Network Services from Iformata Communications

Marshall Eubanks

marshall.eubanks@iformata.com

November, 2007

What is TelePresence ?
And Why Should an ISP Care ?

- Telepresence in this context is very high quality video conferencing combined with an environment designed to produce a feeling of presence between remote participants.
  - It is **not** just HD videoconferencing.
    - But it does depend on high bandwidth video.
  - After a long period of gestation telepresence has finally become a major market.
  - It has major attractions for the Middle East.
  - For more info, see http://www.humanproductivitylab.com/

iformata
COMMUNICATIONS

In the beginning…

- There was "the TeleSuite from TelePort Corporation."
  - In 1995
    - See, e.g., Wired Magazine, January 1996
      http://www.wired.com/wired/archive/4.01/beta.html?pg=6
  - Telepresence over T1's from IBM
- In the last decade, we have gone from Leased Lines and Bundled ISDN Circuits to Nested Full Mesh MPLS.
- I will show how Telepresence Networking is different from the needs of "normal" IP, and how the Telepresence SLA's drive Network design.
- While this talk is focused on network requirements, let's start by looking at Telepresence itself.
  - There are two types of units, full room and screen based.

iformata
COMMUNICATIONS

Polycom RPX

iformata
COMMUNICATIONS

Cisco Telepresence

Photo Courtesy of the Human Productivity Lab

How do you make distance disappear ?

A combination of art and science !

- Art :

    Make the room environments consistent

    Hide the cameras

    Hide the microphones

    A lot of subtle "stagecraft"

- Technology :

    Good cameras

    Low latency

    Low packet loss (or erasure protection)

The network is crucial !

There is one thing more : The MCU

- If all you ever do is point to point meetings, all you need is some telepresence units.
    - This is exceedingly rare in practice.
- If you want to do multi-point meetings, or connect different types of devices, you will need a bridge.
    - In the V/C world, this is called a Multi-Point Control Unit, or MCU.

iformata
COMMUNICATIONS

Network SLAs and QOS

- Service Level Agreements (SLAs) and Quality of Service (QOS) depend on metrics

- What are the QOS metrics on Telepresence ?

- The important parameters include
  - End to end throughput
  - Service Availability
  - Delay (or latency)
  - Delay jitter (or variation)
  - Packet loss

Telepresence needs high bandwidth and high availability

- End to end throughput
  - The original TeleSuite was configured to fit within 1 T1 (1.5 Mbps)
  - Now, with HD, more bandwidth is needed
    - 10 Mbps for 2 screen Polycom RPX's (20 Mbps for 4 Screens)
    - 15 Mbps for Cisco Telepresence (@1080 p)
    - 45 Mbps (DS3) deployment for Teliris

  I strongly recommend a margin over the vendor's recommendations !

- Service Availability
  - These are executive units, and executives expect that they will be available. Except for planned maintenance, they should be available 24x7x365.
  - However, in practice, units are never actually used 24x7
    - This makes it easier to schedule maintenance downtimes.

©2007 Iformata Communications

iformata
COMMUNICATIONS

Network Delay

- Not too important in typical web traffic.

- Quite important in telepresence and videoconferencing

- Excessive delay frustrates human interactions.
  - By ITU-T G.114, 150 milliseconds (msec) or less of end to end delay is acceptable, delays over 400 msec will be considered objectionable
    - (These are one way delays)

- Since light in fiber goes at ~ 0.6 c, your antipode is 111 msec away, *not* counting equipment delays.
  - These will be at least 1 - 3 frames, or 30 to 100 msec

- With good network routing, acceptable delays are possible *in principle* to anywhere on Earth.

The Jitter SLA Metric

- The video codecs do not "care" about the network delay
  - Even though users do…
- BUT, if there is delay jitter > one frame, then the packet will have to be dropped, and that is bad.
- At 30 fps, a frame is ~ 30 msec, so you want jitter to be < 30 msec
- With jitter buffers, you can increase this somewhat, especially for units that are reasonably close to each other.
- 100 msec is a reasonable upper limit.
- Meeting this boils down in practice to avoiding router queues, and thus avoiding congestion.

The packet loss SLA is driven by the video codecs

- For most video, there is a lot of redundancy between successive frames of video.
  - This is especially true with video conferencing, which involves remarkably little basketball.
- But, over time, things do change, and sometimes drastically
  - People and things move around.
  - Lighting changes, etc.
- So, the most efficient codecs encode differences between frames, and estimate how things have moved between frames.
  - Interframe compression
  - Motion Prediction
    - i.e., difference two different areas of the frame to better compress moving objects.
- Over time, these two ideas have been developed quite a bit.

iformata
COMMUNICATIONS

A Brief History of Video Codecs

- MPEG-4 is a late 1990's update to MPEG-2
  - Published 1999

- At the same time, the ITU was working on H.263+ / H.263++ / H.26L standard extensions.

- In 2001, the ITU VCEG and the ISO MPEG joined forces
  - H.264 was published in 2003. It is also MPEG-4 Part 10 (not version 10!).

- H.264 seems to be the codec of choice for Telepresence going forward.
  - The Polycom RPX / HDX
  - Cisco Telepresence
  - HaiVision hai1000 codec

Video Codecs and Network Requirements

- The various codecs in use for Telepresence are all quite similar from a network point of view

- All use the "Group of Pictures" or GOP concept.

- And, this drives network QOS requirements.

- I want to "step through" this to illustrate the needs of Telepresence QOS.

MPEG-X & H.26X

- All of these standards have similar frameworks
  - The fundamental basis for compression is the macroblock (16 x 16 luma pixels or 8x8 chroma pixels), arranged into Slices, and then into frames.
  - All allow the use of previous (or future!) frames to predict the current frame (or macroblock)
  - Encoding is thus the compression of a prediction residual.
  - All allow for motion compensation to improve interframe prediction.
  - All use block based transforms and quantization to low pass filter the residual visual information

iformata
COMMUNICATIONS

The Group of Pictures (GOP)

- There are three kinds of frames:
  - I (intra-coded)
  - P(predictive-coded
  - B (bidirectionally predictive-coded)
- There is one and only one I frame per GOP
  - It is encoded by itself, with no information from other frames
- P frames are encoded using the difference from the last I or P frame.
- B frames are typically not used in videoconferencing
- So, a 20 frame GOP stream looks like
  - IPPPPPPPPPPPPPPPPPPP IPPPPPPPPPPPPPPPPPPP…

Not all MPEG packets are created equal

- An I frame encoding is less efficient, so it might be 10 times as big as a P frame for the same video quality.
  - And the quality of the I frame determines the quality of the entire GOP.
  - Typically in HD an I frame is >> 1 packet. A P frame may not be.
- Unless there is a repair mechanism, packet losses from I and P frames cause video errors that persist until the next I frame.
  - If any one of the I frame packets are lost, there will be errors persisting for the entire GOP.
  - So, without protection, packet losses need to be << 1%

An Example

- At 1 Mbps, 30 fps and an MTU of 1450 bytes, there will be ~ 90 packets / sec (pps). With a GOP of 30 frames (or 1 second), there will be about 20 packets per I frame and 2 per P frame (totaling 78 pps for video).
- Suppose there is a random (Poisson) 1% packet loss rate. With no protection
  - 54 % of all seconds will have video errors
  - The *average* duration of video problems will be 0.47 seconds
- For a loss rate of 0.1%, 5 seconds out of 60 have problems, each with an average duration of 0.47 seconds
- This is for *one* screen.

The Loss SLA Metric

- Packet losses need to be very small, or you need some means of repair them, or both.
  - For a mean time between artifacts of 1 hour, in our example the loss rate should be $< 10^{-5}$ (with no repair).
  - Cisco Telepresence, for example, recommends a loss rate $< 5 \times 10^{-4}$
  - Iformata's goal is $< 10^{-5}$
- This may a problem for wireless LANs, where losses of a few per cent are not unusual in cases of, e.g., RFI.
- For wireline IP, bit error rates due to "thermal noise" might be 10^{-12}, for a packet loss rate of 10^{-8}
- This sounds good, but TCP tends to grab all of the bandwidth available, and uses packet loss to signal congestion

Protection…

- How to protect Telepresence packets ?
- You could use TCP
  - But, then, your conference might stall waiting for an old packet
- With UDP there is
  - Error concealment, which works with P frames, not so well with I frames
  - Forward Erasure Correction (FEC)
    - Iformata is working with Polycom and others to separate FEC from the application, in the FECFRAME Working Group of the IETF
  - Note that any FEC will increase the latency.

…Or Reservation

- Since the problem is sharing TCP traffic with Telepresence traffic, another form of protection is reservation.

  – The old way : Dedicated Circuits

  – The new(er) way : Diffserv is used to protect Telepresence over star topologies or MPLS networks.

  – The way of the future : Nested Full Mesh MPLS networks with Traffic Engineering (TE).

Dedicated Circuits are easy…

- Well, they seem that way. Issues are
  - Obtaining any connectivity in some parts of the world is difficult
  - A full mesh doesn't scale well at all, and is typically quite wasteful as the number of end points goes above ~ 4
  - There are typically latency and efficiency issues with star topologies (telepresence between Hyderabad and Bangalore may have to go through New York!)

iformata
COMMUNICATIONS

MCU Placement is an Issue

- Unless all of your Telepresence is point to point, you will need access to an MCU to bridge multi-point calls.
  - MCU placement will effect latency and scalability.
  - MCU's require bandwidth sufficient to support all calls going through them.
    - A star network for each MCU ?
  - For efficiency, MCU's need to be scheduled well. As MCU's have limited ports, and limited bandwidth, they need Call Admission Control also, and this is frequently more demanding than end point CAC.
  - MCU Placement is thus not trivial and, for example, is continuously re-evaluated by Iformata.

VNOC and POP Map
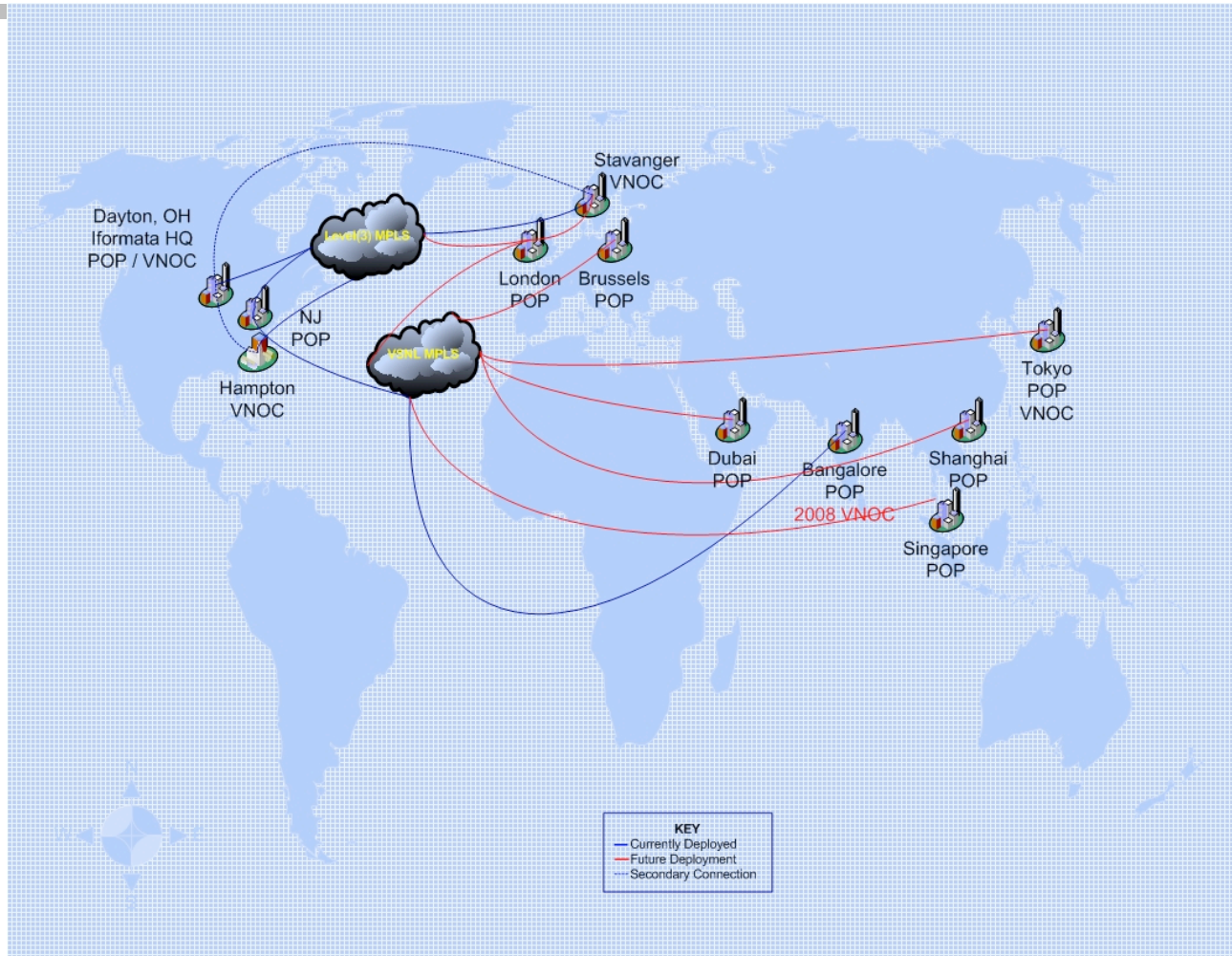
Traffic Engineering

- Iformata uses Diffserv to protect audio, video and control packets on its networks. All other traffic (e.g., data sharing) is given "best effort" priority.

- Diffserv is not enough
  - It only provides good service if you are not overloading your bandwidth.
  - This requires either
    - Always having enough bandwidth for any possible use or
    - Call admission control (CAC) to reserve bandwidth

- RSVP (Intserv) is an automatic standards based solution for Call Admission Control
  - RFC 3209 extensions for MPLS-TE

iformata
COMMUNICATIONS

Full Mesh MPLS

- The evolving industry solution to the issues with point to point circuits involve MultiProtocol Label Switching (MPLS)
- This allows
  - Packets to be tagged so that flows between locations can be scheduled
  - Traffic engineering can be used to reserve / protect bandwidth between end points
  - The network can appear *logically* to be full mesh (connections between all end points) even though *physically* it is not.
  - This requires setting up tunnels between all possible end point pairs
    - For N end points, N (N -1)/ 2 tunnels
- Modern MPLS networks can pick up Diffserv Class of Service Code Point tags applied at the Telepresence unit itself.
  - The MPLS port does not have to be at the Telepresence unit itself.

iformata
COMMUNICATIONS

The Trouble with Full Mesh

- The trouble with Full Mesh MPLS is that the number of tunnels grows quadratically with the number of end sites
    - For 10 End Sites :   45 tunnels
    - For 20 End Sites : 190 tunnels
    - For 30 End Sites : 435 tunnels
- These tunnels require router resources
    - While very large VPN's have been implemented, in my experience the cost / availability of Full Mesh MPLS networks with Telepresence bandwidths effectively limits their size to dozens of end nodes.
- However, the sociology of Telepresence is that
    - Users form Communities
    - There is a lot of communication within communities
    - There is not nearly as much communication between communities
        - In most cases, the community is the customer's company.

©2007 Iformata Communications

iformata
COMMUNICATIONS

Nested Full Mesh MPLS

- The obvious solution is a Nested Full Mesh
- A Full Mesh MPLS for Telepresence within each community.
  - Limited need for admission control
    - MCU's may still need CAC
- Connections between the Full Mesh for each community and
  - Other communities
  - MCU's
  - Network aggregation points.
- Can have, for example, a star topology connecting all of the communities to a central point, or a Full Mesh MPLS network whose end nodes are entry ports into the community Full Mesh networks.
- At this level, CAC will almost certainly be needed.

iformata
COMMUNICATIONS

Multi Protocol Label Switching (MPLS)

- The MASERGY backbone runs MPLS in an entirely native environment.

- MPLS permits true, end-to-end, differentiated Classes of Service.

- Each Class of Service equates to a distinct level of traffic prioritization across the Masergy network.

- In terms of priority, latency intolerant traffic such as video and voice are always transported at the highest level.

Courtesy Masergy Communications

iformata
COMMUNICATIONS

MPLS Routing vs. "Next Hop" Routing

Service Provider Network

Source

Destination

- - - - "Next Hop"
——— MPLS
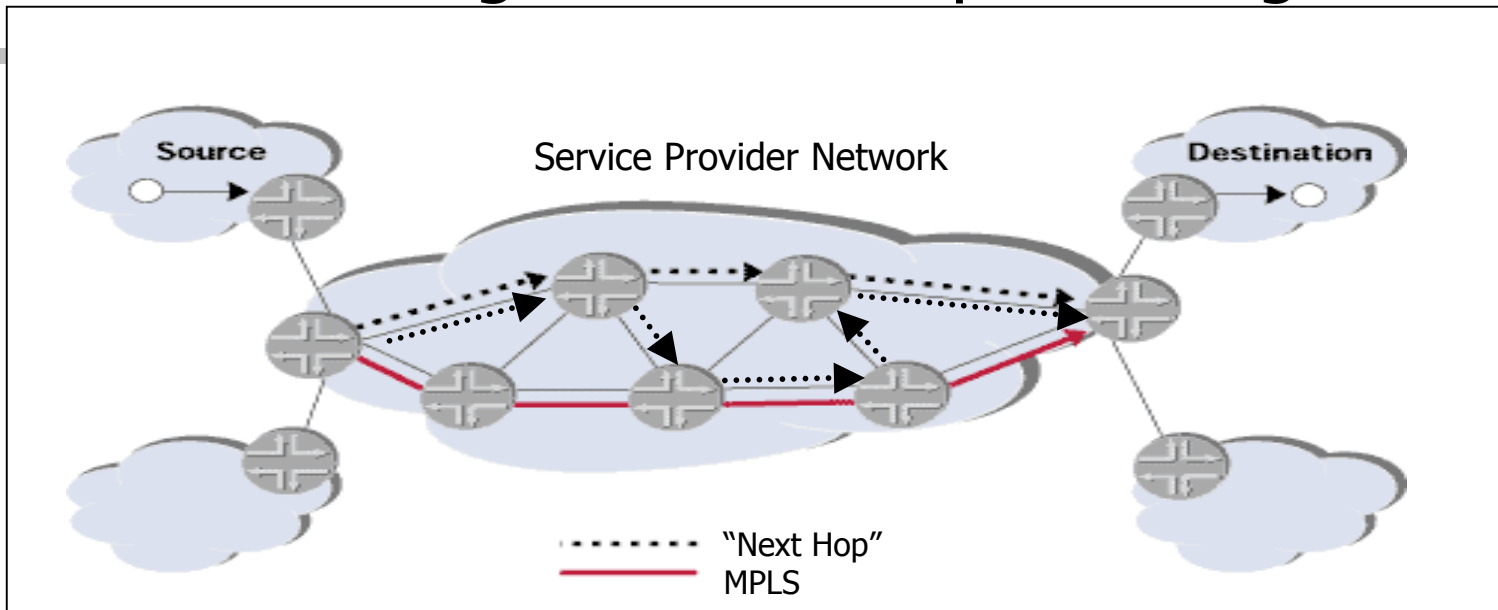
The Masergy MPLS Backbone:

- Routes packets based on a single end-to-end routing decision uniformly applied to all packets bound for the same destination.

- Result: A single physical path between source and destination that delivers video packets in-sequence with uniform latency.

Traditional IP Backbone / "Next Hop" Routing:

- Packets routed individually, router-to-router, based on continuous dynamic congestion analysis.

- Result: Multiple physical paths between source and destination that deliver video packets out-of-sequence with highly variable latency.

Courtesy Masergy Communications

©2007 Iformata Communications

iformata
COMMUNICATIONS

MASERGY NATIVE MPLS/IP BACKBONE
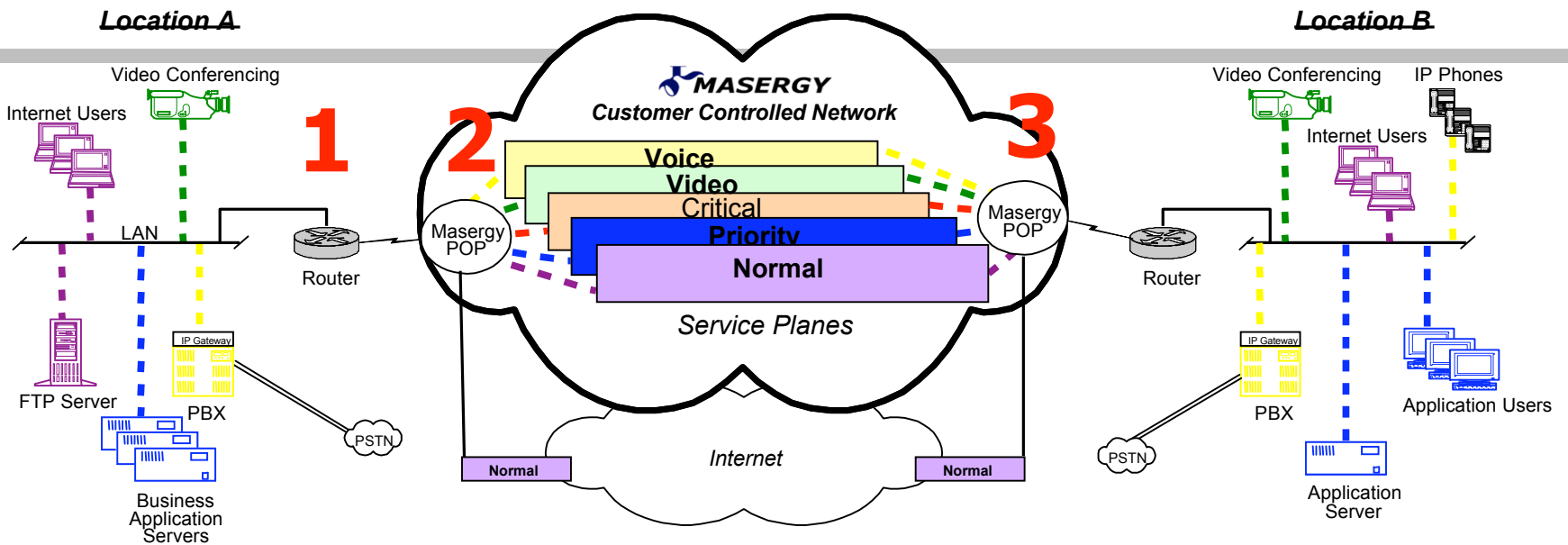
1. Assignment of various traffic types to the appropriate Class of Service (CoS) for WAN transport begins in the premise router which is configured (Class Based Queuing) to prioritize & tag outbound traffic using Type of Service (TOS) markings.

2. As traffic enters the backbone core, TOS tag values are recognized by MASERGY edge equipment and used to map packets to the correct CoS or "Service Plane" for long haul transport with the appropriate level of priority.

3. Enforcement of prioritized routing policies (based on TOS markings) by MASERGY edge equipment applies equally to traffic exiting the backbone core to ingress an individual customer site.

•Result- priority-based routing policies are applied to all traffic, at all points, permitting QoS to be protected for real time traffic and high value business applications.
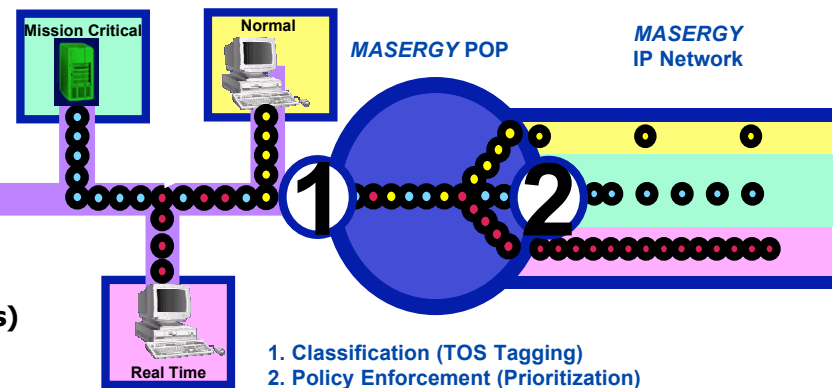
Conclusions

- In this presentation I showed the QOS metrics on Telepresence, and why they are what they are :

    - End to end throughput   ~ 10 Mbps to 45 Mbps
    - Service Availability   << 1 outage / year / site
    - Delay (or latency) $\leq$ 300 msec RTT
    - Delay jitter (or variation)   << 100 msec
    - Packet loss $\leq 10^{-5}$

Sources

- ## Cisco Telepresence :

http://www.cisco.com/en/US/netsol/ns672/networking_solutions_white_paper0900aecd805bbda0.shtml

- ## Telanetix :

http://www.telanetix.com/pdfs/telanetix_datasheetNS2.pdf

http://www.telanetix.com/about/faqstech.html

- ## Wainhouse Report :

http://www.wainhouse.com/files/papers/wr-et4tctp.pdf

- ## HPL Report :

http://www.humanproductivitylab.com/

- ## MASERGY slides from

Chris Carr
Director, Video Markets
Masergy Communications
ccarr@masergy.com
703 846 0498

©2007 Iformata Communications

iformata
COMMUNICATIONS