Scheduling and Managing Tasks with YARN

Overview

Understanding how YARN schedules tasks

Understanding different scheduling strategies i.e. FIFO, Capacity and Fair Scheduler

Configuring multiple queues and setting a job to run on a specific queue

YARN

Yet Another Resource Negotiator

YARN

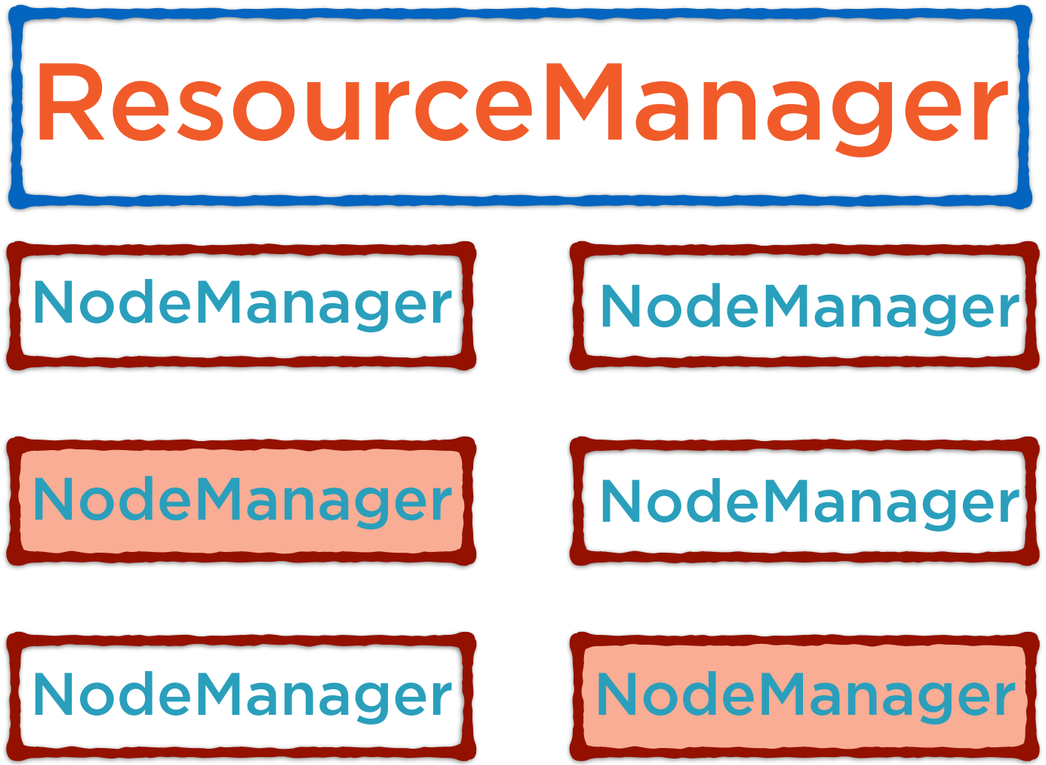
Co-ordinates tasks running on the cluster

Assigns new nodes in case of failure

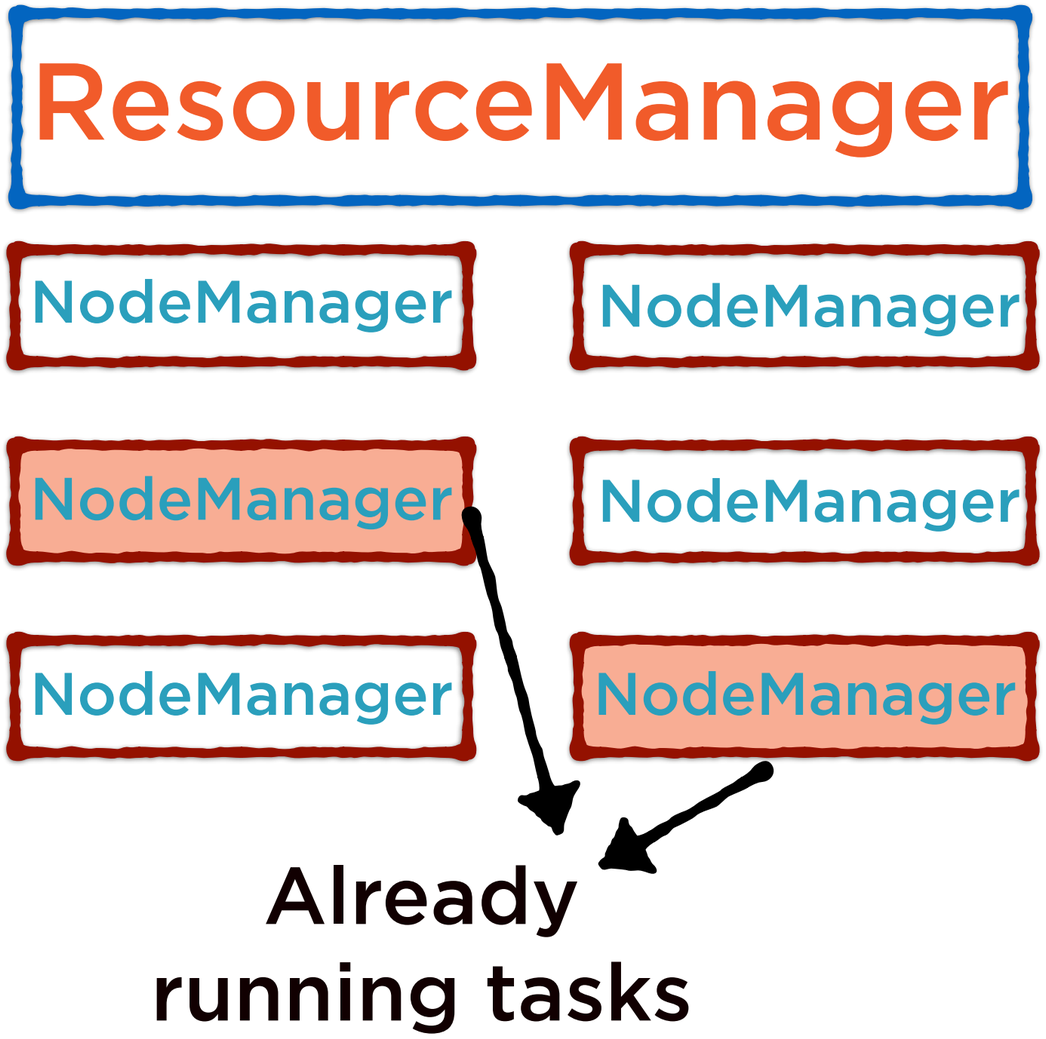
YARN

Resource Manager	Node Manager
Runs on a single master node	Run on all other nodes
Schedules tasks across nodes	Manages tasks on the individual node

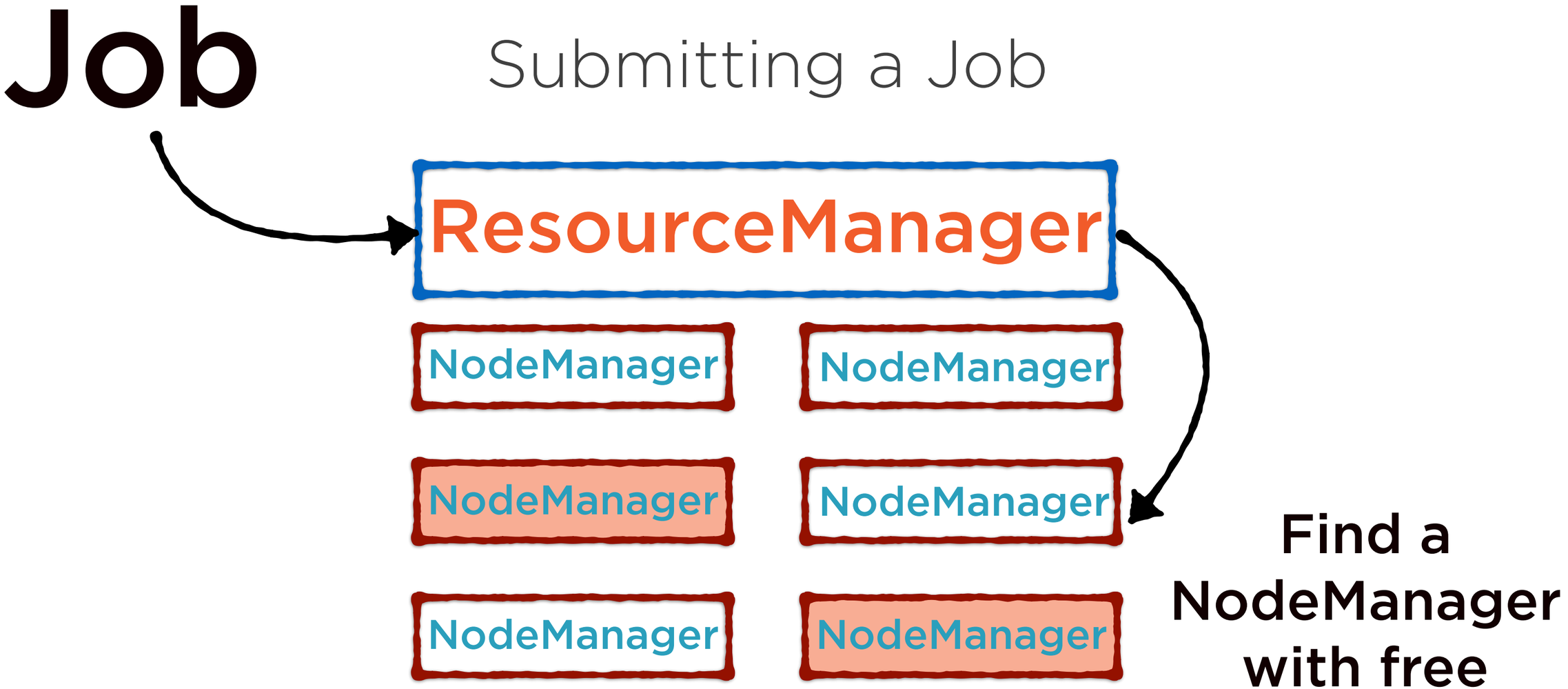
YARN


ResourceManager
NodeManager
NodeManager
NodeManager
NodeManager
NodeManager
NodeManager

YARN

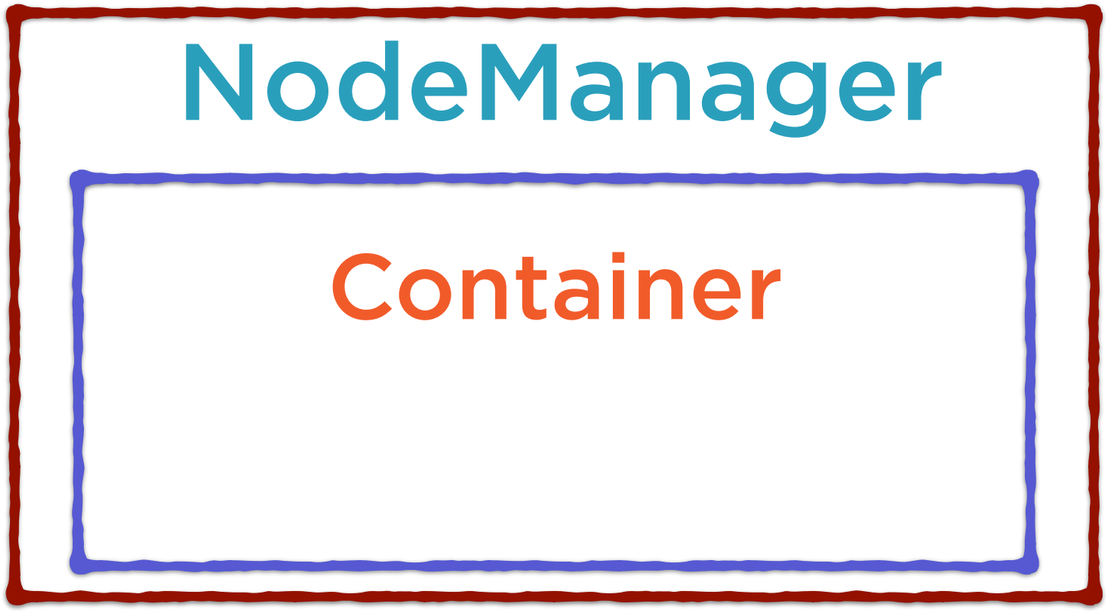

ResourceManager
NodeManager
NodeManager
NodeManager
NodeManager
NodeManager
NodeManager
Already
running tasks

Submitting a Job



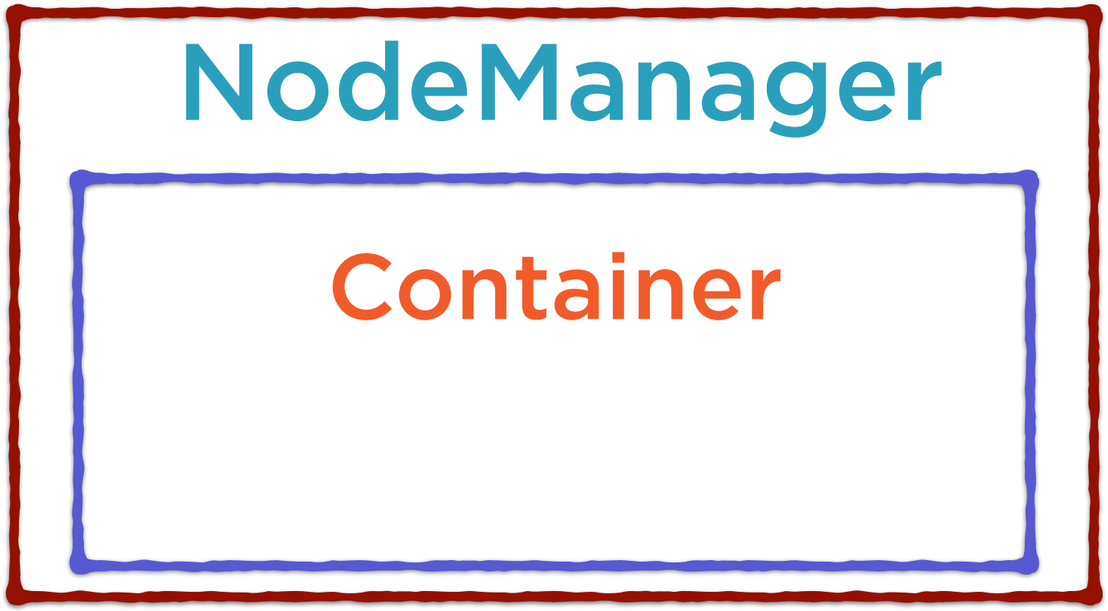
Application Master Process

All processes on a node are run within a container



Application Master Process

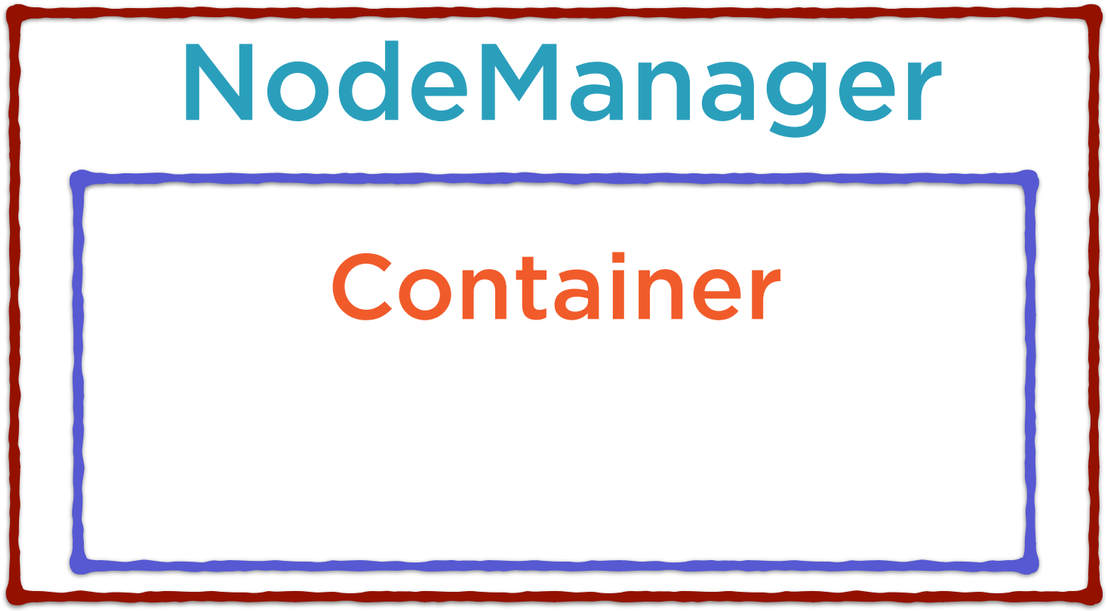
This is the logical unit for resources the process needs - memory, CPU etc



Application Master Process

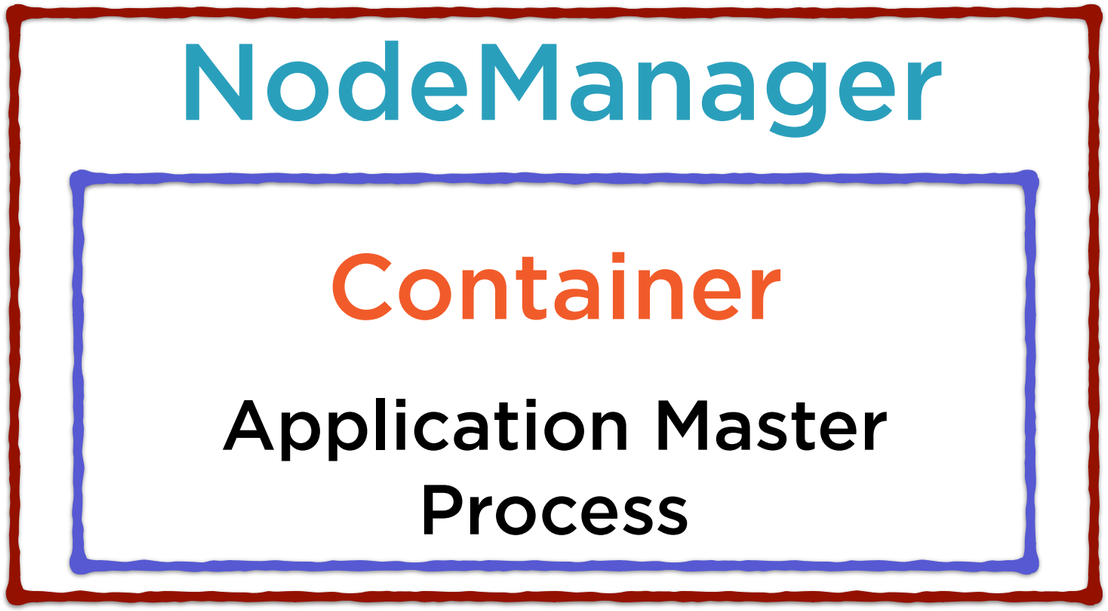
A container executes a specific application

1 NodeManager can have multiple containers



Application Master Process

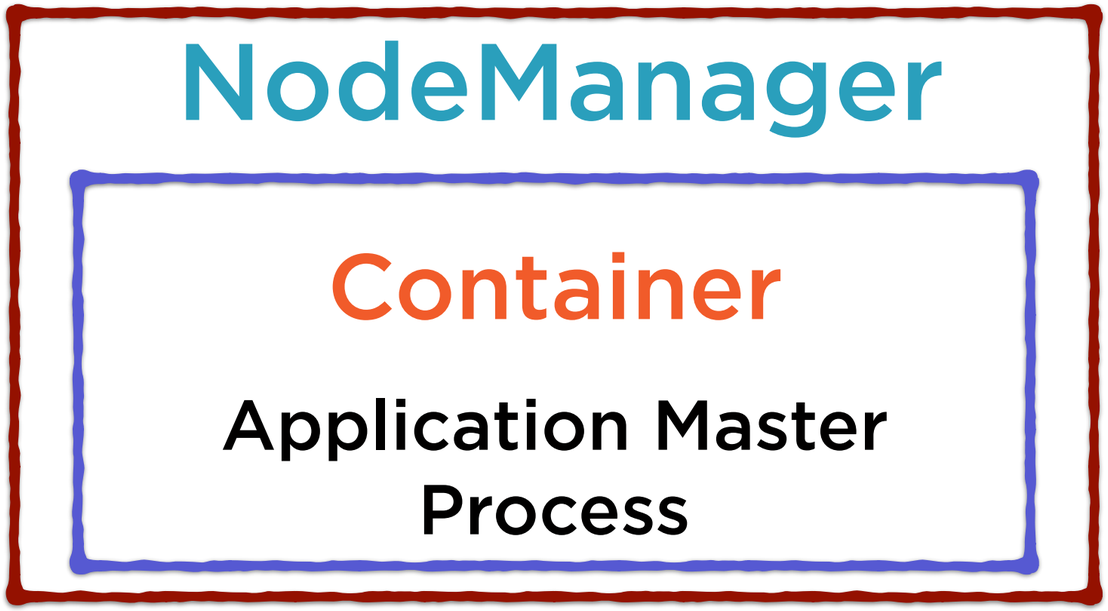
The ResourceManager starts off the Application Master within the Container



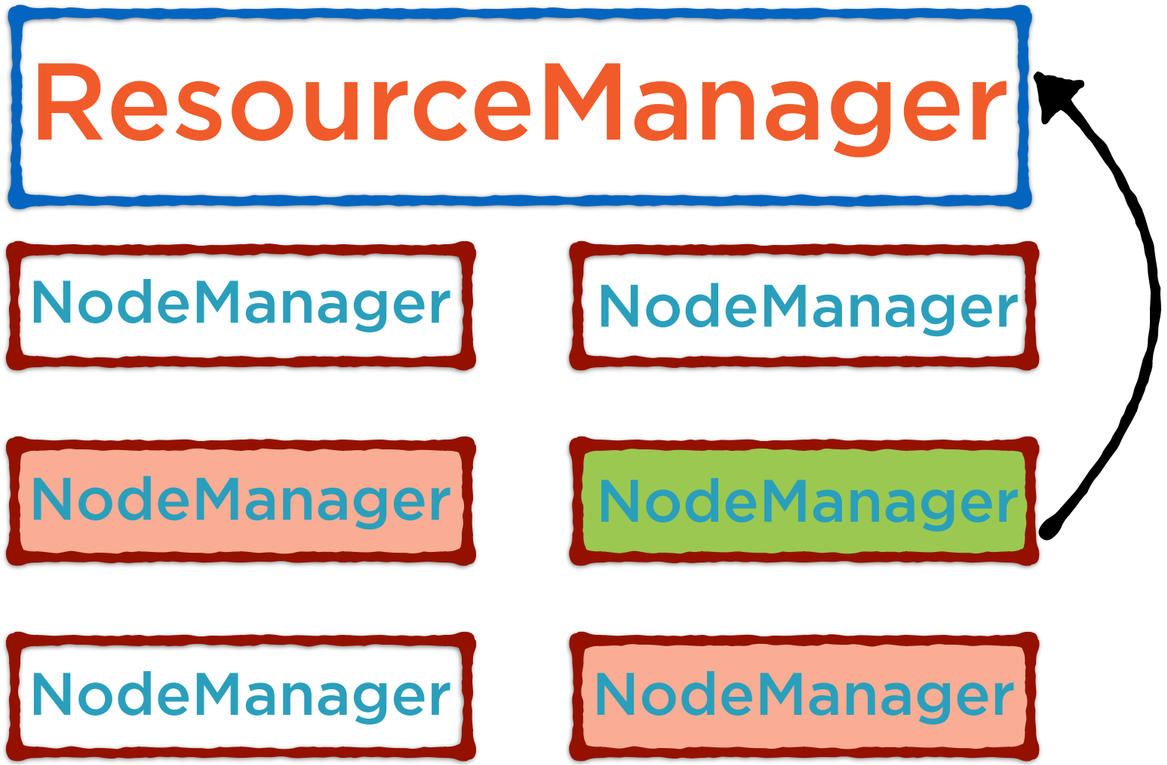
Application Master Process

Performs the computation required for the task

If additional resources are required, the Application Master makes the request



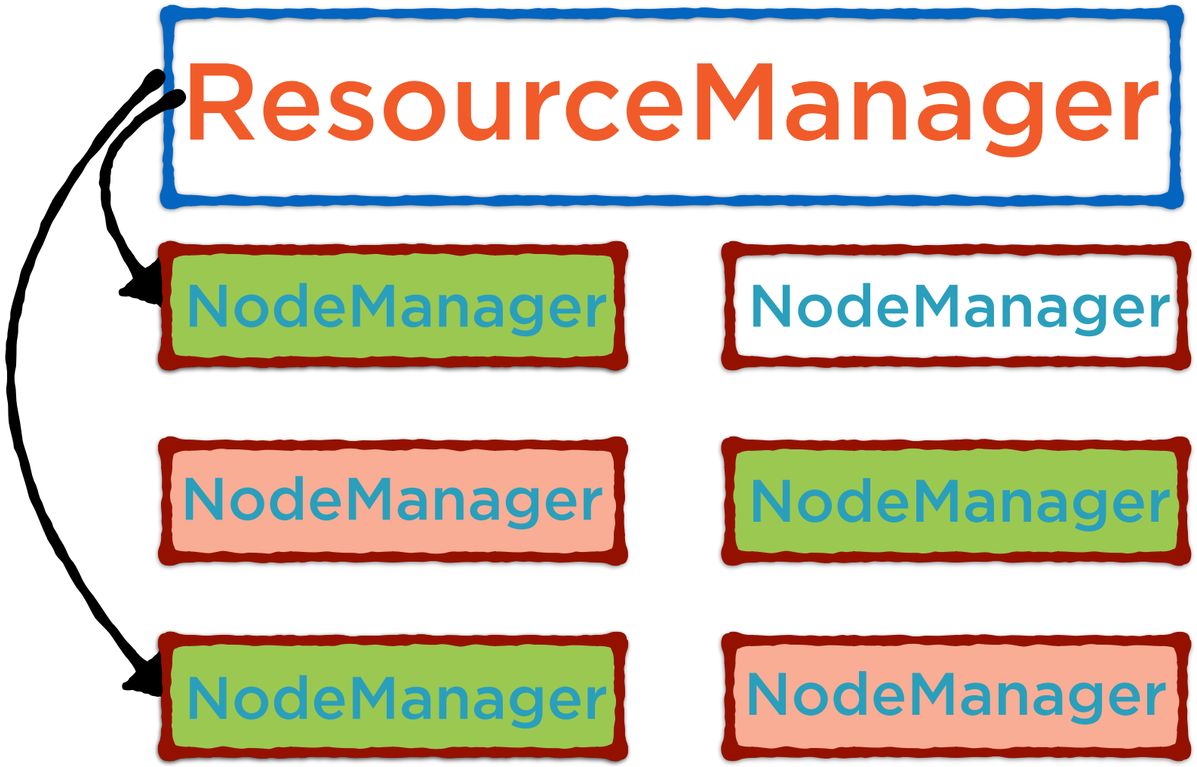
Application Processes



Requests containers for mappers and reducers

Request includes CPU, memory requirement

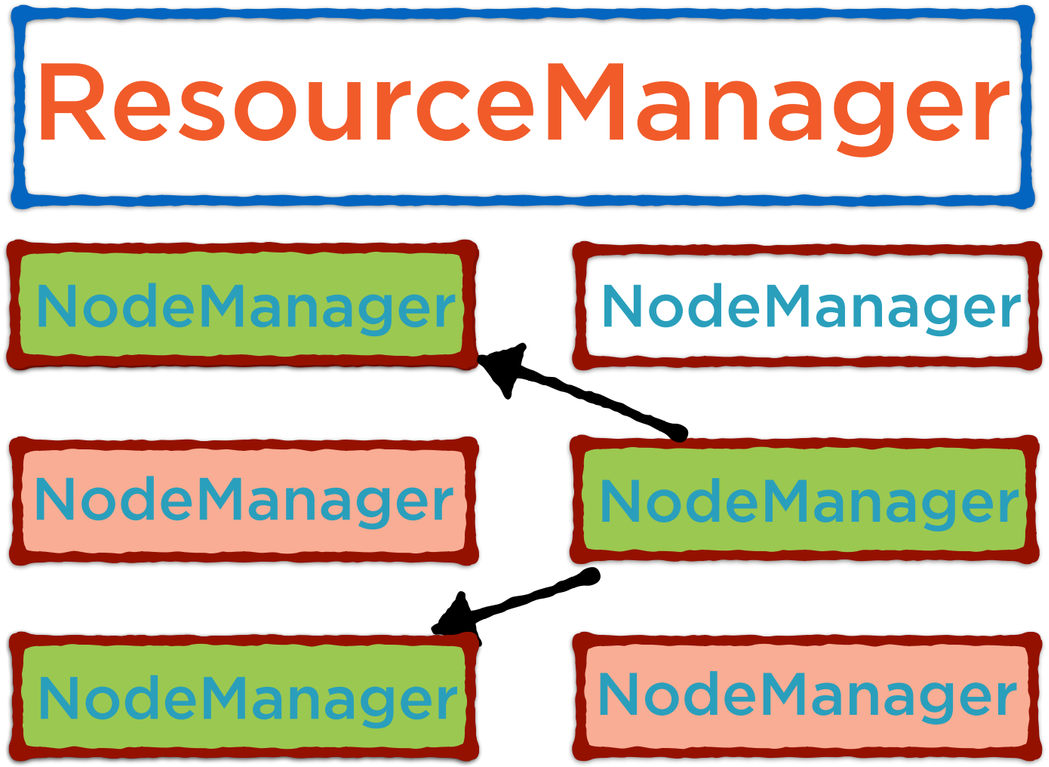
Application Processes



Assigns additional nodes

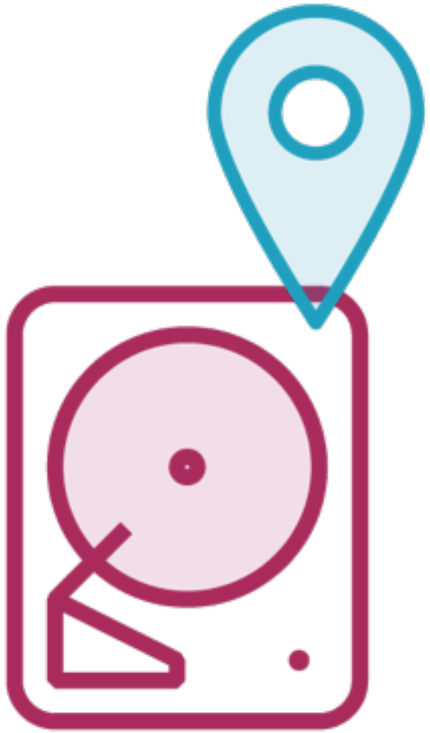
Notifies the original Application Master which made the request

Application Processes



The Application Master on the original node starts off the Application Masters on the newly assigned nodes

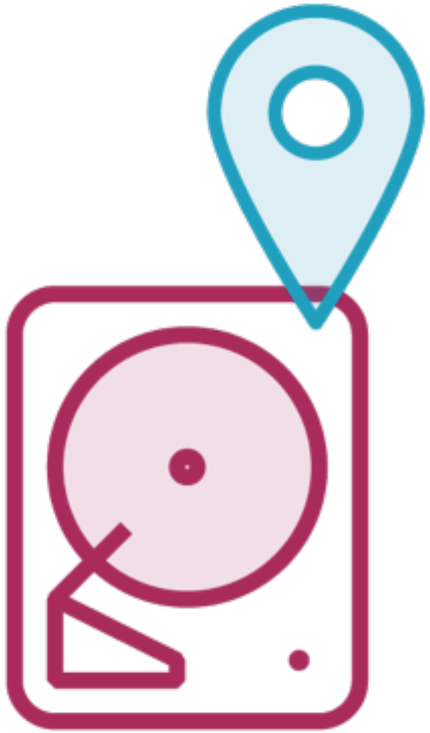
The Location Constraint



Try to minimize write bandwidth

i.e. assign a process to the same node where the data to be processed lives

The Location Constraint



If CPU, memory are not available

wait!

Scheduling Policies

FIFO Scheduler

Capacity Scheduler

Fair Scheduler

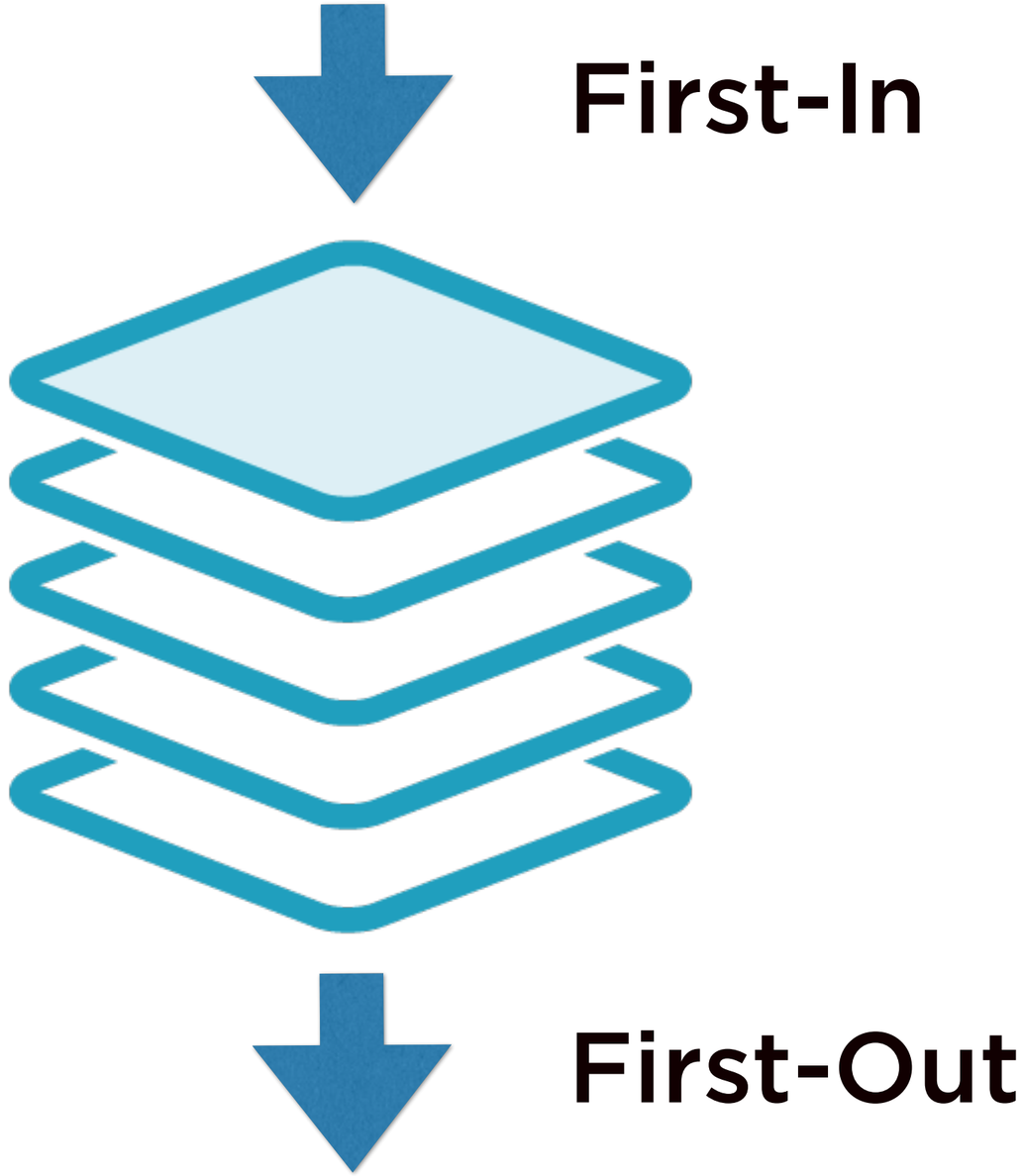
Scheduling Policies

FIFO Scheduler

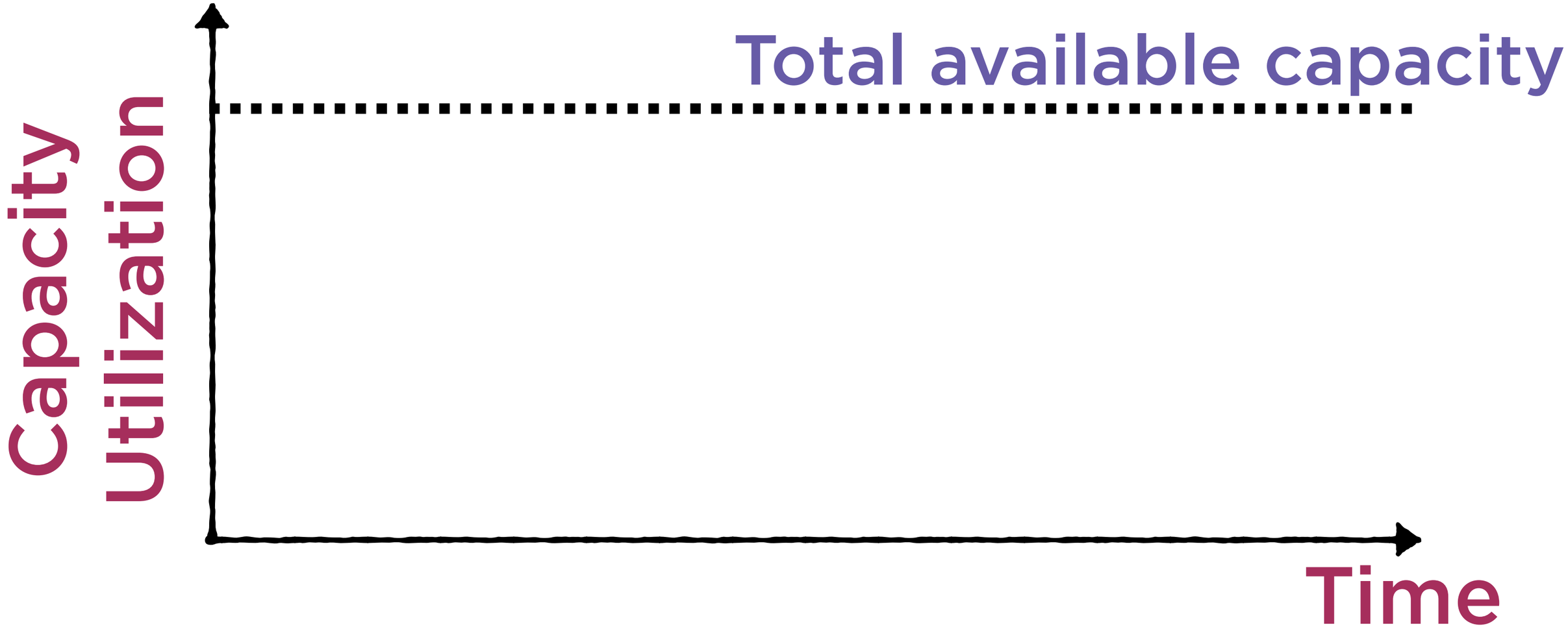
Capacity Scheduler

Fair Scheduler

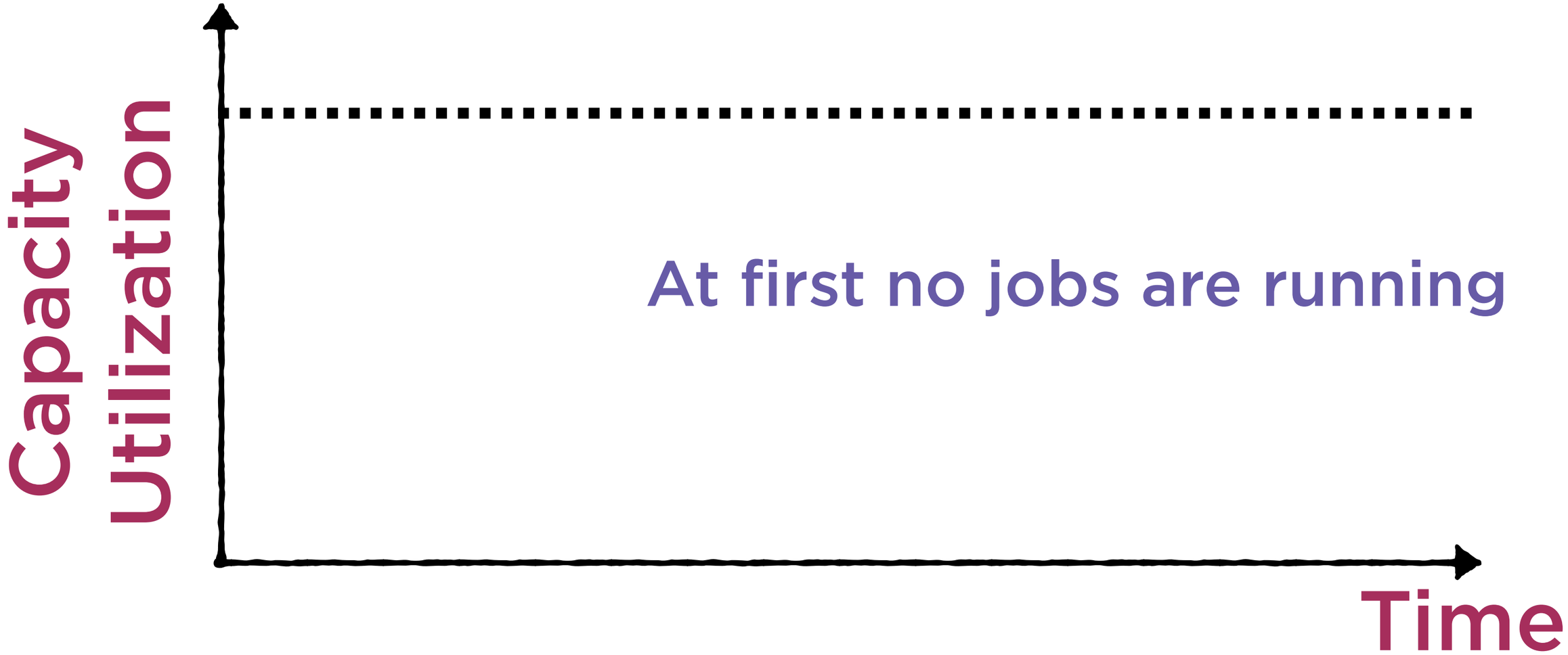
FIFO Scheduler



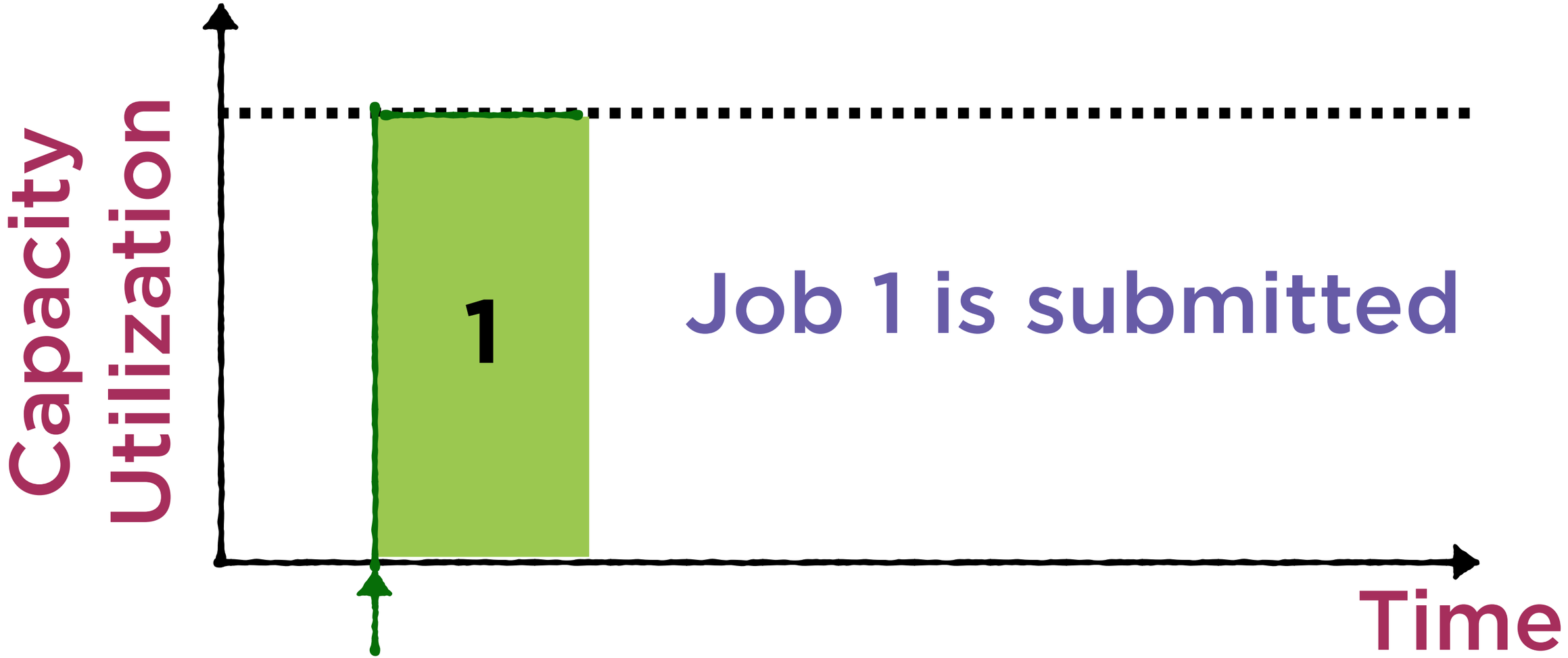
FIFO Scheduler



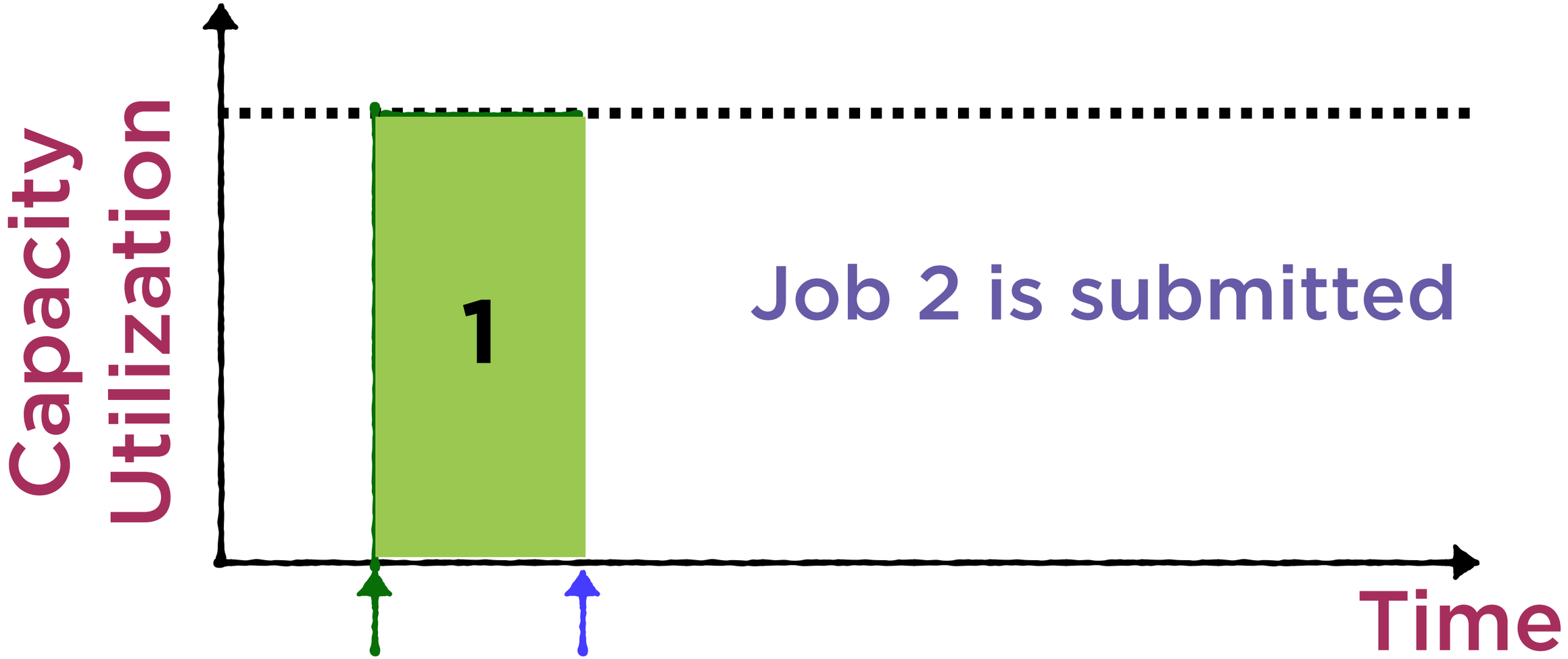
FIFO Scheduler



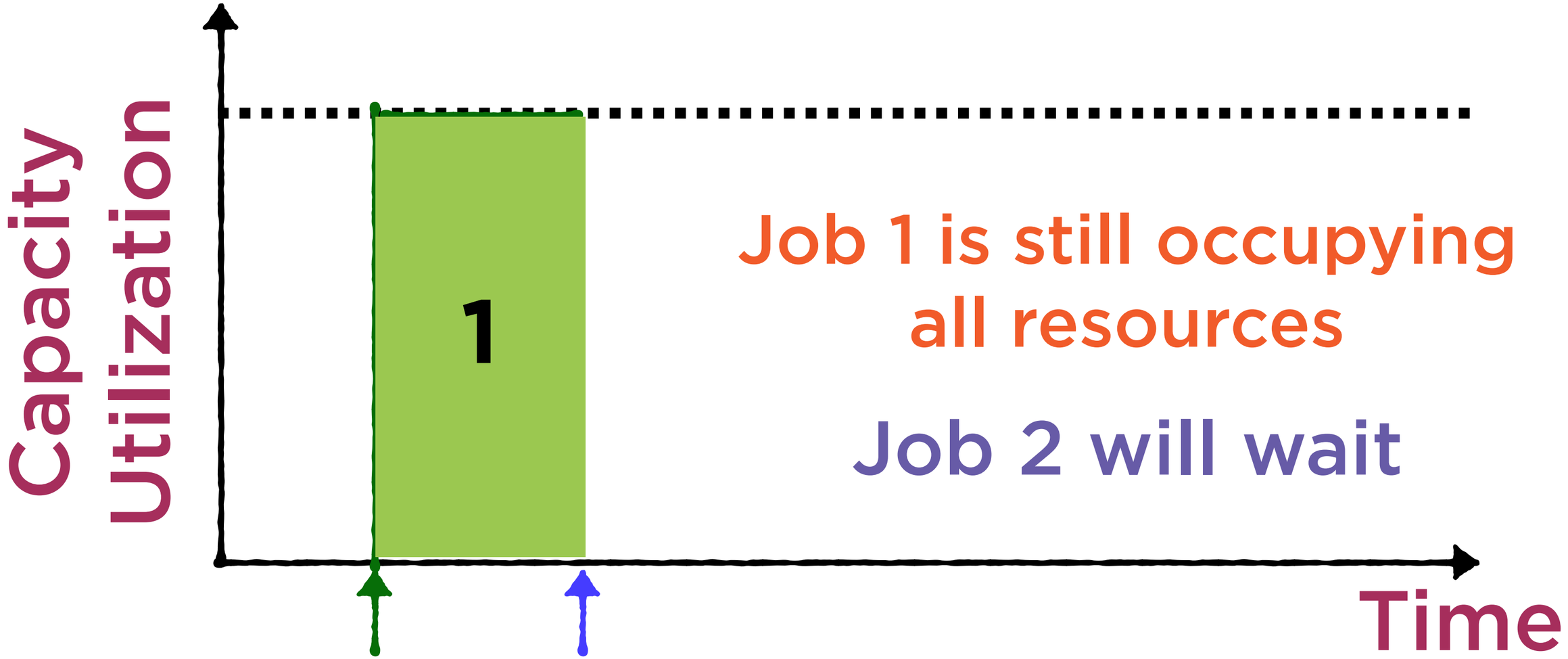
FIFO Scheduler



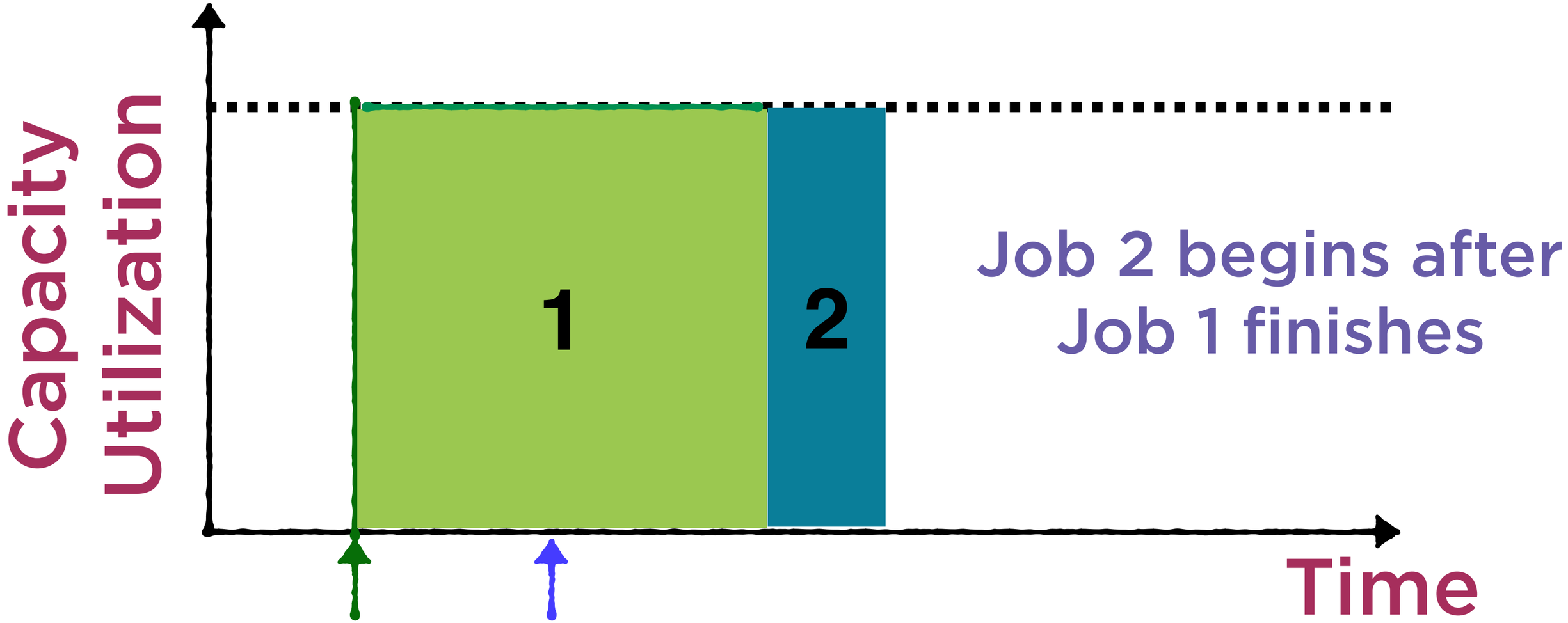
FIFO Scheduler



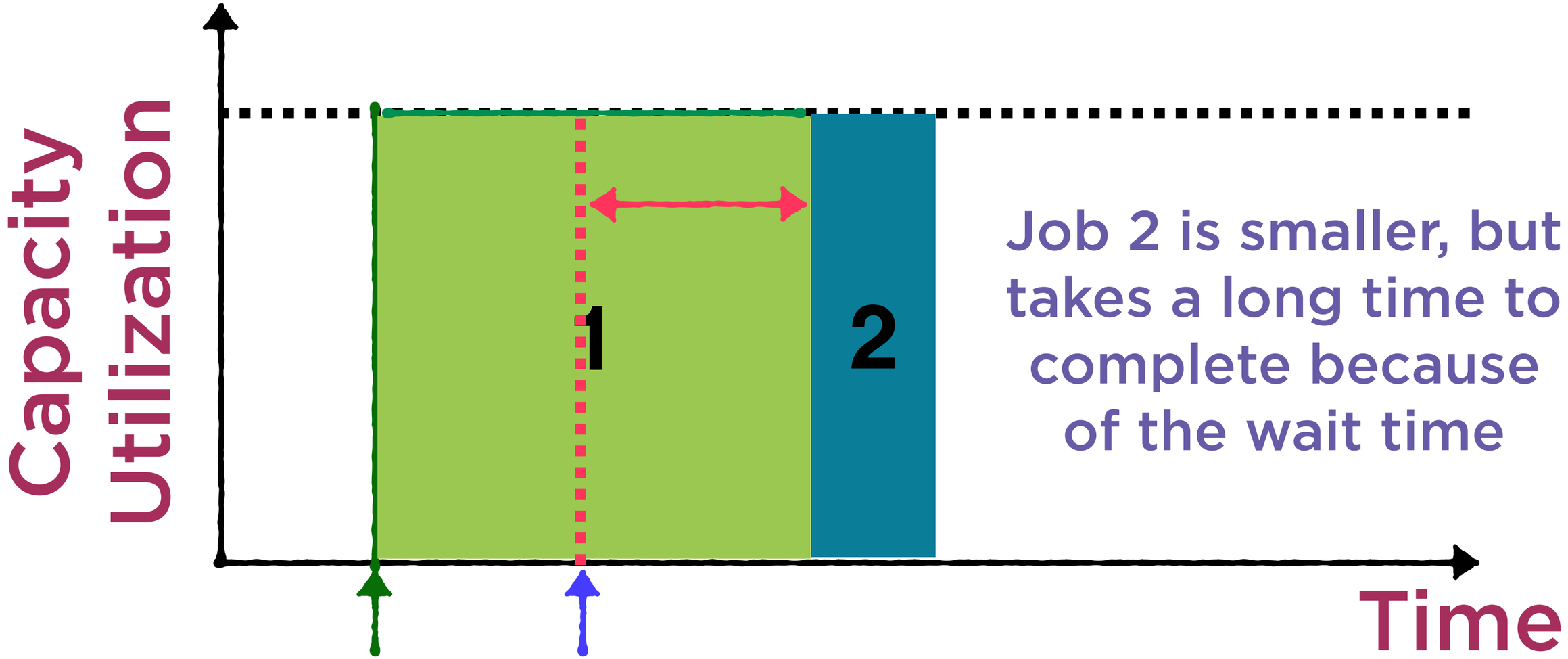
FIFO Scheduler



FIFO Scheduler



FIFO Scheduler



Job 2 is smaller, but takes a long time to complete because of the wait time

FIFO Scheduler

FIFO scheduler is rarely used

Huge wait times!

Scheduling Policies

FIFO Scheduler

Capacity Scheduler

Fair Scheduler

Scheduling Policies

FIFO Scheduler

Capacity Scheduler

Fair Scheduler

Capacity Scheduler

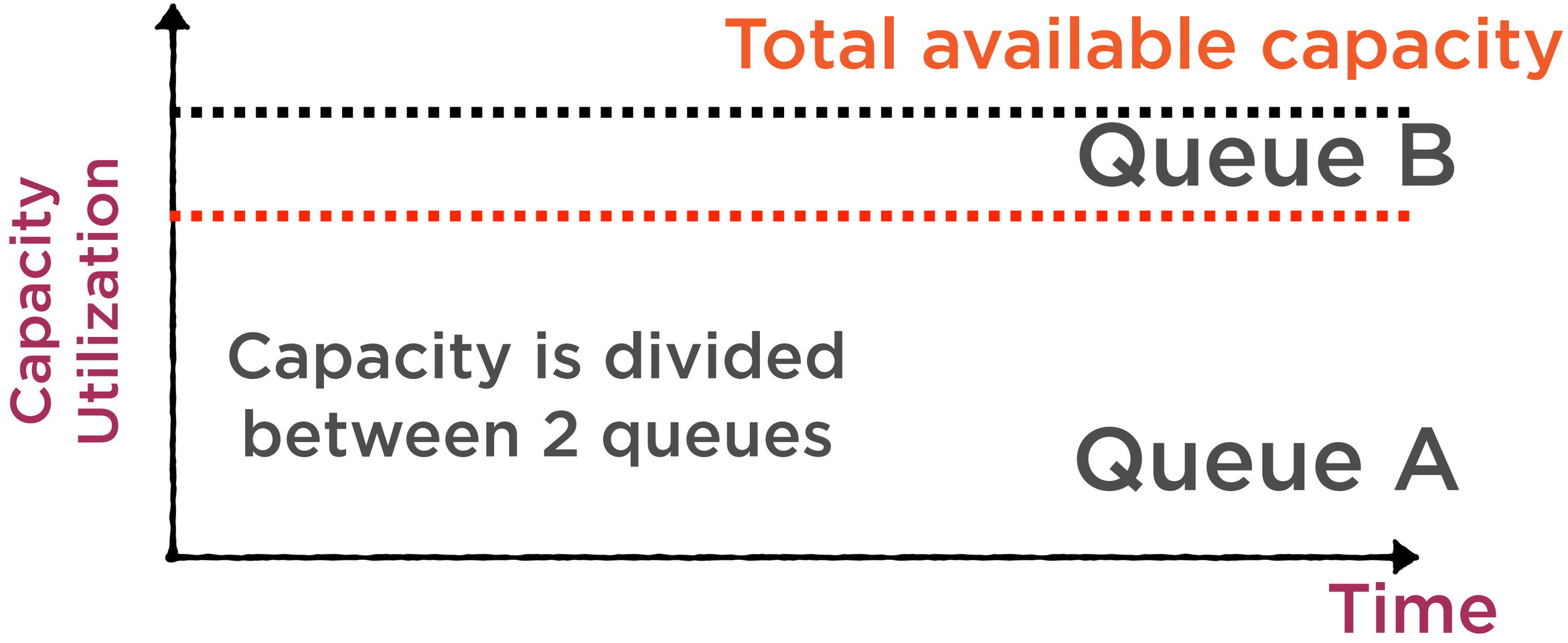
Capacity is distributed among different queues

- Each queue is allocated a share of the cluster resources

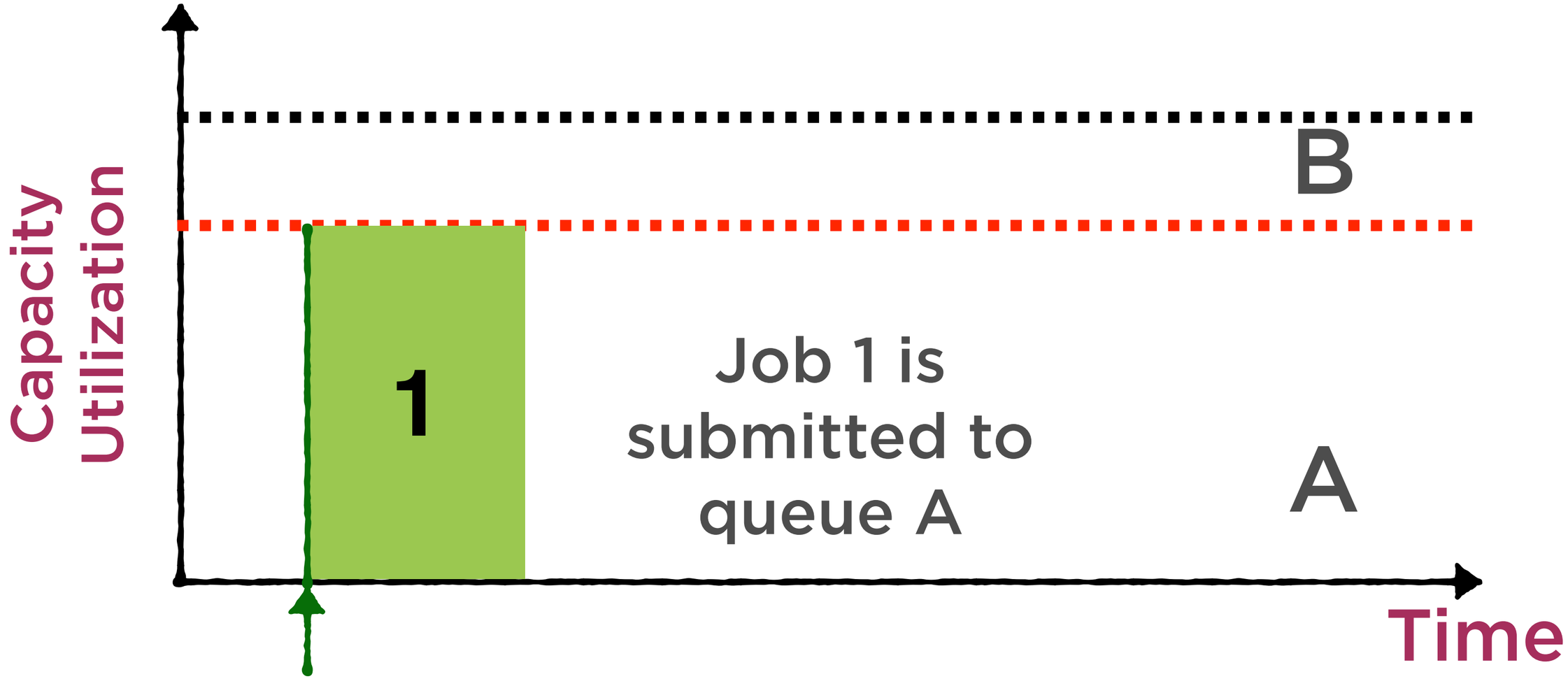
Jobs can be submitted to a specific queue

Within a queue, FIFO scheduling is used

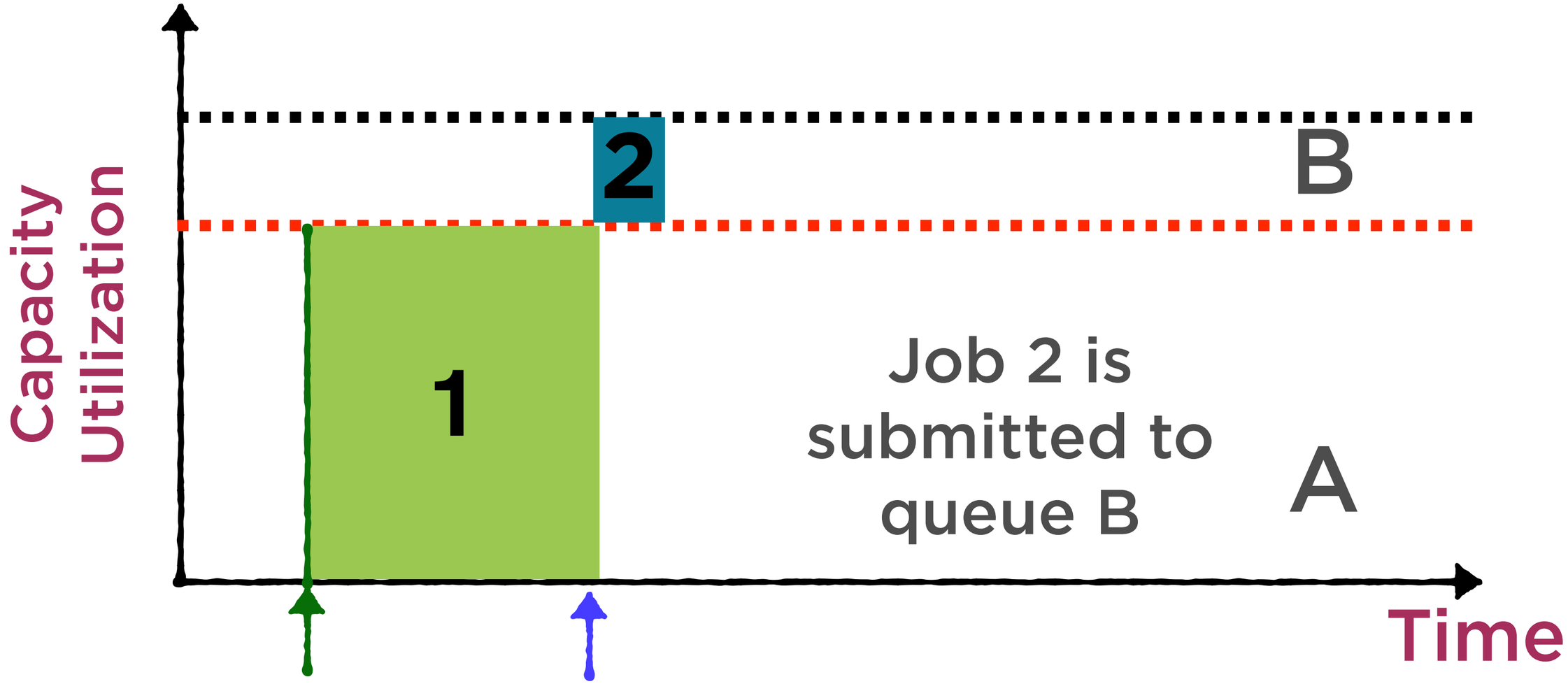
Capacity Scheduler



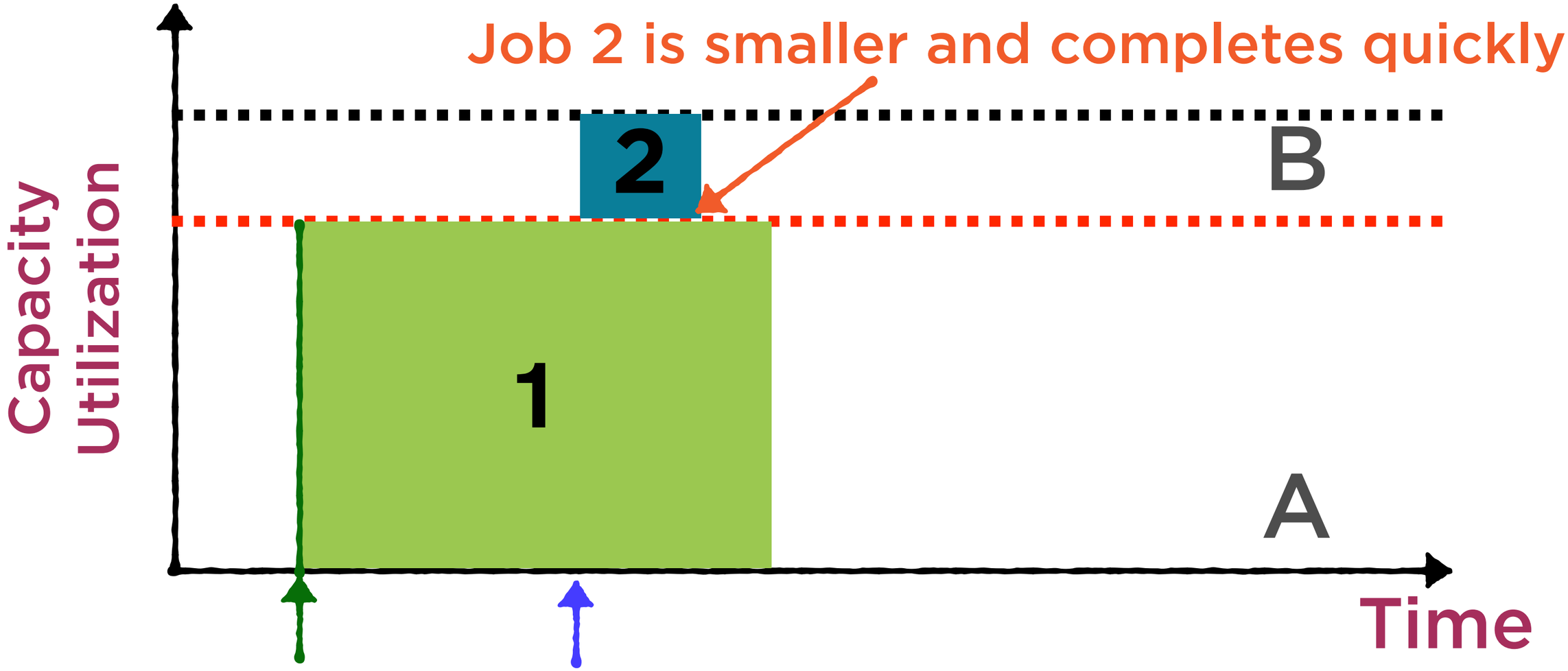
Capacity Scheduler



Capacity Scheduler



Capacity Scheduler



Capacity Scheduler

Pro

Small jobs can finish without getting stuck due to large ones

Con

The cluster might be underutilized if some queues are not occupied

Scheduling Policies

FIFO Scheduler

Capacity Scheduler

Fair Scheduler

Scheduling Policies

FIFO Scheduler

Capacity Scheduler

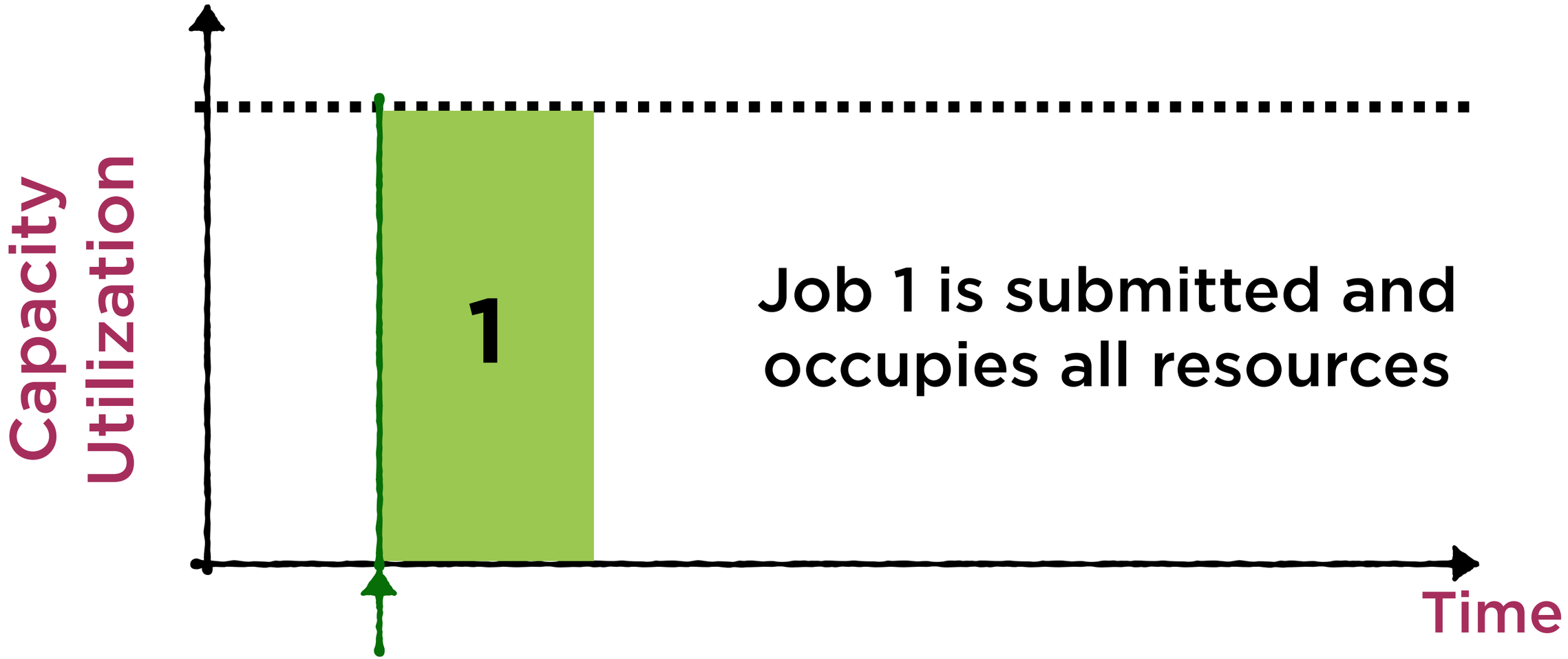
Fair Scheduler

Fair Scheduler

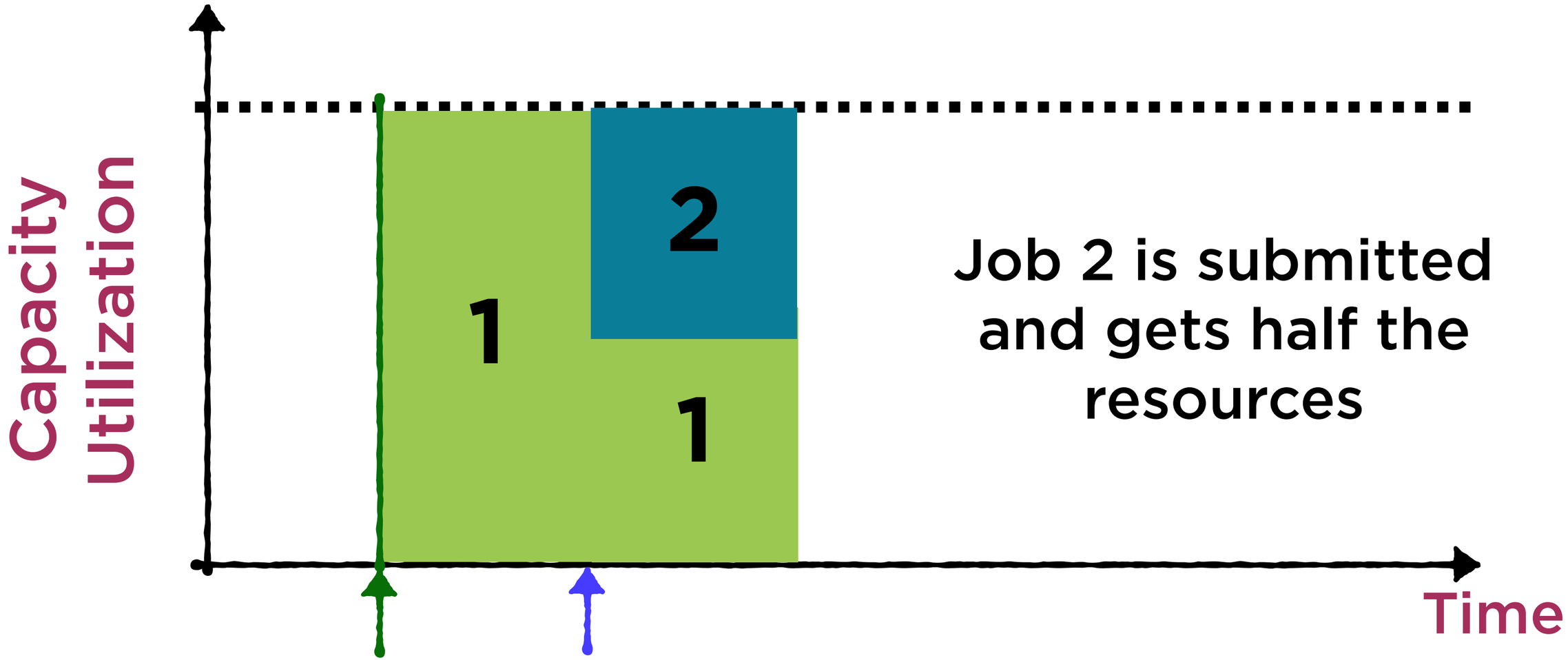
Resources are always proportionally allocated to all jobs

Zero wait time for any job

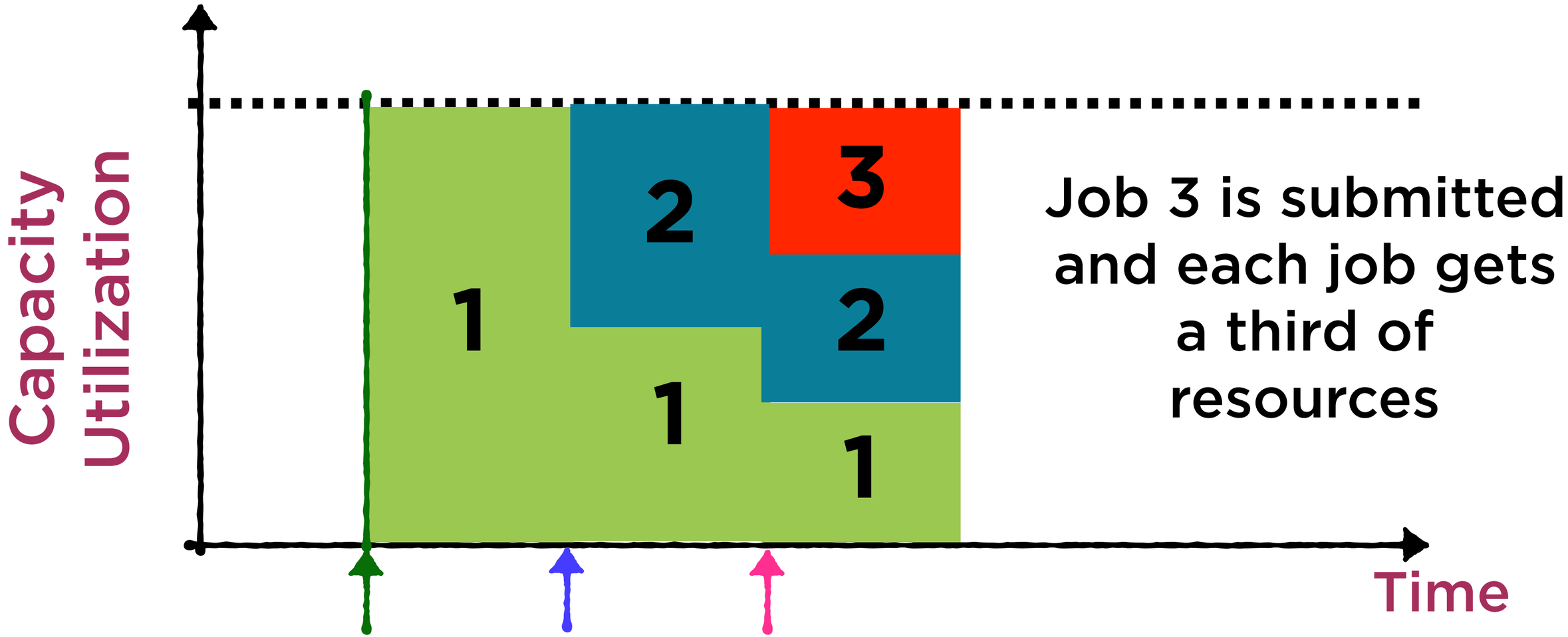
Fair Scheduler



Fair Scheduler

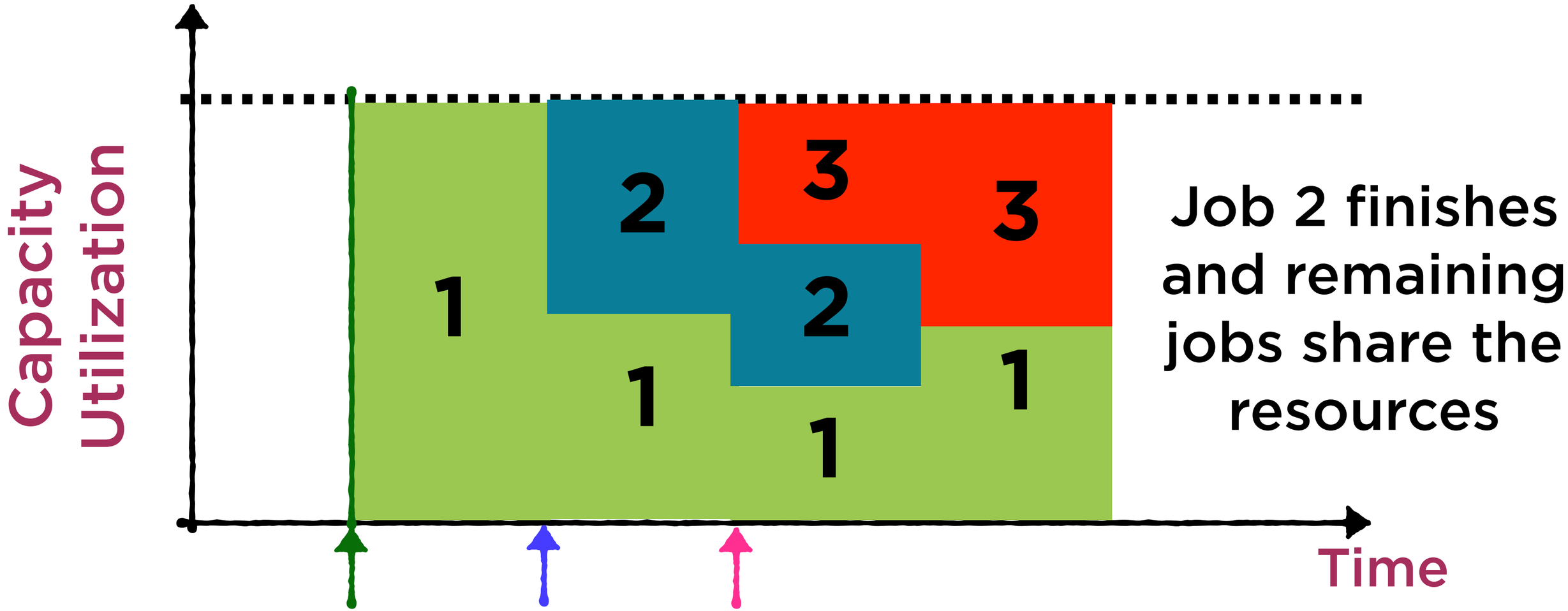


Fair Scheduler

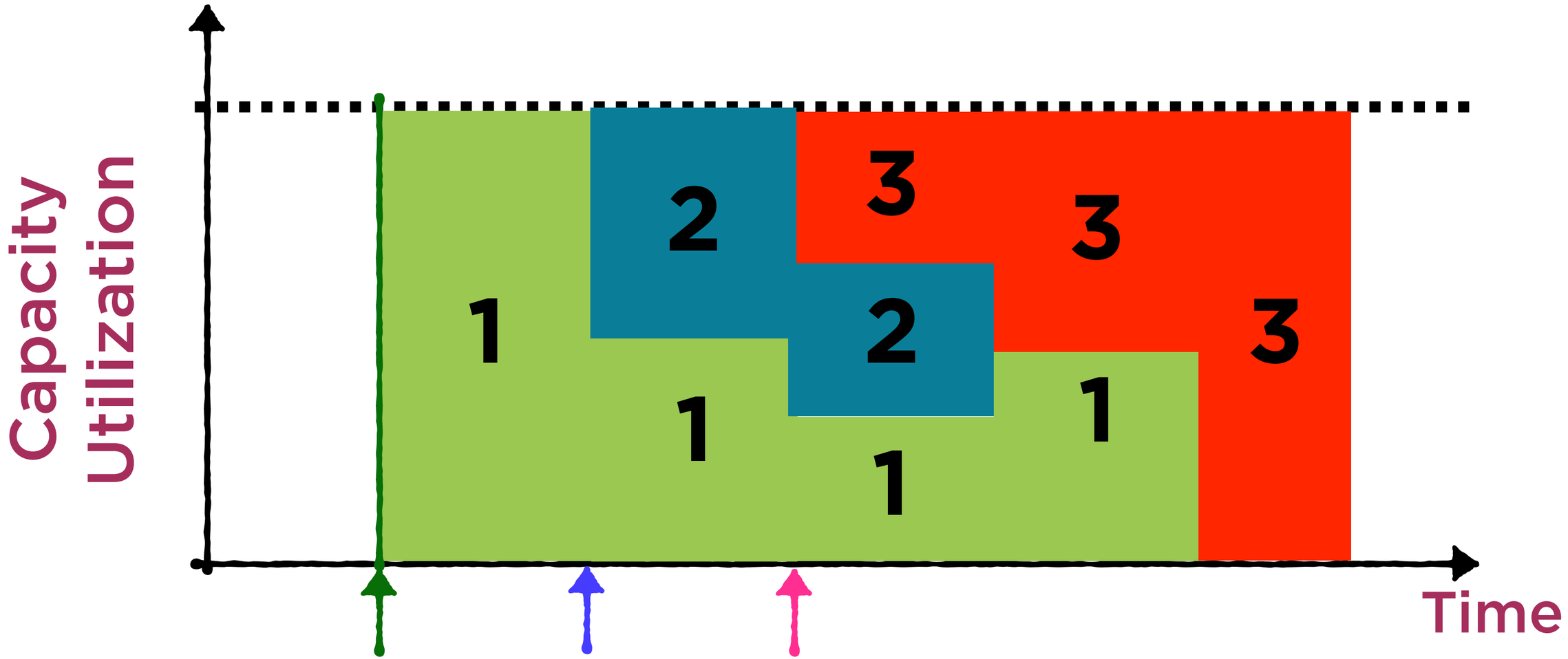


Job 3 is submitted and each job gets a third of resources

Fair Scheduler



Fair Scheduler



Scheduling Policies

FIFO Scheduler

Capacity Scheduler

Fair Scheduler

Scheduling Policies

FIFO Scheduler

Capacity Scheduler

Fair Scheduler

Configuring the Scheduling Policy

Default Policy is Capacity Scheduling

Configuring the Scheduling Policy

Set the policy in yarn-site.xml

```
<configuration>
<property>
    <name>yarn.resourcemanager.scheduler.class</name>
    <value>
    org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler
    </value>
</property>
</configuration>
```

Configuring the Fair Scheduler

The property name to set a scheduling policy

```
<configuration>
<property>
    <name>yarn.resourcemanager.scheduler.class</name>
    <value>
    org.apache.hadoop.yarn.server.resourcemanager.scheduler.fair.FairScheduler
    </value>
</property>
</configuration>
```

Configuring the Fair Scheduler

The complete path to the scheduler class in the Hadoop Java library

Demo

Define different queues for development and production work

Submit a MapReduce job to a specific queue

Monitor the job and the queue using the UI

Overview

Understood how YARN schedules tasks

Understood different scheduling strategies i.e. FIFO, Capacity and Fair Scheduler

Set a job to run on a specific queue