Traffic Engineering with AIMD in MPLS Networks

Jianping Wang*
Stephen Patek**
Haiyong Wang*
Jorg Liebeherr*

*Department of Computer Science
**Department of Systems and Information Engineering
University of Virginia

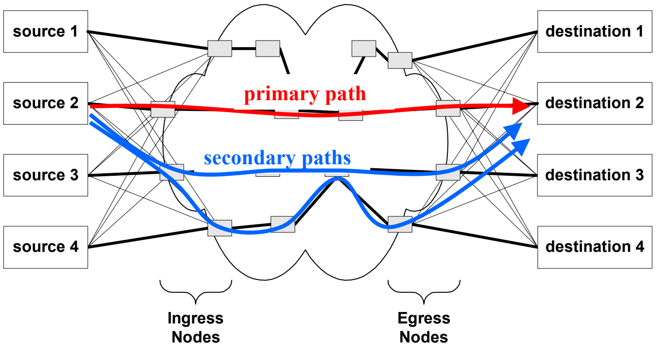
http://mng.cs.virginia.edu

MPLS

Multiprotocol Label Switching (MPLS) offers opportunities for improving Internet services through traffic engineering

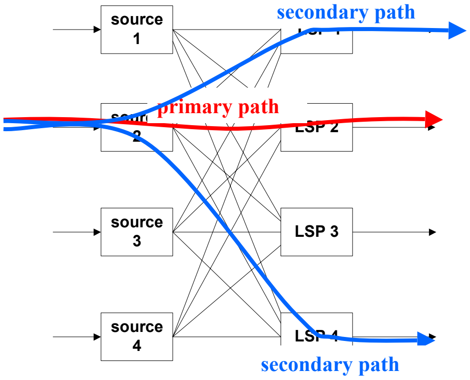
- MPLS makes it possible for network engineers to set up dedicated label switched paths (LSPs) with reserved bandwidth for the purpose of optimally distributing traffic across a given network

MPLS Network



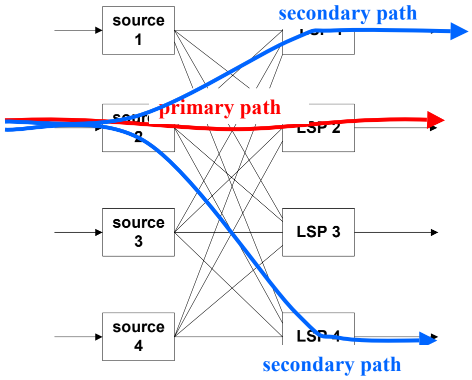
- Flows (traffic between source/destination pairs) may make use of multiple LSPs.
  - Primary vs. Secondary Paths

Simplified MPLS Network



- *N* sources and *N* LSPs
- *LSP i* is the primary path for *source i*. Other *LSPs* $(i \neq j)$ are secondary paths
- *Source i* has a load of λ_i and a throughput of γ_i
- *LSP i* has a capacity of B_i

Simplified MPLS Network



- *Problem: Given* load λ_i and capacity B_i
  Assign flow from source i to primary path and secondary paths by satisfying a given set of objectives

---

Objectives for Flow Assignment

- Efficiency:
  - all resources should be consumed or all sources should be satisfied
- Fairness:
  - Satisfy given fairness criteria
- Primary Path First:
  - Minimize traffic on secondary paths
- Simple and Distributed Allocation:
  - Binary Feedback, Stability

Background

- Binary feedback rate control schemes (AIMD)
  - Jacobson (1988), Jain and Ramakrishnan (1988, 1990, 1996), Chiu and Jain (1989)
- MATE, MPLS Adpative Traffic Engineering
  - Elwalid et al. (2001)
- Optimization-based end-to-end congestion control and fairness
  - Le Boudec (1999), Kelly (1997, 1998), Massoulie and Roberts (1999), Vojnovic et al. (2000)

Outline

1. Fairness and Efficiency
2. PPF Criterion
3. AIMD algorithms
4. NS-2 Experiments
5. Conclusions

Bandwidth allocation

- Two allocation schemes
- Owned Resources: Each source can consume the entire capacity of its primary path (B_i), and it can obtain bandwidth on its secondary paths
- Pooled Resources: The aggregate capacity on all LSPs ($\sum_i B_i$) is distributed across all sources, without regard to the capacity on primary paths

Rate Allocation

- A rate allocation is a relation $R = \{\lambda_i, \gamma_i\}$ $(1 \leq i \leq N)$ such that both $\gamma_i \leq \lambda_i$ and $0 \leq \sum_i \gamma_i \leq \sum_i B_i$
- A rate allocation is efficient if the following hold:
  a) If $\sum_i \lambda_i < \sum_i B_i$ then $\sum_i \gamma_i = \sum_i \lambda_i$
  b) If $\sum_i \lambda_i \geq \sum_i B_i$ then $\sum_i \gamma_i = \sum_i B_i$

  If case b) holds, we say that the rate allocation is saturating

Fairness for pooled resources

- A rate allocation for pooled resources is fair if there exists a value $\alpha^p > 0$ (fair share) such that for each *source i* it holds that

$$\gamma_i = \min \{\lambda_i, \alpha^p \}$$

- The fair share α^p in a network with pooled resources is given by

$$\alpha^p = \begin{cases} \dfrac{\sum_{i=1}^{N} B_i - \sum_{j \in U} \lambda_j}{|O|}, & if\ |O| > 0 \\ \infty & ,otherwise \end{cases}$$

where $U = \{j \mid \lambda_j < \alpha^p \}$ and $O = \{j \mid \lambda_j \geq \alpha^p \}$

Fairness for owned resources

- A rate allocation for owned resources is fair if there exists a value $\alpha^o > 0$ (fair share) such that for each *source i* it holds that

$$\gamma_i = \min \{\lambda_i, B_i + \alpha^o \}$$

- Interpretation: Each source can use all of its primary bandwidth and a fair share of the surplus capacity

- Define:

$U' = \{j \mid \lambda_j < B_i \}$
$O' = \{j \mid \lambda_j \geq B_i \}$
$C' = \sum_{i \in U'} (B_i - \lambda_i)$ (total surplus capacity)
$\lambda_i' = \lambda_i - B_i$, if $i \in O'$

Fair share for owned resources

- The fair share of the surplus is given by

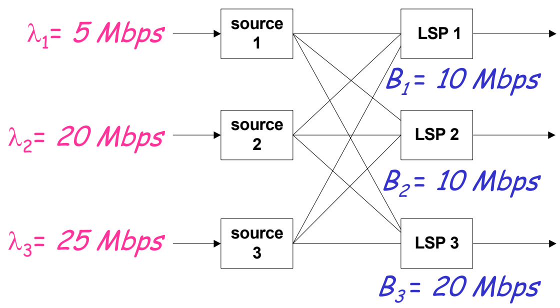
$$\alpha^o = \begin{cases} \dfrac{C' - \sum_{j \in U''} \lambda'_j}{|O''|}, & if\ |O''| > 0 \\ \infty & ,otherwise \end{cases}$$

where $U'' = \{j \in O' \mid \lambda'_i < \alpha^o\}$ and $O'' = \{j \in O' \mid \lambda'_i \geq \alpha^o\}$

- The rate allocation is given by

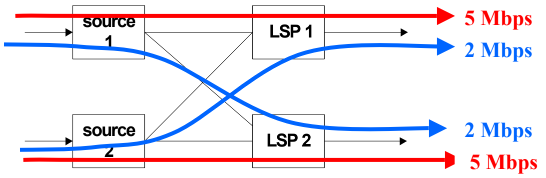
$$\gamma_i = \begin{cases} \lambda_i & , if\ i \in U'\ or\ i \in U'' \\ B_i + \alpha^o & , i \in O'' \end{cases}$$

Example

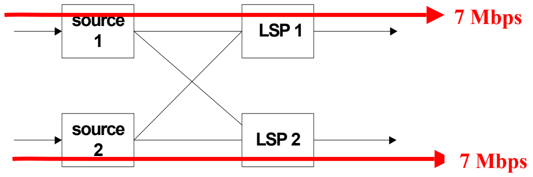


pooled resources	***owned resources***
α^p =17.5	α^o =2.5
γ_1= 5	γ_1= 5
γ_2= 17.5	γ_2= 12.5
γ_3= 17.5	γ_3= 22.5

Primary Path First (PPF)



Sources “spread” the traffic on secondary paths even though there is enough capacity on primary paths



Traffic is concentrated on primary paths

The PPF objective maximizes traffic on primary paths

Primary Path First (PPF)

- Define routing matrix X

  x_{ij} amount of traffic sent by source i on path j.

  $\sum_{i \neq j} x_{ij}$: secondary traffic
  x_{ii} : primary traffic

- A saturating rate allocation is PPF-optimal if it solves the linear program

$$\min \quad \sum_i \sum_{i \neq j} x_{ij}$$

subject to

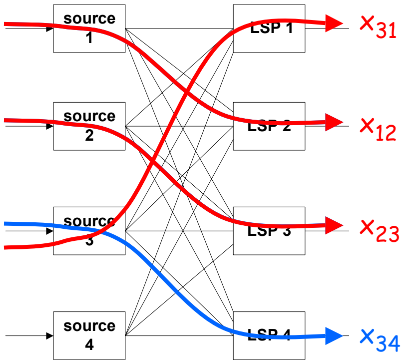
$$\sum_j x_{ij} = \gamma_i \quad , i = 1,2,\ldots,N$$

$$\sum_i x_{ij} = B_j \quad , j = 1,2,\ldots,N$$

$$x_{ij} \geq 0 \quad , i,j = 1,2,\ldots,N$$

Characterizing PPF Solutions

- Chain: $\langle i_1 i_2 \ldots i_k \rangle$, $k > 2$
  $x_{i_1 i_2} > 0$, $x_{i_2 i_3} > 0$,
  $x_{i_3 i_4} > 0, \ldots,$
  $x_{i_{k-1} i_k} > 0$
- Cycle: $\langle i_1 i_2 \ldots i_k \rangle$, $k > 2$, $i_1 = i_k$
  $x_{i_1 i_2} > 0$, $x_{i_2 i_3} > 0$,
  $x_{i_3 i_4} > 0, \ldots,$
  $x_{i_{k-1} i_k} > 0$



Proposition: A routing matrix X is PPF-optimal if and only if there is no chain and no cycle

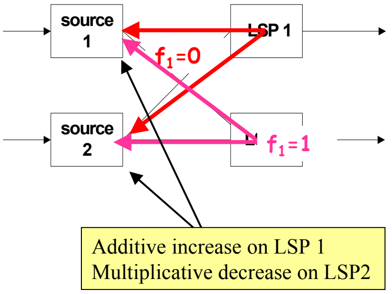
---

Distributed Rate Allocation: Multipath AIMD

Binary Feedback from LSPs:

Each LSP j periodically sends messages to all sources containing a binary signal $f_j = \{0,1\}$ indicating its congestion state

- Utilization = B_j → $f_j = 1$
- Utilization < B_j → $f_j = 0$



Additive increase on LSP 1
Multiplicative decrease on LSP2

Sources adapt rate using AIMD:

- $f_j = 1$ → multiplicative decrease ($0 \leq k_r \leq 1$)
- $f_j = 0$ → additive increase ($k_a \geq 0$)

Multipath-AIMD

For pooled resources:

$$x_{ij} \leftarrow \begin{cases} x_{ij} + k_a & , \mathit{if} \ \sum_{l=1}^{N} x_{il} < \lambda_i \ \mathit{and} \ f_j = 0 \\ x_{ij} & , \mathit{if} \ \sum_{l=1}^{N} x_{il} \geq \lambda_i \ \mathit{and} \ f_j = 0 \\ x_{ij} \cdot (1 - k_r) & , \mathit{if} \ f_j = 1 \end{cases}$$

Multipath-AIMD

For owned resources:

i = *j*: $\lambda_i \leq B_i$

$$x_{ii} \leftarrow \begin{cases} \min\{\, x_{ii} + k_a, \lambda_i \,\} & , \mathit{if} \ x_{ii} \leq \lambda_i \\ x_{ii} \cdot (1 - k_r) & , \mathit{if} \ x_{ii} > \lambda_i \end{cases}$$

$\lambda_i > B_i$

$$x_{ii} \leftarrow \min\{\ x_{ii} + k_a, B_i \ \}$$

$i \neq j$:

$$x_{ij} \leftarrow \begin{cases} x_{ij} & , \mathit{if} \ x_{ii} < B_i \ \mathit{or} \ \left(x_{ii} = B_i, \sum_{l=1}^{N} x_{il} \geq \lambda_i, f_j = 0\right) \\ x_{ij} + k_a & , \mathit{if} \ x_{ii} = B_i , \sum_{l=1}^{N} x_{il} < \lambda_i, f_j = 0 \\ x_{ij} \cdot (1 - k_r) & , \mathit{if} \ x_{ii} = B_i , f_j = 1 \end{cases}$$

Feedback for PPF correction

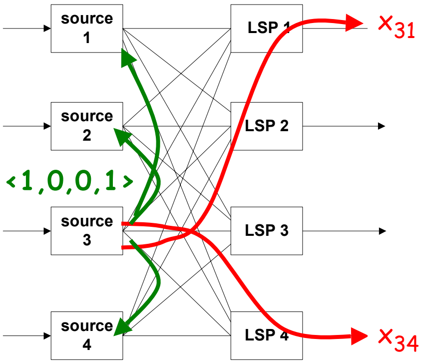
Extra feedback is required to enforce PPF

- Sources exchange bit vectors
- Exchange is asynchronous
- Bit vector of source i :

  $m_i = < m_{ij}, m_{ij}, \ldots, m_{iN} >$

  $m_{ij} = 0, \text{ if } x_{ij} = 0$

  $m_{ij} = 1, \text{ if } x_{ij} > 0$



PPF correction

- After each multipath-AIMD adjustment, sources perform a PPF correction:

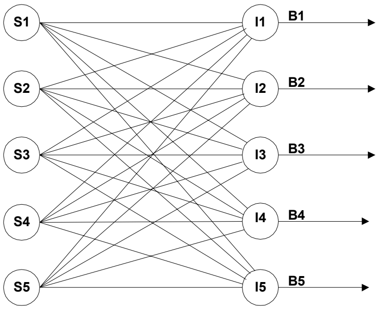
$$x_{ij} \leftarrow \begin{cases} \max\{x_{ij} - K, 0\} & , if \sum_{l \neq i} m_{li} > 0 \\ x_{ij} & , otherwise \end{cases}$$

$$x_{ii} \leftarrow x_{ii} + \sum_{j \neq i} \min\{K, x_{ij}\}$$

Conflict: PPF correction tends to push flow onto primary paths, interfering with the natural tendency of AIMD to arrive at a fair distribution of the load

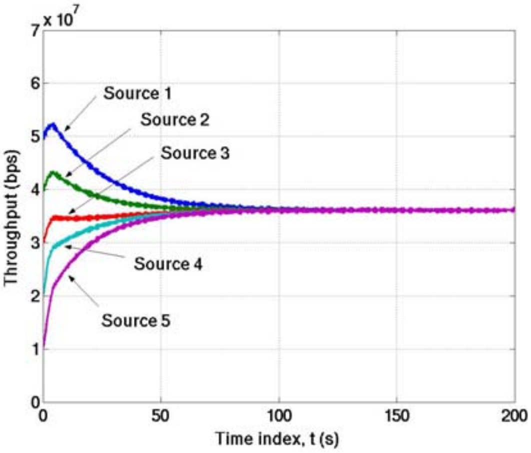
ns-2 simulation

- Packet level simulation
- 5 sources, 5 LSPs
- LSP Capacities
  B_i=(50,40,30,30,30) Mbps



Access link bandwidth: 100 Mbps
Propagation delay: 5 ms
Frequency of
congestion feedback Δ_{LSP} =5ms
source update Δ_{SRC} =5ms
Packet size: 50 Bytes
AIMD parameters:
k_a = 0.1 Mbps
k_r = 0.01

Experiment 1: Basic Multipath-AIMD with Pooled Resources

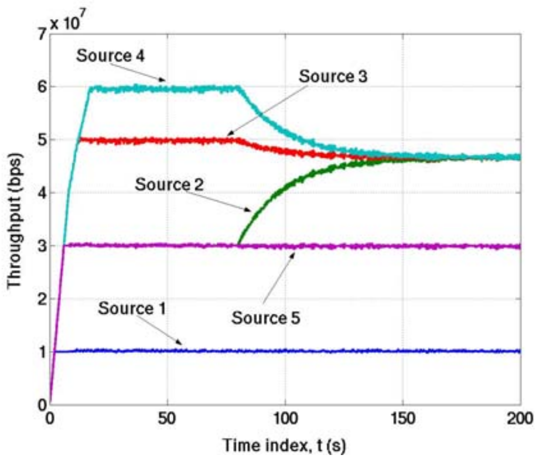


- All sources are always backlogged ("Greedy Sources")
- All sources converge within 90 seconds to the fair-share allocation
- The final routing matrix is not PPF optimal

Experiment 2: Basic Multipath-AIMD

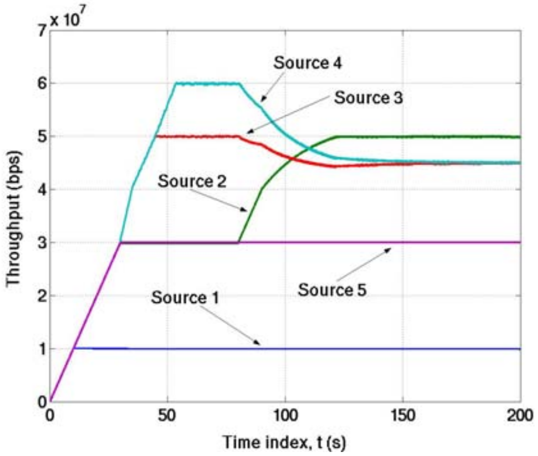
	Initial scenario $0 \le t < 80$ sec			Final scenario $80 \le t < 200$ sec		
Source i	Load λ_i	Tput γ_i pooled	Tput γ_i owned	Load λ_i	Tput γ_i pooled	Tput γ_i owned
1	10	10	10	10	10	10
2	30	30	30	**50**	46.7	50
3	50	50	50	50	46.7	45
4	60	60	60	60	46.7	45
5	30	30	30	30	30	30

Experiment 2: Pooled Resources



- Note convergence to new after load change of source 2 at 80 sec
- Solution not PPF optimal

Experiment 2: Owned Resources



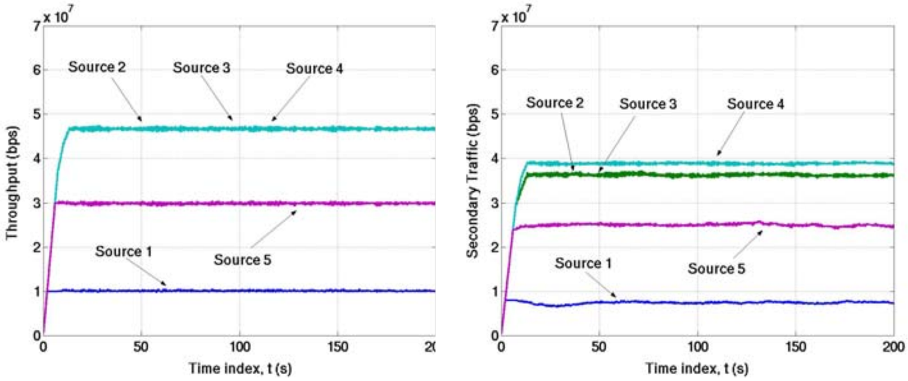
Note convergence to new after load change of source 2 at 80 sec

Solution is not PPF optimal

Experiments with PPF Correction

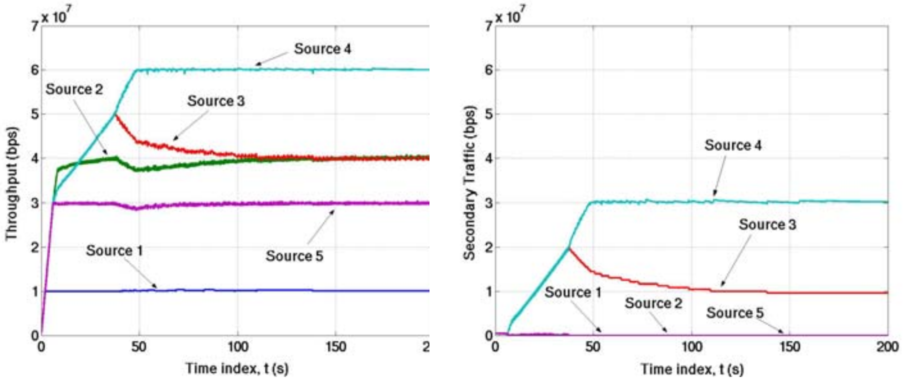
- Loads
  λ_i=(50,40,30,30,30) Mbps
- Resources are pooled
- Sources exchange bit vector over a full-duplex link
  - Bandwidth: 100 Mbps
  - Propagation delay: 1 ms
  - Frequency Δ_{PPF} =5ms
- PPF parameters:
  K = 0.00001 Mbps
  K=0.01 Mbps

Experiment 3: Multipath-AIMD with PPF correction with K = 0.00001 Mbps



- Final allocation is fair, but not PPF-optimal

Experiment 3: Multipath-AIMD with PPF correction with K = 0.01 Mbps



- Final allocation is PPF-optimal, but not fair

Conclusions

- We have proposed **multipath-AIMD** to achieve a fair and PPF-optimal rate allocation to flows in an MPLS network
  - **Multipath-AIMD** seeks to provide a fair allocation of throughput to each source
  - **Multipath-AIMD with PPF Correction** seeks to reduce the volume of secondary path traffic
- Both algorithms rely upon binary feedback information

- **Observation:** Difficult (impossible?) to achieve PPF and fairness objectives simultaneously

- **Open issues:**
  - Relax restrictions on topology
  - (When) is it possible to be both fair and PPF optimal?