Bounding aggregations on phase-type arrivals for performance analysis of clouds

Farah Aït-Salaht
LIP6, Université Paris Ouest Nanterre
4 place Jussieu, 75252 Paris cedex 05, France
Email: farah.aitsalaht@u-paris10.fr

Hind Castel-Taleb
SAMOVAR, CNRS, Télécom Sud Paris
Université Paris-Saclay
9 rue Charles Fourier 91011 Evry, France
Email: hind.castel@it-sudparis.eu

***Abstract*—We evaluate the performance of a cloud system using a hysteresis queueing system with phase-type and batch arrivals. To represent the dynamic allocation of the resources, the hysteresis queue activates and deactivates the virtual machines according to the threshold values of the queue length. We suppose a variable traffic intensity as the client requests (or jobs) arrive by batches, and follow a phase-type process. This system is represented by a complex Markov chain which is difficult to analyze, especially when the size of the state space increases and the length of batch arrival distribution is large. So, to solve this problem, we propose to use stochastic bounds and define bounding systems less complex. We give some results for the performance measures and compare the proposed bounding models with the exact one. The relevance of our methodology is to offer a trade-off between computational complexity and accuracy of the results and provide very interesting solutions for network dimensioning.**

I. Introduction

Over the past years, the cloud computing technology had an impressive increase in popularity in both the software industry and research worlds. Indeed, these systems allow to make a better use of distributed resources, combine them to achieve higher throughput and be able to solve large scale computation. It provides ability for cloud consumers to use or implement flexible and scalable services without having the computing resources installed directly on consumer's system. With this flexibility, the cloud providers can rent the Virtual Machines (VMs) depending on the demand and can gain more profit out of a single Physical Machine (PM).

Performance evaluation of cloud centers is a very crucial research task which becomes difficult due to the dynamic behavior of cloud environments and variability of user demand. In [1] for example, the authors develop an analytical model to evaluate the performance of cloud centers with a high degree of virtualization and Poisson batch arrives. The model of the physical machine with m VMs is based on the $M^{[x]}/G/m/m+r$ queue. To solve the model, they adopt the technique of embedded Markov chain, and derive exact formulas for different performance measures. In [2], the authors analyze the performance of a large scale Iaas Clouds using a multi-level interacting stochastic sub-models approach, where the overall model solution is obtained iteratively over individual sub-models solutions.

In this paper, we propose to represent the system by a queueing model based on queue-dependent virtual machines and analyze quantitatively the dynamic behavior of the data center. The data center is represented by a set of PMs hosting a set of VMs which are instanced according to user demand. The considered queueing model is a multi-server with hysteresis threshold queues [3]. And, in order to represent the variability of customer requests, we suppose that arrivals are bulks and follow a phase-type process.

We model the queueing system by a multidimensional Markov chain. This model is difficult to solve as the size of state space increases with the number of phases and the number of VMs, and it becomes complex when the size of bulk arrivals is high. For the resolution approach of this model, we can use the Quasi Birth and Death (QBD) process or Quasi Simultaneous-Multiple Births and Deaths (QBD-M) process [4] which are used in the well-known web server model. These modeling yields a matrix-geometric steady state solution. However, we remark that for our problem it is not an easy task to compute matrix Z, this observation was also mentioned in [5] when they analyze a threshold-based queueing system with hysteresis and Poisson input process. So in this paper, we propose to have a different analysis approach. In fact, in order to reduce the complexity of this problem and have guarantees on performance mesures, we propose to use stochastic comparisons and derive bounding systems. Different bounding systems are defined: by simplifying the behavior of the exact Markov chain, or by aggregating the batch probability distributions. We prove using stochastic comparisons that these processes provide bounds for performance measures as blocking probabilities and expected buffer length. So, through these models we propose to offer a trade-off between computational complexity and accuracy of results and solve efficiently the network dimensioning problem and satisfy the QoS (Quality of Service) constraints requirements.

The paper is organized as follows: next, we describe the cloud system, and present the queueing model considered for our analysis. In section III, we give some theoretical notions about the stochastic ordering theory and in section IV, we give the bounding models and we prove using the stochastic comparisons that they are bounds of the original system. In section V, we give some numerical results on performance measures. Finally, we discuss the achieved results in the conclusion and we comment about further research.

II. Cloud system and model description

We focus our study on one data center with a set of resources or physical machines which can host a lot of virtual machines, as shown in Fig 1, and we propose to represent it by a stochastic model based on a queueing system.

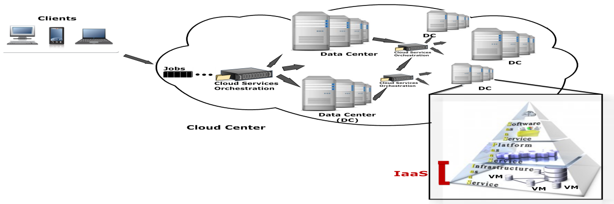

Fig. 1: Cloud center architecture

In order to have a system able to handle the variability of the traffic intensities, the VMs are activated and deactivated according to the system occupancy. In fact, the buffer management is governed by thresholds of the number of customers waiting in the queue, which activate or deactivate the VMs. Clearly, when the number of customers in the queue reaches a forward threshold, a new VM is activated, and when it decreases below the reverse threshold, a VM is deactivated. We present here after in details the queueing model used for our analysis.

We consider a finite buffer capacity with multi-homogeneous servers (VMs). We denote by C the capacity of the queue. We suppose a K multi-server thresholds-based queueing system with hysteresis for which a set of forward thresholds $(F_1, F_2, \dots, F_{K-1})$ and a set of reverse thresholds $(R_1, R_2, \dots, R_{K-1})$ are defined. We assume that $F_i < F_{i+1}$, $R_i < R_{i+1}$, and $R_i < F_i, \forall 1 \le i < K$. The behavior of this system is as follows. We assume that the first VM is always active. If a customer arrives in the system, and finds $F_{i\,(i=1,\dots,K-1)}$ customers in the queue, an additional VM will be activated. When a customer leaves the system with $R_{i\,(i=1,\dots,K-1)}$ customers, then one VM will be removed from the set of active VMs. We assume here that the arrival process is a phase-type process with batch arrivals. This process is defined by Poisson arrivals modulated by phases, with batch size distribution. Indeed, the phase-type process has enough flexibility to describe a wide variety of data flows, and its physical interpretation seems to describe arrival rate fluctuations in many situations. The arrivals are Markovian modulated by a phase process with ℓ states. We denote by $L = \{1, 2, \dots, \ell\}$ the set of phase values. The set of phases corresponds to the different traffic intensity (variability or burstiness). Let us denote by $\mathbf{M}$ the probability transition matrix of this process, where $\mathbf{M}(i, j)$ is the probability transition from phase i to phase j. Let $\{X(t), t \ge 0\}$ be the stochastic process which model the behavior of the hysteresis system. Each state is represented by a 3-tuple (x_1, x_2, ϕ), where x_1 is the number of customers waiting in the queue ($x_1 \in \{0, \dots, C\}$), x_2 is the number of active VMs ($x_2 \in \{1, \dots, K\}$), and ϕ expresses the arrival phase ($\phi \in L = \{1, 2, \dots, \ell\}$). For the traffic arrivals, we suppose that at each phase ϕ, the bulk requests arrive according to a Poisson process with rate $\lambda^{(\phi)}$, and bulk size follows a probability distribution $p_\phi = (p_\phi(1), \dots, p_\phi(k), \dots, p_\phi(n))$, defined as follows, where $E \subset \mathbb{N}$, and $|E| = n$:

$$p_\phi(k) = \Pr[\text{In phase } \phi, \text{the bulk size is } k], \quad \forall k \in E, \forall \phi \in L,$$

Servers (or VMs) have an exponential service time distribution with mean rate $\mu_i = \mu$ $(i = 1, \dots, K)$.

With these assumptions, we deduce that the system $X(t)$ is a Continuous-Time Markov Chain (CTMC) defined over the state space $A = \{(x_1, x_2, \phi) | (x_1, x_2) \in B \,\text{and}\, \phi \in L\}$, where:

$$\begin{aligned} B = \{(x_1,\, x_2) \quad | \quad & 0 \le x_1 \le F_1, \text{ if } x_2 = 1; \\ & R_{i-1} < x_1 \le F_i, \text{ if } x_2 = i \text{ and } 1 < i < K; \\ & R_{K-1} < x_1 \le C, \text{ if } x_2 = K\}. \end{aligned}$$

So, $A = B \times L$.

We suppose that the arrivals take place first in the system, and then we have a phase transition. The evolution equations of $X(t)$ are defined for $i, j = 1, \dots, K-1$, as follows (where $\phi, \phi' \in L$):

$$\begin{aligned} (x_1, x_2, \phi) \rightarrow \quad & (\min\{C, x_1{+}k\}, x_2, \phi'), \text{with rate } \lambda_\phi p_\phi(k)\, \mathbf{M}(\phi, \phi'), \\ & \text{if } (x_1 + k) \le F_j, \text{ and } x_2 = j, \ j = 1, \dots, K-1, \\ \rightarrow \quad & (\min\{C, x_1{+}k\}, K, \phi'), \text{with rate } \lambda_\phi p_\phi(k)\, \mathbf{M}(\phi, \phi'), \\ & \text{if } x_2 = K \text{ or } (x_1 + k) > F_{K-1}, \\ \rightarrow \quad & (\min\{C, x_1{+}k\}, l, \phi'), \text{with rate } \lambda_\phi p_\phi(k)\, \mathbf{M}(\phi, \phi'), \end{aligned}$$

$$\begin{aligned} & \text{if } l = \min\{h | (x_1{+}k \le F_h) \text{ and } x_2{+}1 \le h \le K{-}1\}, \\ \rightarrow \quad & (\max\{0, x_1\text{-}1\}, x_2, \phi), \text{with rate } x_2\mu, \\ & \text{if } (x_1 \ne R_i + 1) \text{ or } (x_1 = R_i + 1 \text{ and } x_2 \ne i + 1) \\ \rightarrow \quad & (\max\{0, x_1\text{-}1\}, \max\{0, x_2\text{-}1\}, \phi), \text{with rate } x_2\mu, \\ & \text{if } x_1 = R_i + 1, \text{ and } x_2 = i + 1. \end{aligned}$$

We note that the threshold-based queue with hysteresis has been already studied in the literature when arrivals follow Poisson and batch Poisson arrival processes [3], [5] and [6]. In [3], the authors propose to use the Green's function method which is not so easy to apply as the formalism is not intuitive. And in [5], the authors use the concept of stochastic complementation to solve the system. They propose to partition the state space in disjoints sets in order to aggregate the Markov chain. However, we notice that in order to be able to capture correlations and burstiness, since real traffic often exhibits these characteristics, the consideration of the Poisson arrival process by batches is not sufficient. Regarding to the context of our study, it was natural to describe the arrival process of data center by phase-type and batch arrivals. Indeed, using modulated arrivals in the threshold-based queue with hysteresis seems more appropriate and allows to capture better the properties of the network traffic such as self-similarity and long-range dependence. And we specify that these characteristics have significant impact on network performance. Moreover, beyond investigating this aspect, we propose in this paper to define bounding models instead of performing an exact resolution of the system which can be often very difficult. Before introducing our bounding models, we give first the methodology used to solve the studied threshold hysteresis model with batch phase-type arrivals.

A. Numerical Analysis

For the numerical analysis of the underlying CTMC $\{X(t),\, t \geq 0\}$, we propose to uniformize the chain and derive a Discret Time Markov Chain (DTMC). Let $\mathbf{Q}$ be the infinitesimal generator of $X(t)$, we consider here the uniformization of $\mathbf{Q}$ denoted by matrix $\mathbf{P}$ such that:

$$\forall\, q \geq \max_{x\in A} |\mathbf{Q}(x,\, x)|,\;\; \mathbf{P} = \mathbf{I} + \frac{1}{q}\mathbf{Q}. \tag{1}$$

Considering a phase-type arrival process with batches, we note that the Markov chain of the hysteresis model $\{X(t),\, t \geq 0\}$ is defined as a block structured matrix. Indeed, the structured matrix of the underlying DTMC can be expressed as functional Kronecker products in the theory of stochastic automata networks [7], [8].

To observe this, we can consider two states $(x_1,\, x_2,\, x_3)$ and $(y_1,\, y_2,\, y_3)$ of the DTMC, and the probability transition matrix $\mathbf{F}_\phi$ of the system when the arrivals are in phase ϕ. We denote that the probability transition matrix of the studied Markov chain is defined as follows:

$$\mathbf{P}((x_1,\, x_2,\, x_3), (y_1,\, y_2,\, y_3)) = \mathbf{M}(x_3,\, y_3)\mathbf{F}_{x_3}((x_1,\, x_2),\, (y_1,\, y_2)).$$

Considering the ordinary lumpability and the rich structure associated to the Kronecker representation, the authors in [8] have developed an iterative steady-state solution method which is able to compute performance measures for DTMCs. So, based on this, we state the following property:

Property 1: The lumpability of the Markov chain for the hysteresis model with batch phase-type arrivals can be investigated among the partition defined by the phase of the arrival process.

We define by φ the steady-state probability distribution of the probability transition matrix $\mathbf{M}$. And let $\pi^{(a)} = (\pi_1,\, \pi_2, \ldots,\, \pi_\ell)$ denotes the transient probability distribution for the probability transition matrix $\mathbf{P}$ at time a, with ℓ the number of phases. Due to the lumpability of $\mathbf{P}$, there exist ℓ vectors π_i of size $|B|$ such that π has the following block decomposition:

$$\pi^{(a)} = (\varphi(1)\pi_1^{(a)} \,|\, \varphi(2)\pi_2^{(a)} \,|\, \varphi(3)\pi_3^{(a)} \,|\, \ldots \,|\, \varphi(\ell)\pi_\ell^{(a)}).$$

To compute the steady-state solution of the model, we use the Iterative Aggregation Disaggregation (IAD) algorithm specialized for lumpable matrices published in [8]. Through this algorithm, we obtain successive values of vectors $\pi_i^{(a)}$. We denote by $\pi_i^{(a)}$ the vectors computed at iteration a.

From the steady state distribution π of the CTMC $\{X(t),\, t \geq 0\}$, we can derive some performance measures as the blocking probability and expected buffer length.

a) Blocking probability. This metric is computed as follows: $Bp = \sum_{x_2,x_1,k} \sum_\phi \pi(x_1,\, x_2,\, \phi)\, p_\phi(k)\, \mathbb{1}_{\{x_1+k\geq C\}}$.

b) Expected buffer length. The expected buffer length is expressed as follows: $\mathbb{E}[\pi] = \sum_{x_1} \left(\sum_{x_2} \sum_\phi \pi(x_1,\, x_2, \phi)\right) x_1$.

Before introducing the proposed bounding models, we describe briefly the stochastic ordering theory.

III. Stochastic Ordering Theory

We refer to Stoyan's book [9] for theoretical issues of the stochastic comparison method. We consider the state space G endowed with a partial order denoted as $\preceq$. Let X and Y be two discrete random variables taking values on G, with cumulative probability distributions F_X and F_Y, and probability distribution vectors p and q, where $p(x) = Prob(X = x)$, and $q(x) = Prob(Y = x)$, for $x \in G$. Next, we give the definition of the strong stochastic ordering $\preceq_{st}$:

Definition 1: We can define the $\preceq_{st}$ ordering as follows:

- **generic definition:** $X \preceq_{st} Y \iff \mathbb{E}f(X) \leq \mathbb{E}f(Y)$, for all non decreasing functions $f : G \rightarrow \mathbb{R}^+$ whenever expectations exist.

Notice that we use interchangeably $X \preceq_{st} Y$, and $p \preceq_{st} q$.

We give also the definition for stochastic comparisons of stochastic processes. Let $\{X(t), t \geq 0\}$ and $\{Y(t), t \geq 0\}$ be stochastic processes defined on G.

Definition 2: We say that $\{X(t), t \geq 0\} \preceq_{st} \{Y(t), t \geq 0\}$, if $X(t) \preceq_{st} Y(t), \forall t \geq 0$.

When the processes are defined on different states spaces we can compare them on a common state space using mapping functions. Let $\{X(t), t \geq 0\}$ (resp. $\{Y(t), t \geq 0\}$) be defined on A (resp. B), g (resp. h) be a many to one mapping from A to S, (resp. $B \rightarrow S$). Next, we compare the mapping of the process $\{X(t), t \geq 0\}$ (resp. $\{Y(t), t \geq 0\}$) by the mapping function g (resp. h), which means $g(X(t))$ (resp. $h(Y(t))$), on the common space S, endowed with the partial order $\preceq$.

The stochastic comparisons of processes by mapping functions is defined as follows [10]:

Definition 3: We say that $\{g(X(t)), t \geq 0\} \preceq_{st} \{h(Y(t)), t \geq 0\}$, if $g(X(t)) \preceq_{st} h(Y(t)), \forall t \geq 0$

We can use the coupling method for the stochastic comparison of the processes. As presented in [10], it remains us to define two CTMCs: $\{\widehat{X}(t),\; t \geq 0\}$ and $\{\widehat{Y}(t),\; t \geq 0\}$ governed by the same infinitesimal generator matrix respectively as $\{X(t), t \geq 0\}$, and $\{Y(t), t \geq 0\}$, representing different realizations of these processes with different initial conditions. The following theorem establishes the $\preceq_{st}$-comparison using the coupling [10]:

Theorem 1: $\{g(X(t)), t \geq 0\} \preceq_{st} \{h(Y(t)), t \geq 0\}$, if there exists the coupling $\{(\widehat{X}(t), \widehat{Y}(t)),\; t \geq 0\}$ such that:

$$g(\widehat{X}(0)) \preceq h(\widehat{Y}(0)) \Rightarrow g(\widehat{X}(t)) \preceq h(\widehat{Y}(t)),\; \forall t > 0.$$

IV. Bounding Systems

We propose to define different bounding systems which are easier to solve. Different ways to simplify the exact system are used. The first bounding systems are defined by reducing the size of the bulk arrivals and aggregating the probability distribution of bulk arrivals. The second bounding models are obtained by taking the same sequences of forward and reverse thresholds. We describe next the proposed bounding models.

A. Hysteresis system with aggregated bounding arrivals

We derive from the original system ($\{X(t),\, t \geq 0\}$) bounding models (upper and lower bounds) which are equivalent to the original one except the probability distribution of size bulk arrivals of each phase in the system. The arrival process of the bounding models are defined as follows: for each phase $\phi \in L$, the arrivals of bulks follow a Poisson process with the rate λ_ϕ, and the size batches will follows a probability distribution p^u_ϕ (resp. p^l_ϕ) for the upper bound (resp. the lower bound). The probability distributions of the batches for the bounds are obtained by aggregations, according to the following relation: $p_\phi \leq_{st} p^u_\phi$ and $p^l_\phi \leq_{st} p_\phi$.

If p_ϕ is defined on a state space of size n, then p^u_ϕ (resp. p^l_ϕ) is defined on a state space of size m, such that $m \ll n$. Moreover, p^u_ϕ and p^l_ϕ are computed to be the closest distributions with m states [11], according to an increasing reward function. Intuitively, the probability distribution p^u_ϕ (resp. p^l_ϕ) has been obtained by removing some states of p_ϕ and by adding their probabilities into higher states (resp. lower states). The optimality of the computed bounding distributions, proved in [11], helps to obtain tight bounds on the results.

We denote by $X^u(t)$ (resp. $X^l(t)$) the threshold-hysteresis system built with the bulk arrival probability distribution p^u_ϕ (resp. p^l_ϕ), where $\phi \in L$. We note that the aggregation of the batch arrival distribution, allow to reduce significantly the number of non-zero element in the transition matrix, allowing therefore to reduce the computational complexity of the model. Next, we prove that these Markov chains represent stochastic bounds for $X(t)$.

1) Stochastic comparison of the systems: We define the many to one mapping function $g : A \to S$, such that for $x = (x_1, x_2, \phi) \in A$, $g(x) = (x_1, \phi)$, where $x_1 \in \{0, \ldots, C\}$. In the state space S, if we consider $u, v \in S$, where $v = (v_1, v_2)$ and $w = (w_1, w_2)$ we define the following partial order: $\forall v, w \in S,\ v \preceq w \ \Leftrightarrow v_1 \leq w_1,\ v_2 = w_2$.

Note that for states v (resp. w), v_1 (resp. w_1) represents the number of customers in the queue, and v_2 (resp. w_2) is the phase. So, the order is defined for states in the same phase, and compare the number of customers in the systems.

Theorem 2: We have the following relations:

- $g(X(0)) \preceq_{st} g(X^u(0)) \Rightarrow g(X(t)) \preceq_{st} g(X^u(t)),\ t > 0.$
- $g(X^l(0)) \preceq_{st} g(X(0)) \Rightarrow g(X^l(t)) \preceq_{st} g(X(t)),\ t > 0.$

Proof: We use Theorem 1 based on the coupling of the processes. We begin with the first relation of Theorem 2, in order to establish that $\{X^u(t), t \geq 0\}$ is really an upper bound. For the proof, we suppose that at time t, $g(X^u(t)) = v = (v_1, v_2)$, and $g(X(t)) = w = (w_1, w_2)$, where $v_2 = w_2 = \phi$. The proof is by induction. We suppose that the order is verified at time t ($v \preceq w$), and we prove that at time $t + dt$ the order is still verified. We denote by $g(X(t + dt)) = v'$ and $g(X^u(t + dt)) = w'$. We consider the two kinds of events:

- *Arrivals:* if we have a batch arrival of size k in $X(t)$ such that at time $t + dt$, $v'_1 = v_1 + k$, then we can have also a transition from w to w' such that $w'_1 = w_1 + l$, and $k \leq l$, as $p_\phi \leq_{st} p^u_\phi$. So $v' \preceq w'$, and the order is still verified at time $t + dt$.
- *Services:* if we have a service for $X^u(t)$ such that at time $t + dt$, $w'_1 = w_1 - 1$, then we can have also a service in $X(t)$ such that at time $t + dt$, we have $v'_1 = v_1 - 1$, as the transition rates are the same in the two systems.

For the lower bound $\{X^l(t), t \geq 0\}$, the proof is similar as we consider the stochastic ordering between the batch arrival probability distributions: $p^l_\phi \leq_{st} p_\phi$, and the same service rates in the two systems. Then, the second relation of Theorem 2 is verified. Note that as the stochastic comparison of the processes is made by the mapping g, then it allows to compare the processes from the number of customers waiting in the system. So, this generates the comparison on performance measures as the mean number of customers waiting in the system, and blocking probabilities.

2) Bounds on performance measures: Let $\pi^{(t),u}$ (resp. $\pi^{(t),l}$) be the transient distribution of $\{X^u(t),\ t \geq 0\}$ (resp. $\{X^l(t),\ t \geq 0\}$). From Theorem 2, we have the following proposition:

Proposition 1: $\forall\, a \in \{0, \ldots, C\}$, and $\forall\, \phi \in L$, we have:

$$\sum_{x_1 \geq a} \sum_{x_2} \pi^{(t)}(x_1,\, x_2,\, \phi) \leq \sum_{x_1 \geq a} \sum_{x_2} \pi^{(t),\,u}(x_1,\, x_2,\, \phi)$$

$$\text{and} \sum_{x_1 \geq a} \sum_{x_2} \pi^{(t),\,l}(x_1,\, x_2,\, \phi) \leq \sum_{x_1 \geq a} \sum_{x_2} \pi^{(t)}(x_1,\, x_2,\, \phi).$$

As in proposition 1 the inequalities are considered for each phase, then we have the following inequality (for the upper bound):

$$\sum_{x_1 \geq a} \sum_{x_2} \sum_{\phi} \pi^{(t)}(x_1,\, x_2,\, \phi) \leq \sum_{x_1 \geq a} \sum_{x_2} \sum_{\phi} \pi^{(t),\,u}(x_1,\, x_2,\, \phi).$$

Obviously, for the lower bound, the inequality on the probability distributions is reversed. As the expectation of the number of customers is an increasing function, we deduce that the relation is maintained. Idem for blocking probabilities.

B. Bounding systems with equal forward & reverse sequence

Considering the same forward and reverse thresholds vectors, we derive upper and lower bounding models for the threshold queueing system with hysteresis. For the upper bound, we take $(F_1, \ldots, F_{K\text{-}1})$ as a forward and reverse thresholds. And, for the lower bound, we take $(R_1, \ldots, R_{K\text{-}1})$ as forward and the reverse thresholds. This modification allows us to reduce the size of the state space of $X(t)$.

The behavior of each of these systems are represented by CTMCs defined on S. We denote by $Y(t)$ the CTMC associated to the upper bounding model.The evolution equation of this model is given as follows (where ϕ, $\phi' \in L$ and $F_0 = 0$):

$$\begin{aligned}
(x,\, \phi) \quad &\to \quad \min((C,\, x_1 + k), \phi'), && (2)\\
&\qquad \text{with rate } \lambda_\phi p_\phi(k)\, \mathbf{M}(\phi, \phi'),\ \ \forall\, k \in E \\
&\to \quad \max((0,\, x_1 - 1), \phi), \text{with rates:} && (3)\\
&\qquad \bullet\ i\mu, \text{ if } F_{i-1} < x_1 \leq F_i,\ \forall i = 1 \ldots K-1 \\
&\qquad \bullet\ K\mu, \text{ if } F_{K-1} < x_1 \leq C
\end{aligned}$$

In the same way, we define by $Z(t)$ the CTMC which represents the lower bound model. In this case, the above equations ((4)-(5)) are also available by changing the sequence $F_{i,\, i=1\ldots K-1}$, by the sequence $R_{i,\, i=1\ldots K-1}$.

Next, we will prove that $Y(t)$ (resp. $Z(t)$) is a stochastic upper bound (resp. lower bound) for $X(t)$.

1) Stochastic comparison proofs: We have the following:

Theorem 3: If $X(t)$, $Y(t)$, and $Z(t)$ represent the systems defined previously, then we have:

1) $g(X(0)) \preceq_{st} Y(0) \Rightarrow g(X(t)) \preceq_{st} Y(t), \forall t > 0$
2) $Z(0) \preceq_{st} g(X(0)) \Rightarrow Z(t) \preceq_{st} g(X(t)), \forall t > 0$

Proof: We begin by the first equation of Theorem 3. For the proof, we apply Theorem 1. In our case, we have $X(t)$ defined on A, g a mapping function $A \to S$, $Y(t)$ defined on S, and h is the identity function. The proof is done by induction: we consider $X(t) = (x_1, x_2, \phi)$, so $g(X(t)) = x = (x_1, \phi)$ and $Y(t) = y = (y_1, \phi)$, such that $g(x) \preceq y$ and we prove that for any event, at time $t + dt$, if $X(t + dt) = x' = (x'_1, x'_2, \phi')$, and $Y(t + dt) = y'$, then $g(x') \preceq y'$. In fact, for states x and y such that $g(x) \preceq y$, we consider the two events for the proof of the stochastic comparison by the mapping function g:

- *Arrivals:* if we have a transition from x to x' such that $x'_1 = x_1 + k$ (for $k > 0$) then we are sure that we can have also a transition from y to y' such that $y' = y + l$ (for $l > 0$), because the arrival rates in the phase ϕ are the same in the two systems. So, we deduce that $g(x') \preceq y'$.
- *Services*: if we have a transition from y to y' such that $y'_1 = y_1 - 1$, then we are sure that we have also a transition from x to x' such that $x'_1 = x_1 - 1$, because the rate for decreasing in $Y(t)$ is lower than in $X(t)$ (as the threshold for deactivation is lower in $X(t)$ than in $Y(t)$). So, $g(x') \preceq y'$.

So, we deduce that for any events, we have: $g(X(t+dt)) \preceq Y(t + dt)$, and from Theorem 1, we deduce that in Theorem 3, equation 1 is verified. For the second equation of Theorem 3, the proof is similar to deduce that $Z(t)$ is a lower bound.

2) Bounds on performance measures: Let denote by π, π_Y and π_Z the steady state distributions of $\{X(t), t \geq 0\}$, $\{Y(t), t \geq 0\}$ and $\{Z(t), t \geq 0\}$). From the Theorem 3, we have the following proposition:

Proposition 2: $\forall \phi \in L$, $\forall a \in \mathbb{N}^+$ and $\forall x_2 \in \{1, \ldots, K\}$, we have: $\sum_{x_1 \geq a} \sum_{x_2} \pi^{(t)}(x_1, x_2, \phi) \leq \sum_{x_1 \geq a} \pi^{(t)}_Y(x_1, \phi)$, and

$$\sum_{x_1 \geq a} \pi^{(t)}_Z(x_1, \phi) \leq \sum_{x_1 \geq a} \sum_{x_2} \pi^{(t)}(x_1, x_2, \phi).$$

We deduce from the Proposition 2 the comparison on performance measures.

C. Other bounding models

An other simplification of previous bounding models consists to aggregate the batch arrival distribution of each phase for $Y(t)$ and $Z(t)$. We denote by $Y^u(t)$ (resp. $Z^l(t)$) the Markov chain with batch arrival distribution p^u_ϕ (resp. p^l_ϕ), for all $\phi \in L$.

Theorem 4: We have the following relations:

- $Y(0) \preceq_{st} Y^u(0) \Rightarrow Y(t) \preceq_{st} Y^u(t),\ t > 0.$
- $Z^l(0) \preceq_{st} Z(0) \Rightarrow Z^l(t) \preceq_{st} Z(t),\ t > 0.$

The proof is similar to Theorem 2.

By transitivity, from Theorem 3 and Theorem 4, we deduce that $Y^u(t)$ (resp. $Z^l(t)$) represents an upper bound (resp. lower bound) of $X(t)$.

To illustrate the relevance of proposed models, we present next some numerical results.

V. Numerical examples

We propose in this section to illustrate the impact of our bounding models on satisfying the QoS constraints and reducing the computational complexity of a threshold-based queue with hysteresis and batch phase-type arrival process. For this example, we propose to compute some performance measures as blocking probabilities and expected buffer length.

We consider a threshold-based queue with hysteresis, two phases-type input process with batches and 50 homogeneous servers. For the phase-type arrivals, we propose to illustrate the batch probability distribution of the phases by considering the Facebook cluster trace, namely the Statistical Workload Injector for MapReduce (SWIM) [12]. Typically, this is the workload replay scripts to generate the real-life workloads from a Facebook production system. This realistic data set synthesized day-long workloads, namely Facebook trace 02 (*FB-2010 samples 24 times 1hr 0.tsv*) is studied. The set contains 24 historical traces sampled on a 600-machines cluster. For a sampling period of 60 seconds, Figure 2 shows the trace of data size input (input size jobs in bytes) per slot.

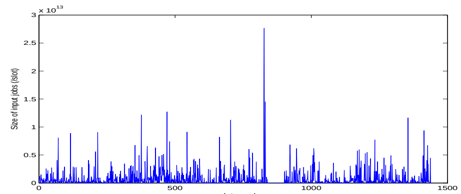

Fig. 2: Traffic trace of day-long workload from a Facebook production system (*FB-2010 samples 24 times 1hr 0.tsv*).

From this trace, we propose to distinguish two phases, phase 1 with a high arrival traffic rate and small data size, and phase 2 with low arrival rate and larger data size. So, we consider that for data size less than 800 GB (Gigabytes) the arrival phase is 1 otherwise the arrival phase is 2. The probability transition matrix for phase modulation is obtained as follows:

$$\mathbf{M}(i, j) = \frac{\text{number of transitions from phase } i \text{ to phase } j}{\text{number of batch in phase } i}, \forall i, j \in L.$$

Therefore, the resulting transition matrix of phases $\mathbf{M}$ is:

$$\mathbf{M} = \begin{pmatrix} 0.7758 & 0.2233 \\ 0.7164 & 0.2836 \end{pmatrix}, \text{ where } L \in \{1, 2\}.$$

In order to analyze this model, we consider a queue with service rate of $5\,10^3$ jobs/mn, an arrival rates λ_1=100 jobs/mn and λ_2=50 jobs/mn with an utilization of 80% of the system, and we vary the buffer size from 10 TB to 100 TB. According to the buffer size, we use the following equation to define respectively the forward and the reverse threshold vectors: $F = (\lfloor \frac{C}{K} \rfloor, 2 \times \lfloor \frac{C}{K} \rfloor, \ldots, (K\text{-}1) \times \lfloor \frac{C}{K} \rfloor)$ and $R_i = F_i\text{-}\lfloor \frac{C}{2K} \rfloor$, for $i = 1, \ldots, K\text{-}1$. In this example, we introduce a data unit which gathers a constant number of bytes (say D) in order to reduce the sizes, and hence the complexity of underlying models. So, for these numerical computations, we have taken a data unit equal to 1Gb which is the batch unit. Depending on the value of the buffer size (C), we present in tables I, II the blocking probabilities and the expected buffer length obtained for the different studied models at steady state. We report also in Table III the computation times needed to solve these models. We note that for the models: $X^l(t)$, $X^u(t)$, $Z^l(t)$ and $Y^u(t)$, the reduction applied on the distributions of the data size for the two phases is $m = 10$.

From tables I and II, we remark that the proposed models provide clearly bounds on performance measures of $X(t)$ (lower bounds for the models $X^l(t)$, $Z(t)$ and $Z^l(t)$ and upper bounds for the models $X^u(t)$, $Y(t)$ and $Y^u(t)$). We can also see that the results provided by our models are very accurate (the bounds are close to the exact results).

(a)

C (GB)	$X(t)$	$X^l(t)$	$X^u(t)$
10^4	0.000381892	0.000342812	0.000456201
$3\,10^4$	7.64789 e-5	6.96193 e-5	8.00854 e-4
$5\,10^4$	2.82538 e-5	1.42637 e-5	4.09931 e-5
10^5	5.82711 e-6	5.51383 e-6	6.23808 e-6

(b)

C	$Z(t)$	$Z^l(t)$	$Y(t)$	$Y^u(t)$
10^4	0.0003619	0.0003328	0.0004020	0.0004762
$3\,10^4$	6.2792 e-5	5.3860 e-5	8.2775 e-5	8.91217 e-5
$5\,10^4$	2.3361 e-5	1.4314 e-5	3.2376 e-5	4.3175 e-5
10^5	5.1265 e-6	5.5008 e-6	5.9516 e-6	6.2713 e-6

TABLE I: Blocking probabilities versus buffer size.

(a)

C	$X(t)$	$X^l(t)$	$X^u(t)$
10^4	8.6014e+3	8.40617e+3	8.72384e+3
$3\,10^4$	1.60545e+4	1.59461e+4	1.61962e+4
$5\,10^4$	2.56954e+4	2.5586e+4	2.58423e+4
10^5	4.83833e+4	4.82183e+4	4.86226e+4

(b)

C	$Z(t)$	$Z^l(t)$	$Y(t)$	$Y^u(t)$
10^4	8.6009e+3	8.4057e+3	8.6067e+3	8.725e+3
$3\,10^4$	1.6037e+4	1.5985e+4	1.6197e+4	1.6348e+4
$5\,10^4$	2.5632e+4	2.5602e+4	2.5947e+4	2.6141e+13
10^5	4.8190e+4	4.8052e+4	4.8820e+4	4.9222e+4

TABLE II: Expected buffer length versus buffer size (in GB).

C (GB)	$X(t)$	$X^l(t)$	$X^u(t)$	$Z(t)$	$Z^l(t)$	$Y(t)$	$Y^u(t)$
10^4	136	20	23	99	11	107	21
$3\,10^4$	277	34	39	122	21	100	29
$5\,10^4$	406	51	53	242	34	259	35
10^5	1465	111	119	928	76	1001	97

TABLE III: Computation times (seconds) versus buffer size.

Regarding to the execution times (Table III), we observe that the bounding models are less complex than the original one, particularly for the models with an aggregation of the batch arrival distributions.

Indeed, we notice that the models $X^l(t)$, $X^u(t)$, $Z^l(t)$ and $Y^u(t)$ allows to reduce the computational time of the model $X(t)$ by 4, 6 and even by 15 depending on the size of the buffer. To conclude, we can say that the exact performance metrics of the original model can be bounded using the proposed models with relatively small computation complexity. And, we recall also that the purpose of using these bounding models is to offer to the user an interesting trade-off between accuracy of the results and the computational complexity in order to satisfy the required QoS constraints.

VI. Conclusion

We propose in this paper to model the behavior of a data center in a cloud system by a queue-dependent multi-server VMs with hysteresis. The relevance of this model is to represent the dynamicity of the resource according to the queue occupation, and the variable intensity of the demand by a phase-type process arrival with batches. As this system could be difficult to analyse when the state space increases, we propose here to use bounding techniques in order to derive bounds on the performance measures. We show clearly in this paper the relevance of the proposed bounding models in term of accuracy of the results and reduction of computational complexity. As a future work, we try to adapt our methodology to analyze hysteresis model with more general input process and we also investigate methods which allow us to define an optimal thresholds vectors in order to optimize the performance of cloud systems.

Acknowledgment

This work was supported by grant ANR MARMOTE (ANR-12-MONU-0019).

References

[1] H. Khazaei, J. V. Misic, and V. B. Misic, "Performance of cloud centers with high degree of virtualization under batch task arrivals," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 12, pp. 2429–2438, 2013.

[2] R. Ghosh, F. Longo, V. K. Naik, and K. S. Trivedi, "Modeling and performance analysis of large scale iaas clouds," *Future Generation Computer Systems*, vol. 29, no. 5, pp. 1216–1234, 2013.

[3] O. C. Ibe and J. Keilson, "Multi-server threshold queues with hysteresis," *Perform. Eval.*, vol. 21, no. 3, pp. 185–213, 1995.

[4] T. V. Do, R. Chakka, and J. Sztrik, *Advanced Computational Methods for Knowledge Engineering*. Heidelberg: Springer International Publishing, 2013, pp. 131–142.

[5] J. C. S. Lui and L. Golubchik, "Stochastic complement analysis of multi-server threshold queues with hysteresis," *Performance Evaluation*, vol. 35, pp. 19–48, 1999.

[6] F. Aït-Salaht and H. Castel-Taleb, "Stochastic bounding models for performance analysis of clouds," in *The 15th IEEE International Conference CIT*, 2015, pp. 1–8.

[7] W. Stewart, *Introduction to the numerical Solution of Markov Chains*. New Jersey: Princeton University Press, 1995.

[8] O. Gusak, T. Dayar, and J.-M. Fourneau, "Iterative disaggregation for a class of lumpable discrete-time stochastic automata networks." *Perform. Eval.*, vol. 53, no. 1, pp. 43–69, 2003.

[9] A. Müller and D. Stoyan, *Comparison methods for stochastic models and risks*, ser. Wiley series in probability and statistics. J. Wiley & sons, 2002.

[10] M. Doisy, "A coupling technique for stochastic comparison of functions of markov processes," *JAMDS*, vol. 4, no. 1, pp. 39–64, 2000.

[11] F. Aït-Salaht, H. Castel-Taleb, J. Fourneau, and N. Pekergin, "Stochastic bounds and histograms for network performance analysis," in *The 10th European Workshop, EPEW*, ser. Lecture Notes in Computer Science, vol. 8168. Springer, 2013, pp. 13–27.

[12] Y. Chen, A. Ganapathi, R. Griffith, and R. Katz, "The case for evaluating mapreduce performance using workload suites," in *The 19th IEEE Annual International Symposium MASCOTS*, 2011, pp. 390–399.