LifeSize™

High Definition Audio Conferencing:

Critical Success Factors for Technology Decision Makers

There have been very few meaningful technology breakthroughs in the field of audio conferencing for more than a decade. However, this field is now poised for significant leaps in technological innovation. Why is this important? It matters because technology advancements in audio conferencing represent an opportunity to drive significant productivity gains with clear implications for enterprise profitability and success.

As Wainhouse Research contends in one recent report on the "competitive advantage" associated with conferencing solutions, "New business communication tools deliver productivity improvements in two different, but intertwined, ways. The first is that many collaborative solutions can reduce the need for travel, and travel reductions have many benefits. Besides the obvious elimination of airfare, taxi and hotel expenses, eliminating travel eliminates many hours of downtime and days away from the office. The second major benefit [generated by audio conferencing and other rich media communication solutions] accrues from faster, more effective communications and problem solving."

New technological innovation can enhance the quality of conference calls, improve communication and collaboration, and have a significant impact on productivity. As Wainhouse suggests, such productivity gains will rise in relation to the elevated quality and the growing frequency of audio conference calls.

Why Audio Conferencing Technology is Important

Due to the active use of audio conferencing in for-profit and non-profit sectors, the state of this conferencing technology is a relevant issue from the standpoint of organizational effectiveness. While it is important at the outset to recognize productivity factors, this white paper will focus on the issues to be considered when assessing audio conferencing solutions *from a technical perspective*.

To maximize productivity benefits through audio conferencing, technology decision makers must recognize a set of criteria that are standard to the industry and consider how different vendors stack up against these criteria. While this white paper makes the case for a new solution from LifeSize Communications, readers are encouraged to compare and consider the different vendors in the audio conferencing industry to ensure they are making the best decision for their organizations.

Critical Success Factors When Assessing Audio Conferencing Solutions

Several key criteria for assessing audio conferencing technology are explored in this white paper. Among the criteria explored here are: *input quality; output quality;* and *connectivity*.

INPUT QUALITY. This criterion revolves around the quality of sound captured by an audio conferencing system. Ultimately, the high definition audio conferencing experience we explore here is determined by one's ability to pick up voices and other relevant audio signals with great clarity, while eliminating irrelevant noise. Here are some of the key factors that influence input quality:

Architecture and Performance. Almost all conference phones on the market today use multiple microphones in order to provide audio pickup over a broad area. Most of these older microphone designs use a number of directional microphones (anywhere from 3 on the low end up to as many as 9 depending on whether you add an extra microphone pod (mic pod) to the product). However, for several reasons, typically only one of those microphones is active at any one time. The result is that each signal transmitted to the far end comes from of a single microphone on the sending side (at any one time).

By contrast, LifeSize has introduced an entirely different, "omni-directional" architecture that differs markedly from these old style conference phone designs. The LifeSize Phone™ consists of a circular microphone array of 16 microphones, arranged around the circumference of the phone.

In the LifeSize microphone array, the signals from all of the mic elements are active all of the time. In engineering terms, the signal-to-noise ratio of a given system is increased by a factor of 3dB for each doubling of the number of sensors. So, if you go from 1 active mic to 2 active mics, then you gain an additional 3dB of signal-to-noise performance. Likewise, if you go from 2 mics to 4 mics, then you gain another 3dB on top of that.

The fact that the LifeSize array has 16 microphones, all of which are able to contribute to the signal all of the time, means that the LifeSize architecture gives an additional overall boost to the system's signal-to-noise ratio of almost 12dB. That's a much better signal-to-noise performance than any other conference phone on the market today, at any price point. That corresponds to clearer audio and more productive conference calls.

As a result of this capability, one doesn't need extra mic pods to provide full room coverage (as is often the case with other systems). That means less expense, fewer wires on the table, a much more elegant solution. With a single phone, LifeSize can cover about 2X the linear distance (which means almost 4X the amount of square footage depending on the room) of any other competing audio conference solution. Additionally, a LifeSize Phone can be daisy chained (up to 2 phones) for very large conference rooms, again without additional mic pods and associated wires.

Frequency Response. This factor heavily influences the clarity of communication and our ability to understand participants on the other end of a conference call. There is no question that we communicate most effectively when we are face-to-face and in the same location. When we communicate remotely, however, the effectiveness of the interaction depends on the degree to which the signals from the remote site match this ideal face-to-face interactive experience. When the signal that comes to us from the far end is missing some key component (such as the high or low frequencies), then the quality and effectiveness of the communications experience can suffer greatly. That is not to say that we cannot communicate at all in such a situation; just that a low quality experience makes it more difficult. In such a situation, the conference participants are more likely to experience fatigue, as they may be required to repeat themselves several times or to listen very closely in order to understand the person speaking on the other end of the conference call. This is especially taxing when participants speak with different accents in today's global communication environment.

The range of the audio frequencies transmitted is extremely important in terms of conference quality. Frequencies on the low end of the audio range give you a sense that the person on the other end is actually in the room with you. It's an issue of *presence*. On the high end of the range, what you get is clarity. It is an issue of *intelligibility*.

Some of the most expensive audio conferencing products available today are actually restricted in software from transmitting any frequencies below about 200 Hz. The reason for this restriction is that the microphones are highly susceptible to vibration noise – someone tapping a foot on the conference table leg or a truck driving by outside, for instance. In order to prevent this noise from being transmitted to the far end, the system must eliminate it before it is encoded and transmitted. Unfortunately, this means that any of the desired audio in that frequency range is also removed from the transmitted signal.

On the high frequency side, many of the microphones in these same designs are unable to pick up what normally would be clearly audible frequencies (e.g., above 5kHz) unless the person speaking is seated directly in front of the microphone. The electronics of the phone may very well be capable of transmitting such frequencies, but if the microphone cannot pick them up, then they can never make it to the far end.

LifeSize relies on a new design that is insensitive to vibration and other low frequency noise. With a special mic suspension system that isolates vibration noise down to well below 100 Hz, LifeSize can demonstrate that its mics are about 20 decibels better than the best competing solutions on the market. On the high end of the frequency range, the omni-directional mic system enables the LifeSize Phone to consistently outperform competing designs by a factor of 10 and generate high definition audio signals – even as audio conference participants freely move about a room.

Signal-to-Noise Ratio. This measure of audio quality expresses the ratio between the magnitude of a meaningful signal and the magnitude of the undesired noise in a system. There is some amount of noise present in all audio systems. In addition to unwanted structural vibration, powerline hum and radio frequency interference (RFI) noise sources, there are other significant contributors to this unwelcome noise. These other contributors include acoustic sources, such as air conditioning vent noise, computer and/or projector fans, as well as the electrical noise inherent in the microphone system itself.

All microphone systems exhibit some amount of "self-noise" which is the signal that shows up at the system output even if there is no acoustic energy present in the room itself. The LifeSize audio system uses specially designed microphone elements that exhibit very low self-noise as well as very high sensitivity. This latter attribute means that the LifeSize microphones can pick up even very faint acoustic signals (such as those coming from someone whispering across the room, for example) more easily than the microphones that are used in standard conferencing phones. When considered individually, each of the LifeSize microphone elements is about 4 times (around 12dB) more sensitive than those used for many of the competing designs.

As mentioned, the LifeSize circular array architecture provides an overall boost to the LifeSize system's signal-to-noise ratio of almost 12dB. This signal-to-noise performance, which is superior to all other audio conference systems in the market today, contributes to the value of the LifeSize phone as a productive tool.

Directional Designs. Directionality is important to ensure human voices are picked up differently than other background noise (such as the air conditioning system or a projector on the table). If all microphones pick up signals

MEASURING PERFORMANCE:
The Directionality Index vs. Array Gain

When one has a system where the microphones' directional pattern is fixed, then the Directionality Index (DI), which is expressed in decibels, can be used as an effective performance indicator. However, DI is not a true measurement of how well the system can isolate desired signals from undesired signals. For example, the DI measurement really only holds true if the system is set up in an ideal scenario where the desired signal is exactly in front of the active microphone and it has a very directional characteristic, whereas all of the interfering noise sources are evenly spread out over all other angles. This condition is rarely the case for a real system. Quite the opposite in fact; interfering noise sources are almost always very directional in nature. Thus, the often quoted DI metric is typically not really a reliable indicator of real world system performance.

When you want to characterize a more practical situation, where the noise sources themselves have directional characteristics (just as the desired signals do), you need a more dependable method for measuring real world performance. The metric by which the effectiveness of these adaptive systems is assessed is called the "array gain", which is measured in dB, in a similar manner to the much simpler DI metric. One of the major differences between these two measurement mechanisms is that the array gain metric takes into account the geometry of real sources as well as other environmental factors, such as reflections, in order to determine the actual performance of the system. An interesting comparison can be drawn in the method by which the two measurements are made. The DI of a microphone is usually measured in an anechoic chamber (not a typical location in which to hold an audio conference). The array gain of a system is typically measured in the same real world environment in which the system is designed to operate.

A major part of LifeSize's performance advantage is linked to its ability to steer a 20 dB null across a broad frequency range and keep it within 5 degrees of the interfering source. This is accomplished with patented "beam forming" technology.

equally from all directions, then the fan noise would be just as loud as someone speaking at the same distance.

In a typical conference room, there are a number of undesired noise sources (such as the projector fan mentioned earlier) that can potentially interfere with the desired speech signal. Thus, it is desirable to have some way for the phone to isolate the speech signal from the interfering noise sources. This is typically accomplished by using a number of directional microphones in different orientations.

The LifeSize Phone's omni-directional microphone however, use a sophisticated mathematical technique called "beam forming" to produce a number of "virtual" directional signals. Much of this technology was pioneered by scientists, who were designing sonar arrays for nuclear submarines and phased array radar installations for detection of aircraft at great distances. These systems were designed for people whose lives depended on how well they could detect and distinguish the direction of arrival of faint signals that were far away and hidden among a number of interfering signals.

Just like these advanced military systems, the omni-directional pattern of the LifeSize circular array is entirely software controlled. Thus, there are no "best" or "worst" directions, and the directional pattern can be varied to suit the real environment. For example, consider the case of a conference room with an inconveniently located noise source, like a projector fan at the head of the table. A software-steered microphone array like the LifeSize system can automatically adapt to the situation and steer a sharp "null" towards the offending noise source. This kind of flexibility is simply not possible with the older style fixed directionality conferencing systems, where the best thing that one can do is to physically move the microphone units in order to attempt to minimize the impact of such annoyances. LifeSize's beam forming technology, however, enables audio conference participants to ensure high definition sound quality is generated.

OUTPUT QUALITY. While it's critical to be able to effectively capture audio input in order to generate a high definition signal, the quality of the conference call is experienced in the audio output – the sound generated by the speaker. Here are some of the key factors that influence output quality:

Distortion. A simple definition of distortion is the addition of extra components to a signal that are not present in the original. If you don't receive all of the original signal in the first place, that's bad enough. But if you then add insult to injury by introducing additional signals that are not present in the original, then it just makes it that much harder for the listener to piece together the original information at the far end. But in another sense – a much more measurable sense – distortion also has an impact on the performance of an audio conference. That problem lies in the heart of one of the last real technological innovations to appear in the audio conferencing world: the Acoustic Echo Canceller (AEC).

To put it in simple terms, the AEC is responsible for searching for and canceling out the signal that is being sent to the loudspeaker whenever it shows up in the signal that is being received from the microphone. This is what keeps a given signal from looping back and forth between these two components in an endless cycle in the audio conference system (a condition that is commonly known as "feedback" or "howl"). The more effectively the AEC operates, the clearer the audio conference will be. However, the challenge lies in removing distortion components without eliminating the desired signals that are present in the conference.

The LifeSize system contains a specially designed (patent pending) loudspeaker system that exhibits distortion that is two orders of magnitude lower than those of many competing products. As a result, the AEC need not take a "heavy handed" approach to produce a clean signal that can then be passed on to the far end of an audio conference. The result is that the AEC operates more effectively, producing a much clearer, more natural sounding communications experience. Indeed, the LifeSize speaker design reduces distortion in some cases more than 10X over competing providers of audio conference systems.

Compression Algorithms. Existing audio conferencing phones use standard algorithms such as G.711 and G.722 as well as other proprietary approaches for compression. Most of these algorithms are limited to 3.5 kHz or 8 kHz bandwidth. In the case of proprietary algorithms, the benefits of higher bandwidths are limited to systems connected to similar proprietary vendor systems.

LifeSize uses a different approach for its super wide bandwidth audio system codec. The LifeSize Phone uses a variant of the industry standard MPEG AAC (Advanced Audio Compression) algorithm from Germany's Fraunhoffer Institute which can support the full 20+kHz audio bandwidth. This is the same algorithm that is used in Apple's popular iPod® audio player. If you download a song from iTunes®, you get an AAC file with a 128kb/s data rate. Since the LifeSize AAC codec normally runs in single channel mode, then the effective data rate is half that of the iTunes song (64kbps). But, other than that, it is using the same basic algorithm and the same compression ratio as iTunes.

The optimized LifeSize implementation of AAC for two-way communications provides higher bandwidth. In the case of an audio only call using IP, use of the LifeSize Phone can range from 100 Hz to 16 kHz wideband audio. In the case of using the LifeSize Phone when integrated with the LifeSize Room video communications systems, the transmitted audio can range from 100 Hz up to 16 kHz super wideband audio. In the case of using LifeSize Phone over PSTN, the system can deliver 200 Hz to 3.3kHz audio based on the limitation of analog bandwidth.

Additionally, the LifeSize version of AAC includes special optimizations in order for the algorithm to produce a consistently low delay between the systems' input and output. The actual delay induced by the LifeSize AAC algorithm is below 20 milliseconds (ms), as opposed to other implementations of the AAC algorithm, where the audio delay is limited to a minimum of around 43 ms. Thus, the LifeSize super wideband codec produces incredibly transparent results with an extraordinarily small delay.

Connectivity. This criterion involves the ability to flexibly connect to any type of network that is in use by the enterprise or organization. With the rapid adoption of Voice over Internet Protocol (VoIP), the ability to connect to wideband IP communications and interface to an IP PBX is a growing requirement for audio conferencing devices. Although there is significant growth in this networking trend, there is a great deal of PSTN connectivity to the conference room and executive office that may also be a requirement for connection to the audio conference phone. Lastly, in the case of using video communication systems, full integration with an audio conference device can allow the system to leverage super wideband audio to provide very high quality audio for use during a videoconferencing session (or in stand alone mode).

Some of today's conferencing phones provide an either/or option in terms of connectivity, usually either PSTN or VoIP. LifeSize approached the development of an audio conferencing device with the goal of providing one architecture that provides connectivity via VoIP, PSTN or integration with the LifeSize video communications systems. Additionally, LifeSize integrates with standard SIP implementations and IP PBX or IP service provider implementations. In the case of a PSTN implementation, LifeSize Phone utilizes a phone interface module for power and PSTN connection. In the IP implementation, the LifeSize Phone can receive Power over Ethernet (PoE).

The Next Era of Audio Conferencing

In the audio conferencing field, there has been little change in more than a decade – and thus, little need for technology decision makers to play a leading and consequential role in the assessment of new options and opportunities. The LifeSize Phone delivers up to 2X the room coverage of existing phones, a far wider frequency response and up to 10X lower speaker distortion than existing audio conferencing phones. Finally, the LifeSize solution provides unmatched connectivity that ensures audio linkages to all kinds of networks – whether PSTN or VOIP.

With the introduction of LifeSize Phone, an entrenched and static technology has been "disrupted" by a new high definition solution. Now, there's an opportunity to revisit old assumptions and explore new ways to drive productivity gains throughout the extended enterprise and global organization. The result with LifeSize is that everyone can be heard.

901 S. Mopac Expressway
Building 3, Suite 300
Austin, Texas 78746

USA Tel: +1 512 347 9300
Fax: +1 512 347 9301
Email: info@lifesize.com