

Revalidação de análise de performance (Mbips e iops) entre
lustre e sistemas de armazenamento compatíveis com s3¹
*Revalidation of performance analysis (Mbips and Iops)
between Lustre and S3-compatible storage systems*

João Pedro Abreu de Souza²

Lúcia Maria de Assumpção Drummond³

Resumo

Utilizando dois benchmarks padrão(io-500 e md-workbench) para analisar a evolução da performance relativa entre sistemas de armazenamento típicos de hpc (lustre, um sistema de arquivos paralelo) e sistemas de armazenamento baseados nas APIs rest do s3 e na API da libs3, busquei verificar se houve melhora desde 2021, ano que foi publicado um artigo realizando uma comparação entre ambos os sistemas de armazenamento. A reavaliação se sustenta no fato do artigo ter 2 anos, e nesse tempo o s3 recebeu várias mudanças e manteve a mesma API, logo sendo possível avaliar se a distância entre ele e o lustre permanece a mesma. O lustre utilizado foi o provisionado pelo serviço fsx da aws e o s3 sendo o serviço da própria aws. O resultado foi ...

Palavras-chaves: HPC; Lustre; S3; AWS.

Abstract

Using two standard benchmarks(io-500 and md-workbench) to analyze the evolution of relative performance between typical hpc storage systems (lustre, a parallel file system) and storage systems based in the rest APIs of s3 and the libs3 API, I sought to check whether there was an improvement since 2021, the year in which an article was published a comparison between both storage systems. The reevaluation is based on the fact that the article is 2 years old, and in this Over time, s3 received several changes and maintained the same API, making it possible to assess whether the distance between it and the chandelier remains the same. The lustre used was the one provisioned by the aws fsx service and the s3 being the aws service itself. The result was ...

¹ Trabalho de conclusão de curso apresentado ao curso de Bacharelado em Ciência da Computação da Universidade Federal Fluminense como requisito parcial para conclusão do curso.

² Graduando do Curso de Ciência da Computação - UFF, jp_abreu@id.uff.br

³ Orientador - Instituto de Computação - UFF, lucia@ic.uff.br

Keywords: HPC; Lustre; S3; AWS.

Aprovado em: dd/mm/aaaa. **Versão Final em:** dd/mm/aaaa

1. INTRODUÇÃO

HPC é utilizado para computar uma grande gama de problemas, que mesmo uma única máquina poderosa não seria capaz em tempo hábil (i.e. Resolução de sistemas de equações parciais em uma grande malha para simulação). Normalmente as aplicações HPC são escritas em C ou Fortran, e rodam em super-computadores dedicados a essa função, como o mistral, super-computador alemão.

Desde a criação da AWS, em 2006, as plataformas de computação com modelo elastico, pagando pelo uso, vulgo nuvem, se tornaram a norma no desenvolvimento de aplicações, chegando até mesmo a areas altamente reguladas como o setor de pagamentos. Por conta disso, pesquisadores buscavam verificar se a nuvem poderia suportar os workflows de aplicações HPC. Nesse contexto, o artigo de 2021, "analyzing the performance of the S3 Object Storage API for HPC Workloads", foi tomado como base para realização desse trabalho, buscando avaliar as mudanças nos resultados que possam ter advindo nesses dois anos. Para isso, utilizei o io500, benchmark ior e mdtest, dado que o find, outro benchmark fornecido pelo io500, não é suportado para uso com o s3.

2. REPRODUÇÃO DA AVALIAÇÃO DE DESEMPENHO DO S3 PARA APLICAÇÕES HPC

2.1. EXECUÇÃO USANDO S3

O artigo utilizou o lustre e múltiplas implementações da API do S3

2.2. EXECUÇÃO USANDO LUSTRE

O artigo utilizou o lustre via FSX com as seguintes configurações :

2.3. ANALISE DOS RESULTADOS

... ..

3. EXPLORANDO PARAMETROS DO LUSTRE

Os parametros a serem avaliados serão

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875
[INVALID]

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 1 – IO500 Benchmark Results

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 2 – IO500 Benchmark Results

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 3 – IO500 Benchmark Results

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 4 – IO500 Benchmark Results

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

4. RESULTADOS EXPERIMENTAIS E ANALISE

Comparando o quociente entre a leitura do lustre nos melhores parametros

4.1. LUSTRE - BASELINE

Como visto na figura 2., a arquitetura utilizada contém 1 sub-redes, pública, com as maquinas ec2 nestas acessando o lustre provisionado via o serviço fsx da

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 5 – IO500 Benchmark Results

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 6 – IO500 Benchmark Results

própria AWS. Isso garante que o teste não será influenciado por falhas no provisionamento do cluster. O teste foi executado com script runTcc.sh, com seu código no repositório indicado . Os resultados foram obtidos variando-se 3 parâmetros : número de instâncias t3.large executando o io500, tamanho do arquivo utilizado na execução dos testes

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

4.2. S3

Como visto na figura 3., a arquitetura utilizada contém 2 sub-redes, uma privada e uma pública, com as máquinas da subnet privada acessando o bucket s3 provisionado via o serviço s3 da própria AWS. Essa arquitetura foi pensada para

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 7 – IO500 Benchmark Results

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 8 – IO500 Benchmark Results

manter a mesma estrutura do teste com lustre, de forma que possamos comparar apenas a mudança de armazenamento. O teste foi executado com script runTcc.sh, com seu código em anexo. Os resultados foram

Score: Bandwidth 0.000931 GiB/s : IOPS 0.003778 kiops : TOTAL 0.001875 [INVALID]

5. CONCLUSÃO E TRABALHOS FUTUROS

Com esse trabalho pudemos verificar

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 9 – IO500 Benchmark Results

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 10 – IO500 Benchmark Results

Referências

Test	Throughput	IOPS	Time (s)	Unit
ior-easy-write	0.002189	-	303.523	GiB/s
mdtest-easy-write	-	0.005847	302.666	kIOPS
timestamp	-	0.000000	0.001	kIOPS
ior-hard-write	0.000130	-	301.501	GiB/s
mdtest-hard-write	-	0.001975	302.595	kIOPS
find	-	0.000000	0.000 [INVALID]	kIOPS
ior-easy-read	0.010056	-	66.307	GiB/s
mdtest-easy-stat	-	0.006040	292.609	kIOPS
ior-hard-read	0.000262	-	150.044	GiB/s
mdtest-hard-stat	-	0.006040	100.022	kIOPS
mdtest-easy-delete	-	0.002928	601.830	kIOPS
mdtest-hard-read	-	0.003012	198.905	kIOPS
mdtest-hard-delete	-	0.002956	203.265	kIOPS

Tabela 11 – IO500 Benchmark Results

APÊNDICE A - Lista de ilustrações

Figura 1 – Arquitetura do cluster usando lustre Fonte: fornecido pelo autor.	13
Figura 2 – Arquitetura do cluster usando lustre Fonte: fornecido pelo autor.	14
Figura 3 – Arquitetura do cluster usando s3 Fonte: fornecido pelo autor.	15

APÊNDICE B - Lista de tabelas

Tabela 1 – IO500 Benchmark Results	3
Tabela 2 – IO500 Benchmark Results	3
Tabela 3 – IO500 Benchmark Results	4
Tabela 4 – IO500 Benchmark Results	4
Tabela 5 – IO500 Benchmark Results	5
Tabela 6 – IO500 Benchmark Results	5
Tabela 7 – IO500 Benchmark Results	6
Tabela 8 – IO500 Benchmark Results	6
Tabela 9 – IO500 Benchmark Results	7
Tabela 10 – IO500 Benchmark Results	7
Tabela 11 – IO500 Benchmark Results	8

APÊNDICE C

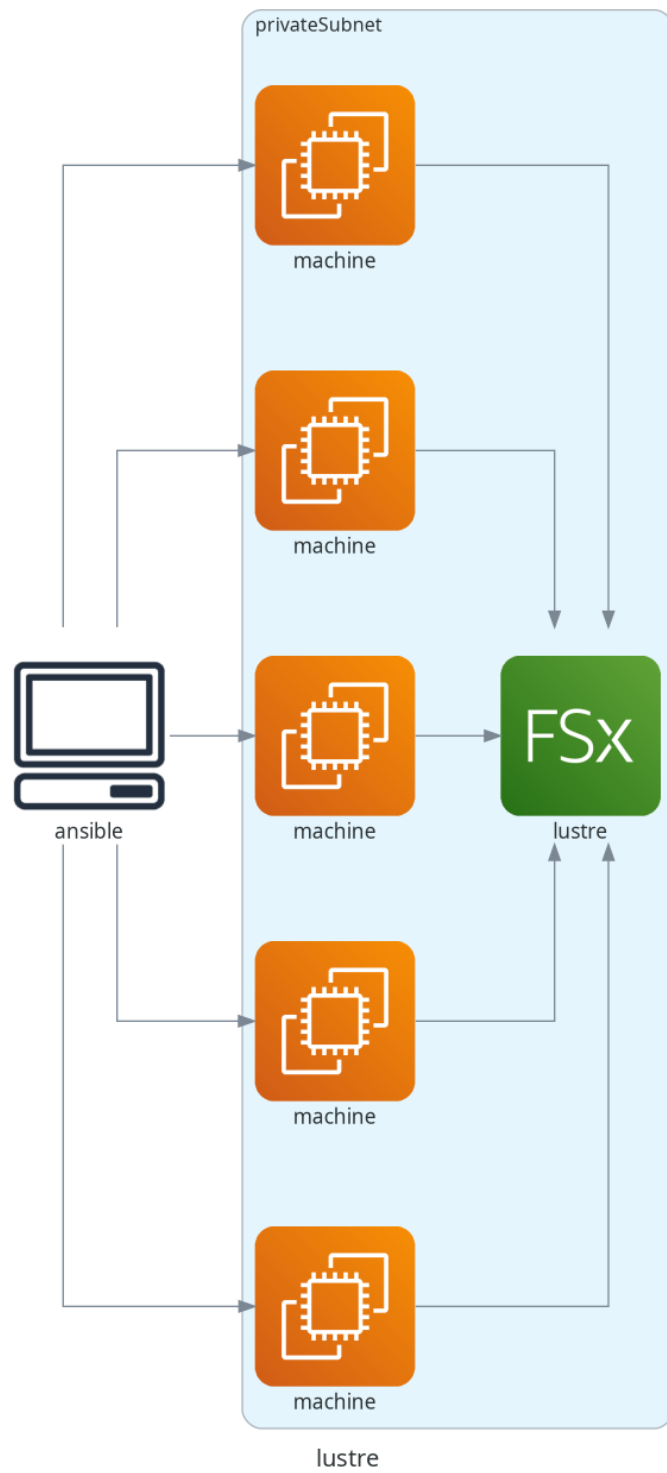
Os apêndices A e B devem conter a lista de ilustrações e tabelas, respectivamente. Os apêndices seguintes podem ser usados para apresentar textos elaborados pelo próprio autor, a fim de complementar a sua argumentação. Por exemplo, no caso de desenvolvimento de sistemas, podem constar nos apêndices os artefatos gerados da análise de sistemas, como protótipos de telas, casos de uso, diagramas UML, ou a lista de requisitos de negócio e de usuário.

ANEXO A

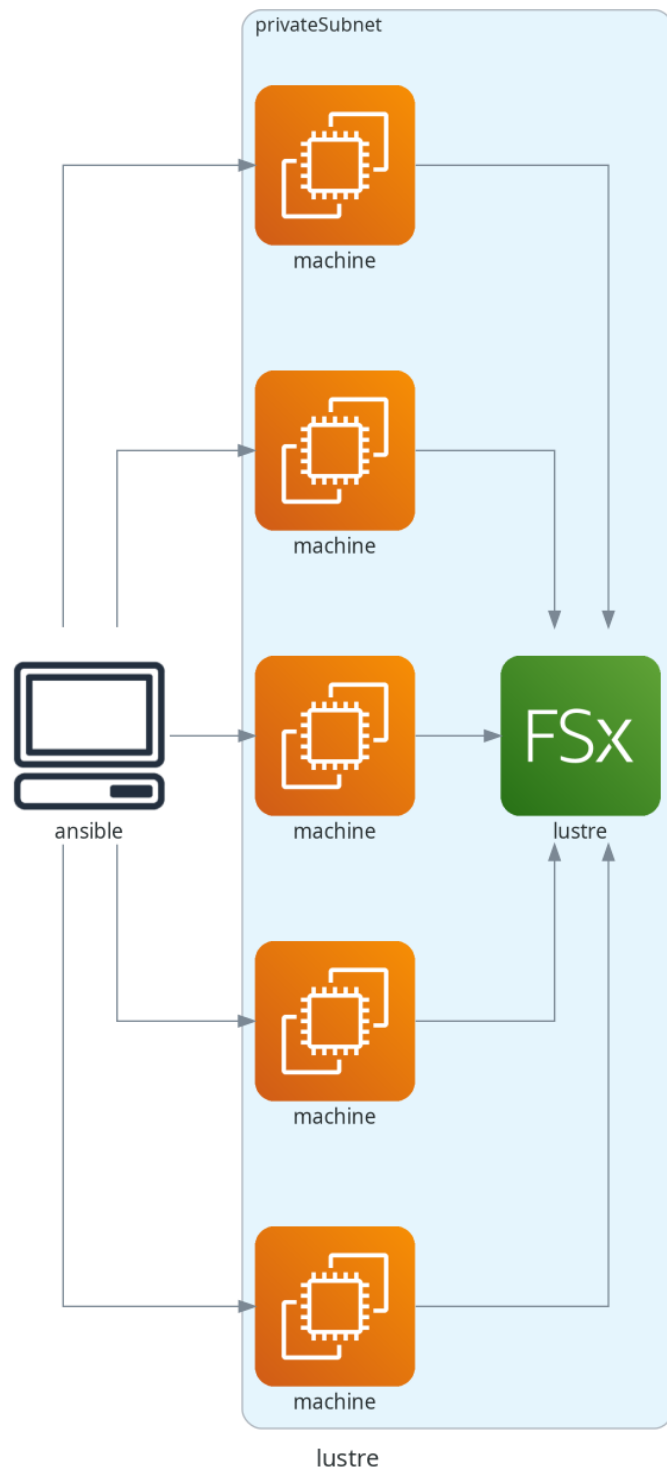
Seção opcional. Anexos são os documentos de autoria externa, ou não elaborados pelo autor, que servem de fundamentação, comprovação ou ilustração.

Agradecimentos

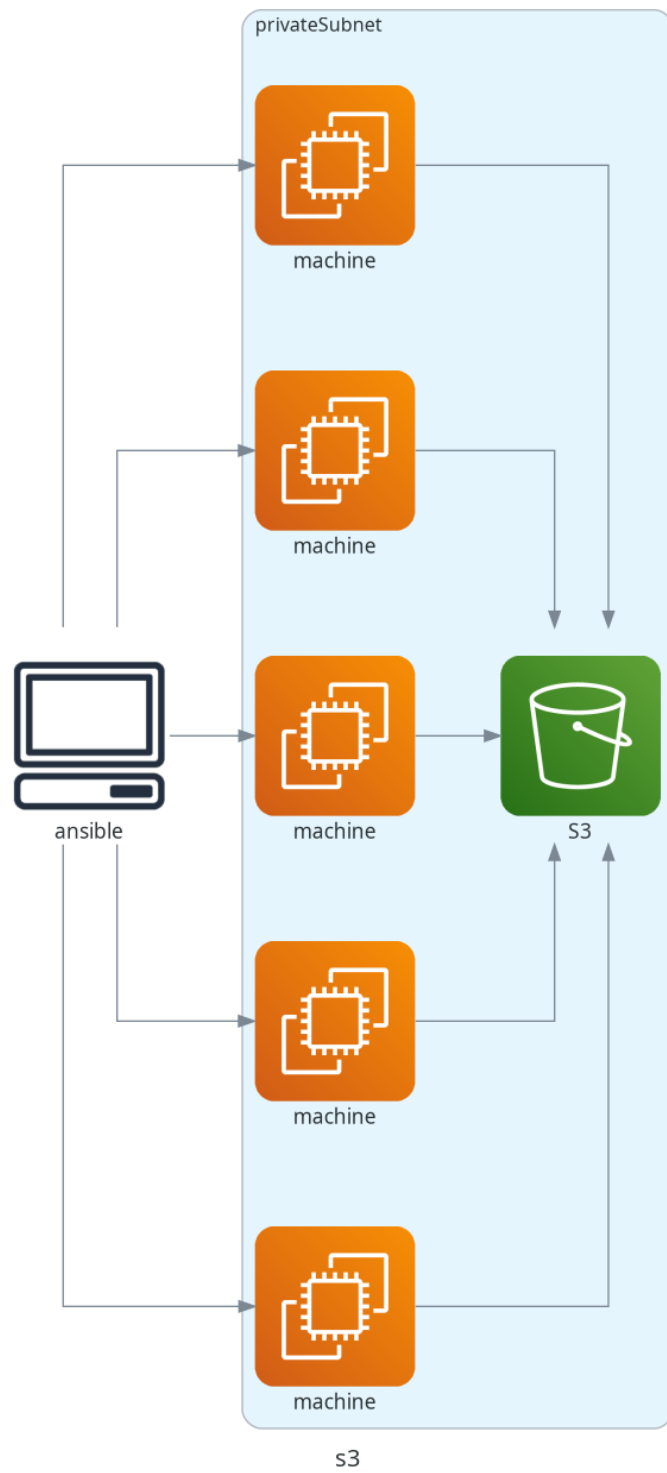
Agradeço primeiramente a Deus pelo suporte do universo, a minha noiva pela compreensão e por ser meu norte, que me impediu de surtar, e por fim mas não menos importante, a minha orientadora que foi extremamente compreensiva, e por fim a marlene.



13
Figura 1 – Arquitetura do cluster usando lustre Fonte: fornecido pelo autor.



14
Figura 2 – Arquitetura do cluster usando lustre Fonte: fornecido pelo autor.



15
Figura 3 – Arquitetura do cluster usando s3 Fonte: fornecido pelo autor.