Architectural Design for Efficient Hardware Upgrade of High Performance IP Routers

Elizabeth Suet Tse, PhD
esta@ieee.org
Office: 732-427-2804
Cell: 732-778-7791
Overview

- Problem Definition
- Solution Approach
- Details of Solution
  - Method to Unify Router Hardware (HW) Architecture
  - Candidate Switch Fabric Selection
  - Special Processing Agent
  - Interface Module (~ Line Card)
- Solution Summary
- Beyond Upgrade
Problem Definition

- **Practical Problem**
  - Multifaceted: Internet Services Providers (ISPs) / Router Vendors or Designers / end users
  - Originated: Common Backbone Systems Engineering Org. → AT&T WorldNet® Services

- **AT&T WorldNet® Services**: AT&T’s high speed IP network → contains 1000s of IP routers

- **ISP Network Background Synopsis**
  - 3 basic types of IP routers, in hierarchical layers → edge / hub / backbone (bb)
  - Large ISP networks: >= 1 hub / bb layers
  - Very different functionalities / speeds
  - Usually can’t quickly/easily interchange routers amongst layers, e.g., can’t quickly take an “edge” and make it to become a “backbone”, vice versa

⇒ Find a way to address this issue as part of addressing the multifaceted problem
Problem Definition (cont.)

- Multifaceted Problem

- bandwidth-intensive QoS / CoS applications in IP network

- **ISPs**: frequent need router **upgrade** to accommodate new end-user needs – speed range / capacity / features → many network layers: $$$$$$ / time-consuming

- **Router vendors or designers**: challenge of designing **flexible router HW architectures** to adapt to frequently changing requirements
  e.g., hard to design router to accommodate wide range of connection technologies / speeds, especially “on-demand”. Many design constraints
  - *mechanical* constraint – can’t ↑ # of I/O card slots “on-demand” without also expanding capacity of router switch fabrics
  - *environmental* constraint – more mechanical components (e.g., I/O card slots), harder to cool router
Problem Definition (cont.)

- Multifaceted Problem (cont.)
  - *End users*: problem of facing *costly* and uncertain network *downtime*

- Objective
  - Common Solution for efficient router HW upgrade:
    1. Flexible HW
    2. Fast upgrade
    3. Avoid frequent “fork lift” upgrades
    4. network layer independence upgrades
Solution Approach

Solution: Router architectures with 3 capabilities for efficient HW upgrade
- HW *scalability* to incrementally expand / adapt
- Flexible *adaptation* to QoS / protocol features changes
- (Bidirectional HW) *reconfigurability* to perform different roles & functions seamlessly wherever it is physically placed: *edge, hub or backbone*

Approach:
- **basic router HW components / functionalities**
- **generic** packet processing tasks performed by different routers
  - KEY QUESTION: What are the generic tasks a router must do as an Edge? Hub? Backbone?
- **router HW architectural evolution & switch fabric (SF) designs**
  - design advantages / limitations
  - existing architectural principles for HW scalability
    - A pre-requisite for reconfigurability
    - Principles for HW scalability    ➔ Principles for HW reconfigurability

* Architectural attribute that allows a router to function “downward” as “edge”, or “upward” as “hub / backbone”.
Details of Solution

- **Contributions**
  - Principles for HW scalability / bi-directional reconfigurability
  - 6 basic router / SF functional requirements & Primary SF selection criterion
  - Methodology for HW Architectural Unification
  - Set of router HW architectures capable of: **scalability, adaptation, reconfigurability**
  - a scalable reconfigurable IP router

- **Detail – Methodology for HW Architectural Unification**
  - KEY observation and principle for reconfigurability
    - Unified router HW architecture across all layers of the network hierarchy
    - easy conversion amongst ALL layers
  - TRICK: How to accomplish functional change in a router HW architecture to perform as edge, hub or backbone *on-demand*, with *off-the-shelf technologies*?
  - Must 1st answer KEY question: what does a router do with an incoming packet that makes it an edge, hub or backbone?
    - Router Functionality Comparison
    - Examination of the packet processing functions
## Router Functionality Comparison

<table>
<thead>
<tr>
<th>Plane of Action</th>
<th>Edge Router</th>
<th>Hub Router</th>
<th>Backbone Router</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Plane</strong></td>
<td>Flexible adaptation to different connection technologies and speeds</td>
<td>Port expansion/traffic concentration</td>
<td>Packet transfer at high speed</td>
</tr>
<tr>
<td></td>
<td>Packet forwarding e.g., filter/differentiation/queue/ schedule packets based on header “tags”</td>
<td>Packet forwarding (possibly based on packet header “tags”)</td>
<td>Packet forwarding (possibly based on packet header “tags”)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reserves bandwidth when asked to do so by a network server (“passively” providing QoS/CoS for packets)</td>
<td>Reserves bandwidth when asked to do so by a network server (“passively” providing QoS/CoS for packets)</td>
</tr>
</tbody>
</table>
Packet Processing Functions

Open Systems Interconnection (OSI) layer 2 Functions

OSI Layer 3 Functions Begin

Packet Header Filtering (1)

Packet Header Classification
Control: Management/administration of network resource and policy (re-) configuration (2)

Packet Differentiation (3)

Packet Forwarding (Table Lookup) (4)

Packet Queuing (5)

Packet Scheduling (6)

Packet Switching (7)

Packet Queuing

OSI Layer 3 Functions Ends
1. **Identify** Distinguishing Functions (previous 2 slides)

2. **Group** Together these identified functions into a *separate* Functional Unit
   - SPA: Special Processing Agent, for processing packets with special needs
   - *Leave* the set of packet forwarding/data plane functions – as a generic part of the (core) router architecture

3. **Use** identical General-Purpose Interface Module (~ Line Card) **Slots** in all routers

- Overall view: SPA & router
- Every router has connection to SPA
- OPEN/CLOSE connection approach enables speedy HW reconfigurability “on-demand” according to router’s physical placement
- Architecture allows new QoS / CoS to be added without necessarily requiring any HW/SW changes at core router
Basic Conceptual Router Architecture with SPA Inside Router

LEGEND:
——— = Logical component interconnection
Main Router HW Components for HW Scalability and Bi-directional Reconfigurability

- Switch Fabric
- SPA
- Interface Module (IM)

How each contributes to our goals of a scalable reconfigurable router with the 3 capabilities for efficient HW upgrade:

- HW *scalability*
- Flexible *adaptation*
- (Bidirectional HW) *reconfigurability*
Switch Fabric (SF) Selection: Functional Requirements & Selection Criterion

1. Efficient uplink and route processor access
2. Efficient downlink access
3. Preservation of small packet delivery delay variation
4. Efficient tree-based algorithm embedding capability
5. Fault-tolerance
6. Ease of incremental (HW) expandability and contractibility (*) – SF’s HW flexibility to scale according to traffic volume demands

(*) Primary Switch Fabric selection criterion

$$SF \text{ Compliance Score} = \sum_{\text{reqt.1}}^{\text{reqt.}N} \text{Impor\ tance}_{\text{reqt.}} \times Conformance_{SF}$$
Ten (10) Surveyed Switch Fabrics and associated scores

Scatter Plot of Overall Compliance Scores for the Ten Switch Fabrics

Switch Fabric Type

Compliance Score
Three (3) Candidate Switch Fabrics for a Scalable Bi-directionally Reconfigurable IP router

3-D Torus Mesh

- = the Switch / Processing Elements within a router/computer

Commercial Examples:
- CRAY T3D/T3E Supercomputer
- Avici Terabit Switch Router

MIN- Multistage Interconnection Network
(e.g., $2^3 \times 2^3$ Delta)

- = Switch / Processing elements within router

Commercial Example:
- Pluris Teraplex 20® router

Commercial Examples:
- Alcatel’s 7770
- Caspian’s Apeiro
- Hyperchip’s PBR (PetaBit Router)
SPA for Efficient IP Router Hardware Upgrade

1. Responsible for making UPDATED PACKET HEADERS w/ new routing info/“tags” for headers of packets with special needs

2. Communication: betw. router’s IM & SPA – request/response
   - **Request**: header of a packet w/ classification or priority reqts., e.g., in the TOS Byte or the 5-tuples
   - **Response**: contains updated packet header with new tag
     + tags: header labels → assign packet processing/forwarding priority
     + assigned by matching header instructions (e.g., TOS Byte/5-tuples) per SPA’s internal classification rules / policies
     + for multi- /broad- cast services: SPA produces enough # of updated/tagged headers needed for the multi-/broad-cast session

3. Also communicates with router’s processor to get updated topology information needed for use by the SPA’s internal rules & policies

4. All internal fxnl units: reprogrammable HW, pipelined, for fast update

**Effect of architectural arrangement**

1. SPA ~ sophisticated packet header classifier, provisions/ administers / processes headers w/ special reqts. router’s data plane functionality to bare minimum (e.g., processes packet per updated header “tags”)

2. Unifies router’s internal HW architecture by having only generic data plane functions for ALL routers
An Example IM

that compliments the SPA

Highlights

- FE: all functional units are dynamically reprogrammable
  - HW-based
    - pipeline format & self-sufficient
- General purpose IM slots: media – / speed – neutral
- HW Reconfigurability via open/close connection to SPA
- Multiple paths for packet processing operations

SPA Path

(to/from SPA)

Layer 3 - Forwarding Engine (FE)
  - The data plane
    - (“fast” path – through data plane / FE to the destination IM via SF)
    - (“slow” path – to the CPU via SF)

Layer 2 Function

(Interface Module - IM)

(General Purpose Interface Module slot)
Solution Summary

- Switch Fabric can incrementally expand / contract \(\rightarrow HW\ Scalability\ (and\ also\ reconfigurability)\)
- Use of reprogrammable HW-based Functional Units \(\rightarrow\) Flexible adaptation
- Use of SPA in all router architectures \(\rightarrow\) bi-directional HW reconfigurability

MAY SIGNIFICANTLY

- **extend** interval between router HW upgrades for ISPs, also allows efficient upgrade (or n/w emergency response)
- **reduce** complexity of designing flexible HW architectures to adapt to changing functionality/speed reqts for router vendors/designers
- **minimize** possibility of network upgrade downtime uncertainties for end users

- **Such Architectures**
  - Can serve as the basis for developing next generation IP routers
  - Directly applicable to the emerging concept of a single-layer IP network architecture*

* Single-layer architectures have the advantage of ↓ the # of network layers, leading to less complex networks with fewer overall connections and fewer devices to manage.
LOOKING AHEAD → Beyond Upgrade: Recent Areas of NSF Research this Solution can support

- Embedded systems
- Sensor networks
- Surveillance/Monitoring /Analysis
- Applicable for Military
- Multicore systems for high end computing: very large scale apps (climate science, weather modeling/forecasting)

All needs speedy, high fidelity, high validity transmission of data, in real time as needed
This architectural design fits in perfectly
THANK YOU

Elizabeth Suet Tse, PhD
esta@ieee.org
Office: 732-427-2804
Cell: 732-778-7791