

# Overcoming Failures:

## Fault-tolerance and Logical Centralization in Clean-Slate Network Management

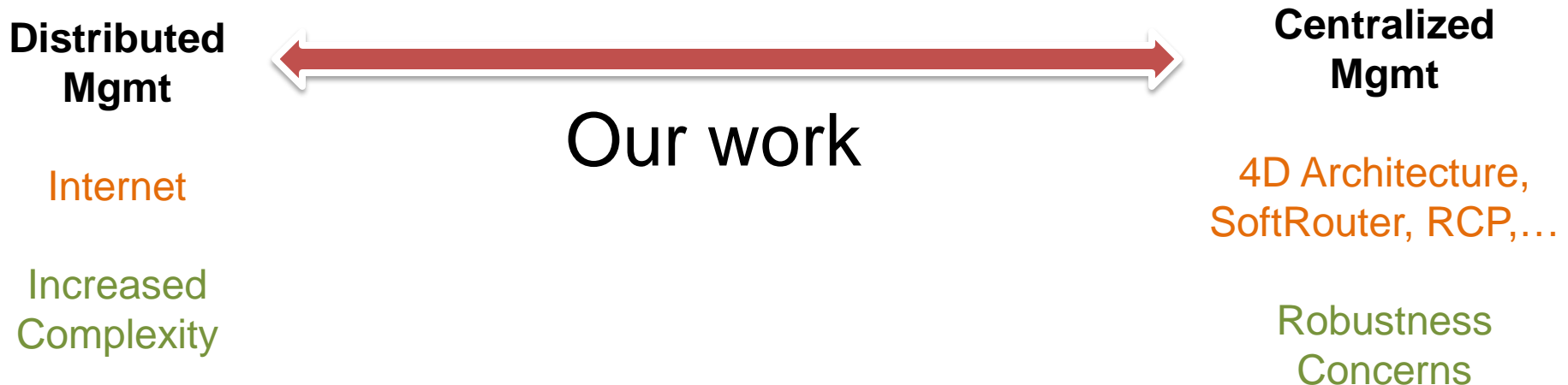
**Hammad Iqbal**  
**Taieb Znati**

**University of Pittsburgh, USA**

March 15, 2010

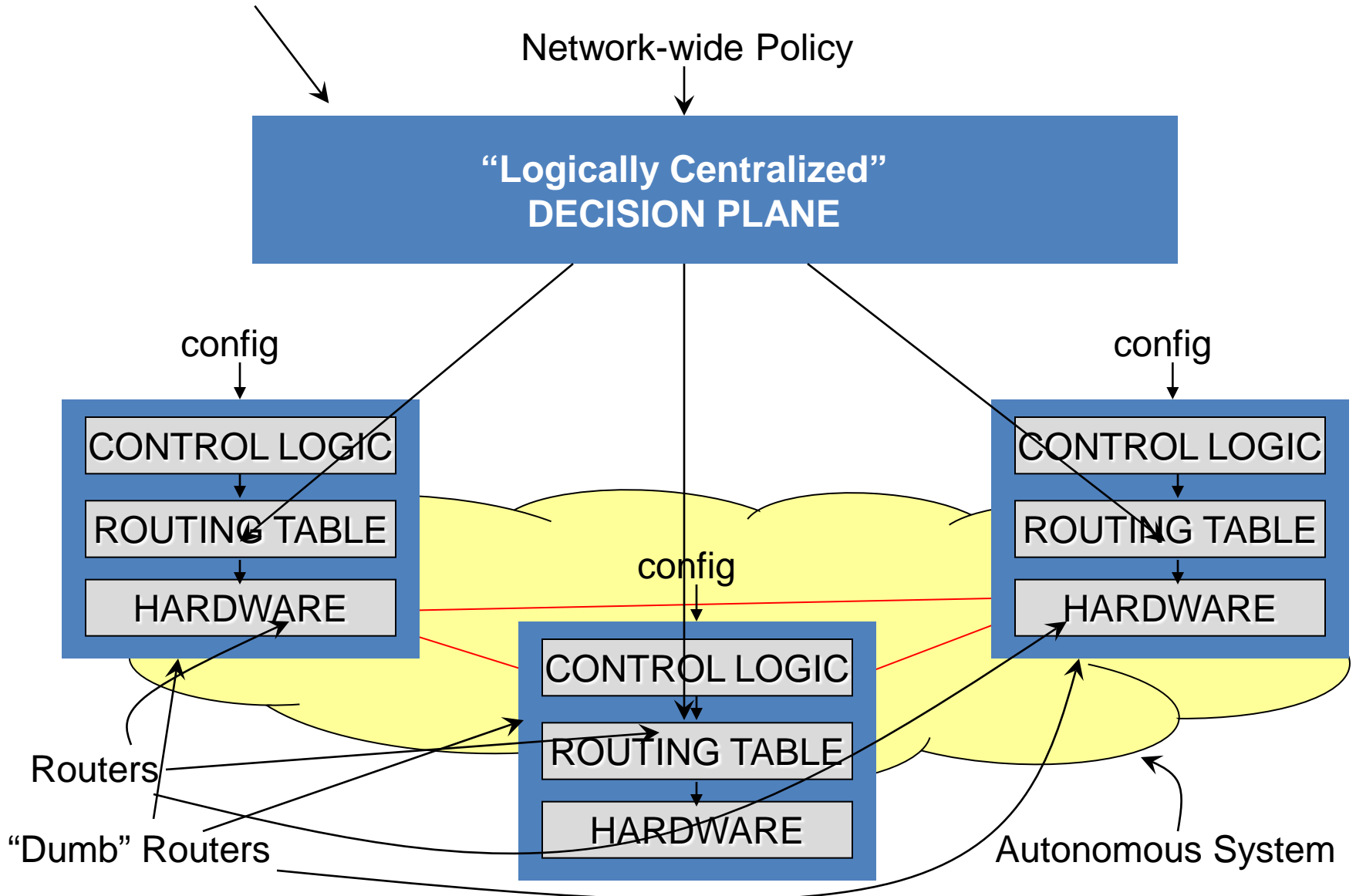
# Context

- Network Control and Management – A challenge for large IP networks
  - Increased likelihood of mis-configuration/errors
  - High O&M costs
  - ...
- Existing Work:

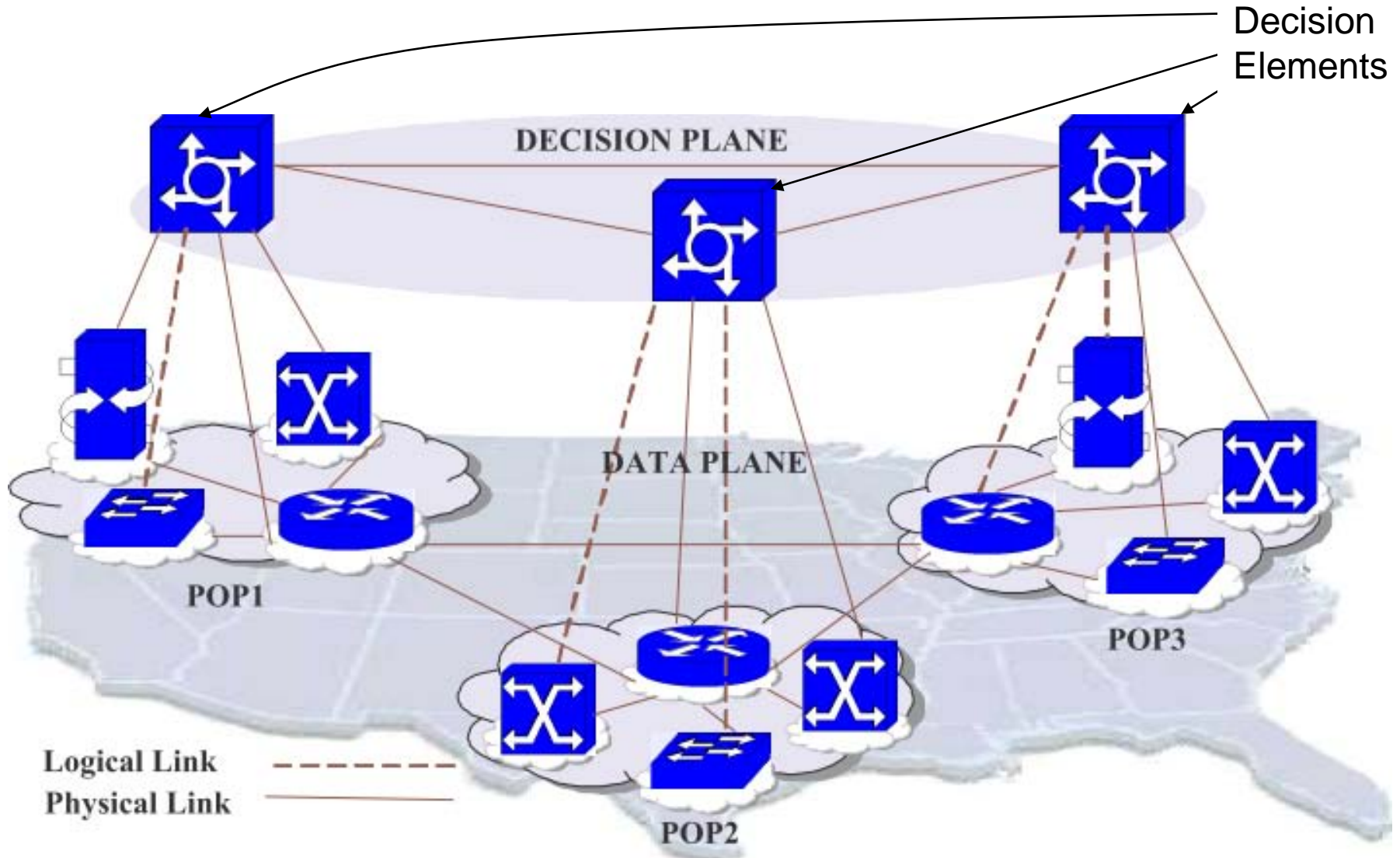


# Design Overview

Set of **Decision Elements (DE)**



# Network Example



# Fault Tolerance at Decision Plane

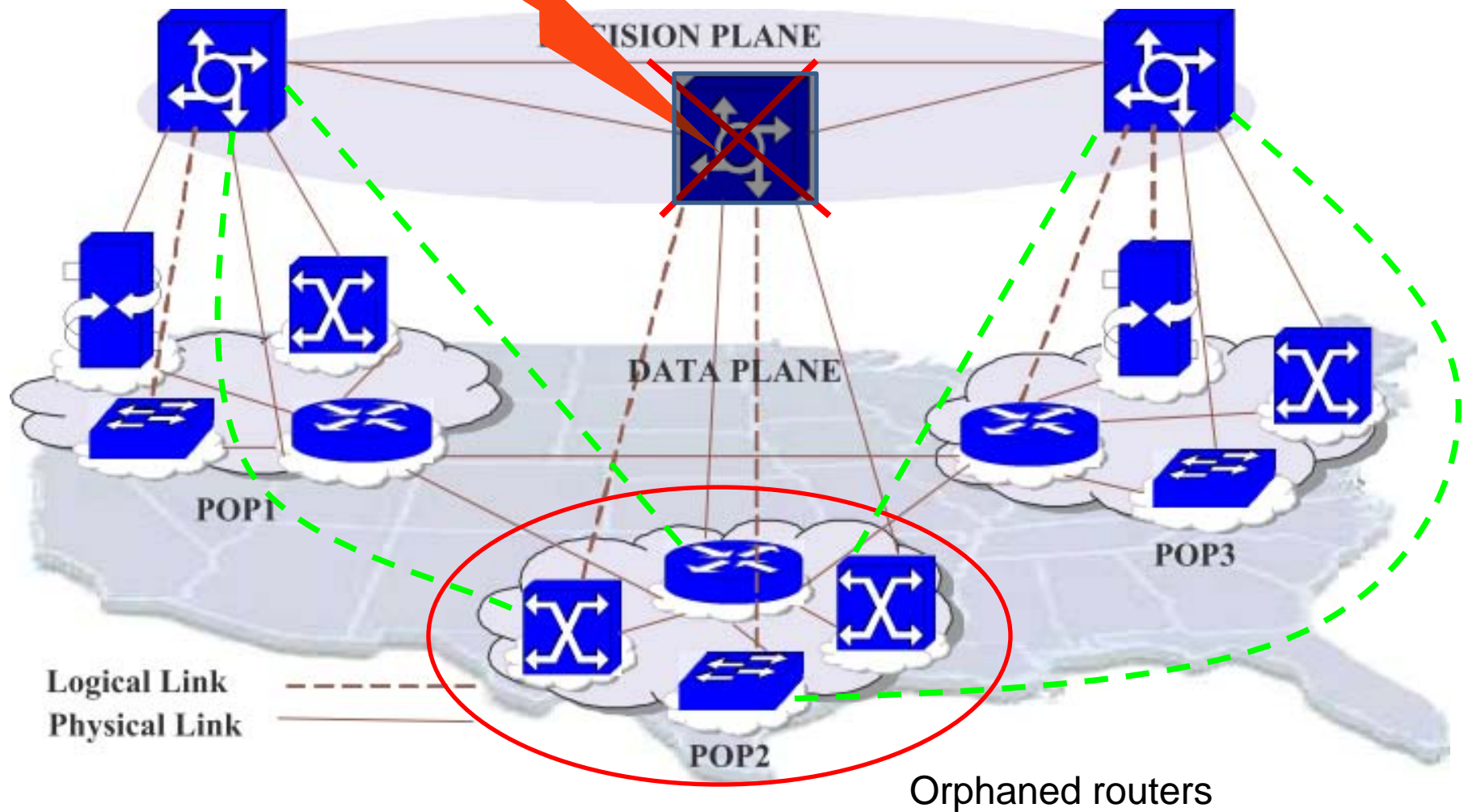
How to ensure continuity of network control with DE failures?

- Routers do not have any control logic
- A Decision Element (DE) failure will result in “orphaned” routers
  - Static routes

Router Assignment Schemes

- Static
- **Adaptive**

# Fault Tolerance at Decision Plane (cont.)



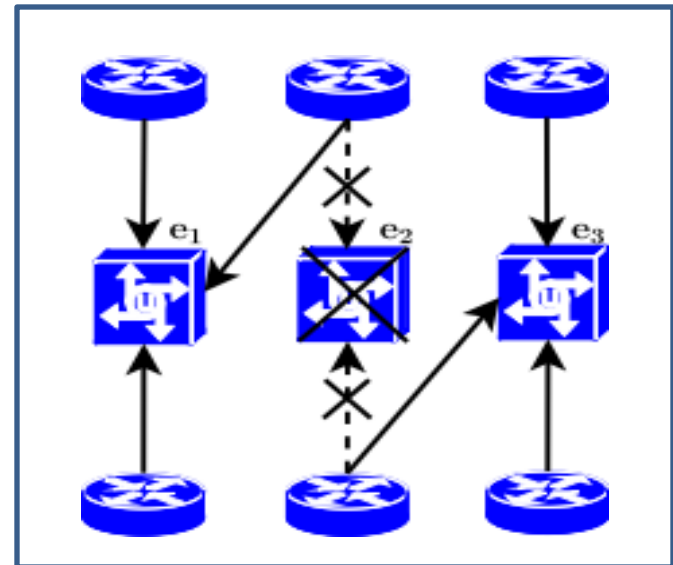
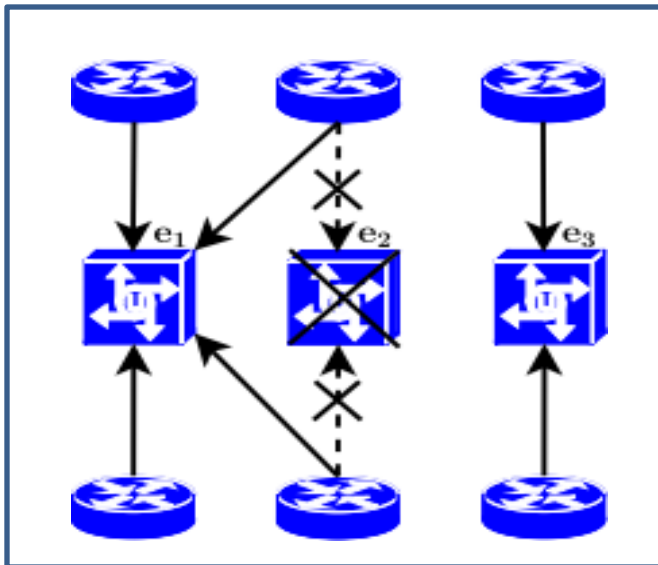
# Adaptive Router Assignment

Goal: Optimal assignment of routers to DEs

Minimize Decision plane response time

- Router → closest (min-delay) DE

Load balancing at Decision plane



# Adaptive Router Assignment Problem Definition

**Given:** Network topology, delay measurements, DE capacities

**Output:** Router assignments for all DEs

Minimize  $\sum_{e_j \in E} \sum_{r_i \in R} d(r_i, e_j) x(r_i, e_j)$   $\longleftarrow$  Aggregate router-DE delay minimization objective

s.t.

$$\sum_{e_j \in E} x(r_i, e_j) = 1$$

$$\sum_{r_i \in R} x(r_i, e_j) - Q_j \leq 0 \quad \longleftarrow \text{DE capacity constraint}$$

$$L_j \leq \lceil \Delta L_{avg} Q_j \rceil \quad \longleftarrow \text{Load balancing constraint}$$

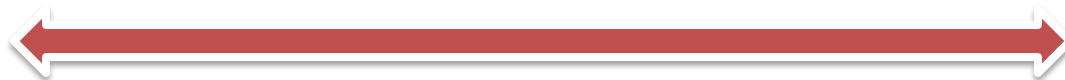
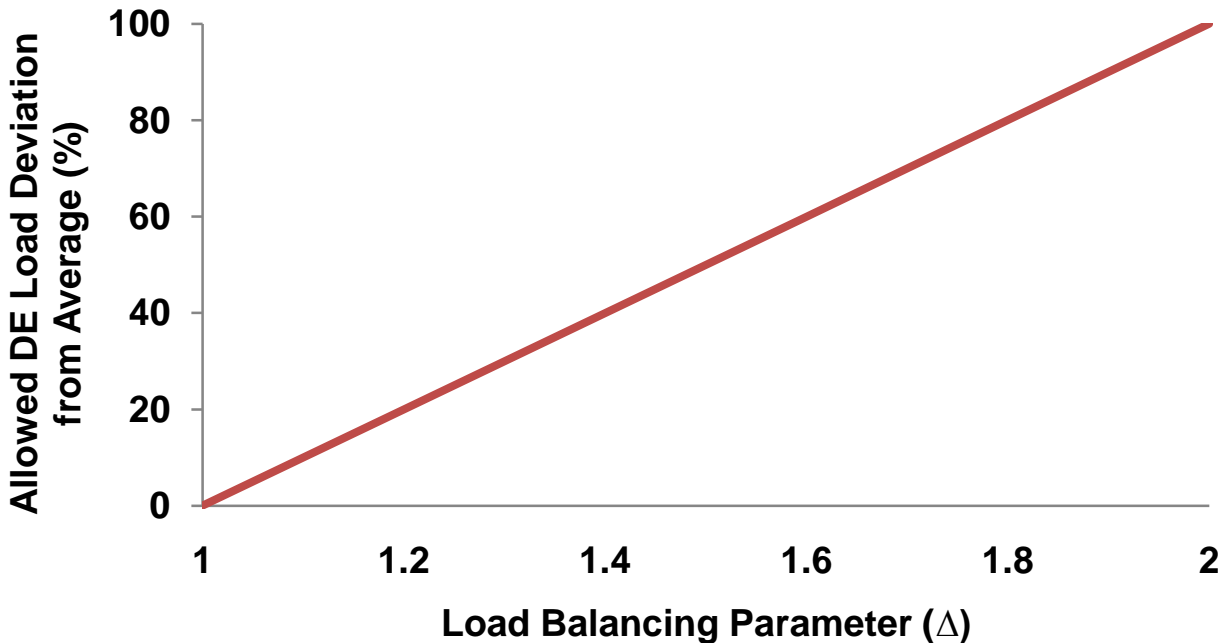
$$\sum_{r_k \in N(r_i)} x(r_k, e_j) \geq x(r_i, e_j)$$

$$x(r_i, e_j) \in \{0, 1\}$$

Parameterized by  $\Delta$   
(continued on next slide)

# Adaptive Router Assignment Load Balancing Approach

Allowed deviation from normalized average workload =  $(\Delta - 1) * 100\%$



Strict load balancing

Delay minimization

# Two Phase Exact Algorithm

## Greedy Phase

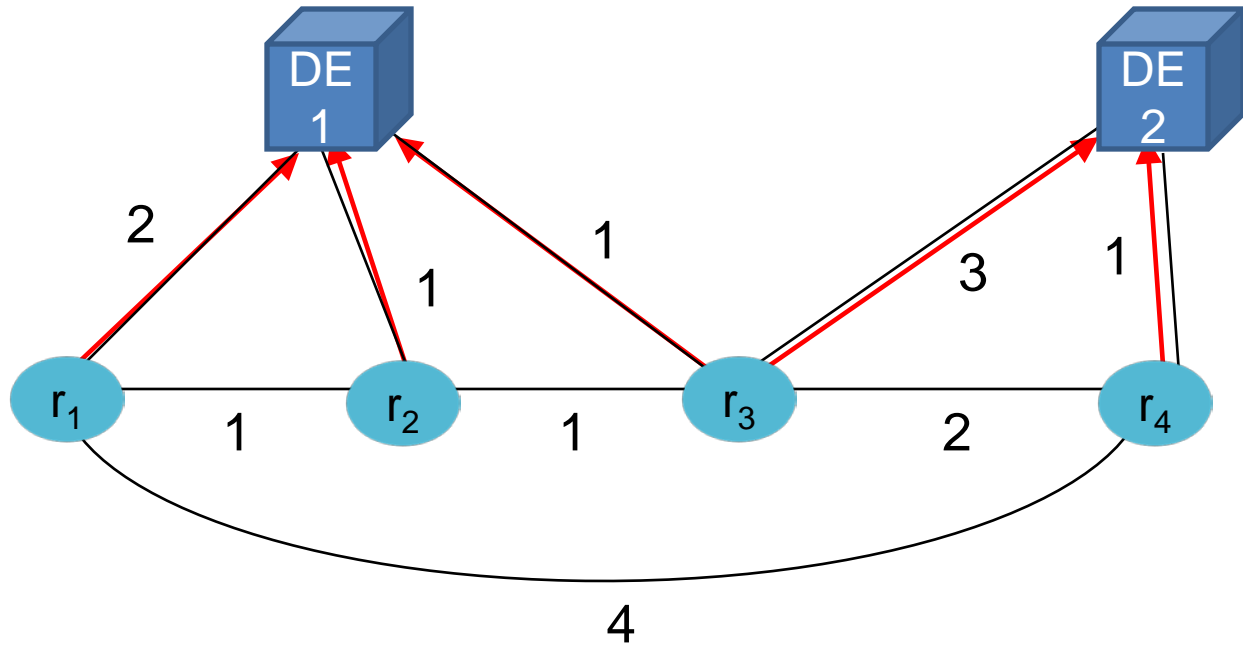
- Assign routers to closest (min-delay) feasible DEs

## Branch Exchange Phase

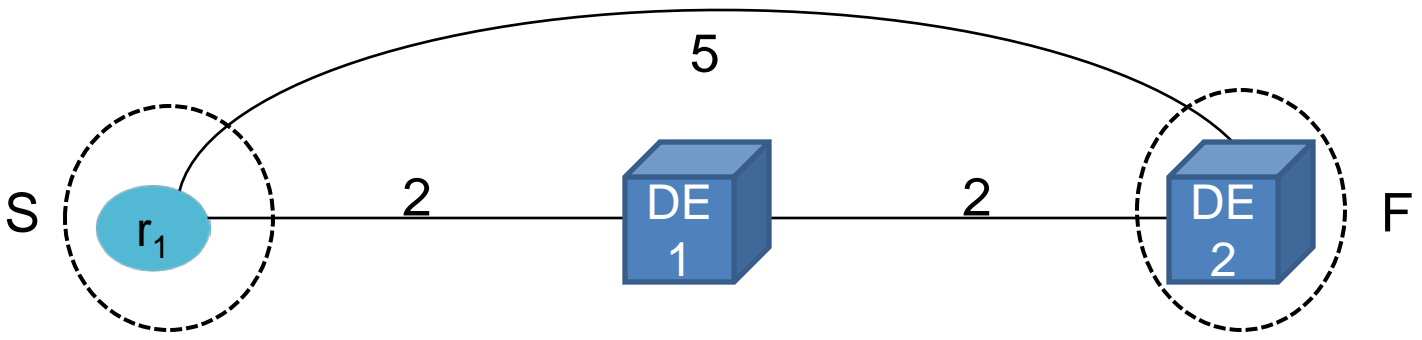
- Iterate for any unassigned routers:
  - Construct auxiliary graph of feasible router assignment exchanges
  - Use shortest-path algorithm to find the min-delay assignment

$\Delta = 1$  (Strict Load Balancing)

Greedy Phase



Exchange Phase



# Complexity

m: # of routers

n: # of DEs

$k \ll m$  (in practical network topologies)

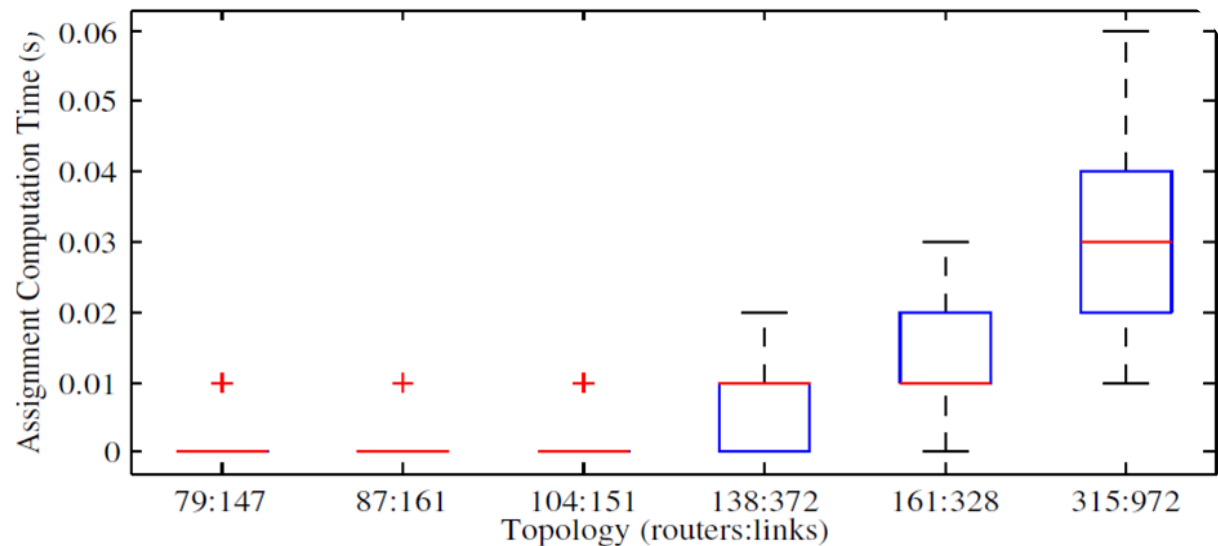
○ Worst-case:  **$O(mn^2)$**

○ Average case:  **$\Theta(m+kn^2)$**

# Evaluation

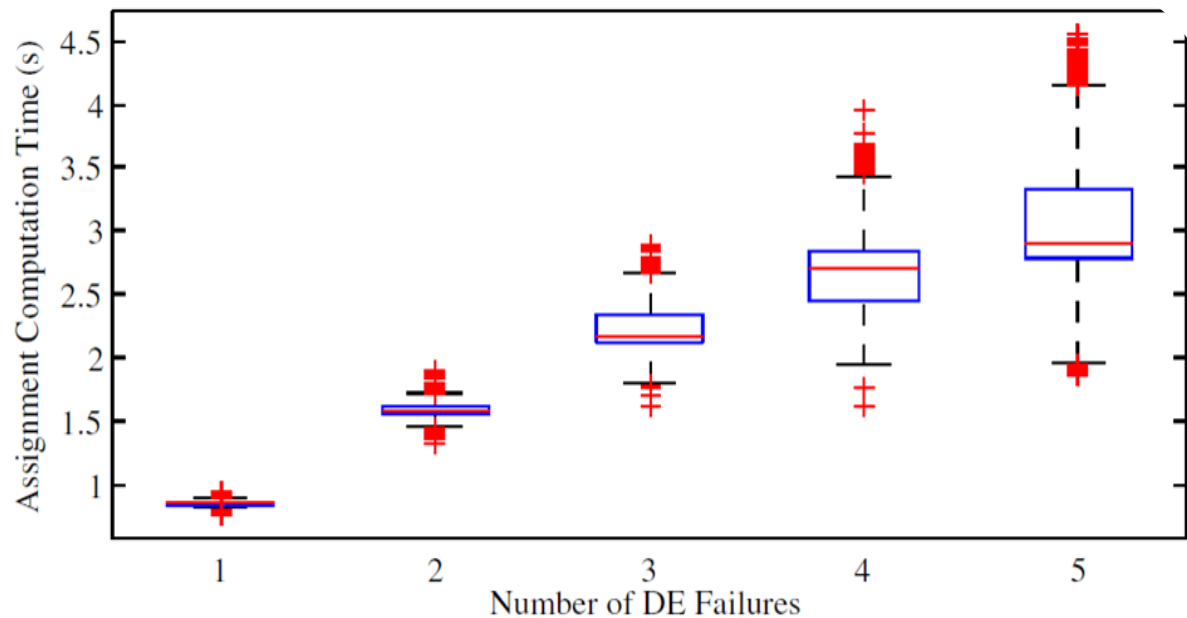
$\Delta=1$  (Strict Load Balancing)  
3.6 Ghz  
4 GB RAM  
64 bit Linux

## Rocketfuel Topologies



## BRITE Topologies

m=1500  
n=15



# Conclusion

## Logical Centralization of Network Control and Management

- ISP PoPs, Large enterprise networks, ...

## Fault-tolerant Decision Plane Design

- Adaptive router assignment algorithm
- Sub-second computation times on real-world topologies

# Questions



hiqbal@cs.pitt.edu