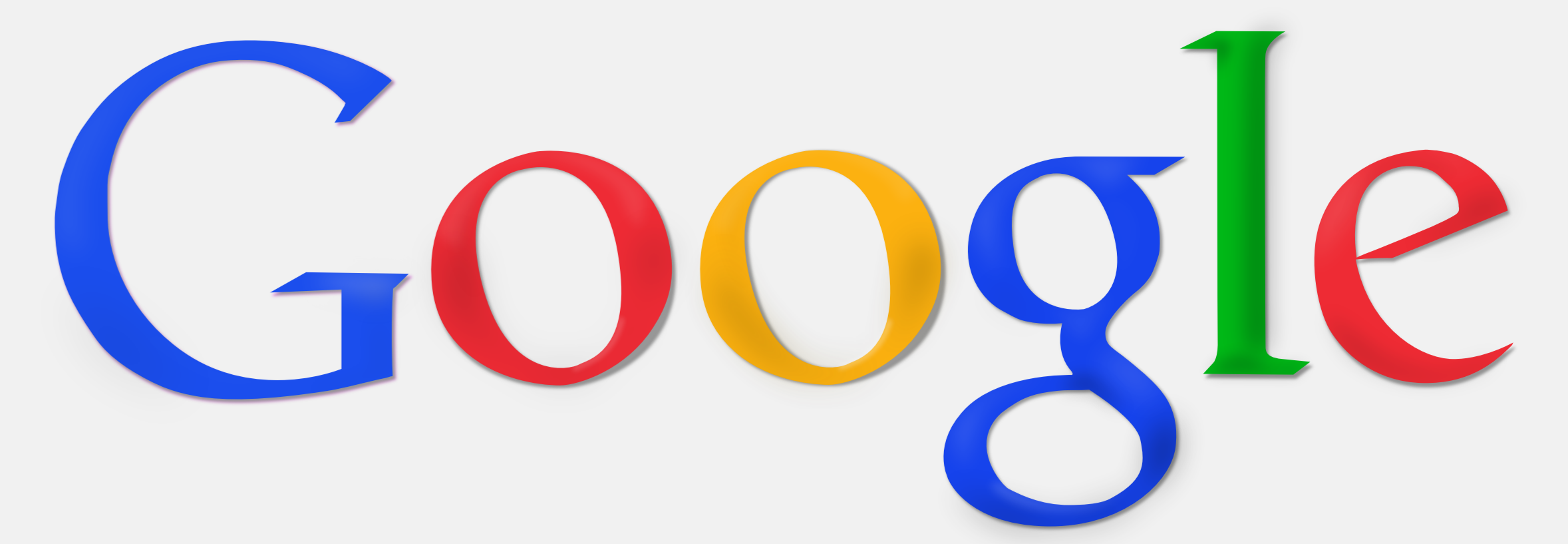


Omega: flexible, scalable schedulers for large compute clusters

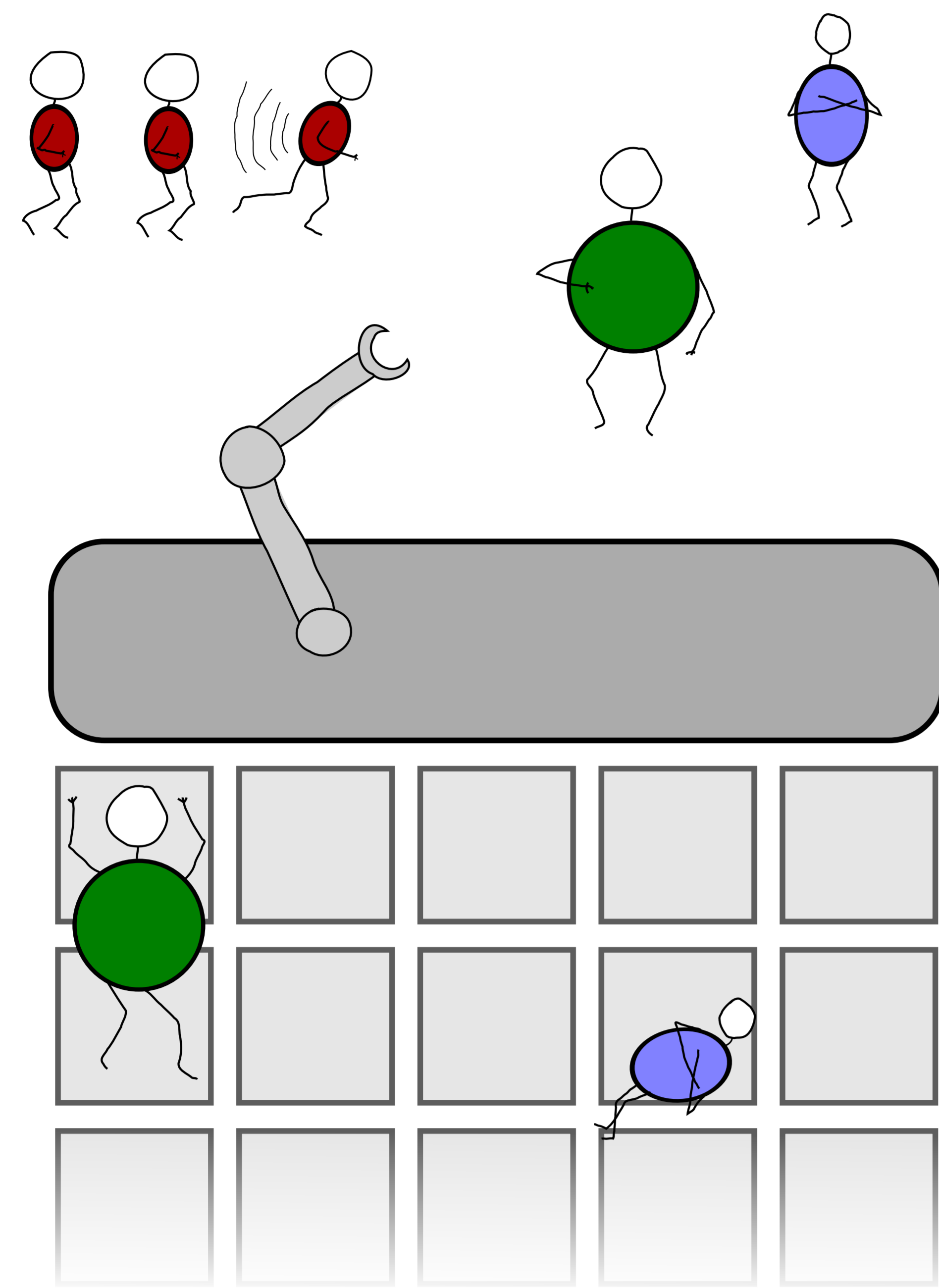


Malte Schwarzkopf
Cambridge Computer Laboratory

Andy Konwinski
University of California, Berkeley

Michael Abd-El-Malek, John Wilkes
Google, Inc.

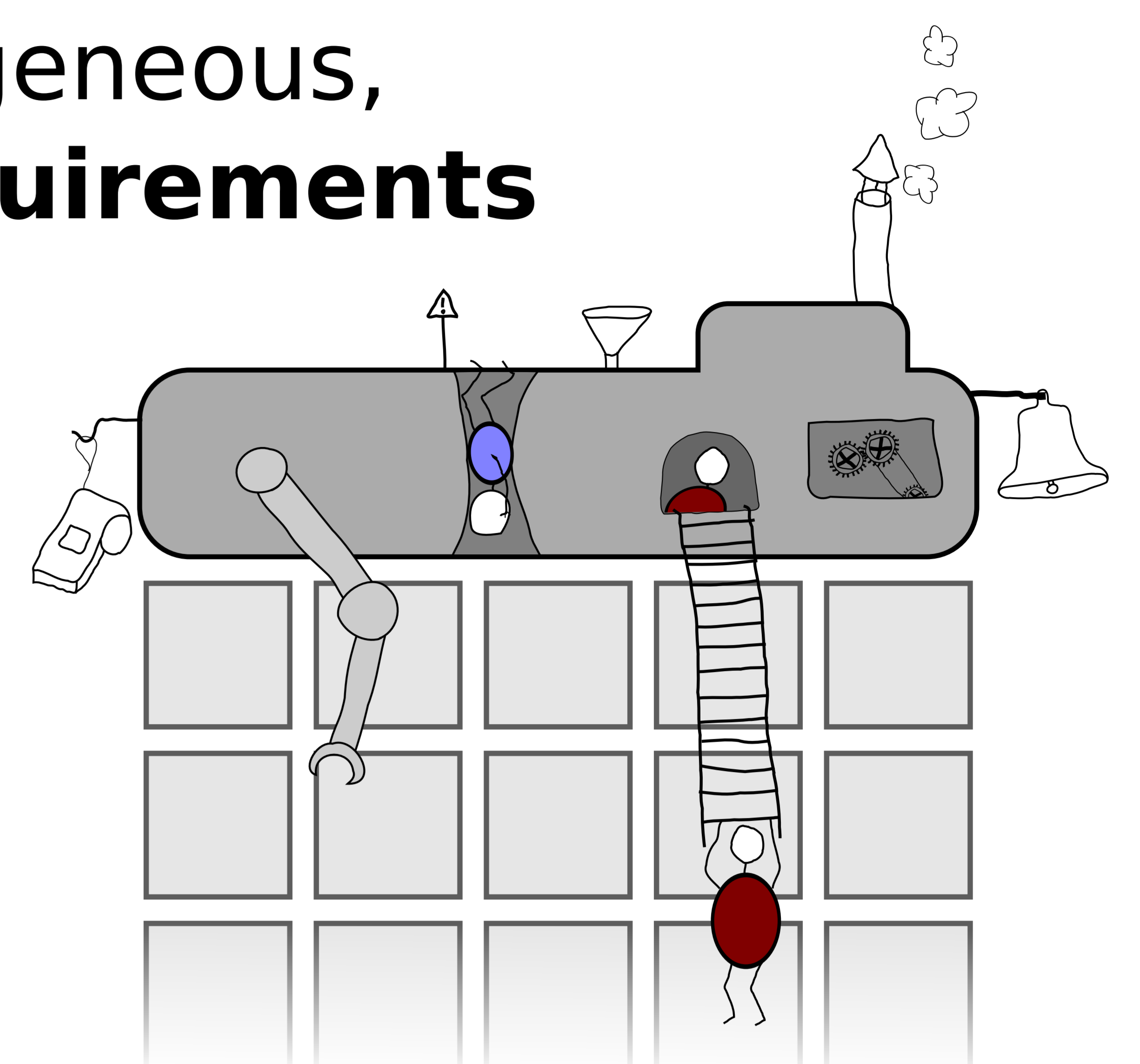
THE PROBLEM



Shared cluster workloads are heterogeneous, and have **different scheduling requirements**

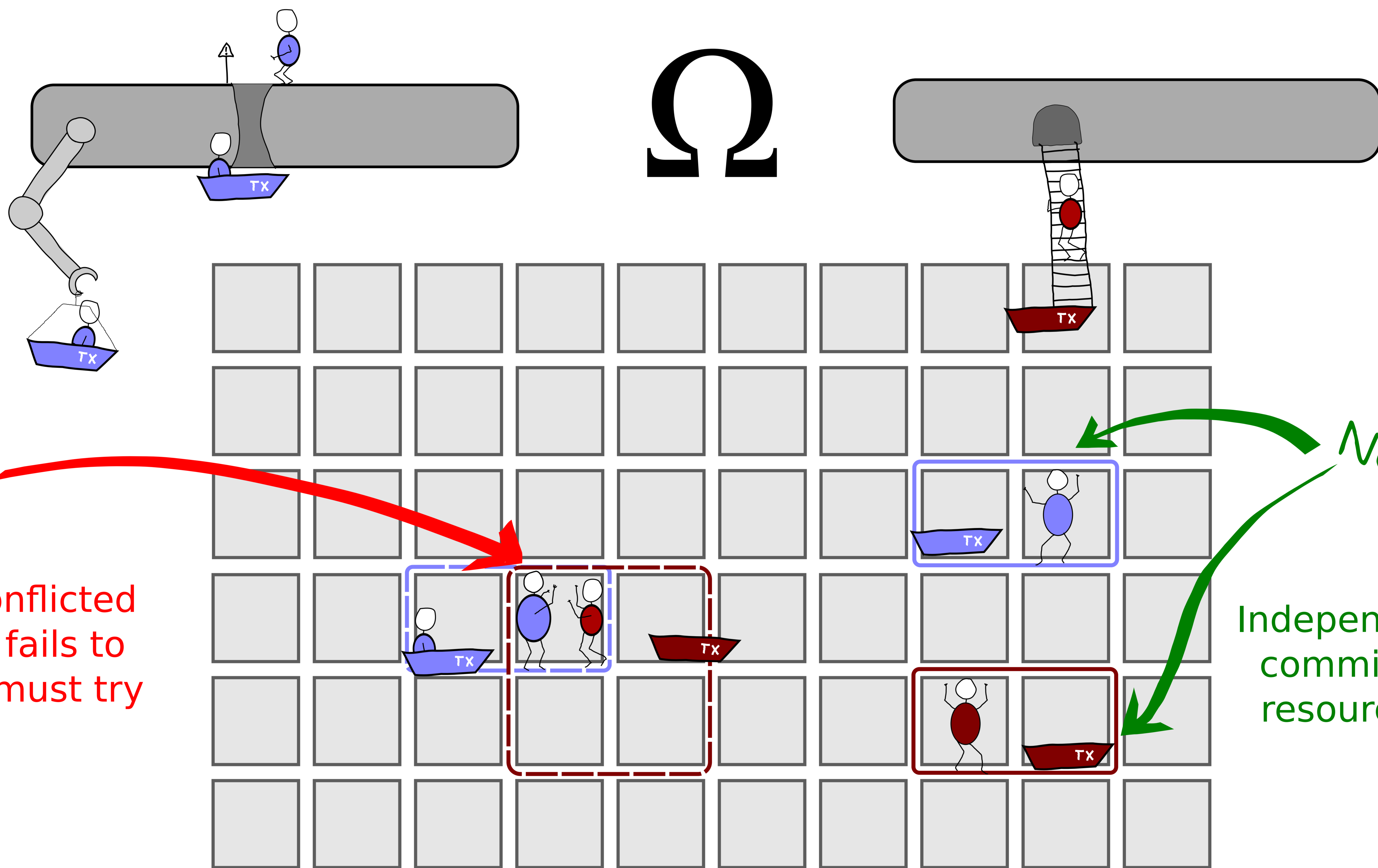
hence

over time, monolithic cluster schedulers become **too complex**.



Independent schedulers with optimistically concurrent transactions against shared state

THE SOLUTION



Conflict

One of the conflicted transactions fails to commit, and must try again

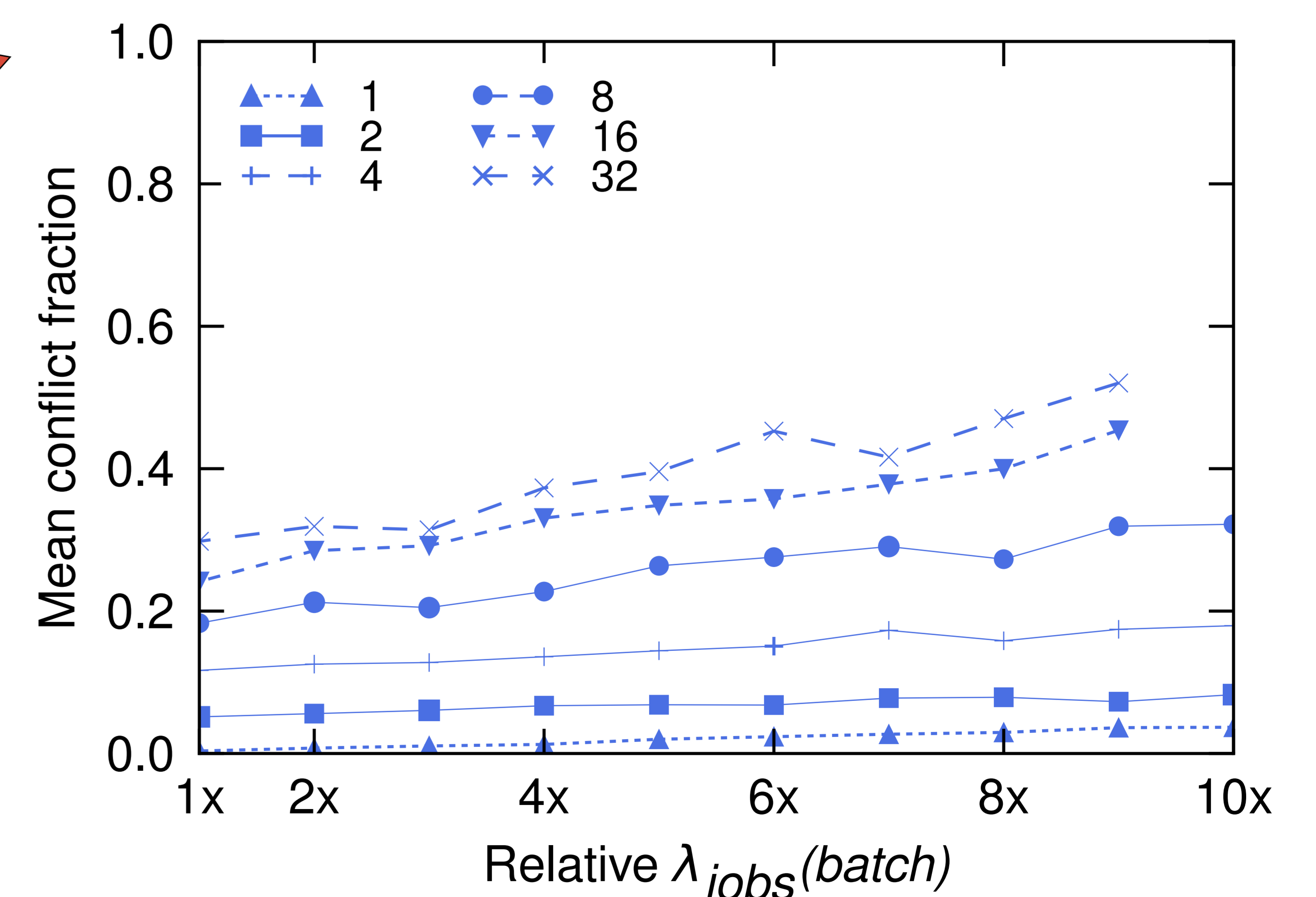
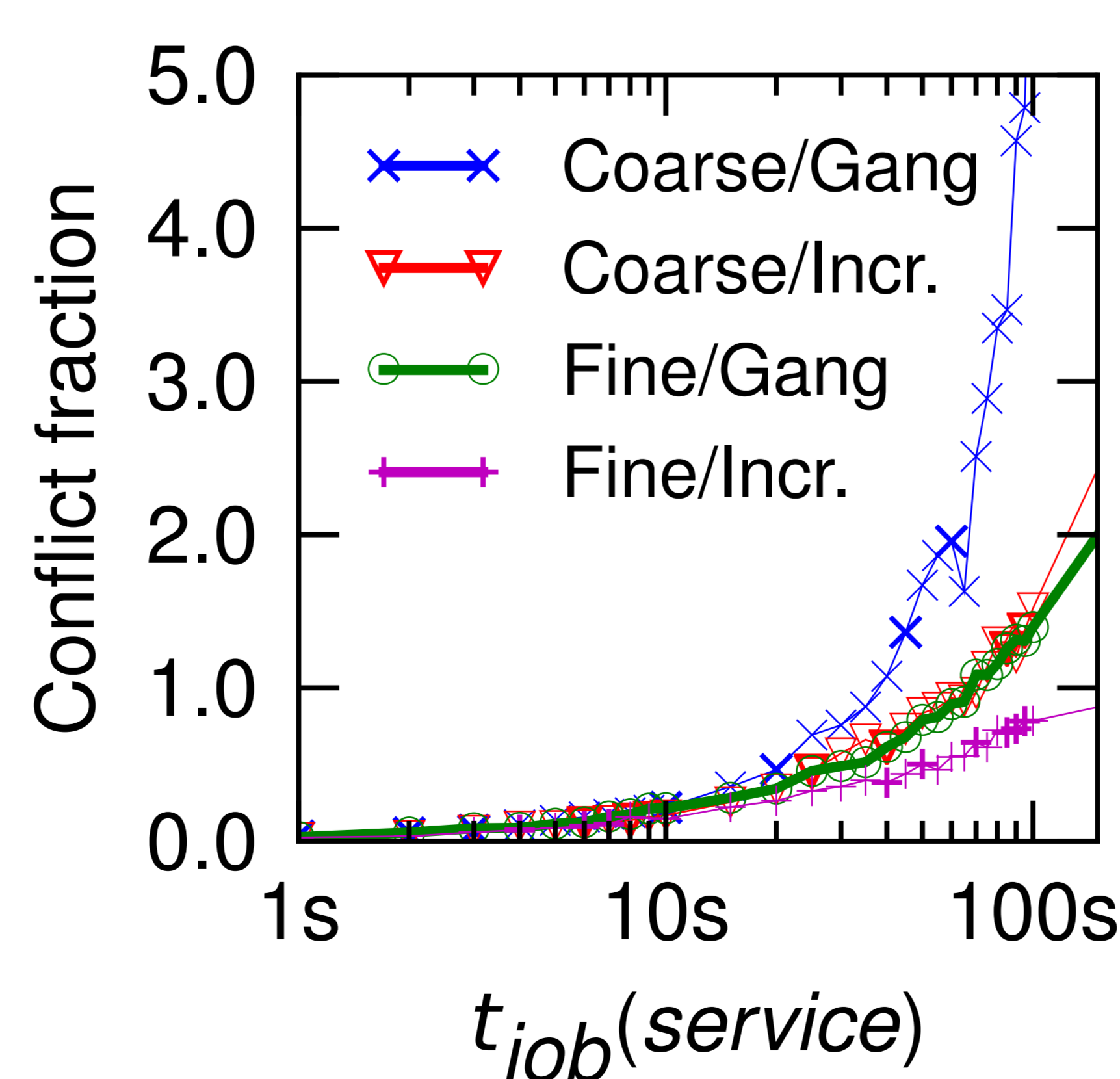
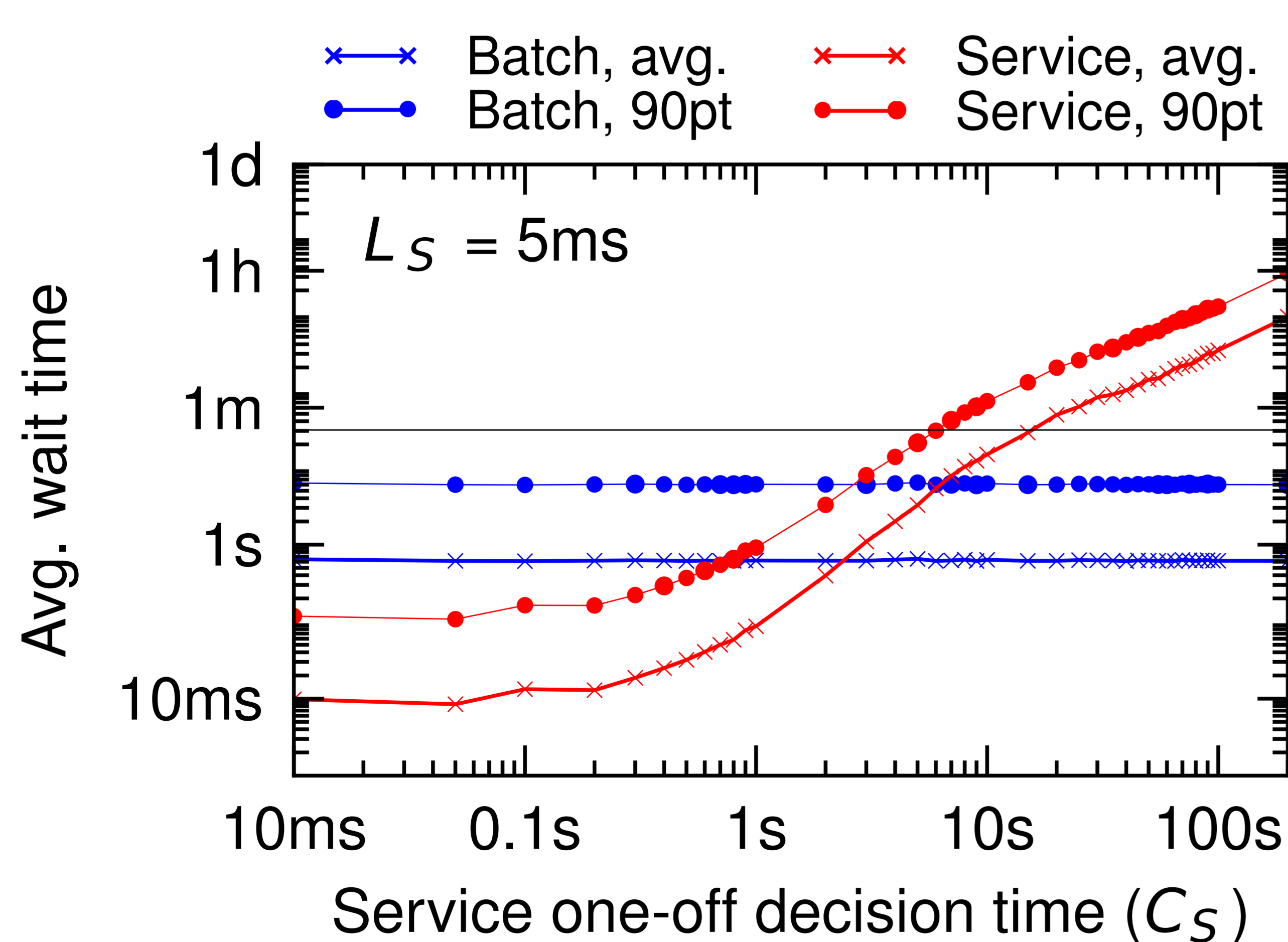
No conflict

Independent transactions commit in parallel, and resources are reserved

RESULTS

How often do conflicts occur?

How many schedulers can we have?



Decision times up to 10s per job viable for current workloads; incremental TX and optimizations help.

32 schedulers at 10x current workload still viable!