



HAL
open science

Performance of a Server Cluster with Parallel Processing and Randomized Load Balancing

Thomas Bonald, Céline Comte

► **To cite this version:**

Thomas Bonald, Céline Comte. Performance of a Server Cluster with Parallel Processing and Randomized Load Balancing. [Research Report] Telecom ParisTech. 2016. hal-01306343

HAL Id: hal-01306343

<https://hal.science/hal-01306343v1>

Submitted on 22 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Performance of a Server Cluster with Parallel Processing and Randomized Load Balancing

T. Bonald and C. Comte

April 22, 2016

Abstract

We consider a cluster of servers where each incoming job is assigned d servers chosen uniformly at random, for some fixed $d \geq 2$. Jobs are served in parallel and the resource allocation is balanced fairness. We provide a recursive formula for computing the exact mean service rate of each job. The complexity is polynomial in the number of servers.

1 Model

We consider a cluster of S servers, each with service rate μ . Jobs arrive according to a Poisson process with intensity λ . Each incoming job is assigned d servers chosen uniformly at random, for some fixed $d \geq 2$. Each job is processed in parallel by its assigned servers, the overall service capacity being allocated according to balanced fairness [1].

There are $N = \binom{S}{d}$ classes of jobs, defined by the assigned servers. Let $\mathcal{I} = \{1, \dots, N\}$ be the set of classes. We denote by $\mathcal{S}_i \subset \{1, \dots, S\}$ the set of servers assigned to each job of class i and, for each set $\mathcal{A} \subset \mathcal{I}$ of classes, we denote by $\mathcal{S}(\mathcal{A})$ the set of servers assigned to jobs whose class belongs to \mathcal{A} . Let $x = (x_i)_{i \in \mathcal{I}}$ be the network state, where x_i is the number of ongoing class- i jobs. We denote by $\phi_i(x)$ the total service rate of class- i jobs in state x . The corresponding vector lies in the capacity set:

$$\mathcal{C} = \left\{ \phi \in \mathbb{R}_+^N : \forall \mathcal{A} \subset \mathcal{I}, \quad \sum_{i \in \mathcal{A}} \phi_i \leq \mu |\mathcal{S}(\mathcal{A})| \right\},$$

where $|\mathcal{A}|$ is the cardinal of the set \mathcal{A} . The capacity set is a polymatroid, and it follows from [3] that balanced fairness is Pareto-efficient. Specifically,

$$\forall i \in \mathcal{I}, \quad \phi_i(x) = \begin{cases} \frac{\Phi(x - e_i)}{\Phi(x)} & \text{if } x_i > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where the function Φ is defined by the recursion $\Phi(0) = 1$ and, using the notation $\mathcal{A}_x = \{i \in \mathcal{I} : x_i > 0\}$,

$$\forall x \in \mathbb{N}^N \setminus \{0\}, \quad \Phi(x) = \frac{\sum_{i \in \mathcal{A}_x} \Phi(x - e_i)}{\mu |\mathcal{S}(\mathcal{A}_x)|}. \quad (1)$$

Under the stability condition $\lambda < S\mu$, the stationary distribution of the system state is given by

$$\forall x \in \mathbb{N}^N, \quad \pi(x) = \pi(0) \Phi(x) \left(\frac{\lambda}{N} \right)^{|x|}, \quad (2)$$

where $|x| = \sum_{i \in \mathcal{I}} x_i$ is the total number of jobs. This stationary distribution is insensitive to the job size distribution beyond the mean.

2 Performance

We are interested in the mean service rate, defined by

$$\gamma = \frac{\mathbb{E}(\sum_{i \in \mathcal{I}} \phi_i(X))}{\mathbb{E}(\sum_{i \in \mathcal{I}} X_i)},$$

where X is a random variable distributed according to the stationary distribution π . Observe that, by symmetry, γ is the mean service rate of any job (whatever its class), and that $\gamma \leq d\mu$. Moreover, by work conservation,

$$\gamma = \frac{\lambda}{\mathbb{E}(\sum_{i \in \mathcal{I}} X_i)}. \quad (3)$$

In particular, it follows from Little's law that $1/\gamma$ is the mean job duration.

There is no explicit formula for computing γ with a low complexity. In particular, the recursive formula of de Veciana and Shah [3] does not apply because the capacity set is *not* a symmetric polymatroid. We use the recent results of Gardner et. al. [2] to derive an explicit recursive formula, whose complexity is linear in SN (thus polynomial in S). We denote the system load by

$$\rho = \frac{\lambda}{S\mu}.$$

Proposition 1 *We have $\gamma = G/F$ with*

$$G = \sum_{n=0}^N \sum_{m=0}^S G_{n,m} \quad \text{and} \quad F = \sum_{n=0}^N \sum_{m=0}^S F_{n,m},$$

where $G_{n,m}$ and $F_{n,m}$ are given by the recursions $G_{0,0} = 1, F_{0,0} = 0$,

$$G_{n,m} = \frac{\frac{\rho S}{N}}{m - n \frac{\rho S}{N}} \left(\left[\binom{m}{d} - n + 1 \right] G_{n-1,m} + \sum_{r=1}^{\min(d,m)} \binom{S-m+r}{r} \binom{m-r}{d-r} G_{n-1,m-r} \right),$$

$$F_{n,m} = \frac{\frac{\rho S}{N}}{m - n \frac{\rho S}{N}} \left(\left[\binom{m}{d} - n + 1 \right] F_{n-1,m} + \sum_{r=1}^{\min(d,m)} \binom{S-m+r}{r} \binom{m-r}{d-r} F_{n-1,m-r} \right) + \frac{1}{\lambda} \frac{m}{m - n \frac{\rho S}{N}} G_{n,m}$$

if $d \leq m \leq S$ and $\lceil \frac{m}{d} \rceil \leq n \leq \binom{m}{d}$, $G_{n,m} = F_{n,m} = 0$ otherwise.

Proof. By symmetry, we have for any $i \in \mathcal{I}$,

$$\gamma = \frac{\lambda/N}{\mathbb{E}(X_i)}.$$

Let

$$G(\lambda_1, \dots, \lambda_N) = \sum_{x \in \mathbb{N}^N} \Phi(x) \prod_{i \in \mathcal{I}} \lambda_i^{x_i}$$

and

$$G = G\left(\frac{\lambda}{N}, \dots, \frac{\lambda}{N}\right).$$

In view of (2), we have $\gamma = G/F$, with

$$F = \frac{\partial G}{\partial \lambda_i} \left(\frac{\lambda}{N}, \dots, \frac{\lambda}{N} \right).$$

We first prove the recursion for computing G . For any $\mathcal{A} \subset \mathcal{I}$, let

$$G_{\mathcal{A}}(\lambda_1, \dots, \lambda_N) = \sum_{x \in \mathbb{N}^N: \mathcal{A}_x = \mathcal{A}} \Phi(x) \prod_{i \in \mathcal{I}} \lambda_i^{x_i}$$

and

$$G_{\mathcal{A}} = G_{\mathcal{A}}\left(\frac{\lambda}{N}, \dots, \frac{\lambda}{N}\right).$$

Observe that

$$G = \sum_{\mathcal{A} \subset \mathcal{I}} G_{\mathcal{A}}.$$

Now let $\mathcal{S}(\mathcal{A})$ be the set of servers that can serve jobs of classes in \mathcal{A} . Let $n = |\mathcal{A}|$ be the number of active classes and $m = |\mathcal{S}(\mathcal{A})|$ be the number of busy servers. In view of (1), we have

$$G_{\mathcal{A}} = \frac{\sum_{i \in \mathcal{A}} \frac{\lambda}{N} G_{\mathcal{A} \setminus \{i\}}}{m\mu - n \frac{\lambda}{N}}. \quad (4)$$

For all $n = 0, 1, \dots, N$ and $m = 0, 1, \dots, S$, let

$$G_{n,m} = \sum_{\substack{\mathcal{A} \subset \mathcal{I} \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} G_{\mathcal{A}}.$$

Observe that

$$G = \sum_{n=0}^N \sum_{m=0}^S G_{n,m}$$

Moreover, $G_{0,0} = 1$ and $G_{n,m} = 0$ unless $d \leq m \leq S$ and $\lceil \frac{m}{d} \rceil \leq n \leq \binom{m}{d}$. For such a pair n, m , we deduce from (4) that

$$G_{n,m} = \frac{\frac{\rho S}{N}}{m - n \frac{\rho S}{N}} \sum_{\substack{\mathcal{A} \subset \mathcal{I} \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} \sum_{i \in \mathcal{A}} G_{\mathcal{A} \setminus \{i\}}. \quad (5)$$

We can rewrite the sum as follows:

$$\begin{aligned} \sum_{\substack{\mathcal{A} \subset \mathcal{I} \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} \sum_{i \in \mathcal{A}} G_{\mathcal{A} \setminus \{i\}} &= \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A} \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} G_{\mathcal{A} \setminus \{i\}}, \\ &= \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{B} \subset \mathcal{I}, i \notin \mathcal{B} \\ |\mathcal{B}|=n-1, |\mathcal{S}(\mathcal{B} \cup \{i\})|=m}} G_{\mathcal{B}}, \\ &= \sum_{r=0}^{\min(d,m)} \sum_{\substack{\mathcal{B} \subset \mathcal{I}, \\ |\mathcal{B}|=n-1, |\mathcal{S}(\mathcal{B} \cup \{i\})|=m \\ |\mathcal{S}(\mathcal{B})|=m-r}} \sum_{i \in \mathcal{I} \setminus \mathcal{B}} G_{\mathcal{B}}, \\ &= \left[\binom{m}{d} - n + 1 \right] G_{n-1,m} + \sum_{r=1}^{\min(d,m)} \binom{S-m+r}{r} \binom{m-r}{d-r} G_{n-1,m-r}, \end{aligned}$$

which is the announced result.

We now show the recursion for F . For any $\mathcal{A} \subset \mathcal{I}$ and $i \in \mathcal{I}$, let

$$F_{\mathcal{A}} = \frac{\partial G_{\mathcal{A}}}{\partial \lambda_i} \left(\frac{\lambda}{N}, \dots, \frac{\lambda}{N} \right).$$

Observe that

$$F = \sum_{\mathcal{A} \subset \mathcal{I}} F_{\mathcal{A}}.$$

For all $n = 0, 1, \dots, N$ and $m = 0, 1, \dots, S$, we can define by symmetry,

$$F_{n,m} = \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A}, \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} F_{\mathcal{A}} = \frac{1}{N} \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A}, \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} F_{\mathcal{A}}$$

and we have

$$F = \sum_{n=0}^N \sum_{m=0}^S F_{n,m}.$$

Similarly, $F_{0,0} = 0$ and $F_{n,m} = 0$ unless $d \leq m \leq S$ and $\lceil \frac{m}{d} \rceil \leq n \leq \binom{m}{d}$.

For any $i \in \mathcal{I}$ and $\mathcal{A} \subset \mathcal{I}$ with $|\mathcal{A}| = n$ and $|\mathcal{S}(\mathcal{A})| = m$, we have

$$F_{\mathcal{A}} = \frac{1}{\mu \left(m - n \frac{S\rho}{N} \right)} \left\{ G_{\mathcal{A}} + G_{\mathcal{A} \setminus \{i\}} + \frac{\lambda}{N} \sum_{j \in \mathcal{A}, j \neq i} F_{\mathcal{A} \setminus \{j\}} \right\}$$

so that

$$\begin{aligned} F_{n,m} = \frac{1}{N} \frac{1}{\mu \left(m - n \frac{\rho S}{N} \right)} & \left\{ \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A}, \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} G_{\mathcal{A}} + \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A}, \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} G_{\mathcal{A} \setminus \{i\}} \right. \\ & \left. + \frac{\lambda}{N} \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A}, \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} \sum_{j \in \mathcal{A}, j \neq i} F_{\mathcal{A} \setminus \{j\}} \right\} \end{aligned} \quad (6)$$

The first term of the sum is

$$\sum_{\substack{\mathcal{A} \subset \mathcal{I} \\ |\mathcal{A}|=n \\ |\mathcal{S}(\mathcal{A})|=m}} \sum_{i \in \mathcal{A}} G_{\mathcal{A}} = n \sum_{\substack{\mathcal{A} \subset \mathcal{I} \\ |\mathcal{A}|=n \\ |\mathcal{S}(\mathcal{A})|=m}} G_{\mathcal{A}} = n G_{n,m}.$$

By equation (5), the second term of the sum is

$$\sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A}, \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} G_{\mathcal{A} \setminus \{i\}} = \sum_{\substack{\mathcal{A} \subset \mathcal{I} \\ |\mathcal{A}|=n \\ |\mathcal{S}(\mathcal{A})|=m}} \sum_{i \in \mathcal{A}} G_{\mathcal{A} \setminus \{i\}} = \frac{m - n \frac{S\rho}{N}}{\frac{\rho S}{N}} G_{n,m}.$$

Thus the first two terms of equation (6) are equal to

$$n G_{n,m} + \frac{m - n \frac{\rho S}{N}}{\frac{\rho S}{N}} G_{n,m} = \frac{mN}{\rho S} G_{n,m}.$$

Now the third term of (6) is

$$\begin{aligned}
& \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i \in \mathcal{A} \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} \sum_{j \in \mathcal{A}, j \neq i} F_{\mathcal{A} \setminus \{j\}}, \\
&= \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}, j \neq i} \sum_{\substack{\mathcal{A} \subset \mathcal{I}, i, j \in \mathcal{A} \\ |\mathcal{A}|=n, |\mathcal{S}(\mathcal{A})|=m}} F_{\mathcal{A} \setminus \{j\}}, \\
&= \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}, j \neq i} \sum_{\substack{\mathcal{B} \subset \mathcal{I}, i \in \mathcal{B}, j \notin \mathcal{B} \\ |\mathcal{B}|=n-1, |\mathcal{S}(\mathcal{B} \cup \{j\})|=m}} F_{\mathcal{B}}, \\
&= \sum_{r=0}^{\min(d,m)} \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{B} \subset \mathcal{I}, i \in \mathcal{B} \\ |\mathcal{B}|=n-1, |\mathcal{S}(\mathcal{B})|=m-r}} \sum_{j \in \mathcal{I} \setminus \mathcal{B}, |\mathcal{S}(\mathcal{B} \cup \{j\})|=m} F_{\mathcal{B}}, \\
&= \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{B} \subset \mathcal{I}, i \in \mathcal{B} \\ |\mathcal{B}|=n-1, |\mathcal{S}(\mathcal{B})|=m}} \left[\binom{m}{d} - n + 1 \right] F_{\mathcal{B}}, \\
&+ \sum_{r=1}^{\min(d,m)} \sum_{i \in \mathcal{I}} \sum_{\substack{\mathcal{B} \subset \mathcal{I}, i \in \mathcal{B} \\ |\mathcal{B}|=n-1, |\mathcal{S}(\mathcal{B})|=m-r}} \binom{m-r}{d-r} \binom{S-m+r}{r} F_{\mathcal{B}}, \\
&= N \left[\binom{m}{d} - n + 1 \right] F_{n-1,m} + N \sum_{r=1}^{\min(d,m)} \binom{m-r}{d-r} \binom{S-m+r}{r} F_{n-1,m-r}.
\end{aligned}$$

Summing the two parts of equation (6) gives the announced result. \square

References

- [1] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo. A queueing analysis of max-min fairness, proportional fairness and balanced fairness. *Queueing Syst.*, 53(1-2):65–84, 2006.
- [2] K. Gardner, S. Zbarsky, M. Harchol-Balter, and A. Scheller-Wolf. Analyzing response time in the redundancy-d system. Technical report, Technical Report CMU-CS-15-141, 2015.
- [3] V. Shah and G. de Veciana. Performance evaluation and asymptotics for content delivery networks. In *Proceedings of IEEE Infocom*. IEEE, 2014.