Stability criterion of a MAP/PH-multiserver model with simultaneous service

Alexander Rumyantsev and Evsey Morozov

Institute of Applied Mathematical Research,
Karelian Research Centre, Russian Academy of Sciences

28.06.2016

The Inspirator

High performance computing cluster

homogeneous machine with multiple CPUs utilizing parallel computing

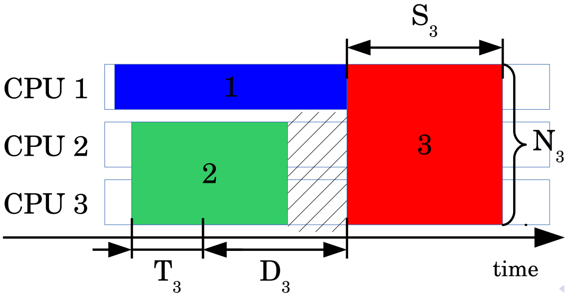
- shared by many users
- job/task uses multiple CPUs
- high utilization level
- very expensive (build, run)

- Memory/storage allocation in a PC
- Wireless channel occupation

The $MAP/M/c$-type Model

Queueing system:

- c identical servers, FCFS;
- interarrival times $T_i, i \geqslant 1$ defined by MAP (D_0, D_1) with k states (intensity $\lambda = \theta D_1 \mathbf{1}$);
- exponential service times S_i (intensity μ);
- customer i requires N_i servers at once, aka *rigid job* (distribution $\{p_j = \mathsf{P}(N = j), 1 \leqslant j \leqslant c\}$).



Idle servers with non-empty queue!

The History

1. *Kim S.S. M/M/s Queueing System Where Customers Demand Multiple Server Use: Ph. D. Dissertation, 1979.*
   Reference to the Neuts ergodicity condition.
2. *Brill P., Green L. Queues in which customers receive simultaneous service from a random number of servers: A system point approach. Management Science, 1984. Vol. 30. No. 1. P. 51–68.*
   Stability criterion for M/M/2-type system (without proof).
3. *D. Filippopoulos, H. Karatza. An M/M/2 parallel system model with pure space sharing among rigid jobs. Mathematical and Computer Modelling, 2007. Vol. 45, No. 5–6. P. 491–530.*
   Stability criterion for M/M/2-type system.
4. *S.R. Chakravarthy, H.D. Karatza. Two-server parallel system with pure space sharing and Markovian arrivals. Computers and Operations Research, 2013. Vol. 40, No. 1. P. 510–519.*
   Stability criterion for MAP/M/2-type system.

Main Result

Continuous-time QBD process $\{\Theta(t) = \{\nu(t), m(t), \varphi(t)\}, t \geqslant 0\}$,

level ν — number of customers in the system,

macrostate $m = (m_1, \ldots, m_n) \in \{1, \ldots, c\}^c =: \mathcal{M}$, m_i is a number of servers required by i-th oldest customer in the system,

MAP-phase $\varphi \in \{1, \ldots, k\}$.

Stability criterion

QBD process is positive recurrent iff

$$\rho = \frac{\lambda}{\mu} C < 1, \quad C := \sum_{m \in \mathcal{M}} \frac{\prod_{j=1}^{c} p_{m_j}}{\sigma(m)}, \tag{1}$$

where $\sigma(m) = \max_{i \leqslant c} \left\{ \sum_{j=1}^{i} m_j \leqslant c \right\}$.

Discussion

Case $c = 2$: a known earlier result

$$\frac{\lambda}{\mu} < \frac{2}{2 - p_1^2}.$$

Case $k = 1, p_1 = 1$ (classical $M/M/c$ system):

$$\frac{\lambda}{\mu} < c.$$

Case $k = 1$ (exponential arrivals):
[*A. Rumyantsev and E. Morozov. Stability criterion of a multiserver model with simultaneous service. Annals of Operations Research, 2015*]

Proof Sketch

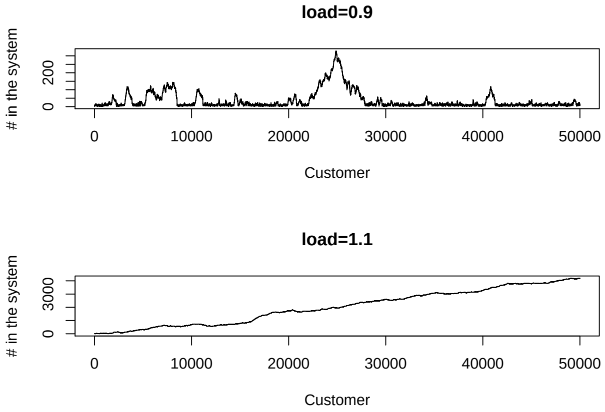
- Define infinitesimal generator for $\{\Theta(t)\}$ with finite number of phases kc^c at high levels $\nu(t) > c$, use Kroenecker sums/products and the properties of $M/M/c$-type model
- Apply the Neuts ergodicity condition $\gamma A_2 \mathbf{1} > \gamma A_0 \mathbf{1}$, where $\gamma = \alpha \otimes \theta$
- The vector α is defined componentwise

$$\alpha_m = C^{-1} \frac{\prod_{i=1}^{c} p_{m_i}}{\sigma(m)}, \quad m \in \mathcal{M},$$

  and comes from the $M/M/c$-type model
- Recall θ comes from MAP as a solution $\theta D = 0,\ \theta \mathbf{1} = 1$, where $D = D_0 + D_1$.

On $MAP/PH/c$ Model

Numerical experiments show the validity of the result for PH-type service (and even for Pareto service!). For a PH (τ, T) define $\mu = (-\tau T^{-1}\mathbf{1})^{-1}$.



Unfortunately, the proof is still in progress.

Necessary Stability Condition

It can be shown, that the system with *batches of customers* of size N_i, each having service time S_i, is minorant to HPC model. The necessary condition follows

$$\frac{\lambda}{\mu}\mathsf{E}N < c. \qquad (2)$$

One may use (2) to easily check the *instability* of the system model.

Accelerated Verification

An equivalent representation

$$C = \sum_{i=1}^{c} \frac{1}{i} \sum_{j=i}^{c} {p_j^*}^i \sum_{t=c-j+1}^{c} p_t, \qquad (3)$$

the summation is done over

- number i of customers at service,
- number j of servers serving customers,
- number t of servers required by the customer at the head of the queue.

For $c = 5000$ on my laptop: 20 sec.

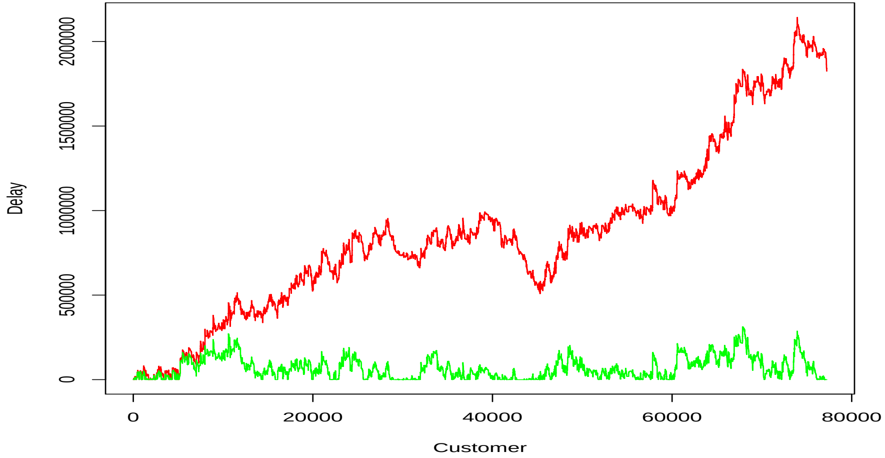
Application Example

Upgrade an unstable cluster: given $\lambda, \mu, p_1, \ldots, p_c$, find $c' > c$ s.t. $\lambda C/\mu < 1$.

Example: Cornell Theory Center (CTC) IBM SP2 cluster

- $s = 336$ processors, 77221 tasks from Workload Archive
- EASY Backfill scheduler: stable, but high delays (mean 25540 sec, max 7231000 sec)
- $\rho \approx 1.14$ (unstable *under FIFO*)

Upgrade CTC SP2 cluster



Original *(under FIFO)* $c = 336$ and Upgraded: $c' = 372$

Thank you for attention!

Rumyantsev Alexander
Institute of Applied Mathematical Research
Karelian Research Centre RAS
ar0@krc.karelia.ru

ResearcherID: L-1354-2013
ORCID: orcid.org/0000-0003-2364-5939
ScopusID: 36968331100

ResearchGate: https://www.researchgate.net/profile/Alexander_Rumyantsev