[Telecommunications Research Center Vienna]

Dynamic Traffic Engineering for Future IP Networks

Ivan Gojmerac, Thomas Ziegler, Fabio Ricciato, Peter Reichl
Telecommunications Research Center Vienna (ftw.)

4. Würzburger Workshop
"IP Netzmanagement, IP Netzplanung und Optimierung"

Würzburg, 27 July 2004

Outline

- Introduction to traffic engineering
- Adaptive Multi-Path (AMP) algorithm
- Performance evaluation and results
- Summary and outlook

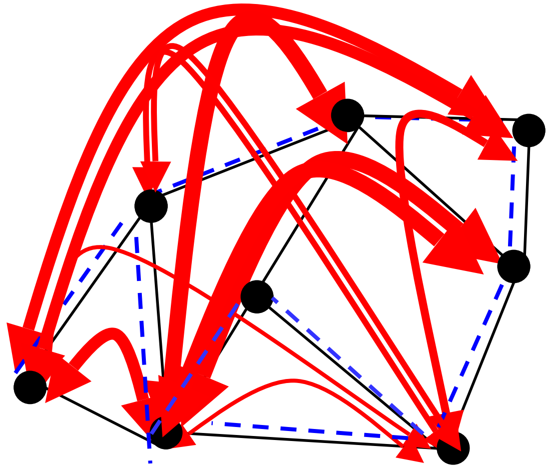
What is "Traffic Engineering" (TE)?

- Traffic engineering is defined as performance optimization of operational networks (IETF)
  - Consider the traffic at the macroscopic level
  - Consider the network as a set of *limited* resources
    - Transmission bandwidth, switching throughput
- Traffic engineering tries to optimally match **traffic demands** with the available **network** resources by acting on **routing**

Traffic Demands

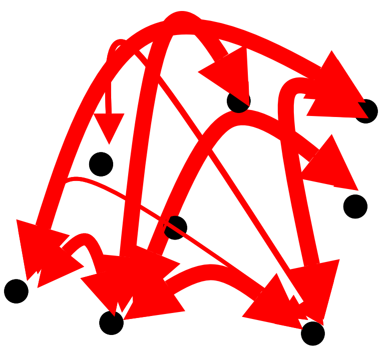
Network

Routing



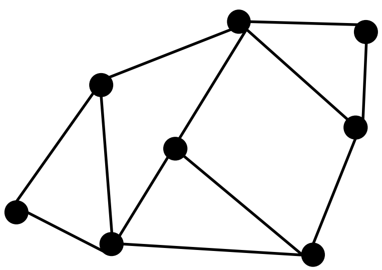
- Traffic engineering methods for IP networks:
  - Link weight optimization in native IP networks
  - Optimization of Multi-Protocol Label Switched (MPLS) networks
  - Algorithmic approaches (dynamic routing in the ARPAnet, OMP)

Example of Connection-Less TE: Link Weight Optimization

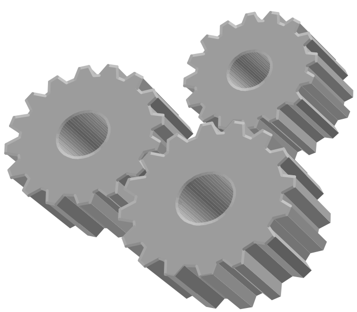


Traffic Demands

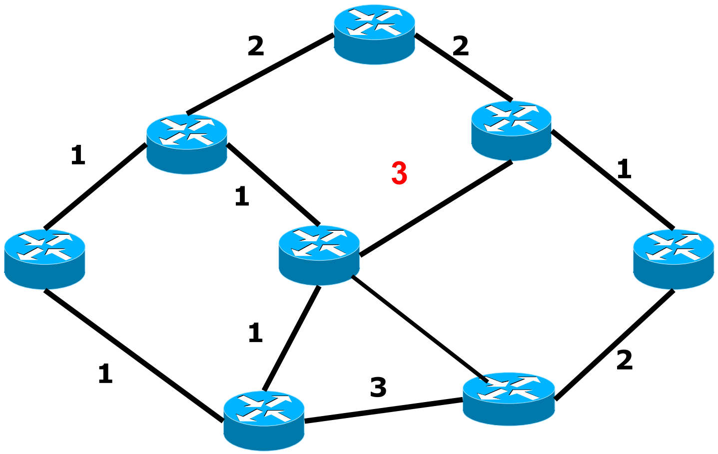
Network



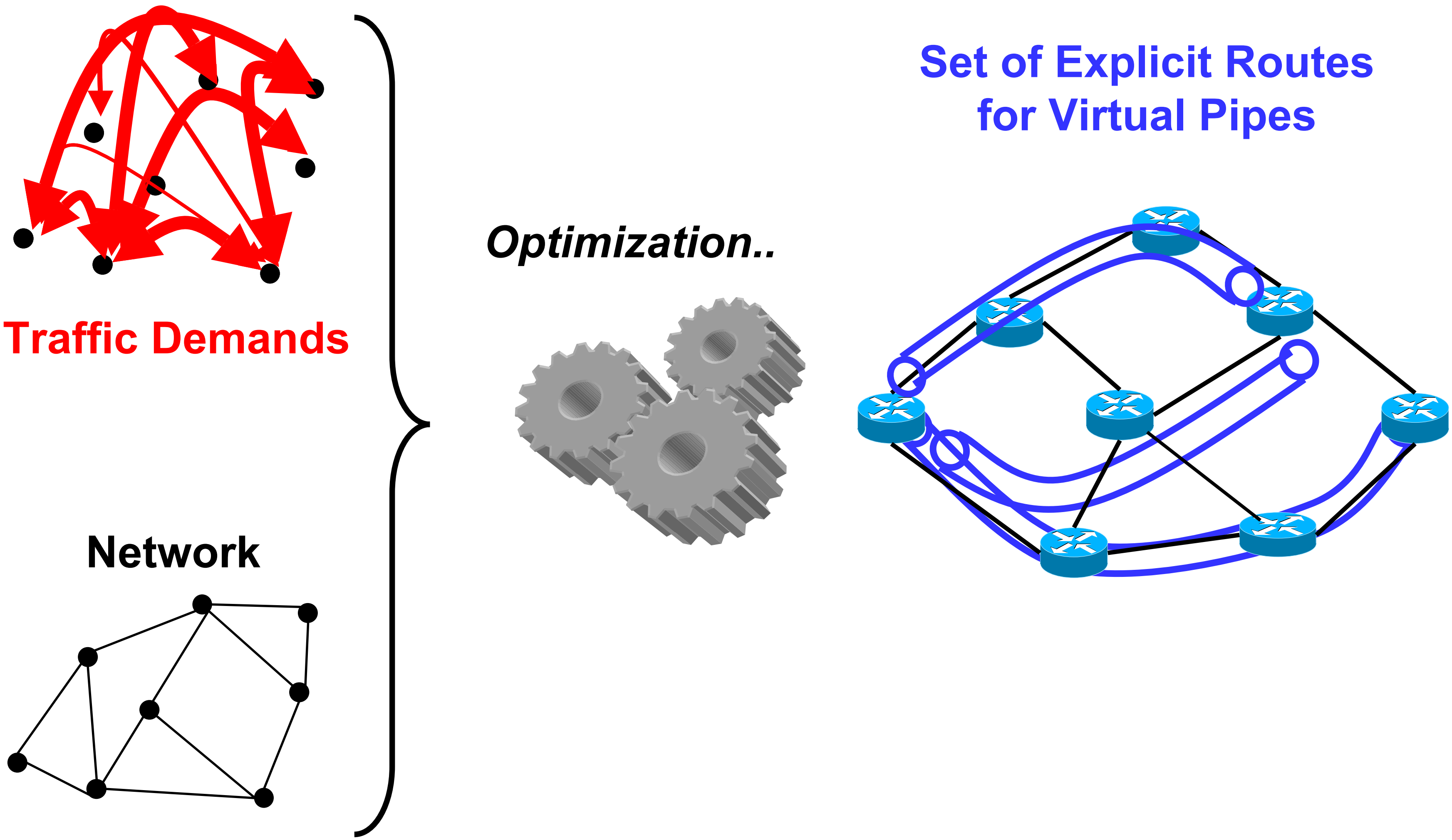
Optimization..



Set of Link Weights



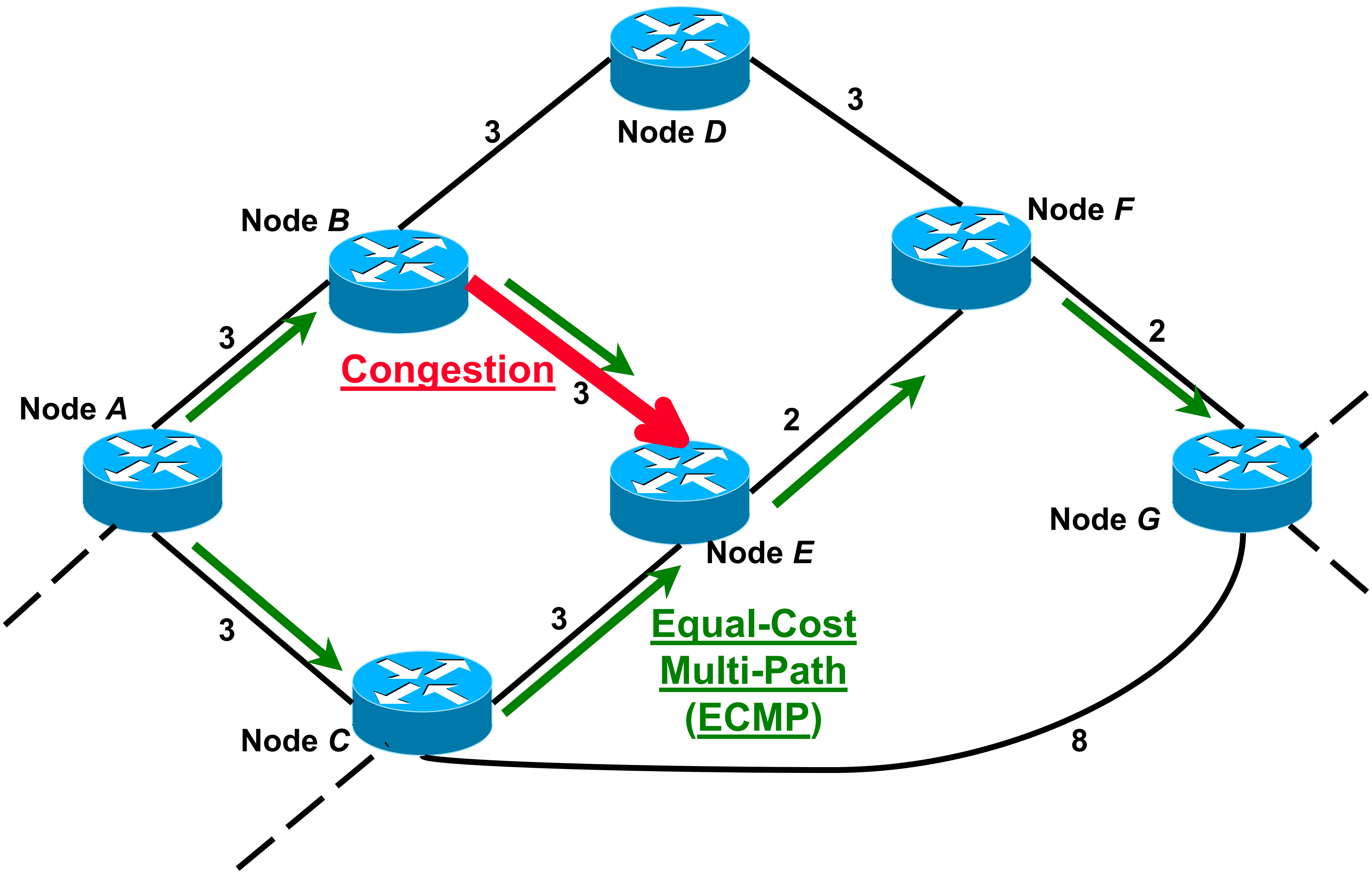
Example of Connection-Oriented TE: Explicit-Routing Optimization



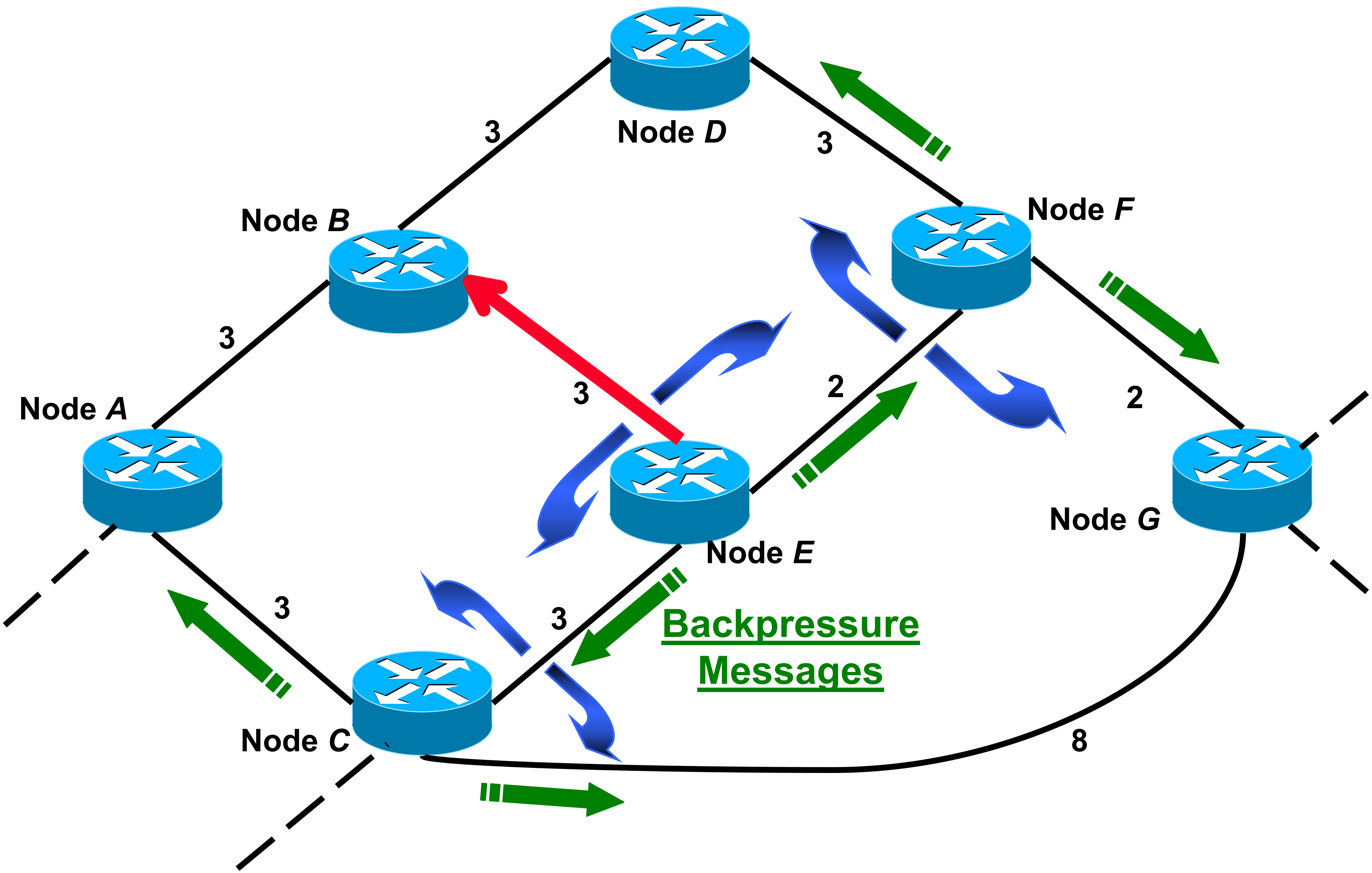
Traffic Engineering in IP Networks

- Existing traffic engineering methods have important disadvantages:
    - MPLS and *link weight optimization* require additional network management
    - Unpredictable signaling overhead with Optimized Multi-Path (OMP)
- Our objective:
    - Autonomous and continuous load distribution in the network
    - Low overhead in terms of memory and bandwidth consumption
- Proposal: Adaptive Multi-Path Algorithm (AMP)

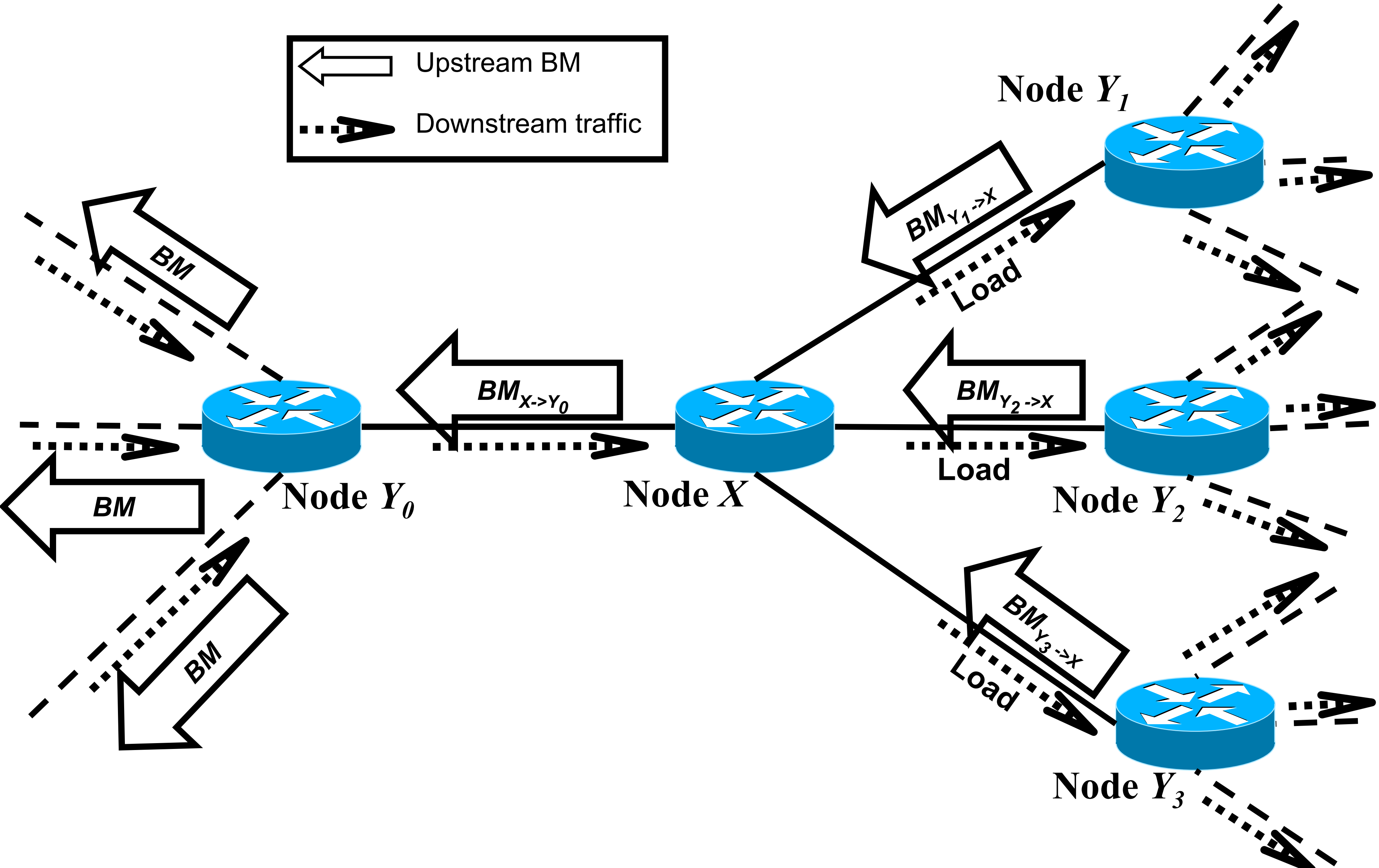
Current IP Routing



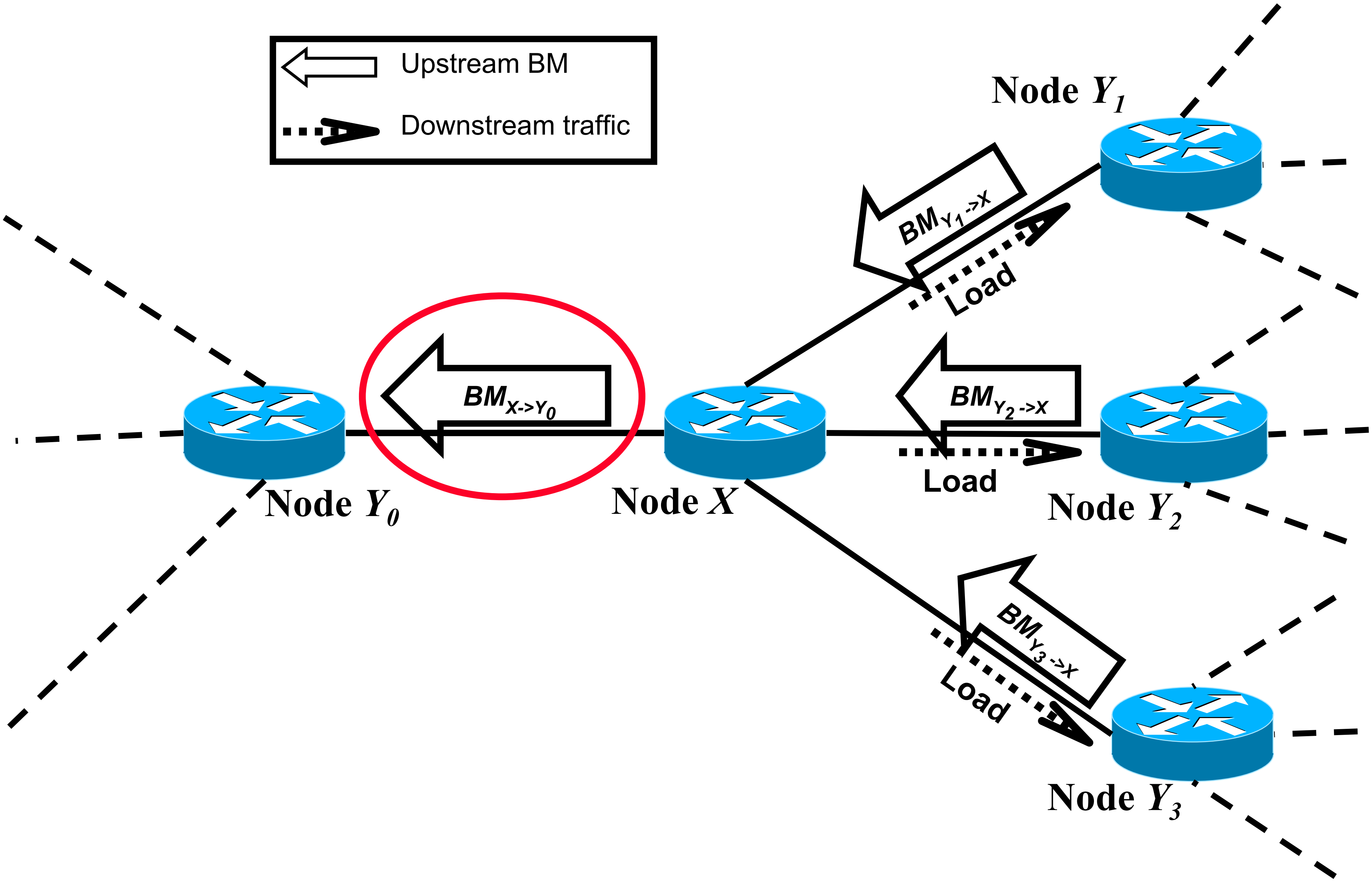
AMP – Basic Operation



AMP – Signaling



AMP – Signaling



AMP – Signaling

$$BM_{X->Y_0} = f(Load_{\overline{XY_1}}, \ldots, Load_{\overline{XY_n}}, BM_{Y_1->X}, \ldots, BM_{Y_n->X})$$

Quasi-recursive structure of backpressure messages

⇨

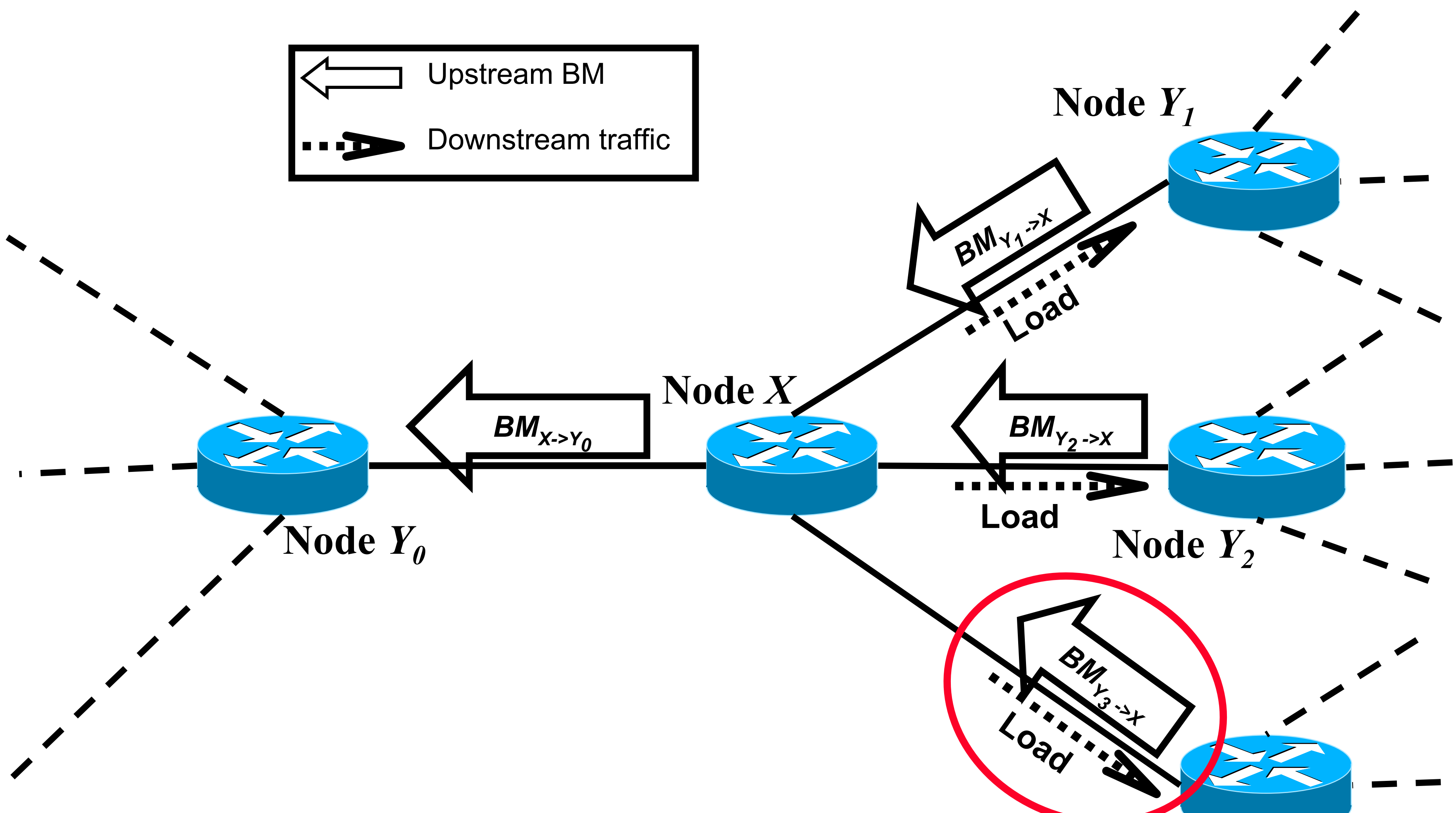
GLOBAL PROPAGATION OF LOAD INFORMATION THROUGH LOCAL EXCHANGE OF SIGNALING MESSAGES

AMP – Signaling

$$BM_{X->Y_0} = f(Load_{\overline{XY_1}}, \ldots, Load_{\overline{XY_n}}, BM_{Y_1->X}, \ldots, BM_{Y_n->X})$$

Summarization of the number of parameters

AMP – Signaling



One parameter per link: $g_i = \max(Load_{\overline{XY_i}}, BM_{Y_i->X})$

AMP – Signaling

$$BM_{X->Y_0} = f(Load_{\overline{XY_1}}, \ldots, Load_{\overline{XY_n}}, BM_{Y_1->X}, \ldots, BM_{Y_n->X})$$

⇓ Reduction of the number of parameters

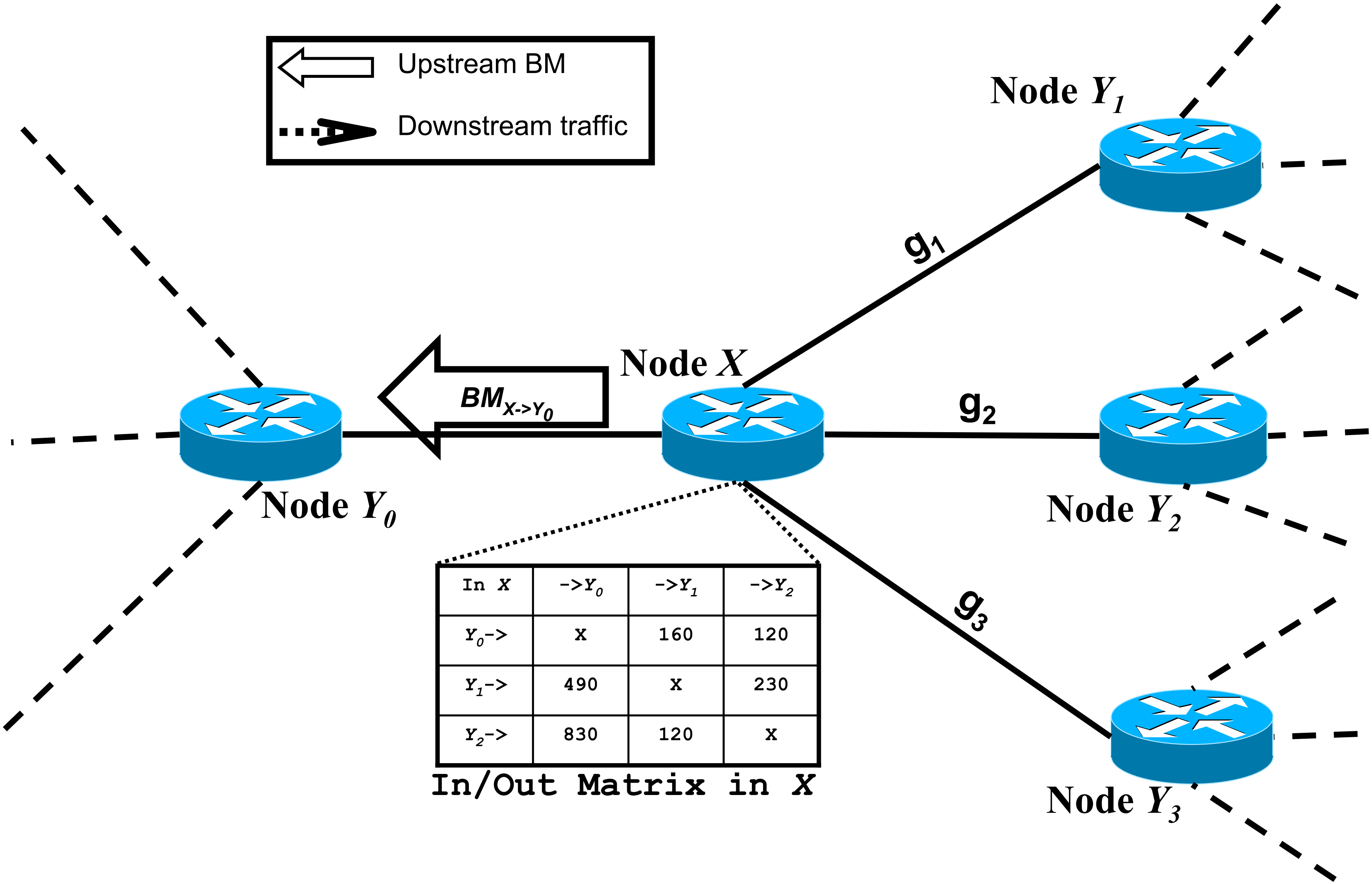
$$g_i = \max(Load_{\overline{XY_i}}, BM_{Y_i->X})$$

⇓

$$BM_{X->Y_0} = f(g_1, g_2, \ldots, g_n)$$

?

AMP – Signaling



In X	-> Y_0	-> Y_1	-> Y_2
Y_0 ->	X	160	120
Y_1 ->	490	X	230
Y_2 ->	830	120	X

In/Out Matrix in X

AMP – Signaling

$$BM_{X->Y_0} = f(Load_{\overline{XY_1}}, \ldots, Load_{\overline{XY_n}}, BM_{Y_1->X}, \ldots, BM_{Y_n->X})$$

⇓ Reduction of the number of parameters

$$g_i = \max(Load_{\overline{XY_i}}, BM_{Y_i->X})$$

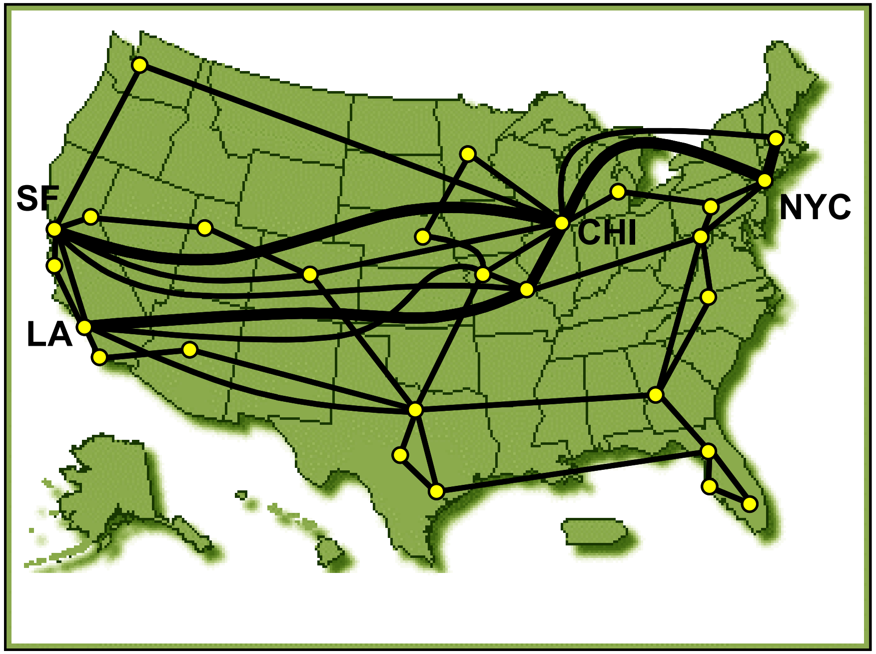
⇓

$$BM_{X->Y_0} = f(g_1, g_2, \ldots, g_n)$$

$$= \sum_{Y_i \in \Omega_X \setminus Y_0} \frac{\beta_{\overline{Y_0 X Y_i}}}{\beta_{\overline{X Y_i}}} \cdot g_i$$

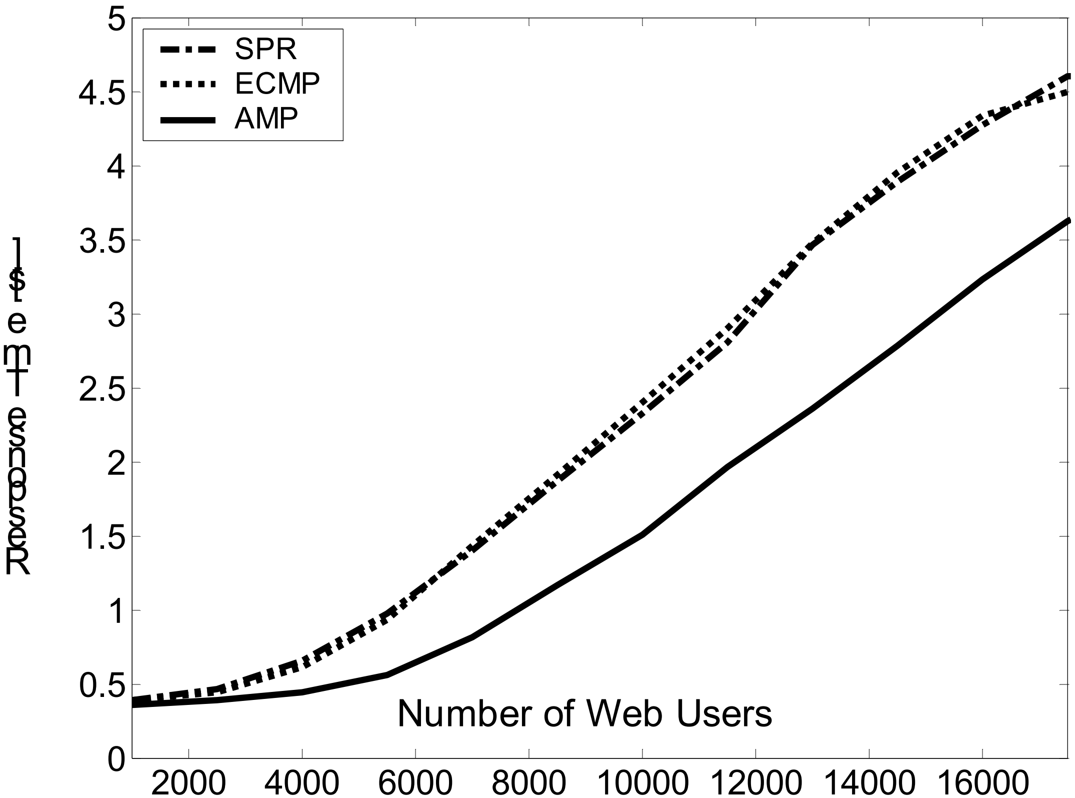
weights for congestion contributions

AMP Performance Evaluation



- Implementation of AMP in Network Simulator (ns-2)
- Simulated topology:
  - AT&T-US Network of 27 nodes and 47 links
  - Link capacities of 2.4 and 9.6 Gbit/s (scaled down to 15 and 60 Mbit/s in our simulations)
- Simulated traffic:
  - Web traffic according SURGE model
  - Traffic distribution according to the gravity model
  - Linear scaling of the number of Web users

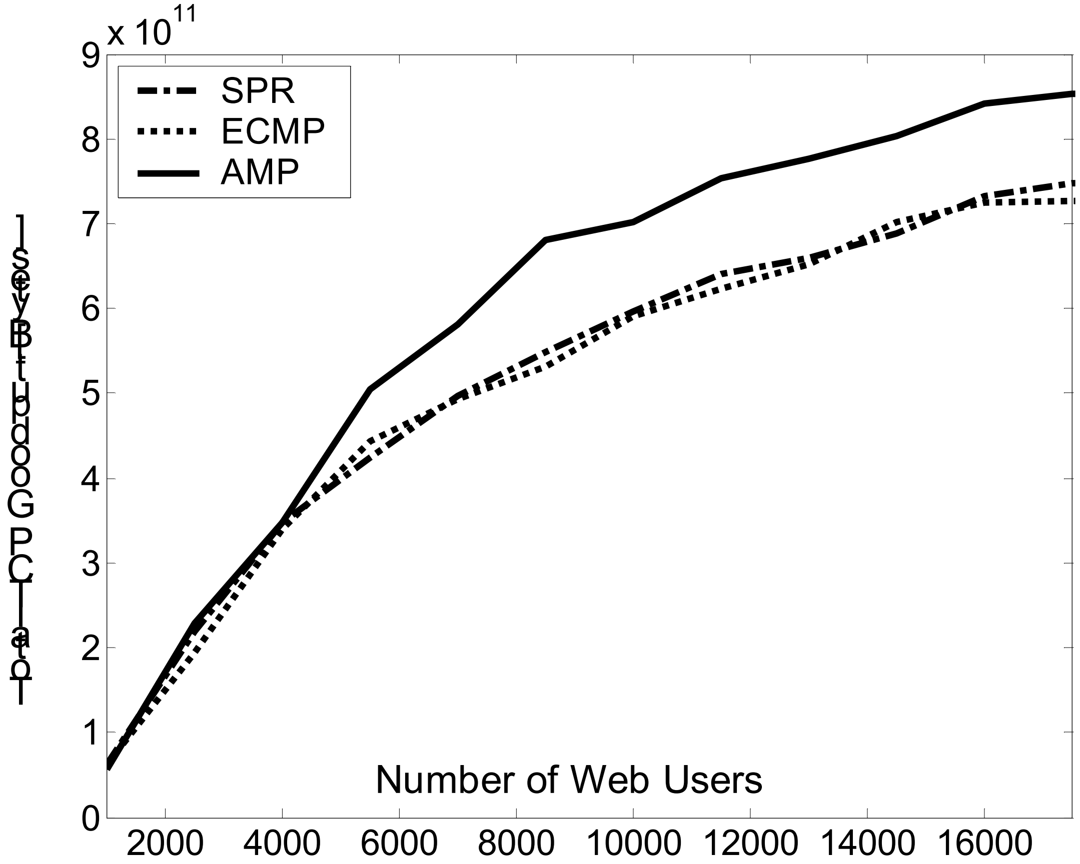
AMP Performance Evaluation – Average Web Page Response Time



- Web page response time most important metric from the user´s perspective
- Significant reductions in Web page response times throughout investigated scenarios (up to 43%)

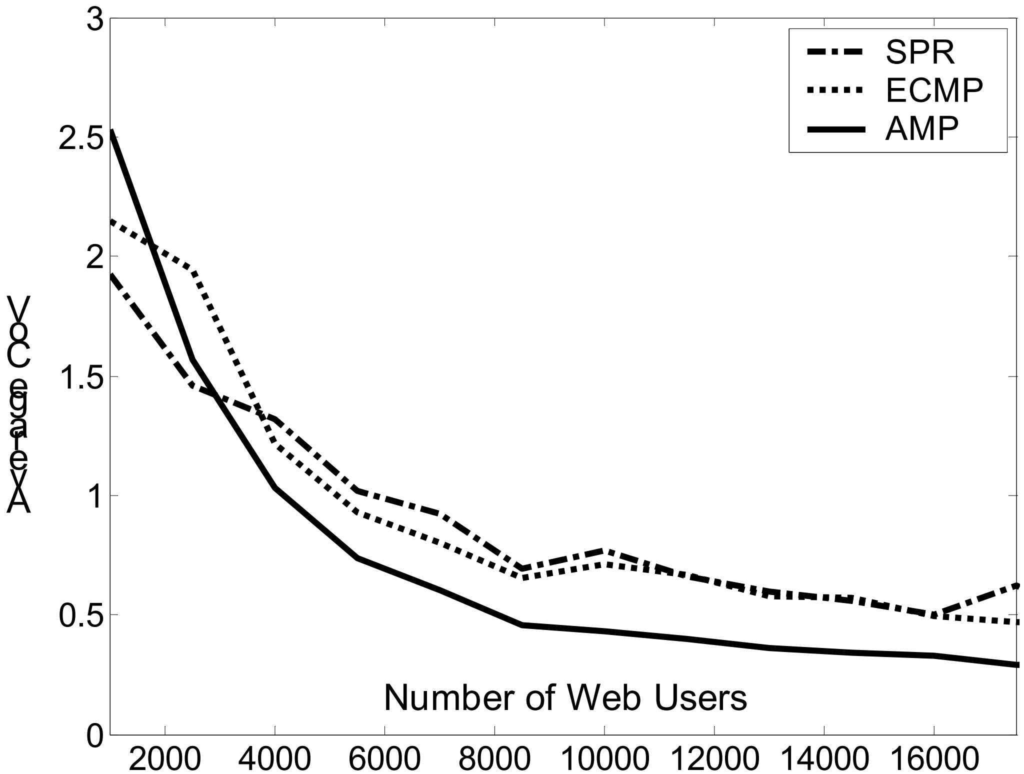
- SPR– Shortest Path Routing
- ECMP – Equal-Cost Multi-Path Routing

AMP Performance Evaluation – Total TCP Goodput



- Improved efficiency of resource utilization
- Total TCP goodput consistently higher with AMP compared to SPR and ECMP in our simulations (improvements of up to 28%)

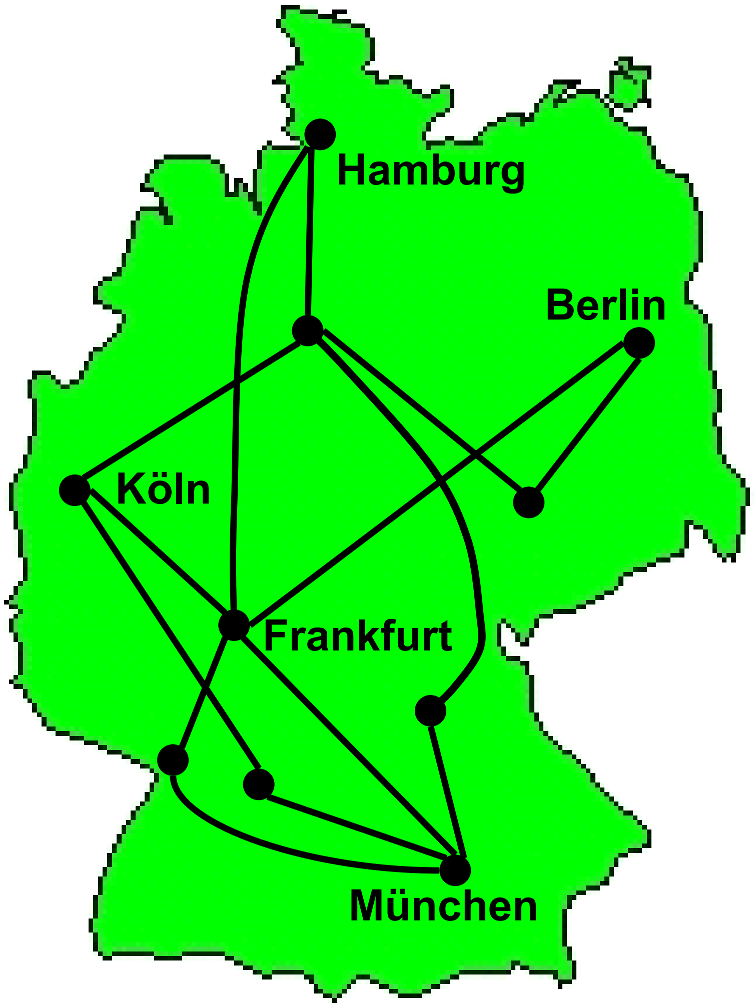
AMP Performance Evaluation – Average CoVs of Link Load



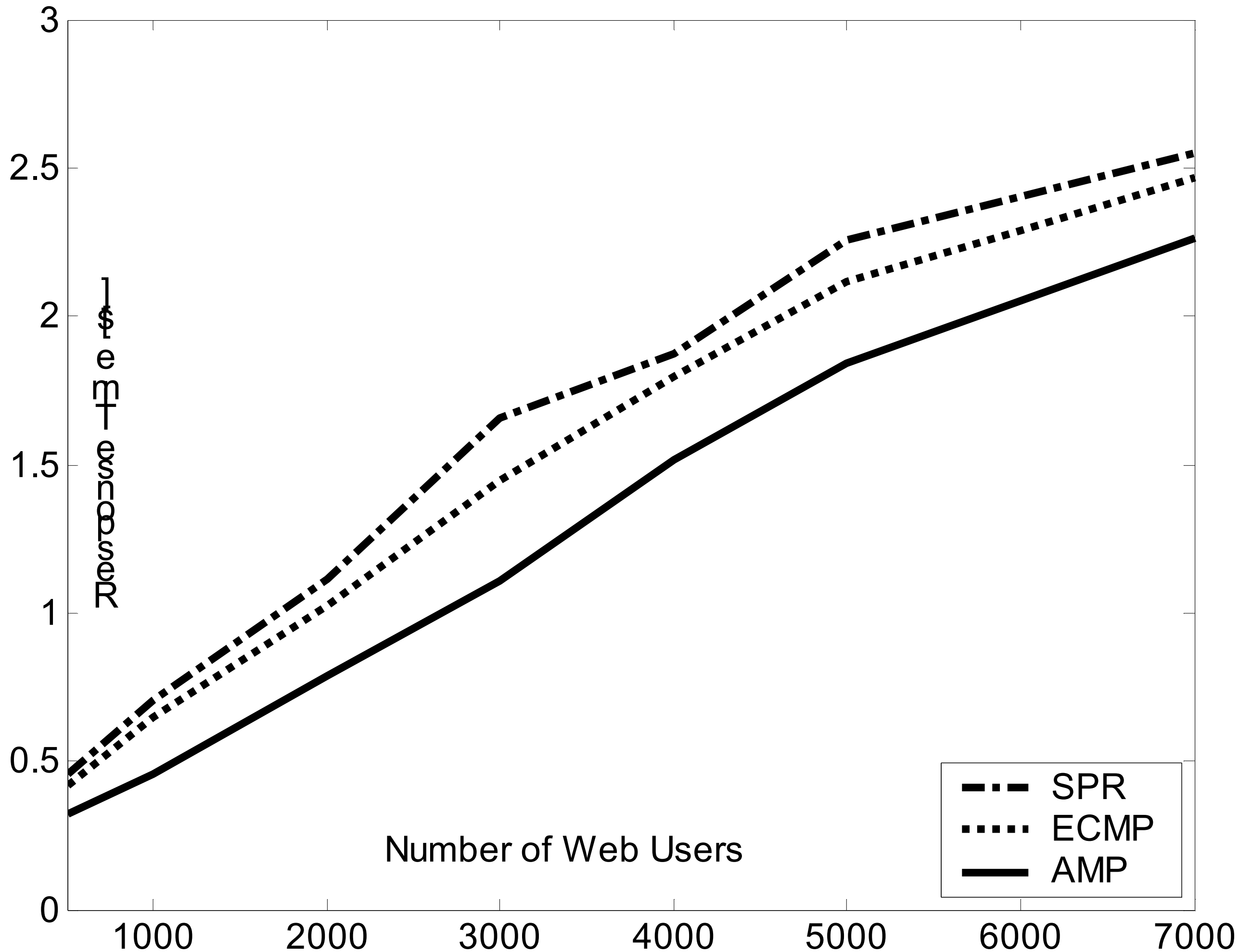
- Similar average Coefficient of Variation (CoVs) of all link loads for the three routing strategies

$\Rightarrow$ stability of AMP load balancing

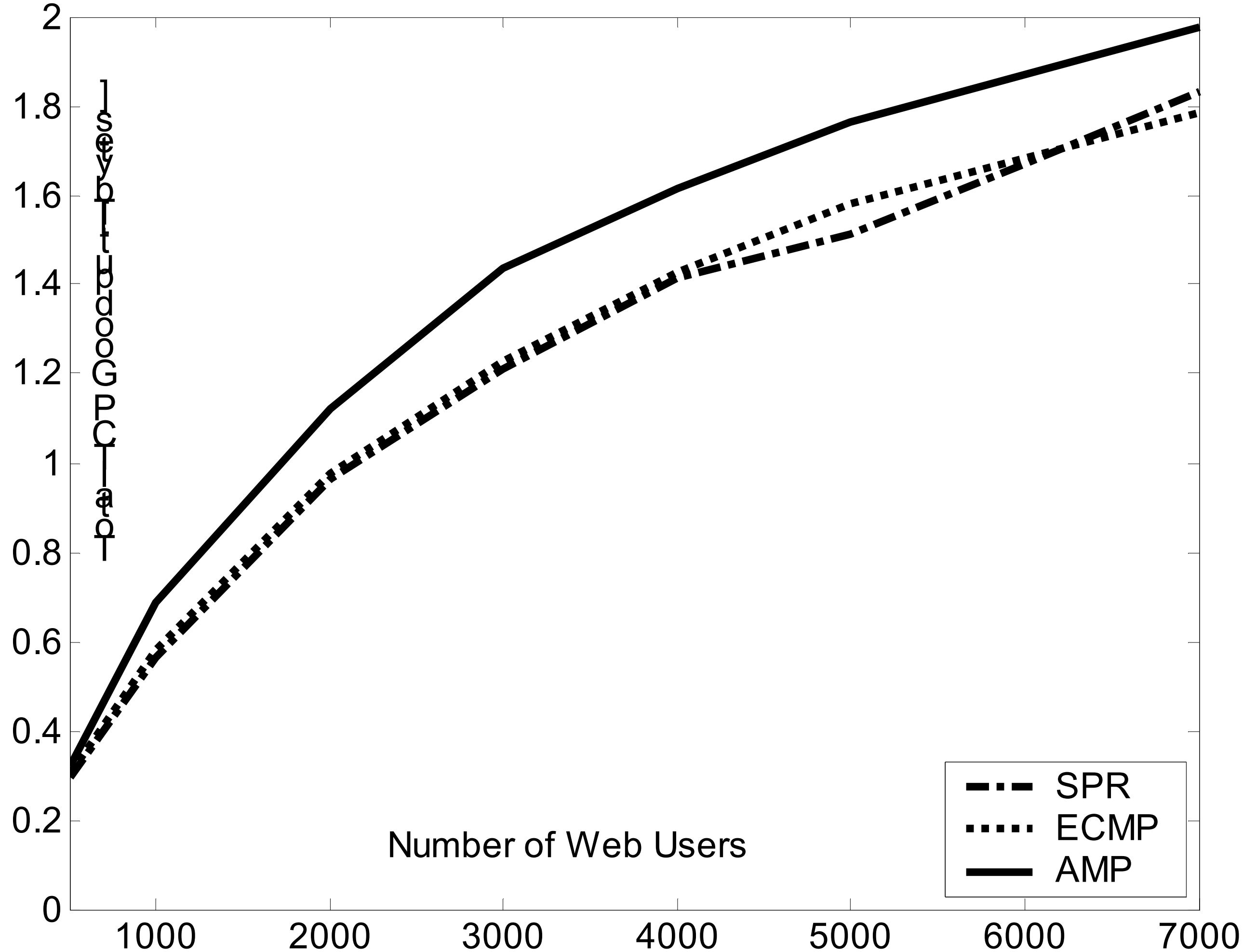
AMP Performance Evaluation



AMP Performance Evaluation – Average Web Page Response Time



AMP Performance Evaluation – Total TCP Goodput



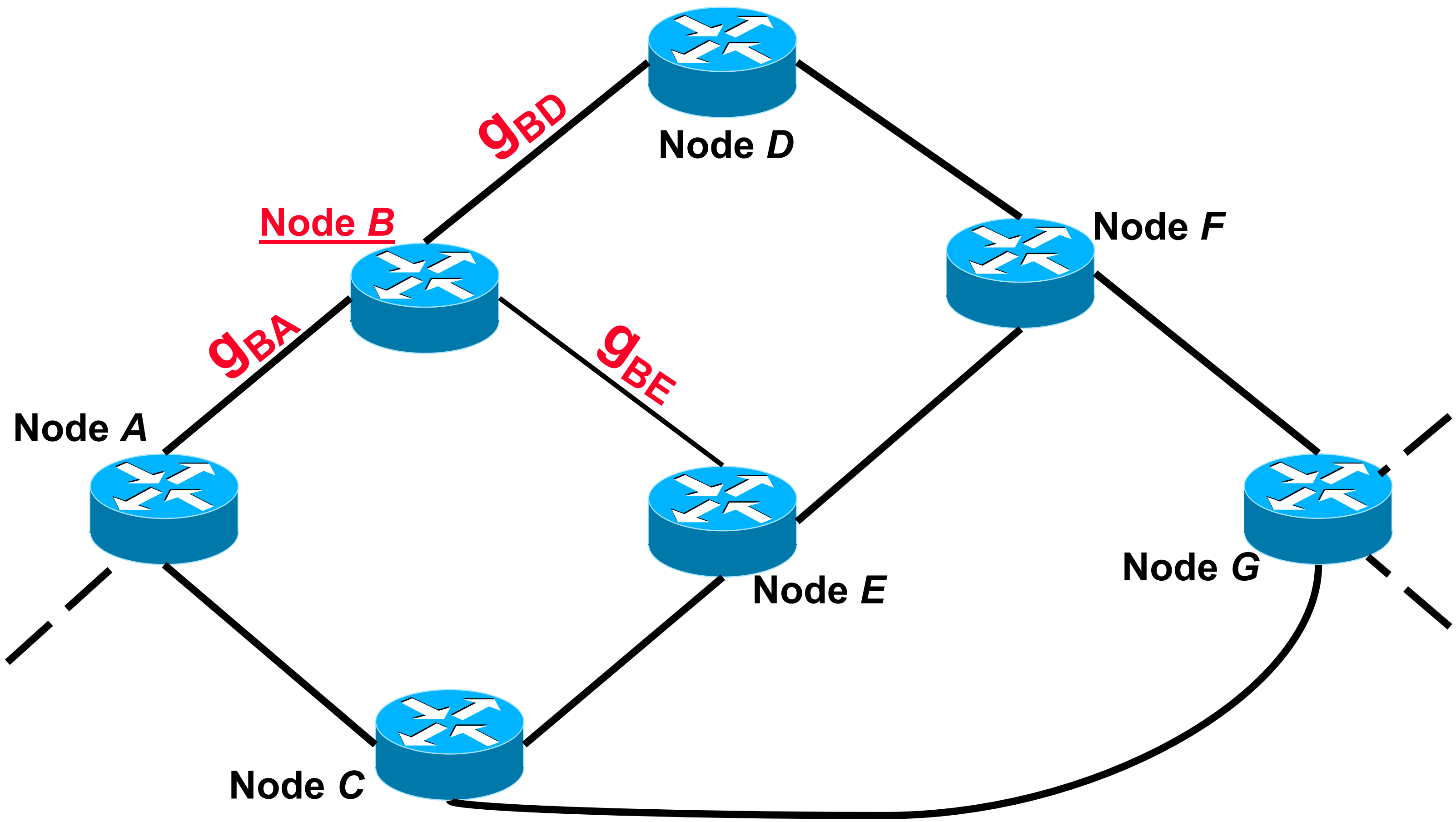
Summary & Outlook

- AMP Summary:
  - Load balancing within the framework of routing
  - No management overhead, minimal signaling overhead
  - Implementation in Network Simulator (ns-2)
  - Significant performance improvements
- Future research:
  - AMP and network resilience
  - AMP fluid simulation

Thank you for your attention!

gojmerac@ftw.at

AMP – Load Balancing



- The goal of the load balancing mechanism in every node is to equalize the values of *g* on all output links

AMP – Load Balancing

- In order to avoid packet disordering:
  - => the unit for load balancing is a microflow aggregate
  - => packets are assigned to an aggregate by applying a *CRC-16* hash-function on their source and destination IP addresses
- The *CRC-16* solution space [0, 65535] is divided among the viable next hops

161.53.101.8
173.42.78.55

CRC-16 → 13217

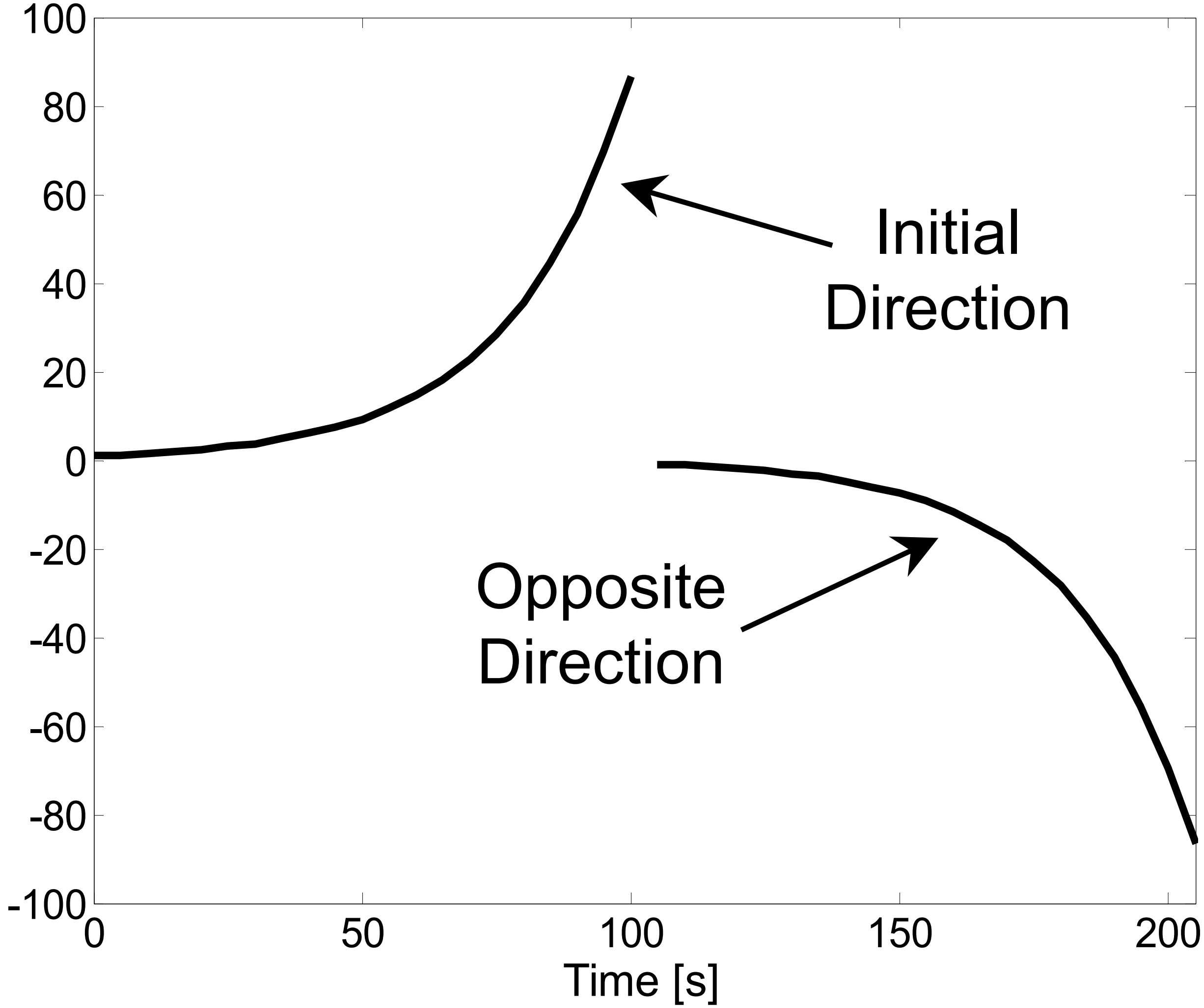
AMP – Load Balancing

- Example routing table in Node *B* – the hash-space boundaries are defined for every reachable destination

Destinations (in Node B)	Next hop: Node *A*	Next hop: Node *D*	Next hop: Node *E*
Node *A*	[0 – 65535] (ALL PACKETS)		
Node *C*	[0 – 23723]		[23724 – 65535]
Node *D*		[0 – 65535] (ALL PACKETS)	
Node *E*			[0 – 65535] (ALL PACKETS)
Node *F*		[0 – 34447]	[34448 – 65535]
Node *G*		[0 – 52142]	[52143 – 65535]

AMP – Load Balancing

- Conservative load balancing mechanism –
  the size of load adjustment steps is changed dynamically



Publications

- I. Gojmerac, T. Ziegler, P. Reichl: *Adaptive Multipath Routing Based on Local Distribution of Link Load Information*. Proc. QoFIS'03, Stockholm, October 2003.
- I. Gojmerac, T. Ziegler, F. Ricciato, P. Reichl: *Adaptive Multipath Routing for Dynamic Traffic Engineering*. Proc. GLOBECOM'03, San Francisco, November 2003.