Internet Topology

Seminar of Distributed Computing

Andrea Weisskopf <weandrea+dcs@student.inf.ethz.ch>

January 2004

1 Introduction

With the steadily growing importance and widespread of the internet, many researchers focused their activity into the direction of internet related things. For either proofing some theoretical results or gaining new insights on real world behaviour it is important to have a comprehensive snapshot of the internet. This static snapshot lets the researcher concentrate on their real tasks and prevents them from dealing with a dynamic, changing environment as the internet is one.
So there is a new topic for other scientists to work with: how to capture a precise as possible picture of the internet. Both of the presented papers here cover this problem and chose a different approach.
Both papers were written at CAIDA[1] and make use of the data gathering techniques developed there. CAIDA defines it self as a "collaborative effort among organizations in the commercial, government, and research sectors aimed at promoting greater cooperation in the engineering and maintenance of a robust, scalable global Internet infrastructure". At CAIDA a modified version of `traceroute` called `skitter` was developed. With about 20 monitors around the world running `skitter` a large data collection emerged.

2 Internet topology: connectivity of IP graphs [ipgraph]

2.1 Summary

This large data collection needs post processing, which may be a motivation for this paper. In a period of about a month, more than 220M hosts were probed and 655K of them sent a response back. The graph constructed out of these responses is the basis for all the later analysis in this paper.
A lot of concepts and definitions[2] are introduced. In the largest part of the paper, the authors try to find subgraphs with some given properties or try to group graph properties and calculate the groups distribution.

2.2 Critical evaluation - Inconsistencies of the paper

Starting with a missing research motivation and ending with a promised, non existing follow up paper, it is far from being a good example of how internet research should be presented. The authors begin by explaining how the 655K node graph is constructed. But there is no single word about the reason for that. It sounds like: "Wow, we have a 655K node graph. But what should be do with it. Hmm, lets start with counting graph properties and draw some nice plots out if it..." For a reader this is not reproducible, one reaches the end of the paper still waiting for an explanation for all the stuff seen so far.

3 Traceroute and BGP AS Path Incongruities [bgp]

3.1 Summary

To build an internet graph you can either use the collected IP traceroute paths from `skitter` or use BGP routing tables. The main advantage of BGP data is, that you do not need to query millions of hosts, instead all the needed data can be fetched from the closest BGP router. This paper tries to answer whether this approach is comparable to a long and painful use of `skitter` monitors. The incongruities of the resulting IP traceroute paths and BGP AS paths are analysed and some explanations are given.

Path collection

With the 3 `skitter` monitors `sjc`, `k-peer` and `m-root`[3] about 580K hosts where probed and the corresponding BGP AS paths where extracted from a nearby BGP router.

[1] Cooperative Association for Internet Data Analysis
[2] like cones, trees, stripping, placeholder graph, complementary cumulative distribution function, Weibull distribution, etc.
[3] located in San Jose, Amsterdam and Tokyo

Path simplification

1. About one third off all the hosts were not answering and therefore not usefull for the study. The remaining IP traceroute paths needed to be converted to IP AS paths. This is done by longest prefix matching of the encountered IP addresses with the announced subnet prefixes of the BGP router and the corresponding AS numbers.

2. For redundancy reasons the base data set for the `skitter` probing contained a lot of hosts located in the same AS. So it is necessary to eliminate all the redundant IP AS paths - BGP AS paths pairs.

3. After that - having just unique pairs - we can start identifying incongruent pairs. The percentage of incongruities varies between 20% and 99% depending on the `skitter` host location.

	`sjc`		`k-peer`		`m-root`	
probed hosts	301752		143193		143193	
completed traceroutes	220088	73%	89677	63%	89317	62%
non-redundant IP AS path - BGP AS path pairs	60271	20%	36950	26%	38527	27%
incongruent paths	11279	4%	36888	26%	38460	27%

Path incongruities

Incongruities are categorised in three different parts: incongruities from IX ASes, incongruities from ASes under the same ownership and all the remaining, unexplainable incongruities.

1. **IX ASes:** Most of the internet traffic peering[4] and transit[5] is done at Internet eXchange points, which is a cause for a lot of IX ASes appearing in IP traceroute AS paths. Due to the fact that `k-peer`'s and `m-root`'s location is near AMS-IX and NSPIXP more than 50% of all incongruities are explainable like that.

2. **ASes under same ownership:** In theory one company just need one AS, in fact most of the ISPs own more than one AS. This arose due to business mergers and acquisitions. AS under the same ownership are responsible for just a small fraction of about 5% to 20% of the incongruities. But how can business mergers influence IP AS paths so that they differ from their corresponding BGP AS path? It is a pity that this interesting thing is not explained in the paper. It is stated correctly that BGP analysis cannot be done without taking peering policies into account, but this is no explanation for the AS path incongruities.

3. **Remaining:** Depending on the data collection point, still more than 40% of all incongruities remain unexplainable. The only clear obseravation is, that in general IP AS paths are longer than BGP AS paths. Other metrics like editing distance and length differences are also considered.

3.2 Critical evaluation

When comparing these two internet models it is important to keep in mind for what kind of research you need it. There is no approach which is better in general, you have to decide from case to case. In a perfect world, with all the necessary data available to the collector, the IP traceroute model is superior to a BGP AS path model. But since we do not live in a perfect world, we have to lower our expectations and the IP traceroute model loses much of it is attraction. Using `ICMP` messages for path discovering, we have just a one way view of most of the paths. So we have a model with most of the paths but some missing bidirectional connectivity. There are also some erroneous hosts because of multihomed or aliased network interfaces. An other problem arises from unresponding hosts, which makes it impossible to combine multiple IP traceroute paths at this junction point. Mostly this limitations are too restrictive.
The situation with BGP AS paths faces the same problem. With just the entries from the routing tables you will not come close enough to the reality because of all the routing policies between the different service providers. One solution to get rid of this, is to study and parse the import and export section in the whois database. Most of the entries are regrettably outdated and cannot be considered. The dataset we get is coming close to the internet but has unfortunately to much inter AS connections. Depending on the kind of study one is carrying out this can be a big problem.

References

[ipgraph] A. Broido, kc claffy: *Internet topology: connectivity of IP graphs*; ITCom 2001

[bgp] Y. Hyun, A. Broido, kc claffy: *Traceroute and BGP AS Path Incongruities*; 2003

[4] exchange of internet traffic without charge
[5] billed traffic exchange