Distributed Learning and You

Ruksi Laine
@ruk_si
https://valohai.com
@valohaiai

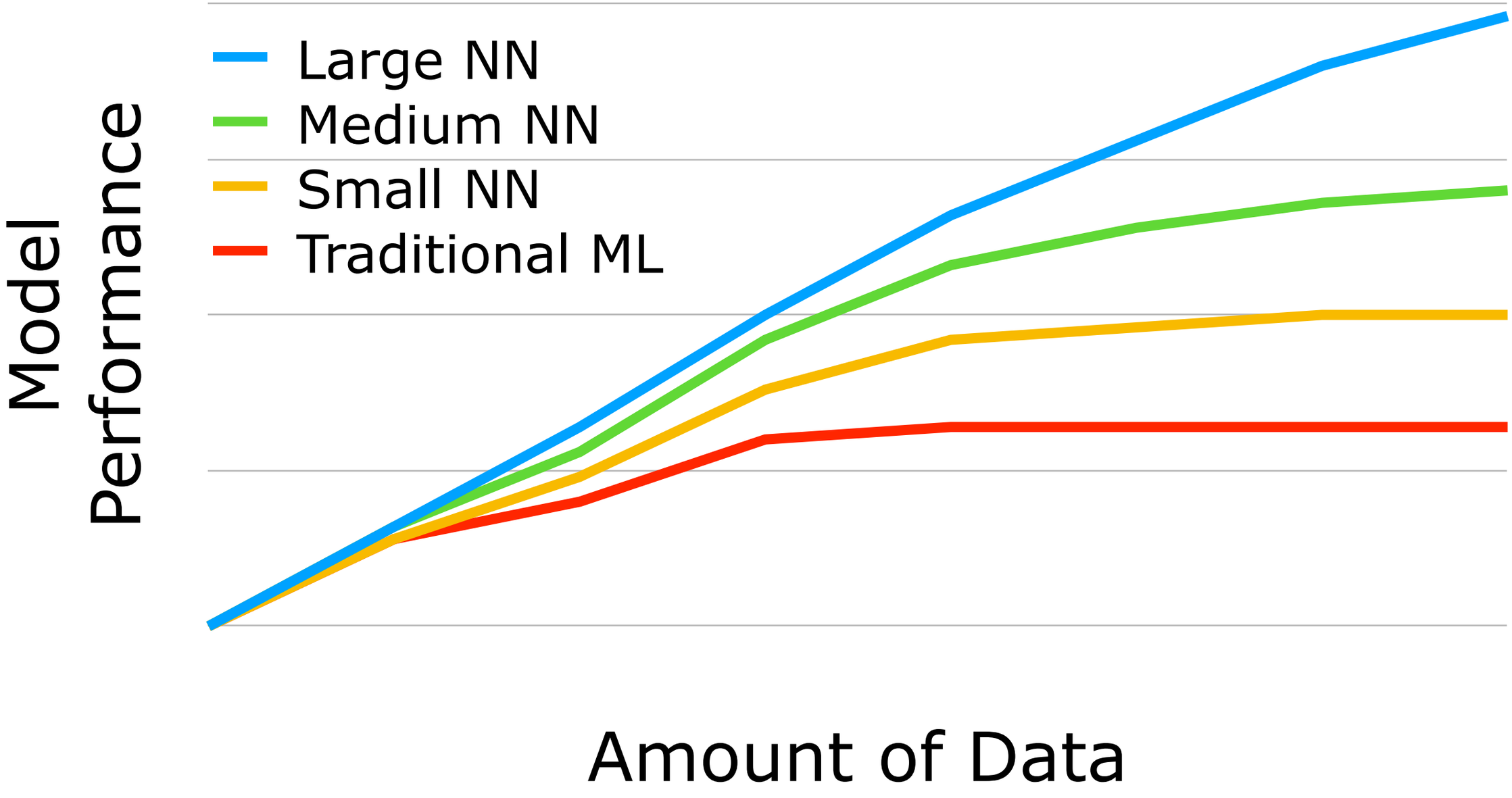
Me?

Hi, I'm Ruksi.

Machine Learning Engineer
https://ruk.si
@ruk_si

VALOHAI @valohaiai

Scale Drives Deep Learning Process

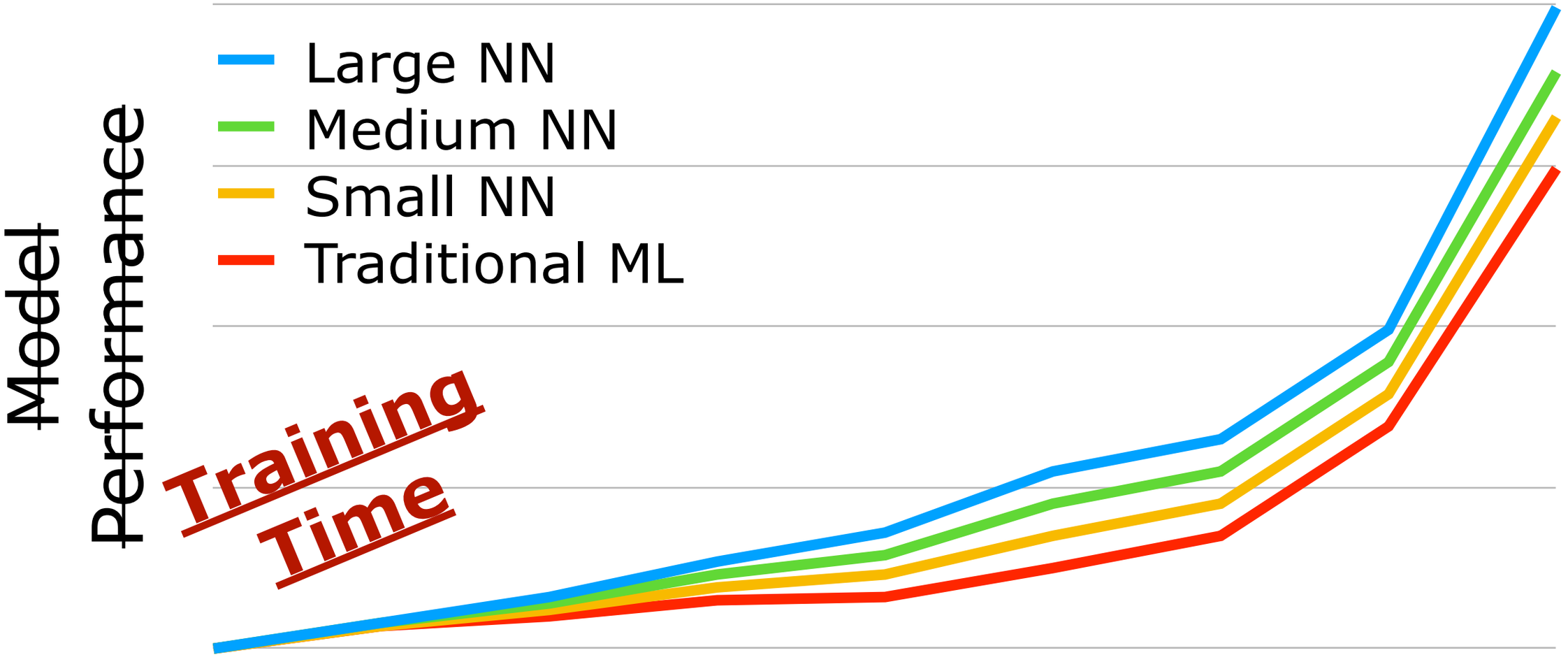


VALOHAI

original: Andrew Ng Coursera lectures

Scale Makes Everything Slow

~~Scale Drives Deep Learning Process~~

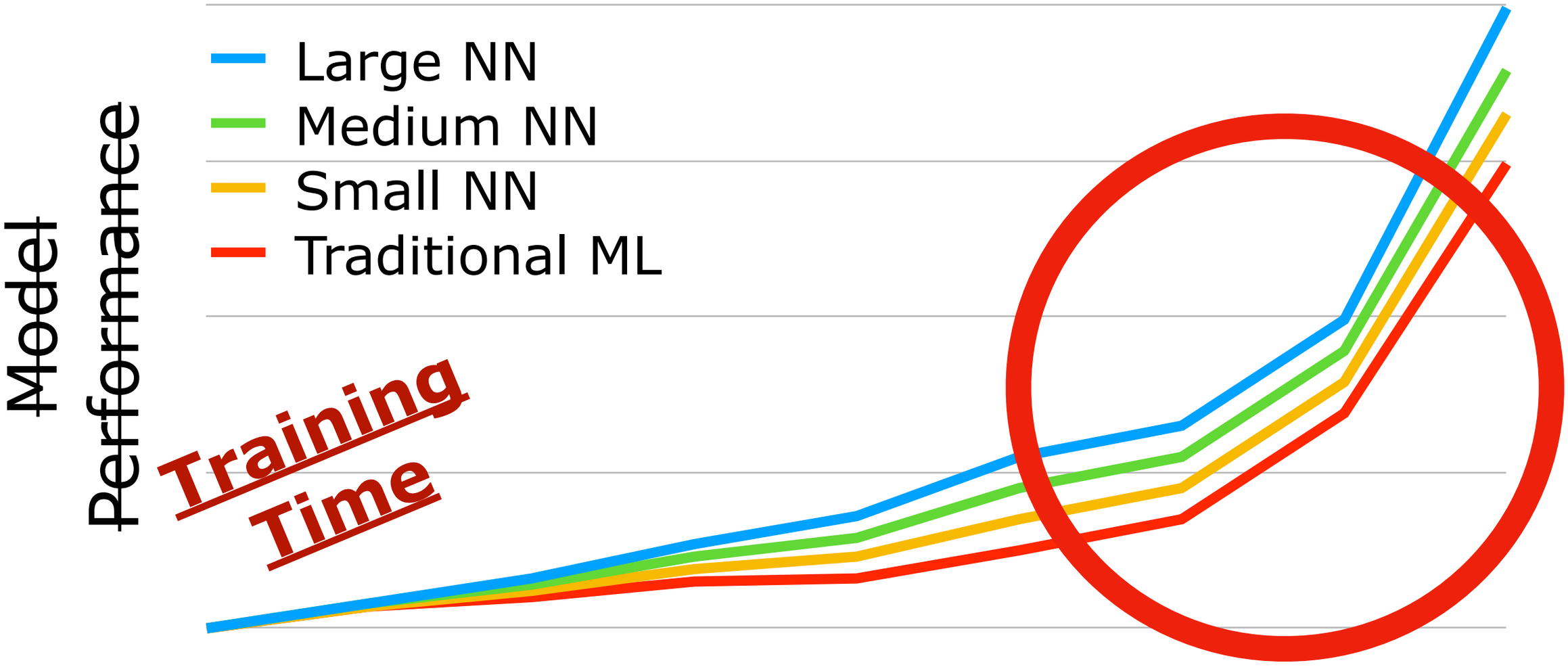


VALOHAI

original: Andrew Ng Coursera lectures

Scale Makes Everything Slow

~~Scale Drives Deep Learning Process~~

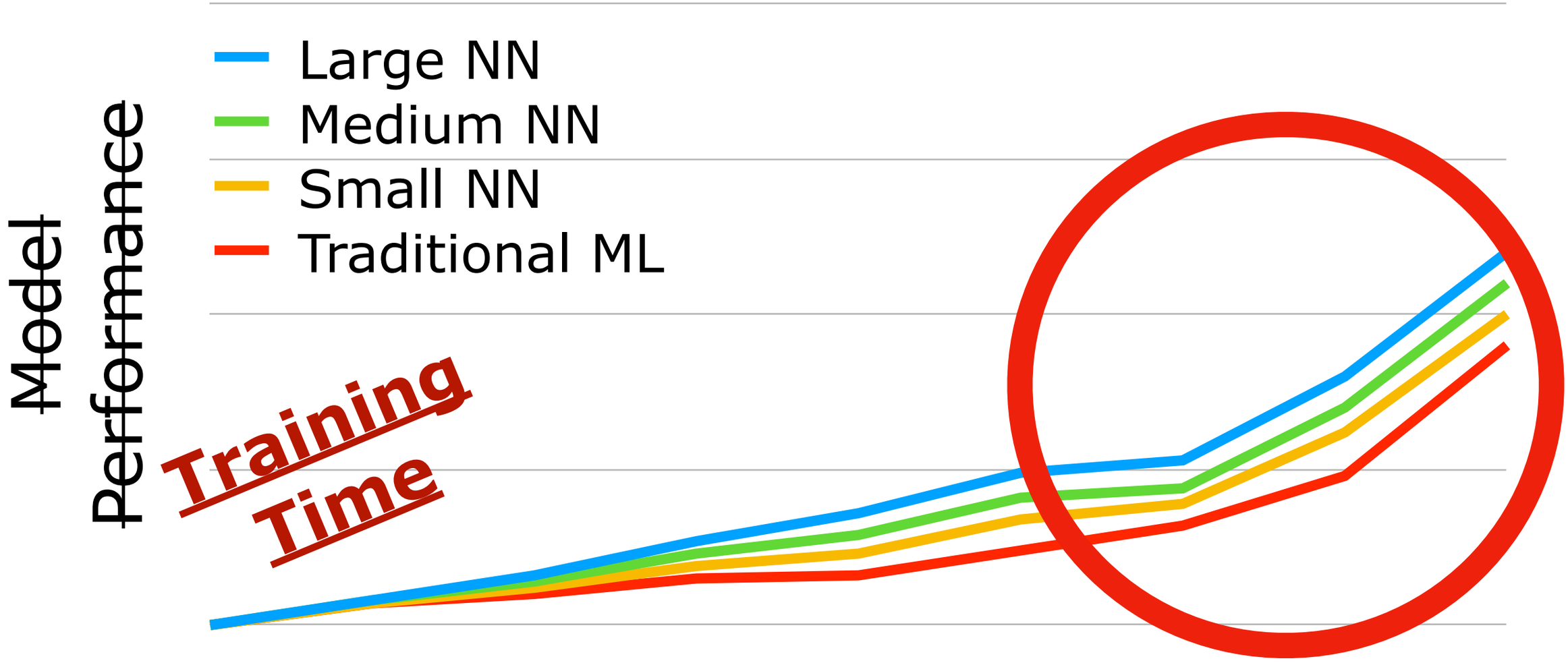


VALOHAI

original: Andrew Ng Coursera lectures

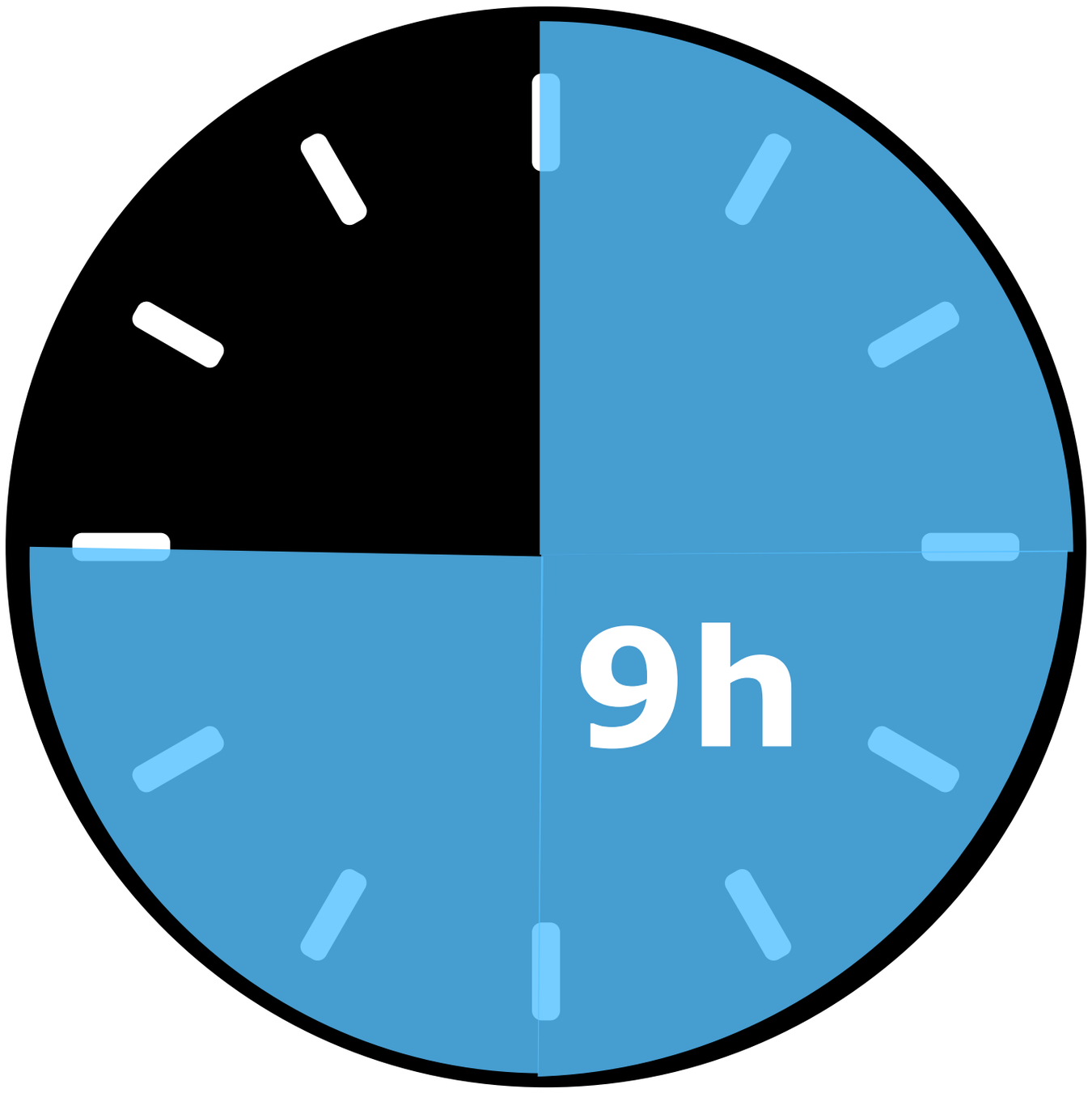
Scale Makes Everything Slow

~~Scale Drives Deep Learning Process~~



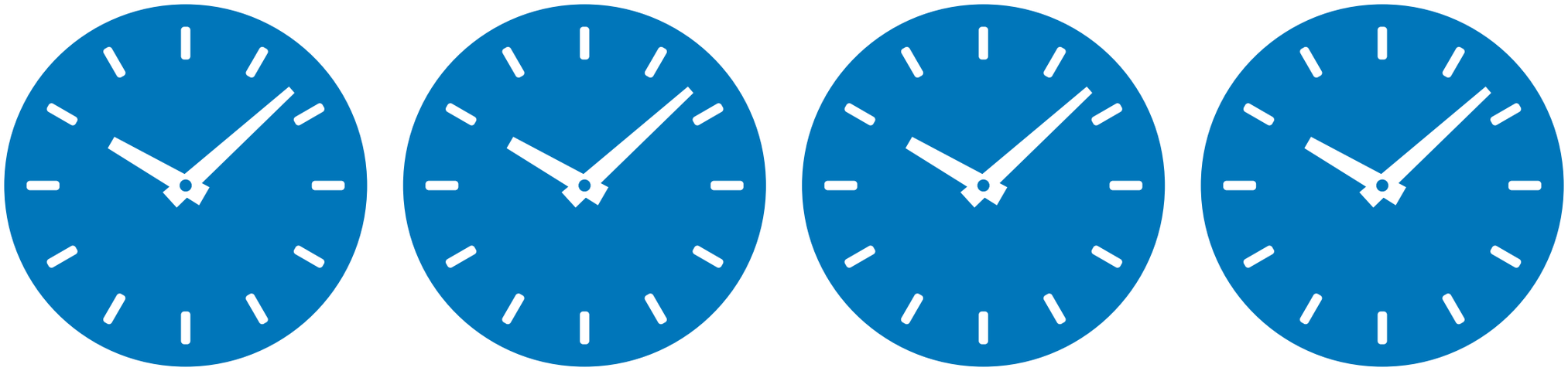
VALOHAI

original: Andrew Ng Coursera lectures



VALOHAI

VALOHAI



VALOHAI

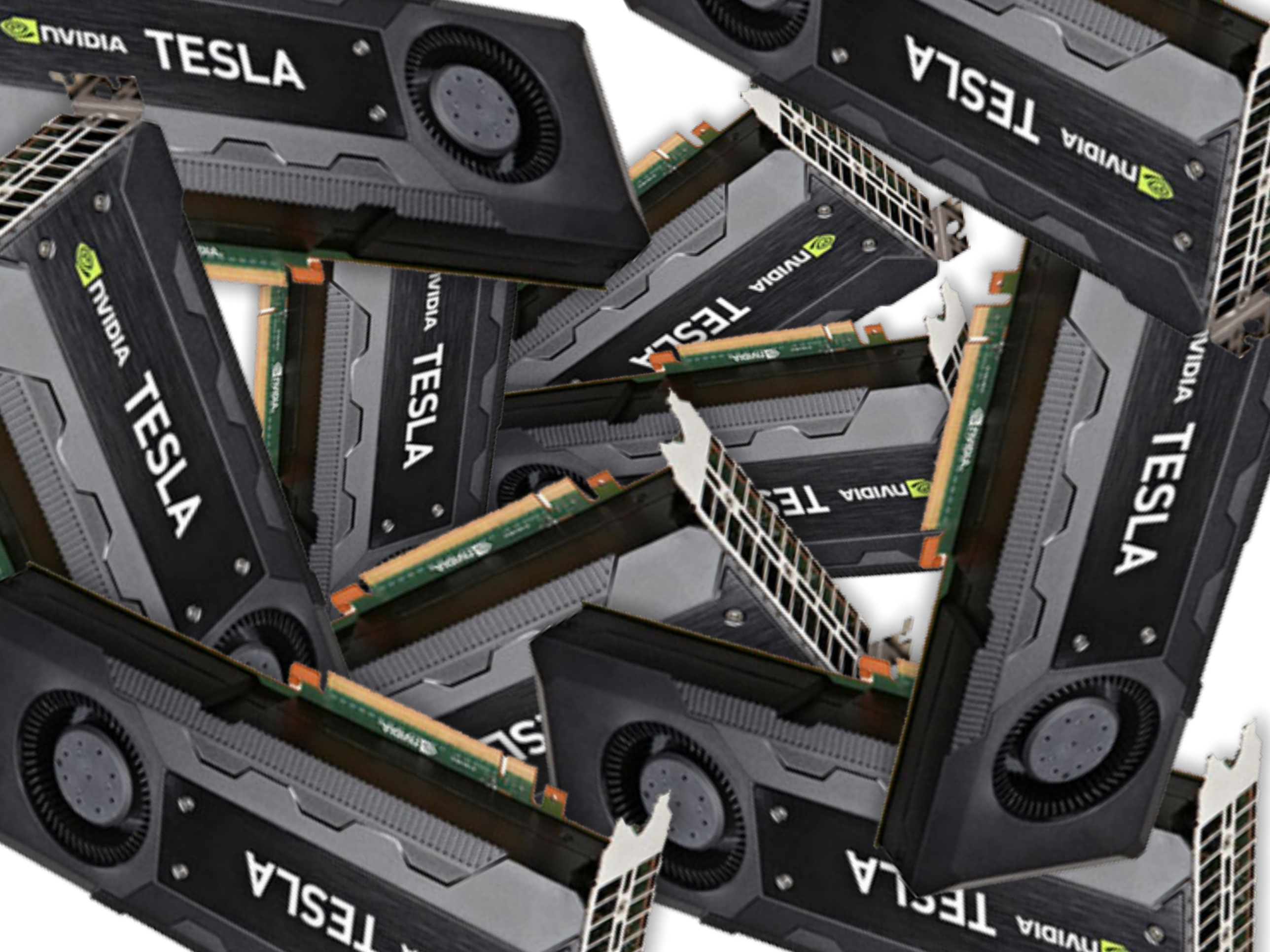
VALOHAI

VALOHAI

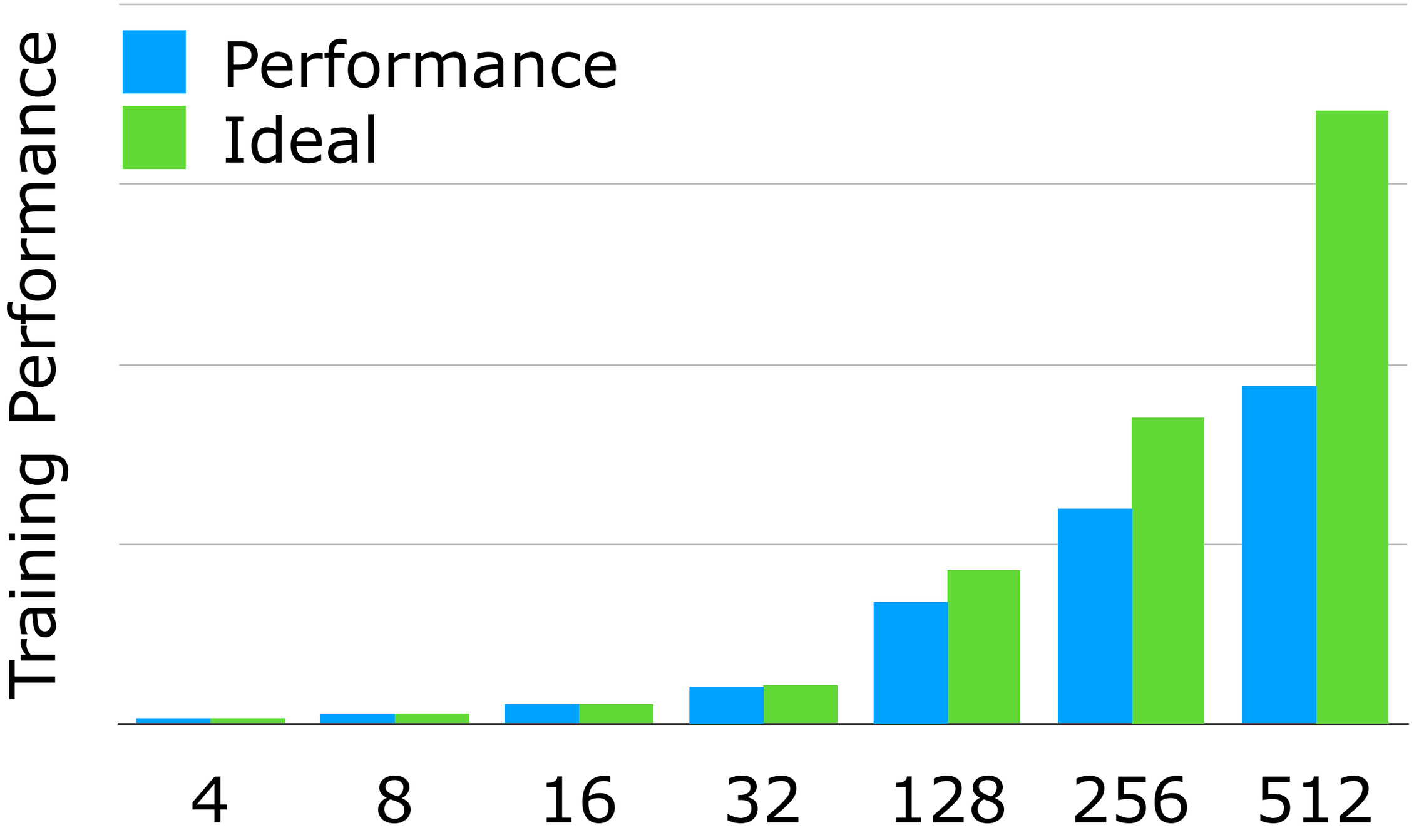
VALOHAI

VALOHAI

VALOHAI


NVIDIA TESLA

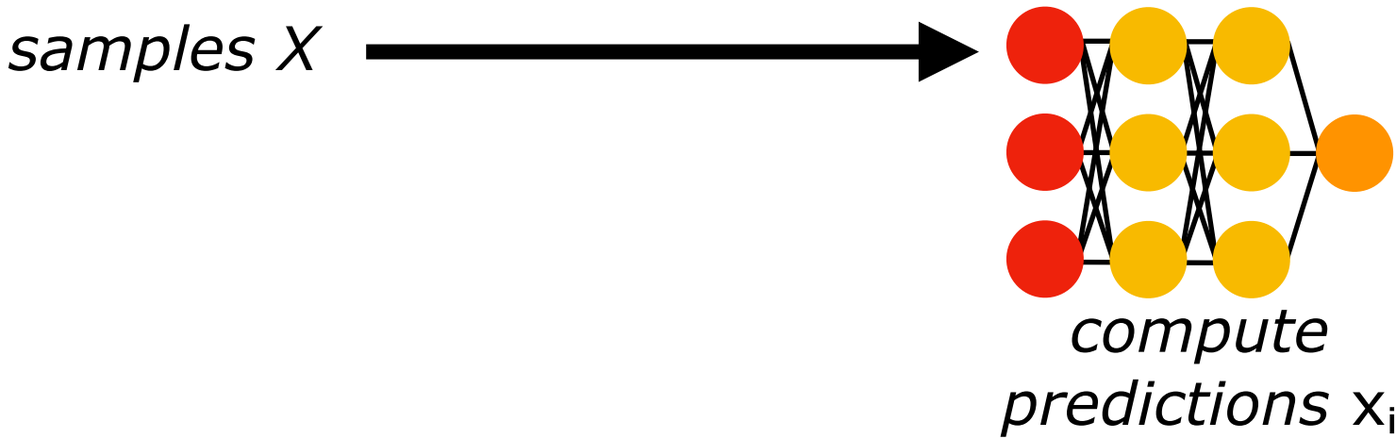
Scaling Efficiency



VALOHAI

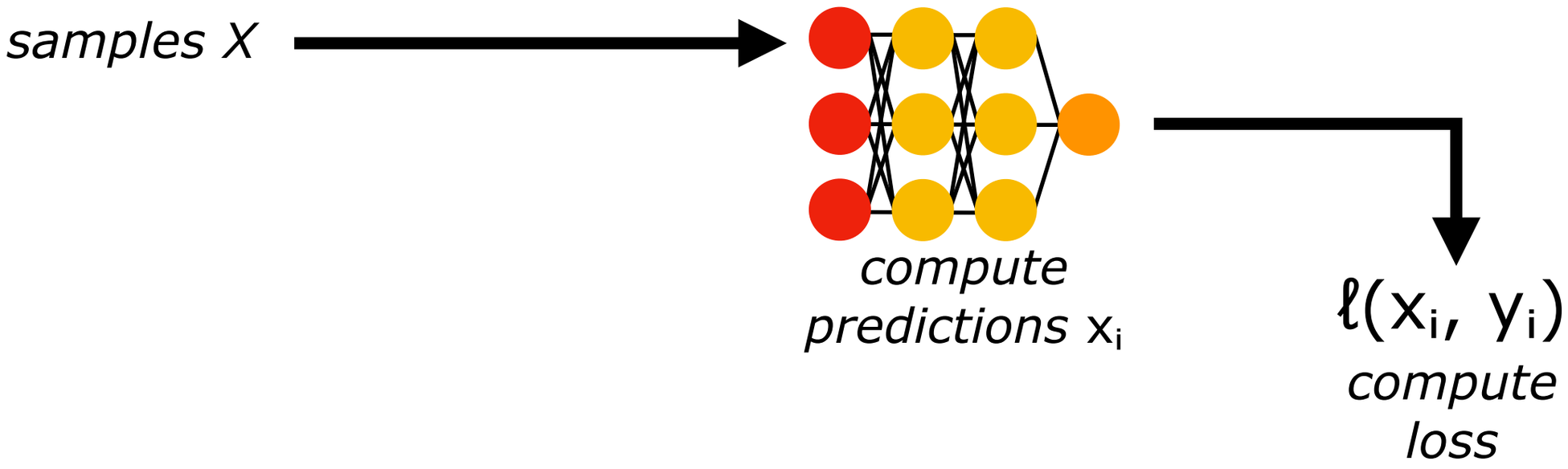
Deep Learning is Innately Sequential

Training Loop



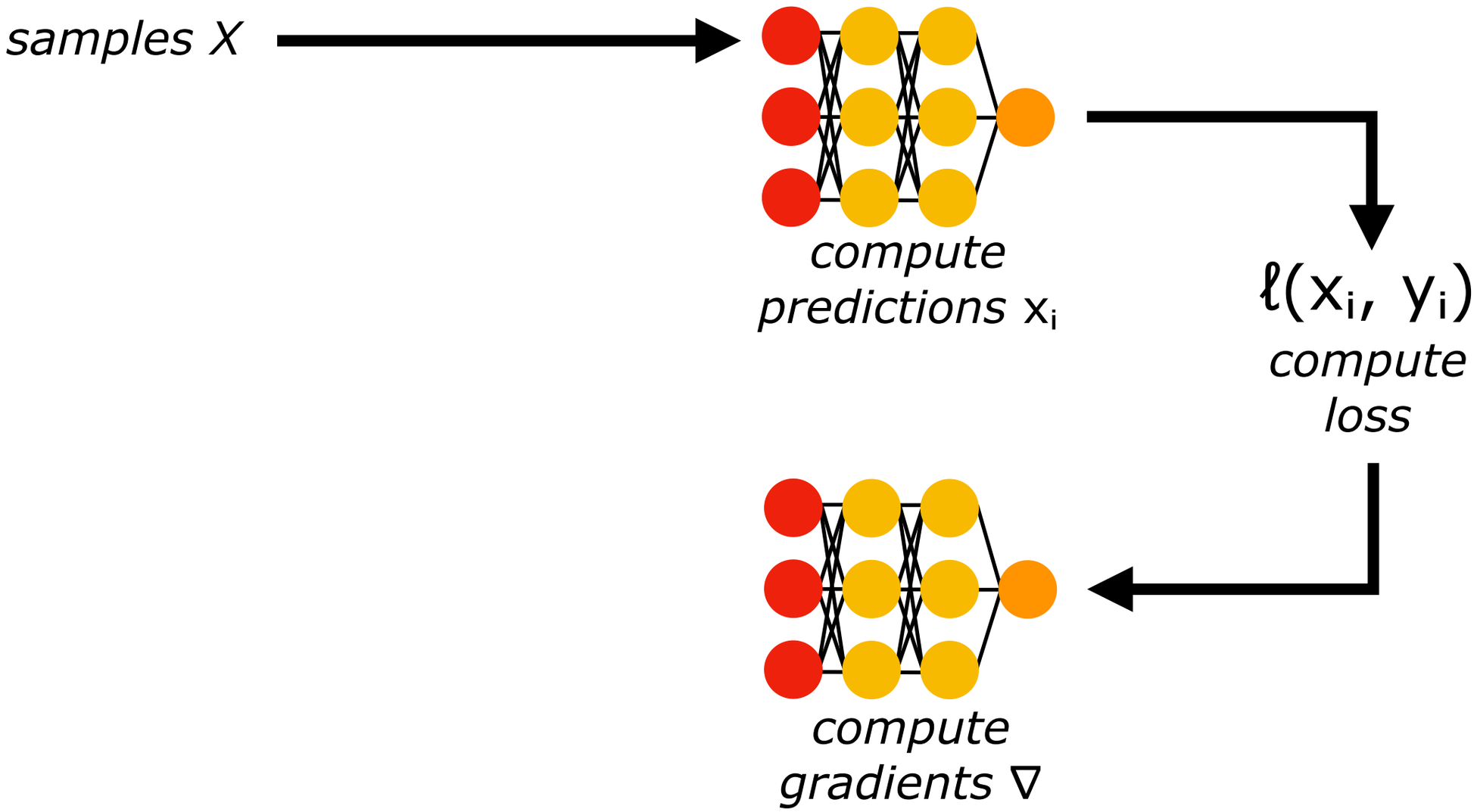
VALOHAI

Training Loop



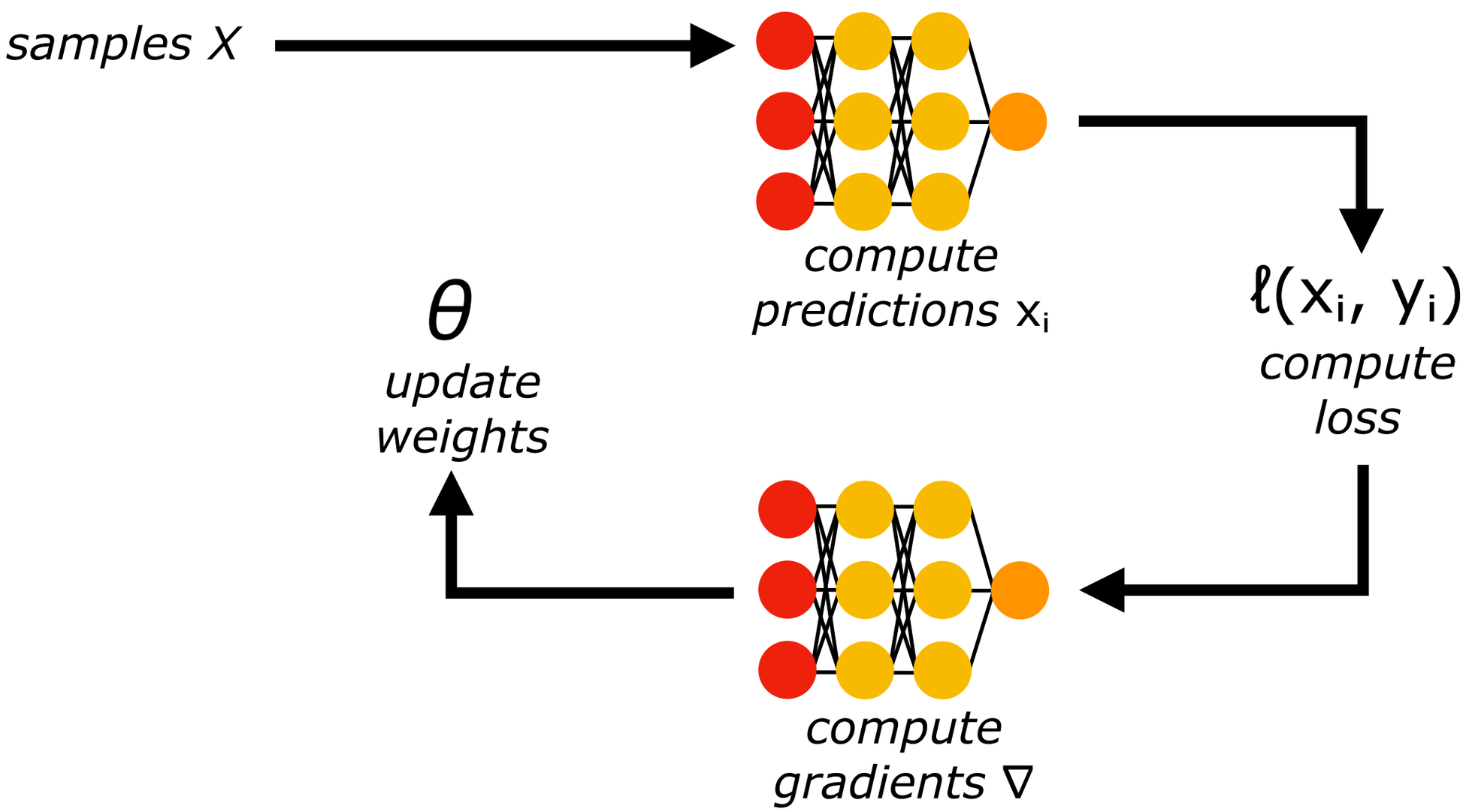
VALOHAI

Training Loop



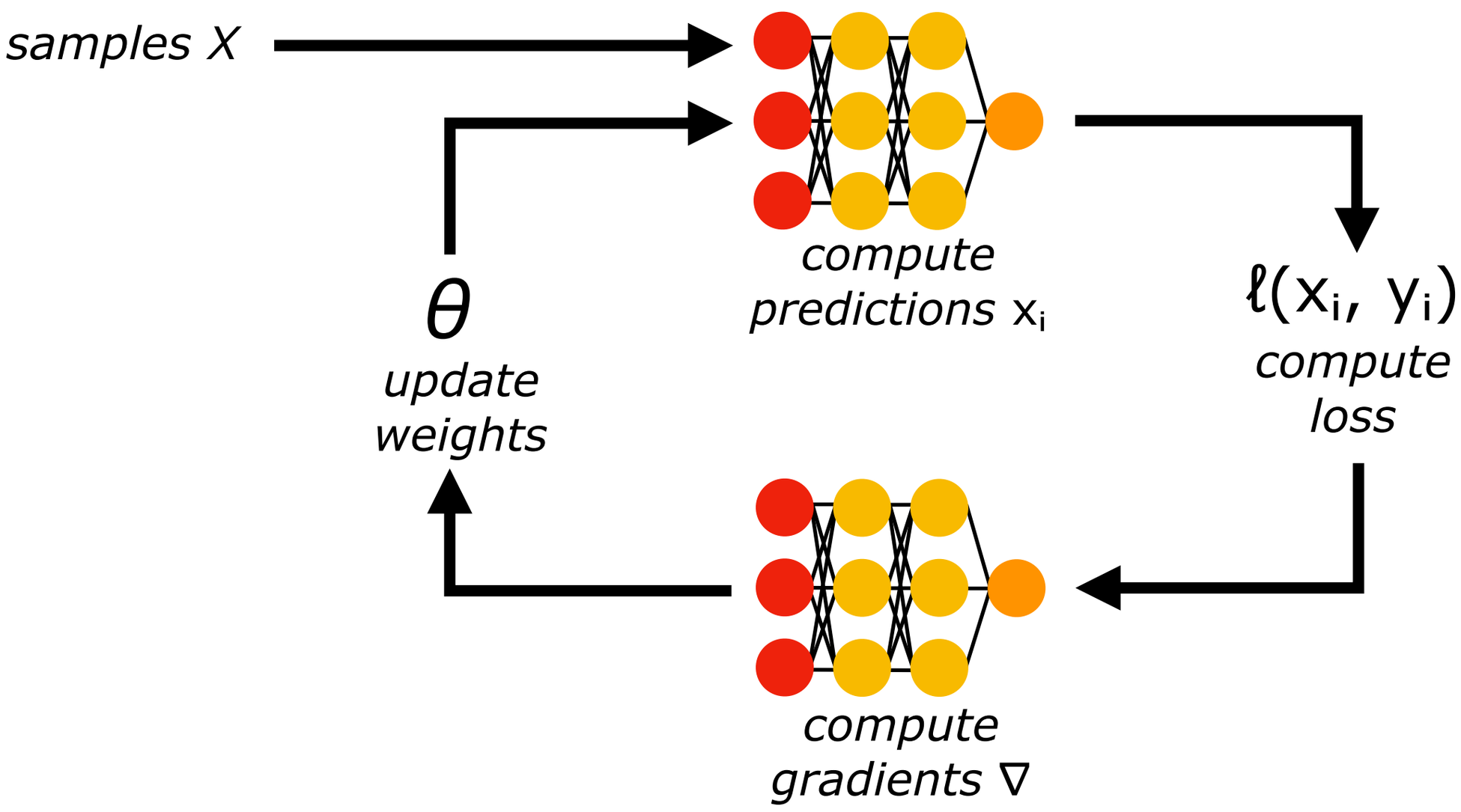
VALOHAI

Training Loop



VALOHAI

Training Loop



VALOHAI

(innately sequential)

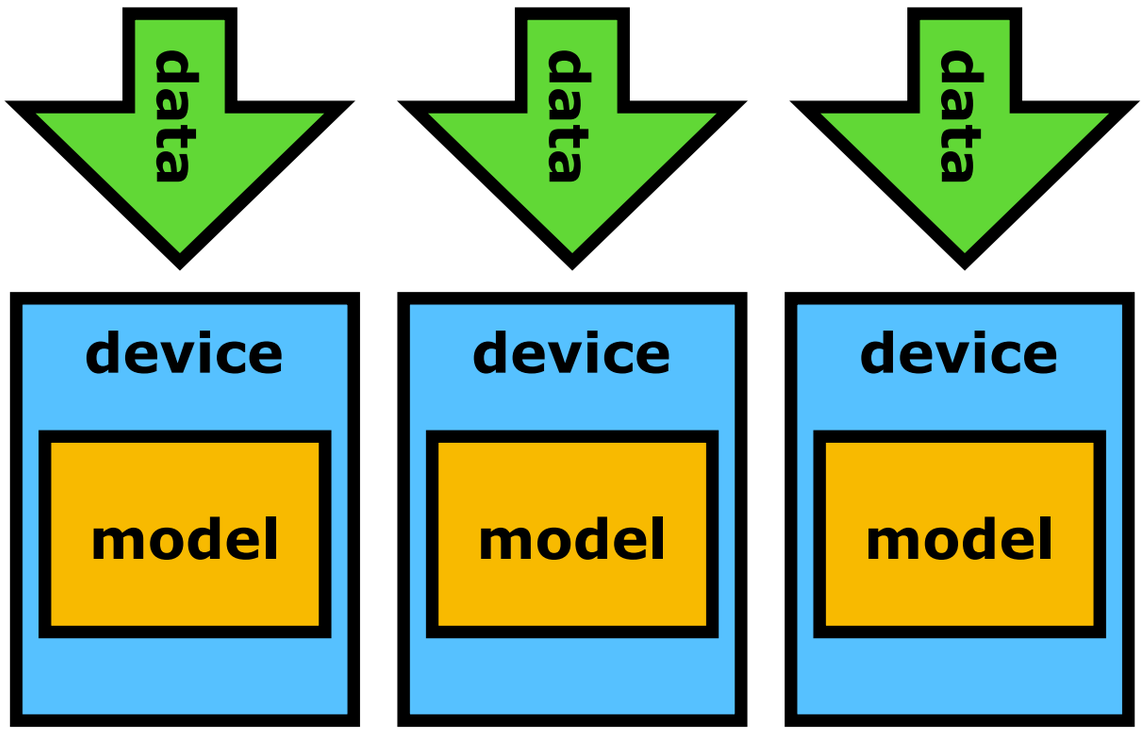
Two Main Branches of Distributed Learning

VALOHAI

Data Parallelization

Model Parallelization

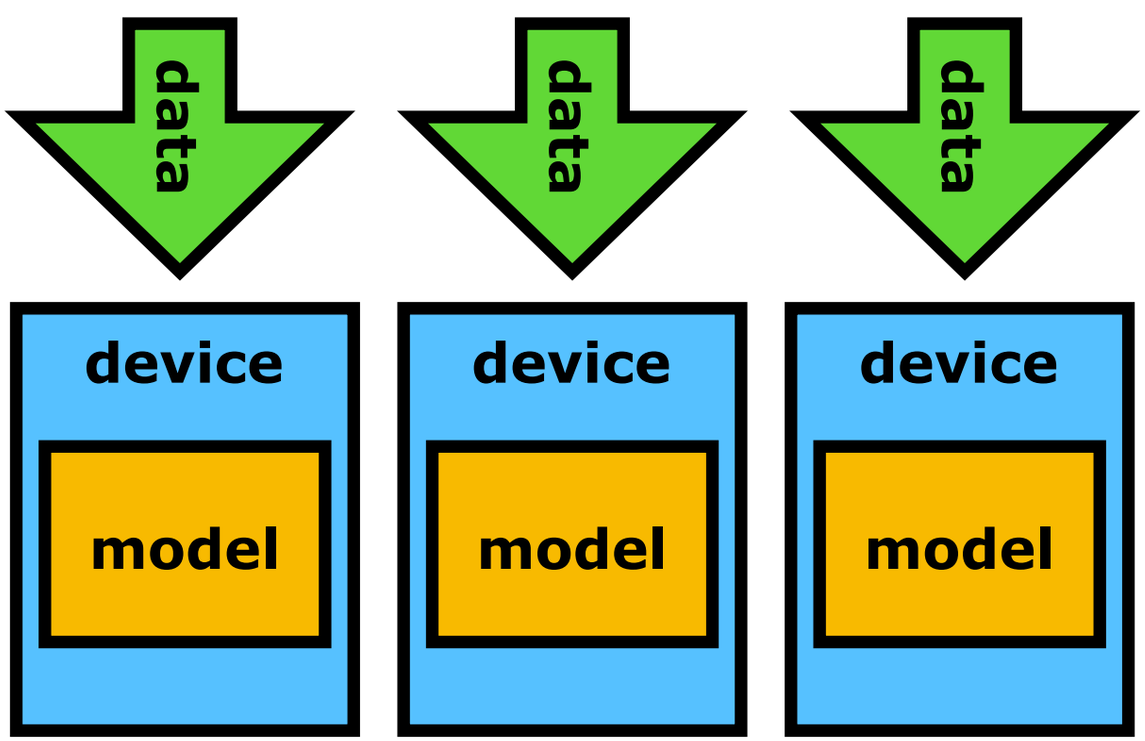
Data Parallelization



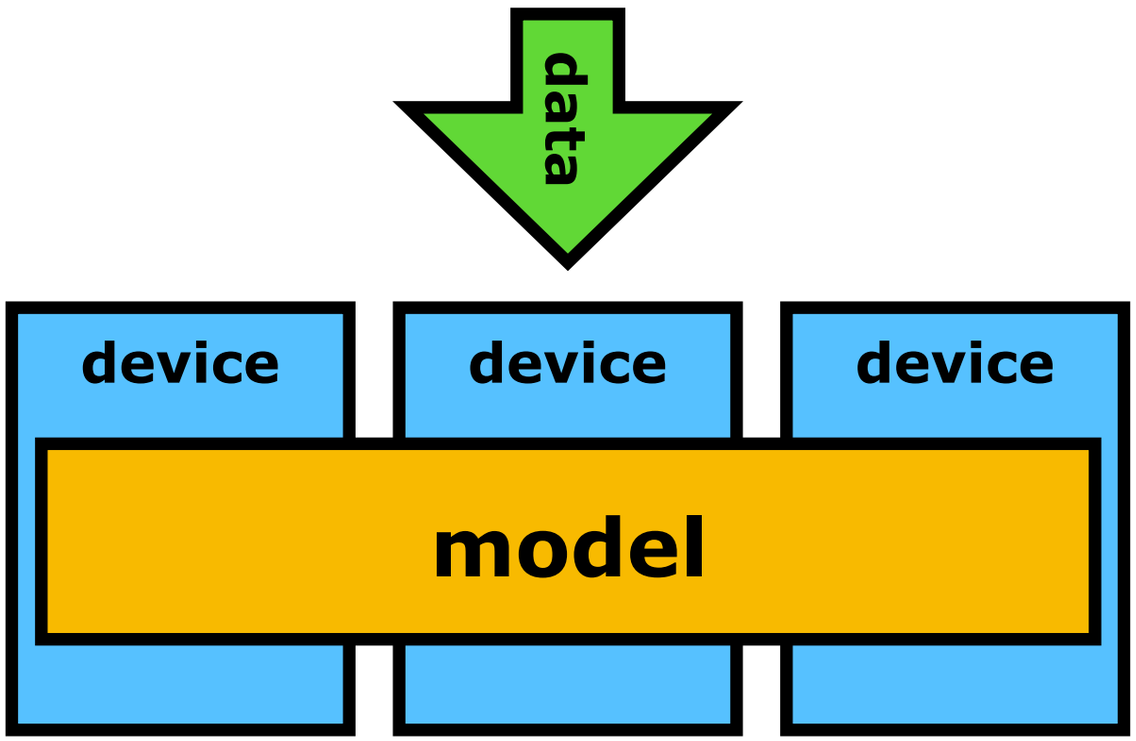
Model Parallelization

VALOHAI

Data Parallelization

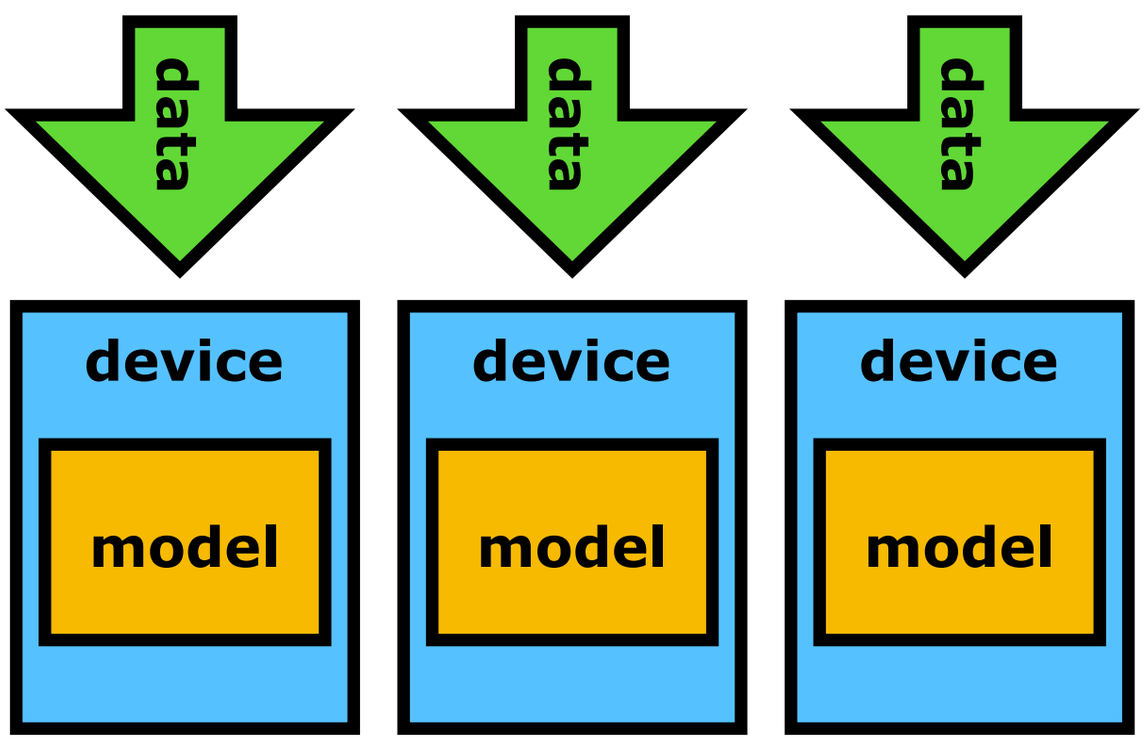


Model Parallelization

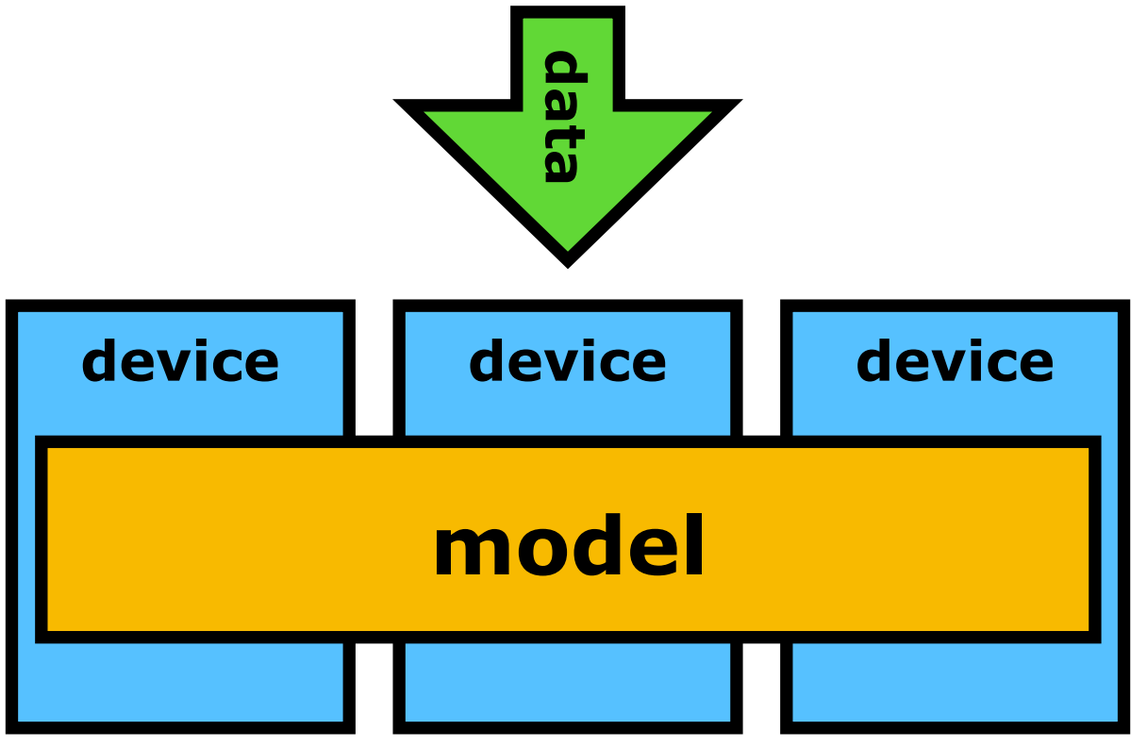


VALOHAI

Data Parallelization



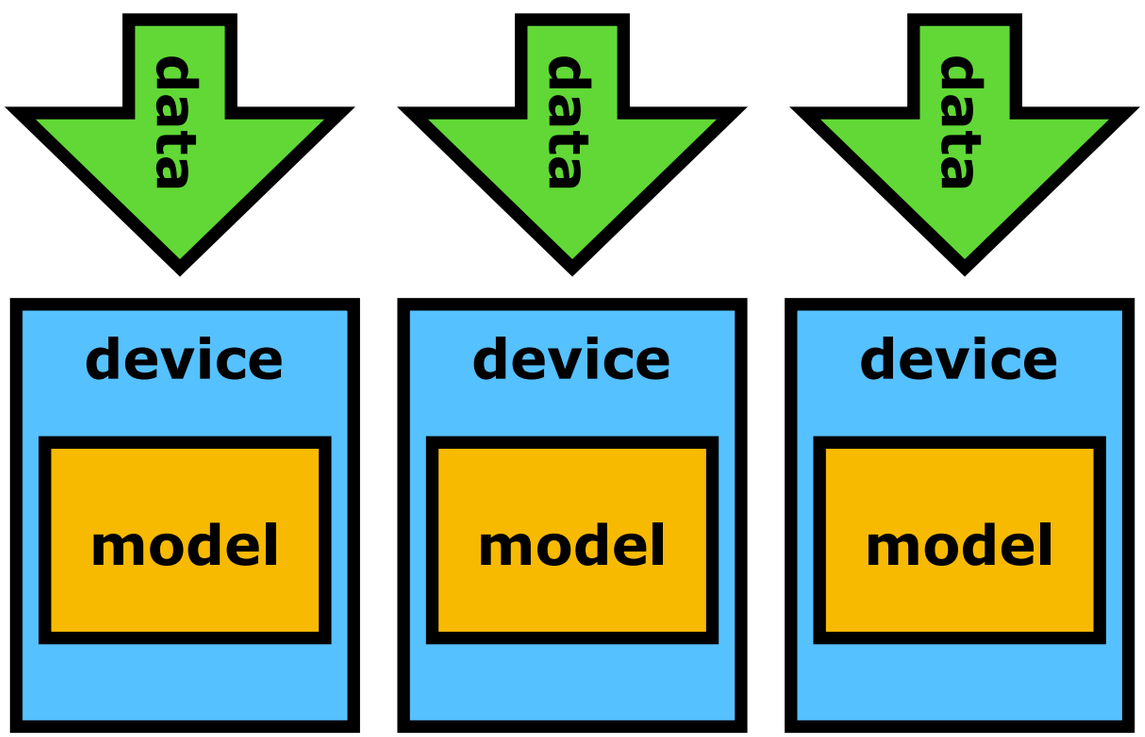
Model Parallelization



Synchronize weights θ to
train models fast.

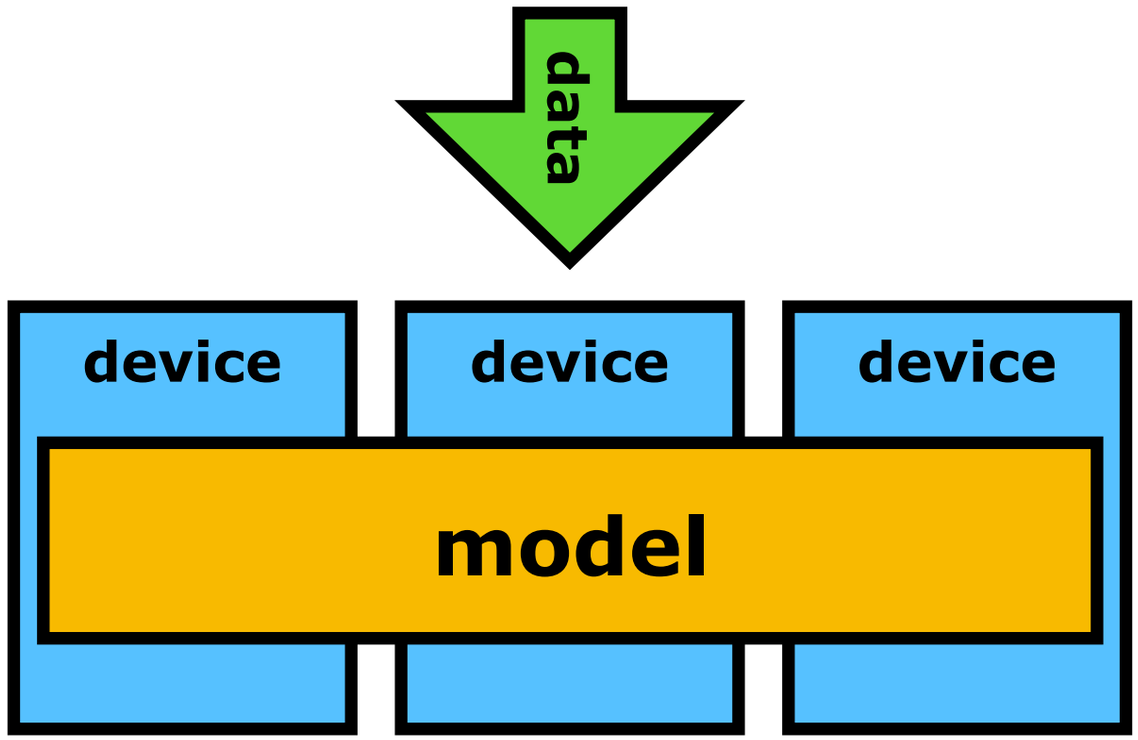
VALOHAI

Data Parallelization



Synchronize weights θ to
train models fast.

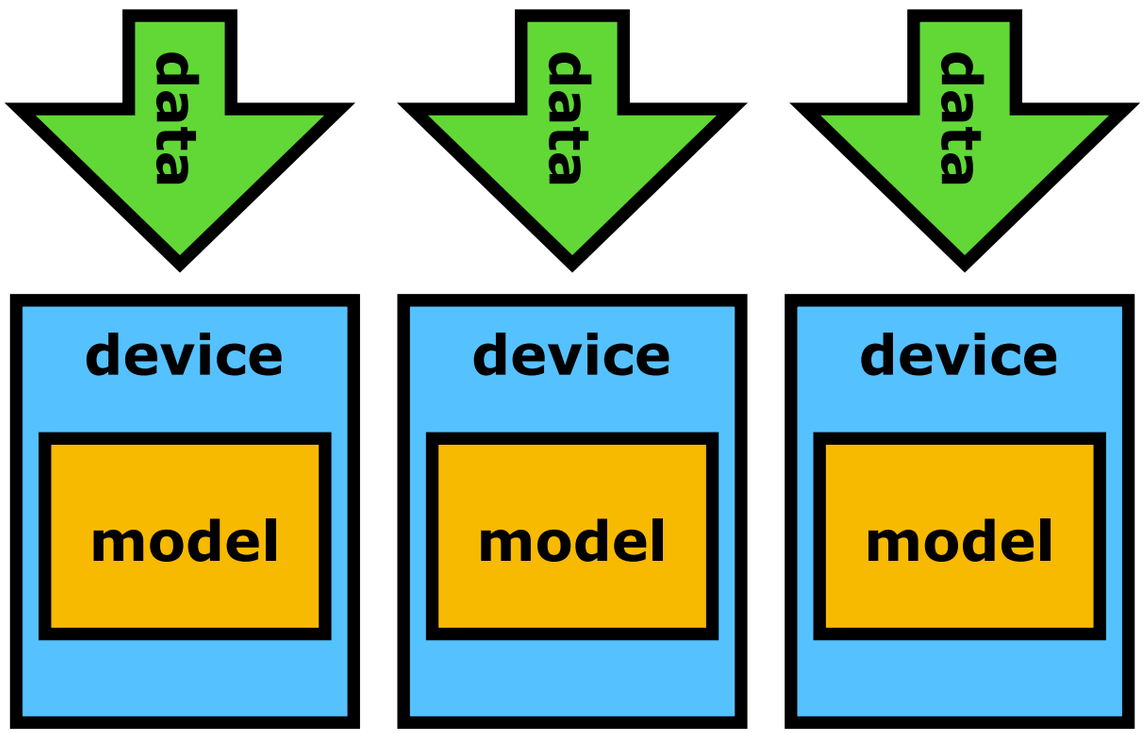
Model Parallelization



Partition a model to
build huge models.

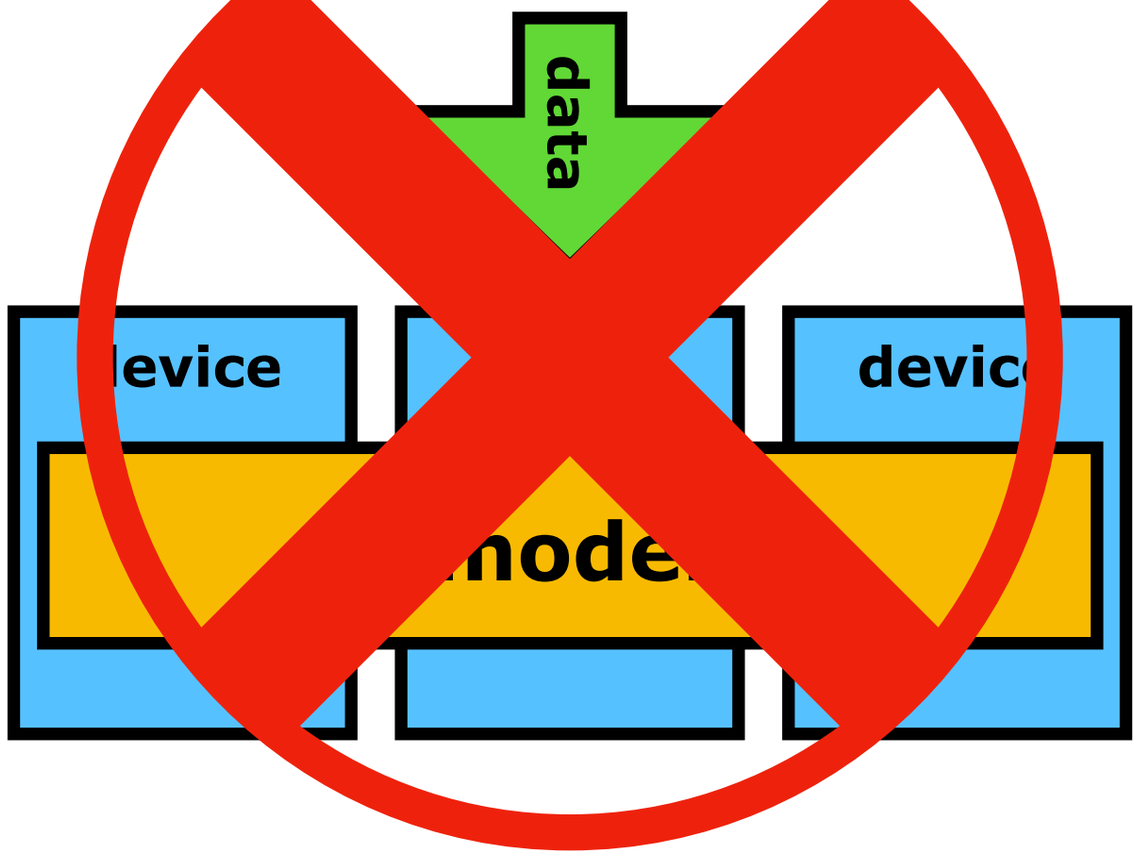
VALOHAI

Data Parallelization



Synchronize weights θ to
train models fast.

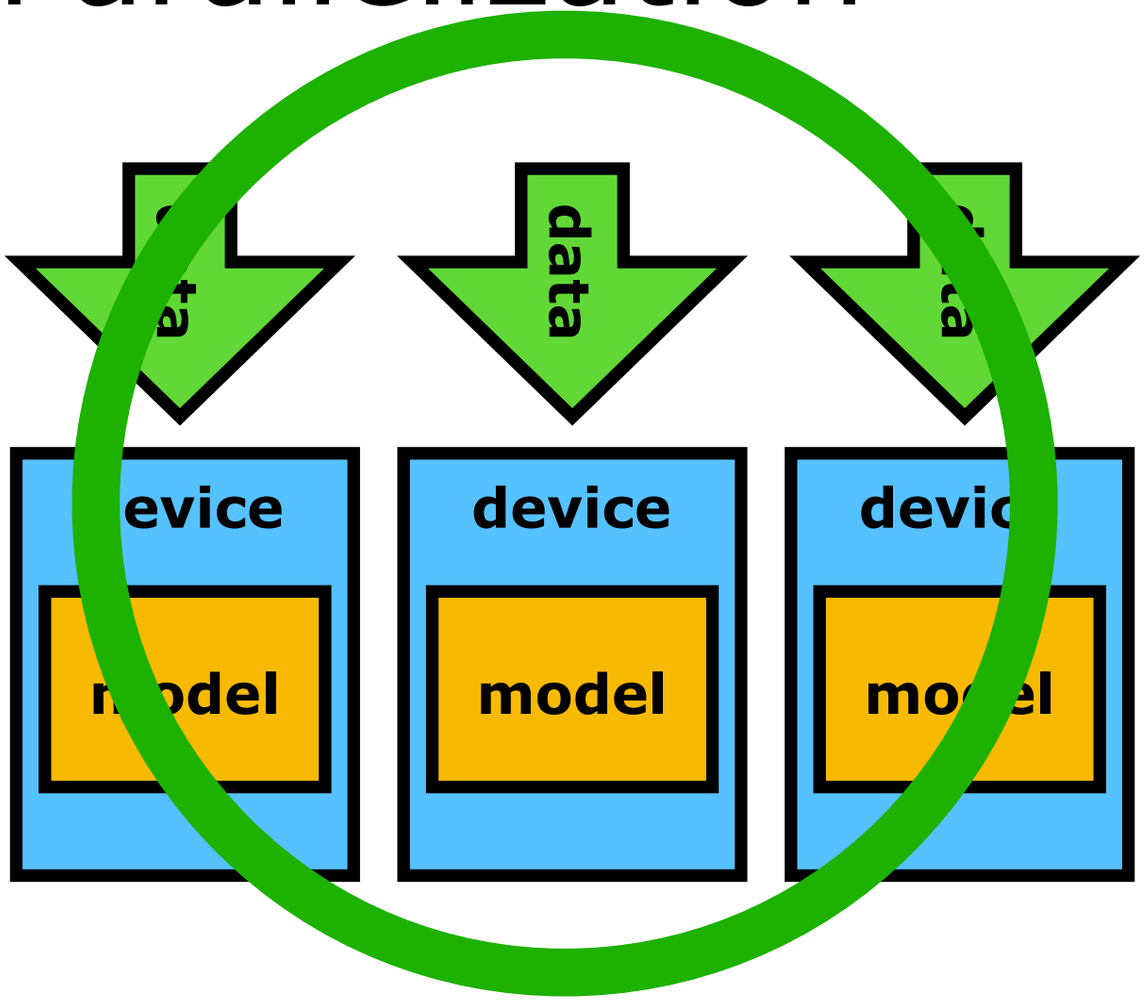
Model Parallelization



Partition a model to
build huge models.

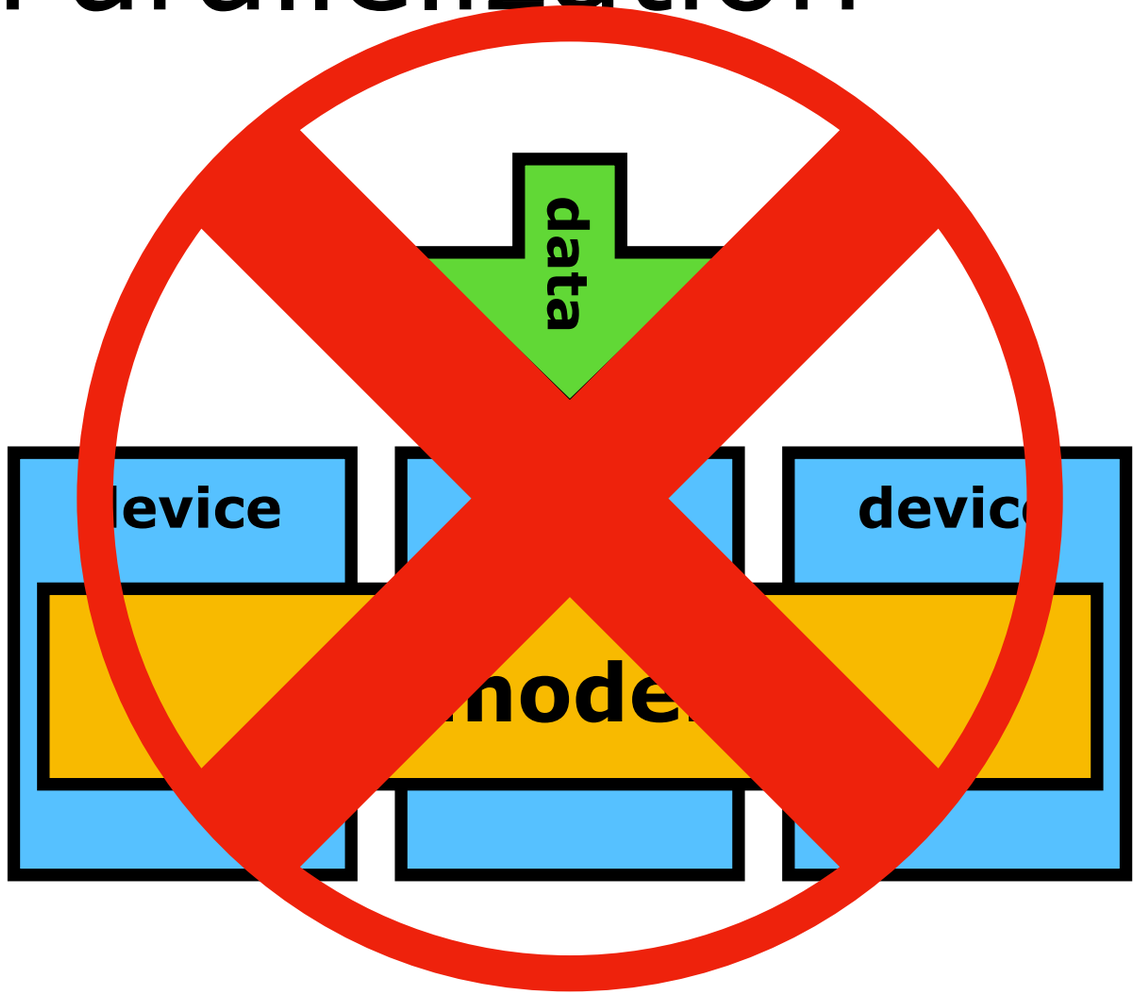
VALOHAI

Data Parallelization



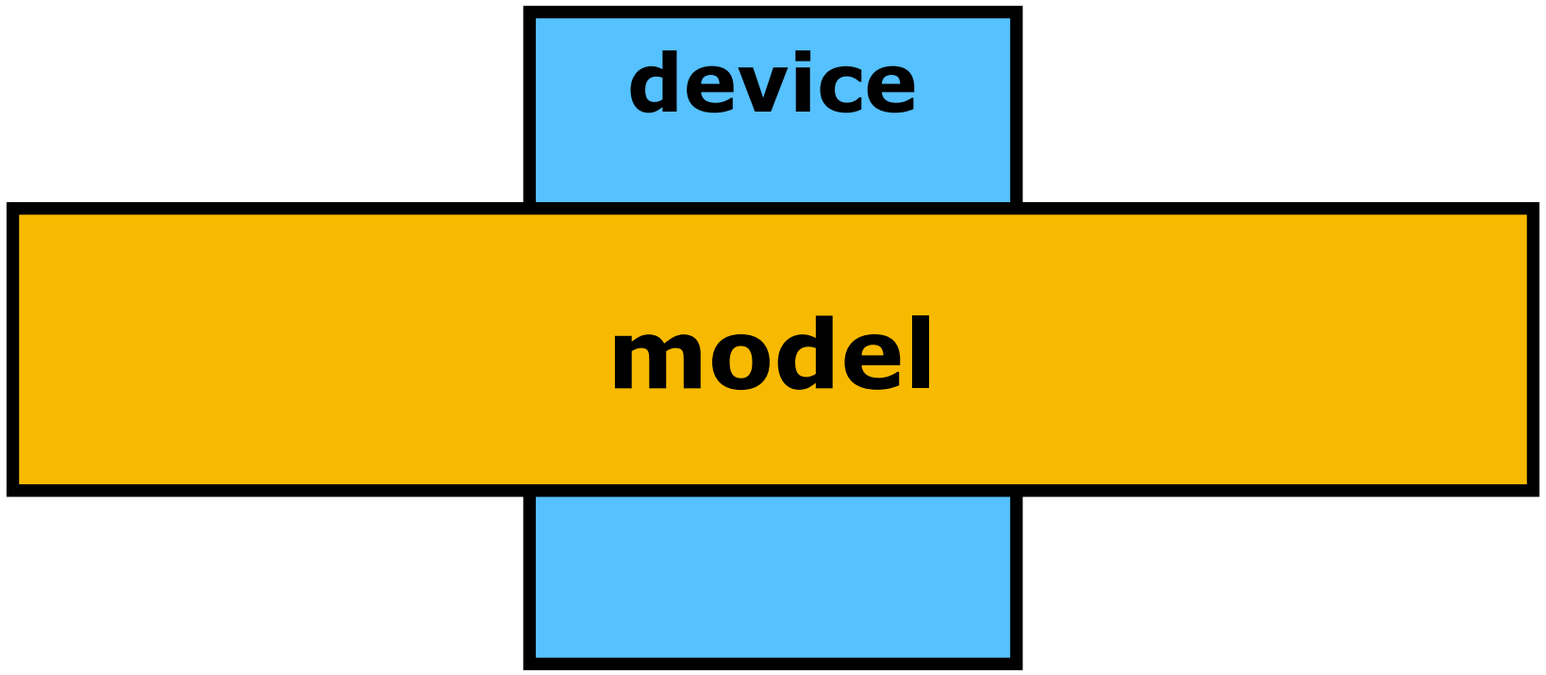
Synchronize weights θ to
train models fast.

Model Parallelization

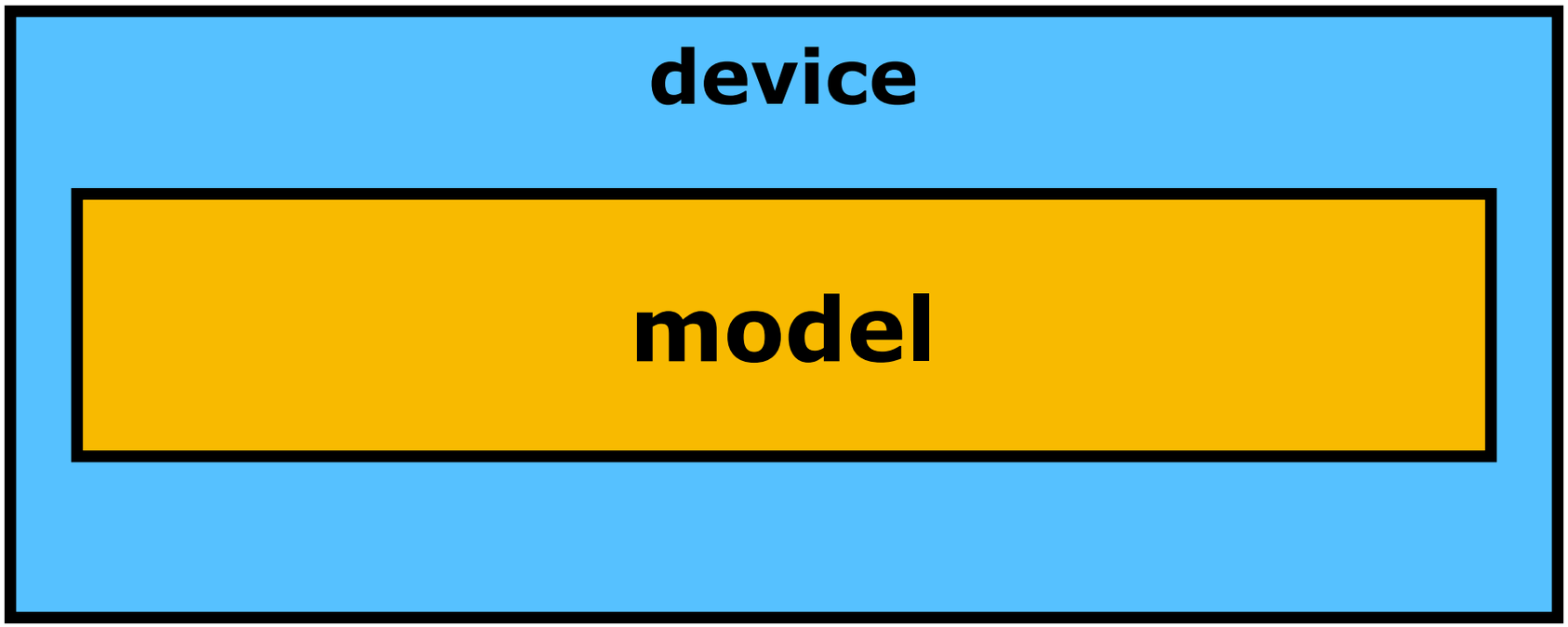


Partition a model to
build huge models.

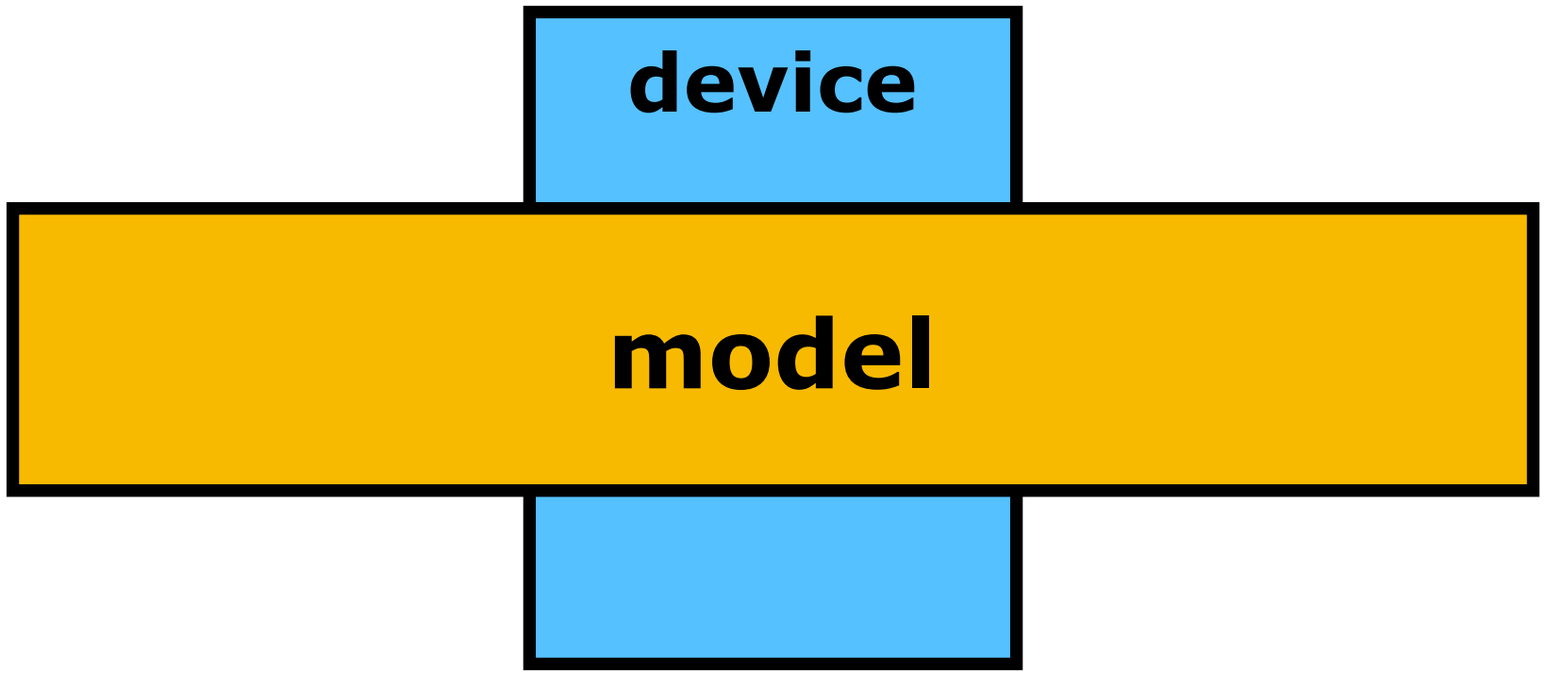
VALOHAI



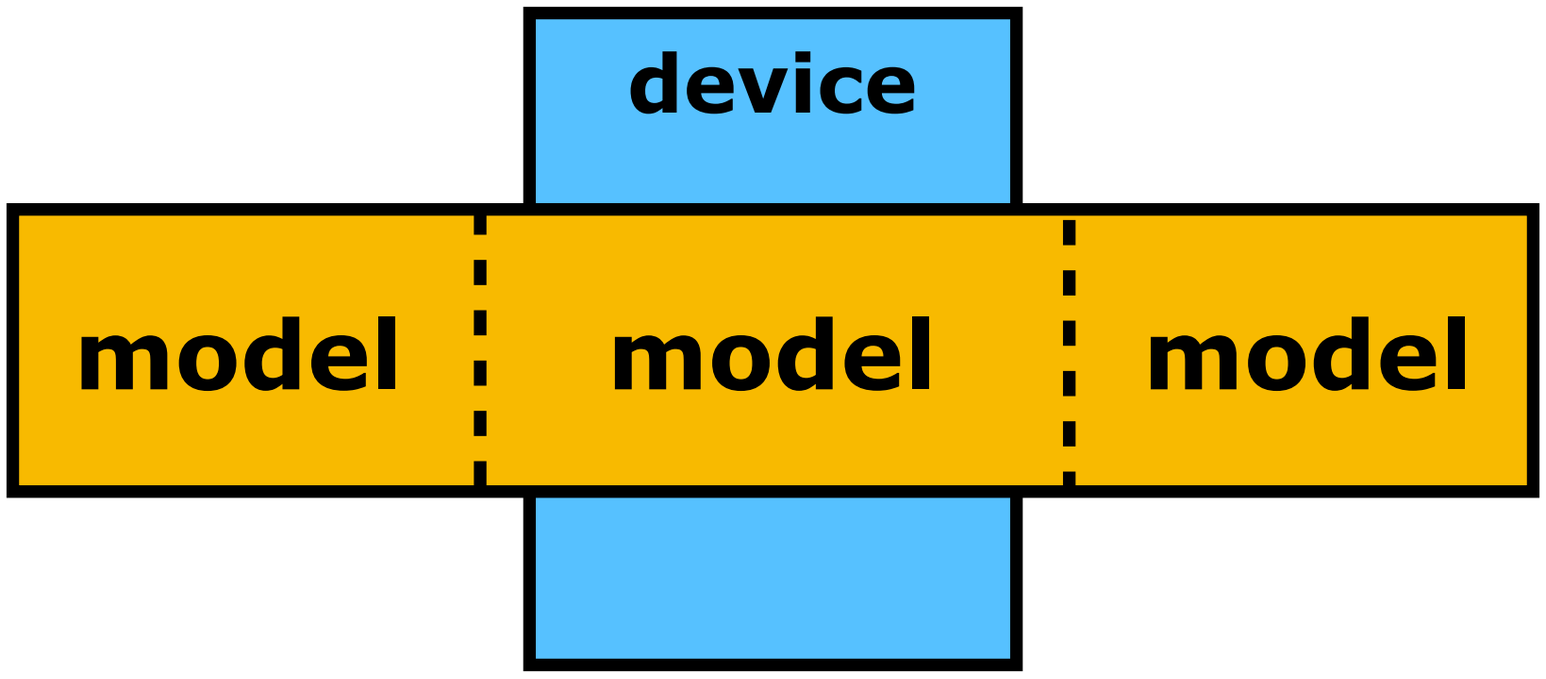
VALOHAI



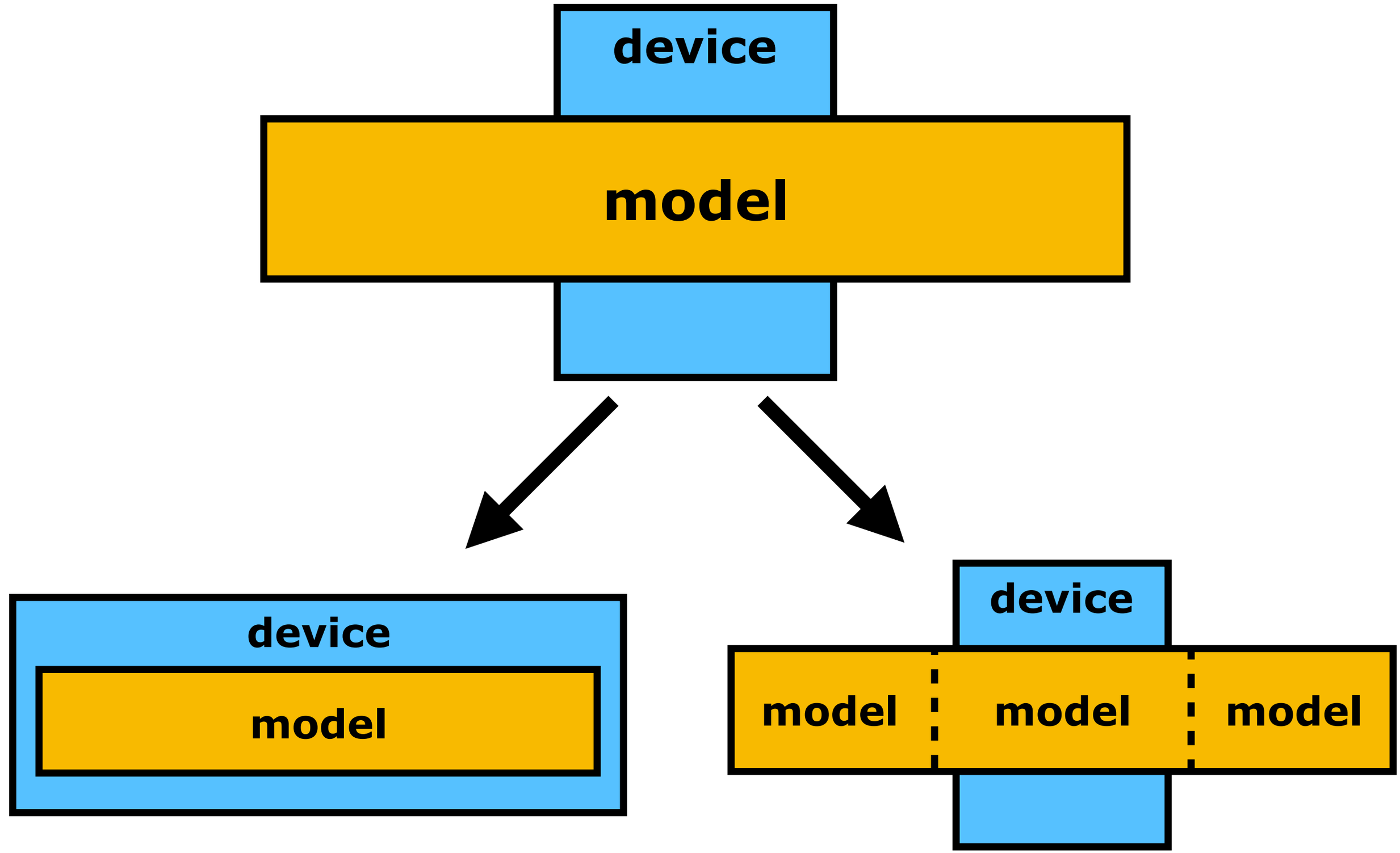
VALOHAI



VALOHAI



VALOHAI



VALOHAI

some training data

data

device

model

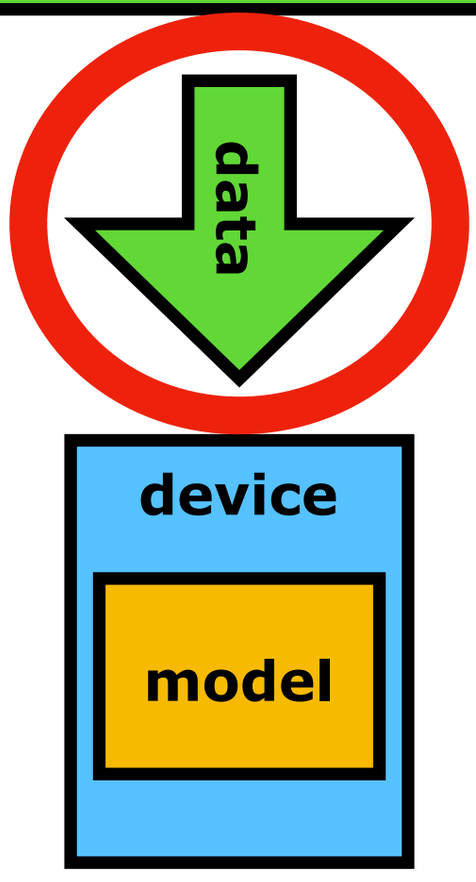
data

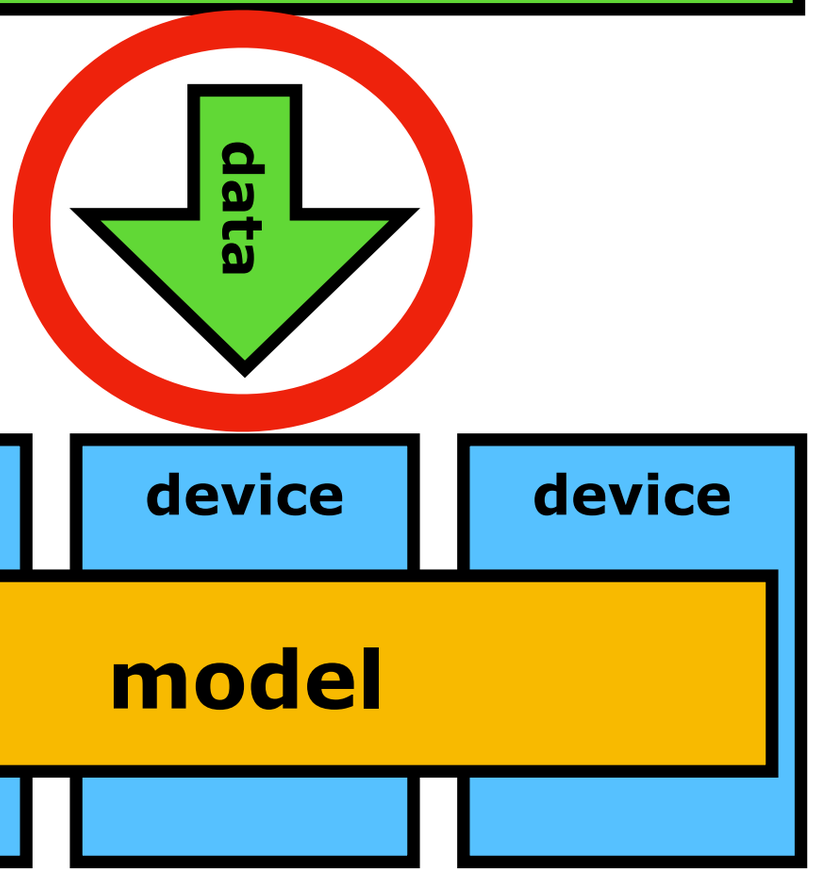
device | **device** | **device**

model

VALOHAI

a lot of training data





VALOHAI

a lot of training data

data

data

data

data

device **model**

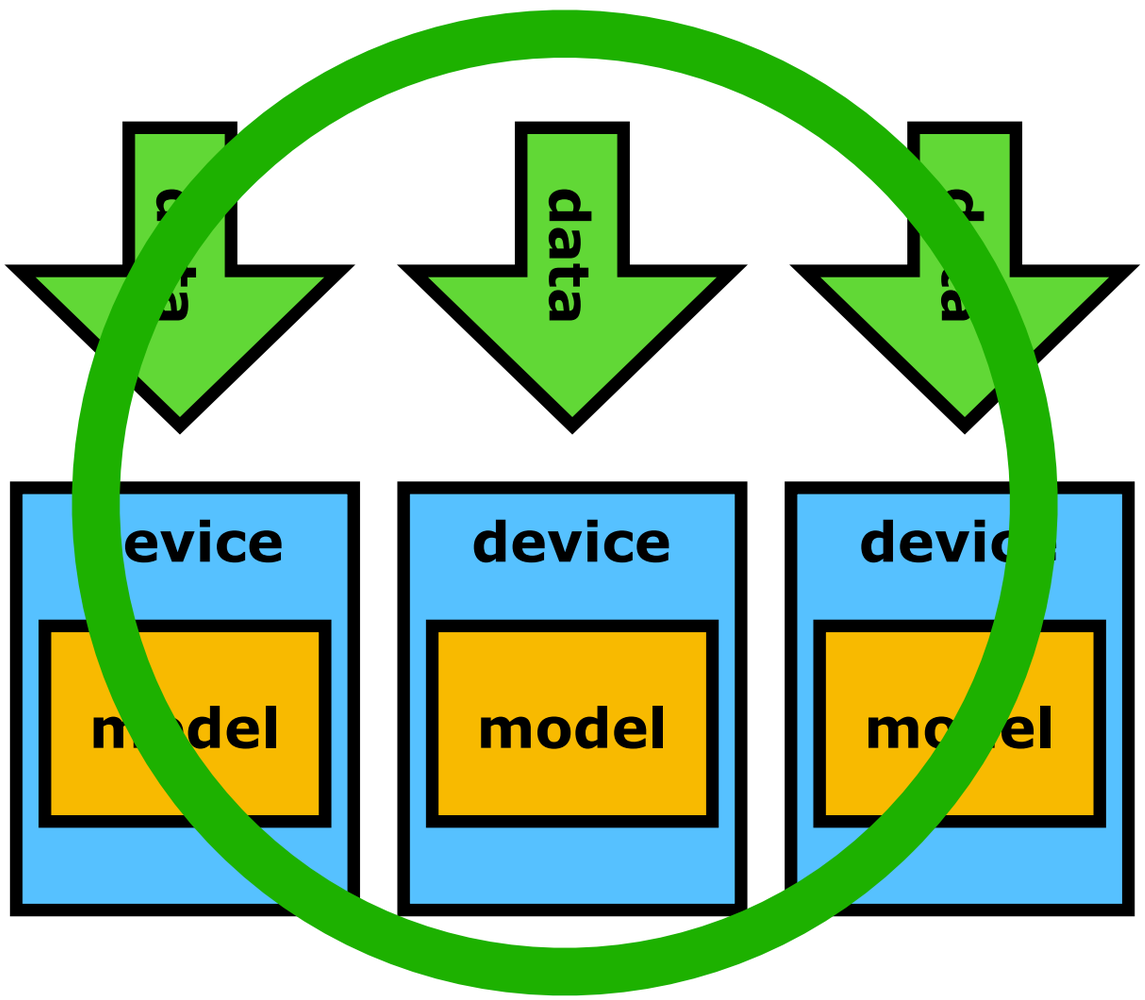
device **model**

device **model**

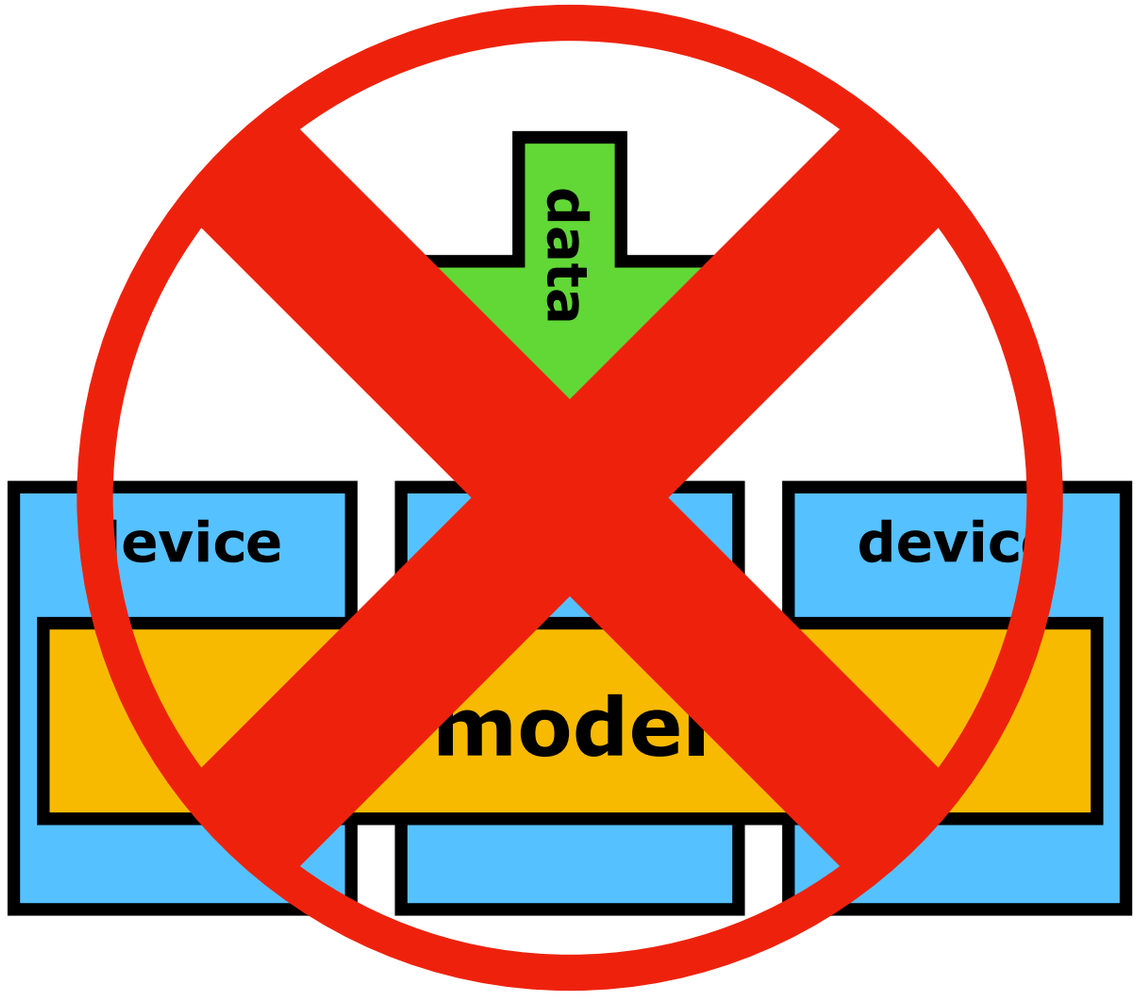
device **device** **device** **model**

VALOHAI

Data Parallelization



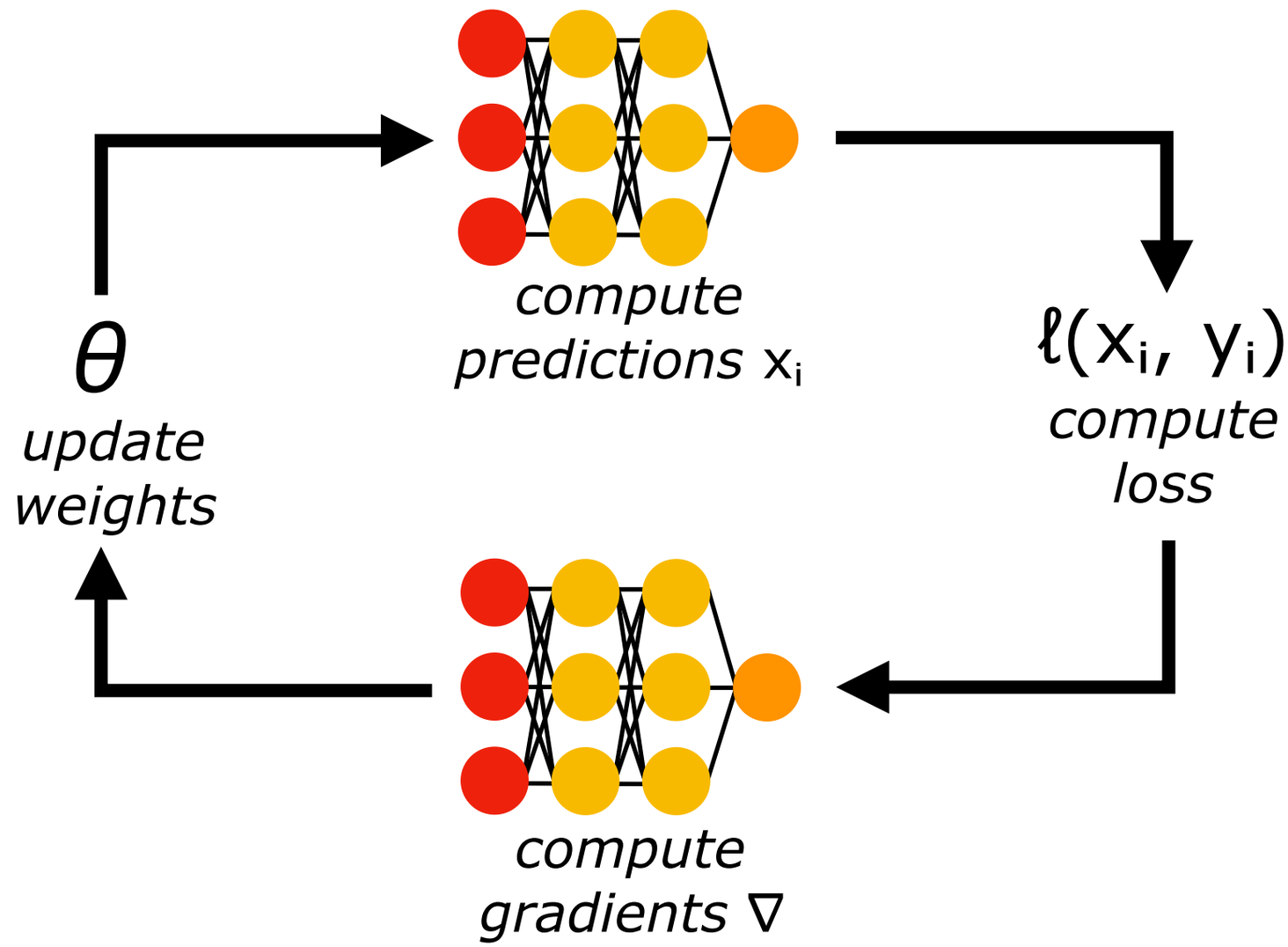
Model Parallelization



VALOHAI

The Brief History of Distributed Learning

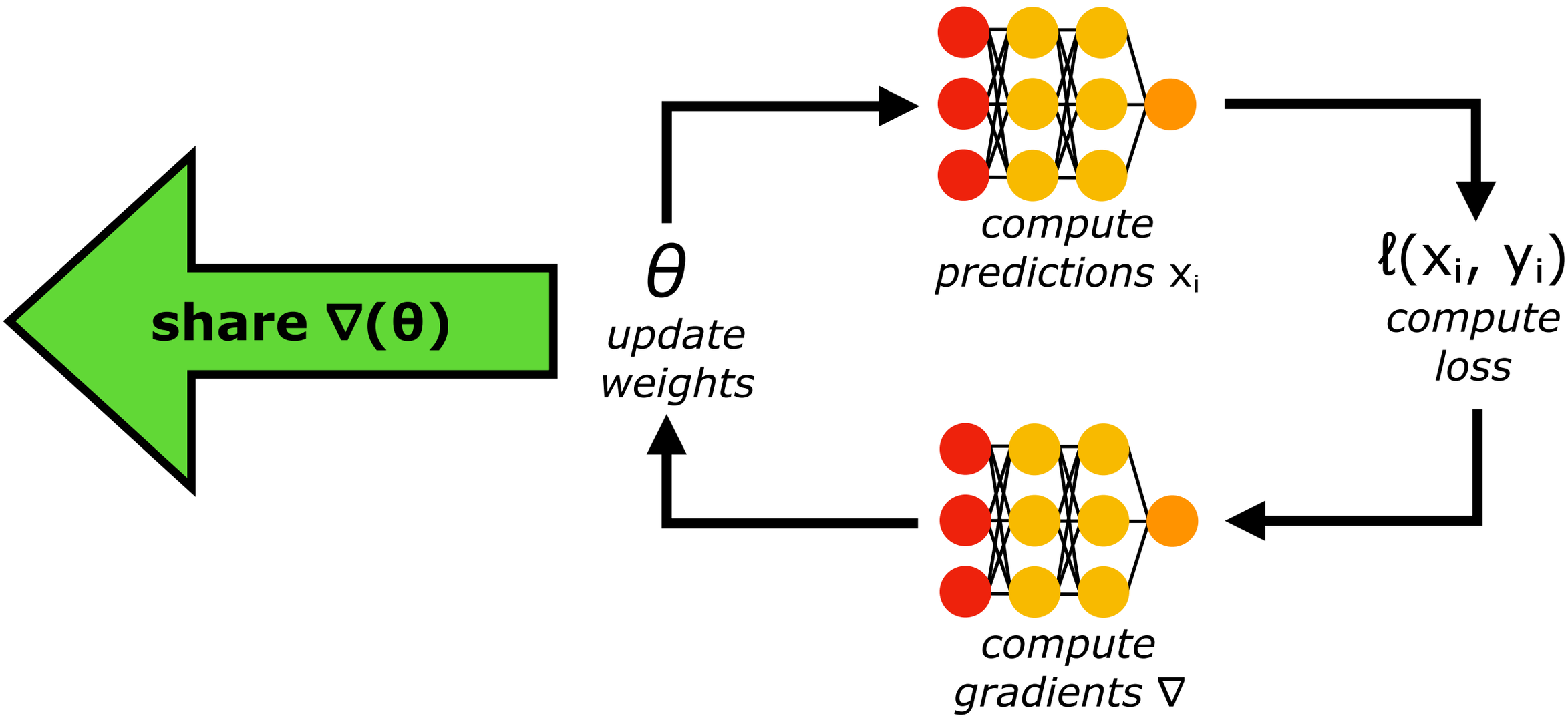
Shared Training Loop



VALOHAI

(innately sequential)

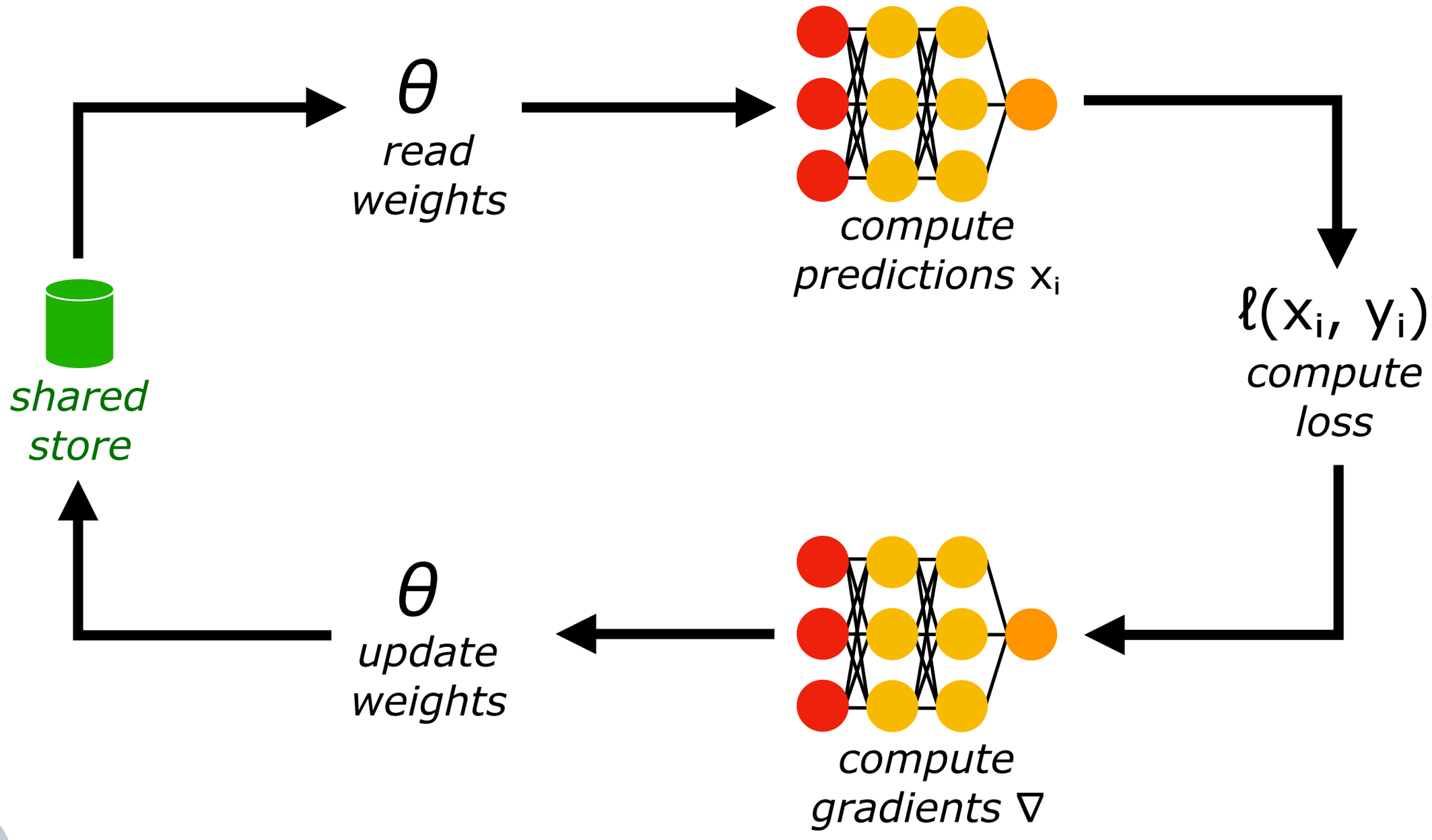
Shared Training Loop



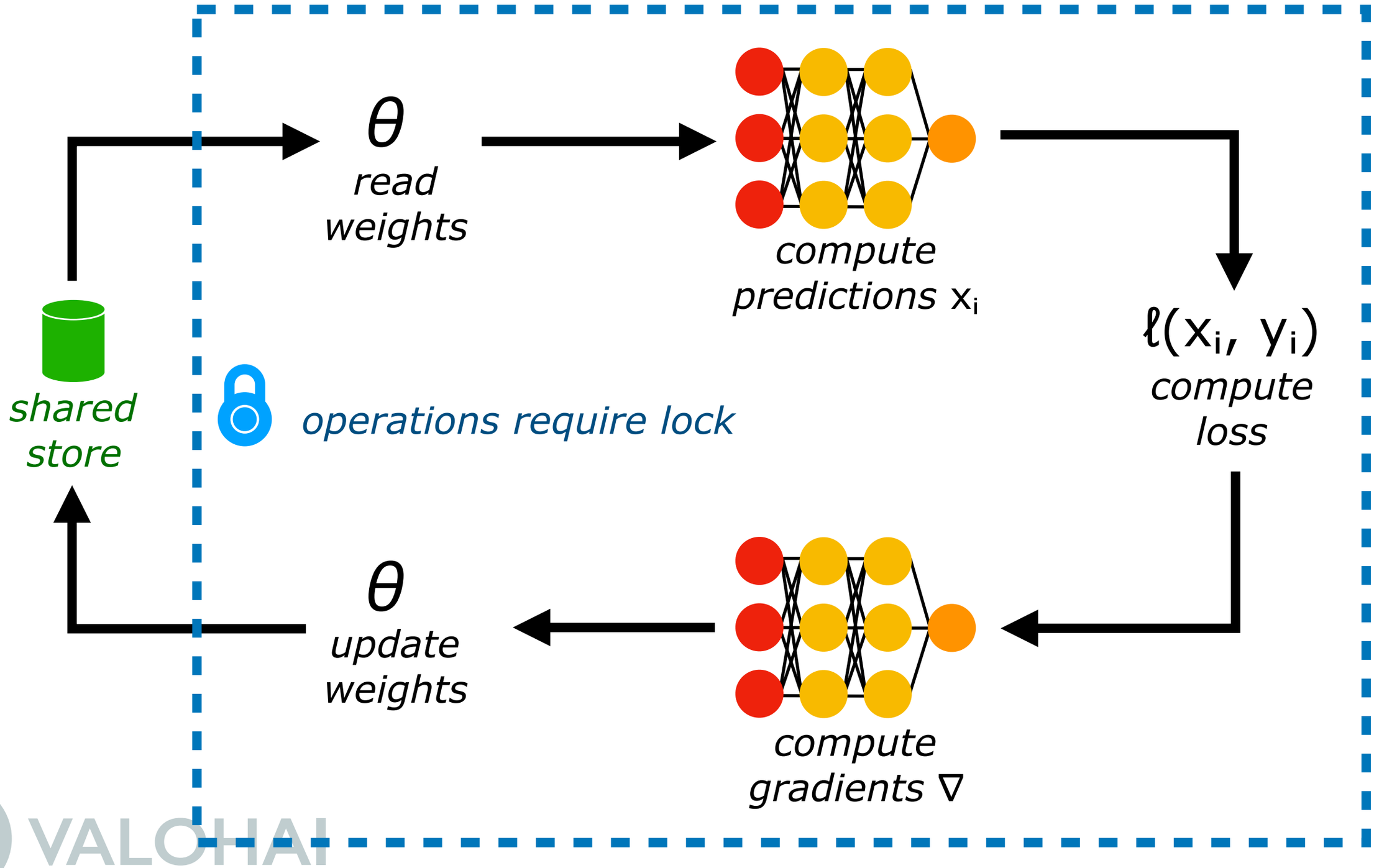
VALOHAI

(innately sequential)

Shared Training Loop

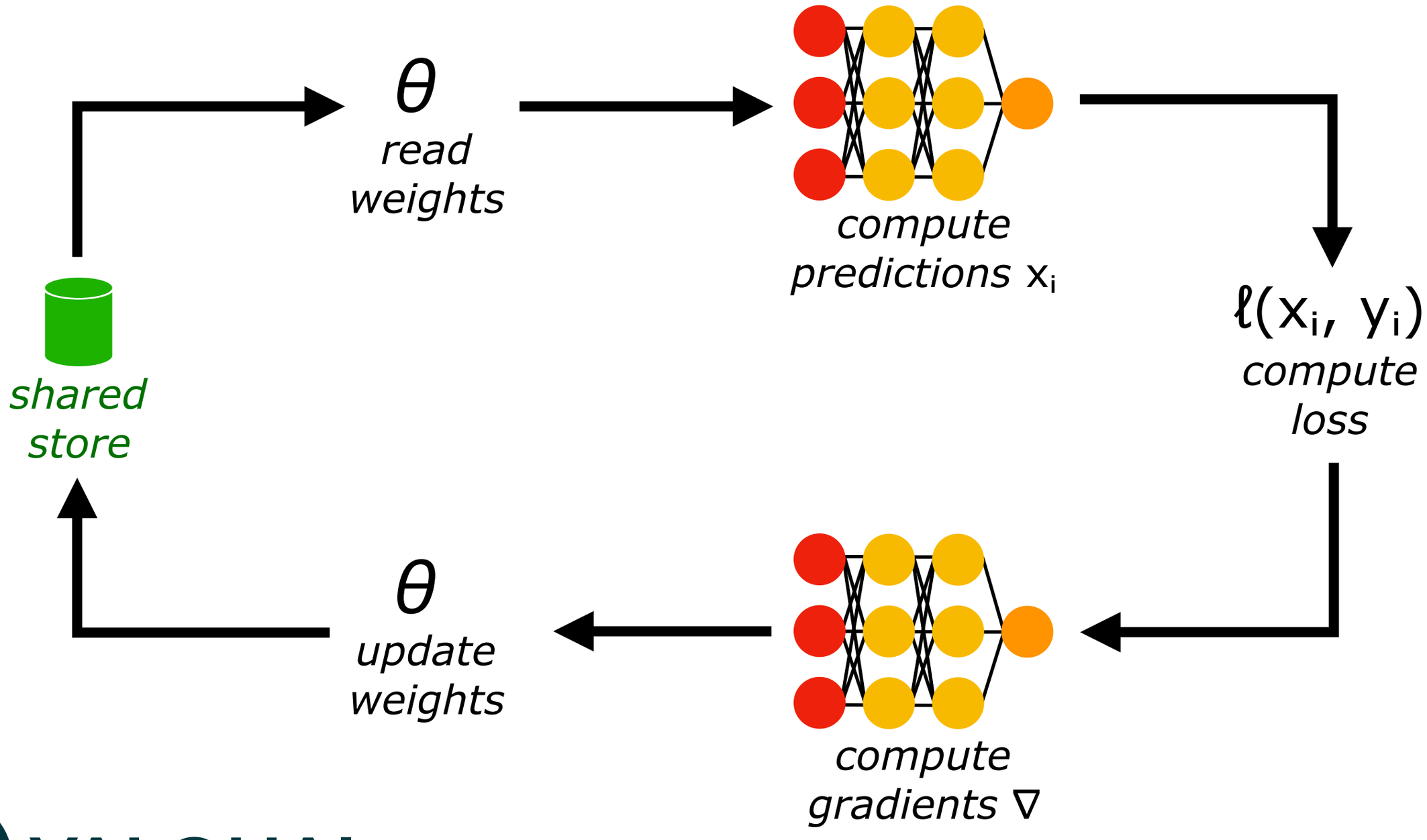


Locking Updates



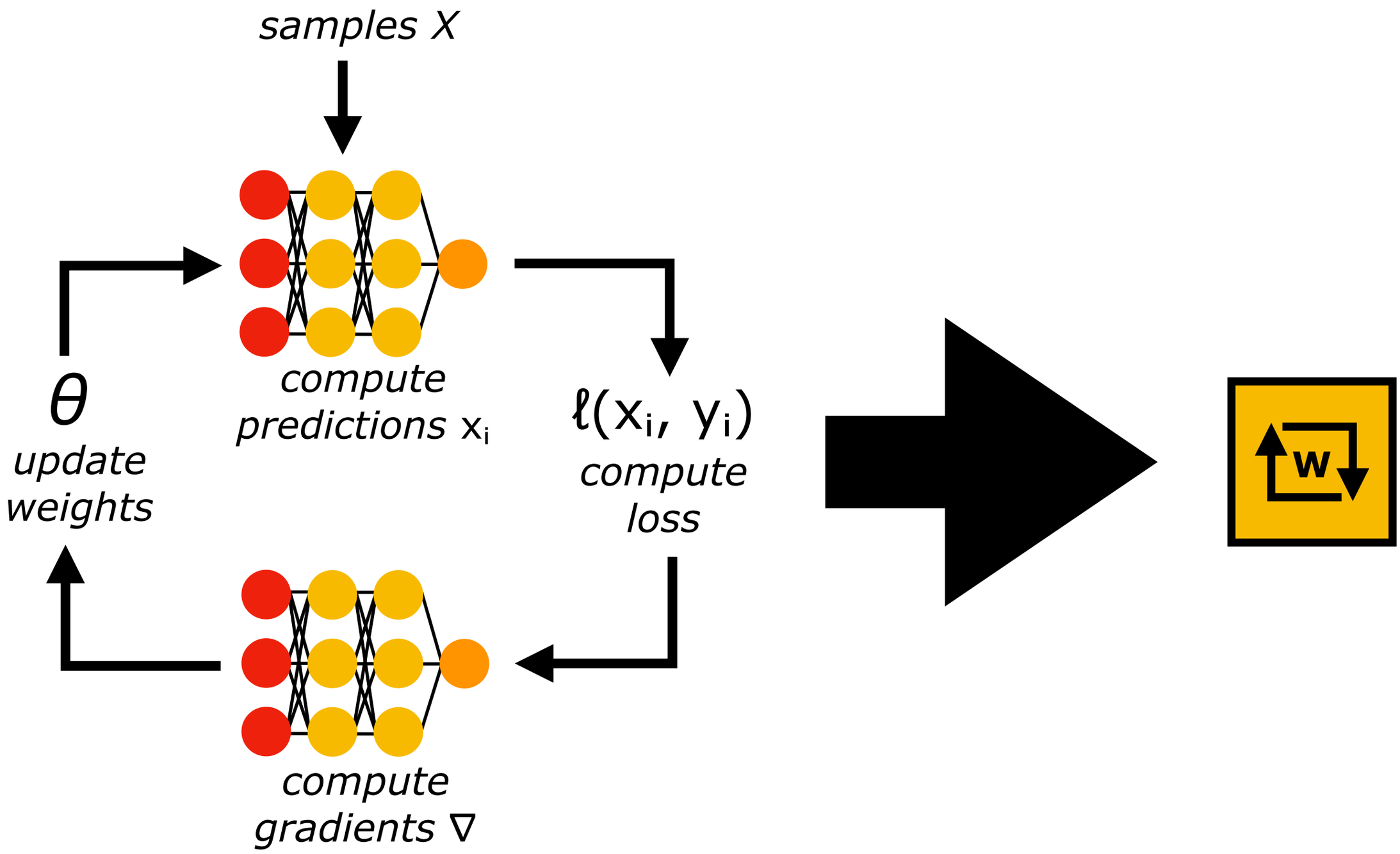
Lock-free Updates

(The Hogwild)



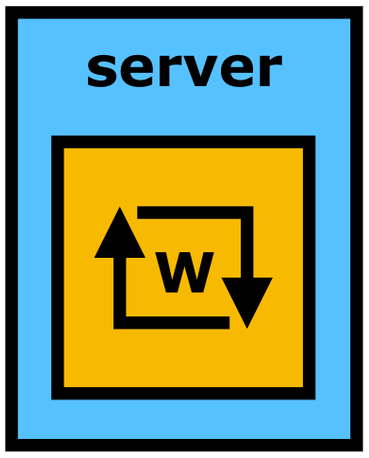
VALOHAI

Notation

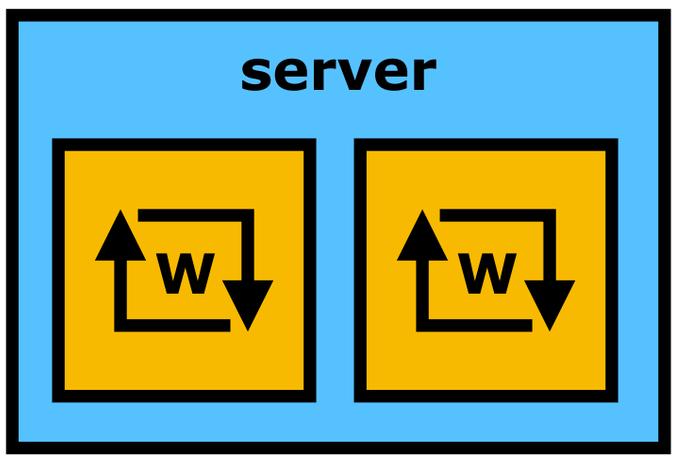


VALOHAI

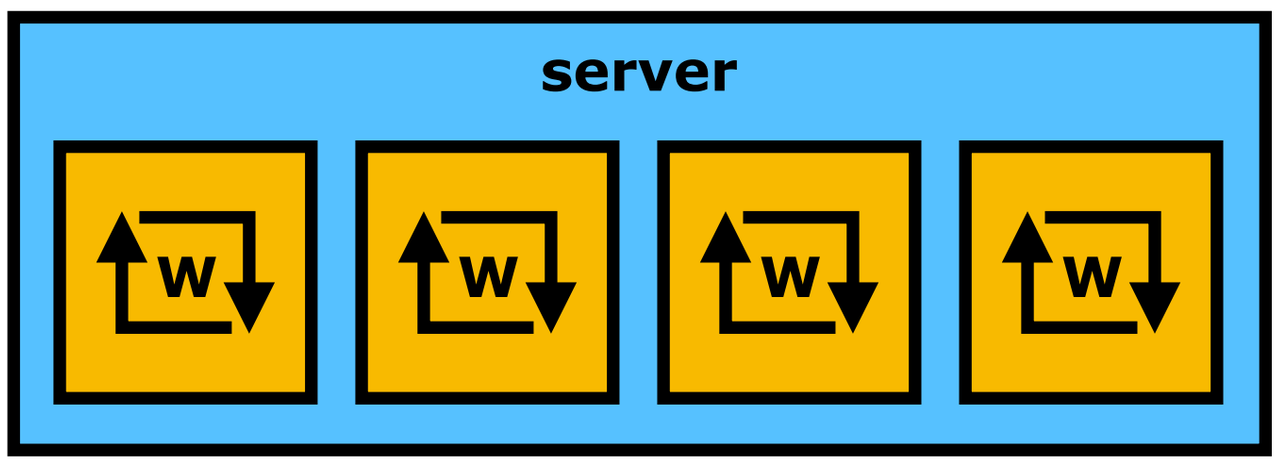
Notation



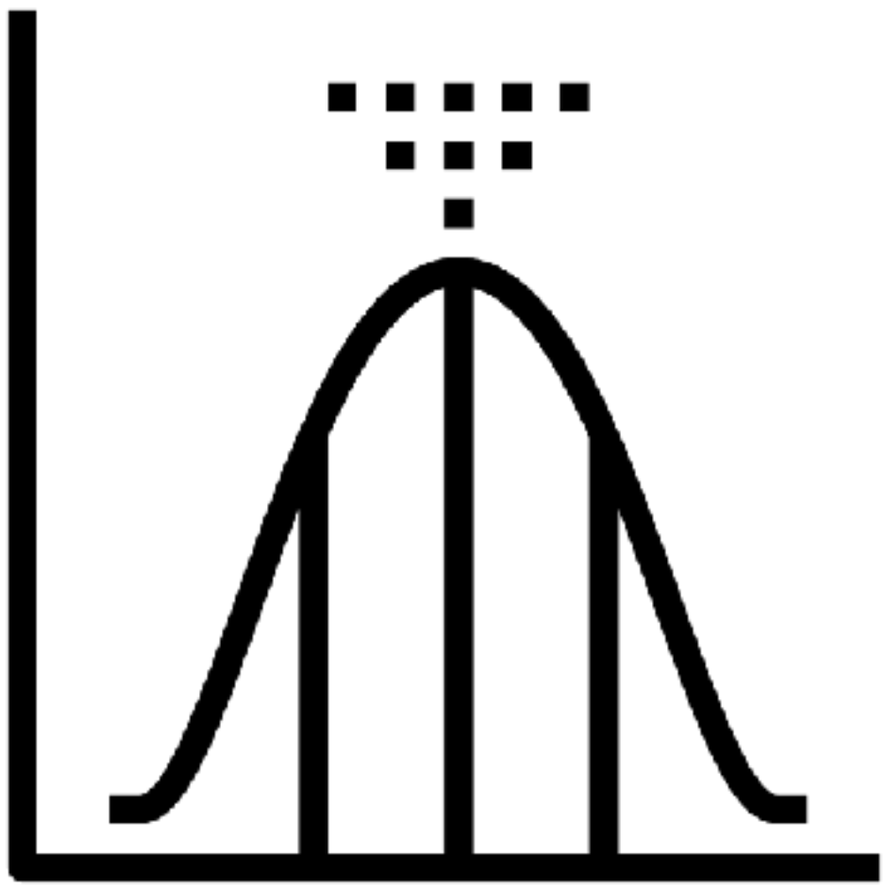
Usually
worker count == number of GPUs

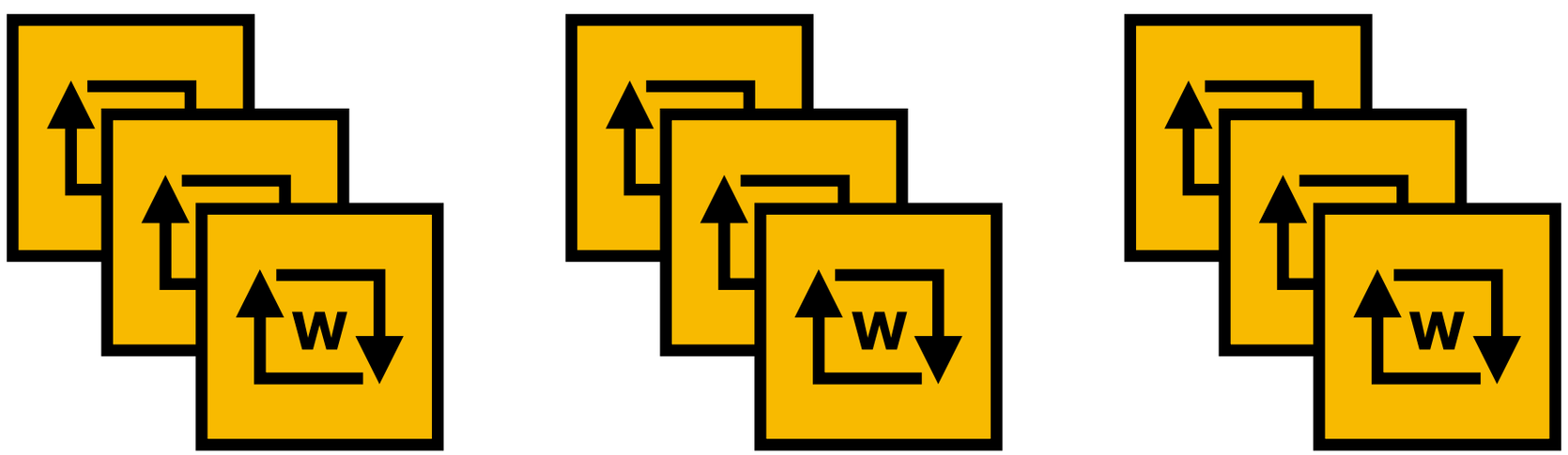


Each server has limited network bandwidth.



VALOHAI

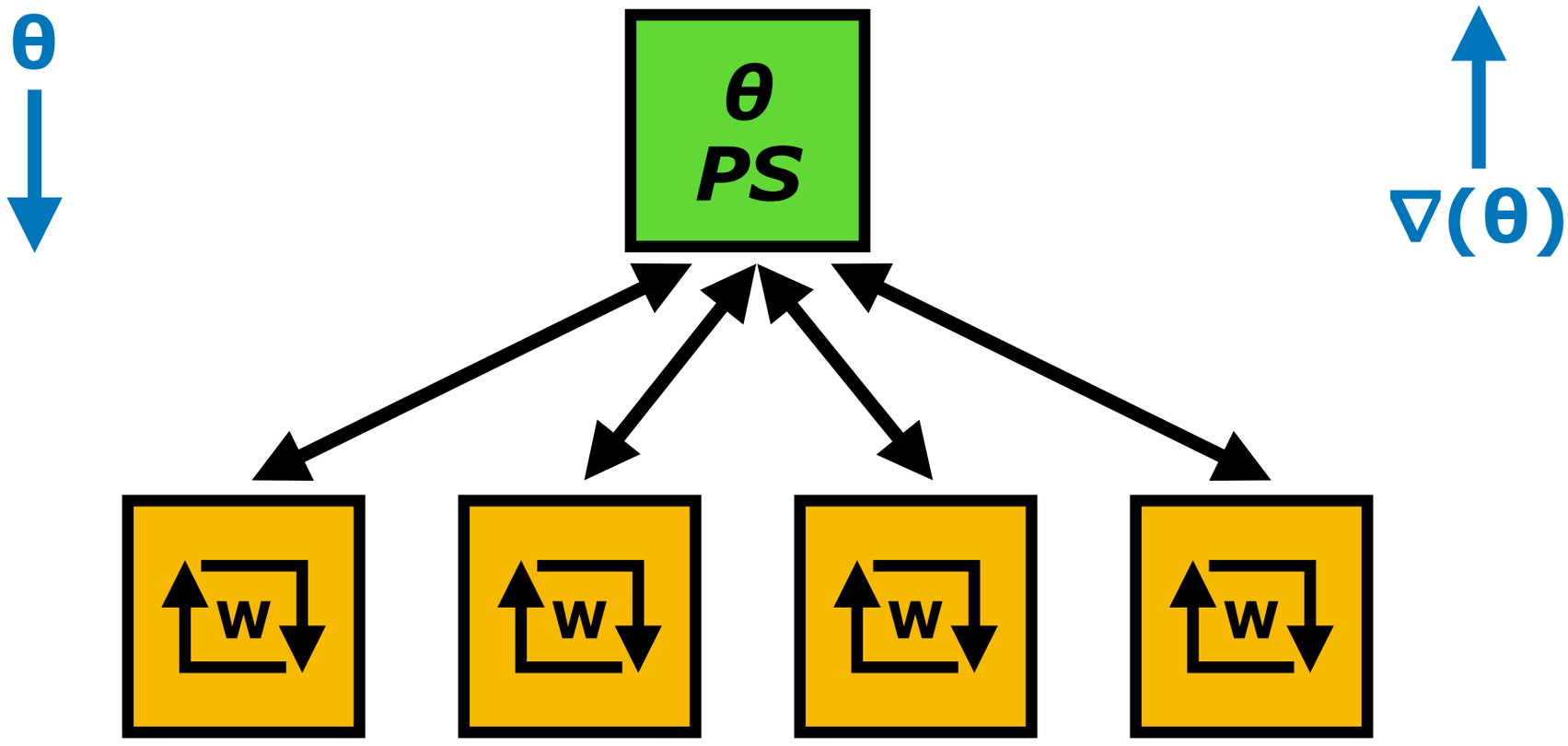




VALOHAI

Asynchronous Stochastic Gradient Descent

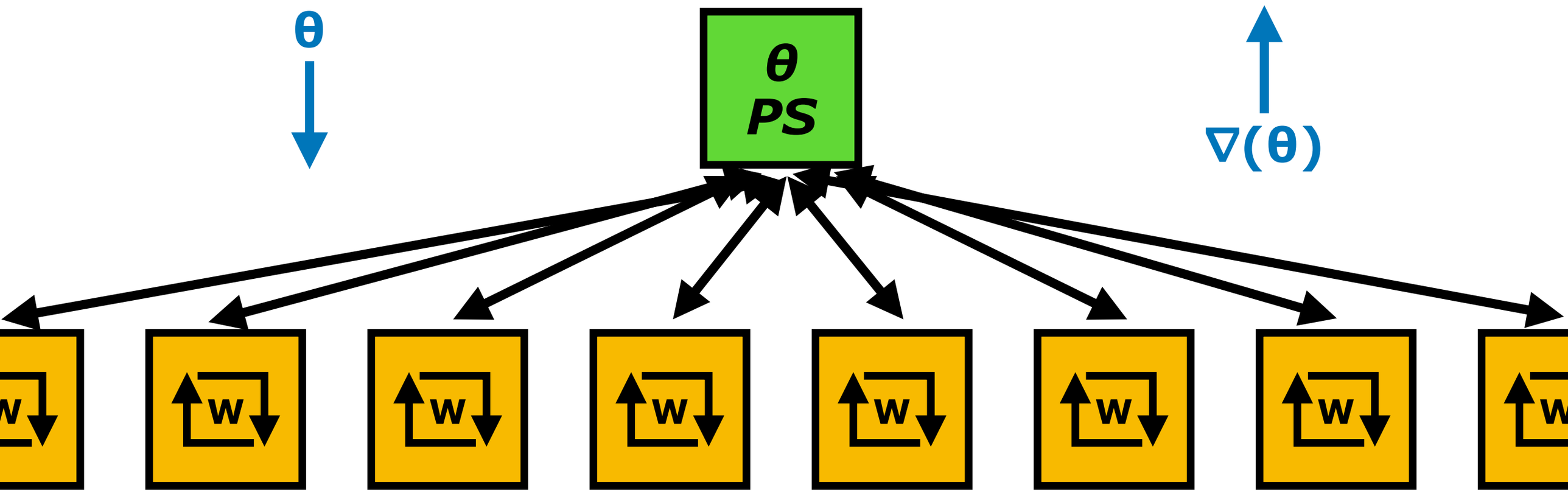
(Parameter Server)



VALOHAI

Asynchronous Stochastic Gradient Descent

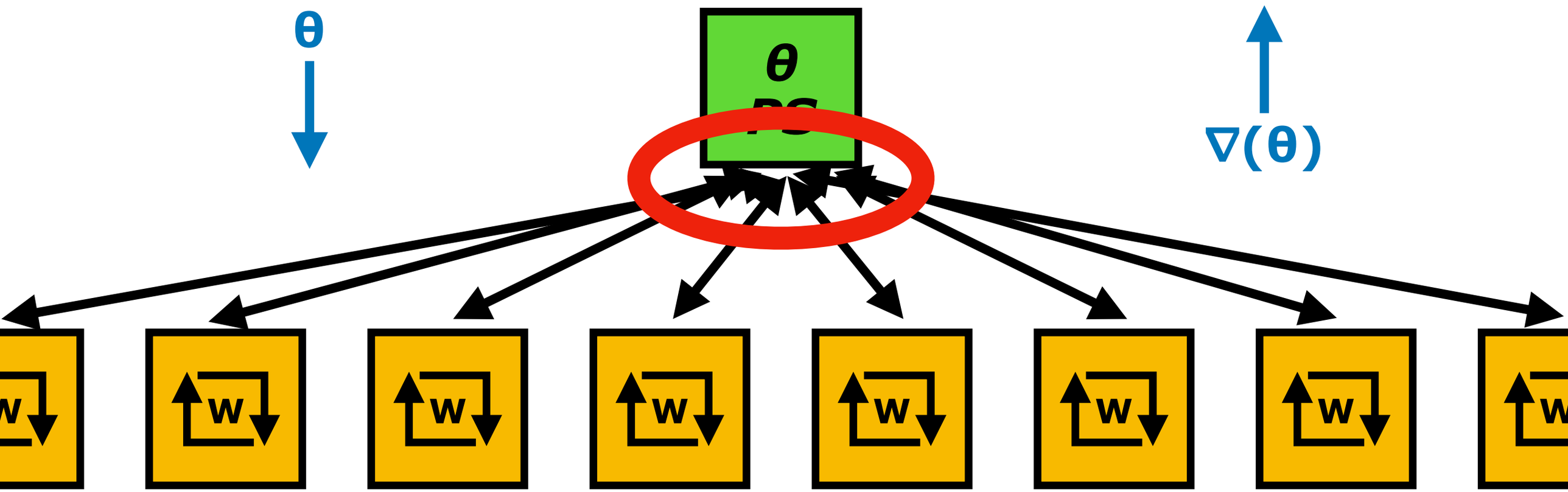
(Parameter Server)



VALOHAI

Asynchronous Stochastic Gradient Descent

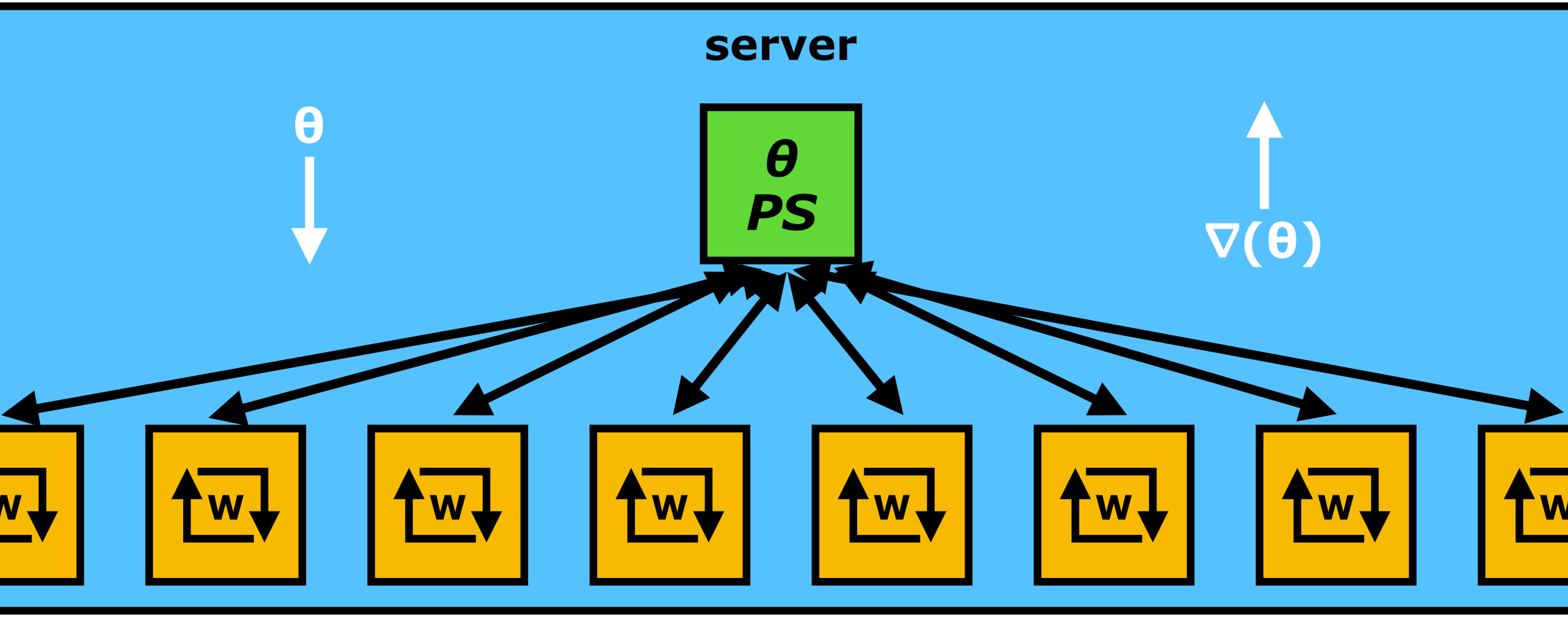
(Parameter Server)



VALOHAI

Asynchronous Stochastic Gradient Descent

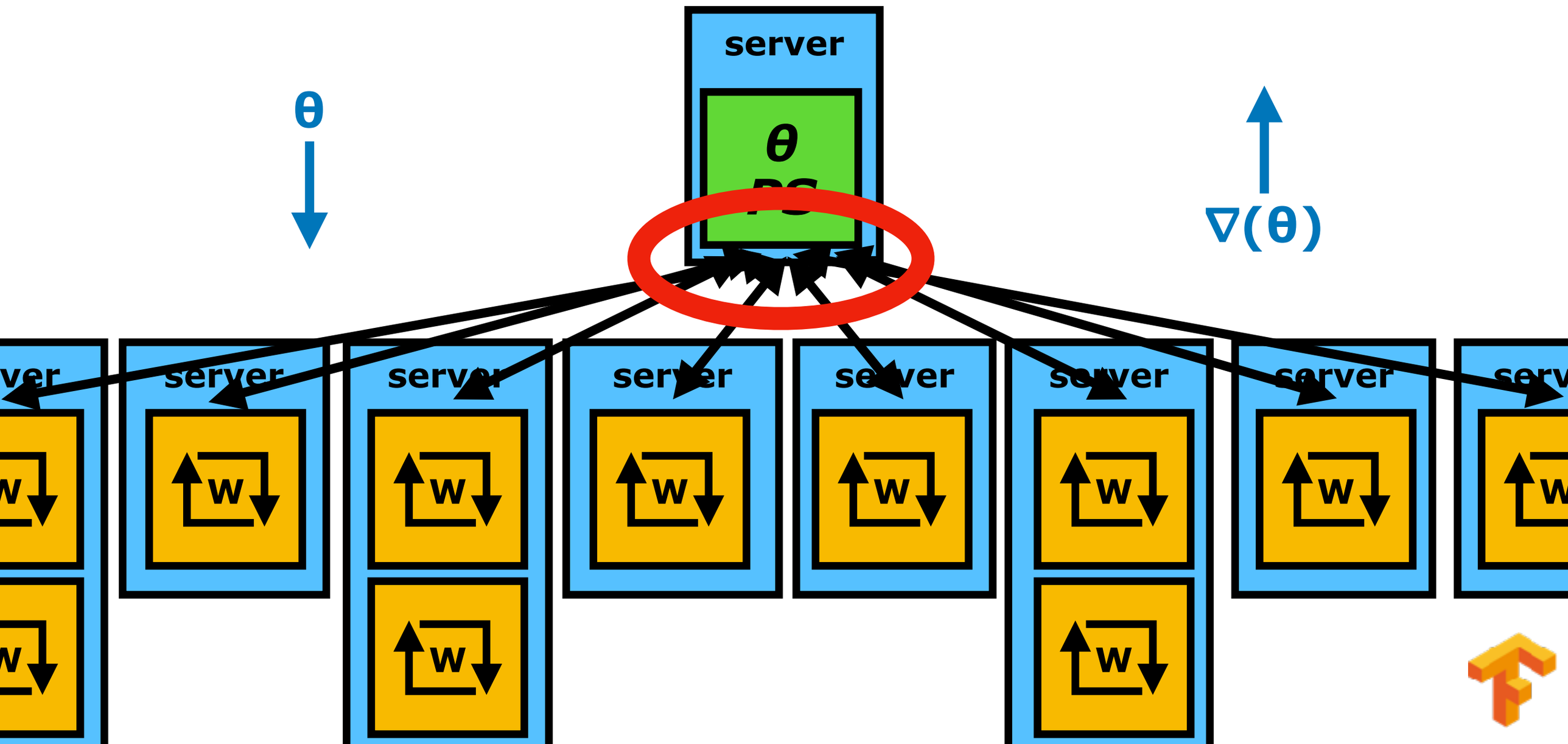
(Parameter Server)



VALOHAI

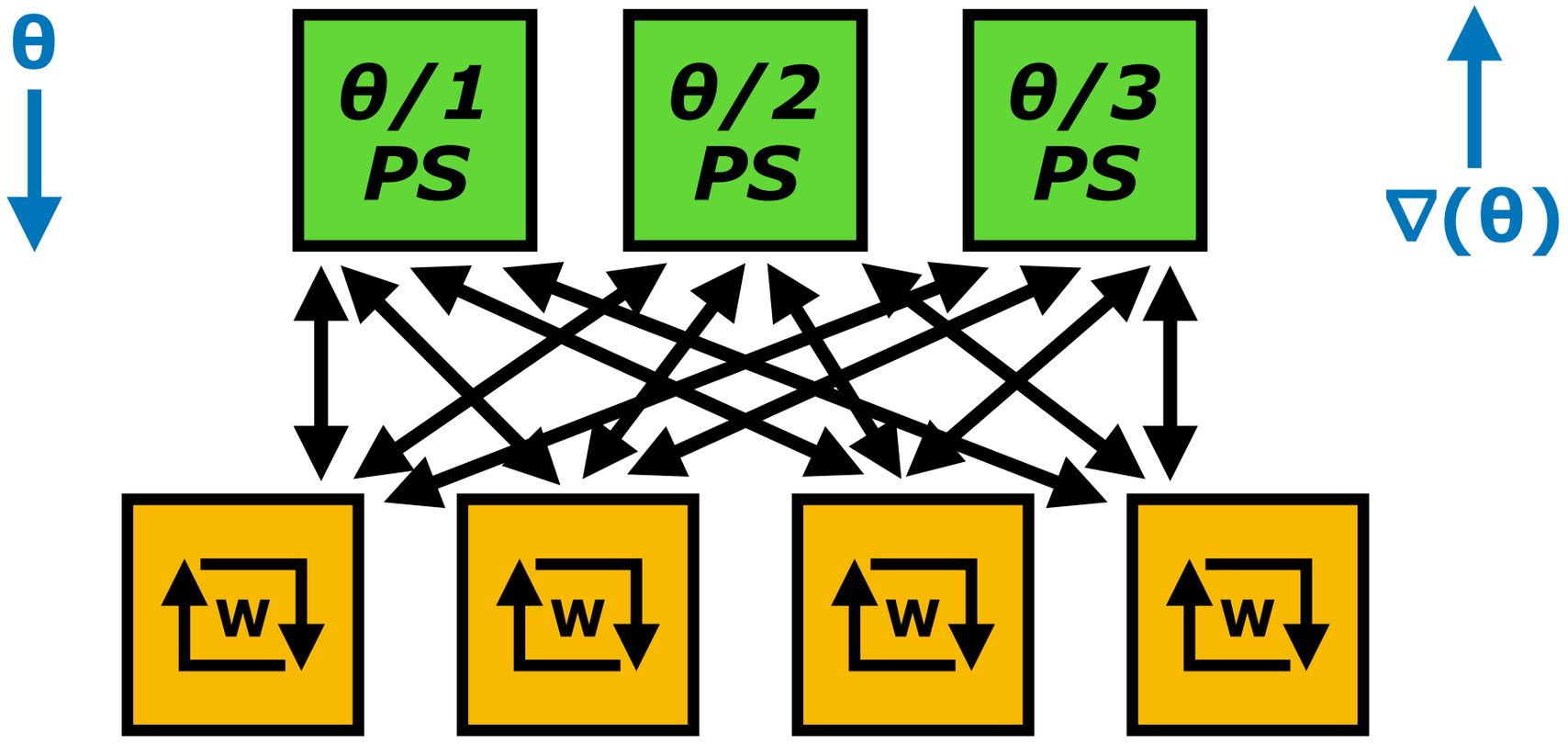
Asynchronous Stochastic Gradient Descent

(Parameter Server)



Asynchronous Stochastic Gradient Descent

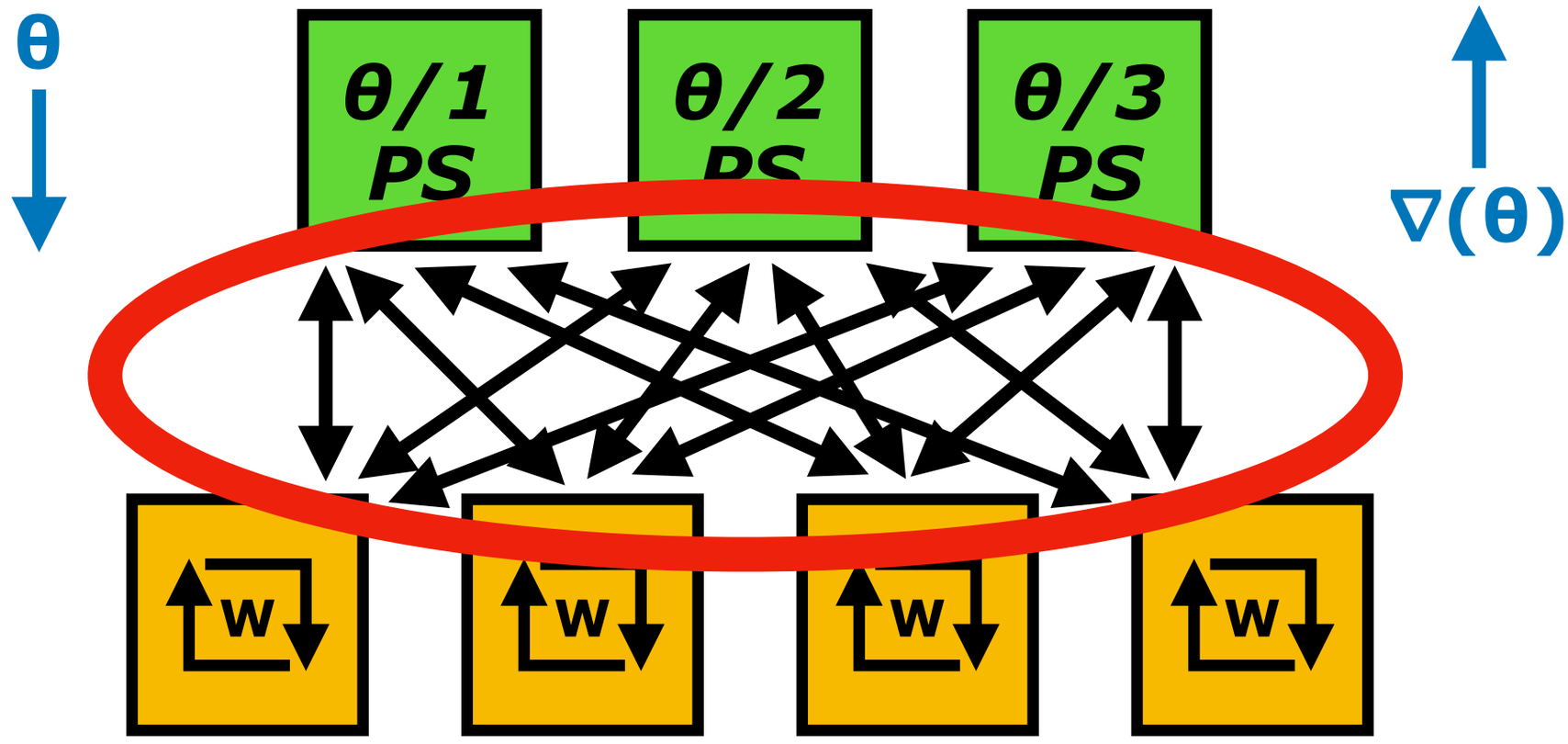
(Parameter Server Cluster)



VALOHAI

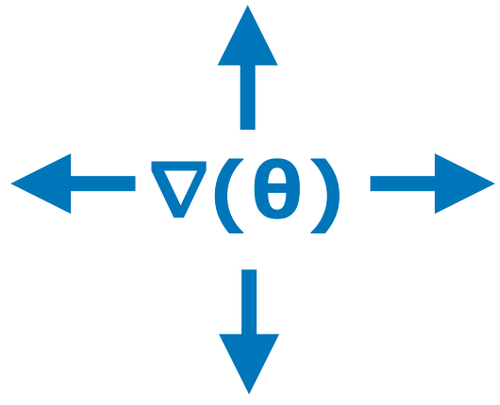
Asynchronous Stochastic Gradient Descent

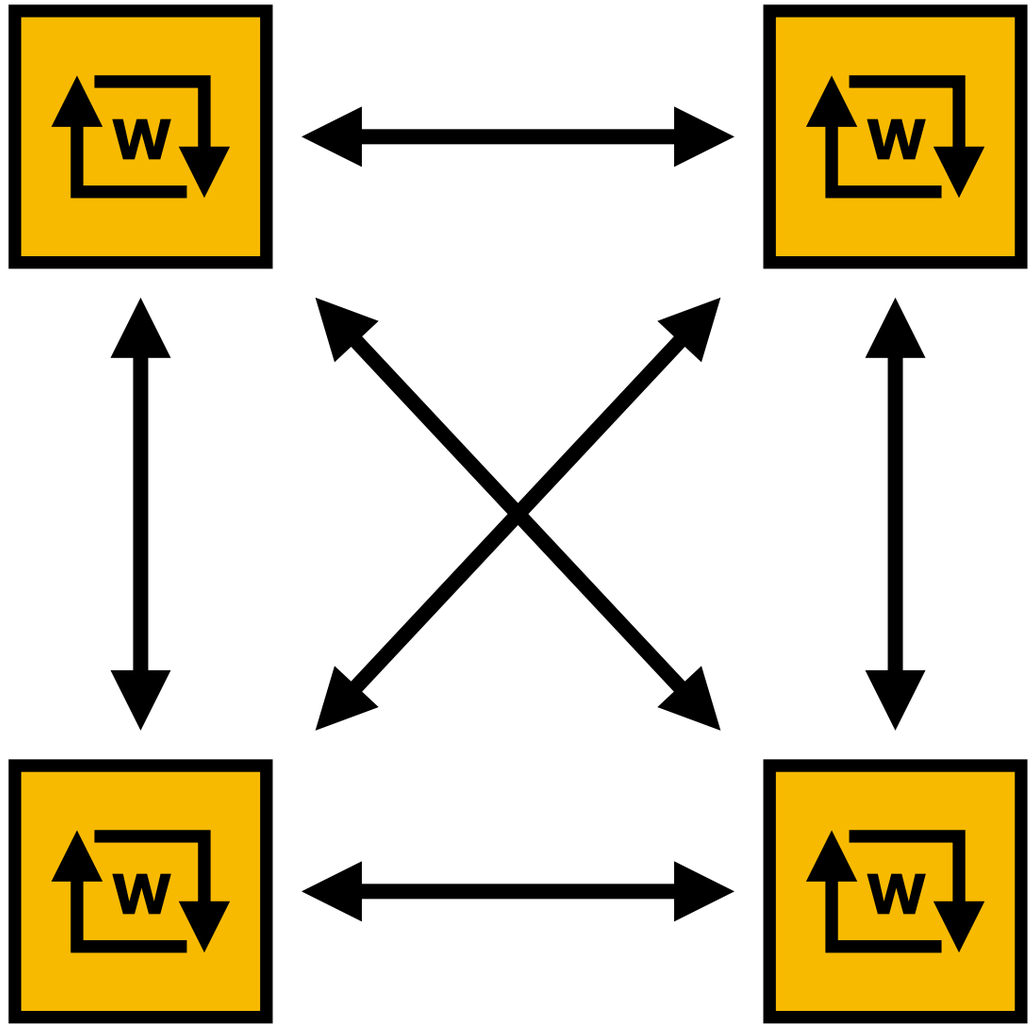
(Parameter Server Cluster)



VALOHAI

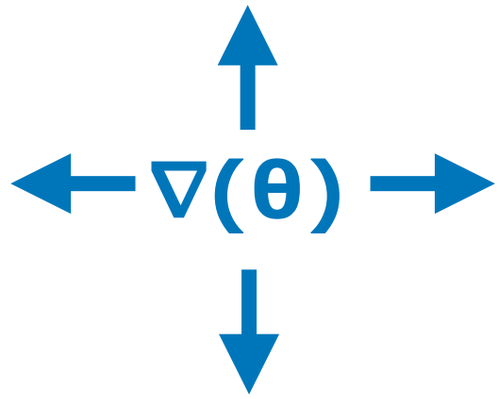
Synchronous Allreduce

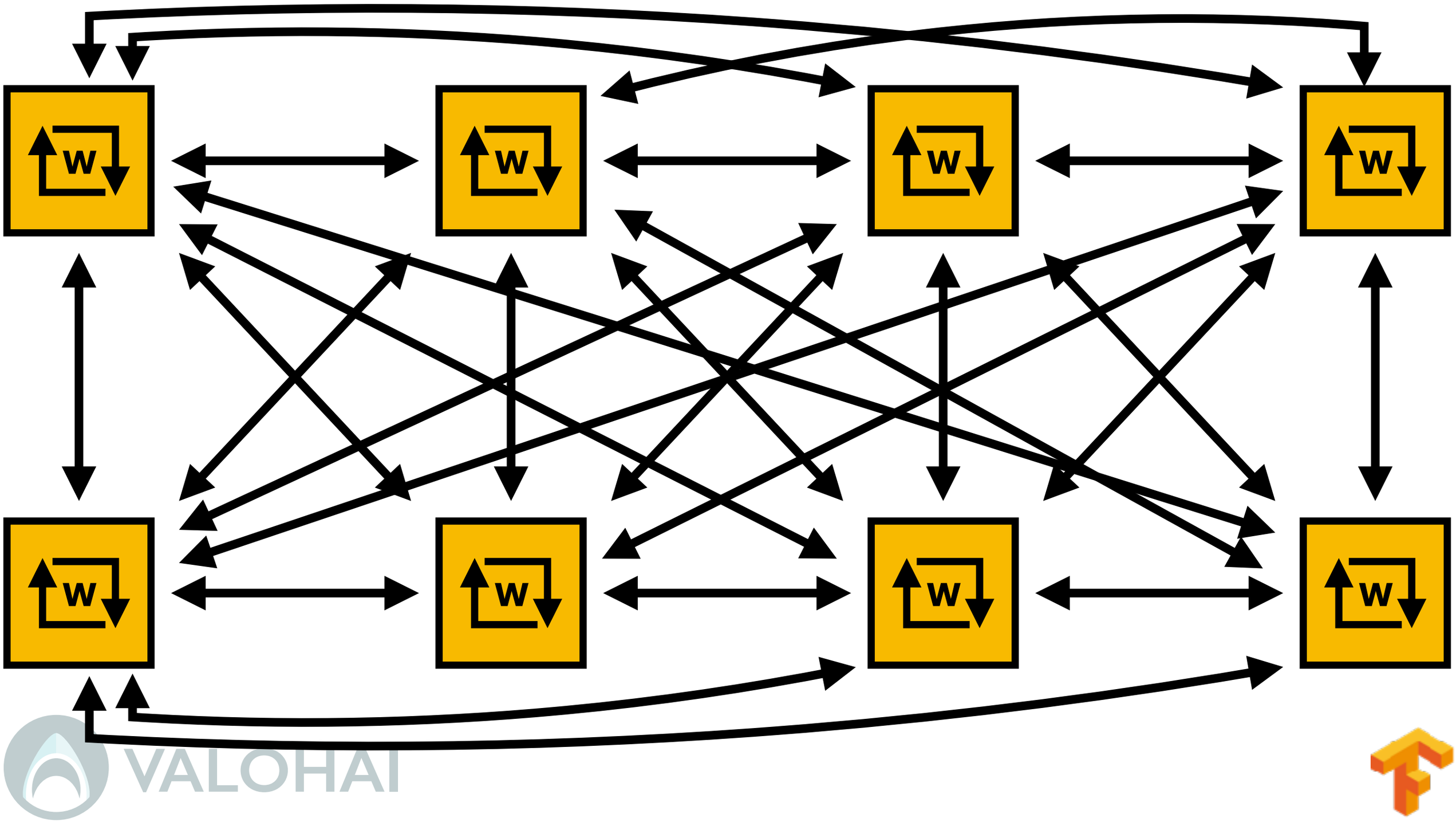




VALOHAI

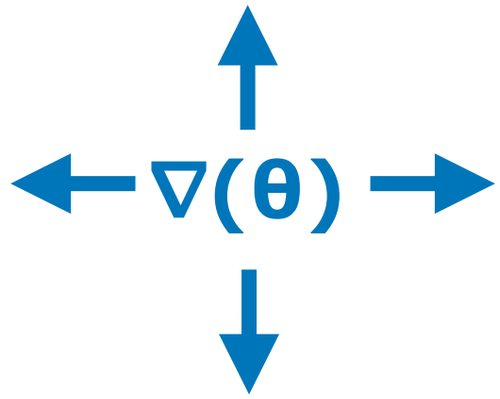
Synchronous Allreduce

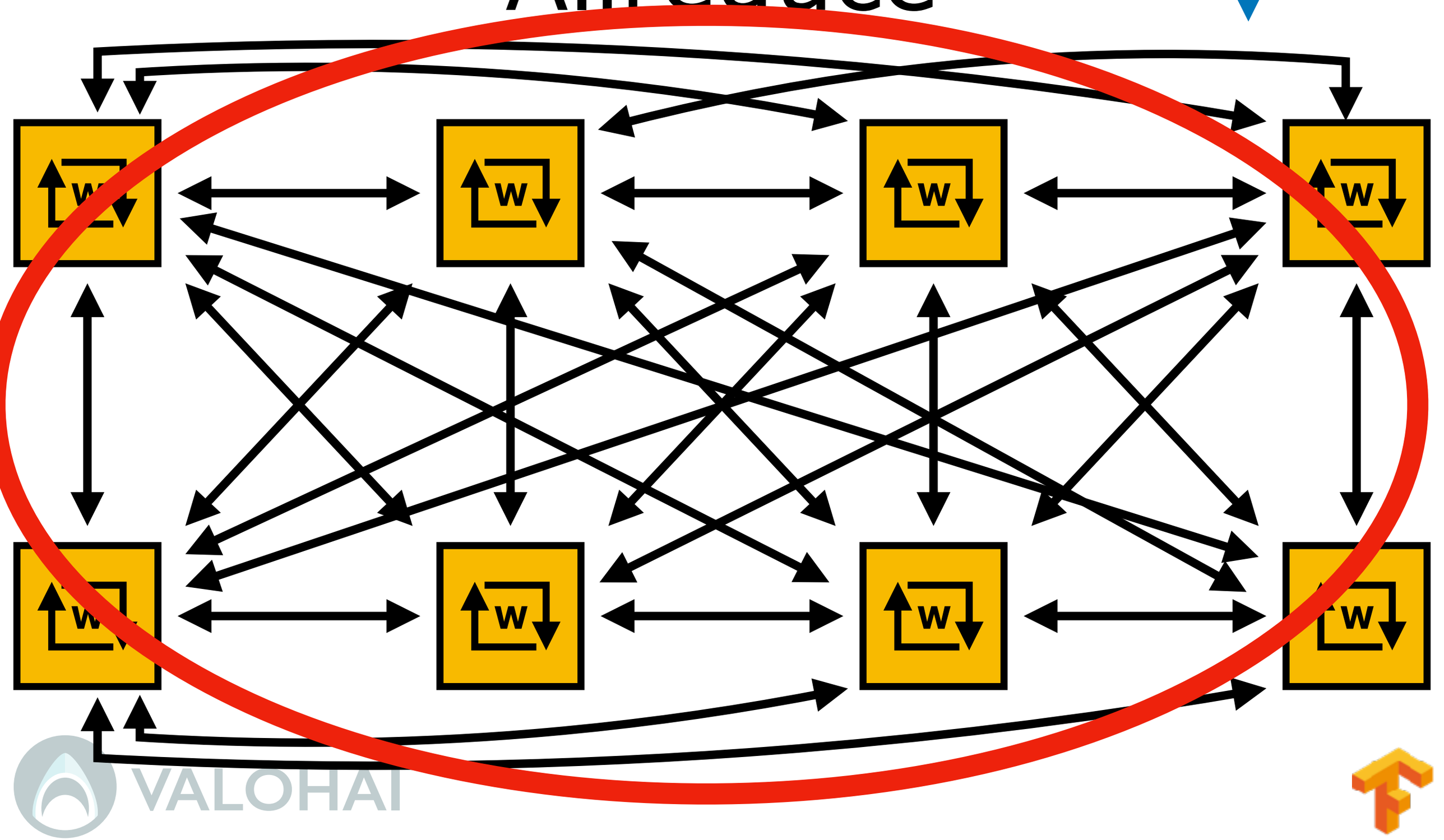




VALOHAI

Synchronous Allreduce





VALOHAI

Butterfly Allreduce

VALOHAI

1

Worker 1 0

Worker 2 1

Worker 3 2

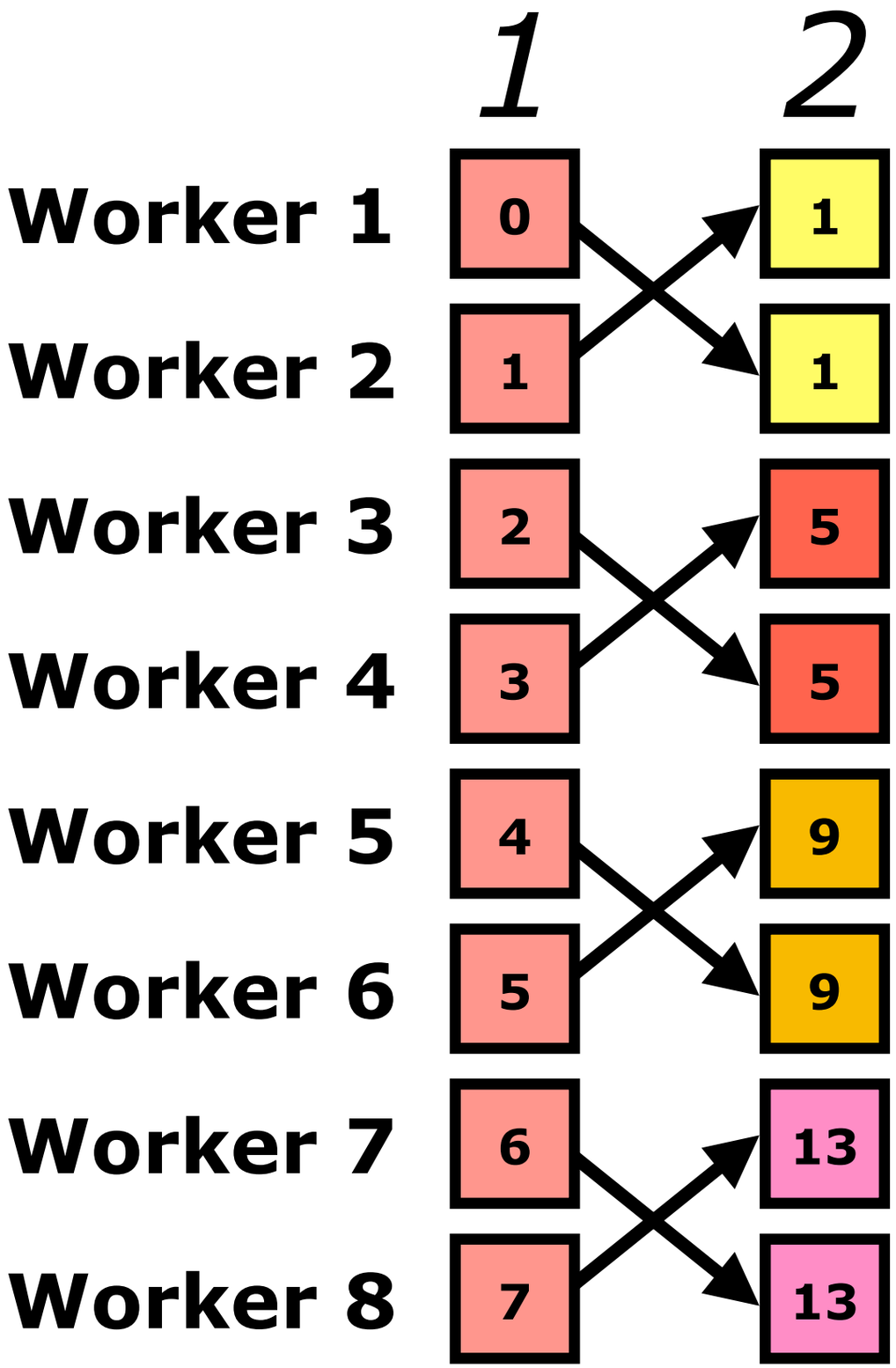
Worker 4 3

Worker 5 4

Worker 6 5

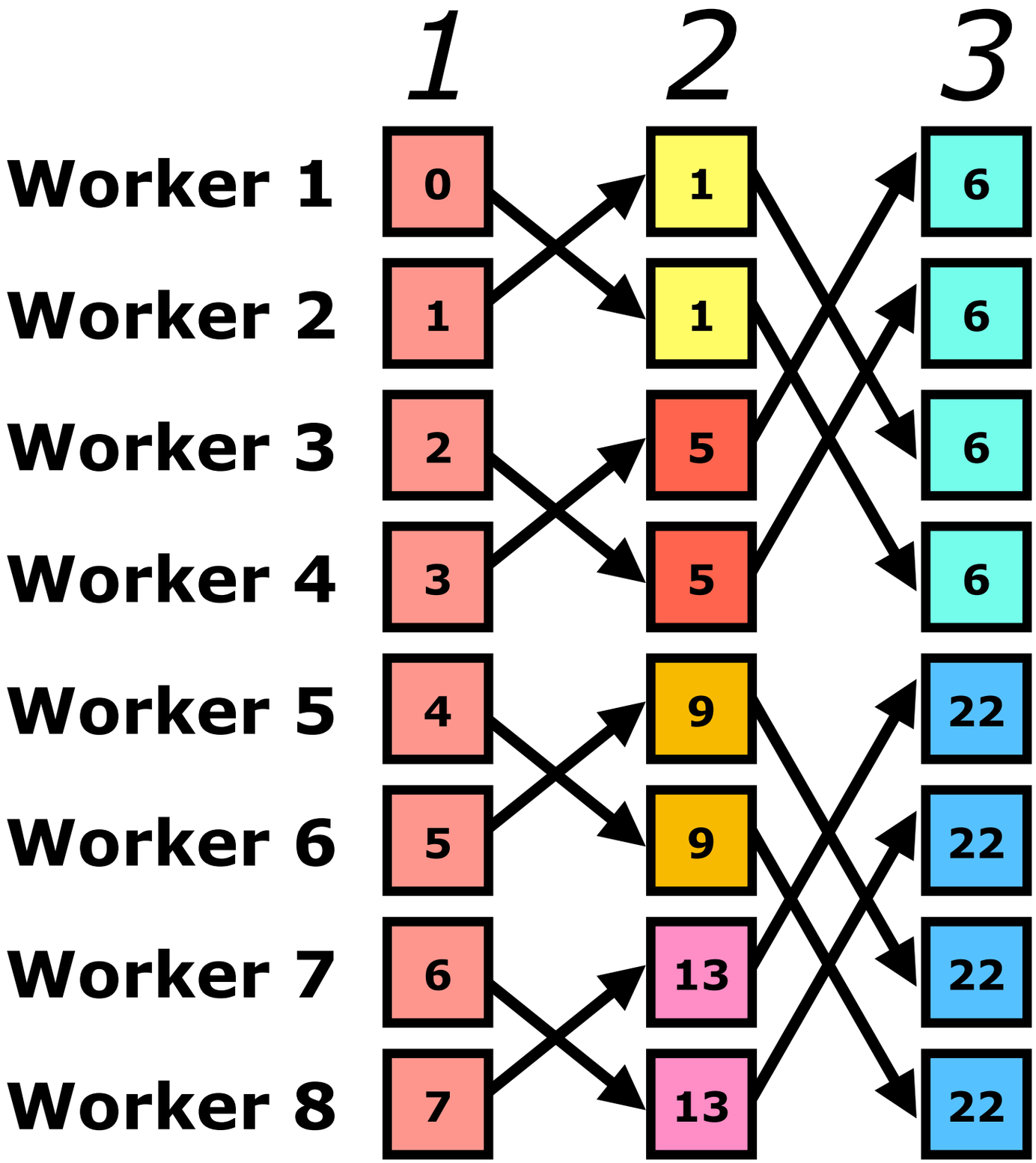
Worker 7 6

Worker 8 7



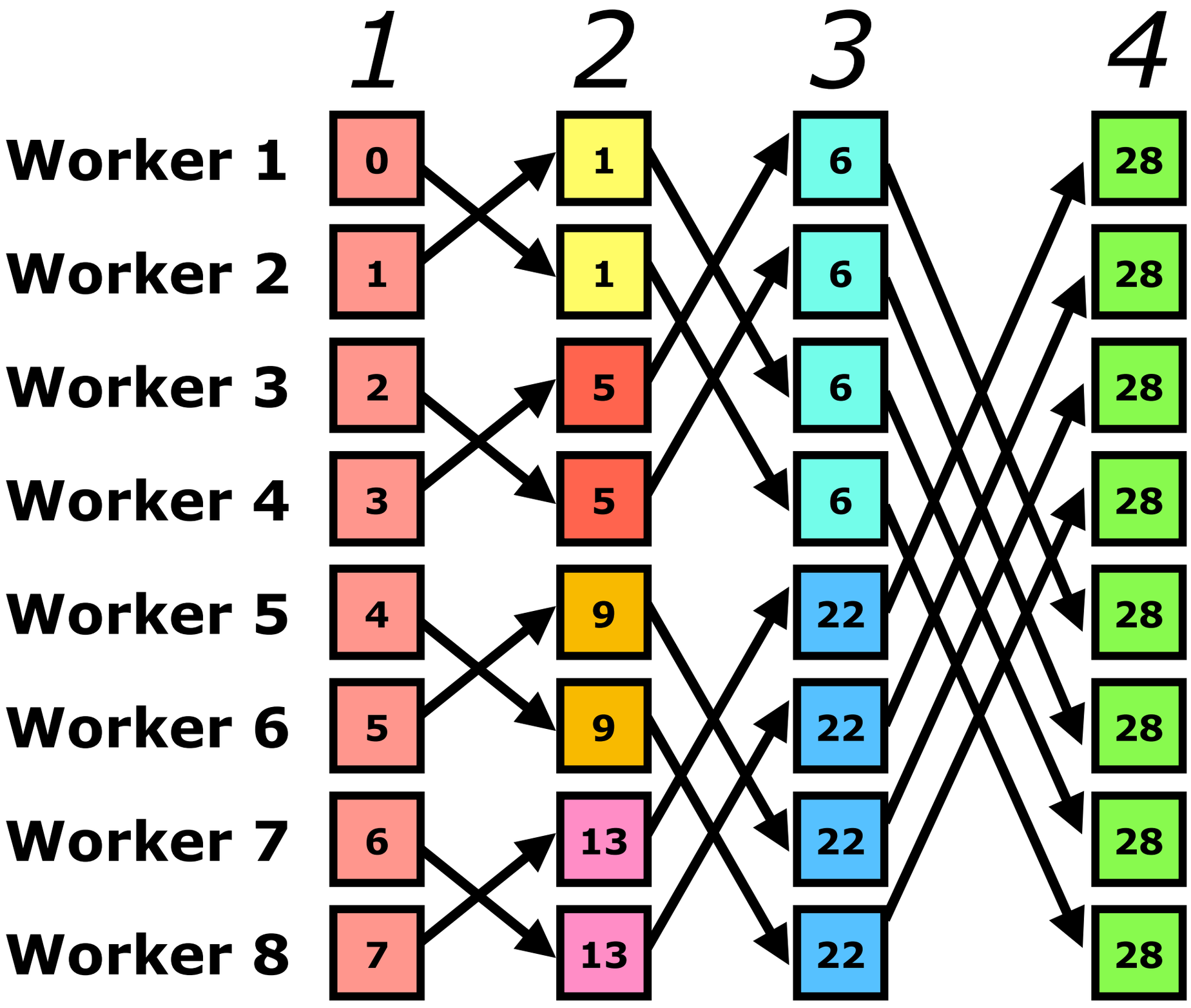
VALOHAI

How Butterfly Allreduce Works



VALOHAI

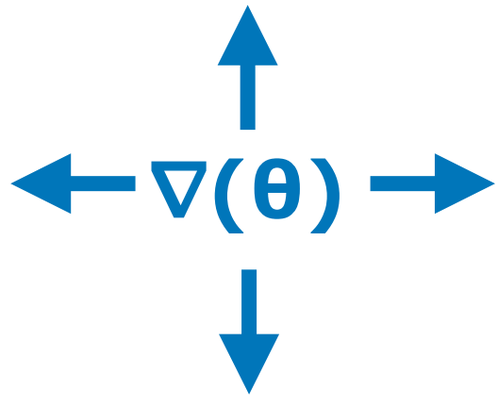
How Butterfly Allreduce Works

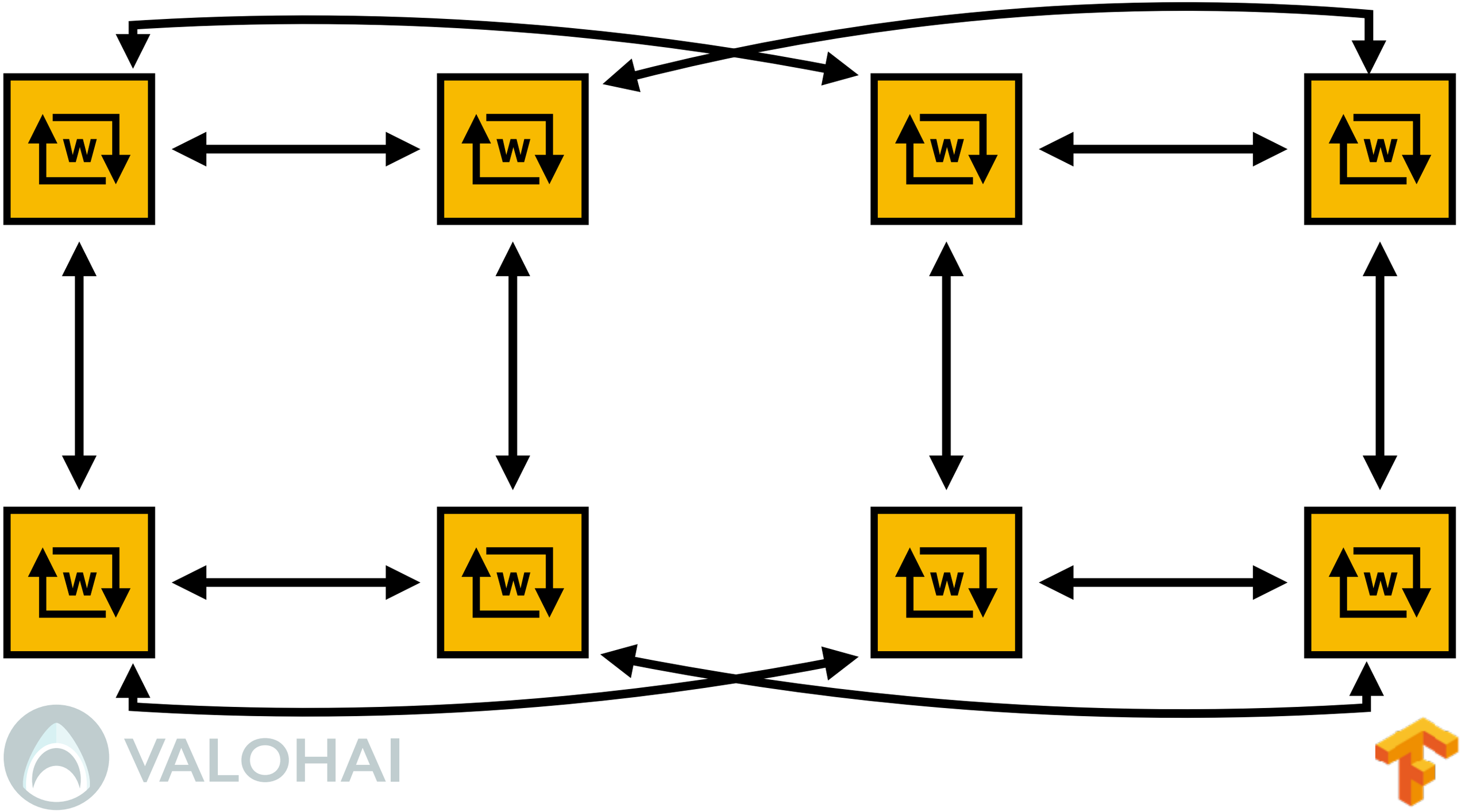


VALOHAI

How Butterfly Allreduce Works

Butterfly Allreduce



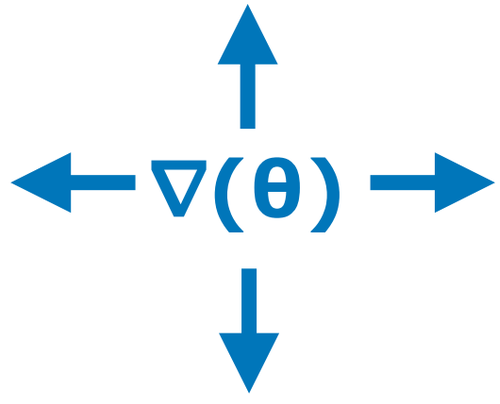


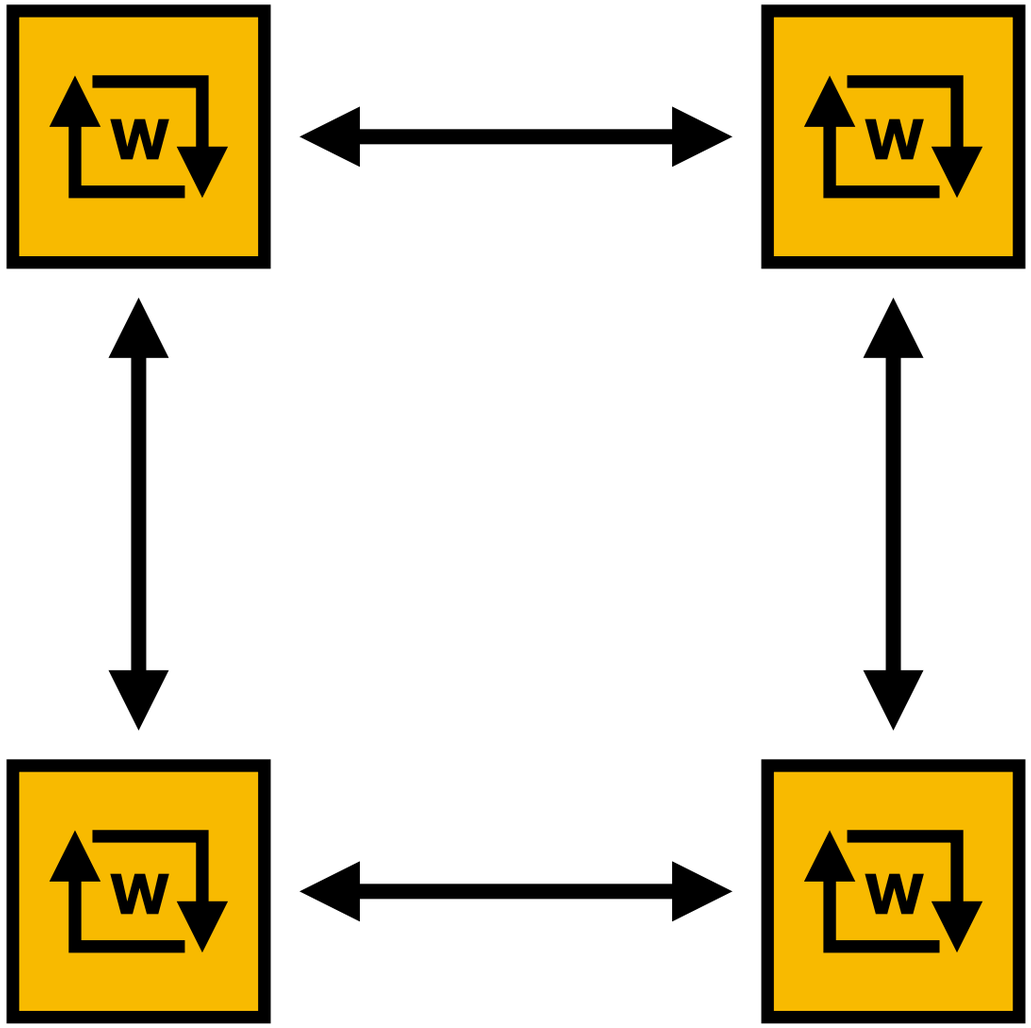
VALOHAI

Ring Allreduce

VALOHAI

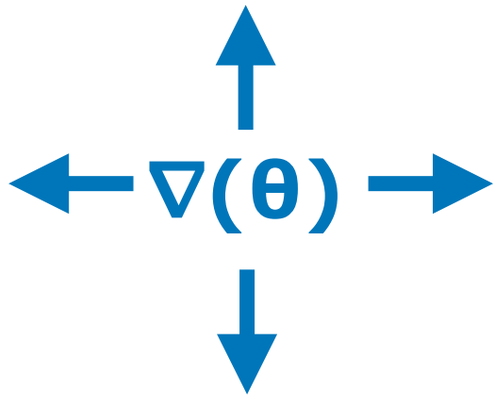
Ring Allreduce

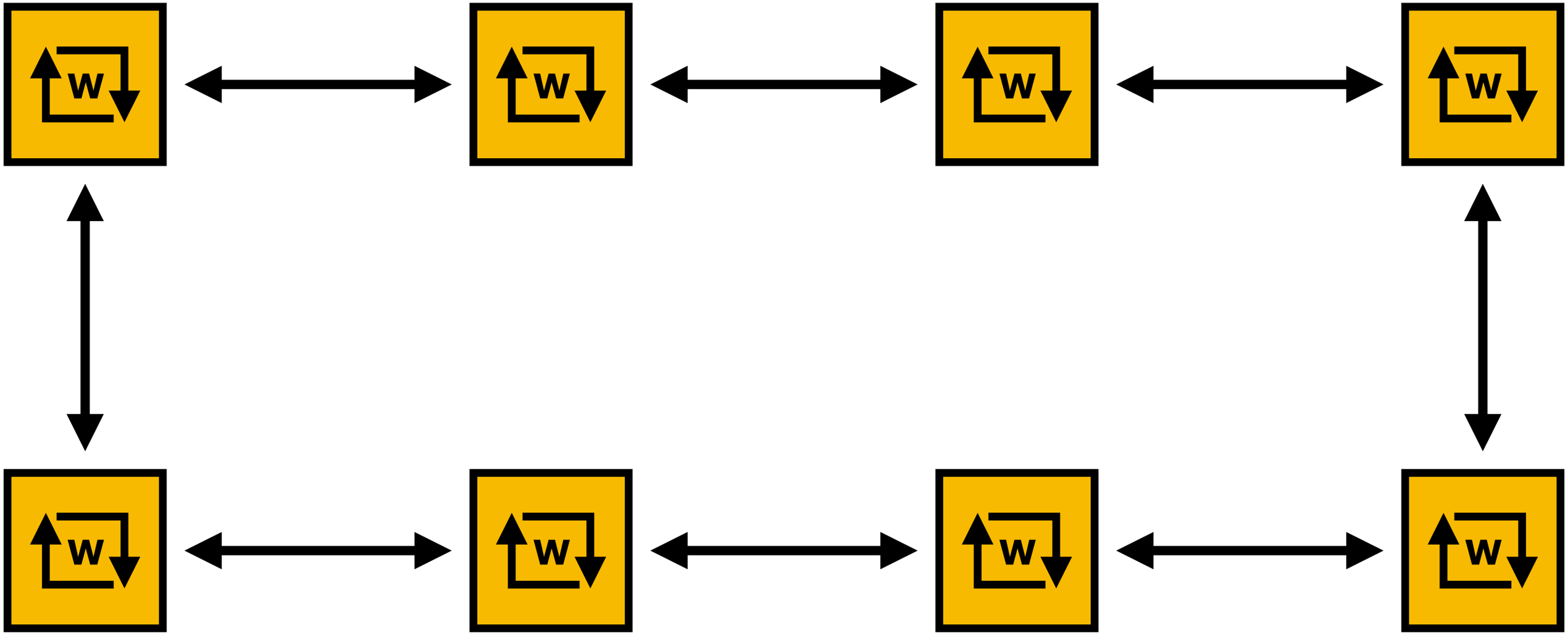




VALOHAI

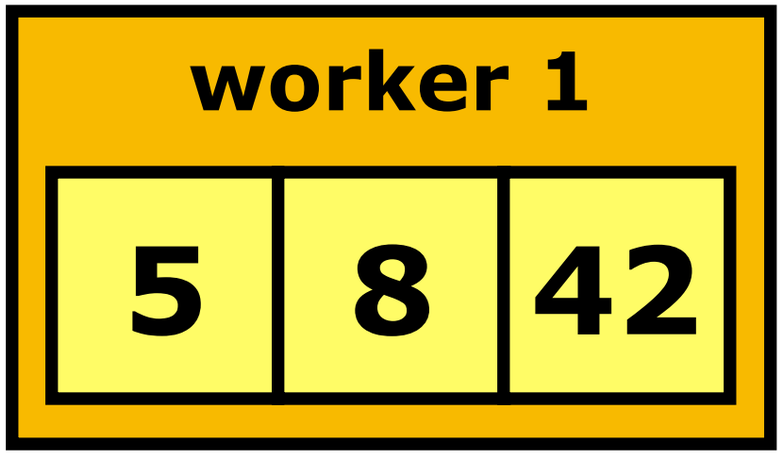
Ring Allreduce

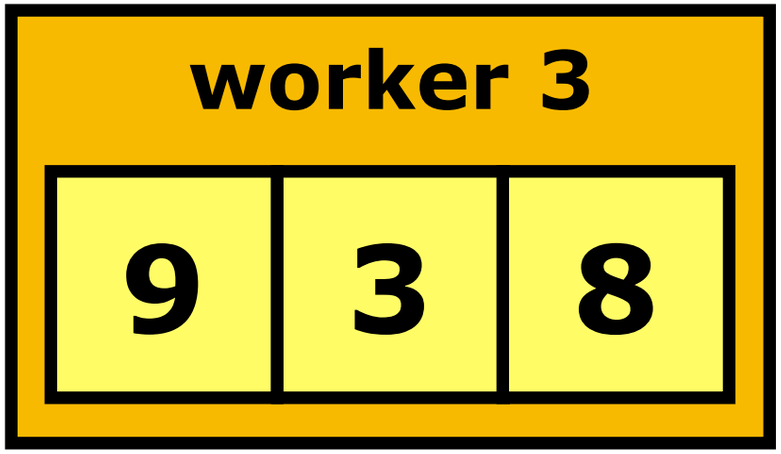


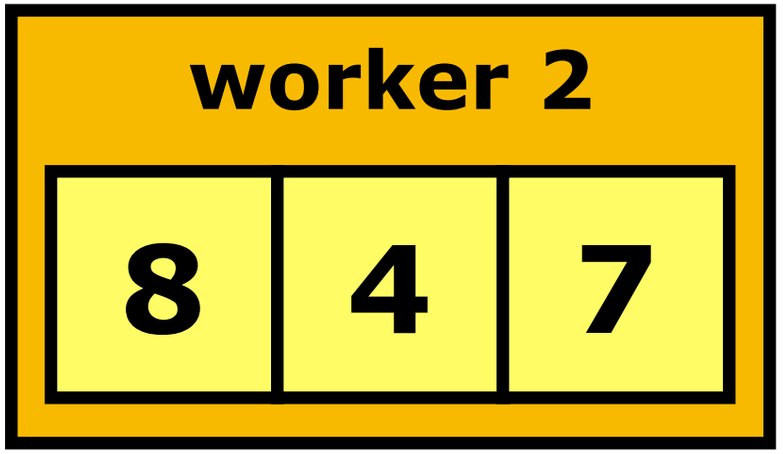


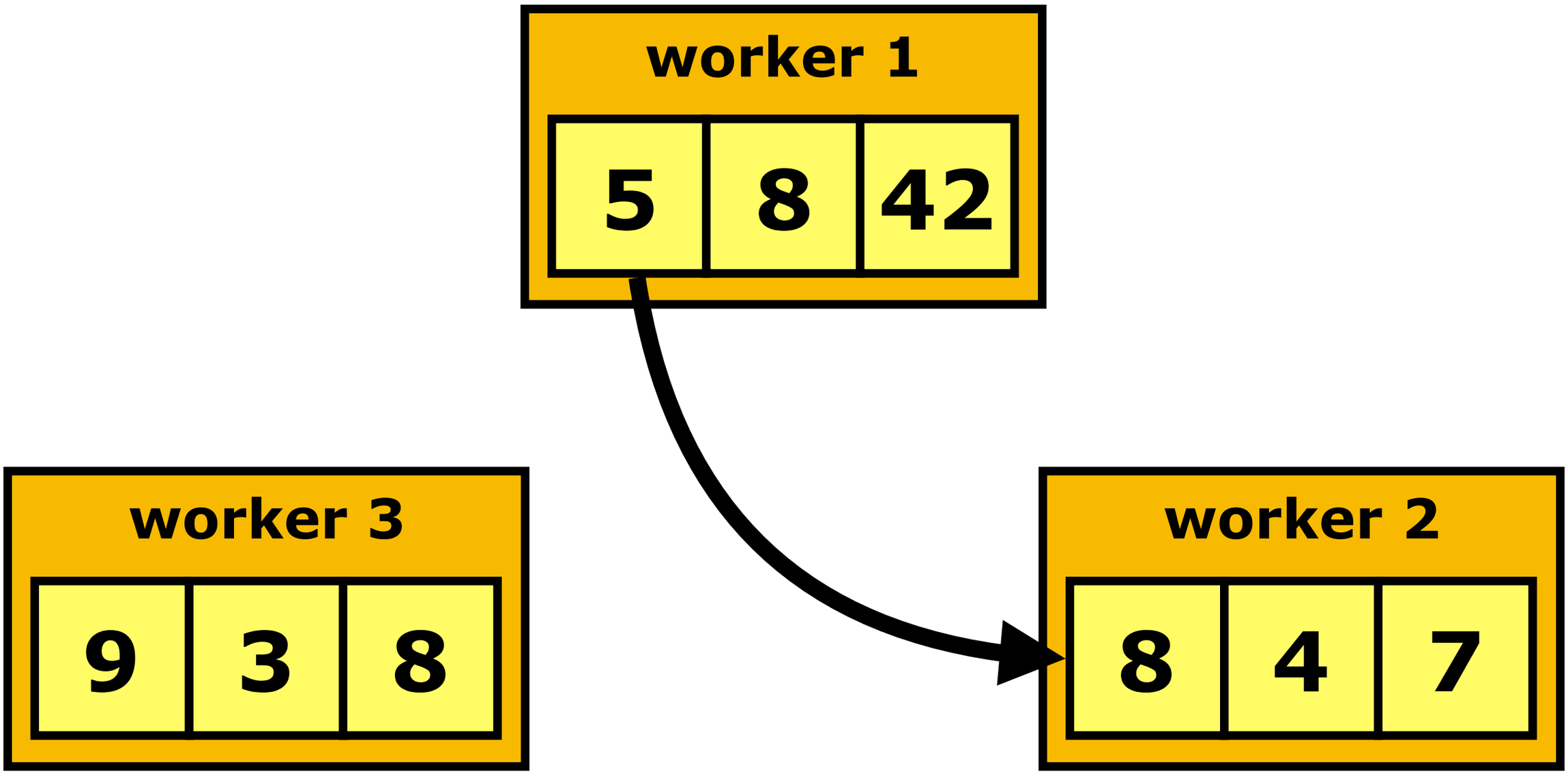
VALOHAI

HOROVOD


worker 1
5
8
42

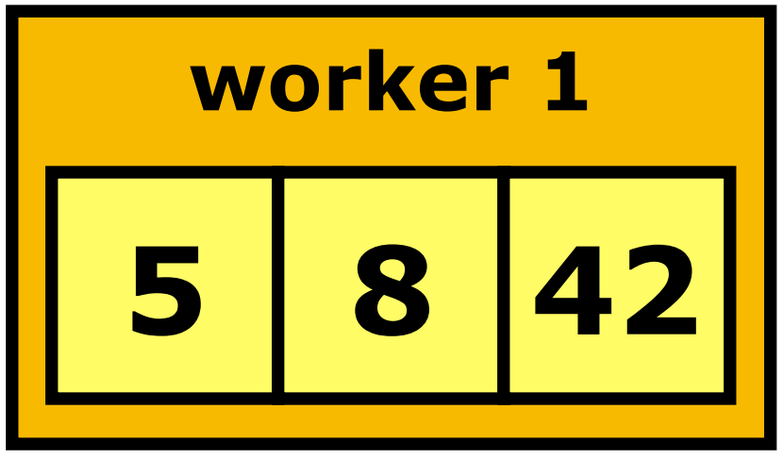

worker 3
9
3
8

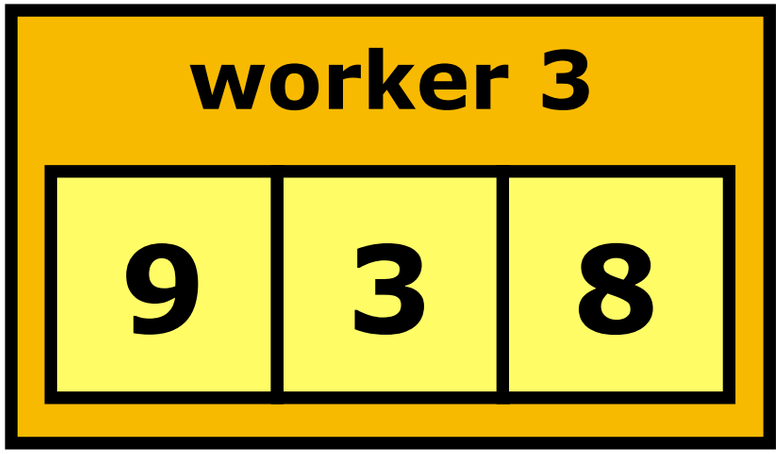

worker 2
8
4
7

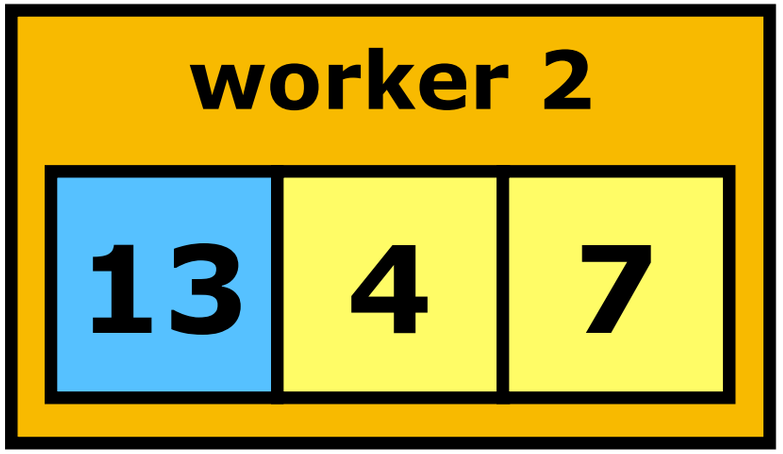


VALOHAI

How Ring Allreduce Works

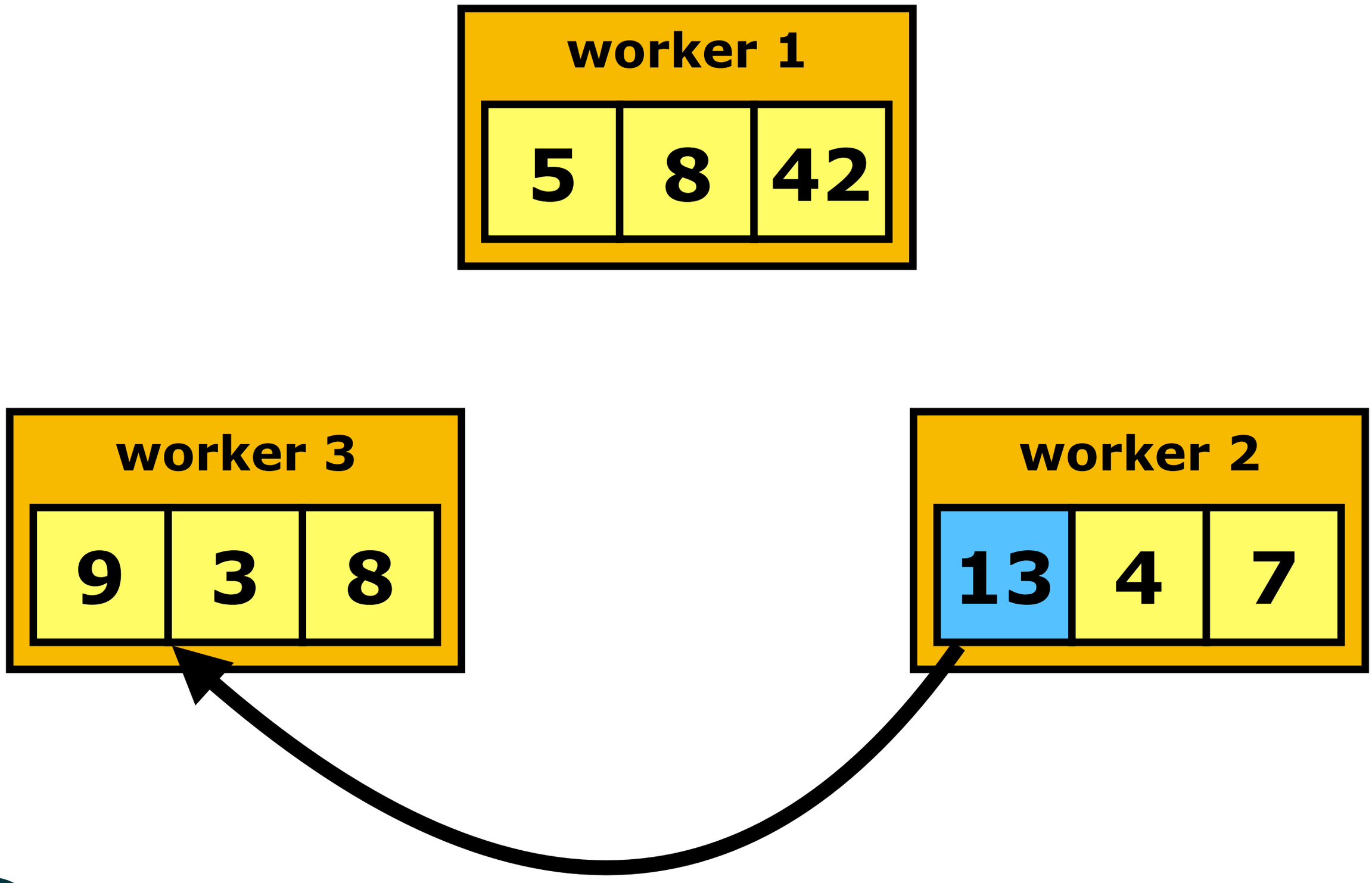






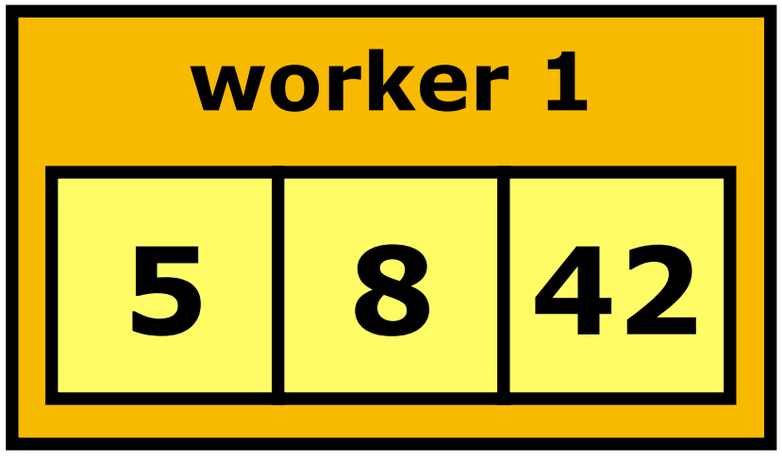
VALOHAI

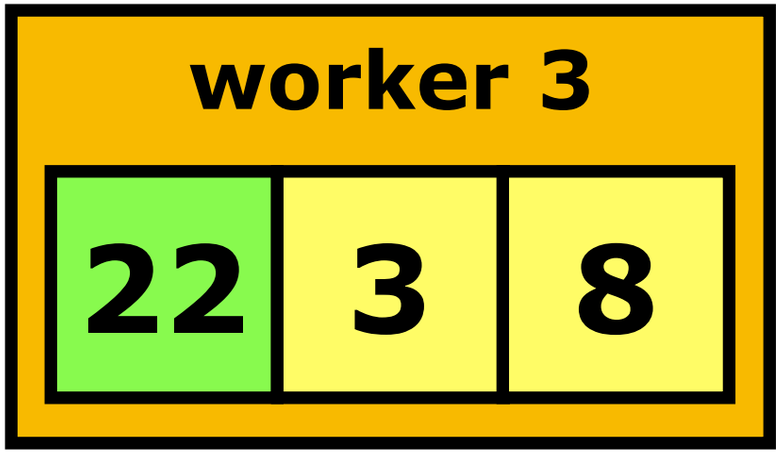
How Ring Allreduce Works

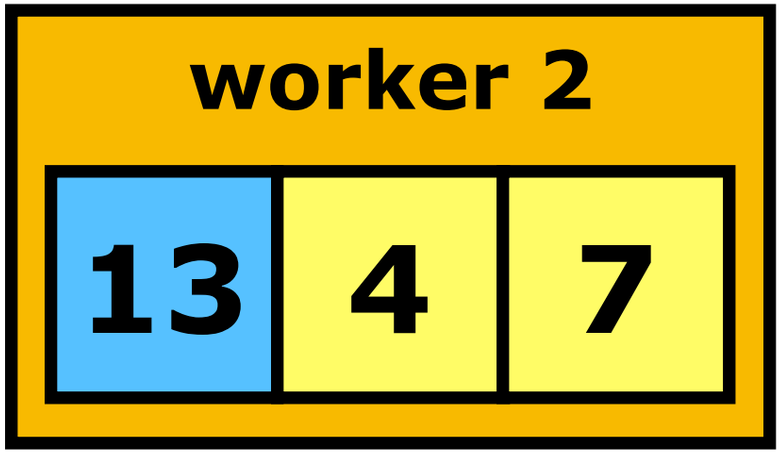


VALOHAI

How Ring Allreduce Works

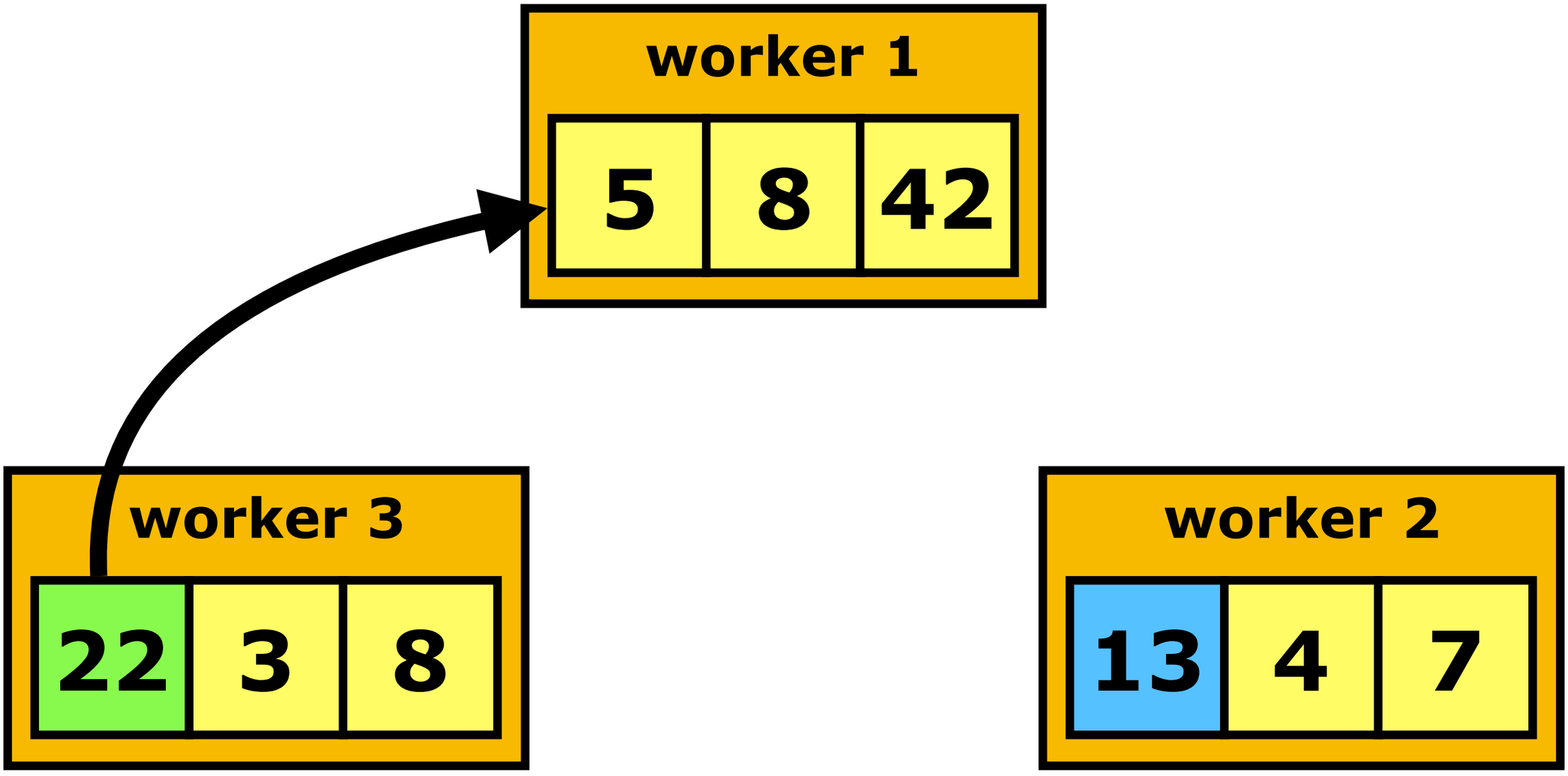






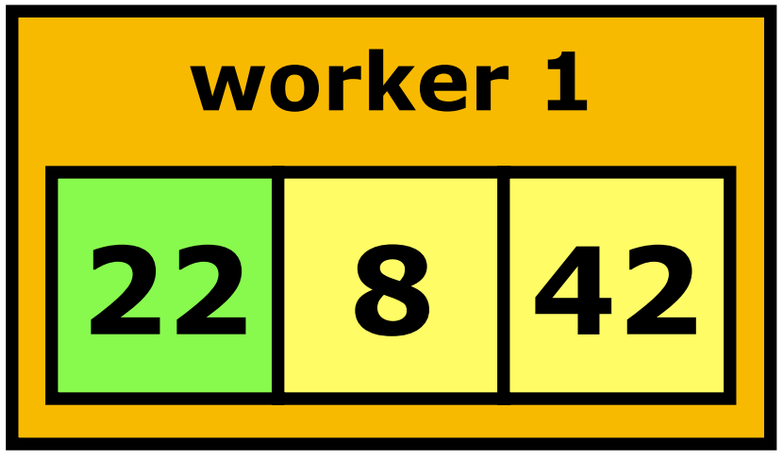
VALOHAI

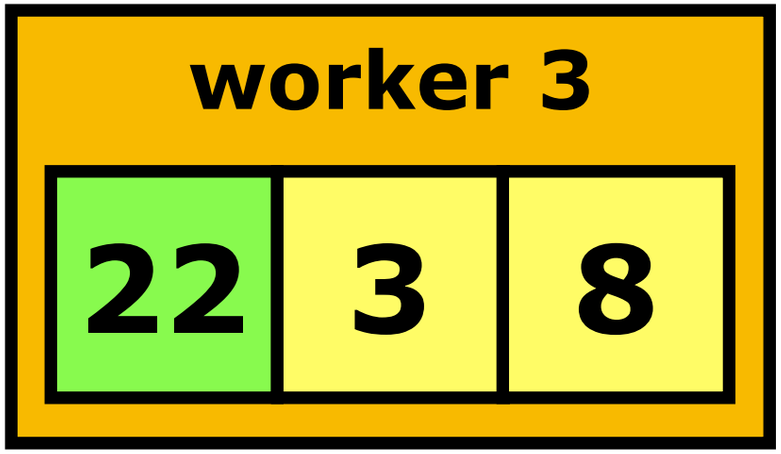
How Ring Allreduce Works

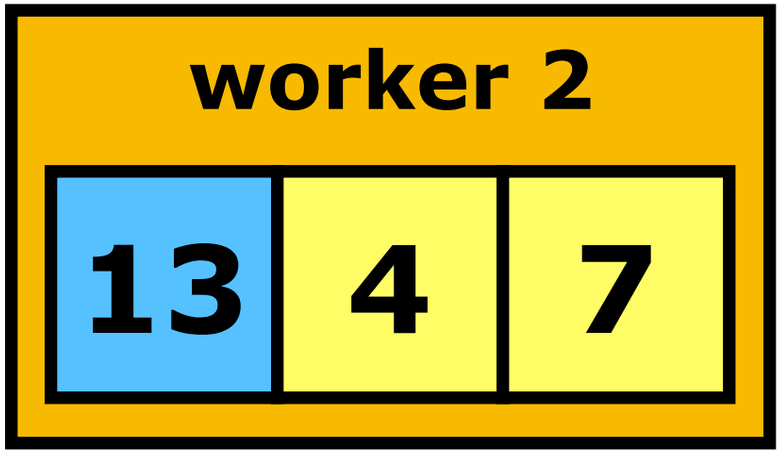


VALOHAI

How Ring Allreduce Works

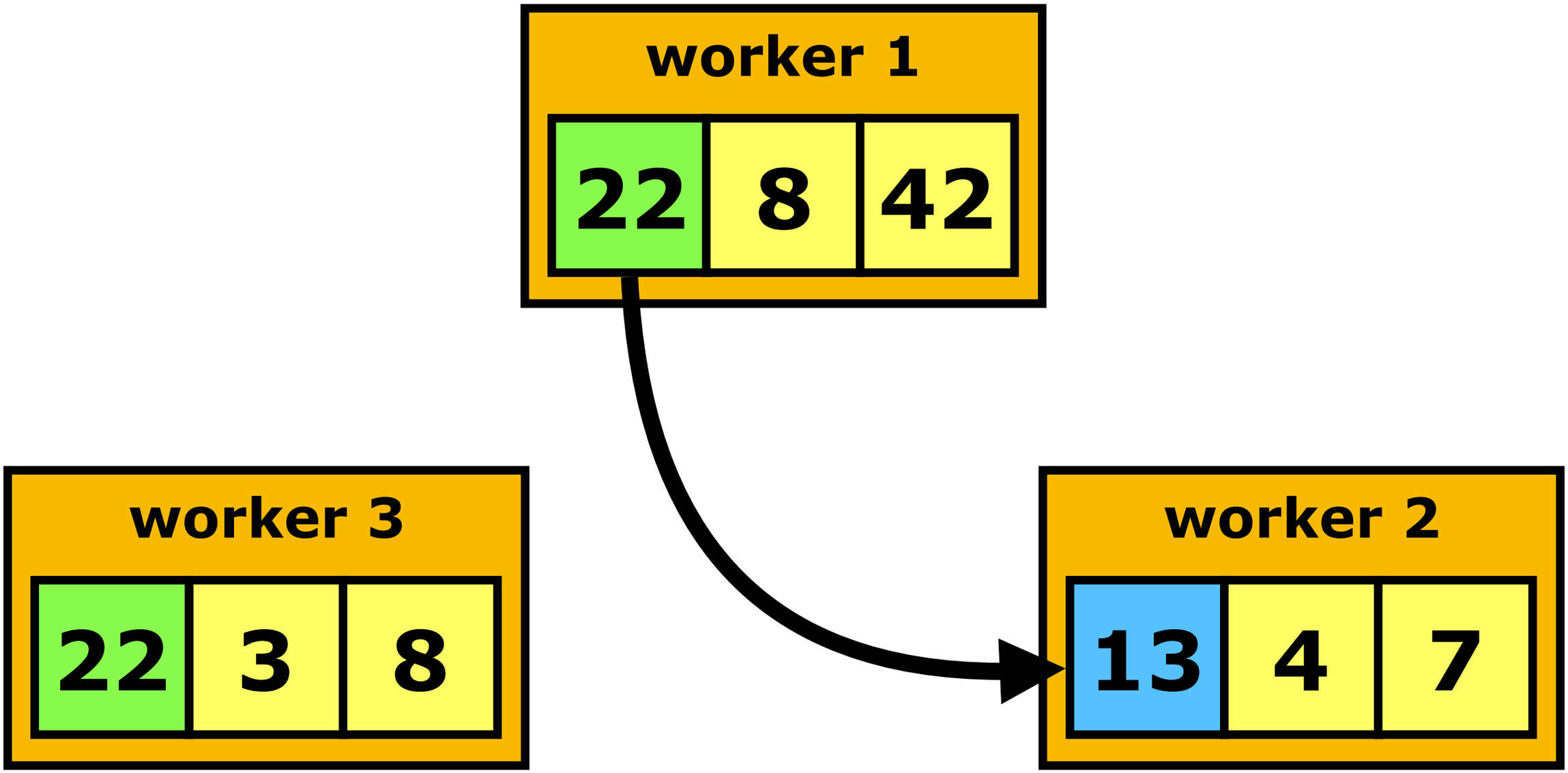






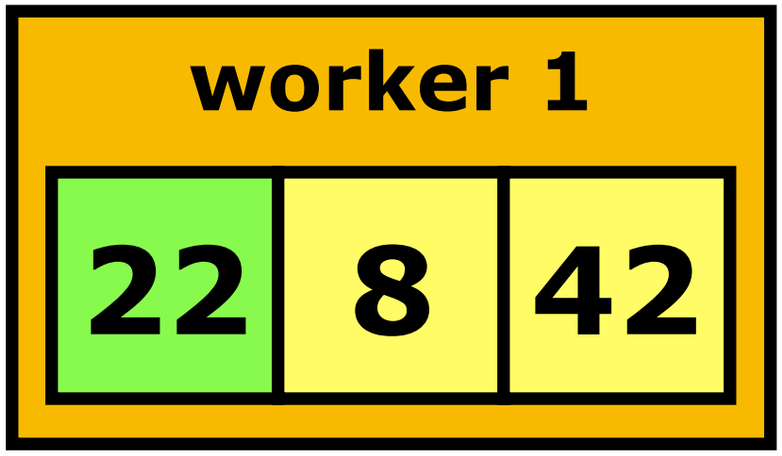
VALOHAI

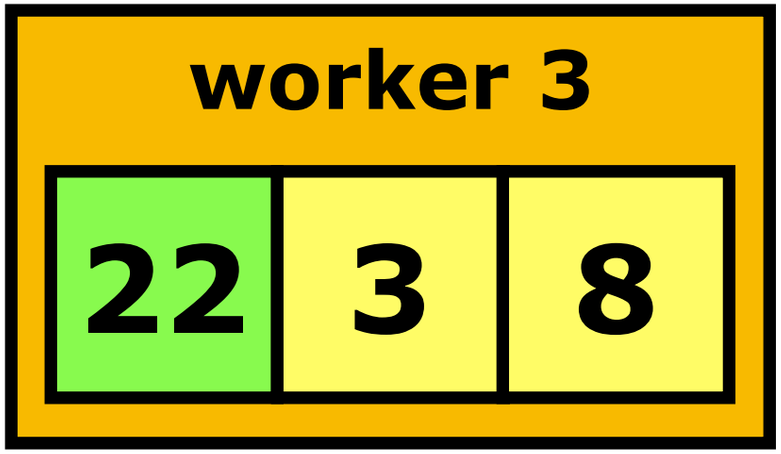
How Ring Allreduce Works

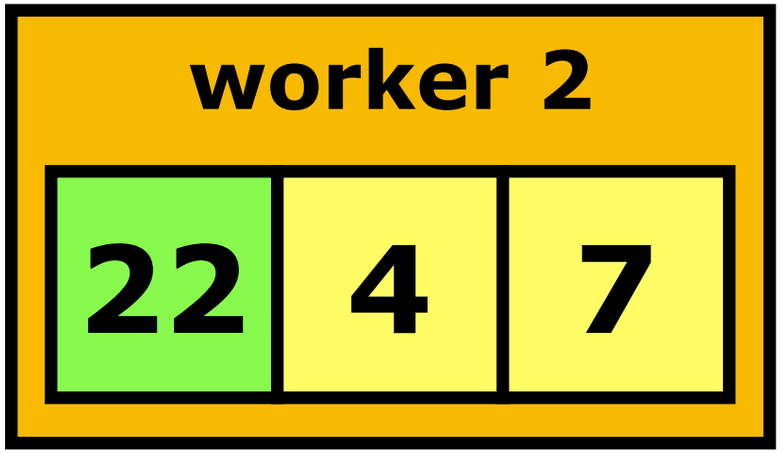


VALOHAI

How Ring Allreduce Works

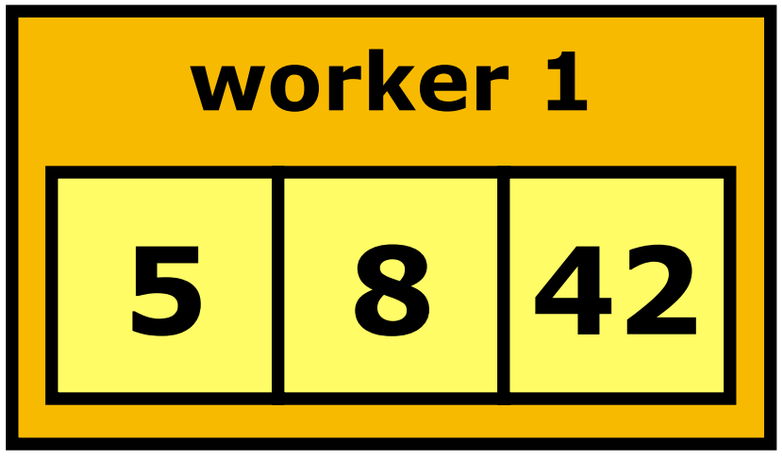


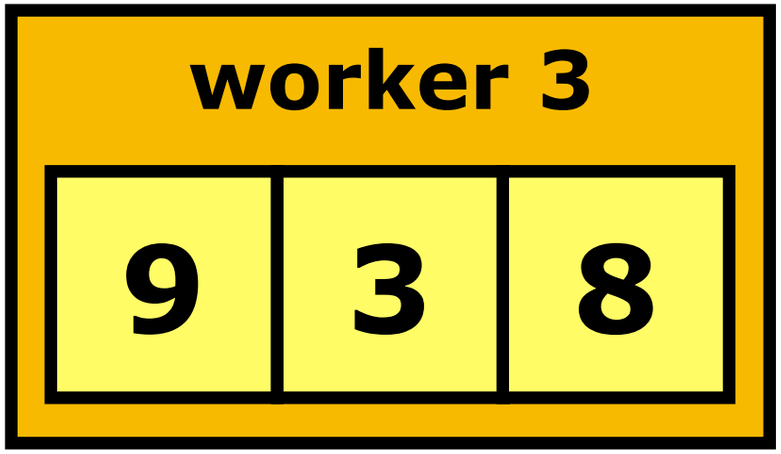


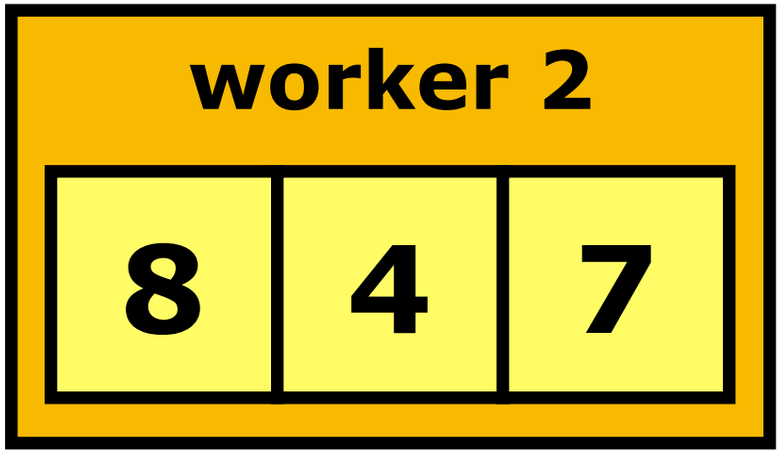


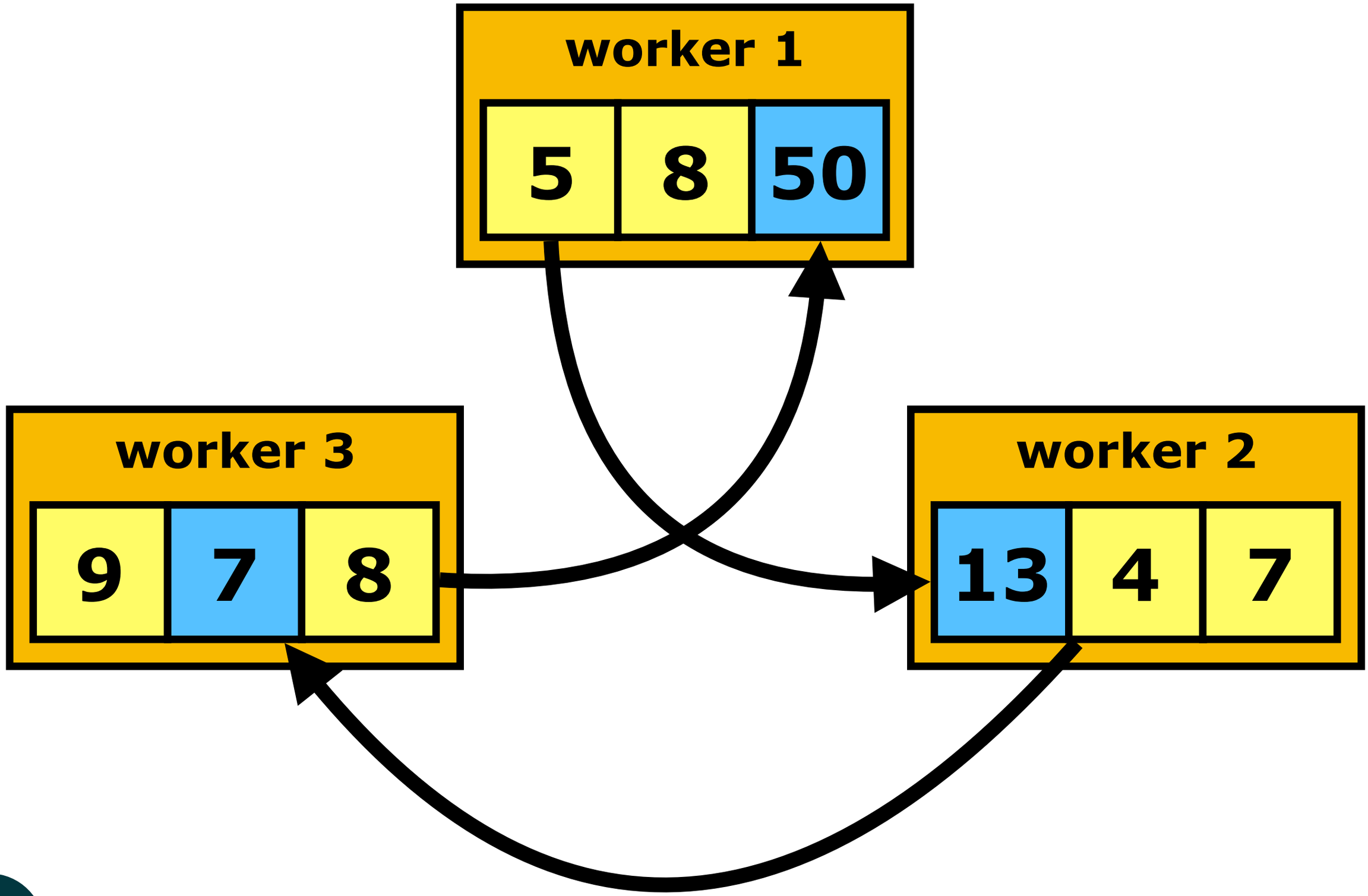
VALOHAI

How Ring Allreduce Works


worker 1
5
8
42

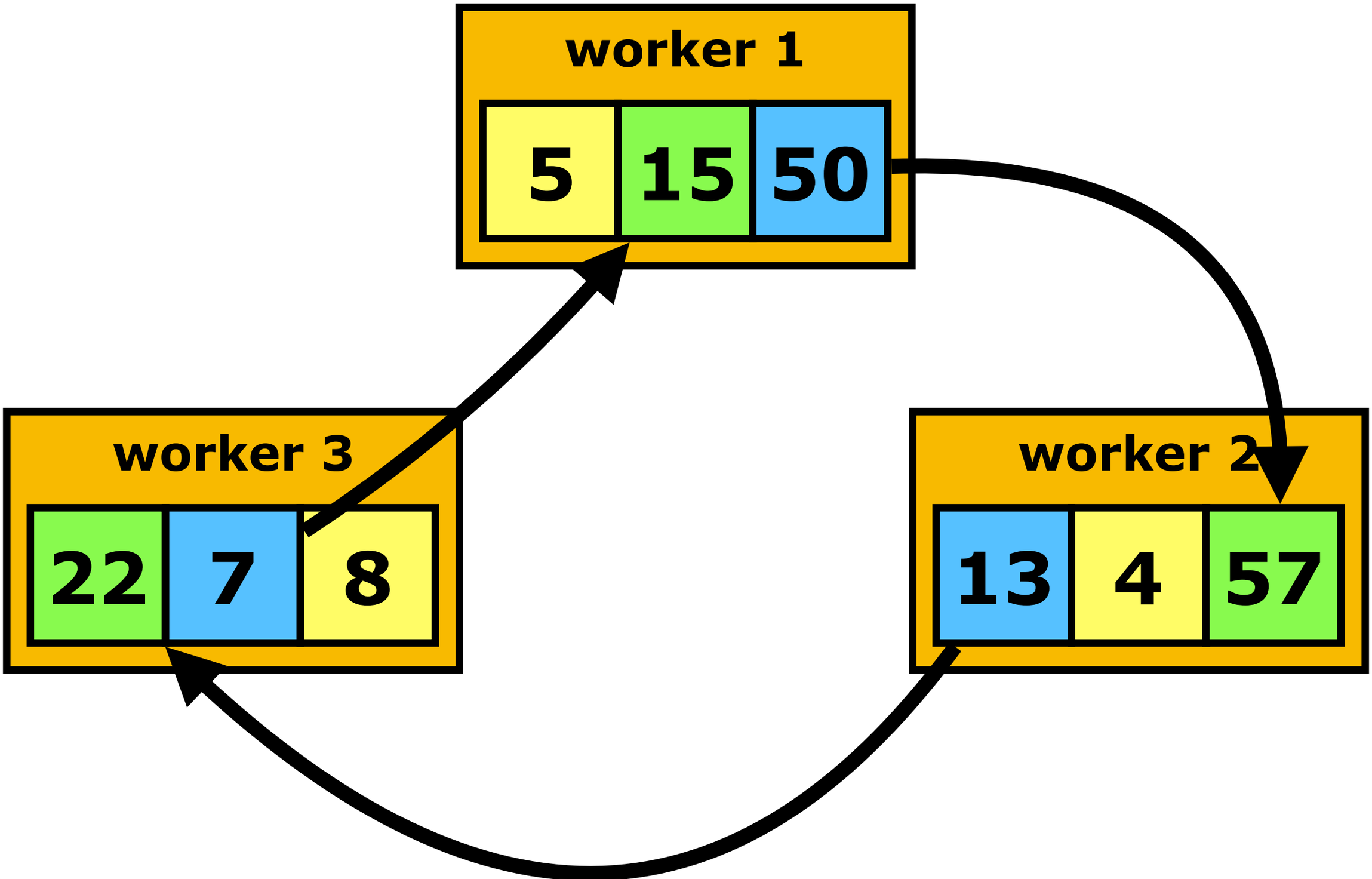

worker 3
9
3
8


worker 2
8
4
7



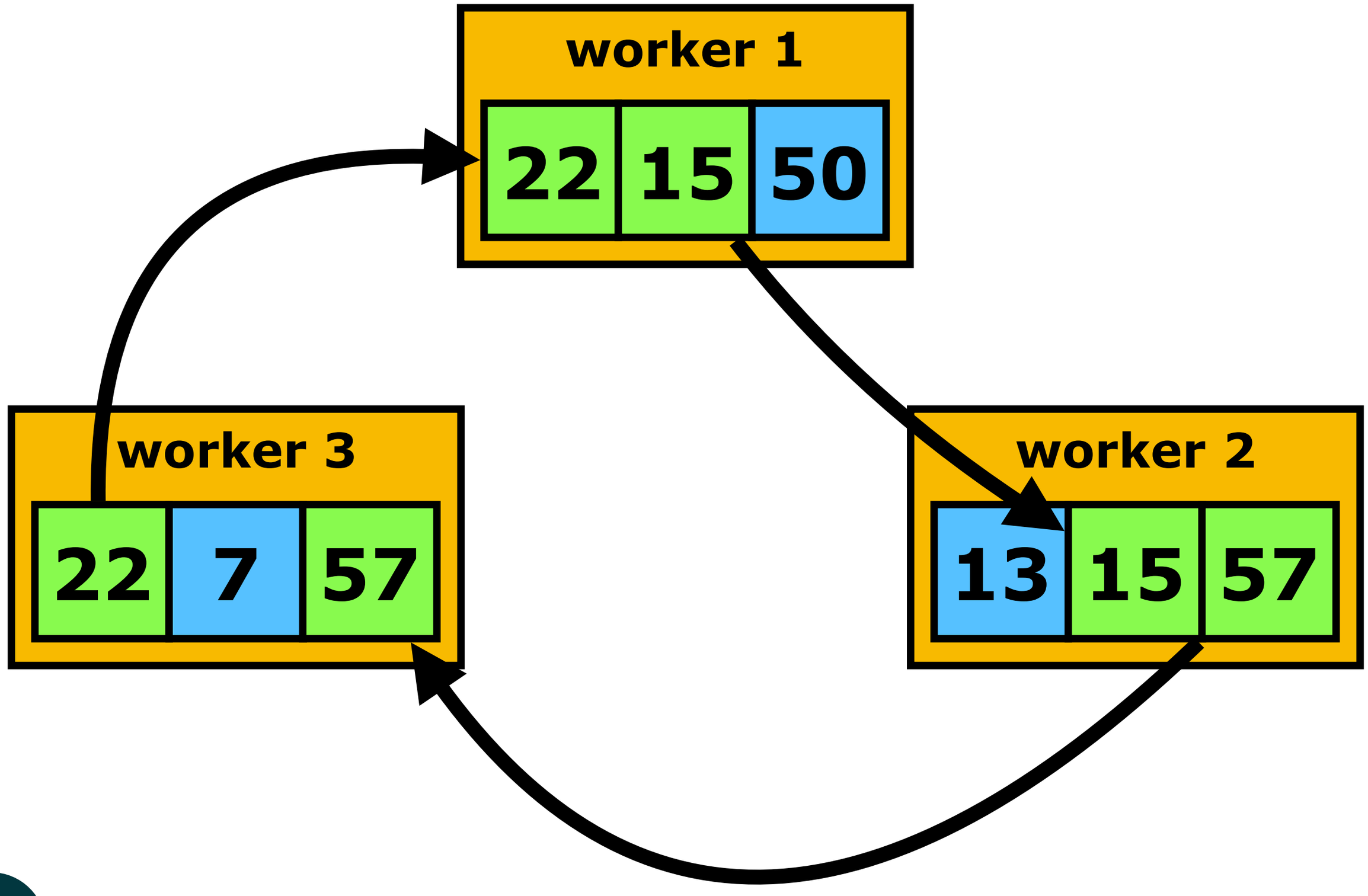
VALOHAI

How Ring Allreduce Works



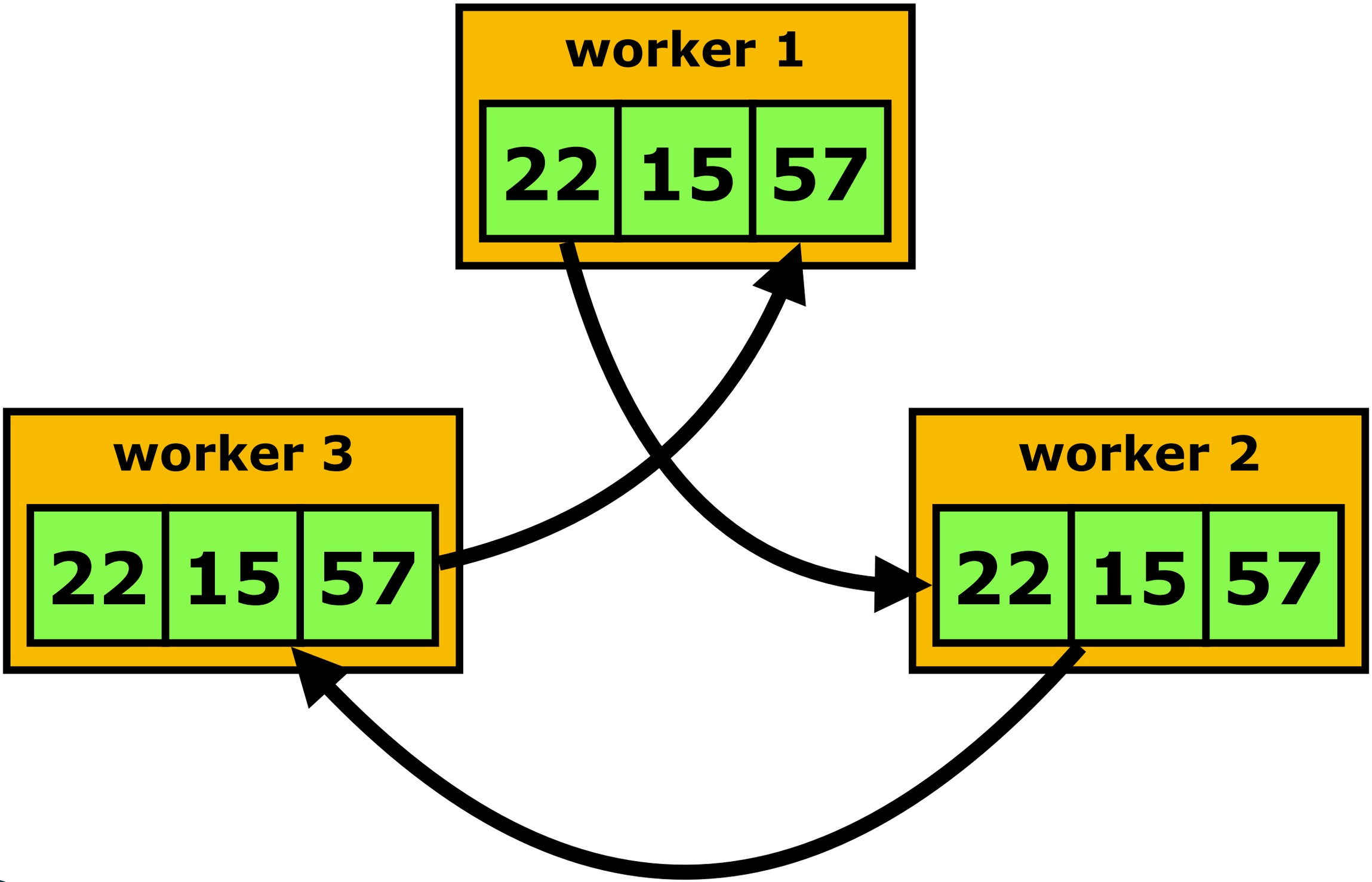
VALOHAI

How Ring Allreduce Works



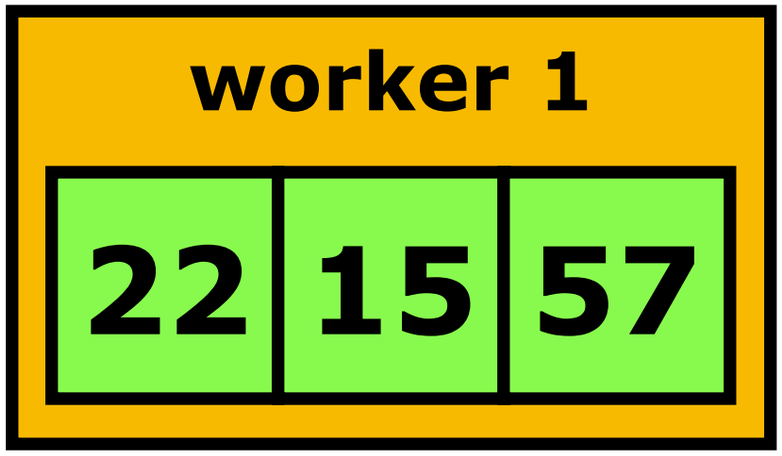
VALOHAI

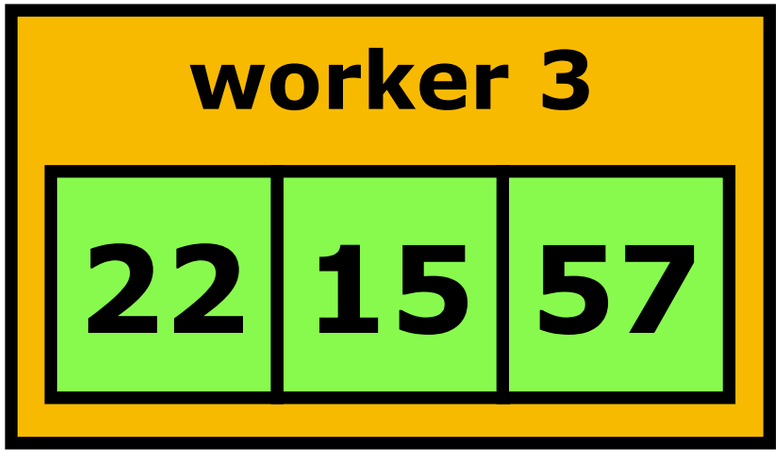
How Ring Allreduce Works

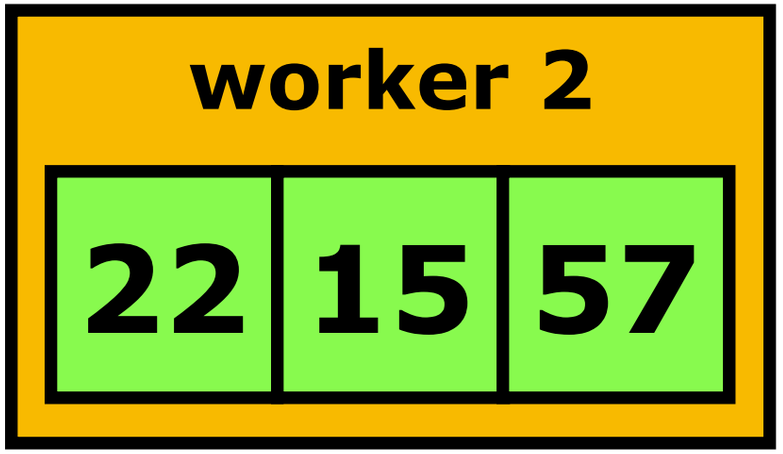


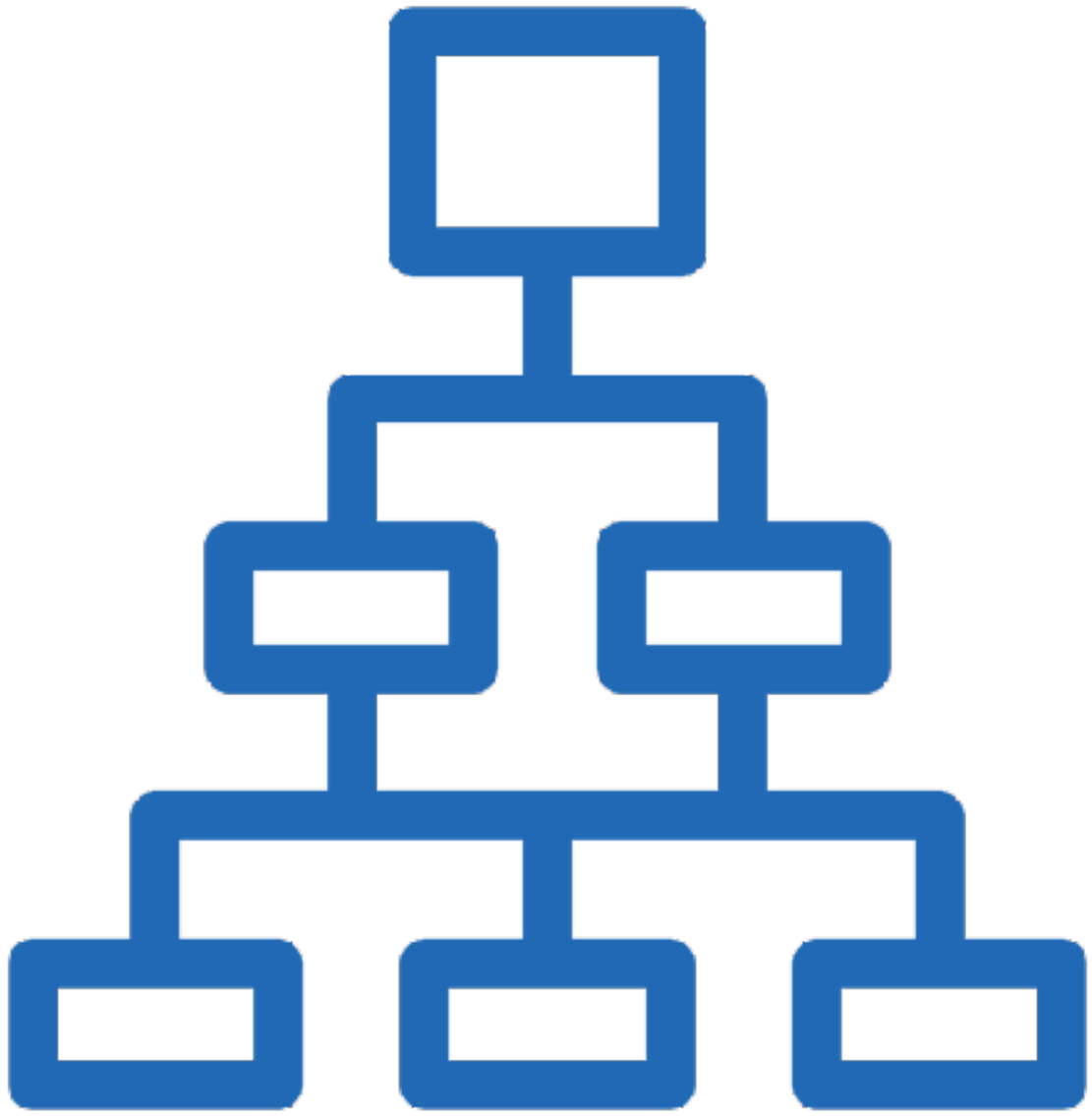
VALOHAI

How Ring Allreduce Works


worker 1
22
15
57


worker 3
22
15
57

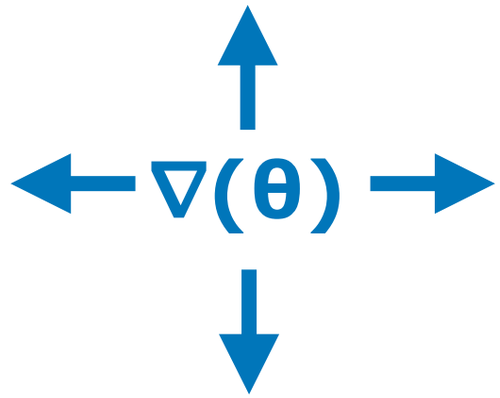

worker 2
22
15
57

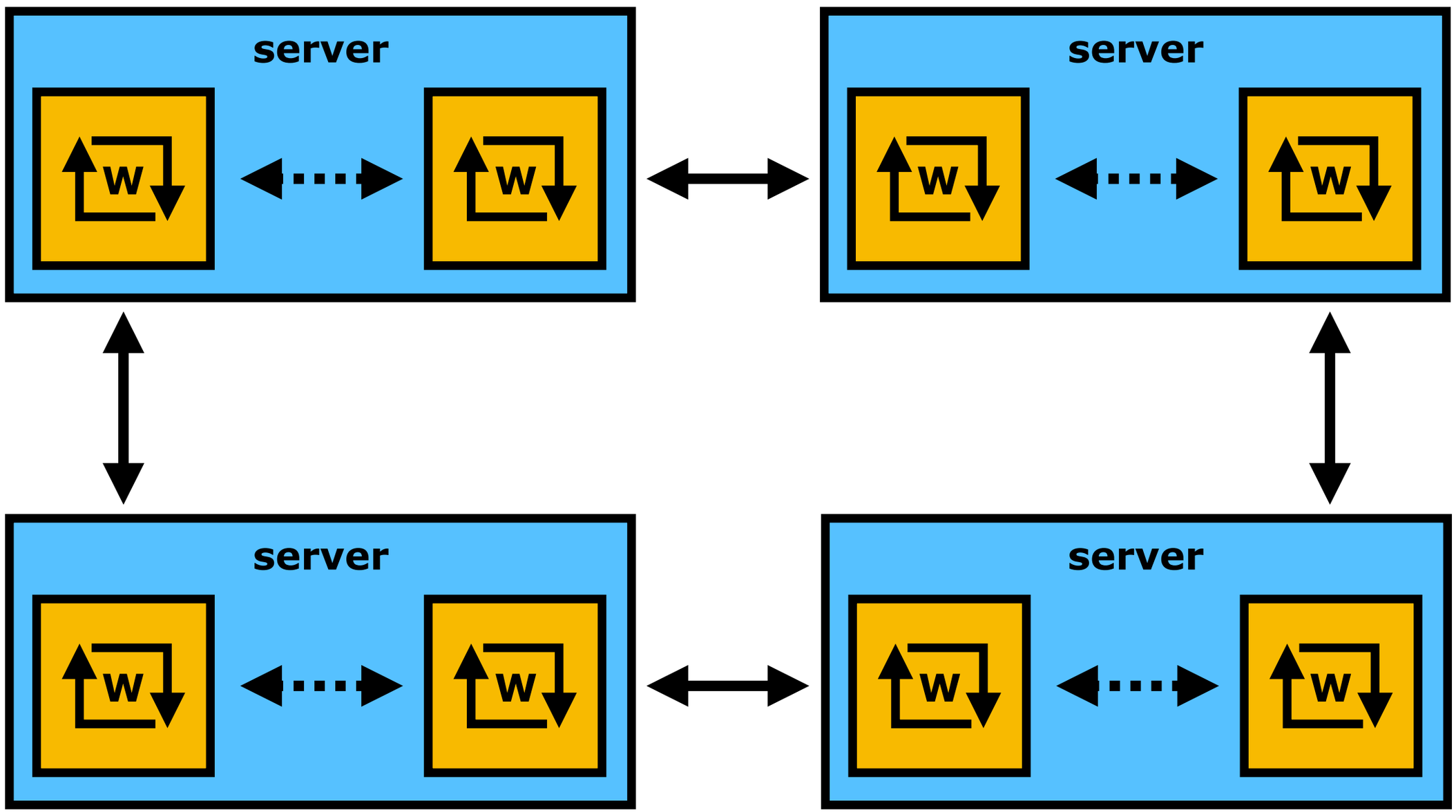


Hierarchical Allreduce

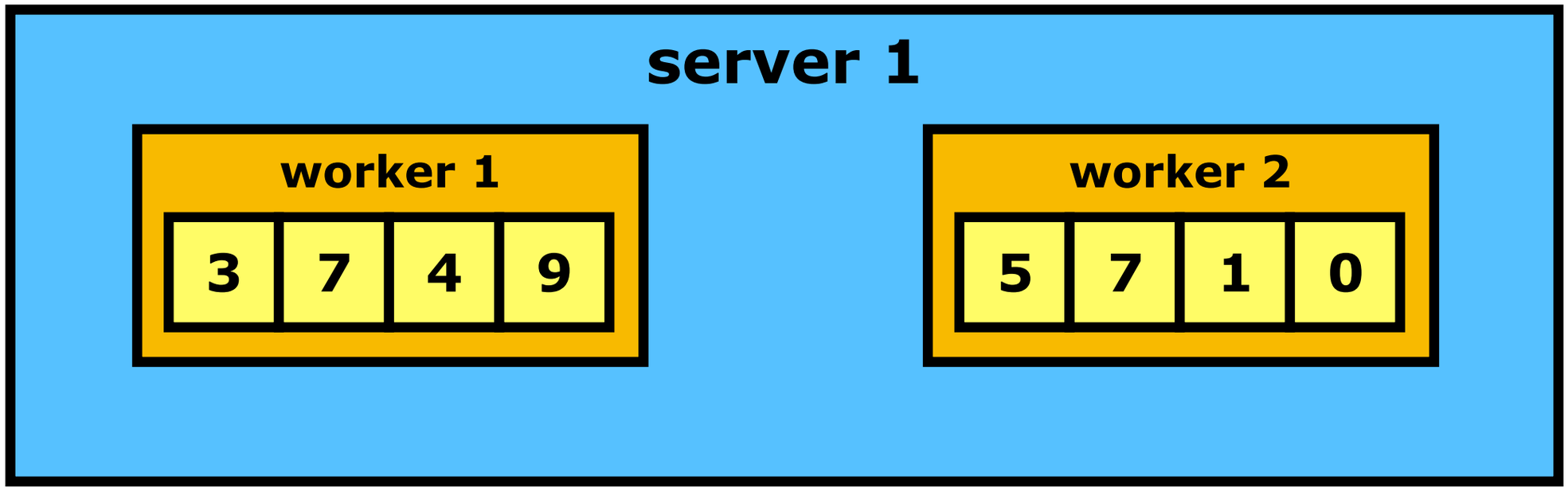
VALOHAI

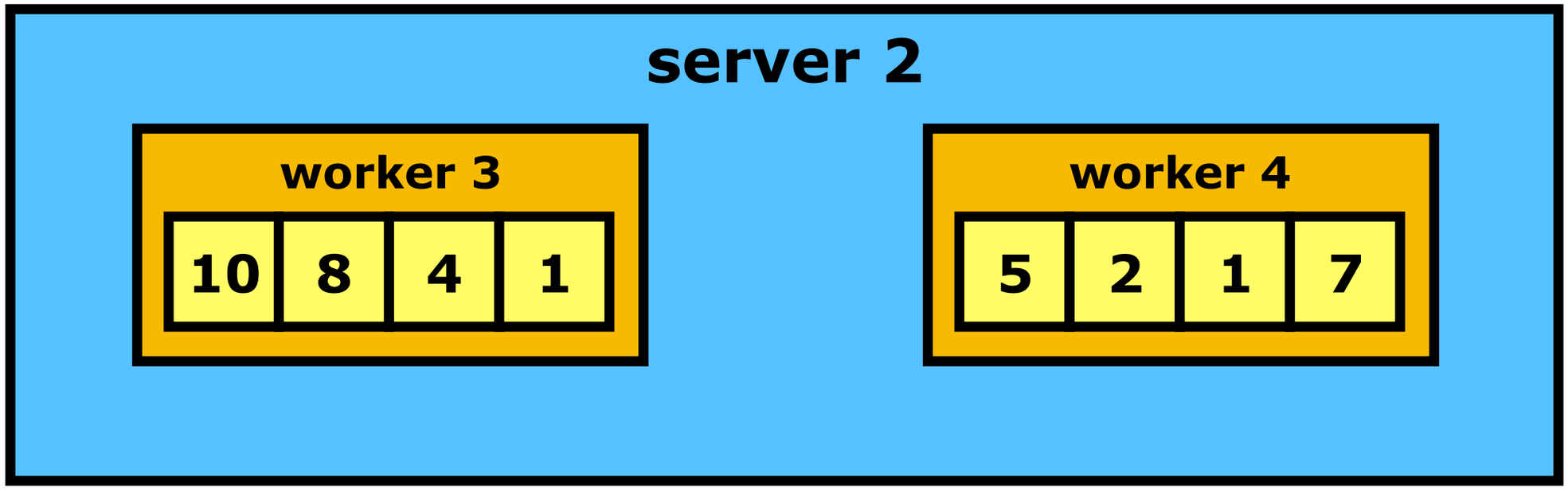
Hierarchical Allreduce

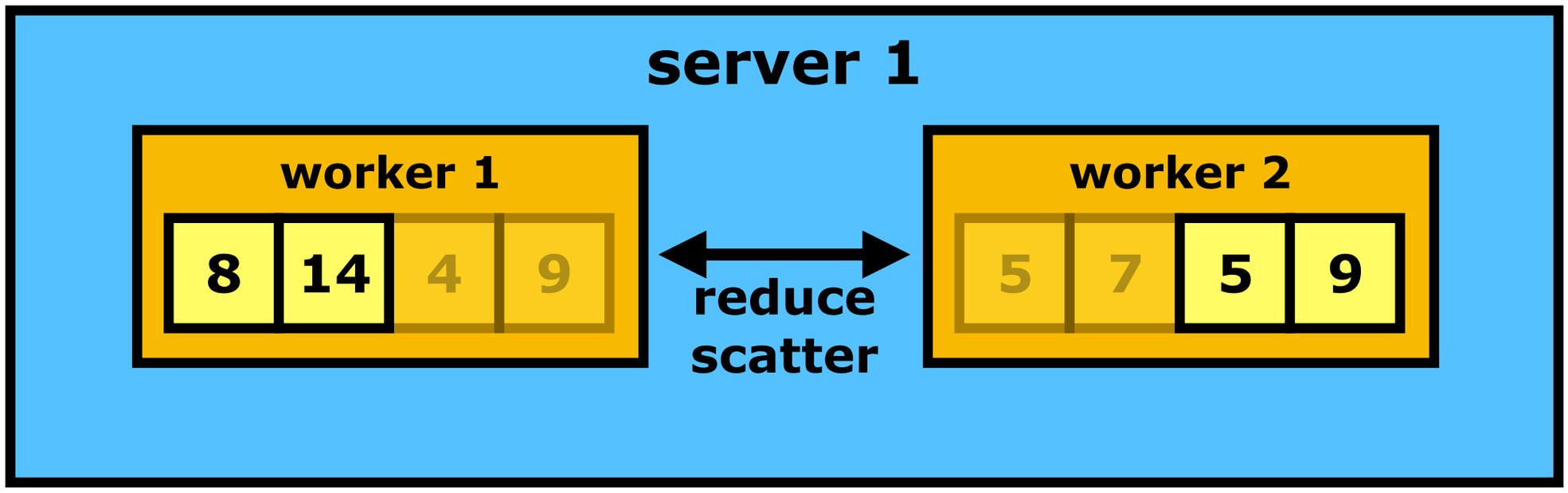


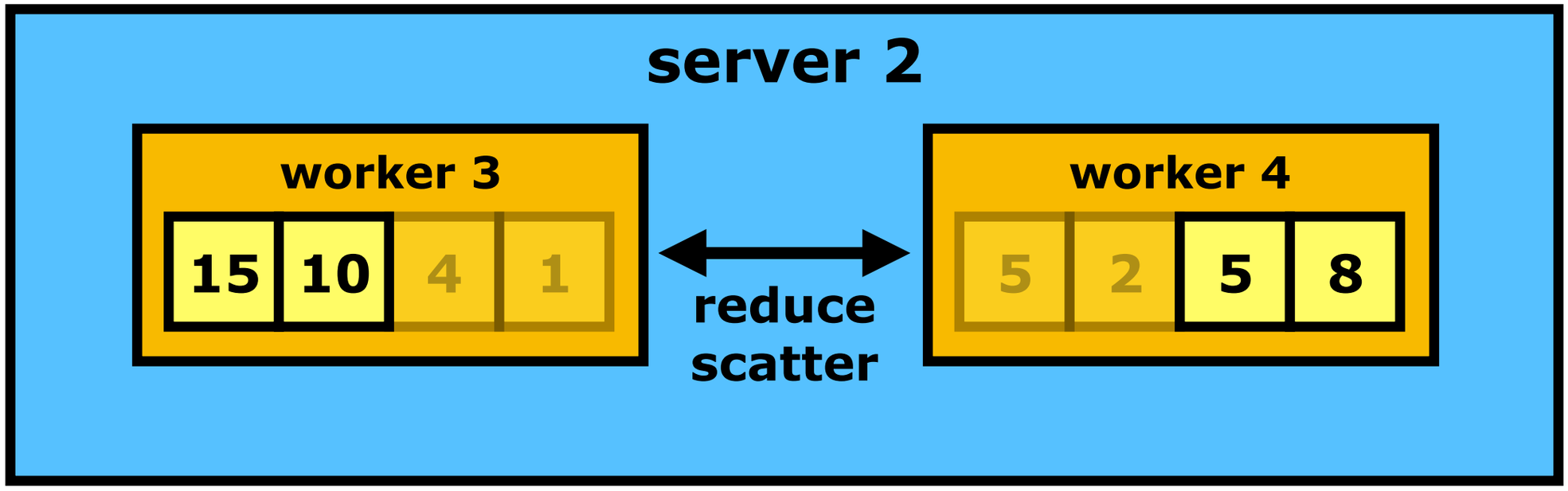


VALOHAI


server 1
worker 1
3
7
4
9
worker 2
5
7
1
0

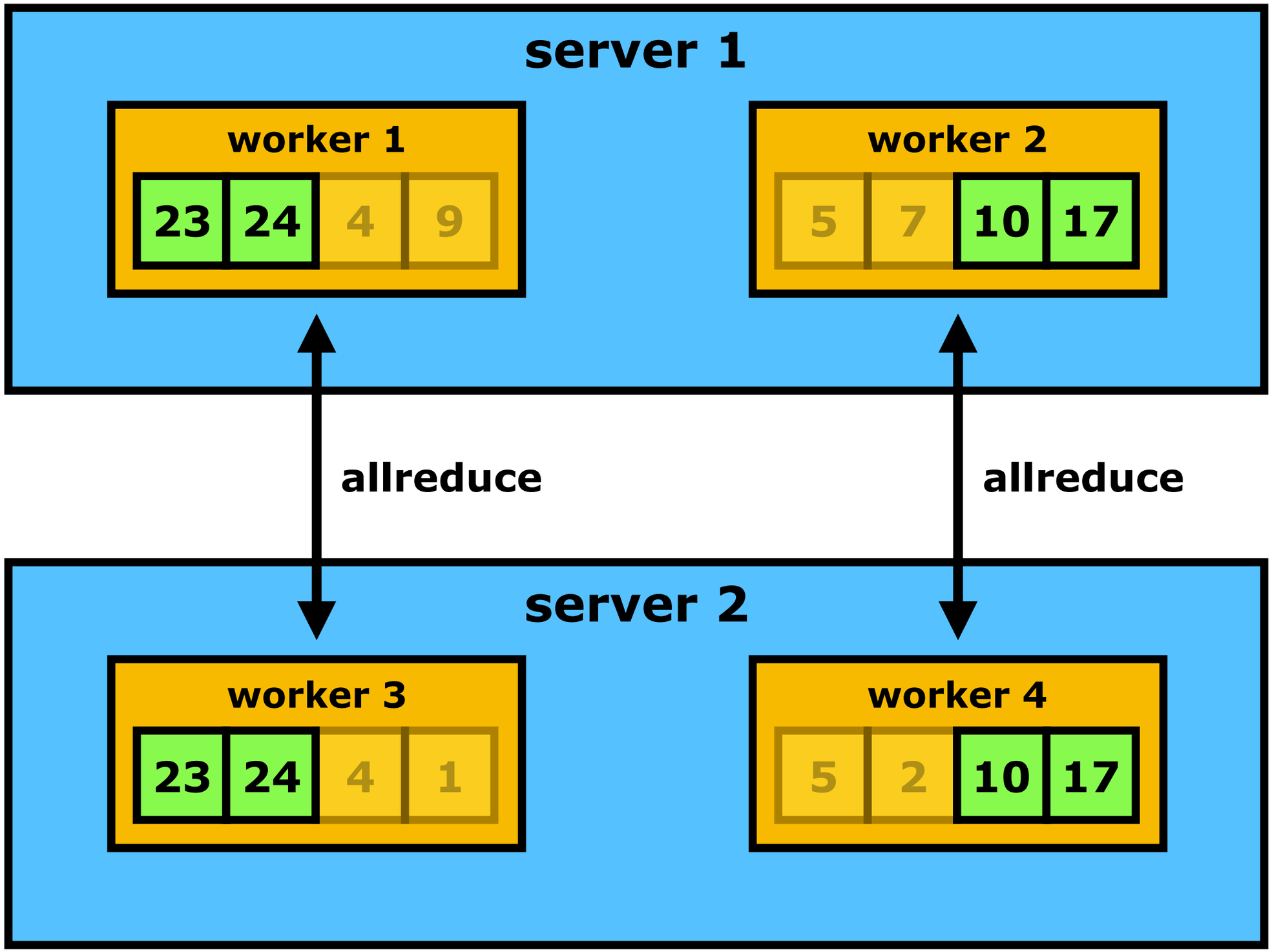

server 2
worker 3
10
8
4
1
worker 4
5
2
1
7





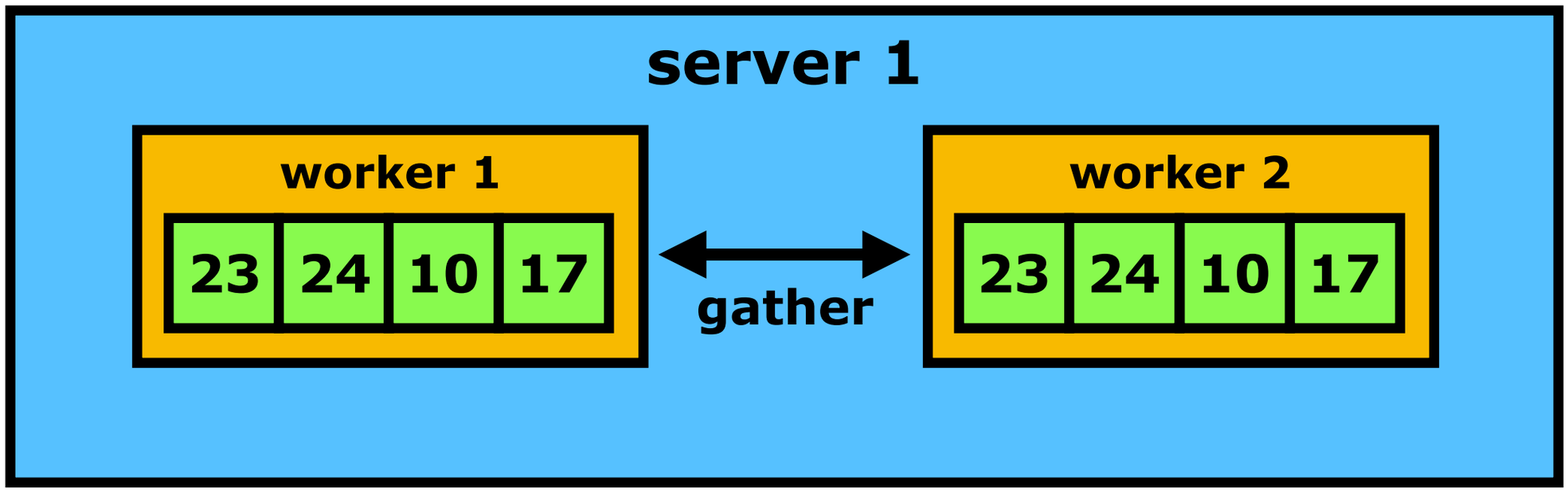
VALOHAI

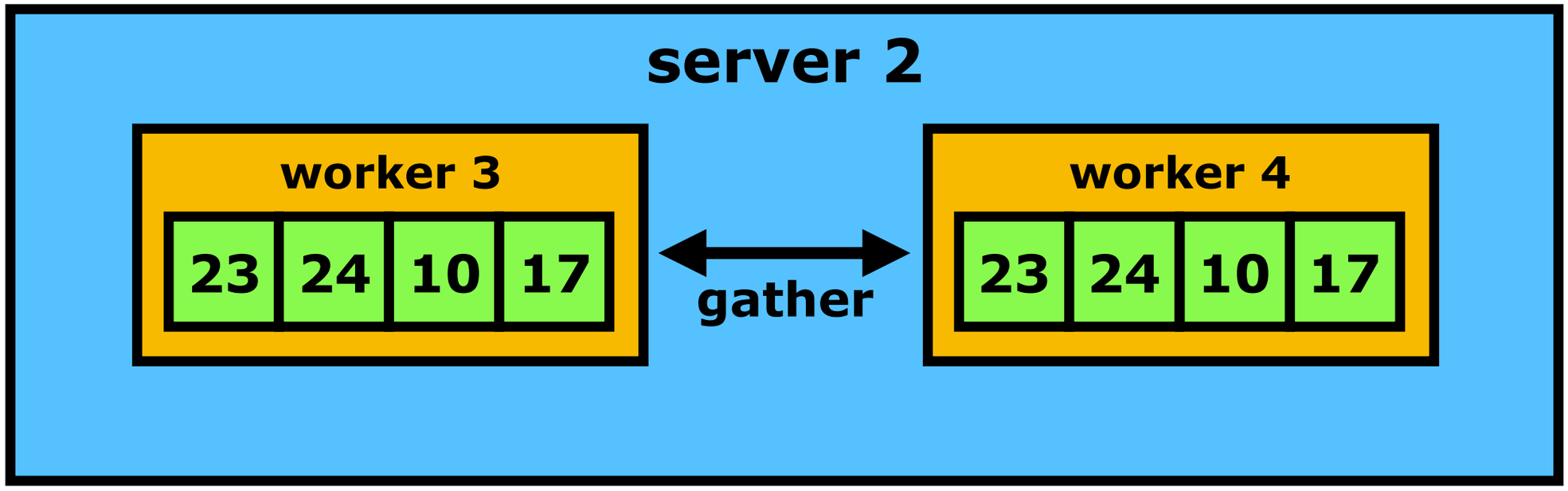
How Hierarchical Allreduce Works



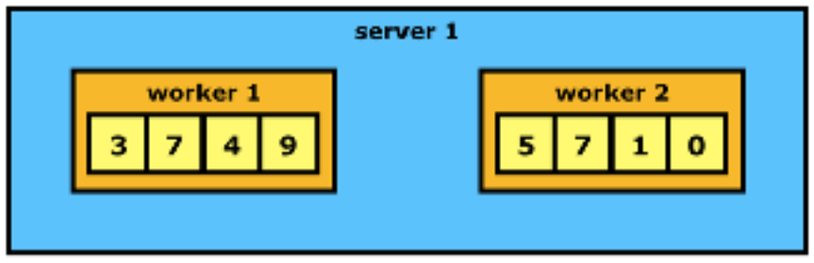
VALOHAI

How Hierarchical Allreduce Works

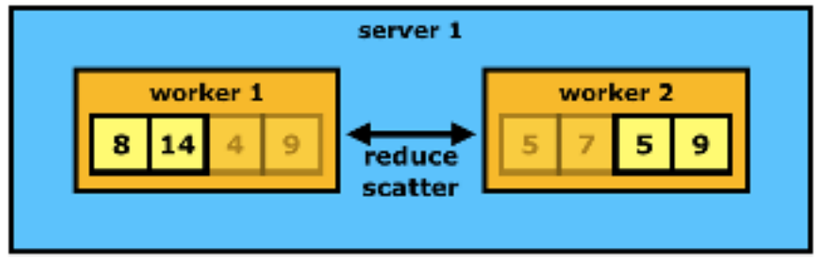

server 1
worker 1
23
24
10
17
gather
worker 2
23
24
10
17


server 2
worker 3
23
24
10
17
gather
worker 4
23
24
10
17

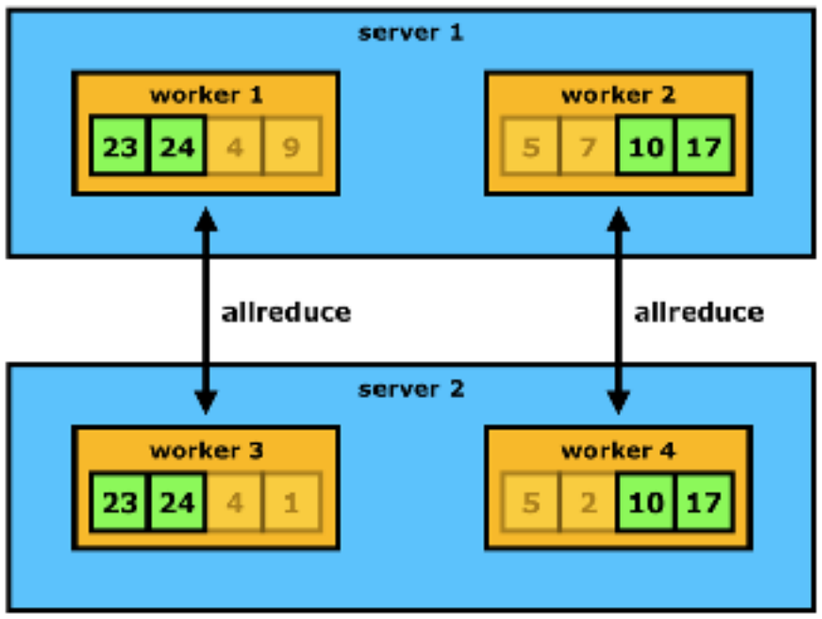
1. Setup



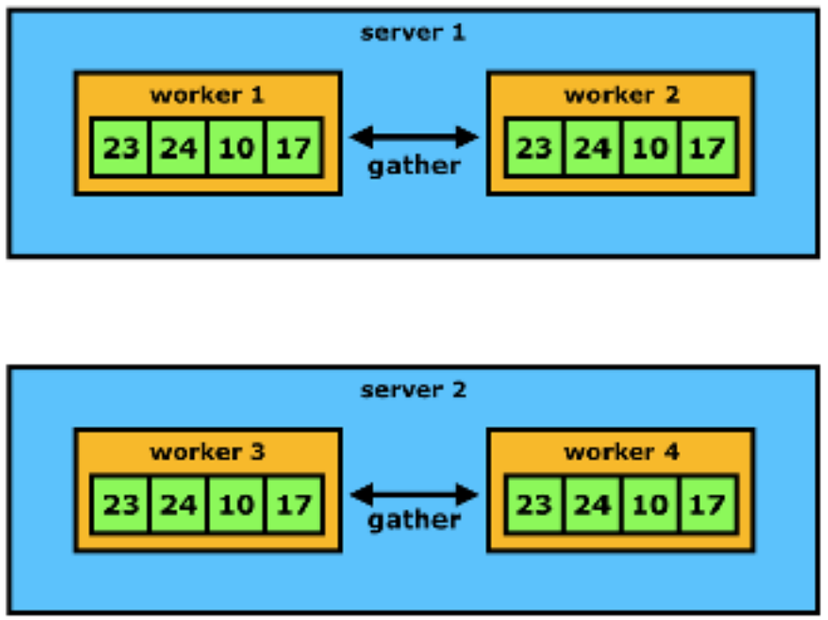
2. Local ReduceScatter



3. Global Allreduce



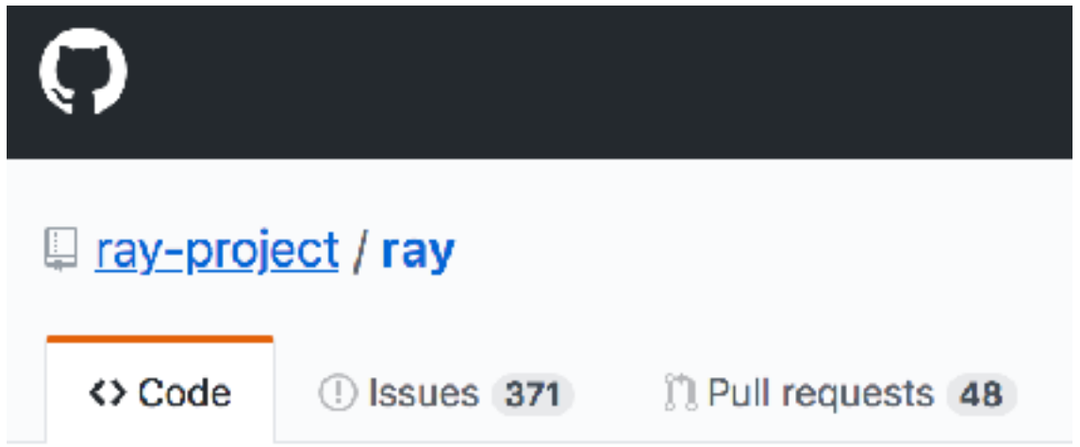
4. Local Gather



VALOHAI

How Hierarchical Allreduce Works

Tools


ray-project / ray
Code
Issues 371
Pull requests 48


HOROVOD


VALOHAI

VALOHAI

Deep Learning Management Platform

Sources and Further Reading

HOGWILD!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent
Feng Niu, Benjamin Recht, Christopher Ré, Stephen J. Wright

Large Scale Distributed Deep Networks
Jeffrey Dean, Greg S. Corrado, Rajat Monga, Kai Chen, et al.

Bandwidth Optimal All-reduce Algorithms for Clusters of Workstations
Pitch Patarasuk, Xin Yuan

Slim Fly: A Cost Effective Low-Diameter Network Topology
Maciej Besta, Torsten Hoefler

Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour
Priya Goyal, Piotr Dollar, , Ross Girshick, et al.

Horovod: Distributed Deep Learning in 5 Lines of Python
Uber Open Summit 2018

Distributed TensorFlow Documentation
TensorFlow

VALOHAI

Thank You!
and Q&A

Ruksi Laine
@ruk_si
Machine Learning Engineer
@valohaiai
https://valohai.com/