An Information-Theoretic Approach to Routing Scalability

Gábor Rétvári, Dávid Szabó, András Gulyás,
Attila Kőrösi, János Tapolcai

Budapest Univ. of Technology and Economics
Dept. of Telecomm. and Media Informatics
{retvari,szabod,gulyas,korosi,tapolcai}@tmit.bme.hu

Hotnets XIII, October 27–28 2014, Los Angeles, CA, USA

MŰEGYETEM 1782

The mysterious compressibility of IP forwarding tables

- Take the IPv4 forwarding table of some Internet router
  - orders next-hops to address prefixes
- Represent each distinct next-hop with a unique label
- Take individual IPv4 addresses and write down the corresponding next-hop labels one by one
  - result is a string of $2^{32} = 4G$ symbols
  - naive representation of our forwarding table
- Compress this string

The mysterious compressibility of IP forwarding tables

Dest. IP Address	Next-hop IP address	Next-hop label
0.0.0.0	blackhole	0
0.0.0.1	blackhole	0
⋮	⋮	⋮
80.92.12.254	149.11.10.9	17
80.92.12.255	149.11.10.9	17
80.92.13.0	213.248.79.185	18
⋮	⋮	⋮
152.66.244.111	195.111.97.83	41
152.66.244.112	195.111.97.83	41
152.66.244.113	195.111.97.83	41
⋮	⋮	⋮
255.255.255.255	blackhole	0

The mysterious compressibility of IP forwarding tables

0	0	...	17	17	18	...	41	41	41	...	0

- For a real router in the `HBONE` (`AS1955`)

```
$ fib2str hbone.fib.dump > hbone.bin
$ ls −hs hbone.bin
4.0G hbone.bin
$ bzip2 hbone.bin
$ ls −hs hbone.bin.bz2
???
```

The mysterious compressibility of IP forwarding tables

0	0	...	17	17	18	...	41	41	41	...	0

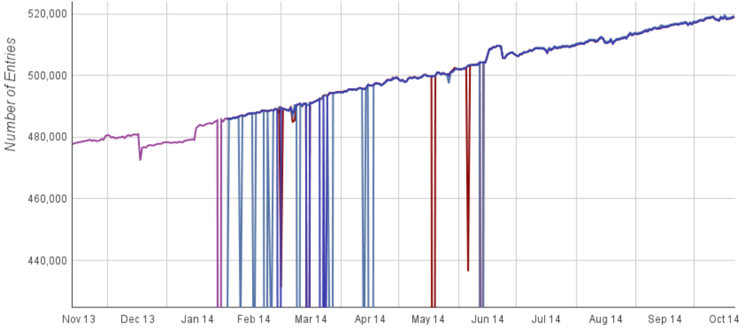
- For a real router in the HBONE (AS1955)

```
$ fib2str hbone.fib.dump > hbone.bin
$ ls -hs hbone.bin
4.0G hbone.bin
$ bzip2 hbone.bin
$ ls -hs hbone.bin.bz2
???
```

- The compressed size is 116 Kbytes
- That's over $37,000$-fold reduction!!

Does hop-by-hop routing scale?

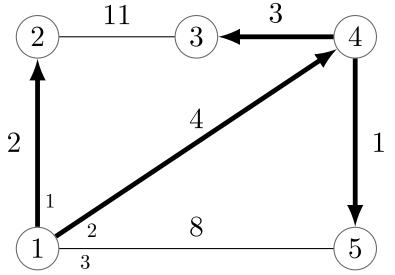
- The key to data plane scalability is forwarding tables
  - involved in every packet lookup
  - routed address space is growing rapidly
- Can we model forwarding tables and reason about size?



Taken from the *Internet Routing Entropy Monitor*, see `http://lendulet.tmit.bme.hu/fib_comp`

The model

- Graph of n nodes
- No address space structure: each node has a random id
- Routing policy (arbitrary) orders each destination node to an outgoing port
- Forwarding table at node v is a string s_v, so that the entry at position u is the next-hop port towards u



$$s_1 = < -, 1, 2, 2, 2 >$$

Modeling power

- How much information **must** be stored at a node to guarantee correct (as of the routing policy) forwarding?
- We can use s_v to answer this question
- **Theorem:** if node ids are assigned randomly, **any** routing scheme must store at least $nH_0(v)$ bits at any node v, where $H_0(v)$ is the Shannon-entropy of the next-hop distribution in s_v
- $nH_0(v)$ bits is attainable, subject to a small error term, with very fast random access *[Ferragina et al., SODA'07]*
- **Routing scalability depends in $H_0(v)$!**

Analysis

- Shortest path routing over the complete graph K_n
- **Bad news:** uniform link weights induce **maximal** forwarding table entropy: $H_0(v) = \lg(n) \underset{n}{\rightarrow} \infty$ bits
- **Good news:** random i.i.d. link weights induce **constant** forwarding table entropy: $\mathbb{E}(H_0(v)) = \lg e \approx 1.44$ bits
- It seems that **heterogeneity** is the key to routing scalability, either topology-wise or policy-wise

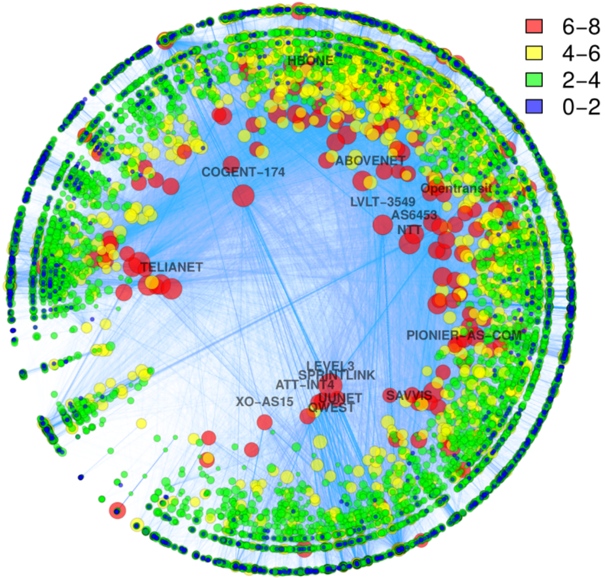
Simulations

- CAIDA AS-level Internet graph with inferred AS-AS business relationships
- Valley-free routing with Gao-Rexford conditions
- Ties broken by shortest AS-path length
- Obtain the per-IP-prefix forwarding table for each AS
- Result is a string of $\sim 500,000$ entries
- Calculate entropy

Validation

- Downloaded IPv4 forwarding tables from two ASes
- `Internet2 (AS11537):` couple of thousand prefixes
  - Reality: $H_0 = 1.3 \ldots 1.7$ bits
  - Simulations: $H_0 = 1.72$ bits
- `HBONE (AS1955):` full-BGP tables with $> 500,000$ prefixes
  - Reality: $H_0 = 1.28$ bits
  - Simulations: $H_0 = 1.26$ bits
- Such a precision is at least suspicious

The Internet scalability map



Discussion

- Forwarding table entropy is surprisingly low
  - below 1 bit at 99% of ASes
  - results 50–70 Kbytes forwarding tables at lower tiers
  - about half a megabyte at the Tier 1
  - 10 million IP prefixes would still yield only 10 Mbytes forwarding tables
- And this is with disregarding address space structure!
- Tier1 pays the price for Internet growth (in terms of entropy)
- Regularity emerges somehow in large-scale forwarding tables

Acknowledgement

- Thanks to `Internet2` and `HBONE` for allowing access to their IP FIBs
- Visit the the *Internet Routing Entropy Monitor* at `http://lendulet.tmit.bme.hu/fib_comp` for daily statistics
- Please, contribute data!