 99.999 percent uptime – especially when the penalty for not delivering is a meager day or two worth of credit. Obviously, penalties of this sort are no compensation for the potential loss in revenue when a company's ... critical applications aren't performing." [6] Your provider must have a strong motive for complying with the SLA you've negotiated. That motive may be either positive (a bonus or additional business) or negative (a substantial monetary penalty).

Your best safeguard when entering into an SLA is a "system of rewards and penalties for compliance and non-compliance." As Mandy Andress of *InfoWorld* notes, "An unenforceable SLA serves little purpose. It is well and good to say that all requests should have a 1-second response time, but if the group responsible for system performance does not incur any penalties for slower response times or reap any rewards for faster response times, then they have no real incentive to comply." [7]

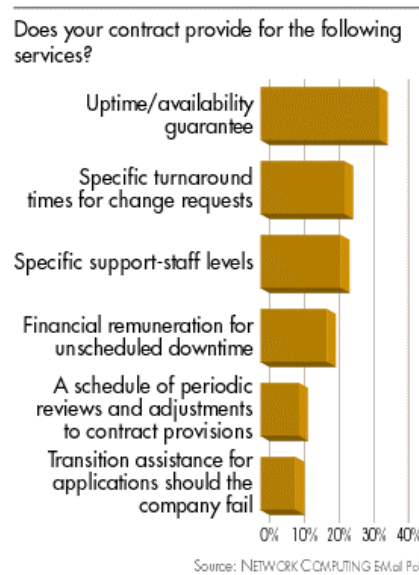
## **9. Discuss transition assistance for services should the service provider fail or suffer a setback.**

Put together a plan that gets you through the lurch if something catastrophic happens with your service provider. This type of situation has unfortunately become more common in recent years. Aside from bankruptcies, service providers face the same scary threats that you do; floods, tornados, malicious attacks by network intruders, and other unpleasant possibilities need to be anticipated and planned for.

## **10. Create a procedure for terminating an SLA contract.**

If you'll pardon the analogy here... Sometimes it just doesn't work out, and an amicable divorce makes sense. Write the pre-nuptial agreement before the marriage, not after it's started to go bad.

Is it reasonable to expect your service provider to agree to all of the sorts of stipulations we've just outlined? Here are the results of a recent survey in *Network Computing*, asking service providers and outsourcers what is covered in their SLA contracts.



**Figure 3. Survey results, showing provisions covered in SLA contracts. [7]**

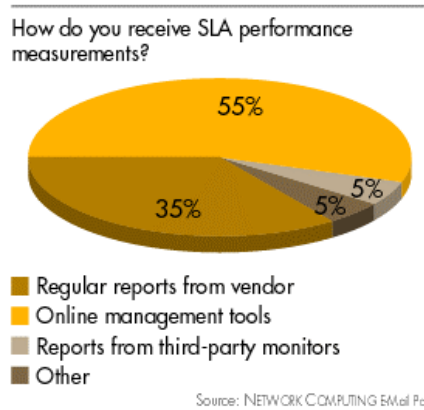
We believe that most boilerplate SLA contracts are probably not good enough for you. They can have lots of holes and exceptions. For example, if a carrier’s sub-carrier goes down, are they still responsible? Depending on the size of your deployment and its geographical scope, there may be a chain of sub-carriers and sub-components to take into account; determine who’s ultimately responsible. Make sure you fully negotiate the contract details with them.

Additionally, consider letting SLA contract quality guide your choice of service provider. You’re now armed with a top ten list of things to include in the negotiations. “If you are choosing between two otherwise-equivalent service providers, if one has a better SLA does that make a difference? And is that more important than past brand experience, than price?” [9] A 2002 survey of enterprises with SLAs found “that not only were SLAs important, but the enterprises were willing to pay a significant premium for verified quality and guaranteed service.” [9]

## Deploy Tools to Measure SLAs

Having deployed your VoIP system, determined the expected performance, and negotiated your SLA contract, you now need to watch the metrics specified in your contract. This means that the performance values are monitored on an ongoing basis, and events are triggered when the target SLA value is about to be crossed.

The monitoring can be done by the service provider, by your team in your enterprise, by a third party, or some combination of these. In any case, the provider will surely be motivated to allow for some comprehensive monitoring of their offerings. And for you, monitoring is even more important. “Enterprises want to outsource their network service offerings and have proof that they’re getting what they’re paying for,” notes Laura Spear, VP of marketing at Trinagy. “That means they need tools to provide proof and credibility back to their customers.” [11]

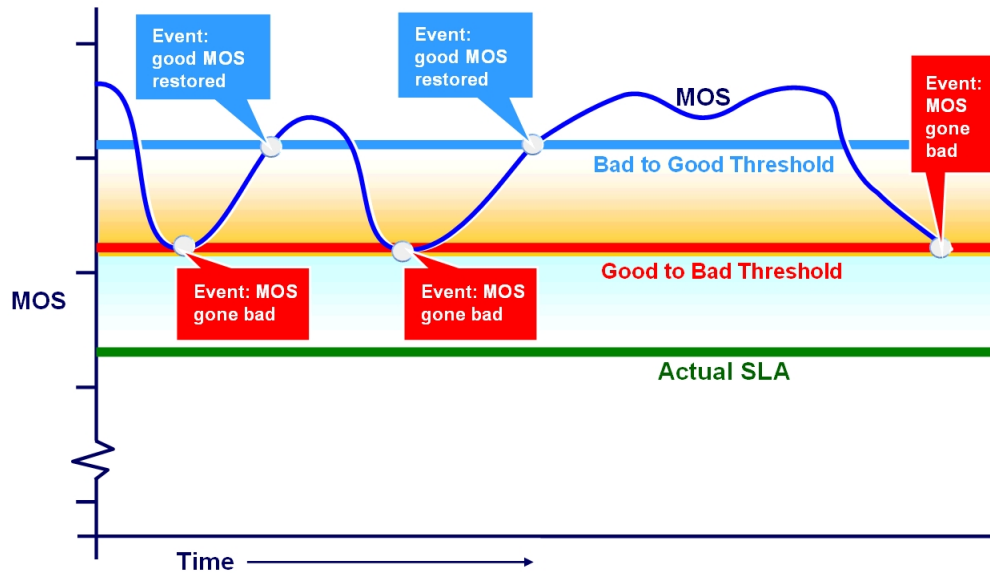


**Figure 4. Survey results, showing how SLA performance measurements are received. [7]**

One reason to perform consistent SLA monitoring is to check SLA compliance. A more important reason, though, is to avoid SLA infractions altogether. This means determining how much early warning you need to deal with problems. While you'd like to know what's going on at any given moment, the closer you get to "real-time monitoring," the greater the amount of data collected and the amount of traffic generated in reporting it. A better method is to set useful thresholds.

For a given SLA measurement, set a pair of thresholds that are stricter than the SLA target. When the second threshold is triggered, take immediate action so that the SLA level is not reached. You want to force action to be taken before an SLA violation, not when the SLA metric has been crossed and it's too late. We call this the "good-to-bad" threshold; when it's crossed, initiate the incident/fault management processes we discussed in previous chapters.

So when setting thresholds, you should set two threshold crossings: crossing on the way down (going from good to bad) and crossing on the way back up again (going from bad to good – indicating that the incident has been resolved). And make sure you allow for some gap between these threshold-crossing values – you don't want a flurry of alarms to besiege your e-mail inbox if the value you're measuring is fluctuating back and forth across this boundary.



**Figure 5. An example of a fixed SLA target for the MOS, along with a pair of thresholds preceding it. Great MOS is at the top, poor MOS is at the bottom.**

Reading up from the bottom of Figure 5, we first encounter the SLA target; if the measured MOS crosses below the line, an SLA violation occurs. Above that is the threshold where, as the MOS declines, you decide that it's gone from good to bad, and you trigger the event and actions necessary to avoid a further decline. You'd like to reset that event when the problem is truly fixed, so the top line is the threshold that's crossed on the way back up; as the MOS improves after having crossed below the "Good to Bad" threshold, it can now be declared good again when it crosses above the "Bad to Good" threshold.

As you get more sophisticated in working with your SLA thresholds, you may consider implementing thresholds that aren't just fixed lines or numerical targets. Thresholds can be intelligent, responding to changes in overall behavior, the time of day, or the number of users.

A wide range of SLA monitoring tools is available to help you monitor your VoIP SLAs. Sterling Research summarizes the different choices available to you: "Some companies offer a limited tool set because they have decided to focus primarily on the monitoring aspect of SLA compliance. ... Others offer a range of tools that treat SLA compliance as an end-to-end process. One tool is initially used to establish baseline performance for a particular service; the next is used to monitor the service on a day-to-day basis; and – finally – simulation tools are used to spin what-if scenarios that calculate the impact on service performance if changes are made to the environment." [10]

## Manage SLA Compliance and Enforcement

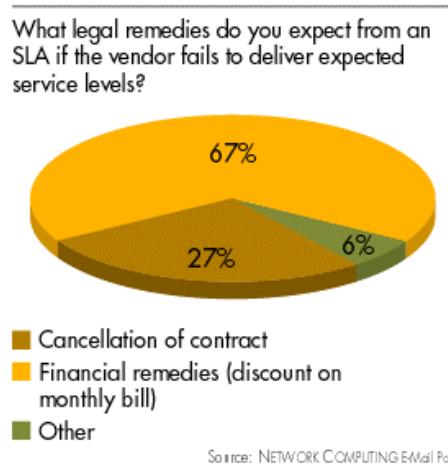
So, suppose it's the end of the month. You review your VoIP SLA reports, and see that one of the SLAs has been violated, that is, too much time has been spent outside the SLA target. What's the transaction that now needs to occur between you and the service provider?

First, don't get in this situation. Overcommunicate with the team fulfilling the SLA responsibilities we described on page 7. All members of the team should be well informed

all along the way. You don't ever really want to get into the enforcement or penalty stage of an SLA contract. Avoiding disputes and legal actions altogether is almost always cheaper and less stressful than pursuing them.

Jared Huizenga of Sage Research believes that enterprises are currently trying to develop “a more proactive SLA” with their providers. [11] In a “proactive” SLA, providers “have to spot, correct, and recompense customers for any problems before customers inform the service providers of the problems.” [11] Huizenga also believes that most enterprises that enter into an SLA with their providers “want to be able to actually monitor” compliance “at their own site” and “receive an automatic credit” if a compliance issue arises.

As discussed above, your SLA contract should establish a system of rewards and penalties for compliance. These are the incentives for the SLA provider to perform well. Rewards for excellent SLA compliance may include things like cash bonuses. The penalties for SLA infractions include automatic credit or reimbursement of your charges, withholding of payment, or cancellation of the contract. Penalties must be stiff enough to have real meaning for a larger provider. A *Network Computing* survey showed that 67% of respondents expected “financial remedies” from their provider if the SLA was breached. [7]



**Figure 6. Survey results, showing expected legal remedies for SLA non-compliance. [7]**

When expectations are not met, any changes for the better should come from the provider. They need to determine how to improve their quality and their processes so expectations are consistently met in the future.

SLAs should be reviewed regularly. An annual review is specified in many contracts. Because of the rapid pace of technology development, user expectations may change frequently – this is especially true of expectations for availability and the response time of business transactions. This means that SLAs must be periodically updated to reflect these changes. Otherwise, SLAs can quickly become outdated, often demanding service levels far below existing technological capabilities.

Contract cancellation may be the most effective penalty to levy in cases of SLA non-compliance. David Kaufman, of Brix Networks, argues that the proactive testing and monitoring with SLA thresholds we described above is appealing to service providers for that very reason. “Having the advance warning that something is beginning to go wrong

with a[n] SLA is really vital because, if there's an SLA outage, you have a one-in-three chance that you've lost that customer," he says. [11]

## Conclusion

This wraps up our softcopy eBook, available on the NetIQ Web site. We have covered a lot of information that we think is "essential" when deploying and managing VoIP. We hope that the information has been useful to you.

As we mentioned in the opening, a revised, bound edition of this book is to be published in autumn 2002 by McGraw-Hill. We appreciate your comments on how to improve it. We'll acknowledge the feedback we use in our printed book.

## References

1. This picture taken from [www.vovida.org/document/Vocal\\_Technology\\_Overview/Technology\\_Overview7a.html](http://www.vovida.org/document/Vocal_Technology_Overview/Technology_Overview7a.html).
2. This value was taken from a discussion in [www.att.com/network/standrd.html#call](http://www.att.com/network/standrd.html#call).
3. "Policy-Based SLA Compliance for Hosted Applications: Identifying the Weakest Link in Next Generation Services," Response Networks brochure, [www.cisco.com/warp/public/765/emea/aip\\_asp/documents/Pulsar\\_xSP.pdf](http://www.cisco.com/warp/public/765/emea/aip_asp/documents/Pulsar_xSP.pdf).
4. "Strengthen service-level agreements," Shally Bansal Stanley, *Network World*, May 6, 2002, [www.nwfusion.com/careers/2002/0506man.html](http://www.nwfusion.com/careers/2002/0506man.html).
5. "VoIP – The Gateway for Smarter Communications," a Light Paper by KPNQwest, [www.kpnqwest.com/links/IP\\_Voice\\_Light\\_Paper.pdf](http://www.kpnqwest.com/links/IP_Voice_Light_Paper.pdf).
6. "The Price of 99.999 Uptime," *CommWeb.com*, June 1, 2001, [www.commweb.com/article/COM20010601S0008](http://www.commweb.com/article/COM20010601S0008).
7. "Service Providers & Outsourcing," *Network Computing*, September 3, 2001, [www.networkcomputing.com/1218/1218f15.html](http://www.networkcomputing.com/1218/1218f15.html).
8. "Internal SLAs benefit the entire company," Mandy Address, *InfoWorld*, April 27, 2001, [www.infoworld.com/articles/tc/xml/01/04/30/010430tcintersla.xml](http://www.infoworld.com/articles/tc/xml/01/04/30/010430tcintersla.xml).
9. "Business-to-Business Research," reprinted by Sage Research with permission of *Quirk's Marketing Research Review*, April 2002, [www.sageresearch.com/Quirks.pdf](http://www.sageresearch.com/Quirks.pdf).
10. "Enterprise-Class Service Level Agreements," Samuel M. Alunni, *Sterling Research Market View*, July 1, 1998, [www.sterlingresearch.com/library/library/07\\_98slas.html](http://www.sterlingresearch.com/library/library/07_98slas.html).
11. "No Downtime," Chris Garifo, *XChange*, May 15, 2001, [www.xchangemag.com/articles/152back.html](http://www.xchangemag.com/articles/152back.html).

## About the Authors

John Q. Walker is a director at NetIQ Corporation. He is one of the founders of Ganymede Software, heading the software development team there before it joined NetIQ in spring 2000. In earlier jobs, he managed teams responsible for designing and developing high-speed networking at IBM. Another job stint involved inventing ways to break security on a new computer system, reporting his methods to the development team. He co-authored a book on portable network programming for McGraw-Hill. Dr. Walker holds a Ph.D. in software engineering; his Masters degree focused on system testing. His e-mail address is [johnq@netiq.com](mailto:johnq@netiq.com).

Jeffrey T. Hicks is a technical lead at NetIQ Corporation. He has recently led the development teams for the award-winning Chariot and Chariot VoIP Assessor products. He has been active in the development of VoIP testing and management solutions for the past four years. In earlier jobs, he helped develop innovative network communications software products at IBM. Mr. Hicks holds a Masters of Engineering from North Carolina State University and a B.S. in Computer Engineering from Auburn University. His e-mail address is [jeff.hicks@netiq.com](mailto:jeff.hicks@netiq.com).