Versatile Anycasting with Mobile IPv6

Michal Szymaniak Guillaume Pierre Maarten van Steen

Vrije Universiteit Amsterdam

Anycasting

- Anycast – network addressing and routing scheme:
  - Groups of nodes (anycast groups) addressed with anycast addresses
  - Traffic to anycast address reaches the node closest to the sender
- Intuitively very useful:
  - Proximity-based routing applicable in many distributed systems
  - E.g.: European Web clients should access content at European mirrors
- Still, distributed systems seldom employ anycast in practice:
  - E.g.: DNS servers organized into anycast groups for load sharing
  - Problem: current anycast implementations not flexible enough
  - Result: anycasting achieved with complex combinations of techniques

Agenda

- Anycast Implementation Properties
- Limitations of Current Solutions
- Our Concept: Use Mobile IPv6
- Versatile Anycast
  - Overview
  - Properties
- Conclusions and Future Work

Anycast Implementation Properties

- Perfect anycast implementation:
  - Organizes nodes into anycast groups
  - Provides anycast groups with anycast addresses
  - Routes traffic according to metrics defined by anycast groups
    - Not just one fixed network distance metric
  - Tolerates sudden changes in anycast group composition
    - Each node might leave its anycast group at any moment
  - Supports connection-oriented communication (TCP)
    - Used by most popular Internet services
  - Incurs low communication overhead
  - Is easy to deploy in the Internet

Limitations of Current Solutions

- Reverse proxy (frontend, as in clusters)
  - Anycast address == frontend address
  - Frontend implements anycasting when forwarding traffic to anycast nodes
  - Full traffic control, TCP supported, etc.
  - But: high communication overhead over WAN

- Client-side software
  - Let's implement anycasting completely on the client side
  - Full traffic control and no communication overhead
  - But: client-side modifications not always possible (Web browsers)

Limitations of Current Solutions ctd.

- Routing-based anycasting (IP anycast, current standard)
  - Nodes in anycast group share the same IP address
  - Each node advertises this IP to the Internet routing infrastructure
  - Each Internet router maps the IP address to the nearby anycast node
  - But(1): traffic controlled by third-part routers
  - But(2): routing is packet oriented, risk of breaking TCP connections
- DNS redirection
  - Nodes in anycast group share the same DNS name
  - DNS name resolved to addresses of different nodes as necessary
  - Full control and TCP support
  - But: very slow updates of group membership

Is there really no way..

- ..to implement anycast such that all required properties are there?

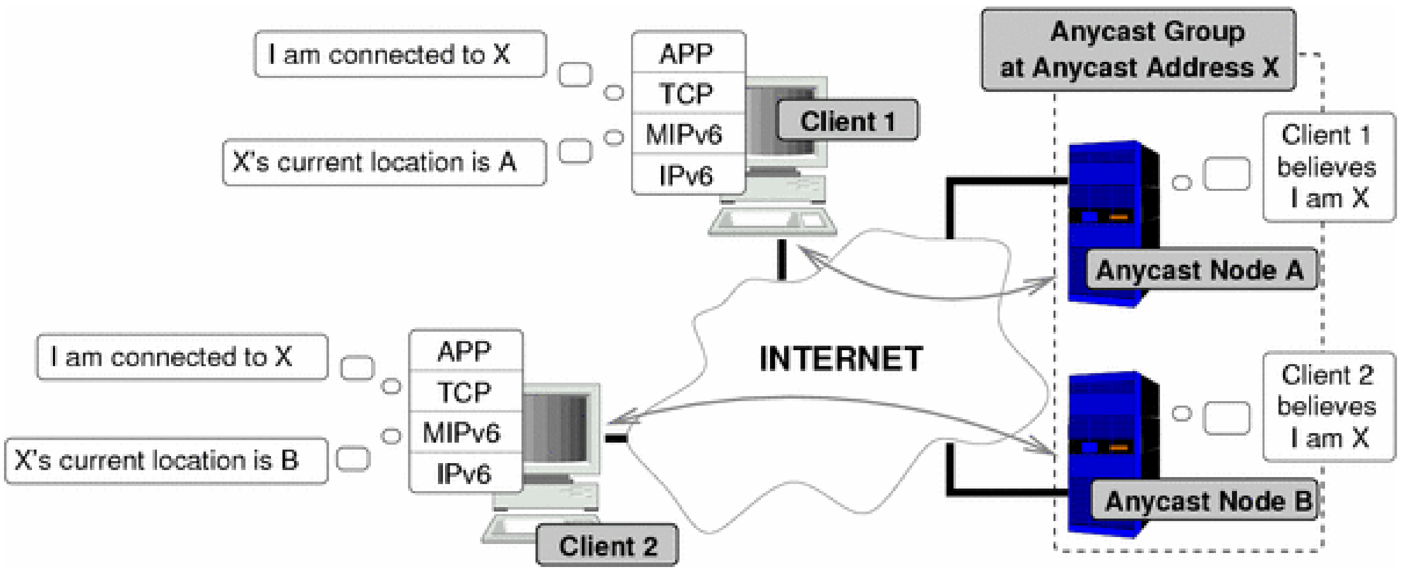
- We say there is..
- ..and we employ Mobile IPv6 for that.

Mobile IPv6 Overview

- Mobile nodes reachable while away from home networks
  - Any node talking to mobile node is called correspondent node (CN)
- Two addresses assigned to each mobile node:
  - Home address - identifies mobile node, never changes
  - Care-of address - represents mobile node's current location
- Mobile nodes inform correspondent nodes about care-of addresses
  - CNs translate addresses in traffic exchanged with mobile nodes
  - Address translation results in traffic switching
  - Transparency: TCP and applications only see home addresses
- Essential: switching is controlled remotely by mobile nodes

Versatile Anycasting: Overview

- Anycast GROUP pretends to be a SINGLE mobile node:
  - Clients == correspondent nodes
  - Anycast address == home address
  - Physical node addresses == care-of addresses
  - Anycast communication == clients switched between anycast nodes



Versatile Anycasting: Properties

- Full traffic control:
  - Each anycast group decides when to switch, and to which node
- Each node can leave at any moment:
  - It just needs to switch all its clients to another node within its anycast group
- TCP supported:
  - Our prototype implementation preserves connections upon switching
- Low overhead:
  - Switching time is very short (about 2 x client-to-anycast node RTT)
- Easy deployment:
  - Mobile IPv6 is a standard network protocol (to be) supported by major OSes

Conclusions and Future Work

- In short:
  - Anycasting could be attractive for various distributed systems
  - Current implementations are not flexible enough
  - Our idea is to exploit client-side traffic switching provided by Mobile IPv6

  - Versatile anycast presents each anycast group as a single mobile node
  - Anycast communication implemented with MIPv6 traffic switching
  - Switching controlled by protocols for announcing changes in location

- Future work:
  - Big: Exploit versatile anycast to develop wide-area distributed servers
  - Small: Client take-over upon ungraceful machine departures

Thank you!

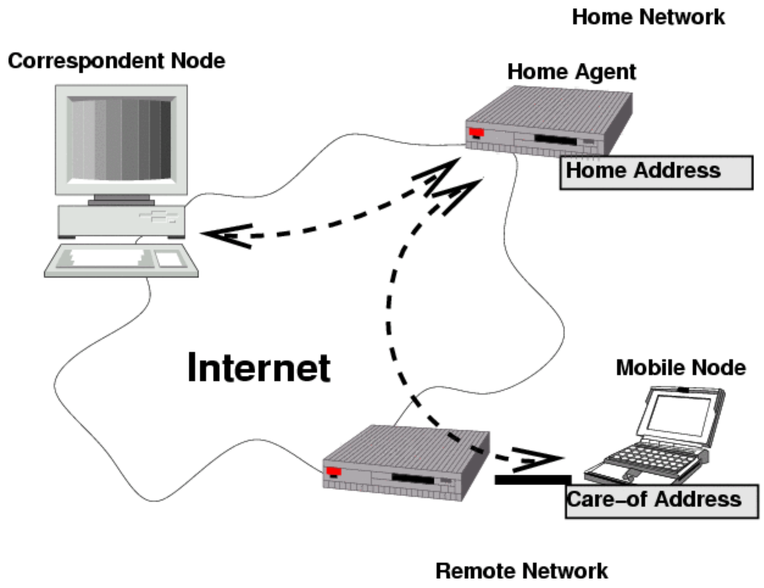
?

Mobile IPv6 Overview

- Mobile nodes reachable while away from home networks
  - Any node talking to mobile node is called correspondent node
- Routers in home networks represent mobile nodes
  - Any such router is called home agent
- Two addresses assigned to each mobile node:
  - Home address - identifies mobile node, never changes
  - Care-of address - represents mobile node’s current location
- Goal: mobile nodes always reachable at their home addresses

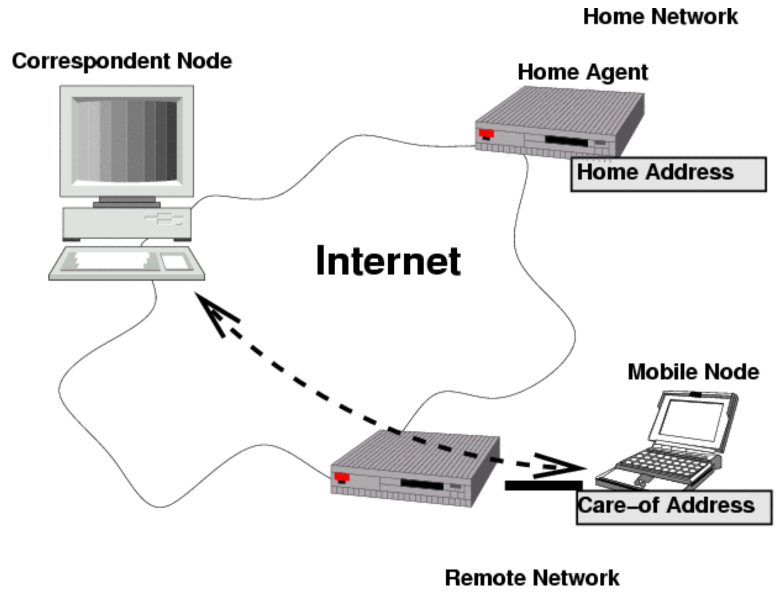
MIPv6: Tunneling

- When away, mobile node (MN) reports its current care-of address to its home agent (HA)
- HA tunnels traffic between MN's home address and MN's care-of address
- Transparent to correspondent nodes
- But:
  - Suboptimal routing
  - HA can become bottleneck



MIPv6: Route Optimization

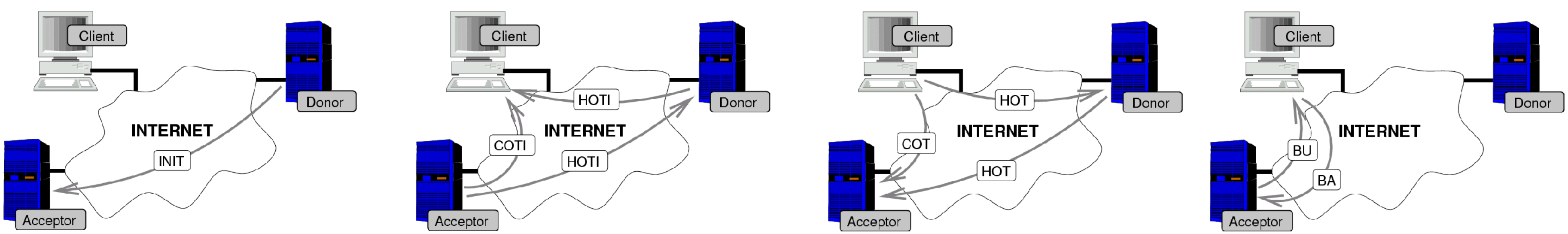
- MN reveals its care-of address to correspondent node (CN)
- CN creates a translation mapping
  - Home address <=> Care-of address
  - Address translation in CN's IP layer
  - Higher layers see home address only
- Result:
  - Direct MN-CN communication..
  - ..with MN movements transparent to applications running on CN



Wide-Area Handoff

- Anycast group must control client-side address translation
  - Translation mappings updated during route optimizations
  - Anycast group mimics the route optimization protocol (b, c, d)
  - Slang: donor handoffs client; acceptor takes over client

a) b) c) d)



a) - Please take over this client

b) - Hi Client, here is my new care-of address

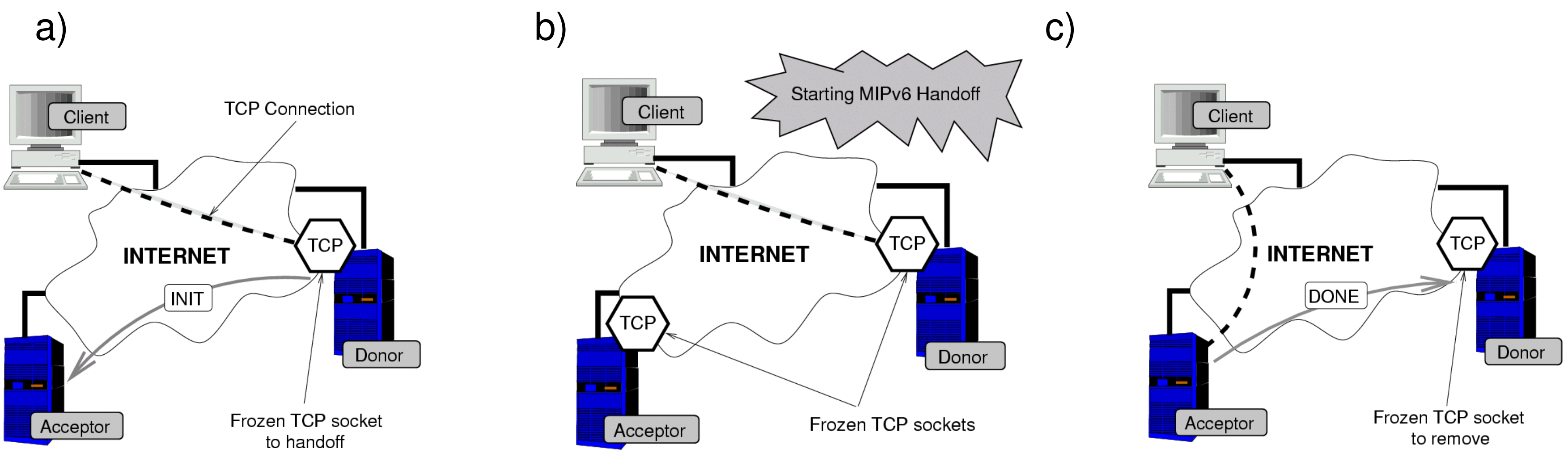
c) - Really? Let me verify it then

d) - Verification OK, please update your mappings

- Client now talks directly to acceptor on IP level, but..

Wide-Area Handoff ctd.

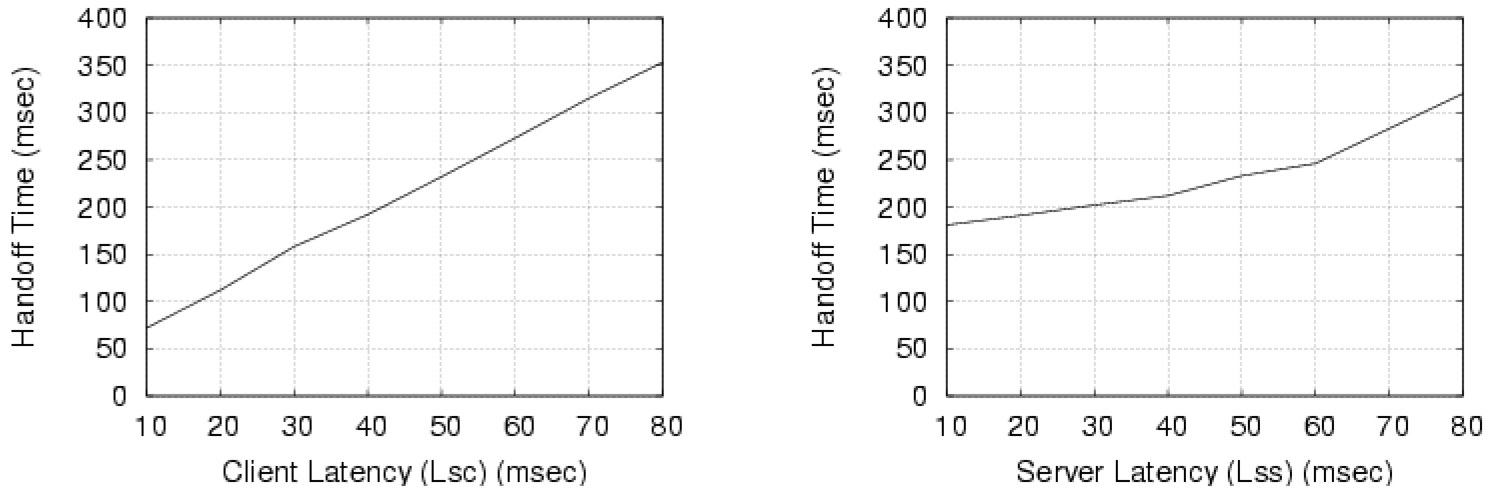
- ..we have just broken TCP connection :-(
- TCP connection state must be transferred to acceptor as well
  - Server-side TCP socket frozen to avoid changes in connection state
  - Two frozen socket instances to avoid accidental connection reset



- Client is now connected to acceptor while believing it is donor :-)

Client-observed Handoff Time

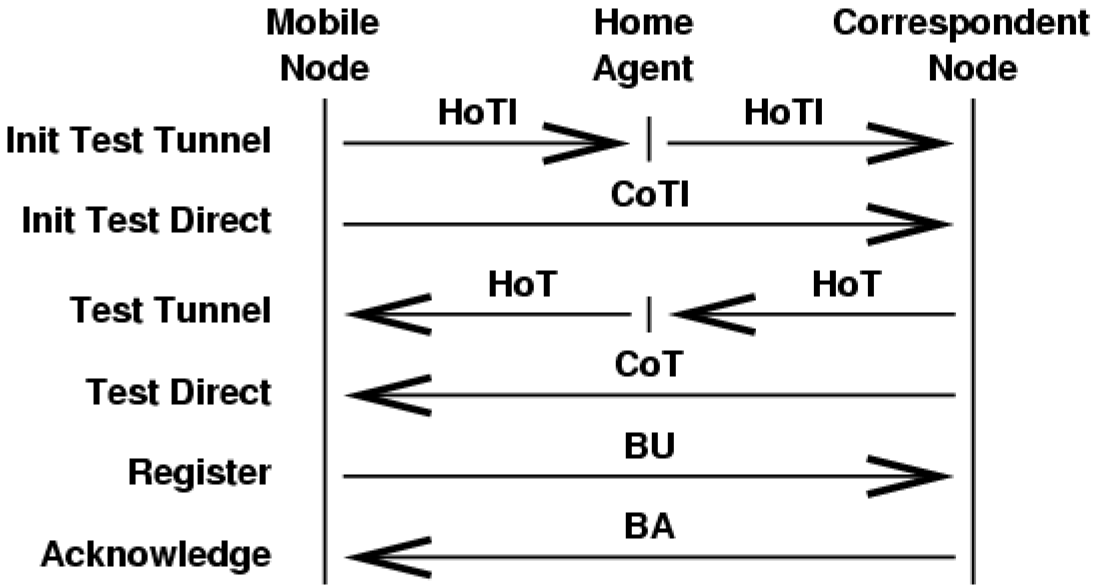
- Delay between receiving data from donor and from acceptor
- Lss – one-way latency between donor and acceptor
- Lsc – one-way latency between client and either donor or acceptor
- After all optimizations: handoff time = Lss + 4 * Lsc



- Some optimizations assume low Lss; worst case: 3 * Lss + 6 * Lsc

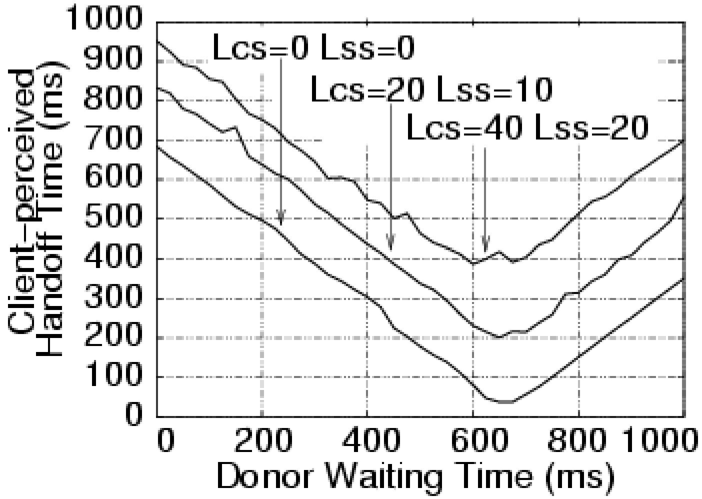
Route Optimization Protocol

- Tests prove that care-of address matches home address
- BU contains combined values of HoT and CoT
- Cryptography all over the place



State Transfer Optimization

- Server-side TCP socket might contain unsent/unacknowledged data
- Such data must be transferred to acceptor as well
- Better wait until socket buffers become empty:



Handoff Time Optimization

- Some messages (HoTI/CoTI and HoT/CoT) exchanged in advance
- Result: (3 * Lss + 6 * Lcs) reduced to (Lss + 4 * Lcs)
  - as long as messages are exchanged before actual handoff starts