USC

Information Sciences Institute

A Recursive Network Architecture

Joe Touch

Postel Center Director

USC/ISI

Research Assoc. Prof.
USC CS and EE/Systems Depts.

USC Viterbi School of Engineering

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Outline

- Towards future network architectures
- Background on X-Bone VNs
- RNA
  - Intro.
  - Design
  - Implementation
  - Implications
  - Related work

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Towards future network architectures

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

What makes an architecture new?

- Shaking the Hourglass (CCW 08)
  - All exchanges are 1 packet
  - Collosograms > RTT*delay
  - No LANs? (L2 is only pt-pt)
- What defines success?
  - Fixing what's 'broken'
  - Doing something new/different
  - The Internet / circuits as a degenerate case

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Internet Architecture

Accused of ossification, but:

- Ossification = stability
- Flexibility is abundant:
  - Shim layers:
    - HIP, SHIM6, IPsec, TLS
  - Muxing layers:
    - SCTP, RDDP, BEEP
  - Connections:
    - MPLS, GRE, IKE, BEEP, SCTP
  - Virtualization:
    - L2VPN, L3VPN/X-Bone/RON/Detour, L7-DHTs

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Motivation

- Desire to support new capabilities
  - Interlayer cooperation, dynamic layer selection, layering created by virtualization
- Desire to support emerging abstractions
  - Overlay layers don't map to 1-7
  - Support for recursive nodes (BARP, LISP, TRILL)
- Desire to coordinate services in diff. places
  - Security, soft-state, pacing, retransmission

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Observations

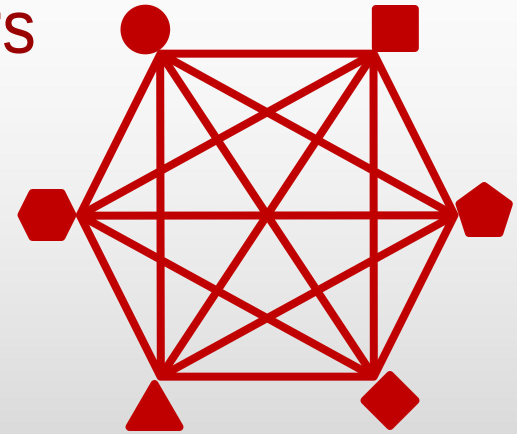
- Networking is *groups of interacting parties*
  - Groups are heterogeneous
  - All members want to interact
  - Groups can be dynamic (*i.e.*, virtual)
- Need an architecture that supports:
  - Heterogeneity
  - Interaction
  - Virtualization

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

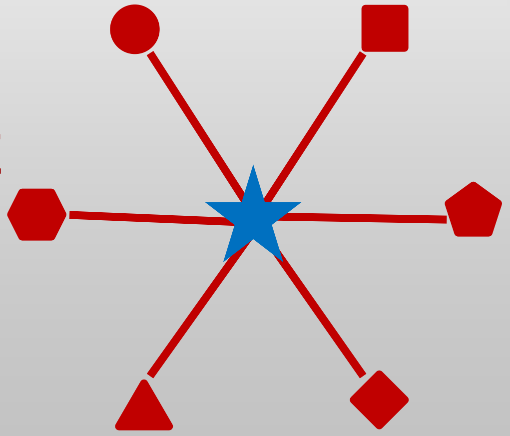
Heterogeneity leads to layering

- M different interacting parties need
  - M^2 translators



or

  - M translators + common format



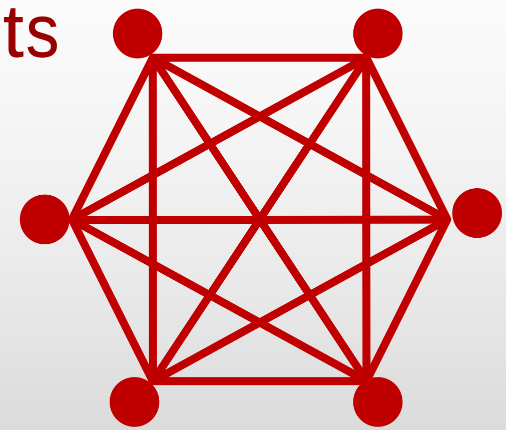
... *i.e.*, a layer

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

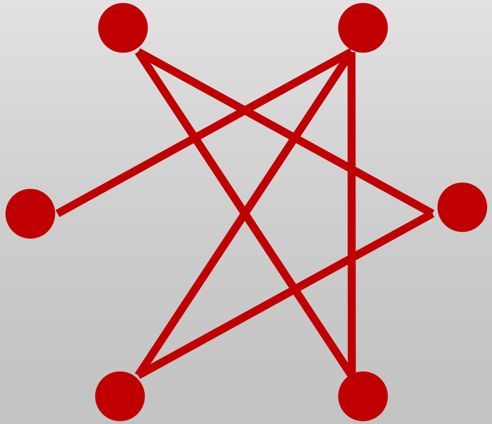
Interaction leads to forwarding

- N parties need
  - N^2 circuits



or

  - O(N) links + forwarding



Copyright 2009, USC/ISI. All rights reserved.

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

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

What makes this an architecture?

- Abstraction for virtualization
  - Tunnel as link
  - Partitioned router as virtual router
  - Partitioned host + internal router as virtual host
- Abstractions for recursion
  - Recursive router implemented as a network of vrouters with vhosts at the router interfaces
  - Recursion within the protocol stack
- General template (metaprotocol + MDCM)
  - Instantiates as different layers or forwarding

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

X-Bone Virtual Nets

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Virtual Net Req'ts

- **Internet-Compliant Architecture**
  - Hosts add/delete headers
  - Routers transit (constant # headers)
- **Supports New Capabilities**
  - Concurrence (multiprocessing)
  - Revisitation (multiple roles in one net)
  - Recursion (to hide topology and/or mgt.)

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

VN Principles

- *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

Copyright 2009, USC/ISI. All rights reserved.

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

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

VN Architecture

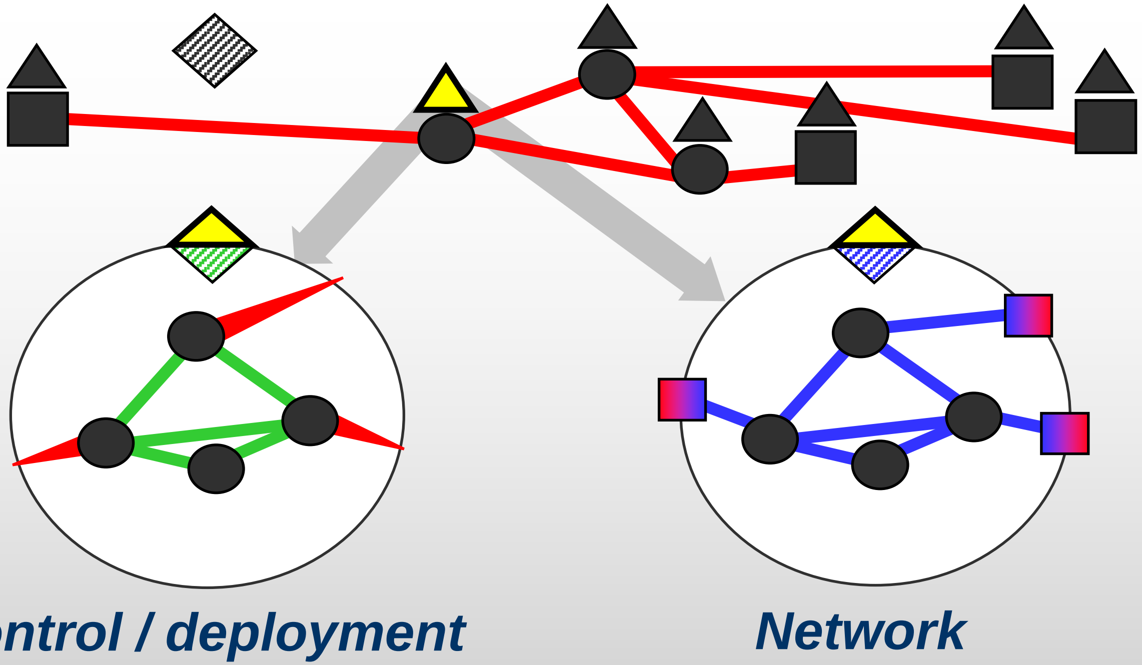
- Components:
  - VH -> hosts include a hidden router
  - VL -> 2 layers of encaps. (strong link, weak net)
  - VR -> partitioned forwarding
- Capabilities:
  - Revisitation -> multihoming for VNs
  - Recursion -> router as network, i.e., Rbridges, LISP

>> RUNNING CODE (FreeBSD, Linux, Cisco)

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Recursive Internet



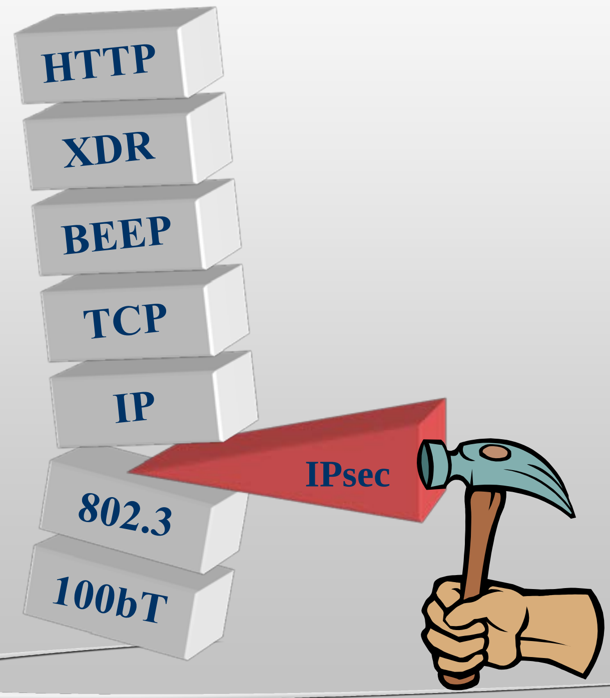
- Recursion as a router (vs. ASes)
- Network recursion examples
  - L3 = BARP (X-Bone), LISP (IRTF)
  - L2 = Rbridges/TRILL

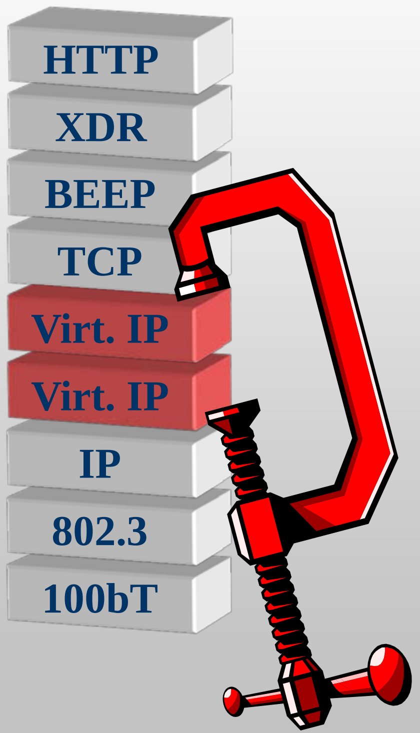
Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Recursion requires new layers – where? Why?

- Wedge between (IPsec, left)
  or replicate (virtualization, right)





Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Challenges of Layering

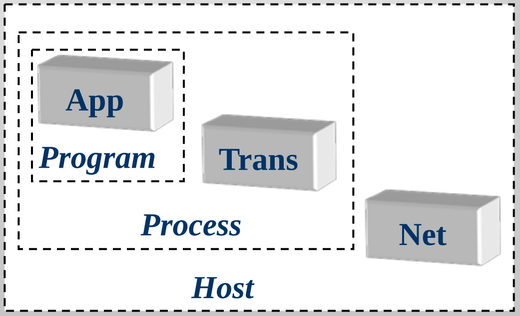
- Which to add...
  - IPv4/IPv6, TCP/DCCP/SCTP
- When to add...
  - Security, muxing, cong. control
- Real vs. virtual
  - What's the difference?

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Scope defines a layer

- Its endpoints
  - A “hop” @layer N = E2E extent of layer N-1
- The layer above
  - What services this layer provides
- The layer below
  - What services this layer requires
- E.g.: Shared state at diff. layers for diff. services
  - Application binding
  - Transport delivery
  - Net security



The difference is scope

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA Intro.

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Motivation for RNA

- Layers of a stack becoming more similar
  - Security, soft-state, pacing, retransmission
- Desire to support new capabilities
  - Interlayer cooperation, dynamic layer selection
- Desire to support emerging abstractions
  - Overlay layers don't map to 1-7
  - Support for recursive nodes (BARP, LISP, TRILL)

Is layering more than a coding artifact?

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Observations

1. Services are relative
2. A template can avoid recapitulation
3. Composition requires coordination

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Recapitulation

- Component services repeat:
  - handshake / state management
  - security
  - policy (admission control, filtering)
  - multiplexing and demultiplexing
  - retransmission
  - reordering
  - pacing / congestion control
  - switching / forwarding
- Compounded by virtualization
  - Layer on layer on layer

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Composition Requires Coordination

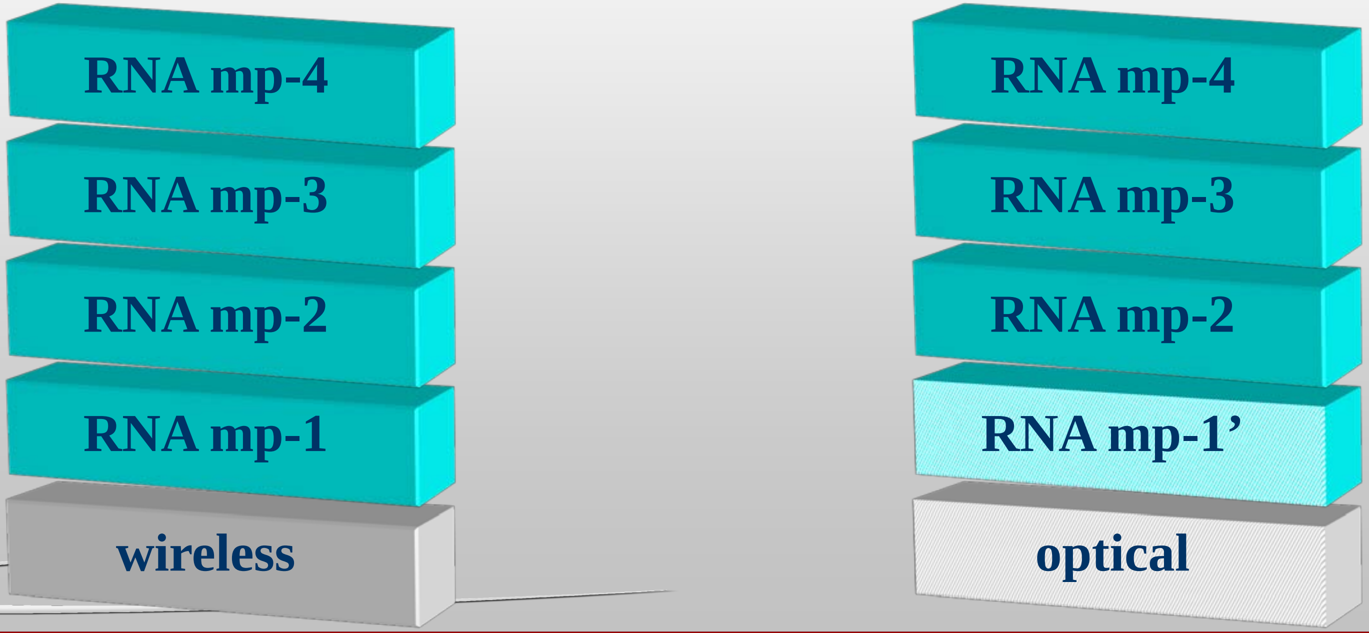
- Many services integrate layers
  - Congestion control
  - Message boundaries
  - Security
  - State establishment
- Current interlayer interface is limited
  - Defined by each layer
  - No general security, state, etc. interface

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA Stack

- One MP, many instances
  - Needed layers, with needed services
  - Layers limit scope, enable context sensitivity
  - Scope defined by reach, layer above, layer below



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

What does RNA enable?

- Integrate current architecture
  - ‘stack’ (IP, TCP) *vs.* ‘glue’ (ARP, DNS)
- Support needed improvements
  - Recursion (AS-level LISP, L3 BARP, L2 TRILL)
  - Revisitation
- Supports “old horses” natively
  - Dynamic ‘dual-stack’ (or more)

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA Design

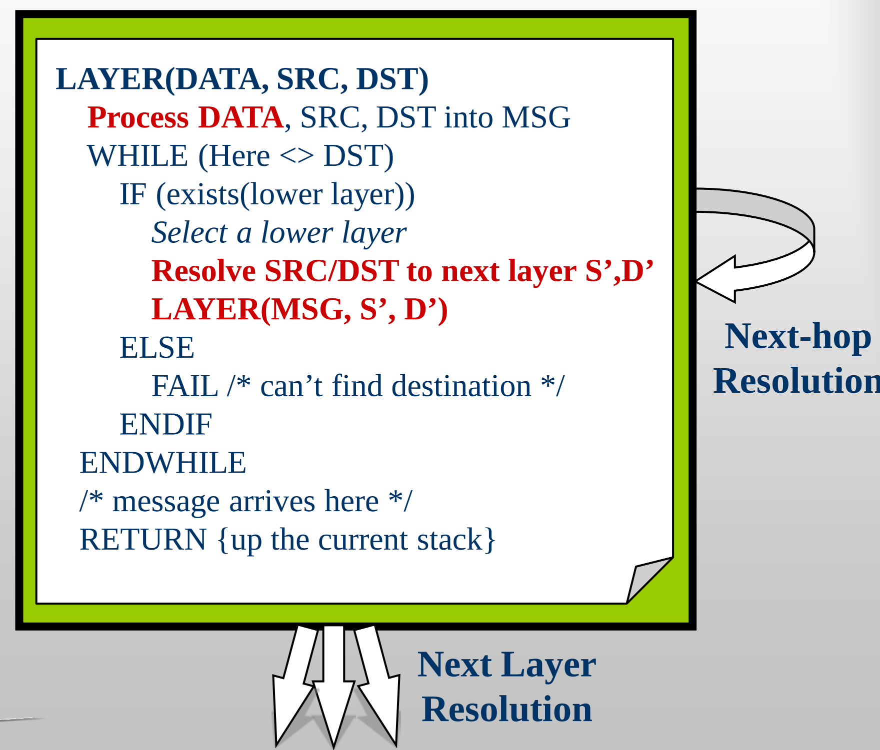
Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

MDCM from *Choices*

Structured template w/plug-in functions

- 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

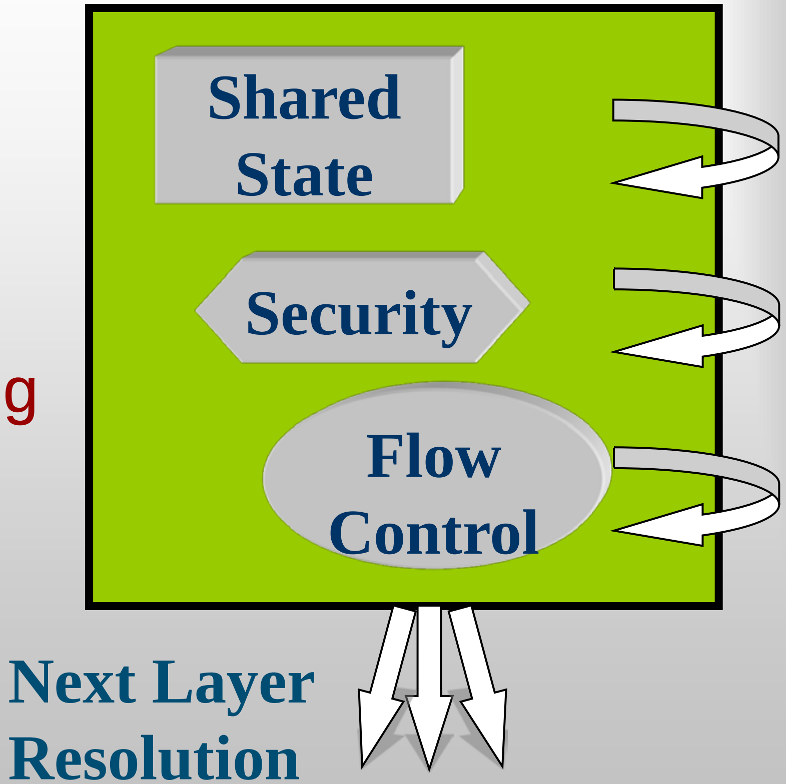


Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA Metaprotocol

- Template of basic protocol service:
  - Establish / refresh state
  - Encrypt / decrypt message
  - Apply filtering
  - Pace output via flow control
  - Pace input to allow reordering
  - Multiplex/demultiplex
    - includes switching/forwarding



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Components of RNA MP

Instantiate MDCM's "Process DATA"

- Establish / refresh state
- Encrypt / decrypt message
- Apply filtering
- Pace output via flow control
- Pace input to allow reordering
- Multiplex/demultiplex as indicated
  - includes switching/forwarding

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA Implementation

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA MP Template

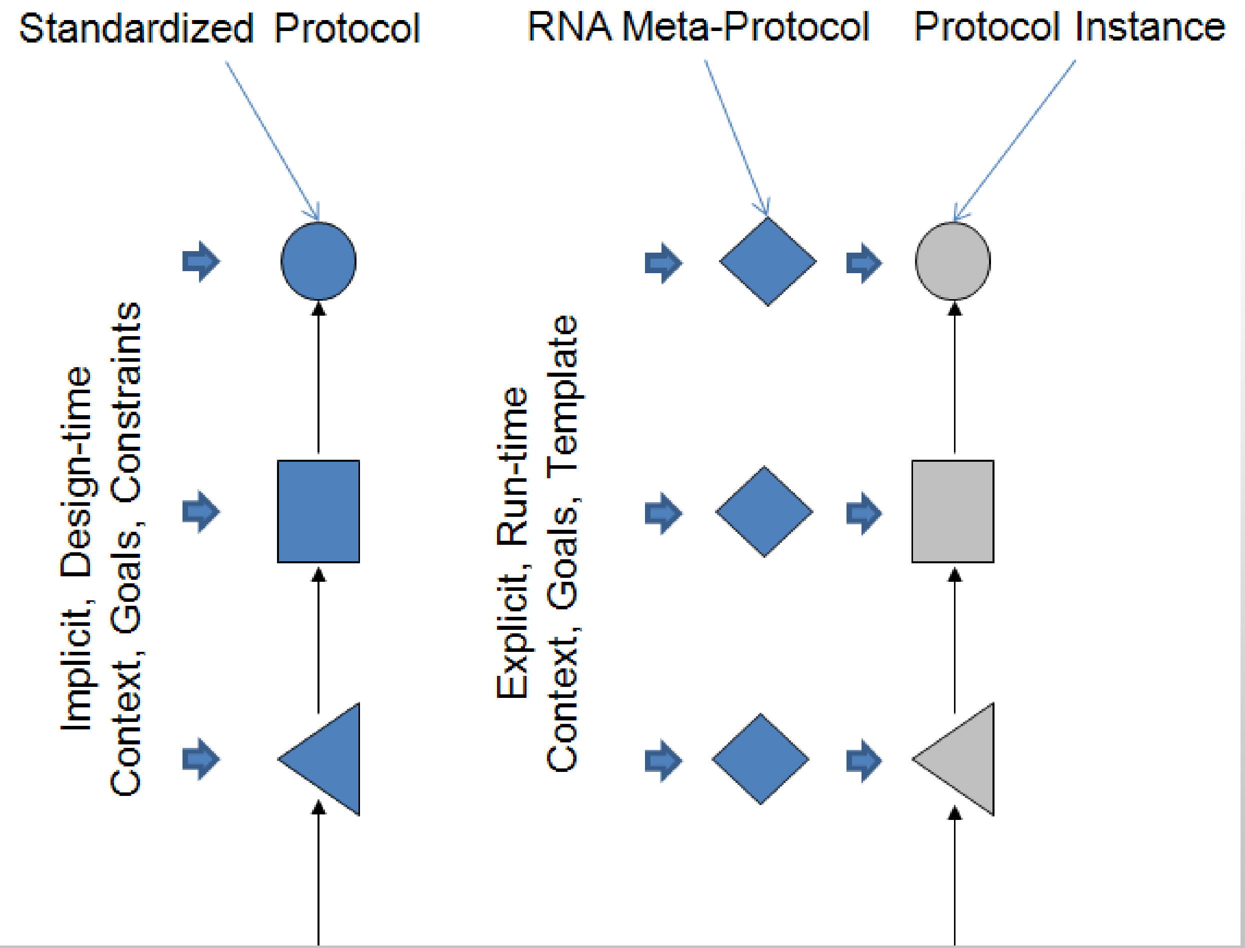
```
START PATTERN MIN

# This simply specifies a buffer. no reodering etc.
PATTERN MIN
    REQ MUST BUFFER 1
    ARG BUFFER 1 VAR size 1000
    LINK ADD SELF 0 BUFFER 1
   ...
# Next use this pattern if MIN is successful
PATTERN ORDERED_DELIVERY
    FOLLOWS MIN
    REQ MUST REORDERING 1
    LINK DEL ….
    LINK ADD ….
…
# If reordering successful, try more stuff…
PATTERN ENCRYPTED_ORDERED_DELIVERY
    FOLLOWS ORDERED_DELIVERY
    REQ MUST ENCRYPTION 1
    ARG ENCRYPTION 1 VAR algo des
    ARG ENCRYPTION 1 VAR keysize 512
    ....
```


Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

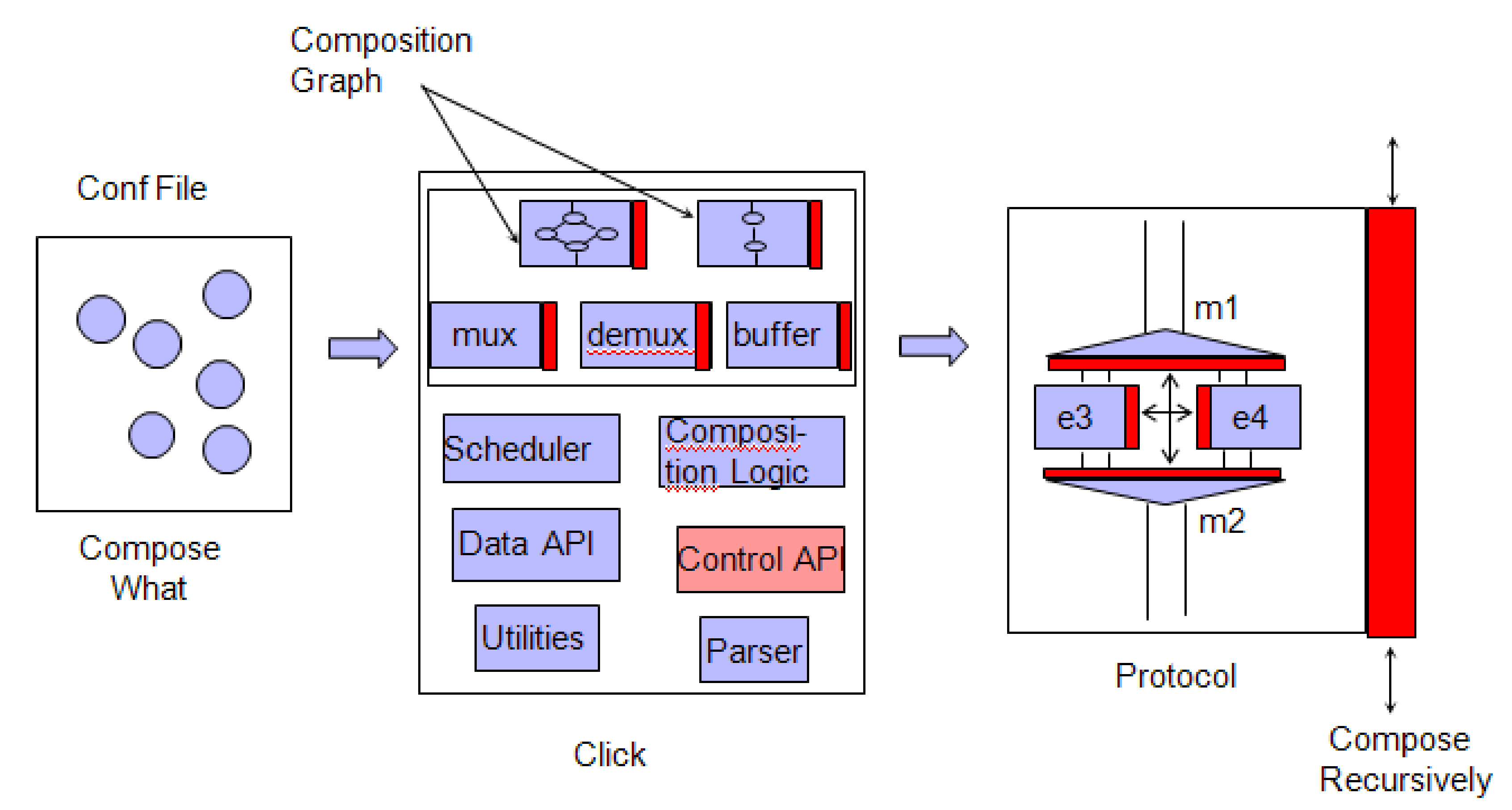
Instantiation



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

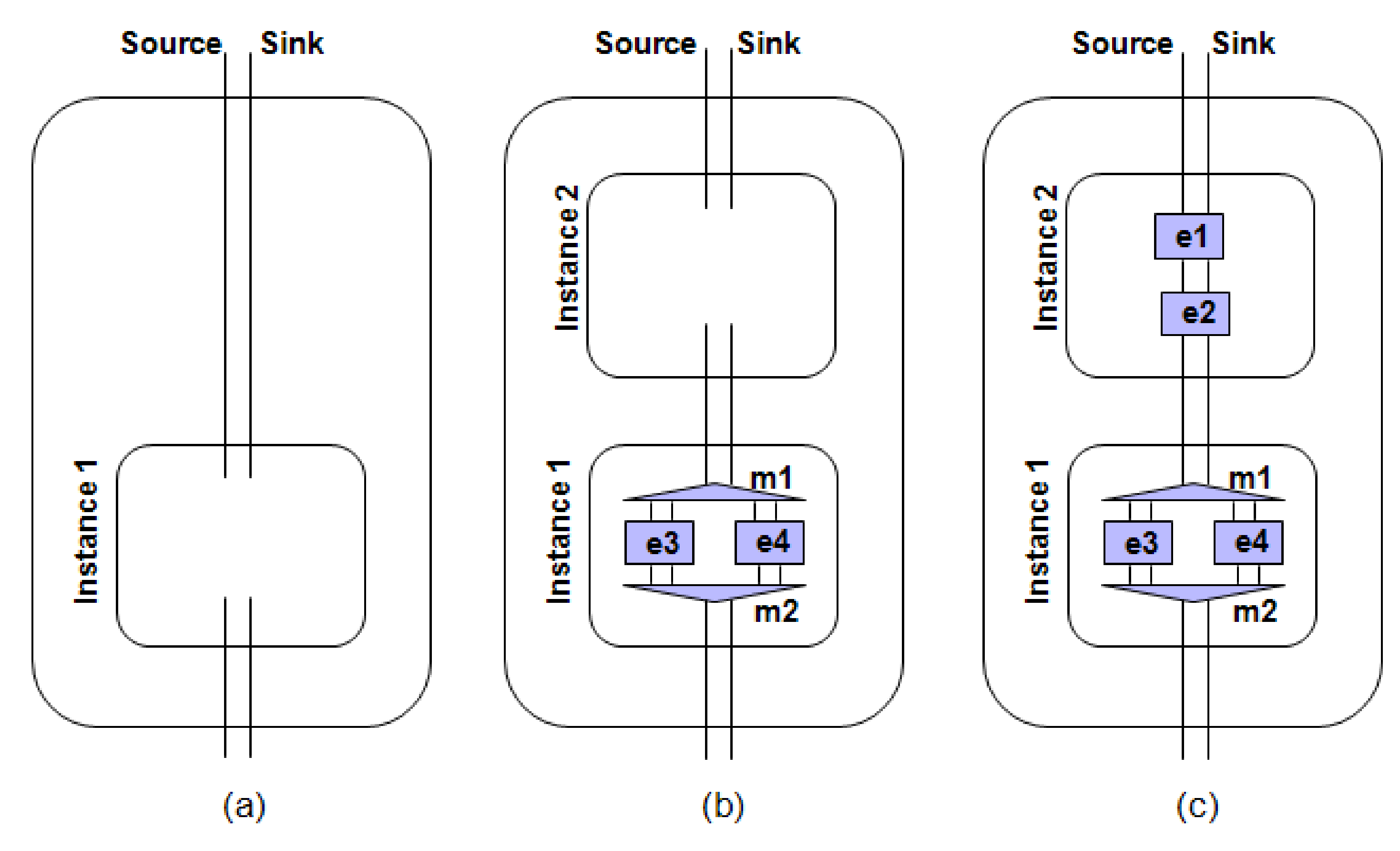
Click Implementation



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

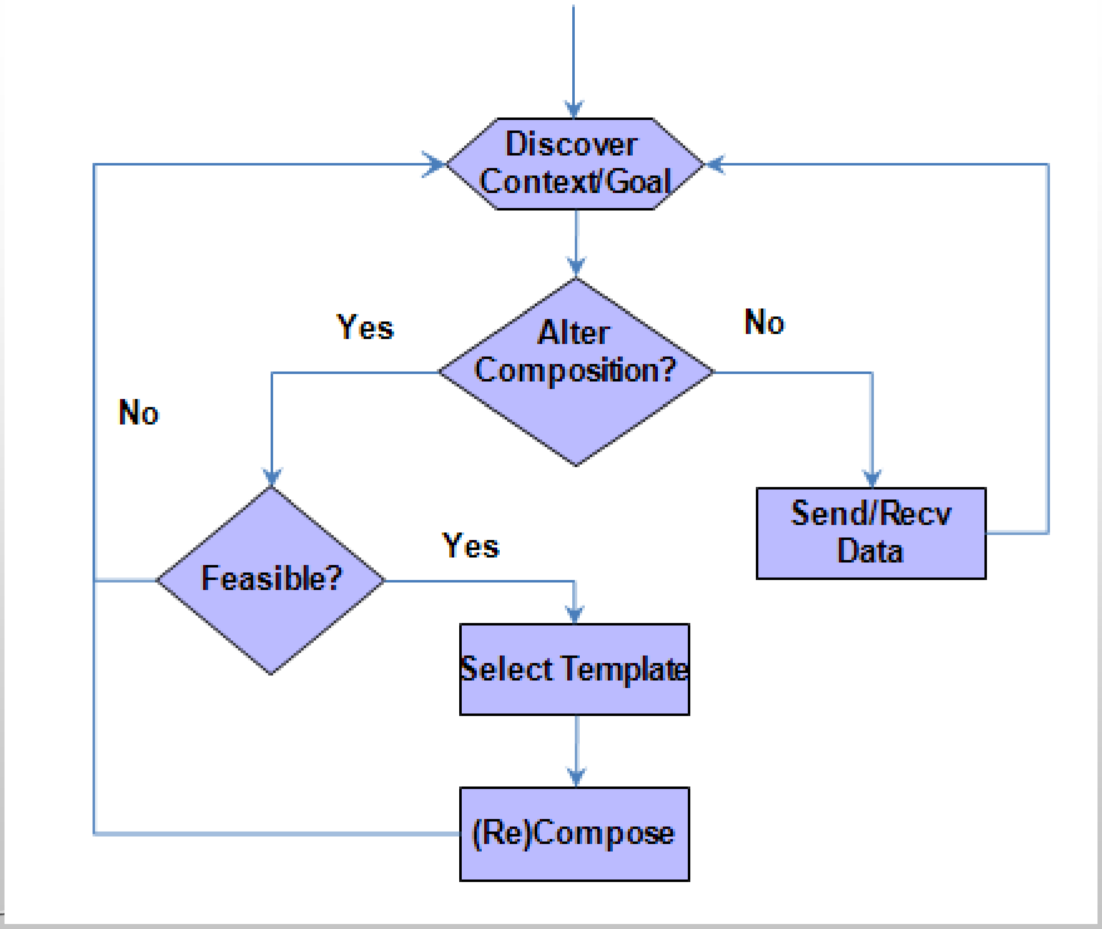
Building a Stack



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Composition Process



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

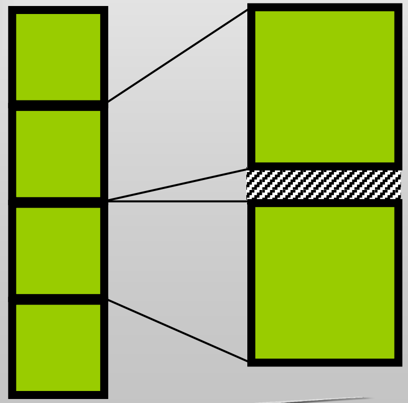
RNA Implications

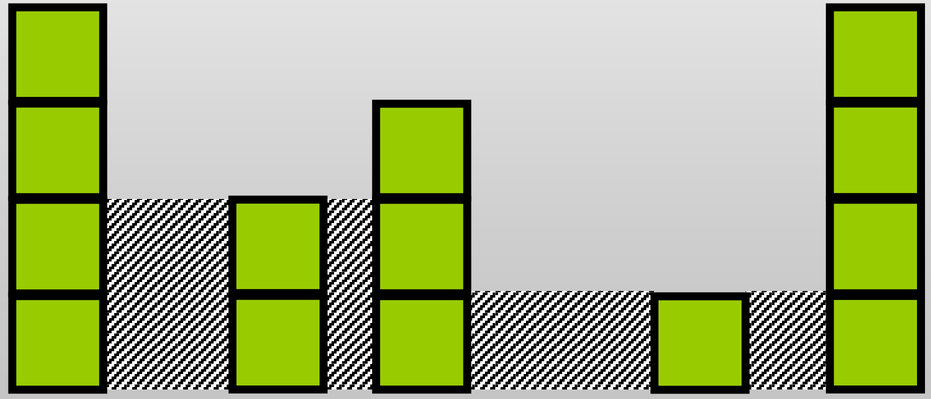
Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

RNA – fills the gaps

- Between layers (left, from *Choices*)
  - Affects next-layer
- Between stacks (right, from Padlipsky)
  - Affects next-hop



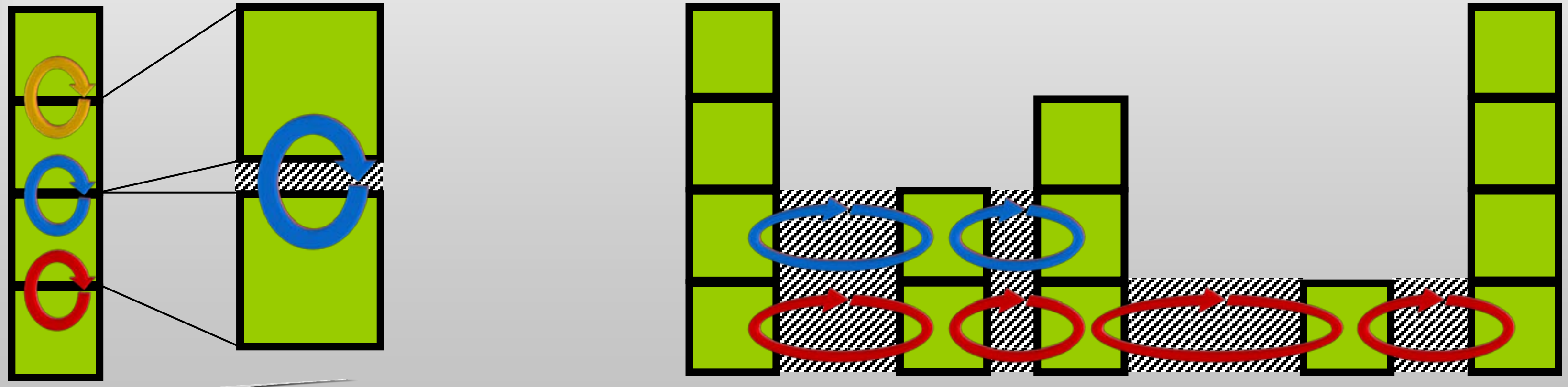


Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Recursion supports Layering and Forwarding

- Layering (left)
  - Heterogeneity via O(N) translators
  - *Requires successive recursive discovery*
- Forwarding (right)
  - N^2 connectivity via O(N) links
  - *Requires successive iterative discovery*



Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Challenges

- MP design
  - Building a sensible, generic template
- Stack management
  - Supporting instantiation and composition
- Supporting interlayer coordination
  - Designing a sensible, recursive API
  - Makes it easier to interface (to yourself, *e.g.*, LEGO)
- Supporting context sensitivity
  - Detecting environment and autotuning

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Related Work

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Related Work Summary

- Recursion in networking
  - X-Bone/Virtual Nets, Spawning Nets, TRILL, Network IPC, LISP
  - *RNA natively includes resolution and discovery*
- Protocol environments
  - Modular systems: Click, x-Kernel, Netgraph, Flexible Stacks
  - Template models: RBA, MDCM
  - *RNA adds a constrained template with structured services*
- Context-sensitive components
  - PEPs, Shims, intermediate overlay layers, etc.
  - *RNA incorporates this into the stack directly*
- Configurable über-protocols
  - XTP, TP++, SCTP
  - *RNA makes every layer configurable, but keeps multiple layers.*

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

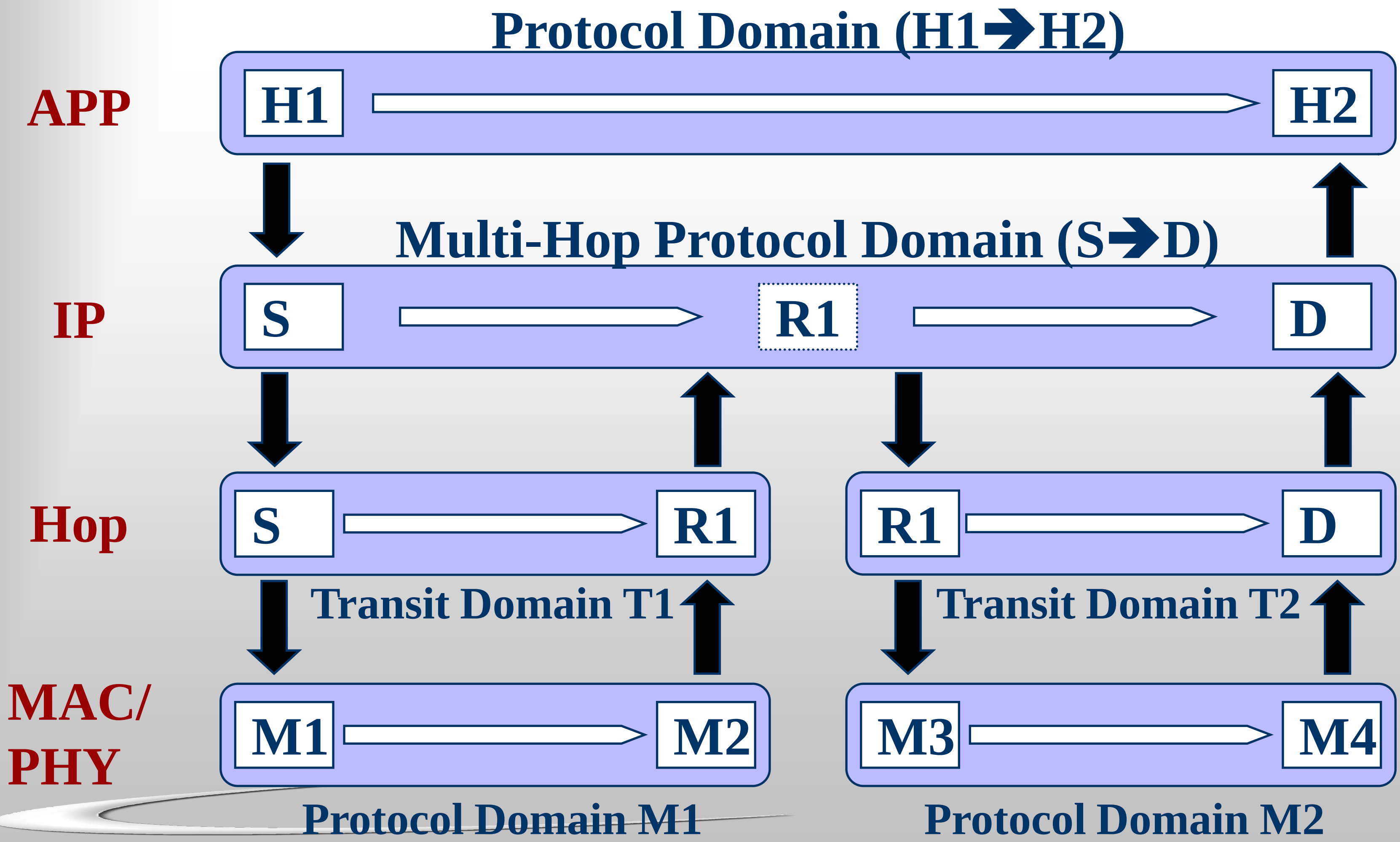
RNA and Network IPC

- Similarities
  - Recursive protocol stack
  - Unified communication mechanism
  - Focus on process-to-process interaction
- Differences
  - RNA uses MDCM to define IPC as combining a Shannon-style channel with namespace coordination
  - RNA provides a detailed (and demonstrated) mechanism that achieves unification and recursion
  - RNA supports both recursion and forwarding in a single mechanism

Copyright 2009, USC/ISI. All rights reserved.

Information Sciences Institute

Protocol & Transit Domains



Copyright 2009, USC/ISI. All rights reserved.

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**

Copyright 2009, USC/ISI. All rights reserved.