

Controle de Versionamento:

Versão	Autores	Data
1.0	<p>Universidade Federal do Rio Grande do Sul</p> <ul style="list-style-type: none">• Rafael Port da Rocha• Sônia Elisa Caregnato• Rene Faustino Gabriel Junior• Caterina Groposo Pavão• Samile Andréa de Souza Vanz• Paula Caroline Schifino Jardim Passos• Ana Maria Mielniczuk de Moura <p>Universidade Federal de Rio Grande</p> <ul style="list-style-type: none">• Eduardo Nunes Borges• Adornete de Almeida Martins Junior	12/09/2020

ACESSO ABERTO A DADOS DE PESQUISA NO BRASIL

Armazenamento Elástico no Dataverse - Perguntas Mais
Frequentes (FAQ) - Nível 2



Coordenador

Rafael Port da Rocha (Universidade Federal do Rio Grande do Sul)

Coordenadora adjunta

Sônia Elisa Caregnato (Universidade Federal do Rio Grande do Sul)

Pesquisadores - Universidade Federal do Rio Grande do Sul

Rene Faustino Gabriel Junior

Caterina Groposo Pavão

Samile Andréa de Souza Vanz

Paula Caroline Schifino Jardim Passos

Ana Maria Mielniczuk de Moura

Pesquisadores - Universidade Federal do Rio Grande

Eduardo Nunes Borges

Adornete de Almeida Martins Júnior

Contato

e-mail: dadosdepesquisa@rnp.br - site: <https://dadosdepesquisa.rnp.br>

Rua Ramiro Barcelos, 2777 - Campus Saúde Sala: 106 - Anexo 1 - Brasil -

Porto Alegre - RS - CEP 90.035-007 - Telefone: +55(51)3308.5942



Como citar: MARTINS JUNIOR, Adornete de Almeida; BORGES, Eduardo Nunes; GABRIEL JUNIOR, Rene Faustino; ROCHA, Rafael Port da; VANZ, CAREGNATO, Sônia Elisa; PAVÃO, Caterina Groposo; Samile Andrea de Souza; PASSOS, Paula Caroline Schifino Jardim; MOURA, Ana Maria Mielniczuk de. Acesso aberto a dados de pesquisa no Brasil: Armazenamento Elástico no Dataverse - Perguntas Mais Frequentes (FAQ) - Nível 2.

Resumo

Este relatório apresenta um conjunto de perguntas e erros mais frequentes em relação à configuração do armazenamento elástico no Dataverse.

Palavras-chave: *Dataverse*, Armazenamento Elástico, S3, FAQ..

Sumário

2. PERGUNTAS E RESPOSTAS	5
2.1 Quais serviços de armazenamento elásticos são compatíveis com o Dataverse?	5
2.2 Posso configurar o armazenamento de cada Dataverse lógico para um bucket diferente?	5
2.3 Qual versão da AWS CLI devo utilizar?	5
2.4 Como realizar testes de conexão da AWS CLI com o serviço?	6
2.5 Onde encontro a documentação completa dos comandos da AWS CLI?	6
2.6 Onde posso encontrar mais soluções para problemas de configuração dos serviços de armazenamento S3?	6
2.7 Onde ficam os arquivos de log que podem me ajudar a resolver algum problema?	6
3. ERROS MAIS COMUNS E SUAS SOLUÇÕES	7
3.1 Ao executar a configuração da AWS CLI retorna um erro?	7
3.1.1 “No module named util.ssl_”	7
3.1.2 “RequestTimeTooSkewed”	8
3.2 Erro ao realizar upload de arquivos na Dashboard do Dataverse após configuração do S3?	9
3.2.1 Upload unsuccessful (413: Request Entity Too Large)	9
3.2.2 Upload unsuccessful (500: Internal Server Error)	10
3.2.3 Upload unsuccessful (0: Network Error)	11

1. INTRODUÇÃO

Este documento tem como objetivo apresentar os principais problemas e suas soluções para a equipe de suporte técnico nível 2 durante a configuração do serviço do Dataverse. Relaciona as questões de armazenamento elástico compatível com S3 para grande volume de dados e configurações de segurança do proxy apache.

2. PERGUNTAS E RESPOSTAS

2.1 Quais serviços de armazenamento elásticos são compatíveis com o Dataverse?

O Dataverse utiliza o padrão S3 compatible, logo todos serviços que oferecem esse padrão tendem a ser compatíveis. Durante este projeto foram realizado testes em serviços como AWS S3, Redhat Ceph e Minio.

2.2 Posso configurar o armazenamento de cada Dataverse lógico para um bucket diferente?

Sim! A partir da versão 4.20, o Dataverse suporta múltiplos locais de armazenamento elástico.

2.3 Qual versão da AWS CLI devo utilizar?

A versão 1 é a correta, utilizada durante os testes deste projeto em conjunto com o *Dataverse* 4.20.

2.4 Como realizar testes de conexão da AWS CLI com o serviço?

A CLI possui uma opção `--debug` que retorna uma visão detalhada de cada passo durante a comunicação com o servidor configurado, é de extrema utilidade ao tentar solucionar um problema de autenticação ou conexão.

2.5 Onde encontro a documentação completa dos comandos da AWS CLI?

A Amazon disponibiliza a lista de comandos e como cada um funciona através do link <https://docs.aws.amazon.com/cli/latest/index.html>.

2.6 Onde posso encontrar mais soluções para problemas de configuração dos serviços de armazenamento S3?

Existe um Google Groups da comunidade de usuários do Dataverse que possui diversas publicações de problemas comuns encontrados e possíveis soluções para eles. Pode acessar através do endereço <https://groups.google.com/u/0/g/dataverse-community/>.

2.7 Onde ficam os arquivos de log que podem me ajudar a resolver algum problema?

Passo 1: Revise o log geral do Glassfish

```
$ sudo tail -500 /usr/local/glassfish4/glassfish/domains/domain1/logs/server.log
```

Passo 2: Revise o log de erros do Apache SSL

```
$ sudo tail -100 /etc/httpd/logs/ssl_error_log
```

Passo 3: Caso tenha instalado, revise o log do módulo de segurança do Apache.

```
$ sudo tail -100 /var/log/httpd/modsec_audit.log
```

3. ERROS MAIS COMUNS E SUAS SOLUÇÕES

3.1 Ao executar a configuração da AWS CLI retorna um erro?

3.1.1 “NO MODULE NAMED UTIL.SSL_”

```
[adornete@sp-dadosdepesquisa-beta dataverse-centos]$ sudo -u glassfish aws
configure
Traceback (most recent call last):
  File "/bin/aws", line 19, in <module>
    import awscli.clidriver
  File "/usr/lib/python2.7/site-packages/awscli/clidriver.py", line 17, in <module>
    import botocore.session
      File "/usr/lib/python2.7/site-packages/botocore/session.py", line 28, in
<module>
    import botocore.configloader
      File "/usr/lib/python2.7/site-packages/botocore/configloader.py", line 19, in
<module>
    from botocore.compat import six
      File "/usr/lib/python2.7/site-packages/botocore/compat.py", line 27, in
<module>
    from urllib3 import exceptions
  File "/usr/lib/python2.7/site-packages/urllib3/__init__.py", line 10, in <module>
    from .connectionpool import (
```

```
File "/usr/lib/python2.7/site-packages/urllib3/connectionpool.py", line 31, in
<module>
  from .connection import (
    File "/usr/lib/python2.7/site-packages/urllib3/connection.py", line 45, in
<module>
  from .util.ssl_ import (
ImportError: No module named util.ssl_
```

Solução: Atualização de pacotes do Python

Passo 1: Atualizar módulo pip

```
sudo python -m pip install --upgrade pip
```

Passo 2: Atualizar módulo urllib3

```
$ sudo python -m pip install --upgrade urllib3
```

3.1.2 "REQUESTTIMETOOSKEWED"

```
ClientError: An error occurred (RequestTimeTooSkewed) when calling the
ListObjectsV2 operation: The difference between the request time and the
server's time is too large.
2020-06-30 14:07:32,298 - MainThread - awscli.clidriver - DEBUG - Exiting with
rc 255
```

```
An error occurred (RequestTimeTooSkewed) when calling the ListObjectsV2
operation: The difference between the request time and the server's time is too
large.
```

Solução: Atualizar o horário e o fuso horário de ambos os sistemas, cliente e servidor. Esse erro ocorre quando estes estão fora de sincronia.

Passo 1: Instalar ntpd

```
$ sudo yum install -y ntpd
```


Passo 2: Iniciar o serviço e habilitar para iniciar com o sistema

```
$ sudo systemctl start ntpd  
$ sudo systemctl enable ntpd
```

3.2 Erro ao realizar upload de arquivos na Dashboard do Dataverse após configuração do S3?

3.2.1 UPLOAD UNSUCCESSFUL (413: REQUEST ENTITY TOO LARGE)

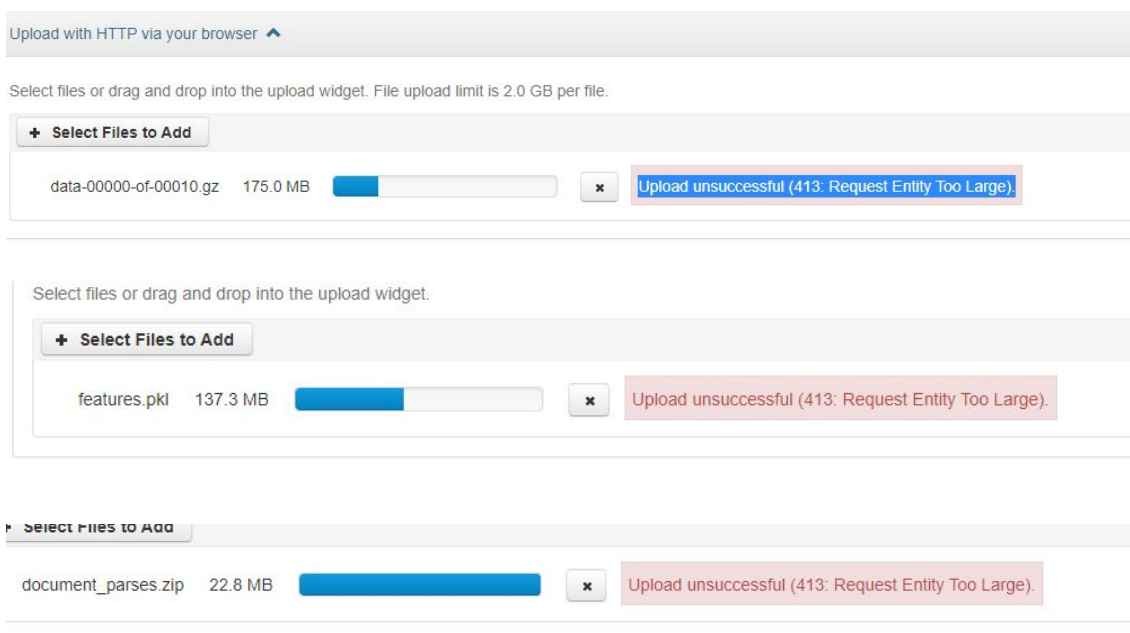


Figura 1 - Request Entity Too Large

Solução: Ajustar limites de upload do módulo de segurança do Apache**Passo 1:** Acessar arquivo de configuração.

```
$ sudo nano /etc/httpd/conf.d/mod_security.conf
```

Passo 2: Altere no trecho abaixo os limites em bytes utilizados pelo Apache, neste exemplo foi utilizado 2GB.

SecRequestBodyLimit **2147483648**
 SecRequestBodyNoFilesLimit **2147483648**
 SecRequestBodyInMemoryLimit **2147483648**

3.2.2 UPLOAD UNSUCCESSFUL (500: INTERNAL SERVER ERROR)

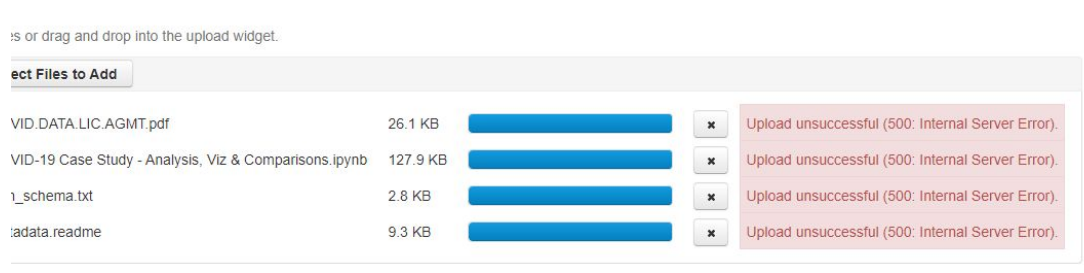


Figura 2 - Internal Server Error

Solução: Ajustar o upload por partes do módulo de segurança do Apache

Passo 1: Acessar arquivo de configuração.

```
$ sudo nano /etc/httpd/conf.d/mod_security.conf
```

Passo 2: Localize as duas linhas abaixo e comente com “#” no início da linha desabilitando a regra *MULTIPART_UNMATCHED_BOUNDARY*.

```
# SecRule MULTIPART_UNMATCHED_BOUNDARY "!@eq 0" \
# "id:'200003',phase:2,t:none,log,deny,status:44,msg:'Multipart parser detected
a possible unmatched boundary.'"
```

3.2.3 UPLOAD UNSUCCESSFUL (0: NETWORK ERROR)

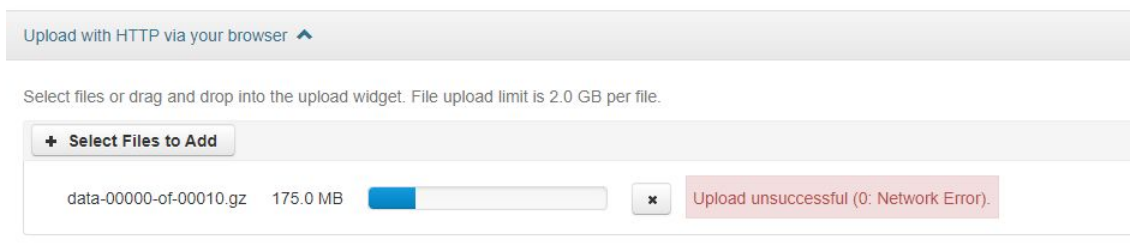


Figura 3 - Network Error

Solução: Ajustar configuração de timeout do proxy do Shibboleth no Apache

Passo 1: Acessar arquivo de configuração.

```
$ sudo nano /etc/httpd/conf.d/ssl.conf
```

Passo 2: Localizar o trecho abaixo e acrescentar a parte em negrito.

```
# pass everything else to Glassfish  
ProxyPass / ajp://localhost:8009/ timeout=6000
```