Fairness Issues in Software Virtual Routers

Norbert Egi

Computing Dept., Lancaster University, UK: M. Hoerdt, L. Mathy

Dept. of Computer Science, UCL, UK: A. Greenhalgh, F. Huici, M. Handley
Introduction

• **Goal:**
  – Build *high-performance and flexible virtual router* on *x86 commodity hardware*

• **Virtual Routers:**
  – *What is a VR?*
    • Concurrent router instances using the same physical resources
  – *Why do we want to build VRs?*
    • Research world
    • New business models
    • Enabler for Internet innovation
    (We are not limited to IP as the L3 protocol)
Introduction

• **Motivation:**
  – Powerful x86 architectures
    • Multi-core, Multi-CPU
    • Larger-and-larger CPU caches
    • High-speed I/O (PCI-Express)

• **Issues:**
  – Forwarding power
  – I/O virtualization

• **Software:**
  – Click in Linux kernel
  – Xen VMM
Virtual Routers

Can conventional techniques for server virtualization be applied to network router virtualization?
Forwarding in the virtual machines/routers

![Graph showing Forwarding in the virtual machines/routers](image)

- **Below Saturation**
- **At Saturation**
- **Native Performance**
- **Aggregated Virtual Router Performance**

---

*Fairness Issues in Software Virtual Routers, PRESTO'08*
Forwarding in the virtual machines/routers

Can conventional techniques for server virtualization be applied to network router virtualization? **Not really.**

**Forwarding in the VMs**

**Forwarding in the Driver Domain only**
Virtual Routers

What to do then?

PERFORMANCE

ISOLATION  FLEXIBILITY
Virtual Routers

VIRTUALIZED FORWARDING PLANE
(FORWARDER DOMAIN)

CONTROL PLANES

PERFORMANCE

ISOLATION

FLEXIBILITY
Performance issues

- **Bottleneck analysis:**
  - 7.1 Mpps min. sized packets (3.6Gbps)
  - With larger packets (>190 bytes) the lines get saturated (12Gbps)
  - On multi-core systems the bottleneck is the main memory latency
  - Caching matters, avoid packets switching cores
  - Approach similar to S. Ratnasamy et. al.

- In the followings we use only min. sized packets
Packet classification

- *DeMux packets to their VRs:*
  - Software classification
    - Overhead (Have to avoid “unauthorized” packets getting into the memory)
    - Insufficient fairness
    - No hardware support is needed
  - Hardware classification
    - No overhead
    - Fairness guaranteed
    - Multi-queue support is needed
    - Prerequisite for virtual routers
Core allocation + Scheduling

1.) Allocating forwarding paths to cores
- *Global load-balancing* (based on cost and resource entitlement of each forwarding path)

2.) FP scheduling (per core)
- *Fair resource usage*
- Include memory accesses into the scheduler (Future work)
Scenario 1: Static virtual forwarding plane
Scenario 1: Static virtual forwarding plane

3 identical forwarding paths are sharing a single CPU core

Flows have different arrival rates

Actual service rate is the same
Scenario 1: Static virtual forwarding plane

3 cores hosting 6 FPs that are given weighted priorities
Scenario 2: Configurable virtual forwarding plane
Scenario 2: Configurable virtual forwarding plane

3 forwarding paths (2 identical, 1 more expensive) are sharing a single CPU core (Default Click Scheduler)

Flows got a fair share, but they shouldn’t!!!
Scenario 2: Configurable virtual forwarding plane

3 forwarding paths (2 identical, 1 more expensive) are sharing a single CPU core (Extended Click Scheduler)

Expensive flow gets lower throughput
Scenario 3: Costumisable virtual forwarding plane
Scenario 3: Costumisable virtual forwarding plane

Baseline: 3 FPs in dom0

3 FPs in dom0 and 3 domUs forwarding via Xen I/O-channel, each FP and domU posseses a separate core

domU forwarding
Scenario 3: Costumisable virtual forwarding plane

3 FPs in dom0 and 3 domUs forwarding via Xen I/O-channel, the first FP in dom0 and the first domU are sharing a core, the other domains using separate ones.

3 FPs in dom0 and 3 domUs forwarding via Xen I/O-channel, each FP and domU posseses a separate core.
Conclusion

• Modern commodity HW is a viable platform for well-performing software virtual routers

• Forwarding paths with *trusted* elements
  - Adequate isolation and fairness
  - Negligible performance loss

• Forwarding paths with *untrusted* elements
  - Co-exist with trusted forwarding paths
  - Some performance drop for the untrusted forwarding path

• HW assistance and proper mechanisms are needed to overcome the novel system issues
  - Multi-queue NIC
  - Multi-queue support in the virtualization technique
  - NUMA
Q & A