NORTEL NETWORKS
BUSINESS WITHOUT BOUNDARIES

A Survey Of Standards Efforts On Traffic & Congestion Management In Ethernet Networks

David W. Martin

Congestion Management SG

May 24-25, 2004

Long Beach, CA

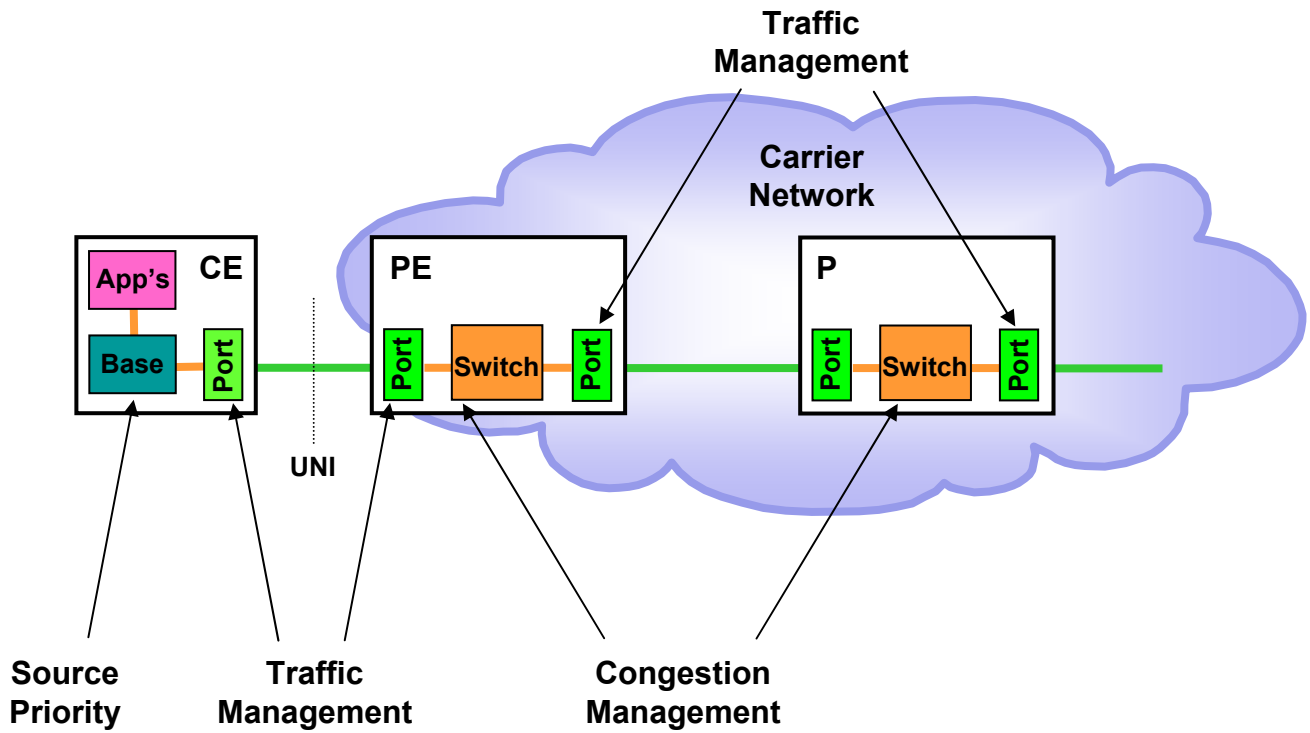
Agenda

- Terminology Alignment
- MEF
- P802.1ad
- ITU-T SG13, Q4
- Summary

Terminology

- **Source Priority**
  - **tagging**
- **Traffic Management:**
  - **traffic shaping**
    - **scheduling, strict priority, round-robin, weighted round-robin, leaky bucket**
  - **classification**
  - **metering**
    - **token bucket**
  - **marking**
  - **dropping** (of Red frames)
- **Congestion Management:**
  - **congestion avoidance**
    - **backpressure, queue mgmt, random early detection, dropping of Yellow frames (Drop Precedence)**
  - **congested state (game over)**
    - **tail drop, dropping of Green frames**

Terminology (cont'd)

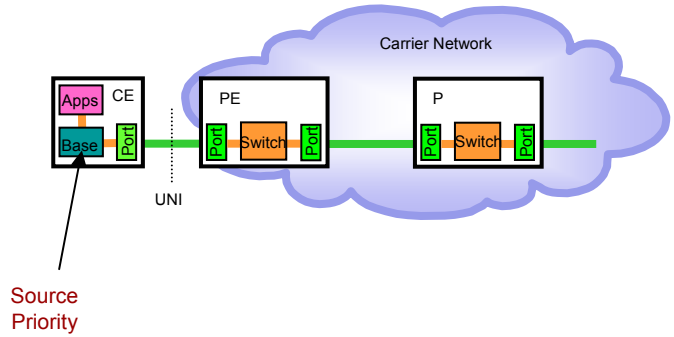


CE: Customer Edge equipment (e.g., a firewall router)

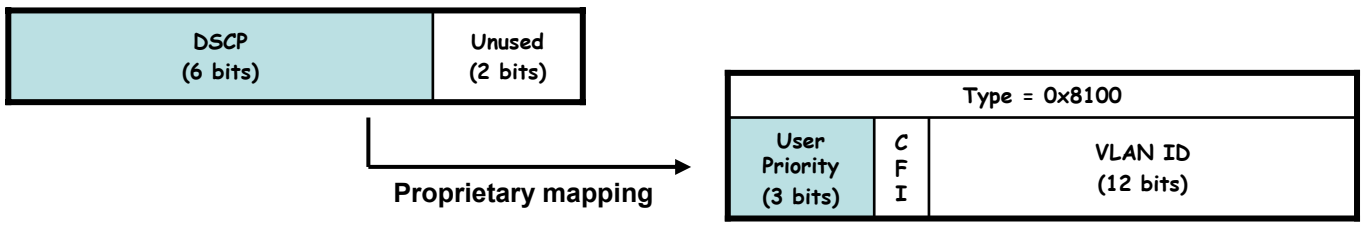
PE: Provider Edge equipment (e.g., an Ethernet switch)

P: Provider interior equipment (e.g., an Ethernet switch)

Host Source Priority

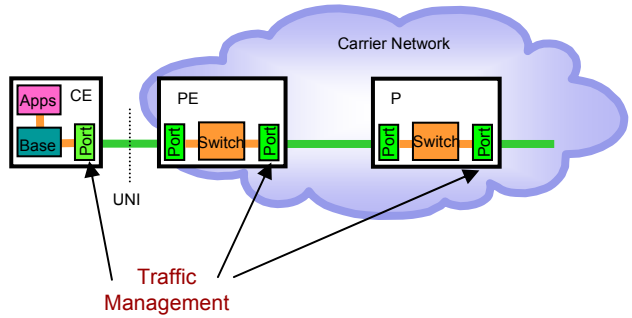


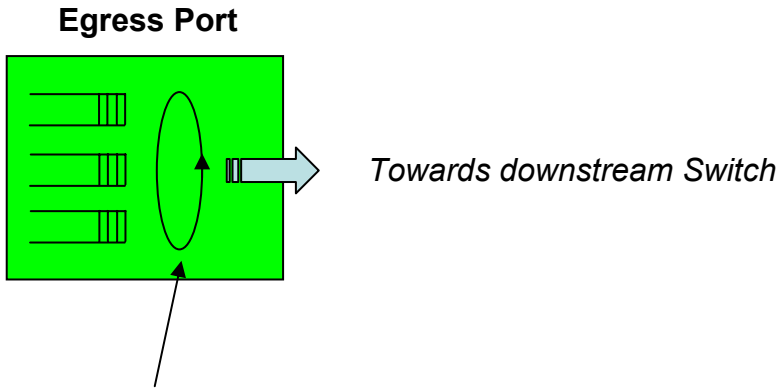
- **Application example: VoIP**
- **DiffServ Code Point set to value X**
- **DSCP value X mapped to an 802.1Q user_priority Y**
- **No standard for DSCP → 802.1Q mapping**



RFC 2474: Definition of the Differentiated Service Field (DS Field) in the IPv4 and IPv6 Headers
RFC 2597: Assured forwarding PHD Group
RFC 3246: An Expedited Forwarding PHB
802.1Q: Virtual Bridged Local Area Networks

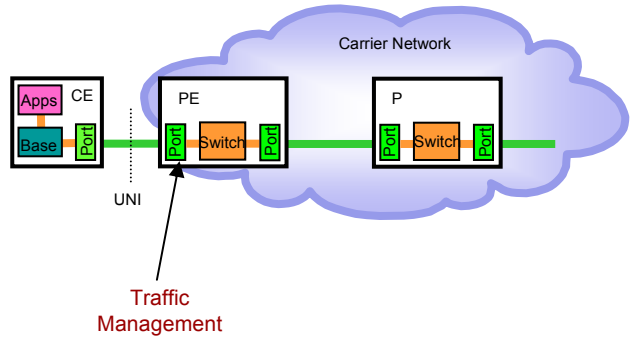
Egress Port Traffic Mgmt





- **Scheduler: strict priority, round-robin, weighted round-robin**
- **Possibly a Leaky Bucket too**
- **For CE, the goal is to shape traffic to minimize number of frames marked Red and tossed at the PE**
- **For PE or P, the goal is to shape traffic to fit the assigned outgoing connection**

Ingress Port Traffic Mgmt



CE-VLAN ID to EVC Mapping

CE-VLAN CoS Grouping

Ingress

UNI

Ethernet Virtual Circuit 1

CoS 7,6,5

CoS 4,3,2

CoS 1,0

Bandwidth Profile per CoS (group)

Ethernet Virtual Circuit 2

Ethernet Virtual Circuit 3

Bandwidth Profile per EVC

*Towards **Meter** function*

Bandwidth Profile per UNI

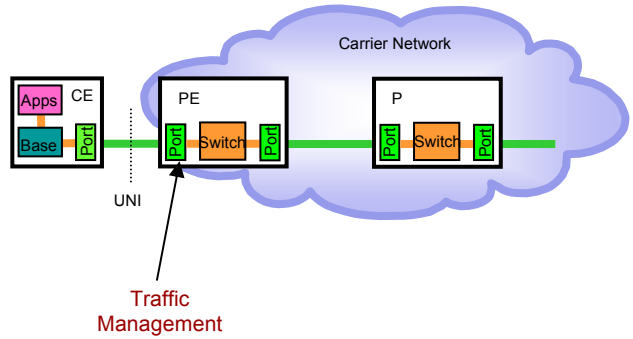
Classifier

Bandwidth Profile:
<CIR, CBS, EIR, EBS, S>

CIR: Committed Information Rate
CBS: Committed Burst Size

EIR: Excess Information Rate
EBS: Excess Burst Size
S: Switch parameter for yellow frames

Ingress Port Traffic Mgmt



Frame of length ℓ_j arrives at t_j

$B_c(t_j) = \min\{CBS, B_c(t_{j-1}) + CIR \times (t_j - t_{j-1})\}$
$B_e(t_j) = \min\{EBS, B_e(t_{j-1}) + EIR \times (t_j - t_{j-1}) + S \times \max[0, B_c(t_{j-1}) + CIR \times (t_j - t_{j-1}) - CBS]\}$

Meter

(Color Blind mode OR frame marked green) AND $\ell_j = B_c(t_j)$ — Yes → Declare frame green $B_c(t_j) = Bc(t_j) - \ell_j$

No

(Color Blind mode OR frame not marked red) AND $\ell_j = B_e(t_j)$ — Yes → Declare frame yellow $B_e(t_j) = B_e(t_j) - \ell_j$

No

Declare frame red

Green frames carried per agreed Delivery Performance

Marker

Yellow frames are marked and carried conditionally

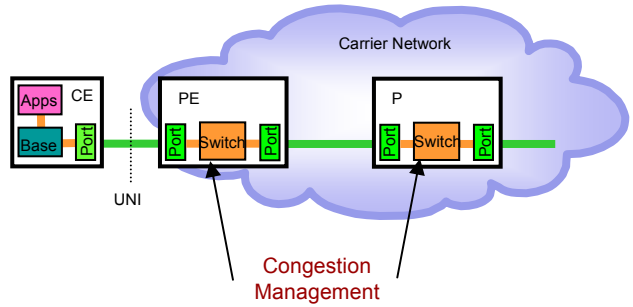
Towards Switch Fabric

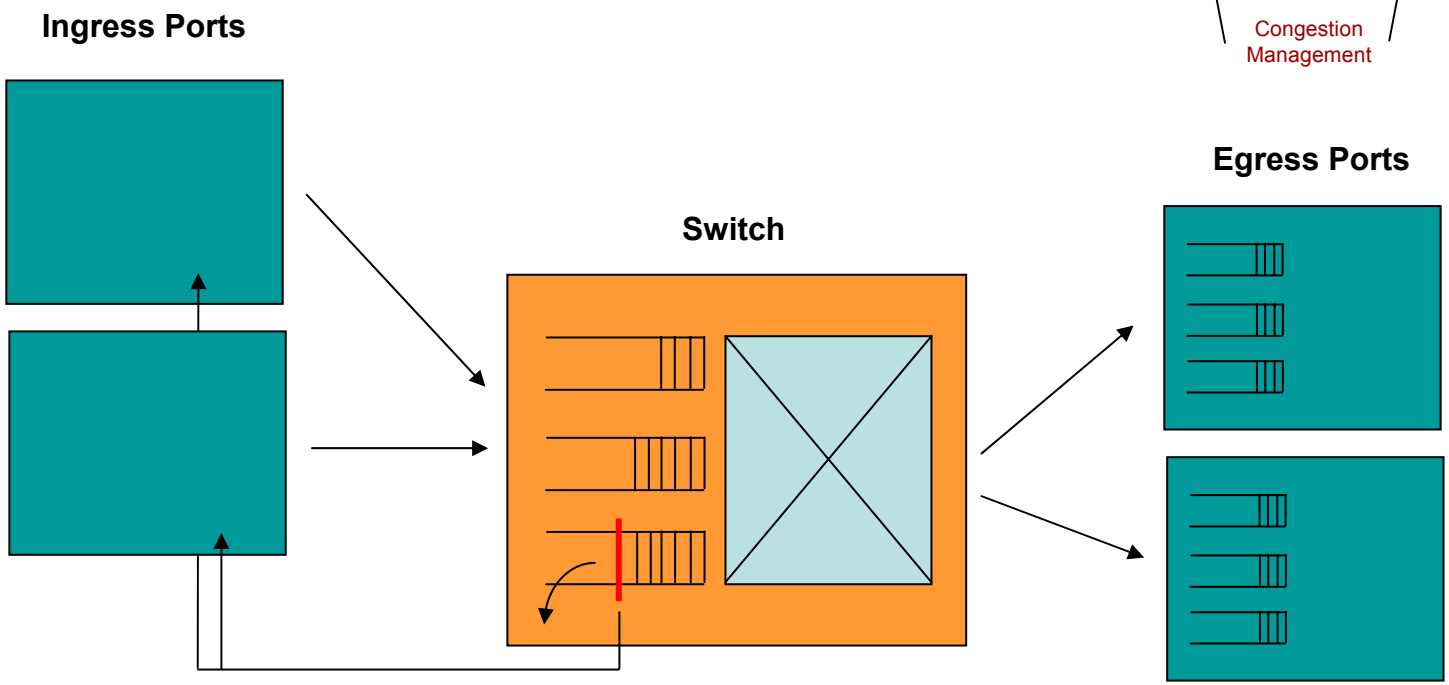
Red frames dropped

Dropper

Delivery Performance: <Frame Delay, Frame Jitter, Frame Loss, EVC Availability>

Congestion Mgmt





Congestion Avoidance

1. Backpressure to ports to reduce offered load
2. *Yellow* frames are dropped – identified by Drop Precedence marking

Congested State

3. *Green* frames are dropped too

Drop Precedence = Discard Precedence = Discard Eligibility = Discard Priority = Drop Priority = Drop Eligibility
(e.g., uses CFI and / or user_priority – still TBD)

MEF

- **Developing suite of Ethernet Service & Traffic Mgmt spec's**
- **Do not spec the Marker encoding**
- **Not currently working on Congestion Mgmt**

Ethernet Services Model, Phase 1 – MEF 1, 2003
Ethernet Services Definition, Phase 1
Traffic Management Specification, Phase 1

P802.1ad Provider Bridges

- An amendment to 802.1Q-1998
- To enable a Service Provider to offer the equivalent of separate LAN segments, bridged or virtual bridged LANs, to a number of users
- Will spec the Marker encoding (CFI / user_priority - TBD)
- Not currently working on Congestion Mgmt

http://www.ieee802.org/1/pages/802.1ad.html

ITU-T SG13, Q4

- **Just starting work on Traffic Mgmt**
- **Will reference the MEF and P802.1ad outputs**
- **Not currently working on Congestion Mgmt**

Summary

- **No other standards body currently addressing Ethernet Congestion Mgmt**

Food For Thought (?)

- **Consider developing an 802.1 standard for per-VLAN ID (and / or per-VLAN CoS?) backpressure mechanism**
- **The Key Issue: What is the impact on higher layer Congestion Mgmt schemes (i.e., TCP)?**