Understanding the Impact of Network Infrastructure Changes using Large-Scale Measurement Platforms

Vaibhav Bajpai

(Prelude) PhD Proposal Presentation

Computer Networks and Distributed Systems
Jacobs University Bremen
Bremen, Germany

April 2013

overview

- introduction
- state of the art
- research statement
- proposed approach
- preliminary work
- conclusions

introduction

- large-scale broadband measurement use case:

introduction

- **large-scale broadband measurement use case:**
  - **ISP**
    - identify, isolate, fix problems in the access network (or at the last-mile)
    - evaluate the QoE of the user
    - benchmark and look into competitor insights

introduction

- **large-scale broadband measurement use case:**
  - **ISP**
    - identify, isolate, fix problems in the access network (or at the last-mile)
    - evaluate the QoE of the user
    - benchmark and look into competitor insights
  - **consumers**
    - does the network adhere to SLAs?
    - diagnose impaired components in the network path

introduction

- **large-scale broadband measurement use case:**
  - **ISP**
    - identify, isolate, fix problems in the access network (or at the last-mile)
    - evaluate the QoE of the user
    - benchmark and look into competitor insights
  - **consumers**
    - does the network adhere to SLAs?
    - diagnose impaired components in the network path
  - **regulators**
    - need datasets to compare multiple broadband providers
    - frame better policies to help regulate the broadband industry* **

* http://www.fcc.gov/measuring-broadband-america

** http://maps.ofcom.org.uk/broadband/

state of the art

- early studies
  - Dischinger et al.* inject packet trains to infer broadband link characteristics

* M. Dischinger, et al.: Characterizing Residential Broadband Networks, IMC 2007

state of the art

- **early studies**
  - Dischinger et al.* inject packet trains to infer broadband link characteristics
- **software-based solutions**
  - speedtest.net, a flash tool to measure broadband throughput
  - DIMES**, a software agent that performs ping and traceroute tests
  - Glasnost***, a java-applet that detects any traffic shaping enforced by the ISP.
  - netalyzer****, a java-applet that performs DNS, NAT, HTTP, IPv6-based tests
  - Fathom*****, a firefox extension to perform DNS, HTTP, UPnP-based tests

***** M. Dhawan, et al: Fathom: A Browser-based Network Measurement Platform, IMC 2012
**** C. Kreibich, et al: Netalyzr: Illuminating the Edge Network, IMC 2010
*** M. Dischinger, et al: Glasnost: Enabling End Users to Detect Traffic Differentiation, NSDI 2010
** Y. Shavitt, et al: DIMES: Let the Internet Measure Itself, ACM CCR 2005
* M. Dischinger, et al.: Characterizing Residential Broadband Networks, IMC 2007

state of the art

- **large-scale measurement platforms**
  - SamKnows
  - BISmark
  - RIPE Atlas
  - Google' Measurement Lab (M-Lab)
  - CAIDA' Archipelago (Ark)
- **LMAP and IPPM standardization**
  - LMAP BoF at IETF 86, Orlando
  - LMAP framework
    - controller protocols: NETCONF, HTTP REST, IPFIX, Alto ...
    - data-models for message exchange: YANG-based, JSON-based ...
    - metrics definition: IPPM registry for used metrics, model-based metrics
    - regulatory implications
    - standards body competition: IETF, BBF and IEEE

research statement

- understanding the impact of network infrastructure changes **using large-scale measurement platforms**

research statement

- understanding the impact of network infrastructure changes **using large-scale measurement platforms**
    - measure the broadband performance from the residential gateway.
    - help regulators sketch better policy decisions.

research statement

- understanding the impact of network infrastructure changes **using large-scale measurement platforms**
    - measure the broadband performance from the residential gateway.
    - help regulators sketch better policy decisions.
- **understanding the impact of network infrastructure changes** using large-scale measurement platforms

research statement

- understanding the impact of network infrastructure changes **using large-scale measurement platforms**
  - measure the broadband performance from the residential gateway.
  - help regulators sketch better policy decisions.
- **understanding the impact of network infrastructure changes** using large-scale measurement platforms
  - transition to IPv6*:
    - can we identify a CGN from a home gateway?
    - can we identify multiple layers of NAT from a home gateway?

* V. Bajpai, et al.: Flow-based Identification of Failures Caused by IPv6 Transition Mechanisms, AIMS 2012

research statement

- understanding the impact of network infrastructure changes **using large-scale measurement platforms**
  - measure the broadband performance from the residential gateway.
  - help regulators sketch better policy decisions.

- **understanding the impact of network infrastructure changes** using large-scale measurement platforms
  - transition to IPv6*:
    - can we identify a CGN from a home gateway?
    - can we identify multiple layers of NAT from a home gateway?
  - IPv6 today:
    - ~~how rapidly is v6 getting adopted today?~~
    - how does the performance of v6 compare to that of v4?

* V. Bajpai, et al.: Flow-based Identification of Failures Caused by IPv6 Transition Mechanisms, AIMS 2012

research statement

- understanding the impact of network infrastructure changes **using large-scale measurement platforms**
    - measure the broadband performance from the residential gateway.
    - help regulators sketch better policy decisions.

- **understanding the impact of network infrastructure changes** using large-scale measurement platforms
    - transition to IPv6*:
        - can we identify a CGN from a home gateway?
        - can we identify multiple layers of NAT from a home gateway?
    - IPv6 today:
        - ~~how rapidly is v6 getting adopted today?~~
        - how does the performance of v6 compare to that of v4?
    - blend of network centralization and decentralization:
        - to what extend web services centralize on CDNs?
        - to what extend does web experience depend on regionalization?

* V. Bajpai, et al.: Flow-based Identification of Failures Caused by IPv6 Transition Mechanisms, AIMS 2012

proposed approach

proposed approach

- **what do we need?**
  - access to a large-scale measurement platform
    - SamKnows and Jacobs University are partners of the Leone Consortium*
  - address allocations from regional registries (RIR)
  - prefix and AS path information from BGP route views.

* http://www.leone-project.eu

proposed approach

- **what do we need?**
  - access to a large-scale measurement platform
    - SamKnows and Jacobs University are partners of the Leone Consortium*
  - address allocations from regional registries (RIR)
  - prefix and AS path information from BGP route views.

- **what do we need to do?**
  - define metrics targeted to our research questions
  - implement measurement tests that adhere to the metric definition
  - deploy measurement tests on our partner' infrastructure
  - conglomerate the measurement result data from multiple MAs
  - integrate measurement results with data from RIRs and route views
  - prepare data analysis tools that can mine the multi-dimensional data
  - uncover insights to answer the research questions

* http://www.leone-project.eu

preliminary work

- global measurements
  - metrics and measurement tests
    - happy eyeballs
    - webpage similarity search
  - data collection framework
    - using git
    - using IPFIX
    - using REST over HTTP
  - measurement trials
    - using EMANICS nodes
    - using amazon EC2 instances
    - using custom MAs
- local management
  - NETCONF as a controller protocol
    - ncclient support for v1.1
    - NETCONF server on SK probes
    - NETCONF over TLS (cont.)

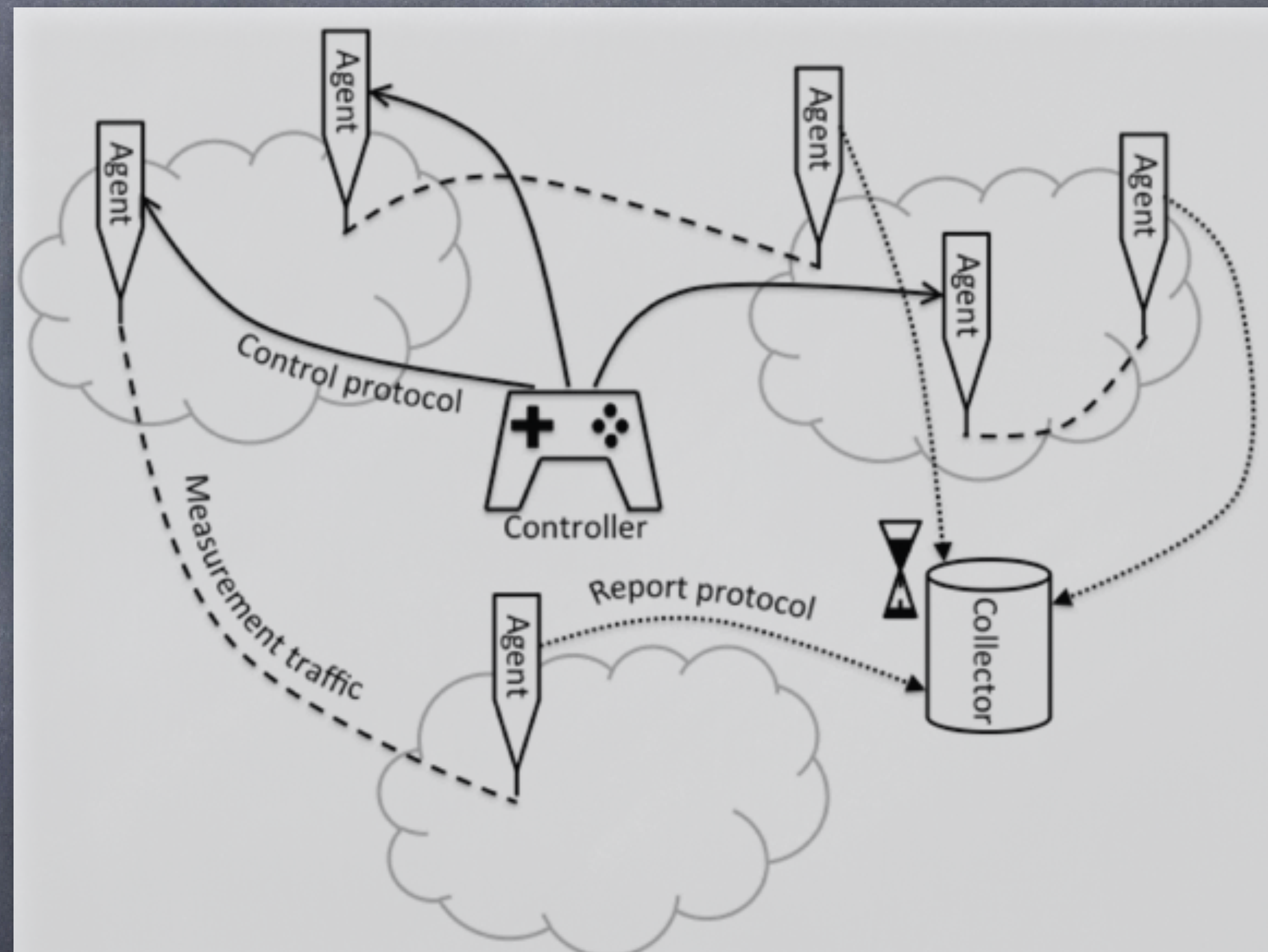
preliminary work

- global measurements
  - metrics and measurement tests
    - happy eyeballs
    - webpage similarity search
  - data collection framework
    - using git
    - using IPFIX
    - using REST over HTTP
  - measurement trials
    - using EMANICS nodes
    - using amazon EC2 instances
    - using custom MAs
- local management
  - NETCONF as a controller protocol
    - ncclient support for v1.1
    - NETCONF server on SK probes
    - NETCONF over TLS (cont.)

data collection framework using HTTP REST

Naming:

- collector must have a unique URI [RFC 3986]
- agents are identified by their mac address tied to their certificate identity
- each measurement cycle is identified by a UUID [RFC 4122]
- UUID is obtained from the collector when starting the cycle.



*

* M. Bagnulo, et. al: Standardizing Large-Scale Measurement Platforms, ACM CCR, July 2013

data collection framework using HTTP REST

HTTP GET	/uuid	debugging purposes
HTTP GET	/msmpoint	list available msmpoints
HTTPS POST	/msmpoint @data	create a new msmpoint
HTTP GET	/msmpoint/<msmid>	get the metadata of <msmid>
HTTPS PUT	/msmpoint/<msmid> @data	update the metadata of <msmid>
HTTPS DELETE	/msmpoint/<msmid>	delete <msmid> (cascading effect)
HTTP GET	/<test>	list msmpoints running <test>
HTTP GET	/<test>/<msmid>	list result metadata of <test> from <msmid>
HTTP GET	/<test>/<msmid>/result	get results of <test> from <msmid>
HTTPS POST	/<test>/<msmid>/result @data	send results of <test> from <msmid>

- HTTP {GET, HEAD, OPTIONS} may or may not be encrypted
- HTTP {POST, PUT, DELETE, PATCH} must be encrypted
- HTTP over TLS [RFC 5246] use peer-authentication using X.509 certificates [RFC 5280]
- The data exchanged between the parties is encoded in JSON format [RFC 4627]
- The data model that define the semantics of encoded data need definition (I-D needed)

data collection framework using HTTP REST

Examples:

```
$ curl -X GET http://<fqdn>/msmpoint
[
    { "msmid"      :     "b4ac4ec087e211e2b10ffff08147c934" }
  , { "msmid"      :     "bdf8bc9a8eeeeee2a74900e08147c934" }
]
```

```
$ curl -X GET http://<fqdn>/msmpoint/<msmid>
{
     "name"        :     "John Doe"
  ,  "asn"         :     "6080"
  ,  "mac"         :     "00:aa:b9:bb:51:cc"
  ,  "location"    :     "New York"
  ,  "msmid"       :     "385054687e311e2b10d00e08147c934"
  ,  "email"       :     "j.doe@email.org"
}
```

```
$ curl -X GET http://<fqdn>/<test>/<msmid>
{
     "start"       :     1361462435
  ,  "end"         :     1363956617
  ,  "result"      :     "/<test>/<msmid>/result"
  ,  "resultcount" :     1018209
}
```

```
$ curl -X GET http://<fqdn>/<test>/<msmid>/result
[
  {
       "status"      :     "OK"
     , "endpoint"    :     "2a03:2880:10:cf01:face:b00c::4"
     , "service"     :     "www.facebook.com"
     , "timestamp"   :     1362650067
     , "msmid"       :     "3eeeeee887e311e2b10d00e08147c934"
     , "version"     :     "HAPPY.0"
     , "time"        :     167380
     , "port"        :     80
  },
  ...
]
```

happy eyeballs

- getaddrinfo(...) behavior:
  - returns list of endpoints in an order that prioritizes IPv6 upgrade path
  - order is dictated by [RFC6724] and `/etc/gai.conf`
  - if IPv6 is broken, application is unresponsive in order of seconds
- happy eyeballs algorithm:
  - initiate a connection with the first endpoint, give it 300ms
  - switch over to a different address family otherwise
  - the competition runs fair after 300ms

happy

- simple TCP happy eyeballs [RFC 6555] probing tool
  - it probes all endpoints of a service name though
    - helps compare different happy eyeball algorithms
  - uses `getaddrinfo(...)` to resolve service names to endpoints
  - uses non-blocking `connect(...)` to connect to all endpoints of a service
  - uses a short-delay between connection attempts to avoid SYN floods
  - returns connection timings to each endpoint
  - produce human-readable or machine-readable output
  - file locking capability
  - cross-compiled for openWRT, running on SK boxes.

```
>> ./happy -q 1 -m www.google.com www.facebook.com
HAPPY.0;1360681039;OK;www.google.com;80;173.194.69.105;8626
HAPPY.0;1360681039;OK;www.google.com;80;2a00:1450:4008:c01::69;8884
HAPPY.0;1360681039;OK;www.facebook.com;80;2a03:2880:10:6f01:face:b00c::8;170855
HAPPY.0;1360681039;OK;www.facebook.com;80;31.13.72.39;26665
```

service names list

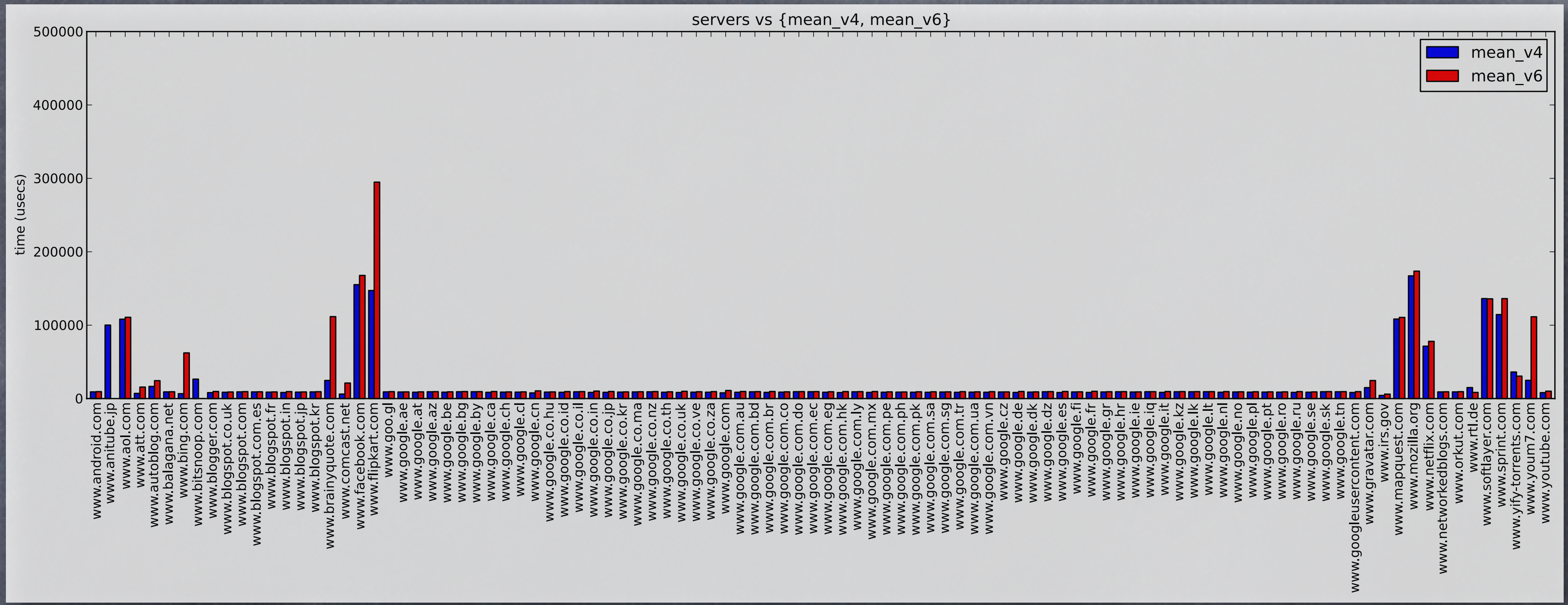
- what is a good set of service names?
  - HE.net maintains a list of top 100 dual-stacked service names
    - they use 1M service names from alexa
    - prepare a top 100 dual-stacked service name list from this input
    - some domains we expect are missing from the list
    - some services only provide a IPv6 endpoint on prepending a www
    - HE.net does not follow CNAMEs
  - amazon has made 1M service name list public
    - we use it and script it ourselves to explicitly follow CNAMEs

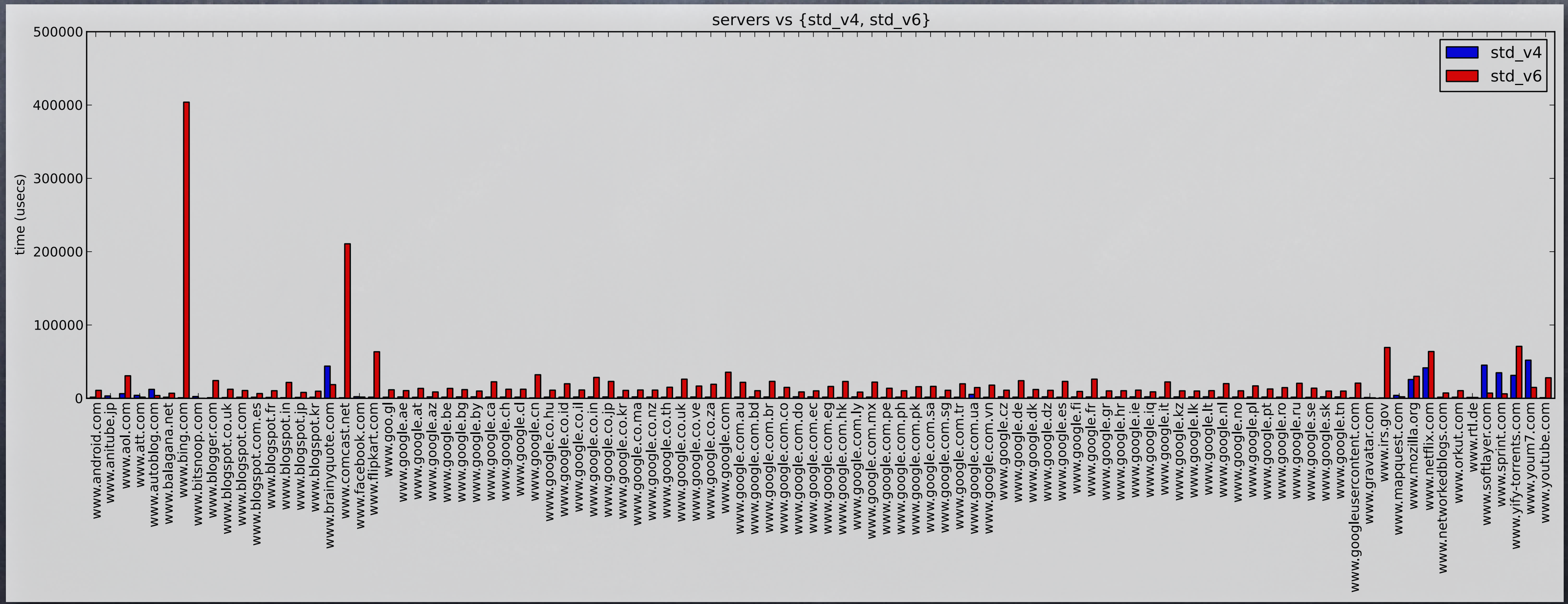
measurement trial

- dual-stacked service name list:
  - custom generated from amazon' publicly available top 1M service names
- measurement points:
  - native IPv6, 6in4, IPv6 Teredo, 4in4, native IPv4
  - located at Bremen, Amsterdam, Braunschweig
  - tunnels popping out at Frankfurt, Berlin, New York
- runtime frequency:
  - every 3 minutes and collect data in a local file
- data collection:
  - sync the local file with a remote data collection endpoint
- measurement cycle length:
  - 1 week

how does IPv6 compare in performance to IPv4?

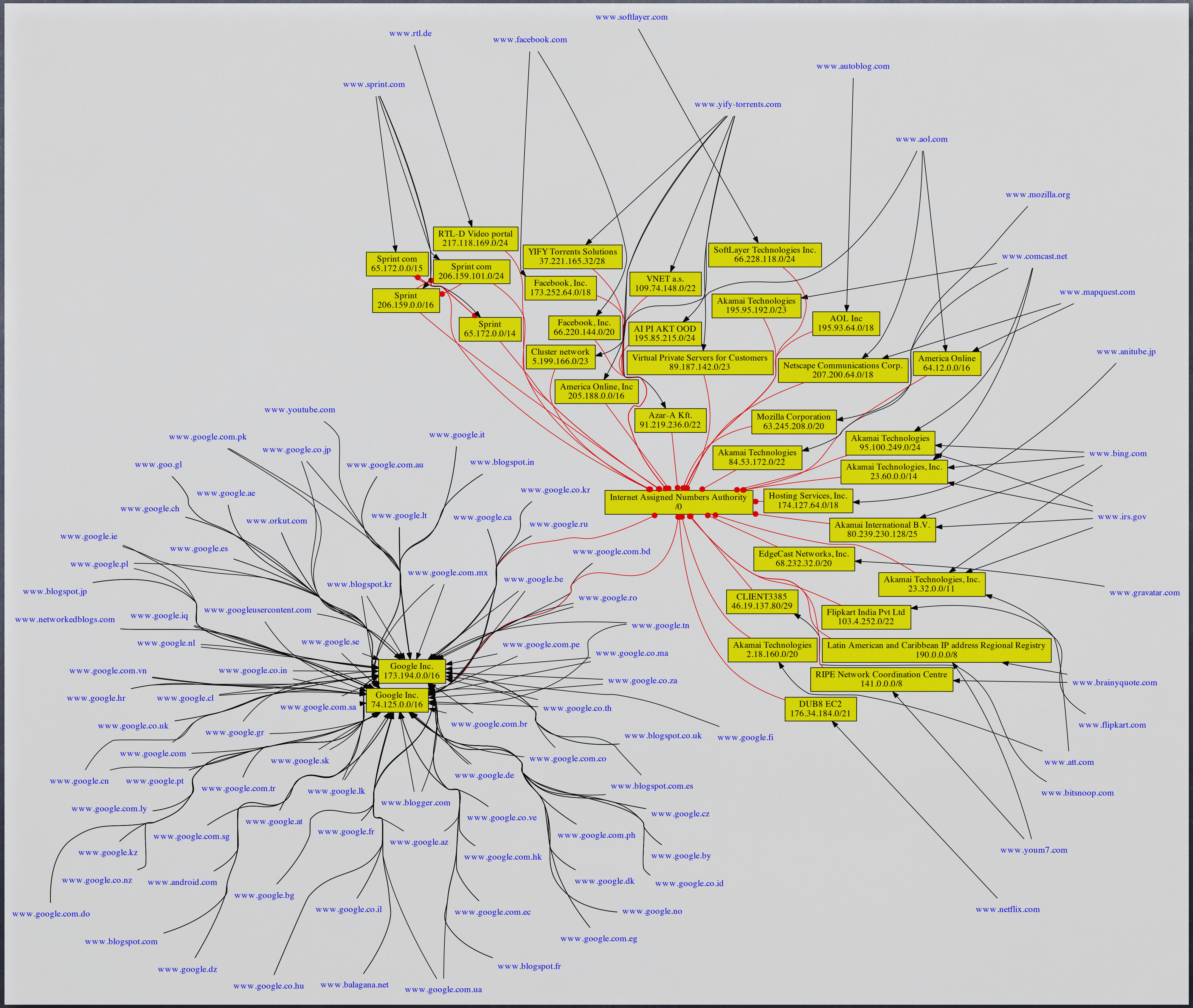
connection times (mean/std)



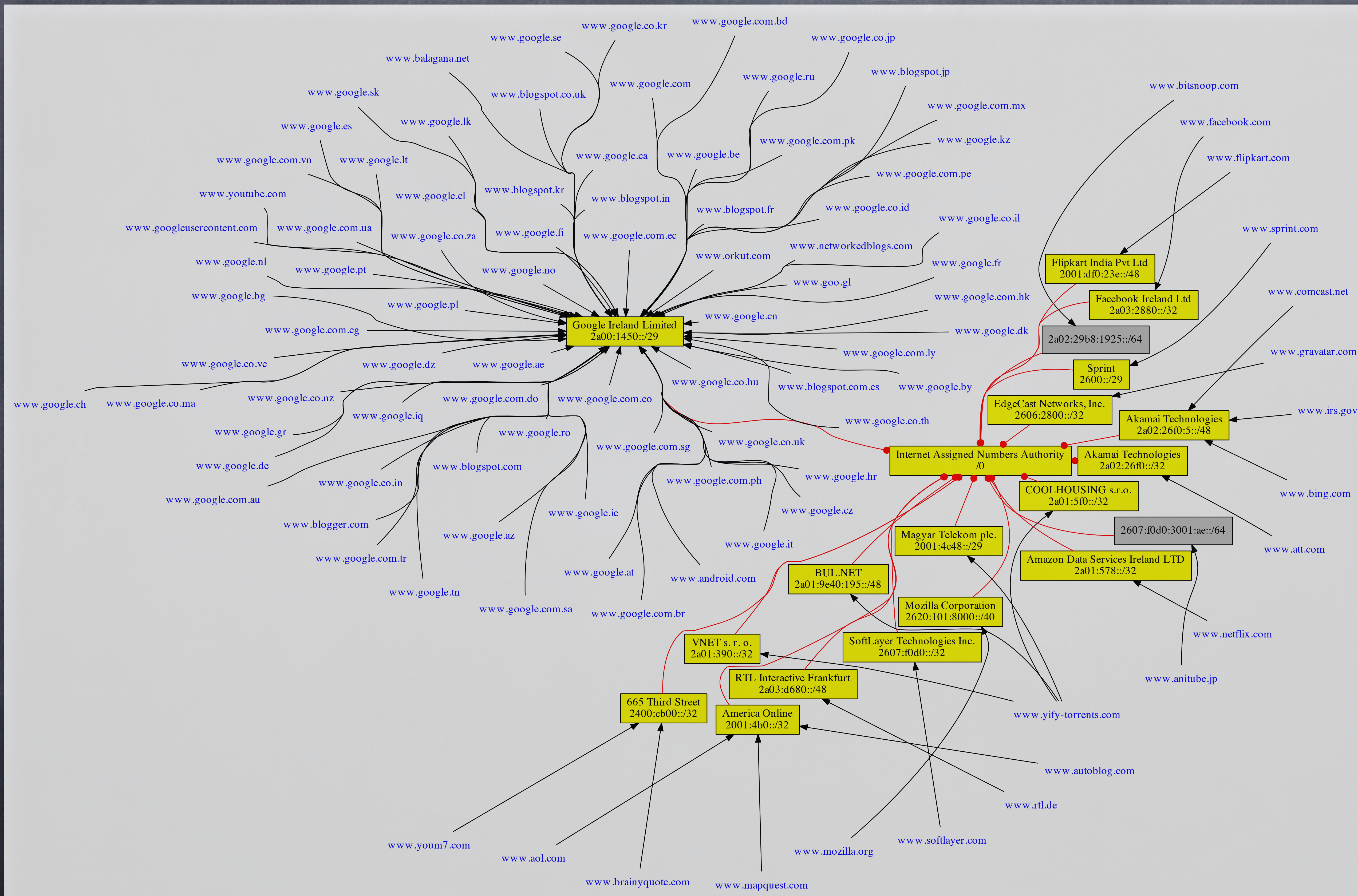


do a major portion of services centralize on CDNs?

IPv4 aggregation cloud



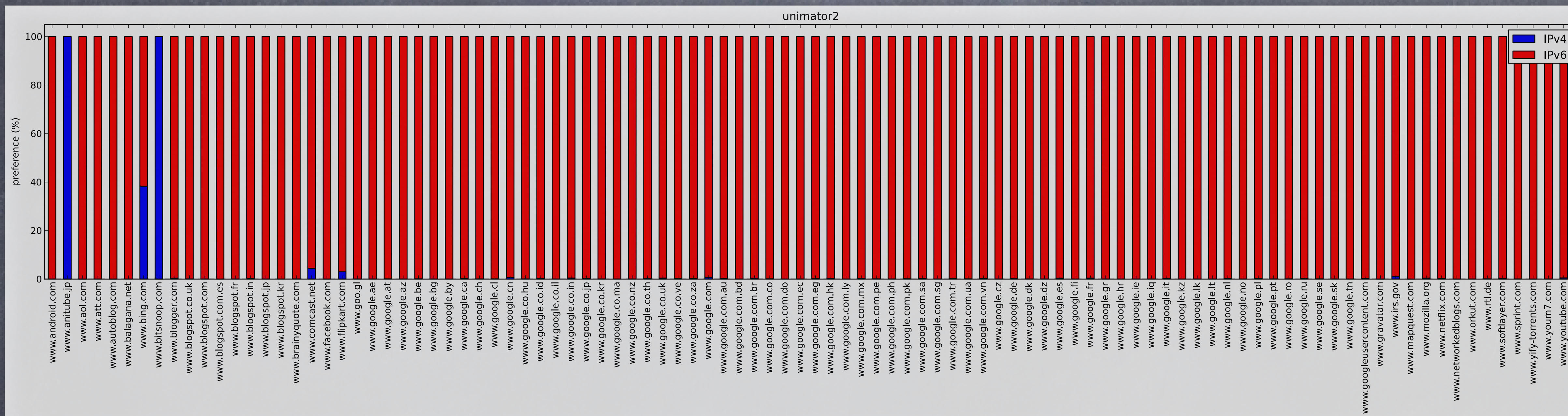
IPv6 aggregation cloud



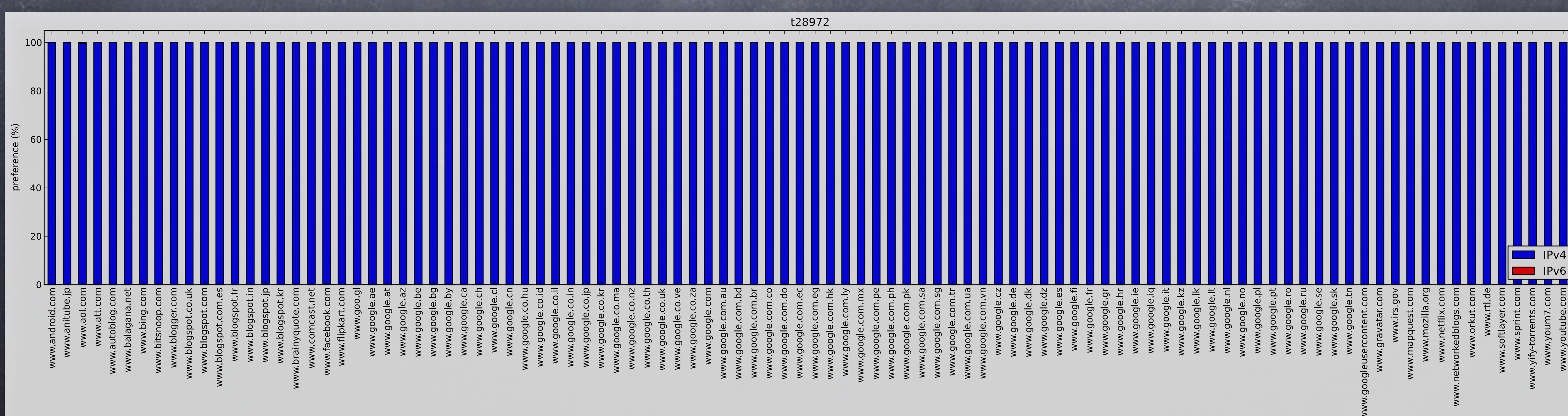
to what extent is IPv6 preferred when connecting to a dual-stacked service?

IPv6 preference %

Native IPv6



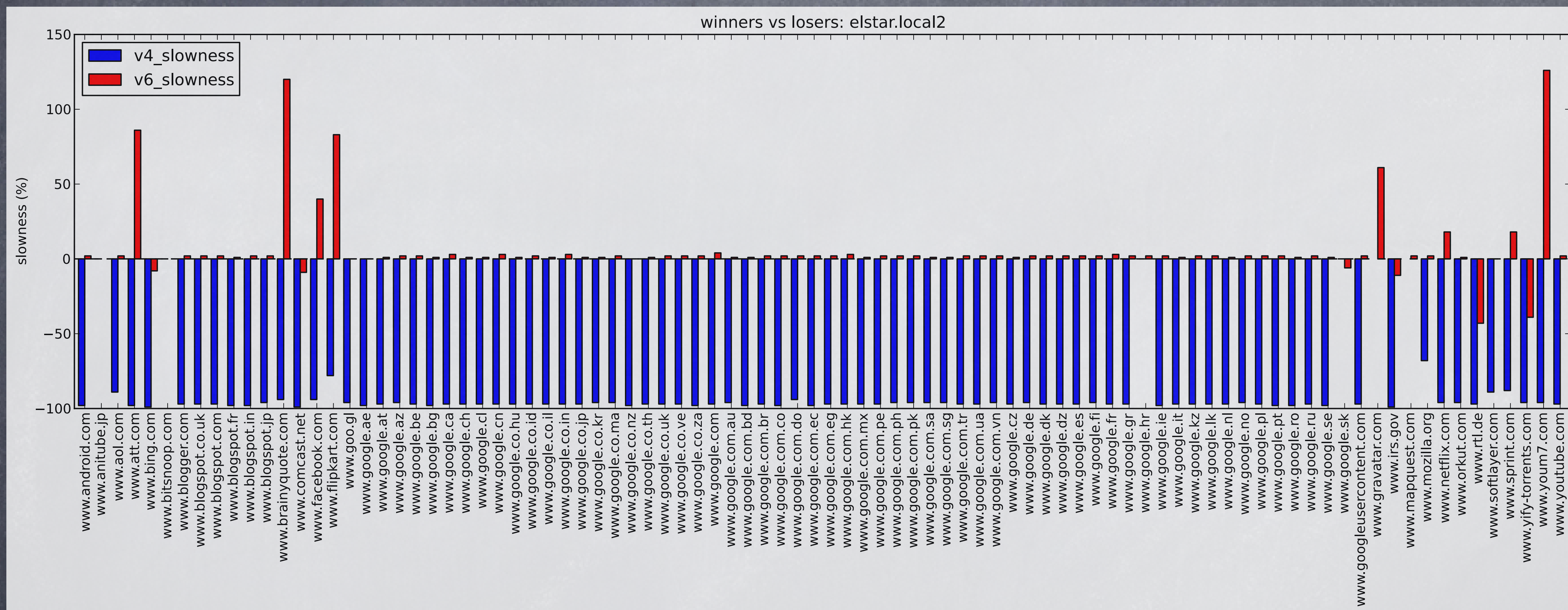
IPv6 teredo tunnel



how slow is the happy eyeballed winner to that of the loser?

IPv6 preference %

Native IPv6



lessons learned

- **happy:**
  - higher connection times and variations over v6
  - application never uses Teredo except when v4 connectivity is broken
  - a 300ms advantage leaves a dual-stacked host only 1% chance to prefer v4
  - a dual-stacked host pays a v6 tax; rarely is the happy eyeballed winner faster

    happy eyeballs algorithm ironically hampers your eyeballs

- **REST with LMAP MA**
  - peer-authentication using X.509 certificates is not well-supported
  - REST over HTTP seems to work well in practice
  - curl and python flask are wonderful implementations