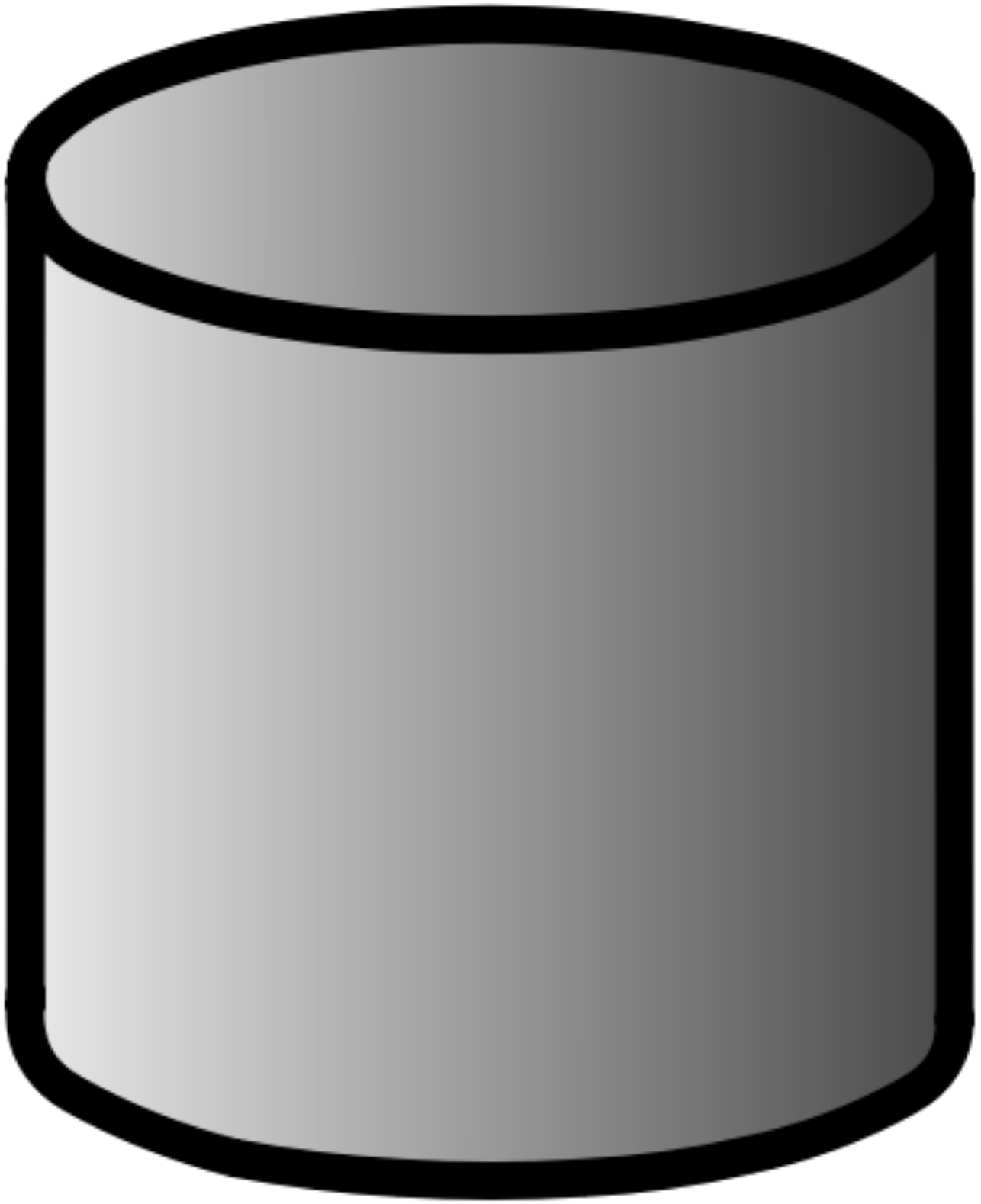
Robust Data Transformation

The Curios Case of Databases & Horses!

Alekh Jindal
CSAIL, MIT



Database

5

How do we

Start

Preparing?

Workload

Workload

Workload

Problem!

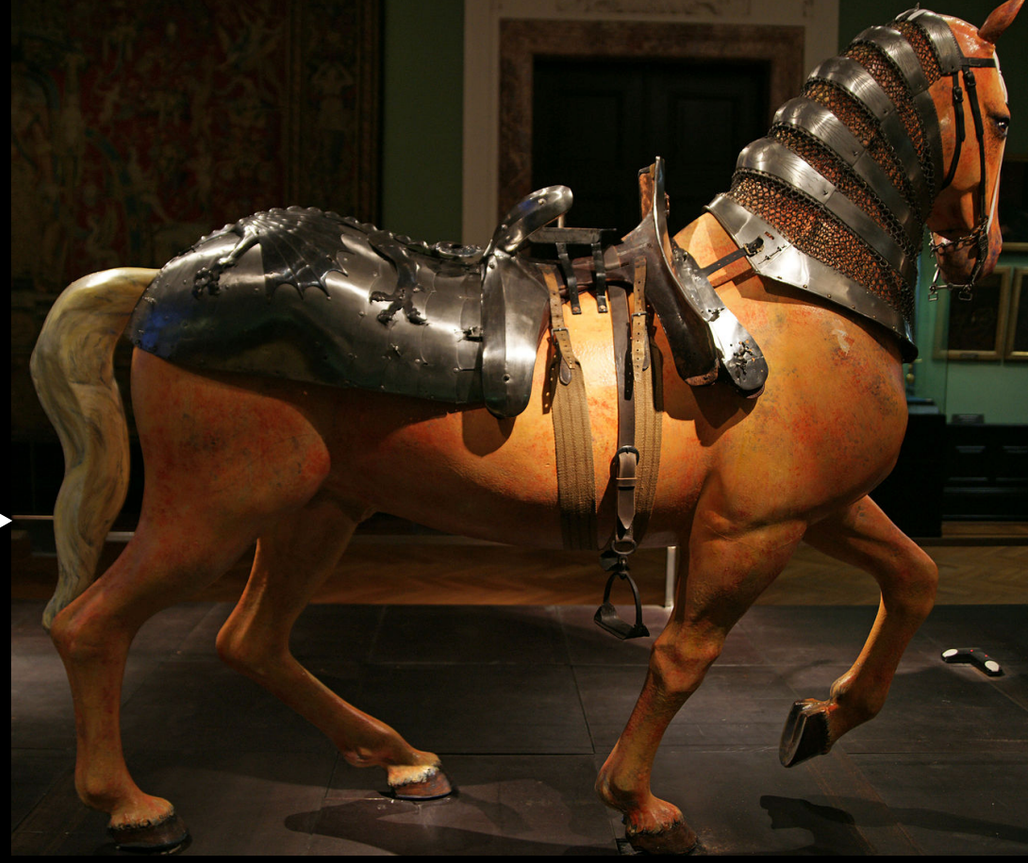
Changing Workload!

Time

Several Workloads!

No Workload!

Adapt?



Adapt?

Adapt?

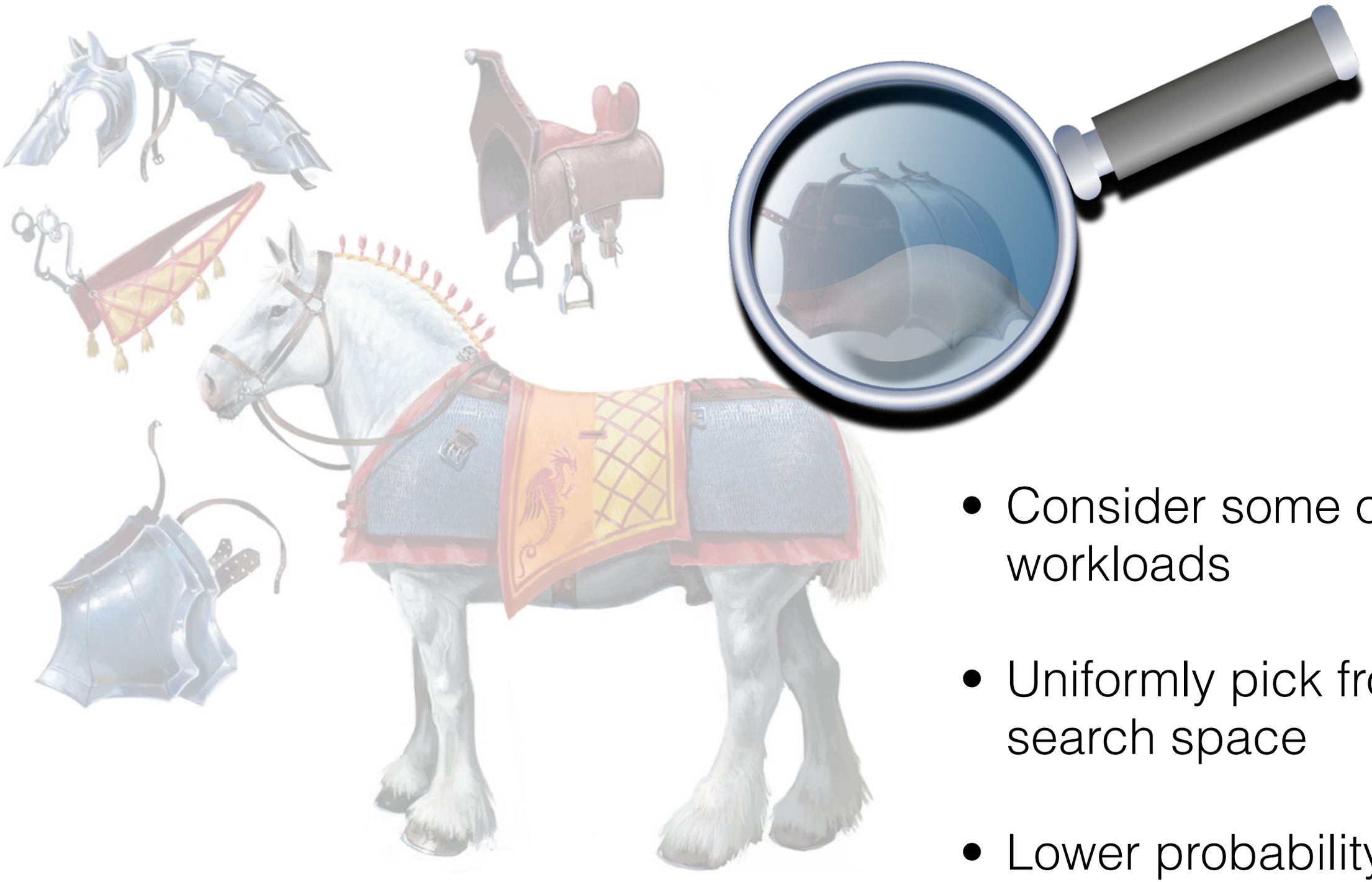
Workload-based Approach!

Robust Approach!

Prepare Exhaustively!

- Consider all possible workloads
- Could be expensive
- But, bad performance not a problem anymore

Prepare Selectively!



- Consider some of the workloads
- Uniformly pick from the search space
- Lower probability of bad performance

Prepare Partially!

- Something for all, everything for none
- No full benefits for a given workload
- Significantly better performance for several

Example: data layouts

- Prepare Exhaustively: create row, column and all hybrid (vertically partitioned) layouts

- Prepare Selectively: uniformly pick some of the layouts based on their tuple width

- Prepare Partially: create different layouts for different subsets of the data

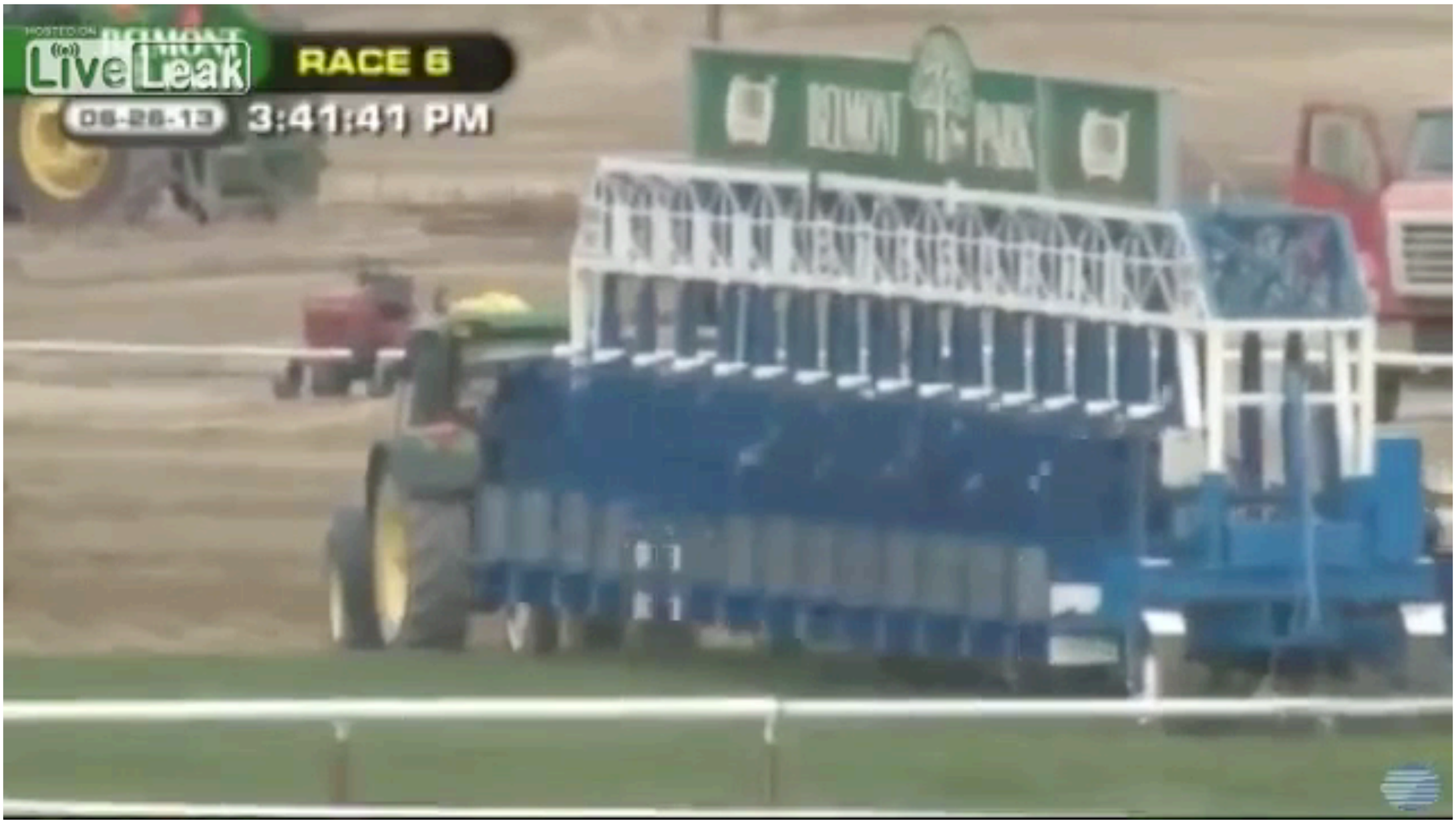
Advantages

No need to find the optimal

Can Observe Performance

Ad-hoc Workload

Performance of a system (horse) without workload



Thank You!

Alekh Jindal
people.csail.mit.edu/alekh