Mufasa: Robust Meta-Workflow Management

Ben Lyons and Douglas Thain, University of Notre Dame

eScience 2022 Conference, Salt Lake City

This work was supported in part by NSF grant OCI-1931348

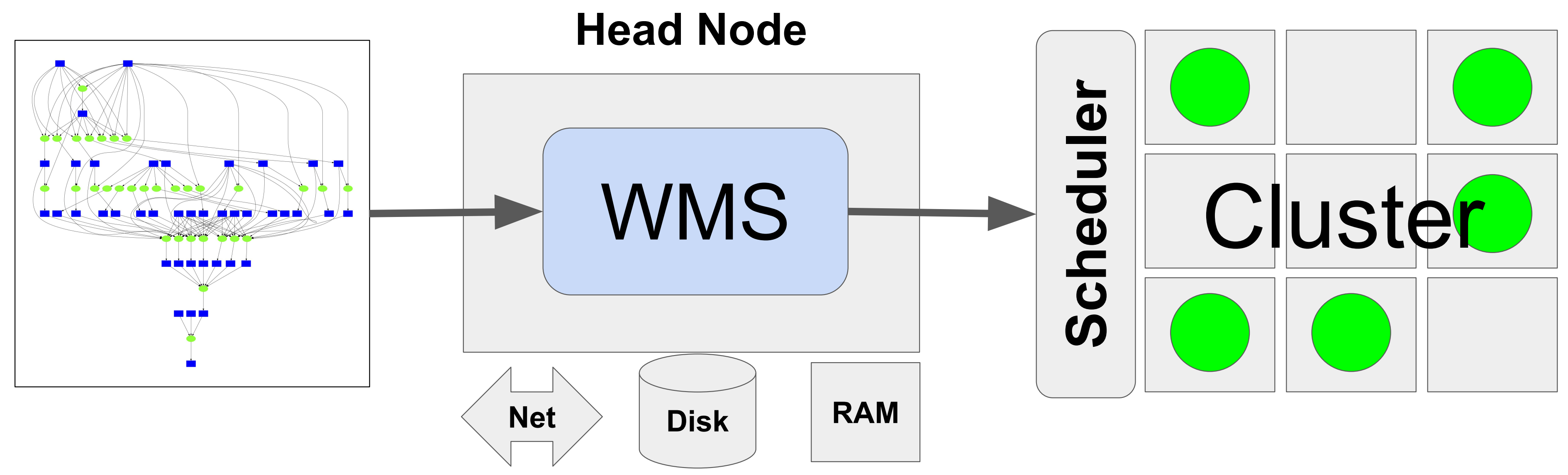
CCTools

http://cctools.readthedocs.io

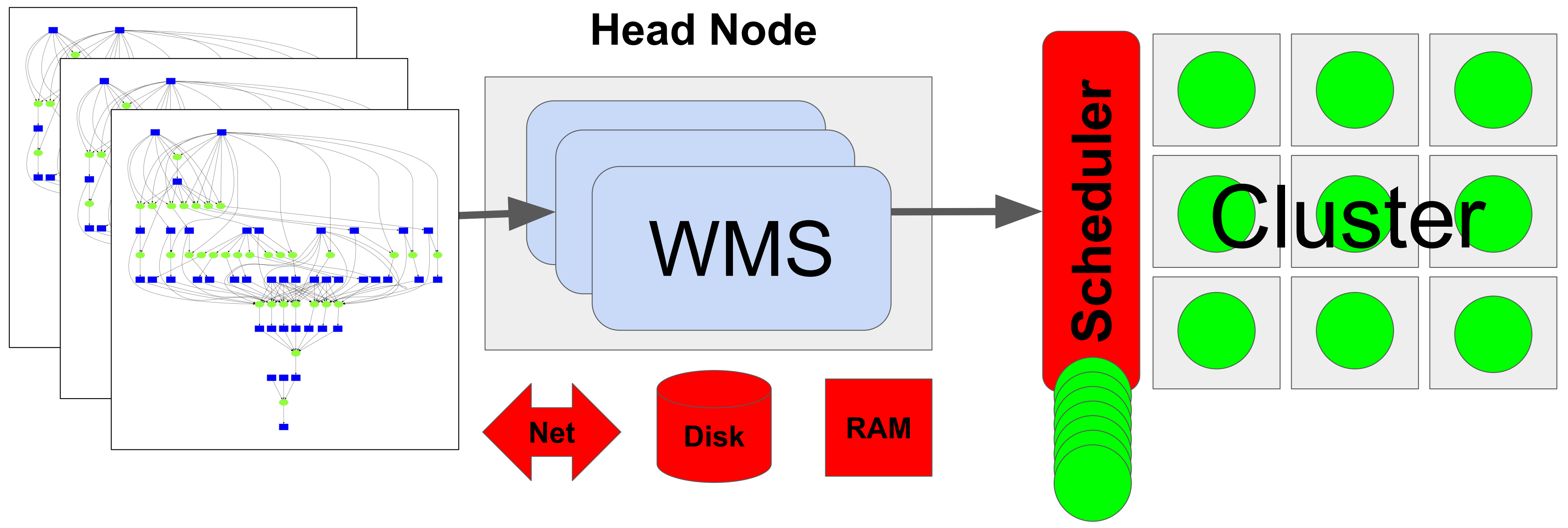
Abstract

Workflow management systems are commonly used in scientific computing to process large amounts of data in parallel. These workflows often interact with a batch system to allocate resources on a cluster. However, a research campaign rarely relies on one workflow alone: rather, it is common to run an ensemble of workflows to accomplish a large quantity of work. Managing an ensemble can be surprisingly challenging because each WMS consumes resources on both the head node and the cluster, and it is commonly the case that the total resource consumption is unknown until runtime. As a result, the naive approach of processing data as soon as it arrives may create a stampede of workflows that overwhelms either the head node or the cluster. This necessitates a meta-workflow scheduling algorithm that manages the ensemble and prevents stampeding. We have created Mufasa as a prototype meta-workflow manager to address these resource management issues. We evaluate the effectiveness of Mufasa by stressing different resources, and show that resource consumption does not exceed any of the system limits.

A WMS Manages One Workflow

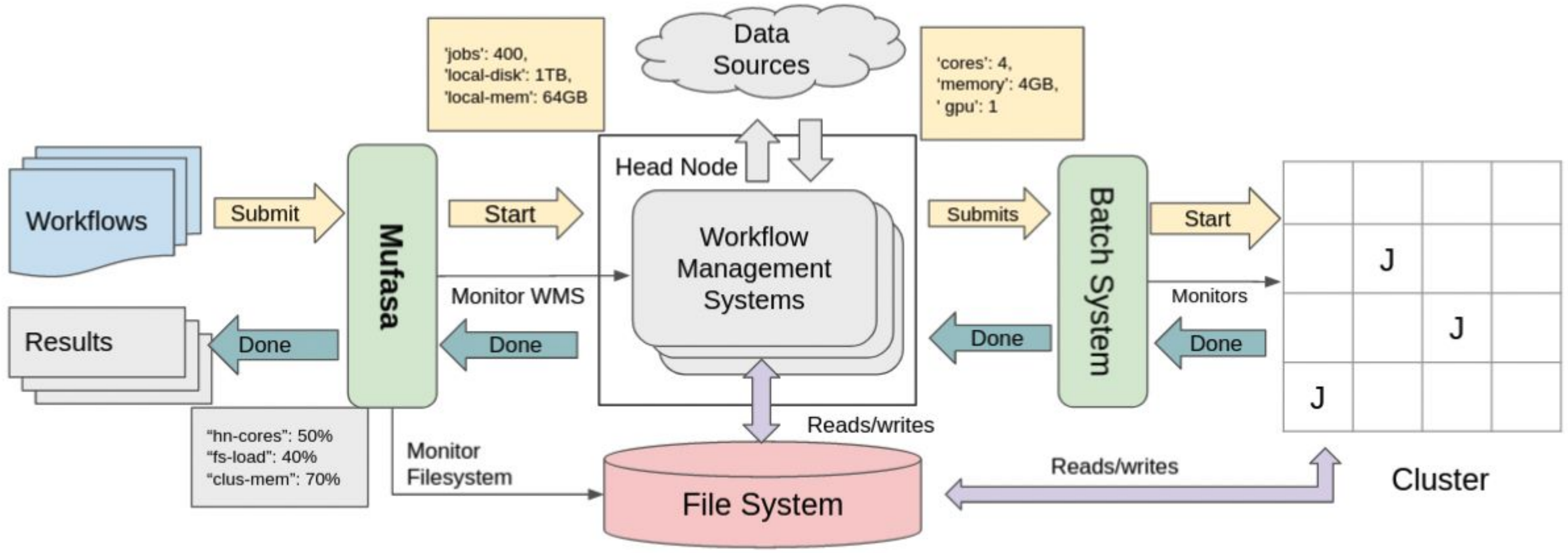


What Happens with Many Workflows?

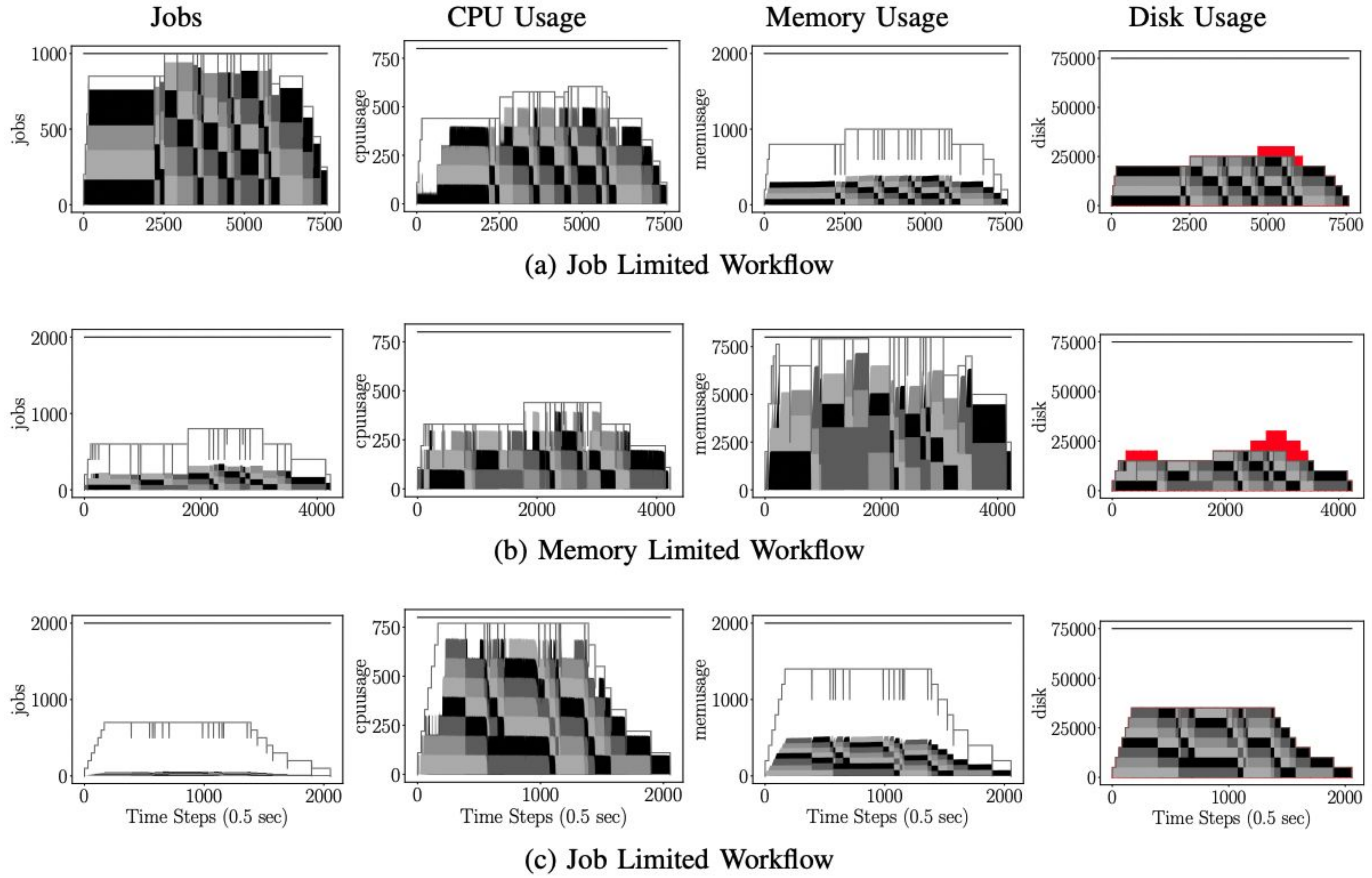


Overloading the Head Node and Scheduler (Even Deadlock!)

Mufasa is a Scheduler for Many Workflows



Robust Global Resource Management



(a) Job Limited Workflow

(b) Memory Limited Workflow

(c) Job Limited Workflow