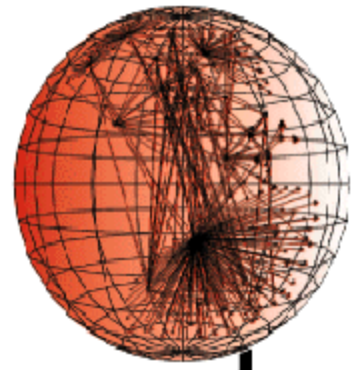


bad data better than no data?

observations on our (in)ability to accurately predict, analyze or even measure conditions on the global Internet

3 oct 2000

the so–called science of poll–taking is not a science at all but a mere necromancy.
people are unpredictable by nature, & though you can take a nation's pulse,
you can't be sure that the nation hasn't just run up a flight of stairs.
––E. B. White New Yorker, Nov 1948.

kc@caida.org
ucsd/sdsc/caida

Internet's resistance to modeling/measurement

evolution-based (good!) reasons

- protocols, technologies, applications
  - independently developed and deployed
  - by no means synergistic
  - by all accounts rapid
  - 'punctuated' but no equilibrium
  - "have done fine without modeling so far"

(let's wait till modeling cheaper than bandwidth)

but simulation/analysis validation (& lately other stuff) needs data

- right granularities hard to come by
- measurement technology just not there
- argument for it also not there
- "helps everyone", but who pays?

Internet's resistance to measurement

measurement tools lack

- well-defined traffic metrics
  - e.g supporting SLAs, QOS, billing
- uniformly applied methodologies
  - varied topologies, equipment, ISP practices
- scalability
- ability to explain phenomena
  - topology changes, routing loops, black holes
- relevance to actual ISP problems or mechanisms for fixing
- communication of useful results

Internet measurement taxonomy

- topology (circulatory/respiratory)
- performance (physiology/psychology)
- workload (cardiovascular/GI)
- routing (neuroscience)

correlation essentially non-explored
.....(holistic Internet measurement?)

topology: caida's skitter

- track/depict topology cross-sections
  - 22 monitors (inc. some root name servers)
  - forward IP path and round-trip delay
  - tens of thousands of dst (multiple lists)
  - remove targets that complain
- architecture
  - continuous, parallel 52-byte ICMP probes
  - depending on dst list size, 0.3 to 200 probes/day/dst
  - kernel time stamping
- correlate path perf. w events, e.g. BGP
- identify critical pieces of infrastructure

other active (probed) data sets

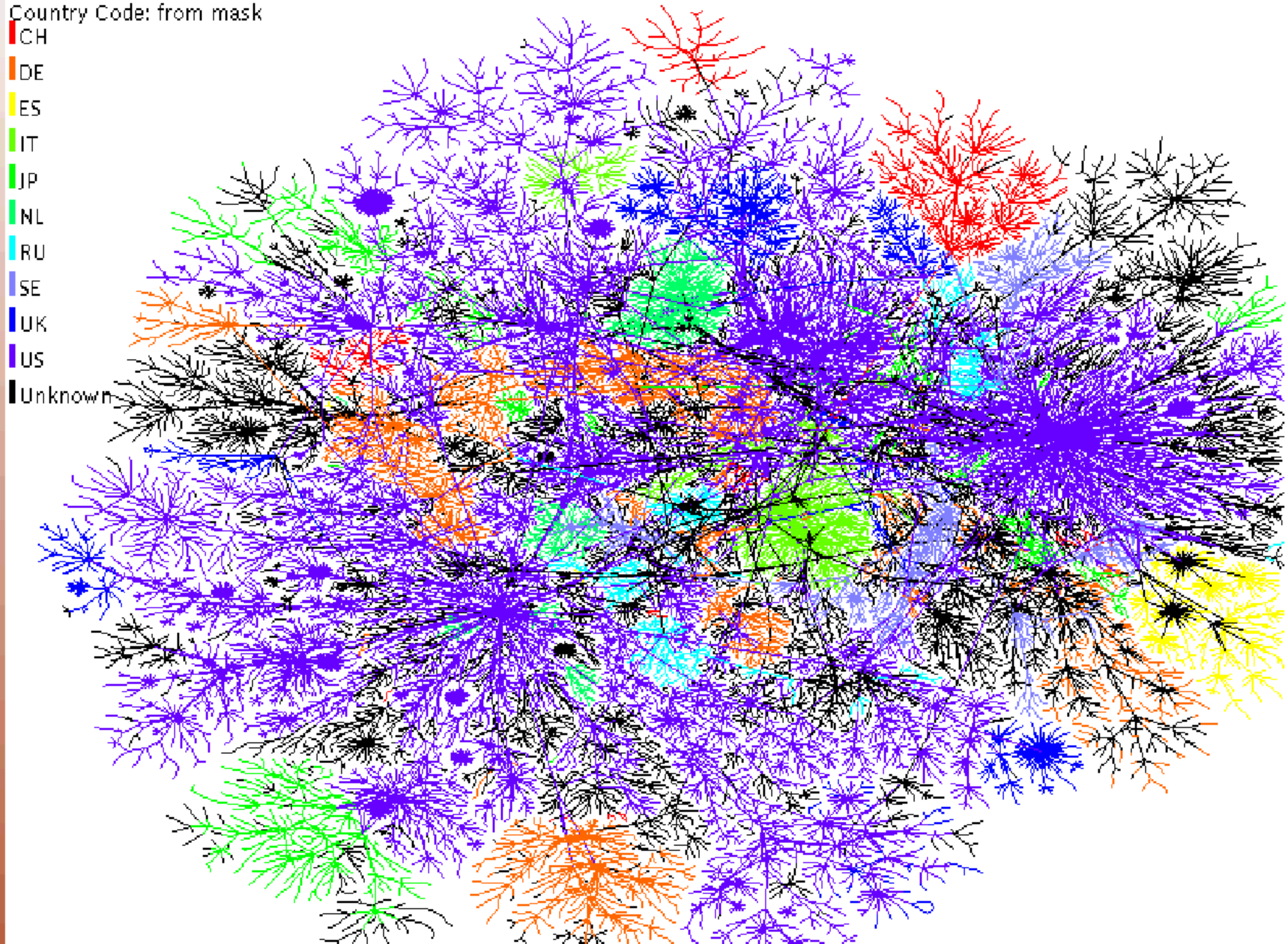
- MOAT: http://amp.nlanr.net
  - HPC sites
  - RTT, traceroutes

- I2's Surveyor:
  http://www.advanced.org/surveyor/
  - I2 sites
  - one-way delay, paths

- vBNS
  http://www.vbns.net:8080/stats/

- SLAC, NIMI, XIWT, RIPE, others (too many)

skitter: colored by countries

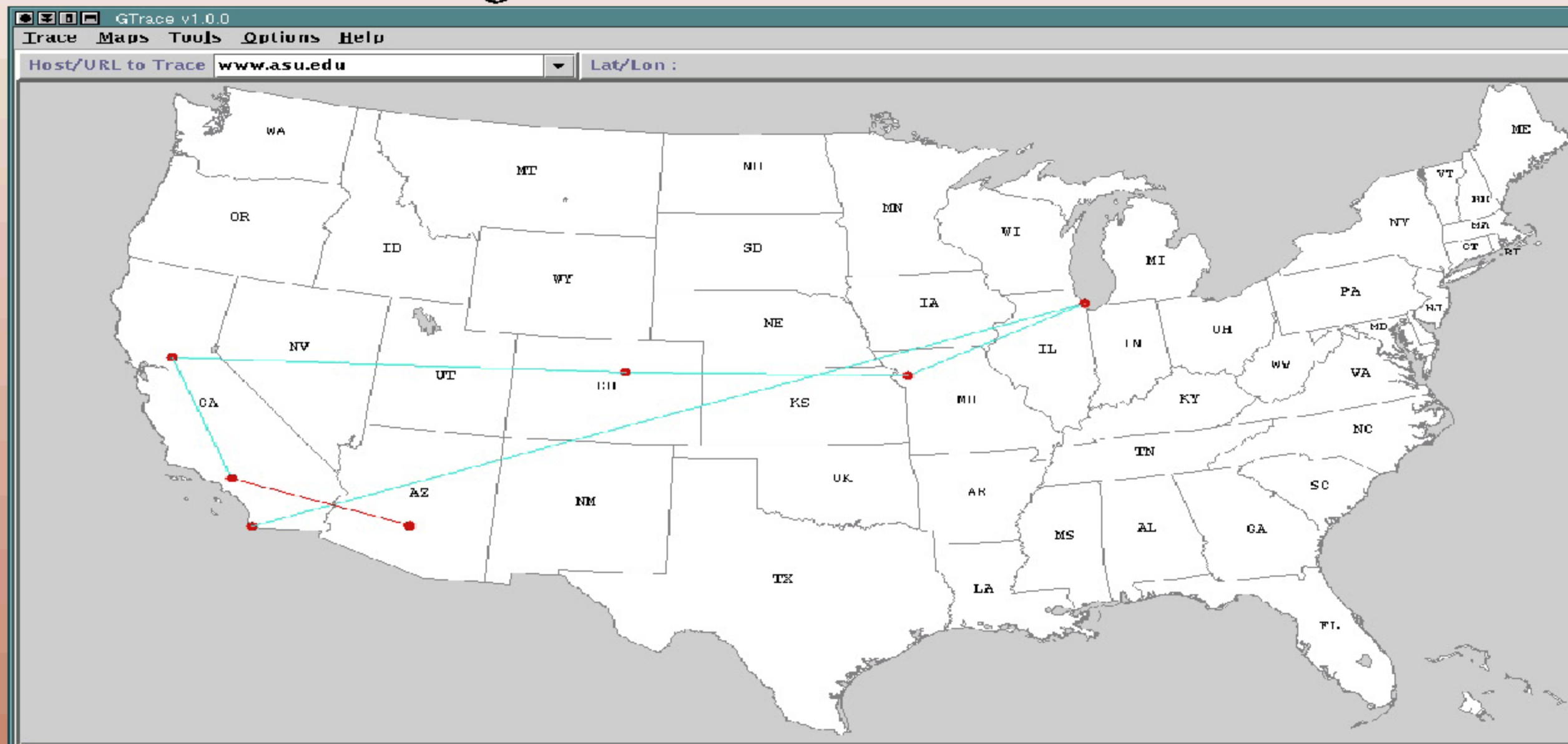


topology vis: geographic mapping

- difficult data analysis
  - requires mapping of thousands (millions?) of nodes to latitude/longitude coordinates
- NetGeo service designed to help
  - http://netgeo.caida.org
- backbones require company–specific heuristics
- DNS registry growth is problematic
  - no common data formats

GTrace: geographic traceroute

www.caida.org/Tools/GTrace/



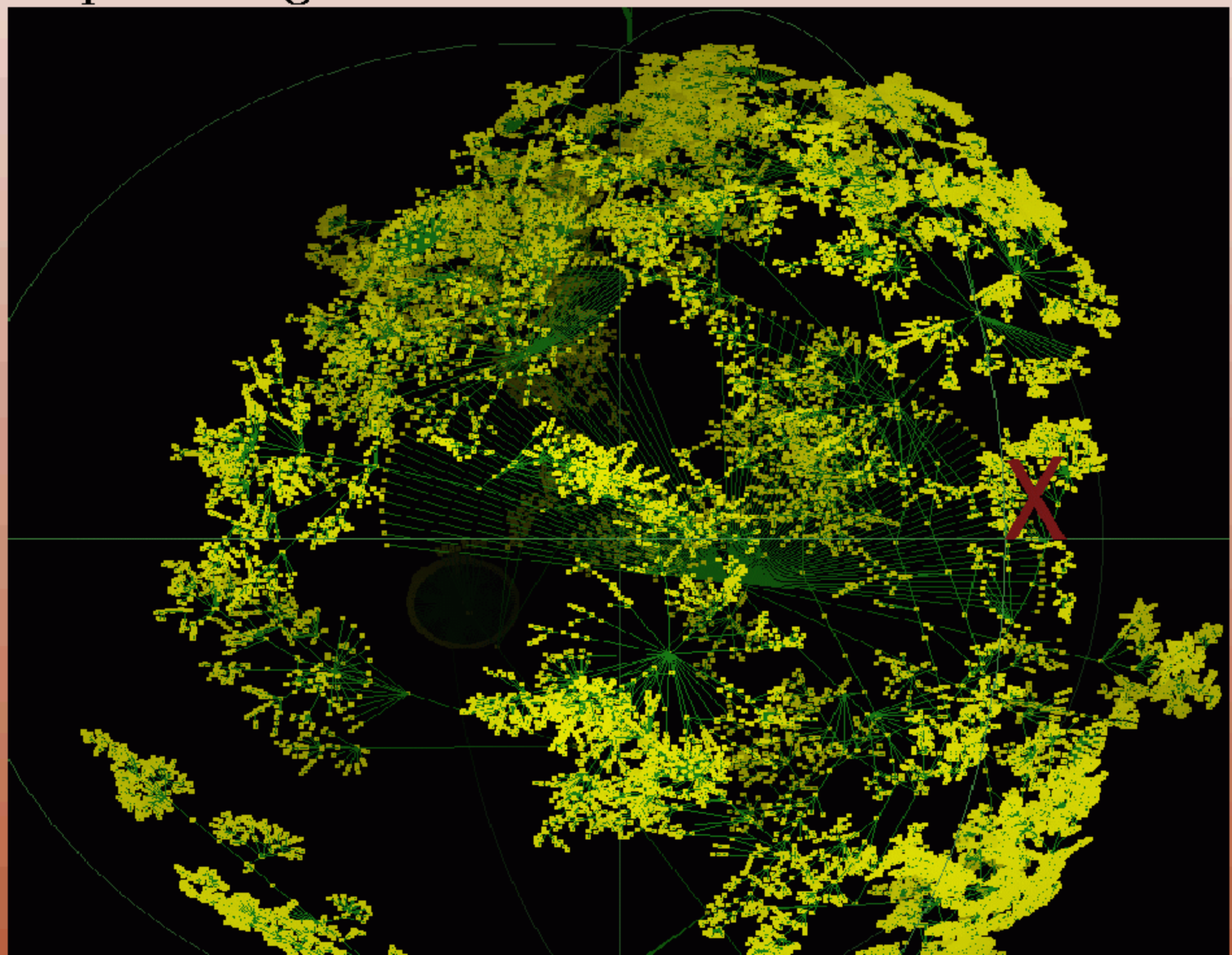
Plot	Hop	IP Address	Interface Name	Location	RTT (ms)
☑	1	(192.172.226.65)	pinot-fe2-0-0	San Diego,California,US	0.777
☑	2	(198.17.46.43)	cs-f-vbns.sdsc.edu	San Diego, CA	0.74
☑	3	(204.147.135.154)	gsr-atm0-0-2.dng.vbns.net	Chicago,IL,US	51.875
☐	4	(204.147.128.14)	204.147.128.14	Reston,Virginia,US	56.175
☑	5	(198.32.8.5)	kscy-ipls.abilene.ucaid.edu	39.6 N, 94.57 W	56.132
☑	6	(198.32.8.13)	denv-kscy.abilene.ucaid.edu	Denver,CO,US	56.367
☑	7	(198.32.8.1)	scrm-denv.abilene.ucaid.edu	38.53 N, 121.46 W	56.268
☑	8	(198.32.8.18)	losa-scrm.abilene.ucaid.edu	Los Angeles,CA,US	56.459
☑	9	(129.219.250.13)	129.219.250.13	Tempe,Arizona,US	60.4
☑	10	(149.169.0.13)	asu1-main.inre.asu.edu	Tempe, AZ	135.991
☑	11	(129.219.152.4)	eca1_fddi.inre.asu.edu	Tempe, AZ	101.21
☑	12	(129.219.13.117)	wsd1.inre.asu.edu	Tempe, AZ	150.031
☑	13	(129.219.117.196)	web2.asu.edu	Tempe, AZ	147.477

topology mapping: interface merging

- 26 sept 2000, 18 hours
- 360k interfaces, 505k probes
- responses
  - joins: 29893
  - new i/fs: 2692
  - nodes with >1 i/fs: 18556
  - i/faces on nodes: 48005
  - single interfaces: 311663

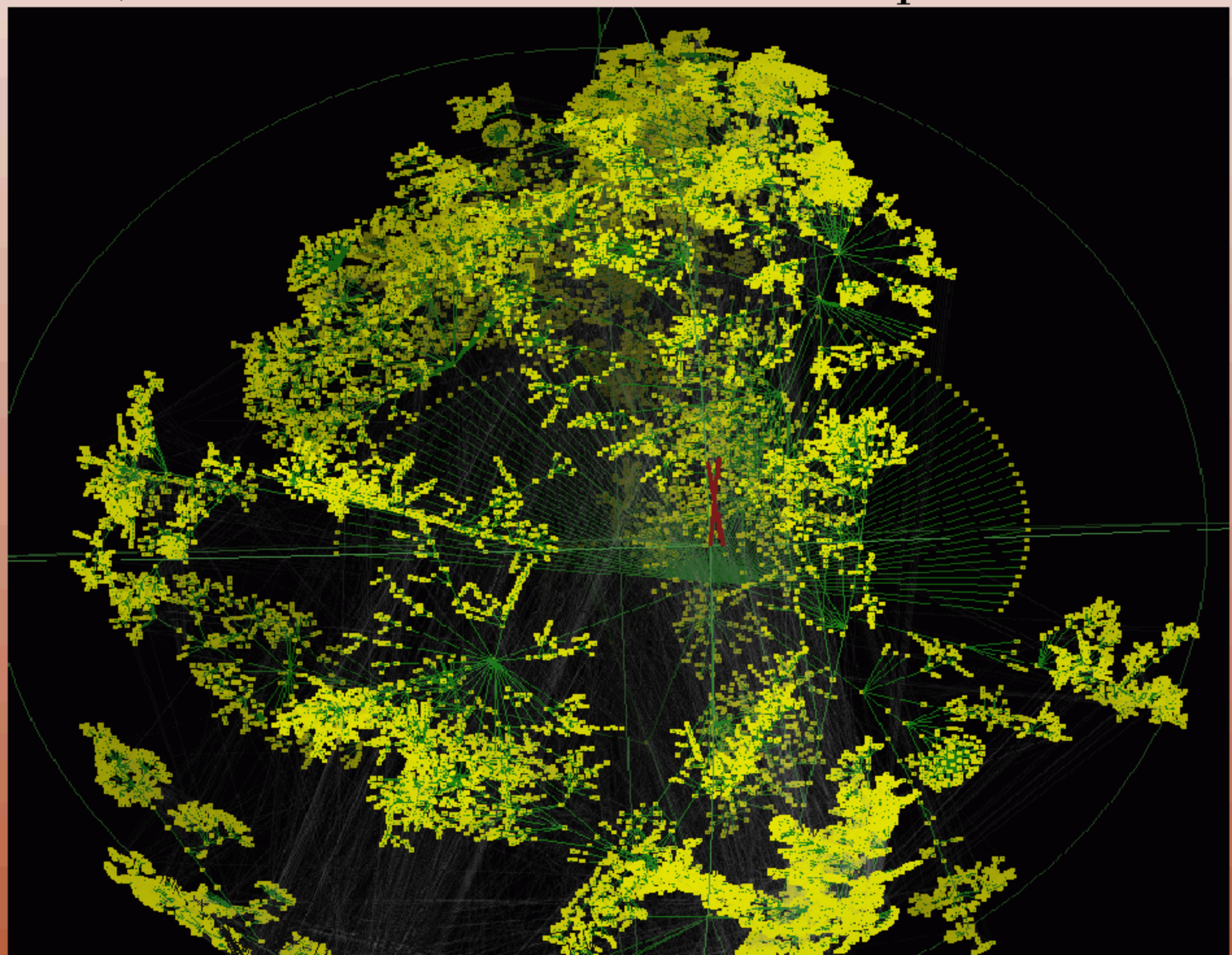
hyperbolic viewer (java 3D, 100,000s nodes)

from riesling skitter monitor in san diego
54,893 nodes, 54,892 tree links
spanning tree from src



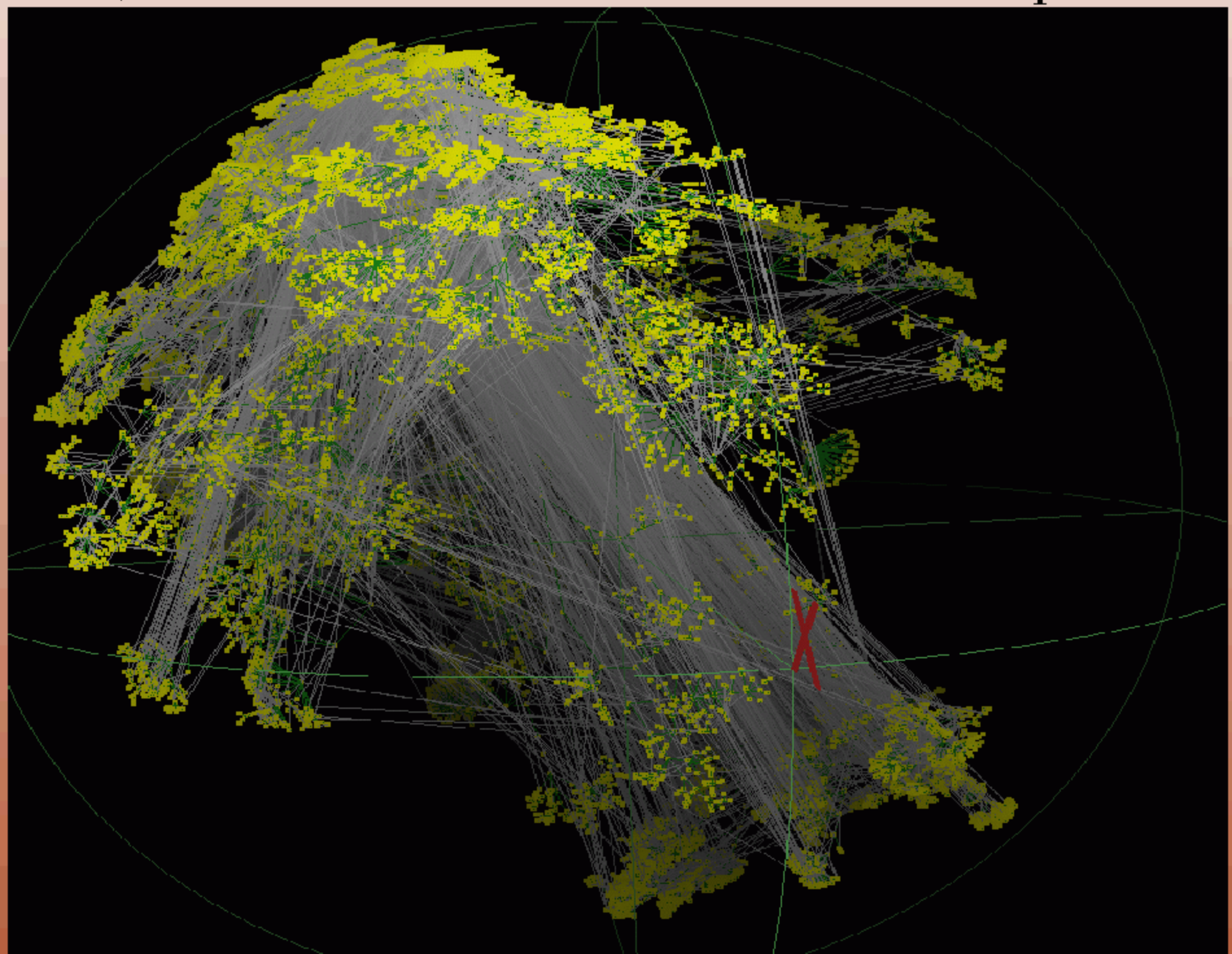
hyperbolic viewer (java 3D, 100,000s nodes)

from riesling skitter monitor in san diego
54,893 nodes, 54,892 tree links
24,517 non-tree links (transparent)



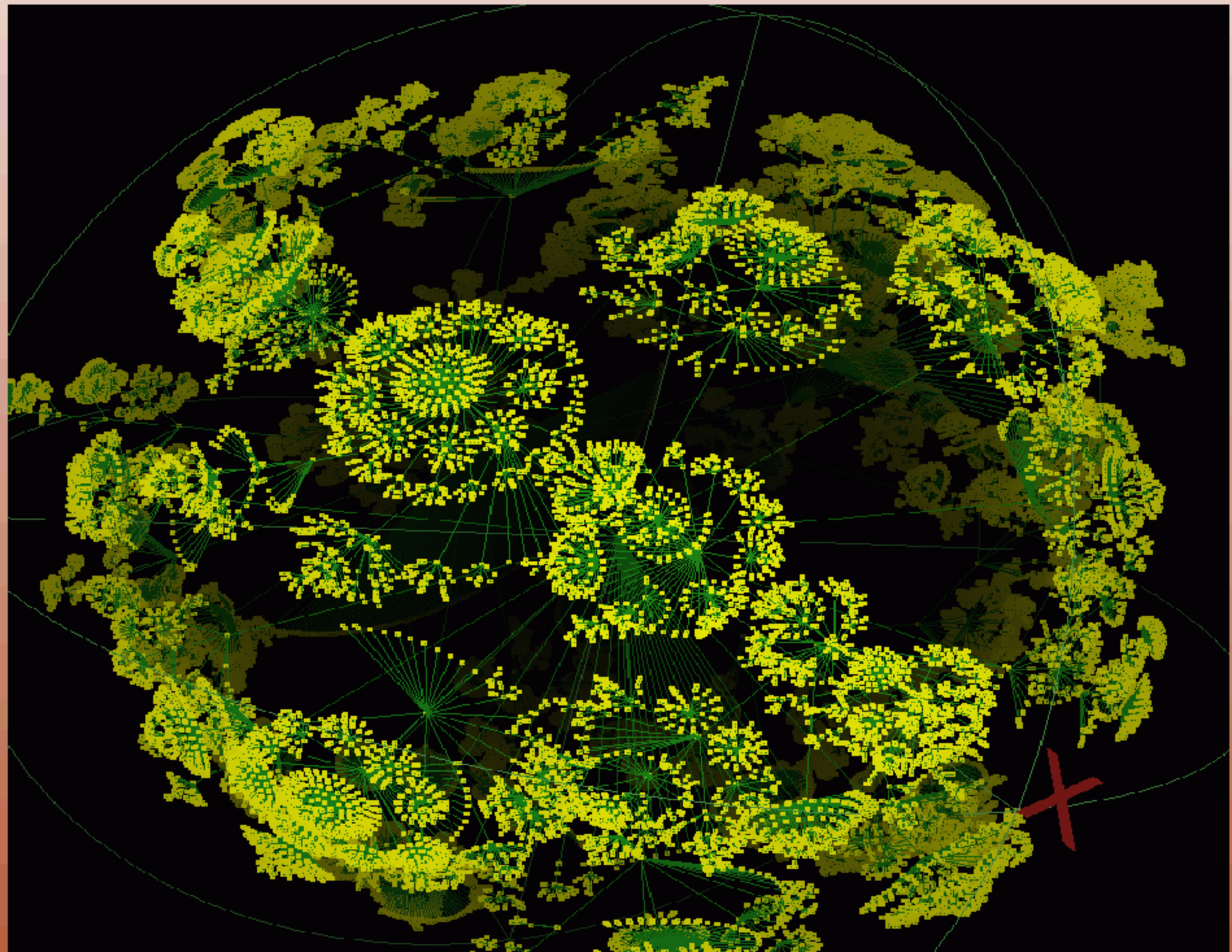
hyperbolic viewer (java 3D, 100,000s nodes)

from riesling skitter monitor in san diego
54,893 nodes 54,892 tree links
24,517 non-tree links (less transparent)



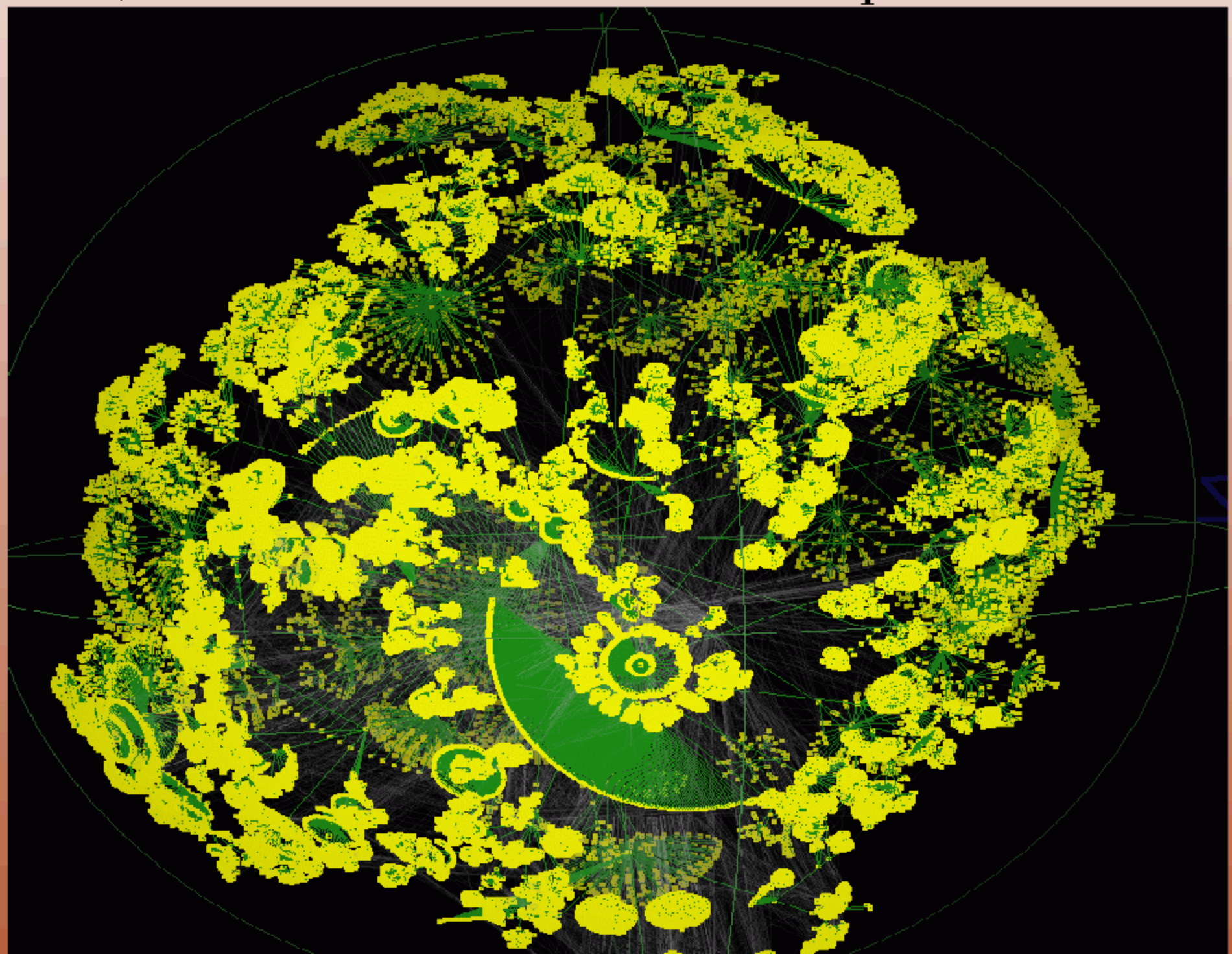
hyperbolic viewer (java 3D, 100,000s nodes)

from london skitter monitor in
535,102 nodes, 535,101 tree links
66,577 nontree links



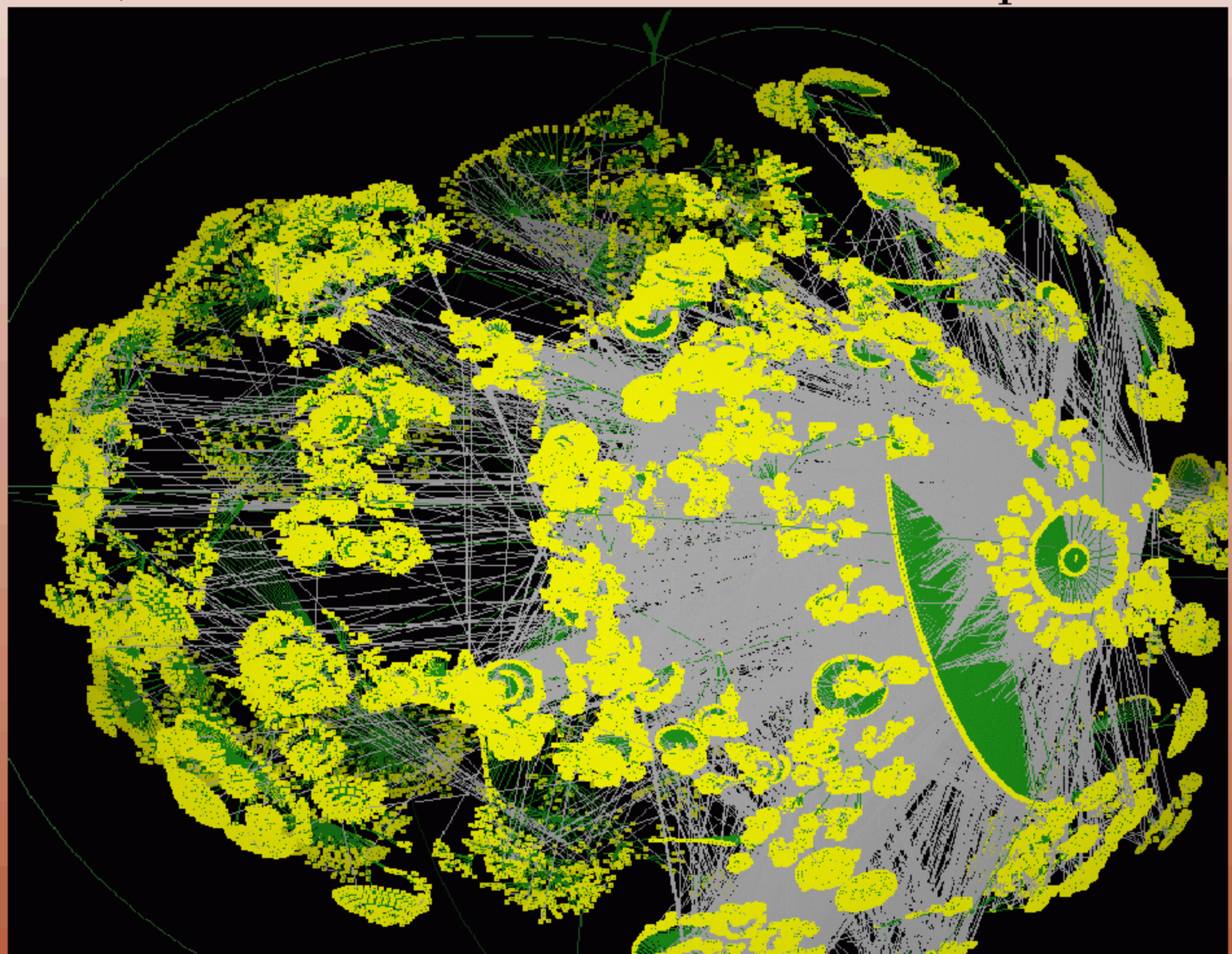
hyperbolic viewer (java 3D, 100,000s nodes)

from london skitter monitor in
535,102 nodes, 535,101 tree links
66,577 nontree links (transparent)



hyperbolic viewer (java 3D, 100,000s nodes)

from london skitter monitor in
535,102 nodes, 535,101 tree links
66,577 nontree links (less transparent)



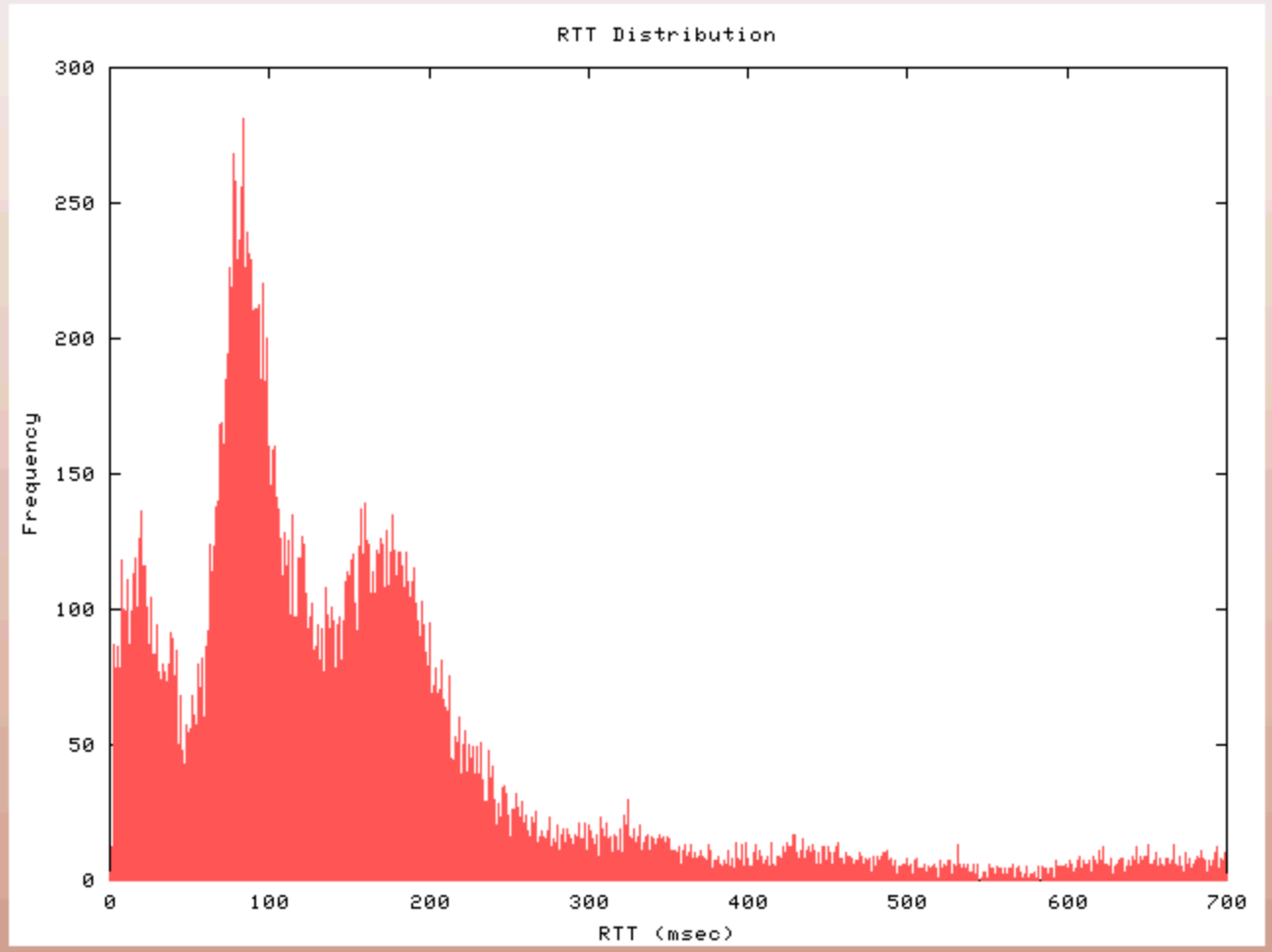
skitter case study: DNS roots

- RSSAC, DNS technical advisory committee to ICANN
- goal: optimize root nameserver location
  - co–locate skitter hosts w root servers
  - demonstrate root server performance in serving target community
  - develop techniques for evaluating architectual optimality for root server placement
  - visualization to correlate data sources/types
- collaborative project to encourage proactive participation (network operators, researchers, others)

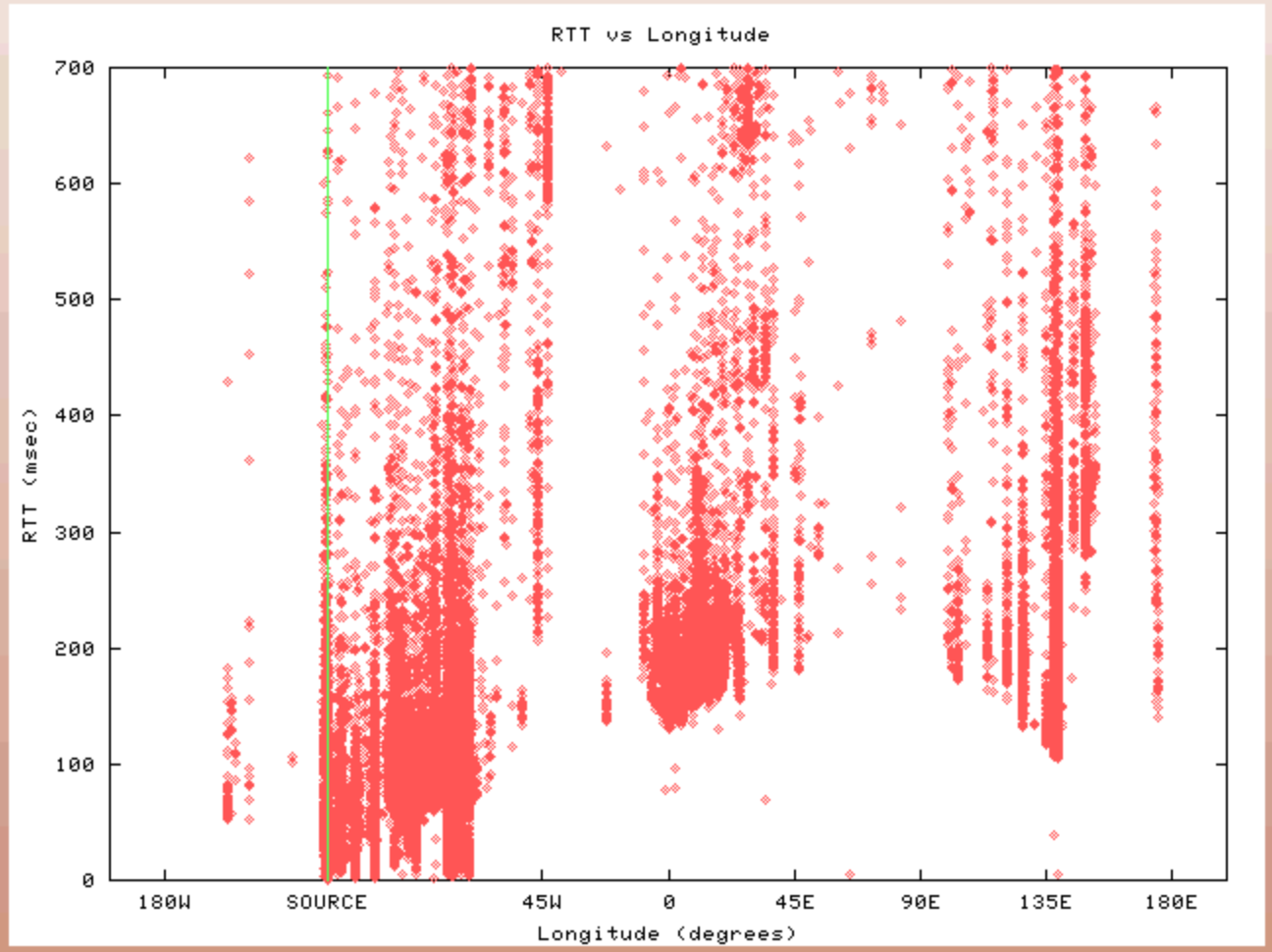
skitter case study: DNS roots

- get roots instrumented
- gather/analyze client lists
- correlation among different sources
- determine of connectivity metrics
  - closeness
  - redundancy
  - persistence of paths
- how many clients not secondaries
- skitter to client sets from non–root sources

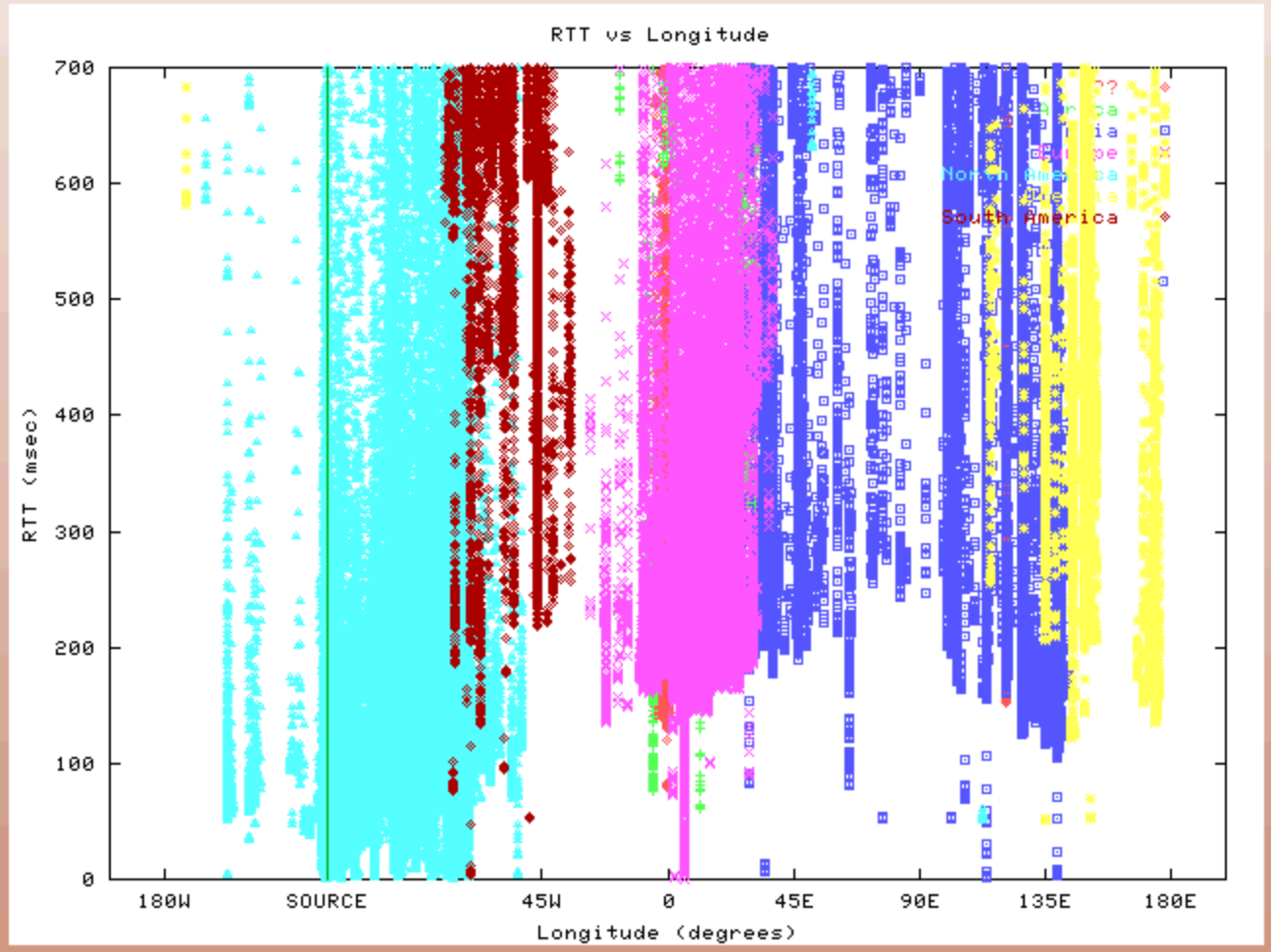
skitter: rtt distribution: tri-modal



skitter: rtt vs longitude (light cone)



skitter: rtt vs longitude (light cone)



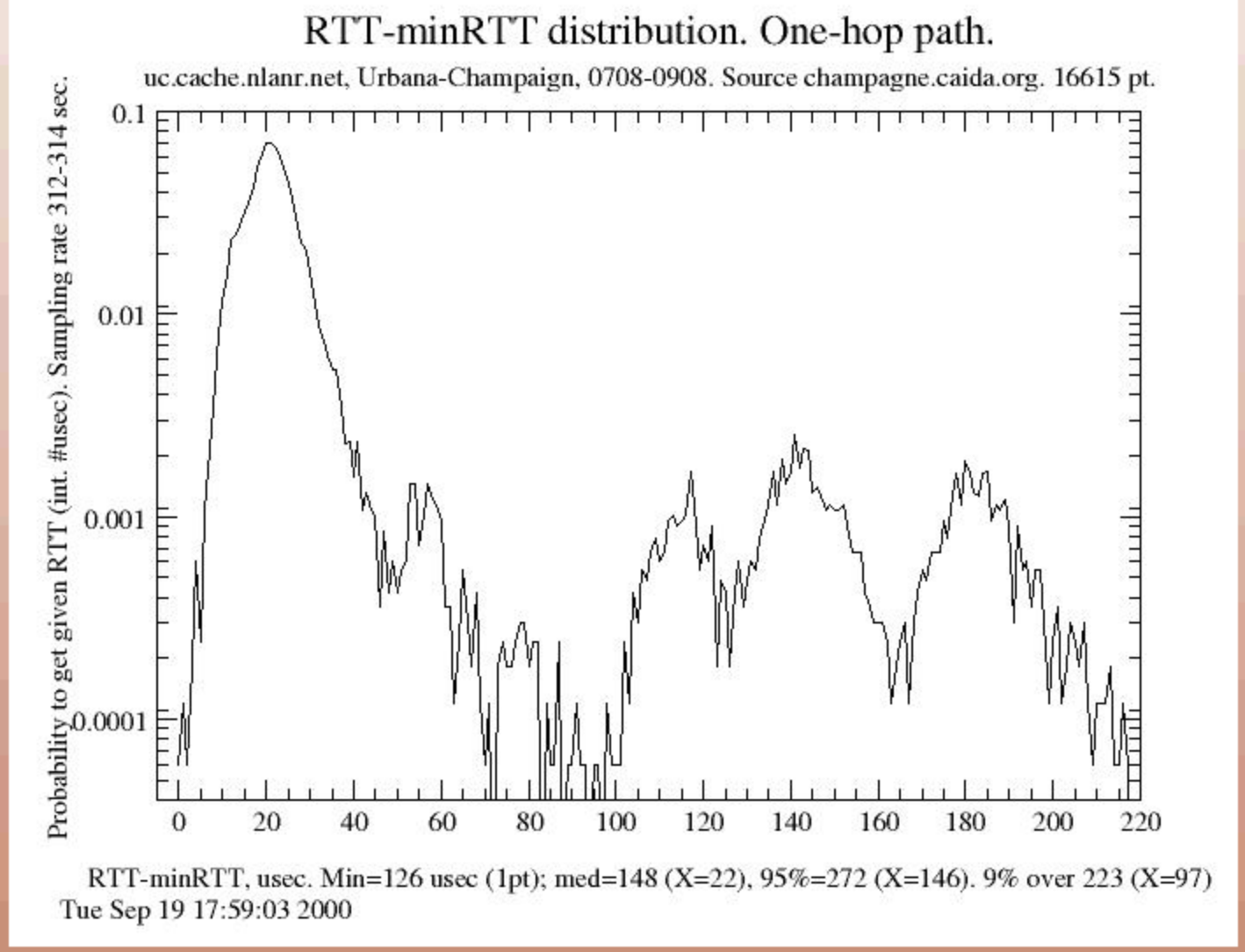
skitter: modeling rtt along a single path

champagne.caida.org (UIUC) monitor

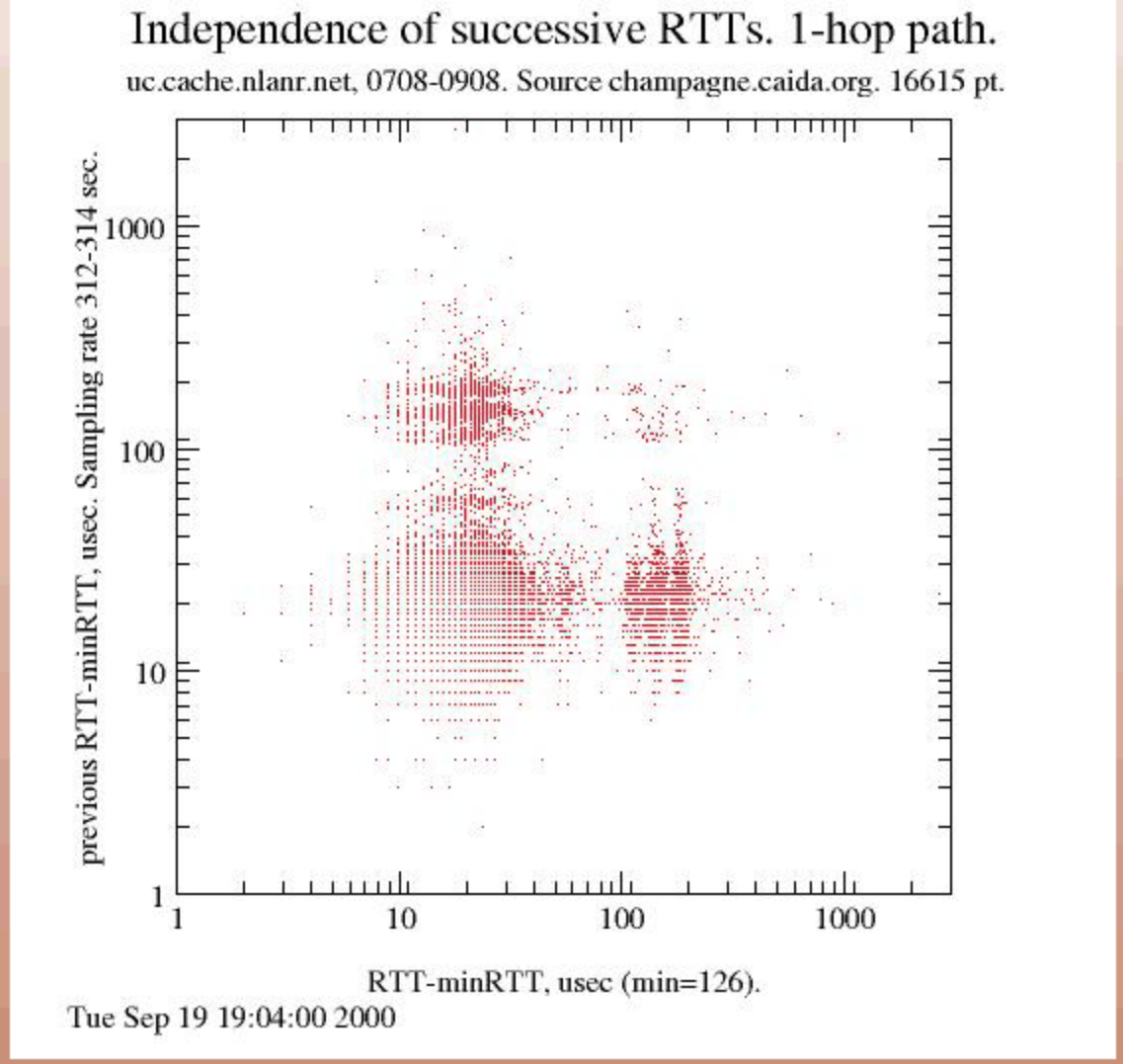
- fewest destinations,
- most probes/dst per day, 1/5min/dst (270)
- 63 days 08 July – 08 Sept 2000
- exactly 9 weeks, sun–sat
- only use stable subsets of paths
- 75% responsive, 33% of paths same per day
- min 5000 paths same

--> gives us 168 pairs to work with

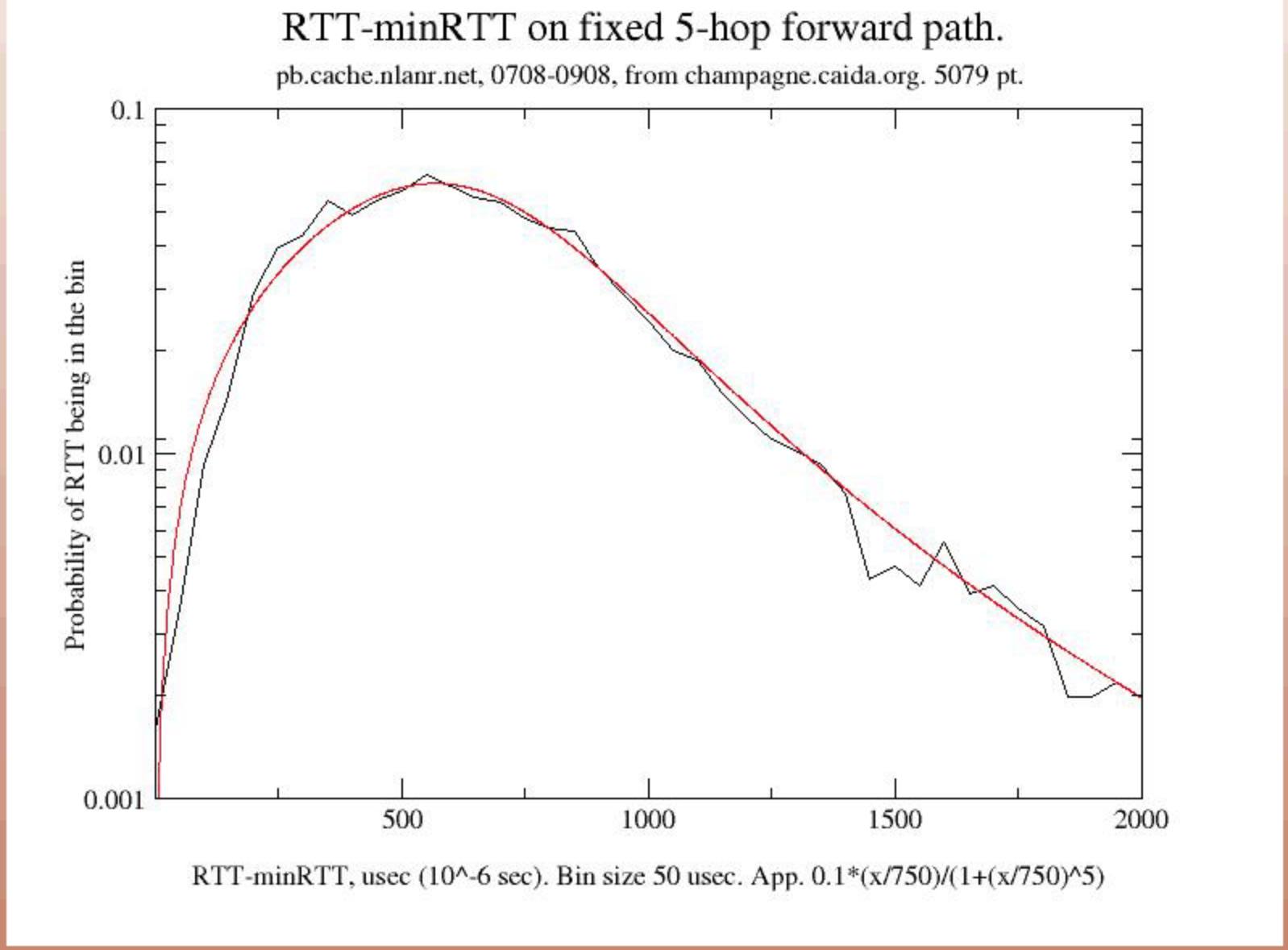
skitter: modeling rtt along a single path



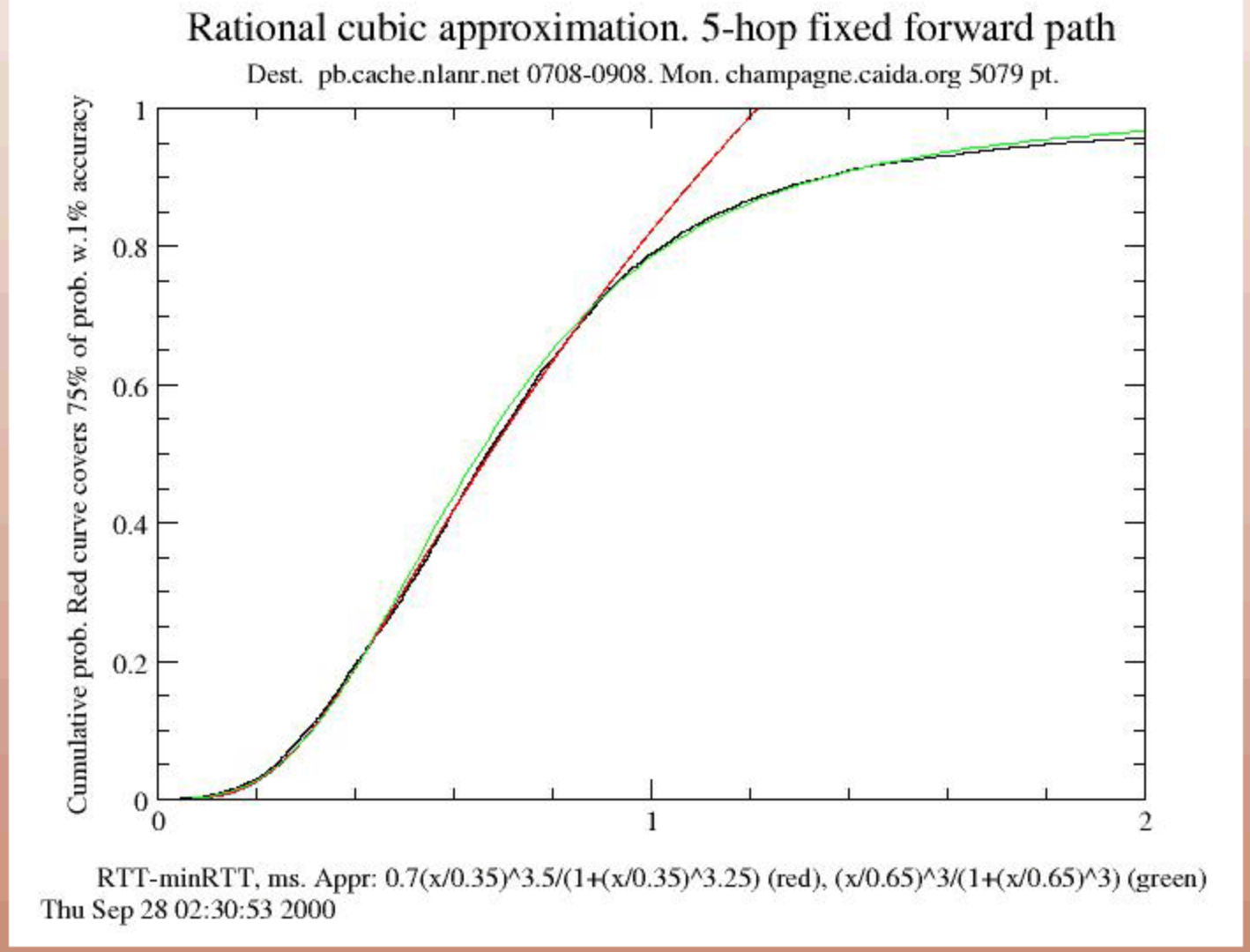
skitter: modeling rtt along a one-hop path



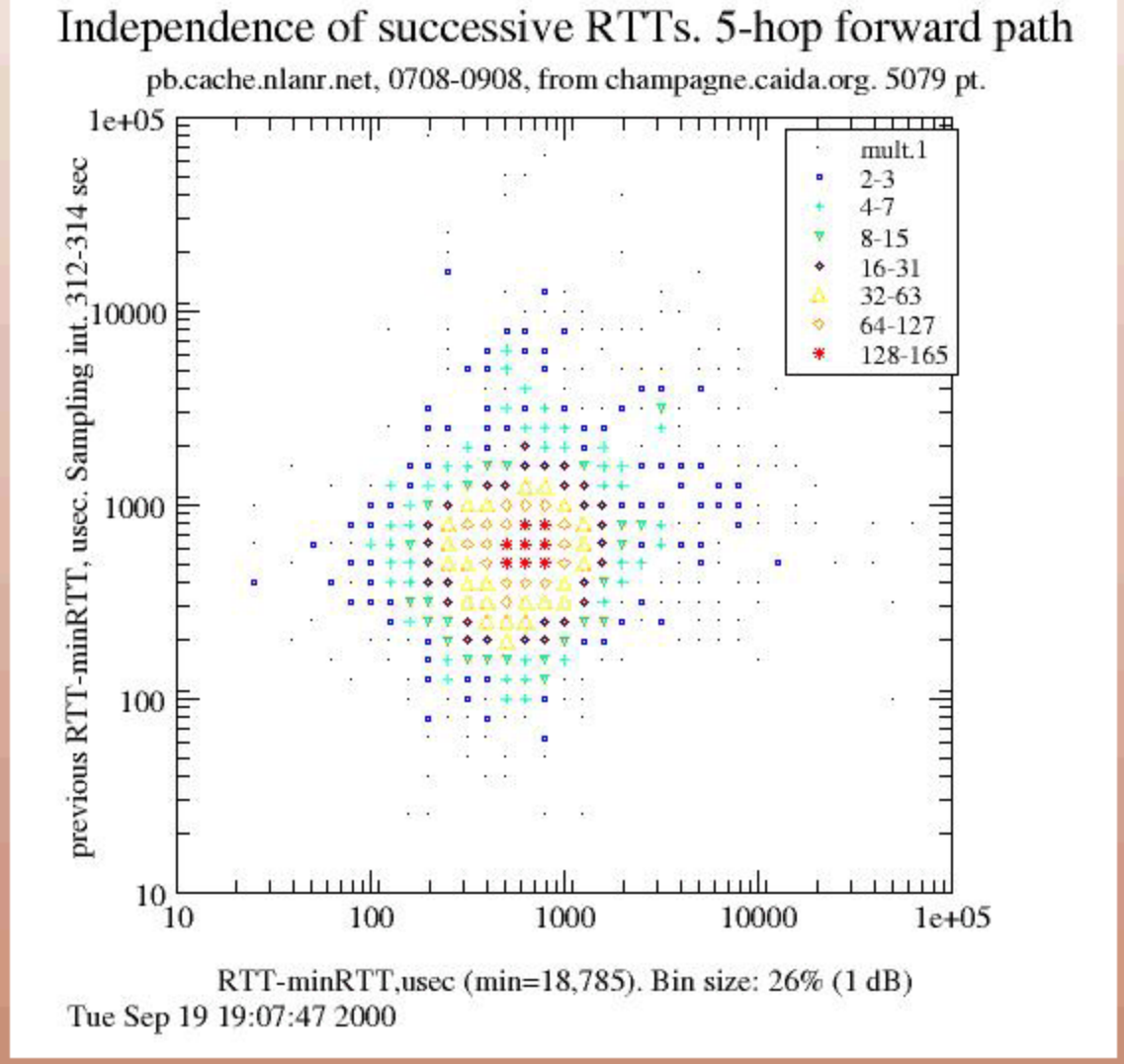
skitter: modeling rtt along a five-hop path



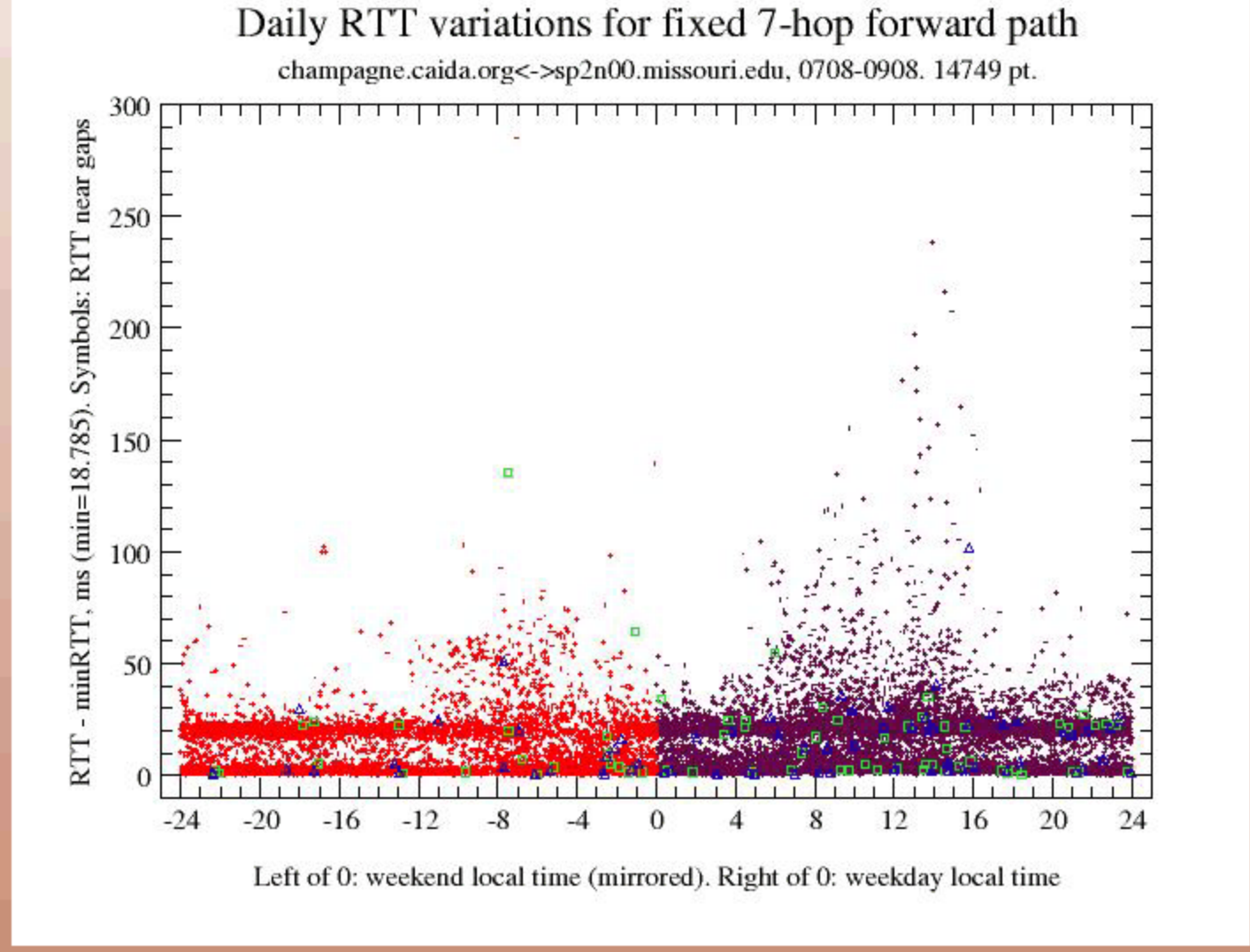
skitter: modeling rtt along a five-hop path



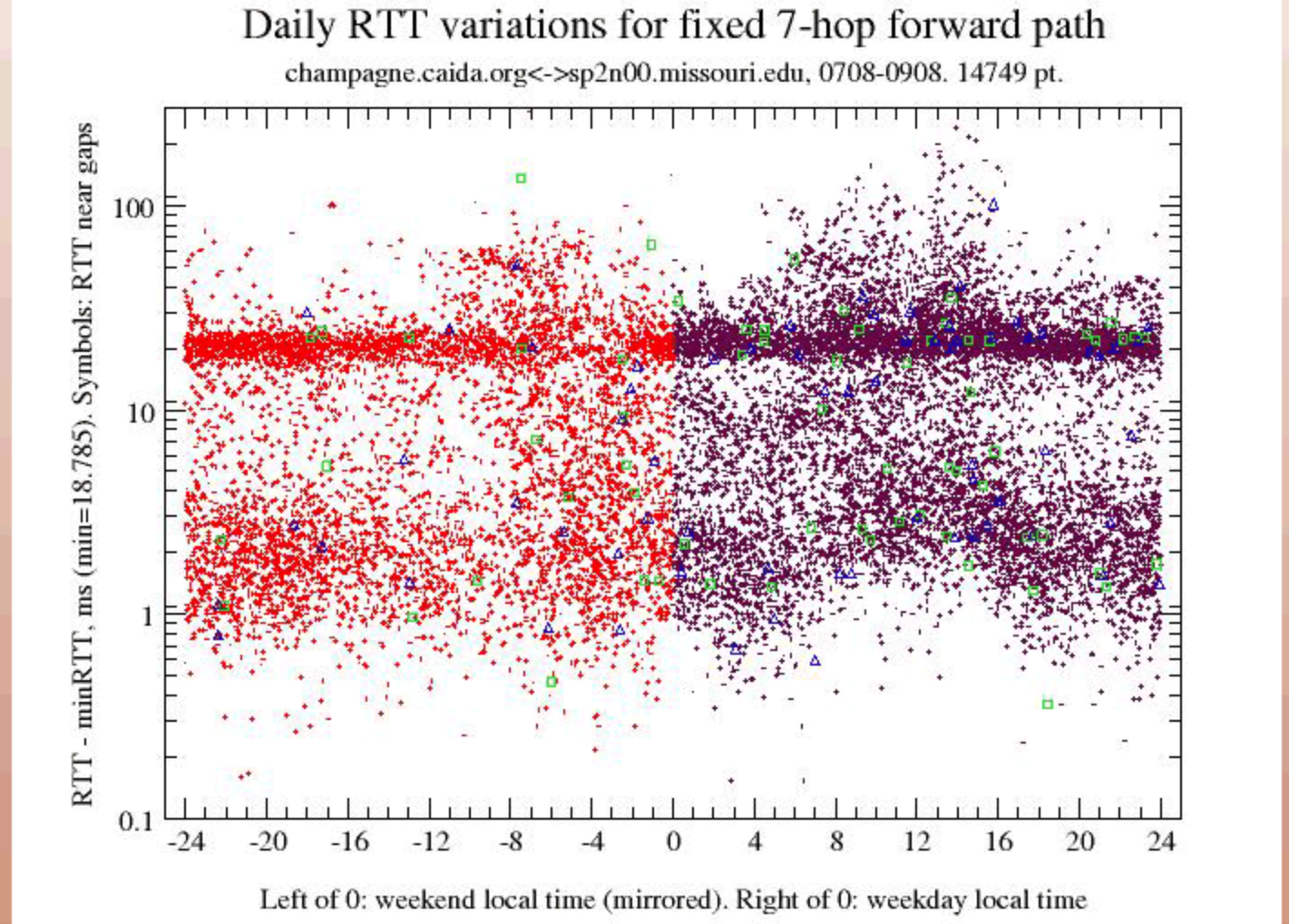
skitter: modeling rtt along a five-hop path



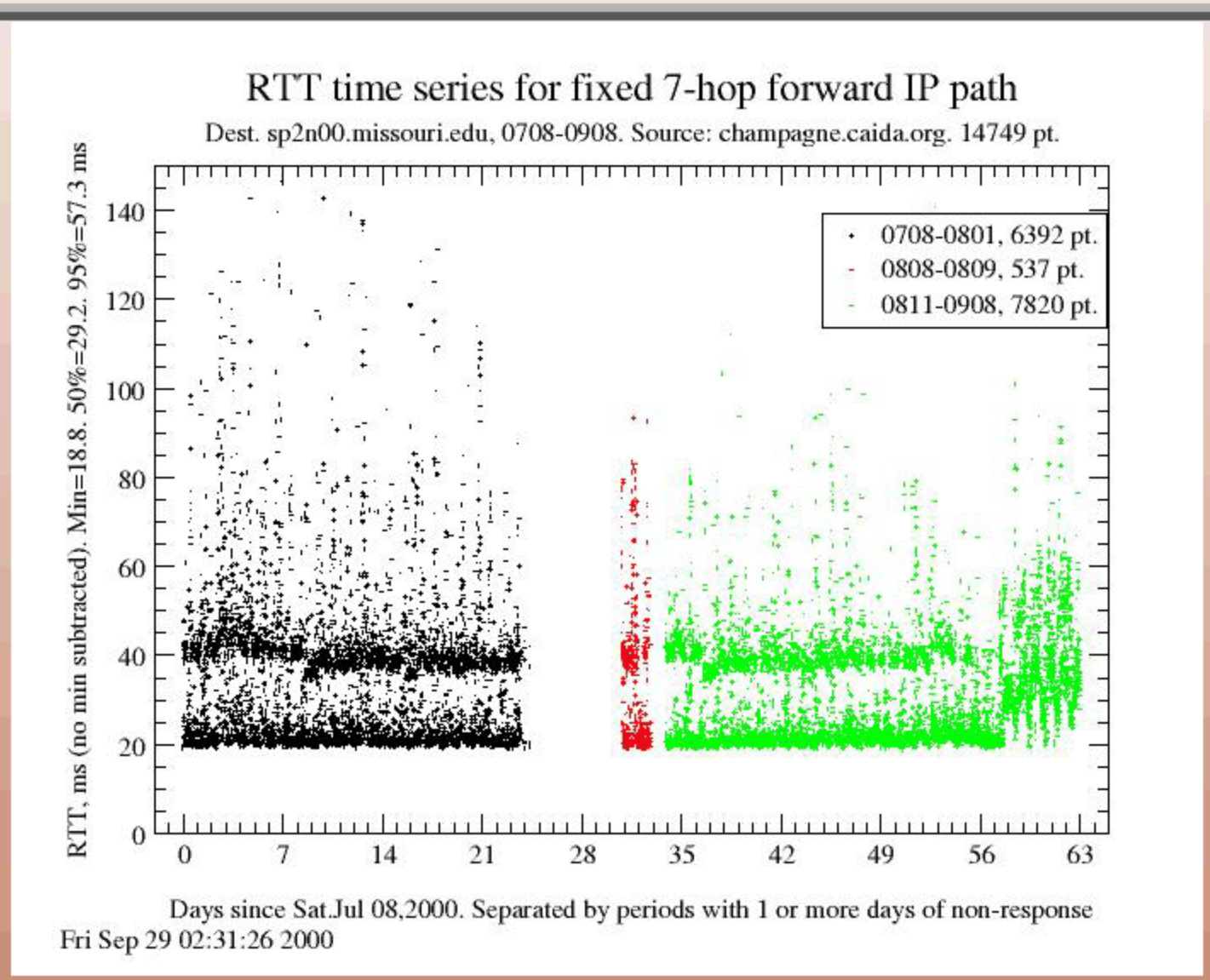
skitter: bimodal rtt along a 7-hop path



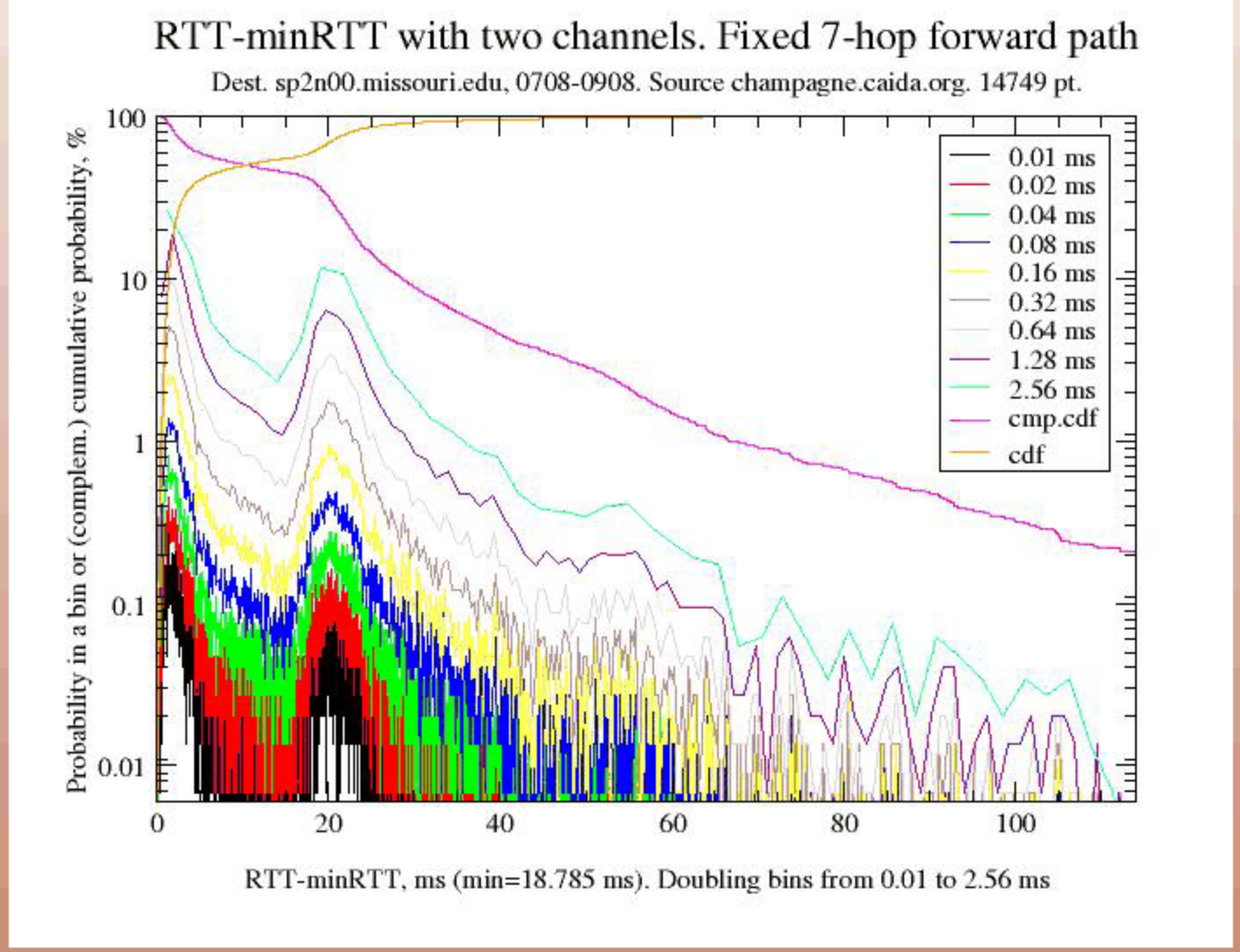
skitter: bimodal rtt along a 7-hop path



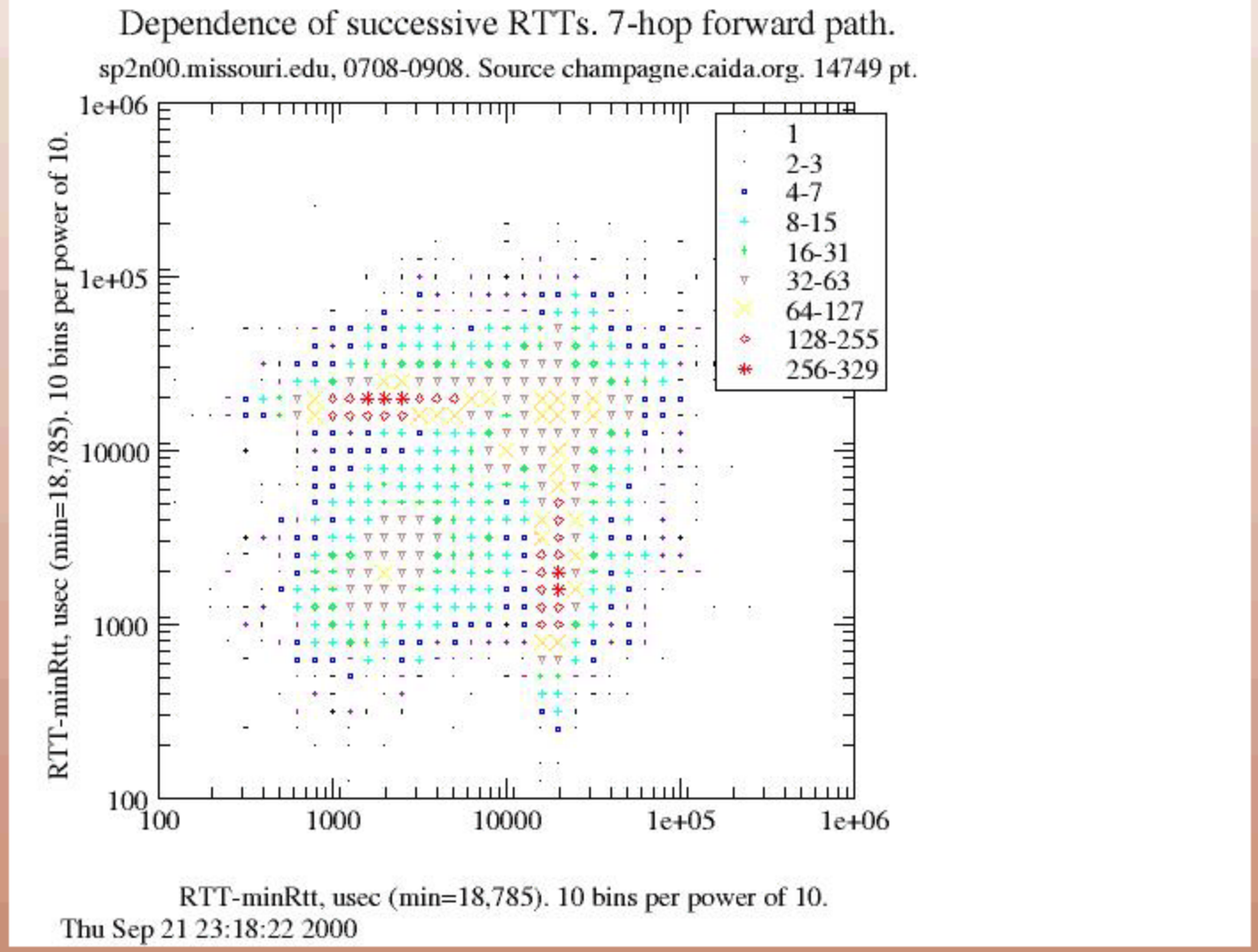
skitter: bimodal rtt along a 7-hop path



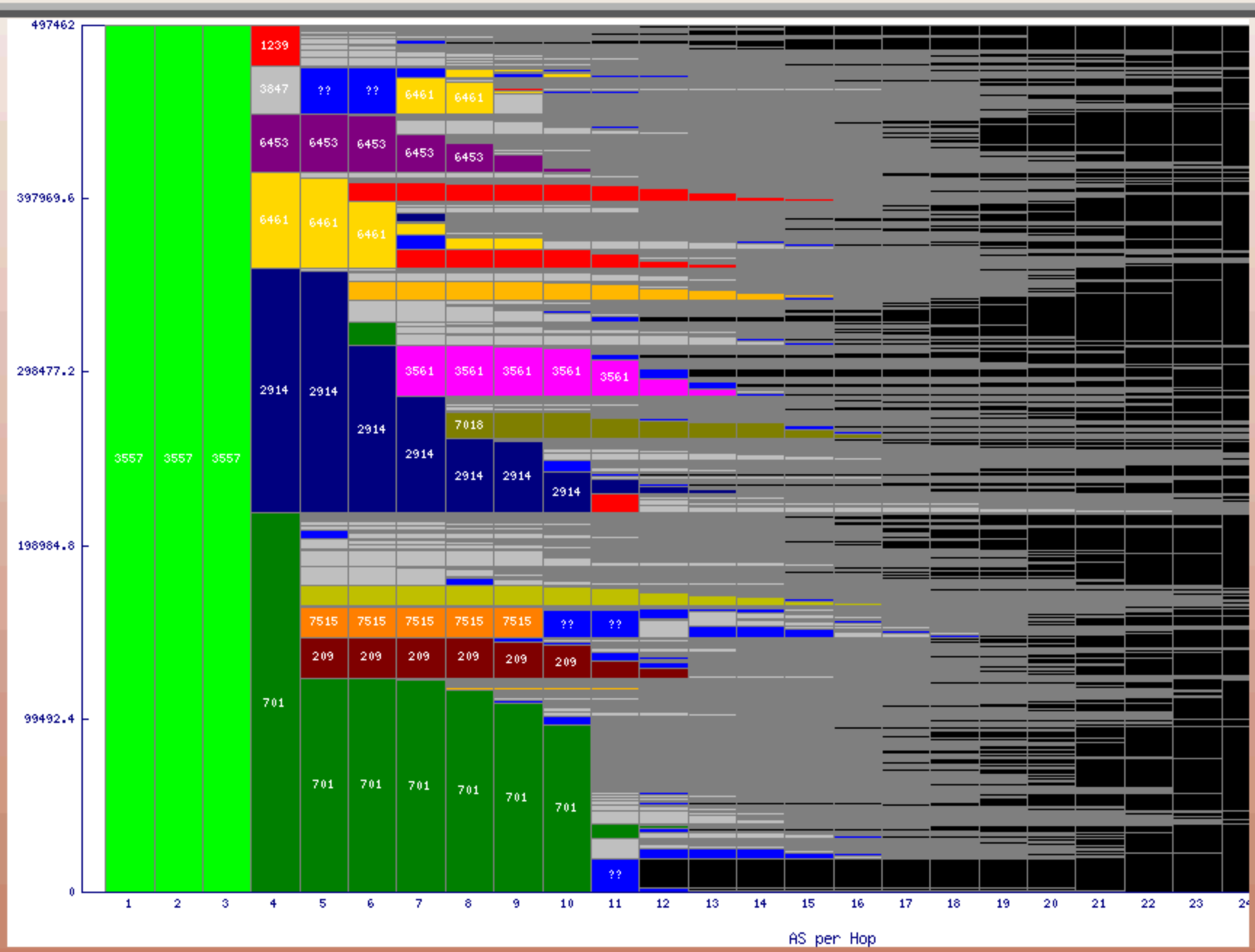
skitter: bimodal rtt along 7-hop path



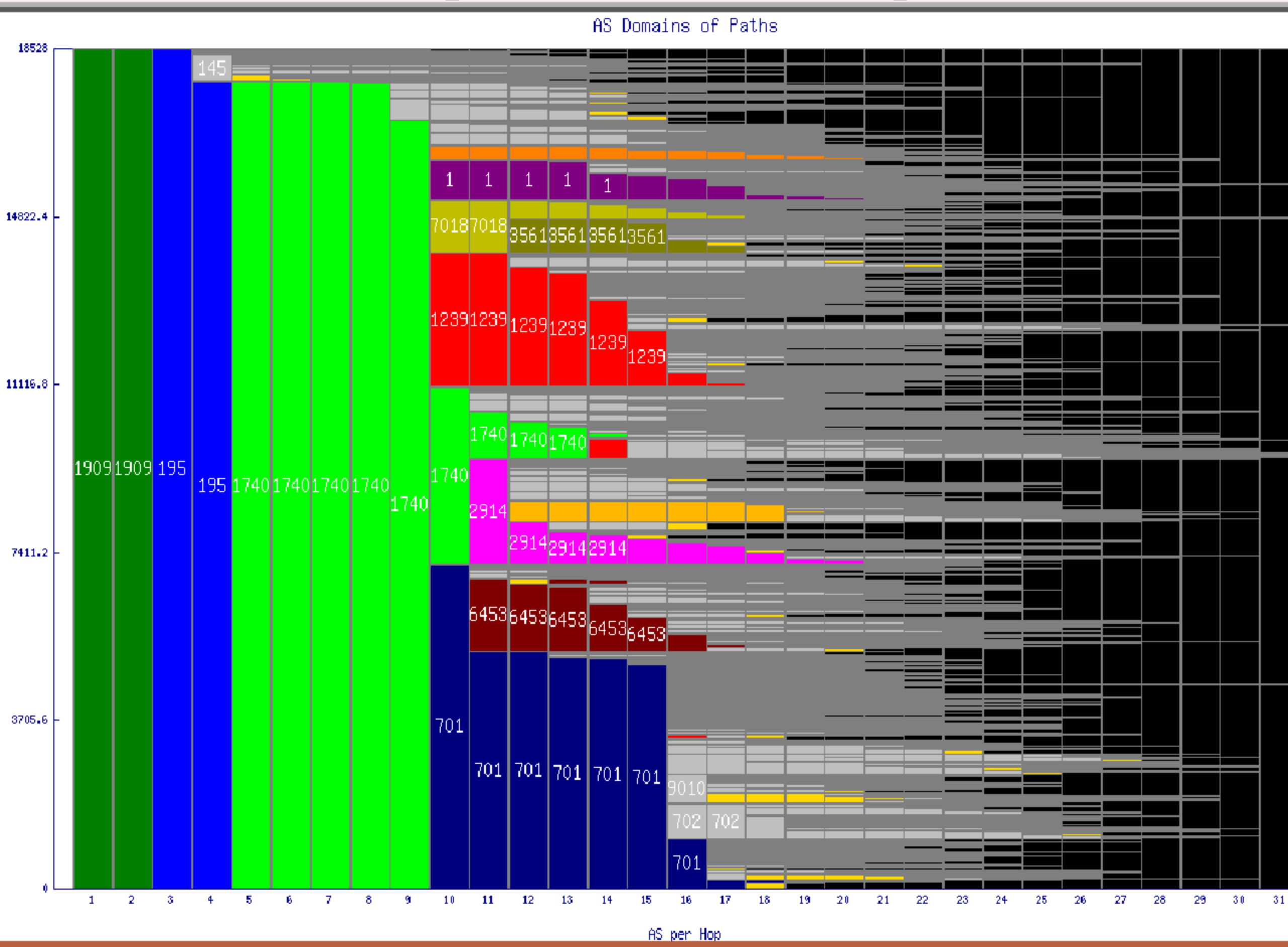
skitter: non-independence of rtt for 7-hop path



skitter: dispersion among ASes across paths



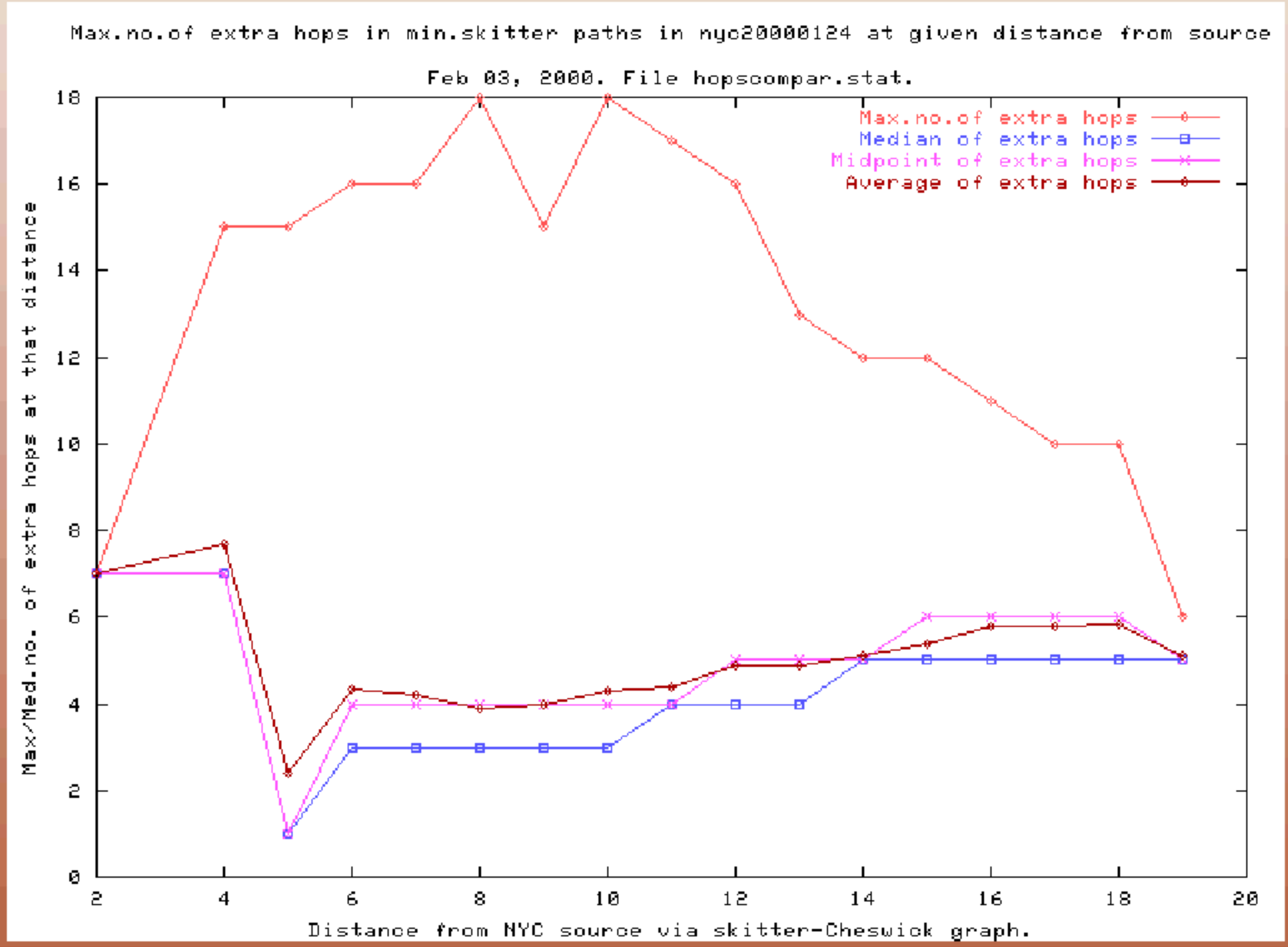
skitter: AS dispersion across paths (sdsc)



skitter: other interesting studies

non-shortestPath-ness of topology

- idealized topology (skit+ches graph)
- median/max # extra hops taken vs optimal
- worst case: actual path 18 hops longer than shortest path



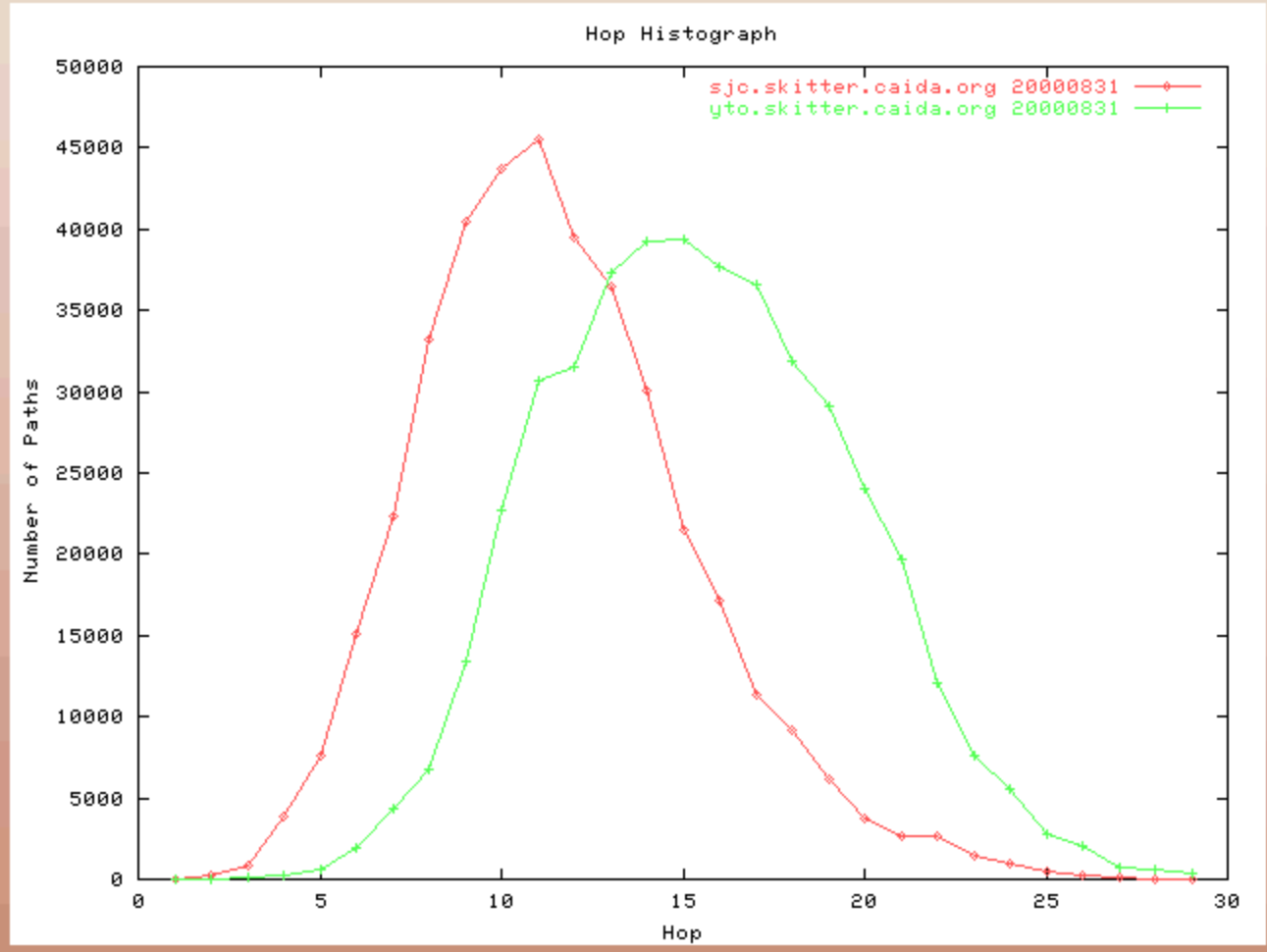
skitter on-going daily summaries

http://www.caida.org/tools/measurement/skitter/summa

- path length (in IP hops) distribution
- RTT distribution
- RTT versus longitude,
- path dispersion
  - AS & country granularity

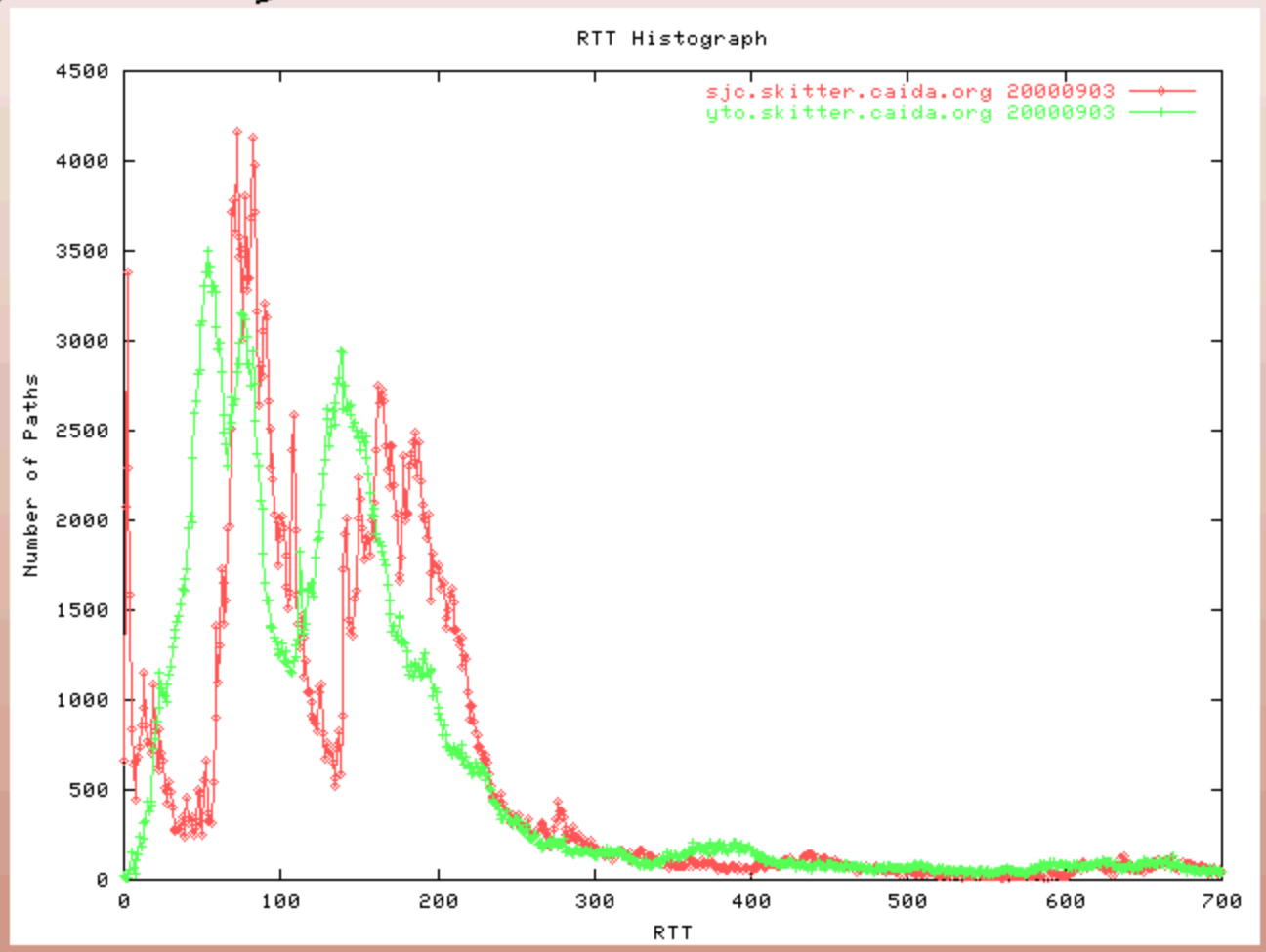
skitter on-going daily summaries

path lengths from yto/sjc.skitter to 35k dsts



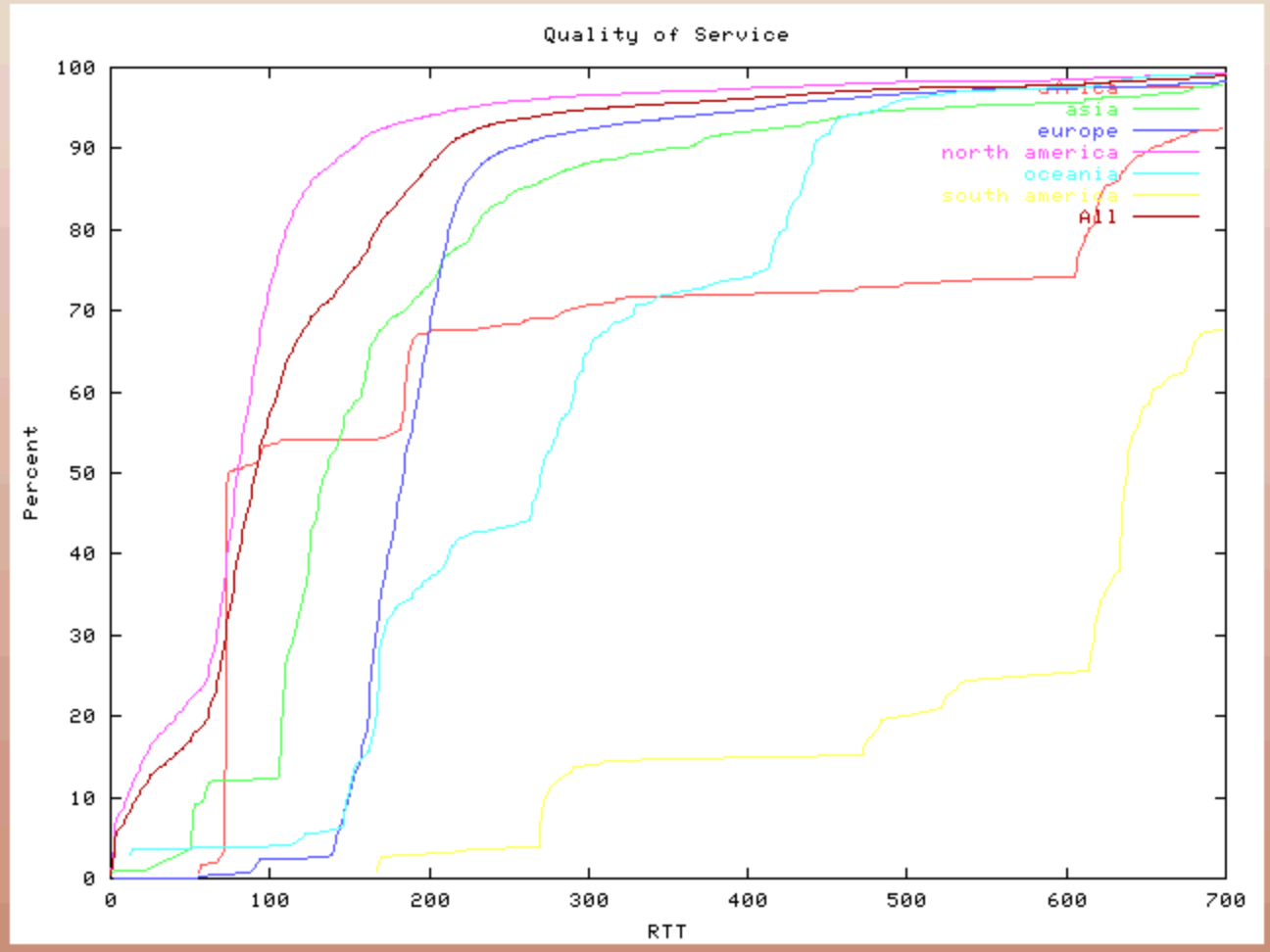
skitter on-going daily summaries

rtts from yto/sjc.skitter to 35k dsts



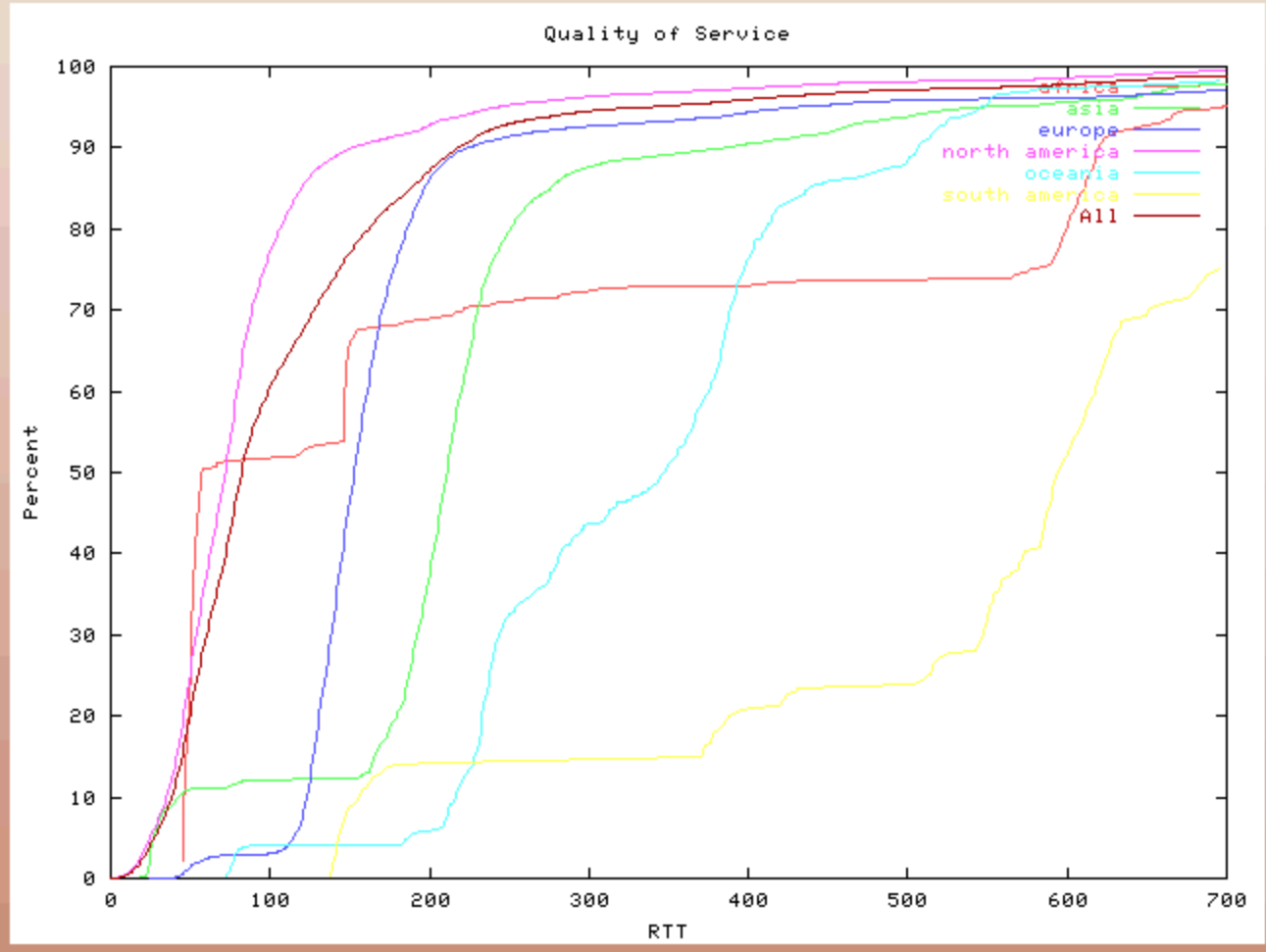
skitter on-going daily summaries

rtt by region from sjc.skitter to 35k dsts



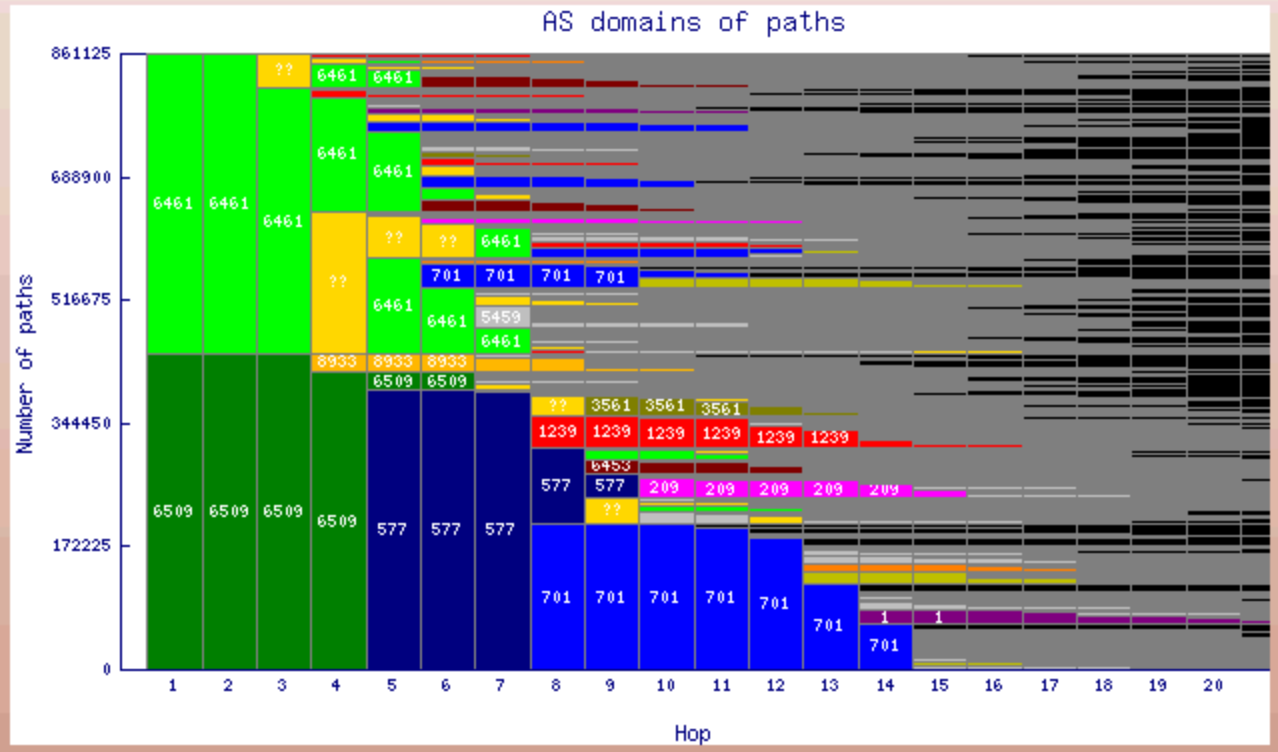
skitter on-going daily summaries

rtt by region from yto.skitter to 35k dsts



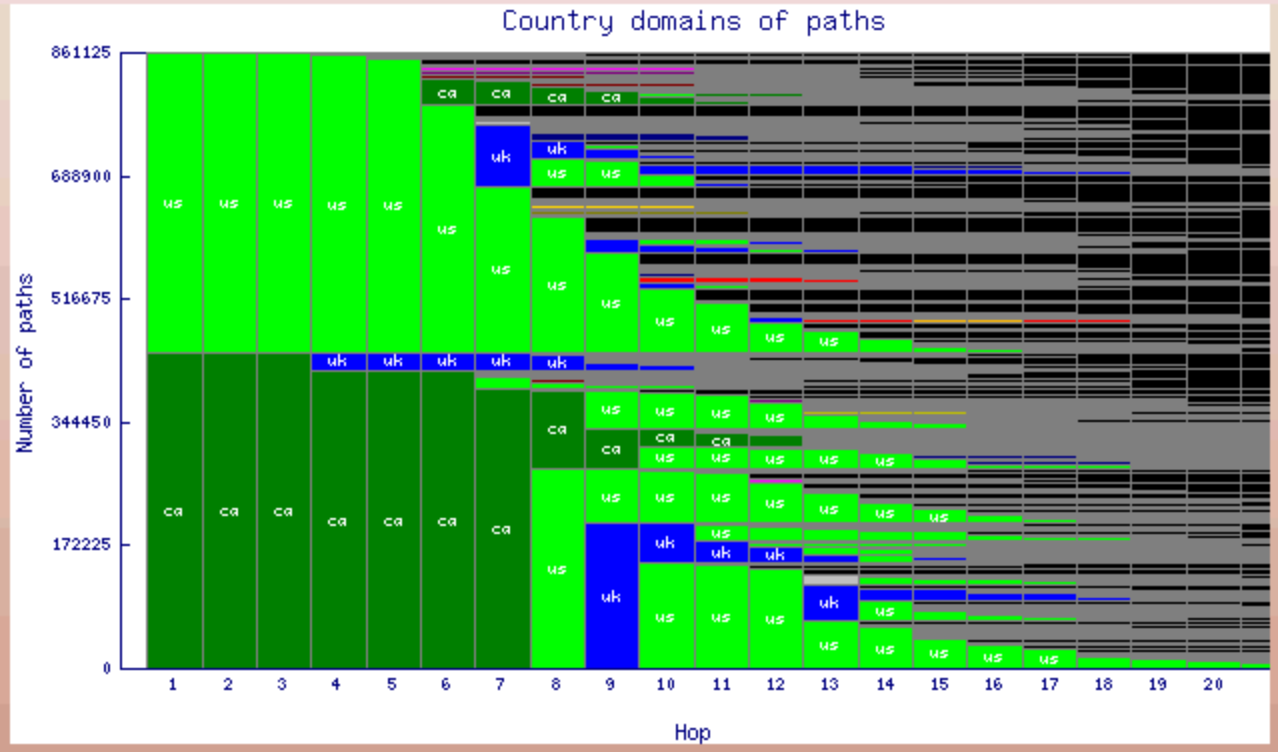
skitter on-going daily summaries

as dispersion graph from sjc/yto to 35k dsts, 090300



skitter on-going daily summaries

country dispersion graph from sjc/yto to 35k dsts, 090300



Internet workload

- many uses
  - capacity planning
  - performance and QOS assurance across ISPs
  - accounting/billing
  - security management
- measurement tools
  - router-based (cflowd, netflow)
  - stand-alone monitors (coral,skitter)
- visualization huge challenge
  - too much data
  - noone correlates across/with much
- evolution requires use
  - envisioning new methods?
  - better data correlation tools are essential

available data: (passive) header traces

coral: oc3/oc12 ‘real’ networks

- HPC sites:

  http://moat.nlanr.net/Traces/

tcpdump: campus/corporate sites

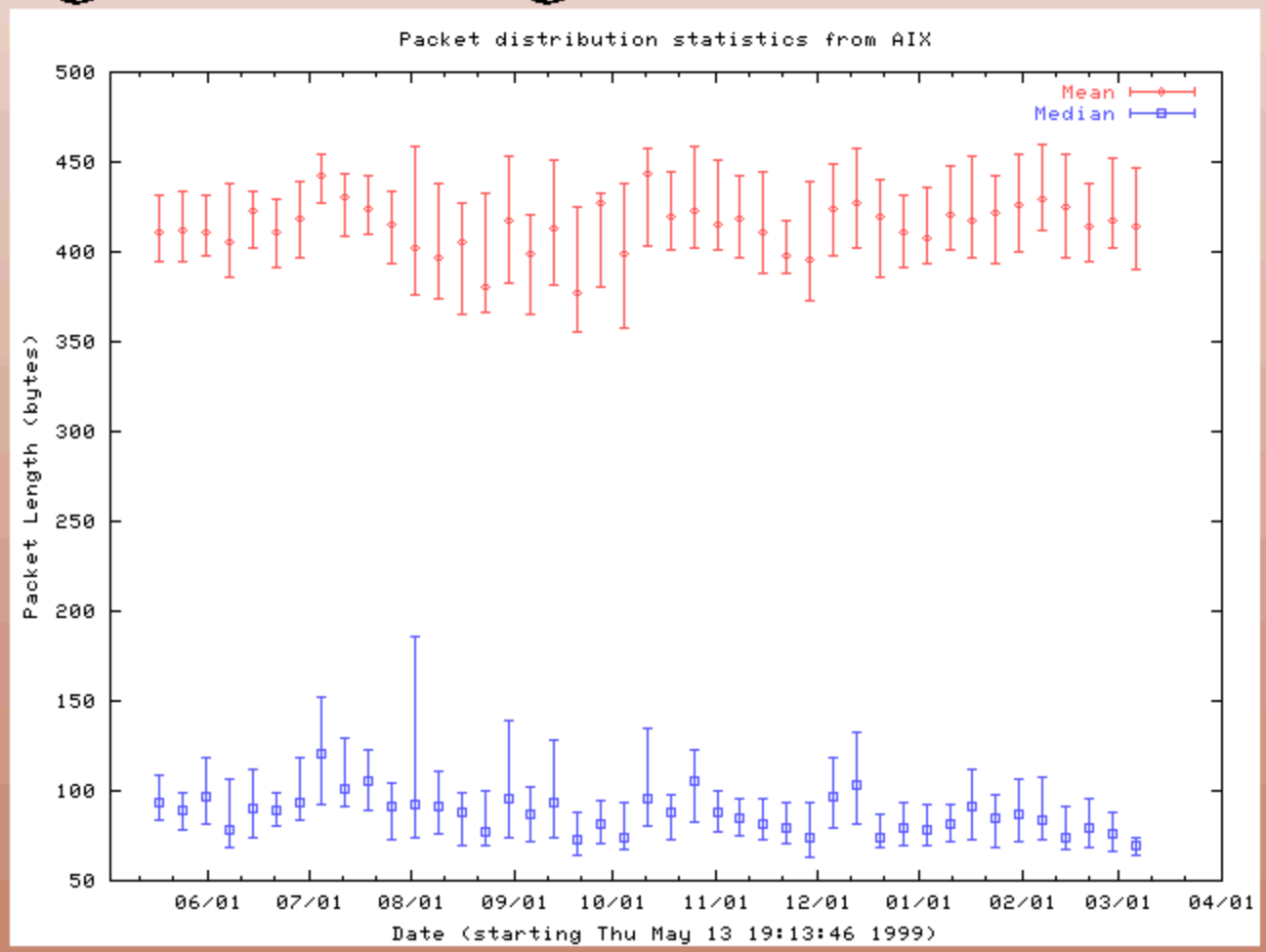
http://ita.ee.lbl.gov/html/traces.htm

have given to NDA’ed researchers

vBNS, ATT

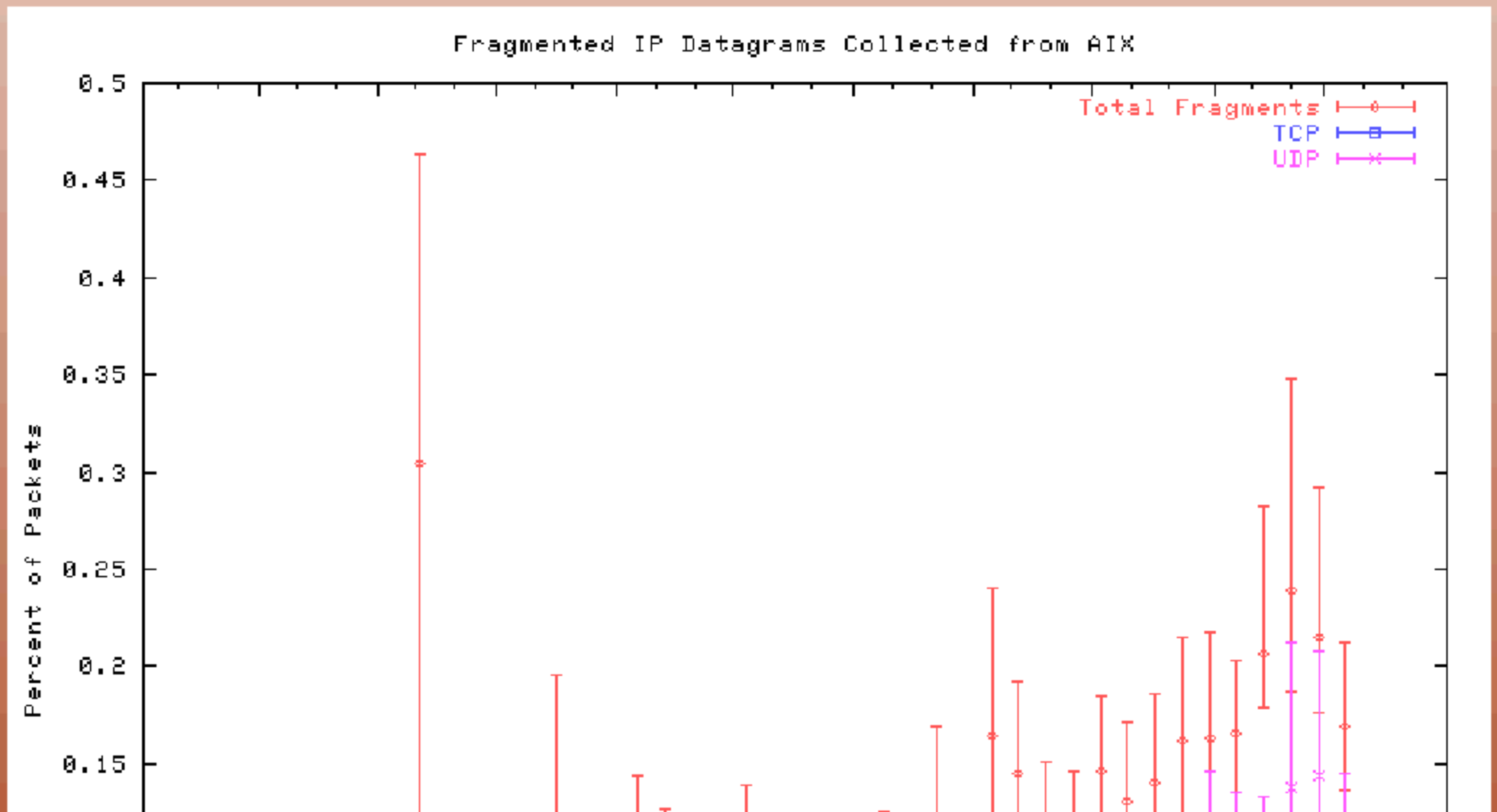
AMES packet size mean/median trend

little change over 9 months
unsurprising as long as TCP dominates



workload: AIX–MAEW fragmentation

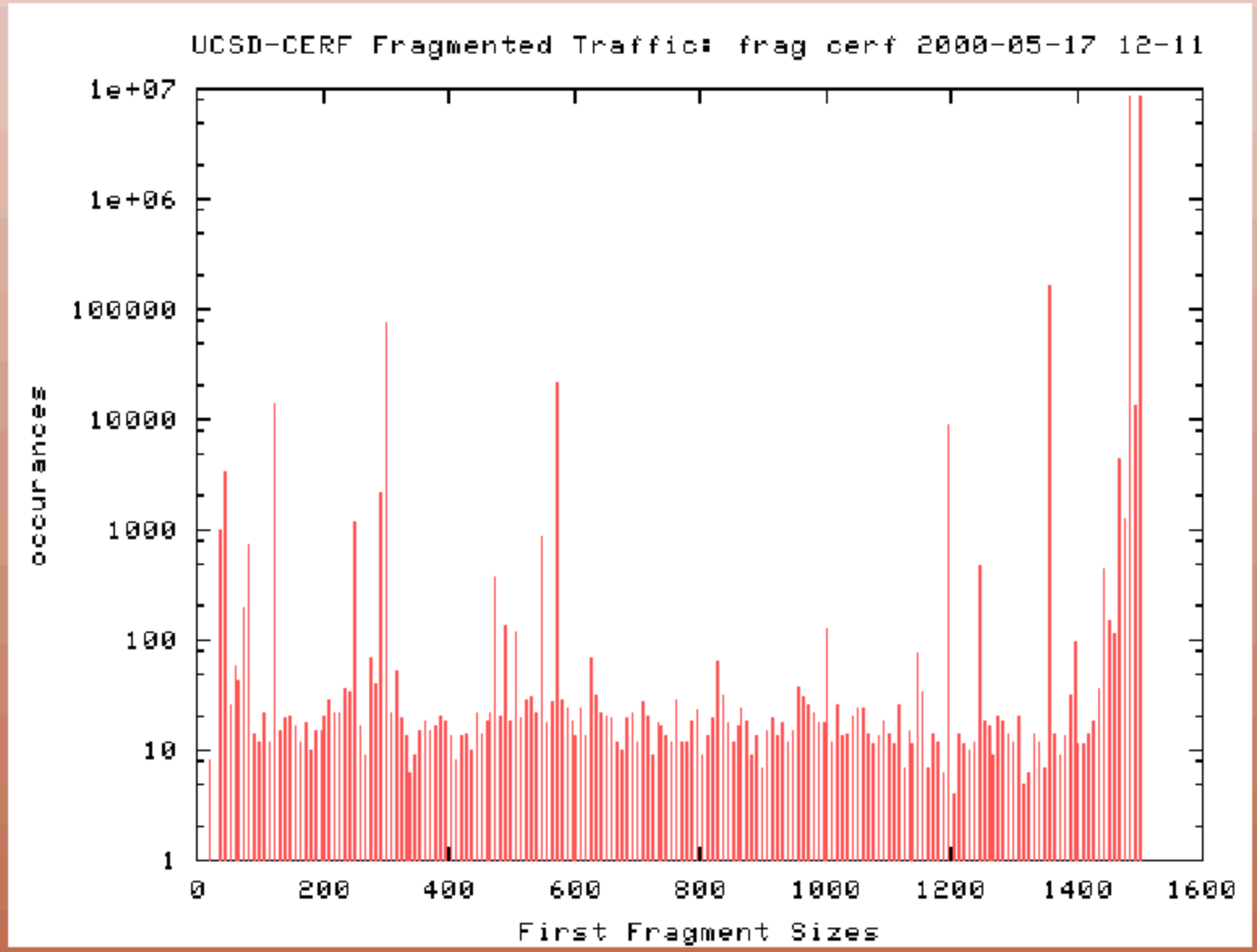
relevant to recent IP traceback techniques [Savage00]
definitely on rise (from UDP) at AIX
almost no TCP frags (MTU disc + small pkts)



workload: CERFNET fragmentation

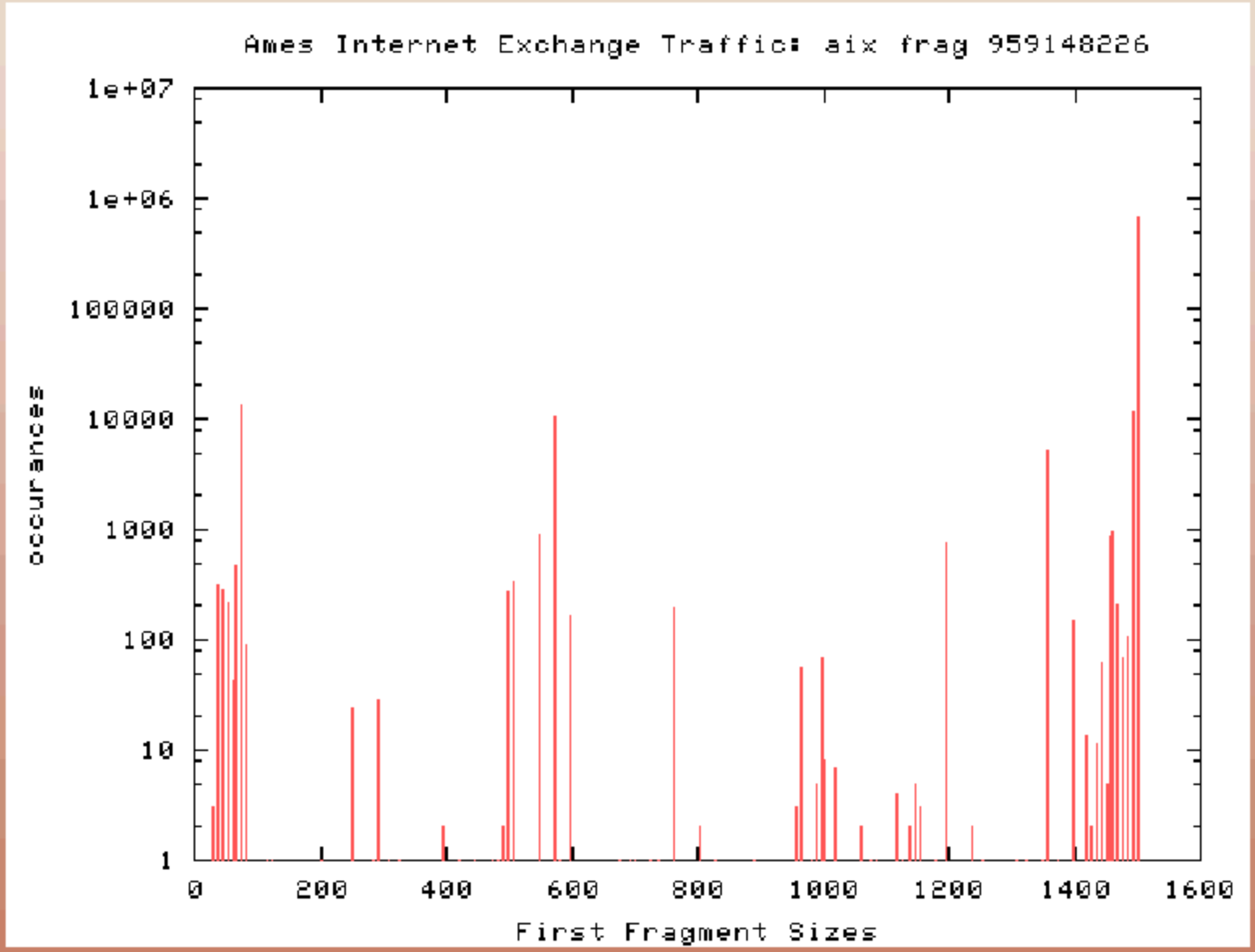
size of first fragment seen in series on UCSD link

component of distribtion is uniform/random

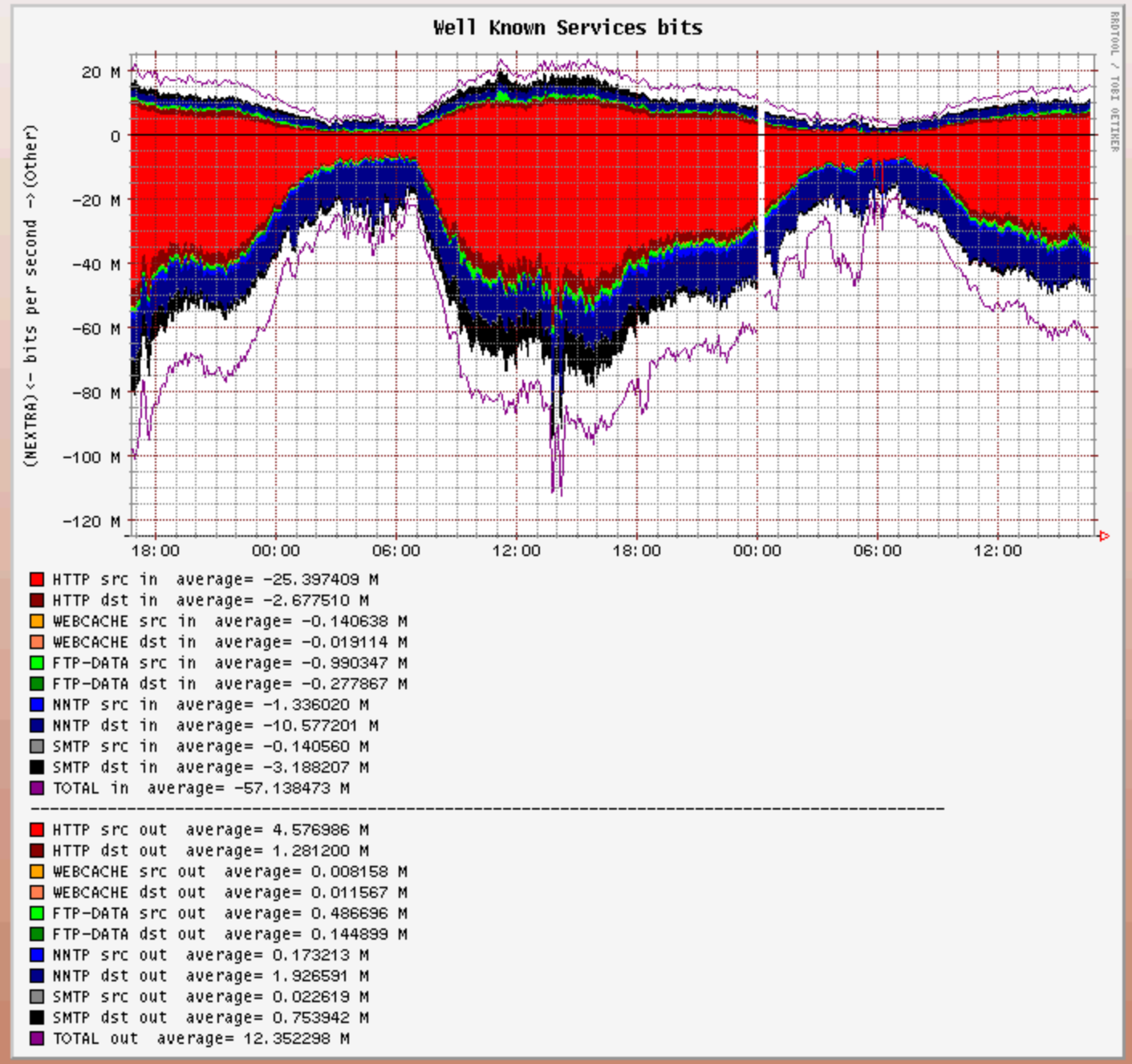


workload: AIX fragmentation

size of first fragment seen in series on AMES AIX link



workload: top applications in/out



workload: other recent work

- UWisconsin flowscan (dave plonka)
  - http://net.doit.wisc.edu/data/flow/size/
- netramet, (nevil brownlee, nz)
  - http://www.caida.org/analysis/workload/netrame
- nextra (alexander kunz, .de)
  - http://flowstats.nextra.de/graphs/
- coralreef (david moore, ucsd)
  - https://anala.caida.org/CoralReef/Demos/

workload data: meta-challenges

- splintered & competitive core
  - limited access to data
  - so difficult to argue 'representativeness'

- network performance impact
  - higher b/w increasing difficult to measure
  - faster speeds and changing transport technologies complicate data acquisition and processing
  - e.g. monitor gone when AIX converts to POS

- user privacy volatile issue
- hard to get data in researchers hands

CAIDA's UCSD/CERFnet link monitor

workload data: challenges

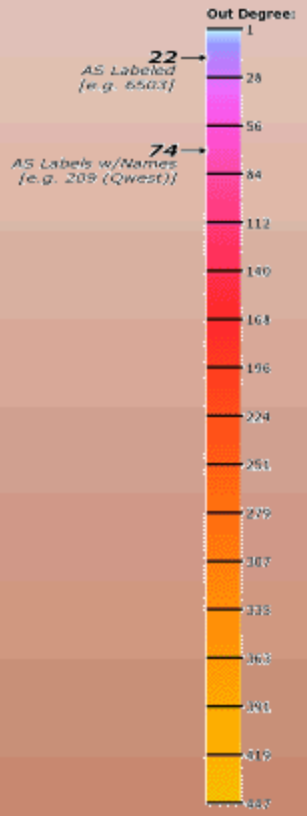
- id and present ‘useful’ workload metrics, particularly given persistence of fire-fighting environment
- id significant patterns, timeframes, correlations
  - vary by user need
  - change as technologies and ’net change
- methodology has many weaknesses
  - dynamic port negotiation (napster)
  - tons of ‘other’ ports unmapped
  - ports not really assurance/unique anyway
  - IPSEC blows away ports anyway
  - need traffic profiling

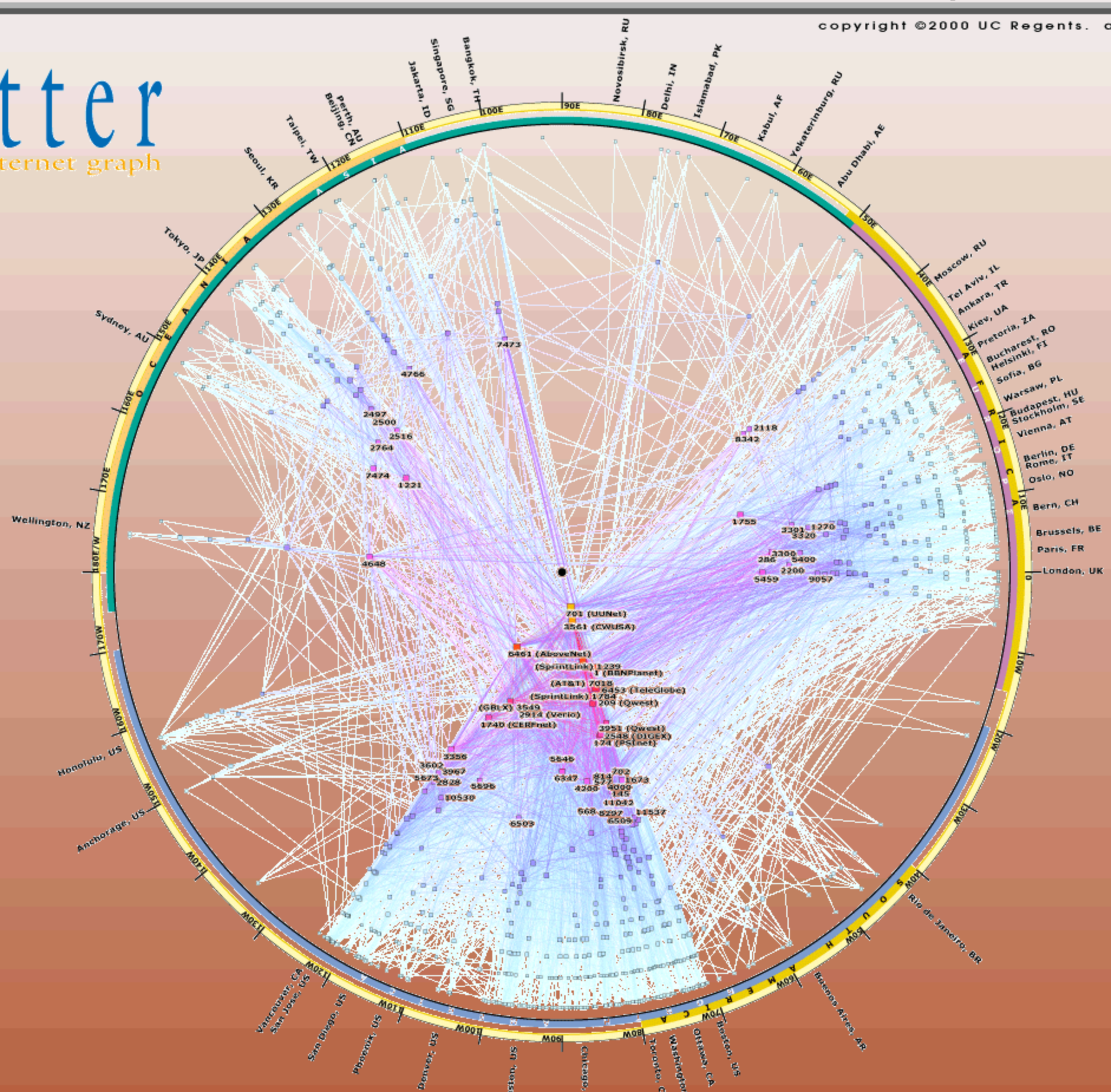
routing & addressing data

- things getting slightly better (data-wise)
  - but BGP/infrastructure getting worse faster
- not much real-time instrumentation on routers
- UO's route-views http://www.antc.uoregon.edu/route-view
- Merit's IPMA http://www.merit.edu/ipma/

skitter: AS interconnectivity

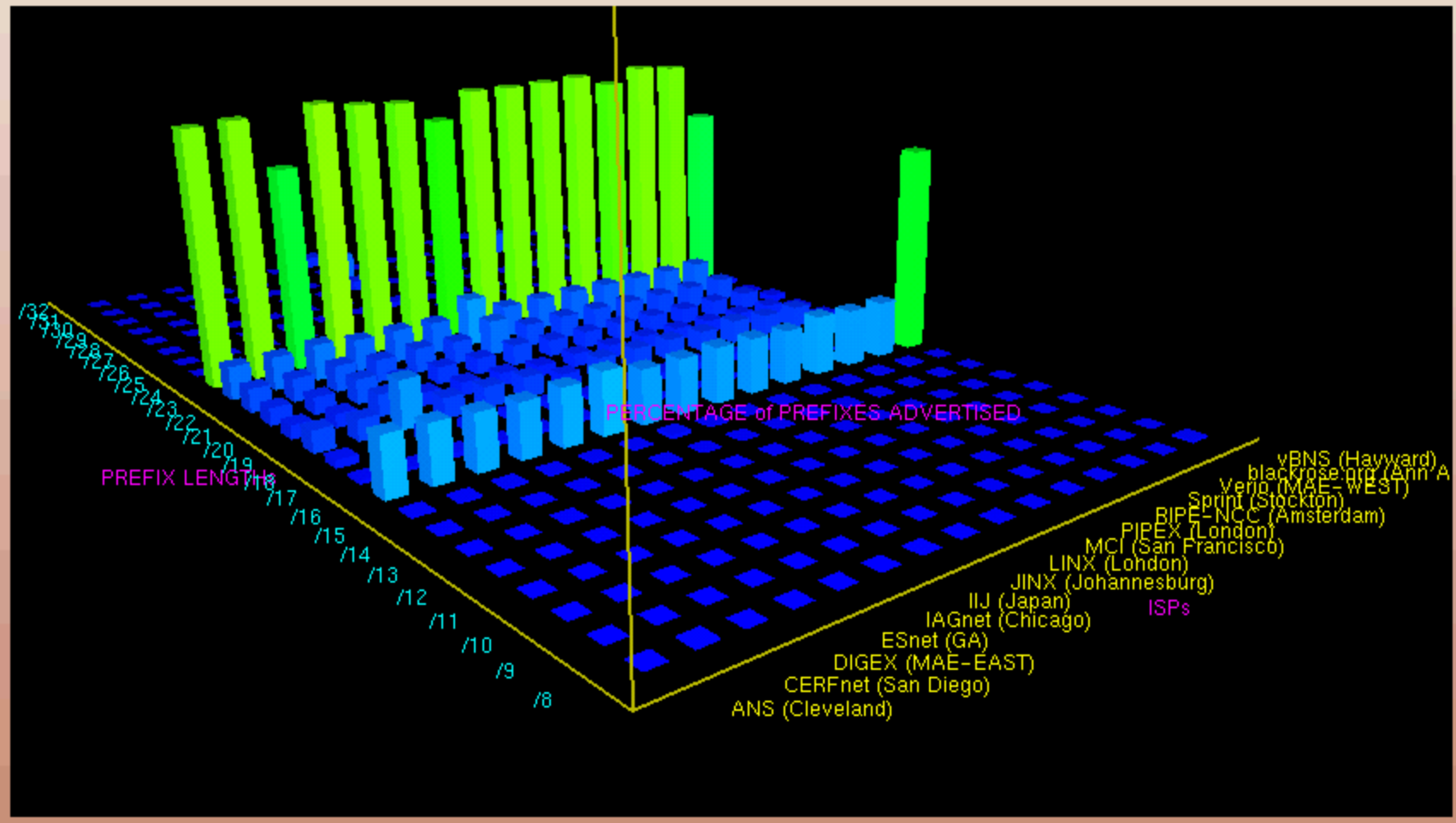
skitter
core AS internet graph





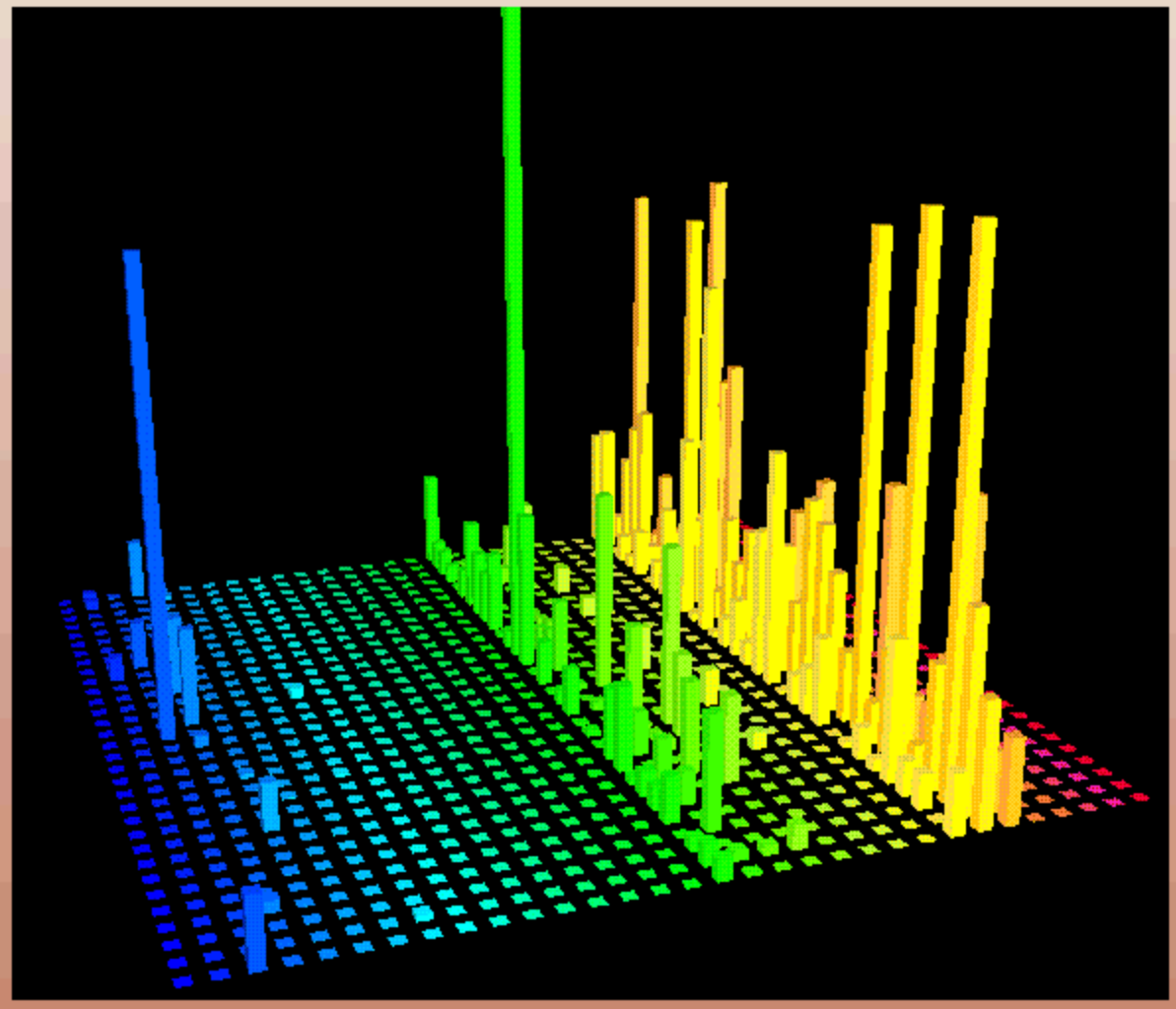
routing: address consumption

prefix length distribution for routes announced by core ISPs, 1–6/1998 (courtesy NLANR/MOAT, Jeff Brown)



routing: address usage of *traffic* sample

32x32 ‘bitmap’ matrix of address space
height is % packets with src IP in that address block



routing: research priorities

- better IP routing instrumentation
- real–time analysis without interfering with performance
- realistic inter–domain routing models

- tasks
  - identification/vis of flaps, outages, critical paths
  - correlation performance with some measure of path ‘length’
  - comparison of forward path with
    - BGP path
    - shortest path
    - reverse path
  - effects of unicast/multicast incongruities?

routing: research obstacles

- routes may change faster than ability to measure or analyze
  - sometimes on purpose (load–balancing)
- poorly instrumented infrastructure (new tools needed)
- prudent security dictates inhibiting research
- mapping IP address to anything (deja vu)

now what?

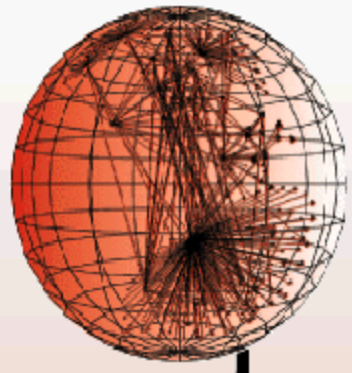
- the ideal:
  - well-instrumented infrastructure
  - seamless integration of variety of data sources
  - important for simulation/prediction (& lately, operations)
  - but unlikely for the foreseeable future
- tools still need:
  - interpret of vast quantities of data in real-time
    - geographically & logically distributed
  - user-friendly integration with network utilities and control systems
  - inter- & intra-ISP feature detection
  - new methods for data collection, reduction, aggregation, and mining (GByte or Tbyte datasets)

setting expectations

rule 1: no magic data sets (not so far anyway)

the so-called science of poll-taking
is not a science at all
but a mere necromancy.
people are unpredictable by nature,
& though you can take a nation's pulse,
you can't be sure that the nation
hasn't just run up a flight of stairs.
--E. B. White

New Yorker, Nov 1948.



www.caida.org/Presentations/

kc claffy
UCSD/SDSC/CAIDA
kc@caida.org
www.caida.org