



# (Some) Implementation Issues for Virtual Routers

Laurent Mathy (Lancaster U)

with

Norbert Egi and Mickael Hoerdts (Lancaster U)

Adam Greenhalgh and Mark Handley (UCL)

# Introduction

- **Why? What for? (Motivation)**
- **Why now? (Enabling technologies)**
- **HOW???** (Design)

# Motivation (1)

- **Conventional router architecture**
  - ✗ Closed router software market
  - ✗ Little potential for innovation
- **One box fulfils the role of multiple routers**
  - ✓ Independent and flexible management
  - ✓ Resource sharing
  - ✓ Lower hardware and support cost
    - E.g. Small businesses within the same building sharing the same router, each managing its own VR

## Motivation (2)

- **Multiple concurrent network architectures**
  - ✓ Innovation in network architecture design
  - ✓ Decoupling the providers from the physical network
- **Excellent platform for experimentation**
  - ✓ Rolling out new and unstable solutions without risk
- ✘ **Single point of failure (pairing is needed)**

# Enabling technologies (1)

- **Virtual machines on regular x86 PC hardware**
  - XEN, VMware, etc.
  - Both Intel and AMD are adding virtualization extensions to their processors
- **Inexpensive and powerful PC hardware**
  - Forwards several Gbps even with min. sized packets
  - Sufficient capacity for most small and medium size businesses
- **Future hardware trend**
  - Massively Multi-core architecture
  - Advanced multi-queue NICs

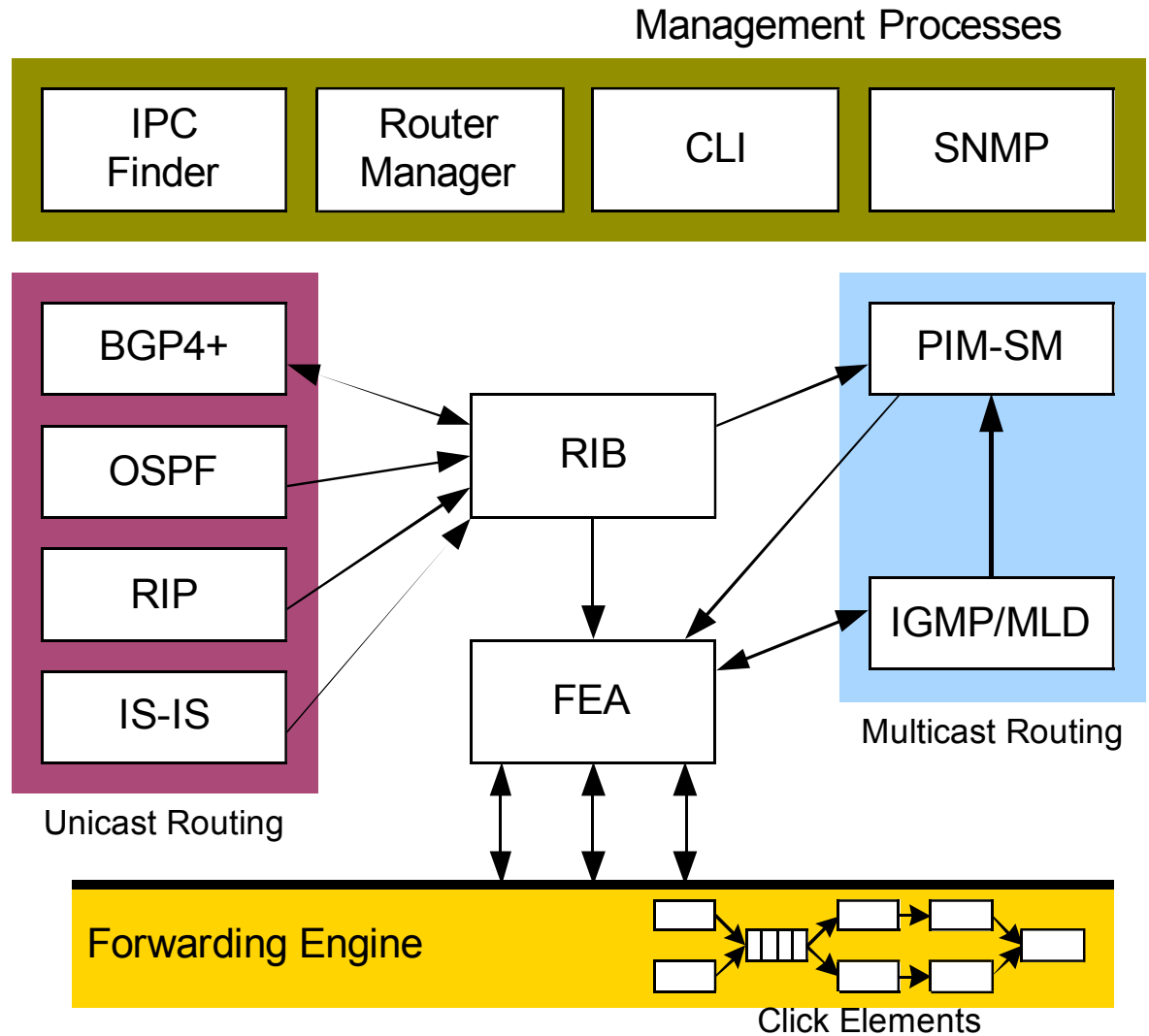
# Enabling technologies (2)

- **XEN**
  - Para-virtualization
  - VMs with performance close to native hardware
  - Excellent hardware support
  - Exploits Intel's and AMD's virtualization technologies
- **Click Modular Router**
  - Running in the kernel
  - Simple and small elements
  - Plenty of elements ready to use
  - Implement new elements is simple
  - High performance (outperforms even native Linux)

# Enabling technologies (3)

- **XORP**
  - Extensibility (XRL mechanism)
  - Flexible interface for the forwarding path (FEA)
  - Supports most of the widespread routing and management protocols
  
- **HEN (Heterogeneous Experimental Network)**
  - ~100 nodes
  - Wide range of network experiments
  - Remote power control

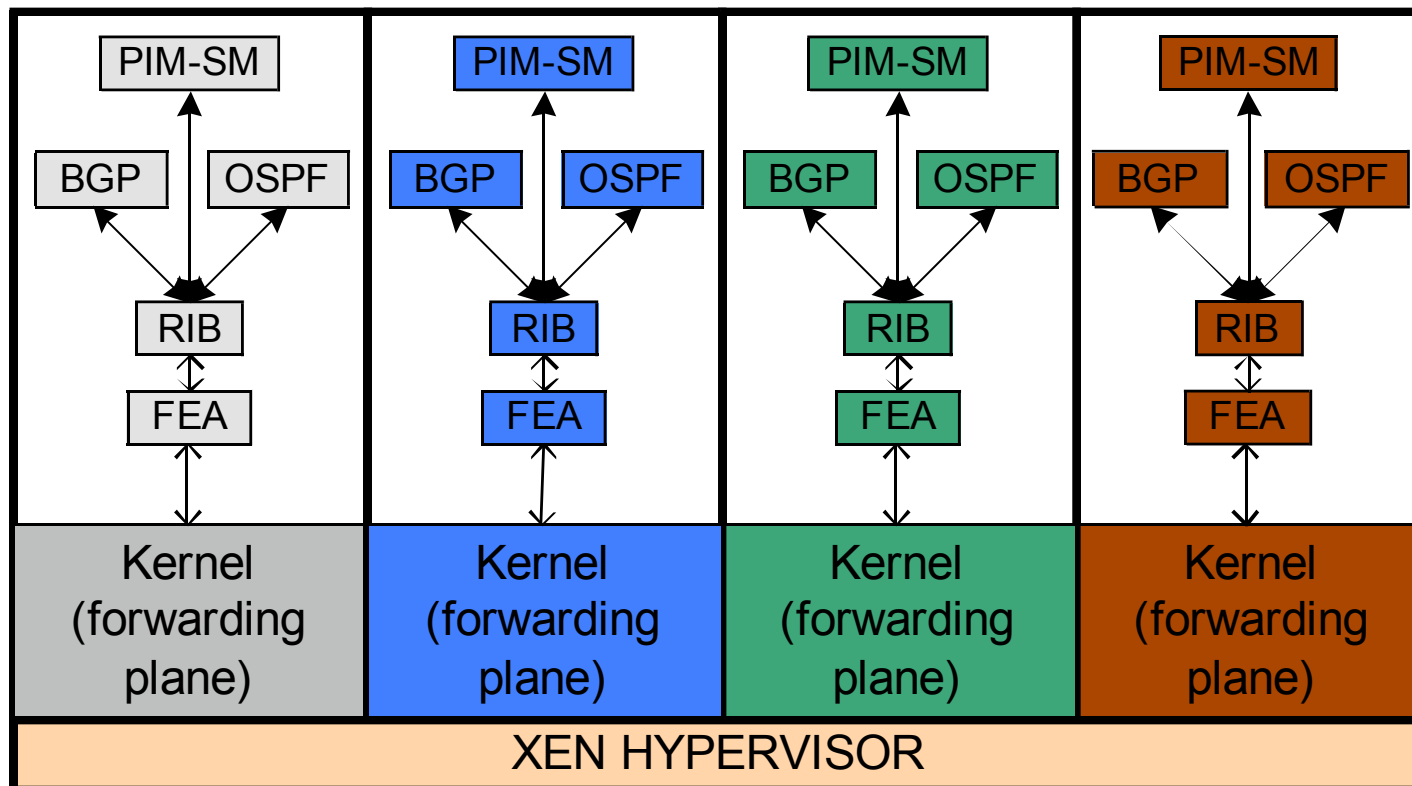
# SW router: XORP + Click





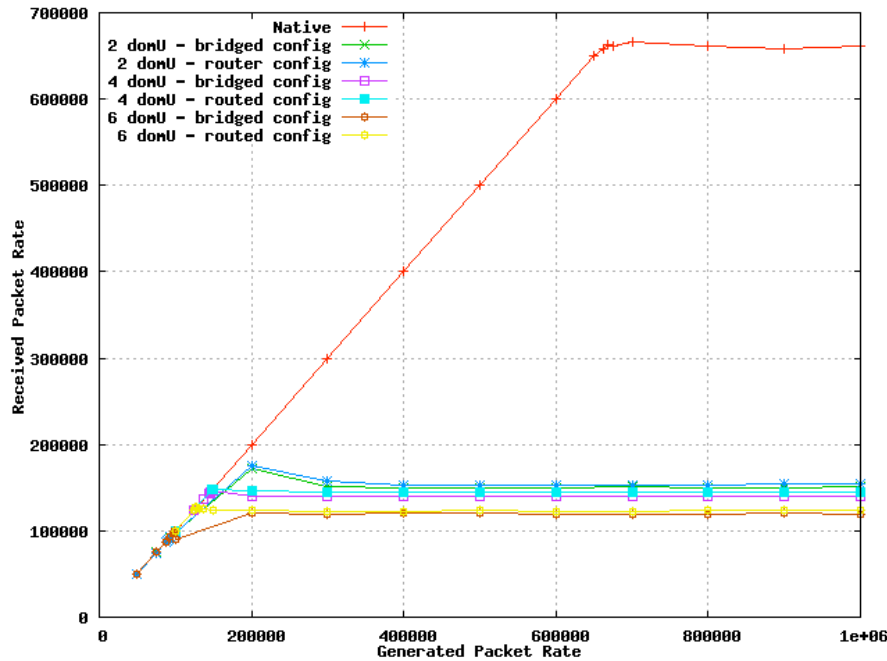
# Router virtualization (1)

**Can conventional techniques for server virtualization be applied to network router virtualization?**

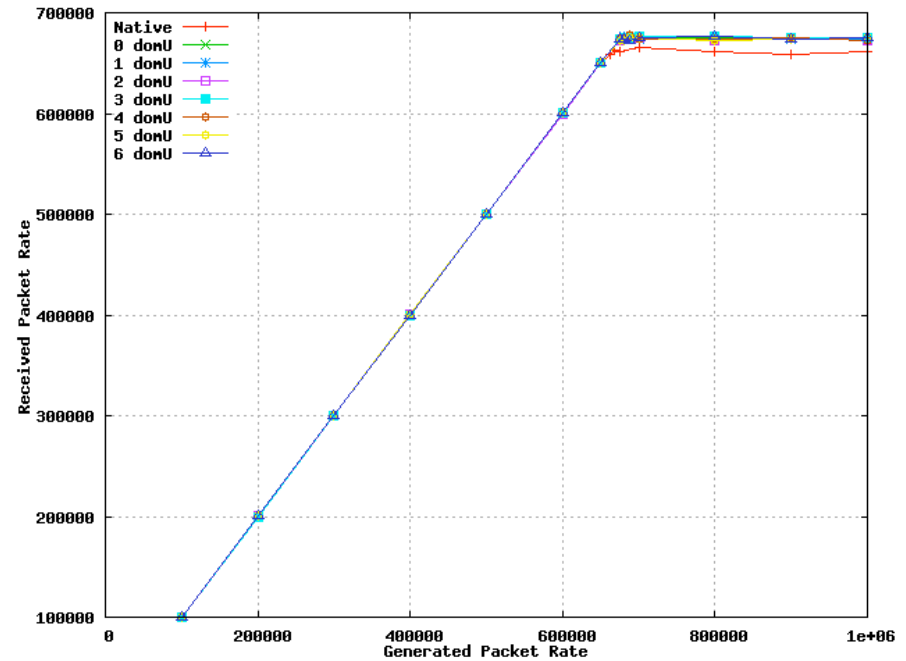


# Router virtualization (2)

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



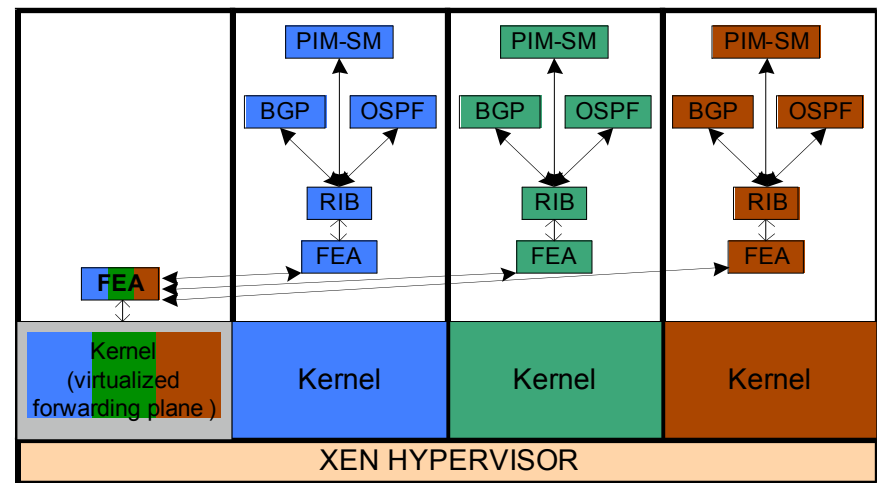
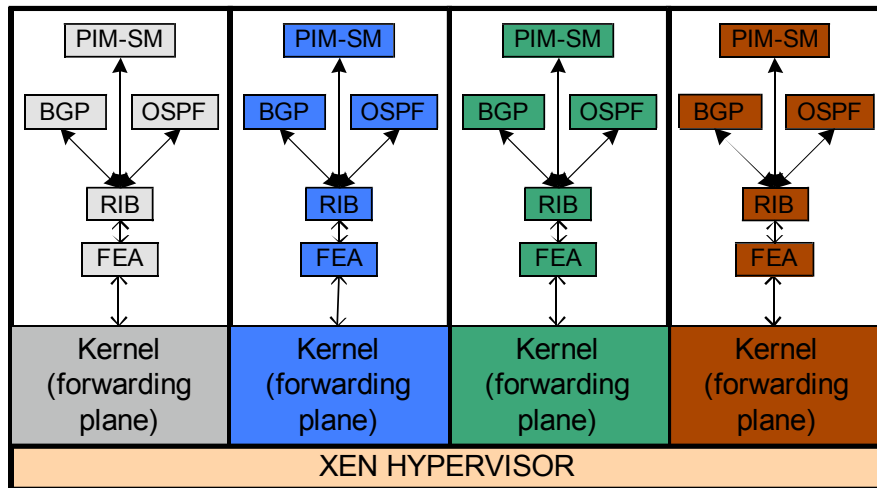
Forwarding in the DomUs



Forwarding in Dom0 only

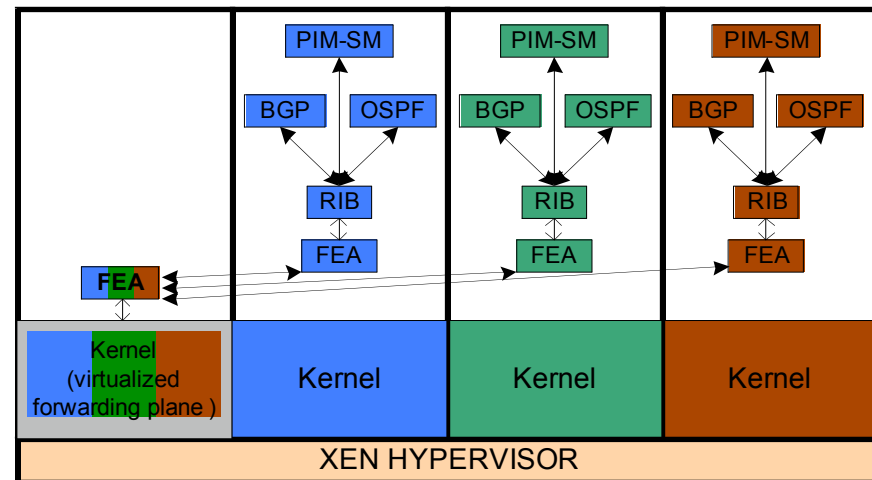
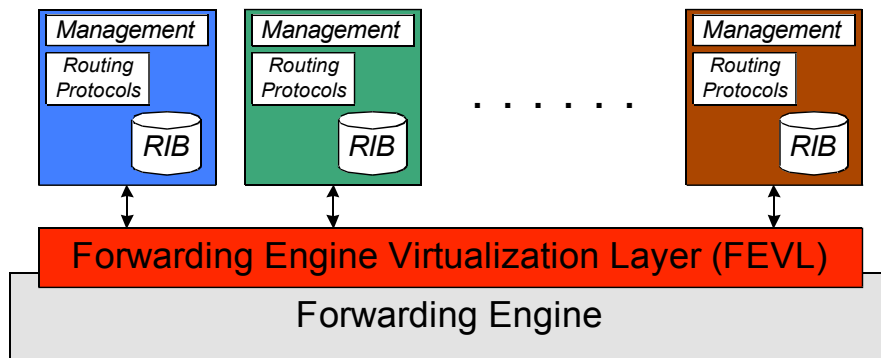
# Virtual Routers

## What to do then?



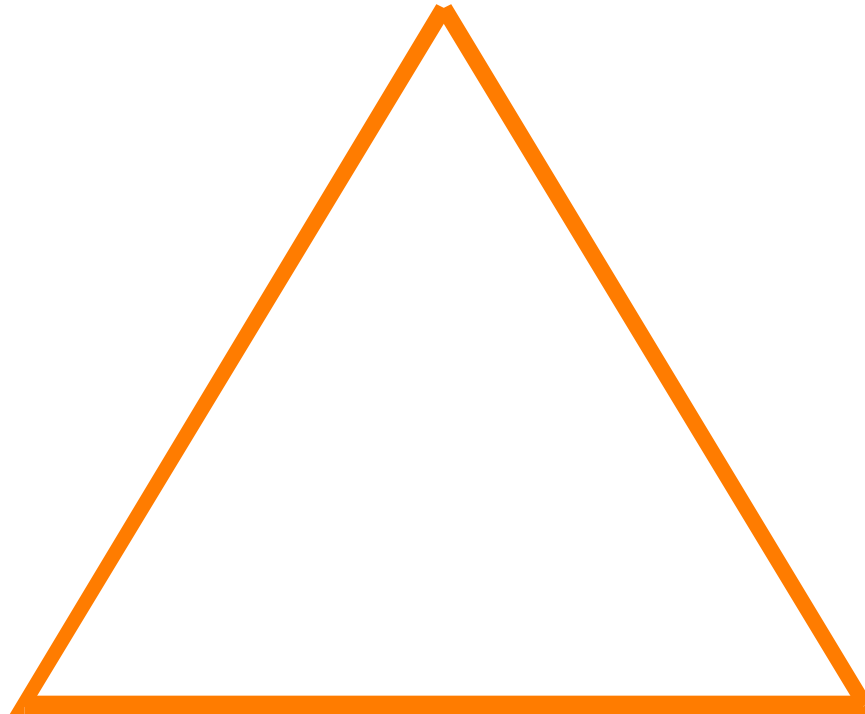
# Virtual Routers

## What to do then?



# Virtual Routers

**PERFORMANCE**



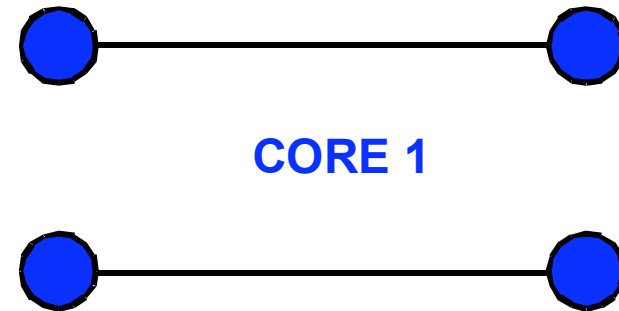
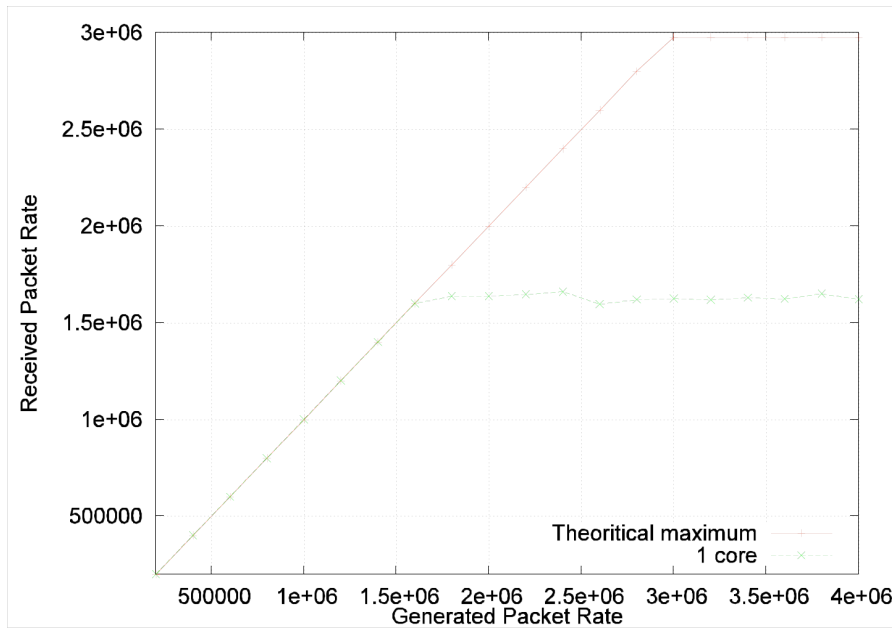
**ISOLATION**

**FLEXIBILITY**

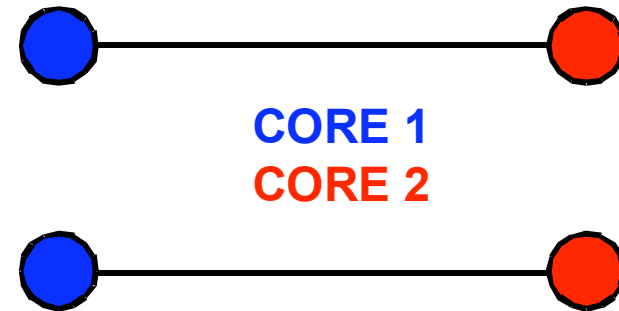
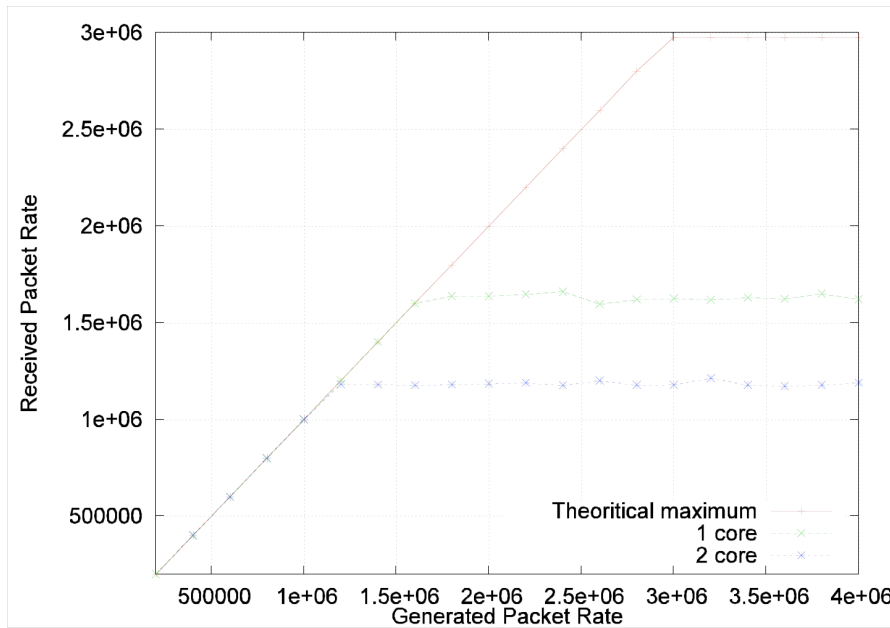
# Virtual Routers

- **Forwarding Path Virtualization**
  - Forwarding decision
    - Different views of the same network
    - Large memory requirement for separated forwarding tables
    - Merge them into a single table residing in the physical forwarding path
  - Classification
    - Determine to which virtual router the packet belongs
    - Similar issues to forwarding
  - Queuing
    - No issue on memory size, but significant increase in complexity
  - Scheduling
    - Provide fairness and isolation among the virtual routers
    - Ensure that packets are handled the same way as if only a single virtual router occupied the whole physical hardware

# Multi-core CPU cache issues

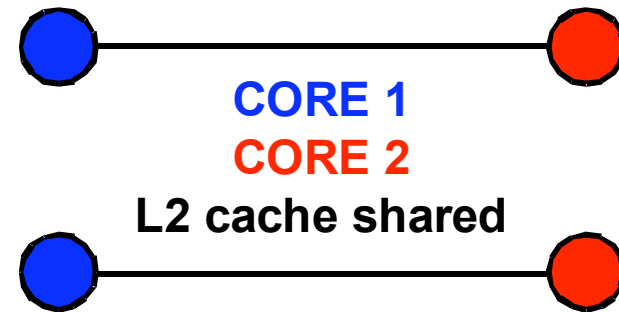
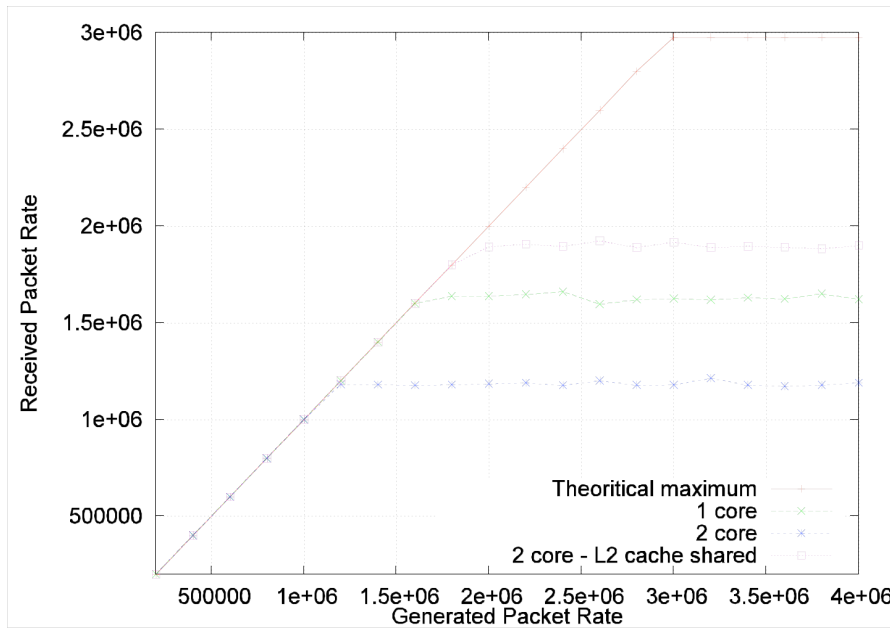


# Multi-core CPU cache issues

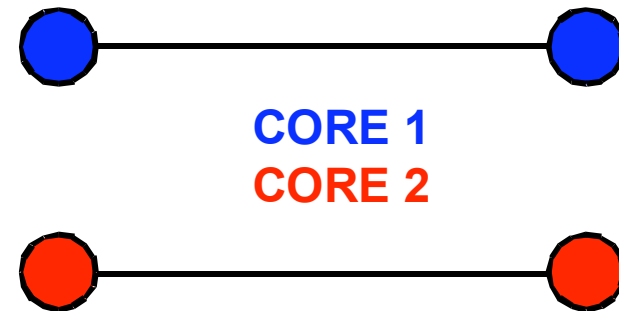
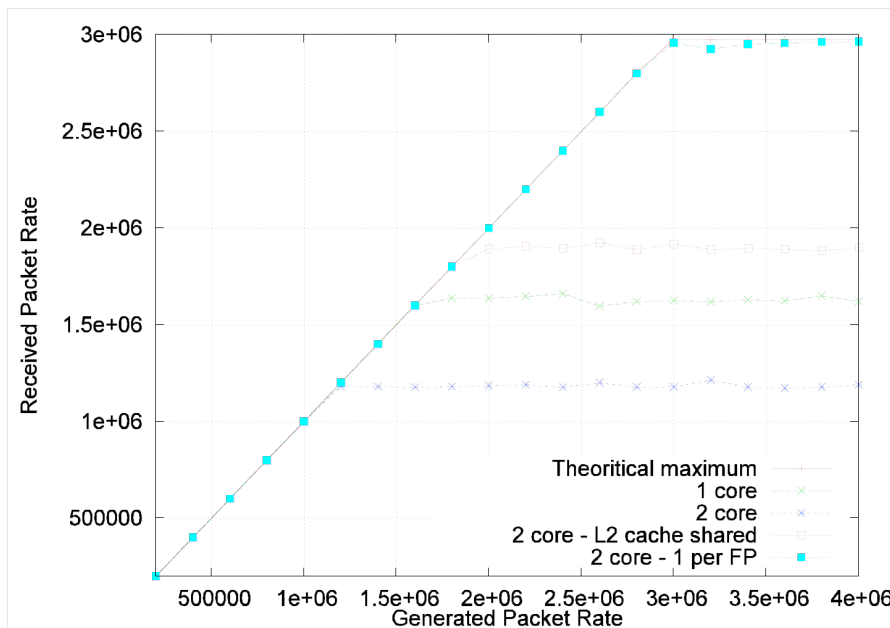




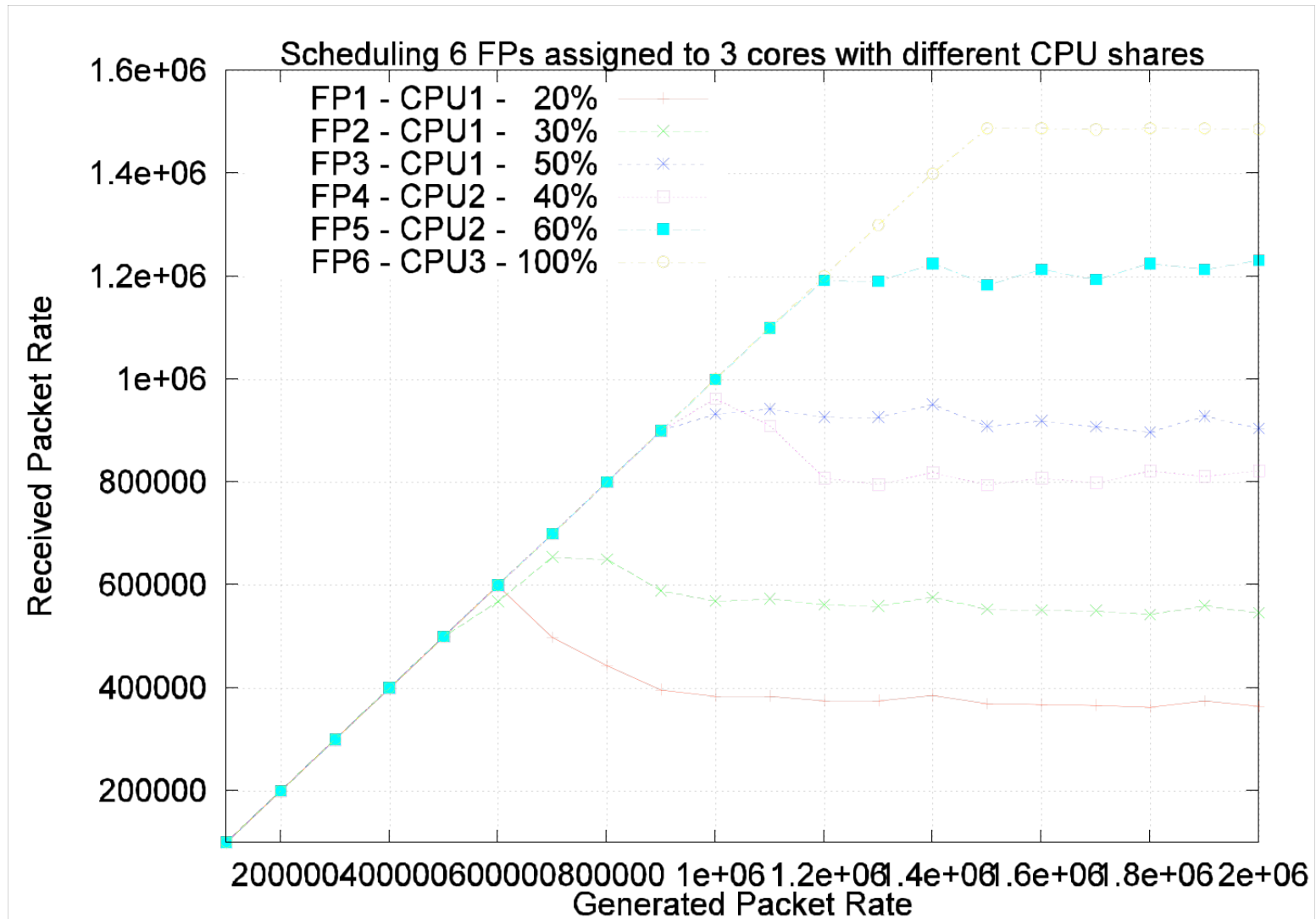
# Multi-core CPU cache issues



# Multi-core CPU cache issues



# Example: CPU isolation



# Some Conclusions

- **Forwarding in Dom0 for performance**
  - As close as possible to HW
  - Loose Xen isolation!
    - Must do it yourself
      - But Click can't do it across CPUs yet...
- **In general, must rethink functionality mapping in multi-core context**
  - But first must understand implications
- **Classification vs virtual Q support on NIC**
- **Many more open issues...**
  - Output bias to avoid losing packets on “empty queues”, etc



# Thank you for your attention!



© Scott Adams, Inc./Dist. by UFS, Inc.