The CASE of FEMU: Cheap, Accurate, Scalable and Extensible Flash Emulator

Huaicheng Li, Mingzhe Hao, Michael Hao Tong,
Swaminatahan Sundararaman*, Matias Bjørling+, Haryadi S. Gunawi

THE UNIVERSITY OF CHICAGO

Parallel M *

CNEXLABS +

What SSD platforms are used?

What SSD platforms are used?

Simulator

Emulator

Hardware Platform

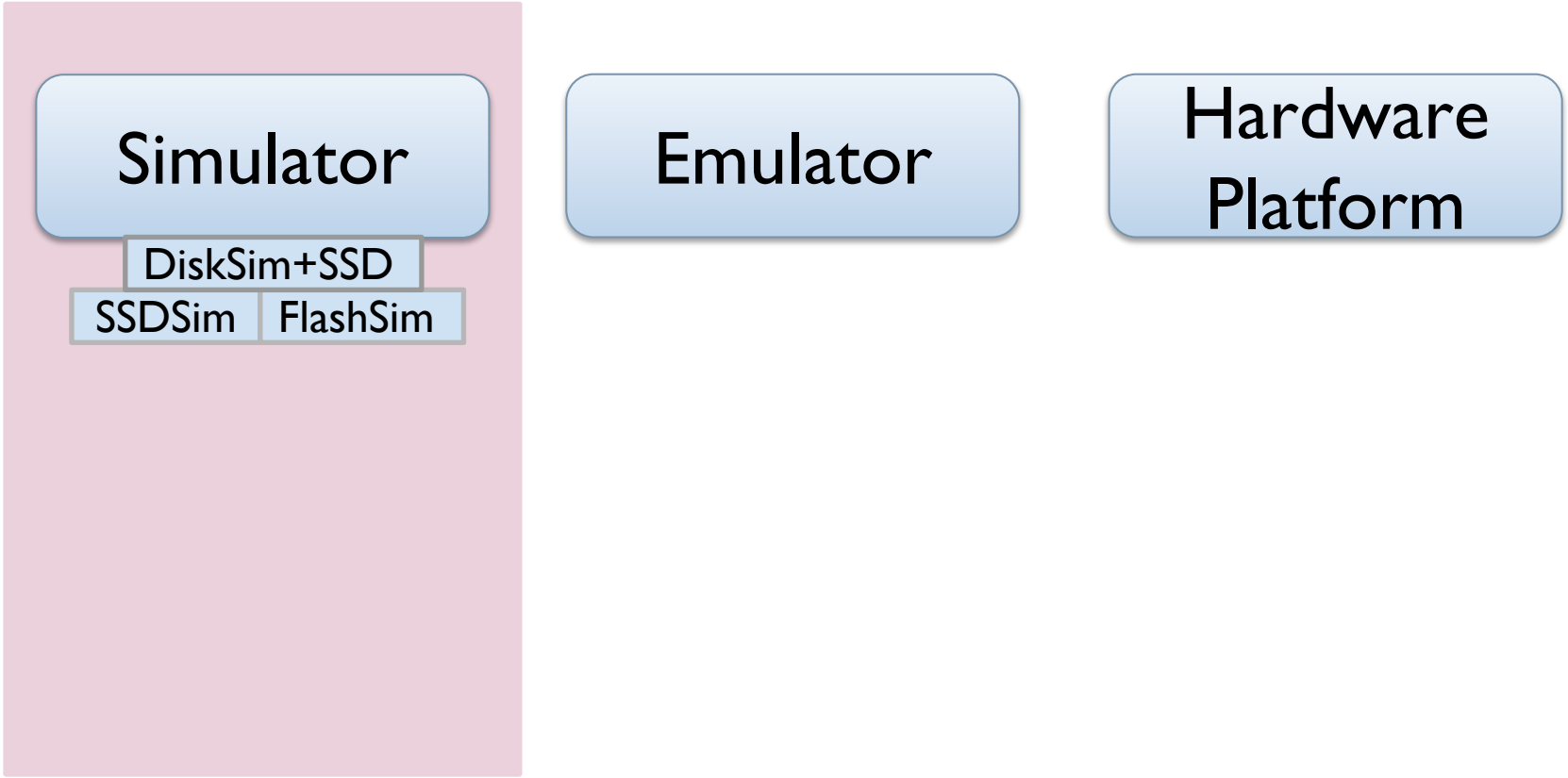
What SSD platforms are used?

Simulator

Emulator

Hardware Platform

What SSD platforms are used?



What SSD platforms are used?

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving

Emulator

Hardware Platform

What SSD platforms are used?

Simulator

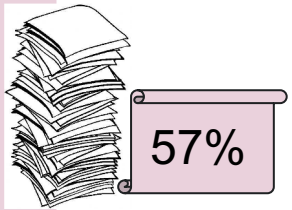
DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving

Emulator

Hardware Platform



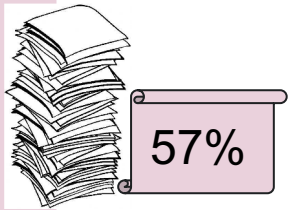
What SSD platforms are used?

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only



Emulator

Hardware Platform

What SSD platforms are used?

Simulator

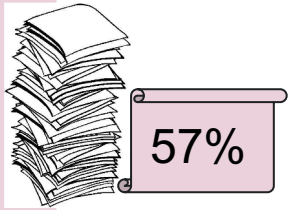
DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator

Hardware Platform



≠

Trends

- Software-Defined Flash
- Split-Level Architecture

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator

Hardware Platform

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator

Hardware Platform

OpenSSD OpenChannel-SSD

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator

Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate

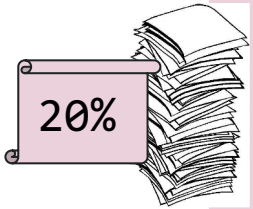
Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator



Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate

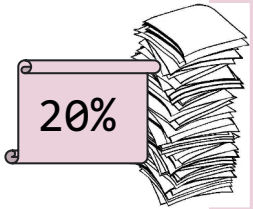
Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator



Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

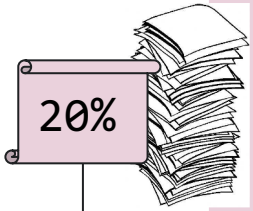
Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator



19% Single SSD

Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

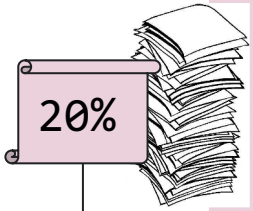
Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator



19% Single SSD

1% Distributed SSDs

Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator

Hardware Platform


OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator

LightNVM's QEMU

VSSIM | FlashEm

Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Emulator


Guest OS

LightNVM's QEMU

VSSIM | FlashEm

- Fullstack Research
- Cheap

Hardware Platform


OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

Simulator

DiskSim+SSD

SSDSim | FlashSim

- Simple
- Time-saving
- Trace driven
- Internal-research only

Guest OS

Emulator

LightNVM's QEMU

VSSIM | FlashEm

- Fullstack Research
- Cheap
- Poor Scalability
- Poor Accuracy

Hardware Platform

OpenSSD OpenChannel-SSD

- Full-stack Research
- Accurate
- Expensive
- Complex to use
- Wear-out

The "*CASE*" of FEMU

FEMU: QEMU/Software based Flash Emulator

The "*CASE*" of FEMU

FEMU: QEMU/Software based Flash Emulator

- **C**heap: $0, https://github.com/ucare-uchicago/femu

The "*CASE*" of FEMU

FEMU: QEMU/Software based Flash Emulator

- **C**heap: $0, https://github.com/ucare-uchicago/femu
- **A**ccurate: 0.5-38% error rate in latency
  - 11% average at microsecond level

The "*CASE*" of FEMU

FEMU: QEMU/Software based Flash Emulator

- **C**heap: $0, https://github.com/ucare-uchicago/femu
- **A**ccurate: 0.5-38% error rate in latency
  - 11% average at microsecond level
- **S**calable: support 32 channels/chips

The "*CASE*" of FEMU

FEMU: QEMU/Software based Flash Emulator

- **C**heap: $0, https://github.com/ucare-uchicago/femu
- **A**ccurate: 0.5-38% error rate in latency
  - 11% average at microsecond level
- **S**calable: support 32 channels/chips
- **E**xtensible
  - modifiable interface
  - modifiable FTL

What is FEMU?

Typical Fullstack Research

FEMU Fullstack Research

What is FEMU?

Typical Fullstack Research

App

Host OS

Hardware Platform

FEMU Fullstack Research

What is FEMU?

Typical Fullstack Research

App

Host OS

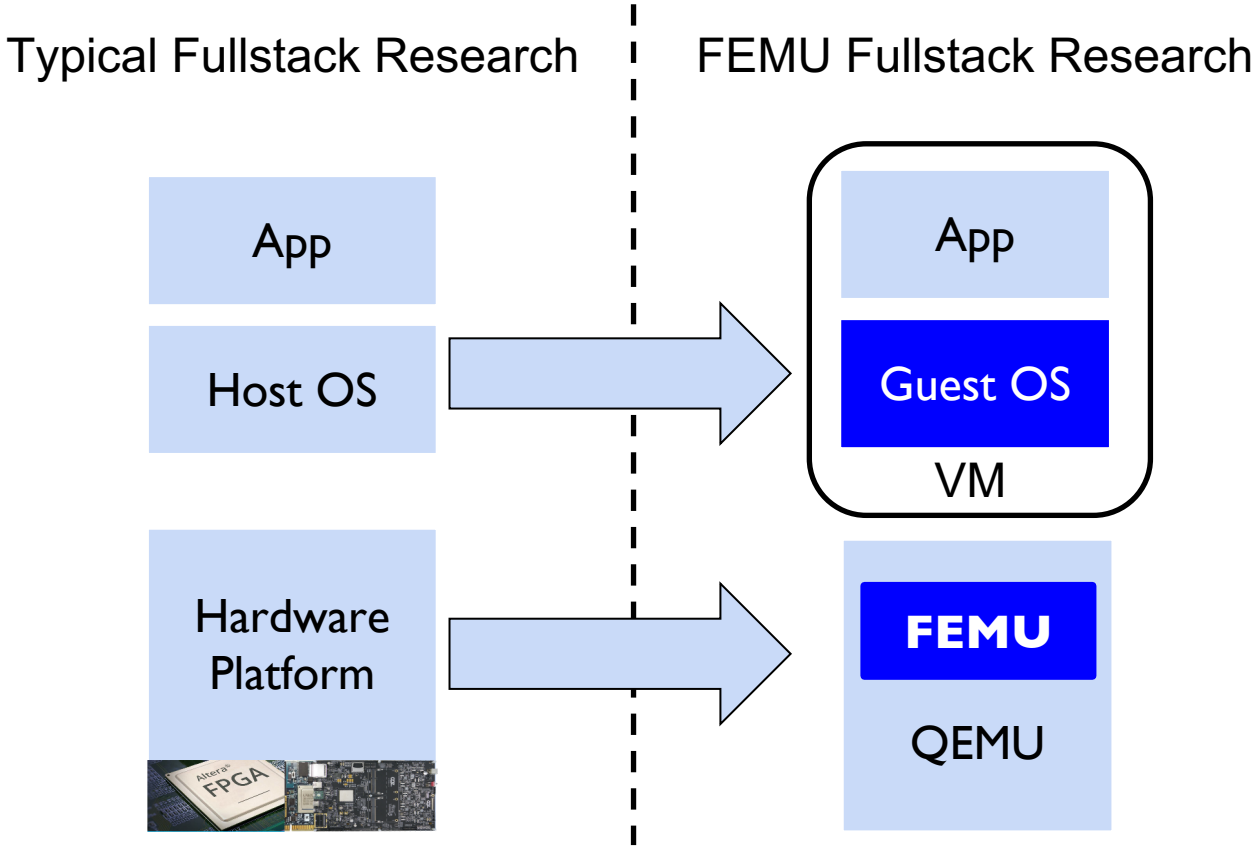
Hardware Platform

FEMU Fullstack Research

FEMU

QEMU

What is FEMU?


Typical Fullstack Research
FEMU Fullstack Research
App
Host OS
Hardware Platform
Altera FPGA
App
Guest OS
VM
FEMU
QEMU

What is FEMU?

Typical Fullstack Research

App

Host OS

Hardware Platform

FEMU Fullstack Research

App

Guest OS

VM

FEMU

QEMU

Supported research:

What is FEMU?

Typical Fullstack Research

App

Host OS

Hardware Platform

FEMU Fullstack Research

App

Guest OS

VM

FEMU

QEMU

FTL changes ✓

Supported research:

What is FEMU?

Typical Fullstack Research

App

Host OS

Hardware Platform

FEMU Fullstack Research

App

Guest OS

VM

NVMe

FEMU

QEMU

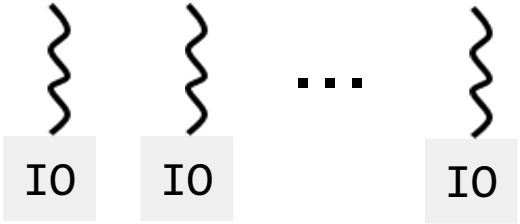
Supported research:

Kernel changes ✓

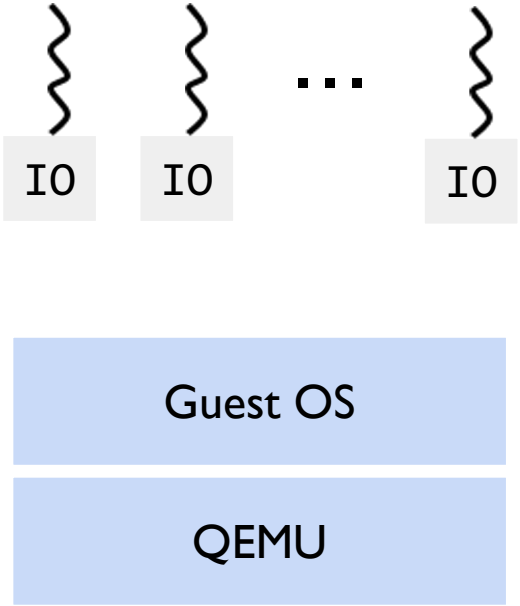
Interface changes ✓

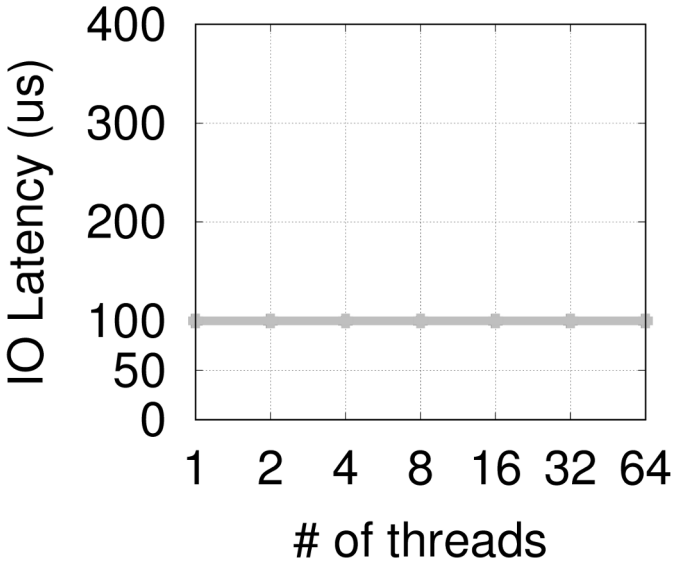
FTL changes ✓

QEMU Scalability

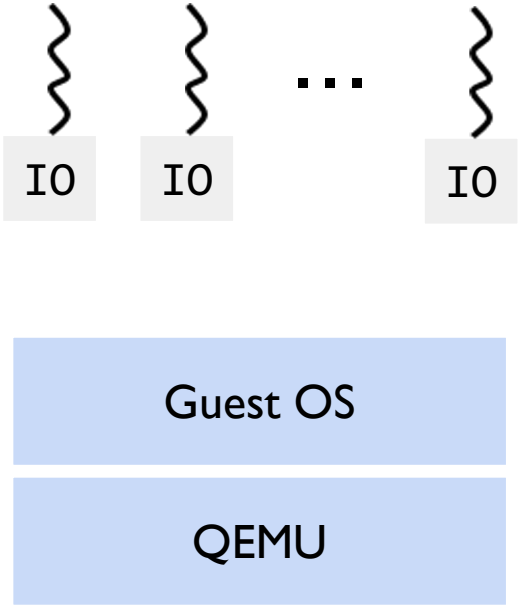


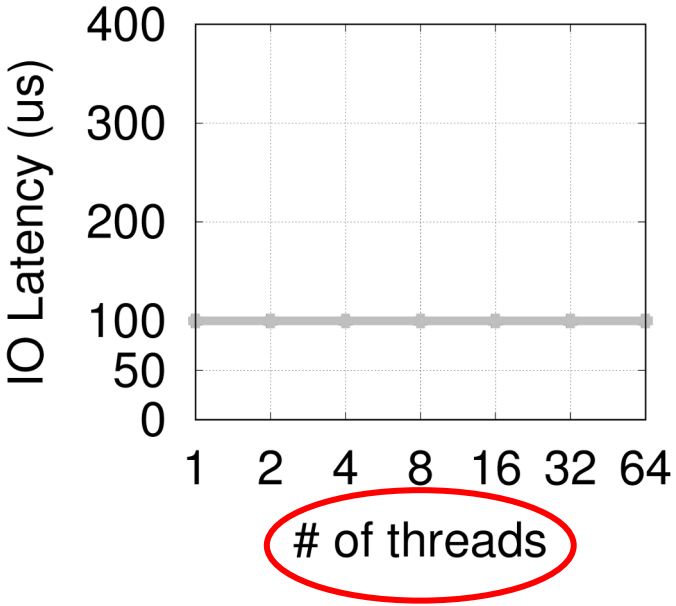
QEMU Scalability



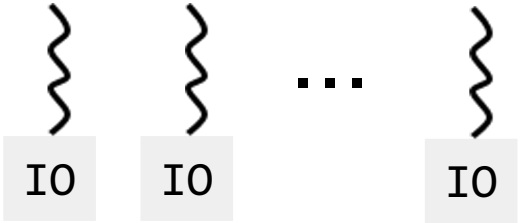


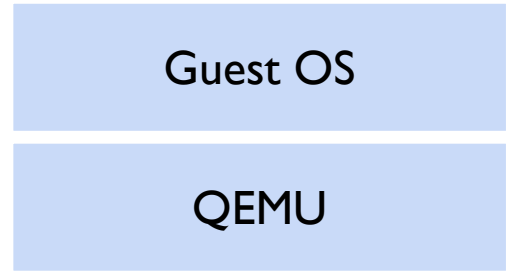
QEMU Scalability

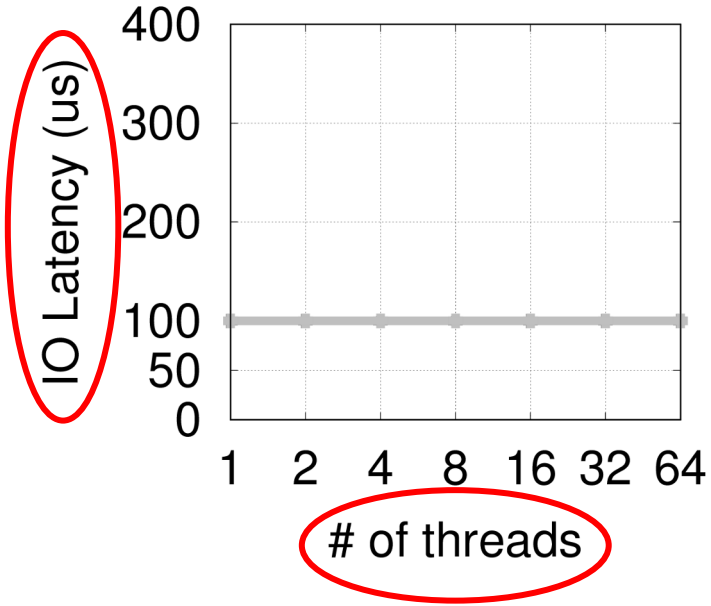




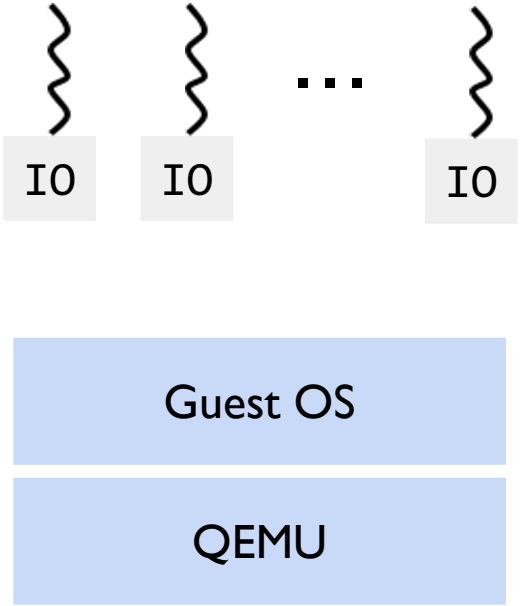
QEMU Scalability

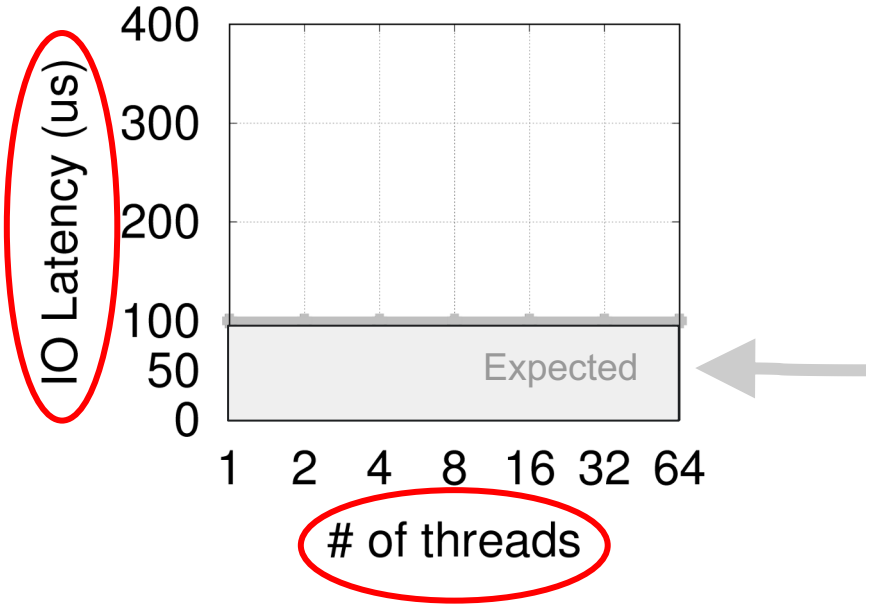




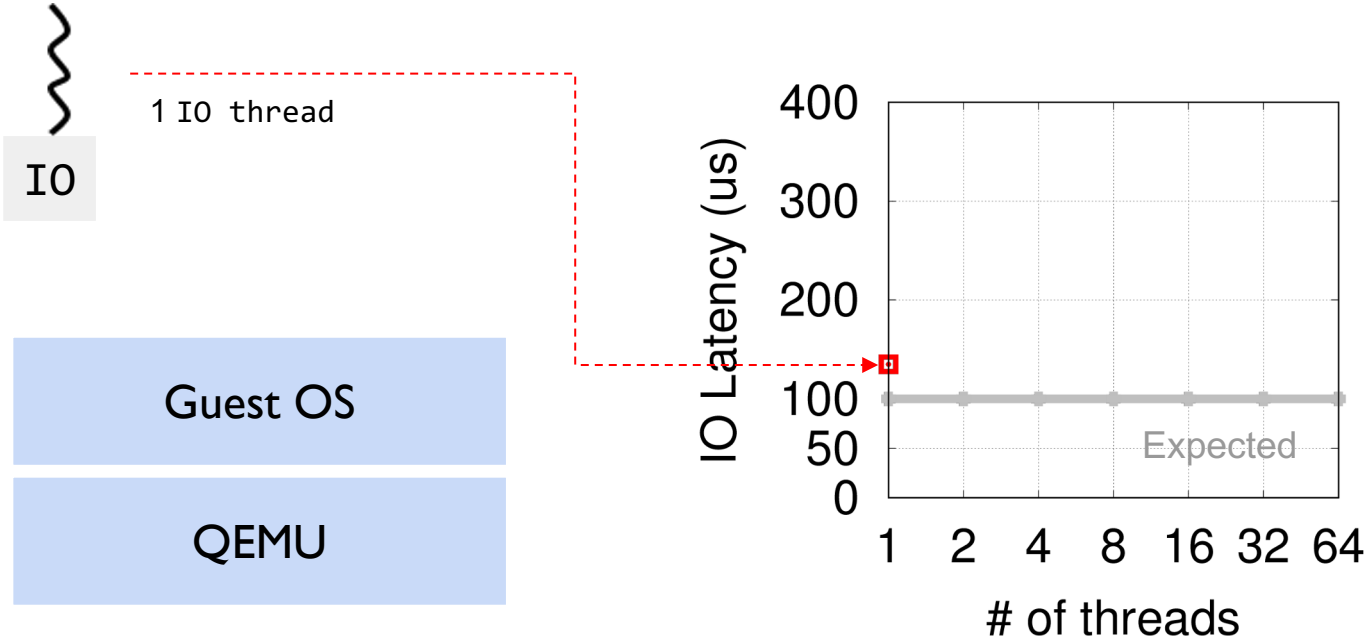


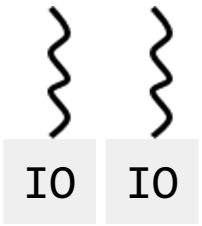
QEMU Scalability





QEMU IDE Scalability

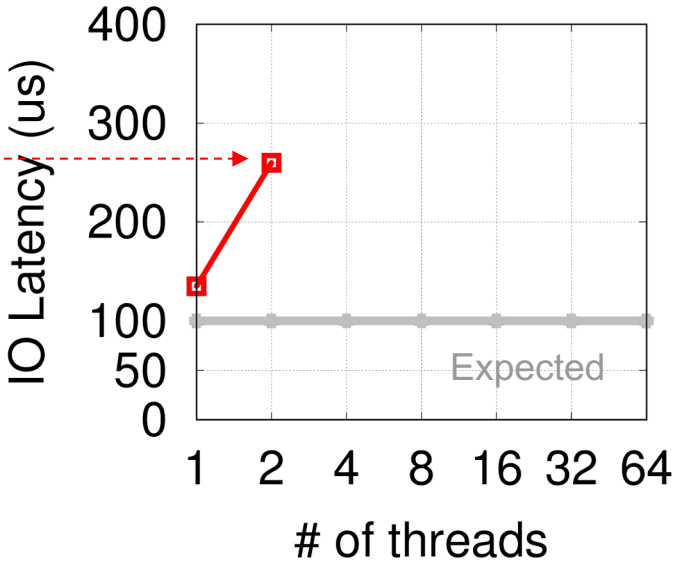


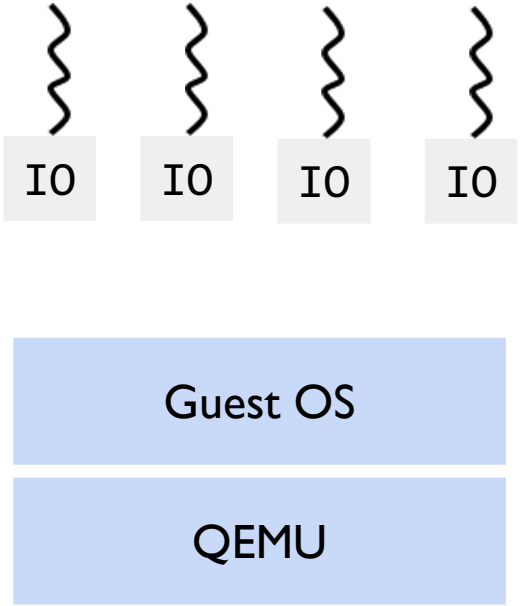


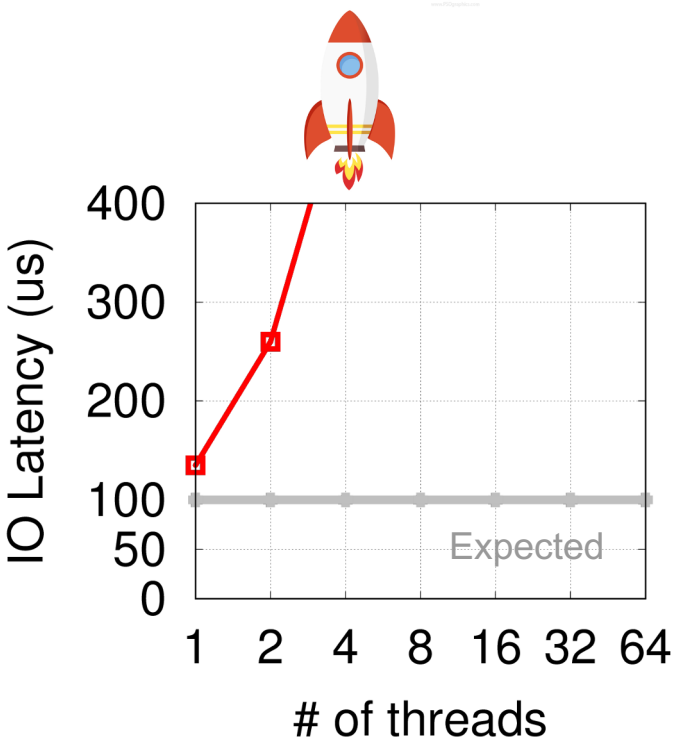
2 IO threads

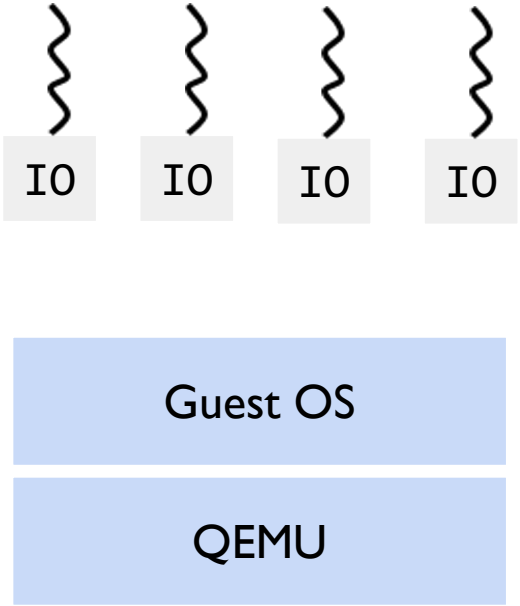
Guest OS

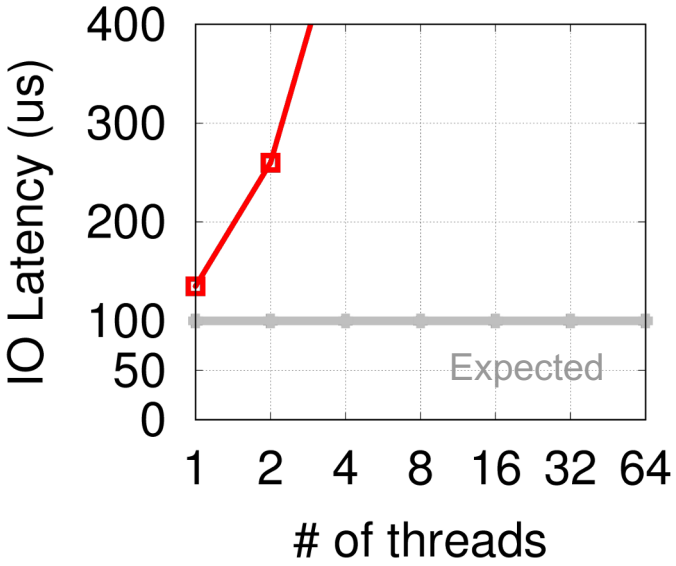
QEMU

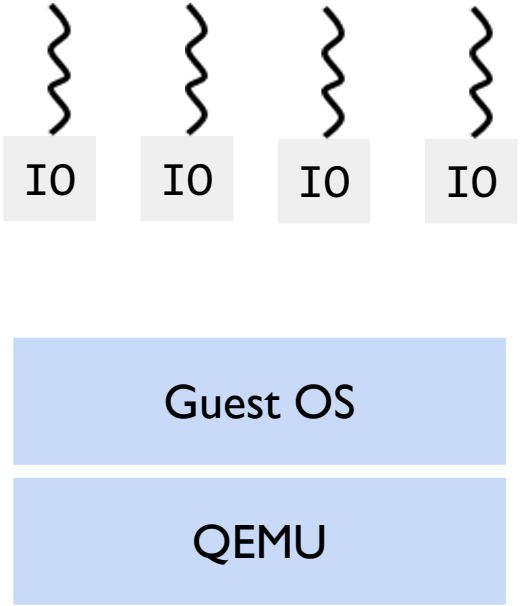


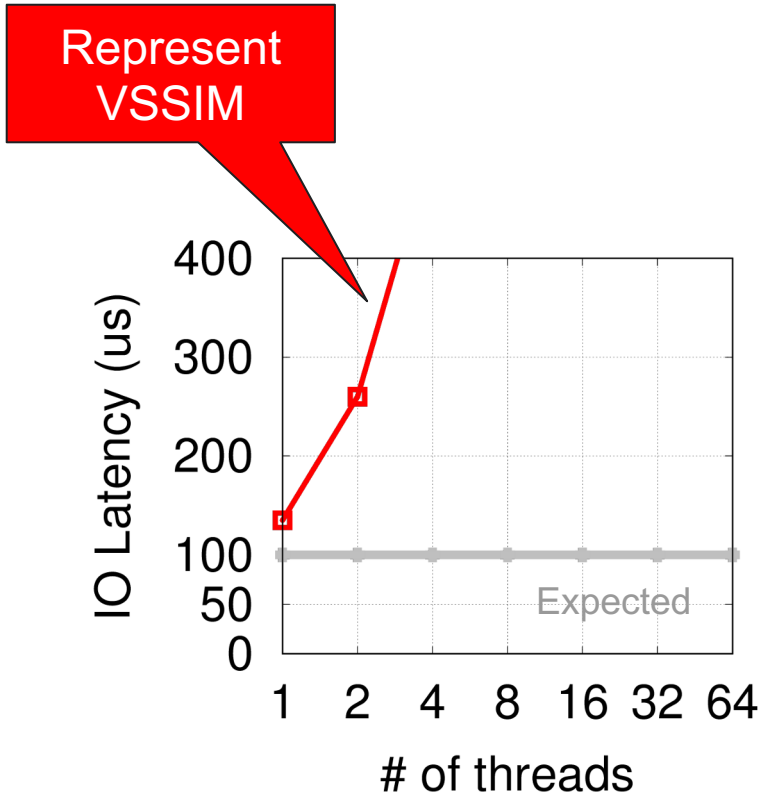

IO
IO
IO
IO
Guest OS
QEMU


IO Latency (us)
400
300
200
100
50
0
Expected
1
2
4
8
16
32
64
# of threads

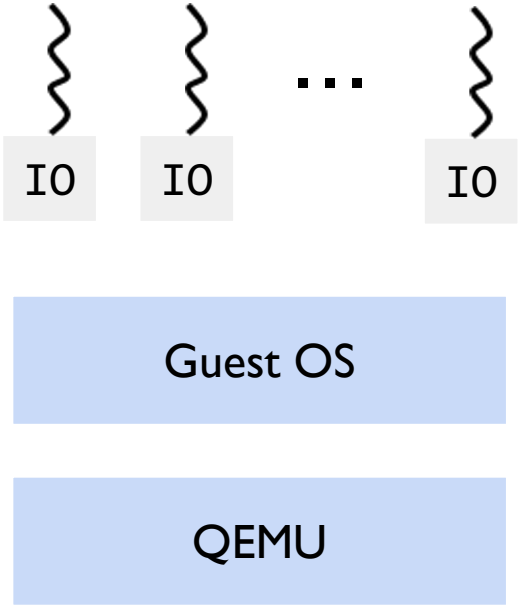

IO
IO
IO
IO
Guest OS
QEMU

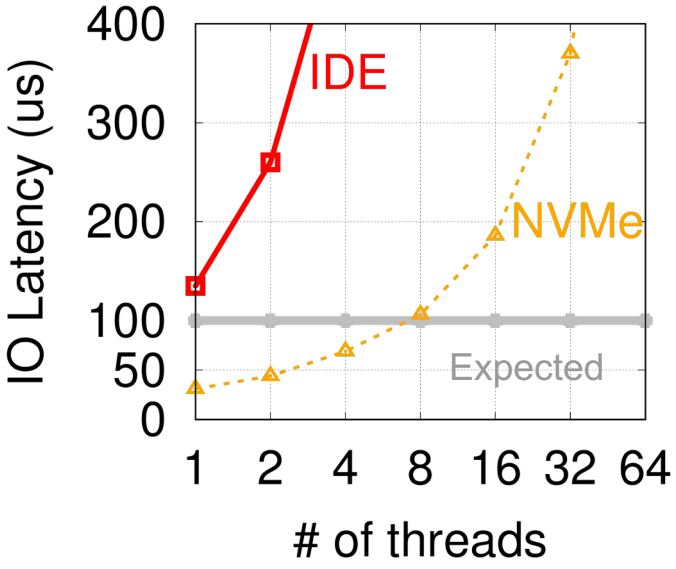

IO Latency (us)
400
300
200
100
50
0
Expected
1
2
4
8
16
32
64
# of threads


IO
IO
IO
IO
Guest OS
QEMU

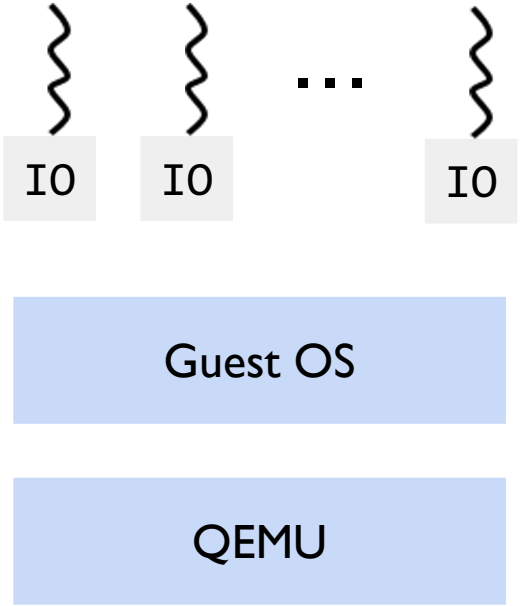

Represent VSSIM
IO Latency (us)
400
300
200
100
50
0
Expected
1
2
4
8
16
32
64
# of threads

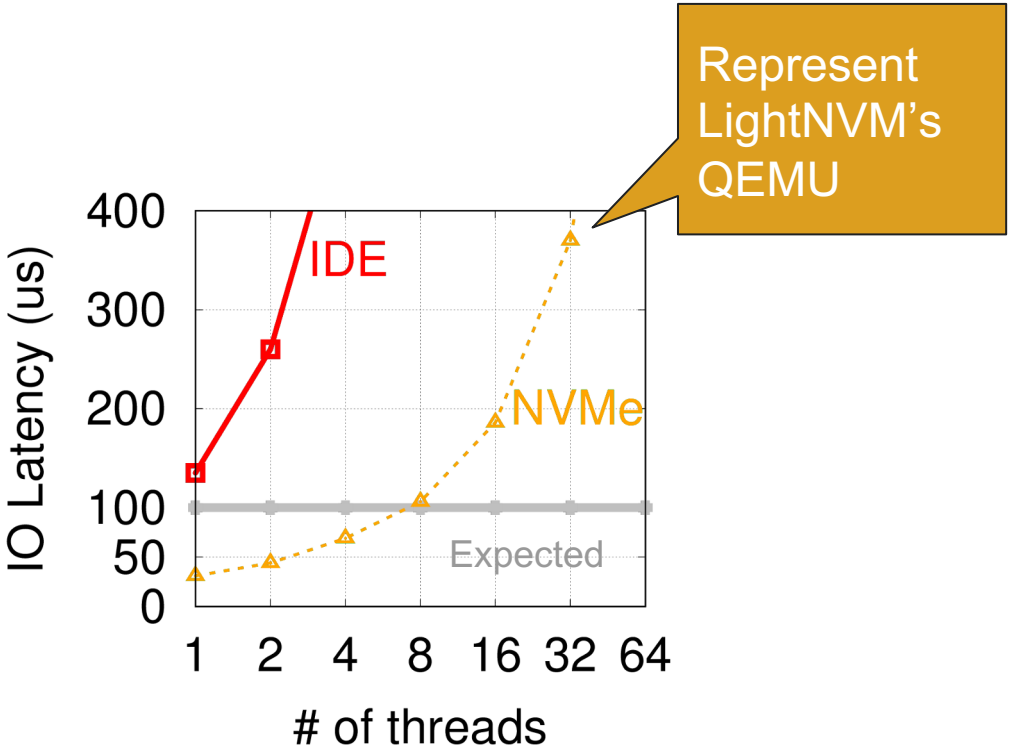
QEMU NVMe Scalability



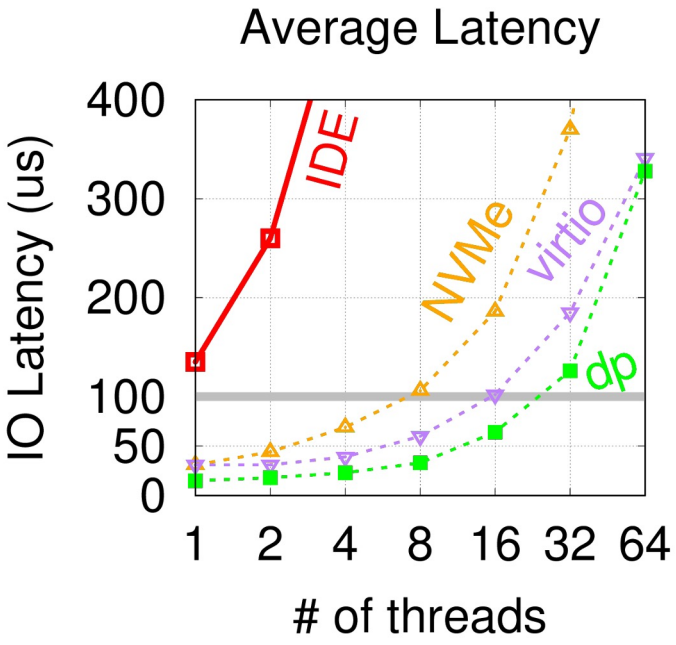


QEMU NVMe Scalability



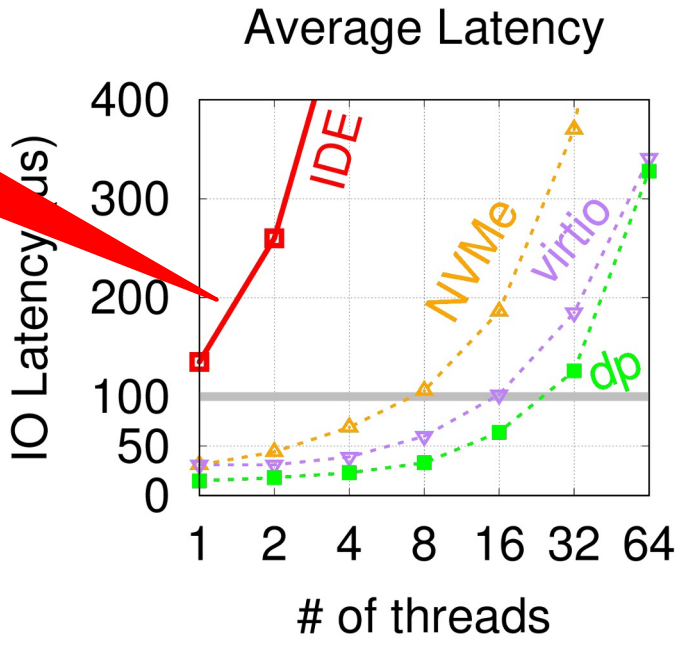


QEMU Scalability



QEMU Scalability

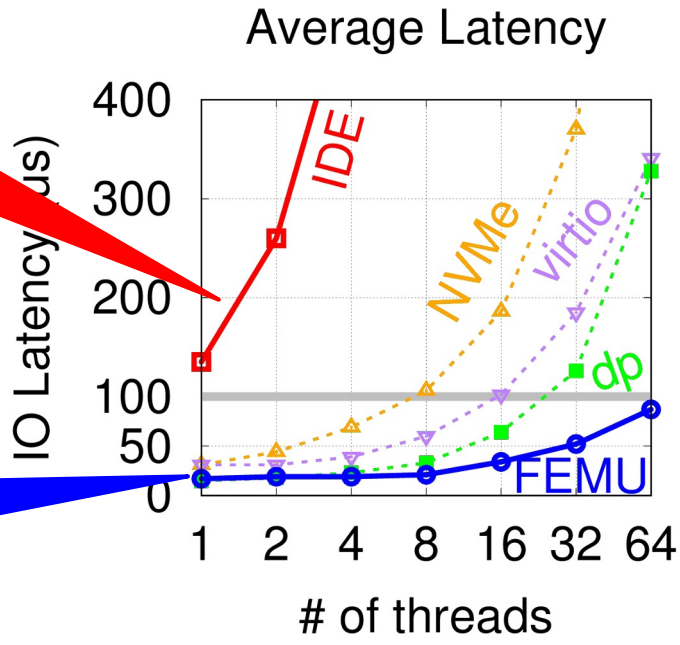
QEMU and existing emulators are NOT Scalable !



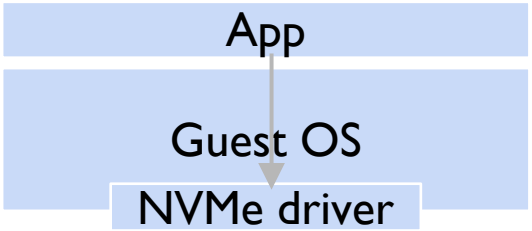
QEMU Scalability

QEMU and existing emulators are NOT Scalable !

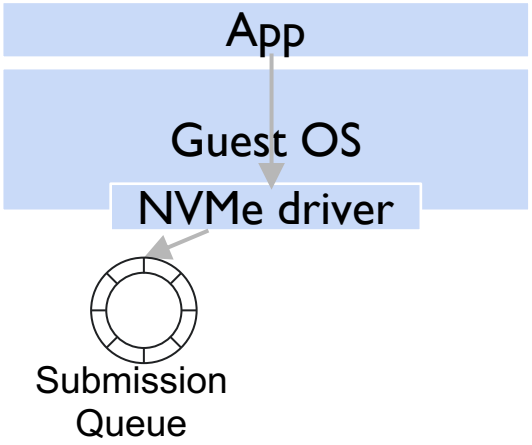
FEMU is Scalable !



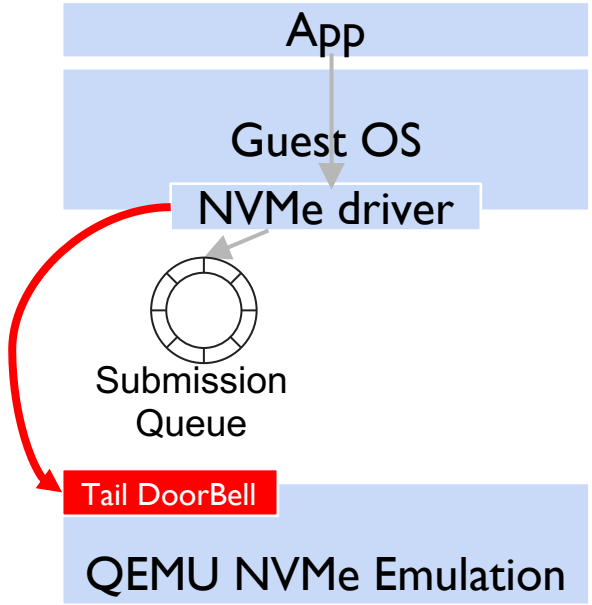
Scalability Root Causes & Solutions (1)



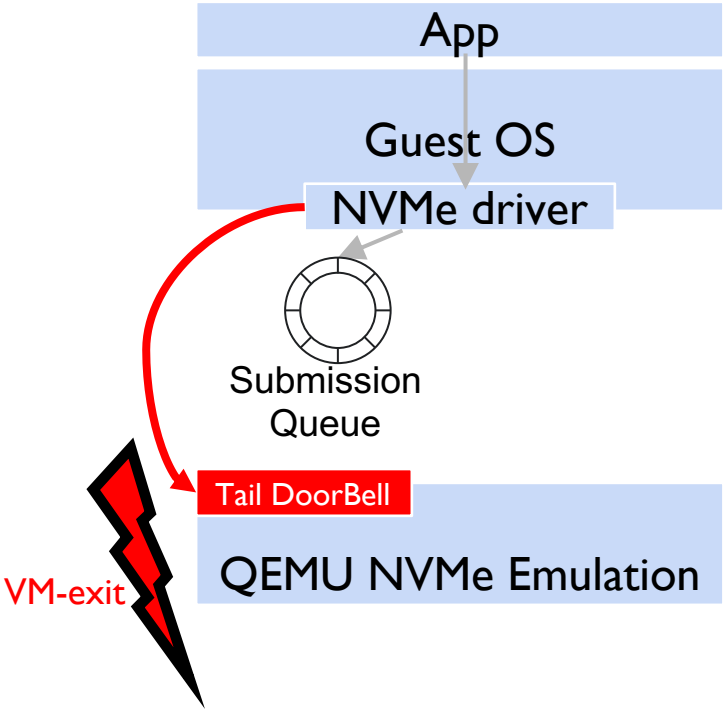
Scalability Root Causes & Solutions (1)



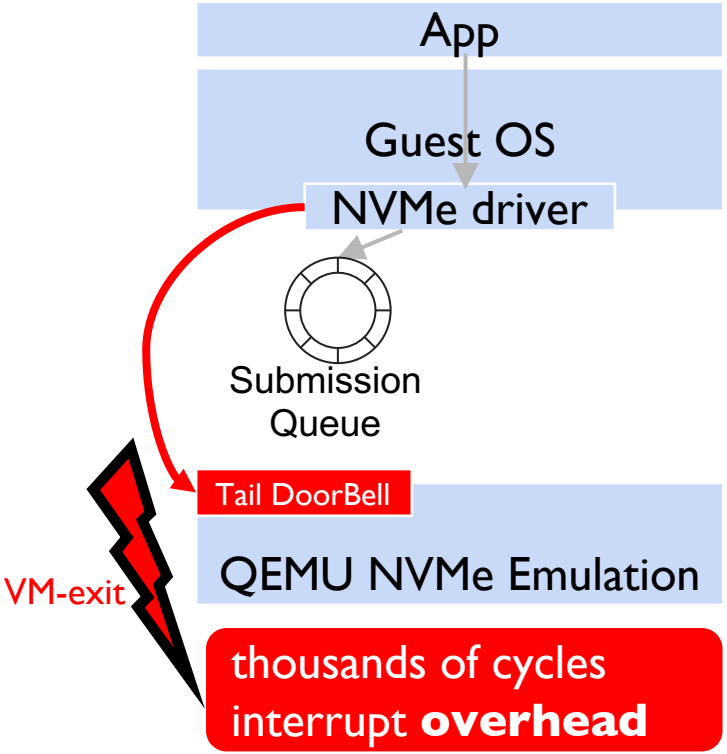
Scalability Root Causes & Solutions (1)



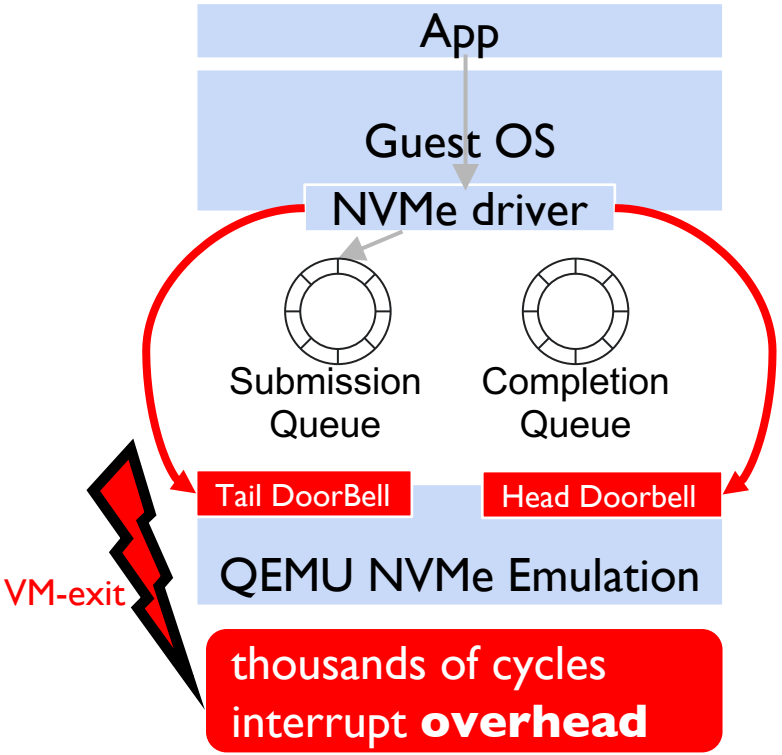
Scalability Root Causes & Solutions (1)



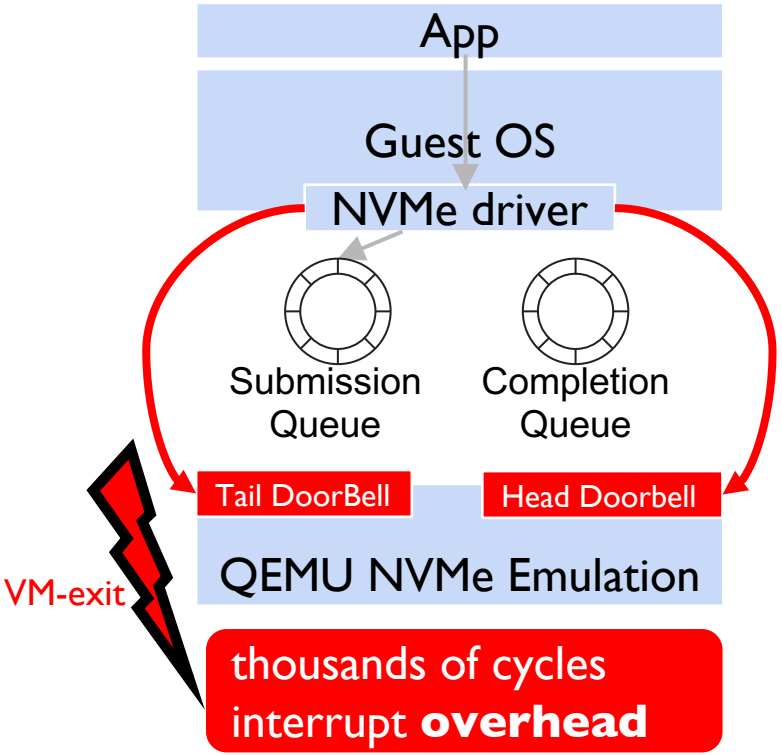
Scalability Root Causes & Solutions (1)

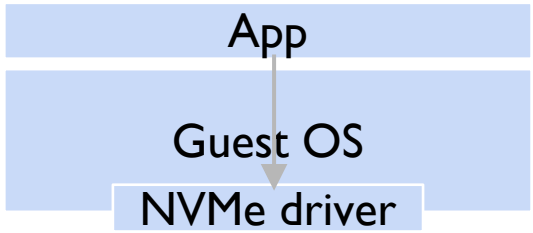


Scalability Root Causes & Solutions (1)

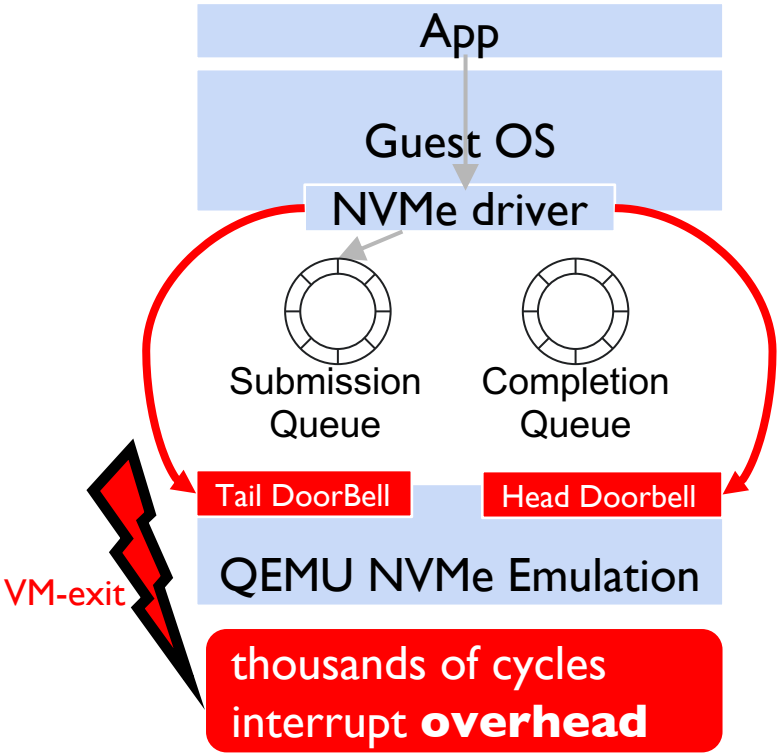


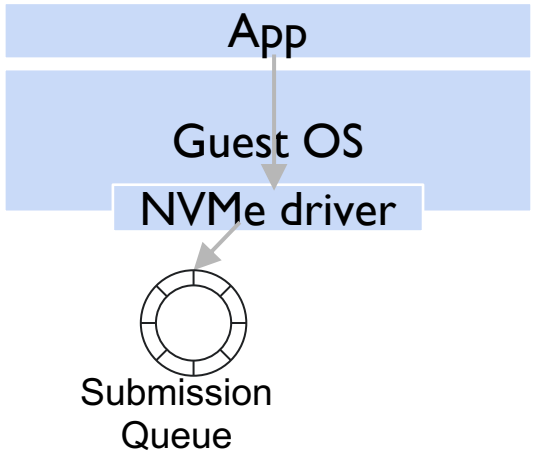
Scalability Root Causes & Solutions (1)



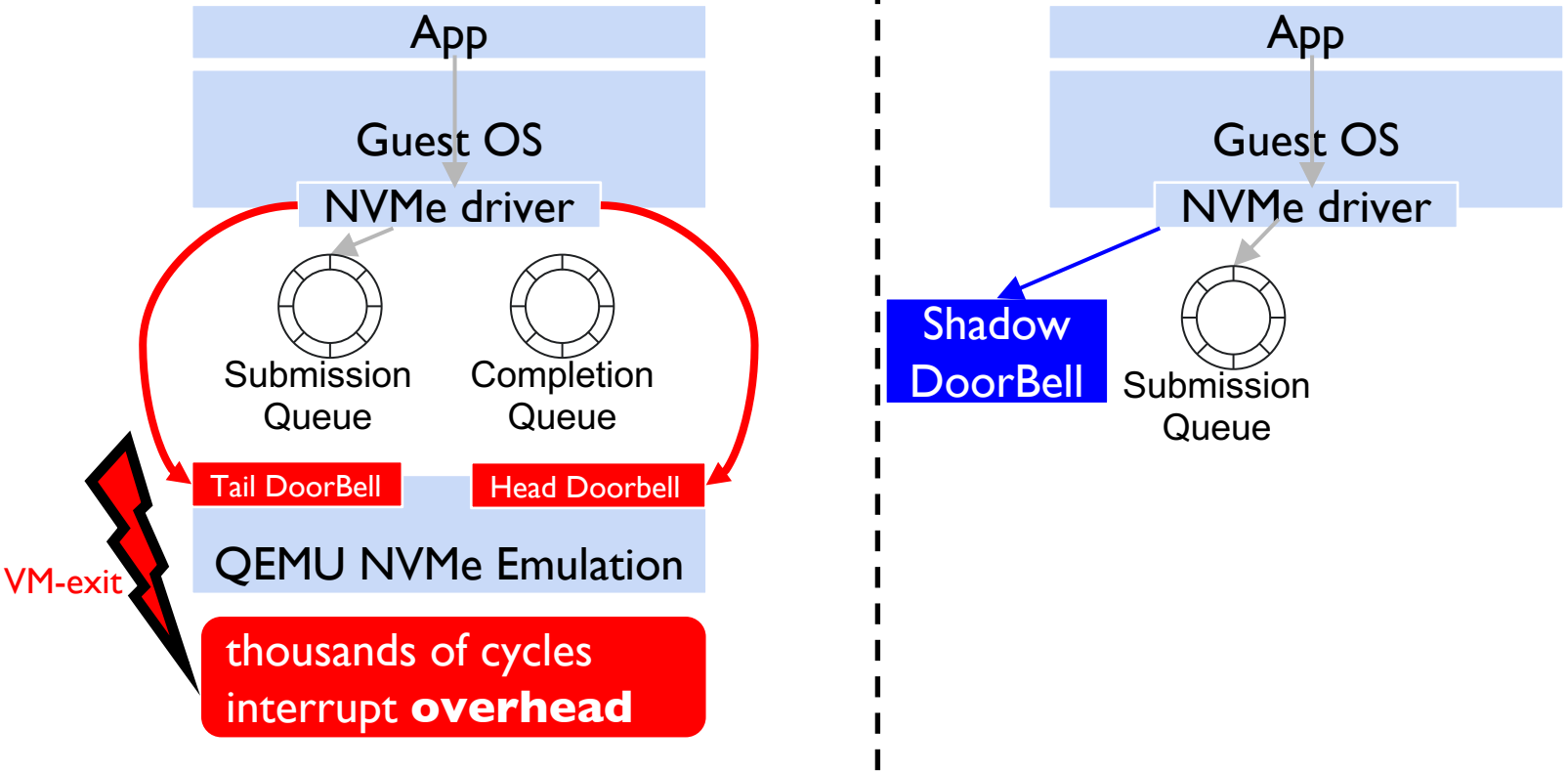


Scalability Root Causes & Solutions (1)

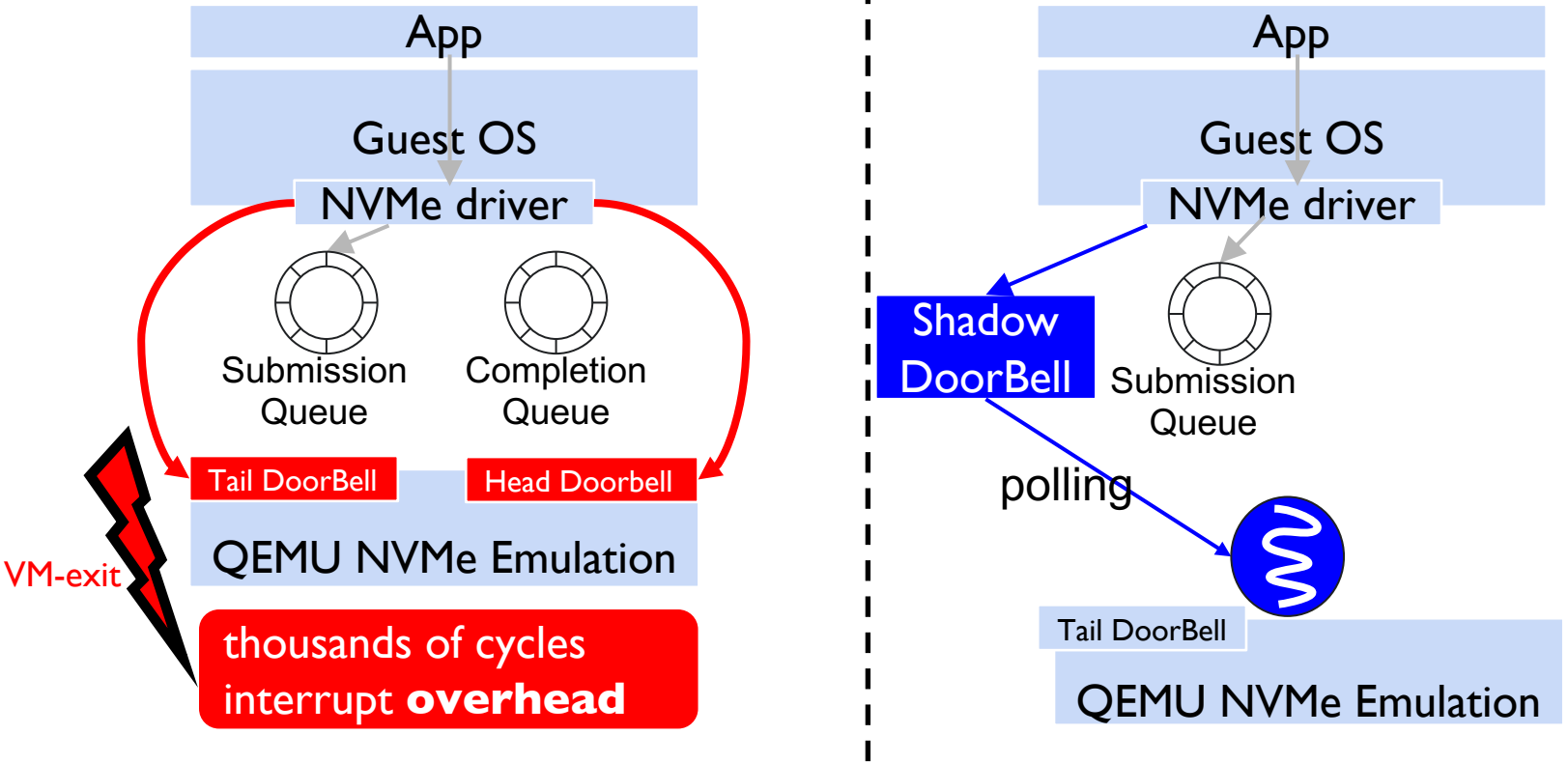




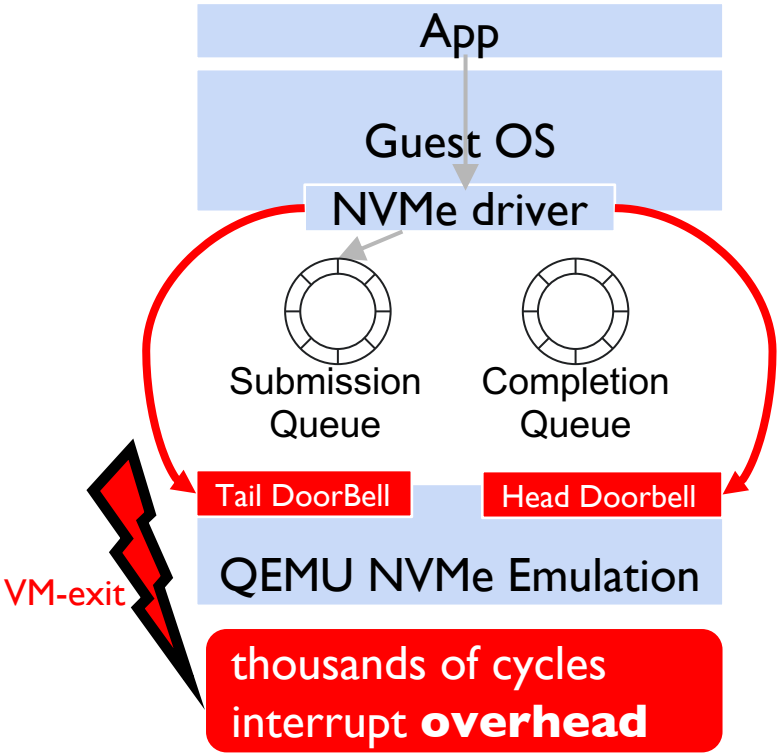
Scalability Root Causes & Solutions (1)

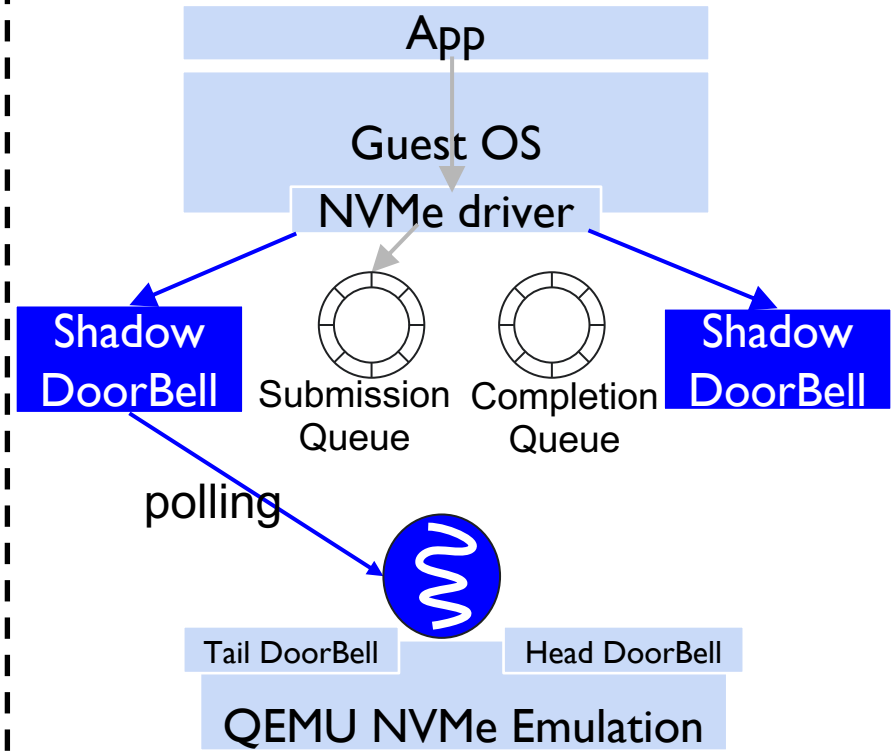


Scalability Root Causes & Solutions (1)

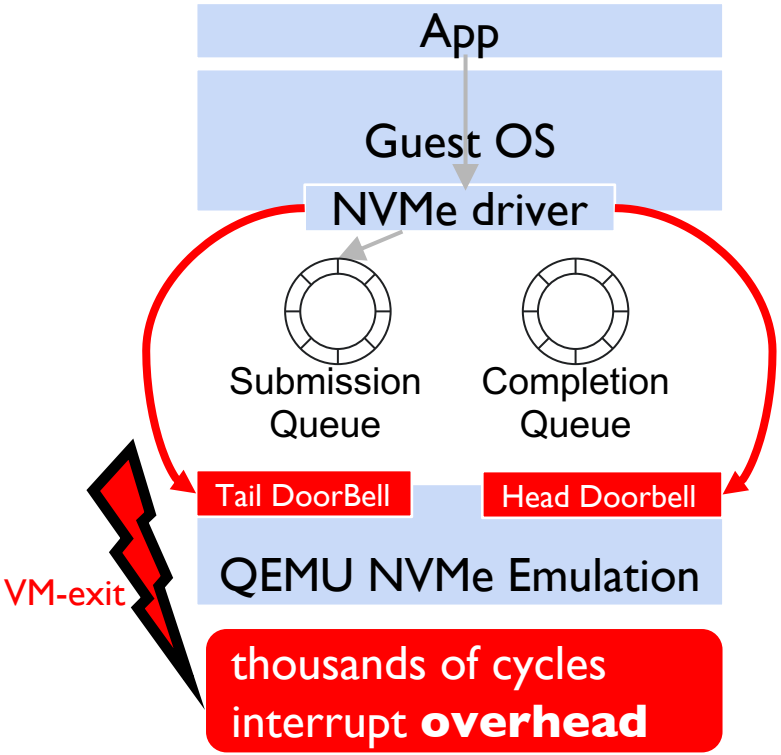


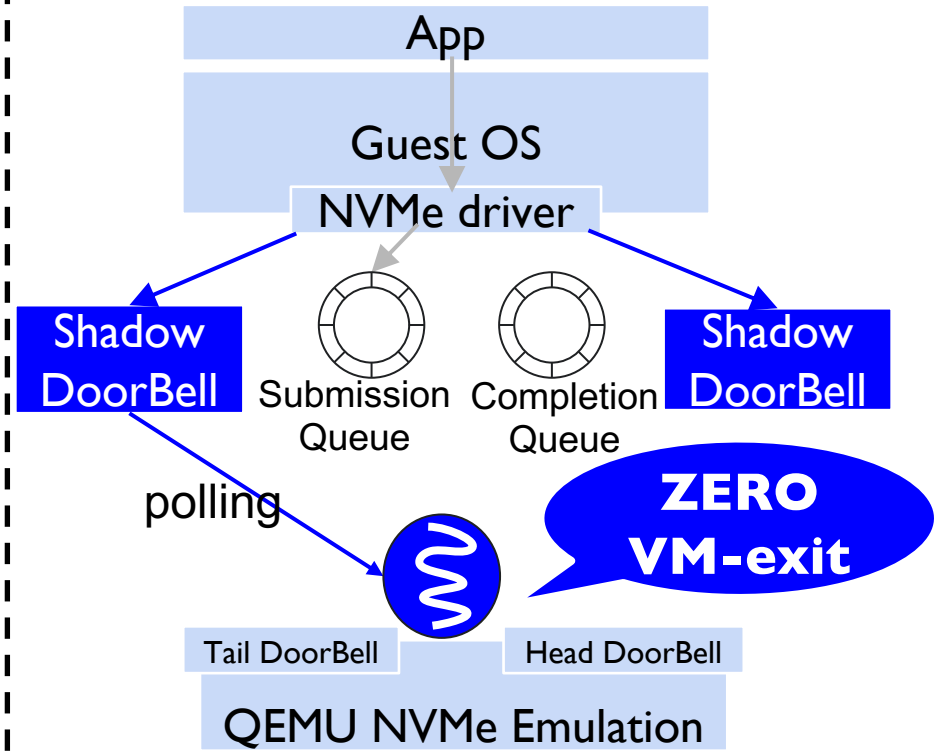
Scalability Root Causes & Solutions (1)





Scalability Root Causes & Solutions (1)

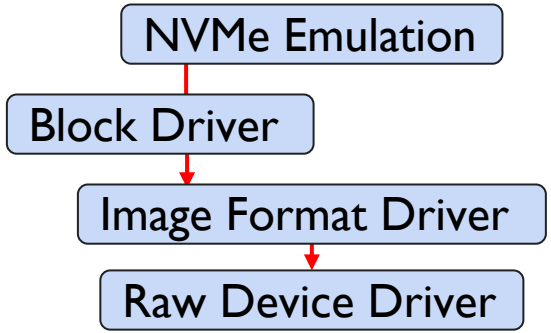




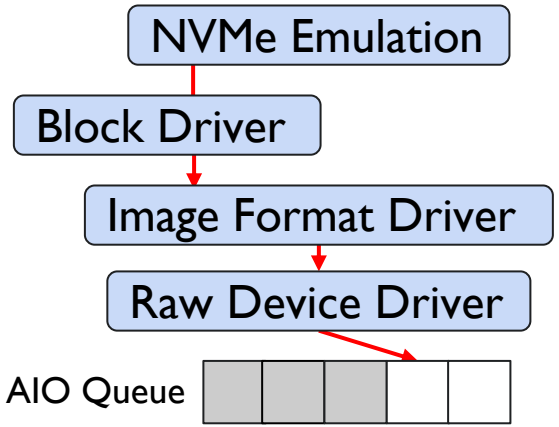
Scalability Root Causes & Solutions (2)

NVMe Emulation

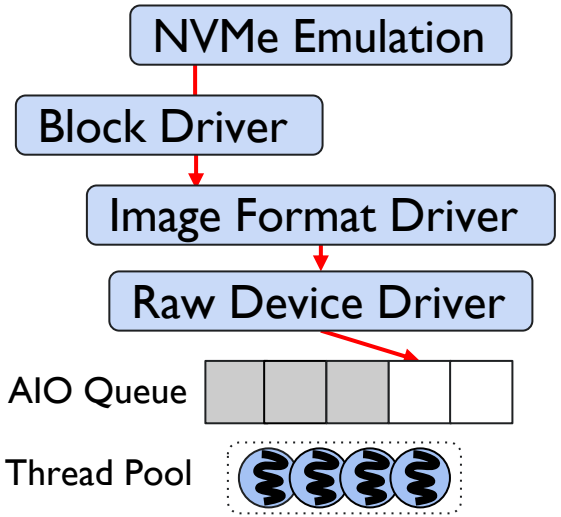
Scalability Root Causes & Solutions (2)



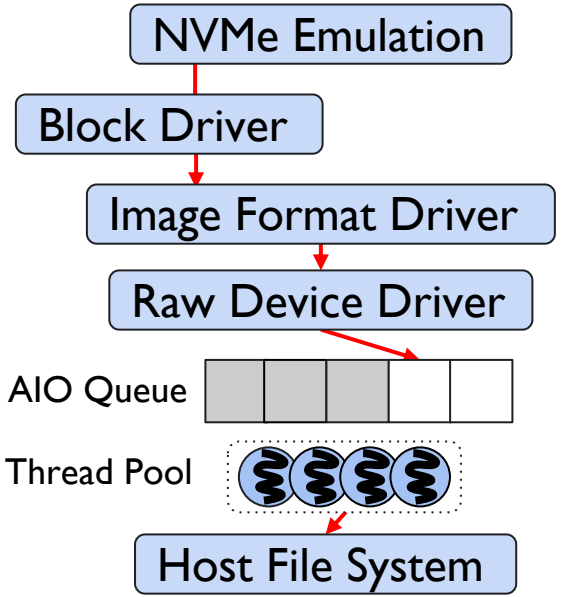
Scalability Root Causes & Solutions (2)



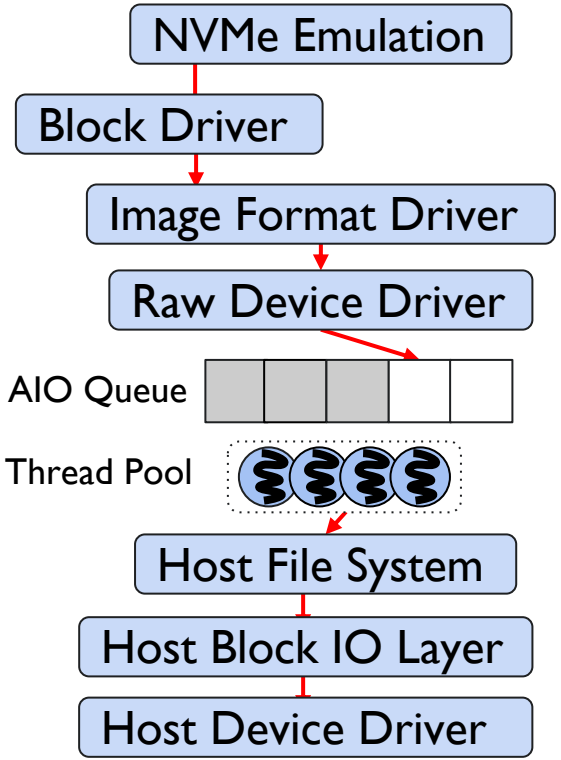
Scalability Root Causes & Solutions (2)



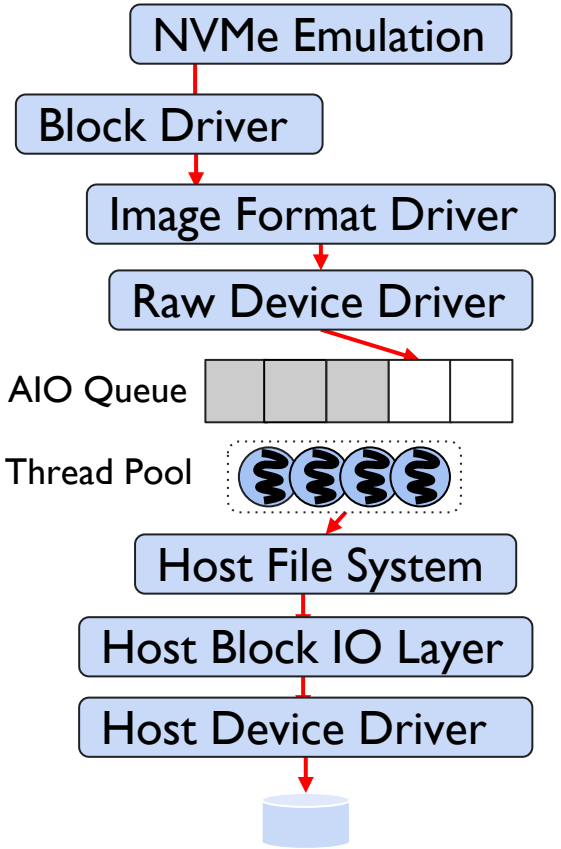
Scalability Root Causes & Solutions (2)



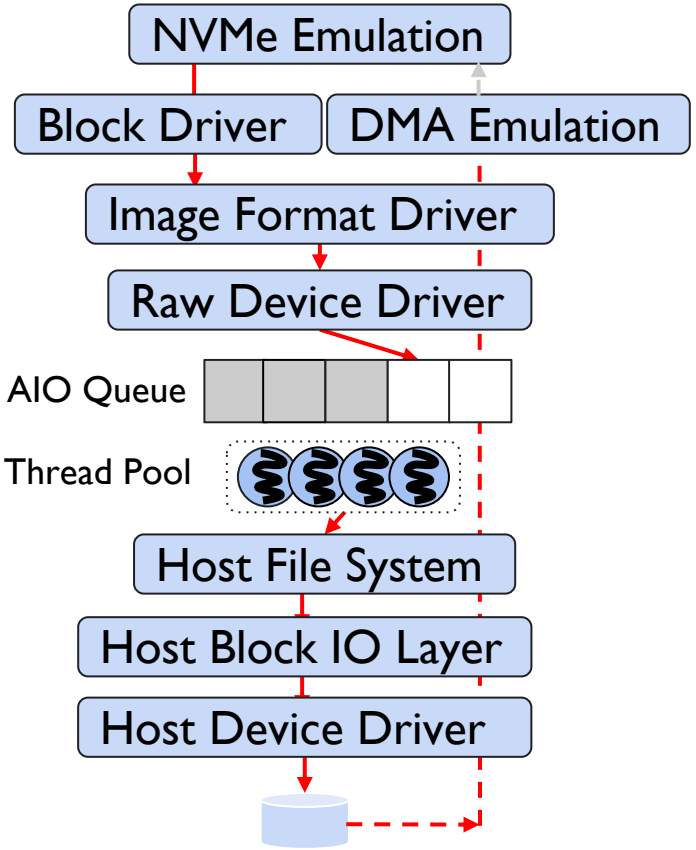
Scalability Root Causes & Solutions (2)



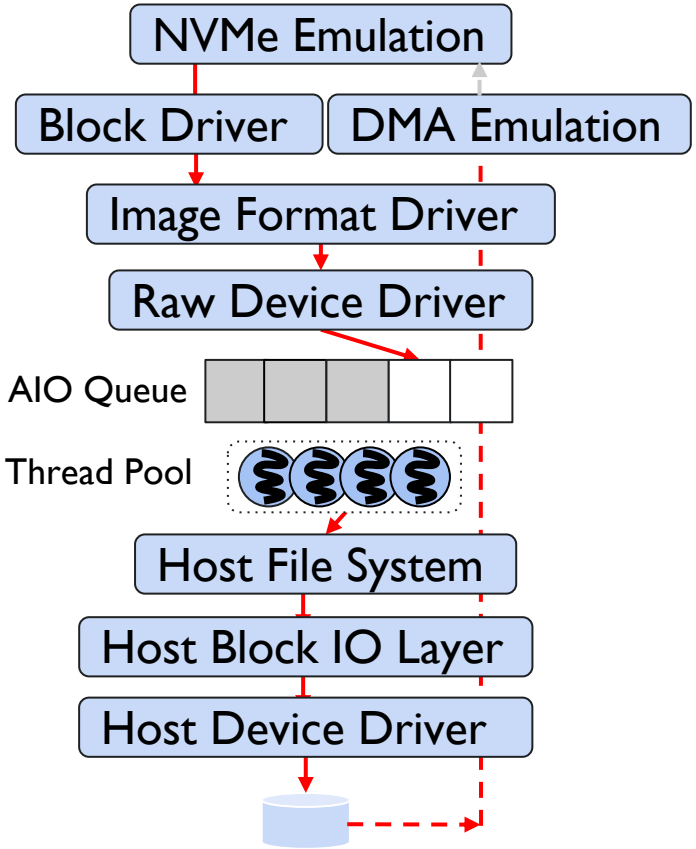
Scalability Root Causes & Solutions (2)

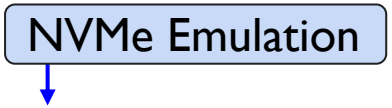


Scalability Root Causes & Solutions (2)

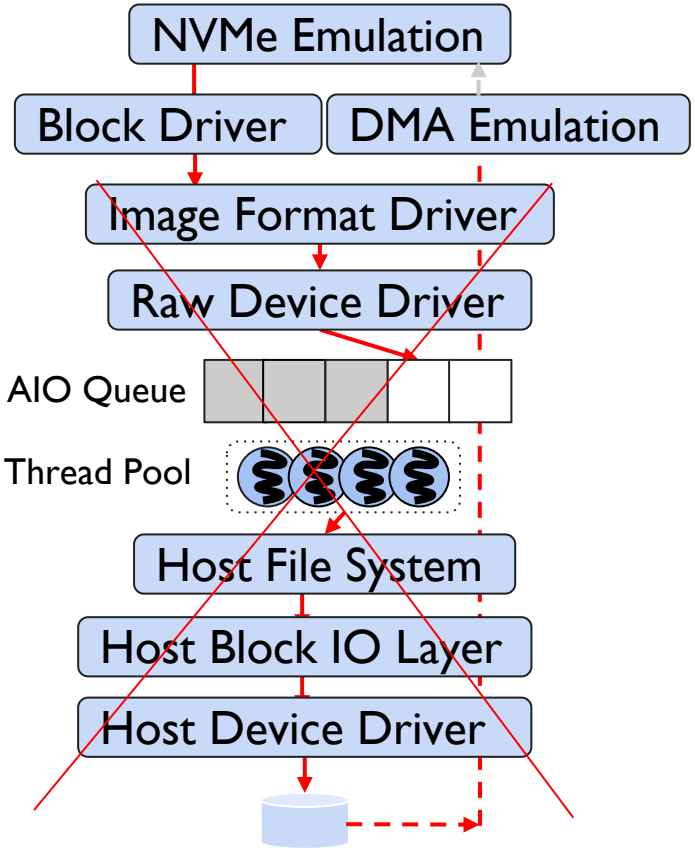


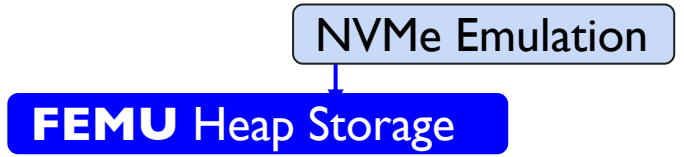
Scalability Root Causes & Solutions (2)



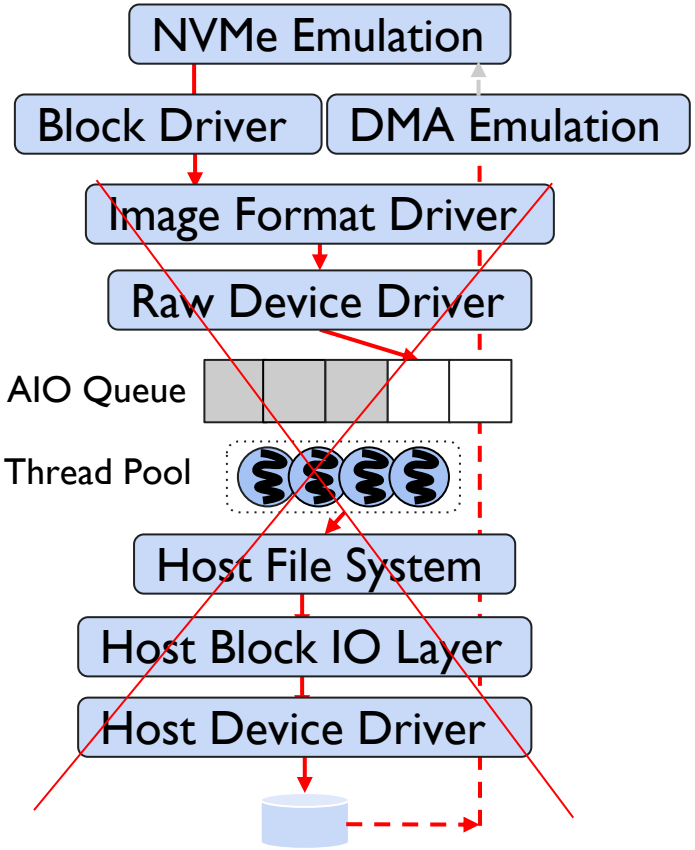


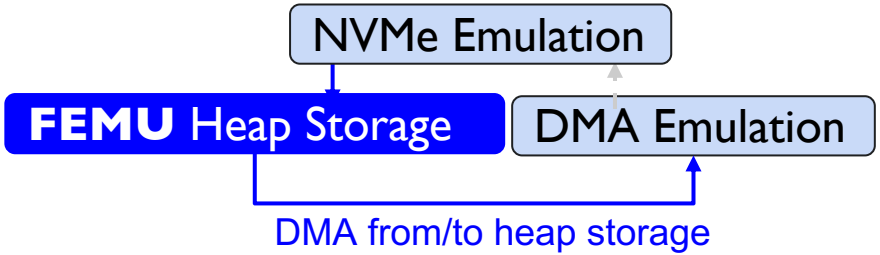
Scalability Root Causes & Solutions (2)



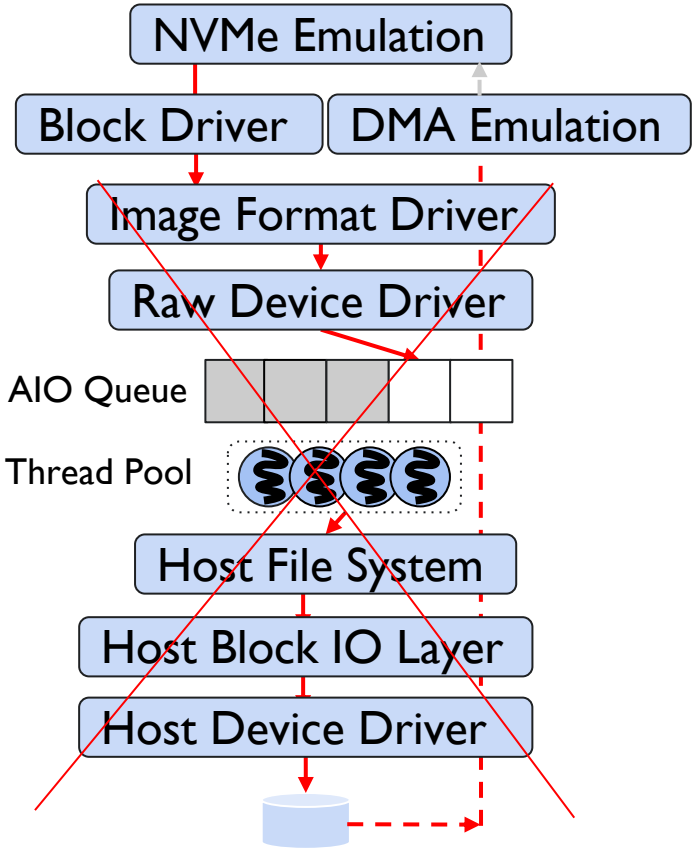


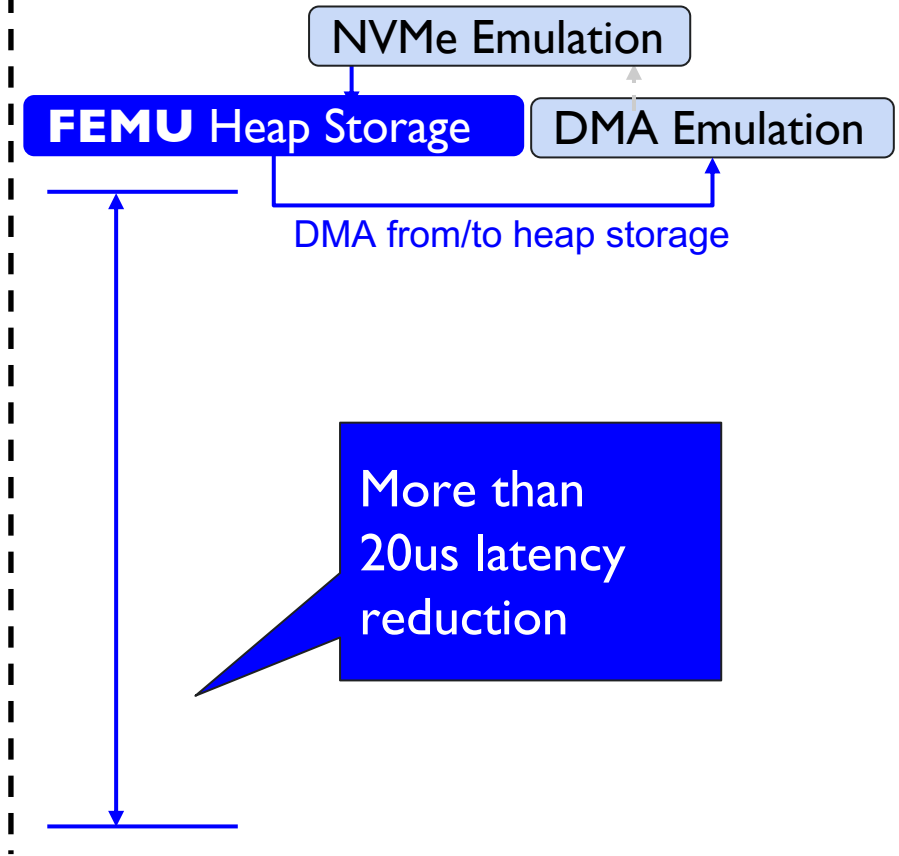
Scalability Root Causes & Solutions (2)





Scalability Root Causes & Solutions (2)





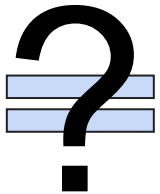
FEMU Accuracy

FEMU

OpenChannel-SSD

FEMU Accuracy

FEMU

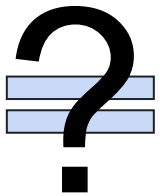


OpenChannel-SSD

FEMU Accuracy

App

FEMU

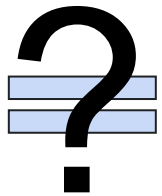


OpenChannel-SSD

FEMU Accuracy

App

FEMU



OpenChannel-SSD

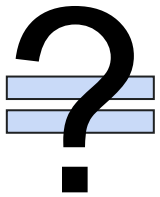
L_{femu}

L_{oc}

FEMU Accuracy

App

FEMU

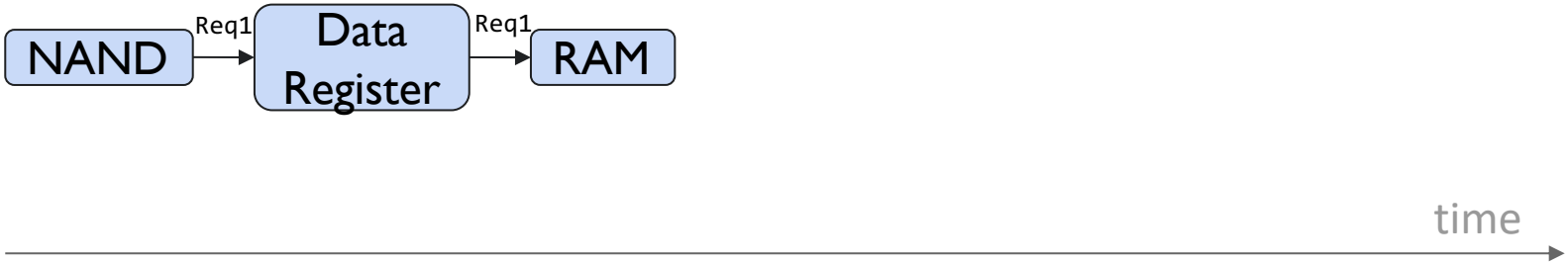


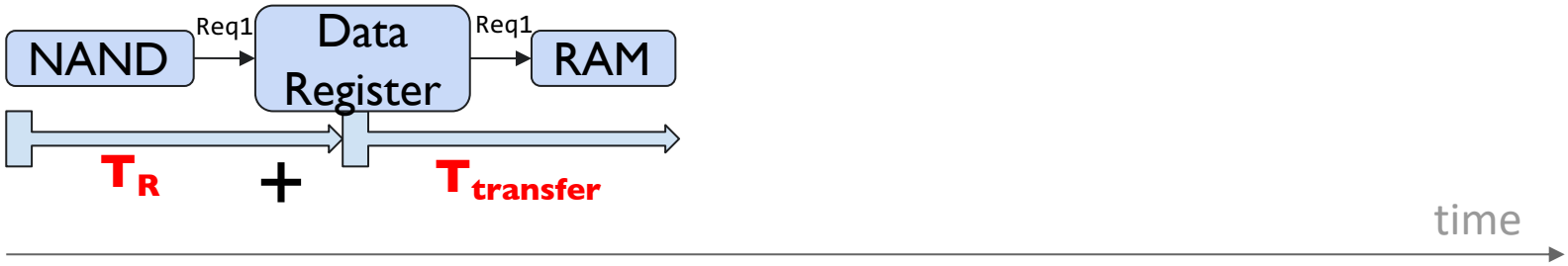
OpenChannel-SSD

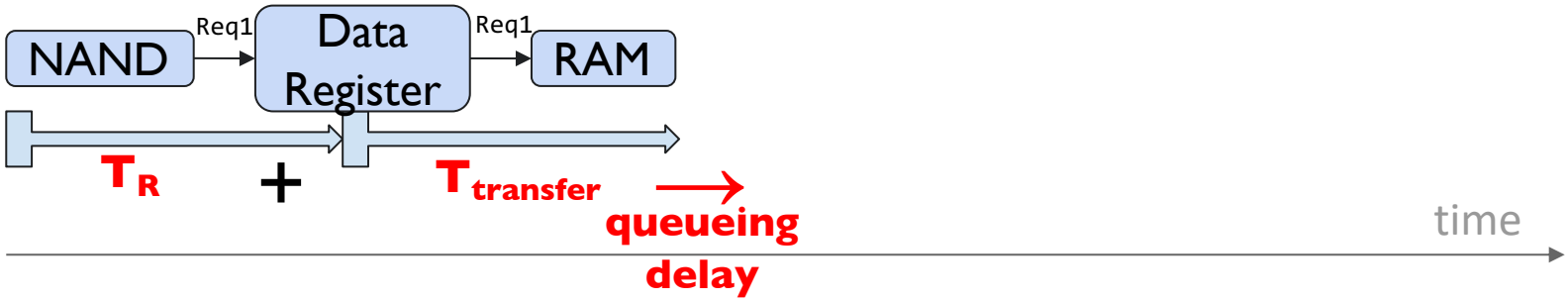
L_{femu} L_{oc}

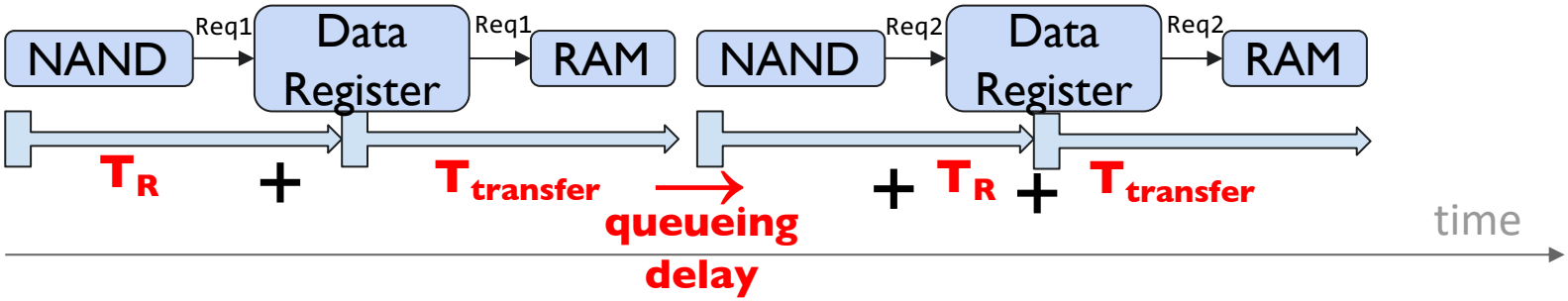
$$Error = |L_{femu} - L_{oc}| / L_{oc}$$


NAND
Req1
Data Register
time

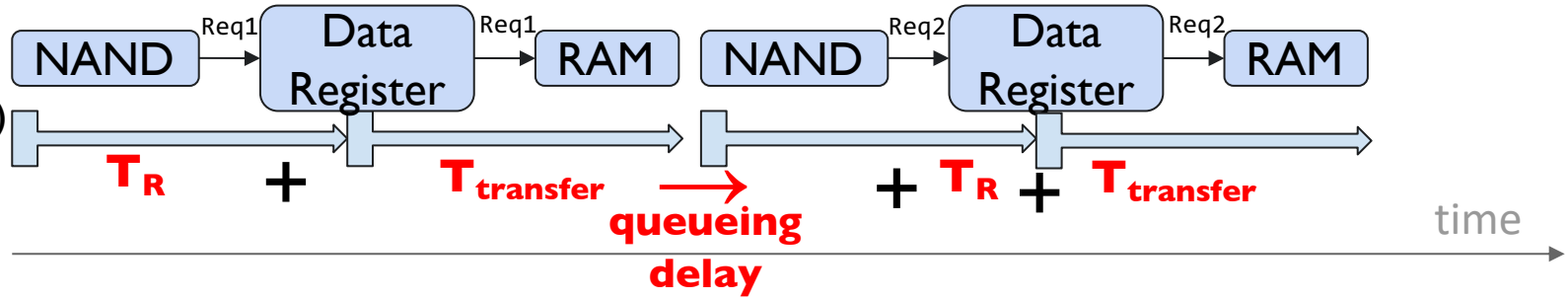

NAND
Req1
Data Register
Req1
RAM
time


NAND
Req1
Data Register
Req1
RAM
T_R
+
$T_{transfer}$
time

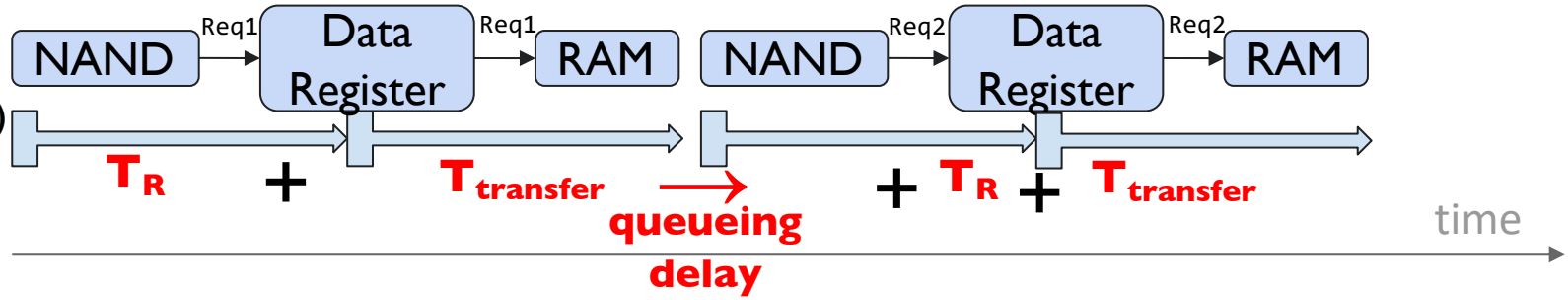

NAND
Req1
Data Register
Req1
RAM
T_R
+
$T_{transfer}$
queueing delay
time

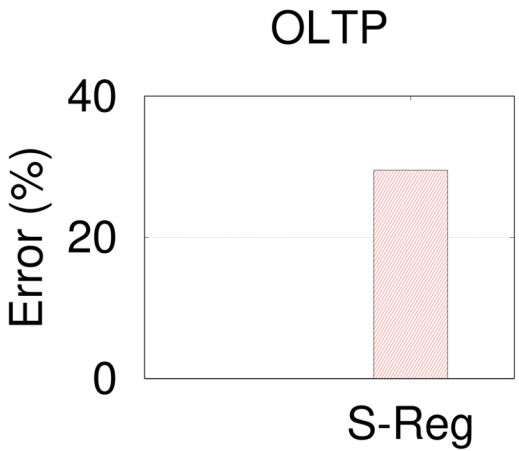

NAND
Req1
Data Register
Req1
RAM
NAND
Req2
Data Register
Req2
RAM
T R
+
T transfer
queueing delay
+
T R
+
T transfer
time

Single-Register model (**S-Reg**)

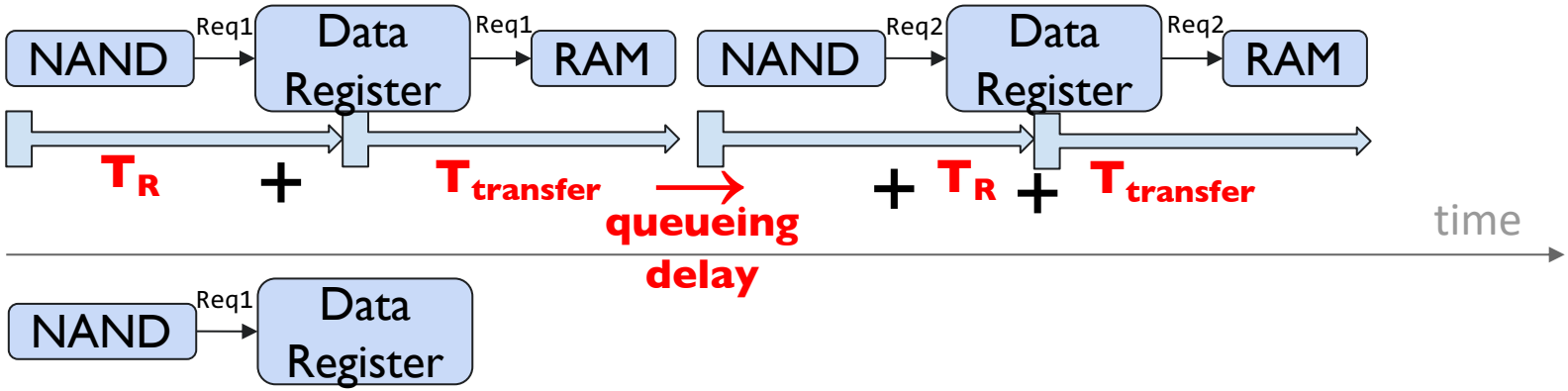


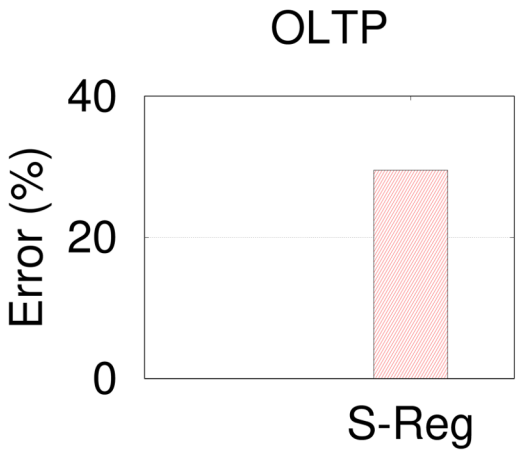
Single-Register model (**S-Reg**)



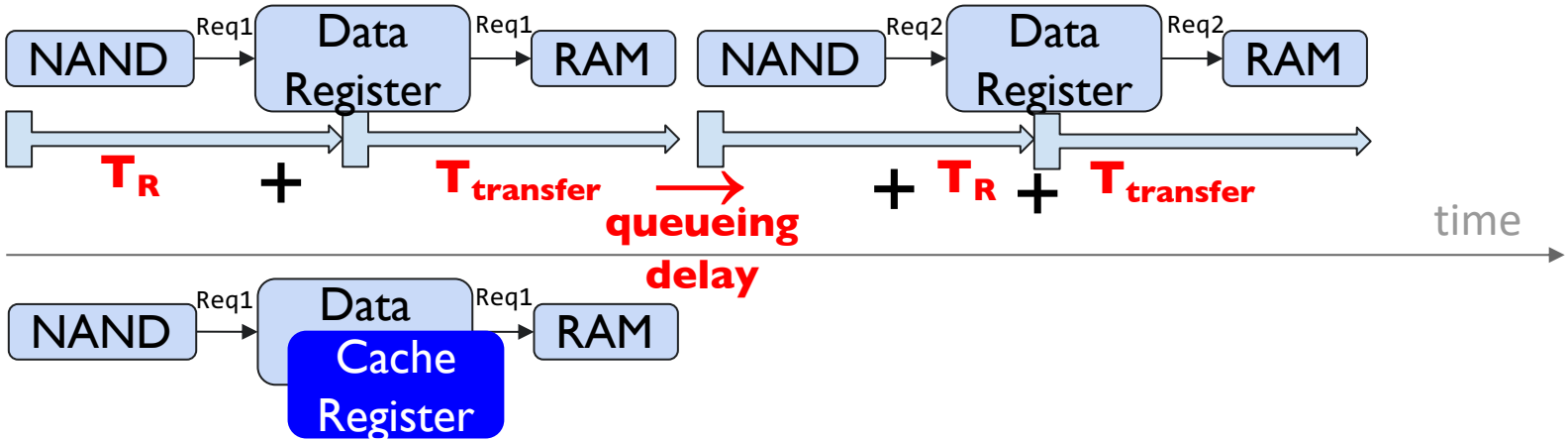


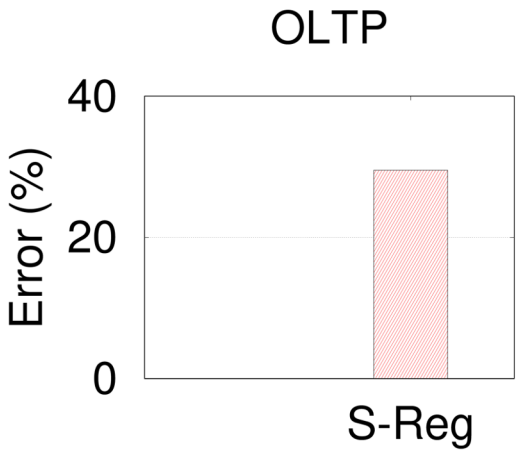
Single-Register model (**S-Reg**)



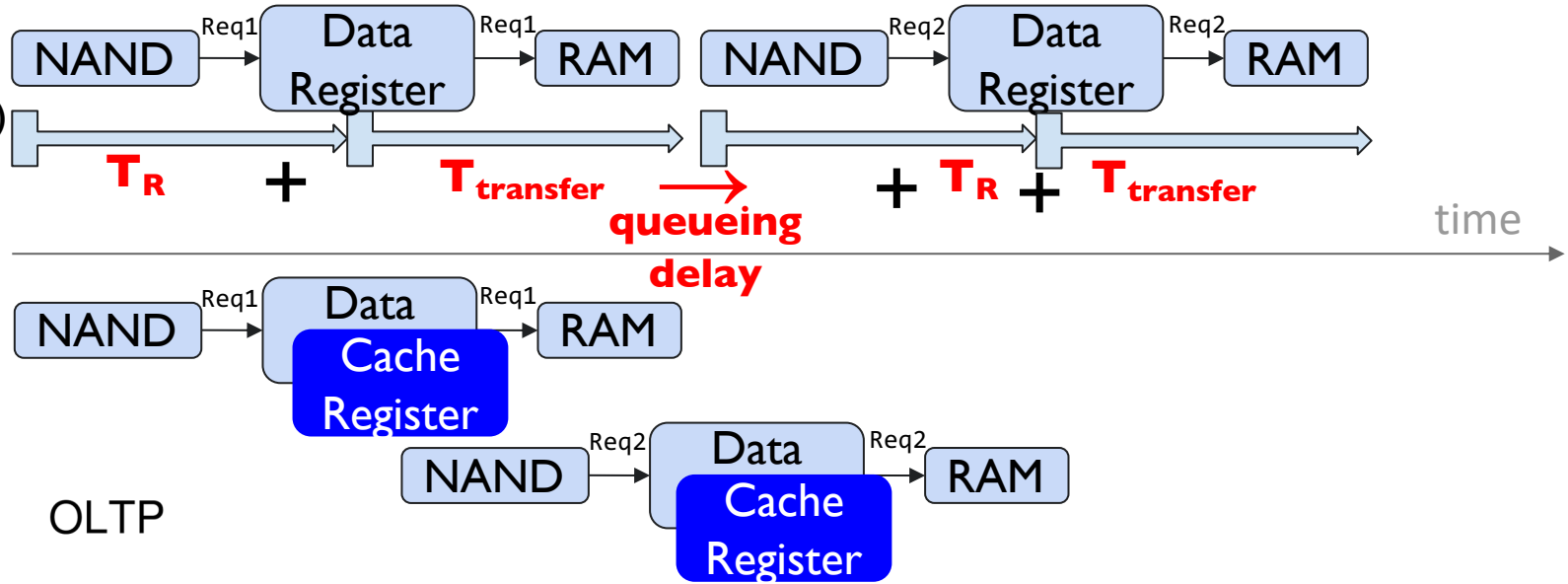


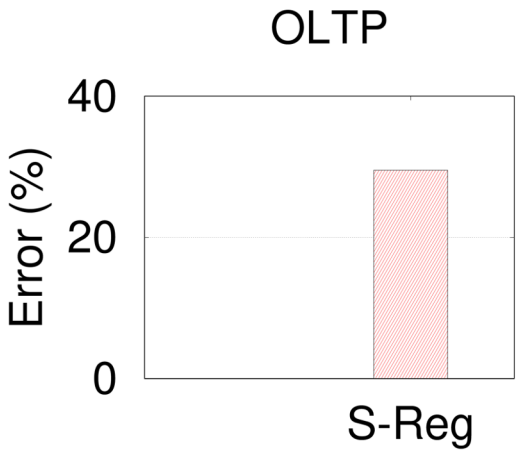
Single-Register model (**S-Reg**)



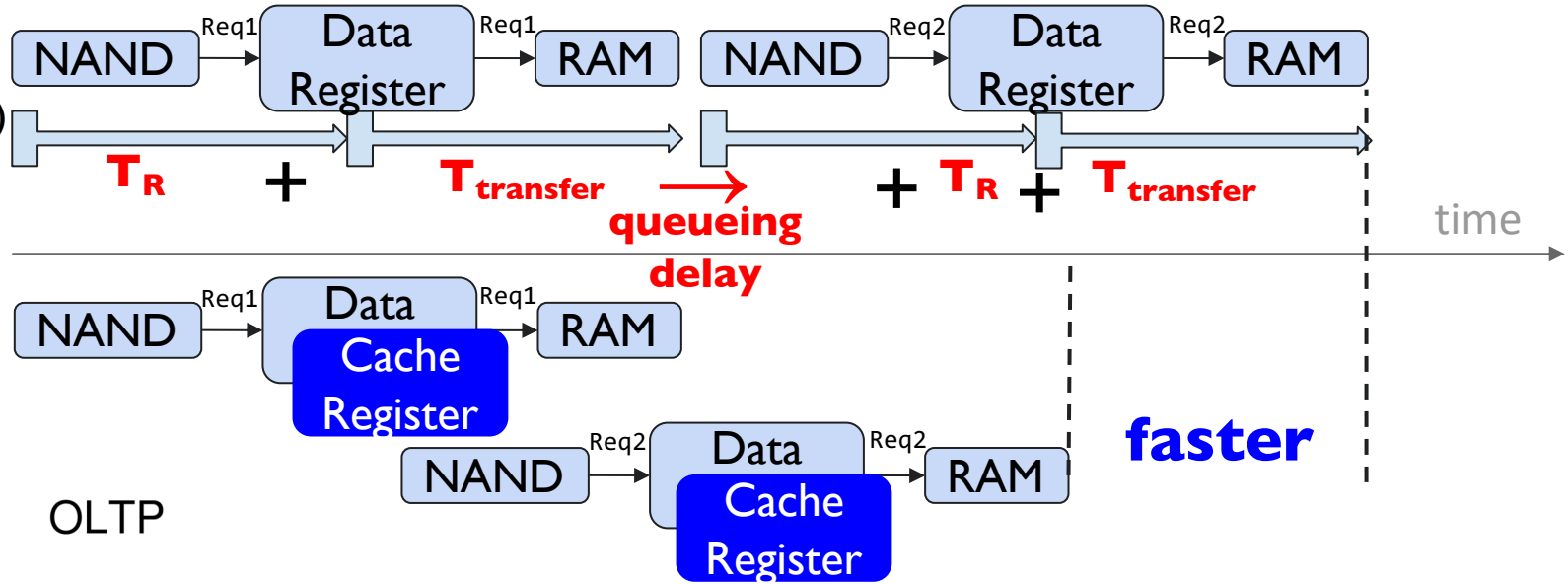


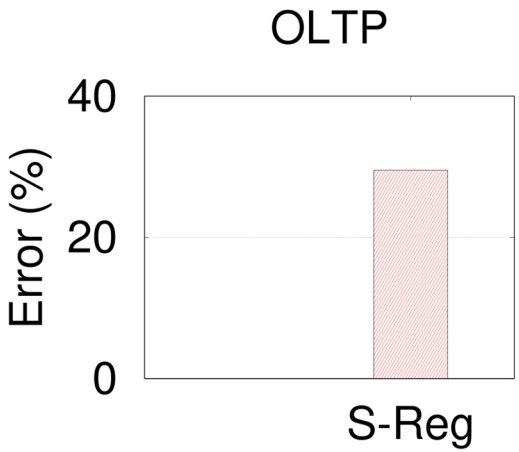
Single-Register model (**S-Reg**)

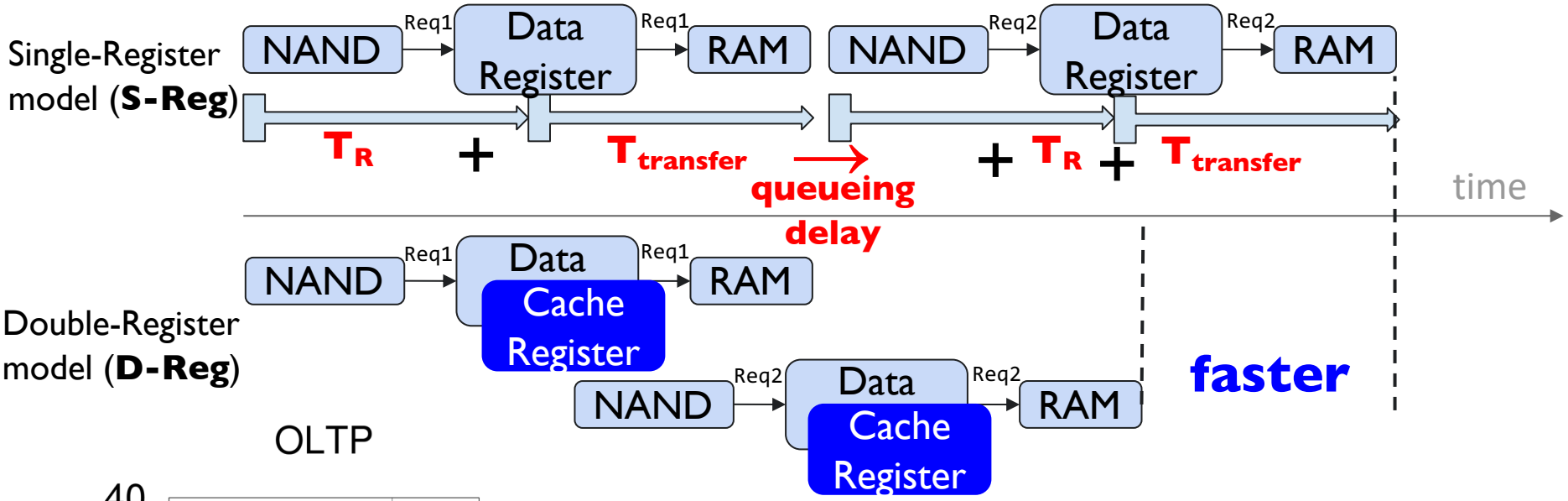


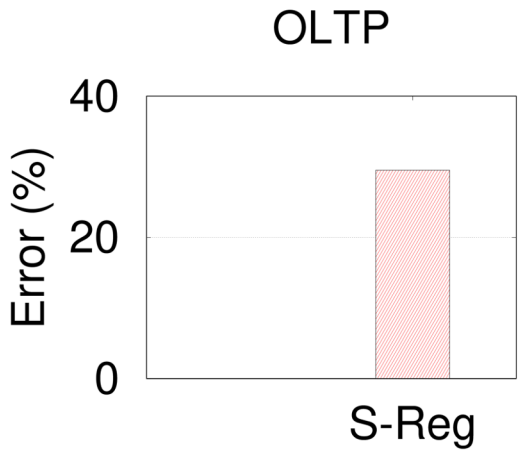


Single-Register model (**S-Reg**)

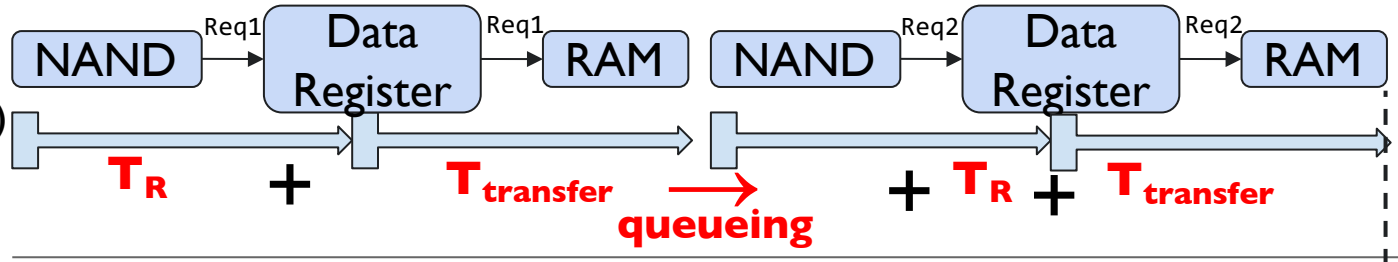




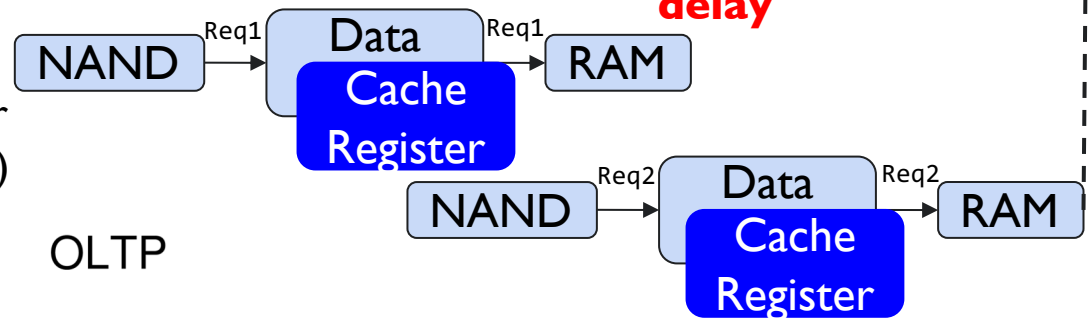

Single-Register model (S-Reg)
NAND
Req1
Data Register
Req1
RAM
NAND
Req2
Data Register
Req2
RAM
T R + T transfer
queueing delay
+ T R + T transfer
time
Double-Register model (D-Reg)
NAND
Req1
Data Cache Register
Req1
RAM
NAND
Req2
Data Cache Register
Req2
RAM
faster


OLTP
Error (%)
40
20
0
S-Reg

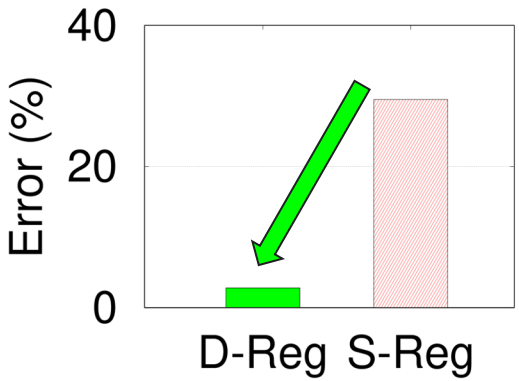
Single-Register model (**S-Reg**)



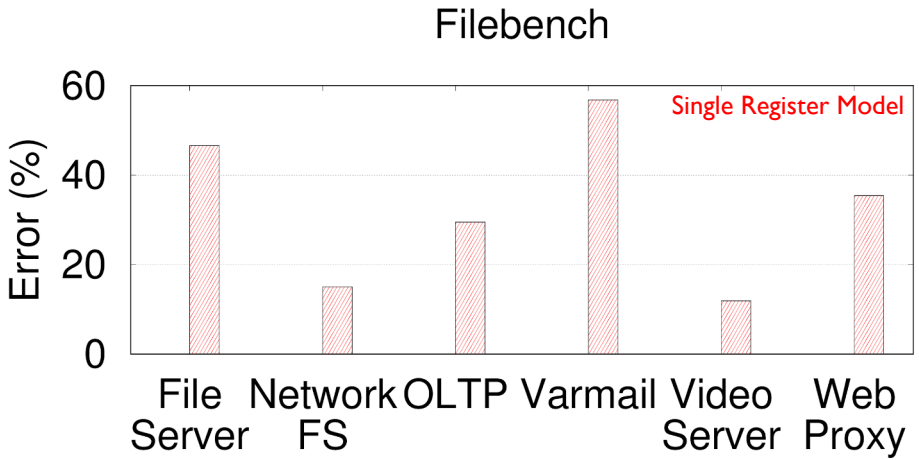
Double-Register model (**D-Reg**)



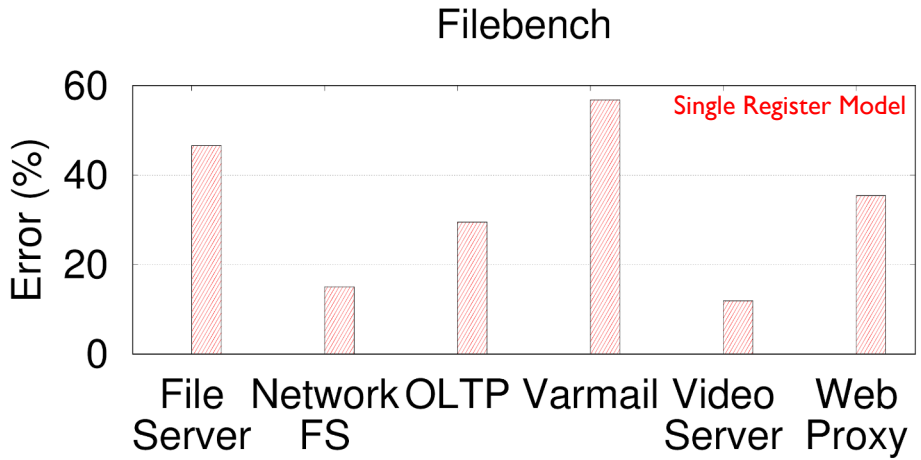
OLTP



FEMU Accuracy



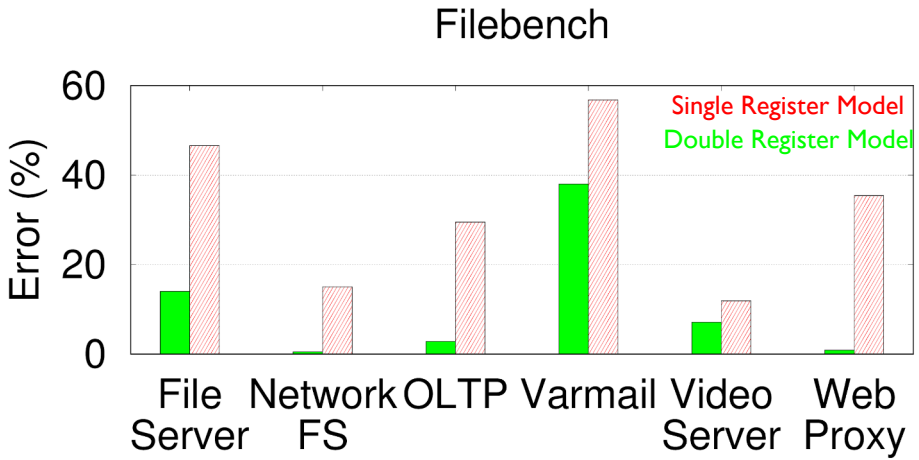
FEMU Accuracy



Latency Error: 11-57%

Single Register Model (**S-Reg**)

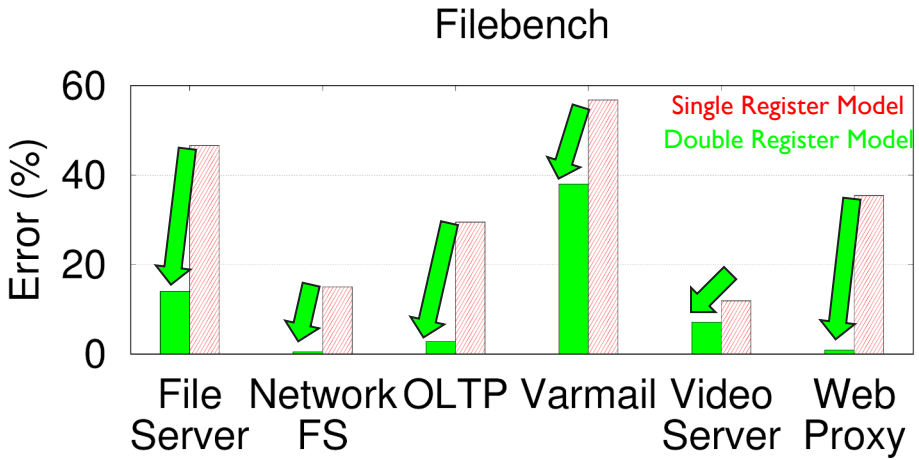
FEMU Accuracy



Latency Error: 11-57%

Single Register Model (**S-Reg**)

FEMU Accuracy

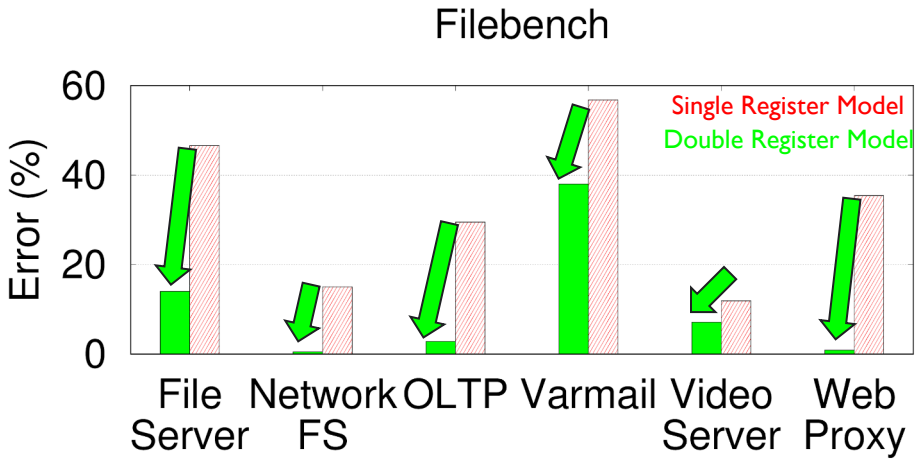


Latency Error: 11-57% ⇒ 0.5-38%

Single Register Model (**S-Reg**)

Double Register Model (**D-Reg**)

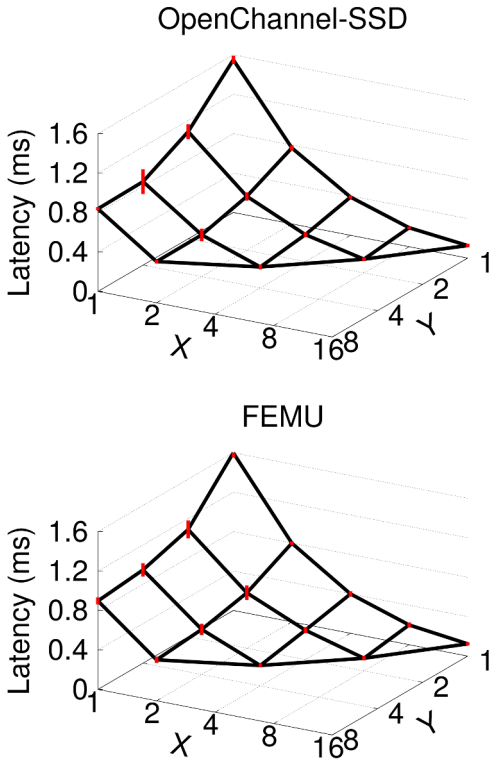
FEMU Accuracy



Latency Error: 11-57% ⇒ 0.5-38%

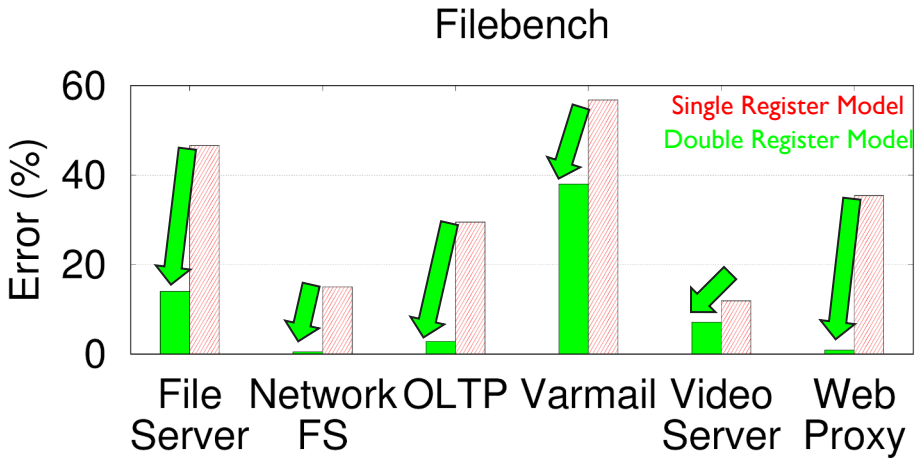
Single Register Model (**S-Reg**)

Double Register Model (**D-Reg**)



X: # of channels
Y: # of planes per channel

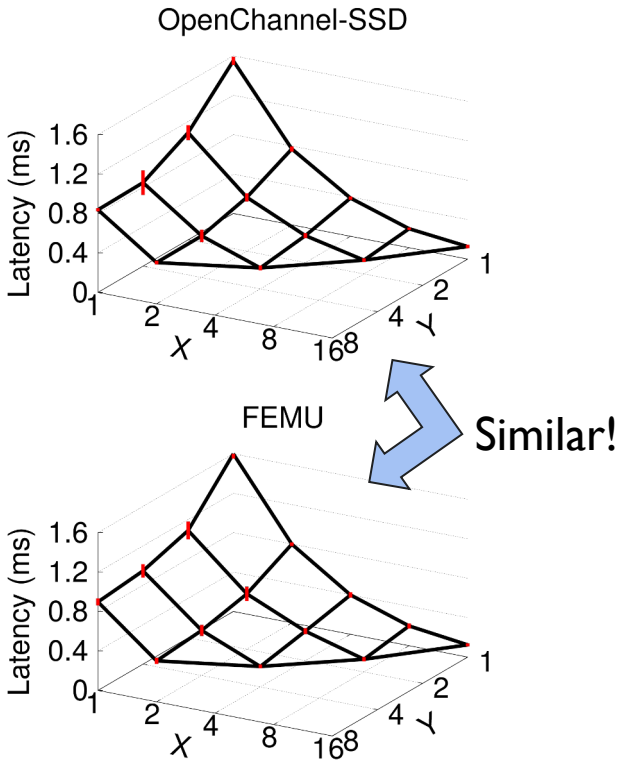
FEMU Accuracy



Latency Error: 11-57% ⇒ 0.5-38%

Single Register Model (**S-Reg**)

Double Register Model (**D-Reg**)



X: # of channels
Y: # of planes per channel

FEMU Limitations

- Further optimizations to support higher parallelism (more scalable)
- Accuracy can be improved
- Not able to emulate large-capacity SSD
- No persistence

Conclusion


FEMU
150mg
Cheap
Accurate
Scalable
Extensible
Order Now
https://github.com/ucare-uchicago/femu

Conclusion


FEMU
150mg
Installing, and using FEMU can cause side effects including headache, nausea, agitation, and depression. If your research condition does not improve after using FEMU for a week, please talk to us, your advisor, or your doctor immediately.
Cheap
Accurate
Scalable
Extensible
Order Now
https://github.com/ucare-uchicago/femu

Thank you!
Questions?

FEMU: https://github.com/ucare-uchicago/femu

Huaicheng Li
huaicheng@cs.uchicago.edu

CERES

UCARE

http://ucare.cs.uchicago.edu

THE UNIVERSITY OF CHICAGO