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 $\rightarrow$

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 the same goal

**Implication:**
CoFlow cannot finish until last flow finishes.
CoFlow Scheduling

• CCT: CoFlow Completion Time
• CoFlow scheduling problem
  • Minimize overall CoFlow Completion Time
• CoFlow scheduling problem is NP Hard
Shortest-Job-First (SJF): optimal in minimizing overall completion time

Process Scheduling

\[ P_3 \quad P_2 \quad P_1 \]

First

Last

\[ P_1 < P_2 < P_3 \]
Online Approximation to SJF

Process durations

Priority queues
(Higher Priority = more CPU time)

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
Aalo: CoFlow Queue Transition

Local Node-1

Q2 (1GB)
Q1 (100MB)
Q0 (10MB)

Local Node-2

Data sent

C2 C1

C2 C1

C2 C1

C2 C1
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

**Intra-queue scheduling:** FIFO
Use Thin CoFlow First
Graviton: Observation 2

Thin CoFlow First Always good? \(\rightarrow\) No!

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!