

Cheleby:

Subnet-level Internet Mapper



ISMA 2010 AIMS-2

Workshop on Active Internet Measurements

Talha Oz, Hakan Kardes, Mehmet Gunes
University of Nevada, Reno

Build an efficient system that produces a map of the Internet such that

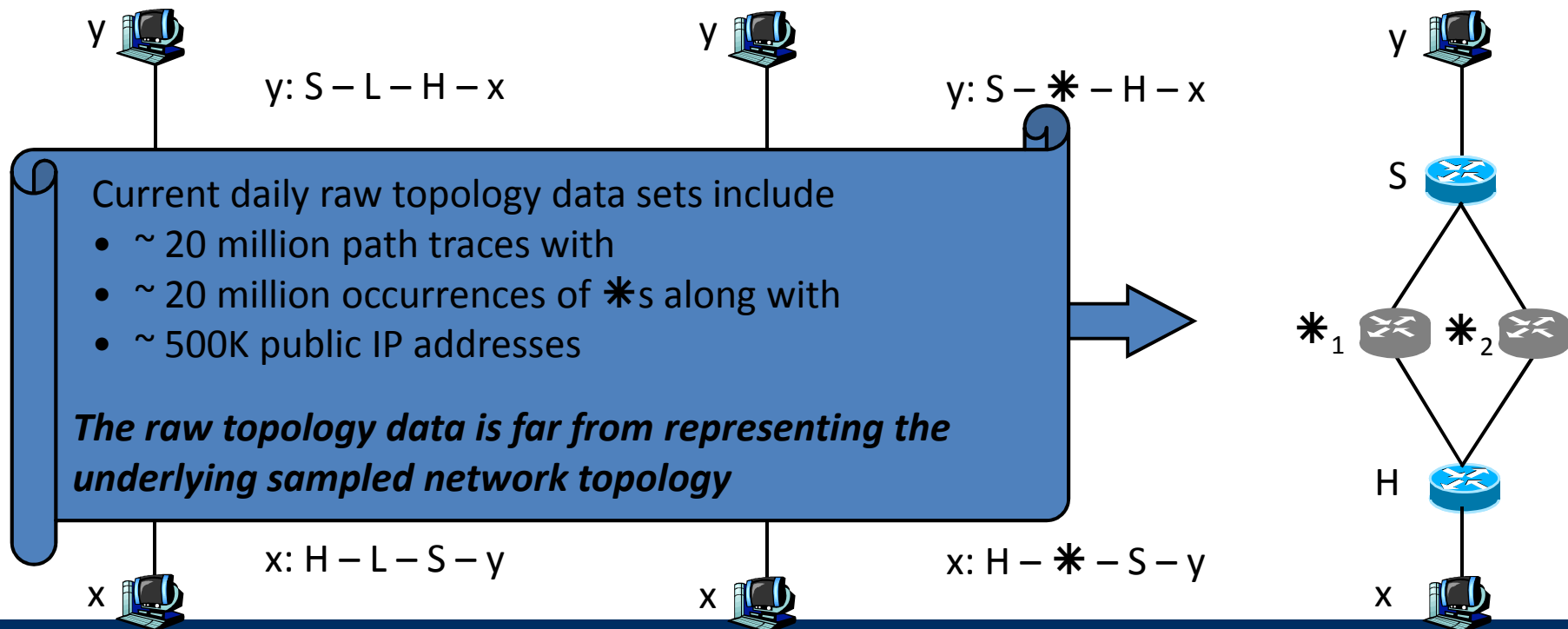
- Alias IP addresses that belong to the same router,
- Star (*) occurrences that stand for the same router,
- IPs that belong to the same subnet are identified.

- Goal
 - Subnet-level Internet Mapping
- Issues
 - Anonymous Routers Resolution
 - Structural Graph Indexing
 - Subnet Inference
 - Distance Preservation
 - Alias IP Addresses Resolution
 - Ally, Analytical & Probe based (APAR)
- Cheleby
 - Mapping System
 - Outer Space 3D Visualization

N

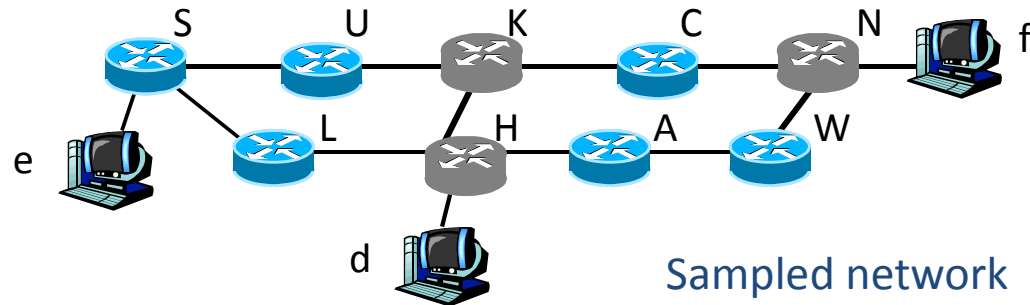
Anonymous Routers

- Anonymous routers do not respond to traceroute probes and appear as * in traceroute output
 - Same router may appear as * in multiple traces.



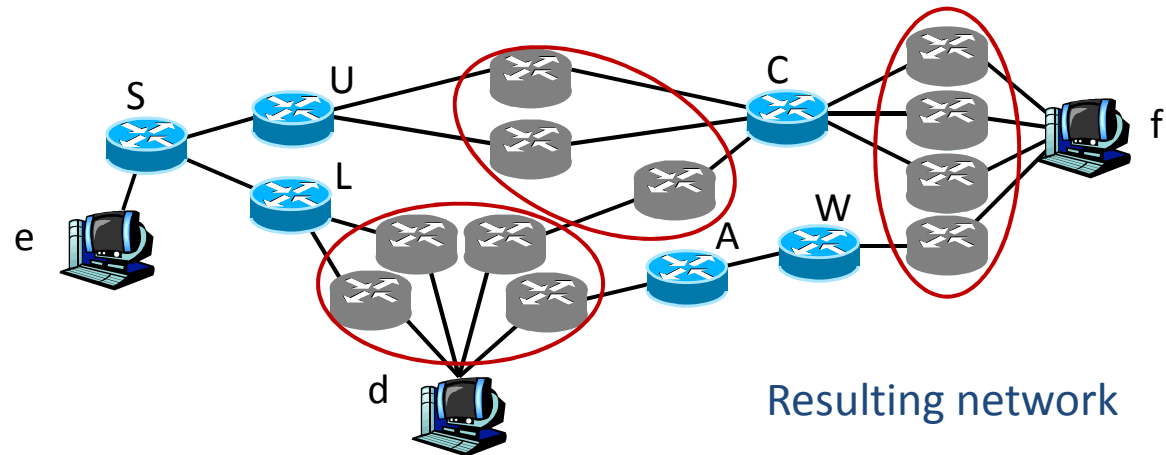
N

Anonymous Router Resolution



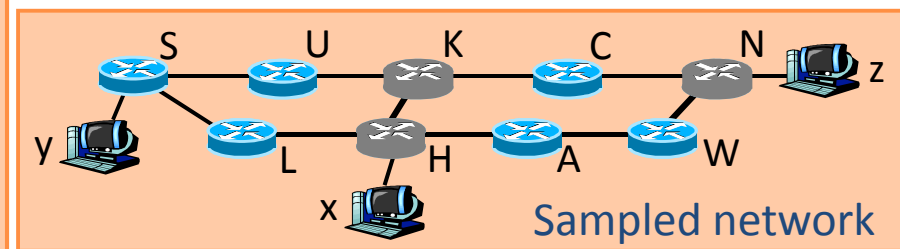
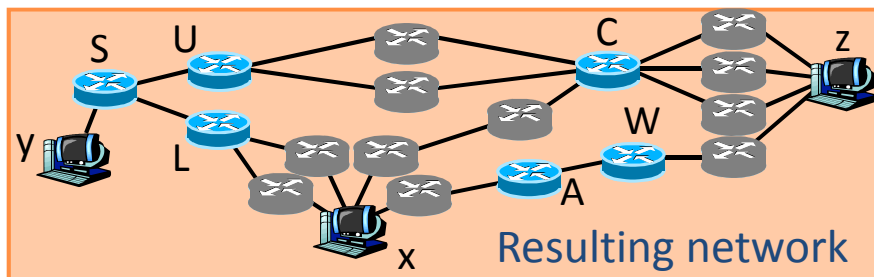
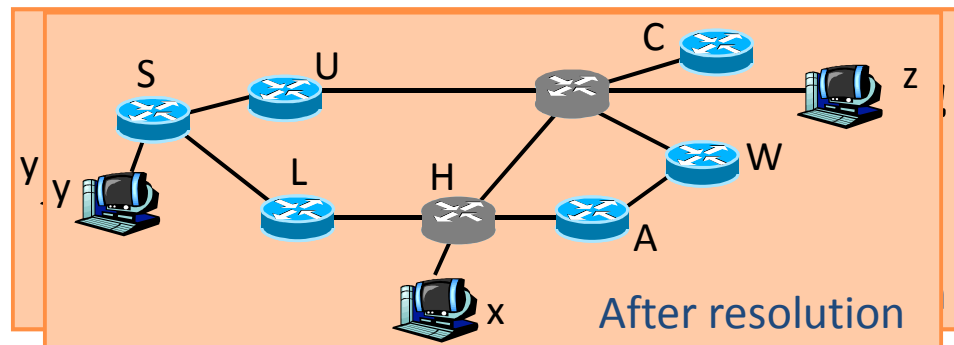
Traces

- d - * - L - S - e
- d - * - A - W - * - f
- e - S - L - * - d
- e - S - U - * - C - * - f
- f - * - C - * - * - d
- f - * - C - * - U - S - e



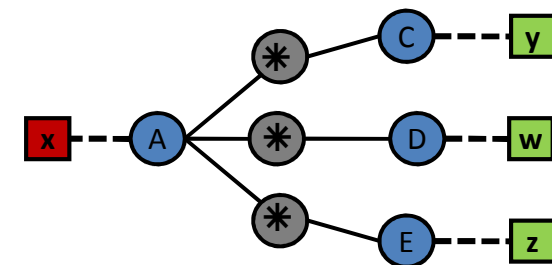
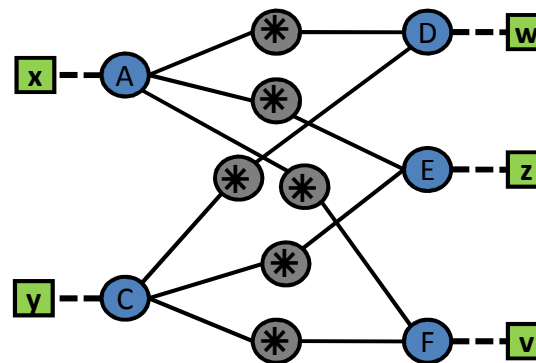
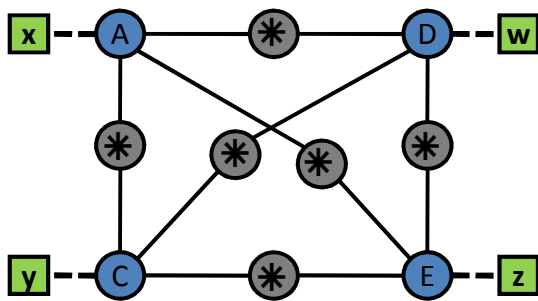
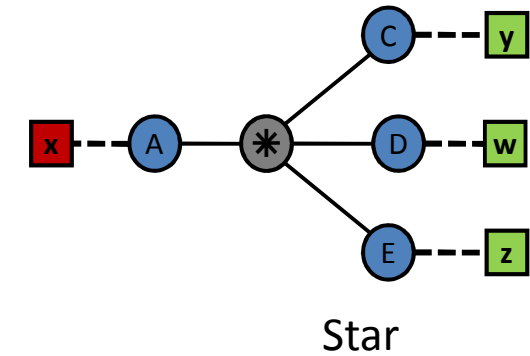
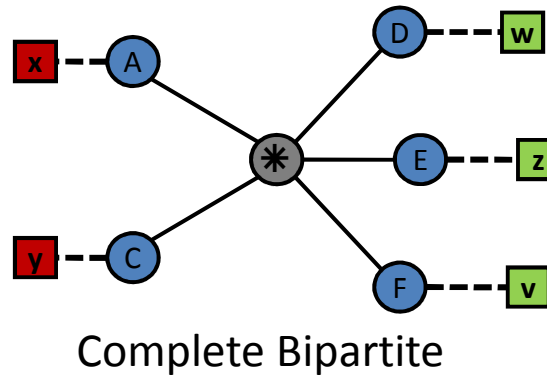
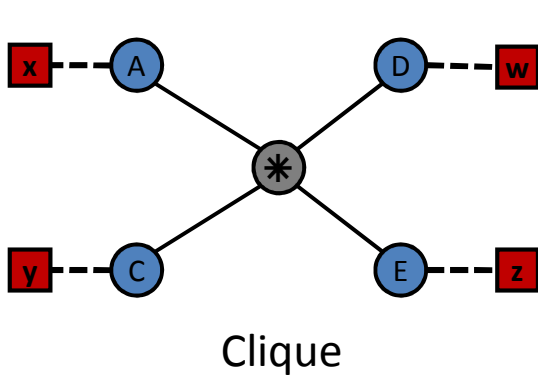
- **Basic heuristics**

- **IP:** Combine anonymous nodes between same known nodes [Bilir 05]
 - Limited resolution
- **NM:** Combine all anonymous neighbors of a known node [Jin 06]
 - High false positives



- More theoretic approaches
 - **Graph minimization** [Yao 03]
 - Combine *s as long as they do not violate two accuracy conditions:
 - (1) Trace preservation condition and (2) distance preservation condition
 - High complexity $O(n^5)$ – n is number of *s
 - **ISOMAP based dimensionality reduction** [Jin 06]
 - Build an $n \times n$ distance matrix then use ISOMAP to reduce it to a $n \times 5$ matrix
Distance: (1) hop count or (2) link delay
 - High complexity $O(n^3)$ – n is number of nodes
 - **Semisupervised Spectral Clustering** [Shavitt 08]
 - Clustering algorithm based on semi-supervised spectral embedding of all the nodes followed by clustering of the anonymous nodes in the projected space.
 - A node will not be chosen to be an unknown root if it shares two or more neighbors with an unknown root.

- Structural Graph Indexing
 - A graph data mining technique
 - Index all pre-defined substructures in a graph data
- Use of SGI for anonymous router resolution
 - Apply SGI to collected path traces
 - Merge anonymous routers using identified structures
 - Trace Preservation Condition
 - Don't merge anonymous routers within the **same trace**
 - Subnet distance as tie-breaker





Graph Indexing based Resolution

Indexing Phase

parallel



star



bipartite



clique

Resolution Phase

parallel



clique



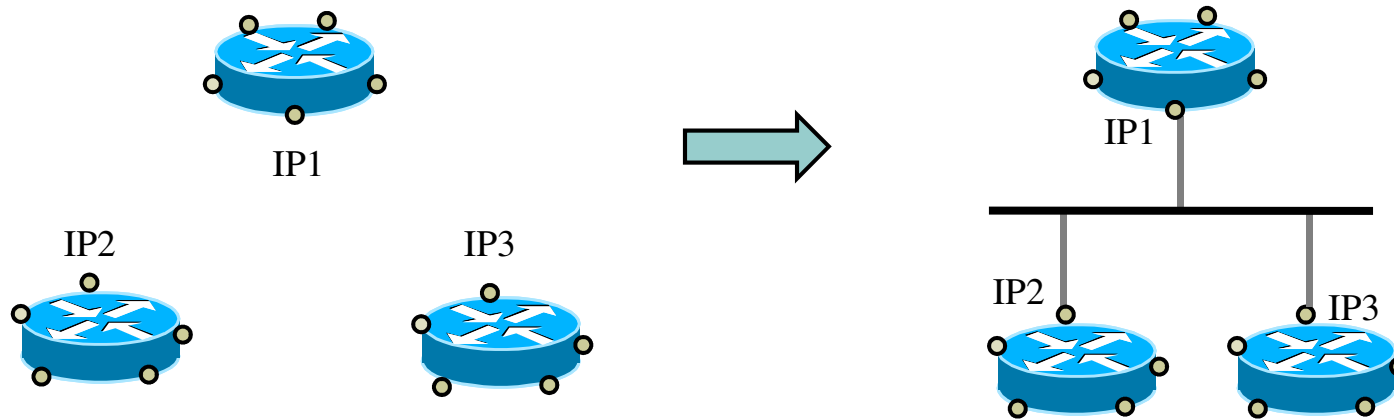
bipartite



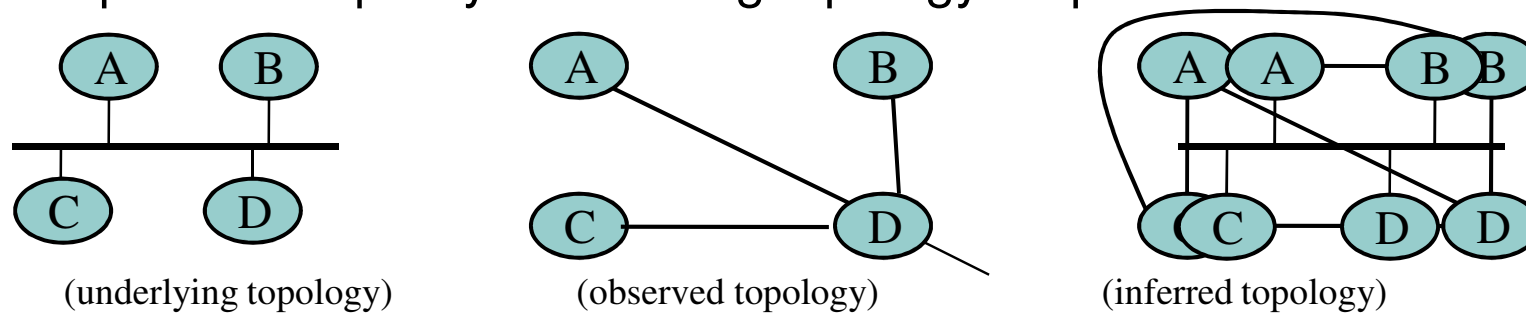
star

- Goal
 - Subnet-level Internet Mapping
- Issues
 - Anonymous Routers Resolution
 - Structural Graph Indexing
 - Subnet Inference
 - Distance Preservation
 - Alias IP Addresses Resolution
 - Ally, Analytical & Probe based (APAR)
- Cheleby
 - Mapping System
 - Outer Space 3D Visualization

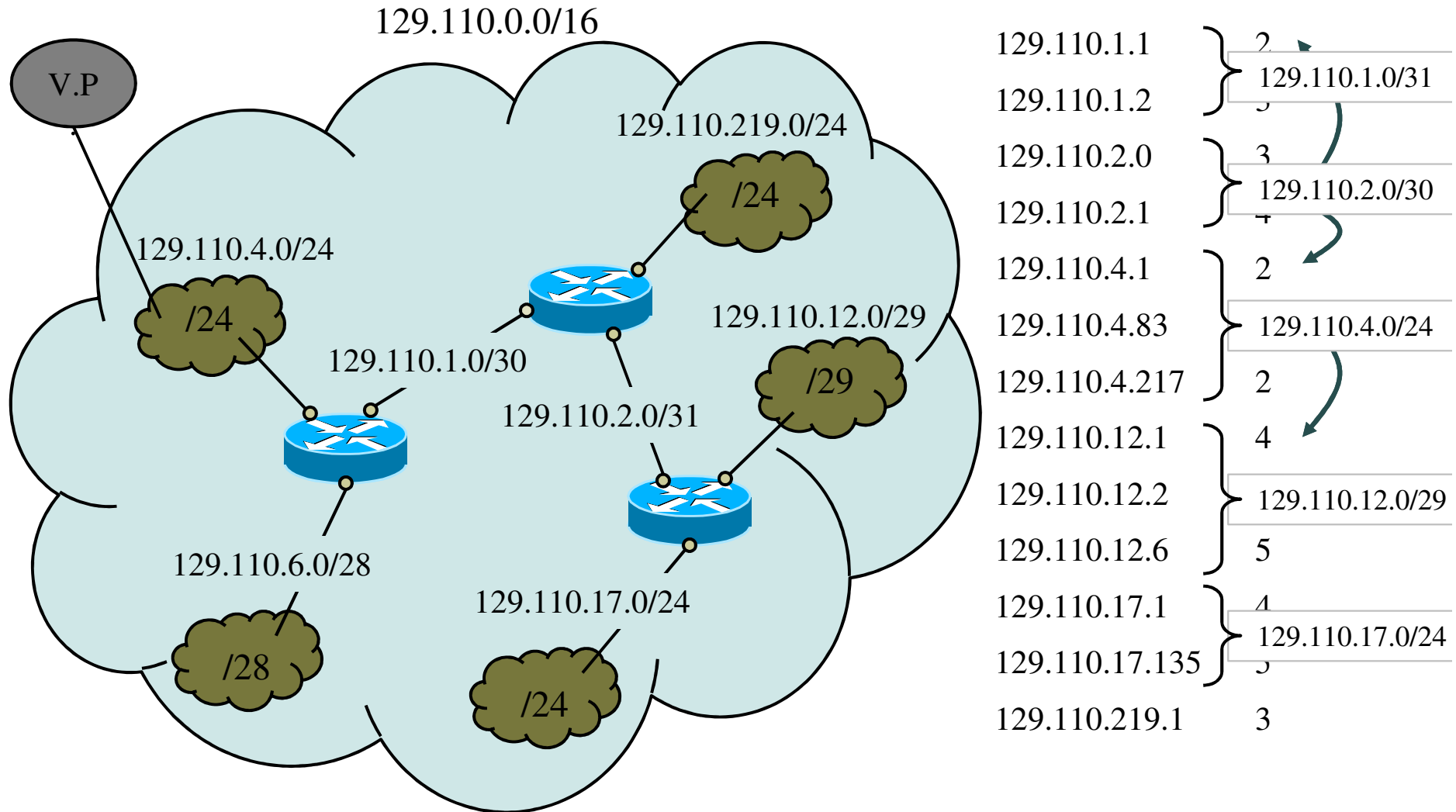
- Subnet resolution
 - Identify IP addresses that are connected over the same medium



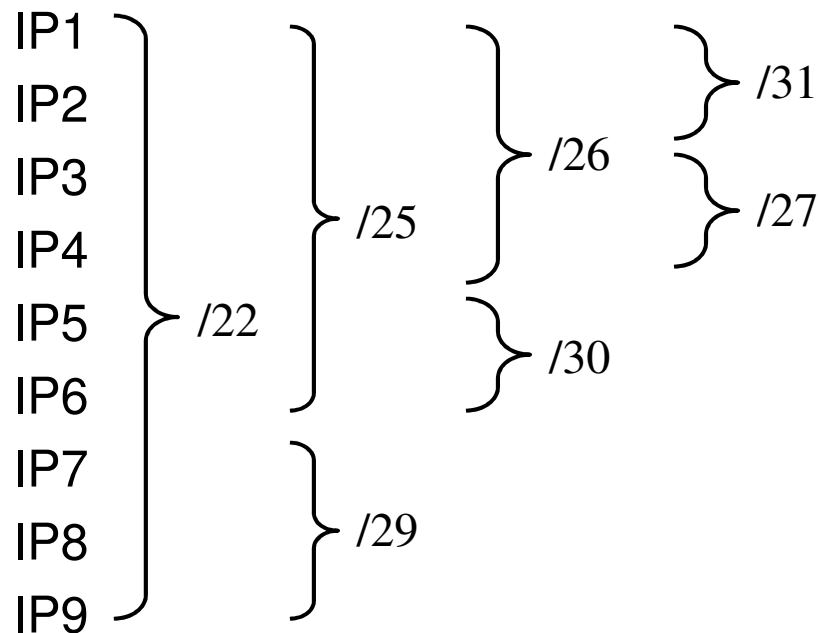
- Improve the quality of resulting topology map



Subnet Inference Approach



- Inferring Subnets
 - Cluster IP addresses into maximal subnets up to a given size (e.g. /22)
 - Distance analysis on candidate subnets to break them down as necessary



A /27 subnet can have up to 2^5 IP addresses.

- *Completeness*: Ignore candidate subnets that have less than one quarter of their IP addresses present
 - after additional probing

N

Inference with Distance Matrix

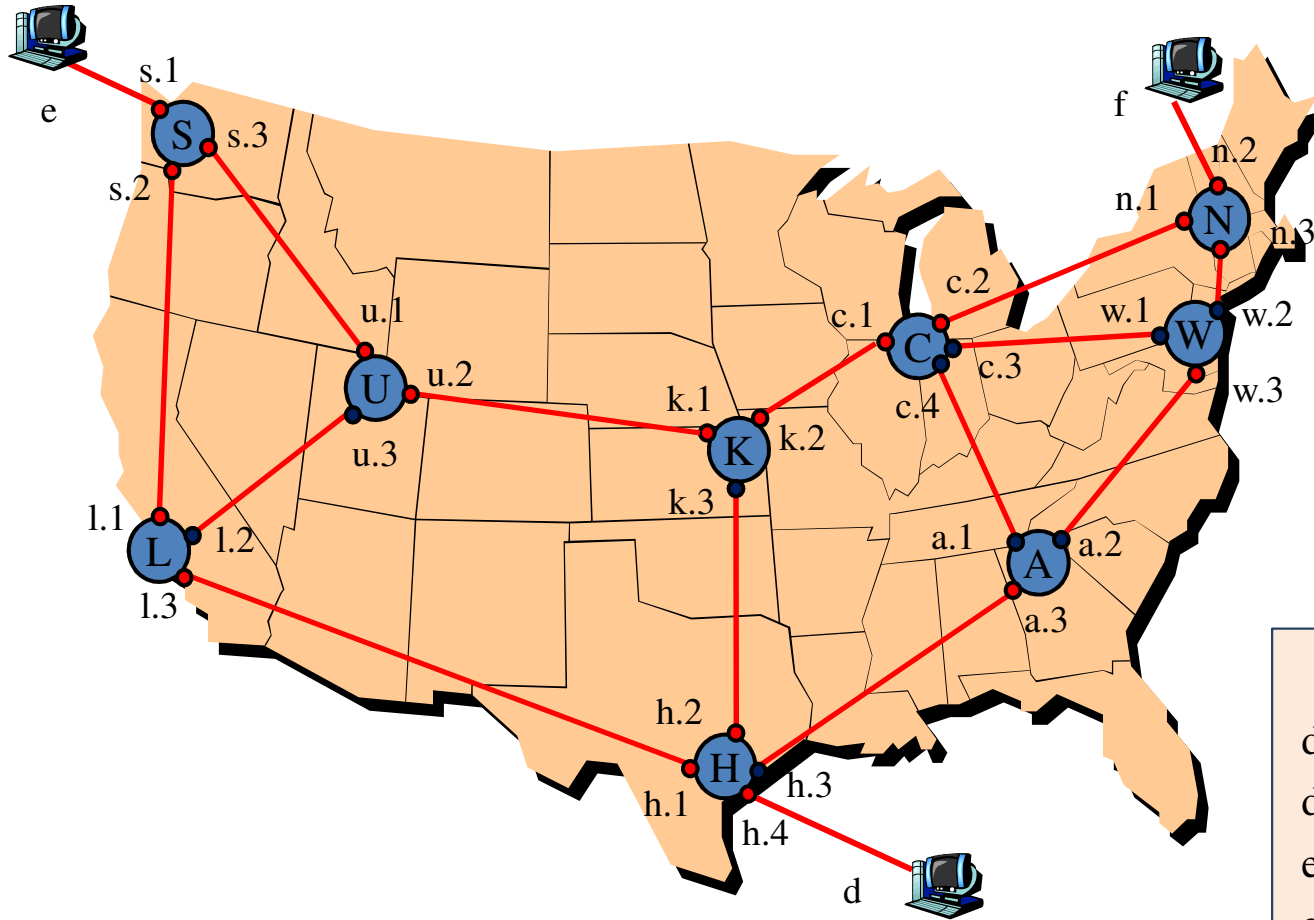
- Obtain distance of each IP from 8 vantage points (VP)
- Only one IP at a subnet might be at a distance 'hop-1' per VP
- IPs after per-destination and per-packet load-balancers
 - Get minimum hop (seen at any ICMP Paris Traceroute) of an IP per VP
 - IP hops after a LB has lower trust
 - Two rounds of computations
 - Compensate for diamond asymmetry if per-destination LB

VP:	1	2	3	4	5	...	672
IP1	0	5	4	0	0	...	7
IP2	0	0	3	5	0	...	7
IP3	2	5	0	4	0	...	6
...							

- Goal
 - Subnet-level Internet Mapping
- Issues
 - Anonymous Routers Resolution
 - Structural Graph Indexing
 - Subnet Inference
 - Distance Preservation
 - Alias IP Addresses Resolution
 - Ally, Analytical & Probe based (APAR)
- Cheleby
 - Mapping System
 - Outer Space 3D Visualization

N

IP Alias Resolution



Traces

d - h.4 - l.3 - s.2 - e

d - h.4 - a.3 - w.3 - n.3 - f

e - s.1 - l.1 - h.1 - d

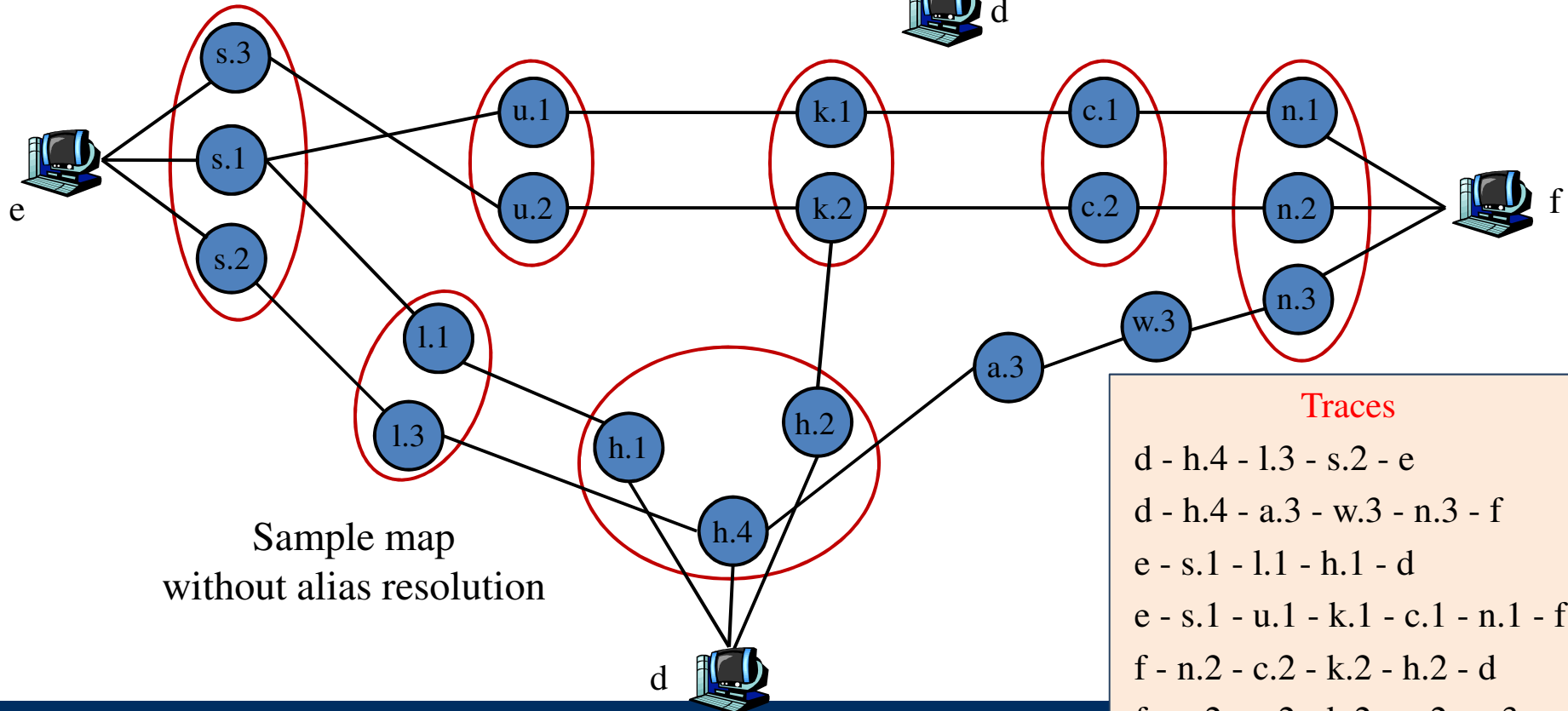
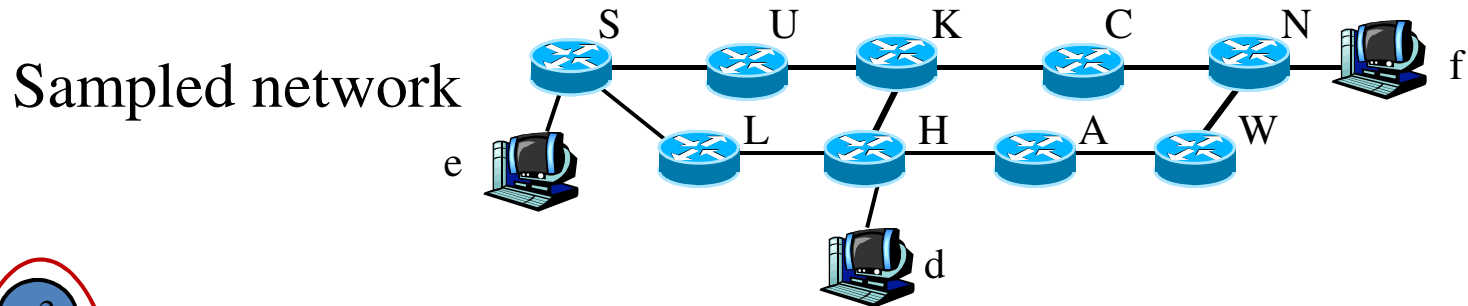
e - s.1 - u.1 - k.1 - c.1 - n.1 - f

f - n.2 - c.2 - k.2 - h.2 - d

f - n.2 - c.2 - k.2 - u.2 - s.3 - e

N

IP Alias Resolution



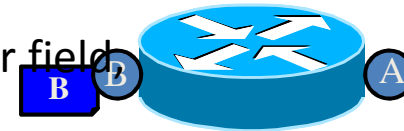
Traces

d - h.4 - l.3 - s.2 - e
 d - h.4 - a.3 - w.3 - n.3 - f
 e - s.1 - l.1 - h.1 - d
 e - s.1 - u.1 - k.1 - c.1 - n.1 - f
 f - n.2 - c.2 - k.2 - h.2 - d
 f - n.2 - c.2 - k.2 - u.2 - s.3 - e

Previous Approaches

- Source IP Address Based Method [Pansiot 98]
 - Relies on a particular implementation of ICMP error generation.

- IP Identification Based Method (ally) [Spring 03]
 - Relies on a particular implementation of IP identifier field.
 - Many routers ignore direct probes.

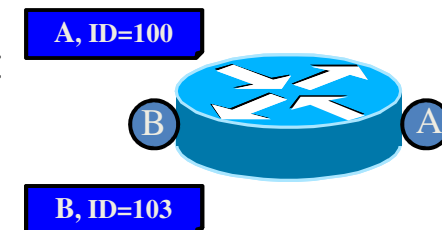


- DNS Based Method [Spring 04]
 - Relies on similarities in the host name structures



sl-bb21-lon-14-0.sprintlink.net
sl-bb21-lon-8-0.sprintlink.net

Works when a systematic naming is used.



- Record Route Based Method [Sherwood 06]
 - Depends on router support to IP route record processing



Analytical Alias Resolution



UTD

no response

129.110.95.1

no response

129.110.5.1

206.223.141.74

206.223.141.73

206.223.141.69

206.223.141.70

198.32.8.33

198.32.8.34

198.32.8.65

198.32.8.66

198.32.8.85

198.32.8.84

192.5.89.10

192.5.89.9

192.5.89.89

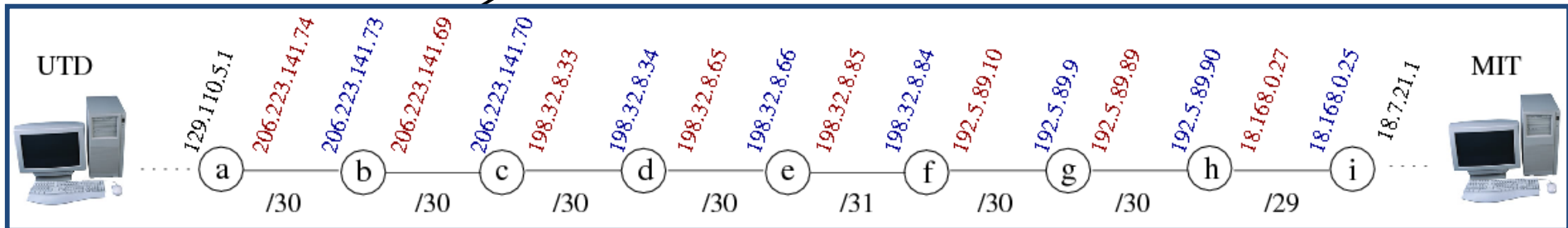
Aliases

129.110.5.1 - 206.223.141.74

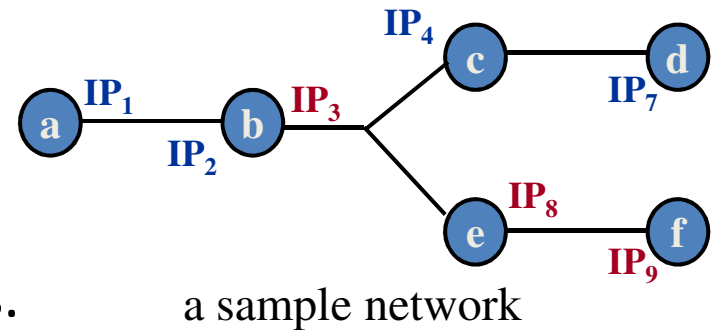
206.223.141.73 - 206.223.141.69

206.223.141.70 - 198.32.8.33

...



- There is possibility of
 - incorrect subnet assumption,
 - Two /30 subnets assumed as a /29,
 - incorrect alignment of path traces.
 - IP_4 and IP_8 are thought of as aliases.

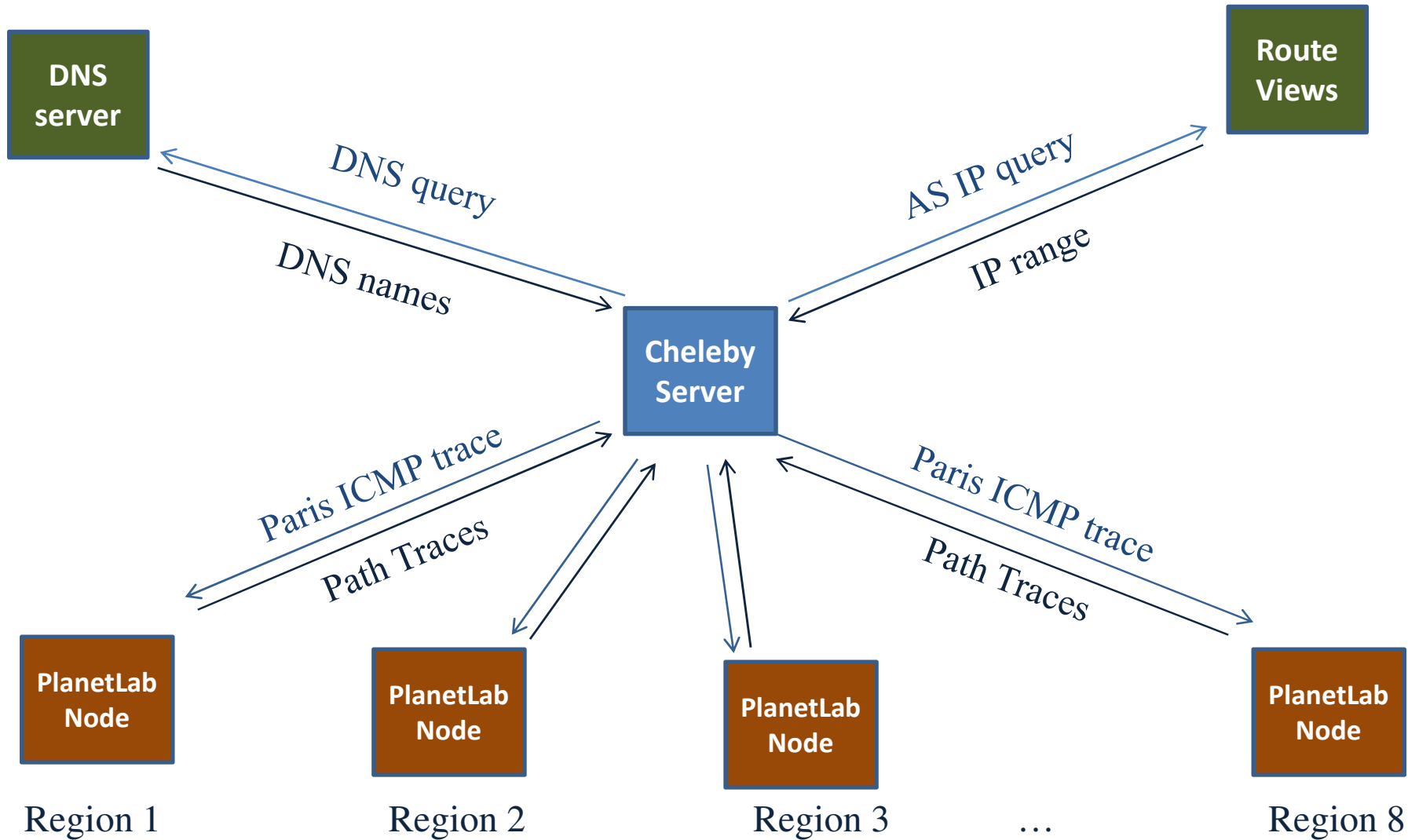


- To prevent false positives, some conditions are defined
 - Trace preservation,
 - Distance preservation (probing component of APAR),
 - Completeness,
 - Common neighbor.

- Goal
 - Subnet-level Internet Mapping
- Issues
 - Anonymous Routers Resolution
 - Structural Graph Indexing
 - Subnet Inference
 - Distance Preservation
 - Alias IP Addresses Resolution
 - Ally, Analytical & Probe based (APAR)
- **Cheleby**
 - Mapping System
 - Outer Space 3D Visualization



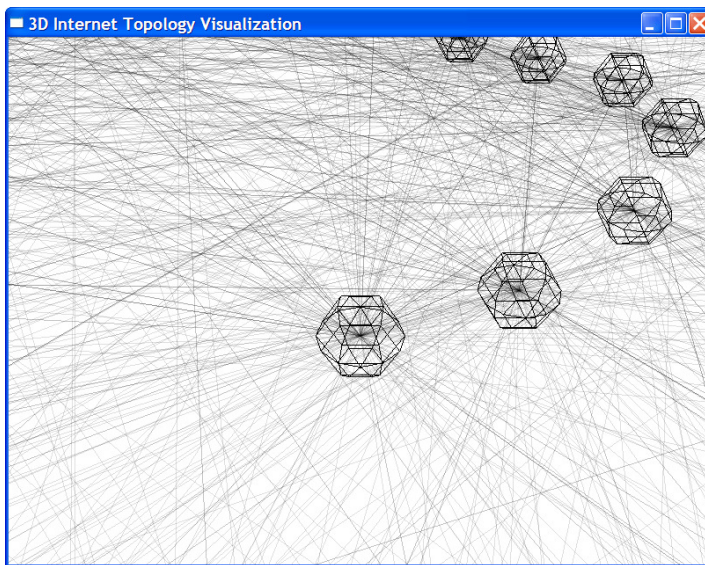
Cheleby Mapping System



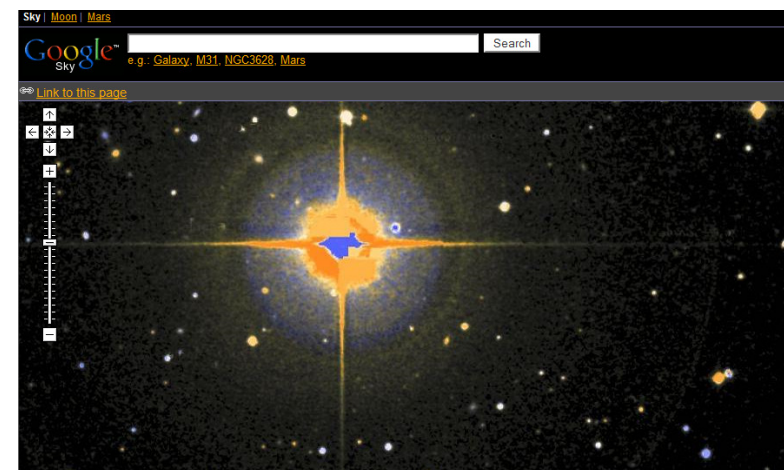
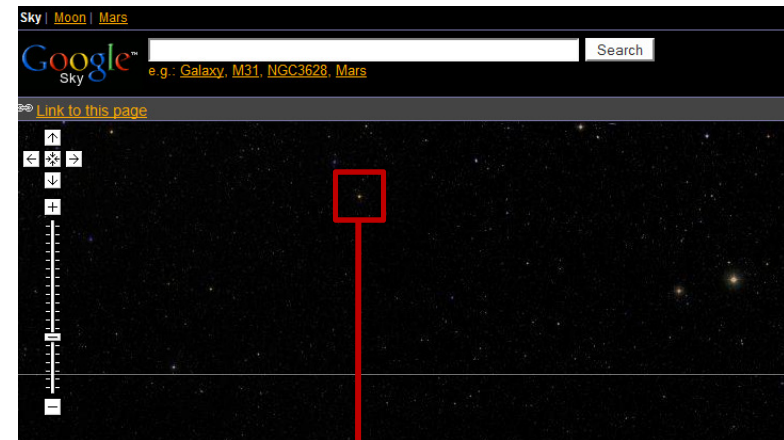
work by David Shelley

– Multiple zoom levels

- Autonomous System-level
- Router-level
- Subnet-level



idea





Questions

