

LANDER Datasets

John Heidemann

joint work with Joe Bannister, Gen Bartlett, Ramesh Govindan, Xinming He, Alefiya Hussain, Christos Papadopoulos (co-PI), Jim Pepin, Yuri Pryadkin
USC: ISI, CS Dept, ISD

27 September 2005



LANDER Dataset Status / 27 September 2005

1

LANDER Data Collection

- packet traces from Los Nettos
 - fully anonymized
 - host-only anonymized
- Internet topology scans
- Internet address-space scans
- phase 2: DNS (Domain Name System) data

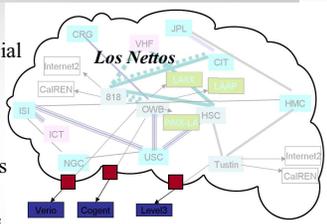


LANDER Dataset Status / 27 September 2005

2

Packet Traces from Los Nettos

- mix of academic, industrial, commercial traffic
- several **monitoring points**
- observing ~800Mb/s of data
 - for months at a time



LANDER Dataset Status / 27 September 2005

3

Current Packet Traces in PREDICT

- DoS_traces-20020629
 - sample of 3 anonymized denial-of-service attacks
- Bottleneck_traces-20041202
 - paired 5-minute traces each hour for 24 hours (anonymized)
- usc_lander_active_tracing_scrambled
 - anonymized packet traces
 - done only on demand
- uses
 - denial-of-service studies
 - background traffic in simulations
 - traffic evolution studies
- current traces use prefix-preserving anonymization
- phase 2 packet traces will optionally support host-only anonymization

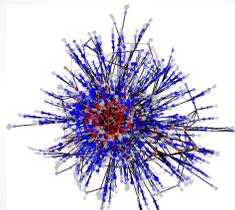


LANDER Dataset Status / 27 September 2005

4

Internet Topology Scans

- internet_router_map_p lanetlab-20030412
- router-level probes
 - includes alias resolution
- uses:
 - simulation parameterization
 - topology studies



[Cable and Wireless's network; from an earlier router scan]

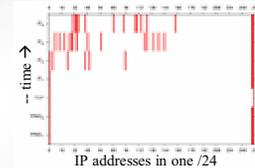


LANDER Dataset Status / 27 September 2005

5

Address Space Scans

- internet_address_survey_i cmp1-20030601 and 7 other scans
 - data taken 2003 through 2005
- pings to *all* routable addresses
- uses:
 - address allocation studies
 - worm propagation studies
 - etc.



ex: a subnet (/24) that was mostly unused, then transitioned to dhcp



LANDER Dataset Status / 27 September 2005

6

Other Project Progress

- software releases
 - DAG anonymization software
 - stand-alone spectral analysis tool
 - trace collection processing scripts (in progress)
 - see <http://www.isi.edu/ant/software/>
- writeup of tracing infrastructure presented at Minenet 2005
 - “Experiences with a Continuous Network Tracing Infrastructure”
- examined 10Gb/s link tracing
 - current costs: ~\$55k/link



LANDER Dataset Status / 27 September 2005

7

Phase 1 Status

- seeing lots of demand for traces
 - considering out-of-band distribution
 - ready for portal to be open
- finishing up first round of formal agreements
 - still pending round of recent changes from RTI



LANDER Dataset Status / 27 September 2005

8

Phase 2 Plans

- packet level traces
 - support for customized traces
 - support for host-only anonymization
 - *for both, need feedback from potential users*
- continuing address space scans (~1/month)
- DNS root server traffic
 - plan to include anonymized source address but query contents
 - pending approval of relevant parties



LANDER Dataset Status / 27 September 2005

9

Conclusions

- phase 1 goals completed
 - basic traces and stable infrastructure
- starting phase 2
 - customized anon. traces; anon DNS

<http://www.isi.edu/ant/lander/>



LANDER Dataset Status / 27 September 2005

10