

What's hot on  
Twitter?



9:00 PM

#GoldenGlobes

#ThingsNerdsSay

Manchester by the Sea

#RHOA

Joey Porter

#ILaughAboutItBecause

#SherlockPBS

Stranger Things

Ivan Rabb

Beauty and the Beast

9:15 PM

#GoldenGlobes

Moonlight

#ThingsGhostsProbablyDo

#ThingsNerdsSay

#ILaughAboutItBecause

#RODEOHOUSTON

Joey Porter

Manchester by the Sea

Beauty and the Beast

Ivan Rabb

9:30 PM

Meryl Streep

Moonlight

#HiddenFences

Casey Affleck

Donald Glover

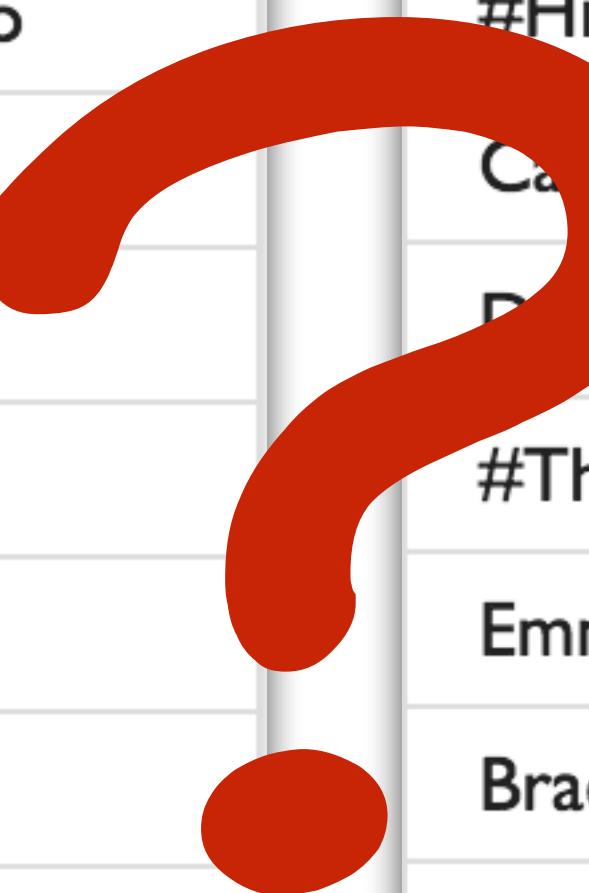
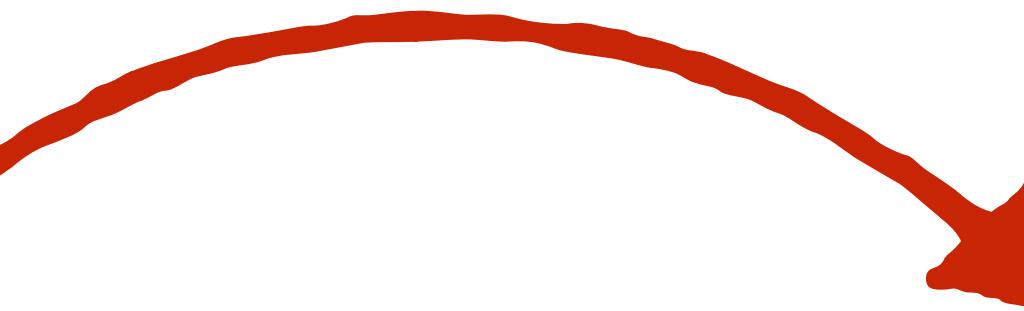
#ThingsGhostsProbablyDo

Emma Stone

Brad Pitt

#ILaughAboutItBecause

#ThingsNerdsSay



9:20 PM

9:15 PM

#GoldenGlobes

Moonlight

#ThingsGhostsProbablyDo

#ThingsNerdsSay

#ILaughAboutItBecause

#RODEOHOUSTON

Joey Porter

Manchester by the Sea

Beauty and the Beast

Ivan Rabb



9:30 PM

Meryl Streep

Moonlight

#HiddenFences

Casey Affleck

Donald Glover

#ThingsGhostsProbablyDo

Emma Stone

Brad Pitt

#ILaughAboutItBecause

#ThingsNerdsSay



9:20 PM



9:21-9:29 PM

"Take your broken heart, make it into art" -- [#CarrieFisher](#) via [#MerylStreep](#)

find someone who looks at you the way everyone looks at **meryl streep**

Can't stop watching this lady having NONE of **Meryl Streep**. 😂😂😂  
[#GoldenGlobes](#) 🏆

**Meryl Streep**'s speech at the [#GoldenGlobes](#) 🏆 is definitely one of the biggest highlights of the night.

Celebrities couldn't contain their love for **Meryl Streep** after [#GoldenGlobes](#) 🏆 speech [huff.to/2j8aj3Z](http://huff.to/2j8aj3Z)

Goodbye [#MerylStreep](#) you just lost 60 million fans. Get use to movie flops.

9:21-9:29 PM

"Take your broken heart, make it into art" -- **#CarrieFisher** via **#MerylStreep**

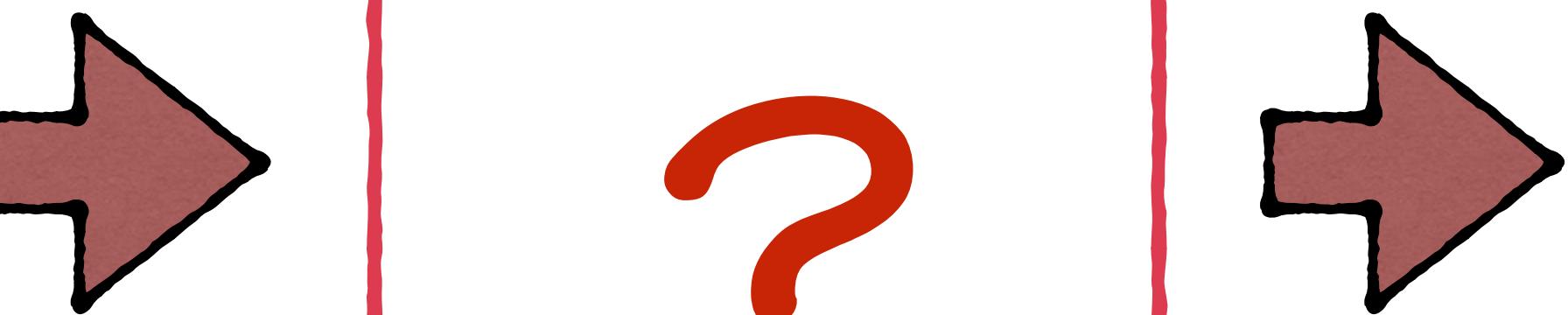
find someone who looks at you the way everyone looks at **meryl streep**

Can't stop watching this lady having NONE of **Meryl Streep**. 😂😂😂  
**#GoldenGlobes** 🎩

**Meryl Streep**'s speech at the **#GoldenGlobes** 🎩 is definitely one of the biggest highlights of the night.

Celebrities couldn't contain their love for **Meryl Streep** after **#GoldenGlobes** 🎩 speech [huff.to/2j8aj3Z](http://huff.to/2j8aj3Z)

Goodbye **#MerylStreep** you just lost 60 million fans. Get use to movie flops.



9:30 PM

**Meryl Streep**

**Moonlight**

**#HiddenFences**

**Casey Affleck**

**Donald Glover**

**#ThingsGhostsProbablyDo**

**Emma Stone**

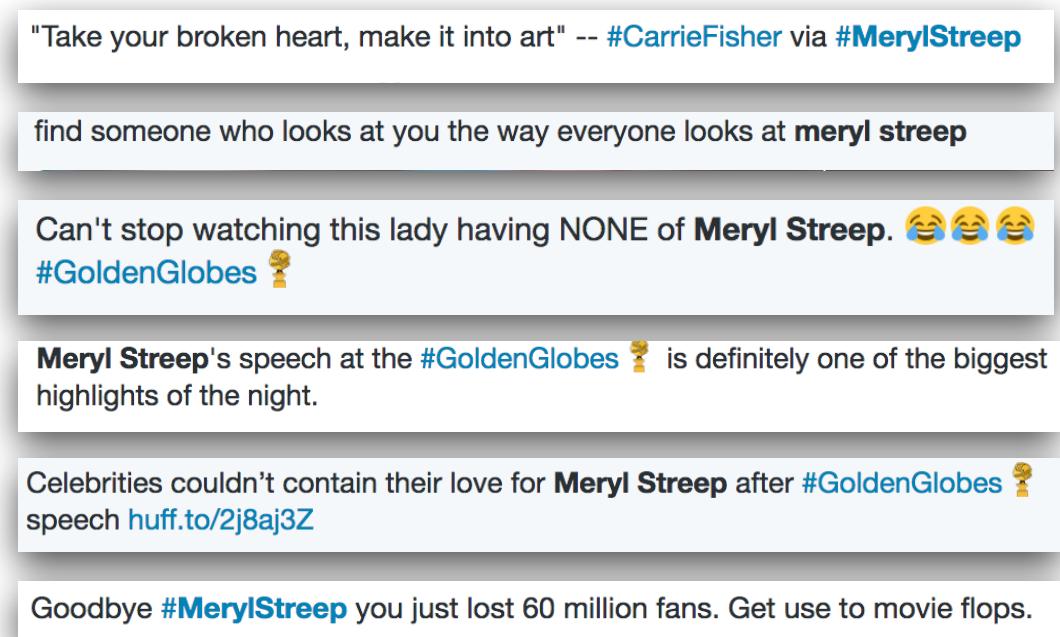
**Brad Pitt**

**#ILaughAboutItBecause**

**#ThingsNerdsSay**

Top keywords  
in last 10 mins

9:21-9:29 PM



## Batch processing

Process all tweets in the last 10 mins at a time

## Stream processing

Process each tweet as it comes in



- Start at 9:00 PM
- Collect tweets for 10 mins
- Process tweets for 2 mins
- First trends at 9:12 PM
- Next update at 9:24 PM

## Batch processing

Process all tweets in the last 10 mins at a time

Stream processing

Process each tweet as it comes in



- Start at 9:00 PM
- Collect tweets for 10 mins
- First trends at 9:10 PM
- As each tweet comes in, trends are updated

Batch processing

Process all tweets in the  
last 10 mins at a time

Stream processing

Process each tweet  
as it comes in

# Stream processing

Get the most recent (real-time) view

React to events immediately

Data  
Source



Events (Tweets)

Stream  
Processing  
App

P7

P6

P5

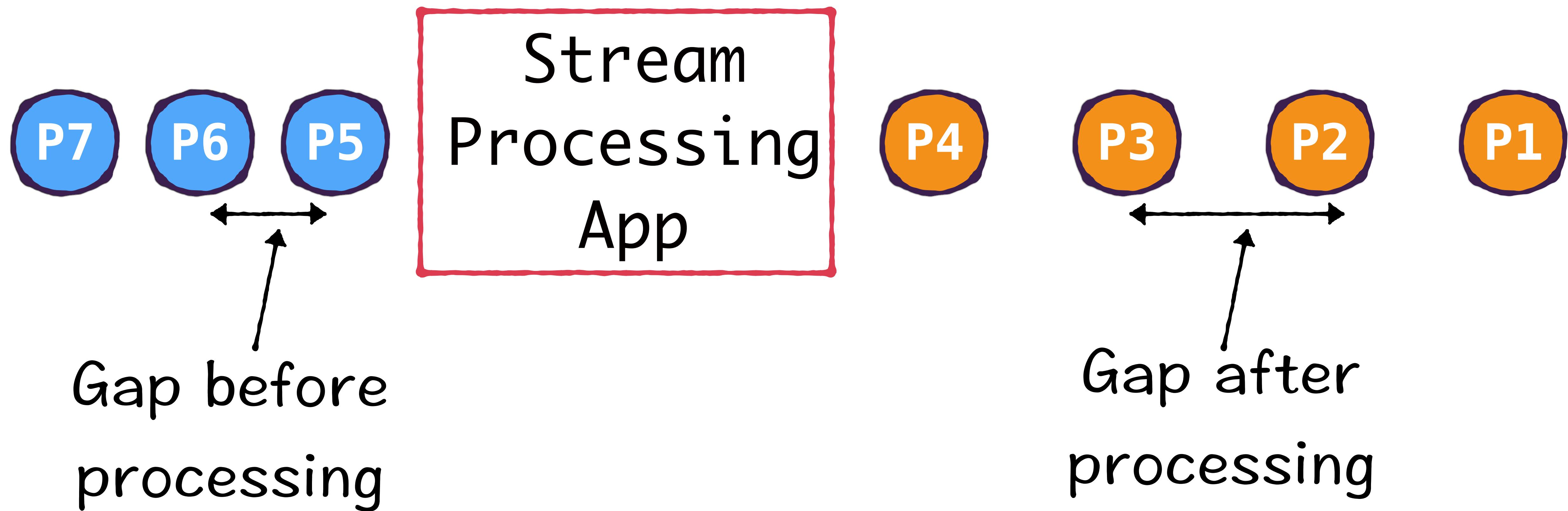
P4

P3

P2

P1

# Gap increased during processing time



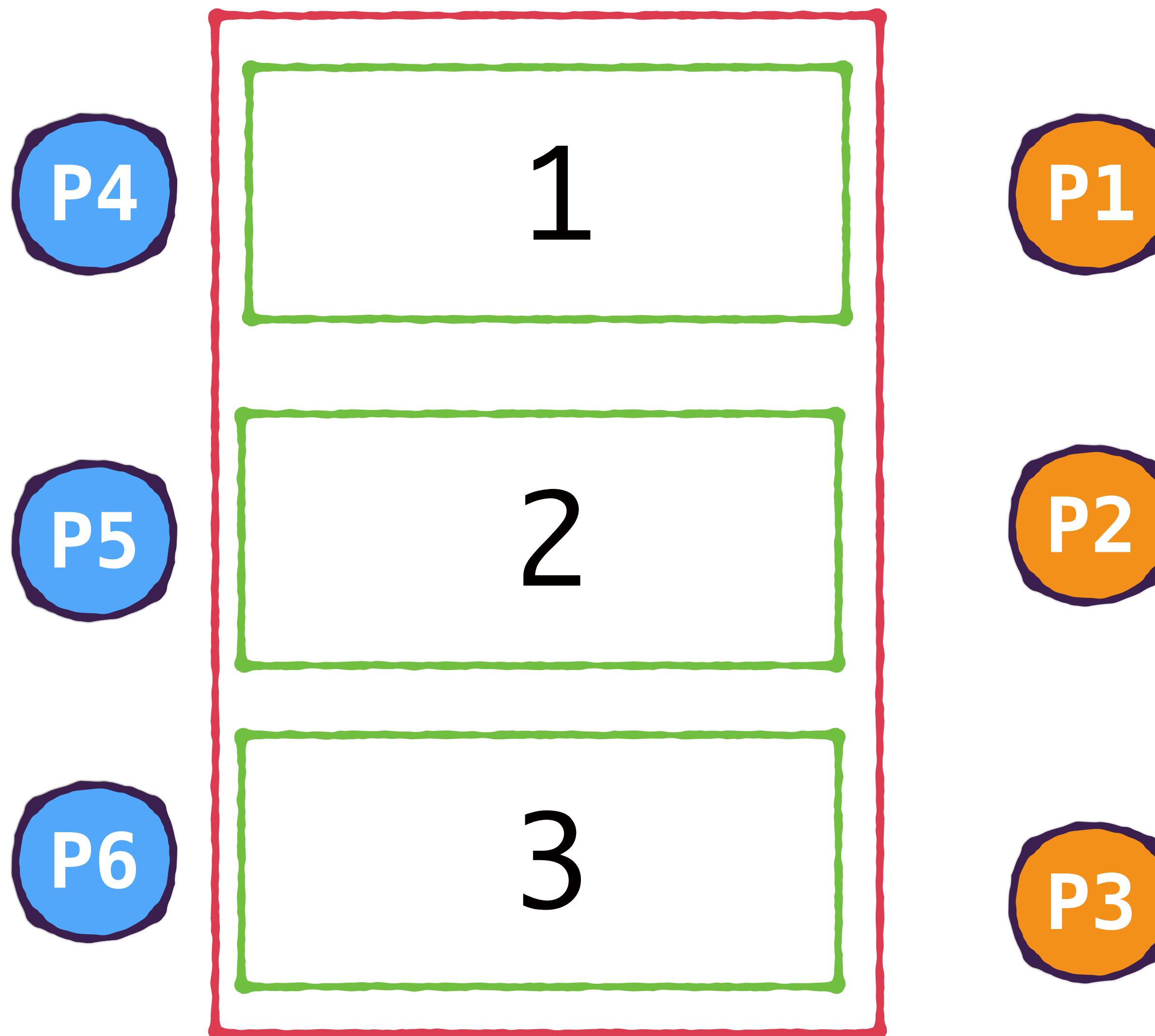
If processing time >> event arrival rate



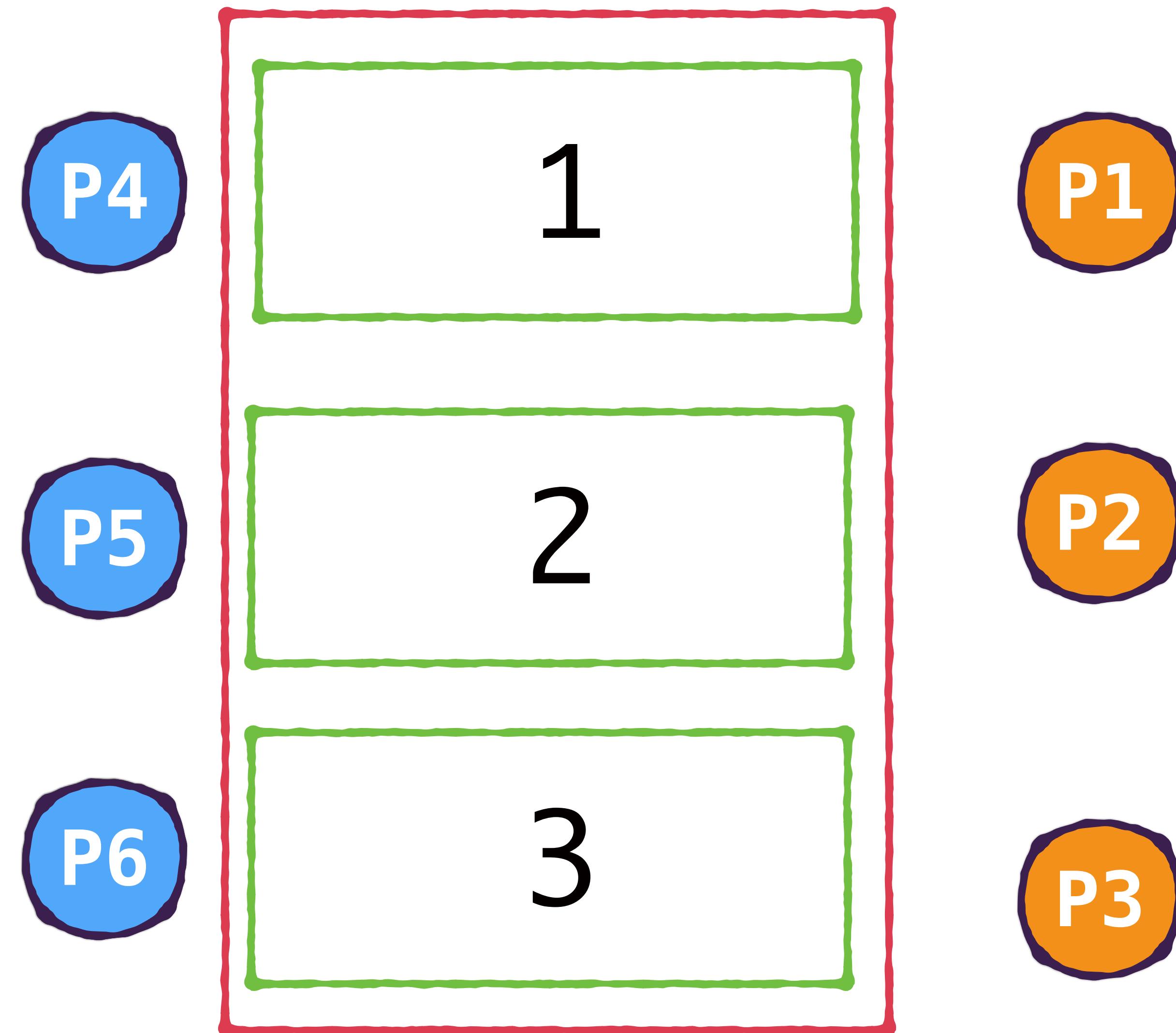
# How do we reduce processing time?



# Distribute the stream to multiple components

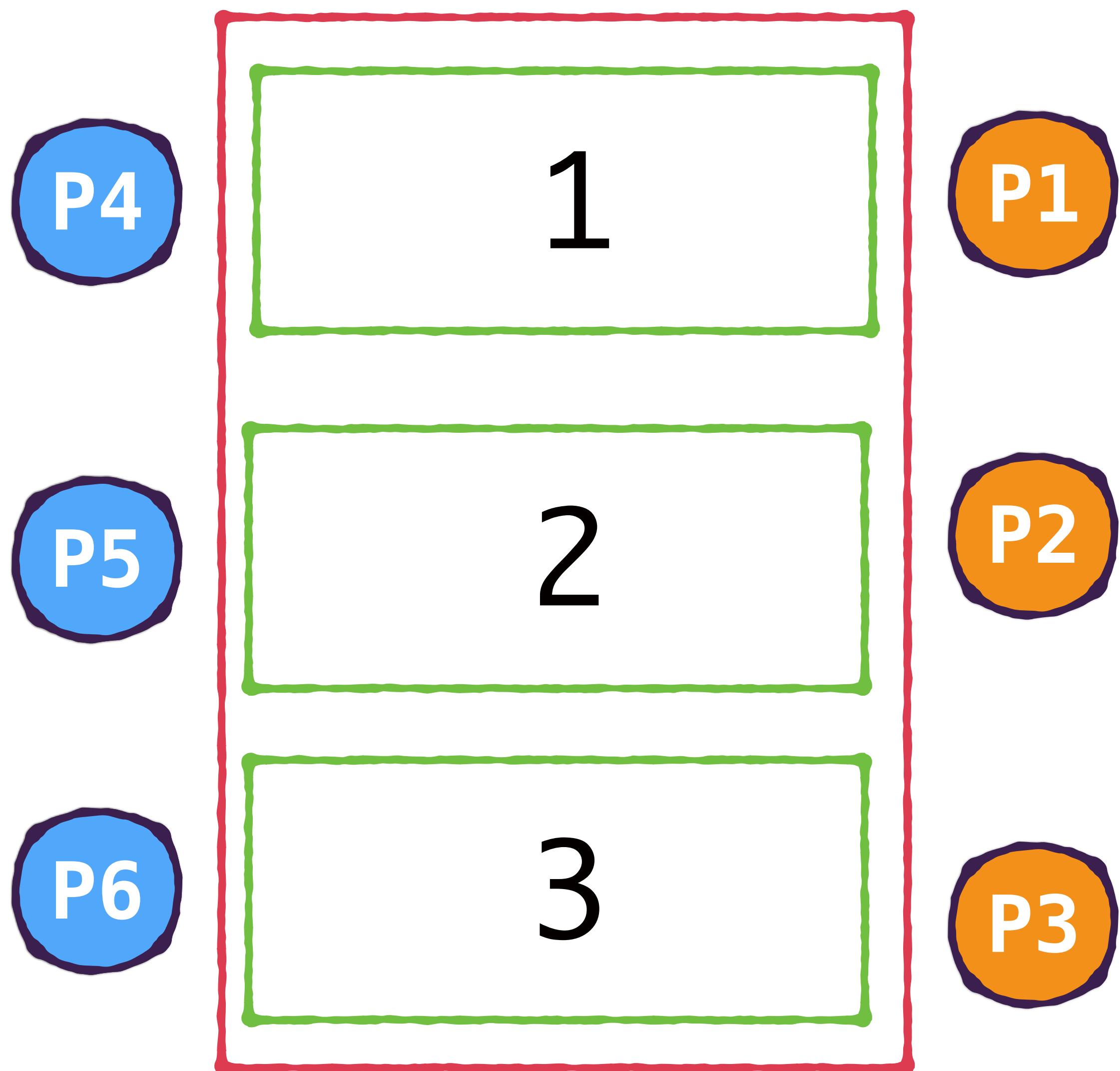


# Apache Storm



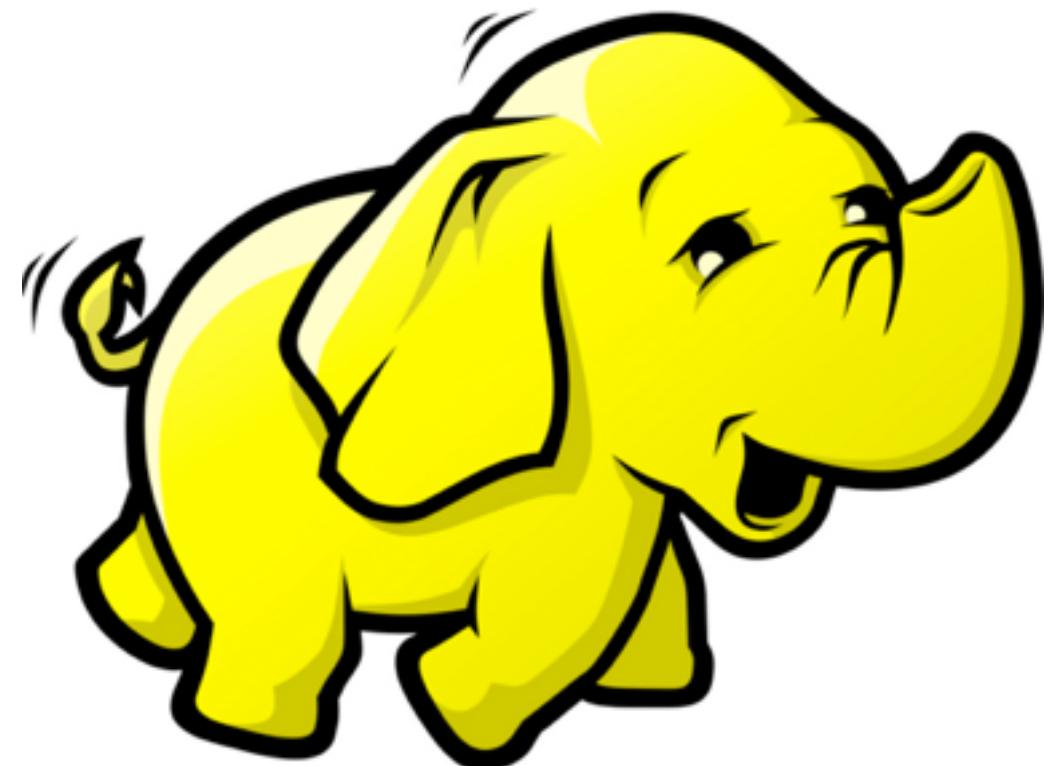
# Apache Storm

A framework for  
distributed  
stream processing



# Hadoop & MapReduce

A framework for  
distributed  
batch processing

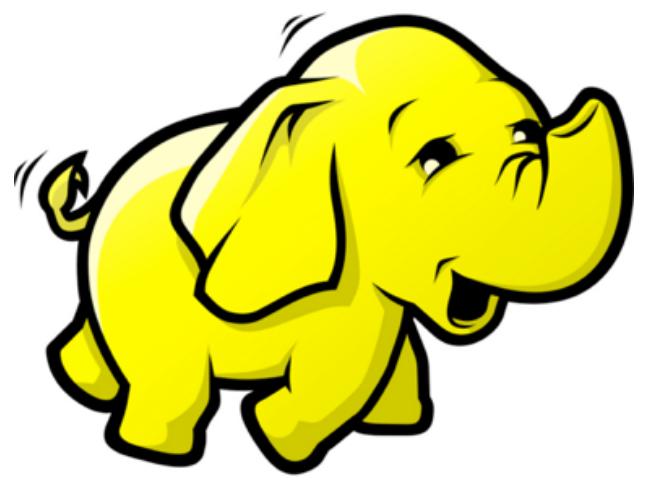


# Apache Storm

A framework for  
distributed  
stream processing



Hadoop &  
MapReduce



batch  
processing

Spark  
Streaming



batch +  
stream  
processing

Apache  
Storm

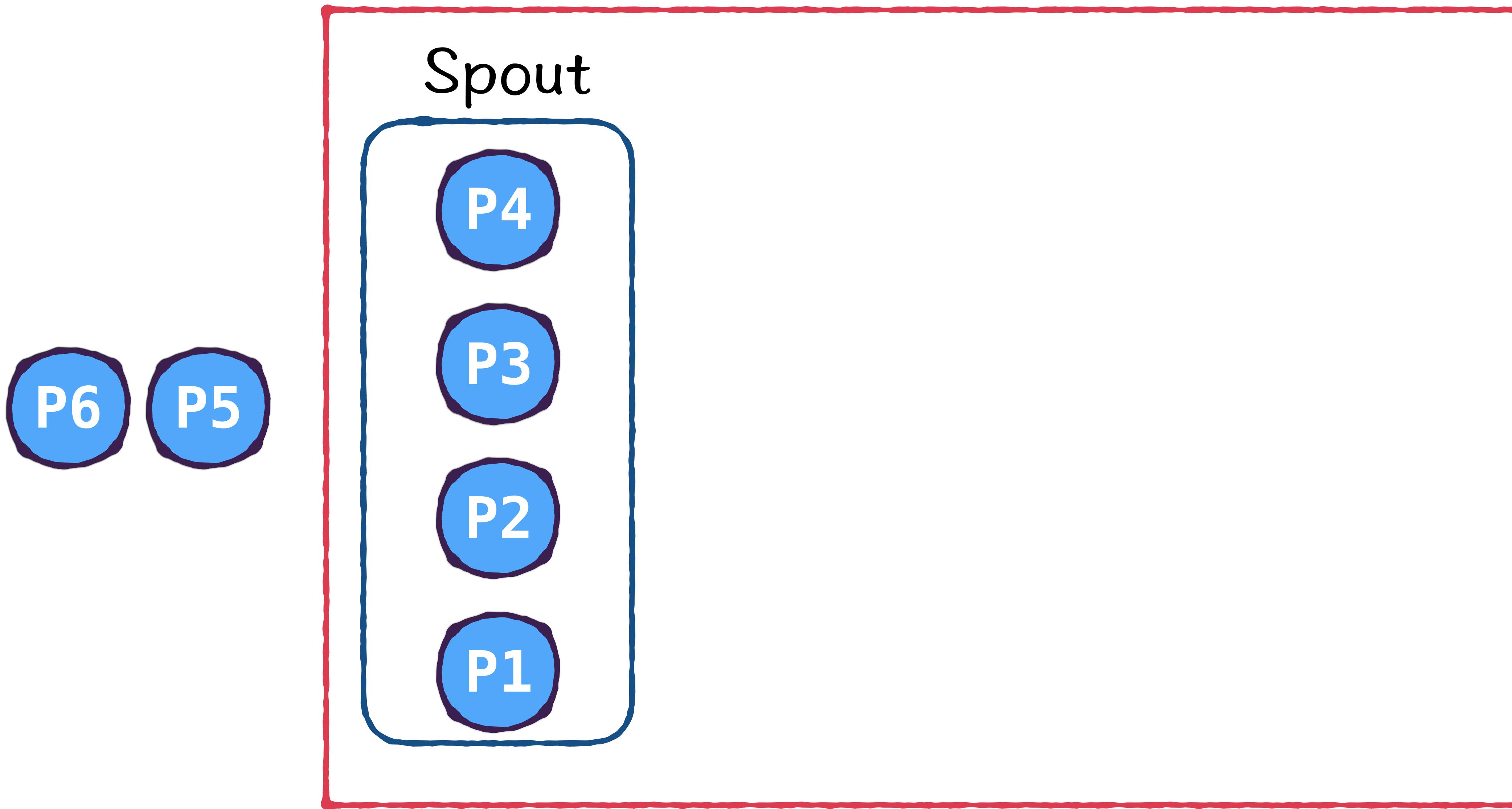


stream  
processing

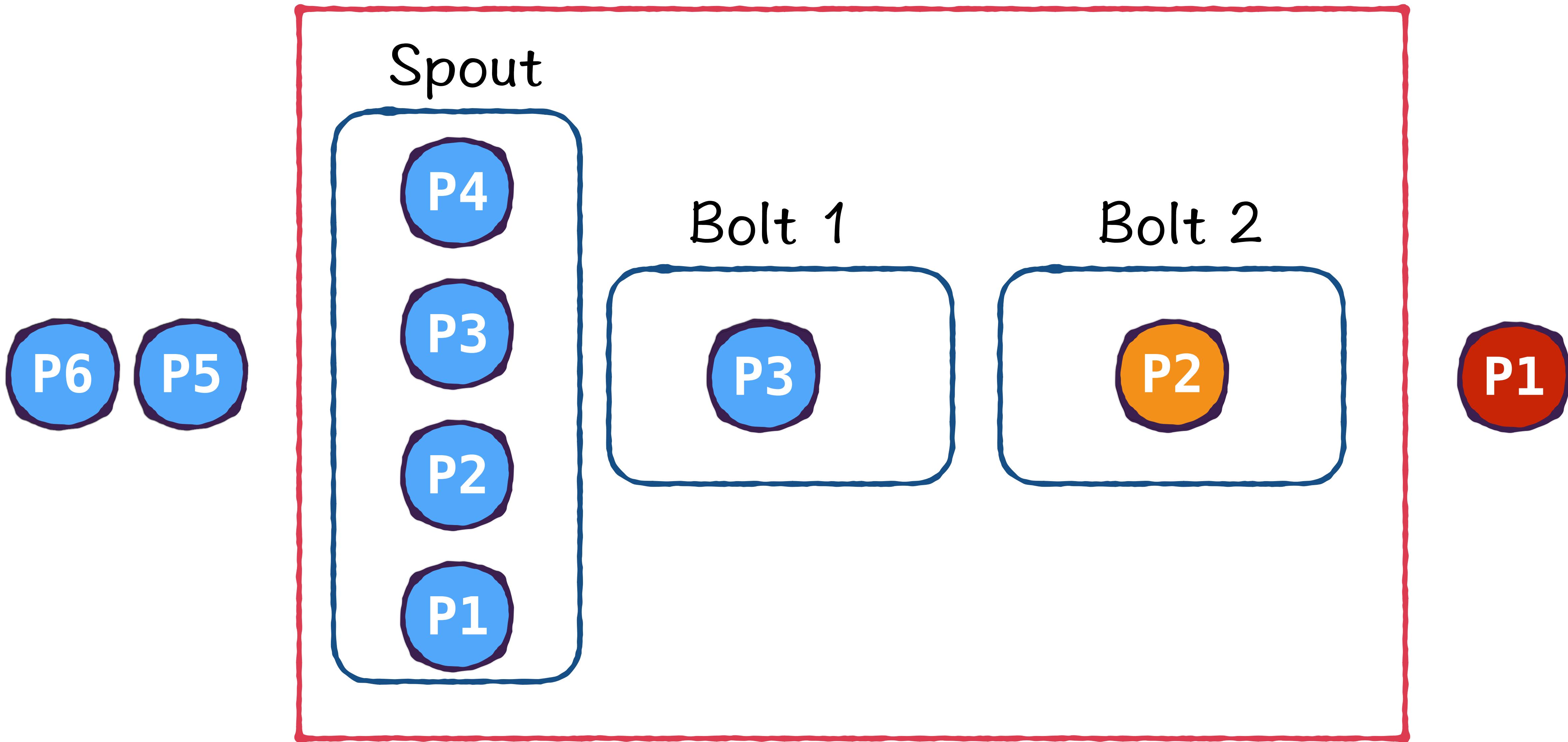
# Apache Storm



# Storm Topology



# Storm Topology



9:21-9:29 PM

"Take your broken heart, make it into art" -- **#CarrieFisher** via **#MerylStreep**

find someone who looks at you the way everyone looks at **meryl streep**

Can't stop watching this lady having NONE of **Meryl Streep**. 😂😂😂  
**#GoldenGlobes** 🎩

**Meryl Streep**'s speech at the **#GoldenGlobes** 🎩 is definitely one of the biggest highlights of the night.

Celebrities couldn't contain their love for **Meryl Streep** after **#GoldenGlobes** 🎩 speech [huff.to/2j8aj3Z](http://huff.to/2j8aj3Z)

Goodbye **#MerylStreep** you just lost 60 million fans. Get use to movie flops.



9:30 PM

**Meryl Streep**

**Moonlight**

**#HiddenFences**

**Casey Affleck**

**Donald Glover**

**#ThingsGhostsProbablyDo**

**Emma Stone**

**Brad Pitt**

**#ILaughAboutItBecause**

**#ThingsNerdsSay**

Top keywords  
in last 10 mins

# App

→ Split tweets into words

"Take your broken heart, make it into art" -- #CarrieFisher via #MerylStreep

# App

GoldenGlobes 10000

MerylStreep 15000

CarrieFisher 2000

- Split tweets into words
- Find counts for each word in last 10 mins

# App

GoldenGlobes 10000

MerylStreep 15000

CarrieFisher 2000

- Split tweets into words
- Find counts for each word in last 10 mins
- Sort words by count

# App

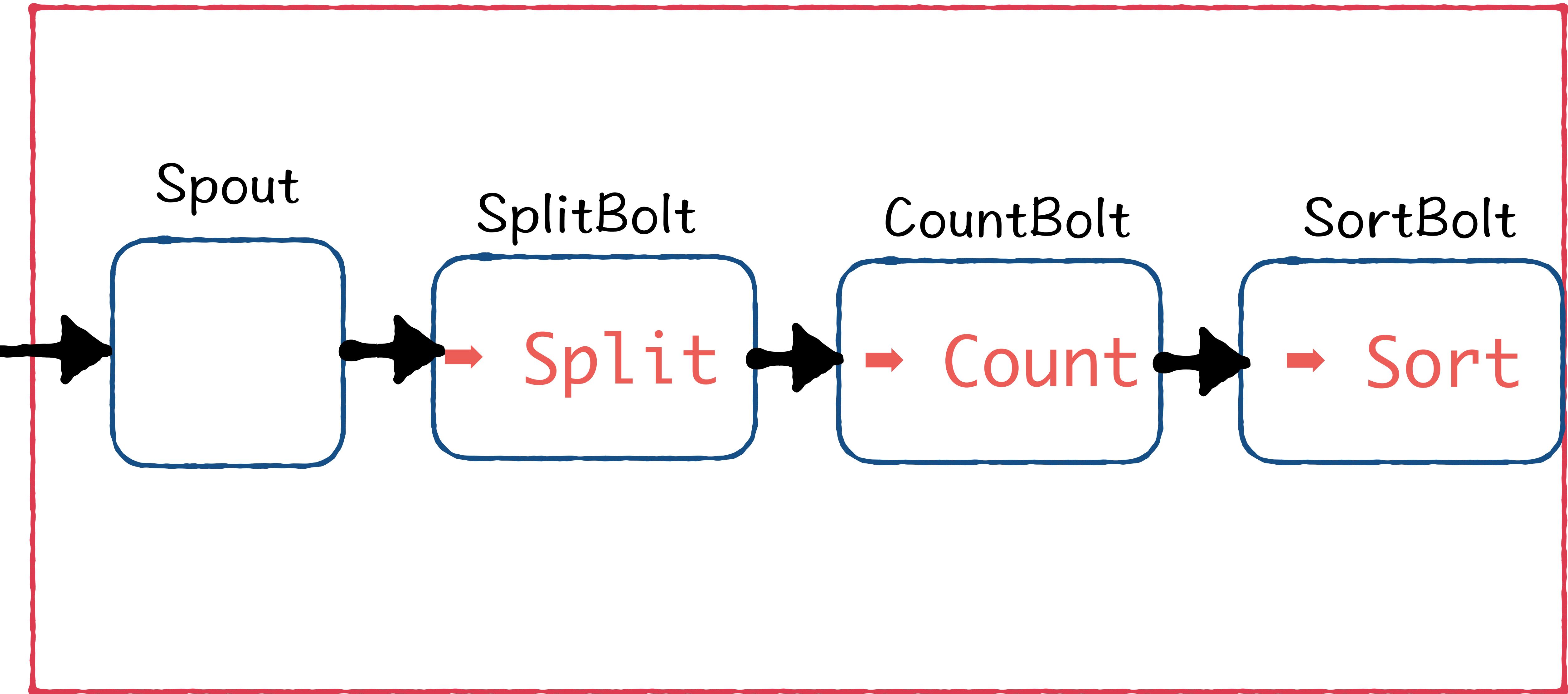
MerylStreep 15000

GoldenGlobes 10000

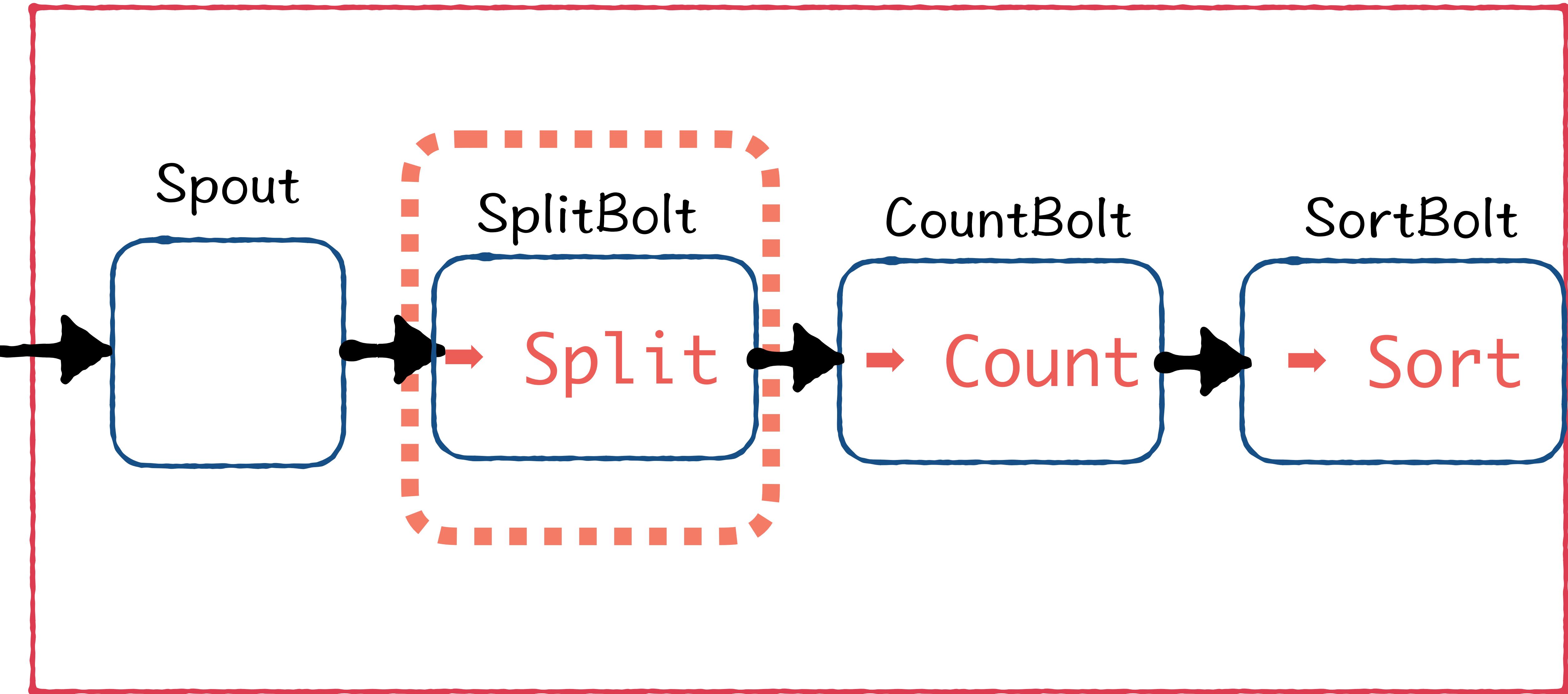
CarrieFisher 2000

- Split tweets into words
- Find counts for each word in last 10 mins
- Sort words by count

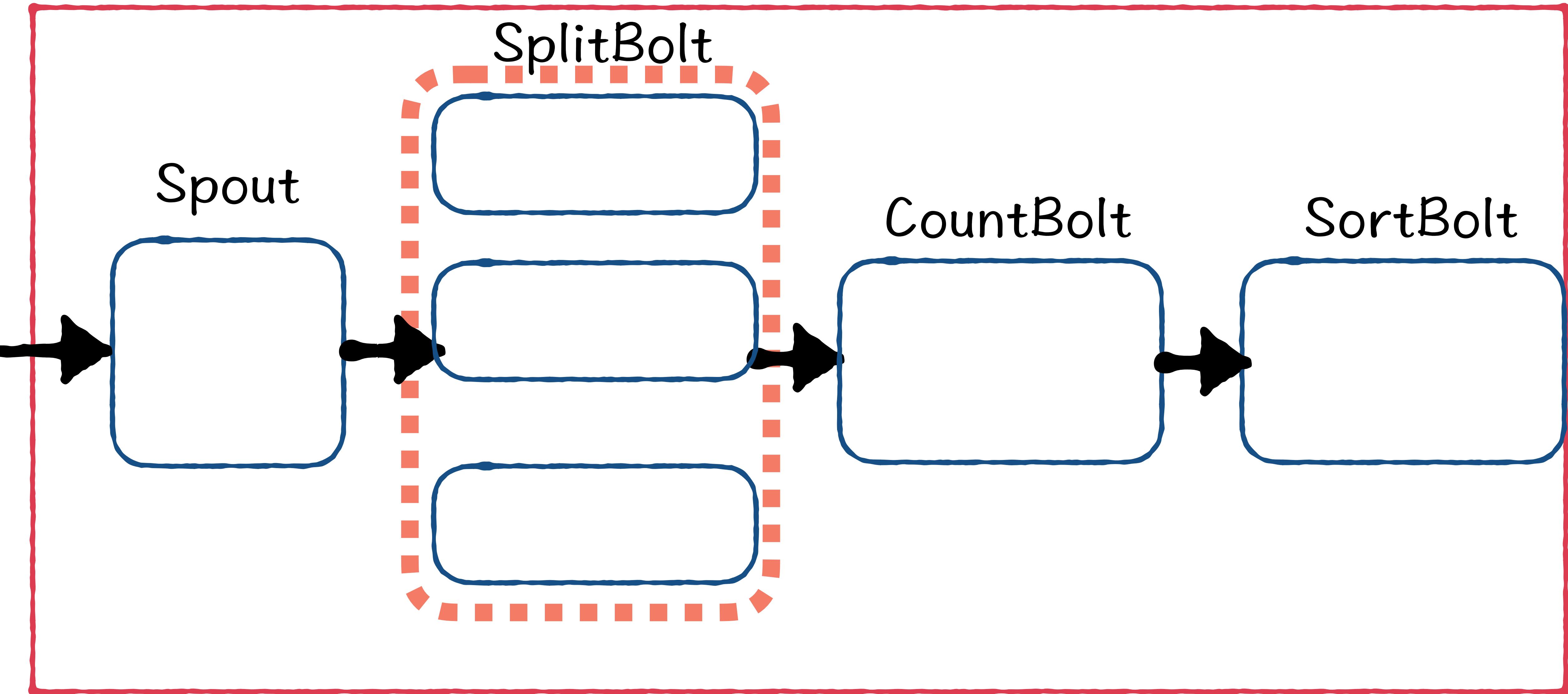
# Trending Tweets Topology



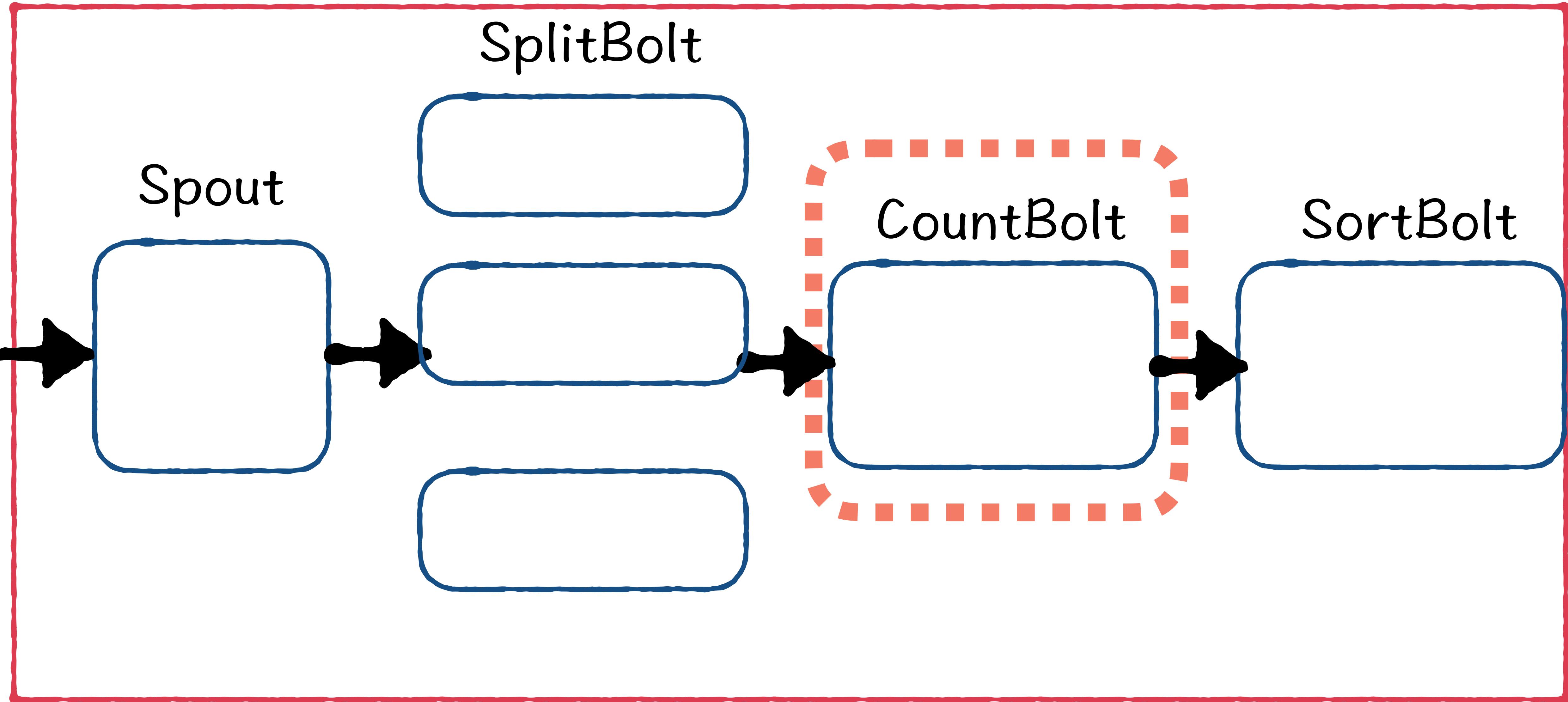
# Trending Tweets Topology



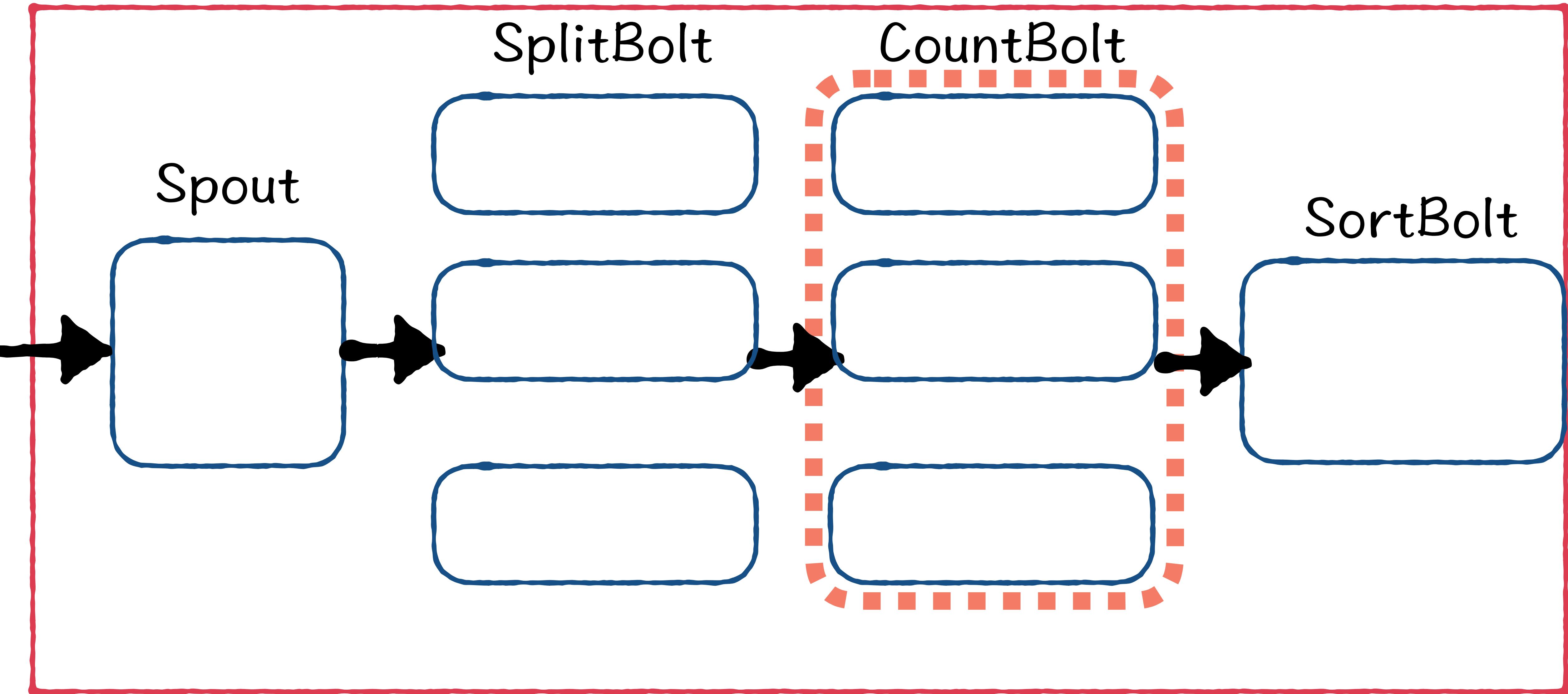
# Trending Tweets Topology



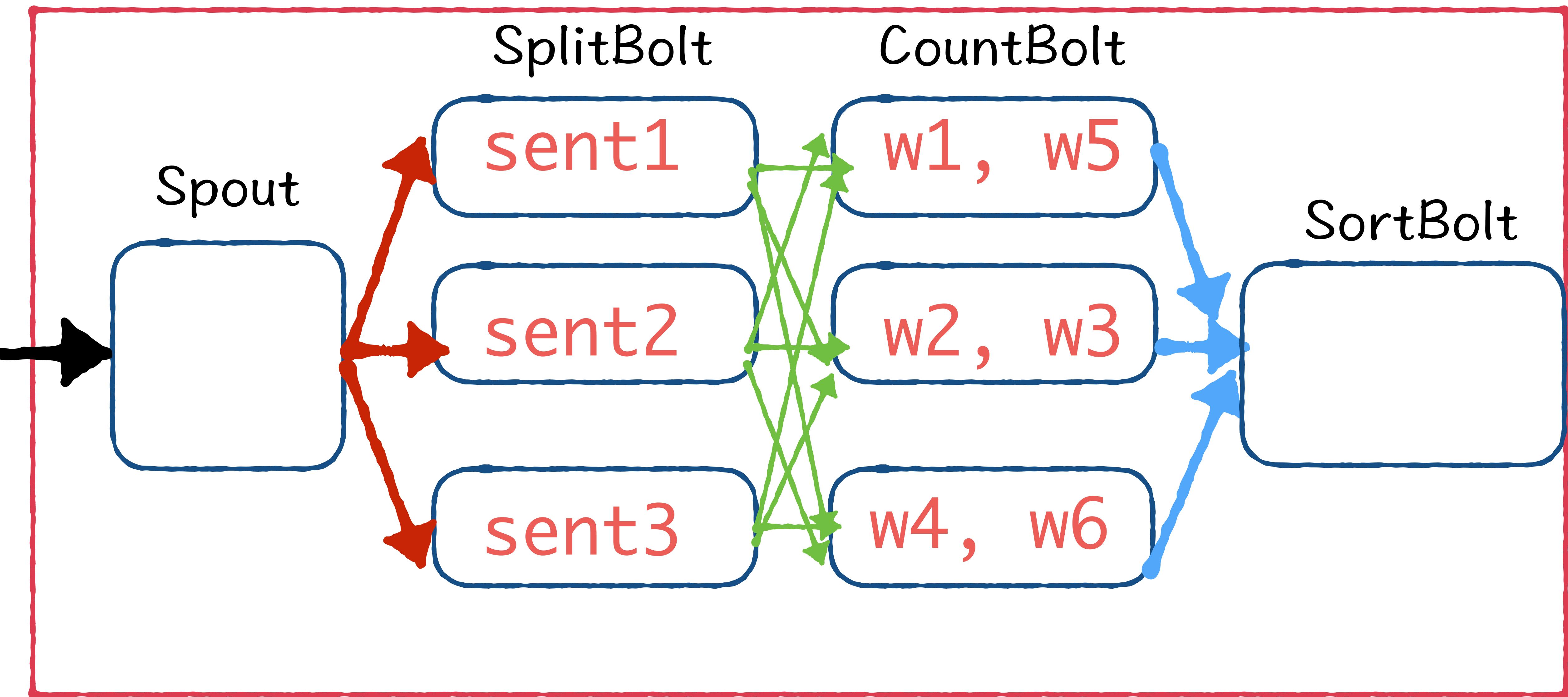
# Trending Tweets Topology



# Trending Tweets Topology



# Trending Tweets Topology



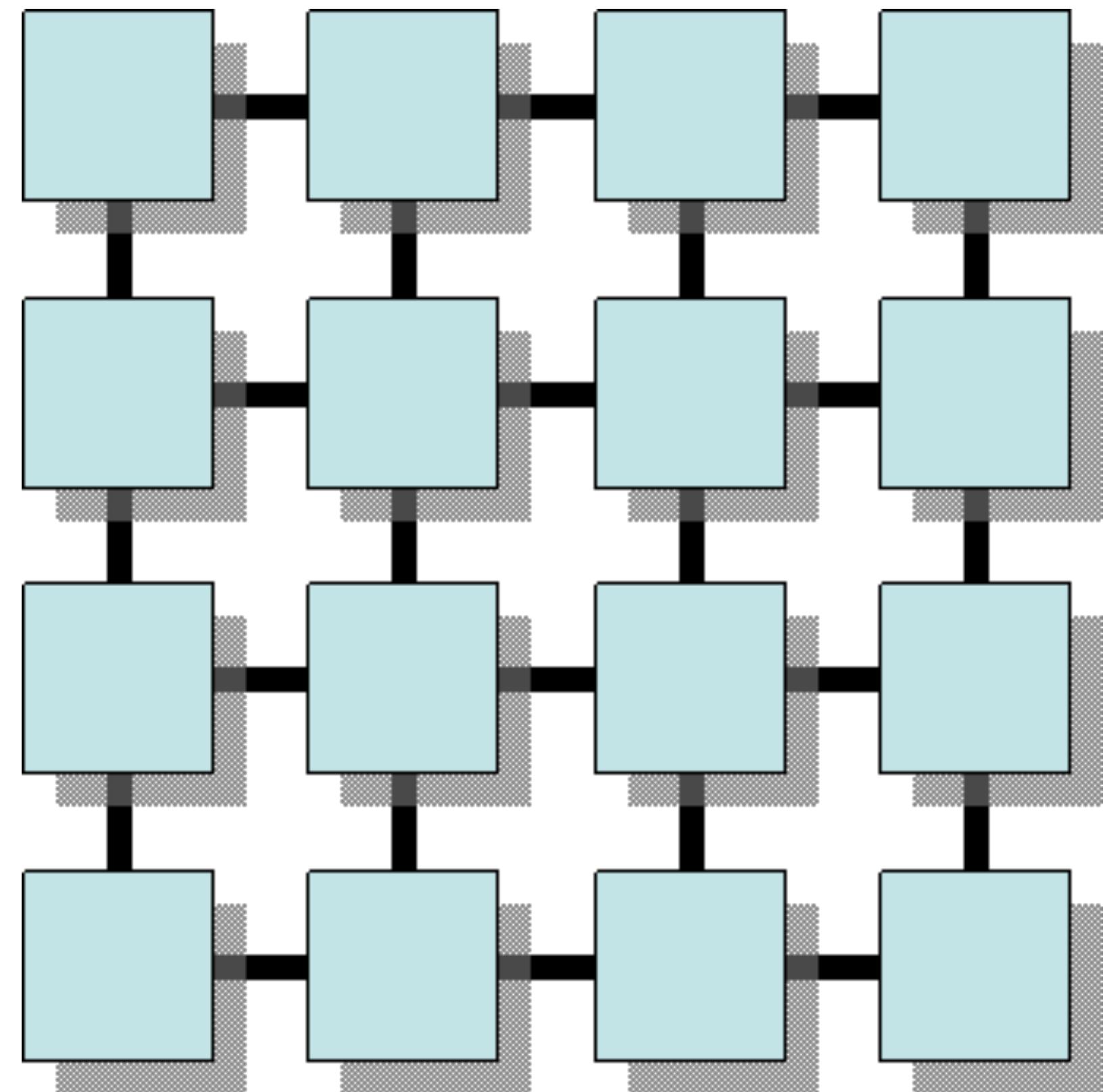
# Apache Storm



Storm topologies run on a Storm Cluster

# Apache Storm

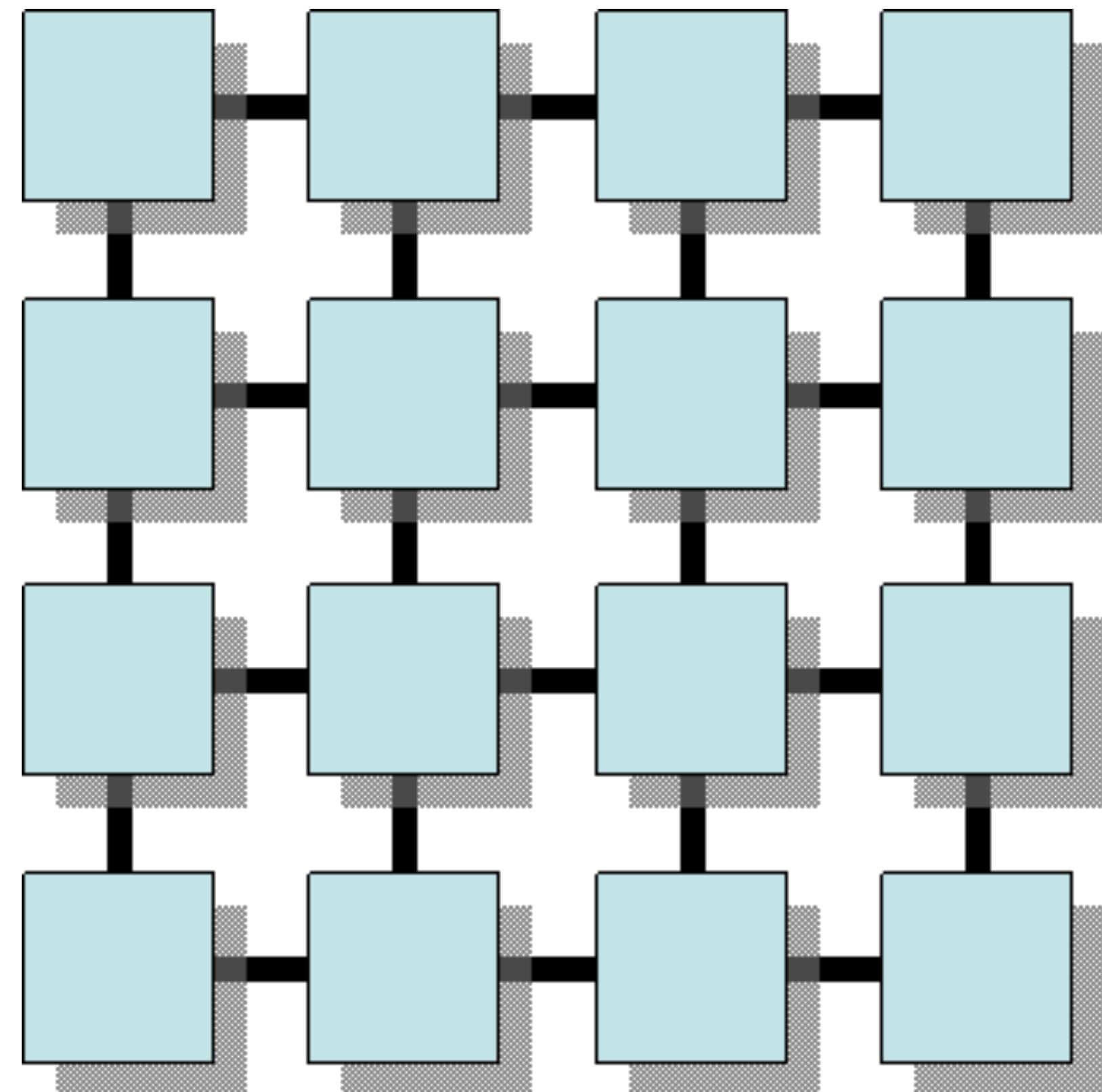
A storm cluster is a group  
of processes which are  
**centrally co-ordinated**



# Apache Storm

The processes can run  
on a single machine

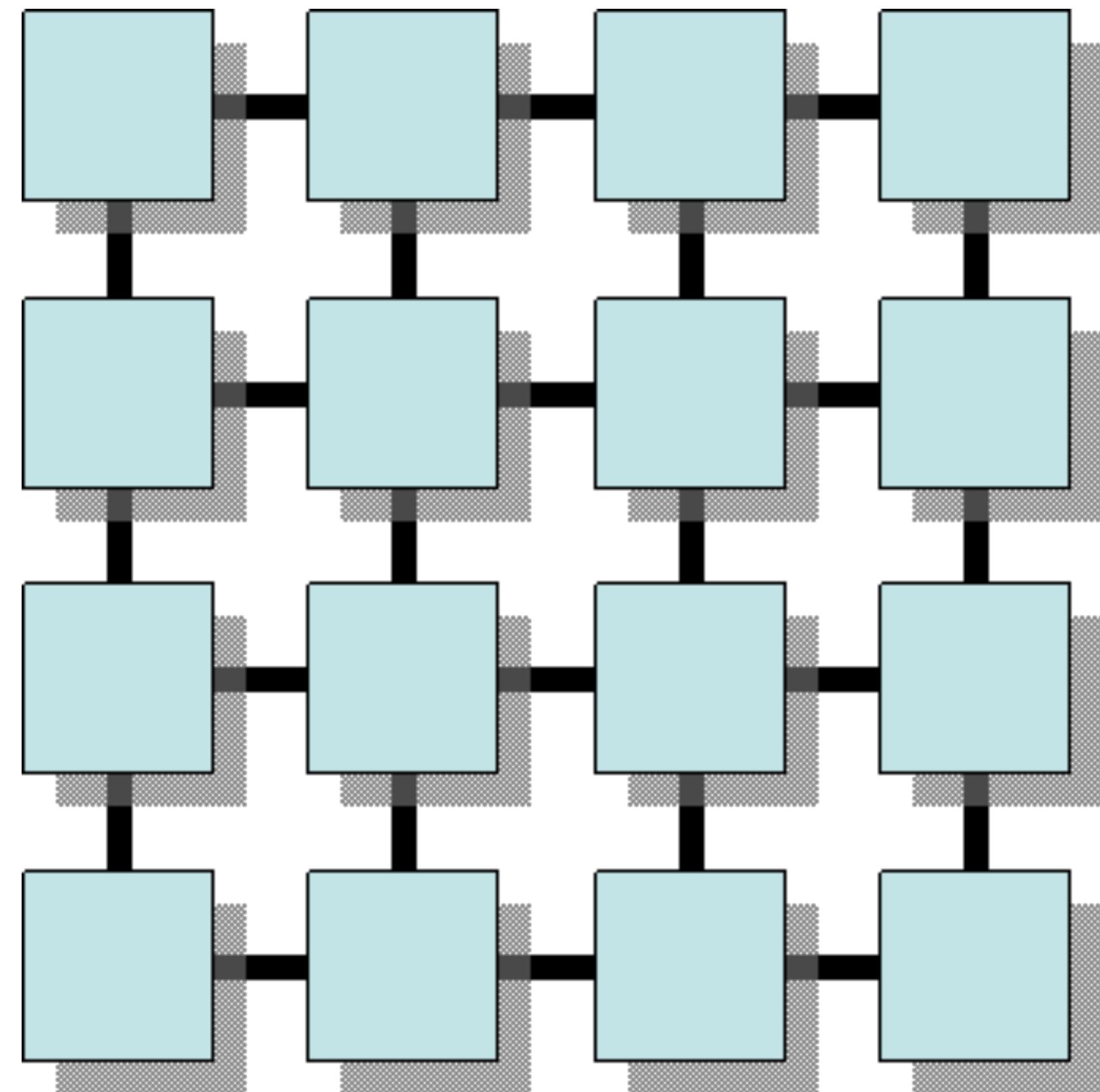
Local Mode



# Apache Storm

The processes can run  
on different machines

## Remote Mode



# Apache Storm

Nimbus

Central co-ordinating  
process

# Apache Storm

Nimbus

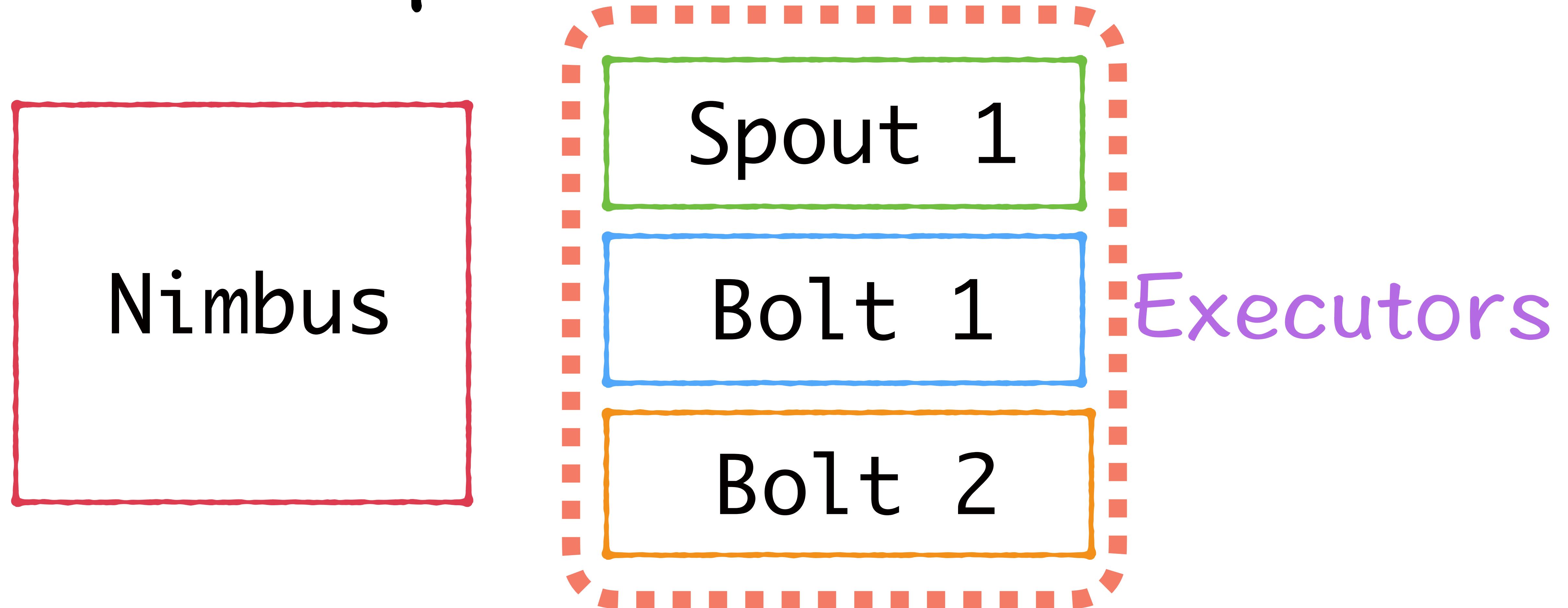
Storm topologies are submitted to the Nimbus

# Apache Storm

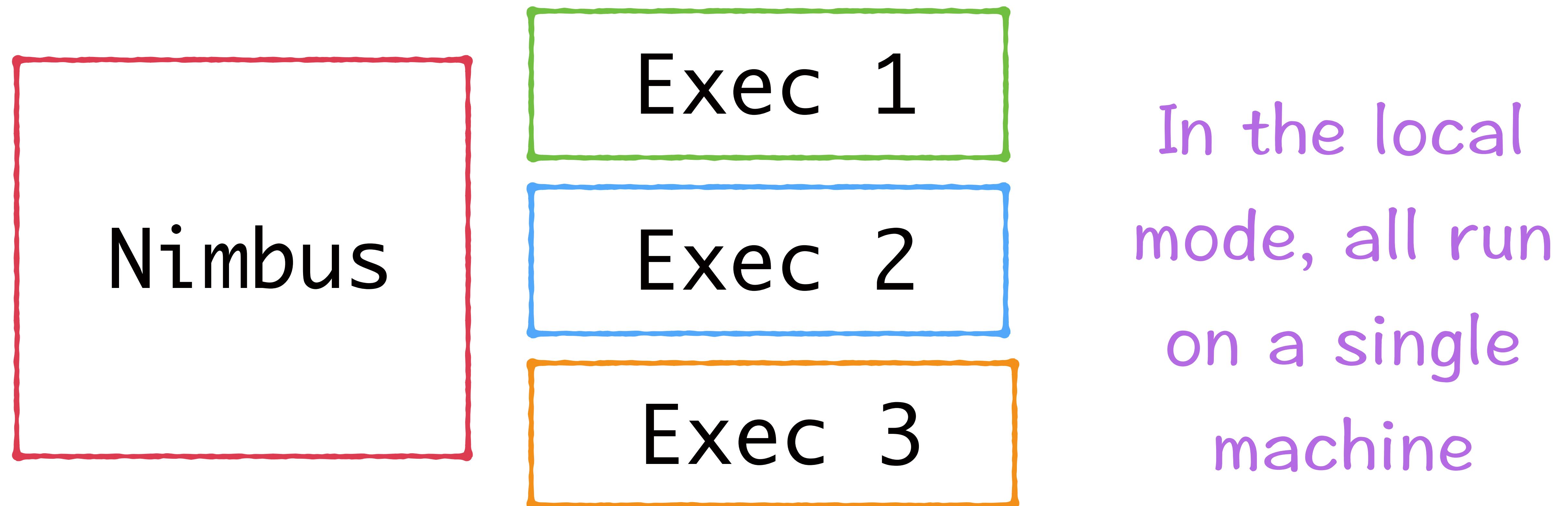
Nimbus

The Nimbus will start up a bunch of processes for Spouts and Bolts

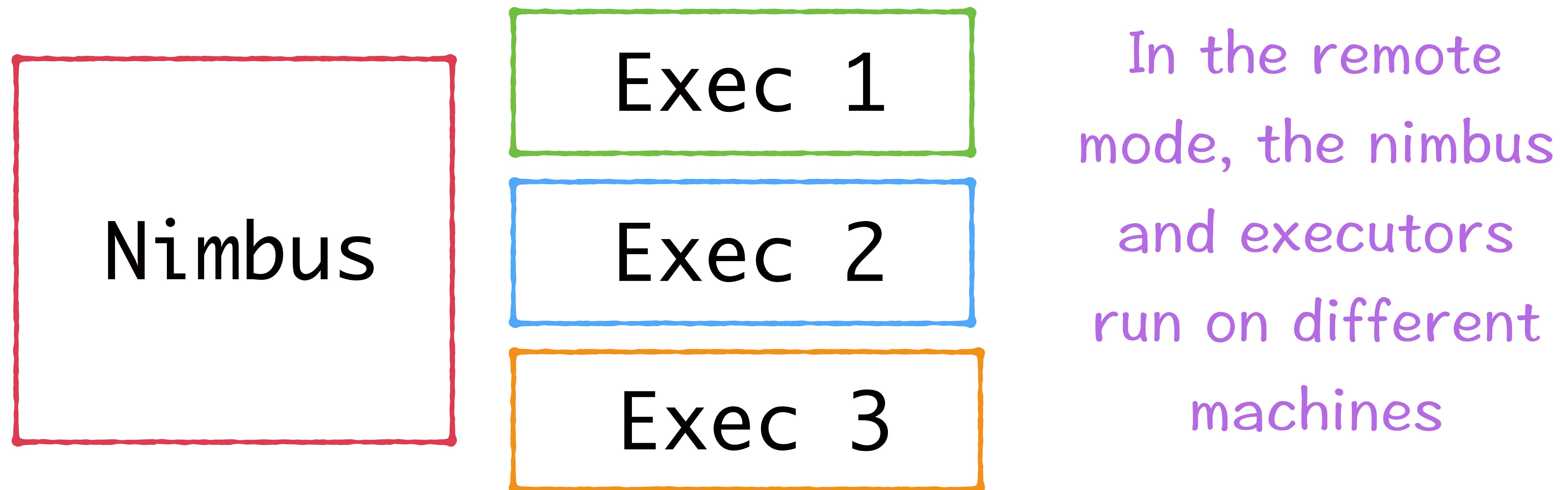
# Apache Storm



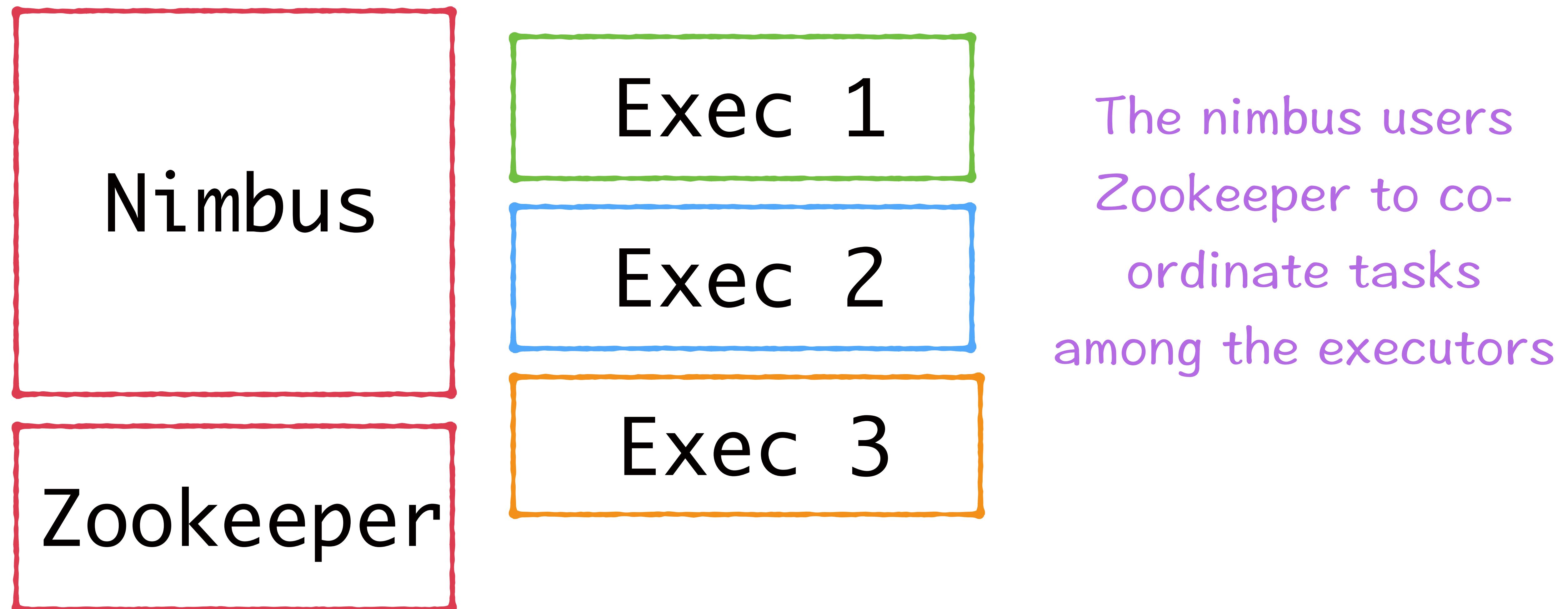
# Apache Storm



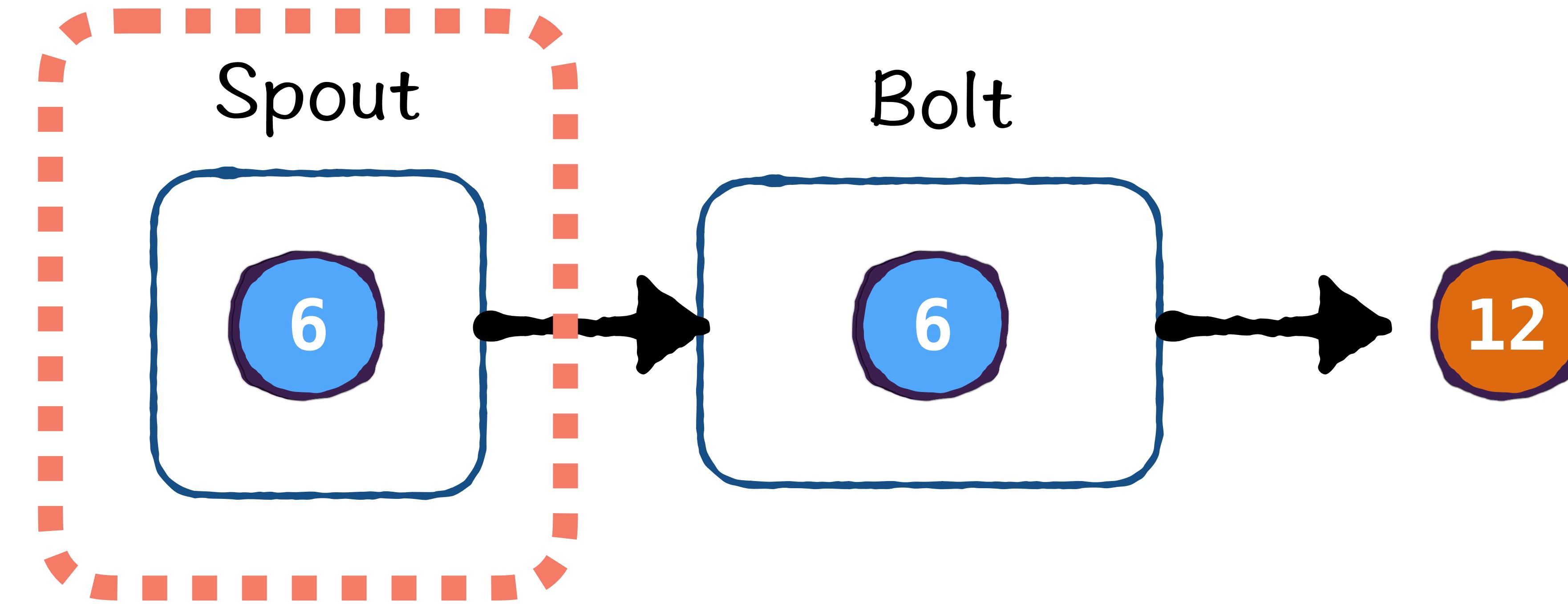
# Apache Storm



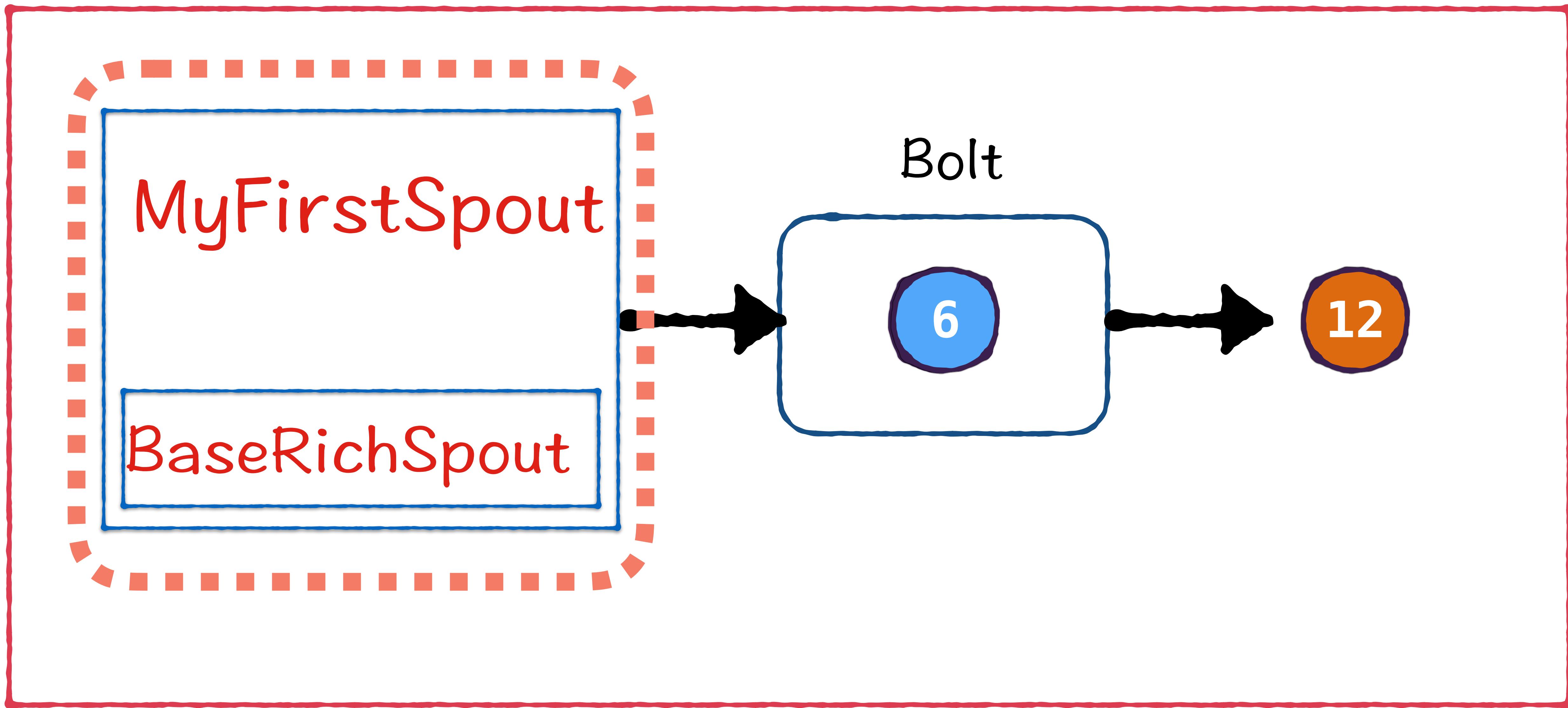
# Apache Storm



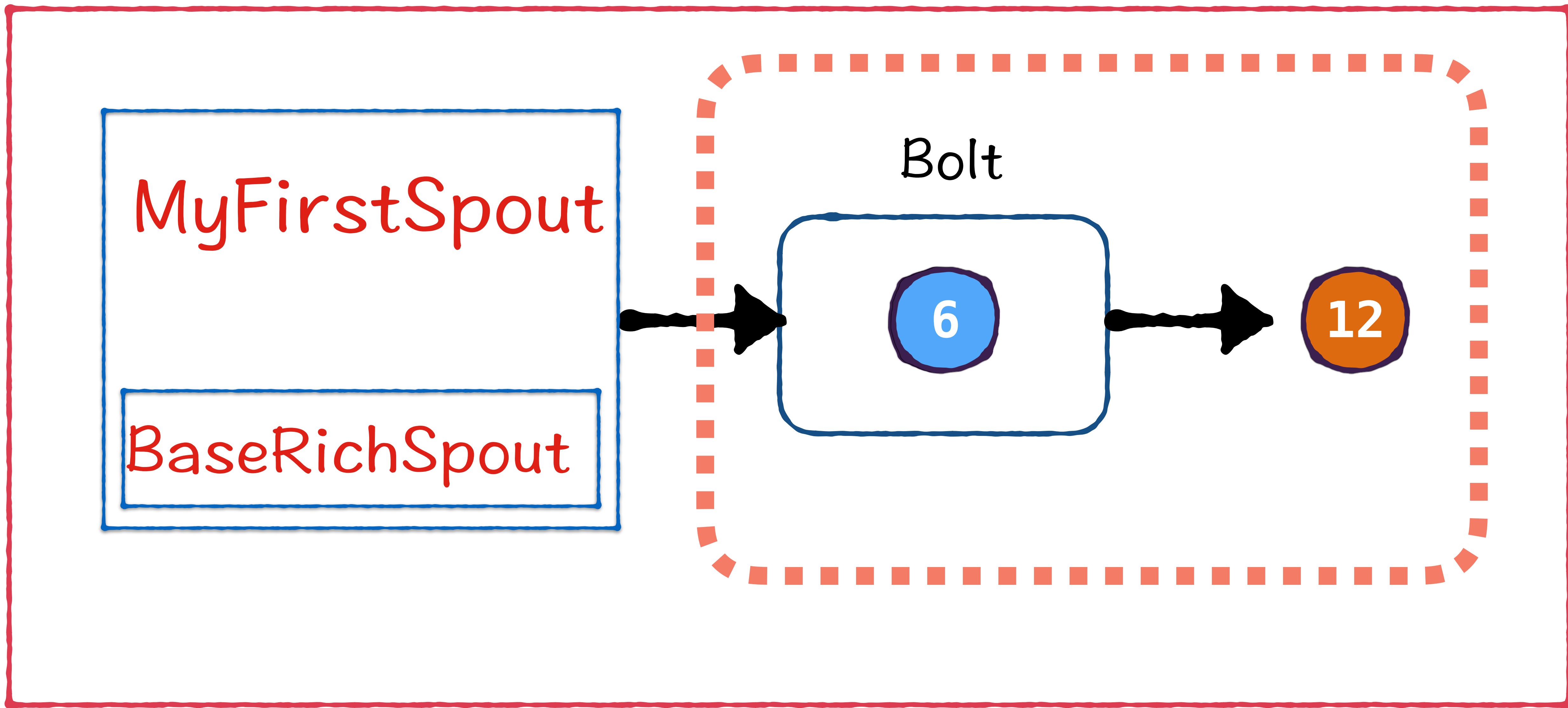
# Hello World Topology



# Hello World Topology



# Hello World Topology



# Hello World Topology

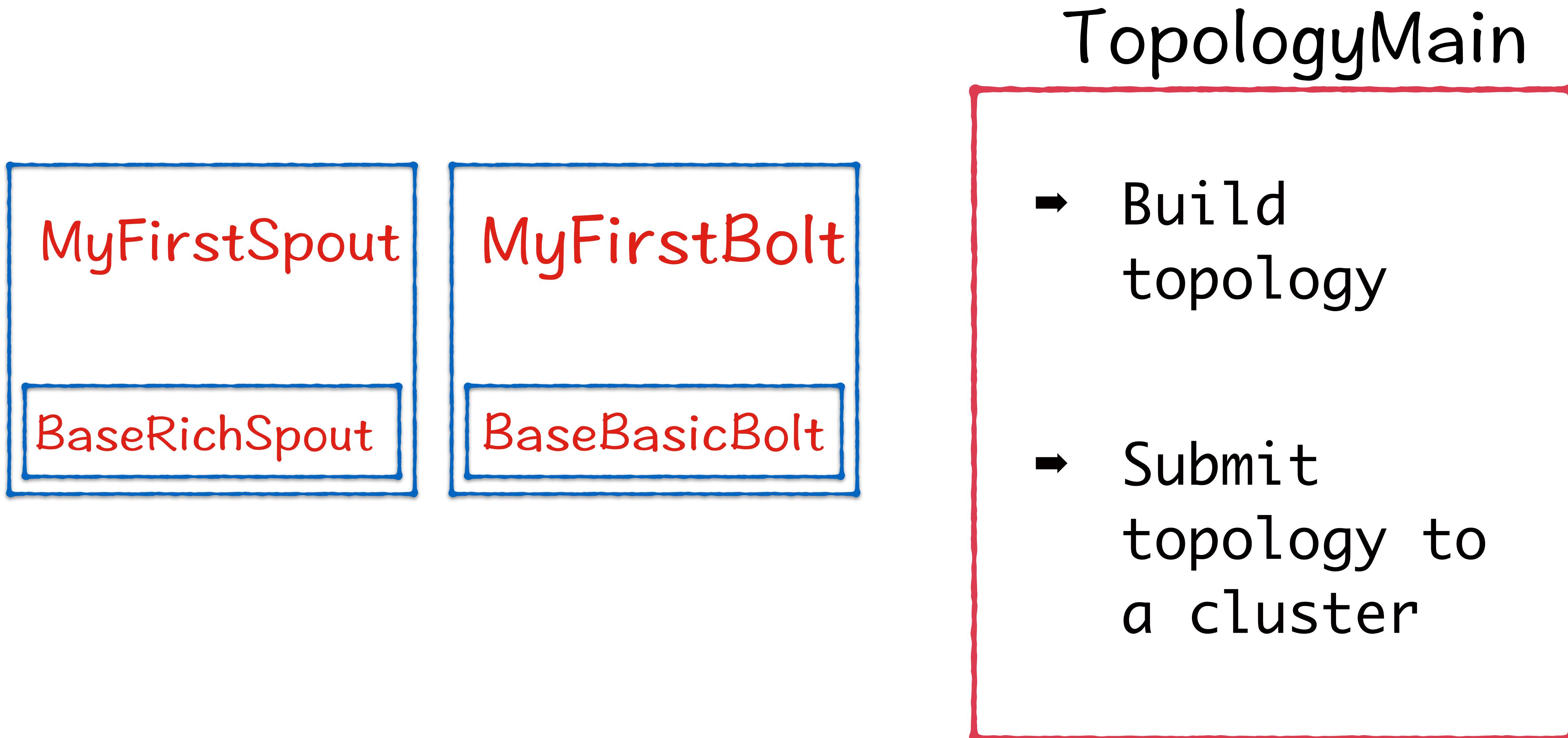
MyFirstSpout

BaseRichSpout

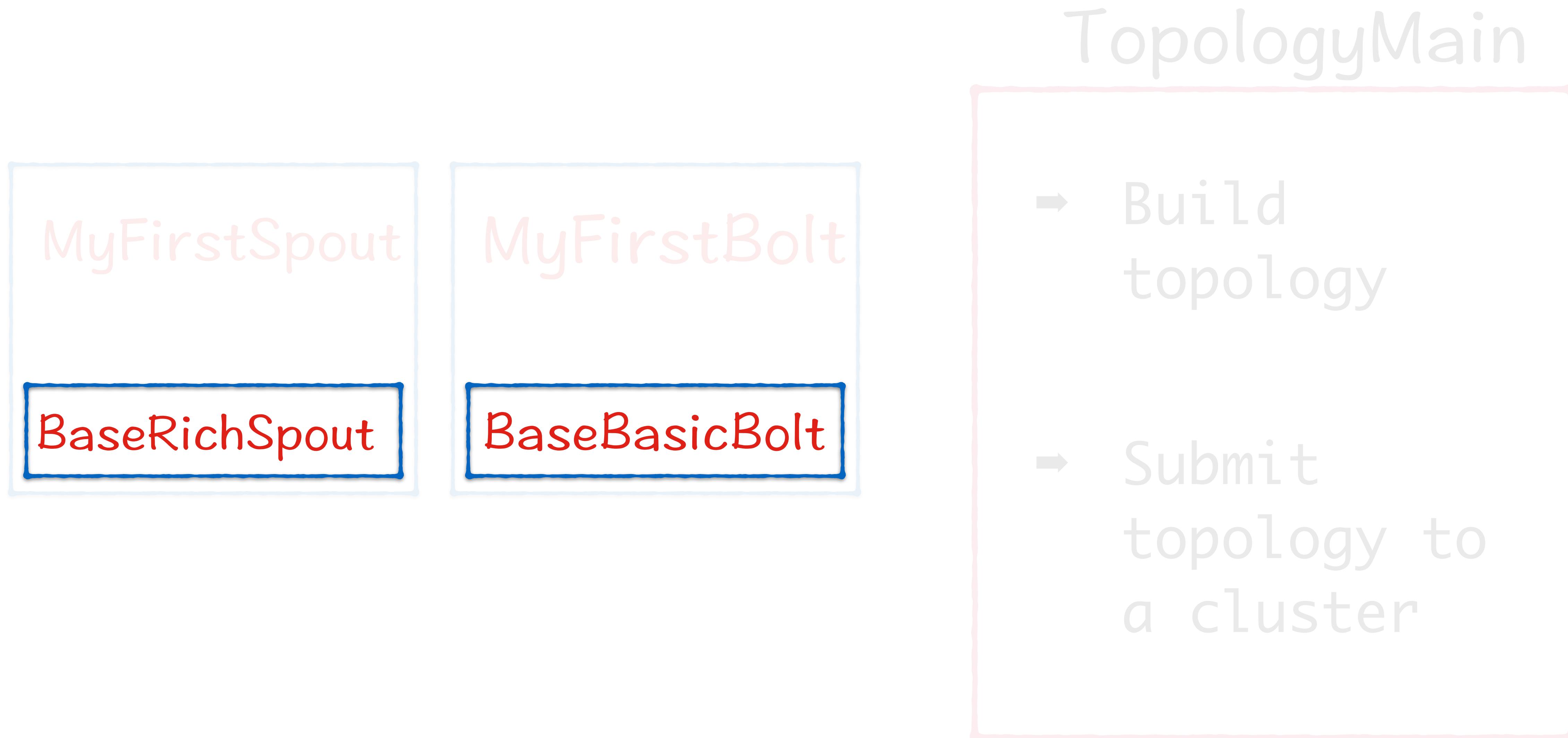
MyFirstBolt

BaseBasicBolt

# Hello World Topology



# Hello World Topology



IComponent

IRichSpout

IBasicBolt

BaseRichSpout

BaseBasicBolt

