IP Quality of Service: Theory and best practices

Vikrant S. Kaulgud

Why are we here?

- Understand need for Quality of Service.
- Explore Internet QoS architectures.
- Check QoS best practices.
  - Be vendor neutral, you can map the practices to products anytime!

Share our experiences. Participation is the key!

Session I: QoS Essentials.

SETLabs

Do we need QoS?

- Why do networks exist?
- Is customer satisfaction on your list?
- Is resource utilization on your list?

Quality of Service?

- ISO definition
  - Quality of Service (QoS) is a "set of qualities related to the collective behavior of one or more objects.“

» Source: ISO 95 QoS Framework, ISO/IEC/JTC1/SC21/WG1 N9680.

Quality of Service?

- An operational perspective
  - It is the ability of the network to service an application effectively, without affecting its performance and functionality.
  - Satisfactory user experience.

Applications need different handling

	Voice	FTP	ERP and Mission-Critical
Bandwidth	Low to Moderate	Moderate to High	Low
Random Drop Sensitive	Low	High	Moderate To High
Delay Sensitive	High	Low	Low to Moderate
Jitter Sensitive	High	Low	Moderate

QoS – An operational perspective

- What affects an application performance or functionality?
  - Applications themselves?
  - Operational environment?
  - Servers – hardware & operating system?
  - Internetwork topology?
  - Links?
  - Interneworking components?

It depends on all !!!

QoS – A business perspective

- Setting up a network costs *$$$*!
- Each network element is a resource
  - Tangible: routers, switches, links, servers …
  - Intangible: packets, frames …
- What affects the intangible assets?
  - Packet loss, delay, jitter …
- Does this impact your business?

QoS – A network perspective

- We will focus on the network aspects of QoS.
- QoS is actually managing network's intangible assets and factors affecting them!

Have we come a full cycle?

A brief look at congestion

- Root cause for congestion is *(dynamic)* lack of bandwidth.
  - Demand for bandwidth is greater than capacity.
  - Sudden surge in demand.
  - Unexpected traffic flowing into the links due to routing. .
  …

Everyday experiences of congestion?

How does congestion affect intangibles?

- Delay
  - Packets start queuing up at the router interfaces.
  - Take more time to exit the router.
- Packet loss
  - Queue buffers exhaust, routers start dropping packets!
- Jitter
  - Packets in the same flow routed to links having variable delay.

Congestion scenarios: Speed Mismatch

- LAN – WAN interconnect.
- Interconnection of high bandwidth LAN links to low bandwidth links.
- Problems:
  - Traffic from high bandwidth links gets choked on entering low bandwidth links.
  - Buffer exhaustion on devices.

Congestion scenarios: Aggregation

- Traffic from multiple links aggregates into a single link of lesser bandwidth than the aggregate.

- Problems
  - Similar to speed mismatch. Here aggregation is the reason for the perceived speed mismatch.
  - Aggregate link is choked.
  - Buffer exhaustion on devices.

Congestion scenarios: Transit networks

- Traffic between two core networks transits through the transit network.

- Problems
  - Transit network acts as the choke point.
  - Poor performance of the core networks.

Flash points in the network

- Congestion causes the flash points!

Will such flash points be static or dynamic?

They will be dynamic! No one can predict with accuracy where congestion will next occur!

Managing flash points – Provision QoS

- Throw bandwidth at the problem!
- Manage the intangibles!

Make your choice!

Throwing bandwidth at the problem?

- Easiest way is to over provision the network.
- Over-provisioning is static.
  - Bandwidth cannot be carried to a new flash point in the network.
  - Over provisioned section *may not* face *congestion*!
- Over-provisioning does not always make business sense!

Managing rather than over provisioning.

- Treat network resources as precious!
- Ensure fair usage of resources by all.
- But, provide for priority access to resource for some.

How will it help in providing quality of service?

Benefits of *managed* QoS

- Enterprise networks
  - Priority service to mission critical application traffic.
  - Non critical traffic does not burden precious bandwidth.
  - Helps in mitigating effects of denial of service (DoS) attacks.

- Service Providers
  - IP QoS is a key cornerstone.
  - Application level SLAs can be built and offered as a premium service ($$$!).

Questions

Session 2: QoS architectures

SETLabs

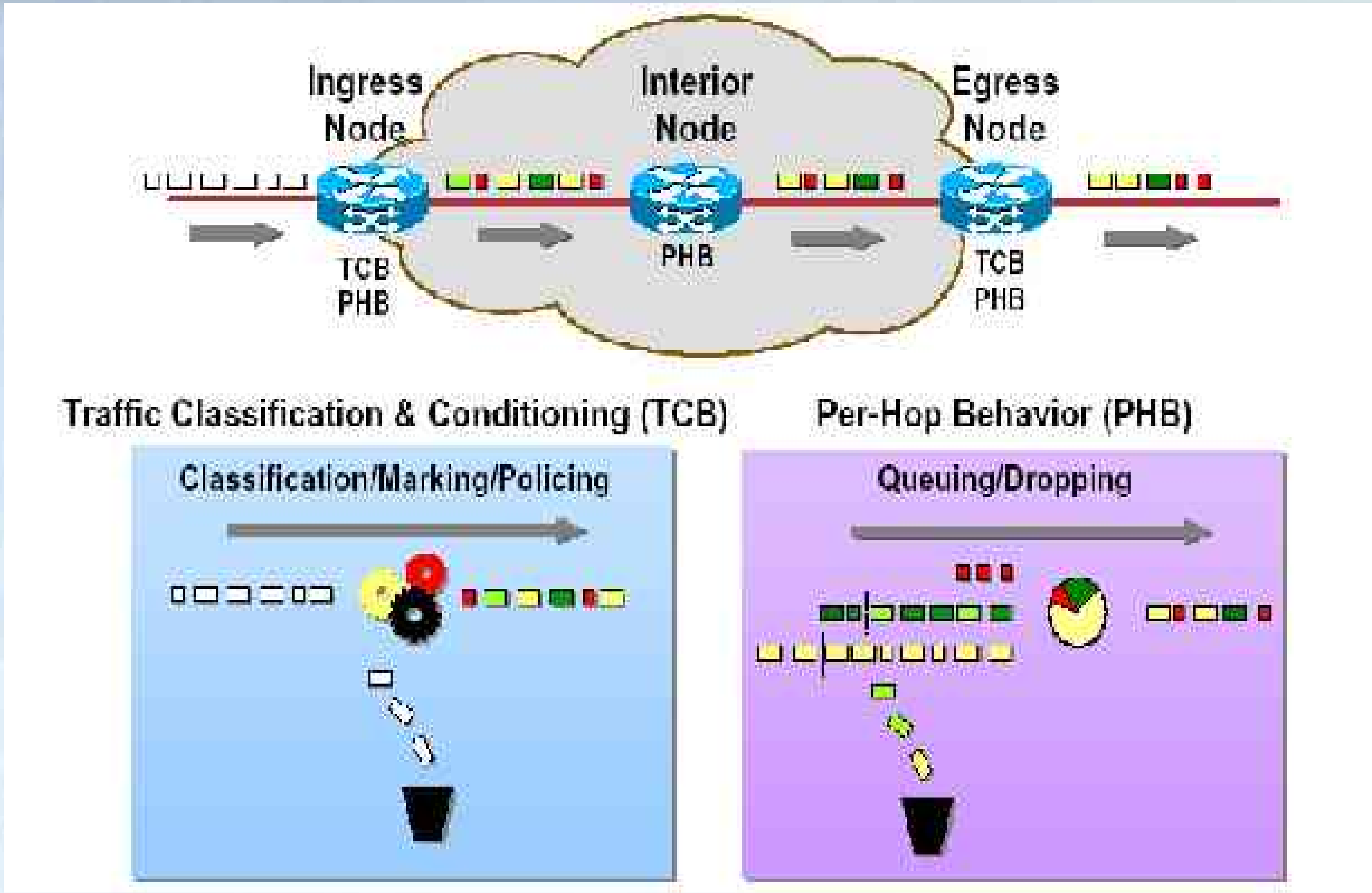
Standards – An evolution

- 1981 – RFC 791
  - Best Effort Service.
- 1997
  - Integrated Services (IntServ)
- 1998 – RFCs 2474, 2475, 2597, 2598.
  - Differentiated Services (DiffServ)
- Now
  - DiffServ-Aware Traffic Engineering (DS-TE).

The DiffServ Architecture

SETLabs

DiffServ Architecture



DiffServ: Terminology

- DS Domain (e.g. ISP, intranet).
- DS Boundary Node (Egress & Ingress).
- DS Interior Node.

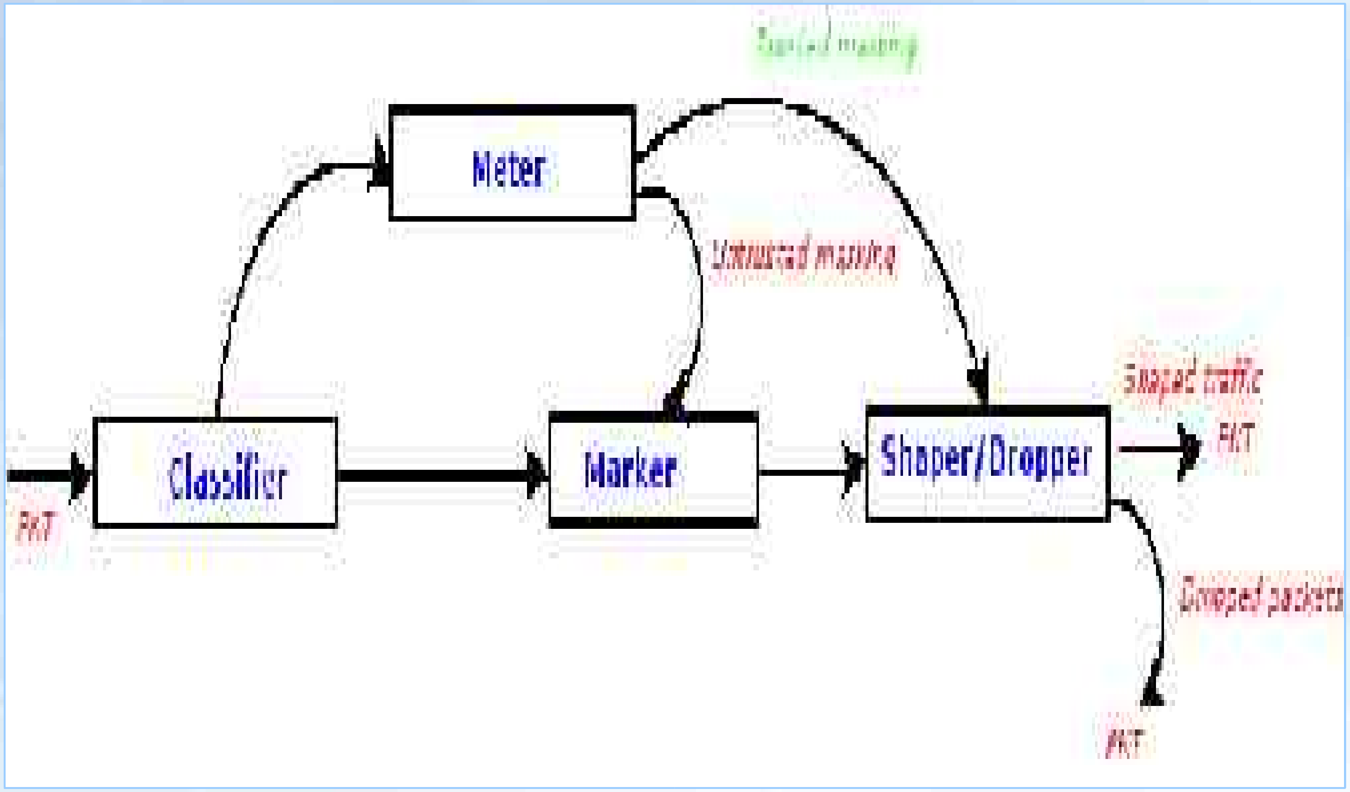
- Per Hop Behavior (PHB).
- DS Codepoint (DSCP).
- DS Behavior Aggregate (BA).

How does DiffServ work? - 1

- Simple idea
  - Offer various service levels e.g. gold, silver, bronze …
  - Insert expected service level in the packet as a “code point”.
  - DiffServ refers to the service level as a “class”.
  - Each router participates in providing a packet its class of service. This is called as “Per Hop Behaviour (PHB)”.

RFC 2474 defines service as ‘some significant characteristics of packet transmission in one direction across a set of one or paths in a network’.

How does DiffServ work? - 2



Classification of packets

- What parameters can be used for classification?
  - Source/Destination IP addresses, Port numbers.
  - Incoming/Outgoing interface.
  - IP precedence values, DSCP value.
  - …

- Two types of classification
  - BA classifier: based on behaviour aggregate.
  - MF classifier: based on multiple fields in the packet header or even the payload.

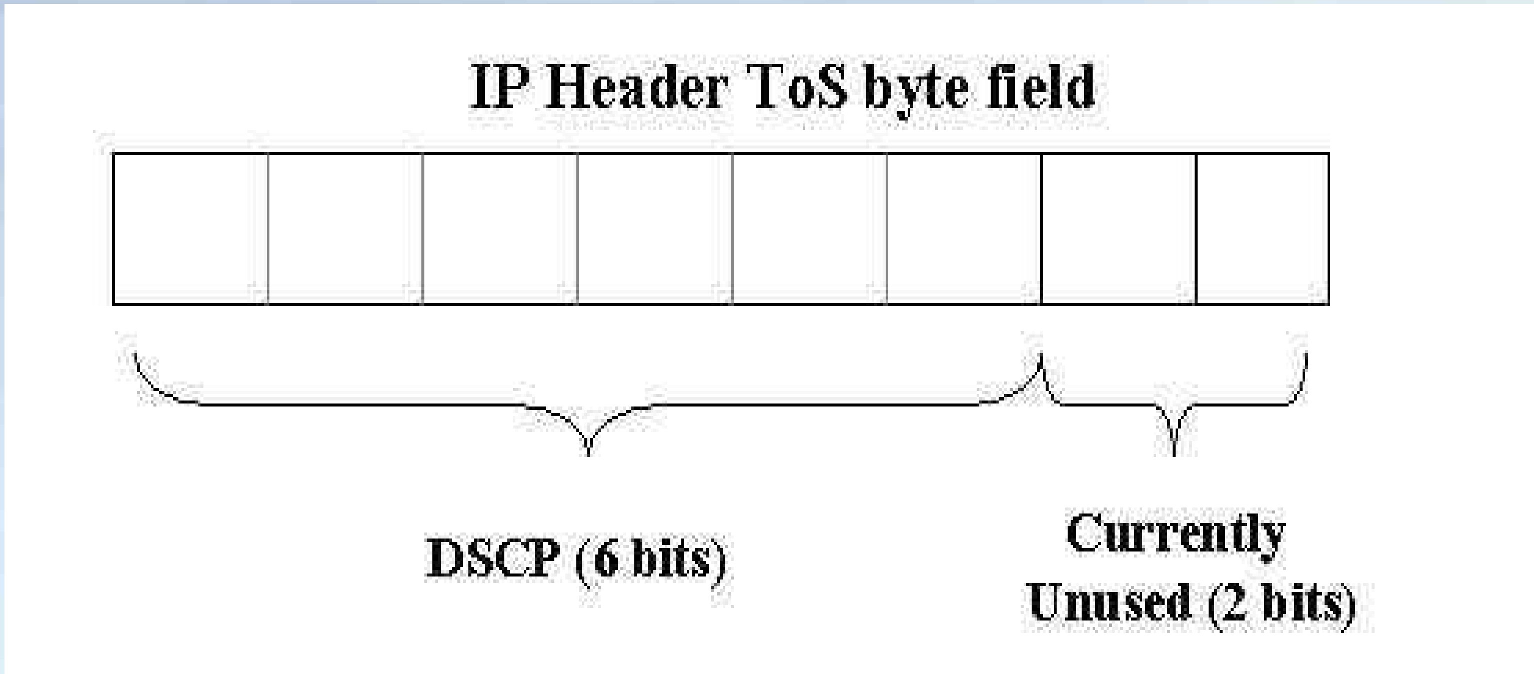
Marking

- Adding service level identification to the frames or packets.

- Marking can be done at L2 or L3
  - IP TOS field.
  - DSCP field.
  - MPLS EXP bits.
  - ATM CLP bit.
  - Frame relay DE bit.
  - IEEE 802.1/q bits.

Marking: DSCP way

SETLabs

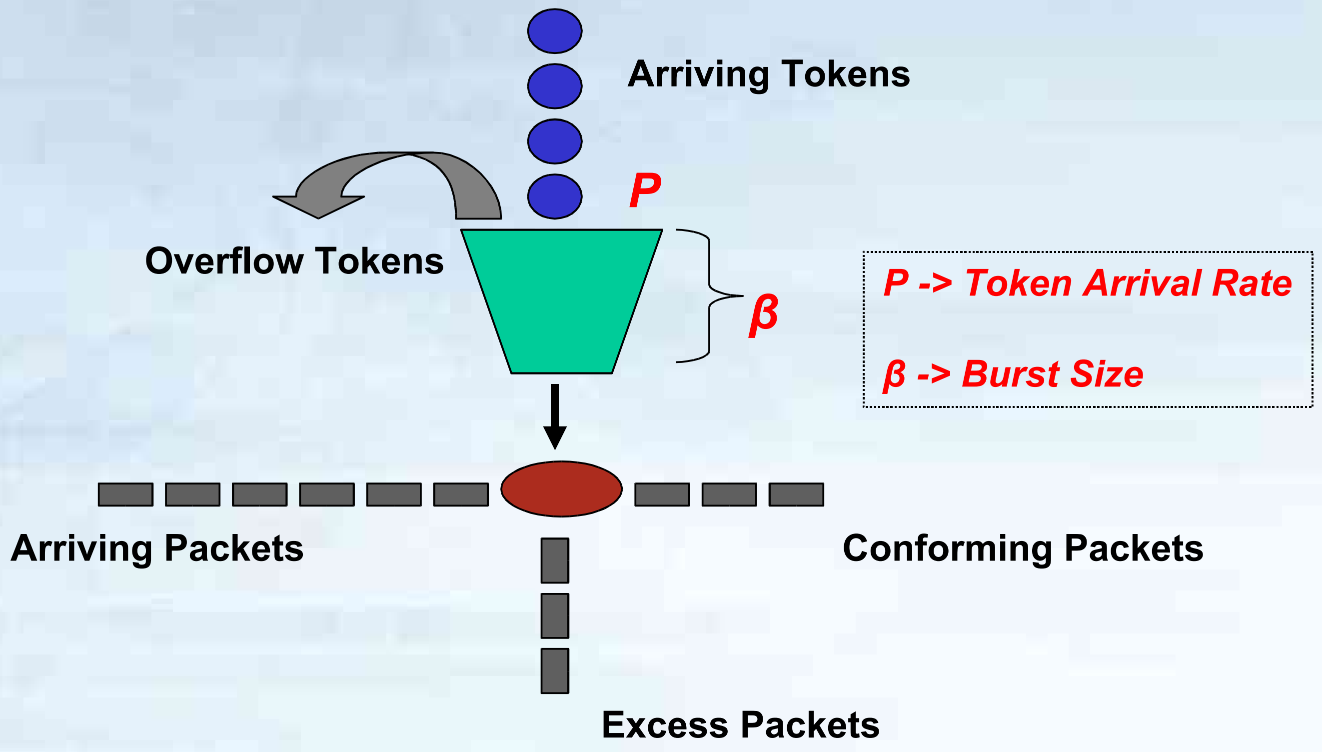


- IP ToS field redefined in DS standard.
- 6 bits used for codepoints (i.e. marking).

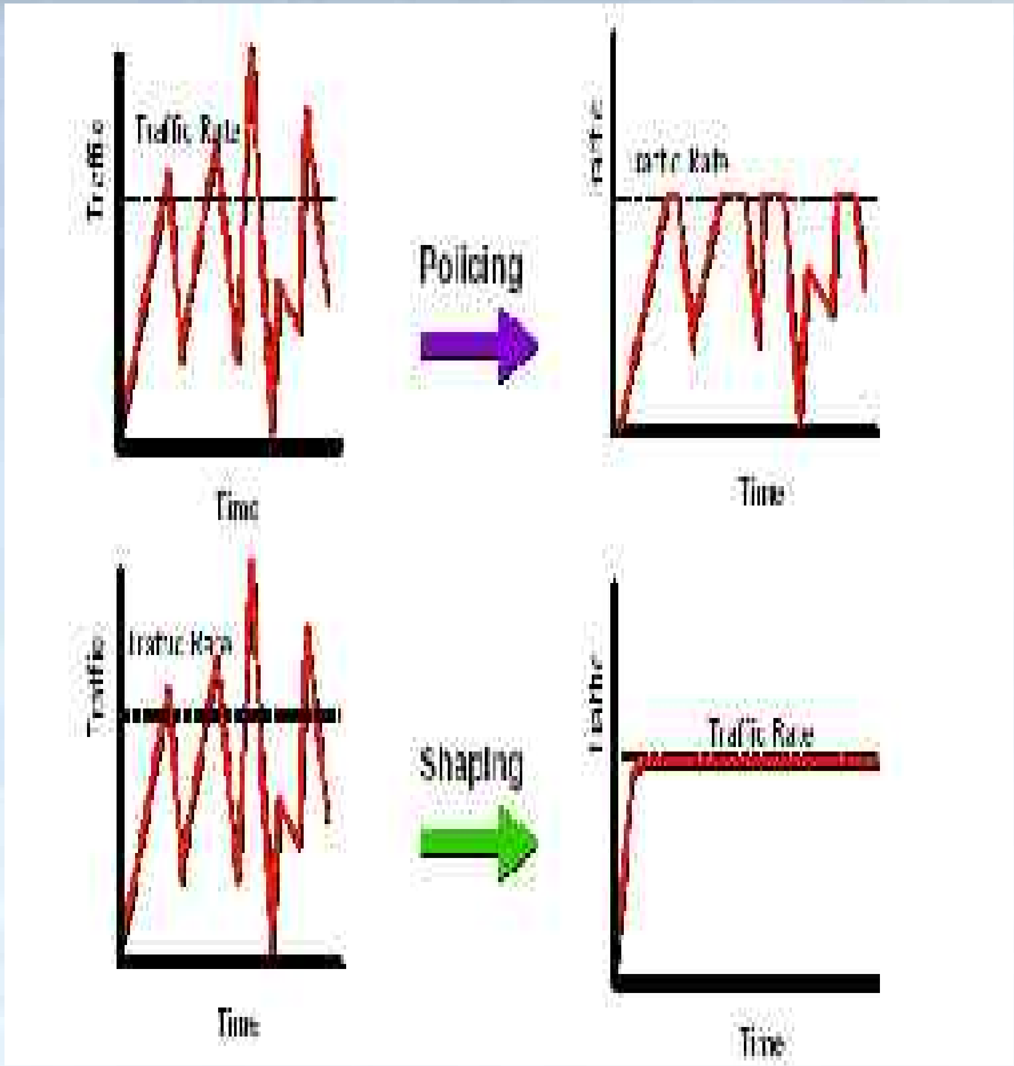
Infosys

Metering

- Remember it’s an optional service.
- Typically uses a Token Bucket (TB).



Policing or Shaping



- This is used with metering.
- Policing
  - Drop non-conformant packets.
  - *Re-classify non-conformant* packets for the next hop to discard them.
  - Aggressive.
- Shaping
  - Buffer and schedule packet egress as per policy.
  - Has an effect of smoothening traffic flow.
  - Typically used for speed-mismatch scenarios.

Queue and/or drop

- Queuing
  - Buffer packets when an interface (link) is congested.
  - Schedule egress of packets out of the buffer using a scheduling algorithm (FIFO, CBQ, WFQ …).
- Dropping
  - Drop packets that cannot be buffered or are non-conformant.
  - Dropping can happen at the edge or the core.

Which of the two is better?

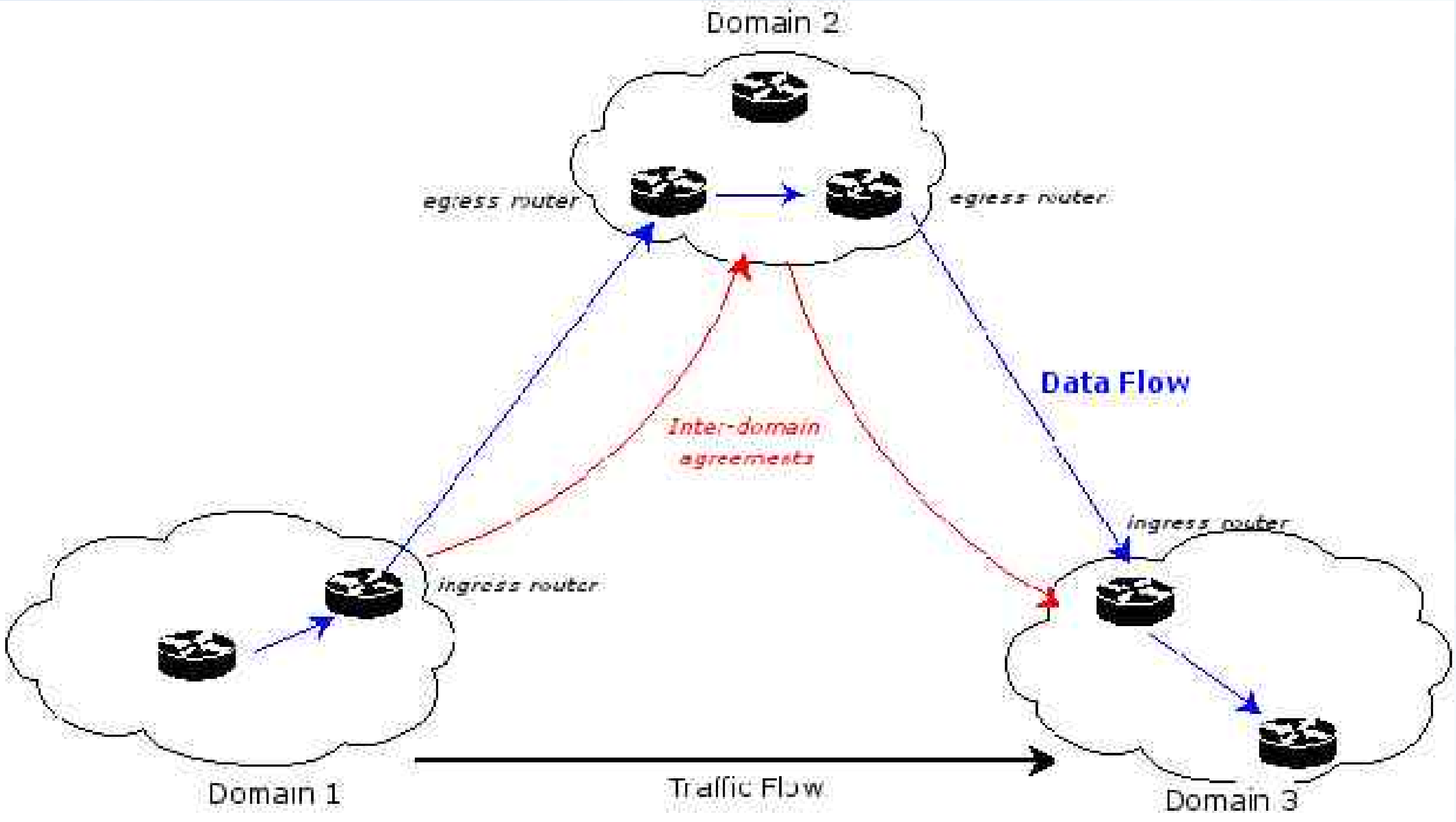
Dropping works believing that sources will back-off!

Common PHBs

No.	PHB	Behaviour
1	**EF (Expedited forwarding)**	**Very low delay, low jitter and assured bandwidth**
2	**AF (Assured Forwarding)**	**Assured amount of bandwidth** **4 IETF defined sub classes**
3	**Default**	**Best effort**
4	**CS (Class Selector)**	**Backwards compatible with IP precedence values.** **Used for Forwarding Probability (FP)**

Can you compare this with a mail service?

Interdomain DiffServ operation



Interdomain agreements have to be brokered to ensure end-to-end QoS

Interdomain DiffServ operation - 2

- Bandwidth Broker (BB) typically used for interdomain negotiation.

- BBs use SLAs and TCAs for negotiation
  - **Service Level Agreement (SLA):** A set of parameters and their values which together define the service offered to a traffic stream by a DS domain.

  - **Traffic Conditioning Agreement (TCA):** A set of parameters and their values which together specify a set of classifier rules and traffic profiles.

Interdomain DiffServ operation - 3

- **Assumption:** Needs of domain 1 towards domain 3 are satisfied by a 64kb/s flow of premium traffic.

- Steps in brief:
  - BB1 negotiates a SLA with BB2.
  - BB2 admits the SLA provided resources are available.
  - BB1 then negotiates the TCA with BB2.
  - Negotiated TCA is used to configure appropriate routers.

  - BB2 may negotiate with BB3 for premium services if required prior to admitting SLA request from BB1.

The IntServ Architecture

SETLabs

Why IntServ? - 1

- Analogy of telephone call.
  - Caller requests for resources from the telco for setting a session with receiver.
  - Telco admits or rejects the call depending on available resources.
  - Once admitted, allocated resources *remain* allocated till the call is terminated by either end-point.

Try explaining this using the DiffServ concept?

Why IntServ? - 2

- Certain applications expect uniform service level for the entire duration of the call/session/flow.

- DiffServ does not have a concept of a “call” (or session / flow)
  – DiffServ is incapable of handling flows.

- Other limitations include lack of admission control.

IntServ Basics - 1

- Defines two service classes
  - Controlled Load service
    - No fine grained guarantees provided.
    - Bandwidth reservation necessary. (limited)
    - Additional packets receive best effort service.
  - Guaranteed Service class
    - Provides firm bounds on throughput and deterministic upper bounds on packet delay.
    - Designed for intolerant real time applications (CBR, rt-VBR, interactive multimedia)

IntServ Basics - 2

- Applications need to know the characteristics of the traffic before hand.
- Hosts "signal" the network to request for resources to meet traffic requirements.
- The network performs admission control and either accepts or rejects the resource reservation request.

IntServ

- IntServ provides QoS guarantees to individual application sessions or flows.

- Three components
  - Sender specification (T_{spec}).
  - Receiver specification (R_{spec}).
  - Signaling by sender and receiver to network components.

IntServ signaling

- Key component.

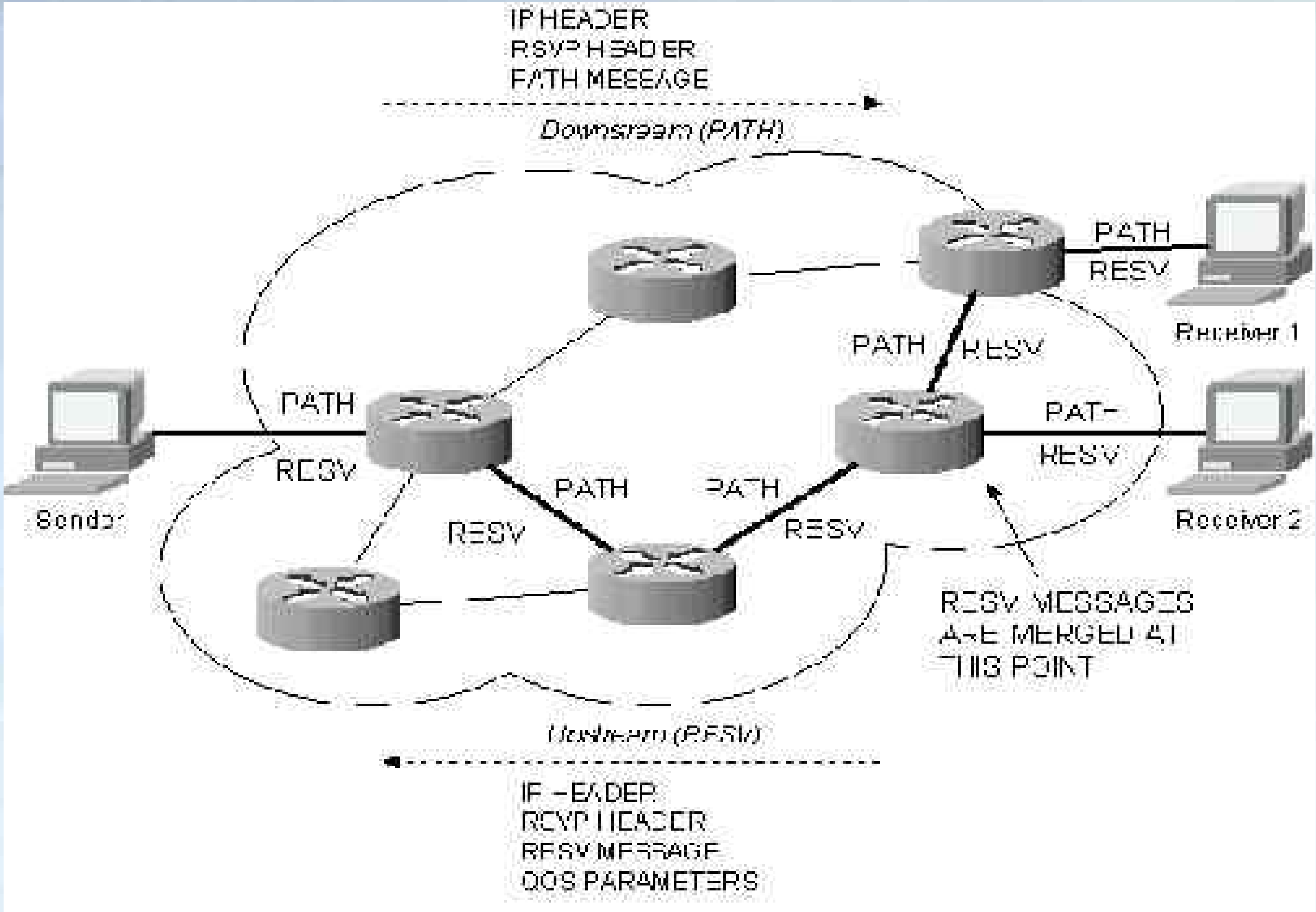
- Done using ReSerVation Protocol (RSVP)
  - Signaling resources for a call.
  - Maintaining and tearing down resources during and after the call respectively.

RSVP signaling

- RSVP signaling uses following messages
  - PATH
  - RESV
  - PATH TEAR
  - RESV TEAR

- PATH and RESV messages include T_{spec} and R_{spec} respectively.

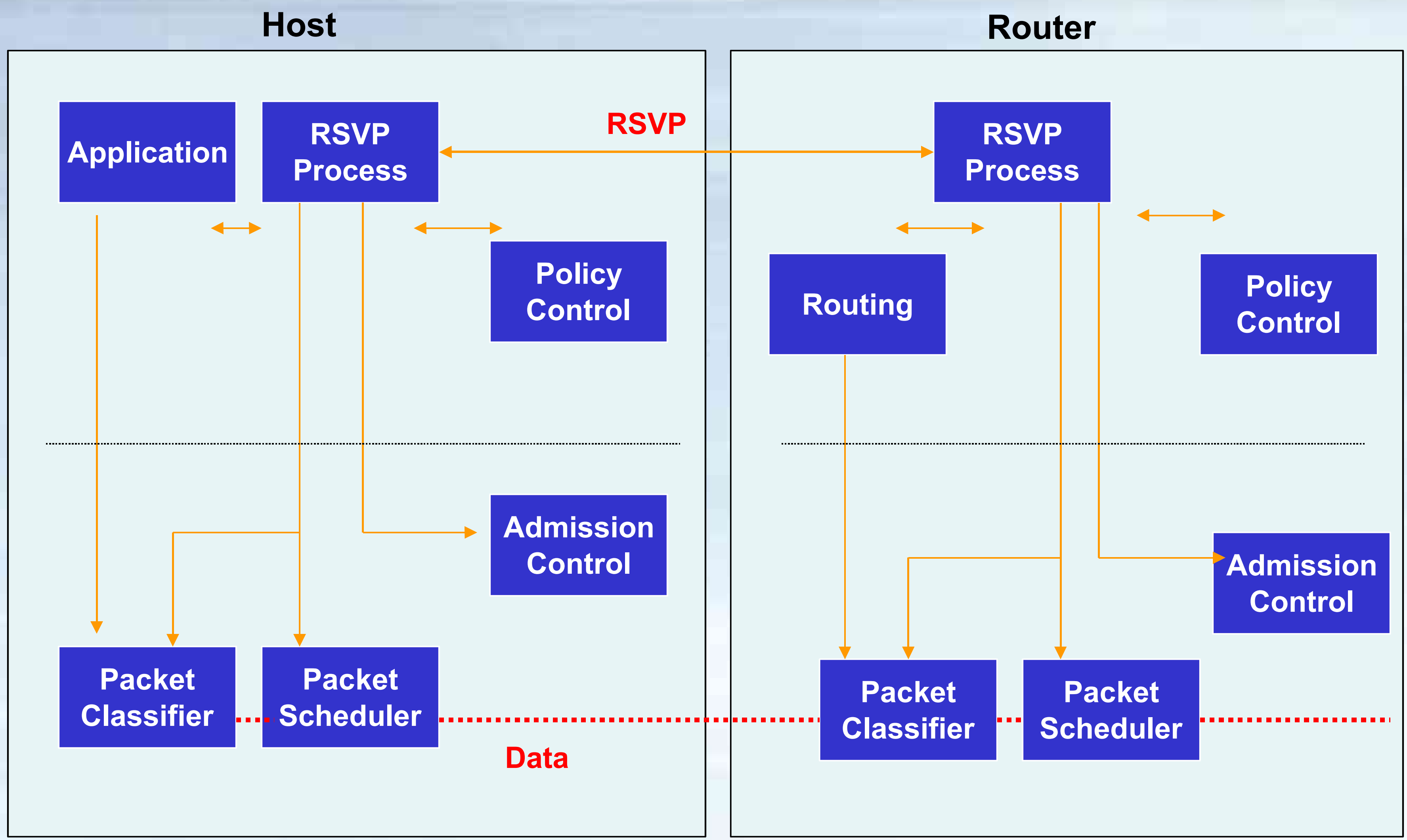
RSVP signaling



RSVP: Effect on routers

- RSVP signaling causes each router in the path to allocate required resources for the flow.
  - This state information has to be maintained for the during of the flow.
  - When the flow ends, the state information is removed and the resources de-allocated.

RSVP: Logical components



Benefits and drawbacks

- Benefits
  - Policy based deployment simple using COPS.
  - Largely automatic operation due to RSVP.
  - Flow level granularity.

- Drawbacks
  - Signaling overheads in a global network are high.
  - Operational overheads are large for large number of flows.

IntServ vs. DiffServ

Parameter	IntServ	DiffServ
Coordination for service differentiation	End to end	Per hop
Scope of service differentiation	Unicast or multicast path	Anywhere in the network or in specific paths.
Scalability	Limited by number of flows	Limited by number of classes of service
Network accounting	Based on flow characteristics and QoS requirement	Based on class usage
Network Management	Similar to circuit switching	Similar to IP networks
Inter domain deployment	Multilateral agreements	Bilateral agreements

Choosing between the two

- Choose IntServ for:
  - Guaranteed bandwidth, end-to-end QoS.
- Choose DiffServ for:
  - High scalability
- But don't we require all this!
  - Can't we use the best of both worlds?
  - IntServ is obsolete.

Deployment options

- Mostly router based
  - E.g. Cisco platforms like 26xx, 36xx, 72xx …
  - Linux based solutions for DiffServ (exciting option for low-cost deployments, experimental setups).

DS Aware MPLS Traffic Engineering

MPLS : A micro introduction!

- Specifies mechanisms to manage traffic flows between different hardware, machines, or applications.
- Maps IP addresses to simple, fixed-length labels
- Interfaces to existing routing protocols such as RSVP, OSPF etc.
- Supports the IP, ATM, and frame-relay Layer-2 protocols

MPLS operation: Steps

- Label creation and distribution
- Table creation at each router
- Label-switched path creation
- Label insertion/table lookup
- Packet forwarding

MPLS benefits

- Improves packet-forwarding performance in the network.
- Supports QoS and CoS for service differentiation.
- Supports network scalability.
- Helps builds interoperable networks.

DS and MPLS: Aggregation at the edge

- DS Aggregation
  - Packets in the same flow are marked with a common DSCP.
- MPLS aggregation
  - Packets in the same flow are marked with a common Forwarding Equivalent Class (FEC) in the MPLS label/

DS and MPLS: Processing at the core

- DS core processing
  - PHB (dropping & queuing) based on the DSCPs.
- MPLS core processing
  - Packets are processed in the core based on labels.

Power of two!

- DiffServ aware Traffic Engineering relies on both DS and MPLS for effective operation.
- Problems:
  - Make MPLS aware of the DiffServ DSCP value.
- Solution: Use the EXP field in the MPLS header
  - E-LSP
    - Queue” inferred from Label and EXP field
    - “Drop priority” inferred from label and EXP field

Traffic engineering

- Process that enhances overall network utilization by attempting to create a uniform or differentiated distribution of traffic throughout the network.

The need for TE

- IP routing (Destination address based best /shortest path selection)
  - Over utilization of certain paths while others are under utilized.

- Basic traffic engineering
  - Find and set up best path to a destination with certain characteristic.

QoS and Routing

- IP QoS is “routing-unaware”.
- IP QoS focuses on resource allocation, while routing focuses on path selection.
- Constraint-based routing
  - Select path that satisfies resource constraints, e.g. bandwidth

MPLS traffic engineering - 1

- MPLS provides constraint-based routing.

- Uses a L3 source routing approach.
  - Ingress routers set up path across the network using Forwarding Equivalent Class for resource allocation.
  - Such traffic engineered Label Switched Paths are called as “traffic engineering tunnels”
  - The LSPs are created independently, specifying different paths that are based on user-defined policies

MPLS Traffic Engineering - 2

- Complex relationship between MPLS-TE components.
  - IGP for advertising link capacity and other information.
  - Constraint based Routing selects links that satisfy the constraint specified for the traffic flow.
  - 'RSVP' used for admission control.

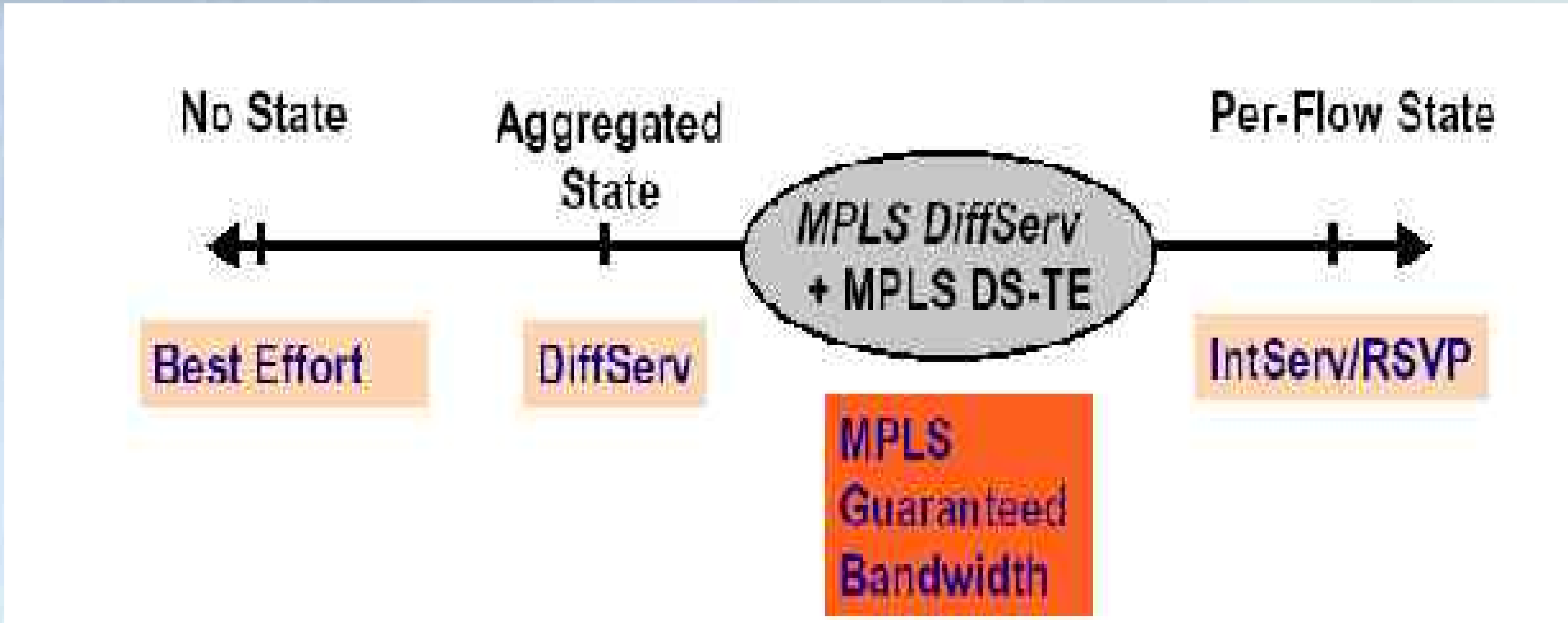
  - LSPs used for forwarding.

MPLS TE and QoS

- MPLS TE supports aggregate behaviour.
  - Does not provide granularity to a DS class level.

- Tight constraints can be provided if:
  - Constraint based routing is provided per class.
  - Admission control is provided per class.

Guaranteed bandwidth services

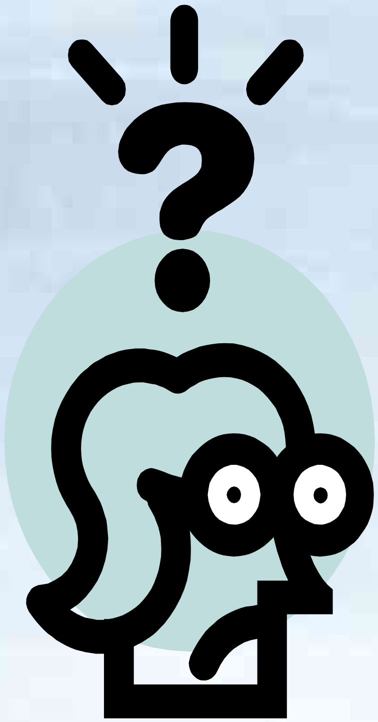


- How are we benefited?
  - Efficiency of DS.
  - Admission control as in IntServ.
  - Guaranteed bandwidth.

Summary

- Varying QoS architectures.
- DiffServ is the dominant candidate for a global and scalable deployment.
- DiffServ with MPLS and traffic engineering is a potent combination.

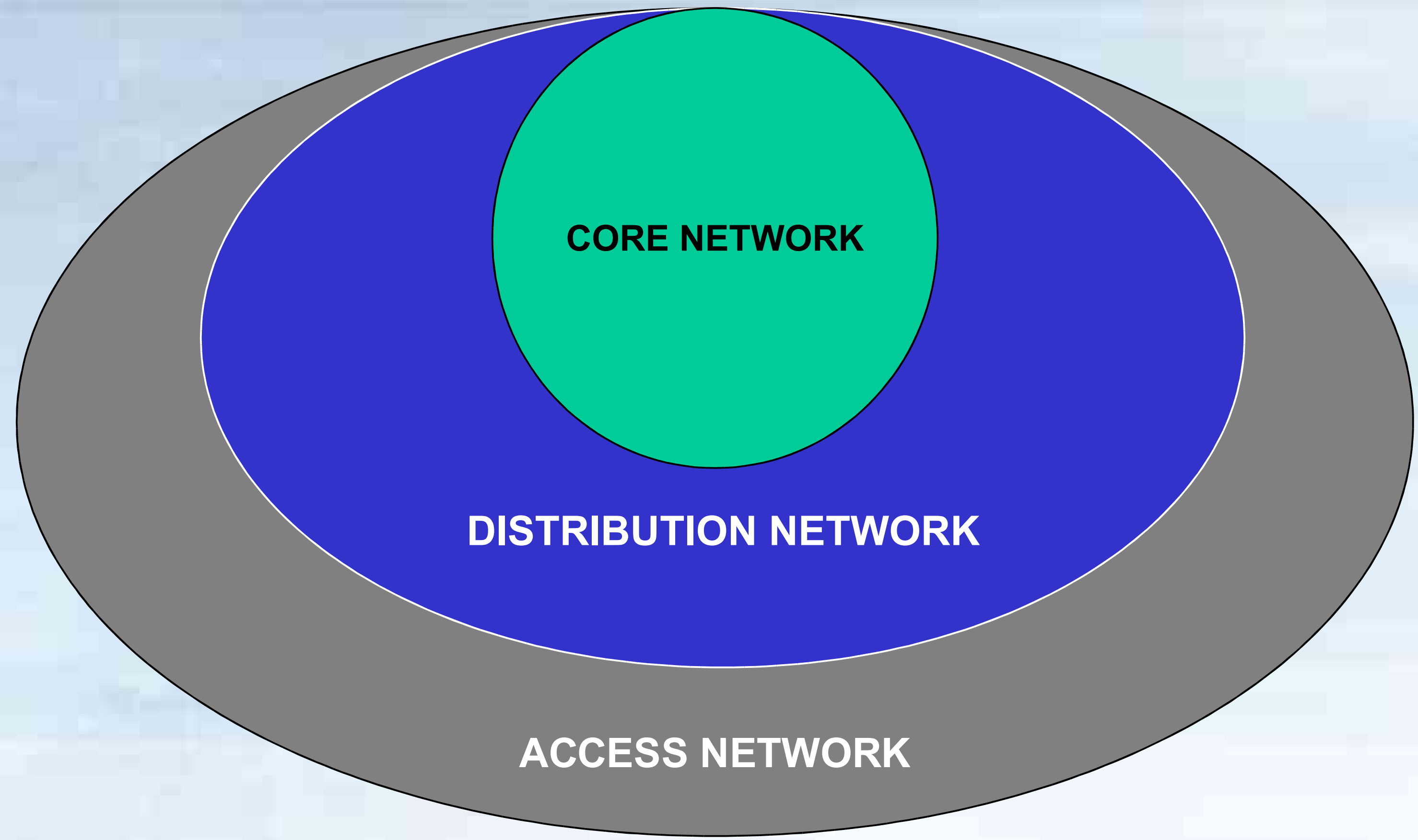
Questions



Session 3: Deploying DiffServv –
Best Practices

SETLabs

Network architecture – 1



Network architecture - 2

No.	Layer	Characteristic
1	**Core**	▪ **High-speed switching backbone** ▪ **Designed to switch packets as fast as possible**
2	**Distribution**	▪**Address or area aggregation** ▪**Departmental or workgroup access** ▪**Broadcast/multicast domain definition** ▪**Virtual LAN (VLAN) routing** ▪**Any media transitions that need to occur** ▪**Security**
3	**Access**	▪**Shared bandwidth** ▪**Switched bandwidth** ▪**MAC layer filtering** ▪**Microsegmentation**

DiffServ deployment

- Align DiffServ deployment with each layer's characteristics
  - Do not break the structure by assigning wrong DiffServ responsibilities to network layers.
  - Remember the primary DiffServ functions are:
    - Packet classification.
    - Packet marking.
    - Queuing and/or Dropping.
    - Policing and Shaping with optional metering.

Latency of DiffServ functions

No	Function	What it does?	Latency
1	**Classification**	**Analyze each packet, map packet to classes**	**High**
2	**Marking**	**Insert class identification in each packet**	**Medium**
3	**Queuing / Dropping**	**Continually buffer packets and schedule egress as per queuing discipline**	**Low to medium**
4	**Policing / Shaping**	**Identify non-conformant packets and drop them or shape egress traffic**	**Medium to high**

DiffServ functions and network hierarchy

No	Function	Network Hierarchy
1	Classification	Access
2	Marking	Access, Distribution
3	Queuing / Dropping	Access, Distribution, Core
4	Policing / Shaping	Access, Distribution

Deploying QoS

- Understand application requirements
- Define QoS policy
- Test, test, test .....
- Fine-tune policy (trash and restart if required)
- Deploy QoS
- Monitor flash points and continually tune the QoS

QoS Policy: A more detailed look

Specific Tasks
▪ Define the classes into which services on your network must be divided.
▪ Define filters for the classes
▪ Define flow-control rates for measuring traffic
▪ Define DS codepoints or user priority values to be used in the QoS policy.
▪ If applicable, set up a statistics-monitoring plan for traffic flows on the network.

QoS Policy: Defining classes

- If your company offers SLAs, analyze them thoroughly.
  - Its possible that same applications have been offered to customers with different priorities.
- Is your network carrying “disruptive” traffic?
- What mission critical applications does your network support?
- Please verify everything with measurements.
  - This will reduce fine-tuning efforts in the future.

QoS Policy: Designing filters

- Normally you will have one filter per class.
- Consider use of in-bound and out-bound filters for special applications, e.g. ftp.
- Use advanced tools for application recognition.

QoS Policy: Defining the flow control

- Use metering if required for specific classes
  - Not all classes need to be metered.
  - Metering will put extra overhead on the routers.

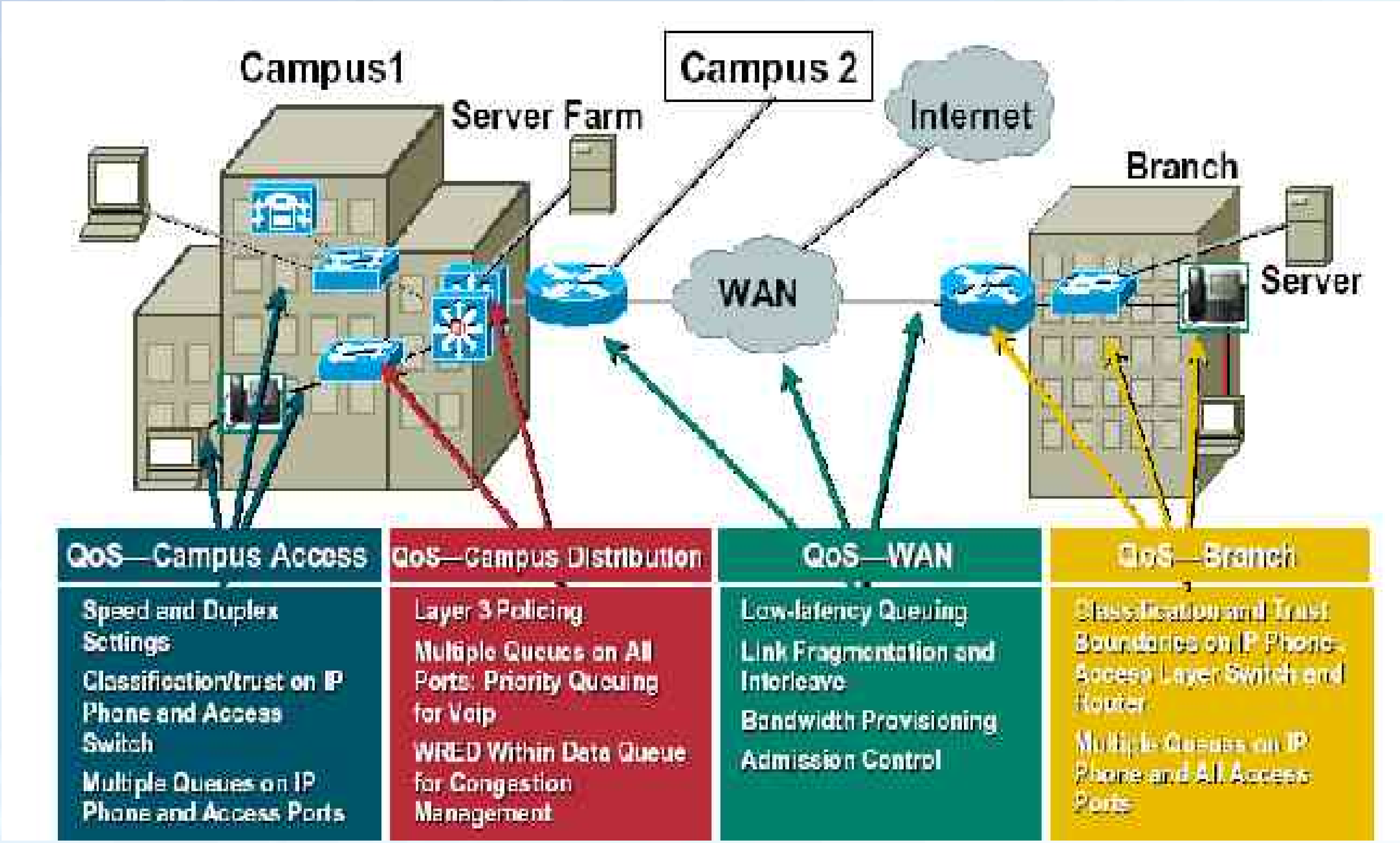
- Metering is useful if
  - SLA guarantees a network load dependent service to a class
  - Traffic from a lower category class tends to flood the network.

- Define the metering (token bucket) parameters in conformance with your policies.

QoS Policy: Using the DS Codepoints

- Use the DS Codepoints judiciously.
  – E.g. EF will be assigned to the highest priority traffic and so on..

QoS Deployment



Configuring QoS for Voice over IP

SETLabs

Why use voice as an example?

- VoIP is being widely deployed on the enterprise and Internet scale too.
- Voice is delay and jitter sensitive.
- Lost / inaudible voice is more irritating than a jittery video clip!

Scenario

- Consider a large enterprise network with a large number of VoIP users, say 10s of thousands
  - Enterprise has mission critical applications also running on the same backbone.
  - Also, the usual non-critical disruptive applications are vying on bandwidth!

Process

- Use the best practices flow we discussed earlier.

- Key information to have before proceeding:
  - Bandwidth required by mission critical applications.
  - Average and minimum bandwidth required by voice.
  - Access layer technologies.

- We will assume that we can classify voice traffic using appropriate filters (port numbers, …)

General guidelines - 1

- Marking
  - Do not mark voice packets in a manner that would cause them to be dropped later.
    - E.g. if you mark a voice packet with DE = 1, there is a great chance of the packet being dropped somewhere down the line.

General guidelines – 2

- Queuing
  - Put voice traffic in a high priority queue.
  - Ensure bandwidth allocated to voice is more than the average aggregate.
  - Use a low latency queuing algorithm e.g. strict priority queuing.
    - Using algorithms like WRED is not suggested.

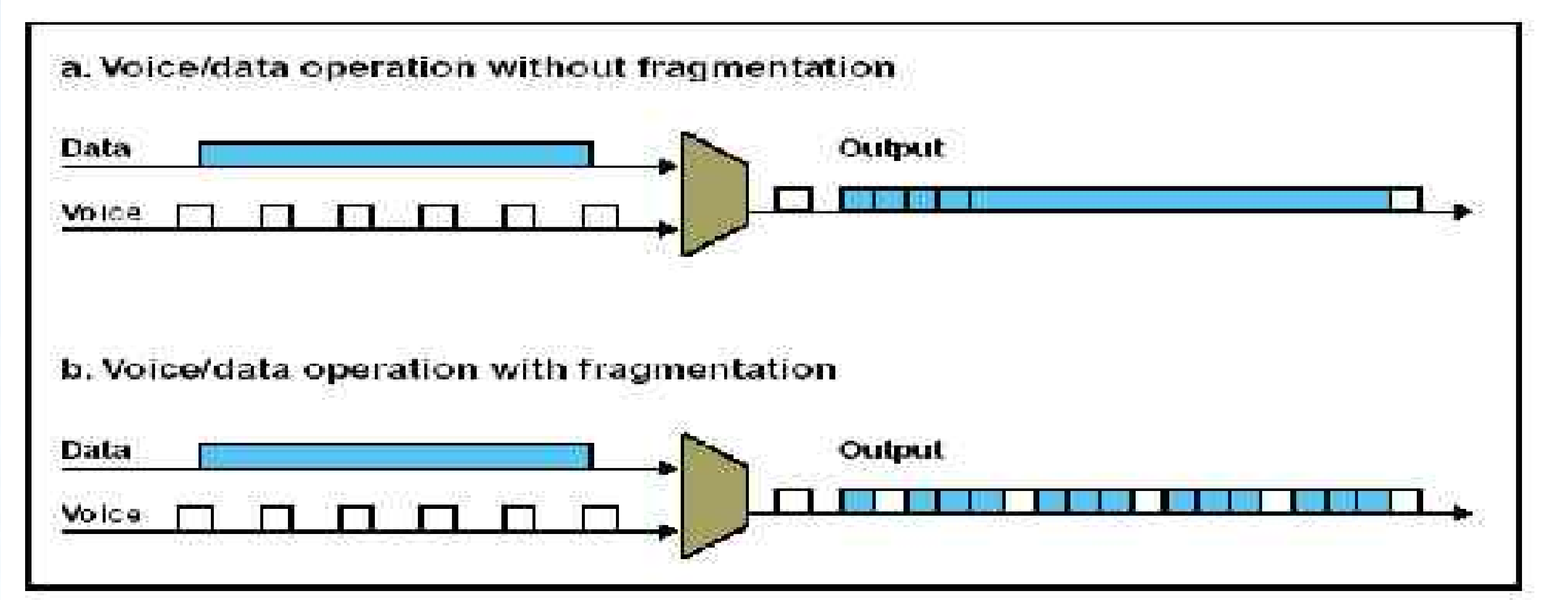
General guidelines - 3

- Handling voice
  - What will happen if voice packets get fragmented?
  - Voice packets should never be fragmented (ideally)!
    - Use appropriate fragmentation size on access links.

- Configuring link layer protocols
  - E.g. frame relay itself has primitive QoS functions (e.g. CIR, Bc etc.) configure them appropriately.
  - Ensure link layer efficiency is high to handle voice.

General guidelines - 4

- Be aware of the serialization delay on slow access links.
- Use link layer fragmentation and interleaving for optimal voice performance.



General guidelines - 5

- Understand VoIP signaling well. You may have to integrate VoIP signaling and QoS signaling together.

- **Case**: there is not sufficient bandwidth to allow a new VoIP call of satisfactory quality
  - How do inform the VoIP gateway of this condition?
  - Routers typically offer integration of RSVP and H.323/SIP for integrated VoIP call setup and QoS. Use such features.

General guidelines - 6

- Are you using an ISP to carry your VoIP traffic?
  - You should have appropriate SLAs with your ISPs.
  - Ensure that voice traffic belongs high(est) priority class provided by the ISP.
  - Ensure that voice traffic does not violate traffic policies.

- Are you using a VPN for the transit through the ISP?
  - Ensure that QoS markers (e.g. DSCP) are copied into the VPN protocol header as well. (e.g. IPSec headers).

After this

- Implement the network and reap the benefits of VoIP.

Summary

- We discussed the need to map QoS deployment to the network architecture.
- Always plan the QoS deployment in details. It saves patch-work in the future.
- Post implementation monitoring is essential.
- If possible choose platforms that provide Policy-based Management of QoS deployment.

Questions

Contact information

Software Engineering and Technology Labs,

Infosys Technologies Ltd

Electronics City, Hosur Road

Bangalore – 566100, India.

Email: vikrant_kaulgud@infosys.com

Phone: +91-80-25660261 ext: 53710