Packet Design

Traffic & Peering Analysis

or how I learned to stop worrying and love route hijacking

Pete Crocker
pete@packetdesign.com

Agenda

- Alternate methods of traffic / peering analysis
- Traffic Matrices – Pros & Cons
- Route Flow Fusion
- Solving Traffic Analysis Scaling Problems
- Troubleshooting Global Routing Events
- Conclusion & Acknowledgements

Packet Design

Delivering Traffic with Service Assurance

- Traffic measurements
  - Traffic volume measurements
    - e.g. link utilizations to identify congested links
  - Traffic demand measurements
    - move x bytes from a to b

- Netflow Analysis
  - Dictates where and how traffic flows through network

Classical Traffic Analysis

- Interface byte and packet counters
  - Near real time interface utilization statistics
  - Counters say nothing about who they came from, where they're going or why they're there
- City (or PoP) pair byte and packet counters
  - Counters on ATM circuits or on MPLS tunnels
  - An approximation of traffic demands
  - Useful for IGP tuning, not as useful for inter-domain

Traffic Analysis via netflow

- Traffic Flows
  - Routers sample packets (1 in 1000 typical)
  - Flow is identified by interface number, source/destination addresses, ports, CoS/Diffserv
  - Router maintains a byte count of each flow until timeout or tcp fin/rst, then exports it to a collector
  - Useful for diagnostics, e.g. who is causing congestion, where it came from, which application
  - Can be used to generate various traffic demands for both intra- and inter-domain tuning

Traffic Matrices

- Routing tells which flow (i.e. how much traffic) is going to which:
  - neighbour
  - destination
  - transit AS
  - BGP community, etc.
- Current technologies combine Netflow w. BGP to create matrices
- Off-The-Shelf Products
  - Compuware (formerly Adlex Flowtracker)
  - Network Signature BENTO

(Part of) A customer-transit outbound traffic matrix

(% of total traffic)	level3	cogent	qwest	wiltel	row total
All	63.9	18.3	16.9	0.5	
ucb	5.5	3.4	1.8	0	10.7
ucla	7.4	1.0	1.3	0.2	9.9
ucsd	5.9	0.5	1.4	0.1	7.9
csunet	4.7	1.4	1.2	0	7.4

- Outbound transit traffic demand of top four (of 112) customers.
- Table was automatically computed from 24 hours of CENIC Netflow and BGP data.

* All data shown in this presentation is courtesy of CENIC (www.cenic.org) and used by permission.

What a Traffic Matrix Can’t Do

- A traffic matrix is primarily an engineering tool. It’s an O(n^2) analysis that’s extremely useful for optimization & capacity planning but:
  - Can’t answer operational questions like “who filled up this link?” (requires an O(n^3) “A to B via C” analysis).
  - Can’t answer strategic planning questions like “where does customer traffic go when it leaves here?” (requires an O(n!) path analysis).

- Conventional wisdom says this scaling makes most operational & strategic questions too expensive to answer. But conventional wisdom is wrong ...

Packet Design

Route-Flow Fusion

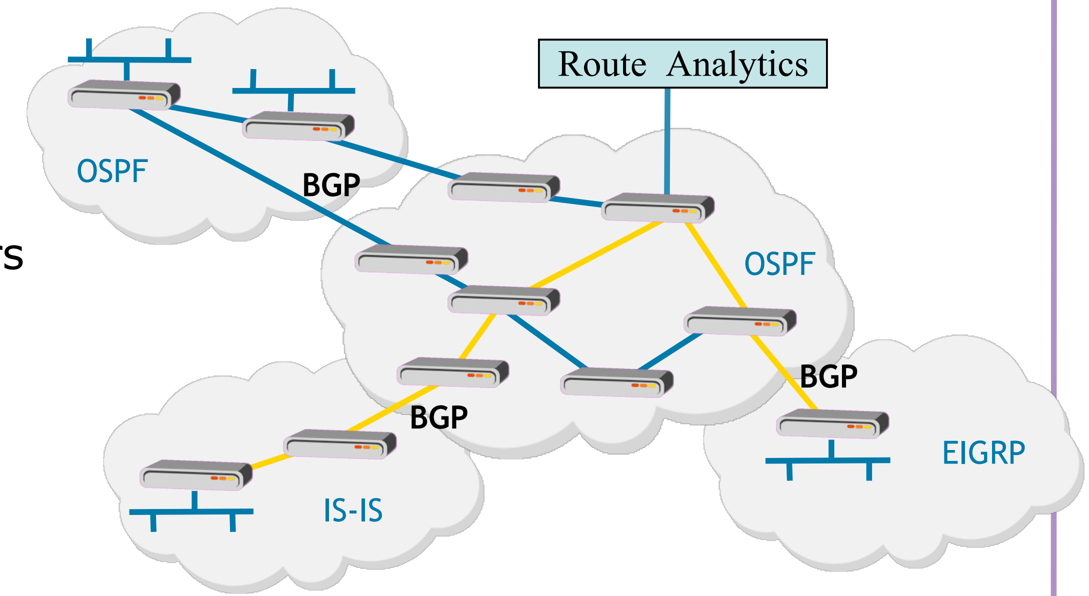
- Separately record route (IGP & BGP) and flow information
- Demand-driven data fusion of route and flow information
- Result gives you aggregate data rate & traffic volume induced by selected flows on each link they traverse
- Also, allows modification of the routing model to see how traffic is affected

Route-Flow Fusion Uses

- Converts point measurements to path measurements
- Helps scale flow data storage
- Routing provides natural aggregation boundaries for various traffic measurements

Route Analytics
Collects Routing Data

- Listen passively to routing updates
- Create a real-time network map
  - As up to date as routers
- Analyze paths
  - Paths are computed using the same procedures as routers
- A historical view with breakdown of instability
  - Full routing event history/forensic audit trail
  - Flapping links, prefixes
  - Ability to look at state of routing at any point in recorded history



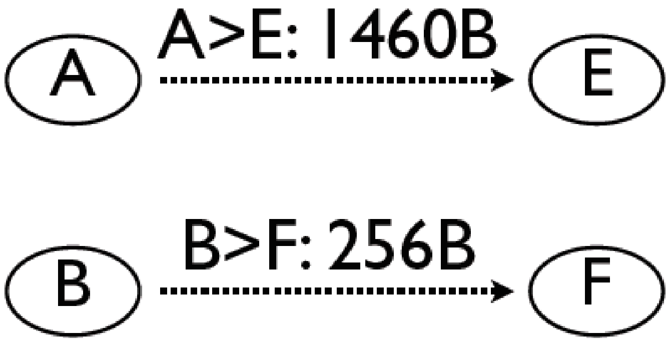
Works with all types of protocols:
- OSPF, IS-IS, EIGRP
- BGP, RFC 2547bis / RFC 4026

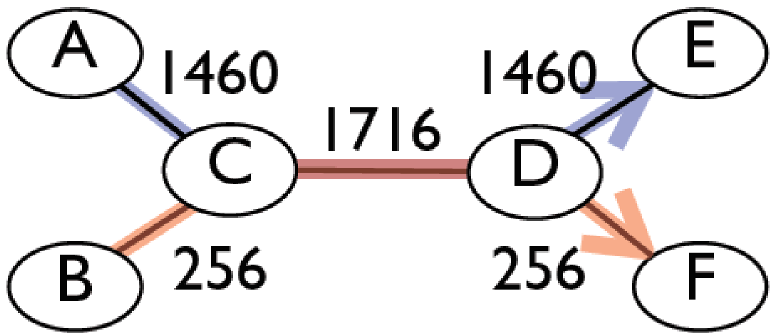
Why Routing is So Helpful

- Routing contains all the meta-information needed to classify and aggregate flow information:
  - IGP prefix and BGP prefix & last-hop AS# maps source and dest addresses to higher level units (network, organization, etc.).
  - BGP first-hop AS# identifies customers, transit providers & peers (BGP community attributes tell you which is which).
  - IGP & BGP next-hop show where external entities attach to internal topology.

Path measurements

- Routing maps a point measurement to a path measurement





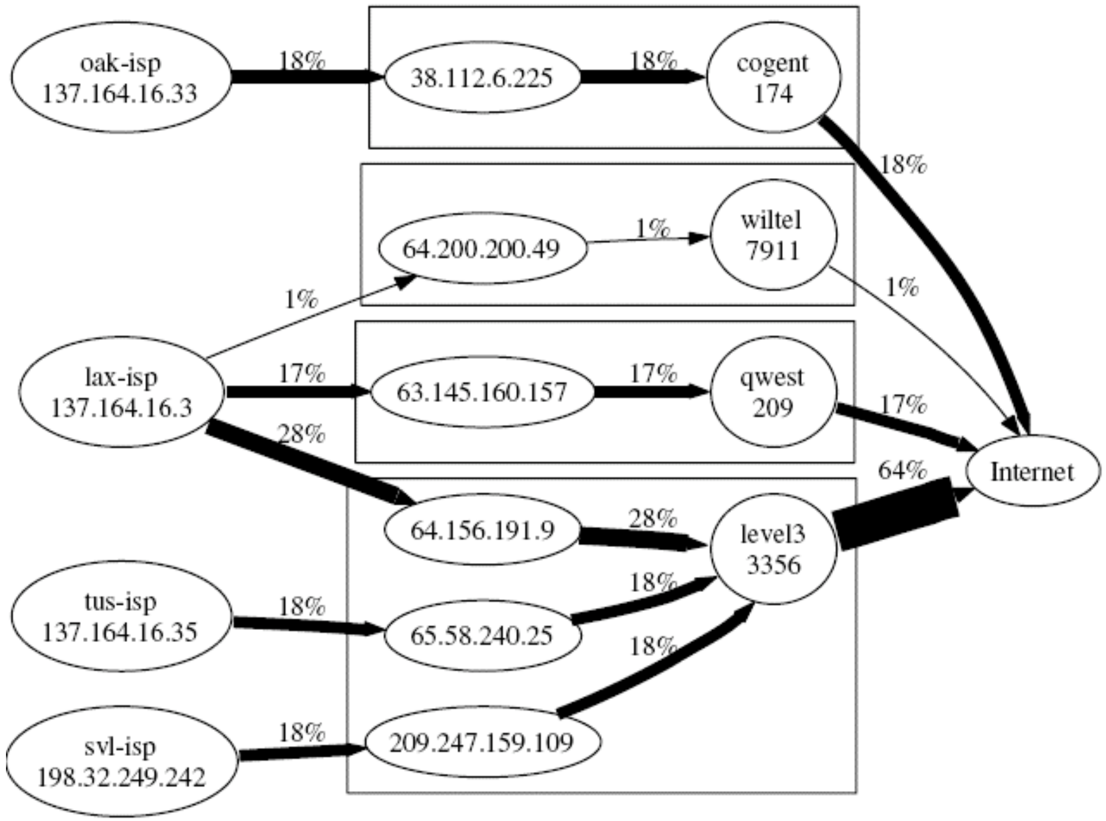
- Routing also maps the reverse, which flows are on link C-D

Scaling Flow Storage

- Routing works at prefix level
- Aggregate flows into “macro” flows from the source prefix to the destination prefix
  - Does not change traffic demand matrices
- Perhaps maintain longer prefixes for large flows so that the new prefixes can be created to divide these mega flows

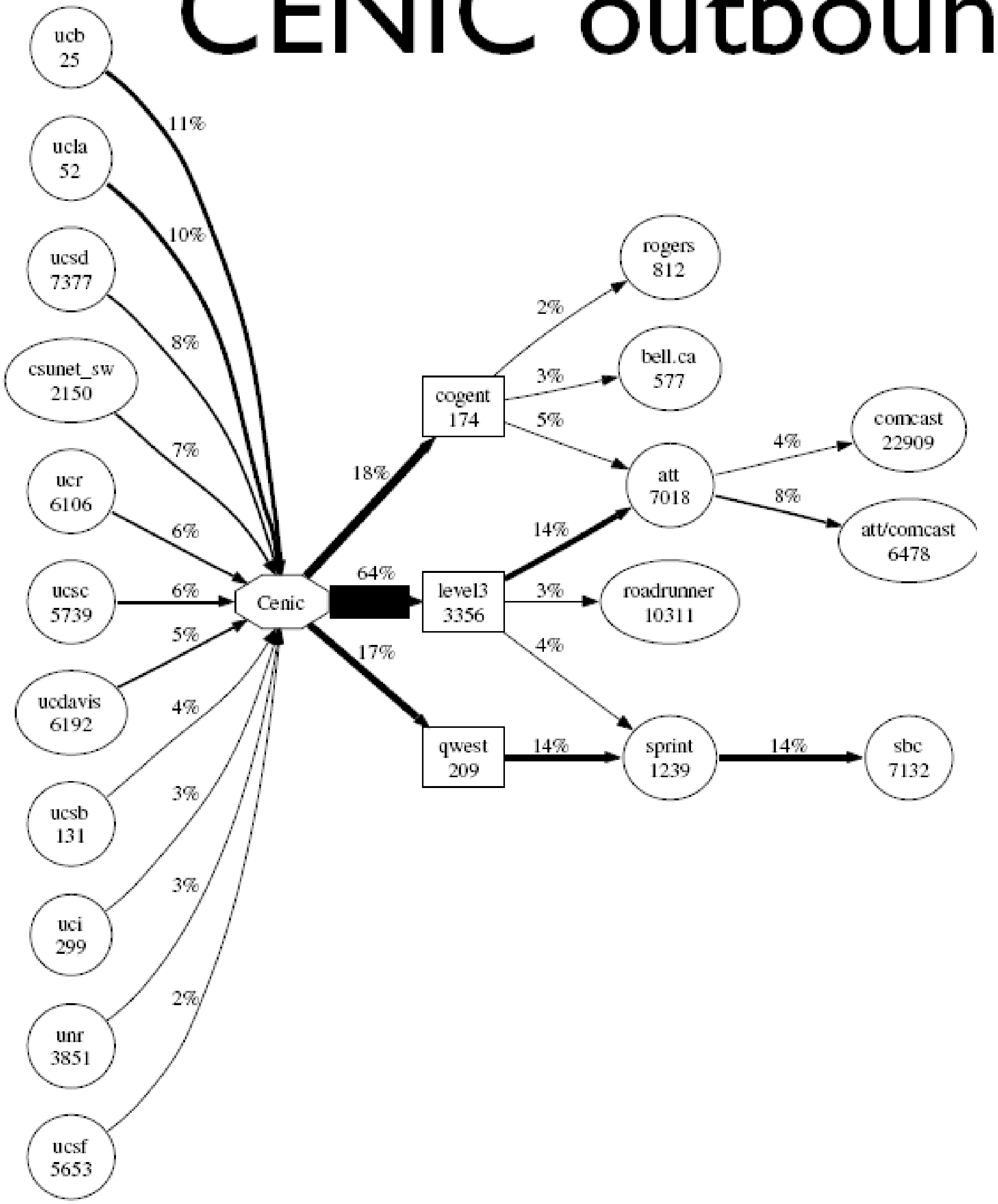
Solving the Scaling Problem

- Basis of scaling problem:
  - The number of places where data *might go* is huge. But at any particular time the number of places where it actually *does go* is small.



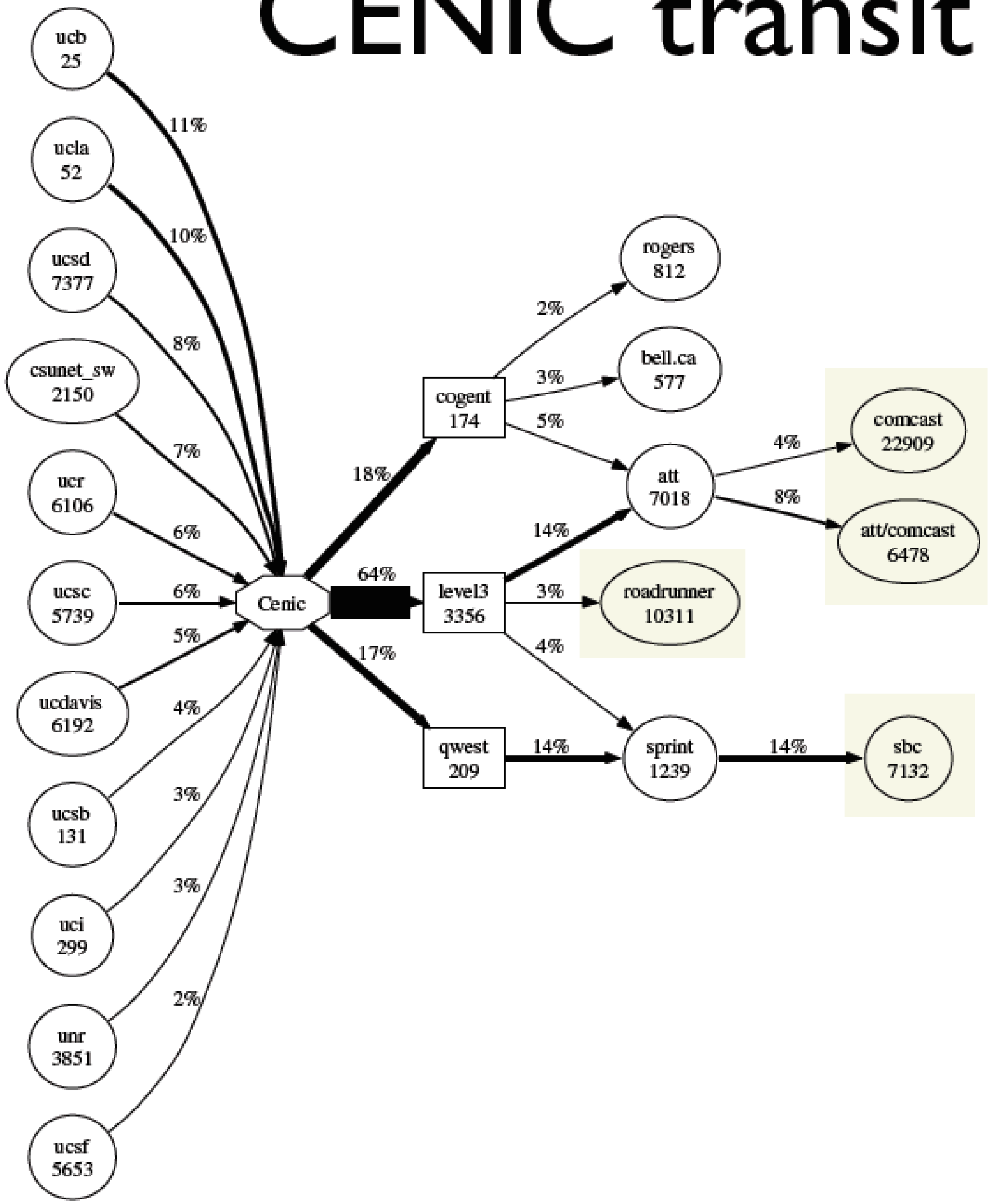
CENIC transit traffic

CENIC outbound transit traffic



- Same data as the traffic matrix on slide 7
  - Customers are on the left
  - Transit providers are the rectangles
  - Edge thickness shows traffic volume
  - Edges carrying less than 1% of the traffic are pruned
- The **only** manual input needed to create this picture were two BGP community tags (customers and transits)

CENIC transit traffic (cont.)



The computational cost of this view of the data is the same as a traffic matrix but this contains more operational and business information.

For example, note that a third of the total traffic goes to residential providers (comcast, roadrunner, sbc) or that 80% of the traffic sent to qwest is destined for sbc.

Additional Benefits

- Route Analytics allow modelling routing changes
  - link/routing failures
  - adding links / routers / bgp-peerings
  - local-pref/med changes, as path prepending, etc.
- Route-Flow fusion shows the effects of these changes on traffic
  - what link failure causes the most congestion
  - can I save $$$ by direct peering with AS X

Packet Design

Troubleshooting Global Routing Problems

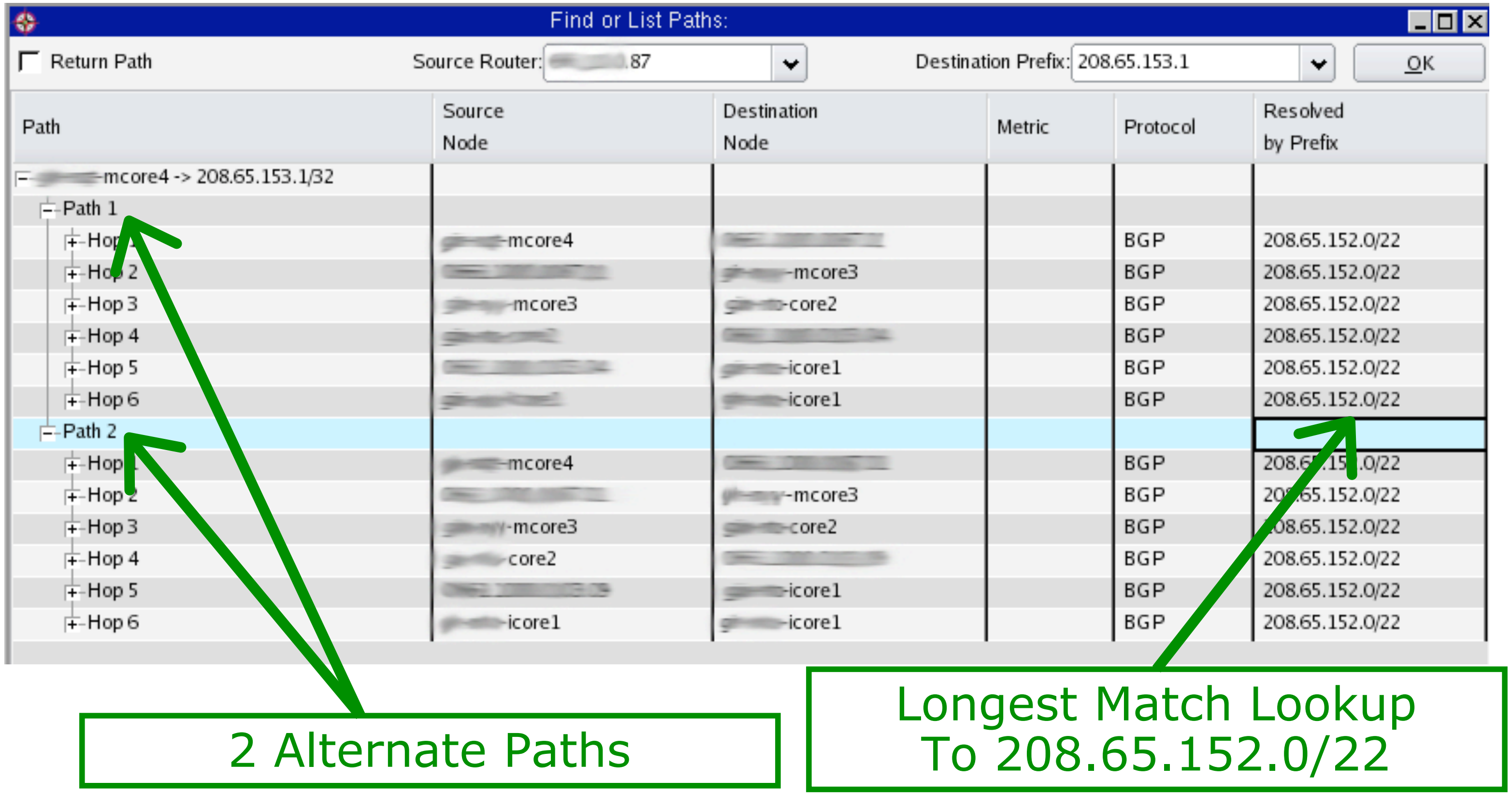
- YouTube incident on February 24, 2008
- Major drop in traffic to my network
- What happened?

More Specific Announcement

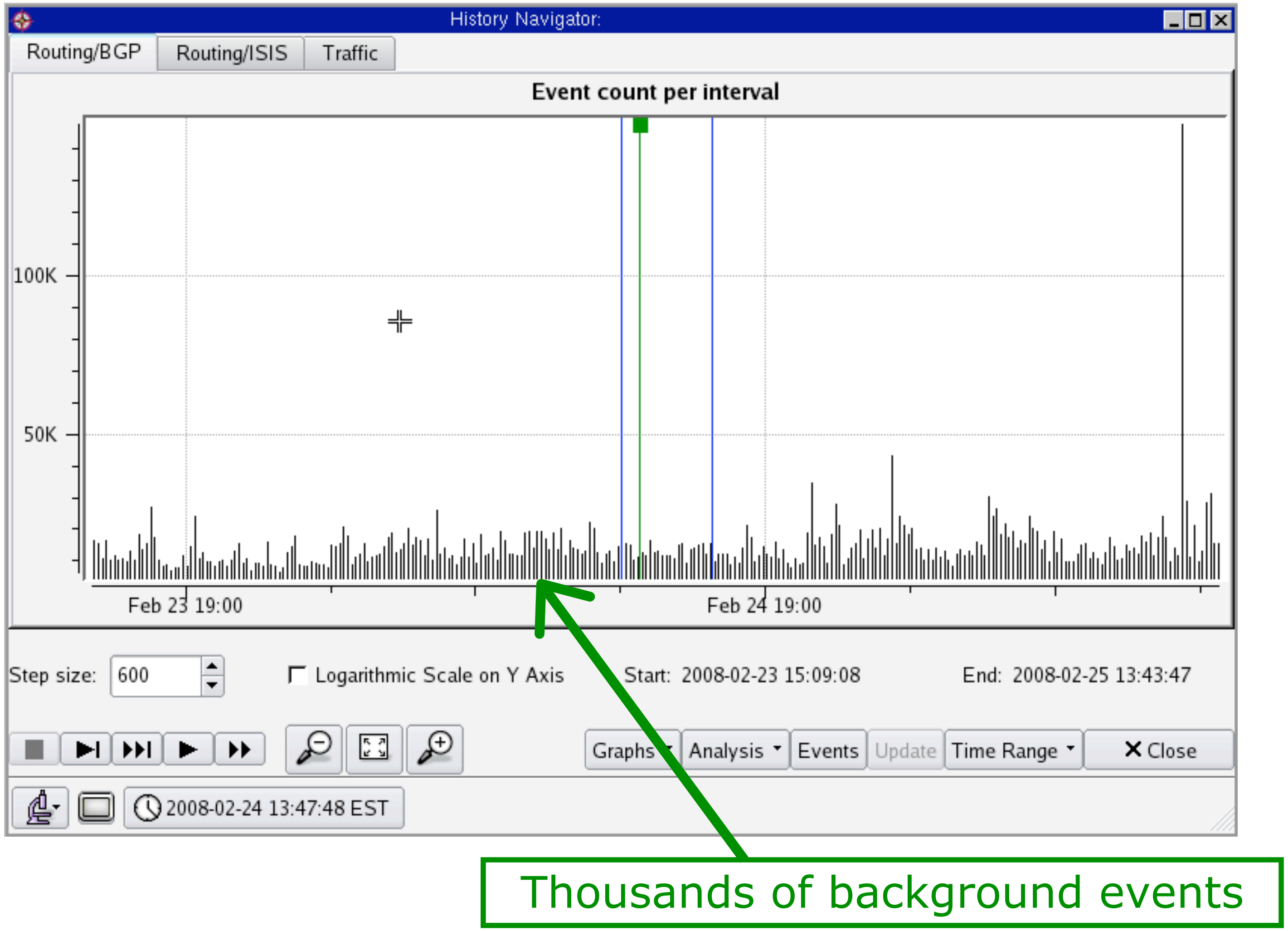
- Pakistan Telecom (AS17557) announces 208.65.153.0/24 at 18:47:48 UMT

Normal Routing to YouTube

- 2 Redundant Paths to a /22

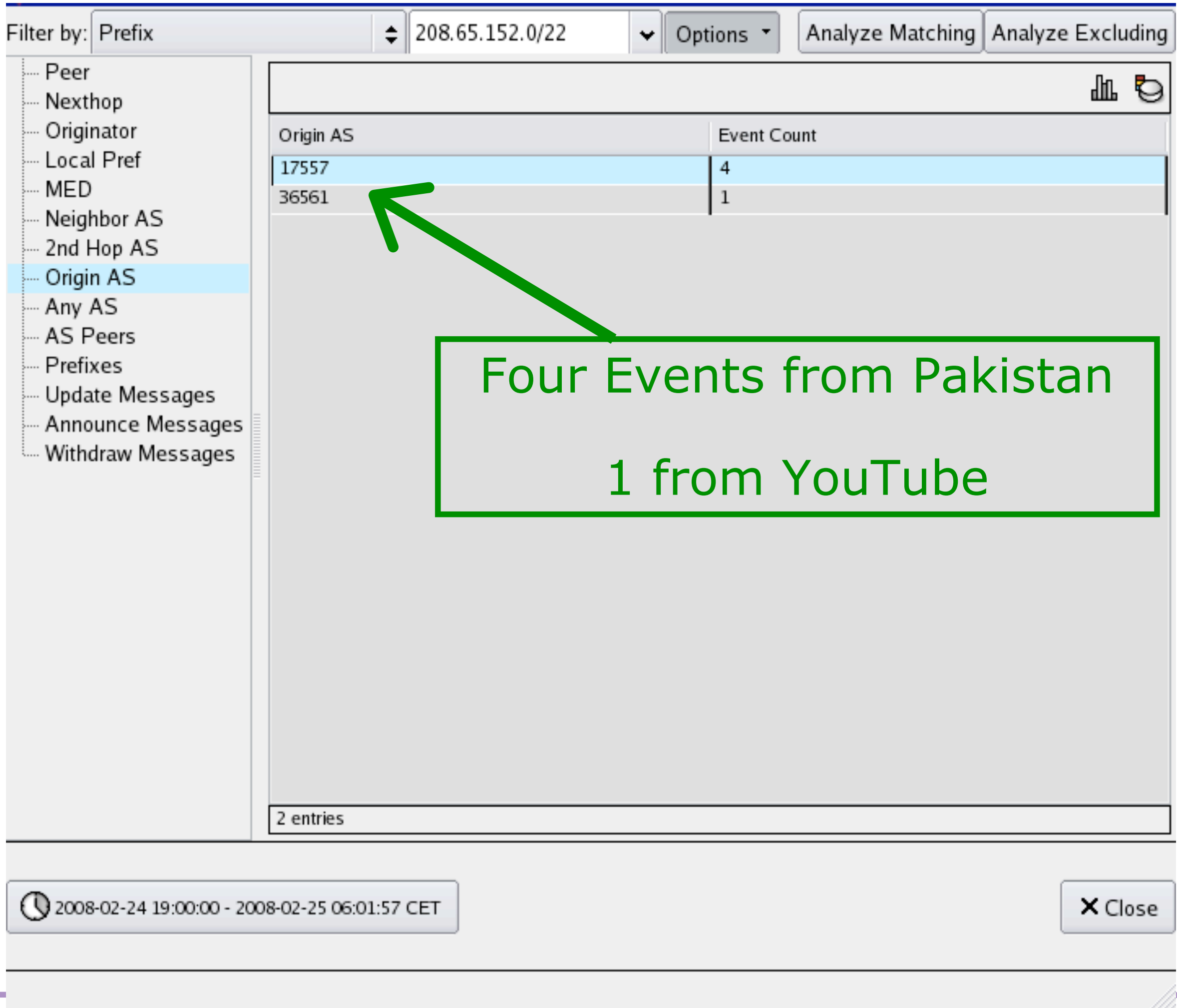


Background Noise

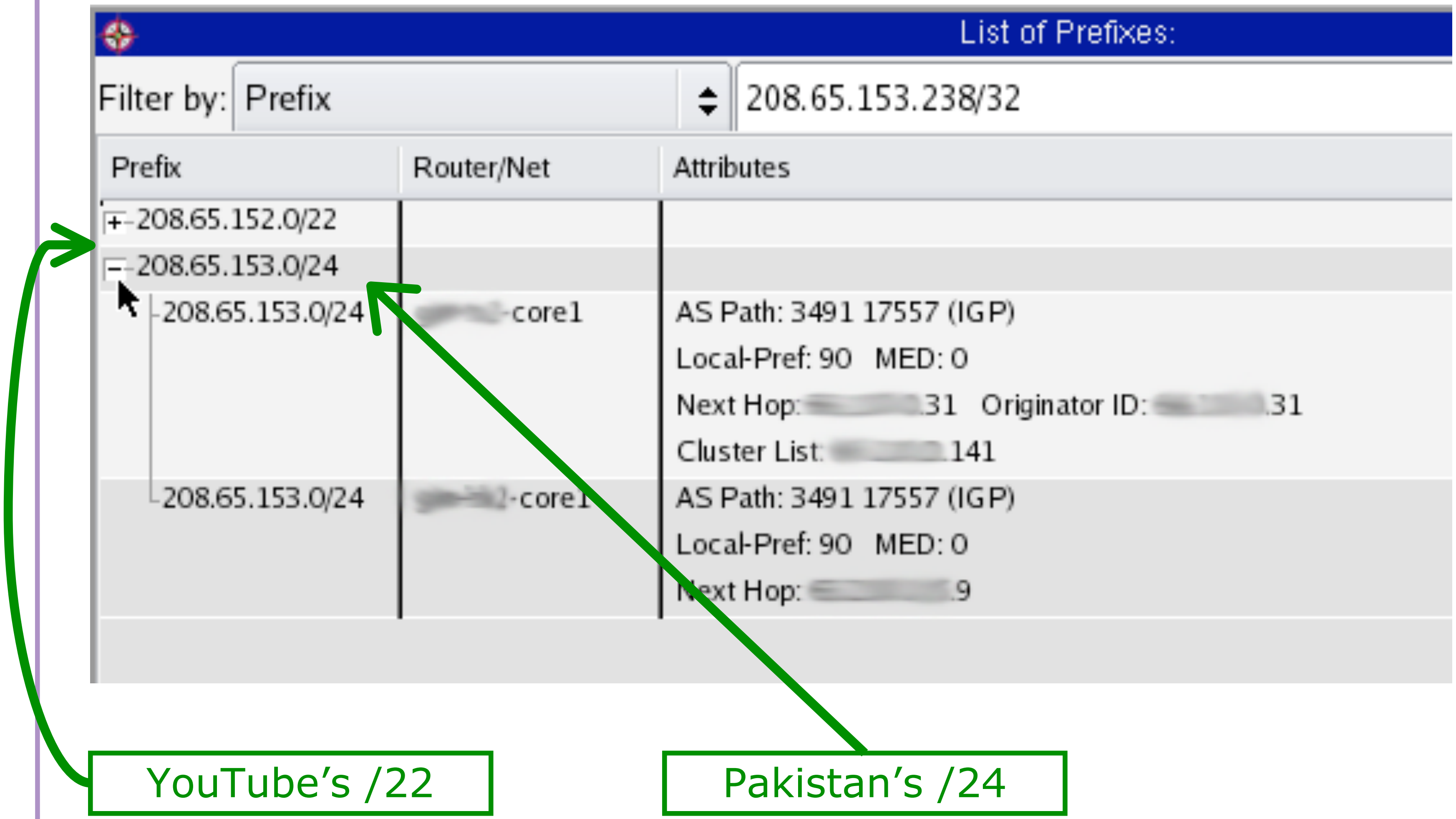


Events

- We can quickly narrow down the events:



Routes After



Single Path After

- Traffic

Find or List Paths:

Return Path — Source Router: [illegible].87 — Destination Prefix: 208.65.153.1 — OK

Path	Source Node	Destination Node	Metric	Protocol	Resolved by Prefix
[illegible]-mcore4 -> 208.65.153.1/32					
Hop 1	[illegible]-mcore4	[illegible]-mcore4		BGP	208.65.153.0/24
Hop 2	[illegible]-mcore4	[illegible]-core1		BGP	208.65.153.0/24
Hop 3	[illegible]-core1	[illegible]-core3		BGP	208.65.153.0/24
Hop 4	[illegible]-core3	[illegible]-core2		BGP	208.65.153.0/24
Hop 5	[illegible]-core2	[illegible]-core1		BGP	208.65.153.0/24
Hop 6	[illegible]-core1	[illegible]-core1		EBGP	208.65.153.0/24

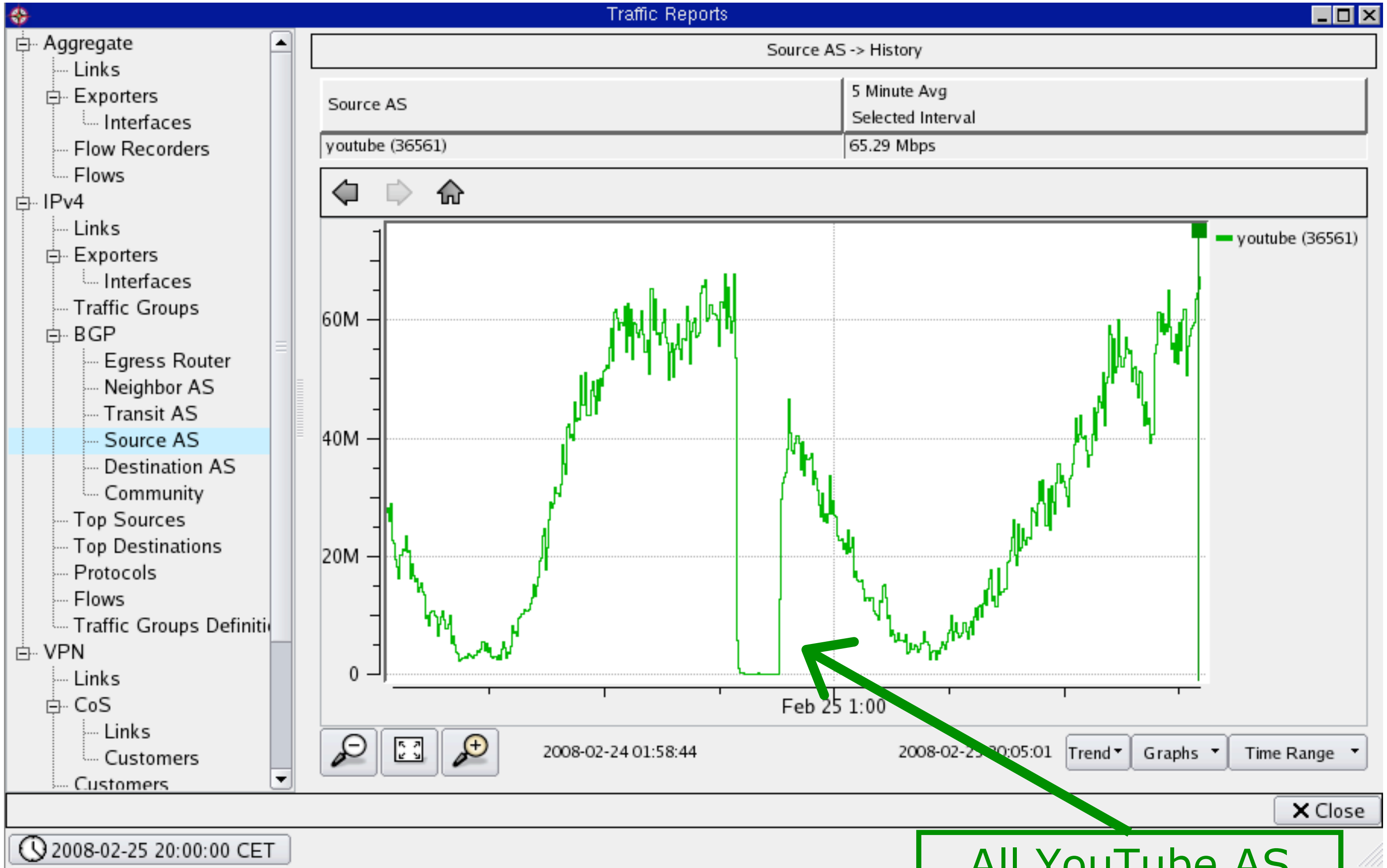
Traceroute to YouTube
Now Matches /24 route

Routing Events

Time	Router	Operation	Neighbor/ Prefix	Attributes
2008-02-24 13:47:48.697005 Pakistan Announces /24	[redacted].141	New Announce	208.65.153.0/24	AS Path: 3491 17557 (IGP) Local-Pref: 90 MED: 0 Next Hop: [redacted].31 Originator ID: [redacted].31 Cluster List: [redacted].141
2008-02-24 13:47:52.991634 Learned from other RR	[redacted].141	Announce	208.65.153.0/24	AS Path: 3491 17557 (IGP) Local-Pref: 90 MED: 0 Next Hop: [redacted].63 Originator ID: [redacted].63 Cluster List: [redacted].141
2008-02-24 15:51:05.457009 Pakistan Withdrawals /24	[redacted].141	Withdraw	208.65.153.0/24	AS Path: 3491 17557 (IGP) Local-Pref: 90 MED: 0 Next Hop: [redacted].63 Originator ID: [redacted].63 Cluster List: [redacted].141
2008-02-24 15:51:05.516320 Pakistan Re-Announces /24	[redacted].141	New Announce	208.65.153.0/24	AS Path: 3491 17557 (IGP) Local-Pref: 90 MED: 0 Next Hop: [redacted].31 Originator ID: [redacted].31 Cluster List: [redacted].141
2008-02-24 15:51:28.604227 Pakistan Withdrawals /24	[redacted].141	Withdraw	208.65.153.0/24	AS Path: 3491 17557 (IGP) Local-Pref: 90 MED: 0 Next Hop: [redacted].31 Originator ID: [redacted].31 Cluster List: [redacted].141

Return Traffic from YouTube

Packet Design



All YouTube AS
Return Traffic
Stops for 2 hours

Packet Design

Solving Peering Management Challenges

- BGP’s configuration parameters are relatively coarse-grained
- Very difficult to understand all the effects of a peering change
- Route-flow fusion provides thorough analysis of the effect of peering changes on:
  - link utilizations
  - destination AS, BGP community and exit router traffic
- Identifies any potentially harmful, collateral traffic shifts that would occur as a result of a new peering

Packet Design

Conclusions

- Route Flow Fusion provides much more operationally relevant insight than simple traffic matrices
- Provides network-wide view of link utilization, traffic composition (by CoS, Service, per MPLS VPN), while only requiring flow record collection from edges of the network

Acknowledgments

- CENIC for allowing to share these results
- I would also like to thank Van Jacobson, Cengiz Alaettinoglu, Haobo Yu, Bruce Mah, and Alex Henthorn-Iwane for their contributions to this talk.