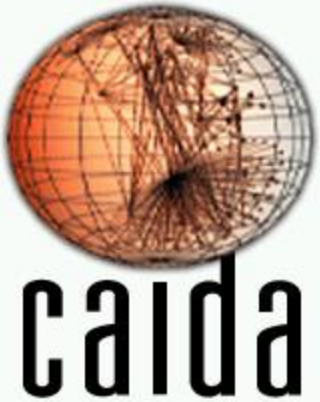


university of california, san diego (ucsd)
san diego supercompter center (sdsc)
cooperative association for Internet data analysis (caida)

kc@caida.org
www.caida.org

outline

- measurements
  - hardware
  - data
  - analysis
- results
  - traffic by applications
  - traffic by ASes
  - traffic by countries
  - ‘geographical’ 3−D plots
- conclusions

measurements: 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)

measurements: data

- data provided by Waikato Applied Network Dynamics group
  - http://wand.cs.waikato.ac.nz/
- collected at Metromedia Fiber Network (MFN) backbone, 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

measurements: analysis

use CoralReef software suite

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

obtain quantitative parameters of captured traffic:

- Byte rates and Packet rates
- Flows
  - Flow = (src IP, src port, dest IP, dest port, protocol)

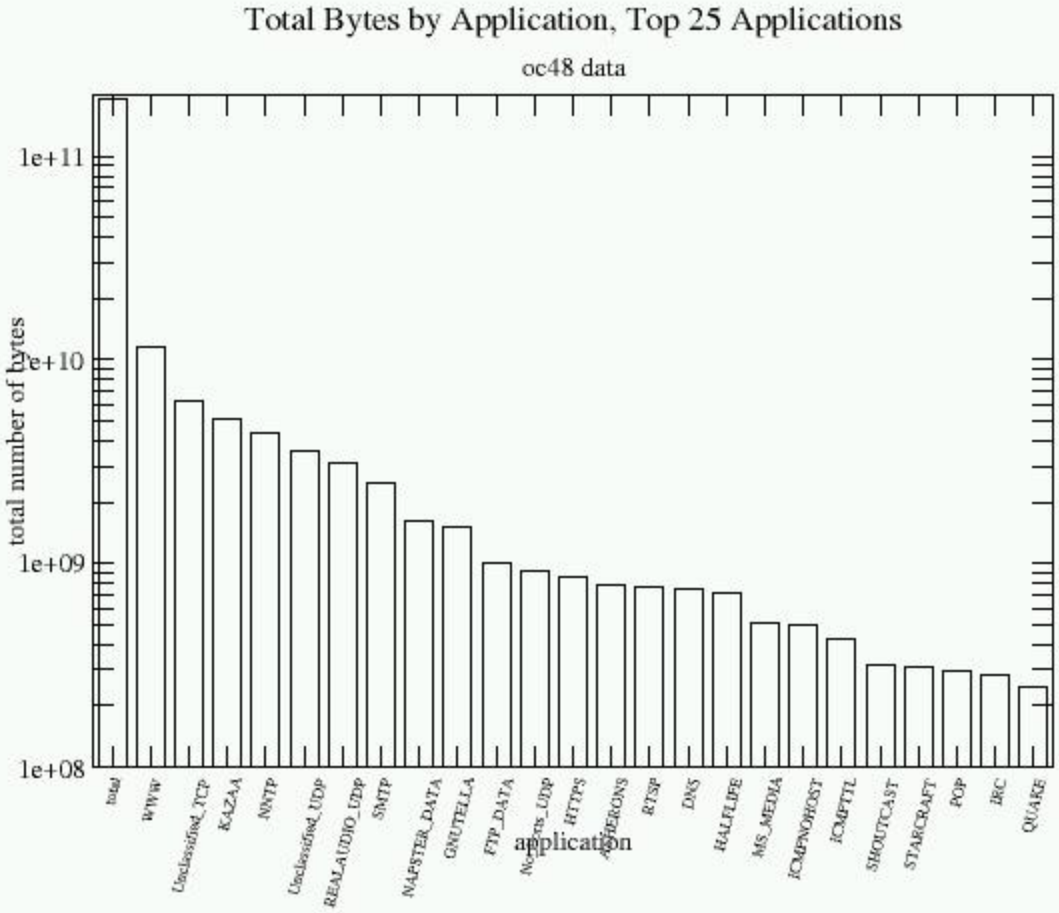
use NetGeo tool to map src/dst IP addresses to ASes and countries

- 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:

results: characteristics of traffic by ASes

map IP addresses to their origin Autonomous Systems

consider distribution of bytes, packets, flows
– by source and destination ASes

- top source ASes: Microsoft–AS–Block
  - also: Shawfiber, JNIC–ASN, Abovenet, ACTTG, Telus
- top destination ASes:
  - Hanaro (Korea), Chinalink (China), Dacomnet (Japan), Thrunet (Korea)
  - Abovenet, AOL–Primehost, Chinanet–core–wan–north, Backbone–Guangdone–AP
  - Hotmail–AS: 5th by flows, 17th by bytes

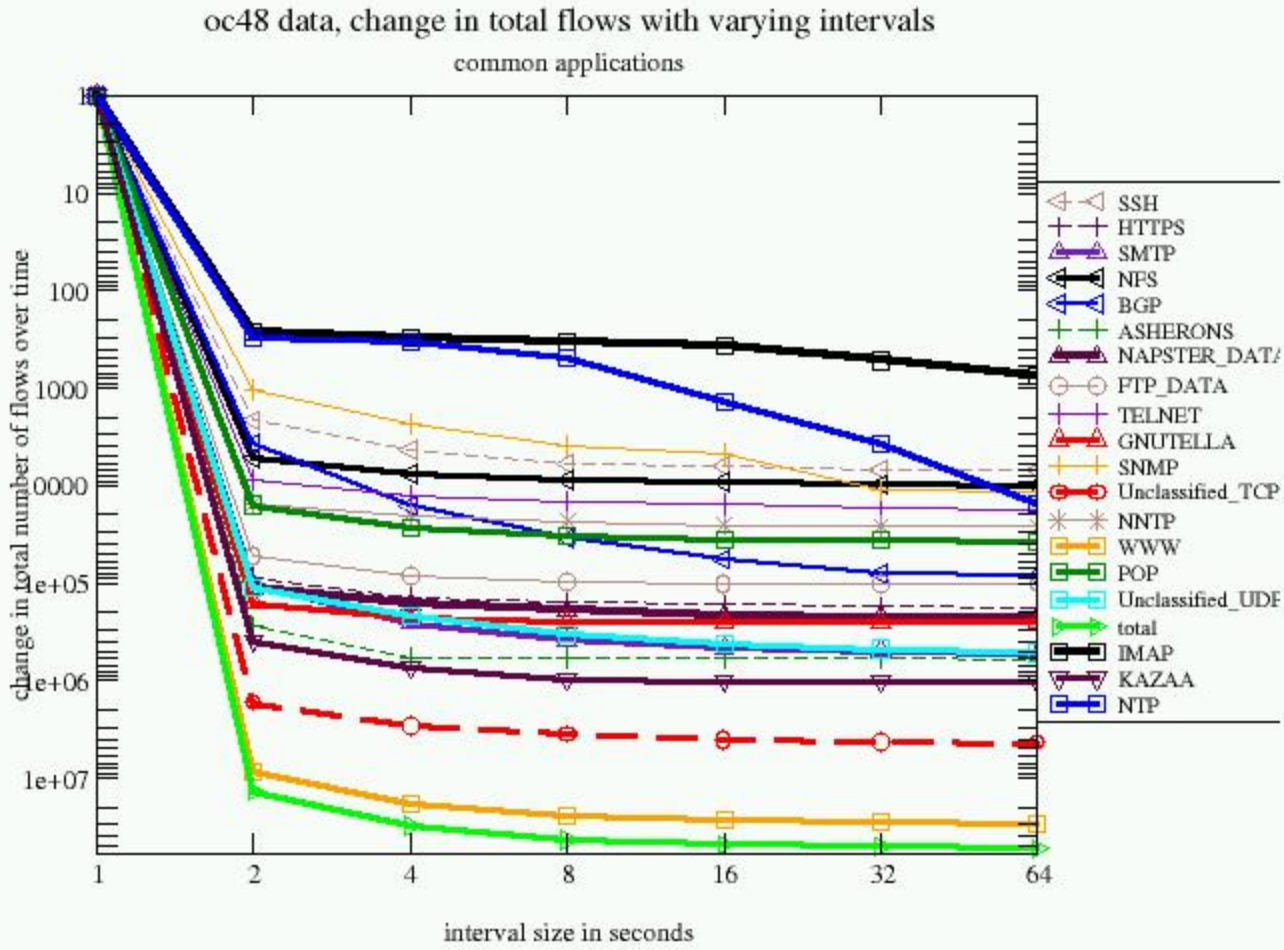
results: traffic by countries/continents

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

AS analysis and geographical findings both reflect nature of traffic passing through MFN backbone: mostly directed to asia

results: flow interval spacing

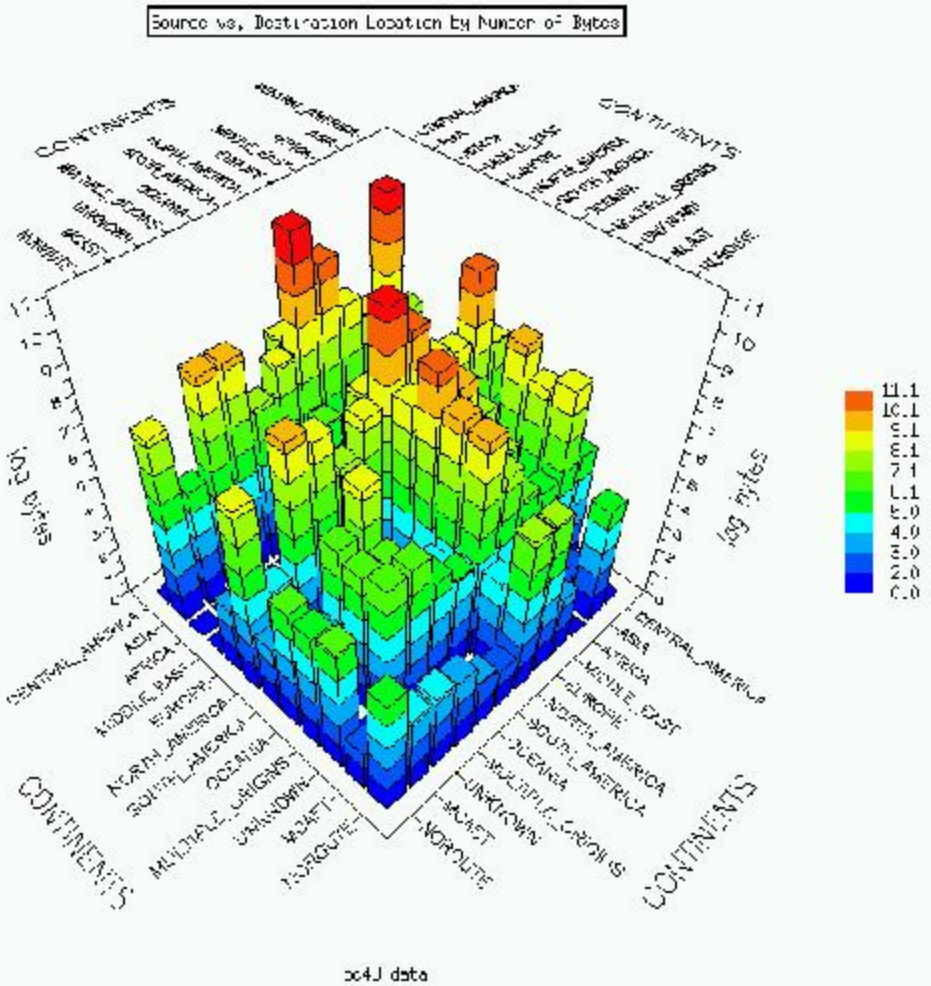


- packet interarrival time is _always_ less than 2s
- 4s timeout sufficient to capture most application flows

results: 3D traffic matrices

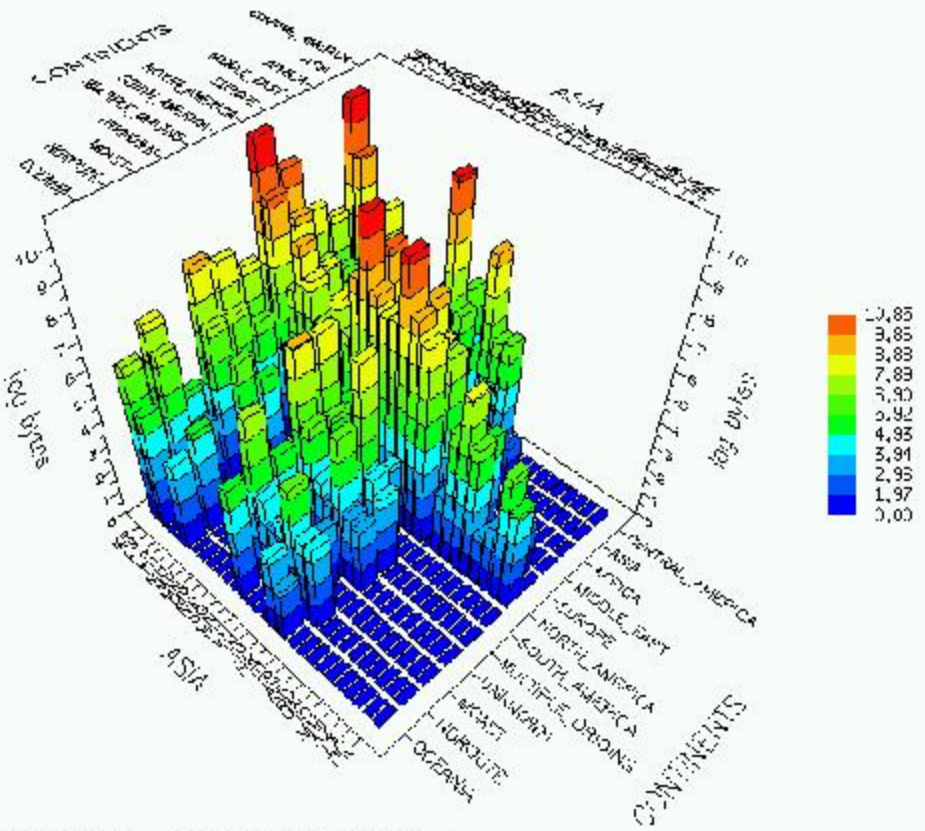
- source/destinations pairs aggregated by continents/regions
- 3D plots use XRT−based tool
  - http://www.caida.org/tools/utilities/graphing/graph_xrt3d.xml
  - x/y axes − source/destination locations
  - z−axis − logarithmic scale traffic volume (bytes, packets, or flows)
- examples follow (number of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

results: 3D traffic matrices



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

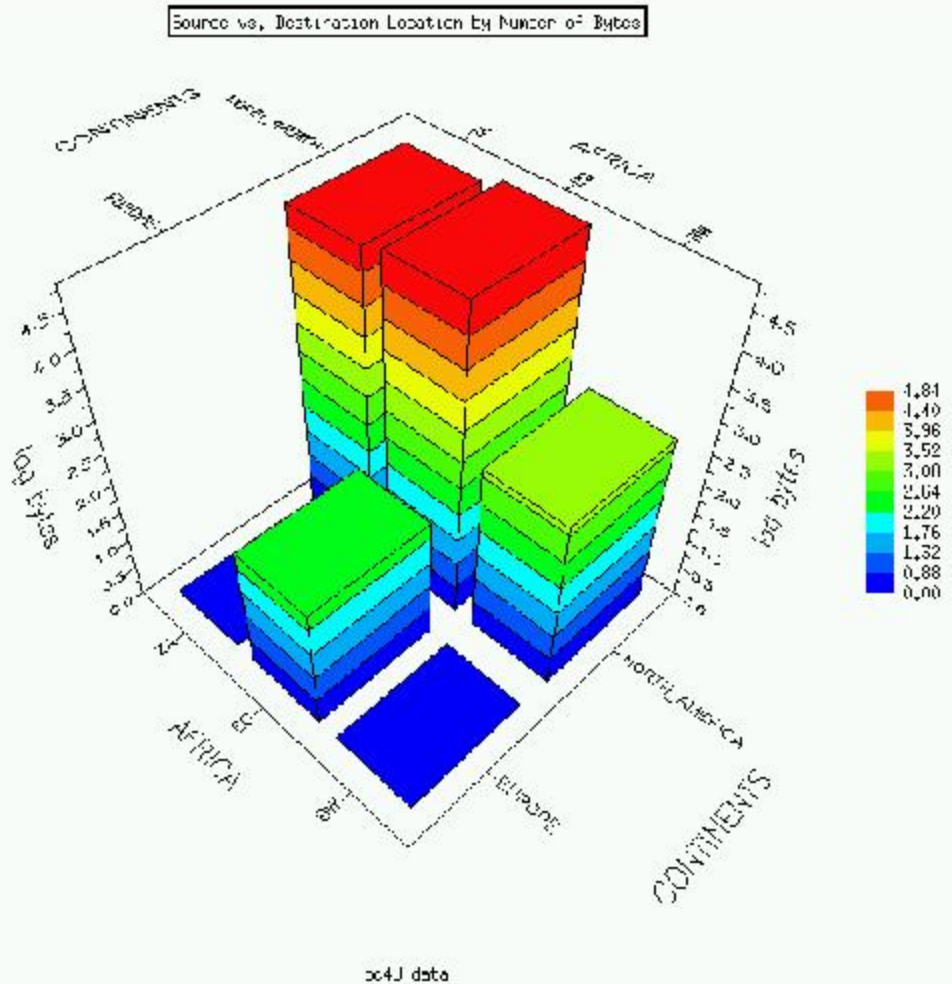
results: 3D traffic matrices



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

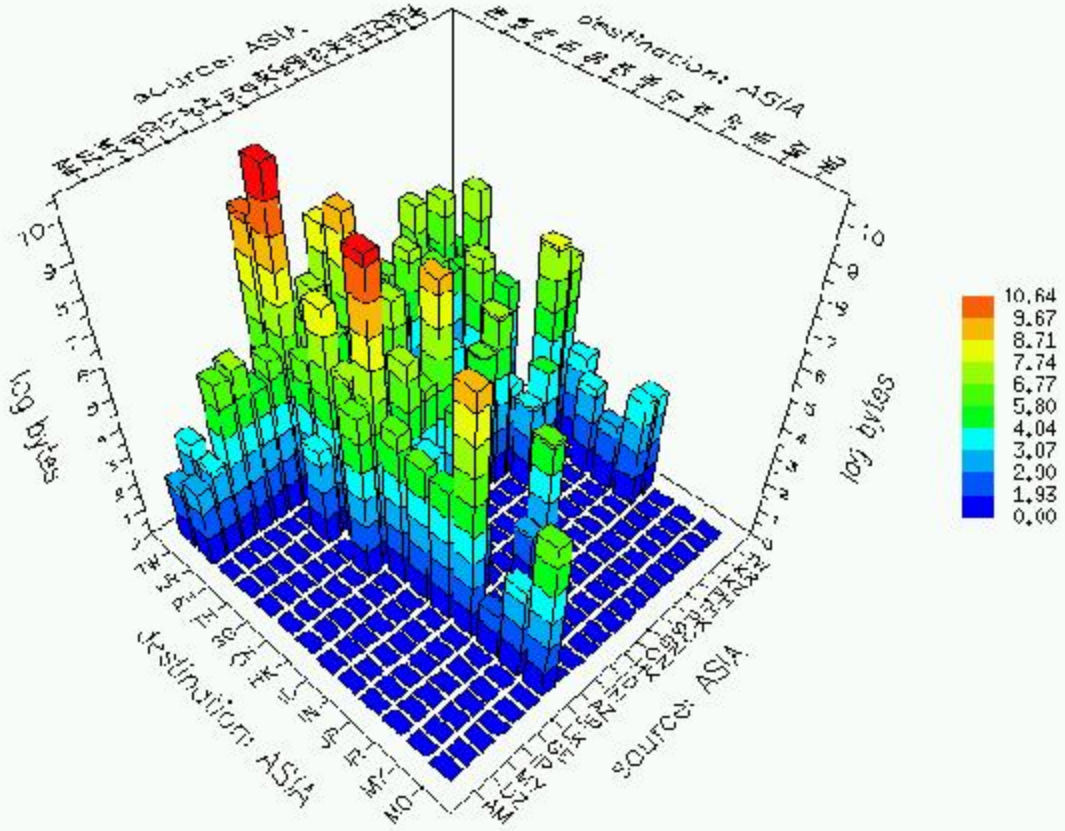
results: anomalies

we see the following unexpected traffic at our measurement locatio n 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)

c

gs

g applications

- traffic destinations
  - dominated by East Asia (KI
- significant traffic 'anoma
  - western europe – western E
  - eastern asia – eastern asia

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
  - currently no such project exists