

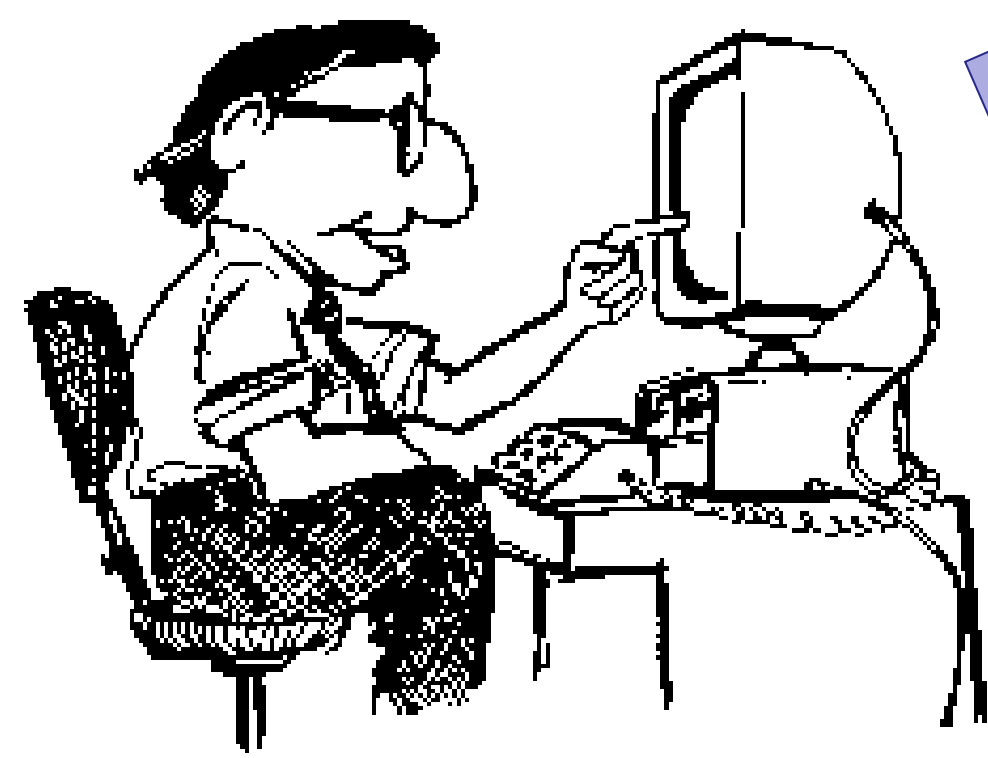
Enhancing Cloud Elasticity Via Software Performance Test Automation and Learning

Md Abu Naser Bikas, Abdullah Alourani and Mark Grechanik

Department of Computer Science, University of Illinois, Chicago

1 Motivation

In cloud computing, consumers can acquire resources on demand and pay only for the resources used by the application. Understanding when and how to reallocate resources is a hard problem, since provisioning resources precisely and automatically requires the knowledge of specific application behaviors for given inputs.



http requests (inputs)

Provisioning Strategy

These performance rules are used to automatically search for application-specific provisioning strategies during performance testing. These provisioning strategies guide the cloud to (de)allocate resources to applications to improve their throughputs.

Rule	Provisioning Strategy
R1	$sres(P,3,VMi) \wedge sinst(2,VMn)$
R2	$sres(R,1,VMi) \wedge dinst(2,VMn)$
R3	$sres(P,2,VMi) \wedge sres(R,2,VMi)$

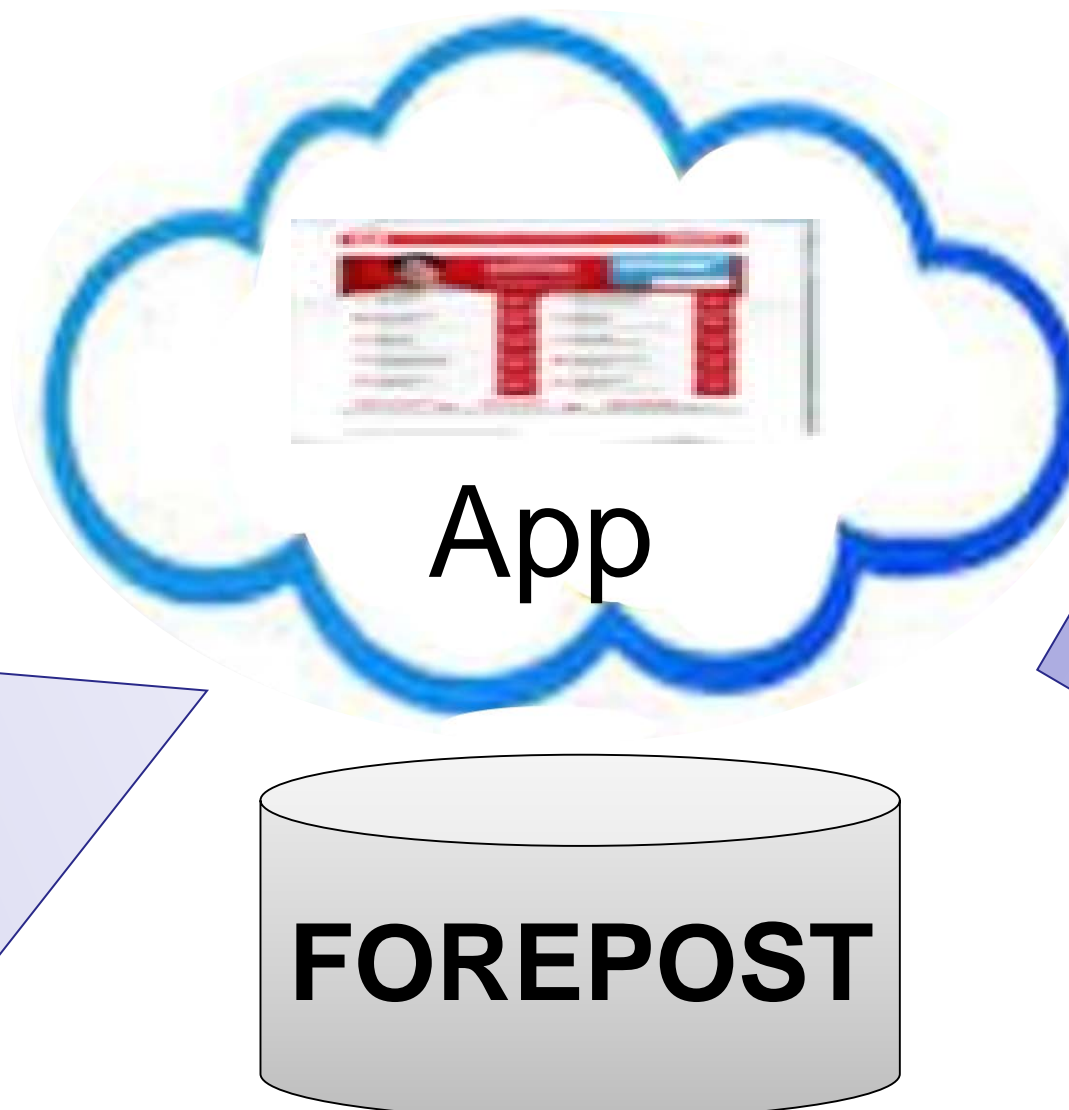
Learning provisioning strategy

2 Problem Statement

How to precisely know what type of and what quantity of resources to allocate to an application in the cloud for certain types of input data and their combinations in the load and when to deallocate what resources to enable elasticity with a high degree of precision.

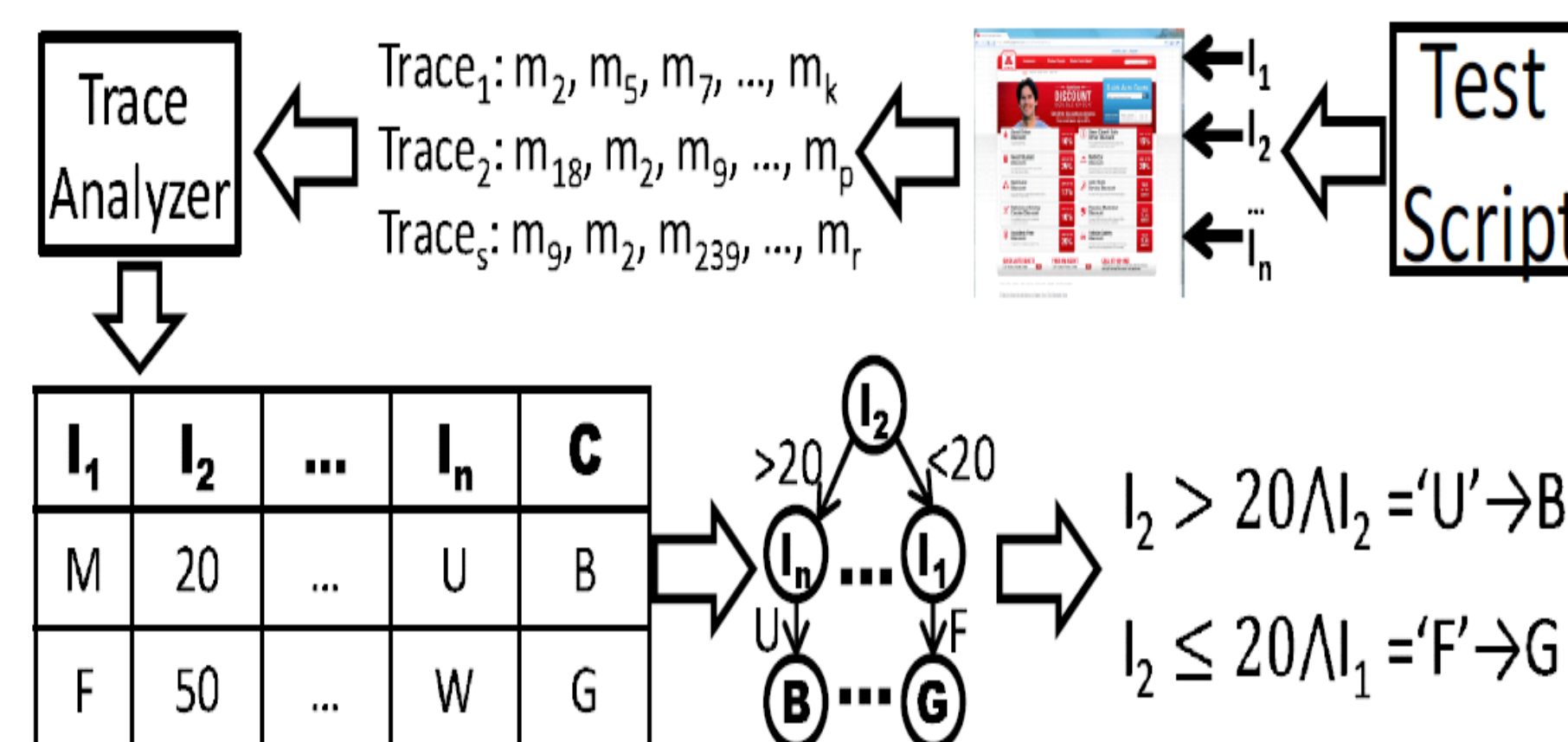
Example Rules
$(viewCtgr_CATS \geq 25) \text{ and } (viewCtgr_FISH \geq 10)$
$(addItem_EST-7 \geq 5) \text{ and } (viewCtgr_DOGS \geq 20)$

performance rules



FOREPOST

We use our tool for finding performance problems automatically by learning and using rules that describe classes of input data that lead to intensive computations.



Learning performance rules

Learning Provisioning strategy

update

Hot-add feature

add resources on-the-fly while the virtual machine is running ("hot") to meet an application's demand

inform load-balancer

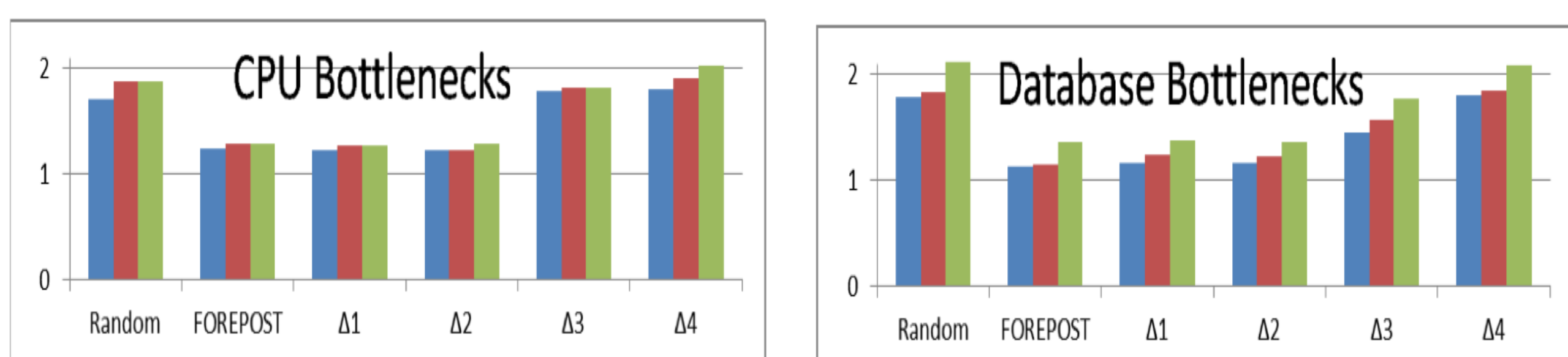
Load Balancer

distribute load

Cloud

responses

4 Experimental Results



Throughputs for JPetStore using Cloudstack and PRESTO

3 Our idea is to improve the cloud elasticity to match the application's demand

We propose an approach, coined as Provisioning Resources with performance Software Test automation (PRESTO), to automatically learn behavioral models of software applications during performance testing in order to synthesize provisioning strategies that guide the cloud to (de)allocate resources to these applications.

With PRESTO, not only we can improve cloud elasticity and to save customers money and to improve the performance of their applications, but also we can test for violations when over-provisioning (i.e., allocating more resources than required) and SLA violation occur at the same time in the cloud, which called the **Double-Whammy** problem.