

# Otimização de custo de um sistema de modelagem atmosférica em nuvens computacionais

*Instituto de Computação  
Universidade Federal Fluminense*

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Mateus S. de Melo  
Lúcia M. A. Drummond  
Roberto P. Souto  
06 de Novembro de 2025

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Roteiro

## Introdução

## Conceitos Preliminares

## Reescalonamento em Spot

## Reescalonamento em On-Demand

## Análise de arquitetura

## Avaliação Experimental

## Conclusão e Trabalhos Futuros

### Introdução

### Conceitos Preliminares

### Reescalonamento em Spot

### Reescalonamento em On-Demand

### Análise de arquitetura

### Avaliação Experimental

### Conclusão e Trabalhos Futuros

## Introdução

### Conceitos Preliminares

### Reescalonamento em Spot

### Reescalonamento em On-Demand

### Análise de arquitetura

### Avaliação Experimental

### Conclusão e Trabalhos Futuros

- A utilização da computação em nuvem para a execução de aplicações de HPC tem sido amplamente discutida
- Devido a isso, vários provedores buscaram oferecer serviços voltados para computação de alto desempenho
- Principal desvantagem: custo elevado de sua utilização

- Propor uma abordagem para reduzir os elevados custos financeiros associados à utilização de *clusters* em uma plataforma de nuvem pública
- São apresentadas duas principais estratégias:
  - ① Utilização exclusivamente de um mercado menos custoso, porém sujeito a revogação de acesso (*Spot*).
  - ② Algoritmo de seleção de máquinas virtuais de menor custo aliado ao reescalonamento em um mercado que não está sujeito a revogação de acesso (*On-Demand*).
- Será utilizado o modelo BRAMS (*Brazilian developments on the Regional Atmospheric Modelling System*).

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

**Conceitos  
Preliminares**

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Supercomputador Santos Dumont



- O Santos Dumont (SDumont) é um supercomputador localizado no LNCC
- Capacidade de processamento de 20 Petaflop/s

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

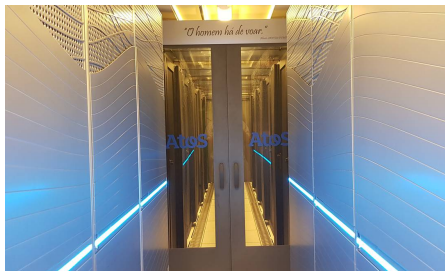
Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Supercomputador Santos Dumont



- Nós são conectados por uma rede *Infiniband*
- Dois sistemas de arquivos: Lustre e DDN Exascaler
  - Capacidade agregada de armazenamento da ordem de 4 PB

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros



# Supercomputador Santos Dumont

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

O SDumont foi escolhido como referência *On-Premises* nos testes devido à sua importância no cenário da pesquisa científica brasileira.

# Amazon Elastic Compute Cloud (EC2)



Amazon  
EC2

- Disponibiliza uma variedade de tipos de VMs (instâncias) de maneira escalável e sob demanda
- Instâncias com diversas configurações
- São categorizadas em famílias, cada uma possuindo características comuns específicas

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

## Famílias:

- **Uso Geral:** configurações equilibradas, adequada para uma variedade de aplicações
- **Otimizadas para Computação:** ideais para aplicações que realiza processamento intensivo
- **Otimizadas para Memória:** ideais para aplicações intensivas em memória
- **Otimizadas para Armazenamento:** adaptadas para cargas de trabalho que requerem alto acesso sequencial de leitura e escrita a grandes conjuntos de dados
- **Computação Acelerada:** equipadas com aceleradores de *hardware*, ou coprocessadores
- **Otimizadas para HPC:** suporta tarefas intensivas em computação, típicas de aplicações HPC

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Três principais mercados de instâncias:
  - **Reserved:** contratação por longos períodos, que podem variar entre 1 ou 3 anos
  - **On-Demand:** pagamento ocorre com base no uso, tendo custo fixo por hora/segundo
  - **Spot:** custo variável por hora/segundo, geralmente são muito mais baratas e podem ser revogadas pelo provedor a qualquer momento

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros



- É uma ferramenta de código aberto que simplifica a criação e o gerenciamento de *clusters* de HPC na AWS
- Ferramenta gratuita, apenas os serviços AWS utilizados são cobrados

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Utilização de apenas instâncias *Spot* para diminuir o custo
- Procedimento desenvolvido utilizando **contador de reinicializações do Slurm**
- Procedimento responsável por monitorar **reinicializações de instâncias** devido a **falhas/revogações** e configuração de uma execução do BRAMS a partir de um **checkpoint**
- Provê ao usuário a utilização de um ambiente tolerante a falhas sem a necessidade de sua interação.

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Supercomputador Santos Dumont:
  - 2 processadores Intel Xeon Cascade Lake Gold 6252
  - 24 núcleos cada, totalizando 48 núcleos computacionais
  - 384GB de memória principal
- AWS:
  - Instância r5n.12xlarge (Head Node e Compute Nodes)
  - Intel Xeon Cascade Lake Platinum 8259
  - 48 vCPUs
  - 384GB de memória principal
  - Instâncias *On-Demand* x *Spot*

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros



Introdução

Conceitos  
Preliminares

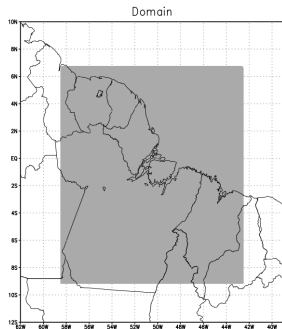
Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros



OADS/COLA

2024-08-21-06:12

- Resolução espacial de 10Km x 10Km
- Grade de 180 x 180 pontos
- 35 níveis verticais
- Tempo de previsão de 24h
- *Timesteps* de 30s (total de 2880 timesteps)

- Execução:
  - 1, 2 e 4 nós computacionais
  - 48 processos em cada nó
- Analisado:
  - Tempo total de execução sem checkpoint
  - Tempo total de execução com checkpoint a cada 1, 2 e 3 horas de previsão
  - Tempo total de execução com diferentes quantidades de revogação de instâncias Spot
  - Custo de utilização dos testes anteriores

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Execução na nuvem e no SDumont

## Introdução

## Conceitos Preliminares

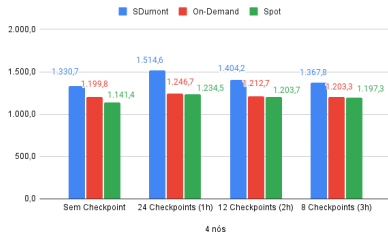
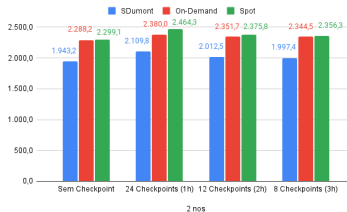
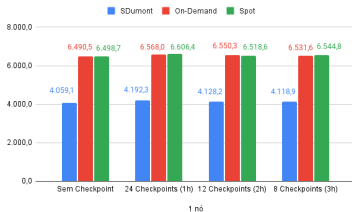
## Reescalonamento em Spot

## Reescalonamento em On-Demand

## Análise de arquitetura

## Avaliação Experimental

## Conclusão e Trabalhos Futuros



- Tempo de execução no SDumont com 1 nó sem *checkpoint* foi menor comparado às execuções na AWS.
- Pouca diferença entre os testes com diferentes quantidades de *checkpoints*

# Nós	Sem checkpoint		24 Checkpoints (1h)		12 Checkpoints (2h)		8 Checkpoints (3h)	
	<i>On-Demand</i>	<i>Spot</i>	<i>On-Demand</i>	<i>Spot</i>	<i>On-Demand</i>	<i>Spot</i>	<i>On-Demand</i>	<i>Spot</i>
1	\$12,98	\$9,69	\$13,13	\$9,85	\$13,10	\$9,77	\$13,06	\$9,76
2	\$6,85	\$4,54	\$7,12	\$4,87	\$7,04	\$4,70	\$7,02	\$4,65
4	\$5,97	\$3,36	\$6,21	\$3,63	\$6,04	\$3,59	\$5,99	\$3,54

- Foi cobrado \$3,576 por hora de uso das instâncias *On-Demand* e em média \$1,744 por hora de uso das instâncias *Spot*
- Diminuição de custo ao utilizar instâncias *Spot*.
- Baixo custo da sobrecarga de gravação de *checkpoint*

# Execução com simulação de revogações

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Testes com falhas usando instâncias *Spot* e feitos em 2 e 4 nós
- Falhas inseridas utilizando um simulador de interrupções seguindo a distribuição de Poisson
- *Checkpoint* sendo feito a cada 1 hora de previsão
- Testes com 3 execuções com diferentes quantidades de falhas

# Execução com simulação de revogações

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Execução	# Falhas			Tempo	Head Node	Custo	
	Nó 1	Nó 2	Total			Compute Nodes	Total
1°	1	2	3	4.012	\$3,99	\$3,89	\$7,92
2°	1	0	1	3.207	\$3,19	\$3,11	\$6,33
3°	4	1	5	6.493	\$6,45	\$6,29	\$12,82

Execução	# Falhas				Total	Tempo	Head Node	Custo	
	Nó 1	Nó 2	Nó 3	Nó 4				Compute Nodes	Total
1°	1	2	2	4	9	7.744	\$7,69	\$15,01	\$22,80
2°	0	1	1	0	2	3.069	\$3,05	\$5,95	\$9,03
3°	1	0	0	0	1	2.233	\$2,22	\$4,33	\$6,57

- Falhas aumentam consideravelmente o tempo de execução
- Pior cenário com 9 falhas utilizando 4 nós
  - Execução *On-Demand*: \$6,21

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Seleção do Tipo de Instância e Mercado

- Foi desenvolvido um Algoritmo para selecionar instâncias com a melhor relação custo-benefício
- Motivação:
  - Diferenças de preço e desempenho das instâncias
  - Variações de preço de instâncias *Spot* ao longo do dia
- O algoritmo retorna três sugestões de instâncias

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros



Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Para evitar múltiplas interrupções:

- Criação de um *cluster* com três filas:
  - *Spot*
  - *On-Demand*
  - Controladora

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Instância ( <i>Head Node</i> )	Tempo (s)	Custo <i>Head Node</i>	Custo <i>Compute Nodes</i>	Custos Instâncias
t2.micro	826,1	\$0,003	\$2,50	\$2,50
c6in.4xlarge	835,3	\$0,210	\$2,53	\$2,74
c6in.12xlarge	823,7	\$0,623	\$2,49	\$3,11

- Diferença de aproximadamente 10 segundos entre o maior e o menor tempo
- Diferença não é significativa
- Redução de até 20% no custo total com t2.micro
- Uso dessa configuração nos próximos testes

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Análise de *Compute Node*

## Introdução

### Conceitos Preliminares

### Reescalonamento em Spot

### Reescalonamento em On-Demand

### Análise de arquitetura

### Avaliação Experimental

### Conclusão e Trabalhos Futuros

Família de instâncias	Arquitetura da CPU	Fabricante da CPU	Tipo de instância
Otimizada para Computação	x86	Intel	c6i.12xlarge
			c6in.12xlarge
	ARM	AMD	c6a.12xlarge
		Graviton	c6g.12xlarge
			c6gn.12xlarge
			c7g.12xlarge
Otimizada para Memória	x86	Intel	r6i.12xlarge
			r5n.12xlarge
	ARM	AMD	r6a.12xlarge
		Graviton	r6g.12xlarge
			r7g.12xlarge
Otimizada para HPC	x86	AMD	hpc7a.12xlarge
			hpc7a.24xlarge
	ARM	Graviton	hpc7g.16xlarge

- Objetivo: identificar as instâncias que oferecem bom desempenho em relação ao seu custo financeiro

Arquitetura	Instância	# Nós	# Processos MPI	Preço Instância (USD\$/h)	Tempo (s)	Custo <i>Compute Node</i>	Custo Instância
x86	c6a.12xlarge	4	192	\$1,84	779,6	\$1,59	\$1,59
	c6i.12xlarge	4	192	\$2,04	835,8	\$1,89	\$1,90
	c6in.12xlarge	4	192	\$2,72	826,1	\$2,50	\$2,50
	r6i.12xlarge	4	192	\$3,02	854,6	\$2,87	\$2,87
	r6a.12xlarge	4	192	\$2,72	822,8	\$2,49	\$2,49
	r5n.12xlarge	4	192	\$3,58	1199,8	\$4,77	\$4,77
	hpc7a.12xlarge	4	96	\$7,20	480,8	\$3,85	\$3,85
	hpc7a.24xlarge	2	96	\$7,20	485,5	\$1,94	\$1,94
ARM	c6g.12xlarge	4	192	\$1,63	684,9	\$1,24	\$1,24
	c6gn.12xlarge	4	192	\$2,07	683,4	\$1,57	\$1,58
	<b>c7g.12xlarge</b>	<b>4</b>	<b>192</b>	<b>\$1,74</b>	<b>436,8</b>	<b>\$0,84</b>	<b>\$0,85</b>
	r6g.12xlarge	4	192	\$2,42	678,7	\$1,82	\$1,83
	r7g.12xlarge	4	192	\$2,57	436,3	\$1,25	\$1,25
	<b>hpc7g.16xlarge</b>	<b>3</b>	<b>192</b>	<b>\$1,68</b>	<b>512,5</b>	<b>\$0,72</b>	<b>\$0,72</b>

- Instâncias Graviton apresentam os menores tempos e custos de execução
- Custo de uso de instâncias ARM ligeiramente inferior a USD\$1,00 (*c7g.12xlarge* e *hpc7g.16xlarge*)

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Avaliar o BRAMS integrado à abordagem proposta de seleção e reescalonamento de instâncias
- Conjunto de instâncias dos testes anteriores
- Regiões: us-east-1, us-east-2, us-west-1, us-west-2 e sa-east-1
- Dois conjuntos de instâncias: x86 e x86 com ARM
- Domínio menor x domínio maior

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Testes sem Revogação

Arquitetura	# Exec	Instância	Preço Spot	Preço On-Demand	Tempo	Custo On-Demand	Custo Spot
x86	1	c6a.12xlarge (us-east-2)	\$0,65	\$1,84	893,3	\$1,59	\$0,64
	2	c6a.12xlarge (us-east-2)	\$0,70	\$1,84	885,1	\$1,59	\$0,68
	3	c6a.12xlarge (us-east-2)	\$0,66	\$1,84	916,9	\$1,59	\$0,67
	4	c6a.12xlarge (us-east-2)	\$0,66	\$1,84	878,5	\$1,59	\$0,64
	5	c6a.12xlarge (us-east-2)	\$0,66	\$1,84	885,8	\$1,59	\$0,65
	6	c6a.12xlarge (us-east-1)	\$0,83	\$1,84	861,3	\$1,59	\$0,80
	7	c6i.12xlarge (us-east-1)	\$0,91	\$2,04	965,3	\$1,89	\$0,97
	8	c6a.12xlarge (us-east-1)	\$0,92	\$1,84	847,5	\$1,59	\$0,86
	9	c6a.12xlarge (us-east-1)	\$0,78	\$1,84	893,4	\$1,59	\$0,78
	10	c6a.12xlarge (us-east-1)	\$0,93	\$1,84	852,2	\$1,59	\$0,88
x86 e ARM	1	c7g.12xlarge (us-east-2)	\$0,47	\$1,73	490,4	\$0,84	\$0,26
	2	c6gn.12xlarge (us-west-2)	\$0,21	\$2,07	740,4	\$1,57	\$0,17
	3	c6g.12xlarge (sa-east-1)	\$0,30	\$2,51	766,3	\$1,91	\$0,25
	4	c6g.12xlarge (sa-east-1)	\$0,30	\$2,51	754,3	\$1,91	\$0,25
	5	c7g.12xlarge (us-west-2)	\$0,27	\$1,74	492,4	\$0,84	\$0,15
	6	c7g.12xlarge (us-west-2)	\$0,27	\$1,74	496,0	\$0,84	\$0,15
	7	c7g.12xlarge (us-west-2)	\$0,27	\$1,74	489,5	\$0,84	\$0,15
	8	c7g.12xlarge (us-west-2)	\$0,27	\$1,74	495,1	\$0,84	\$0,15
	9	c6gn.12xlarge (us-west-2)	\$0,21	\$2,07	741,4	\$1,57	\$0,18
	10	c6gn.12xlarge (us-west-2)	\$0,21	\$2,07	741,4	\$1,57	\$0,18

- Todas as instâncias escolhidas no mercado *Spot*
- Segunda sugestão do algoritmo a partir da sexta execução com arquitetura x86
- Redução de até 89% do custo no melhor caso (instância c6gn.12xlarge na região us-west-2)

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros



Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Execuções com interrupções no início, meio e fim
  - Início: antes da conclusão do primeiro *timestep* até 25% do total de *timesteps*
  - Meio: de 25% a 75% do total de *timesteps*
  - Fim: de 75% do total de *timesteps* até o último *timestep*

# Testes simulando revogação

- Arquivos relacionados ao cálculo da radiação estavam sendo recriados no reescalonamento
- Recalculado o tempo de execução excluindo a o tempo de recriação dos arquivos

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Testes simulando revogação

Tabela: Com recriação dos arquivos

# Exec	Instância	Tempo Spot	Tempo OD	Tempo	Tempo Somente OD	Custo Spot	Custo OD	Custo Instâncias	Custo Somente OD
1	c7g.12xlarge (sa-east-1)	348,0	446,0	794,0	436,8	\$0,10	\$1,32	\$1,43	\$1,30
2	c7g.12xlarge (sa-east-1)	410,0	390,0	800,0	436,8	\$0,13	\$1,15	\$1,29	\$1,30
3	c7g.12xlarge (sa-east-1)	471,0	321,0	792,0	436,8	\$0,14	\$0,95	\$1,10	\$1,30

Tabela: Sem recriação dos arquivos

# Exec	Instância	Tempo Pré-proc	Tempo Spot	Tempo OD	Tempo	Custo Spot	Custo OD	Custo Instâncias	Custo Somente OD
1	c7g.12xlarge (sa-east-1)	271,8	348,0	174,2	522,2	\$0,11	\$0,52	\$0,63	\$1,30
2	c7g.12xlarge (sa-east-1)	270,3	410,0	119,7	529,7	\$0,13	\$0,35	\$0,49	\$1,30
3	c7g.12xlarge (sa-east-1)	270,3	471,0	50,7	521,7	\$0,14	\$0,15	\$0,30	\$1,30

- Redução significativa no tempo e no custo total das execuções
- Comparativo de custos da estratégia:
  - Redução de até 72,78% no melhor cenário
- Comparativo custo somente On-Demand:
  - Redução de até 76,92% no melhor cenário

Introdução

Conceitos Preliminares

Reescalonamento em Spot

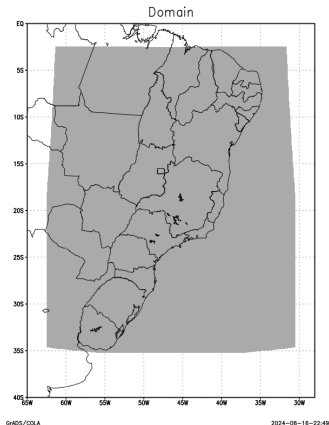
Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

# Testes com um domínio espacial maior



- Avaliar a estratégia proposta utilizando um domínio espacial maior
- Verificar revogações por parte da AWS
- Resolução espacial de 5Km x 5Km
- Grade de 680 x 740 pontos
- 40 níveis verticais
- Tempo de previsão de 48h
- *Timesteps* de 10s (total de 17.280 *timesteps*)

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

# Testes com um domínio espacial maior na AWS

# Exec	Instância	Tempo (s)	Preço Spot	Preço On-Demand	Custo Spot	Custo Somente On-Demand	Custo Total	Custo Total Somente On-Demand
1	c6g.12xlarge (sa-east-1)	15.119,2	\$0,25	\$2,51	\$4,22	\$42,08	\$4,89	\$42,75
2	c6g.12xlarge (sa-east-1)	15.573,5	\$0,25	\$2,51	\$4,36	\$43,35	\$5,05	\$44,04
3	c7g.12xlarge (sa-east-1)	13.526,1	\$0,28	\$2,67	\$4,18	\$39,88	\$4,60	\$40,30
4	c7g.12xlarge (sa-east-1)	11.691,2	\$0,27	\$2,67	\$3,58	\$34,45	\$4,10	\$34,96
5	c7g.12xlarge (sa-east-1)	13.124,5	\$0,28	\$2,67	\$4,12	\$38,69	\$4,70	\$39,27
6	c7g.12xlarge (sa-east-1)	13.722,3	\$0,27	\$2,67	\$4,11	\$40,46	\$4,71	\$41,07
7	c6gn.12xlarge (us-west-2)	14.895,9	\$0,55	\$2,07	\$9,14	\$34,18	\$9,57	\$34,62
8	c6gn.12xlarge (us-west-2)	14.887,7	\$0,23	\$2,07	\$3,77	\$34,16	\$4,20	\$34,59
9	c7g.12xlarge (us-east-2)	12.926,4	\$0,61	\$1,73	\$8,83	\$24,79	\$9,19	\$25,15
10	c7g.12xlarge (us-east-2)	13.102,7	\$0,61	\$1,73	\$8,88	\$25,13	\$9,25	\$25,50

- Primeira sugestão do algoritmo nas seis primeiras execuções
- Não houve revogações por parte da AWS
- Redução de custo de quase 90% em relação a execução em *On-Demand*

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

# Testes com um domínio espacial maior no SDumont

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Exec	Tempo Espera	Tempo Execução (s)
1	33h 44min 15s	18.603,1
2	84h 28min 39s	18.570,1
3	9h 2min 51s	18.877,4

- Tempo de execução no SDumont ligeiramente maior
- Tempo de espera na fila Slurm muito alto
- Tempo de *turnaround* do SDumont maior mesmo se houvesse revogação na nuvem

# Roteiro

Introdução

Conceitos Preliminares

Reescalonamento em Spot

Reescalonamento em On-Demand

Análise de arquitetura

Avaliação Experimental

Conclusão e Trabalhos Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Ao migrar as aplicações HPC para a nuvem, é necessário considerar desempenho e custo
- O uso de instâncias *Spot* auxilia na diminuição do custo



- Algoritmo de seleção de instâncias apresentou reduções de custo significativas (89% no melhor caso de teste)
- Instâncias ARM tende a ter custos mais baixos que x86
- Reescalamento em instâncias On-Demand se mostrou vantajoso com revogações no meio e no final

Introdução

Conceitos  
Preliminares

Reescalamento  
em Spot

Reescalamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Nuvem x *On-Premises*
  - Tempo menor na nuvem
  - Tempo de espera (*turnaround*) elevado em ambiente *On-Premises*
- O uso de um cluster na nuvem combinado com as estratégias apresentadas para executar o modelo BRAMS mostra-se promissor tanto em termos de tempo de *turnaround* quanto de custo financeiro.

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

- Monitorar as variações de preço das instâncias *Spot*
- Desenvolver uma estratégia para a reutilização dos arquivos relacionados ao cálculo da radiação
- Investigar mais a fundo as eventuais variações de desempenho obtidas, o que levou a *speedups* elevados

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

Os resultados deste trabalho foram divulgados por meio de uma publicação no **Simpósio em Sistemas Computacionais de Alto Desempenho** e um artigo no periódico *Concurrency and Computation: Practice and Experience*. Além disso, os resultados foram apresentados ao **Comitê Científico do Model for Ocean-land-Atmosphere prediction (MONAN)**, um programa institucional do INPE e Ministério da Ciência e Tecnologia (MCTI) para desenvolvimento do novo modelo comunitário brasileiro do sistema terrestre.

Introdução

Conceitos  
Preliminares

Reescalonamento  
em Spot

Reescalonamento  
em On-Demand

Análise de  
arquitetura

Avaliação  
Experimental

Conclusão e  
Trabalhos  
Futuros

# Agradecimentos



Os autores agradecem o Laboratório Nacional de Computação Científica (LNCC/MCTI, Brasil) por fornecer recursos de HPC do supercomputador SDumont, que contribuíram para os resultados da pesquisa relatados neste artigo. URL: <http://sdumont.lncc.br>. Este trabalho também foi desenvolvido utilizando recursos do Projeto Universal/CNPq número 404087/2021-3.

Obrigado!