

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

Fernando de Almeida Silva

HPC-SBC: Um esforço experimental para avaliar o armazenamento em configurações de computação de alto desempenho usando uma abordagem contextual

Juiz de Fora

2024

Fernando de Almeida Silva

HPC-SBC: Um esforço experimental para avaliar o armazenamento em configurações de computação de alto desempenho usando uma abordagem contextual

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Computação da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Ciências da Computação. Área de concentração: Ciência da Computação

Orientador: Prof. Dr. Mário Antônio Ribeiro Dantas

Coorientador: Prof. Dr. José Maria Nazar David

Juiz de Fora

2024

Ficha catalográfica elaborada com os dados fornecidos pelo autor

Silva, Fernando de A..

HPC-SBC: Um esforço experimental para avaliar o armazenamento em configurações de computação de alto desempenho usando uma abordagem contextual / Fernando de Almeida Silva. – 2024.

59 f. : il.

Orientador: Mário Antônio Ribeiro Dantas

Coorientador: José Maria Nazar David

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Programa de Pós-Graduação em Ciências da Computação, 2024.

1. Contexto de Armazenamento. 2. Desempenho de Armazenamento. 3. Memória Secundária. 4. Arquitetura de Ambientes Distribuídos Paralelos. 5. Armazenamento. 6. HPC.

I. Dantas, Mario, orient. II Nazar, José Maria, coorient. III. Título.

Fernando de Almeida Silva

HPC-SBC: Um esforço experimental para avaliar o armazenamento em configurações de computação de alto desempenho usando uma abordagem contextual

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Computação, da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Ciência da Computação. Área de concentração: Ciência da Computação.

Aprovada em 25 de março de 2024.

BANCA EXAMINADORA

Prof. Dr. Mario Antonio Ribeiro Dantas - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. José Maria N. David - Coorientador

Universidade Federal de Juiz de Fora

Prof^a. Dra. Bárbara Melo Quintela

Universidade Federal de Juiz de Fora

Prof. Dr. Eduardo Camilo Inácio

Universidade Federal de Santa Catarina

Juiz de Fora, 13/03/2024.



Documento assinado eletronicamente por **mario antonio ribeiro dantas, Usuário Externo**, em 26/06/2024, às 16:18, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Camilo Inacio, Usuário Externo**, em 26/06/2024, às 16:36, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jose Maria Nazar David, Professor(a)**, em 26/06/2024, às 21:34, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Barbara de Melo Quintela, Professor(a)**, em 16/07/2024, às 11:59, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1749315** e o código CRC **AF0CF53C**.

Dedico este trabalho aos meus falecidos pais, Joaquim Fernandes da Silva e Maria Elza de Almeida que foram exemplos de perseverança e minha força motivadora nos momentos mais difíceis.

AGRADECIMENTOS

Agradeço à minha esposa Adriana e minha irmã Cátia, pelo amor, carinho e apoio emocional.

Gostaria de agradecer ao Professor Jean-François Méhaut da Université Grenoble Alpes pela atenção e pelo uso do ambiente de pesquisas Grid'5000.

Agradeço aos meus colegas e amigos, como Walkíria Garcia de Souza Silveira, João Pedro de Souza Jardim da Costa e demais colegas pela troca de conhecimento nos projetos e a experiência adquirida ao longo desta jornada.

Este trabalho foi realizado com apoio do Programa de Apoio à Qualificação (Proquali) da UFJF, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, CNPq, FAPEMIG e INESC P&D Brasil.

Agradeço especialmente ao meu orientador, Mário Dantas, e ao coorientador, José Maria Nazar, pelo tempo e conhecimento compartilhados. Ambos semelhantes ao farol, que no meio da escuridão, indicam o caminho a trilhar.

“If I have seen further it is by standing on the shoulders of Giants”
(Newton, 1676)¹

¹ “Se eu vi mais longe, foi por estar de pé sobre ombros de gigantes” (Tradução Nossa).

RESUMO

Ambientes de alto desempenho, muitas vezes, enfrentam gargalos de entrada e saída (input output - I/O). Um fator é a lacuna entre o poder de processamento e o desempenho de armazenamento. Isso se dá, pois as tecnologias de armazenamento evoluem a uma velocidade mais lenta que a de processamento, contribuindo para o aumento do problema de desempenho de I/O. Assim, se fazem cada vez mais relevantes contribuições para mitigar lacunas de I/O. Esta dissertação, portanto, apresenta um esforço experimental em avaliar o desempenho da memória secundária dos nodos computacionais em ambientes de alto desempenho, ao utilizar como exemplo o ambiente Grid'5000. Dessa forma, obtém-se de forma mais técnica e científica informações sobre a quantidade, tipo e desempenho dos dispositivos de armazenamento dos nodos em ambientes distribuídos paralelos. Em complemento a isso, uma proposta foi desenvolvida para análise de armazenamento destes ambientes de alto desempenho, que foi denominada como HPC-SBC (*HPC Storage Benchmark Context*). Esta abordagem computacional pode suprir as limitações observadas nas ferramentas disponíveis no ambiente computacional Grid'5000, por fornecer de forma diferenciada informações sobre o desempenho dos dispositivos de memória secundária disponíveis em cada nodo deste ambiente. Esta ferramenta permite analisar a taxa de transferência de cada dispositivo encontrado e ainda fornecer aos usuários de forma automatizada as informações sobre o desempenho dos dispositivos de armazenamento no contexto de *cluster*. Esta abordagem contextual proveu resultados iniciais que indicam o sucesso da pesquisa, ao apontar dentre os nodos analisados no ambiente de processamento de alto desempenho (High Performance Computing - HPC), aqueles dispositivos de armazenamento com o melhor desempenho.

Palavras-chave: Contexto de Armazenamento; Desempenho de Armazenamento; Memória Secundária; Arquitetura de Ambientes Distribuídos Paralelos; Armazenamento; HPC.

ABSTRACT

High-performance environments often face input and output (I/O) bottlenecks. One factor is the gap between processing power and storage performance. This is because storage technologies evolve at a slower speed than processing, contributing to the increase of I/O performance problems. Thus, increasingly relevant contributions are being made to mitigate I/O gaps. This dissertation, therefore, presents an experimental effort to evaluate the performance of the secondary memory of computing nodes in high-performance environments, using the Grid'5000 environment as an example. In this way, information about the quantity, type and performance of node storage devices in parallel distributed environments is obtained in a more technical and scientific way. In addition to this, a proposal was developed to analyze the storage of these high-performance environments, which was called HPC-SBC (HPC Storage Benchmark Context). This computational approach can overcome the limitations observed in the tools available in the Grid'5000 computing environment, by providing differentiated information about the performance of the secondary memory devices available in each node in this environment. This tool allows you to analyze the transfer rate of each device found and also provide users with automated information about the performance of storage devices in the cluster context. This contextual approach provided initial results that indicate the success of the research, by pointing out among the nodes analyzed in the high-performance processing environment (High Performance Computing - HPC), those storage devices with the best performance.

Keywords: Storage Context; Storage Performance; Secondary Memory; Architecture of Parallel Distributed Environments; Storage; HPC.

LISTA DE ILUSTRAÇÕES

Figura 1 – Visão geral do fluxo das atividades da pesquisa.....	19
Figura 2 – Um modelo típico de infraestrutura de ambientes distribuídos HPC	25
Figura 3 – Visão geral da arquitetura do ambiente computacional Grid'5000	30
Figura 4 – Nodos de <i>cluster</i> do <i>site</i> Grenoble	31
Figura 5 – A indicação da disponibilidade de nodos na ferramenta	32
Figura 6 – Dispositivos de armazenamento de nodos de <i>cluster</i> do <i>site</i> Lille	33
Figura 7 – Exibição parcial de nodos e seus dispositivos de armazenamento dos existentes no <i>site</i> de Grenoble	34
Figura 8 – Mensagem ao tentar listar discos disponíveis dos nodos de alguns <i>sites</i>	35
Figura 9 – Diagrama de alto nível da solução proposta	40
Figura 10 – Visão sistemática dos componentes do <i>software</i> HPC-SBC.....	42
Figura 11 – Gráfico com o limite superior (MAX) e limite inferior (MIN) dos dispositivos de armazenamento por <i>site</i>	52
Figura 12 – Gráfico com o limite superior (MAX) e limite inferior (MIN) de dispositivos de armazenamento por <i>cluster</i>	53

LISTA DE TABELAS

Tabela 1 – Valores de limite inferior (MIN) e limite superior (MAX) dos dispositivos.	48
Tabela 2 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos SSD	49
Tabela 3 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos NVMe...	49
Tabela 4 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos NVMe com maior diferença no mesmo nodo	50
Tabela 5 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos SSD com maior diferença no mesmo nodo	50
Tabela 6 – Alcance de cobertura dos testes no ambiente computacional Grid'5000.	54

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
CSV	Comma-separated values
FTP	File Transfer Protocol
HPC	<i>High-performance computing</i>
I/O	<i>Input/Output</i>
IOPS	<i>Input/Output per Second</i>
MAX	Máximo (limite superior)
MB	Megabyte
MB/s	Megabytes per Second
MIN	Mínimo (limite inferior)
NVMe	Non-Volatile Memory Express
OAR	Gerenciador de recursos e tarefas para <i>clusters</i> HPC
PCI	Peripheral Component Interconnect
PCIe	Peripheral Component Interconnect Express
SATA	Serial ATA
SQL	Structured Query Language
SSD	Solid State Drive
SSH	Secure Shell

SUMÁRIO

1	INTRODUÇÃO	13
1.1	MOTIVAÇÃO	13
1.2	OBJETIVO GERAL.....	14
1.3	OBJETIVOS ESPECÍFICOS	14
1.4	CONTRIBUIÇÕES DO TRABALHO	15
1.5	ESTRUTURA DO TRABALHO	16
2	MATERIAIS E MÉTODOS	18
2.1	INSTRUMENTOS DE PESQUISA.....	20
2.2	DELIMITAÇÃO DAS INFORMAÇÕES OBTIDAS.....	20
2.3	TERMOS RELACIONADOS	20
2.3.1	Armazenamento.....	21
2.3.2	Computação distribuída.....	21
2.4	CONSIDERAÇÕES FINAIS DO CAPÍTULO	23
3	TRABALHOS RELACIONADOS	25
3.1	CARACTERÍSTICAS DE ARMAZENAMENTO EM AMBIENTES HPC.....	25
3.2	CONSIDERAÇÕES FINAIS DO CAPÍTULO	28
4	AMBIENTE EXPERIMENTAL	29
4.1	AMBIENTE COMPUTACIONAL HPC GRID'5000	30
4.2	FERRAMENTAS DISPONÍVEIS.....	31
4.3	FERRAMENTAS NO CONTEXTO DE ARMAZENAMENTO.....	33
4.4	DESENVOLVIMENTO EXTERNO DE FERRAMENTAS.....	36
4.5	CONSIDERAÇÕES FINAIS DO CAPÍTULO	36
5	PROPOSTA	38

5.1	VISÃO INICIAL	38
5.2	IMPLEMENTAÇÃO.....	41
5.3	CONFIGURAÇÃO EXPERIMENTAL.....	44
5.4	CONSIDERAÇÕES FINAIS DO CAPÍTULO	45
6	RESULTADOS EXPERIMENTAIS.....	47
6.1	ANÁLISE INICIAL	47
6.2	ANÁLISE DOS DISPOSITIVOS DE ARMAZENAMENTO DO AMBIENTE ..	48
6.3	ANÁLISE COMPARATIVA DE ARMAZENAMENTO POR <i>SITE</i> E POR <i>CLUSTER</i>	51
6.4	ANÁLISE DA COBERTURA DOS TESTES	53
6.5	CONSIDERAÇÕES FINAIS DO CAPÍTULO	55
7	CONCLUSÕES E TRABALHOS FUTUROS	57
	REFERÊNCIAS	59
	APÊNDICE A – PUBLICAÇÕES COMO AUTOR PRINCIPAL	63
	APÊNDICE B – PUBLICAÇÕES COMO CO-AUTOR	65
	APÊNDICE C – HPCSBC (HPC STORAGE BENCHMARK CONTEXT)	66

1 INTRODUÇÃO

A crescente produção e consumo de dados mudou a forma como lidamos com a informação. Porém, estudos e pesquisas não exigem apenas alto poder de processamento, mas também geram um enorme volume de dados. Assim, a aquisição, armazenamento, análise e visualização de dados desempenham um papel cada vez mais importante. Rivas-Gomez (2019) argumenta sobre a importância dos dispositivos de armazenamento, ou seja, que a memória secundária pode afetar no desempenho dos experimentos.

Esta diferença entre os elementos de processamento e o desempenho de Entrada e Saída (E/S), ou *Input Output (I/O)*, é chamada de lacuna de I/O como descreve Neuwirth (2019). Como os avanços na redução da latência parecem ser muito limitados, uma possível solução para aumentar o desempenho é explorar algum tipo de paralelismo. Por exemplo, Lößer, Raskob e Scheuermann (2023), em sua pesquisa, utilizam técnicas de múltiplas operações de I/O em paralelo, em vez de executá-las sequencialmente.

Inclusive nesse cenário, a escolha dos dispositivos de armazenamento com maior desempenho pode ser mais vantajoso do que escolher aleatoriamente. Por exemplo, em ambientes de alto desempenho que possuem nodos disponíveis em quantidade superior ao da necessária para um experimento, permitiria uma escolha mais criteriosa através do desempenho da memória secundária dos mesmos.

Dentre estes ambientes, pode-se citar o Grid'5000, que é um ambiente de testes flexível e em grande escala para pesquisas baseadas em experimentos em todas as áreas da ciência da computação, com foco em computação paralela e distribuída.

1.1 MOTIVAÇÃO

Ao explorar de forma preliminar as ferramentas disponíveis no ambiente Grid'5000 para avaliar o desempenho da memória secundária, não foi possível obter as informações sobre o desempenho de armazenamento dos dispositivos de armazenamento dos nodos e *clusters* nas ferramentas observadas. Tal situação observada apresentou-se como uma oportunidade de melhoria para as ferramentas do ambiente.

Por isso questionou-se: Como poderia ser desenvolvida uma abordagem computacional que fornecesse a situação das configurações de armazenamento dos nodos de um ambiente computacional de alto desempenho?

1.2 OBJETIVO GERAL

Este trabalho caracteriza-se como um esforço experimental em mostrar as diferenças das configurações de armazenamento em ambientes de alto desempenho no contexto de *cluster*. Em adição, promover o desenvolvimento de uma abordagem computacional, denominada como HPC-SBC (*HPC Storage Benchmark Context*), que tem como objetivo fornecer aos usuários, informações sobre o desempenho de cada dispositivo de armazenamento dos nodos, a fim de auxiliar na escolha destes em experimentos que demandem um armazenamento diferenciado. Assim sendo, tem a possibilidade de ampliar as informações fornecidas pelas ferramentas disponíveis no ambiente.

1.3 OBJETIVOS ESPECÍFICOS

A seguir são apresentadas a etapas necessárias para atingir o objetivo geral:

- a) Revisar a literatura sobre ambientes de alto desempenho no intuito de esclarecer conceitos e ideias, com vista na formulação de problemas e hipóteses de forma mais precisa;
- b) Compreender as configurações e dispositivos de armazenamento em ambientes de alto desempenho para delimitar um escopo comparativo ao analisar um exemplo real de ambiente HPC;
- c) Analisar as ferramentas disponíveis do ambiente Grid'5000 que possibilitam a visualização das configurações de armazenamento, como por exemplo a quantidade, tipo e desempenho dos dispositivos de armazenamento de cada nodo;
- d) Desenvolver uma abordagem computacional que forneça mais informações sobre os dispositivos de armazenamento dos nodos computacionais nos diversos *clusters* do ambiente em estudo, de forma a ampliar as funcionalidades das ferramentas disponíveis;

- e) Testar a ferramenta desenvolvida com a finalidade de produzir dados estruturados sobre as configurações de armazenamento do ambiente de alto desempenho Grid'5000.
- f) Analisar os dados obtidos pela ferramenta para avaliar os dados sobre a quantidade, tipo e desempenho dos dispositivos de armazenamento dos nodos de cada *cluster* do ambiente. Isso permite uma comparação com os dados fornecidos pelas ferramentas disponíveis no ambiente computacional de alto desempenho.
- g) Examinar, a partir dos dados obtidos pelo conjunto de experimentos, a amplitude da cobertura dos testes, para estimar o alcance da ferramenta na obtenção de métricas comparativas de desempenho no ambiente Grid'5000, devido à heterogeneidade dos tipos de dispositivos de armazenamento nos *clusters* do ambiente.

1.4 CONTRIBUIÇÕES DO TRABALHO

Esta pesquisa contribui para o melhor entendimento das configurações de armazenamento no contexto de ambientes computacionais de alto desempenho, ao utilizar para isso o ambiente Grid'5000, como exemplo. Para além do objetivo central desta investigação, visa-se detalhar o seu funcionamento, as ferramentas disponíveis e como utilizá-las. Assim, visa-se também promover o uso de ambientes HPC em meios acadêmicos, já que se trata de uma plataforma multifuncional para experimentos científicos. Desta forma, pretende-se contribuir para um melhor entendimento em termos de armazenamento, já que ambientes computacionais de alto desempenho possuem particularidades quando comparados com ambientes computacionais convencionais.

Ao explorar este ambiente computacional de alto desempenho, o estudo também aponta as dificuldades e limitações encontradas ao se trabalhar neste ambiente. Desta forma, permite-se que pesquisas futuras se debrucem sobre as ferramentas disponíveis para o aprimoramento do funcionamento e usabilidade destas. Ao apontar limitações e potencialidades das ferramentas no ambiente Grid'5000, este estudo também pode indicar direções de melhorias em outros ambientes HPC semelhantes. Isso posto, as observações obtidas neste estudo em termos de armazenamento podem ser implementadas em outros ambientes e,

consequentemente, melhorando a experiência dos usuários e a qualidade dos experimentos.

Como principal contribuição desta pesquisa, o desenvolvimento da ferramenta chamada HPC-SBC (*HPC Storage Benchmark Context*) visou prover um diferencial em obter informações sobre as configurações de armazenamento. Com isso, teve-se o intuito de mitigar a lacuna de I/O em experimentos que demandam um armazenamento diferenciado.

A abordagem de *software* desenvolvida possibilita de forma diferenciada, encontrar os dispositivos de armazenamento disponíveis em cada nodo computacional. Em adição, auxilia na análise da taxa de transferência de cada dispositivo de armazenamento encontrado. As tarefas são executadas simultaneamente em vários nodos do ambiente. Desta forma, os usuários finais obtêm uma lista de forma automatizada com as informações, sobre o desempenho de cada dispositivo de armazenamento de cada nodo computacional. Informações essas que não foram obtidas pelas ferramentas analisadas no ambiente Grid'5000, o que diferenciou a ferramenta HPC-SBC das outras soluções disponíveis.

1.5 ESTRUTURA DO TRABALHO

O esforço experimental desta pesquisa é descrito nesta dissertação foi separado em seis capítulos: 1. Introdução, 2. Materiais e Métodos, 3. Trabalhos Relacionados, 4. Ambiente Experimental, 5. Proposta e Resultados Experimentais e 6. Conclusões e Trabalhos Futuros.

Por conseguinte, para sintetizar as etapas da pesquisa nesta dissertação, o Capítulo 2 apresenta os materiais e métodos utilizados no esforço experimental. O Capítulo 3 aborda os trabalhos relacionados sobre ambientes HPC, de forma a examinar os conceitos relacionados a este tipo de ambiente para um melhor entendimento das configurações de armazenamento deste ambiente a fim de compor um arcabouço teórico. O Capítulo 4 apresenta detalhes sobre o ambiente HPC, ao abordar como exemplo o Grid'5000, suas principais características e ferramentas. Além disso, esta seção apresenta os desafios e limitações em obter informações sobre o armazenamento dos nodos e seus respectivos desempenhos. O Capítulo 5 explica a abordagem de *software* proposta para preencher a lacuna encontrada no ambiente HPC em relação às configurações de armazenamento. Em

adição, no Capítulo 6 os experimentos realizados com a ferramenta desenvolvida HPC-SBC são descritos e os resultados obtidos são analisados de forma textual e gráfica. Por fim, o Capítulo 7 traz discussões, limitações e potencialidades desta pesquisa, bem como possibilidades para trabalhos futuros e acréscimos às considerações finais.

2 MATERIAIS E MÉTODOS

O raciocínio indutivo de desenvolvimento da hipótese desta pesquisa se fundamenta em um esforço experimental e comparativo particular ao universo de armazenamento em ambientes de alto desempenho no contexto de *cluster*. Utilizou-se para isso o ambiente computacional Grid'5000, como exemplo. Desta forma, foi proposto o desenvolvimento de uma ferramenta denominada HPC-SBC, para avaliar o desempenho de armazenamento deste ambiente e comparar com as informações obtidas, com as fornecidas pelo ambiente Grid'5000.

O Grid'5000 é um ambiente de testes flexível e de grande escala para pesquisas baseadas em experimentos em todas as áreas da ciência da computação, com foco em computação paralela e distribuída, incluindo nuvem, HPC, *Big Data* e IA. Este ambiente permite realizar os experimentos necessários para esta pesquisa, sobre o armazenamento em ambientes de alto desempenho no contexto de *cluster*.

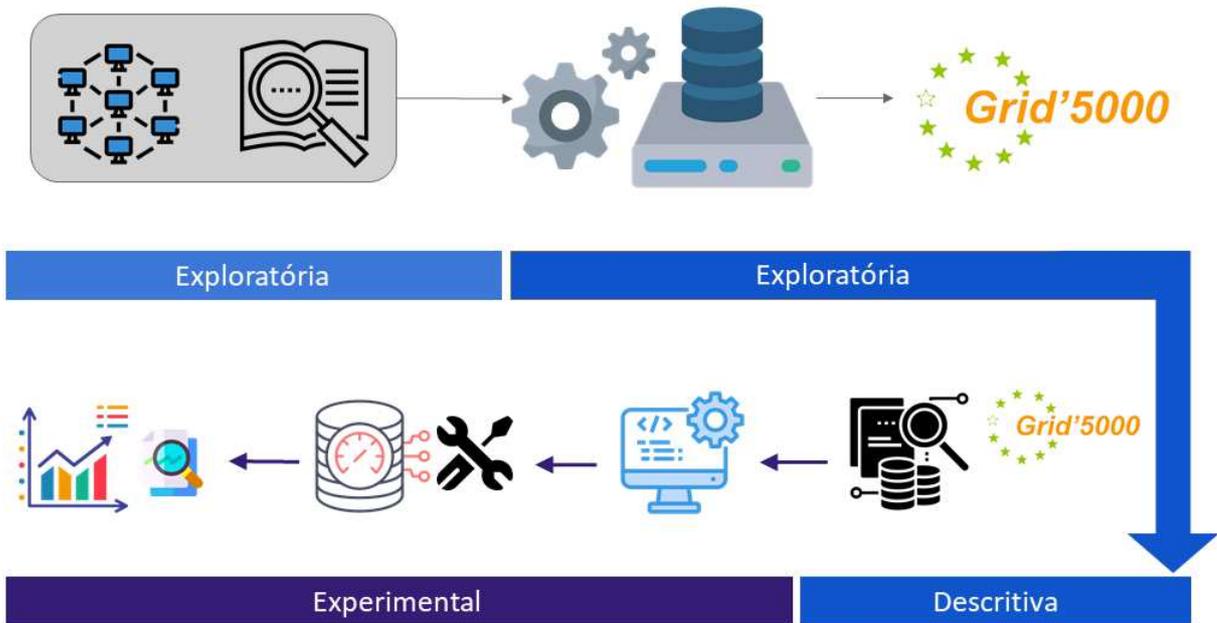
Pelo fato de o Grid'5000 ser um ambiente computacional complexo quando comparado com ambientes computacionais convencionais, esta investigação se caracteriza como exploratória. Pelo fato de objetivar a formulação de problemas mais precisos ou hipóteses precisas para estudos futuros, a pesquisa também se enquadra como um estudo de caso exploratório e descritivo, uma vez que visa caracterizar o funcionamento do ambiente e as ferramentas disponíveis nele.

Em adição, trata-se de um estudo de modelagem computacional, cuja análise dos resultados é de natureza quali-quantitativa, com o objetivo de obter os dados sobre a quantidade, tipo e taxa de transferência dos dispositivos de armazenamento dos nodos dos *clusters* no ambiente, por meio de uma abordagem computacional denominada como HPC-SBC (*HPC Storage Benchmark Context*). Através das informações obtidas pela ferramenta, pôde-se comparar e avaliar o desempenho dos dispositivos de armazenamento do ambiente computacional Grid'5000. De forma semelhante, a aplicação deste esforço experimental, em relação às configurações de armazenamento em ambientes de alto desempenho no contexto de *cluster*, poderia ser extrapolada para outros ambientes semelhantes.

Para o desenvolvimento deste trabalho, foi possível agrupar os materiais e métodos utilizados nesta dissertação em fases que se relacionam aos três tipos de

pesquisas que foram utilizadas: exploratória, descritiva e experimental. A Figura 1 ilustra o fluxo de atividades da pesquisa.

Figura 1 – Visão geral do fluxo das atividades da pesquisa



Fonte: Elaborada pelo autor (2024).

Na primeira fase deste esforço experimental exposta no Capítulo 3, realiza-se o levantamento do conhecimento teórico sobre armazenamento de dados em ambientes computacionais de alto desempenho, como por exemplo, o Grid'5000. No Capítulo 4, aborda-se a segunda fase, que foi realizada a partir da exploração da documentação e das ferramentas disponíveis no ambiente com ênfase no armazenamento. Esta teve, o objetivo de avaliar possíveis melhorias. Na terceira fase, no Capítulo 5, explana-se sobre a abordagem computacional HPC-SBC, seus principais objetivos e as pesquisas realizadas de forma iterativa e incremental para implementar os recursos desejados. Complementa-se esta fase no Capítulo 6, ao apresentar as análises das informações obtidas através do conjunto de experimentos no ambiente computacional HPC, ao utilizar a abordagem computacional proposta.

Explanam-se, a seguir, os métodos utilizados para obtenção do conhecimento teórico desta dissertação.

2.1 INSTRUMENTOS DE PESQUISA

Para construir a base de conhecimento, utilizada no desenvolvimento deste documento sobre o armazenamento de dados no ambiente Grid'5000, foi feito um levantamento de estudos relacionados. Através de buscas nas bases de dados IEEE Xplore e Google *Scholar*, com objetivo de encontrar trabalhos relacionados com a área em questão.

2.2 DELIMITAÇÃO DAS INFORMAÇÕES OBTIDAS

Para delimitar o universo da amostra, foram utilizadas palavras-chave como: “I/O”, “HPC”, “grid5000”, “*storage*”. Além disso, os resultados das buscas foram filtrados para trabalhos publicados desde 2018, resultando em 150 documentos encontrados. No intuito de delimitar o *corpus* de leitura, foram analisados fatores relevantes que estivessem de acordo com o projeto desta dissertação por meio da leitura do título, resumo, conclusões e métodos utilizados, nesta ordem. Em seguida, esses documentos foram comparados com publicações de referência citadas pelos orientadores, Meffe, Mussi e Mello (2006) e Dantas (2005), que estão relacionados à proposta da dissertação. A abordagem utilizada foi uma comparação das características dos trabalhos encontrados com as publicações utilizadas como base acima mencionadas. A partir desta filtragem, obteve-se cerca de 20 documentos que foram estudados. Destes, seis trabalhos relacionados à proposta da dissertação foram abordados, que formaram a base de conhecimento teórica para compor a revisão da literatura necessária para a fase seguinte.

2.3 TERMOS RELACIONADOS

Nesta seção apresentamos uma breve visão de alguns termos, que foram abordados no esforço experimental da proposta desta dissertação. Tanto sobre armazenamento, incluindo alguns tipos de dispositivos analisados, suas características e diferenças relevantes entre estes. Da mesma maneira, tratou-se além de alguns termos relacionados à computação distribuída, que é utilizada em ambientes HPC. Tal fundamentação se faz necessária para contextualizar a

pesquisa, mostrar seu embasamento e explicar a relevância dos conceitos escolhidos, além de contribuir para uma melhor compreensão deste trabalho acadêmico.

2.3.1 Armazenamento

Um dispositivo de armazenamento é um componente de *hardware* usado para armazenar e recuperar informações digitais. Os dispositivos tradicionais de armazenamento magnético, como unidades de disco rígido (HDDs), embora tenham prevalecido por décadas, dependem de movimento físico e campos magnéticos, impactando negativamente seu desempenho. Assim, com o objetivo de melhorar o desempenho dos dispositivos de armazenamento, avanços recentes introduziram tecnologias que aproveitam a memória baseada em semicondutores para armazenar os dados eletronicamente. Estes dispositivos chamados *Solid State Drive* ou unidade em estado sólido (SSDs) permitem um acesso mais rápido e eficiente aos dados, pois elimina a necessidade de movimentos físicos no interior do dispositivo, e isto permite tempos de acesso mais rápidos, menor latência e maior durabilidade devido à ausência de partes móveis (Rivas-Gomez, 2019).

Entre as interfaces diferentes usadas para dispositivos de armazenamento SSD, pode-se citar as SATA e *PCI Express* (PCIe). Os SSDs SATA por meio da interface SATA, oferecem transferência de dados de alta velocidade, mas limitada pela largura de banda SATA. Por outro lado, os SSDs NVMe aproveitam o barramento *PCI Express* de alta velocidade para melhorar o desempenho. Portanto, embora o SSD SATA seja comum e eficiente, a utilização de SSD NVMe, por utilizar o barramento *PCI Express* pode aprimorar ainda mais o desempenho do armazenamento, especialmente em ambientes de computação de alto desempenho (Neuwirth, 2024).

2.3.2 Computação distribuída

A computação distribuída envolve a utilização de vários computadores interconectados para resolver problemas complexos, oferecendo assim recursos de computação de alto desempenho (Kyaw e Phyu, 2020). Tanto a computação em *grid* quanto a computação em *cluster* exemplificam os paradigmas de computação

distribuída e desempenham um papel crucial na facilitação da pesquisa científica em ambientes computacionais de alto desempenho.

Acerca de *cluster*, Kyaw e Phyu (2020) descrevem como um grupo de dispositivos computacionais chamados de nodos, que funcionam em conjunto como um único sistema. Cada nodo do *cluster* é conectado com os demais, permitindo a transferência de dados entre si. Em adição, os *clusters* são projetados para fornecer computação e disponibilidade de alto desempenho, superior ao que seria fornecido por um único nodo. Além disso, descrevem nodo computacional como uma unidade discreta de um sistema de computador que executa seu próprio sistema operacional.

Sobre *Grid* Computacional, Meffe, Mussi e Mello (2006) propõem um conceito de uma plataforma de fornecimento de serviços computacionais que apresenta as determinadas características. Contudo, a ausência de alguma das características não desqualifica automaticamente uma determinada plataforma como *Grid*:

a) Heterogeneidade: A infraestrutura é caracterizada pela diversidade de seus componentes, que inclui recursos computacionais de diferentes tipos e uma ampla variedade de instrumentos e serviços;

b) Dispersão geográfica: Podem alcançar uma escala global, que integra serviços localizados em diversas regiões do mundo;

c) Compartilhamento: Não podem ser exclusivamente dedicados a uma única aplicação por um período determinado. Isto impacta o desenvolvimento de aplicações sobre essa infraestrutura;

d) Múltiplos domínios administrativos: Reúnem recursos de diversas instituições, o que pode resultar em diferentes políticas de acesso e uso de serviços, de acordo com as diretrizes de cada domínio participante;

e) Controle distribuído: Permite que cada instituição implemente suas políticas locais sem interferir diretamente nas políticas de acesso das outras instituições.

Em adição, na computação distribuída, os papéis que os nodos desempenham são cruciais na facilitação de uso das várias funções dentro do ambiente, servindo a propósitos específicos, como por exemplo, a execução de tarefas, serviços de rede e gerenciamento de recursos (Bertot; Nussbaum e Margery, 2020). Entre estes papéis dos nodos em ambientes de computação de alto desempenho, pode-se citar o *worker node* e o *farmer node*. Sendo que os nodos no papel de *worker* são responsáveis por executar tarefas e gerenciar recursos com eficiência. Enquanto os nodos no papel de *farmer* lidam com tarefas de alocação e

provisionamento de recursos. A interação harmoniosa entre os nodos em ambos os papéis é essencial para alcançar o desempenho ideal em sistemas de computação distribuída (Pereira, 2023).

2.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Este capítulo detalhou os materiais e métodos utilizados para esta pesquisa, que visou avaliar e comparar o desempenho de armazenamento em ambientes HPC no contexto de *clusters*, utilizando a ferramenta HPC-SBC no ambiente Grid'5000 como exemplo.

A escolha do Grid'5000 como ambiente experimental se mostrou adequada, dado seu caráter de pesquisa exploratória e descritiva, permitindo a formulação de hipóteses precisas para estudos futuros. A pesquisa forneceu esclarecimentos sobre a interação entre diferentes tipos de dispositivos de armazenamento e o desempenho destes em ambientes computacionais de alto desempenho, contribuindo para o avanço do conhecimento na área.

A pesquisa foi estruturada em fases distintas, incluindo a coleta de conhecimento teórico, exploração do ambiente Grid'5000, desenvolvimento iterativo da ferramenta HPC-SBC e análise dos dados obtidos por meio do conjunto de experimentos utilizando esta ferramenta. Esta teve um caráter exploratório, descritivo e experimental, focado na coleta e análise quali-quantitativa de dados sobre os dispositivos de armazenamento dos nodos.

Na primeira fase deste esforço experimental, foi realizado um levantamento teórico sobre armazenamento em ambientes HPC, seguido de uma exploração da documentação e ferramentas disponíveis no Grid'5000. Na segunda fase, foram identificadas possíveis melhorias no ambiente, e na terceira fase, foi desenvolvida a HPC-SBC, uma abordagem computacional iterativa e incremental para implementar os recursos desejados.

Os métodos utilizados envolveram uma abordagem quali-quantitativa para obter dados detalhados sobre os dispositivos de armazenamento. A comparação das informações obtidas pela ferramenta HPC-SBC, incluindo quantidade, tipo e taxa de transferência com as informações fornecidas pelo ambiente Grid'5000 permitiu uma avaliação do desempenho de armazenamento, destacando as capacidades e limitações dos diferentes tipos de dispositivos. Estas informações proporcionaram

dados detalhados e comparativos que podem ser úteis para aprimorar experimentos futuros e otimizações de configurações de armazenamento em ambientes HPC.

O próximo capítulo revisará trabalhos relacionados com foco nas características de armazenamento em ambientes HPC, analisando abordagens, tecnologias e soluções na literatura, destacando inovações e desafios. Esta revisão fornecerá uma base sólida para compreender práticas e tendências atuais, identificando lacunas para nossa pesquisa. Em busca de uma compreensão crítica do estado da arte, de contribuições significativas e oportunidades para avanços futuros.

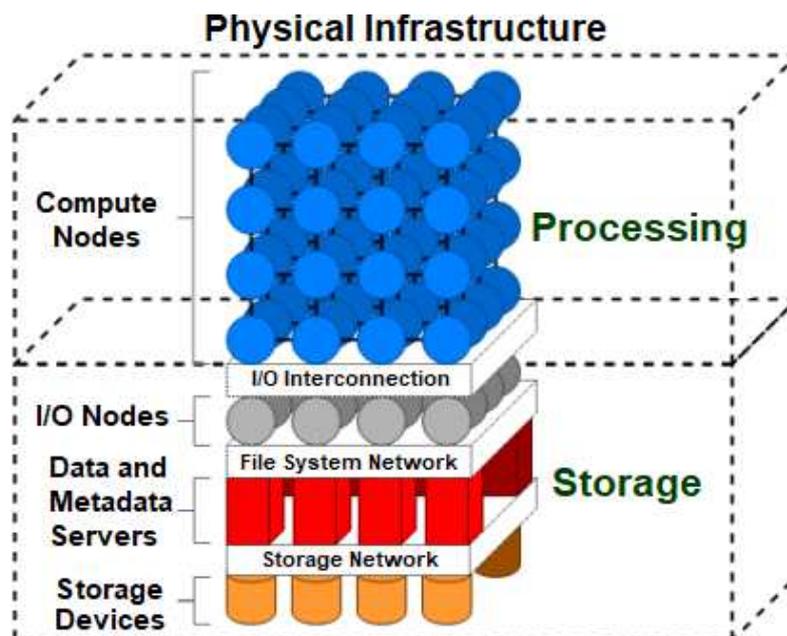
3 TRABALHOS RELACIONADOS

A partir das bases de dados citadas na seção 2.1, foram feitas buscas sobre trabalhos relacionados à proposta deste projeto. As palavras-chave utilizadas foram “I/O”, “HPC”, “grid5000”, “storage” e combinações das mesmas. Deste conjunto, alguns trabalhos foram selecionados para serem discutidos neste capítulo.

3.1 CARACTERÍSTICAS DE ARMAZENAMENTO EM AMBIENTES HPC

Pioli *et al.* (2020) descrevem os ambientes distribuídos dos sistemas HPC como uma combinação muito especializada e complexa de elementos de *hardware* e *software*. Estes ambientes se concentram principalmente em fornecer alto poder de processamento para aplicações paralelas e distribuídas em larga escala. Os autores destacam inclusive que a baixa latência de comunicação e acesso a dados são considerados requisitos cada vez mais relevantes. A Figura 2 apresenta uma visão geral de um modelo de infraestrutura geral comum em muitos ambientes HPC. Mesmo ao utilizar uma ilustração simplificada, a complexidade desses ambientes pode ser percebida, dado que demonstra a interação de diferentes elementos de forma simultânea.

Figura 2 – Um modelo típico de infraestrutura de ambientes distribuídos HPC



Fonte: Adaptado de Pioli *et al.* (2020, p. 2096).

Ambientes de alto desempenho, como sistemas HPC, muitas vezes enfrentam gargalos de I/O. Um fator que aumenta este problema é a lacuna entre o poder de processamento e o desempenho de armazenamento. Como as tecnologias de armazenamento evoluem a uma velocidade mais lenta que o processamento, isso também contribui para o aumento do problema de desempenho de I/O. Além disso, os aplicativos de alto desempenho, podem mover ou gerar exabytes de dados computados entre nodos, e seu desempenho depende do processo de I/O. Normalmente, uma arquitetura em *cluster*, disponível no ambiente de computação Grid'5000, é usada para executar esses aplicativos. Portanto, gargalos de desempenho de I/O são uma importante área de estudo para encontrar soluções que melhorem o desempenho de I/O (Pioli; Ströele e Dantas, 2023).

Especialmente nos campos científico e industrial, o armazenamento é um elemento chave, porque normalmente essas aplicações empregam uma enorme quantidade de dados. O desempenho dessas aplicações complexas está diretamente ligado ao tempo de execução das operações de entrada e saída (I/O). Pioli, Ströele e Dantas (2020) apresentam em seu trabalho uma pesquisa que visa melhorar o desempenho do armazenamento para aplicações de computação de alto desempenho e computação escalável com uso intensivo de dados. Eles avaliam algumas melhorias nesse contexto para abordar questões relacionadas à camada de I/O. Os experimentos foram conduzidos no ambiente distribuído Grid'5000, que é adequado para compreender os desafios enfrentados por aplicações HPC. Os resultados dos *benchmarks* de I/O demonstram melhorias no desempenho, especialmente na redução da latência das operações de I/O.

Em seu trabalho Inacio e Dantas (2018) destacam que o processo de avaliação de desempenho de I/O, prova ser um esforço complexo. Apesar da importância deste assunto, a comunidade de investigação sobre I/O ainda procura uma abordagem unificada, fácil de utilizar e flexível para resolver os problemas bem conhecidos. Além disso, propõe uma ferramenta experimental denominada IORE (IOR-EXTENDED), que é uma ferramenta aprimorada de avaliação de desempenho de I/O. Além de um gerador de carga de trabalho sintético, o IORE foi projetado para reduzir o tempo de resposta em experimentos e análises de desempenho de I/O. Em seu trabalho utilizam o ambiente de computação Grid'5000, para executar os experimentos relacionados ao IORE de avaliação de desempenho de I/O. A ferramenta se demonstra promissora. Entretanto segundo os autores, observa-se a

oportunidade na ferramenta de recursos adicionais, como a variação do tempo de espera entre solicitações, simulação de clientes assíncronos ou múltiplos usuários e interface para sistemas de armazenamento baseados em objetos.

Igualmente, Da Silva *et al.* (2020) em seu trabalho utilizam o ambiente de computação Grid'5000, para executar experimentos com o objetivo de testar o provisionamento de recursos de armazenamento de forma mais eficiente energeticamente. Sua análise demonstrou que o consumo de energia não está linearmente relacionado à utilização da CPU e que as operações de I/O impactam significativamente o consumo de energia. Inclusive propuseram um modelo de consumo de energia que leva em conta as operações de I/O, que inclui o impacto da espera pela conclusão dessas operações e das execuções simultâneas de tarefas em nodos de computação. Em adição, os autores destacam a necessidade, em trabalhos futuros, de avaliar o custo energético dos nodos de computação que utilizam diferentes tecnologias de armazenamento como HDDs, SSDs, NVMe, por exemplo. Segundo os autores, o impacto de I/O no consumo de energia pode apresentar comportamentos diferentes, ao utilizar diferentes tecnologias de armazenamento.

Em sua pesquisa Delamare e Nussbaum (2021) abordam que experimentos de teste muitas vezes necessitam da coleta de métricas sobre o sistema, serviço ou aplicativo em teste e seu ambiente circundante. Entre os diferentes tipos de métricas, destacam-se as métricas sobre o ambiente experimental, que refletem aspectos que podem influenciar o experimento. Em adição, descreve a proposta Kwollect, que é um serviço projetado e desenvolvido no contexto do Grid'5000, para coletar métricas de infraestrutura e expô-las aos experimentadores. O Kwollect é dimensionado para coleta de métricas, que podem ser personalizadas pelo experimentador para centenas de nodos. A ferramenta Kwollect, em termos de consumo de energia, permitiu uma avaliação parcial de alguns *clusters* nos *sites* Lyon, Grenoble, Nancy. No entanto, esta possui recursos valiosos para coleta de métricas e monitoramento sobre consumo de energia dos nodos de todo o ambiente.

Em adição, às informações obtidas sobre o impacto no consumo de energia dos diferentes tipos de dispositivos de armazenamento dos nodos foram limitadas. Isso se deu, porque a carga de trabalho solicitada era pequena o suficiente para caber na memória principal do sistema, que impediu mensurar o impacto de I/O e desempenho dos dispositivos de armazenamento. Tais informações sobre o impacto

no consumo de energia poderiam ser úteis na escolha de nodos que possuam desempenho de armazenamento semelhante. Estas poderiam funcionar como critério de desempate na escolha dos nodos para experimentos.

3.2 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo foram considerados trabalhos que apresentam abordagens relacionadas à nossa proposta. Estes tratam das características de ambientes de alto desempenho e as configurações de armazenamento nesses ambientes e consideram possíveis melhorias no desempenho de armazenamento nesse ambiente computacional.

O Capítulo 3 aborda de forma exploratória um exemplo de ambiente de alto desempenho, o Grid'5000. Explica sobre estrutura, especificações, serviços e ferramentas disponíveis.

O capítulo seguinte visa apresentar as principais características de um ambiente computacional de alto desempenho, utilizando o Grid'5000 como exemplo. Ao abordar a respeito deste a sua estrutura, especificações, serviços e algumas ferramentas disponíveis, em especial sobre as configurações de armazenamento dos nodos.

4 AMBIENTE EXPERIMENTAL

A computação de alto desempenho tem se tornado mais heterogênea, ao combinar processadores tradicionais com unidades de processamento gráfico em nodos de computação, segundo Nesi *et al.* (2020). Os autores destacam que a análise de desempenho tem se tornado cada vez mais relevante no desenvolvimento e otimização de aplicações paralelas. Porém, a complexidade das aplicações e plataformas computacionais torna esse processo de análise trabalhoso e difícil. De forma semelhante, o armazenamento tem se tornado cada vez mais heterogêneo nos nodos de computação, combinando diferentes tipos e tecnologias. Desta forma, para explorar completamente o *hardware* nesse misto de padrões computacionais, atualmente se carece de ferramentas de análise de desempenho acessível a um público não especializado (Pereira, 2023).

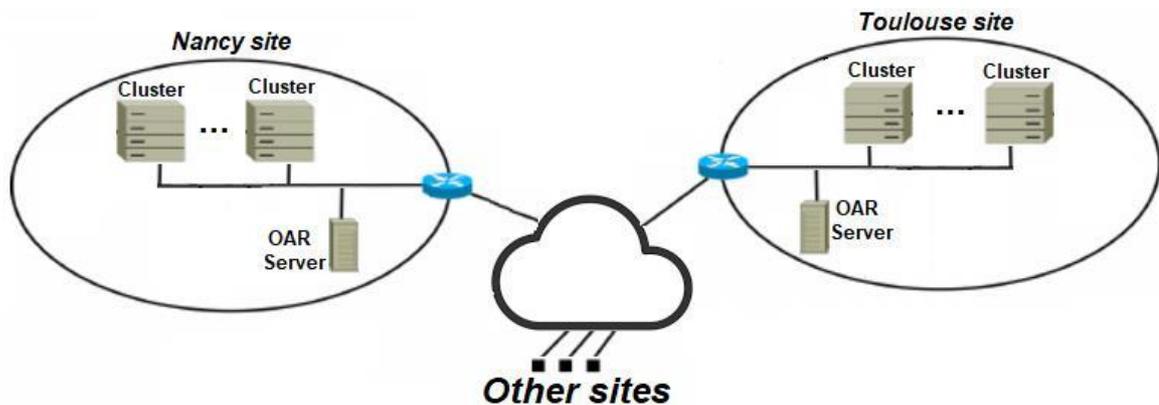
Inclusive, devido à heterogeneidade de recursos da *Grid* Computacional em ambientes HPC, a seleção dos nodos de armazenamento de dados pode afetar o atraso de transmissão da rede. Assim, descobrir os recursos disponíveis e selecionar um subconjunto desses recursos apropriado para a aplicação é muito importante para alcançar alto desempenho ou reduzir o custo dos experimentos (Dong e Akl, 2006).

Por esta razão, neste capítulo abordam-se os principais resultados da fase exploratória deste esforço experimental que apresentou características relevantes de um ambiente computacional de alto desempenho, utilizando o Grid'5000 como exemplo. Além disso, apresentam algumas ferramentas deste ambiente, em especial, as relacionadas com as configurações de armazenamento dos nodos. Para isto, foram realizadas pesquisas adicionais diretamente na documentação oficial do Grid'5000 sobre estrutura, especificações, serviços e ferramentas disponíveis. Isto foi feito com o objetivo de obter, por meio das ferramentas disponíveis, mais informações operacionais sobre as configurações de armazenamento do ambiente de alto desempenho. Destarte, foi possível avaliar as diferenças em relação às configurações de armazenamento do ambiente de alto desempenho. A seção 4.1 aborda esta fase exploratória do ambiente Grid'5000.

4.1 AMBIENTE COMPUTACIONAL HPC GRID'5000

O Grid'5000 é um ambiente de testes flexível e em grande escala para pesquisas baseadas em experimentos em todas as áreas da ciência da computação, com foco em computação paralela e distribuída, que inclui *Cloud*, HPC, *Big Data* e IA. O ambiente possui centenas de servidores, ou nodos, espalhados por nove locais na França e em Luxemburgo. Atualmente estão agrupados em nove locais: Grenoble, Lille, Luxemburgo, Lyon, Nancy, Nantes, Rennes, Sophia e Toulouse. A Figura 3 ilustra a arquitetura dos servidores no ambiente computacional.

Figura 3 – Visão geral da arquitetura do ambiente computacional Grid'5000



Fonte: Adaptado de Khalil *et al.* (2008, p. 7).

No Grid'5000, cada *site* possui um ou mais *clusters*, compostos por um ou mais nodos que são considerados homogêneos entre si em termos de memória e elementos de processamento. Porém, eles são heterogêneos quando comparados com nodos de *clusters* diferentes, sejam eles do mesmo *site* ou de outros. A heterogeneidade é ainda mais evidente entre os dispositivos de armazenamento do ambiente, pois um mesmo nodo pode possuir diferentes tipos de dispositivo de armazenamento (Vasconcelos; Cordeiro e Dufossé, 2022).

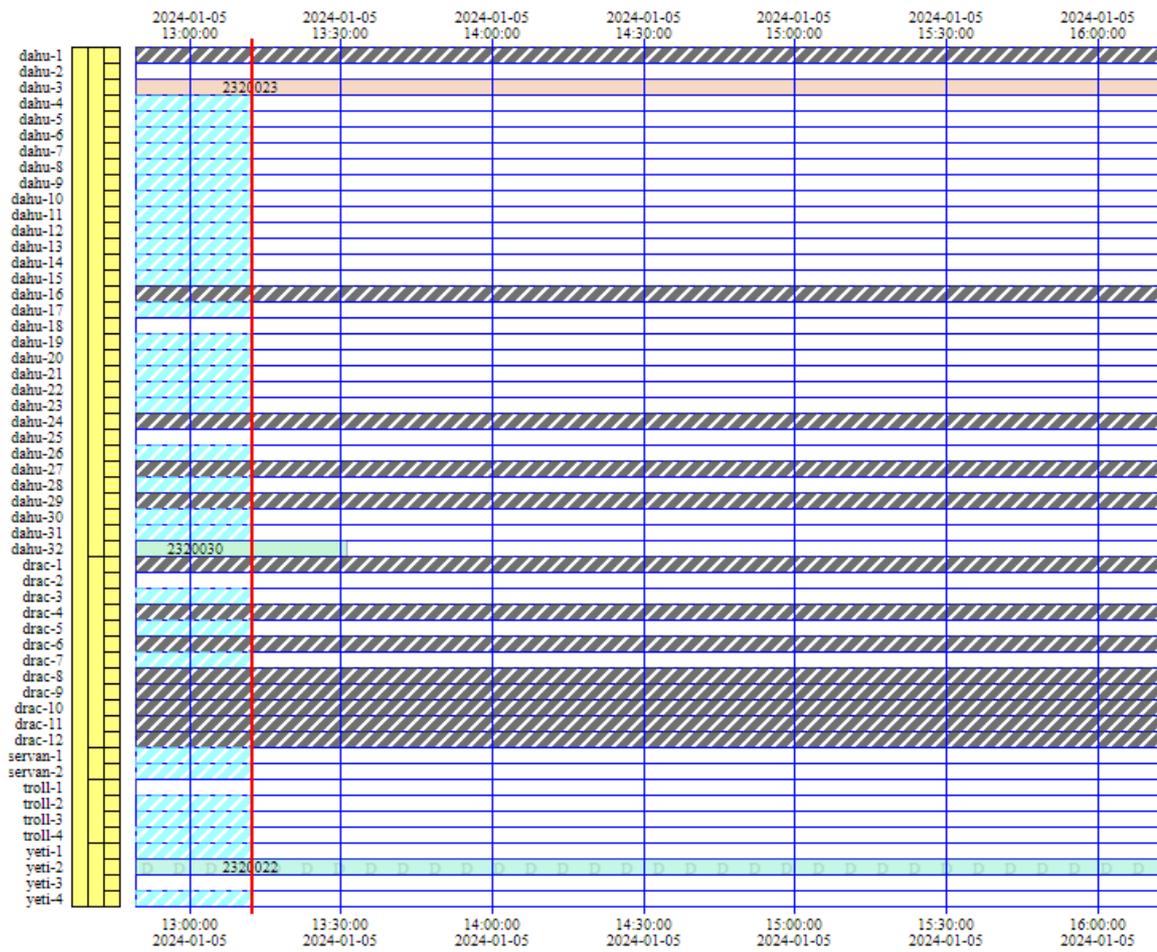
A infraestrutura de controle é baseada no OAR Server, que é um gerenciador de recursos HPC e ferramentas relacionadas. Cada *site* executa uma instância independente do OAR, permitindo reservar recursos que podem variar de um *cluster* inteiro a um único elemento de processamento. Os utilizadores que procuram recursos específicos, podem consultar através da página *web* do Grid'5000 os *clusters* (conjuntos de nodos homogêneos) disponíveis em diferentes *sites* (conjuntos de *clusters*) (Bertot; Nussbaum e Margery, 2020).

4.2 FERRAMENTAS DISPONÍVEIS

A realização de experimentos em sistemas distribuídos geralmente requer instrumentação complexa para automatizar experimentos que envolvem um grande número de máquinas. No Grid'5000, os usuários devem ser cadastrados para ter acesso às ferramentas do ambiente. Geralmente, eles se conectam via SSH aos *front-ends* dos *sites* e utilizam os comandos disponíveis para gerenciar recursos.

No ambiente, dentre as ferramentas amigáveis de interface *Web*, se destaca a Drawgantt que exibe na forma de um gráfico de Gantt os nodos do *site* em tempo real, independente do *cluster*. Ela pode ser acessada em [[https://intranet.grid5000.fr/oar/\[Site\]/drawgantt-svg/](https://intranet.grid5000.fr/oar/[Site]/drawgantt-svg/)], mediante autorização, conforme explicado no parágrafo supracitado. Por exemplo, a Figura 4 ilustra, os nodos do *site* de Grenoble, ao acessar [<https://intranet.grid5000.fr/oar/Grenoble/drawgantt-svg/>].

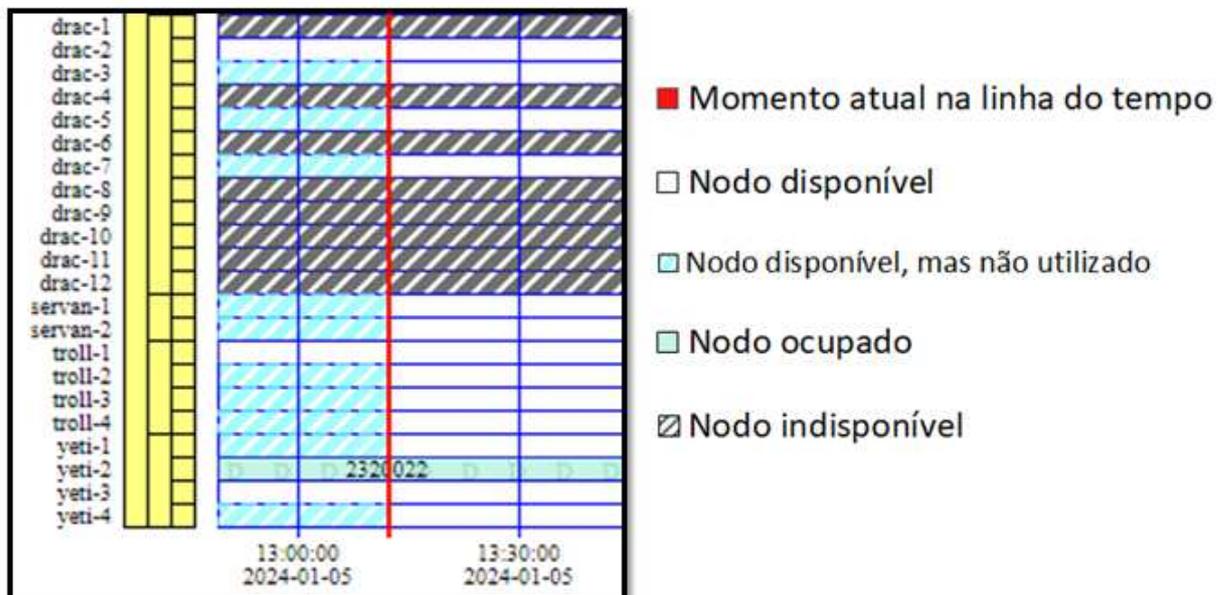
Figura 4 – Nodos de *cluster* do *site* Grenoble



Fonte: Elaborada pelo autor (2024).

A ferramenta Drawgantt representada pela Figura 4 exibe os nodos, disponíveis ou não, de cada *cluster* do *site* solicitado de uma forma rápida e visual por meio de um gráfico de Gantt. O eixo vertical desta ferramenta lista os nodos dos diferentes *clusters* de cada *site*. O eixo horizontal representa o tempo, com marcações de datas e horas. As linhas horizontais indicam diferentes estados dos nodos ao longo do tempo. Para esmiuçar os possíveis estados dos nodos, a Figura 5 aborda detalhes relevantes no gráfico por meio de uma legenda descritiva.

Figura 5 – A indicação da disponibilidade de nodos na ferramenta



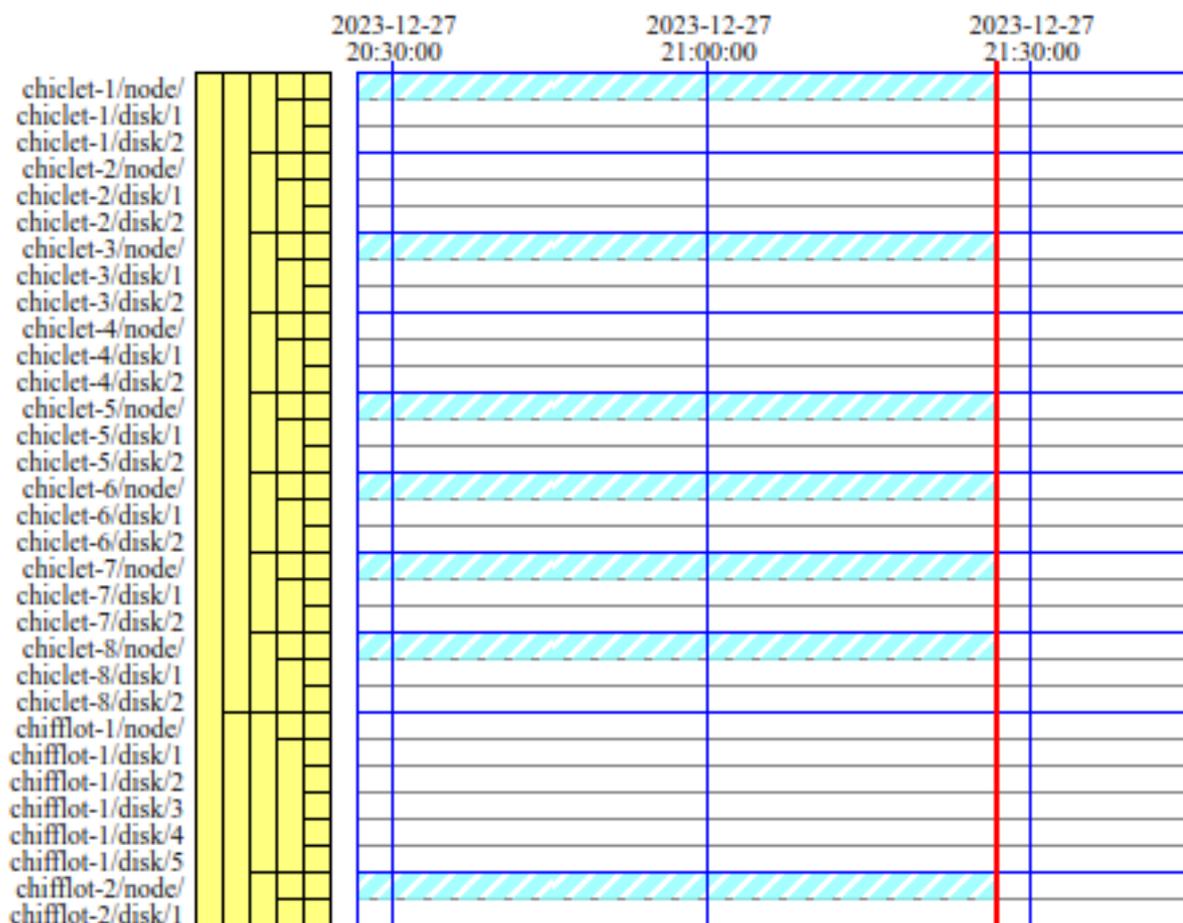
Fonte: Elaborada pelo autor (2024).

Na Figura 5 que é um recorte da Figura 4 representa, por exemplo, que no momento atual, indicado pela linha vermelha vertical, todos os nodos do *cluster* “troll” estão atualmente disponíveis para experimentos. Isto é indicado pelas faixas brancas horizontais à direita da linha vertical vermelha. No entanto, no *cluster* “yeti” apenas três dos quatro nodos estão disponíveis: Yeti-1, Yeti-3, Yeti-4. Assim a ferramenta Drawgantt por exibir de forma rápida e visual a disponibilidade dos nodos, se destaca e atende as demandas dos usuários neste quesito.

4.3 FERRAMENTAS NO CONTEXTO DE ARMAZENAMENTO

Nesta seção, irá se apresentar as principais ferramentas disponíveis no ambiente computacional Grid'5000, relacionadas às configurações de armazenamento do ambiente. A ferramenta que pode ser acessada em [[https://intranet.grid5000.fr/oar/\[Site\]/drawgantt-svg-disks/](https://intranet.grid5000.fr/oar/[Site]/drawgantt-svg-disks/)] é visualmente semelhante à analisada na seção 4.2. Esta exibe por meio de diagramas de Gantt, os nodos e seus respectivos dispositivos de armazenamento. Ela permite, então, identificar visualmente os nodos, que podem ou não ser reservados para experimentos através do OAR Server. Por exemplo, a Figura 6 aborda a interface da ferramenta, ao acessar o endereço [<https://intranet.grid5000.fr/oar/Lille/drawgantt-svg/>].

Figura 6 – Dispositivos de armazenamento de nodos de *cluster* do *site* Lille



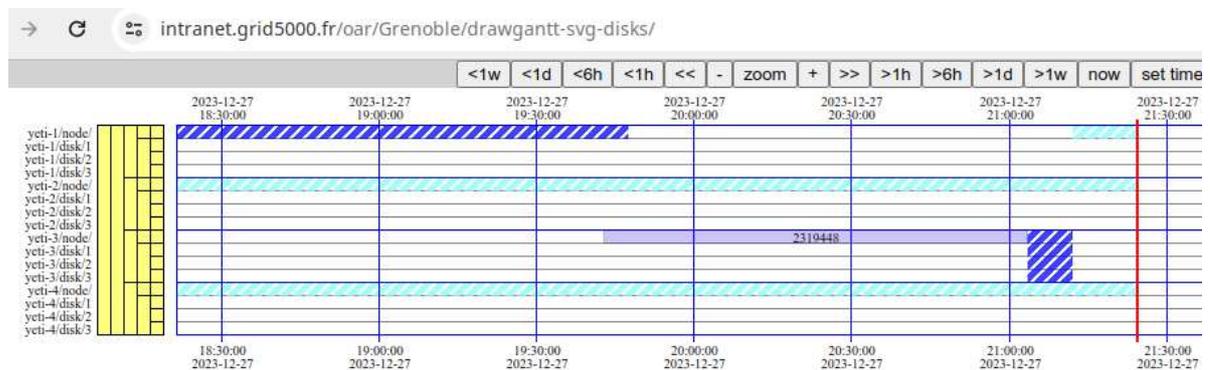
Fonte: Elaborada pelo autor (2024).

A Figura 6, que é um recorte da interface da ferramenta, mostra os nodos dos *clusters* chiclete e chifflot no *site* Lille e seus respectivos dispositivos de

armazenamento no eixo vertical. Semelhante a ferramenta da Figura 4, a linha vermelha vertical representa o momento atual, o eixo horizontal representa o tempo e as linhas horizontais indicam diferentes estados dos nodos ao longo do tempo.

No entanto, ao contrário da ferramenta observada na seção anterior, esta apresenta algumas limitações. Ao consultar determinados *sites*, apresenta-se um número reduzido dos nodos deste e não a sua totalidade. A Figura 7, por exemplo, apresenta o resultado ao consultar os *clusters* do *site* de Grenoble. A figura também mostra que a ferramenta exibe 7,41%, ou seja, quatro do total de 54 nodos no *site* de Grenoble [<https://www.grid5000.fr/w/Grenoble:Hardware>]. Esses nodos pertencem apenas a um dos cinco *clusters* do *site*, sem exibir os demais. Além disso, ao consultar alguns *sites*, nenhum nodo de nenhum *cluster* é exibido.

Figura 7 – Exibição parcial de nodos e seus dispositivos de armazenamento dos existentes no *site* de Grenoble



Fonte: Elaborada pelo autor (2024).

Ao passo que, a Figura 8 exibe a mensagem que aparece quando se tenta obter informações sobre nodos e dispositivos de armazenamento de *clusters* do *site* de Luxemburgo, por meio do endereço [<https://intranet.grid5000.fr/oar/Luxembourg/drawgantt-svg-disks/>].

Figura 8 – Mensagem ao tentar listar discos disponíveis dos nodos de alguns *sites*



Fonte: Elaborada pelo autor (2024).

A Figura 8 apresenta o retorno observado na ferramenta ao pesquisar quatro dos nove *sites* disponíveis do ambiente Grid5000, sendo estes os *sites* Luxembourg, Nantes, Sophia e Toulouse. Estes *sites* não exibem nenhuma informação sobre armazenamento dos nodos dos dispositivos existentes. Verificações adicionais foram realizadas em distintos momentos, no sentido de examinar tal ocorrência e a situação persistiu. Além disso, foram feitas conferências do endereço utilizado no navegador e a priori não foi encontrada nenhuma divergência. Acrescenta-se, que durante as pesquisas não se deparou, na documentação disponível, com a justificativa de tal situação da ferramenta ao pesquisar estes quatro *sites*.

Entretanto, o aspecto observado nas ferramentas e considerado o mais relevante para justificar esta pesquisa, é que nenhuma destas fornece o desempenho dos dispositivos de armazenamento dos nodos. Ao planejar a implementação dos experimentos no ambiente, pode ser relevante considerar essas informações, especialmente em experimentos que utilizam intensamente dispositivos de armazenamento. Estas informações poderiam evitar ou mitigar lacunas de I/O, que podem influenciar no desempenho em experimentos. Assim, nota-se a necessidade de uma abordagem computacional que forneça informações em relação às configurações de armazenamento de todos os nodos do ambiente, a fim de auxiliar na escolha destes em experimentos que demandem um armazenamento diferenciado.

4.4 DESENVOLVIMENTO EXTERNO DE FERRAMENTAS

Uma das características no desenho da estrutura do ambiente Grid'5000, além dos recursos disponíveis, foi a atenção dada à possibilidade de colaboradores externos desenvolverem novos recursos. Visto que tanto a API de referência quanto o OAR fornecem APIs que podem ser utilizadas para automatizar tarefas complexas, de forma a superar lacunas encontradas. Por exemplo, uma deficiência encontrada no Grid'5000 é a sua natureza distribuída: cada *site* tem a sua própria instância do gestor de recursos, o que pode dificultar a reserva de um grande número de recursos em vários *sites* ao mesmo tempo. Então, nesta conjuntura, um usuário do Grid'5000 desenvolveu uma ferramenta chamada Funk que usa a API *Benchmark* e OAR e planejamento de reservas para encontrar recursos que correspondam a uma especificação desejada e um intervalo de tempo que corresponda ao necessário (Margery *et al.*, 2014).

4.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO

O Grid'5000 é um poderoso e flexível ambiente para experimentos em todas as áreas da ciência da computação, com foco em computação paralela e distribuída. Possibilita testes em grande escala para pesquisas, pois o ambiente possui centenas de servidores, ou nodos, espalhados por nove locais denominados *sites*. Cada *site* possui um ou mais *clusters*, compostos por um ou mais nodos.

Porém, os nodos são heterogêneos quando comparados aos de *clusters* diferentes, sejam eles do mesmo *site* ou de outros. Essa diferença é ainda mais evidente em dispositivos de armazenamento, pois durante a fase exploratória foi constatado que um mesmo nodo pode possuir diferentes tipos de dispositivos de armazenamento.

As principais ferramentas disponíveis no ambiente computacional Grid'5000, possibilitam reservar recursos de um *cluster* inteiro a um único elemento de processamento. Além disso, permitem monitorar os nodos de forma amigável por meio de interface *Web* e exibe em tempo real a disponibilidade dos nodos por *site*.

Entretanto, observou-se a inexistência de uma ferramenta que trouxesse o desempenho dos dispositivos de armazenamento dos nodos. Tal fato estimulou o desenvolvimento da ferramenta proposta. Uma das características do Grid'5000 é a

possibilidade de que colaboradores externos possam desenvolver as suas próprias ferramentas. Inclusive, a documentação oficial do ambiente aborda exemplos como a ferramenta Funk.

Portanto, de forma semelhante ao desenvolvimento da ferramenta Funk, foi feito um esforço experimental de uma ferramenta que obtivesse do ambiente e fornecesse aos usuários, de forma diferenciada informações das configurações de armazenamento do ambiente de alto desempenho.

O próximo capítulo mostra o esforço envolvido para desenvolver e testar no ambiente, esta abordagem de *software* para obter e comparar as informações coletadas sobre os dispositivos de armazenamento dos nodos em ambientes de computação distribuídos, semelhantes ao Grid'5000.

5 PROPOSTA

Durante a fase exploratória deste trabalho, abordado no Capítulo 4, foi verificada a heterogeneidade dos nodos, inclusive em armazenamento. Verificou-se nos diferentes *clusters*, que um nodo pode possuir um ou mais dispositivos HDD, ou SSD, ou até ambos dispositivos de armazenamento no mesmo nodo. Esta situação pode dificultar pesquisadores na escolha de quais nodos selecionar para os experimentos. Isto ocorreria especialmente para aqueles que têm menos familiaridade nas especificidades dos dispositivos de memória secundária em ambientes de alto desempenho.

Diante desta situação e da possibilidade de contribuir para melhorias no ambiente HPC em investigação, procurou-se obter de forma mais acessível informações sobre a quantidade, tipo e desempenho dos dispositivos de armazenamento de nodos em ambientes distribuídos paralelos. Para isso, uma proposta desenvolvida, denominada como HPC-SBC (*HPC Storage Benchmark Context*), teve o intuito de produzir uma avaliação de armazenamento no ambiente em estudo no contexto de *cluster*. Esta visa comparar o desempenho dos diferentes dispositivos de memória secundária destes ambientes de alto desempenho, a fim de auxiliar na escolha destes em experimentos que demandem um armazenamento diferenciado.

Na seção a seguir, são apresentadas as vantagens e possíveis utilizações da ferramenta HPC-SBC e das informações obtidas por meio desta ao avaliar os dispositivos de armazenamento em ambientes de alto desempenho.

5.1 VISÃO INICIAL

A finalidade principal da ferramenta HPC-SBC é fornecer aos usuários de forma automatizada as informações sobre o desempenho de cada dispositivo de armazenamento dos nodos computacionais no contexto de *cluster*. Esta pode auxiliar pesquisadores na escolha mais assertiva dos nodos para experimentos computacionais em ambientes HPC. Idem, na escolha dos dispositivos de armazenamento dos nodos, caso haja mais de um dispositivo de memória secundária no mesmo.

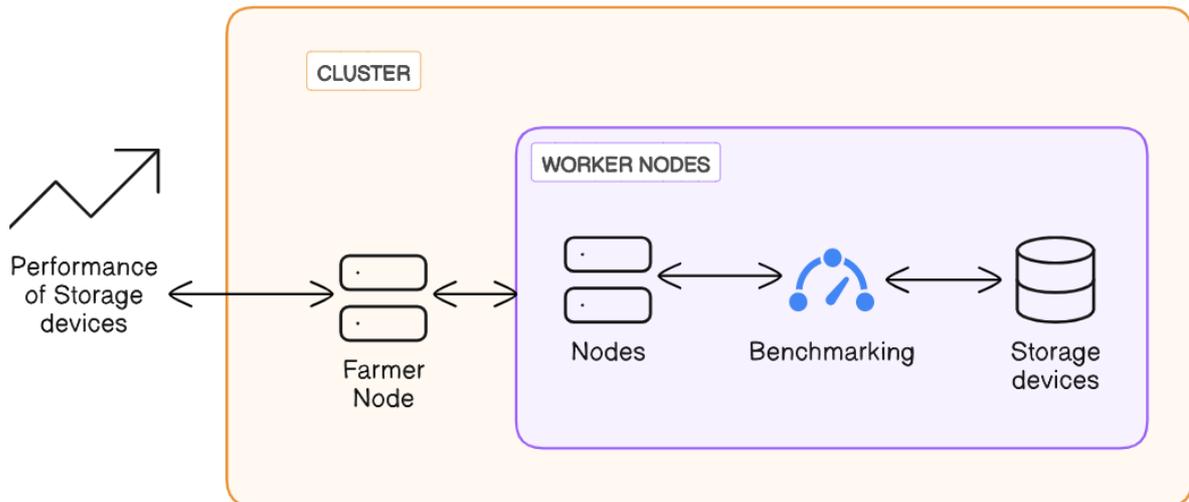
Em adição, o uso da ferramenta HPC-SBC permite aos pesquisadores realizar de forma consciente, experimentos em diferentes configurações de armazenamento. É possível executá-los tanto no melhor, quanto no pior cenário, previamente identificados por meio do HPC-SBC. Assim, ao mensurar em diferentes condições os resultados obtidos, um pesquisador se torna capaz de vislumbrar possíveis otimizações e ajustes necessários para o experimento.

A ferramenta HPC-SBC pode ser de grande ajuda para equipes de infraestrutura computacional, por avaliar com base no desempenho quais dispositivos de armazenamento selecionar, para futuras atualizações e ampliações dos *clusters* de ambientes HPC. Inclusive, ao utilizar em conjunto com ferramentas de eficiência energética como o Kwollect, por exemplo, que foi abordado na seção 3.1 no trabalho Delamare e Nussbaum (2021). A atuação destas duas em conjunto tornaria viável a averiguação de quais dispositivos possuem o melhor custo-benefício, ao considerar desempenho e eficiência energética.

As equipes citadas no parágrafo anterior poderiam ampliar as capacidades das ferramentas de ambientes HPC por meio da integração da ferramenta HPC-SBC nestes. Assim como a ferramenta Drawgantt, de forma visual e nativa, indicar em tempo real o *ranking* de desempenho de armazenamento dos nodos disponíveis no ambiente.

Assim, uma das principais contribuições dessa pesquisa é o desenvolvimento de uma abordagem computacional, cujo objetivo é obter as informações sobre o desempenho de cada dispositivo de armazenamento dos nodos de ambientes de alto desempenho de forma simultânea, mais completa e diferente da que já é feita. A Figura 9 mostra um diagrama de alto nível da solução proposta.

Figura 9 – Diagrama de alto nível da solução proposta



Fonte: Elaborada pelo autor (2024).

A Figura 9 ilustra, por meio de um diagrama simplificado, a solução projetada para monitorar e avaliar o desempenho de dispositivos de armazenamento em ambientes de alto desempenho no contexto de *cluster*. A arquitetura é composta por um nó no papel de *Farmer* e os demais como *Workers*, ambos os papéis apresentados na seção 2.3.2.

O nó designado como *farmer*, é responsável por solicitar ao ambiente HPC nodos no papel de *worker* dentre os disponíveis no ambiente para criar uma rede virtual a partir destes. Ele também envia simultaneamente os comandos de testes de desempenho de armazenamento para a execução simultânea destes em todos os nodos desta rede virtual. Após isto, guarda os resultados para tratá-los, organizá-los e fornecê-los de forma unificada aos usuários.

Os nodos no papel de *worker* são responsáveis por receber os comandos e executar os testes de desempenho em seus dispositivos de armazenamento. Os dados são obtidos por meio de *benchmarking*, onde testes e medições são realizados para avaliar o desempenho dos dispositivos de armazenamento de cada nó. Os resultados obtidos por cada nó *worker* são enviados para o nó designado como *farmer*.

Assim, a finalidade principal da abordagem computacional HPC-SBC, e exemplificada pelo diagrama da proposta, é fornecer aos usuários de forma automatizada as informações sobre o desempenho de cada dispositivo de armazenamento dos nodos computacionais, a fim de auxiliar na escolha destes em

experimentos que demandem um armazenamento diferenciado. Assim, ela possibilita a ampliação das informações fornecidas pelas ferramentas disponíveis em ambientes HPC.

Na seção a seguir, descreve-se em detalhes o desenvolvimento da abordagem computacional proposta, para isso utilizou-se uma linguagem de programação amplamente utilizada no meio científico que pode ser empregada em diferentes ambientes HPC. Inclusive, são descritas as escolhas durante a implementação e as justificativas para tais, com base no conceito de reprodutibilidade.

5.2 IMPLEMENTAÇÃO

No início do desenvolvimento do HPC-SBC foi levado em consideração que o desempenho dos dispositivos de armazenamento pode ser medido em MB/s ou IOPS (operações de entrada/saída por segundo). Dentre as ferramentas e comandos para avaliação de desempenho estão: Fio, dd, hdparm, gnome-disks, rsync e outros.

Além disso, para desenvolvimento no ambiente Grid'5000, segundo Bertot e Nussbaum (2021) estão disponíveis as bibliotecas Execo e EnOSlib para criação de ferramentas para automatização de experimentos. A exemplo, Execo é uma biblioteca *Python* para experimentos, que permite o controle assíncrono de processos locais ou remotos, autônomos ou paralelos (Chardet *et al.*, 2020). Da mesma forma, EnOSlib é uma biblioteca para linguagem de programação *Python* dedicada à pesquisas e experimentos em computação distribuída.

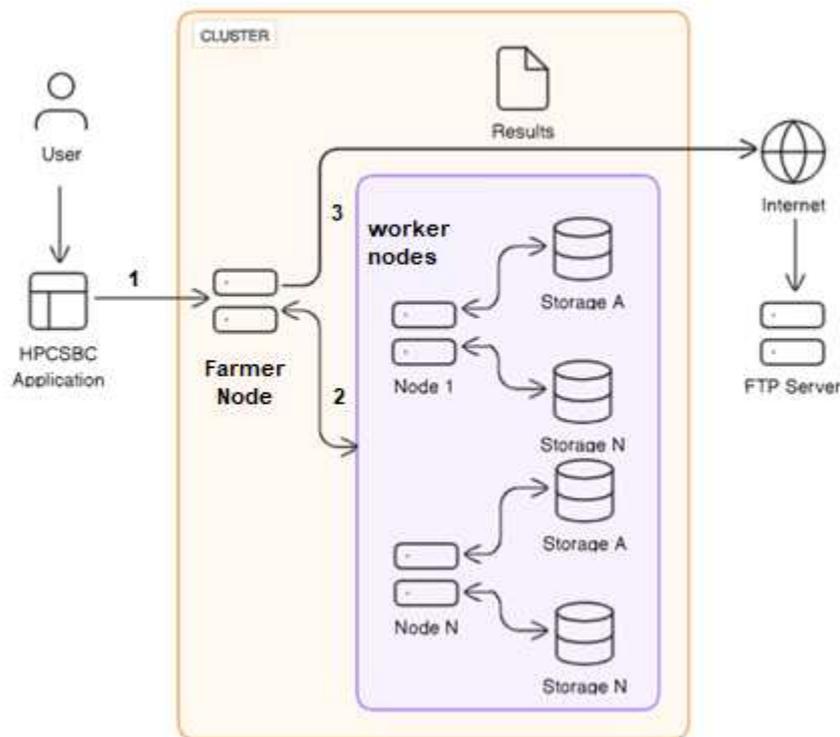
A escolha da biblioteca EnOSlib se justifica tecnicamente, pois possui características relevantes como abstrações para ajudar os pesquisadores a implementar artefatos extensíveis e replicáveis através de provedores para outros ambientes de computação distribuída além do Grid'5000. O provedor é uma função que toma como argumento uma descrição declarativa dos recursos esperados naquele ambiente computacional. Na versão atual, a EnOSlib possui provedores para: Vagrant, Grid'5000, Vmon, OpenStack, Distem e FIT IoT-Lab (Chardet *et al.*, 2022).

A abstração de recursos permite que os desenvolvedores escrevam códigos praticamente independentes de recursos concretos, enquanto os provedores fazem

o trabalho árduo de adquirir recursos do ambiente. Isso permite que as ferramentas sejam executadas em diferentes ambientes computacionais. Em adição, Chardet *et al.* (2022) também destacam que o EnOSlib leva em consideração as melhores práticas de experimentação e aproveita *kits* de ferramentas modernos em sistemas automáticos de implantação e configuração. Com isto, permite-se integrar serviços de última geração que são úteis na maioria das atividades orientadas para experimentos, como modelagem de tráfego de rede, tecnologias de contêineres e monitoramento de pilha, por exemplo.

Para o desenvolvimento do HPC-SBC foi utilizado um poderoso recurso da biblioteca EnOSlib, que é suportar emulação de rede para uso rápido. Desta forma, permite-se criar uma rede virtual em poucos segundos a partir de nodos selecionados. E essa rede virtual pode, então, replicar a execução paralela dos mesmos comandos para todas as máquinas e retornar os resultados de forma unificada. A Figura 10 ilustra o fluxo de trabalho na rede virtual criada pela ferramenta HPC-SBC e como as tarefas são distribuídas e processadas no contexto de *cluster*.

Figura 10 – Visão sistemática dos componentes do *software* HPC-SBC



Fonte: Elaborada pelo autor (2024).

A Figura 10 exemplifica o fluxo de trabalho quando o usuário utiliza a ferramenta HPC-SBC. Esta cria uma rede virtual composta por um nodo *Farmer* e vários *Workers*. O nodo *Farmer* envia comandos de teste de desempenho simultaneamente para todos os nodos *Workers* e armazena os resultados para posterior análise e disponibilização aos usuários. Enquanto isso, os *Workers* recebem os comandos, realizam os testes de desempenho em seus dispositivos de armazenamento e enviam os resultados de volta para o *Farmer*. Os resultados tratados e organizados pelo nodo *Farmer* são então enviados para um servidor FTP. A seguir são descritos detalhes utilizados na implementação da proposta durante a pesquisa e as justificativas para tais escolhas no momento da implementação. Isto foi feito com o objetivo de que outros pesquisadores possam replicar o processo de implementação desta abordagem computacional.

Dentre os comandos a serem executados para avaliação de desempenho, o “*hdparm*” aplica-se, pois é uma ferramenta integrada de testes de I/O nas máquinas Linux, que pode ser ativada via linha de comando. O trabalho de Godinho *et al.* (2023) destaca que “*hdparm*”, além de exibir estatísticas de dispositivos de armazenamento, também se pode utilizar parâmetros relacionados ao *hardware*, o que se mostra útil na análise de diferentes tipos de dispositivos de armazenamento.

Como o ambiente possui dispositivos de armazenamento heterogêneos entre nodos de *clusters* diferentes, foi incluído o comando “*df -h*” para detectar os dispositivos. Este verifica a disponibilidade dos dispositivos conectados nos *data centers* e é citado entre as principais ferramentas e comandos nativos para dispositivos de armazenamento (Huawei, 2022). Além disso, Findlay (2024) destaca em seu trabalho a importância deste comando para identificar os tipos de dispositivos de armazenamento no nodo, a forma como estão montados e a quantidade de espaço utilizado e disponível em cada um.

Ao receber os resultados, o HPC-SBC organiza estes dados e cria um arquivo estruturado com a extensão *.csv, que é nomeado com o nome do *cluster* avaliado e a data e hora da avaliação. A seguir, a ferramenta envia o arquivo com os resultados obtidos, de forma automatizada para uma pasta compartilhada na Internet em um servidor FTP. Ela permite, assim, acesso remoto dos resultados obtidos tanto pelo usuário da ferramenta como por outras partes interessadas. Similar ao trabalho de Duraipandian, Abhishek e Dipsijo (2023), o processo de *upload* de arquivos para o

servidor FTP utiliza a biblioteca *Python* `ftplib` para conectar-se ao servidor FTP e transferir os dados obtidos. Entre as vantagens desta biblioteca, se destacam alguns conceitos implementados nela, por exemplo, a movimentação de conteúdo, documentos emparelhados e tratamento de erros.

A ferramenta HPC-SBC, semelhante a solução de Rifai e Ramadhani (2023), envia o arquivo com os resultados obtidos para o servidor FTP DLPTTest. Entre as vantagens deste é que o mesmo é um servidor público, gratuito e aberto a todos. Além disso, o servidor foi escolhido devido à sua infraestrutura existente, com fácil gerenciamento de arquivos e exclusão automatizada periódica de arquivos. Contudo, caso se deseje alterar o servidor FTP, suas configurações são facilmente identificadas no Código fonte.

Para executar o HPC-SBC, o Grid'5000 oferece a ferramenta de laboratório chamada Jupyter, que fornece uma interface *web* para enviar trabalhos ao Grid'5000 e executar código *Python*. Assim, tal ferramenta permite acompanhar a evolução do experimento durante a fase exploratória enquanto evolui o desenvolvimento do código. O Jupyter funciona em conjunto com o JupyterHub que possui dois *kernels*, um para código `python3` e outro para código `bash`. Ela permite aos usuários, através do comando “`pip`”, instalar qualquer biblioteca *python* que o usuário necessite e disponibilizá-la no ambiente do laboratório virtual (Bertot e Nussbaum, 2021).

Após o desenvolvimento da abordagem computacional proposta, um conjunto de experimentos foi realizado, utilizando a ferramenta desenvolvida no ambiente computacional de alto desempenho Grid'5000, e descritos na seção a seguir.

O código-fonte do HPC-SBC (*HPC Storage Benchmark Context*) está disponível no GitHub (github.com/fernandojf1/hpcsbc) e no Apêndice C. Bem como os resultados obtidos nos testes, que estão descritos no próximo capítulo.

5.3 CONFIGURAÇÃO EXPERIMENTAL

No Grid'5000, além do JupyterHub, a ferramenta HPC-SBC, que é uma ferramenta de linha de comando, pode ser executada remotamente via SSH. Além disso, no ambiente Grid'5000 também através de uma interface *web* que está disponível em [[https://intranet.grid5000.fr/shell/\[site\]/](https://intranet.grid5000.fr/shell/[site]/)]. A ferramenta pode ser executada sem parâmetros, por exemplo: `python hpcsbc.py`. A ferramenta foi

configurada para, por padrão, executar sempre testes em dois nodos do *cluster* neowise que pertence ao *site* Lyon.

Ao utilizar parâmetros de comando na ferramenta, ela consegue atingir todo o seu potencial, por permitir ao usuário avaliar nodos de qualquer outro *cluster* desejado nos *sites* do ambiente. Ela também altera a quantidade de nodos avaliados simultaneamente, até o limite de nodos disponíveis no *cluster*. O número de nodos disponíveis, em tempo real, pode ser obtido pela ferramenta mostrada na Figura 4. O HPC-SBC permite avaliar *clusters* de qualquer *site*, mesmo se o usuário estiver acessando remotamente um nodo de outros *sites* do ambiente.

No conjunto de experimentos utilizando a ferramenta HPC-SBC, foi possível avaliar simultaneamente 31 nodos do *cluster* paravance no *site* de Rennes, avaliando, assim, 21% dos dispositivos de armazenamento do *cluster* em poucos minutos. Isso indica a robustez da ferramenta em diferentes volumes de análise.

Em outro momento, dois outros testes foram executados simultaneamente para avaliar os nodos do *cluster* uvb, por meio de duas máquinas diferentes acessadas remotamente. Cada teste avaliou 10 nodos disponíveis, totalizando 20 nodos de um mesmo *cluster*. Isto indicou a capacidade de utilização da ferramenta de forma simultânea.

Ao final das avaliações, semelhante ao experimento realizado por Nakayama *et al.* (2022), os arquivos de cada experimento foram combinados em um único arquivo por meio do comando Linux “cat”, denominado Results2023.csv. Tanto os arquivos obtidos nos testes, quanto o arquivo resultante da combinação destes, estão disponíveis no GitHub (github.com/fernandojf1/hpcsbc/blob/main/results2023).

5.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO

A proposta apresentada neste trabalho destacou a heterogeneidade dos nodos em ambientes de alto desempenho, com variações significativas nos dispositivos de armazenamento, incluindo HDDs, SSDs, ou uma combinação de ambos. Este cenário pode dificultar a escolha dos nodos apropriados para experimentos, especialmente para pesquisadores menos familiarizados com as especificidades dos dispositivos de memória secundária. Para mitigar este problema e melhorar a acessibilidade das informações sobre dispositivos de armazenamento, foi desenvolvida a ferramenta HPC-SBC (*HPC Storage Benchmark Context*).

A HPC-SBC visa automatizar a coleta de informações sobre o desempenho dos dispositivos de armazenamento dos nodos em ambientes HPC, facilitando a escolha mais assertiva para experimentos que exigem um armazenamento específico. A ferramenta também auxilia na realização consciente de experimentos em diferentes configurações de armazenamento, permitindo a identificação de possíveis otimizações e ajustes necessários. Adicionalmente, a abordagem computacional HPC-SBC pode ser utilizada por equipes de infraestrutura para avaliar quais dispositivos de armazenamento selecionar para futuras atualizações e ampliações de *clusters*, considerando o desempenho e a eficiência energética.

A ferramenta HPC-SBC, utiliza nodos no papel de *Farmer* e *Workers* em uma rede virtual, que permite a execução simultânea de testes de desempenho de armazenamento dos diferentes dispositivos destes nodos. Os resultados são organizados e fornecidos de forma unificada aos usuários, facilitando a análise comparativa. Este desenvolvimento computacional proporciona uma abordagem diferenciada na obtenção de informações sobre o desempenho dos dispositivos de armazenamento, ampliando as capacidades das ferramentas disponíveis em ambientes HPC como o Grid'5000, por exemplo.

O próximo capítulo aborda a análise dos dados obtidos pela ferramenta, como por exemplo, a quantidade, tipo e desempenho dos dispositivos de armazenamento dos nodos dos *clusters*. Faz-se isto comparando-se com os dados fornecidos pelas ferramentas disponíveis no ambiente computacional de alto desempenho, em conjunto com a análise de cobertura dos testes, com o intuito de verificar se o objetivo da proposta foi alcançado.

6 RESULTADOS EXPERIMENTAIS

Neste capítulo, apresenta-se a análise das informações obtidas através do conjunto de experimentos no ambiente computacional HPC denominado Grid'5000 ao utilizar a abordagem computacional proposta HPC-SBC (*HPC Storage Benchmark Context*).

Os dados estão agrupados no arquivo dos testes Results2023.csv, descritos no capítulo anterior. O arquivo foi tratado inicialmente em um aplicativo de texto puro, para fazer a substituição de qualquer ocorrência da palavra “ MB/sec” por “” (vazio). Adicionalmente, foi realizada a substituição de qualquer ocorrência do caractere “.” por “,”. Este segundo passo foi realizado para o arquivo ser compatível com a versão brasileira do aplicativo de planilhas eletrônicas. Então, após esse tratamento dos dados obtidos, o arquivo foi visualizado no Microsoft Excel para uma primeira análise.

6.1 ANÁLISE INICIAL

Durante a análise inicial dos dados obtidos pela ferramenta HPC-SBC, constatou-se que foram avaliados 717 partições dos 144 dispositivos de armazenamento de 137 nodos. Dentre os dispositivos de armazenamento avaliados, um apresentou limite superior (MAX) 36 vezes maior que o dispositivo com limite inferior (MIN).

A Tabela 1 compara o desempenho de dois nodos, uvb-5 e servan-1, localizados em *clusters* de diferentes *sites*. Cada linha da tabela apresenta informações sobre o nome do nodo, o *site* onde está localizado, o dispositivo de armazenamento analisado, a taxa de transferência obtida em megabytes por segundo (MB/s) e uma classificação qualitativa do desempenho. Esta comparação permite uma avaliação direta das diferenças de desempenho entre os dispositivos de armazenamento em termos de velocidade de transferência.

Tabela 1 – Valores de limite inferior (MIN) e limite superior (MAX) dos dispositivos

Nodo	Site	Dispositivo de armazenamento	Taxa de transferência (MB/s)	Desempenho
uvb-5	Sophia	sda3	78,49	Inferior
servan-1	Grenoble	nvme1n1p1	2836,82	Superior

Fonte: Elaborada pelo autor (2024).

Sobre os limites “Inferior (MIN)” e “Superior (MAX)”, descritos tanto na Tabela 1 quanto nas demais tabelas, usou-se a taxa de transferência real obtida pelo dispositivo de armazenamento no conjunto de experimentos ao utilizar a ferramenta HPC-SBC para classificar o desempenho de armazenamento dos dispositivos de memória secundária. Por exemplo dos 137 nodos analisados, um dispositivo de armazenamento localizado no nodo uvb-5 obteve a menor taxa de transferência de 78 MB/s e respectivamente um dispositivo de armazenamento no nodo servan-1 obteve a maior taxa de transferência de 2836 MB/s.

É importante ressaltar que os limites inferior e superior obtidos nas taxas de transferência são influenciados por diversos fatores apresentados na seção 2.3.1, como o tipo de dispositivo de armazenamento (disco rígido, SSD, etc.), a interface de conexão (SATA, PCIe, etc.), além da qualidade dos componentes e a eficiência do sistema de gerenciamento de dados. Assim, uma abordagem computacional como o HPC-SBC é de grande ajuda para avaliar o desempenho real dos dispositivos de armazenamento dos nodos de diferentes *clusters* e *sites* de ambientes HPC.

Adicionalmente, fez-se uma análise mais aprofundada das informações obtidas no conjunto de experimentos, para tentar encontrar e compreender tendências dos fatores citados no parágrafo anterior e a possível influência destes no desempenho dos dispositivos de armazenamento. Tais análises são abordadas a seguir.

6.2 ANÁLISE DOS DISPOSITIVOS DE ARMAZENAMENTO DO AMBIENTE

Após a análise inicial, o arquivo que contém as informações dos experimentos foi importado para um banco de dados para análise mais aprofundada, para utilizar comandos SQL. O uso destes tipos de comandos teve o objetivo de mitigar

possíveis erros humanos no processamento dos dados obtidos. Além disso, os valores foram arredondados para evitar poluição visual.

A inserção neste banco de dados permitiu, entre outros aspectos, inferir, por exemplo, sobre o nodo com a maior diferença entre os limites superior e inferior entre os dispositivos de armazenamento, em diferentes cenários descritos a seguir.

Como mencionando, as taxas de transferência são influenciadas por fatores como o tipo de dispositivo de armazenamento ou a interface de conexão, por exemplo. Assim fez-se necessário decompor as informações obtidas para filtrar e comparar com o objetivo de evidenciar o maior e o menor valor, por tipo de dispositivo.

Dentre os dispositivos de armazenamento avaliados, ao se considerar exclusivamente os do tipo SSD, um apresentou limite superior (MAX) cerca de 14 vezes maior que o dispositivo com limite inferior (MIN), conforme Tabela 2.

Tabela 2 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos SSD

Nodo	Site	Dispositivo de armazenamento	Taxa de transferência (MB/s)	Desempenho
uvb-5	Sophia	sda3	78	Inferior
chiclet-4	Lille	sda3	1073	Superior

Fonte: Elaborada pelo autor (2024).

Dentre os dispositivos de armazenamento avaliados do tipo NVMe, um dispositivo apresentou um valor de desempenho como limite superior (MAX), 94% superior, em comparação com o valor do limite inferior (MIN), conforme Tabela 3.

Tabela 3 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos NVMe

Nodo	Site	Dispositivo de armazenamento	Taxa de transferência (MB/s)	Desempenho
servan-2	Grenoble	nvme1n1p1	1460	Inferior
servan-1	Grenoble	nvme1n1p1	2837	Superior

Fonte: Elaborada pelo autor (2024).

A análise das informações, tanto da Tabela 2 quanto da Tabela 3, indicaram diferenças relevantes de desempenho dos dispositivos de armazenamento quando

agrupados apenas por tipo de dispositivo. Assim sendo, se fez necessário averiguar se em um mesmo nodo, ou seja, com a mesma configuração de *hardware*, ocorrem diferenças de desempenho em dispositivos do mesmo tipo e mensurar esta diferença.

Tabela 4 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos NVMe com maior diferença no mesmo nodo

Nodo	Site	Dispositivo de armazenamento	Taxa de transferência (MB/s)	Desempenho
servan-1	Grenoble	nvme1n1p4	1479	Inferior
servan-1	Grenoble	nvme0n1p4	2773	Superior

Fonte: Elaborada pelo autor (2024).

Esta situação também foi observada em dispositivos do mesmo nodo, ou seja, com a mesma configuração de *hardware* existem diferenças de desempenho entre os dispositivos de armazenamento do tipo NVMe. Na Tabela 4, evidencia que um dispositivo apresentou o valor de desempenho, ou limite superior (MAX), 88% maior, em comparação com o valor classificado como limite inferior (MIN).

Tabela 5 – Valores Limite inferior (MIN) e superior (MAX) dos dispositivos SSD com maior diferença no mesmo nodo

Nodo	Site	Dispositivo de armazenamento	Taxa de transferência (MB/s)	Desempenho
dahu-15	Grenoble	sdb2	189	Inferior
dahu-15	Grenoble	sda5	531	Superior

Fonte: Elaborada pelo autor (2024).

De forma semelhante, entre os nodos apresentaram diferença relevante entre os dispositivos de armazenamento do tipo SSD de um mesmo nodo. A Tabela 5 mostra que no nodo dahu-15, que um dispositivo obteve o valor de desempenho cerca de três vezes maior que o outro, mesmo ambos sendo SSD.

As diferenças relevantes no desempenho entre os dispositivos de armazenamento do mesmo tipo e no mesmo nodo indicam que os testes desempenho não ocorreram na memória principal. Posto que provavelmente os

valores seriam iguais ou muito próximos entre si. Assim, presumivelmente a ferramenta conseguiu mensurar o desempenho dos dispositivos de memória secundária.

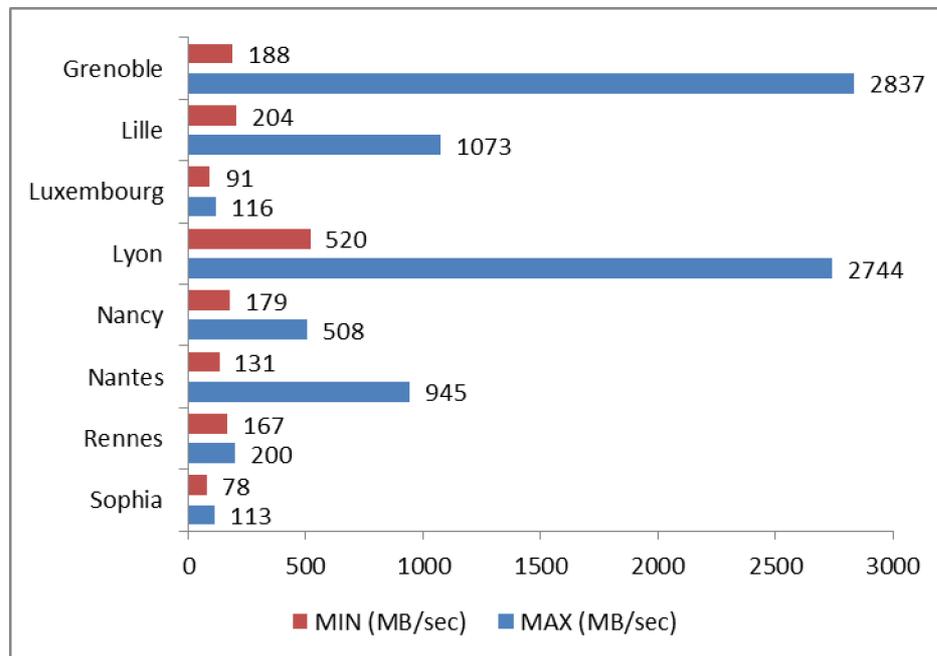
Em adição, tanto as informações da Tabela 2 quanto da Tabela 3, evidenciaram que o desempenho dos dispositivos de armazenamento do mesmo tipo e com a mesma configuração de *hardware* podem variar. Pois outros fatores também podem influenciar no desempenho, como a qualidade dos componentes e outros detalhes no projeto do dispositivo de memória secundária.

Inclusive, esta situação observada pode confundir os pesquisadores na escolha de quais nodos selecionar para os experimentos, especialmente para aqueles que têm menos familiaridade nas especificidades dos dispositivos de armazenamento. Sendo que a ferramenta HPC-SBC pode auxiliar nisto, por automatizar a coleta de informações sobre o desempenho dos dispositivos de armazenamento dos nodos em ambientes HPC. Além de permitir por meio das informações obtidas, comparar o armazenamento dos diferentes *clusters* e *sites* do ambiente, que são abordados na seção a seguir.

6.3 ANÁLISE COMPARATIVA DE ARMAZENAMENTO POR *SITE* E POR *CLUSTER*

Dentre as análises dos dados obtidos pela ferramenta HPC-SBC, a Figura 11 demonstra o gráfico com o levantamento obtido com o limite superior (MAX) e limite inferior (MIN) dos dispositivos de armazenamento por *site*.

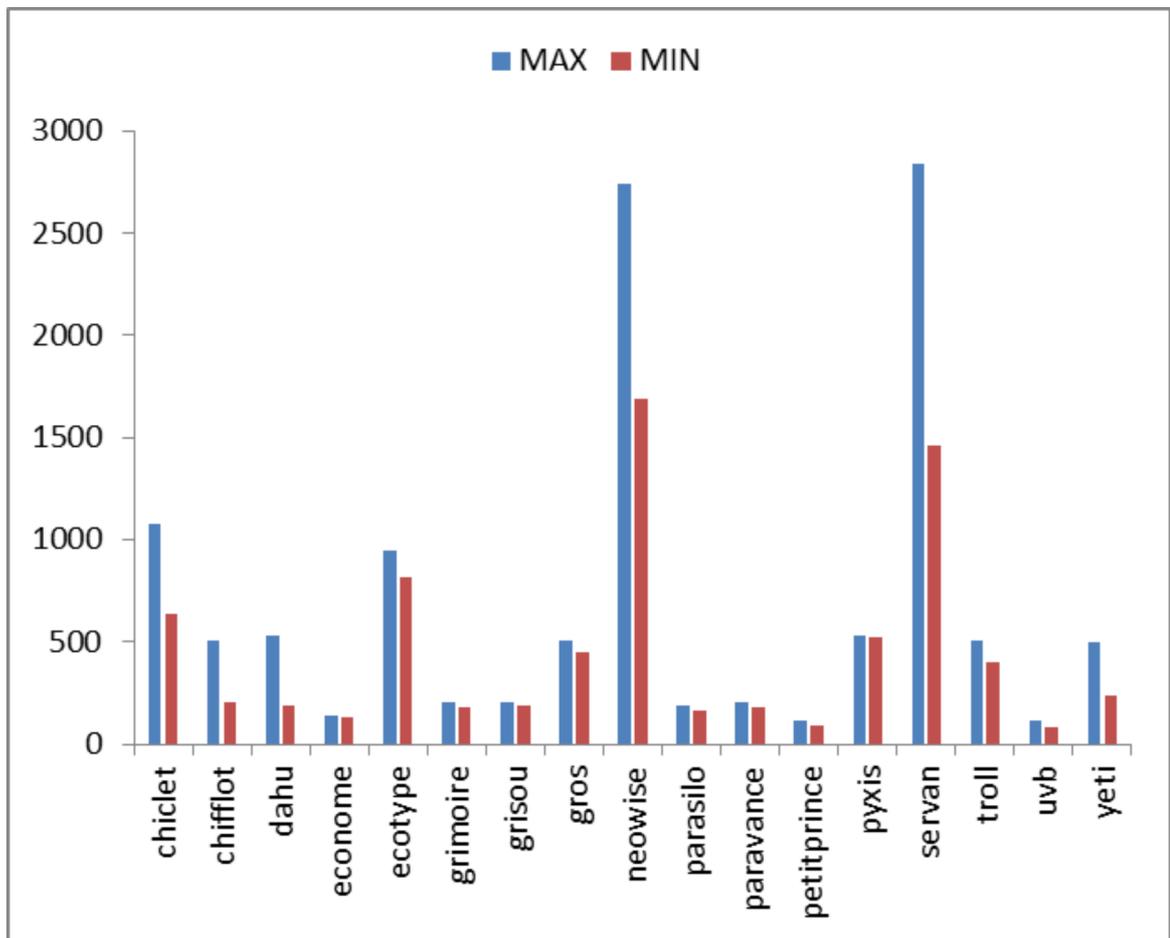
Figura 11 – Gráfico com o limite superior (MAX) e limite inferior (MIN) dos dispositivos de armazenamento por *site*



Fonte: Elaborado pelo autor (2024).

No Gráfico da Figura 11, observa-se que os *sites* Grenoble e Lyon destacam-se no limite superior (MAX). Porém, o limite inferior (MIN) está próximo dos demais *sites*, o que indica a necessidade de procurar em quais *clusters* se apresenta o maior desempenho. Portanto, foi realizada uma análise para obtenção do limite superior (MAX) e limite inferior (MIN) dos dispositivos de armazenamento por *cluster*, conforme a Figura 12 do gráfico que contem essas informações.

Figura 12 – Gráfico com o limite superior (MAX) e limite inferior (MIN) de dispositivos de armazenamento por *cluster*



Fonte: Elaborado pelo autor (2024).

O Gráfico da Figura 12 mostra que o *cluster* neowise do *site* Lyon e o *cluster* servan do *site* Grenoble se destacam em termos dos valores de limite superior (MAX) e limite inferior (MIN) dos dispositivos de armazenamento, quando comparados aos demais *clusters* avaliados. Além disso, o limite inferior (MIN) de ambos é maior que o limite superior (MAX) dos demais *clusters* avaliados.

6.4 ANÁLISE DA COBERTURA DOS TESTES

Por meio da ferramenta HPC-SBC, foram avaliados 144 dispositivos de armazenamento de 137 nodos. Estas informações permitiram, uma análise comparativa para obter o percentual de cobertura dos testes em relação ao *hardware* do ambiente Grid'5000. Ao se comparar com a documentação oficial do

ambiente que está disponível em [<https://www.grid5000.fr/w/Hardware#Summary>], conforme apresentado na Tabela 6.

Tabela 6 – Alcance de cobertura dos testes no ambiente computacional Grid'5000

Estrutura do ambiente Grid'5000	Quantidade	Analisado(s)	Cobertura (%)
<i>sites</i>	9	8	89%
<i>clusters</i>	69	17	25%
nodos	770	137	18%
Dispositivos de armazenamento	1468	144	10%

Fonte: Elaborada pelo autor (2024).

A Tabela 6 mostra de forma sintética, informações relevantes sobre a cobertura dos testes na utilização da ferramenta HPC-SBC no ambiente computacional. Entre elas o alcance da cobertura de todos os *sites*, exceto o *site Toulouse*. Em adição, dos dispositivos de armazenamento de todos os *sites*, aproximadamente metade deles foram analisados.

Todavia, no conjunto de experimentos, a cobertura dos testes dos nodos e *clusters* foi limitada por causa de alguns fatores. Entre eles, a indisponibilidade de alguns nodos e até de *clusters* inteiros, por estarem fora do ar ou sendo utilizados em outros experimentos. Além disso, nodos de alguns *clusters* possuem particularidades quando comparadas com outros, sendo inclusive citadas na documentação oficial do ambiente. Por exemplo, na página de documentação disponível em [https://www.grid5000.fr/w/Production:Getting_Started], menciona que no ambiente alguns *clusters* são classificados com exóticos, pois têm uma arquitetura diferente da x86 ou são específicos o suficiente para serem classificados assim. Inclusive, sendo necessário o uso de comandos adicionais para acessá-los, pois esses recursos computacionais não são selecionados automaticamente pelo OAR, que é o gerenciador de recursos e tarefas para HPC's.

Mesmo com algumas limitações, o esforço experimental alcançou um nível satisfatório em sua proposta alinhado à questão central da pesquisa, com uma maior completude das informações disponíveis sobre os nodos e *clusters* do ambiente em estudo em comparação com informações providas pelas ferramentas disponíveis no Grid'5000. Dessa forma, permite-se uma avaliação dos dispositivos de

armazenamento dos nodos disponíveis de maneira diferenciada para os usuários do ambiente.

Além disso, a revisão da literatura sobre ambientes de alto desempenho possibilitou um melhor entendimento sobre os conceitos e princípios destes, o que permitiu uma melhor compreensão, por meio da pesquisa, deste ambiente computacional não convencional. Da mesma forma, compreender as configurações e dispositivos de armazenamento em ambientes HPC, permitiu criar um arcabouço de critérios para analisar um exemplo real deste ambiente, o ambiente Grid'5000. Pesquisar a documentação e as ferramentas disponíveis no Grid'5000 auxiliou no desenvolvimento das habilidades necessárias para utilizar esse ambiente HPC, o que também pode ser útil a futuros pesquisadores.

Em adição, ao fazer um teste de viabilidade da ferramenta HPC-SBC e ao analisar os dados obtidos, foi possível demonstrar que existem caminhos viáveis de melhoria para as ferramentas já disponíveis do ambiente. Como, por exemplo, fornecer aos usuários mais informações sobre o desempenho dos dispositivos de armazenamento dos nodos, informações estas que podem contribuir na mitigação de lacunas de I/O nos experimentos deste ambiente.

6.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Ao observar, no ambiente de computação de alto desempenho as limitações das ferramentas no contexto de armazenamento, verificou-se a necessidade de desenvolver uma abordagem computacional para contornar essas limitações. A proposta da ferramenta HPC-SBC foi a de obter as informações sobre o desempenho de cada dispositivo de armazenamento dos nodos de ambientes de alto desempenho de forma simultânea e diferenciada.

Neste capítulo foram detalhadas as escolhas técnicas que foram feitas ao adicionar funcionalidades à ferramenta. Isto foi executado, com o objetivo de justificar e explicar os detalhes no desenvolvimento da ferramenta HPC-SBC, considerando tanto o contexto de armazenamento como o ambiente computacional Grid'5000.

Na seção resultados experimentais, foram abordados os resultados obtidos pelo uso da ferramenta HPC-SBC e como estes foram tratados previamente. Além disso, explanou-se por meio de análises, tabelas e gráficos as informações de

desempenho dos dispositivos de armazenamento do ambiente. Também comparou-se os dispositivos de armazenamento de diferentes nodos, *clusters* e *sites*. Deste modo, infere-se que esta abordagem computacional proveu resultados iniciais que indicam o sucesso da pesquisa, ao apontar, entre nodos analisados no ambiente de alto desempenho, aqueles que possuem os dispositivos de armazenamento com o melhor desempenho.

Em adição, o levantamento de informações sobre a cobertura dos testes, utilizando a ferramenta HPC-SBC no ambiente computacional indicou oportunidades futuras. Como por exemplo, implementar na ferramenta, abordagens adicionais para contornar algumas particularidades do ambiente, e dessa forma, ampliar o alcance de cobertura dos testes.

7 CONCLUSÕES E TRABALHOS FUTUROS

O objetivo primário deste trabalho foi o de explorar as possibilidades de uma plataforma de execução de testes de desempenho de armazenamento em nodos computacionais em ambientes HPC no contexto de *cluster*. O primeiro passo, na fase exploratória, foi avaliar as principais ferramentas disponíveis no ambiente computacional Grid'5000. Entre estas ferramentas, investigou-se as que possibilitam monitorar de forma amigável e em tempo real, a disponibilidade dos nodos por *site*. Além das ferramentas do ambiente no contexto de armazenamento, suas características e limitações foram motivadores para o desenvolvimento da abordagem computacional proposta.

O trabalho buscou também demonstrar que a grande quantidade de dados criados, aliada à crescente lacuna entre o poder de processamento e a latência de armazenamento, representa um desafio em termos de transferência e armazenamento de dados.

Assim, nesta dissertação apresentaram-se os desafios e esforços para avaliar o desempenho da memória secundária de nodos no ambiente computacional HPC, como o Grid'5000. Por isto, foi desenvolvida a ferramenta HPC-SBC (*HPC Storage Benchmark Context*), que é um esforço experimental para avaliar o armazenamento em configurações de alto desempenho por meio de uma abordagem de contexto. Com isto, demonstrou-se ser pertinente o questionamento inicial e, também, viável a proposta desenvolvida.

A ferramenta HPC-SBC obteve informações sobre o desempenho de cada dispositivo de armazenamento dos nodos do ambiente de alto desempenho de forma simultânea e diferenciada. Os resultados experimentais obtidos permitiram, graças às análises das informações obtidas através do conjunto de experimentos, avaliar o desempenho dos dispositivos de armazenamento do ambiente HPC, comparando-os entre os diferentes nodos, *clusters* e *sites*. Desta forma, evidenciaram-se quais nodos, *clusters* e *sites* tiveram desempenho de armazenamento superior e inferior aos demais. Em adição, às informações obtidas no conjunto de experimentos utilizando a ferramenta permitiram uma análise comparativa sobre o percentual de cobertura dos testes em relação ao *hardware* do ambiente Grid'5000. Tais informações podem ser relevantes para experimentos com

altas demandas de armazenamento, para mitigar lacunas de I/O. Destarte, os resultados iniciais da ferramenta HPC-SBC indicam o sucesso da pesquisa.

No intuito de se contribuir com o avanço do conhecimento na temática, parte da pesquisa e dos resultados apresentados nesta dissertação, foi reunida em um artigo e que foi aceito para o evento *The 38th International Conference on Advanced Information Networking and Applications (AINA-2024)*.

Como trabalhos futuros, sugere-se que outros ambientes possam ser avaliados. Dessa forma, será possível considerar ajustes na ferramenta HPC-SBC para aumentar o alcance de cobertura desta em diferentes cenários. Exemplos disto seriam a abordagem de outros tipos de dispositivos de armazenamento e a avaliação do desempenho por intermédio de técnicas de paralelismo. Estas últimas permitem utilizar vários dispositivos de armazenamento, que estão virtualmente unidos no mesmo nodo computacional.

Da mesma maneira, pesquisas adicionais poderiam ser conduzidas para desenvolver soluções de *software* de modo a fornecer sugestões de combinações de dispositivos em seus melhores cenários de armazenamento. Inclusive, ao utilizar em conjunto com ferramentas de eficiência energética para averiguar quais dispositivos possuem o melhor custo-benefício, ao considerar tanto o desempenho quanto a eficiência energética dos dispositivos de armazenamento disponíveis no ambiente.

Em adição, durante a pesquisa visualizou-se uma oportunidade de contribuição futura de estender a biblioteca EnOSlib, com os recursos desenvolvidos na abordagem computacional HPC-SBC. Assim como, ampliar as capacidades das ferramentas de ambientes HPC por integrar recursos da ferramenta HPC-SBC, com o objetivo de indicar em tempo real o *ranking* de desempenho de armazenamento dos nodos disponíveis no ambiente de forma visual e nativa.

REFERÊNCIAS

BERTOT, Luke; NUSSBAUM, Lucas. Leveraging Notebooks on Testbeds: the Grid'5000 Case. *In: IEEE CONFERENCE ON COMPUTER COMMUNICATIONS WORKSHOPS IEEE (INFOCOM WKSHPS)*, 34., 2021, Vancouver, Canadá. Vancouver, Canadá: IEEE, 2021, p. 1-6.

BERTOT, Luke; NUSSBAUM, Lucas; MARGERY, David. Implementing SFA Support on an Established HPC-flavored Testbed: Lessons Learned. *In: IEEE CONFERENCE ON COMPUTER COMMUNICATIONS WORKSHOPS (INFOCOM WKSHPS)*, 33., 2020, Toronto, Canadá. Toronto, Canadá: IEEE, 2020, p. 830-835.

CHARDET, Maverick.; COULLON, Helène; PÉREZ, Christian; PERTIN, Dimitri; SERVANTIE, Charlene; ROBILLARD, Simon. Enhancing Separation of Concerns, Parallelism, and Formalism in Distributed Software Deployment with Madeus. *Journal of System and Softwares*, França: Inria, local. 1-28, 2020. No prelo. Disponível em: <https://inria.hal.science/hal-02737859> . Acesso em: 07 jan. 2024.

CHERRUEAU, Ronan-Alexandre; DELAVERGNE, Marie; VAN KEMPEN, Alexandre; LEBRE, Adriene; PERTIN, Dimitri; BALDERRAMA, Javier Rojas; SIMONET, Anthony; SIMONIN, Matthieu. EnOSlib: A Library for Experiment-Driven Research in Distributed *Computing*. *Transactions on Parallel and Distributed Systems*, [S.l.]: IEEE, v. 33, n. 6, p. 1464-1477, 2022. Disponível em: <https://hal.science/hal-03324177/>. Acesso em: 07 jan. 2024.

DA SILVA, R. F., CASANOVA, H., ORGERIE, A. C., TANAKA, R., DEELMAN, E., & SUTER, F. (2020). Characterizing, modeling, and accurately simulating power and energy consumption of i/o-intensive scientific workflows. *Journal of computational science*, 44, 101157.

DANTAS, Mário. **Computação distribuída de alto desempenho: redes, clusters e grids** computacionais. [S.l.]: Axcel books, 2005.

DELAMARE, Simon; NUSSBAUM, Lucas. Kwolect: Metrics collection for experiments at scale. *In: IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2021. p. 1-6.

DONG, Fangpeng; AKL, Selim G. **Technical Report No. 2006-504**. Scheduling algorithms for *grid computing*: State of the art and open problems. Ontario, Canadá: **School of Computing:Queen's University**, 2006. 55 p. Disponível em: <https://research.cs.queensu.ca/home/akl/techreports/GridComputing.pdf>. Acesso em: 09 fev. 2024

DURAIKANDIAN, M.; ABHISHEK, N.; DIPSIJO, M. P.. Security Auding System in Cloud *Computing*. *Journal of Xi'an Shiyou University*, v. 19, n. 4, p. 646-650, 2023.Natural Science Edition.

FINDLAY, Ben. Techniques and methods for obtaining access to *data* protected by linux-based encryption—A reference guide for practitioners. **Forensic Science**

International: Digital Investigation, [S.l.]: Elsevier, v. 48, n. 1, p. 1-12, 2024.

Disponível em:

<https://www.sciencedirect.com/science/article/pii/S2666281723001816>. Acesso em: 07 jan. 2024.

GODINHO, António; ROSADO, José; SÁ, Felipe Alexandre; CARDOSO, Felipe. IoT single board computer to replace a home *server*. In: **IBERIAN CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGIES - CISTI**, 18., 2023, Aveiro, Portugal. Aveiro, Portugal: IEEE, 2023, p. 1-6.

HUAWEI TECHNOLOGIES CO., Ltd. OpenStack. **ICT: Cloud Computing** Technology. Singapore: Springer, 2022. p. 251-294. cap. 6.

INACIO, Eduardo C.; DANTAS, Mario AR. Iore: A flexible and distributed i/o performance evaluation tool for hyperscale storage systems. In: **2018 IEEE Symposium on Computers and Communications (ISCC)**. IEEE, 2018. p. 01026-01031.

INRIA. CNRS. Summary. Grid'5000 **Hardware**. Versão X.X. França: Inria:CNRS, 2023. Disponível em: <https://www.grid5000.fr/w/Hardware#Summary>. Acesso em: 02 jan. 2024.

KHALIL, Fadi.; MIEGEMOLLE, Bernard; MONTEIL, Thierry; AUBERT, Herve; COCCETTI, Fabio; PLANA, Robert. Simulation of micro Electro-Mechanical systems (MEMS) on *grid*. In: **INTERNATIONAL MEETING HIGH PERFORMANCE COMPUTING FOR COMPUTATIONAL SCIENCE - VECPAR**, 8., 2008, Toulouse, França. Toulouse, França: [s.n.], 2008, p. 2095-2100. Disponível em: https://www.researchgate.net/publication/200021517_Simulation_of_Micro_Electro-Mechanical_Systems_MEMS_on_Grid/link/0c960523c0a9ec004600000/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19. Acesso em: 01 fev. 2024.

KYAW, Lett Yi; PHYU, Sabai. Scheduling methods in hpc system. In: **IEEE CONFERENCE ON COMPUTER APPLICATIONS (ICCA)**, n. X, 2020. [S.l.]. [S.l.]: [s.n.], 2020. p. 1-6.

LÖßER, Ansgar; RASKOB, Florian; SCHEUERMANN, Björn. Lazy Read: Asynchronous Execution of Synchronous File I/O. In: **IEEE INTERNATIONAL CONFERENCE ON BIG DATA (BIGDATA)**, n., 2023, Sorrento, Itália. Sorrento, Itália: Editora, 2023. p. 2311-2318. Disponível em: <https://ieeexplore.ieee.org/document/10386429>. Acesso em: 13 fev. 2024

MARGERIE, David; MOREL, Emily; NUSSBAUM, Lucas; RICHARD, Olivier; ROHR, Cyril. Resources Description, Selection, Reservation and Verification on a Large-scale Testbed. In: **INTERNATIONAL CONFERENCE ON TESTBEDS AND RESEARCH INFRASTRUCTURES - TridentCom**, 9., 2014, Guangzhou, China. **Testbeds and Research Infrastructure: Development of Networks and Communities**. Guangzhou, China: Springer International Publishing, maio 2014, p. 239-247.

MEFFE, C.; MUSSI, EO de P.; MELLO, LR de. **Guia de estruturação e administração do ambiente de *cluster* e *grid***. Brasília, DF: Ministério do Orçamento, planejamento e Gestão, 2006. versão 1.0.

NAKAYAMA, H., OHNUKI, H., NAKAHARA, M., NISHIDA-FUKUDA, H., SAKAUE, T., FUKUDA, S., KUSUMOTO, C. Inactivation of axon guidance molecule netrin-1 in human colorectal cancer by an epigenetic mechanism. **Biochemical and Biophysical Research Communications**. 611, p. 146-150, 2022.

NESI, Lucas Leandro; PINTO, Vinicius Garcia; MILETTO, Marcelo Cogo; SCHNORR, Lucas Mello. StarVZ: *Performance Analysis of Task-Based Parallel Applications*. **Journal of Statistical Software**, v. VV, n. II, local. 1-30, 2020.

NEUWIRTH, Sarah Marie. **Accelerating Network Communication and I/O in Scientific High Performance Computing Environments**. 2019. 264 f. Tese (Doutorado em Ciências Naturais) – Combined Faculty of the Natural Sciences and Mathematics, Ruprecht-Karls University. Heidelberg, Alemanha. 2019. Disponível em: https://archiv.ub.uni-heidelberg.de/volltextserver/25757/1/dissertation_sneuwirth_publish.pdf. Acesso em: 13 fev. 2024.

NEWTON, Sir Isaac. [**Carta enviada para Robert Hooke**]. Destinatário: Robert Hooke. Pensilvânia, EUA, 5 fev. 1676. 1 carta.

PEREIRA, Romain. **Efficient Use of Task-based Parallelism in HPC Parallel Applications**. 2023. 129 f. Tese (Doutorado em Informática) – École Doctorale N°512 INFOMATHS Informatique Mathématiques, École normale supérieure de Lyon-ENS:Université de Lyon, França. 2023. Disponível em: https://theses.hal.science/tel-04466797v1/file/PEREIRA_Romain_2023ENSL0097_These.pdf. Acesso em: 13 fev. 2024

PIOLI, Laercio; INACIO, Eduardo C.; MACEDO, Douglas D. J. de; STRÖELE, Victor; DAVID, José Maria N.; MÉHAUT, Jean-François. Characterization research on I/O improvements targeting *DISC* and HPC applications. *In: ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY - IECON*, 46., 2020. Singapura. Singapura: IEEE, 2020, p. 2095-2100. Disponível em: <https://ieeexplore.ieee.org/document/9255087/authors#authors>. Acesso em: 02 fev. 2024.

PIOLI, Laércio; STRÖELE, Victor; DANTAS, Mario Antonio Ribeiro. Research characterization on I/O improvements of storage environments. *In: Advances on P2P, Parallel, Grid, Cloud and Internet Computing: Proceedings of the 14th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2019)* 14. Springer International Publishing, 2020. p. 287-298.

PIOLI, Laercio; STRÖELE, Victor; DANTAS, Mario A.R. An effort to characterize enhancements I/O of *storage* environments. **International Journal of Grid and Utility Computing**, [S.l.]: [s.n.], v. 14, n. 1, p. 51-61, 2023.

RIFAI, Ardian Dwi; RAMADHANI, Erika. Implementation of Logging Feature in Android Payment SDK using Scrum Method. **International Journal of Information System and Technology - IJISTECH**, [S.l.]: [S.n.], v. 5, n. 4, p. 346-353, 2021.

RIVAS-GOMEZ, Sergio. **High-Performance I/O Programming Models for Exascale Computing**. 2019. 135 f. Tese (Doutorado em Ciências da Computação) – School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology. Estocolmo, Suécia, 2019. Disponível em: <https://www.diva-portal.org/smash/get/diva2:1367264/FULLTEXT01.pdf>. Acesso em: 13 fev. 2024.

VASCONCELOS, Miguel Felipe Silva; CORDEIRO, Daniel; DUFOSSÉ, Fanny. Indirect network impact on the energy consumption in multi-clouds for follow-the-renewables approaches. *In*: **INTERNATIONAL CONFERENCE ON SMART CITIES AND GREEN ICT SYSTEMS**, 11., 2022, [S.l.]: [S.l.]: SCITEPRESS-Science and Technology Publications, 2022, p. 44-55.

APÊNDICE A – PUBLICAÇÕES COMO AUTOR PRINCIPAL

SILVA, Fernando de Almeida; DANTAS, Mario Antonio Ribeiro. Um Estudo Experimental de Avaliação de Armazenamento em Memória Secundária com Foco em Alto Desempenho. In: Anais da VII Escola Regional de Alto Desempenho do Rio de Janeiro. SBC, 2021. p. 1-5.

DOI: <https://doi.org/10.5753/eradri.2021.18557>

Resumo: Este trabalho de pesquisa apresenta um esforço experimental que tem por objetivo elaborar e apresentar um estudo com foco na modicidade do custo para transferência e armazenamento de dados em dispositivo de armazenamento secundário, focando na utilização nas camadas de fog, com foco em alto desempenho. Nossos resultados experimentais iniciais ilustram a viabilidade técnica da proposta e o efeito diferencial da mesma, inclusive em computadores que não seriam considerados como ambientes computacionais adequados.

SILVA, Fernando de Almeida; SOUZA SILVEIRA, Walkíria Garcia de; DANTAS, Mario. A Differentiated Approach Based on Edge-Fog-Cloud Environment to Support e-Health on Rural Areas. In: International Conference on Advanced Information Networking and Applications. Cham: Springer International Publishing, 2022. p. 421-432.

DOI: https://doi.org/10.1007/978-3-030-99587-4_36

Abstract: The population aging and the concern to maintain primary health care for the elderly population in rural areas requires a differentiated approach to tackle this challenge. In this article, we propose a model for remote control of elderly people with difficult to access internet considering an Edge-Fog-Cloud structure. In other words, the goal is to provide an appropriate environment with an emphasis on prepare and present a study focusing on low cost of data transferring and storage, utilizing a secondary storage device, focusing on the use of Edge-Fog-Cloud environment, and with a focus on data high-performance *computing*. Our initial experimental results indicate the feasibility of the present proposal and its differential effect, even on computers that initially would not be considered adequate as computational environments.

Qualis: A3

SILVA, Fernando de Almeida; COSTA, João Pedro de Souza Jardim da; DAVID, José Maria N.; DANTAS, Mário A. R. HPC-SBC: An Experimental Effort to Evaluate Storage in High-Performance *Computing* Configurations Using a Context Approach. In: International Conference on Advanced Information Networking and Applications. Cham: Springer International Publishing, 2024. local. 1-12. No prelo.

Abstract: In this article we present the challenges and efforts to evaluate the secondary memory performance of nodes from cluster architectures, for utilization in cluster-like environments in HPC, or data centers, such as the Grid'5000 *computing* environment. The proposal is a tool for general analysis of these experimental environments and including a preliminary study of the tools available to evaluate the performance of the nodes storage. There were some limitations in evaluating tools for secondary memory performance analysis. As a result, the proposal tool, called HPC-SBC (HPC *Storage Benchmark Context*), is presented. The proposal is an experimental effort to evaluate storage in high-performance *computing* configurations adopting a context approach. The initial results indicate the success of the proposal, as it is shown the nodes with the best storage performance.

Qualis: A3

APÊNDICE B – PUBLICAÇÕES COMO CO-AUTOR

COSTA, João Pedro de S. J. da; DANTAS, Mário A. R.; DAVID, José Maria N.; **SILVA, Fernando de Almeida**. Functionalities as a Service - An Approach to Conciliate Interoperability and Data Reduction in E-Health. In: SIMPÓSIO EM SISTEMAS COMPUTACIONAIS DE ALTO DESEMPENHO (WSCAD), 24. , 2023, Porto Alegre/RS. Porto Alegre: Sociedade Brasileira de Computação, 2023 . p. 193-204.

DOI: <https://doi.org/10.5753/wscad.2023.235668>.

Abstract: Interoperability and data reduction have been proven beneficial to health and medical applications that deal with large datasets. Still, the conflicts between these qualities turned out to be a problem for their conciliation. This paper presents an Edge-Fog-Cloud architecture that offers functionalities as a service. These functionalities can guarantee certain qualities depending on the necessities of the client, such as, in our case, interoperability and data reduction. With the use of context simulators, we found that it was possible to significantly increase the output of data delivered to the servers, and decrease the size of the data that transitions in the network and is stored in the servers' databases, without interfering with the syntactic interoperability.

Qualis: B4

APÊNDICE C – HPCSBC (HPC Storage Benchmark Context)

Arquivo: hpcsbc.py

Código-fonte:

```
# -----
#   HPCSBC (HPC Storage Benchmark Context)
#   A tool to evaluate the performance of storage in secondary memory.
# -----
#
# by Fernando de Almeida Silva
#
# https://www.grid5000.fr/w/Hardware#Clusters:
# site = [cluster1, cluster2, ...]
# Grenoble = [dahu, drac, troll, yeti, servan]
# Lille = [chiclet, chifflot]
# Luxembourg = [petitprince]
# Lyon = [neowise, nova, taurus, sagittaire, hercule, pyxis, orion, gemini,
sirius]
# Nancy = [gros, grvingt, grisou, graoully, grappe, grele, graffiti,
grimoire, grimani, graphique, grue, gruss, grouille, grat]
# Nantes = [ecotype, econome]
# Rennes = [paravance, parasilo, roazhon5, roazhon13, paranoia, abacus25,
roazhon6, roazhon8, roazhon9, roazhon10, roazhon12, abacus22, roazhon11,
abacus1, abacus2, abacus3, abacus4, abacus5, abacus8, abacus9, abacus10,
abacus11, abacus12, abacus14, abacus16, abacus17, abacus18, abacus19,
abacus20, abacus21, roazhon1, roazhon2, roazhon4]
# Sophia = [uvb]
# Toulouse = [estats, montcalm]
#
# updates: https://www.grid5000.fr/w/Hardware
# See nodes (authentication required):
https://intranet.grid5000.fr/oar/[site]/drawgantt-svg/

import logging
from pathlib import Path
from datetime import datetime, timezone
from ftplib import FTP

import enoslib as en
import sys
```

```

def send_file_ftp():
    # https://dlptest.com/ftp-test/
    ftp_server = 'ftp.dlptest.com'
    ftp_user = 'dlpuser'
    ftp_password = 'rNrKYTX9g7z3RgJRmxWuGHbeu'

    try:
        # Connect to the FTP server
        with FTP(ftp_server) as ftp:
            ftp.login(ftp_user, ftp_password)

            # Open the local file in binary read mode
            with open(file_name, 'rb') as file:
                # Send the file to the FTP server
                ftp.storbinary(f'STOR {file_name}', file)

            print(f"File '{file_name}' successfully sent to the FTP server.")

    except Exception as e:
        print(f"An error occurred while sending the file: {e}")

def get_zulu_time():
    # Get current time in UTC
    current_time_utc = datetime.now(timezone.utc)
    zulu_time_day = current_time_utc.strftime("%Y-%m-%dT%H:%M:%S.%f")
    return zulu_time_day[:23] + "Z"

en.init_logging(level=logging.INFO)
en.check()

job_name = Path(__file__).name

send = False
hasHDD = False
hasSSD = False
hasNvme = False

cluster_name = "neowise"

```

```

quant_nodes = 2

if len(sys.argv) >= 2:
    cluster_name = sys.argv[1]
    if int(sys.argv[2]) > 0:
        quant_nodes = int(sys.argv[2])

if (quant_nodes) == 1:
    conf = (
        en.G5kConf.from_settings(job_type="exotic", job_name=job_name,
        walltime="0:02:00")
        .add_machine(roles=["compute", "control"],
        cluster=cluster_name.lower(), nodes=1)
    )

if (quant_nodes) >= 2:
    conf = (
        en.G5kConf.from_settings(job_type="exotic", job_name=job_name,
        walltime="0:02:00")
        .add_machine(roles=["compute", "control"],
        cluster=cluster_name.lower(), nodes=1)
        .add_machine(roles=["compute"], cluster=cluster_name.lower(),
        nodes=(quant_nodes-1))
    )

# This will validate the configuration, but not reserve resources yet
provider = en.G5k(conf)

# Get actual resources
roles, networks = provider.init()

zulu_time = get_zulu_time()
file_name = f"{cluster_name}_Results_{zulu_time.replace(':',
'').replace('.', '')}.csv"
content = "node; cluster; storage; message; work load and time; result;
transfer rate;"

with open(file_name, 'w') as file:
    file.write(content)

```

```

# Display Zulu time and number of nodes
print("Zulu Time:", zulu_time)
print("Quant Nodes:", quant_nodes)

# test disk type
results = en.run_command("df -h", roles=roles)
for result in results:
    # host = result.host.replace(result.host, result.host.split('.', 1)[-1])
    host = result.host.replace(".grid5000.fr", "")
    host = host.replace(".grid5000.fr", "")

    # print(result.payload["stdout"])
    if "nvme" in result.payload["stdout"]:
        hasNvme = True
        print(host + " has Nvme!")
    if "/sd" in result.payload["stdout"]:
        hasSSD = True
        print(host + " has SSD!")

if hasSSD:
    # Check SSD performance on all hosts
    results = en.run_command("hdparm -t /dev/sd[a-b][0-9]*", roles=roles)
    for result in results:
        host = result.host.replace(".grid5000.fr", "")
        host = host.replace(".grid5000.fr", "")
        host = host.replace(".", ";")
        send = True

        result_message = result.payload['stdout']
        result_message = result_message.replace(":", ";")
        result_message = result_message.replace("=", ";")
        result_message = result_message.replace(";\n", "")
        result_message = result_message.replace("/sec", "/sec;")
        result_message = result_message.replace("/dev/", host + "; /dev/")
        result_message = result_message.replace("/sec;\n", "/sec;")
        result_message = result_message.replace(" Timing", "; Timing")

    print("")
    print(f"\n{result.host}: \n{result.payload['stdout']}")
    with open(file_name, 'a') as file:

```

```

        file.write(result_message)

if hasNvme:
    # Check NVMe performance on all hosts
    results = en.run_command("hdparm -t /dev/nvme[0-9]*[0-9]*[0-9]*",
roles=roles)
    for result in results:
        host = result.host.replace(".grid5000.fr", "")
        host = host.replace(".grid5000.fr", "")
        host = host.replace(".", ";")
        send = True

        result_message = result.payload['stdout']
        result_message = result_message.replace(":", ";")
        result_message = result_message.replace("=", ";")
        result_message = result_message.replace(";\n", "")
        result_message = result_message.replace("/sec", "/sec;")
        result_message = result_message.replace("/dev/", host + "; /dev/")
        result_message = result_message.replace("/sec;\n", "/sec;")
        result_message = result_message.replace(" Timing", "; Timing")

    print("")
    print(f"\n{result.host}: \n{result.payload['stdout']}")
    with open(file_name, 'a') as file:
        file.write(result_message)

if send:
    # Send the file via FTP
    send_file_ftp()

# Release all Grid'5000 resources
provider.destroy()

```