CHAPTER 12

Initial Network Planning and Design

If you live (or at least work) in the real world, then the success of your voice/data integration project(s) will be judged by more than whether or not the technical solution is eventually achieved. The path to that solution, in terms of time and money, is critical. If you have a cavalier approach to the initial planning phases, there is an increased likelihood that you will be surprised later in the project. And I can tell you from experience that the surprise is not usually good!

Most organizations want to save time and money. For non-technical project managers, this may translate to a reduction in discovery, analysis, and planning phases: "Let's not waste any time—we needed this done last week and we are paying by the hour!"

As a technical project manager, it is your job to resist the intuitive—albeit wrong—reaction of cutting corners on the initial phases. The extra time you spend at the beginning of a project is well rewarded by an easier implementation. The implementation phase is the wrong time to discover fundamental flaws with the design, incompatible interfaces, unsupported features, and so on. It can be very expensive and awkward to correct such problems when a project is far along the implementation path. According to an old adage, "A stitch in time saves nine." This may be updated for our purposes to, "A dollar spent on planning and design, in implementation, will save nine."

Before you place an order for new hardware, there are a number of project steps that you must complete:

- Gathering requirements and expectations for voice services
- Gathering telephony interface and signaling information
- Selecting a VoX technology
- Planning trunk and bandwidth requirements
- Selecting hardware platforms
- Reviewing proposed solutions in terms of requirements

Gathering Requirements and Expectations for Voice Services

It is important that you work with all concerned parties at the beginning of the project to determine the requirements for the project and the measures of success. Specifically, you must address the issues of calling patterns, voice-mail services, voice applications such as call centers and ACD groups, and the voice quality expectations of all concerned parties.

To understand why calling patterns are a concern, consider the following scenario:

> You call the main telephone number of Acme Corporation, and the receptionist transfers you to a person in sales. When you speak to the person in sales, you realize that you really need to speak to someone in marketing. Nobody answers the telephone when you are transferred to marketing, so you are automatically forwarded to voice mail.

This is a fairly normal business communication process, and companies expect their voice systems to accommodate these types of transactions. It is very likely that an integrated voice/data network would perform poorly when placed in this environment. Some may claim that the technology is not yet mature, but it is more commonly a problem with the design.

The inherent problem in this case is that the company may have departments in different locations. This means that every time a call is transferred between departments, it crosses a PBX/router interface and incurs another tandem encoding. Voice quality rapidly degrades as the signal undergoes coding/decoding cycles—current low bit-rate codecs tolerate a maximum of two encoding cycles in real networks. The previous call scenario may require four encodings. Now imagine that the person placing the call is using a mobile telephone; another low bit-rate codec is in the audio path. The best solution in this case is to use a G.726 or G.711 codec, which greatly increases the bandwidth consumption on the network, but preserves voice quality during tandem encodings.

Many network designers would not even consider this issue before implementing the network. The integrated network would be installed, and preliminary tests would indicate that voice quality is good. But they would forget to test call transfers between sites. After about a month, end users complain bitterly.

The point of the preceding story is to "think outside the box" when designing these networks. Do not assume you understand all of the requirements ahead of time. Talk in detail with the end users of the network to understand how they will use it. The project will be more successful if you understand the requirements, voice your concerns in a timely fashion, and address them properly. Among other issues, be sure to address voice mail from the earliest phases of the project. This is almost always a show-stopper issue if there are problems, so get it right the first time. If a centralized voice-mail system is used, you generally cannot use low bit-rate codecs and preserve good voice quality. If you are selling the project based on 8-kbps calls, you better rework your strategy.

In general, be wary of situations that require tandem encoding. Common examples include centralized call accounting, centralized voice-mail systems, and proprietary PBX features supported via transparent common channel signaling (CCS). Possible solutions include using higher bit-rate codecs, and possibly using Q.SIG instead of transparent CCS.

Here is one last thought with respect to voice/data network designs that can cause problems. Some corporations have a hierarchical data network design, based on the classic concept of a corporate headquarters, regional hubs, and satellite offices. All is well in the data-only world, with regional e-mail hubs and data centers. But this design may introduce problems for company-wide, real-time applications. Consider what happens if users in a satellite office call users in a distant satellite office. The audio path must traverse four WAN circuits:

- Originating satellite office to originating regional hub
- Originating regional hub to corporate headquarters
- Corporate headquarters to destination regional hub
- Destination regional hub to destination satellite office

The combined effect of transmission delays and variable queuing/serialization delays can severely impact voice quality. Even though the data network is well designed, it is not designed well for the addition of voice traffic. There are at least three solutions to this problem:

- Collapse the hierarchy to a hub-and-spoke topology from the corporate headquarters.
- Add extra circuits from satellite offices to corporate headquarters.
- Install circuits between offices with identified high-traffic patterns.

Each of these solutions has technical, logistical, and economic implications that are nontrivial. There are no easy fixes for these types of problems, but you are in a better position to manage them if you identify them early in the process.

You have read some of the ways your project can go awry. Hopefully, you will not repeat the mistakes presented here. Think through the consequences of your design decisions, and talk with others about issues of concern.

Gathering Telephony Trunk and Signaling Information

Early on in the project, you must evaluate your role. If you are working with a strong telecom department that has specific ideas about how to do things, then you might not want to dictate designs and requirements to them. If you alienate these folks, then your job will be more difficult. On the other hand, some organizations very badly need someone to take charge and figure out what needs to be done. These extremes are important to identify because it affects how you gather the required telephony information.

If you have a strong telecom department, you can rely on them to provide the information you need, such as what type of PBX ports will be connected to the router, and what signaling types to use. Make sure to provide feedback to them on key issues, such as how to divide number blocks for the address plan.

Be prepared to make decisions if you have a lot of unanswered questions, or you get responses such as, "We can do either," or "What do you want to do?" This is a more common scenario when you are working with telecom technicians who are not responsible for the whole network. They usually program the basic switch options using information provided to them, and leave many options in default settings. In these cases, you can choose which of the available PBX ports to use, how to provision the signaling, who provides clocking, and so on.

Refer to Chapter 2, "Enterprise Telephony Signaling," for a detailed discussion of the traditional telephony signaling types. From a project-planning perspective, you can gather the required information in phases:

- Type of equipment and physical interface connecting to the router
- Software configuration options for interface connecting to the router

It is usually easy to identify at an early stage whether a router will connect to POTS telephones, a key system (KSU), a private branch exchange (PBX), or a voice trunk provided by a carrier. It may be more onerous to determine what interfaces are available on the phone switches at each site, such as analog station cards or trunk cards. If your telecom contacts are competent, then you can work with them to determine the information you need. At least try to determine whether analog or digital interfaces will be used for routers connecting to phone switches or carrier-provided voice circuits. Make sure the telecom contacts know that proprietary digital phone sets cannot be connected to the voice ports on the router. A channel bank can interface to such phones, and connect to serial interfaces on the router for circuit emulation. Table 12-1 summarizes the types of interfaces that you may connect to Cisco multiservice routers.

Table 12-1 *Traditional Telephony Interfaces, and Corresponding Cisco Router Voice Interfaces*

If the PBX or other interface is a . . .	Then connect to this Cisco voice interface:
POTS telephone (residential phone)	FXS
Fax	FXS
Carrier provided: analog loop start/ground start (1 MB,and so on)	FXO
Carrier provided: digital T1/E1—CAS (DID, both-way, and so on)	Digital T1/E1
Carrier provided: digital T1/E1—PRI (to support E911, and so on)	Digital T1/E1—PRI

Table 12-1 *Traditional Telephony Interfaces, and Corresponding Cisco Router Voice Interfaces (Continued)*

If the PBX or other interface is a . . .	Then connect to this Cisco voice interface:
PBX/Key: analog station card (normally for fax or POTS phone)	FXO
PBX/Key: analog trunk card (normally for loop start/ground start to CO)	FXS
PBX/Key: analog E&M tie line (normally to remote PBX via CO)	E&M
PBX/Key: digital T1/E1—CAS or CCS	Digital T1/E1
PBX/Key: digital T1/E1—PRI	Digital T1/E1—PRI
Vendor-specific business set (proprietary digital phone)	none
Proprietary digital phone via telco channel bank (or other CSU)	Non-voice T1/E1 (use CES or PBX)

Often, you will have to ask questions of people who have no idea what you are talking about. When you try to find a more technical contact, you might get responses such as, "Our phone switch has been running fine for the last five years. I think the company that installed it went out of business." In these situations, your best bet is to partner with a PBX service company that has experience with the specific model of PBX or key system. In the absence of such partners, you might have to read many manuals. Worse yet, you might have to solve the enigma black box. If you are working with an ancient and questionable phone switch, it may be less expensive to purchase a new one than to spend a lot of time or money to connect it to a router. You can burn hour after hour troubleshooting, only to determine that the old phone switch is not behaving properly. "It seemed so close! It was working yesterday . . . " On the other hand, make sure you have configured everything properly (and reset interfaces) before you blame an old phone switch.

TIP If you are considering a new phone switch, it might not be a bad time to look at the Cisco Call Manager. The solution supports traditional calling features and new applications such as voice-mail/e-mail integration, using native IP telephones and simple GUI web interfaces for configuration.

Assuming you can work with the telecom contacts, and the existing phone switches seem functional, then you can proceed with the data collection. This is easier to manage if you complete a form for each site (or router). Using a form ensures that you do not forget to ask for key pieces of information, and provides written documentation of the project progress.

If problems arise during the subsequent implementation, this form may help identify what went wrong.

You are likely to have problems if you are not involved in the form-completion process. It is interesting to note that no matter how detailed, simple, or precise you make the telecom information form, you will still receive baffling results from some sources. Even though it seems like a good place to save time, your time is well spent if you work with each contact to complete the form. Your project will stay on track if you schedule time to work with each contact for this purpose. If you make each telecom contact responsible for completing the form, you will have several weeks of silence from some (or many) sources, followed by a form with blank spots and vague answers to multiple-choice questions. Do not ask how—it just happens.

Making Early Estimates

You may be pressured for cost estimates and budgetary approvals before you have enough information to develop an accurate bill of materials. If this is the case, you can determine budgetary pricing (but not an accurate bill of materials) before you know the exact PBX interfaces with which the routers must connect. You still need to know the basic router platforms, which VoX technology will be used, the number of telephony interfaces required, and whether the interfaces are analog or digital. Because gathering the exact PBX interface information is often time-consuming, you can hasten the budget-approval process by pricing the analog hardware with arbitrary analog voice interface (for example, FXS). The equipment order should not be placed until the actual bill of materials is ready—after the PBX interface information has been gathered. You must take pains to ensure that nobody places an equipment order based on the budgetary bill of materials. I have seen it happen for 50 routers.

Selecting a VoX Technology

There are at least two ways to decide whether VoFR, VoATM, or VoIP is appropriate for your network: (1) compare the relative merits and shortcomings of each technology and weigh these factors based on their importance in your network, or (2) just pick VoIP.

If you choose the more thought-intensive approach, you should consider the following criteria for the decision:

- Reliability
- Scalability
- Quality of Service
- Cost and Complexity
- Feature Support
- Existing WAN Environment

VoFR, VoATM, and VoIP are compared against each of these criteria in the following subsections.

Reliability

Because both VoFR and VoATM are link-layer technologies, they are sensitive to circuit failures. Redundant circuits increase the chance of successful call connections, but any active VoFR or VoATM calls are terminated when a circuit fails. VoIP performs better in this respect because it operates at the network layer. IP packets may be rapidly rerouted around a failed circuit, without causing active VoIP calls to be dropped.

While VoIP is resilient to circuit failures, it is sensitive to routing problems and configuration errors. Both the signaling and audio portions of VoIP rely on the existing routing protocols in the IP network. If a new device that falsely advertises routes is added to the network, then valid destinations may become unreachable. Any routing problems in the network may impact VoIP connections, even if the problems are not caused by VoIP routers. VoFR and VoATM are less likely to be affected by changes in unrelated parts of the network.

If your IP network is stable, then VoIP offers the best overall reliability.

Scalability

The main scalability issue for enterprises with VoX technologies is the management of dial peers and call routing. There are two aspects to managing dial peers that must be considered as networks grow:

- End-to-End Versus Hop-by-Hop Peering
- Number of Dial Peers per Router

End-to-End Versus Hop-by-Hop Peering

Even with default dial peers, VoFR and VoATM require more configuration and maintenance than VoIP. VoFR and VoATM dial peers must be configured at every router along a call path, whereas VoIP only requires the end routers associated with the call to have dial peers. VoIP is simpler in this respect because it builds on the services offered by IP routing protocols that are already a part of the network.

The preceding discussion assumes that PVCs are used for ATM, which is not the most scalable option. The ATM PNNI protocol can operate with E.164-based ATM addresses when creating SVCs, which allows dynamic call routing. However, this approach does not appear to be making significant advances in the market. Most companies that use ATM in the WAN have PVCs; few companies use SVCs between sites over carrier-provided Permanent Virtual Paths (PVPs). Companies can implement an E.164-based SVC solution between their own sites, but generally these networks are isolated from other companies and the PSTN as a whole. Because the technology is not ubiquitous, scalability is limited by market adoption.

Number of Dial Peers per Router

For large networks, it is not feasible to statically configure dial peers in each router. It is difficult to manage the consistency of dial-peer configurations across multiple boxes, the router configuration memory is limited, and pattern searches become inefficient as the number of dial peers increase. Hub-and-spoke networks can extensively use default call routing, but the central site routers may still be burdened. Hierarchical networks can distribute the dial peers, but still require maintenance of the dial peers in many different routers.

For VoFR and VoATM, the hierarchical approach (of which the hub-and-spoke design is a subset) enables call-routing summarization, and is the best option to scale. VoIP can take advantage of H.323 gatekeepers or SIP proxy servers, which enable dial peers to be hierarchical without forcing the VoIP signaling and audio paths across the same hierarchy. The signaling and audio paths of the call are still optimized according to IP routing, so there is no performance degradation for centralizing the dial-peer databases (other than the delay for the lookup request). The ENUM Working Group of the Internet Engineering Task Force (Transport Area) is currently defining a distributed database standard that enables mapping between E.164 telephone numbers (hence the name ENUM) and URLs, followed by a DNS resolution to an IP address. The ENUM telephone number mapping, and the Telephony Routing over IP (TRIP) protocol, will enable VoIP–call-routing scale to the entire Internet.

Quality of Service

Frame Relay offers rudimentary mechanisms to provide quality-of-service assurances. The Committed Information Rate (CIR) provides a working bandwidth guarantee, assuming the carrier does not oversubscribe its backbone too much. However, there is no guarantee for delay or delay variation. During periods of network congestion, latency across a Frame Relay PVC may increase by a factor of 20 or more (based on personal observation of the author). A consequence of this behavior is that enterprises are not empowered to provide end-to-end QoS for VoFR. You can ensure that your equipment is configured properly to enable the best performance for VoFR, but if you have an uncooperative carrier, then voice quality may suffer. Be wary of situations where VoFR traffic must cross multiple carriers' Frame Relay networks, because there is less accountability for poor performance. If you pursue a VoFR solution, you must closely monitor your frame relay provider to ensure that you are receiving the contracted CIR with a reasonable latency.

ATM offers a clear and compelling quality-of-service solution. QoS is a pillar of ATM network design, offering guarantees for bandwidth, delay, and delay variation. ATM is an excellent technology to provide real-time quality of service along with standard traffic requirements. Both VoATM and VoIP can take advantage of the QoS features of ATM.

Current IP QoS solutions meet the needs of VoIP in the enterprise. It is important to realize, however, that end-to-end QoS in an IP network relies heavily on link-layer QoS. It does not matter how high the IP precedence field is set if the packet is stuck in a congested Frame Relay network. For this reason, you must be careful when selecting a link layer for the WAN. The best options are currently ATM or leased lines. In the LAN, excess bandwidth has traditionally been the best form of QoS, but IEEE 802.1p and the RSVP Subnet Bandwidth Manager (SBM) standards now provide other options.

QoS options for the Internet are rapidly evolving. Internet telephony service providers (ITSPs) now offer VoIP-enabled backbones for their customers, but the contracts between service providers are still developing. Clearinghouse and settlement services allow ITSPs to significantly extend the reach of their QoS network, which enables subscribers to place native VoIP calls to more destinations.

Cost and Complexity

VoFR, VoATM, and VoIP all have similar requirements for interfacing with traditional telephony equipment. The differences in cost and complexity are mainly attributed to assuring QoS.

Frame Relay is an economical WAN technology for voice, provided that QoS requirements can be met. Cisco hardware options are available for small installations, and products can scale to large enterprise deployments. For companies that want to switch to VoIP, Cisco products offer a software-only upgrade path from VoFR solutions to VoIP solutions. The conceptual issues for Frame Relay QoS are somewhat challenging, but the actual router configurations are less difficult.

ATM for voice applications is more expensive than Frame Relay, but you get what you pay for. The built-in QoS guarantees across the WAN reduce the likelihood of stumbling blocks during the implementation phase. If the routers are configured correctly and the network is designed properly (that is, there are minimal tandem codecs), then VoATM quality should be good.

VoIP solutions require more router configuration commands than VoFR or VoATM solutions. In addition to configuring link-layer QoS (that is, FRF.12 fragmentation or ATM CoS parameters), VoIP requires network and transport layer QoS options such as RSVP, special queuing techniques, IP precedence, WRED, and so on. While these features require additional experience from a design and implementation perspective, they offer flexibility beyond what VoFR or VoATM can provide. A unified VoIP network can span heterogeneous link-layer technologies.

Feature Support

It is important to consider not only the traditional PBX features, but also new features and applications that are emerging as communication technologies advance.

Traditional Telephony Features

There are few inherent differences between VoFR, VoATM, and VoIP with respect to telephony feature support. Integrated voice and data networks in general support fewer telephony features than proprietary single-vendor telecom environments. Integrated voice/data networks often sacrifice some features across the WAN in favor of reducing toll costs. If your company already uses PBX equipment from a variety of vendors, many features are not available regardless of voice/data integration.

Emerging MultiService Applications

While VoFR and VoATM will remain as technologies for integrating legacy voice networks, VoIP is developing as a native end-to-end solution. Traditional telephony features are supported in hardware and software versions of IP telephones, and multiservice applications are being integrated with these products. Companies that use VoIP to integrate their legacy voice networks with their data networks will be positioned to take advantage of new multimedia applications and communications tools. There is a clear, long-term, strategic advantage to using VoIP for voice/data integration.

Existing WAN Environment

ATM, ISDN, and clear channel TDM facilities provide good performance for VoX technologies. Because these services are circuit-switched technologies, they ensure consistent delay characteristics, which is important for real-time traffic. Frame Relay can work well for real-time traffic, but it is a risky proposition. You have no assurances that performance will remain good for VoFR or VoIP across Frame Relay.

Consideration of the existing WAN environment was very important before VoIP supported FRF.12 and hardware options matured. Still, your WAN environment should heavily influence your selection of a VoX technology. The following WAN types are considered here:

- Frame Relay
- ATM
- Clear Channel TDM
- ISDN

Frame Relay

When this book was conceived, VoFR was the only reasonable option for networks with a Frame Relay WAN. Since Cisco has incorporated support for FRF.12 fragment/interleaving into VoIP platforms, this advantage has disappeared. VoIP can now provide the same level of QoS as VoFR across frame relay networks.

ATM

VoATM and VoIP are available for networks with an ATM WAN. There is no difference in quality of service between VoATM and VoIP over ATM, but there is an important bandwidth and efficiency consideration. If IP RTP header compression is not available for ATM interfaces, then VoATM is more bandwidth-efficient than VoIP across an ATM network. This is because 40 bytes of IP/UDP/RTP header information are included with the data in each VoIP packet, which means that at least two ATM cells are required to transmit each VoIP packet. VoATM encapsulates the voice coder output directly in a single ATM cell.

Another bandwidth issue arises because ATM cells carry a fixed 48-byte payload (minus 1- or 2- byte AAL headers). If the transported data does not fill the cell, then the additional payload space is padded. Considering that a typical VoIP packet (with two G.729 samples) is 60 bytes long without header compression, then it must be segmented into two ATM cells. The second cell carries about 12 to 14 bytes of data, with more than 40 padded bytes! This yields an ATM payload efficiency of 62.5 percent for a VoIP packet. VoIP does support a configurable number of codec samples per packet, and ATM payload efficiency climbs to 94 percent when VoIP is properly tuned. For example, five samples of G.729 at 10 bytes each, along with 40 bytes of IP/RTP/UDP header, yield 90 bytes of data in two ATM cells. Using the G.723.1 codec with VoIP, which would ostensibly save bandwidth, results in an ATM payload efficiency of only 73 percent. Because G.723.1 uses a 30-ms frame, placing multiple frames into a packet to improve the ATM payload efficiency is not feasible because of the additional packetizing delay incurred.

You must address the ATM payload efficiency issue for both VoIP and VoATM, or you might waste much of your WAN bandwidth, support fewer simultaneous calls, and telephony users will experience reduced Grade of Service (GoS).

Leased-Line TDM

For TDM circuits (for example, T1/E1/J1/Y1), any VoX option may be used. Private ATM across leased lines is usually not a good option, because ATM cell headers consume much of the bandwidth. You should only consider VoATM on these circuits if you also have applications such as video using ATM-CES. VoFR was a good option when VoIP hardware options were limited, but now there is no compelling reason to use it. VoIP is a good option because it has low overhead (when using RTP header compression), requires fewer dial peers, and is flexible to integrate with other WAN technologies. On low-bandwidth TDM circuits, VoIP requires Frame Relay or multilink PPP encapsulation to provide fragment/interleaving.

If other parts of your network use a given VoX technology, you should match that technology on leased-line TDM circuits. This will allow you to keep an integrated dial plan with compatible dial peers. When call passing between VoX technologies is available, this will not be necessary. By the time you read these passages, it may already be available.

ISDN

Many companies use ISDN as a backup solution to their primary frame relay, ATM, or clear channel circuits. In some countries—Japan for example—permanent ISDN connections are available. The only VoX technology that is reasonable for these environments is VoIP, which can run with PPP encapsulation. VoIP is a very attractive option when redundancy is required, because ISDN can economically provide fault tolerance. VoFR and VoATM are not compatible with ISDN.

Table 12-2 summarizes the attributes of VoFR, VoATM, and VoIP with respect to the criteria discussed in the preceding sections.

Table 12-2 *General VoX Attributes*

	Vox Technology		
Attributes	**VoFR**	**VoATM**	**VoIP**
Reliability	OK	OK	Good
Scalability	Poor	OK	Good
Quality of Service	Poor	Good	Good
Cost	Good	OK	Good
Complexity	Good	OK	Poor
Support of Telephony Features	OK	OK	Good
Emerging Applications	Poor	OK	Good

Planning Voice Trunk and Bandwidth Requirements

The object of telecom traffic planning is to determine an optimal number of voice trunks to a destination, such that a certain call success rate is achieved during peak traffic intervals. The standard models used in the telecom industry are statistical models developed by A.K. Erlang at the beginning of the 20th century:

- Erlang B
- Extended Erlang B
- Erlang C

Model Assumptions and Applicability

Each of these models assumes that time between received call attempts is random with a Poisson distribution. These models do not apply to environments that receive spikes of call traffic, such as radio call-in contests, or ticketing vendors when concerts come to town. Each of the models differs with respect to how calls are handled when the trunks in question are busy.

The *Erlang B* model is appropriate when there is an overflow path for busy trunks. For example, a PBX may reroute calls to the PSTN if the VoX trunks are all busy. This example assumes that the remote location has Direct Inward Dial (DID) trunks to facilitate direct PSTN rerouting, or else functionality is compromised (for example, a call destined for a specific person may reach an operator when routed via the PSTN). Another example is a PBX with standard tie lines as a backup to the VoX trunks.

The *Extended Erlang B* model is appropriate when there is no overflow path, and the caller hears a busy tone when the desired VoX trunks are busy. This scenario is common when a site PBX has VoX trunks for interoffice dialing, and no alternate routes through the PSTN (that is, remote sites do not have DID and operator intervention is not acceptable). This model accounts for the fact that many users will immediately redial when a call fails, which increases the amount of incoming traffic.

The *Erlang C* model is appropriate if calls are placed in a queue when the VoX trunks are busy. This model is applicable to call centers, which strive to maintain high utilization of call center agents and trunk facilities. Call centers are outside the scope of this book.

Using the Models

The Erlang models require that you provide some of the variables to solve for an unknown variable. Table 12-3 provides definitions for the variables and measurement units that are commonly associated with these models.

Table 12-3 *Definition of Common Terms for Telecom Trunk Planning*

Term	Definition
Erlang	A measure of call volume equal to 1 hour of aggregate traffic. Three calls of 20-minute duration yield 1 Erlang of call traffic.
Centi-Call Seconds (CCS)	100 seconds of calling traffic. 36 CCS = 1 Erlang. (Not to be confused with Common Channel Signaling CCS.)
Lines	The number of provisioned voice trunks that carry traffic. Each analog port is 1 line; a full T-1 CAS port is 24 lines.
Busy Hour Traffic (BHT)	Amount of call traffic (in Erlangs) that must be supported during a peak-traffic reference hour. Use high estimates for conservative trunk planning.
Blocking	Percentage of calls that cannot be accommodated because of busy trunks. A typical blocking design goal is 1 to 3 percent.
Recall Factor	When there is no overflow path for blocked calls, this is the percentage of calls that are immediately retried (for example, the end user redials the destination).

To determine the optimal number of Lines, the Erlang B model requires the Busy Hour Traffic (BHT), measured in Erlangs, and the Blocking fraction for calls attempted during the busy hour. The blocking fraction is a measure of Grade of Service (GoS). The Extended Erlang B requires the same input and the Recall Factor, which indicates how many people redial after hearing a busy tone.

The Erlang C model operates with slightly different variables. Instead of using the BHT measure, the model uses the number of calls per hour and the average call length. Instead of measuring a blocking factor, the model considers how long the callers must wait before speaking with an agent. Table 12-4 summarizes the form and application for each model type.

Table 12-4 *Applicability and Required Information for the Erlang Trunk Planning Models*

Model	Equation Form	Response When VoX Trunks Are Busy
Erlang B	Lines = f(BHT,Blocking)	Overflow to standard trunks or PSTN
Extended Erlang B	Lines = f(BHT,Blocking,Recall)	Terminate the call (user hears busy tone)
Erlang C	Lines = f(CallsPerHour,Duration,WaitTime)	Call added to a waiting queue (ACD system)

The equations to solve these models are not pretty, so network planners use reference tables that contain presolved values for the different traffic models. A few years ago, it would have been appropriate to include such a table as an appendix to this book. Now, there are numerous online tools that perform the calculation for you. Search the Web for Erlang calculator.

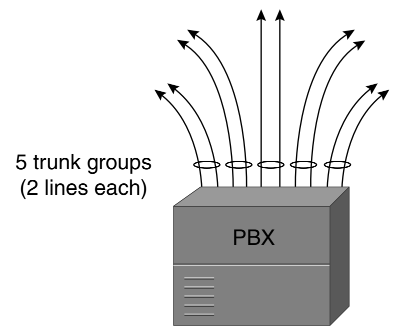
Adjusting the Models for VoX

The way these models apply to real networks is different for VoX networks than for traditional telecom networks. In traditional voice networks, each remote location is reachable via dedicated tie lines. There is no sharing between the lines (tandem switching is excluded for the moment to make a point). In VoX networks, the lines for all remote sites are pooled and connected to a router. Instead of having two lines to each of five different locations, ten lines are all connected to a router, which can appropriately direct the traffic. Figure 12-1 illustrates the difference between tie-line connections in traditional versus VoX networks:

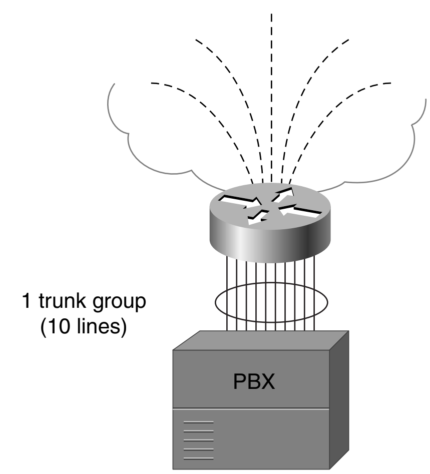
The difference between independent and pooled lines affects the total number of lines required during peak loads. Consider a company with 2 lines dedicated to 50 different offices. A third call to a given office will fail, even if there are 98 idle lines! In a VoX network, all 100 lines could be provisioned for calls to any of the offices. A traditional network may require 3 lines to each office, a total of 150 lines, to accommodate the occasional peak. In the VoX environment, 100 lines may be more than sufficient to meet the total traffic requirements. The assumption here is that the peak traffic load to every destination does not occur at the same time. The ability to reduce the number of lines per site as more sites are added *(statistical multiplexing)* is a hidden benefit of integrated voice/data networks. The companies that benefit most are those with high traffic between many sites.

Figure 12-1 *Independent Tie Lines for Traditional Networks; Pooled Tie Lines for VoX Networks*

a)



b)



There is one concern caused by the pooled line strategy. Consider a hub-and-spoke network, where each of five remote PBXs have three lines connecting to a router, and the central site PBX has ten lines connecting to a router (taking advantage of statistical multiplexing). The first three calls between the central site and a remote site proceed normally, but all successive calls are routed in a suboptimal fashion. Figure 12-2 illustrates what happens for every call when the remote lines are busy:

1 The central PBX receives a call for the remote site and forwards it to the router, because there are free lines.

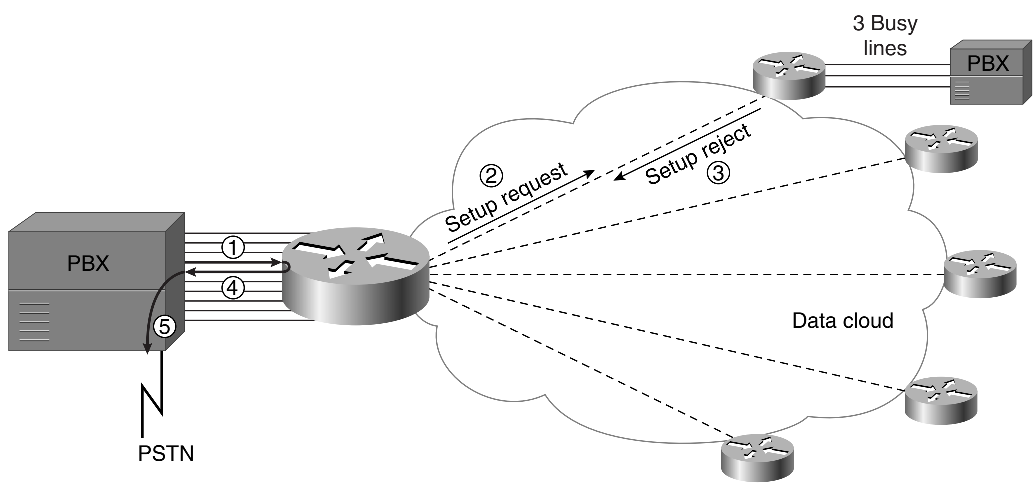
2 The central router accepts the call from the PBX, receives the digits, and attempts to establish the call with the remote router identified by the VoX dial peer.

3 Because the remote router does not have a free voice port (that is, they all are busy), it rejects the call setup request.

4 To prevent dropping the call, the central router should have an alternate dial peer pointing back to the PBX. The same digits cannot be sent back to the PBX because a routing loop would occur between the PBX and the router, which would seize all of the lines. The dial peer that points back to the PBX should add a prefix such as 91555 to the dialed digits, such that the PBX completes the call through the PSTN.

5 The PBX completes the call from the router to the PSTN. The call now occupies two lines between the router and PBX for the duration of the call.

Figure 12-2 *The Hairpin, Trombone, or Boomerang Effect When All Lines at a Remote Site Are Busy*



All calls that exceed the planned VoX capacity must exit the central PBX to the router, and immediately return to the PBX before routing to the PSTN. This situation is called a *hairpin*, *trombone*, or *boomerang*. It is undesirable because it uses two lines between the router and PBX for each call. There is no performance degradation from multiple codec cycles, because calls between local voice ports on a Cisco router are not compressed. But the wasted lines are an efficiency concern.

The solution to this problem is to have more communication between the router and the PBX, such that the path can be dynamically optimized (for example, cut the router out of the loop). This is a common feature between PBXs from the same vendor, but it has not historically been supported between vendors. QSIG is designed to fill this gap—to provide a standard telephony signaling protocol for vendor interoperability. QSIG must be supported in the PBX and in the router to take advantage of interoperable signaling.

Using the trunk or PLAR-OPX connection modes on Cisco routers is another way to solve this problem. The local router is able to refuse connection requests from the PBX when it knows the status of the remote VoX sites (via the PLAR-OPX or trunk signaling). This solution comes at the expense of dedicating voice ports on the router to specific remote destinations, which may not be the most efficient use of the voice ports.

Back to the trunk-planning issue. For VoX networks, the Erlang models should not be applied separately for traffic to each remote site. Rather, apply the models to the total traffic volume to all sites during the combined busy hour. Keep in mind that the busy hour for the combined traffic to all sites may not match the busy hour to any specific site. Applying the Erlang models to the pooled traffic (as opposed to independently for each site) yields a fewer number of required lines, in harmony with the statistical multiplexing advantage. This does not account for the hairpin effect, where two lines are required between the router and PBX for each call that is rerouted to the PSTN. To rigorously account for the hairpin effect, the Erlang traffic models should be modified.

In practice, you can observe the traffic that reroutes to the PSTN during the combined busy hour (for example, on a Cisco router, show dial peer for the peer that matches rerouted traffic, or review the call detail records), and perform an Erlang calculation on this traffic to determine the number of additional lines required for hairpinning. An alternative method is to make an educated guess about the number of extra lines required for traffic rerouted to the PSTN during the combined busy hour.

Converting Number of Trunks to Bandwidth

The Erlang calculations tell you how many voice ports you need on the router and PBX, but you must still determine the amount of bandwidth that each call consumes. This varies depending on the flavor of VoX and the codec. ATM efficiency is variable depending on the payload size. Refer to the section "Selecting a VoX Technology" in this chapter for a discussion of ATM payload efficiency.

The bandwidth calculations are not included here for every combination of codec type, samples per frame, VoX technology, and WAN transport. Instead, tools are provided to help you calculate these values yourself. You can calculate the overall bandwidth per the following equation:

$$(actual\_bandwidth) = (codec\_bandwidth) \times \frac{(payload\_length + encapsulation\_length)}{(payload\_length)}$$

This equation provides the amount of bandwidth required for each call, including the encapsulation overhead. The codec bandwidth values are provided in Table 12-5.

Table 12-5 *Bandwidth Requirements Per Call for Different Codecs*

Codec	Bandwidth (kbps)
G.711	64
G.723.1	6.3/5.3
G.726	16/24/32/40
G.728	16
G.729	8
G.729A	8

The configurable payload length must be an integer multiple of the codec sample size. This controls how many codec samples are placed in each cell, frame, or packet. Because the default values used in Cisco IOS may change, they are not provided here. You should adjust the default value if you need to:

- Increase ATM payload efficiency
- Decrease encapsulation overhead
- Reduce the effects of high cell/frame/packet loss rates

Table 12-6 summarizes the header lengths for various VoIP implementations. Add the number of bytes in the link-layer header to the number of bytes in whichever VoIP packet option you are using (for example, IP/UDP/RTP, CRTP with UDP checksums, CRTP without UDP checksums) to determine the total number of overhead bytes.

Table 12-6 *Header Length (Bytes) for VoIP over Different WAN Technologies*

WAN Technology	Link Layer	IP/UDP/RTP	CRTP	CRTP
			(UDP Checksums)	(No UDP Checksums)
HDLC	6-8	40	4	2
ML-PPP	7-9	40	4	2
Frame Relay	4	40	4	2
FRF.11 Annex C	9	40	4	2
FRF.12	8	40	4	2
ATM-AAL1	6-12	40	4	2
ATM-AAL5	13-18	40	4	2

ATM header requirements actually vary with the VoIP packet size. For all data segmented into AAL5, each cell introduces a 5-byte header, with the final cell requiring an 8-byte trailer. VoIP packets that fit into a single ATM cell payload require 5 + 8 bytes of ATM headers, and VoIP packets that fit into two ATM cell payloads require 5 + 5 + 8 bytes of ATM headers. This does not include any padding that is necessary to fill the payload field. The values in Table 12-6 assume that the IP header and codec data fit within one ATM cell when CRTP is used, and within two ATM cells when CRTP is not used.

Table 12-7 summarizes the header lengths for various VoFR implementations.

Table 12-7 *Header Length for VoFR in Different Configurations*

	Overhead (Bytes)
FRF.11 (Annex C fragmentation)	9
FRF.11 (FRF.12 fragmentation)	8
Cisco proprietary (voice-encap method)	6

VoATM is currently supported on the Cisco MC3810 router using AAL5 and AAL2. In addition to the 5-byte ATM header, there is a 4-byte VoATM header before the codec data. The ATM header lengths (for VoATM and VoIP) do not consider the payload padding, which increases the effective size of the header. You cannot determine the effective header length until you choose the payload size. Be sure to add the payload pad to the header length before calculating the required bandwidth.

Selecting Hardware to Meet Requirements

The surest way to make a book obsolete before it is published is to refer to the capabilities of specific hardware models. That being said, traditional telephony interfaces are available for the following Cisco routers as of this writing:

- 1750 Modular Access Router
- MC3810 Multi-access Concentrator
- 2600 series Modular Access Routers
- 3600 series Modular Access Routers
- AS-5300, AS-5800, and Access-Path Solutions
- 7200VXR series Core Routers
- 7500 series Core Routers

High-end ATM switches are not included here because the services they offer for voice are usually of the ATM-CES variety as opposed to VoATM. The distinction is made because the ATM-CES service provides a T1/E1 type of service, with no telephony intelligence required.

Small-scale implementations can support up to 4 analog ports (FXS/FXO/E&M) or 48 to 60 digital ports (2 T1/E1) in the 2600 series Modular Access Routers. The 3600 series offers up to three times the port density of the 2600 series, and the 7200 and 7500 routers provide higher density aggregation. The Access Server router line (for example, AS-5300, and so on) provides high-density T1/E1 telephony connections for VoIP, with value-adds like SS7 adjuncts and Interactive Voice Response (IVR).

As of this writing, there are various hardware caveats for supporting transparent pass-through of common channel signaling (CCS) and interpreted QSIG. For current hardware support and feature comparisons, you should look at Cisco Connection Online (CCO): www.cisco.com/. If you have not spent a lot of time there already, you are missing a truly incredible amount of free documentation. Cisco deserves a lot of credit for maintaining and constantly updating a wide range of product and technology information.

Reviewing Proposed Solutions in Terms of Requirements

After you have spent time developing the high-level integrated network design (that is, selecting the VoX technology, interfaces between routers and phone switches, bandwidth and trunk requirements, and hardware platforms), you must take a step back and verify whether you have met the design goals that were initially identified. Presumably, you have been mindful of the design goals at all phases of the early design, but it is a worthwhile task to review the requirements again before you commit to an equipment purchase.