

# Coflow

A Networking Abstraction  
For Cluster Applications

Mosharaf Chowdhury  
Ion Stoica



# Cluster Applications

## Multi-Stage Data Flows

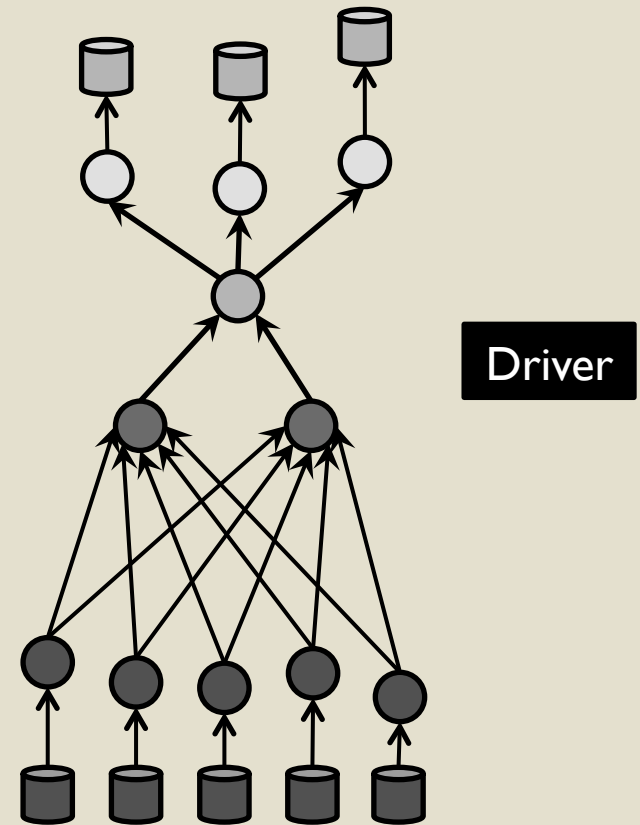
- » Computation interleaved with communication

## Computation

- » Distributed
- » Runs on many machines

## Communication

- » Structured
- » Between machine groups



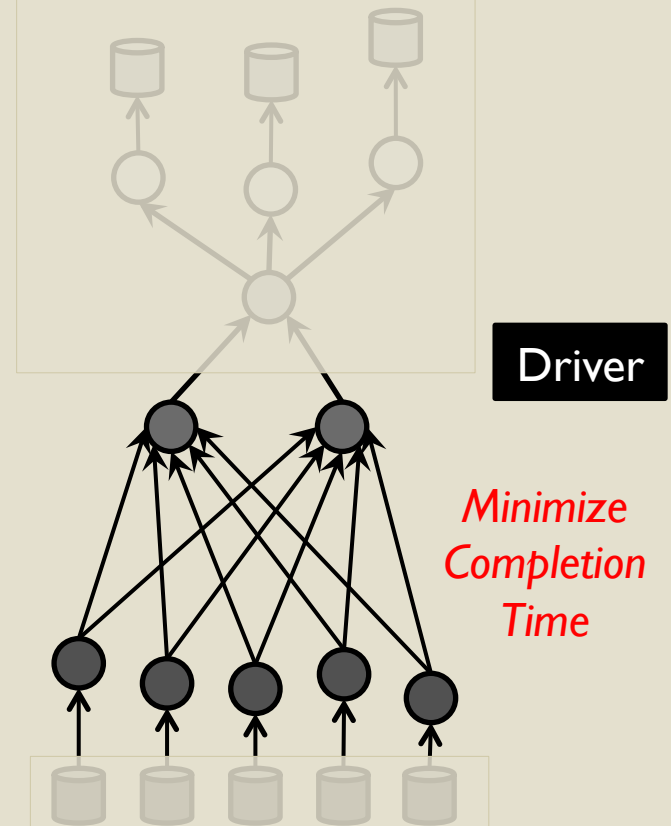
# Communication Abstraction

## A Flow

- » Sequence of packets
- » Independent
- » Often the unit for network scheduling, traffic engineering, load balancing etc.

## Multiple Parallel Flows

- » Independent
- » **Yet**, semantically bound
- » Shared objective



# Coflow

“ A **collection** of flows between two groups of machines that are **bound together by application-specific semantics** ”

## Captures

1. Structure
2. Shared Objective
3. Semantics

# *We Want To...*

Better schedule the network

- » **Intra-coflow**
- » **Inter-coflow**

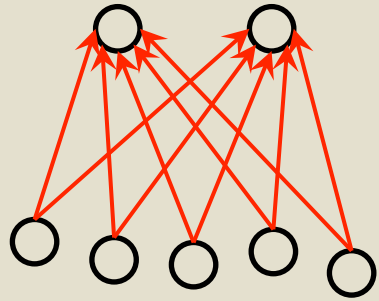
Write the communication layer of a new application

- » **Without reinventing the wheel**

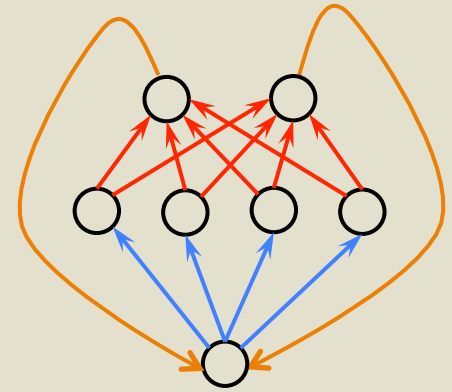
Add unsupported coflows to an application, or

Replace an existing coflow implementation

- » **Independent of applications**



Cluster  
Applications



# Coflow API

**The Network**

(Physically or Logically Centralized Controller)

# Coflow API

## Goals

1. Separate intent from mechanisms
2. Convey application-specific semantics to the network

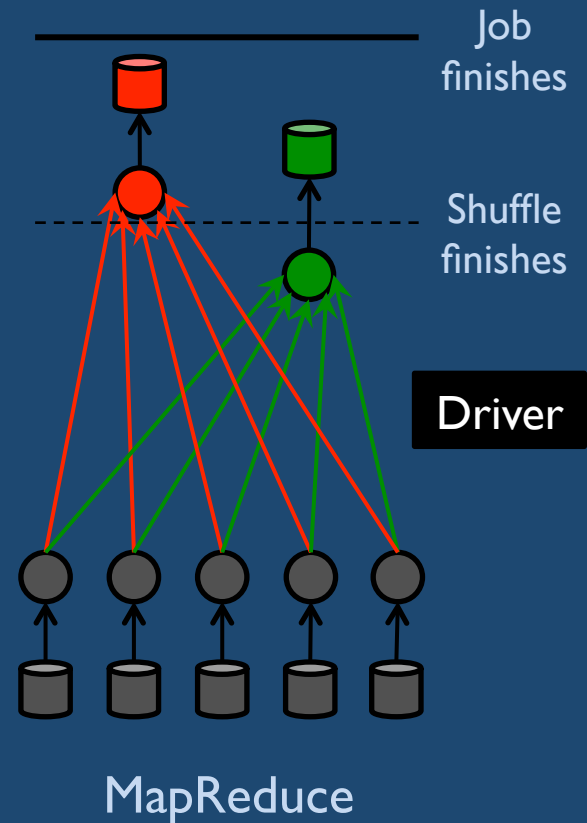
# Coflow API

`terminate(handle)`

`get(handle, id) → content`

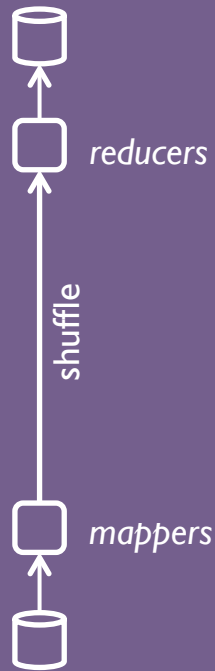
`put(handle, id, content)`

`create(SHUFFLE) → handle`





# Coflow Flexibility



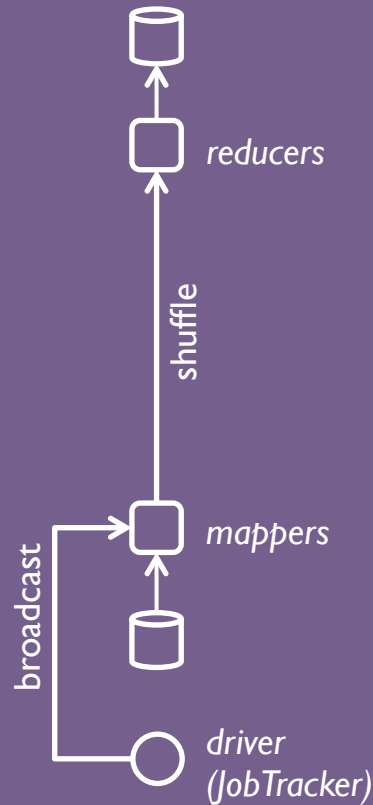
Choice of algorithms

- » Default
- » WSS<sup>1</sup>

Choice of mechanism

- » App vs. Network layer
- » Pull vs. Push

# Coflow Flexibility



**@driver**

$b \leftarrow \text{create}(\text{BCAST})$

...

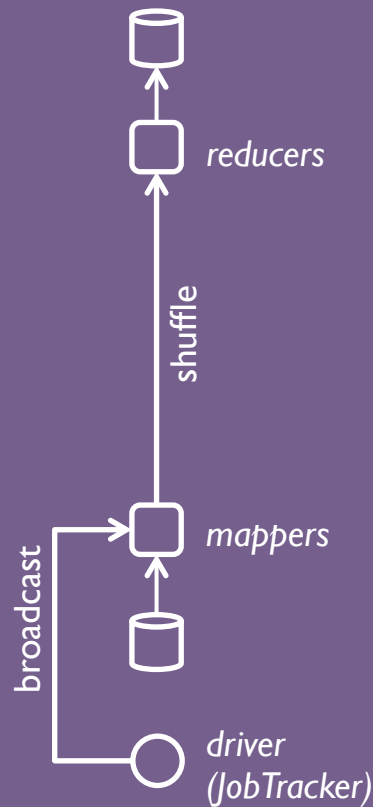
$\text{put}(b, id, \text{content})$

**@mapper**

$\text{get}(b, id)$

...

# Coflow Flexibility



## @driver

```
b ← create(BCAST)
s ← create(SHUFFLE,
           ord=[b ~> s])
```

```
put(b, id, content)
```

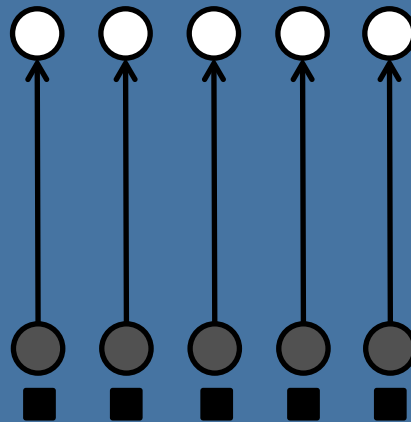
```
...
terminate(b)
terminate(s)
```

## @mapper

```
get(b, id)
put(s, ids1)
```

```
...
```

# Throughput-Sensitive Applications

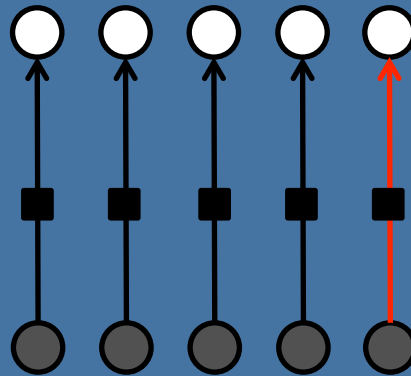


After 2 seconds

*Minimize Completion Time*

# Throughput-Sensitive Applications

After 4 seconds



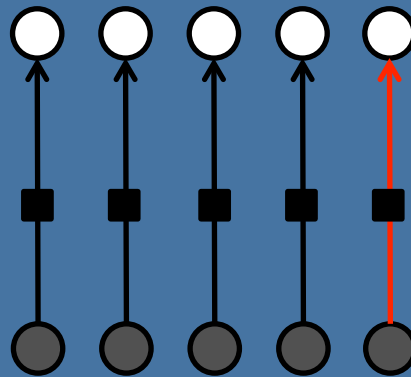
After 7 seconds

After 2 seconds

*Minimize Completion Time*

# Throughput-Sensitive Applications

Free up resources  
**without hurting**  
application-perceived  
communication time

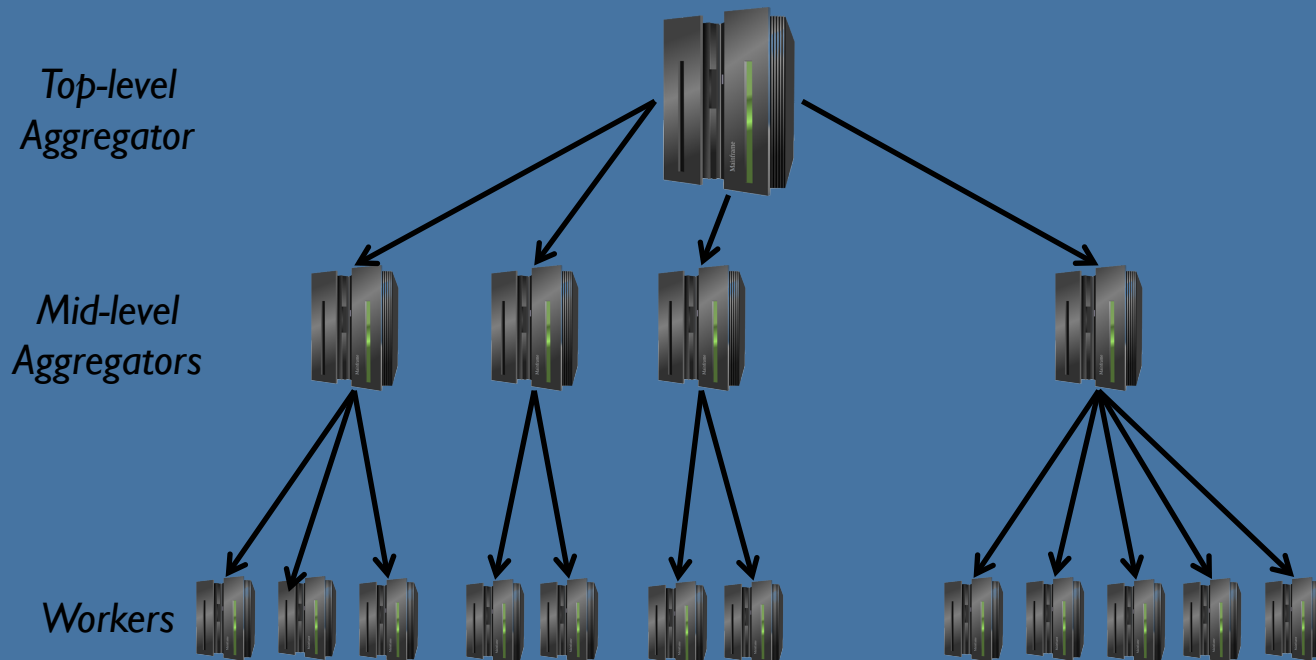
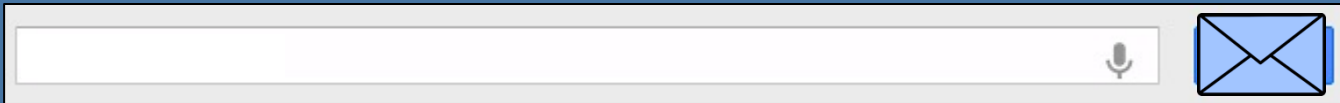


After 7 seconds

After 2 seconds

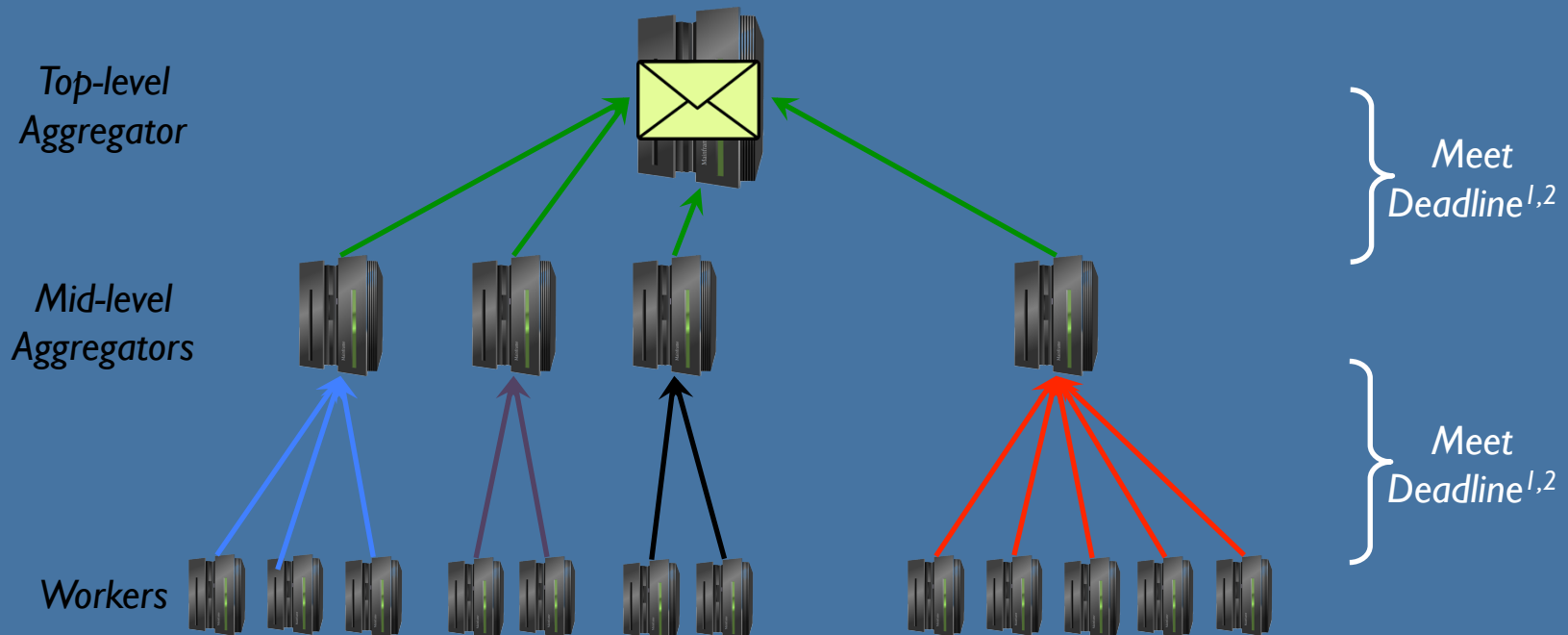
Minimize Completion Time

# Latency-Sensitive Applications



# Latency-Sensitive Applications

HotNets 2012



1. D3, SIGCOMM'2011
2. PDQ, SIGCOMM'2012



# Latency-Sensitive Applications

HotNets 2012



## [HotNets-XI: Home Page](#)

[conferences.sigcomm.org/hotnets/2012/](http://conferences.sigcomm.org/hotnets/2012/)

The Eleventh ACM Workshop on Hot Topics in Networks (HotNets-XI) will bring together people with interest in computer networks to engage in a lively debate ...

## [HotNets Workshop | acm sigcomm](#)

[www.sigcomm.org/events/hotnets-workshop](http://www.sigcomm.org/events/hotnets-workshop)

The Workshop on Hot Topics in Networks (HotNets) was created in 2002 to discuss early-stage, creative ... HotNets-XI, Seattle, WA area, October 29-30, 2012.

## [HotNets-XI: Call for Papers](#)

[conferences.sigcomm.org/hotnets/2012/cfp.shtml](http://conferences.sigcomm.org/hotnets/2012/cfp.shtml)

The Eleventh ACM Workshop on Hot Topics in Networks (HotNets-XI) will bring together researchers in computer networks and systems to engage in a lively ...

## [Coflow accepted at HotNets'2012](#)

[www.mosharaf.com/blog/2012/09/.../coflow-accepted-at-hotnets201...](http://www.mosharaf.com/blog/2012/09/.../coflow-accepted-at-hotnets201...)

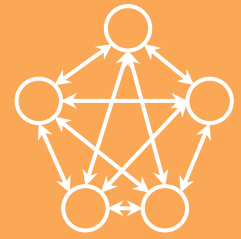
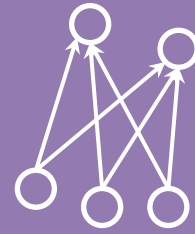
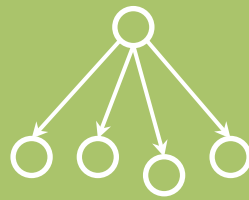
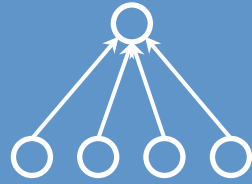
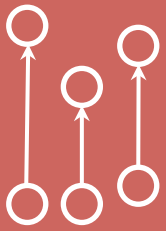
Sep 13, 2012 – Update: Coflow camera-ready is available online! Tell us what you think! Our position paper to address the lack of a networking abstraction for ...

1. D3, SIGCOMM'2011
2. PDQ, SIGCOMM'2012

*Limit impact to as few requests as possible*

# *One More Thing...*

1. Critical Path Scheduling
2. OpenTCP
3. Structured Streams
4. ...



# Coflow

A semantically-bound collection of flows

Conveys application intent to the network

- » Allows better management of network resources
- » Provides greater flexibility in designing applications

Mosharaf Chowdhury  
<http://www.mosharaf.com/>

