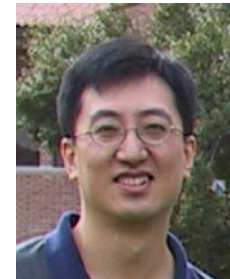# Graviton: Twisting Space and Time to Speed-up CoFlows
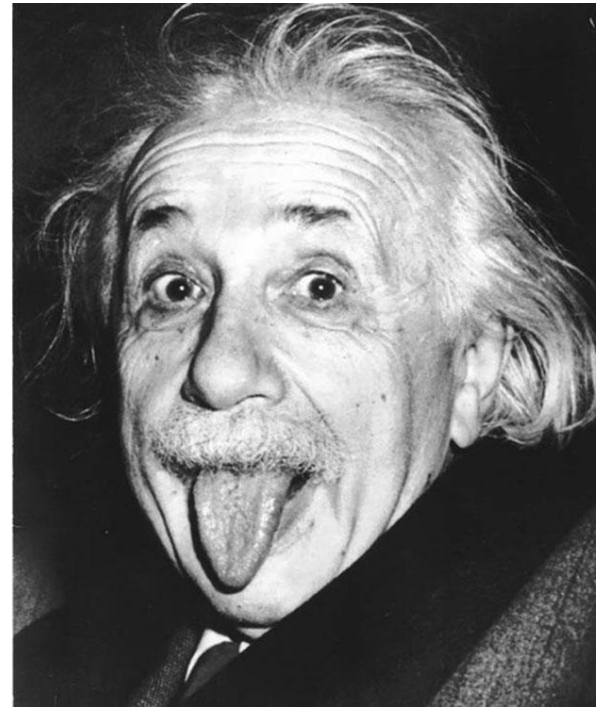
Akshay Jajoo   Rohan Gandhi   Y. Charlie Hu

Graviton is a hypothetical elementary particle that mediates the force of gravitation in the framework of quantum field theory. If it exists, the graviton is expected to be and must be a spin-2 boson.

Graviton in CoFlow Scheduling

Analytics Jobs in Big Data

- Analytics jobs in data-centers
  - process huge amount of data
  - Collected from various resources
  - Driven by applications like Ads, health care

Communication Stage Is Important for Job Performance

Facebook jobs spend 33% time in communications!

In-memory data based computation proliferating →

Networks increasingly becoming bottlenecks

Outline

- CoFlow abstraction
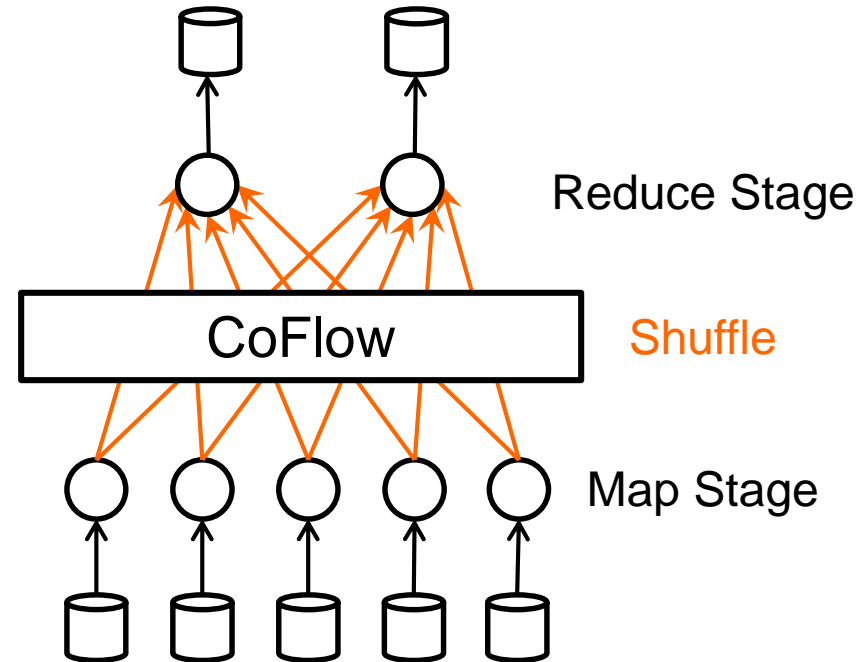- Scheduling in Aalo and its implications
- Graviton key ideas
- Evaluation

CoFlow Abstraction

CoFlow:

Collection of all flows that share same goal

Implication:

CoFlow cannot finish until last flow finishes

Reduce Stage

CoFlow

Shuffle

Map Stage

CoFlow Scheduling

- CCT: CoFlow Completion Time
- CoFlow scheduling problem
  - Minimize overall CoFlow Completion Time
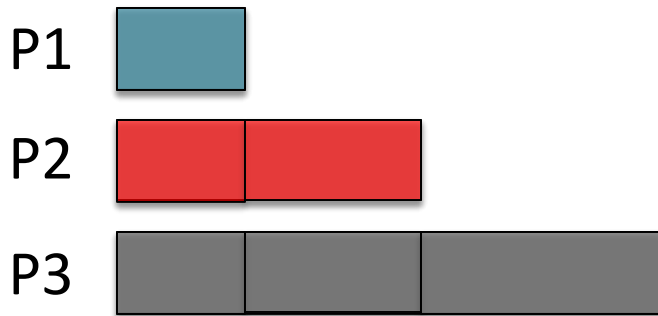- CoFlow scheduling problem is NP Hard

Scheduling 101

Shortest-Job-First (SJF): optimal in minimizing overall completion time
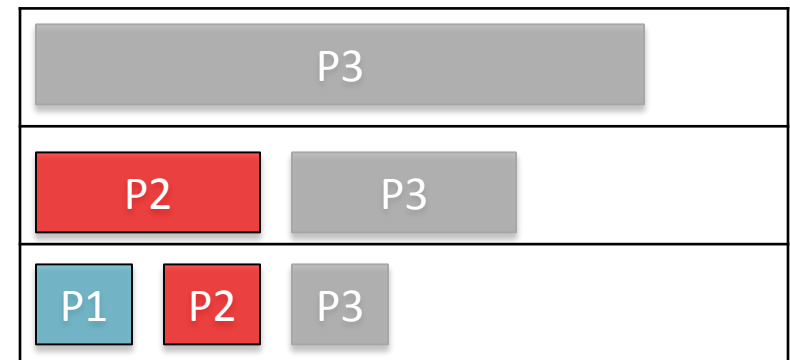
Last   First

Process
Scheduling

| P_3 | P_2 | P_1 |

$$P_1 < P_2 < P_3$$

Online Approximation to SJF

Process durations

Priority queues
(Higher Priority = more CPU time)

P1

P2
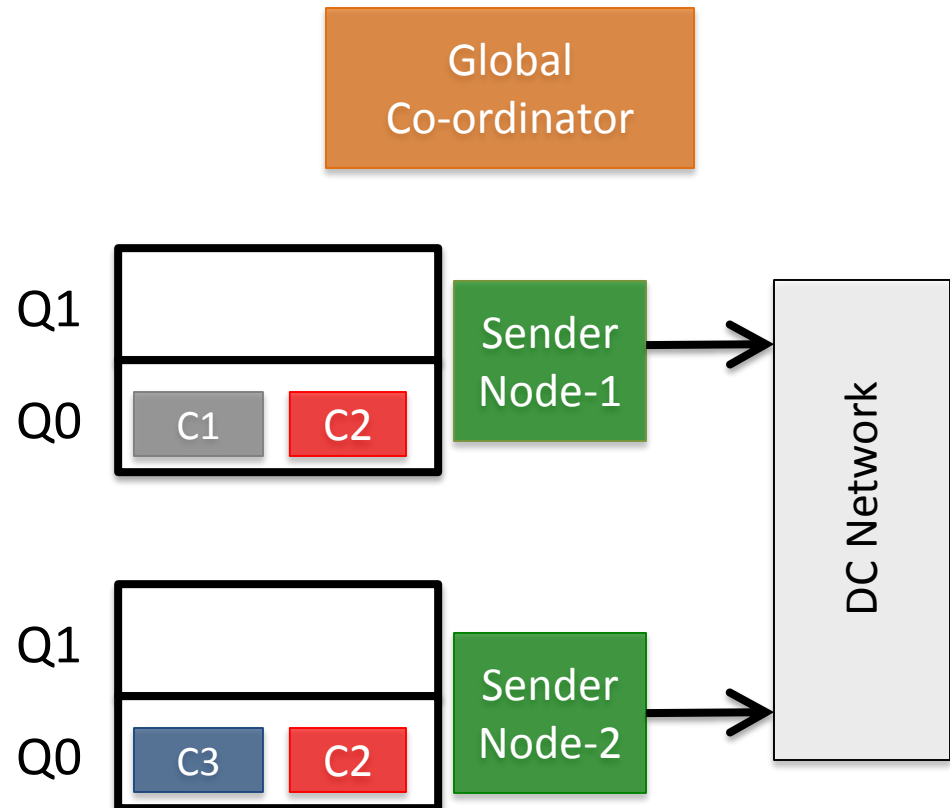
P3

Low

Q_2

Q_1

High Q_0

P3

P2 P3

P1 P2 P3

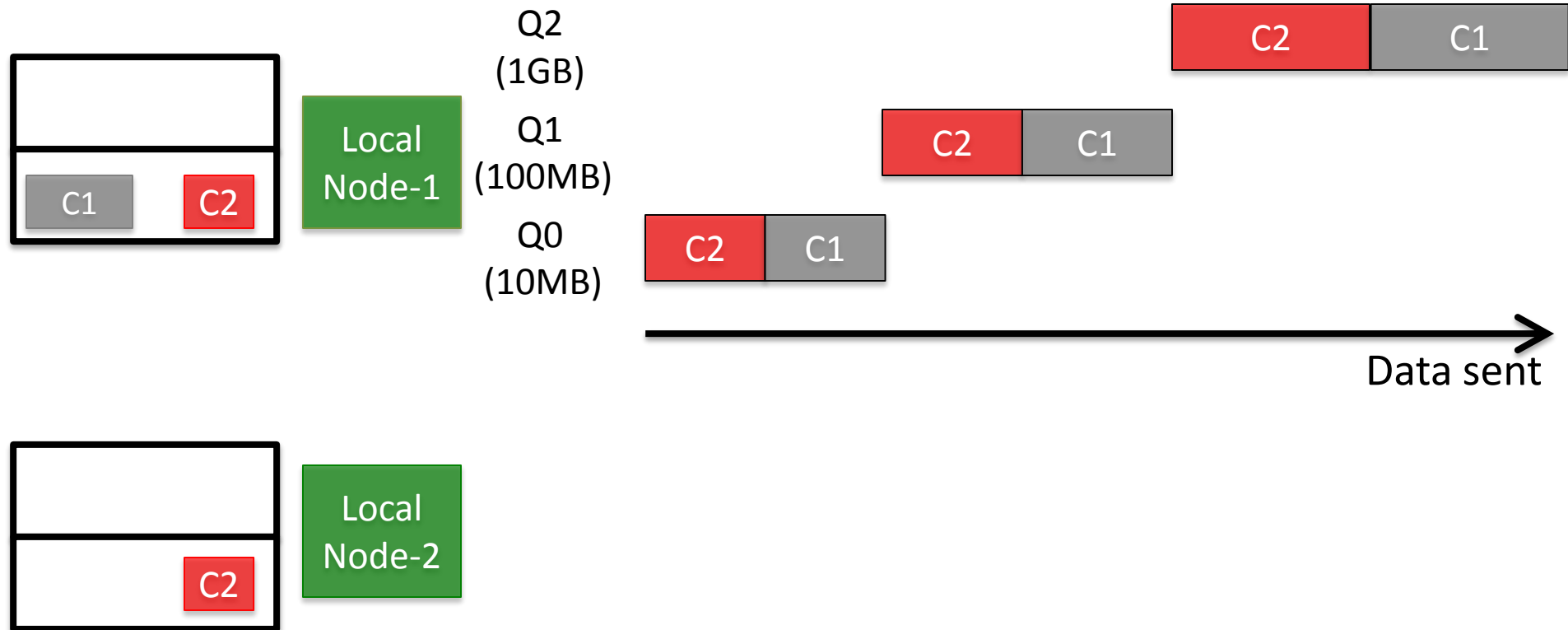Shorter processes finish in High priority queues

Aalo: Online CoFlow Scheduler

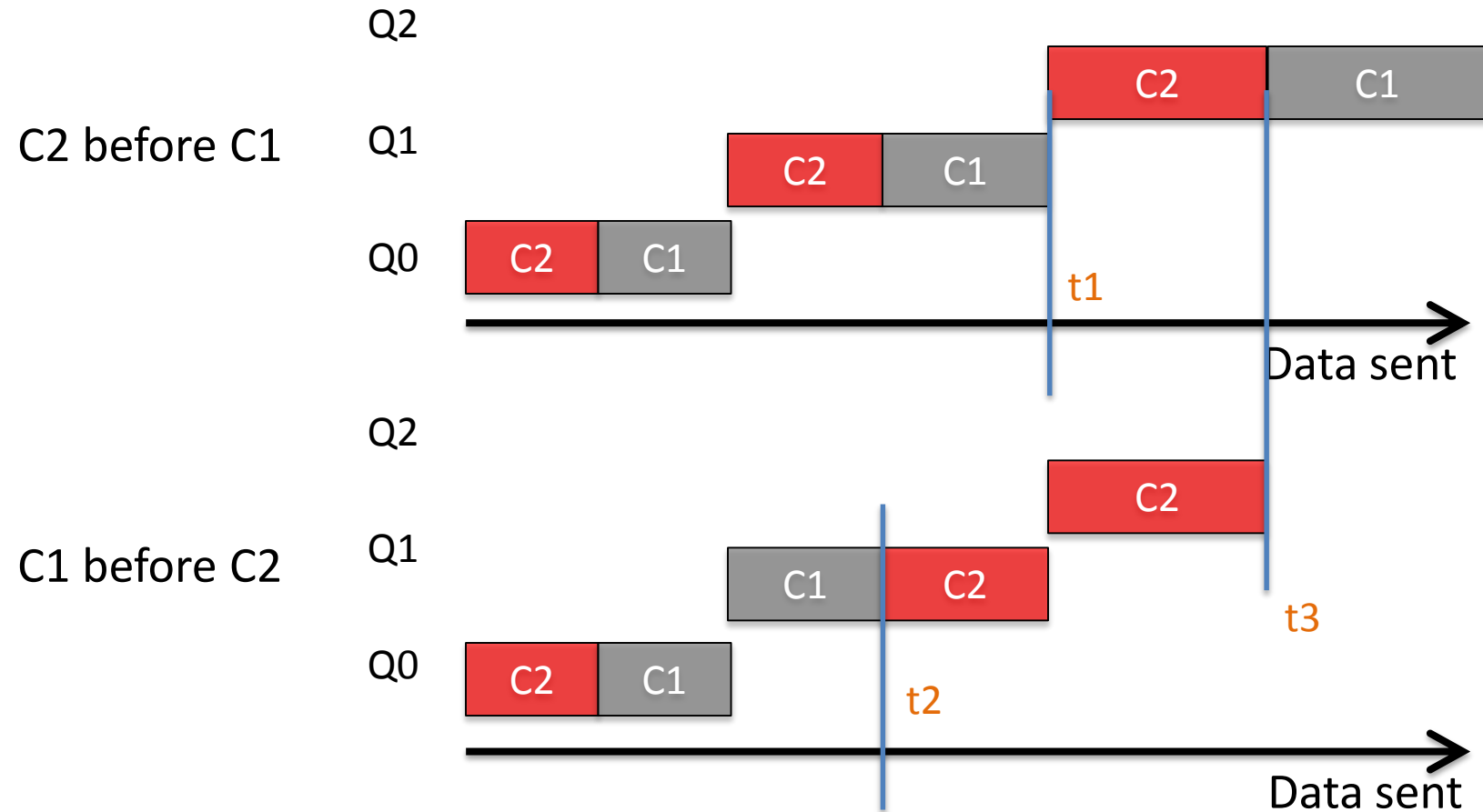CoFlow has spatial dimension → Many flows

How to approximate SJF?

1) Replicate Priority queues
2) Assign CoFlows use: Total bytes sent
3) Intra-queue: Use FIFO

Global Co-ordinator

Q1

Q0   C1   C2

Sender Node-1

Q1

Q0   C3   C2

Sender Node-2

DC Network
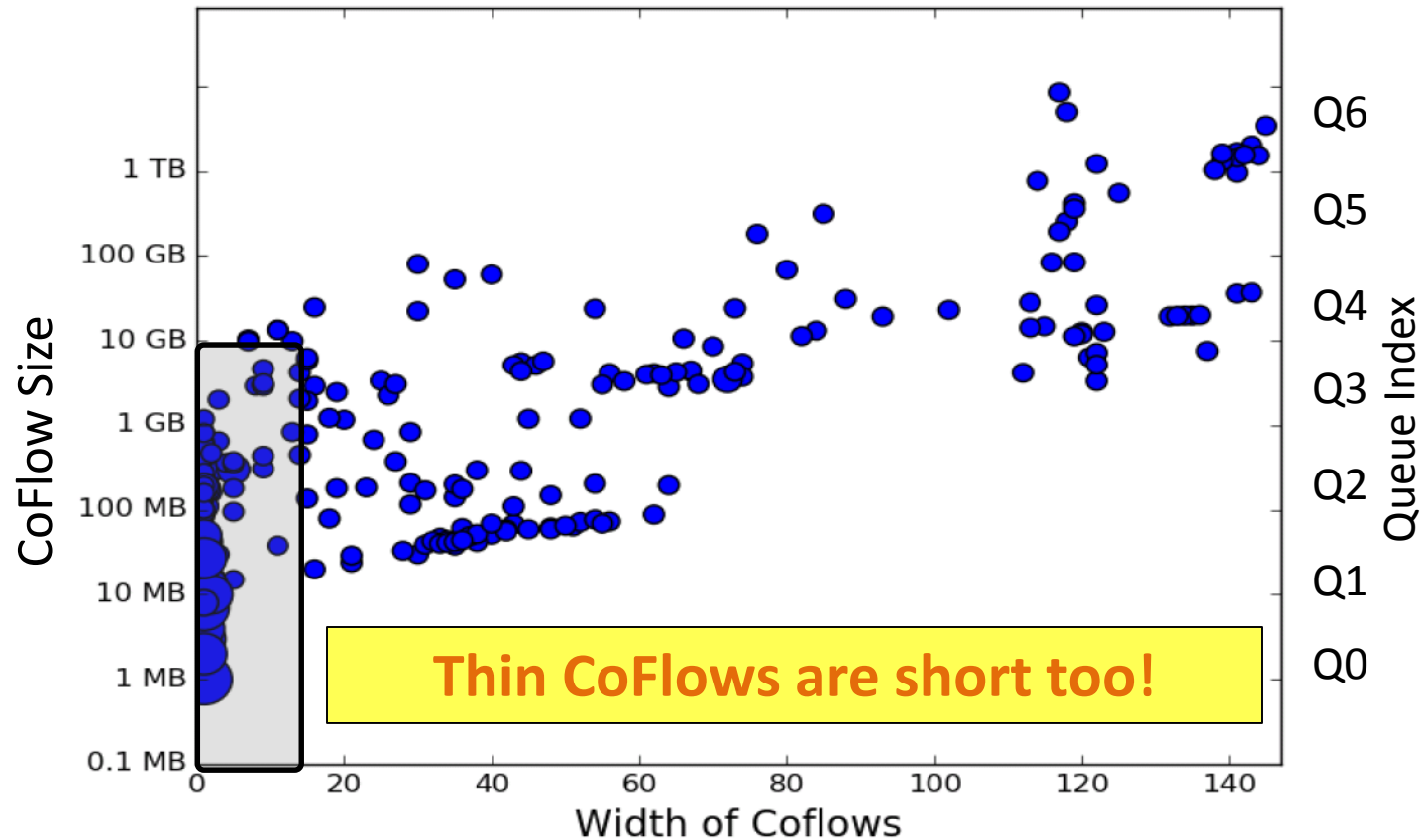
Aalo: CoFlow Queue Transition

Aalo Limitation due to FIFO

Intra-queue: Aalo does not facilitate CoFlows finishing in that queue

Graviton:
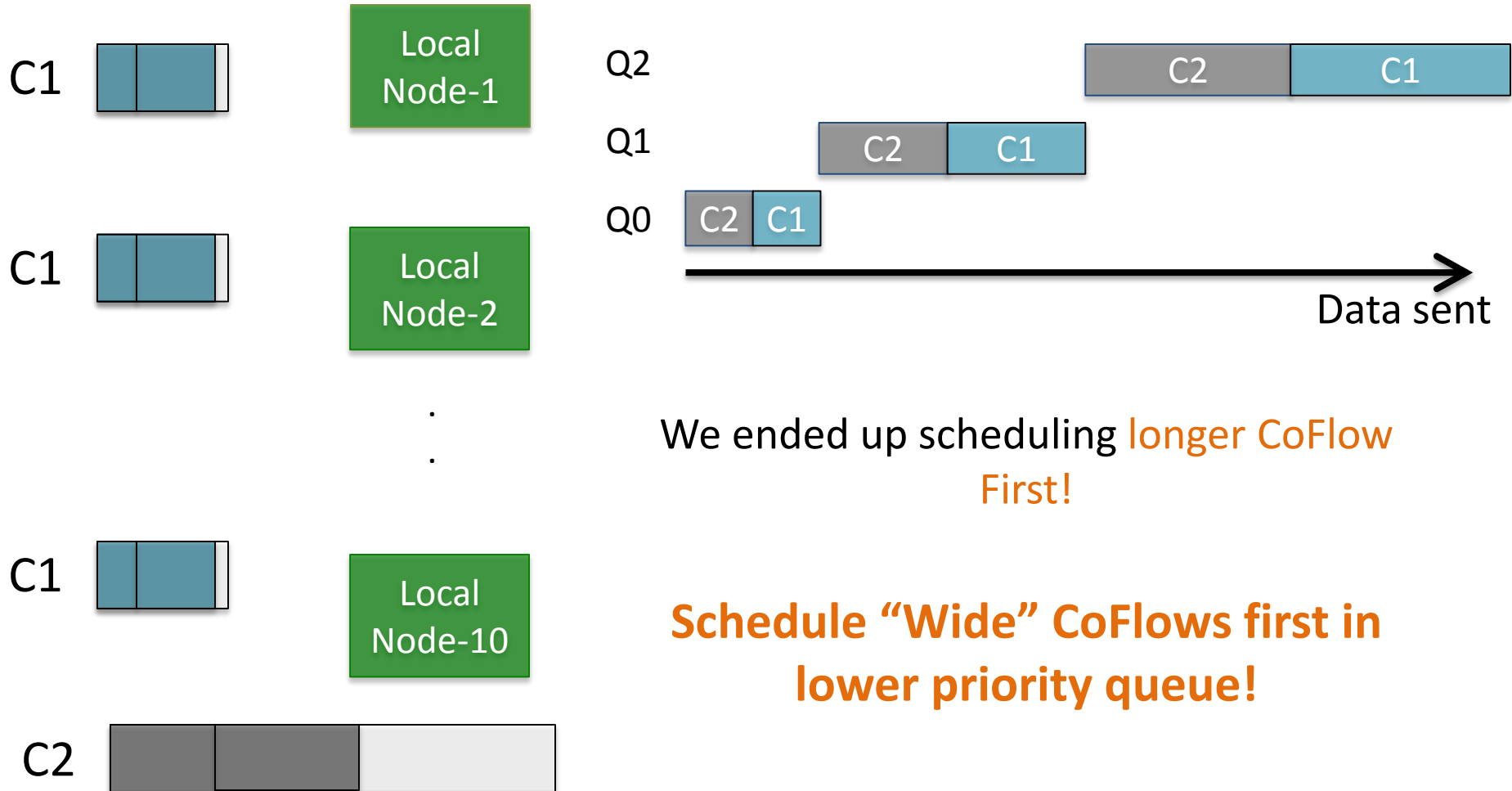## Improving over FIFO in Intra-queue CoFlow Scheduling

Graviton Observation

Thin CoFlows are short too!

Intra-queue scheduling: ~~FIFO~~
Use Thin CoFlow First

Graviton: Observation 2

Thin CoFlow First Always good? → No!

C1

Local Node-1

C1

Local Node-2

.
.
.

C1

Local Node-10

C2

Q2   C2   C1

Q1   C2   C1

Q0   C2   C1

Data sent

We ended up scheduling longer CoFlow First!

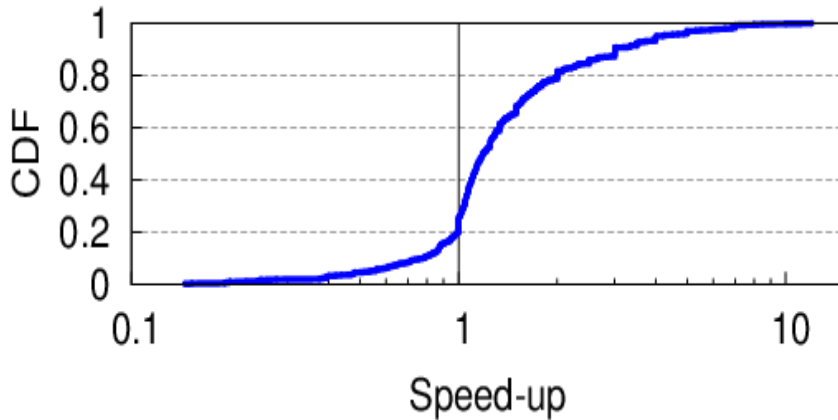**Schedule "Wide" CoFlows first in lower priority queue!**

Graviton Scheduling Summary

- High priority queues: Thin CoFlow First

- Low priority queues: Wide CoFlow First
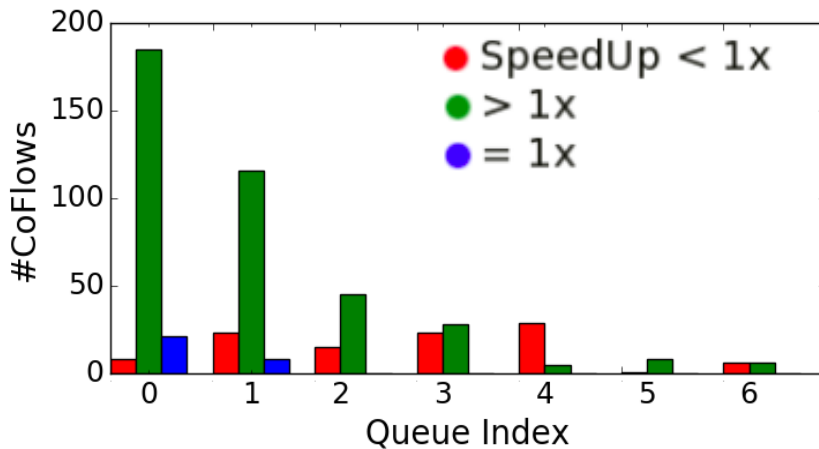
- Lowest priority queue: FIFO

Evaluation Methodology

1. CoFlow-level Simulations
2. Trace: 526 CoFlows, 150 nodes
3. Min. CoFlow size: 1MB
4. Max. CoFlow: 1+ TB

Evaluation
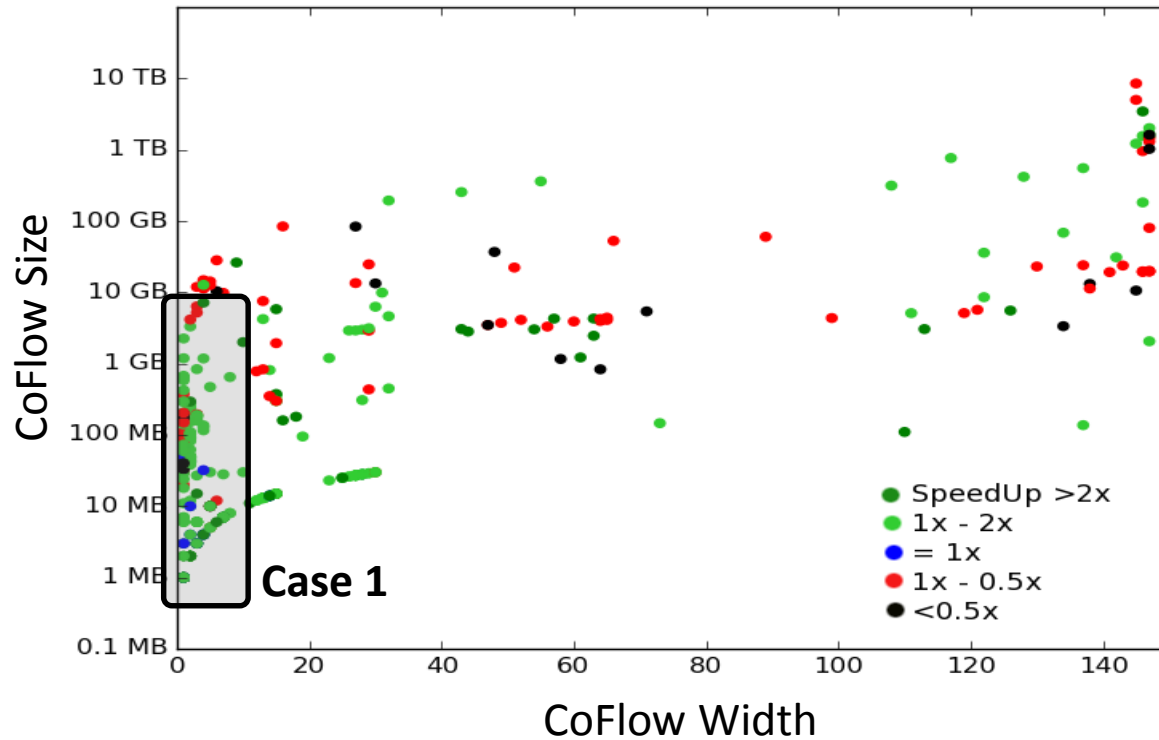
80.2% CoFlows: speed-up >1
Median = 1.25x; P90 = 8x

86.4% small CoFlows: speed-up >1

Evaluation (2)

For CoFlows with width<10 & size<10GB

~80 % have speedup >1

~10 % have speedup =1

Conclusion

- **Aalo limitation:**
  - Ignores spatial dimension
  - Local node: FIFO, which has no reminiscence of SJF
- **Graviton:**
  - Fuses spatial dimension in CoFlow scheduling
  - Different scheduling policies depending on CoFlow width
  - Evaluation: CCT improvement: 1.25x (P50), 8x (P90)

Thank you!