Hot Potatoes Heat Up BGP Routing

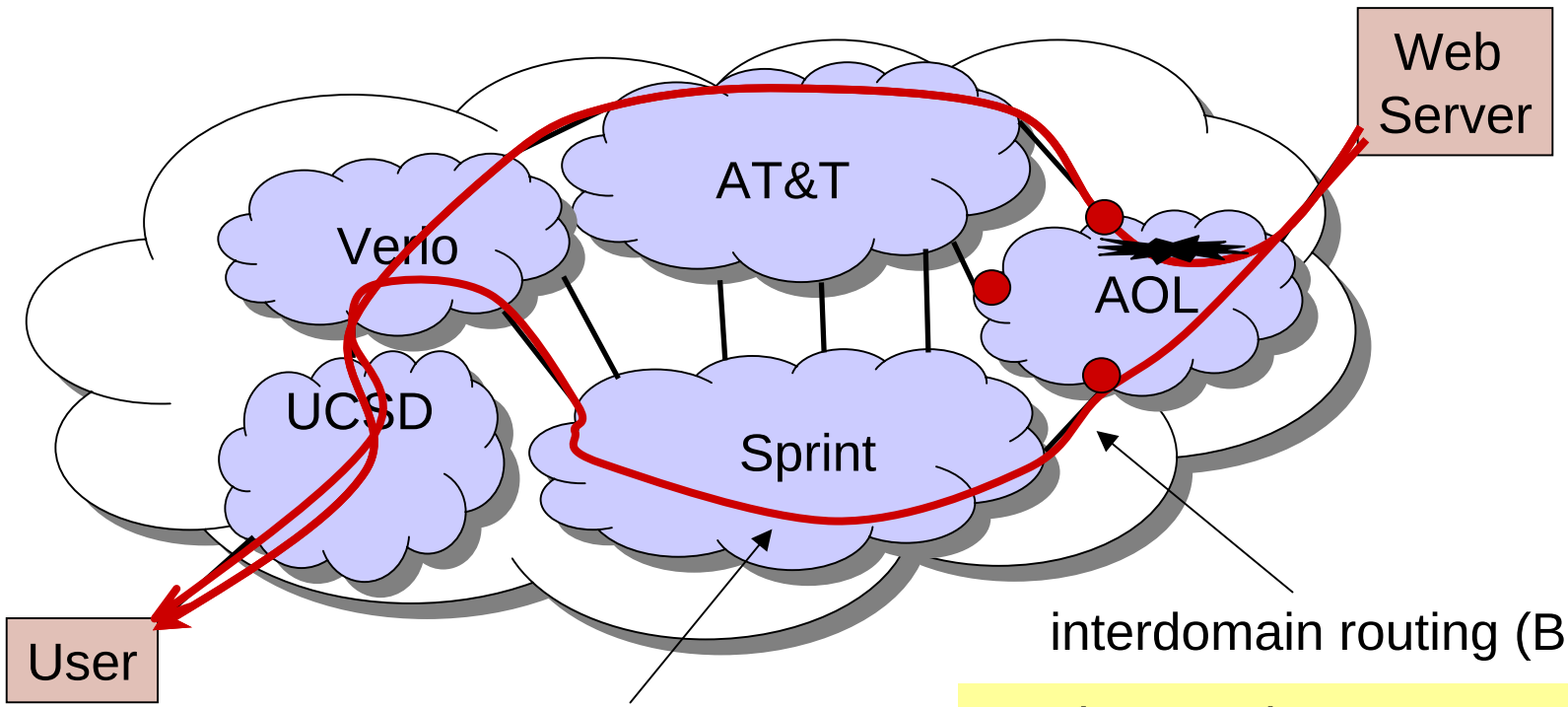
Renata Teixeira

Laboratoire d'Informatique de Paris 6

Université Pierre et Marie Curie

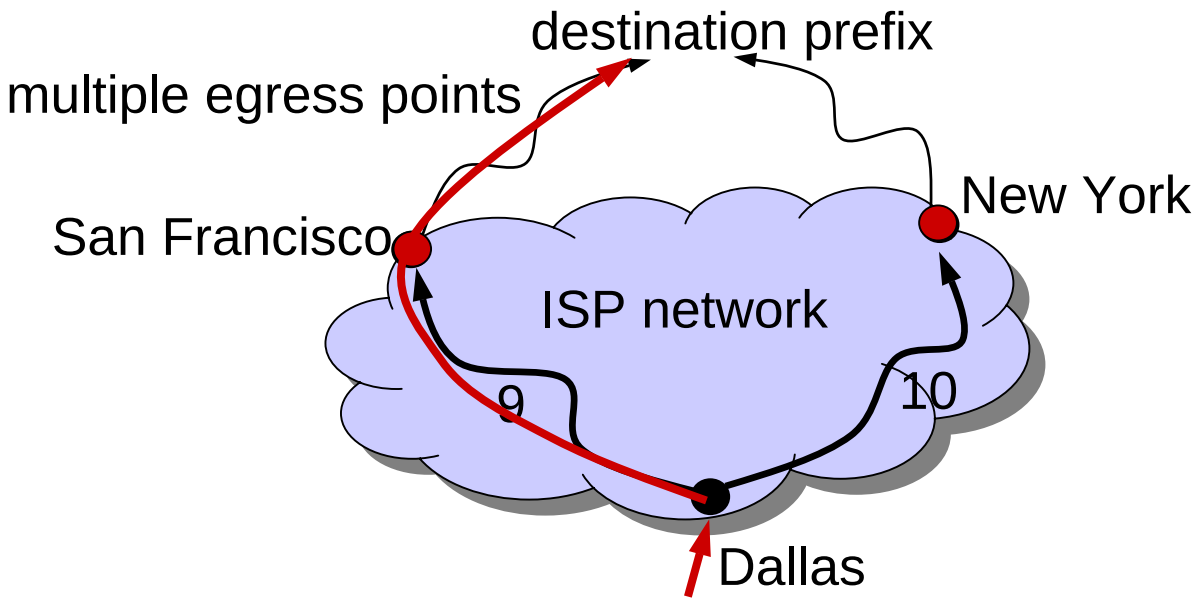
RIPE 51 – Amsterdam

Internet Routing Architecture



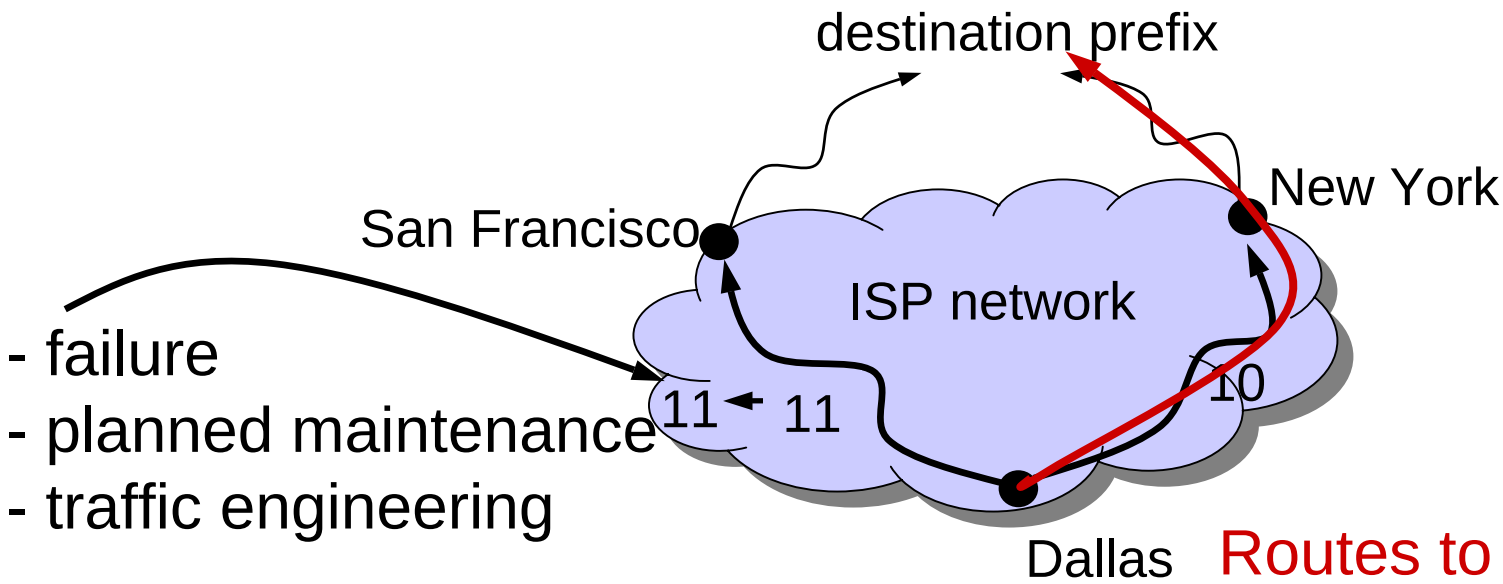
Changes in one AS
may impact traffic
and routing in other ASes

Interaction between IGP and BGP



Hot-potato routing = select closest egress point when there is more than one route to destination

Impact of Internal Routing Changes



- failure
- planned maintenance
- traffic engineering

Routes to thousands of prefixes switch egress points!!!

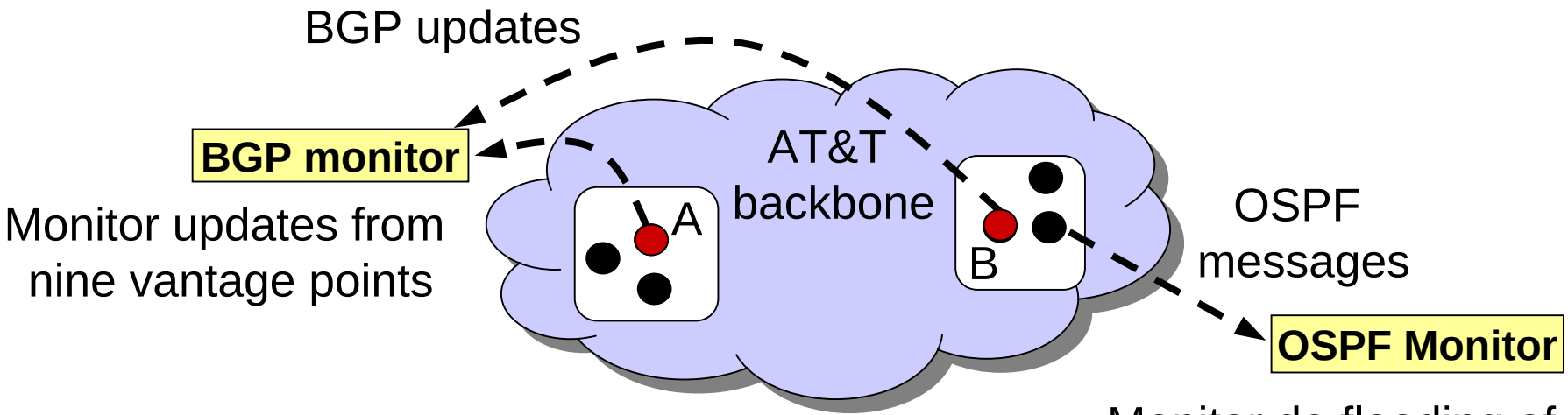
Consequences:

- Transient forwarding instability
- Traffic shift
- Interdomain routing changes

Outline

- Measurement methodology
  - Collection of OSPF and BGP data of AT&T
  - Identification of hot-potato routing changes
- BGP impact
- Traffic impact
- Minimizing hot-potato disruptions

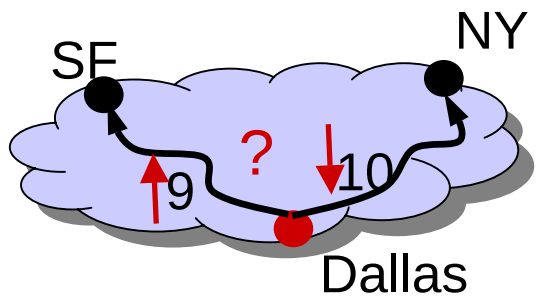
Collecting Input Data



Replay routing decisions
from vantage point A and B
to identify hot-potato changes

Algorithm for Correlating Routing Changes

- Compute distance changes
  - Group OSPF messages close in time
  - Compute distance changes from each vantage point
- Classify BGP changes by possible OSPF cause
  - Group updates close in time
  - Compare old and new route according to decision process
- Determine causal relationship
  - Consistent BGP and OSPF changes
  - Close in time

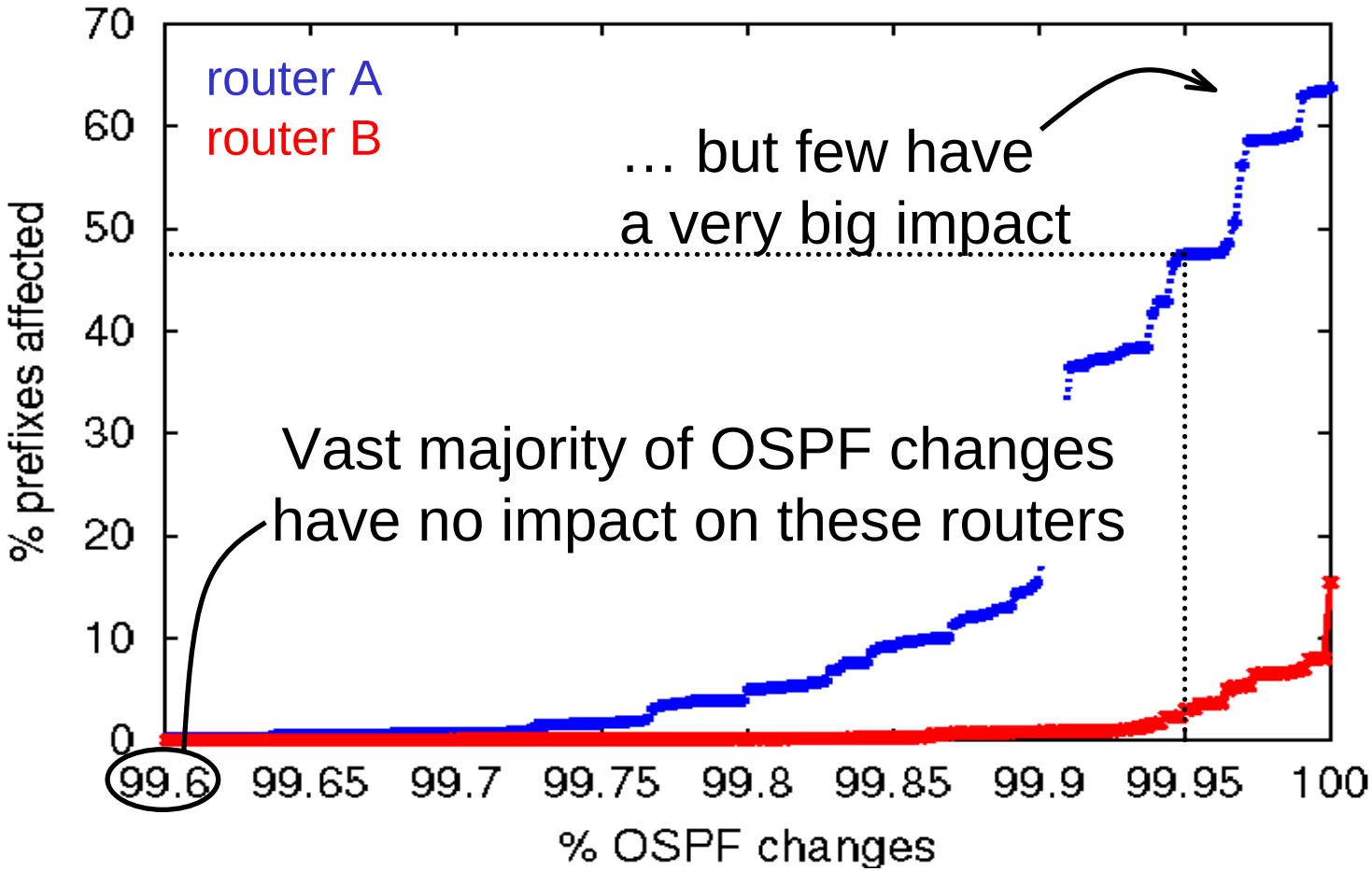


BGP update: SF → NY

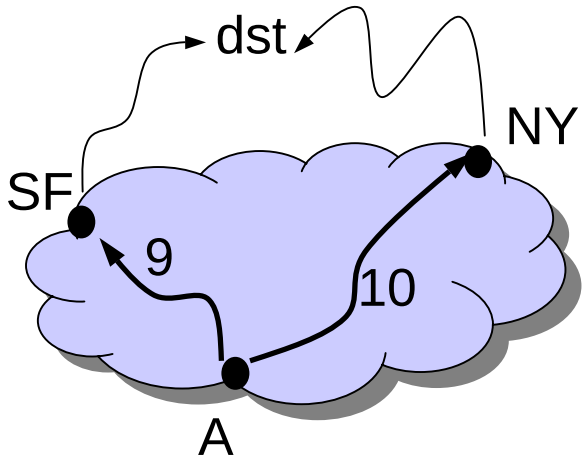
Outline

- Measurement methodology
- BGP impact
  - How often do hot-potato changes happen?
  - Which fraction of prefixes do they affect?
- Traffic impact
- Minimizing hot-potato disruptions

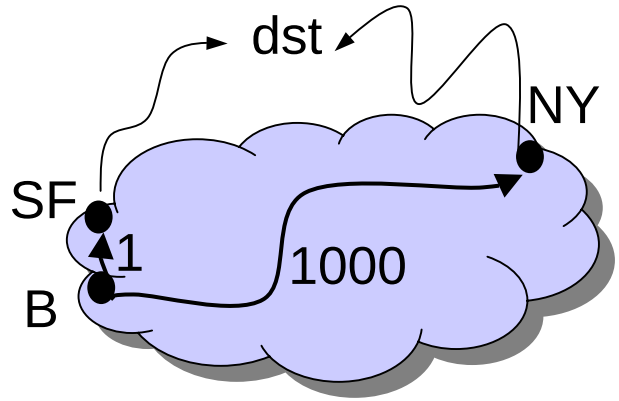
BGP Impact of an OSPF Change



Variation across Routers



Small changes will make router A switch egress points to dst



More robust to intradomain routing changes

Significance of hot-potato routing depends on network design and router location.

Outline

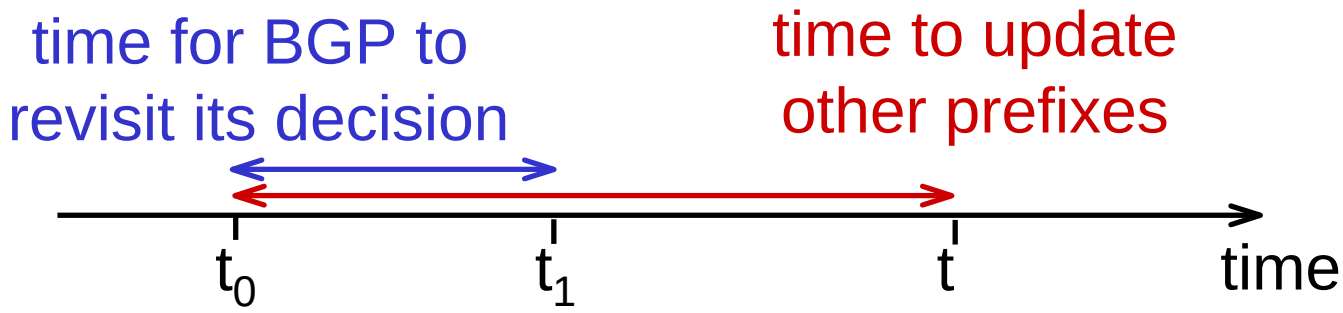
- Measurement methodology
- BGP impact
- Traffic impact
  - How long are convergence delays?
  - What is the impact in the traffic matrix?
- Minimizing hot-potato disruptions

Delay for BGP Routing Change

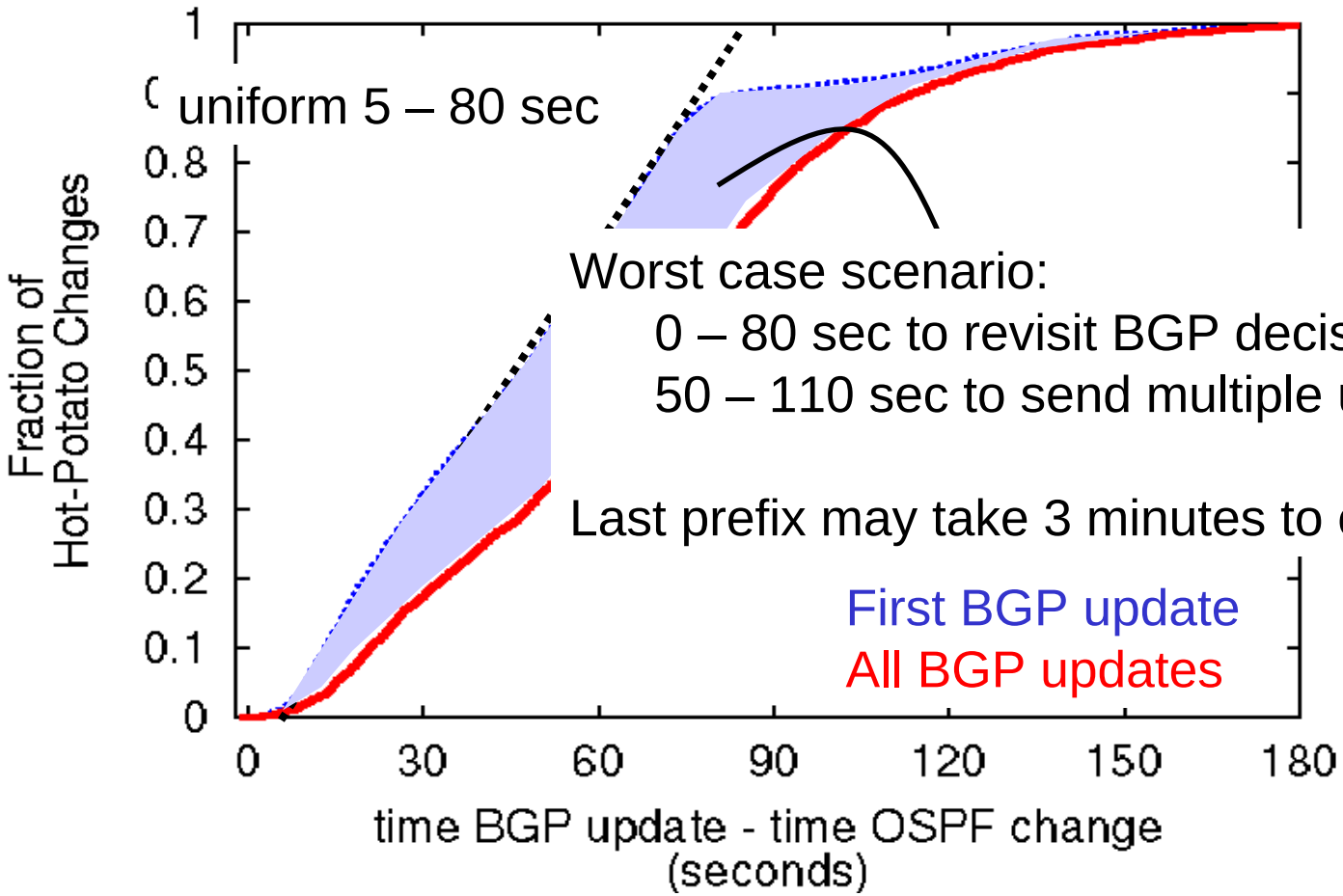
- Steps between OSPF change and BGP update
  - OSPF message flooded through the network (t_0)
  - OSPF updates distance information
  - BGP decision process rerun (timer driven)
  - BGP update sent to another router (t)
    - First BGP update sent (t_1)
- Metrics

OSPF monitor

BGP monitor



BGP Reaction Time

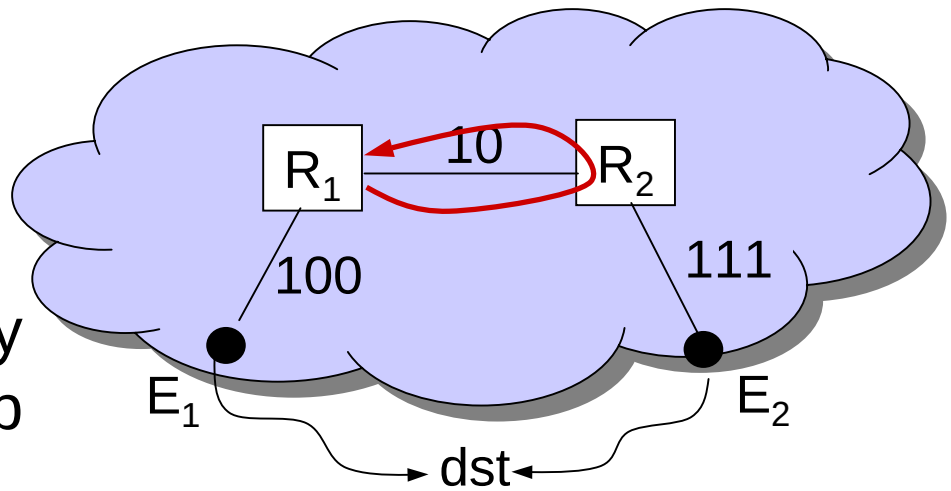


Transient Data Disruptions

1 – BGP decision process runs in R_2

2 – R_2 starts using E_1 to reach dst

3 – R_1's BGP decision can take up to 60 seconds to run

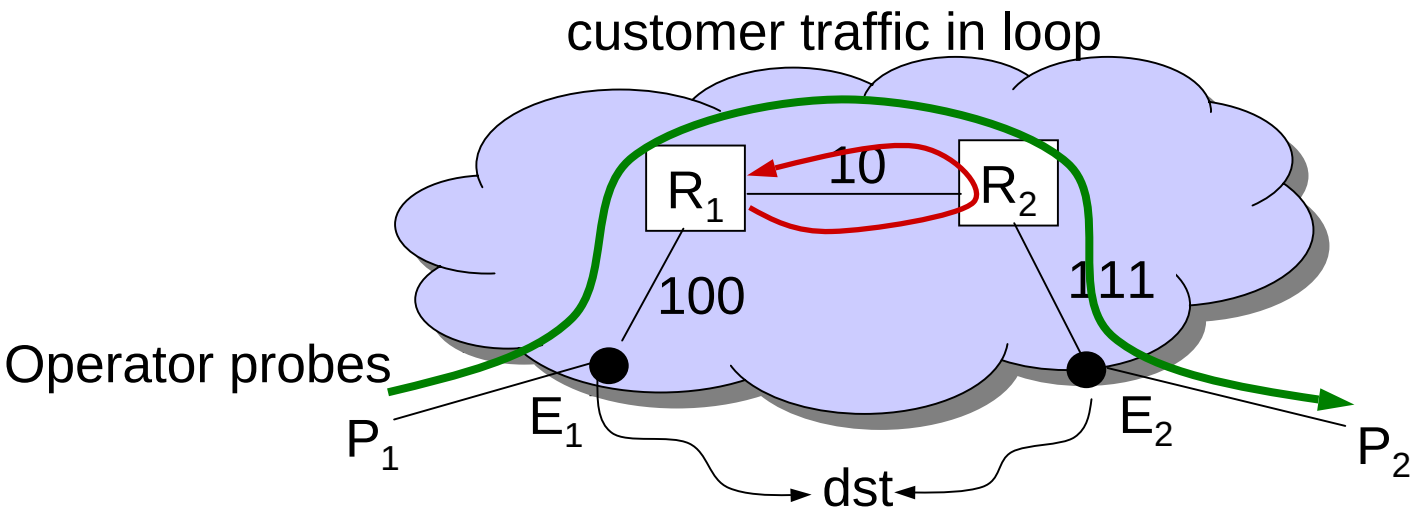


Packets to dst may be caught in a loop for 60 seconds!

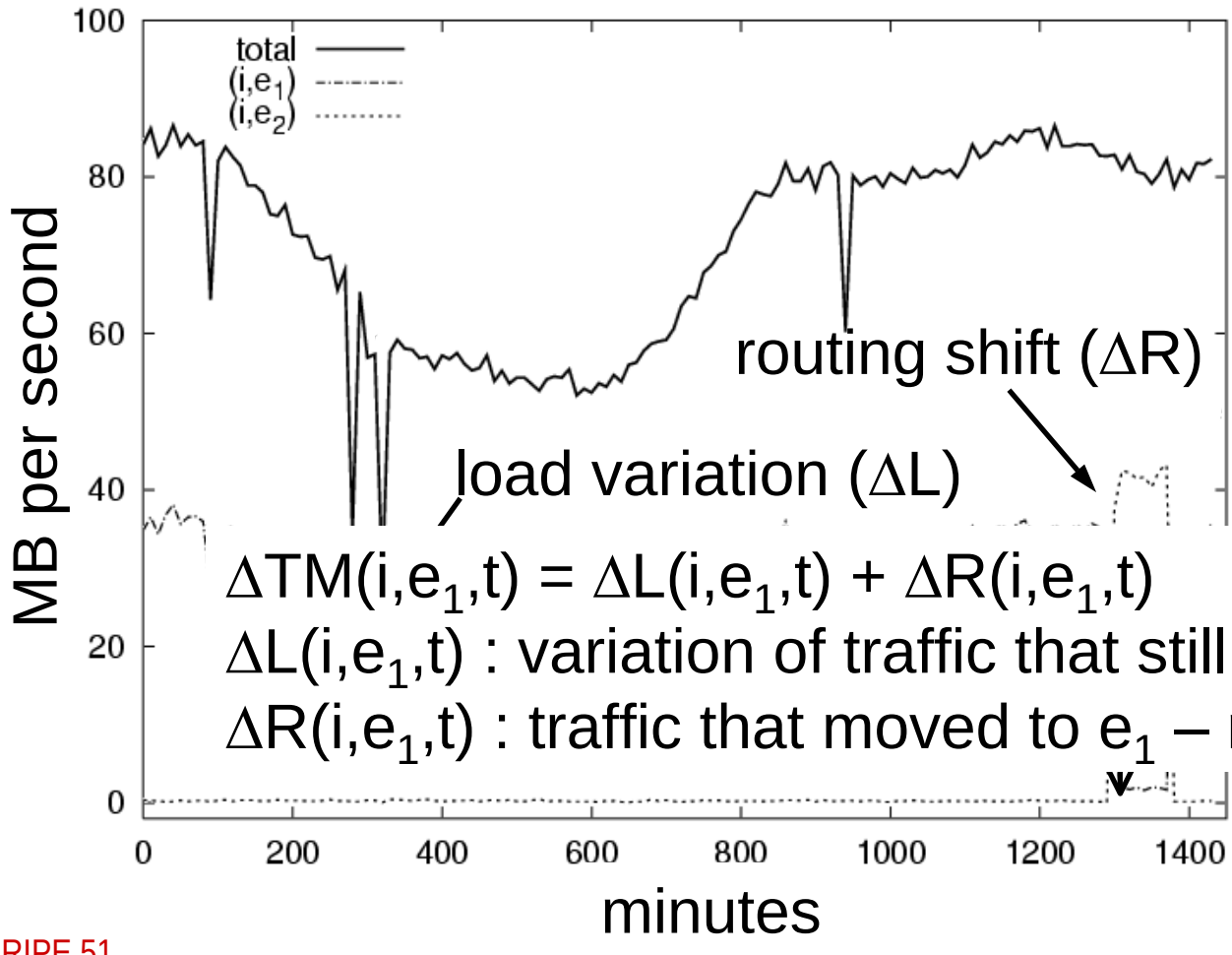
Disastrous for interactive applications (VoIP, gaming, web)

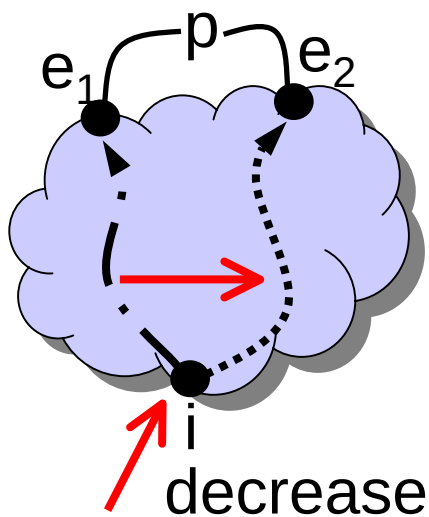
Challenges for Active Measurements

- Problem: Single-homed probe machines
  - Probes do not experience the loop
  - Probes do not illustrate the customer experience



Traffic Shifts



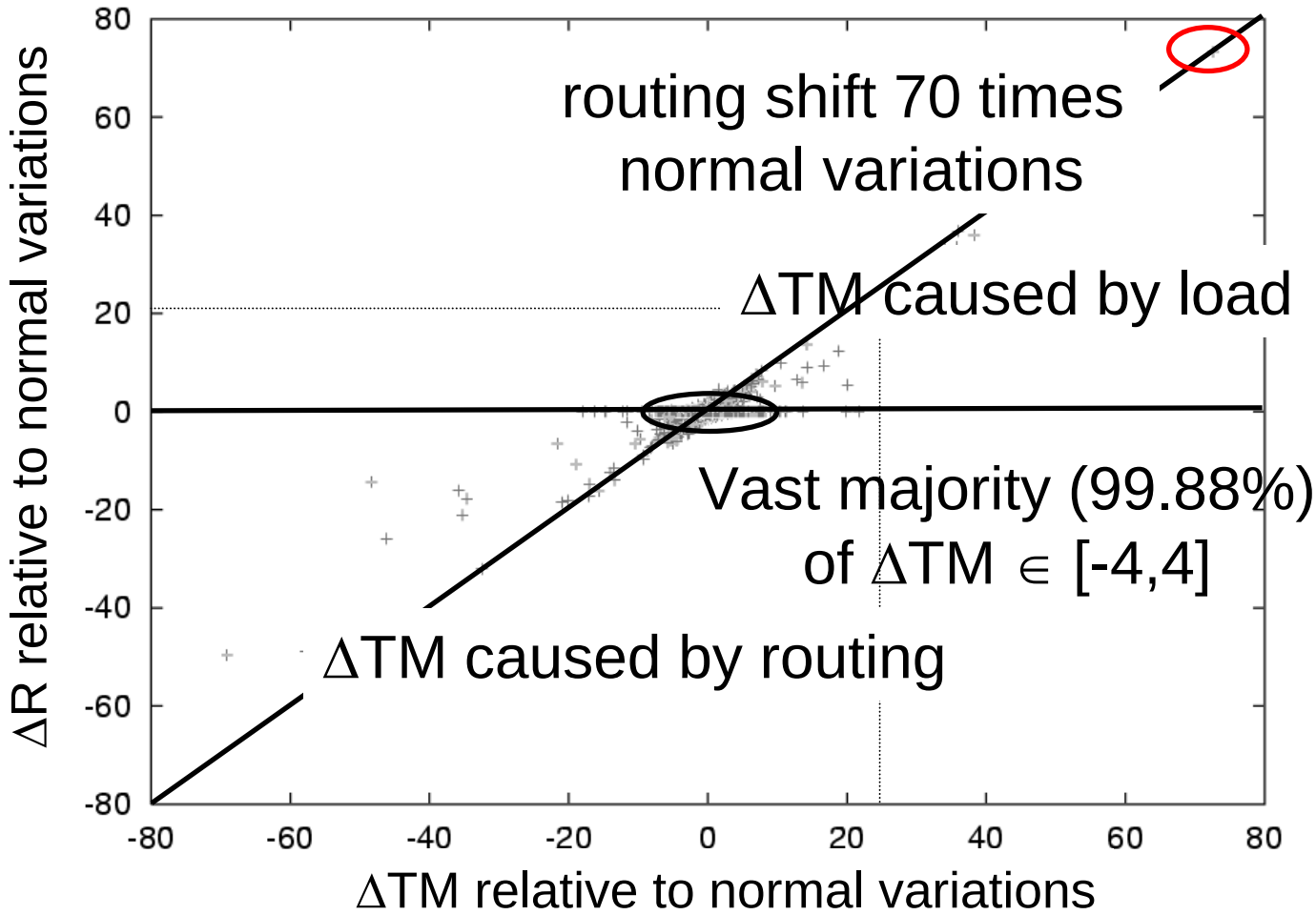


$\Delta TM(i,e_1,t) = \Delta L(i,e_1,t) + \Delta R(i,e_1,t)$

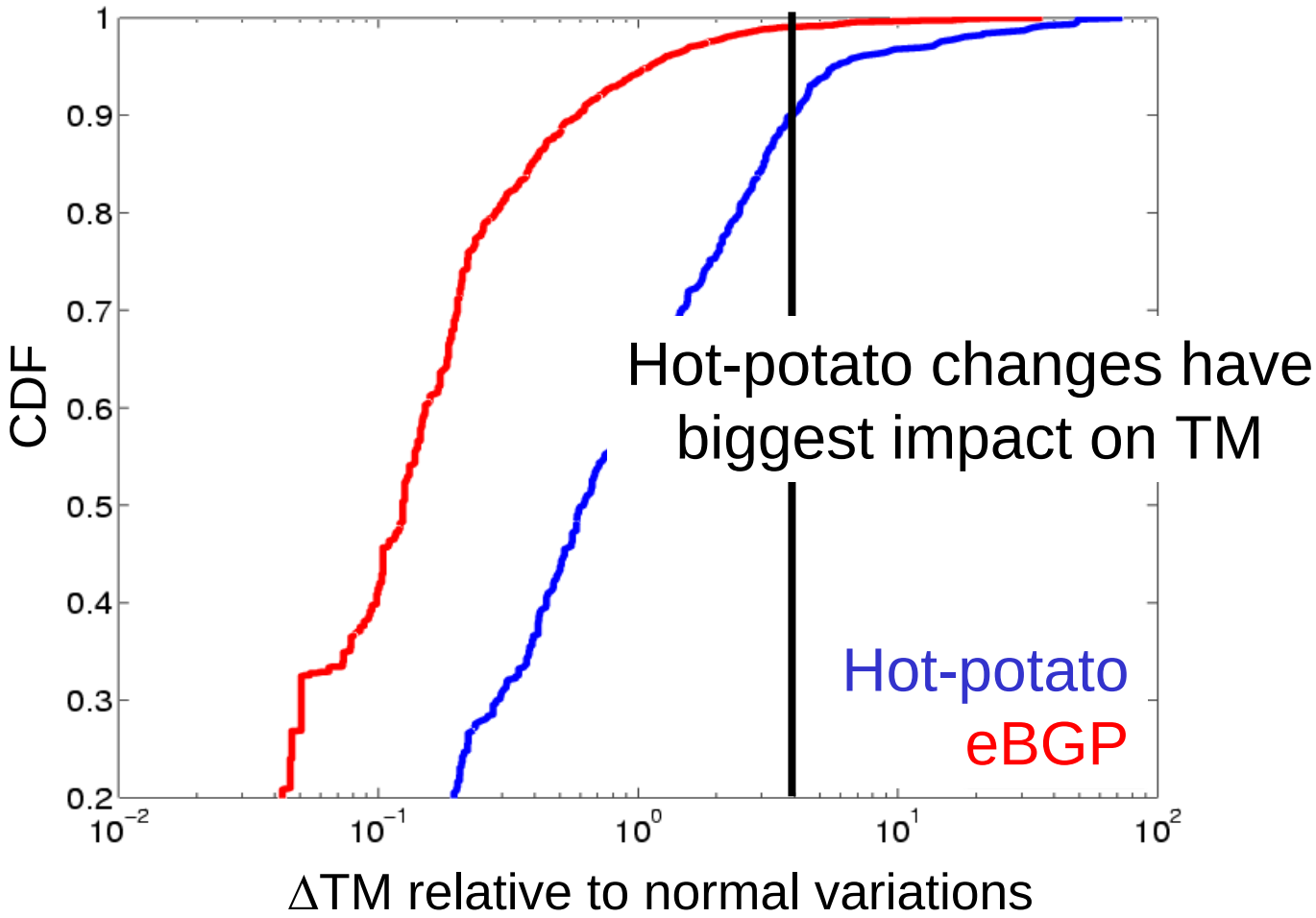
$\Delta L(i,e_1,t)$: variation of traffic that still uses e_1

$\Delta R(i,e_1,t)$: traffic that moved to e_1 – moved out of e_1

Large Shifts Caused by Routing Changes

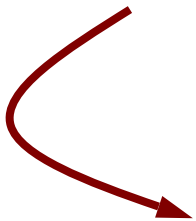


Hot-potato vs. External BGP Routing Changes



Summary of Measurement Analysis

- Convergence can take minutes
  - Forwarding loops, leads to packet loss and delay
  - Fixes: event-driven implementations or tunnels
- Frequency of hot-potato changes depends on location
  - Once a week on average for more affected routers
- Internal events can have big impact
  - Some events affect over half of a BGP table
  - Responsible for largest traffic variations



Implications

- End users: Transient disruptions and new end-to-end path characteristics
- Network administrators: Instability in the traffic matrix

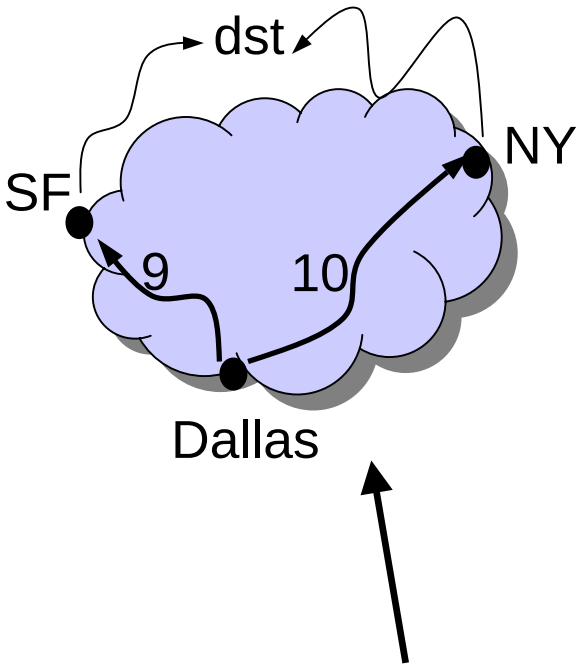
Outline

- Measurement methodology
- BGP impact
- Traffic impact
- Minimizing hot-potato disruptions
  - What can operators do today?

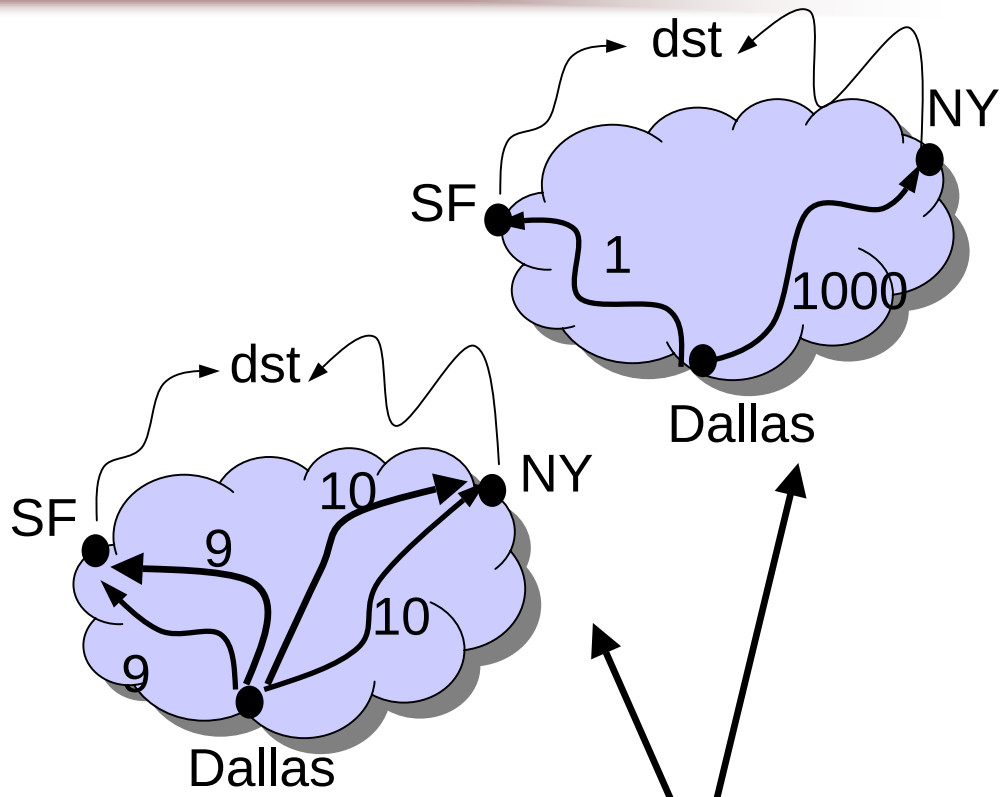
What can operators do today?

- Network design
  - Design networks that minimize hot-potato changes
  - Implement a fixed ranking of egress points (e.g., MPLS tunnels injected in IGP)
- Maintenance
  - Plan maintenance activities considering the impact of changes on BGP routes
- Monitoring
  - Deploy measurement infrastructure that captures disruptions caused by hot-potato routing

Comparison of Network Designs

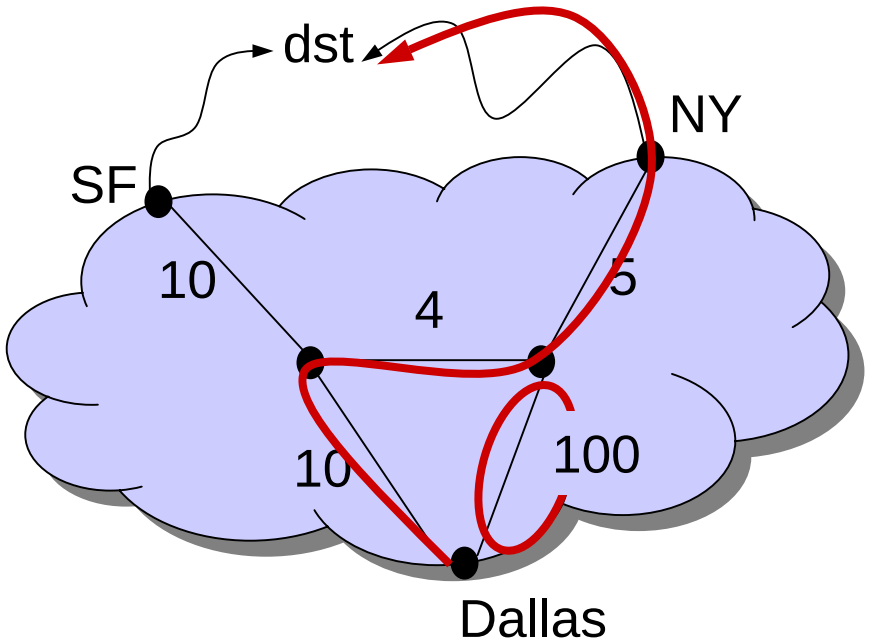


Small changes will make Dallas switch egress points to dst



More robust to intradomain routing changes

Careful Cost in/out Links



Conclusion

- Hot-potato routing is too disruptive
  - Small changes inside an AS can lead to big disruptions on BGP and transit traffic
- In addition, hot potato is…
  - Too restrictive: Egress selection mechanism dictates a policy
  - Too convoluted: IGP metrics determine BGP egress selection
- Introduce more flexible egress selection mechanism
  - TIE: Tunable Interdomain Egress selection

More Info

http://rp.lip6.fr/~teixeira

- BGP impact
  - R. Teixeira, A. Shaikh, T. Griffin, and J. Rexford, "Dynamics of Hot-Potato Routing in IP networks", in proceedings of ACM SIGMETRICS, June 2004.
- Traffic impact
  - R. Teixeira, N. Duffield, J. Rexford, and M. Roughan, "Traffic Matrix Reloaded: Impact of Routing Changes", in proceedings of PAM, March 2005.
- Model of network sensitivity to IGP changes
  - R. Teixeira, T. Griffin, A. Shaikh, and G.M. Voelker, "Network Sensitivity to Hot-Potato Disruptions", in proceedings of ACM SIGCOMM, August 2004.
- New egress selection mechanism
  - R. Teixeira, T. Griffin, M. Resende, and J. Rexford, "TIE Breaking: Tunable Interdomain Egress Selection", in proceedings of CoNext, October 2005.