Hashes, Caches, Puzzles, and the IXP

Wu-chang Feng

Damien Berger
Abdelmajid Bezzaz
Francis Chang
Jin Choi
Wu-chi Feng
Ashvin Goel
Kang Li
Antoine Luu
Deepa Srinivasan
Jonathan Walpole

OGI SCHOOL OF SCIENCE & ENGINEERING

OREGON HEALTH & SCIENCE UNIVERSITY

Outline

- A quick tour of our work on...
  - Mapping Bit Vector onto the IXP
  - Exact packet classification cache architectures
  - Approximate packet classification caches
  - TCPivo: High-performance packet replay
  - IXP networking practicum course
- Followed by...
  - The Case for IP Puzzles

Packet classification algorithm mapping

- Motivation
  - Packet classification is an inherent function of network devices
    - Many algorithms for single-threaded software execution
    - Many hardware-specific algorithms
    - Not a lot for programmable multi-processors
- Our study
  - Examine algorithmic mapping of a hardware algorithm (BitVector) onto the IXP
    - Pipelined (4 dimensions on 3 μ-engines, 1 combo, 1 ingress, 1 egress)
    - Parallel (complete lookup on 4 μ-engines, 1 ingress, 1 egress)

Packet classification algorithm mapping

- Initial results
  - Hard to generalize
    - Depends on workload, rulesets, implementation
  - Trie lookups bad for μ-engine health
    - Frequently forced into aborted state due to branching
      - Linear search: ~10-11%,
      - Pipelined Bit-Vector: ~17%
      - Parallel Bit-Vector: ~22%
    - Impacts device predictability and algorithm/compiler design
      - Avoid branches, utilize range-matching?
  - Memory bottleneck favors parallel over pipelined in IXP1200
    - Pipelined slightly worse than parallel due to multiple header parsing
    - Will change with IXP2xxx next-neighbor registers

Deepa Srinivasan, "Performance Analysis of Packet Classification Algorithms on Network Processors", OGI MS Thesis, May 2003 (submission planned)

Exact Packet Classification Caching

- Motivation
  - Caching essential for good performance
  - Impacted by traffic and address mix
  - Recent work on analyzing..
    - Internet address allocation
    - Traffic characteristics of emerging applications such as games and multimedia
- Our study
  - How does recent work impact design of caches?
    - Hash function employed in cache (IXP hash unit vs. XOR)
    - Replacement policies (LFU vs. LRU)

Exact Packet Classification Caching

- Initial results
  - Address allocation policies allow μ-engine based XOR-hashes to outperform stronger hashes (i.e. centralized IXP hash unit)
  - LFU provides only marginal improvement over LRU with multimedia traffic

Kang Li, Francis Chang, Damien Berger, Wu-chang Feng, "Architectures for Packet Classification Caching", *in Proceedings of International Conference on Networks*, Sept. 2003.

Approximate Packet Classification Caching

- Motivation
  - Large # of fields and large headers
    - Forcing caches to grow (and become slow)
  - Reducing entries degrades performance
    - Classic space-time trade-off in cache performance
- Our study
  - Throw a wrench into space-time trade-off
  - Examine another axis: *accuracy*
    - Quantify the space-time benefits of reducing cache accuracy
    - Understand the implications of using network devices that are not always “correct”
  - Similar to Intel's probabilistic computing
    - See recent interviews from Borkar, Tennenhouse

Approximate Packet Classification Caching

- Results
  - Order of magnitude space savings for an error rate of one in a billion
  - Analytical model for controlling misclassification rate
  - Poster outside....

Francis Chang, Kang Li, Wu-chang Feng, "Approximate Caches for Packet Classification", *in ACM SIGCOMM 2003 Poster Session,* Aug. 2003. **Poster**

Francis Chang, Kang Li, Wu-chang Feng, "Approximate Caches for Packet Classification", *in submission*. **Paper**

TCPivo: High-Performance Packet Replay

- Motivation
  - Require accurate, high-performance packet replay with IP addresses intact to evaluate network devices
  - Must be cheap (commodity hardware, open-source software)
- TCPivo
  - Linux x86-based tool for accurate replay above OC-3
    - Trace management
    - Timer management
    - Low transmission overhead
    - Proper scheduling and pre-emption
  - Software available, poster outside...

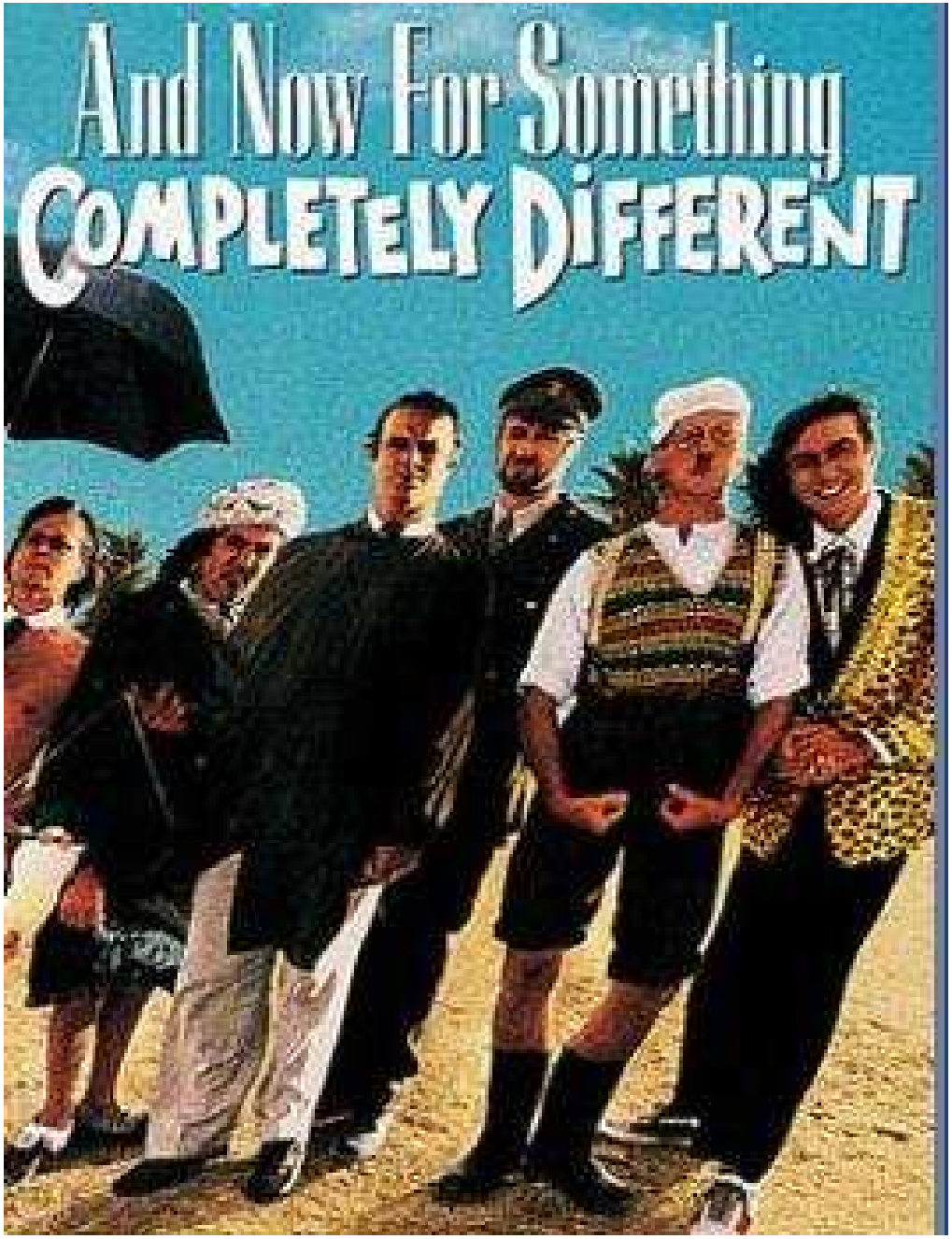
Wu-chang Feng, Ashvin Goel, Abdelmajid Bezzaz, Wu-chi Feng, Jonathan Walpole, "TCPivo: A High-Performance Packet Replay Engine", *in Proceedings of ACM SIGCOMM Workshop on Models, Methods, and Tools for Reproducible Network Research (MoMeTools)* August 2003.

IXP Network Practicum course

- Contents
  - 10-week quarter, 3-hour laboratory per week
  - Basics of the ACE framework
  - Command-line development
    - µ-code assembler development of microblock components
    - C development of core components
  - Workbench development
    - µ-engine C development of microblock components
    - IXP simulator
- Projects and assignments
  - Packet and protocol counters
  - Load-balancing switches, FragRouter, Token bucket markers
  - Content filters (SMTP viagra reset, WWW bomb reset)

IXP Network Practicum course

- Spring 2003
  - Instructor: Me
  - TA: Francis Chang
  - Enrollment: 16 fairly satisfied students (3.4 out of 4.0)
    - Limited by our hardware resources
    - Forced to turn back students
    - 12 out of 16 with `intel.com` e-mail addresses
  - Several external inquiries for student list
    - Feeling very conflicted about this...
- Fall 2003
  - Offering course again to those unable to take it
  - Instructor: Francis Chang
  - Will start in two weeks....

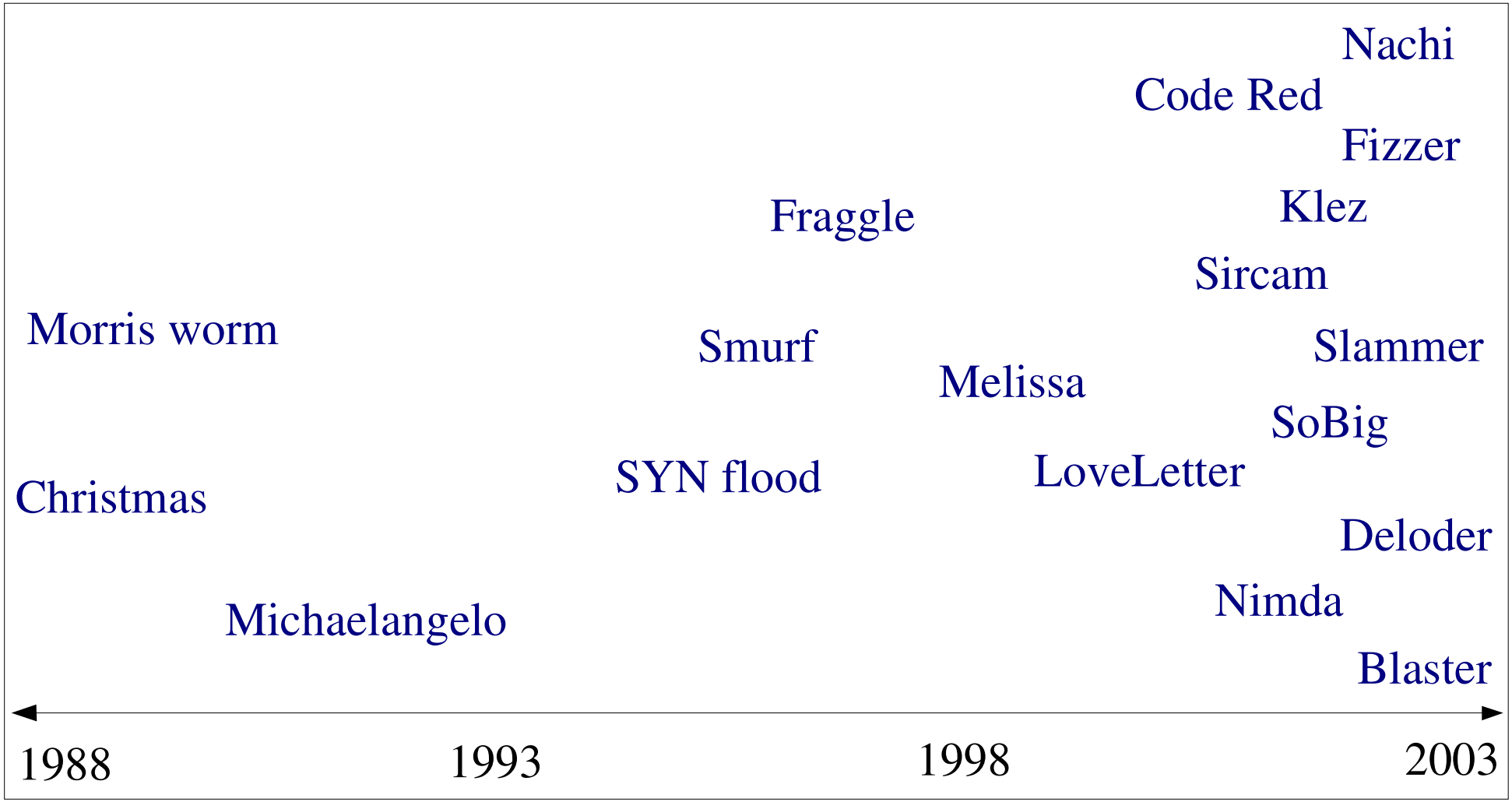

And Now For Something
COMPLETELY DIFFERENT

The Case for IP Puzzles

Motivation

- A quick look back on 15 years of not so “Good Times”

  SMTP, TCP, ICMP, UDP, FastTrack, SMB, finger, SSL, SQL, etc.



Puzzles

- An interesting approach for mitigating DoS activity...
  - Force client to solve a problem before giving service
  - Currently for e-mail, authentication protocols, transport layers
  - Fundamentally changes the Internet's service paradigm
    - Clients no longer have a free lunch
    - Clients have a system performance incentive to behave
- A contrast in approaches
  - Leave doors open and unlocked, rely on police/ISPs
    - Centralized enforcement (not working)
  - Give everyone guns to shoot each other with
    - Distributed enforcement (may not work either)
    - Harness the infinite energy of the global community to fight problem
    - Promising anecdotal evidence with spamming the spammers...

Posit

- Puzzles can only be effective if placed at the IP layer

Why are IP puzzles a good idea?

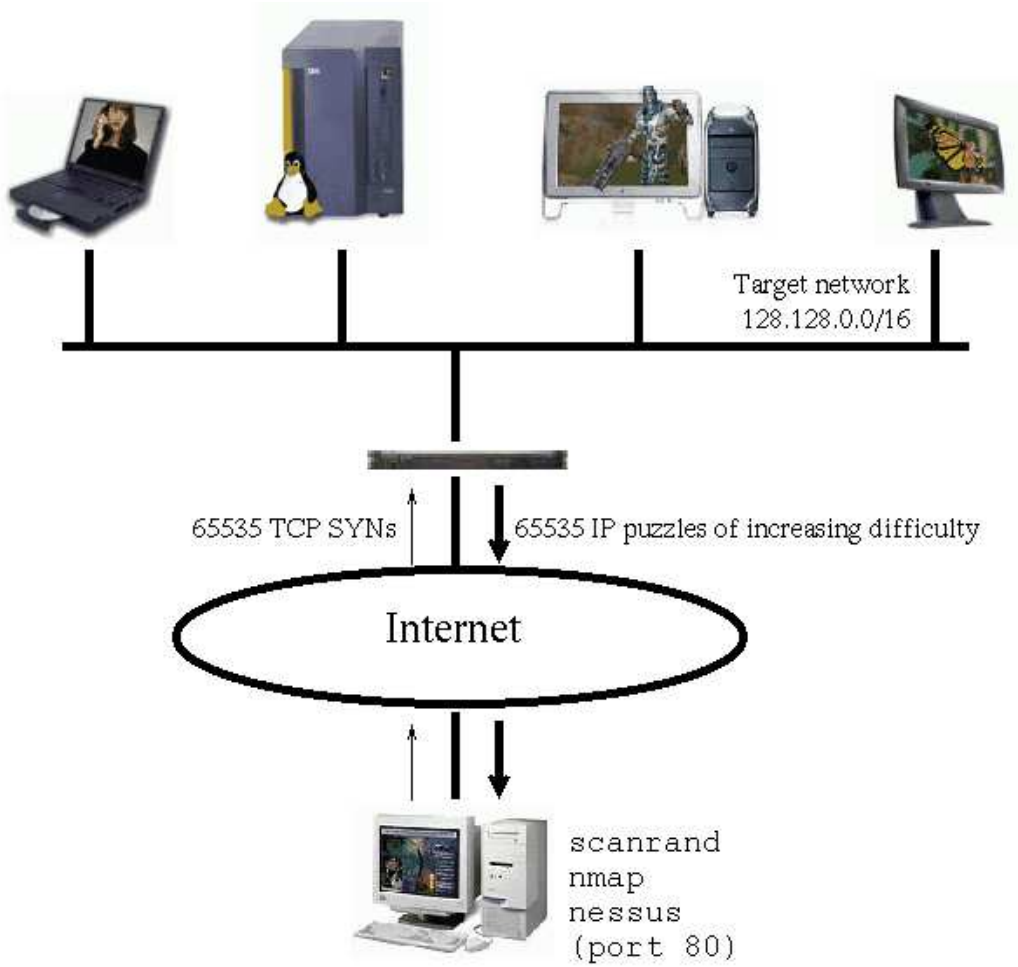
- "Weakest link" corollary to e2e/waistline principles
  - Put in the common waistline layer functions whose properties are otherwise destroyed unless implemented universally across a higher and/or lower layer
  - DoS prevention, congestion control destroyed if any adjacent or underlying layer does not implement it
    - TCP congestion control thwarted by UDP flooding
    - DoS-resistant authentication protocols thwarted by IP flooding
  - Until puzzles are in IP, it will remain one of the weakest links

IP puzzle scenario #1

- Port and machine scanning
  - Instrumental to hackers and worms for discovering vulnerable systems
  - The nuclear weapon: `scanrand`
    - Inverse SYN cookies and a single socket
    - Statelessly scan large networks in seconds
      - 8300 web servers discovered within a class B in 4 seconds
    - Technique not used in any worm....yet
      - Forget Warhol
      - “American Pie” worm => done in 15 seconds?
      - Finally, a grand networking challenge!

IP puzzle scenario #1

- Mitigation via a “push-back” puzzle firewall



Why are IP puzzles a bad idea? (What are the research challenges?)

- Tamper-resistance
- Performance
- Control
- Fairness
- Deployment

Tamper-resistance

- A tool to both prevent and initiate DoS attacks
  - Disable a client by...
    - Spoofing bogus puzzle questions to it
    - Spoofing its traffic to unfairly trigger puzzles against it
  - Disable a router or server by...
    - Forcing it to issue loads of puzzles
    - Forcing it to verify loads of bogus puzzle answers
    - Replaying puzzle answers at high-speed
  - Probably many more....

Performance

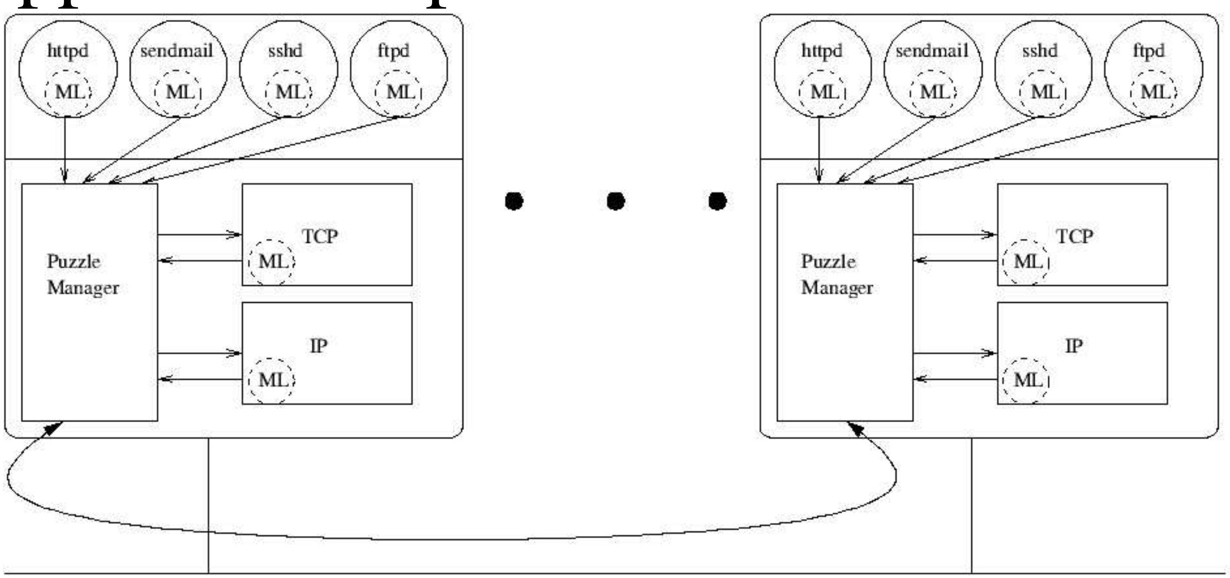
- Must support low-latency, high-throughput operation
  - Must not add latency for applications such as on-line games
  - Must support high-speed transfers
- At what granularity should puzzles be applied?
  - Per byte(s)?
  - Per packet(s)?
  - Per flow(s)?
  - Per flow aggregate?
  - Driven by performance and level of protection required

Control

- Puzzles require control algorithms to maintain high utilization and low loss
  - Mandatory, multi-resolution ECN signals that can be given at any time granularity
  - Can apply ideas from TCP/AQM control
    - Adapt puzzle difficulty within network based on load
    - Adapt end-host response to maximize throughput while minimizing system resource consumption (natural game theoretic operation)
  - Hypothesis
    - Easier to design puzzle controllers versus those used in TCP/AQM

Fairness

- Mechanism in a layer with minimal information
  - Can support bandwidth-based puzzle delivery
  - Can support some differentiation to deter Smurf/Fraggle
- But.. 202.183.197.116 - - [02/Jun/2003:02:08:29 -0700] "GET /default.ida?XXXXXXXXXXXXX XXX%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%u8190%u00c3%u0003%u8b00%u531b%u53ff%u0078%u0000%u00=a HTTP/1.0" 404 306 "-" "-"
- Need a "puzzle manager" to drive IP-layer puzzle policy based on application input



Fairness

- Enables "Reputation-based networking"
  - Software vendors
    - Making "trustworthy computing" mandatory (not marketing)
    - Long-term, computational tax for poorly designed software
  - System administrators and IT practices
    - Making responsible system management mandatory
    - Disturbing pervading notion: "cheaper to leave infected than patch"
    - Long-term, computational tax on poorly administered systems
  - End-users
    - Making users choose more secure software and adopt better practices
    - Punish users behaving "badly"
    - Long-term, computational tax on ignorance and maliciousness
  - "Nothing is certain but death and taxes." - B. Franklin

Deployment

- Can be transparently and incrementally deployed via puzzle firewalls/proxies
- Application-driven puzzle manager requires more intrusive changes
- Financial incentive to change is present
  - Lost productivity (see last several weeks)
  - Lost revenue, services (WWW, power, ATM, etc.)
    - SoBig.* author laughing all the way to the bank (Grrrr....)
  - Change may need a kick from the government or industry?

Why is this good for Intel?

- Keeping the Internet healthy
- Drives a whole new market for faster CPUs
  - Make the incompetent, the lazy, and the malicious “pay” for use of the Internet
  - Computational tax for running insecure software paid directly to Intel
- Demand for a whole new class of network devices
  - Puzzle proxies and firewalls based on IXP network processors

Status

- `netfilter/iptables` implementation
  - Tamper-proof operation (must be along path to deny service)
  - Puzzle generation ~1μs
  - Puzzle verification ~1μs, constant amount of state
  - Fine-grained puzzle difficulty adjustment
  - 100,000 puzzles/sec on commodity hardware
  - 1Gbs+ for per-packet puzzles with MTU packets
  - Small packet overhead
    - Puzzle question ~40 bytes
    - Puzzle answer ~20 bytes
  - Puzzle proxy and puzzle firewall implemented
    - Can set up demo upon request
    - Can play puzzle-protected Counter-strike transparently

Where's the IXP implementation?

- Big issue: IXP1200 is not built for security
  - Pseudo-random number generator can be predicted
  - Internal hash unit cyptographically weak
- Have a very short wish-list of functions
  - IXP 2850? Overkill, but we'll take one...

20000 NW Walker Road

Beaverton, OR 97006

Wu-chang Feng, "The Case for TCP/IP Puzzles", *in Proceedings of ACM SIGCOMM Workshop on Future Directions in Network Architecture (FDNA-03)*

Wu-chang Feng, Antoine Luu, Wu-chi Feng, "Scalable Fine-Grained Control of Network Puzzles", *in submission*

Questions?

- Packet Classification

  http://www.cse.ogi.edu/sysl/projects/ixp

- TCPivo

  http://www.cse.ogi.edu/sysl/projects/tcpivo

- CSE 506: Networking Practicum material

  http://www.cse.ogi.edu/~wuchang

  http://www.cse.ogi.edu/~francis/cse506

- PuzzleNet and Reputation-based Networking

  http://www.cse.ogi.edu/sysl/projects/puzzles

Extra slides

Fairness

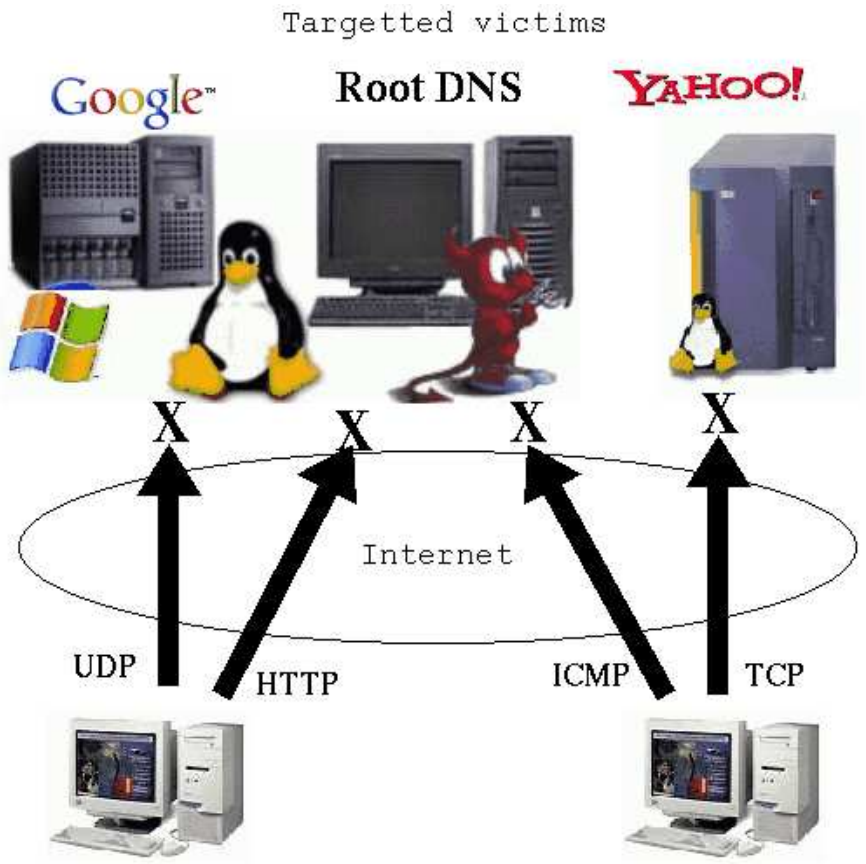
- Inserting a “trust” estimator into the knowledge plane
  - Answer the “WHO” question?
    - Who is a likely source of a future DoS attack?
  - No keys, no signatures, no centralized source
  - Based on time-varying distributed view of client behavior
  - Similar to GeoNetMap's “confidence” measure

IP puzzle scenario #2

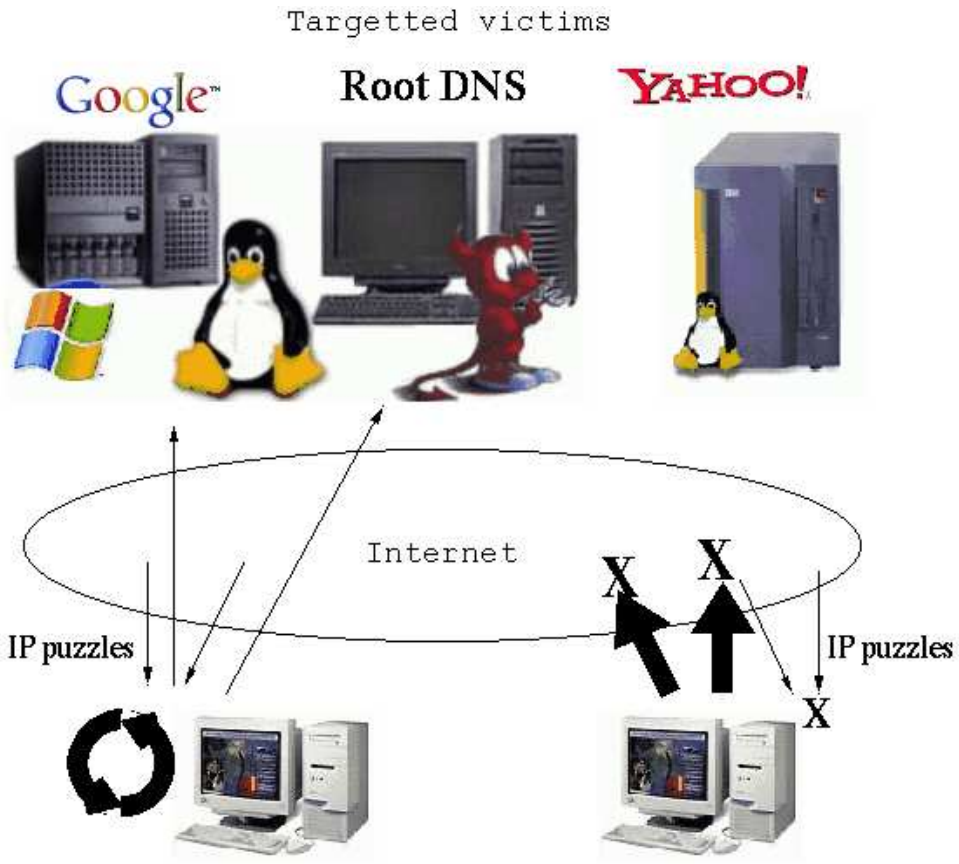
- Coordinated DDoS: simultaneous attacks against multiple sites from the same set of zombie machines
  - Mafiaboy (2000)
  - Have zombies initiate low bandwidth attacks on a diverse set of victims to evade localized detection techniques (such as `mod_dosevasive`)

IP puzzle scenario #2

- Mitigation using IP puzzles



Zombie participants in a coordinated DoS attack



Zombie participants in a coordinated DoS attack