

DIGERINDO DADOS COM



APACHE NiFi



ELIÉZER ZARPELÃO

CURRÍCULO ACADÊMICO

- x Técnico
 - o Informática
 - ênfase em Sistemas de Informação
- x 2003–2005



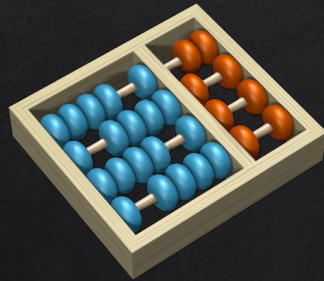
CURRÍCULO ACADÊMICO

- ✕ Graduação
- ✕ Sistemas de Informação
- ✕ 2007-2012



CURRÍCULO ACADÊMICO

- ✕ Especialização
- ✕ Engenharia de Software
- ✕ 2014

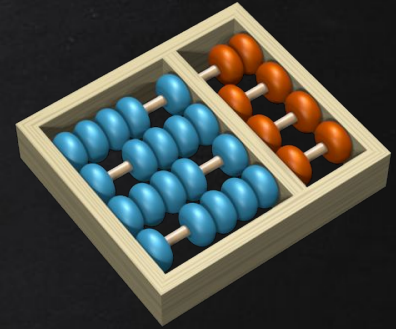


CURRÍCULO ACADÊMICO

- Mestrado (especial)
- Ciência da Computação
- Futebol + Machine Learning



FEF



PROFISSIONAL

- ✕ Manager Tech Architect – Avanade (Habemus Vagas)
- ✕ Arquiteto de Soluções – ZarpSystem
- ✕ Analista de Sistemas – FEM – Unicamp
- ✕ Docente/Coordenador de Curso – IESCAMP
- ✕ Docente – UNAERP
- ✕ Tecnologia desde 2005 (1996)

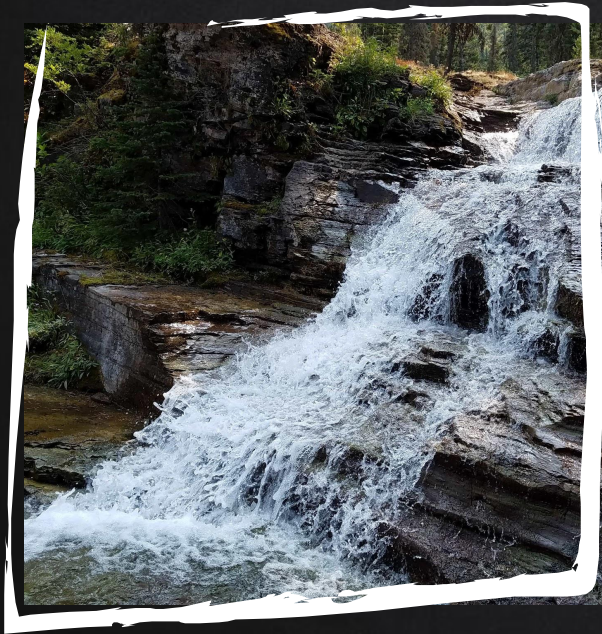




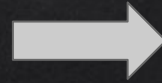
O que são? Onde vivem? O que comem?



DATAFLOW



Mover conteúdo de “A” para “B”
Logs, HTTP, XML, CSV, Imagens,
Vídeos, etc e etc



SUAVE... #SQN

HOW STANDARDS PROLIFERATE:

(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION:
THERE ARE
14 COMPETING
STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.



SOON:

SITUATION:
THERE ARE
15 COMPETING
STANDARDS.



DESAFIOS

Dados

Padrões

Formatos

Protocolos

Veracidade

Validade

Schemas

Particionamento

Infraestrutura

Entrega “Exactly Once”

Segurança

Gestão de acesso

Rede

Pessoas

Compliance

Pessoa / Time / Grupo

Mudança nos consumidores

Alterações de requisitos

Entrega “Exactly Once”



“Provides a way to prioritize data flows more effectively and get rid of artificial delays in identifying and transmitting critical information”

Joseph L. Witt



APACHE NIFI

- ✗ Plataforma open-source para construção e gestão de fluxos de dados (low-code)
- ✗ Escalável, segura e tolerante a falhas
- ✗ Interface Web e API para gestão e monitoramento dos fluxos
- ✗ Fluxos de dados podem ser alterados em tempo de execução

FLOW FILE

- ✗ Objeto que se move através do fluxo
- ✗ Coração do NiFi
- ✗ Atributos (chave/valor)
- ✗ Apontador para conteúdo

HTTP HEADER

```
Request URL: https://www.google.com/  
Request Method: GET  
Remote Address: 172.217.29.164:443  
Status Code: 200 OK  
Version: HTTP/2
```

HTTP PAYLOAD

```
<!doctype html><html itemscope="" itemtype="http://schema.org/WebPage" lang="en-BR"><head><meta  
document.documentElement.addEventListener("submit",function(b){var a;if(a=b.target){var c=a.getE  
var e=window.performance;var g=function(a,b,c,d){a.addEventListener?a.removeEventListener(b,c,d)  
var b=[];function(){google.tick&google.tick("load","dcl")};google.dcl=function(a){b.length?b.pu  
var b=[];google.jsc={xx:b,x:function(a){b.push(a)},mm:[]};m=function(a){google.jsc.mm.length||(g  
var f=this||self,h=Date.now;  
  
var x={};var aa=function(a,c){if(null==c)return!1;if("contains"in a&&1==c.nodeType)return a.co  
var I=function(){this.h=this.a=null;K=function(a,c){var d=J;d.a=a;d.h=c;return d};I.prototype.g  
var O=function(){this.g=[];this.a=[];this.g=[];this.j=[];this.h=null;this.i=[];N(this," custom")  
4=b.button[b.shiftKey]?e:"clickmod";else{var lb,which[b.keyCode];ca663=l&&(l=13);if(l3)=l&&  
e.tagName).toLowerCase(),(e=32==(b.which[b.keyCode]&&"CHECKBOX")?e):(e=(b.n.e.tagName,tolUpp  
m)||"mouseout"==b.type&&"mouseleave"==m)||e&&(e===k||aa(k,e))}l.action="",l.actionElement=null;e  
window.rvt=function(){return!0;}}.call(this);(function(){(function(){window.jsarvt=function(){return!1;}}).  
var a=window.location,b=a.href.indexOf("#");if(0<=b){var c=a.href.substring(b+1);/(^|&)q=/.test(  
c)}break a;c=c[e];a=a[a.length-1];d=c[a];b=b(d);b!=d&&null!=b&&h(c,a,{configurable:!0,writable:  
try{  
/*
```

FlowFile

DETAILS

ATTRIBUTES

Attribute Values

```
atributo1  
valor-atributo-1  
  
atributo2  
valor-atributo-2  
  
filename  
0fd809ac-69ee-4011-9d99-ea5ae68e97fe  
  
path  
./  
  
uuid  
0fd809ac-69ee-4011-9d99-ea5ae68e97fe
```

View as: original

Filename: 0fd809ac-69ee-4011-9d99-ea5ae68e97fe
Content Type: text/plain

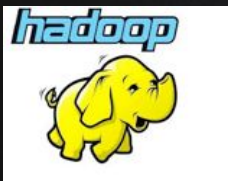
1	CONTEUDO DO FLOW FILE
---	-----------------------

PROCESSOR

- ✗ Executa um trabalho
- ✗ Roteamento, transformação...
- ✗ Mais de 200 componentes...

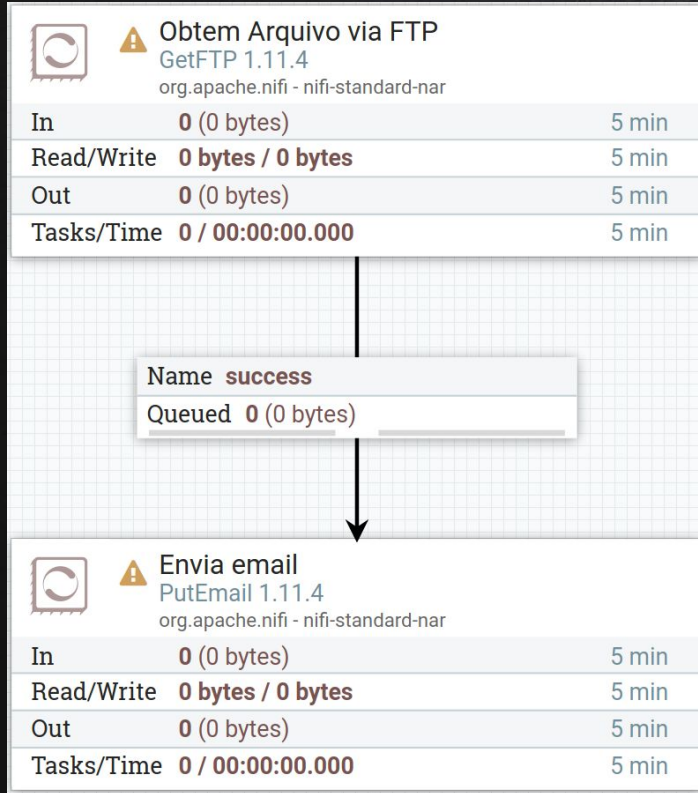
  Obtem Arquivo via FTP GetFTP 1.11.4 org.apache.nifi - nifi-standard-nar		
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

PROCESSORS



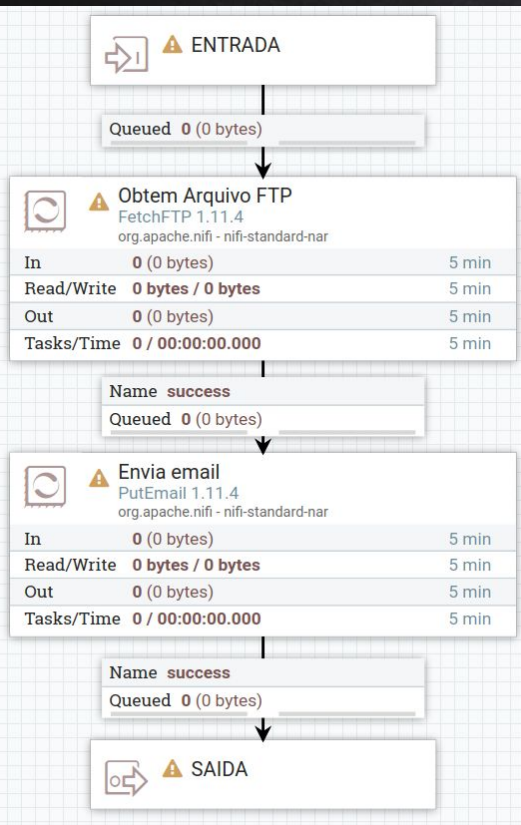
E MAIS...

CONNECTIONS

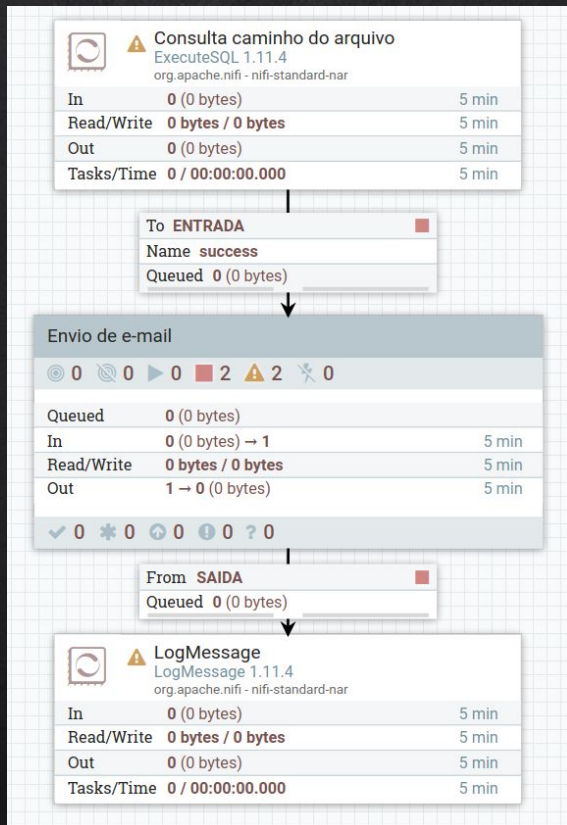


- ✗ Conexão entre Processors
- ✗ Atua como fila

VISÃO INTERNA DO PROCESS GROUP



VISÃO EXTERNA DO PROCESS GROUP



PROCESS GROUP

- ✗ Conjunto de processadores e connections
- ✗ Envia e recebe dados via portas

CONTROLLER SERVICE

x Recursos compartilhados no fluxo

CONEXÃO COM
BANCO DE
DADOS

GESTÃO DE
CACHE
(REDIS)

CREDENCIAIS
CLOUD

CONTEXTMAP
HTTP
(API)

...

DATA PROVENANCE

x Caminho percorrido pelo FlowFile

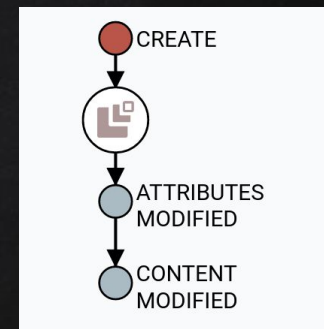
NiFi Data Provenance

Displaying 3 of 3
Oldest event available: 06/26/2020 12:29:22 UTC

Showing the events that match the specified query. [Clear search](#)

Filter by component name

	Date/Time ▾	Type	FlowFile Uuid	Size	Component Name	Component Type	
i	06/26/2020 12:50:58.339 ...	CONTENT_MODIFIED	26205ee2-f787-4262-8e5d...	1 bytes	EvaluateXQuery	EvaluateXQuery	🔗 →
i	06/26/2020 12:50:58.325 ...	ATTRIBUTES_MODIFIED	26205ee2-f787-4262-8e5d...	11 bytes	UpdateAttribute	UpdateAttribute	🔗 →
i	06/26/2020 12:50:58.315 ...	CREATE	26205ee2-f787-4262-8e5d...	11 bytes	GenerateFlowFile	GenerateFlowFile	🔗 →



Operadores: rastreamento, ciclo de vida

Governança: Auditoria

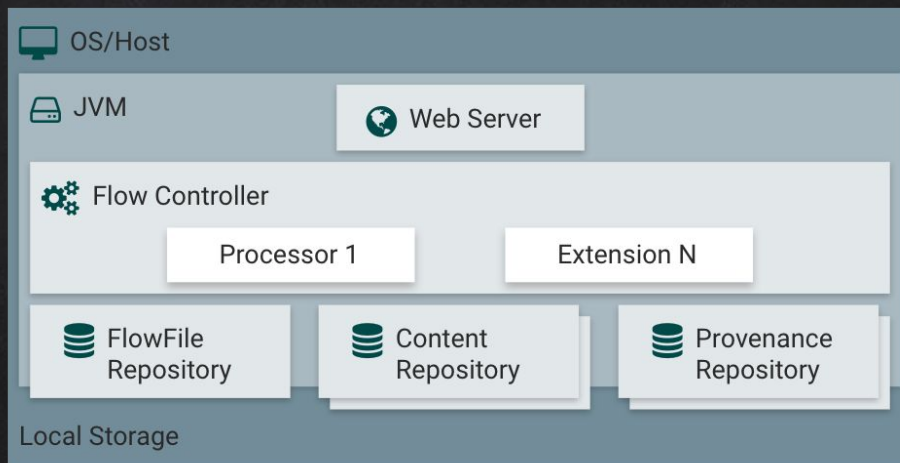
Negócio: Fontes de valor, Valores de investimento

FLOW CONTROLLER

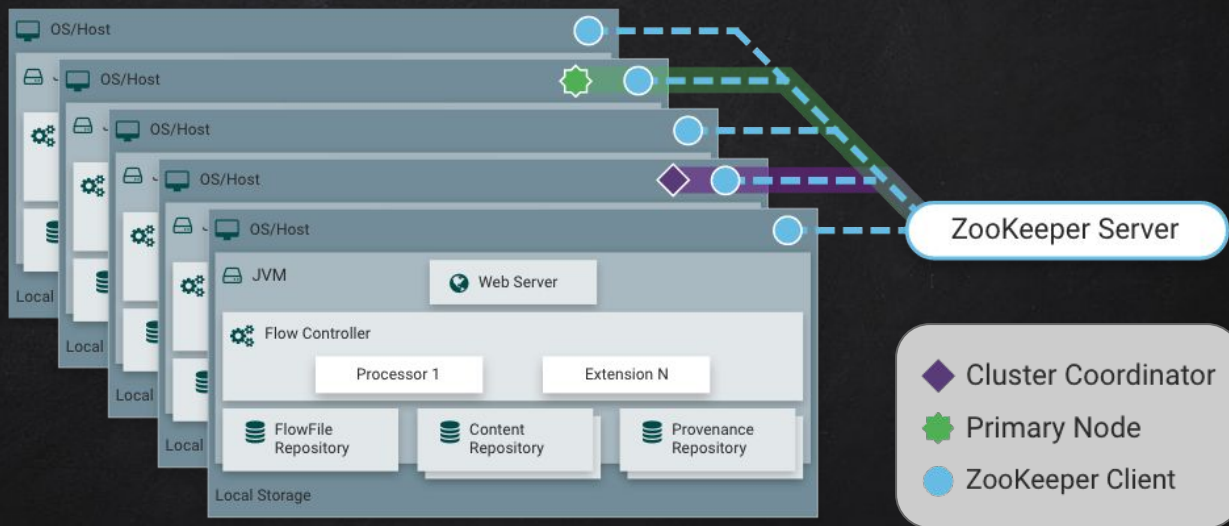
- ✗ Orquestra como os processos se conectam e gerencia todas alocações que os processos usam.
- ✗ Atua como broker facilitando a troca de FlowFiles entre os Processors



ARQUITETURA



ARQUITETURA - CLUSTER





**KEEP
CALM
AND
DEMO
ON**



FUNCIONALIDADES

GARANTIA DE
ENTREGA

BACK PRESSURE

PRIORIZAÇÃO
DE FILAS

TEMPLATES DE
FLUXOS

DATA
PROVENANCE

EXTENSIBILIDADE

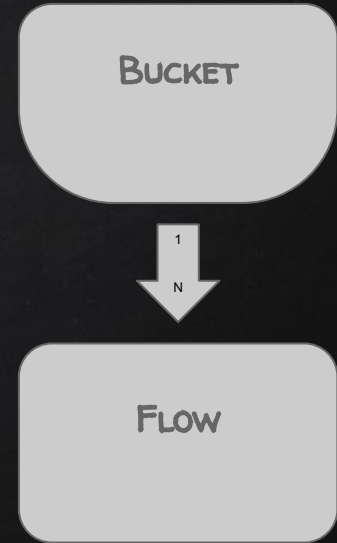
ESCALABILIDADE

SEGURANÇA

NIFI REGISTRY

The screenshot displays the NiFi Registry web interface. At the top, the header shows 'NiFi Registry / All' and the user 'anonymous'. Below the header, there is a search bar and a 'Sort by: Name (a - z)' dropdown. The main content area shows details for a flow named 'ABCD - Test Flow'. It includes a 'VERSIONS' section with '1' version and an 'ACTIONS' dropdown. The 'DESCRIPTION' section states 'Generate flowfiles and store attributes'. The 'CHANGE LOG' section shows 'Version 1 - 5 minutes ago by anonymous' and 'Initial save Jan-16-2018 at 2:48 PM'.

DESCRIPTION	CHANGE LOG
Generate flowfiles and store attributes	Version 1 - 5 minutes ago by anonymous
	Initial save Jan-16-2018 at 2:48 PM



APRENDA MAIS

- ✕ Site Oficial: <https://nifi.apache.org/>
- ✕ Idéias e sugestões: <https://issues.apache.org/jira/browse/nifi>
- ✕ Twitter: @apachenifi
- ✕ Telegram: <https://t.me/nifibrasil>



OBRIGADO!

Perguntas?

<https://github.com/elizarp/campdata-2021>



CONTATO

- x eliezer.zarpelao@gmail.com
- x <https://www.linkedin.com/in/eliezerzarpelao/>
- x <http://eliezerzarpelao.eti.br>
- x @eliezerzarpelao (Twitter/Instagram)



CREDITS

Special thanks to all the people who made and released these awesome resources for free:

- ✕ Presentation template by [SlidesCarnival](#)
- ✕ Photographs by [Unsplash](#)