Big Data Framework Interference In Restricted Private Cloud Settings

Stratos Dimopoulos
Prof. Chandra Krintz
Prof. Rich Wolski

Computer Science Department
University of California Santa Barbara

IEEE BigData 2016
Small Clouds. Multi-Analytics.

• Analytics now part of smaller clouds!
• Resource Managers (Yarn, Mesos)
• Workflow Managers (ex: Musketeer)
• Batch + Streaming
Offers (Slave₁, CPU, Mem; ...; Slaveₙ, CPU, Mem)

Accept (Slaveₓ, Tasks) / Reject
Multi-tenant Performance
Hadoop Vs SparkFG

![Bar chart showing total runtime comparison between Single Tenant and Multi-Tenant for Hadoop and SparkFG.](chart.png)
Conclusions

• Absence of effective Resource revocation and admission control lead to:
  – Violation of Fair Share
  – Resource under-utilization
  – Deadlocks

• Check our paper for more results on:
  – Interference of:
    • SparkCG and Hadoop
    • Batch and Streaming
  – Startup Overheads on Mesos
Thank You! Questions?

• Stratos Dimopoulos (stratos@cs.ucsb.edu)
• The UCSB Lab for Research on Adaptive Computing Environments (RACELab)