July 2007

Using Xen For Router Virtualisation

Adam Greenhalgh, Mark Handley,Tim Schooley
University College of London

Norbert Egi, Mickaël Hoerdt, Laurent Mathy
University of Lancaster

Outline

- Motivations
- Xen network internals overview
- Experiments
- Results
- Conclusions and further work

Motivations

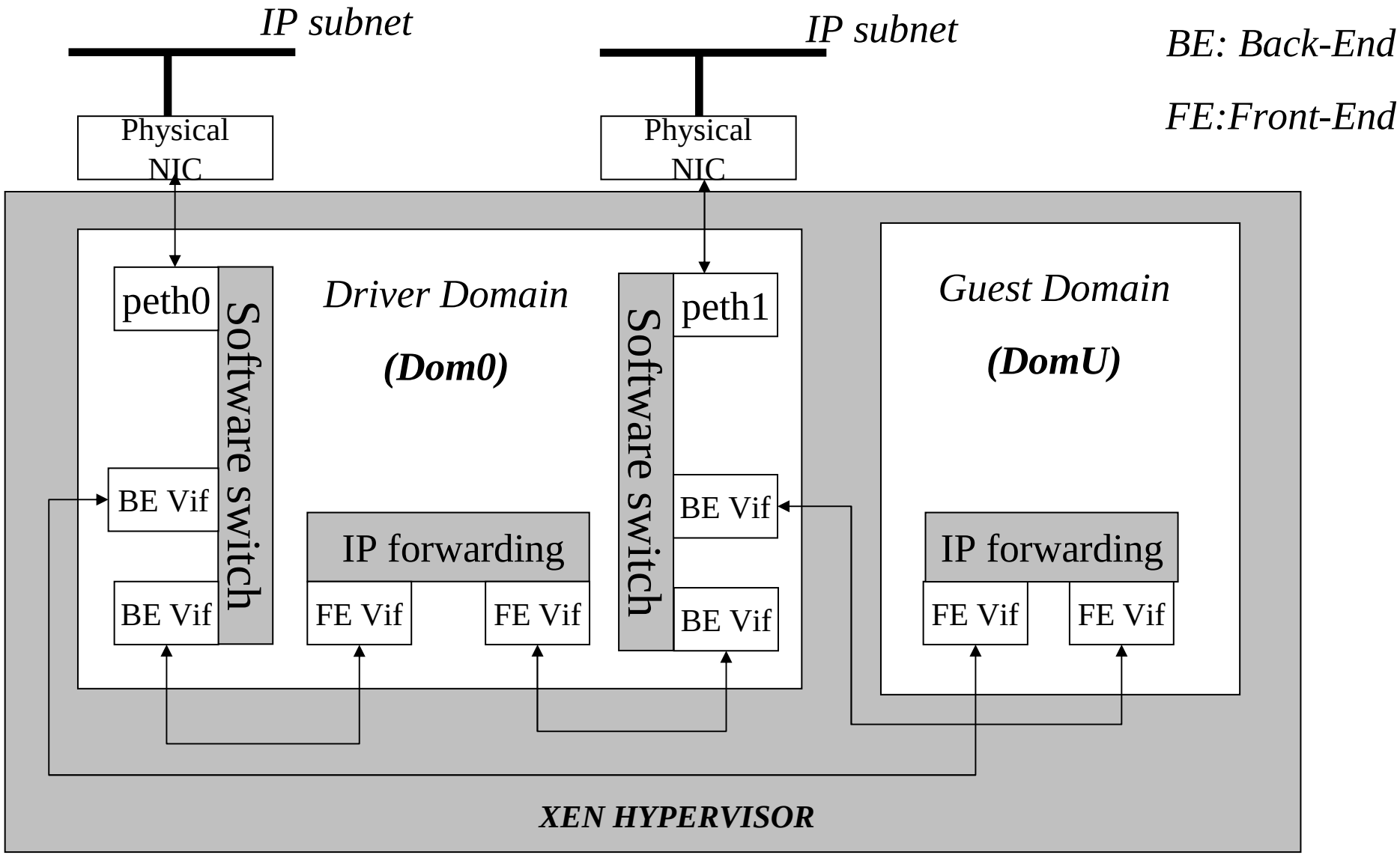
- Virtual Routers: Why?
  - One box can play the role of multiple independent routers.
    - Resources sharing, management flexibility
  - Multiple organizations sharing a single physical router.
    - Small businesses within one building.
    - Internet eXchange Points.
  - Entire physical network can be shared.
    - Virtualize the routers.
    - Tunnel between virtual routers over shared IP infrastructure.
    - And then you have an Inter-domain overlay
  - Great platform for experimentation (cf. VINI).

Motivations

- Enabling technologies:
  - Click Modular Router
    - Enable dynamically reprogrammable forwarding plane
  - Xorp Extensible Open Router platform
    - Enable an extensible open source control plane
  - Xen Virtual Machine Monitor
    - Hardware support on commodity PCs
- What could we build with that ?
- Can we achieve good performances ?

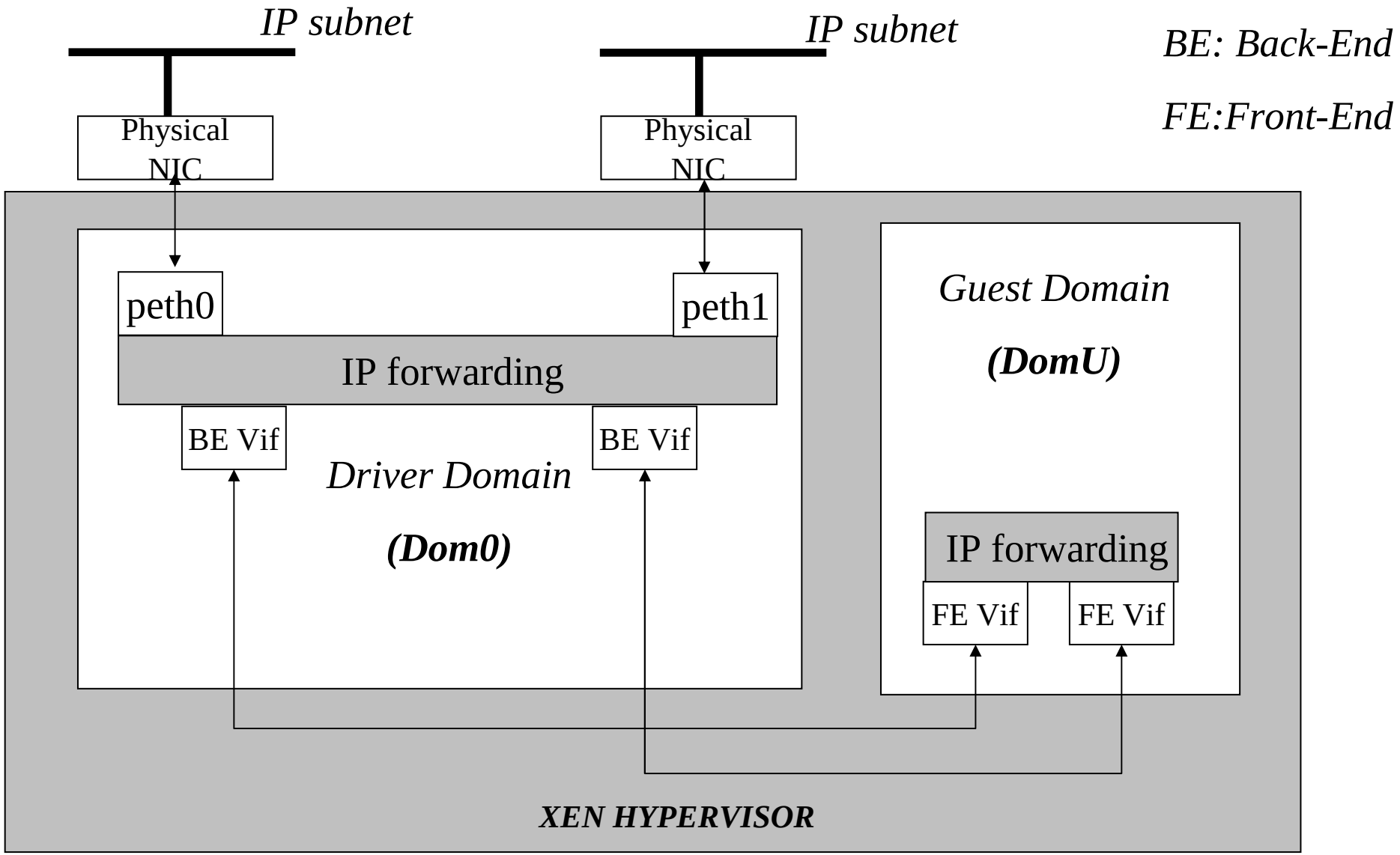
Xen network internal overview

- Xen classical bridged setup



Xen network internal overview

- Xen classical routed setup



Opened questions

- Question : Is it worth pushing the forwarding plane of a domU down to dom0 ?
  - How does dom0 forwarding performances compare with native linux performances ?
  - What is the impact of increasing the number of DomU's on dom0 forwarding performances ?
  - What is the impact of the routed and bridged classical Xen setup on the forwarding performances ?
  - How does the forwarding performances of dom0 compare to the forwarding performances of a domU ?

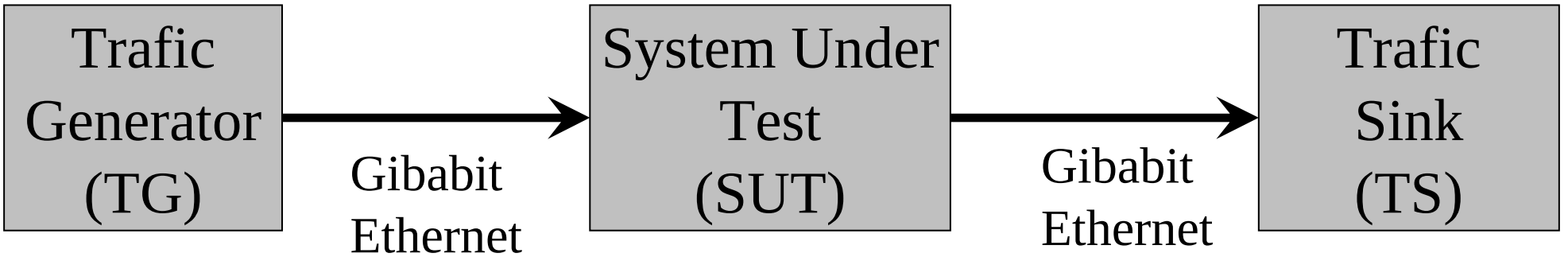
Experimental setup

- A 3 nodes topology :

Dell powerEdge 1950
- Two 2.66 Ghz dual core Intel® Xeon processors
- 2Gb of memory
- Gigabit Ethernet Intel® Pro/1000 NICS on PCI- Express x4 slot
- Linux 2.6.16.33 + Click pooling

Sun Fire X4100
- one 2.2 Ghz AMD Opteron™ processors (single core)
- 2Gb of memory
- Gigabit Ethernet Intel® Pro/1000 NICS on PCI-X 100Mhz slot
- Vanilla Linux 2.6.16.33 + NAPI
- **OR**
- Xeno Linux + Xen 3.0.4-1 + NAPI

Dell powerEdge 1950
- Two 2.66 Ghz dual core Intel® Xeon processors
- 2Gb of memory
- Gigabit Ethernet Intel® Pro/1000 NICS on PCI- Express x4 slot
- Linux 2.6.16.33 + Click pooling



TG : Generate an 10s CBR trafic with a rate ranging from 100 kpps to 1000 kpps, packet size 64 bytes, granularity 100kpps, using native linux or dom0 or/and domUs as routers

SUT: forward

TS: Measure the rate

Results: Dom0 only

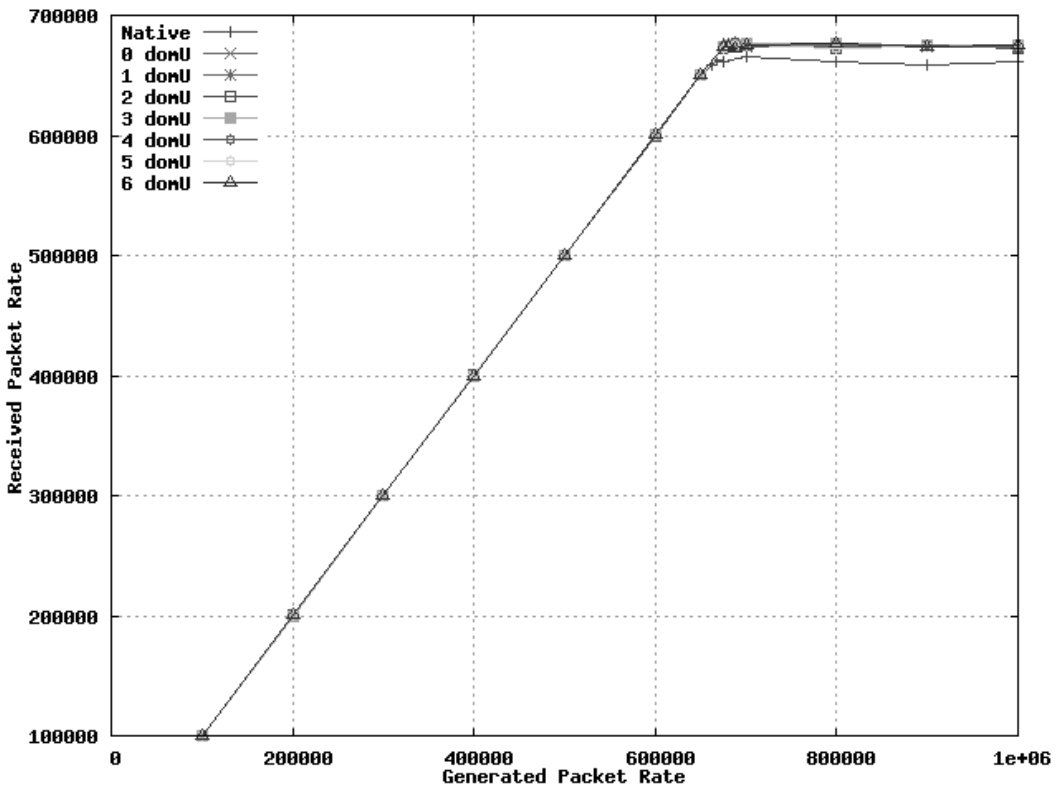


Fig 1: Dom0 forwarding performances in **routed** setup with different number of domUs vs native linux

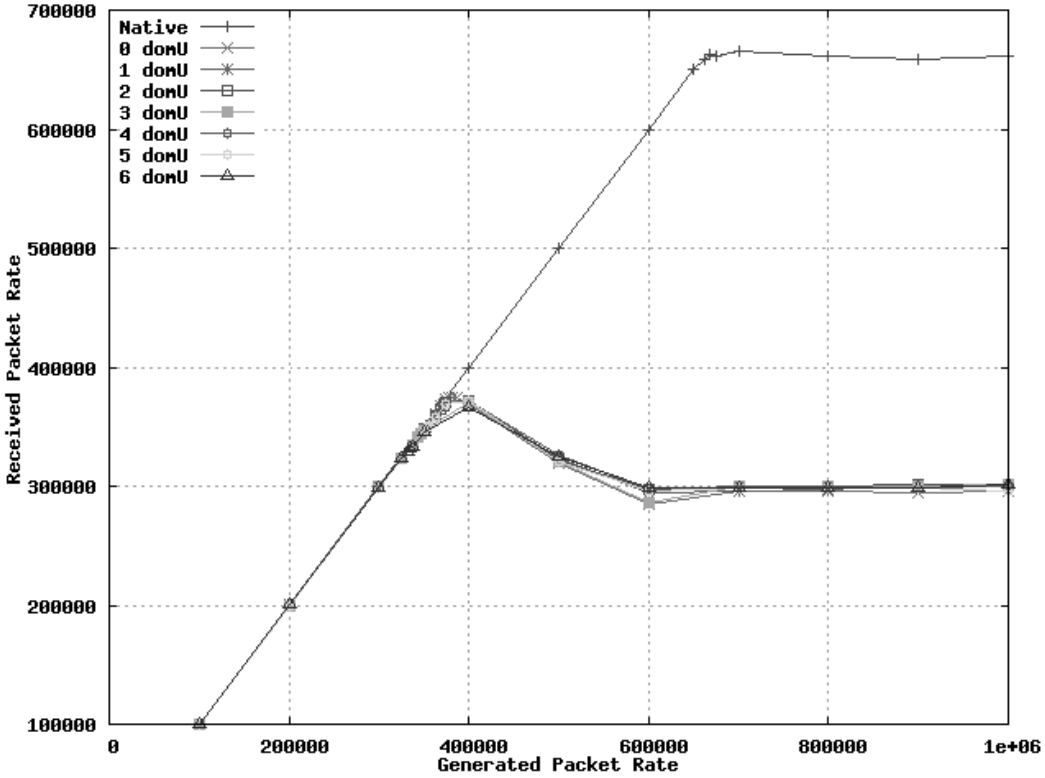
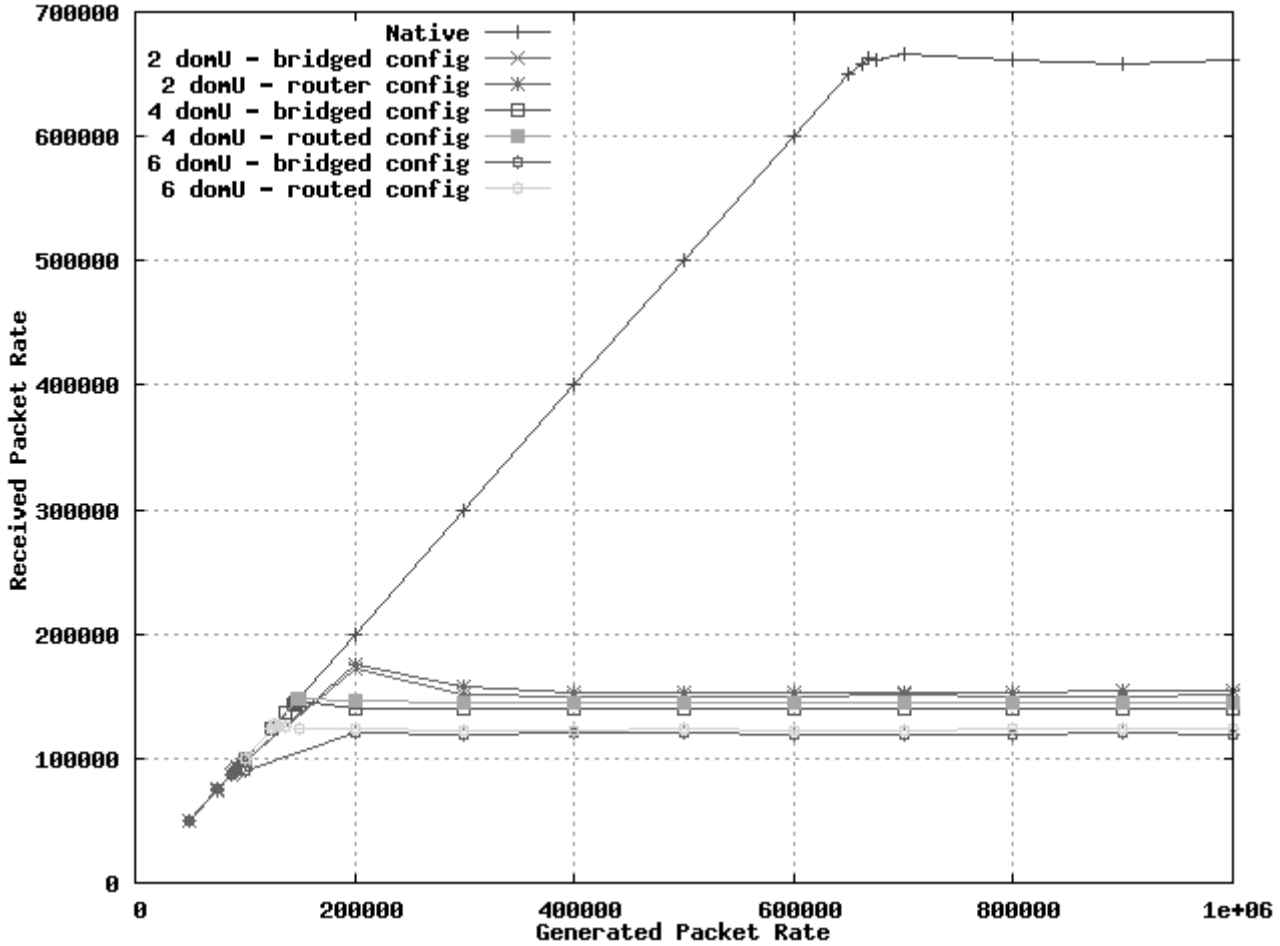


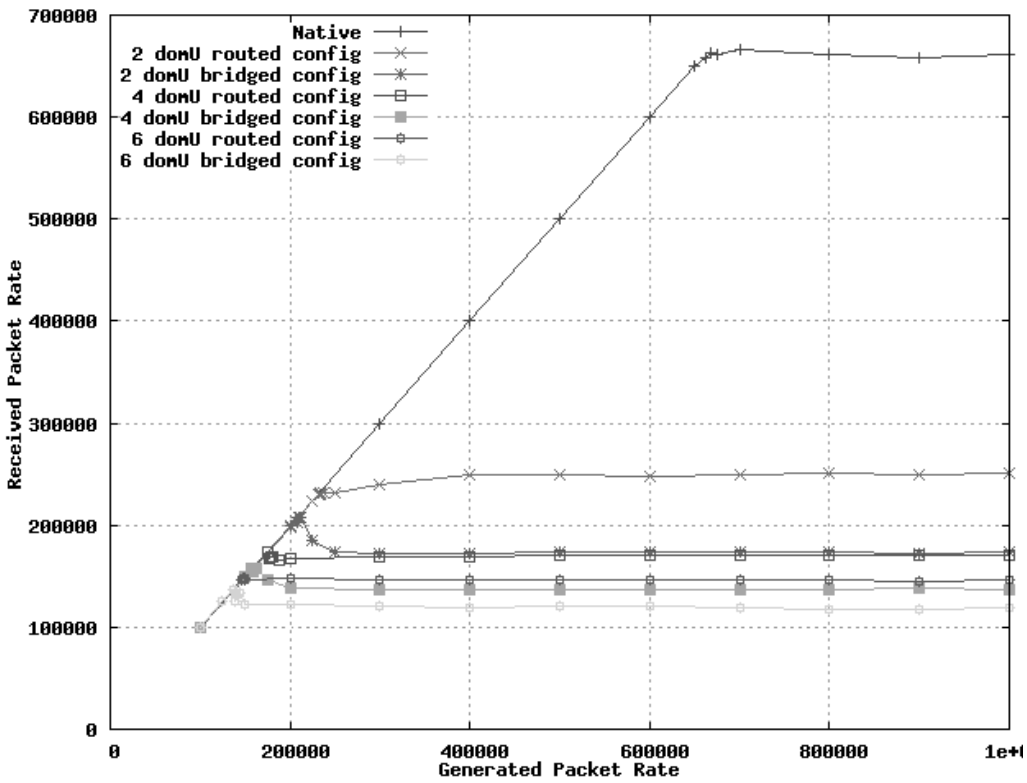
Fig 2: Dom0 forwarding performances in **bridged** setup with different numbers of domUs vs native linux

Results: DomUs only

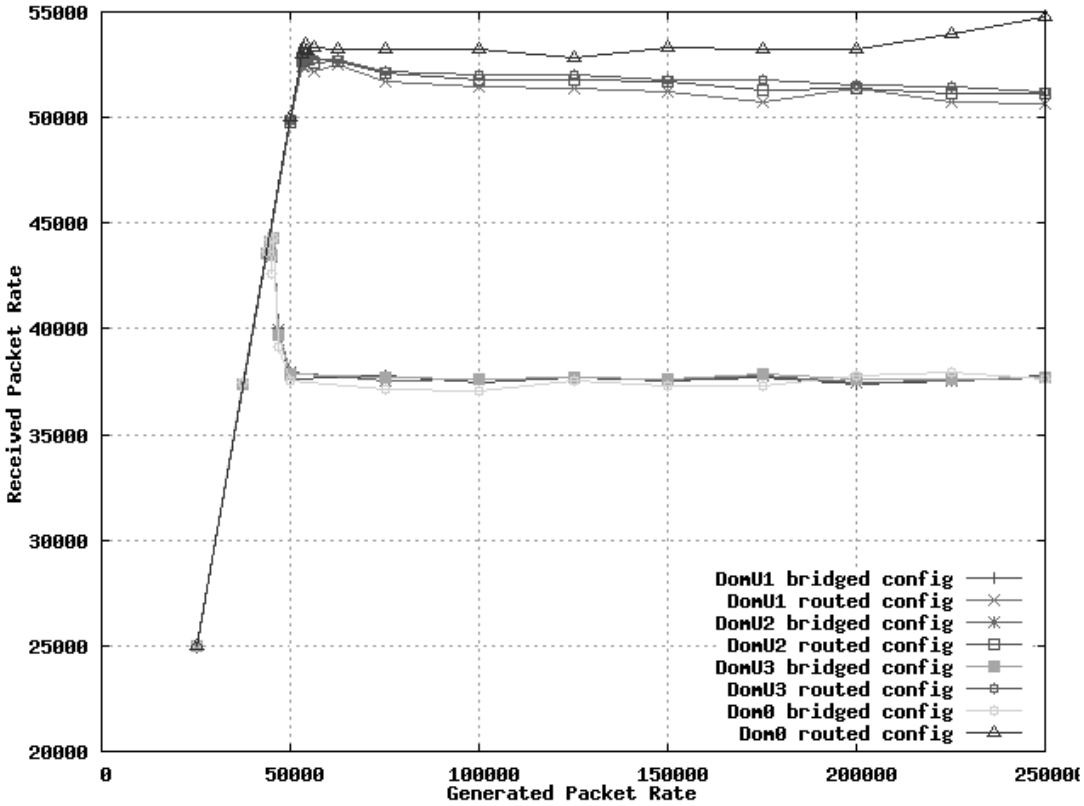


DomUs only aggregated forwarding performances in **bridged** and **routed** setup with different number of domUs vs native linux

Results: Dom0 and DomUs



Dom0 and DomU agregated forwarding performances in **bridged** and **routed** setup with different numbers of domUs vs native linux



Dom0 and DomUs flow breakdown forwarding performance in **bridged** and **routed** setup with 3 domUs

Conclusions and Further work

- Conclusions:
  - Dom0 forwarding performances are good compared to native Linux!
  - *But* DomU's forwarding performances aren't.
  - *And* DomU's forwarding impact Dom0 performances badly!
- Further work
  - Use Click and Xorp to design a programmable shared forwarding plane running in dom0 on the behalf of the domUs

Results: 6 DomUs+Dom0 flow breakdown

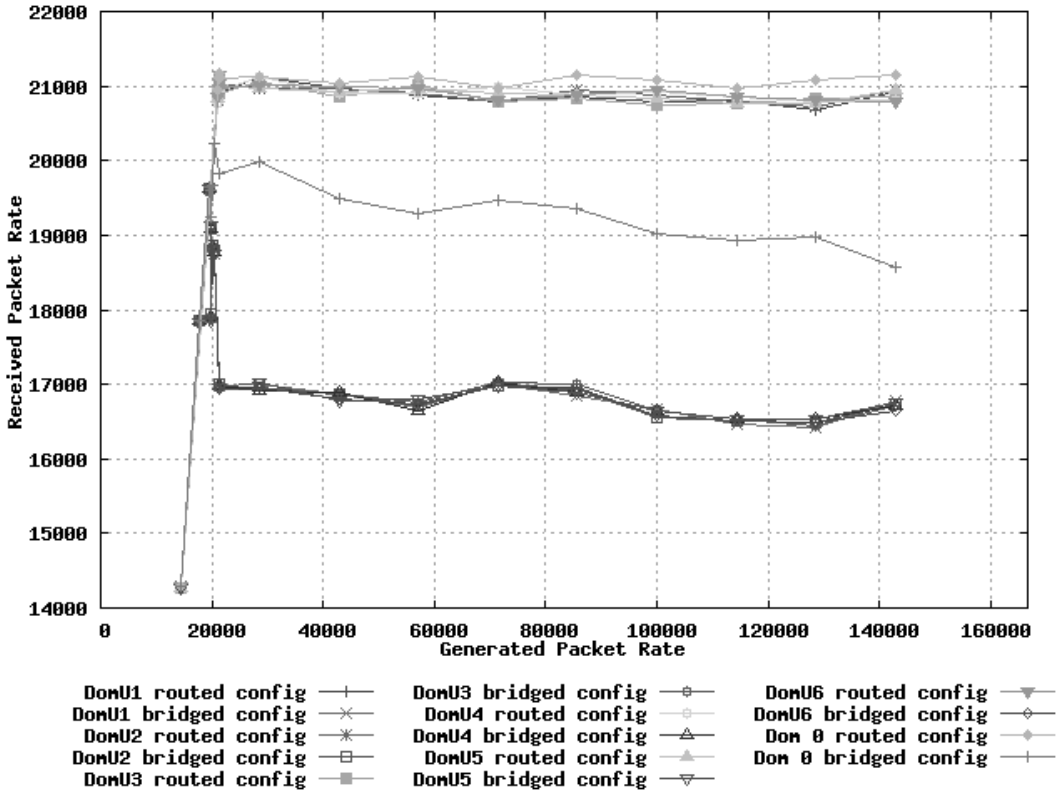


Fig 1: Sun Fire X4100 Dom0 and domU flow breakdown forwarding performances in **bridged** and **routed** setup with 6 domUs

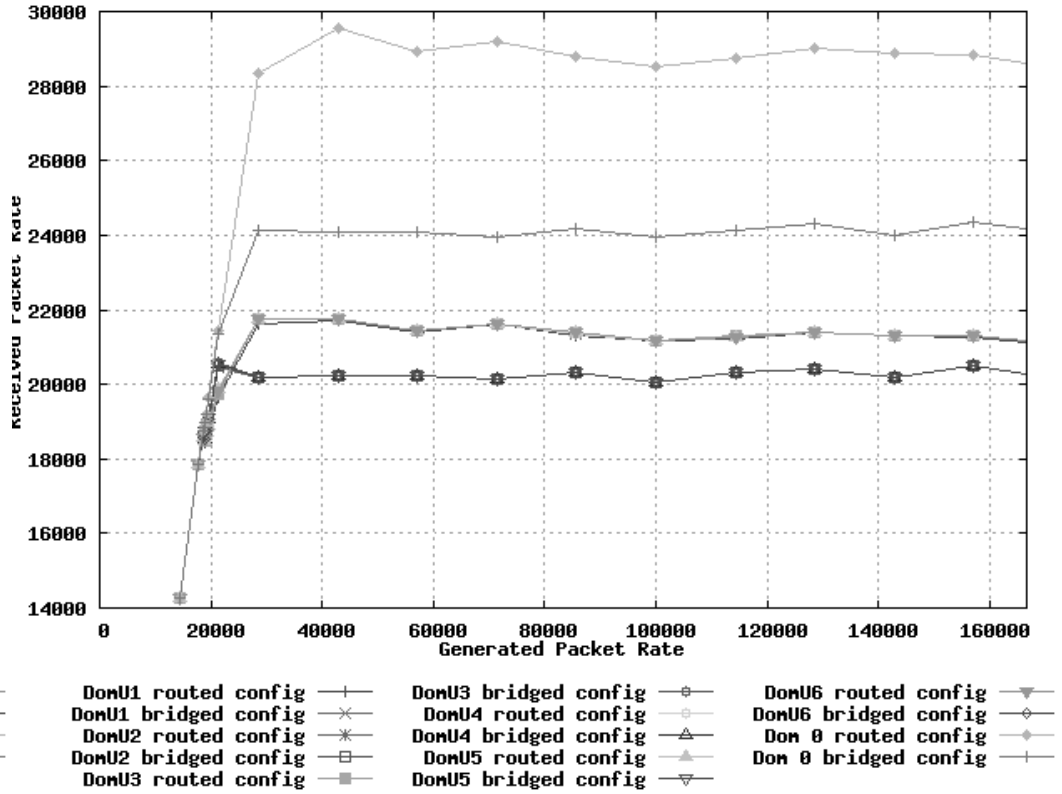
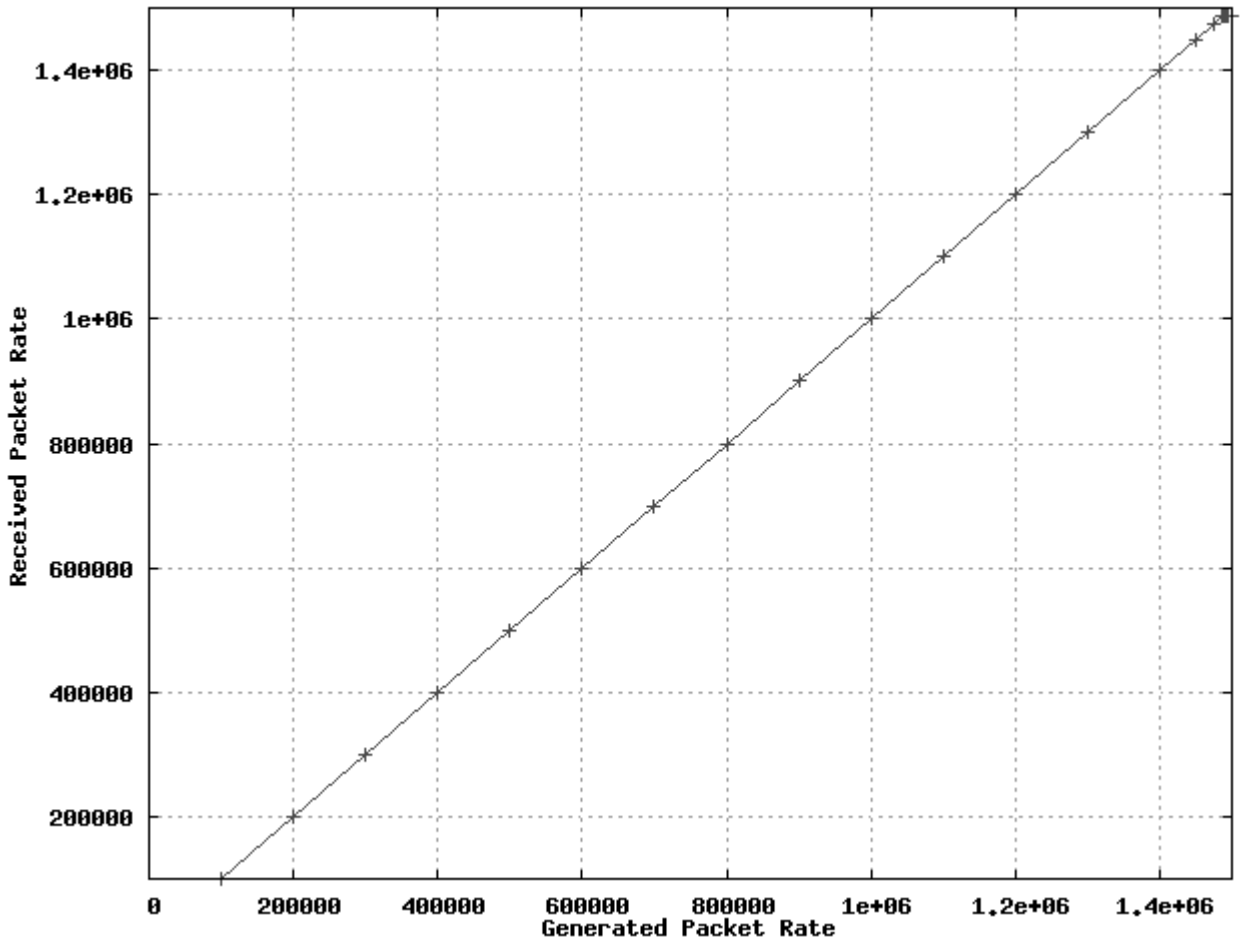


Fig 2: Dell 1950 Dom0 and domU flow breakdown forwarding performances in **bridged** and **routed** setup with 6 domUs

Experimental setup

- Checking TG to TS performances:



TG to TS IP link performances