

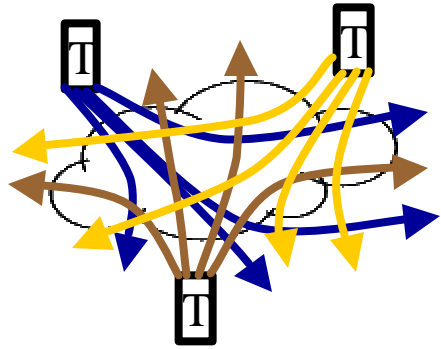
Measuring ISP topologies with Rocketfuel

Ratul Mahajan

Neil Spring

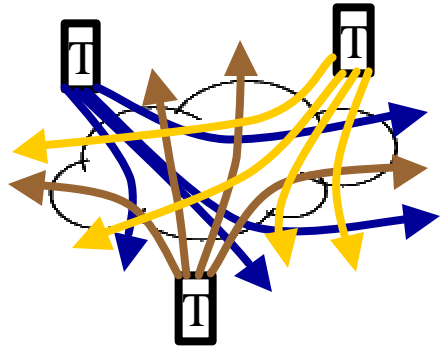
David Wetherall

University of Washington



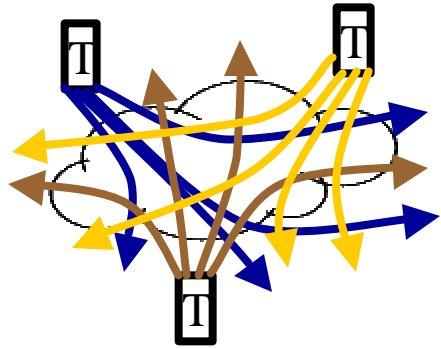
What are we after?

- ◆ High fidelity router-level map of ISPs
 - Routers
 - backbone
 - gateway
 - directly connected customers and peers
 - Links between these routers
 - Understanding the structure of backbone and PoPs
- ◆ ISP vs AS
 - customer networks not included
 - dialup and broadband access networks not included



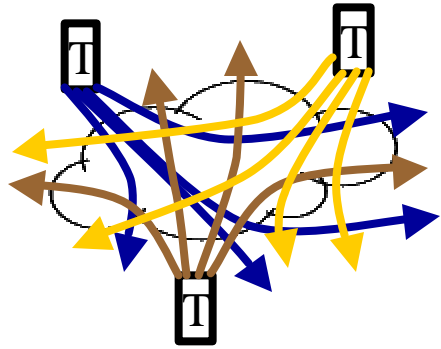
Why measure?

- ◆ Network topologies closely guarded
- ◆ Inaccurate or incomplete official versions
- ◆ Want higher fidelity than current mapping efforts
 - not just backbone
 - we found 7 times more routers and links than Skitter by focusing on one ISP at a time



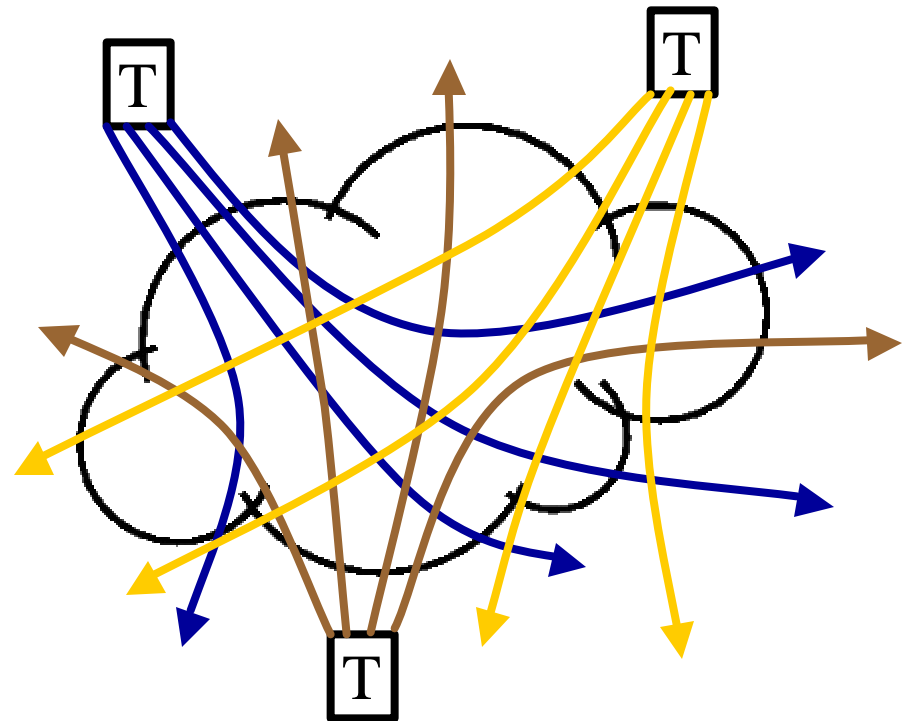
But, what is it good for?

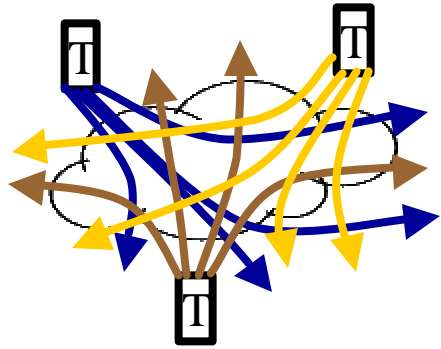
- ◆ Realistic topologies for design and simulation
 - e.g., studying intra-domain routing protocols
- ◆ Generating synthetic topologies
 - Synthetic ISP maps + Synthetic AS-level map = Synthetic Internet map
- ◆ Understanding design and engineering principles behind these networks
- ◆ Understanding how ISPs connect to customers and other ISPs



Methodology

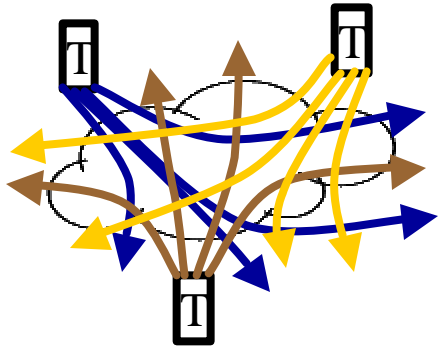
- ◆ Traceroute
 - known limitations
 - but pretty much the only tool
 - need multiple vantage points
- ◆ Public traceroute servers
 - 1000s of servers out there
 - we used 300 servers that provided 800 vantage points
 - no deployment required





Challenges

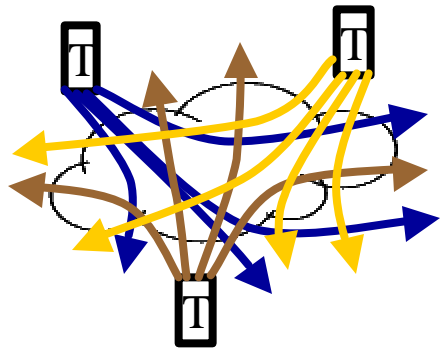
- ◆ Software engineering
 - understanding the language spoken by each server
 - system architecture
- ◆ **Reducing measurements**
 - traceroute servers are a public resource
 - network admins do not like getting too many probe packets
 - to make it even possible
- ◆ Identifying ISP routers
- ◆ Alias resolution
- ◆ Geographical and role information of routers



Reducing measurements (making it possible)

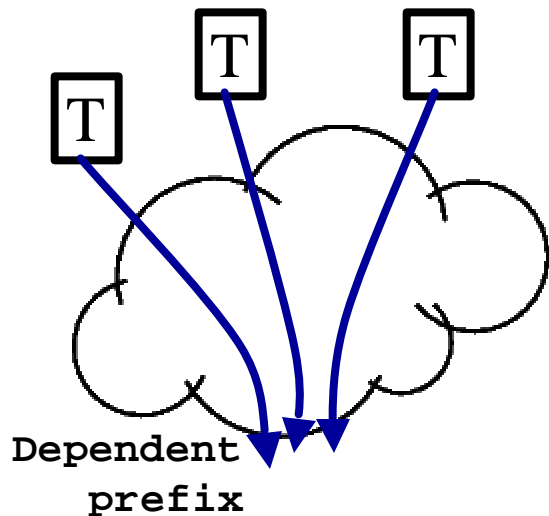
- ◆ Goal: reduce the number of traces as much as possible with as little loss of detail as possible
- 1. Identify traces that are likely to transit the ISP
 - *BGP directed probing*: leverage information in the BGP routing tables
- 2. Identify likely duplicate traces
 - *Path reductions*: exploit properties of IP routing

Another helpful trick: find a live host within a prefix to speed up the traceroute



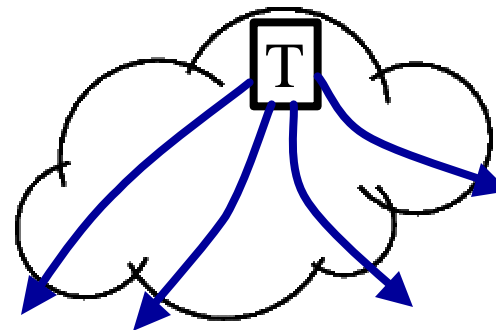
BGP directed probing

Using the BGP tables identify traces that are likely to transit the ISP

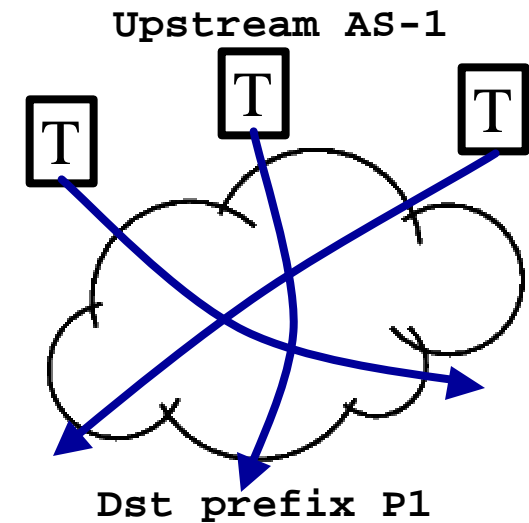


1. Traces from anywhere to *dependent prefixes*

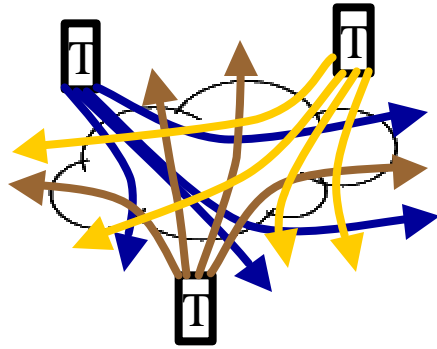
Insider vantage point



2. Traces from *insiders* to anywhere

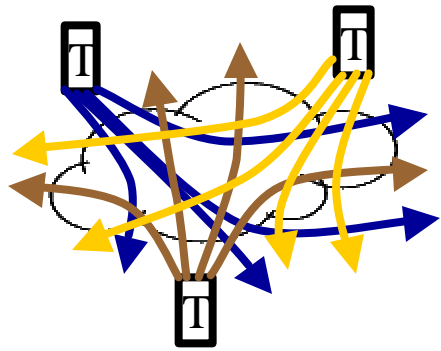


3. Traces from upstream ASes that use the ISP to reach some prefix



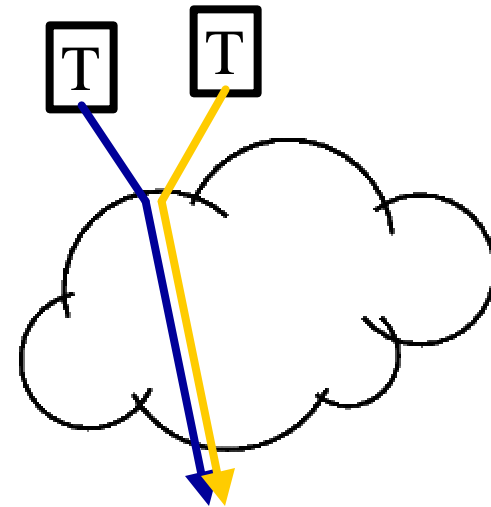
Path reductions

- ◆ Interested only in part of the trace that traverses the ISP
- ◆ Avoid duplicate traces
- ◆ Exploit properties of IP routing to determine if paths through the ISP is likely to be identical
 1. Ingress reduction
 2. Egress reduction
 3. Next-hop AS reduction



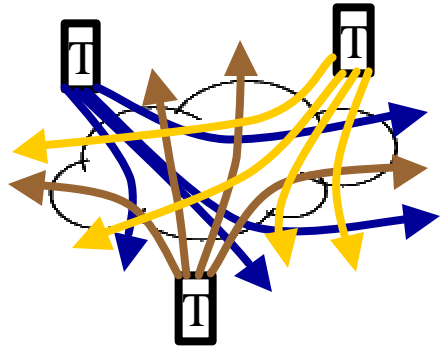
Ingress reduction

- ◆ Path taken by a packet in the network is dependent only on the destination
- ◆ Shared ingress points



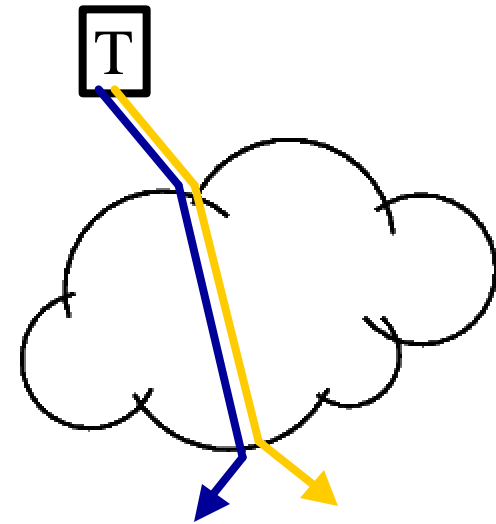
Ingress reduction: when two traces to the same destination enter the ISP at the same point, only one of them is required

Also helps to load balance across traceroute servers

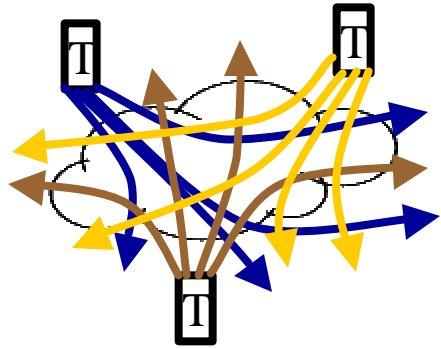


Egress reduction

- ◆ Dependent subprefixes attach to the ISP at a unique point
- ◆ Some egresses would be shared
- ◆ Requires *egress discovery*
 - finding possibly multiple egresses for a dependent prefix
 - done locally

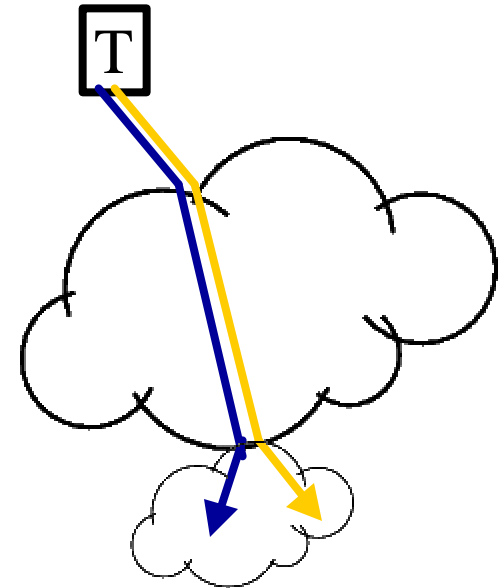


Egress reduction: when two traces from the same ingress to destinations that attach to the ISP at the same point, only one of them is required

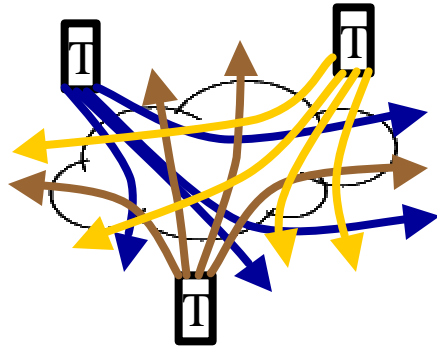


Next-hop AS reduction

- ◆ “Hot potato” routing
 - traces from same ingress to destinations behind the same next-hop AS exit the ISP at the same point
- ◆ Next-hop AS information present in the BGP table

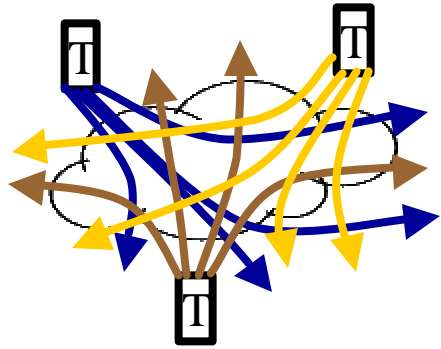


Next-hop AS reduction: when two traces from the same ingress cross over to the same next-hop AS, only one of them is required



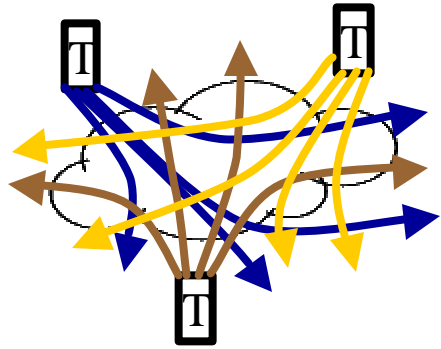
Measurement reduction summary

- ◆ BGP directed probing
 - every vantage point to dependent prefixes
 - insiders to all prefixes
 - upstream ASes to downstream prefixes
- ◆ Path reductions (composable)
 - ingress reduction
 - egress reduction
 - next-hop AS reduction
- ◆ Helped reduce the number of measurements by three orders of magnitude



Challenges

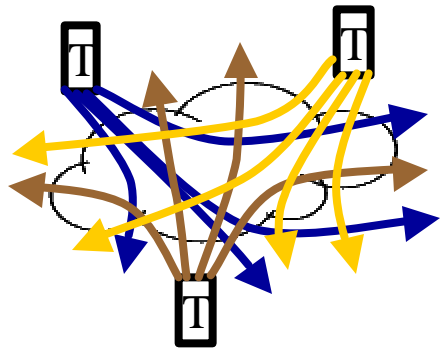
- ◆ Software engineering
 - send request to each server in a language it understands
 - understanding their replies
- ◆ Reducing measurements
 - traceroute servers are a public resource
 - network admins do not like getting too many probe packets
 - to make it even possible
- ◆ Identifying ISP routers
- ◆ Alias resolution
- ◆ Geographical and role information of routers



Identifying ISP routers

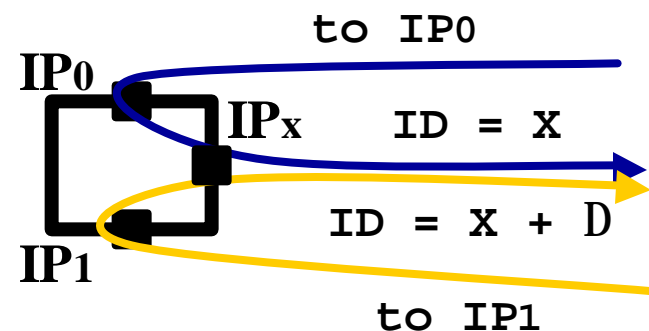
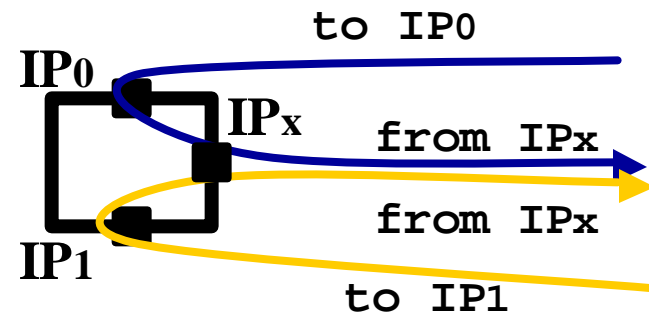
What part of the trace belongs to the ISP?

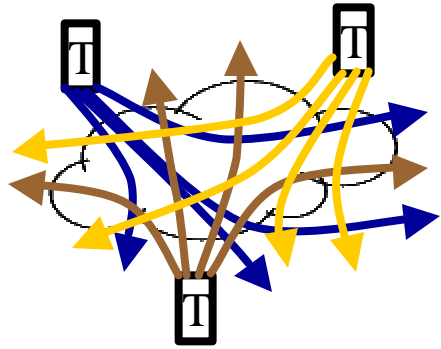
- ◆ Use DNS names
 - all Verio routers are *.verio.net
 - prune out cable modems, DSL and dial-up modem pools
 - customer networks have different DNS names
- ◆ Use BGP tables in absence of names
 - address space advertized by the ISP



Alias resolution

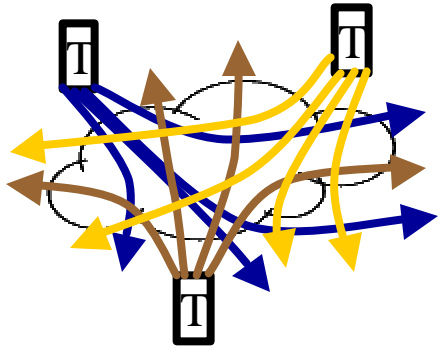
- ◆ Resolving multiple IP addresses of the same router
- ◆ Return IP address
 - the return address for packets generated by the router is that of the outgoing interface
- ◆ IP Identifier
 - closely spaced responses from the same router will have close IP ID
 - finds 3 times more aliases than return IP address approach



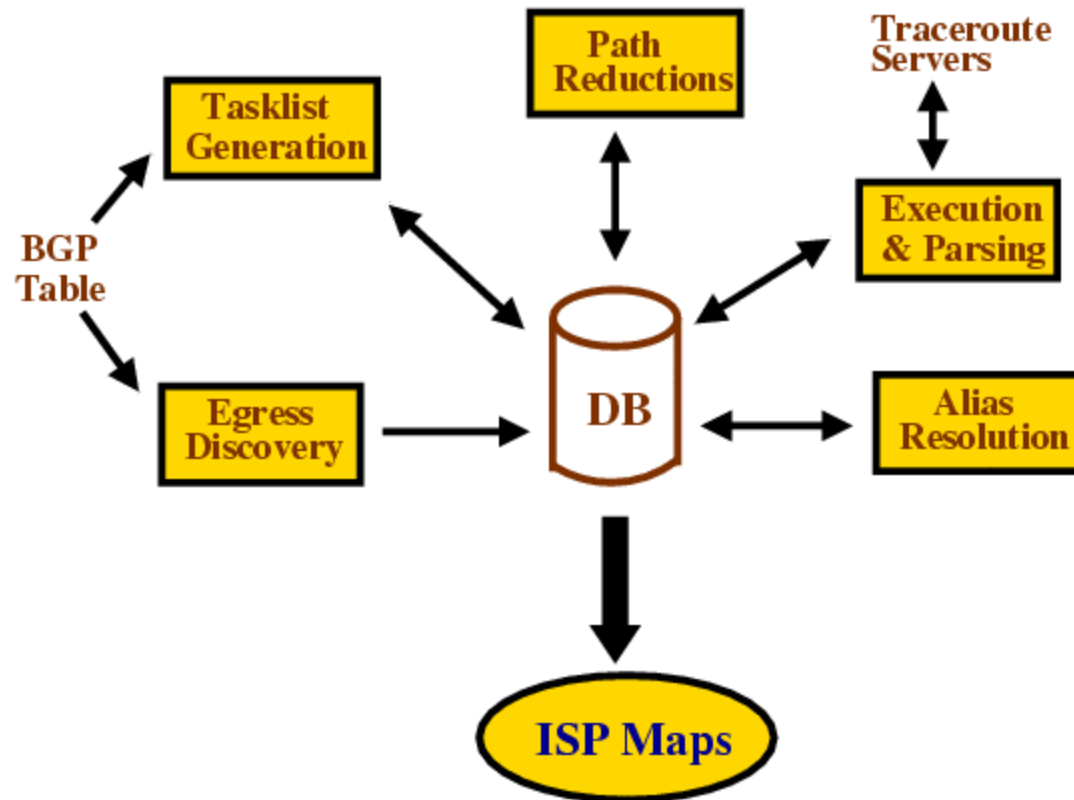


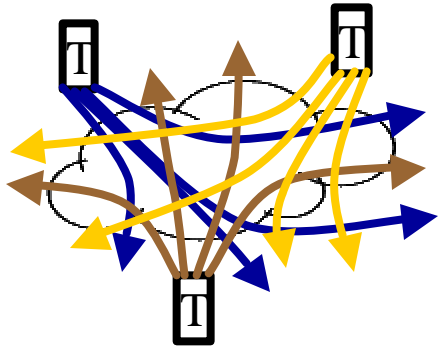
Location and role discovery

- ◆ Where is this router located?
 - use DNS names
 - **s1-bb11-nyc-3-0.sprintlink.net** is a Sprint router in New York City
 - use connectivity information
 - if a router connects only to routers in Seattle, it probably is in Seattle
- ◆ What role does this router play in the topology?
 - only backbone routers connect to other cities
 - use DNS names
 - **s1-gw2-sea-3-1.sprintlink.net** is a Sprint gateway router



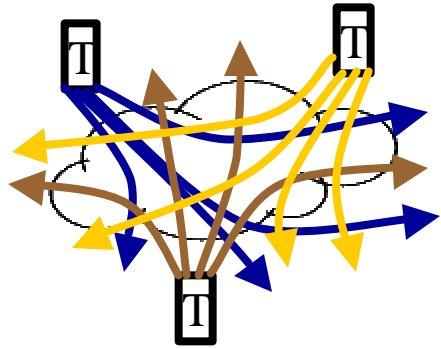
Rocketfuel architecture





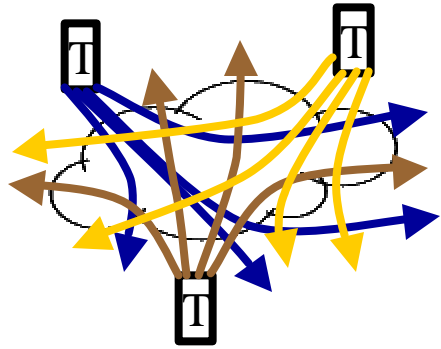
Summary statistics

ASN	Name	Routers	Links	PoPs
1221	Telstra (Aus)	4,440	4996	54
1239	Sprint	11,889	15,263	25
1755	Ebone (Eur)	438	1,192	26
2914	Verio	7,574	19,175	103
3257	Tiscali (Eur)	618	839	52
3356	Level3	2,064	8,669	44
3967	Exodus	688	2,166	22
4755	VSNL (India)	664	484	8
6461	Abovenet	843	2,667	22
7018	AT&T	13,993	18,083	109

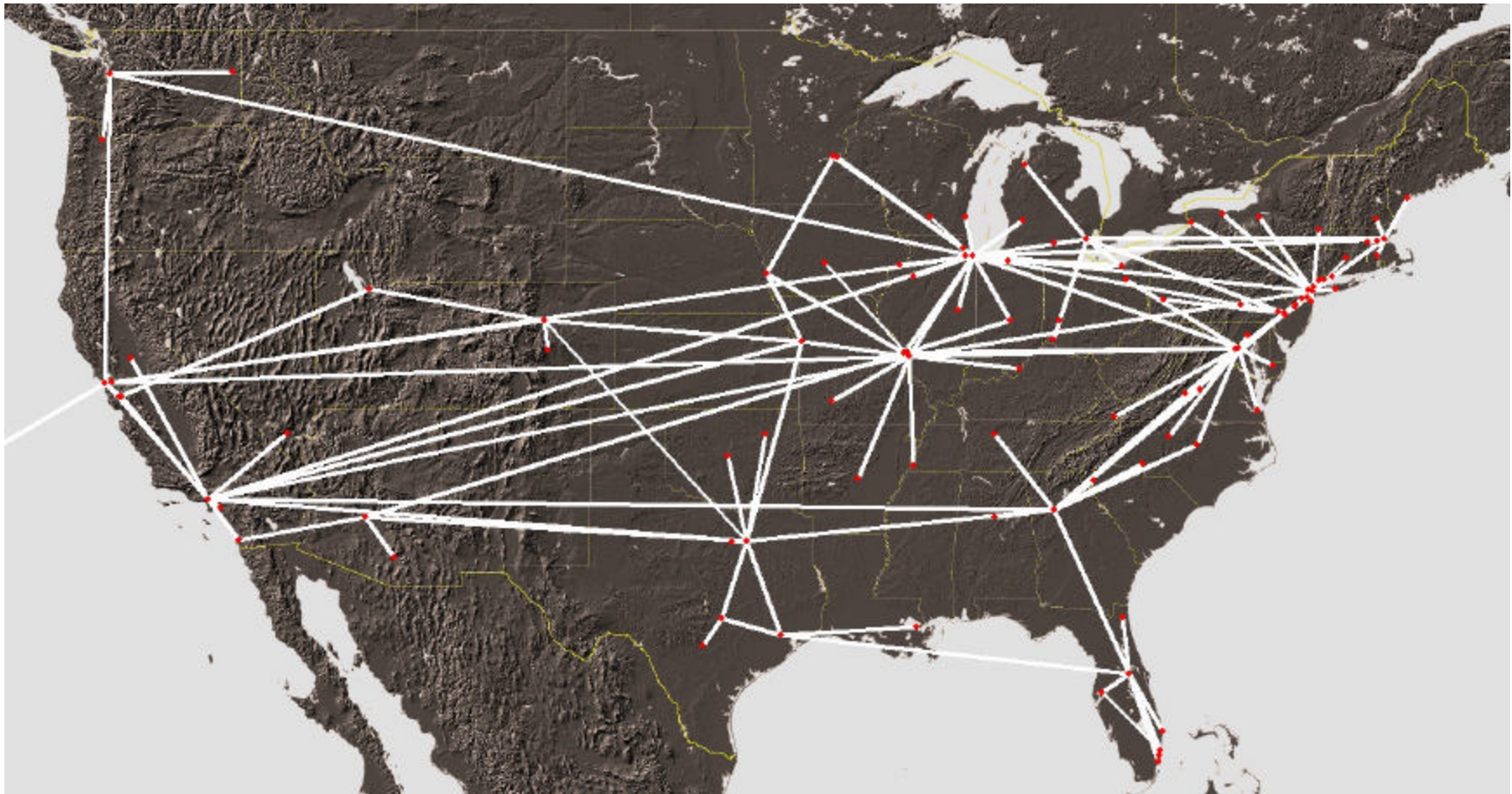


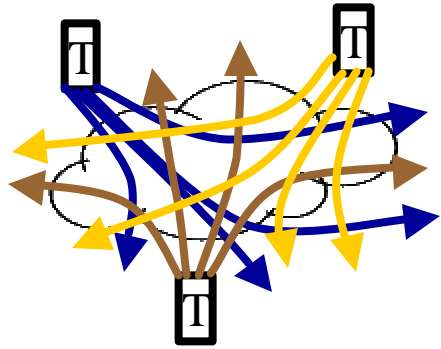
Validation

- ◆ Talk to ISPs
 - complete city-level interconnection (backbone); PoPs “looked fine”
 - spurious links (tools lie) ☹
- ◆ Comparison with peering in BGP data
 - number of router-level peerings observed
 - >90% for most ISPs; worst case was Sprint with >70%
- ◆ Comparison with Skitter
 - we have 7 times more routers and link
- ◆ Conclusions
 - our maps are fairly complete
 - very useful for qualitative inferences
 - our maps are not *authoritative*

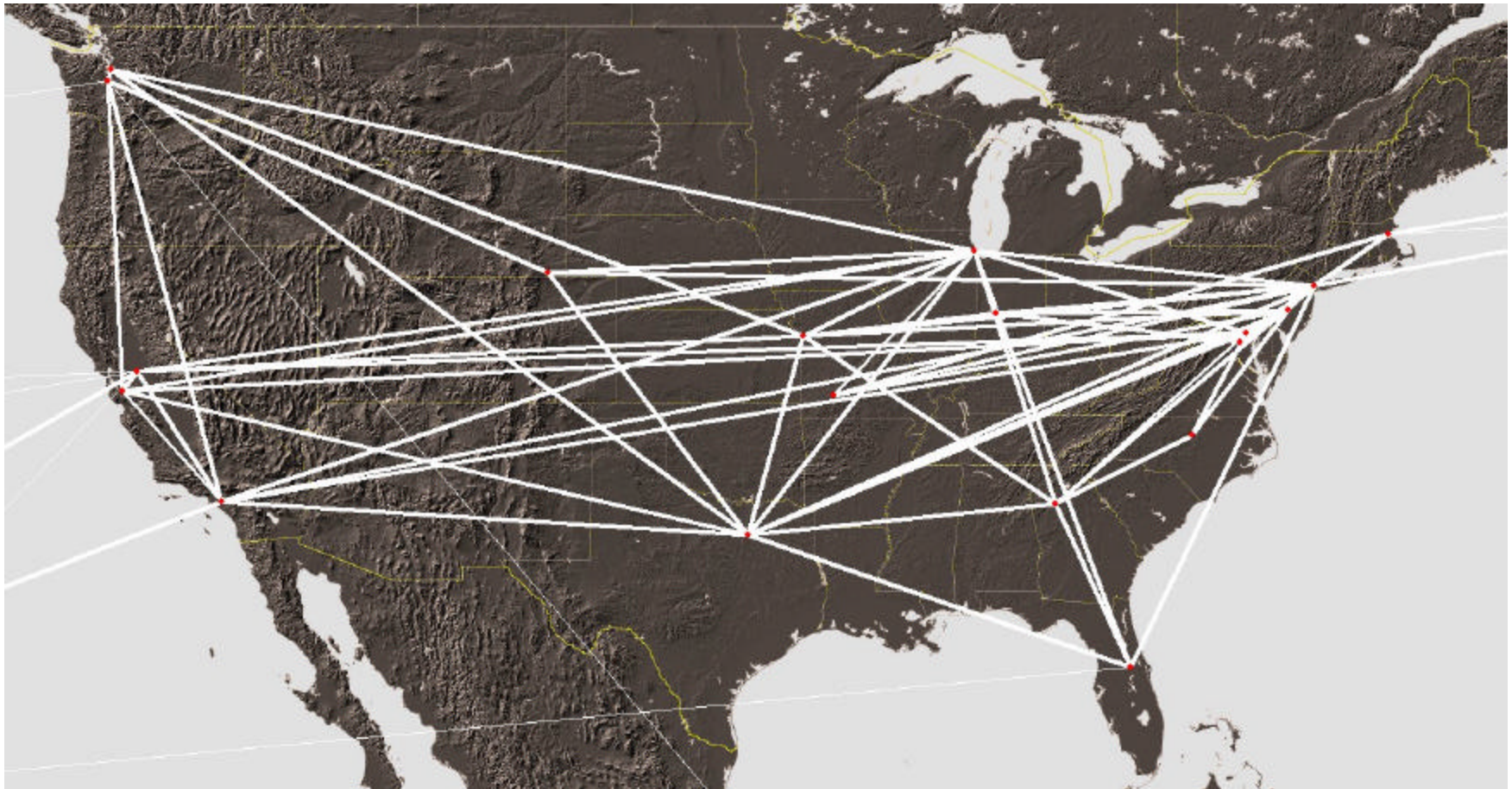


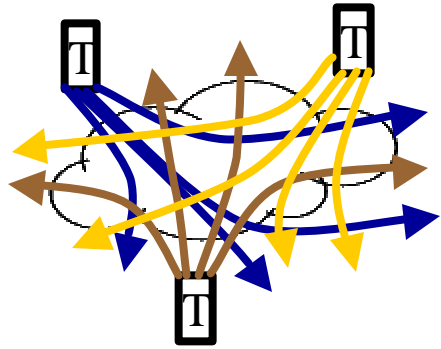
AT&T (AS-7018)



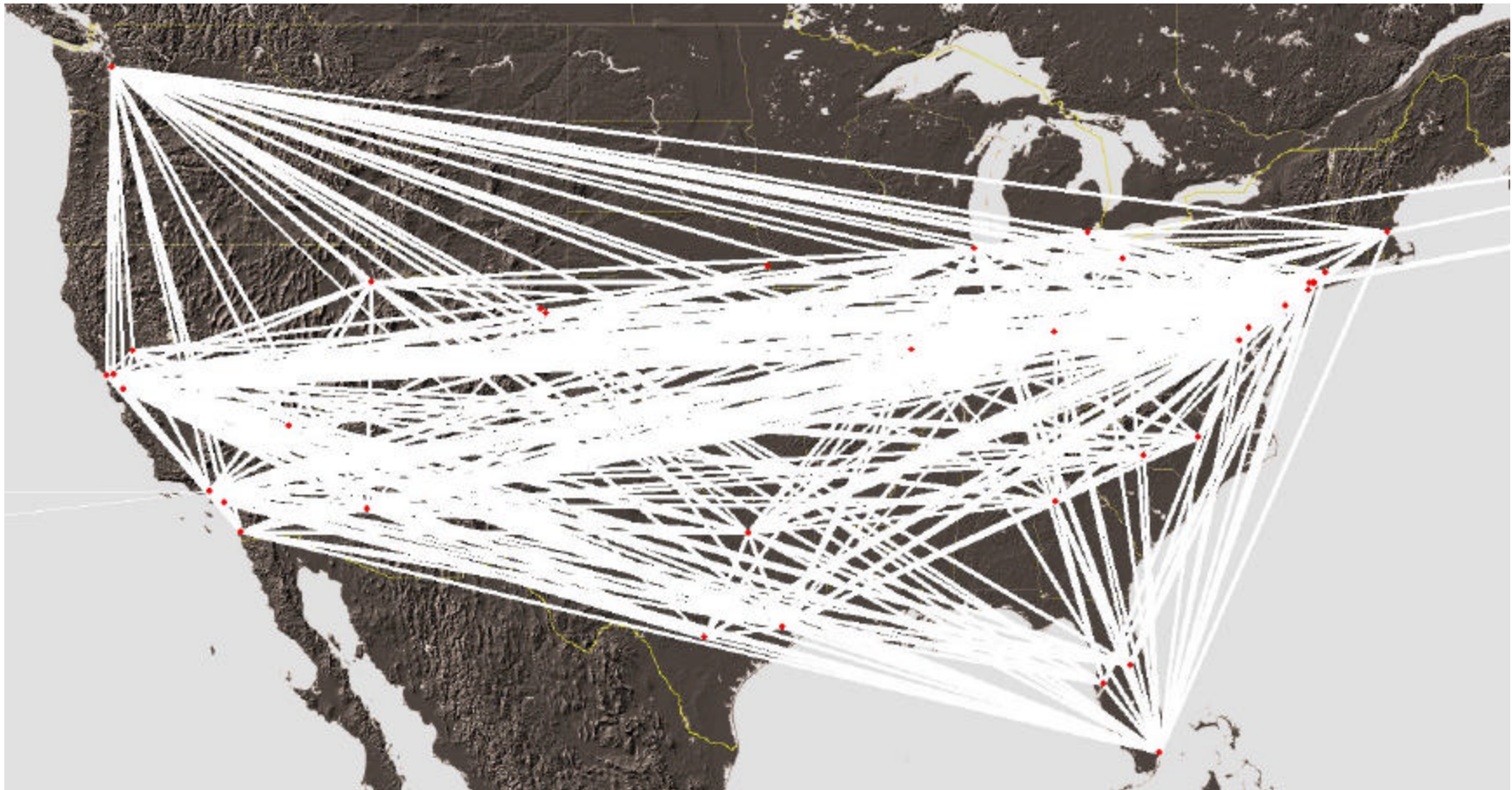


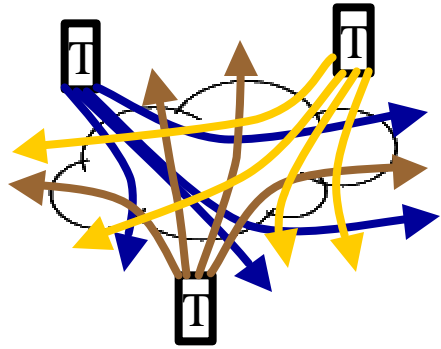
Sprint (AS-1239)



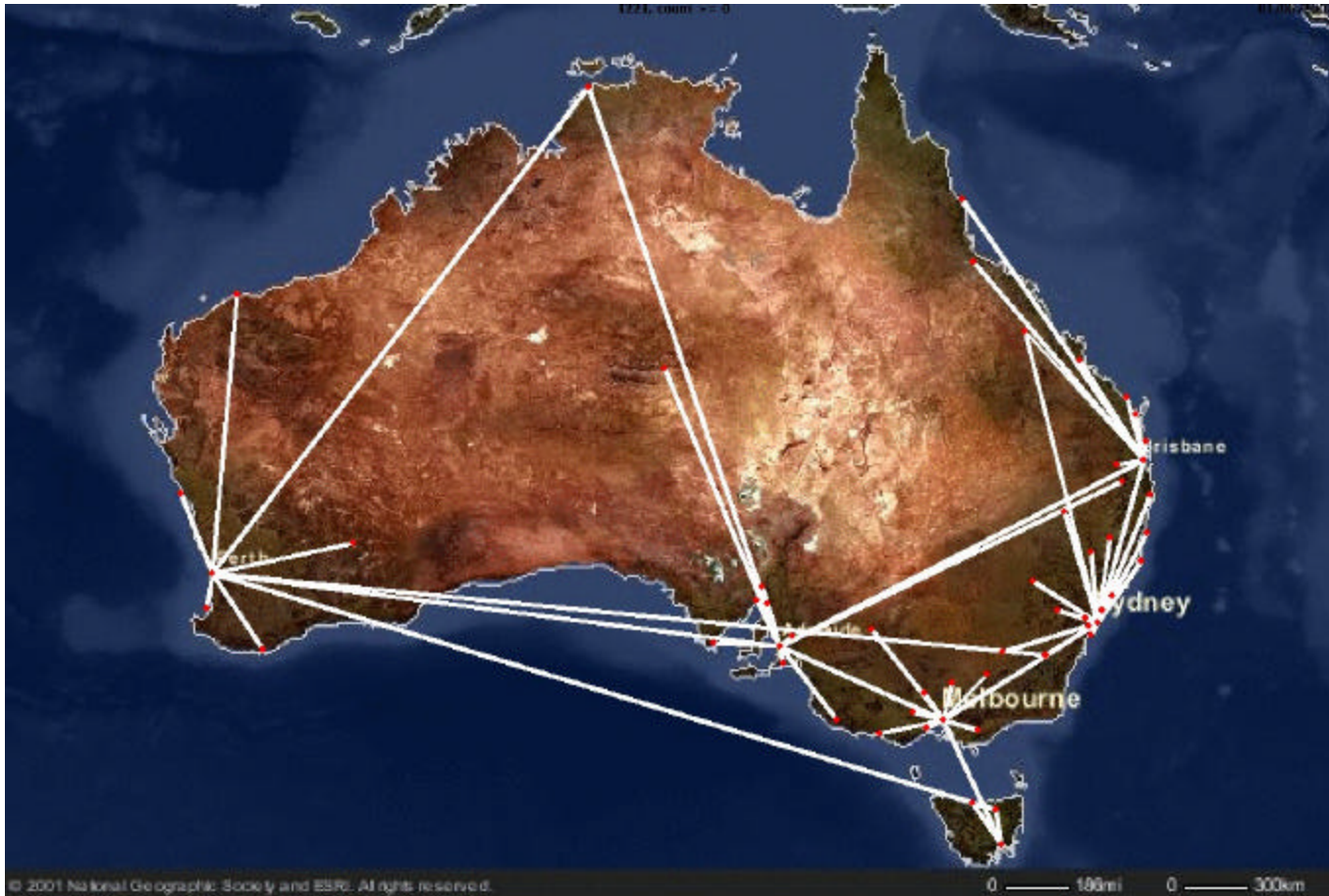


Level3 (AS-3356)

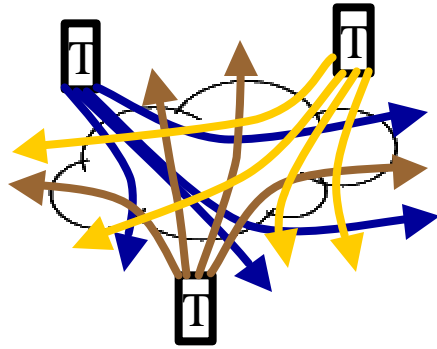




Telstra (AS-1221)

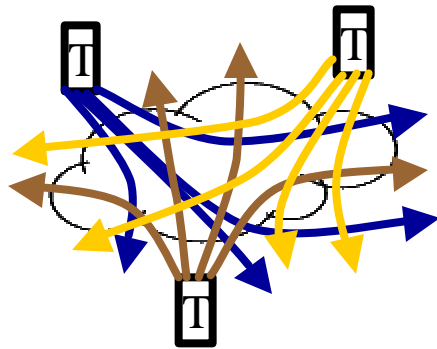


Background image courtesy National Geographic

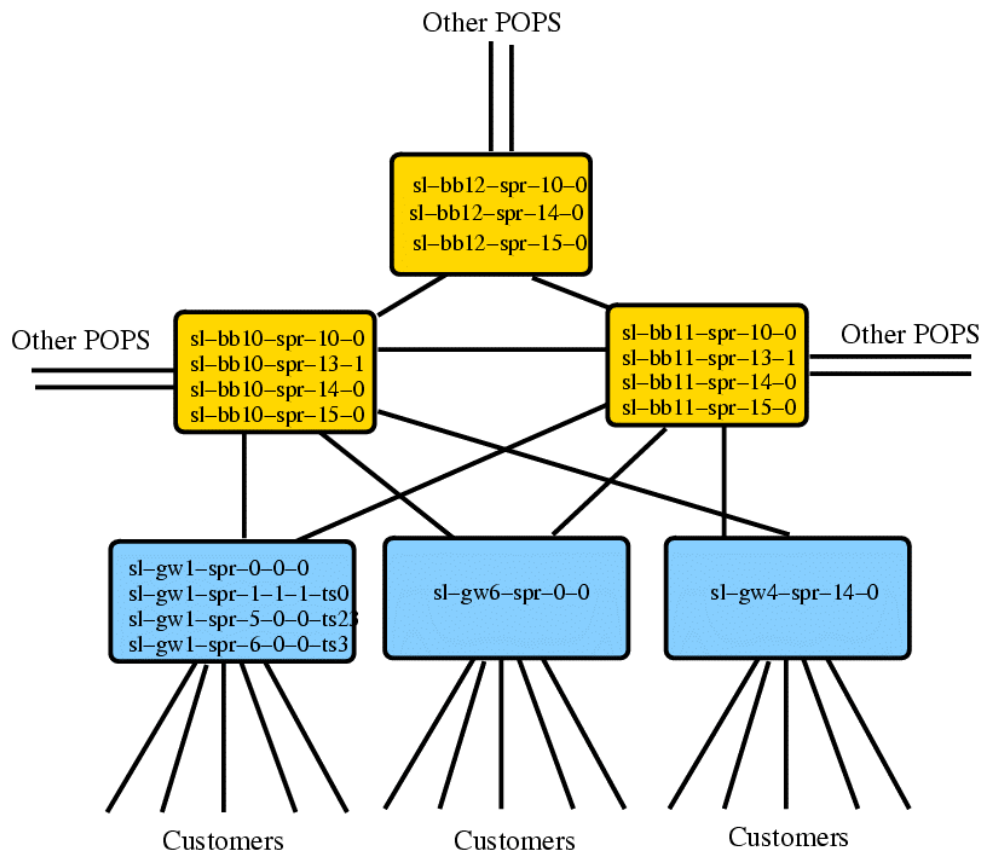


Backbone structure

- ◆ Different across ISPs
 - AT&T and Telstra - hubs and spokes
 - Sprint - fewer, bigger PoPs
 - Level3 - circuits

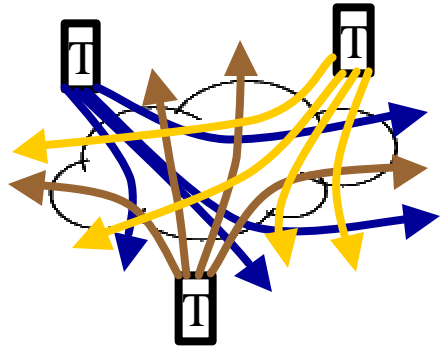


PoP structure



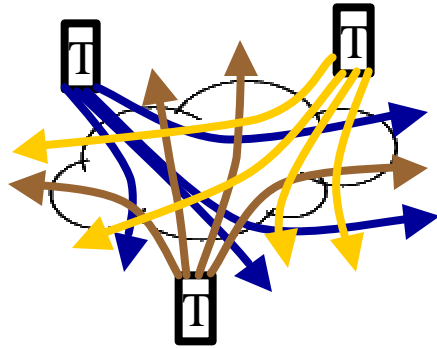
Sprint PoP in Springfield, MA

- ◆ Backbone routers within a PoP are densely connected
- ◆ Gateway routers connect to two backbone routers
- ◆ Customers connect to gateway routers
- ◆ High variability in PoP sizes
 - more than two orders of magnitude
 - [not a power law]



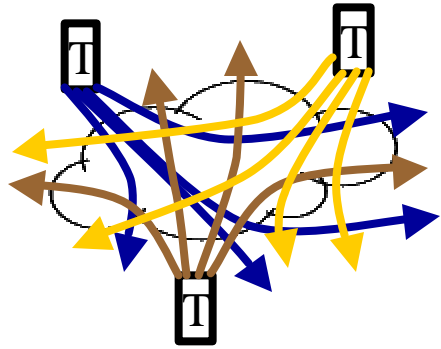
Ongoing and future work

- ◆ Releasing maps and data to the community
- ◆ Analysis
 - peering structure
- ◆ Alias resolution
- ◆ Map annotations
 - latency, bandwidth, link weights
- ◆ Synthetic generation



Conclusions

- ◆ 1000s of public traceroute servers can be used for high-fidelity map collection
 - measurement reduction is the main challenge
 - routing tables and properties of IP routing can be exploited
 - DNS names are very useful for studying various aspects of collected maps
- ◆ ISPs have different backbone structure
- ◆ PoPs have similar designs; two orders of magnitude difference in PoP sizes within an ISP
- ◆ More information: `{ratul,nspring,djw}@cs.washington.edu`



Acknowledgements

- ◆ Mapped ISPs, especially those who validated our maps
- ◆ Public traceroute servers, RouteViews
- ◆ Steve Bellovin, Randy Bush, Christophe Diot and Ramesh Govindan
- ◆ CAIDA - visualization tools and Skitter data
- ◆ Venkat and Lakshmi - GeoTrack city codes