

# Towards an Accurate, Geo-Aware, PoP-Level Perspective of the Internet's Inter-AS Connectivity

Reza Motamedi, Amir Farzad, Hannah Pruse, Reza Rejaie

Department of Computer and Information Science, University of Oregon

## 1- Motivation and Background

- The Internet is a network of networks, it consists of many inter-connected Autonomous Systems (Ases).
- The topology of each AS consists of a collection of connected routers, each with multiple interfaces. A group of tightly-connected routers that directly connect to routers of other Ases are called Point-of-Presence (or PoP) of an AS.
- Capturing the Internet topology is essential to investigate a wide range of problems including:
  - Investigating Internet vulnerabilities and security issues.
  - Designing and re-engineering the Internet protocols at different layers.
  - Building realistic models and formally investigate the Internet topology.

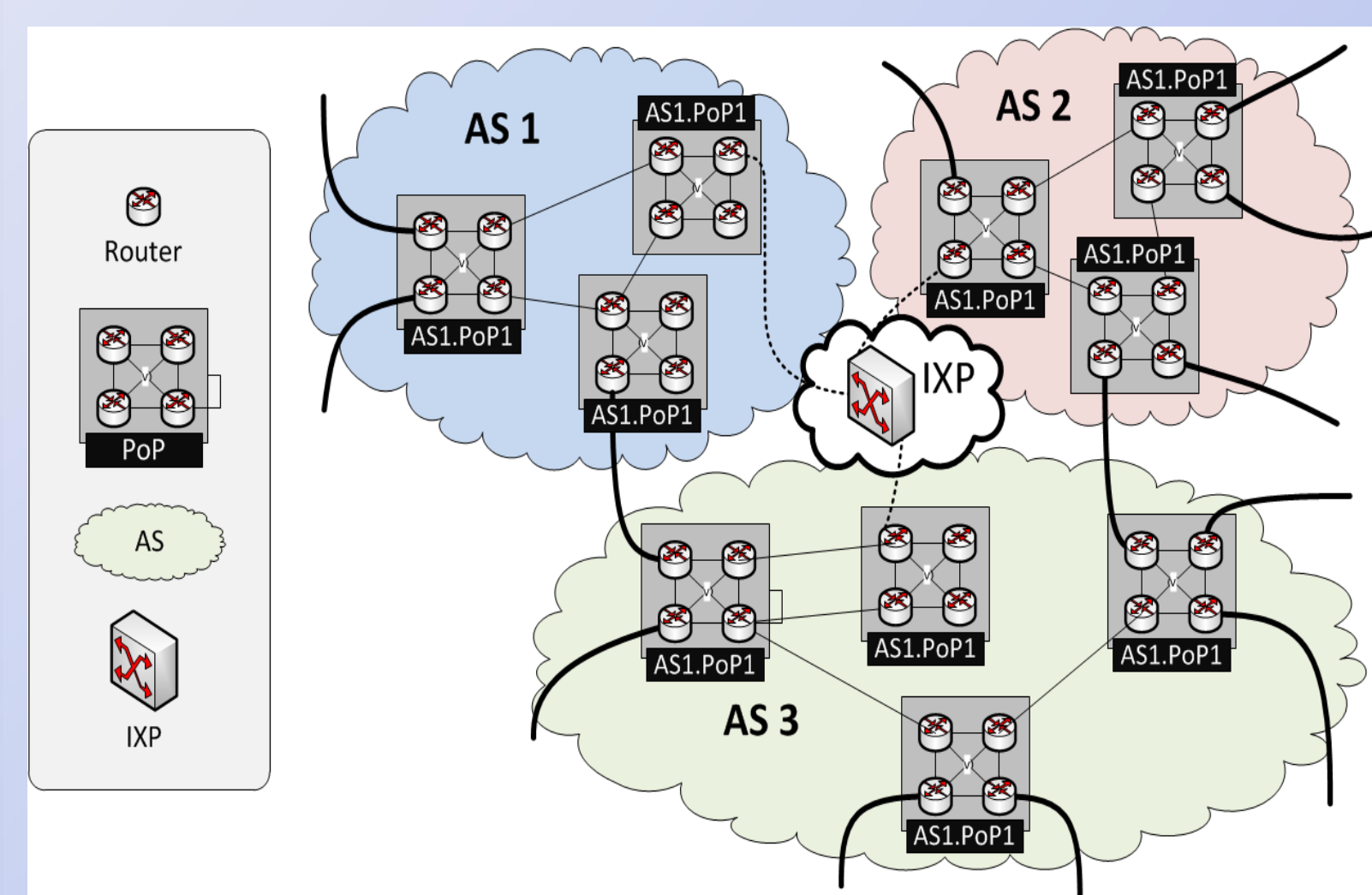


Figure 1. Four levels of the internet topology. Routers are connected through IP interfaces. PoPs are physical locations with a set of routers. They are connected to other PoPs in the same Autonomous System (AS) entity or other ASes. Different ASes may also connect through Internet exchange points (IXPs).

- Internet topology can be viewed at different resolutions, from coarse- to fine-grain: AS level, PoP level, and router level
- Capturing AS-level topology is feasible but it provides a logical view that lacks sufficient details
- A router-level view offers lots of details and thus it is very difficult to accurately capture.
- PoP level topology offers the proper resolution as it is feasible to capture and provides a sufficient level of details to tackle most topology-related problems.
- A useful PoP-level topology should specify the geographical (geo) location of individual PoPs

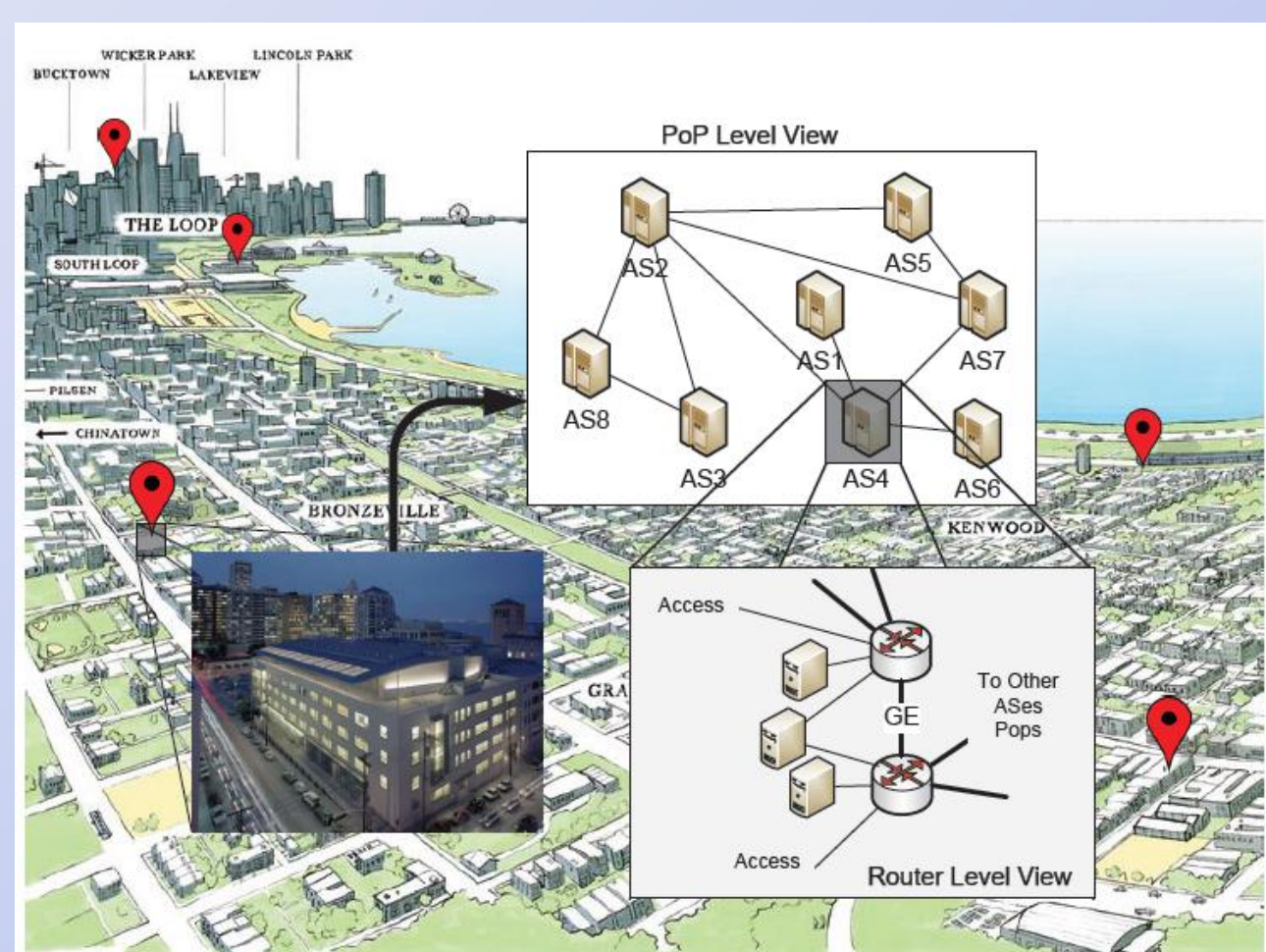


Figure 2. This figure demonstrates different view of connectivity by sketching hypothetical example with multiple colo-facilities in Chicago, inter-AS connectivity (PoP-level view) within a single colo, and intra-AS connectivity (router-level view) for a single PoP

## 2- Problem Definition and Goals

- Given an AS, our goal is to accurately identify (i) the geo location of all its PoPs, and (ii) all of its inter-AS connections through each PoP (considering both direct inter-AS connection and those through Internet eXchange Points, IXPs)
- Most of the network equipment and connections between entities are in PoPs, IXPs and Colocation facilities.

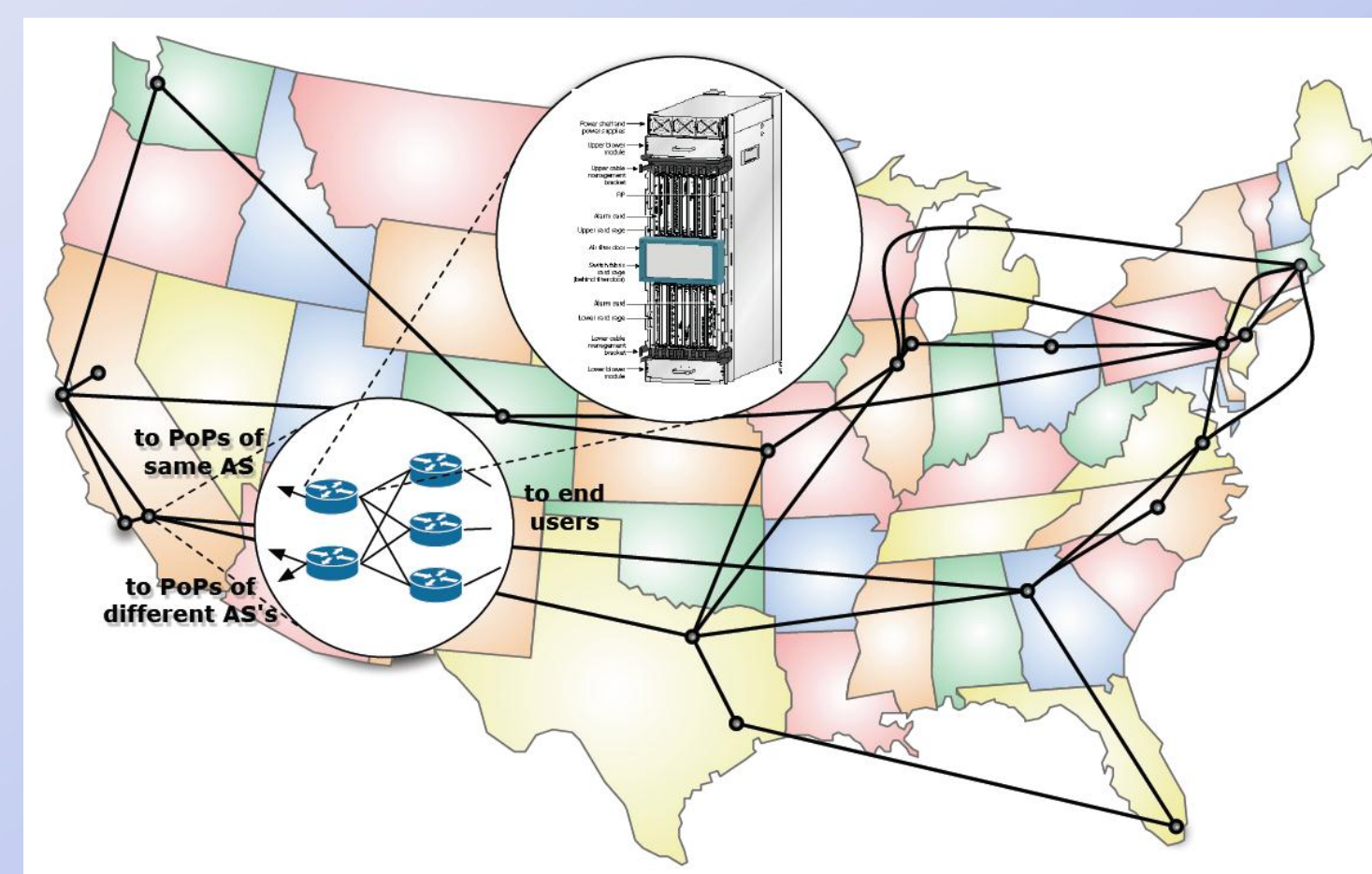


Figure 3. building geo-aware internet map: a hypergraph

- Key Observation: a significant fraction of PoPs of all Ases are located in a relatively small number of colocation facilities (or colo)
- Colo facilities are managed by companies that provide a great deal of information about each facility including its specific address and (in some cases) list of ASes present in a facility
- By "mapping a colo", we can identify ASes that have a PoP in that colo and their inter-AS connections at that colo.
- The key challenge is to conduct a *traceroute* measurement that crosses a target colo. We use the idea of "targeted probing" to achieve this goal.

## 3- Methodology

- Collecting publicly available information about colo facilities from online sources:
  - The international colo and interconnection marketplace is populated by a number of global companies (e.g., Equinix [www.equinix.com](http://www.equinix.com)) and numerous more region-specific businesses (e.g., Interxion [www.interxion.com](http://www.interxion.com) in Europe). They provide information about their services on their web site to attract new ASes.
- Conducting *Targeted traceroute* Probing for High-Yield Discovery of PoPs:
  - We select the source and destination of our probe to be close to the target colo facility with respect to geo distance and AS-level network distance.
  - A traceroute measurement between such a pair of nodes is most likely to cross the target colo facility and reveal the presence of PoPs and inter-AS connections for some ASes at that colo

## Methodology (cont.)

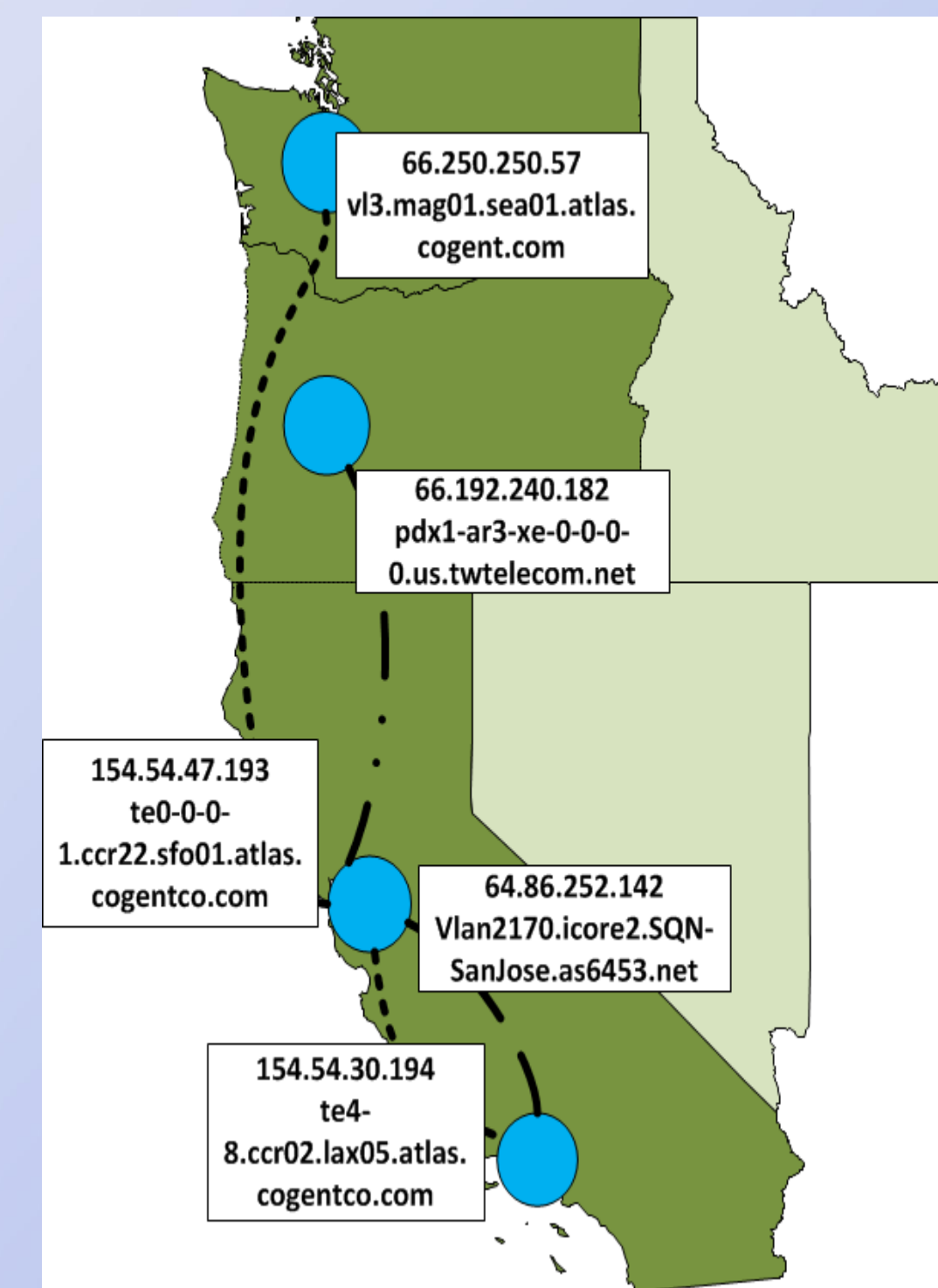


Figure 4. Targeted traceroute from one interface in *cogent* AS to *ttelecom* AS. The trace goes through PoPs in physical facilities. The trace reveals the map for this part of the network

- Dealing with private colocation facilities:** Some large companies place their PoPs in their own private colo instead of publicly known colo facility. The location of these private colos are not publicly known.
- These large companies are often ISPs that provide Internet connection to end users. To find the physical location of their colos, we infer the density of their user population across the globe using **Kernel Density Estimation (KDE)** method.
- The pronounced peaks in their user density indicate the likely location of their private colo facility

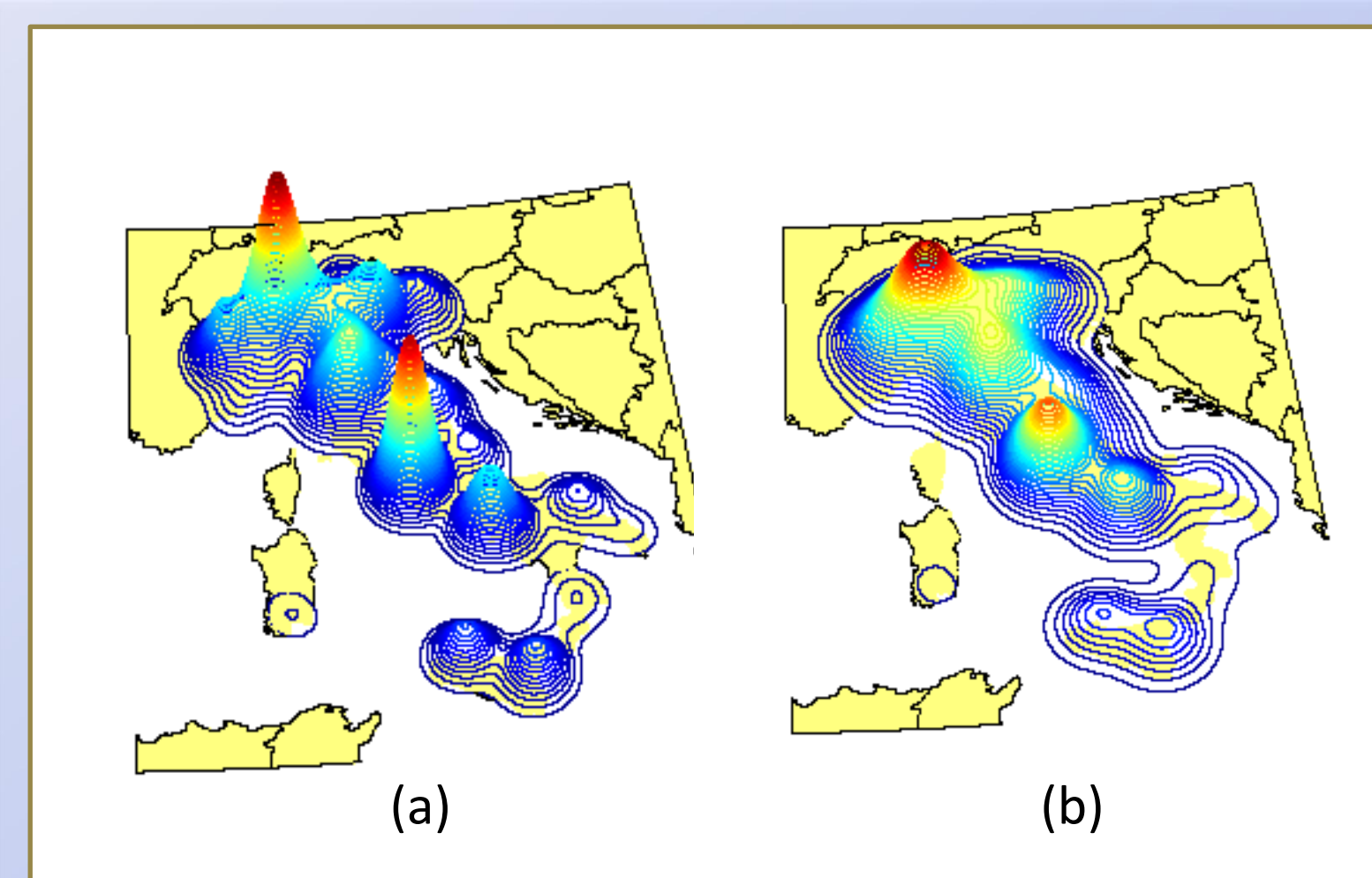


Figure 5. 3D visualization of user density from KDE method for AS 3269.  
a): city level KDE, b): state level KDE

- Dealing with Content Providers:** It is difficult to identify valid destination for traceroute measurement if an AS is associated with a large content provider (e.g. **Google, Amazon, Yahoo**).
- We use the idea of Application Level Probing to identify valid destination IP address within these content providers.
- The idea is to request a service from these content providers and record the IP address of the server that serves the request.
- We then use the discovered IP address as destination for our targeted traceroute probes.

## Methodology (cont.)

- Using traceroute to identify PoPs and Inter-AS Connections at each colo
  - traceroute provides traverse interfaces by a probe. We map each interface to its corresponding AS and create the AS-level path for each probe.
  - The AS level path reveals the presence of inter-AS connection between two relevant ASes.
  - Alias resolution techniques can be used to group the interfaces associated with a single router.
- Mapping Discovered PoPs to Colo Facilities or locations
  - The DNS name of each traversed interface could reveal the geo location (and/or colo) of the interface, its corresponding router and PoP.
  - Putting this info for major colo facilities lead to a geo-aware, PoP-level view of AS\_level topology for the Internet.

## 4- Case studies:

As an example, we have exploited our methodology to map three ASes in US.

- collocation facility mapping:

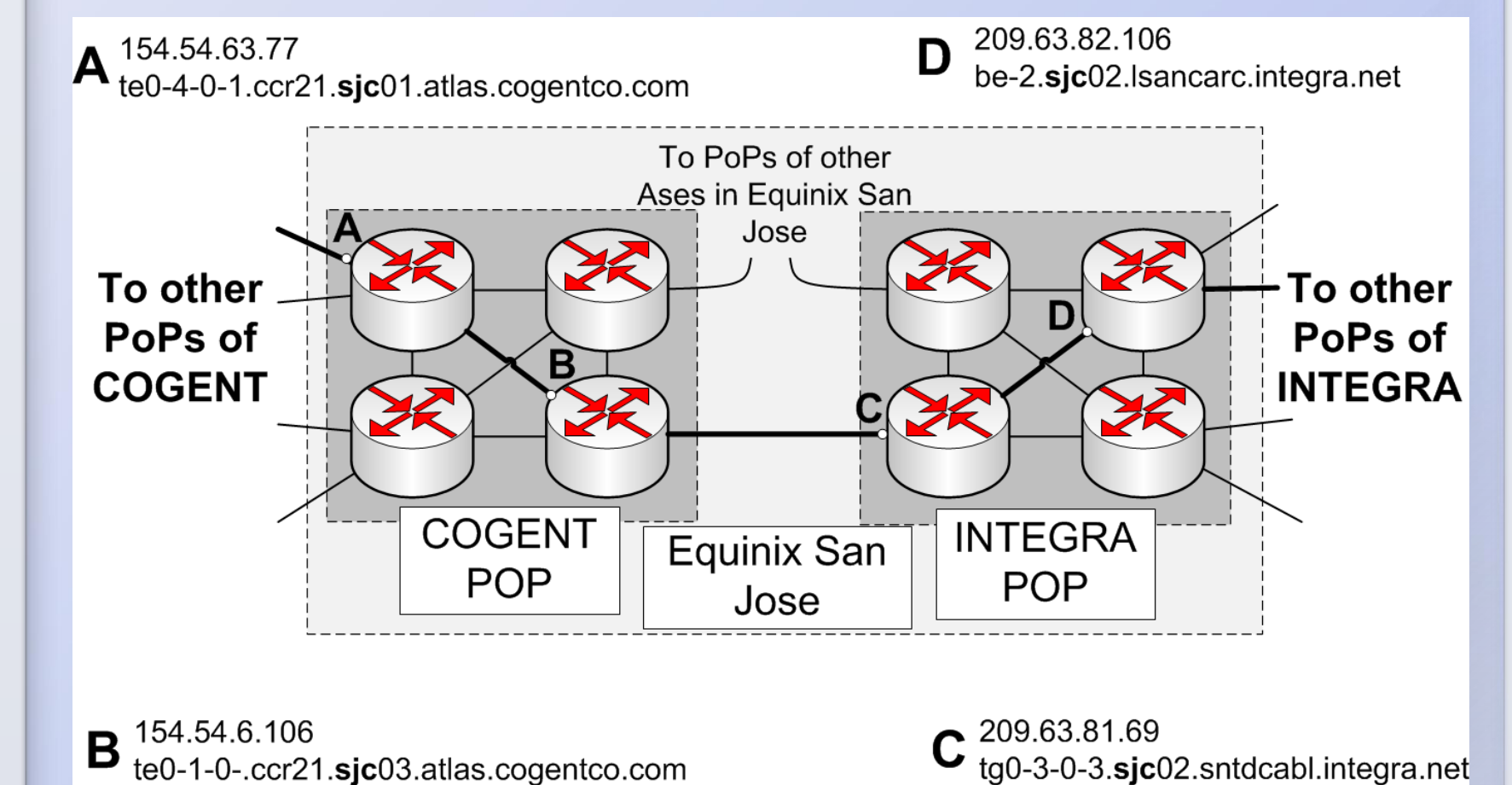


Figure 6. This figure illustrates the PoP of two ASes (Cogent and Integra) along with inter- and intra-PoP connectivity at the Equinix colo facility in San Jose. The darker link marks the path that a targeted traceroute probe may traverse to reveal these PoPs and their inter-PoP connection at this facility. This map has been discovered using our active measurement probing method

Geo-aware map of the ASes and colo facilities using ARC-GIS

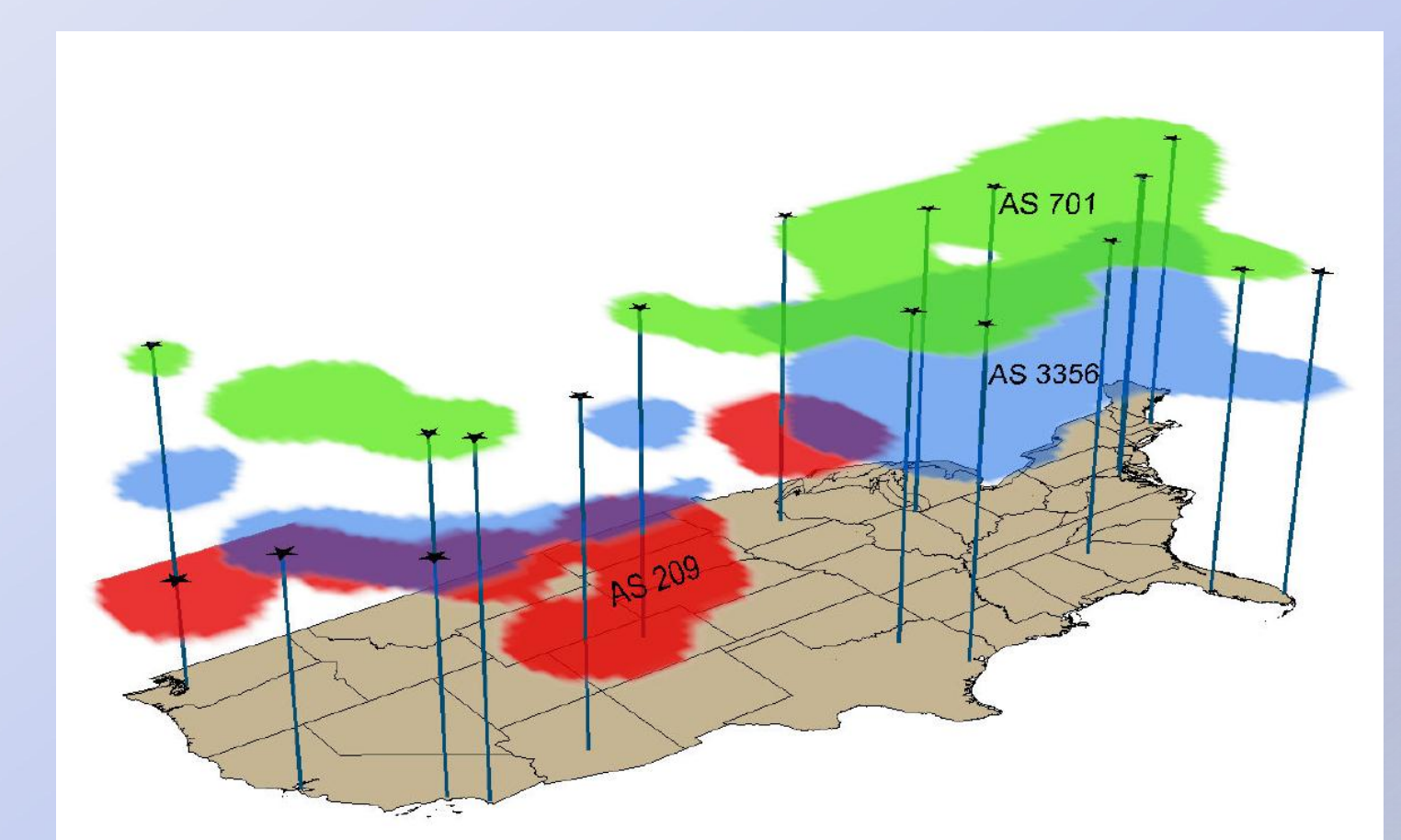


Figure 7. Sample view of geo-footprint for multiple ASes in our portal. The vertical lines indicate the city where PoP for an AS is located.

## 5- Ongoing and Future Work

- Creating a hyper graph of Internet PoP-level topology
- Realistic modeling of AS-level reachability and its application
- Building a GIS-based online portal of Internet maps.