Shim6: Multihoming for IPv6

Sébastien Barré

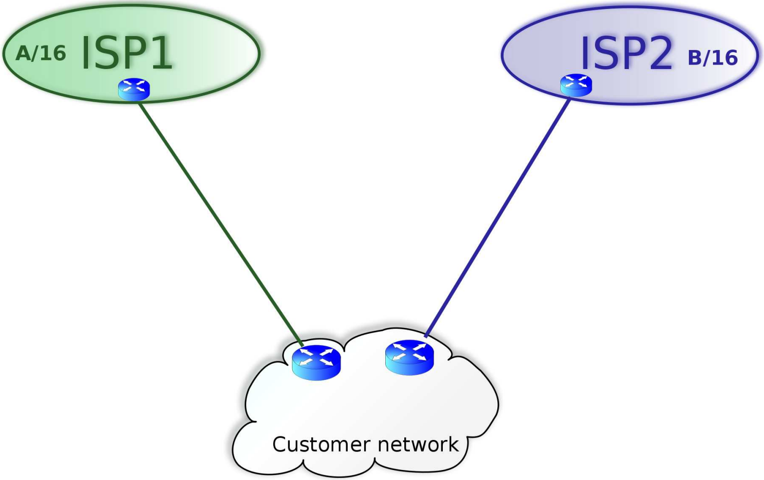
Université catholique de Louvain
http://inl.info.ucl.ac.be

Nov. 18th, 2008

INGI Research Meeting

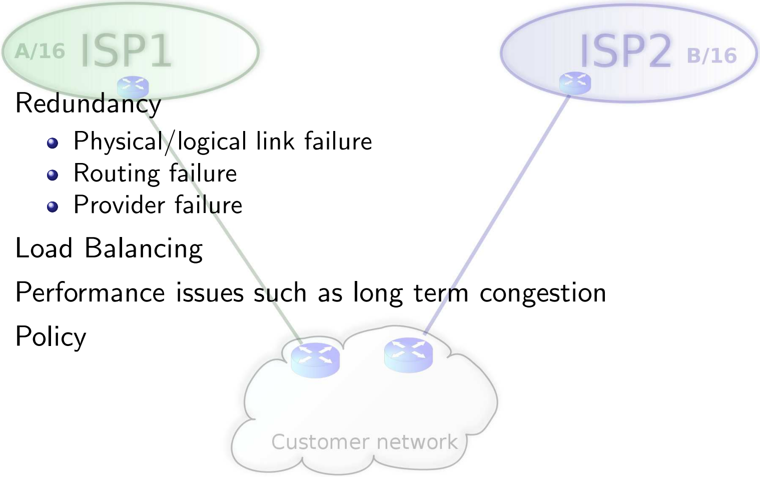
What is multihoming ?

What is multihoming ?



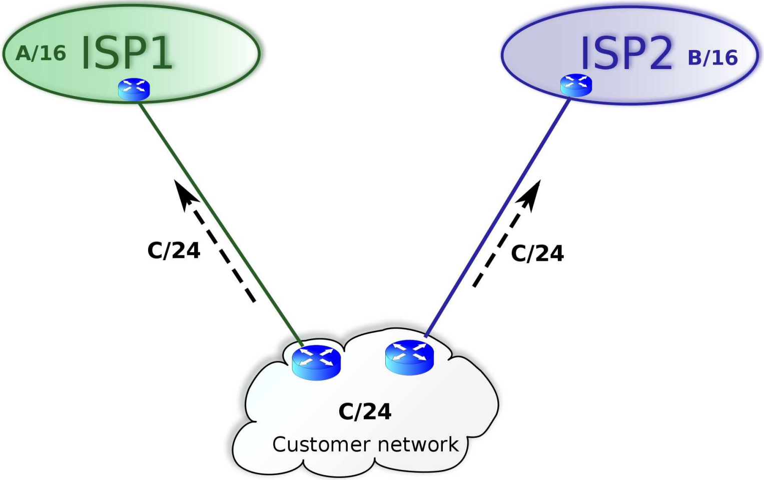
Motivations for Multihoming

- Redundancy
  - Physical/logical link failure
  - Routing failure
  - Provider failure
- Load Balancing
- Performance issues such as long term congestion
- Policy

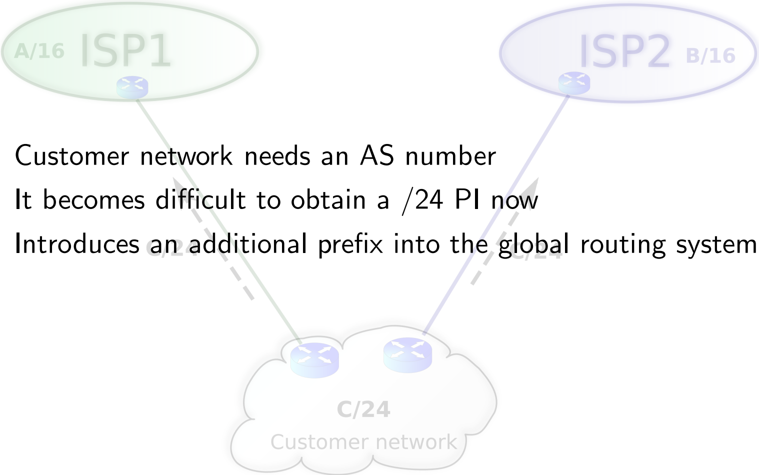


1. Introduction
   - Multihoming with IPv4
   - Motivations for IPv6
   - IPv6 addresses
2. The Shim6 protocol
3. LinShim6 implementation for Linux
4. Conclusion

Using a Provider Independent (PI) IPv4 address block



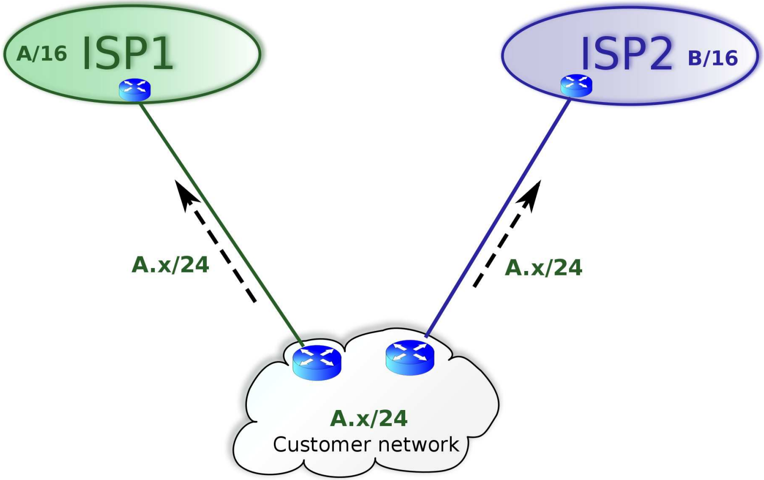
Using a Provider Independent (PI) IPv4 address block



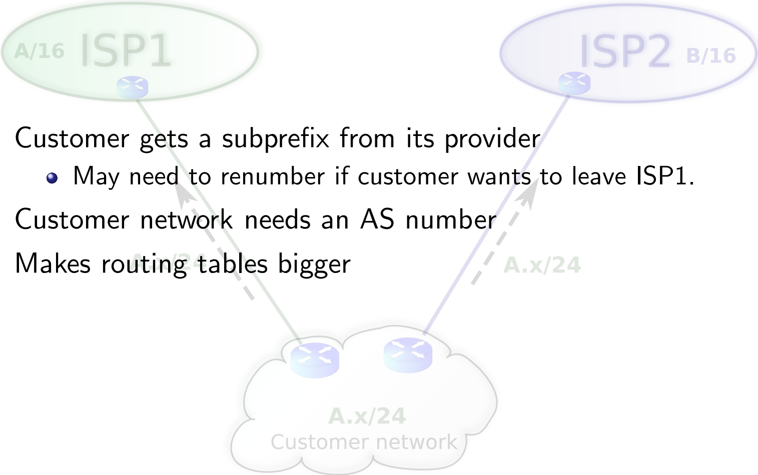
- Customer network needs an AS number
- It becomes difficult to obtain a /24 PI now
- Introduces an additional prefix into the global routing system

Réf.: Abley et al., RFC4116, *IPv4 multihoming practices and limitations*

Using a Provider Aggregatable (PA) IPv4 address block



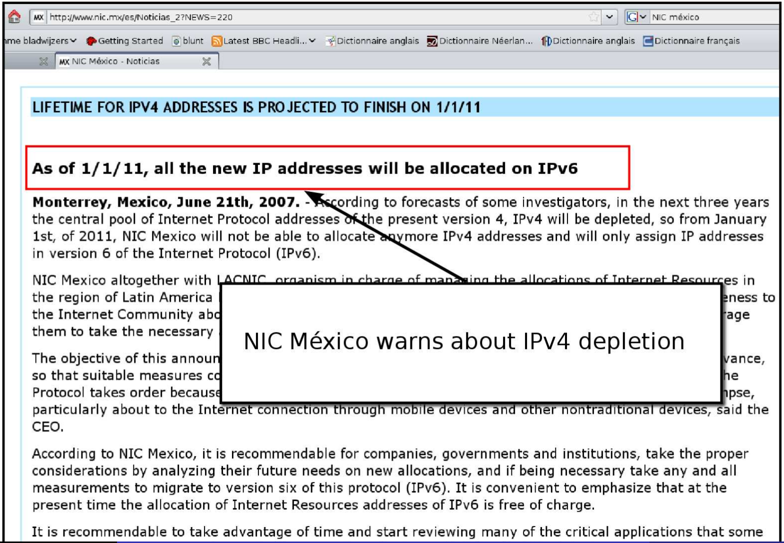
Using a Provider Aggregatable (PA) IPv4 address block



- Customer gets a subprefix from its provider
  - May need to renumber if customer wants to leave ISP1.
- Customer network needs an AS number
- Makes routing tables bigger

Réf.: Abley et al., RFC4116, *IPv4 multihoming practices and limitations*

Mainly: IPv4 address depletion



http://www.nic.mx/es/Noticias_2?NEWS=220

NIC México - Noticias

LIFETIME FOR IPV4 ADDRESSES IS PROJECTED TO FINISH ON 1/1/11

As of 1/1/11, all the new IP addresses will be allocated on IPv6

Monterrey, Mexico, June 21th, 2007. - According to forecasts of some investigators, in the next three years the central pool of Internet Protocol addresses of the present version 4, IPv4 will be depleted, so from January 1st, of 2011, NIC Mexico will not be able to allocate anymore IPv4 addresses and will only assign IP addresses in version 6 of the Internet Protocol (IPv6).

NIC Mexico altogether with LACNIC, organism in charge of managing the allocations of Internet Resources in the region of Latin America ... eness to the Internet Community abo... age them to take the necessary ...

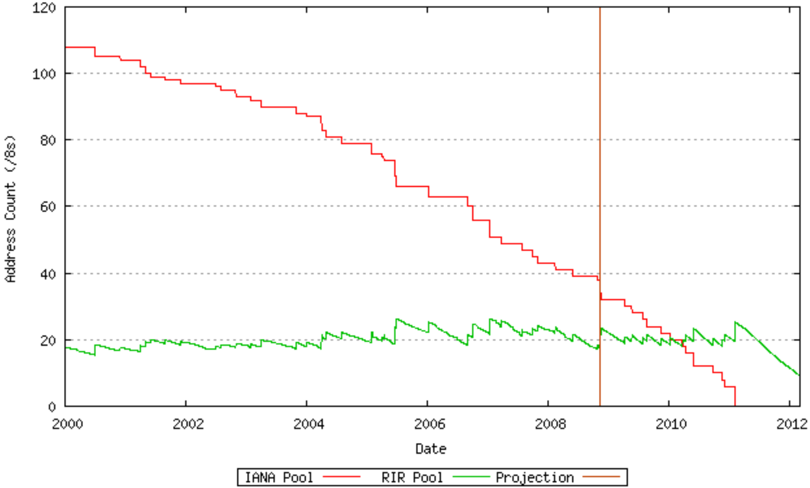
The objective of this announ... vance, so that suitable measures co... he Protocol takes order because ... pse, particularly about to the Internet connection through mobile devices and other nontraditional devices, said the CEO.

According to NIC Mexico, it is recommendable for companies, governments and institutions, take the proper considerations by analyzing their future needs on new allocations, and if being necessary take any and all measurements to migrate to version six of this protocol (IPv6). It is convenient to emphasize that at the present time the allocation of Internet Resources addresses of IPv6 is free of charge.

It is recommendable to take advantage of time and start reviewing many of the critical applications that some

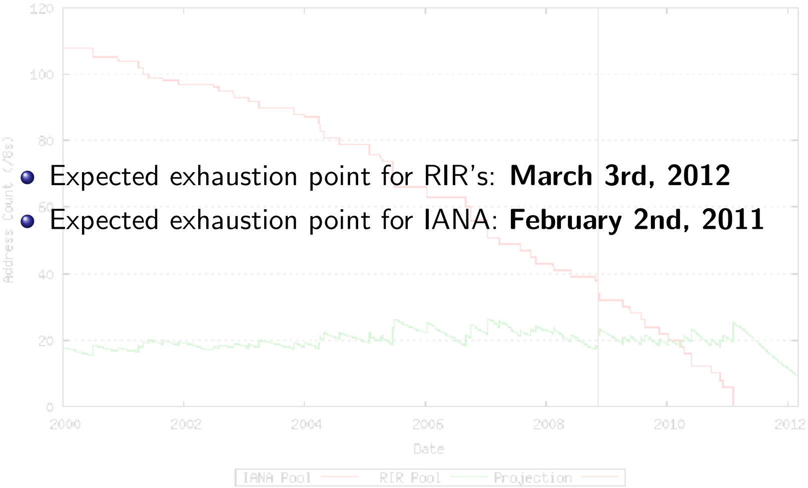
NIC México warns about IPv4 depletion

Mainly: IPv4 address depletion



Source: `http://www.potaroo.net/tools/ipv4/index.html`

IPv4 address depletion: latest informations



- Expected exhaustion point for RIR's: **March 3rd, 2012**
- Expected exhaustion point for IANA: **February 2nd, 2011**

Source: `http://www.potaroo.net/tools/ipv4/index.html`

Other expectations for IPv6

- Lower load of Internet routing tables
- Less packet processing in the core of the Internet
    - Push state towards the edges
- No more NATs: IP address for everyone
- Improved security, mobility and **multihoming**

Other expectations for IPv6

- Lower load of Internet routing tables
- Less packet processing in the core of the Internet
  - Push state towards the edges
- No more NATs: IP address for everyone
- Improved security, mobility and **multihoming**

Other expectations for IPv6

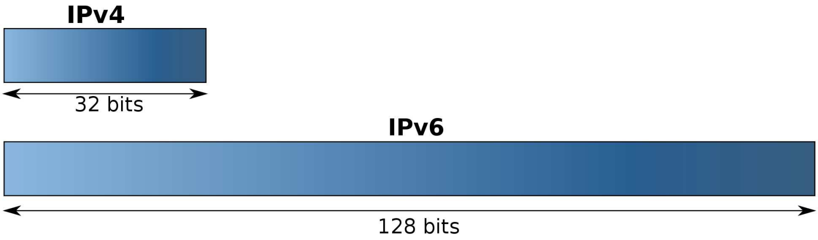
- Lower load of Internet routing tables
- Less packet processing in the core of the Internet
  - Push state towards the edges
- No more NATs: IP address for everyone
- Improved security, mobility and **multihoming**

Other expectations for IPv6

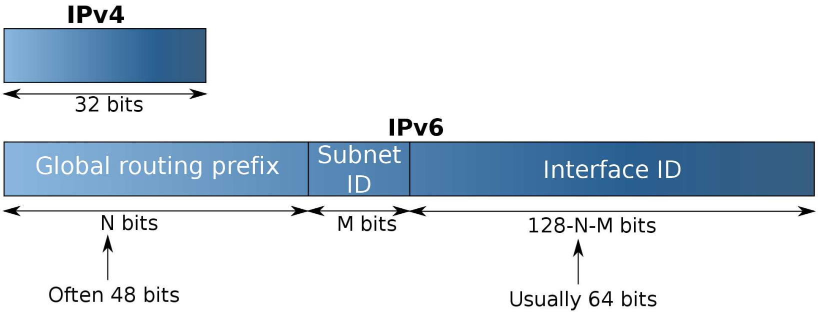
- Lower load of Internet routing tables
- Less packet processing in the core of the Internet
  - Push state towards the edges
- No more NATs: IP address for everyone
- Improved security, mobility and **multihoming**

1. Introduction
   - Multihoming with IPv4
   - Motivations for IPv6
   - IPv6 addresses
2. The Shim6 protocol
3. LinShim6 implementation for Linux
4. Conclusion

IPv6 address format



IPv6 address format



PA vs PI addresses

- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?

      ➔ Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

PA vs PI addresses

- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?
      → Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

PA vs PI addresses

- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?

      → Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

PA vs PI addresses

- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?

      → Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

PA vs PI addresses

- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?

    → Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

PA vs PI addresses

- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?

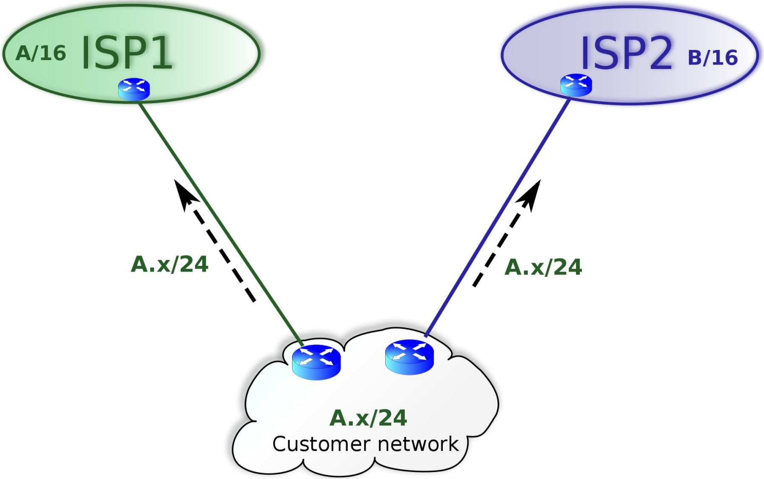
      → Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

PA vs PI addresses

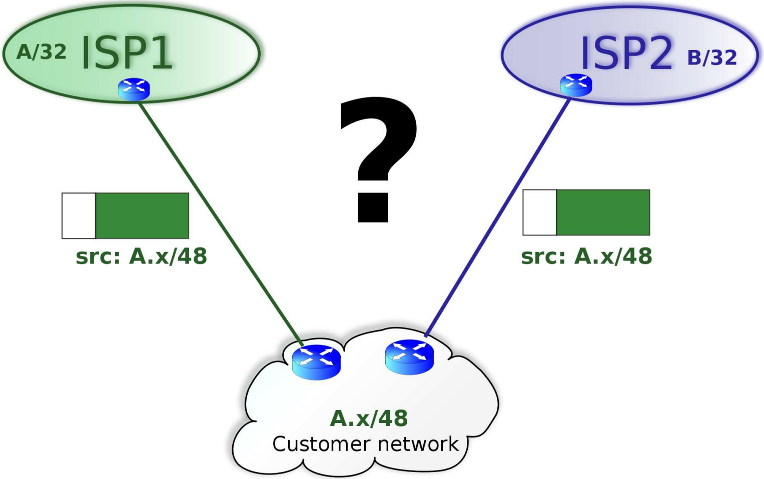
- **PI**: Provider Independent addresses
  - The site announces its PI address set through BGP
  - If multihomed: multiple BGP annoucements
    - Global announcements of PI prefixes
    - What if many sites get multihomed ?

      → Scalability problem
- **PA**: Provider Aggregatable addresses
  - The site receives a subset of its provider's addresses
  - Only the provider announces its address set through BGP
  - If multihomed : The site receives **several** address blocks

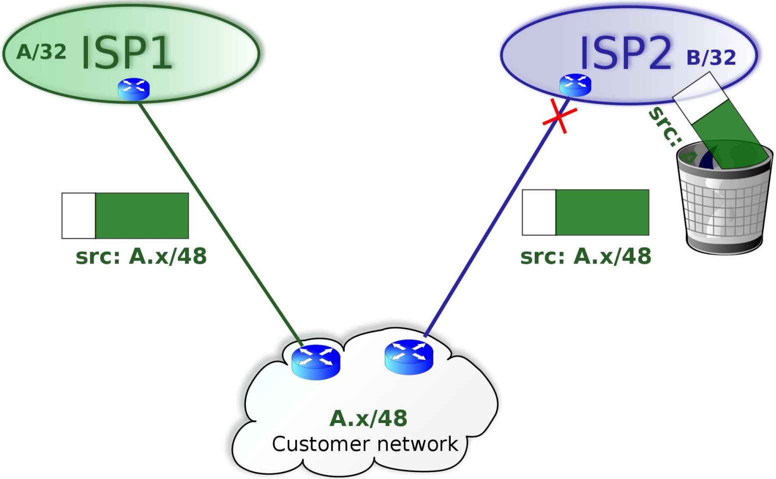
More about PA - Reminder: IPv4 PA



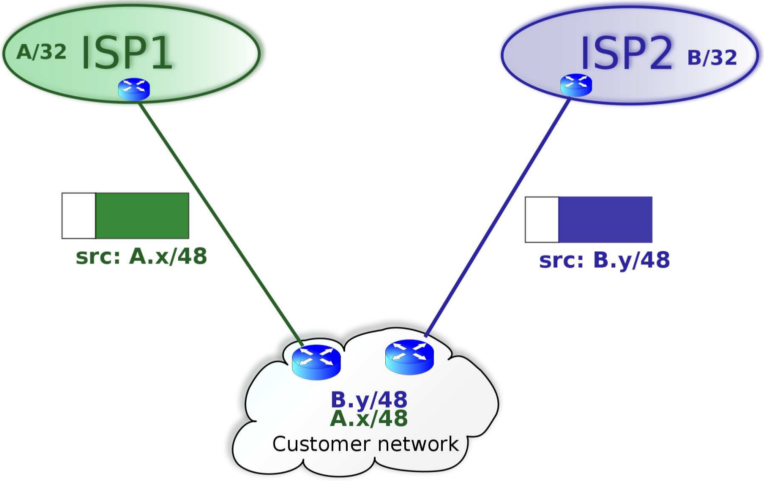
More about PA - And so... IPv6 PA ?



More about PA - And so... IPv6 PA ?



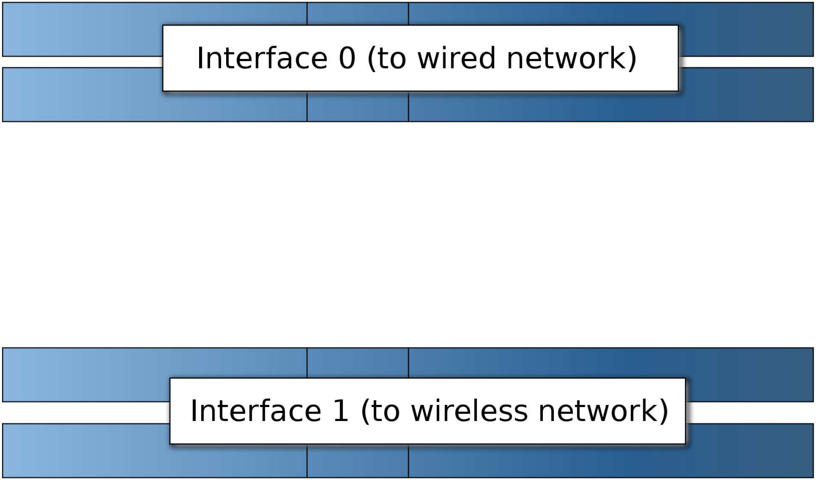
More about PA - And so... IPv6 PA ?



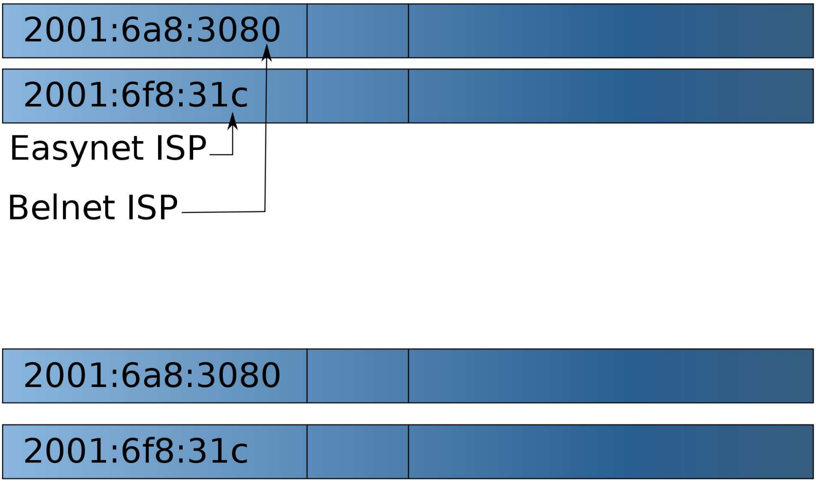
The case of UCLouvain

- Two providers, thus two global routing prefixes:
  - `2001:6a8:3080`: Provider is Belnet
  - `2001:6f8:31c`: Provider is Easynet
- Several subnetworks:
  - 2: Staff
  - 3: Servers
  - 4: Experiments
  - 2001: Wifi staff

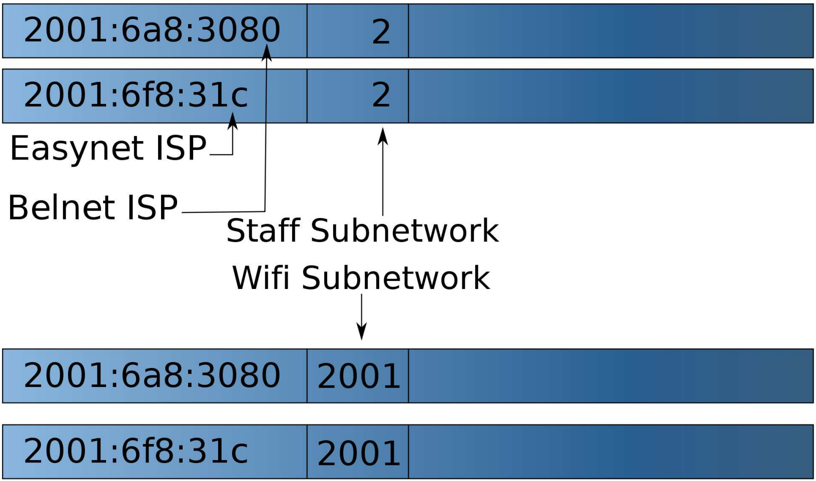
A typical laptop in our department



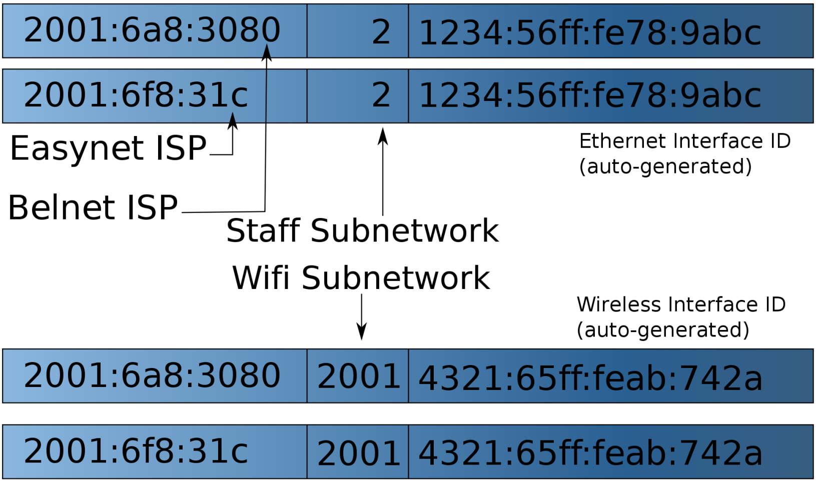
A typical laptop in our department



A typical laptop in our department



A typical laptop in our department



PA implications

- PA addresses reduce the load for the BGP system…
- …But it pushes new responsibilities to the end system
  - ➡ Failover from one address to another working one
  - ➡ Load balancing
- Those are completely managed by the network in v4.
- In v6, **part** is now managed by the end-system
  - ➡ We need to upgrade the end-hosts !

PA implications

- PA addresses reduce the load for the BGP system...
- ...But it pushes new responsibilities to the end system
  - ➔ Failover from one address to another working one
  - ➔ Load balancing
- Those are completely managed by the network in v4.
- In v6, **part** is now managed by the end-system
  - ➔ We need to upgrade the end-hosts !

PA implications

- PA addresses reduce the load for the BGP system. . .
- . . . But it pushes new responsibilities to the end system
  - ➔ Failover from one address to another working one
  - ➔ Load balancing
- Those are completely managed by the network in v4.
- In v6, **part** is now managed by the end-system
  - ➔ We need to upgrade the end-hosts !

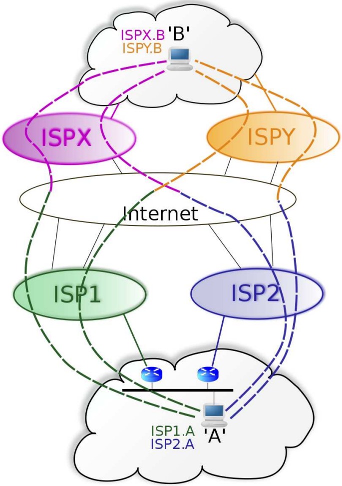
PA implications

- PA addresses reduce the load for the BGP system...
- ...But it pushes new responsibilities to the end system
  - ➔ Failover from one address to another working one
  - ➔ Load balancing
- Those are completely managed by the network in v4.
- In v6, **part** is now managed by the end-system
  - ➔ We need to upgrade the end-hosts !

PA implications

- PA addresses reduce the load for the BGP system...
- ...But it pushes new responsibilities to the end system
  - ➔ Failover from one address to another working one
  - ➔ Load balancing
- Those are completely managed by the network in v4.
- In v6, **part** is now managed by the end-system
  - ➔ We need to upgrade the end-hosts !

PA implications



End-host upgrade: the problem

- Current applications assume one <src,dest> address pair for a given communication
- They also assume that the network ensures failover if a problem happens somewhere.

  ➠ **How to manage failover in the end-host without changing applications ?**

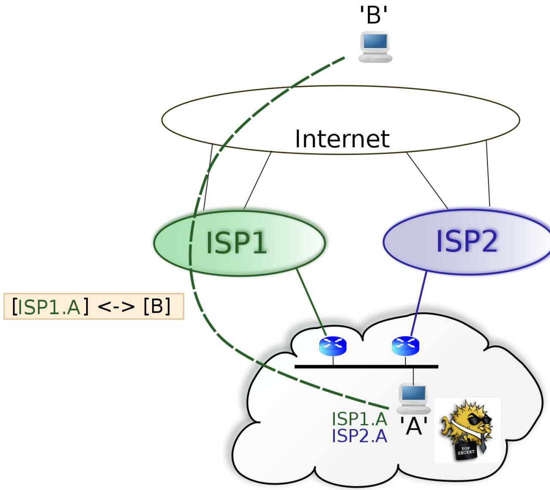
End-host upgrade: the problem

- Current applications assume one <src,dest> address pair for a given communication
- They also assume that the network ensures failover if a problem happens somewhere.
  ➔ **How to manage failover in the end-host without changing applications ?**

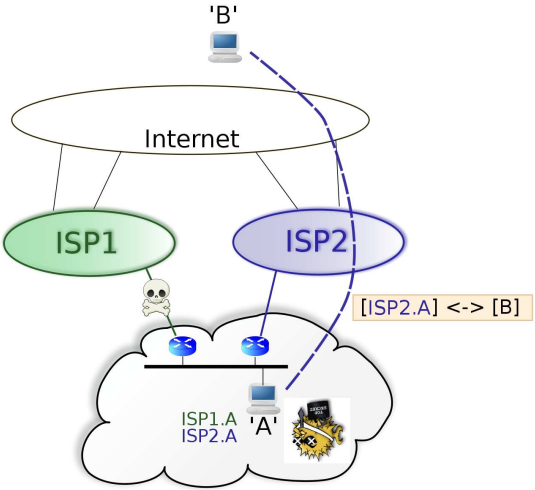
End-host upgrade: How to do it ?

- To detect failures: Monitor the communications
- To switch to a working path: Change the current address pair

End-host upgrade: a solution ?



End-host upgrade: a solution ?



Why ?

- An IP address has currently a double semantics: **Locator** and **Identifier**
  - Locator: The IP address is used to forward the packet towards its destination.
    ➡ Changing the IP address has the effect of changing the path.
  - Identifier: The IP address is used as part of the TCP context identifier
    ➡ **Changing the IP address has the effect of breaking TCP connexions**

Why ?

- An IP address has currently a double semantics: **Locator** and **Identifier**
  - Locator: The IP address is used to forward the packet towards its destination.

    → Changing the IP address has the effect of changing the path.
  - Identifier: The IP address is used as part of the TCP context identifier

    → **Changing the IP address has the effect of breaking TCP connexions**

Why ?

- An IP address has currently a double semantics: **Locator** and **Identifier**
  - Locator: The IP address is used to forward the packet towards its destination.
    ➔ Changing the IP address has the effect of changing the path.
  - Identifier: The IP address is used as part of the TCP context identifier
    ➔ **Changing the IP address has the effect of breaking TCP connexions**

The Shim6 proposal

- Separate the two semantics
- The transport and application layer see an identifier
- The network and data link layer see a locator.
- A new Shim layer **rewrites** identifiers to replace them with locators
- The same Shim layer **rewrites** locators to replace them with identifiers

The Shim6 proposal

- Separate the two semantics
- The transport and application layer see an identifier
- The network and data link layer see a locator.
- A new Shim layer **rewrites** identifiers to replace them with locators
- The same Shim layer **rewrites** locators to replace them with identifiers

The Shim6 proposal

- Separate the two semantics
- The transport and application layer see an identifier
- The network and data link layer see a locator.
- A new Shim layer **rewrites** identifiers to replace them with locators
- The same Shim layer **rewrites** locators to replace them with identifiers

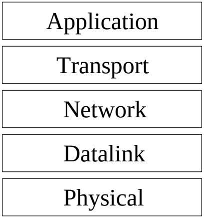
The Shim6 proposal

- Separate the two semantics
- The transport and application layer see an identifier
- The network and data link layer see a locator.
- A new Shim layer **rewrites** identifiers to replace them with locators
- The same Shim layer **rewrites** locators to replace them with identifiers

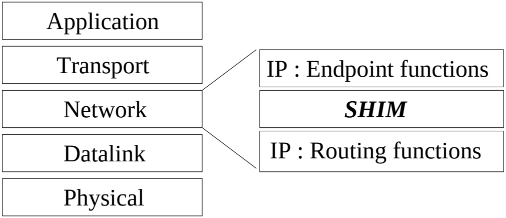
The Shim6 proposal

- Separate the two semantics
- The transport and application layer see an identifier
- The network and data link layer see a locator.
- A new Shim layer **rewrites** identifiers to replace them with locators
- The same Shim layer **rewrites** locators to replace them with identifiers

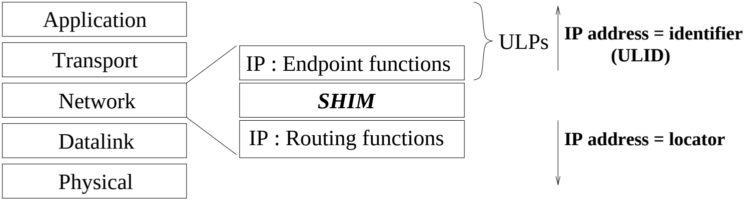
Locators vs Identifiers (ULIDs)



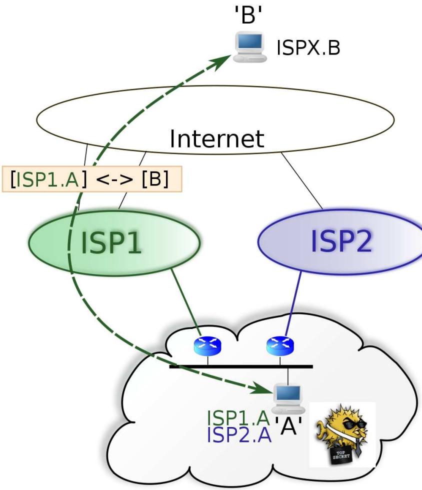
Locators vs Identifiers (ULIDs)



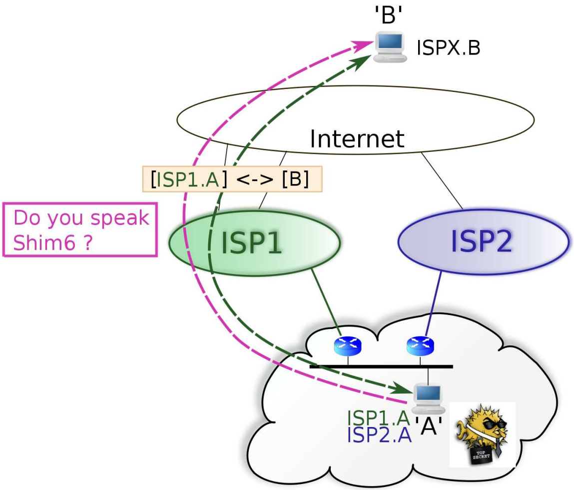
Locators vs Identifiers (ULIDs)



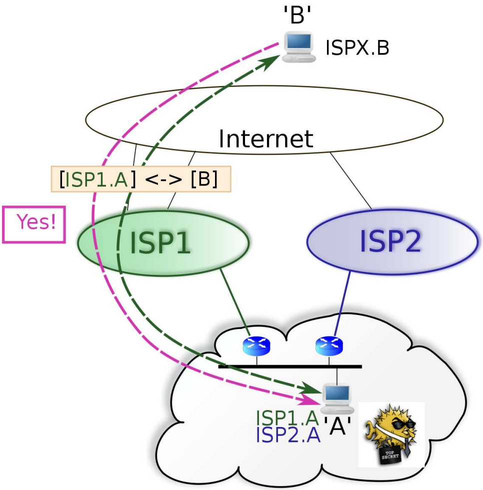
Shim6 operation



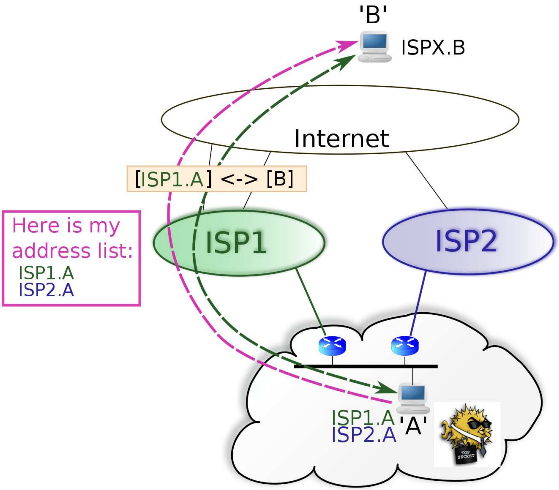
Shim6 operation



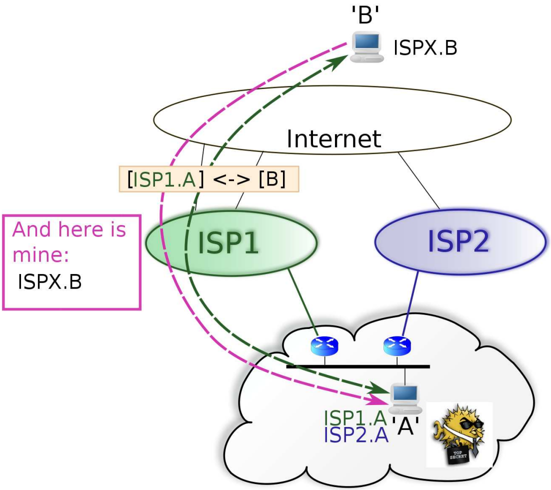
Shim6 operation



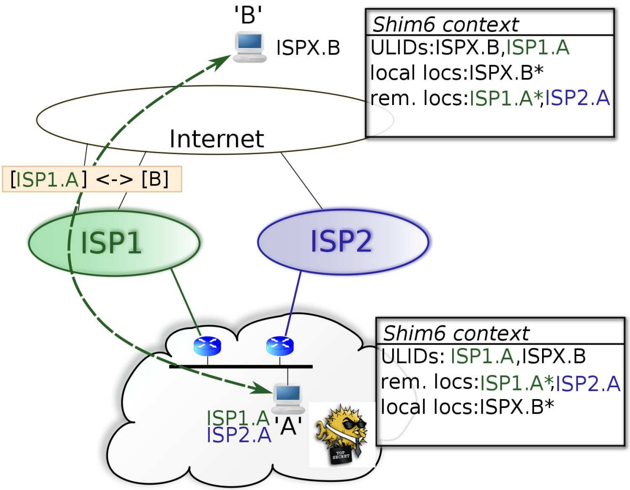
Shim6 operation



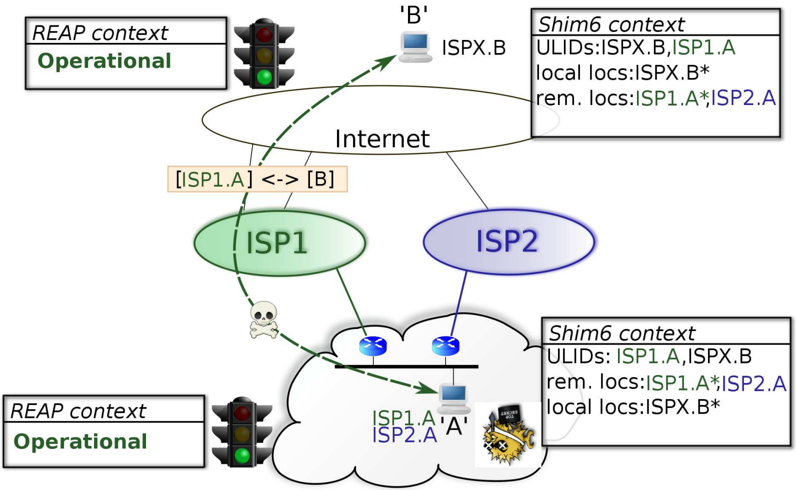
Shim6 operation



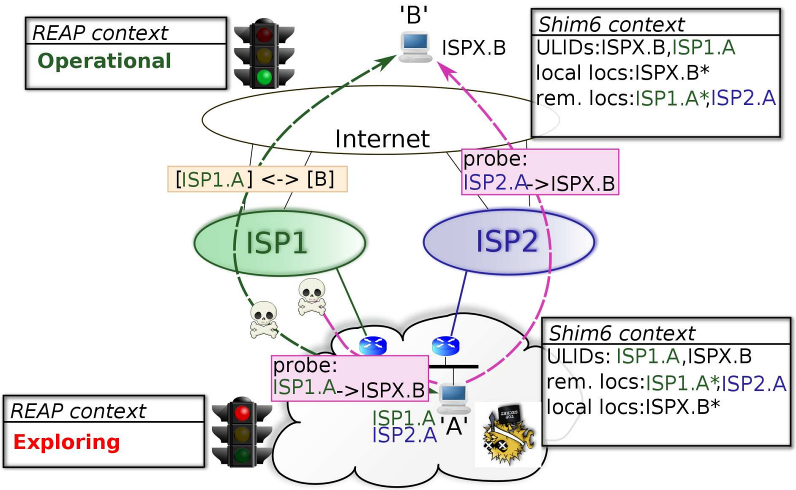
Shim6 operation



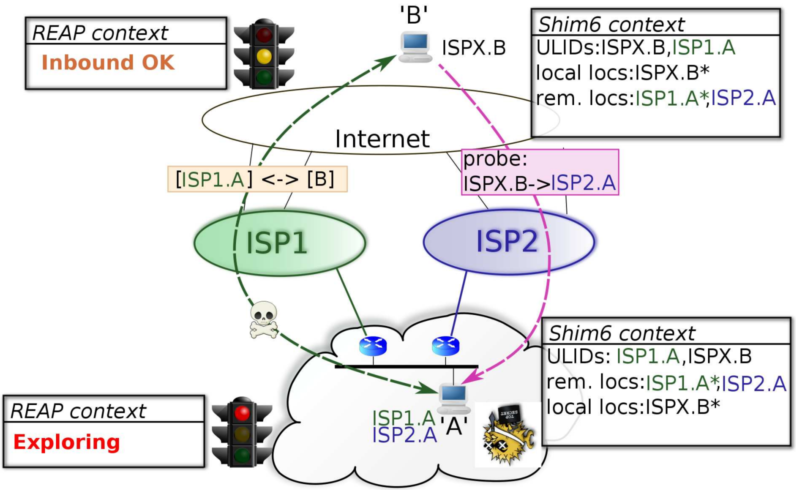
REAP operation



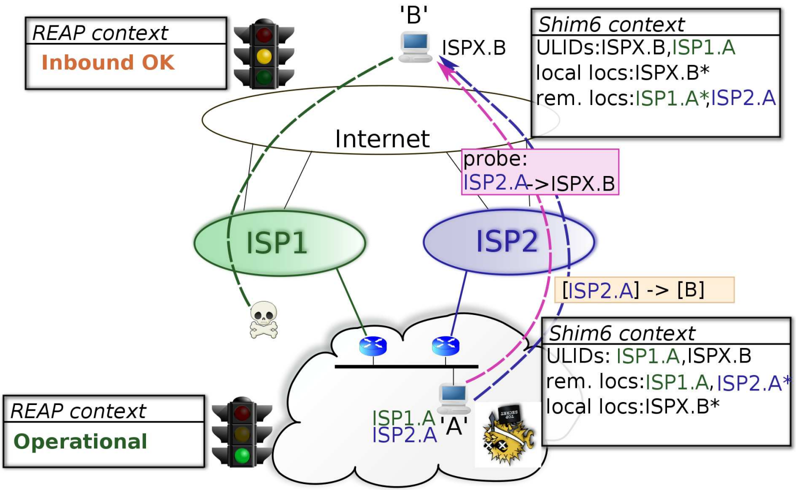
REAP operation



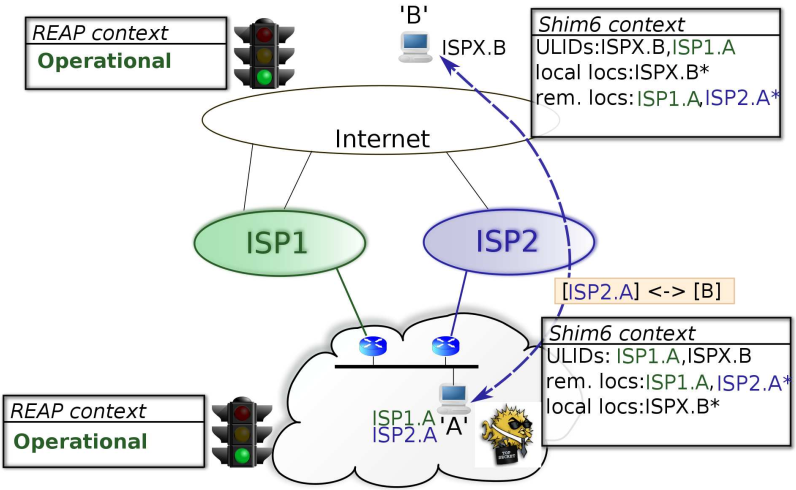
REAP operation



REAP operation



REAP operation



TCP connection survival

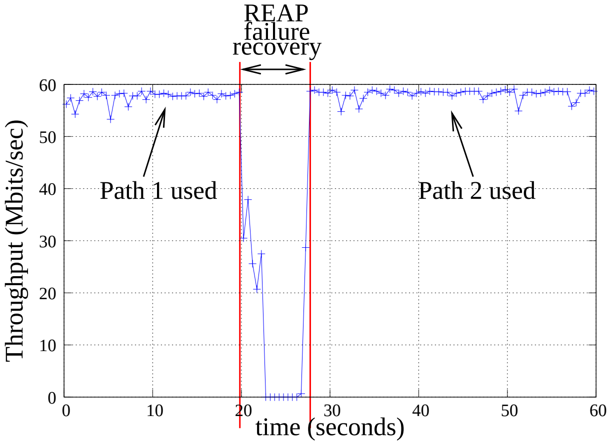
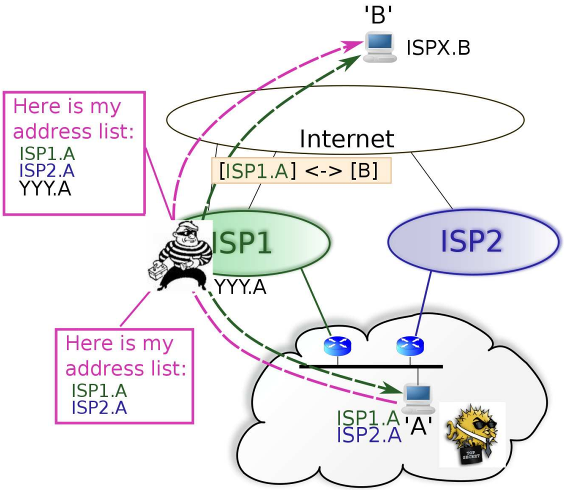


Figure: Evolution of throughput for an iperf TCP session

New solutions - new problems: the time shifting attack



How to avoid that ?

- Sign the message with a private key
- Put the public key in the message
- The receiver verifies the signature thanks to the provided public key.

How to ensure that the public key is not replaced by the attacker ?

How to avoid that ?

- Sign the message with a private key
- Put the public key in the message
- The receiver verifies the signature thanks to the provided public key.

How to ensure that the public key is not replaced by the attacker ?

How to avoid that ?

- Sign the message with a private key
- Put the public key in the message
- The receiver verifies the signature thanks to the provided public key.

How to ensure that the public key is not replaced by the attacker ?

How to avoid that ?

- Sign the message with a private key
- Put the public key in the message
- The receiver verifies the signature thanks to the provided public key.

How to ensure that the public key is not replaced by the attacker ?

How to ensure public key authenticity ?

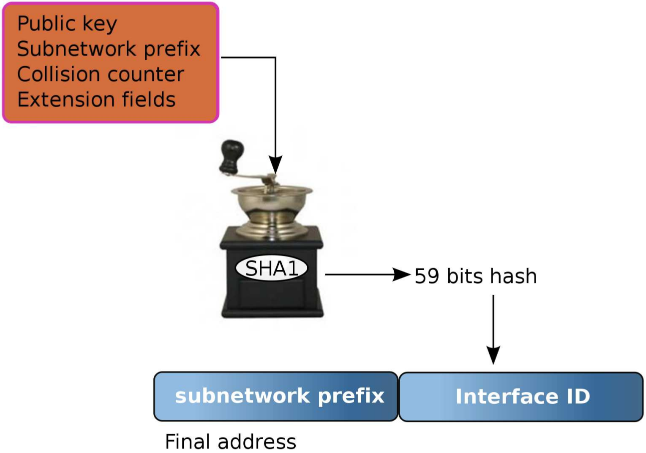
- Classical solution: Use a certificate, signed by a trusted third-party
  ➔ We cannot give a certificate to everyone in the Internet !
- We have long addresses anyway: let's embed the public key inside the address. . .

How to ensure public key authenticity ?

- Classical solution: Use a certificate, signed by a trusted third-party
  ➡ We cannot give a certificate to everyone in the Internet !
- We have long addresses anyway: let's embed the public key inside the address...

Cryptographically Generated Addresses (CGAs)

A first proposal:



Cryptographically Generated Addresses (CGAs)

- 59 bits is too short a hash to ensure that it won't be broken.
- Solution: *artificially* extend the hash length
  - Compute a second hash, with an additional input called *modifier*
  - Require that n bits be 0 in the result
  - increment the modifier and retry the hash computation until n bits are zero
    - ➡ Brute-force attack of our own address...
    - ➡ But we are $\mathcal{O}(2^{59})$ in advance over our attacker !

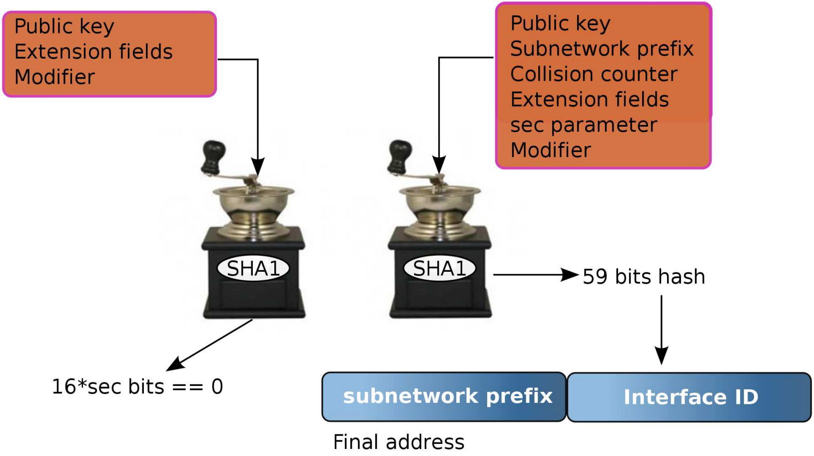
Cryptographically Generated Addresses (CGAs)

- 59 bits is too short a hash to ensure that it won't be broken.
- Solution: *artificially* extend the hash length
  - Compute a second hash, with an additional input called *modifier*
  - Require that n bits be 0 in the result
  - increment the modifier and retry the hash computation until n bits are zero

→ Brute-force attack of our own address...

→ But we are $\mathcal{O}(2^{59})$ in advance over our attacker !

Cryptographically Generated Addresses (CGAs)



Cryptographically Generated Addresses (CGAs): the cost

- Generation (owner): $\mathcal{O}(2^{16*sec})$
- Breaking the address (attacker): $\mathcal{O}(2^{59+16*sec})$
- Verification (receiver): $\mathcal{O}(1)$
  → Two hash computations

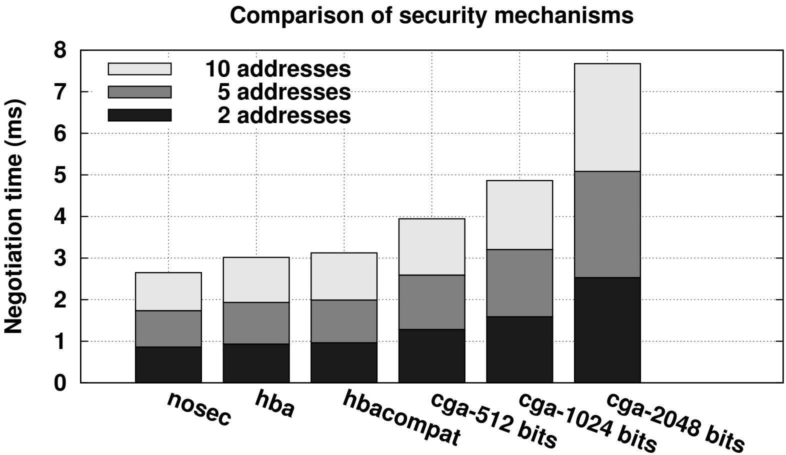
Hash Based Addresses

- Similar to CGA addresses, but lighter.
- Same input as for CGAs
- Public key is a random number
- Extension field is the list of prefixes.
- No signature needed, addresses validated by the fact that they are all bound together.

HBA vs CGA

- CGA is computationnally more expensive (key generation, signature).
- But HBA does not allow adding addresses later
  - All prefixes are included in the hash
  - Adding one prefix results in changing **all** addresses
- Tradeoff: CGA-compatible HBAs
  - A public key is used for generation, but the multi-prefix extension is included
  - Initial address set is announced through HBA
  - Additional addresses can be generated and announced using CGA.

HBA vs CGA: efficiency



Want to play ?

- LinShim6 can be downloaded at `http://inl.info.ucl.ac.be/LinShim6`
- Currently the implementation that best supports the specification
- Allows using CGA/HBA/CGA-compat HBAs.
- Can be easily installed in Ubuntu thanks to .deb packages.
- No special configuration needed (except for special purposes).

Shim6 challenges

- Bootstrap problem: **both** ends need to support Shim6 in order to get any benefit
  - If you install Shim6 now, almost no peer will know about it...
  - But if it gets installed in standard distributions, the whole world would have it at once.
- Load balancing: The end-host is now responsible for part of it. How to give control back to the network ?
  - Use a central server that hints the end-hosts ? (IDIPS)
  - Allow routers to re-rewrite Shim6 packets to enforce network policy ?

Shim6 challenges

- Renumbering: All ongoing communications are broken in case of renumbering
  - We would probably need a separate identifier space to solve that.
- Transport level multipath: Extending Shim6 to make it a path manager for transport protocols ?
- Mobility: To be investigated

Questions ?