



Hostcount++

Reimplementation of the Hostcount Project

Design team:

- Lee Wilmot
- Henk Uijterwaal
- Brett Carr
- Bruce Campbell

Presented by: Timur Bakeyev



Original Mission Statement

- To enumerate number of hosts in the service region
- Estimate size and growth of Internet
- Reflect distribution of the hosts on per-country basis



Look Behind

- First time Hostcount statistics was published in RIPE-20 document
- October 2, 1990 is the birthday of the project
- Work was done by Marten Terpstra and Daniel Karrenberg
- Totally 19 countries were summarized - AT, BE, CH, DE, DK, ES, FI, FR, GR, IE, IL, IS, IT, NL, NO, PT, SE, UK, YU
- Total number of hosts – 31,724
- Total number of 2nd level domains - 258



Look Behind

- RIPE-76 was last published Hostcount document
- In November 1992 total amount of hosts reached 271,795 and number of CC TLDs reached 38
- Starting from that date Hostcount runs on a regular basis, stats are published on FTP site and sent to the LIRs
- Currently stats are at:
<http://www.ripe.net/hostcount/>



Methodology

- Current implementation of Hostcount described at:
<http://www.ripe.net/hostcount/method.html>
- Hostcount runs every month
- Use of the DNS query tool host written by the Eric Wassenaar at NIKHEF
- Recursively applied to each ccTLD
- Local collectors for a few ccTLDs – more precise data



Technical Problems

- Connectivity problems with certain hosts
- Precision of the data – it's impossible to retrieve all the zone files for any ccTLD
 - Blocked access to certain zones
- Correspondence between ccTLD and geographical country nowadays isn't straightforward
- More than 5 million domains have to be queried on site



Administrative Problems

- Access to raw data requires signing of AUP
- Commercial usage of the raw data due vague AUP, loosely defined usage cases
- AUP not transitional for local ccTLD data collectors – to access their own data they need to sign AUP as well
- There is no explicit permission from local collectors to let RIPE NCC redistribute collected data
- Privacy issues



Proposed Changes

- Add more data sources to increase precision:
 - Forward DNS(exists)
 - Reverse DNS(collected, but not in use)
 - BGP view on availability of prefixes(RIS data)
- Consistent date of publication
- Scalability
- DIY kit for those, who want to to collect stats by themselves

Proposed Changes

- Show trends of Internet development
 - Give more detailed and advanced statistics on the web
 - Correlation between all three data sources
- Nicer graphs 😊
- Not going to ship raw data
 - May give precooked data out(open to discussion)

Questions?

