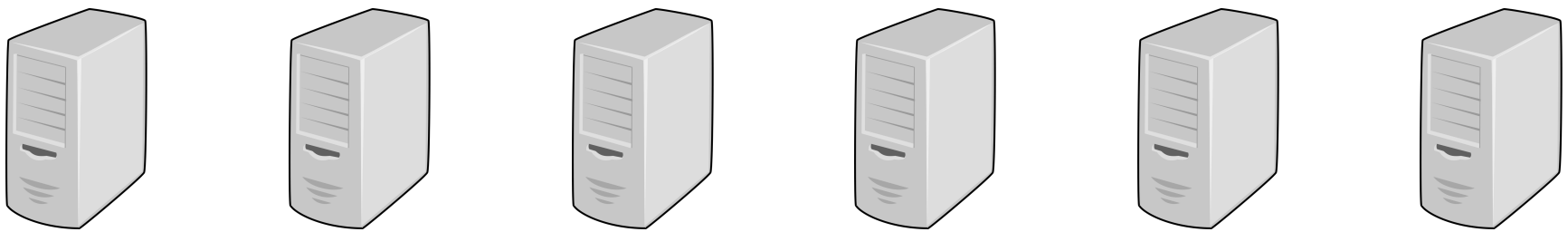
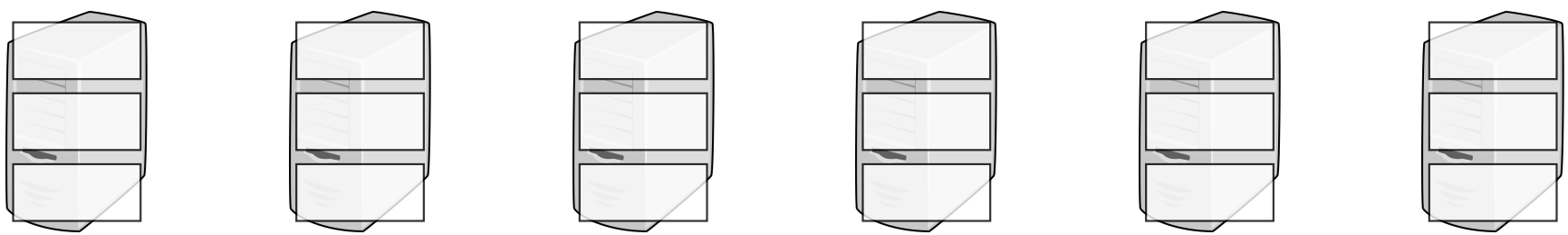
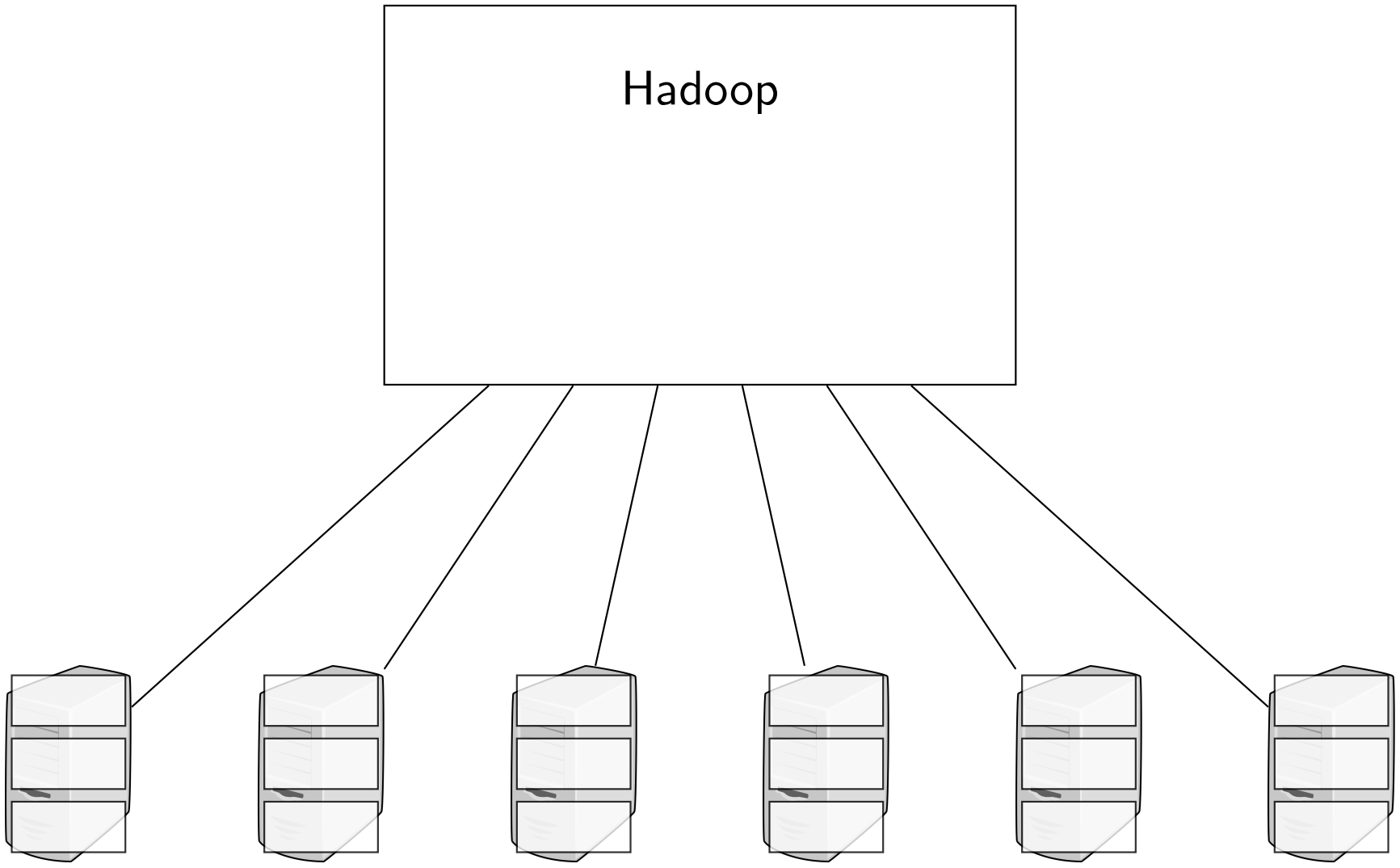
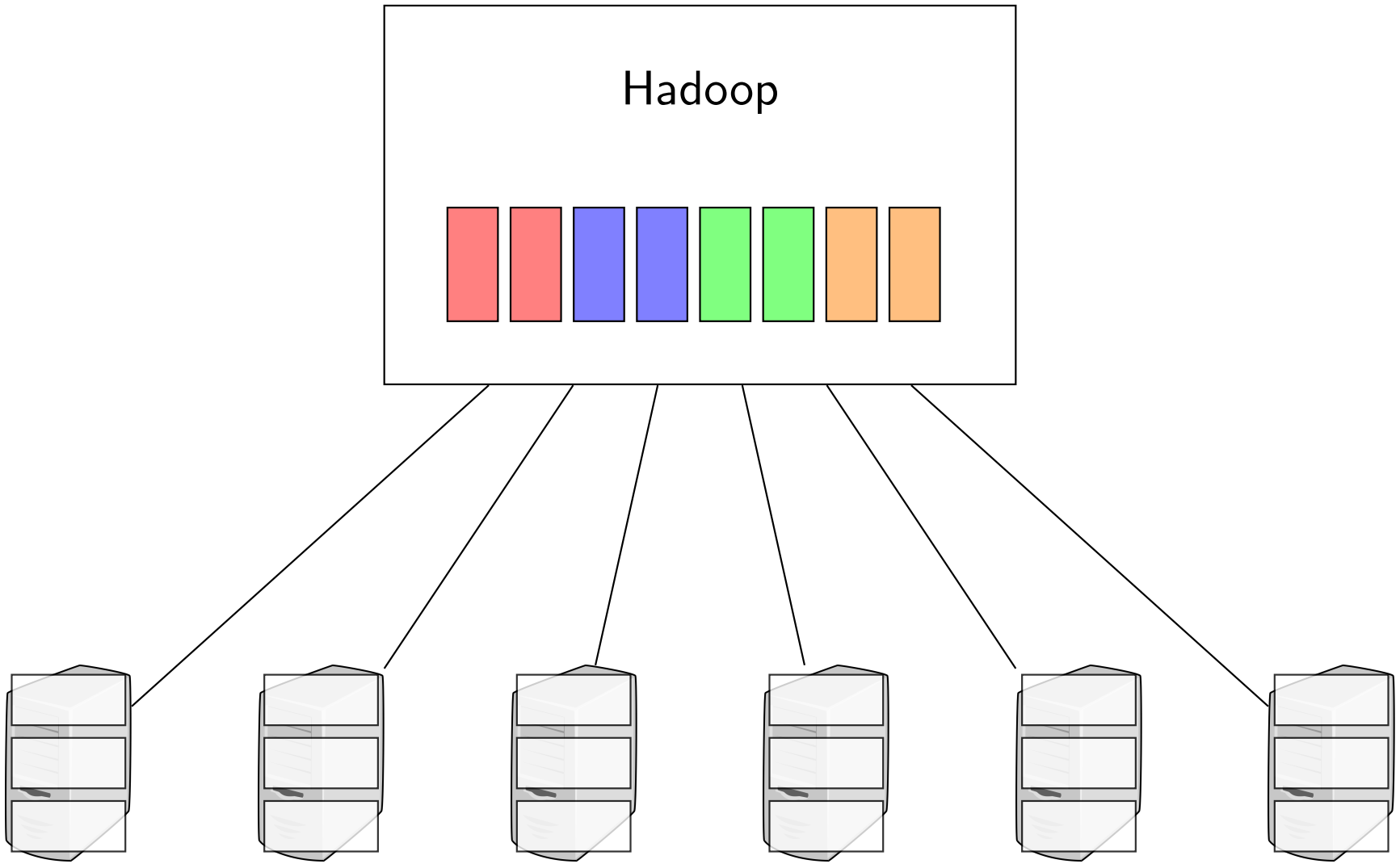
MESOS

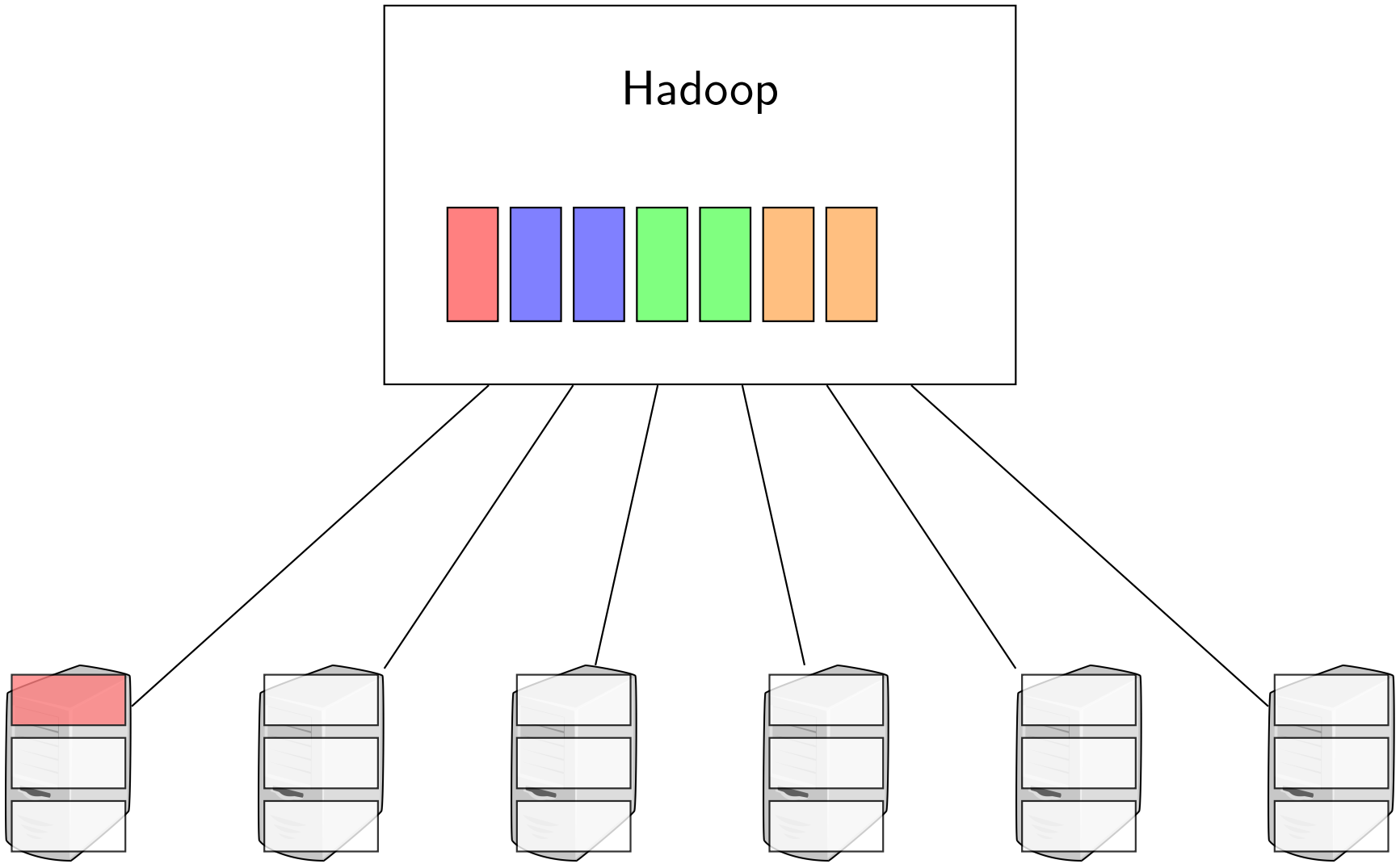
Michael Whittaker

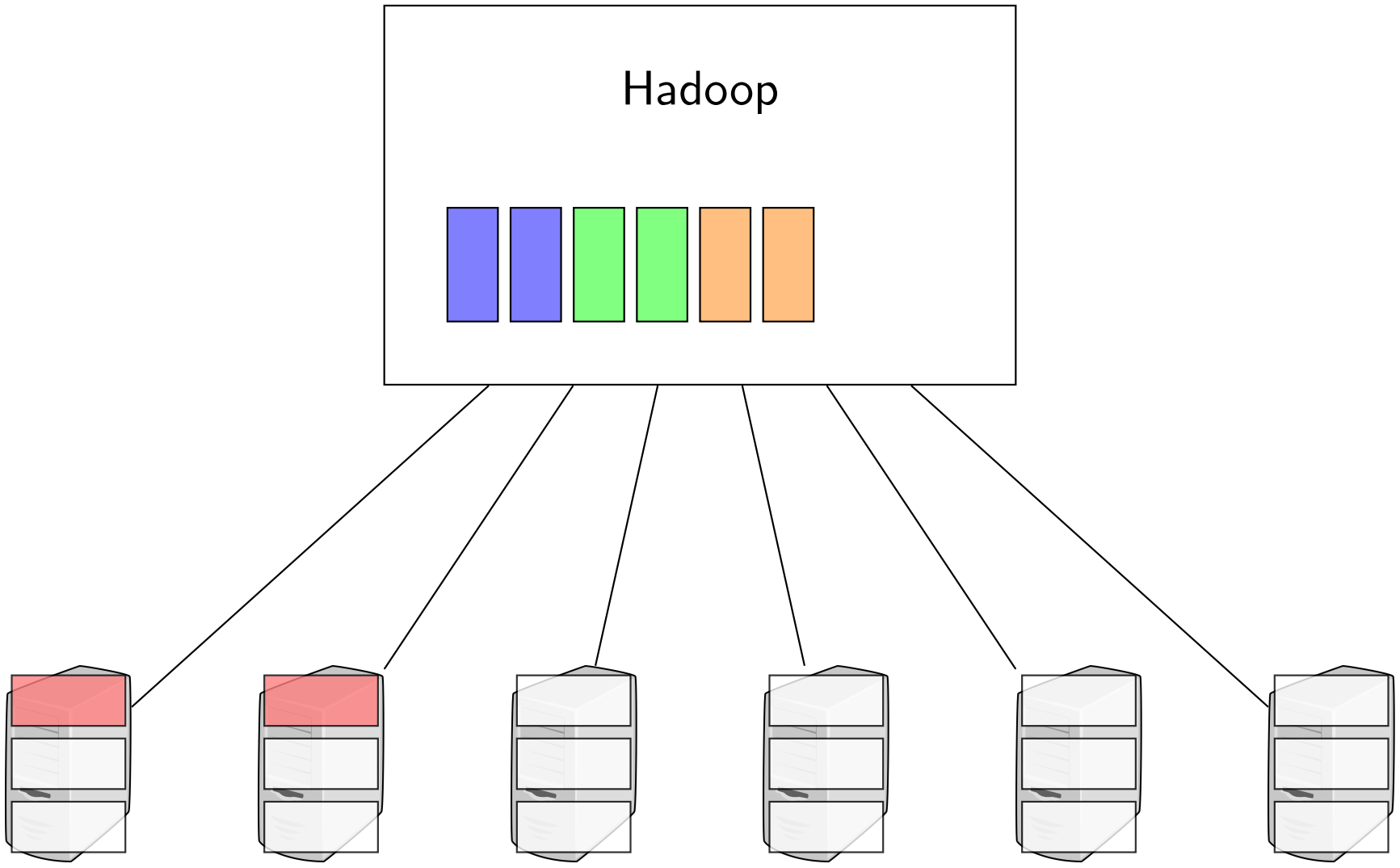


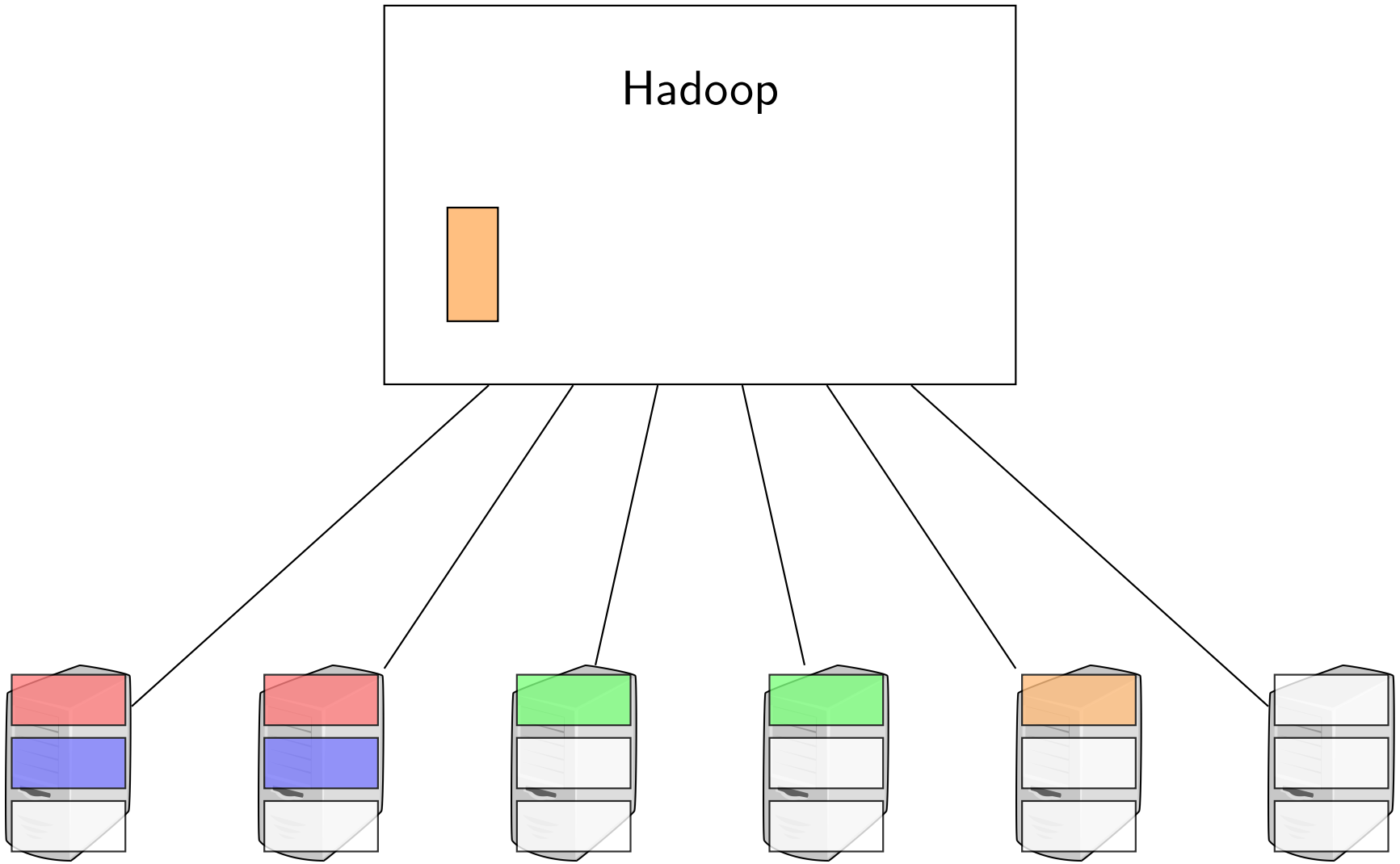


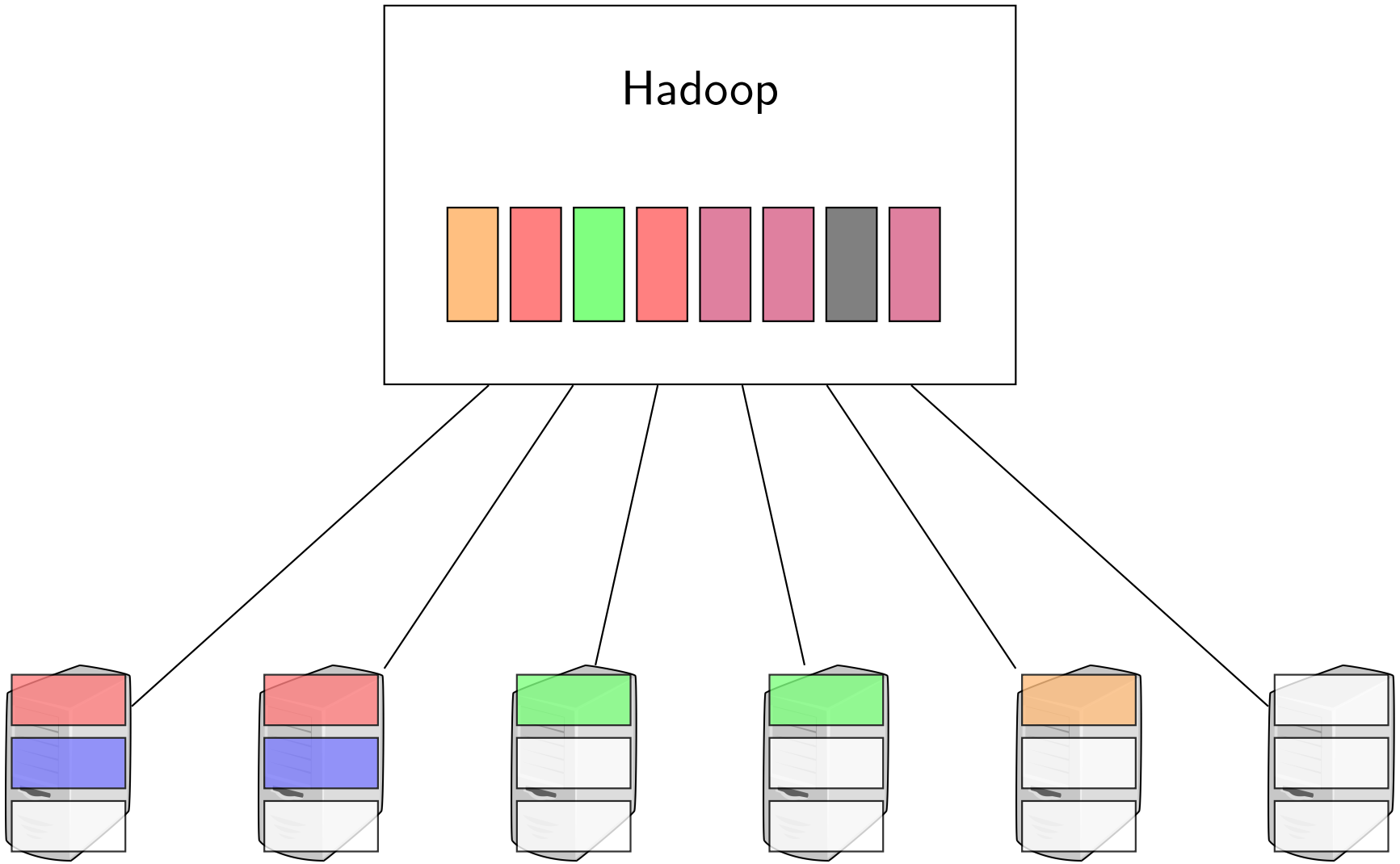

Hadoop

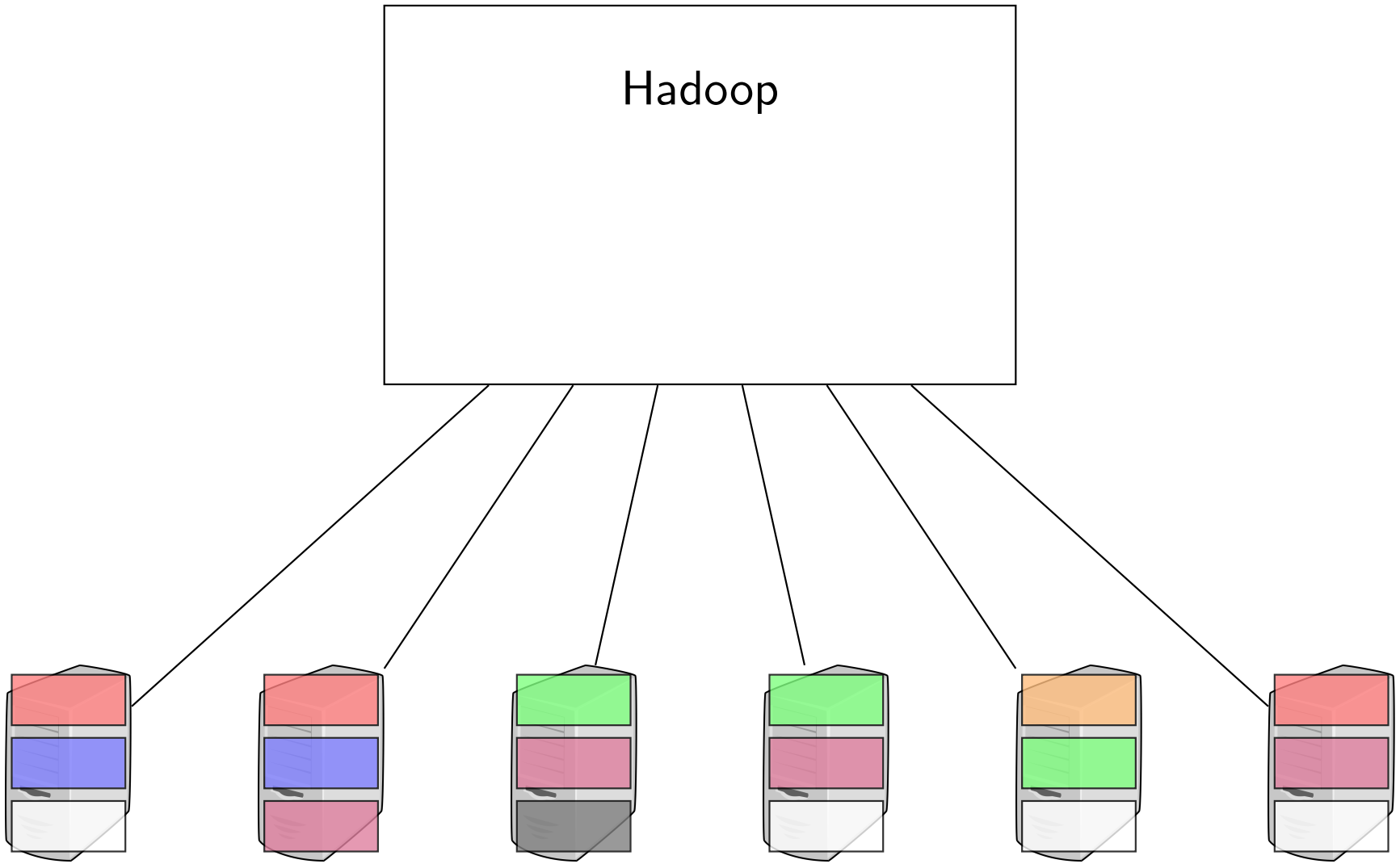

Hadoop


Hadoop


Hadoop

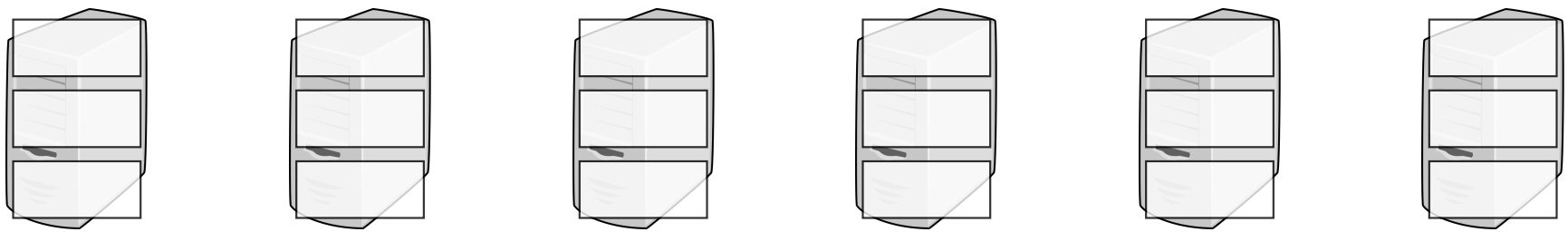

Hadoop

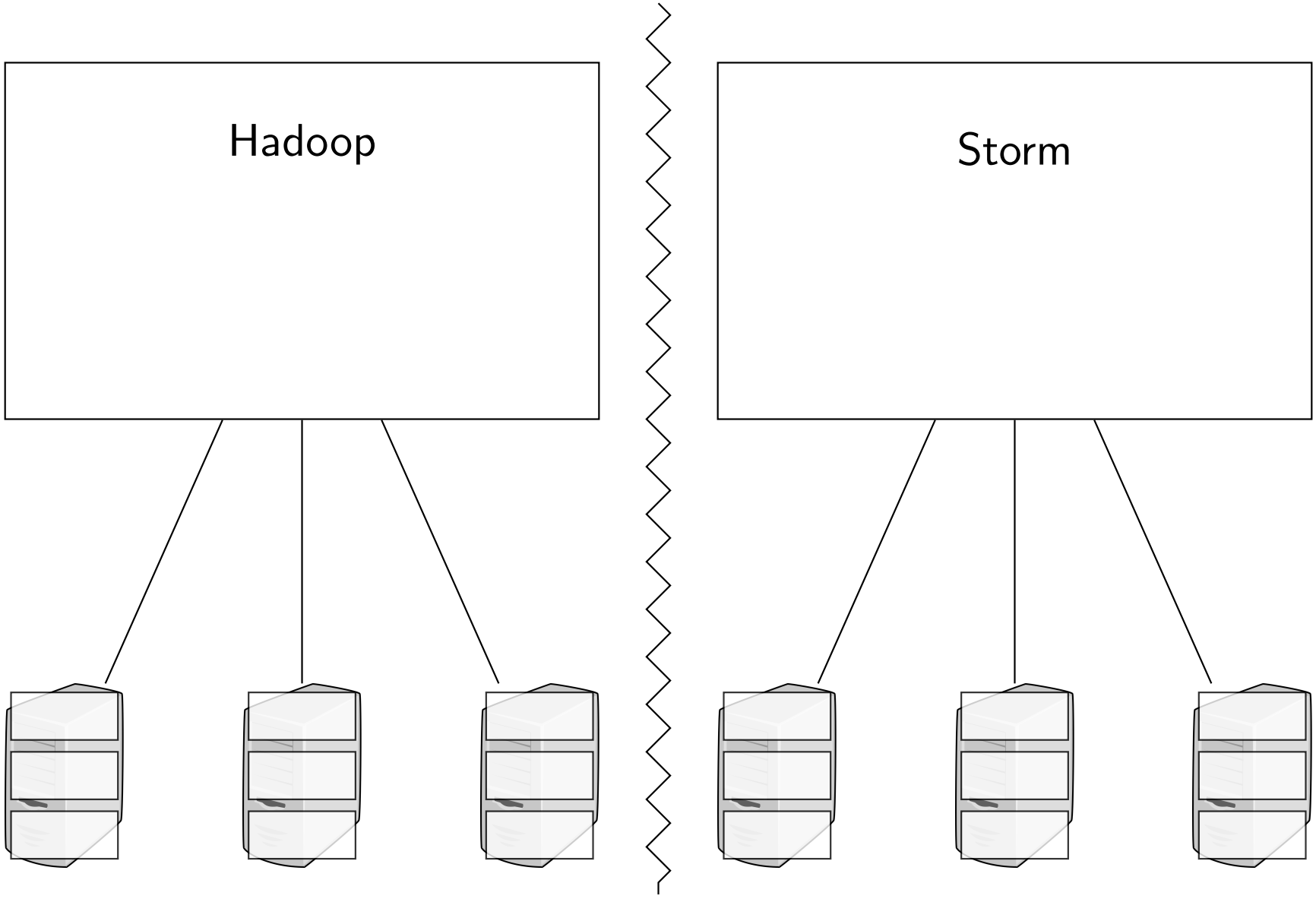

Hadoop

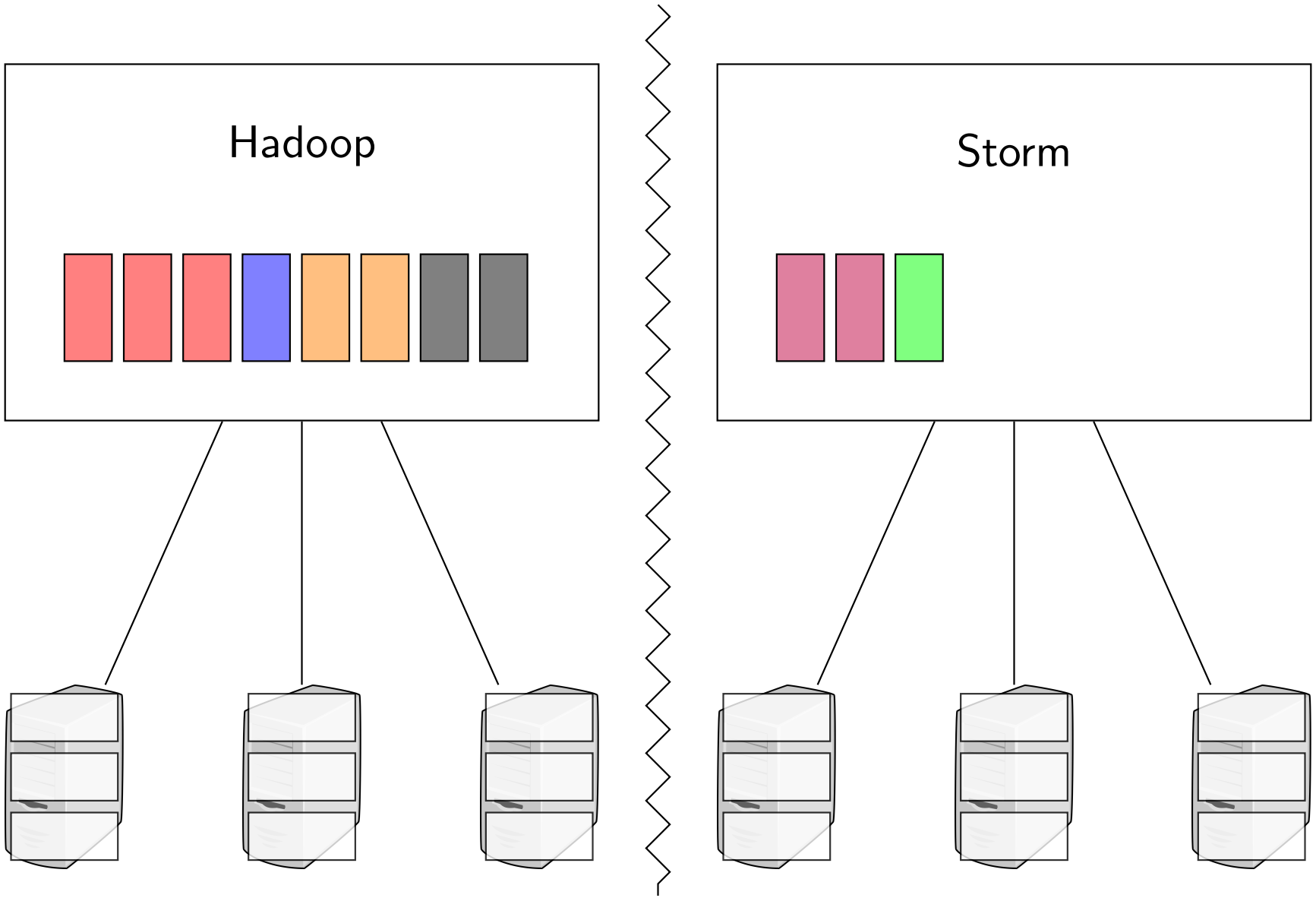

Hadoop

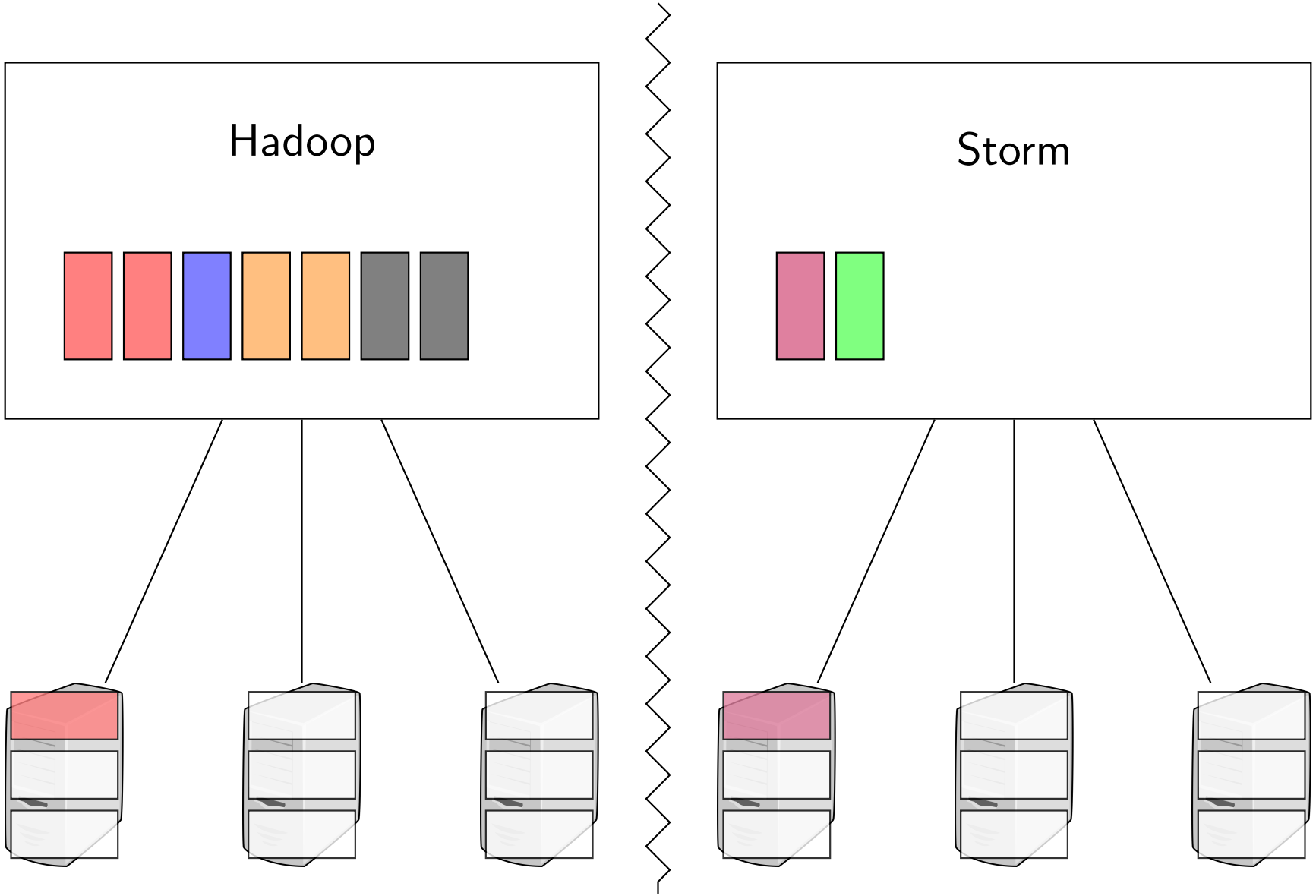
Hadoop

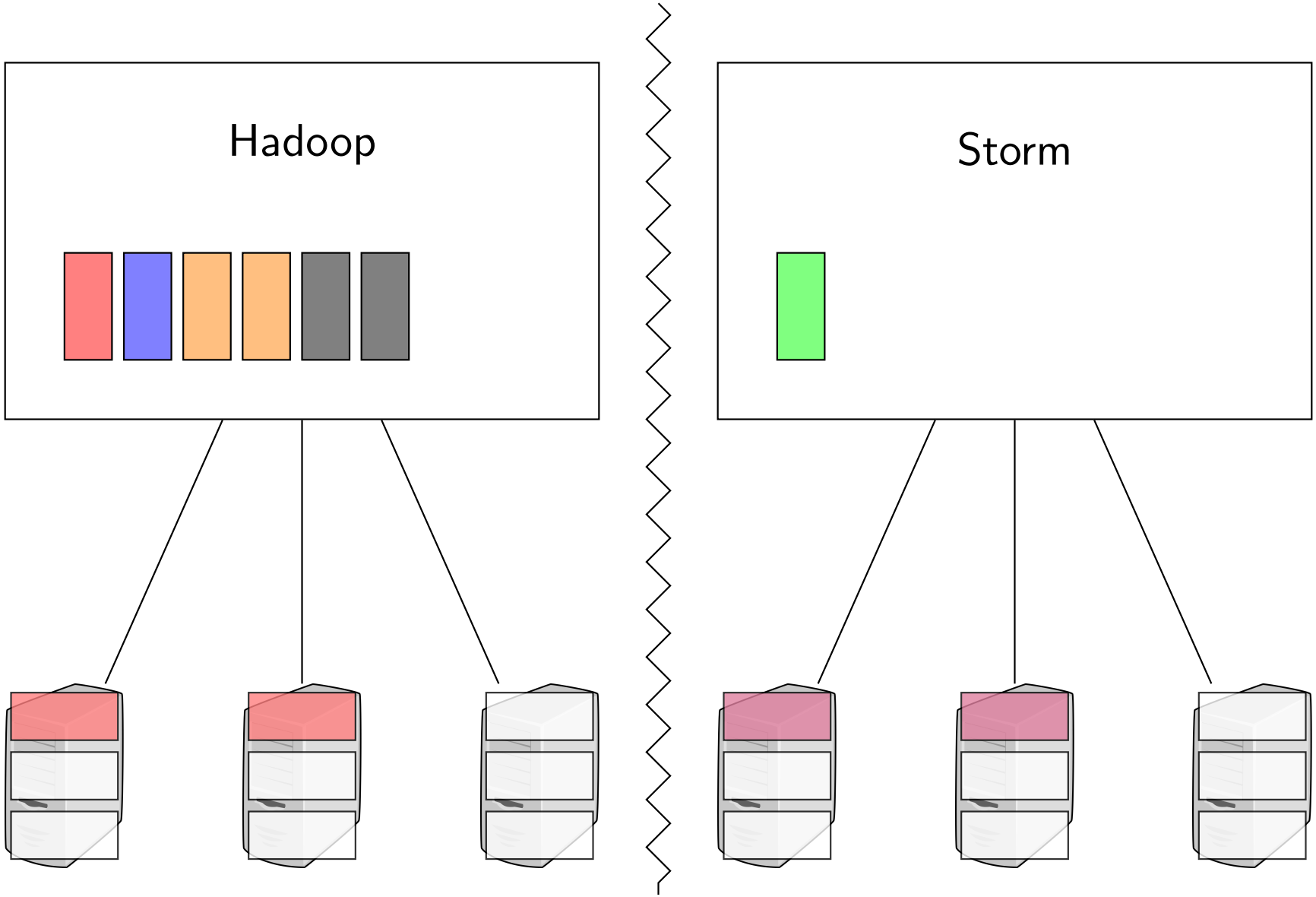
Storm

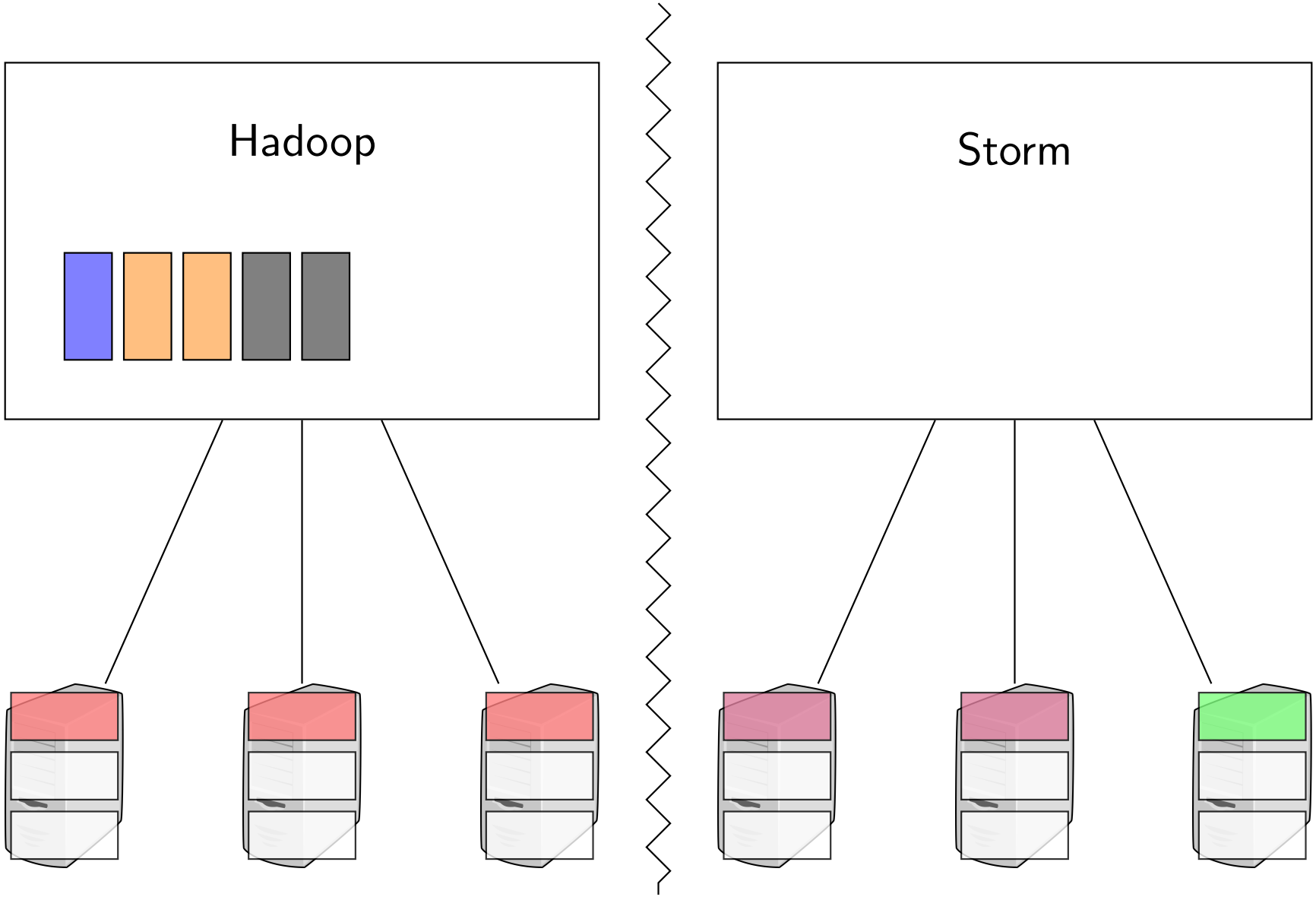


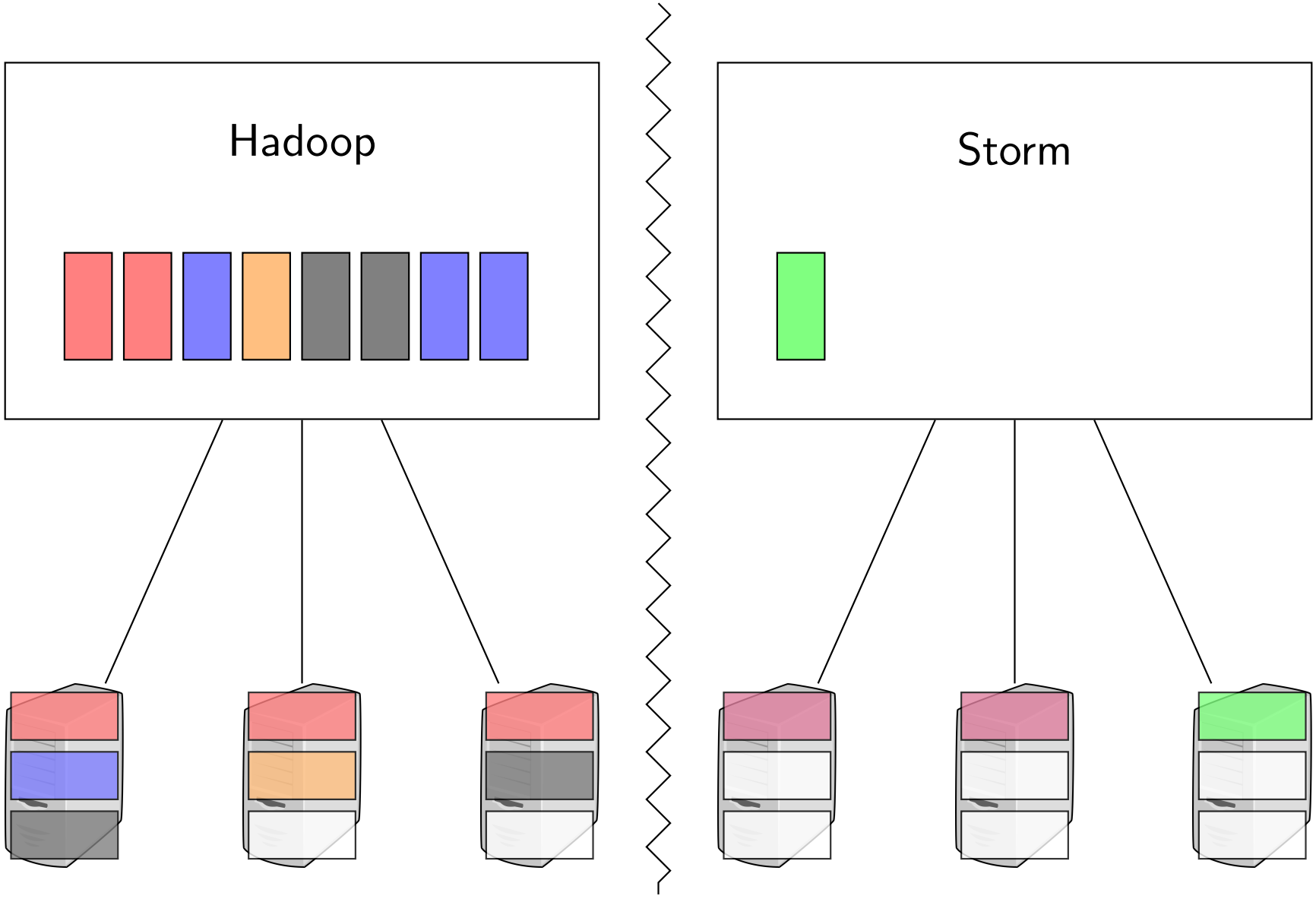

Hadoop
Storm

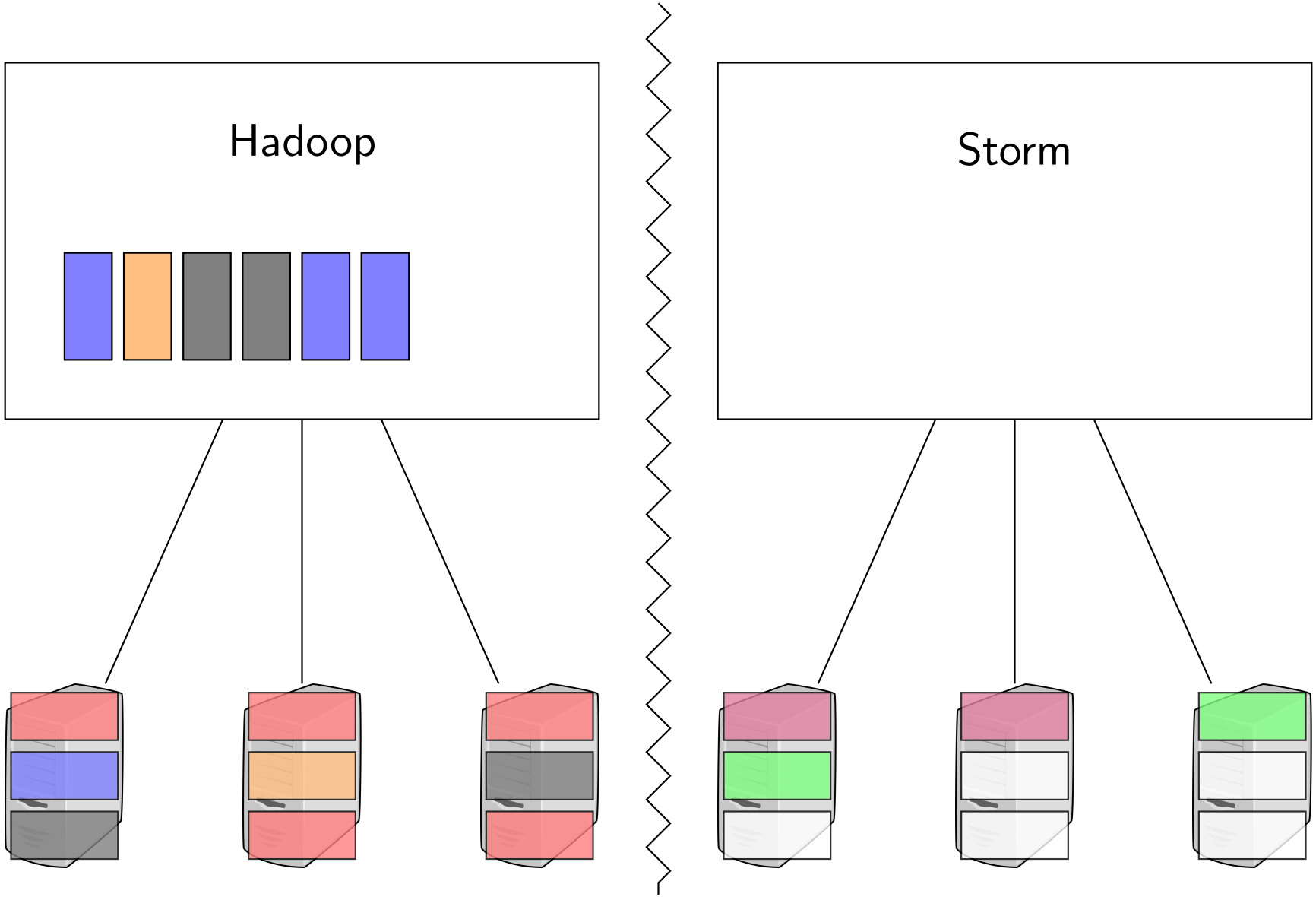

Hadoop
Storm

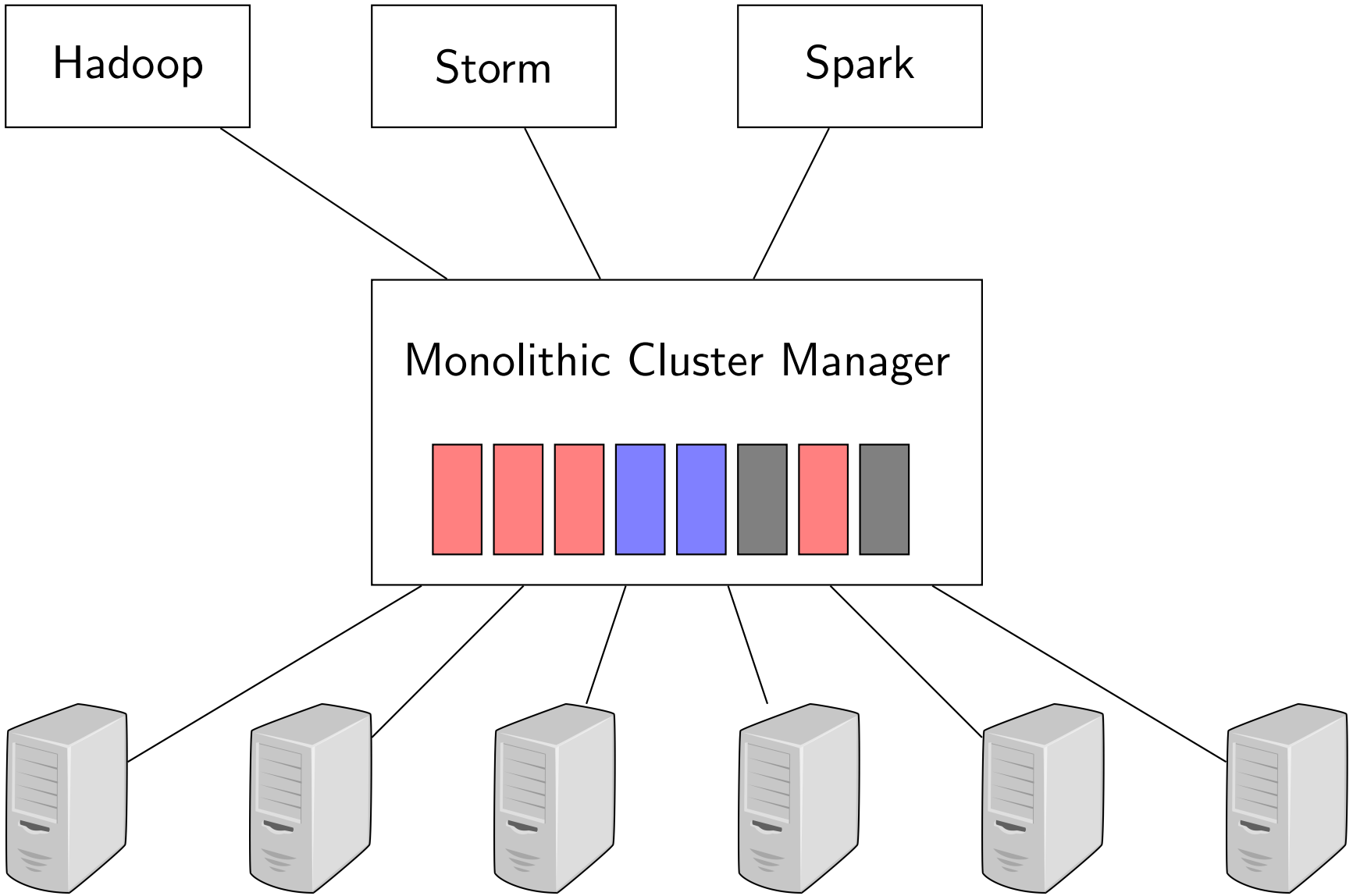

Hadoop
Storm

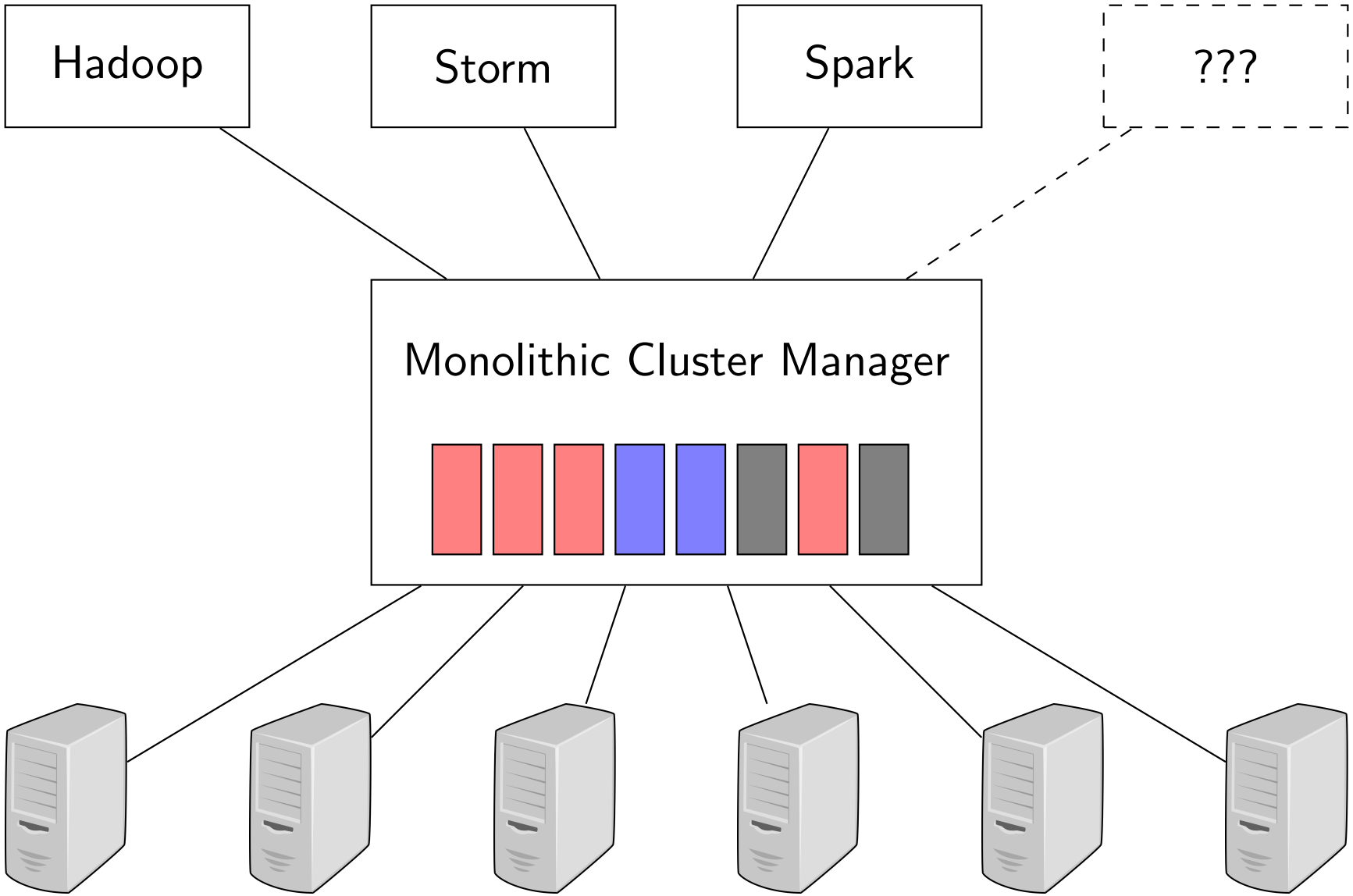

Hadoop
Storm

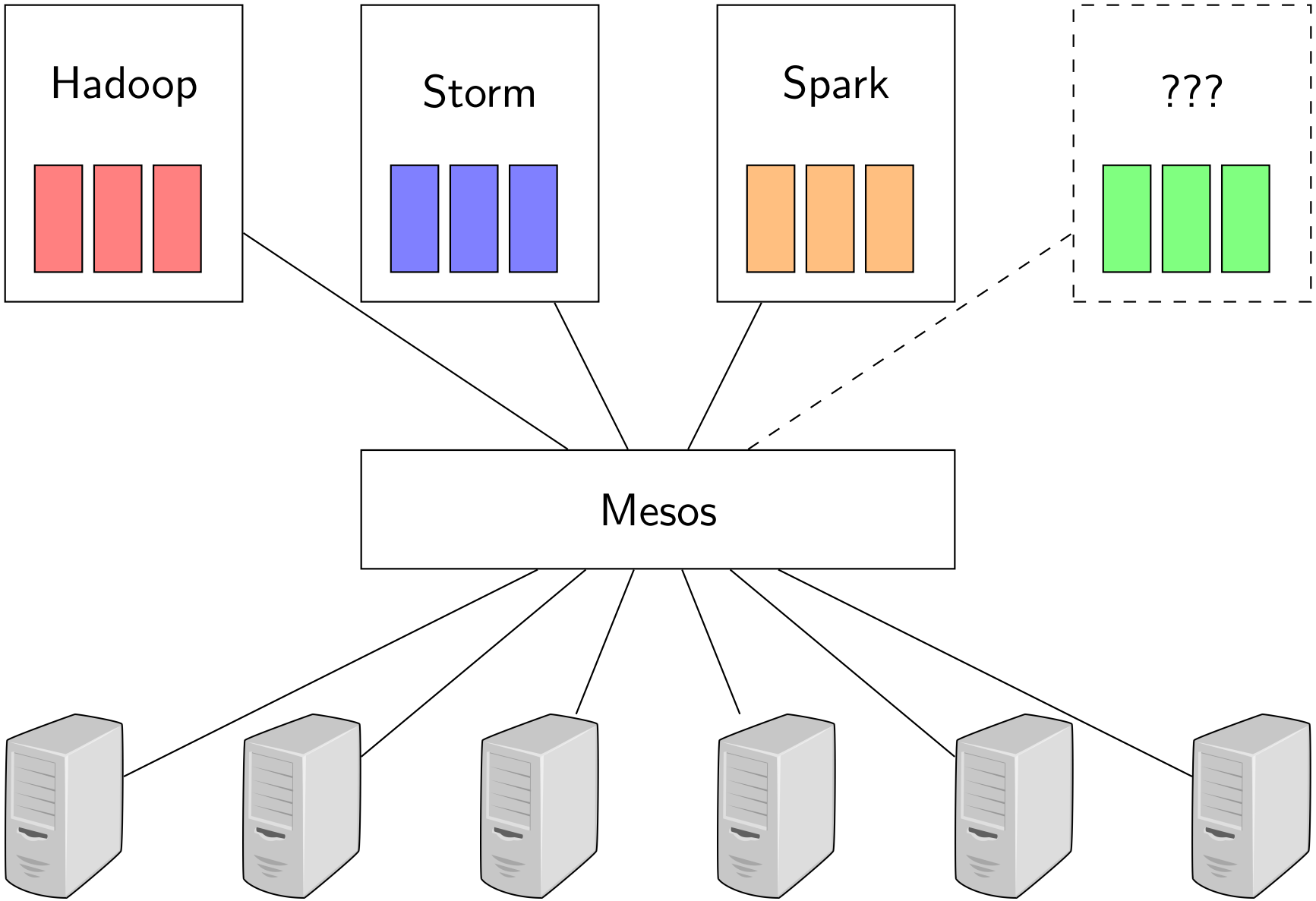

Hadoop
Storm

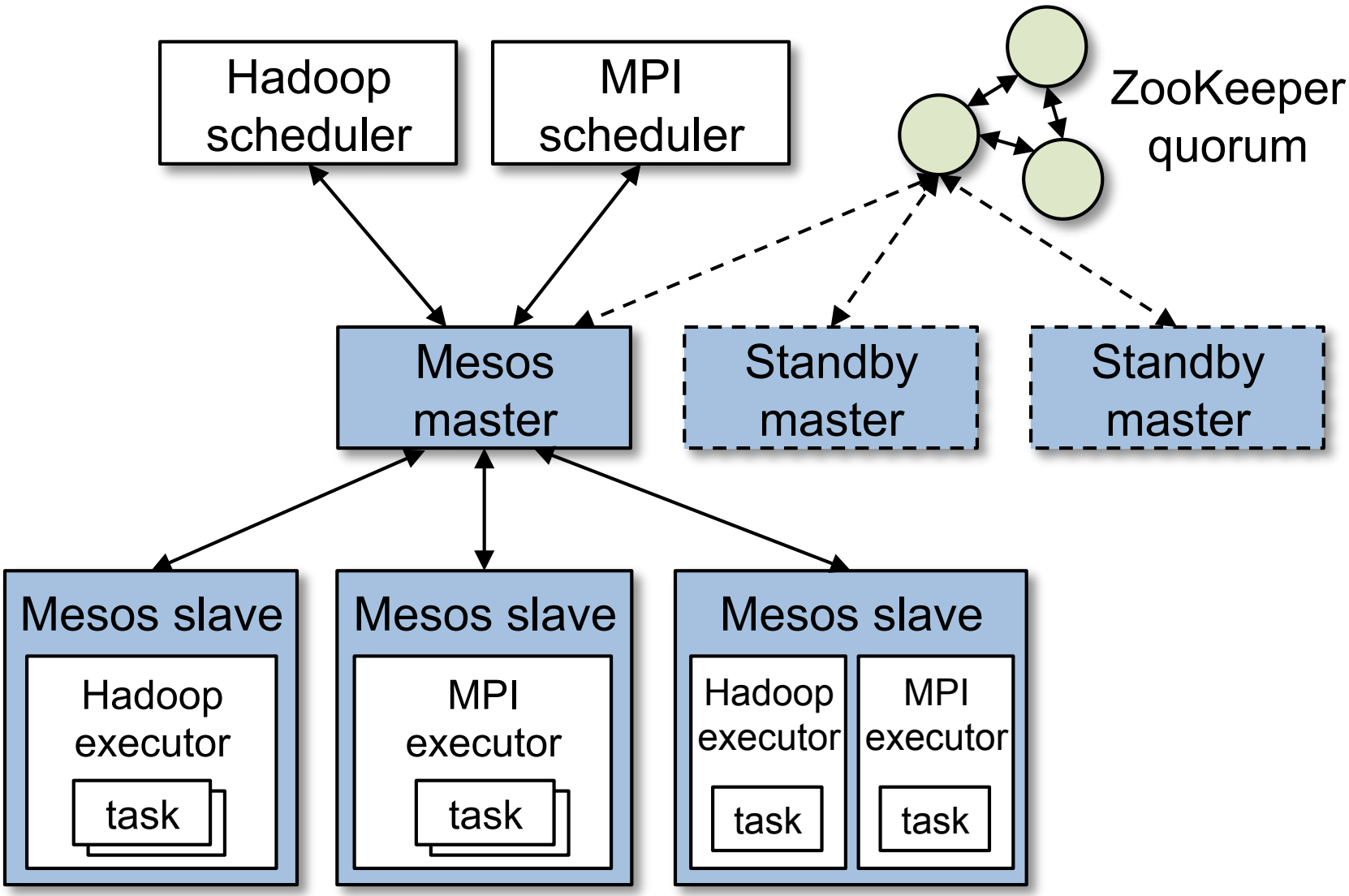

Hadoop
Storm


Hadoop
Storm

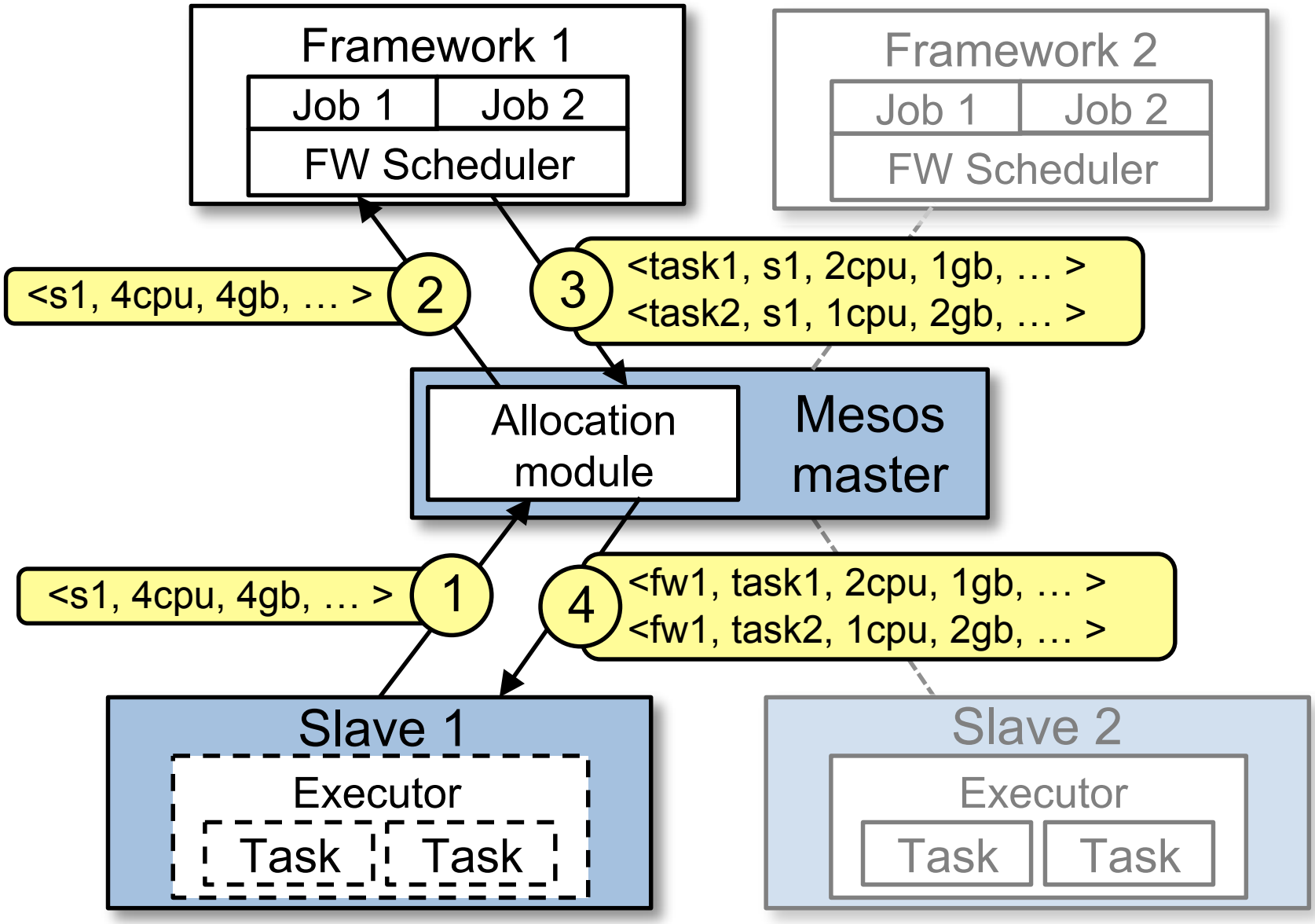

Hadoop
Storm
Spark
Monolithic Cluster Manager


Hadoop
Storm
Spark
???
Monolithic Cluster Manager


Hadoop
Storm
Spark
???
Mesos



Source: *Mesos: A Platform for Fine-Grained Resource Sharing in the Data Center*



Source: *Mesos: A Platform for Fine-Grained Resource Sharing in the Data Center*

Resource Allocation

- *Q: How are resources offered to frameworks?*

Resource Allocation

- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.

Resource Allocation

- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.
- *Q: When are resources offered to frameworks?*

Resource Allocation

- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.
- *Q: When are resources offered to frameworks?*
- A: Mesos assumes tasks are short and offers resources when tasks end.

Resource Allocation

- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.
- *Q: When are resources offered to frameworks?*
- A: Mesos assumes tasks are short and offers resources when tasks end.
- *Q: What if tasks aren't short?*

Resource Allocation

- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.
- *Q: When are resources offered to frameworks?*
- A: Mesos assumes tasks are short and offers resources when tasks end.
- *Q: What if tasks aren't short?*
- A: Mesos can kill tasks, giving the framework a grace period for cleaning up.

Resource Allocation

- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.
- *Q: When are resources offered to frameworks?*
- A: Mesos assumes tasks are short and offers resources when tasks end.
- *Q: What if tasks aren't short?*
- A: Mesos can kill tasks, giving the framework a grace period for cleaning up.
- *Q: What if jobs don't want to die?*

Resource Allocation

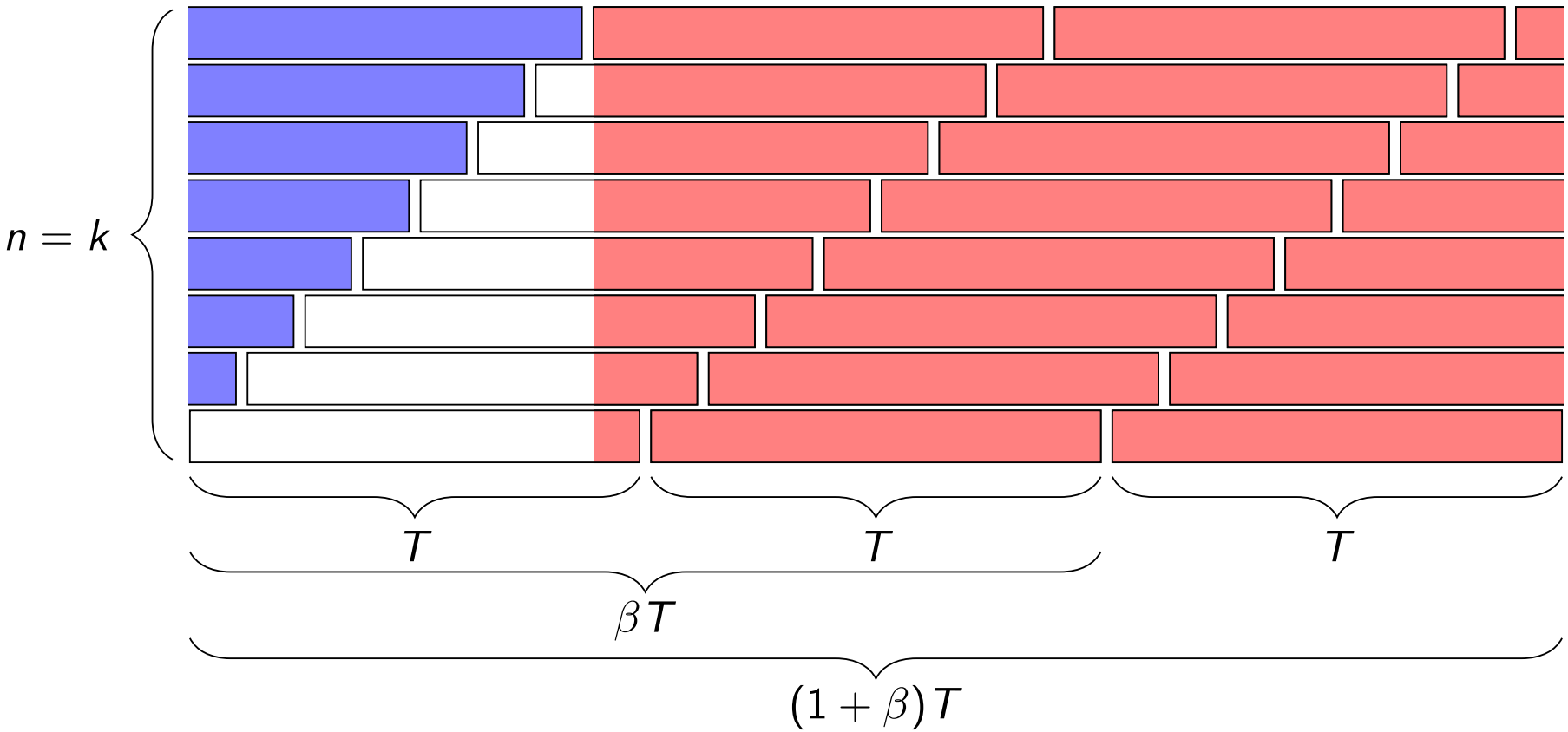
- *Q: How are resources offered to frameworks?*
- A: Pluggable allocation module determines how resources are offered to frameworks.
- *Q: When are resources offered to frameworks?*
- A: Mesos assumes tasks are short and offers resources when tasks end.
- *Q: What if tasks aren't short?*
- A: Mesos can kill tasks, giving the framework a grace period for cleaning up.
- *Q: What if jobs don't want to die?*
- A: Mesos provides each framework with a guarenteed allocation. So long as framework uses less than it's guarenteed allocation, it's jobs won't be killed.

Demo

Mesos performs best with *elastic* frameworks and *homogenous* task durations.

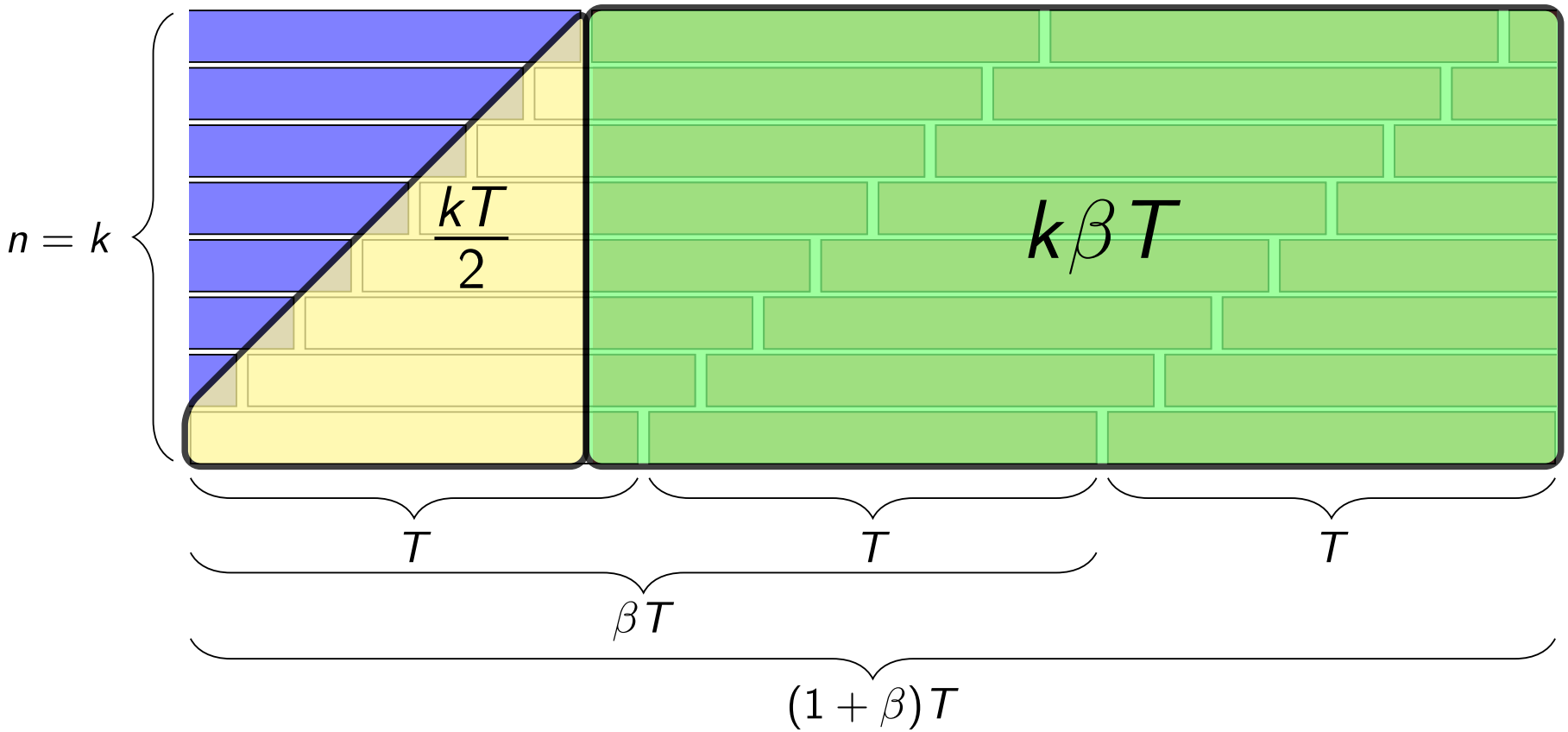
Rigid Framework

- Ramp-up time: T
- Completion time: $(1+\beta)T$
- Utilization: $\frac{\beta}{\frac{1}{2}+\beta}$



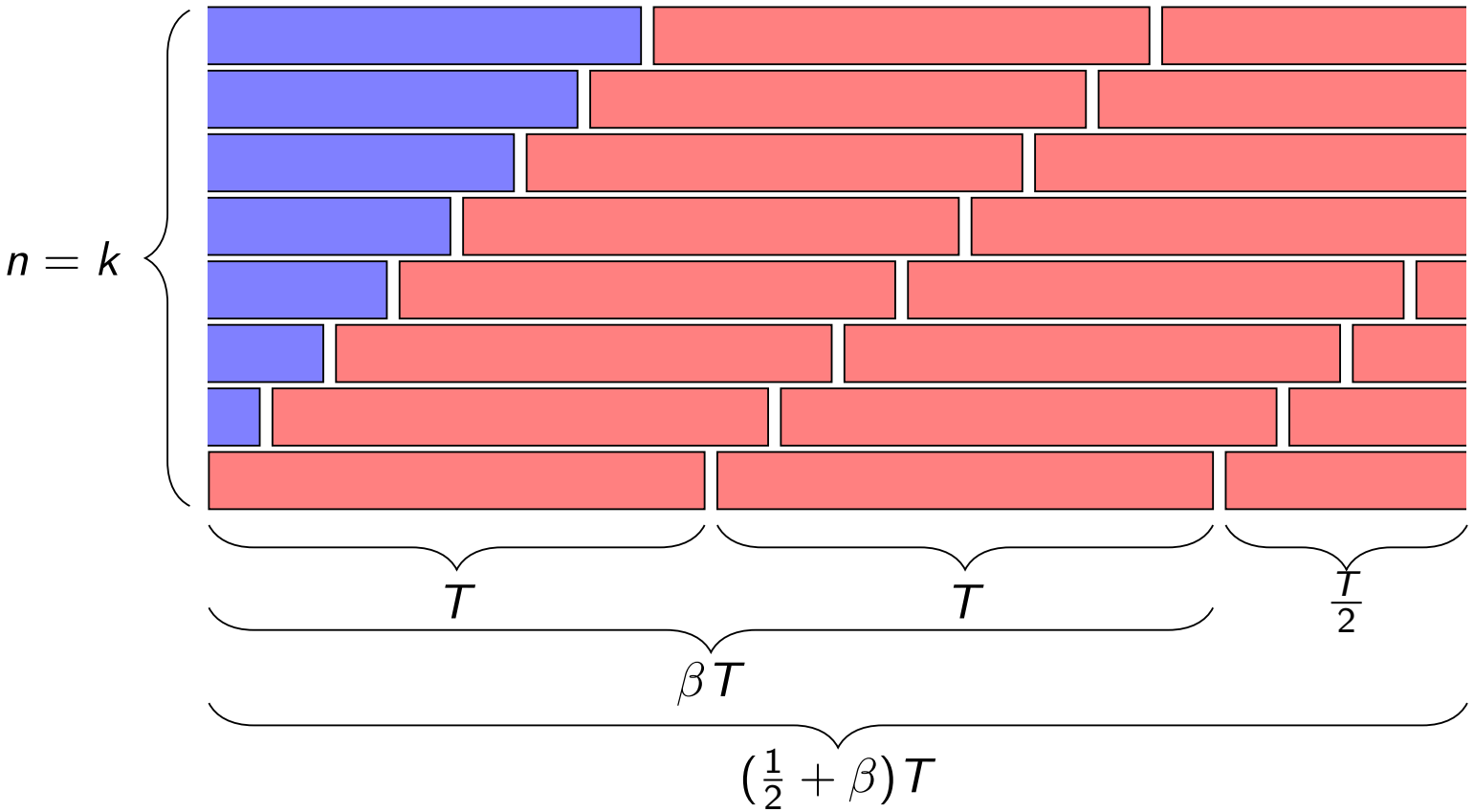
Rigid Framework

- Ramp-up time: T
- Completion time: $(1+\beta)T$
- Utilization: $\frac{\beta}{\frac{1}{2}+\beta}$



Elastic Framework

- Ramp-up time: T
- Completion time: $(\frac{1}{2} + \beta)\,T$
- Utilization: 1



Implementation

- 10,000 lines of C++
- Supported Hadoop, Torque, and Spark
- Lots of impressive performance benchmarks

Lessons

- Clusters need schedulers to improve utilization
- Schedulers should form a narrow waist between frameworks and the cluster
- Be simple
- Adhere to the end-to-end argument

Questions

Extras

Isolation

https://goo.gl/dqQ6lG

Scalability and Robustness

- Frameworks can install filters with the Mesos master.
- Offered resources count against a framework's allocation.
- If a frameworks is slow to respond, Mesos can rescind offers.

Fault Tolerance

- Mesos uses soft state derivable from slaves and frameworks.
- Hot standby replicas.
- Each framework can install multiple schedulers.

Placement Preferences

Placement preferences can be achieved with delay and lottery scheduling.

Heterogenous Tasks

If the number of slots on each machine is big, the chances that a machine will be filled completely with long lived tasks is small.