

# Virtual Internets

Lars Eggert

[larse@isi.edu](mailto:larse@isi.edu)

USC Information Sciences Institute

# Talk Outline

- ▶ Virtual Internets
  - ▶ definitions + motivation
  - ▶ constraints + principles
  - ▶ consequences
- ▶ related projects at ISI
  - ▶ X-Bone, DynaBone, TetherNet, DataRouter

# Definitions

▶ network = hosts +  
routers + links

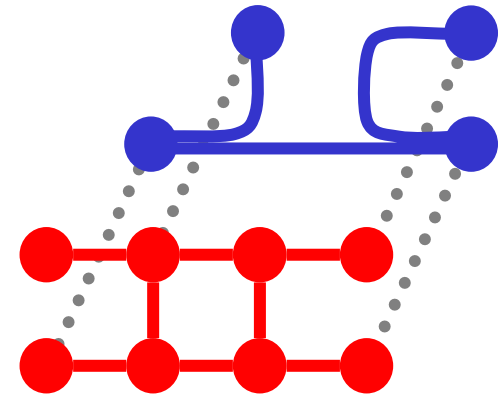
▶ virtual network =

▶ virtual host → packet source/sink

▶ virtual router → packet gateway

▶ virtual link → tunnel X over Y

▶ **virtual Internet**:  $X = \text{IP}$ ,  $Y = \text{IP}$



# Motivation

- ▶ unified, consistent architecture
  - ▶ VPNs, overlay nets, peer nets
  - ▶ isolation for concurrency + sharing
- ▶ topology-based services
  - ▶ adapt topologies to applications
  - ▶ emulate larger/different nets
  - ▶ DHT, geographic fwd, string-rewriter fwd
- ▶ layer-based services
  - ▶ customized routing, fault tolerance, security

# Constraints

- ▶ Internet-like
  - ▶ route (link up) vs. provision (link add)
  - ▶ use IP + provide IP → can recurse
- ▶ complete end-to-end system
- ▶ pass virtual net “Turing Test”
  - ▶ can’t tell it’s virtual from the inside
- ▶ support existing protocols, OS, apps

# Use Existing Mechanisms

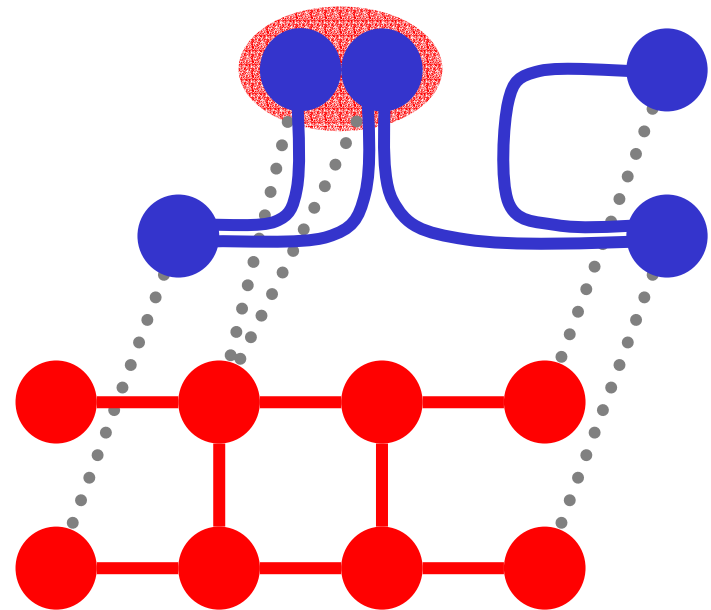
- ▶ no new protocols
  - ▶ combine existing pieces in new ways
- ▶ IP is **the** interoperability layer
- ▶ use IP mechanisms
  - ▶ ubiquitous deployment
- ▶ provide complete IP network
  - ▶ any IP app/protocol just works
  - ▶ including dynamic routing, etc.

# Principles

- ▶ node behavior
  - ▶ host: add/remove headers
  - ▶ router: transit → static # headers
- ▶ addresses indicate network context
  - ▶ separate address spaces
- ▶ complete isolation from physical
  - ▶ concurrency, recursion, revisitation

# Revisitation

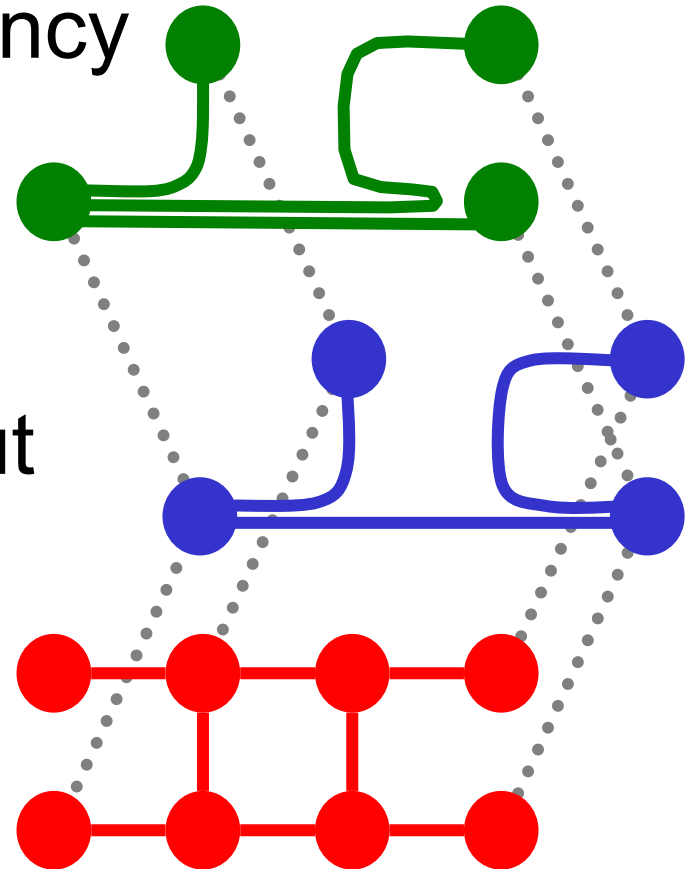
- ▶ node participates multiple times in **same** virtual network
  - ▶ possibly in different roles
- ▶ decouple virtual/physical topologies
- ▶ virtual network > physical network





# Recursion

- ▶ Virtual Internet **inside** Virtual Internet
  - ▶ indicator of consistency
- ▶ transparent reconfiguration
  - ▶ change outer without affecting inner
  - ▶ fault tolerance



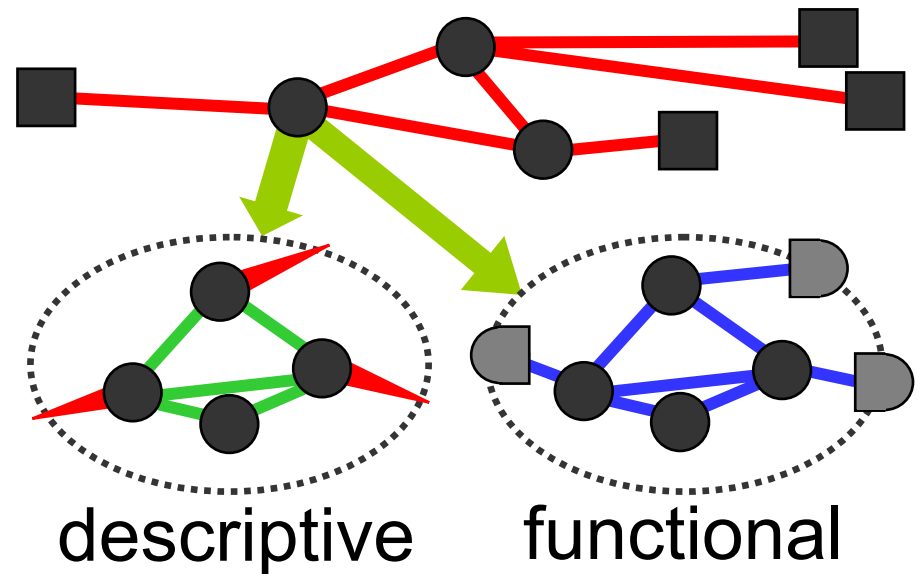
# Recursion Variants

- ▶ two kinds of virtual recursion
  - ▶ descriptive: purely syntax (macro)
  - ▶ functional: net-inside-net

- ▶ model functional

- ▶ inner virtual network acts + appears as virtual router

- ▶ BARP protocol



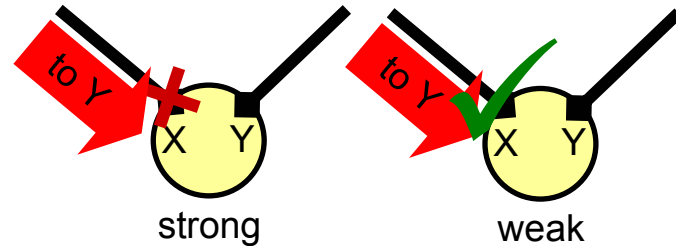
# Consequences

- ▶ double tunneling
  - ▶ Internet = link + network layers
- ▶ input context for forwarding
- ▶ implicit multi-homing
  - ▶ phantom router in all hosts
- ▶ virtualize OS + network stacks
  - ▶ INADDR\_ANY & friends
  - ▶ integrate with jail, vserver, VMware

# Two-level IPIP Encapsulation

## ▶ Internet: network of networks

- ▶ has strong link + weak network layer
- ▶ RFC1122



## ▶ two-level IPIP encapsulation

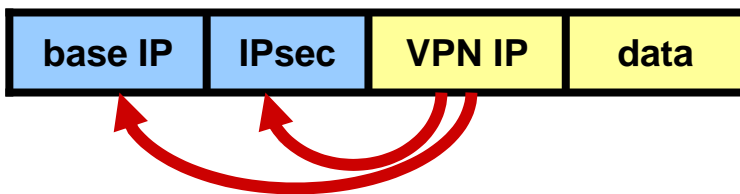


- ▶ outer tunnel: virtual link layer
  - ▶ inner tunnel: virtual network layer
- ## ▶ enables advanced capabilities
- ▶ revisitation + recursion

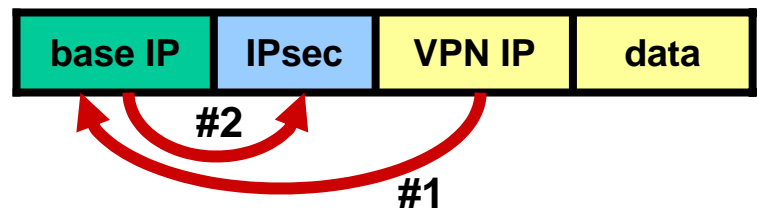
# Hop-by-hop Security

- ▶ security is link property
  - ▶ decoupled from topology
  - ▶ coexist with end-to-end security

IPsec tunnel mode



IPIP tunnel + IPsec transport mode



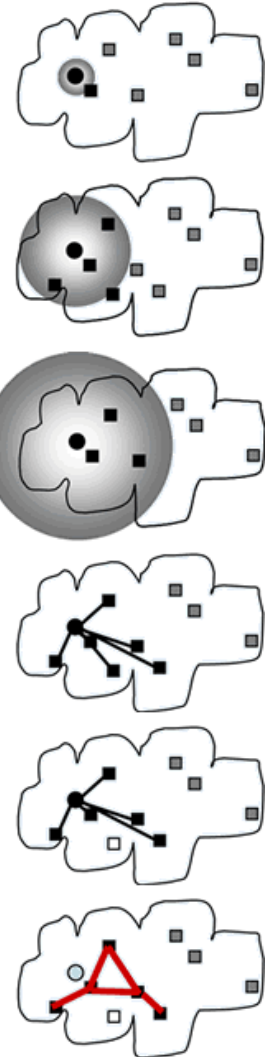
- ▶ IPIP tunnels + IPsec transport mode
  - ▶ modular tunnel mode equivalent
  - ▶ draft-touch-ipsec-vpn-05.txt

# Related Projects at ISI

- ▶ X-Bone [DARPA FTN]
  - ▶ deployment + management system
- ▶ DynaBone [DARPA FTN]
  - ▶ spread-spectrum fault tolerance
- ▶ TetherNet
  - ▶ rent real Internet behind firewall + NAT
- ▶ Others
  - ▶ X-Tend, NetFS, GeoNet, DataRouter

# X-Bone

- ▶ deployment + management system
  - ▶ programs → standardized API
  - ▶ humans → web interface
- ▶ high-level XML description
  - ▶ express virtual topology + services
- ▶ collaborating, distributed daemons
  - ▶ multicast expanding-ring discovery
  - ▶ distributed resource reservation
  - ▶ instantiate + manage virtual network



# X-Bone Screenshots

X-Bone Overlay Creation - Mozilla

X-Bone Overlay Creation

You are logged in with these credentials (taken from your X.509 certificate):

**User** Yu-Shun Wang <yushunwa@isi.edu>  
**Location** Marina del Rey, CA, US  
**Organization** USC Information Sciences Institute, Div 7

This page allows you to create a new overlay. Please fill out **all remaining red fields**.

**Overlay-Wide Properties**

**Name**  Name of the new overlay. Suffix ".xbone.net" will be added automatically. If "use DNS" is checked below, the overlay name will also become part of the DNS names of your overlay nodes.

**DNS**  use DNS If you check "use DNS", the overlay manager will assign DNS names in the OM's domain to the nodes of the new overlay. If unchecked, no DNS entries are created, and you will need to use IP addresses directly to reach overlay nodes.

**Search Radius**  Multicast search radius limiting the region in which the overlay manager will look for X-Bone hosts willing to participate in setting up the new overlay.

**Topology**  Linear These topologies are available for new overlays:  
  
Linear Ring Star

**Dynamic Routing**  use Dynamic Routing This option will determine whether to use Static Routing or Dynamic Routing within the overlay. **Only dynamic routing with RIP running GateD are supported.**

**Application Deployment**  Deploy Application Automatically deploy and start an application after the overlay has been set up. You need to specify the complete URL of the deployment script, eg. `http://`, `file://`, or (anonymous) `ftp://`.

**Host Properties**

**Number of Hosts**  Number of hosts in the overlay. (Hosts are overlay nodes that do not route packets.)

**Host Operating System**  FreeBSD  
 Linux  
 Solaris  
 NetBSD Operating system requirements for the hosts. Only hosts of the checked operating systems will be picked for the new overlay.

**Router Properties**

**Number of Routers**  Number of routers in the overlay. (Routers are overlay nodes that route packets.)

**Router Operating System**  FreeBSD  
 Linux  
 Solaris  
 NetBSD Operating system requirements for the routers. Only routers of the checked operating systems will be picked for the new overlay.

**Link Properties**

**Authentication** (None) IPsec authentication algorithm used to authenticate all overlay traffic.

**Encryption** (None) IPsec encryption algorithm used to encrypt all overlay traffic.

**Dummynet (FreeBSD only)**  
 100 ms Per-link transmission delay in milliseconds.  
 10 Kbps Per-link bandwidth limit.  
 100 Bytes Per-hop queue length limit.  
 0 % Per-hop loss probability.

Create this Overlay Reset

X-Bone Overlay Status - Mozilla

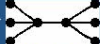
X-Bone Overlay Status

You are logged in with these credentials (taken from your X.509 certificate):

**User** Yu-Shun Wang <yushunwa@isi.edu>  
**Location** Marina del Rey, CA, US  
**Organization** USC Information Sciences Institute, Div 7

**Overlay Parameters**

**Name** line-test.xbone.net

**Topology** linear 

**Overlay Properties** Authentication Encryption Dynamic Routing Dummynet  
none 3des No No

**Creator** Yu-Shun Wang <yushunwa@isi.edu>

Role	Resource	Daemon	Local	Tunnel	End	Remote	Tunnel	End	Status
Router	cmn.isi.edu		172.26.1.2			172.26.1.1			up
	128.9.160.76	FreeBSD/KAME	172.26.1.6			172.26.1.5			up
Router	hbo.isi.edu		172.26.1.14			172.26.1.13			up
	128.9.160.75	FreeBSD/KAME	172.26.1.17			172.26.1.18			up
Host	mtv.isi.edu		172.26.1.9			172.26.1.10			up
	128.9.160.79	FreeBSD/KAME							
Router	sci.isi.edu		172.26.1.10			172.26.1.9			up
	128.9.160.93	FreeBSD/KAME	172.26.1.18			172.26.1.17			up
Host	tjc.isi.edu		172.26.1.1			172.26.1.2			up
	128.9.160.31	FreeBSD/KAME							
Host	tnn.isi.edu		172.26.1.5			172.26.1.6			up
	128.9.168.57	FreeBSD/KAME							

Back to the [main X-Bone page](#).



# X-Bone Status

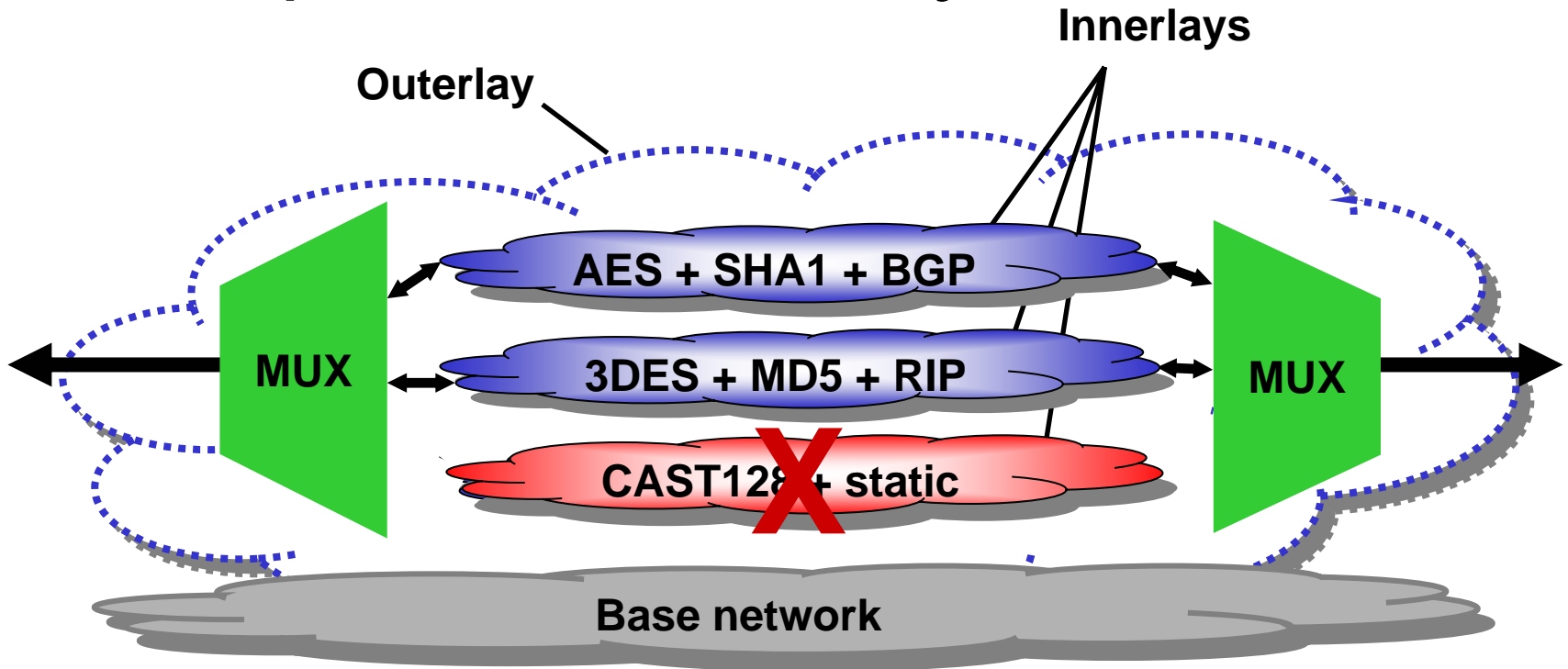
- ▶ current release: 3.0
  - ▶ mature: 5+ years availability
- ▶ platforms: FreeBSD, Linux
  - ▶ unofficial: NetBSD, Cisco
- ▶ actively maintained (X-Tend)
  - ▶ IPv6, Linux 2.6 IPsec, DDNS, DNSSEC
- ▶ widely used
  - ▶ UCL, UPenn, Army, Navy, Aerospace, DOD Canada, Sinica Taiwan, etc.

# DynaBone

- ▶ architecture
  - ▶ multiple, parallel inner virtual networks
  - ▶ algorithmic and protocol diversity
  - ▶ spread-spectrum multiplexer
  - ▶ wrapped inside outer virtual network
- ▶ innerlay: gracefully disconnectable
  - ▶ attacker-like parallelism as a defense
- ▶ outerlay: hide innerlays from apps
  - ▶ allow transparent restoration

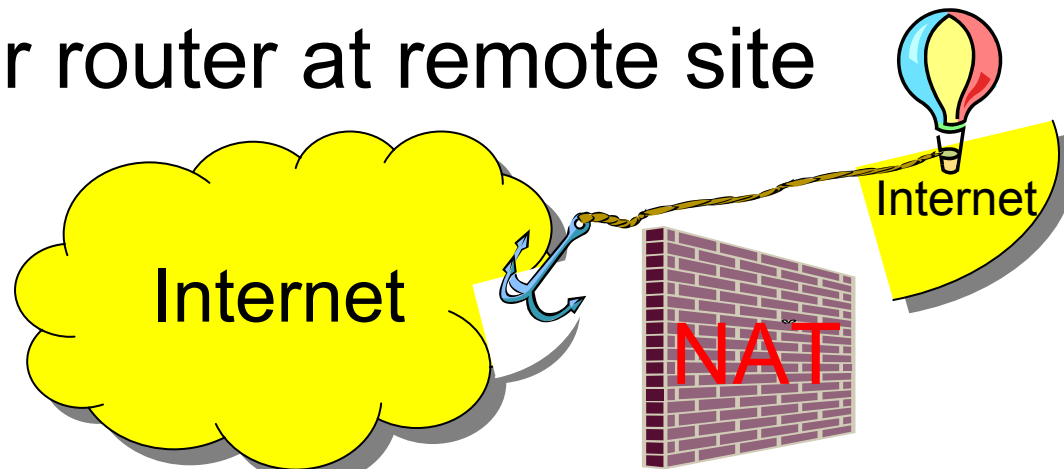
# DynaBone Overview

- ▶ mux transparently disconnects compromised innerlay



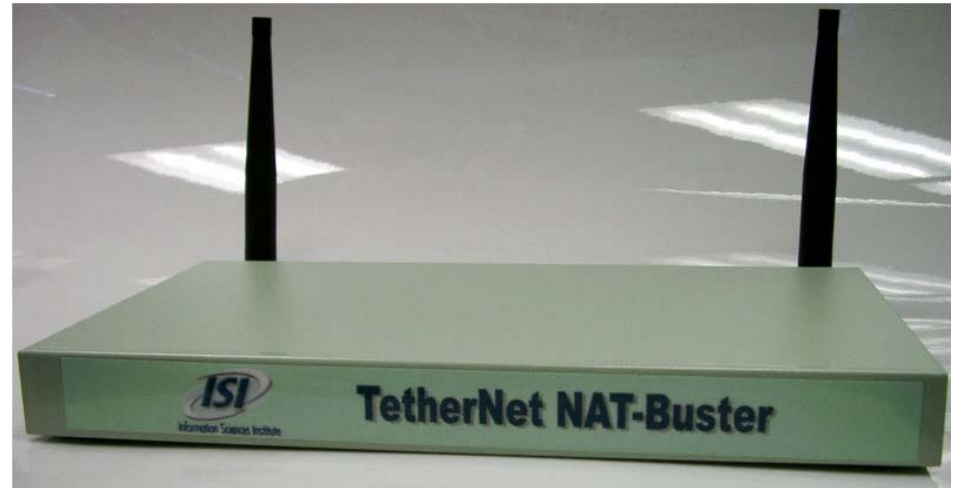
# TetherNet

- ▶ issue: firewalls, NATs, clueless ISPs
  - ▶ broken end-to-end connectivity
- ▶ solution: relocate **real** Internet subnet
  - ▶ real = routable IP + DNS + no fw + ...
  - ▶ tunnel subnet from anchor router to tether router at remote site



# TetherNet Features

- ▶ true Internet behind NATs and firewalls
  - ▶ IPv4 + IPv6
  - ▶ multicast
  - ▶ fwd/rev DNS
  - ▶ traffic shaping
  - ▶ 802.11b AP
  - ▶ secure: IPsec for traffic, X.509 for user auth
  - ▶ web interface configuration
- ▶ U.S. patent filed, talks with licensees



# TetherNet Screenshots

## TetherNet Rental

### Required rental parameters:

Rental Site	<input type="text" value="Marina del Rey, USA"/> <input type="text" value="198.32.16.91"/>	Pick a preconfigured TetherNet rental site close to you, or specify the IPv4 address of a custom one.
Subnet Size	<input type="text" value="9"/> hosts	Effective usable subnet size of the new TetherNet. Choose a large enough size for the planned number of client end hosts.
Access Code	<input type="text"/>	Some TetherNet rental sites require access privileges. If you have been provided with an access code for a rental site enter it here, otherwise leave empty.

Start TetherNet Service

### Optional rental features:

Relay Type	<input type="radio"/> TCP <input checked="" type="radio"/> UDP <input type="radio"/> IPv4	Local Port: <input type="text" value="auto"/> Remote Port: <input type="text" value="auto"/>	Use specified relay method for TetherNet. Local and remote ports are only meaningful for TCP and UDP relays, and may be set to <i>auto</i> if no specific port setting is required to pass middleboxes.
Relay Encryption	<input type="checkbox"/> encrypt with	<input type="text" value="aes"/>	Optionally, the traffic between the TetherNet box and the rental site can be encrypted - <u>this does not provide end-to-end security.</u>

### Optional advanced networking features:

IPv6	<input checked="" type="checkbox"/> enable	Enable IPv6 routing on the TetherNet, including autoconfiguration. IPv6-aware end hosts receive IPv6 addresses automatically through router solicitation.
Multicast	<input checked="" type="checkbox"/> enable IPv4 <input checked="" type="checkbox"/> enable IPv6	Configure IPv4 and/or IPv6 multicast connectivity for the TetherNet.
DHCP Server	<input checked="" type="checkbox"/> enable Range: <input type="text"/>	Start a DHCP server on the LAN interface, enabling end hosts to dynamically request IPv4 addresses. The <i>Range</i> field specifies how many IP addresses at the bottom of the allocated subnet block are handed out via DHCP (the rest are available for static assignment.)

## Rental Server Response

### Rental information:

Rental Server	anchor.postel.org <larse@isi.edu>
Organization	USC/ISI, TetherNet
Location	Marina del Rey, CA, US
Local Time	Tue Sep 17 15:13:00 2002

### Rented network block:

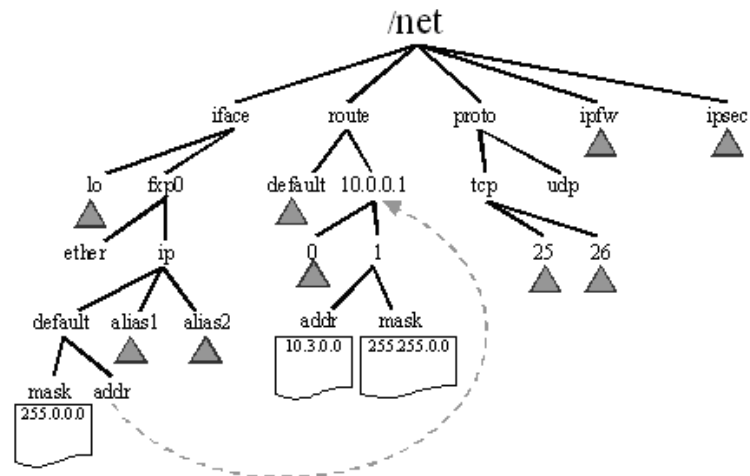
IP Block	206.117.27.16
Size	16

### TetherNet properties:

Rental Site	198.32.16.91
LAN Size	206.117.27.16/28, 9 hosts, IP addresses 206.117.27.22 - 206.117.27.30
DNS Suffix	tethered.net
Tunnel Type	UDP (local port 35770, remote port 34213)
DHCP Service	on, handing out the range from 206.117.27.22 to 206.117.27.22
Tunnel Encryption	rijndael-cbc
IPv6	on, allocated prefix is 3ffe:825:117:27:16::/64
IPv4 Multicast	on
IPv6 Multicast	on

# Other Projects

- ▶ X-Tend [NSF]
  - ▶ maintain + extend X-Bone as tool for research + education
- ▶ DataRouter [internal]
  - ▶ source-route-like fwd based on payload data
- ▶ GeoNet [internal]
  - ▶ geographically-addressed overlays
- ▶ NetFS [NSF]
  - ▶ access control for the network stack via pseudo file system



# Questions

`larse@isi.edu`

`http://www.isi.edu/larse/`



# FAQ

- ▶ why not VPN, P2P or other?
  - ▶ most net-level is incremental, partial, etc.
  - ▶ app- level recapitulates network + doesn't compose
- ▶ isn't this more complex?
  - ▶ AS-like management encapsulation (multi-level)
  - ▶ can make application view simpler
- ▶ isn't this suboptimal?
  - ▶ so is VM: like VM, OOB info + direct measurements can help
  - ▶ layering implies increasing coarseness
- ▶ wasn't this done in X before?
  - ▶ this is uniform, consistent + *implemented*