Network Reliability and Interoperability Council VI
Focus Group 3

Network Interoperability

Final Report

November 2003

Table of Contents

1. Executive Summary

The telecommunications industry of the United States is undergoing a fundamental technology shift. Traditional Public Switched Telephone Network (PSTN) circuit-switched networks are converging with Internet Protocol (IP) packet-switched networks. This is occurring across the spectrum of wireline to wireless media. This convergence is due to a number of factors:

- The economics of providing telephony and data across a common underlying packet-switched networking infrastructure is increasingly compelling to the industry and consumers.
- Consumers, service providers, network operators, original equipment manufacturers, and independent software vendors see the possibilities of new or enhanced services and features in this convergence.
- The U.S. government and regulatory bodies desire to provide all users with seamless and transparent interoperability and access between and across circuit- and packet-switched networks.

The Network Reliability and Interoperability Council (NRIC VI) Focus Group 3 is chartered to:

> "… prepare analyses and, where appropriate, make recommendations for improving interoperability among networks to achieve the objectives that are contained in Section 256 of the Telecommunications Act of 1996, with particular emphasis on ensuring 'the ability of users and information providers to seamlessly and transparently transmit and receive information between and across telecommunications networks.'"

The recommendations and best practices included in this report address the interoperability of Voice over Internet Protocol (VoIP) and the Public Switched Telephone Network (PSTN). The purpose of this report is to inventory existing and in-process standards and industry best practices against a set of basic telephony features and functions to determine:

- Existing and in-work standards that address interoperability.
- Gaps in standards and best practices that standards bodies or industry are recommended to address to achieve full VoIP-PSTN interoperability.
- Industry best practices that have been identified.

The scope of these recommendations and best practices are VoIP-to-VoIP and VoIP-to-PSTN calls between service providers.

Because FG3 is addressing interoperability "between and across telecommunications networks," these recommendations and best practices do not address interoperability or protocols within a service provider's network, VoIP end devices, nonconsumer voice features (e.g., Centrex), or emerging transport

technologies such as Voice over Asynchronous Transfer Mode (VoATM) or Voice over Multiprotocol Label Switching (VoMPLS) networks.

To arrive at a set of recommendations, the focus group drafted an interrelationship diagram of the edge components included in service providers' VoIP and PSTN networks (see section 2.2, figure 2). By mapping these relationships, the focus group was able to identify gaps and overlaps in standards activities and industry best practices that, through experience, members have found necessary to achieve full, seamless, and transparent access across circuit- and packet-switched networks for voice services. The intent of this report is to bring attention to issues such as these.

The interoperability topics addressed within this report are

- Signaling architectures.
- Call control architectures.
- Voice over wireless.
- Inter-provider interfaces.
- Directory services.
- Safety and security features.

Areas of Attention

There are several significant challenges to interoperability. The most significant is the overlap in standards for VoIP. Two sets of standards bodies have been developing signaling protocol specifications that perform similar functions, but do not directly interoperate. Specifically, the ITU-T first developed the H.323 set of VoIP standards based largely on Integrated Services Digital Network (ISDN) while the IETF has developed a set of standards based on Session Initiation Protocol (SIP). The ITU approach is network-based, while the IETF approach is end-system based. In some cases, this fundamental variation in approaches creates significant interoperability challenges. Ultimately, either every service provider will need to support both sets of standards or the industry will eventually pick one interoperable set of standards.

Another significant gap is the need for U.S. Government policy decisions regarding the administration and international standards position on the mapping of VoIP electronic numbers (ENUM) to traditional telephone numbers. Without a consistently administered, common database of records accessible to all service providers and enterprises (such as the one implemented for local number portability), VoIP interoperability may not occur.

Communications industry experts have also identified network management controls as an area of attention for the industry. The network outages that occurred in the early 1990's were a result of the cascading effect of software messages spreading through the network and the network elements not being able to protect against the rogue messages. In recent years, there have been a number of outages due to excessive traffic being sent from wireless networks to wireline networks. With the

expected increase in traffic (VoIP traffic) on both wireless and IP networks, network management controls need to be appropriately implemented among the various network types. In absence of these, the networks might experience outages at the network element level, or in some cases, cascading outages within a network as well as among networks.

Other interoperability gaps are also highlighted within this report, such as a specific means for handling E911 calls for mobile VoIP devices, wireless authentication and access, Quality of Service (QoS) between IP networks, support of CALEA, and adoption of VoIP encoding conventions that support all traditional PSTN features (e.g., tone-based services, facsimile, TTY).

Best Practices

Focus Group 3 is also recommending an initial set of Best Practices to help facilitate interoperability between packet-switched and circuit-switched networks for telephony services (see appendix E). These recommendations are limited due to the fact that VoIP is an emerging technology. Hence, the telecommunications industry does not have a wide body of experience from which to derive best practices. We recommend that best practices for PSTN and packet-switched network interoperability continue to be a focus area for future NRIC charters.

2. Background and Scope of Focus Group 3

In this section, we review the overall structure of NRIC VI and describe the position and objectives of FG3 based on the NRIC VI charter (see appendix B). We also recognize the many individuals who contributed to this effort.

2.1 Structure of NRIC VI

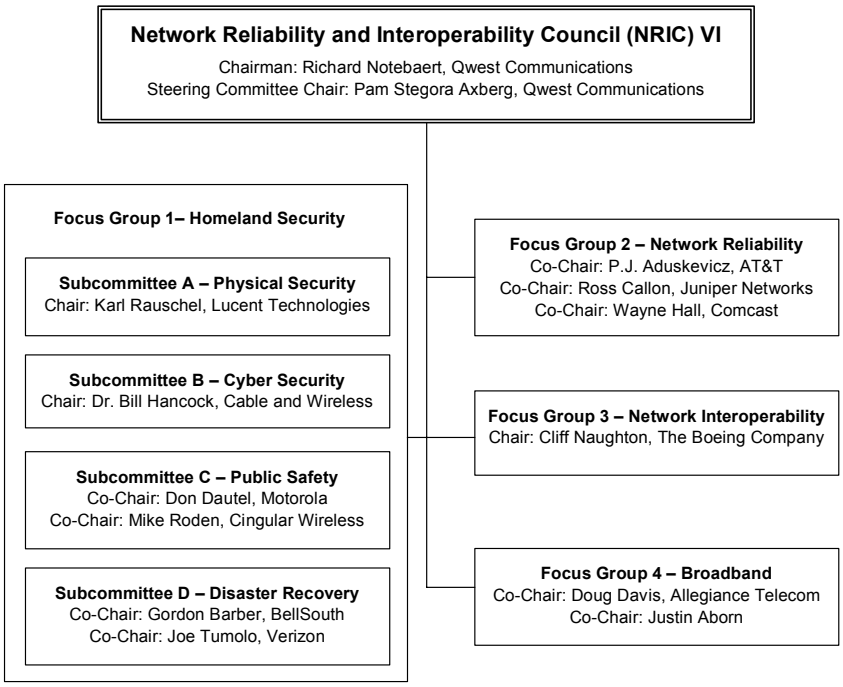

Figure 1. Structure of NRIC VI

2.2 Scope of NRIC VI FG3 Interoperability Effort

The purpose of this report is to inventory existing and in-progress standards efforts related to network interoperability and to analyze these against a set of basic features and functions to see if any gaps exist in standards or industry best practices. The emphasis is on inter-provider signaling; however, where applicable, the perspective of an enterprise is also considered in the analysis. Take, for example, the mapping of VoIP numbers to telephone numbers. In order for VoIP users to have the same ubiquitous access as telephone users, there need to be common policy and protocol agreements for translating VoIP numbers to IP routing information in order to place a call. The situation is analogous to the need for local

number portability (LNP) in the 1990s. The intent of this report is to bring attention to currently unresolved issues such as these.

The scope of the NRIC VI FG3 interoperability effort is focused on VoIP support for a basic set of features and functions between service providers and, in some cases, enterprises. The scope covers VoIP-PSTN calls as well as direct VoIP-VoIP calls. Discussion of interoperability issues covers the technical, operational, and/or regulatory space. The interoperability between the existing circuit-switched networks and the VoIP networks within a single service provider network is out of scope for FG3.

Figure 2 illustrates the overall scope of the NRIC VI FG3 interoperability report. At the bottom of the figure are the users of voice, data, and VoIP. In the middle are the various access networks and technologies used to access public telecommunication services—the PSTN and IP networks—which are shown at the top of the figure. The scope and focus of this effort are on the network interconnection points shown in this figure as shaded ellipses. This figure illustrates two general principles used in determining what is in and what is out of scope. Only interfaces and protocols between networks or service providers are in scope. Interfaces and protocols between a network and a user/subscriber or interfaces internal to a network or service provider are out of scope. This report summarizes these interfaces and protocols, along with any gaps, but does not state how implementation of these would be regulated or agreed to between service providers.

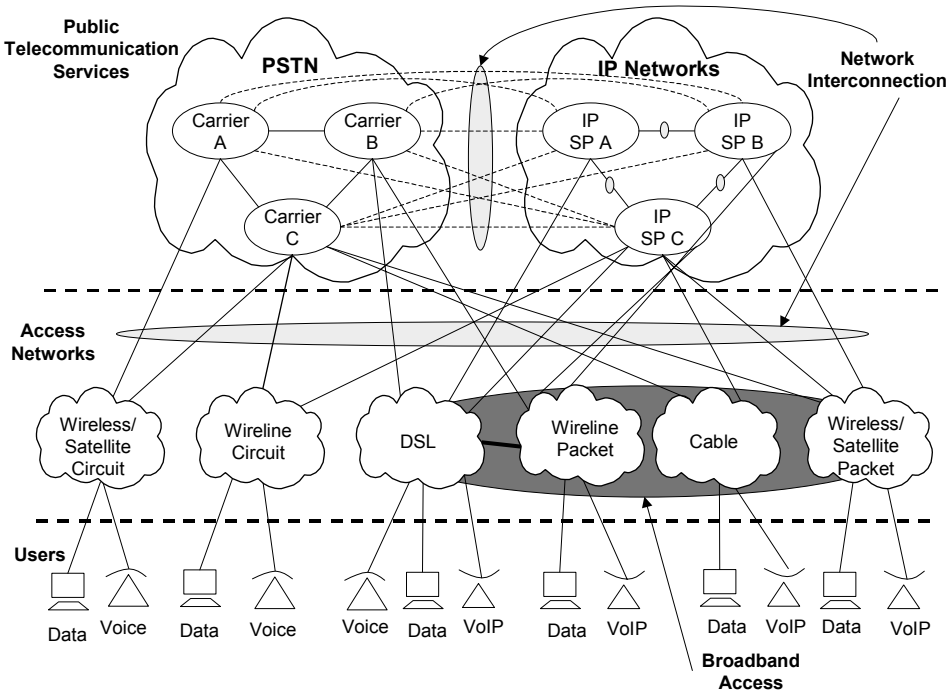

Figure 2. Scope of VoIP Interoperability Report

The following bullet points further detail what is in and what is out of scope in this report.

The following functions are **in scope**:

- Allow any voice user to place and receive calls from other users that are identified by an assignment of a North American Numbering Plan (NANP) number to that user.
- Support portability of NANP numbers across providers for numbers assigned to VoIP users.
- Provide the means for a VoIP device to access basic PSTN functions, such as using a telephone keypad to interact with a voice mail system.
- Support essential voice services, such as E911 and teletype technology (TTY/TDD).
- Provide minimum interoperation of VoIP service when features are not available on the other networks.
- Identify VoIP signaling protocols used between service provider (and some enterprise) networks.
- Identify methods to achieve VoIP calls of acceptable quality and delay.
- Identify VoIP protocol standards that could be used to support consumer telephony features (e.g., caller ID, call waiting, hold).
- Provide support for the Communications Assistance for Law Enforcement Act (CALEA).

A number of topics are **not explicitly in scope but are included** in this report because there is a need to understand the associated functions and standards in order to achieve interoperability. These include

- Voice coding standards for user VoIP devices and gateways within and between service provider and/or enterprise networks.
- Signaling protocols for user VoIP devices and gateways within and between networks (e.g., ITU-T H.323 and IETF Session Initiation Protocol [SIP]).
- Quality-of-service (QoS) requirements.
- Support over certain types of access networks (e.g., satellite, IEEE 802.11 Wireless Local Area Networks).
- Specifics of vendor interoperability, which are achieved through service provider interoperability.

The following topics are considered **out of scope** in this report. This does not mean that these issues are unimportant or irrelevant; they may not have been addressed because of limited time and resources:

- Current PSTN and time-division multiplexing (TDM) network interoperability.
- VoIP end-device (e.g., SIP phone) portability between service providers.
- “Best effort VoIP with no service provider involved” (e.g., intra-enterprise or between VoIP devices over the Internet that do not involve a service provider).

- Protocols and interfaces used within a service provider's network (e.g., Media Gateway Control, Megaco).
- Nonconsumer voice features (e.g., Centrex, Government Emergency Telecommunications Service [GETS]).
- Voice over Asynchronous Transfer Mode (VoATM) and Voice over Multiprotocol Label Switching (VoMPLS).

The diagram shown in figure 3 represents what FG3 has determined to be in scope and out of scope within this document in terms of the generic network interconnection between PSTN and IP networks in the context of the Public Telecommunication Services (upper part) shown in figure 2. All connections between IP and PSTN service providers are in scope and labeled with a green dot. Protocols and interfaces that are out of scope are represented with a red X. In general, protocols and interfaces within a service provider or to subscribers are out of scope for this document. The following text briefly introduces these protocols and their use as background to this section. Details about these protocols are presented in section 3.

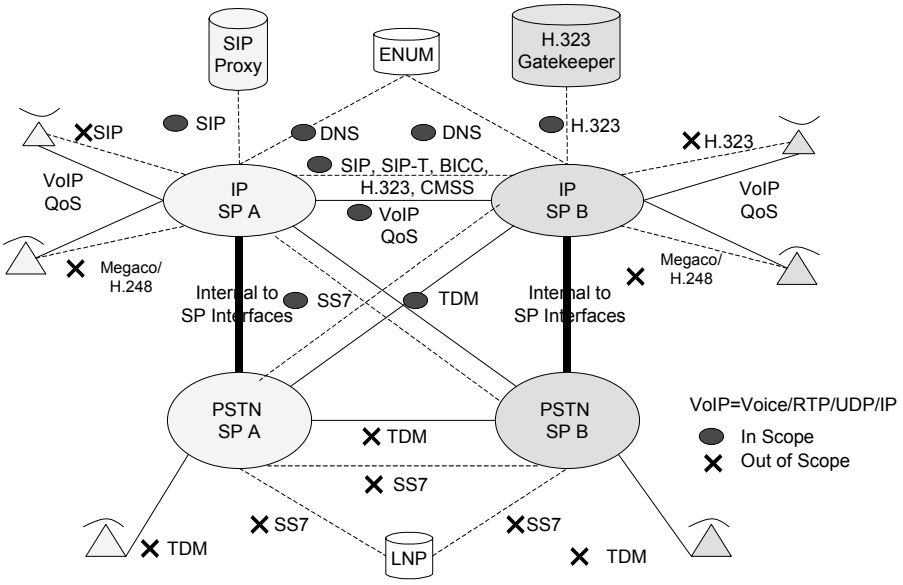

Figure 3. Basic VoIP Interoperability Reference Model

Well-established and interoperable protocol connections between PSTN service providers such as Signaling System 7 (SS7) and digital TDM voice are out of scope. Current PSTNs employ SS7 signaling to set up and manage calls and to deliver advanced intelligent network/intelligent network (AIN/IN) services, such as LNP. On the other hand, such SS7/TDM protocols, when used between an IP and PSTN in different service providers, are within scope.

The SIP is an Internet Engineering Task Force (IETF) standard protocol that supports both VoIP and more advanced multimedia (integrated voice, data, video, and graphics) services on an IP-based infrastructure. The ITU-T has defined a set of VoIP specifications as described in Recommendation H.323. Interoperation between SIP and H.323 VoIP end devices is an important interoperability consideration. When used as subscriber signaling, these protocols are out of scope, but they are in scope when used between service providers.

Various sets of standard protocols exist for VoIP-based networks; however, the standards for the integration of multiple protocols are in varying stages of development and deployment. In Voice over Packet environments, Bearer Independent Call Control (BICC) is an International Telecommunication Union (ITU) standard signaling protocol that supports narrowband voice-oriented services over a broadband packet-based network. BICC is based on SS7/ISDN User Part (ISUP), has multiple capability releases, and is seen as a practical solution to ease the transition towards next generation network (NGN) architectures. Another protocol that transparently conveys SS7/ISUP over SIP is called SIP for Telephones (SIP-T). Both BICC and SIP-T are in scope because they will potentially be used between service providers.

VoIP call processing architectures have a feature server (called a “proxy” in SIP and a “gatekeeper” in H.323) that uses the native protocol as an interface. Because the provider of this server may be different than that of the service provider, the SIP or H.323 protocol interaction is in scope. VoIP protocols use names similar to e-mail addresses instead of phone numbers. In order to interoperate with phones connected to the PSTN, there is a need to map telephone numbers (as defined in ITU-T Recommendation E.164) to VoIP names. The Domain Name System (DNS) protocol is used to access an E.164 number (ENUM) database for this purpose; hence, this is an important part of VoIP interoperability that is within scope.

The Megaco/H.248 is a standard protocol in joint development by the IETF/ITU-T for communication between a media gateway and a media gateway controller, which may be located on a subscriber premise or internal to a service provider network. These protocols are considered out of scope because they are used only internally within a service-provider network, or between a network and a subscriber.

While the networks evolve to NGN architecture, many different protocols are going to coexist, so it is critical to determine how they are going to interoperate in order for companies to begin to deploy IP networks effectively.

To accomplish the integration and evolution from PSTN- to IP-based networks (and/or the interoperation of VoIP services and PSTNs), QoS issues must be addressed. QoS metrics include transit delay (latency), delay variation (jitter), and packet loss. In order to meet these QoS metrics, different mechanisms may be employed based upon access network technology or by agreement between service providers. In a manner analogous to other scope decisions, QoS between a subscriber and a network are out of scope, while QoS interactions between IP service providers are in scope.

2.3 FG3 Team Members

The following participants served as authors.

Participant	Company
Franklyn Athias	Comcast Corporation
John Border	Hughes Network Systems
Jamal Boudhaouia	Qwest Communications
Rick Canaday	AT&T
Greg Carras	The Boeing Company
John Chapa Jr.	SBC Operations
Robert Dianda	Sprint
Thomas R. Helmes	Verizon
Percy Kimbrough	SBC
Denis Kuwahara	The Boeing Company
Jim Lankford	SBC
Chris Liljenstolpe	Cable and Wireless
Marc Linsner	Cisco Systems
Dr. Anil Macwan	Lucent Technologies
Dr. David E. McDysan	MCI
Michael McInnis	The Boeing Company
Cliff Naughton	The Boeing Company
Mark Neibert	Intelsat
Art Reilly	Cisco Systems
Kent Shuey	The Boeing Company
Jim Turner	ATIS
Robert M. Wienski	VeriSign (formerly Illuminet)
Mark Willborn	Allegiance Telecom
Dr. Eric Yam	ECTEL
Albert Young	Cox Communications

The following participants served as reviewers.

Reviewer	Company
Justin Aborn	Genuity
Ron Bath	VoiceStream
Jane Builder	VoiceStream
Adam Dunstan	Avici Systems
Randall Hemauer	Sprint
Mike Holmes	Lucent Technologies
John Jennings	Nortel Networks
Rick Kemper	CTIA (Cellular Telephone and Internet Assoc.)
Tom Kuba	Lockheed Martin
Sam Phillips	BITS
Rod Raglan	Hughes Network Systems
Gary Roboff	BITS
Marty Schulman	Juniper Networks
Dan Schutzer	BITS (Citigroup)
Iyad Tarazi	Nextel Communications
Chris Wallace	Nokia
Heather Wyson	BITS

3. VoIP Interoperability Gap Analysis

Overview

The amount of data traffic is now surpassing the amount of voice traffic on U.S. telecommunications networks. Continued growth of data traffic makes the transition to an IP-based infrastructure economically attractive for several reasons (e.g., common, shared technology infrastructure; common operations and support organizations). Service providers are seeking technology solutions to help them deploy IP-based voice services in addition to (or instead of, in some cases) the traditional PSTN-based voice services.

Industry standards are required in order to ensure interoperability between both vendor equipment and individual networks, as well as to ensure that end-to-end performance and reliability, scalability, and security objectives can be met. Complicating the standards issues are the different requirements, markets, and businesses that are currently deploying VoIP networks.

This section provides an overview, analysis, and any FG3 recommendations concerning the protocols considered to be in scope of this document (also see section 2.2). Not all providers need implement every protocol or function. However, at least pairwise agreement between providers is needed on which of several protocols to implement. Hopefully, as occurred in the telephone industry, a smaller set of protocols will eventually become the de facto industry standard.

Research and Analysis

Interoperability needs to be addressed at every point of interconnection of network components such as protocols, vendor implementation, carrier interoperability, and services interoperability. The standardization of VoIP protocols and the development of Profiles and Implementation Agreements will facilitate vendors in getting products to market quickly and cost-effectively and will enable carriers to deploy flexible services.

Current Practices

In the past, the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) defined H.323-based networks as the preferred VoIP architecture, but emerging IP-based voice services have been made possible with the Session Initiation Protocol (SIP). The new VoIP voice-based applications and services, based predominantly on SIP, can be rapidly deployed, generate revenue for the service provider, and at the same time, can integrate web-based applications for full multimedia service interoperability.

SIP is the IETF's next-generation protocol for multimedia services and service control. SIP was first standardized by the IETF in March 1999 and updated in June of

2002. It is currently becoming the predominant standard for the development of new VoIP services.

Current Work Items and Standards Development Organizations

Current work items include

- The ongoing standardization of VoIP protocols.
- The specification of Profiles and Implementation Agreements between the network elements.
- Providing a forum for vendors to test the interoperability of their hardware. This forum will ensure that the vendors uniformly interpret the standards and that all network elements interwork without interoperability issues.

The key standards bodies currently working on VoIP interoperability are the IETF, the ITU, ANSI, and the SIP Forum.

3.1 Signaling Architectures

The following three areas are addressed in this section:

1. Call control protocols. Call control protocols (e.g. SS7, SIP, SIP-T, BICC) cover the establishment, release, and modification of calls.

2. Signaling transport. Signaling transport provides a reliable transport of signaling messages between signaling endpoints (e.g., between two switches). Functions provided include detection and recovery of lost or corrupted information and the detection of loss of communication between signaling endpoints.

3. Network management. Network management provides controls for maintaining network performance and security during overload (e.g. because of a mass calling event).

3.1.1 Signaling System 7

Overview

SS7 consists of multiple parts, including the following:

- ISDN User Part (ISUP) is the call control part of the SS7 protocol. ISUP determines the procedures for setting up, coordinating, and taking down trunk calls on the SS7 network.

- Message Transfer Part (MTP) is the part of SS7 that is used to

  – Place formatted signaling messages into packets.

  – Strip formatted signaling messages from packets.

  – Send or receive packets.

- Transaction Capability Application Part (TCAP) is the application layer protocol of SS7. TCAPs in the SS7 suite are functions that control non-circuit-related information transferred between two or more signaling nodes (e.g., in database queries).

- Network management capabilities are used during traffic overload conditions.

Analysis

For circuit-switched networks, SS7 is a mature protocol and is widely used. BICC is the part of SS7 that addresses VoIP. See section 3.1.4 for details on BICC.

Wireline SS7 signaling networks can invoke different forms of network management when networks become congested, as when a natural disaster occurs. Two forms of network management controls are typically used, protective and expansive. Protective controls remove traffic from the network during overload conditions. Expansive controls reroute traffic from routes experiencing overload to other, less congested routes.

Gaps Identified

None for circuit-switched networks.

Recommendations

SS7 network management controls keep overload conditions from propagating across the public network. VoIP networks signaling networks can use the network management controls of SS7. VoIP network operators should give serious consideration to implementing and using these controls within their networks.

3.1.2 Session Initiation Protocol

Overview

The SIP is an IETF signaling protocol for establishing real-time calls and conferences and is typically carried over IP networks (IETF RFC 3261). Each session may include different types of data, such as audio and video communication. Telephone calls are considered a type of multimedia session where only audio is exchanged. As a traditional text-based Internet protocol, SIP resembles Hypertext Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP). SIP uses the Session Description Protocol (SDP) for media description.

SIP supports five facets of establishing and terminating multimedia communications:

- User location: determination of the end system to be used for communication.
- User availability: determination of the willingness of the called party to engage in communications.
- User capabilities: determination of the media and media parameters to be used.
- Session setup: "ringing," establishment of session parameters at both called and calling party.
- Session management: including transfer and termination of sessions, modifying session parameters, and invoking services.

SIP is independent of the packet layer. It has been designed to be a general-purpose protocol. SIP is an open standard and is extensible. As a basic feature, SIP enables personal mobility by providing the capability to reach a called party at a single, location-independent Uniform Resource Identifier (URI), which is similar in form to an e-mail address.

The basic architecture of SIP is client/server in nature. The main entities in SIP are the User Agent, the SIP Proxy Server, the SIP Redirect Server, and the Registrar.

The User Agents, or SIP endpoints, function as user agents as clients (UAC) when initiating requests and as user agents as servers (UAS) when responding to requests. User Agents communicate with other User Agents directly or through an intermediate server. The User Agent also stores and manages call states.

SIP intermediate servers have the capability to behave as proxy or redirect servers. SIP Proxy Servers forward requests from the User Agent to the next SIP server or User Agent within the network and also retain information for billing and accounting purposes. SIP Redirect Servers respond to client requests and inform them of the address of the requested server. Numerous hops can take place before the data reaches the final destination. The flexibility of SIP allows the servers to contact external location servers to determine user or routing policies. Therefore, the user is not bound into only one scheme to locate users. In addition, to maintain scalability, the SIP servers can either maintain state information or forward requests in a stateless fashion.

The third entity that comprises SIP is the SIP Registrar. The User Agent sends a registration message to the SIP Registrar and the Registrar stores the information. This registration information associates the URI of the SIP user with the current IP address in a location service by means of a non-SIP protocol. Once the information is stored, the Registrar sends the appropriate response back to the user agent.

A module performing the mapping between the PSTN SS7 ISDN User Part (ISUP) protocol and SIP is usually referred to as a media gateway controller (MGC), although the terms “soft switch” or “call agent” are also sometimes used. An MGC has logical interfaces facing both networks, the network carrying ISUP and the network carrying SIP. The MGC also has some capabilities for controlling the voice path; there is typically a media gateway (MG) with E1/T1 trunking interfaces (voice from PSTN) and with IP interfaces (VoIP). The MGC and the MG can be merged into one physical box or kept separate.

These MGCs are frequently used to bridge SIP and ISUP networks so that calls originating in the PSTN can reach IP telephone endpoints and vice versa. This is useful when PSTN calls need to take advantage of services in the IP world, when IP networks are used as transit networks for PSTN-to-PSTN calls, for architectures in which calls originate on desktop “softphones” but terminate at PSTN terminals, and for many other similar next-generation telephone architectures.

Analysis

SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions (conferences) such as VoIP calls. SIP can also invite participants to already-existing sessions, such as multicast conferences. Media can be added to (and removed from) an existing session. SIP transparently supports name mapping and redirection services, which supports personal mobility; users can maintain a single externally visible identifier regardless of their network location.

SIP is not a vertically integrated communications system. SIP is rather a component that can be used with other IETF protocols to build a complete multimedia architecture. Typically, these architectures will include protocols such as the Real-time Transport Protocol (RTP) for transporting real-time data and providing quality of service (QoS) feedback (RFC 1889), the Real-Time Streaming Protocol (RTSP) for controlling delivery of streaming media (RFC 2326), and the Session Description Protocol (SDP) for describing multimedia sessions (RFC 2327). Therefore, SIP must be used in conjunction with other protocols in order to provide complete services to the users. However, the basic functionality and operation of SIP do not depend on any of these protocols.

SIP does not provide services. Rather, SIP provides primitives that can be used to implement different services. For example, SIP can locate a network object (e.g., a user, a voice mailbox) and deliver an opaque object to its current location. If this primitive is used to deliver a session description written in SDP, for instance, the endpoints can agree on the parameters of a session. If the same primitive is used to deliver a photo of the caller as well as the session description, a "caller ID" service can be easily implemented. As this example shows, a single primitive is typically used to provide several different services.

SIP does not offer conference control services such as floor control or voting and does not prescribe how a conference is to be managed. SIP can be used to initiate a session that uses some other conference control protocol. Because SIP messages and SIP sessions can pass through entirely different networks, SIP cannot and does not provide any kind of network resource reservation capabilities.

The nature of the services provided makes security particularly important. To that end, SIP provides a suite of security services, which include denial-of-service prevention, authentication (both user-to-user and proxy-to-user), integrity protection, and encryption and privacy services.

Gaps Identified

As identified above, SIP provides a large portion of the signaling required to establish and tear down telephony calls. SIP does not provide a mechanism for QoS, billing, network maintenance, or other aspects of operating a network. These issues are either individual implementation issues or left to other protocol definitions covering the operation of an IP network.

Although pure SIP has all the requisite instruments for the establishment and termination of calls, it does not have any baseline mechanism to carry any midcall

information, such as the ISUP information/information request (INF/INR) query, along the SIP signaling path during the session. This midcall information does not result in any change in the state of SIP calls or the parameters of the sessions that SIP initiates. SIP does provide the INFO method (RFC 2976) for midcall messages which should be used for this purpose, but interpretation of these messages is dependent on endpoint implementation.

Recommendations

The IETF working groups are actively extending the functionality of the baseline SIP protocol definitions to cover features offered by extended PSTN providers. These activities are ongoing and appear to be adequate at this time. It is believed that the natural process within the IETF will cover the currently identifiable gaps.

3.1.3 SIP to PSTN Interworking

Overview

SIP is an application-layer protocol for establishing, terminating, and modifying multimedia sessions. It is typically carried over IP. Within SIP, telephone calls are considered a type of multimedia session where only audio is exchanged.

ISUP is a layer 4 protocol used in SS7 networks. It typically runs over MTP, although it can also run over IP (see Stream Control Transmission Protocol [SCTP], IETF RFC 2960). ISUP is used for controlling telephone calls and for maintenance of the network (e.g., blocking circuits, resetting circuits).

A functional module performing the mapping between these two protocols is usually referred to as an MGC, although the terms “soft switch” or “call agent” are also sometimes used. An MGC has logical interfaces facing both networks, the network carrying ISUP and the network carrying SIP. The MGC also has some capabilities for controlling the voice path; there is typically an MG with E1/T1 trunking interfaces (voice from the PSTN) and with IP interfaces (VoIP). The MGC and MG are often merged into one physical box, though they can be kept separate.

These MGCs are frequently used to bridge SIP and ISUP networks so that calls originating in the PSTN can reach IP telephone endpoints and vice versa. This is useful when PSTN calls need to take advantage of services in the IP world, in architectures that have calls originating on desktop softphones but terminating at PSTN terminals, and in many other similar next-generation telephone architectures.

As described in section 3.1.2, SIP is one of the key protocols used to implement VoIP, but a VoIP network will most likely not exist in isolation from traditional telephone networks; therefore, it is vital for a SIP network to interoperate with the PSTN. SIP-T (IETF RFC 3372) is a set of mechanisms for interfacing traditional telephone signaling with SIP. The purpose of SIP-T is to provide protocol translation and feature transparency across points of PSTN-SIP interconnection. The actual mapping of ISUP messages into SIP is described in IETF RFC 3398. Both of these

mechanisms are intended for use where a VoIP network interfaces with the PSTN. At a SIP-ISUP gateway, SIP-T encapsulates SS7 ISUP messages so that information necessary for services is not discarded in the SIP request. SIP-T also translates critical routing information from an ISUP message into corresponding SIP headers for intermediaries such as proxy servers.

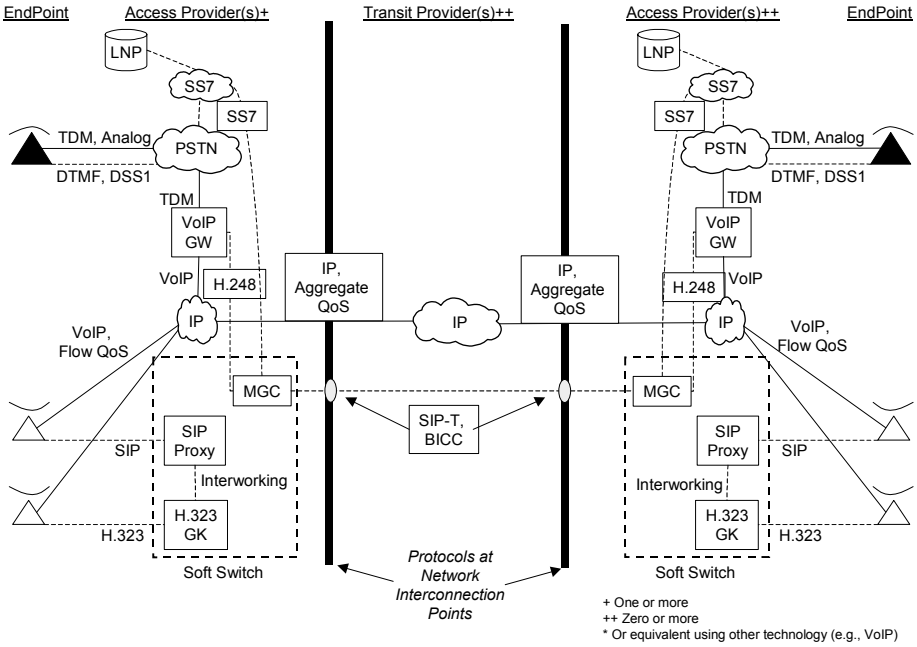

Figure 4. SIP-T Network Connection Points

Analysis

An important characteristic of any SIP network is feature transparency with respect to the PSTN. Traditional telecom services (e.g., call waiting, toll-free numbers) implemented in PSTN protocols such as SS7 should be offered by a SIP network in a manner that precludes any debilitating difference in user experience while not limiting the flexibility of SIP. One compelling need to do so arises from the fact that certain networks use proprietary SS7 parameters to transmit certain information through their networks. On the one hand, it is necessary that SIP support the primitives for the delivery of such services where the terminating point is a regular SIP phone rather than a device that is fluent in SS7. However, it is also essential that SS7 information be available at gateways, the points of SS7-SIP interconnection, to ensure transparency of features not otherwise supported in SIP. If possible, SS7 information should be available in its entirety and without any loss to trusted parties in the SIP network across the PSTN-IP interface.

Another important characteristic of a SIP telephony network is routability of SIP requests. A SIP request that sets up a telephone call should contain sufficient information in its headers to enable it to be appropriately routed to its destination by

proxy servers in the SIP network. Most commonly, this requires the parameters of a call (e.g., the dialed number) to be carried over from SS7 signaling to SIP requests.

In the progression from the PSTN model on nonintelligent end devices to the Internet model of intelligent end devices, it is necessary to analyze these different architectures to determine any interoperability gaps. A large number of PSTN class features that use the PSTN network can and will be replicated within the intelligent SIP device; therefore, communication requests for such features are not necessary for the terminating SIP endpoint. Although the PSTN uses the network to carry feature requests, pure SIP does not have any provision or need for carrying any midcall control information that is generated during a session, other than in SIP-T when a PSTN-to-PSTN call transits a SIP network. SIP does provide the INFO method (RFC 2976) for midcall messages, which should be used for this purpose. This midcall information does not result in any change in the state of SIP calls or the parameters of the sessions that SIP initiates. Note, however, that INFO is not suitable for managing overlap dialing at this time. Work is ongoing within the IETF to handle this need. Also, note that the use of INFO for signaling midcall Dual-Tone Multi-Frequency (DTMF) signals is not recommended because there are other mechanisms within SIP for this function. (See IETF RFC 2833 for a recommended mechanism.)

SIP provides a large portion of the signaling required to establish and tear down telephony calls. SIP-T does not provide a mechanism for QoS, billing, network maintenance, or other aspects of operating a network. These issues are either individual implementation issues or left to other protocol definitions covering the operation of an IP network.

The SIP framework, as described in RFC 3372 and RFC 3398, provides a mechanism for SIP-to-ISUP interworking when it is desired that the bearer channel is also controlled by SIP. SIP-T is not intended to provide a transport only for all layers of SS7 networks because SIP-T does not handle network layer issues like MTP error detection and recovery. A better choice for this application is Stream Control Transmission Protocol (SCTP) as described in RFC 2960 and the corresponding Message Transfer Part 3 (MTP3) User Adaptation Layer (M3UA) protocol defined in RFC 3332.

SS7 MTP3/ISUP network maintenance and management messages and network overload messages have impact on the SIP network only to the extent that established calls may get dropped because of reset or blocking messages, or call setups may get denied because of overload conditions. It is the responsibility of the MGC or MG to react to these SS7 network messages. SIP will react to the corresponding UA/UAS messages that the MCG or MG generates, based on the SS7 network message.

SIP offers a wide feature set with many different ways to accomplish a task. For example, transmitting DTMF digits during a call could be done within the bearer channel or transmitted in an additional RTP session outside of the voice channel. Of course, when a SIP call crosses different service provider networks, the mechanisms used to accomplish a task need to be agreed on by the service providers. This agreement involves identifying the application profiles (a set of agreed-upon mechanisms) that will be used to provide continuity of features. This issue is

currently being worked within the TIA TR41.4 group as it updates the TIA-811 standard to include VoIP.

Gaps Identified

SIP is a broad protocol that provides the primitives for call establishment and teardown. As noted above, there are multiple ways to accomplish different call features. Without an agreed-upon set of profiles for these feature mechanisms, there could be a gap in service provider interoperability.

Also, as identified above, SIP-T does not define a mechanism to respond to overlap dialing, and it supports only en bloc dialing. En bloc dialing is the standard mechanism in use within the United States.

Recommendations

The IETF working groups are actively extending the functionality of the baseline SIP protocol definitions to cover features offered by extended PSTN providers. These activities are ongoing and appear to be adequate at this time.

Industry forums such as the SIP Forum (www.sipforum.org) track the resolution of issues concerning interoperability between SIP implementations and other protocols.

In the near term, as the protocol matures, interoperability can be achieved by means of bilateral agreements between service providers. Industry forums and standards development organizations will be better positioned to create best current practices as they gain experience with the new technology.

3.1.4 Bearer Independent Call Control

Overview

BICC is an ITU-T protocol suite designed to allow PSTNs to offer the complete set of PSTN/ISDN services, including all supplementary services, over a variety of intervening data networks (e.g., IP and ATM). The initial focus of BICC was the transport of narrowband ISDN over an intervening ATM broadband network (Capability Set 1 [CS1]). Capability Set 2 (CS2), now nearing completion, added support for IP bearers and a multitude of interworking scenarios. Capability Set 3 (CS3), currently under development, with releases scheduled for late 2003 and for 2004, adds support mechanisms for end-to-end QoS control and multimedia capabilities.

Basically, BICC provides for the carriage of call-level signaling between PSTN and IP gateways, termed Interface Serving Nodes in ITU terminology and MGCs in a decomposed gateway model. As such, BICC can be considered functionally equivalent to SIP-T, although it utilizes different mechanisms and architecture concepts (i.e., BICC is based on SS7 ISUP rather than on the IETF SIP).

Analysis

The BICC protocol suite is at an advanced stage of development in the ITU-T Study Group 11, with North American input from the ATIS-T1 Common Channel Signaling (T1S1.3) Working Group and the Services Architecture and Control (T1S1.7) Working Group. Its greatest strength is its complete support for existing PSTN networks and services, being based from its inception on current ITU-T PSTN protocols (mainly SS7 ISUP).

Numerous standards are complete and approved, a few of which are listed below:

- Q.765.5. SS7 Application Transport Mechanism – Bearer Independent Call Control (BICC) (CS1).
- Q.1901. Bearer Independent Call Control Protocol (CS1).
- Q.1902.1 to Q.1902.6. Bearer Independent Call Control Protocol (CS2). Status: Approved and in force.
- Q.1912.1 to Q.1912.4. Interworking between BICC and Other Signaling Systems. Status: Approved and in force.
- Q.1903 series Recommendations. BICC CS3 parameters, messages, and requirements under development as part of ITU-T Study Group 11 Question 11/11. These recommendations are planned for release in late 2003 and in 2004.

As stated above, BICC was designed from the start as a means to extend end-to-end PSTN connectivity and services over packet networks. As such, its ability to support all legacy PSTN services over intervening IP networks is basically ensured. However, although BICC is being offered in vendor equipment and has been deployed somewhat throughout the world, it appears that the trend in the industry is to migrate IP-based networks toward a SIP-based signaling infrastructure. This suggests that SIP-T, designed in the IETF for the carriage of ISUP across an (SIP-based) intervening IP network, may be a better solution. Therefore, it is not clear that there will be significant deployment of BICC in the future. Interworking may still be required between legacy or new BICC implementations and evolving SIP-T networks (e.g., at an IP-based carrier interconnection point) but this should not lead to any significant interoperability difficulties because both protocols perform essentially the same function—the carriage of ISUP information elements across IP networks, which allows for a straightforward mapping of information elements.

Gaps Identified

No interoperability gaps have been identified, primarily because of the legacy of BICC as an ITU-T protocol suite specifically targeted at providing end-to-end PSTN service capabilities across intervening IP networks.

Recommendations

None.

3.1.5 H.323 to PSTN Interworking

Overview

The H.323 protocol suite is the international standard developed by the ITU-T for multimedia communications over packet-based networks, including the convergence of voice, video, and data communications. H.323 standardization work continues within ITU-T Study Group 16, with H.323 version 5 scheduled for approval in the near future. Originally approved in 1996 with an emphasis on multimedia LAN capabilities, including the extension of PSTN connectivity over LANs, it has subsequently been extended (and used) to cover wide-area IP network connectivity. As such, it is a suitable protocol basis for providing VoIP capabilities between PSTN service providers or between PSTN and pure IP endpoints.

H.323 is an umbrella document describing the use of a number of specific protocols, including H.225.0 (for call signaling and remote authentication—Registration, Admission, and Status Protocol [RAS]), H.245 (for end-to-end capability negotiation), H.235 (for security aspects), as well as a number of other extensions, including H.246 for PSTN interworking. It uses a number of IETF protocols, including RTP for real-time transport of audio and video over packet networks and the Uniform Resource Locator (URL) concept for identifying endpoints.

The main components of an H.323 system consist of gateways, terminals, multipoint control units, and an optional gatekeeper. Gateways can be either integrated or decomposed into a separate control function and media processing function (decomposed gateways use the Megaco/H.248 control protocol).

Analysis

H.323 was designed from the start as a means to extend PSTN connectivity over LANs in addition to providing multimedia capabilities between terminals directly connected to the LAN. As such, PSTN-to-VoIP packet interworking capabilities are basically ensured. Additional details are specified in H.246 as well. And while being originally designed for LAN applications, it has in fact frequently been used for wide-area packet connectivity and has been enhanced numerous times in its transition from version 1 (in 1996) to version 5 (due shortly) to support scalable, wide-area PSTN service provider connectivity (as well as other enhancements).

Gap Analysis

No interoperability gaps have been identified, primarily because of the H.323 legacy as an ITU-T protocol suite specifically targeted at providing PSTN interoperability and extension across LANs.

Recommendations

None.

3.1.6 H.323 to SIP Interworking

Overview

The H.323 protocol suite is the international standard developed by the ITU-T for multimedia communications over packet-based networks, including the convergence of voice, video, and data communications. H.323 standardization work continues within ITU-T Study Group 16, with H.323 version 5 scheduled for approval in the near future. As such, it is a suitable protocol basis for providing VoIP capabilities between PSTN service providers or between PSTN and VoIP endpoints.

SIP was developed by the IETF. Although it was designed as the basis for general multimedia IP-based communications networks, it is also seen as the primary candidate protocol to serve as the basis for VoIP. SIP is not a vertically integrated communications system like H.323. Therefore, SIP must be used in conjunction with other protocols in order to provide complete services to the end users.

Analysis

H.323 is an umbrella standard that takes a classical telephony/telecommunications approach by providing a complete, well-defined system architecture as well as implementation guidelines that cover the entire call set-up, call control, and media used in the call. SIP, on the other hand, takes the IETF approach of defining individual components or building blocks rather than complete systems. SIP is therefore not as strictly defined or as complete a system as H.323. Many aspects of the SIP architecture are left open to interpretation or deemed to be “implementation issues.”

Both H.323 and SIP (with its complementary IETF protocols) provide similar QoS and comparable functionality using different mechanisms. Although SIP promises to be more flexible and scalable, H.323 offers better network management and interoperability because of its well-defined nature. The differences between the two have been diminishing with each new version.

Although H.323 has been widely deployed throughout the world, both in the enterprise and in the wide area, the trend in the industry appears to be to migrate toward a SIP-based network infrastructure in the future rather than continue to expand H.323 networks. Legacy H.323 networks will likely remain in place and be somewhat extended, and some new H.323 networking will be used by certain providers. For these reasons, interworking between SIP-based networks and H.323-based networks is an important issue.

Work is relatively advanced in the IETF to address this SIP/H.323 interworking. The current draft document (draft-agrawal-sip-h323-interworking-reqs-05.txt, June 28,

2003, expires December 2003) describes the requirements for the logical entity known as the SIP-H.323 Interworking Function, which will allow the interworking between SIP and H.323.

Gaps Identified

Although the SIP/H.323 interworking requirements draft is technically quite advanced in the IETF, it is still not a working group draft, and no decision has been issued yet from the Internet Engineering Steering Group (IESG) with respect to its review for consideration as a Proposed Standard RFC (a necessary step to progress it along a standards track). This is expected to happen, however, and to be noncontroversial, although no dates have been set. In any case, the current draft is entirely usable as a basis for vendor implementations.

Recommendations

The industry should address the SIP-to-H.323 interworking draft within the IETF to ensure that it progresses on the standards track in a timely manner.

The reason for this recommendation is that, as various PSTN service providers evolve their PSTN-VoIP interworking capabilities, SIP-to-H.323 interfacing will increasingly be needed. This is partly due to the need to interwork with legacy H.323 networks and partly due to the fact that H.323 will continue to be deployed to a certain extent for some end users and enterprises. This will necessitate more H.323-to-SIP interconnectivity as PSTN providers continue to deploy and enhance their SIP-based VoIP networks.

3.1.7 Signaling Transport

Overview

The Signaling Transport Working Group of the IETF is in the process of developing a set of RFCs that define a means of transporting packet-based PSTN signaling across an IP network. So far they have attended to many of the various signaling applications that currently use SS7 for transport. The stated goal is to provide transport functionality and performance over IP for these signaling applications that equal the functionality and performance of the currently used packet transport mechanisms. Signaling Transport (SigTran) defines gateway configurations as well as end-to-end IP transport between two PSTN signaling points.

Analysis

The RFCs released by the Signaling Transport Working Group specify a group of new protocols, which work over IP to replace the first two or first three layers of the SS7 protocol stack. The stated goal of the working group is to provide all the

functionality and performance in SigTran that the current SS7 transport protocol provides. Higher layers of the SS7 stack will pass untouched across the IP network.

Gaps Identified

Because SigTran replaces only the transport layers and leaves the application layers intact, there should not be any interoperability gaps.

Recommendations

None.

3.1.8 Network Management Control Between Different Networks and Applications (e.g., Wireline and Wireless)

Overview

Network management is a set of real-time procedures aimed at optimizing network performance when the network is under stress caused by overload conditions. Network management provides and operates control and surveillance features that aid in maintaining network integrity and stability during overloads and failures.

See appendix D for a detailed discussion of network management procedures.

Analysis

Network management between wireless and wireline networks is incompatible.

Wireline SS7 signaling networks can invoke different forms of network management when networks become congested, as when a natural disaster occurs. Two forms of network management controls are typically used, protective and expansive. Protective controls remove traffic from the network during overload conditions. Expansive controls reroute traffic from routes experiencing overload to other, less congested routes. Some wireless networks use TIA/EIA IS-41 for signaling. IS-41 supports intersystem operations for wireless networks. IS-41 is a upper layer application that supports X.25 or SS7 call setup and transport and with some recent standards development, supports IP. Wireless systems have moved away from X.25 and are using SS7.

IS-41 networks can use the network management controls of SS7. Use of SS7 network management controls is an implementation/operational decision of the wireless operator. There is nothing inherent in IS-41 that prevents it from using SS7 network management controls. It is left to the network operator to decide how he will incorporate IS-41 or use those controls within his networks.

Gaps Identified

Initially, wireless networks used IS-41 signaling. These networks connected to the wireline SS7 networks directly. Some IS-41 networks do not use SS7 network management controls (even though they can do so). As a result, they cannot inter-operate seamlessly with wireline networks.

Some wireless carriers have migrated from IS-41 networks to SS7 networks. As a result, the network management capabilities now work between the wireless and wireline network. Some wireless companies have also migrated to SS7. The industry has several examples of network management controls working properly between wireless and wireline networks when both networks are using SS7.

Recommendations

SS7 network management controls keep overload conditions from propagating across the public network. Wireless Service Providers (WSP) who have deployed IS-41 signaling networks (code division multiple access [CDMA], Global System for Mobile Communications [GSM], and time-division multiple access [TDMA]) can use the network management controls of SS7. WSPs are encouraged to give serious consideration to implementing and using these controls within their networks. These WSPs are also encouraged to implement and use SS7 network management controls within their networks.

3.2 Call Control Architectures

This section describes two examples of wireline IP call control architectures that use a master/slave approach appropriate to relatively nonintelligent endpoints. A call control architecture provides a functional (or logical) architecture for a switching system or network. In addition to defining the functional components, the architecture can include specification of the interface between the different functional components, as well as the external interfaces. Each functional component can be implemented in separate physical components, or multiple functional components can be implemented in a single physical component.

SIP and H.323 may be considered to contain call control architectures; they are peer-to-peer architectures suited to relatively intelligent endpoints. Because they have been described in detail in section 3.1, they will not be mentioned here.

3.2.1 Packet Tandem Architectures

Overview

The term “packet tandem” is not an official name for any standard architecture or technology grouping. It is, rather, used here to name a group of architectures that have emerged, primarily in the long-distance industry. These architectures are all characterized by a distributed set of components that connect using largely proprietary protocols on top of IP. Also, these components provide standard time-division multiplexing (TDM) interfaces to connect to PSTN switches. The call management components of these architectures were the first to use the term “soft switch.” These architectures resemble a disaggregated circuit switch. Although these architectures do not have formal standards support, some understanding can be found by studying the ongoing work on the web sites of the International Packet Communications Consortium (IPCC) at www.packetcomm.org and the Multiservice Switching Forum at www.msforum.org.

Analysis

Though the number of components and functionality of each component varies among vendors, there is always a component with call control functionality that is known variously as the “soft switch,” “call agent,” or “media gateway controller.” There is always a component that provides a media gateway function between standard TDM interfaces compatible with the PSTN and IP networks. ISDN primary rate interfaces (PRI) are common, as well as SS7 interfaces. These architectures are closed or self-contained, meaning their only network-to-network interfaces are the standard PSTN interfaces mentioned.

Gaps Identified

Because these are closed IP systems with standard PSTN interfaces, there should not be any interoperability gaps in the near term. In the longer term, these networks will need to be opened for direct interconnection to IP networks of other types in order to minimize delay and quality impairments caused by multiple encoding/decoding or transcoding (see section 3.4.3). Standard signaling interfaces and protocols such as those described in section 3.1 will be required. All of the gaps and recommendations listed for the included protocols will apply.

Recommendations

Service providers who have deployed closed tandem architectures using proprietary signaling protocols should prepare to open these networks for direct interconnection with other IP networks using standard signaling protocols.

3.2.2 PacketCable™ Architectures

Overview

"PacketCable" is the name given to a suite of interface specifications created by collaboration between Cable Television Laboratories (CableLabs® at www.packetcable.com), member Multiple System (cable TV) Operators (MSO), and vendors from the data and telecommunications technology industries. The PacketCable specifications form an architecture that allows multiple forms of electronic communication media to be carried on top of the CableLabs' Data-Over-Cable Service Interface Specifications (DOCSIS), more commonly known as "cable modems," with a high degree of QoS and security. PacketCables 1.0, 1.1, and 1.2 are heavily focused on delivery of telephony services. PacketCable multimedia exposes QoS and security functionality to other service architectures. The hallmarks of PacketCable 1.x include a master/slave orientation that allows end users to use standard "black phone" customer premise equipment and provides full E911 functionality as well as Justice Department accepted ("safe harbor") Communications Assistance for Law Enforcement Act (CALEA) support architecture. Protocols used within PacketCable are either commonly used, standard protocols or are "profiles," slightly modified or extended versions of commonly used or standard protocols.

PacketCable specifications have been submitted to the Society of Cable Telecommunications Engineers (SCTE) for adoption as a North American standard and accepted under the name "IPCablecom." PacketCable specifications have been submitted to ITU-T and accepted as IP Cablecom-approved specifications in the J.16x and J.17x series.

Analysis

The PacketCable 1.x architecture is a fully defined, stand-alone architecture that uses standard interfaces to the PSTN (e.g., SS7, CAS, ISDN PRI). Mechanisms are included to provide high-quality voice and to support E911 functionality in standard ways. CableLabs has incorporated support for CALEA and proactively promoted the methodology to the FBI and Justice Department, gaining “safe harbor” status. The specifications define several logical components but allow the functionality of these components to be combined in virtually any combination of physical components. Two of the defined functions are that of media (trunk) gateway and signaling gateway. These two functions provide standard PSTN TDM interfaces. Interoperability to the PSTN is a function of the quality of implementation of the PSTN standard interfaces on individual gateway components.

PacketCable 1.2 includes specifications that define interfaces between PacketCable networks operated by different service providers. One specification describes an internetwork signaling protocol named Call Management Server Signaling (CMSS). CMSS is a profile of SIP, as described in RFC 3261, with several extensions to make it more robust. CableLabs personnel are active in the IETF, proposing the SIP extensions as RFCs. Over time, the distinction between CMSS and SIP will blur and possibly disappear.

The PacketCable specifications provide support for CALEA on calls between two PacketCable networks. There are cases where calls require participation by both networks and messaging between them to support CALEA. One such case is a call inbound from a foreign network to a surveillance subject on a PacketCable network who has forwarded his calls to a directory number on a foreign network. The PacketCable call management server loses visibility of the call as soon as the redirect is accomplished. This case is covered with signaling in CMSS for two PacketCable networks. This case is covered with gateway requirements for PacketCable to PSTN interoperability. There are likely other call scenarios that need to be addressed in the CALEA area.

Interoperability between a PacketCable network and the PSTN or between two PacketCable networks is well defined by the specifications. Interoperability between a PacketCable network and other VoIP networks is less defined.

Gaps Identified

The PacketCable specifications do not adequately address interconnection of a PacketCable network with a non-PacketCable VoIP network. These networks will need to be opened for direct interconnection to IP networks of other types in order to minimize delay and quality impairments caused by multiple encoding/decoding or transcoding (see section 3.4.3). Standard signaling interfaces and protocols such as those described in section 3.1 will be required. All of the gaps and recommendations listed for the included protocols will apply.

Recommendations

Service providers who deploy PacketCable networks should be prepared to open these networks for direct interconnect to other IP networks as they evolve using standard signaling protocols.

3.3 Voice Over Wireless

The wireless industry continues striving to augment or even replace the wired local loop. With the proliferation of the Internet Protocol (IP), the need for wireless services to support data connectivity and voice services has become the driving force for all wired and wireless technologies.

Wireless and satellite communication technologies rely on the transmission of voice and data services through electromagnetic radiation across free space. Wireless communications are highly sensitive to atmospheric fading conditions, physical `blocking' obstacles, multi-path interference, and interference from other wireless transmitters. Wireless communications are also subject to interception by unauthorized recipients, leading to more stringent requirements for encryption or other means of securing voice and data in free space transit, to prevent fraud and unlawful intrusion.

For many years, telecommunications networks have employed wireless technology in the form of microwave radio communications. Microwave radio communications is a mature wireless technology whose engineering practices are well understood. Whether terrestrial or by means of "bent pipe" satellite links, these wireless communications links have generally provided back-end transport connections, linking highly controlled and managed nodes of a communication network together.

Wireless and satellite technologies have evolved to provide digital connectivity between end stations (e.g., mobile phones, laptop computers) and the communications network. As voice and data networks converge, wireless networks are becoming a key delivery vehicle for voice and data to the end station.

Overview

Wireless Personal Area Networks (WPAN)

A wireless personal area network (WPAN) is a wireless network of interconnecting devices centered around an individual person. Typically, a WPAN uses some wireless technology that permits communication within about 10 meters, in other words, a very short range. The objective is to facilitate seamless operation among home or business devices and systems. Each device in a WPAN is able to connect to any other device within the same WPAN, provided they are within range of one another. One WPAN technology is Bluetooth, which was used as the basis for a new WPAN standard by the IEEE 802.15 WPAN Working Group. Variations of the IEEE 802.15 standard include 802.15.1, 802.15.3, and 802.15.4.

WPAN technology is in its infancy and is undergoing rapid development. Within the IEEE 802.15 WPAN Working Group, ultra-wideband (UWB) radio technology has been proposed to increase WPAN data speeds to over 100 Megabits per second.

Wireless Local Area Networks (WLAN)

A wireless local area network (WLAN) is a network where wireless devices interconnect with a wired LAN through a network element called an access point. The connections between the WLAN devices and the access point are wireless. Typically, a WLAN uses some wireless technology that permits communications within about 100 meters. One such technology, currently the most common, is IEEE 802.11. Variations of the IEEE 802.11 standard include 802.11a, 802.11b, and 802.11g.

The technology for WLANs is undergoing rapid development. Within the IEEE 802.11 WLAN Working Group, 802.11g was recently released and a next-generation WLAN standard is currently being worked on within the 802.11n Task Group to make enhancements to the 802.11 WLAN standard to achieve throughputs of at least 100 megabits per second.

WLANs are expected to be widely deployed in public locations such as airports, restaurants, hotels, and coffee shops. Cellular operators commonly believe that they must provide a seamless user experience between cellular coverage areas and these WLAN hotspot areas.

Wireless Wide Area Networks (WWAN) and Wireless Metropolitan Area Networks (WMAN)

A wireless wide area network (WWAN) is a geographically dispersed telecommunications network where wireless devices interconnect with a wired voice and data network, which may include Internet Service Provider (ISP) resources. The term distinguishes a broader telecommunication structure than is provided from a wireless metropolitan area network (WMAN) and wireless local area network (WLAN). The wireless wide area network term usually connotes the inclusion of public (shared user) network elements.

A wireless metropolitan area network (WMAN) is a wireless network that interconnects users to wired Internet service provider (ISP) resources in a geographic area larger than that usually covered by a wireless local area network (WLAN) but is smaller than an area covered by a wireless wide area network (WWAN). The term is usually applied to the interconnection of a number of networks in a city (metropolitan area) into a single larger network, which may then also offer connection to a wide area network. It is also sometimes used to mean the interconnection of several local area networks by bridging them together with backbone private or leased lines.

The IEEE 802.16 wireless metropolitan area network (WMAN) group of broadband wireless communications standards were developed by a working group of the Institute of Electrical and Electronics Engineers (IEEE). The original 802.16 standard, published in December 2001, specified fixed point-to-multipoint broadband wireless systems. An amendment, 802.16a, approved in January 2003, specified non-line-of-sight extensions, delivering up to 70 Mbps at distances up to 31 miles. Officially called the WirelessMAN™ specification, 802.16 standards are expected to enable multimedia applications with wireless connection and, with a range of up to 30 miles, provide a viable last mile technology.

Wireless mobile cellular WWAN communication is one of the most prolific voice communications platforms that have been deployed within the last two decades. Within the United States technologies providing cellular concept services include: advanced mobile system (AMPS), digital-AMPS, total-access communication system (TACS), Code Division Multiple Access (CDMA) 2000, global system for mobile communication (GSM), and integrated dispatch enhanced network (iDEN) systems.

The concept of cellular radio was initially developed by AT&T at Bell Laboratories to provide additional radio capacity for a geographic customer service area. In 1979, the first commercial cellular phone system began operation in Tokyo, Japan. In 1981, Motorola and American Radio Phone began a U.S. cellular radio-phone system test in the Washington D.C.- Baltimore area. By 1982 the FCC authorized commercial cellular phone service in the United States. A year later, the first commercial cellular phone service to begin operation in the United States was an Advanced Mobile Phone Service (AMPS) cellular phone system offered in Chicago by Ameritech.

ITU wireless activities include the establishment of a set of interdependent ITU Recommendations called the International Mobile Telecommunications-2000 (IMT-2000). This is the global standard for third generation (3G) wireless communications. IMT-2000 provides a framework for worldwide wireless access by linking diverse terrestrial and/or satellite based networks. ITU activities on IMT-2000 comprise international standardization, including frequency spectrum and technical specifications for radio and network components.

Wireless Local Loop (WLL) systems use many platforms similar to cellular. A WLL system differs from a cellular system in its application, which is to provide fixed services rather than mobile services. Primarily, a WLL system connects a subscriber to the local telephone company using a radio link as its transport medium instead of copper wires. A fixed wireless service is often referred to as either a local multipoint distribution system (LMDS), a fixed wireless point-to-multipoint (FWPMP) system, a multichannel multipoint distribution system (MMDS), an instructional television fixed service (IFTS), or a multipoint distribution service (MDS) system.

Satellite

A satellite is a specialized wireless receiver and transmitter that is launched by a rocket and placed in orbit around the earth. There are hundreds of satellites currently in operation. They are used for such diverse purposes as weather imaging and forecasting, television broadcast, radio broadcast, amateur radio communications, Internet communications, and location determination based on the Global Positioning System (GPS).

There are three types of communications satellite systems. They are categorized according to the type of orbit they follow.

A geostationary satellite orbits the Earth directly over the equator, approximately 22,000 miles above the Earth’s surface. At this altitude, one complete trip around the Earth (relative to the sun) takes 24 hours. Thus, the satellite remains over the same spot on the Earth's surface at all times, and stays fixed in the sky. Any point on the surface from which it can be seen is commonly referred to as its footprint. A single geostationary satellite has a footprint that covers approximately 40 percent of the earth's surface. Three such satellites, spaced at equal intervals (120 angular degrees apart), can provide coverage of the entire civilized world.

A low-earth-orbit (LEO) satellite system employs some number of satellites, each in it's own circular or elliptical orbit around the Earth at an altitude of a few hundred miles, providing its own footprint on the Earth's surface. LEO orbits take the satellites over, or nearly over, the geographic poles. A LEO satellite system operates in a manner similar to the way a cellular telephone system functions. The main difference is that the satellite transponders, the wireless receivers and transmitters, are moving rather than fixed, and are in space rather than on the earth. A well-designed LEO system makes it possible for anyone to place a phone call or access the Internet from any point on the planet using a wireless device.

A medium earth orbit (MEO) satellite is one with an orbit within a range of a few hundred miles to a few thousand miles above the earth's surface. Satellites of this type orbit higher than low earth orbit (LEO) satellites, but lower than geostationary satellites.

Because MEO satellites are closer to the earth than geostationary satellites, earth-based transmitters with relatively low power and modest-sized antennas can access the system. Because MEO satellites orbit at higher altitudes than LEO satellites, the average footprint is greater for each MEO satellite.

Telecommunications carriers such as VoIP service providers and Internet service providers use satellite links for voice and data communications network delivery where additional capacity, route diversity, or delivery to remote areas is required. The ubiquitous nature of satellite communications makes it an ideal candidate for providing VoIP services.

Standards organizations contributing to Intelsat Earth Station Standards (IESS) include the ITU and the IETF.

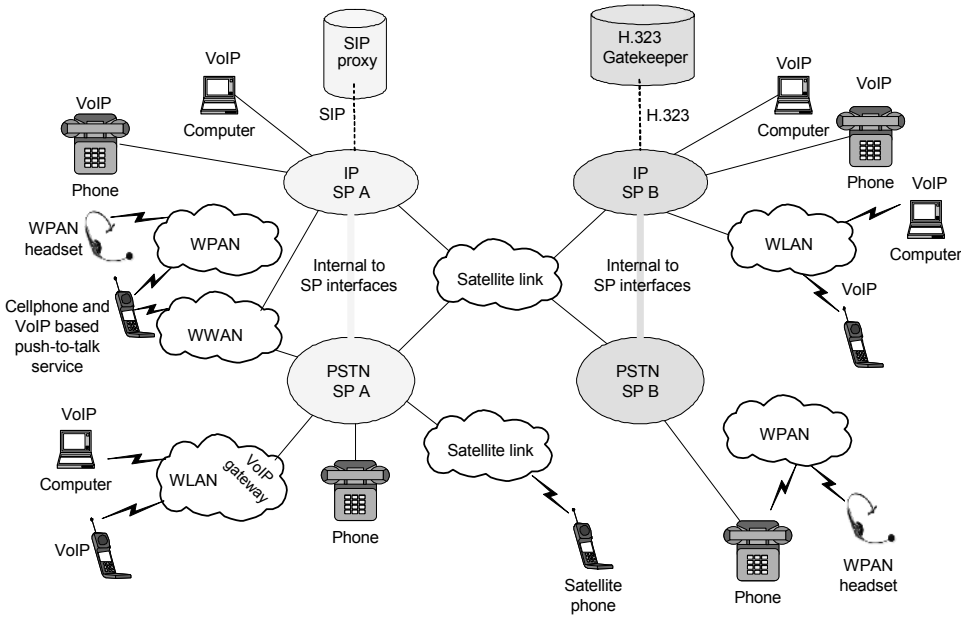

Figure 5. Wireless Interoperability Network Diagram

Analysis

Many different wireless systems and architectures exist today, ranging from wireless personal area networks (WPAN) to wireless local area networks (WLAN), fixed terrestrial wireless local loop (WLL) networks, wireless mobile cellular networks, and satellite systems.

IP, which is already a universal network-layer protocol for wireline packet networks, is becoming a universal network-layer protocol over most wireless systems. An IP device with multiple radio interfaces or a software radio could roam between different wireless networks if they all support a common IP network layer.

A key challenge for all-IP wireless networks is how to support seamless mobility between different wireless architectures. Seamless mobility is the ability of the wireless systems to support fast wireless data handoffs between normally inoperable wireless network elements (e.g., cellular network base stations and WLAN access point radios) with low delay and minimum to zero packet loss.

An additional challenge is to provide for IP-based authentication, authorization, and accounting (AAA) on a WLAN. When a mobile user attempts to access a public WLAN, the access point must make sure the mobile user is authorized to access the WLAN and can be properly charged for services rendered. Simultaneously, the mobile user must make sure that the WLAN is trustworthy and is certified by his or her service provider. Also, both the user and the WLAN must make sure that the transmission between them is secure, so that no one can fake the user's identity to gain unauthorized access.

Wireless propagation between mobile devices and Earth-bound wireless networks, or between Earth ground stations and mobile satellite devices using satellites in geostationary, LEO, or MEO orbits introduce a communication delay which is greater than experienced through wired networks.

Wireless communications links may also exhibit high bit-error rates. In some cases, this problem may be partially solved at lower communication protocol layers. However, practically all VoIP call control mechanisms on the Internet interpret packet loss as a sign of congestion, so this can pose a real problem.

Wireless communications links also suffer from a large bandwidth delay product. Data that is transmitted but not yet acknowledged by the receiver is considered “in flight.” The transmitter often waits for packet acknowledgements to return from the receiver prior to sending additional packets. Therefore, devices using IP cannot fully utilize the data throughput capability of the wireless link.

For VoIP protocols, the delay induced by voice compression algorithms, network communication protocol stacks, and wireless signal propagation can be in the range of between 150-800 msec, larger than the ITU recommendation of a maximum 150 msec for VoIP connections. The caller (or client) may experience a response time delay which considerably degrades the interactivity of the call or service. From this perspective, selecting SIP rather than H.323 for VoIP wireless applications may help minimize delays.

The major factor in perceived speech and service quality is from delay induced by wireless signal propagation time. This factor is independent of the VoIP protocol employed.

Hybrid satellite-terrestrial routes, when compared to terrestrial-only networks, may have relatively longer delays, variable error rates, and lower speeds. However, VoIP by means of satellite can overcome most of these limitations through appropriate protocol selection, call control, and echo cancellation techniques.

Care should be taken to minimize network congestion points, which can lead to increased round-trip delay. Provisioning sufficient equipment to handle maximum call volumes minimizes queue-processing time in VoIP access devices.

Gaps Identified

The inability of users to roam across WLAN hotspots as they can with cell phones today highlights the need for a common WLAN hotspot architecture that is based on open standards and that is acceptable to the various WWAN and WLAN service provider communities. Such an architecture must also be flexible enough to accommodate users with a variety of mobile device form factors and login credential types as well as enable service providers to implement a variety of billing models.

Intra-city roaming for WLAN users will be required if providers are to expand the use of their WLAN hotspots. Unless a common roaming framework is deployed, WLAN hotspot deployment in urban areas is unlikely to be monopolized by individual operators or operator communities, which would limit the available footprint for WLAN users.

Mobility management within WLAN networks is currently achieved through proprietary access-point-to-access-point handoff protocols. IP-based mobility management involves redirecting IP packet flow to the mobile's current point of attachment whether WLAN or WWAN based. The goal of an IP-based handoff scheme between WLAN and WWAN networks is seamless mobility–the ability of the WWAN and WLAN networks to support fast wireless data handoff between normally inoperable wireless network elements with low delay and minimum to zero packet loss.

Recommendations

Wireless architectures (WWAN, WLAN, and satellite) need to evolve to a common IP platform that fully supports end-to-end IP connectivity and integration with a variety of other wireless IP-based network architectures.

The industry needs to develop a common WLAN hotspot architecture that all wireless architecture (WWAN, WLAN, and satellite) operator types can embrace.

Authentication mechanisms and authentication, authorization, and accounting (AAA) signaling between the WLAN hotspot and the various back-end authentication systems of different wireless architecture (WWAN, WLAN, and satellite) operator types must be compatible.

A goal of an IP-based handoff scheme between WLAN and WWAN networks should be seamless mobility—the ability of WWAN and WLAN networks to support fast wireless data handoff between normally inoperable wireless network elements (i.e. cellular base stations and WLAN access points) with low delay and minimum to zero packet loss.

During emergency situations, communications within the emergency response community (police, fire, rescue, local government) and with the public are vital. State and local government officials should develop emergency communication plans using wireless systems. Because satellite footprints cover broader areas than WLANs and WWANs, satellite systems offer a critical component to an emergency communication plan.

3.4 Inter-Provider Interfaces

The scope of this document states that inter-provider connections can be made by means of traditional SS7 or ISDN telephony protocols, possibly with extensions (e.g., BICC) or through VoIP protocols (e.g., SIP or H.323). The telephony protocols are well understood and widely deployed. However, the promise of packet switching is diminished when a packet call has to be converted to TDM to interconnect with another service provider. This section identifies the gaps that may occur where various service providers connect at a packet level.

In the circuit world, a bearer channel is dedicated to a call for its duration. In the packet world, the bearer channel is shared by many calls. In the circuit world, signaling such as SS7 and PRI are constantly aware of the state of the bearer channel (whether active or disconnected). Currently in the packet world, the state of a call may or may not be known.

Connecting at a packet level creates challenges for ensuring end-to-end QoS, traditional billing settlement, and signaling that normally stays within the bearer channel. These issues can be mitigated with inter-vendor service agreements between service providers or through industrywide agreement to a common standard, similar to the industry adoption of SS7.

3.4.1 Quality of Service

This section will discuss the inter-provider issues of QoS metrics and mechanisms. If a provider has contracted QoS for some particular traffic across its domain to the next inter-provider handoff, then the provider must offer it to achieve interoperability. Otherwise, discussions of intra-provider or subscriber signaling for QoS are out of the scope of this document. This includes the use of integrated services (Intserv) and the Resource Reservation Setup Protocol (RSVP) and DOCSIS when used as forms of subscriber signaling for QoS on a per-flow basis.

A **QoS metric,** as described in section 3.4.1.1, is a specific performance or quality goal to be achieved on either a network interface or on an end-to-end basis across a set of networks.

A **QoS mechanism**, as described in section 3.4.1.2, is a method to classify specific network traffic (e.g., VoIP) for specific treatment (e.g., queuing behavior, constrained routing) to enable it to achieve specific QoS metrics.

This document assumes that providers will implement “aggregated QoS” mechanisms and policies on their inter-domain links. In other words, providers will use QoS mechanisms to provide certain levels of performance to classes of traffic, (e.g., VoIP) as defined by the QoS metrics and not employ QoS mechanisms on a per-flow basis. However, a provider may (and in some cases should or even must) implement per-flow QoS mechanisms on access networks, especially if the access network is of lower capacity or can be congested.

The proper functioning of QoS mechanisms that carriers deploy is dependent on the overall availability and reliability of the network. For example, a denial of service (DoS) attack on a provider's underlying IP network route processors, if not mitigated, could render QoS treatment ineffective.

3.4.1.1 QoS Metrics (Aggregate)

Overview

QoS metrics characterize the quality level of a certain aspect of a service in terms of quantified values. QoS metrics can be used by service providers to manage and improve their service offering. They can also be used by the customers (end users or partner providers) in service-level agreements (SLA) to ensure the quality level they expect.

Telecom standards organizations, such as ATIS T1 committees, ITU-T study groups, and the Quality of Service Development Group (QSDG), have been working extensively on quality-assessment methodologies and metrics for traditional voice-band services (voice, fax, and modem) over PSTN. The objective is to produce QoS metrics that are meaningful, validated as accurate, and standardized for industrywide use. These standards are now being enhanced, and new metrics are being developed to meet the interoperability requirements of the emerging, converged networks that use new technologies (e.g., IP, wireless) and offer new types of services (e.g., streaming media, web browsing, e-mail). Collaborative efforts continue among the ATIS, ITU, and other standards groups such as the IETF.

Categories of QoS and Network Performance Metrics

QoS metrics can be primary parameters that are determined by direct measurement of call events, such as noise, echo, packet loss, delay variation, or signaling release cause. Alternatively, QoS metrics can be derived from a collection of primary parameters, for instance

- Statistical calculation (e.g., call completion rate to a given destination for a day).
- Opinion modeling (e.g., Call Clarity Index calculated from call measurements).

Survey of Standardized QoS Metrics

This section provides a survey of existent standardized QoS metrics.

Network Performance Parameters

Building on the initial work of the ATIS T1A1 committee, the new ITU-T Recommendation Y.1540 defines a set of parameters for characterizing IP network performance for network-segment or end-to-end applications. The parameter set includes IP packet transfer delay (IPTD); IP packet delay variation (IPDV), sometimes called "jitter"; IP packet loss ratio (IPLR); and IP packet error

rate (IPER). In conjunction with the accompanying recommendation, Y.1541, for QoS classes, these network performance parameters are useful for supporting SLA management at the inter-provider level as well as at the end-user level.

Call/Session Setup Success

This metric relates to the rate of success in reaching the called party for each call setup attempt, as normally indicated by the signaling release causes. Meaningful statistical metrics can be derived over an aggregate of calls (e.g., calls to a given destination per hour through a given route). ITU-T Recommendations E.425 and E.600 provide definitions of the commonly used answer-to-seizure ratio and network effectiveness ratio. A similar statistical metric is used to characterize the session-setup success rate for the generic IP-based services.

Call/Session Setup Delay

This metric relates to the waiting time to get to the called party after the initial setup request. For PSTN, this is represented by the commonly used post dialing delay (PDD), which is the time between the last dialed digit and the beginning of ring-back, or newer post gateway answer delay (PGAD) as defined in ITU-T Recommendations E.431 and E.437, respectively. Target values for call setup delay are specified in ITU-T Recommendation E.721. For the new IP-based networks, generic “session setup delay” is similarly defined.

Conversation and Voice Quality

This metric relates to the conversation or voice quality during the call, after the call connection is established. Conversation or voice quality can be affected by parameters such as noise, echo, talker volume, latency delay, and impairments caused by voice compression, packet loss, and delay variation. In particular, two-way interactive conversation quality is critically affected by latency delays. For the IP-based networks, the achievable voice quality is critically determined by the available bandwidth associated with types of voice codec used for transmission and their corresponding robustness with respect to IP-domain impairments such as packet loss and jitter (see section 3.4.3). ITU-T Recommendation G.113, table I.1, summarizes the achievable voice quality in terms of equipment impairment factor (*Ie*) for a number of commonly used voice codecs at different operating rates (see section 3.4.3, table 3).

A number of standards that relate to conversation or voice quality.

- **Subjective Evaluation:** The most direct way to assess voice quality is through subjective evaluation methods, as specified in ITU-T Recommendations P.800 and P.831, using a mean opinion score (MOS, 1 = bad to 5 = excellent). Because subjective evaluation is costly and time-consuming in practice, objective psycho-acoustic models are often used to estimate user-perceived MOS.

- **Call Clarity Index:** ITU-T Recommendation P.561 defines in-service non-intrusive measurement devices (INMD) for measuring voice-grade parameters (speech level, noise, echo, and delay) from live calls. ITU-T

Recommendation P.562 describes the Call Clarity Index (CCI), a conversation opinion model that transforms call parameters into two MOS indices to characterize the two-way conversation quality.

- **Transmission Rating R-factor:** ITU-T Recommendation G.107 ("The E-model") generates a transmission rating R-factor (0 to 100) based on parameters pertaining to the characteristics of voice circuit, packet transmission, and voice encoding. An accompanying recommendation (P.833) provides guidance on impairment effects caused by various voice encoders.

- **Perceptual Evaluation of Speech Quality:** ITU-T Recommendation P.862 provides a standardized psycho-acoustic model (PESQ) for assessing speech listening quality (MOS) in a test call, capable of detecting impairment effects of compression, packet loss, and delay variation. New non-intrusive objective models are also being evaluated in the ITU for assessing speech quality in live calls.

Fax Transmission Quality

Fax transmission quality is important for business applications, especially in an international environment. Fax transmission QoS parameters are defined in ITU-T Recommendations E.4xx (e.g., E.458 for figure of merit of fax transmission, E.459 for non-intrusive fax transmission performance metrics, and E.460 for specific fax performance metrics for V.34 Group 3 fax).

QoS Classes and Performance Objectives

Classes of QoS have been defined to facilitate QoS management for service and business applications. The following are examples of QoS class definitions provided by standards organizations:

1. VoIP SLA Classes: ETSI TIPHON TS 101329-2, "Definition of Speech Quality QoS Classes," provides guidelines for narrowband VoIP QoS classes (4 = high, 3 = medium, 2 = acceptable, and 1 = best-effort/no-guaranty) in terms of transmission rating R-factor, speech quality (equivalents of known voice-codec quality), and end-to-end delay. A new QoS class has been recently added for the wideband voice service.

2. End-User Multimedia QoS Categories: A new ITU-T recommendation (Recommendation G.1010) specifies different multimedia QoS categories from the end user's perspective. Performance considerations are addressed in terms of three parameters (delay, delay variation, and information loss) for different service applications, including

    - **Audio:** Conversational voice, voice messaging, high-quality streaming audio.

    - **Video:** Videophone, one-way video.

- **Data (Interactive or Delay Sensitive):** Web-browsing (HTML), transaction (e-commerce, ATM), command and control, interactive games, and remote access (such as telnet, SSH).

- **Data (Asynchronous):** Bulk data, image transfer, e-mail, Usenet, fax.

One note is that G.1010 seems to be a good framework for discussion, but some of the datapoints on bandwidth use may need further input from the operator community.

Analysis

In the past few years, the industry as a whole has invested significant efforts in developing and enhancing QoS metrics for the emerging converged networks, building upon the vast experiences gained from the traditional voice-band services. For example, ITU-T has designated its Study Group 12 (SG12) to be the lead QoS Group, supported by other study groups such as SG2, SG9, and SG13, with a clear focus on VoIP quality and VoIP/TDM interoperability. The IETF OPS area and T1A1 deal with network performance and QoS issues, and they have driven much of the ITU progress; also, they have addressed IP-related network reliability and restoration issues, which have gone essentially untreated in the ITU to date. There also has been an increased collaboration among ATIS, the ITU, and the IETF on QoS metrics standardization. The release of the latest ITU-T Recommendations Y.1540 and Y.1541 represents a significant milestone in specifying a useful framework for the IP QoS parameters and performance targets for different QoS classes. Understandably, such a framework will be continuously enhanced and perfected as the industry gains more experience from the new and dynamically evolving IP-based services.

In summary, it is fair to say that the industry is basically on track regarding QoS metrics standardization for the emerging interoperability requirements between TDM and IP networks.

Gaps Identified

One gap that has been identified is in the development of inter-domain metric mechanics and common data sets. This may be an issue for each bilateral relationship to negotiate privately, as IP network measurement is widely disparate between carriers, but an effort should be made to see if some standardization activity (such as a standard data interchange format) is possible in the appropriate standards and operational forums (such as the ITU, the IETF, and North American Network Operators' Group [NANOG]).

As pointed out in the preceding section, the industry is basically on-track on the standardization of QoS metrics to support interoperability between TDM and IP networks. The basic framework will be continuously enhanced as the industry gains more experience from the existing and emerging services. Regarding interoperability between wireline and wireless networks, however, it has been pointed out in ATIS T1P1 that the QoS-related 3rd Generation Partnership Project (3GPP) specifications currently under consideration are not compatible with ITU-T QoS specifications such

as Recommendation Y.1541 and will therefore hinder interoperability between the two. These concerns need to be addressed.

Recommendations

Standards bodies such as ATIS, the ITU, the IETF, and the 3GPP must collaborate closely to harmonize the views of the telephony and packet segments as well as the wireline and wireless segments of the industry.

The operational and standards bodies should investigate the possibility of creating a standardized inter-domain QoS metric data interchange format. For example, the extended Real-Time Control Protocol (RTCP) from the IETF could be employed to report on loss and delay variation between service provider and/or enterprise-controlled VoIP gateways.

Further work is needed for the harmonization of QoS specifications for wireline networks (e.g., in the ITU and ATIS) and those for wireless networks (e.g., 3GPP).

3.4.1.2 QoS Mechanisms (Aggregate)

Overview

Although QoS metrics provide the means for providers to determine whether negotiated inter-domain QoS requirements are being met, the mechanisms are the tools that will actually provide the ability to meet those requirements on an inter-domain link.

There are two basic approaches to implementing QoS mechanisms on a given link: one is through provisioning and the other one is through technical means.

1. Provisioned QoS Mechanism

Some providers may find that it is more efficient to provision the QoS requirements for the most stringent subset of traffic rather than classify the traffic and treat each class of traffic differently.

Provisioning is simply providing enough bandwidth on the network that queuing effects are within the QoS metrics for the most stringent case under all expected operating conditions. On a link basis, base QoS metrics (latency[1], loss, and delay variation) can be derived from the queuing characteristics of the link in question. Queuing characteristics are defined by the speed of the link and interface, distribution of the packet size in the offered traffic, and percentage of use of the

[1] Latency can also be affected by forwarding performance of the link termination gear (switch or router). However, in most current-generation equipment, this performance is close to, if not matching, the line rate of the circuits constituting the link. This minimizes the effect of forwarding latency on the link.

link or component. The first two factors contribute to serialization delay (the speed in which a given packet can be emitted onto the link in question) and the last contributes to the likelihood of a given queue being occupied by a packet when another is presented for transmission. This use can be referred to as the "effective bandwidth" of the link or component.

If a provider provisions an effective bandwidth that meets all of the contracted QoS metric requirements across the network, then QoS mechanisms are not required.

Because the provisioning mechanism is inherently an intra-provider activity, interoperability is moot and therefore it is not necessary to give it further consideration in this document, other than to caution network providers that a network is not static. Hence, the provisioning of the network needs to be periodically reevaluated to ensure that the provisioned bandwidth is sufficient to meet all contracted QoS requirements.

2. Technical QoS Mechanisms

All technical QoS mechanisms involve specification actions taken by VoIP equipment and some or all intermediate routers. For an IP or Multiprotocol Label Switching Protocol (MPLS) network, several standards-based approaches can be taken, two of which are differentiated services (Diffserv or DS) and Voice over MPLS (VoMPLS).

Differentiated Services

RFC 2475 defines the Diffserv architecture in terms of characteristics of packet transmission in one direction across a set of one or more nodes within a network. Therefore, Diffserv is inherently asymmetric. Characteristics can be statistically defined by throughput, delay, delay variation, and measures of loss and of relative priority. The approach taken for Diffserv involves a component involved with forwarding data that is separate from that employed by control components, such as routing, policy administration, and configuration.

The Diffserv architecture defines a unique set of terminology, as illustrated in figure 6. As defined in RFC 2473, a DS-compliant node uses the differentiated services code point (DSCP), the first 6 bits of the type of service (TOS) byte in the IPv4 header or the traffic class byte in the IPv6 packet header, to determine which externally observable per-hop behaviors (PHB) to apply to a packet. A DS domain is a set of contiguous nodes that implement a common set of PHBs, provisioned in a common manner to deliver a per-domain behavior (PDB) (RFC 3086). A DS region is a set of contiguous DS domains that offer differentiated services. A DS boundary node connects by means of a DS boundary link to another DS domain or a non-DS-capable domain. With reference to a particular traffic flow, as shown in figure 6, the DS domain that sends the flow is said to be upstream, while the DS domain that receives the flow is said to be downstream. The upstream DS domain boundary node that transmits traffic is called a DS egress node, while the downstream DS domain boundary node that receives traffic is called a DS ingress node.

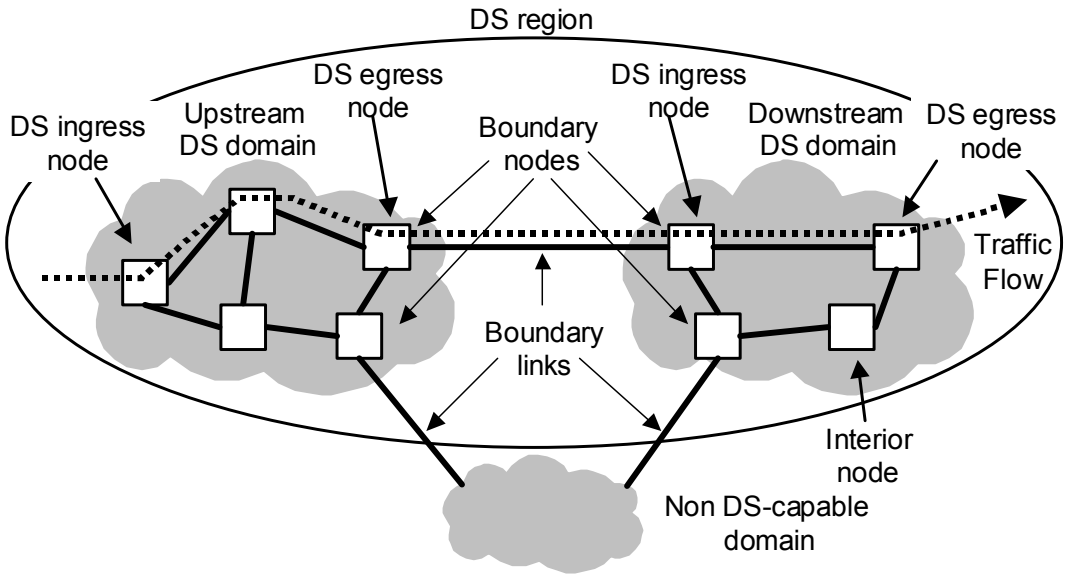

Figure 6. Diffserv Terminology and Reference Model

Typically, ingress, interior, and egress DS nodes perform different functions. These functions include a small set of forwarding PHBs, packet classification, and traffic conditioning functions, including metering, marking, shaping, and policing. In fact, a fundamental tenet of the Diffserv architecture is that scalability is achieved by implementing complex multifield classification and traffic conditioning functions at the edge and then applying the appropriate PHBs within the core solely on the basis of the Diffserv field. The following summarizes some other Diffserv-specific terminology from RFC 2475:

- A DS behavior aggregate (BA) is a collection of packets with the same DSCP value crossing a link in a particular direction.

- A PHB is the externally observable forwarding behavior applied at a DS-compliant network device to a DS BA. At the time of this writing, the IETF had defined 22 PHBs: 1 for expedited forwarding (RFC 2598), 12 for assured forwarding composed in four classes, each with three drop precedence levels (RFC 2597), 8 that operate on a class selector (RFC 2474), and 1 default or best effort (RFC 2474). A PHB group is a set of one or more PHBs that can only be meaningfully specified and implemented simultaneously, for example, the drop priorities of the assured forwarding (AF) PHB.

A PDB is the expected treatment that an identifiable or target group of packets will receive from one edge to another of a DS domain (RFC 3086). A particular PHB (or, possibly, a set of PHBs) and traffic conditioning requirements are associated with each PDB. No PDBs have yet been standardized, but several have been proposed, including an assured rate PDB based on the AF PHB, a virtual wire PDB based on the expedited forwarding (EF) PHB that strives to replace dedicated circuits, and a bulk-handling PDB that is effectively a "less than best effort" class of service.

Analysis of DiffServ

The Diffserv aggregate QoS mechanism is the only standard defined (or in progress) that can scale to support large numbers of VoIP flows. Some industry experts believe that per-flow signaling may be necessary to support end-to-end QoS; however, the current industry direction is to use the Diffserv aggregate method. Operational experience will determine whether this approach meets the QoS metrics needed for VoIP. Per-flow mechanisms, such as those defined in the IETF RFC 2210, "Integrated Services" (i.e., RSVP) or next steps in signaling (NSIS) require the processing of too many messages and retention of too much state in order to scale efficiently. However, such per-flow mechanisms may be used in access networks (e.g., wireless) that are capacity constrained and have a need for tight admission control. In the future, such per-flow mechanisms may be used between providers, for example in packet-wireless network interfaces to wired packet networks. Such signaling may be associated with the path (e.g., RSVP) or may not be associated with the path (e.g., subscription verification and authentication).

If the VoIP packet stream is encrypted (e.g., using IPSec tunnels), then other means to prioritize packets (e.g., port number range) cannot be used. Diffserv avoids this problem as long as the tunnel header uses the DSCP from the tunneled packet.

Diffserv reduces the potential impact of traffic overload DoS attacks; however, if too many Diffserv-marked packets arrive at a network interface, it too will become congested. The Diffserv standards allow a service provider to remark the DSCP. In order for Diffserv to be used across multiple provider networks, service providers would need to agree to not remark the DSCP (or do so in a compatible way) so that subsequent networks in the direction of packet flow can use the DSCP to perform prioritized queuing.

An important distinction between the IP Diffserv architecture and traditional voice or connection-oriented models is the absence of numerical values for QoS parameters because the stated objective of Diffserv is to provide only differentiated performance. Nonetheless, an IP service provider could assign numerical IP performance parameters to a DS domain, and the performance of a concatenation of such domains may be meaningful. The ITU has attempted to quantify IP QoS along these lines, with the results planned for Recommendation Y.1540.

ITU-T Recommendation Y.1541, "Network Performance Objectives for IP-based Services," defines six IP QoS classes (0 through 5) that could be used as a basis for Diffserv classes in a carrier that has chosen technical QoS mechanisms to satisfy its contracted obligations. This ITU-T recommendation defines these classes from the network perspective on the basis of

- Applications (from "real-time, delay variation-sensitive, highly interactive" to "traditional application of default IP networks").

- Node mechanism (from "separate queue with preferential servicing, traffic grooming" to "long queue, drop priority").

- Network techniques (from "constrained routing and distance" to "any route/path").

For each QoS class, IP network performance objectives are defined in terms of value ranges (upper bound) of measured IP network parameters: IPTD, IPDV, IPLR, and IPER. Although six classes may be excessive for most carriers, the recommendation could serve as a common starting point for definition of the supported classes of service for interdomain technical QoS.

Gaps Identified on Diffserv

There is no standard for the DSCP (and associated PHB) nor for the PDB that should be used for VoIP.

Recommendations for Diffserv

Best practices should be developed to at least identify Diffserv DoS attacks and describe a means to mitigate them.

A best practice should also be developed for service provider remarks that are not configured to the default Diffserv interface (i.e., best effort) DSCP.

An attempt should be made to standardize DSCP PHB and PDB for VoIP to be used at the inter-domain boundary. The ITU-T has recommended the EF PHB for VoIP in Recommendation Y.1541. However, this recommendation should be coordinated with the IETF, and operational testing should be performed. Also, the operational testing of Diffserv should be tracked to ensure that it is able to deliver the required end-to-end QoS metrics defined in section 3.4.1.1.

3.4.2 Inter-Provider Usage Metering (Reciprocal Compensation)

This section deals with the ability to exchange billing records between two carriers connected through traditional TDM (PSTN) and VoIP connections. The need to exchange billing records is often required by regulation or bilateral agreements. The support for specific end-user billing functions by a particular provider that places a requirement on another provider is out of the scope of this document, as described in section 2.

Overview

The 1996 Telecommunications Act requires that incumbent local exchange carriers (ILEC) and competitive local exchange carriers (CLEC) interconnect for the purpose of exchanging traffic between the two networks. Currently, this is established in two ways: interconnection at the end office and interconnection at the tandem.

Also, state public utilities commissions have established mechanisms for carrier-to-carrier compensation. This is accomplished through “bill and keep” or where each carrier bills the other for terminating each other’s traffic. In the case of bill and keep, each carrier bills its end users and keeps the revenue. There is no need for carrier-to-carrier compensation and thus there is no need for each carrier to send the other any billing records other than for determining the access charges.

The other scenario however, requires both interconnecting carriers to do the following:

- Declare the percentage of local use.
- Measure the amount of traffic being exchanged in minutes of use.

In the current TDM networks, the originating carrier is able to record and generate a billing record or a call detail record (CDR) in the originating class 5 switch. This record is then formatted to the Ordering and Billing Forum (OBF) standard and exchanged between the two carriers in order to determine the compensation amount.

When a traditional PSTN (assume an ILEC) and a VoIP provider are interconnecting, there may be a need for carrier-to-carrier compensation. The ILEC is able to capture and record both the billing record and the CDR. The VoIP provider should have the capability to do so also. However, several questions need to be answered concerning billing for VoIP. What is a CDR in a VoIP network? Is it the IP address or URI of the device generating the VoIP call; is it the SIP proxy; or is it a phone number? What does “minutes of use” translate to? How can a carrier determine the jurisdiction of the VoIP call for proper billing?

It is possible to use a SIP proxy or an H.323 gatekeeper to provide a billing record in the OBF format. It is also possible to have SIP act as a finite state machine as opposed to a stateless protocol. This configuration could be used by an ILEC or CLEC to produce a CDR for an IP network that is as reliable as a TDM CDR.

Analysis

Currently, neither the standards bodies or any regulatory agency require the VoIP provider to capture these records.

The Internet Protocol Detail Record (IPDR) organization (www.ipdr.org) is a standards body defining detailed records in the Network Data Management - Usage (NDM-U) format. This body has defined detailed records for the following services:

- VoIP.
- Content settlement.
- Wireless application roaming.

In addition, the ATIS OBF has a working group formed to translate the IPDR record into the Exchange Message Interface (EMI) record. The EMI record is the most common method for carrier-to-carrier compensation.

The settlement protocols and testing protocols are out of scope for this document.

Gaps Identified

Currently, there is no standard for mapping specific events or timepoints (e.g., location, calling time, calling number, disconnect time) of a VoIP protocol into a billing record.

Recommendation

NRIC VI Focus Group 3 recommends that the industry support the ATIS OBF working group to translate the IPDR record into the EMI record.

NRIC VI Focus Group 3 recommends that the OBF establish guidelines for a mechanism as well as for the format required for the exchange of these records.

3.4.3 VoIP Encoding (PCM/TDM)

Overview of IP Encoding Standards

VoIP encoding encompasses the standards that define how analog voice is encoded into a digital stream, which can then be placed into IP packets. Pulse Code Modulation (PCM) is the oldest such standard, as specified in ITU-T Recommendation G.711; however, several other packetized voice coding standards (table 3) (as defined by the ITU-T or other bodies identified in the column “Standard”) are either in use or are being considered. Although historically, the ITU-T has been the owner of voice coding standards, the best VoIP codecs may not come from the ITU-T, as indicated in the table. Important attributes of the coding standard are the peak bit rate and the algorithmic delay, which are included in the table. Another important factor, which is not included in the table, is whether the implementation supports silence suppression, which often reduces the average bit rate to half that of the peak rate.

The subjective perception of packetized voice is dependent upon the choice of algorithm as well as on the performance of the underlying IP network. In particular, the packet loss average rate and burstiness and delay can have an impact on the subjective perception of voice quality.[2] Operators need to consider implementing VoIP encoding standards and putting in place a means to monitor performance (e.g., loss, delay, delay variation) such that subjective quality is acceptable.

Table 1. Voice Coding Standards

Acronym	Name	Standard	Peak bit rate (Kbps)	Algorithmic delay, ms	Equipment impairment factor, *Ie*
PCM	Pulse Code Modulation Differential	G.711	64	0.125	0
ADPCM	Adaptive Pulse Code Modulation Differential	G.726	40	0.125	2
ADPCM	Adaptive Pulse Code Modulation	G.726	32	0.125	7
ADPCM	Adaptive Pulse Code Modulation	G.726	24	0.125	25
ADPCM	Adaptive Pulse Code Modulation	G.726	16	0.125	50

(continued)

[2] ITU-T Recommendation G.107, "The E-Model, a computational model for use in transmission planning," March 2003.

Acronym	Name	Standard	Peak Bit Rate (Kbps)	Algorithmic Delay (ms)	Equipment Impairment Factor (*Ie*)
LD-CELP	Low-Delay Code Excited Linear Prediction	G.728	16	2.5	7
CS-ACELP	Conjugate-Structure Algebraic-Code Excited Linear Prediction	G.729	8	10	10
MP-MLQ	Multi Pulse–Maximum Likelihood Quantizer	G.723.1	6.3	30	15
ACELP	Algebraic Code-Excited Linear Prediction	G.723.1	5.3	30	19
VSELP	Vector Sum Excited Linear Prediction	IS-54	8	—	20
ACELP	Algebraic Code-Excited Linear Prediction	IS-641	7.4	—	10
QCELP™	Qualcomm Code Excited Linear Prediction	IS-96a	8	—	21
RCELP	Residual Code-Excited Linear Prediction	RS-127	8	—	6
VSELP	Vector Sum Excited Linear Prediction	Japanese PDC	6.7	—	24
RPE-LTP	Regular Pulse Excited Linear Predictive Coding using Long Term Prediction	GSM 06.20, Full-Rate	13	—	20
VSELP	Vector Sum Excited Linear Prediction	GSM 06.10, Half-Rate	5.6	—	23
ACELP	Algebraic Code-book Excited Linear Prediction	GSM 06.60 Enhanced Full Rate	12.2	—	5
ILBC	Internet Low Bit Rate Codec	GIPS – IETF Draft –avt-rtp-ilbc-02	13.3	30	—
ILBC	Internet Low Bit Rate Codec	GIPS – IETF Draft –avt-rtp-ilbc-02	15.2	20	—

Analysis

In general, an encoding standard with a lower bit rate has lower quality as compared with an encoding standard at a higher bit rate. The choice of coding standard also impacts the ability to deliver other nonspeech but voice-band signals, such as tones, modem, fax, and TDD/TTY. The G.711 coding standard supports all of these nonspeech voice band signals. However, many of the other coding standards do not support these nonspeech signals and require an "out-of-band" protocol to transfer them, in particular, modem tones and fax machines, as described in section 3.6.2. Standards efforts are under way for out-of-band signaling for TDD/TTY, as described in section 3.6.2. This is important because the Americans with Disabilities Act (ADA) and FCC Docket 94-102 require TDD/TTY capability over wireline and wireless telephone networks.

One important objective of VoIP is achieving convergence of voice and data on a shared network. The protocol stack carrying VoIP has 40 bytes of overhead per packet (20 for IPv4, 8 for UDP, and 12 for RTP). To achieve better efficiency, the caller must incur additional delay at the transmitter to collect a string of encoded samples and also at the receiver to allow for playback of the voice packets. However, perceived quality degrades if the total of all contributions to delay exceeds approximately a tenth of a second. Therefore, use of an appropriate voice packet size is an important consideration in achieving acceptable quality VoIP at reasonable efficiency. It is also possible to use RTP/UDP header compression on an access network to achieve better efficiency in a bandwidth-limited access network. In summary, the contributions to overall VoIP delay are

- Voice coding algorithmic delay (see table 3).
- Packetization delay—the time to fill a packet with samples for transmission over the packet network. Typically, this is between 5 and 20 milliseconds.
- Serialization or store and forward—the time required to transmit the packets on links, which can be significant on a low-speed access line.
- Switching and queuing delay encountered by voice packets traversing the IP network, which can be many milliseconds.
- Propagation delay, which depends on the distance between the communicating parties.
- Playback buffer (or jitter absorption) delay, which accounts for delay variation caused by the IP network. Packets traversing one or more IP networks will experience variable latency, but the decoder requires packets at a constant rate for smooth playback, so this delay is necessary.

Furthermore, converting from one coding standard to another (sometimes called transcoding) at some intermediate point in a service provider network (e.g., a media proxy) adds another decoding and coding delay, and if the coding standard is not G.711 PCM, then quality also degrades. Therefore, to the extent possible, it is desirable to prevent or at least to minimize the number of coding and

decoding operations in a VoIP call. If transcoding occurs, then performance may not be acceptable.

VoIP signaling protocols support negotiation of the coding standard and its associated parameters, but there needs to be a minimum coding standard and parameters that all VoIP implementations would support to achieve interoperability.

Gaps Identified

Too many standards exist for the packetized coding of voice, as evidenced by the long list in table 3. In some cases, standards, industry consensus, and support for a minimum interoperable subset of voice coding standards with associated parameters need to be established.

Another approach to interoperability between end devices that do not support the same coding algorithm and/or parameters is to decode VoIP packets and then re-encode them in the other format. This has two disadvantages: increased delay and decreased quality. If a minimum interoperable subset cannot be achieved, we need some further definition in signaling protocols or in guidelines for their use in order to enable networks to minimize the number of such VoIP recodings.

Impairments in an IP network (e.g., packet loss, delay, delay variation) can degrade subjective perception of voice quality. Objectives need to be set for these impairments, and there should be a means in place to at least perform sample measurements of these impairments.

Recommendations

As identified above, to achieve interoperability, there is a need for service providers to agree on a minimum interoperable subset for these coding standards (for example, G.711 using a 20-millisecond sample without silence suppression). In order to implement such a recommendation, implementations should always announce support for the default in codec negotiation. TIA 811 is being rewritten to have G.711 be the default codec.

T1A1 should develop a guideline for service providers that minimizes the number of transcodings; otherwise, quality will be degraded and delay increased, potentially to unacceptable levels. If this recommendation were to be adopted, then the urgency of augmenting inter-network protocols to minimize the occurrence of multiple transcodings may not be as urgent.

There should be IETF SIP and ITU-T H.323 signaling standards to indicate that transcoding has occurred for use in codec algorithm selection by intermediate media proxies. Furthermore, there should be standards such that codec transcoding is recorded in call records for purposes of troubleshooting and complaint resolution.

3.4.4 Interoperability With PSTN Station Signaling (e.g., FLASH, DTMF Digits, Point of Sale)

Overview

During a voice call, some communication devices are controlled by the use of tones in the hearing frequency range or switch hook flashes. Thus, VoIP networks have to be able to accurately reproduce these tones or switch hook flashes in order for these types of devices to continue to work over IP networks.

Examples of such applications that use tones are

- Modems, where a subscriber is connecting to the Internet using an analog modem, and an IP network is in the call path.
- Fax machines, where a fax is sent over a network that uses IP components.
- Dual-tone multi-frequency (DTMF-, or touchtone-) controlled voice mail, where a subscriber is accessing his voice mailbox, and part of the call traverses an IP network.
- DTMF interactive voice response systems, where a calling card subscriber dials the access number from a phone to place a long distance call, and the call traverses an IP network.
- Point-of-sale devices, where a customer is purchasing an item at a store that uses a point-of-sale verification device and the verification call traverses an IP network. Point-of-sale devices use modem technology (V.150.1) or ISDN.

Analysis

Many of these types of communication devices have had issues working in a VoIP environment. As a result of these problems, standards development organizations (SDO) have developed solutions to mitigate these issues. Table 2 lists devices and the accompanying standards used to resolve issues in a VoIP environment.

Table 2. Standards for Communication Devices

Device	Standard
Modem	V.150.1
Fax machine	T.38
DTMF	RFC 2833

Gaps Identified

Standards V.150.1, T.38, and RFC 2833 have been presented as solutions; however, SIP was not included as a possible solution for some of the control and signaling features for VoIP. SIP is an important protocol that is becoming widely deployed. SIP is a catalytic protocol that delivers key signaling elements. These elements can turn a VoIP network into a true IP communications network, a network capable of delivering next-generation converged services.

Recommendations

VoIP devices used in networks where circuit-switched phones will be used must support the above standards where appropriate to ensure that these types of devices will continue to interoperate. Service providers should publish the interface requirements for this type of service in order for end users to identify which customer premises equipment is compatible.

3.5 Directory Services

Modern computing and telecommunications networks use many types of directories. For the network to be useful, people and resources on the network must be able to locate each other to establish communication. A directory is a network service that associates two or more pieces of information about people or resources on the network. Directories may be used to locate resources, authenticate or authorize users, or route calls. It is this last use with which we are concerned here—the use of directories by a telephone system to facilitate the establishment of communication.

The most basic directory is the traditional phone book. The white pages associates the names of people or businesses to numbers that have meaning to the phone system. This simple directory is used by people to map a name to a number that they can dial. Once they dial the number, other directories within the phone system interpret the dialed number and map it to the connections and protocols required to route the call. In the traditional PSTN, this function is coded into the physical circuits of switches. With the advent of number portability, a separate directory maps ported numbers to the physical switch circuits.

In TCP/IP networks such as the Internet, a directory called the Domain Name System (DNS) maps a human-readable name (such as www.fcc.gov) to a number, which is the IP address of a web server. This function is not much different from that of the white pages, although the organization of the information differs greatly.

Regardless of their technology, all communication systems require directories to locate stations and route calls. As the PSTN and TCP/IP networks converge, with calls being routed between them, the establishment and maintenance of synchronized directories are among the most basic requirements for interoperability of these disparate networks.

3.5.1 Local Number Portability, North American Numbering Plan

Overview

Local number portability (LNP) is a network capability that allows an end user to change service provider, location, and/or service type without having to change his telephone number. Today LNP is accomplished by using location routing number (LRN) capability. The LRN is a 10-digit number used to uniquely identify a switch or point of interconnection in an LNP environment. The LRN for a particular switch must be in the same format as a native numbering plan area (NPA)-NXX assigned to the service provider for that switch. Essentially, LRN assigns a unique 10-digit number to each switch in a defined geographic area. The LRN serves as a network address. Carriers routing telephone calls to end users that have ported their telephone numbers from one carrier to another perform an SS7-based database query to obtain the LRN that corresponds to the local switch of the dialed telephone number. The database query is performed for all calls where the NPA-NXX of the called number has been marked as portable. The NPA-NXX portion of the LRN is used to route calls to numbers that have been ported. The three types of LNP are

- **Service provider portability** allows an end user to change local SP while retaining his telephone number.

- **Location (geographic) portability** (beyond or outside rate center) allows an end user to change from one geographic area to another while retaining his telephone number.

- **Service portability** allows an end user to change from one service to another (e.g., CENTREX to POTS) while retaining the same telephone.[3]

Today the FCC has limited LNP to service provider portability within a given rate center as designated by a state regulatory authority.

Additional information on number portability can be found in the footnotes.[4]

1. Service Provider Portability

The first type of number portability, service provider portability, is made technically feasible in the PSTN by the LRN method, as described above. This type of portability is confined to within the rate center that the end user's telephone number has been designated. Either the North American Numbering Plan Administrator (NANPA) or the Pooling Administrator (PA) allocates numbers to the service provider based on the specific rate center requested.

VoIP technology uses a database to convert an end-user-dialed E.164 telephone number into a format that can be transported by means of IP. Thus, any call from a VoIP device on the Internet can call any PSTN number through a gateway or vice versa. A VoIP subscriber could port his number to an Internet service provider (ISP) within the rate center that the telephone number has been assigned and connect anywhere on the Internet to place and receive calls using his telephone number. The PSTN still places calls to a ported number in the same manner as it has, using the LRN of the new local provider's switch described above, and then the VoIP networks carry the call to the rate center where the called number was geographically assigned.

2. Geographic Portability

The second type of portability is geographic portability. This gives a subscriber the ability to move outside of a defined rate center to a larger predefined service area (e.g., NPA, statewide, or anywhere in the country) and keep his phone number.

For example, if countrywide geographic portability were authorized by the FCC, a New York number ported to California would have a PSTN caller from California being directed to New York and then the PSTN would carry the call from New York back to California. This can create significant backhaul for some

[3] See FCC Report and order 96286 at www.fcc.gov/Bureaus/Common_Carrier/Orders/1996/fcc96286.txt

[4] www.ported.com/index3.html; www.nanpa.com

scenarios. Also, for this scenario and others, 9-1-1[5] issues are created. Geographic number portability is out of scope for this report.

3. **Service Portability**

   The last type of number portability is service portability. Service portability allows a subscriber to retain his directory number when changing type of service. An example would be changing from plain old telephone service (POTS) to ISDN. Service portability is out of scope for this document.

Number Assignments to Carriers

Numbers are assigned to service providers either by the NANPA or the PA. Telephone number format follows the North American Numbering Plan (NANP). NANP conforms to E.164 guidelines, with North America having a country code of 1. A company is eligible to obtain numbers from NANPA or the PA by providing appropriate certification and facilities readiness per FCC and industry requirements.[6]

State governments establish rate center boundaries for a given geographic location. NANPA then allocates codes to service providers for use within rate centers. The LRN must be selected and assigned from a valid NPA-NXX that has been uniquely assigned to the service provider by NANPA, and that LRN must be published in the Local exchange Routing Guide. [7]

Analysis

Ported calls between the PSTN and the Internet will continue to work as long as the ISP obtains PSTN numbers from a carrier. The established tools and procedures used by carriers for the LRN method will work for VoIP.

Currently, only certified carriers (CLEC, IXC, ILEC, CMRS) obtain numbers from NANPA. Issues would arise if an ISP were to create E.164-like numbers for its subscribers' use outside of the PSTN. Examples of issues are concurrence of routing databases (e.g. ENUM, LNP, toll free), Enhanced 911 (E911) location, and interoperability between PSTN and ISP subscribers same number could be assigned to two different subscribers. This is why a carrier cannot indiscriminately assign telephone numbers outside of rate center designation if there is to be any inter-working between PSTN and the Internet-based service. See sections 3.5.2 and 3.6.3 for more details.

---

[5] "Geographic Portability and 9-1-1," www.ported.com/geopor~1.rtf

[6] Numbering Resource Optimization Report and Order and Further Notice of Proposed Rulemaking; FCC 00-104, adopted March 31, 2000, paragraph 91.

[7] NANP, www.nanpa.com
ATIS committee INC, www.atis.org/atis/clc/inc/inchom.htm
RFC 3482 Number Portability, www.ietf.org

Gaps Identified

No gaps exist as long as ISPs obtain their end-user telephone numbers from carriers who meet the assignment criteria, or the ISP becomes a qualified carrier and follows the same assignment guidelines and LNP carrier requirements.

Recommendations

As long as ISPs obtain their telephone numbers from carriers or qualify as carriers and obtain their telephone numbers from the NANPA or PA, there is no impact on service provider portability.

3.5.2 ENUM/DNS

Overview

The Domain Name System is a distributed database accessed by a simple query-response protocol. DNS can be used for a variety of purposes. Its most common use (for which it was created) provides name-to-number and number-to-name mapping for Internet hosts using the TCP/IP communication protocol. The ubiquitous ".com" in web URLs is a DNS construct.

The DNS database is hierarchical in nature, and it is commonly described as a tree, with a single root and many branches. Each branch is called a domain. The "leaves" of the tree are end systems with unique domain names and other attributes such as IP addresses. The DNS database is organized into administrative divisions called zones. A DNS zone is a set of connected domains under a common administrative authority.

ENUM is a scheme that uses DNS domain names to represent E.164 telephone numbers. The numbers are used as indexes to information in the DNS database. ENUM-aware devices can use this DNS information to establish a connection to a device serviced by that number.

ENUM operates in a manner similar to a number portability database. Just as number portability maps an E.164 number to a physical circuit ID, ENUM uses the DNS database to map an E.164 number to a connection specification. But unlike NP, the DNS database is distributed across many servers, each with authority for a small branch of the overall DNS "tree."

An ENUM-enabled client device wishing to initiate a call makes a DNS query for the E.164 number. The query is sent to a DNS server specified by the entity with administrative authority for the client device, which may be a different entity than that with authority for the information requested. The DNS server receiving the query processes it on behalf of the client. If the nameserver processing the query is not authoritative for the zone of DNS data in which the information resides, it may generate additional queries to other DNS nameservers that can answer authoritatively for that information. This process, called iterative name resolution, presents the client device with the illusion of a single, unified DNS database, when in

fact the DNS data is distributed over thousands of servers, each with authority for a subset of the entire namespace.

If the DNS server operating on the client's behalf finds a valid, authoritative answer to the client's query, that information is returned to the client. In the standard Internet DNS, this information is typically an IP address (DNS record type A). In ENUM, the information returned is a record of type Naming Authority Pointer (NAPTR), containing a set of parameters, which the calling device can use to determine where to direct the call and the protocol to use for the connection. Because DNS can map multiple pieces of information to a single domain name, the response may contain multiple NAPTR records, offering a choice of multiple destinations and protocols. ENUM deployment scenarios assume that the user of an E.164 number (or his service provider) will be able to manipulate the NAPTR records for that number to indicate his preferred contact methods. In some deployment scenarios, the DNS information may lead the calling device to initiate a direct connection to the IP address of the called device. In other scenarios, the DNS may point the calling device to a proxy device that mediates the connection.

Implementation of a standardized DNS database supporting ENUM is viewed as a key enabler for VoIP interoperability. Just as the Internet DNS provides a unified global database for the location of network services, the ITU and telecom industry envision ENUM as a global database that could be used by all VoIP devices worldwide for call setup. Other, alternative directory services are also possible, but if ENUM fulfills its vision, those alternative deployments must interoperate with the public ENUM rooted in **e164.arpa**.

Analysis and Summary of Current Activities

1. ITU Activities

ITU-T Recommendation E.164, "The International Public Telecommunication Numbering Plan," defines the numbering system that ENUM implements in DNS. ITU-T Study Group 2 (study period 2001 – 2004) is the focus of ENUM activity in the ITU. Included in this activity is a series of ENUM deployment trials being conducted in various countries around the world.[8]

2. U.S. Activities—Public Sector

The United States Government has not yet "opted in" to the public ENUM system rooted in **e164.arpa**. However, various departments of the executive branch are active in ENUM affairs.

Policy liaison between the United States and ITU is provided by the U.S. Department of State, Bureau of Economic and Business Affairs, International Communications and Information Policy, Office of Multilateral Affairs (EB/CIP/MA). CIP also maintains

[8] Reports from the ITU-sponsored ENUM trials, along with many other ENUM resources, are available on the ITU ENUM web site at www.itu.int/osg/spu/enum/index.html.

an International Telecommunication Advisory Committee (ITAC).[9] ITAC advises the Department of State in the preparation of U.S. positions for meetings of international treaty organizations, develops and coordinates proposed contributions to international meetings as U.S. contributions, and advises the Department on other matters to be undertaken by the United States at these international meetings. The Telecommunications Standardization sector of ITAC (ITAC-T) deals specifically with international telecommunication positions for the United States to be taken at these meetings. The ITU-T deals with standards such as ENUM.

The U.S. Department of Commerce, through the National Telecommunications and Information Administration (NTIA), has been involved in policy issues surrounding ENUM. The NTIA conducted a Roundtable on Convergence of Communications Technologies in August 2002, in which ENUM was prominently featured.[10]

The FCC has a number of network convergence-related activities, of which NRIC VI is one. Most of the FCC's ENUM policy work is focused in the Office of Strategic Planning and Policy Analysis, which has several presentations on the subject.[11]

In February 2003, in letters to the Department of State CIP group, both the Department of Commerce and FCC endorsed the use of public ENUM.[12] These letters recommended a formal "opt-in" by the United States to the public ENUM system rooted in **e164.arpa** and outlined a set of "principles to guide domestic implementation of ENUM." Private sector work on ENUM deployment for the United States is driven by these principles.

3. U.S. Activities—Private Sector

The IETF (described below) and ENUM Forum are the focus of U.S. private sector activity on ENUM. The ENUM Forum was created in accordance with the recommendation of the July 6, 2001, report developed by ITAC-T, Study Group A Ad Hoc on ENUM.[13] The ENUM Forum is an open industry group whose membership comprises companies with an interest in VoIP and ENUM. The primary mission of the ENUM Forum is to develop the implementation framework for deploying ENUM for E.164 numbers within the United States and a potential common implementation with other countries served by the NANP.

The ENUM Forum has a number of task groups addressing various issues raised by ENUM deployment. In March 2003 the ENUM Forum released a major document,

[9] See U.S. State Department International Telecommunication Advisory Committee web site at www.state.gov/e/eb/adcom/c668.htm.

[10] See NTIA Roundtable on Convergence of Telecommunications Technologies at www.ntia.doc.gov/forums/enum2002/index.html.

[11] See presentations on ENUM by J. Scott Marcus, FCC Senior Advisor for Internet Technology:
"A Perspective on ENUM," www.fcc.gov/opp/enum.ppt and
"Challenges of Convergence," www.fcc.gov/opp/challenge.ppt.

[12] Full text of both letters at www.fcc.gov/commissioners/powell/gross_enum_letter-021303.pdf.

[13] See ENUM Forum home page at www.enum-forum.org.

"Specifications for US Implementation of ENUM."[14] This document is the baseline specification for ENUM deployment in the United States.

The ENUM Forum continues its work on the many issues surrounding provisioning and management of U.S. ENUM information. The U.S. Government acknowledged this ongoing work in an August 2003 joint letter to the ENUM Forum from the FCC, Department of Commerce, and Department of State.[15]

Although there are as yet no commercial implementations of ENUM, many companies are researching it and participating in its development through the IETF, the ENUM Forum, and other industry bodies. There is currently no ITU-recognized national ENUM trial in the United States, but a number of private trials are under way as the industry refines the technology and vendors prepare product offerings.

5. Internet Standards Activities

ENUM is a specialized extension to the Internet DNS protocols. Like all protocols used on the Internet, DNS is defined by the Internet standards process.[16] The basic DNS protocol has been used on the Internet since the late 1980s. Over time, DNS protocol extensions have added features relating to data management and security, some of which may apply to ENUM deployment. The ENUM protocol extensions are a more recent addition. All of these extensions are the subject of IETF working groups. Some working groups of particular relevance to ENUM and its DNS implementation are

- DNS Extensions (DNSEXT). [17]
- Domain Name System Operations (DNSOP). [18]
- Telephone Number Mapping (ENUM).[19]
- Provisioning Registry Protocol (PROVREG). [20]

These working groups are taking various ENUM-related RFCs through the Internet standards process. The Internet Architecture Board (IAB) is collaborating with the ITU on ENUM issues and has recommended the use of the DNS domain **e164.arpa**

---

[14] Document available at www.enumf.org/documents/6000_1_0.pdf.

[15] Full text of joint letter at www.ntia.doc.gov/ntiahome/ntiageneral/enum/enumletter_08132003.pdf.

[16] For a full explanation of this process, see RFC 2026, "The Internet Standards Process," at www.ietf.org/rfc/rfc2026.txt. The DNS is standardized by Internet Standard 13, which is composed of RFC 1034 (www.ietf.org/rfc/rfc1034.txt) and RFC 1035 (www.ietf.org/rfc/rfc1035.txt).

[17] IETF DNS Extensions Working Group web site at www.ietf.org/html.charters/dnsext-charter.html.

[18] IETF DNS Operations Working Group web site at www.ietf.org/html.charters/dnsop-charter.html.

[19] IETF Telephone Number Mapping Working Group web site at www.ietf.org/html.charters/enum-charter.html.

[20] IETF Provisioning Registry Protocol Working Group web site at www.ietf.org/html.charters/provreg-charter.html.

for ENUM provisioning.[21] This creates the basic Internet DNS structure necessary for standardized, interoperable ENUM deployment. Although the ITU has not formally accepted the IAB recommendation, as of mid-2003, 13 ITU member nations have "opted in," committing themselves to the use of **e164.arpa** for ENUM representation of the E.164 numbers under their country codes. The United States is not one of these, but it appears to be heading in this direction.

Many open issues exist concerning the management of the information to be stored in the **e164.arpa** ENUM domain and the coordination of information between ENUM registries and "alternative deployments" of other ENUM domains and number mapping databases. Although the PROVREG Working Group is addressing the underlying protocols for communicating information between registries, many of the open issues are outside the scope of the IAB and the Internet standards process and are being worked in other forums. These are discussed in more detail below.

Analysis of ENUM Deployment Issues

ENUM presents many complex deployment and provisioning issues. Most of these have nothing to do with the ENUM or DNS technology itself but rather with the administrative processes required to manage the information contained in the ENUM DNS database. Because of its implications for privacy and security, there is also increasing interest in ENUM by private groups involved in the creation of public policy. The industry is responding to these concerns in the ENUM Forum, IETF, and in other public and private forums.[22] A full analysis of these issues is beyond the scope of this report, which confines itself to the issues directly affecting interoperability.

In the following discussion, it is important to draw a distinction between types of ENUM deployments. The IAB-recommended ENUM deployment using specified DNS protocol extensions and a DNS tree rooted at **e164.arpa** is called the "public ENUM." Any deployment of the ENUM protocol using any other DNS tree, or not directly connected to the public **e164.arpa** tree, is a "private ENUM" deployment. Any alternative deployment that provides ENUM functionality but does not use the DNS protocol specified for ENUM is an "ENUM-like" deployment.

Provisioning and Data Management Issues

The Internet DNS is a single tree, with a single root domain. Control of that root and the domains immediately below it (the top-level domains, [TLD]) rests with ICANN,

---

[21] For background on this decision, see RFC 3245, "The History and Context of Telephone Number Mapping (ENUM) Operational Decisions: Informational Documents Contributed to ITU-T Study Group 2 (SG2)," March 2002 at www.ietf.org/rfc/rfc3245.txt.

[22] Some examples of the public policy concerns raised by ENUM may be found at
www.cdt.org/standards/enum/030428analysis.pdf,
arxiv.org/ftp/cs/papers/0110/0110018.pdf, and
www.ietf.org/internet-drafts/draft-ietf-enum-privacy-security-01.txt.

The ENUM Forum's "Specifications for US Implementation of ENUM" document also addresses many of these issues.

[22] See ICANN home page at www.icann.org.

The Internet Corporation for Assigned Names and Numbers.[23] ICANN delegates administrative authority for TLDs to other administrative bodies. This delegation of authority is both an administrative action, whereby the responsible organization is identified, and a technical implementation per the DNS protocol, in which specific IP addresses are identified as authoritative nameservers for the domain. These nameservers are responsible for all information in the DNS database for that domain and all its subdomains and must respond to DNS queries for that information.

Administrative authority for the TLD **arpa** rests with the IAB. Authority for the public ENUM domain **e164.arpa** has been delegated by the IAB to the RIPE Network Coordination Center (NCC).[24] ENUM implementation architectures identify tiers of responsibility for managing ENUM information. In this hierarchy, the RIPE-NCC is the Tier 0 Registry (responsible for **e164.arpa** ENUM TLD). The ENUM architecture defines the Tier 1 Registry as the responsible party for managing DNS ENUM information for a specific country code, or portion thereof. It is expected that Tier 1 subdomains of the **e164.arpa** ENUM domain will be delegated by RIPE to various national authorities in accordance with the country codes defined by E.164. Interim procedures for this delegation have been established between the ITU and RIPE-NCC. These procedures are intended to verify that any requested delegation of a country code in **e164.arpa** has been requested by the national regulatory authority of the country in question. It is expected that the interim procedures will eventually be replaced by an ITU-T Recommendation.[25]

Management of the DNS data in the delegated Tier 1 subdomains of **e164.arpa** will be the responsibility of the designated national regulatory authorities, in accordance with international telecommunications agreements, and local laws and policies. In most cases these national authorities have yet to be identified or their management processes defined.

The reference architecture assumes that Tier 1 Registries will delegate authority for Tier 2 subdomains to various entities who will have the responsibility for actually managing the DNS information on behalf of the users whose numbers fall within those subdomains. In the United States, the Tier 2 registries would manage the ENUM data for the U.S.-based NPA codes under Country Code 1 and the number blocks within those NPAs.

Of particular concern to U.S. deployment, the Tier 1 entity (or entities) that will manage the public ENUM information for the NANP (Country Code 1) has yet to be identified. In the traditional telephone system, a numbering plan administrator is designated for each country code. The NANPA has this responsibility for Country Code 1. The United States shares Country Code 1 with a number of other nations,

---

[23] RIPE is one of the four Regional Internet Registries that manage IP addresses worldwide. A description of the Regional Internet Registries and their role in Internet management is available at www.iana.org/ipaddress/ip-addresses.htm.

[24] For details on the delegation of **e164.arpa** to RIPE, see
Joint IAB-ITU statement announcing the decision (May 2002),
at www.iab.org/Documents/enum-pr.html.

IAB statement on liaison to RIPE-NCC concerning management of e164.arpa (Sept. 2002),
at www.iab.org/Documents/sg2-liaison-e164-sep-02.html.

and management of the DNS information under the domain **1.e164.arpa** and its subdomains must be coordinated between them. The relationship of the current NANPA to the Tier 1 Registry for **1.e164.arpa** is still to be determined.

The interoperability of ENUM with the PSTN will be governed by the extent to which these (as yet undefined) entities are able to coordinate their activities with each other and the carriers who manage E.164 numbers for the PSTN.

A key provisioning issue is the ability of DNS servers to process the DNS protocol extensions used by ENUM. These extensions include NAPTR records and DNS Security Extensions (DNSSEC). Both are relatively recent extensions to the DNS protocol, and a modern version of DNS code is required in order to process them. DNS servers used by VoIP devices should run a DNS implementation that supports NAPTR records and DNSSEC. Wide deployment of ENUM-enabled VoIP devices may require some network managers to upgrade their DNS servers to provide this support.

Another provisioning concern is related to DNS performance. DNS nameservers with authority for ENUM domain information should be provisioned so as to make the service continuously available and process queries in a timely manner. The IETF has published Best Current Practices, which provide guidelines for provisioning of critical DNS nameservers.[26] Many of these guidelines apply to provisioning of ENUM servers as well. However, because of the distributed nature of DNS, the response time seen by a client is highly dependent upon local factors. These include

- The DNS provisioning for the local network where the client resides.
- The robustness of the connectivity between that local network and the network where authoritative ENUM servers reside (e.g., the Internet). This includes factors like bandwidth, link utilization, and latency.

Under some circumstances, these factors may impact VoIP devices to the point where DNS lookup delays may cause calls to fail. To prevent this, any network where VoIP devices reside should be engineered to provide robust and highly available DNS performance.

Alternative Deployments (General Discussion)

Another risk to interoperability is posed by the potential fragmentation of the ENUM namespace. Some countries have expressed a desire to manage their ENUM information in private DNS domains separate from the designated public domain **e164.arpa**. Some commercial entities are advocating use of alternative domains as well. Although it is technically possible to put private ENUM and ENUM-like data in any DNS domain, any approach that attempts to fragment ENUM data into multiple domains will increase the difficulty of presenting end users with a single, unified directory for VoIP.

[26] Reference the following IETF Best Current Practices:
BCP 40 (RFC 2870), “Root Name Server Operational Requirements” at www.ietf.org/rfc/rfc2870.txt.
BCP 16 (RFC 2182), “Selection and Operation of Secondary DNS Servers” at www.ietf.org/rfc/rfc2182.txt.

This risk is increased by some of the U.S. Government's own positions on ENUM. To illustrate, two statements in the February 2003 letter from the Department of Commerce to the State Department CIP should be noted. These are two of the "principles to guide domestic implementation of ENUM":

> **Preserve opportunity for alternative deployments:** The implementation of ENUM within the United States must not preclude alternative deployments of ENUM or other solutions that may provide competitive alternatives to ENUM.
>
> **Allow for interoperability:** In order to support competition and the emergence of alternative technologies and networks, the implementation of ENUM within the United States should accommodate alternative deployments' interconnection with the ENUM tree.

These two principles, while not directly contradictory, may act in opposition to each other. Because of the nature of the Internet DNS, "alternative deployments" may impede interoperability and undermine the viability of ENUM as a global directory for VoIP. This is because a key requirement for ensuring interoperability of telephone systems is for the interoperating systems to use a common directory service.

Every telephony system requires a directory database in order for the calling party to locate the called party and route the call to its destination. For telephony systems to interoperate, the database must be implemented and used consistently by all parties to a call. If different telephony deployments use different directory services, they *cannot* interoperate unless (1) they are able to use each other's directories or (2) their respective directories are synchronized. Failure to accommodate this will create "islands" of service whose boundaries are defined by the directory service they use. Communication between those islands is possible only if there is a common directory between them or if they share their directories.

In the traditional PSTN, call routing is a function of the numbering plan and the interconnecting switches of the physical circuits. Numbers are geographically assigned to switches, and each switch has tables that contain its numbering plan and the numbers assigned to switches to which it is connected. In effect, the global PSTN directory database is the physical network itself.

With the advent of number portability, the call routing information is moved to an external database, which is maintained by a central authority (the number portability administrator). The database maps ported numbers to LRNs. Each call to a ported number causes the switch to query the number portability database and route the call to the LRN identified in the response to the query.

VoIP introduces yet another abstraction. VoIP call routing requires a directory database to determine the IP address to route a call to. This IP address may be that of the called party or some intermediate device such as a proxy. In the first generation of commercial VoIP products, this database is contained within the VoIP system itself, making it applicable only to the local implementation and non-interoperable with other implementations. To provide global interoperability for VoIP, public ENUM moves the call routing information to a single, global database—the Internet DNS.

Alternative Deployments (ENUM-Based)

The interoperability of the Internet DNS depends on the use of a single unified DNS tree. Public ENUM, being a DNS-based service, must fit within this tree. The designated public ENUM domain, **e164.arpa,** is but one branch of this tree. The use of **e164.arpa** for public ENUM data does not preclude "alternative deployments" from using other domains, and several such deployments already exist. However, all public deployments must be branches of the global Internet DNS tree.

Organizations may, for various reasons, wish to create private implementations of ENUM or ENUM-like services for use within private networks. Such private deployments are out of scope for this document, but they may present interoperability issues if users wish to make VoIP calls outside their private network or to receive VoIP calls originating outside that network. A truly private network has no requirements to exchange VoIP traffic with other networks, but a private network with such requirements must maintain some type of public ENUM information. Such networks may be referred to as "public/private."

For an example of a public/private network, consider a service provider or enterprise that wishes to shield its users' private information from the public but still allow inbound and outbound VoIP calls. This provider's public ENUM DNS might direct calls originating outside its network to contact a SIP proxy on a firewall with a public IP address. The proxy would then use a private ENUM DNS to direct the call to an actual user within the private network. For calls within the private network, the provider might use its private ENUM DNS but refer to the public ENUM DNS for outbound calls to numbers other than its own.

Regardless of the DNS domain actually used for a provider's ENUM data, global interoperability would be ensured if all E.164 numbers are presented to the global public ENUM system as a single, unified namespace. The IAB has recommended **e164.arpa** as the root of that namespace. Although it is possible for ENUM data to reside in any DNS tree, any number mapping information maintained in databases outside of the public **e164.arpa** ENUM DNS tree must be made visible in some fashion to users of the **e164.arpa** tree if the implementer of the alternative database intends for its users to interact with users in the public space. In practice, this will require providers of ENUM services that are not based on the **e164.arpa** tree to make arrangements with the various Tier 2 Registries to populate the corresponding **e164.arpa** subdomains with their information. The ENUM Forum Specifications document refers to this general approach as "interconnected registries" or "referrals."[27] If this approach is taken, interoperability depends upon the degree of coordination between the provider of the alternative deployment and the applicable Tier 1 or 2 ENUM Registries.

The ENUM Forum Specifications document outlines several other possible techniques for providing interoperability between the public **e164.arpa** ENUM domain and other, private ENUM domains.[28] These require either specialized DNS resolver code on ENUM-enabled clients or specialized configurations of the DNS

[27] See www.enumf.org/documents/6000_1_0.pdf, Annex B.

[28] See www.enumf.org/documents/6000_1_0.pdf, Annexes B and C.

servers that service client DNS queries. Not all VoIP users may be able to implement such specialized configurations on their clients or DNS servers. Therefore, alternative deployments that rely on these client-side techniques for resolution of ENUM information in DNS trees outside **e164.arpa** may present risks to interoperability.

Alternative Deployments (Non-ENUM)

The above discussion assumes that the "alternative deployments" are also based on the ENUM DNS. The "principles to guide domestic implementation of ENUM" also specify the need to accommodate alternative deployments that are ENUM-like but are not based on the ENUM DNS protocol extensions. These present a completely different set of interoperability issues.

There is no facility in DNS to allow non-ENUM "alternative deployments" to "interconnect with" the DNS tree. No ENUM-based system can place calls to a non-ENUM system unless its numbers are mapped into the ENUM DNS. Any implementation of any alternative database that must interoperate with ENUM requires that the information from that database be mapped into ENUM so that ENUM-only systems can locate the users of that alternative deployment.

This presents a database synchronization problem, which grows in size with the number of "alternative deployments." Any non-ENUM database must be synchronized with the public ENUM DNS tree to be visible to systems based on ENUM. Lacking this synchronization, equipment vendors must provide support for every possible directory system in their products, and end users or their service providers must select which of these directory services to use for any given call.

LNP Synchronization

A similar data consistency issue exists between ENUM and the LNP database. In order for users to move between PSTN carriers and VoIP providers, the LNP database must be synchronized with the public ENUM. For a VoIP user, the LNP database is used to direct calls originating on the PSTN to the appropriate VoIP gateway for the user's provider, while the public ENUM is used to direct calls originating on VoIP networks. If the user changes to another VoIP provider or back to a PSTN carrier, both the LNP and public ENUM databases must be updated to reflect this change. If they are not synchronized, calls from either the PSTN or VoIP networks will fail to be routed correctly. Flawless synchronization of LNP and the public ENUM is required to ensure interoperability. If additional "alternative deployments" to ENUM are introduced, those must also be synchronized with LNP and the public ENUM. Other number portability databases will have similar issues as well.

In summary, although alternative deployments (both ENUM-based and ENUM-like) are possible, the larger the number of such alternative deployments, the more the data synchronization issues become a barrier to interoperability.

Gaps Identified

As summarized in the analysis above, there are four major gaps in the deployment of ENUM:

1. *Lack of global agreement on use of the IAB-designated **e164.arpa** DNS domain for management of ENUM information.* Universal agreement on a common DNS structure for ENUM is optimum for interoperability. Although the United States has not formally opted in to the use of **e164.arpa** for public ENUM, statements issued to date indicate a strong preference for this course of action. If alternative deployments implement private ENUM in different domains, then interoperability between the **e164.arpa** implementations and the alternative deployments will require the employment of various methods to ensure data synchronization or coordination of their information.

2. *Unresolved provisioning issues for management of the public ENUM DNS data.* For any subdomain of **e164.arpa** corresponding to telephone numbers under the NANP, authoritative DNS nameservers must be provisioned and supported so as to be continuously available without service interruption. The entities responsible for this task have yet to be identified, as do the processes for populating the DNS database, keeping it current, and synchronizing it with alternative private deployments of ENUM in other DNS trees.

3. *Mapping of ENUM to alternative directory schemes.* Interoperation of the public ENUM with any deployment using a private ENUM-like directory requires a method of mapping the private data into the public ENUM. Lacking this, equipment vendors must provide support for all possible directory services, or service providers must implement methods for translating information between their respective directories at call setup time.

4. *Synchronization between LNP and the public ENUM.* To ensure accurate call routing between the PSTN and ENUM-based systems, the LNP database and the public ENUM DNS must be kept synchronized as users move between providers. Other, future NP databases will also have similar requirements.

Recommendations

It is evident that ENUM technology is still evolving, as is its deployment and support infrastructure. The industry is making progress on resolution of many outstanding issues. NRIC FG3 makes the following recommendations:

1. The U.S. Government should continue to encourage, support, and participate in ENUM deployment at the technical interchange level.

2. The United States should formally opt in to the **e164.arpa** global public ENUM DNS domain.

3. Government and industry should work together to assign responsibility for administration of the subdomains of **1.e164.arpa** corresponding to NANP and U.S. telephone numbers. The Tier 1 Registries for Country Code 1 should be identified, and processes for managing this data should be established.

4. Providers implementing alternative deployments to the **e164.arpa** public ENUM should ensure that their deployments provide methods for maintaining the synchronization of their data with the public ENUM and the LNP database, as well as with other applicable Number Portability databases in the PSTN.

5. Managers of networks containing clients who use ENUM should provision their local DNS servers with a modern DNS implementation that supports NAPTR records and DNSSEC.

6. Providers implementing authoritative DNS servers for ENUM domains should provision those servers per the IETF Best Current Practices (currently BCP 40 and BCP 16).

3.6 Safety and Security

Several safety and security systems have evolved over the hundred years of existence of the PSTN. Public safety (E911), CALEA, and TTY capabilities are examples discussed in this section. These systems must be supported or replaced as networks converge and/or evolve to packet networks from circuit networks. Backward-compatible issues arise because of the shift from a finite state signaling system like SS7 to a stateless signaling system like SIP. Backward-compatibility issues also arise because of the shift from the bearer channel being nailed up for the duration of a call to the bearer channel being shared between many calls.

A signaling-compatibility issue could be a call being monitored by CALEA that disconnects without the disconnect message being received at the originating end. In such a case, the CALEA circuit would remain connected even though the monitored call is no longer active.

An example of a bearer channel backward-compatibility issue would be TTY, where using a codec other than G.711 would cause the TTY signals to become garbled and prevent a hearing-impaired person from calling 911 and communicating clearly, using a TTY device.

The solutions to issues like those described above can be resolved through standards development organizations or through policy changes. The intent of this section is to identify these types of issues that pertain to safety and security and identify where the issues are being worked.

3.6.1 Support of CALEA

Overview

The Communications Assistance for Law Enforcement Act, enacted in 1994, was passed to preserve the ability of law enforcement agencies to conduct electronic surveillance in light of changing technology. Electronic surveillance includes interception of communications content (wiretaps) and acquisition of dialing information used to identify origin and termination of a call. CALEA seeks to ensure that carriers will have the technical capability and sufficient capacity to fulfill obligations to assist law enforcement.

Analysis

To achieve compliance with CALEA, carriers must ensure that equipment, facilities, and services used for communications are capable of interception and call identification.

Support for CALEA must be balanced with the protection of privacy interests and the promotion of the development of new technologies and services.

Gaps Identified

No gaps are identified at this time. The ANSI T1.678 Working Group is currently finalizing the requirements for CALEA support.

Recommendations

The use of a session border control function has been suggested as a means of providing control and content session replication for the purpose of supporting CALEA. Session border control involves a media proxy that can replicate the RTP/UDP/IP media stream in response to commands from a signaling controller. The types of signaling controllers that need to support CALEA include SIP proxy, H.323 gatekeeper, media gateway controller, and call management system (CMS). The signaling controller provides access to call control information and the media proxy provides access to call content information. The session border control function must be present in either ingress or egress network and may be present elsewhere.

Also, based on ANSI T1.678 draft, an edge router (or a device attached to it) must replicate a copy of VoIP signaling and media streams because an intruder (hacker) could detect or avoid a signaling and/or media proxy (i.e., session border controller). Also, a VoIP server (i.e., conferencing) may be required to perform replication.

ANSI T1S1 and TIA TR45 are working on developing a standard recommendation for compliance with CALEA. This standard recommendation is targeted for packet-based networks and is numbered J-STD-025B.

3.6.2 Teletype Technology (TTY/TDD)

Overview

From the teletype technology of the mid-1960s, the TTY was developed in the United States in 1964 out of personal need by a deaf physicist named Robert Weitbrecht. By coupling existing teletypewriters to the PSTN, Weitbrecht made the first TTY. Although the code, frequencies, and speed of data transmission would be considered old and slow by today's standards, the TTY in the 1960s nevertheless provided a very dependable tool. The TTY was successful also because it was accessible to non-hearing-impaired persons as well.

TTY technology has for the most part not changed since its inception. Select carriers have added relay services, and the hardware is smaller and easier to manage because of smaller circuit boards. The underlying communication protocols are much the same now as they were in the 1960s. As TTYs got smaller and more inexpensive, public TTY ports gradually became available at schools, airports, shopping malls, and even roadside rest stops. Nowadays, public phones as well as cellular phones can access TTY recipients.

The most current problem in using the TTY to its full potential has been the fact that it uses frequencies, codes, and data transmission speeds that are completely different from those used by personal computers. In other words, personal computer modems are not compatible with TTY modems. Therefore, a computer modem cannot be used to operate a TTY.

Analysis

A major attraction in building VoIP networks in comparison to traditional TDM networks is the potential for bandwidth savings as a result of low bit-rate codec technology. It is this low bit-rate technology that degrades the use of TTYs on a VoIP network. The same is true of computer modems and fax machines. The standards development organizations (SDO) have established standards for the use of computer modems and fax machines over a VoIP network by converting those modem tones to data at the ingress to the VoIP network, carrying that conversation by means of a data stream, and converting back to modem tones at the egress of the VoIP network. Currently, the SDOs have not established the same type of facility for TTY communication, although this is work in progress. ITU-T is working on V.ToIP with the goal to complete a Text Relay standard by February 2004; TIA/TR30.1 is working on TIA-1001, a U.S. interim standard for the transport of TIA-825A (Baudot code) TTY/TDD signals over IP networks. The IETF RFC 2833 is used to convert DTMF tones to text messages, and they are discussing the possibility of adding the two Baudot code tones to this RFC to cover TTY. In addition, according to the TIA, TTYs can operate over a VoIP network that does not employ low bit-rate codecs and is engineered to support ITU-T Y.1541 Class 0 or 1 networks (which results in very low packet loss).

Although it is expected in the long term that TTY technology will be replaced by a newer mechanism for text conversation, it is still necessary to support the large embedded base of TTY users. This user community is dependent on TTYs not only for their personal communication. In times of emergency, TTYs can be used to contact emergency services. All 911 public safety answering points (PSAP) are capable of communicating with TTYs. It should also be noted that, currently, local exchange carriers and digital wireless carriers are mandated to support TTY transmission over their respective networks.

Gaps Identified

The use of standardized low bit-rate codecs to encode and transmit TTY modem tones over a VoIP network could inhibit the use of TTYs on a VoIP network that employs low bit-rate codecs. The majority of the VoIP terminals in use today have no mechanism, manual or automatic, to recognize TTY tones and shift to a TTY-friendly codec to accommodate the TTY call.

Recommendation

It is recommended that the SDOs finish their work in the area of TTY transmission on VoIP networks. This is currently work in progress, with an expectation of completion

within 12 months. Each SDO has raised the priority of this issue within its organizations and has already held several meetings this year on the subject matter. Once this work is complete, a review of this issue for further action will be necessary.

When supporting TTY on a VoIP network, service providers should use a codec that encodes TTY to a performance level equivalent to or better than G.711 until the appropriate SDOs have finished their work.

3.6.3 E911 VoIP Interoperability

Overview

The three-digit telephone number "9-1-1" has been designated as the "universal emergency number" for citizens throughout the United States to request emergency assistance. It is intended as a nationwide telephone number and gives the public fast and easy access to a PSAP.

Enhanced 911, or E911, is a system that routes an emergency call to the 911 center closest to the caller and automatically displays the caller's phone number and address.

Analysis

The 911-network infrastructure within the United States was established many years ago with the then-current technologies and practices. Unfortunately, for the most part, the technologies that were used have not changed significantly. Still in use today are analog centralized automatic message accounting (CAMA) trunks (using in-band signaling) and external databases to link the caller's phone number to a physical location. A few advances have been made to update the trunking to digital while still supporting the external database for location lookup.

The key for providing the caller's location is the lookup into the external database (automatic location database) using the caller's number (automatic number identification) that was provided during call setup. The updating of the external database has been a limiting factor to affording new technologies functional access to the 911 networks because of dependence by PSAPs for that information.

The influx of wireless telephones and subsequent connection to the 911 networks has afforded some technological advances. Because wireless telephones are mobile, the use of a static database for location reference did not work, so the wireless industry developed a mechanism to provide real-time information about the caller's location.

The National Emergency Number Association (NENA) has developed a proposed mechanism to support a mobile (roaming) telephone within an enterprise environment while preserving the existing static database architecture in use at the PSAP. In NENA's model legislation for multiline telephone systems, it has suggested using a static NANPA number to describe a geographic location and use the private

call switching mechanism to manipulate the outgoing calling party number information to reflect this geographic location (www.nena.org/9-1-1TechStandards/TechInfoDocs/MLTS_ModLeg_Nov2000.PDF).

NENA is suggesting that the geographic location be known as an Emergency Response Location (ERL) and the corresponding NANPA number be known as an Emergency Location Identification Number (ELIN). This NENA-proposed mechanism does require the enterprise to use dynamic calling party number-capable trunks to the local exchange carrier for E911 calls and for the enterprise to employ an intelligent private switch that has the capability to manipulate the CPN on outgoing E911 calls. This architecture is currently employed and has been successful in early VoIP implementations. Currently, the VoIP SDOs have either outlined architectures based on the NENA proposal (TIA TSB-146) or are working on developing standards-based mechanisms to achieve results similar to those in the current wireless industry (IETF GeoPriv & SIPPING Working Groups).

Gaps Identified

VoIP is similar in some respects to the wireless architecture because VoIP terminals and users can be mobile. As the wireless industry discovered, the use of an external database does not work for the mobile user, with the exception of the controlled enterprise environment described by the NENA model legislation. Similar to wireless phone technologies, VoIP protocols do not make provision for sending the caller's location, as described in the document. There is also the lack of a mechanism for updating the automatic location identification database in a timely enough fashion to support using the automatic number identification as a key into the database for location information for a mobile user.

Recommendations

Several standards bodies and user groups have on-going efforts to design technologies and protocols to meet and/or exceed the current functionality of the 911 networks. The IETF is defining a protocol to pass a user's geographic location information, possibly at call setup. As outlined above, NENA has defined mechanisms to circumvent the current database issues so that mobile VoIP phones can coexist on an enterprise network.

It is believed that the VoIP industry will not only match the functionality of the current 911 infrastructure but will provide a means to enhance that functionality. Short-term mandates may impede longer term enhancements in this area and are not recommended. The challenge of this added functionality is the updating of the existing 911 infrastructures and the funds to do so, in order to access these added functions.

3.6.4 Network Address Translation (NAT)

Overview

Network address translation (NAT) is most often implemented by an entity using private IPv4 addressing, as outlined in IETF RFC 1918. Private IP addresses are defined as blocks of IPv4 addresses that are reserved by the Internet Assigned Numbers Authority (IANA) and not used on the public Internet. By using private addressing, an entity can increase address space on an internal network without fear of overlapping with public addresses. NAT is the mechanism that allows these privately addressed machines to access the public Internet. A NAT gateway (commonly a router or firewall) performs the IP packet header transformation from a private address to a public address so that the return packet can be routed on the public network. The NAT gateway also tracks the state of the header information so that it can perform the appropriate transformation when traffic flows back through the Internet to the private address. Through the use of port address translation (PAT), the NAT gateway can also perform an address-sharing function so that the internal private addresses outnumber the publicly available addresses. The result is that a large number of privately addressed machines can access the Internet using a single public address or a group of public addresses.

NAT/PAT is widely used within private IP networks because of the impending exhaustion of IPv4 address space and because of the flexibility it affords to the administration of internal network addressing. In addition, some view NAT as a security function because the internal private addresses cannot be seen or targeted from the public Internet (except for those tracked by the NAT gateway). This attribute has led some to believe that NAT is the only firewall technology required, which is a false assumption.

IPv6 will prevent the exhaustion of addresses and alleviate the need for private addresses (and NAT). Some believe that private IP addressing has other benefits. The IETF is currently studying these benefits to determine if private addressing (site local addressing) should be allowed in IPv6.

Analysis

Some higher layer applications use communication schemes that cause NAT functions to fail. Some applications embed IP addresses within the upper layer information, where it is normally not examined by a router. If a machine with a private address is operating such an application, the application will fail to communicate properly with a receiver located on or across the public network. For this reason, NAT gateways need to be “application aware.” NAT gateways must examine packets for this type of implementation and change the upper layer information as it traverses the NAT gateway. This function is sometimes called an application gateway. Manufacturers of NAT gateways are continually updating NAT software to recognize application packets with embedded addresses.

Using address-sharing functionality (PAT), a NAT gateway normally allows only traffic to traverse that has been initiated from within the private network. If an

unsolicited packet is received on the public interface of a NAT gateway, the gateway will not know where to deliver the packet on the internal private network. This attribute will cause applications such as VoIP to fail, as the internal machine cannot be reached ad hoc from the public network. In these instances, an application gateway can be used to determine the internal destination of the packet and to set up the appropriate path to the internal address.

Using another addressing scheme, some applications allow the receiver of the initial application request to contact the initiator with a different layer 4 port number than the initiator originally used. In this case, the application will fail because the NAT gateway will not recognize this new layer 4 port number. However, NAT gateways (or application gateways) can also be made aware of applications that perform this way and allow this type of communication. Again, NAT gateway manufacturers are continually updating their products as new applications are implemented.

SIP, H.323, RTP, and other VoIP signaling protocols may use these problematic addressing schemes. This has forced NAT gateway providers to become application aware in order to perform header transformations without detection by other entities on the Internet.

Gaps

An entity that uses private addresses must also use an intelligent NAT gateway or application gateway for VoIP to work properly en route to and from the public network.

Recommendations

Users need to consider the effects on applications when using NAT.

3.6.5 Firewalls

Overview

Most enterprises and many consumers deploy either a separate device or software as a firewall between their site and the Internet. In some cases, a default firewall configuration may block certain IP-related communications that are necessary to provide VoIP. For example, a firewall may block all UDP traffic and hence block VoIP RTP/UDP/IP media streams.

A firewall that is deployed as part of service provider Internet access or deployed by the end enterprise or consumer is out of the scope of this document. However, a service provider may deploy a firewall on an interface with another provider and, therefore, proper configuration of firewalls and/or support of automatic discovery protocols may be appropriate.

Analysis

Because signaling protocols (e.g., SIP, H.323) usually employ TCP and use well-known standard port numbers, there is usually not a firewall issue with these protocols. Of course, firewalls used between service providers must leave these ports open for the protocols to interoperate.

If the firewalls between service providers block UDP, then the RTP/UDP/IP media stream will be blocked and no VoIP service can be provided. In this case, UDP ports belonging to sessions authenticated by the signaling protocol (e.g., SIP or H.323) must be opened.

Gaps

At this time, there are no gaps as long as the entity implementing a firewall facing another service provider uses an implementation that opens the UDP ports for the media stream based on the signaling information received.

Recommendations

None.

4. Acknowledgements

Many individuals and organizations contributed to the FG3 effort. A list of FG3 participants can be found in section 2.3. In addition, FG3 members asked peer SMEs to review and provide feedback to their efforts. The following organizations and individuals generously volunteered their time, effort, and expertise.

Organization	Reviewers
AT&T	Percy Tarapore, Steve Lind
ATIS	Charles Bailey, SBC Chris Daniel, Leapstone Systems Chuck Dvorak, AT&T Labs Fred Iffland, Bell South Hui-Lan Lu, Lucent Technologies Steve Norby, Qwest Gary Sacra, Verizon Rajiv Shah, Alcatel R. Wohlent, SBC
Cisco	Patrik Fältström
Federal Communications Commission	Jeffery Goldthorp
Lucent Technologies	Bernie Cyr, Terry Jacobson, Andre Beck, Cheryl Blum, Kevin Patfield, Stu Goldman
MCI	Robert Schafer, Karen Mulberry, Henry Sinnreich
Qwest	Phil Linse, Connee Moffatt, Ron Egan, Steve Norby, James Adams, Ben Johnson
SBC	Phyllis Anderson, Alexander Huang, Randolph Wohlert, David Wolter
Cisco	Patrik Fältström

5. Appendixes

Appendix A List of Acronyms

Appendix B Network Reliability and Interoperability Council VI Charter

Appendix C FG3 Mission Statement

Appendix D Automatic Network Management Controls

Appendix E NRIC VI Network Interoperability Best Practices

Appendix F References

Appendix A List of Acronyms

3G	third generation
3GPP	3rd Generation Partnership Project
A6	DNS Resource Record used to look up 128-bit IPv6 Address
AAA	authentication, authorization, and accounting
ACE	ASCII Compatible Encoding
ACELP	algebraic code excited linear prediction
ADA	Americans With Disabilities Act
ADPCM	adaptive differential pulse code modulation
AIN	advanced intelligent network
ALI	Automatic Location Identification
AMPS	Advanced Mobile Phone Service
ANI	Automatic Number Identification
ANSI	American National Standards Institute
APNG	Asia Pacific Networking Group
ATIS	Alliance for Telecommunications Industry Solutions
BA	behavior aggregate
BER	bit error rate
BICC	Bearer Independent Call Control
CALEA	Communications Assistance for Law Enforcement Act
CAMA	Centralized Automatic Message Accounting
CANT	cancel to
CAS	channel-associated signaling
CCI	Call Clarity Index
CCS	common channel signaling
CDMA	code-division multiple access
CDR	Call Detail Record
CGC	circuit group congestion

CIP	Communication and Information Policy
CLEC	competitive local exchange carrier
CMSS	call management server signaling
CPN	calling party number
CS	Capability Set
CS-ACELP	conjugate-structure algebraic code excited linear prediction
DCC	destination code cancellation
DES	Data Encryption Standard
DIG	Domain Internet Groper
DNS	Domain Name System
DNSOP	Domain Name System Operations
DNSEXT	DNS Extensions
DNSSEC	DNS Security Extensions
DOC	U.S. Department of Commerce; dynamic overload control
DOCSIS	Data Over Cable Systems Interface Specification
DOS	denial of service
DS	differentiated services (Diffserv)
DSCP	differentiated services codepoint
DSL	digital subscriber line
DTMF	Dual Tone Multi-Frequency
E911	Enhanced 911
ELIN	Emergency Location Identification Number
EMI	Exchange Message Interface
ENUM	IETF Telephone Number Mapping Working Group and resultant protocol
ERL	Emergency Response Location
ETSI	European Telecommunications Standards Institute
FBI	Federal Bureau of Investigation
FCC	Federal Communications Commission
FEC	forward error correction

FG3	Focus Group 3
GETS	Government Emergency Telecommunications Service
GIC	Group Identification Code
GK	gatekeeper
GPS	Global Positioning System
GSC	group signaling congestion
GSM	Global System for Mobile Communications
GW	gateway
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IAB	Internet Architecture Board
IANA	Internet Assigned Numbers Authority
ICANN	The Internet Corporation for Assigned Names and Numbers
IEEE	Institute of Electrical and Electronics Engineers
IESG	Internet Engineering Steering Group
IESS	Intelsat Earth Station Standard
IETF	Internet Engineering Task Force
ILBT	Internet Low Bit Rate Codec
ILEC	incumbent local exchange carrier
IMT	International Mobile Telecommunications
IN	intelligent network
INF	information
INMD	in-service non-intrusive measurement devices
INR	information request
IP	Internet Protocol
IPCC	International Packet Communications Consortium
IPDR	Internet Protocol Detail Record
IPDV	IP packet delay variation
IPER	IP packet error rate

IPLR	IP packet loss ratio
IPSAT	Internet Protocol satellite
IPSec	Internet Protocol Security
IPTD	IP packet transfer delay
ISDN	Integrated Services Digital Network
ISM	industrial, scientific, and medical
ISP	Internet service provider
ISUP	interconnect support; ISDN User Part
ITAC	International Telecommunication Advisory Committee
ITAC-T	Telecommunications Standardization (sector of ITAC)
ITU	International Telecommunication Union
ITU-T	Telecommunication Standardization Sector
IWF	interworking function
IXC	inter-exchange carrier
LAN	local area network
LDAP	Lightweight Directory Access Protocol
LD	long distance
LD-CELP	low-delay code excited linear prediction
LEC	local exchange carrier
LNP	Local Number Portability
LRN	Location Routing Number
M3VA	Message Transfer Part 3 – User Adaptation Layer
MA	Office of Multilateral Affairs
MAP	Mobile Application Part
MG	media gateway
MGC	media gateway controller
MOS	mean opinion score
MPLS	Multiprotocol Label Switching
MP-MLQ	Multi Pulse-Maximum Likelihood Quantizer

MRTG	Multi Router Traffic Grapher
MSF	Multiservice Switching Forum
MSO	Multiple System Operator
MTP	Message Transfer Part
MTP3	Message Transfer Part 3
NANOG	North American Network Operators' Group
NANP	North American Numbering Plan
NANPA	North American Numbering Plan Administrator
NAPTR	Naming Authority Pointer (RFC 2915)
NAT	network address translation
NCC	Network Coordination Center
NDM-U	Network Data Management for Usage of IP-based services
NENA	National Emergency Number Association
NGN	next-generation network
NIST	National Institute of Standards and Technology
NOTIFY	extension to DNS protocol defined in RFC 1996
NPA	Numbering Plan Area
NRIC	Network Reliability and Interoperability Council
NSC	national switching congestion
NSIS	next steps in signaling
NTC	national trunk congestion
NTIA	National Telecommunications and Information Administration
OBF	Ordering and Billing Forum
PA	pooling administrator
PAT	port address translation
PDC	personal digital communications
PDB	per-domain behavior
PCM	Pulse Code Modulation
PDD	Post Dialing Delay

PESQ	Perceptual Evaluation of Speech Quality
PGAD	Post Gateway Answer Delay
PHB	per-hop behavior
POTS	plain old telephone service
PRI	Primary Rate Interface
PROVREG	Provisionary Registry Protocol
PSAP	Public Safety Answering Point
PSTN	The Public Switched Telephone Network
QCELP	Qualcomm code-excited linear prediction
QoS	Quality of Service
QSDG	Quality of Service Development Group
RAS	Registration, Admission, and Status Protocol
RBL	Realtime Blackhole List
RCELP	residual code-excited linear prediction
RF	radio frequency
RFC	request for comments
RIPE	Réseaux IP Européens
RPE-LTP	regular pulse excited linear predictive coding using long term prediction
RR	reroute
RSVP	Resource Reservation Setup Protocol
RTCP	Real-Time Control Protocol
RTP	Real-time Transport Protocol
RTSP	Real-Time Streaming Protocol
RTT	Round Trip Time
SCPC	single channel per carrier
SCTE	Society of Cable Telecommunications Engineers
SCTP	Steam Control Transmission Protocol
SDO	Standards Development Organization
SDP	Session Description Protocol

SEC	switching equipment congestion
SG	Study Group
SigTran	Signaling Transport
SIP	Session Initiation Protocol
SIP-T	Session Initiation Protocol for Telephone
SLA	service level agreement
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SP	service provider
SS7	Signaling System 7
SSH	Secure Shell
TCAP	Transaction Capability
TCP	Transmission Control Protocol
TDM	time-division multiplexing
TDMA	time-division multiple access
TDD	telecommunication display device
TGC	trunk group control
TIA	Telecommunications Industry Association
TLD	top-level domain
TOS	type of service
TR	trunk reservation
TTY	teletype technology
TV	television
UAC	User Agents as clients
UAS	User Agents as servers
UDP	User Datagram Protocol
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
VoATM	Voice over Asynchronous Transfer Mode

VoIP	Voice over Internet Protocol
VoMPLS	Voice over Multiprotocol Label Switching
VSELP	vector sum excited linear prediction
WG	Working Group
WLAN	wireless local area network
WSP	wireless service provider
WWAN	wireless wide area network

Appendix B Network Reliability and Interoperability Council VI Charter

A. The Committee's Official Designation

The official designation of the advisory committee will be the "Network Reliability and Interoperability Council."

B. The Committee's Objective and Scope of Its Activity

The purposes of the Committee are to give telecommunications industry leaders the opportunity to provide recommendations to the FCC and to the industry that, if implemented, would under all reasonably foreseeable circumstances assure optimal reliability and interoperability of wireless, wireline, satellite, and cable public telecommunications networks. This includes facilitating the reliability, robustness, security, and interoperability of public telecommunications networks. The scope encompasses recommendations that would ensure the security and sustainability of public telecommunications networks throughout the United States; ensure the availability of adequate public telecommunications capacity during events or periods of exceptional stress due to natural disaster, terrorist attacks or similar occurrences; and facilitating the rapid restoration of telecommunications services in the event of widespread or major disruptions in the provision of telecommunications services. The Committee will address topics in the following areas:

1. Homeland Security

(A) Prevention. The Committee will assess vulnerabilities in the public telecommunications networks and the Internet and determine how best to address those vulnerabilities to prevent disruptions that would otherwise result from terrorist activities, natural disasters, or similar types of occurrences.

(1) In this regard, the Committee will conduct a survey of current practices by wireless, wireline, satellite, and cable telecommunications services providers and Internet service providers that address the Homeland Defense concerns articulated above.

(2) By December 31, 2002, the Committee will issue a report identifying areas for attention and describing best practices, with checklists, that should be followed to prevent disruptions of public telecommunications services and the Internet from terrorist activities, natural disasters, or similar types of occurrences.

(B) Restoration. The Committee will report on current disaster recovery mechanisms, techniques, and best practices and develop any additional best practices, mechanisms, and techniques that are necessary, or desirable, to more effectively restore telecommunications services and

Internet services disruptions arising from terrorist activities, natural disasters, or similar types of occurrences.

(1) The Committee will report on the viability of any past or present mutual aid agreements and develop, and report on, any additional perspectives that may be appropriate to facilitate effective telecommunications services restorations. The Committee will issue this report within six (6) months after its first meeting.

(2) The Committee will issue a report containing best practices recommendations, and recommended mechanisms and techniques (including checklists), for disaster recovery and service restoration. The Committee will issue this report within twelve (12) months of its first meeting.

(3) The Committee will prepare and institute mechanisms for maintaining and distributing contact information for telecommunications industry personnel who are, or may be, essential to effective telecommunications service and Internet restoration efforts within six (6) months of the first meeting of the Committee.

(C) Public Safety. The Committee will explore and report on such actions as may be necessary or desirable to ensure that commercial telecommunications services networks (including wireless, wireline, satellite, and cable public telecommunications networks) can meet the special needs of public safety emergency communications, including means to prioritize, as appropriate, public safety usage of commercial services during emergencies.

2. Network Reliability

(A) The Committee will prepare and provide recommended requirements for network reliability and network reliability measurements for wireline, wireless, satellite, and cable public telecommunications networks, and for reliability measurements for the Internet, for reporting within twelve (12) months of the Committee's first meeting.

(B) The Committee will evaluate, and report on, the reliability of public telecommunications network services in the United States, including the reliability of router, packet, and circuit-switched networks.

(C) During the charter of a previous Committee, interested participants recommended that the FCC adopt a voluntary reporting program in conjunction with the National Communications System, to gather outage data for those telecommunications and information service providers not currently required to report outages to the Commission, and voluntary reporting was initiated. The Committee shall: (i) analyze the data obtained from the voluntary trial; and (ii) report on the efficacy of that process and the information obtained therefrom.

(D) Should the Commission initiate an inquiry or rulemaking with respect to any of the above-mentioned issues, the Committee will make formal recommendations as a part of such proceeding(s).

3. Network Interoperability

The Committee will prepare analyses and, where appropriate, make recommendations for improving interoperability among networks to achieve the objectives that are contained in Section 256 of the Telecommunications Act of 1996, with particular emphasis on ensuring "the ability of users and information providers to seamlessly and transparently transmit and receive information between and across telecommunications networks."

4. Broadband Deployment.

The Committee will make recommendations concerning the need for technical standards to ensure the compatibility and deployment of broadband technologies and services, and will evaluate the need for improvements in the reliability of broadband technologies and services.

5. Other Topics

(A) The Committee will make recommendations with respect to such additional topics as the Commission may specify. These topics may include requests for recommendations and technical advice on interoperability issues that may arise from convergence and digital packet networks, and how the Commission may best fulfill its responsibilities, particularly with respect to national defense and safety of life and property (including law enforcement) under the Communications Act.

(B) The Committee will assemble data and other information, perform analyses, and provide recommendations and advice to the Federal Communications Commission and the telecommunications industry concerning the foregoing.

C. Period of Time Necessary for the Committee to Carry Out its Purpose

The Committee will require two years to carry out the purposes for which it has created.

D. Official to Whom the Committee Reports

The Committee will report to the Chairman, Federal Communications Commission.

E. Agency Responsible for Providing Necessary Support

The Federal Communications Commission will provide the necessary support for the Committee, including the facilities needed for the conduct of the meetings of the committee. Private sector members of the committee will serve without any

government compensation and will not be entitled to travel expenses or per diem or subsistence allowances.

F. Description of the Duties for Which the Committee is Responsible

The duties of the Committee will be to gather the data and information necessary to prepare studies, reports, and recommendations for assuring optimal network reliability and restoration of damaged, or impaired, telecommunications services within the parameters set forth in Section B, above. The Committee will also monitor future developments to ensure that network interoperability and network reliability are not at risk.

G. Estimated Annual Operating Costs in Dollars and Staff Years

Estimated staff years that will be expended by the Committee are three (3) for the FCC staff and 12 for private sector and other governmental representatives. The estimated annual cost to the FCC of operating the committee is $200,000.

H. Estimated Number and Frequency of Committee Meetings

The Committee will meet at least two times per year. Informal subcommittees may meet more frequently to facilitate the work of the Committee.

I. Committee's Termination Date

The Committee will terminate January 6, 2004.

J. Date Original Charter Filed

January 6, 1992.

Appendix C FG3 Mission Statement

The mission of the NRIC VI Focus Group 3 is to

> "… prepare analyses and, where appropriate, make recommendations for improving interoperability among networks to achieve the objectives that are contained in Section 256 of the Telecommunications Act of 1996, with particular emphasis on ensuring 'the ability of users and information providers to seamlessly and transparently transmit and receive information between and across telecommunications networks.'"

To achieve this mission, FG3 is recommending the implementation of a set of industry best practices and existing or in-progress standards that address the interoperability of VoIP and PSTN wireless and wireline service provider networks. FG3 will identify gaps in standards or industry best practices against the basic features and functions of telecommunications services.

Appendix D Automatic Network Management Controls

Automatic network management controls respond dynamically to switching office and trunk group congestion and failures. When call attempts in a telephone network rise beyond the capacity of that network, the overall performance of the network degrades. Automatic network management provides real-time surveillance and control techniques to minimize this degradation, optimize call-carrying capacity, and maintain network integrity during periods of stress caused by either traffic overload or failure conditions. Network management centers (NMC) or the SS7 network can perform this function. Several types of NM tools are available:

- Dynamic overload controls (DOC) (code controls).
- Protective trunk group controls (cancel-to, cancel-from, and skip).
- Expansive trunk group controls (reroute).
- Manual network management controls.
- Automatic congestion controls (ACC).

Dynamic Overload Controls (Code Controls)

Code controls limit traffic to destination codes. Code controls are most effective for controlling focused overload, a condition characterized by a surge of traffic from many parts of the network to a single office or destination code.

Protective Trunk Group Controls

Protective controls can be used to inhibit the spread of congestion in the network by restricting normal trunk group access and overflow. Protective trunk group controls include trunk group cancel and skip controls.

Expansive Trunk Group Controls

Expansive controls are used to exploit routing beyond the normal in-chain routes when in-chain routes are busy or have failed and there exists idle capacity in out-of-chain routes. The control that accomplishes this is called a reroute control.

Manual Network Management Controls

Manual network management controls supplement and augment automatic network management controls. Manual controls also provide more flexibility in coping with situations that require human judgment. Manual controls, such as reroutes, can be expansive in nature. Alternatively, they are protective by canceling or blocking traffic that cannot be completed. Manual controls can be activated and deactivated at

NMCs through the system that supports the operation of the NMC or through an on-site NM capability.

There are several types of manual NM tools:

- Code controls
- Call gapping, which regulates the maximum rate at which calls are released toward a destination code.

Common Channel Signaling Network Management

The CCS NM feature provides the basic NM components for CCS:

- CCS DOC/ACC.
- Group signaling congestion (GSC) control.
- Manual trunk group control (TGC).
- ACCs.
- Enhancements to the existing TGCs avoid overflow of calls to the source office, allow for reroute of previously rerouted calls, allow for reroute of inbound international calls, and to automatically cancel hunt for certain elements of spray reroute for a period of time upon receiving a national trunk congestion (NTC) or national switching congestion (NSC) indication.
- Process a GSC indication in a manner similar to a SKIP.
- Provide data on prevailing CCS controls and switching congestion conditions.

The CCS NM feature provides the following NM controls for the CCS7 ISUP protocol signaling:

- DOC.
- Enhancements to existing TGCs to avoid overflow of calls to the source office, to allow for reroute of previously rerouted calls, to allow for reroute of inbound international calls, and to automatically cancel hunt for certain elements of spray reroute controls for a period of time upon receiving a circuit group congestion (CGC) or switching equipment congestion (SEC) indication in the national network.
- Data on prevailing CCS controls.

Signaling capabilities on the SS7, packet switching unit (PSU)-based signaling platform include

- ACC.
- Trunk reservation (TR).

- Cancel to (CANT).
- Cancel from (CANF).
- SKIP.
- Reroute (RR).

Alternate Route Cancellation

The alternate route cancellation (ARC) feature is implemented at the switch that has direct trunk groups to the switch experiencing congestion. The ARC is office selective instead of trunk-group selective. The ARC can provide the following two controls:

1. CANF: Traffic that terminates in a congested switch is not allowed to alternate route through other switches to reach that switch. However, the traffic that is switched through the congested switch is allowed to alternate route. This control restricts calls of a selected level (routine or all levels of precedence) terminating in the congested switch from overflowing from the direct route. The CANF control is provided to reduce the spread of the congestion.

2. CANT: Traffic that does not terminate in the congested switch is not allowed to access the direct trunks to that switch. This control prevents calls from being alternate routed through the switch in congestion to reach their destination offices. Therefore, through-traffic bypasses the direct trunk to the traffic-congested office. This control relieves an overloaded office of traffic that can probably complete by another route.

Both the CANF and CANT controls affect the routing of a call. They can be initiated for traffic of all levels of precedence. They can be removed for either precedence or all traffic. Either one or both of the two controls can be activated to the same office at the discretion of the network manager.

Destination Code Cancellation

The DCC control limits traffic to particular destination codes that are difficult or impossible to reach. With this control, specific calls are routed to a special announcement to free up resources for calls that are more likely to be completed. The DCC is an effective control for a focused overload where a large volume of calls is directed toward one destination.

The DCC control is NNX (first three digits of a telephone number) selective. The DCC can be implemented whether or not a direct trunk group to the affected office exists. The DCC control blocks the call at the point where the control is implemented, before trunk group hunting begins. When a switching office is detected to be in trouble, the DCC may be applied at all other connected switches. This allows calls to be blocked at or near their originations.

A DCC control can be applied for traffic of routine or all levels of precedence. It can also be removed from routine or from all traffic. The blocking applied to the destination office should not be total unless the destination office is completely disabled through disaster or equipment failure.

The DCC code blocking allows the controlled code to be NPA (area code), NNX, NPA-NNX, or NPA-NNX-XXXX. It also allows the network manager to control the rate at which calls are permitted to be sent to the affected code. The code-blocking capability allows simultaneous existence of up to 64 code controls in a switch.

Therefore, the network management personnel can establish a DCC control specifying the maximum rate at which calls are released toward a problem destination code. When the DCC control exists, each call's terminating code is compared with the codes being controlled. If a match occurs, the call is terminated to an announcement, depending on its precedence and the controlled traffic rate.

The network management controls of this feature include:

- ACC: an automatic, prehunt restrictive control, affecting both the switch in congestion and adjacent (connected) switches. It serves to restrict traffic sent to or through a switch, when that switch is in an overload condition.
- TR: an automatic, pre-hunt restrictive trunk group control, having functionality in only the switch where it is activated. TR serves to limit access to outgoing trunks on two-way trunk groups (TG), when the TG is nearly full. TR helps to reduce call volume on a distant switch by shifting traffic away from selected trunk groups.
- CANT: a manual, prehunt restrictive control, having functionality in the switch in which it is implemented.
- CANF: a manual posthunt restrictive control.
- SKIP: a manual prehunt restrictive control.
- RR: a manual, posthunt expansive trunk group control, having functionality in the switch in which it is applied.

Appendix E NRIC VI Network Interoperability Best Practices

The convergence of traditional telephony networks with IP networks such as the Internet also requires a convergence of the engineering practices by which those networks are implemented. Engineering practices for IP networks are established by the Internet Engineering Task Force (IETF), using the Internet Standards Process as described by RFC 2026. This RFC describes the process by which Internet protocols and practices are codified in RFCs.

The Internet Standards Process defines a category of RFCs called Best Current Practices (BCP). These are not standards or directives, but, rather, they are intended as common guidelines for policies and operations for the diverse operators of the interconnected set of IP networks known as the Internet.

Many of the recommendations in these IETF BCPs relate to reliability and security, and as such they are outside the scope of FG3. Certain NRIC Best Practices that relate to interoperability have been distilled from the IETF's BCPs, as well as from other RFCs. However, the IETF is the highest authority on all matters pertaining to the Internet and other IP networks. The Internet and its protocols evolve much more rapidly than do the NRIC Best Practices. Therefore, FG3 makes a general recommendation that all operators and users of IP-based networks, protocols, and applications implement in accordance with current IETF guidelines.

BP Number	Best Practice
6-P-0762	Network Operators should engineer networks supporting VoIP applications to provide redundant and highly available application-layer services. Examples of such services include DNS and other directory services, SIP, H.323, and other application-level gateways. To ensure interoperability, all implementations of such IP-based application protocols should conform to the applicable IETF standards for those protocols.
6-P-0763	Service Providers implementing DNS servers in support of VoIP applications such as ENUM should provision those servers per the IETF Best Current Practices for operation of DNS nameservers: BCP 40 (RFC 2182) and BCP 16 (RFC 2870).
6-P-0764	Network Operators and Service Providers implementing protocols for the transport of VoIP data on IP networks should implement congestion control mechanisms such as those described by RFC 2309, RFC 2914, and RFC 3155.
6-P-0765	To optimize the performance of TCP/IP data transport for VoIP over 2.5G and 3G wireless networks, Network Operators and users of such networks should configure their TCP algorithm parameters according to RFC 3481.

6-P-0766	To achieve interoperability and support all types of voiceband communication (e.g., DTMF tones, facsimile, TTY/TDD), Service Providers should consider using a minimum interoperable subset for VoIP coding standards (for example, TI 811 mandates the use of G.711) in a VoIP-to-PSTN gateway configuration.
6-P-0767	Service Providers implementing a SIP-signaled VoIP network should consider using media gateway controllers according to IETF RFC 3372 BCP 63, "Session Initiation Protocol for Telephones (SIP-T): Context and Architectures," in order to achieve interoperability with SS7/ISUP-signaled TDM voice networks.
6-P-0768	Service Providers implementing a SIP-signaled VoIP network should consider using media gateway controllers that map ISUP-to-SIP and SIP-to-ISUP messages according to IETF RFC 3398, "Integrated Services Digital Network (ISDN) User Part (ISUP) to Session Initiation Protocol (SIP) Mapping" in order to achieve a consistent interpretation of ISUP-to-SIP messaging industrywide.
6-P-0769	Service Providers implementing a BICC-signaled network should consider implementing ITU-T Recommendation Q.1912.5, “Interworking between Session Initiation Protocol (SIP) and Bearer Independent Call Control Protocol or ISDN User Part,” or 3GPP TS 29.163, “Interworking between the IP Multimedia (IM) Core Network (CN) subsystem and Circuit Switched (CS) networks,” to achieve interoperability between an SS7/ISUP-signaled TDM voice network and a SIP-signaled VoIP network.
6-P-0770	Wireless Service Providers who have deployed IS-41 or GSM Mobility Application Part (MAP) signaling networks should consider implementing and using the network management controls of SS7 within their networks.

Appendix F References

Organization	Web Address and Content
Alliance for Telecommunications Industry Solutions (ATIS)	www.atis.org T1A1 Committee Technical and Operations Council (TOPS)
ENUM Forum	www.enum-forum.org
Federal Communications Commission (FCC)	www.fcc.gov
Institute of Electrical and Electronics Engineers, Inc. (IEEE)	www.ieee.org wireless local area networks (WLAN) wireless personal area networks (WPAN) wireless wide area networks (WWAN)
International Telecommunication Union Telecommunication Standardization Sector (ITU-T)	www.itu.int/ITU-T/ SS7, H.323, BICC
Internet Assigned Numbers Authority (IANA) Regional Internet Registries	www.iana.org/ipaddress/ip-addresses.htm
Internet Corporation for Assigned Names and Numbers (ICANN)	www.icann.org
Internet Engineering Task Force (IETF)	www.ietf.org IP, SIP, SigTran, CMSS, DiffServ, DNS, ENUM Request for Comments (RFC) Best Current Practices (BCP)
National Emergency Number Association (NENA)	www.nena.org
National Telecommunications and Information Administration (NTIA) Roundtable on Convergence of Telecommunications Technologies	www.ntia.doc.gov/forums/enum2002/index.html
Network Reliability and Interoperability Council (NRIC)	www.nric.org

North American Numbering Plan Administrator (NANPA)	www.nanpa.com
PacketCable	www.packetcable.com
SIP Forum	www.sipforum.org
Telecommunications Industry Association (TIA)	www.tiaonline.org
U.S. State Department International Telecommunication Advisory Committee (ITAC)	www.state.gov/e/eb/adcom/c668.htm