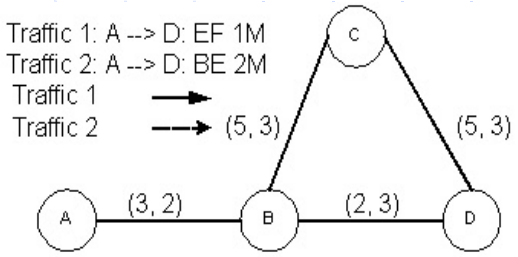
Multi-class routing in DiffServ networks

Peng Zhang
pgzhang@netlab.hut.fi
Networking Laboratory
January 8th 2004, IRoNet seminar

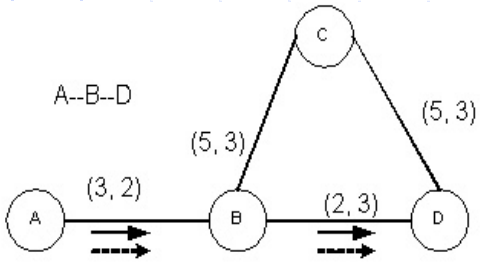
Motivation

- Best-effort oriented IP networks are expected to migrate into multi-class networks, e.g.,
  - DiffServ networks
  - MPLS networks
- Multi-class networks are expected to support traffic with
  - Various traffic characteristics, and
  - Various QoS demands
- Traditional routing belongs to single-class routing (SCR) scenario, which calculates same routing for all kinds of traffic classes.
- In contrast, we propose multi-class routing (MCR), which calculates different routing for different traffic classes.

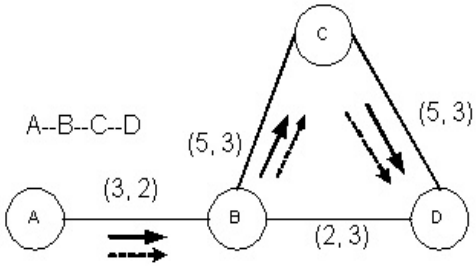
A simple example



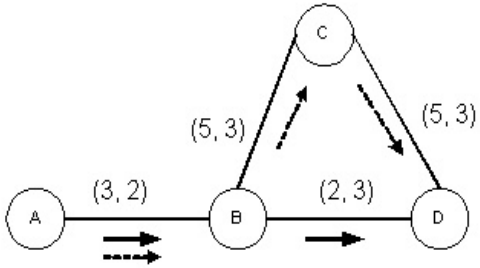
(a) Network and traffic



(b) SCR: SP



(c) SCR: BSP



(d) MCR: (SP, BSP)

- Two traffic from A to D
  - Traffic 1: EF 1M
  - Traffic 2: BE 2M
- Goals:
  - Minimum delay for EF
  - Maximum throughput for BE
- A link is denoted by (bandwidth, delay)

Only MCR can achieve routing optimization of each class.

Multi-class routing (MCR)

- Clarification
  - MCR is related but not same as QoS routing (QoSR)
    - MCR may use QoS routing algorithms
  - MCR for MPLS is referred to PERD (a legacy name), which is based on requests.

- We propose MCR mainly because of the following reasons:
  - Different classes may have different QoS requirements, e.g., video/voice traffic conveyed by EF class requires minimum delay and delay jitter while WWW/FTP traffic conveyed by BE class requires maximum throughput.
  - Migrating SCR into MCR requires only a few changes, which is much simpler than MPLS. Memory cost is not significant in modern routers.
  - DiffServ might be a policy based network, where MCR can be used to compute optimal routes that are downloadable to each router.
  - Our previous simulation results prove that MCR can improve the performance of not only high priority classes but also low priority classes.

MCR algorithm

MCR scheme:

Begin
Initialization {
Set routing metric and method for each class,
i.e., set $D = \{d_i \mid i \in C\}$
Initialize routes for each class,
i.e., $P = \{P_i \mid i \in C\}$
}
Start{
Step 1:
Started with the highest priority class, i.e., i=0,
Compute P_i *with* d_i
Step 2:
Proceed to lower priority class, i.e., i++;
Compute P_i *with* d_i
Step 3:
If this is not the lowest priority class,
go to step 2.
}
End.

Notations:

$C = \{i \mid i = 0,1,2...\}$, the set of traffic classes. These classes are ordered from the highest priority to the lowest priority.

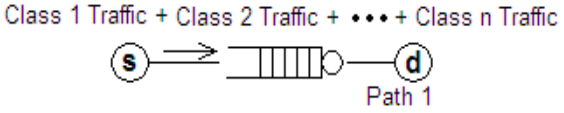
$M = \{m_j \mid j = 0,1,2...\}$, the set of routing metrics.

$R = \{r_k \mid k = 0,1,2...\}$, the set of routing methods.

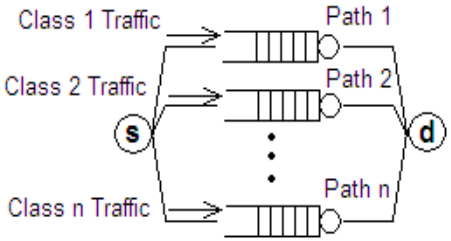
$d_i = \{m_j, r_k \mid m_j \in M, r_k \in R\}$, a duple of routing metric and routing method for traffic class *i*.

P_i, the set of routes for traffic class *i*.

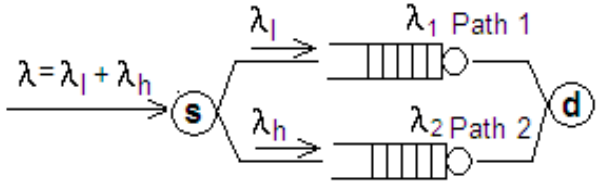
MCR modeling



(a) SCR



(b) MCR



$$\Delta T = T_{P1}^{MCR} - T_{P1}^{SP}$$

$$= \frac{\lambda_1 + \lambda_l}{\mu_{P1}(\mu_{P1} - \lambda_1 - \lambda_l)} - \frac{\lambda_1 + \lambda}{\mu_{P1}(\mu_{P1} - \lambda_1 - \lambda)}$$

$$= -\frac{\lambda_h}{(\mu_{P1} - \lambda_1 - \lambda_l)(\mu_{P1} - \lambda_1 - \lambda)}$$

- In normal case, we get delta T < 0, which means MCR achieves smaller cost (average processing time) than SCR.

Optimization problem of BE class

- Assumptions:
  - Two classes: EF class and BE class
  - The performance of BE class is mainly concerned because
    - BE class takes the biggest portion of all network traffic
    - EF class can be guaranteed due to its high priority.
- BECOP (BE class optimization problem)
  - Given a DiffServ network G(V,E),
    - Each pair of nodes is associated with an EF traffic pair with equal rate.
    - Each pair of nodes is associated with a BE traffic pair with equal rate.
  - Given an EF rate and a routing algorithm r, with the increase of BE rate, the BE rate when some link is saturated is denoted by B_s
  - B_{max} = max{B_{sr1}, B_{sr2}, ...} for BE class, when applying different routing algorithms r1, r2, ...
  - BECOP is to find optimal routing algorithm combination of MCR(r_{ef}, r_{be}) and Bmax

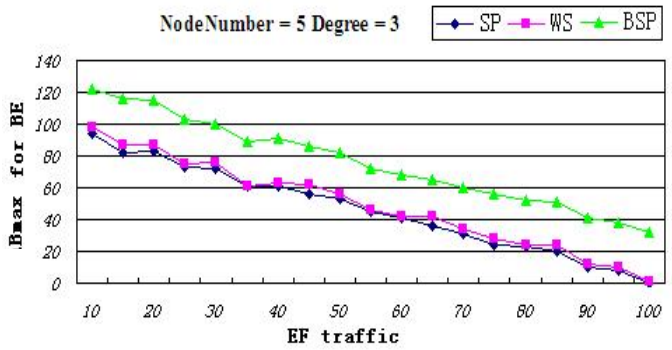
Simulation study

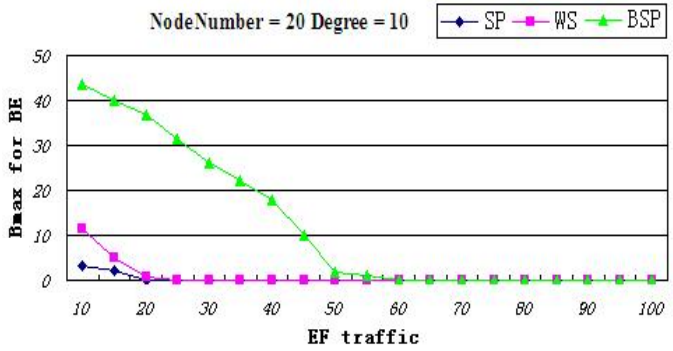
- Topology generator: thousands of topologies
  - Number of nodes N: 5, 12, 20
  - Variance index for the degree of each node D_{var} : 3, 6, 10
  - Variance index of link capacities C_{var} : 10

  Given a specific value of D_{var}, we generate a random number within $[1, D_{var}]$ for each node as its degree. And given a specific value of C_{var}, we generate a random number within $[100, 100 \times C_{var}]$ for each link as its capacity.

  1000 topologies for each configuration
- Algorithms for EF class and BE class:
  - SP (shortest path)
  - WSP (widest shortest path)
  - BSP (bandwidth-inversion shortest path)

Results – B_{max} under SCR

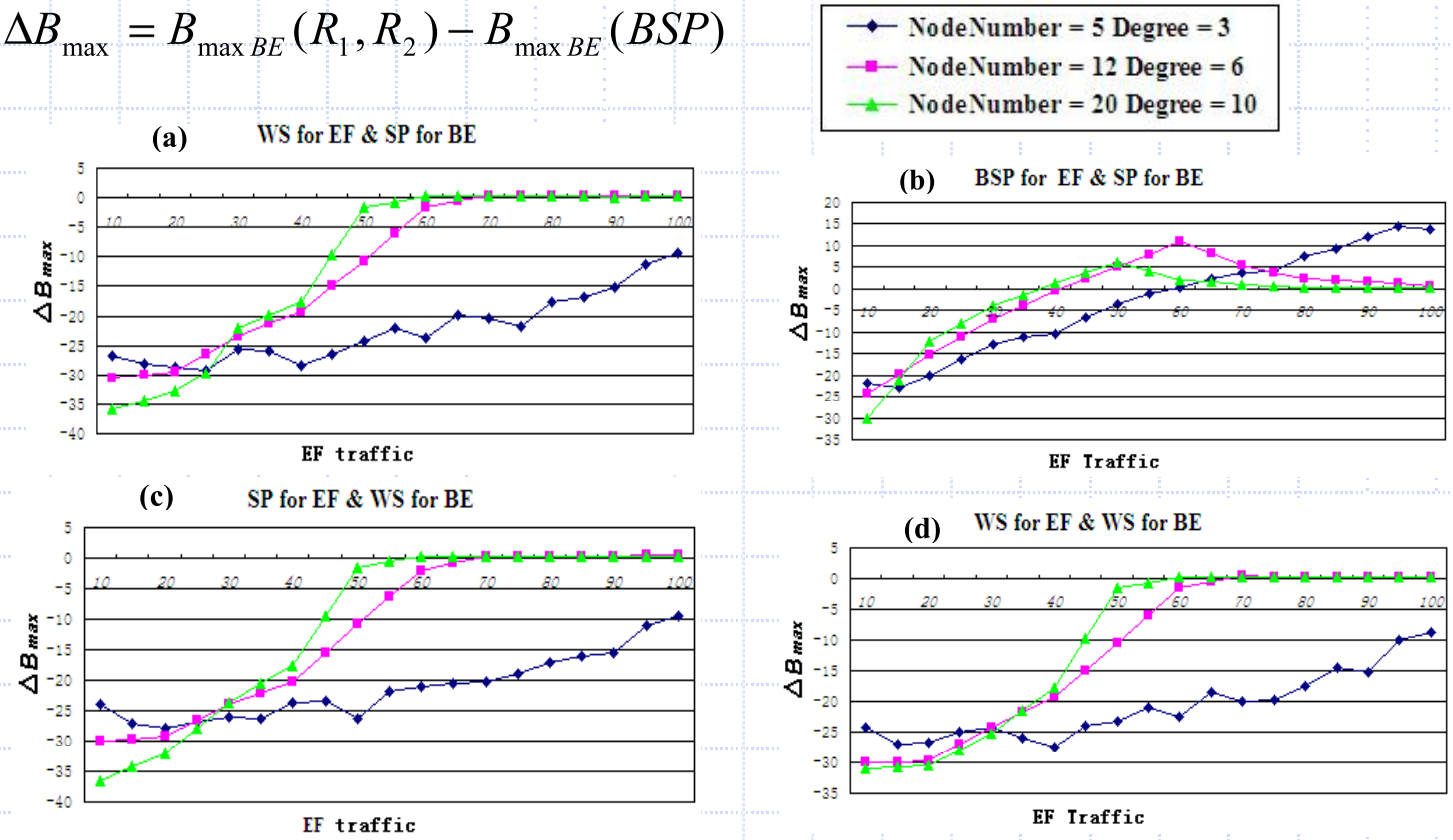




- With the increase of EF traffic, B_{max} decreases.
- Given same EF traffic, B_{max} decreases significantly with the increase of network complexity and size.
- BSP algorithm achieves the best B_{max} for BE class under SCR scenario.

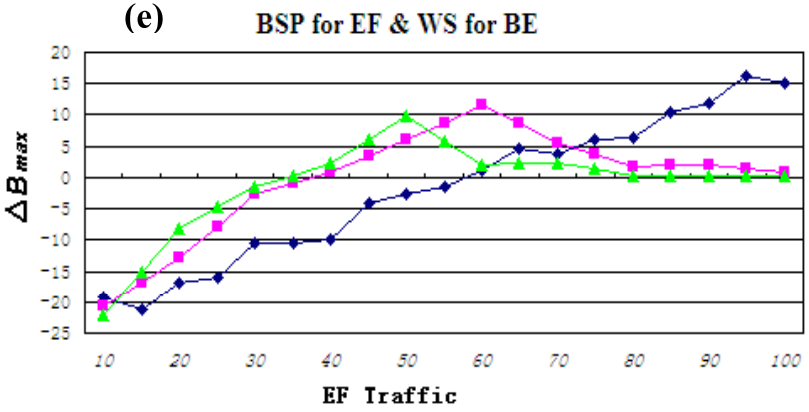
Results – MCR vs. SCR(BSP)

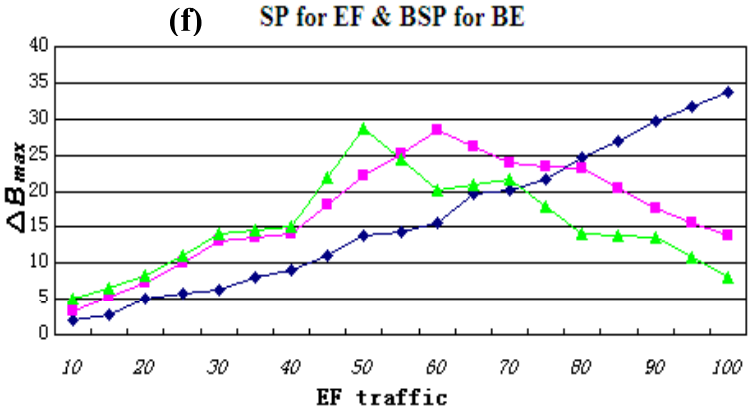
$$\Delta B_{\max} = B_{\max BE}(R_1, R_2) - B_{\max BE}(BSP)$$

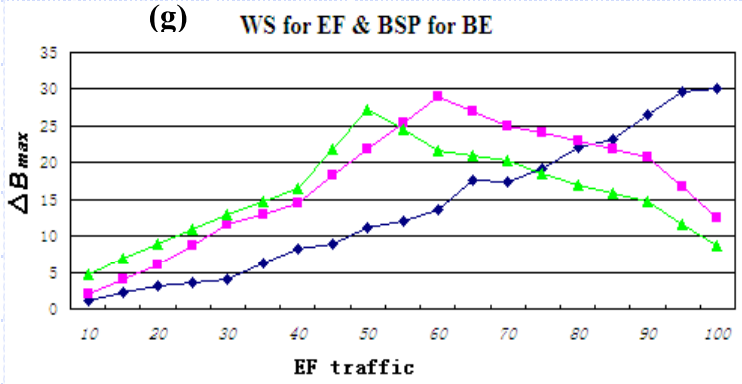


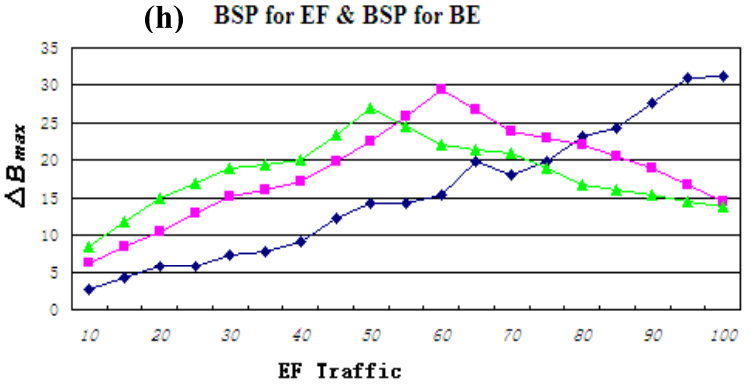
- These routing combinations of MCR achieve worse performance than SCR(BSP)

Results (cont'd)









- When BSP is used for BE class, MCR achieves much better performance than SCR(BSP).
- With the increase of EF traffic, MCR(x, BSP) is becoming better than SCR(BSP)
- When network becomes more complicated, MCR is much better than SCR.

Prototyping of MCR

- Based on Zebra open source and major modifications including
  - Single-class routing table → multi-class routing table
  - Normal link state database → TE link state database according to RFC3630
  - Single class route computation → multi-class route computation
- RFC3630 – Traffic engineering (TE) extension OSPF version 2
  - Became RFC in September 2003
  - Does not specify any QoS routing algorithm, but provides the TE extension to link state database which will become a TE link state database.
  - Has a LSA that can contain different link state information (i.e., available bandwidth) for upto 8 priority classes. This exactly favors the requirement of MCR which needs a link state protocol to exchange link state information for different classes

Future work

- Experimental study
- MCR for intra-area study
- Extend BSP to consider EF traffic ratio on a link

Reference:

Master thesis: Investigation and analysis of optimal multi-class routing algorithms for IP network. June 2003. HUT

Thank you! Any comments?