The eXpressive Internet Architecture: From Architecture to Network

Peter Steenkiste

Dave Andersen, David Eckhardt, Sara Kiesler, Jon Peha, Adrian Perrig, Srini Seshan, Marvin Sirbu, Hui Zhang
Carnegie Mellon University

Aditya Akella, University of Wisconsin

John Byers, Boston University

Future Internet Forum,
Seoul, South Korea, Sept 11, 2012

How do you Improve on the Internet?

- The Internet has been tremendously successful
  - Supports very diverse set of applications and services
  - Integral part of our society and economy
  - But there are also many challenges …
- Lots of exciting research on how to improve Internet
  - Security, routing, wireless/mobile, management, …
  - But Internet architecture constrains what can be modified
- Future Internet Architecture frees researchers to go beyond today's IP architecture and infrastructure
  - Multi-phase, NSF-funded research program
  - Five teams building full scale networks

2

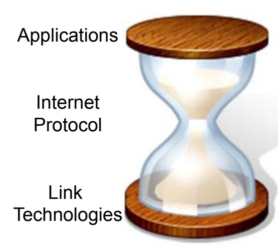
NSF Future Internet Architecture

- Program focuses on new architectural features for the Internet - address challenges in fundamental way
  - Want to keep the good features of today's network
- Four teams were selected in the second phase:
  - Named Internet Architecture: content centric networking - data is a (the) first class entity
  - Mobility First: mobility as the norm rather than the exception – generalizes delay tolerant networking
  - Nebula: Internet centered around cloud computing data centers that are well connected
  - eXpressive Internet Architecture: focus on trustworthiness, evolvability

3

"Narrow Waist" of the Internet Key to its Success

- Has allowed Internet to evolve dramatically
- But now an obstacle to addressing challenges:
  - No built-in security
  - New usage models a challenge
  - Limited interactions edge-core



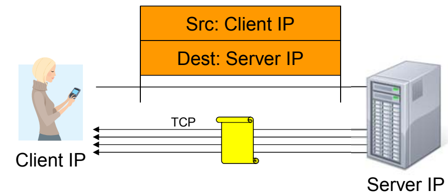
- XIA exploring three concepts to address issues:
  - Diverse types of end-points
  - Intrinsic security
  - Flexible addressing

Outline

- Background
- XIA principles
- XIA architecture
- Building XIA
- Ongoing research
- Conclusion

5

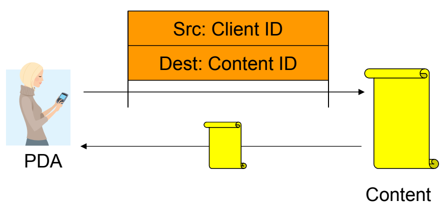
Today's Internet



- Client retrieves document from a specific web server
  - But client mostly cares about correctness of content, timeliness
  - Specific server, file name, etc. are not of interest
- Transfer is between wrong principals
  - What if the server fails?
  - Optimizing transfer using local caches is hard
    - Need to use application-specific overlay or transparent proxy – bad!

6

eXpressive Internet Architecture



- Client expresses communication intent for content explicitly
  - Network uses content identifier to retrieve content from appropriate location
- How does client know the content is correct?
  - Intrinsic security! Verify content using self-certifying id:
    hash(content) = content id
- How does source know it is talking to the right client?
  - Intrinsic security! Self-certifying host identifiers

7

Three Key Principles

- An set of principals allows direct identification of the intended communicating entities
  - Not having to force communication at a lower level (hosts in today's Internet) reduces complexity and overhead
- Set up principals can evolve over time
  - Adapt to changes in usage models
  - Support custom requirements of specific deployments
- Intrinsic security guarantees security properties as a direct result of the design of the system
  - Do not rely on external configurations, actions, data bases

8

Other XIA Principles

- Narrow waist for all principals
  - Defines the API between the principals and the network protocol mechanisms
- Narrow waist for trust management
  - Ensure that the inputs to the intrinsically secure system match the trust assumptions and intensions of the user
  - Narrow waist allows leveraging diverse mechanisms for trust management: CAs, reputation, personal, …
- All other network functions are explicit services
  - Keeps the architecture simple and easy to reason about
  - XIA provides a principal type for services (visible)

Look familiar?

9

XIA: eXpressive Internet Architecture

- Each communication operation expresses the intent of the operation
  - Also: explicit trust management, APIs among actors
- XIA is a single inter-network in which all principals are connected
  - Not a collection of architectures implemented through, e.g., virtualization or overlays
  - Not based on a "preferred" principal (host or content), that has to support all communication

10

The XIA Research Team

- Principles do not make an architecture!
- Meet the initial core XIA team …

Fahad Dogar | Dongsu Han | Hyeontaek Lim | Ashok Anand | Michel Machadoy | Boyan Li | Wenfei Wu

- … and some new researchers who joined in year 2

Matthew Mukerjee | David Naylor | Junchen Jiang | Suk-Bok Lee (postdoc)

11

Outline

- Background
- XIA principles
- XIA architecture
  - Multiple principals
  - DAG-based addressing
  - Intrinsic security
- Building XIA
- Ongoing research
- Conclusion

12

Multiple Principal Types

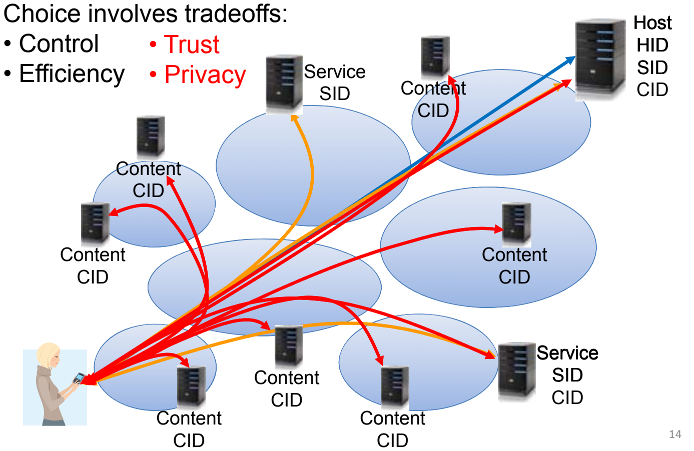
- Associated with different forwarding semantics
  - Support heterogeneity in usage and deployment models
  - Set of principal types can evolve over time
- Hosts XIDs support host-based communication – *who?*
- Service XIDs allow the network to route to possibly replicated services – *what does it do?*
  - LAN services access, WAN replication, …
- Content XIDs allow network to retrieve content from “anywhere” – *what is it?*
  - Opportunistic caches, CDNs, …
- Autonomous domains allow scoping, hierarchy

“XIA: An Architecture for an Evolvable and Trustworthy Internet”, ACM Hotnets 2012

13

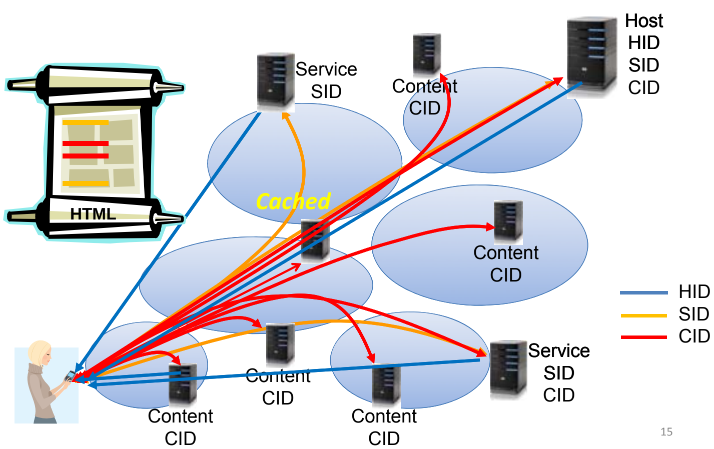
Multiple Principal Types

Choice involves tradeoffs:
- Control
- Efficiency
- Trust
- Privacy



14

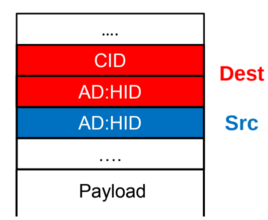
Content-centric Optimizations



15

Supporting Evolvability

- Introduction of a new principal type will be incremental – no “flag day”!
  - Not all routers and ISPs will provide support from day one
- Creates chicken and egg problem - what comes first: network support or use in applications
- Solution is to provide an *intent* and *fallback* address
  - Intent address allows in-network optimizations based on user intent
  - Fallback address is guaranteed to be reachable



16

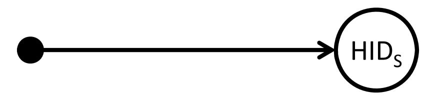
Addressing Requirements

- Fallback: intent that may not be globally understood must include a backwards compatible address
  - Incremental introduction of new XID types
- Scoping: support reachability for non-globally routable XID types or XIDs
  - Needed for scalability
  - Generalize scoping based on network identifiers
  - But we do not want to give up leveraging intent
- Iterative refinement: give each XID in the hierarchy option of using intent

17

Our Solution: DAG-Based Addressing

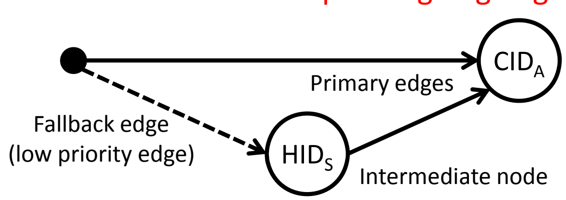
- Uses direct acyclic graph (DAG)
  - Nodes: typed IDs (XID; expressive identifier)
  - Outgoing edges: possible routing choices
- Simple example: Sending a packet to HID_S



18

Support for Fallbacks with DAG

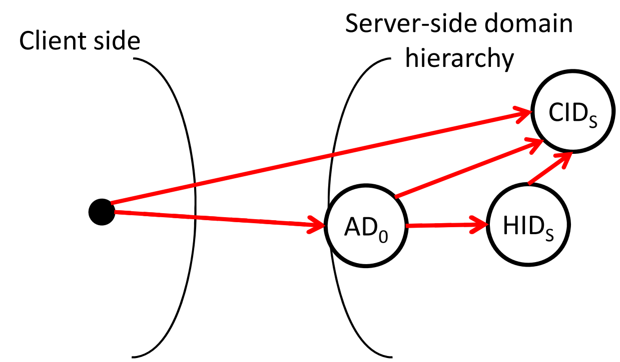
- A node can have multiple outgoing edges



- Outgoing edges have priority among them
  - Forwarding to HID_S is attempted if forwarding to CID_A is not possible – Realization of fallbacks

19

DAGs Support Scoping and Iterative Refinement



"XIA: Efficient Support for Evolvable Internetworking", NSDI 2012

20

DAG Addressing Research Questions

- DAG addressing supports is flexible …
  - Fallback, binding, source routing, mobility, ..
- … but many questions remain:
  - Is it expensive to process?
  - How big will the addresses be?
  - How do ISPs verify policy compliance?
  - Can they be used to attack network?
  - Can it be deployed incrementally?

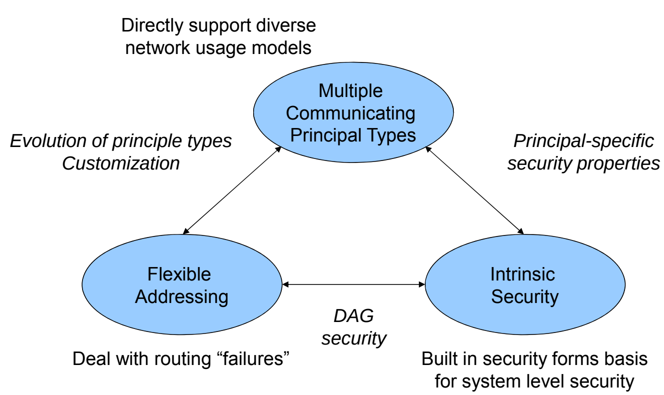
21

Intrinsic Security in XIA

- XIA uses self-certifying identifiers that guarantee security properties for communication operation
  - Host ID is a hash of its public key – accountability (AIP)
  - Content ID is a hash of the content – correctness
  - Does not rely on external configurations
- Intrinsic security is specific to the principal type
- Example: retrieve content using …
  - Content XID: content is correct
  - Service XID: the right service provided content
  - Host XID: content was delivered from right host

22

XIA Dataplane Concepts Revisited

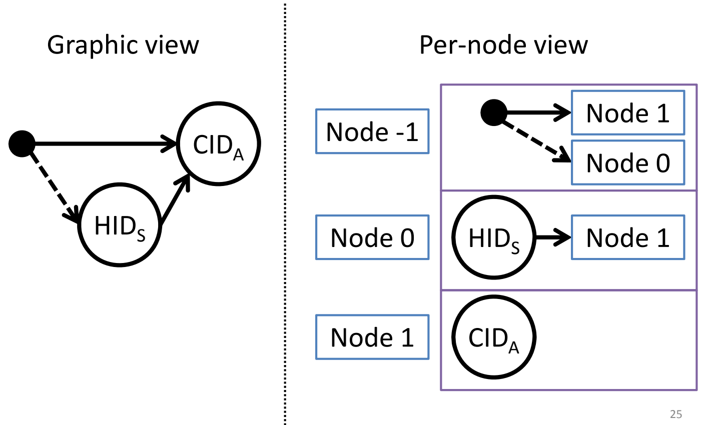


Outline

- Background
- XIA principles
- XIA architecture
- Building XIA
  - Forwarding packets
  - Building a network
  - Prototype
- Ongoing research
- Conclusion

24

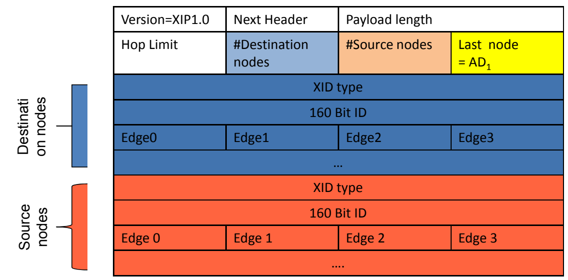
Putting DAG Addresses into Packet Headers



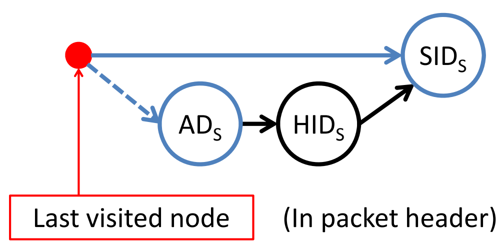
25

XIP Packet Header

- DAGs represent source and destination addresses
- Array of nodes with pointers
- Maintains a *LastNode* field in the header
  - Routers to know where to begin forwarding lookups



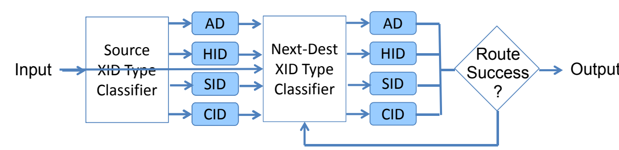
Router's View on Packet Forwarding



1. Forward to SID$_S$ if possible
2. Otherwise, forward to AD$_S$
   - If router is AD$_S$ itself, update last visited node to AD$_S$

27

Packet Processing Pipeline



- Principle-independent processing defines how to interpret the DAG
  - The core XIA architecture
- Principle-dependent processing realizes forwarding semantics for each XID type
- Optimizations possible: fast path processing, packet level and intra-packet parallelism

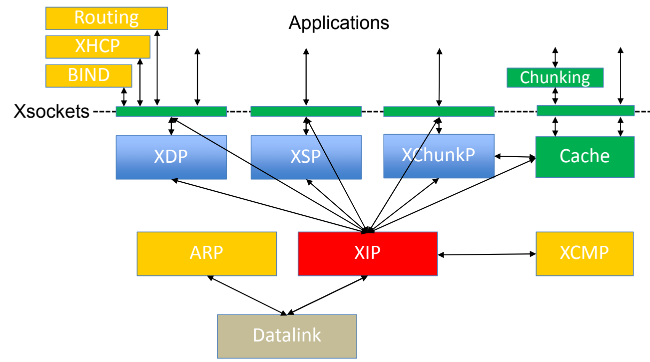
28

XIA Prototype

- Full stack for routers, caches, and end-points
- Based on Click protocol framework
  - User-level/in-kernel, native/overlay
- XIA forwarding engine was used in performance study
- Expanded to support applications, services
  - "xsocket" programming interface
  - basic transport: datagrams, streaming, content
  - Routing, naming, diagnostics, ...

29

XIP Protocol Stack



30

Open Source XIA Release

- XIA Prototype released in May 2012
  - Includes full XIA protocol stack and utilities
  - Support for GENI and VM-based experiments
  - Improve over time with research results
  - More info: http://www.cs.cmu.edu/~xia
- Being used to support applications, services
  - Working towards permanent XIA deployment
- Prototype good platform for collaboration
  - We can provide support to users and developers

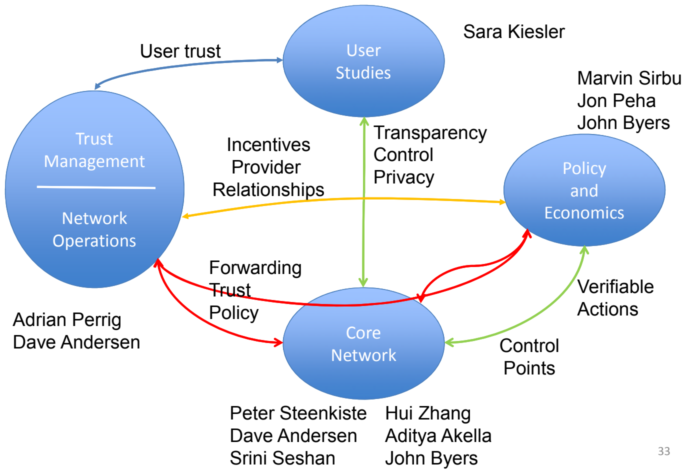
31

Outline

- Background
- XIA principles
- XIA architecture
- Building XIA
- Ongoing research
- Conclusion

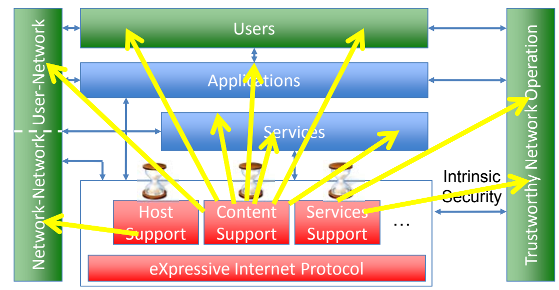
32

XIA Project Structure



33

XIA Components and Interactions



34

A Broad Research Agenda

- Applications and services
  - Web, CDNs, video delivery, teleconferencing, games, mobility services, …
- Protocols and network infrastructure
  - Security, transport protocols, naming, mobility, routing, service deployment, principal types, network operations, diagnostics, video,DTNs, ..
  - XIA forwarding, services, caching, intrinsic sec., …
- Targeted deployments
  - Use XIA to optimize unique networks, e.g. wireless access, Scada, sensors, “ad hoc”, data center, …

35

Examples of Ongoing Research

- Path selection using Scion
- XIA forwarding performance
- Transport protocols
- Incremental deployment using 4ID
- Mobility
- Middleboxes and in-path services

36

SCION Architectural Goals

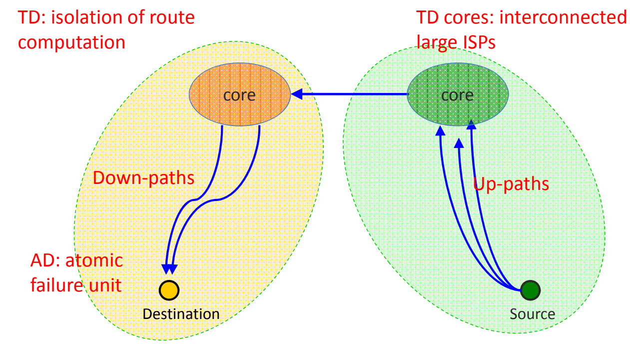

Adrian Perrig

- High availability in presence of malicious parties
- Explicit trust for network operations
- Minimal TCB: limit number of entities that need to be trusted for any operation
  – Strong isolation from untrusted parties
- Operate with mutually distrusting entities
  – No single root of trust
- Enable route control for ISPs, receivers, senders
- Simplicity, efficiency, flexibility, and scalability

37

Hierarchical Decomposition

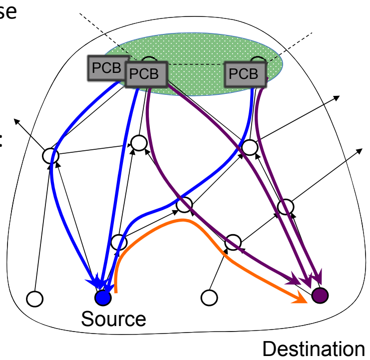
- Split network into a set of trust domains (TD)



38

Path Selection in SCION Architecture Overview

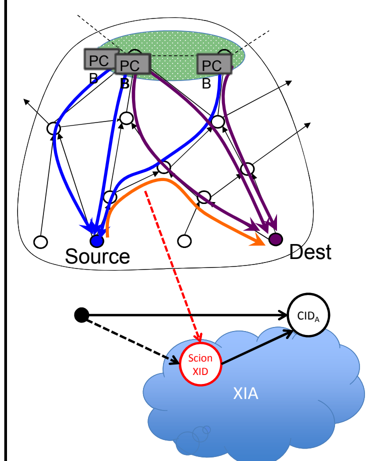
- Source/destination can choose among up/down hill paths
  - Path control shared between ISPs, receivers, senders
- Desirable security properties:
  - High availability, even in presence of malicious parties
  - Explicit trust for operations
  - Minimal TCB: limit number of entities that must be trusted
  - No single root of trust
  - Simplicity, efficiency, flexibility, and scalability



Adrian Perrig presentation in GFIS, Friday afternoon

39

Scion over XIA Dataplan



- Integration of Scion-XIA started this semester
- Store paths generated by Scion into a new type of XID
  – Sequence of MACs
  – Can be combined with other principal types
- XIA network supports both path and destination-based forwarding

40

Examples of Ongoing Research

- Path selection using Scion
- XIA forwarding performance
- Transport protocols
- Incremental deployment using 4ID
- Mobility
- Middleboxes and in-path services

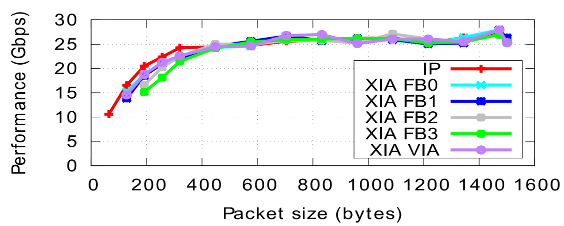
Evaluation Setup


Dongsu Han


Hyeontaek Lim

- Use packet generator to evaluate throughput
- Software:
  - PacketShader I/O Engine
  - Click modular router – multithreaded(12 threads)
- Hardware:
  - 10Gbit NIC : 4 ports (multi-queue support)
  - 2x 6 Core Intel Xeon @ 2.26GHz
- Optimizations apply: fast path processing, packet level and intra-packet parallelism

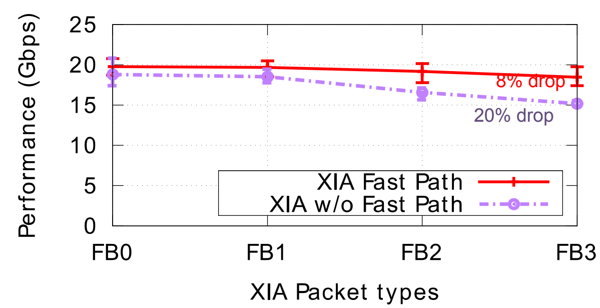
Forwarding Performance Comparison



351K FIB entries
Workload: Identifiers generated using Pareto distribution

XIP forwarding is fast!
@128 byte FB0 is 8% slower than IP
@192 byte FB3 is 26% slower than IP

Fast Path Performance



Look-aside cache of 1024 entries

Using fast-path processing, the gap between FB0 and FB3 is reduced significantly !

Summary

- XIA packet forwarding cost is reasonably competitive compared with IP!
- Inter-packet parallelism and fast-path can be applied to get high-speed XIA forwarding on software routers
- Intra-packet parallelism can be used for further speedup in hardware implementations

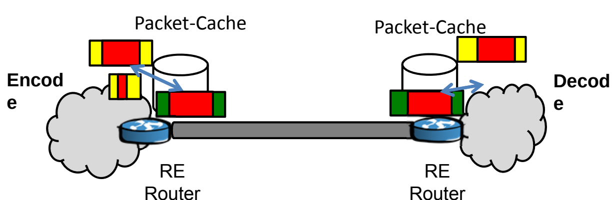
Rethinking Transport Protocols

- How can we improve on the well established error and congestion control mechanisms?
- Error control based on retransmission or FEC
  - FEC avoids retransmission delays - (soft) RT
  - Can we reduce overheads of redundancy by leveraging content-aware networking ideas?
- Congestion control is either end-point or network based
  - How can we accommodate preferences of both?

46

Content-aware Networks: New environment

Content-aware networks = caching + content-aware processing to remove duplicates



Caching effectively eliminates redundant data transfer.

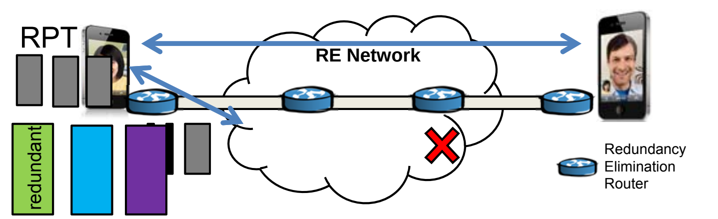
Redundancy elimination (RE) networks [Spring 00, Anand 08]

Content-Centric Networking (CCN) [Jacobson 09]

47

Re-designing Redundancy

- Redundant Packet Transmission–Introducing redundancy in a way that the network understands



Questions/Challenges

- How do we retain the robustness benefits?
- How much redundancy is needed?
- How does it compare with FEC?

"RPT: Re-architecting Loss Protection for Content-Aware Networks", Dongsu Han, Anand Ashok, Aditya Akella, Srini Seshan, NSDI, 2012.

48

E2E Performance: Video Quality

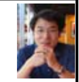
Metric: Average Peak Signal to Noise Ratio (PSNR)

RPT(3) Overhead ~6% | FEC(10,9) Overhead ~10% | Naïve UDP

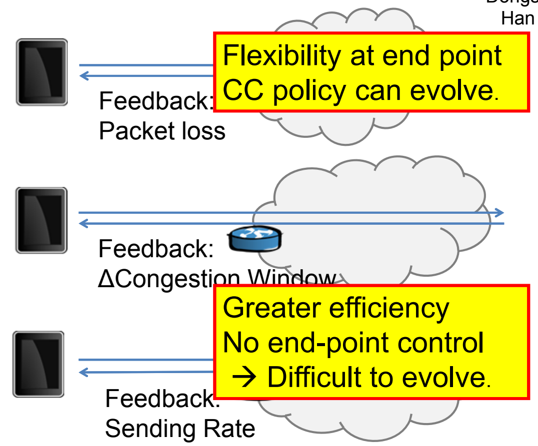
1.8dB ~ 3dB difference in quality (Peak Signal to Noise Ratio)

49

Congestion Control



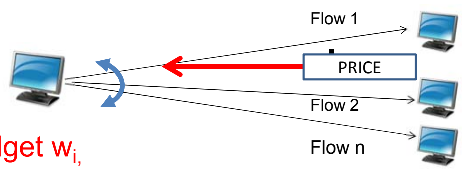
- TCP and friends
- XCP
- RCP



Can we achieve both?

Sender | Network | Receiver

Budget W ($/sec)



Flow i's budget w_i, subject to $W \geq \Sigma w_i$

Flow i's price: P_i ($/Byte)

Flow i's rate: R_i = budget/price = w_i / P_i (Byte/sec)

Questions/Challenges
- Can we achieve high efficiency?
- Can we implement different CC policies?
- How to deal with feedback errors due to delay?

"A Framework for an Evolvable Transport Protocol" Dongsu Han, Robert Grandl, Aditya Akella, Srini Seshan, Under submission/TR.

51

Examples of Ongoing Research

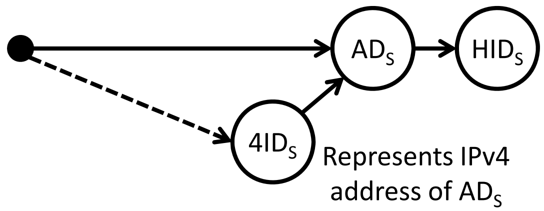
- Path selection using Scion
- XIA forwarding performance
- Transport protocols
- Incremental deployment using 4ID
- Mobility
- Middleboxes and in-path services

52

Incremental Deployment of XIA

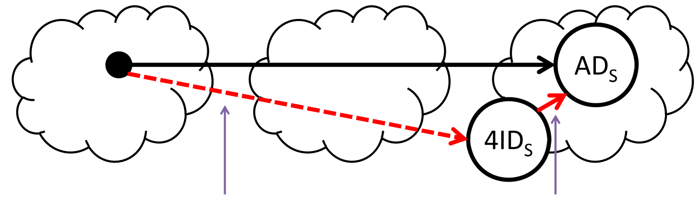

Matthew Mukerjee

- 4ID: IPv4 address as an XID
  - IPv4 encapsulation between XIA network islands
  - Leverages fallback for legacy networks
- No need for statically configured tunnels!



53

4ID in Action: Partially Deployed XIA Networks



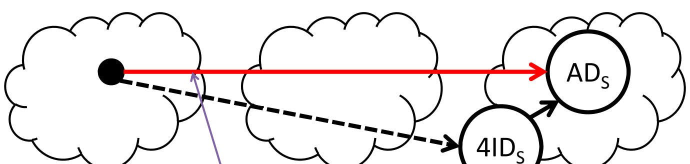
Entering IPv4 network: Encapsulate XIA packet with IP header

Entering XIA network: Remove IP header for native XIA packet processing

Dynamic encapsulation: no static tunnels

54

4ID in Action: Fully Deployed XIA Networks



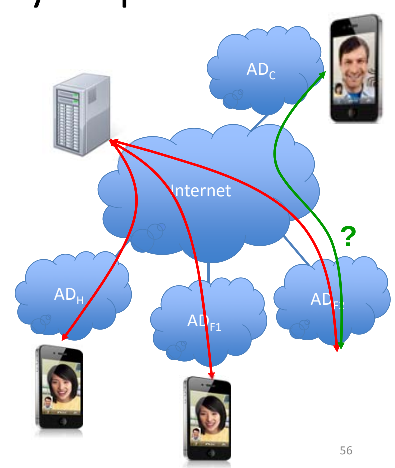
Use native XIA forwarding and ignore fallback

- Seamless incremental deployment of XIA
- Leverages routing in each network

55

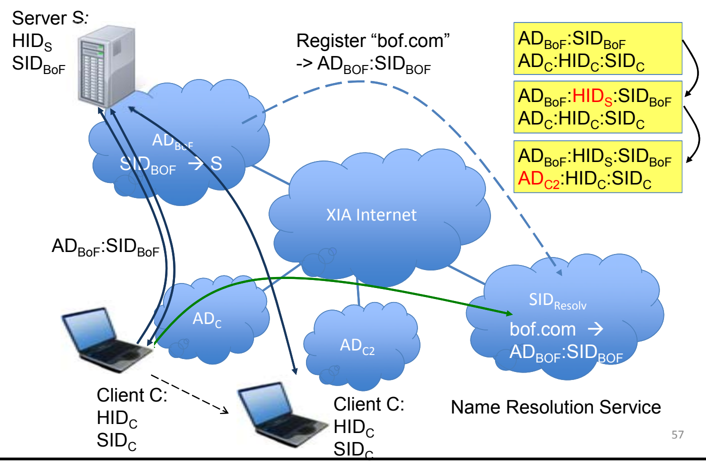
Mobility is a Key Requirement

- Inter-domain mobility remains a challenge in today's Internet
  - Active sessions
  - New sessions
- Requires separation of identifier and locator – for XIA:
  - Identifier = HID
  - Locator = DAG



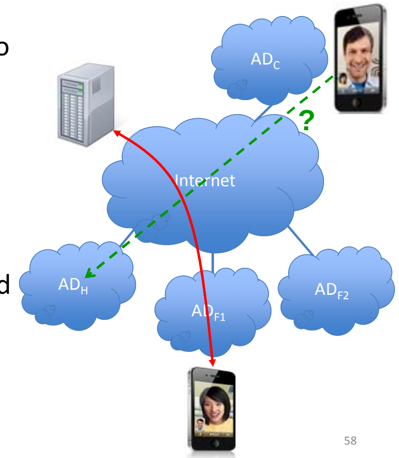
56

Maintaining a Mobile User Session



Finding Mobile Devices

- Must map identifier into an up-to-date locator
  - Challenges include scalability, security, deployment complexity, latency, overheads, ….
  - Lots of previous work
- Exploring solution based on locator services
  - Keeps track of user's current location
  - User can choose service



Examples of Ongoing Research

- Path selection using Scion
- XIA forwarding performance
- Transport protocols
- Incremental deployment using 4ID
- Mobility
- Middleboxes and in-path services

Our Love-Hate Relationship with Middleboxes

- Middleboxes are everywhere
  - NATs, firewalls, virus scanners, mobility services, media gateways, …
  - We cannot live without them!
- But they are a source of many concerns
  - Viewed as ugly given an end-to-end philosophy
  - Source of silent application failure modes, e.g., caused by faith sharing
  - We cannot live with them!

Fahad
Dogar

Support for Mobility Services

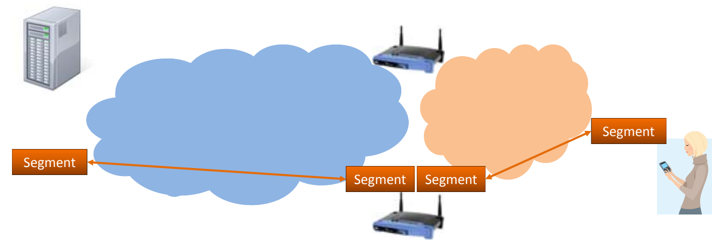
- Increasing network heterogeneity
  - Paths are no longer homogeneous
- Topology control
  - Handoff, multi-path
- Heterogeneous devices, usage
  - Relaxed end-point synchronization
- Diverse network services
  - Content retrieval, mobility services

Unbundle Transport Function

Leverage in-network functionality

61

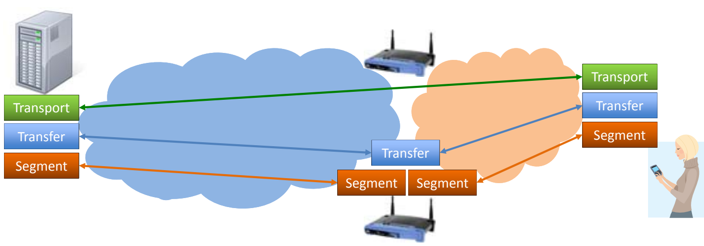
Transfer Access Points



- Tapa supports visible middleboxes (TAPs) that break up end-end connections in homogeneous segments
- Provide natural insertion point for diverse services
- Segments support best effort delivery of "chunks"
  - Can use custom protocols on segment

62

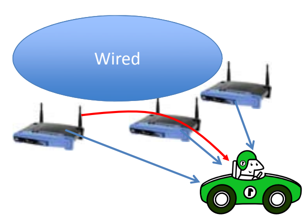
Unbundling the Transport Layer

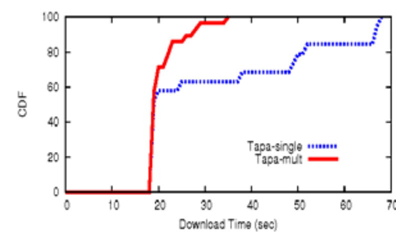


- Transfer layer glues segments into e-e path
  - Kind of like IP, but across segments, not hops
  - Naturally supports insertion of network services
- Thin end-to-end transport supports e-e semantics
  - Also flow, error, congestion control across segment path
  - Must account for failures of TAPs, segment breaks, etc.

63

Vehicular Example

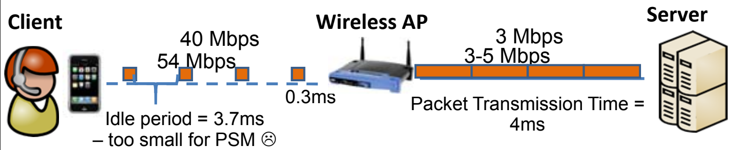




- Vehicle-infrastructure suffers from frequent interruptions, short periods of connectivity
- Vehicle optimizes transfers by explicitly managing server-TAP and TAP-vehicle transfers
  - Leverages self-certifying content identifiers

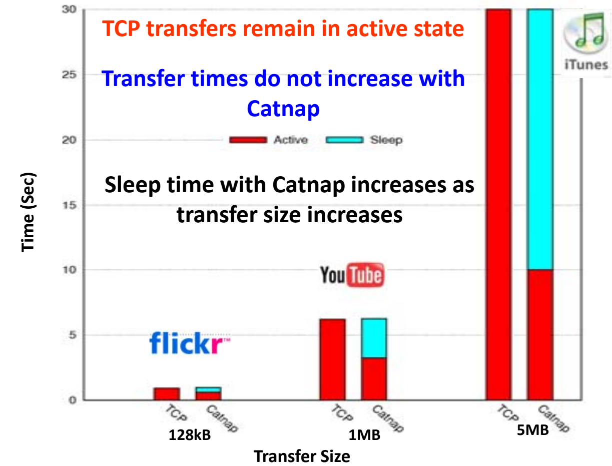
64

Bandwidth Discrepancy in End-to-end Transfers



- Catnap uses this opportunity to save energy
- TAP buffers incoming packets while client sleeps
- Scheduler schedules burst transfer to maximize energy savings while avoiding e-e delay
  - Estimates bandwidth in wired and wireless segments

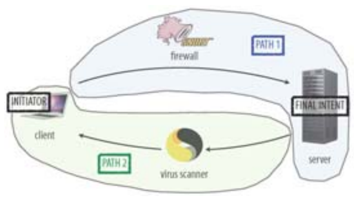
How Much Can the NIC Sleep?



Using In-Path Services


David Naylor

- Use XIA to better support in-path services
  - Builds on the Tapa transport architecture
- Raises research questions in many areas



  - What type of DAGs are effective and for what services?
  - How do transport protocols and services interact?
  - What are the intrinsic security properties of a session?
  - How can DAGs be safely modified during a session?

Conclusion

- XIA supports evolution, expressiveness, and trustworthy operation.
  - Multiple principal types, intrinsic security, and flexible addressing
  - Open source prototype available online:
    http://www.cs.cmu.edu/~xia
- Looking for collaborators on broad research agenda applications, protocols, and deployments
  - Use XIA to fundamentally improve the network: transport protocols, trust management, applications, services, …
  - Use flexibility to target demanding network deployments
    - Customize without giving up interoperability