Distributed Router Architectures

Markus Hidell, KTH IMIT/LCN
mahidell@imit.kth.se
Router Architecture Research at KTH

Winternet – Router & Switch Architectures

KTH IMIT/LCN
- Networking Laboratory
  - Wallenberg
  - Intel
  - Switchcore
  - Xelerated

Motivations
- Packet processing, new functionalities, technology trends, control plane/data plane
Packet Processing

- New services require per-packet processing
  - Transition to IPv6
    - Tunneling over IPv4, header translation, ...
  - Quality of Service
    - Classification, policing, shaping, monitoring
  - Virtual Private Networks (VPN)
  - Overlay networks
  - Mobility
  - ...

Router Data Path Dilemma

- Typically two data paths, *fast* and *slow* path
- Fast path is expensive
  - ASICs
  - Field upgrade is very costly
    - In conflict with packet processing trends
- Slow path is slow
  - Software executed by microprocessor
Technology Trends

Trends

- User traffic: 2x / 12 months
- Line capacity: 2x / 7 months (since late 90s)
- DRAM random access time: 1.1x / 18 months
- Router capacity: 2x / 18 months

(Source: Prof McKeown, Stanford)

Transmission > Processing > Buffering

Processing pushed to line cards (before aggregation of traffic)
Approach

- Modularised functionality with well-defined interfaces
  - Opposed to traditional monolithic architecture
  - Distributed architecture allows for physical separation
- IETF ForCES and GSMP, Multiservice Switching Forum, Network Processing Forum
- New business models
  - Switch and router "components"
  - Modularization in router industry!
- What mechanisms (protocols, interfaces) are needed in order to modularize a router?
LCN Networking Lab

KTHLAN

wilgot (main server)

blaster (compile server)

Production Network
172.31.212.0/24

Terminal Servers
(1 per rack)

Serial lines

Console Network

Smartbits Traffic Generator

Payload Network

1U rack mounted CPUs
(~10 per rack)

Control Network

Payload Network

60 1U rack mounted PCs
30 Network processors
20 Switches
Traffic generator/analyzer

Backplane Network

Serial lines
Lab Production Network

- Rack 1 holds lcn-gw, servers, etc...
- Server wilgot can be accessed from outside
- Rack 2-6 have general configurations
- Rack 7 holds network management platform (still self-contained)
- Rack Hector holds WLAN prototype platform (still self-contained)

![Diagram of Lab Production Network](image-url)
Rack Layout

Available rack space

- Netgear FS 524: 10/100 sw
- Netgear FSM 726: VLAN capable 10/100 sw
- Netgear GSM 712: VLAN capable GE sw

Serial line access

- console02
- r2n1
  - r2n1 \( rmn: \) rack \( n \) node \( m \)
- r2n2
- r2n12

Networks

- Production Network
- Control Network
- Backplane Network

- 10/100 Eth
- 10/100 Eth
- 10/100/1000 Eth
Monolithic Router Architecture

- config software
- routing software
- mgmt software
- user
- kernel
- netlink
- forwarder
- RIB
- FIB
- IF
- External links
Modularized Architecture

CE – Control Element
FE – Forwarding Element
Virtual Router Model

VFIB – Virtual Forwarding Information Base
VTM – Virtual Traffic Manager
VIF – virtual interface

IB – Forwarding Information Base
TM – Traffic Manager
IF - interface
Physical Separation CE/FE

Experimental set-up:
- 2 processes
  - zorg_ce (1 per CE)
  - zorg_fe (1 per FE)
- Forz protocol
  - CE-FE communication
  - Discovery of CEs/FEs
  - Encapsulates netlink messages
Forz Protocol – Current Status

Association phase

Association established

FE

JOIN REQ

JOIN REP

HEARTBEATS

GETLINK

NEWLINK

NEWROUTE

Report IF

Report IF

Set FIB entry

CE

Association phase

Association established

Create VIF

Create VIF
FEs Studied at LCN

- FE Reference Software Implementation (C)
- UNIX Software FE (BSD/Linux)
  - Routing Socket/Netlink/Patricia Tree FE
- NP: Intel IXP
  - Jing Fu (PhD Student)
- NP: Xelerated’s X10
  - Said Kashoob (Msc Thesis)
- Switchcore
- FPGA (Celoxica)
Conclusions

- New services kept back by current router architectures
  - Inflexibility in packet processing path
  - Management, signalling, forwarding, etc, are all integrated into a monolithical system
- Modularization as a means to improve functionality, performance and flexibility
- Experimental research on distributed routers in the Networking Laboratory at KTH/LCN