Decentralized Internet Resource Trust Infrastructure

Bingyang Liu, Fei Yang, Marcelo Bagnulo, Zhiwei Yan, and Qiong Sun

Huawei

UC3M

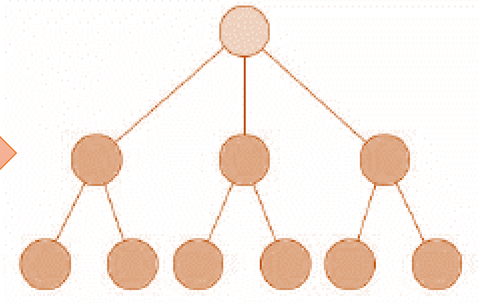
CNNIC

China Telecom

Critical Internet Trust Infrastructures are Centralized

RPKI	IP addresses and ASNs
DNSSEC	Domain names
PKI	Identities

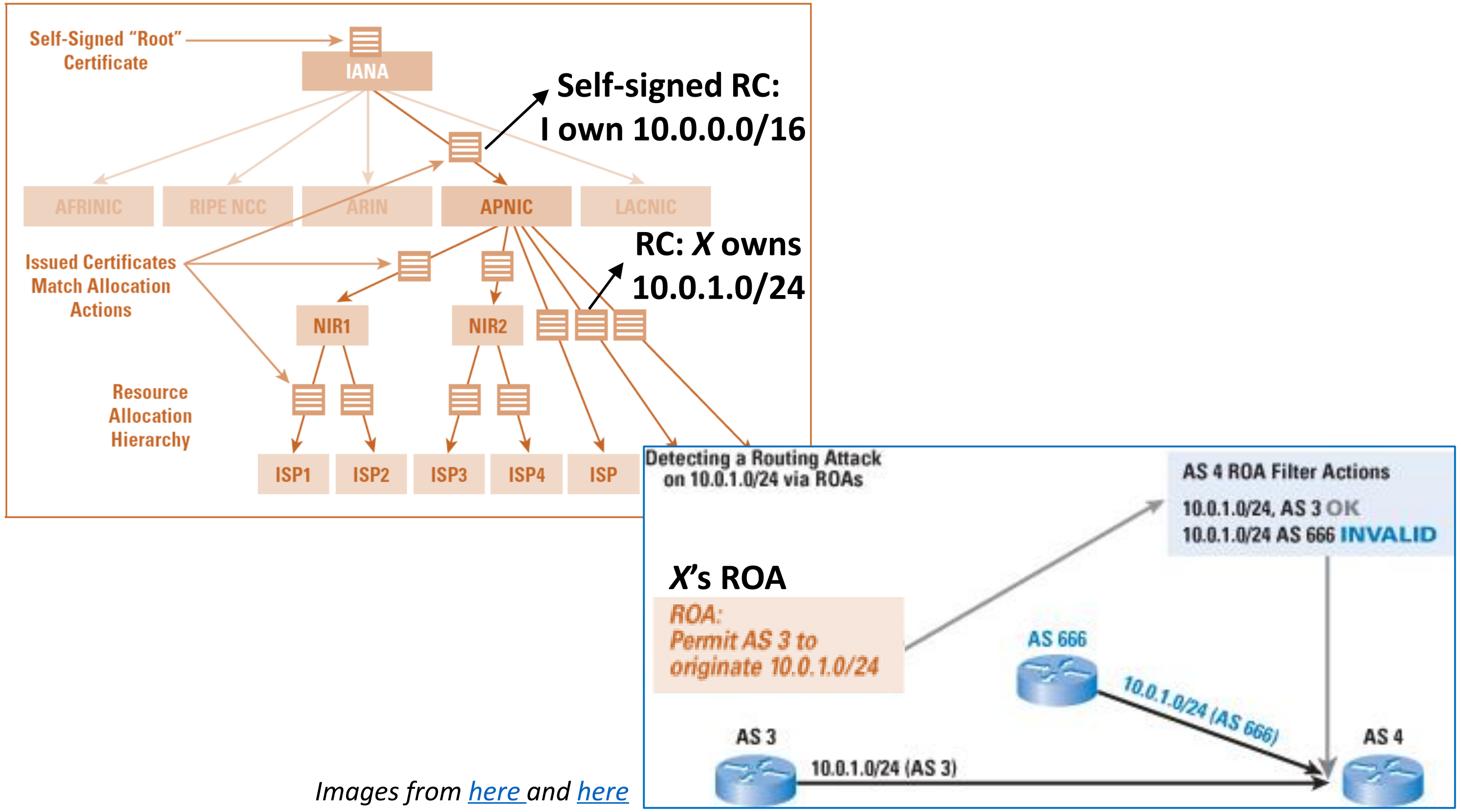
They all have centralized/hierarchical structures



Root nodes often have privilege over sub-trees. Malicious or misconfigured roots can cause problems to sub-trees.

This talk is focused on RPKI

1. Resource PKI follows the **hierarchy** of IP address allocation. IANA, RIRs and NIRs are roots of (sub-)trees
2. Parent node signs a resource certificate (**RC**) to child for **address ownership**
3. Address owner signs a route origin authentication (**ROA**) to **map prefixes to ASNs**
4. **BGP** routers rely on ROAs to detect route origin hijack (fake prefix->ASN mapping)



Self-Signed "Root" Certificate
IANA
Self-signed RC:
I own 10.0.0.0/16
AFRINIC
RIPE NCC
ARIN
APNIC
LACNIC
Issued Certificates Match Allocation Actions
RC: *X* owns
10.0.1.0/24
NIR1
NIR2
Resource Allocation Hierarchy
ISP1
ISP2
ISP3
ISP4
ISP
Detecting a Routing Attack on 10.0.1.0/24 via ROAs
AS 4 ROA Filter Actions
10.0.1.0/24, AS 3 OK
10.0.1.0/24 AS 666 INVALID
X's ROA
ROA: Permit AS 3 to originate 10.0.1.0/24
AS 666
10.0.1.0/24 (AS 666)
AS 3
AS 4
10.0.1.0/24 (AS 3)

Images from here and here

Misbehaving RPKI Authorities Cause Risks to BGP

- **The flipped threat model: BGP route is legitimate while RPKI is at fault.**
  1. *On the Risk of Misbehaving RPKI Authorities [2014 IRTF ANRP]*
  2. *From the Consent of the Routed: Improving the Transparency of the RPKI [SIGCOMM 14]*

- **Misbehaving authority can unilaterally takedown descendant’s valid routes,** by adding or wracking ROAs, by revoking, deleting, overwriting RC/ROA objects.
  - [Mis-add an ROA] Dec 13, 2013: a new ROA was (mis-)added to the production RPKI rooted at ARIN, authorizing prefix 173.251.0.0/17 with maxlength 24 to AS 6128. This caused a large portion of the address space to downgrade from “unknown” to “invalid”, including several legitimate /24 routes.
  - [Mis-delete an ROA] Dec 19, 2013, a ROA for (79.139.96.0/24, AS 51813), for a network in Russia, was (mis-)deleted from the production RPKI. Meanwhile, since at least November 21, the RPKI also had a covering ROA mapping 79.139.96.0/19-20 to another Russian ISP, AS 43782. The covering ROA caused the route corresponding to the whacked ROA to downgrade from valid to invalid.

Root cause is an entity does not independently owns its address space. Instead, its parent or ancestor have privilege to manage its RC and ROA.

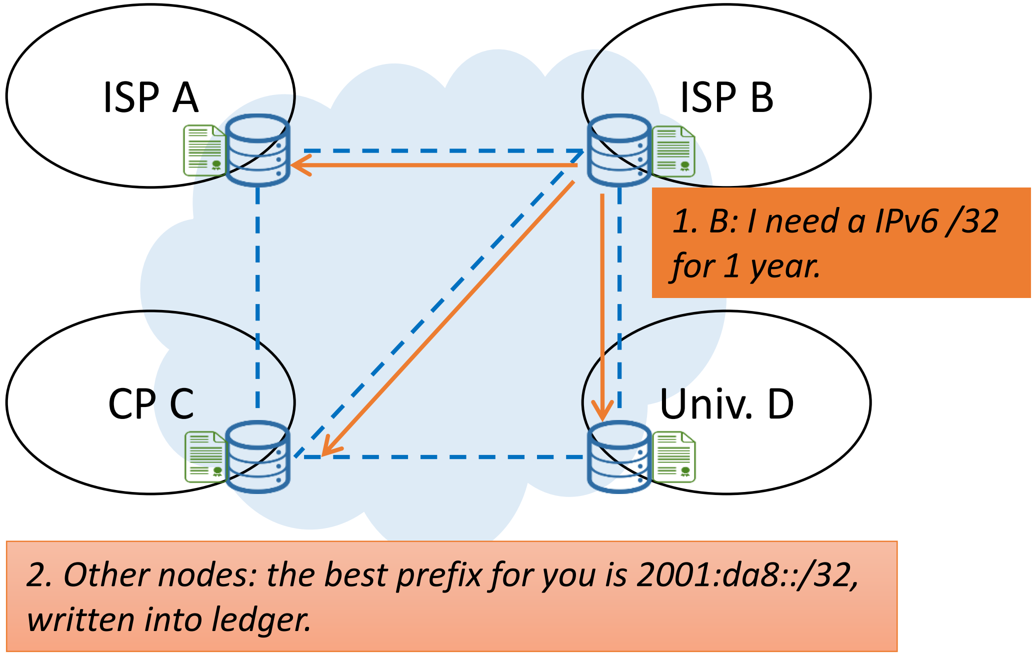
Design Goals

- **Top Goals**:
  - Organization (ISP, CP, enterprise) independently owns its resources.
  - The validity of resource ownership and mapping only depends on the owner itself, instead of any third party.
- **Other goals**:
  - Prevent address exhaustion
  - Enforce prefix aggregation
  - Enforce organization-level traceability and admission control
- We will deal with IPv6 address allocation, and IPv6/IPv4 address transferring.

System Design (1): Address Ownership

- Eligible organizations run a **decentralized ledger** for consistent **prefix ownership** and **prefix-to-ASN mapping**

- **Smart contract** is used to ensure **unique and aggregated prefix allocation**
  1. ISP B sends a request for a IPv6 /32 prefix and pays annual fee in the transaction.
  2. Smart contract calculates a continuous prefix for B from available address pool and writes the transaction into ledger.
  3. If B doesn't renew the prefix before it expires, smart contract will be triggered and the prefix is returned back to the pool



1. B: I need a IPv6 /32 for 1 year.

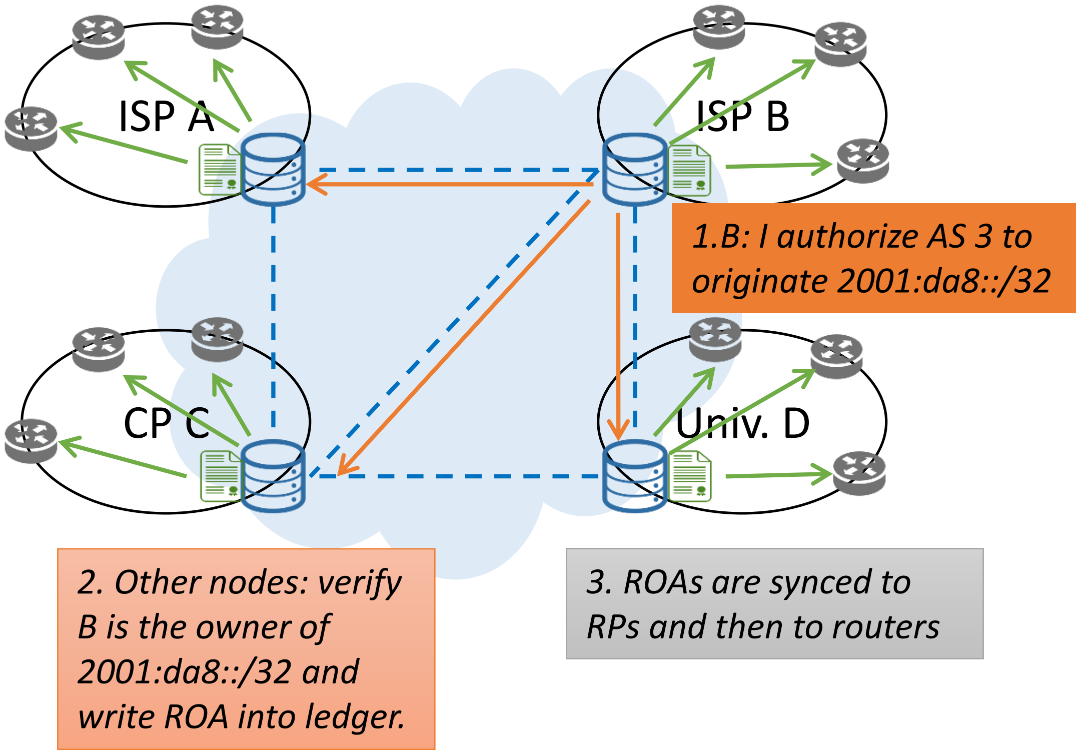
2. Other nodes: the best prefix for you is 2001:da8::/32, written into ledger.

System Design (2): Prefix-to-AS Mapping (ROA)

1. Address owner initiates an ROA as a transaction

2. Smart contract verifies the address ownership, and writes it into ledger

3. Relying parties get updated ROAs from the ledger, and sync to BGP routers, which then verify BGP routes



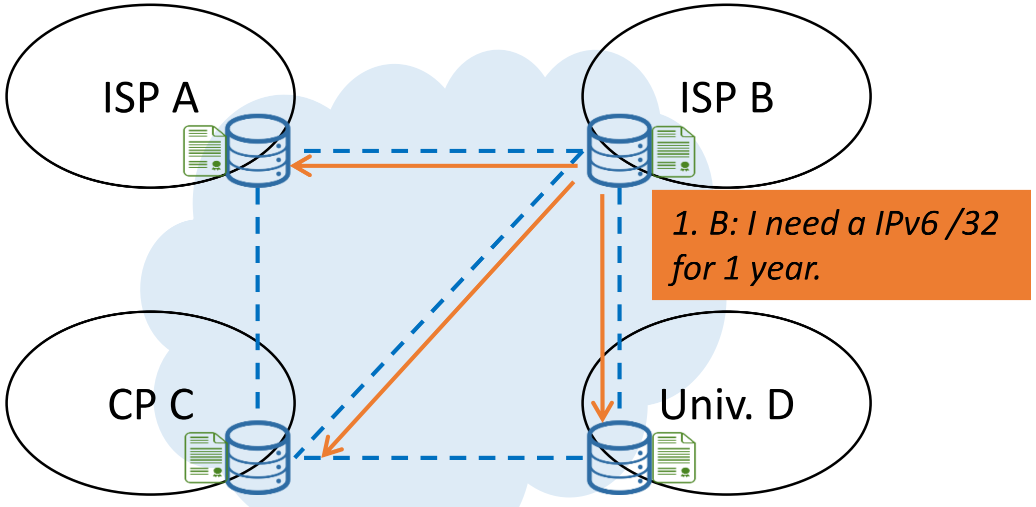
1.B: I authorize AS 3 to originate 2001:da8::/32

2. Other nodes: verify B is the owner of 2001:da8::/32 and write ROA into ledger.

3. ROAs are synced to RPs and then to routers

System Design (3): Prevent Address Exhaustion

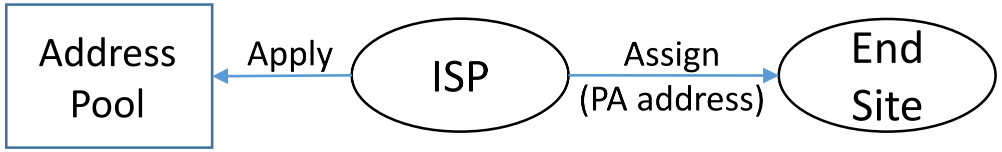
- End-sites can get smaller address space, e.g., /48. ISPs can get a new /32 if its host density ratio (RFC4692) is over the threshold
  - HD-ratio = log(#_Assigned_/56) / log(#_Allocated_/56)
  - Assignment of PA addresses is also logged in ledger
  - Smart contract can then calculate HD-ratio before agreeing on the /32 allocation



1. B: I need a IPv6 /32 for 1 year.

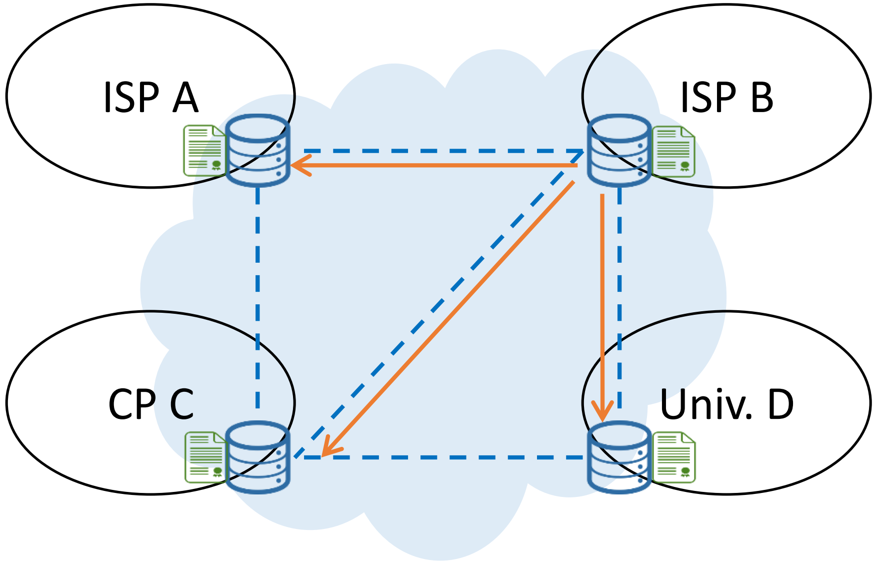
*2. Other nodes: **verify HD-ratio**, and calculate the best prefix for you is 2001:da8::/32, written into ledger.*

- Today, RIR annual fee for /32 is $1000 ~ $2500, and /48 is $100 ~ $800 (more expensive per /56). If $2000 annual fee is applied to a /32:
  - $\$2000 * 2^{32} = \$8*10^{12}$~ 10.5% world GDP, making exhaustion attack impractical
  - Although not entire address space is unicast, longer prefixes are more expensive and /32 requires HD-ratio, the cost still efficiently prevents exhaustion attack
  - *Money can be given to miners and IETF?*



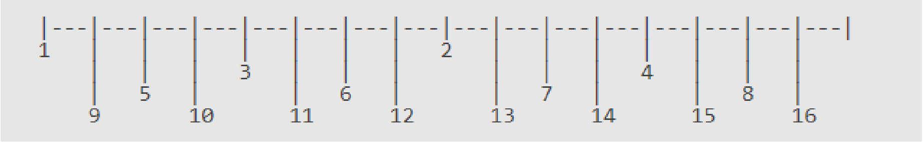
System Design (4): Prefix Aggregation

- An entity cannot decide which prefixes it gets. Instead, it can only request the size of address space, and smart contract will calculate the best prefix for it
  - “Best” is in the sense of prefix aggregation



- Smart contract runs sparse delegation algorithm used by RIRs. It allows address owner to grow, and avoids fragmentation
  - Sparse address for the new user.
  - Adjacent address for the same user.

Sparse Delegation Sequence



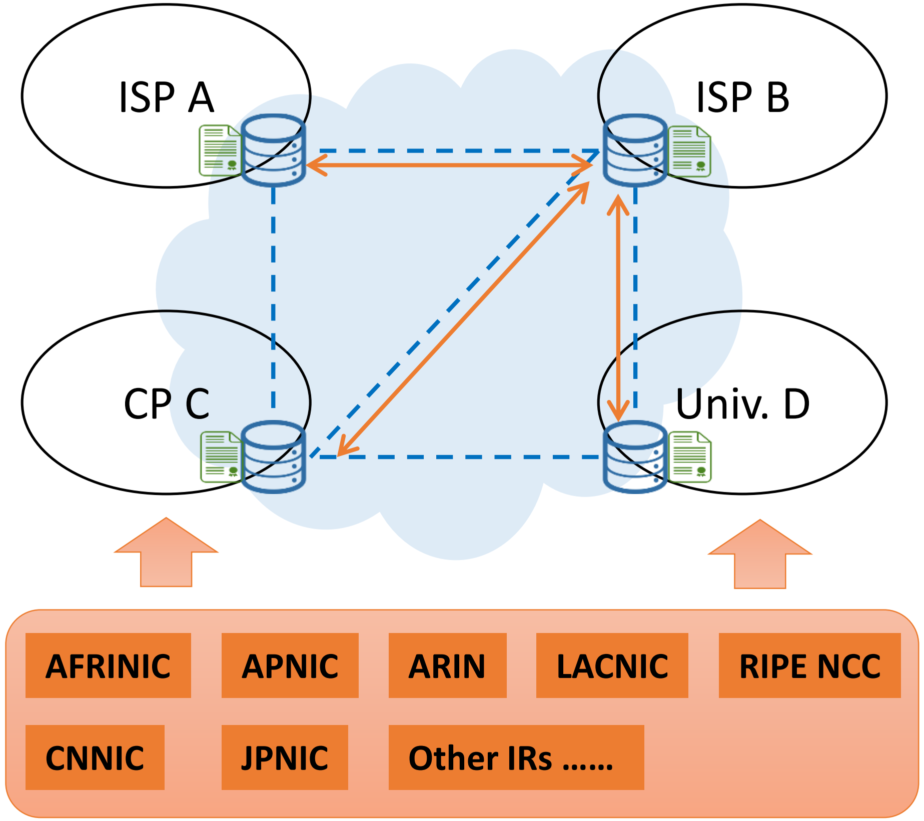
```
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
1   |   |   |   |   |   |   |   2   |   |   |   |   |   |   |
    |   |   |   3   |   |   |       |   |   |   4   |   |   |
    |   5   |       |   6   |       |   7   |       |   8   |
    9       10      11      12      13      14      15      16
```

System Design (5): Admission Control

- For purposes of security and traceability, only organizations authorized by RIRs, NIRs or LIRs are eligible to the ledger
  - Identity information is registered for accountability and traceability (like WHOIS)

- So we use **permissioned ledger**. Only entities whose identities are endorsed by IRs are permitted.

- Unlike today, **XIRs are only endorsers. They do not own or control resources. RIRs & NIRs are equal and independent**



Open Question

- Interdependency between BlockChain and BGP
  - The decentralized ledger is a P2P network built upon underlying routing (BGP)
  - It is still an open chicken-egg problem. Actually, RPKI has the same problem
- Consensus algorithm
  - We are implementing a permissioned Ethereum, which supports POW and POS.
  - However, eventually we may need a best algorithm for our application. SCP?
- How to get started
  - Request for an unsinged /20 IPv6 address space to do experiment, so that the solution will not have conflict with RPKI.
  - After real-world experiments, the address should be kept as ordinary address

Thanks, and welcome to join in us!

liubingyang@huawei.com