Empowered by Innovation NEC

ClickOS and the Art of Network Function Virtualization

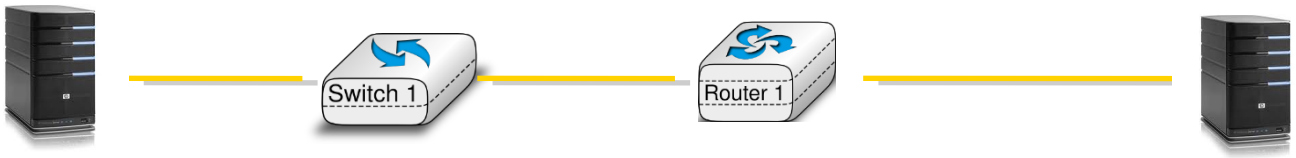
Joao Martins*, Mohamed Ahmed*, Costin Raiciu§, Roberto Bifulco*, Vladimir Olteanu§, Michio Honda*, Felipe Huici*

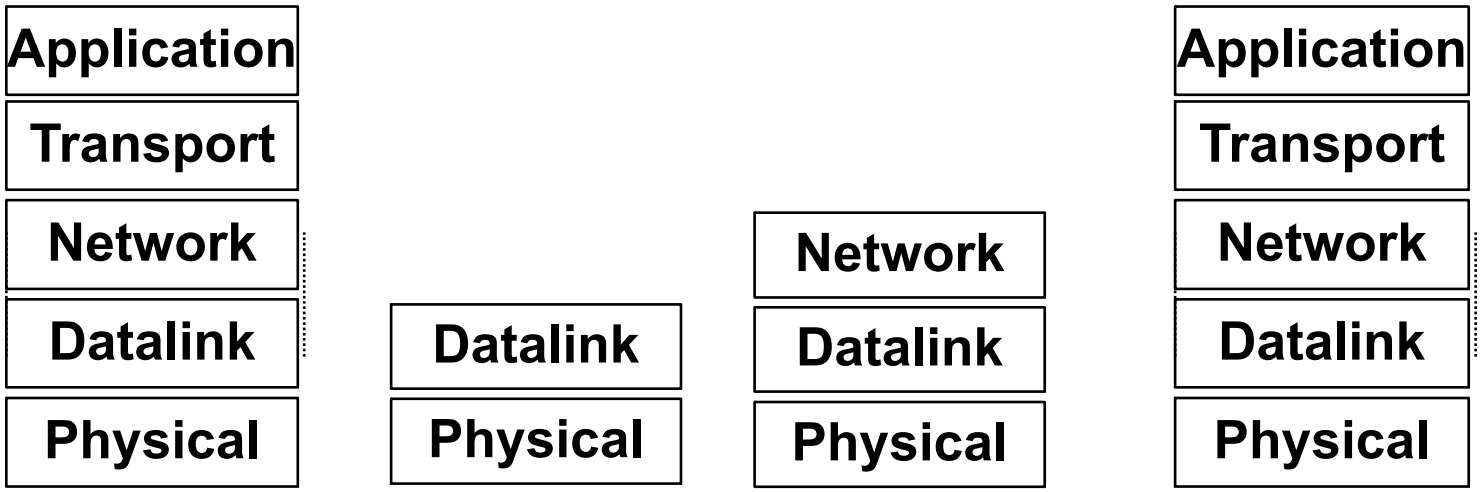
*** NEC Labs Europe, Heidelberg, Germany**

§ University Politehnica of Bucharest

firstname.lastname@neclab.eu, firstname.lastname@cs.pub.ro

The Idealized Network





A Middlebox World

ad insertion

carrier-grade NAT

BRAS

WAN accelerator

transcoder

session border controller

IDS

load balancer

DDoS protection

firewall

QoE monitor

DPI

Hardware Middleboxes - Drawbacks

- Expensive equipment/power costs
- Difficult to add new features (vendor lock-in)
- Difficult to manage
- Cannot be scaled on demand (peak planning)

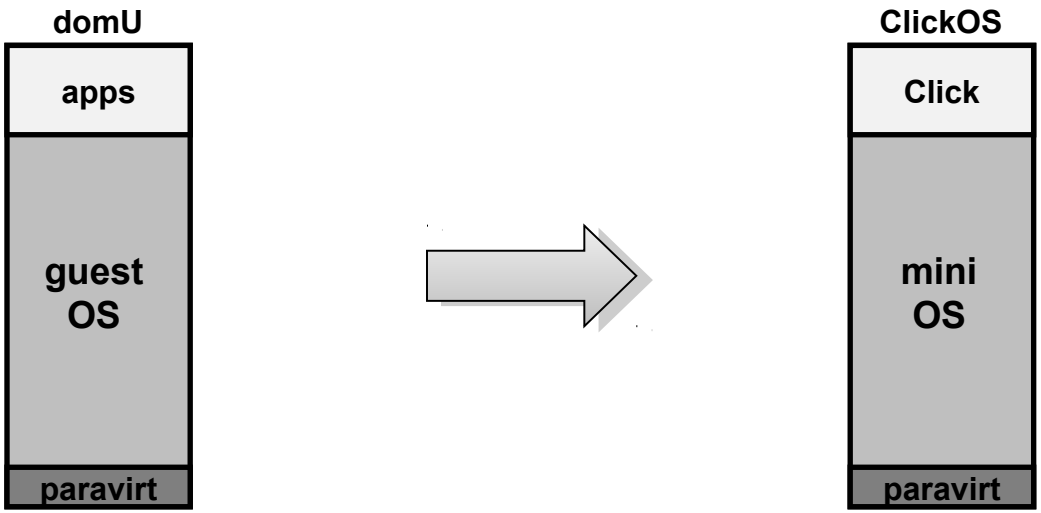
Shifting Middlebox Processing to Software

- Can share the same hardware across multiple users/tenants
- Reduced equipment/power costs through consolidation
- Safe to try new features on a operational network/platform
- **But can it be built using commodity hardware while still achieving high performance?**
- **ClickOS:** tiny Xen-based virtual machine that runs Click

From Thought to Reality - Requirements

	ClickOS
Fast Instantiation	 **30 msec boot times**
Small footprint	 **5MB when running**
Isolation	 **provided by Xen**
Performance	 **10Gb/s line rate* 45 µsec delay**
Flexibility	 **provided by Click**

What's ClickOS ?

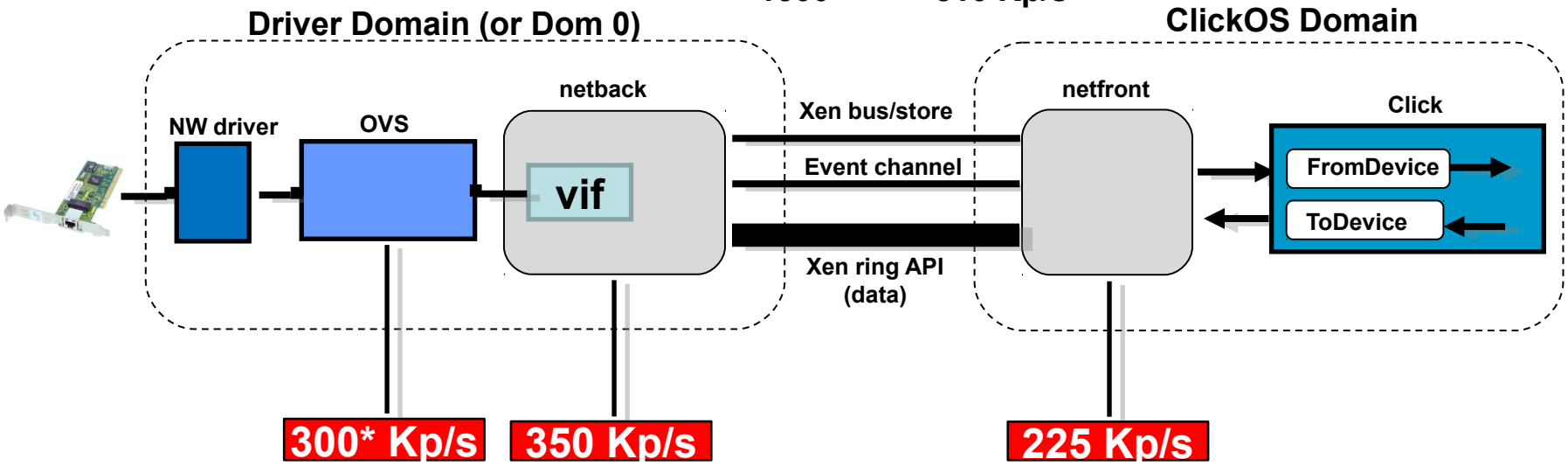


Work consisted of:

- Build system to create ClickOS images (5 MB in size)
- Emulating a Click control plane over MiniOS/Xen
- Reducing boot times (roughly 30 milliseconds)
- Optimizations to the data plane (10 Gb/s for almost all pkt sizes)
- Implementation of a wide range of middleboxes

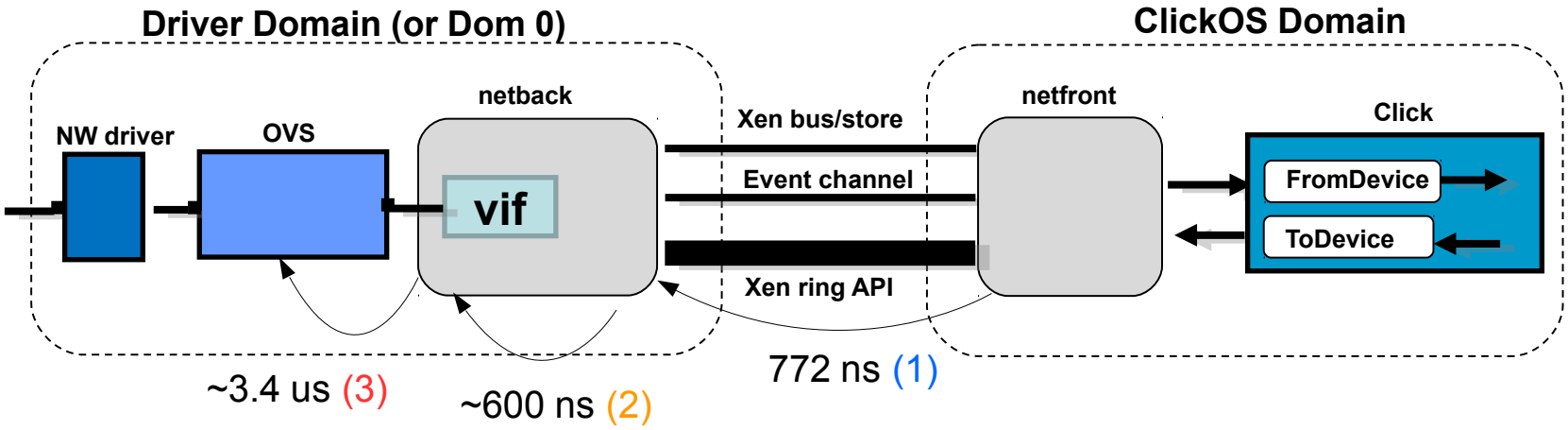
Performance analysis

packet size (bytes)	10 Gbit/s rate
64	14.88 Mp/s
128	8.4 Mp/s
256	4.5 Mp/s
512	2.3 Mp/s
1024	1.2 Mp/s
1500	810 Kp/s



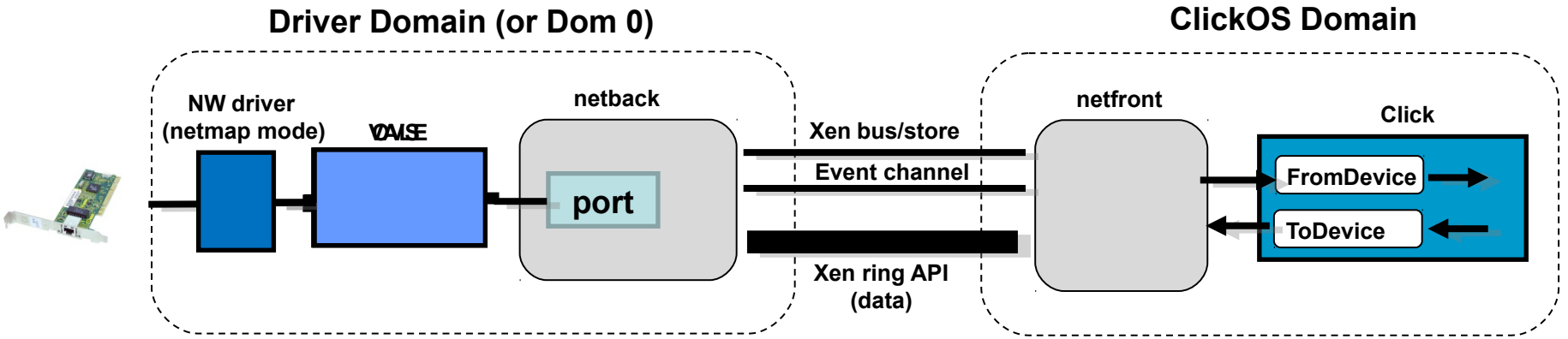
* - maximum-sized packets

Performance analysis

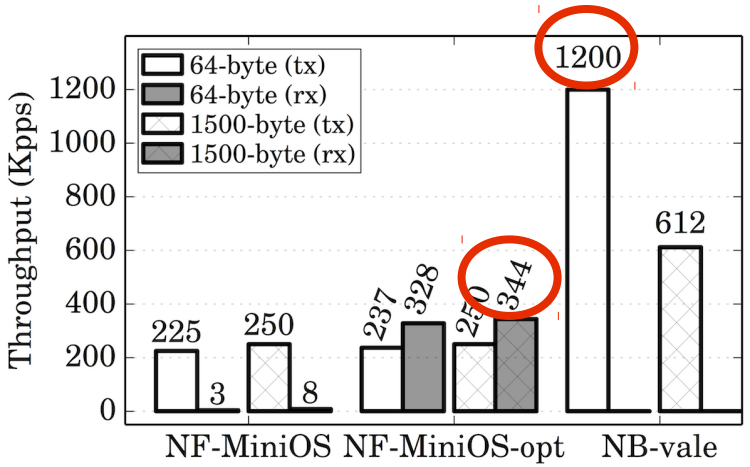


- Copying packets between guests greatly affects packet I/O (1)
- Packet metadata allocations (2)
- Backend switch is slow (3)
- MiniOS netfront not as good as Linux

Optimizing Network I/O – Backend Switch

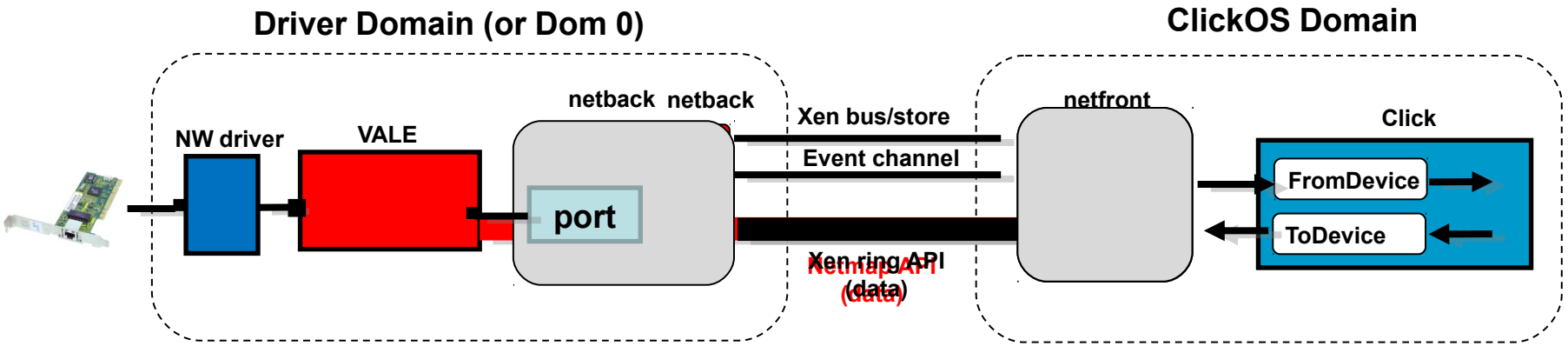


- Reuse Xen page permissions (frontend)
- Introduce VALE[1] as the backend switch
- Increase I/O requests batch size



[1] VALE, a switched ethernet for virtual machines, ACM CoNEXT'2012
Luigi Rizzo, Giuseppe Lettieri
Universita di Pisa

Optimizing Network I/O



- Minimal memory requirements
  - **For max. throughput a guest only needs 4 MB of memory**
- Breaks other (non-MiniOS) guests
  - But we have implemented Linux netfront driver

slots	KB (per ring)	# grants (per ring)
64	135	33
128	266	65
256	528	130
512	1056	259
1024	2117	516
2048	4231	1033

ClickOS Prototype Overview

- Click changes are minimal ~600 LoC
- New toolstack for fast boot times
- Cross compile toolchain for MiniOS-based apps
- `netback` changes comprise ~500 LoC
- `netfront` (Linux/MiniOS) around ~600 LoC
- VALE switch extended to:
  - Connect NIC ports and modular switching

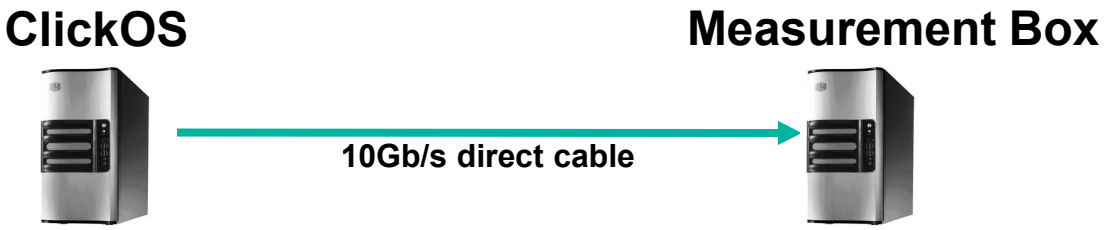
EVALUATION

Experiments

- ClickOS Instantiation
- State reading/insertion performance
- Delay compared with other systems
- Memory footprint

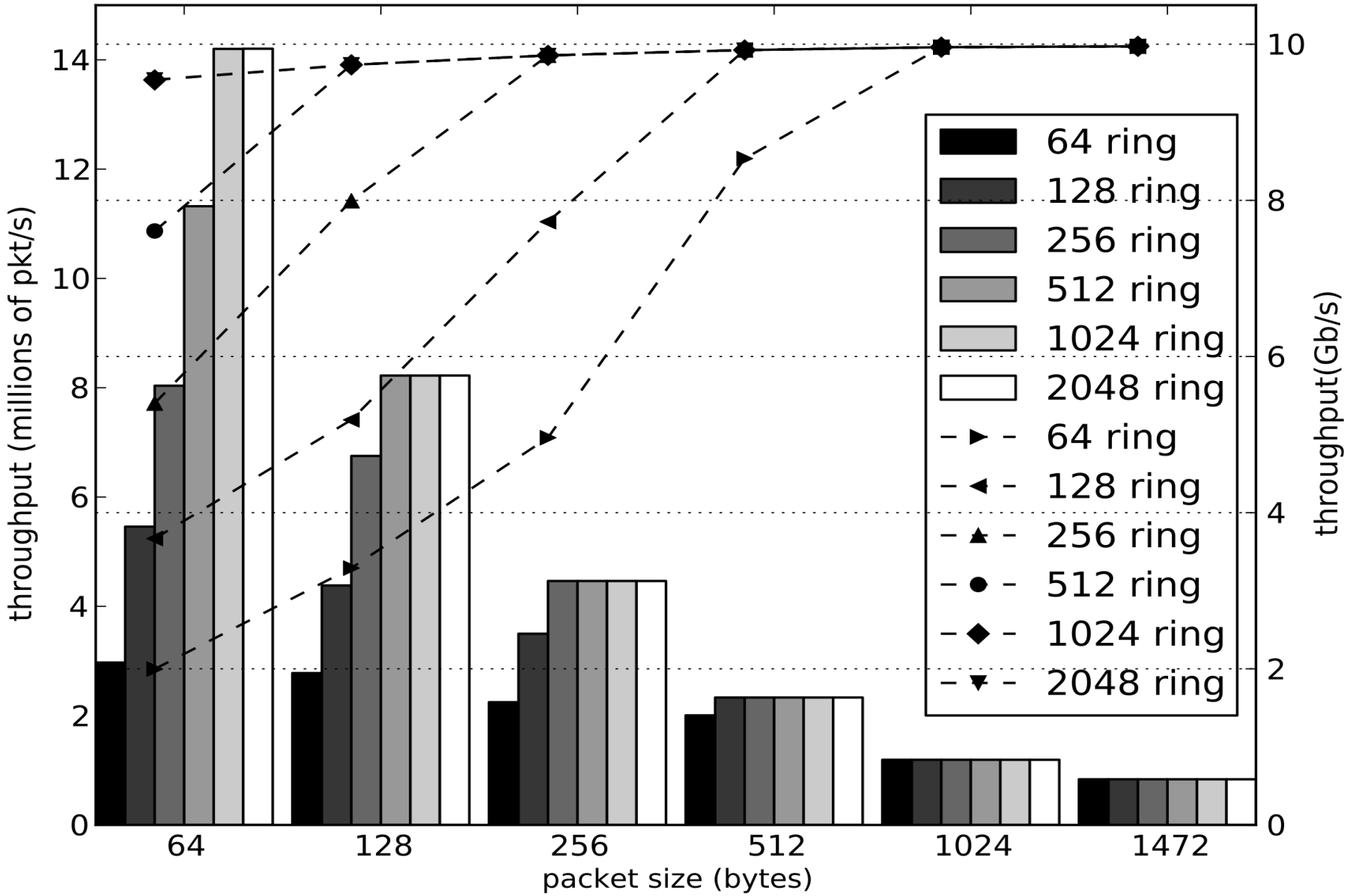
- Switch performance for 1+ NICs
- ClickOS/MiniOS performance
- Chaining experiments
- Scalability over multiple guests
- Scalability over multiple NICs
- Implementation and evaluation of middleboxes
- Linux Performance

ClickOS Base Performance



Intel Xeon E1220 4-core 3.2GHz (Sandy bridge)
16GB RAM, 1x Intel x520 10Gb/s NIC.
One CPU core assigned to VMs, the rest to the Domain-0
Linux 3.6.10

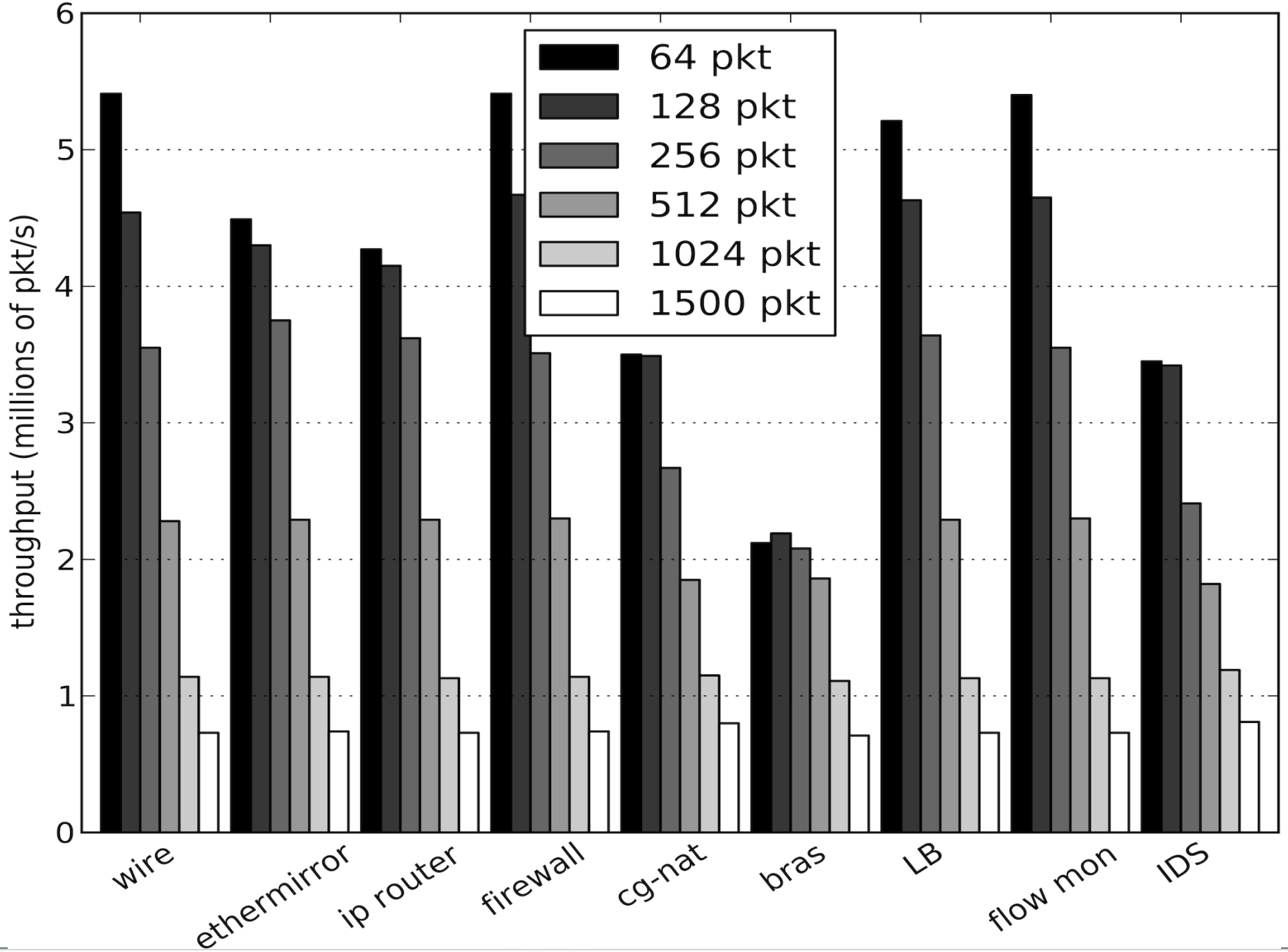
ClickOS Base TX Performance



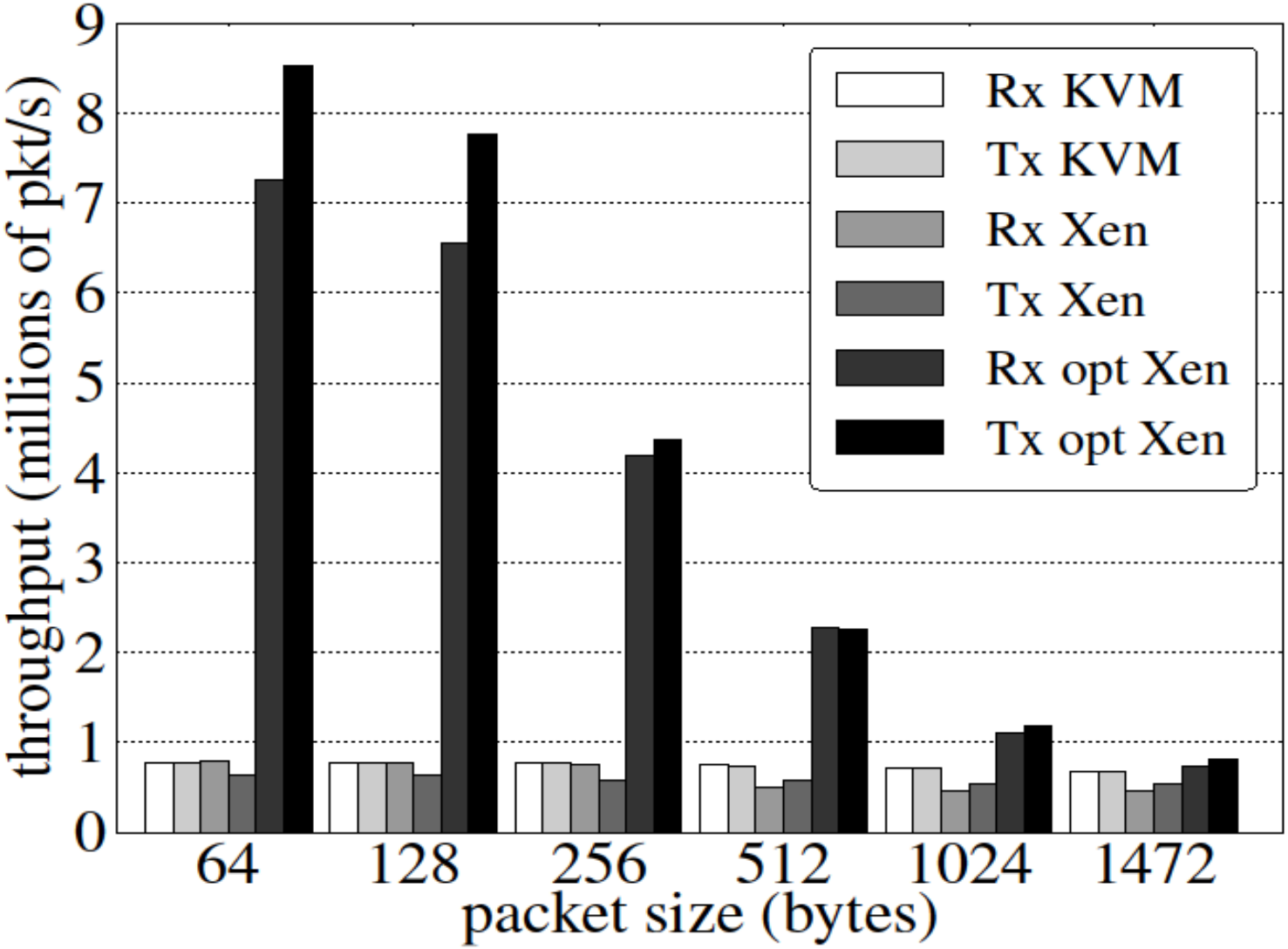
ClickOS (virtualized) Middlebox Performance

Intel Xeon E1220 4-core 3.2GHz (Sandy bridge)
16GB RAM, 2x Intel x520 10Gb/s NIC.
One CPU core assigned to Vms, 3 CPU cores Domain-0
Linux 3.6.10

ClickOS (virtualized) Middlebox Performance

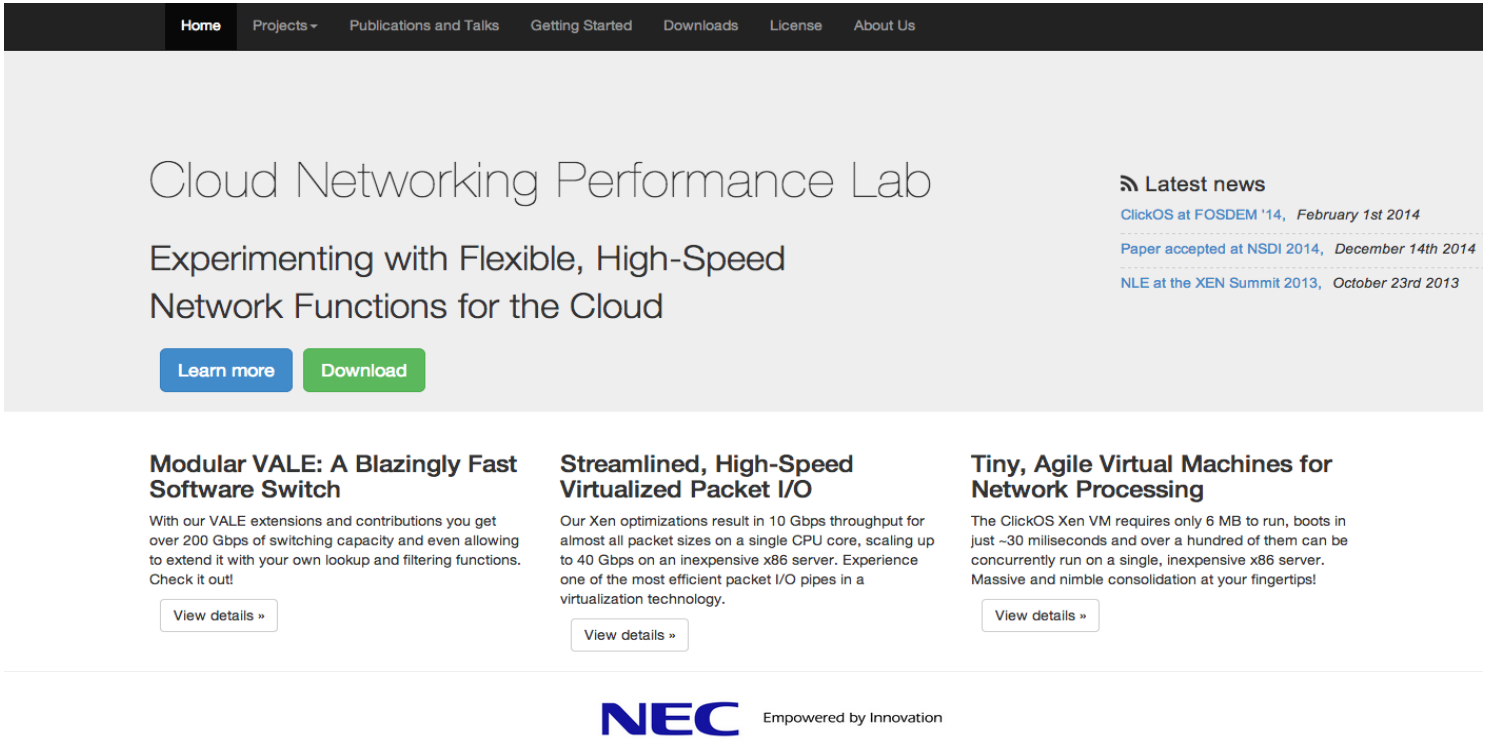


Linux Guest Performance



- Note that our Linux optimizations apply only to netmap-based applications

It's Open Source!



Checkout http://cnp.neclab.eu

- **ClickOS, Backend Switch, Xen optimizations and more!**
- **Github (** https://github.com/cnplab **)**
- **Tutorials**
- **Better performance!**

Conclusions

Virtual machines can do flexible high speed networking

ClickOS: Tailor-made operating system for network processing

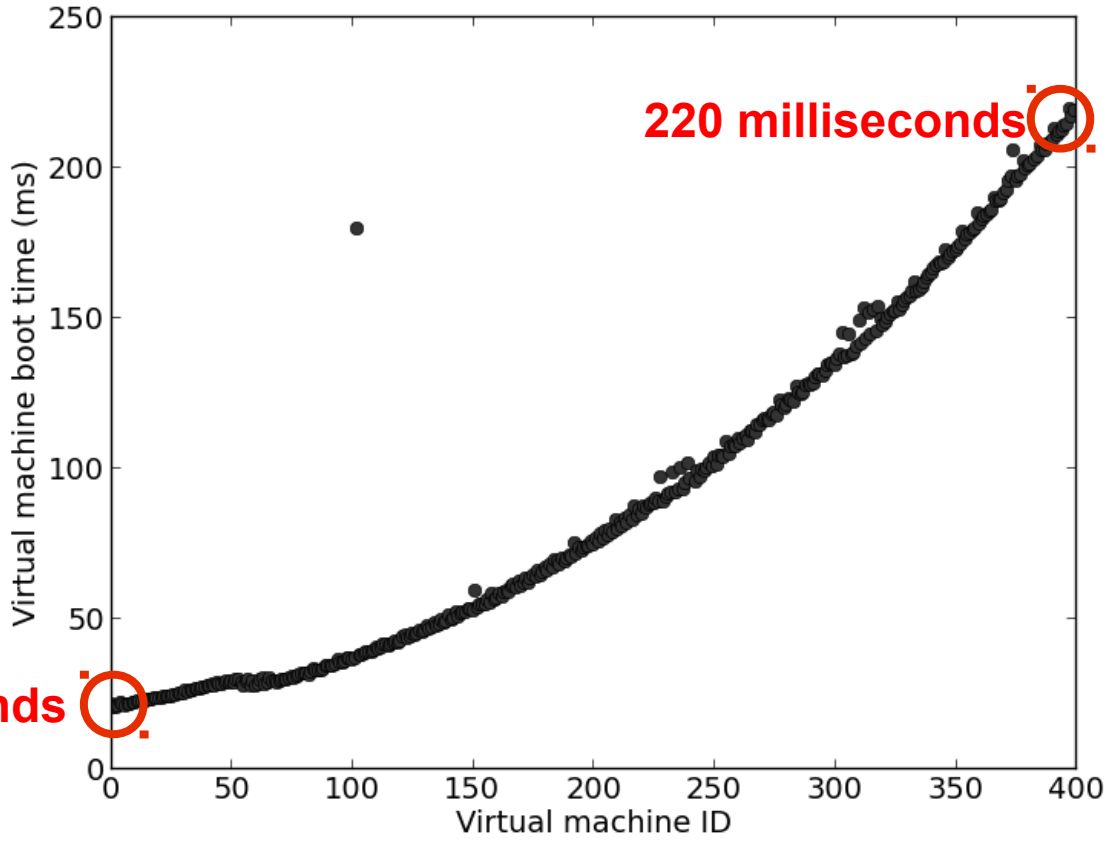
- **Small is better:** Low footprint is the key to heavy consolidation
- **Memory footprint:** 5MB
- **Boot time:** 30ms

Future work:

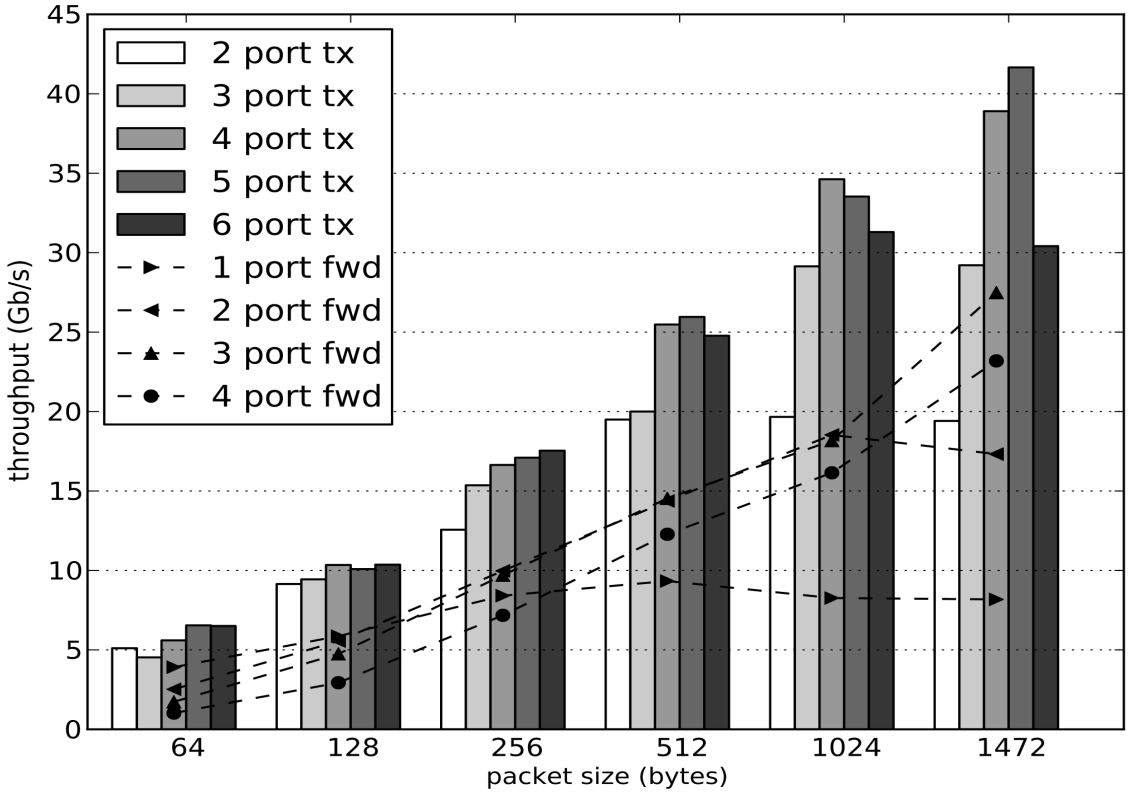
- Massive consolidation of VMs (thousands)
- Improved Inter-VM communication for service chaining
- Reactive VMs (e.g., per-flow)

Empowered by Innovation

ClickOS Boot times

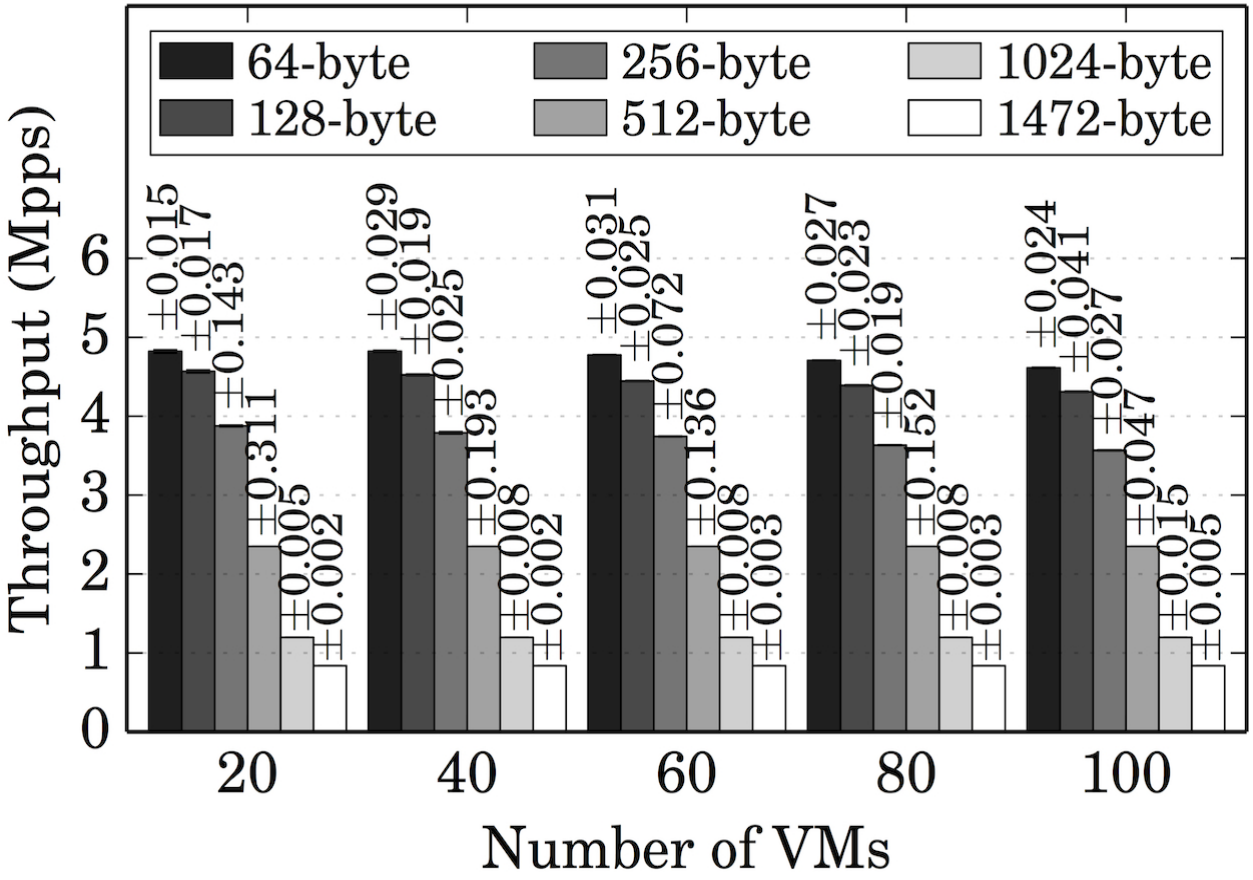


Scaling out – Multiple NICs/VMs



Intel Xeon E1650 6-core 3.2GHz, 16GB RAM, dual-port Intel x520 10Gb/s NIC.
3 cores assigned to VMs, 3 cores for dom0

Scaling out – 100 VMs Aggregate Throughput



Intel Xeon E1650 6-core 3.2GHz, 16GB RAM, dual-port Intel x520 10Gb/s NIC.
3 cores assigned to VMs, 3 cores for dom0

ClickOS Delay vs. Other Systems