Architectures For Convergence

Gary Audin

Gary Audin is president of Delphi, Inc. consultancy. He has been an independent communications and security consultant for 25 years.

Which definition of Convergence do you accept? All information media on one network? All applications on one network? Wired combined with wireless transmission? Public combined with private networks? Or business combined with entertainment traffic? Pick a vendor or service provider and you will discover their personal view—one that will probably not match any other's definition.

In 1992, John C. Malone, CEO of Tele-Communications Inc., used the term to predict an all-cable-company vision for digital transmission for business and the home. Richard C. Notebaert, then CEO of Ameritech (now CEO of Qwest), gave a broader view of convergence in the *New York Times* in 1998, "Conventional wisdom holds that convergence—gradual blurring of telecommunications, computers and the Internet—is primarily about the inevitable clash of voice and data networks."

Convergence has now evolved into the concept of one network for all applications: data, voice and video. Even this view will change to cover all applications converged into one IT environment. Past attempts at convergence have been application-agnostic and could not support application or user convergence (see "The Past Determines the Future"). A converged network is a stepping-stone to the final evolution of convergence.

Convergence Defined

Convergence can best be defined as one structure, one architecture to support all forms of information media on all forms of network technology. The massive adoption of the TCP/IP protocol suite has brought together multiple network technologies, local and wide area networks (LANs and WANs), wired and wireless, public and private networks, and access to a near-infinite number of applications. Even though the earlier attempts at convergence did not foresee the dominance of TCP/IP, convergence is now generally acknowledged to be based on the IP protocol. IP is the glue that binds the networks and applications together.
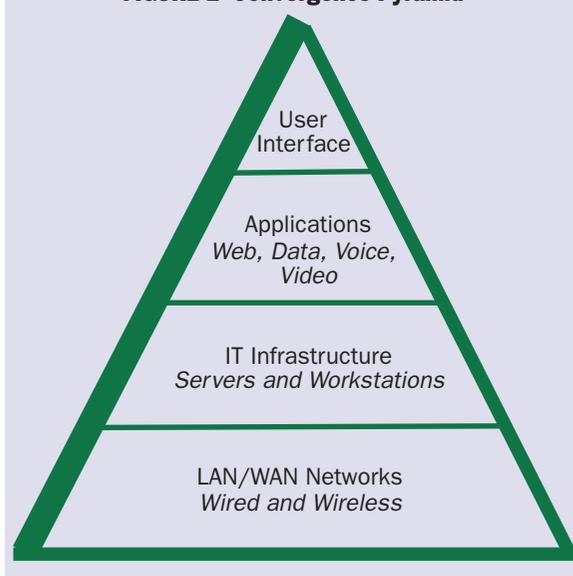
The Convergence Pyramid (Figure 1) shows the four levels of convergence: **Network convergence** (bottom level) of LAN, WAN, wired and wireless transmission using the IP protocol suite, delivers common information transport. This level is independent and agnostic to the applications except for network performance considerations. The primary devices and facilities are LAN switches, carrier facilities, wireless LANs (WLANs), cellular networks and routers.

The next level up is the **IT infrastructure** of servers, workstations, phones and video terminals. These devices can be application-specific like a telephone, or can be an application generic PC or server. Application support protocols like FTP, SMTP, Telnet, HTTP, RTP and RTSP are also part of this level. This level can provide software-based services to the application programs at the next-higher level.

The third level up from the bottom contains the **applications** to support the end user. These include software for application and Web servers, codec and compression technology for voice and video and other forms of input/output devices (point-of-sale

FIGURE 1 Convergence Pyramid

User Interface

Applications
Web, Data, Voice, Video

IT Infrastructure
Servers and Workstations

LAN/WAN Networks
Wired and Wireless

This article was printed in *Business Communications Review* in October, 2004

terminals, bar code readers, scanners etc.). This level also brings together the multiple interfaces to the end user. This is where application convergence is implemented.

The top level, **user interface**, presents common audio, visual and physical (keyboard, mouse, etc.) interfaces that offer the end user access to the converged applications.

These four levels have to be evaluated further on many attributes, as seen in "Architecture Attributes" (p.6).

Centralized Architecture

The centralized architecture (Figure 2) places all the applications, management and network connections in a single point. The IBM System Network Architecture (SNA) and the PBX are good examples of this design. This worked well when networks and computers were expensive and the IT staff was small. It was initially used by large organizations.

As the cost decreased, smaller organizations could afford this approach. A single-site organization may still find this centralized approach valid. It can scale to the maximum capacity of the centralized system (hundreds to thousands of devices) but cannot be easily expanded beyond this point.

This approach, however, does not address most of the issues defined in "Architecture Attributes." There are limitations in flexibility, reliability, disaster recovery and load balancing. The connecting network is used to access the applications: Direct device-to-device communications is not possible. The end

devices are so specialized that it is impossible to support converged applications; in many instances, terminals are application-specific.

A more modern version of this configuration is a group of PCs connected to a single server or cluster of servers supporting only data and Web applications.

Horizontal Architecture

The Horizontal (homogenous) Architecture (Figure 3) adopts the concept that all devices are equal and should

FIGURE 2 Centralized Architecture
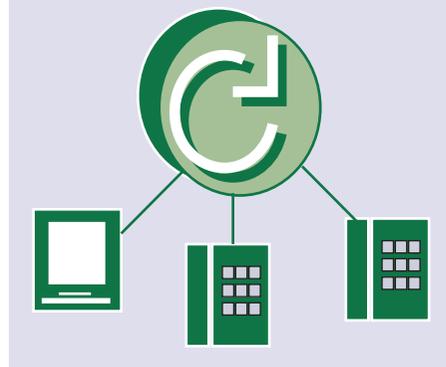
The Past Determines The Future

Convergence has been a dream for network architects for decades. The first two attempts were directed toward the convergence of the transmission: ISDN (Integrated Services Digital Network) and ATM (Asynchronous Transfer Mode).

In the mid-1980s, telephone companies offered ISDN, which provided 64-kbps transmission paths on a digital circuit-switched network. There are two types of ISDN:
■ The Basic Rate Interface (BRI) consists of two 64-kbps transmission paths (channels) with one 16-kbps X.25 packet-switched signaling path that can be used for low-speed data transmission.
■ The Primary Rate Interface (PRI) consists of 23 64-kbps channels and one signaling-only channel. The 23 64-kbps paths are information– and data protocol-agnostic. Any digital information can be carried. This means that voice, data and video can be carried on the channels. All forms of information are provided with only one form of quality of service (QOS), namely, guaranteed bandwidth, constant and short delay and no packet loss. Channels can be bonded together for video and Internet access to provide paths that operate at 128, 256, 384 and 512 kbps.

The ISDN network provides only OSI Layer 1 convergence. It does not deliver any level of protocol, application or user convergence.

The second attempt at convergence, ATM, became available in the early 1990s.

ATM was also called Broadband ISDN, even though ATM was a completely different approach than traditional ISDN.

ATM is a packet-switched model (the packets are called cells) for network convergence. ATM is a switching technology designed to work with multiple forms of transmission facilities: copper wire, coaxial cable and fiber optics.

ATM requires that a specific cell-based protocol be used to access the ATM network. Any cell can carry any information—voice, data or video—at speeds greater than traditional ISDN, including multi-megabit transmission.

Five ATM Adaptation Layer (AAL) interfaces were developed for supporting different types of application requirements. AAL 1 and AAL2 were designed for constant and variable bit-rate voice and video. AAL 3 and 4 were designed for data and LAN transmission, but were found to be inefficient and expensive. AAL 5 is a simpler and more efficient interface for data.

Carriers and some large enterprises and government agencies implemented ATM networks. Its main attractions were the ability to support multiple forms of QOS and connect to high-speed links like T3 and SONET facilities. ATM is also application-agnostic and does not define protocols above OSI Layer 2.

ATM is one step closer to true convergence, but it still does not address internetworking, application and user convergence□

Architecture Attributes

Any organization moving to a converged environment will have to evaluate the vendor and service provider offerings with a set of general, technology-independent requirements. Many vendors offer converged solutions but go on to define convergence their way, which promotes their particular approach. Many go only as far as the network and leave the application convergence issue to be solved by the customer or another vendor. Others offer a complete solution. In any case, there are specific attributes that any solution should provide:

■ **Flexibility—**Organizations have to change. Stagnation leads to elimination. An architecture must be capable of being modified easily, rapidly and with low risk.

■ **Longevity—**The customer wants real future-proofing of the architecture, not a change every three to four years. Retaining as many assets over time as possible is a major goal.

■ **Reliability/Availability—**The architecture should limit failures as well as restoring service rapidly. This is an issue of architecture hardware and software design, as well as vendor support. Products designed for the carrier space and those that have been on the market for some time generally meet these requirements.

■ **Disaster Recovery—**This is different than reliability/availability. The architecture should be able to sustain a total loss of network and/or application elements and still provide service to the customer.

■ **Common Services—**The goal of convergence is to deliver services that are available to all users and customers, and which look and behave the same regardless of what the service is or where it is delivered.

■ **Management—**Managing local and remote resources should be simple, intuitive, secure and support reliability, availability and disaster recovery goals.

■ **Load Balancing—**No two sites in an organization will be identical in network traffic, and different sites will probably vary in their use of applications. The architecture design should be able to shift work and traffic so that the best balance of resources can be accomplished.

■ **Expansion/Scalability—**The more modular the hardware and software offered, the better the customer can implement a solution that meets his or her requirements without over-purchasing components because the vendor tries to make one size fit all products□

have the ability to connect to any other device. It is a network of peers, islands of services. This concept is exemplified by the design of Ethernet, X.25 packet and IP-based networks. The use of routers provides converged transport architecture. IP-based routers do not offer IT, applications or user convergence. Routers are, by design, application–ignorant.

If the boxes in Figure 3 contain applications, then applications and user convergence can be delivered. An example would be multiple IP-PBXs networked together. Each IP-PBX can offer all the applications to a common user interface for its local users, and relay calls and data sessions to other IP-PBXs and their users. It can scale to the maximum capacity of a single system (tens to hundreds of devices). Adding sites multiplies the number of devices to hundreds and thousands for the entire network.

This approach, however, does not address all the issues defined in "Architecture Attributes." There are limitations in managing multiple equal sites (especially when the number of sites goes beyond five or six); and disaster recovery and load balancing among multiple sites can be complex. Expansion may require a second IP-PBX at one of the sites if the existing IP-PBX cannot scale to a large user population. On the plus side, this architecture is more flexible; it allows sites to be added easily; the loss of one site does not bring down the operation of other sites; and the modular approach can produce a longer-lasting solution.

Hierarchical Architecture

The Hierarchical Architecture (Figure 4) resembles an organizational chart. Large enterprises like banks, insurance companies, retail chains, real estate firms, government agencies, travel agencies and schools with off-campus facilities all fit this model. They have many small sites of 5 to 25 devices; a few medium sites of 100+ devices and probably a single large (main) site of hundreds to more than 1,000 devices.

This architecture addresses shortcomings of the previous two examples. The Centralized Architecture requires a great deal of communications facilities that would be inefficiently used to support the device distribution and population described in Figure 4. Similarly, the Horizontal Architecture systems may be too large for the very small sites and not large enough for the main site.

In the Hierarchical Architecture, there should be two sizing levels for the sites: one size that can support 4 to 100 devices and a second that can support hundreds to thousands. The few medium sites of 100 or more devices could be implemented with a scaled-down version of the large configuration. The large (main) site could use a carrier-based IP-PBX, while the smaller sites could use an enterprise IP-PBX.

This architecture provides the best solution for organizations with many sites to support. Applications can be supported at each site and/or can be located at the large site. The user interface can be common to all

locations. The architecture design is flexible and can last a long time. The loss of any size site will not bring down the network, since all sites could support all applications.

The architecture provides distributed intelligence to the users where needed. If one or more of the small or medium–site servers fails, the large (main) site can provide services as long as the network is functioning, providing disaster recovery. Load balancing and expansion are made easier with the adaptability of the systems for different sizes. Management of all sites can be centralized at the main site.

An added capability of the hierarchical structure is the ability to support functions locally, such as local phone calls, without having to signal through the whole network. This reduces bandwidth requirements and speeds execution of the call request. This does not prevent calls from moving up the hierarchy when appropriate. The reverse direction will also work; the central site can pass calls down to the local site where employees, acting as call center agents, can handle the overflow from a central call center.

It Can't Be All One Vendor

No single vendor can provide every application for convergence. The architecture should support third-party software and devices. This is especially true for vertical markets like retail, hotels, transportation and some government agencies. Interoperability is one of the emerging issues as businesses explore convergence options. How the architecture vendor works with other vendors' software, systems and devices is a measure of the support that will produce a successful convergence implementation.

Any new architecture will have to be inter-operable with the customer's legacy net-works and applications as well as carrier networks. There should be no limitations on the legacy interfaces, nor should special interface designs be required.

All three architectures are capable of interoperating with other networks. The real difference in interoperability will be the vendor's approach when producing their products. The customer should be able to transition to the new architecture at their own pace and still interoperate with legacy technologies.

The Business Case For The Architecture Decision

The architecture selected for convergence must satisfy the business requirements before it meets the technical requirements. The architecture must improve the business's operational effectiveness both for internal users and external customers. This means business processes must be simplified. The business continuity (reliability) of the converged environment must satisfy the requirements of the critical applications and telephone services. Finally, as more critical functions move to converged operation, the security requirements of the applications, data and multimedia services increase. Without security, does anyone want to use the features and functions offered?

The three architectures are compared in Table 1. Each attribute is evaluated for the three architectures, rating it on its ability to
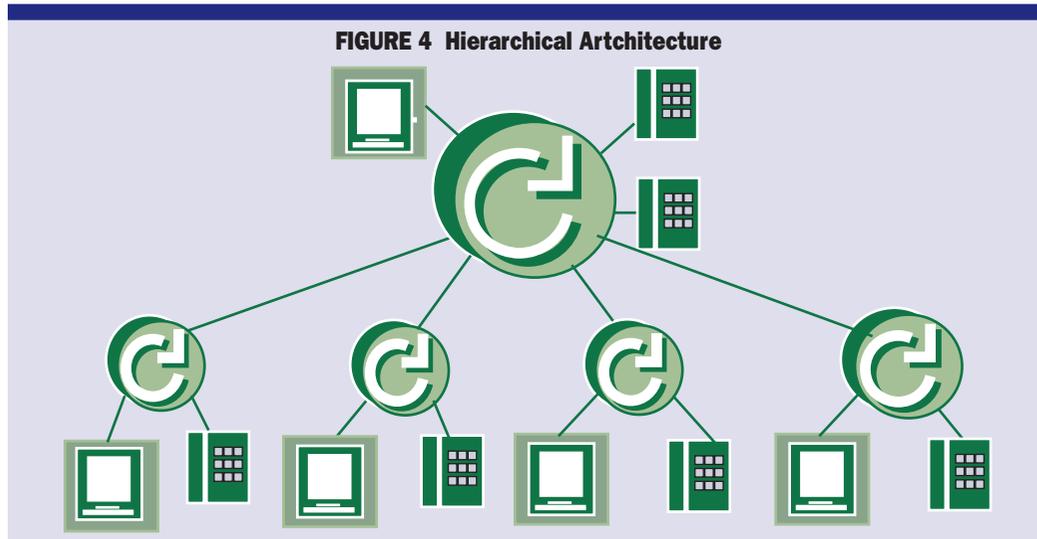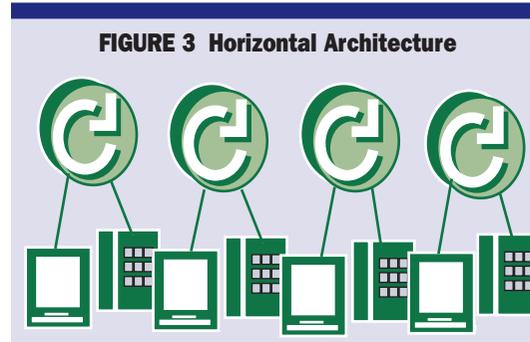

FIGURE 3 Horizontal Architecture

FIGURE 4 Hierarchical Artchitecture

TABLE 1 Comparing Architectures

Architecture Attributes	Centralized	Horizontal	Hierarchical
FlexibilityLow	Low	High	
LongevityHigh	Low	High	
Reliability/Availability	Low	Low	High
Disaster Recovery	Low	Low	High
Common Services	High	Moderate	High
Management	High	Low	High
Load Balancing	Does not exist	Low	High
Expansion/Scalability	Limited	Low	High

deliver the attribute. A "Low" rating means that the attribute is not a strong point for that architecture. "High" means that the attribute is a major value. For example, "High" longevity for a centralized architecture can best be observed with the long life of the legacy PBX and IBM's SNA products. Another example: The disadvantage of horizontal architecture for management is that it is segregated, not unified like the other two architectures.

Cost reduction is paramount. Cost containment is also desirable. The more the cost can be allocated as a fixed fee, the better the user can predict the expenses that will be incurred. There should be more visible, accurate and timely reporting of the operating costs.

Whether a centralized, horizontal or hierarchical architecture is selected, the management and administration expenses should also decrease. All these elements lead to a lower total cost of ownership (TCO).

Finally, the architecture choice should increase productivity. The user interface should meet the following goals:
1. Consistent—It is dependable and performs the same for every use.
2. Obvious—It should be similar to other experiences that a user may have encountered, limiting training expenses and reducing errors. If an IP phone looks and behaves like the legacy phone with additional functions and features, then the new collaborative voice applications will be readily accepted.
3. Intuitive—The interface should be easy to learn and should encourage the user to continuously explore for new functions and features and be rewarded by higher productivity when they find them.

These three productivity goals will then strengthen both the user and customer interactions.

It's Not Just The Network
For many years, network architects drove convergence goals. These architects strove to produce a network structure that could carry all forms of information. They considered the application devices as subscribers to the network. The primary consideration for the subscriber (applications) was to add quality of service (QOS) support so that applications could be provided with the performance required. Otherwise, the network was independent of the applications and traffic flowing across it.

The evolution of the PC, server, desktop operating systems, richer application software, handheld computers and digital signal processors invited architecture designers to move into collaborative multimedia support. Users became familiar with point-and-click operation and color displays of all sizes. Both users and application designers from multiple disciplines started to imagine what could be accomplished if there was no boundary between data, voice and video.

Collaboration became the new theme. The architectures for collaboration must address the IT infrastructure, both what is already in place and what can be implemented in the future. Vendors that deliver collaborative operation, especially between voice and data applications, will deliver the applications. Finally, the user interface will blend the features and functions available, removing the distinctions among data, voice and video.

An architecture that focuses only on the network is no longer adequate to support the future of IT and communications. The IT organization must implement an applications and user interface architecture that enhances the business mission. Choose the next architecture for convergence wisely. You will have it for most of your career□