A KelCor Thought Leadership Paper

Creating the “It Just Works” Video Network

Enterprise Video That Meets People’s Expectations

November 2013

E. Brent Kelly, Ph.D.
Principal Analyst

Sponsored by:

Disclosure on Editorial Control

KelCor has been compensated by Sonus Networks to prepare this position document, but the thoughts and opinions are those of the author alone, and KelCor has maintained full editorial control throughout. Sonus specifically sought KelCor's unbiased views on key challenges organizations face when using video and what can be done to mitigate these challenges. This document is the result of that effort.

Making Video Work for People

From the boardroom to the browser, video has become a valuable business tool. When video communications systems work properly, people become more interactive, creative, and productive. Conversely, when video systems fail, they cause delays along with angst and frustration in the people trying to use them.

People using video want and expect solutions that "just work." And, "just working" may mean different things to different people, depending upon their immediate circumstances. For example,

1. A traveling executive may want video from her tablet coming in over a virtual private network (VPN) connection to display well on the telepresence screen the board of directors is viewing.
2. Following an acquisition, an employee using a Cisco Jabber video endpoint wants to communicate with someone in the acquired company who has Microsoft Lync.
3. During a meeting with partners or suppliers, video is used to strengthen relationships. This video needs to securely traverse the network firewall.
4. A customer needs help and wants to use video on his smartphone to interface with someone so that he can properly fill out an application or locate the right service part.
5. Engineering teams working at remote locations need to meet regularly for progress reviews, but there is limited wide area network bandwidth between.

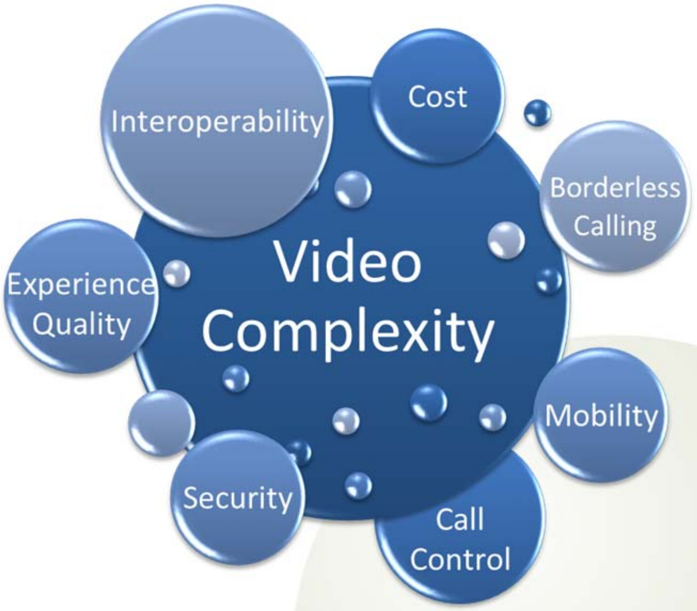
The Video Challenge for IT

These and many other strong business use cases illuminate how people are using video; they also highlight some of the challenges video systems that "just work" must overcome. For example,

1. In the case of the traveling executive, video from her tablet must be made interoperable with the telepresence video system; furthermore, her video image must be rendered with sufficient resolution so that it so that it can be viewed on the telepresence screens without distortion.
2. Different desktop communications systems, such as those from Cisco, Microsoft and Avaya, use different signaling mechanisms, and they need a signal translator between them in order to be interoperable.
3. Meeting with people outside of the organization raises security issues for IT including secure firewall traversal and the use of proper encryption mechanisms.
4. The smartphone user traverses the public Internet and the firewall, and his video must be securely routed to the right party within the organization in order to receive the help requested.
5. The engineering teams cannot saturate the network with just video or other key business process that rely on the network may cease to function; hence, call admission and bandwidth control are required to ensure a high quality experience within the confines of limited network bandwidth capacity.

While users may “just want video to work”, the complexities behind making it work can be challenging. The good news for business managers is that video and video infrastructure have evolved to the point such that with some foresight and planning, “making video work” is readily achievable.

Figure 1. Challenging adding complexity to how video works (source: KelCor).



The Infrastructure Required to Make Video Work

Behind successful video deployments one will encounter up to five infrastructure components. These are often logically separated, appearing as different devices or servers, but not always.

A Simple Video Network

In a video network in which video endpoints communicate only with other endpoints using the same codecs and protocols (see the box below for a discussion of codecs and protocols), only two infrastructure components are required: an **MCU** and a **gatekeeper** or **SIP proxy**. The MCU provides multipoint capabilities by combining video streams from multiple endpoints into a single video stream and returning the combined video image back to the endpoints. In practice, a person does not see his own video image in the mixed stream; the mixing removes his stream, sending back only the images of other participants in the call.

H.323 gatekeepers and SIP proxies both help with setting up and routing video calls. The H.323 gatekeeper can also be programmed to provide call admission control (CAC) so that calls are not initiated if sufficient network bandwidth is not available. A gatekeeper can also be programmed with a dialing plan to make calling using video easier. The SIP protocol does not natively provide call admission control capability[1]. In some organizations, the PBX has assumed the functions of the gatekeeper and/or SIP proxy server[2].

[1] Call admission control is not part of the SIP standard. Other network devices often provide CAC such as session border controllers and/or PBXs.
[2] Most PBXs today support the SIP protocol. Some PBXs also support H.323 or a derivative of H.323 (like Cisco’s SCCP protocol).

A network with only a MCU and either a H.323 Gatekeeper or a SIP proxy will support video calls between similar systems with the same protocol. If the MCU supports both SIP and H.323 then calls from these types of systems can be bridged together. However, this simple video system has no mechanism to securely support calls that traverse the network boundary. Because most endpoints reside behind a firewall and a NAT[3], pinholes will need to be punched into the firewall to allow firewall traversal by the video packets[4]. Furthermore, this simple video network cannot handle calls between two endpoints with slightly different signaling mechanisms unless the MCU is used. For example, a desktop endpoint using the SIP protocol will not be able to communicate directly with an endpoint running H.323 even if they are using the same video compression algorithm internally (i.e. H.264) without using up two ports on an expensive video bridge[5].

Normalization vs. Transcoding

In order to transmit video over an IP network, the video must be compressed using a compression/decompression algorithm (often called a codec for short). The sender's device captures a person's or group's video images, digitizes them, and then runs the digitized video through the codec to compress it so that these video images can traverse the network using a reasonable amount of bandwidth. The receiver's video device takes the compressed video stream from the sender's device, decompresses it, and displays it on the video screen.

There are a number of compression algorithms that exist. Some of the most popular are H.264, H.263, and VP8. Most of the video in enterprise-oriented video conferencing devices is based on the H.264 or H.263 codecs. Skype uses VP8, and the emerging WebRTC standard is determining whether to use VP8, H.264, or both.

Besides just compressing the video, video devices must use some type of signaling protocol in order to initiate, control, and terminate a call. The most common video signaling protocol standards are Session Initiation Protocol (SIP) and H.323. The H.323 standard is leaves little room for private interpretation – devices using this protocol are generally interoperable. However, the SIP standard leaves a lot of room for private interpretation, and as a consequence, systems from two different vendors that are both labeled "SIP compliant" many not interoperate with one another even if they are using the same video compression algorithm.

Because there are differing compression standards and variations of signaling protocols, video sometimes needs an intermediary to make it interoperate. The easier case is where two different devices use the same video compression algorithm, H.264 for example, but they use slightly differing versions of SIP signaling or one uses SIP and the other uses H.323. In this case, the intermediary device does signaling **normalization** between the video devices, modifying the information in the signaling packets so that the other device can accept the signaling control messages. The great benefit here is that the compressed video does not need to change nor does each device consume an MCU port.

A much more difficult case is where one device uses one video codec, say H.264, and the other device uses a different video codec, VP8 for example. In this case, the compressed video stream must be modified. This is a process called **transcoding**. In the H.264-VP8 example, the transcoding gateway or MCU must convert H.264 video to VP8 by decompressing the H.264 video stream, recompressing it in VP8 format, and then sending it to the endpoint that uses VP8. The process is reversed for the VP8 video: the transcoding device must first decompress the VP8 video stream, recompress it in the H.264 format, and then send it on to the H.264 device.

Transcoding is much more complex than signal normalization; hence, transcoding gateways or MCUs often require more computing power and are usually done using hardware-based devices with specialized DSPs. Signal normalization, however, is much less compute intensive and can usually be done on off-the-shelf servers.

[3] NAT is an abbreviation for a Network Address Translation device. Among other functions, this device hides the IP address of internal devices from devices outside of the firewall.

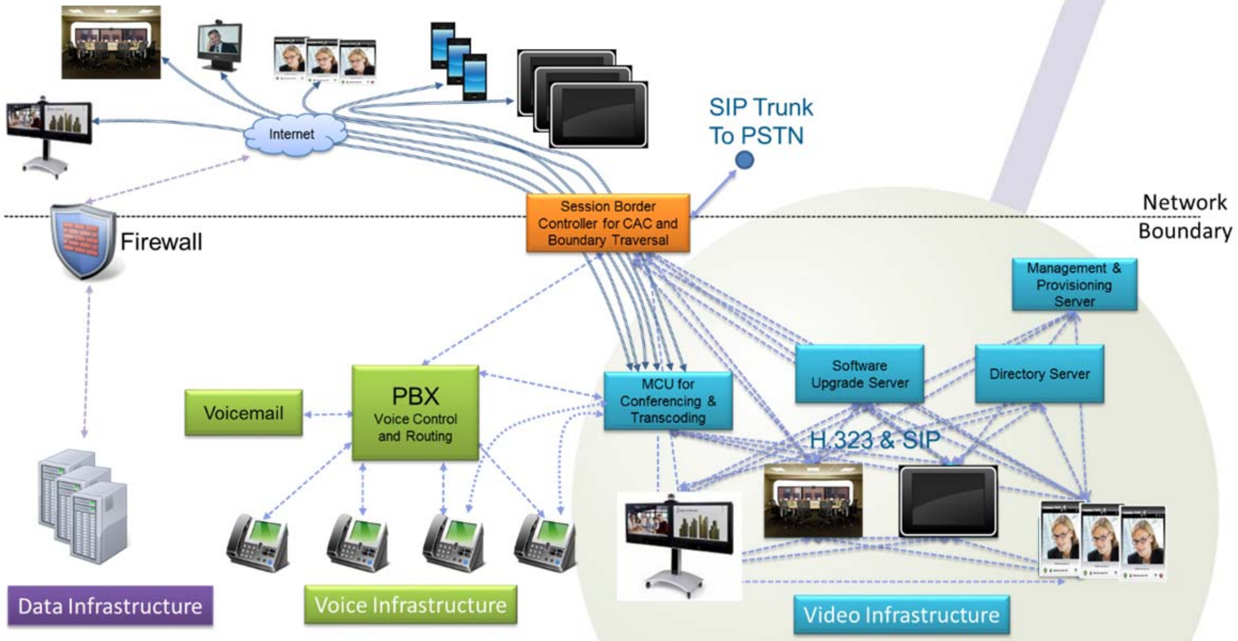
[4] There are several ways such a network could be set up to handle calls outside of the firewall. One way would be to put two different network cards in an MCU and have one face the outside world and the other on the Local Area Network (LAN). Others involve punching pinholes in the firewall using application layer gateways (ALGs). Frankly, these methods are not commonly found to be that useful nor scalable. Most organizations place a session border controller between the inside and the external networks.

[5] MCU ports typically cost $1,000 - $2,000 per port.

The Real World Video Network

In a real world video network, an organization will likely have video endpoints from several manufacturers. Some will support only H.323 while the newer endpoints will support both SIP and H.323. Desktop endpoints will likely only support SIP. All of these can be interconnected with an MCU for multipoint calling.

Figure 2. The complexity of a real world video deployment with internal and external communications (source: KelCor).



Real world organizations often have voice systems (PBXs) from a variety of manufacturers; hence, trying to centralize call control and SIP proxy/H.323 gateway capabilities in any one of them will prove difficult. In addition, each voice manufacturer has a slightly different version of the SIP protocol and each supplies its own desktop unified communications client. Some examples would include

1. Cisco Unified Communication Manager for call control and Cisco Jabber as the desktop/tablet UC client.
2. Avaya Aura for call control and Avaya one-X or Avaya Flair as the desktop/tablet UC client.
3. NEC Univerge 3C for call control and Univerge 3C UC Client as the desktop/tablet client.
4. Unify (formerly Siemens) OpenScape Voice call control and OpenScape UC as the desktop/tablet client.
5. Microsoft Lync Server for IM/presence and call control and the Lync client as the desktop/tablet UC client.

In an organization with multiple PBXs and UC clients, different desktop/tablet UC clients may be spread throughout the organization that do not interoperate with one another even though they all are “SIP compliant”.

Alternatively, an organization may have one or more PBX systems for voice communications, but it chooses to deploy a single desktop client not manufactured by any of the PBX vendors. Microsoft Lync or IBM Sametime would be examples of UC clients that can be used with different types of telephony systems. The net result is that hybrid solutions often exist in which each voice vendor has its own signaling method for phones while the UC client with its video capabilities uses a different signaling protocol[6]. An example would be an organization using Cisco for its PBX and group/telepresence video solutions but Microsoft Lync on the desktop.

A situation occurring frequently is the need to communicate and collaborate with people outside the firewall. Examples include employees working from home or in a hotel and communicating via video with partners and customers who at their own locations. In these scenarios, there can be a multitude of different video devices used such as desktop clients, tablet clients, group or telepresence video systems. For these use case to actually work, the video needs to first traverse the firewall securely, and then it needs to be made interoperable with the video from all of the other endpoints.

In summary, many organizations are faced with interworking challenges because they have multi-vendor communications systems, hybrid unified communications environments, video endpoints ranging in size and quality from telepresence systems to smartphones, and users who need remote access.

Parallel Boundary Elements

In real world video networks, two additional infrastructure components, **firewalls**[7] and **session border controllers** (SBCs), are critical. Understanding how they work is vital to successfully setting up a video solution that “just works” in the manner people expect it to.

In a real-time multimedia-enabled communications network, firewalls and SBCs often work in parallel. Firewalls handle the regular IP data traffic while SBCs handle the real-time communications traffic.

Firewalls are designed to protect or isolate the computing resources in one network from those in another. The firewall works by examining each incoming or outgoing packet and determines whether to forward it on. By their nature, the SIP and H.323 communications protocols are firewall-unfriendly for two reasons:

1. Devices from outside the firewall may send unsolicited invitations intended for someone inside the firewall to join in a real-time voice and/or video communications session. Because no device on the inside of the firewall requested the invitation, the firewall discards these unsolicited invitation packets.
2. The IP addresses of where to send the multimedia packets are embedded within the IP packets. When these packets are required to traverse the firewall, the firewall often cannot resolve them and simply discards these packets.

SBCs on the other hand, understand media protocols and can work in parallel with the firewall. They may be thought of as a real-time communications-aware firewall, providing the functionality to make a real world video implementation “just work,” along with those critical policy capabilities that allow a video deployment to “just work” securely.

[6] We should point out that all of the PBX manufacturers as well as Microsoft Lync all support “standard SIP”. The issue is that within the SIP protocol, there is lots of “wiggle room” in how the protocol is actually implemented. Consequently, even if two SIP-based solutions are labeled “SIP compliant”, they may not actually interoperate with one another.

[7] Firewalls and NATs are often found in the same device. As mentioned earlier, NATs hide internal IP addresses from external devices, providing an additional measure of security protection beyond what the firewall does on its own. For simplicity, the term “firewall” in this discussion includes NATs as well.

The Business Value of the Session Border Controller in a "Just Works" Video Network

Session border controllers typically sit at the edge of the network, providing a clear demarcation point between video endpoints inside the trusted network and those outside the trusted network. In a sense, the term "border" in the name "session border controller" implies edge capabilities, but this is really a misnomer. SBCs provide session control and security, whether the session is between endpoints inside and outside the trusted network or whether sessions are between endpoints entirely contained within the trusted network.

SBCs provide organizations with a number of advantages, but in terms of making the video network "just work" for people, they deliver three main business benefits:

1. Session Management,
2. Endpoint interoperability, and
3. Security/Governance.

Session Management

From a network topology perspective, the session border controller is the ideal element in a complex network to enforce call admission control and to do it on a session by session basis. Multiple UC and video devices can access the SBC for call admission control (CAC) and quality of experience parameters to make enable audio and video to traverse the network with as much quality as possible.

CAC helps provide optimal end user video experiences by regulating the number of endpoints allowed on the network at any given time. This assures that there is sufficient bandwidth allocated to each individual video stream, and if there is not enough bandwidth to support a high quality call, the SBC can independently, session by session, negotiate call quality and codec encoding parameters between devices to assure the highest possible video quality given the network constraints. If sufficient bandwidth is not available for a quality video call, the SBC may optionally fall back to audio-only, or it can be programmed to not even let the call start up at all. These decisions are done proactively at the beginning of session negotiation to ensure that established sessions are not impacted.

CAC works hand-in-hand with quality of service (QoS) mechanisms so that video network packets can be marked to have higher network traversal priorities than other packets, such as http-based web browsing sessions. For example, a board meeting using very high quality telepresence video endpoints may have its packets marked with a higher QoS level than other video or data streams because the board meeting may be considered more important. Audio calls made to 911 numbers can also be prioritized by the SBC above all other voice and video sessions.

Call admission control and QoS mechanisms also work together to protect data applications. On a low bandwidth WAN link, for example, the CAC will employ intelligent network saturation policies so that video streams on the network do not negatively impact critical application data flows over the same link.

SBC Sessions

It is important to understand a key principle behind how SBCs work. Session border controllers are built around a concept called the back-to-back user agent. In laymen's terms, this means that all calls are terminated at the session border controller.

For example, if a video endpoint located outside of the firewall wants to communicate with a video endpoint inside of the firewall, there are actually two legs to the call: one leg between the external video endpoint and the SBC and a second leg between the SBC and the internal endpoint. The same would be true for video between endpoints located within the firewall.

It is this ability to provide governance and control on each individual session that makes an SBC such a powerful and useful infrastructure element in the "just works" video-enabled network.

Organizations with multiple PBXs or video deployments will gain great benefit from moving all of the CAC functionality and media QoS markings into the session border controller[8]. The advantage of using the SBC for CAC and QoS is that the session attributes determined from the SIP signaling allow for finer grained policy compared to typical low-level routing techniques that are not session aware.

Endpoint Interoperability

In many organizations, video endpoints from different manufacturers have been deployed. Some of these systems support multiple video codecs, and the SBC can negotiate with each video device so that the same video codec is used, thus enabling interoperability.

Even if all of the endpoints in a video call use the same video codec, the signaling protocols used by Cisco, Microsoft, Avaya, Polycom, and others are just different enough that the video will still not connect. When tablets and smartphones are added to the list of possible endpoints, interoperability becomes a real issue in making video “just work”.

To help solve this issue, session border controllers are often able to modify the signaling information contained within the SIP signaling packets so that endpoints from different manufacturers connected to different UC systems can communicate with each other. For example, Cisco has its version of the SIP protocol, and Cisco Jabber clients connected to a Cisco UC system will use this protocol. Likewise, Microsoft has its version of the SIP protocol, and Lync endpoints connected to the Lync server will use Microsoft’s protocol. Putting an SBC between these systems allows the video to interoperate because the SBC can normalize the SIP signaling between them.

What this does for an organization is that it allows them to keep existing investments in hardware and software while making video solutions from different systems interoperable. The investment protection is tremendous in some cases, negating the need for a forklift upgrade in order to have a homogeneous solution from a single vendor.

Some SBCs go even farther by allowing systems using different protocols to interoperate. For example, it is possible to connect a SIP video endpoint with an H.323 video endpoint using a session border controller to do the normalization between the protocols.

Security/Governance

Session border controllers allow organizations to apply policies to video sessions. One of the more critical policies is security.

The call admission control capabilities SBCs deliver also provides an element of network security. For example, those using voice or video to call into the organization from outside the firewall must be properly authenticated before any video crosses the network boundary. Furthermore, the SBC can assure that any malformed SIP or H.323 packets sent by hackers are discarded, thus avoiding security breaches. Even for endpoints can call into an organization without using a VPN, like Cisco Jabber or Microsoft Lync, the SBC plays a critical role by assuring that external users are authenticated and that incoming video is bandwidth properly controlled.

SBCs can also help deflect denial of service (DoS) and/or distributed DOS attacks. They can also encrypt both the signaling and the video media that pass through them in order to assure that the contents of a video conversation are kept private by providing a secure connection.

[8] Cisco, Polycom, and Avaya/Radvision) all have their own video management solutions. For Cisco and Avaya/Radvision, video signaling can be controlled by the PBX or it can be controlled by these companies’ own software-based video management systems.

In a typical enterprise, there may be three separate systems competing for call control: the PBX, the video conferencing management system, and an IM/presence solution (i.e. Microsoft Lync) hat also supports voice calling. The SBC provides the capability to harmonize all call control and routing logic throughout the organization, into a single device. As a consequence, centralized call control and routing through an SBC make it less expensive for organizations to operate their video as well as their voice-over-IP networks.

SBCs also benefit organizations by providing billing capabilities. Because the SBC works on individual video streams, it can track all of the video calls made or received, and it does so from a centralized location. Organizations that bill for network usage can use these call detail records to bill back to the departments or groups who use video on the network. In addition, companies can use the SBC’s call records to see if video is being used and by whom as part of an internal marketing effort to fully realize the benefits of the investment in video communications endpoints and infrastructure. In addition, session border controllers allow organizations to review the quality metrics in the call detail records as a way of monitoring service quality.

Actionable Recommendations

People want video systems that work, and they want them to work regardless of which video-enabled device they happen to be using at any particular time. Video helps people be more creative, interactive, engaged, and productive.

Designing a video solution that “just works” does not need to be hard, but it does require forethought and planning to provide the most flexibility and utility for the people that use it. Here are some actionable recommendations that can help your organization:

1. Carefully evaluate where call admission control, call routing, and QoS packet markings should be done. These can be done by PBXs, video management systems, or SBCs. You will likely find that putting these into the SBC will provide more flexibility, lower maintenance costs, and deliver better system performance and interoperability. The communications market is seeing a general trend toward moving these functions into the SBC and out of the PBX, particularly in companies with multiple brands of PBXs and with varying UC desktop clients, video endpoints and infrastructure.

2. Decide where you want video signaling normalization and transcoding to occur in your organization. Both functions can occur on an SBC or on an MCU. If most of your video endpoints solutions support a common codec, such as H.264 AVC, then signaling normalization should occur on the session border controller because this will not consume expensive MCU ports. Note that an MCU requires one port per endpoint connected, so costs can escalate rapidly if you choose the MCU option for normalization. Transcoding requires more computing power and should probably be done on an MCU. If transcoding is done on the MCU, the SBC will still provide value by delivering CAC, QoS, security, and billing.

3. Plan on deploying an SBC if you will allow video calls to traverse the firewall/NAT before entering or exiting the network. Even for video systems that provide “secure” video conferencing without a VPN (i.e. Microsoft Lync and Cisco Jabber), an SBC is an important element for a secure and scalable solution.

4. If you plan to use Skype as part of your video solution, also plan on using a third-party video bridging service provider (i.e. Blue Jeans Networks, VidTel, etc.) to deliver video interoperability with your non-Skype video endpoints. The possible exception to this may be if you run Microsoft Lync, which may soon interoperate natively with Skype video.

5. If people want to use Apple FaceTime, they must understand that it is not interoperable with any other business-type video solution, and that there is no support for multiparty video.
6. Not all SBCs are created equal in terms of functionality and processing power. You will be well served by generating real use cases on how people in the organization want to use video, and then follow both the signaling and the media flow through the network and through all intermediary devices to make sure that you have the right infrastructure, including SBCs, MCUs, etc. Video bandwidth can grow quickly as the take up rate increases. It is imperative that SBC capacity can accommodate the ramp in network bandwidth without compromising features and performance.
7. There are a number of SBCs on the market, and choosing between them is not always easy. Before buying any SBC, put it into your network and run through your own use cases to make sure it will do the session management your organization needs. Furthermore, verify that the SBC will provide all of the signaling normalization, interoperability, and security functions your organization and its people require.

About the Author

Dr. Brent Kelly is Principal Analyst at KelCor, Inc. and Vice President and Principal Analyst at Constellation Research, Inc. He focuses on the intersection of technologies comprising unified communications, social business, video, cloud services and mobility. Dr. Kelly provides strategy and counsel to key client types: Chief Information Officers, Chief Technology Officers, investment analysts, VCs, technology policy executives, sell side firms and technology buyers.

Previously, Dr. Kelly served for ten years as a partner at Wainhouse Research where he was the primary author of many unified communications reports and forecasts. Click here for a complete listing.

Expertise

Dr. Kelly has experience as the Vice President of Marketing for Sorenson Vision, an early innovator in the IP communications space, and he has served as the chief executive in a privately held manufacturing company. Prior to this, Dr. Kelly was part of the team at Schlumberger that built the devices Intel used to test Pentium microprocessors. He also led teams developing real-time data acquisition and control systems, and adaptive intelligent design systems in several Schlumberger Oil Field services companies including 4 1/2 years doing R&D in France.

Dr. Kelly has worked as a research engineer for Conoco, implementing more efficient mathematical convergence methods for oil reservoir simulators, and as a process engineer for Monsanto. He has also worked as an assay technician in the mining industry.

Media Influence

Dr. Kelly is a regular presenter at Enterprise Connect (formerly VoiceCon), the communications industry trade show where his well respected half-day tutorials have covered topics such as hosted and managed unified communications services, Microsoft OCS and Lync technical deep dives, and IBM Lotus Sametime architectural reviews. He has also taught seminars in North America, Europe, Australia, and South America.

Education

Dr. Kelly has a Ph.D. in engineering from Texas A&M University specializing in thermodynamics and a B.S. in engineering from Brigham Young University. He serves as an elected official in his community.

About KelCor

KelCor (www.kelcor.com) is a specialized consulting and analyst firm with a passion for providing client satisfaction through product and service excellence. We have laser focus on the business communications market, emphasizing those products and services that are proven to accelerate an organization's business processes.

We provide value to our end-user and vendor clients by offering an unbiased, 360° view of the unified communications and collaboration marketplace. We prepare research reports, vendor profiles, market forecasts, white papers, case studies, and presentations designed to inform, educate, and assist vendors with strategy, tactics, market approaches, and end user attitudes to help them identify and capitalize on opportunity. We help our end-user clients understand options, strategies, competitive vendor offerings, and best practices engaging our collaborative process engineering expertise, all designed to improve organization efficiency while increasing top line revenues or bottom line profits.

KelCor data and reports provide more detail and analysis about collaborative markets, offerings, and strategies than anyone. The depth of our reports and our ability to discern key market trends significantly differentiates us from any other consulting and analyst firm you've ever worked with.