



WORKSHOP MODCS 2017.2

Proposta

Proposição de Modelos Estocásticos de Desempenho Para Suportar o Planejamento de Aplicações de Nuvem Móvel em Infraestruturas de Nuvem Pública

Thiago Felipe da Silva Pinheiro
tfs3@cin.ufpe.br

Orientador: Dr. Paulo Romero Martins Maciel



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO

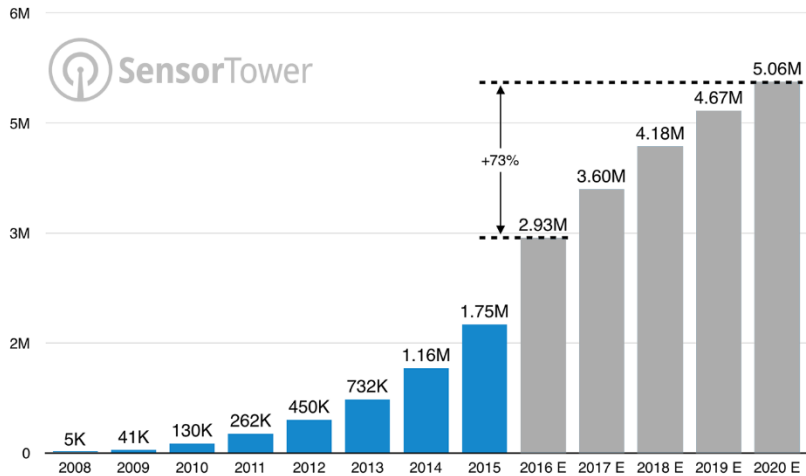
70 ANOS
TEMPOS TRANSVERSOS



Agenda



Total Number of Active Apps on the App Store by 2020 - Worldwide

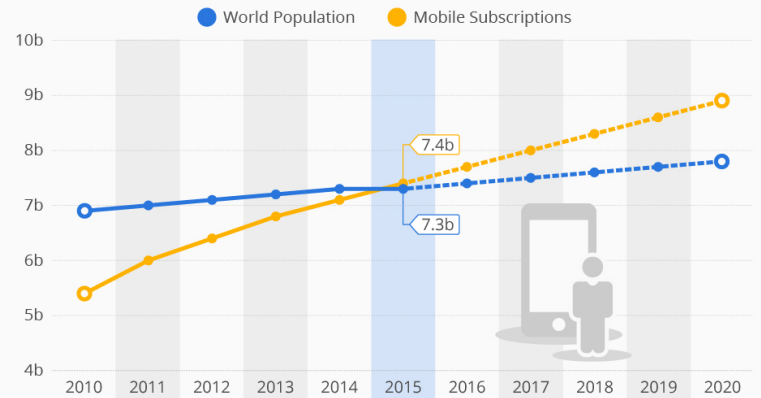


SensorTower Data That Drives App Growth

sensortower.com

Mobile Subscriptions to Outnumber the World's Population

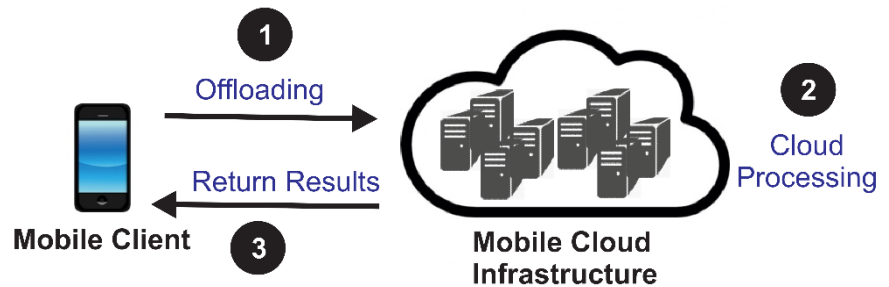
World population vs. estimated number of worldwide mobile subscriptions



StatistaCharts

Sources: Ericsson, United Nations

statista



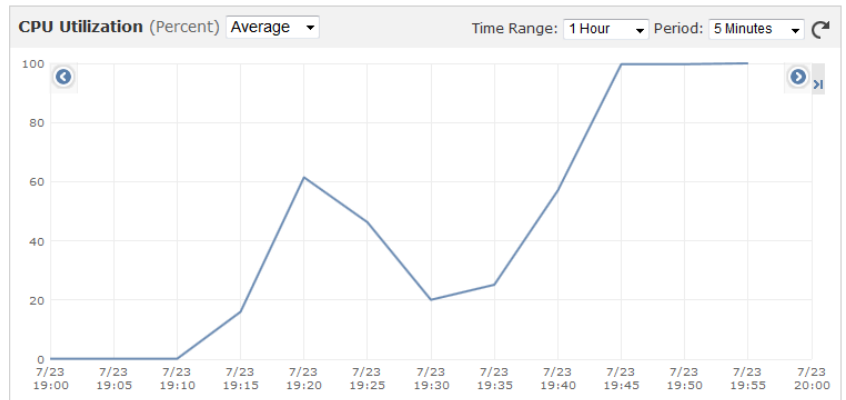


Contexto



CloudWatch Monitoring Details

Cancel X



Monitored Instances: ■ i-91331dfe

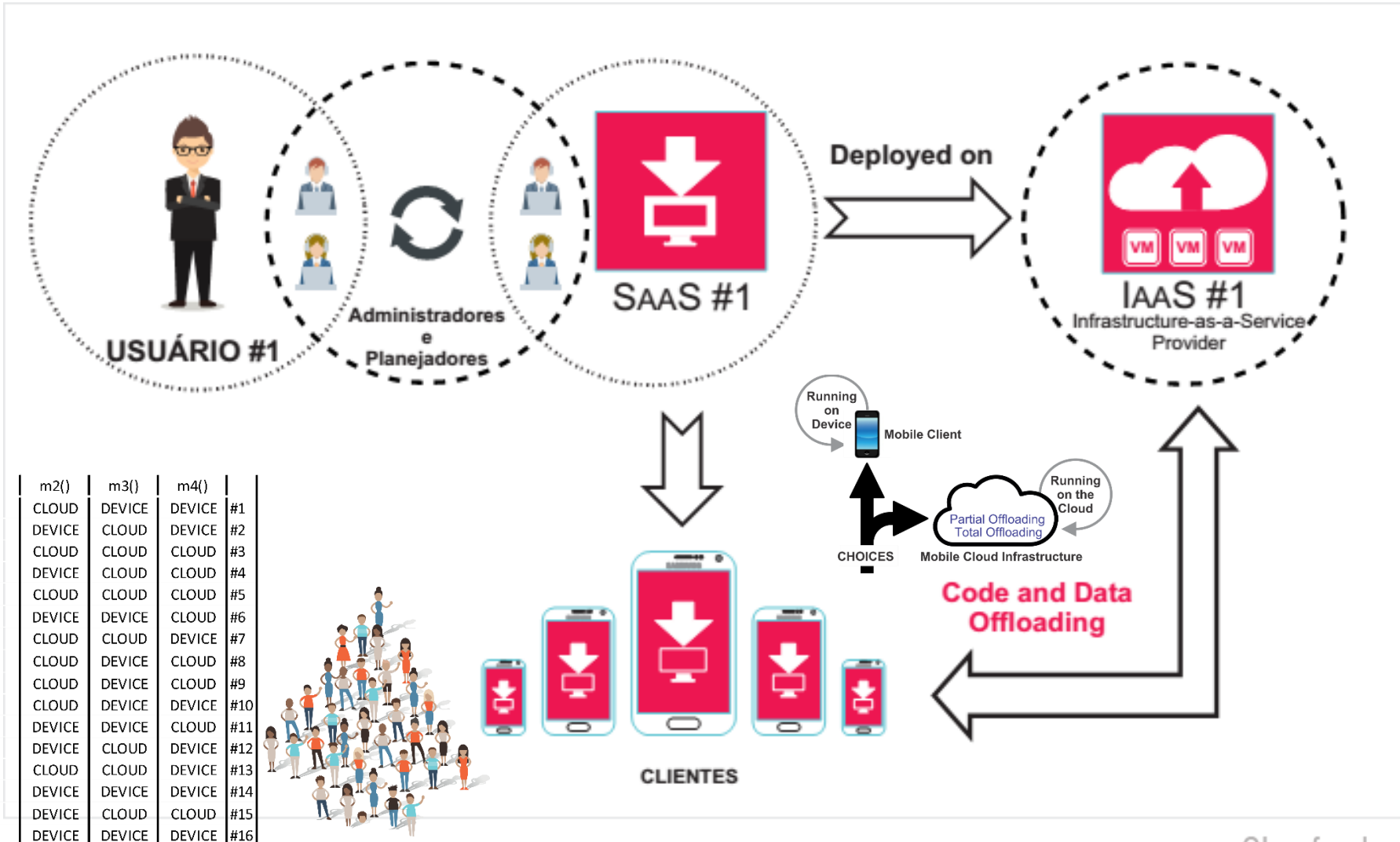
Times are displayed in UTC.

Note: datapoints are plotted at the start of the period.

Close



Motivação





Recursos Considerados

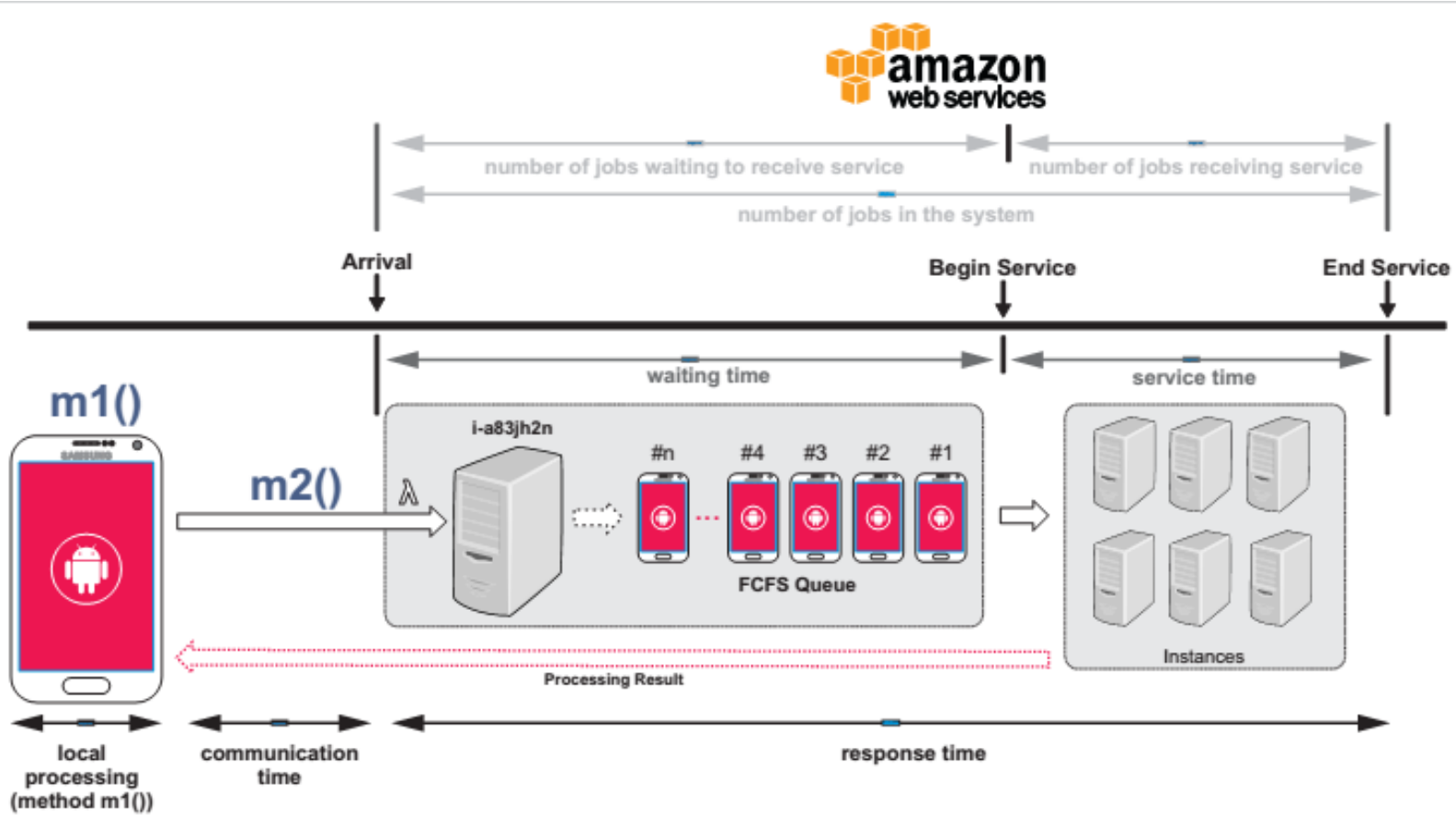
Uso geral – Geração atual

	vCPU	Mem (GiB)	On Demand
t2.nano	1	0.5	\$0.0059 por hora
t2.micro	1	1	\$0.012 por hora
t2.small	1	2	\$0.023 por hora
t2.medium	2	4	\$0.047 por hora
t2.large	2	8	\$0.094 por hora
t2.xlarge	4	16	\$0.188 por hora
t2.2xlarge	8	32	\$0.376 por hora
c4.large	2	3.75	\$0.1 por hora
c4.xlarge	4	7.5	\$0.199 por hora
c4.2xlarge	8	15	\$0.398 por hora
c4.4xlarge	16	30	\$0.796 por hora
c4.8xlarge	36	60	\$1.591 por hora

Data Transfer OUT To Internet	Price/GB
First 10 TB / month	\$0.155
Next 40 TB / month	\$0.115
Next 100 TB / month	\$0.090
Next 350 TB / month	\$0.065



Problema



	m1()	m2()
#1	device	device
#2	device	cloud
#3	cloud	device
#4	cloud	cloud

Application's execution time ...

... needs to attend the SLA

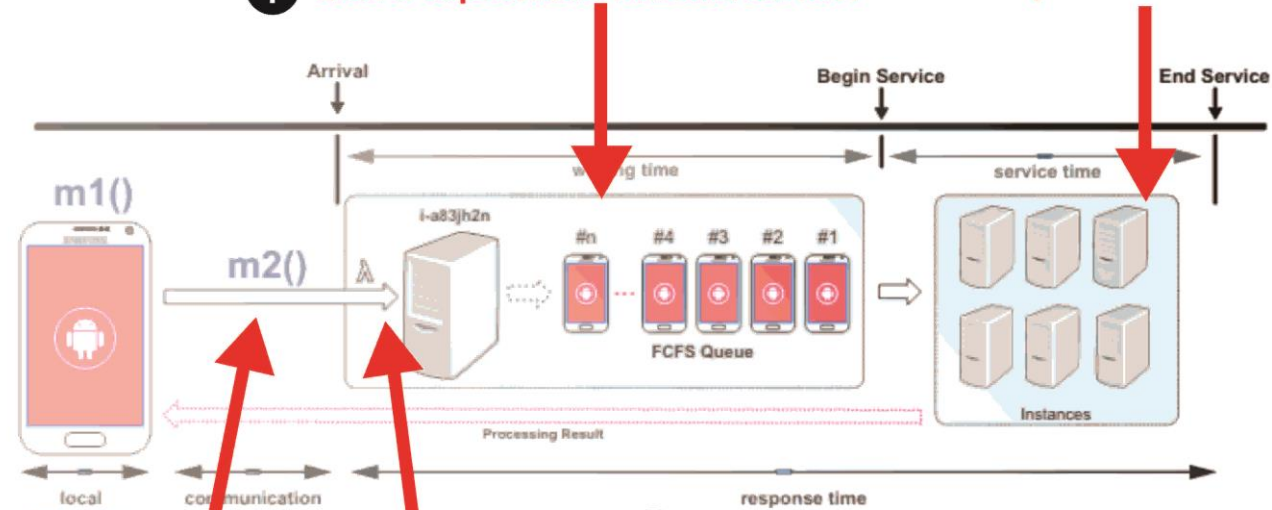
SLA

36,000 ms



Problema

- ? É aplicável uma fila?
- ? Qual a capacidade máxima da fila?
- ? Quantas instâncias usar?
- ? Quantos trabalhos alocar por instância?



? Quais as configurações que atenderão ao SLA?

	m1()	m2()
#1	device	device
#2	device	cloud
#3	cloud	device
#4	cloud	cloud

Application's execution time ...
... needs to attend the SLA

SLA
36,000 ms

? Considerando a demanda, qual a estratégia de offloading vai acarretar no menor consumo de recursos

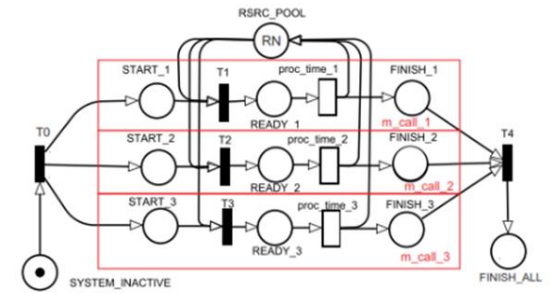
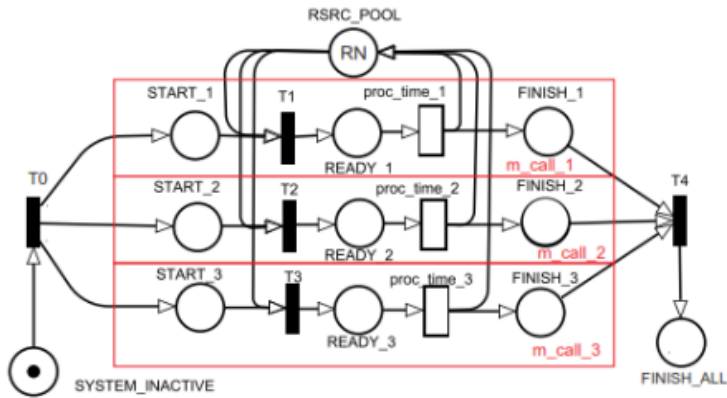
? Qual o volume de dados transferidos por cada estratégia de offloading?

Instance Type	vCPU	Memory (GiB)	Reserved	On-Demand (per hour)
t2.nano	1	0.5	\$36.84	\$0.0059
t2.micro	1	1	\$73.56	\$0.012
t2.small	1	2	\$147.12	\$0.023
t2.medium	2	4	\$294.36	\$0.047
t2.large	2	8	\$588.72	\$0.094
t2.xlarge	4	16	\$1177.32	\$0.188
t2.xlarge	8	32	\$2354.64	\$0.376

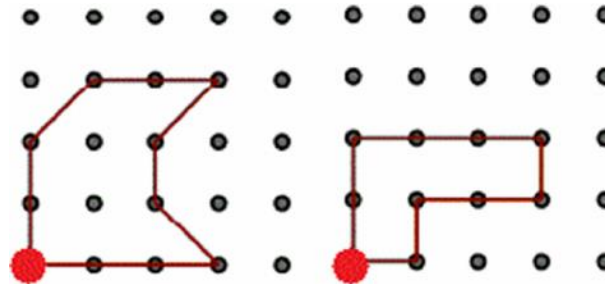
? Quais os tipos de instâncias utilizar?



Solução



GRASP



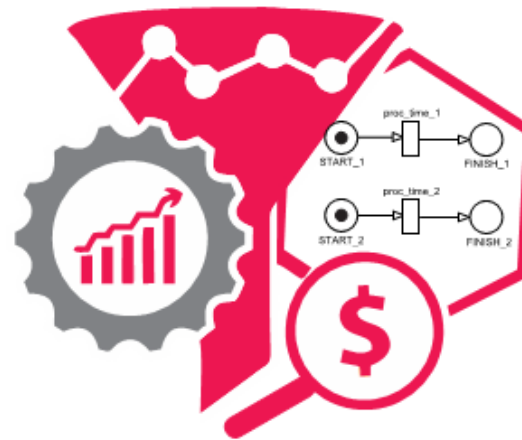


Solução

1. Cenários de offloading a serem considerados
2. Desempenho estabelecido em SLA
3. Número estimado de usuários do sistema
4. Taxas de chegadas
6. Tipos de instâncias
6. Um valor máximo à pagar
8. Período de uso do sistema (meses ou anos)
9. Tempo médio de comunicação
9. Bytes trafegados por requisição

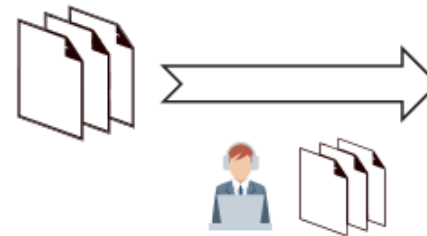
1

**Collecting
Input
Parameters**



2

**Generate Scenarios and
Find out an
Optimal Scenario
using an
Optimization Algorithm**



IaaS #1

3

**Application configuration
on IaaS**



Métricas

1. Mean Time To Execute (MTTE)
2. Mean Communication Time
3. Throughput
4. Mean Response Time



Modelos de Custos

A proposta vai considerar como custo:

1. O custo para transferir um volume de dados gerado pelo offloading para a nuvem pública; e
2. O custo do uso de VMs na nuvem pública em um determinado período.

- **Custo por Volume de Dados Trafegados:**

$$DataTransferCost = \sum_{ur=1}^n TransferredGBs_{ur} \times PricePerGB_{ur}$$

- **Custo por Uso de VMs:**

$$VmCost = \sum_{type=1}^n NumberOfVMs_{Type} \times PricePerHour_{type} \times UsageHours$$

- **Custo Total:**

$$TotalCost = DataTransferCost + VmCost$$



Próximos Passos

1. Finalizar o testbed;
2. Executar estudos de casos
3. Criar modelos;
4. Criar algoritmos de otimização



