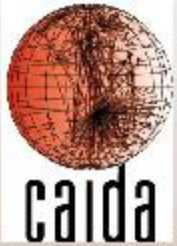


Internet measurement: state of DeUnion

no aphorism is more frequently repeated...
than that we must ask Nature few questions,
or ideally, one question at a time.
...this view is wholly mistaken.
Nature will best respond to a logically
and artfully thought out questionnaire;
indeed if we ask her a single question, she will often refuse
to answer until some other topic has been dicussed.
-- perspectives in medicine and biology 1973 sir ronald a fisher

15 nov 01
kc claffy, UCSD/SDSC/CAIDA
kc@caida.org
www.caida.org

abysmal but unsurprising

- little capacity to predict, depict, or even measure traffic behavior on current and advanced networks
- few tools to engineer/operate networks or identify traffic anomalies in real time
- doesn't stop researchers from building junk
- doesn't stop random users from doing random junk (no dearth of activity)

increasing [opportunity] cost to infrastructure

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 is 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?
- losing battle?

Internet's resistance to measurement

many would benefit

- vendors, users, researchers, ISPs

ISPs would bear cost

- multiple media: atm, pos, dwdm, mpls
- logistics/management
- privacy implications
- analysis/research obsolete after (before) done

.....how to justify/accomplish measurement?

(when market forces are torqued)

alternatives:

1) tools that positively affect an ISP's balance sheet
2) regulatory intervention

payoffs

- insights for vendors re next generation hw/sw requirement
- calibration for users, e.g., monitoring service level agreements
- diagnostic and planning tools for ISPs
- windows into the infrastructure for researchers

measurement tools lack

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

so what happened instead

andrew odlyzko's excellent "myth of Internet growth" study (nov 2000)

- 'traffic doubling every 90 days'
  - maybe for a few months in 1995–1996
  - in reality, no real data since 1995 (nsfnet sunset)
  - more like every 12–18 months for rest of 1990s
- financial markets (at least in US) believed (bubbly!) estimates
- over 6 years, that means a factor of 16 million
  - assume (generously) 500M users, 1.5Mbps per user around the clock
  - and yet we're mostly still using 28k modems, & only for an hr/day, and ave 5k bits/sec then
  - the math just does not work out
- it took 5 years for true traffic growth data to finally manifest itself
  (since providers would not release data, if they even had it)
  via other metrics (hardware and bandwidth sales)

that's actually an embarrassingly pathetic willingness
to just ignore real data (or just invent it)

Internet traffic growth myths (cont.)

- costs
  - tech stock bubble?
    - really takes new technologies a decade to penetrate
    - web was exception, Internet is not
  - retarded technical developments
  - negligence of what users what, and likely to get
    - community gets mired in sub-necessary QOS hubbub, ATM, GMPLS
- benefits
  - unparalled platform for innovation
  - open standards, rapid development of new services
  - big empty pipes was key factor in supporting revolution
- lessons
  - 25 year contracts for pipes should be amortized over 3
  - capacity planning impossible
  - simplify engineering (atm/sonet --> IP over WDM, GigE)

three 'waves' of application

- first wave: shared (remote) use of computers
  - telnet, email, ftp
- second wave: client/server model, formatted lang.
  - web
- third wave: collaborative, peer-to-peer, interactive
  - napster, imesh, kazaa, gaming, video

emergence of third wave will require more real-time interaction with and reaction to network status

the growth of these applications will be self-limiting
(by user frustration with performance) unless we have either:

- a better grip on measurement
  - either done by the applications themselves (e.g., vat)
  - or via some other middleware aspect of the infrastructure)
- or no service-affecting queueing anywhere in the network
  - (seems unlikely)

four areas of measurement

- workload characterization (passive)
- topology (mapping, path dynamics)
- performance evaluation (active, passive)
- routing (dynamics)

caida focuses on

- measurement tools
- macroscopic analyses
- identifying priorities and obstacles

workload characterization

- workload profiling (s/w & h/w design, architecture optimizing, capacity planning)
- security
- performance analysis
  - delay, loss, jitter?
- QOS assurance across ISPs
- accounting/billing

- tools: netramet, netflow, cflowd, coral
  - some suck less? ...evolution requires use

workload characterization: priorities

- coral/ocXmons (OC3,12,48,192 gigE)
- persistent, real–time, full–frame collection
- dynamic packet filtering triggered by attack precursors
- security policy
  - compliance auditing (passive)
  - enforcement (active)

obstacles
- hardware expensive
- privacy issues
- IPsec

wkld char. on OC48 backbone link: hardware

DAG – PCI network monitoring cards:

- project at University of Waikato (New Zealand) computer science department
  - http://dag.cs.waikato.ac.nz/
- DAG4 card – ATM and PoS capture at OC48c
  - 2.5 GBit/sec link rate
  - exceeds PCI bus bandwidth, requires filtering & compression
  - provides highly accurate timestamping
  - timestamp sync across boards available via cable and GPS
- CAIDA/U.Waikato collaboration (subcontract)

workload on OC48 backbone link: data

- data collected by CAIDA and Waikato Applied Network Dynamics group
  - http://wand.cs.waikato.ac.nz/
- collected on OC48 commercial backbone link in San Jose, CA
- oc48mon2 link, one direction only
- duration: 76 minutes total
  - 20:00 – 21:16 (PDT), 5 Aug 2001
- volume: 32 GB of data

workload on OC48 backbone link: analysis

use coralreef software suite

- http://www.caida.org/tools/measurement/coralreef/
- public domain

obtain quantitative parameters of captured traffic:

- packet and byte rates
- flows (src IP, src port, dest IP, dest port, protocol)

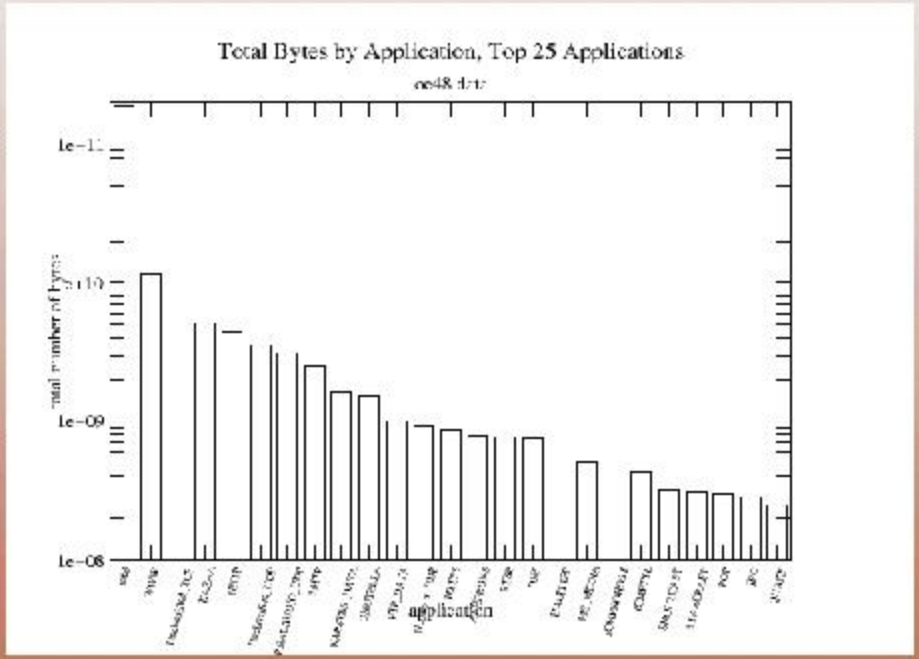
use NetGeo tool to map src/dst IP addresses to ASes and countries

- http://www.caida.org/tools/utilities/netgeo/

consider various aggregations of traffic:

- applications
- ASes
- countries

results: application characteristics



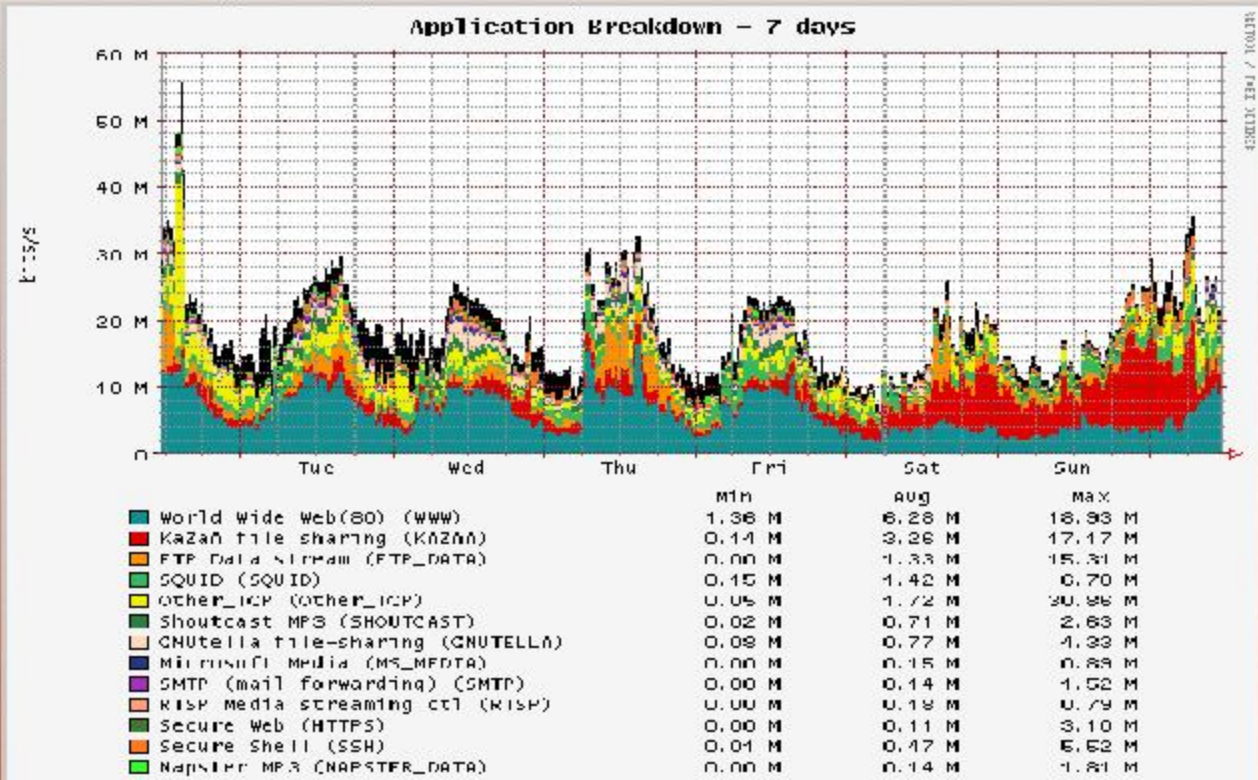
– well–known applications are determined from port numbers
– plot distributions of bytes, packets, and flows by applications

results: traffic by (top 10) applications

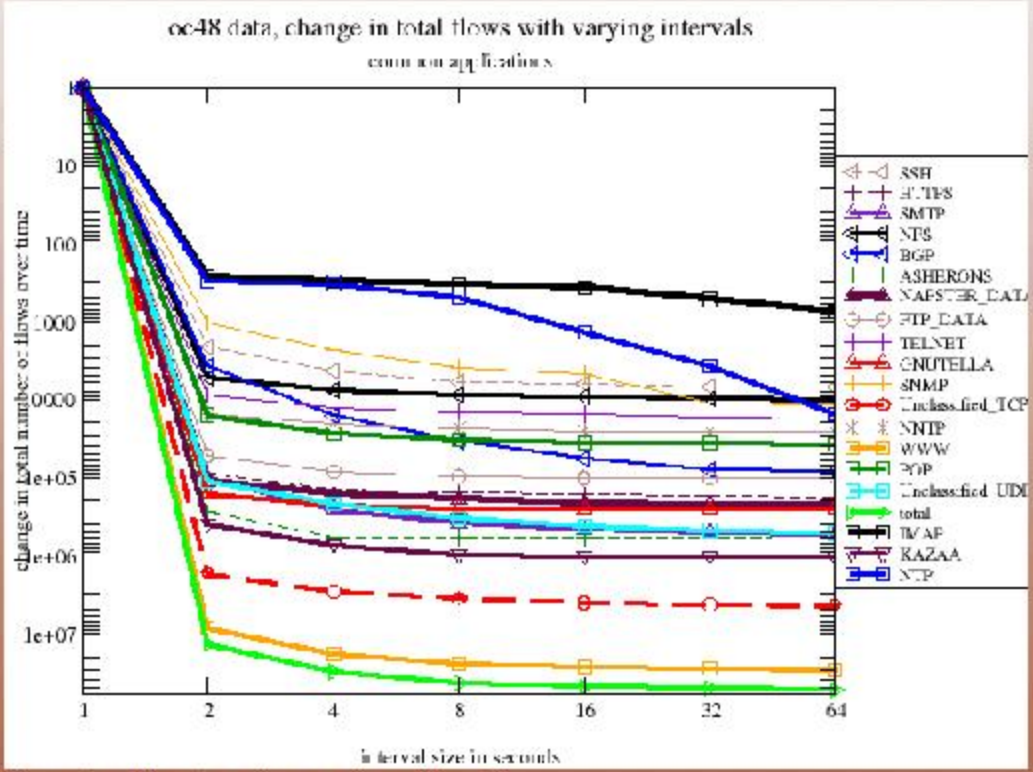
- by bytes
  - www (79%)
  - unclassified TCP (4%)
  - kazaa – peer-to-peer file sharing system (for music) (2%)
  - 1% or less:
    - nntp – netnews; unclassified UDP; realaudio, smtp, napster, gnutella, ftp
- by packets
  - www (67%)
  - unclassified TCP/UDP (7%/3%)
  - 1% or less:
    - Halflife (game); ICMP (e.g., ping – small pkts); smtp; kazaa (large pkts), dns, realaudio, Starcraft (game)
- by flows
  - www (69%)
  - ICMP (16%)
  - dns (3%)
  - unclassified TCP (2%)
  - asherons (game) (2%)
  - 1% or less:
    - smtp, unclassified UDP, ftp, https (secure http), Halflife

workload char.: new killer application?

compare to university commodity link (UCSD)
3–12 nov 01 (post–napster.demise)



results: flow interval spacing



- packet interarrival time is _always_ less than 2s
- 4s timeout sufficient to capture most application flows
- cannot do any such analysis via sampling

results: traffic by countries/continents

distribution of bytes, packets, flows by source and destination countries

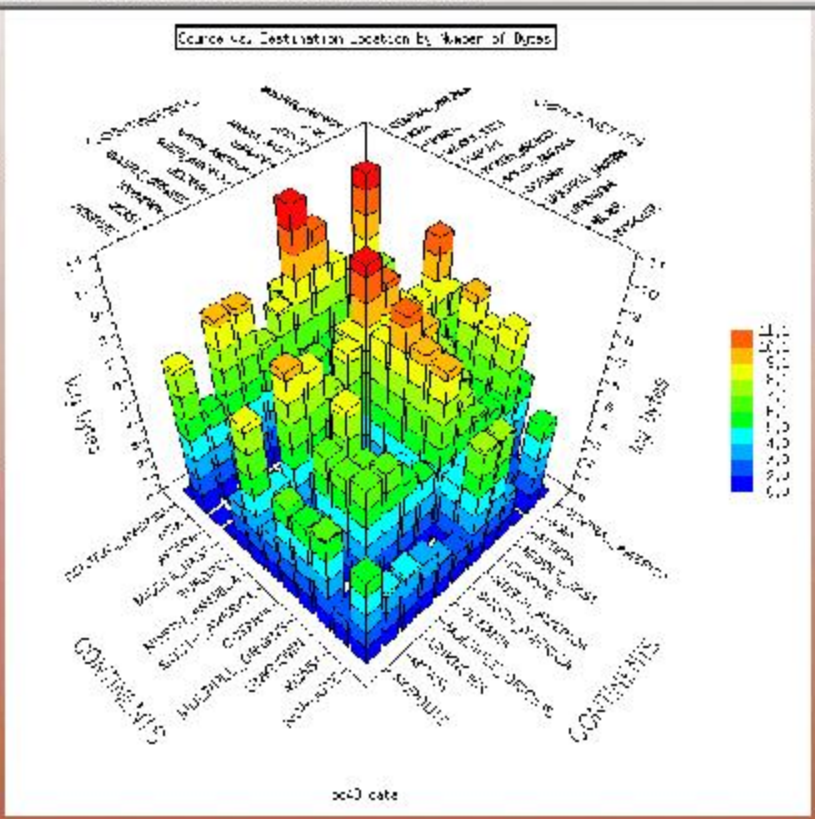
- top source countries:
  - US – 1st by bytes and packets, 2nd by flows
  - Japan – 1st by flows, 2nd by bytes and packets
  - also: Canada, United Kingdom, Hong Kong, Denmark (but: 10 times less bytes, packets, or flows than US or JP)
- top destination countries:
  - Korea, US, China
  - also: Japan, New Zealand, Taiwan, Australia

suggests that routing policy for this link directs most traffic to asia

results: 3D traffic matrices

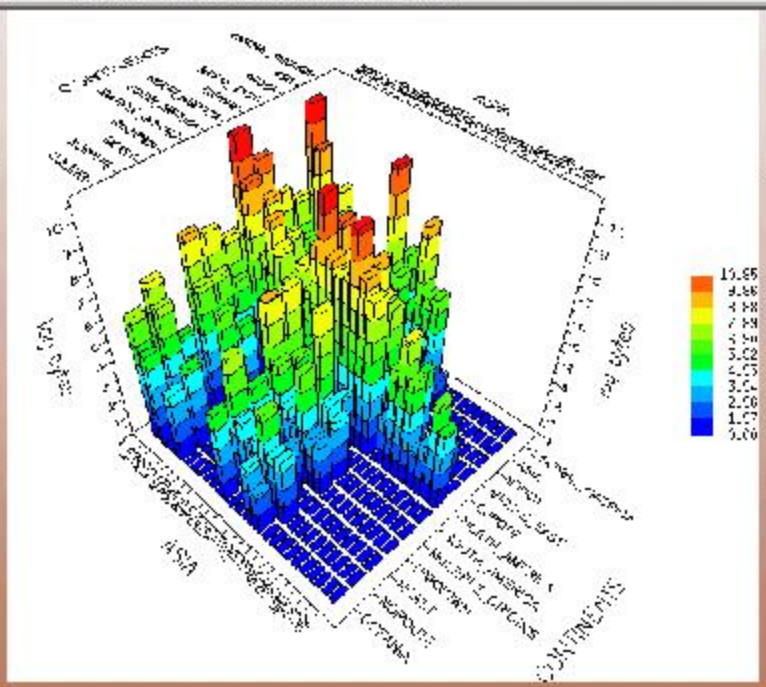
- source/destinations pairs aggregated by continents/regions
- 3D plots use XRT-based tool
  - http://www.caida.org/tools/utilities/graphing/graph_xrt3d.xml
  - x/y axes - source/destination locations
  - z-axis - log scale traffic volume (bytes, packets, or flows)
  - examples follow (byte volume)

results: 3D traffic matrices



peaks: asia/n.amer/eur. –> asia; n.amer/eur. –> n.amer; n.amer/eur –> oceania

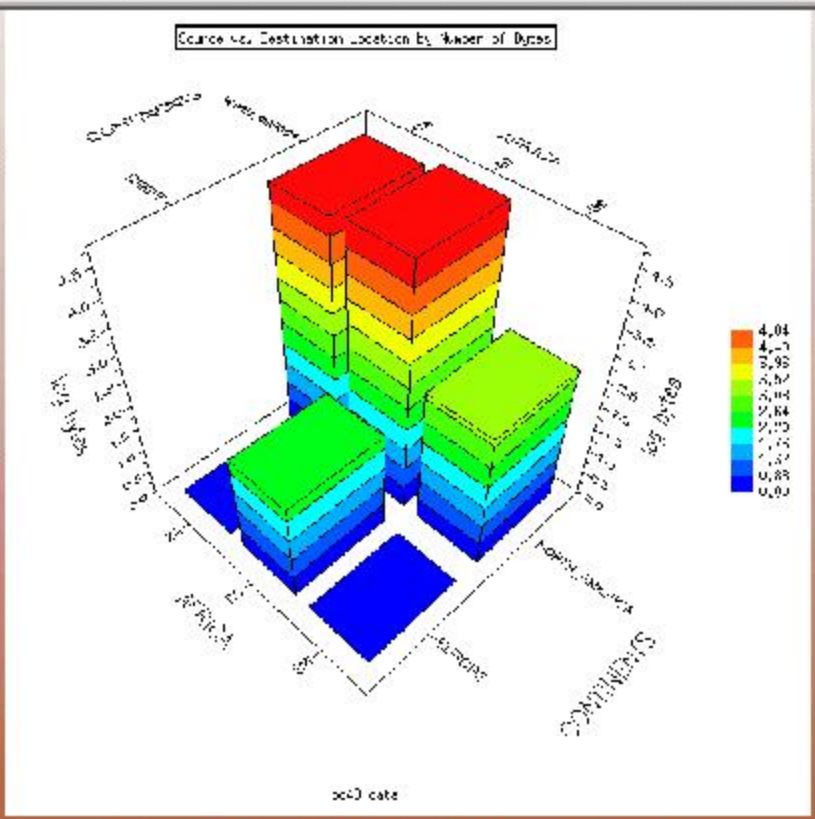
results: 3D traffic matrices



continents–to–asia: primary srcs – n.amer/asia/europe
primary dsts – south korea, china, japan, taiwan

east asia locations are primary traffic destinations for this link

results: 3D traffic matrices (continents)



continents–to–africa: only 3 countries receive (little) traffic: .za, .eg, .bw

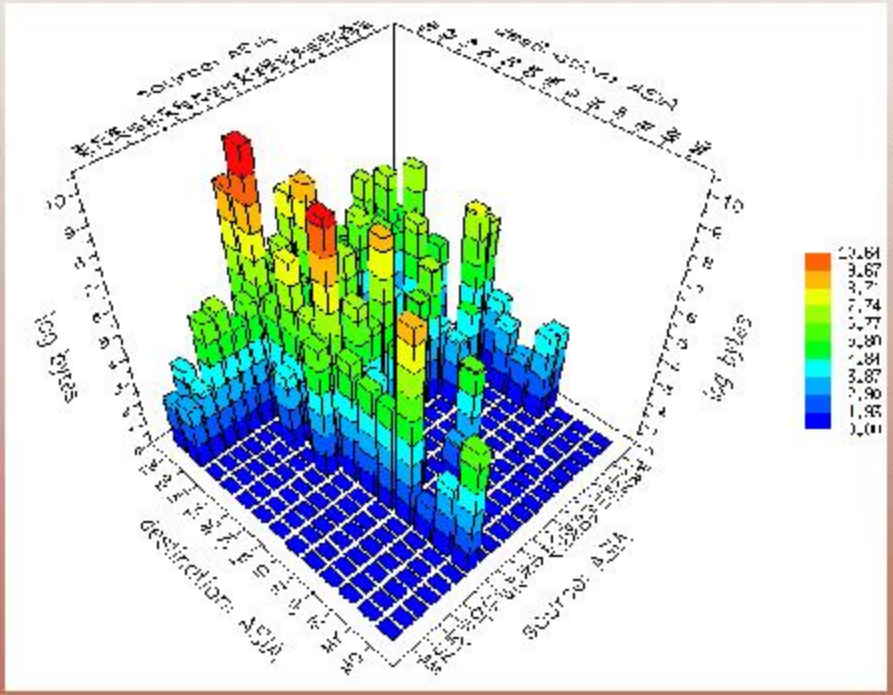
results: anomalies

we see the following unexpected traffic at our measurement location in San Jose, California, US:

- asia to asia – rather significant amount

- .uk to .eg (egypt)
- .uk to .tu (turkey)
- .fr to .fr and .uk to .uk
- .se (sweden) to .es (spain)

results: anomalies (example)



significant amount of asia–to–asia traffic passes through San Jose!
includes even same country traffic (e.g., .jp–>.jp, .tw–>.tw)

workload on OC48 backbone link: conclusions

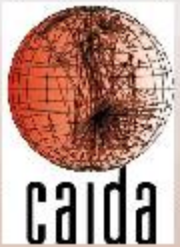
- we infer applications using port mappings
  - dominated by WWW, filesharing, and gaming applications
- flow duration
  - relatively short timeouts (i.e. 4 seconds) may be adequate
- traffic destinations
  - dominated by East Asia (KR, TW, CN, JP) and US
- significant traffic 'anomalies' through San Jose
  - western europe – western Europe
  - eastern asia – eastern asia
  - likely typical, routing policies a function of economic and regulatory realities

workload on OC48 backbone link: conclusions

- unique
  - first and only OC48 flow monitor worldwide
  - caida's public tools analyze data without modification
- software implemented
  - CoralReef, NeTraMet, custom routines (CAIDA)
  - custom routines by U. of Waikato, others
  - darpa/nsf/caida members funded
- software, data analysis, viz tools all prototypes
- backbone core now needs oc192/oc768 monitoring

now what?

- 'seamless': no such thing
- measurement tools/infrastructure
  - well-considered
  - strategically deployed
  - collaboratively maintained
- more infrastructure-relevant research on resulting data
  - feedback into tool design
- correlation among data sources/types, simulation, visualization
- proactive participation
  - top-down (app developers scope constraints)
  - bottom-up (ISP cooperation)



www.caida.org/outreach/presentations/

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