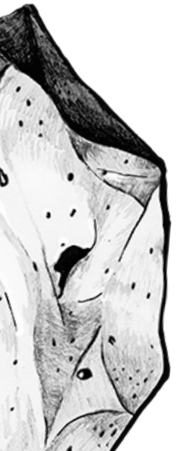
Managing Peering as we Scale

Andy **Davidson**, andy@asteroidhq.com

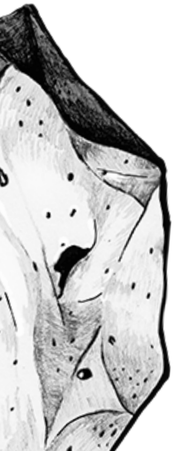
Life is too short for manual peering

I just joined this exchange, would you like to peer?



I just joined this exchange, would you like to peer?

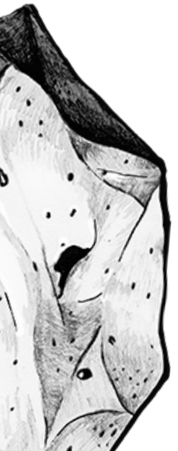
Sorry for the delay to my reply, sure, whats your peering IP?



I just joined this exchange, would you like to peer?

Sorry for the delay to my reply, sure, whats your peering IP?

Sorry for the delay to my reply, it is 198.51.100.20

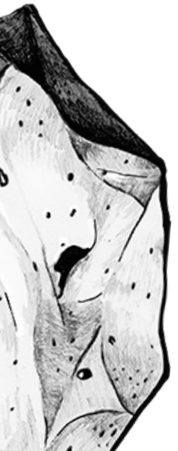


I just joined this exchange, would you like to peer?

Sorry for the delay to my reply, sure, whats your peering IP?

Sorry for the delay to my reply, it is 198.51.100.20

That is my IP address.



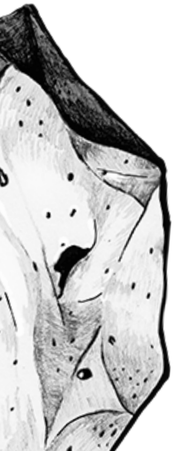
I just joined this exchange, would you like to peer?

Sorry for the delay to my reply, sure, whats your peering IP?

Sorry for the delay to my reply, it is 198.51.100.20

That is my IP address.

Oh, Lol, Soz, it’s 198.51.100.30. Sessions are waiting.



I just joined this exchange, would you like to peer?

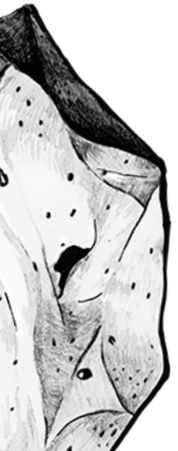
Sorry for the delay to my reply, sure, whats your peering IP?

Sorry for the delay to my reply, it is 198.51.100.20

That is my IP address.

Oh, Lol, Soz, it's 198.51.100.30. Sessions are waiting.

Hello?



I just joined this exchange, would you like to peer?

Sorry for the delay to my reply, sure, whats your peering IP?

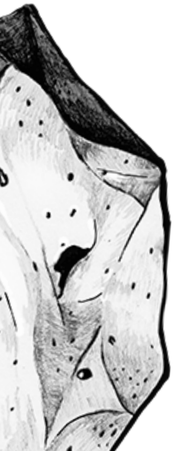
Sorry for the delay to my reply, it is 198.51.100.20

That is my IP address.

Oh, Lol, Soz, it's 198.51.100.30. Sessions are waiting.

Hello?

Sorry for the delay to my reply, Bill has now left the company.
Does this ticket still require action?



I just joined this exchange, would you like to peer?

Sorry for the delay to my reply, sure, whats your peering IP?

Sorry for the delay to my reply, it is 198.51.100.20

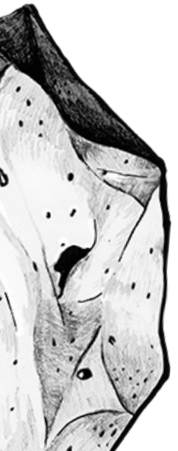
That is my IP address.

Oh, Lol, Soz, it's 198.51.100.30. Sessions are waiting.

Hello?

Sorry for the delay to my reply, Bill has now left the company.
Does this ticket still require action?

How much traffic is this session actually "worth"?



Imagine

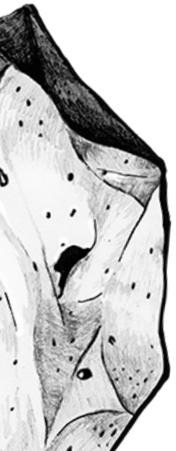
Imagine a world where computers would help us make realtime peering decisions

… With known and predictable outcomes

… And error free provisioning of peering

… For free

… Today



For this to happen we need:

Our Networks:
> Software to help us understand the network and our traffic (cost/volume, latency, value)
> Software to reliably orchestrate

Open data:
> Peering network intentions
> Authoritative information about exchange data

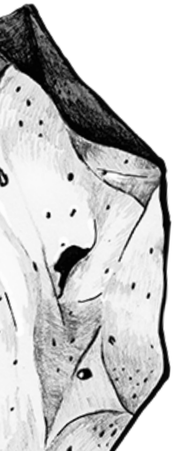
Ecosystem:
> Low friction data exchange norms
> More peering opportunity

Our Networks

> Software to help us understand the network and our traffic (cost/volume, latency, value)

Cost and Volume is largely solved now, assuming you can export Netflow

Pmacct is the most pluggable store of sorted volume data



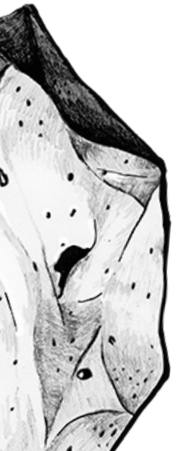
pmacct

http://www.pmacct.net

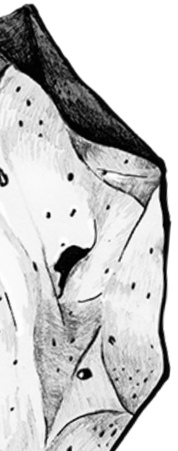
Open Source
High performance, high scale, powerful
Most flexible, most configuration required

Collector -> own reports

```
sfacctd_port: 2100
sfacctd_as: sflow
sfacctd_renormalize: true
!
plugins: print[testprint]
!
aggregate[testprint]: in_iface, out_iface, proto, peer_src_ip, peer_dst_ip, peer_dst_as, peer_src_as, src_as, dst_as
!
print_output_file[testprint]: /path/to/spool/blabla-$peer_src_ip-%Y%m%d-%H%M.txt
print_output[testprint]: csv
print_output_separator[testprint]: ;
print_refresh_time[testprint]: 60
print_history[testprint]: 1m
print_history_roundoff[testprint]: m
```



SRCAS;DSTAS;PEER_SRC_AS;PEER_DST_AS;PEER_SRC_IP;PEER_DST_IP;IN_IFACE;OUT_IFACE;PROTOCOL;PACKETS;BYTES
4x0;224;4x0;2603;x.x.x.253;x.x.x.246;3;4;tcp;2048;151552
4x0;15169;4x0;15169;x.x.x.253;x.x.x.246;3;4;tcp;10240;880640
4x0;50247;4x0;24724;x.x.x.253;x.x.x.246;3;4;tcp;2048;167936
4x0;9269;4x0;1273;x.x.x.253;x.x.x.237;3;3;tcp;2048;135168
4x0;3356;4x0;1273;x.x.x.253;x.x.x.237;3;3;tcp;32768;2375680
4x;209;4x0;1273;x.x.x.253;x.x.x.237;3;3;udp;2048;2940928



```
plugins: mysql[5mins], mysql[hourly]

sql_optimize_clauses: true
sql_dont_try_update: true
sql_multi_values: 1024000

sql_history_roundoff[5mins]: m
sql_history[5mins]: 5m
sql_refresh_time[5mins]: 300
sql_table[5mins]: acct_bgp_5mins

sql_history_roundoff[hourly]: h
sql_history[hourly]: 1h
sql_refresh_time[hourly]: 3600
sql_table[hourly]: acct_bgp_1hr

plugin_buffer_size: 10240
plugin_pipe_size: 1024000
aggregate: tag, src_as, dst_as, peer_src_as, peer_dst_as, peer_src_ip,
                                peer_dst_ip, local_pref, as_path
```

```
mysql> SELECT * FROM int_tm-20130803_1400 LIMIT 10;
+----------+-------------+-------------+-------------+------------------+-------+
| iface_in | peer_ip_src | peer_ip_dst | peer_dst_as | stamp_inserted   | bytes |
+----------+-------------+-------------+-------------+------------------+-------+
|      212 | 10.0.0.107  | 10.0.0.3    |       65000 | 03-08-2013 14:00 |   859 |
|      212 | 10.0.0.107  | 10.0.0.253  |       65001 | 03-08-2013 14:00 |  5358 |
|      212 | 10.0.0.107  | 10.0.0.234  |       65002 | 03-08-2013 14:00 |  6181 |
|      212 | 10.0.0.107  | 10.0.0.251  |       65003 | 03-08-2013 14:00 | 27002 |
|      205 | 10.0.0.107  | 10.0.0.233  |       65004 | 03-08-2013 14:00 |  1200 |
|      258 | 10.0.0.107  | 10.0.0.240  |       65005 | 03-08-2013 14:00 |   560 |
|      212 | 10.0.0.107  | 10.0.0.252  |       65006 | 03-08-2013 14:00 | 62682 |
|      212 | 10.0.0.107  | 10.0.0.234  |       65007 | 03-08-2013 14:00 |  3843 |
|      212 | 10.0.0.107  | 10.0.0.17   |       65008 | 03-08-2013 14:00 | 21074 |
|      205 | 10.0.0.107  | 10.0.0.254  |       65009 | 03-08-2013 14:00 |  2023 |
```

```
mysql> SELECT * FROM int_tm-20130803_1400 LIMIT 10;
+----------+-------------+-------------+-------------+------------------+-------+
| iface_in | peer_ip_src | peer_ip_dst | peer_dst_as | stamp_inserted   | bytes |
+----------+-------------+-------------+-------------+------------------+-------+
|      212 | 10.0.0.107  | 10.0.0.3    |       65000 | 03-08-2013 14:00 |   859 |
|      212 | 10.0.0.107  | 10.0.0.253  |       65001 | 03-08-2013 14:00 |  5358 |
|      212 | 10.0.0.107  | 10.0.0.234  |       65002 | 03-08-2013 14:00 |  6181 |
|      212 | 10.0.0.107  | 10.0.0.251  |       65003 | 03-08-2013 14:00 | 27002 |
|      205 | 10.0.0.107  | 10.0.0.233  |       65004 | 03-08-2013 14:00 |  1200 |
|      258 | 10.0.0.107  | 10.0.0.240  |       65005 | 03-08-2013 14:00 |   560 |
|      212 | 10.0.0.107  | 10.0.0.252  |       65006 | 03-08-2013 14:00 | 62682 |
|      212 | 10.0.0.107  | 10.0.0.234  |       65007 | 03-08-2013 14:00 |  3843 |
|      212 | 10.0.0.107  | 10.0.0.17   |       65008 | 03-08-2013 14:00 | 21074 |
|      205 | 10.0.0.107  | 10.0.0.254  |       65009 | 03-08-2013 14:00 |  2023 |
```

Limit to specific interfaces

Group by src/dst ASN and time

Sum the bytes

Our Networks

> Software to help us understand the network and our traffic (cost/volume, latency, value)

I am not sure that latency sensitive potential peers and high company value potential peers so simple to measure

- performance problem scraping from trouble ticket data?
- internal audit?
- expectation that volumes follow value?

Discussion point for tonight!

Our Networks

> Software to reliably orchestrate

Largely a solved problem

If tin doesn't have an API, do not buy it.
- Make sure the API is not simply a front end to the CLI

Developing confidence
- Start by writing 'read only' utilities
- Install virtual environments for your network assets
- Automated testing

Open Data

> Peering Network Intentions

- PeeringDB is the best resource we have, and good for location based data
- Coarse specification of peering openness
- No common way to express peering policy or approach programatically
  - “We prefer PNI in Frankfurt”
  - “Public only below 1Gbps”
- Not authoritative for IXP “owned” data elements

Open Data

> Authoritative information about exchange data

- Euro-IX JSON schema
- IXPs authoritative for IX elements, e.g. peering IP addresses
- Complete, accurate, automatic

GET https://peeringdb.com/api/net/8736

Body Cookies Headers (8) Tests

Pretty Raw Preview JSON


```
    "status": "ok"
  },
  "name": "TREX Regional Exchanges Oy",
  "aka": "TREX",
  "website": "http://www.trex.fi/",
  "asn": 29432,
  "looking_glass": "",
  "route_server": "",
  "irr_as_set": "AS29432:AS-PEERING",
  "info_type": "Non-Profit",
  "info_prefixes4": 7,
  "info_prefixes6": 5,
  "info_traffic": "0-20 Mbps",
  "info_ratio": "Balanced",
  "info_scope": "Regional",
  "info_unicast": true,
  "info_multicast": true,
  "info_ipv6": true,
  "notes": "TREX Tampere is an IXP in Finland.",
  "policy_url": "http://www.trex.fi/peering.html",
  "policy_general": "Open",
  "policy_locations": "Not Required",
  "policy_ratio": false,
  "policy_contracts": "Required",
  "netfac_set": [
    {
      "id": 17956,
      "name": "Tampere Hallituskatu 14",
      "city": "Tampere",
      "country": "FI",
      "fac_id": 2333,
      "local_asn": 29432,
      "created": "2015-09-23T00:00:00Z",
      "updated": "2016-03-14T20:25:04Z",
      "status": "ok"
    }
  ],
  "netixlan_set": [
    {
      "id": 24674,
      "ix_id": 328,
      "name": "TREX: Unicast Peering",
      "ixlan_id": 328,
      "notes": "",
      "speed": 1000,
      "asn": 29432,
      "ipaddr4": "195.140.192.9",
      "ipaddr6": "2001:7f8:1d:4::72f8:1",
      "is_rs_peer": true,
      "created": "2015-09-01T00:00:00Z",
      "updated": "2016-07-01T06:47:04Z",
      "status": "ok"
    },
```


GET http://www.trex.fi/memberlist.json

Body Cookies Headers (11) Tests

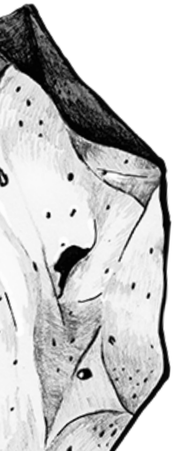
Pretty Raw Preview JSON

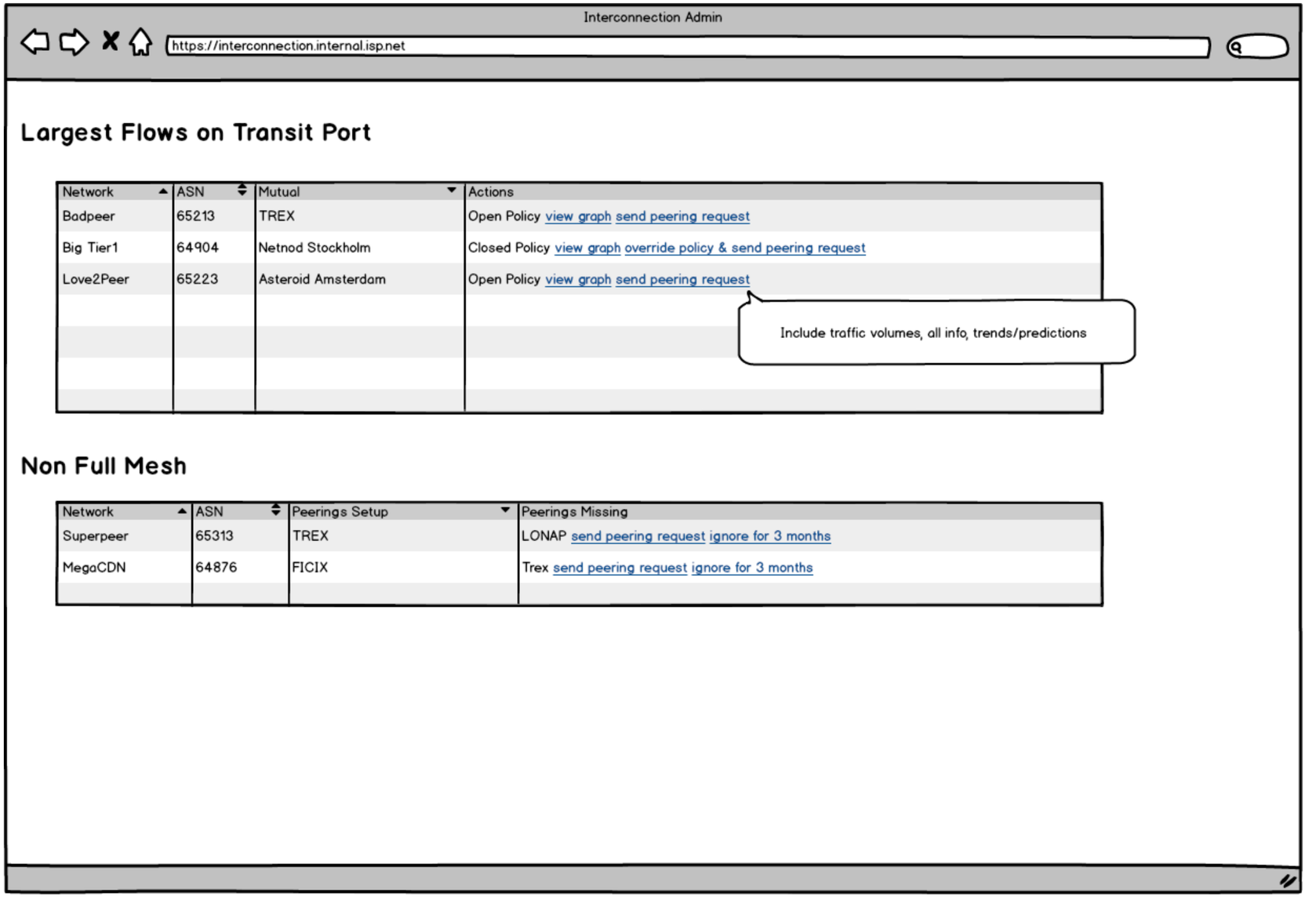

```
  {
    "connection_list": [
      {
        "vlan_list": [
          {
            "vlan_id": 4,
            "ipv4": {
              "max_prefix": 1500,
              "as_macro": "AS-EUNETIP",
              "address": "195.140.192.13"
            },
            "ipv6": {
              "max_prefix": 1500,
              "address": "2001:7f8:1d:4::1a0b:1",
              "as_macro": "AS-EUNETIP"
            }
          }
        ],
        "ixp_id": 46
      }
    ],
    "name": "Elisa/Saunalahti",
    "url": "http://www.elisaip.net/",
    "member_type": "peering",
    "asnum": 6667
  },
  {
    "connection_list": [
      {
        "ixp_id": 46,
        "vlan_list": [
          {
            "ipv4": {
              "address": "195.140.192.11",
              "as_macro": "AS-MMD"
            },
            "ipv6": {
              "as_macro": "AS-MMD",
              "address": "2001:7f8:1d:4::723b:1"
            },
            "vlan_id": 4
          }
        ]
      }
    ],
    "asnum": 29243,
```

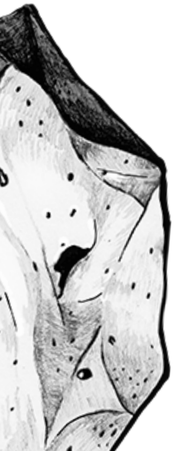

So what do we sort of have so far?

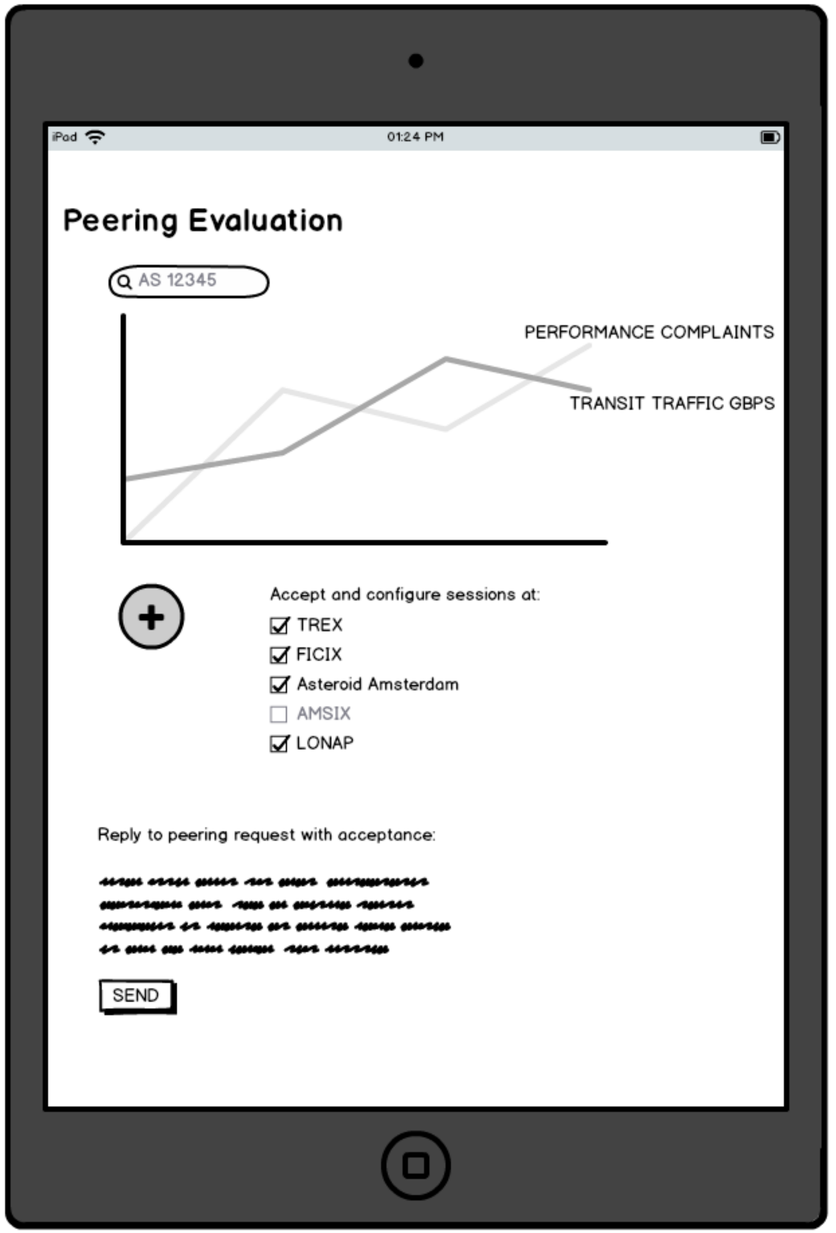
Building blocks:

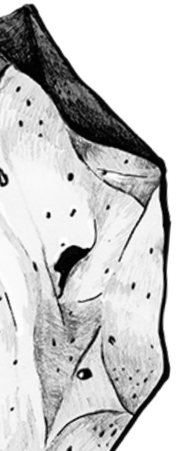
- Lots of data from our own networks about volumes
- Work to do to get data about quality, performance, availability, latency…
- Good orchestration capability
- Excellent coverage about network peering intentions
- Good, and growing coverage of exchange data




Interconnection Admin
https://interconnection.internal.isp.net
Largest Flows on Transit Port
Network
ASN
Mutual
Actions
Badpeer
65213
TREX
Open Policy view graph send peering request
Big Tier1
64904
Netnod Stockholm
Closed Policy view graph override policy & send peering request
Love2Peer
65223
Asteroid Amsterdam
Open Policy view graph send peering request
Include traffic volumes, all info, trends/predictions
Non Full Mesh
Network
ASN
Peerings Setup
Peerings Missing
Superpeer
65313
TREX
LONAP send peering request ignore for 3 months
MegaCDN
64876
FICIX
Trex send peering request ignore for 3 months




iPad
01:24 PM
Peering Evaluation
AS 12345
PERFORMANCE COMPLAINTS
TRANSIT TRAFFIC GBPS
Accept and configure sessions at:
TREX
FICIX
Asteroid Amsterdam
AMSIX
LONAP
Reply to peering request with acceptance:
SEND



Ecosystem

> Low friction data exchange norms

Not just for participant data

For service instantiation, management, augmentation, removal…

“API First”

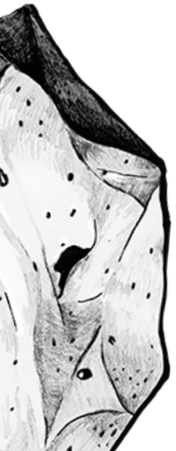
Transparent reporting

My view is your view

Ecosystem

> More peering opportunity

Watch this space :-)



Asteroid

Questions and comments?

andy@asteroidhq.com // @andyd // @asteroidhq