The Past, Present, and Future of Virtual Networks

Joe Touch
Postel Center Director
USC/ISI
Research Associate Prof.
USC CS & EE/Systems Depts.

Information Sciences Institute

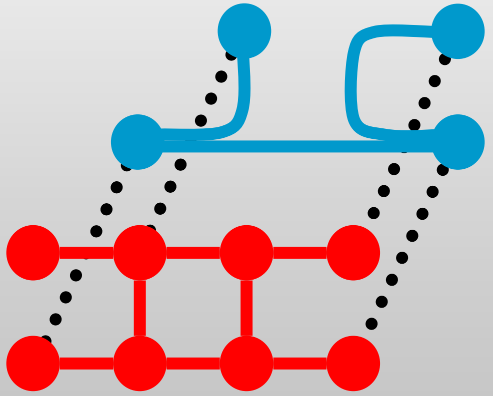
Outline

- Background
  - Definitions & uses
- Past
  - Origins & some accomplishments
- Present
  - Current uses & Caveats
- Future
  - VNs to drive unification

Information Sciences Institute

VN– definition

- Virtual Network is network composed of:
  - Virt. hosts, virt. routers, virt. links (**tunnels**), i.e., an end-to-end system
  - provides at least the same services as any NA
  - in a virtual context
- First-principles extension
  - More than a patch
  - More than interim



Information Sciences Institute

What is a VN?

- *TENET 1. Internet-like*
  - VIs = VRs + VHs + tunnels
  - Emulating the Internet
- *TENET 2. All-Virtual*
  - Decoupled from their base network
- *TENET 3. Recursion-as-router*
  - Some of VRs are VI networks

Information Sciences Institute

VN Corollaries

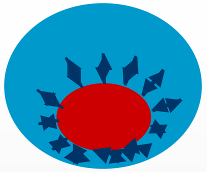
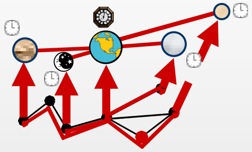
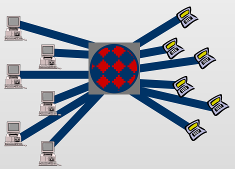
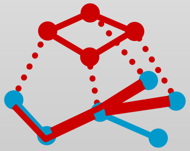
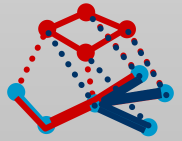
- Behavior:
  - VH adds/deletes headers
  - VRs transit (constant # headers)
- Structure:
  - VIs support concurrence
  - VIs support revisitation
- Each VI has its own names, addresses
  - Address indicates overlay context

How are VNs different?

- Nets deployed/managed over a net
  - Enables new levels of automation/mgt
- Nets not 1:1 to physical devices/topology
  - Logical topology
  - Nodes can be emulated

Information Sciences Institute

Potential Uses

- **Isolate**
  
  - Testbeds, privacy
- **Deploy**
  
  - Dynamic routing, proxylets, security
- **Emulate**
  
  - Overlapping nets, add delay & loss
- **Scale**
  
  - Simplify view of topology
- **Abstract**
  
  - Added level of recovery

Information Sciences Institute

The Past…

- Cronos (1982, RFC-824)
  - Added layer between IP and link *ABSTRACT*
- Operational:
  - M-Bone – multicast *ISOLATE*
  - 6-Bone – IPv6 *ISOLATE*
- Testbed:
  - A-Bone – Active Networks *ISOLATE*
  - Q-Bone – QoS *ISOLATE*
- VPNs *ISOLATE*

Information Sciences Institute

1996-1999 VN Origins

- Planned:
  - Supranet – L1-7 *EMULATE*
  - MorphNet – L1-7 *EMULATE*
  - VONs – "stackable" *SCALE*
  - Genesis – active nets, recursion *SCALE*
- Developed for experiments:
  - Detour/RONs – L3, alternate routing *ABSTRACT*
  - Netscript VANs – L2, active nets, QoS *ABSTRACT*
  - Darwin – QoS *ABSTRACT*
- Deployed:
  - X-Bone – L3 *(any)*

Information Sciences Institute

What changed?

- Virtual interfaces
  - Decoupling address from interface
- Encapsulation as a link
  - No need for new tunnel protocols
  - No need for immediate adjacency
- Use of the base net as OOB channel
  - Allows net management to deploy new nets

Information Sciences Institute

Virtual Interfaces

- Allow device sharing
  - More than one address on a single physical device
- Allow overloading
  - More than one L3 address on a single L2 address
- Revise without reboot
  - No need to restart OS to change addresses
  - (Happened prior to VIFs, but esp. with VIFs)

Information Sciences Institute

Encapsulation as Link

- Custom layering – *one time only*
  - VPN IDs
  - Source routing
- Generic layering – *can be repeated*
  - IP in IP
  - GRE
  - Ethernet in Ethernet

Information Sciences Institute

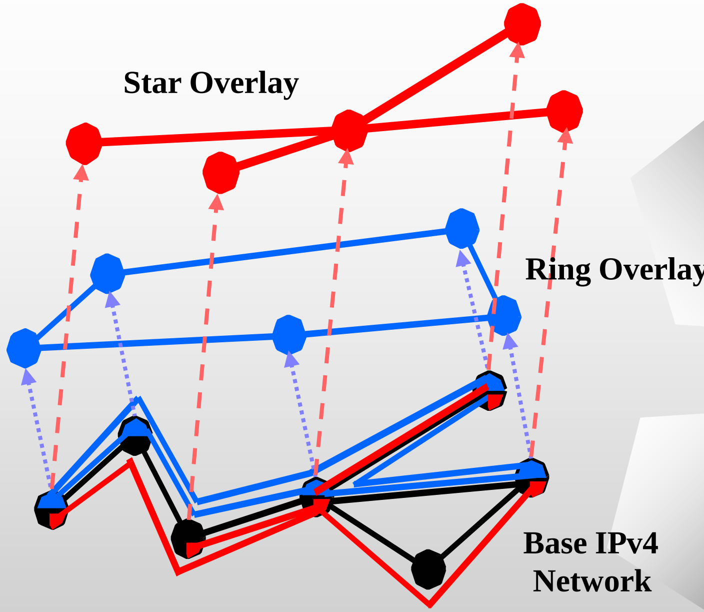
Base OOB Channel Use

- “Base” networks require non-network management
  - Can’t assume a control channel
  - Treat provisioning as separate from operation
- VNs always have a base network
  - Assumed control channel encourages automation
  - Automation encourages increased optimization and monitoring

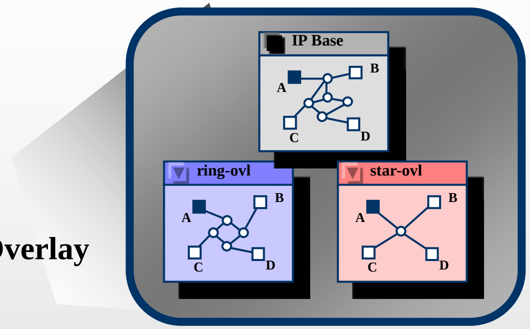
Information Sciences Institute

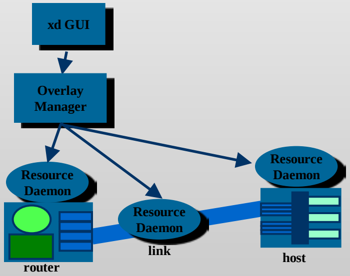
X-Bone Overlay System

Web GUI

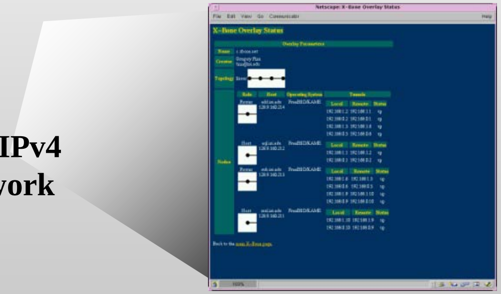


Multiple views





X-Bone system



Automated monitoring

Information Sciences Institute

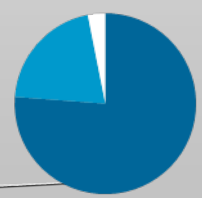
X-Bone Aspects

- Network management over a network
  - DWIM, GUI-based network deployment
  - XML language for describing overlays
- Robust distributed system
  - Idempotent commands
  - Transactions with rollback and recovery
  - Persistent state (save to disk)
- Overlay advances
  - See later slide...

Information Sciences Institute

Timeline

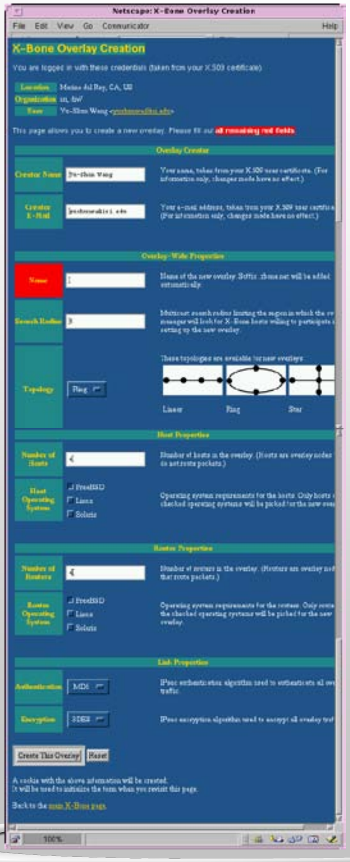
- 1997 – first whitepaper
- 1998-2001 – X-Bone (DARPA)
  - IP overlays with revisitation, recursion (LISP)
  - 2000 – running code (FreeBSD, Linux)
  - 2000 – application deployment
  - 2001 – TetherNet “NAT-buster” to support demos
- 2001-2004 – DynaBone (DARPA)
  - 800-way spread-spectrum parallel overlays
  - 15-level deep overlays
- 2001-2003 – NetFS (NSF)
  - File system configuration of network properties
- 2002-2005 – X-Tend (NSF)
  - X-Bone for testbed uses
- 2003-2005 – DataRouter (int.)
  - Support for overlay P2P forwarding
- 2005-2006 – Agile Tunnels (NSA)
  - Partial overlays for DDOS safety
- 2006-2009 – RNA (NSF)
  - Extending X-Bone Choices model to general protocol stack architecture

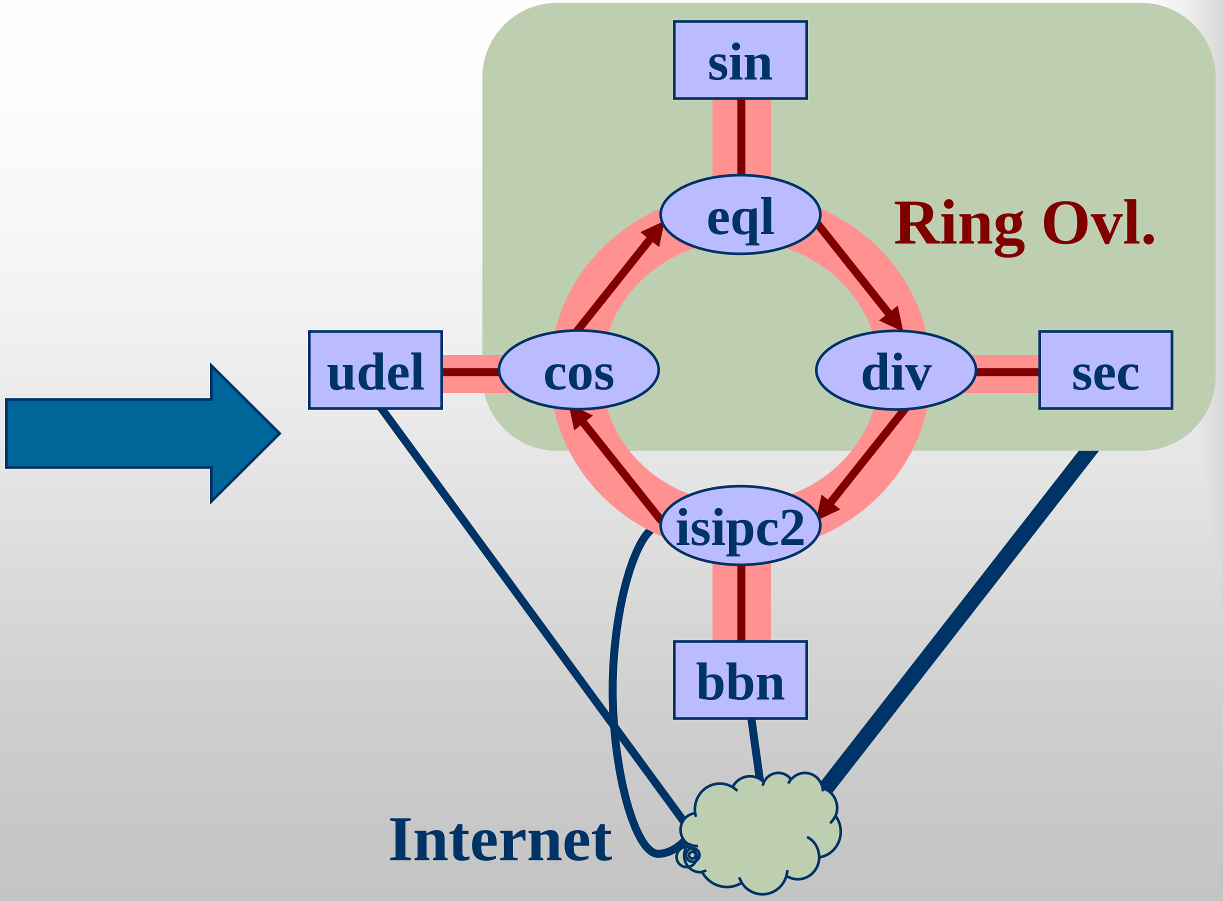


Information Sciences Institute

Creating a Ring

Request





Information Sciences Institute

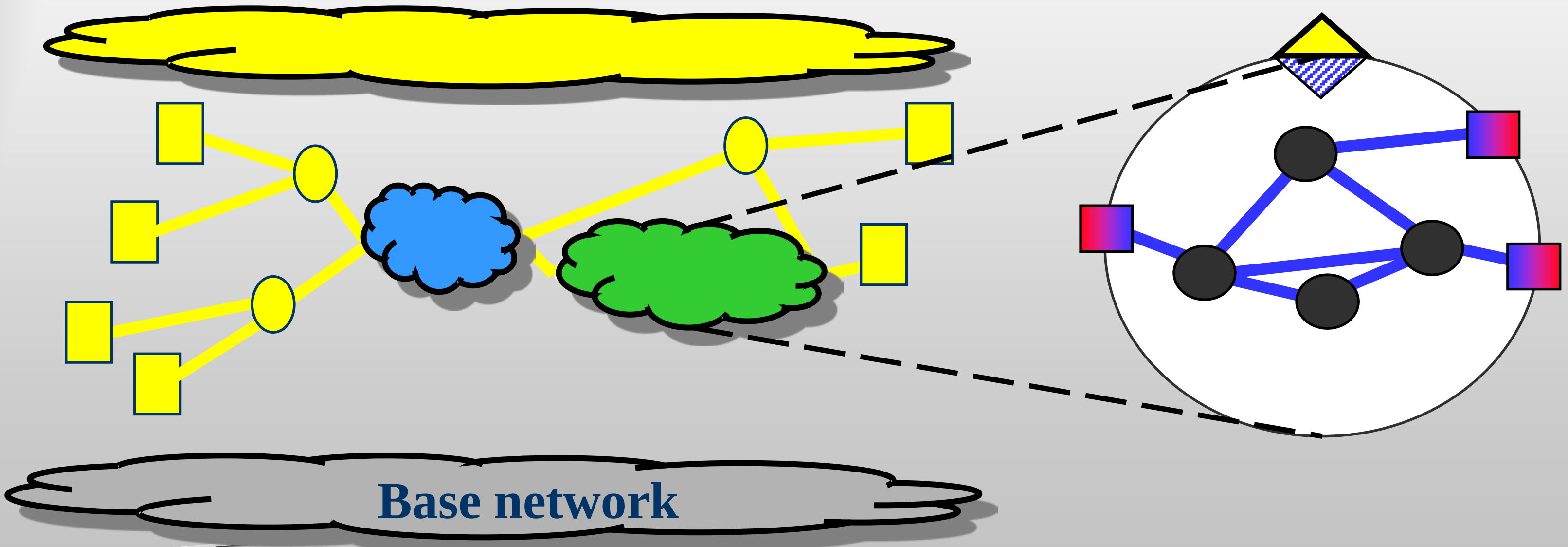
X-Bone Constraints

- Internet-based
  - Routing (link up) vs. provisioning (link add)
  - *…one header to bind them all…*
    (use IP & provide IP = recursion)
- Complete E2E system
  - All VNs are E2E
- VN "Turing Test"
  - A net can't tell it's virtual
- Use existing protocols, OSs, apps.

Information Sciences Institute

Recursion-as-Router

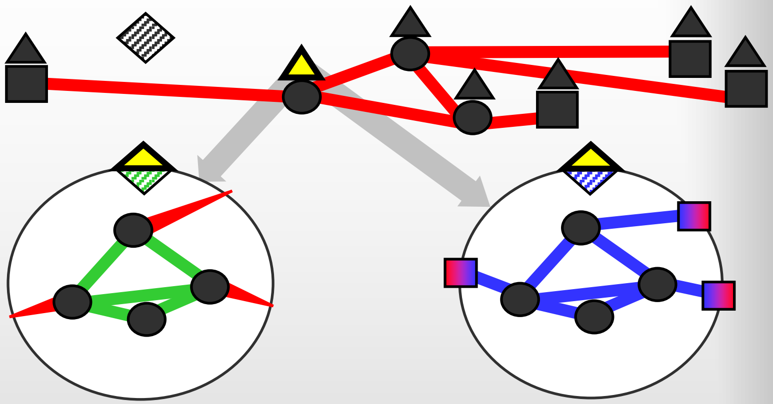
- **Sub-overlays look like routers**
  - L3 version of *rbridges (IETF TRILL WG)*
  - Similar to *LISP*



Information Sciences Institute

X-Bone Enables (1)...

- Recursion
  - Control (like BGP AS's)
  - Network (like LISP/NERD)
    - BARP (label distrib)

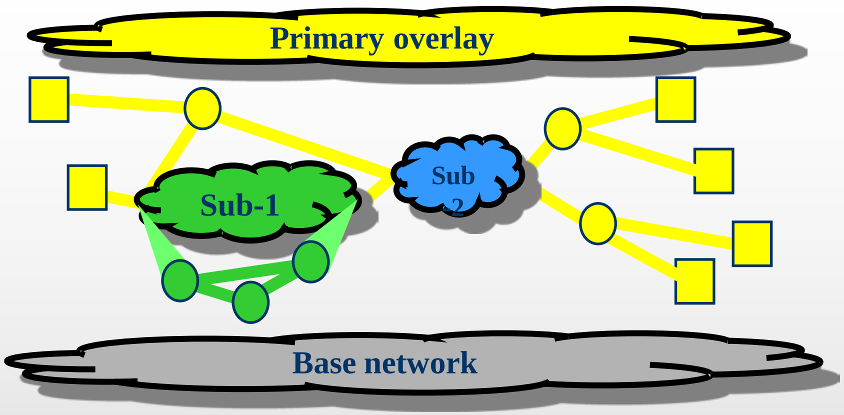


- Revisitation
- Integration of resolution, choices
  - Shims and glue layers as fundamental
- Service for deploying & managing VIs
  - Language for describing VIs

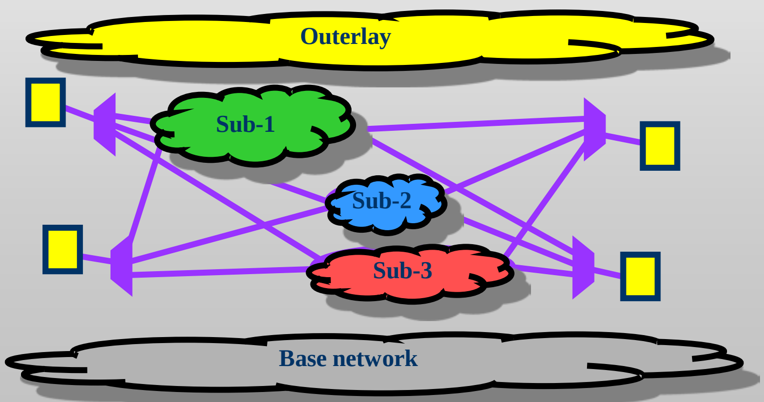
Information Sciences Institute

X-Bone Enables (2)...

- Compose:
  - DTN, Plutarch



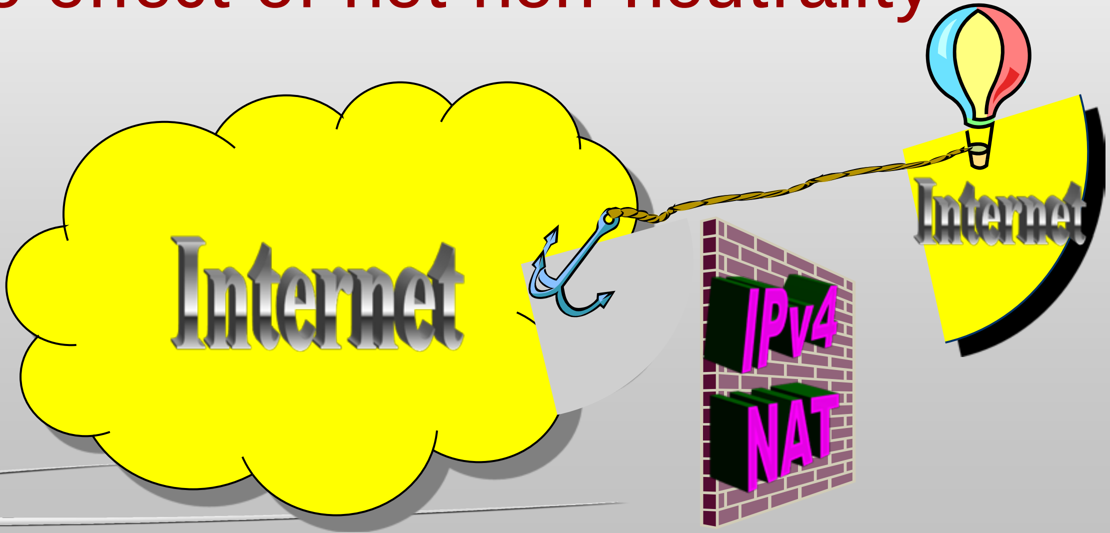
- Alternate:
  - Control Plane, FEC, Boosters,
  - Dynabone



Information Sciences Institute

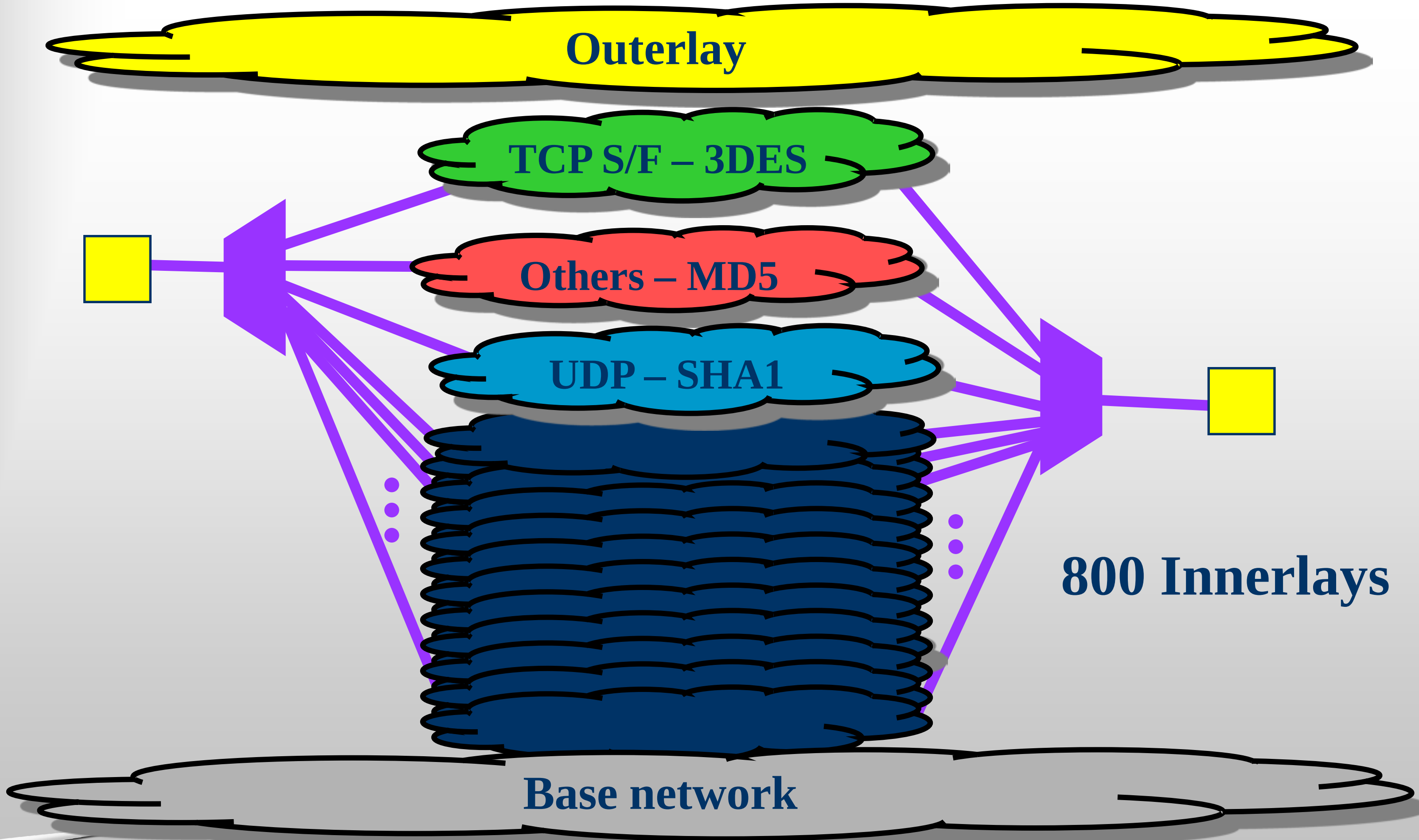
TetherNet

- Rents a block of addresses
  - Auto-configures secure tunnel
- Undoes effect of NAT/NAPT
  - Also effect of net non-neutrality



Information Sciences Institute

DynaBone: Spread Spectrum

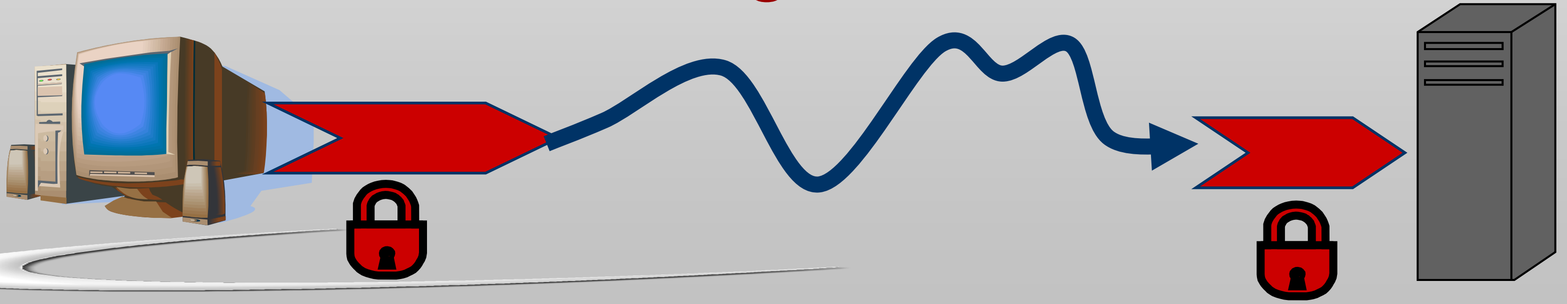


Information Sciences Institute

Agile Tunnel Protocol (ATP)

- Client
  -> tunnel head @client
  -> roaming tunnel tail
  -> server (hidden)

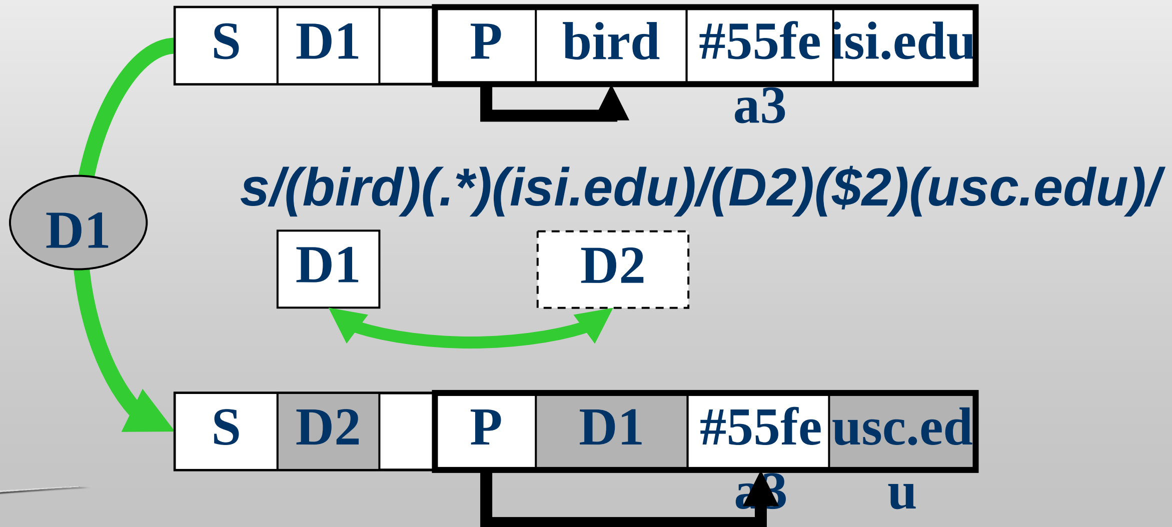
- Works like a floating tunnel:



Information Sciences Institute

DataRouter for P2P

- P2P re-implements network arch.
- Need app.-layer forwarding at net layer
- Add string-based forwarding



Information Sciences Institute

X-Bone Contributions

- Host model
  - Embedded router
  - Socket as unit of overlay isolation
- Recursion model
  - Subnet as router
- Revisition architecture
  - Requires 2-layer tunnels
- Routing / IPsec integration architecture
  - Requires embedded intermediate interfaces

Information Sciences Institute

Observations

- Virtualization *changes* the architecture
  - Hosts are really processes,
    everything else is really a router or system
  - Devices aren't localized
    - Subnet as a router
    - NAT as a host front-end
  - Link and net layers are tightly coupled
- Core concepts from previous glue/shims
  - A single model yields layering, forwarding, routing, and dynamic composition

Information Sciences Institute

The Present…

- Testbeds
  - GENI *ISOLATE/EMULATE*
  - AKARI *ISOLATE/EMULATE*
  - FIRE *ISOLATE/EMULATE*
- Routing infrastructure
  - Rbridges/TRILL *SCALE*
  - LISP *SCALE*

Information Sciences Institute

What VNs Currently Do

- Keep "ships" separate
  - No sibling interference
  - No parent-child interference
  - Establish sibling "relative" QoS ("at most")
- PEP-style enhancements
  - Dynamic routing
  - FEC, Multipath

Information Sciences Institute

What VNs Cannot Do

- Enforce performance constraints
  - Fixed BW, latency
  - Provisioning-style, e.g., “at least” QoS
- Enhance app. interactions
  - Needs networking, i.e., multihop forwarding
  - Grid/Cloud Computing is single hop E2E

Information Sciences Institute

Potholes

- Confusing virtual provisioning with routing
  - Establishing tunnel = provisioning
  - Selecting from a set of tunnels = routing
- Optimizing to an underlying network
  - It could be virtual!
- Tunnel problems
  - MTU issues, signalling issues
  - Security/protection (IP ID wrap, checksum)

Information Sciences Institute

E.g.: New Tunnels

- SEAL (Templin, I-D 2009)
  - Augments IP ID number space
  - Adds checksum
  - Adds PMTUD / PLPMTUD
  - Adds ingress-egress signalling

Information Sciences Institute

Current Efforts

- IRTF NetVirt BOF / VNRG mailing list
  - Preparing charter for IRTF RG
  - Focusing on network issues (host arch., net arch.)
  - was "NVRG"
- Future Internet meetings
  - ICCCN 2008 "FIAPP" (future Internet arch & protos.)
  - CoNext 2008-9 "ReArch" (re-architecting the Internet)
  - ICCCN 2009 "NAP" (net arch & protocols)
  - Globecom 2008-9 FutureNet

Information Sciences Institute

The Future: Unified Architecture

- VN as basis of unification
  - Unify layering and forwarding
  - Unifying different layers
- Examples:
  - RNA
  - Network IPC (Day)

Information Sciences Institute

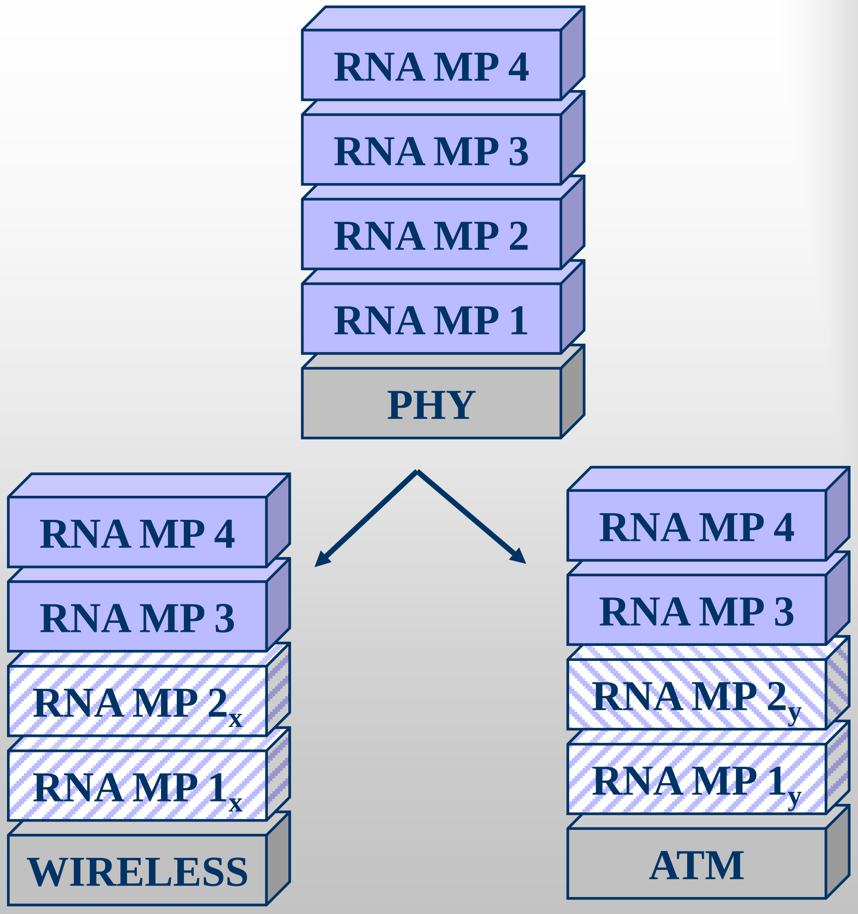
What if...

- Über-protocols are the right idea...
  - A single configurable protocol with
    - Hard/soft state management
    - Congestion control, error management
    - Security
  - *E.g.*, XTP, TP++
- But they went too far...
  - Keep layering – because of first principles

Information Sciences Institute

Recursive Net Arch

- Layering as more than software engr.
  - Layers defined by scope, context
- Create a one layer 'stem cell' protocol
  - Integrate resolution, "choices" from X-Bone
  - Template of basic functions, ala J. Day



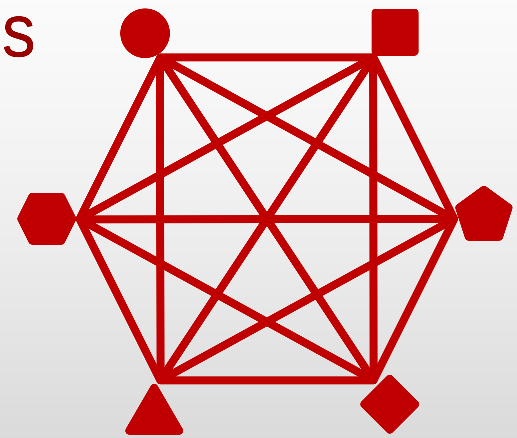
Information Sciences Institute

Exploring Invariants

- Networking is *groups of interacting parties*
  - Groups are heterogeneous
  - All members want to interact
  - Groupings are dynamic (*i.e.*, virtual)
- Thus, need an architecture that supports:
  - Heterogeneity
  - Interaction
  - Virtualization

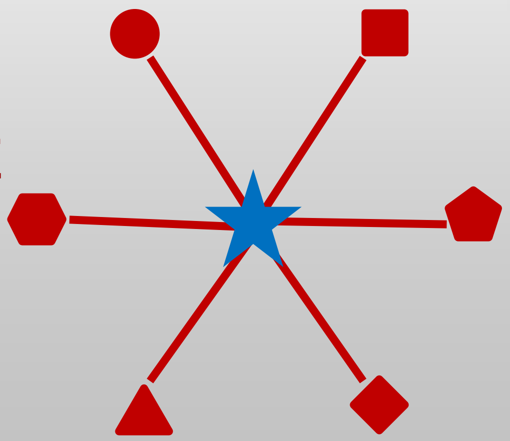
Heterogeneity leads to layering

- M different interacting parties need
  - M^2 translators



or

  - M translators + common format

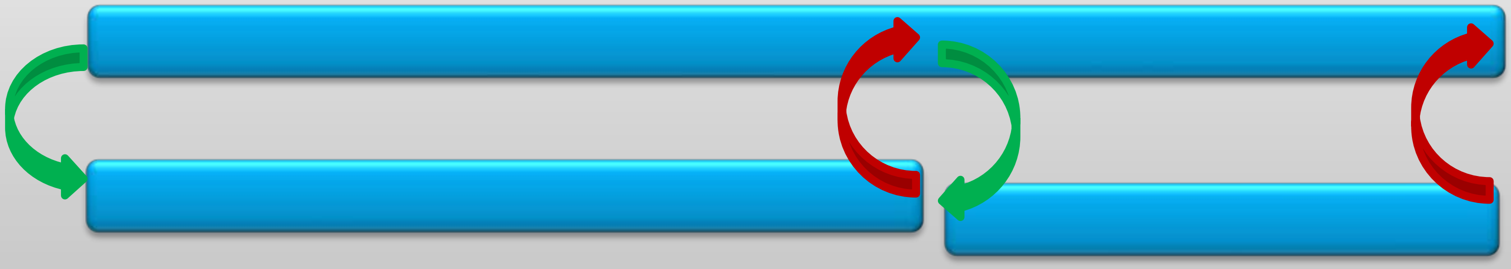


... *i.e.,* a layer

Information Sciences Institute

Layering leads to resolution

- IDs are local to a layer
  - Whether names, paths, locations
- Need to resolve IDs between layers
  - Google, DNS, ARP, LISP encap tables



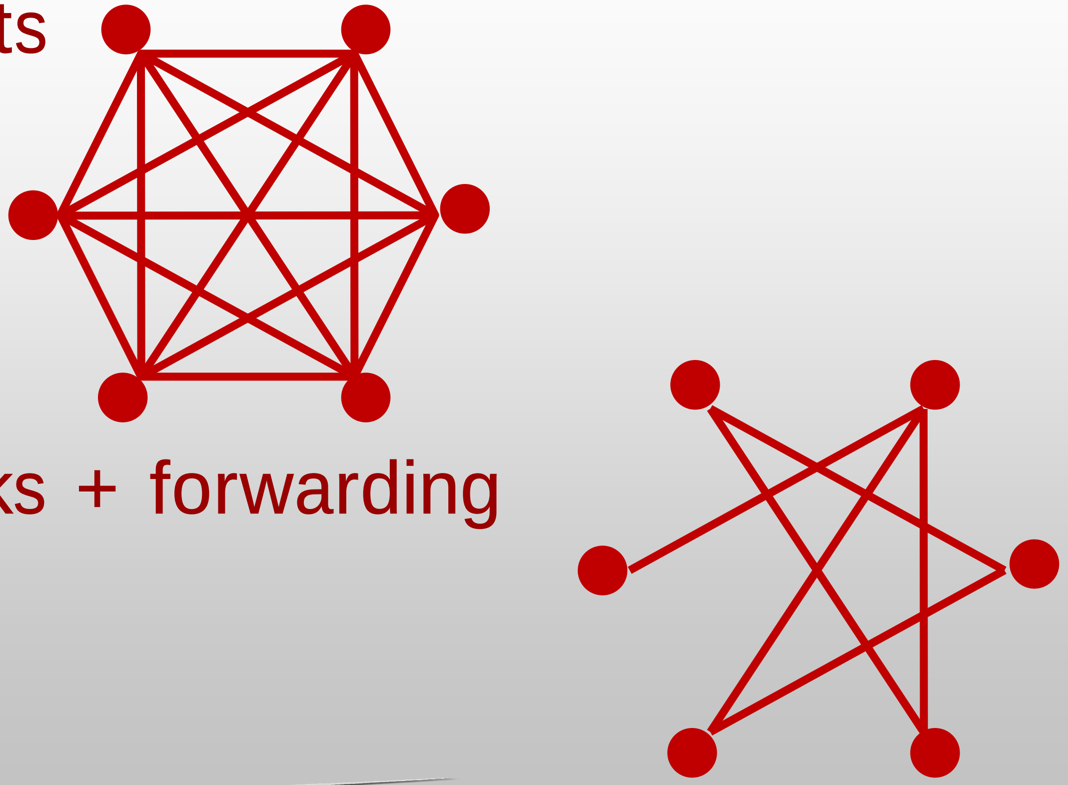
Information Sciences Institute

Interaction leads to forwarding

- N parties need
  - N^2 circuits

or

  - O(N) links + forwarding

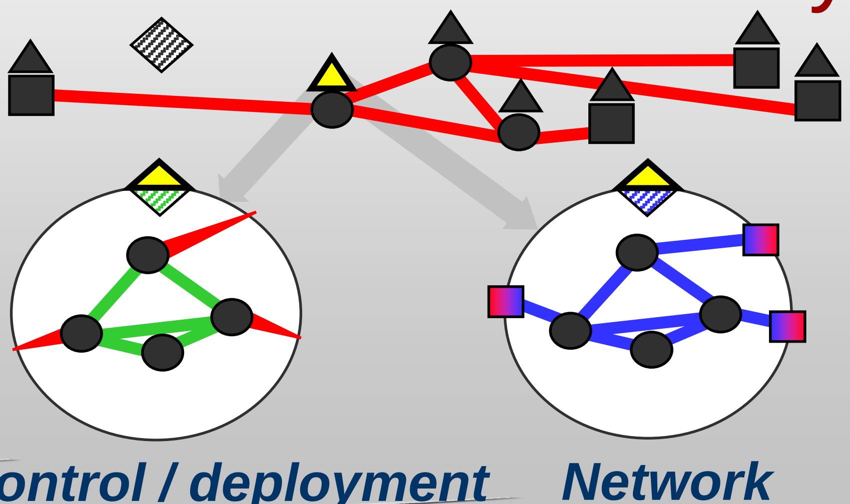


Information Sciences Institute

Virtualization leads to recursion

- N parties want to group in arbitrary, dynamic ways.
  - ... such groups are inherently virtual

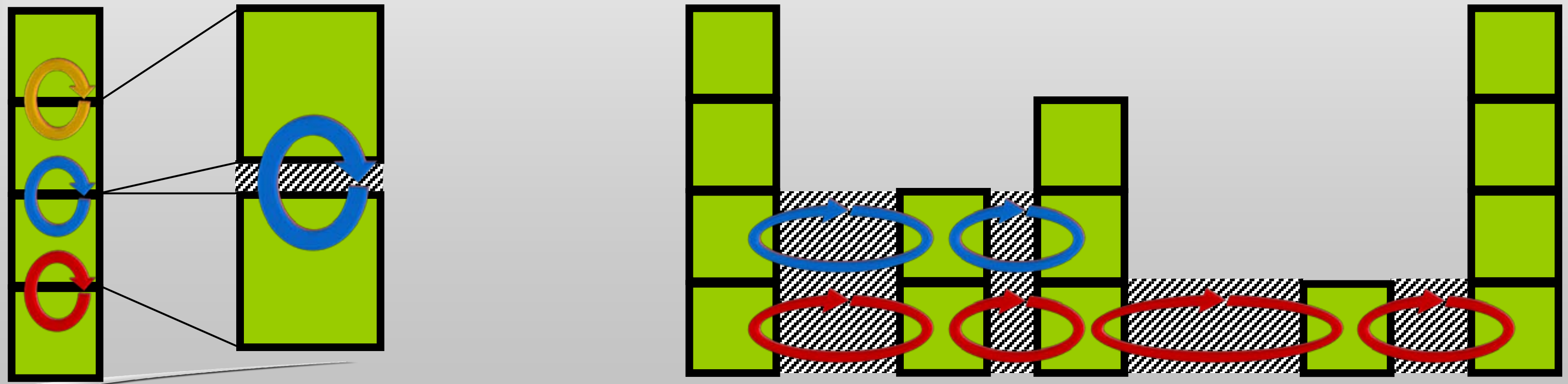
... and virtualization is inherently recursive



Information Sciences Institute

Recursion unifies layering, forwarding, & resolution

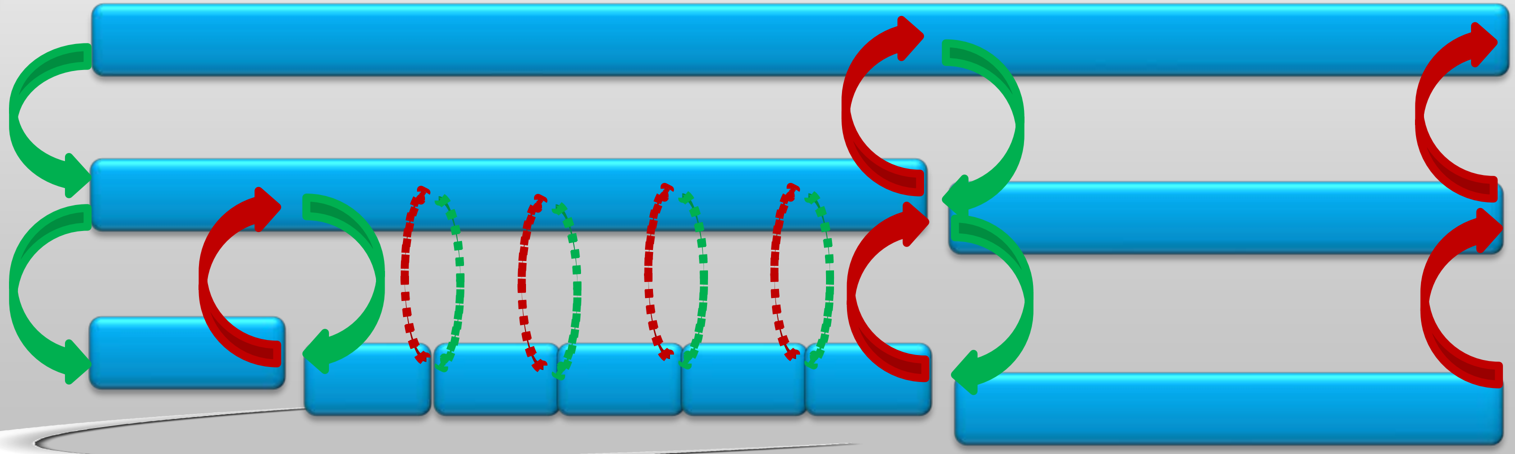
- Layering (left)
  - Heterogeneity via O(N) translators
  - *Supported by successive recursive resolution*
- Forwarding (right)
  - N^2 connectivity via O(N) links
  - *Supported by successive iterative resolution (tail recursion)*



Information Sciences Institute

RNA

- One metaprotocol, many instances
  - Needed layers, with needed services
  - Layers limit scope, enable context sensitivity
  - Scope defined by reach, layer above, layer below
  - Resolution connects the layers (red/green)



Information Sciences Institute

RNA MP Unifies…

- "Resolve" unifies:
  - Layer address translate/resolution
    - ARP, IP forwarding lookup
    - BARP/LISP/TRILL lookup
  - Layer alternates selection
    - IPv4/IPv6, TCP/SCTP/DCCP/UDP
  - Iterative forwarding
    - IP hop-by-hop, DNS recursive queries
- "Process data" unifies:
  - Shared state, security, management
  - Flow control, error control

```
LAYER(DATA, SRC, DST)
   Process DATA, SRC, DST into MSG
   WHILE (Here <> DST)
      IF (exists(lower layer))
         Select a lower layer
         Resolve SRC/DST to next layer S',D'
         LAYER(MSG, S', D')
      ELSE
         FAIL /* can't find destination */
      ENDIF
   ENDWHILE
   /* message arrives here */
   RETURN {up the current stack}
```

Next-hop Resolution

Next Layer Resolution

Information Sciences Institute

What does RNA enable?

- Explains and details invariants
  - Layering as more than a SW Engr. artifact
- Integrate current architecture
  - ‘stack’ (IP, TCP) *vs.* ‘glue’ (ARP, DNS)
- Support needed improvements
  - Recursion (AS-level LISP, L3 BARP, L2 TRILL)
  - Revisitation (X-Bone)
  - Concurrence (VPNs, multipath TCP)
- Supports “old horse” challenges natively
  - Dynamic ‘dual-stack’ (or more)

Information Sciences Institute

Conclusions

- Virtualization requires recursion
- Recursion supports layering
- Recursion supports forwarding

One recurrence to bind them all…

- *Recursion is a native network property*
  - Integrates and virtualization, forwarding and layering **in a single mechanism**

Information Sciences Institute

Acknowledgements

- X-Bone, DynaBone, X-Tend
  - Lars Eggert, Yu-Shun Wang, Greg Finn, Steve Hotz, Oscar Ardaiz-Villanueava, Norihito Fujita
- NetFS
  - Josh Train
- DataRouter
  - Venkata Pingali
- RNA
  - Yu-Shun Wang, Venkata Pingali