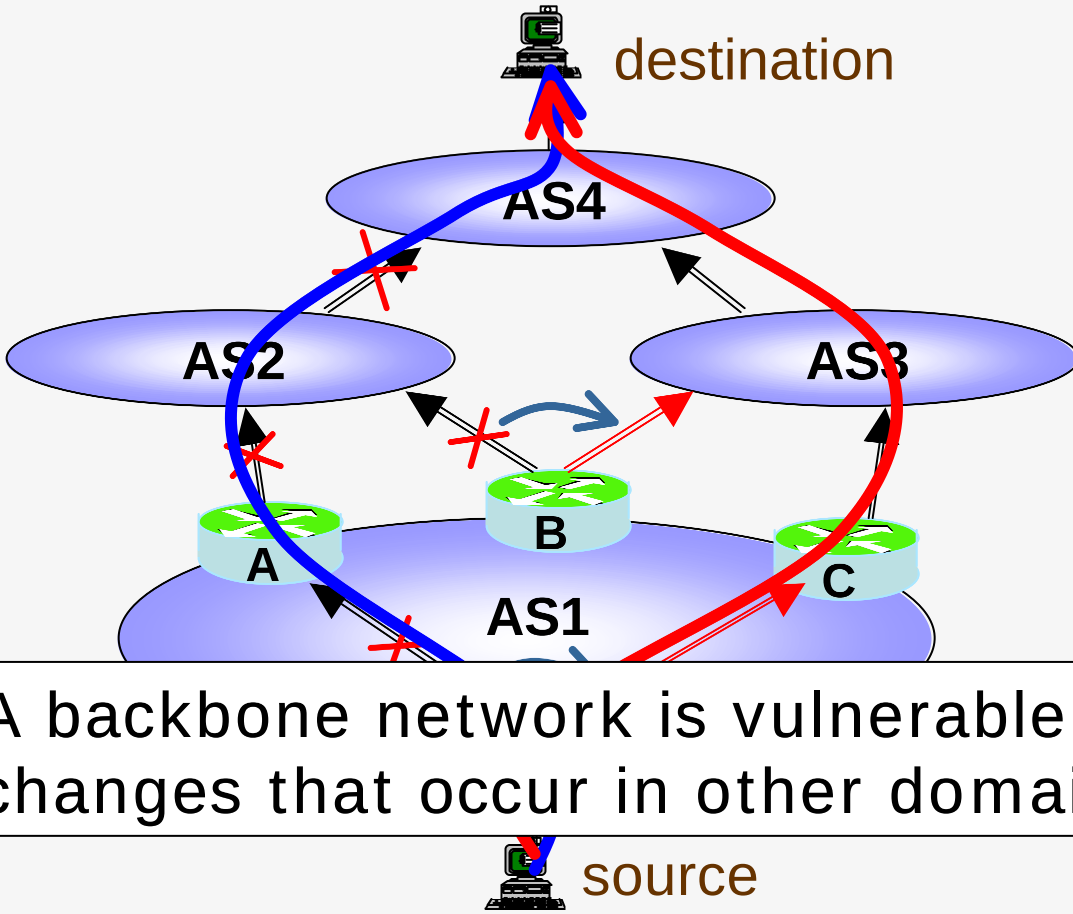
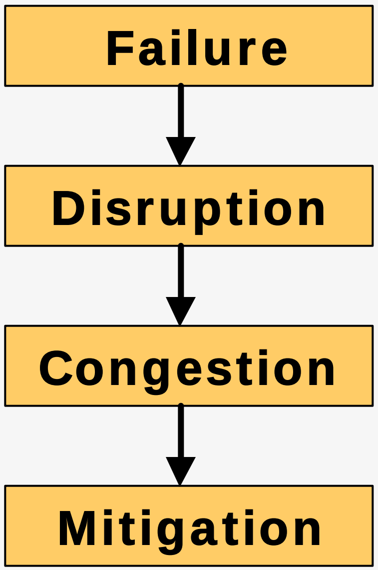
Finding a Needle in a Haystack: Pinpointing Significant BGP Routing Changes in an IP Network

Jian Wu (University of Michigan)
Z. Morley Mao (University of Michigan)
Jennifer Rexford (Princeton University)
Jia Wang (AT&T Labs Research)

Motivation





A backbone network is vulnerable to routing changes that occur in other domains.

Goal

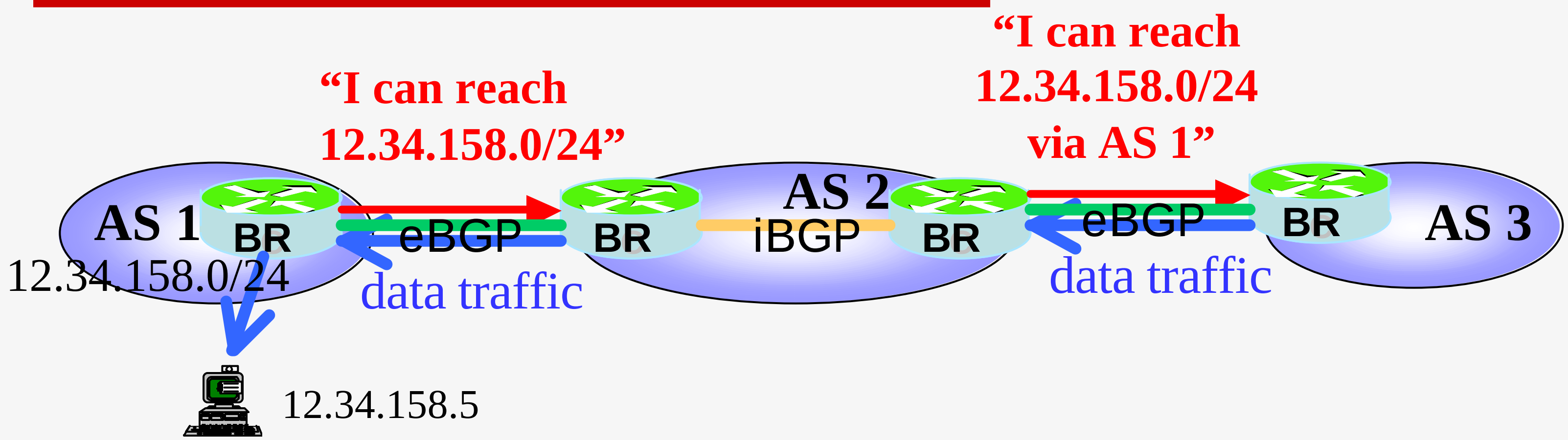
- Identify important routing anomalies
  - Lost reachability
  - Persistent flapping
  - Large traffic shifts

Contributions:

•Build a tool to identify a small number of important routing disruptions from a large volume of raw BGP updates in real time.

•Use the tool to characterize routing disruptions in an operational network

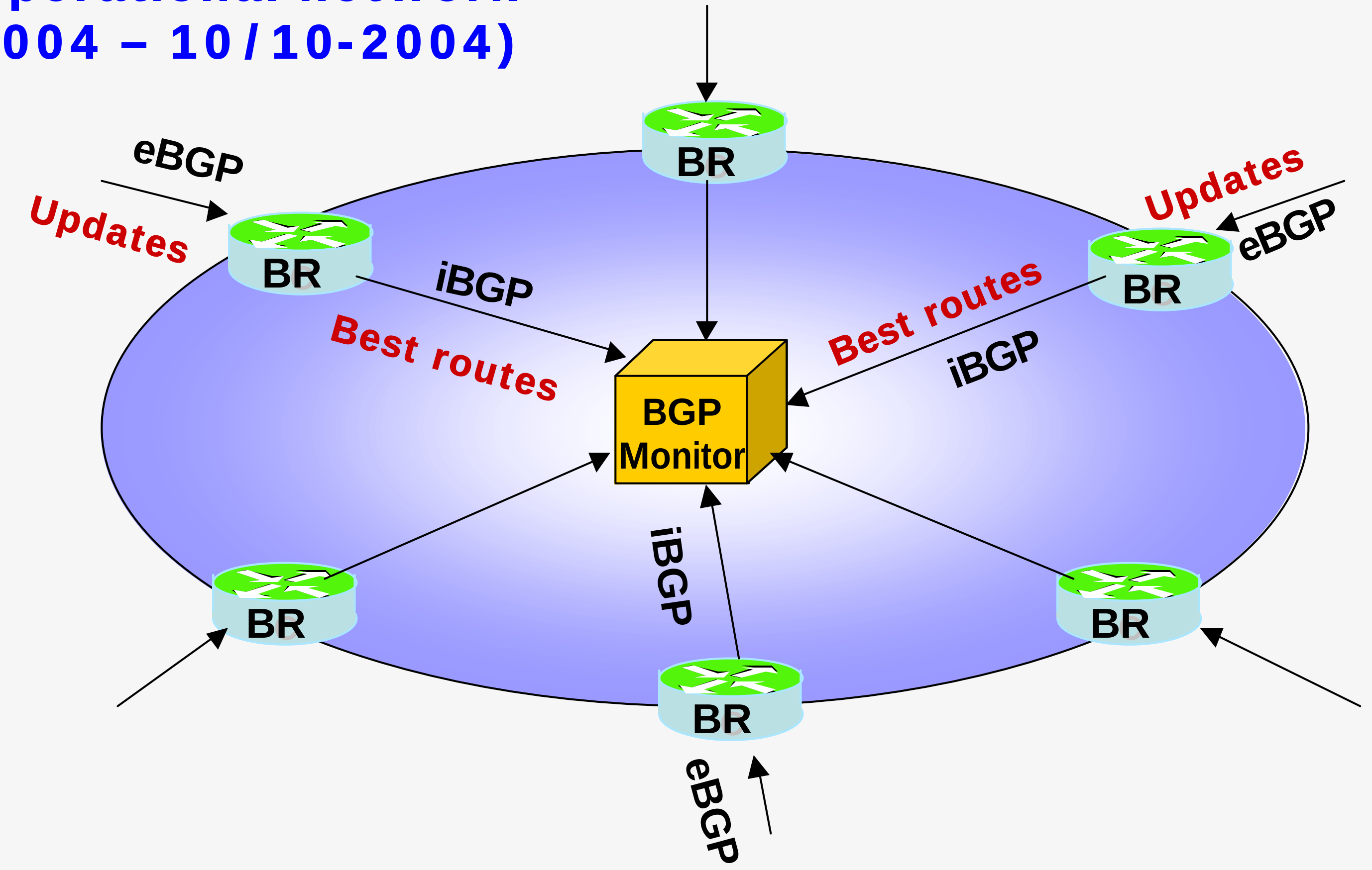
Interdomain Routing: Border Gateway Protocol



- Prefix-based: one route per prefix
- Path-vector: list of ASes in the path
- Incremental: every update indicates a change
- Policy-based: local ranking of routes

Capturing Routing Changes

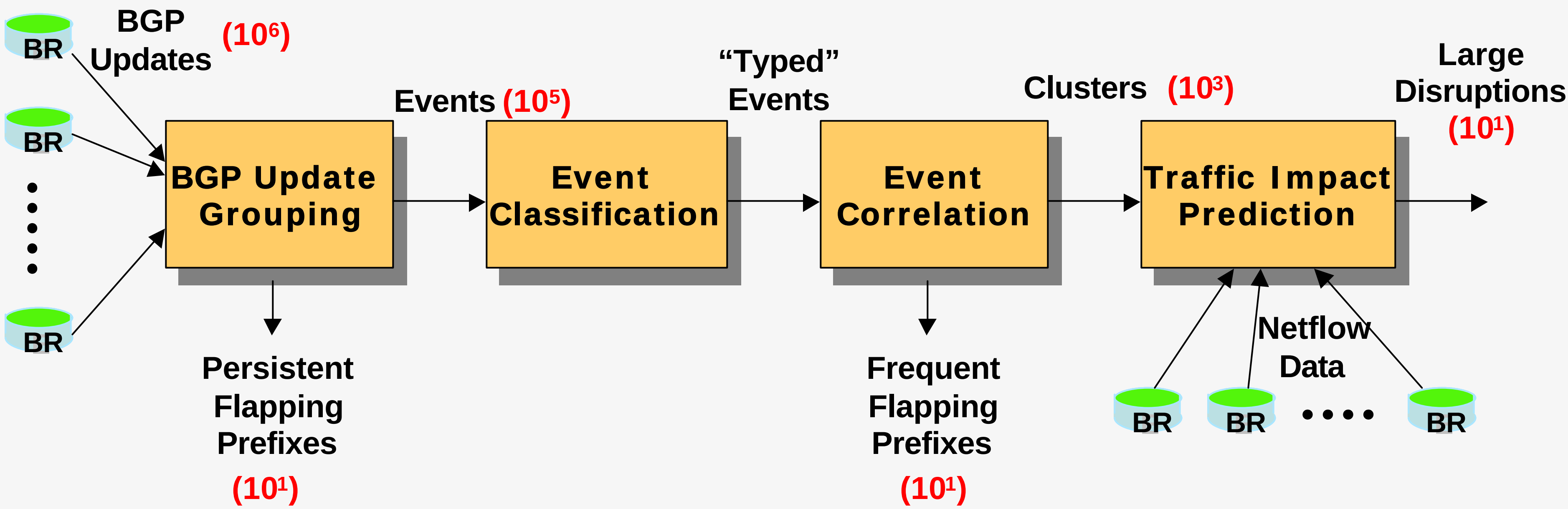
A large operational network (8/16/2004 – 10/10-2004)



Challenges

- Large volume of BGP updates
  - Millions daily, very bursty
  - Too much for an operator to manage
- Different from root-cause analysis
  - Identify changes and their effects
  - Focus on actionable events rather than diagnosis
  - Diagnose causes in/near the AS

System Architecture

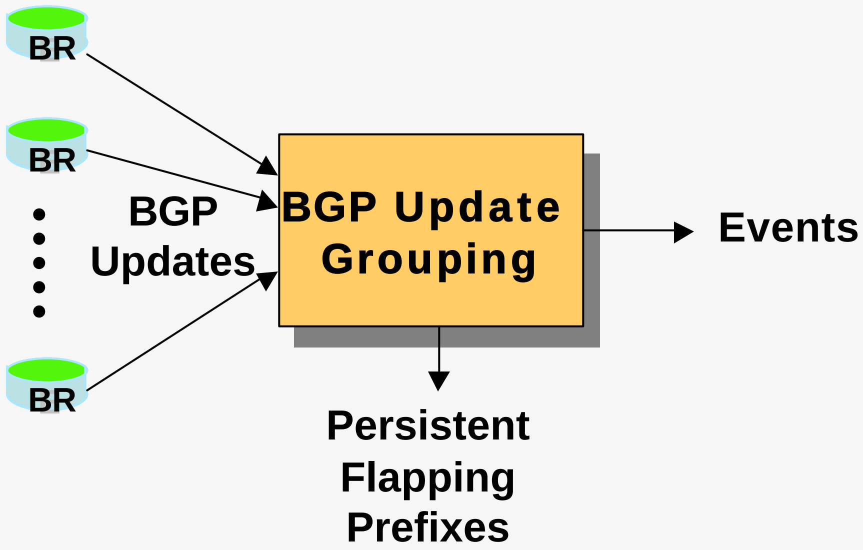


From millions of updates to a few dozen reports

Grouping BGP Update into Events

Challenge: A single routing change

- leads to multiple update messages
- affects routing decisions at multiple routers



Approach:

- **Group together all updates for a prefix with inter-arrival < 70 seconds**
- **Flag prefixes with changes lasting > 10 minutes.**

Grouping Thresholds

- Based on our understanding of BGP and data analysis
- Event timeout: 70 seconds
  - 2 * MRAI timer + 10 seconds
  - 98% inter-arrival time < 70 seconds
- Convergence timeout: 10 minutes
  - BGP usually converges within a few minutes
  - 99.9% events < 10 minutes

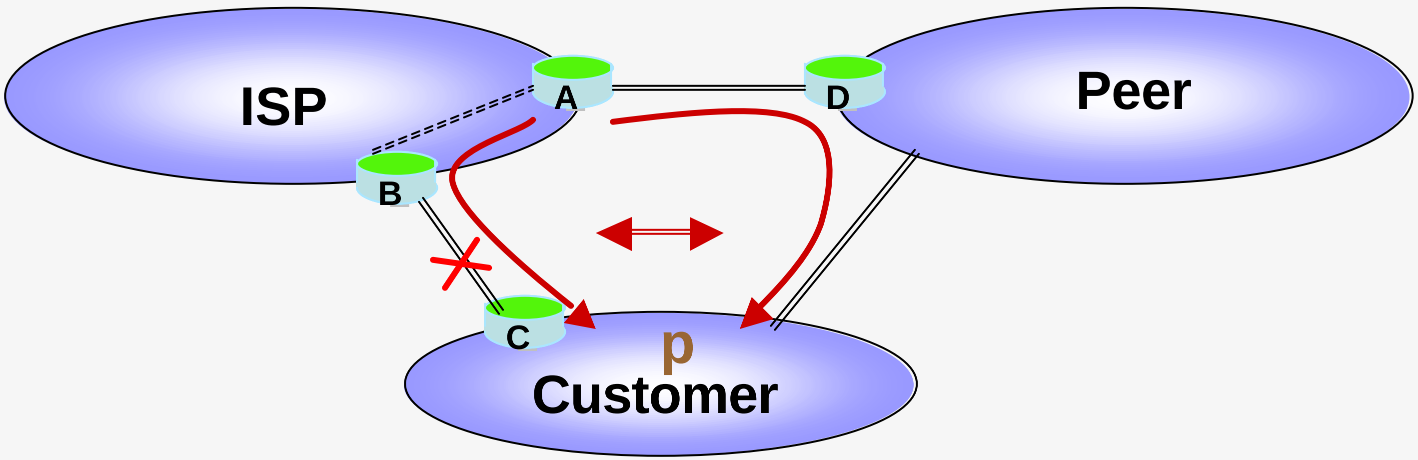
Persistent Flapping Prefixes

A surprising finding:

15.2% of updates were caused by persistent-flapping prefixes even though flap damping is enabled.

- Types of persistent flapping
  - Conservative damping parameters (78.6%)
  - Protocol oscillations due to MED (18.3%)
  - Unstable interfaces or BGP sessions (3.0%)

Example: Unstable eBGP Session

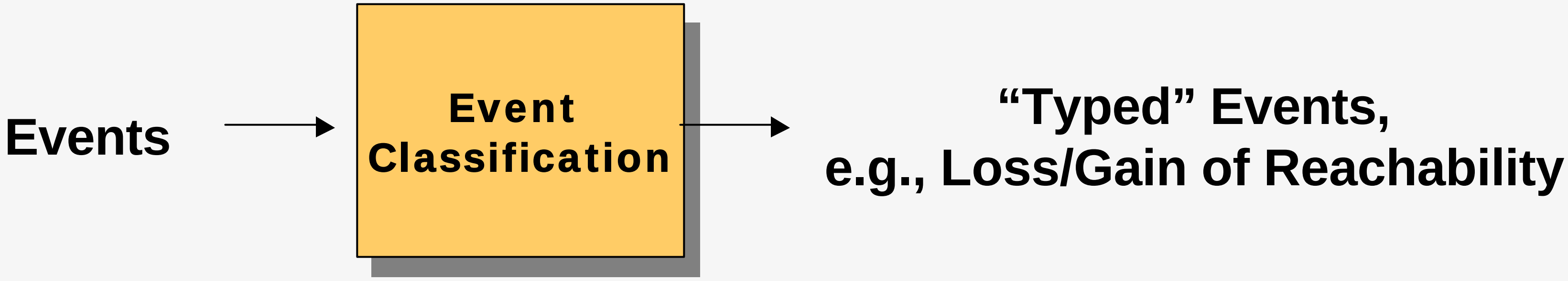


- Flap damping parameters is session-based
- Damping not implemented for iBGP sessions

Event Classification

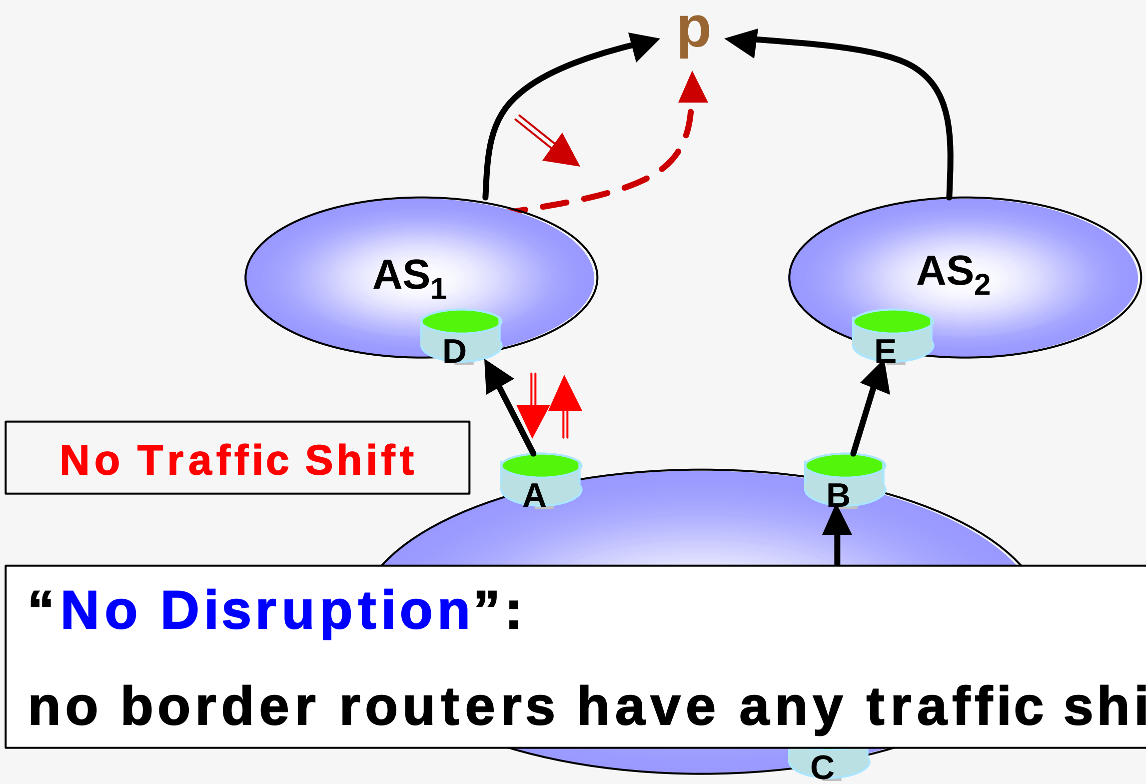
Challenge: Major concerns in network management

- Changes in reachability
- Heavy load of routing messages on the routers
- Change of flow of the traffic through the network



Solution: classify events by severity of their impact

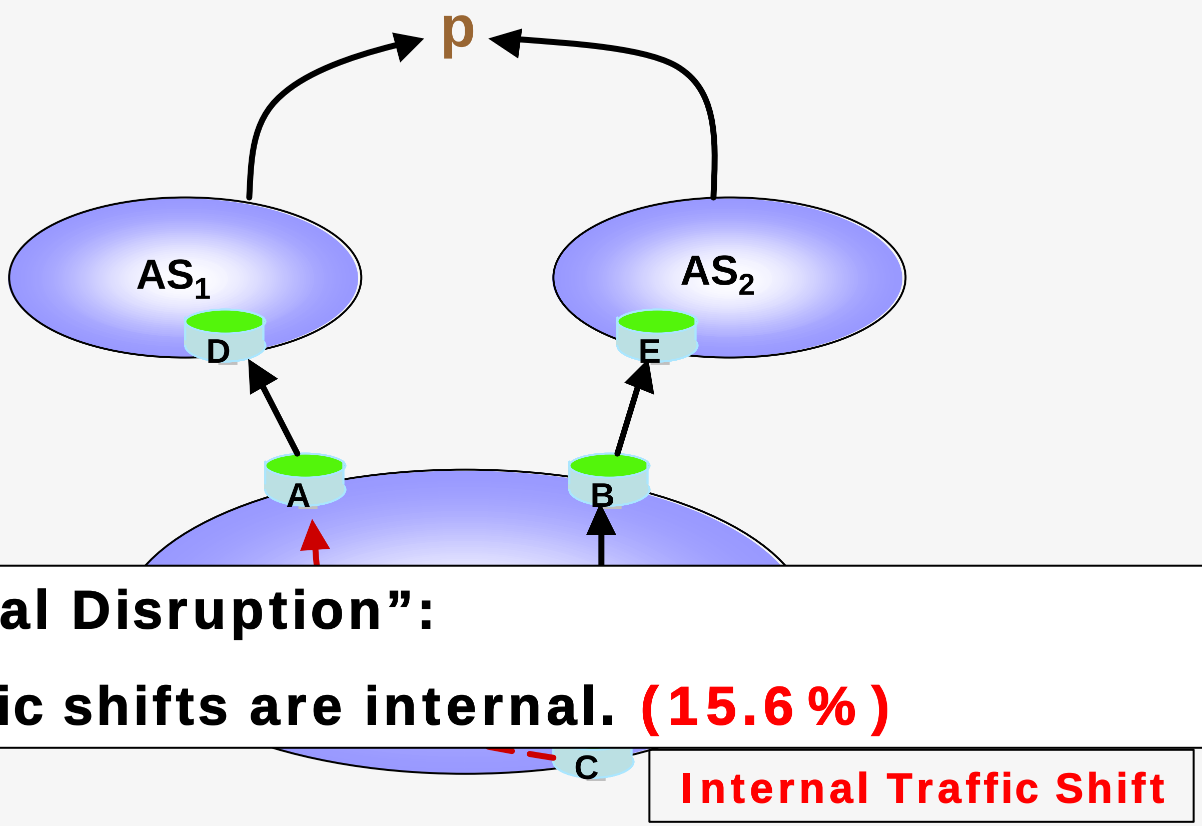
Event Category – “No Disruption”



“**No Disruption**”:

no border routers have any traffic shift. (50.3%)

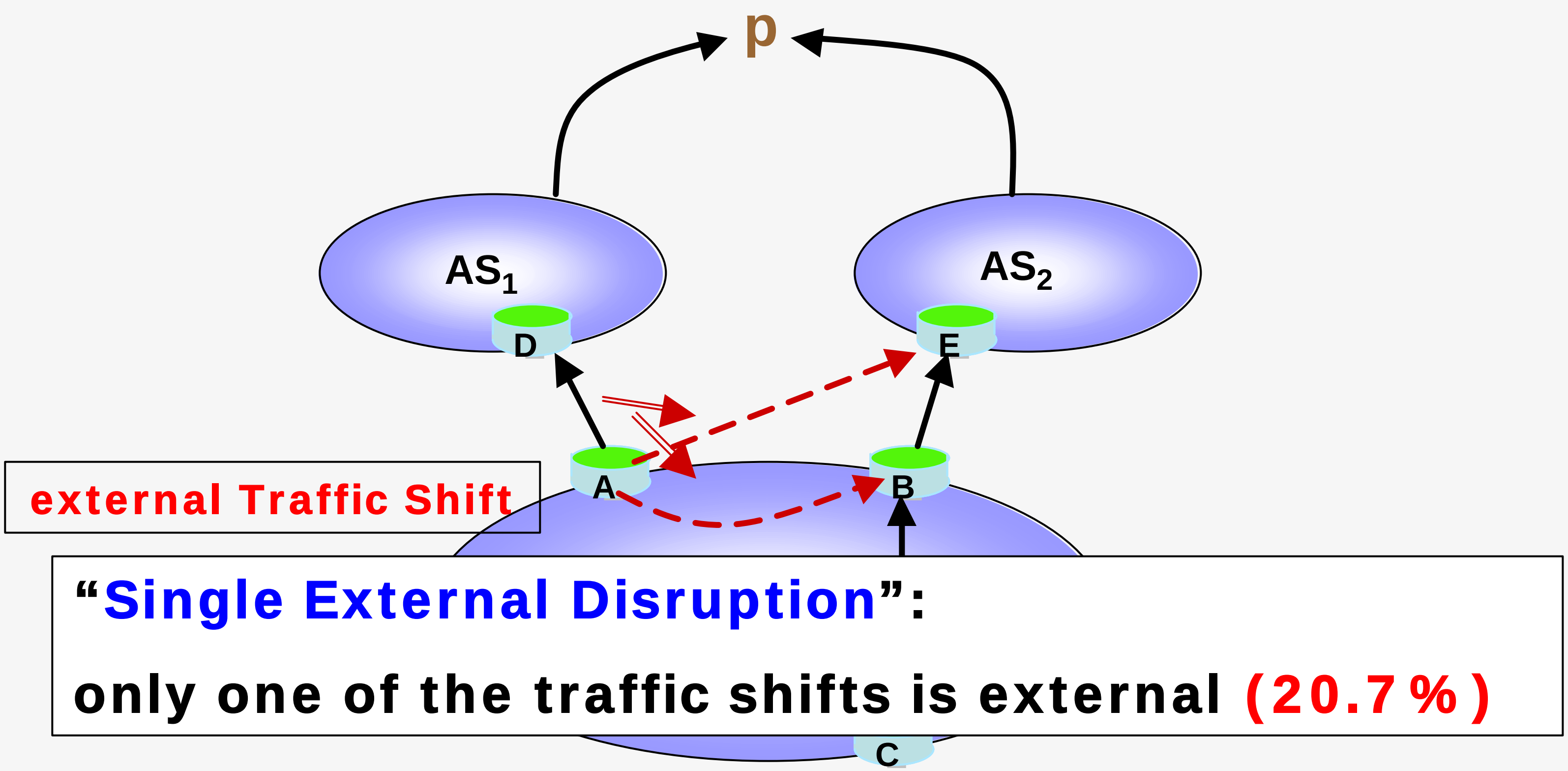
Event Category – "Internal Disruption"



"Internal Disruption":

all traffic shifts are internal. (15.6%)

Event Category – "Single External Disruption"



"Single External Disruption":

only one of the traffic shifts is external (20.7%)

Statistics on Event Classification

	Events	Updates
No Disruption	50.3%	48.6%
Internal Disruption	15.6%	3.4%
Single External Disruption	20.7%	7.9%
Multiple External Disruption	7.4%	18.2%
Loss/Gain of Reachability	6.0%	21.9%

- First 3 categories have significant day-to-day variations
- Updates per event depends on the type of events and the number of affected routers

Event Correlation

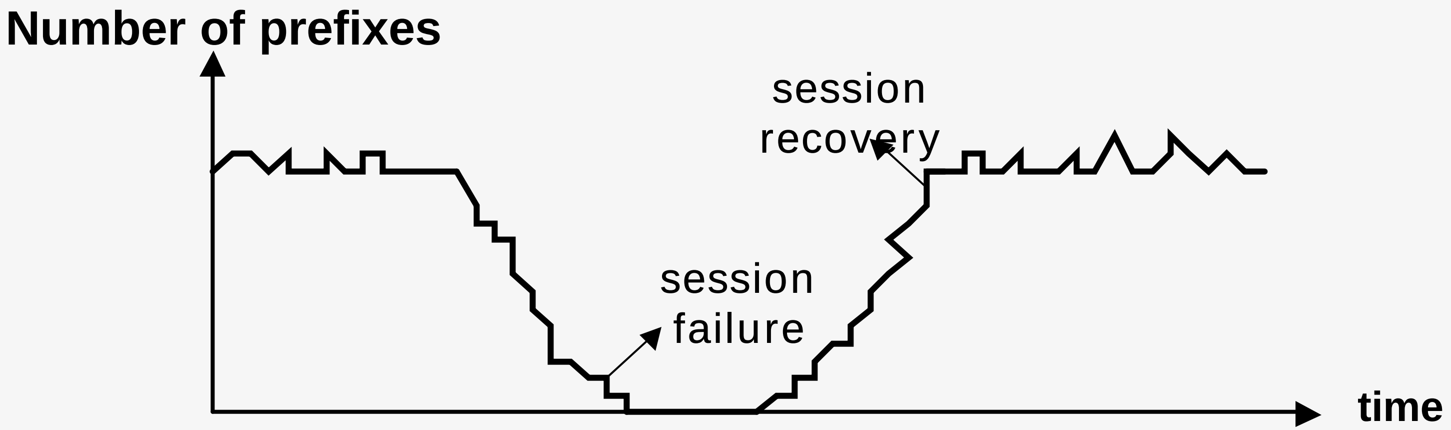
Challenge: A single routing change

- affects multiple destination prefixes

Solution:
group the same-type, close-occurring events

EBGP Session Reset

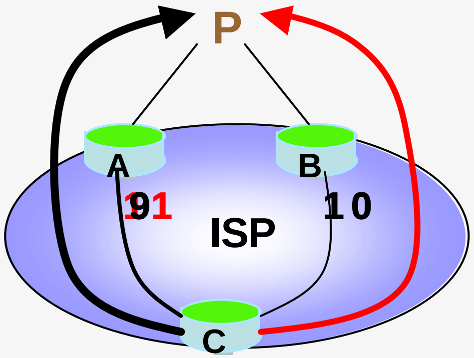
- Caused most of “single external disruption” events
- Check if the number of prefixes using that session as the best route changes dramatically



- Validation with Syslog router report (95%)

Hot-Potato Changes

- Hot-Potato Changes

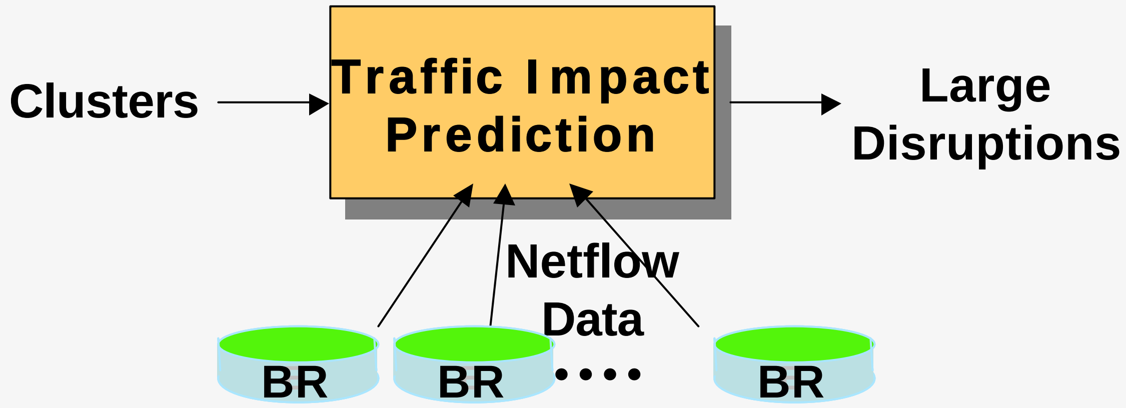


"Hot-potato routing" = route to closest egress point

- Caused "internal disruption" events
- Validation with OSPF measurement (95%) [Teixeira *et al* – SIGMETRICS' 04]

Traffic Impact Prediction

Challenge: Routing changes have different impacts on the network which depends on the popularity of the destinations



Solution: weigh each cluster by traffic volume

Traffic Impact Prediction

- Traffic weight
  - Per-prefix measurement from netflow
  - 10% prefixes accounts for 90% of traffic
- Traffic weight of a cluster
  - the sum of “traffic weight” of the prefixes
  - A small number of large clusters have large traffic weight
  - Mostly session resets and hot-potato changes

Performance Evaluation

- Memory
  - Static memory: “current routes”, 600 MB
  - Dynamic memory: “clusters”, 300 MB
- Speed
  - 99% of intervals of 1 second of updates can be process within 1 second
  - Occasional execution lag
  - Every interval of 70 seconds of updates can be processed within 70 seconds

Measurements were based on 900MHz CPU

Conclusion

- BGP troubleshooting system
  - Fast, online fashion
  - Operators' concerns (reachability, flapping, traffic)
  - Significant information reduction
    - millions of update → a few dozens of large disruptions
- Uncovered important network behavior
  - Hot-Potato changes
  - Session resets
  - Persistent-flapping prefixes