



Proposta

Suporte para o Planejamento de Ambientes Elastic Cloud para Transcodificação de Vídeo.

Iure de Sousa Fé - (isf2@cin.ufpe.br)

Orientador : **Paulo Maciel** (prmm@cin.ufpe.br)

- Contextualização;
- Problema;
- Contribuição;
- Pesquisa;
- Modelo;
- Validação;
- Otimização;



CONTEXTUALIZAÇÃO - MOTIVAÇÃO

- A crescente expansão da utilização de vídeo na Internet, deixando de **ocupar 67% tráfego** total em 2015 para ocupar **82% 2020**;
- O compartilhamento de vídeo é um dos mais utilizados, só o youtube recebe **um terço dos usuários de internet**.
- W3C recomenda utilização dos **containers suportados** na maioria dos browsers e dispositivos (MP4,Ogg,WebM);
- Transcodificação é uma atividade computacionalmente exaustiva e na em muitas vezes possui carga variável;
- **Computação em nuvem** pode ser utilizada para fornecer a computação necessária para atender a demanda variável através da **elasticidade**;

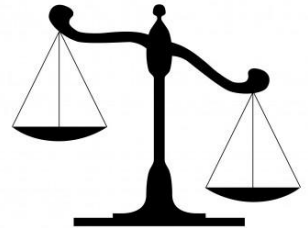
- Para uma nuvem, o tempo médio de resposta, vazão e custo (Aluguel ou Elétrico) varia de acordo com diversos parâmetros configuráveis pelo gerente do sistema.
- Identificar os valores desses parâmetros não é uma tarefa trivial, a escolha inadequada leva o sistema ao não cumprimento do SLA e/ou custo elevado.

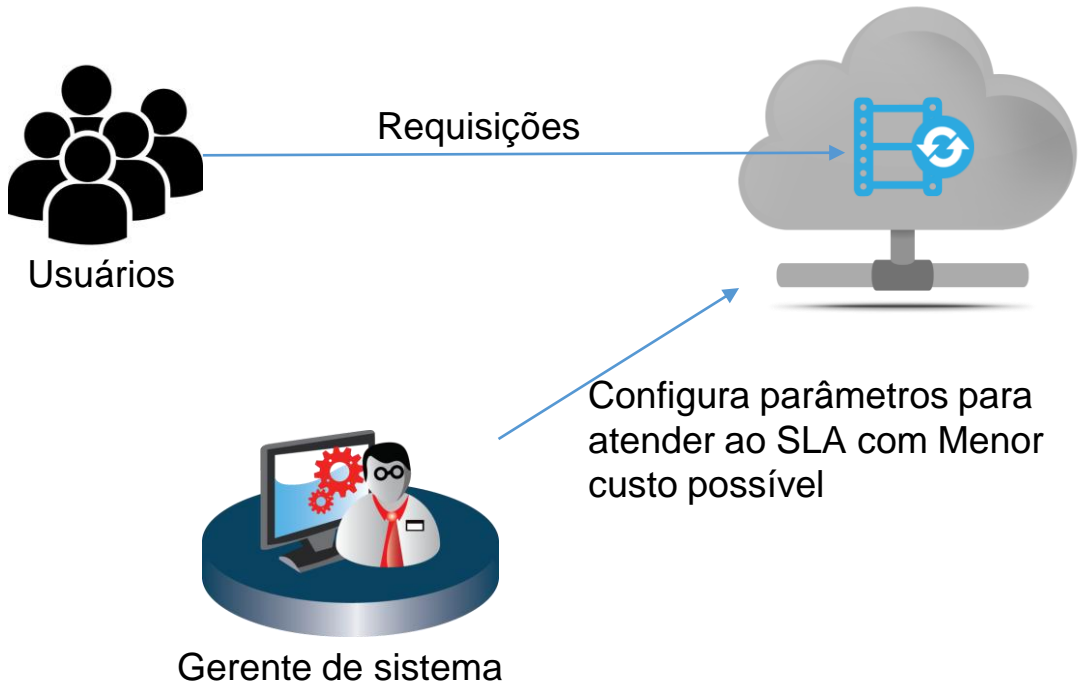




CONTRIBUIÇÃO ESPERADA

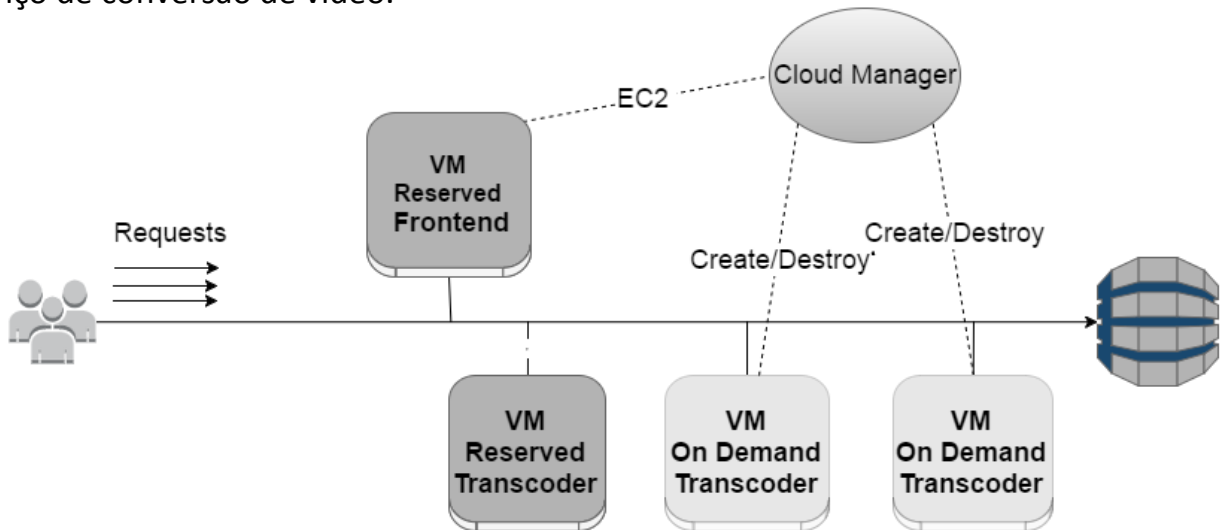
- Modelagem de performance e Avaliação de Custo em Ambientes Elastic Cloud Para Conversão de vídeo;
- Identificar através de algoritmo de otimização configurações que respeitem o SLA e minimizem o custo Elétrico.





A proposição de modelos de desempenho para serviço de conversão de vídeo com *autoscaling*. Com objetivo de dar suporte ao planejamento de infraestruturas de nuvens públicas e privadas.

Propor/Adaptar um mecanismo para definição de configuração apropriada para suportar o serviço de conversão de vídeo.

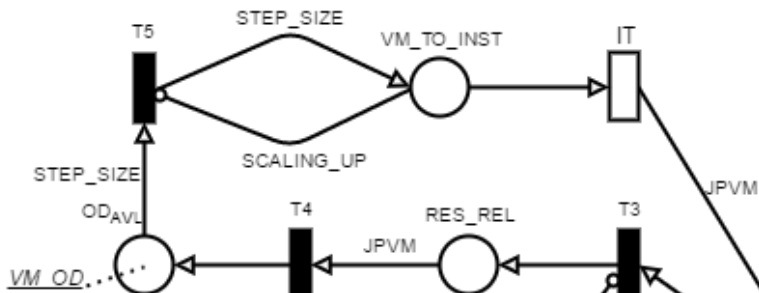


- Qual será a melhor configuração para manter uma vazão e tempo médio de resposta ao menor preço ?
 - Qual o **tipo** de VM utilizada ?
 - Quatas VMs por tipo de **contrato** ?
 - Qual o **threshold** de **instanciação** ?
 - Qual **threshold** de **destruição** ?
 - Qual o **step_size** ?
 - Quantas **conversões** por **VM**?
 - Quantos nós de computação ?
 - Quando ligar e desligar esses nós;

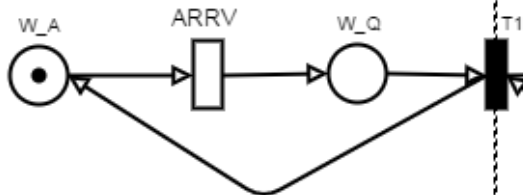


MODELO PROPOSTO – NUVEM PUBLICA

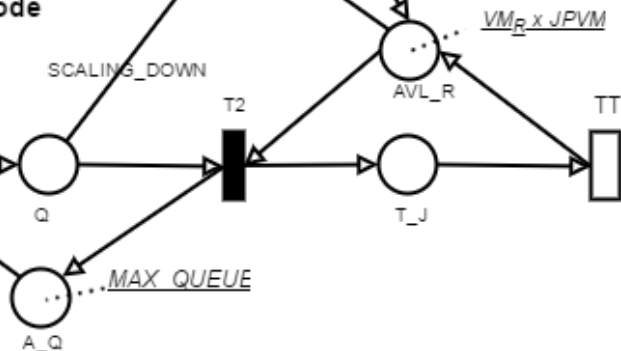
Auto-Scaling



Admission



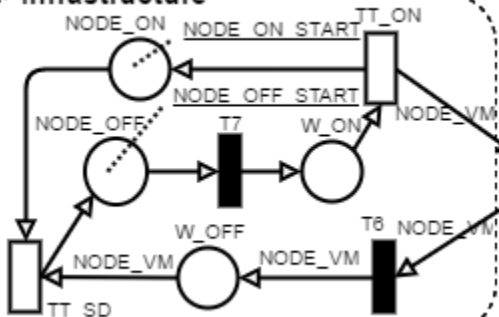
Transcode



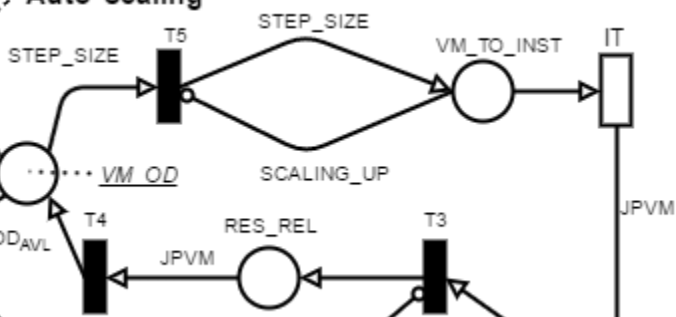


MODELO PROPOSTO – NUVEM PRIVADA

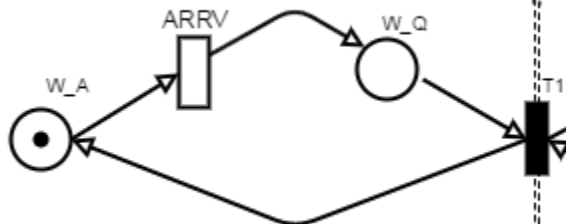
Infrastructure



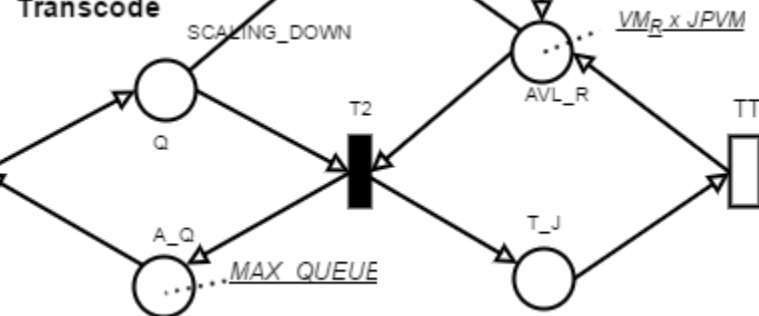
Auto-Scaling



Admission



Transcode





MODELO PROPOSTO

Scaling UP

```

IF(#Q >= ((#VM_TO_INST+
  ((#AVL_R + #T_J)/JPVM)) × THR_INST)):
  (#VM_TO_INST + 1)
ELSE
  (0)

```

Scaling Down

```

IF(VM_OD_INITIAL > #OD_AVAL):
  (((#ALV_R + #T_J + #RES_REL)/JPVM) - 1)
  × THR_DEST)
ELSE
  (0)

```

Transition	Type	Server Semantic	Weight	Priority	Enabling Function
ARRV	Timed	Single Server	-	-	
TT	Timed	Infinite Server	-	-	
IT	Timed	Infinite Server	-	-	
TT_ON	Timed	Infinite Server	-	-	
TT_SD	Timed	Infinite Server	-	-	
T1	Immediate	-	1	1	
T2	Immediate	-	1	1	
T3	Immediate	-	1	1	
T4	Immediate	-	1	1	
T5	Immediate	-	1	1	
T6	Immediate	-	1	1	$((\#OD_{avl}) + (\#W_OFF)) \geq \text{ThresholdShutdownNode}$
T7	Immediate	-	1	1	$(\#NODE_OFF < \text{node_off_start})$
					$((\#OD_{avl}) + (\#W_ON)) \leq \text{ThresholdStartNode}$



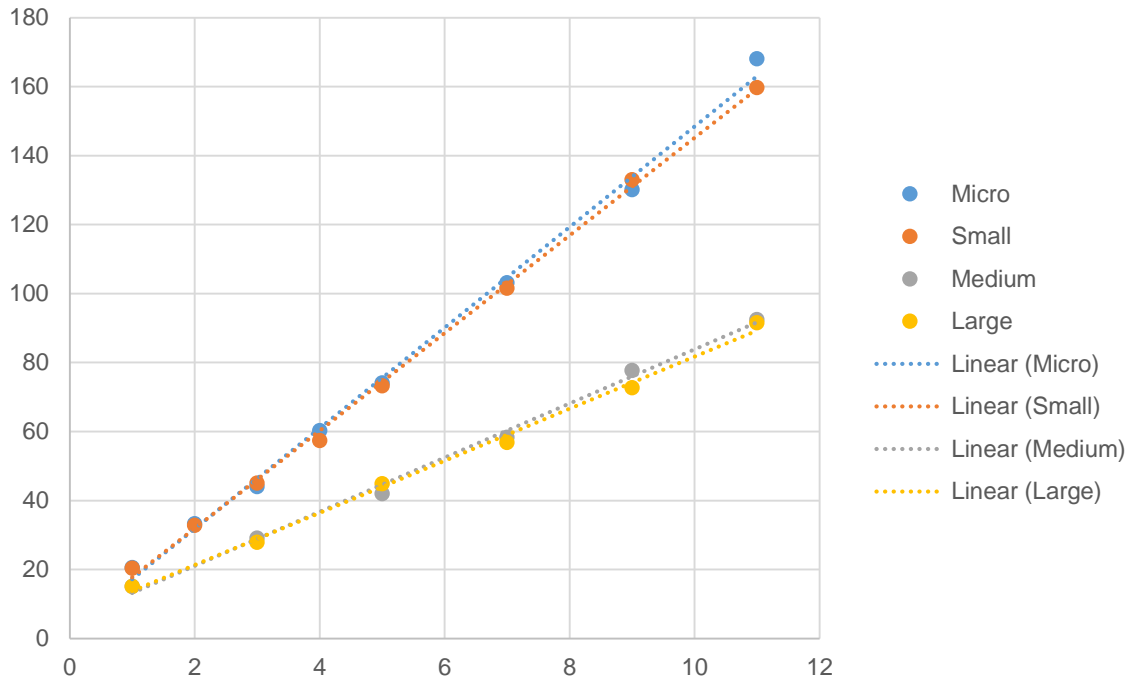
MODELO PROPOSTO

Parâmetro	Descrição
Size Infra	Quantos Nós de Computação
Node ON Start	Quantos Nós Ligados no início
Threshold Start Node	Limite Para Ligar Nó
Threshold Shutdown Node	Limite Para Desligar Nó
VM Type	Tipo de VM
Threshold Instantiate VM	Momento de Ligar uma nova VM
Threshold Destroy VM	Momento de Destruir uma VM
JPVM	Quantidade de Trabalhos por V<
VM on Start	Numero mínimo de VMs ligadas
stepsize	Quantas VMs devem ser ligadas
ArrivalTime	Tempo entre de chegadas

Parâmetro	Descrição
Vazão	(Trasnc/segundo)
Tempo médio de resposta	(segundos)
Custo	Custo Elétrico (Kwh)



NUVEM PRIVADA – TEMPO TRANSCODIFICAÇÃO





NUVEM PRIVADA – TEMPO INSTANCIAÇÃO

Tipo de VM	Tempo (s)
Micro	24,24
Small	27,13
Medium	32,50
Large	35,66



NUVEM PRIVADA – TEMPO NÓ

Ação	Tempo (s)
Ligar nó	100,38
Desligar Nó	50,30



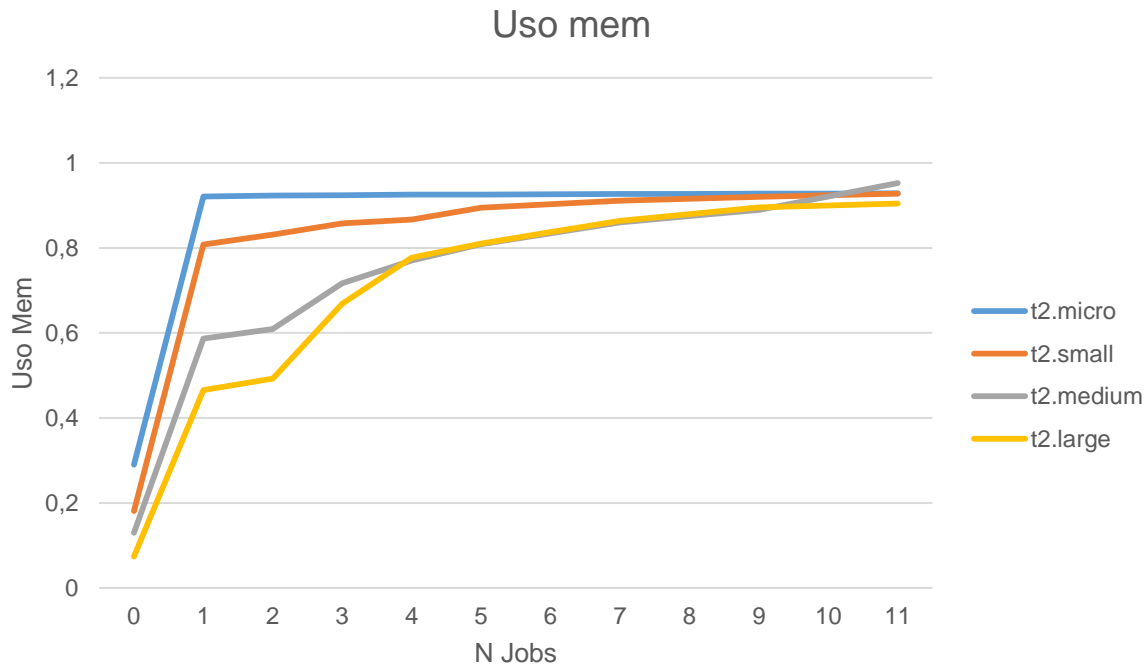
NUVEM PRIVADA

- O custo não está mais associado ao tipo de contrato, ou valor por hora e sim ao consumo elétrico;
- Existem diversos modelos matemáticos para previsão de custo elétrico de VMs;

<ul style="list-style-type: none">• Energy-efficient Virtual Machine Provision Algorithms for Cloud Systems	$P = \left[\frac{c}{C} (1 - \alpha) + \alpha \right] P_p$
<ul style="list-style-type: none">• VMeter: Power Modelling for Virtualized Clouds	$P_{(total)} = P_{\{baseline\}} + \sum_{k=1}^N P_{\{domain(k)\}}$ $P_{\{domain(i)\}} =$ $\alpha(a_2 p_{CPU(i)} + a_3 p_{cache(i)}) + \beta(a_5 p_{DRAM(i)} + a_6 p_{disk(i)})$
<ul style="list-style-type: none">• An Online Power Metering Model for Cloud Environment;• Prediction Model for Virtual Machine Power Consumption in Cloud Environments;	$P_{server} = \alpha \cdot \sum_{k=1}^n U_{CPU}(k) + \beta \cdot \sum_{k=1}^n U_{MEM}(k)$ $+ \gamma \cdot \sum_{k=1}^n U_{IO}(k) + n \cdot e + E_{baseline}$
<ul style="list-style-type: none">• An intelligent power consumption model for virtual machines• under CPU-intensive workload in cloud environment	$p_{vm}(u_p) = \alpha \cdot u_p^\beta,$

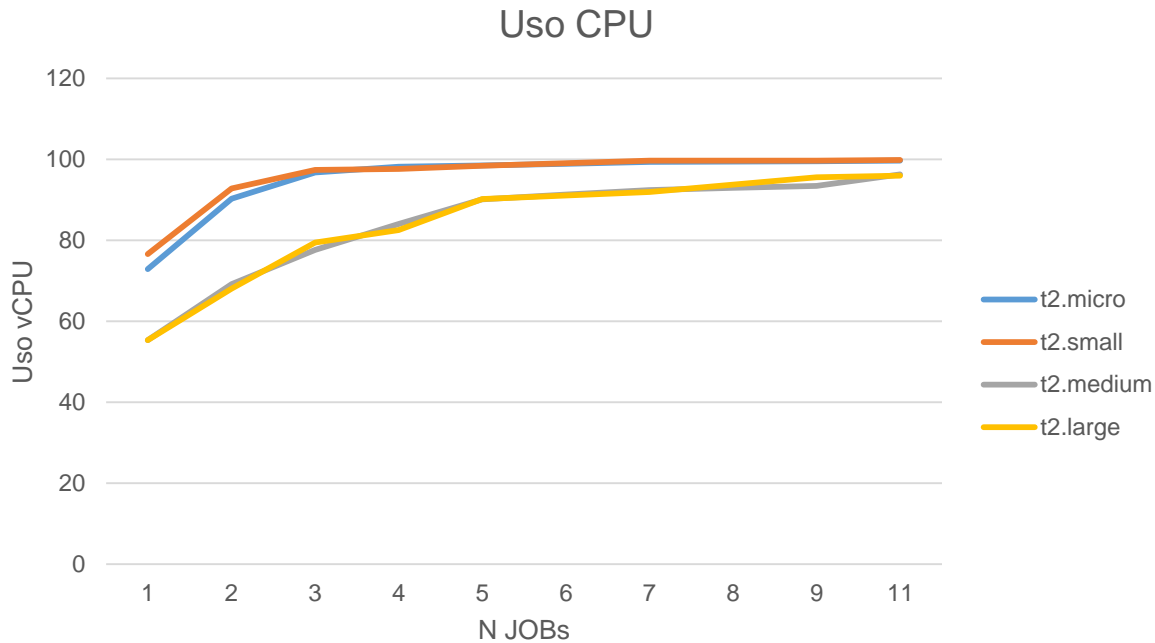


NUVEM PRIVADA – USO MEM





NUVEM PRIVADA – Uso CPU





NUVEM PRIVADA – Uso CPU

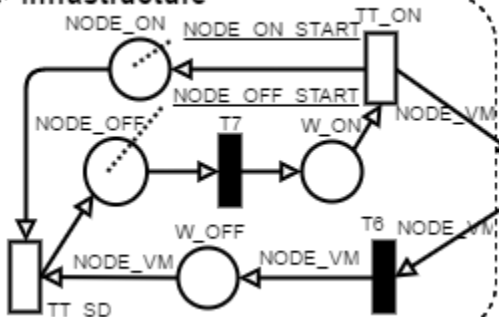
$$P_{server} = \alpha \cdot \sum_{k=1}^n U_{CPU}(k) + \beta \cdot \sum_{k=1}^n U_{MEM}(k) + \gamma \cdot \sum_{k=1}^n U_{IO}(k) + n \cdot e + E_{baseline}$$

	e	α	β	Ebaseline
Micro	<u>1,429</u>	<u>20,21</u>	<u>1,259</u>	<u>37,7</u>
Small	1,241	20,012	0,649	37,7
Medium	3,379	32,245	0,85	37,7
Large	2,56	33,298	2,573	37,7

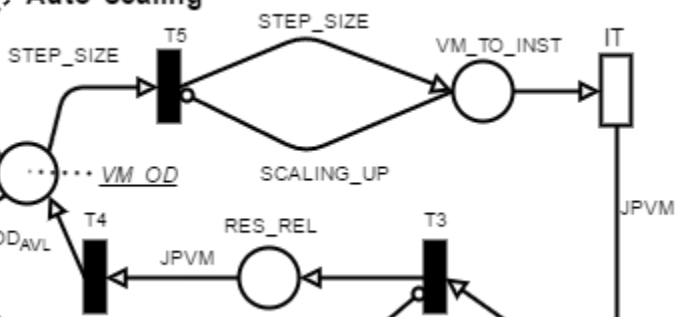


MODELO PROPOSTO

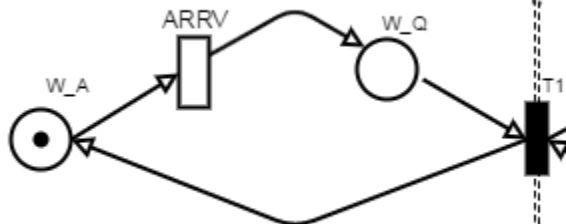
Infrastructure



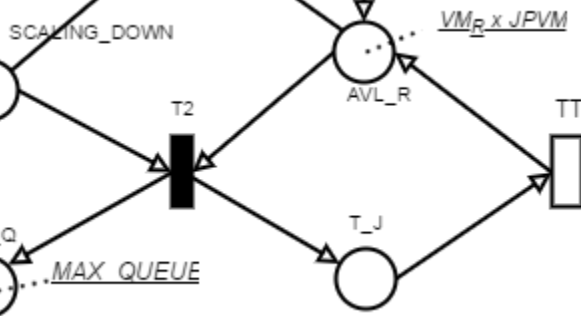
Auto-Scaling



Admission



Transcode





NUVEM PRIVADA

$$Vm_P_use = \sum_{i=1}^n P\left(\frac{\#P2 + \#P3}{W} = i\right) \times \left(\sum_{k=0}^{i \times W} P(\#P2 = k) \times \left(\sum_{t=k}^{k+i} Fe\left(\left[\frac{t}{i}\right]\right) \right) \right)$$

$$(P\{((\#P2+\#P3)/2)=1\})^*(((P\{\#P2=0\})^*(0))+((P\{\#P2=1\})^*(1))+((P\{\#P2=2\})^*(2))) +$$

$$((P\{((\#P2+\#P3)/2)=2\})^*(((P\{\#P2=0\})^*((0)+(0))) + ((P\{\#P2=1\})^*((0)+(1))) + ((P\{\#P2=2\})^*((1)+(1))) + ((P\{\#P2=3\})^*((1)+(2))) + ((P\{\#P2=4\})^*((2)+(2)))))) + \dots$$

$$((P\{((\#P2+\#P3)/2)=3\})^*(((P\{\#P2=0\})^*((0)+(0)+(0))) + ((P\{\#P2=1\})^*((1)+(0)+(0))) + ((P\{\#P2=2\})^*((1)+(1)+(0))) + ((P\{\#P2=3\})^*((1)+(1)+(1))) + ((P\{\#P2=4\})^*((2)+(1)+(1)))))) + \dots$$



NUVEM PRIVADA – CONSUMO TRANSIÇÕES

Tipo de VM	Consumo (wh)
Micro	0,126306
Small	0,151381
Medium	0,202545
Large	0,229436

Ação	Consumo (wh)
Ligar nó	1,21
Desligar Nó	0,567199



NUVEM PRIVADA – CUSTO ELÉTRICO

$$\text{Custo } E = (VM_{E_{USE}} + TP_{inst} * CE_{inst} + TP_{PowerONNode} * CE_{PowerONNode} + TP_{PowerOFFNode} * CE_{PowerOFFNode}) * \text{Tempo}$$

Taxa entre Chegadas (exp)	Medida	média	Real	Modelo (Análise estacionária)
40 s	Tempo para converter 30 vídeos (s)	1220,7	(1154,1; 1287,3)	1200
	Potencia (W)	86,679	(86,003; 87,355)	86,1475132

Obs: Não há evidências para refutar a normalidade dos resultados do experimento, com 95% de confiança.



GRASP - OTIMIZAÇÃO

Parâmetro/SLA	15 s	30 s	45 s	60 s
Size Infra	3	3	3	3
Node ON Start	2	1	1	1
Threshold Start Node	2	0	0	0
Threshold Shutdown Node	8	6	4	6
VM Type	t2.medium	t2.medium	t2.medium	t2.micro
Threshold Instantiate VM	5	2	7	4
Threshold Destroy VM	1	7	3	1
JPVM	1	1	2	1
VM on Start	3	2	2	1
stepsize	1	3	1	1
ArrivalTime (s)	10	10	10	10

Métrica/SLA	15 s	30 s	45 s	60 s
Vazão (transcod/seg)	0.09999	0.0999	0.0999	0.0999
Tempo médio de resposta(s)	14.6327	23.458	43.983	48.032
Custo (Kwh)	963.59880	602.622	557.903	507.237



GRASP - OTIMIZAÇÃO

