# Advanced Preferential IP and Telephony Services in the Internet

**Fred Baker**

**Cisco Fellow**



© 2004 Cisco Systems, Inc. All rights reserved.

# Lessons learned 11 September 2001



- There were many simultaneous individual failures
  - Some could have been avoided by using better procedures
  - Many were automatically routed around or recovered from
- On the whole, the Internet is what worked that day
  - The Internet is reliable because it is distributed in control and operation
  - Even if individual parts perform in new and exciting ways
- The Internet is useful for emergency services such as police, fire, and medicine, and for military use.
- The Internet is now critical national communication infrastructure.


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues we would like to cover today



- Existing Public Switched Telephone Network (PSTN) voice services
  - MultiLevel Precedence and Preemption (MLPP)
  - Government Emergency Telecommunications Service (GETS)
- Extensions to the Internet
  - Application to voice, video, and real time services
  - Application to file transfer, email, and messaging
- Key concept:
  - Preferential service offered to an authenticated user
- This is not e-911
  - That is preferential service offered to a random user of an identified service


© 2004 Cisco Systems, Inc. All rights reserved.

# MultiLevel Precedence and Preemption (MLPP)



- In military communications, a priority scheme
  - for assigning one of several precedence levels to specific calls or messages so that the system handles them in a predetermined order and time frame,
  - for gaining controlled access to network resources in which calls and messages can be preempted only by higher priority calls and messages,
  - that is recognized only within a predefined domain, and
  - in which the precedence level of a call outside the predefined domain is usually not recognized.

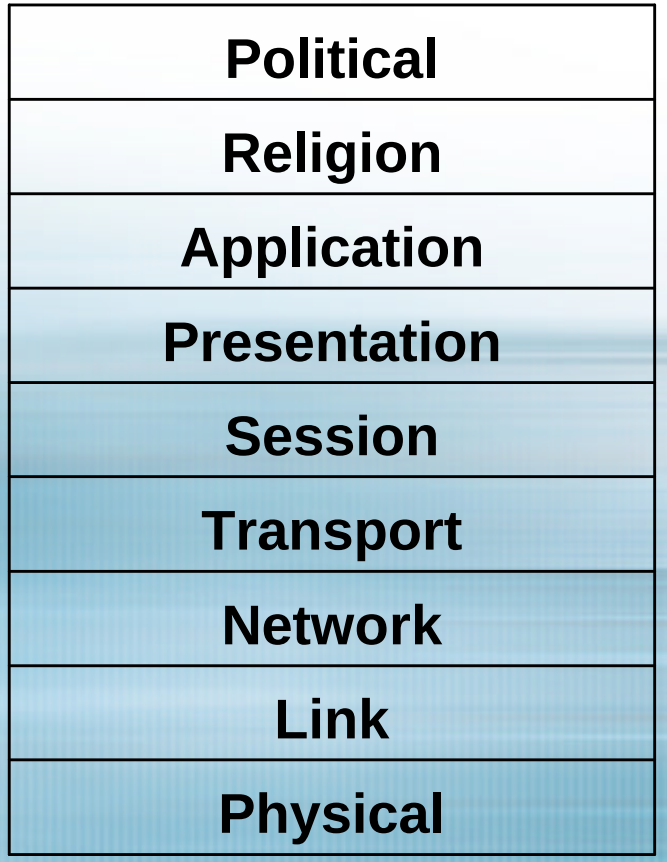
http://www.its.bldrdoc.gov/fs-1037/dir-023/_3416.htm


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues that have to be addressed



- **Issues at several layers**
  - **Politics: Who gets to use preferential access?**
  - **Application layer: is there a difference in queuing or other resources?**
  - **Session Layer: do we let script kiddies use this? How do we keep them out?**
  - **Can a transport provide predictable throughput?**
  - **Can the Network layer provide guarantees to the packets?**





© 2004 Cisco Systems, Inc. All rights reserved.

# RFC 1633: Types of applications



- **Two broad classes:**
  - **Those with fixed requirements of the network**
    - **"Real Time" applications include voice, video, telemetry, and any other application with an *irreducible lower bound* on its bandwidth requirements or an upper bound on the jitter and loss it will accept**
  - **Those that will adapt to a wide range of network environments**
    - ***"Elastic" applications adapt* themselves to any bandwidth as long as it is consistently available. Examples include (but are not limited to) TCP and SCTP-based applications.**


© 2004 Cisco Systems, Inc. All rights reserved.

# Messages from DARPA* and NCS**



- Statistical services like the Internet provide statistical guarantees
- Sometimes, people need non-statistical guarantees:
  - A certain amount of bandwidth between here and there, no excuses
  - Move a file across the network in a fixed amount of time
  - Deliver a message to a set of people within a fixed period of time

* DARPA: Defense Advanced Research Projects Agency
**NCS: Department of Homeland Defense, National Communication Service


© 2004 Cisco Systems, Inc. All rights reserved.

# Preferential Real Time (Voice/Video) Services

© 2004, Cisco Systems, Inc. All rights reserved.

# Important operational characteristics



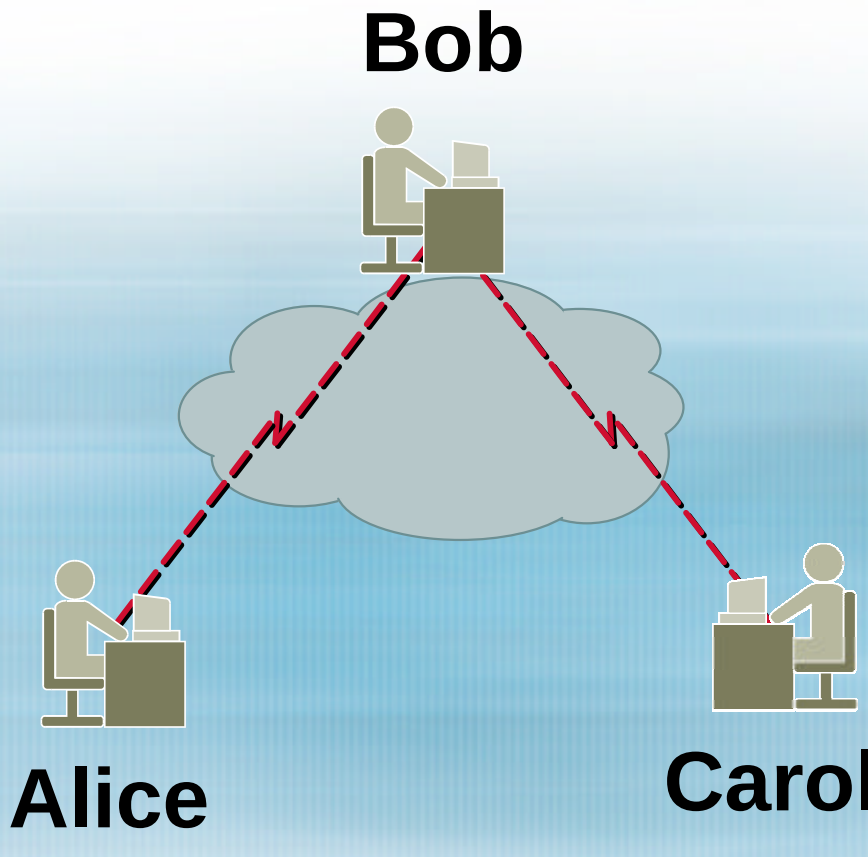
- Need to be able to authenticate and authorize certain calls
- MLPP: Need to be able to preempt calls
  - At handset – incoming call preempts existing call
  - In network – new bandwidth requirement preempts lower precedence usage of the bandwidth
- Need to signal to users using standard signals
  - Chime/tone indicating preemption
- Need to preserve existing calls at all precedences


© 2004 Cisco Systems, Inc. All rights reserved.

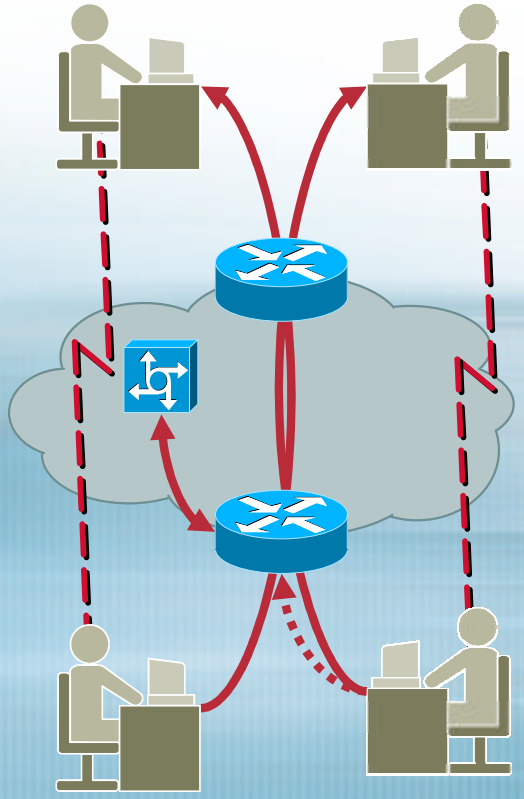
# Two fundamental preemption cases:



- Preemption at handset




- Preemption in network

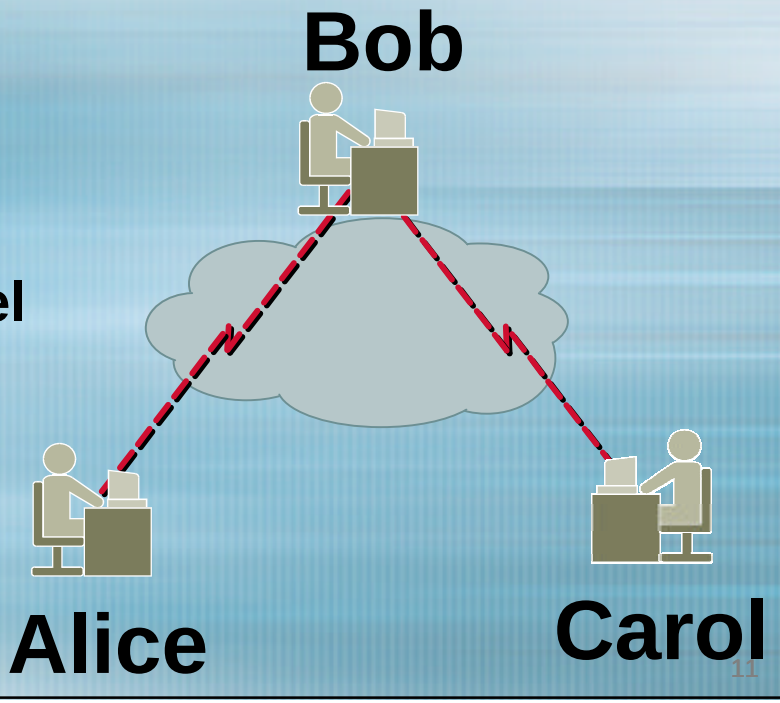



© 2004 Cisco Systems, Inc. All rights reserved.

# End system preemption



- **Deals with case where call with elevated precedence is placed to a handset that is in use**
  - **Alice calls Bob who is talking with Carol at lower precedence**
- **H.323:**
  - **H.460.mlpp**
- **SIP Solution:**
  - **SIP Resource Priority Header**
    - **Label call with precedence level**
  - **SIP Call Failure Reason Code**
    - **Reason = “Call Preempted”**





© 2004 Cisco Systems, Inc. All rights reserved.

# Network admission/preemption: Call Admission Control (CAC) solutions



- **Heuristics based on loss**
  - **Differential loss among different calls**
  - **Calls experiencing high loss might self-drop**
- **Server-based call admission**
  - **Call counting**
- **Network bandwidth admission**
  - **Using signaling with Server CAC**


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues with heuristics based on loss



- In a word, not deterministic
- In overload case, all calls at a precedence see roughly equal loss
  - All calls at a level hear a lot of noise
  - Users tend to re-call “due to bad circuit”
- If instruments apply algorithms to measure and respond to loss by dropping
  - Result is indeterminate
  - May drop too many calls
  - May incorrectly present signals to users


© 2004 Cisco Systems, Inc. All rights reserved.

# Loss and jitter are issues…

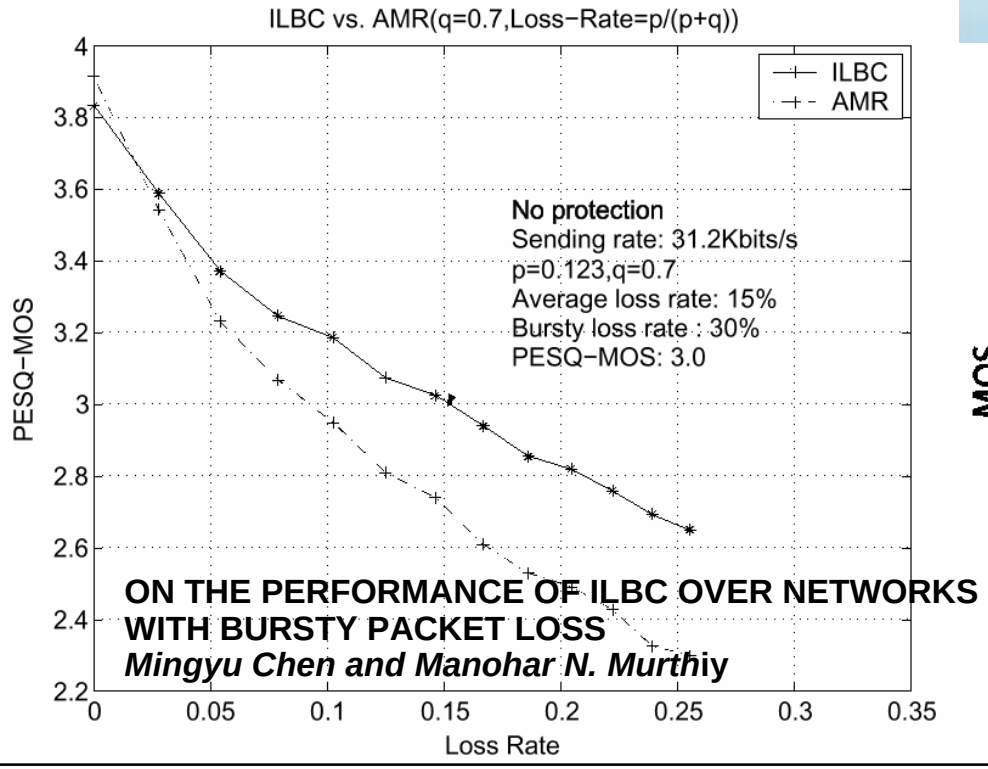


- …because jitter and loss directly affect signal quality

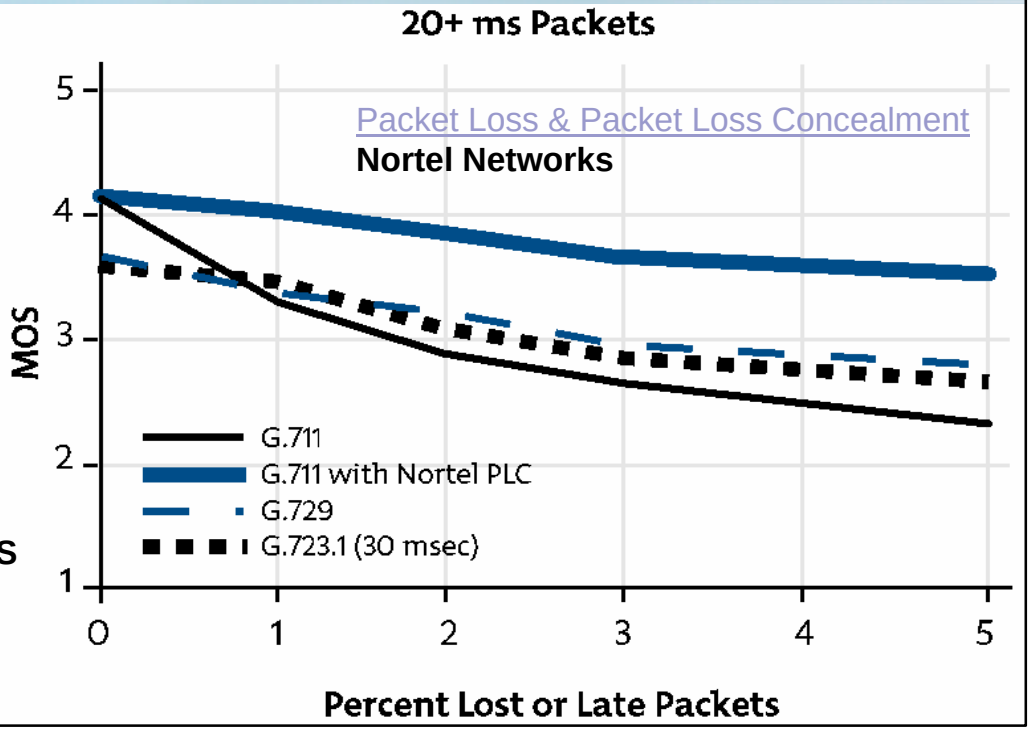
- In the end:

  service preserves higher precedence calls

  All calls that are not preempted should enjoy reasonable sound quality




**ON THE PERFORMANCE OF ILBC OVER NETWORKS WITH BURSTY PACKET LOSS**
***Mingyu Chen and Manohar N. Murthiy***



# Server-based call admission



- Server based admission implies that server knows the network
  - It has knowledge of its own domain and neighboring domains, and policies regarding codecs and permitted number of simultaneous calls
  - It does not know routing or ambient traffic loads
- Servers are often located away from tactical deployments and specific offices
  - And specifically not in routers


© 2004 Cisco Systems, Inc. All rights reserved.

# Network bandwidth admission

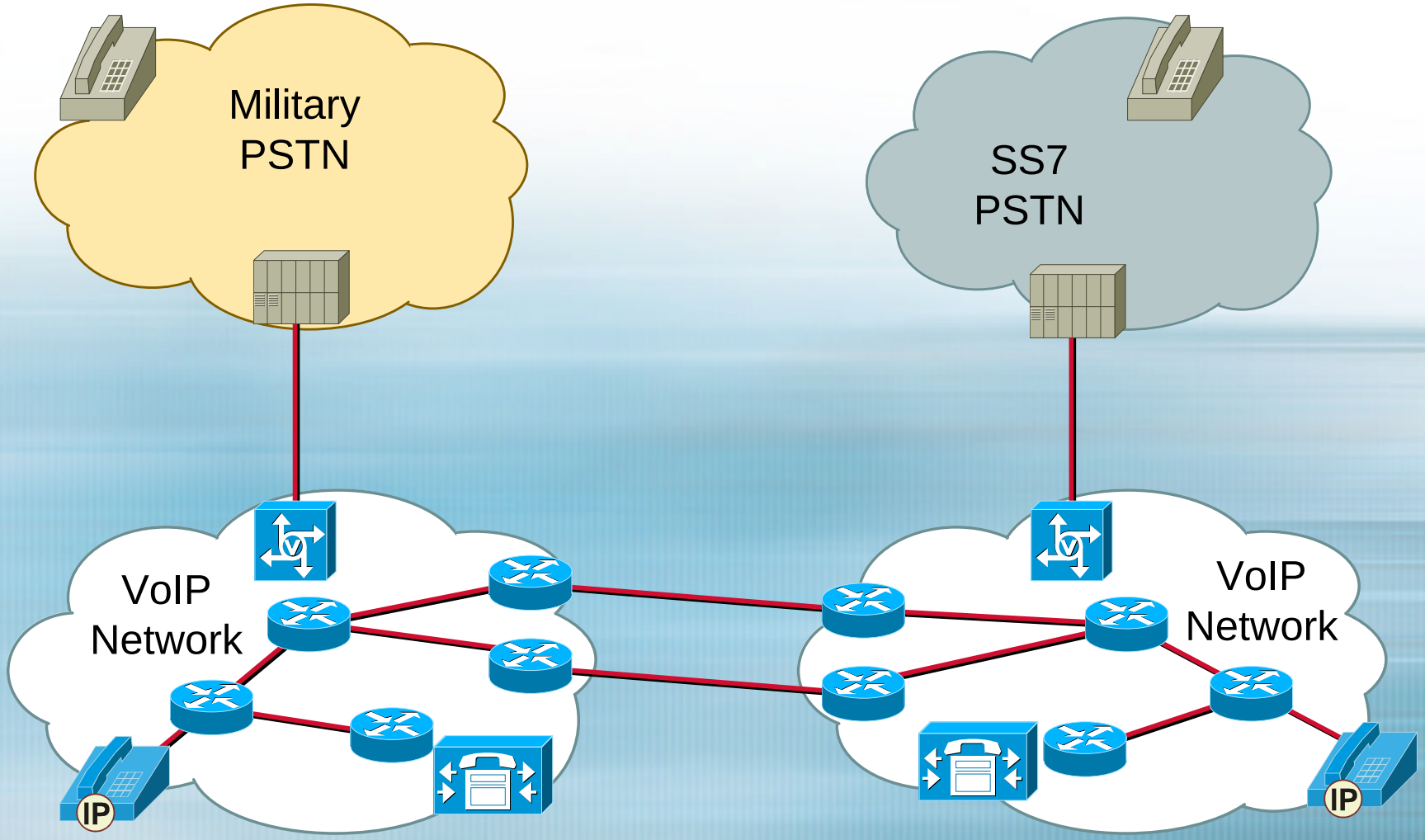


- It is possible, using RSVP to
  - Use control plane signaling to deterministically authorize/preempt bandwidth (RFC 2205)
  - Enable SIP to start data transmission only when network bandwidth has been authorized (RFC 3312)
  - Aggregate calls within backbones (RFC 3175)
  - Limit state to systems with bottleneck links
- NSIS is a possible future
  - IETF “Next Steps in Signaling” WG; work in progress


© 2004 Cisco Systems, Inc. All rights reserved.

# Control paths may not follow data paths



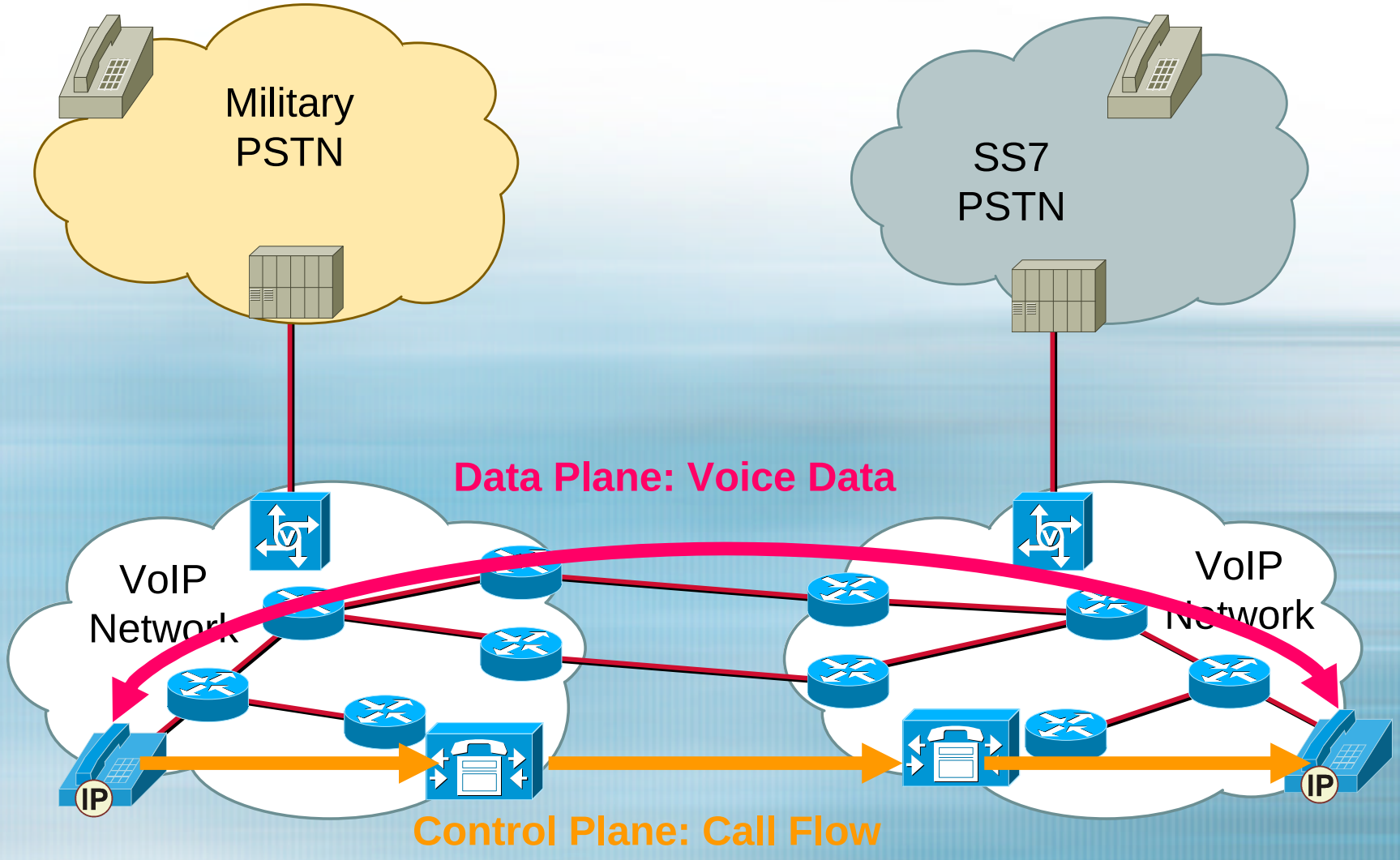




© 2004 Cisco Systems, Inc. All rights reserved.

# VoIP Call Control and Data



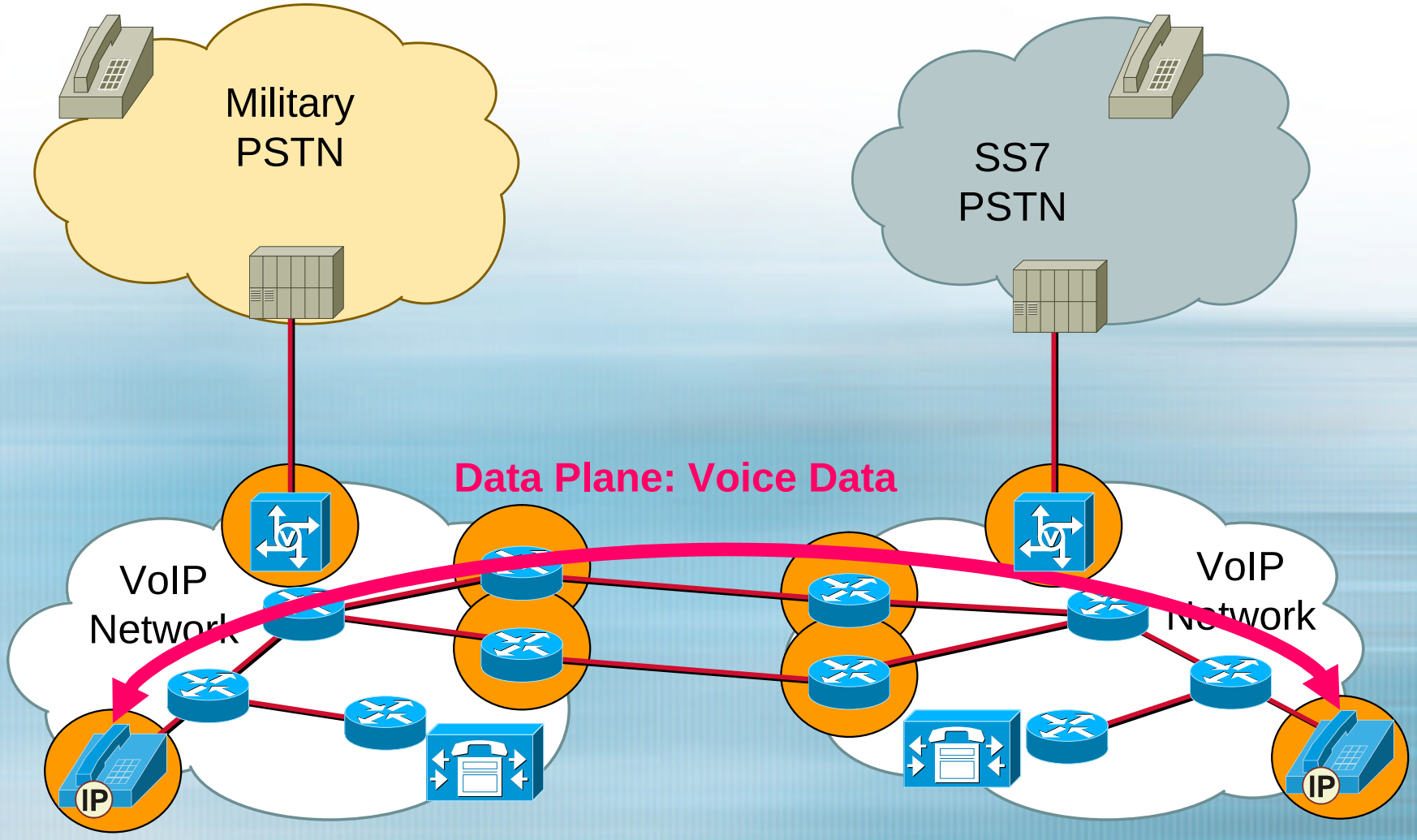




© 2004 Cisco Systems, Inc. All rights reserved.

# Where to configure signaling







© 2004 Cisco Systems, Inc. All rights reserved.

# Session AAA procedures in RSVP



- RSVP session types
  - IPv4/IPv6: RFC 2205 {IP addresses, IP Protocol, UDP/TCP/SCTP Ports}
  - IPSEC: RFC 2207 {IP addresses, IPSEC Association}
  - Aggregated: RFC 3175 {IP Addresses, DSCP}
- Policy Data Object
  - RFC 2750 RSVP Extensions for Policy Control
  - RFC 2753 A Framework for Policy-based Admission Control
  - RFC 2872 Application and Sub Application Identity Policy Element for Use with RSVP
  - RFC 3182 Identity Representation for RSVP
- Policy data object identifies:
  - Requestor, by personal identity or some aggregate identity
  - Priority of reservation as a numeric value


© 2004 Cisco Systems, Inc. All rights reserved.

# Proposal

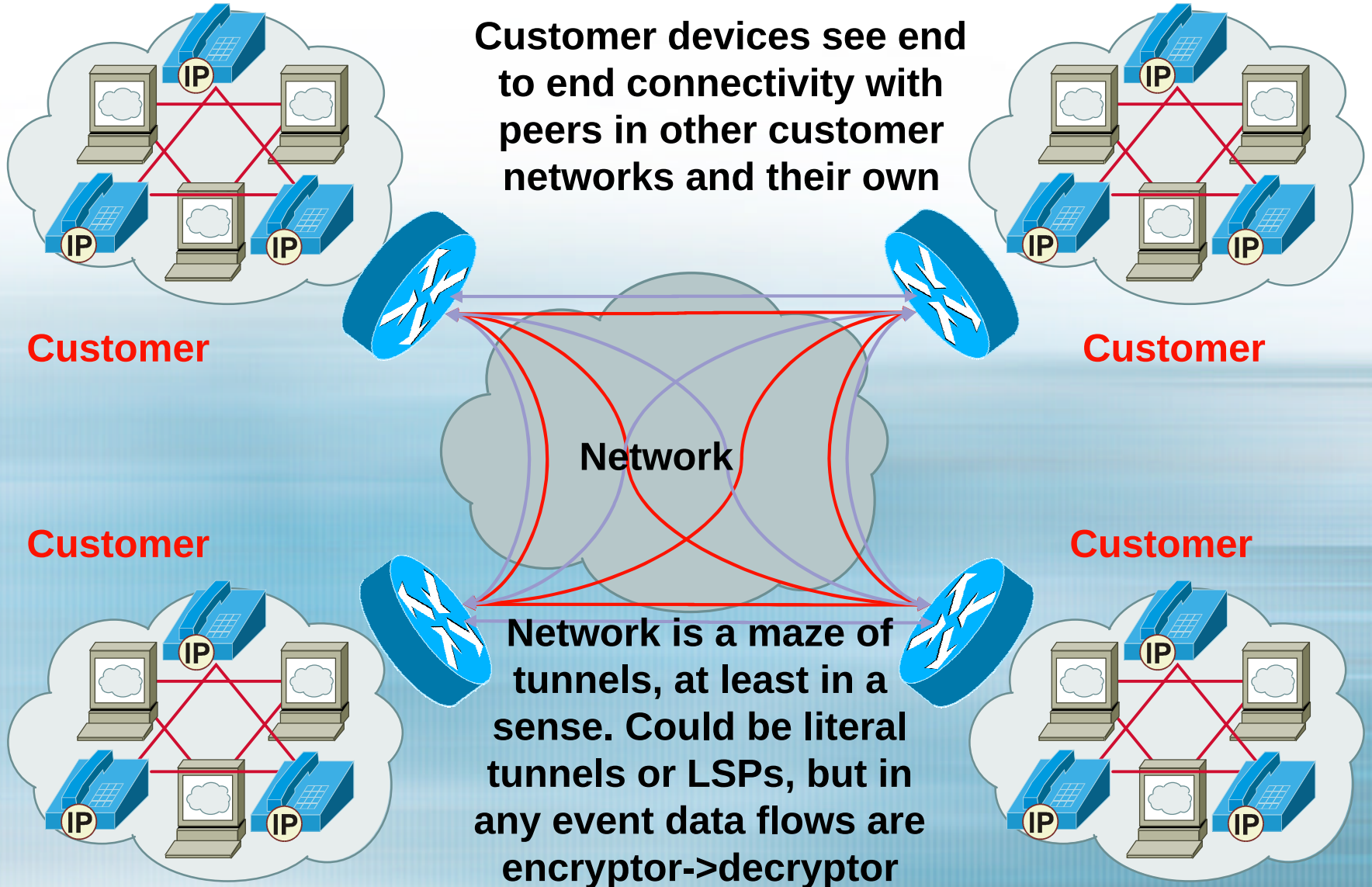


- **Use RSVP signaling in control plane**
  - **RFC 2205, 2207, or 3175**
  - **Primarily configured at interchange points**
    - **Only on systems with bottleneck interfaces**
    - **With AAA/policy/preemption capabilities as described in RFCs**
- **Use EF PHB for VoIP traffic data plane**
  - **RFC 3246/3247**
  - **Flag voice with DSCP, apply priority and simple policing to sum of bandwidths of calls**


© 2004 Cisco Systems, Inc. All rights reserved.

# Mental model of a VPN network







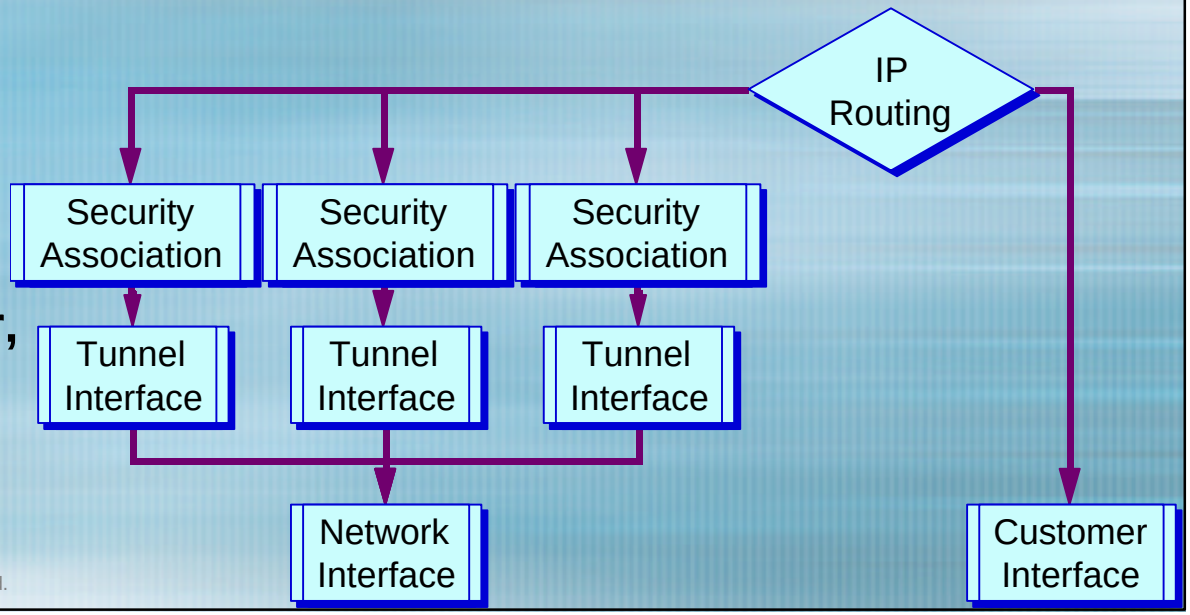
© 2004 Cisco Systems, Inc. All rights reserved.

# Structure of VPN Edge Router



- **Customer Message**
  - Received from customer Network
  - Routed to network
  - Encrypted, adding IPSEC ESP
  - Encapsulated in network IP header
  - Routed in tunnel
- **Network Message**
  - Received in tunnel
  - Decrypted
  - Routed, usually to customer, but also possible to route to another tunnel

- **Example:**
  - Configuration in VPN Network
  - Tunnels set up via Routing
  - Keys exchanged by combination of configuration and key exchange under “Unity” security architecture
  - Encrypt/decrypt performed in box labeled “security association” using relevant key





© 2004 Cisco Systems, Inc. All rights reserved.

# RFC 2207: RSVP for IPSEC data flows



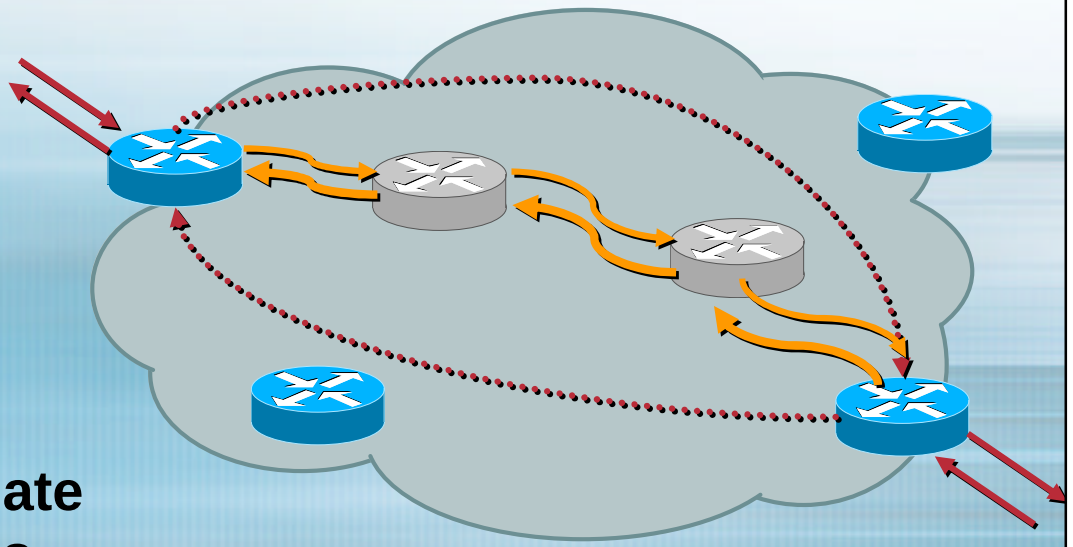
- Enables RSVP to signal QoS Flow Specification per Security Association per IP source/destination address pair
  - Two systems may have multiple security associations, and therefore SPIs
  - Data may be identified with same DSCP, as this is in the control plane
- Flow characteristics of a tunnel are the sum of the flow characteristics of the traffic in the tunnel


© 2004 Cisco Systems, Inc. All rights reserved.

# RFC 3175: RSVP Aggregation



- **Designed to permit aggregation of reservations**
  - **Service Provider Voice…**
- **Ways to identify affected traffic**
  - **IP Source/Destination/DSCP**
  - **Tunnels and LSPs**
- **Mechanism:**
  - **RSVP reservation for aggregate is the sum of the reservations using it**
  - **Traffic in aggregate identified/services by DSCP**

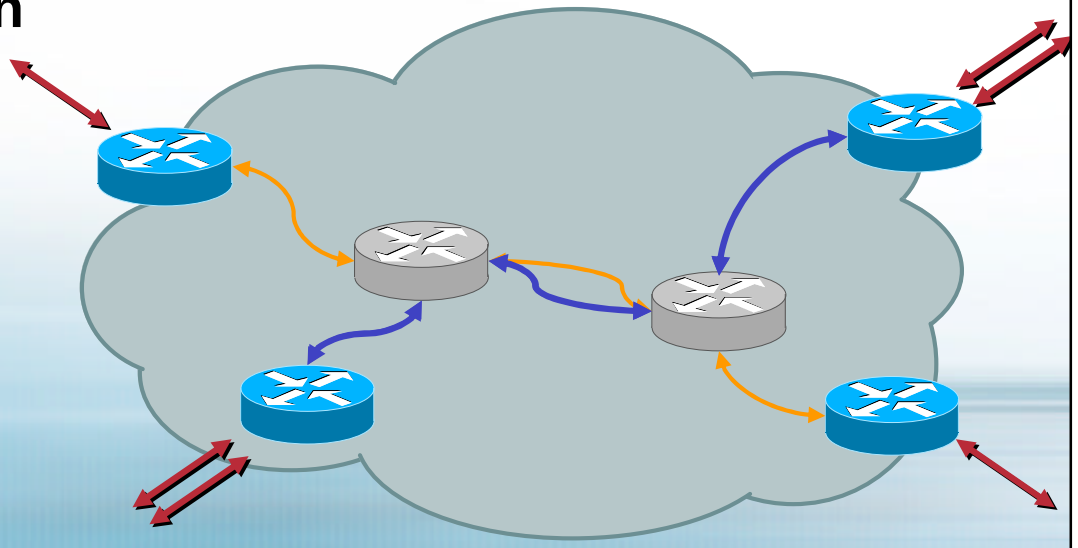



© 2004 Cisco Systems, Inc. All rights reserved.

# Preemption within an aggregate



- Procedure for preemption within VPN core
  - New call arrives at advanced priority
    - May result in new aggregate reservation for one call or increase in bandwidth of existing reservation
  - RESV_ERROR message along routine aggregate
    - Instructs de-aggregator to reduce bandwidth
    - MLPP policy used to decide which customer call to drop
  - Lower priority aggregate
    - Reinstated at lower bandwidth or removed



- Key implication:
  - Customer and network sides of VPN router have to exchange information


© 2004 Cisco Systems, Inc. All rights reserved.

# Key information leakage observation:



**Traffic in the network can be measured by any device in the network.**

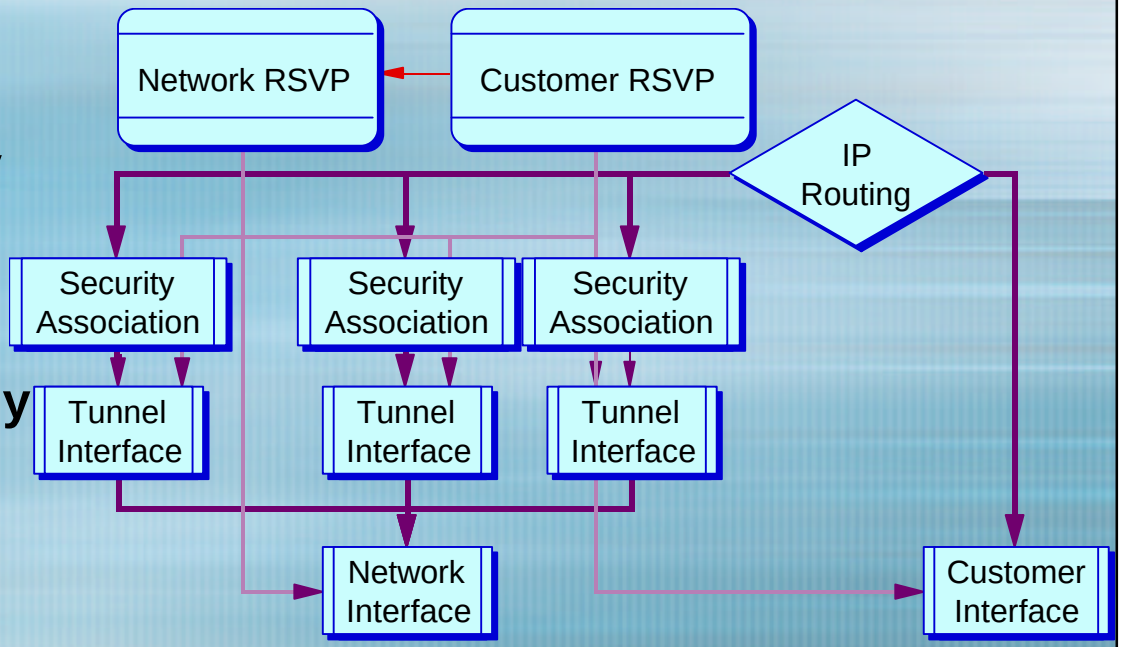
**Meta-statements about that traffic, such as reservation information or MPLS-TE bandwidth commentary, conveys no information that could not be obtained by measurement.**


© 2004 Cisco Systems, Inc. All rights reserved.

# Proposal



- **Customer:**
  - **Same as in normal network**
  - **as far as it knows, it is a normal network**
- **Network:**
  - **Advertise PATH/RESV for sum of sessions using a tunnel**
  - **If new network reservation refused, apply MLPP policy to refuse or preempt customer reservations**
  - **If old network reservation preempted, repeatedly apply MLPP policy to customer reservations until network reservation is accepted**
- **Use RSVP 2207/3175**
  - **Reservations for IPSEC security associations, one SA for best effort traffic and one for each level of VoIP traffic.**
  - **Bandwidth of a tunnel £ sum of mean rates of sessions using tunnel**



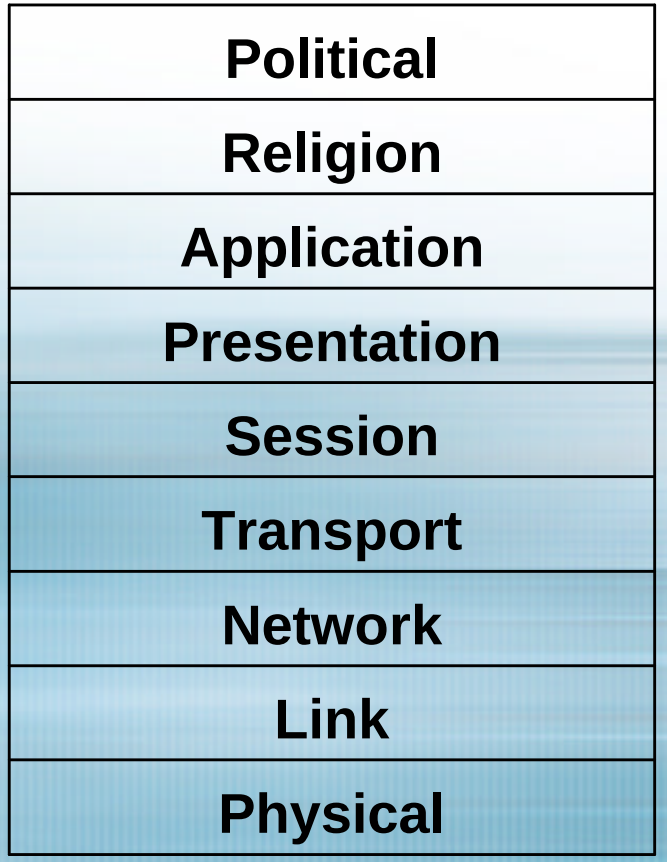


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues addressed in this solution



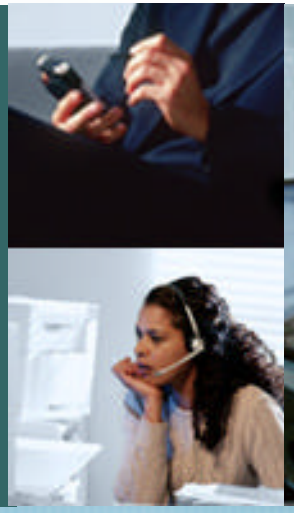
- **Who gets authenticated to use preferential access?**
  - **That is a policy question, not addressed in the technology**
- **Application layer issues:**
  - **Endpoint preemption, Authentication, Authorization in SIP**
- **Session Layer controls:**
  - **RSVP provides bandwidth/route-aware CAC**
  - **Session AAA issues handled in policy**
- **Network layer guarantees?**
  - **Yes, now that we know we have available bandwidth**





© 2004 Cisco Systems, Inc. All rights reserved.

# Preferential Elastic Services



© 2004, Cisco Systems, Inc. All rights reserved.

# Preferential Services for Elastic Applications



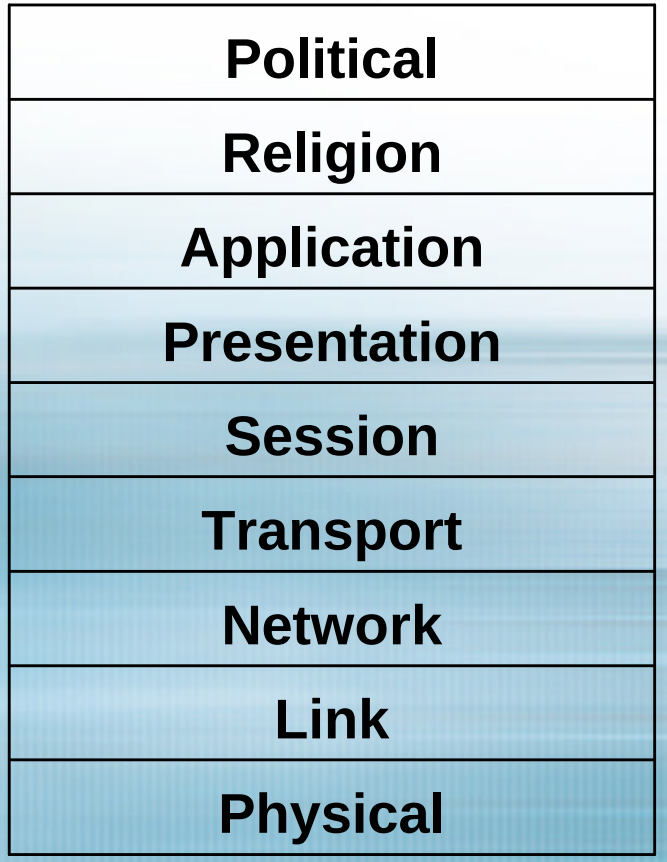
- **Example applications**
  - **File Transfer**
  - **Transaction applications**
  - **Instant Messaging**
  - **Electronic Mail**
- **Issue: Predictable access during congestive collapse**
  - **Large proportion of network out of service**
  - **Ongoing electronic attacks**
  - **Heavy internet traffic from users**
  - **Emergency management needs access**


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues that have to be addressed



- **Issues at several layers**
  - **Who gets authenticated to use preferential access?**
  - **Application layer: is there a difference in queuing or other resources?**
  - **Session Layer: do we let script kiddies use this? How do we keep them out?**
  - **Can a transport provide predictable throughput?**
  - **Can the Network layer provide guarantees to the packets?**





© 2004 Cisco Systems, Inc. All rights reserved.

# Network layer procedures

- Political
- Religion
- Application
- Presentation
- Session
- Transport
- **Network**
- Link
- Physical


© 2004, Cisco Systems, Inc. All rights reserved.

# Two approaches to making guarantees to IP traffic



- **Drop precedence**
  - **Permit traffic with multiple DSCPs to use the same queue**
  - **Make some DSCPs compete more effectively for queue bandwidth**
- **Class provision**
  - **Divide traffic with differing DSCPs into different queues**
  - **Give different guarantees – or no guarantees – to various queues**


© 2004 Cisco Systems, Inc. All rights reserved.

# Drop precedence effects



- RFC 2581 TCP, or similar SCTPs, respond to loss by slowing down
  - Therefore, loss is a signal to a sending TCP or SCTP
- RFC 2597 Assured Forwarding
  - Assures at least minimal service to each competing session
  - “Excess” users tend to absorb drops, and therefore slow down
  - Tends to guarantee fairness
- Assigned drop precedence (“routine”, etc)
  - Preferred sessions still compete with lower priority sessions
  - No guarantees to preferred sessions beyond “less loss”


© 2004 Cisco Systems, Inc. All rights reserved.

# Effect of a class definition



- **Similar to TDM definition of a multiplexing channel:**
  - **Guarantees a minimum level of bandwidth if there is traffic to use it**
- **But "work conserving":**
  - **Unused bandwidth is redistributed among other classes**
- **Implication:**
  - **Cost of an unused class is nil**
  - **Makes bandwidth guarantee to traffic should traffic show up for it**
  - **Radical effect on traffic: traffic in another class that is using redistributed bandwidth suddenly reduced by amount that preferred class uses**


© 2004 Cisco Systems, Inc. All rights reserved.

# Building preferential classes



- **Services defined:**
  - **GETS requires one “emergency traffic” class**
  - **MLPP requires several graded traffic classes, each for all elastic traffic at that class**
- **My suggestion:**
  - **Define classes and apply bandwidth to them**
- **Issue in this:**
  - **How do we ensure that only authorized users use these classes?**


© 2004 Cisco Systems, Inc. All rights reserved.

# Security: who is authorized to get preferential access?

Political

Religion

Application

Presentation

**Session**

Transport

Network

Link

Physical


© 2004, Cisco Systems, Inc. All rights reserved.

# Identification, Authentication, Authorization



- **GETS/MLPP Policy:**

  **Provides service to an *authenticated user* who *requests* the service**

- **Need to provide ways to**

  **Authenticate the user**

  **Let him/her request the service**

  **Authorize the usage**

  **Provide access to the authenticated user**


© 2004 Cisco Systems, Inc. All rights reserved.

# Authenticate the user



- **This uses an existing AAA function:**
  - **TACACS, Radius, Diameter, etc**
  - **Uses standard AAA facilities and procedures**
- **Requirements:**
  - **Simple/scalable to deploy to hundreds of thousands of users**
  - **Strong enough to prevent unauthorized access**
  - **Examples: one-time password, cryptographic authentication with PKI, Kerberos**


© 2004 Cisco Systems, Inc. All rights reserved.

# Authorize the usage



- On presentation of request and credential, must be simple to implement (such as updating ACL)
- Effect is to provide privileged access to the authenticated user


© 2004 Cisco Systems, Inc. All rights reserved.

# File Transfer Services

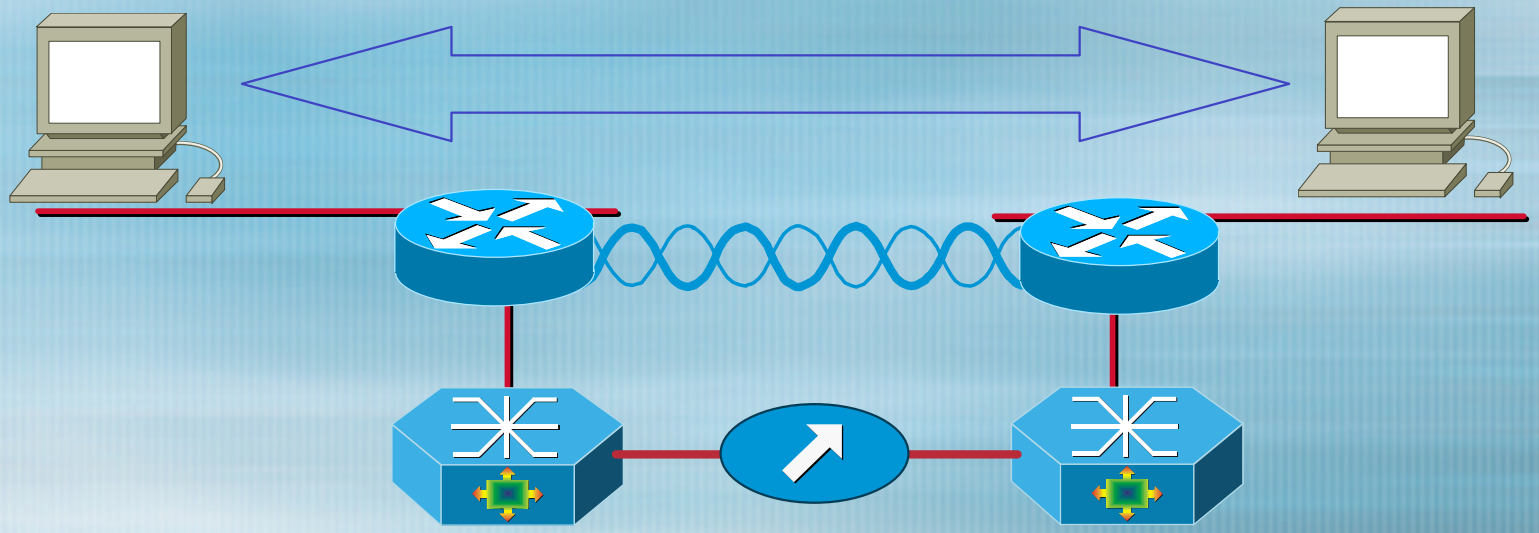
- Political
- Religion
- Application
- Presentation
- Session
- **Transport**
- Network
- Link
- Physical


© 2004, Cisco Systems, Inc. All rights reserved.

# File transfer is particularly part of this



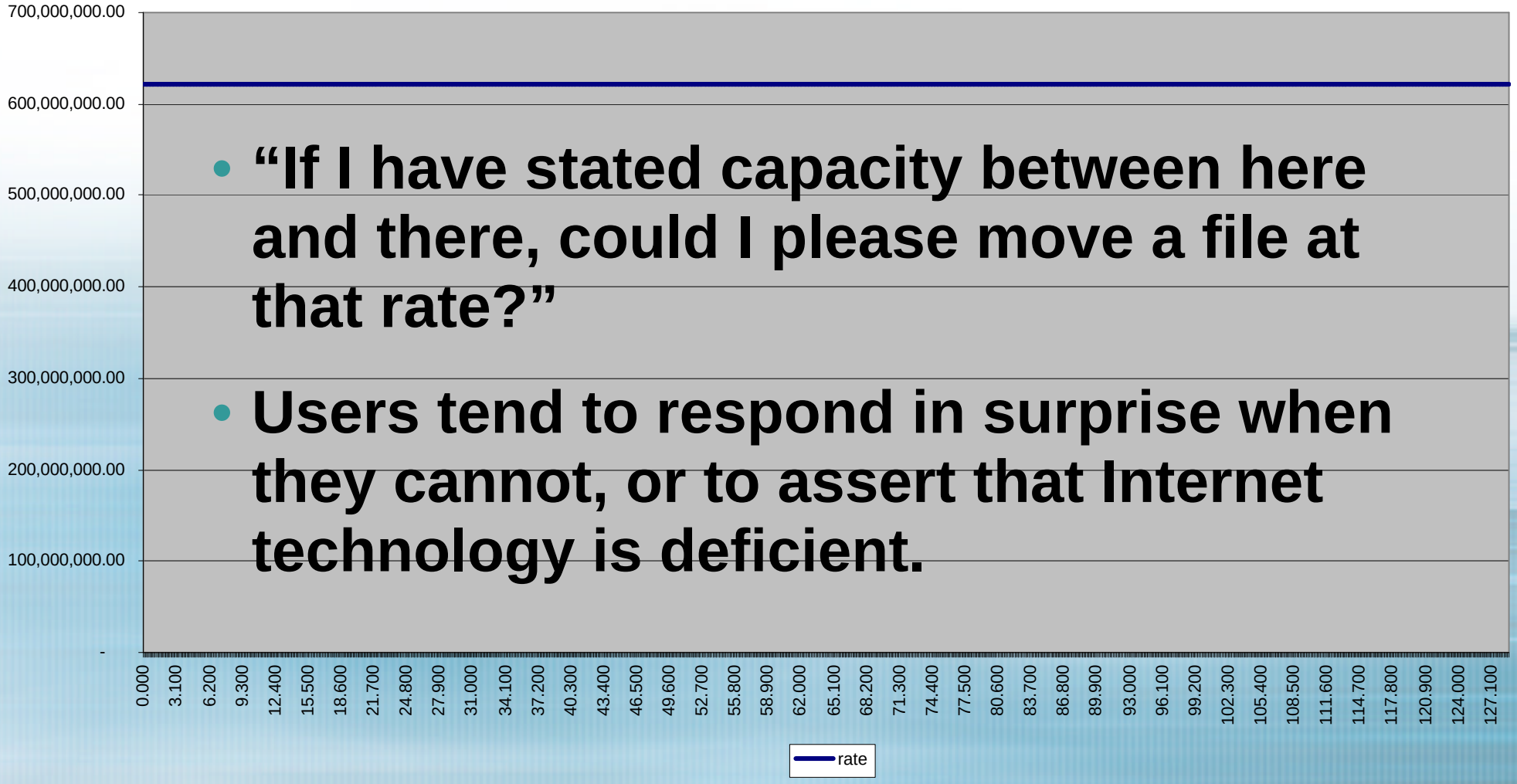
- **Drivers for current research agenda**
  - **Military requirements for predictability**
    - **Large delay*bandwidth and large error rates**
  - **Research requirements for large file transfer**
    - **Large delay*bandwidth**
  - **VPNs on the Internet**
    - **Maximizing servicability on the backbone**




© 2004 Cisco Systems, Inc. All rights reserved.

# What users would like…



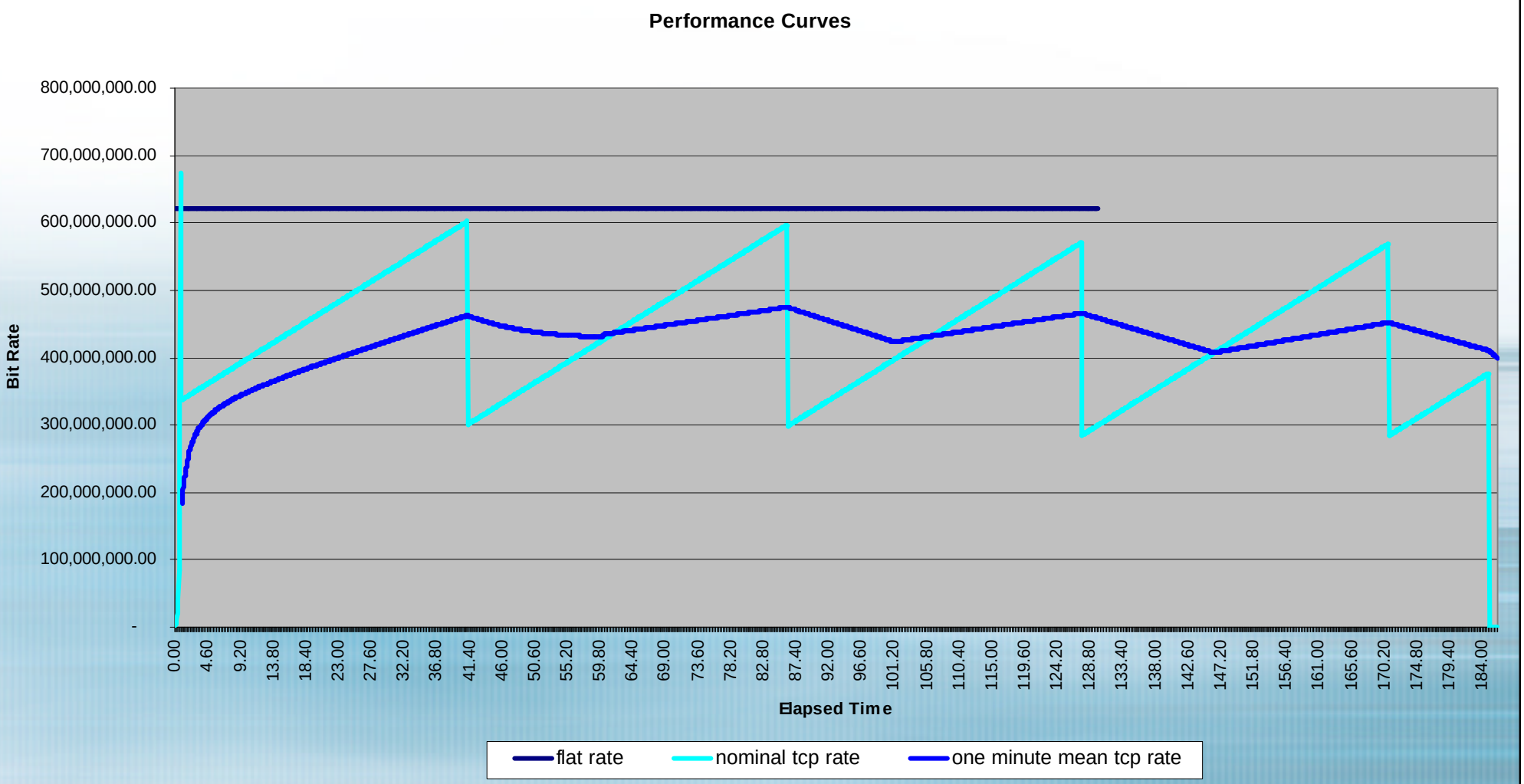



- **“If I have stated capacity between here and there, could I please move a file at that rate?”**
- **Users tend to respond in surprise when they cannot, or to assert that Internet technology is deficient.**


© 2004 Cisco Systems, Inc. All rights reserved.

# TCP or SCTP actual performance







© 2004 Cisco Systems, Inc. All rights reserved.

# Australian Defense solution



- **Simple File Transfer:**

  **Divide file into enumerated blocks**

  **Place blocks into transmission set**

  **While transmission set is not empty**

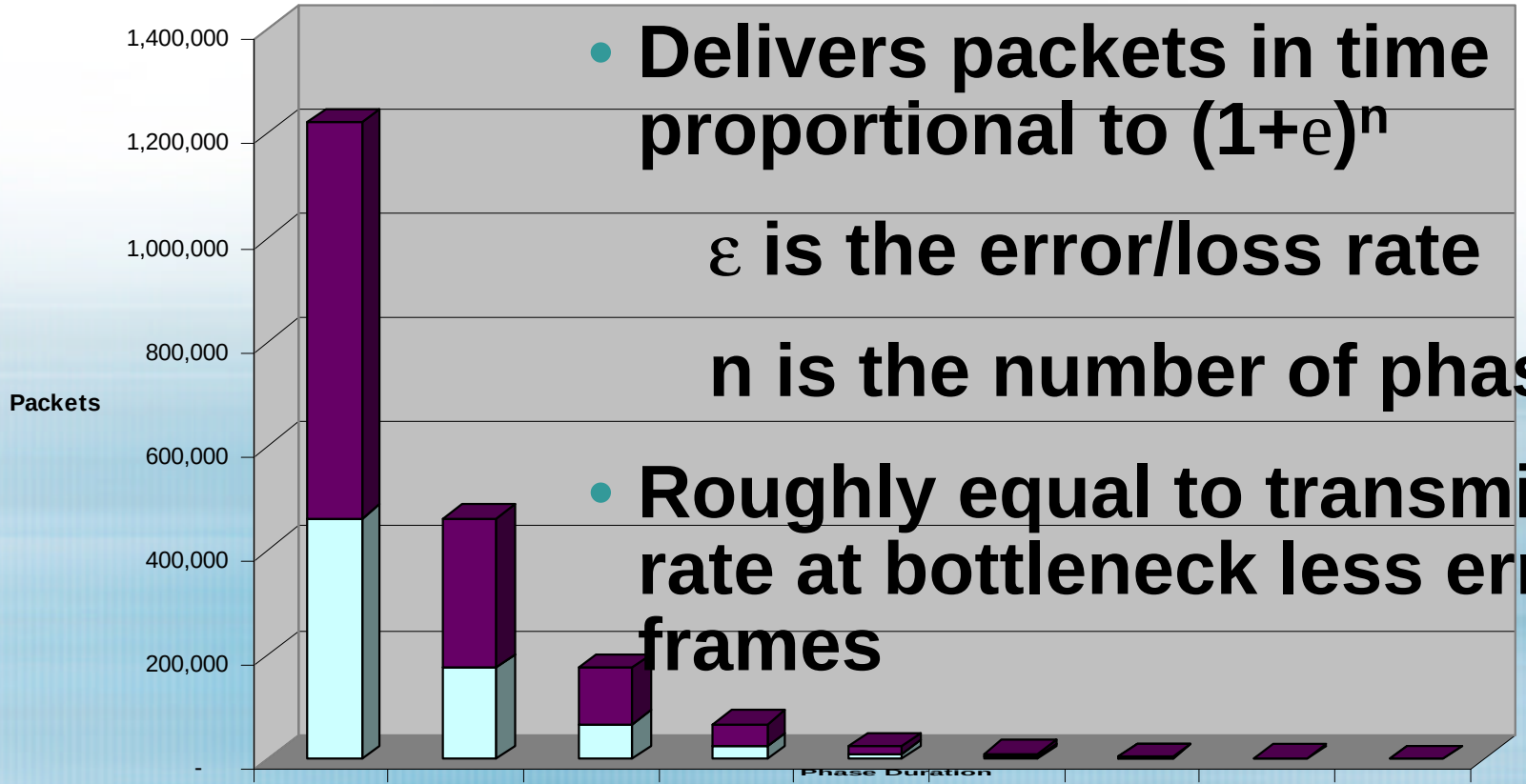
    **Send enumerated blocks to receiver**

  **Receiver acknowledgements remove blocks from transmission set on receipt at sender**


© 2004 Cisco Systems, Inc. All rights reserved.

# Behavior of that solution in time






- Delivers packets in time proportional to $(1+\varepsilon)^n$

  $\varepsilon$ is the error/loss rate

  n is the number of phases

- Roughly equal to transmission rate at bottleneck less errored frames


© 2004 Cisco Systems, Inc. All rights reserved.

# Calculating the maximum rate at which TCP or SCTP sends



- **RFC 2581 (Reno and NewReno) "standard" TCP or SCTP**

  **Heuristics adjust window against loss**

$$cwnd_{i+1} = \begin{cases} loss: & cwnd_i/2 \\ no\ loss: & cwnd_i + mss \end{cases}$$

$$rate = \frac{effective\ window}{round\ trip\ time}$$

- **CalTech FAST TCP or SCTP**

  **Heuristic adjusts window against varying delay**

$$cwnd_{i+1} = cwnd_i \frac{least\,RTT}{mean\,RTT} + mss$$


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues and approaches in TCP or SCTP performance: Measuring available capacity



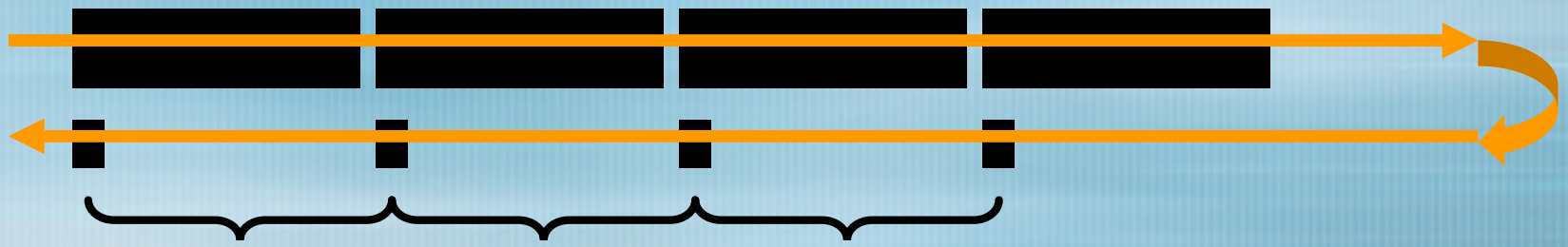
- Can we generate a filtered mean rate rather than playing with heuristics applying to the window?
- Ack pairs constitute packet pairs

$$mean\ rate = mean\left(\frac{amount\ acknowledged}{acknowledgement\ interval}\right)$$

- Treat this as an upper bound rate

  Maximum number of octets sent during relatively small interval shaping traffic to minimize queues

- How does rate-control management interact with "friendliness"?




Keshav, Packet-Pair Flow Control
http://www.cs.cornel.edu/skeshav/doc/94i2-17.ps


© 2004 Cisco Systems, Inc. All rights reserved.

# Upper bound rate



- **We can therefore calculate the rate at which to send**

$$mean\ available\ rate = moving\ average\left(\frac{amount\ acknowledged}{acknowledgement\ interval}\right)$$

- **And as a result calculate the window size to use instead of estimating it**

$$window\ upper\ bound = mean\ available\ rate * RTT$$


© 2004 Cisco Systems, Inc. All rights reserved.

# Issues and approaches in TCP or SCTP performance: Loss, and retransmission



- Use rate measurement for congestion avoidance
- Retransmit on loss
  - Detect loss by timeout: mean RTT + mean DRTT
- SACK with windows O(N*delay*bandwidth) more effective
  - Redefine “effective window” as mean RTT * mean available rate, and keep that amount outstanding


© 2004 Cisco Systems, Inc. All rights reserved.

# We should therefore…



- **Be able to implement TCP or SCTP in a manner that**
  - **Delivers traffic at a predictable rate**
  - **Is compatible with current generation TCPs and SCTPs**
  - **Interoperates well with those TCP and SCTPs**


© 2004 Cisco Systems, Inc. All rights reserved.

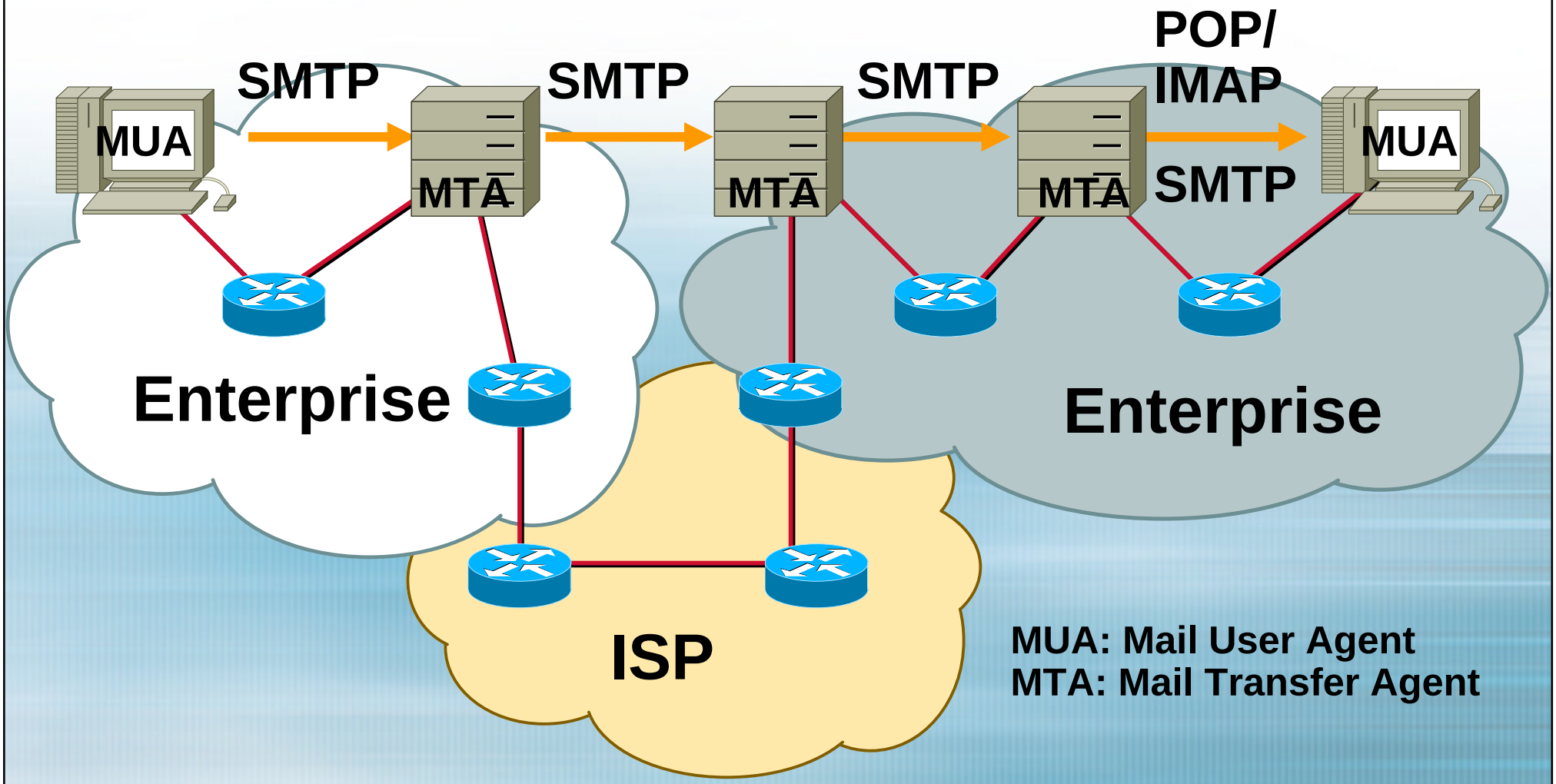
# Issues in the application: Electronic mail as an example

- Political
- Religion
- **Application**
- Presentation
- Session
- Transport
- Network
- Link
- Physical


© 2004, Cisco Systems, Inc. All rights reserved.

# Electronic mail as a privileged service







© 2004 Cisco Systems, Inc. All rights reserved.

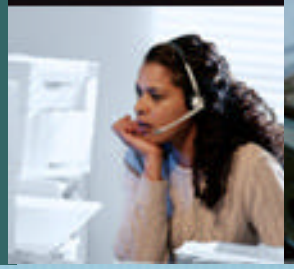
# Issues to consider



- In MTA->final MUA
  - Need a “push” protocol, not a “pull” protocol
  - Need notification to user to read priority message
- In initial MUA and each MTA
  - Need recognition of priority messages
  - Changes in policy:
    - Do not use exponential backoff on failure
    - Quick send/resend
    - Quick “bounce” on failure
- Is this for everyone?
  - Probably just for authenticated users: SMTPAUTH
  - Probably just when requested: SMTP Priority


© 2004 Cisco Systems, Inc. All rights reserved.

# Building preferential services

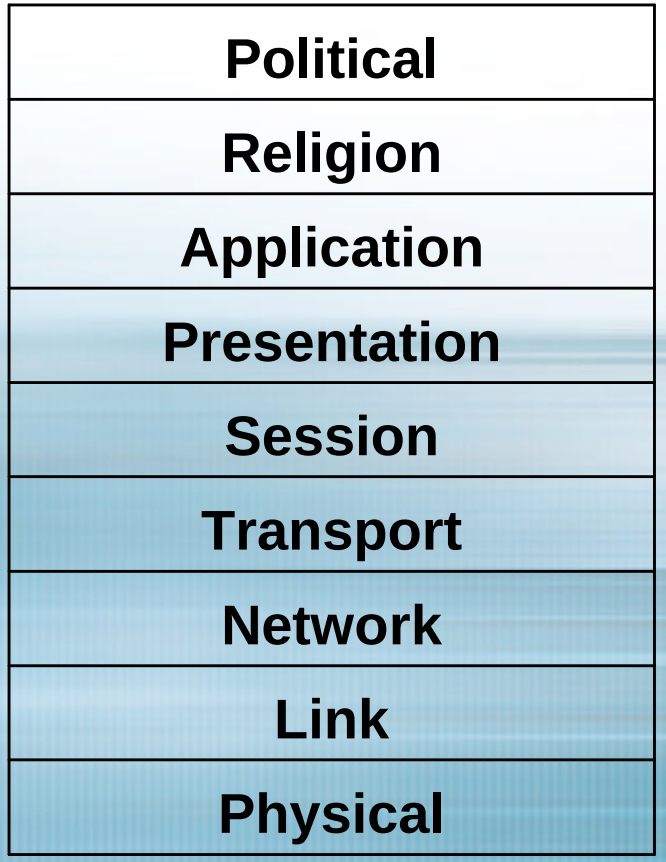


© 2004, Cisco Systems, Inc. All rights reserved.

# Key things to note



- **Building a preferential service requires technology in many layers**
  - **Application**
    - **Identification and implementation issues in services**
    - **Choice of protocols and features**
  - **Session**
    - **Strong authentication of individual users or functional groups of users**
  - **Transport**
    - **Appropriate procedures in end system implementations**
  - **Network**
    - **Application of bandwidth to the right sessions according to an authenticated policy**



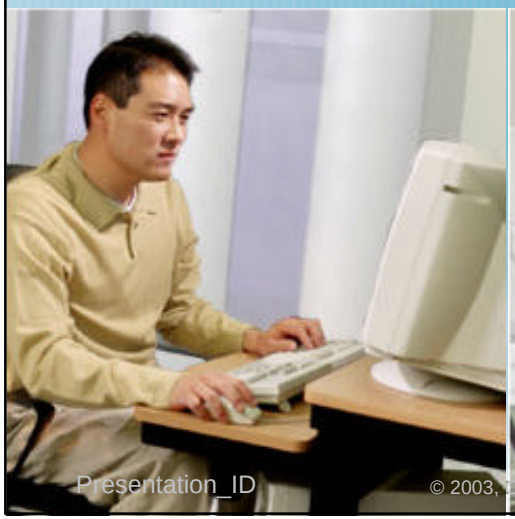
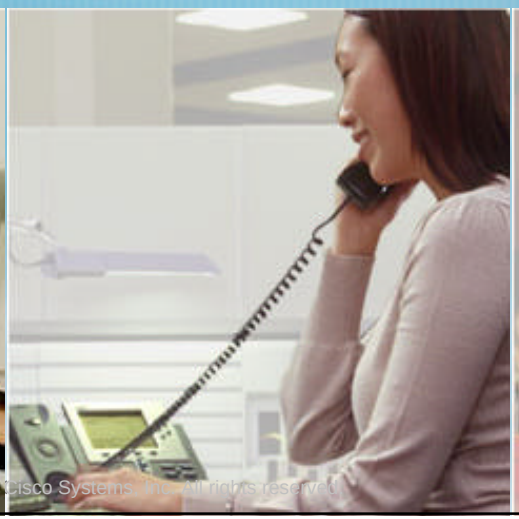


© 2004 Cisco Systems, Inc. All rights reserved.

- **“Everything is very simple in war, but the simplest thing is difficult. These difficulties accumulate and produce a friction, which no man can imagine exactly who has not seen war.**
- **So in war, through the influence of an ‘infinity of petty circumstances’ which cannot properly be described on paper, things disappoint us and we fall short of the mark.”**

**Carl von Clausewitz**

**On War, Chapter VII, “Friction in War”**

© 2003, Cisco Systems, Inc. All rights reserved.

# Preferential services: Not hard, not trivial



- Technology is mostly in place, but not in use for this purpose
  - Obvious next step: harness it to the purpose
  - Many specific recommendations
- We must approach the problem with our eyes open
  - Careful service definition at each layer
  - Consideration of the specific needs of users one wishes to give preference to


© 2004 Cisco Systems, Inc. All rights reserved.

# Advanced Preferential IP and Telephony Services in the Internet

**Fred Baker**

**Cisco Fellow**


© 2004 Cisco Systems, Inc. All rights reserved.