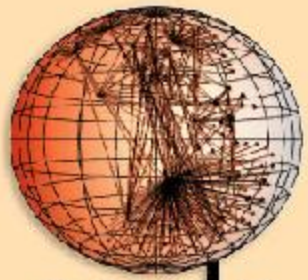


priorities and challenges in Internet measurement, simulation, and analysis

// somebody has to do something, and it's just incredibly pathetic that it has to be us. //
-- jerry garcia

9 january 2003
nsf pi meeting reston virginia
kc

Jan 2003 on Internet mailing list

- http://www.postel.org/pipermail/end2end-interest/2003-January/002720.html

I believe (based solely on a single long-ago observation that an M/M/1/K queuing model seemed to predict the measured behaviour of core routers with short buffers pretty accurately) that the speculation in your third sentence above may in fact be true. To relate this to another recent thread on this list, however, this seems like one of those things that should require no speculation since it is not so difficult to measure, yet I know of no good-quality data from "typical" core circuits which has been published anywhere. The network we've built is constructed with insufficient instrumentation to enable us to understand what it is we've built with any certainty, so we speculate.

---- dennis@juniper.net

Internet windmills needing data support

- topology, topography
  - connectivity and stability
- routing
  - dynamics
- performance
  - diagnostics
  - bandwidth estimation
  - congestion avoidance
  - qos
- workload characterization
  - traffic flow analysis & modeling
- security
  - anomaly detection
  - IP traceback
- macroscopic level measurement
  - dos attacks
  - dns system

- middleware layers
  - feedback to system at various granularities (dynamics and trends)
- new protocol development: routing, transport, IP
- feedback to future operational measurement methodologies

why is it so hard to do research w real data?

- **short answer: there hasn't been any since 1995**
- **long answer: there's way too much data floating around**
  - disadvantage: inappropriate data can be distracting or worse
  - advantage: publishing inappropriate data can incent people to offer you better data (`desperate times call for desperate measures' methodology)
- 2 outstanding talks about problems w Internet data
  - vern's talk aug2001 www.icir.org/vern/talks/vp-nrdm01.ps.gz
  - david's talk apr2002 www.caida.org/outreach/presentations/

two approaches to measurement (vern)

(1) research-driven

- for specific goal
- notion may be flawed
- may design data to support conclusion

(2) for-the-sake-of-measurement

- needed for longitudinal studies
- may foster serendipity
- often useless data, rife with errors
  (problem with public measurement repositories)

unsettling admissions about dealing w data

(courtesy vern paxson, david moore)

- bizarre behavior, misconfigurations, non-RFC, attacks, `impossible' behavior
- measurement tools lie
  - **(packet filters drop, reorder, replicate, miss due to routing)**
  - **clocks can be arbitrarily off/moving, timestamps don't know accuracy, applied differently**
  - **app-level measurement tools miss hidden network stuff (middleware, socket buffer parameters)**
- asymmetric paths
- measurements made 2 different ways always disagree (anisotropic)
- even a single measurement may disagree (not atomic): routing tables, traceroute
- events ripple through network along trajectory that is unlikely fully instrumented
- measurements carry no indication of quality
- measurements lack meta-info (e.g., hostnames)
- representative data points - there is no typical on the Internet
- analysis results not reproducible
- lack a culture of calibration
- large-scale measurements required for repr/longitudinal analysis overwhelm our current methods
  - **archived data often ad hoc, corrupt, truncated, poorly documented. unnavigable**
  - **lack of historical data renders it difficult to assess trends**
  - **alas, people do it anyway, see kc's myths talk (or any trade rag)**
- Internet measurement, although too hard, is too easy
  - **not enough data and too much data**
- we don't yet know how to measure real traffic in the core
  - **speed, sampling, anonymization**
- can't keep up with media in core (oc12 monitor arrives right after upgrade to oc48)

intra-scientist irreproducibility

missing: systematic approach to reduction/analysis

- e.g., paper trail for analysis train, especially for bugs

what we don't yet require of ourselves and each other:

- keep master script of analysis results
- keep intermediate forms of data
- keep notebook of what was done
- version control for scripts and notebook
- ways to visualize what's changed in analysis results after re-run

(recently funded meta-repository project will try to faciliate the cultivation of a culture of calibration. long road.)

in spite of it all, amazing stuff has emerged

- topology mapping/inference: caida/ucsd, uw, and icir
- simulation: ns (icir,isi...), renesys bgp/ssfnet
- bgp/security measurement and analysis: renesys
- tcp analysis: icir, cambridge (frank kelly, glenn vinnicombe), ucla (paganini), caltech
- dns: gatech (ewz), mit (hari), caida
- passive measurement/analysis: sprint, icir, caida, att
  - **dragonflies/tortoises, self-similarity**
- macroscopic analysis: dns roots, dos attacks
- invariance of rtt spectrum
- evolution of bgp system at AS, prefix, and IP granularities
  - **introduce semi-global prefix and related taxonomy**
- dispel myths of bgp growth and churn (rates and contributors)
- route dampening considered harmful
  - **using bgp beacon (unused prefix announced and withdrawn at well-known times)**
- lots about topology/policy/event inferences from bgp data
  - **patrignani, gao, schulzrinne, nicol, feamster**
- bgp convergence and scalability (anja&olaf)
- udp/tcp performance during bgp update activity (avi)
- comparison of routeviews and ripe data for bgp analysis (agilent, caida)
  - **data sets seem congruent, both need improvement/coordination**
- traffic modeling

amazing++: ISPs even sort of care now

(can't say the recession hasn't helped raise awareness of the value of data)

example nanog meeting, oct 2002 *(see nanog.org videoarchives)*

- route dampening considered harmful (again)
- damage in the DNS system
- bgp behavior under stress (lixia et al)
  - http://www.cs.ucla.edu/~lanw/paper/imw02-bgp.ps
- bgp enhancements to prevent persistent route oscillations
  - http://www.nanog.org/mtg-0210/ppt/sue.pdf
- feds want to secure cyberspace (office of cyberspace security)
  - http://www.whitehouse.gov/pcipb/
  - `how to own the Internet'
- panels on measurement, complexity/robustness spiral
  - http://www.nanog.org/mtg-0210/complexity.html
  - http://www.nanog.org/mtg-0210/measurement.html
  - raised more questions than answered (<- a good thing)
- scriptroute: 'public measurement facility' proposal
  - www.scriptroute.org
- now specifically solicit proposals from research community

nanog (oct2002): ISP/researcher interaction

bgp enhancements to prevent persistent route oscillations

- major factors for route oscillation:
  - dependency of IBGP updates
    - **sometimes circular dependency**
  - partial information by RR or confederation
    - **withdraws (over reduction) amplifies the issue**
  - partial order (due to MED) in route selection
- suggestions: modify route reflection spec, allow advertisement of multiple paths

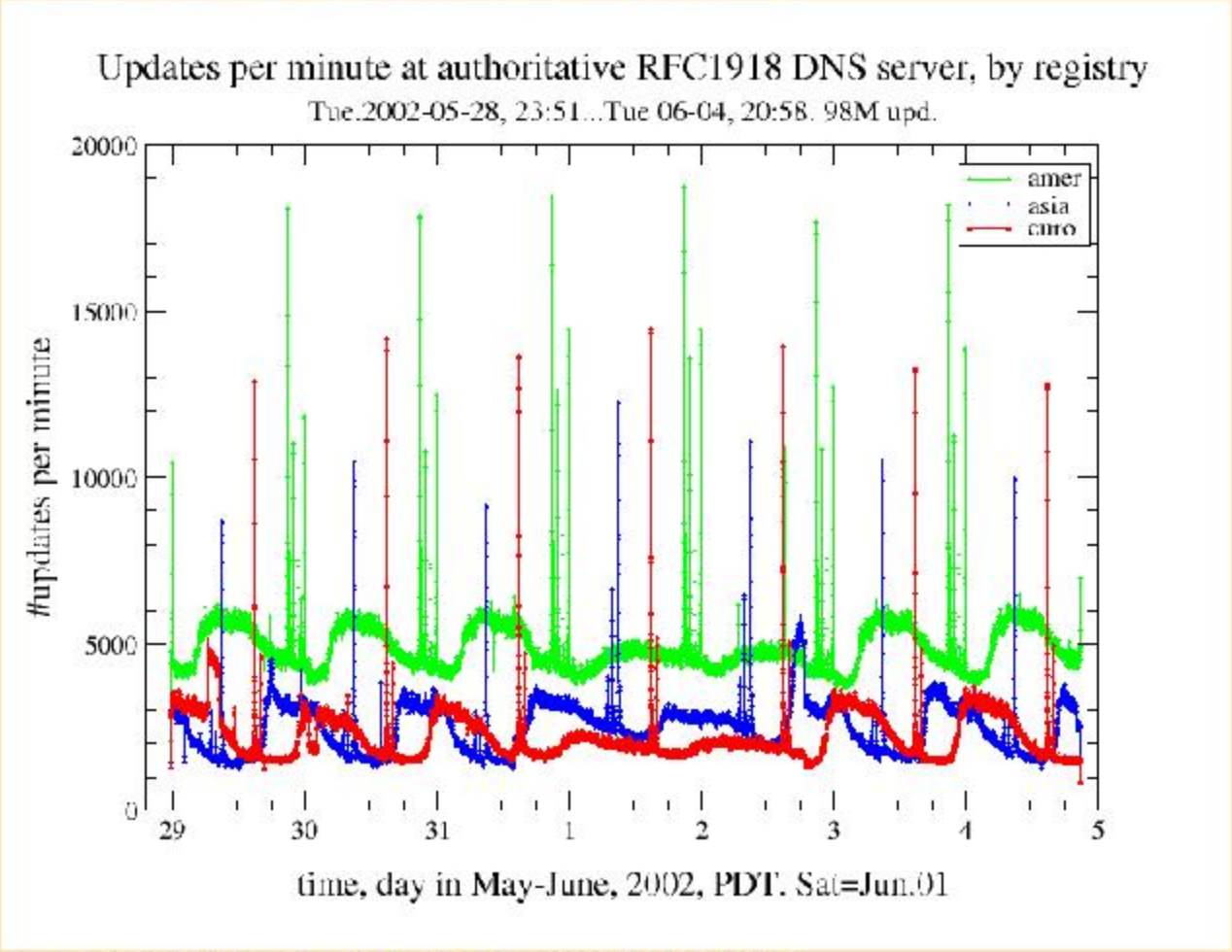
issues:

- non-deterministic routing can be demonstrated within full mesh topology
- MEDs prevent reasoning about routing policies
  - though existing references encourage use of MEDs to influence inbound policies
- how much are we facing limitations of distance vector protocol and inherent limitations of same
- can't tell what tie-breakers cause route selection..

vendor example: measurement reveals bugs

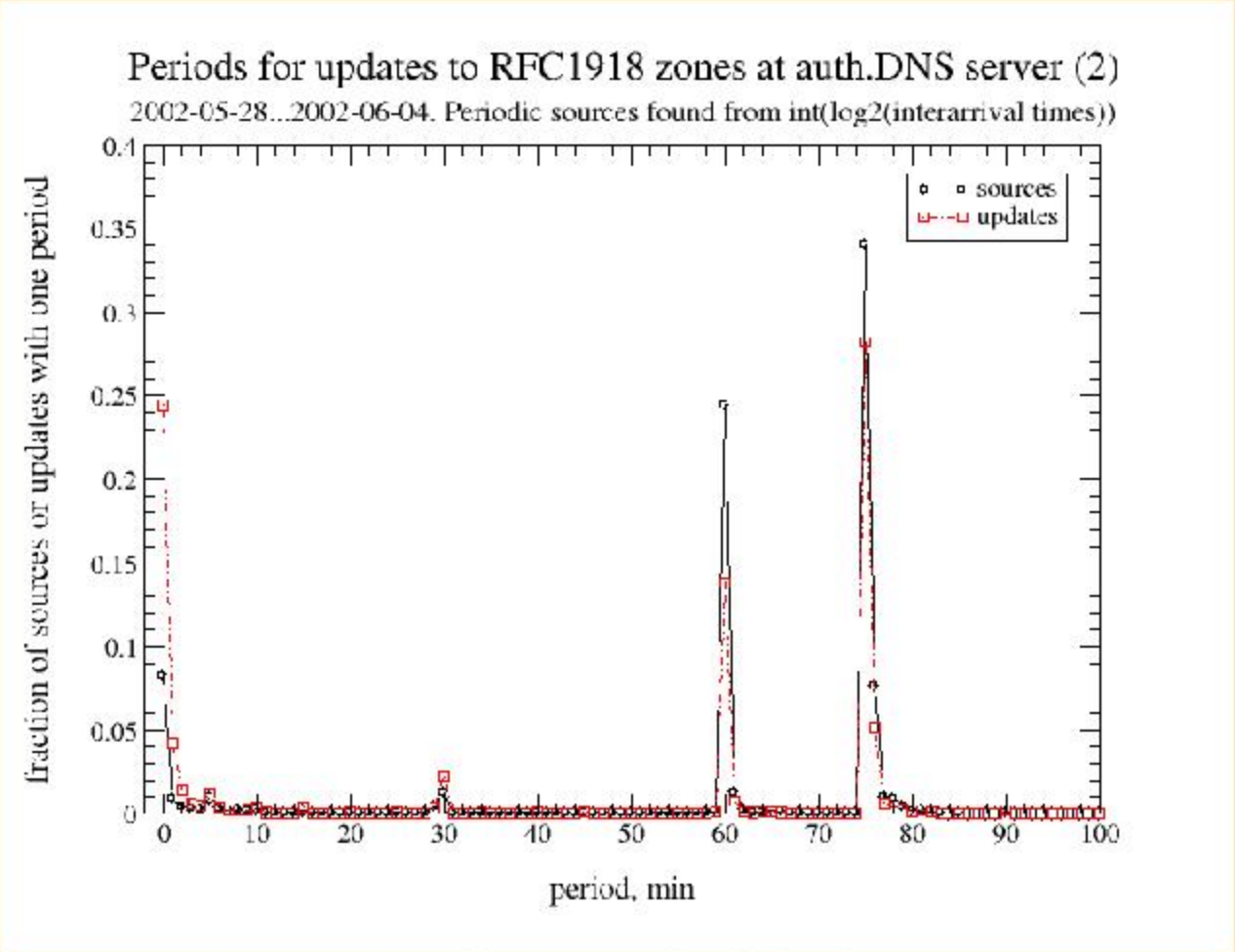
dns updates for private address space leaking up to roots

- spectroscopy analysis of RFC1918 updates
- RFC1918 updates coming from DHCP/nameservers
- millions a day getting to root name servers (whee)
  - **51.4M updates in 86.5 hr = 10,000 per minute = 165 per second**



  - **weekday, weekend patterns; weird spikes at midnight local time**

... global RFC1918 damage in DNS system



- rare to get macroscopic Internet data so radically broken
- who is trying to update the roots anyway?
  - **dsl, cablemodem, small population providers, developing countries**
- verified that vast majority derive from two OSes: Windows 2000 and Windows XP
  - **majority of updates from sources that send them constantly**
  - **bulk of workload from contributions of medium size, not mice/elephants**
  - **most source IP addresses are of home and small business users (owned by individuals, not organizations)**

...global threat arising from single vendor

- combination of Microsoft software features & misconfigurations essentially causing a slowly paced massive distributed denial of service (DDOS) attack on the root name server system
- current state of fielded desktop software poses substantial & increasing burden on (if not threat to) the robustness of the global Internet
- software and setups affecting global systemic Internet stability must be designed more carefully wrt potential effects of:
  - software engineering decisions
  - misimplementations
  - misconfigurations
- measurement can make a huge difference

optimism (no, really)

still the case: security, performance, configuration, and fault management lack both effective solutions, as well as an apparent lack of people able to state a concise problem to be solved.

[ok so that wasn't the optimistic part...]

researchers do have a valuable role to play
(but we have to get out of our labs)

measurement needs (logistics)

- performance
  - e2e measurements must be relevant to the user
  - bandwidth estimation tools
    - single source vs two end point
    - available bw vs capacity
    - calibration against real paths w cross-traffic (requires link and other info from multiple providers)
    - infrastructural heterogeneity rendering harder by the minute
- workload
  - longer traces
    - tcp/ip headers, w options, other headers if possible
    - at least 24 hours
    - concurrent at several places
    - several across time for trend analysis
    - (buy stock in RAID vendors now)
  - traffic matrices
    - same constraints as above
  - diverse locations
    - disa, enterprise, university, backbone, peering point, cdn's (e.g., akamai)
- topology
  - large scale traceroute coverage
  - overcoming problems in traceroute methodology
  - tracking topology at other layers
  - IPv6

measurement needs (logistics)

- routing
  - strategically designed interdomain routing data collection
    - **constant, sufficiently large, diverse, and representative set of peers**
    - **stability across time**
  - IGP updates (requires cooperation/release from ISPs)
  - configured software parameters from ISPs, e.g MRAI
  - other implementation details from vendors
- security
  - anomaly detection, traceback
- wireless
  - headers, location, signal strength
  - needs further requirements analysis
- sensor
  - johnh's going to cover
- macroscopic monitoring
  - [how to do] macroscopic monitoring of DNS, BGP system
  - network telescopes for dos attack tracking
- correlation
  - need repository support

also needed: analyses of causes of damage

no rigorous study exists of root causes of Internet performance problems/outages. anecdotal survey (courtesy sean donelan nanog post):

1. network engineers (what's this command do?)
2. power failures (what's this switch do?)
3. cable cuts (backhoes, enough said)
4. hardware failures (what's that smell?)
5. congestion (more bandwidth! Captain, I'm giving you all she's got!)
6. attacks (malicious, you know who you are)
7. software bugs (your call is very important to us....)

// *``I prefer the wicked rather than the foolish.*
the wicked sometimes rest." - alexandre dumas //

panelist questions

(1) how would you characterize past progress and current efforts to characterize, understand or model behaviors of large-scale systems, such as the Internet, where the artifacts being studied are complex and require more detailed understanding of traffic behavior?

(2) it has been argued that `good data outlives bad theory'. what are impediments to acquiring a representative set of data points that capture Internet behavior in some significant way and what long-term vision is needed to overcome these impediments?

(3) the challenge of scale, coupled with the lack of configuration-specific knowledge, calls for a more precise formulation of the measurement challenge and associated research directions. what role would the use of inference techniques and simulation play in yielding insights into behavior of current or future Internet protocols (transport, routing, etc)? what type of infrastructure must be available for resarchers not only to address current questions but new questions that cannot be adequately addressed by current tools?

action items for research community

- culture of & passion for sound measurement, as science & discipline
  - measurements need pedigrees describing them, how to navigate
  - audit trails, portable analysis scripting language to support reproducibility
  - well-managed meta-data
  - understand sampling implications and technology better
  - anonymization tools & reduction agents
- simulation
  - needs way to calibrate against real data (broken record)
  - safety tip: still has mostly no respect from providers
- analysis
  - find ways to assess opex versus capex of any new idea
    - **or at least don't render it impossible to do so later**
  - as scott encourages 'let's try to formalize some of what we (and providers) know'
  - as dave clark encourages 'it's about the $$$, stupid'
  - safety tip: providers do lose their patience w this research @#$%
- continued/increased interaction with providers and vendors
  - nanog, ietf
  - switch, router, measurement hardware vendors

take advantage of the industry's regrouping efforts and inclination to listen. who knows how long it will last.

action items for funding/publishing agents

- willingness to invest $$$$ in measurement and databases
  - software, storage, communications for collaboration, data distribution
  - frustrating [perception] that it takes funding away from `real' research
  - encourage tech transfer from other disciplines who have done it
- willingness to bounce manuscripts that don't include raw data + scripts
- willingness to publish `mundane' work on measurement management
- emphasis on reproducible results

- encourage research projects that involve provider/vendor cooperation
- encourage researchers to attend industry meetings (nanog, ietf)

good news: there has never been a better time for this

- overheard 'the recession is stay of execution for the bgp routing system'

challenges in Internet measurement

results of breakout session on 11am fri 10 jan 2003

(next 5 slides)

challenges in Internet measurement

motivating vision: self-aware network

- cultivate culture of sound measurement as science & discipline
  - measurements need pedigrees describing them, how to navigate
  - audit trails, portable analysis scripting language to support reproducibility
  - well-managed meta-data
  - understand sampling implications and technology better
  - anonymization tools & reduction agents
- more strategic measurement, guided by rather than constraining research questions
  - what data is missing and how do we strategically optimize the return on investment in data collection and instrumentation
  - recognizing that we don't always know what questions will be asked next year
    - **routing, topology, passive, active**
  - improved standardized interface to data archives

challenges in Internet simulation & analysis

- mathematical frameworks to find structure/patterns in traffic
  - a la scott's encouragement to `formalize some of what we (and providers) know'
  - macroscopic as well as microscopic
  - theory of joing spatial/temporal locality
  - spectroscopy, tomography
  - ietf/ippm has been trying for a few years, but without dedicated funding
- source modeling (for realistic inputs into simulations, models)
  - extract a set of source models from an aggregate trace
    - **feature extraction problem**
    - **10,000 gnutella port numbers are not 10,000 flows**
  - ultimate goal: augment libraries of source level models w generation of own
  - calibrate models by evaluating their power for prediction
- empirically validated simulation of a significant aspect of the Internet
  - already much work in large-scale simulations, but no recognized empirically validated simulation of any signficant piece of the Internet.
  - requires cooperation from providers and vendors to get default and configured parameters of OSes and algorithms. NSF should shepard/foster this cooperation
    - **(note: large scale means in size as well as # of protocols)**

challenges in Internet meas. & analysis

don't forget the real world

- analyses must incorporate expense ratios (opex/capex) into tradeoffs where possible
  - need to get/keep relevant to providers
- systematic studies of outages
  - assessment of various causes of damage
- tools and techniques to tie user-perceived performance w/ network measurement, with statistically significant results
  - function of control plane, routing plane, server on other end.
    - **including bugs in routers, servers, client stacks**
  - correlating user experience with events that happen in the network
  - incorporate into network performance models
  - validation of tools still doesn't exist
  - use at truly large scale still doesn't exist
  - dave clark's `why?' tool (non network geek compatible) still doesn't exist

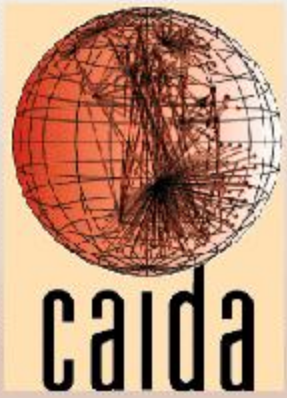
other challenges in Internet measurement

- discovering pervasive hidden bugs
  - any modeling or analylsis must also handle the impact of this huge component of traffic
- how does measurement affect/support security goals
  - infer bgp, firewall, and virus spread behavior
  - how do you get networks to share security-related information
  - protection of measurement infrastructure from security compromises
- measurement specific to optical, wireless and sensor networks
  - especially assessment of new application domains (for wireless/sensor)
- encouragement of strategic measurement in new networks
  - based on what we learned from what we did wrong in old networks

payoffs to Internet measurement

- improve accuracy, validity, repeatability of network research
- provide reference points or baselines for simulation and model validation
  - other fields, e.g., architecture, have had this for years
- build a solid understanding of network behavior
  - including subtleties not otherwise detected
  - including damage not otherwise detected
- accelerate present and future modeling, simulation, and analysis efforts
  - avoid duplication effort

// scientific apparatus offers a window to knowledge,
but as they grow more elaborate,
scientists spend ever more time washing the windows.
-- Isaac Asimov //



// disorder increases with time
because we measure time in the
direction in which disorder increases //
-- stephen hawking

www.caida.org