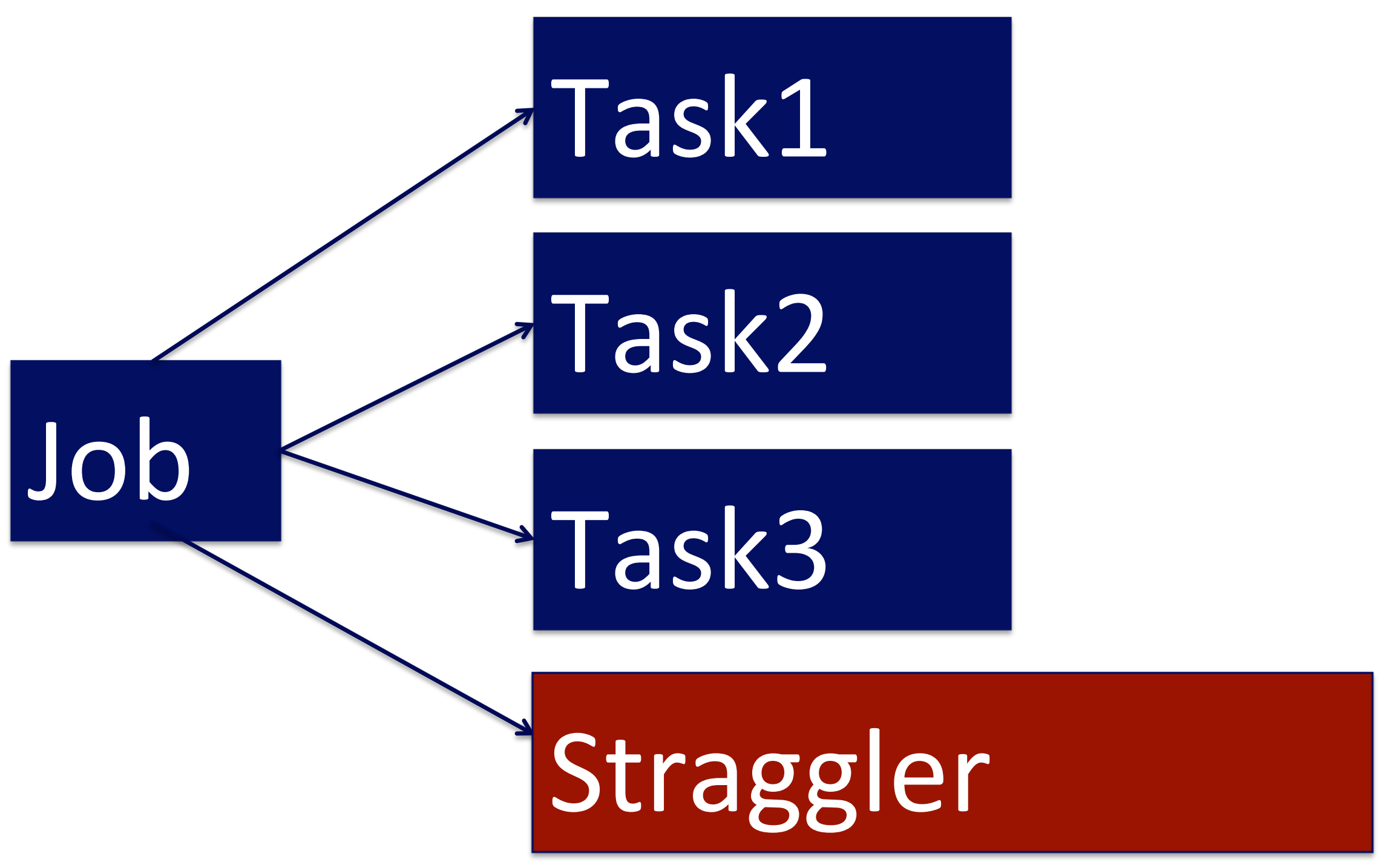
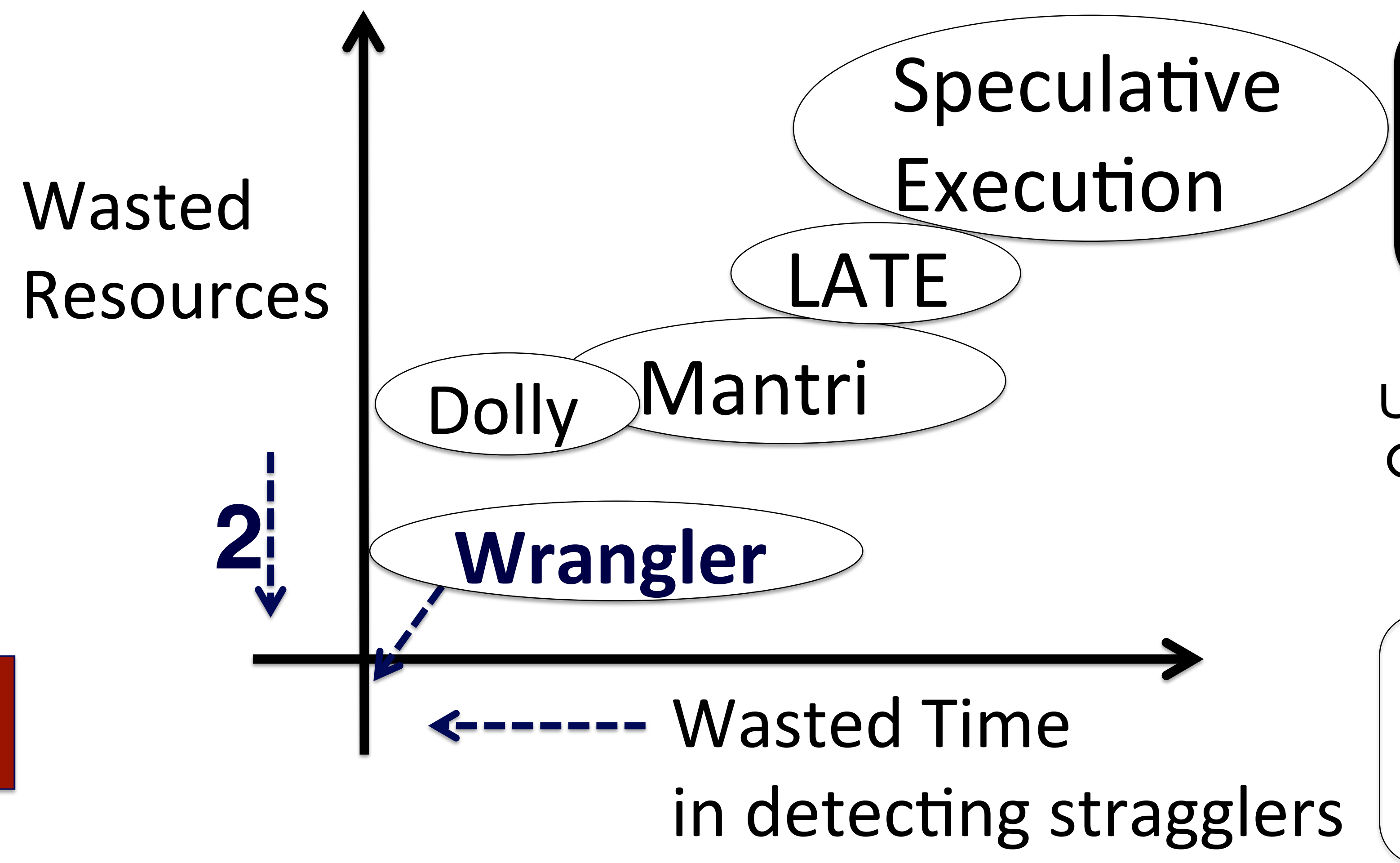
Wrangler: Predictable and Faster Jobs in Distributed Processing Systems using Machine Learning

Neeraja J. Yadwadkar, Bharath Hariharan, Ganesh Ananthanarayan, Joseph Gonzalez, and Randy Katz

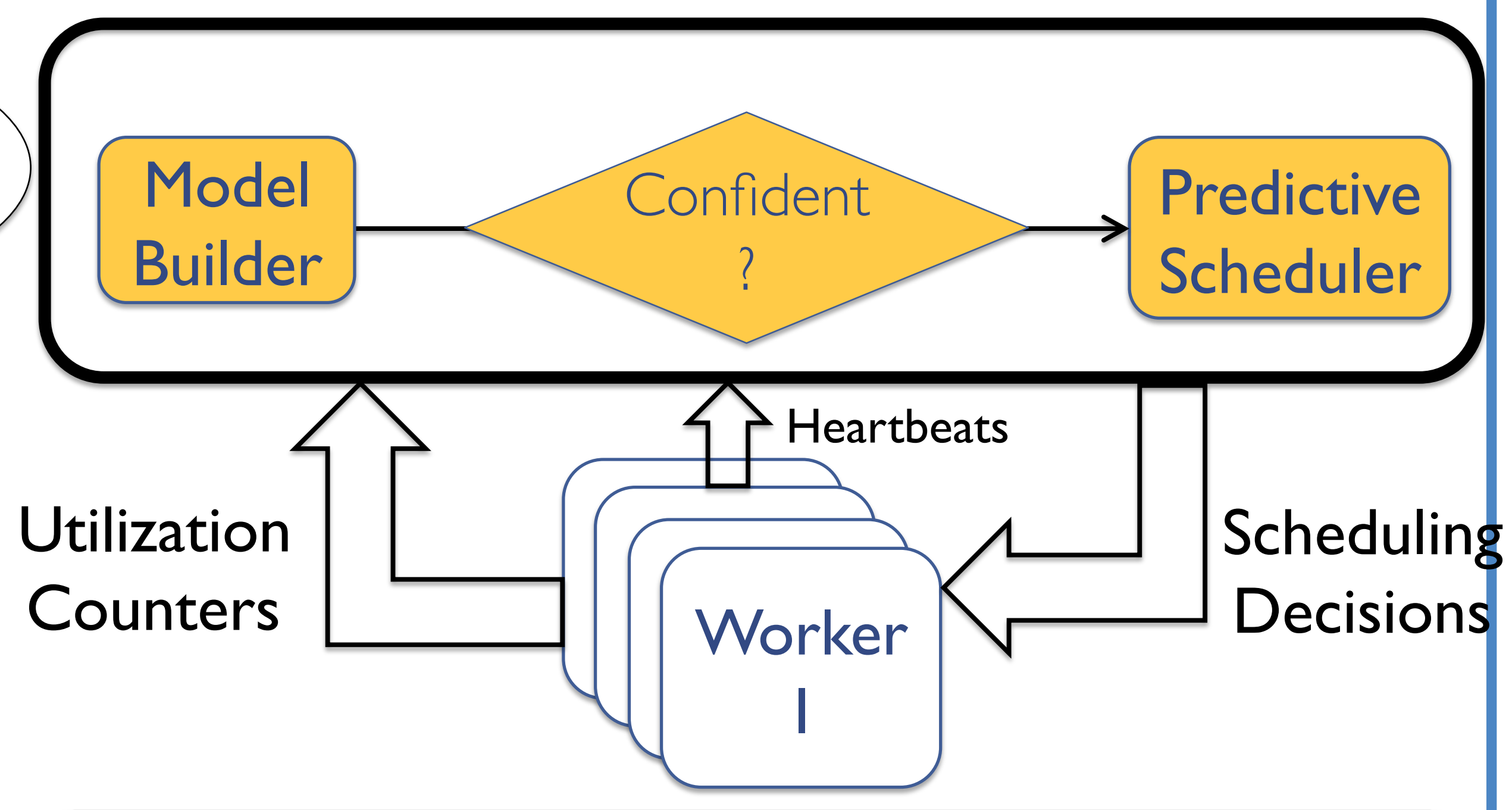
Parallel Data analytics and stragglers



Design Space

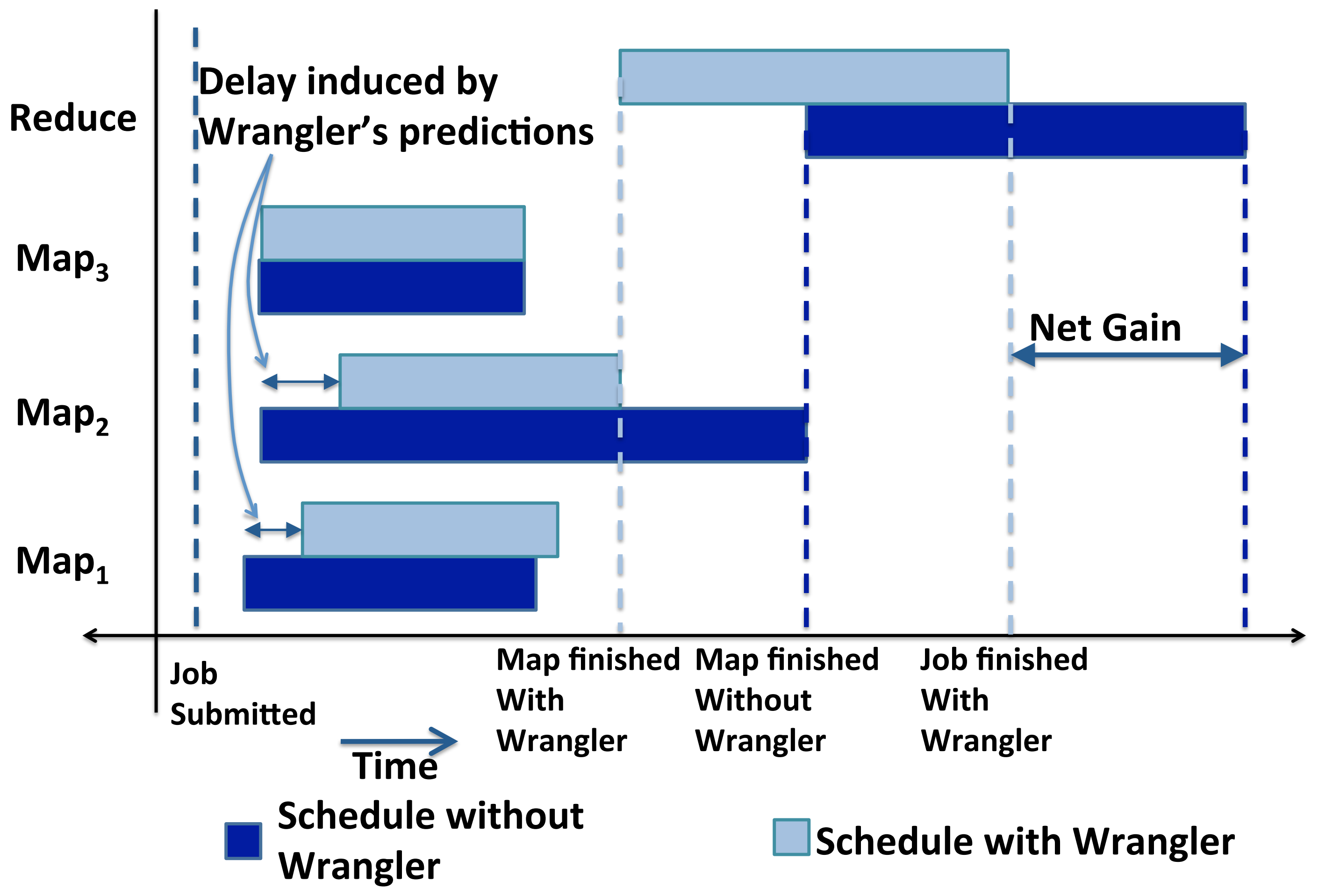


Wrangler: Architecture

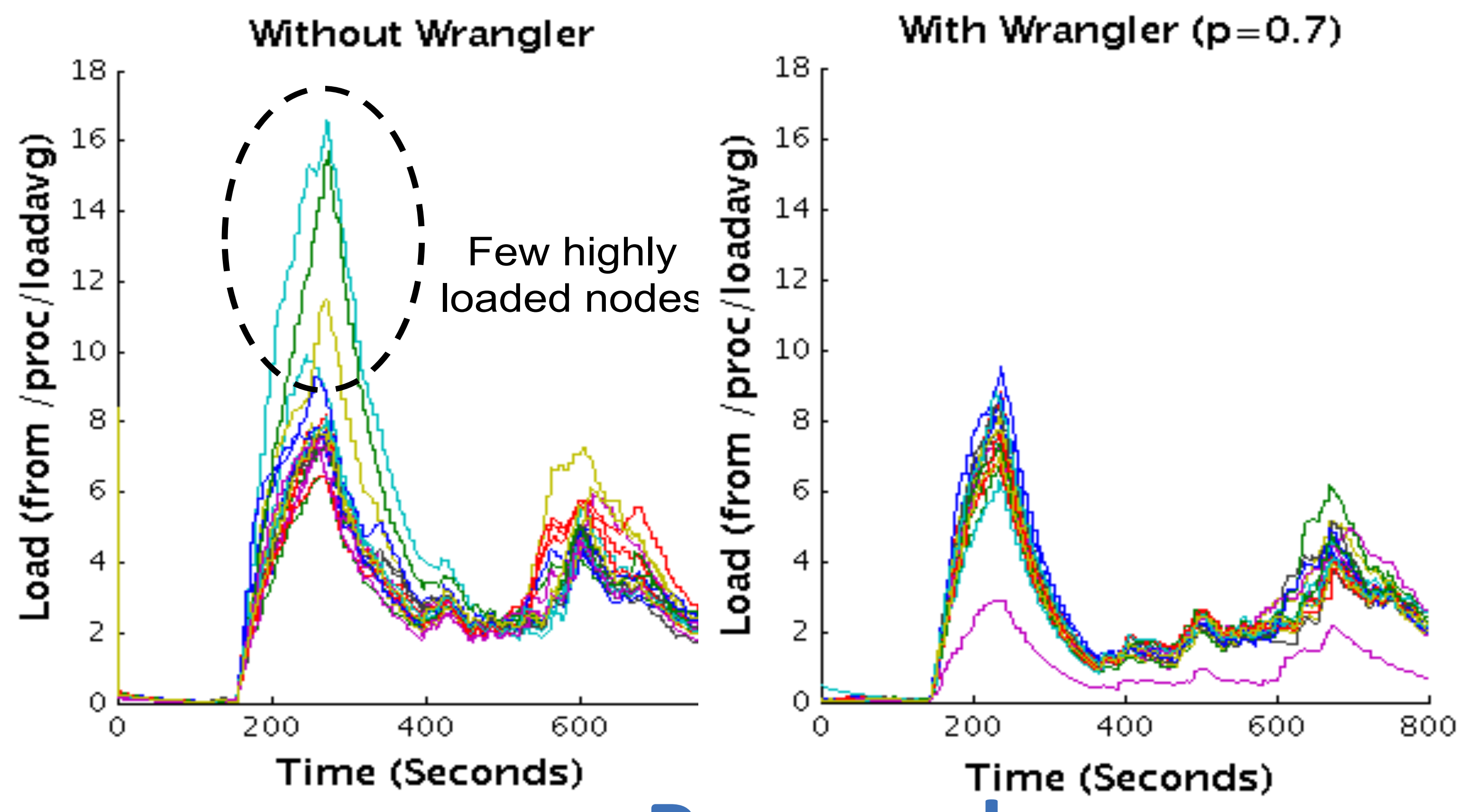


Approach: Binary Classification
Input: Perf. counters at launch time
Label: Yes/No

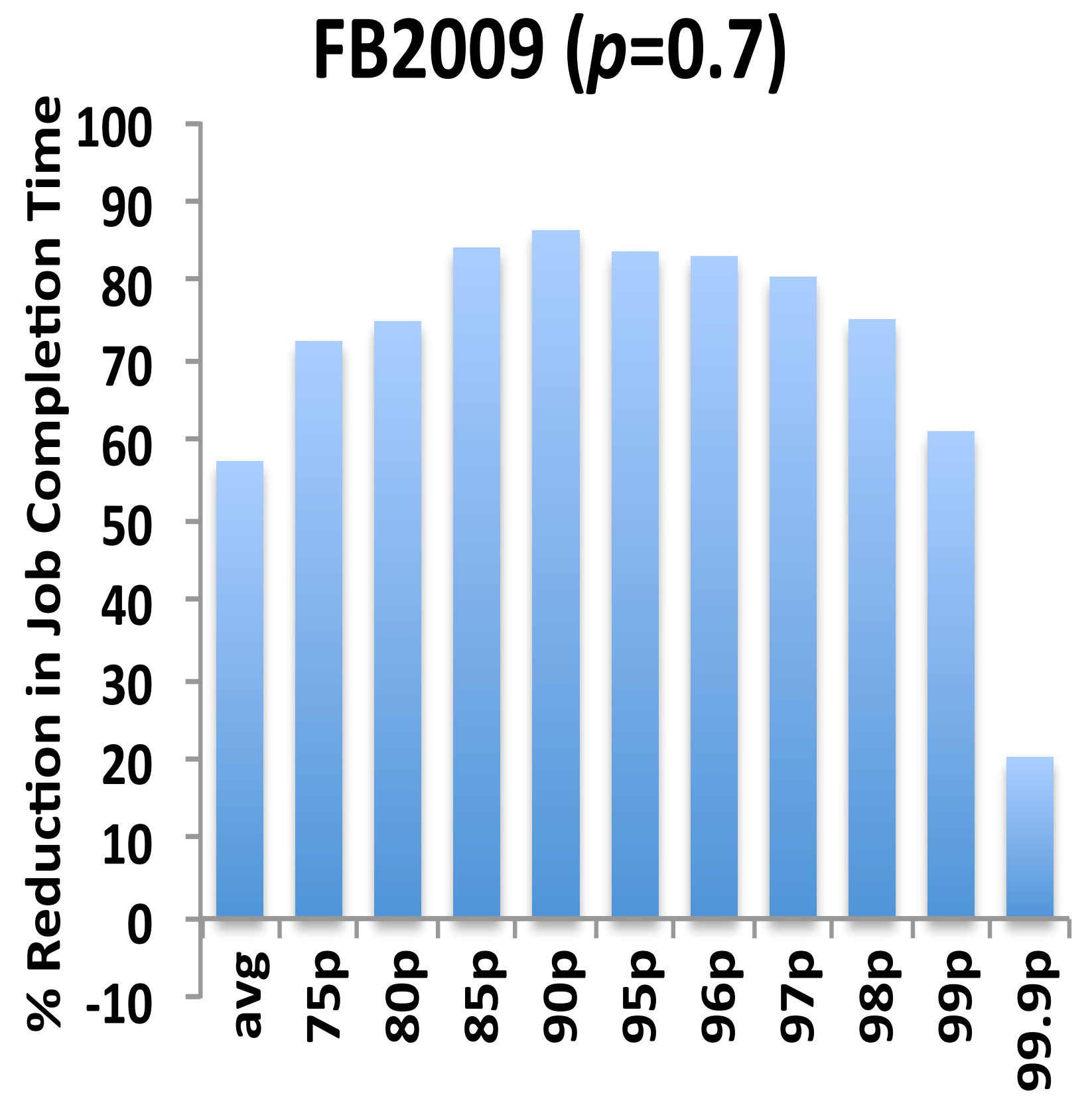
Intuition


Reduce
Map3
Map2
Map1
Delay induced by Wrangler's predictions
Net Gain
Job Submitted
Time
Map finished With Wrangler
Map finished Without Wrangler
Job finished With Wrangler
Schedule without Wrangler
Schedule with Wrangler

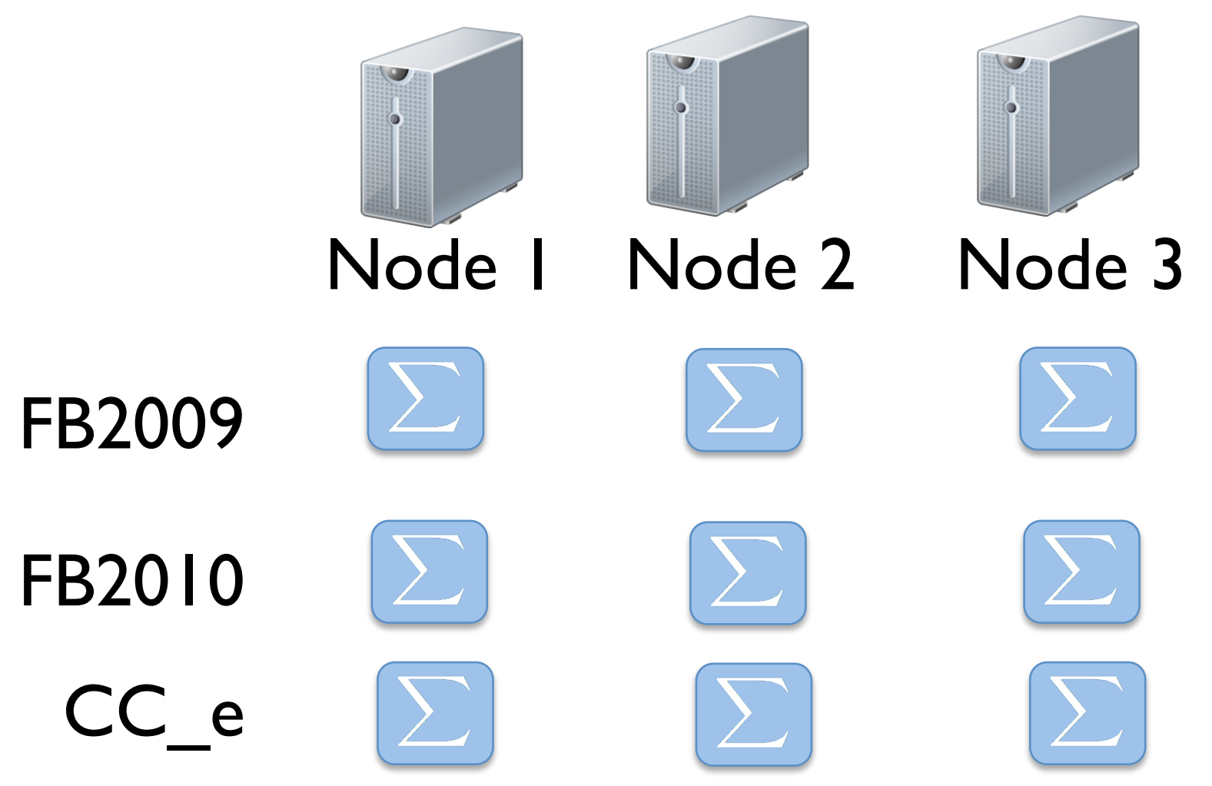
Load-Balancing



Faster Job Completion



Model Builder



Prediction Accuracy: 70-80%

Scalability!

Train too many models separately

Why? Heterogeneity across nodes and tasks

Prohibitively long data capture time

Proposal

- Underlying modeling task remains the same
- Learning from other similar tasks
  - Reduce training data capture time
  - Improve accuracy by generalizing better

Share data across nodes and workloads: Multi Task Learning

Regularized MTL [KDD'04]:

$$\mathbf{w}_t = \mathbf{w}_0 + \mathbf{v}_t$$

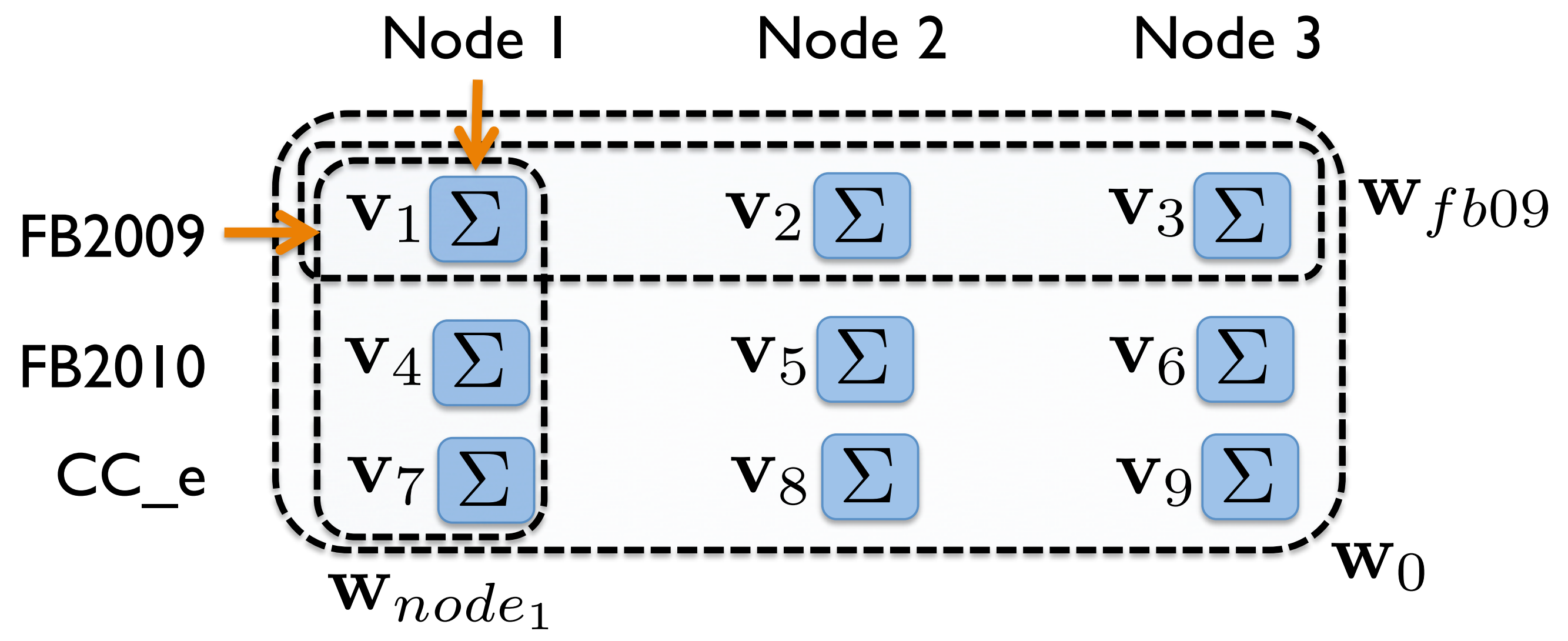
Common across all the learning tasks

Specific for a learning tasks, *t*

Our Formulation:

$$\mathbf{w}_t = \mathbf{w}_0 + \mathbf{v}_t + \mathbf{w}_{g(t)}$$

Common across the tasks in a group, denoted by g



$$\mathbf{w}_1 = \mathbf{w}_0 + \mathbf{w}_{node_1} + \mathbf{w}_{fb09} + \mathbf{v}_1$$

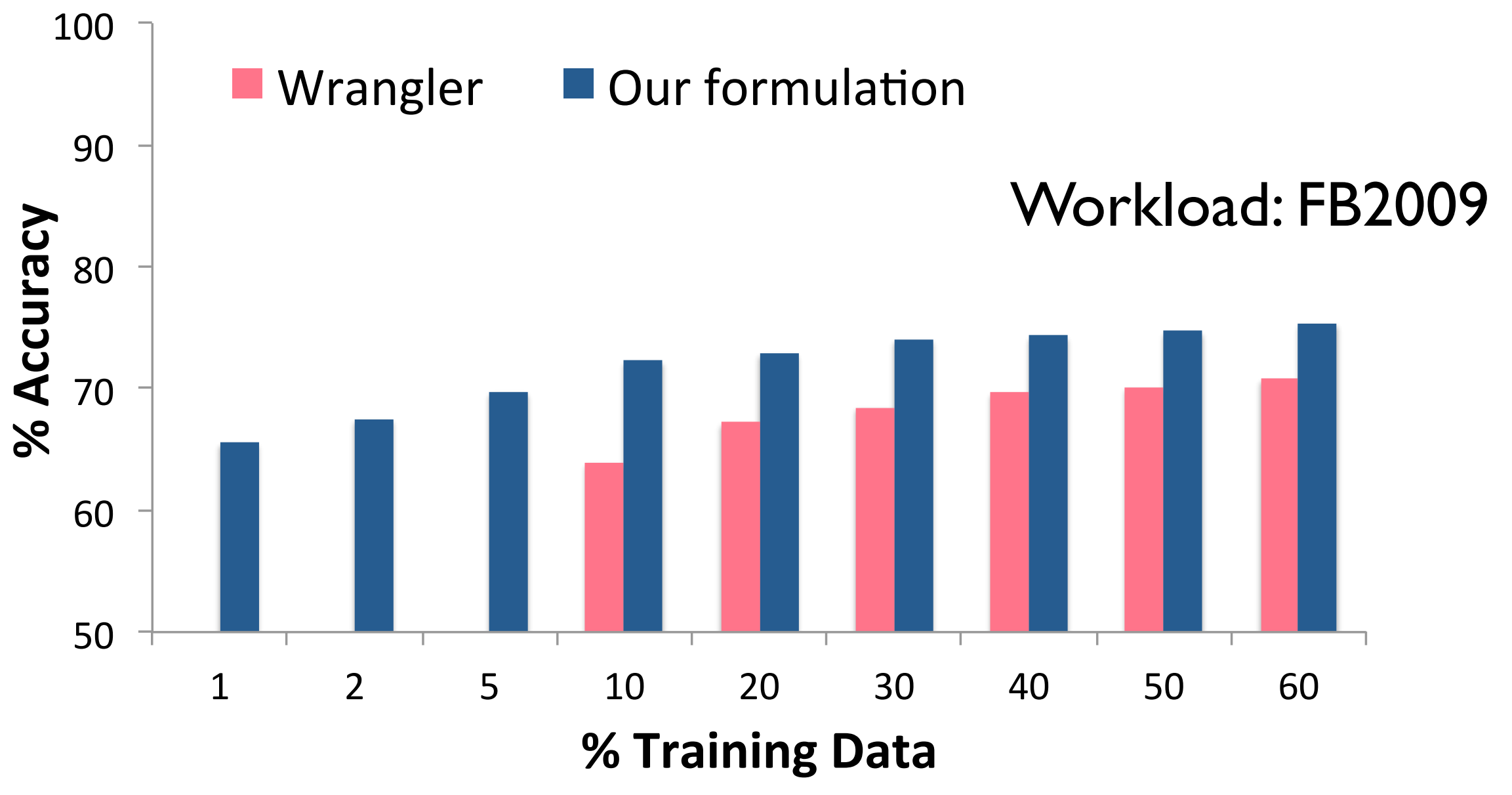
Training Problem

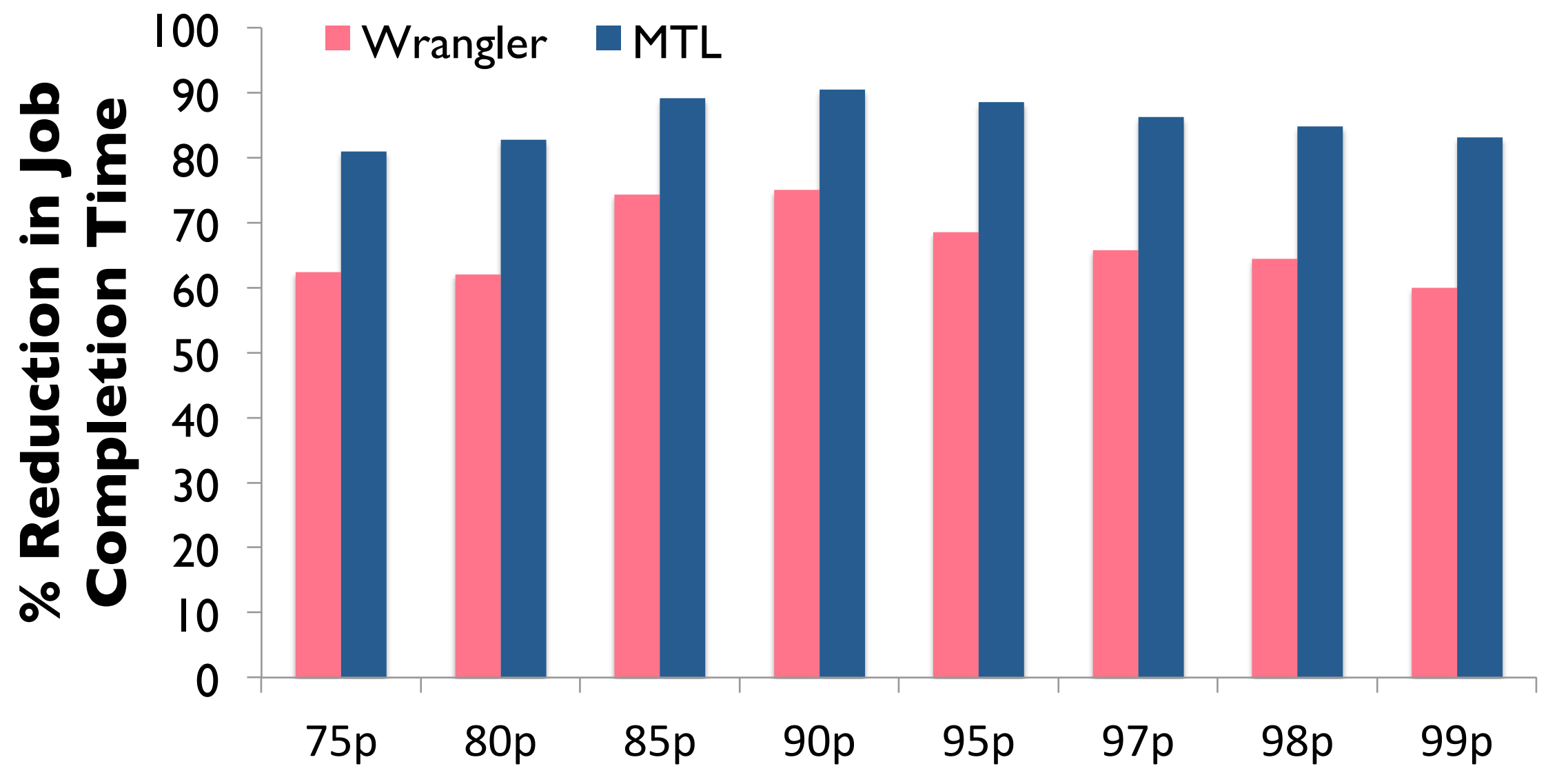
$$\min_{\mathbf{w}_0,\mathbf{v}_t,b} \lambda_0\|\mathbf{w}_0\|^2 + \frac{\lambda_1}{T}\sum_{t=1}^{T}\|\mathbf{v}_t\|^2 + \sum_{t=1}^{T}\sum_{i=1}^{m_t}\xi_{it}$$

$$\text{s.t} \quad y_{it}\left((\mathbf{w}_0+\mathbf{v}_t)^T\mathbf{x}_{it}+b\right) \geq 1-\xi_{it} \quad \forall i,t$$

$$\xi_{it} \geq 0 \quad \forall i,t$$

Evaluation





Intel Science & Technology Center for Cloud Computing