CAIDA participation in PREDICT

- **Provider role:** what data are we collecting, how are we curating and serving?

- **Host role:** what data are we hosting, who are we trying to recruit?

- **Researcher role:** anything security-related and/or useful being done w the data?

- **Fires burning brightest:** which ones is PREDICT trying to douse?

what data are we collecting?

- OC192 backbone: 734GB (DITL2008)
- OC48 traces: 1.7TB (same old 2004 traces)
- UCSD telescope: 80.2TB (backsctr, worm)
- topology: 4.2TB (232GB routed ipv4), ipv6 as of jan08

   Total: 86.8TB (as of 31 March 2008)

how are we curating the data?

- OC192 backbone: still cleaning. sensitive.
- OC48 traces: strip payload/L1/L2, anonymized w (prefix-preserve) cryptopan
- UCSD telescope: clean out suspected user data, curate backscatter and worm data separately.
- topology: see cybersecurity project

(http://www.caida.org/home/legal)

how many requests for the data?

- OC192 backbone: not announced
- OC48 traces: 15 (370 cumulative)
- UCSD telescope: 33 (413)
- topology: 46 (552)

Total: 2.43TB (1335 requests cum.)

how are we serving the data?

- **OC192 backbone:** hopefully report generator up soon, traces same as OC48
- **OC48 traces:** to academics who sign
  http://www.caida.org/data/passive/anon_internet_traces_request.xml
- **UCSD telescope:** to academics who sign
  http://www.caida.org/data/passive/network_telescope.xml#access
- **topology:** 4.2TB (232GB routed ipv4):
  http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml

IRB: first draft of application, 2 may 08
(http://www.caida.org/data/statistics/)

how is the data being used?

- **OC192 backbone:** hopefully report generator up soon, traffic classification

- **OC48 traces:** traffic classification, modeling, monitoring, filtering, generation, locality
http://www.caida.org/data/publications/bydataset/index.xml#OC48

- **UCSD telescope:** worm monitoring
http://www.caida.org/data/publications/bydataset/index.xml#backscatter

- **topology:** pkt traceback, marking. DOS defense. topo and routing modeling, discovery, metrics, improvements
http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml

(http://www.caida.org/data/publications/)

how are we using the data?

- OC192 backbone: dispel myths in .gov, .edu

- OC48 traces: traffic classification, modeling, p2p,
(also http://www.caida.org/data/realtime/passive/?monitor=sdnap)

- UCSD telescope: traffic classification, real-time
monitor   http://www.caida.org/data/realtime/telescope/

- topology: doug's cybersecurity project
http://www.caida.org/research/topology/

(www.caida.org/publications/papers/)

what other data are we seeking?

- OC192 backbone: longer traces, other sites

- Internet2: better netflow, pkt traces, report gen.
  http://www.caida.org/data/realtime/passive/?monitor=sdnap)

- CENIC: hoping same as I2, also supporting COMMONS

- COMMONS: participating community networks
  http://www.caida.org/projects/commons

(www.caida.org/publications/papers/)

concerns i have about PREDICT
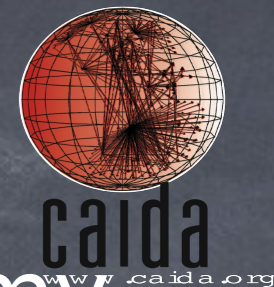
- **anonymization situation:** we have met enemy
- **response to ohm07 paper:** crucial
- **no govt use of data:** needs clarification
- **no networks that serve public:** needs repair
- **no usable metadata catalog:** e.g., datcat
- **metrics for success:** what are we solving? top 10 list
- **need more lawyers:** aaron, simson's papers a good start, need wiki, outreach, bofs at conferences
- **privacy impact statement:** needs repair
- **improved PR:** how to raise awareness (blog.caida.org)

IPv4 INTERNET
TOPOLOGY MAP
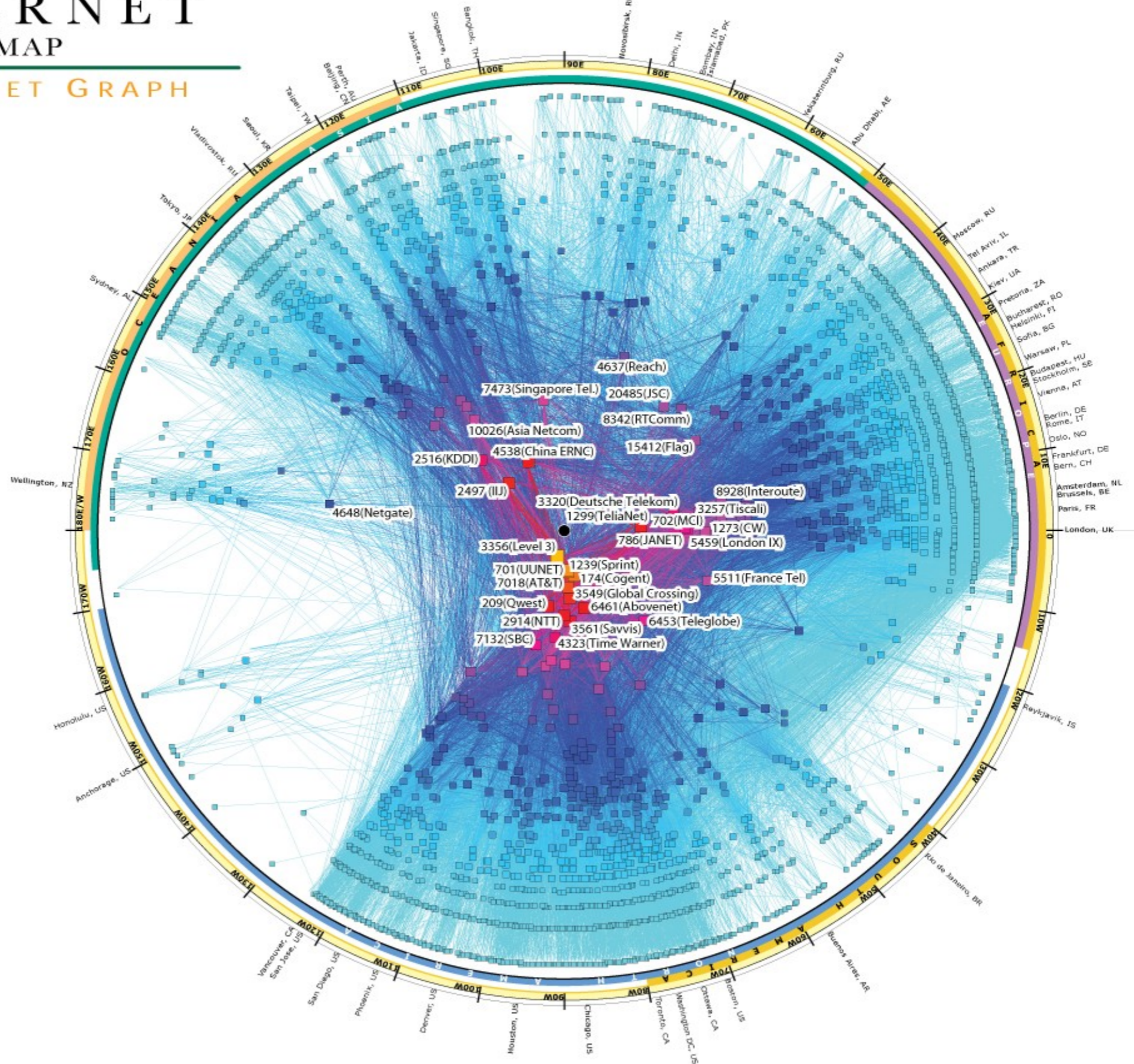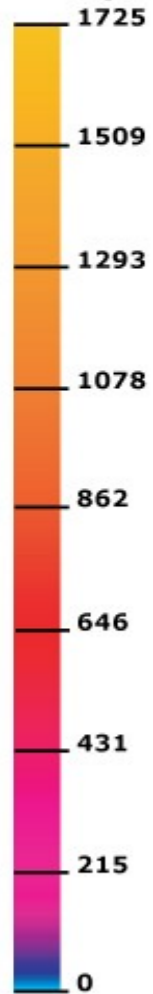### AS-level INTERNET GRAPH

Peering:
OutDegree

1725
1509
1293
1078
862
646
431
215
0

4637(Reach)
7473(Singapore Tel.)   20485(JSC)
8342(RTComm)
10026(Asia Netcom)
15412(Flag)
4538(China ERNC)
2516(KDDI)
2497, (IIJ)
8928(Interoute)
3320(Deutsche Telekom)   3257(Tiscali)
4648(Netgate)   1299(TeliaNet)   702(MCI)
1273(CW)
786(JANET)   5459(London IX)
3356(Level 3)
1239(Sprint)
701(UUNET)   174(Cogent)
7018(AT&T)   5511(France Tel)
3549(Global Crossing)
209(Qwest)   6461(Abovenet)
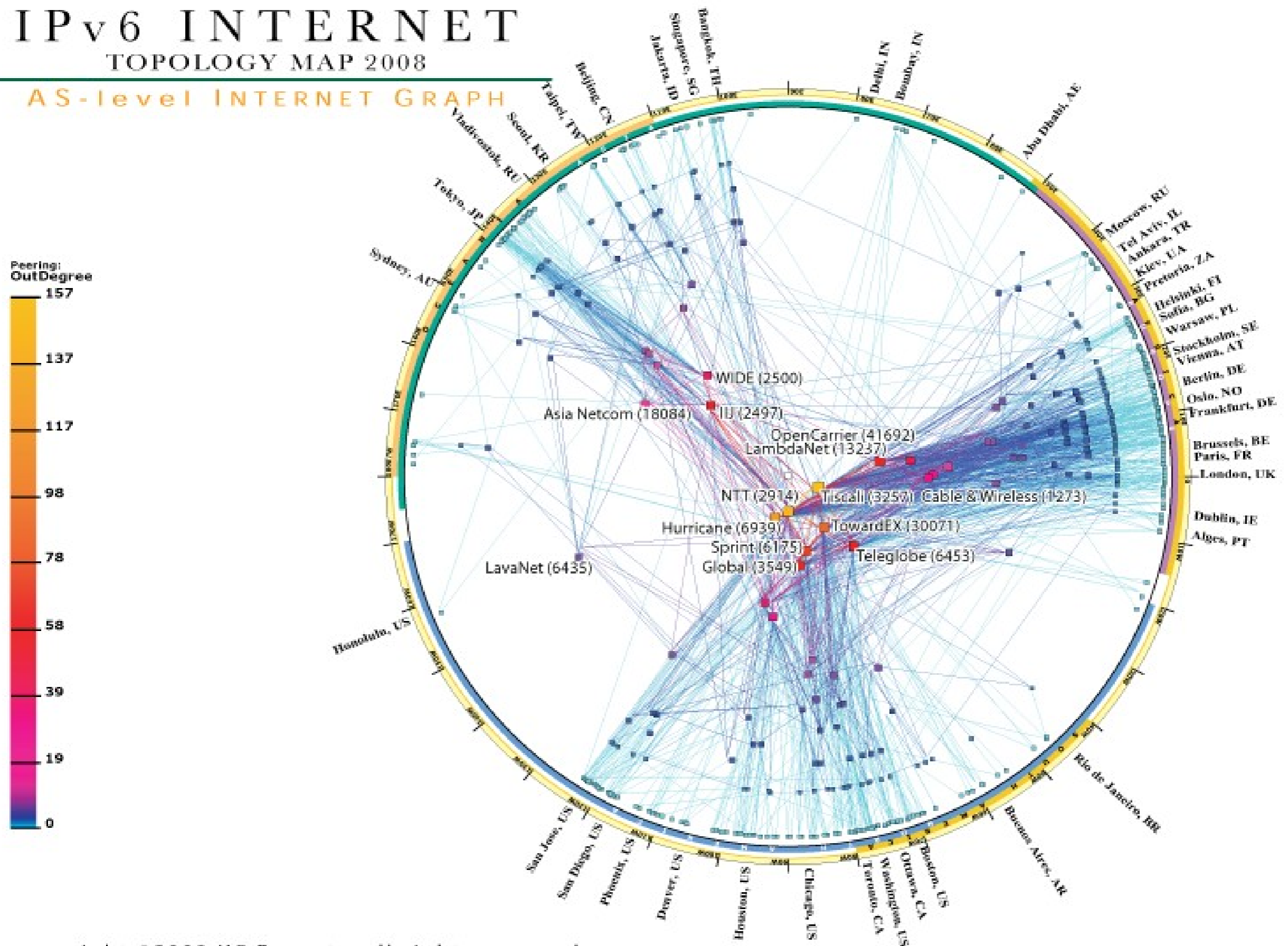2914(NTT)   6453(Teleglobe)
3561(Savvis)
7132(SBC)   4323(Time Warner)

IPv6 INTERNET
TOPOLOGY MAP 2008
AS-level INTERNET GRAPH

visualizing topology data