Overlay techniques in the underlay

1st Conference Seminar
Massively Distributed Systems
Winter Term 2006/2007

Elmar Hoffmann <elho@elho.net>

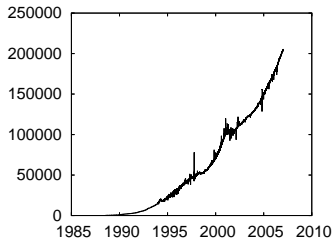RWTH Aachen University

February 16, 2007

Overview

Motivation

- Addresses being both identifiers and locators is a bad idea
  - what you want vs. where to find it
  - mobility (Laptops, ...)
  - re-homing
  - multi-homing
- Limited scalability of existing routing algorithms
  - Internet routing
    - depends on hierarchy to scale
    - >200 000 routes
    - (fast) router memory is limited
  - Mesh routing
    - flooding
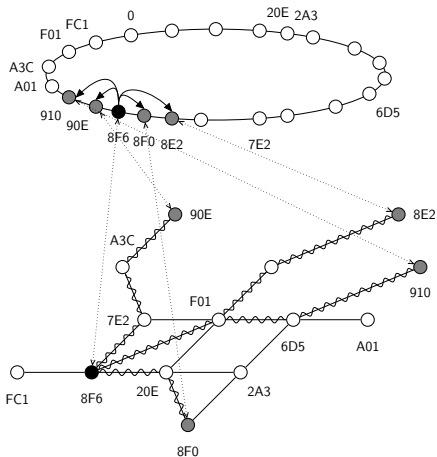   - location based addresses

Overview

VRR: Concept

- Node identifiers
  - ▸ random unsigned integers
  - ▸ location independent
- Nodes are arranged into virtual ring
- Nodes maintain
  - ▸ Virtual neighbour set (*vset*)
  - ▸ Physical neighbour set (*pset*)
    - ■ one-hop neighbours
    - ■ two-hop neighbours
      (as an optimisation)
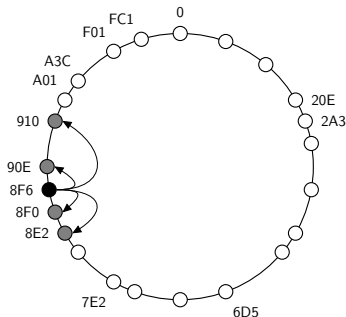
VRR: The big picture

VRR: Routing

- Provides DHT functionality
- Routing table
  - ▶ *vset-paths* to *vset* members
  - ▶ *vset-paths* that run across the node
  - ▶ physical neighbour paths to *pset* members
- Routing table entry
  - ▶ both endpoints
  - ▶ next hop towards both endpoints
  - ▶ next next hop towards originating endpoint (as an optimisation)
  - ▶ path id (doubling as preference)

Failure detection and repair

- *Symmetric failure detection*
  - ▸ *hello* messages
  - ▸ physical neighbour state
- *vset-path* repair
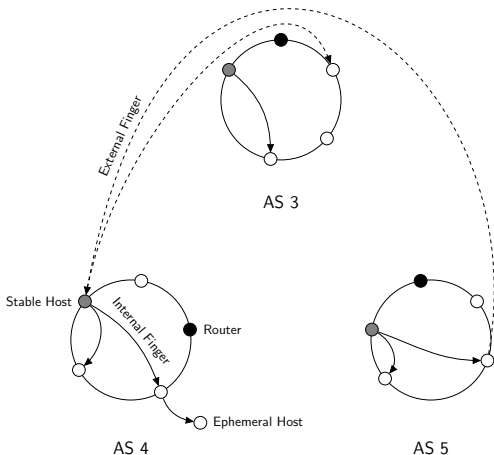 - ▸ teardown and setup
- *local repair*
  - ▸ constant cost

Overview

ROFL: Concept

- Node identifiers
  - location independent
  - self-certifying (public key cryptography)
- Node types
  - *routers*
  - *stable hosts*
  - *ephemeral hosts*

ROFL: The big picture

- Nodes are arranged into virtual rings
  - ▸ internal ring — intra-domain routing
  - ▸ external rings — inter-domain routing
- Nodes have *successor* pointer(s) into rings

External Finger

AS 3

Stable Host

Internal Finger

Router

Ephemeral Host
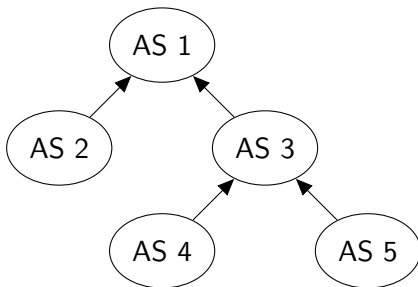
AS 4

AS 5

Distributed Systems

ROFL: Routing

- Provides DHT functionality
- *Routers* maintain *virtual node* on behalf of the hosts
- Routers additionally have *pointer cache*

ROFL: Inter-domain routing

- Node potentially joins ring at each level of its *up-hierarchy*
- *isolation property* assures routing through least-common ancestor.

ROFL: Inter-domain routing

- Node potentially joins ring at each level of its *up-hierarchy*
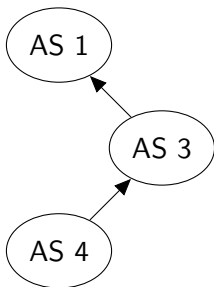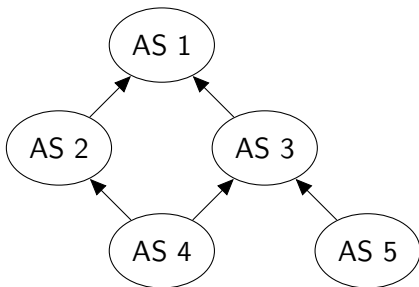- *isolation property* assures routing through least-common ancestor.

ROFL: Inter-domain routing

- Node potentially joins ring at each level of its *up-hierarchy*
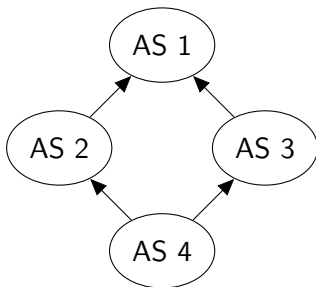- *isolation property* assures routing through least-common ancestor.

ROFL: Inter-domain routing

- Node potentially joins ring at each level of its *up-hierarchy*
- *isolation property* assures routing through least-common ancestor.

Overview

Summary

- Virtual Ring Routing
  - efficient mesh routing protocol
    - optimisations to take shortcuts reduce stretch
    - *symmetric failure detection* and *local repair* make it robust
    - low per-packet overhead and lack of flooding make it scalable
- Routing on Flat Labels
  - intra- and inter-domain routing protocol for the Internet
    - anycast, multicast and mobility
    - peering, multi-homing and routing policy
  - proof of concept
    - needs large ($>$20M entries) *pointer cache* to perform acceptably
    - demonstrates feasability of flat Internet routing

Further Reading

- M. Caesar, M. Castro, E. B. Nightingale, G. O'Shea, and A. Rowstron. Virtual ring routing: network routing inspired by DHTs.

  In *SIGCOMM '06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications*,

  pages 351–362, New York, NY, USA, 2006. ACM Press.

- M. Caesar, T. Condie, J. Kannan, K. Lakshminarayanan, and I. Stoica. ROFL: routing on flat labels.

  In *SIGCOMM '06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications*,

  pages 363–374, New York, NY, USA, 2006. ACM Press.

- Overlay techniques in the underlay

  `http://www.elho.net/pub/`

Distributed Systems

Questions?