Automated Control in Cloud Computing: Challenges and Opportunities

Harold C. Lim[1], Shivnath Babu[1], Jeffrey S. Chase[2], Sujay S. Parekh[2]
Duke University, NC, USA[1], IBM T.J. Watson Research Center[2]
ACDC '09 Proceedings of the 1st workshop on Automated Control for Datacenters and Clouds
Pages 13-18

Tópicos em Computação em Nuvem

Bruno Bastos Rodrigues

brodrigues@larc.usp.br

Roteiro

- Introdução
- Fontes, trabalhos relacionados e derivados
- Pontos altos/citações
- Principais contribuições do artigo
- Análise crítica & Conclusões

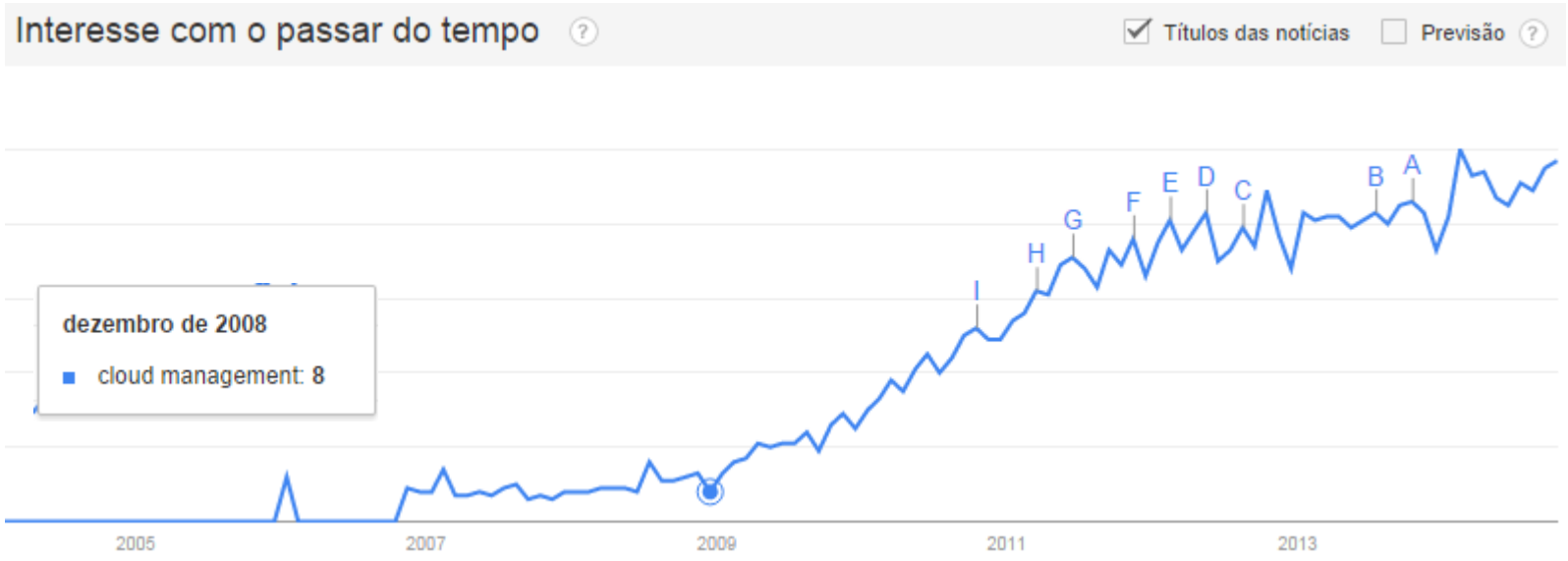
Introdução

- Qual é o problema a ser resolvido?

Construção de um controlador eficiente para locação de recursos de um provedor de IaaS utilizando a técnica "proportional thresholding".

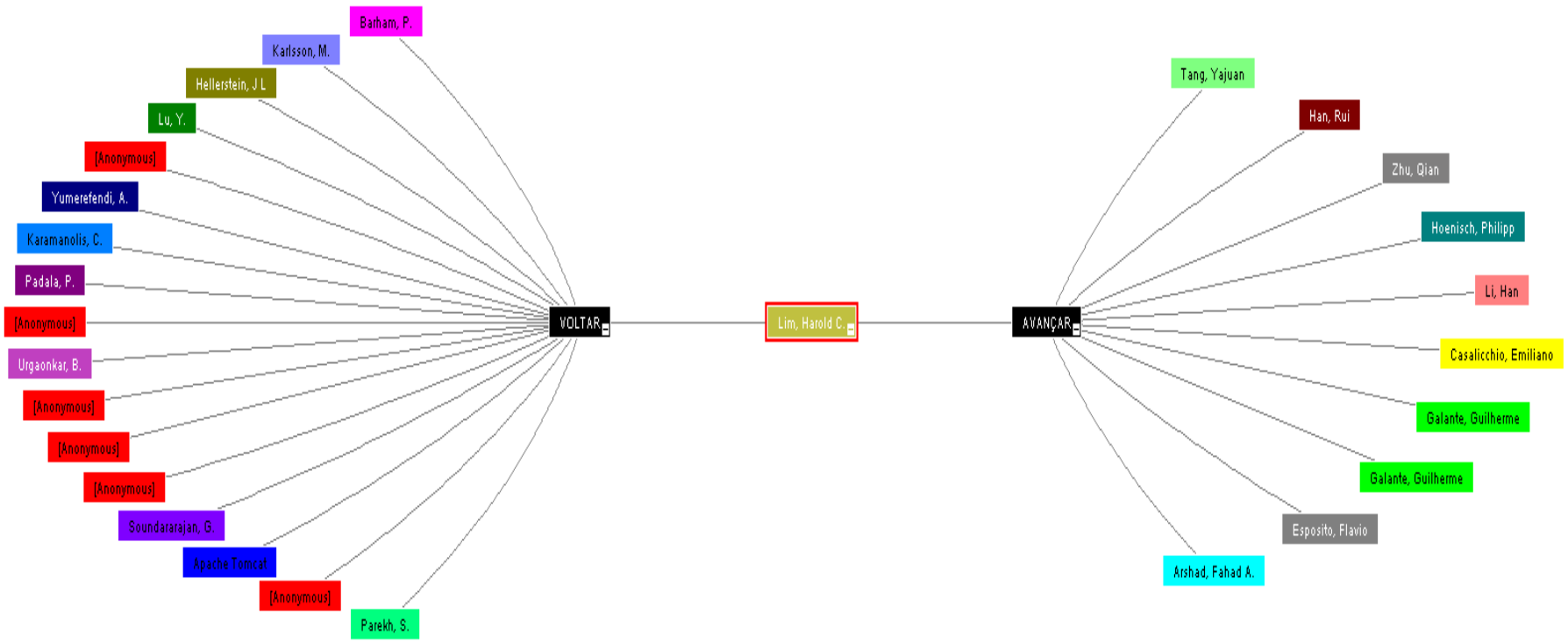
Introdução

- Google trends
- Palavras-chave sugeridas não retornam resultados específicos para o tema:
  - “automated control”, “feedback control”
- Termo utilizado*: “cloud management”*



Fontes e trabalhos relacionados

- *Referências e Citações: Web of Science*



17 Referências

10 Citações

Fontes e trabalhos relacionados

- Citações:
  - *Web of Science / ACM / Google Schoolar*

2010 ◀	2011	2012	2013	2014 ▶	Total	Média de citações por ano
0	0	3	6	2	11	1.83

Web of Science

Bibliometrics

- Downloads (6 Weeks): 42
- Downloads (12 Months): 315
- Downloads (cumulative): 3,605
- Citation Count: 32

ACM

Automated control in **cloud computing**: **challenges** and **opportunities**
HC Lim, S Babu, JS Chase, SS Parekh - ... workshop on **Automated control** ..., 2009 - dl.acm.org
Abstract With advances in virtualization technology, virtual machine services offered by **cloud** utility providers are becoming increasingly powerful, anchoring the ecosystem of **cloud** services. Virtual **computing** services are attractive in part because they enable customers ...
Citado por 155 Artigos relacionados Todas as 17 versões Citar Salvar

Google Schoolar

Storyline do Artigo

- 1. Introduction
- *2. Feedback Control in a Cloud Computing Infrastructure*
  - *2.1. Decoupled Control*
  - *2.2. Control Granularity*
  - *2.3. Other API Restrictions*
- *Proportional Thresholding*
  - *3.1. Design of Proportional Thresholding*
- *4. Example*
  - *4.1. Prototype Control System*
  - *4.2. Control Parameters*
- *5. Evaluation*
- *Conclusion*

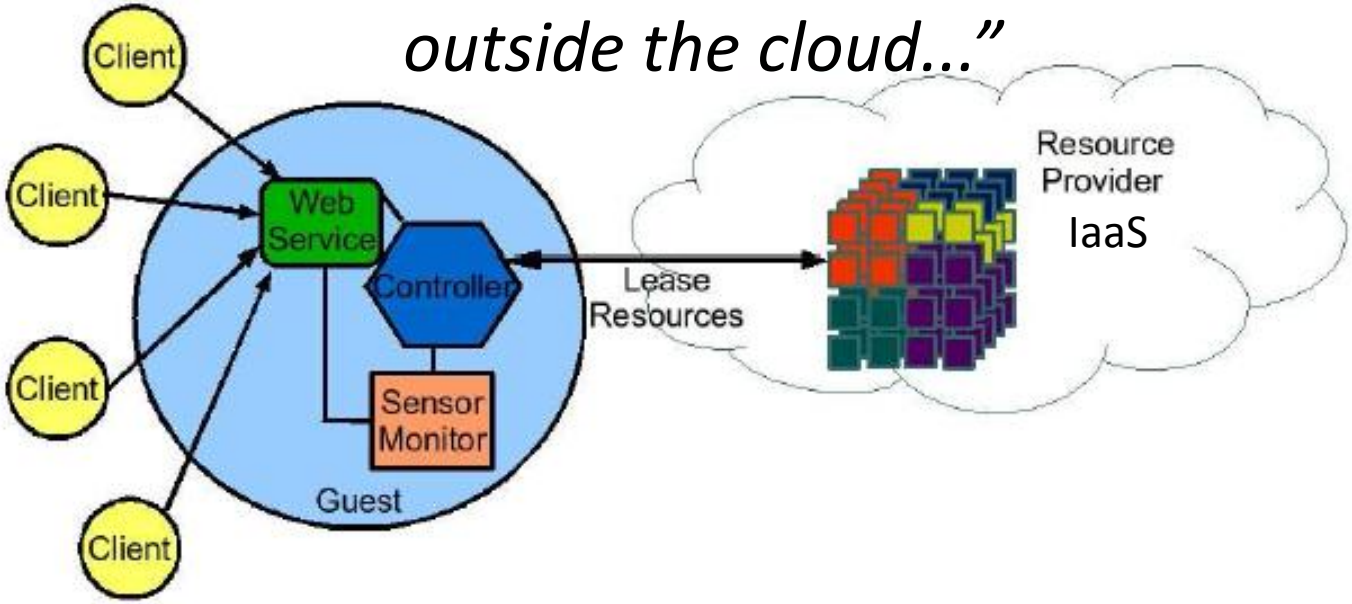
Pontos altos/citações

1. Introduction

Esclarecimento do problema a ser atacado

"This paper focuses on cloud computing infrastructure services"

"...building an effective controller as a customer add-on outside the cloud..."



Pontos altos/citações

2. Feedback Control in a Cloud Computing Environment

2.1. – Decoupled Control

Requisitos:

"...application control should be factored out of the cloud platform and left to the guest..."

"A principled layering offers the best potential for guests to innovate in their control policies..."

Pontos altos/citações

2. Feedback Control in a Cloud Computing Environment

2.2. – Control Granularity

"There may be a gap between the sensors and actuators desired by the slice controllers and those exposed by the resource providers. Providers may hide useful information to preserve their flexibility...."

"The slice controller must make the best of whatever sensors and actuators are available, and whatever control intervals and granularity the provider allows"

Requisito: uma política de controle não deve considerar que a nuvem ofereça informações muito precisas.

Pontos altos/citações

2. Feedback Control in a Cloud Computing Environment

2.3. – Other API Constraints

Sensores

"...the slice controller depends on the cloud platform for accurate sensor measurements at the resource level..."

"...[cloud platform] expose well-specified sensors that are suitable for stable control..."

Atuadores

"Request routing solutions may involve programmable network elements installed by the host provider

Pontos altos/citações

3. Proportional Thresholding

Static thresholding

"only one target goal"

Proportional thresholding

"...the goal is turned into a target range, defined by a high and a low threshold. Thus, the system is considered on target when the sensor measurement falls inside the target range"

Pontos altos/citações

3. Proportional Thresholding

3.1. Design of Proportional Thresholding

Static thresholding

$$u_{k+1} = u_k + K_i \times (y_{ref} - y_k)$$

Proportional thresholding

$$u_{k+1} = \begin{cases} u_k + K_i \times (y_h - y_k) & \text{if } y_h < y_k \\ u_k + K_i \times (y_l - y_k) & \text{if } y_l > y_k \\ u_k & \text{otherwise} \end{cases}$$

***uk+1**: novo valor do atuador*

***uk**: valor atual do atuador*

***Ki**: parâmetro integral de ganho (constante)*

***Yref**: valor alvo do sensor*

***Yk**: valor atual do sensor*

Yh: limite superior para o valor alvo do sensor

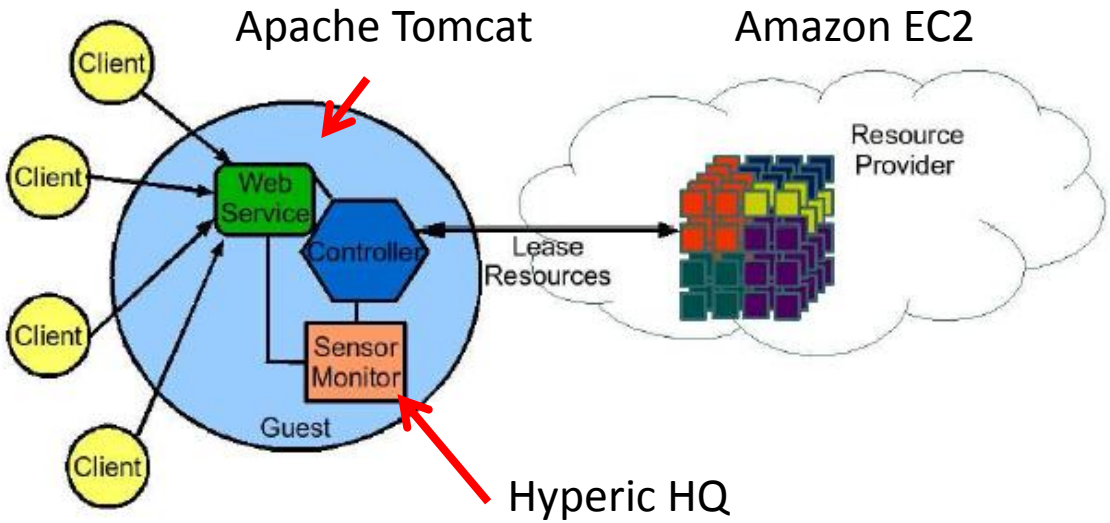
Yl: limite inferior para o valor alvo do sensor

Pontos altos/citações

4. Example

4.1. Prototype Control System

- Configuração do protótipo:
  - Tomcat Server Cluster
  - API fornecida pela **Amazon EC2**
  - **Hyperic HQ** para coletar o uso de CPU de todas VMs alocadas



Pontos altos/citações

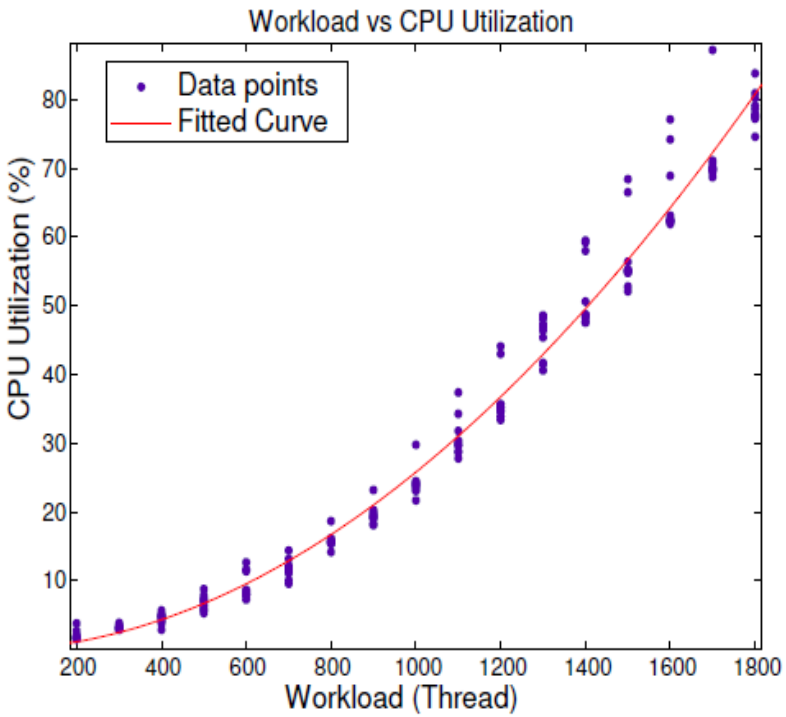
4. Example

4.2. Control Parameters

"we empirically measured the CPU utilization of a single machine under various workloads to formulate the equation for Yl"

Yl: limiar inferior

Yh: limiar superior



$$y_{k+1} = .8819y_k - .5892u_k$$

$$K_i = -.07$$

"From this data, we find the best-fit curve:

$$\mathtt{CPU} = (3.869 \times 10^{-5}) \times \mathtt{workload}^{1.947}.$$

Dado um limite superior Yh... Encontrar Yl:

$$\mathtt{workload}_{est} = \left(\frac{y_h}{3.869 \times 10^{-5}}\right)^{\frac{1}{1.947}}$$

$$\mathtt{workload}_{tot} = \mathtt{workload}_{est} \times (\mathtt{currVM} - 1).$$

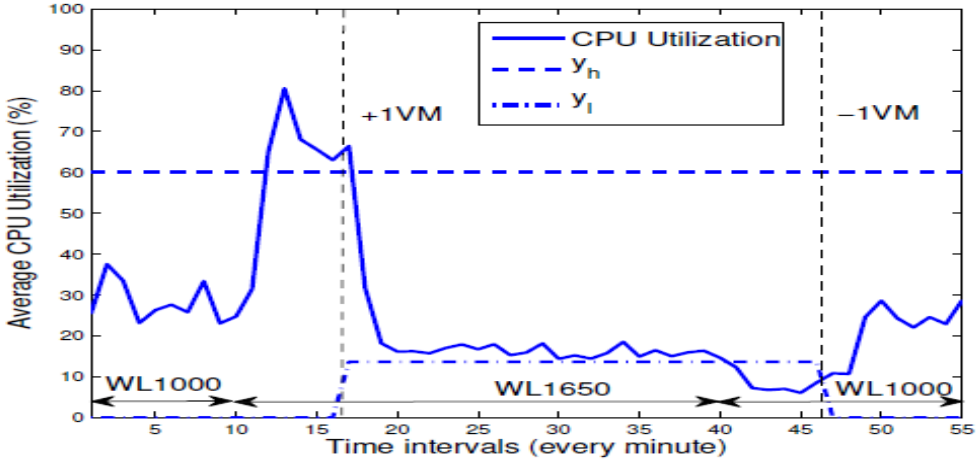
$$\mathtt{workload}_{low} = \mathtt{workload}_{est} \times \frac{\mathtt{currVM} - 1}{\mathtt{currVM}}$$

$$\boxed{y_l = y_h \times \left(\frac{\mathtt{currVM} - 1}{\mathtt{currVM}}\right)^{1.947}}$$

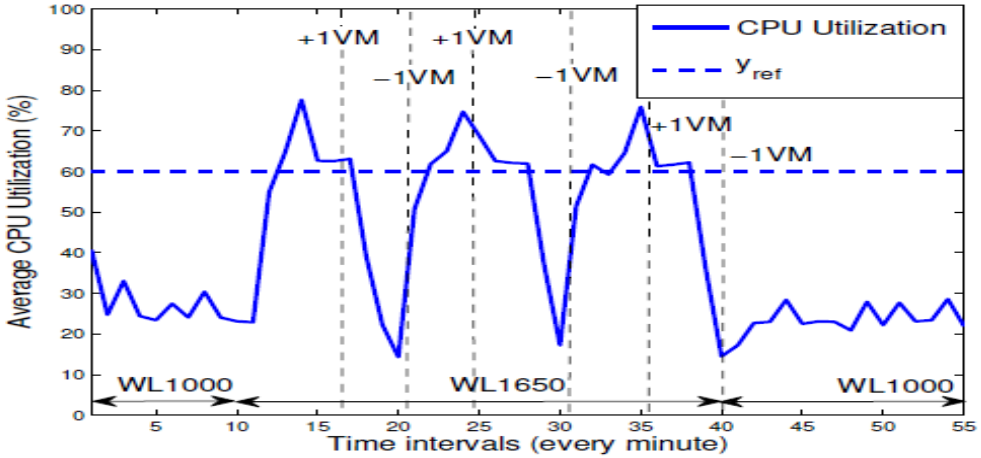
Pontos altos/citações

5. Evaluation

Proportional Thresholding
X
Integral Control



(a) Proportional Thresholding

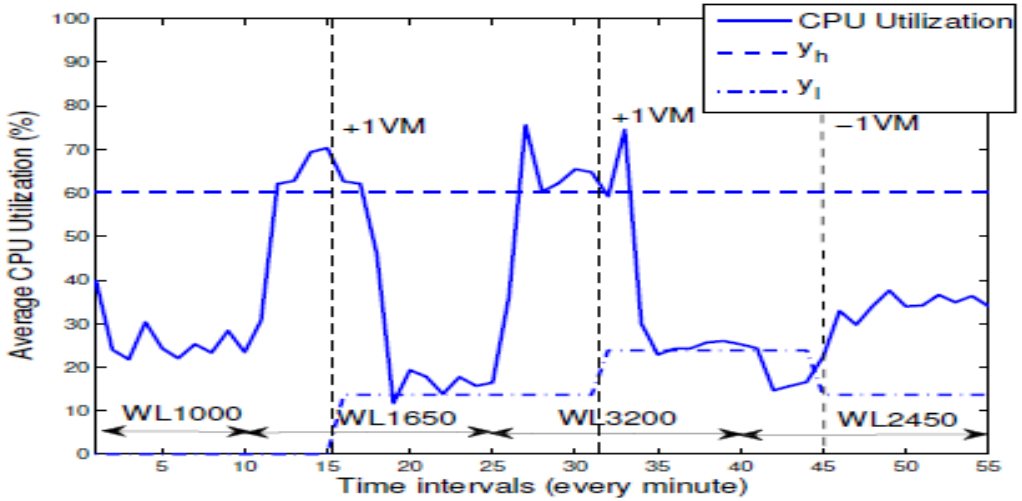


(b) Integral Control

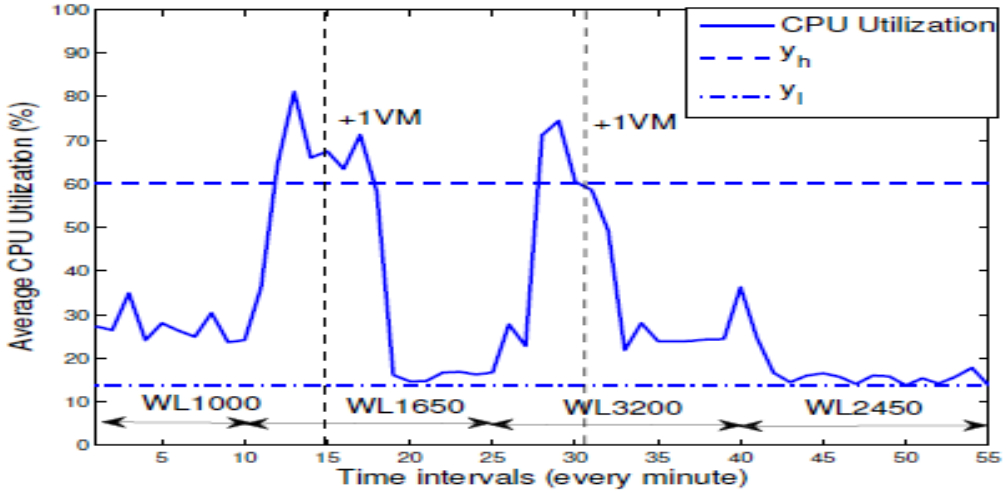
Pontos altos/citações

5. Evaluation

Proportional Thresholding
X
Static Thresholding



(a) Proportional Thresholding



(b) Static Thresholding

Análise crítica e Conclusões

- A solução possui diversos requisitos:
  - Controle de aplicação descentralizado do controlador da nuvem
  - API da nuvem deve fornecer sensores e atuadores para o controlador da aplicação
  - Ambiente de teste e configuração não estão bem esclarecidos.
- Por que é uma publicação de destaque?
  - Ataca um grande problema: “como utilizar os recursos de modo eficiente?”
  - A solução em si apesar de específica, foi inovadora.

Discussão