Carnegie Mellon CyLab

CONFIDENCE FOR A NETWORKED WORLD

Carnegie Mellon CyLab
4616 HENRY STREET
PITTSBURGH, PA 15213
PH: (412) 268-7195
FX: (412) 268-7196
www.cylab.cmu.edu

SCION:

Scalability, Control and Isolation On Next-Generation Networks

Current main team: Soo Bum Lee, Hsu-Chun Hsiao, Hyun Jin Kim, Yue-Hsun Lin, Sangjae Yoo, Adrian Perrig, Virgil Gligor

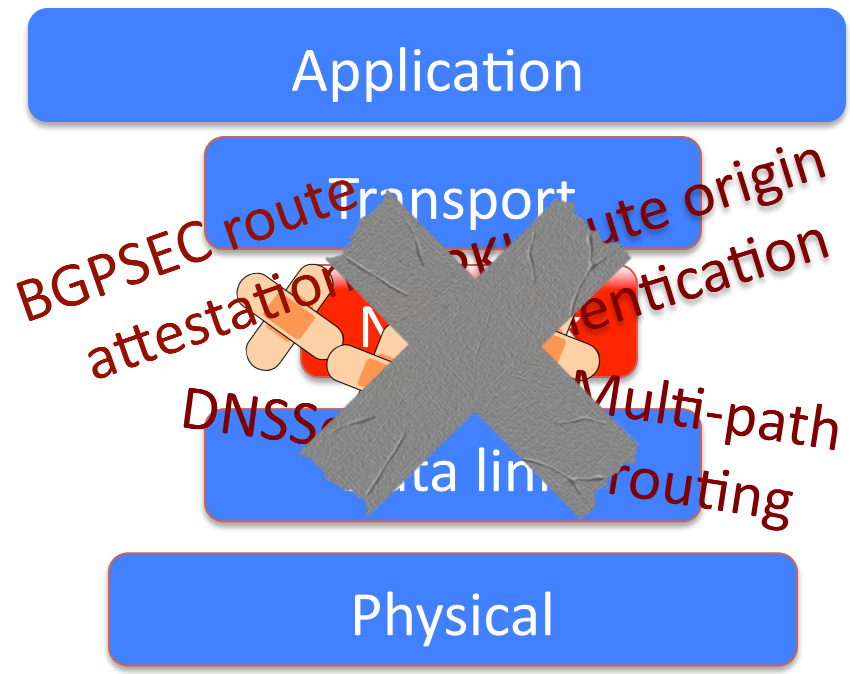
Previous members: Xin Zhang, Geoff Hasker, Haowen Chan, David Andersen

After years of patching, the Internet is still neither Reliable nor Secure!

Feb 2008: Pakistani ISP hijacks YouTube prefix

Apr 2010: A Chinese ISP inserts fake routes affecting 15% of global traffic

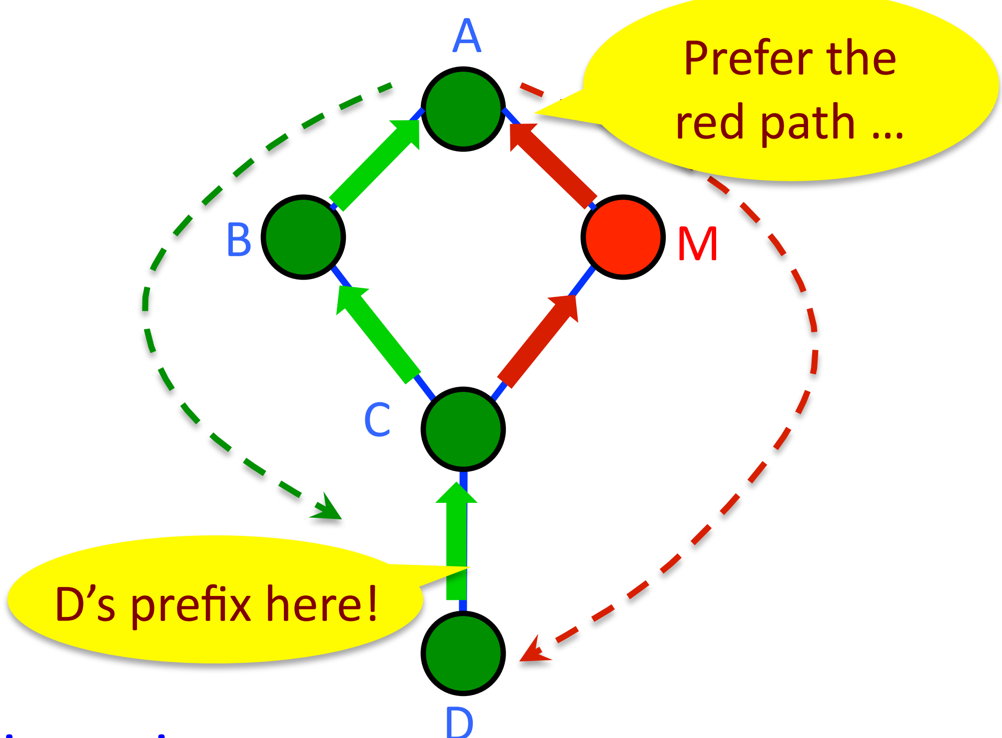
Nov 2010: 10% of Internet traffic 'hijacked' to Chinese servers due to DNS Tampering.



- Fixes to date – ad hoc, patches
- Inconvenient truths
  - S-BGP: delayed convergence
  - Global PKI: single root of trust

Fundamental BGP Limitations

- Destination or ISP have no control over inbound paths



- Route inconsistencies
  - Forwarding state may be different from announced state

Fundamental BGP Limitations (cont'd)

- Lack of routing isolation
  - A failure/attack can have global effects
  - Global visibility of paths is not scalable
- Slow convergence / route oscillation
- Large routing tables
  - Multi-homing / flat namespaces prevent aggregation
- Lack of route freshness
  - Current (S-)BGP cannot prevent replay of old paths

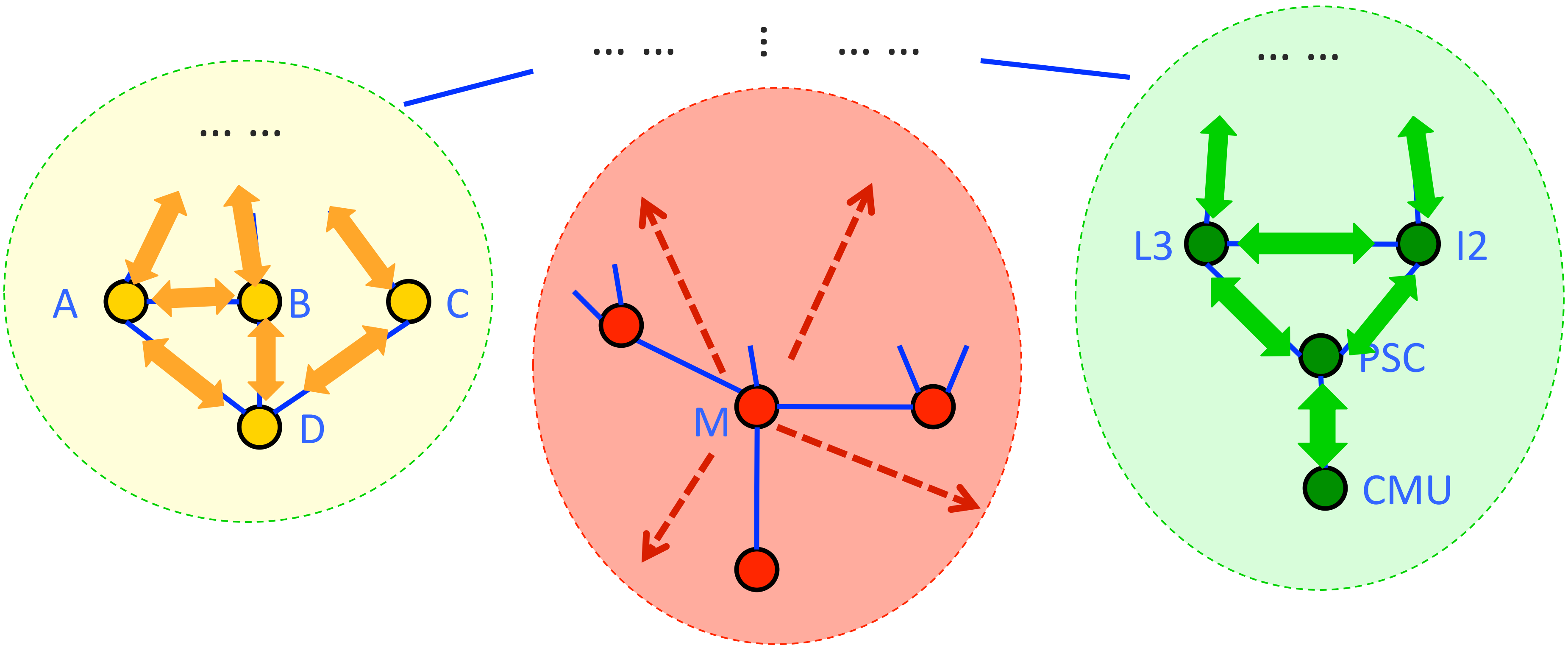
Note that these issues are fundamental to (S)-BGP, they cannot be easily fixed by small changes!

S-BGP Limitations

- Slow convergence
- Router outage causes high overhead
- Circular dependency between UPDATE message and connectivity with RPKI server
  - Route Origin Authentication (ROA), prefix certificate and BGPSEC router certificate needs to be downloaded to validate UPDATE message!
  - Rebooting Internet would be very slow as initial UPDATE messages cannot be validated
- Route flap dampening can be misused
  - Ensure an AS's updates are ignored
  - Prevent updates to fix a path
  - Potential to create a loop that persists

Wish List (1): Isolation

- Isolation of attacks and faults
- Scalable and reliable routing updates
- Operate with mutually distrusting entities without a global single root of trust: enforceable accountability

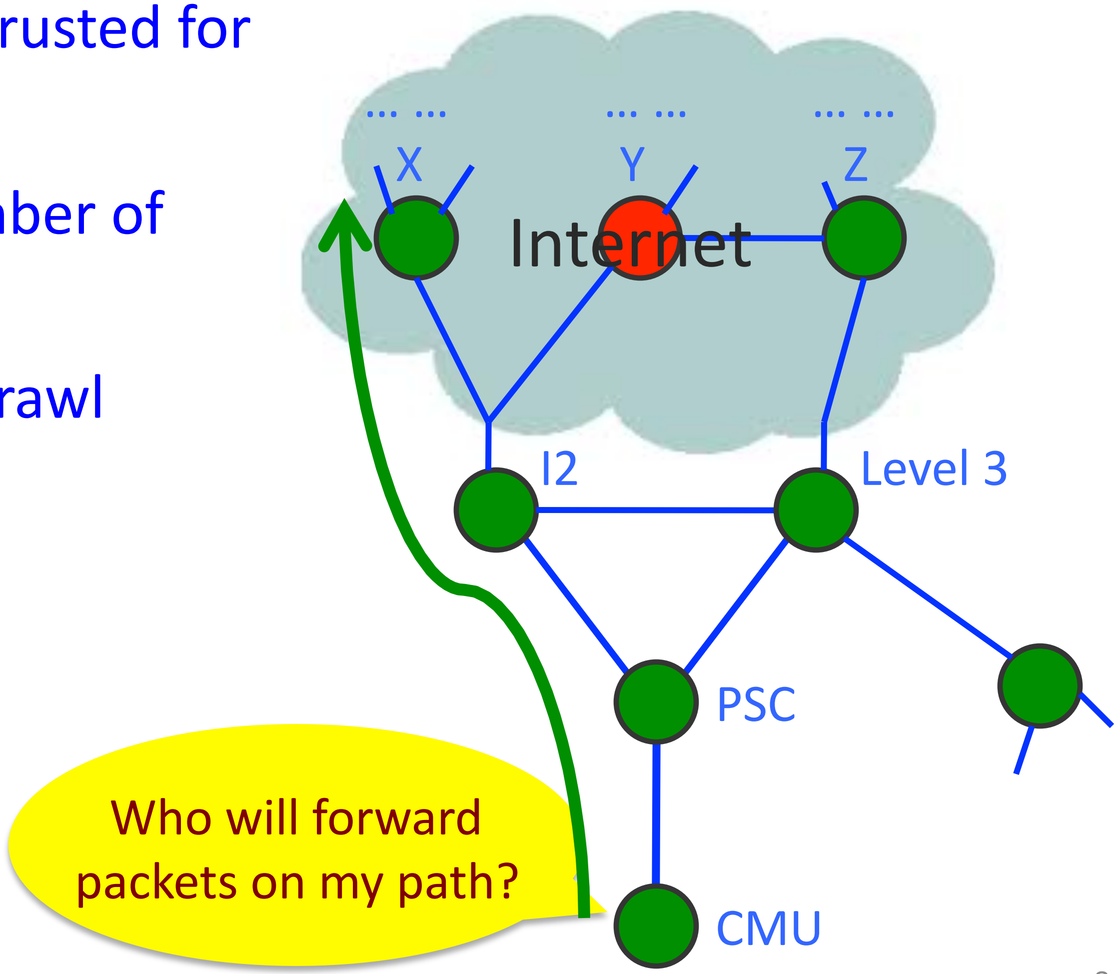


Wish List (2): Balanced Control

- ISPs, source and destination all need path control
- Challenges
  - What granularity of control is appropriate?
  - How to prevent attacks based on sender / receiver control?

Wish List (3): Minimal / Explicit Trust

- Clear who needs to be trusted for network operations
- Small TCB: minimal number of trusted entities
- Avoid transitive trust sprawl

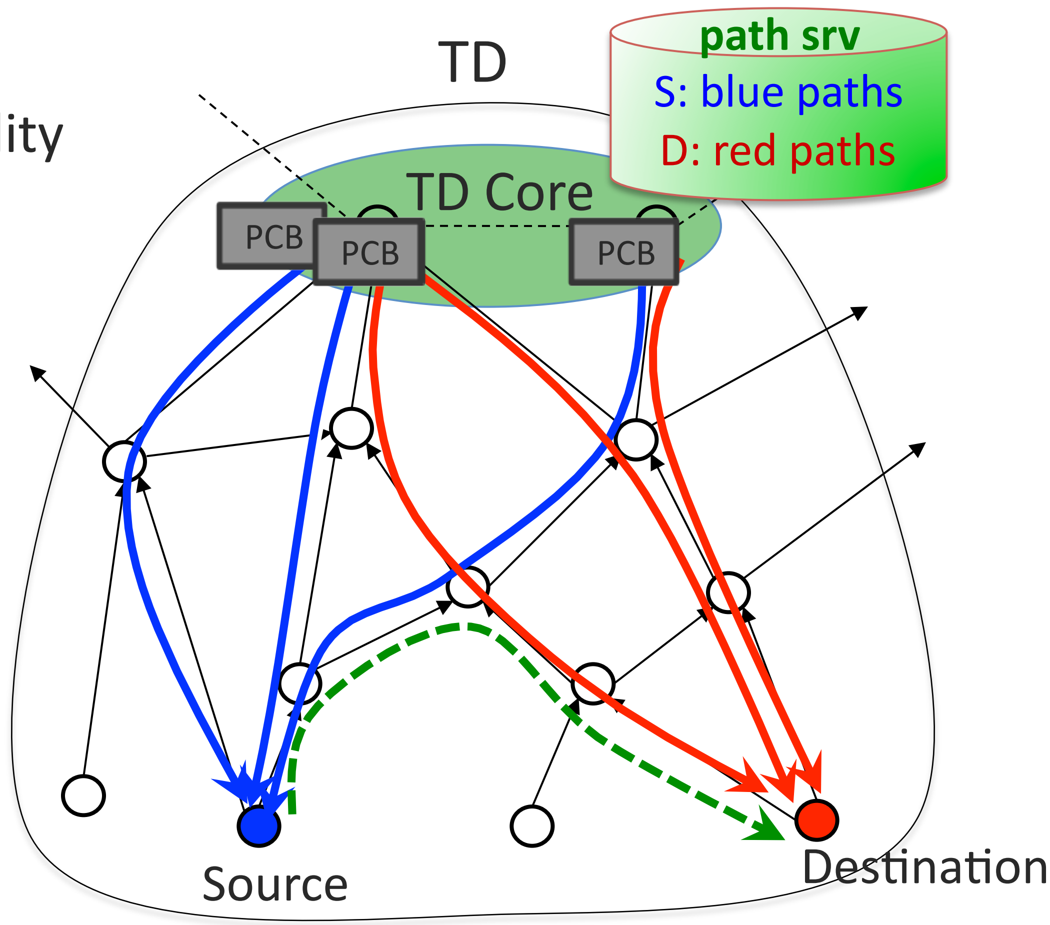


SCION Architectural Goals

- High availability, even for networks with malicious parties
  - Communication should be available if attacker-free path exists
- Explicit trust for network operations
- Minimal TCB: minimize trusted entities for any operation
  - Strong isolation from untrusted parties
- Operate with mutually distrusting entities
  - No single root of trust
- Balanced route control for ISPs, receivers, senders
- No circular dependencies during setup: enable rebootability
- Simplicity, efficiency, flexibility, and scalability

SCION Architecture Overview

- Trust domain (TD)s
  - Isolation and scalability
  - Enforceable accountability
- Path construction
  - Path construction beacons (PCBs)
- Path resolution
  - Control
  - Explicit trust
- Route joining (shortcuts)
  - Efficiency, flexibility



Trust Domain Decomposition

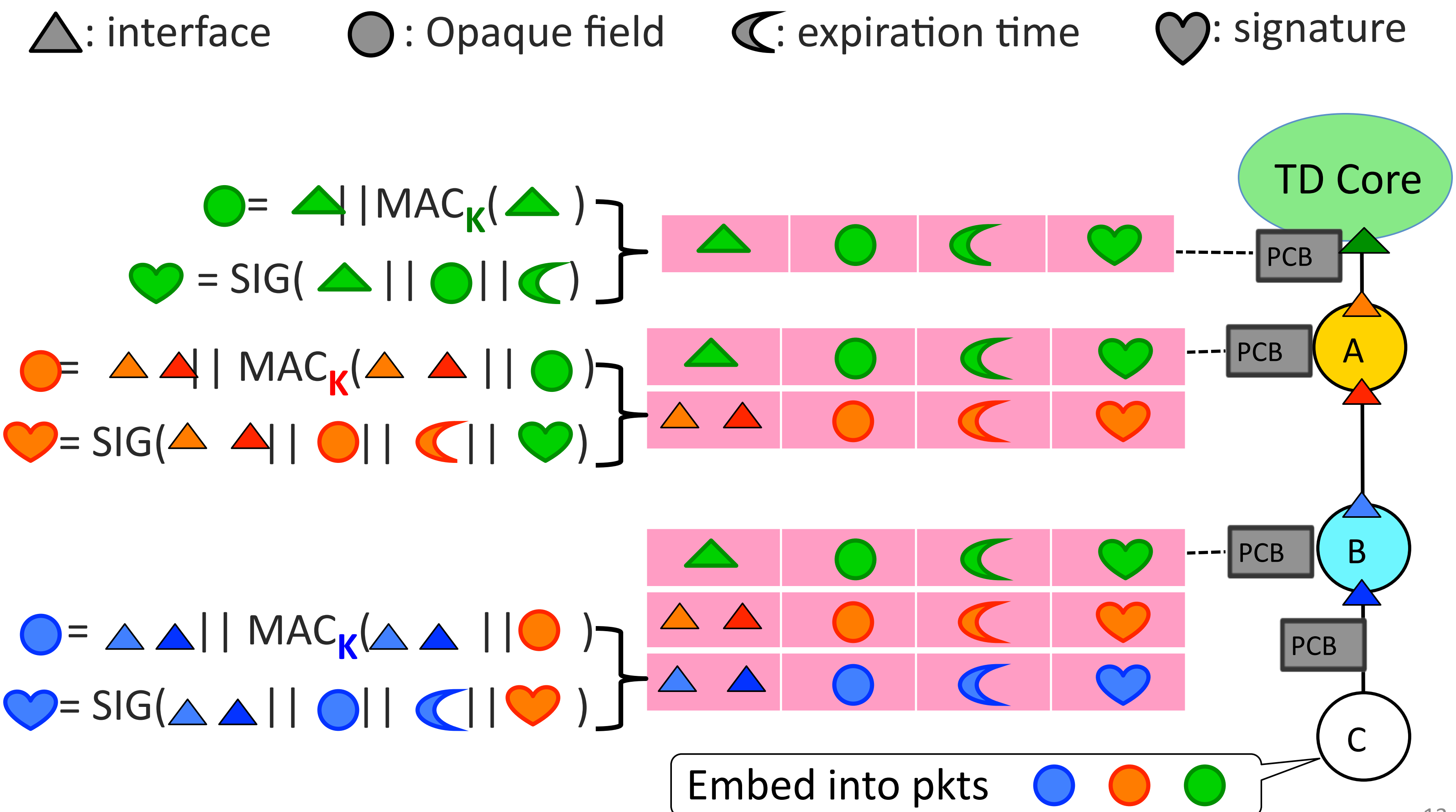
- Global set of TD (Trust Domains)
  - ✓ Map to geographic, political, legal boundaries
  - ✓ Usually corresponds to a jurisdiction
  - ✓ Provide enforceable accountability
- TD Core: set of top-tier ISPs that manage TD
  - ✓ Route to other TDs
  - ✓ Initiate path construction beacons
  - ✓ Manage Address and Path Translation Servers
  - ✓ Handle TD membership
  - ✓ Root of trust for TD: manage root key and certificates
- AD: Autonomous Domain
  - ✓ Transit AD or endpoint AD

Path Construction

Goal: each endpoint learns multiple verifiable paths to its core

- Discovering paths via Path Construction Beacons (PCBs)
  - ✓ TD Core periodically initiates PCBs
  - ✓ ADs asynchronously propagate PCBs
- ADs perform the following operations
  - ✓ Collect PCBs
  - ✓ For each neighbor AD, select which *k* PCBs to forward
  - ✓ Update cryptographic information in PCBs
- Endpoint AD receives at least *k* PCBs from each neighbor AD, selects *k* down-paths to advertise

Path Construction (simplified)



Path Construction

Interfaces: I(i) = previous-hop interfaces || local interfaces

Opaque field: O(i) = local interfaces || MAC over local interfaces and O(i-1)

Signature: Σ(i) = sign over I(i), T(i), O(i), and Σ(i-1), with cert of pub key

TC→A: $I(TC)$: ϕ ||{ϕ, $TC1$}

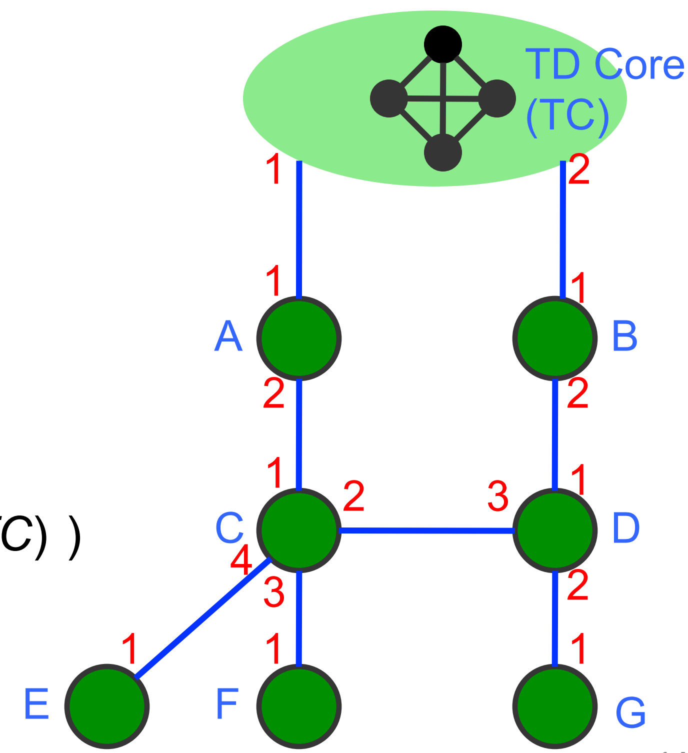
$O(TC)$: {ϕ, $TC1$} ||MAC_{Ktc}({ϕ, $TC1$} || ϕ)

Σ(TC): Sign($I(TC)$ || $T(TC)$ || $O(TC)$ || ϕ)

A→C: $I(A)$: $I(TC)$|| {$A1$, $A2$}

$O(A)$: {$A1$, $A2$} || MAC_{Ka}({$A1$, $A2$} ||$O(TC)$)

Σ(A): Sign($I(A)$ || $T(A)$ || $O(A)$ ||Σ(TC))



Path Construction

Interfaces: I(i) = previous-hop interfaces || local interfaces

Opaque field: O(i) = local interfaces || MAC over local interfaces and O(i-1)

Signature: Σ(i) = sign over I(i), T(i), O(i), and Σ(i-1), with cert of pub key

C? – One PCB per neighbor

C→E: *I*(*C*): *I*(*A*)|| {*C*1, *C*4}

O(*C*) : {*C*1, *C*4} || MAC_{Ka}({*C*1, *C*4} || *O*(*A*))

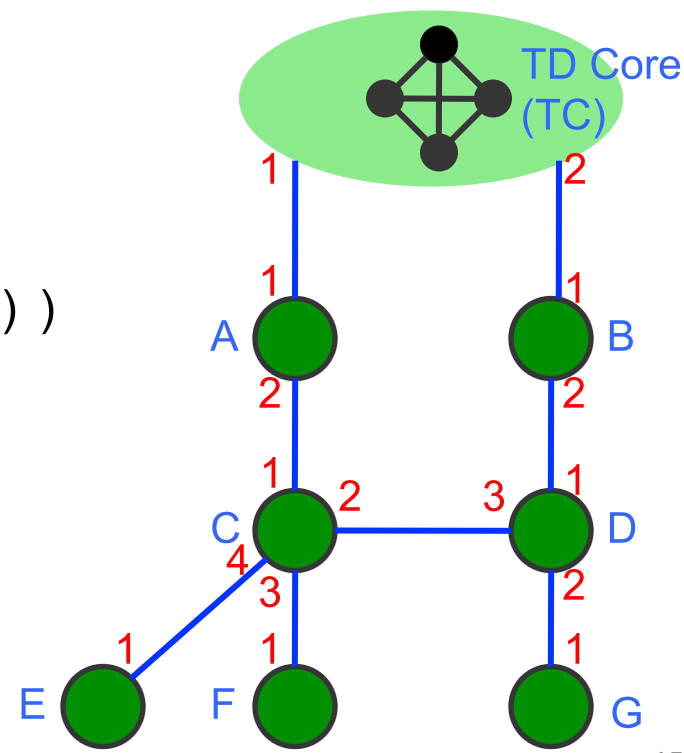
Σ(C): Sign(*I*(*C*) || *T*(*C*) || *O*(*C*) ||Σ(A))

Also include peering link!

$I_{C,D}$(C): {*C*4,*C*2} || TD || AID_D

$O_{C,D}$(C): {*C*4, *C*2} ||MAC_{Kc}({*C*4, *C*2})

$\Sigma_{C,D}$(C): Sign($I_{C,D}$(C) || $T_{C,D}$(C) || $O_{C,D}$(C) || *O*(*C*))

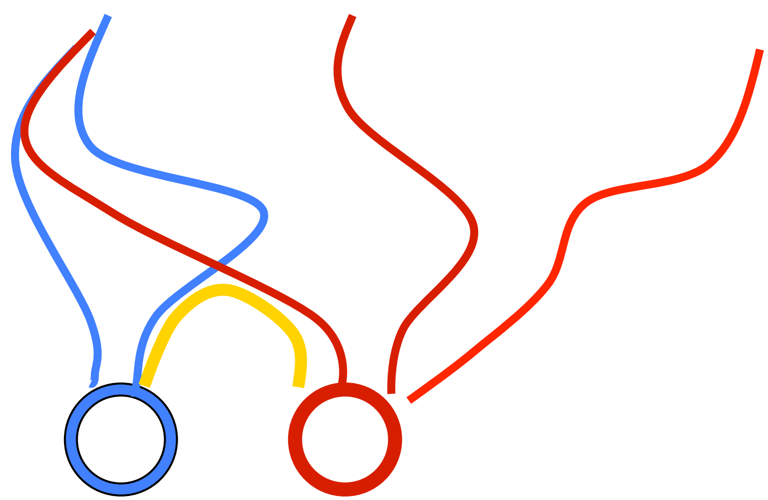


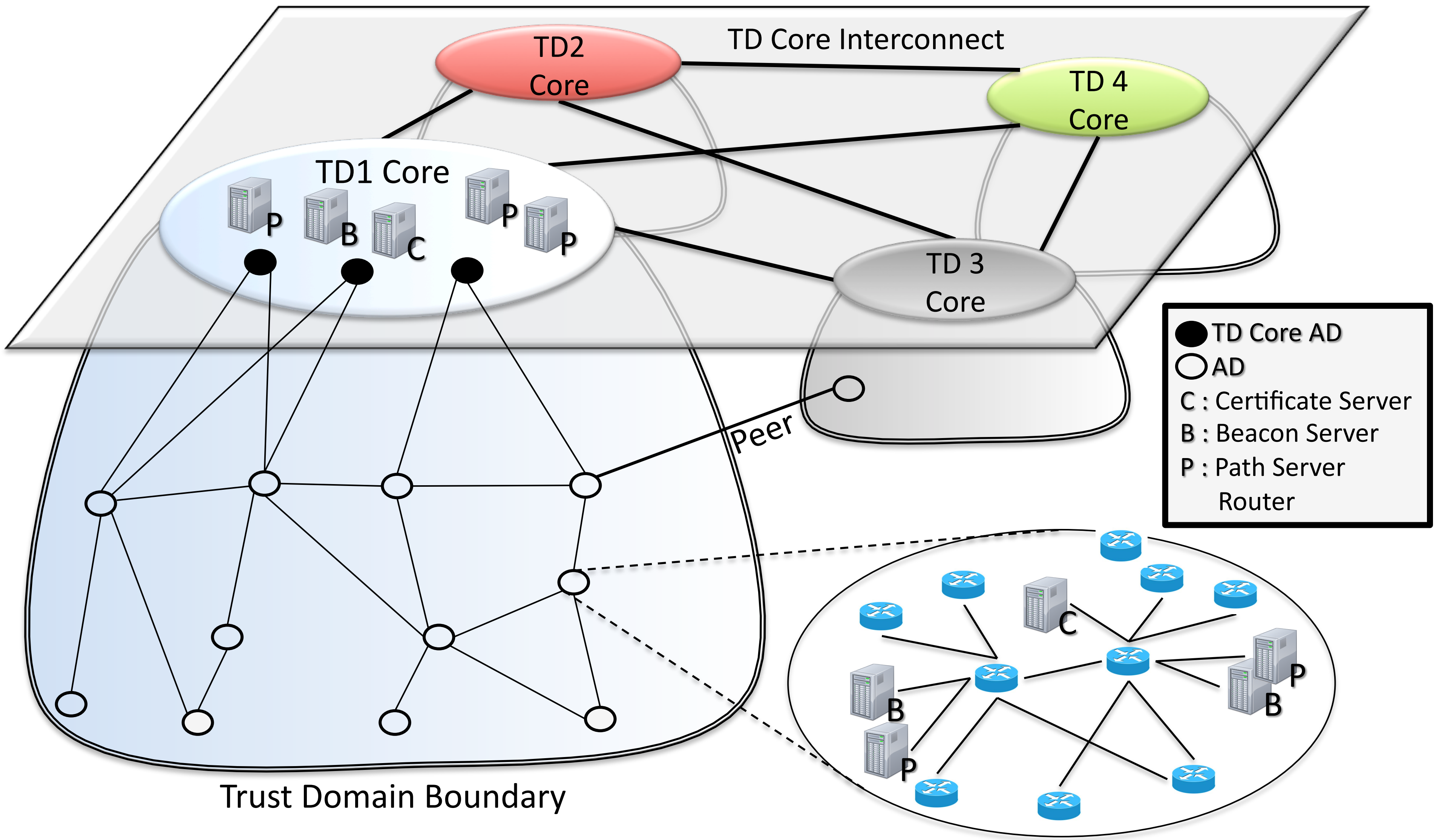
Address-to-Path Resolution

- TD core path server provides address/path resolution
- Endpoints can have arbitrary addresses
  - IPv4, IPv6
  - Public key [AIP 2008]
- Each AD registers AD ID / k down paths at TD Core path server
  - Uses up-path to reach TD core
- Two stages for name resolution
  - Name → EID, AD ID
  - AD ID → k down paths

Route Joining

- Sender obtains receiver's *k* down-paths from path server
- Sender intersects its up-paths with receiver's down-paths
- Sender selects preferred routes based on k^2 options




TD Core Interconnect
TD2 Core
TD 4 Core
TD1 Core
TD 3 Core
P
B
C
P
P
Peer
TD Core AD
AD
C : Certificate Server
B : Beacon Server
P : Path Server
Router
Trust Domain Boundary

Inter-TD Forwarding

- TD Cores recursively execute SCION beaconing to create paths amongst each other
  - Each TD core initiates PCB to neighboring TD cores
  - Propagates TD Core PCBs to create inter-TD-core path
- Endhosts can request paths to reach any other TD Core
- Endhosts combine up path + inter-TD-core path + down path
  - Provides explicit trust, as entire path is known

SCION Advantages

- Security
  - Isolation of data plane from control plane
    - Data plane still usable even if control plane disrupted
  - Cryptographic validation of packet header
  - Trust agility: local & selectable roots of trust (no global root of trust)
  - Avoidance of BGP / IP attacks (blackhole, wormhole, etc.)
  - No single point of failure
  - Explicit trust for packet forwarding, small Trusted Computing Base (TCB)
- Reliability
  - Isolation between mutually untrusted network domains
  - Multi-path forwarding, dozens of potential paths available
  - ISP / sender / receiver controllable paths
  - Instant convergence of routing protocol
  - No route-flap dampening necessary
- Efficiency
  - Scalability: routing overhead independent on # of destinations
  - Low energy forwarding: no TCAM for routing tables
  - No routing / forwarding tables
  - Low packet overhead

SCION Disadvantages

- Constant update of downpaths
- New protocols, new equipment
- Packet header larger than IP
- Static path binding
  - No automated route failure recovery

SCION Stakeholder Pros and Cons

- Manufacturers
  - ✓ Sale of additional equipment
  - ⚡ Commoditization: routers become simple and inexpensive
- ISPs
  - ✓ New revenue streams through service differentiation
  - ✓ High-availability service offerings, powerful DDoS defenses
  - ✓ Resilient to attacks and configuration errors
  - ✓ Incremental update, only new edge routers needed, inexpensive routers
  - ⚡ New equipment, new protocols
- Consumers
  - ✓ High reliability and availability
  - ✓ Differentiated services, path choice, trading off quality and price
  - ✓ Trust agility
  - ⚡ Software / HW upgrade
- Government
  - ✓ High reliability and availability for critical services
  - ✓ Selectable roots of trust, no single global root of trust
  - ✓ Simple, verifiable router hardware

Resolved BGP / Control Plane Issues

- Lack of fault isolation
  - Error propagation, potentially to entire Internet, disruption of flows outside domain
  - Adversary can attract flows outside domain (blackhole attacks)
  - Black art to keep BGP stable, manual rule sets, unanticipated consequences
- Lack of scalability, amount of work by BGP is O(N), N number of destinations
  - Path changes need to be sent to entire Internet
- S-BGP requires single root of trust for AS and address certificates
- Dramatically higher router overhead during periods of route instability
  - Increased number of routing updates during DDoS attacks
- Slow route convergence
  - Convergence attack
  - Network may require minutes up to tens of minutes to converge
- Lack of freshness for BGP update messages
- Circular dependency of UPDATES and RPKI data
- Route flap dampening-based attacks

Resolved IP / Data Plane Issues

- Complex route table lookup for each packet
- Bursting routing tables
- Lack of predictability for path availability
- Lack of route choice/control by senders and receivers

Resolved IP / BGP / Misc. Issues

- No path predictability due to inconsistency between routing table and BGP updates
- No isolation between control and data planes (routing and forwarding)
  - By attacking routing, prevent forwarding to work correctly
- Huge TCB (entire Internet)
- Single root of trust for DNSsec
- Intermittent routing loops during BGP convergence, need TTL to avoid packet looping

Incremental Deployment

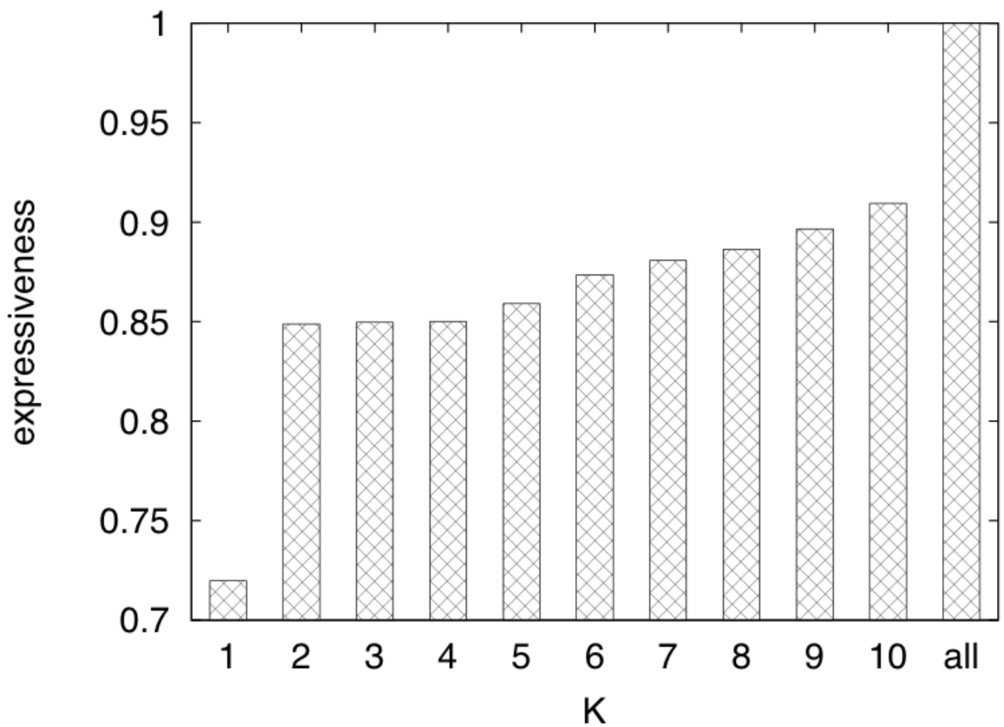
- Current ISP topologies consistent with SCION TDs
- Minor changes for ISPs
  - SCION edge router deployment
  - Beacon / certificate / path server deployment (1 host)
  - Regular MPLS forwarding internally
  - IP tunnels connect SCION to edge routers in different ADs
- Minor changes in end-domains
  - IP routing used for basic connectivity
  - SCION gateway enables legacy end hosts to benefit from SCION network

Evaluation

- Methodology
  - Use of CAIDA topology information
  - Assume 5 TDs (AfriNIC, ARIN, APNIC, LACNIC, RIPE)
  - We compare to S-BGP/BGP
- Metric 1: additional path length (AD hops) compared to BGP
  - *Without* shortcuts: 21% longer
  - *With* shortcuts:
    - 1 down/up- path: 6.7% longer
    - 2 down/up- path: 3.5% longer
    - 5 down/up- path: 2.5% longer

Evaluation (cont'd)

- Metric 2: Expressiveness
  - Fraction of BGP paths available under SCION



Related Work

- Routing security
  - S-BGP, soBGP, psBGP, SPV, PGBGP
  - Only topological correctness; addressed a subset of attacks addressed in SCION
  - H-NPBR provides robustness in Byzantine environments, but efficiency is a concern
- Routing control
  - Multipath (MIRO, Deflection, Path splicing, Pathlet), NIRA
  - Only given control to the source, and/or little security assurance
- Next-generation architectures
  - HLP, HAIR, RBF, AIP, ICING/IGLOO
  - Focusing on other aspects (reducing routing churns and routing table sizes, enforcing routing policies, and providing source accountability)

SCION Conclusions

- Basic architecture design for a next-generation network that emphasizes isolation, control and explicit trust
- Highly efficient, scalable, available architecture
- Enables numerous additional security mechanisms, e.g., network capabilities, DoS defenses