Next Steps in Enabling A Virtual Internet

Joe Touch

USC/ISI

with Lars Eggert, Yu-Shun Wang, Venkata Pingali, Greg Finn, Steve Hotz, Norihito Fujita

Information Sciences Institute

Outline

- Background
- Future Directions:
  - Lessons Learned
  - Issues
  - Impossibilities
  - Opportunities

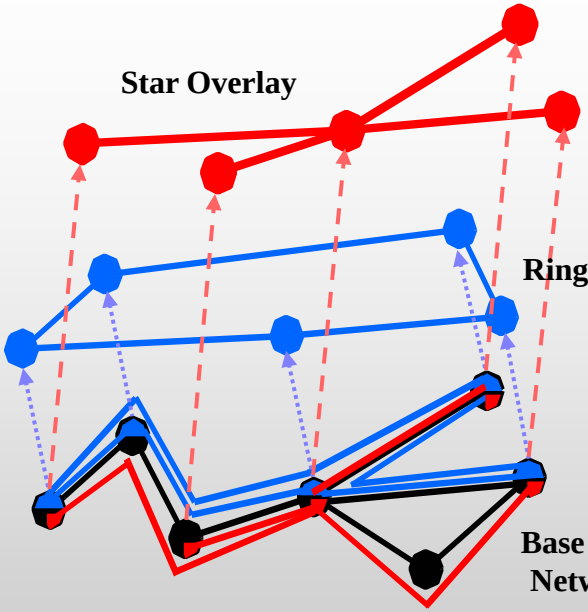
Information Sciences Institute

Background

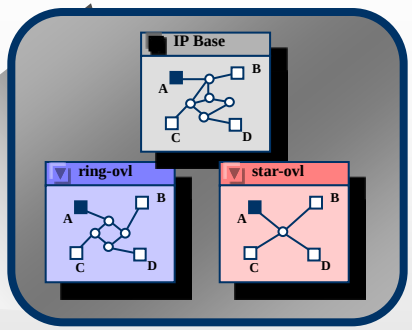
Information Sciences Institute

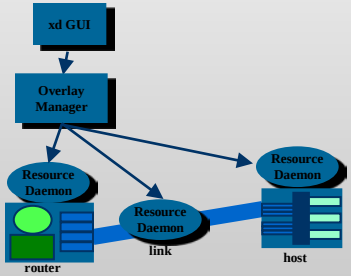
X-Bone Overlay System

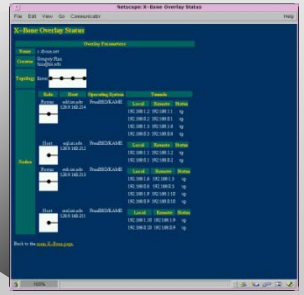
Web GUI



Multiple views





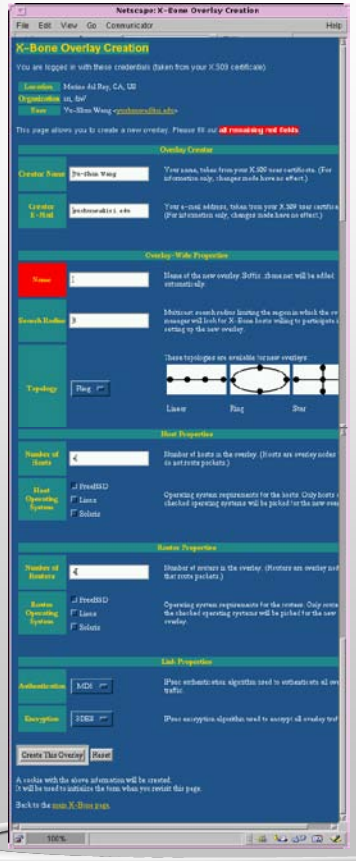


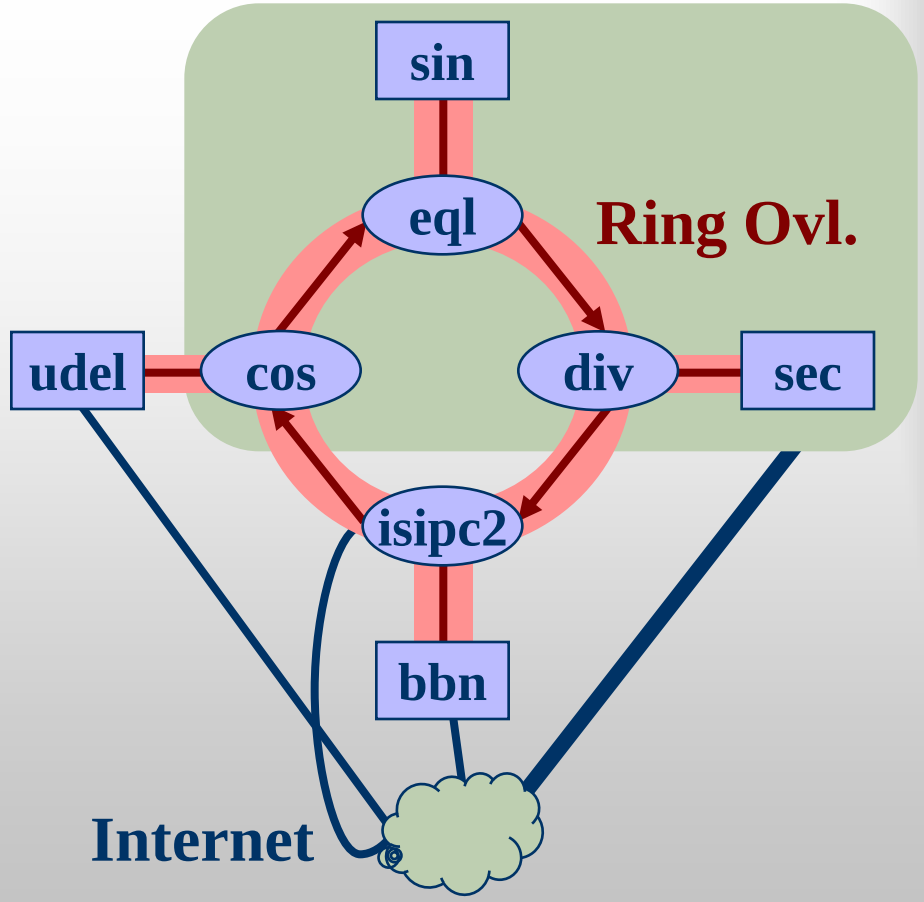
X-Bone system

Automated monitoring

Creating a Ring

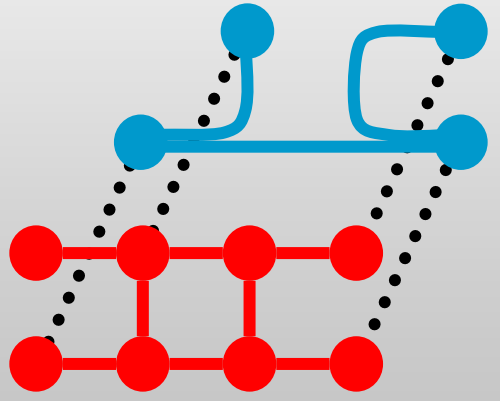
Request





VI – definition

- Virtual Internet 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 IA
  - in a virtual context
- First-principles extension
  - More than a patch
  - More than interim



Information Sciences Institute

VIA Principles

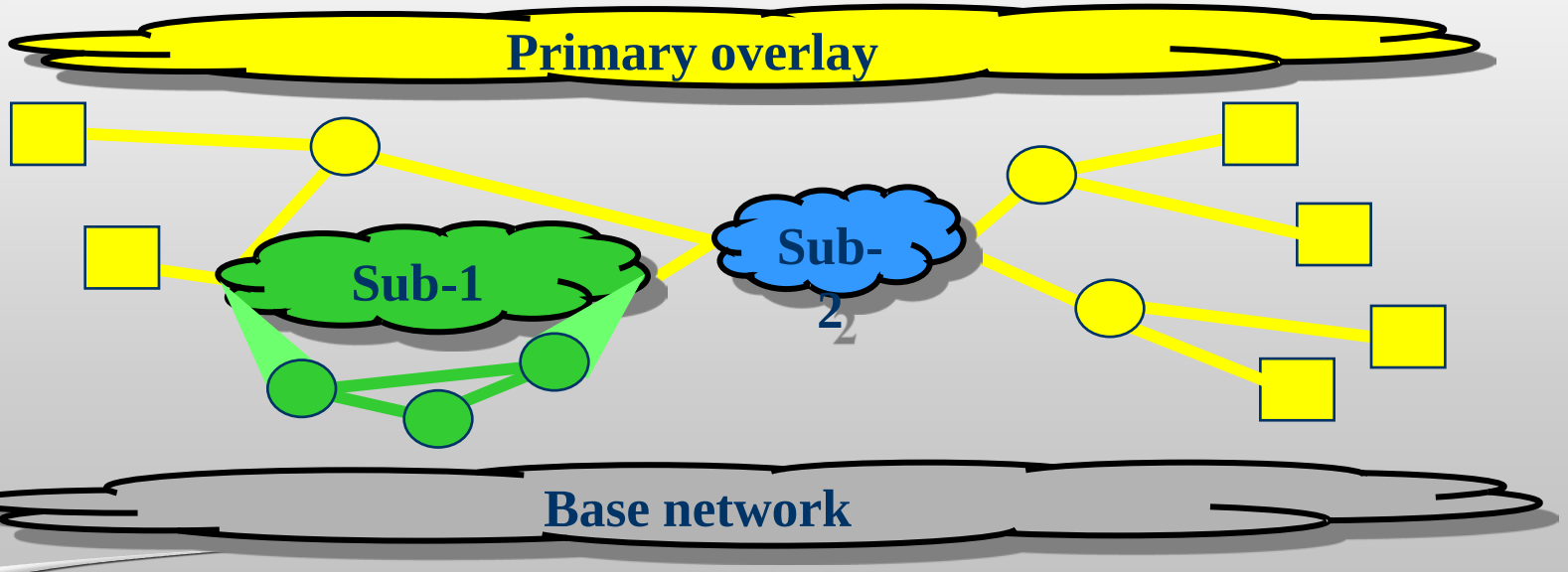
- *TENET 1. Internet-like*
  - VIs = VRs + VHs + tunnels
  - Tunnels are links; separate net addresses
  - Emulating the Internet
- *TENET 2. All-Virtual*
  - decoupled from their base network
- *TENET 3. Recursion-as-router*
  - some of VRs are VI networks

Extra 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.

Recursion-as-Router

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



Information Sciences Institute

Software Features

- Running code since 2000
  - FreeBSD port, Linux RPM, Cisco via buddy host
- Secure
  - TLS control plane, IPsec data plane
  - “Red teamed” software
- IPv4, IPv6 (both with IPsec)
  - Integrated with DNS updates, dynamic routing
- Per socket association to overlays
  - Allows process to bind to multiple overlays
- Application deployment
  - slice configuration, control distributed services
- Supports recursion, revisitation

Information Sciences Institute

What We Don't Do...

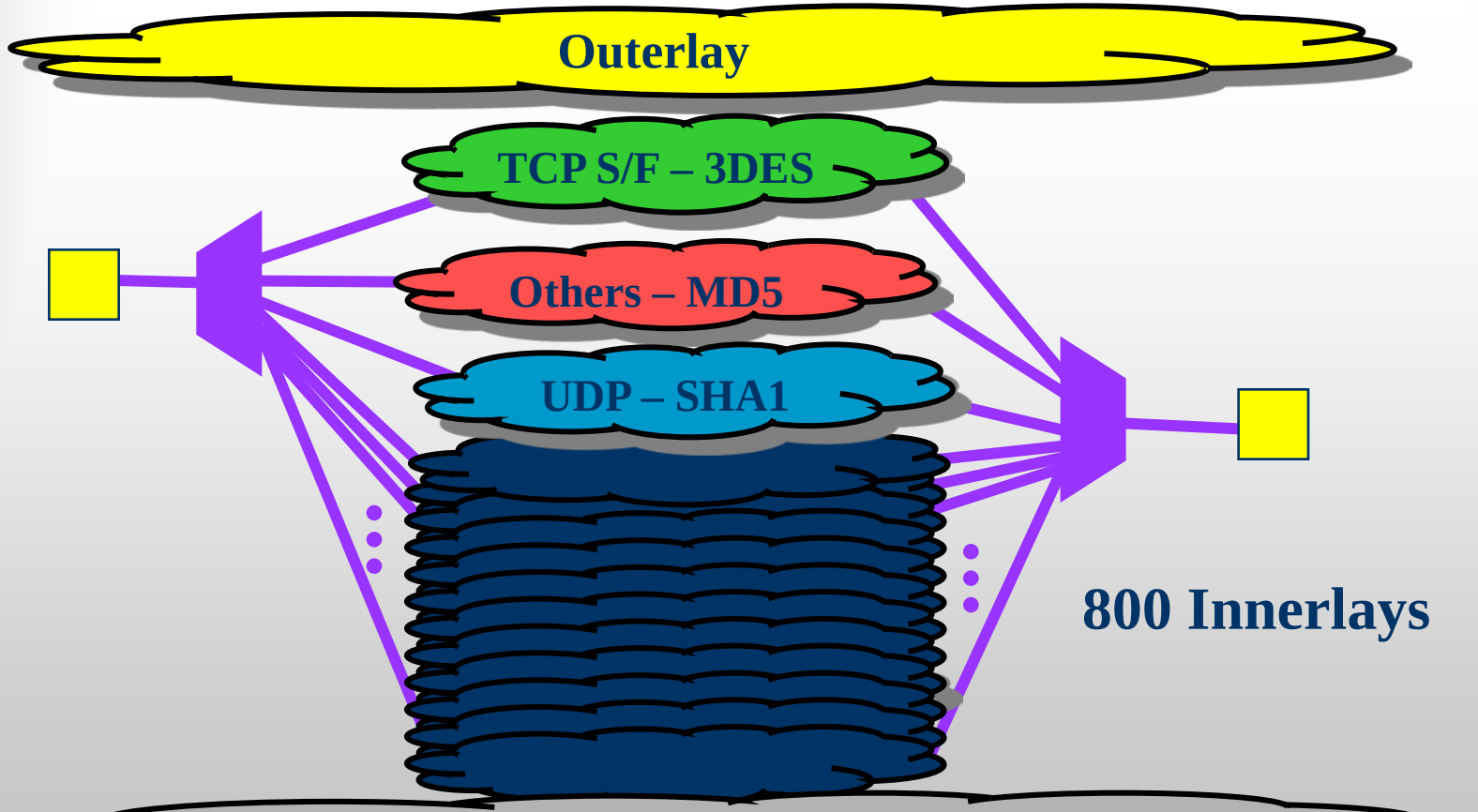
- **Optimize the overlay topology**
  - we use a plug-in module (AI folk can provide)
  - it requires network status (not quite mature)
  - fault tolerance only via ground truth (future work)
  - X-Bone is capability more than performance (now)
- **Non-IP overlays**
  - Single, common interoperation layer
  - IP recurses / stacks nicely
  - "The lowest level at which experimentation is permitted is also the highest level at which experimentation occurs." – J.Touch, 1996

TetherNet

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

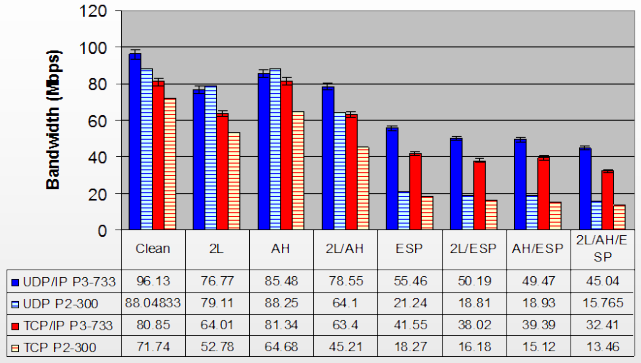
Information Sciences Institute

DynaBone Spread Spectrum

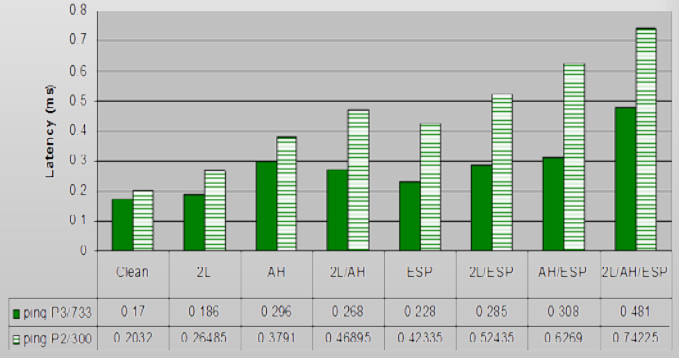


Costs of Encapsulation

- Packet MTU limits
  - Layers eat packet space
  - May stress impls.
- Bandwidth costs
  - 20% (10% IPSEC'd)
- Latency costs
  - 0.02-0.06 msec per hop



	Clean	2L	AH	2L/AH	ESP	2L/ESP	AH/ESP	2L/AH/ESP
UDP/IP P3-733	96.13	76.77	85.48	78.55	55.46	50.19	49.47	45.04
UDP P2-300	88.04833	79.11	88.25	64.1	21.24	18.81	18.93	15.765
TCP/IP P3-733	80.85	64.01	81.34	63.4	41.55	38.02	39.39	32.41
TCP P2-300	71.74	52.78	64.68	45.21	18.27	16.18	15.12	13.46



	Clean	2L	AH	2L/AH	ESP	2L/ESP	AH/ESP	2L/AH/ESP
ping P3/733	0.17	0.186	0.296	0.268	0.228	0.285	0.308	0.481
ping P2/300	0.2032	0.26485	0.3791	0.46895	0.42335	0.52435	0.6269	0.74225

Information Sciences Institute

VI 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

Future Directions

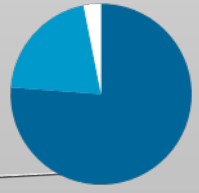
Information Sciences Institute

A Decade of Lessons

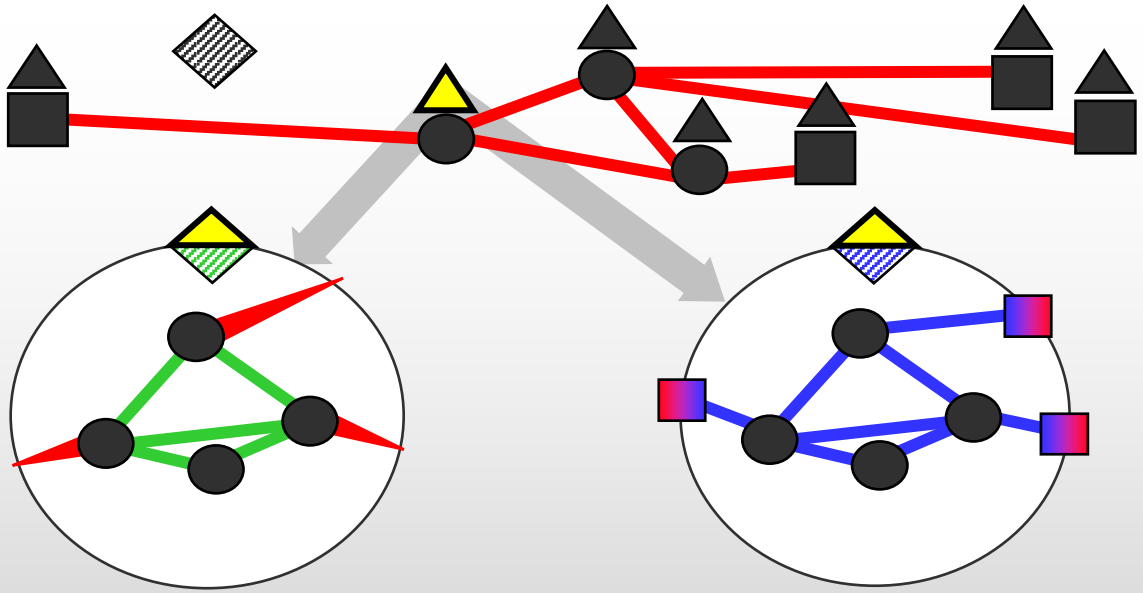
- Revisitation support
  - Two layers – Vnet, Vlink
- Recursion as map-and-encap
  - Subnet as router – TRILL, LISP
- Links as tunnels
  - Signaling interactions
- Hosts as host/router set
  - Router shares within/between overlays

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



Recursive Internet



Control / deployment ***Network***

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

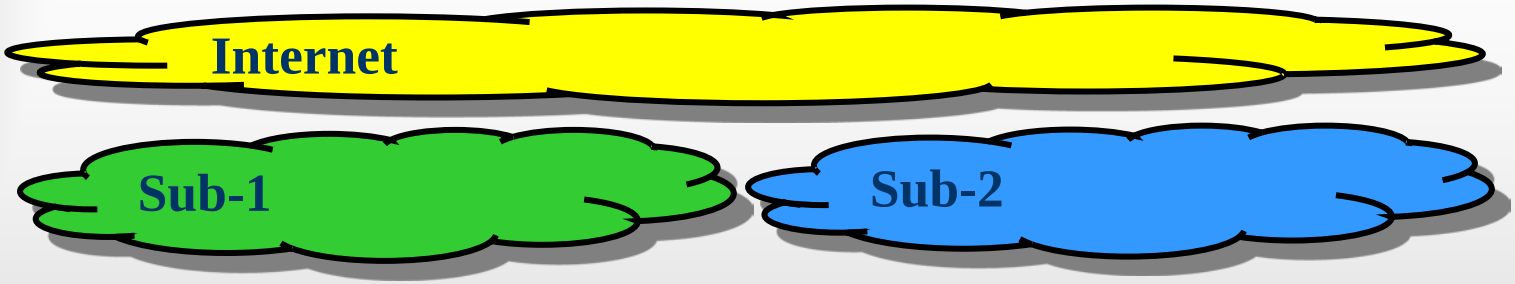
Information Sciences Institute

Issues

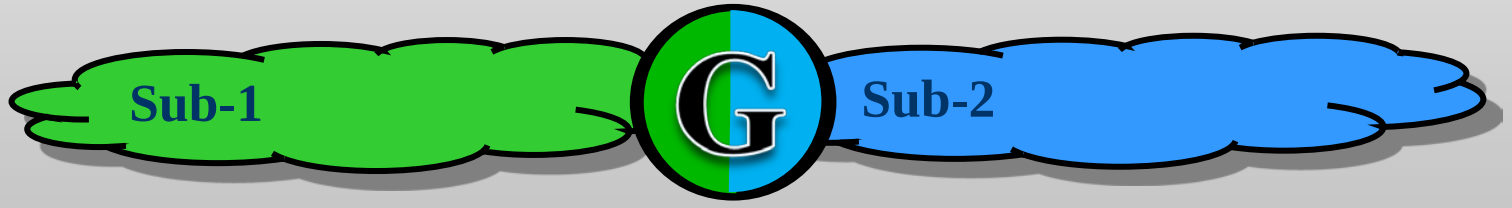
- Binding overlays to hosts
  - Per process (X-Bone), per host/OS (slice)
- Selecting an overlay
  - Impact of naming
- Supporting cross-overlay gateways
  - Translate vs. Internet

Cross-overlay

Internet:



Translate:



Information Sciences Institute

Impossibilities

- Optimization (tuning)
  - Can see path properties
  - Can't know actual path
- QoS (constraints)
  - Can tie to lower-layer resource mgt.
  - Otherwise, can enforce peer QoS only

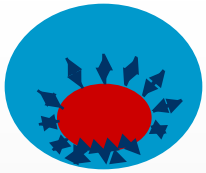
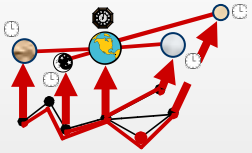
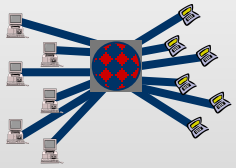
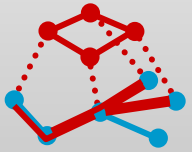
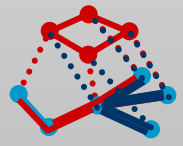
Information Sciences Institute

Opportunities

- VI as VM for daily use
  - Useful ubiquitous services
  - Not just for experiments
- Extend recursion
  - Hints at unifying general model
- Development as a full architecture
  - Host requirements for VI
  - OS extensions

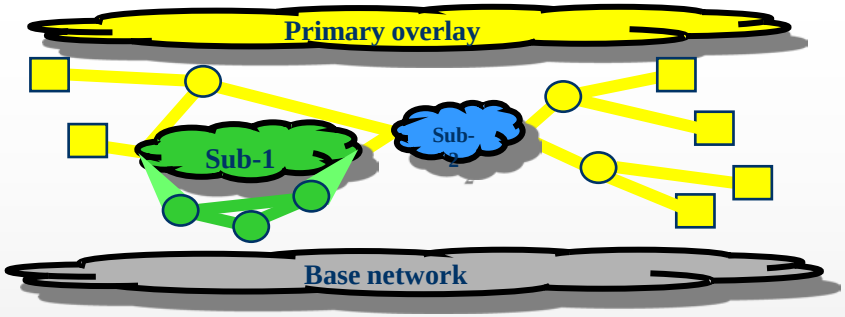
Potential Uses

- **Test new protocols**
  - Test denial-of-service solutions
- **Deploy new services incrementally**
  - Dynamic routing, proxylets, security
- **Increase lab & testbed utility**
  - Overlapping nets, add delay & loss
- **Scale to very many nodes**
  - Simplify view of topology
- **Support fault tolerance**
  - Added level of recovery

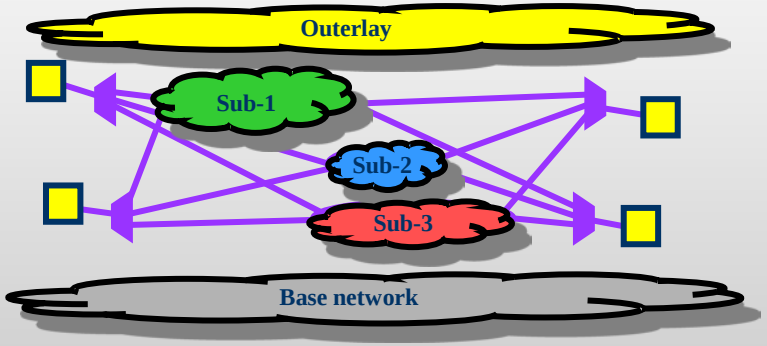






Daily Use of VIs

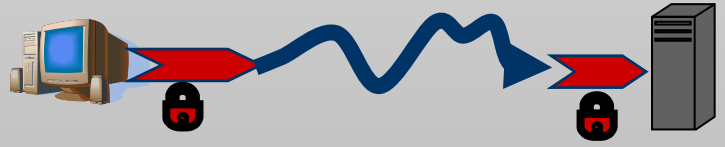
- Compose:
  - DTN, Plutarch



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



- Shift:
  - ATP



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*