Research on Future Internet Architecture

NSF GENI and FIND Program

xiangyang zhang, xiang@cs.queensu.ca
CONTENTS

• Why research on FI?
• Who researches FI?
• Research Topics
• GENI Backbone Design
Why Research on FI?

• A group of computer scientists believe “there are serious problems facing the Internet”
• Most of the users have some unhappy experiences using the Internet
Problems of the Internet

• Security
  – We don’t want to interact with the bad guys
  – We don’t trust the guys we actually interact
  – We don’t believe the software we use

• Mobility
  – Number of Internet hosts as of 2008: 541 mil
  – Number of mobile users as of 2005: 2.1 bil
Problems of the Internet, cont’d

• Availability
  – Four 9s at best, compared to five 9s of PSTN

• Management
  – Hard to configure and trouble-shoot for operators
  – Clumsy to use for layman

• Economic issues
  – QoS vs. network neutrality, inter-ISP QoS
  – Users’ choice for local and LD ISPs
Who Research on FI?

• US: NSF CISE
  – Future Internet Design (FIND)
  – Scientific Foundations for Internet's Next Generation (SING)
  – Next Generation Networked Information (NGNI)
  – Global Environment for Network Innovations (GENI)

• EU: FP7
  – Dozens of small projects

• IETF
  – IRTF Routing Research Group (RRG)
  – IRTF End-to-End Research Group
GENI

• Vision: A national facility to explore radical designs for a future global networking infrastructure
• An NSF Major Research and Equipment Facility Construction (MREFC) project
• Incorporate technologies for networks and distributed services within a 10-20 year time frame, specifically CPU & disk farms, programmable ‘routers’, optical networks, and wireless access
FIND

• Innovate in the large vs. Innovate in the small
  – CISE’s first transformative research initiative
  – Dozens of small projects that are complementary and are focused on future Internet architecture

• Most PIs are from universities
  – MIT, CMU, UC Berkeley, Stanford, etc.
FP7

- The Seventh Framework Programme for research and technological development (FP7) is the European Union's chief instrument for funding research over the period 2007 to 2013

- Future Internet Conference Bled 2008
Research Facilities

• Many countries have government sponsored networks for research and education institutions
  – Internet2, US: 200 edu, 70 com, 45 gov, 50 int’l
  – GEANT2, Pan-EU: 30 national nets, 3500 org
  – APAN, Pan-AP
  – CANARIE, CA

• PlanetLab seems to be the closest thing to GENI
  – Sponsored by NSF
  – Many tests running
CONTENTS

• Why research on FI?
• Who researches FI?
• Research Topics
  • GENI Backbone Design

  • Design Principles
  • Name and Address
  • Mobility
  • Routing
  • Security
  • Multicast and Anycast
New Technologies

• Wireless
  – Moving terminals: mobility, intermittent connectivity
  – Mesh network: stationary, no infrastructure
  – Sensor network: power-constrained

• Optical network
  – Granularity: fiber, color, burst, packet

• Storage
  – Data center
  – P2P file sharing
FIND Research Topics

• Principles: D. Clark
• Name and address: 20+ proposals
  – Separation of name and address
  – Name verification
  – Name lookup
• Routing
  – BGP enhancement
  – Source routing
• Packet and multiplexing
• Network Management
• Virtualization of network

Where are security and mobility?
Internet Design Principles

Principles
1. Communication despite loss of networks
2. Support multiple types of services
3. Support heterogeneous networks
4. Distributed management
5. Cost effective
6. Easy host attachment
7. Accountable

Implications
1. Stateless network, end-to-end argument
2. Separation of TCP and IP
3. Datagram, hourglass model
What Has Changed?

In the Past: **point-to-point** communication between **fixed locations** among the **research community**

Now the Internet is

- Used by mainstream society
  - economy issue, lack of trust
- For all kinds of communication
  - point to point
  - point to a group
  - point to any point of a group
- With fixed and mobile terminals
  - mobility
  - home net, sensor net, etc.
FI Design Principle

• Design to allow the tussles to happen
• Hourglass model
  – fine-grained statistical multiplexing
  – TCP: how many kinds?
• Separation of identity and location
  – Addressing scheme: probably topology-based
  – Naming scheme? Early binding or late binding?
• End-to-end argument
  – Controlled transparency
• Management and security be first-class citizen

No consensus on priority!


Name and Address

• Benefits of separation of identification and location
  – **Mobility**: interact with an entity, not a location
  – **Security**: self-certifying name
  – Multi-homing, change ISP without changing names, etc.

• How to deliver a packet to a name?
  – Carry name or address in packet headers?
  – Down to the bottom, a packet has to be delivered to an address, not a name.

• Name lookup
  – Name is not topology-based (usually address is)
  – Name space may not be hierarchical (flat name space)
Mobility

• Care-of approach
  – Send packets to your mailing address, someone who knows where you are forward it to you
  – Mobile IPv6 and IPv4

• Rendezvous-based approach
  – Send packets to a rendezvous (packets bear name), the rendezvous forwards it to you
  – The rendezvous may be an overlay network (e.g. i3 by Ion Stoica, etc)
  – The rendezvous may forward packets (late binding, e.g. i3), or be a DNS-like name lookup service (early binding)
Internet Indirection Infrastructure

- Each packet is associated an identifier $id$
- To receive a packet with identifier $id$, receiver $R$ maintains a trigger $(id, R)$ into the overlay network
- Host just needs to update its trigger as it moves from one subnet to another

By Ion Stoica, 2003
i3 Overlay

- How to locate the i3 node with name “ident”? 
- Distributed Hash Table 
- Chord chosen for implementation, CAN, Tapestry, Pastry also possible.

Finger Table for N8

<table>
<thead>
<tr>
<th>N8+1</th>
<th>N14</th>
</tr>
</thead>
<tbody>
<tr>
<td>N8+2</td>
<td>N14</td>
</tr>
<tr>
<td>N8+4</td>
<td>N14</td>
</tr>
<tr>
<td>N8+8</td>
<td>N21</td>
</tr>
<tr>
<td>N8+16</td>
<td>N32</td>
</tr>
<tr>
<td>N8+32</td>
<td>N42</td>
</tr>
</tbody>
</table>

1/13/2012

By Ion Stoica, 2003
Self-Certifying Names

- Asymmetric key system
  - Public key, advertised publicly as name
  - Private key, kept secret
- Signature to prove “you are the holder of the name”
- 3rd party to prove “you are the person with the name”
- Append digest to guarantee integrity of message
- Proposals:
  - “A New Approach to Internet Naming and Name Resolution”, Scott Shenker and Ion Stoica, UC Berkeley
  - Accountable Internet Protocol (AIP)’, David Andersen, etc.

• Separate names for verification and lookup or not?
• Lookup in flat name space
• Possibility of a universal name service like DNS?
• Possibility of late binding
Routing

• Reduce routing table size
  – Now ~250k routes, growing 20k per year
  – Reason: not so topology-based, multi-homing

• Secure routing protocol
  – Independent AS and IP address: RIR, new addressing scheme like IPv6
  – Verify identity of BGP router: sBGP, soBGP

• Source routing
  – Give users the choice
Routing Scalability

• Switched Internet Architecture
  – Nirmala Shenoy and Nirmala Shenoy, Institution(s): Rochester Institute of Technology
  – PSTN-like addressing

• Greedy Routing on Hidden Metric Spaces as a Foundation of Scalable Routing Architectures without Topology Updates
  – Dmitri Krioukov, KC Claffy, Kevin Fall; CAIDA/UCSD, Intel, Berkeley
  – Make routing static, no topology updates
Source Routing

• Why it is difficult?
  – ISP unwilling to reveal routing information
  – Economic issue: agreement among ISP, payment by users

• An Internet Architecture for User-Controlled Routes
  – Xiaowei Yang; Institution(s): UC Irvine

• Separate route computation: IETF PCE WG

• Secure routing scheme (BGP)

• New scheme for scalability, like
  no-update routing, traffic diffusion

• Source routing

• Economic issues
Security

• Identity management
  – self-certifying address, introduction (CA)

• Secure routing
  – sBGP, soBGP, RIR

• Secure name resolution
  – DNS SEC extension, anycast, etc.

• Support for end-host security
  – network should assist the host security
Host Security

• Source of attacks
  – Exploit software vulnerability: routers stronger than servers, servers stronger than hosts
  – DoS, DDoS

• Network assistance
  – Firewall
  – Enforce security policy on hosts, e.g. Cisco NAC
  – Communication is no longer transparent, middle boxes have states
An Example

• David Anderson, etc. “Holding the Internet Accountable”, “Accountable Internet Protocol (AIP)”

• Address format: $AD:EID$
  – AD is the hash of public key of the autonomous domain, border routers have private key
  – EID is the hash of public key of the host; hosts have private key

• Routing
  – AD for Interdomain routing, EID for intradomain routing, non-topology based
  – A border router can only advertise EID within its own AD as originator
  – A border router must append signature when announcing routes

• Host identity
  – A host must sign the first packet, the gateway bind EID with MAC
  – Not very secure, but it is impractical for hosts to sign every packet.

self-certifying address
same idea as IPv6
An example, cont’d

- **Multihoming**
  - A host connects to two ADs has address AD1:EID, AD2:EID

- **Mobility**
  - When a host moves, the AD part changes, the EID part remains.

- **Scalability**
  - Size of intradomain routing table: e.g. 100k
  - Size of interdomain routing table: e.g. 1 mil
  - Fixed length key, lookup by cache-hit
  - Flat address, how to build up routing table?

Name and address scheme is a fundamental decision!

- same idea as IPv6
- need to register to HA as IPv6
- lookup ok, routing risky
Multicast

• Multicast failed to emerge as a service. Will it resurrect?
• Economic issue: If ISP charge by volume, multicast reduces ISP's revenue; inter-ISP charging
• Application issue:
  – CDN reduces the need for multicast (non-live contents)
  – Multicast can be implemented in the overlay network
  – Video conference: multicast is not necessary if participants are not many. Overlay network multicast may be better.
  – Multicast has merit in delivering live content in large scale, like IPTV. But face competition from MSO and P2P live TV
• Technical issue
  – Soft-state in the network
  – Traffic management, QoS, accounting
Content Distribution

• Abstraction: point to any point of a group
• Application
  – Popular web service: Yahoo, CNN, etc.
  – Video-on-Demand, Youtube, DNS, etc.
• Benefits
  – Offload original servers’ burden, reduce user perceived latency, reduce network traffic, ...
• Methods on today’s Internet
  – anycast: e.g. DNS root server
  – DNS-based: e.g. Akamai
  – URL rewriting: e.g Akamai
• Possible methods on FI: route by name
Recap: Open Topics

- Name and address schemes
- Mobility, identity
- Route by name: routing and lookup
- Routing: scalability, secure BGP, source routing
- Content distribution scheme
- Not discussed today
  - management, trouble shooting, auto-config
  - forwarding: QoS, small buffer router
  - packet aggregation
  - FL “TCP”
  - spam, worm, and a lot more...
CONTENTS

• Why research on FI?
• Who researches FI?
• Research Topics
• GENI Backbone Design
Resource and Slicing

• Physical network resources
  – Link: probably from National Lambda Rail Network, 2~4 fibers, 25 sites across US
  – Router: need to be developed
  – Disk, CPU, wireless net: opt-in from research institutions

• Slice the physical network resources
  – Virtual link
  – Virtual router
  – Virtual disk/CPU/wireless net
How to Build the Substrate?

• Slicing
  – It is easy to slice Disk, CPU, wireless nets
  – Virtual links could be provisioned, CIR/PIR, or BE
  – But how to slice a router?

• Two proposals for GENI Backbone Network (GBN)
  – “An Architecture for a Diversified Internet”, Jonathan Turner, Patrick Crowley, Sergey Gorinsky and John Lockwood, Washington University in St. Louis
  – “CABO: Concurrent Architectures are Better Than One”, Nick Feamster, Lixin Gao and Jennifer Rexford, Georgia Tech, UMass Amherst and Princeton Univ.
Router Architecture, Distributed

You can do nothing fancy on the packets!
Router Architecture, Centralized
Dr. Turner’s Proposal

- ATCA chassis with power
- Crossbar Switch Fabric
- Line Card (LC)
  - All packets are tagged with meta-router ID and meta-interface ID
  - On receiving a packet, LC lookup tag table to switch the packet to its related PE
  - When receiving or sending a packet, LC performs CAR, so the meta-interface can have provisioned, CIR/PIR, or BE bandwidth
- Processing Engine (PE)
  - Each user is assigned a dedicated PE
  - PE strip off the tag, lookup routing table, add new tag, and forward the packet to LC
  - PE run routing protocols with its peers of virtual routers.
  - Theoretically, users can install their own OS on their PE
- Shelf-Management Module
  - Configure LC and PE
  - Allow users to configure their meta-router and meta-interface
Dr. Turner’s Proposal, cont’d

• Where to get the hardware?
  – All parts are commercially available, but LC, SF, PE must work together.

• How to support large number of researches?
  – The number of PE are limited
  – Slicing LC is extremely difficult if allowing flexibility for users

• Performance is much lower than commercial routers

by Jonathan Turner
GENI Participants

• GENI roles
  – Network Science and Engineering Council (NetSE): Definitive source of “what we need in GENI”, Authors of GENI Research & Education Plan
  – GENI Project Office (GPO): Project management and execution, Authors of GENI facility construction plan.
  – BBN perform the work of the GENI Project Office (GPO).

• Engineering conference and WG meetings every four months
GENI Timeline

Planning Phase
3-4 years

Construction Phase
5 years

Operations Phase
TBD years

June 2007

GENI Engineering Conferences
Solicitations issued for new prototypes & trials

9 months

Early 2008

18-30 months

12 months

www.geni.net
Q&A