



VII Reunião Bimestral de Acompanhamento do Compromisso pela Ciência Aberta no 4º Plano de Ação da OGP Brasil

Relato do Marco 7

Carolina Felicíssimo e Leandro Ciuffo



Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



1. Coordenadora: Iara Machado - RNP
2. Equipe:
 - Leandro Ciuffo, Carolina Felicissimo e Gustavo Dias - RNP
 - Washington Segundo - IBICT
 - Rosana Figueiredo - CNPq
3. Data de Início: 01/01/2019
4. Data de Fim: 30/06/2020
5. Avanços no último bimestre (de 15/08/2019 a 17/10/2019)
 - i. *Habemus Dataverse em Português :-))))*
 - ii. FIOCRUZ com seu *Dataverse* no ar!
 - iii. Aproximações: CBPF, LNCC, EMBRAPA, UFRJ, FGV, AN, GDCC
 - iv. Próximos passos

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



Habemus Dataverse em Português :-))))

conta - Root

dadosderede.mp.br/dataverseuser.xhtml?selectTab=dataRelatedToMe

Dataverse

Pesquisa Guia do usuário Suporte Portugues Carolina Felicissimo

Root > conta

Editar conta

Meus dados Notificações Informação da conta Token da API

Aqui estão todos os dataverses, conjuntos de dados e arquivos aos quais você tem acesso. Você pode filtrá-los por status e funções de publicação.

Pesquisar meus dados ... Localizar

- Dataverses
 - Conjuntos de Dados
 - Arquivos
- Status da Publicação
- Publicado
 - Unpublished
 - Esboço
 - In Review
 - Desacordo
- Papéis
- Admin
 - File Downloader
 - Dataverse + Criador de conjunto de dados
 - Criador do Dataverse
 - Criador de conjunto de dados
 - Colaborador
 - Curador
 - Membro

Sorry, nothing was found for these roles: A
Dataverse Creator, Dataset Creator, Contr

GitHub - GlobalDataverseComm...
github.com/GlobalDataverseCommunityConsortium/dataverse-...
Painel - Confluence

dataverse-language-packs

Repository for language files associated with [Dataverse](#)

Some language packages are only available for older versions, to get the latest version view the see notes below).

API Documentation for deploying language files is still under development.

Available languages:

- English (US), latest develop branch maintained by IQSS Harvard
- French (Canada), latest available 4.17 maintained by Bibliothèques Université de Montréal
- French (France), 4.9.4 maintained by Sciences Po
- German (Austria), 4.9.4 maintained by AUSSDA
- Slovenian, 4.9.4 maintained by ADP, Social Science Data Archive
- Swedish, 4.9.4 maintained by SND, Swedish National Data Service
- Ukrainian, 4.9.4 maintained by The Center for Content Analysis
- Spanish, 4.11 maintained by El Consorcio Madroño
- Italian 4.9.4 maintained by Centro Interdipartimentale UniData
- Hungarian, 4.9.4 maintained by TARKI

Desenvolvido por Dataverse

FIOCRUZ com seu *Dataverse* no ar!

Open
Government
Partnership

Conferência Web



conferenciaweb.rnp.br/rec/c5abaf062293beb17a0f0c2e6e3df49690bf3075-1568823800852/presentation

GT em Rede de Dados de Pesquisa Brasileira



Gravado no Conferência Web



!!! Precisamos de Identificadores Persistentes !!!

FIOCRUZ

Metrics 0 Downloads

Contact Share

The root dataverse.

Search this dataverse... Find Advanced Search + Add Data

- Dataverses (23)
- Datasets (0)
- Files (0)

1 to 10 of 23 Results

Sort

Dataverse Category
Organization or Institution (23)

Metadata Source
FIOCRUZ (18)
Root (5)

Publication Status
Unpublished (23)

Biomanguinhos (FIOCRUZ-RNP) Unpublished Sep 19, 2019	
ICTB (FIOCRUZ-RNP) Unpublished Sep 19, 2019	
COC (FIOCRUZ-RNP) Unpublished Sep 19, 2019	
IGM (FIOCRUZ-RNP) Unpublished Sep 19, 2019	
Gereb (FIOCRUZ-RNP) Unpublished Sep 19, 2019	
ENSP (FIOCRUZ-RNP) Unpublished Sep 19, 2019	





Handle.Net® Registry Invoice
Corporation for National Research Initiatives (CNRI)
1895 Preston White Drive, Suite 100
Reston, VA, 20191, US

BILL TO: REDE NACIONAL DE ENSINO E PESQUISA RNP
Attn: Rene Faustino Gabriel Junior
Rua Lauro Muller, 116, sala 1103
BOTAFOGO - Cep: 22290-906
Rio de Janeiro, RJ, 22290-906, Brazil

Date: 03/06/2019
Invoice Number: HNRT-928956
Payment Due: Net 30 days
Payment Currency: US Dollars

Description	Cost Per Year Per Prefix	Total
One Time Registration fee for prefix <u>20.500.12401</u>	\$50.00	\$50.00
Annual Service Fee at \$50 per year for 1 year: 3/06/19 – 3/31/20 for prefix <u>20.500.12401</u>	\$50.00	\$50.00
	Subtotal	\$100.00
	Amount Paid	\$0.00
	Total Amount Due	\$100.00

CNRI accepts the following methods of payment:

- **Checks drawn on US Banks mail to:**
Corporation for National Research Initiatives (CNRI)
Attn: Business Office
1895 Preston White Drive, Suite 100
Reston, VA 20191-5434
[CNRI Federal Tax ID: 52-1447747]
- **Credit Card Payments** - American Express, Visa, and Master Card
- **Wire Payments:** Please use these instructions to wire funds to Access National Bank via the international SWIFT Network.
FIELD 56A INTERMEDIARY INSTITUTION:
Pacific Coast Bankers' Bank
San Francisco, CA 94104
SWIFT CODE: PCBBUS66
ABA Number: 121042484
FIELD 57D ACCOUNT WITH:
Access National Bank
1800 Robert Fulton Dr., Ste 310
Reston, VA 20191
Routing Number: 056009039
FIELD 59 BENEFICIARY CUSTOMER:
Name: Corporation for National Research Initiatives
Address (Line 1): 1895 Preston White Dr., Ste 100
Address (Line 2): Reston, VA 20191
Account Number: 2952950

Medições 2019 Version 3.0

Rede IPÊ, 2019, "Medições 2019", <https://hdl.handle.net/20.500.12401/dv/323>, RNP - Dados Abertos Versão BETA, V3

Description Dados referentes às medições realizadas nos PoP
Subject Computer and Information Science

- Files
- Metadata
- Terms
- Versions

Search this dataset...

7 Files

- Codebook Medições Ipe.pdf
Adobe PDF - 205.1 KB - 17/04/2019 - 18 Downloads
MD5: 1486868932479c60f146c6a8e50bc95a
- medicoesRedelpe_01-2019.tar.gz
Gzip Archive - 517.6 KB - 08/07/2019 - 9 Downloads
MD5: 2117bfff2da802ddee1691bf9f5dd07
Medições da Rede Ipê referentes ao mês de Janeiro de 2019
- medicoesRedelpe_02-2019.tar.gz
Gzip Archive - 450.3 KB - 08/07/2019 - 1 Download
MD5: 0636568cd4e75d496a10fc2e40d9969e
Medições da Rede Ipê referentes ao mês de Fevereiro de 2019
- medicoesRedelpe_03-2019.tar.gz
Gzip Archive - 464.8 KB - 08/07/2019 - 1 Download
MD5: d59b57854f78dcbefd02d596d3bd52aa
Medições da Rede Ipê referentes ao mês de Março de 2019
- medicoesRedelpe_04-2019.tar.gz
Gzip Archive - 491.8 KB - 08/07/2019 - 1 Download
MD5: d8d751f3ec4e0c3b43c8a6f6e6e29565
Medições da Rede Ipê referentes ao mês de Abril de 2019

redelpê Rede Ipê (RNP)

RNP - Dados Abertos Versão BETA > Rede Ipê > Medições 2019 > medicoesRedelpe_02-2019.tar.gz

Metrics 1 Download

Contact Share Download



medicoesRedelpe_02-2019.tar.gz Version 3.0

This file is part of "Medições 2019". If you use this file, please cite the dataset:

Rede IPÊ, 2019, "Medições 2019", <https://hdl.handle.net/20.500.12401/dv/323>, RNP - Dados Abertos Versão BETA, V3; medicoesRedelpe_02-2019.tar.gz [fileName]

Gzip Archive - 450.3 KB - Last Updated: 08/07/2019
MD5: 0636568cd4e75d496a10fc2e40d9969e
Medições da Rede Ipê referentes ao mês de Fevereiro de 2019

- Metadata
- Versions

Export Metadata

File Metadata

File Persistent ID	hdl:20.500.12401/dv/323/5
Download URL	https://dadosabertos.rnp.br/api/access/datafile/?persistentId=persistentId=hdl:20.500.12401/dv/323/5
MD5	0636568cd4e75d496a10fc2e40d9969e
Publication Date	2019-07-08
Size	450.3 KB
Type	Gzip Archive

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)



Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

The screenshot shows a web browser window displaying a PubMed search result. The address bar shows the URL ncbi.nlm.nih.gov/pubmed/31253105. The page header includes the NCBI logo and navigation links for 'Resources' and 'How To'. The main content area features the article title, authors, and a list of affiliations. On the right side, there are sections for 'Full text links' (including BMC and PMC), 'Save items', and 'Similar articles'.

Analysis of isolates from Bangladesh highlights multiple ways to carry resistance genes in *Salmonella* Typhi.

Lima NCB^{1,2}, Tanmoy AM^{3,4,5}, Westeel E⁴, de Almeida LGP², Rajoharison A⁴, Islam M⁵, Endtz HP^{3,4}, Saha SK^{5,6}, de Vasconcelos ATR⁷, Komurian-Pradel F⁴.

Author information

- 1 Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil.
- 2 Laboratório Nacional de Computação Científica, Petrópolis, Brazil.
- 3 Department of Medical Microbiology & Infectious Diseases, Erasmus MC, Rotterdam, the Netherlands.
- 4 Fondation Mérieux - Laboratoire des Pathogènes Emergents, Lyon, France.
- 5 Child Health Research Foundation, Department of Microbiology, Dhaka Shishu Hospital, Dhaka, 1207, Bangladesh.
- 6 Bangladesh Institute of Child Health, Dhaka Shishu Hospital, Dhaka, 1207, Bangladesh.
- 7 Laboratório Nacional de Computação Científica, Petrópolis, Brazil. atr@lncc.br.

Full text links

Read free full text at **BMC**

PMC **FREE** Full text

Save items

Similar articles

Salmonella enterica Serovar Typhi in Bangladesh: E [MBio. 2018]

Antimicrobial resistance and molecule [Int J Med Microbiol. 2017]

Emergence of an Extensively Drug-Resistant *Salmonella* [MBio. 2018]

Review A systematic review of

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

The screenshot shows a web browser window with the address bar displaying "ncbi.nlm.nih.gov/pmc/articles/PMC659...". The page content includes several sections:

- Additional file 2.** (51K, xlsx)
Table S3. Ampicilin, ciprofloxacin, chloramphenicol, ceftriaxone, and cotrimoxazole resistance. (XLSX 50 kb)
- Additional file 3.** (4.1M, jpeg)
Figure S1. Isolate chromosomes were aligned with ProgressiveMauve. Conserved sequence block can be observed as they are marked with the same color. (JPEG 4269 kb)
- Additional file 4.** (16K, xlsx)
Table S1. Sheet NCBI ID: Accession numbers for sequences deposited on NCBI. Sheet Metadata: metadata for the isolates studied. (XLSX 15 kb)
- Additional file 5.** (257K, jpeg)

Vertical text on the left side of the page reads "enomics" and "BMC Genomics".



Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

1	Table S1. Accession numbers for sequences deposited on NCBI.						
2	Sample Name	Chromosome Accession	pHCM1 Accession	pHCM2 Accession	pK91 Accession	pK43 Accession	pPRJEB21992 Accession
3	311189_201186	CP029958	CP029957				
4	311189_202186	CP029956	CP029955				
5	311189_203186	CP029942	CP029941				
6	311189_204186	CP029954	CP029953				
7	311189_205186	CP029950		CP029951			
8	311189_206186	CP029940	CP029939				
9	311189_207186	CP029938	CP029937				
10	311189_208103	CP029932					
11	311189_208186	CP029949	CP029948				
12	311189_209186	CP029952					
13	311189_210186	CP029946	CP029947				
14	311189_211186	CP029945					
15	311189_212186	CP029944	CP029943				
16	311189_213186	CP029936	CP029935				
17	311189_214186	CP029933	CP029934				
18	311189_215186	CP029930	CP029931				
19	311189_216103	CP029928		CP029929			
20	311189_217103	CP029927					
21	311189_217186	CP029646	CP029645	CP029647			
22	311189_218186	CP029925	CP029926				
23	311189_219186	CP029911	CP029910				
24	311189_220186	CP029886	CP029887				
25	311189_221186	CP029883	CP029884				

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

Search: CP029958 - NCBI | Salmonella enterica subsp. enteri

ncbi.nlm.nih.gov/search/all/?term=CP029958

U.S. National Library of Medicine
National Center for Biotechnology Information

Search NCBI CP029958 Search

Search results for: CP029958

NUCLEOTIDE SEQUENCE

Salmonella enterica subsp. enterica serovar Typhi strain 311189_201186 chromosome, complete genome

Salmonella enterica subsp. enterica serovar Typhi
4,783,542 bp genomic sequence
CP029958.1

FASTA

Results by database
Results found in 6 databases

Literature		Genes		Proteins	
Bookshelf	0	Gene	37	Conserved Domains	0
MeSH	0	GEO DataSets	0	Identical Protein Groups	4,162
NLM Catalog	0	GEO Profiles	0	Protein	0
PubMed	0	HomoloGene	0	Protein Clusters	0
PubMed Central	0	PopSet	0	Sparcle	0
				Structure	0

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

The screenshot shows the NCBI GenBank interface for the complete genome of *Salmonella enterica* subsp. *enterica* serovar Typhi strain 311189_201186. The page displays the FASTA format of the genome sequence, which is a long string of nucleotide characters (A, C, G, T). The interface includes a search bar at the top, navigation links, and a sidebar on the right with various tools and information options.

FASTA

Salmonella enterica subsp. enterica serovar Typhi strain 311189_201186 chromosome, complete genome

GenBank: CP029958.1

[GenBank](#) [Graphics](#)

>CP029958.1 *Salmonella enterica* subsp. *enterica* serovar Typhi strain 311189_201186 chromosome, complete genome

AGAGATTACGCTGGTTGCAAGAGATCATAAACAGGGGAAATTGATTGAAAAATAATATATCGCCAGCAGC
ACATGAACAAGTTTCGGAATGTGATCAATTTAAAAATTTATTGACTTAGGGGGGAGATACTTTAACCAA
TATAGGAATACAAGCAGACAAATAAAAAATGACAGAGTACACAACATCCATGAACCGCATCAGCACACC
ACCATTACCACCATCACCATTACCACAGGTAACGGTGGGGCTGACGCGTACAGGAACACAGAAAAAAG
CCCGCACCTGAAACAGTGCGGGCTTTTTTTCGACACAGAGATCACGAGGTAACAACATGCGAGTGTGAA
GTTTCGGCGGTACATAGTGGCAAAATGCAAGACGTTTCTGCGTGTGCGGATATTCGGAAAGCAATTC
AGGCAAGGGCAGGTAGCGACCGTACTTTCCGCCCCCGGAAAAATTACCAACCATCTGGTGGCGATGATT
AAAAAATATCGCGGCCAGGATGCTTTGCCGAATACAGCGATGCCGAACGATTTTTTCTGACCTGCT
CGCAGGACTTGCAGCGCCAGCCGGGATCCCGCTTGCACGGTTGAAAATGTTGTGCAACAAAGAAATC
GCTCAGATCAAAATGTTTTGATGGTATCAGCCTGCTGGGTCACTGCGGATAGCATCAACGCCCGC
TGATTTGCCGTGGCGAAAAATGTCGATCGCGATATAGCGGGACTCTGAGGGCGCTGGACATCGCGT
CACGGTATGATCCGGTAGAAAAATGCTGGCGGTGGGCCATTACCTTGAATCACCCTCGATATCGCG
GAATCGACTCGCCGTATCGCCCGCAGCAGATCCCGCCGATCAGATGATCCTGATGGCGGGCTTTACT
CGGTAATGAAAAGGGTGAACCTGGTGGTGGTGGCGCTAATGGTTCCGACTATCCCGCCCGCTGTCGG
CGCTGTTTACGCGCTGACTGCTGTGAAATCTGGACTGACGTCGATGGCTGTATACCTGTGACCCCGC
CAGGTGCCGGACCCAGGCTGTGAAATCGATGTCTACCAGGAAGCGATGGAGCTCTTACTTCCGGC
CTAAAGTCTTACCCTCGCACCAATAACGCCTATCGCCAGTTCCAGATCCCTGTGATGAAAAATAC
CGGCAATCCGCGAGGCAGCAAGACGCTGATCGCGCGTCCAGCGACGATGATAATCTGCCGGTTAAAGGG
ATCTCTAACCTTAAACAACATGGCGATGTTAGCGTCTCCGGCCGGGAATGAAAGGGATGATGGGATGG
CGGCGCGTGTTCGCCGCCATGTCTCGCGCCGGGATCTGGTGGTGTCTATTACCAAGTCCCTCTGTA
GTACAGCATCAGCTTCTGTGTGCCGACAGTGAAGTGCAGCGCTGCCCGCTGCGATGCAGGATGAGTTC
TATCTGGAGCTGAAAAGGGGGTGTGGAGCCGCTGGCGGTTACGGAGCGGTGGCGATATCTCTGTTG
TCGGGACGGTATGCGCACCTACGCGGCATTTCAGCGAAATCTTTCGCGCGCTGGCGGGGCTAAAT
CAATATCGTGGCGATCGCTCAGGGATCTCTGAGCGTCCATTTCTGTGGTGGTGAATAACGACGATGCC
ACCACCGCGTGGGGTAACGCACCCAGATGCTGTTCAATACCGATCAGGTGATGAAAGTGTGTTGCTATT
GGCTGGCGGGCTGGCGGGCGGCTACTGGAACAGCTTAAACGTCAGCAAACTGGCTGAAGAAACAAGCA
CATCGATCTACGCGTGTGGCGGTGGCAACTCAAAGCGTGTGCTAACCAATGTGCATGGCTGAATCTG
GACAATGGCAGCGGAACTGGCGCAAGCGAACCGCGCTTCAATCTGGGACGTTTAAATCGCCTGGTGA
AAGAATATCATCTACTCAATCCGGTGAATGTTGATTGATGTAACCTCAGTCAGGCGGGCCGACAGTATGC
CGACTCTCTGCGCAAGGGTTCATGTGGTGGCGGCAACAAAGAAAGGCAACCTCTCGATGGAGTACT
TACCATCAGCTACGTTTCGCCCGCGCAATACGGCGCAAAATCTTGTATGACCAACAGCTCGCGCGC
GTTTCCGGTAAATGAAAACCTGCAAAACCTGCTGAATGCGGGTGAATGCAAAAAATTTTCGGCAT
CTTTCCGGTGCCTCTCTTTATTTTCGGTAAACTGGAAGAGGGGATGAGTCTCTCACAGGGCAGCGCT
TGGCGCGGAGATGGGCTATACCGAACCCGATCCGCGCAGCAGCTTTCCGGTATGGATGGGCGGCTA

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

The screenshot shows a web browser window with the OpenAIRE website. The browser's address bar shows the URL: `explore.openaire.eu/search/dataset?datasetId=r36ad727a71a::96dc771f4c4913a5d2ae32144b972ce9`. The website has a dark blue navigation bar with the following menu items: EXPLORE, PROVIDE, CONNECT, MONITOR, DEVELOP. Below this is a secondary navigation bar with: SEARCH, SHARE, LINK, CONTENT PROVIDERS, SIGN IN. The main content area features the OpenAIRE logo (a blue circle with a plus sign) and the word 'EXPLORE' in orange. The title of the dataset is 'Data Supporting "Metadata for Research Data Discovery and Management"'. Below the title are three tags: DATASET, EN, OPEN. The author is listed as 'Jennings, Lizz; (2017)'. The publisher is 'University of Bath'. The related identifier is 'doi: 10.15125/BATH-00374'. The subject is 'RCUK Research Classifications: Information and Knowledge Management'. On the right side, there is a button 'LINK THIS RESEARCH DATA TO...' and a 'Share - Bookmark' section with icons for email, print, Twitter, Facebook, LinkedIn, and a plus sign. Below that is a 'Download from' section with two options: 'University of Bath Research Data Archive via University of Bath Research Data Archive (Dataset, 2017)' and 'University of Bath'. At the bottom right, there is a 'Cite this research data' section with a dropdown menu for 'select a citation style'. A vertical 'feedback' button is on the far right edge of the page.

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

The screenshot shows a web browser window with two tabs. The active tab is titled "Data Supporting 'Metadata for R...". The address bar shows the URL "researchdata.bath.ac.uk/374/". The page header features the University of Bath logo and navigation links: "Home", "About", "Guide", and "Browse". A search bar is located in the top right corner. The main content area displays the title "Data Supporting 'Metadata for Research Data Discovery and Management'" and a description: "This dataset contains the University of Bath research data metadata schema and mappings to other schemas (EPrints, DataCite, Pure, DDI and JSON-LD Google Datasets) used in the Research Data Archive." Below the description, there are "Keywords" (metadata, research data, rdm) and "Subjects" (Information and communication technologies, Information and Knowledge Management). On the right side, there is a "Cite this dataset as:" section with the citation: "Jennings, L., 2017. Data Supporting 'Metadata for Research Data Discovery and Management'. University of Bath Research Data Archive. https://doi.org/10.15125/BATH-00374." Below this is an "Export" section with a dropdown menu set to "ASCII Citation" and an "Export" button. At the bottom right, there is a QR code.



Special Issue Paper

International Journal of
HIGH PERFORMANCE
COMPUTING APPLICATIONS

The International Journal of High
Performance Computing Applications
1-19

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1094342019868526

journals.sagepub.com/home/hpc



I/O performance of the Santos Dumont supercomputer

Jean Luca Bez¹, André Ramos Carneiro², Pablo José Pavan¹,
Valéria Soldera Girelli¹, Francieli Zanon Boito³,
Bruno Alves Fagundes², Carla Osthoff²,
Pedro Leite da Silva Dias⁴, Jean-François Méhaut³
and Philippe OA Navaux¹

Additionally, we are making all the data we collected and analyzed in this article freely available.¹ Traces from

Abstract

In this article, we study the I/O performance of the Santos Dumont supercomputer, since the gap between processing and data access speeds causes many applications to spend a large portion of their execution on I/O on a large scale expensive supercomputer, it is essential to ensure applications achieve the best I/O performance for efficient usage. We monitor a week of the machine's activity and present a detailed study on the workload at providing an understanding of its workload. From experiences with one numerical simulation, we analyze the performance differences between the MPI implementations available to users. We investigated the phenomenon and narrowed it down to collective I/O operations with small request sizes. For these, we concluded that the customized MPI implementation by the machine's vendor (used by more than 20% of the jobs) presents the worst performance. By investigating the issue, we provide information to help improve future MPI-I/O collective write implementations and practical guidelines to help users and steer future system upgrades. Finally, we discuss the challenge of describing applications I/O behavior without depending on information from users. That allows for identifying the application's I/O bottlenecks and proposing ways of improving its I/O performance. We propose a methodology to do so, and use GROMACS, the application with the largest number of jobs in 2017, as a case study.

Keywords

High-performance computing, supercomputer, storage, parallel I/O, workload characterization

Notes

1. <https://gitlab.com/jeanbez/ijhpc-sdumont>

IJHPCA-SDumont

Project

Repository

Files

Commits

Branches

Tags

Contributors

Graph

Compare

Charts

Locked Files

Issues 0

Jean Bez > IJHPCA-SDumont > Repository

master ijhpca-sdumont / data / file-system

 **include CSTOR and OST data**
Jean Bez authored 10 months ago

Name	Last commit
..	
 20170716	include CSTOR and OST data
 20170717	include CSTOR and OST data
 20170718	include CS  cstor-sdumont3002-201707160000.log.bz2
 20170719	include CS  cstor-sdumont3003-201707160000.log.bz2
 20170720	include CS  cstor-sdumont3004-201707160000.log.bz2
 20170721	include CS  cstor-sdumont3005-201707160000.log.bz2
 20170722	include CS  cstor-sdumont3006-201707160000.log.bz2

 cstor-sdumont3002-201707160000.log.bz2

 cstor-sdumont3003-201707160000.log.bz2

 cstor-sdumont3004-201707160000.log.bz2

 cstor-sdumont3005-201707160000.log.bz2

 cstor-sdumont3006-201707160000.log.bz2

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa

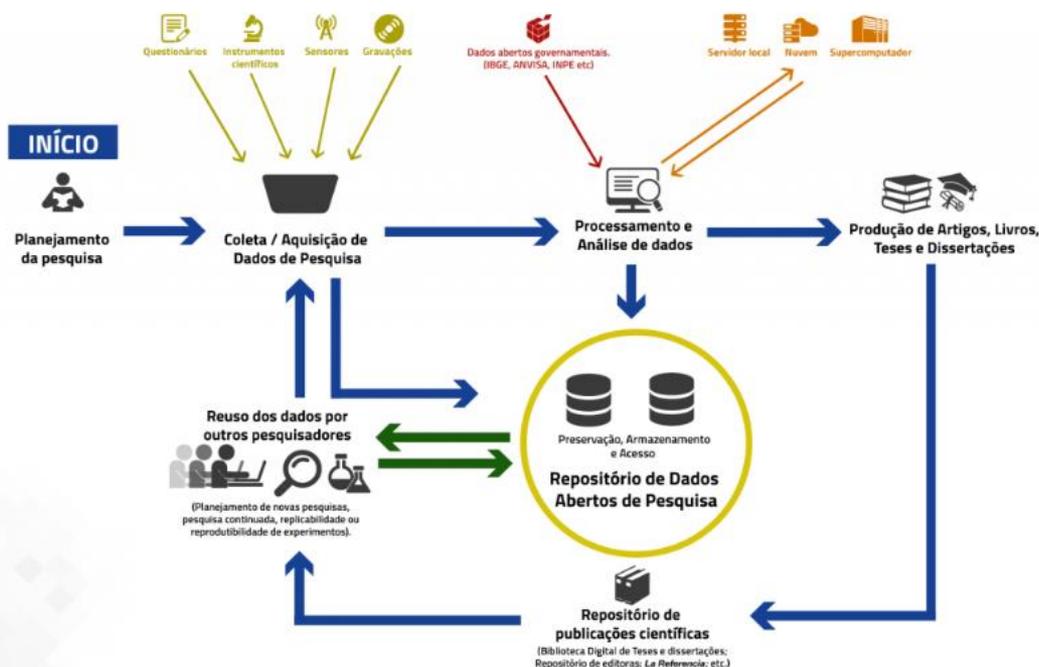


Dados de Pesquisa Abertos Sobre nós Newsletters Dados de pesquisa Contribua Grupos de Trabalho

Dados de pesquisa

Postado por: dadosabertos
Em: março 7, 2018

Na categoria: - No responses



Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação com o LNCC para o estudo de caso do LabInfo (Bioinformática)

Upload a Dataset | IEEE DataPort x +

ieee-dataport.org/submit-dataset

Painel - Confluence

	Submit a Standard Dataset	Submit an Open Access Dataset	Submit a Data Competition
Cost	FREE	\$1,950 (one-time fee) FREE for a limited time Use Promo Code OPENACCESS1 at checkout to upload an Open Access dataset at no cost.	FREE
Access	IEEE DataPort Paid Subscribers	All IEEE DataPort Users	Access Controlled by Data Competition Administrator
DOI	Yes	Yes	Yes
Analysis	IEEE DataPort Users may submit Analysis files	IEEE DataPort Users may submit Analysis files	Competition entries are submitted as Analysis submissions (only accessible to Data Competition Administrator during competition)
File Storage	Unlimited Files up to 2 TB each	Unlimited Files up to 2 TB each	Unlimited Files up to 2 TB each
Cloud Service	<u>Amazon AWS S3</u>	<u>Amazon AWS S3</u>	<u>Amazon AWS S3</u>

f
in
Twitter
Gmail
Email

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



Aproximações:

- **UNICAMP:**
 - Benilton + FAPESP
- **EMBRAPA:**
 - Carla Macario: responsável de TI da EMBRAPA P&D
 - Debora Drucker: Analista de Negócios
- **UFRJ:**
 - Miguel Neto: Chefe de Seção de Informatização das Bibliotecas
- **FGV:**
 - Júlio Chaves: Governança de Infraestrutura de TI para Pesquisas Avançadas
 - Maria do Socorro: Chefe da Seção das Bibliotecas da FGV
- **AN:**
 - Neide de Sordi, Diretora Geral, e Luana Sales, Coordenadora Geral
 - Max: Coordenador de TIC
 - Leandro: Coordenador Geral de Administração
- **CBPF:**
 - Prof. Nilton Alves + Fatima (bibliotecária) + David (TI)
- **LNCC:**
 - Esquadrão da “batalha” CSeC

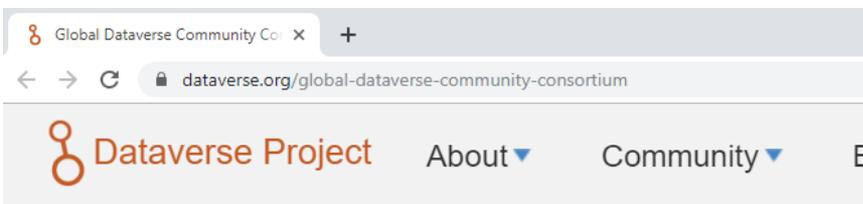


Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Aproximação GDCC:

I ENCONTRO DE GESTÃO E PRESERVAÇÃO DE DADOS DE PESQUISA EM HUMANIDADES



1. Global shared help desk
2. Global team for installation support
3. Global team for data curation support
4. Dataverse Consortium advisory board (this can be a technical advisory board that includes people from outside the Dataverse community; it could also include both a general advisory board and a technical advisory board)
5. Organization of Dataverse Community Meetings in multiple places around the world (in addition to the annual Dataverse Community Meeting at Harvard University)
6. Best practices for leveraging the power of Dataverse, and implementation of policies on data sharing
7. CoreTrust Seal team to investigate trusted repository status
8. Privacy and sensitive data concerns
9. Strategies for metadata sharing or harvesting across installations
10. Collaboration on external tool development and interoperability with Dataverse

Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



**Analistas de Negócios +
Analistas de TI +
Tomadores de Decisão Institucional**

Catálogo



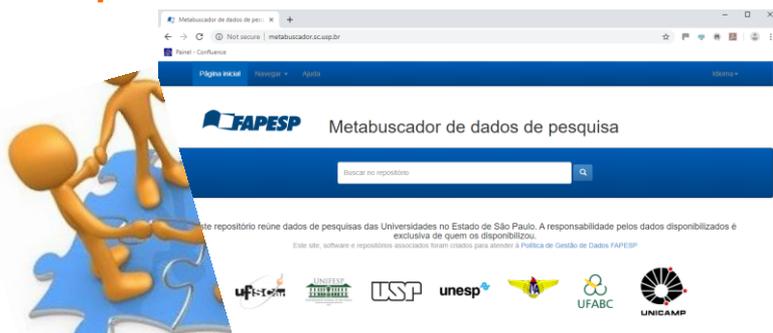
www.digitalbevaring.dk



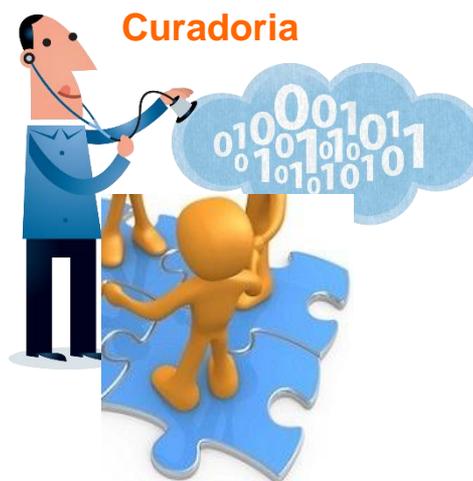
**Armazenamento
Publicação
Preservação**



Compartilhamento



Curadoria



**Identificadores
Persistentes**



Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



- Próximos passos?

→ **Infraestrutura Federada Piloto de Repositórios de Dados de Pesquisa**

→ **Ciberinfraestrutura Federada Nacional de Dados de Pesquisa Científicos**

→ **Ciberinfraestrutura Federada Nacional para Ciência Aberta**



Marco 7: Implantação de infraestrutura federada piloto de repositórios de dados de pesquisa



Referências:

Global Dataverse Community Consortium

<http://dataversecommunity.global/>

<https://dataverse.org/global-dataverse-community-consortium>

GitHub: <https://github.com/GlobalDataverseCommunityConsortium>

Wiki da Ciência Aberta:

<http://url.rnp.br/?OGP-Brasil>

