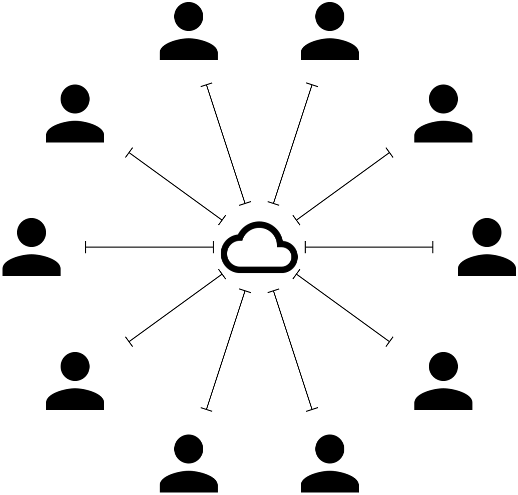
Online Load Balancing of Temporary Jobs

Alex Cann

University of Toronto

March 4, 2021

Motivating example



Model

Definition

- a set of m machines $\{M_1 \ldots M_m\}$
- Each event consists of a Job arriving or leaving
- A Job $\mathcal{J}$ consists of a weight w
- The online algorithm is unaware of both future arrivals and departures
- The load of a machine is the total weight of active jobs on the machine.

Summary of Variations

	Duration		
Machines	Unknown	Known	Permanent
Identical	$2 - \frac{1}{m}$ [5]	?	$2 - \epsilon$
Related	$\Theta(1)$ [7, 6]	$\Theta(1)$	$\Theta(1)$
Restricted	$\Theta(\sqrt{m})$ [6]	$O(\log mT), \Omega(\sqrt{m})$	$\Theta(logm)$
Unrelated	$O(m), \Omega(\frac{m}{\log m})$[1]	$O(\log mT), \Omega(\sqrt{m})$	$\Theta(logm)$

Table: Results for every combination of Machine Model and Job Duration. T is $\frac{\text{largest weight}}{\text{smallest weight}}$. m is the number of machines.

Competitive Ratio

Definition

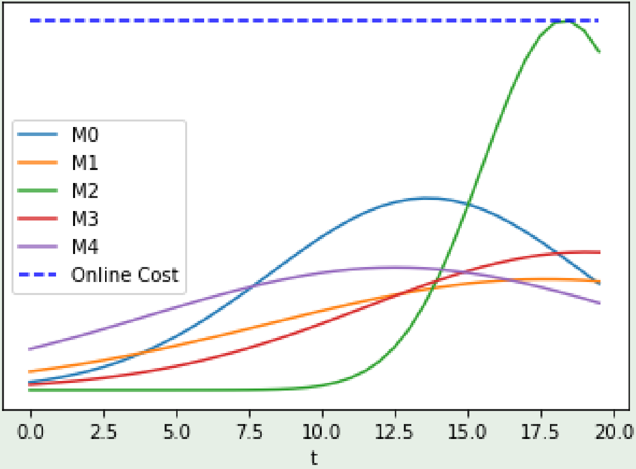
$$load(m, t) = \text{load on machine m at time t}$$

$$load_{OPT}(m, t) = \text{algorithm's load for an optimal offline algorithm}$$

$$r = \frac{\max_{t,m} load(m, t)}{\max_{t',m'} load_{OPT}(m', t')}$$

Competitive Ratio

Example



Momentary Competitive Ratio

Definition

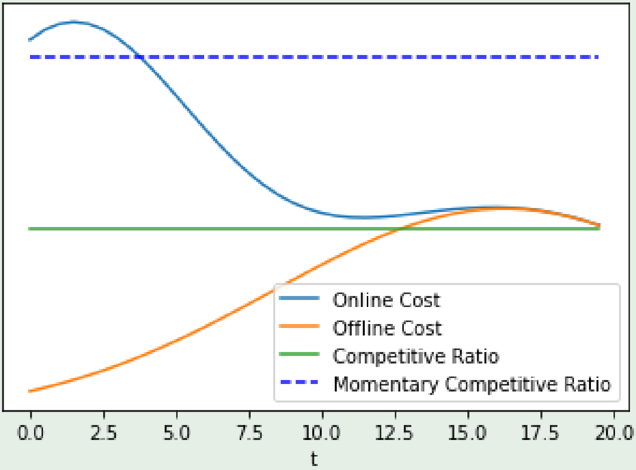
$$load(m, t) = \text{largest difference between online and offline}$$

$$load_{OPT} = \text{load for an optimal offline algorithm}$$

$$r = max_t \left(\frac{max_m load(m, t)}{max_{m'} load_{OPT}(m', t)} \right)$$

Momentary Competitive Ratio

Example



Theorem

There is a $2 - \frac{1}{m}$ competitive algorithm for temporary tasks on identical machines.

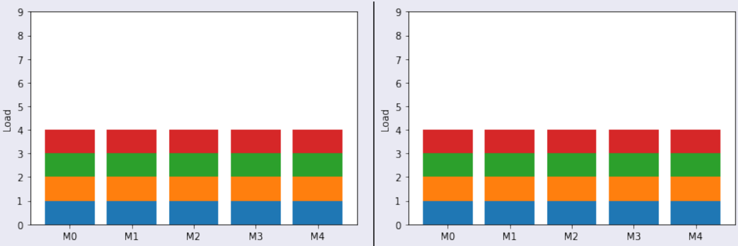
Theorem

There is a $2 - \frac{1}{m}$ competitive algorithm for temporary tasks on identical machines.

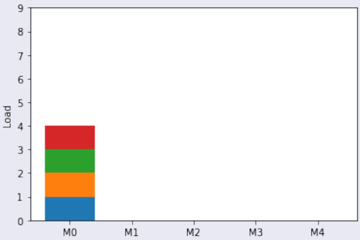
Lower Bound

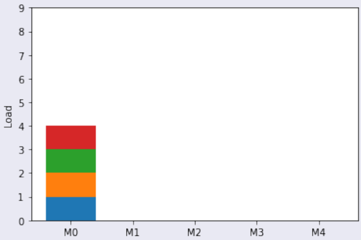
Any deterministic online algorithm for load balancing of temporary tasks has competitive ratio of at-least $2 - \frac{1}{m}$.

Case 1: m(m-1) unit weight jobs arrive

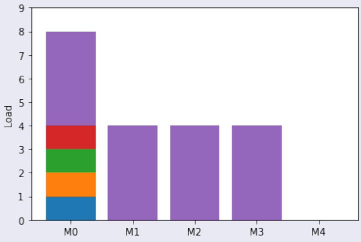


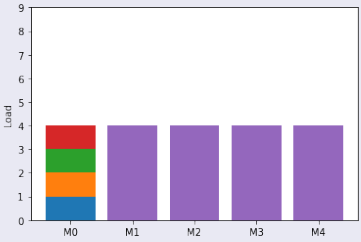
Case 1: All jobs except for m-1 on one machine depart



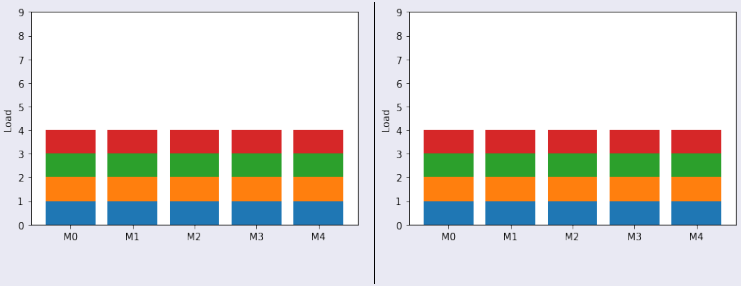


Case 1: m-1 jobs of weight m-1 arrive

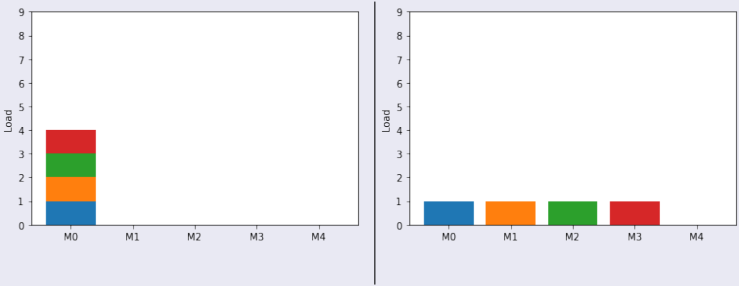




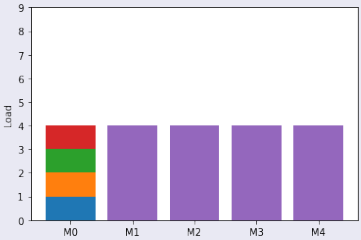
Case 2: m(m-1) unit weight jobs arrive

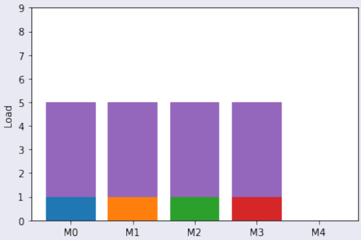


Case 2: All jobs except for m-1 on one machine depart

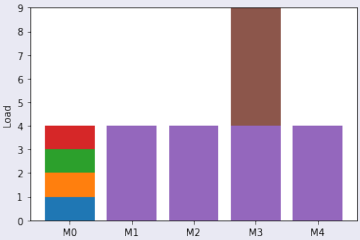


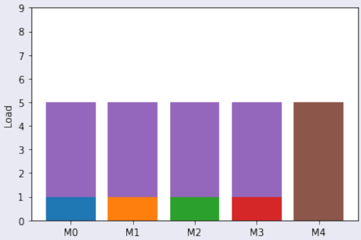
Case 2: m-1 jobs of weight m-1 arrive





Case 2: One jobs of weight m arrives





Theorem

Any lower bound for the deterministic model where the optimal value of the load is known in advance is also a lower bound on randomized algorithms of the same problem.

Lemma (Yao's theorem)

The lower bound for the competitive ratio of deterministic algorithms on any distribution on the input is also a lower bound for randomized algorithms.

Definition

$T = \{\sigma_1, \ldots \sigma_n\}$ be a set of inputs to a deterministic algorithm $\mathcal{D}$ with fixed optimal load.
Consider σ_{worst}, the input corresponding to the worst case cost C_{on}. Then $\sigma_{worst} \in T$.
let $\mathcal{S}$ be a random sequence of inputs separated by periods with no active jobs.Then $COST(\mathcal{S}) = \max_{\sigma \in \mathcal{S}} COST(\sigma)$.

$$P(COST(S) < C_{on}) = P(\sigma_{worst} \notin \mathcal{S}) = (1 - \frac{1}{\mid T \mid})^{len(\mathcal{S})}$$
$$= (1 - \frac{1}{\mid T \mid})^{|T|k}$$
$$\leq e^{-k}$$

Definition

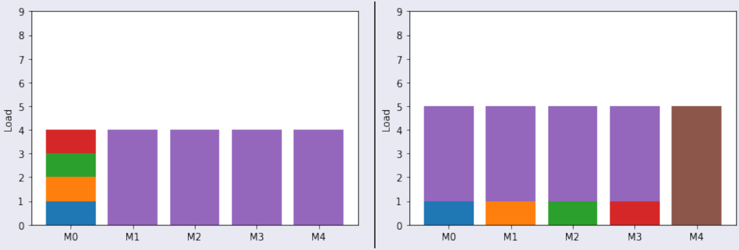
$T = \{\sigma_1, \ldots \sigma_n\}$ be a set of inputs to a deterministic algorithm $\mathcal{D}$ with fixed optimal load.
Consider σ_{worst}, the input corresponding to the worst case cost C_{on}. Then $\sigma_{worst} \in T$.
let $\mathcal{S}$ be a random sequence of inputs separated by periods with no active jobs.Then $COST(\mathcal{S}) = \max_{\sigma \in \mathcal{S}} COST(\sigma)$.

$$\begin{aligned}
\mathbb{E}[COST(\mathcal{S})] &\geq P(COST(S) < C_{on}) * C_{opt} \\
&+ P(COST(S) = C_{on}) * C_{on} \\
&= e^{-k} * C_{opt} + (1 - e^{-k}) * C_{on} \\
&= (0) + (1 - 0) * C_{on} \\
&= C_{on}
\end{aligned}$$

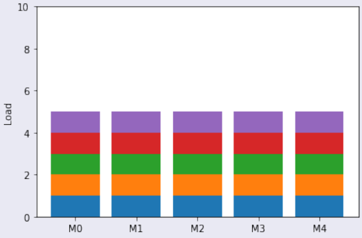
Theorem

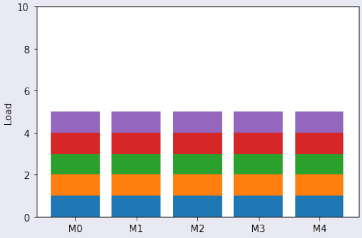
Any deterministic online algorithm for load balancing of temporary tasks has competitive ratio of at least $2 - \frac{2}{m+1}$ *even if the optimal load is known in advance.*

Case 1 and 2 Have different optimal loads

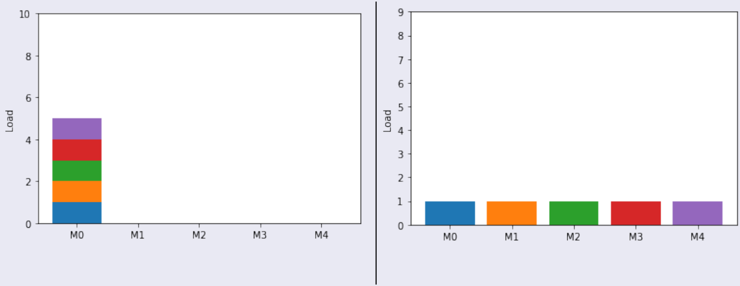


m^2 unit jobs arrive

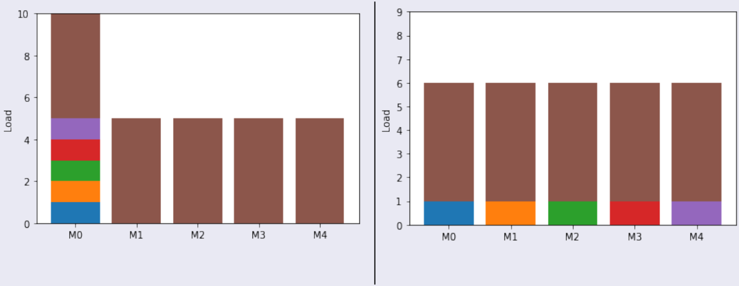




All jobs except for m on one machine depart



m jobs of weight *m* arrive



References I

Amitai Armon et al. "Temporary Tasks Assignment Resolved". In: *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '02. San Francisco, California: Society for Industrial and Applied Mathematics, 2002, pp. 116–124. ISBN: 089871513X.

Yossi Azar. "On-line load balancing". In: *Online algorithms*. Springer, 1998, pp. 178–195.

Yossi Azar, Andrei Z. Broder, and Anna R. Karlin. "On-line load balancing". In: *Theoretical Computer Science* 130.1 (1994), pp. 73–84. ISSN: 0304-3975. DOI: https://doi.org/10.1016/0304-3975(94)90153-8. URL: https://www.sciencedirect.com/science/article/pii/0304397594901538.

References II

Yossi Azar, Amir Epstein, and Leah Epstein. “Load balancing of temporary tasks in the Lp norm”. In: *Theoretical Computer Science* 361.2 (2006). Approximation and Online Algorithms, pp. 314–328. ISSN: 0304-3975. DOI: https://doi.org/10.1016/j.tcs.2006.05.016. URL: https://www.sciencedirect.com/science/article/pii/S0304397506003306.

Yossi Azar and Leah Epstein. “On-Line Load Balancing of Temporary Tasks on Identical Machines”. In: *SIAM Journal on Discrete Mathematics* 18 (Feb. 2005), pp. 347–352. DOI: 10.1137/S0895480197329296.

References III

Yossi Azar et al. "On-Line Load Balancing of Temporary Tasks". In: *J. Algorithms* 22.1 (Jan. 1997), pp. 93–110. ISSN: 0196-6774. DOI: 10.1006/jagm.1995.0799. URL: https://doi.org/10.1006/jagm.1995.0799.

References IV

Amotz Bar-Noy, Ari Freund, and Joseph (Seffi) Naor. "New algorithms for related machines with temporary jobs". In: *Journal of Scheduling* 3.5 (2000), pp. 259–272. DOI: https://doi.org/10.1002/1099-1425(200009/10)3:5<259::AID-JOS47>3.0.CO;2-3. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1099-1425\%28200009/10\%293\%3A5\%3C259\%3A\%3AAID-JOS47\%3E3.0.CO\%3B2-3. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/1099-1425%5C%28200009/10%5C%293%5C%3A5%5C%3C259%5C%3A%5C%3AAID-JOS47%5C%3E3.0.CO%5C%3B2-3.

References V

Tomáš Ebenlendr and Jiří Sgall. "Optimal and Online Preemptive Scheduling on Uniformly Related Machines". In: *J. of Scheduling* 12.5 (Oct. 2009), pp. 517–527. ISSN: 1094-6136. DOI: 10.1007/s10951-009-0119-7. URL: https://doi.org/10.1007/s10951-009-0119-7.