Traffic Engineering and Network Design

Anwar Walid
with: D. Mitra, N. Nithi, I. Saniee, Q. Wang, I.Widjaja

Lucent Technologies

NMS PI Meeting, Miami, Jan. 21, 2004.



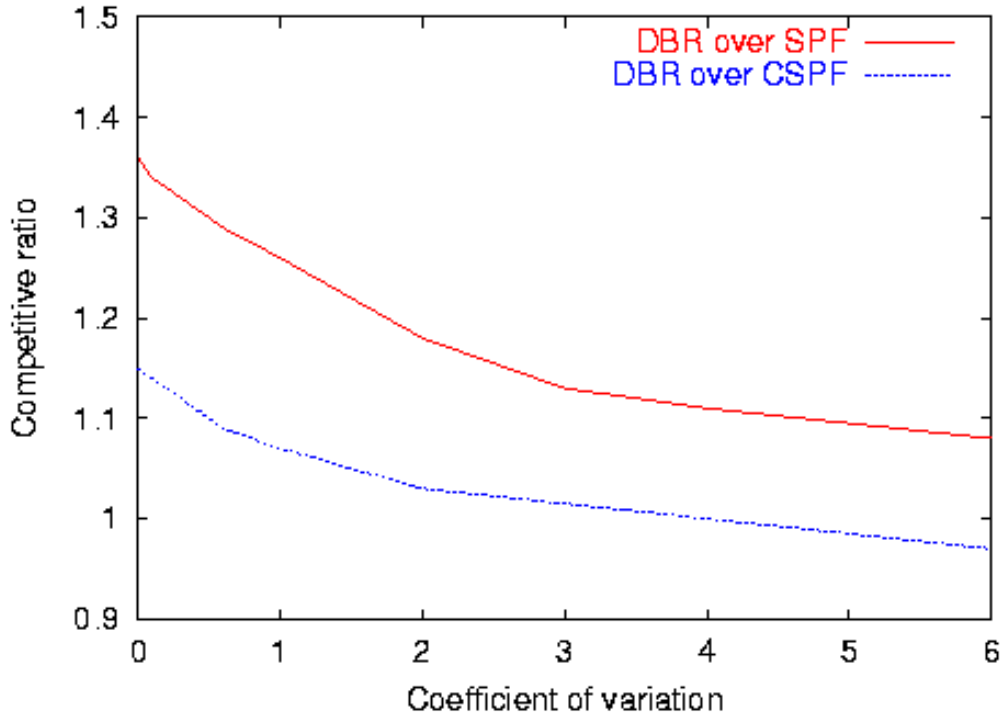
Lucent Technologies
Bell Labs Innovations

Outline

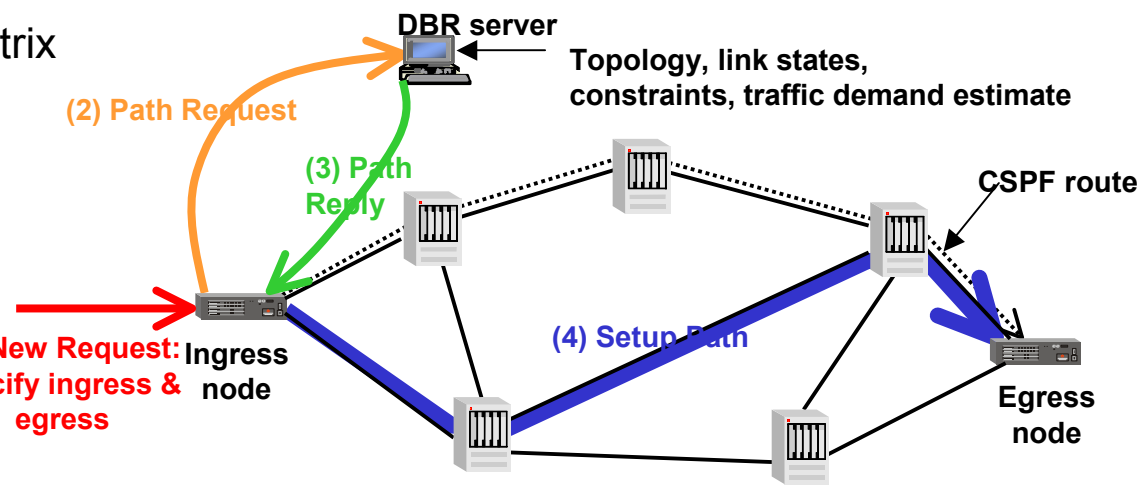
- **Design-Based Routing (DBR): pre-design optimization to assist online routing and protection**
- **Adaptive Multi-Path Load balancing**
- **Multi-Layer Optimization of Routing and Resource Provisioning in Overlay Networks**

DBR Routing Efficiency Compared to SPF & CSPF

- **Online TE with CSPF*:**
  - Connection requests arrive and are routed one-by-one
  - Each path is computed individually using SP with adaptive weights
  - Performance can be far from optimal



- **Online TE with DBR:**
  - Connection requests arrive and are routed one-by-one
  - Online routing assisted with pre-design optimization
  - Approximate traffic demand matrix used
  - Performance is close to optimal

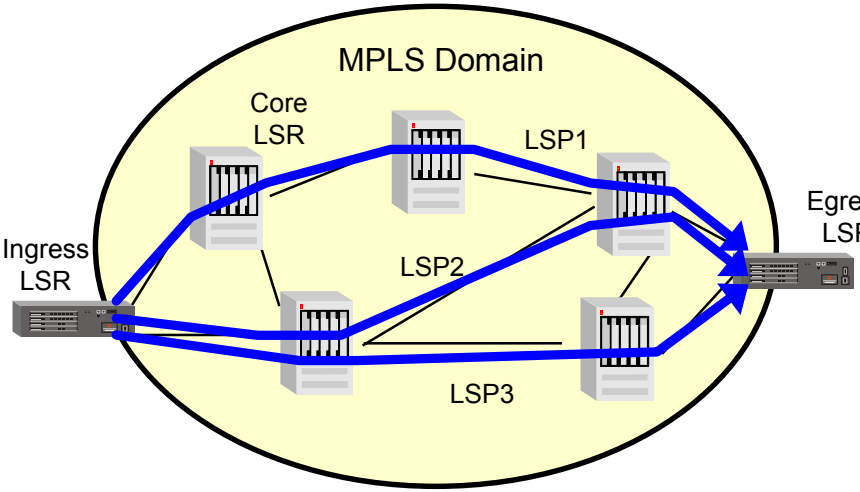


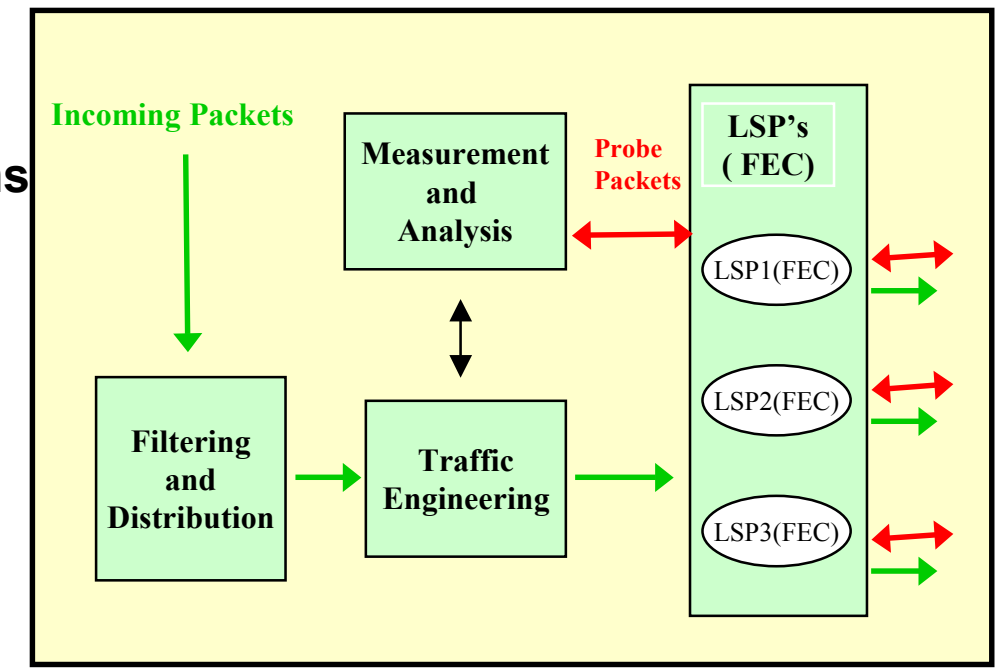
- **A Realization of DBR**

•A paper appeared in J. Lightwave Networks, 2003.

Adaptive Multi-path Load Balancing

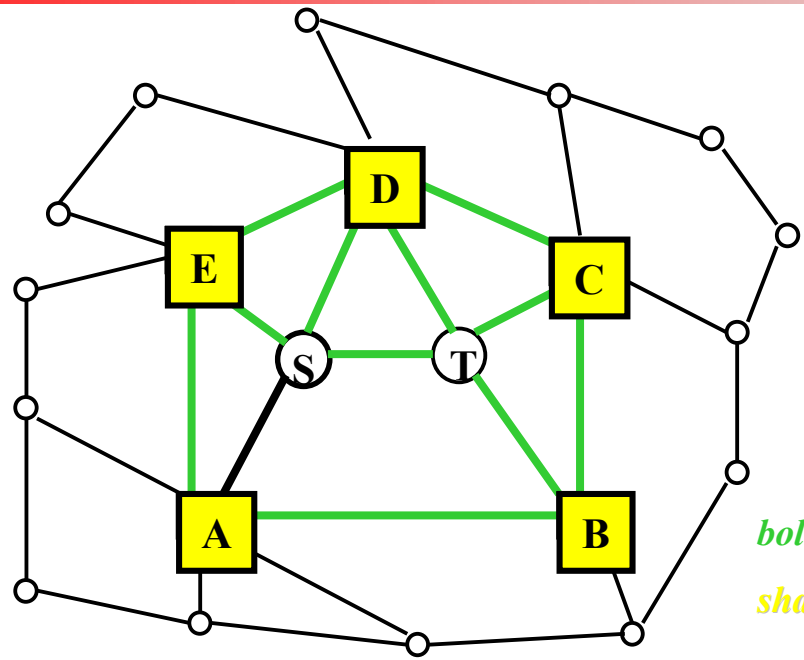
- Features
  - –Online, automated enhancement of throughput in MPLS-based IP networks
  - –Ability to react to dynamic and unpredictable traffic patterns
  - –Load balancing using real time probes and statistical analysis to measure one-way path metrics
  - –Distributed optimization algorithms
  - –End-to-end control with no new requirements at intermediate Routers
  - –Minimal additional packet re-sequencing needed
- Status
  - –Building Network simulation running at 10G





MATE Functions in Ingress LSR

Overlay Networks



bold line: *Core (bottom layer) network links*

shaded box: *Interface between top and bottom layer*

- **Multi-layer networks, bottom layer provides resource to the top layer**
- **Each layer performs its own optimization with minimal information exchange**
- **Applications:**
  - **IP over ATM;**
  - **IP/MPLS over Optical;**
  - **Wireless**

Global Optimization of Entire Network

Formulate the Objective for Optimal Routing:

Extended Concurrent Flow Problem

σ:	pair of end-nodes,
d_σ	total demand between σ
$R(\sigma)$	set of candidate routes between σ,
x_r	traffic carried on route r
$\sum_{r \in R(\sigma)} x_r / d_\sigma$	percentage of traffic carried for demand σ

Let $\delta = \min_{\sigma} \left(\sum_{r \in R(\sigma)} x_r / d_\sigma \right)$

The routing is chosen to maximize:

$$a\delta + \sum_{\sigma} b_\sigma \sum_{r \in R(\sigma)} x_r$$

max-min fairness

ensure a large percentage of traffic is carried for every demand

efficiency

ensure available capacity is always used to carry valuable demand

Let $a >> b_\sigma d_\sigma \ \forall \sigma$ so max-min fairness is of the dominant importance

Global Optimization of Entire Network

$$\max_{\delta,\vec{x},\vec{y}\geq 0} \{ \underbrace{[a\delta + \sum_{\sigma} b_{\sigma} \sum_{r\in R(\sigma)} x_r]}_{\text{optimize routing objective}} - \underbrace{\sum_{l} k_l y_l}_{\text{minimize design cost}} \}$$

l: link index

k_l: unit cost, y_l: capacity (integer)

$$\underbrace{\sum_{r\in R(\sigma):l\in r} x_r \leq C_l}$$

IP link capacity constraint,

C_l: capacity on top layer links

$$\underbrace{\sum_{r\in R(\sigma):l\in r} x_r \leq y_l}$$

Optical link capacity constraint

y_l: capacity on bottom layer links

$$\underbrace{\sum_{r\in R(\sigma)} x_r \geq \delta d_\sigma \qquad 0 < \delta \leq 1}$$

definition and bound of variable δ

Need for Decomposition to Reduce Complexity

- **Complexity of the global optimization**
  - the problem is NP hard
  - may have large number of variables and constraints, at least $O(N^2)$,
  - requires complete knowledge of all layers -- centralized
- **Scalability via decomposition**
  - reduce the overall problem to a set of smaller sub-problems
  - given routing, optimized design is an instance of Knapsack problem
    - no exact solution, but fast efficient heuristic exist
  - given design, optimized routing is an instance of maximum concurrent flow problem
    - solvable by many methods
- **Global optimal can be achieved through iteration between sub-problems**
  - keep information flow between sub-problems parsimonious

Bender's Decomposition

The original problem can be rewritten as

$$\max_{\vec{y} \geq 0} \quad \varphi(\vec{y}) \equiv \{\theta(\vec{y}) - \sum_{l} k_l y_l \mid \vec{y} \text{ integer} \quad \}$$

where

$$\theta(\vec{y}) = \max_{\vec{x},\delta} \rho(\vec{x}, \delta \mid \vec{y}) \equiv \{a\delta + \sum_{\sigma} b_\sigma \sum_{r \in R(\sigma)} x_r \mid$$

$$\sum_{r \in R(\sigma): l \in r} x_r \leq w y_l, \quad \sum_{r \in R(\sigma)} x_r \geq \delta d_\sigma, \quad \sum_{r \in R(\sigma): l \in r} x_r \leq C_l\}$$

The above divides the original optimization into sub-problems

- optimal design problem to maximize φ (by choosing min-cost design)
- optimal routing problem to maximize θ (carried demand)

- computing φ requires knowledge of $\theta(y)$
- optimizing θ requires information on lower layer link capacity (represented by y variables)

Lucent Technologies
Bell Labs Innovations

Decomposition and Topology Hiding

- Bottom layer constructs bandwidth pipes between interface points
- Top layer only sees fully meshed lower layer network, instead of internal structure

Design/configuration optimization at the bottom layer

$$\max_{\vec{y},\vec{z},\vec{f}\geq 0} \varphi(\vec{y},\vec{z},\vec{f}) \equiv \{\theta(\vec{z}) - \sum_{l} k_l y_l \mid \vec{y} \text{ integer } \}$$

$$\sum_{l\in l_n} f_{ijl} - \sum_{l\in h_n} f_{ijl} = \begin{cases} -z_{ij} & i=n \\ z_{ij} & j=n \\ 0 & i,\ j \neq n \end{cases}$$

$$\sum_{(i,j)} f_{ijl} \leq y_l$$

network configuration

create bandwidth pipes, z_{ij}, between interface points (i,j) based on link capacity y_l.
f_{ijl} is the flow between i,j carried on link l

Routing optimization at the top layer (use end-to-end capacity z_{ij} as inputs)

$$\theta(\vec{z}) = \max_{\vec{x},\delta} \rho(\vec{x},\delta \mid \vec{z}) \equiv \{a\delta + \sum_{\sigma} b_\sigma \sum_{r\in R(\sigma)} x_r \mid$$

$$\sum_{r\in R(\sigma):(i,j)\in r} x_r \leq z_{ij},\ \sum_{r\in R(\sigma)} x_r \leq \delta d_\sigma,\ \sum_{r\in R(\sigma):l\in r} x_r \leq C_l \}$$

Iteration and Feedback

top-layer routing optimization: max $\theta_s(x,\delta|z)$ for each scenario s

$$\theta_s \le \Phi_s + \sum_{ij} \lambda_{ijs} z_{ij}$$

linear bounds of θ_s

Φ_s and λ_{ijs} are outputs top-layer optimization

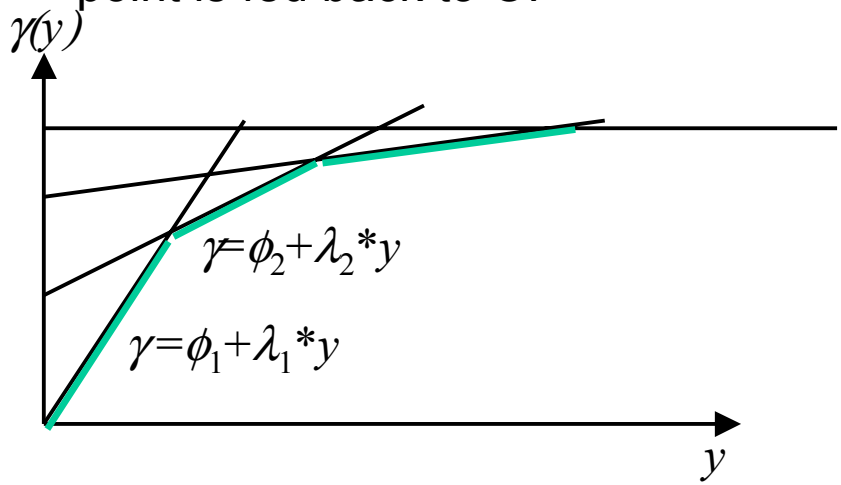
λ_{ijs} can be interpreted as willingness to pay for unit bandwidth between i, j

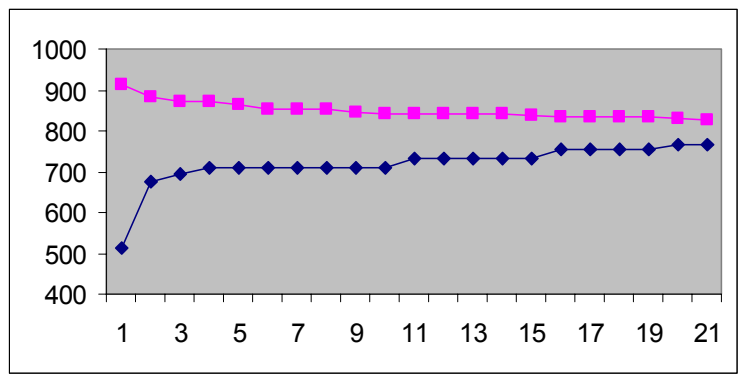
$z_{ij:}$ bandwidth pipes between i,j provisioned to the top layer

bottom-layer design/configuration optimization: max $\varphi(y,z,f)$

Lucent Technologies
Bell Labs Innovations

- IP passes to OP (ϕ_1, λ_1), (ϕ_2, λ_2), ….(ϕ_i, λ_i),……
- OP uses that information to construct envelop function $\gamma(y)$ for the value of capacity, and choose capacity y^* by maximizing $\gamma(y)$-$K(y)$ (where $K(y)$ is capacity cost)
- IP takes y^* as given, optimize routing and admission to generate another sample point, $(\phi_{i+1}, \lambda_{i+1})$, on the value curve $\gamma(y)$. The sample point is fed back to OP





Convergence behavior