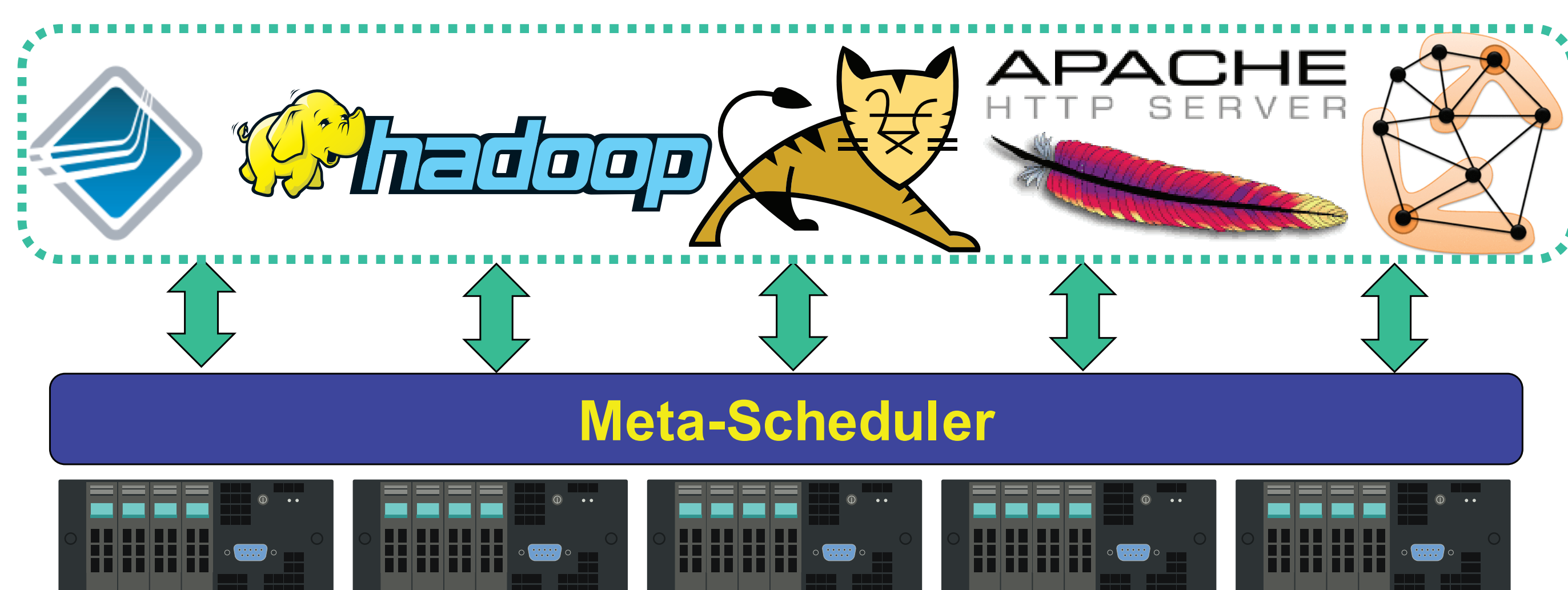


DYNAMIC CHANGE IN SCHEDULING OF HETEROGENEOUS SERVICES

Alexey Tumanov, James Cipar, Elie Krevat, Michael Stroucken, Greg Ganger (CMU),
Mike Kozuch (Intel), Matei Zaharia (UCB), Ion Stoica (UCB)

OVERVIEW

- Large clusters shared by varied workloads
 - Batch frameworks, like Hadoop
 - Elastic services, like web frontends
 - Highly constrained tasks, like experiments
- Need scheduling substrate serving all
 - Exploiting app-specific context
 - Adapting to changes in demand



TOOL: CHANGE TO FIT NEW REQUEST

- Some requests are difficult
 - Big or highly-constrained
- May need to change current allocations
 - Migrate VMs, kill/restart, checkpoint, etc.
- Often many options available → pick best
 - One may not have a recent checkpoint
 - One may be stateless and easily moved
 - One may rely on data locality for efficiency
 - Etc.

INVERSE OFFERS

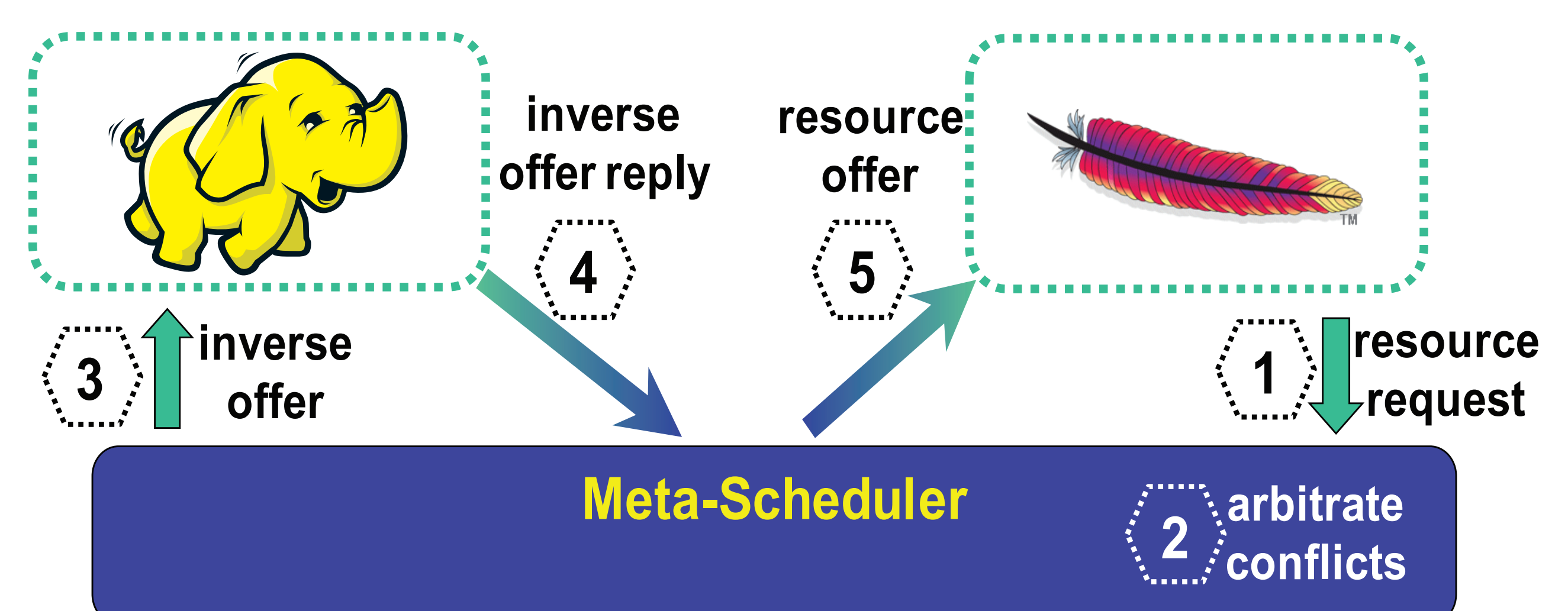
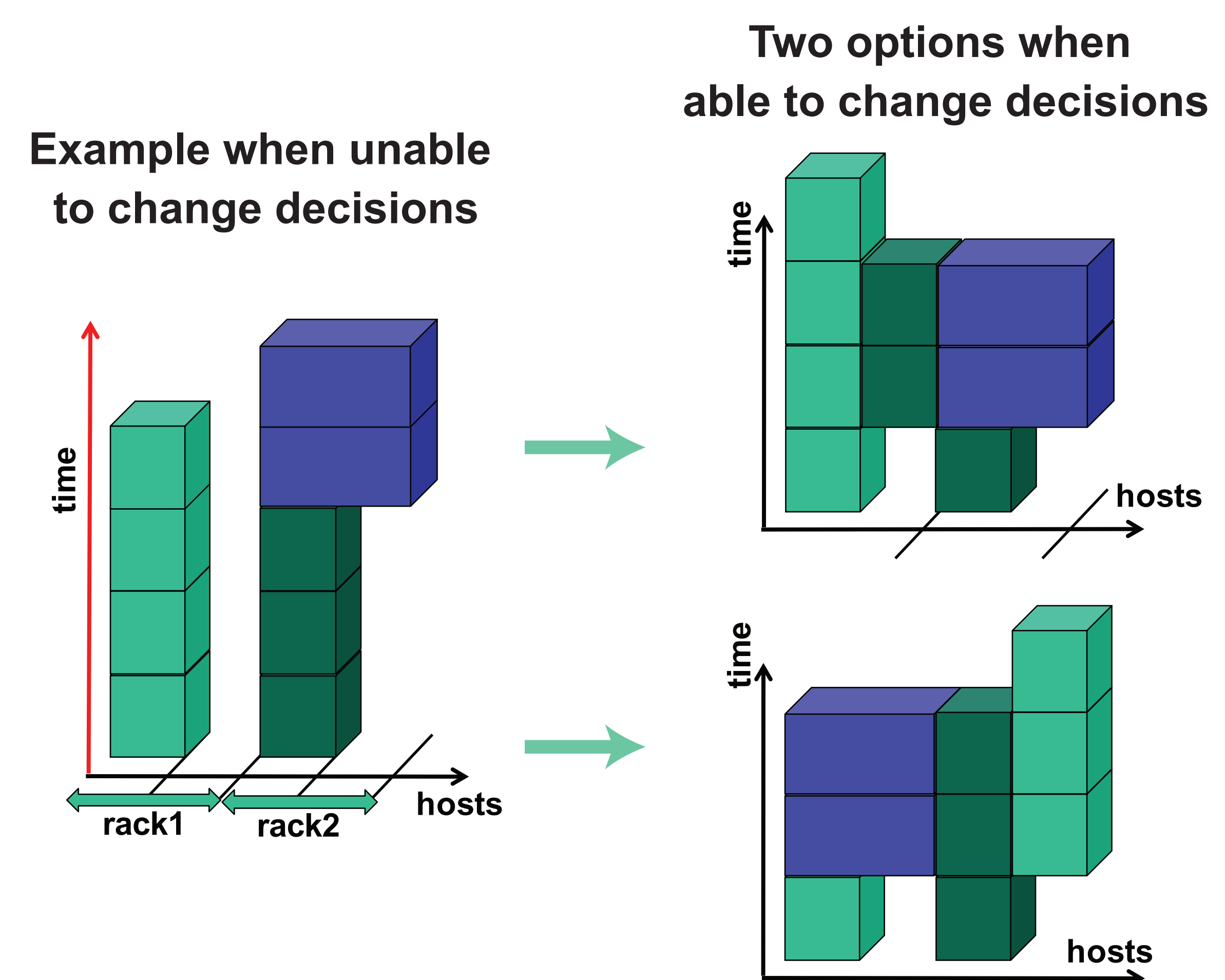
- Mechanism for two-level change decisions
 - Involve both levels and their knowledge
- Meta-scheduler describes options in inverse offer
 - Must give up X of Y where Z
 - May couple with resource offer as trade
- Framework replies with its preferred option
 - Considering its context on costs/effects

WORK IN PROGRESS

- Quantify potential benefits of ability to change
- Cost-benefit analysis of placement changes
 - Including cost of making a change
- Algorithms for making change decisions
 - Identifying factors influencing placement decisions
- Incentivizing flexibility among frameworks/services

TWO-LEVEL SCHEDULING

- Each framework/service requests resources
 - Distributes work among its machines
 - Requests/releases based on demand
- Meta-scheduler arbitrates allocations
 - Determines how much each requester gets
- Example: Resource Offers in Mesos [NSDI'2011]
 - Framework describes need
 - Meta-scheduler exposes options
 - Framework selects from among them



PRELIMINARY ANALYSIS

