Virtual Internets

Lars Eggert
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 = IP, Y = 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**

```
base IP | IPsec | VPN IP | data
```

**IPIP tunnel + IPsec transport mode**

```
base IP | IPsec | VPN IP | data
```

- 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 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

Base network

**Innerlays**

- AES + SHA1 + BGP
- 3DES + MD5 + RIP
- CAST128 + static

**MUX**

**Outerlay**
**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**: Marina del Rey, USA, 198.32.16.91
- **Subnet Size**: 8 hosts
- **Access Code**: Blank

Optional rental features:

- **Relay Type**: TCP/UDP/IPv4
  - **Local Port**: Auto
  - **Remote Port**: Auto
  - **Relay Encryption**: Encrypt with

Optional advanced networking features:

- **IPV6**: Enable IPv6 routing on the TetherNet, including autoconfiguration. IPv6-aware and hosts receive IPv6 addresses automatically through router advertisement.
  - **Multicast**: Enable IPv4 and/or IPv6 multicast connectivity for the TetherNet.
  - **DHCP Server**: Enable MULTICAST

Rental Server Response

Rental information:

- **Rental Server**: anchor.post.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**: 206.117.27.18
- **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-128
- **IPv6**: On, allocated prefix is 3ffe:825:117:27:16::/64
- **IPv4 Multicast**: On
- **IPv6 Multicast**: On

September 2, 2003  Lars Eggert <larse@isi.edu>
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

September 2, 2003 Lars Eggert <larse@isi.edu>
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