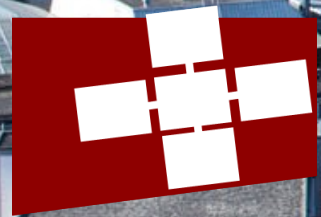


Marcin Copik, Marcin Chrapek, Larissa Schmid, Alexandru Calotoiu, Torsten Hoefler

Software Resource Disaggregation for HPC with Serverless Computing



IPDPS
San Francisco, CA USA
May 27-31, 2024

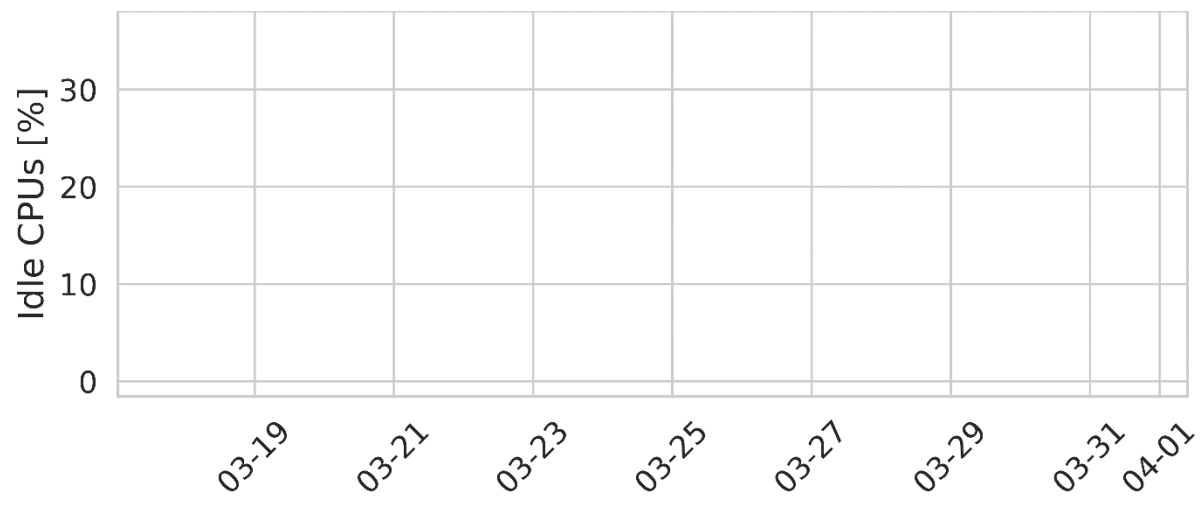
Tracking Wasted Resources in HPC



Tracking Wasted Resources in HPC



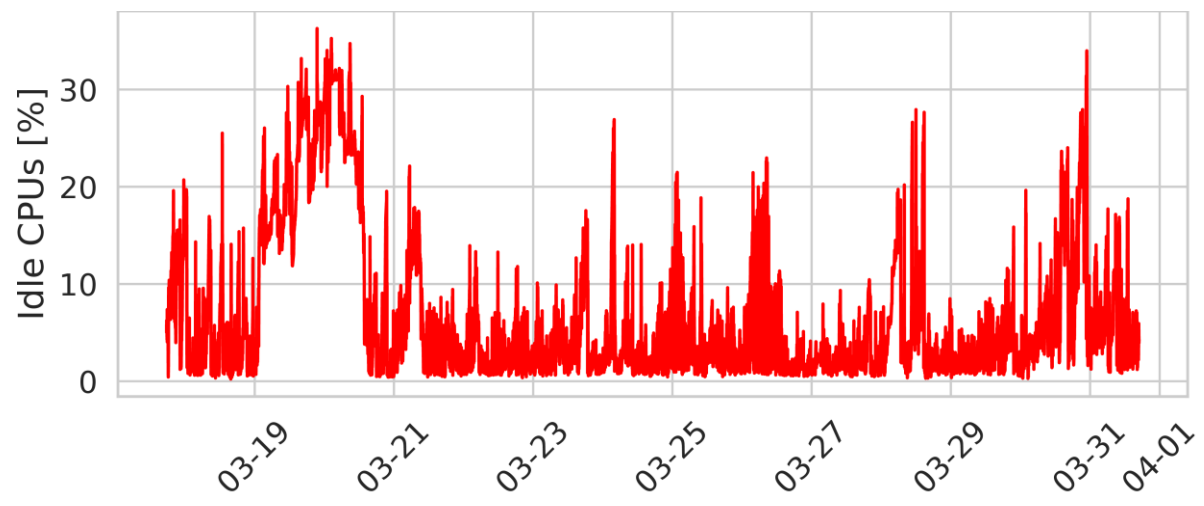
CPU



Tracking Wasted Resources in HPC



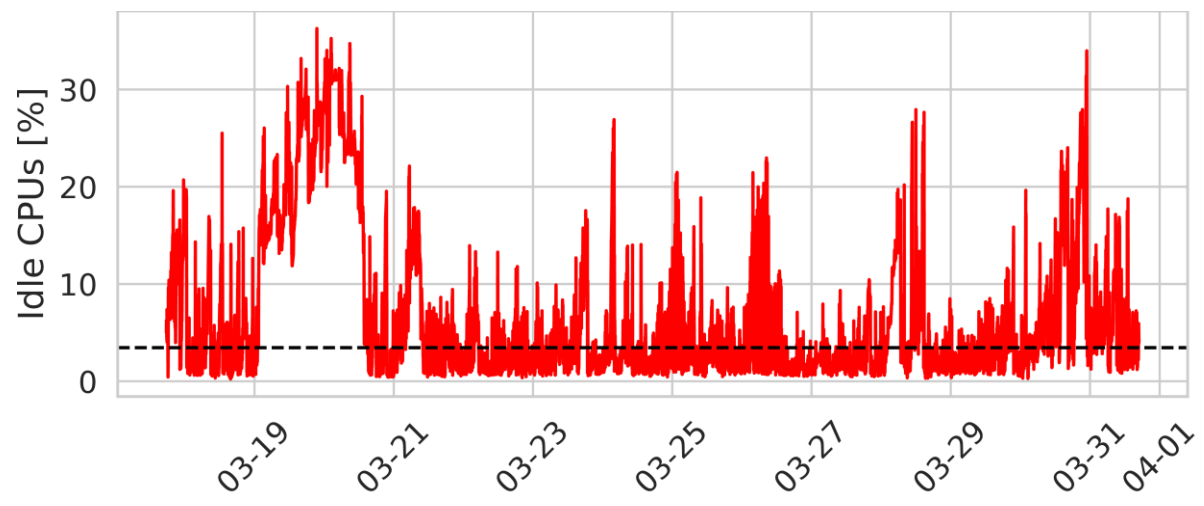
CPU



Tracking Wasted Resources in HPC



CPU

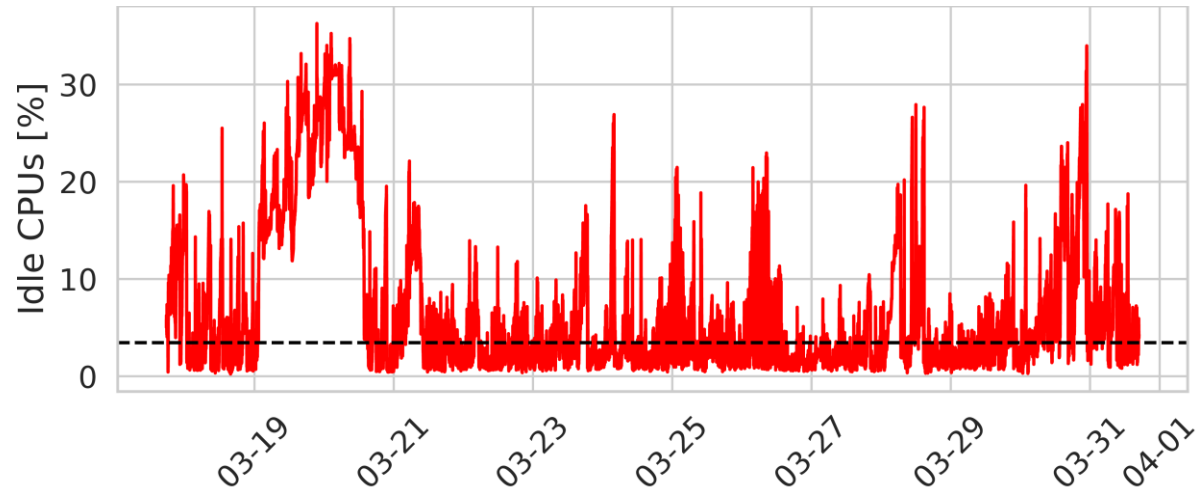


Mean idle CPUs: 6.6%

Tracking Wasted Resources in HPC

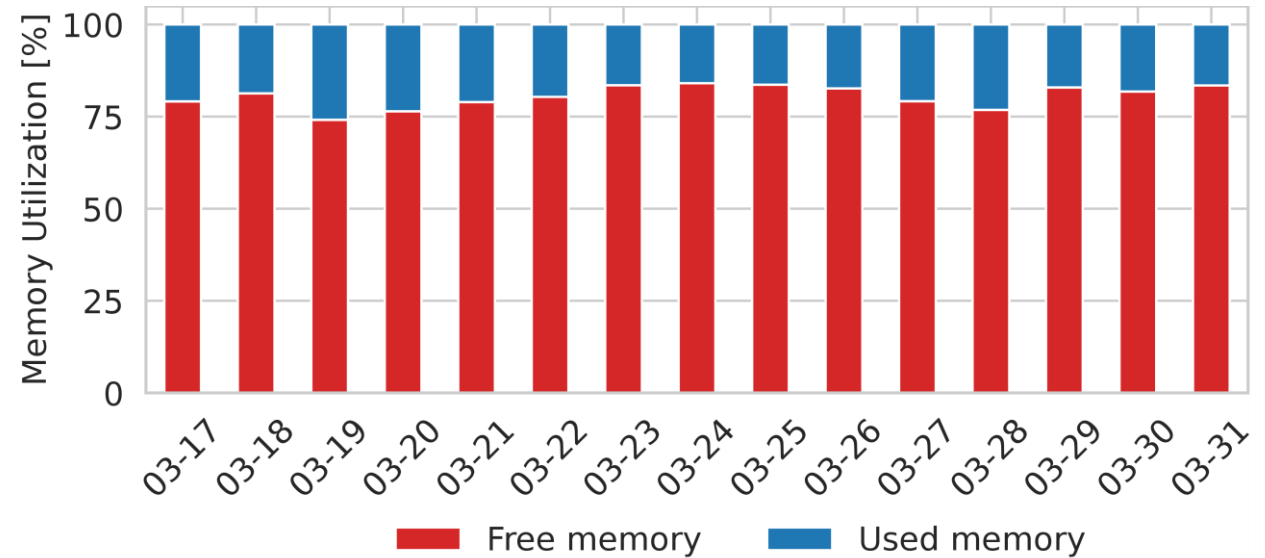


CPU



Mean idle CPUs: 6.6%

Memory



■ Free memory ■ Used memory

Mean free memory: 80.5%

Tracking Wasted Resources in HPC

Learning from Five-year Resource-Utilization Data of Titan System

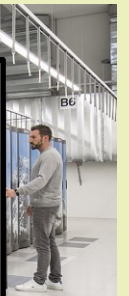
Feiyi Wang*, Sarp Oral†, Satyabrata Sen ‡ and Neena Imam§
Oak Ridge National Laboratory

CLUSTER, 2019

CPU

FINAL REPORT WORKLOAD ANALYSIS OF BLUE WATERS (ACI 1650758)

Matthew D. Jones, Joseph P. White, Martins Innus, Robert L. DeLeon, Nikolay Simakov, Jeffrey T. Palmer, Steven M. Gallo, and Thomas R. Furlani (furlani@buffalo.edu), Center for Computational Research, University at Buffalo, SUNY



Quantifying Memory Underutilization in HPC Systems and Using it to Improve Performance via Architecture Support

Gagandeep Panwar*
Virginia Tech
Blacksburg, USA
gpanwar@vt.edu

Da Zhang*
Virginia Tech
Blacksburg, USA
daz3@vt.edu

Yihan Pang*
Virginia Tech
Blacksburg, USA
pyihan1@vt.edu

Mai Dahshan
Virginia Tech
Blacksburg, USA
mdahshan@vt.edu

Nathan DeBardeleben
Los Alamos National Laboratory
Los Alamos, USA
ndebar@lanl.gov

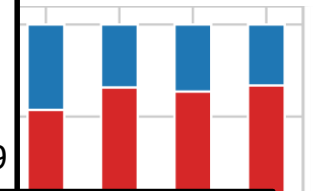
Binoy Ravindran
Virginia Tech
Blacksburg, USA
binoy@vt.edu

Xun Jian
Virginia Tech
Blacksburg, USA
xunj@vt.edu

Enos, and
lications
es.

iv, 2017

MICRO, 2019



A Case For Intra-rack Resource Disaggregation in HPC

GEORGE MICHELOGIANNAKIS, Lawrence Berkeley National Laboratory, USA
BENJAMIN KLENK, NVIDIA, USA
BRANDON COOK, Lawrence Berkeley National Laboratory, USA

A Holistic View of Memory Utilization on HPC Systems: Current and Future Trends

Ivy B. Peng*
peng8@llnl.gov
Lawrence Livermore National
Laboratory
USA

Ian Karlin
karlin1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Maya B. Gokhale
gokhale2@llnl.gov
Lawrence Livermore National
Laboratory
USA

Kathleen Shoga
Shoga1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Matthew Legendre
legendre1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Todd Gamblin
gamblin2@llnl.gov
Lawrence Livermore National
Laboratory
USA

MEMSYS, 2021

Comprehensive Workload Analysis and Modeling of a Petascale Supercomputer

Haihang You¹ and Hao Zhang²

¹ National Institute for Computational Sciences,
Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
² Department of Electrical Engineering and Computer Science,
University of Tennessee, Knoxville, TN 37996, USA
{hyou, haozhang}@utk.edu

JSSPP, 2012

Idle CPUs [%]

Memory Ut

Tracking Wasted Resources in HPC

Learning from Five-year Resource-Utilization Data of Titan System

Feiyi Wang*, Sarp Oral†, Satyabrata Sen ‡ and Neena Imam§
Oak Ridge National Laboratory

CLUSTER, 2019

CPU

FINAL REPORT WORKLOAD ANALYSIS OF BLUE WATERS (ACI 1650758)

Matthew D. Jones, Joseph P. White, Martins Innus, Robert L. DeLeon, Nikolay Simakov, Jeffrey T. Palmer, Steven M. Gallo, and Thomas R. Furlani (furlani@buffalo.edu), Center for Computational Research, University at Buffalo, SUNY

Quantifying Memory Underutilization in HPC Systems and Using it to Improve Performance via Architecture Support

Gagandeep Panwar*
Virginia Tech
Blacksburg, USA
gpanwar@vt.edu

Da Zhang*
Virginia Tech
Blacksburg, USA
daz3@vt.edu

Yihan Pang*
Virginia Tech
Blacksburg, USA
pyihan1@vt.edu

Mai Dahshan
Virginia Tech
Blacksburg, USA
mdahshan@vt.edu

Nathan DeBardeleben
Los Alamos National Laboratory
Los Alamos, USA
ndebard@lanl.gov

Binoy Ravindran
Virginia Tech
Blacksburg, USA
binoy@vt.edu

Xun Jian
Virginia Tech
Blacksburg, USA
xunj@vt.edu

MICRO, 2019

Enos, and
lications
es.
iv, 2017

A Case For Intra-rack Resource Disaggregation in HPC

GEORGE MICHELOGIANNAKIS, Lawrence Berkeley National Laboratory, USA
BENJAMIN KLENK, NVIDIA, USA
BRANDON COOK, Lawrence Berkeley National Laboratory, USA
MIN YEE
LARRY D
KEREN
JOHN S

Idle CPUs [%]
30
20
10
0

A Holistic View of Memory Utilization on HPC Systems: Current and Future Trends

Ivy B. Peng*
peng8@llnl.gov
Lawrence Livermore National
Laboratory
USA

Ian Karlin
karlin1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Maya B. Gokhale
gokhale2@llnl.gov
Lawrence Livermore National
Laboratory
USA

Kathleen Shoga
Shoga1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Matthew Legendre
legendre1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Todd Gamblin
gamblin2@llnl.gov
Lawrence Livermore National
Laboratory
USA

MEMSYS, 2021

Memory Ut

Comprehensive Workload Analysis and Modeling of a Petascale Supercomputer

Haihang You¹ and Hao Zhang²

¹ National Institute for Computational Sciences,
Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

² Department of Electrical Engineering and Computer Science,
University of Tennessee, Knoxville, TN 37996, USA

{hyou, haozhang}@utk.edu

JSSPP, 2012

Tracking Wasted Resources in HPC

Learning from Five-year Resource-Utilization Data of Titan System

Feiyi Wang*, Sarp Oral[†], Satyabrata Sen[‡] and Neena Imam[§]
Oak Ridge National Laboratory

CLUSTER, 2019

“the goal of achieving near 100% utilization while supporting a real parallel supercomputing workload is unrealistic”

FINAL REPORT WORKLOAD ANALYSIS OF BLUE WATERS (ACI 1650758)

Matthew D. Jones, Joseph P. White, Martins Innus, Robert L. DeLeon, Nikolay Simakov, Jeffrey T. Palmer, Steven M. Gallo, and Thomas R. Furlani (furlani@buffalo.edu), Center for Computational Research, University at Buffalo, SUNY

Quantifying Memory Underutilization in HPC Systems and Using it to Improve Performance via Architecture Support

Gagandeep Panwar*
Virginia Tech
Blacksburg, USA
gpanwar@vt.edu

Shan
a Tech
rg, USA
@vt.edu

Enos, and
Frank

Scheduling for Parallel Supercomputing: A Historical Perspective of Achievable Utilization

James Patton Jones¹ and Bill Nitzberg¹

MRJ Technology Solutions
NASA Ames Research Center, M/S 258-6
Moffett Field, CA 94035-1000

jjones@nas.nasa.gov

1999

A Holistic View of Memory Utilization on HPC Systems: Current and Future Trends

Ivy B. Peng*
peng8@llnl.gov
Lawrence Livermore National
Laboratory
USA

Ian Karlin
karlin1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Maya B. Gokhale
gokhale2@llnl.gov
Lawrence Livermore National
Laboratory
USA

Kathleen Shoga
Shoga1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Matthew Legendre
legendre1@llnl.gov
Lawrence Livermore National
Laboratory
USA

Todd Gamblin
gamblin2@llnl.gov
Lawrence Livermore National
Laboratory
USA

MEMSYS, 2021

Characterization and Modeling of a Petascale Supercomputer

Haihang You¹ and Hao Zhang²

¹ National Institute for Computational Sciences,
Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

² Department of Electrical Engineering and Computer Science,
University of Tennessee, Knoxville, TN 37996, USA

{hyou, haozhang}@utk.edu

JSSPP, 2012

Idle CPUs [%]

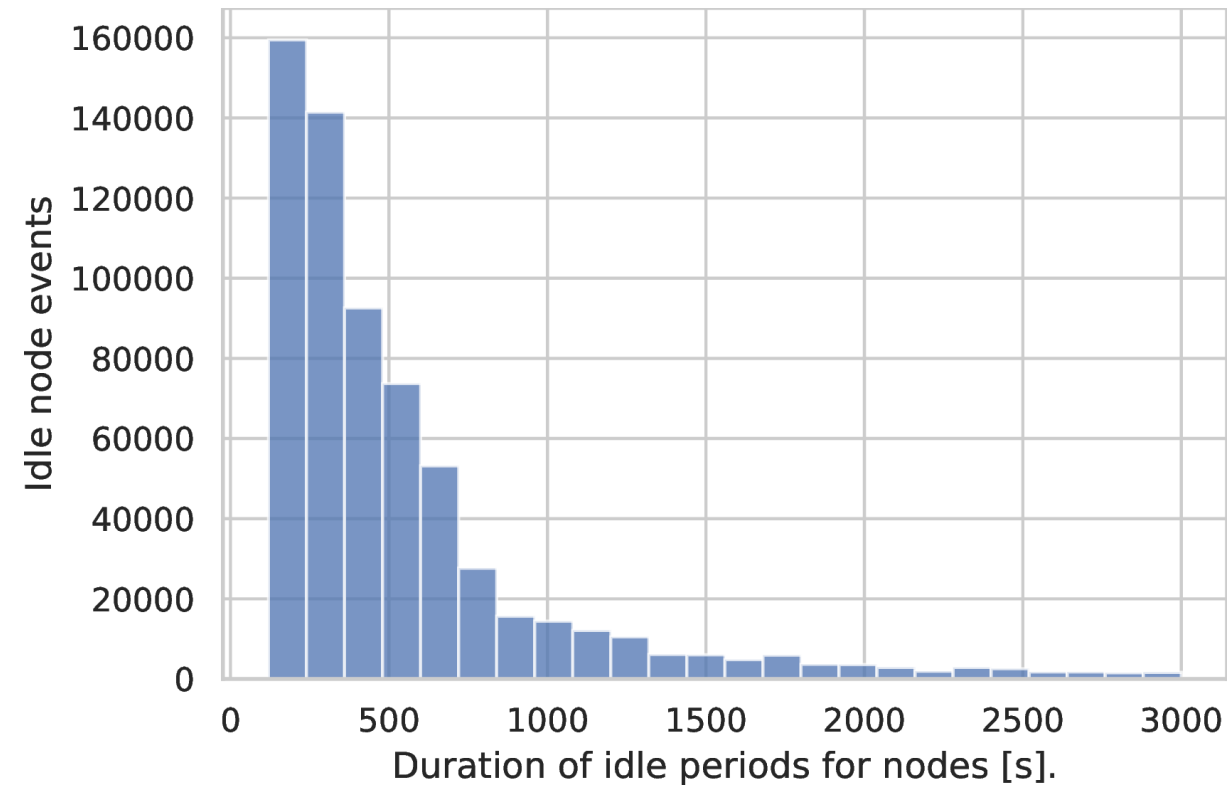


Memory Ut

Tracking Wasted Resources in HPC

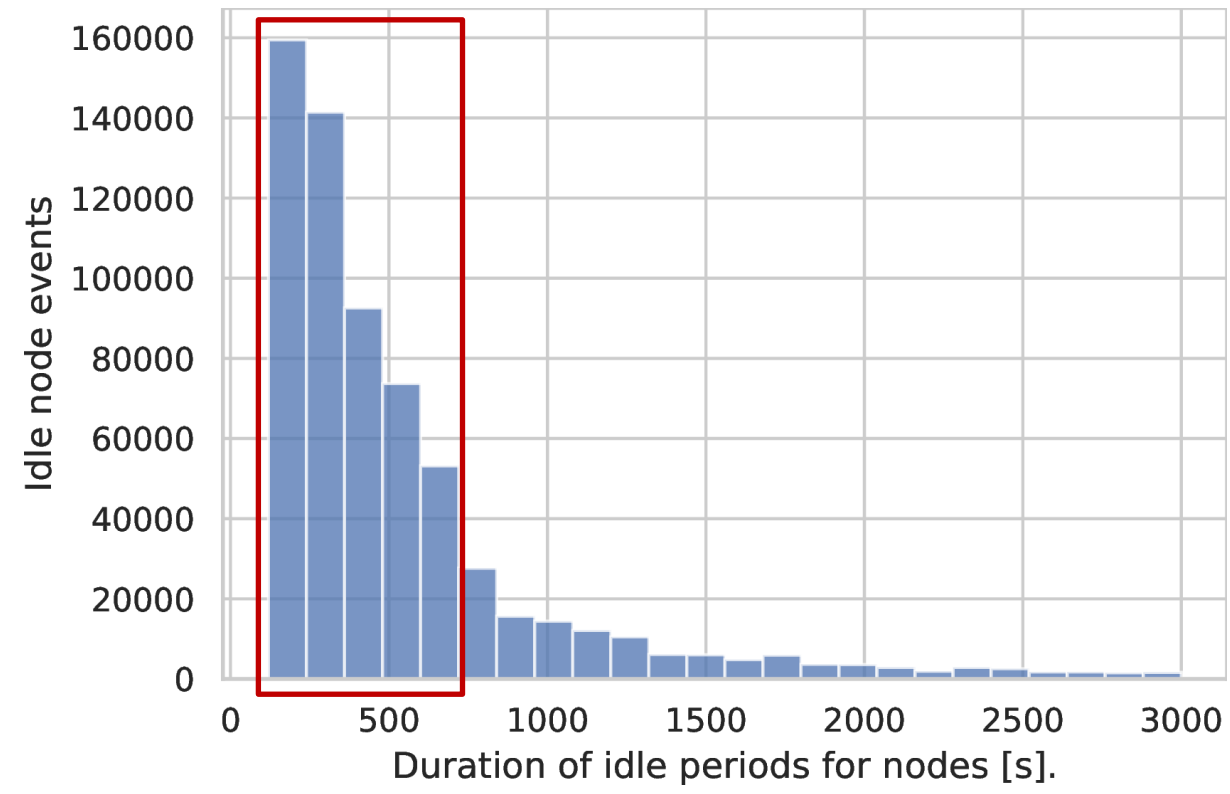
Tracking Wasted Resources in HPC

Duration of Node Idleness



Tracking Wasted Resources in HPC

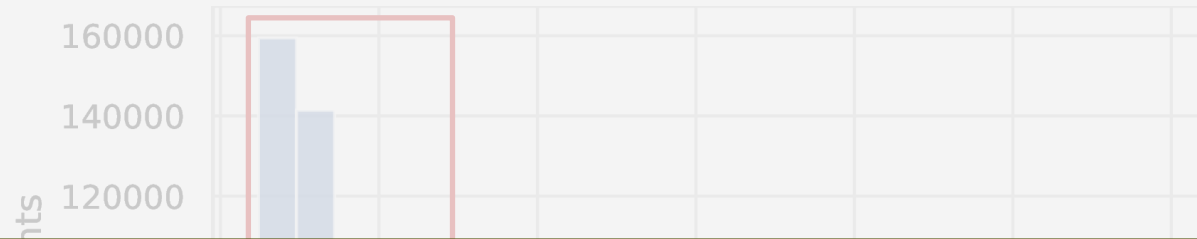
Duration of Node Idleness



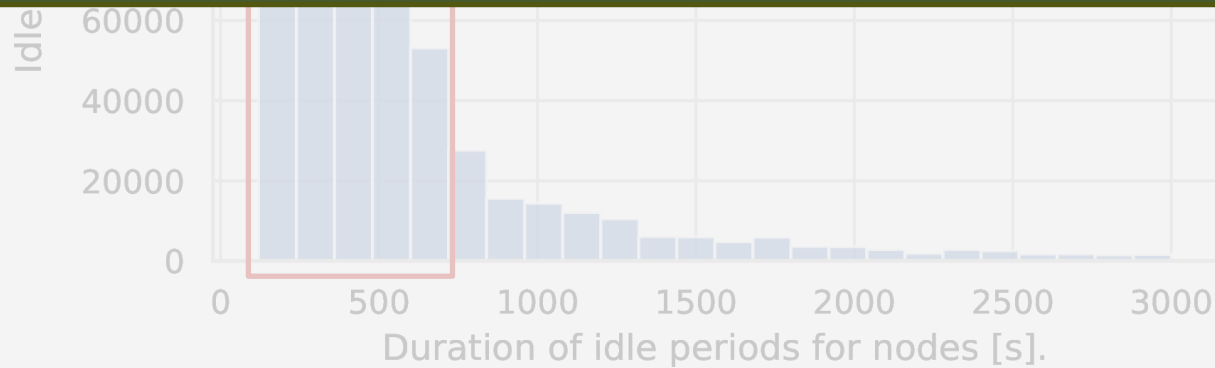
70% of idle node events last less than 10 minutes.

Tracking Wasted Resources in HPC

Duration of Node Idleness

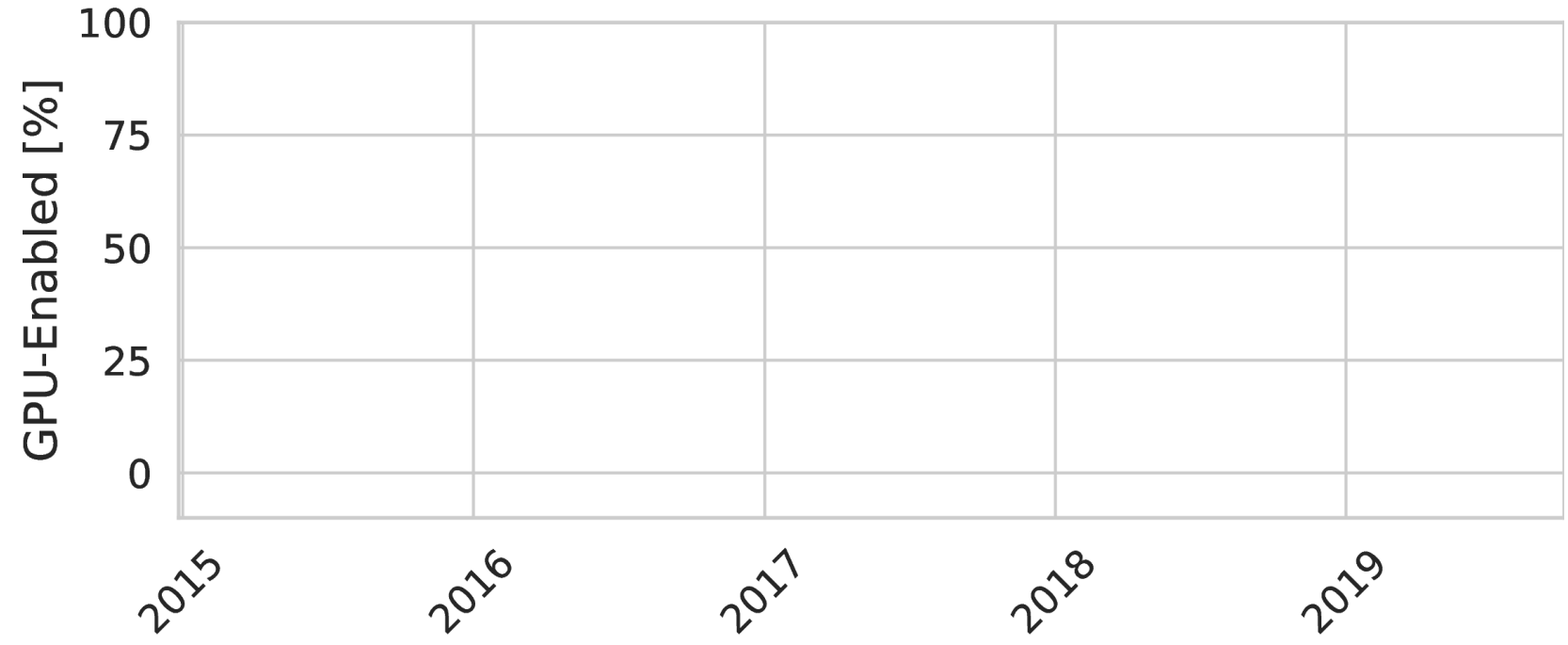


Short-term resource availability requires short-term allocations.



70% of idle node events last less than 10 minutes.

HPC System Utilization - GPU

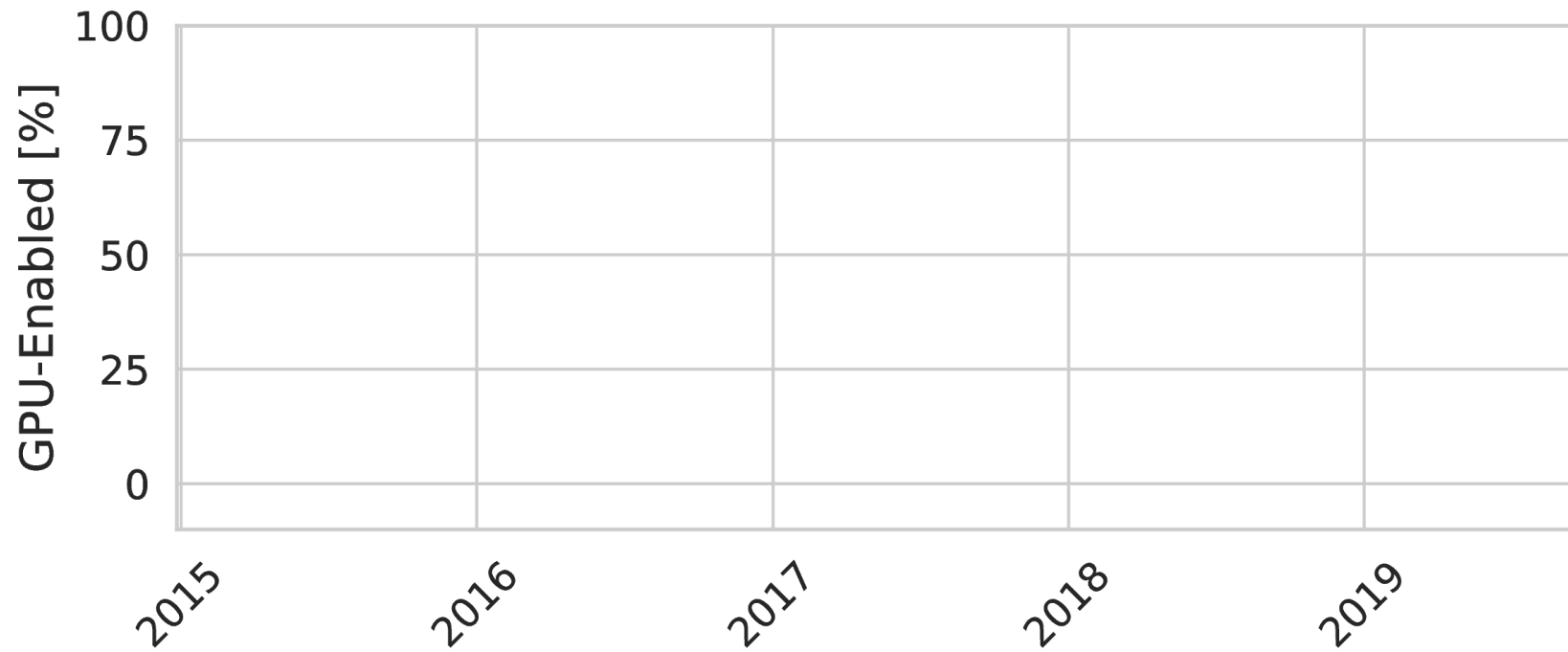


HPC System Utilization - GPU

Learning from Five-year Resource-Utilization Data
of Titan System

Feiyi Wang^{*}, Sarp Oral[†], Satyabrata Sen[‡] and Neena Imam[§]
Oak Ridge National Laboratory

CLUSTER, 2019

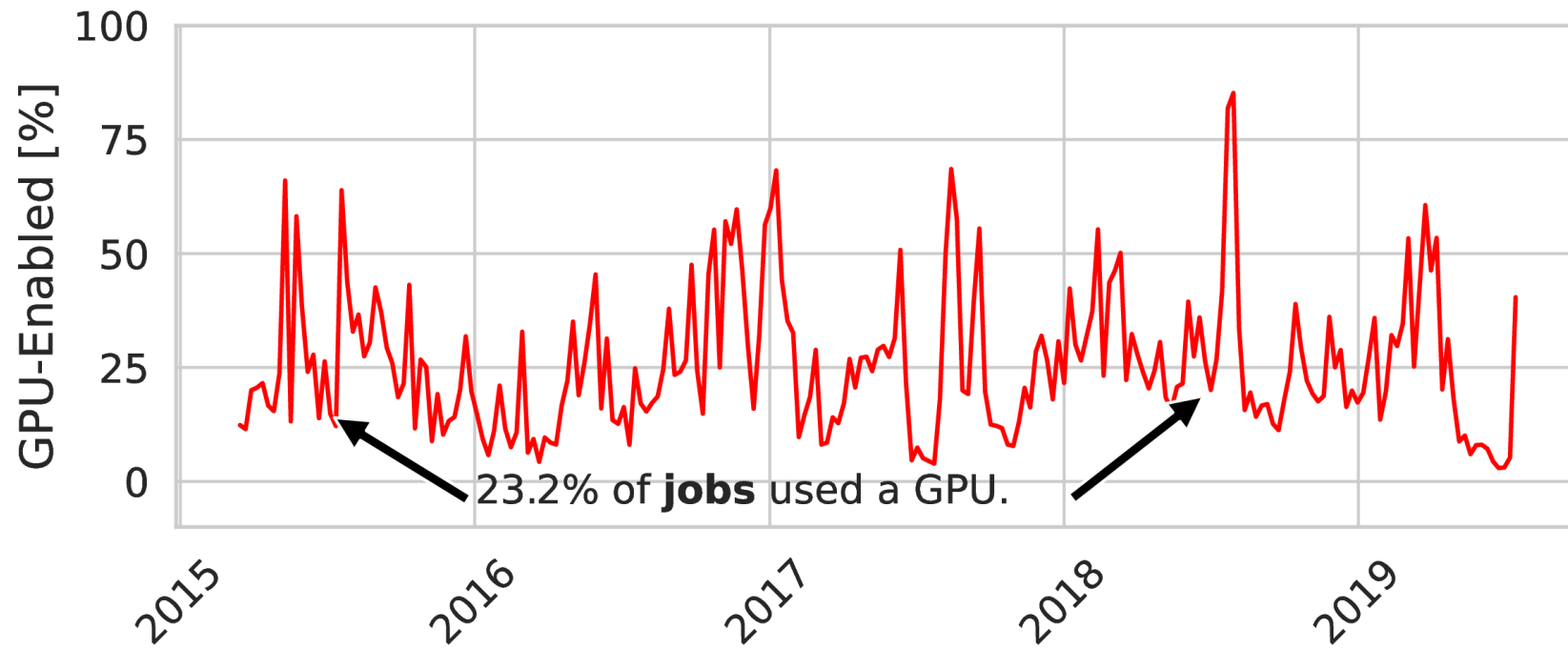


HPC System Utilization - GPU

Learning from Five-year Resource-Utilization Data
of Titan System

Feiyi Wang^{*}, Sarp Oral[†], Satyabrata Sen[‡] and Neena Imam[§]
Oak Ridge National Laboratory

CLUSTER, 2019

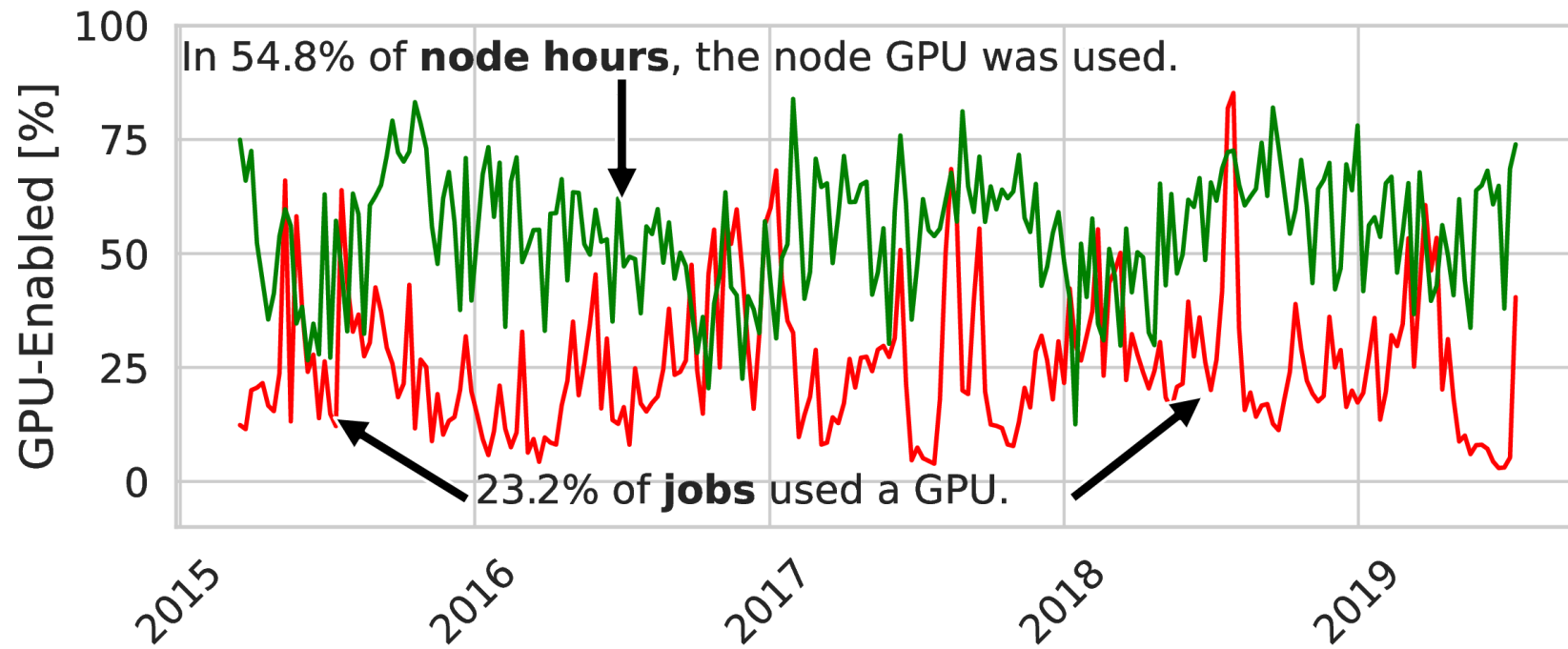


HPC System Utilization - GPU

Learning from Five-year Resource-Utilization Data
of Titan System

Feiyi Wang^{*}, Sarp Oral[†], Satyabrata Sen[‡] and Neena Imam[§]
Oak Ridge National Laboratory

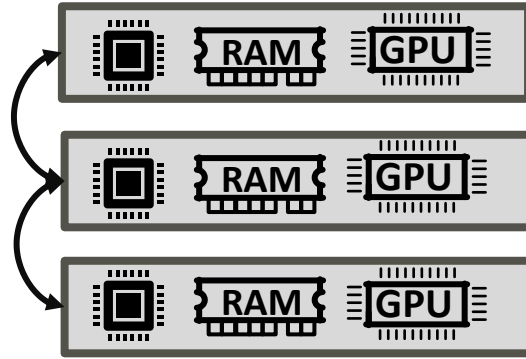
CLUSTER, 2019



Software Solution

Software Solution

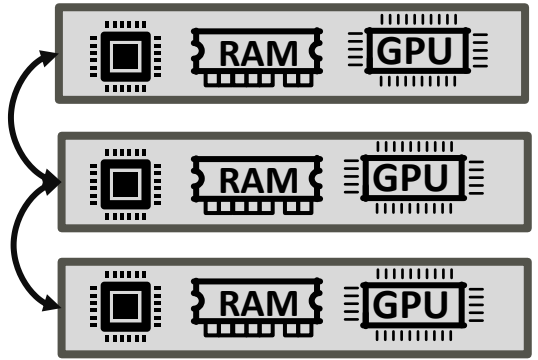
Standard HPC Node



- ✓ High performance
- ✗ Inflexible architecture

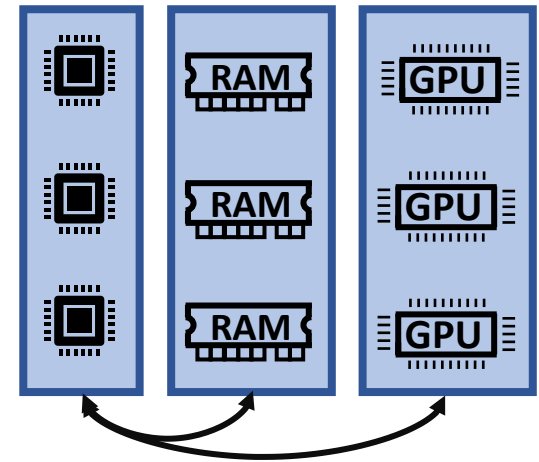
Software Solution

Standard HPC Node



- ✓ High performance
- ✗ Inflexible architecture

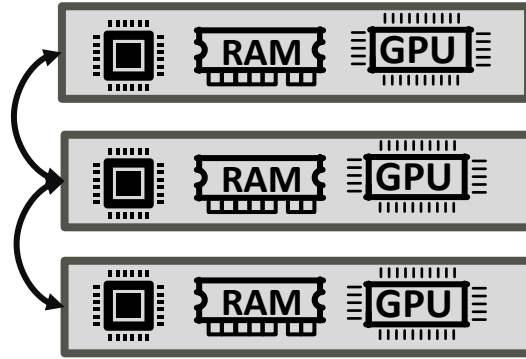
Hardware Disaggregation



- ✓ High efficiency
- ✗ Cost, performance penalty

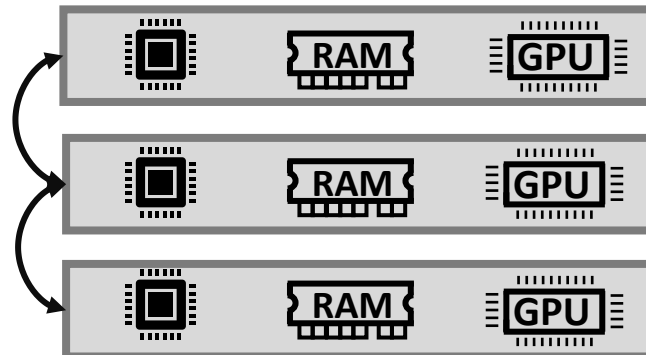
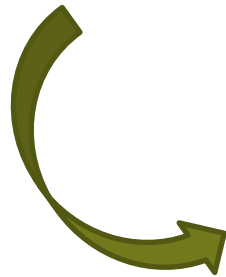
Software Solution

Standard HPC Node

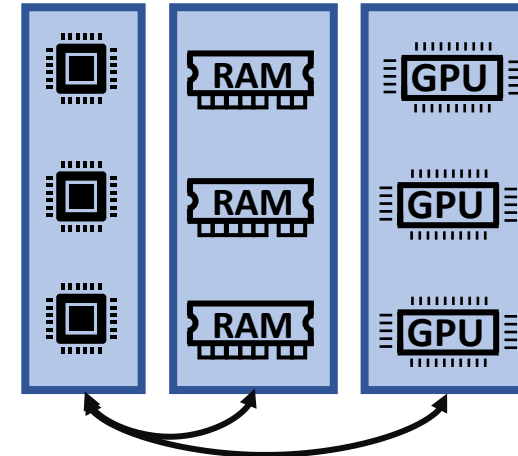


- ✓ High performance
- ✗ Inflexible architecture

Existing Coupled
Hardware Systems



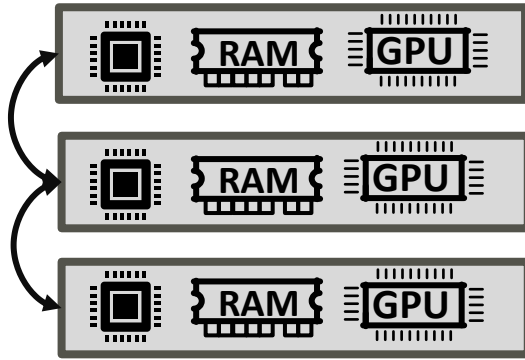
Hardware Disaggregation



- ✓ High efficiency
- ✗ Cost, performance penalty

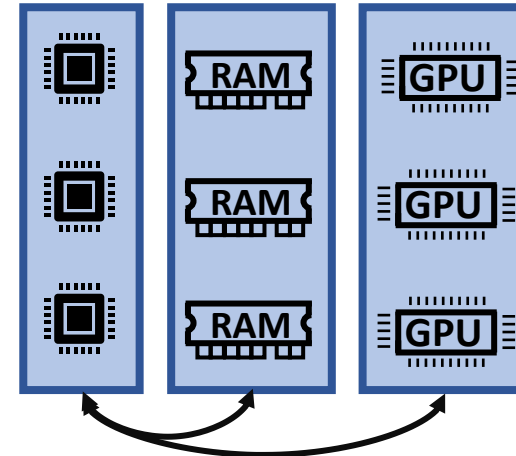
Software Solution

Standard HPC Node



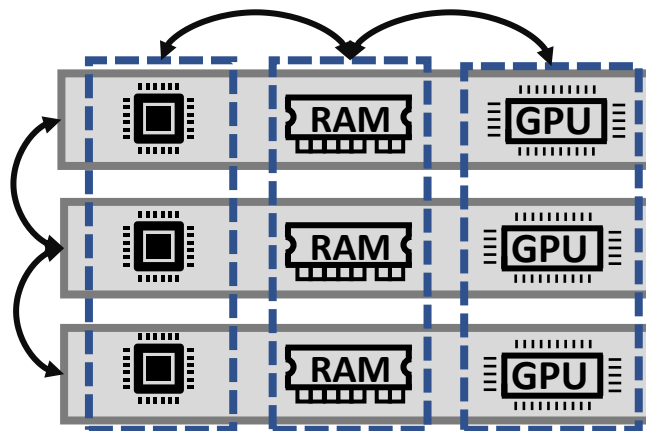
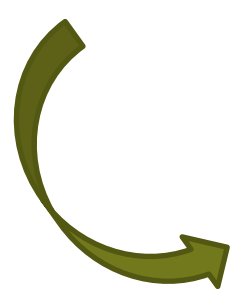
- ✓ High performance
- ✗ Inflexible architecture

Hardware Disaggregation

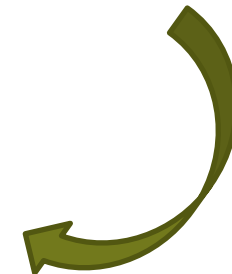


- ✓ High efficiency
- ✗ Cost, performance penalty

Existing Coupled Hardware Systems



Software Abstraction for Disaggregation



We propose a software disaggregation approach to share node resources

We propose a software disaggregation approach to share node resources
between
coarse-grained, long-running, and static batch jobs

We propose a software disaggregation approach to share node resources

between

coarse-grained, long-running, and static batch jobs

and

fine-grained, short-term, and dynamically allocated serverless functions.

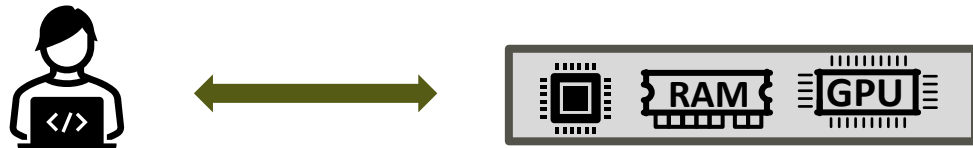
Serverless as an Answer

Hardware Abstraction

Software Abstraction

Serverless as an Answer

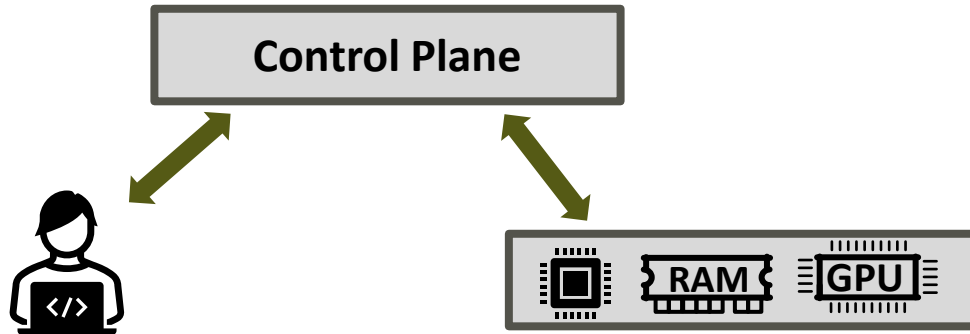
Hardware Abstraction



Software Abstraction

Serverless as an Answer

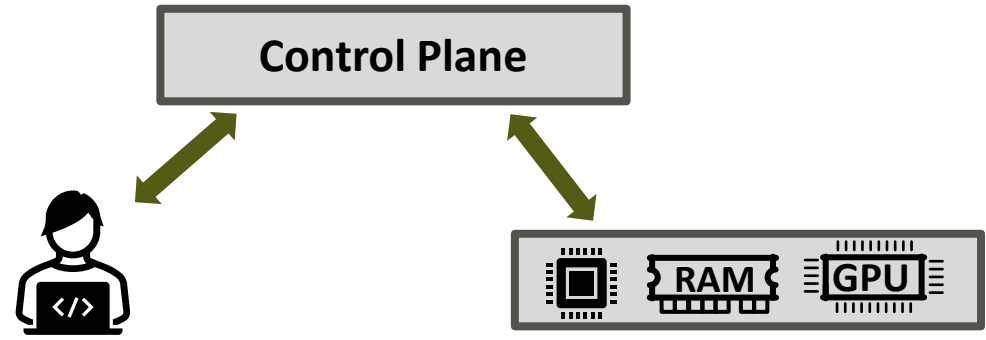
Hardware Abstraction



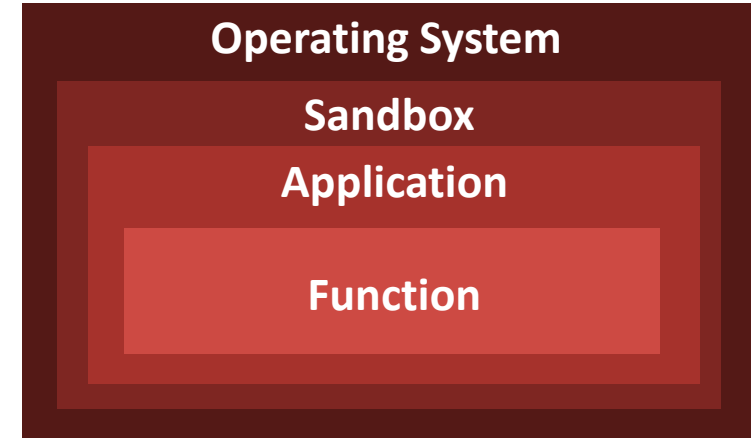
Software Abstraction

Serverless as an Answer

Hardware Abstraction

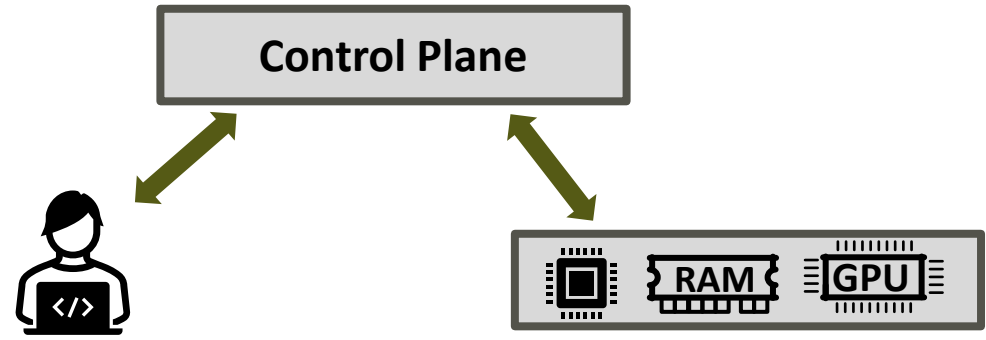


Software Abstraction

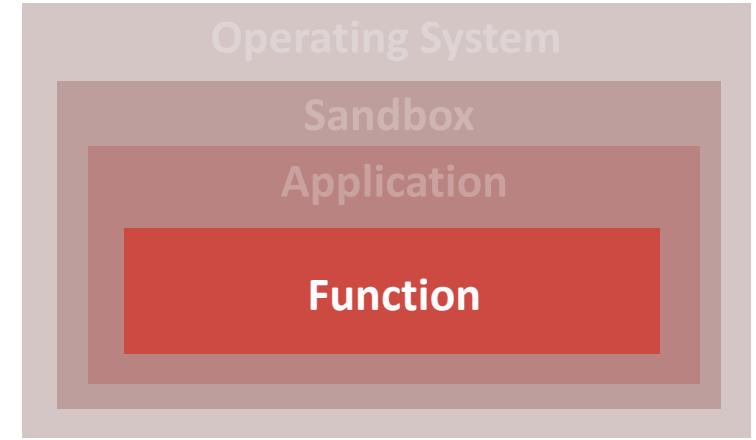


Serverless as an Answer

Hardware Abstraction

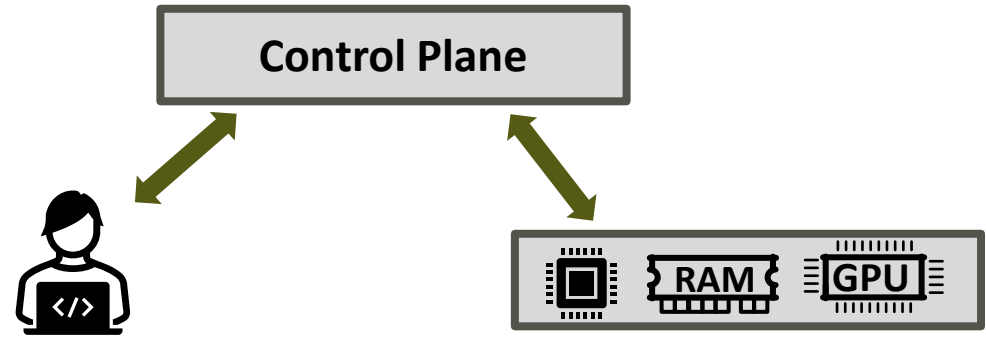


Software Abstraction

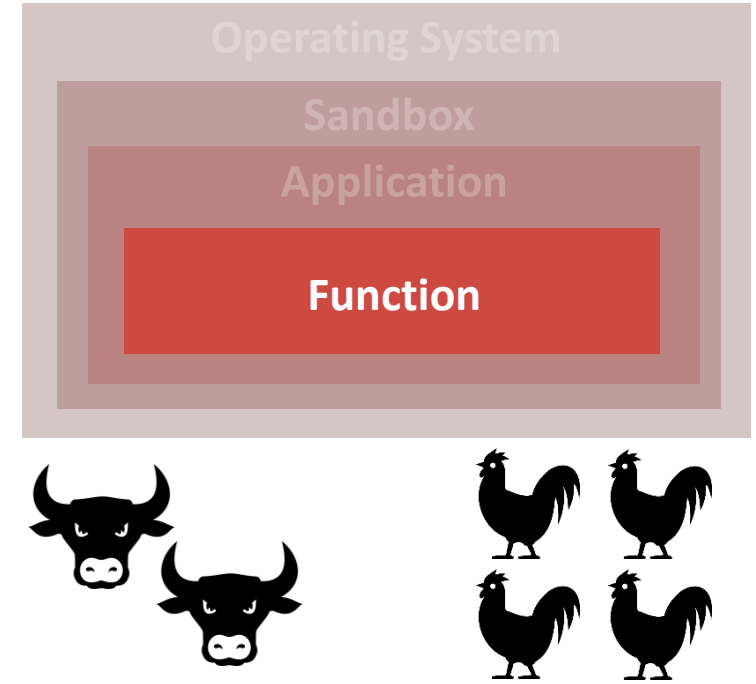


Serverless as an Answer

Hardware Abstraction

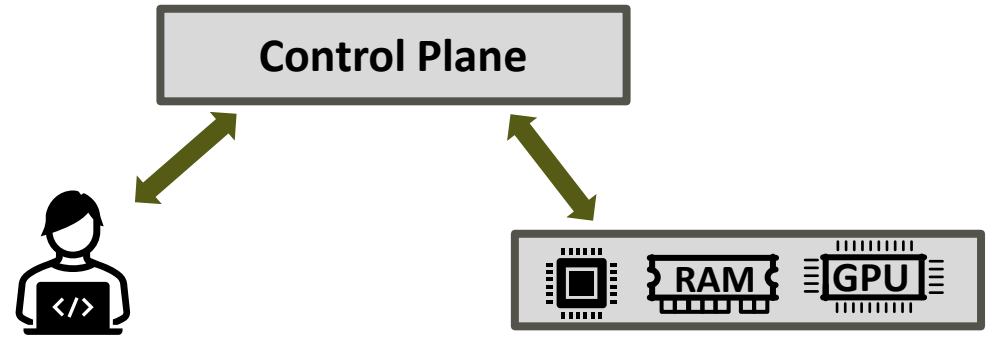


Software Abstraction



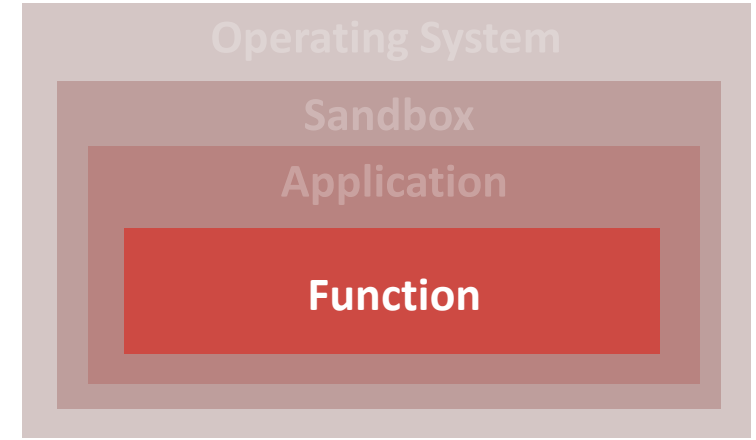
Serverless as an Answer

Hardware Abstraction



Pay-as-you-go billing

Software Abstraction



Granular computing

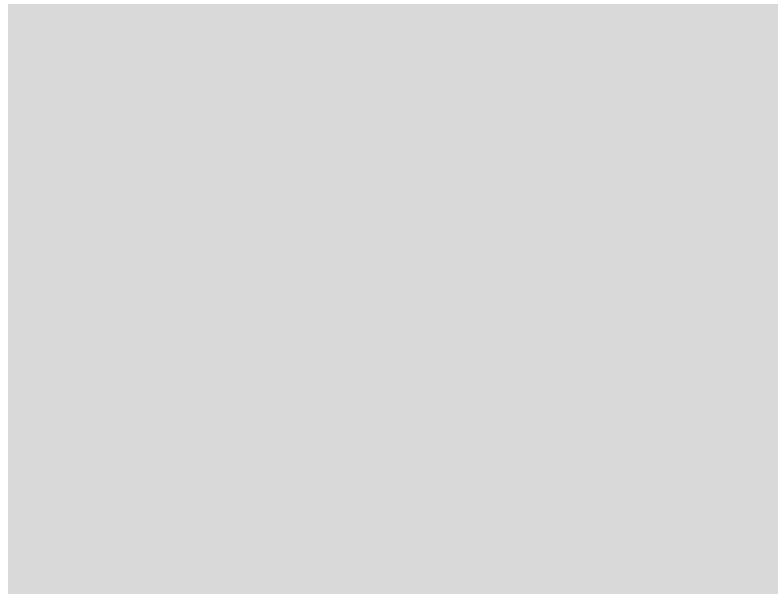
High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

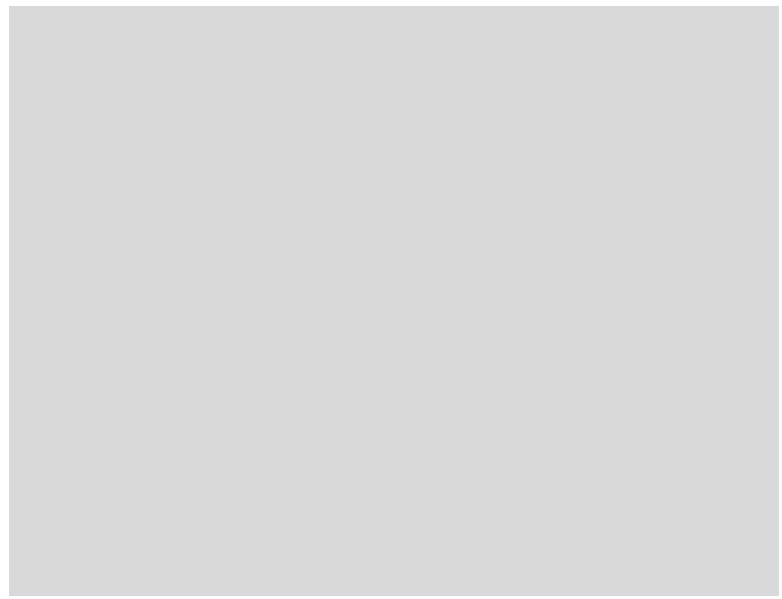
Marcin Copik Konstantin Taranov Alexandru Calotoiu Torsten Hoefler
ETH Zürich ETH Zürich ETH Zürich ETH Zürich

IPDPS, 2023

Cloud Server



Serverless-Enabled HPC Node



High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU

Serverless-Enabled HPC Node

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

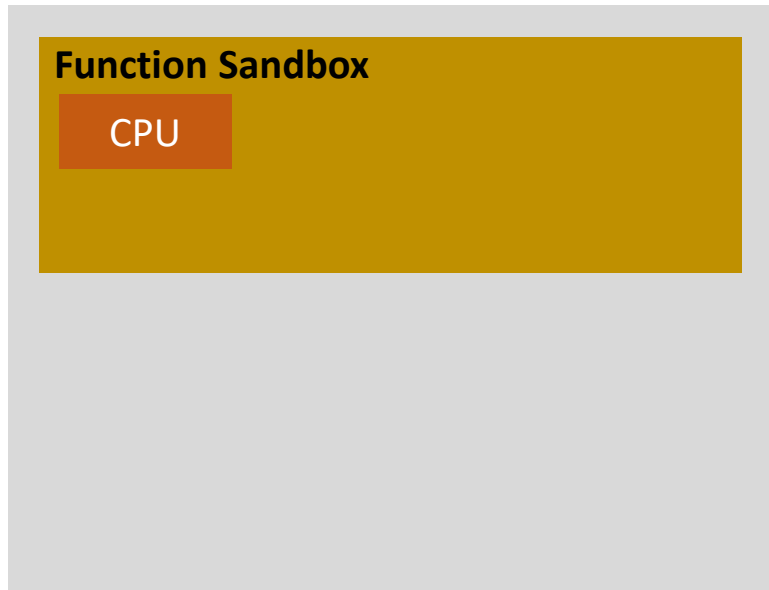
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

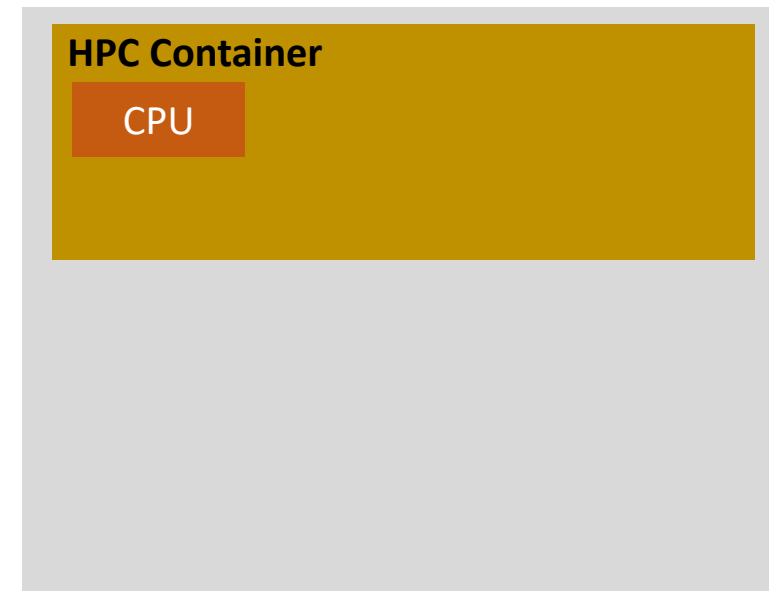
Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server



Serverless-Enabled HPC Node



Singularity, Sarus

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU

Serverless-Enabled HPC Node

HPC Container

CPU

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU



Cloud Services

Serverless-Enabled HPC Node

HPC Container

CPU

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

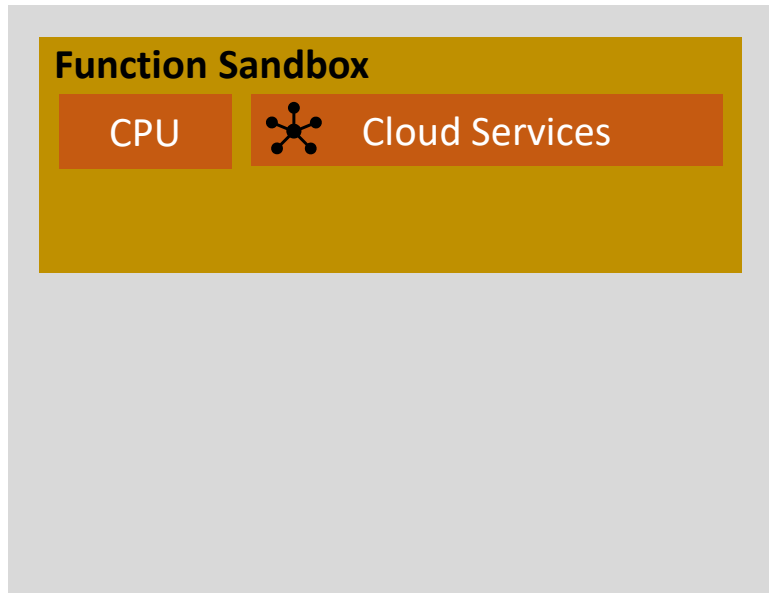
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

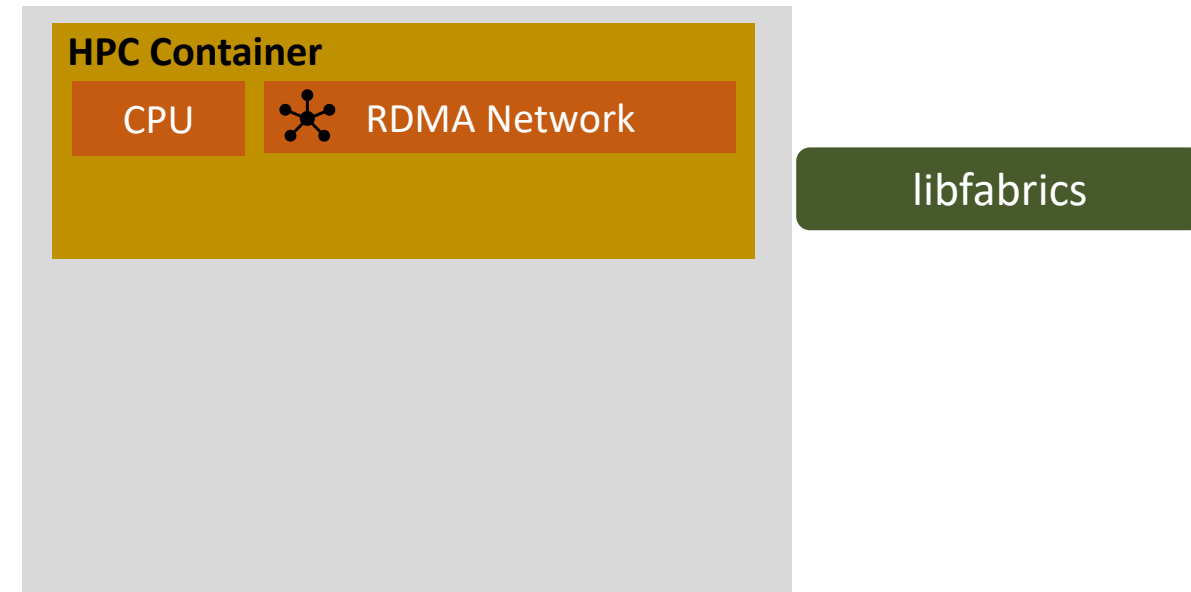
Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server



Serverless-Enabled HPC Node



High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU



Cloud Services

Serverless-Enabled HPC Node

HPC Container

CPU



RDMA Network

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU



Cloud Services



Cloud Storage

Serverless-Enabled HPC Node

HPC Container

CPU



RDMA Network

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU



Cloud Services



Cloud Storage

Serverless-Enabled HPC Node

HPC Container

CPU



RDMA Network



HPC Filesystem

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server

Function Sandbox

CPU



Cloud Services



Cloud Storage

Serverless-Enabled HPC Node

HPC Container

CPU



RDMA Network

GPU



HPC Filesystem

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

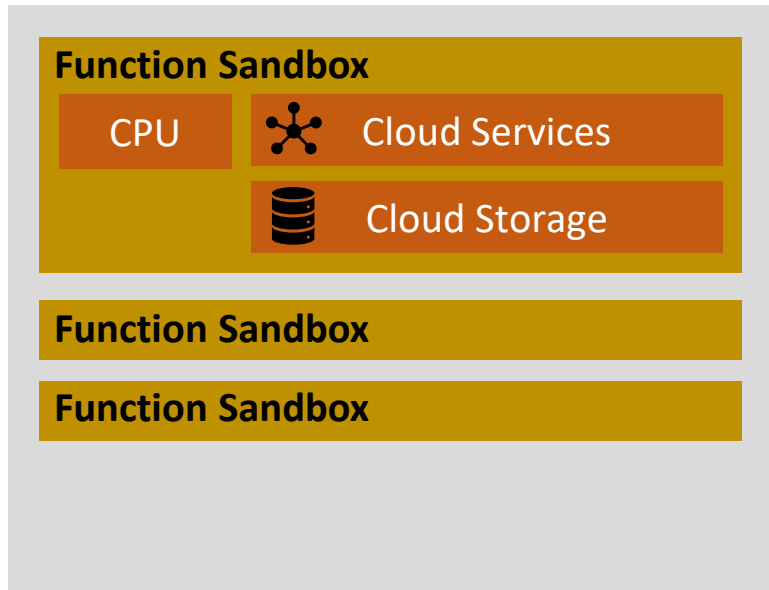
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

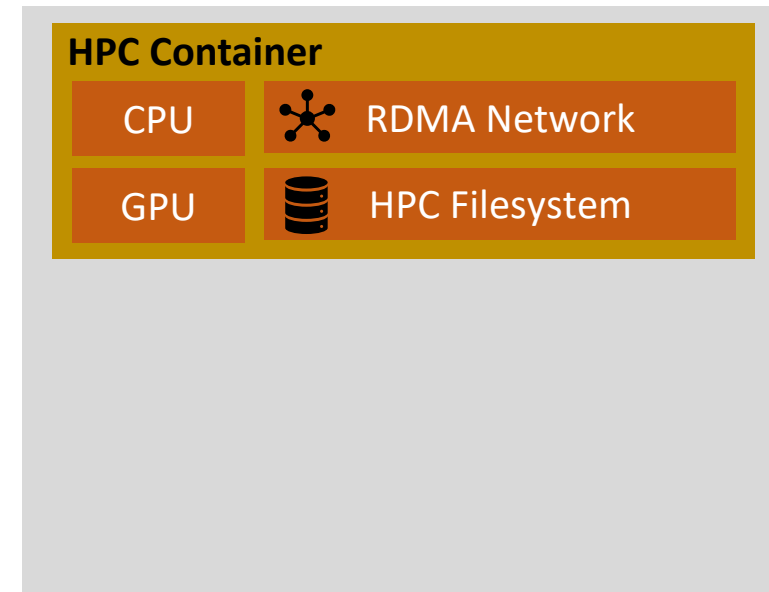
Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server



Serverless-Enabled HPC Node



High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

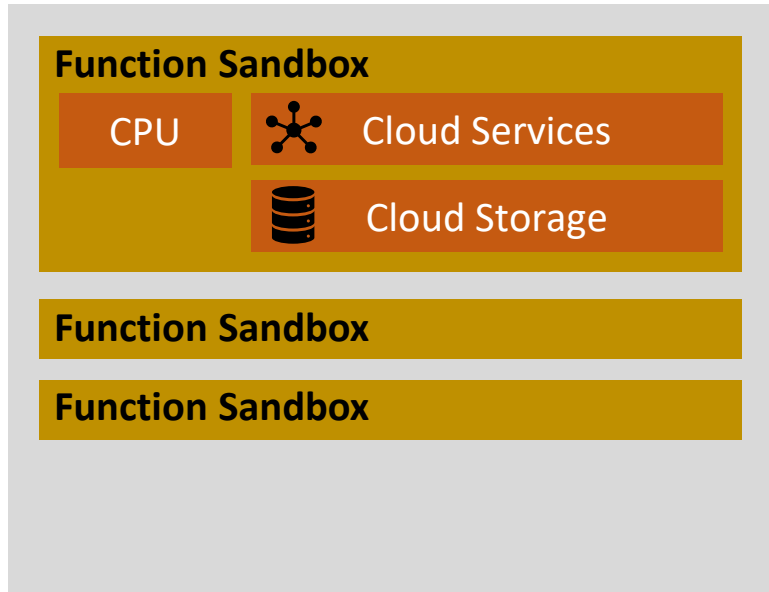
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

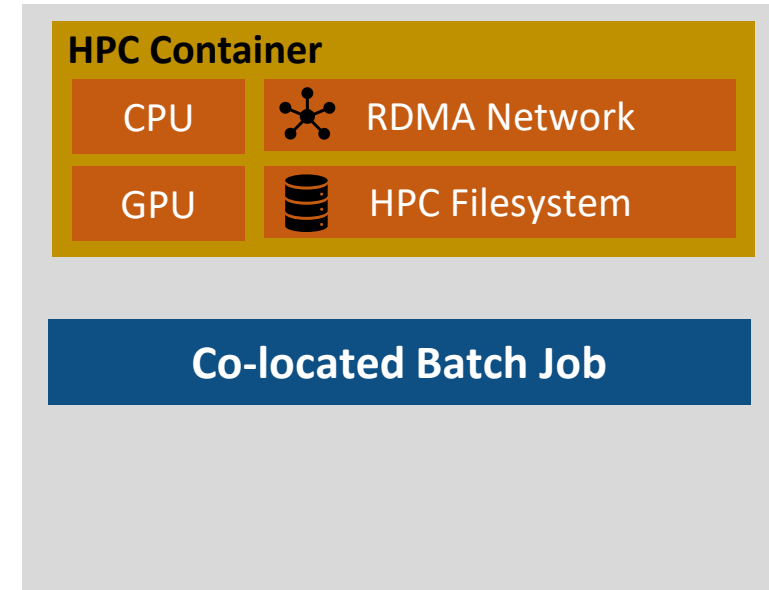
Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server



Serverless-Enabled HPC Node



High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

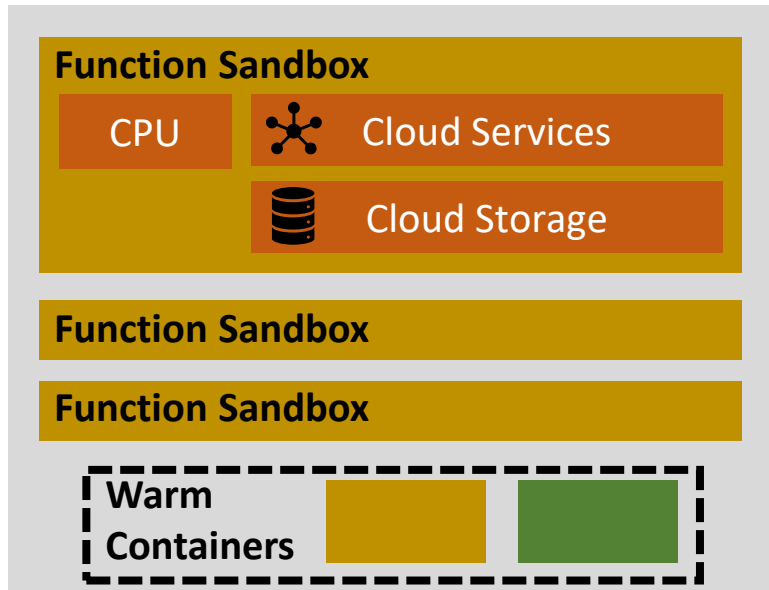
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

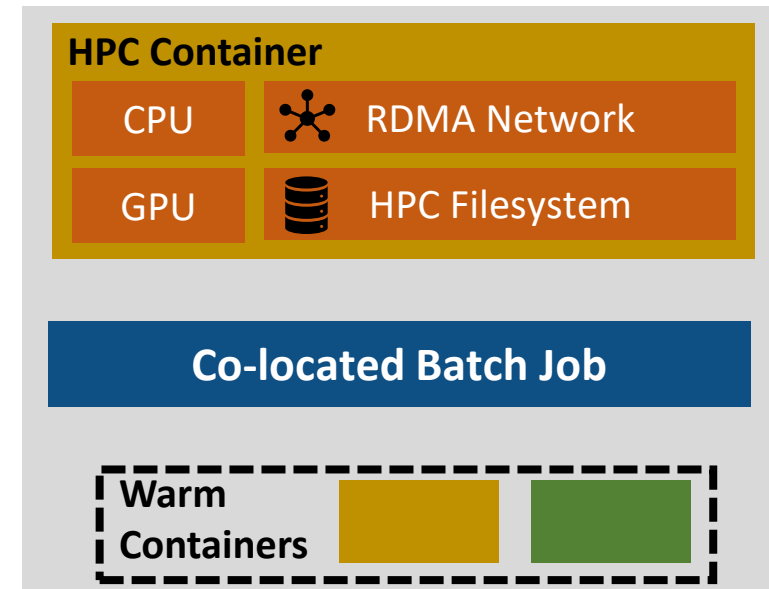
Torsten Hoefler
ETH Zürich

IPDPS, 2023

Cloud Server



Serverless-Enabled HPC Node



High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

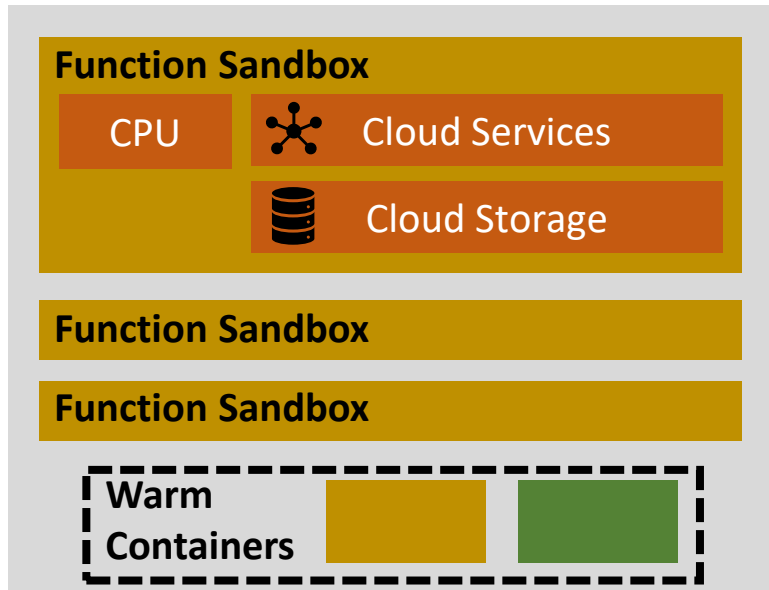
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

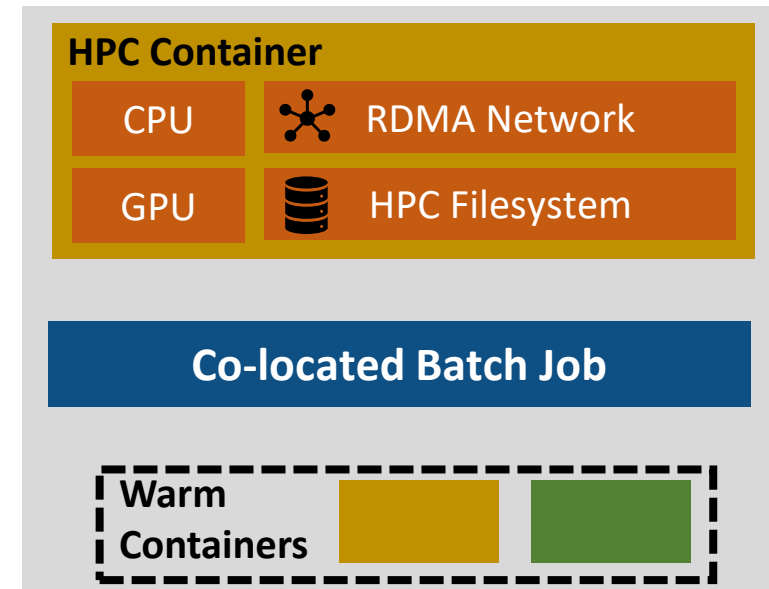
IPDPS, 2023

Cloud Server



Cloud Control Plane

Serverless-Enabled HPC Node



HPC Batch System

High-Performance FaaS with rFaaS

rFaaS: Enabling High Performance Serverless with RDMA and Decentralization.

Marcin Copik
ETH Zürich

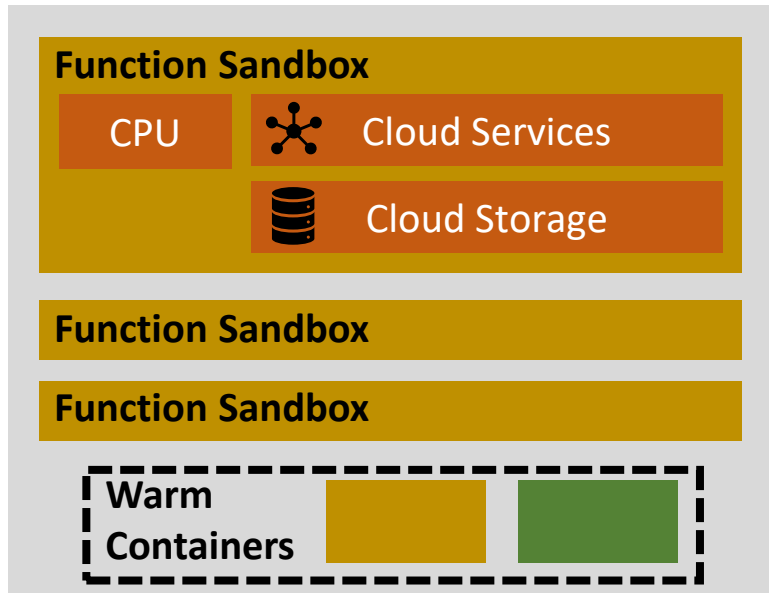
Konstantin Taranov
ETH Zürich

Alexandru Calotoiu
ETH Zürich

Torsten Hoefler
ETH Zürich

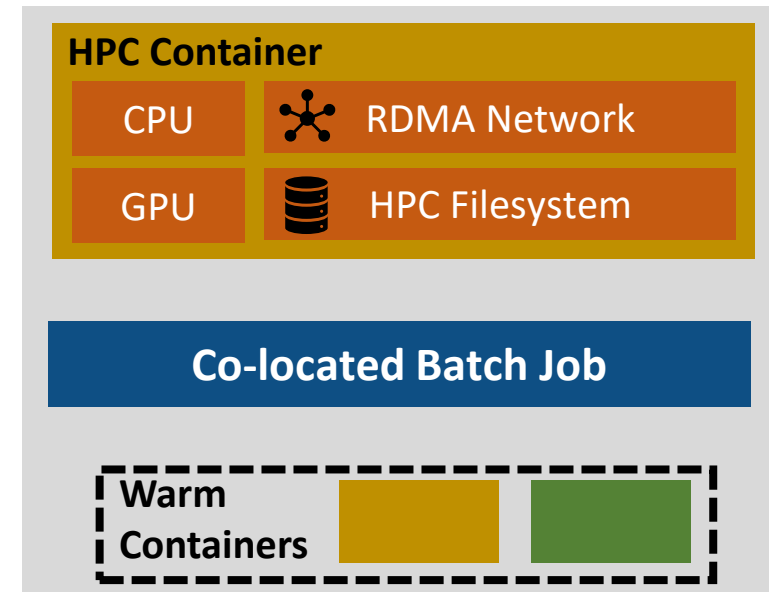
IPDPS, 2023

Cloud Server



Cloud Control Plane

Serverless-Enabled HPC Node



HPC Batch System

rFaaS Resource Manager

Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

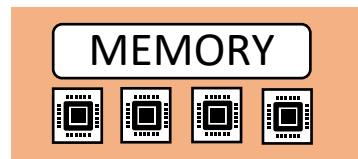
Node

Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

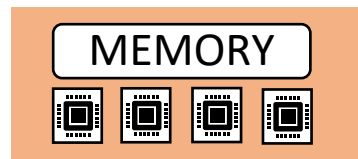
Node



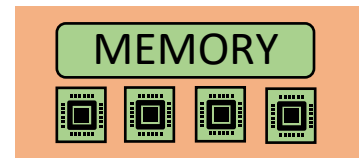
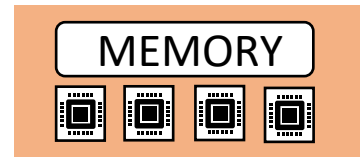
Serverless Disaggregation

Batch jobs

Node



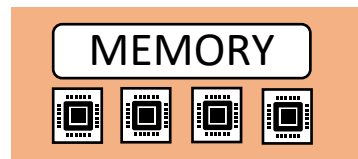
Batch jobs + serverless workloads



Serverless Disaggregation

Batch jobs

Node



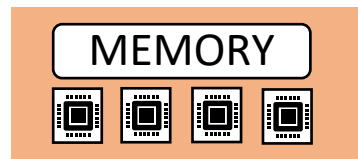
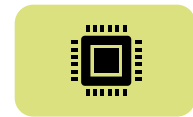
Batch jobs + serverless workloads



Serverless Disaggregation

Batch jobs

Node



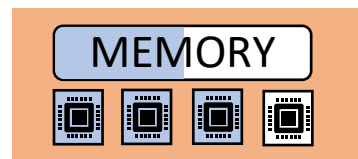
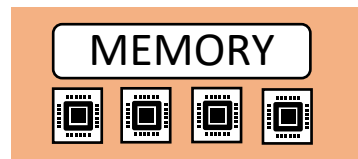
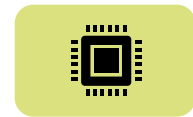
Batch jobs + serverless workloads



Serverless Disaggregation

Batch jobs

Node



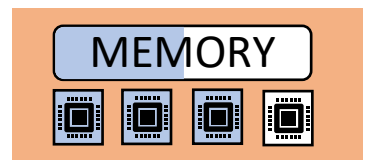
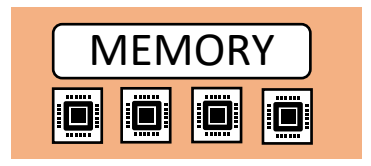
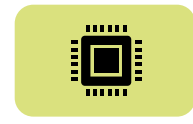
Batch jobs + serverless workloads



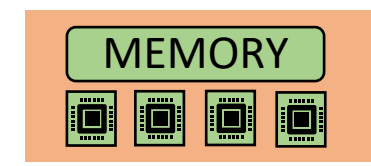
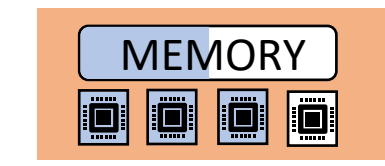
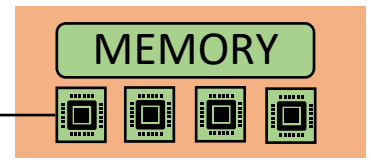
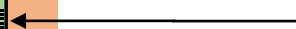
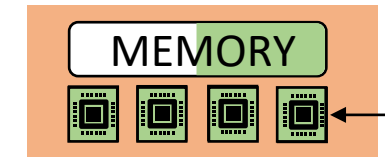
Serverless Disaggregation

Batch jobs

Node



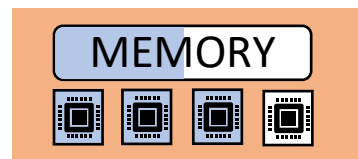
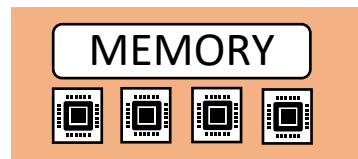
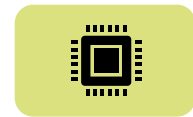
Batch jobs + serverless workloads



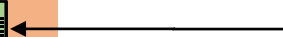
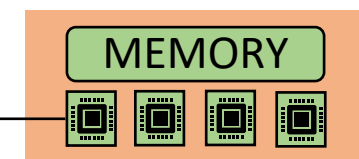
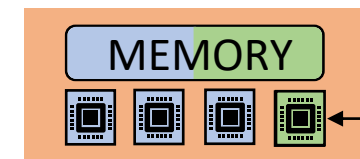
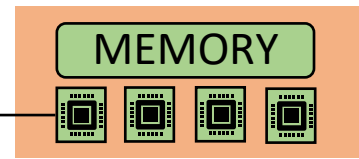
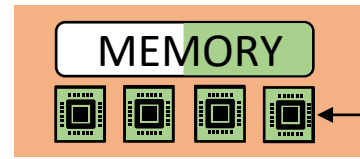
Serverless Disaggregation

Batch jobs

Node



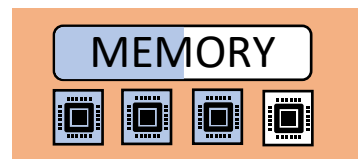
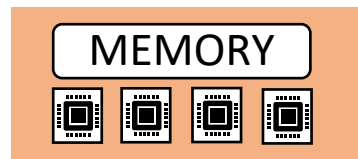
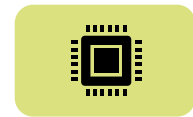
Batch jobs + serverless workloads



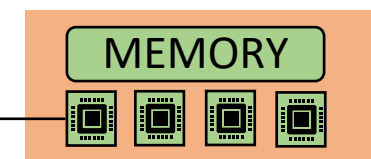
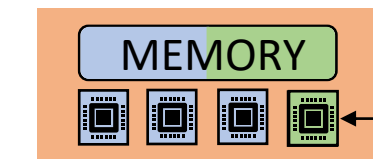
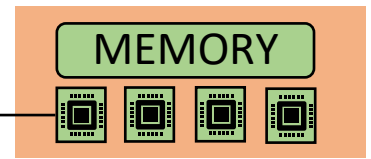
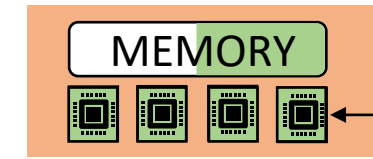
Serverless Disaggregation

Batch jobs

Node



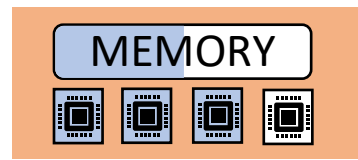
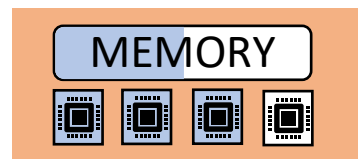
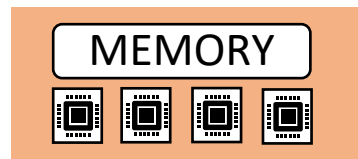
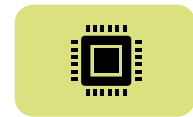
Batch jobs + serverless workloads



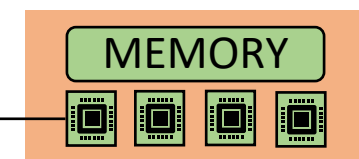
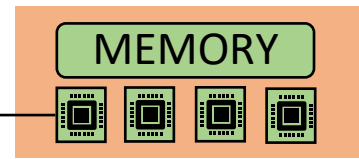
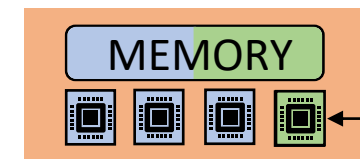
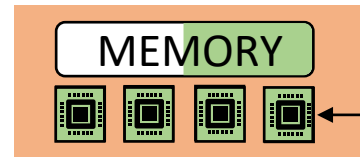
Serverless Disaggregation

Batch jobs

Node



Batch jobs + serverless workloads

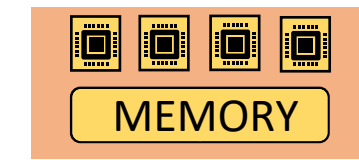
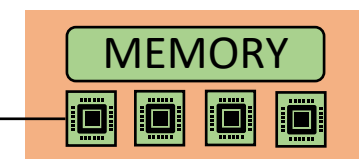
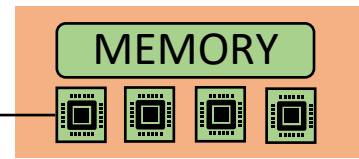
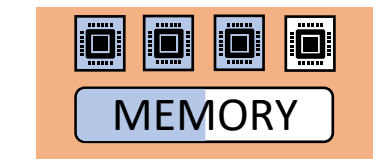
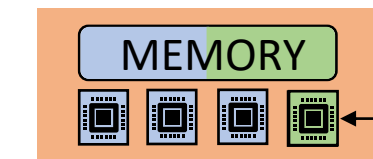
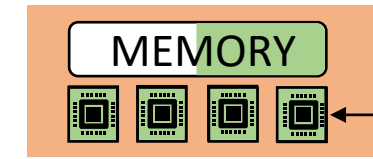
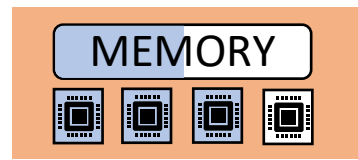
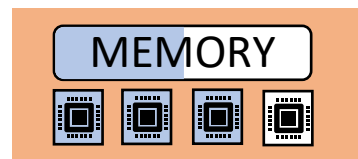
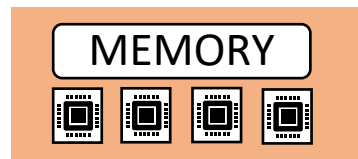
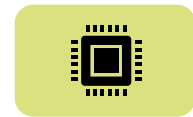


Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

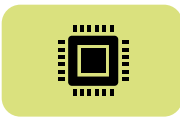

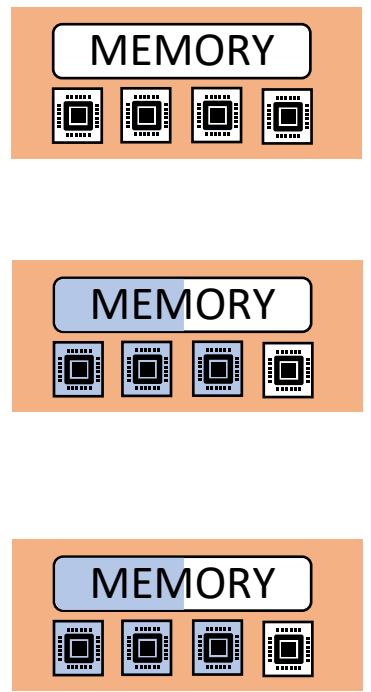
Node



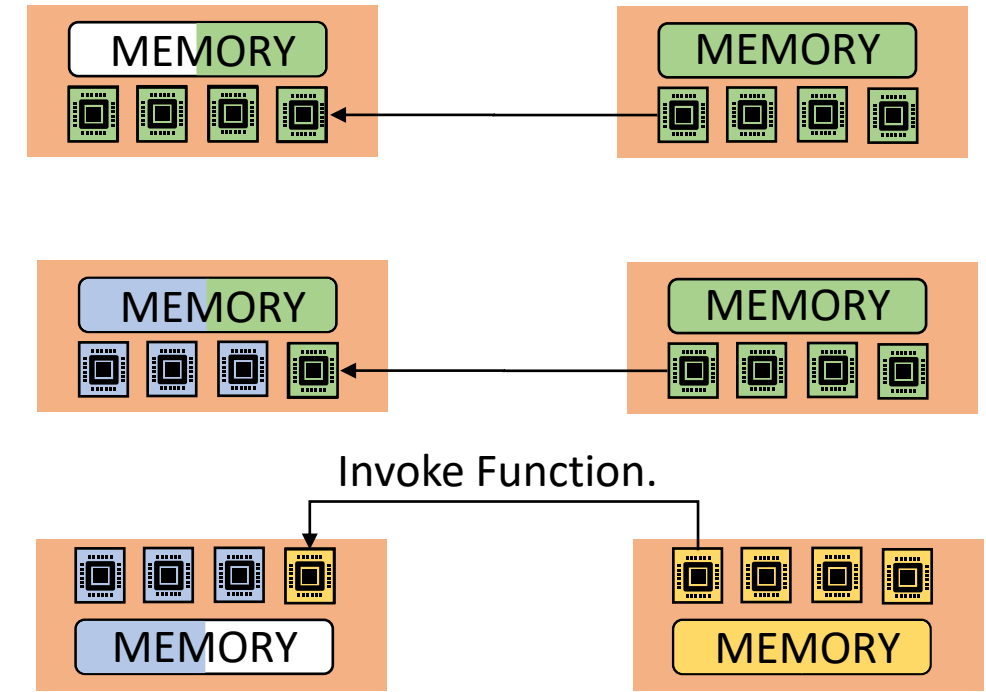
Serverless Disaggregation

Batch jobs

Node

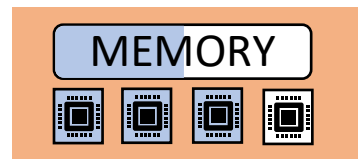
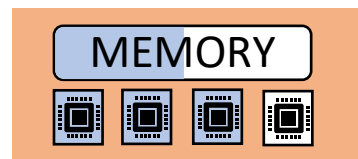
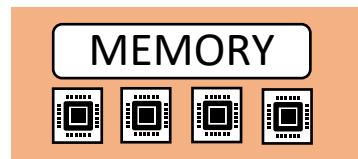
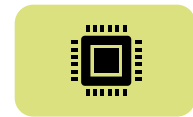
Batch jobs + serverless workloads



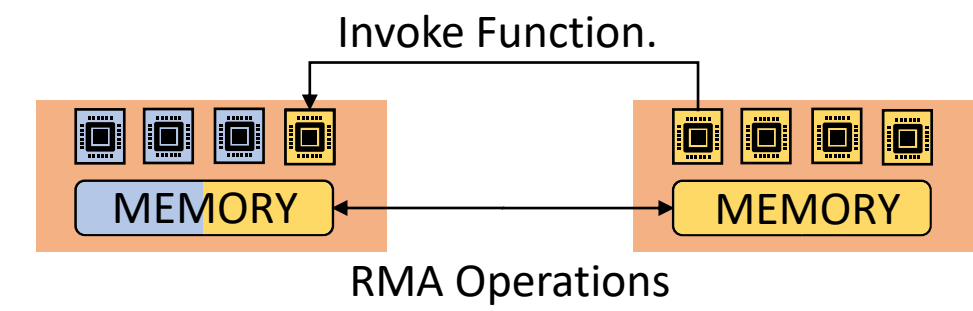
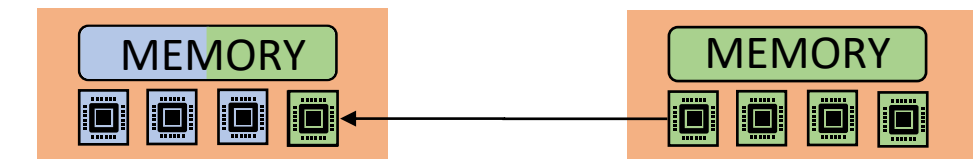
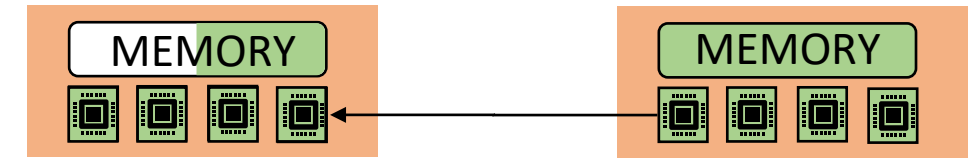
Serverless Disaggregation

Batch jobs

Node

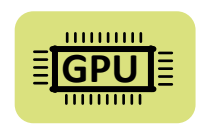
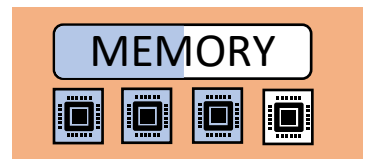
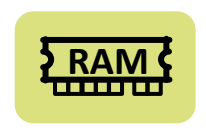
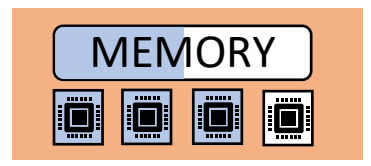
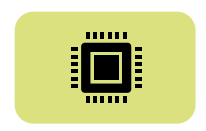
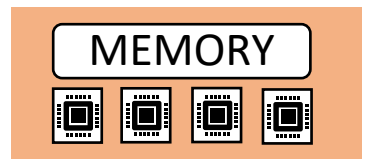
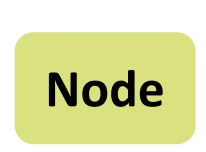


Batch jobs + serverless workloads

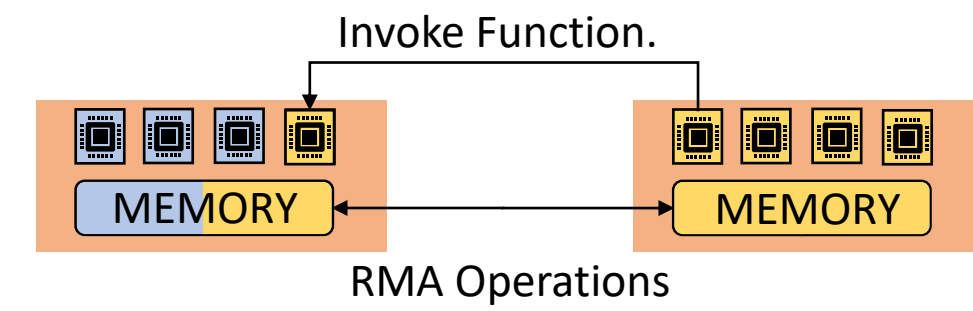
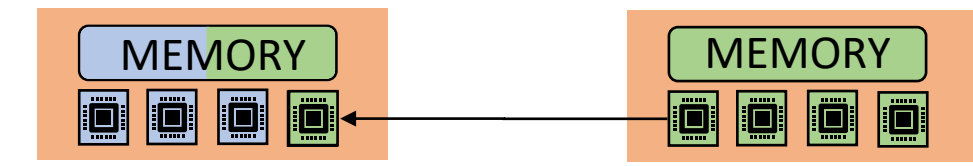
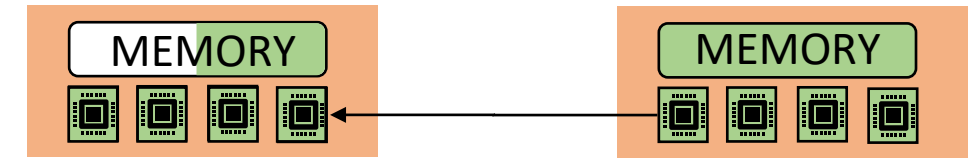


Serverless Disaggregation

Batch jobs



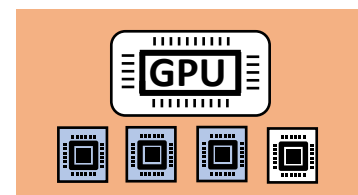
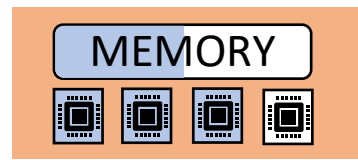
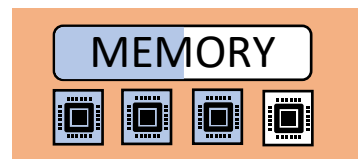
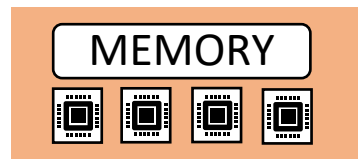
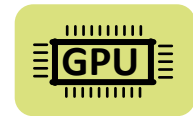
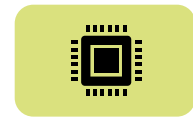
Batch jobs + serverless workloads



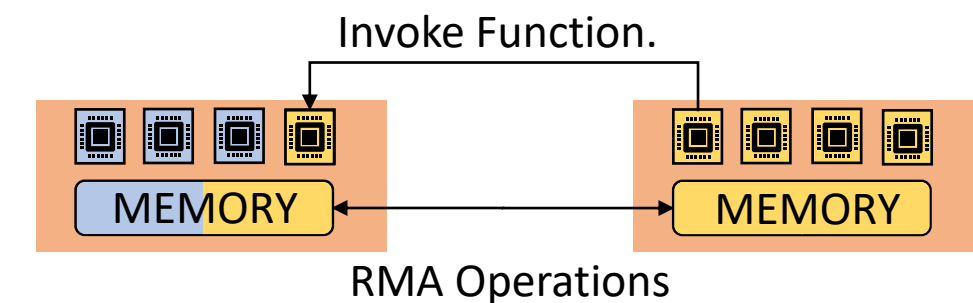
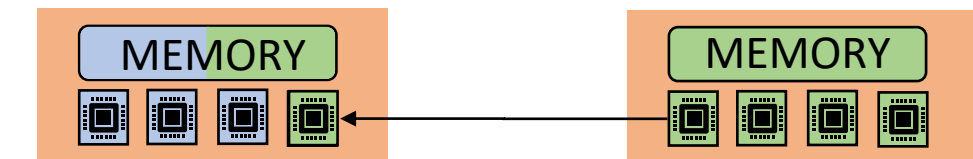
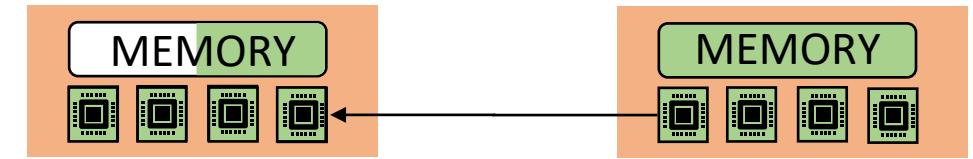
Serverless Disaggregation

Batch jobs

Node



Batch jobs + serverless workloads

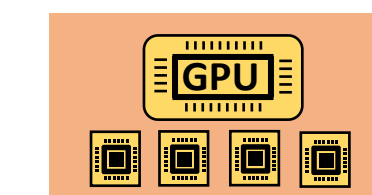
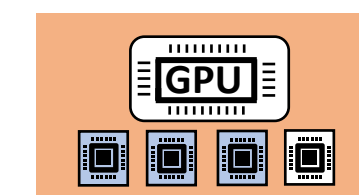
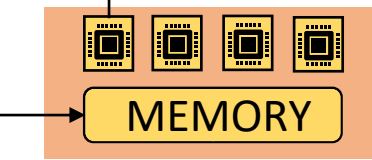
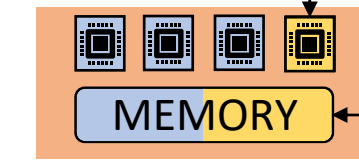
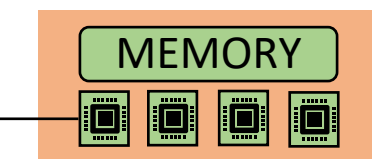
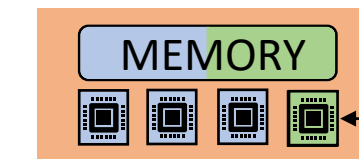
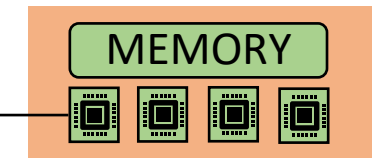
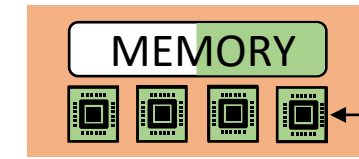
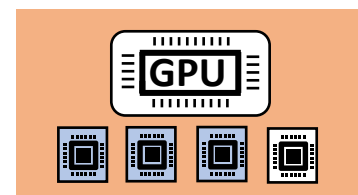
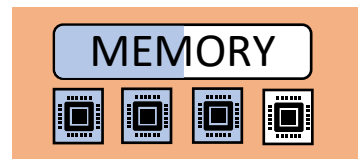
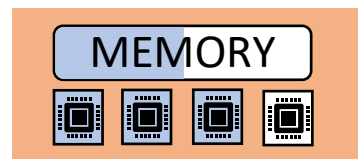
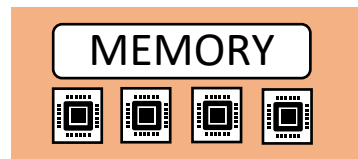
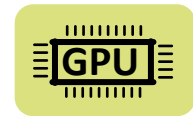
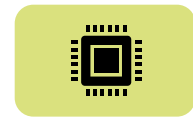


Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

Node



Invoke Function.

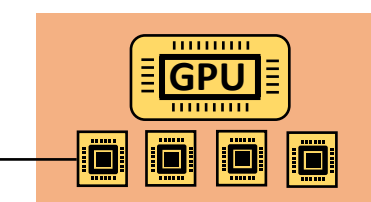
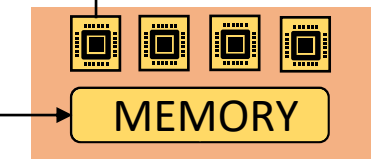
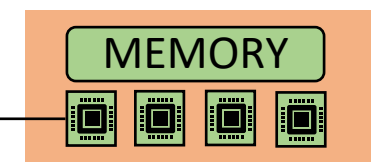
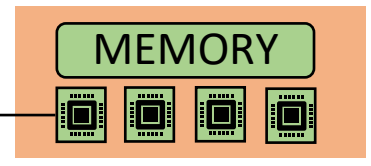
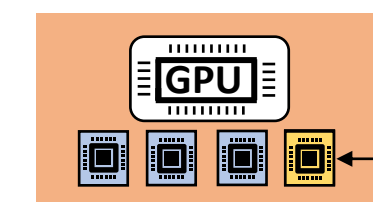
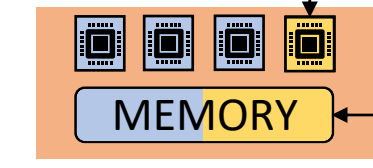
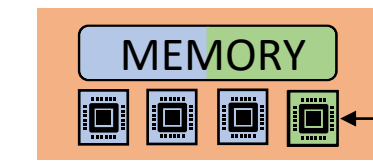
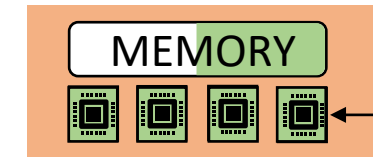
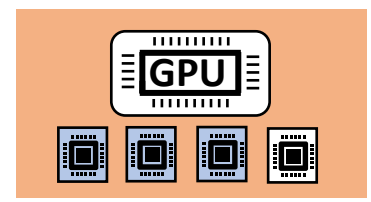
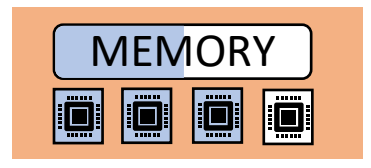
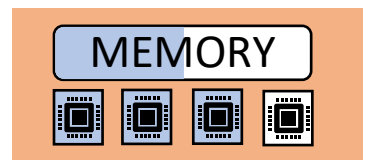
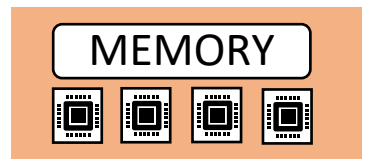
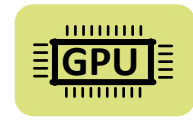
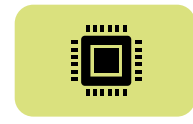
RMA Operations

Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

Node



Invoke Function.



RMA Operations

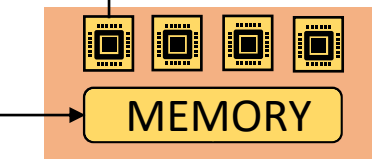
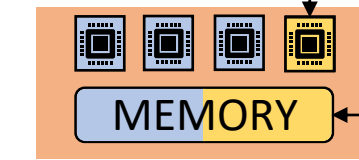
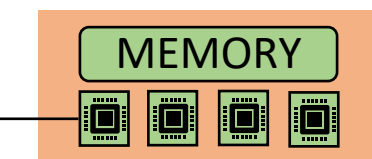
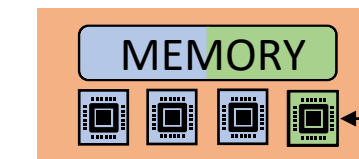
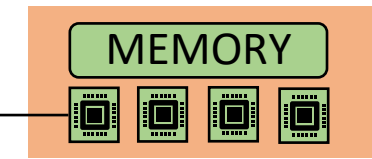
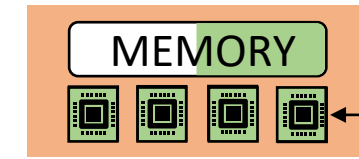
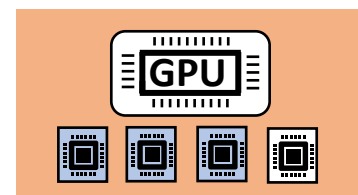
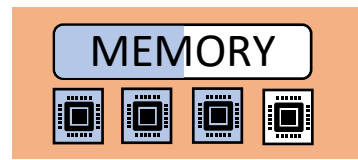
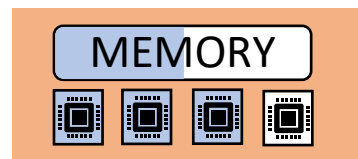
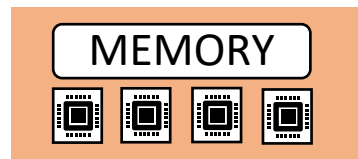
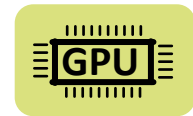
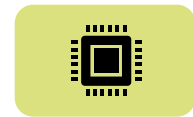


Serverless Disaggregation

Batch jobs

Batch jobs + serverless workloads

Node



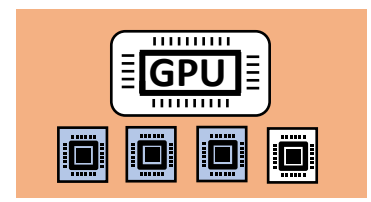
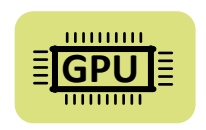
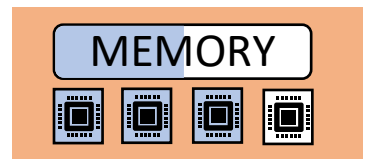
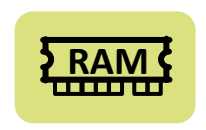
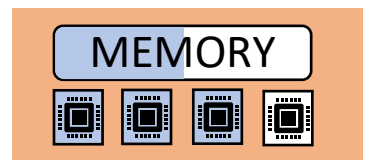
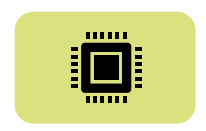
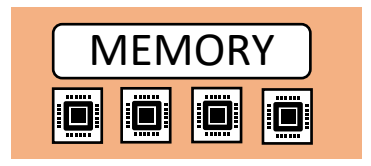
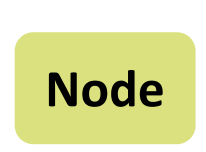
Invoke Function.



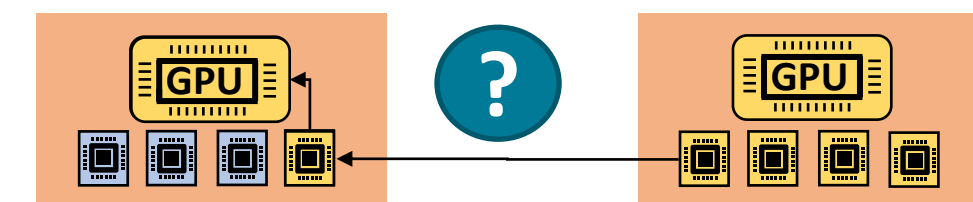
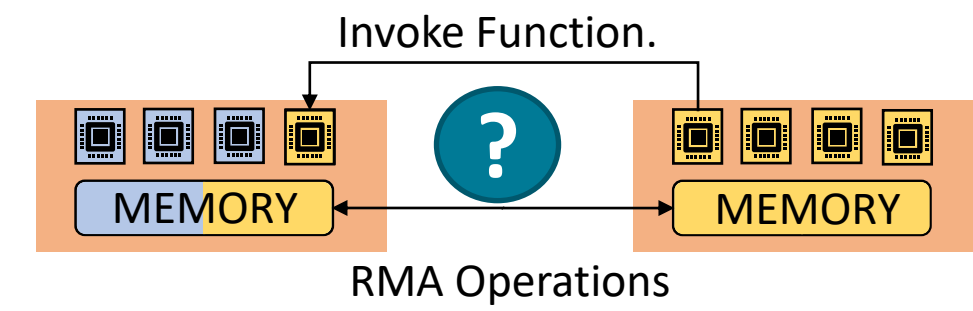
RMA Operations

Serverless Disaggregation

Batch jobs



Batch jobs + serverless workloads



Evaluation

Evaluation



XC50 nodes - 12 CPU cores, GPU, 64 GB memory.

XC40 nodes - 36 CPU cores, 64/128 GB memory.

Cray Aries interconnect.

Evaluation



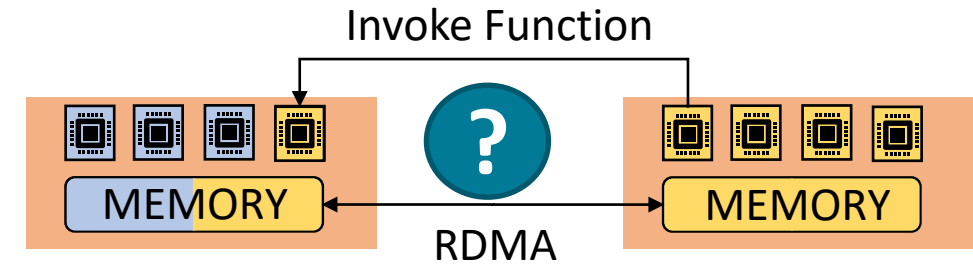
XC50 nodes - 12 CPU cores, GPU, 64 GB memory.

XC40 nodes - 36 CPU cores, 64/128 GB memory.

Cray Aries interconnect.

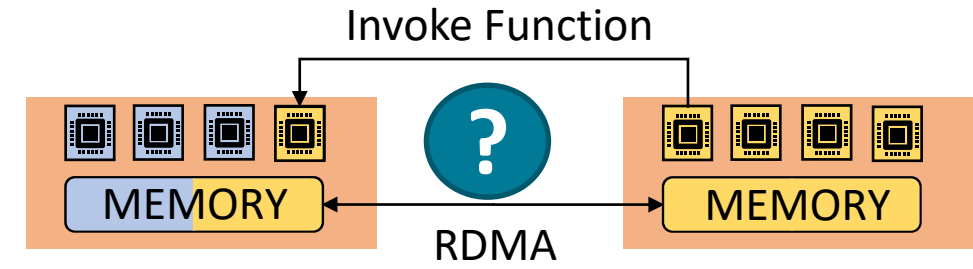
36 CPU cores, 377 GB memory.
Ethernet with RoCEv2 support.

Serving Remote Memory

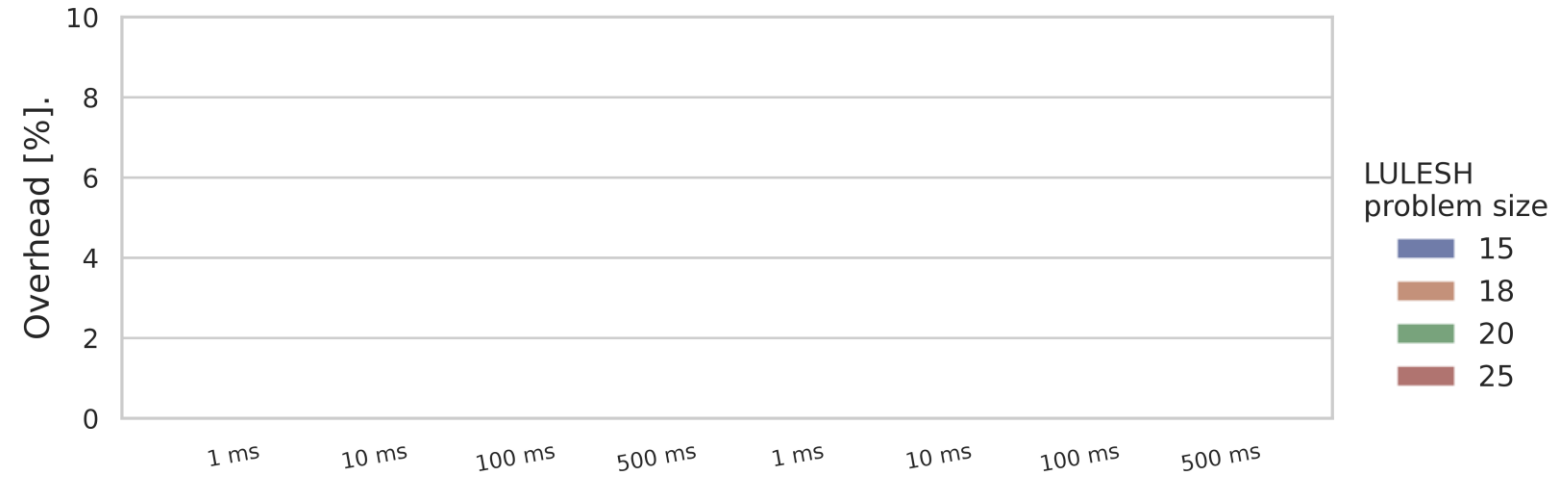


LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

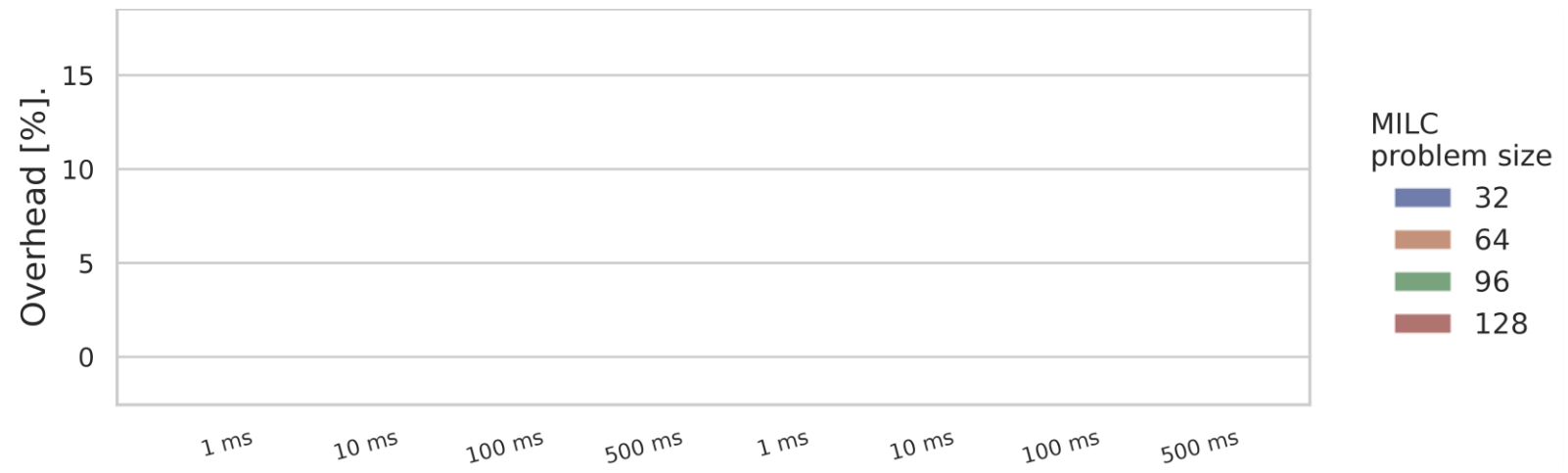
Serving Remote Memory



LULESH
125 ranks

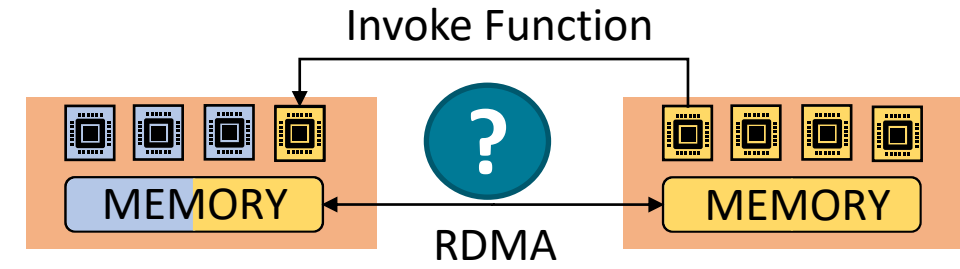


MILC
32 ranks

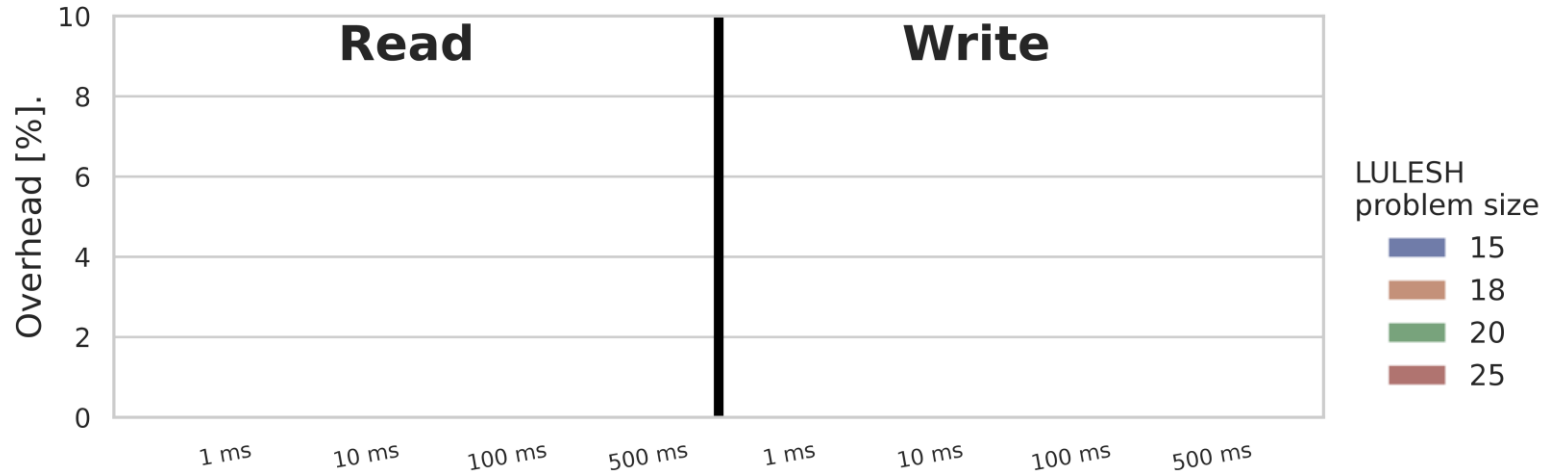


LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

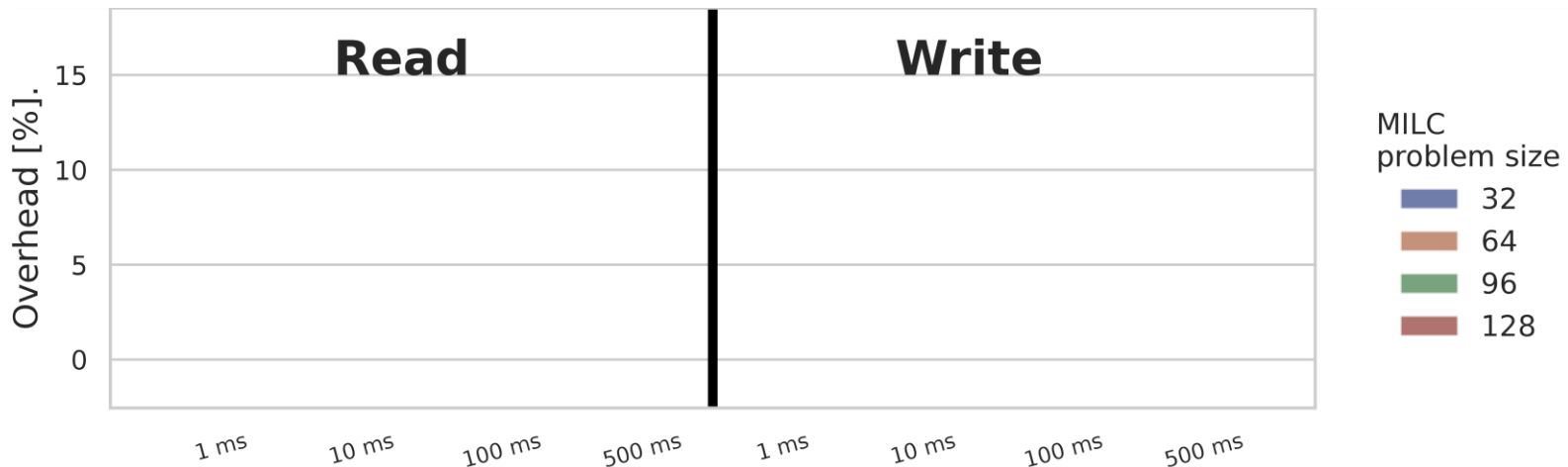
Serving Remote Memory



LULESH
125 ranks

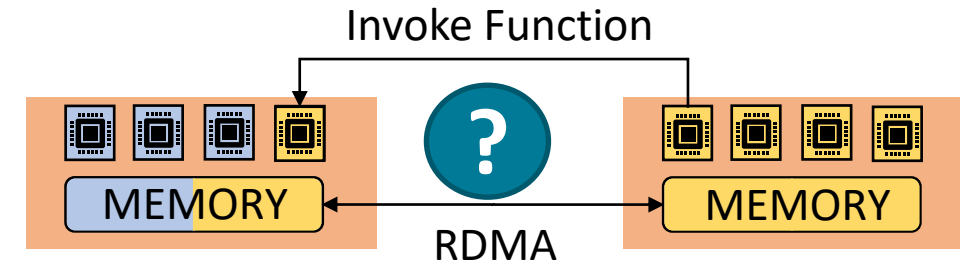


MILC
32 ranks

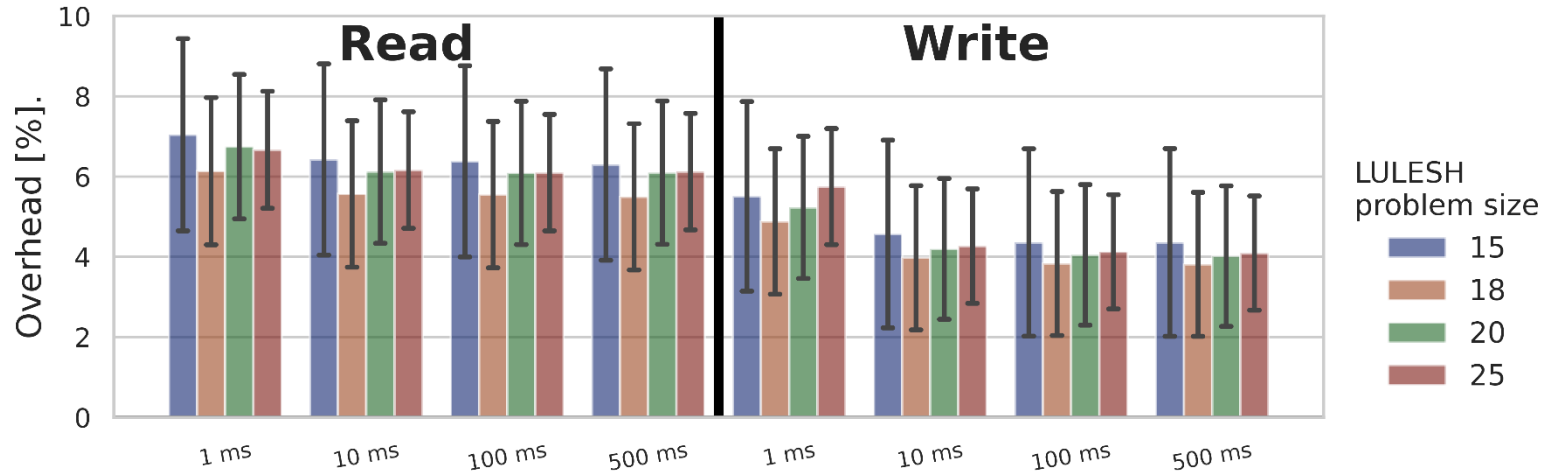


LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

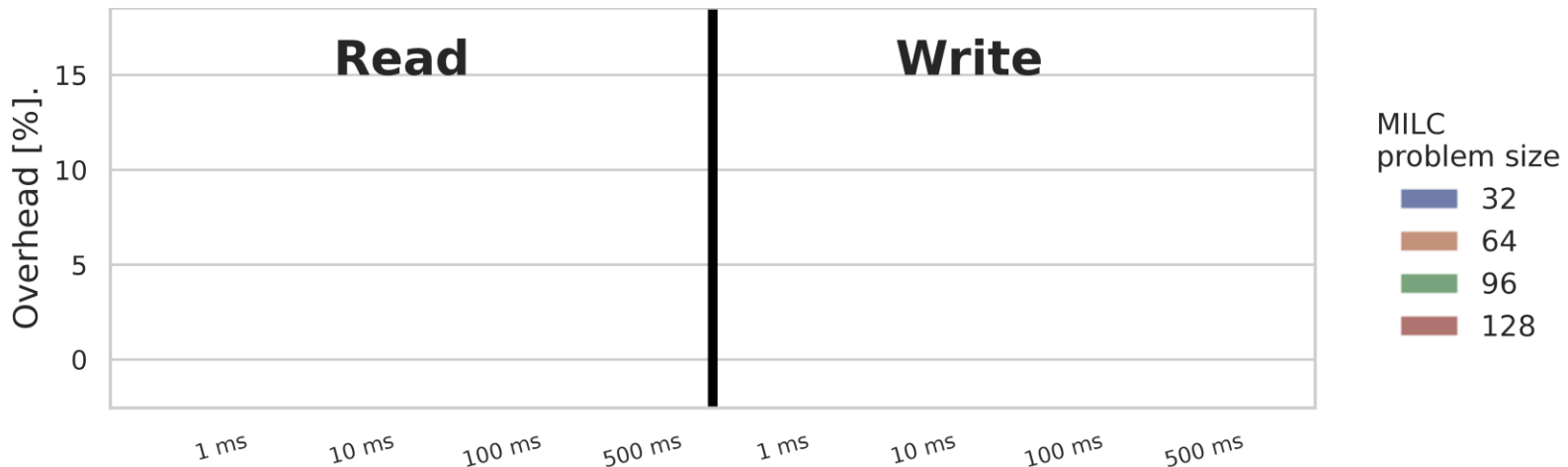
Serving Remote Memory



LULESH
125 ranks

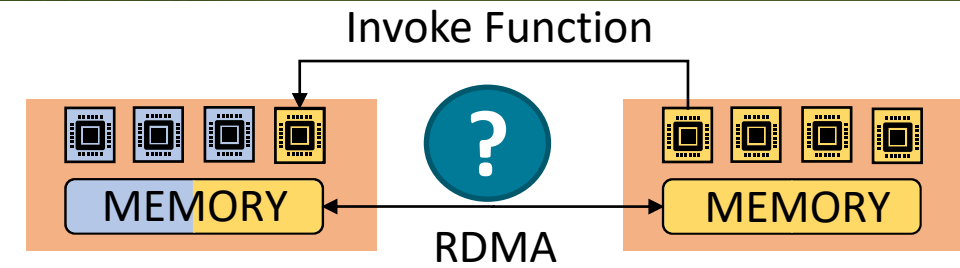


MILC
32 ranks

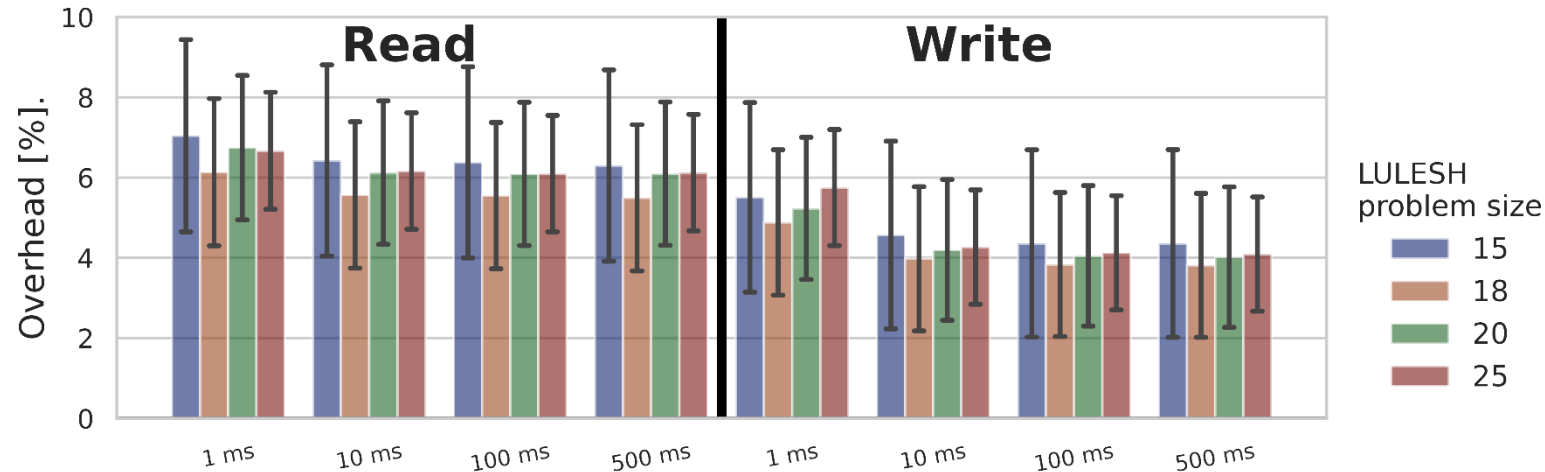


LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

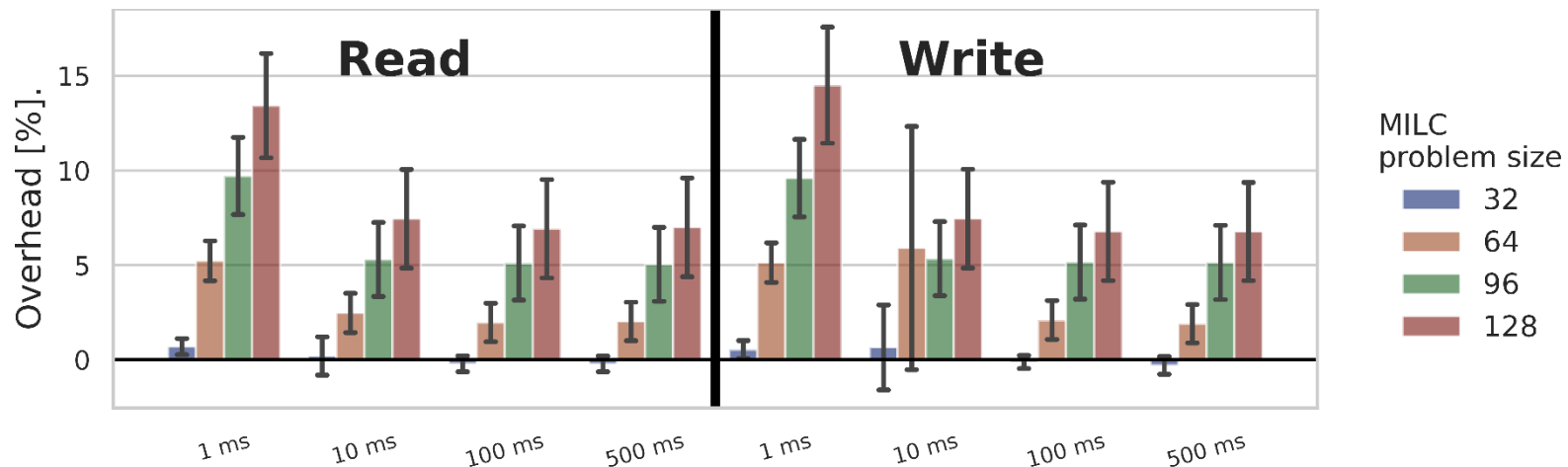
Serving Remote Memory



LULESH
125 ranks

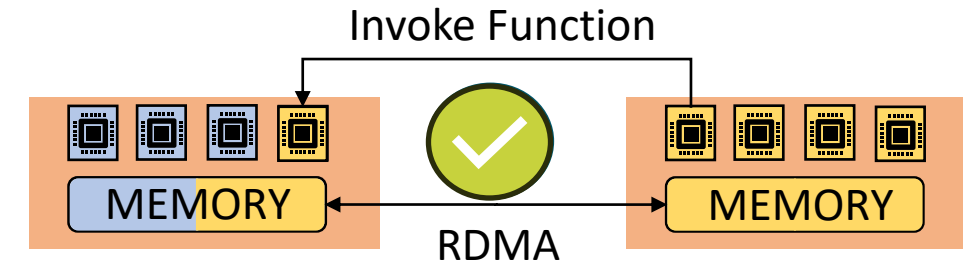


MILC
32 ranks

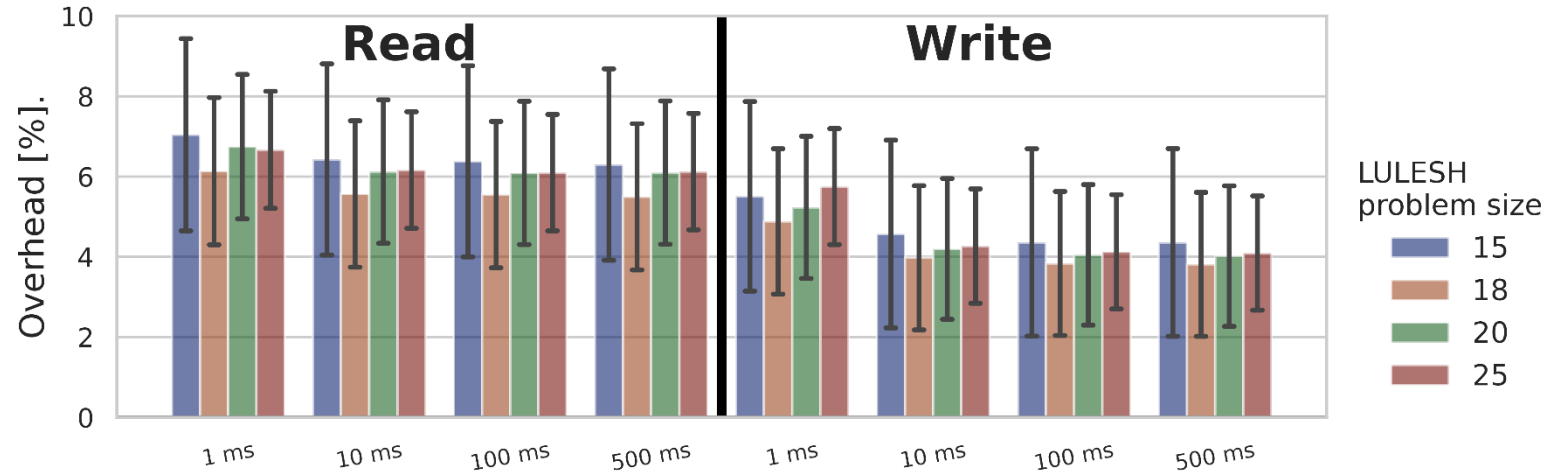


LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

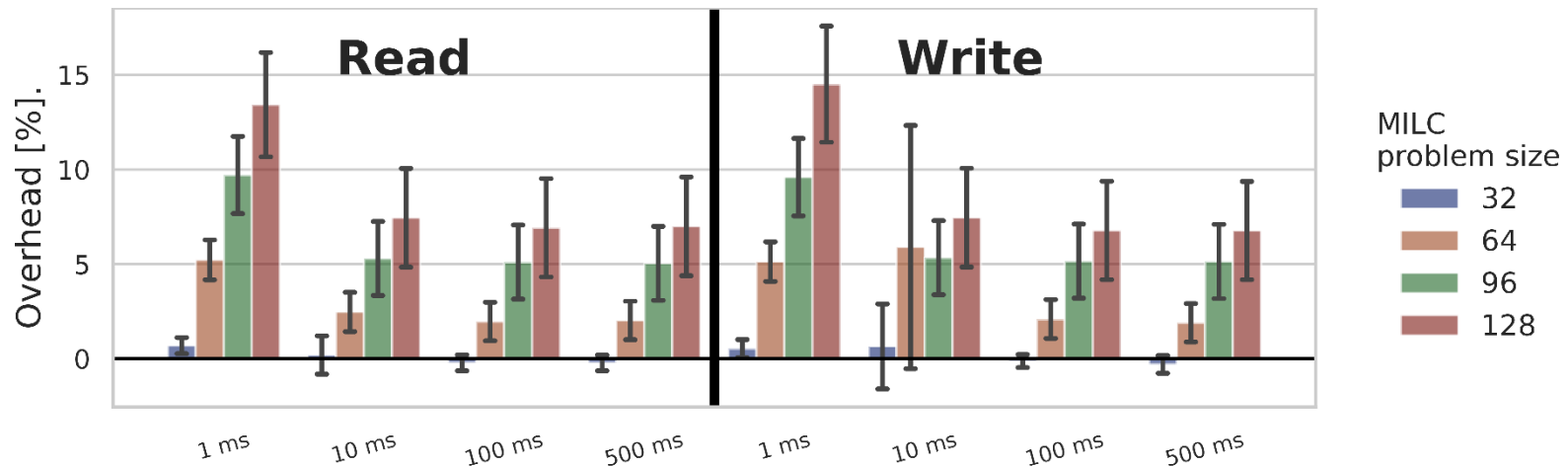
Serving Remote Memory



LULESH
125 ranks

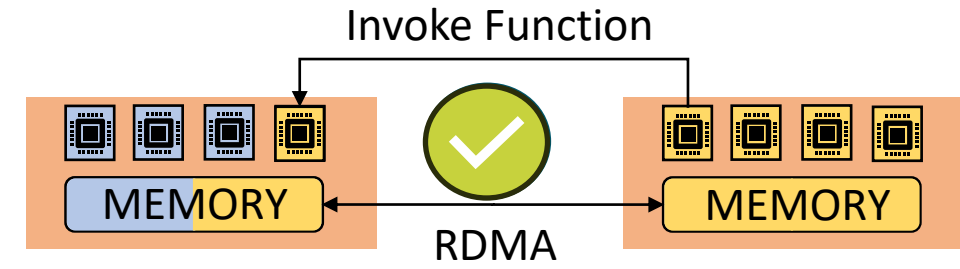


MILC
32 ranks

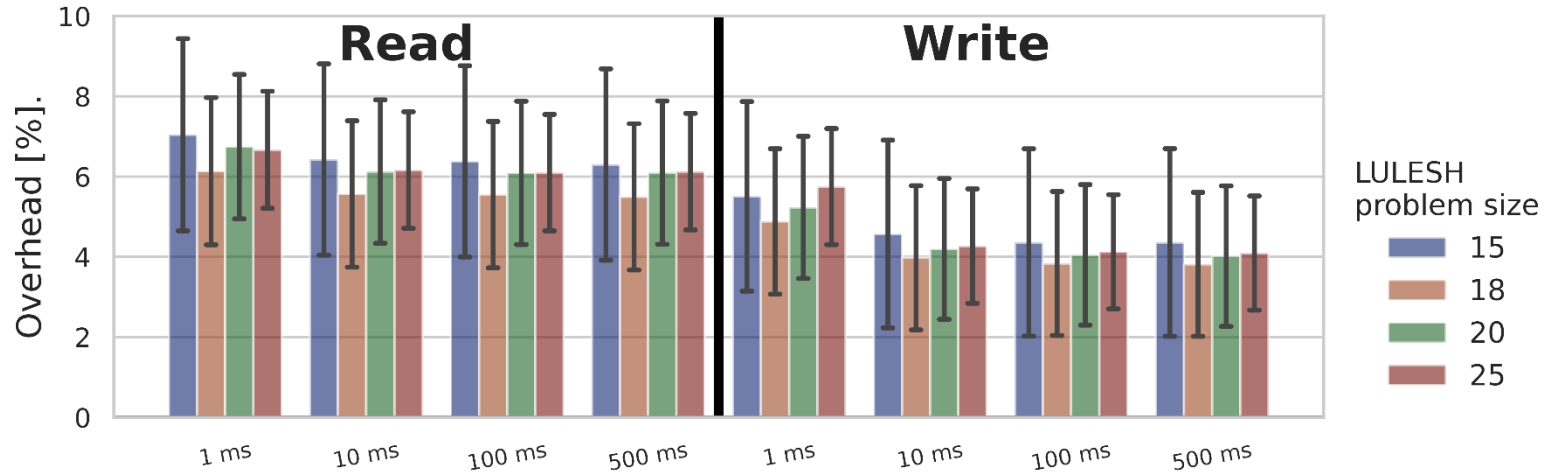


LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

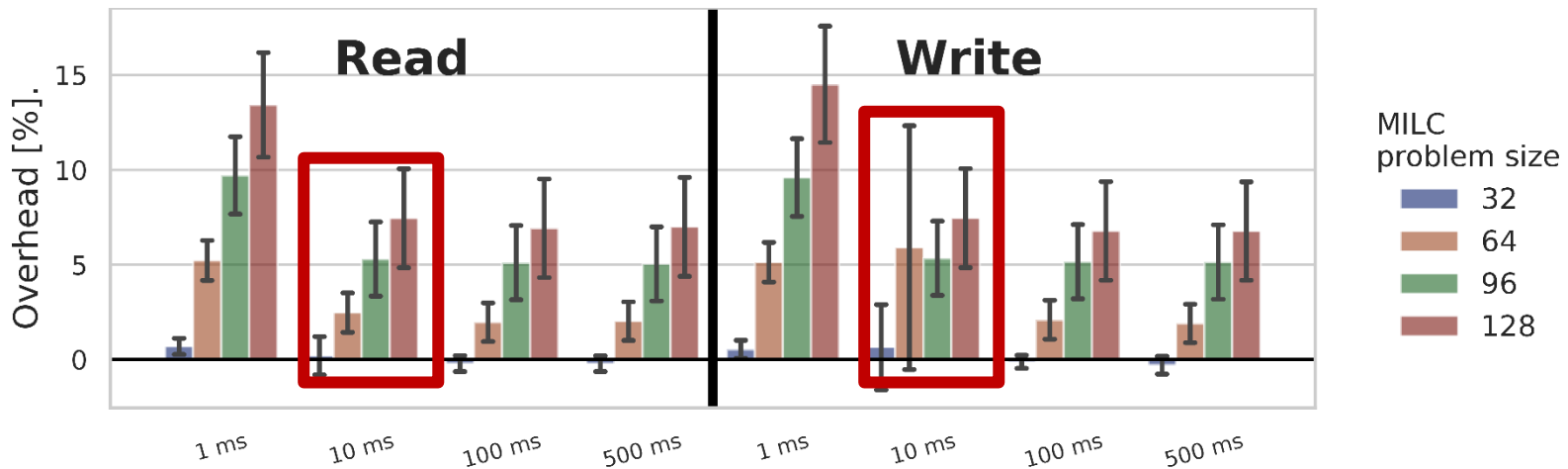
Serving Remote Memory



LULESH
125 ranks



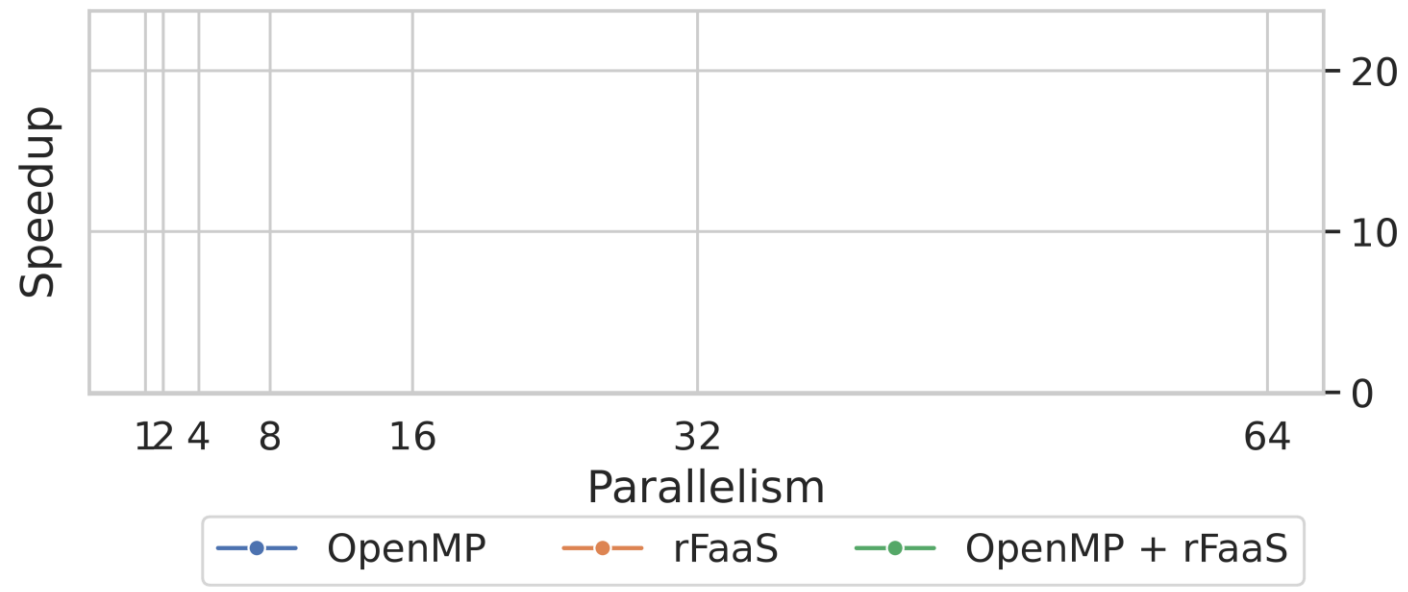
MILC
32 ranks



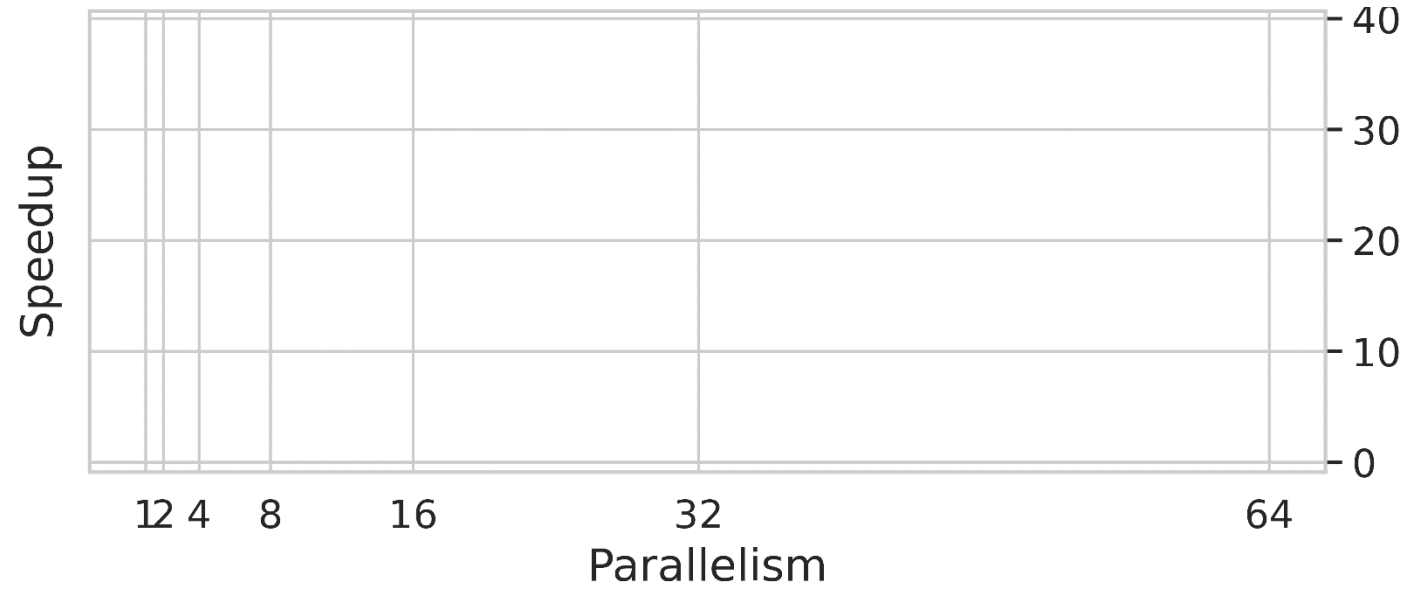
LULESH, MILC – 32 ranks, 1 node, 32 out of 36 cores allocated.

HPC Computations with FaaS: OpenMC

**1,000
Particles**

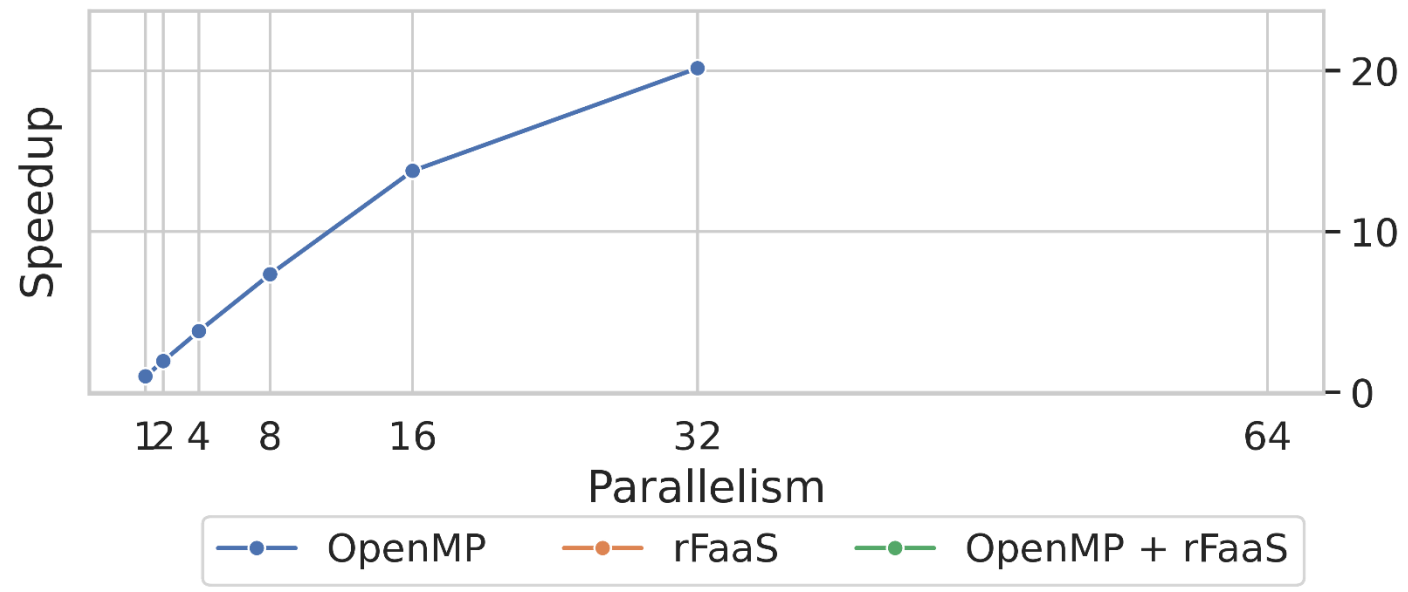


**10,000
Particles**

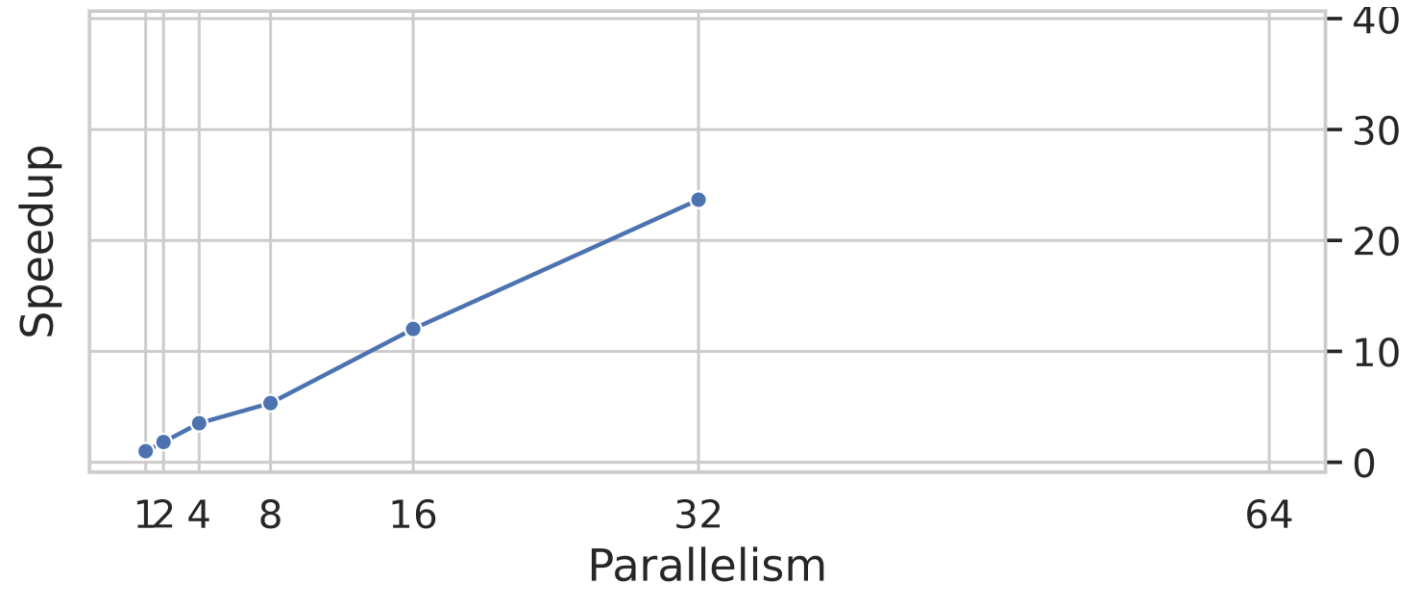


HPC Computations with FaaS: OpenMC

**1,000
Particles**

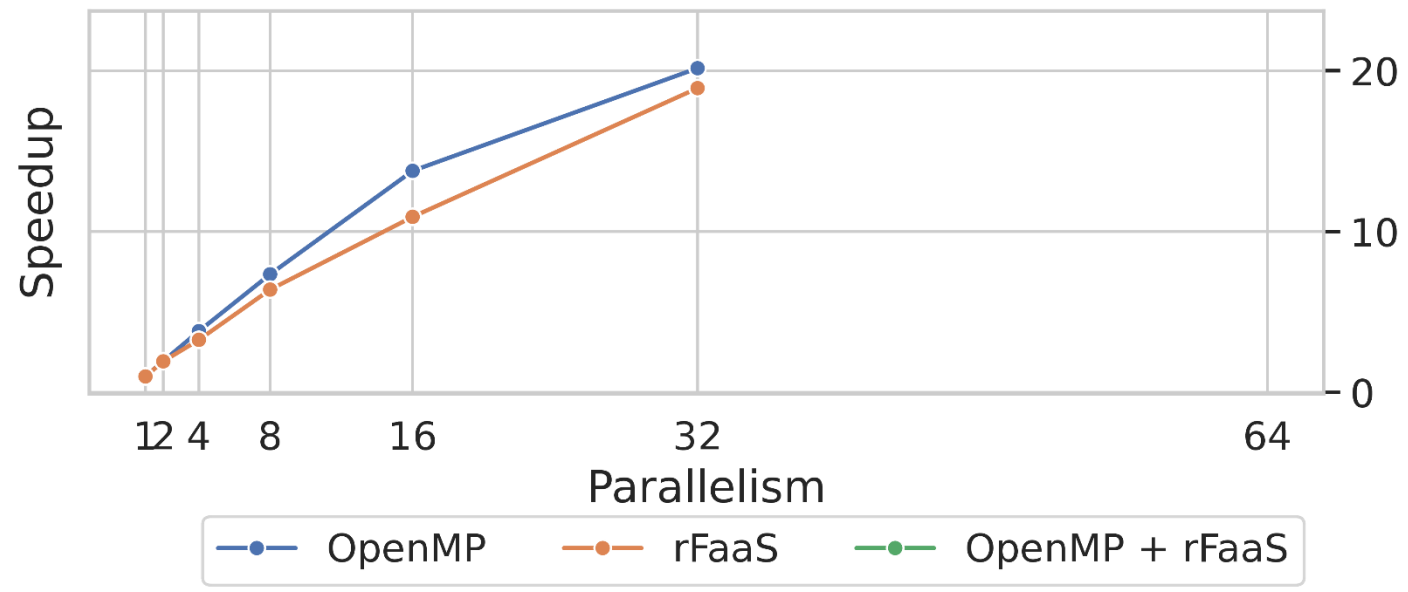


**10,000
Particles**

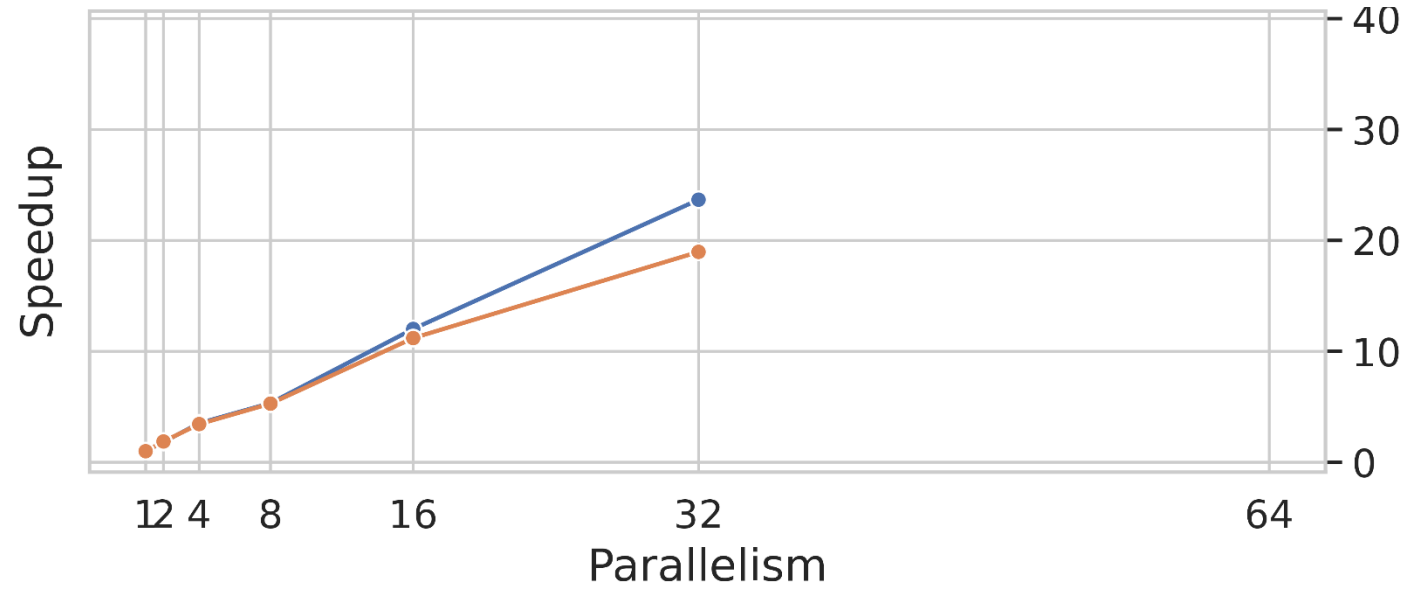


HPC Computations with FaaS: OpenMC

**1,000
Particles**

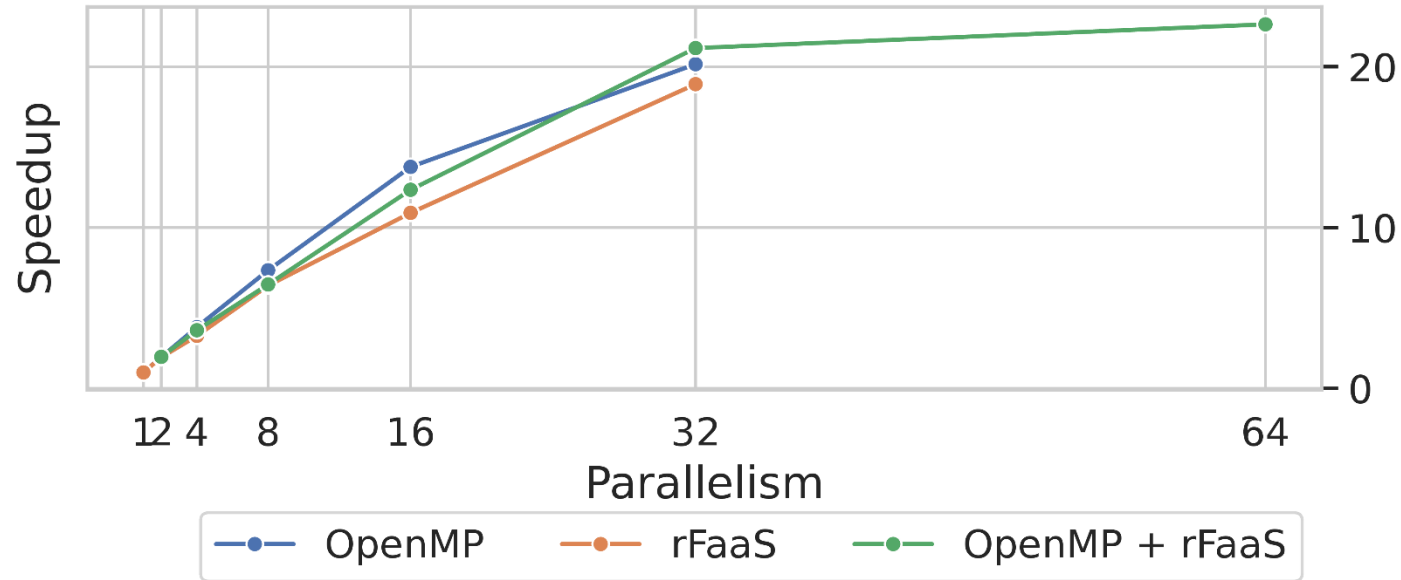


**10,000
Particles**

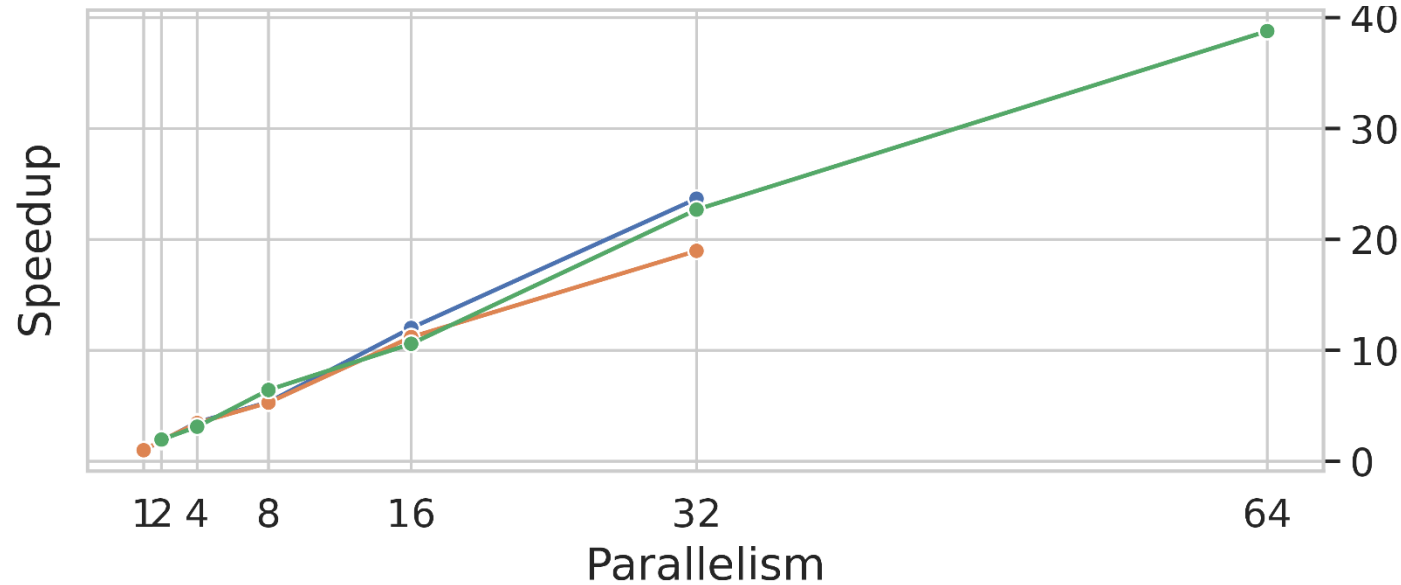


HPC Computations with FaaS: OpenMC

1,000
Particles

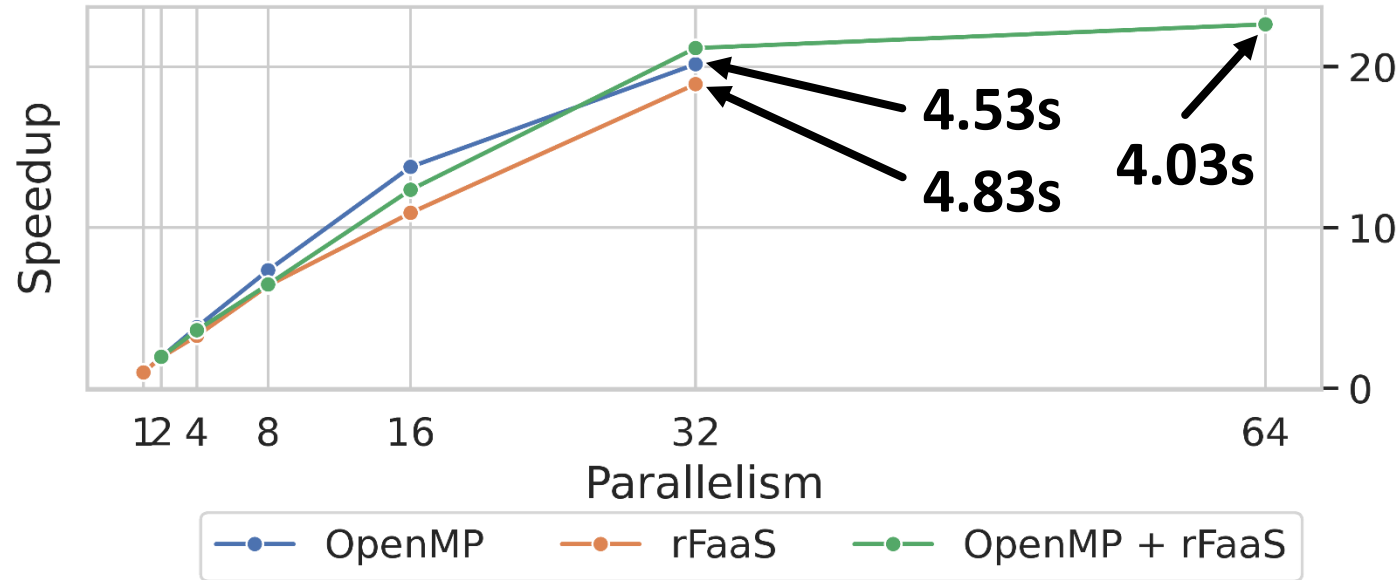


10,000
Particles

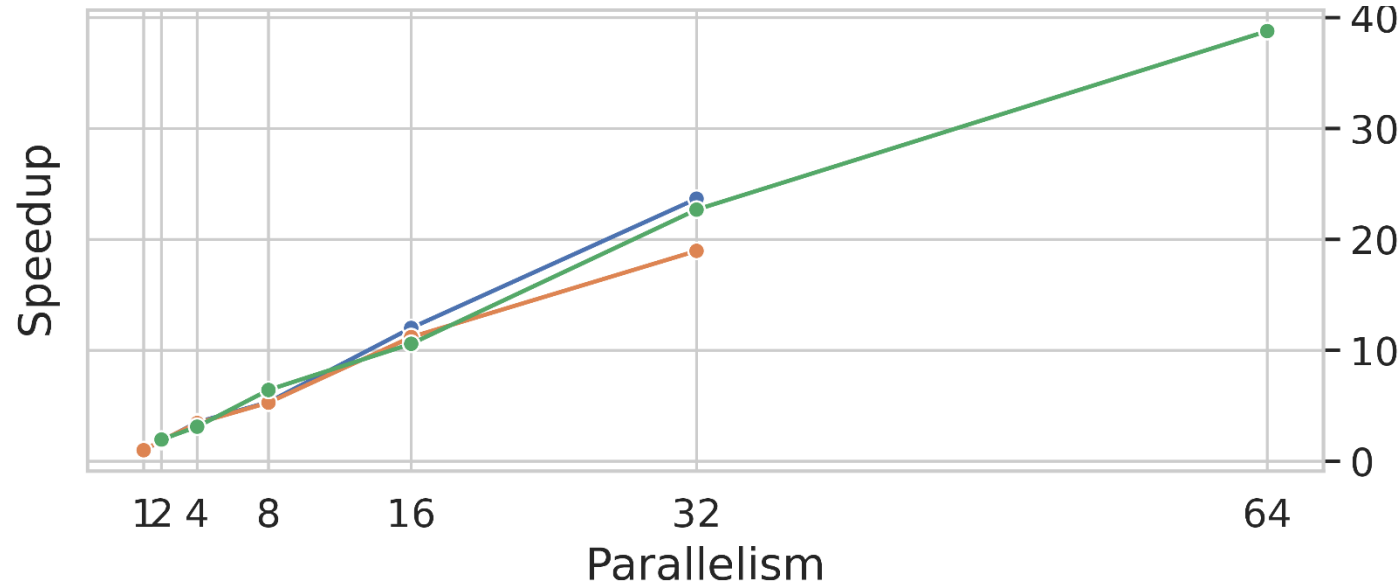


HPC Computations with FaaS: OpenMC

1,000
Particles

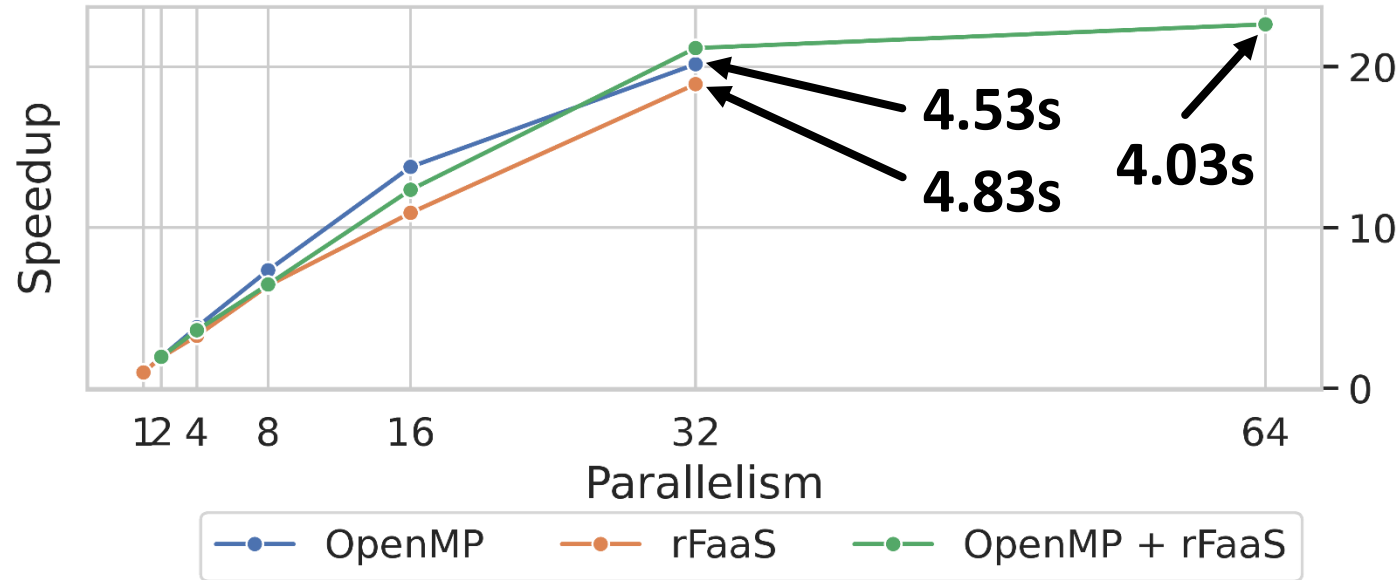


10,000
Particles

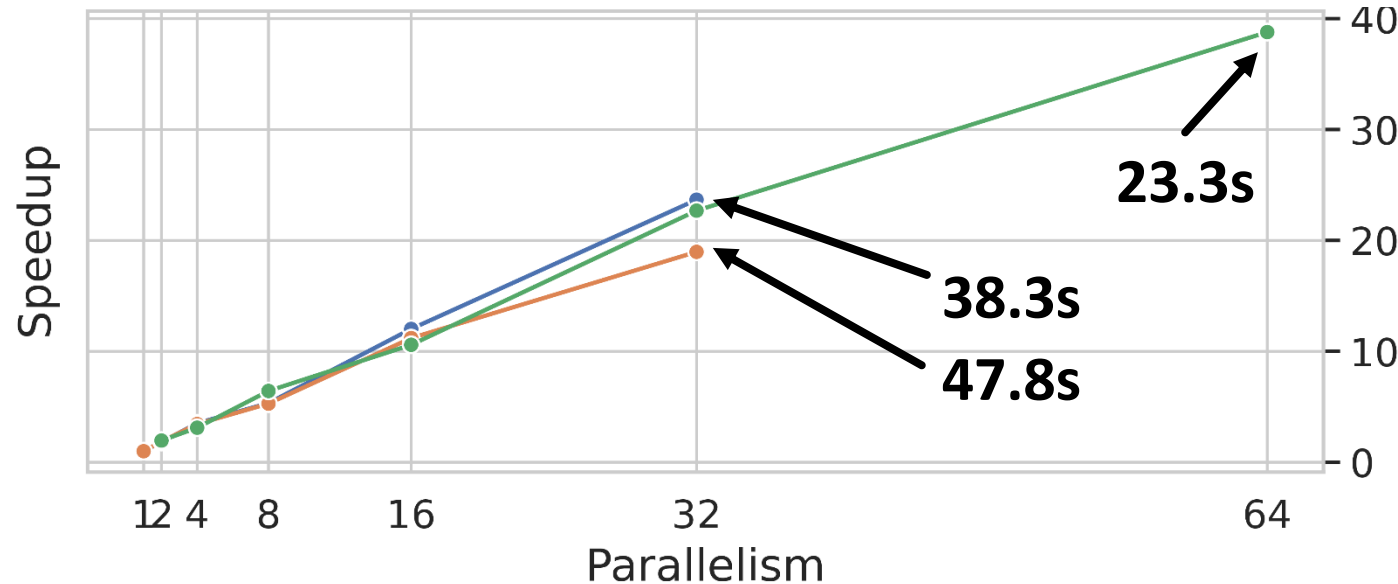


HPC Computations with FaaS: OpenMC

1,000
Particles



10,000
Particles




Serverless Solutions for HPC

Serverless Solutions for HPC



Serverless Solutions for HPC

 [spcl/rFaaS](https://github.com/spcl/rFaaS)

 [spcl/serverless-benchmarks](https://github.com/spcl/serverless-benchmarks)

Serverless Solutions for HPC



[spcl/rFaaS](#)



[spcl/serverless-benchmarks](#)




[spcl/fmi](#)

Serverless Solutions for HPC

 [spcl/rFaaS](#)

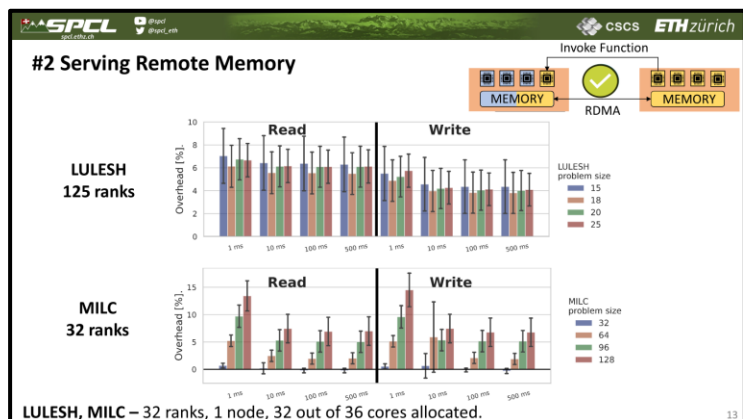
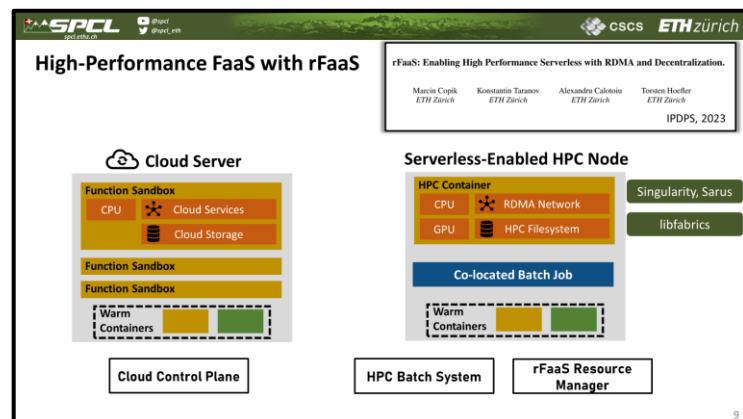
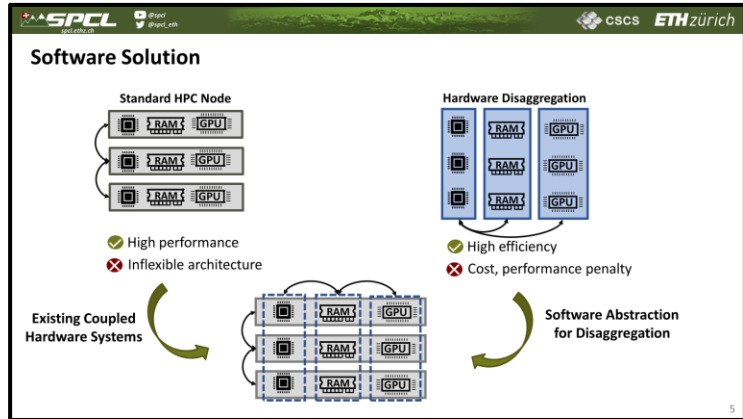
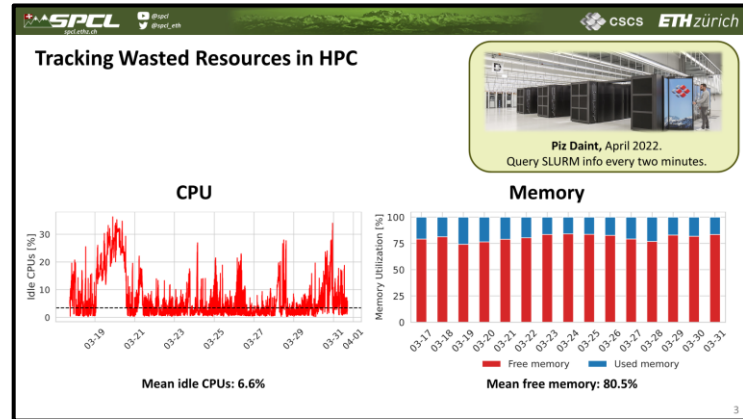
 [spcl/serverless-benchmarks](#)

 [spcl/fmi](#)

 [spcl/PraaS](#)

Conclusions

More of SPCL's research:



youtube.com/@spcl **180+ Talks**

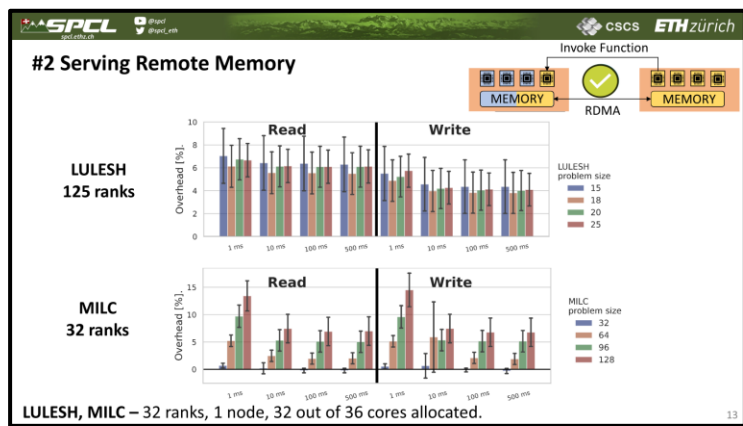
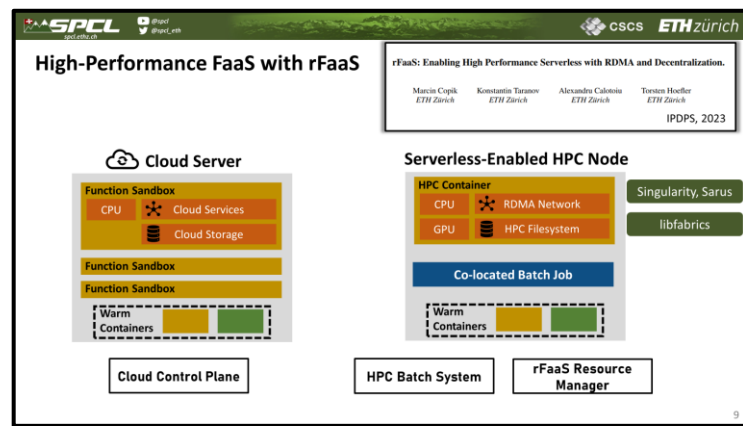
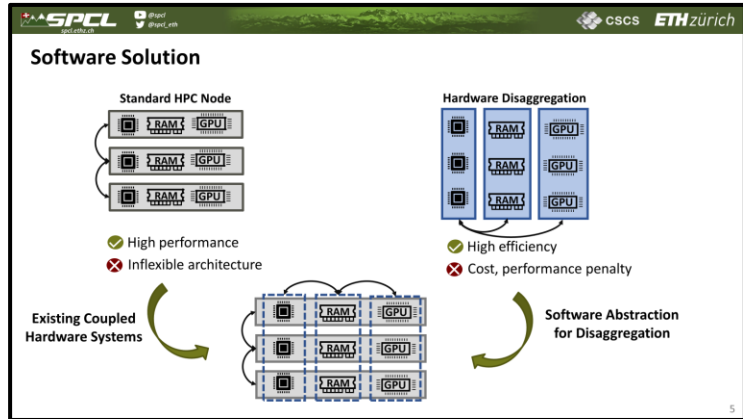
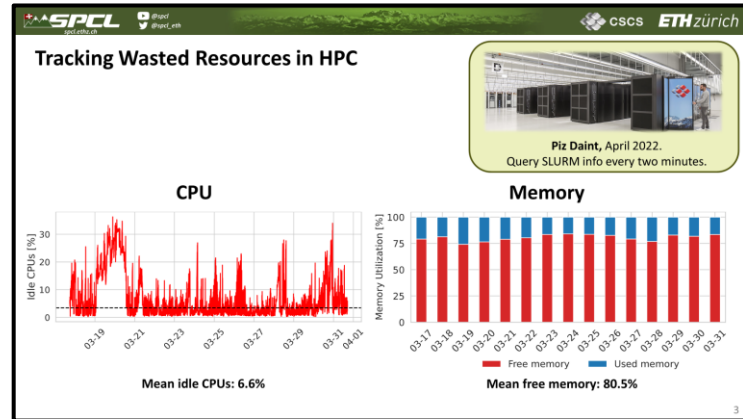
twitter.com/spcl_eth **1.4K+ Followers**

github.com/spcl **3.8K+ Stars**

... or spcl.ethz.ch



Conclusions



More of SPCL's research:

- youtube.com/@spcl **180+ Talks**
- twitter.com/spcl_eth **1.4K+ Followers**
- github.com/spcl **3.8K+ Stars**

... or spcl.ethz.ch



Paper



Projects

