Service-oriented Backbone Router for Information Open Innovation Platform

Keio University Hiroaki Nishi
Co-project with NII, Tsukuba Univ., UEC
HITACHI JTE, NTT
Agenda

- **Purpose:**
  - Why do we need a SoR (Service-oriented router)?

- **Achievement:**
  - Service-oriented switch/router software/hardware emulator
  - Hardware architecture of service-oriented router
    - REGEXP co-processor
    - DBINS co-processor
    - PPDP co-processor

- **Future works:**
  - Where do we have to go next?
Cats and dogs of network

**Service·Application layer**
- Services including search engine are provided by end-hosts, and user can access these services from end-hosts, too.
- Rich services are making continuous progress and big benefit.

**Infrastructure layer**
- Routers or switches always devote themselves into packet forwarding.
- They only provide a few limited services; protocol translation, firewall, QoS, and load balancing.
- Evaluation of them are only done by using throughput or traffic control.

Some serious problems are caused by this gap
- Service providers earn money by exchanging contents over a flat-rate Internet access services.
- Carriers seek reasonable expense of their burdens. = Network neutrality problem
Router/switch should provide not only simple forwarding performance but also a high-level functions by using its superiority in location.

Service-oriented router can provide several services when installed at:
- Core area: New rating services for a search engine can be provided
- Metro area: Access control or security service can be provided
- Edge area: Meticulous service (user behavior analysis) can be provided
Related works

- **CISCO Systems**
  - AXP (Application eXtension Platform), ASR (Aggregation Service Router)
    - Provides PC extension module
    - Permits to design application by using special APIs.
    However, acceptable events is limited; fault detection, etc.
    **Now CISCO seeks what kinds of application can they provide.**

- **OpenFlow**
  - Provides flexible routing
  - Still have restrictions in providing application layer services

- **XML router proposed by Moscola, Cho, Lockwood**
  - Enables Content base routing
  - Inspects and forwards XML messages in several Gbps throughput
  However, it only controls packet routing and is not a truly service-friendly router.

- Why don’t you use the advantage the routers and switches have?
Basic operation of SoR

- Deep inspection of traffic with contents of packets
- SoR inserts REGEX-filtered contents into DB and provide API to access the DB by using SSRQL (Service-oriented Switch and Router Query Language)
- Realtime services independent of realtime communication
  - Freshness of services is the essential for users, and realtime communication is not essential of user requirement.
  - High-throughput and low-latency network is not a fundamentals to enrich services.
  - SoR enhances realtime services

```xml
<book>
  <title>Westlab</title>
  <author>Member</author>
</book>

http://www.westbook.com

<book>
  <title>Westlab</title>
  <author>Member</author>
</book>

http://www.westbook.com

DB

book

author

Member

http://www.westbook.com

DB

book

title

Westlab

http://www.westbook.com
```
Our study targets

- Design Service-oriented Router/Switch Query Language based on SQL
  - Improvement in description and expression ability
  - Can handle packet, stream and transaction separately
- Provide applications
  - Spatial information
    - Optimization of physical path of P2P
    - To provides flexible contents delivery services (CDS): Shift to contents based CDS
  - Temporal information
    - To provide seeds of ranking for search engines according to visiting duration
- Service-oriented database migration
  - To optimize the allocation of data resources by distinguishing attracted information
  - New recommendation services by sharing information between heterogeneous services
- Service-oriented security
  - Contents based network intrusion detection system/service (NIDS)
  - Routing and contents based anti-phishing attack function
- Privacy preserving data publishing
Current state of research

(1) SoR software simulator is now available
   - Stream reconstruction function was proposed to reduce the size of packet buffer
   - The simulator will be implemented on Juniper router and installed on SINET

(2) Hardware-based data insertion architecture was proposed
   - For providing lossless data insertion and highly availability

(3) Privacy preserving hardware is now available
   - FPGA based privacy preserving data publishing function is proposed, and special cache hardware is also proposed to eliminate information loss under the constraint of memory size.

(4) Real SoR will be available on next year end
   - Developing under the support of NICT

(5) Hardware based selection engine is under development

(6) Multiple SoR data sharing mechanism is under development
Hardware-based data insertion

- 3 insertion methods were implemented and evaluated.
  - Simple linked list of insertion time (= network time)
  - Two-dimensional list of both insertion time and user ID for faster selection
    - Above two method can be implemented only on chip SRAM.
  - Memory effective insertion with hierarchical memory management
    - Though off-chip DRAM is required, complex selection can be handled.
  - Embedded processor with special microcodes provides flexible memory management.

<table>
<thead>
<tr>
<th>Packet size</th>
<th>Network Throughput (Gbps)</th>
<th>Microcode Size (KB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On-chip Memory</td>
<td>Off-chip Memory</td>
</tr>
<tr>
<td></td>
<td>50B (min.)</td>
<td>1306B (HTML ave.)</td>
</tr>
<tr>
<td>Simple linked list</td>
<td>52.9Gbps (160MINPS)</td>
<td>67.4Gbps (160MINPS)</td>
</tr>
<tr>
<td>Two-dimensional list of both time and user ID</td>
<td>37.6Gbps (94MINPS)</td>
<td>67.4Gbps (94MINPS)</td>
</tr>
<tr>
<td>Hierarchical Index</td>
<td>0.66Gbps (1.65MINPS)</td>
<td>17.3Gbps (1.65MINPS)</td>
</tr>
</tbody>
</table>

MINPS = Million Insertions Per Second
Hardware Emulator of SoR

- FPGA-base hardware emulator was designed.
  - 3 FPGAs and 2 full-duplex 1G Ethernet ports
  - 4Gbps wire-rate processing (ideal)
  - Fully hardware-based L2 to L7 packet parser
  - DBINS co-processor, μREGEX, PPDP, etc. are evaluated
Privacy Preserving Data Publishing

- Hardware-based PPDP mechanism for wire-rate processing
- Both $k$-anonymity and $l$-diversity are supported.
- This mechanism uses TCAM for parallel data query (about 3,000 times faster than software-based PPDP) but hardware cost of TCAM is very high.
- To reduce the hardware cost, cache-based architecture are proposed.

50 times faster than SRAM based architecture
Thank you

Our web site of this project:
http://www.openinter.net
FAQ

- Should we replace all routers in the Internet into SoR?
  - Statistically, limited information can be regarded as mass behavior. Many rich services can be provided by using limited information.
  - Many application can be provided if SoR is installed the gateway of business area network or access network.

- How about the technological feasibility? The most important requirement of router is throughput!
  - We challenge to overcome the required throughput.
    - Software simulate can denote required resources.
    - FPGA based hardware emulator was established to estimate hardware cost.
    - Low-memory-cost packet stream reconstruction architecture was established.
    - There are many related works and they can be a basement of SoR.

- How about Privacy?
  - Basically, opt-in/opt-out should be taken. We also providing wire-rate PPDP architecture to provide rich contents under preserving privacy.