Regional Opt-In

Matt Mathis

Pittsburgh Supercomputing Center (PSC)

and

Three Rivers Optical Exchange (3ROX)

GEC 4

2-Apr-2009

Regional Opt-In

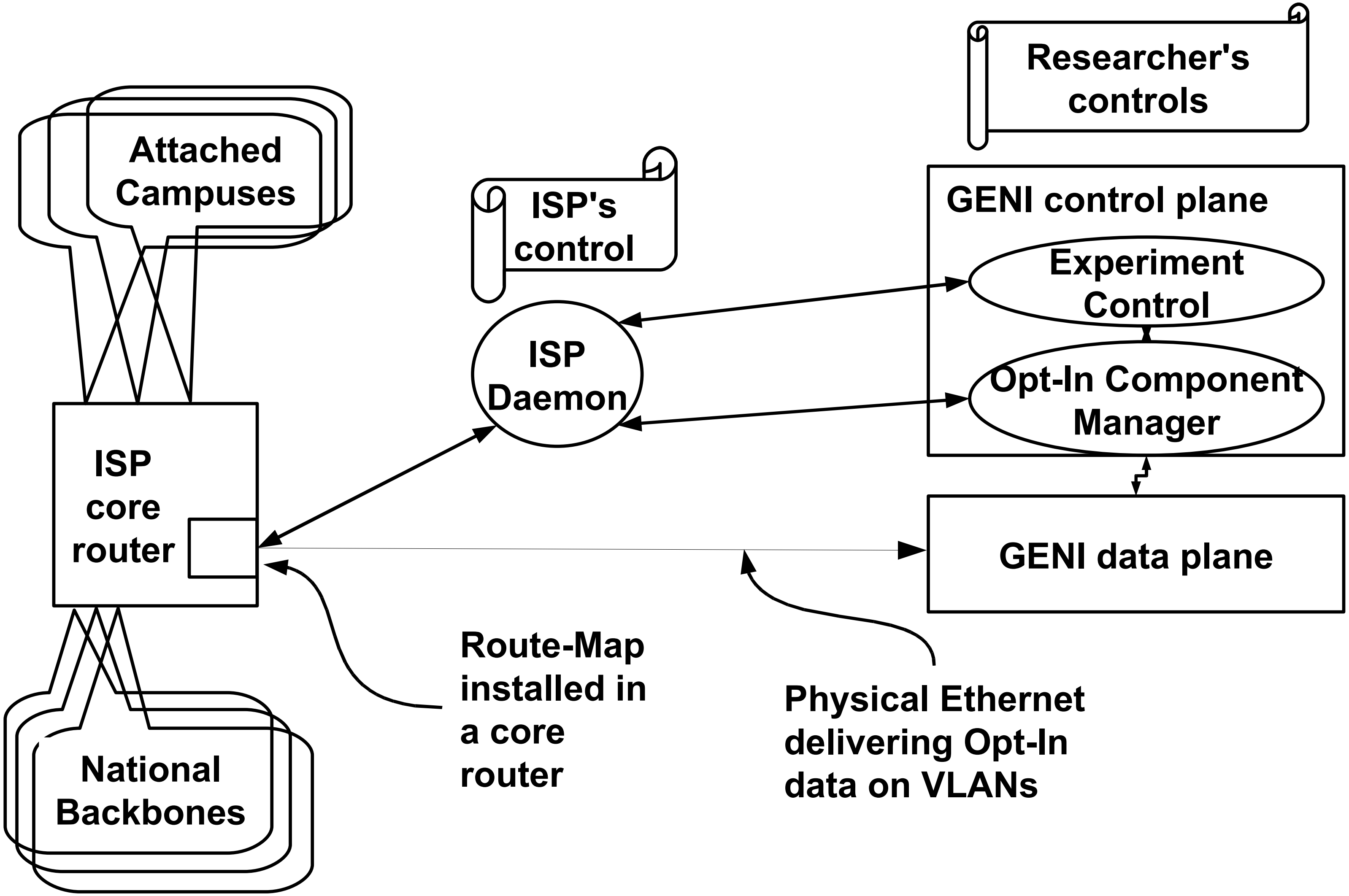
- Interpose GENI at regional interconnects
- Intercept “innocent user” traffic
  - AKA wholesale Opt-in
  - Completely authentic traffic
  - Intrinsically Layer 3
- Want a strongest possible position:
  - (Eventually) Ask NSF to encourage participation
    - Progressive: Specific programs, then CNS, CISE, all NSF
  - Complete control of user impact
    - Minimize unexpected consequences
      - Avoid outages
      - Avoid leaking PII (Personally Identifiable Information)
  - Fully motivate all actors

Outline

- Technical Overview
- Scaling really large
- Constraints as seen from:
  - GENI (Researchers)
  - IRB (Opt-In users)
  - ISP (other users, aka customers)

See .....wiki/RegionalOptIn/OptInReqs.pdf

Regional Opt-In



Implementation at 3ROX

- Non-profit GigaPoP run by PSC
  - CMU, PITT, PSU, WVU, most k-12 in western PA
  - Libraries, museums, etc
  - Some commercial sites
  - Roughly 200k users
- Connections to multiple backbones
  - NLR, I2, (ETF), NLR transit rail
  - Sprint, Global Crossing
- Redundant core routers
  - Approximately $250k each

Opt-In Intercept

- Use OpenFlow, route-map, or firewall-filter
  - ACL style packet header match
    - Implemented in Ternary Content Addressable Memory (TCAM)
  - Applies some action to override regular routing
- Can match many combinations of fields
  - From: CMU to: Stanford
  - Port 53 (DNS)
  - Student housing subnets
  - DSCP/TOS byte
  - Include or exclude individual IP addresses

ISP Daemon

- Provides isolation between ISP and GENI
  - Owned, controlled and audited by ISP
  - GENI does not need direct access to ISP resources
- Facilitates managing risks......

The GENI control plane needs two levels

- Opt-In Component Manager
  - Opt-In looks like a specialized link
  - Connects the Intercept to other GENI resources
- Experiment Control (aka Slice Manager)
  - Responsible for overall integrity of the experiment
    - Primary/preferred experimenter's console
  - Don't enable Opt-In unless the entire slice is ready
    - Inhibit sliver/slice deallocation/preemption
  - Disable Opt-In before shedding resources
  - Liveness checks and monitoring
    - First level “safety controls”
    - Automatic shutdown on failures

Thinking about really big scales

- Might Interpose GENI on all traffic
  - Intercept 100% of US R&E traffic if we want
- Must address broader issues
  - Motivating all actors
  - Managing risk at all levels
- Easiest approach is to think about stakeholders
  - GENI
    - Constituents: Researchers
  - The IRB
    - The innocent users (experimental subjects)
  - The ISP
    - The ISP staff and all users

Stakeholders: GENI & Researchers

- Researchers drive the process
  - Researchers want users
    - Active users: choose to participate e.g. want advanced services
    - Innocent users: did nothing and may be unaware
  - Specify experiments
  - Engage IRB
  - Negotiate intercept pattern matches w/ ISP
- Need to manage their user base
  - Tension between stability and flexibility

Managing the user base

- Core tension between users and researchers
  - Researchers need both innocent and active users”
  - Users want stable (advanced) services
  - Researchers want to change things
- Three example Opt-In scenarios
  - Simple Opt-In
  - Version agility for sustained Opt-In
  - Weaning users from an experimental service

Simple Opt-in Scenarios

- Short running or small scale experiments
- IP address based
  - Individual (enumerated) Opt-In
  - IP prefix block (subnet)
  - IP prefix block (subnet) except individual Opt-out
    - Ultimately need to be able to do dynamic updates

Version agility for sustained Opt-In

- Allocate two long lifetime slices
  - Alpha slice with Individual Opt-In for developers, etc
    - Frequent changes and restarts
  - Beta slice with Wholesale Opt-In
    - One stable version
- Upgrade services by exchanging Opt-In filters
  - Alpha slice becomes new beta w/ Wholesale Opt-In
  - Beta slice disassembled and rebuilt for new alpha
- Claim: Researchers can have full version agility as long as they consider their own internal version compatibility issues.
- Regional Opt-In facilitates gracefully upgrading an experimental services.

Weaning users from an experimental service

- Assume you have a success disaster:
  - Experimental service with limited resources
  - Too many addicted and demanding users
    - They continue to invite their friends to Opt-In by word of mouth
- Convert from wholesale to individual Opt-in
  - Automate the individual (re)Opt-In process
  - Disallow new users
  - Require periodic renewals
    - But make them progressively harder
- Long term goal: fully manage Opt-In for all users

Stakeholders: IRB and Users

- Institutional Review Board
- Supervises all experiments on Human Subjects
  - Explicitly responsible for protecting user interests
- See CFR Title 45, Part 46
  - “Protection of Human Subjects”
  - NSF version: CFR Title 45, Part 690
- Two standard review protocols (or tracks)
  - Social Sciences and Biomedical
- The main rules:
  - Subject has to give informed consent
  - Must protect Personal Identifiable Information (PII)
  - Must balance/justify the risks

Network research, Opt-In and the IRB

- Informed Consent isn't generally feasible
  - Akin to field testing new highway detour signs
  - Exception are permitted but need extra considerations
    - E.g. Public notices, Opt-out instructions
- Primary risks are technical issues
  - Interactions with obscure or experimental services
    - Most users frequently “Opt-In” to new services
  - Unintended PII leaks
    - Inferences about trace data, etc
- These all require accurate risk assessment
  - Technology issues may be more subtle than ethics

The IRB and telecom law

- (I am not a lawyer)
- Strong, IRB supervised, PII protection may be sufficient to placate telecom lawyers

- There are existing procedures to protect IRB supervised studies from subpoena
  - E.g surveys about criminal activities and drug use
  - But the actual wording is completely general

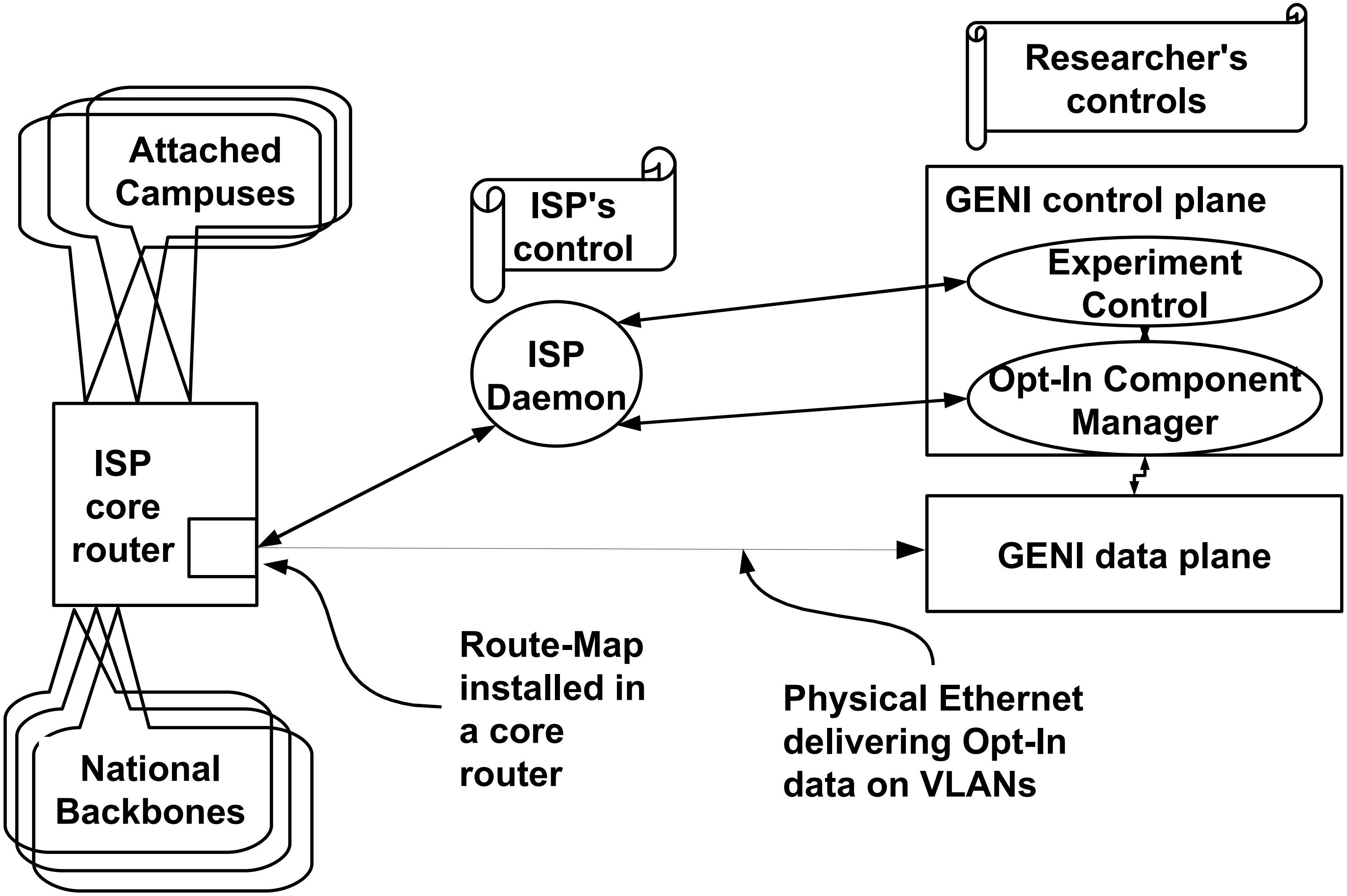
Stakeholders: ISP and other users

- ISP is responsible to its customers
- Opt-In failures hurt everyone
  - Not just the experimental subjects
  - Especially the ISP staff
- Opt-In mechanism itself must be ISP grade
  - Not "GENI grade"
  - Direct price of service is several dollars per second
  - Indirect cost is probably orders of magnitude higher
- Need to strongly manage all risks
  - Opt-In must not expose core routers to rogue actors

ISP Daemon to isolate risks

- Proxy between GENI and core router
  - Owned and managed by the ISP
    - Can be audited/instrumented per the ISP's interests
  - Participates in the ISP's private authentication
  - Participates in the GENI control plane
- ISP must have ultimate control over Opt-In
  - E.g. may veto Opt-In during unrelated failures
  - Otherwise any ISP will refuse to participate
- Enforce IRB "Human Subjects" policies
  - May only be the high bit
  - IRB permissions required for most experiments
    - Automatic exceptions for "self" if not a shared syst

Regional Opt-In


Researcher's controls
Attached Campuses
ISP's control
GENI control plane
Experiment Control
Opt-In Component Manager
ISP Daemon
ISP core router
GENI data plane
Route-Map installed in a core router
Physical Ethernet delivering Opt-In data on VLANs
National Backbones