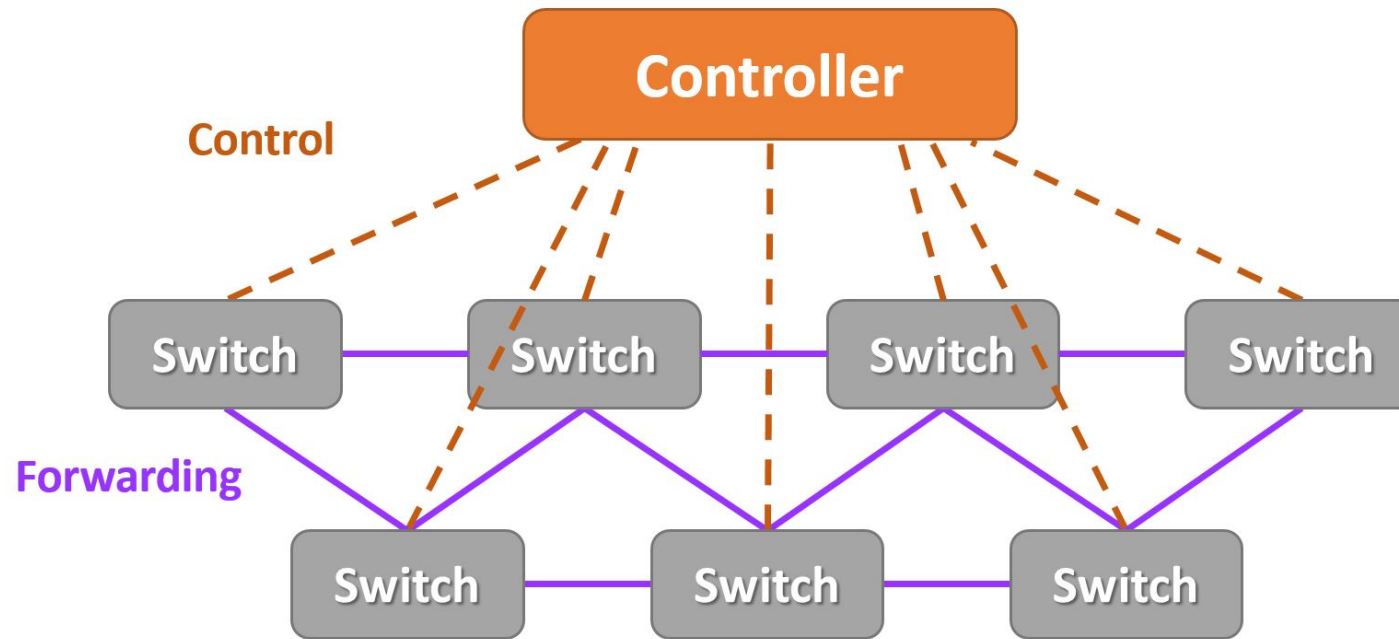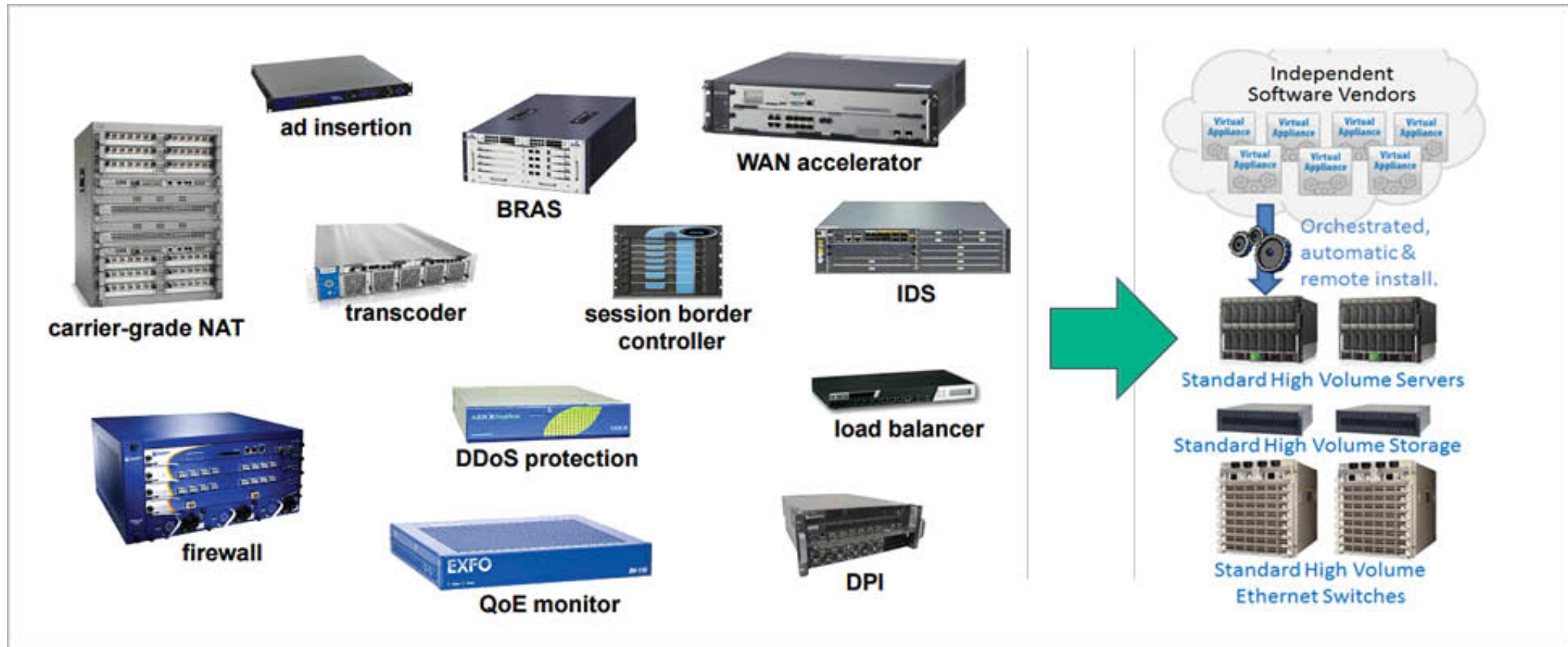# AOS Project

黄宏毅

hhy17@mails.tsinghua.edu.cn

https://github.com/fjzzhongyi/aos_hw
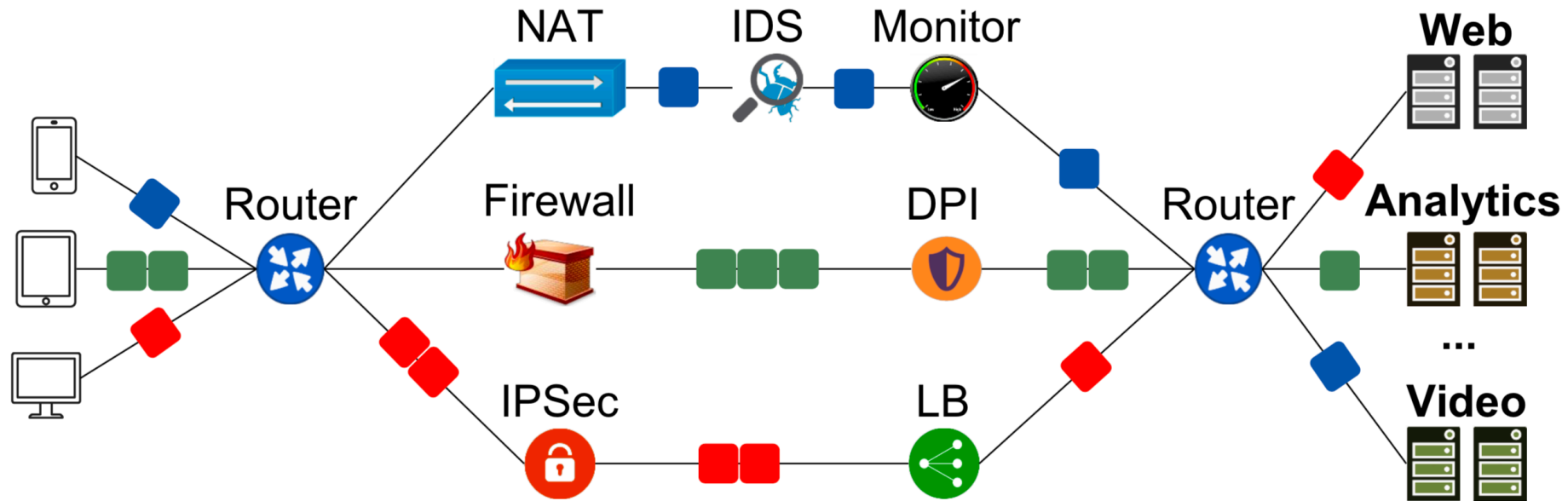
SDN
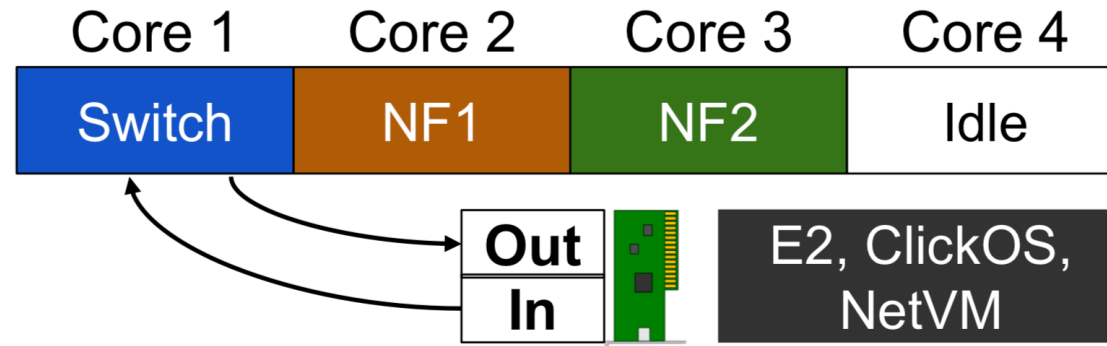
NFV

Metron: NFV Service Chains at the True Speed of the Underlying Hardware

- NF Chain

Dedicated switch

Core 1	Core 2	Core 3	Core 4
Switch	NF1	NF2	Idle

Out
In

E2, ClickOS, NetVM

Pipeline

Rx	Rx	NF1+NF2	Tx

RSS

Out
In

NFP, Flurries

RSS/FD

Rx+NF1	Idle	NF2+Tx	Idle

Flow Director
Or   RSS

Out
In

OpenBox, SNF, FastClick
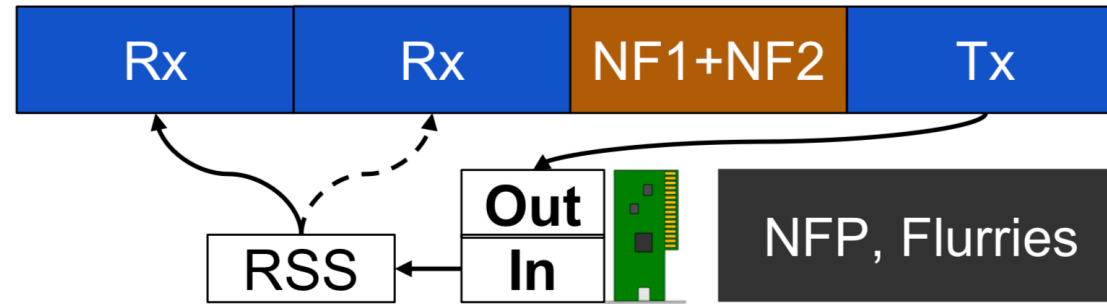
Programmable Switch

match-action

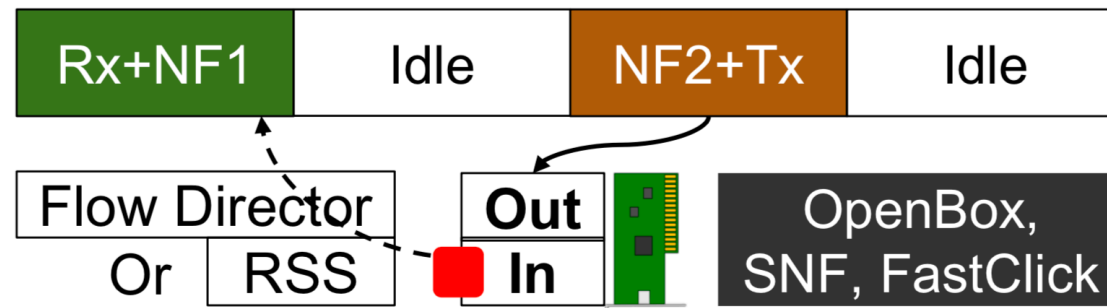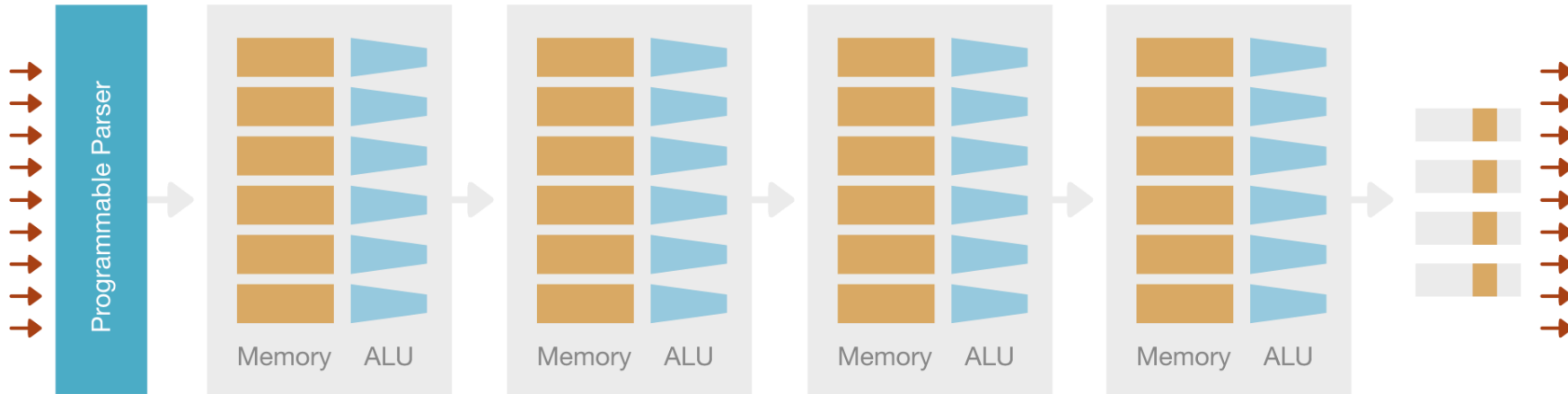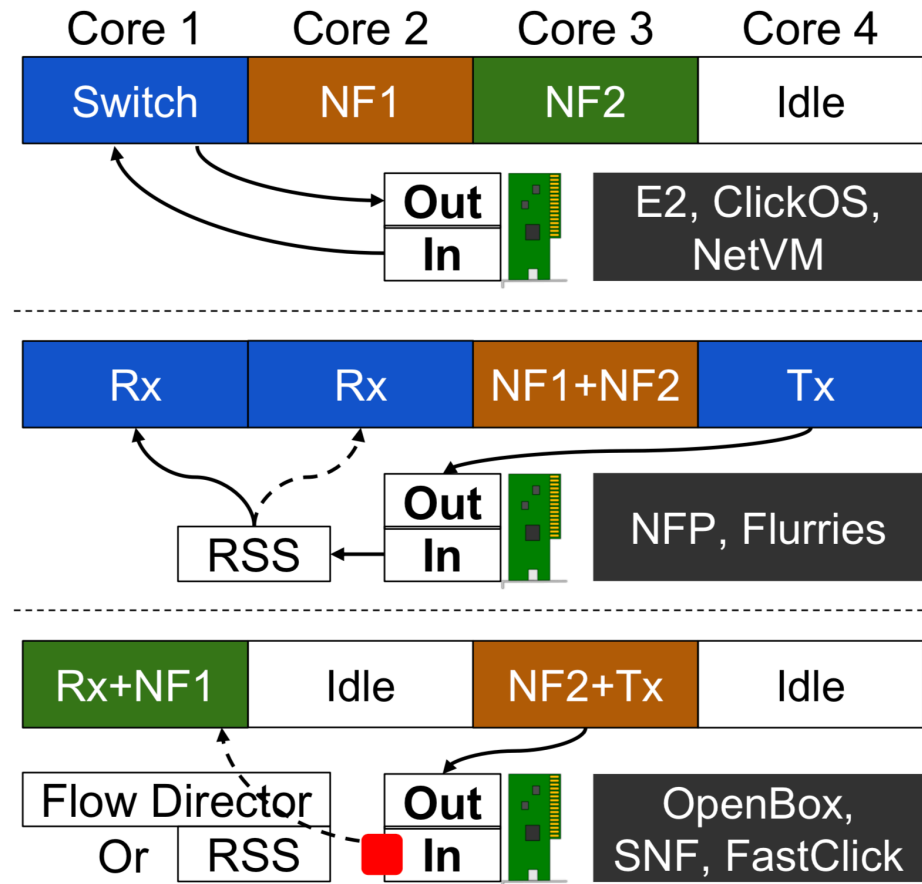
- Stateless read/write operation
- Tagging
- SDN → Dynamic tagging

Comments

- Globally shared state
  - Cross-core communication
  - Lock
- Scaling/Load Balancing
  - State migration

SafeBricks: Shielding Network Functions in the Cloud

- NetBricks
  - Rust
  - "Shared memory"
- Intel SGX
  - Enclave
  - IO optimization
- Compiler
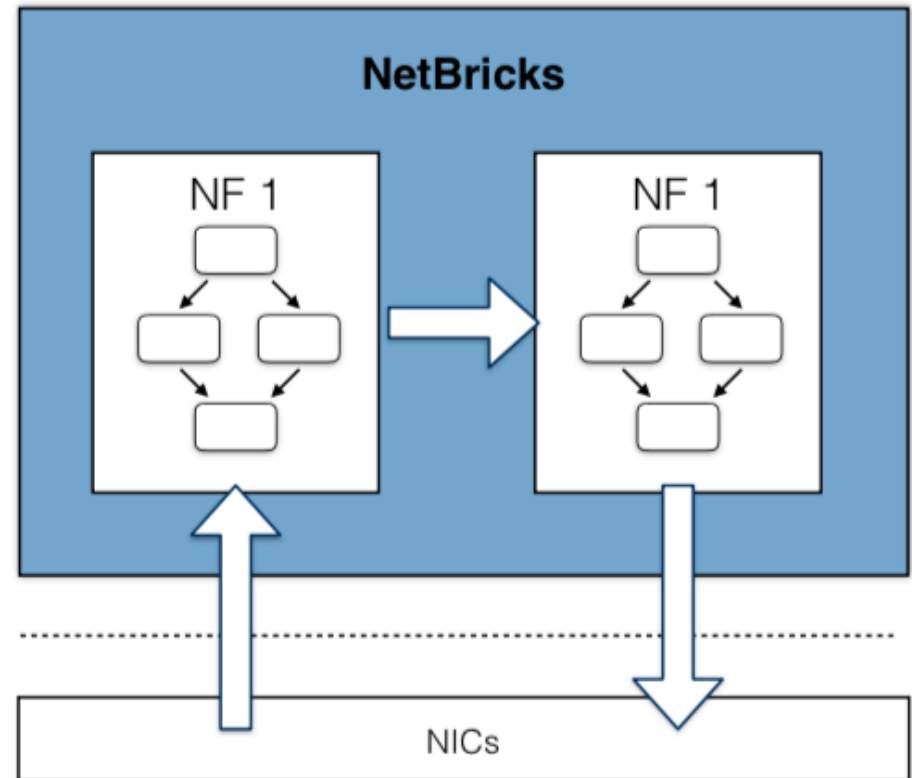
ResQ: Enabling SLOs in Network Function Virtualization

- SLO: service level object
  - Predictable performance & latency

- Last-Level Cache(LLC)
  - Intel Cache Allocation Technology
  - Contention

- Solution: control DMA buffer
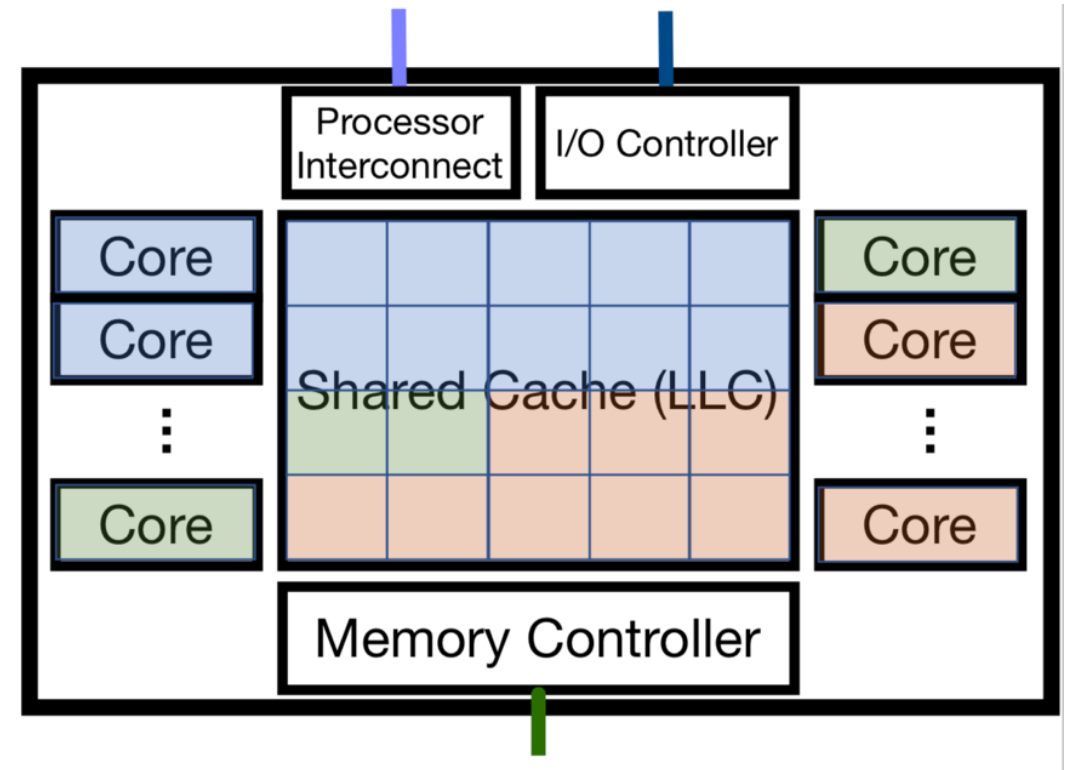
mTCP: A Highly Scalable User-level TCP Stack for Multicore Systems

- Linux TCP stack
  - 缺乏连接局部化
  - 共享的文件描述符空间
  - per-packet处理
  - 系统调用开销

mTCP

- User-level
- RSS
- Local socket
- DPDK
- Batch processing
- Lock-free queue (single consumer, single producer)
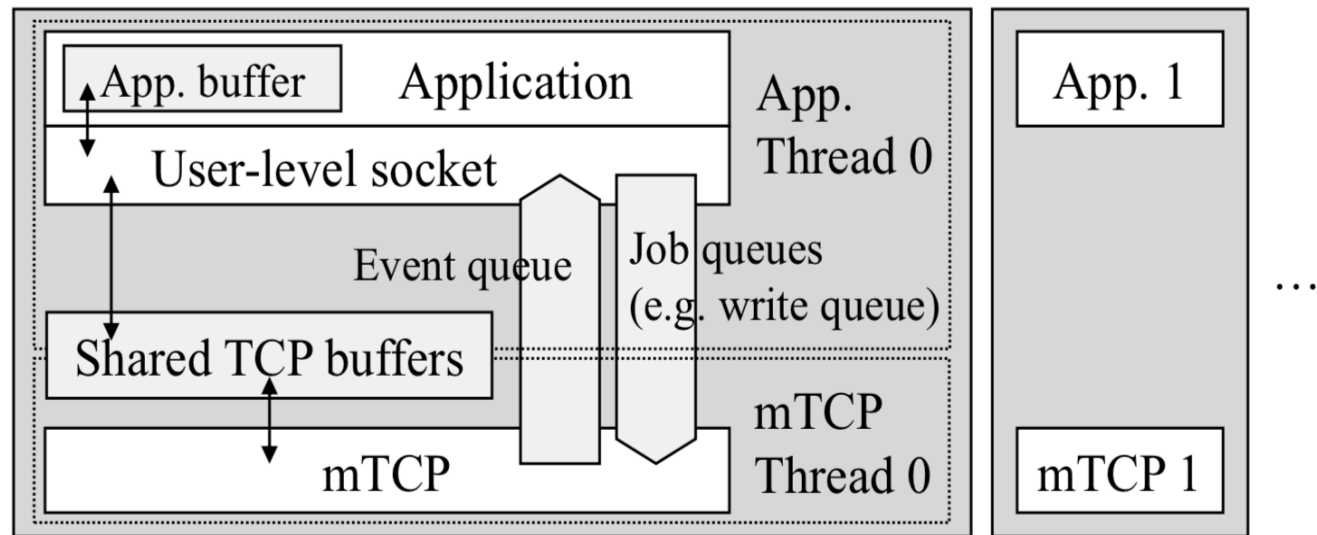

Figure 4: Thread model of mTCP.

Limitations:
1. 内核实现的协议/模块
2. 使用buffer会使copy开销大

FastSocket

- BSD socket
  - Local listen table
  - Local established table
  - VFS

- RFD+hash
  - 流量不均衡

mTCP实验

- PSIO
  - Linux kernel 2.6.*
- DPDK
  - mTCP-dedicated DPDK
  - NIC support
- netmap

https://github.com/fjzzhongyi/mtcp/blob/master/HHY.md

mTCP实验

- Config
  - VMware
  - e1000 virtual NIC
  - Ubuntu 14/16
  - LAN
- Setup
  - Hugepage
  - igb module
  - Configure NIC
    - no ping-pong
  - wget 64B
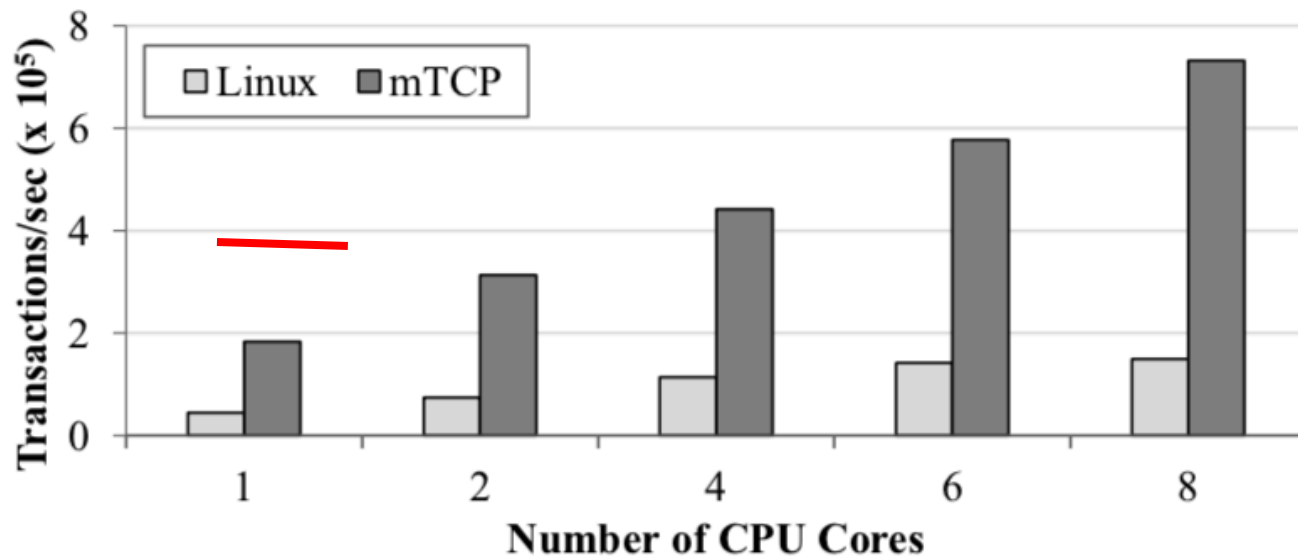
t=0

Figure 10: Performance of ab as a function of the number of cores. The file size is 64B and 8K concurrent connections are used.

https://github.com/fjzzhongyi/mtcp/blob/master/HHY.md

HyperKernel实验

- Ubuntu16.04, 18.04
  - make hv6-verify failed!

- Docker
  - spec.py → sys_close

```
# decrement file refcnt
new.files[fn].refcnt[(pid, fd)]
```

```
===========================================================================
FAIL: test_sys_close (__main__.HV6)
---------------------------------------------------------------------------
Traceback (most recent call last):
  File "o.x86_64/hv6/hv6py/kernel/spec/main.py", line 410, in <lambda>
    self._syscall_generic(syscall)
  File "o.x86_64/hv6/hv6py/kernel/spec/main.py", line 545, in _syscall_generic
    cond, newstate = getattr(spec, name)(self.state, *args)
  File "/home/hv6/o.x86_64/hv6/hv6py/kernel/spec/spec/specs.py", line 794, in sys_close
    new.files[fn].refcnt[(pid, fd)] -= 0
  File "/home/hv6/o.x86_64/hv6/hv6py/base.py", line 60, in __sub__
    return self._imap.__isub__(self._instance, value)
  File "/home/hv6/o.x86_64/hv6/hv6py/base.py", line 115, in __isub__
    return self.isub(instance._fields[self._name], instance._path, value)
  File "/home/hv6/o.x86_64/hv6/hv6py/base.py", line 627, in isub
    assert value == 1
AssertionError

---------------------------------------------------------------------------
Ran 1 test in 8.456s

FAILED (failures=1)
hv6/Makefrag:169: recipe for target 'hv6-verify' failed
make: *** [hv6-verify] Error 1
```