The Internet Architecture: Its Future and Why it Matters

David Cheriton
Computer Science Department
Stanford University

Internet Architecture

- What: principles, protocols and structure for highly scalable digital communication
- Principles
  - Application state at endpoints
    - fate-sharing and otherwise soft net. state (D.Clark88)
  - One thin-waist (IP) for simple end-to-end connectivity
    - Multiple types of traffic
  - No off-path components
  - Liberal in what you receive; conservative in what you send

An amazing accomplishment

no thanks to me

Internet Architecture: what it provides

- Properties
  - Survivability: Intermediate nodes can crash and reboot w/o loss of application state
  - Simple to achieve connectivity for different applications
    - Just implement IP plus transport/application protocols
  - Have path, will communicate
  - Interoperability: not need for perfect implementation
- Applications build/rely on these properties

So, architecture provides properties,
but only if you are faithful to it

The Future

- Internet-enabled devices are everywhere
- Internet connectivity is ubiquitous
- Internet bandwidth is plentiful
- Special-purpose networks go extinct
  - No separate telephone, TV, SCADA networks
- All critical systems on the public Internet
  - Global financial systems
  - Power distribution systems
  - Air traffic control
  - . . .

Triumph: unified general-purpose communication
or is it: a disaster waiting to happen?

Air Traffic Control on the Public Internet!

Crazy? No, because there is no alternative:

- Internet technology: Ethernet switches, IP routers, etc.
  - Market: best, lowest-cost products
  - Staffing: good people know IP, etc.
- Public Internet: Really a collection of ISPs
  - Cost: lowest cost WAN connectivity
  - Availability: expert operators with lots of redundant connectivity

- But how about separation at layer 1?
  - Different colors for Internet ATC (I-ATC)
  - But where does the control plane for the optical run?
    - Single point of failure or public Internet?

I-ATC is inevitable! And frightening

The Internet Architecture: Why it matters?

The architecture allows us to state properties of the Internet and meet application requirements

- E.g. how to configure to meet I-ATC requirements?

If reality departs from architecture, properties are lost or unknown

- E.g. Ad hoc firewalling and NAT break end-to-end connectivity and reliability

If the architecture is wrong - can fail catastrophically

- The largest, most attractive “asset” to attack in the history of mankind

It matters too much to be ignored or wrong

Unfortunately, it is both

Ignored? Many violations of the architecture:

- What connectivity can a new wide-area Internet application assume?
  - Port 80 HTTP where the headers “look like” normal HTTP headers, going through NAT
    - Or maybe nothing because of DDoS, route flaps, etc.
  - No end-to-end addressing or reliability
- Dependences on off-path DNS server, root CA

Wrong?

- Current Internet could not work without the above

A New & Better Internet Architecture is required

Trust and Technologies

- New technologies develop, focused on improving features, performance and cost, however:
- The limit of most technologies is TRUST
- 250 MPH car: can build it, who do you trust to drive?
- Nuclear power plant: most efficient power but limited by trust in who builds and who operates
- GM Foods – we can grow them, will you eat?

Challenge: Internet architecture trusted to support critical infrastructure systems

Internet ATC Requirements

Very high availability, even under attack:

- Multiple disjoint paths between end-systems with fast fail-over
- Protection against DDoS
- Packet trace-ability – what source
- NOT Performance – low data rate
- NOT Confidentiality – in fact, open to be safe!

Other critical systems have same requirements

None supported by current architecture;

Oh, but … the work on Internet security!

You want security, I have a “solution”

It’s just that it:

- Has a single point of failure
- Is not testable
- Relies on negative acks, not positive acks
- Requires a costly complex implementation that is not understandable by most people
- Does not scale

Dead-on-arrival in the Internet community?

No, it just needs good “packaging”

The “Solution”: PKI Certificates

- Single point of failure
  - Loss of secrecy of private key of root CA
  - Flooding attacks
- Is not testable
  - No way to test if a key is secret
- Uses negative acks, not positive acks
  - Send out nacks in CRLs as part of revocation
- Costly complex implementation
  - PKE, signing, X.509, off-line CAs, CRLs, etc.
- Does not scale: off-line root CA for “security”

This is Internet security? I don’t feel secure!

Where did we go wrong

Dictionary: security == safety

- Security was hijacked to mean confidentiality
- Confidentiality was hijacked to mean encryption
  - Same for authentication
- Encryption only "understood" by cryptographers
- So, Internet security delegated to cryptographers
  - Cryptographers are algorithm designers
- Result: Standardized metaprotocols so poor interoperability, no safety, lots of overhead, single point of failure, no useful properties

Secrecy does not scale

A secure system needs a system design

You want e2e reliability, I have a “solution”

It’s just that it:

- Doesn’t provide end-to-end reliability
- Increases exposure to flooding DoS attacks
- Still a design-in-progress after 10 years
- Will take forever to deploy
- Hasn’t been evaluated relative to alternatives

Surely, a non-starter in the Internet community

No, just needs some good marketing, ardent followers and government mandates

The "Solution": IPv6

- No end-to-end reliability for named endpoints
  - Name-to-address binding can change w/ DHCP
- Exposure to flooding DoS attacks
  - Requires double forwarding/lookup bandwidth
- It is still a design-in-progress after 10 years
  - Addressing architecture, renumbering, mobility, flows
- It will take forever to deploy and makes things worse in the mean time – breaks IP thin waist
  - Upgrading 200 million hosts? IPv4<->IPv6 ?
- No evaluation of alternatives
  - Like change the transport checksum computation?

An enormous effort in the wrong direction

Where did we go wrong?

- Back in the 1970s - using IP addresses to identify end-system state
  - an IP address identifies an interface on host on particular (sub)network at a particular time
  - IPv6 – further ties it to a particular ISP's network
  - But state reachable by different interfaces/networks
- Again in the 1990's, by "believing" e2e IP addresses had some useful semantics

Reliability requires semantics;

IP addresses are transient routing tags, nothing more

You want routing: I have a “solution”

It’s just that:

- It depends on global trust and competence
- It must be operated at less than 1/1000th of real speed to be stable
- Forces you to upgrade your router as the Internet grows but provides you no benefit
- You have no control beyond first hop (and last I/F)

Surely, we would never implement . . .

wrong again!

The “Solution”: (secure) BGP

- global trust and competence
  - Shared world model: believe updates from your peers
  - Signed updates so you can “trust” your peers
- Operated at 1/1000th of real speed for stability
  - 30 second damping to avoid oscillations
- Non-scalable cost
  - Every router stores/recomputes all routes after updates
- You have no control beyond first hop
  - Source routing is disabled by ISPs

A large open loop dynamic control system

Defying conventional engineering or …?

Internet WAN Traffic Load

- Total WWW bandwidth, world-wide
  - P. Danzig 2000 estimate: 250 Gbps!
  - P. Danzig 2003 estimate: 250 Gbps!!
  - WWW is half of internet traffic
    - P2P “file sharing” and spam is the rest
- 1/2 single terabit router for entire known universe
- Not an issue even if larger by factor of 10 or more
- Moreover
  - 10 G Ethernet coming down in price
  - lots of dark fiber

Wide-area bandwidth is not the problem

wide-area business models are

This is all very depressing for I-ATC

- The Internet architecture is wrong
- The new developments do not address issues
- Research is focused “elsewhere”
- Critical systems will fail with terrible consequences when a massive Internet failure happens

Can we avoid disaster?

Let’s reboot

Cisco: How to sell a router

- Early days of Cisco: how to get someone to buy a router?
  - Already had connectivity
  - International Ethernets
- Selling point: routers limit Ethernet broadcast storms
  - STP loops, misconfigs would bring down the whole wide-area Ethernet
  - You don’t need a router to forward packets
  - You need it to (selectively) not forward packets

The router as a point of network control

Routing as a Broadcast Overlay

- “Shared world” model of routing – topology info sent everywhere
- Parallel to L2 packet bcast everywhere on unknown address
  - L2 proliferate packet vs. L3 proliferate routing info
  - L2 proliferate packet garbage vs. L3 proliferate routing garbage
- Damage: routing blackhole or grey hole

The router needs to filter out routing misinformation and select the route, without depending on all other routers

Feedback-based Routing

- Each access router
  - Gets potential routes from “broadcast” topology updates
  - Monitors packet traffic over routes plus sends probes to check potential routes
  - Filters out bad routes, only uses routes known to work
  - Controls packet paths with source routing
- Use feedback, like most engineered dynamic control systems

Local control and no need for global trust, assuming source routing

Source Routing

- Control the (loose) route each packet takes
- WRAP: Wide-area Relay Addressing Protocol
  - Specifies loose source route for packet
  - Shim protocol over IPv4
- But also, fosters competition among ISPs
- But also, supports NAT Inter-realm addressing
- But also, more addresses than IPv6
- And most routers and hosts need not change

Keep IPv4, easier to deploy and solves more problems, including …

Network filtering and traceback

- Provides instant packet trace-ability
  - Records the route the packet takes
- Versus other schemes
  - Anti-source spoofing (ingress filtering) is not scalable
  - Statistical techniques do not respond fast enough
- Allows scalable network-based filtering
  - Push filters back along receive path to ingress points
  - Reduces flood attack to portion of bandwidth

Research: Show WRAP/filtering can scale

Ref. K. Argyraki, WRAP, forthcoming Ph.D. thesis

But with source routing and FBR, there's more . . .

Instant fail-over for high availability

- Access router maintains two or more edge-disjoint paths to destination
- Packets sent on each path
  - Recall: lots of capacity
- Duplicate suppression at receiving router
- At least one packets gets through with high probability
- Concurrent recovery of failed paths

Research: Show FBR can scale

Ref. D. Zhu, Feedback-based Routing, HotNets 2002, forthcoming Ph.D. thesis

Name-based Routing

- Route to named endpoints, not addresses
  - That's what really identifies end-system state
- Integrate naming into routing system
  - Routing system is a directory service
    - address to next hop mapping
  - Extend to provide name to next hop
- Routing protocols extended to disseminate name binding together with topology info
- Provide multi-path routing at the naming level
  - Supporting replicated sites

True Internet routing to end-system state, but there's more …

Highly Available Naming System

- If you can name it, you can reach it
  - Naming in routers so no off-path dependence
- Redundancy of naming service matches redundancy of connectivity
  - If K multi-homed, then K separate name servers
- Attack-resistant to DDoS
- Scaling by level of indirection
  - Names to routing aggregates, routing aggregates to next-hop

Research: Show NBR can scale

Ref. M. Gritter, Content-based Routing, USITS 2000, forthcoming Ph.D. thesis

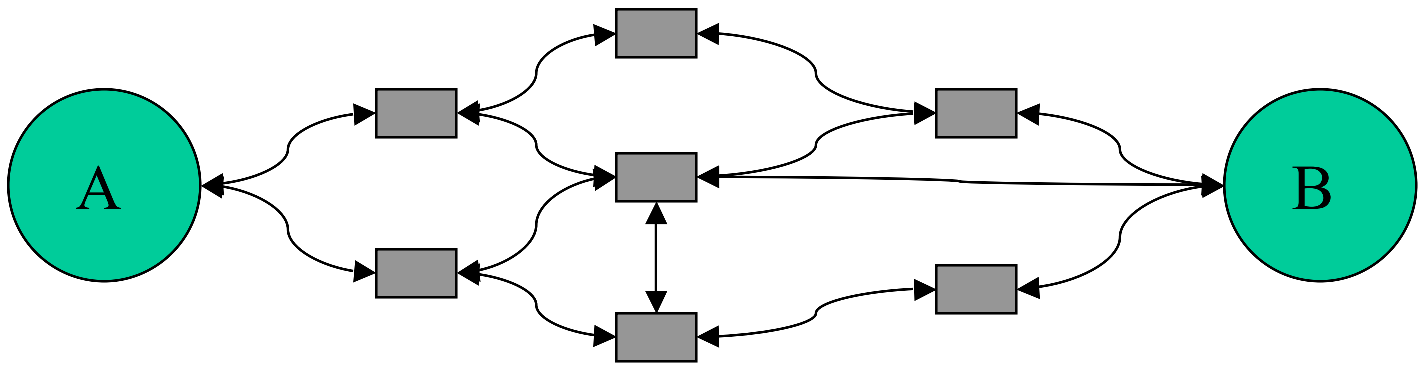
Name-based Connections

- Connection endpoint identified by name, not addr
  - i.e. specify name on connect setup and reconnect
- Name-based checksum
  - Just derive checksum base from end system names
  - Verify packet delivered to right end-system, at same cost
- Works fine with NAT
  - no dependence on addresses
  - Makes NAT state "soft"
- Deployable as a TCP option

Provides true end-to-end reliability,

And allows the Internet to support NAT

I-ATC: Mapping application security onto Physical Security



- True end-to-end reliability to named end systems
- Multiple disjoint redundant paths between nodes
  - Non-stop packet delivery
- Open authentication
  - Multiple messages by independent paths
  - Detection of forged attempts, like ECC
- Clear indication to network operators how to configure

Can we trust this architecture?

So, adversary attacks the I-ATC

- Crack the keys/encryption: sorry, there is none
- Forge a message:
  - Ignored because of trace-ability
  - Detected as a conflict with independent true updates
- Blow up a router:
  - no problem, use an alternative route instantly
- DDoS flooding attack:
  - repelled by network-based filtering

Attack is, at worst, a local failure

Why, why, why ?

- Why is it so hard to make architectural progress
- Named-based Internet proposed in 1991 (RFC 1287)

?

The Internet Religion

True believers do recite:

- The Internet has been very successful so DHCP/IP/TCP/BGP/DNS must be basically right
- Minor technical extensions are the surest means to political agreement
  - DNSsec, secure BGP,
- Political solutions are solutions:
  - There are many possible technical solutions; the hard part is getting agreement, compromise is the key

If you believe it, it will work!

All you need is faith

The Ages of the Internet Architecture

- Age of Pioneers: 1970s
  - Bob Kahn, Vint Cerf, D. Clark, Jon Postel, Len Kleinrock, …
  - Design and build it as “proof of concept”
- Age of Embellishers: 1980s
  - E.g. Deering/Cheriton(IP Multicast)
- Age of Religious Defenders: 1990s+
  - Return to network “transparency” – the flat earth society
  - Defending against excessive (re)invention
- Yes, we need standards, stability, etc but now it’s …

Time for a New Age

The Age of Network Reason

Architectural design based on careful specification on principles and properties

- Semantics
  - E.g. what does "end-to-end reliability" mean?
- Quantitative analysis of scalability

Solid reasoning, not (just) gut instinct, faith and tradition

You may not agree entirely, but …

Hopefully, I’ve convinced you that:

- We need the right architecture and we do not have it now
  - Technical choices do matter
- We need to be faithful to the right architecture
- Many efforts are frightfully off base

So far:

- Students and I identified some of the problems
- Explored some potential solutions
  - And performed preliminary evaluation

There is much more architectural work to do

Conclusions

The Internet architecture:

- is a success
  - Good enough to annihilate the competition
- is a disaster
  - Not good enough to handle critical systems
    - i.e. bad enough to annihilate us!

The future Internet:

- Frightful ad hoc-ery or architecturally faithful

The future Internet architecture

- Political sham “solutions” or science

It matters: I-ATC, You bet your life it does