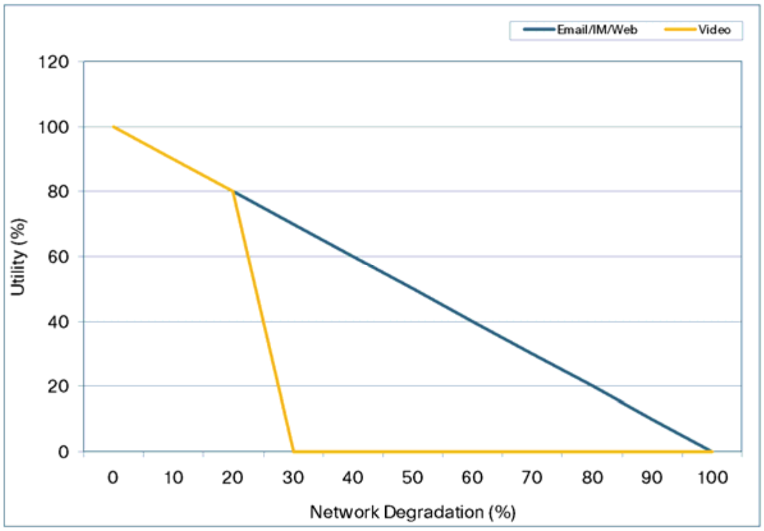
The Dawn of the Medianet

According to the Cisco® Visual Networking Index (2008), global IP traffic will nearly double every 2 years and video alone will account for 90 percent of all consumer traffic by 2012. Consumers are increasingly adopting video for their interaction and entertainment. In response to growing pressures from globalization, competition, and rising energy costs, enterprises are also adopting video technologies to improve collaboration and security, get closer to customers, and transfer knowledge. At Cisco today, more than 60 percent of network traffic is video (and the percentage is rising).

The Cisco Enterprise Technology Advisory Board in September 2008 reported that Cisco enterprise customers are experiencing an average growth of 70 percent per year in video traffic on their networks. The rising demand for video is significantly affecting business and service provider networks. Service providers see video as a way to deliver integrated consumer experiences and create new business models. As video becomes the dominant data type on the network, we need to optimize these networks to support and enhance the use of rich media.

Video is a compelling medium because 64 percent of communication is nonverbal (according to Pearn Kandola's "The Psychology of Effective Business Communications in Geographically Dispersed Teams", Cisco Systems, September 2006). Vision Group Research at the University of Oxford, UK, reports that one-third of the human cortex is dedicated to vision. Therefore, networked video is an experience technology (one that is assimilated in real time and creates the illusion of presence). Experience technologies do not degrade gracefully (that is, utility does not decline linearly with network throughput). When delay, packet loss, and jitter exceed visible thresholds, the usefulness of streamed video quickly drops to zero because it becomes unintelligible and the experience it creates becomes nonexistent. Video is a very demanding application that immediately exposes any weaknesses in the network.

Figure 1. Utility vs. Network Degradation



This growing use of video on networks requires a better intelligent network. Delivering new video experiences will place additional demands on IP networks in terms of performance, adaptability, and manageability. Networks will need to scale and deliver an optimized quality of experience (QoE), introducing additional complexity. Networks that were designed for an era of best-effort delivery, low bandwidth, and high latency do not work for video. Because video is the most demanding media type, a network optimized for interactive video will support many other rich-media types.

Networks have always evolved to respond to new requirements -- from intranet to extranet to Internet and now to medianet -- an intelligent network optimized for rich media.

The medianet is a media-optimized network comprising advanced, intelligent technologies and devices. medianets will emerge within the home, businesses, and service providers, all interlinked to deliver a more visual, more social, and more personal experience for users. The medianet is an evolution of converged IP networks: By adding medianet technologies, we can create a platform optimized for the delivery of rich-media experiences.

A medianet has the following characteristics:

- **Media-aware:** Capable of detecting and optimizing different media types (video, audio, etc.) to deliver an optimal QoE
- **Endpoint-aware:** Capable of detecting and configuring media endpoints automatically
- **Network-aware:** Able to detect and respond to changes in device, connection, and service availability

Endpoints are any devices that capture or display video, including set-top boxes, video surveillance cameras, TelePresence, webcams, digital signs, personal computers, IP video phones, and any other networked video device. A network that is media-, endpoint-, and network-aware ensures that appropriate services are configured automatically when a device connects to it. On a medianet, if the media is not in the right format for the endpoint or it exceeds the capacity of the network, it will be adapted to deliver the best possible experience.

Transmitting video presents several challenges for the network, including ensuring predictability, performance, quality, and security.

Predictability

Video consumes a significant amount of bandwidth. For example, streaming 1 hour of high-definition (HD) video consume tens of gigabytes of bandwidth. Without special handling, video does not permit graceful degradation in the presence of network congestion. In order to guarantee a predictably good experience, it is necessary to ensure that bandwidth is available when producing or consuming video on the network. In a medianet, video endpoints negotiate to obtain the desired bandwidth before they transmit and the network is capable of assigning the required resources. In a traditional IP network, endpoints transmit and hope for the best. When congestion occurs, all transmitting devices are affected. Dynamic Call Admission Control (CAC) ensures that devices can request bandwidth and are informed of changes in available bandwidth so that they can act accordingly. If the endpoint has an adaptive encoding technology, it may be able to reduce the bandwidth required by changing the way it encodes video (by reducing resolution, frame rate, or bits per pixel). H.264 Scalable Video Codec (SVC) is a particularly elegant video encoding standard that encodes video at varying resolutions (called layers), allowing it to degrade much more gracefully than most other video codecs. In the presence of congestion, it is possible to discard some of the layers (reducing visual detail) without dropping frames (which is much more noticeable).

Performance

Packets flowing through a network are prone to delay, loss, and jitter.

Delays arise from physical limits (speed of light) as well as from buffers in routers and gateways that the packets traverse along the way. When delay increases above 400 msec (camera to display), people become aware of it and it starts to impede interactive communications.

In the presence of congestion or errors, IP networks can discard packets (and then retransmit them). In a real-time interaction, it is often impossible to retransmit lost information (it takes too long to get the missing packets). Forward-error-correcting algorithms can embed additional information in the media stream, allowing the endpoint to reconstruct missing or corrupted video data. A medianet is network-aware and can engage these kinds of mechanisms as appropriate.

Jitter is the variability in delay, and buffers are used to smooth out these variations. In order to ensure proper playback, the network must have accurate timing information. All devices must have their clocks synchronized to ensure that smooth playback. Too much buffering in the network will prevent effective interactive video. A medianet relies on precise clock synchronization and end-to-end measurement of network delays to use the least amount of buffering necessary.

Quality

We obtain higher quality in video by using higher resolutions, more colors (increased bits per pixel), spatial audio (multiple audio channels and higher sampling rates), and multiple displays. All of these factors increase demand for bandwidth. We are increasingly engaging in multistream interactions (multiple audio and video streams combined to form a single immersive experience). For example, Cisco TelePresence™ technology consists of multiple HD video and audio streams combined to deliver the illusion of shared presence. Cisco WebEx® conferencing also can include multiple participants on webcams and telephones. These multistream interactions place additional demands on the network because all streams now must be handled as one to ensure a consistent experience (they need to stay perfectly synchronized). All streams must take the same path through the network and must be given the same priority in order to avoid problems such as lip-synch errors (when audio is out of phase with video). It is no longer possible to prioritize network traffic solely by media type (video first, audio second, then text applications); all streams of the same interaction must be handled with the same quality-of-service (QoS) guarantees.

These multistream interactions can be point-to-point, multipoint, or multicast. A point-to-point interaction involves two endpoints, whereas multipoint can involve many endpoints. Multicast interactions are a scalable way of delivering one-to-many interactions (one source, many destinations). When all the endpoints are of the same type, it is a symmetric interaction. If some of the endpoints are different, it is an asymmetric interaction. In the case of asymmetric interactions, media and signaling translation may be required to allow all the endpoints to participate in the same interaction. A medianet provides transcoding services (resizing and changing the format of video) to produce the best possible quality for each endpoint.

Service providers play a particularly important role in the delivery of networked video. They can offer premium IP services to businesses and consumers that ensure video traffic is allocated sufficient bandwidth, with the right service levels to preserve the QoE from end to end.

Security

In addition, voice and video communications are increasingly being encrypted at the endpoint, limiting the ability of the network to automatically infer traffic priority based on an analysis of the packet contents. The medianets therefore must rely on other mechanisms to ensure that the requisite QoS and bandwidth is allocated to rich-media traffic. Metadata is a form of tagging that sends additional contextual information alongside the media stream. This metadata can describe the encrypted stream, indicating its source and importance (as well as other contextual information). This information can then be used to ensure that the right QoS levels are used.

Creating secure point-to-point video communications between companies is particularly challenging because in addition to the standard delay and jitter constraints imposed by video, we also need to resolve encryption, firewall traversal, and addressing concerns (how to find and get at the remote video endpoint). The medianets provide a standard way to handle this scenario, so broader adoption of video in intercompany collaboration is expected.

Differences Between Current IP Networks and medianets

By being media- and endpoint-aware, medianets provide a better experience to the end user, use network resources more efficiently, and minimize complexity for the IT organization. The medianets extend traditional IP technologies by adding protocols and devices to:

- Deliver predictable performance
- Automatically adapt to varying network conditions
- Optimize video playback by adapting signaling and media
- Automate and ensure security

These extensions make medianets a much better platform for video applications than traditional IP networks (Table 1).

Table 1. Current IP Networks vs. medianets

When	Current IP Network	medianet
Plug in a new video endpoint	• Endpoint gets an IP address	• Endpoint gets an IP address • Endpoint is automatically configured and provided with required bandwidth, QoS guarantees, and security
Find video	• Need to know the precise location and name of a video file in order to download or play a video	• File virtualization enables endpoints to refer to video files by name (not necessary to know the location) • Video is streamed or buffered automatically from nearest source, depending on available network bandwidth
View video	• Provides best-effort playback • Displays nothing if endpoint does not support video format • If video format exceeds available bandwidth or device processing power, choppy video results • Video stream is interrupted when roaming	• Provides best-quality playback: Video playback is optimized for endpoint capabilities and currently available bandwidth • Offers fast and secure roaming between access points without interrupting video stream
View video on different types of endpoints	• Some devices can play back video, but some cannot • Video quality is not optimized for endpoint or link speed	• Video is reformatted in the best possible format that each device is capable of displaying • Video quality is determined by display size, supported video formats, and currently available bandwidth • The network optimizes bandwidth usage

Distribute rich media	• Media is pushed from a single source and is available for playback only when the entire file has arrived at the destination, making it difficult to distribute real-time events (debates, sports, and news)	• Topology-optimized peer-to-peer replication accelerates distribution of content and allows for immediate playback with just-in-time caching
Network congestion happens	• Glitches occur in the video playback or video playback stops altogether	• The network constantly monitors and reroutes to avoid congestion from occurring • Forward error correction is used to allow the endpoint to repair packet loss • The network reencodes video streams to reduce bandwidth consumption
Network capacity is oversubscribed	• All network users suffer glitches or dropped connections	• Less-critical applications get “busy signal” or the network reencodes video streams to reduce bandwidth consumption
Link failure occurs	• Poor video quality and service is disrupted during the failure and before convergence occurs	• Milliseconds convergence minimizes disruption • Highly available mechanisms completely eliminate disruption
Send video from one company to another	• Endpoints behind firewalls are difficult to reach	• Offers simple, consistent, and secure way to reach endpoints behind firewalls
Encrypt video streams	• Difficult-to-use QoS prioritizes video traffic	• Metadata sent along with the media stream ensures that video remains prioritized

How Network Devices Will Evolve

Networks today consist of many different devices. These devices will evolve over time to include more and more medianet technologies (Table 2).

Table 2. How Network Devices Will Evolve

Device	medianet Changes
Video endpoints	Additional protocols for service discovery, CAC, QoS, performance monitoring, and security will be added to all video endpoints to provide a standard interface for network services. This standard interface provides a ready-to-use capability when endpoints are added to a medianet (they automatically configure themselves and inform the network of their presence and capabilities). It also provides a feedback mechanism to endpoints so they can respond to network requests. Endpoints will emit metadata streams (to provide contextual information about the media) alongside media streams that will be used by routers and gateways.
Switches	The standardized set of services that are added to endpoints will also be added to switches so that they can provide a proxy for traditional endpoints that will not be upgraded to have these services, making it possible to add medianet capabilities to traditional endpoints but without the ready-to-use and the feedback benefits that medianet video endpoints enjoy.
Routers	Routers will monitor the network for current bandwidth conditions, apply the autodiscovery and autoconfiguration capabilities for the endpoints, and redirect media streams to gateways as required to adapt them to changing traffic conditions or different endpoint requirements. Routers will also monitor the video quality and take appropriate actions to deliver optimal QoE. Routers will increasingly rely on metadata sent along with media streams to make control and experience optimization decisions.
Gateways	Gateways will carry out signal and media conversions to adapt video streams to the requirements of different endpoints and changing traffic conditions.
Storage	Storage appliances at the edge and core of medianets provide local playback of stored media, scaling video playback and saving bandwidth. These devices will perform streaming, caching, replication, and peer-to-peer synchronization of media, letting endpoints refer to media by name (independent of location), thereby simplifying application development. The storage appliances maintain a global name space of all content (they know where everything is) and will stream or cache the requested content on demand.

In addition to adding medianet technologies to existing network devices, new medianet devices will emerge (content virtualization appliances, streaming servers, and media processing engines).

How Customers Will Implement medianets

The increasing adoption of new video applications (for example, TelePresence, video surveillance, or digital signage) will require investments in new network infrastructure. Customers will choose to add medianet-capable devices to their network to increase predictability, performance, quality, and security. The medianet devices will continue to coexist with traditional IP infrastructure for some time. The medianet devices will be added to support the most demanding video applications first. As more video applications are added, the medianet infrastructure will help accelerate the deployment of these new applications, minimize the complexity, and help to scale the infrastructure for the best QoE. Over time, as the traditional devices become obsolete, they will be replaced with medianet devices. When the network integrates more medianet devices, it will progressively become part of the larger medianet network.

To determine which part of their traditional IP networks customers need to upgrade first, Cisco offers a Media Ready Network program. This program comprises a collection of design guides, white papers, documented use cases, as well as a set of services to prepare, plan, design, implement, operate, and optimize a successful and step-wise migration to a medianet.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

Printed in USA

C22-512174-00 12/08