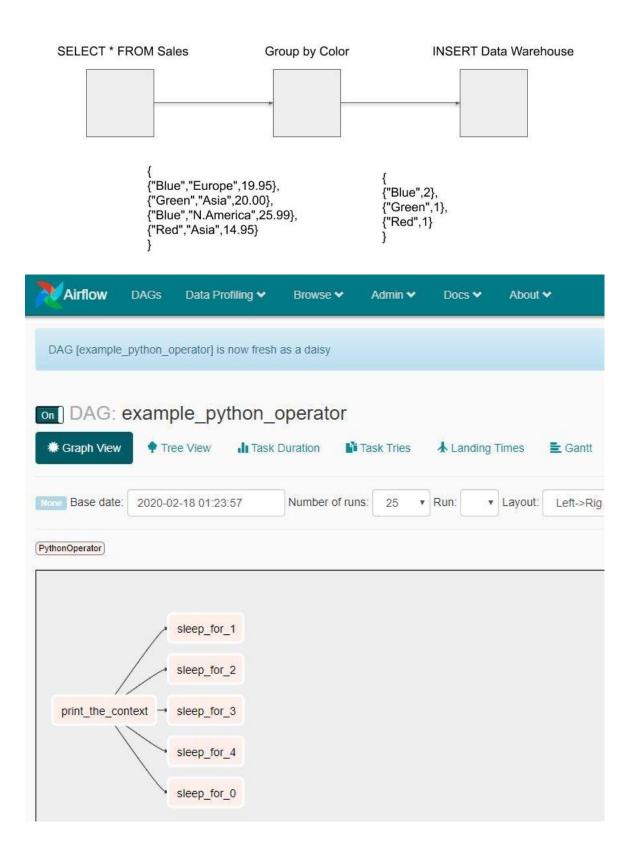


Chapter 1: What Is Data Engineering

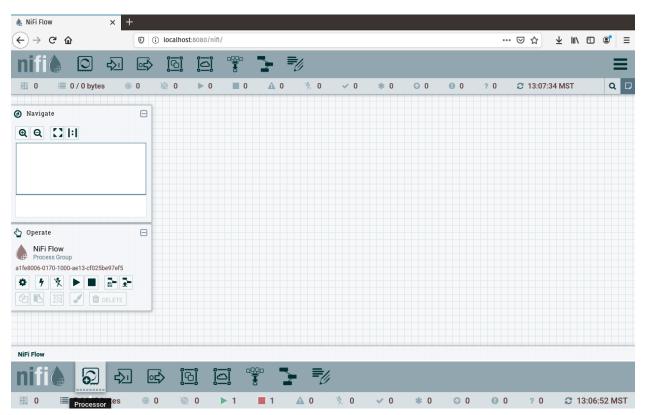
Widget	Region	
Blue	1	
Green	2	
Red	3	

Widget	Blue	
	Green	
	Red	
Region	1	
	2	
	3	

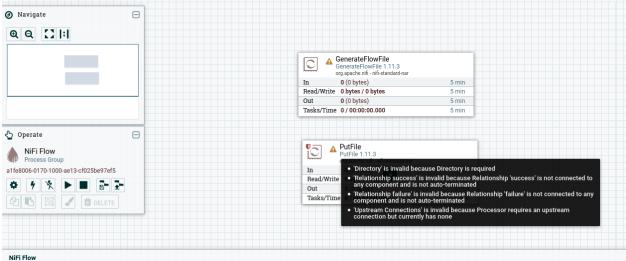


Chapter 2: Building Our Data Engineering Infrastructure

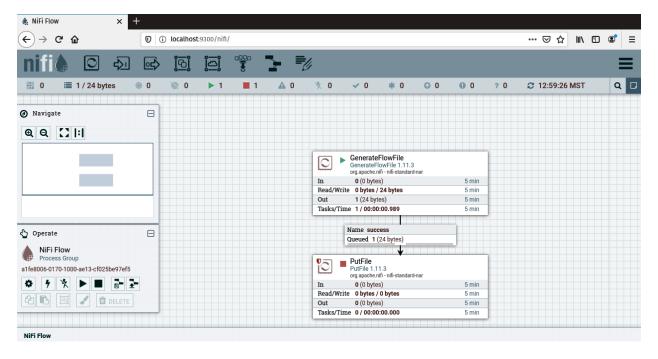




$ \underbrace{ \left\{ \begin{array}{c} Nifi \ Flow & X \end{array} \right. } \\ \underbrace{ \left\{ \begin{array}{c} \leftarrow \end{array} \right\} \rightarrow \ C^{L} & \textcircled{O} & \textcircled{O} \end{array} } $	ocalhost:9300/nifi/			6	ז מו ב) © ≡
	dd Processor					≡
🗱 0 🗎 0 / 0 bytes 💿 0				[]	:09 MST	Q
So	ource	Displaying 284 of 284		Filter		
🕢 Navigate 📃	all groups 🛛 🗸	Туре 🔺	Version	Tags		
		AttributeRollingWindow	1.11.3	rolling, data science, Attribute		
	mazon attributes	AttributesToCSV	1.11.3	flowfile, csv, attributes		
	vro aws consume	AttributesToJSON	1.11.3	flowfile, json, attributes		
	sv database	Base64EncodeContent	1.11.3	encode, base64		
	elete fetch get	CalculateRecordStats	1.11.3	stats, record, metrics		
	ngest ingress	CaptureChangeMySQL	1.11.3	cdc, jdbc, mysql, sql		
	nsert json kafka	CompareFuzzyHash	1.11.3	fuzzy-hashing, hashing, cyber		
	sten logs	CompressContent	1.11.3	Izma, decompress, compress,		09 MST Q
	nessage pubsub	ConnectWebSocket	1.11.3	subscribe, consume, listen, We	data science, Attribute a, csv, attributes a, json, attributes b, json, attributes b, json, attributes c, msql, sql ashing, hashing, cyber lecompress, compress, lecompress, compress, lecompress, lecompress, lecompress, lecompress, lecompress,	
	out record	ConsumeAMQP	1.11.3	flowfile, csv, attributes flowfile, json, attributes encode, base64 stats, record, metrics cdc, jdbc, mysql, sql fuzzy-hashing, hashing, cyber Izma, decompress, compress, subscribe, consume, listen, We receive, amqp, rabbit, get, cons cloud, streaming, streams, eve nlfi-stateful-analysis-nar n Language expression on each FlowFile will be emitted with the count of		
A NiEi Elow	estricted source ext update	ConsumeAzureEventHub	1.11.3	cloud, streaming, streams, eve	er Ing, data science, Attribute wfile, csv, attributes wfile, json, attributes code, base64 ts, record, metrics b, jdbc, mysql, sql zy-hashing, hashing, cyber ha, decompress, compress, bscribe, consume, listen, We veive, ampt, rabbit, get, cons ud, streaming, streams, eve H-analysis-nar expression on each FlowFile ted with the count of t time window. CANCEL ADD	
Process Group Image: Control of the contr		Track a Rolling Window based and add that value to the proce	Version Tags Window 1.11.3 rolling, data science, Attribute V 1.11.3 flowfile, iso, attributes ON 1.11.3 flowfile, iso, attributes Content 1.11.3 encode, base64 rdStats 1.11.3 encode, base64 MYSQL 1.11.3 cdc, jdbc, mysql, sql Hash 1.11.3 jdzy-hashing, hashing, cyber Lent 1.11.3 izma, decompress, compress, Socket 1.11.3 izma, decompress, compress, p 1.11.3 izma, decompress, compress, p 1.11.3 izma, decompress, compress, gWindow 1.11.3 org.apache.nifi - nifi-stateful-analysis-nar ndow based on evaluating an Expression Language expression on each FlowFile to the processor's state. Each FlowFile will be emitted with the count of a laggregate value of values processed in the current time window.			
	ocalhost:9300/nifi/			··· @	7☆ Ⅲ\ Œ) 🔹 =
	الم الم الم					=
🗱 0 🗮 0 / 0 bytes 💿 0 🔞	0 > 0 0	🛕 2 🦎 O	✓ 0 * 0 ○ 0	🕒 0 ? 0 📿 12:	54:40 MST	Q



🎄 NiFi Flow 🗙 🕂				
← → ℃ ŵ	localhost:9300/nifi/?processGroupId=root&component	:lds=ac42a400-0170-1000-8ea	a0-ee5ca5617099 🚥 🗵 🏠 🕪 🗉 🔘	' ≡
nifia o 🎝 🗗	igi isi 🍟 🛼 🗐		_	≡
🇱 0 🗮 0 / 0 bytes 💿 0	Configure Processor		2:56:41 MST	2 🗆
	🔺 Invalid			
Navigate	SETTINGS SCHEDULING PROPERTIES	COMMENTS	_	
	Required field	•	+	
	Property	Value		
	File Size O	0B	_	
II	Batch Size O Data Format O	1 Text	-	
	Unique FlowFiles 0	false	_	
	Custom Text O	This is a file fron nifi		
🖒 Operate 📃	Character Set 0	UTF-8		
GenerateFlowFile	filename 0	NifiFile.txt 🛍		
ac42a400-0170-1000-8ea0-ee5ca5617099				
* * * > = = =				
省 🖪 🗔 🖌 🛍 delete				
		CANCEL APPLY		
NiFi Flow				



		🛈 🛈 localhos	c .9300/1111/				⊚ ☆ Ш\	
			<u>a</u> p	nQQn 🗬	=,			
c1100000								×
SUCCESS	0.00 (
Displaying 2 of 2 (4 Position	UUID		Filename	File Size	Queued Duration	Lineage Duration	Penalized	
0 1		7-4966-b40c-1892f4	NifiFile.txt	24.00 bytes	00:00:22.332	00:00:22.332	No	± © Ľ
0 2	87b095a1-f82	2f-48b5-a47a-129b1	NifiFile.txt	24.00 bytes	00:00:12.312	00:00:12.312	No	± • Ľ
Fi Flow	× +	localhoct.0200/ni6/2nd	rocose Croupld-ro	act ^e componentide-	ac(74551.0120.1000.07b7	c166b70c2265	R A	
→ሮ		localhost:9300/nifi/?pr	rocessGroupId=ro	oot&componentIds=	ac47d551-0170-1000-e7b7	-c166b78a3365	⊠ ☆	III\ 🗉 😂
		<u></u>		= /				
success		FlowFile						×
Displaying 2 of 2 (48.00 b	ytes)							
		DETAILS	ATTRIBUTES		Contont Claim		Penalized	
1 f8	f4fd78-38f7-4966-b40	FlowFile Details			Content Claim Container		No	± © Ľ
0 2 87	b095a1-f82f-48b5-a4						No	± • Ľ
0 2 87	'b095a1-f82f-48b5-a4'	f8f4fd78-38f7-4966-b	040c-1892f47e43	86	default		No	1 () () 1 () () () () () () () () () (
0 2 87	'b095a1-f82f-48b5-a4'		o40c-1892f47e43	186			No	± ⊕ Ľ
0 2 87	b095a1-f82f-48b5-a4	f8f4fd78-38f7-4966-b Filename NifiFile.txt File Size	o40c-1892f47e43	86	default Section 1 Identifier		No	2 © 4
9 2 87	b095a1-f82f-48b5-a4	f8f4fd78-38f7-4966-b Filename NifiFile.txt	o40c-1892f47e43	186	default Section 1		No	* ® Ľ
• 2 87	b095a1-f82f-48b5-a4	f8f4fd78-38f7-4966-5 Filename NifiFile.txt File Size 24 bytes Queue Position No value set	o40c-1892f47e43	186	default Section 1 Identifier 1583438364746-1 Offset 72		No	± ⊕ Ľ
● 2 87	b095a1-f82f-48b5-a4	f8f4fd78-38f7-4966-t Filename NifiFile.txt File Size 24 bytes Queue Position	040c-1892f47e43	186	default Section 1 Identifier 1583438364746-1 Offset		No	± ⊕ Ľ
2 87	b095a1-f82f-48b5-a4	f8f4fd78-38f7-4966-5 Filename NfiFile.txt File Size 24 bytes Queue Position No value set Queued Duration	040c-1892f47e43	186	default Section 1 Identifier 1583438364746-1 Offset 72 Size	VIEW	No	±⊕ <u>₽</u>
0 2 87	b095a1-f82f-48b5-a4	f8f4fd78-38f7-4966-5 Filename NfiFile.txt File Size 24 bytes Queue Position No value set Queued Duration 00:00:46.491 Lineage Duration	040c-1892f47e43	186	default Section 1 Identifier 1583438364746-1 Offset 72 Size 24 bytes	I VIEW	No	2 O E

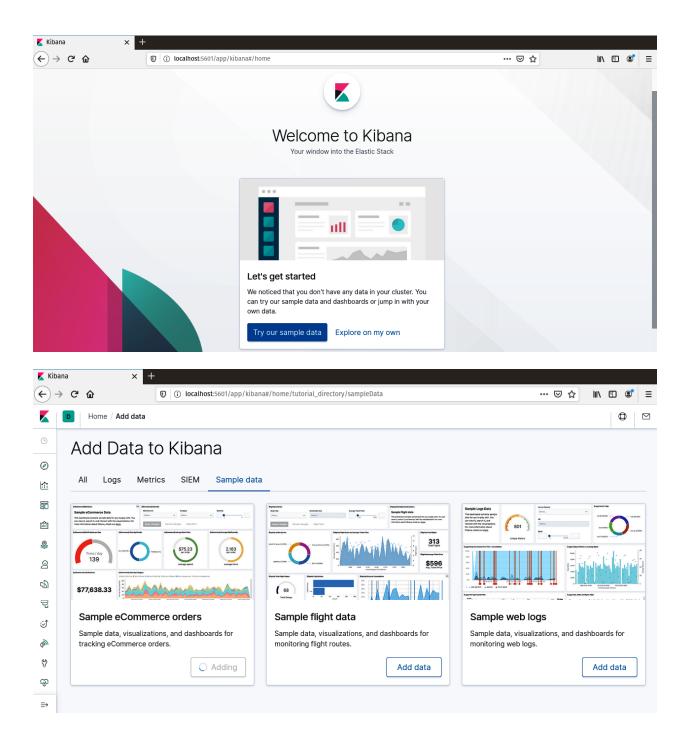


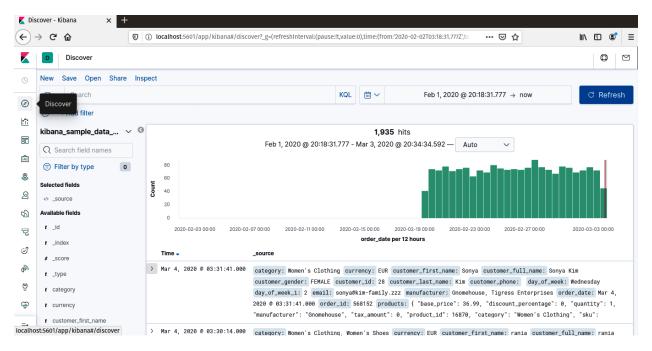
/ _ //_ / / _ / / /
// _/_/ /_/ /_/ /_/ \/\/ / [2020-03-05 16:08:06,657] {initpy:51} INFO - Using executor SequentialExecu
tor
[2020-03-05 16:08:06,823] {scheduler_job.py:1344} INFO - Starting the scheduler
[2020-03-05 16:08:06,824] {scheduler_job.py:1352} INFO - Running execute loop fo
r -1 seconds
[2020-03-05 16:08:06,825] {scheduler_job.py:1353} INFO - Processing each file at
most -1 times
[2020-03-05 16:08:06,825] {scheduler_job.py:1356} INFO - Searching for files in
/home/paulcrickard/airflow/dags
[2020-03-05 16:08:06,844] {scheduler_job.py:1358} INFO - There are 23 files in /
home/paulcrickard/airflow/dags
[2020-03-05 16:08:06,845] {scheduler_job.py:1409} INFO - Resetting orphaned task
s for active dag runs
[2020-03-05 16:08:06,940] {dag_processing.py:556} INFO - Launched DagFileProcess
orManager with pid: 29453
<pre>[2020-03-05 16:08:06,964] {settings.py:54} INFO - Configured default timezone <t< pre=""></t<></pre>
imezone [UTC]>
[2020-03-05 16:08:07,000] {dag_processing.py:758} WARNING - Because we cannot us
e more than 1 thread (max_threads = 2) when using sqlite. So we set parallelism
to 1.

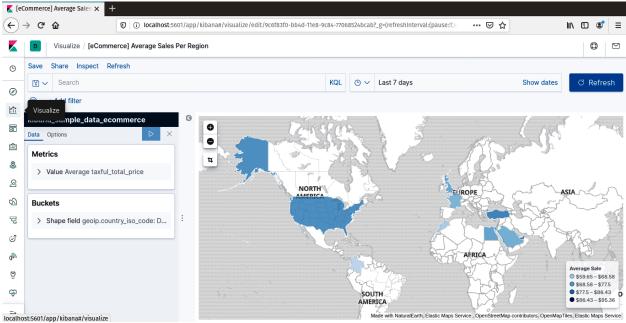
	0	DAG	Schedule	Owner	Recent Tasks	Last Run DAG F	Runs 🔁	Links
3	Off	example_bash_operator	0 0 * * *	Airflow	0000000000	00		・ *:川助水三 + 三○(
ï	Off	example_branch_dop_operator_v3	*/1 * * * *	Airflow			0	¢#.الللهُ≮⊇٢≡۵(
š	Off	example_branch_operator	@daily	Airflow			0	¢≉.اللللله ♦
5	Off	example_complex	None	airflow			0	¢≉ الللهُ الله
ŝ	Off	example_external_task_marker_child	None	airflow			0	₩.IIIIIA≣≯≣C
ï	Off	example_external_task_marker_parent	None	airflow			0	¢≣≁⊒≁≣۵
š.	Off	example_http_operator	1 day, 0:00:00	Airflow			0	*.li@ik≥ / ≘C(
ŝ	Off	example_passing_params_via_test_command	*/1****	airflow			0	°≈+⊒√≦∩
ï	Off	example_pig_operator	None	Airflow			0	*.h@★≥≁≘C(
3	Off	example_python_operator	None	Airflow			0	*山峰大三ヶ三〇(
		G: example_bash_operator	✓ Admin N ✓ Task Tries		About ✔ ng Times E Gantt III Details ∳C	ode 💿 Trigger D		03-05 23:23:42 UTC schedule: 0 0 * * esh ② Delete
	sraph v					- 55		
	Graph V Base c	late: 2020-03-05 23:23:06 Number of	of runs: 25	• Run:	Layout: Left->Right Go			Search for
ti te c	Base o	late: 2020-03-05 23:23:06 Number of DummyOperator	of runs: 25	• Run:	Layout: Left->Right Go	pstream_failed up_for_re	schedule up_for_	

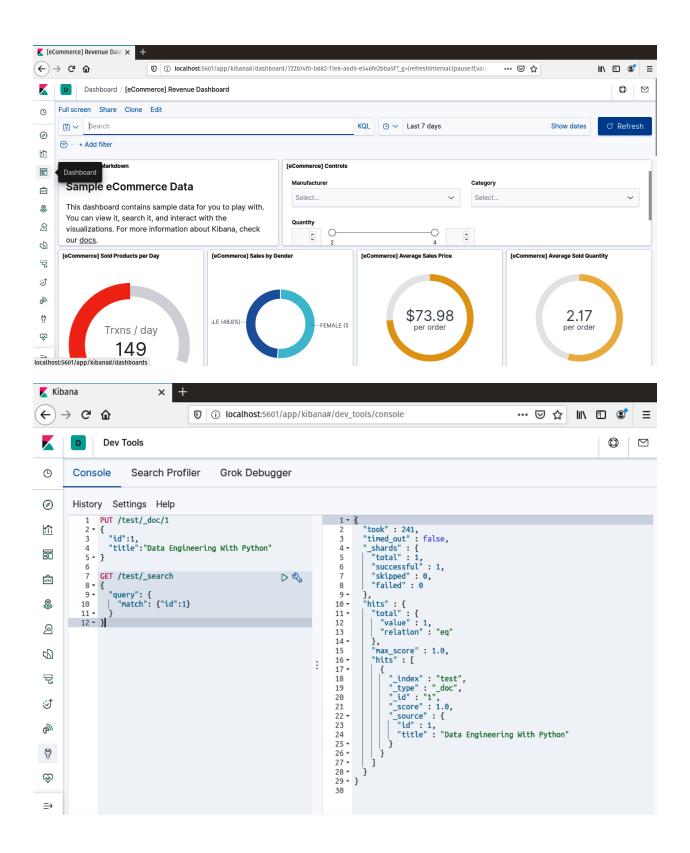
🗙 Airflow - DAGs 🛛 🗙 🕂									
← → C ŵ	(i) localhost:8080/ac	dmin/airflow/tr	ee?dag_id=exampl	e_bash_operator				⊠ ☆	li\ ⊡ © ≡
Airflow DAGs Data Profilir	ng 🗙 🛛 Browse 🗸	Admin 🗸	Docs 🗸 🛛 Ab	out 🗸				2020-03-05	23:43:43 UTC
on] DAG: example_bash	_operator							sc	hedule: 0 0 * * *)
Graph View	Task Duration	Task Tries	★ Landing Time:	Gantt	🔳 Details	✤ Code	O Trigger DAG	C Refresh	Oelete
Base date: 2020-03-05 23:43:10	Number of runs: 2	5 J Go							
BashOperator DummyOperator			suc	cess 📘 running 📕	ailed 🔲 skipped 📘	upstream_faile	d 🔲 up_for_reschedule [up_for_retry 📕 q	ueued 🗌 no_status
IDAGJ runme_0 run_after_loop run_this_last runme_1 Orun_after_loop Orun_after_loop Orun_ater_loop oliso_run_this orun_this_last									
localhost:8080/admin/airflow/trigger?dag_ic	l=example_bash_operat	tor&origin=/adn	nin/airflow/tree?d	ag_id=example_b	ash_operator				
æ		pa	ulcrickard@	pop-os: ~/a	irflow			Q	= 🙁
# The amount of pa # the max number o # on this airflow parallelism = 32	of task in	istance						ines	
# The number of ta dag_concurrency =		ices al	lowed t	o run d	oncurr	ently	by the	schedul	.er
# Are DAGs paused dags_are_paused_at									
# The maximum numb max_active_runs_pe			G runs	per DAG	j				
# Whether to load # get started, but # environment <mark>l</mark> oad_examples = Tr	t you prob								1
# Where your Airfl plugins_folder = /				low/plı	ıgins				
# Secret key to sa	ave connec	tion p	assword	s in th	ie db				

↔ ở ŵ	🛈 🛈 localhost:8	080/admin/			⊌ ☆	II\ ⊡ 💐 =
Airflow DAGs Data Prof	iling 🗸 🛛 Browse	e 🗸 🛛 Admin 🗸	Docs 🗙 About 🗸		2020-03	-06 00:06:45 UTC
DAGs						
			Search:			
DAG	Schedule	Owner	Recent Tasks 🕄 '	Last Run 🖯	DAG Runs 🖯 '	Links
No data available in table						
« · · »					Showi	ng 0 to 0 of 0 entries
Hide Paused DAGs						
localhost:9200/ $\leftarrow \rightarrow C^{\prime}$		 localhost: 	9200	⊘	☆ ⊻ III (• ● =
Save Copy Collapse All Expan						
	d All 🛛 🗑 Filter	JSON				
name:	d All 🛛 🖓 Filter	JSON "OnlyNo	ode"			
	d All 🛛 Filter	"OnlyNo	ode" ngineeringWithPytho	n"		
name:	d All 🛛 🖓 Filter	"OnlyNo "DataEr		n "		
name: cluster_name:	d All 🛛 🖓 Filter	"OnlyNo "DataEr	ngineeringWithPytho	n"		
name: cluster_name: cluster_uuid:	d All 🛛 🖓 Filter	"OnlyNo "DataEr	ngineeringWithPytho Z4dQTeMbh-VV66U4w"	n"		
<pre>name: cluster_name: cluster_uuid: version:</pre>	d All 🛛 🖓 Filter	"OnlyNo "DataEr "hxxcU2	ngineeringWithPytho 24dQTeMbh-VV66U4w"	n"		
<pre>name: cluster_name: cluster_uuid: version: number:</pre>	d All 🛛 Filter	"OnlyNo "DataEr "hxxcU2 "7.6.0"	ngineeringWithPytho 24dQTeMbh-VV66U4w"	n"		
<pre>name: cluster_name: cluster_uuid: version: number: build_flavor:</pre>	d All 🛛 Filter	"OnlyNo "DataEr "hxxcUZ "7.6.0" "defaul "tar"	ngineeringWithPytho 24dQTeMbh-VV66U4w"		e3"	
<pre>name: cluster_name: cluster_uuid: version: number: build_flavor: build_type:</pre>	d All 🛛 Filter	"OnlyNo "DataEr "hxxcU2 "7.6.0" "defaul "tar" "7f634e	ngineeringWithPytho 24dQTeMbh-VV66U4w" .t"	cclda681b0c3b1	e3"	
<pre>name: cluster_name: cluster_uuid: version: number: build_flavor: build_type: build_hash:</pre>	d All 🛛 Filter	"OnlyNo "DataEr "hxxcU2 "7.6.0" "defaul "tar" "7f634e	ngineeringWithPytho 24dQTeMbh-VV66U4w"	cclda681b0c3b1	e3"	
<pre>name: cluster_name: cluster_uuid: version: number: build_flavor: build_type: build_hash: build_date:</pre>	d All 🛛 Filter	"OnlyNo "DataEr "hxxcU2 "7.6.0" "defaul "tar" "7f634e "2020-0	ngineeringWithPytho 24dQTeMbh-VV66U4w" - Lt" 29f44834fbc12724506 02-06T00:09:00.4499	cclda681b0c3b1	e3"	
<pre>name: cluster_name: cluster_uuid: version: number: build_flavor: build_type: build_hash: build_date: build_snapshot:</pre>		"OnlyNo "DataEr "hxxCU2 "7.6.0" "defaul "tar" "7f634e "2020-0 false "8.4.0"	ngineeringWithPytho 24dQTeMbh-VV66U4w"	cclda681b0c3b1	e3"	
<pre>name: cluster_name: cluster_uuid: version: number: build_flavor: build_flavor: build_type: build_hash: build_date: build_snapshot: lucene_version:</pre>	ility_versior	"OnlyNo "DataEr "hxxcU2 "7.6.0" "defaul "tar" "7f634e "2020-0 false "8.4.0"	ngineeringWithPytho 24dQTeMbh-VV66U4w"	cclda681b0c3b1	e3"	









æ	paulcrickard@pop-os: ~ Q	Ξ	8
Pack	age configuration		
	Configuring pgadmin4-apache2 The pgAdmin4 web interface needs an initial user configured. Please enter the email address you want to use as the account name. Initial pgAdmin4 user email: postgres@localhost		
	<0k>		



😝 pgAdmin 4 🗙 🕂				
(\leftarrow) \rightarrow C' \textcircled{a} $\textcircled{0}$ $\textcircled{0}$ $\textcircled{0}$ $\textcircled{0}$	calhost/pgadmin4/browser/#		⊠ ☆	III\ 🗊 🔹 ≡
File - Object - Tools - H	Help 🗸			🎽 postgres@localhost 🗸
Browser Servers	Ashboard Welcome Feature pgAdmin is procedural c Quick Links Password Save	x ion_SSL_SSH Tunnel_Advanced localhost 5432 postgres postgres 	e. It includes a graphical administration inter ppers, DBAs and system administrators alike	¥ face, an SQL query tool, a
G	Postgr	★ Cancel CReset Save	Configure pgAdmin	pmmunity Support

Pg pgAdmin 4 X	+									
← → ୯ ŵ	Iocalhost/pgadmin4	/browser/#					₪	☆	lii\ 🗉) (2)
PgAdmin File - Objec	ct ∽ Tools ∽ Help ∽							8	postgres@	localhos
Browser	🗊 🎟 📷 Dashboard Prope	erties SQL S	Statistics Depen	dencies Dependen	ts					
Servers (1)	Database session 2.00 1.80 Total 1.60 Idle	S				Transactions per second		Λ		
 Sector Solution Sec	1.20					0.0				
 ♥ Schemas (1) ♥ public > \$ Lollations > \$ FS Confi > \$ FTS Confi > \$ FTS Confi > \$ TS Persi > < FTS Terry > \$ FF Foregar 	igurations 0.50			Fuples out 15000 12500 10000 7500 5000 2500 0			Block I/O			
 > (c) Functions > (c) Materializ > (f) Procedure 1.3 Sequence 	create cr	Table Prepared Tra	nsactions					Q Search		2
Tables		User postgres	Application pgAdmin 4 - DB:	dataengineering	Client ::1	Backend start 2020-03-03 16:36:05 MST	State active	Wait event	Blocking	PIDs

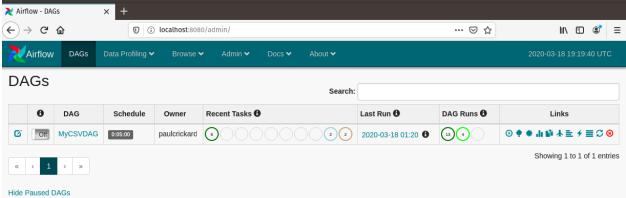
Со	lum	nns						•
		Name	Data type		Length/Precision	n Scale	Not NULL?	Primary key
ľ	Ŵ	name	text	•			Yes	No
ß	Ŵ	id	integer	•			Yes	Yes
ß	Ŵ	street	text	•			Yes	No
ľ	Ŵ	city	text	•			Yes	No
ß	Ô	zip	text	•			Yes	No
i		?				× Can	cel 🗘 Rese	t 🖪 Save
	7	Airflow DAGs Data	Profiling 🗸 🛛 Browse 🗸	Admin 🛩 🛛 Docs 🗸	About 🗸		2020-03	3-13 18:55:47 UTC
		The scheduler does not appear to	o be running. Last heartbeat wa nd new tasks will not be schedu					

🖽 Create - Table

Chapter 3: Reading and Writing Files



æ		paulcricka	rd@pop-os: ~		Q	Ξ	×
Pytho [GCC Type >>> i >>> d	rickard@pop-os:~\$ on 3.7.5 (default, 9.2.1 20191008] c "help", "copyrigh mport pandas as p f=pd.read_csv('da f.head(10)	, Nov 20 2 on linux nt", "cree od			e infor	matio	on.
0 1 2 3 4 5 6 7 An 8 9 K [10 r		23 36 61 29 23 57 42 34 57		-59.009437 58.163167 38.689889 81.272300 11.476733 23.872011 -78.616711 73.918076 63.915574			
>>>	DAGs × +		Airflow - DAGs - Mozi	lla Firefox			8



Airflow - DAGs 🛛 🗙 –	+ <u> </u>								
)→ C' û		080/admin/airflow/t	tree?base_date=&nt	um_runs=5&root=	≈&dag_id=MyCSV	/DAG&_csrf_to	ken=ljZmNzRmY	⊚ ☆	\ ⊡ ©
Airflow DAGs Data F	Profiling 🗸 🛛 Brows	e 🗸 🛛 Admin 🗸	Docs 🗸 🛛 At	bout 🗸				2020-03-1	.8 19:41:35 UTC
Graph View	I Task Duration	🗳 Task Tries	★ Landing Time:	s 🖹 Gantt	🔳 Details	✤ Code	O Trigger DAG	C Refresh	Oelete
se date: 2020-03-18 00:20:00	Number of run	s: 5 • Go							
BashOperator OPythonOperator			suc	ccess 📘 running 📕	failed 🔲 skipped	upstream_faile	d 🔲 up_for_reschedule	e 🔜 up_for_retry 📕	queued 🗌 no_statu
[DAG] O starting O convertCSVtoJson									
Airflow - DAGs X H	-	Air	flow - DAGs - Mozill	a Firefox					
)→ C' û	C	080/admin/airflow/ti	ree?base_date=ν	Im_runs=5&root=8	&dag_id=MyCSV[DAG&_csrf_tok	en=ljZmNzRmY.	⊠ ☆	III\ 🗉 📽
Airflow DAGs Data F	rofiling 🗙 🛛 Browse	e 🗸 🛛 Admin 🗸	Docs 🗸 🛛 Ab	out 🗸				2020-03-1	8 19:44:01 UTC
DAG: MyCSVDA	G							s	chedule: 0:05:00
Graph View		💕 Task Tries	★ Landing Times	s 🖹 Gantt	🔳 Details		• Trigger DAG	C Refresh	🙁 Delete
	II Task Duration		A5						
se date: 2020-03-18 00:35:00 BashOperator O PythonOperator	Number of run: Task_id: conv Run: 2020-03 Operator: Pyt	S: 5 Go rertCSVtoJson I-18T00:15:00+00:00 honOperator	su		failed 🛄 skipped 🛛	upstream_faile	d 🔲 up_for_reschedule	up_for_retry	queued 🗌 no_statu:
se date: 2020-03-18 00:35:00	Number of run: Task_id: conv Run: 2020-03 Operator: Pyt Started: 2202	s: 5 Go rertCSVtoJson -18T00.15.00-00.00 honOperator -03-18T19:42:46.65921 03-18T19:42:46.211794 33Sec	9+00:00		failed 🛄 skipped	upstream_faile	d 🔲 up_for_reschedule	up_for_retry	queued 🗌 no_statue
EashOperator PythonOperator	Number of run Task_id: com Run: 2020-00 Operator: Pyl Stared: 2020 Ended: 2020 Duration: 1.55	s: 5 Go rertCSV/toJson -18700/15/00-00:00 honOperator -03-18719-42:44.65921 03-18719-42:46.211794 33Sec s	9+00:00 ++00:00		failed 🛄 skipped 🛛	upstream_faile	d 🔲 up_for_reschedule	: up_for_retry	queued 🗌 no_status
EashOperator PythonOperator	Number of run Task_id: com Operato: Pyt Started: 2020 Ended: 2020 Duration: 20 Duration: Pyt Started: 2020 Horation: 2020 Duration: 2020 State: succes	s: 5 Go rertCSV/toJson -18700/15/00-00:00 honOperator -03-18719-42:44.65921 03-18719-42:46.211794 33Sec s	9+00:00 ++00:00 Airflow - DAG	iccess 🗖 running 📕 5s - Mozilla Firefox	failed 🛄 skipped 🛛	upstream_faile	d ∎up_for_reschedule ල ☆		queued □ no_status
se date: 2020-03-18 00:35:00 BashOperator PythonOperator [DAG] \bigcirc starting \bigcirc convertCSVtoJson \bigstar Airflow - DAGs × \longleftrightarrow \rightarrow \circlearrowright \bigstar	Number of run Task, Id: com Run: 2020-00 Operator: Py Statted: 2020 Ended: 2020 Duration: 1.5 State: succes	s: 5 Go ertCSVtoJson honOperator 03-18119-42-44.65921 03-18119-42-44.65921 33Sec 5 6 6 6 6 6 6 6 6 6 6 6 6 6	9+00:00 Airftow - DAG lag_id=MyCSVDAG&nun cocs ♥ About ♥	ss - Mozilla Firefox	failed 🗌 skipped 🛛	upstream_faile			× • • •
se date: 2020-03-18 00:35:00 BashOperator PythonOperator [DAG] \bigcirc starting \bigcirc convertCSVtoJson \bigstar Airflow - DAGs × $\overleftarrow{\leftarrow}$ \rightarrow \heartsuit $\overleftarrow{\bigtriangleup}$	Number of run: Task_id: com Run: 2020-00 Operator: Pyt Started: 2020 Ended: 20	s: 5 Go ertCSVtoJson honOperator 03-18119-42-44.65921 03-18119-42-44.65921 33Sec 5 6 6 6 6 6 6 6 6 6 6 6 6 6	9+00:00 Airtlow - DAG lag_id-MyCSVDAG&nun pcs < About < D20-03-18T00:00	ss - Mozilla Firefox n_runs=&root= 0:00+00:00	failed 🔄 skipped 🛾	upstream_faile		III	× ••• ••
se date: 2020-03-18 00:35:00 BashOperator PythonOperator [DAG] \bigcirc starting \bigcirc convertCSVtojson \checkmark Airflow - DAGs × $(\leftarrow) \rightarrow ~ \bigcirc ~ \bigcirc ~ \bigcirc$ \land Airflow DAGs Data The scheduler does not appear to	Number of run: Task_id: com Run: 2020:00 Diratared: 2020 Ended: 2020 Diratared: 2020 Diratared: 2020 Ended: 2020 Uratared: 2020 Ended: 2020 Uratared: 2020 Ended: 2020 Uratared: 2020 Diratared: 2020	s: 5 Go rertCSVtoJson -18T0015.00+000.00 honOperator -03-18T19-42-46.5921 53Sec s s to/admin/airflow/tree?d Admin V D starting V on 24 Task Instance Details Download Log (by atte	9+00:00 Airtlow - DAG lag_id=MyCSVDAG&nun DCS About 020-03-18T00:00 Rendered Ta	ss - Mozilla Firefox n_runs=&root= 0:00+00:00		upstream_faile		lin 2020-03-18 1f	× ••• ••
se date: 2020-03-18 00:35:00 BashOperator \bigcirc PythonOperator DAGJ \bigcirc starting \bigcirc convertCSVtoJson \checkmark Airflow - DAGs × $\bigcirc \bigcirc $	Number of run: Task_id: com Qoprator: Pyl Started: 2020 Duration: 1.5 State: succes Profiling	s: 5 Go rertCSVtoJson -18T00:15:00-000 honOperator -03-18T19:42:46.65921 03-18T19:42:46.211794 53Sec s s ao/admin/airflow/tree?d Admin D starting T on 24 Task Instance Details Download Log (by atted 1	9+00:00 Airflow - DAG lag_id=MyCSVDAG# locs About 020-03-18T00:00 Rendered Ta ampts):	ccess running 5s - Mozilla Firefox n_runs=&root= 1:00+00:00 visk instances vi	iew Log			lin 2020-03-18 1f	 20 €° Ξ 3:05:55 UTC
se date: 2020-03-18 00:35:00 BashOperator ○ PythonOperator [DAG] Starting ConvertCSVtojson Airflow - DAGs × ⓒ → ⓒ ⓒ Airflow DAGs Data The scheduler does not appear The DAGs list may not update, a IDAG: MYCSVDA	Number of run: Task_id: com Operator: Pyt Started: 2020 Ended: 202	s: 5 Go rertCSVtoJson -18T0015.00+000.00 honOperator -03-18T19-42-46.65921 00-18T19-42-46.211794 53Sec s s ao/admin/airflow/tree?d Admin Da starting T on 21 Task Instance Details Download Log (by atted 1 Run Ignore All De	9+00:00 Airflow - DAG lag_id=MyCSVDAG# locs About 020-03-18T00:00 Rendered Ta empts): eps Ignore Task Sta	iccess running 5s - Mozilla Firefox n_runs=&root= 1:00+00:00 vi isk Instances vi ate Ignore Task I	lew Log Deps	r D/	··· 🗵 🏠	lii\ 2020-03-18 18 sche	 20 €° Ξ 3:05:55 UTC
se date: 2020-03-18 00:35:00 BashOperator PythonOperator [DAG] Starting ConvertCSVtojson Airflow - DAGS × ConvertCSVtojson Airflow DAGS Data The scheduler does not appear I The DAGS list may not update, a Con DAG: MyCSVDA Graph View Tree View	Number of run: Task_id: com Operator: Pyt Started: 2020 Ended: 202	s: 5 J Go rertCSVtoJson -18T0015.00+000.00 honOperator -03-18T19-42-44.65921 00-18T19-42-46.211794 53Sec s s s ao/admin/airflow/tree?d v Admin V D starting ♥ on 21 Task Instance Details Download Log (by atted 1 Run Ignore All De Clear Past Fu	9+00:00 Airflow - DAG lag_id=MyCSVDAG# locs About 020-03-18T00:00 Rendered Ta empts): eps Ignore Task Sta ture Upstream	iccess running 55 - Mozilla Firefox n_runs=&root= 1:00+00:00 isk instances Vi ate Ignore Task I Downstream Re	lew Log Deps ccursive Failed		··· 🗵 🏠	III 2020-03-18 18 sche O Delete	 2005:55 UTC dule: 0:05:00
Se date: 2020-03-18 00:35:00 BashOperator ○ PythonOperator DAGJ ○ starting ○ convertCSVtojson Airflow - DAGS × ⓒ → C ☆ Airflow DAGS Data The Scheduler does not appear to The DAGs list may not update, a Graph View ● Tree View Base date: 2020-03-18 00:50:00 ○ BashOperator ○ PythonOperator	Number of run: Task_id: com Operator: Pyt Started: 2020 Ended: 202	S: 5 J Go HertCSVtoJson -18T0015.00+000.00 honOperator -03-18T19-42-46.65921 00-18T19-42-46.211794 53Sec s s ao/admin/airflow/tree?d Admin ✓ Da starting ♥ on 21 Task Instance Details Download Log (by atted 1 Run Ignore All De Clear Past Fu Mark Failed Past	9+00:00 Airflow - DAG lag_id=MyCSVDAG# locs About 020-03-18T00:00 Rendered Ta empts): eps Ignore Task Sta ture Upstream	iccess running is- Mozilla Firefox n_runs-&root- :00+00:00 isk instances vi ate Ignore Task I Downstream Re am Downstream	ew Log Deps cursive Failed		··· ♥ ☆ AG ♥ Refresh	III 2020-03-18 18 sche O Delete	 2005:55 UTC dule: 0:05:00
se date: 2020-03-18 00:35:00 BashOperator ○ PythonOperator	Number of run: Task_id: com Operator: Pyt Started: 2020 Ended: 202	s: 5 J Go rertCSVtoJson -18T0015.00+000.00 honOperator -03-18T19-42-44.65921 00-18T19-42-46.211794 53Sec s s s ao/admin/airflow/tree?d v Admin V D starting ♥ on 21 Task Instance Details Download Log (by atted 1 Run Ignore All De Clear Past Fu	9+00:00 Airflow - DAG lag_id=MyCSVDAG# locs About 020-03-18T00:00 Rendered Ta empts): eps Ignore Task Sta ture Upstream	iccess running is- Mozilla Firefox n_runs-&root- :00+00:00 isk instances vi ate Ignore Task I Downstream Re am Downstream	ew Log Deps cursive Failed		··· ♥ ☆ AG ♥ Refresh	III 2020-03-18 18 sche O Delete	 2005:55 UTC dule: 0:05:00

	Airflow - DAGs - Mozilla Firefox 🛛 😵
🗙 Airflow - DAGs 🛛 🗙	+
$\overleftarrow{\bullet}$ > C $\widehat{\bullet}$	🖸 🛈 localhost:8080/admin/airflow/log?task_id=readCSV&dag_id=MyCSVDAG&execution_date=2020-03-18T00%3A00%3 … 🖂 🏠 📗 🖾 🗮
	ta Profiling ♥ Browse ♥ Admin ♥ Docs ♥ About ♥ 2020-03-18 18:07:26 UTC
2020-03-16 12:00:41,702 2020-03-18 12:00:42,163 2020-03-18 12:00:42,201 2020-03-18 12:00:42,202 2020-03-18 12:00:42,202 2020-03-18 12:00:42,202 2020-03-18 12:00:42,202 2020-03-18 12:00:42,202 2020-03-18 12:00:42,202 2020-03-18 12:00:42,203 2020-03-18 12:00:42,203 2020-03-18 12:00:42,203 2020-03-18 12:00:42,203 2020-03-18 12:00:42,203 2020-03-18 12:00:42,204 2020-03-18 12:00:42,204 2020-03-18 12:00:42,204 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18 12:00:42,205 2020-03-18	<pre>{taskinstance.py:887 INFO - Executing <task(pythonoperator): readgsv=""> on 2020-03-18T00:00:00+00:00 {standard_task_runner.py:53 INFO - Started process 25805 to run task {logging_mixin.py:112 INFO - Running %s on host %s <taskinstance: 2020-03-18t00:00:00+00:00="" [running]="" mycsvdag.readcsv=""> pop-os.localdomain {logging_mixin.py:112 INFO - Arthur Garcia {logging_mixin.py:112 INFO - Arthur Garcia {logging_mixin.py:112 INFO - Arthur Garcia {logging_mixin.py:112 INFO - Faren Hodges {logging_mixin.py:112 INFO - Karen Hodges {logging_mixin.py:112 INFO - Karen Hodges {logging_mixin.py:112 INFO - Karen Hodges {logging_mixin.py:112 INFO - Michael Mack {logging_mixin.py:112 INFO - Michael Mack {logging_mixin.py:112 INFO - Michael Mack {logging_mixin.py:112 INFO - Mark Grimes {logging_mixin.py:112 INFO - Thomas Hall {logging_mixin.py:112 INFO - Thomas Hall {logging_mixin.py:112 INFO - Thomas Hall {logging_mixin.py:112 INFO - ElizabetD Smith {logging_mixin.py:112 INFO - Jon Nichols {logging_mixin.py:112 INFO - Jesica Cooper {logging_mixin.py:</taskinstance:></task(pythonoperator):></pre>
	le e 1.11.3 che.nifi - nifi-standard-nar

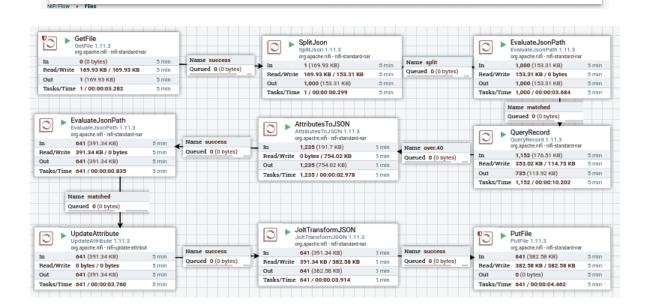
	me 1,000 / 00:00:05.961	5 min		Tasks/Time 635 / 00:00:01.711	5 min
Out	635 (80.07 KB)	5 min		Out 635 (80.07 KB)	5 min
	ite 253.29 KB / 80.07 KB	5 min	Queued 0 (0 bytes)	Read/Write 80.07 KB / 0 bytes	5 min
In	1,000 (126.64 KB)	5 min	Name over.40	In 635 (80.07 KB)	5 min
0	QueryRecord QueryRecord 1.11.3 org.apache.nifi - nifi-standard-nar			ExtractText ExtractText 1.11.3 org.apache.nifi - nifi-standard-ni	
	Queued 0 (0 bytes)			Queued 0 (0 bytes)	
	Name splits			Name matched	
.uond/ 1				Tasks/Time 635 / 00:00:01.288	5 min
	Time 1/00:00:03.012	5 min		Out 635 (80.07 KB)	5 min
Out	1,000 (126.64 KB)	5 min		Read/Write 0 bytes / 0 bytes	5 min
	rite 89.82 KB / 126.64 KB	5 min		In 635 (80.07 KB)	5 min
	 SplitRecord SplitRecord 1.11.3 org.apache.nifi - nifi-standard-na 1 (89.82 KB) 	5 min		UpdateAttribute UpdateAttribute 1.11.3 org.apache.nifi - nifi-update-attr	
	Name success Queued 0 (0 bytes)			Name success Queued 0 (0 bytes)	
Tasks/Ti	me 1/00:00:00.051	5 min		Tasks/Time 635 / 00:00:01.556	5 min
Out	1 (89.82 KB)	5 min		Out 0 (0 bytes)	5 min
	ite 89.82 KB / 89.82 KB	5 min		Read/Write 0 bytes / 0 bytes	5 min
In	0 (0 bytes)	5 min		In 635 (80.07 KB)	5 min
	GetFile 1.11.3 org.apache.nifi - nifi-standard-nar			PutFile 1.11.3 org.apache.nifi - nifi-standard-nar	

			NiFi F	low - Mozilla Firefox			•
🎄 NiFi Flow	× R Online r	egex tester and det $ imes $ +					
↔ ↔ ↔ ↔	00	localhost:9300/nifi/?processGroupId=	e537967f-0170-10	00-273d-2a7a2263f512&component	tlds=e53941bf-0170-1000-9	⊌ ☆	III\ 🗉 🔎 E
nifia 🖸		igi izi 📽 🦫	<i>≡∥,</i>				=
88 0 🔳 1 / 24 bytes		Configure Processor				/:38:30 MDT	Q [
		Stopped					
3 Navigate	8						
Q Q []!		SETTINGS SCHEDULING	PROPERTIES	COMMENTS			
		Required field				rd-nar 5 min	
		Droporty		Value		5 min 5 min	
		Property				5 min	
		Input Directory	0	/home/paulcrickard			
		File Filter	0	data.csv No value set			
		Path Filter	0				
		Batch Size	0	10		.3 te-attribute-nar	
) Operate	8	Keep Source File	0	true		5 min	
Operate		Recurse Subdirectories	0	false		5 min 5 min	
GetFile		Polling Interval	0	0 sec		5 min	
		Ignore Hidden Files	0	true			
53941bf-0170-1000-9cca-3d683a	290ed8	Minimum File Age	0	0 sec			
\$ 9 \$ ▶ ■		Maximum File Age	0	No value set			
		Minimum File Size	0	0 B		_	
🖆 🖪 🔟 🛍 DE	ELETE	Maximum File Size	0	No value set		lard-nar 5 min	
		Maxinflutti File Size	U	ING Funde Set		5 min	
						5 min	
					CANCEL A	S min	

	NiFi Flow - Mozitla Firefox	8
🎄 NiFi Flow 🛛 🗙 🖪 Apa	che Nifi Users List - Rec 🗙 🕂	
← → C' ŵ	🛈 localhost:9300/nifi/?processGroupId=e537967f-0170-1000-273d-2a7a2263f512&componentIds=e53d657c-0170-1000-1 🛛 💀 🔂 📗 🔘	\$ ∃
nifi 🖸 🖓 🖬	ার্টা বিষ্ণ 📲 🚚 🗮	≡
🗱 0 🔳 1 / 24 bytes 💿 0	Configure Processor 3:26:52 MDT	Q 🗆
	A Invalid	
 Ø Navigate Ø Q () : 	SETTINGS SCHEDULING PROPERTIES COMMENTS	
	Required field	
	Property	
	Record Reader No value Record Writer No value	
	Records Per Split Create new service	
🖒 Operate 🖂		
SplitRecord Processor		
e53d657c-0170-1000-188a-b81484eec8a1		
🌣 1/ 🦎 🕨 📑 🛬		
🖆 🖪 🧾 🖌 🛍 delete		
	CANCEL APPLY	
NiFi Flow » Files	Tasks/Time 0 / 00:00:000 Simm	

		NiFi Flow - Mozilla Firefox		<u> </u>
🎄 NiFi Flow 🗙 🖪 Apach	e NiFi UsersList - Rea 🗙	+		
← → ♂ ŵ	i) localhost:9300/nifi/?pr	rocessGroupId=e537967f-0170-1000-273d-2a7a2263f512&componentIds=e53d657c-0170-1000-1	⊠ ☆	III\ 🗊 🔹 ≡
nifi □ ↓ □ 0 1 / 24 bytes 0 0 0 1 / 124 bytes 0		Add Controller Service	27:22 MDT	

CI A		ee ee n00n 🖛	= /			
iles Con	figuration					
GENERA	CONTROLLER SERVICES					
GENERA	L CONTROLLER SERVICES					
GENERA	CONTROLLER SERVICES					
GENERAI	CONTROLLER SERVICES	Туре	Bundle	State	Scope	[
GENERAJ		Type CSVReader 1.11.3	Bundle org.spache.nifi - nifi-record-serializati	State * Disabled	Scope Files	¢ † ±



_		Nif	Flow - Mozilla Firefox		(
🎄 NiFi Flow 🛛 🗙 🚽	F				
← → ♂ ☆	0	localhost:9300/nifi/?processGroupId=e8d1d908-0170	1000-1358-69b6af6cf8ad&componentIds=e8d53948-0170-	-100 ···· 🗵 🏠 🔤	•
nifi 🖸 🖓		igi izi 📽 🛼 🗐//			=
₩ 0 🗎 1,166 / 147.51 KB		Processor Details		2:49:02 MDT	Q
		Running	🕑 ST(DP & CONFIGURE	
∂ Navigate QQ∑ :		SETTINGS SCHEDULING PROPERTIE	COMMENTS	In the second se	5 min 5 min 5 min
		Property	Value	Tasks/Time: 0 / 00:00:00.000 Name matched Quorued 0 (0 bytes)	5 min
			flowfile-attribute auto-detect	DueryRecord CueryRecord 1.1.3 cog apache affi- affi standard ear in (0 bytes)	
			ignore empty string	Read/Write 0 bytes Out 0 (0 bytes) Tasks/Time: 0 / 00:00:00.000	
) Operate	8	-9-	\$.age \$.city		
EvaluateJsonPath Processor		filename	\$.name	PutFile PutFile PutFile 1.11.3 og apache nfl- nfl-standard ear	
8d53948-0170-1000-6460-3eab86cd5c32			\$.name \$.zip	► In 0 (0 bytes) Read/Write 0 bytes / 0 bytes Out 0 (0 bytes) Tasks/Time 0 / 00.00:000	

Chapter 4: Working with Databases

Pg.Admin File v Object v Tools v	Help 🗸							
	D		_	_	_		_	
✓	🔠 use	ers						×
✓ I MyPostgreSQL	Genera	al Columns C	onstraints Ad	vanced Parame	eters Securit	ty SQL		
✓ ■ Databases (2)	Inhorit	ted from table(s)	Select to inh	avit frame				
✓ Set dataengineering	Innen	ted from table(s)	Select to inn	ent from				
> 🐼 Casts > 💖 Catalogs	Colun	nns						+
 Catalogs Catalogs Cent Triggers 								
> D Extensions		Name	Data type	Length/Pr	ecision Scal	le No	t NULL? Pri	mary key?
> 🍧 Foreign Data Wrappers	C ti	name	text			1	Yes	No
> 🤤 Languages			lexi	•				
✓ Schemas (1)	C 🛍	id	integer	•			Yes	Yes
✓ ♦ public		street						
 ABUCOllations ABUCOLLATIONS ABUCOLLATIONS 	ш	Sueer	text	•			Yes	No
 Tornains TS Configurations 	C 🛍	city	text	•		1	Yes	No
 IN FTS Dictionaries 								
> Aa FTS Parsers	C Ó	zip	text	•			Yes	No
> 💽 FTS Templates								
> 🖷 Foreign Tables								
 > (c) Functions > Real Materialized Views 								
King Materialized Views () Procedures	1							
> 1.3 Sequences								
✓ I Tables (1)	i	?			×	Cancel	🗘 Reset	🖺 Save
✓ 目 users		▶ 1100 postg				::1 2		
File Obj Browser Canguages	ect 🗸		p ✓ hboard Pr	operties S0	QL Statis	tics De	pendencies	5
		B		i Q ~	2 -	6 0		т
✓ Schemas (1)		-57	public user	s/dataengine	aring/postg	ree@MvD	l 02etare	
✓ ♦ public		ୟ	public.users	s/uataengine	ening/postgi	les@iviyP	USIGIESQL	
> AU Collatio		Que	ry Editor C	uery History)				
> 🏫 Domair		1	SELECT A	FROM pub	lite users			
Ib FTS Co	-	ions	SELECT A	FROM PUD	LIC.users	•		
> 🕅 FTS Die	tionarie	s Z						
> AaFTS Pa	rsers							
> 🔞 FTS Tei	nplates							
> 🖷 Foreign	-							
> (iii) Functio								
> 💽 Materia								
		CW3						
> (() Proced								
> 1.3Sequen								
✓		Date	a Output E	xplain Mes		otification		
🗸 🖽 use				vpiairi ivies	sages No	Juncation	0	
> 咱	Columns	S	name 🎤 id	1:	street 🎤	city 🎤	zip	
> •4	Constrai	ints	text 🖁 [PK	(] integer 🥤	text 🕈	text 🕈	text "	_
> 🚬	ndexes	1	Big Bird	1	Sesame	Fakeville	12345	

- 🕨 🥅 Rules
- ➤ Iriggers

g Admin File 🗸 🕐	Object 🗸 🛛 Tools 🗸	Help) ~																					×.	postgr	es@localh	ost
rowser	\$ # %	Dash	board	Prop	erties	SQI	L Sta	tistics	D	epende	ncies	Dep	endents	E	III public	.us	ers/	dat	aengi	ineer	ring/p	oos	tgres@	MyF	Postgre	eSQL	
🗸 🔶 public		B	•		Q	~	@ ▼		Û	@ v	T	~	No limit	it	-		►	~	8		~		c (0)		. ~	±	
> 🔒 Coll		ୟ	public	users/d	lataen	gineer	ing/po	stgres(@My	Postgre	SQL																
> 🏫 Dor	nains Configurations	Quer	y Edito	r Que	ery Hist	ory																			Scrat	ch Pad	
	Dictionaries	1	SELE	CT * F	PON 1	oubl.	ie us	0.55																			
> AaFTS		2	SELE		ROM	publ	ic.us	ers																			
> 🔞 FTS	Templates	-																									
> 🖷 For > 🅼 Fun	eign Tables	Data	Outpu	t Exp	lain	Mess	ages	Notifi	catio	ons																	
> 📴 Mat	erialized Views		nan text			/ id [P	 'K] integ	er 🖋	sti tex	reet kt					city text							S	zip text				
> (() Pro > 1.3Seg		1	Big	Bird				1	I Se	same Str	eet			F	Fakeville								12345				
✓ 🛱 Tab		2	Clar	ence Cof	fey			2	2 23	7 Danielle	Spur			E	East Richa	rd							83211				
✓ (1)	.,	3	Mar	garet Ale	xander			3	3 41	682 Wilso	on Squa	re		1	Nelsonpor	t							05479				
>	🛱 Columns	4	Amy	Gordon				4	4 47	03 Vasqu	ez Strea	m		F	Pattersonf	fort							76184				
	► Constraints	5	Sier	ra Smith				5	5 74	28 Goodr	nan Par	kways	Suite 306	1	West Ryan	wille	е						24869				
	ሕ Indexes	6	Kevi	n Smith				6	5 57	5 Werner	Summit	1		١	West Jenn	ifer	rview						83115				
	■ Rules Triggers	7	Tim	othy Fitzp	atrick			7	7 81	9 Boyd Gl	en Apt.	416		ι	Leachhave	en							72697				
	ger Functions	8	Emi	y Thoma	s			8	3 67	21 Steph	ens Alle	y Apt. 2	203	ł	Harrismou	rth							96156				
> 🛅 Typ	· .	9	Kath	nleen Smi	th			9	33	71 Clay C	ourt Sui	te 160		1	Morenobu	rgh							37626				
> 🐻 Viev		10	Carr	ie Cruz				10	25	03 Cheryl	Keys				Joneshave	en							22439				
> 🍔 postgres		11	Sha	nnon Joh	nson			11	35	29 Aman	da Drive	s			Jesushave	en							18214				
> 🚣 Login/Group Role	s	12	Elair	ne Molina	1			12	2 22	948 Stepl	nanie Ho	ollow A	pt. 242	ł	Kennedyha	avei	n						85691				
› 눰 Tablespaces																											

	D Management / Index pa	tterns	Ø
©	Elasticsearch Index Management Index Lifecycle Policies	Index patterns ②	① <u>Create index pattern</u>
	Rollup Jobs Transforms Remote Clusters	Q Search	
â	Snapshot and Restore License Management	Pattern 1	
ଛ ବ୍ର	8.0 Upgrade Assistant	kibana_sample_data_ecommerce Default Rows per page: 10 ~	< 1 >
50	Kibana		-
2	Saved Objects Spaces		
5 P	Reporting Advanced Settings		
θ.			
⇒			
_			

K	D	Discover			0 2	
C	New	Save Open Share Ins	pect			
Ø	• ~	Search		KQL	ਤ Refresh	
企	-	+ Add filter				
50	users'	<u> </u>	source			
¢	🗇 Fil	earch field names Iter by type 0	<pre>> name: Ronald Goodman street: 36949 Wendy Fords Suite 051 city: Brendaburgh zip: 44959 _id: ozYlOHEBxMEH3Xr-kQMc _type score: 0</pre>	: doc _inde	x: users	-
2	selecte		> name: Matthew Lewis street: 78761 Dana Curve city: Wilsonborough zip: 58312 _id: pDY10HEBxMEH3Xr-2QPk _type: doc _in	dex: users	_score: 0	
۲ <u>۵</u>	Availabl t_id		<pre>> name: Ryan Lowe street: 740 Smith Pine Suite 065 city: Lake Jamesberg zip: 38837 _id: pTYoOHEBxMEH3Xr-PgMT _type: doc _score: 0</pre>	_index: us	sers	-
3	t_ind		> name: Christopher Thompson street: 6099 Ian Walk city: Johnsonside zip: 43192 _id: pjYoOHEBxMEH3Xr-PgMT _type: doc _	index: users	s _score: @	0
ş B	# _sc t _typ t city	pe	> name: Evelyn Phillips street: 9075 David Oval city: South Randyhaven zip: 81852 _id: pzYoOHEBxMEH3Xr-PgMT _type: doc _score: θ	_index: use	ers	-
- &	t nar t stre	ne	<pre>> name: Tricia Mcmillan street: 8077 Nancy Mills Apt. 810 city: Jamesberg zip: 63792 _id: qDYoOHEBxMEH3Xr-PgMT _type: c _score: 0</pre>	loc _index:	users	-
						-
2	Airflow	DAGs Data Profilino	y ▼ Browse ♥ Admin ♥ Docs ♥ About ♥ 2	2020-04-04 16	5:41:15 UTC	ł
Ma	rked suc	cess on 1 task instances			×	
Off	DAG	: MyDBdag		sche	dule: 0:05:0	0
	Graph Vi	ew 🕈 Tree View 🏦	Task Duration 🔹 Task Tries 🛧 Landing Times 🖹 Gantt 🔚 Details 🗲 Code 📀 Trigger DAG 📿 Refresh 🤒 D	elete		
runnin	g Base	date: 2020-04-02 01:15:01	Number of runs: 25 VRun: scheduled_2020-04-02T01:15:00+00:00 VLayout: Left->Right V Go St	earch for		
	-					
Python	Operator		success (running) tailed (skipped) (upstream_tailed) (up_for_reschedule) (up_for_reschedule)	p_for_retry que	eued) no_statu	ıs
					C]
			QueryPostgreSQL InsertDataElasticsearch			
R	Airflo	w DAGs Data Pro	filing 🗸 Browse 🗸 Admin 🖌 Docs 🗸 About 🗸 202	:0-04-04 01::	16:38 UTC	
Th	e DAGs	s list may not update, and	new tasks will not be scheduled.			Γ
On) DA(G: MyDBdag		sched	ule: 0:05:00	0
•	Graph \	View 🕈 Tree View	ul Task Duration 🚯 Task Tries 🗼 Landing Times 🖹 Gantt 🔠 Details 🗲 Code ⊙ Trigger DAG 📿	Refresh	Delete	
Base	date:	2020-04-02 01:15:00	Number of runs: 25 Go			
ОРу	/thonOper	ator	success 📕 running 📕 failed 🛄 skipped 📕 upstream_failed 🛄 up_for_reschedule 📕 up_for	retry 📕 queue	d 🗌 no_statu	JS
			the Co. Ma			
Olo	AG]		o ⁶ 2 ^m 0 ⁶ 2 ⁰ 0 ⁷ m ²			
- \	Query	PostgreSQL IsertDataElasticsearch				
	0					

	D Discover	0
٩	frompostgresql* V	2,208 hits
Ø	Q Search field names	
ſ	 Filter by type Selected fields 	> Unnamed: 0: 1 name: Clarence Coffey city: East Richard _id: jza8QnEBxMEH3Xr-SQfP _type: doc _index: frompostgresql _score: 0
50	source	> Unnamed: θ: 2 name: Margaret Alexander city: Nelsonport _id: kDa8QnEBxMEH3Xr-SwdU _type: doc _index: frompostgresql _score: θ
<u>ش</u> چ	Available fields # Unnamed: 0	> Unnamed: 0: 3 name: Amy Gordon city: Pattersonfort _id: kTa8QnEBxMEH3Xr-TAfC _type: doc _index: frompostgresql _score: 0
2	t_id	> Unnamed: 0: 4 name: Sierra Smith city: West Ryanville _id: kja8QnEBxMEH3Xr-Tgc8 _type: doc _index: frompostgresql _score: 0
5	t _index # _score	<pre>> Unnamed: 0: 5 name: Kevin Smith city: West Jenniferview _id: kza8QnEBxMEH3Xr-Twes _type: doc _index: frompostgresql _score: 0</pre>
2	t _type	<pre>Unnamed: 0: 6 name: Timothy Fitzpatrick city: Leachhaven _id: 1Da8QnEBxMEH3Xr-UAeC _type: doc _index: frompostgresql _score: 0</pre>
ڻ ه	t city t name	> Unnamed: 0: 7 name: Emily Thomas city: Harrismouth _id: lTa8QnEBxMEH3Xr-UAfX _type: doc _index: frompostgresql _score: 0
₿ G		<pre>> Unnamed: 0: 8 name: Kathleen Smith city: Morenoburgh _id: lja8QnEBxMEH3Xr-UQdtype: doc _index: frompostgresql _score: 0</pre>
ŵ		> Unnamed: 0: 9 name: Carrie Cruz city: Joneshaven _id: lza8QnEBxMEH3Xr-Ugdc _type: doc _index: frompostgresql _score: 0
⇒		> Unnamed: θ: 1θ name: Shannon Johnson city: Jesushaven _id: mDa8QnEBxMEH3Xr-Uwca _type: doc _index: frompostgresql _score: θ

🇱 0 🗮 1,166 / 147.51 KB 💿 0	▶ 2 ■ 19 🔺 0 🦎 0 🛹 0 🗰 0	○ 0 0 0 ? 0 20:33:48 MDT Q
🔊 Navigate 🖂	ExecuteSQLRecord ExecuteSQLRecord 1.11.3 org.apache.nifi - nifi-standard-nar	
Q Q []!	In 0 (0 bytes)	5 min
	Read/Write 0 bytes / 233.05 KB	5 min
	Out 4 (233.05 KB)	5 min
	Tasks/Time 4 / 00:00:00.512	5 min
	Name success	
	Queued 0 (0 bytes)	
	SplitText SplitText 1.11.3 org.apache.nif - nif-standard-nar	
	In 4 (233.05 KB)	5 min
C Operate	Read/Write 233.05 KB / 0 bytes	5 min
🖞 Operate 📃	Out 4,004 (229.15 KB)	5 min
postgresToelasticseach	Tasks/Time 4 / 00:00:02.670	5 min
Process Group	Name splits	
42f4030d-0171-1000-166d-3a3bdb8d120b	Queued 0 (0 bytes)	
	PutElasticsearchHttp PutElasticsearchHttp TutElasticsearchHttp TutElasticsearch-ntr org.apache.ntfi-ntfi-elasticsearch-nar	
	In 4,004 (229.15 KB)	5 min
	Read/Write 229.15 KB / 0 bytes	5 min
	Out 0 (0 bytes)	5 min
	Tasks/Time 44 / 00:00:36.778	5 min

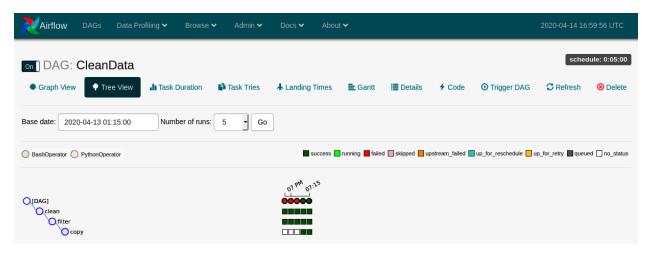
\leftrightarrow > C $$	0 (i) localhost:9300/nifi/?processGroupId=42f4030d-0171-10	00-166d-3a3bdb8d120b&componentIds=42f4a28f-0	71-1000 ••• 🗵	☆ III	\ □ \$ =
nifi 🖸 🕹 🛛	isi izi 🏆 🗣 🗐 🖉				
🎎 0 🗮 1,166 / 147.51 KB	O O Configure Processor			13:26 MDT	Q 🗆
	A Invalid				
Navigate					
Q Q []!	SETTINGS SCHEDULING PROPERTIES	COMMENTS			
	Required field		+		
	Property				
	Database Connection Pooling Service	No value	✓		
	SQL Pre-Query	No value			
	SQL select query	Create new service	1 - I		
	SQL Post-Query				
🖒 Operate	Max Wait Time 0	0 seconds			
•	Record Writer	No value set			
ExecuteSQLRecord Processor	Normalize Table/Column Names 0	false			
42f4a28f-0171-1000-268a-24fa2d49be8b	Use Avro Logical Types O	false			
	Max Rows Per Flow File O Output Batch Size O	0			
	Output Batch Size O Fetch Size O	0			
	Fetch Size	0			
	Tople Come of	CANCEL	APPLY		
NIFI Flow » postgresToelasticseach					

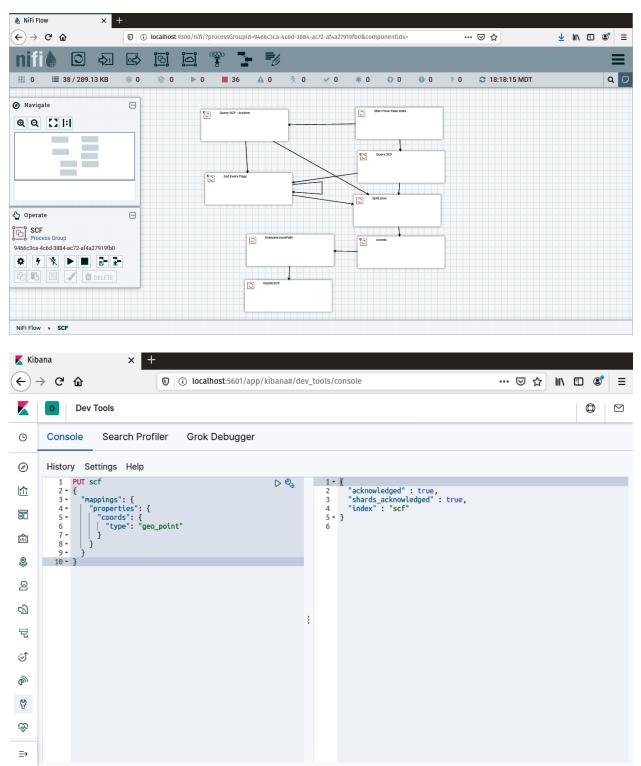
ostgres	Toelasticseach Configu	Controller Service Details				1
GENERA	L CONTROLLER SERVICES	SETTINGS PROPERTIES	COMMENTS		-	
		Required field Property		Value		•
	Name 🔺	Database Connection URL	Θ	jdbc:postgresql://localhost/dataengineering		
2	DBCPConnectionPool	Database Driver Class Name	0	org.postgresql.Driver	esToelasticseach	\$ \$
	JsonRecordSetWriter	Database Driver Location(s)	0	/home/paulcrickard/nifi-1.11.3/drivers/postgresql-42.2	esToelasticseach	0 %
		Kerberos Credentials Service	0	No value set		
		Database User	0	postgres		
		Password	0	Sensitive value set		
		Max Wait Time	0	500 millis		
		Max Total Connections	0	8		
		Validation query	0	No value set		
		Minimum Idle Connections	Θ	0		
		Max Idle Connections	Θ	8		
		Max Connection Lifetime	0	-1		
		Time Between Eviction Runs	0	-1		

ostaresTa	elasticseach Configu	Controller Service Details				1
GENERAL	CONTROLLER SERVICES	SETTINGS PROPERTIES Required field	COMMENTS			
		Property		Value		
	Name 🔺	Schema Write Strategy	0	Do Not Write Schema		
B	DBCPConnectionPool	Schema Cache	0	No value set	esToelasticseach	\$ %
_ D	JsonRecordSetWriter	Schema Access Strategy	0	Inherit Record Schema	esToelasticseach	• *
		Schema Registry	0	No value set		
		Schema Name	0	\${schema.name}		
		Schema Version	0	No value set		
		Schema Branch	0	No value set		
		Schema Text	0	\${avro.schema}		
		Date Format	0	No value set		
		Time Format	0	No value set		
		Timestamp Format	0	No value set		
		Pretty Print JSON	0	false		
		Suppress Null Values	0	Never Suppress		
		Output Grouping	0	One Line Per Object		

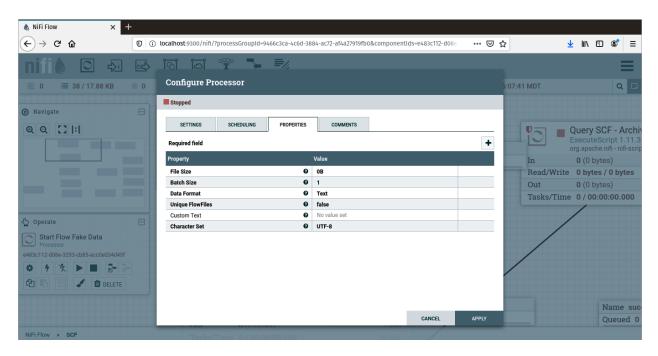
yellow	open	frompostgresql	u3EVJEFJR-eDPtrufvSNkA	1	1	2208	0	250.4kb	250.4kb
yellow	open	fromnifi	NBAL_aLdQ1y4iK90kDFauw	1	1	1001	0	233.8kb	233.8kb
yellow	open	test	VWacjWq_S9a0fW6l-YbcIA	1	1	1	0	3.9kb	3.9kb
green	open	kibana_sample_data_ecommerce	-6sdGU13T12GIjE2rU_ZlQ	1	Θ	4675	0	4.9mb	4.9mb
green	open	.kibana_task_manager_1	1nLCVInfTa6XmmifoGntXA	1	Θ	2	0	13kb	13kb
green	open	.apm-agent-configuration	51toZCyRS2yWlnLJZWF5-A	1	0	Θ	0	283b	283b
green	open	.kibana 1	g5gxKq0xR0y71FCG0rKiMw	1	Θ	66	5	993.3kb	993.3kb
yellow	open	users	8K7w0mQ4S90gxh2TqBybKw	1	1	1003	0	286.6kb	286.6kb
-	•								

Chapter 5: Cleaning and Transforming Data

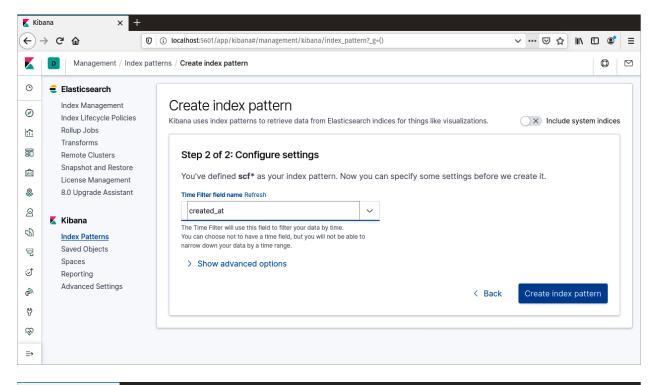




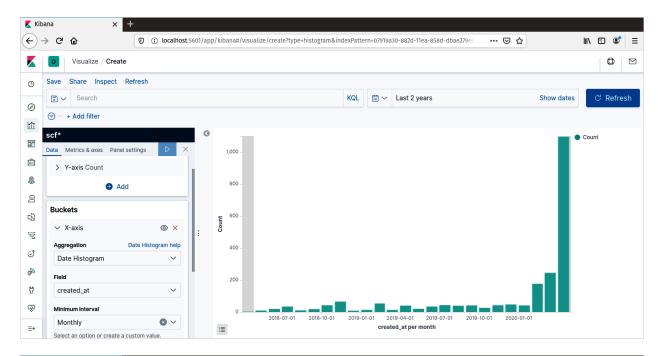
Chapter 6: Building a 311 Data Pipeline



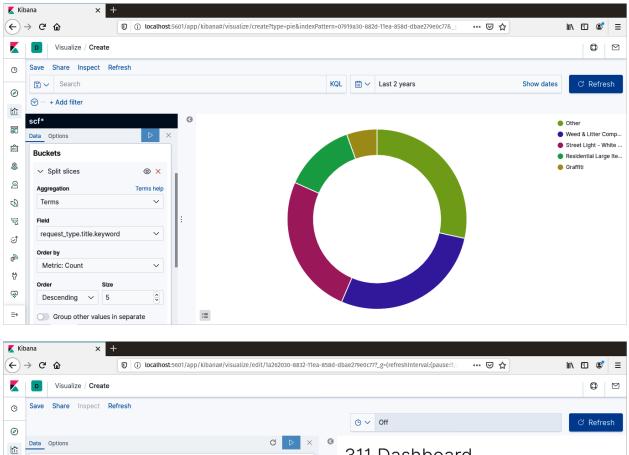
📕 Kib	ana × +	
€-	→ C û	🛛 🛈 localhost:5601/app/kibana#/management/kibana/index_pattern?_g=() 🖂 🏠 📗 🖉 🗧
	D Management / Index pa	atterns / Create index pattern
©	🗧 Elasticsearch	
@ £	Index Management Index Lifecycle Policies Rollup Jobs	Create index pattern Kibana uses index patterns to retrieve data from Elasticsearch indices for things like visualizations.
5	Transforms Remote Clusters Snapshot and Restore	Step 1 of 2: Define index pattern
<u>ه</u>	License Management 8.0 Upgrade Assistant	Index pattern Scf*
<u>ල</u> ක	Kibana	You can use a * as a wildcard in your index pattern. > Next step You can't use spaces or the characters /, ?, *, <, .
ē	Saved Objects Spaces	scf
ୖ	Reporting Advanced Settings	Rows per page: 10 ${\sim}$
8 P		
æ		
⇒		



👗 к	ibana	×	+							
¢	→ C' û		0	 localhost:5601/ap 	op/kibana#/vi	sualize?_g=(refr	eshInterval:(pause:	t,value:0),time:(from:now-2y,to:now)) … 🗵 🛧	lii\ 🗉 📽	≡
	DVi	sualize							Ø	
Ŀ				New Visu	alizatio	n		×		
Ø						··		alization		
盒			_	Q Filter				Select a visualization type		
50			Q	B ه		<u>∽</u> E		Start creating your visualization by selecting a type for that visualization.		
<u>í</u>				Lens	Area	Controls	Data Table	Try Lens, our new, intuitive way to create		
٨				0	6	•0		visualizations.		
2				Gauge	Goal	eo Heat Map	Horizontal Bar	Go to Lens 🖸		
5				00090		Troat map				
7				\mathbb{N}	® 2	$\{\tilde{\mathbf{T}}\}$	8	0		
Í				Line	Maps	Markdown	Metric			
P					0		67	0		
9				œ	<u></u>		۲			



📕 Kib	ana	×	+									
€-	→ Cª	ŵ	I i localhos	:5601/ap	pp/kibana#/visualize/create?type=metric&indexPat	tern=0	7919a30-	882d-11ea-858d-dbae279e0c77 🛛 💀 😒 🏠		111\ (D®	≡
	D	Visualize / Crea	te								٥	
©	Save 3	Share Inspect	Refresh							_		
Ø	•	Search			1	QL	⇒	Last 2 years	Show dates	e	Refre	esh
	;; - +	Add filter										
_	scf*			G								
80	Data O	ptions	\triangleright									
â	Metric	cs										
8	~ M	letric										
2	Aggre	gation	Count help									
50	Cou	nt	~					176				
ē		m label		÷				2,176				
đ	Issu	es						Issues				
P	> A	dvanced										
썁		G A	dd									
ŵ	Bucke	ets										
⇒		• A	dd									



Help 🕑

Markdown

311 Dashboard

* It has coordinates

* Only for Bernalillo County

* Can be filtered by Created Date

This dashboard displays data collected by Nifi at 8 hour intervals. It queries the

SeeclickFix API, and was backfilled with Archived data.

37

ŵ

٩

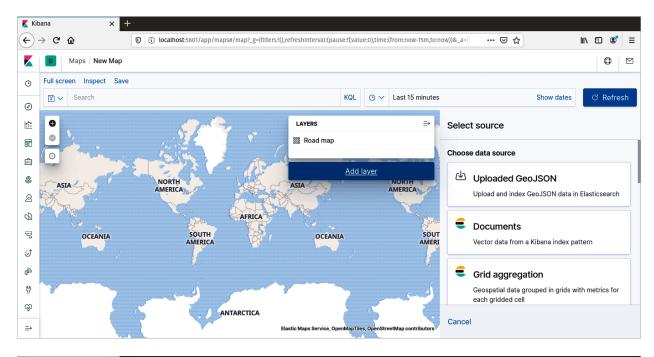
2

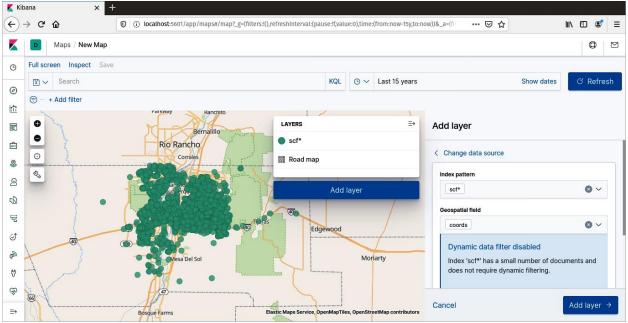
50

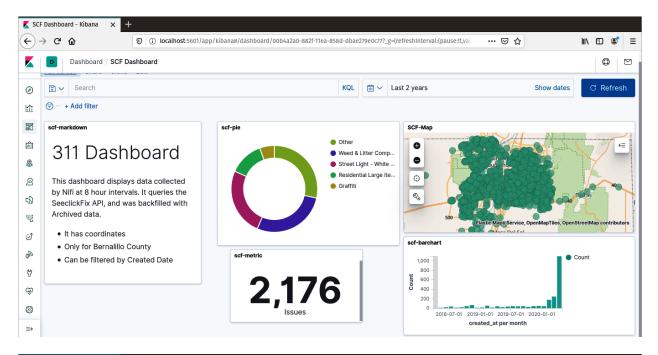
311 Dashboard

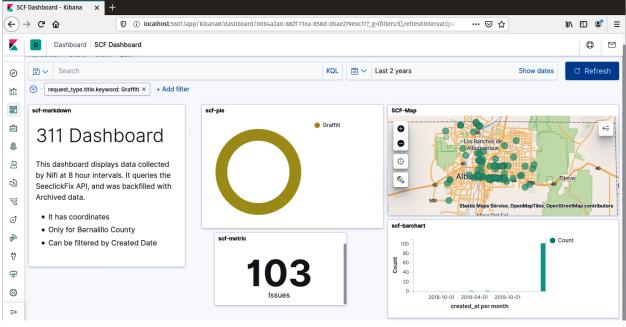
This dashboard displays data collected by Nifi at 8 hour intervals. It queries the SeeclickFix API, and was backfilled with Archived data.

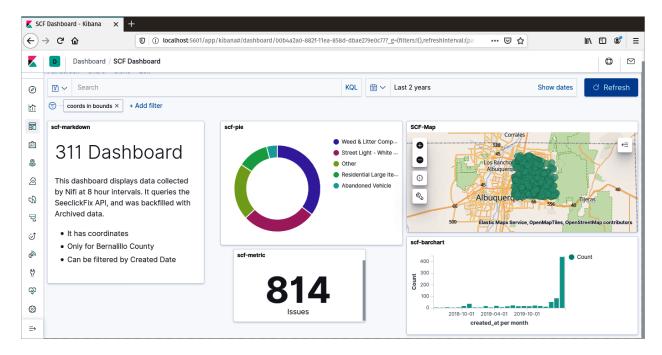
- It has coordinates
- Only for Bernalillo County
- Can be filtered by Created Date

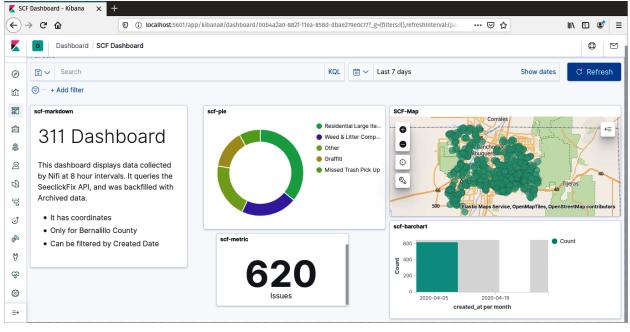


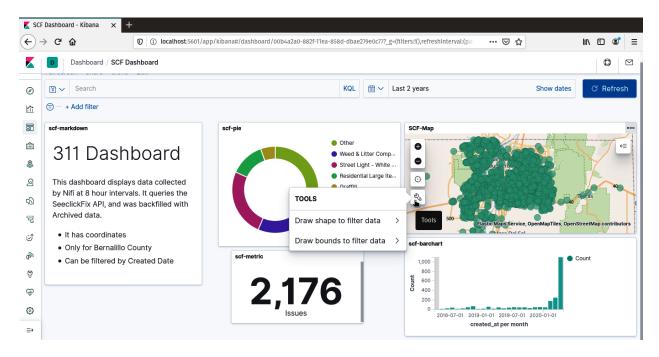


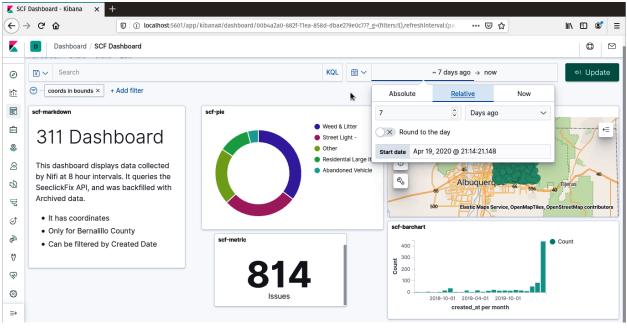




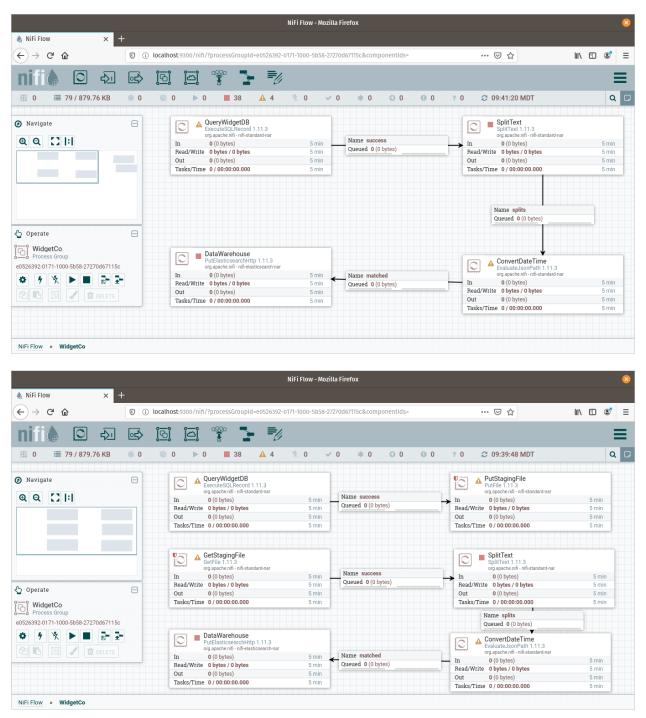




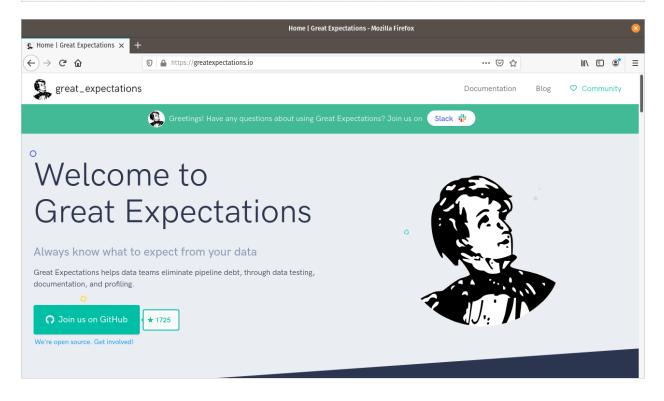


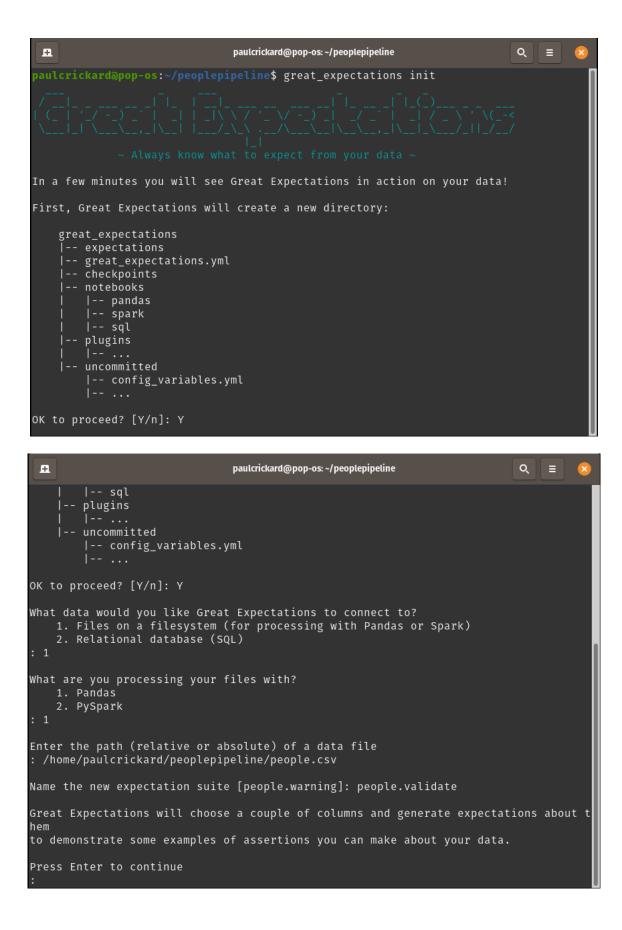


Chapter 7: Features of a Production Data Pipeline



			NiFi Flow -	Mozilla Firefox					8
🍖 NiFi Flow 🗙 🕂	-								
← → ♂ ☆	0) localhost:9300/nifi/?processGroup	old=e0526392-0171-1000-51	58-27270d67115c&componentl	ds=	… ⊘ ☆	∭\ 🖽	۲	≡
nifi 🖸 🕹	œ\$		- - //					Ξ	=
3 0 🗮 79 / 879.76 KB	0	◎ 0 ▶ 0 ■ 39	🔺 6 🦎 O	✓ 0 * 0 ⊙	0 0 0 ? 0	C 10:19:40 MDT		۹	
 Navigate Q Q [] : 	8	Read/Write 0 bytes / 0 bytes Out 0 (0 bytes)	S man S man S man S man	PutSitesjingFile PutSite 1113 organizational statement of the	3 min 3 min 3 min 3 min				
		Pead/Write 0 bytes / 0 bytes 0 to 0 to 0 (0 bytes)	min Decord 0 (0 tytes)	SplitText splittest SplitText splittest SplitText splittest In 0 (2 bytes) Read/write & bytes / 0 bytes / Out 0 (2 bytes) Traike/Time & / 00:00:00:000	5 mm 5 mm 5 mm 5 mm 5 mm	Convertigue Trans Transaction of other states Transaction of the states	S min S min S min S min		
Operate	Θ					Queued @ (0 bytes) StagingDataWarehouse PutElsaticsearchHttp: 1.1.3 ag. aparter of reflexitionercherue Ta 0 (0 bytes) Read/Write 0 bytes	5 min 5 min		
WidgetCo Process Group e0526392-0171-1000-5b58-27270d67115c						Out 0 (Dayles) Out 0 (Dayles) Tasks/Time 0 / 80:00:00.000 Name success Queued 0 (Obyles)	5 min 5 min		
		DataWarehouse ung automatical information of the statistical informating statistical information of the statistinter of the sta	in Queued 0 (0 bytes)	A Validation Evaluate.icooPath 1113 evaluate.icooPath 1113 evaluate.icooPath 1113 evaluate.icooPath 1113 Read/Write 0 bytes / Cot 0 (0 bytes / Tasks/Time 0 / 00 000.000	S min D min D min S min	A QueryStagingDataWarehouse Example 2011 11.3 a 0 (0.0)ref ReadWite 0 bytes / 0 bytes Out 0 (0.0)refs) Tasks/Time 0 / 00:00:00.000	5 mis 5 mis 5 mis 5 mis		
NiFi Flow » WidgetCo									





		Data documentation compile	ed by Great Expectations	- Mozilla Firefox	
Data documentation compi \times +					
-) → C' û	① file:///hom	e/paulcrickard/peoplepipeline/great_expectation	ns/uncommitted/data_do	ocs/local_site/validations/people 🚥 🗵 🕅 🛙) ©
great_expectations	Home / pe	ople.validate / 20200504T195933.61497	6Z / 6f1eb7a06079eb	9cab8de404c6faab62	
Expectation Validation Result Evaluates whether a	-	ew tion Suite: people.validate Succeeded			
batch of data matches	Statistics				
expectations.		Expectations	11		
Actions	Successf	ul Expectations	11		
	Unsucces	sful Expectations	0		
Validation Filter:	Success	Percent	100%		
Show Failed All Only	Show mor	e info			
How to Edit This Suite	Table-	Level Expectations			
Show Walkthrough	Status	Expectation		Observed Value	
	0	Must have between 900 and 1100 rows	ð.	1000	
Table of Contents	I	Must have exactly columns.		8	

Status	Expectation	Observed Value
•	Must have between 900 and 1100 rows.	1000
•	Must have exactly B columns.	8

Must have these columns in this order: name, age, street, city, state, zip, lng, lat ['name', 'age', 'street', 'city', 'state', 'zip', 'lng', 'lat']

age

Status	Expectation	Observed Value
•	values must never be null.	100% not null
•	minimum value must be between 17 and 19.	18
•	maximum value must be between 79 and 81.	80
•	mean must be between 49.151 and 51.151.	50.151

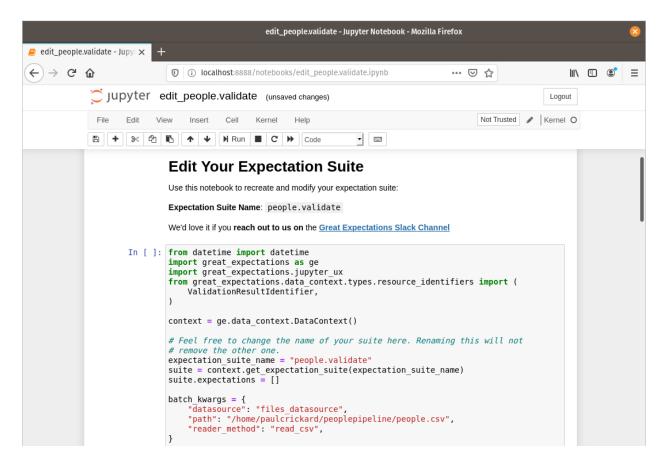
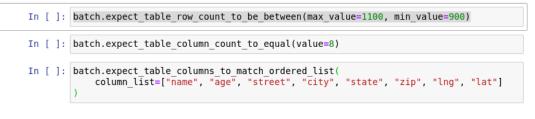
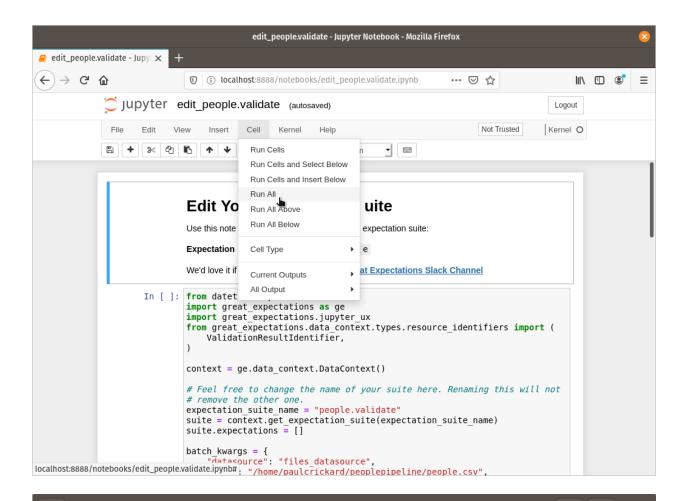


Table Expectation(s)



age

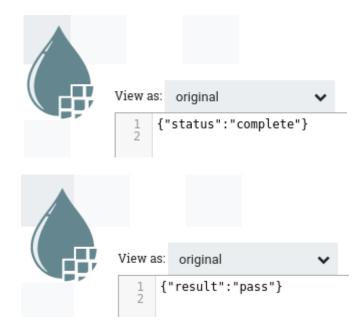
In []: batch.expect_column_values_to_not_be_null("age")
In []: batch.expect_column_min_to_be_between("age", max_value=19, min_value=17)
In []: batch.expect_column_max_to_be_between("age", max_value=81, min_value=79)
In []: batch.expect_column_mean_to_be_between("age", max_value=51.151, min_value=49.151)
In []: batch.expect_column_median_to_be_between("age", max_value=52.0, min_value=50.0)
In []: batch.expect_column_quantile_values_to_be_between(
 "age",
 quantile_ranges={
 ["quantiles": [0.05, 0.25, 0.5, 0.75, 0.95],
 "value_ranges": [[21, 23], [34, 36], [50, 52], [64, 66], [76, 78]],
 },
]



paulcrickard@pop-os: ~/peoplepipeline
Q =
paulcrickard@pop-os: ~/peoplepipeline\$ great_expectations tap new people.validate
peoplevalidatescript.py
This command will be removed in the next major release in favor of a similar fea
ture `tap script`
Enter the path (relative or absolute) of a data file
: /home/paulcrickard/peoplepipeline/people.csv
A new tap has been generated!
To run this tap, run: python /home/paulcrickard/peoplepipeline/peoplepipe

paulcrickard@pop-os:~/peoplepipeline\$ python3 peoplevalidatescript.py
Validation Succeeded!

			NiFi	Flow - Mozilla Firefox					8
🎄 NiFi Flow 🛛 🗙 🕂									
\leftrightarrow > C $$	🚺 🤅 loca	alhost:9300/nifi/?processGroupId	=0171101d-7595-1a	160-d847-b6b18e1534a5	&componentIds=	•••	⊠ ☆	∭\ ≞	: ≡
nifi 🖸 🕹	<u>محکم أو</u>	<u>di ioi 🊏 🏅</u>	. <i>≡</i> //						≡
🇱 0 🗮 79 / 879.76 KB	00	0 🕨 0 📕 44	A 8 🕺	0 ~ 0 *	0 0 0 0	?02	08:37:31 MDT		Q L
Ø Navigate	Θ	Creat 1000 People					Generate FlowFile 1.11.3		
ଷ୍ ସ୍ 🕻 🛛 :		ExecuteStreamCommand organiche.nfi-nfi-andurie In 0 (0 bytes) Read/Write 0 bytes / 0 bytes Out 0 (0 bytes)	1.11.3 # 5 min 5 min 5 min		Name success Queued 0 (0 bytes)		GenerateFlowFile 1.11.3 org.asade.off -off-standard-or In 0 (0 bytes) Read/Write 0 bytes /0 bytes Out 0 (0 bytes)	5 min 5 min 5 min	
		Tasks/Time 0/00:00:00.000 Name output stream	5 min				Tasks/Time 0/00:00:00:000	5 min	
		Queued @ (0 bytes)	1.11.3		Extract Valids				
		erg apacte of or for standards In 0 (0 bytes) Read/Write 0 bytes / 0 bytes Out 0 (0 bytes)	5 min 5 min 5 min	Name output stream Queued 0 (0 bytes)	In 0 (0 bytes) Bead/Write 0 bytes / 0 by Out 0 (0 bytes)	5 min 5 min 5 min			
		Tasks/Time 0/00:00:00.000	5 min		Tasks/Time 0/00:00:00.0				
Coperate	Θ	Patsed PutFile 1.11.3 og apsche offi- off-standarde			Name metcheo Queued 0 (0 b)				
Validate Process Group 0171101d-7595-1a60-d847-b6b18e1534a5		In 0 (0 bytes) Read/Write 0 bytes / 0 bytes Out 0 (0 bytes) Tasks/Time 0 / 00:00:00:000	5 min 5 min 5 min 5 min	Name pass Queued 0 (0 bytes)	RouteOnAttre				
 ♦ ♥ X ► ■ ► Ξ 					In 0 (0 bytes) Read/Write 0 bytes / 0 b Out 0 (0 bytes) Tasks/Time 0 / 00:00:00	5 min			
		PutFie 1.11.3 org. apache.offs. nfi-standarde		Name fail Queued 0 (0 bytes)					
		In 0 (0 bytes) Read/Write 0 bytes / 0 bytes Out 0 (0 bytes) Tasks/Time 0 / 00:00:00.000	5 min 5 min 5 min 5 min						



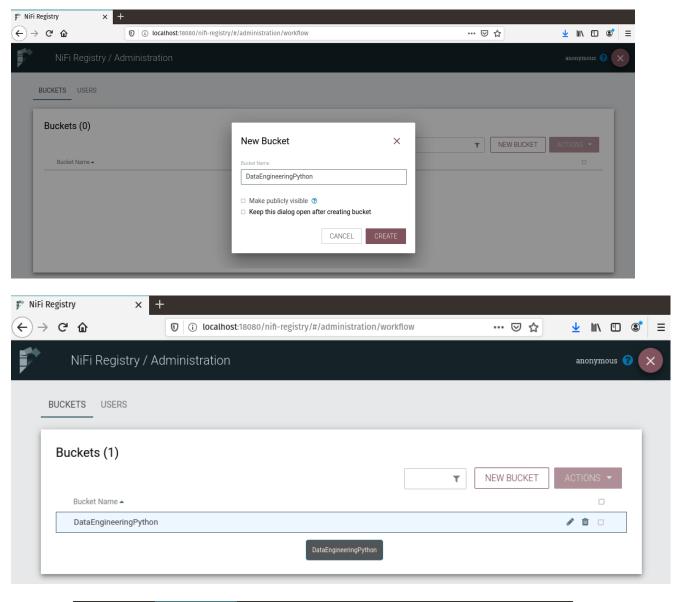
	Da	ta Docs created by Great Expectations - Mozilla Firefox	6
Typical Workflow — great_e 🗙 📔 e	dit_people.validate - Jupy 🗙 🧕 Data Do	ccs created by Great × +	
- → C' û	i) file:///home/paulcrickard/peoplepipelir	ne/great_expectations/uncommitted/data_docs/local_site/index.html 🚥 🗟 🕪 🗈	•
great_expectations			
Data Docs autogenerated using Great Expectations.	Data Docs local_site		
Actions	Expectation Suite	Validation Results (run_id)	
Validation Filter: Show Failed All Only Show Walkthrough	people.validate	 × 20200505T145722.862661Z × 20200505T145528.788769Z 20200505T144559.641868Z 20200505T021322.913702Z 20200505T020553.174834Z 20200505T012823.671147Z 20200505T011831.865106Z 20200504T224236.055072Z 20200504T204504.344872Z 20200504T195933.614976Z 	

		Data documentation compiled by Great Expectations - Mozilla I	Firefox
Typical Workflow — great_e 🗙 🏻 💋	edit_people.validate	- Jupy × 🕵 Data documentation compi × 🕂	
) → C' 🏠	i) file:///home/par	ulcrickard/peoplepipeline/great_expectations/uncommitted/data_docs/local_	site/validations/people 🚥 🖂 🛣 🗊
great_expectations	Home / people	e.validate / 20200505T145722.862661Z / 6f1eb7a06079eb9cab8de	e404c6faab62
Actions	age		
Validation Filter:			
Show Failed	Status	Expectation	Observed Value
All Only	0	values must never be null.	100% not null
☑ How to Edit This Suite	×	minimum value must be between 17 and 19.	1
Show Walkthrough	×	maximum value must be between 79 and 81.	100
	•	mean must be between [49.151] and [51.151].	51.111
Table of Contents	•	median must be between 50.0 and 52.0.	51
Overview			
Table-Level Expect ations	name		
age	Status	Expectation	Observed Value
name	\bigcirc	values must never be null.	100% not null

A NITI Flow			NiFi Fl	ow - Mozilla Firefox				
$ \underbrace{(A, NiFiFlow)}_{A} \rightarrow C \widehat{(A)} $	<u> </u>	i) localhost:9300/nifi/?	processGroupId=9466c3ca-4c6d-38	84-ac72-af4a27919fb0&com	ponentids=	··· 🗵 🕁	III\ 🗉) (2)
nifi 🖸 🖓	<u></u>	الم الم	¥ 🤰 🗾					=
🗱 0 🔳 81 / 882.05 KB	• 0	© 0 ► 3	42 A 11 🕺	0 ~ 0 * 0	0 0	0 ? 0 😂 12:23:31 MDT		Q
🔊 Navigate	Θ	GenerateFlo	vFile 1.11.3			ElasticSCF PutElasticsearchHttp 1.11.3 org.apache.nifi - nifi-elasticsearch-nar		
			nifi-standard-nar	Name success		In 0 (0 bytes)	5 min	
ଷ୍ ଷ୍ 🕻 🛙 🖓		In 0 (0 bytes) Read/Write 0 bytes / 0	5 min - bytes 5 min	Queued 0 (0 bytes)		Read/Write 0 bytes / 0 bytes	5 min	
		Out 0 (0 bytes)	bytes 5 min 5 min 1			Out 0 (0 bytes)	5 min	
		Tasks/Time 0/00:00:0				Tasks/Time 0 / 00:00:00.000	5 min	
				Name suc	cess	Name feilure		
				Queued 1	(2.27 KB)	Queued 1 (2.27 KB)	_	
🗅 Operate	Θ							
SCF Process Group			PutFile 1.11.3 org.apache.nifi - nifi-standard	-nar		PutFile 1.11.3 org.apache.rifi - nifi-standard-nar		
9466c3ca-4c6d-3884-ac72-af4a27919fb0			In 0 (0 bytes)	5 min		In 0 (0 bytes)	5 mi	
🌣 🤊 🦎 🕨 🔳 📑 🚠			Read/Write 0 bytes / 0 bytes	5 min		Read/Write 0 bytes / 0 bytes	5 mi	
			Out 0 (0 bytes) Tasks/Time 0 / 00:00:00.000	5 min 5 min		Out 0 (0 bytes) Tasks/Time 0 / 00:00:00.000	5 mi	
			Tasks/Time 0/00:00:00.000	5 min		Tasks/Time 0700:00:00.000	5 mi	n

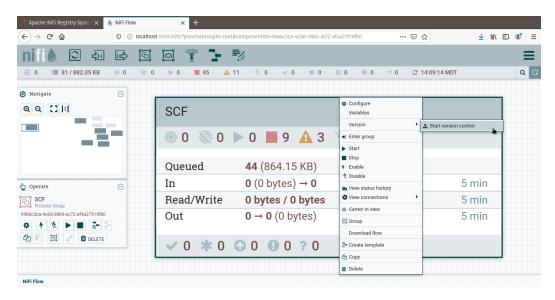
Chapter 8: Version Control Using the NiFi Registry

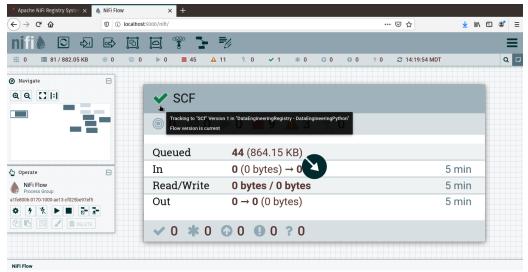
🎄 Apache NiFi - Registry	× +										
↔ → ♂ ŵ	🔽 🔒 htt	ps://nifi. apache.org /r	registry				⊽ ≰	2	١١١	1	' ≡
		Project 🤝	Documentation 👻	Downloads 🤝	Community 👻	Development 🤝	ASF Links 👻 Sub	projects 🤝			
		APACHE NI		ry							
	A subproject of	Apache NiFi to	store and man	age shared re	sources.						
	About										
	Registry—a subproject of Apache NiFi—is a complementary application that provides a central location for storage and management of shared resources across one or more instances of NiFi and/or MiNiFi.										
	Specific goals fo	or the initial thru	ust of the Regis	stry effort inclu	ude:						
🎄 Apache NiFi - Registry	× +										
← → ♂ ŵ	🔽 🔒 htt	ps://nifi. apache.org /r	egistry				⊠ ੯	2	lii\	•	' ≡
		Project 👻	Documentation 👻	Downloads 🤝	Community 👻	Development 🤝	ASF Links 👻 Sub	projects 👻			
	Links										
	 Binaries nifi-regist nifi-regist Release Not 0.5.0 Sources 	YS file can be fou ult the Migration G ry-0.6.0-source-rel ry-0.6.0-bin.tar.gz ry-0.6.0-bin.zip (a	nd here. auidance when upg lease.zip (asc, sha (asc, sha256, sha sc, sha256, sha51	grading. a256, sha512) 1512) 2)	be found here ar	nd here.					
🌮 NiFi Registry	× +										
↔ → ♂ ŵ	0 i loc	alhost: 18080/nifi-regi	stry/#/explorer/grid	-list			⊌ ☆		⊻ III\	•	≡
NiFi Regi	stry / All 🔻								anonymo	us 💡	æ
								T Sort by:	Name (a - z) 🔻	
				No results match	h this query.						



🖊 Apache NiFi Registry System 🗙	🎄 NiFi Flow	× +	
↔ ↔ ↔ ↔	I i localhos	£9300/nifi/ 😇 ✿	¥ II\ 🗊 📽 Ξ
nifil 🖸 Ð	i <u>়ি</u> (ত	i 🚏 🤰 🗾	
2 0 🔳 81 / 882.05 KB	◎ 0	▶ 0 📕 45 🛕 11 🔅 0 🛩 0 🔹 0 💿 0 0 0 ? 0 🎜 13:56:02 MD	23 Counters
Ø Navigate	Θ		Bulletin Board
€ € []		SCF	L [®] Data Provenance
		 ◎ 0 ◎ 0 ▶ 0 ■ 9 ▲ 3 × 0 Queued 44 (864.15 KB) 	Controller Settings Parameter Contexts Flow Configuration History Templates Help
👌 Operate	Θ	In 0 (0 bytes) $\rightarrow 0$	About Smin
SCF Process Group		Read/Write 0 bytes / 0 bytes	5 min
9466c3ca-4c6d-3884-ac72-af4a27919fb0		Out $0 \rightarrow 0 (0 \text{ bytes})$	5 min
🕐 🖪 🗹 🏛 delete		✓ 0 * 0 0 0 0 ? 0	
NiFi Flow			

1	Apache NiFi Registry Syster 🗙 🎄 NiFi Flow 🗙	+		
(→ C û (i) localhost:9300/nifi/?pro	ccessGroupId=root&componentIds=9466c3ca-4c6d-3884-ac72-af4a27919fb0	··· v 🕸 🚽 🕪 🗈 📽 🗉	-
ſ		0n 🗮 💻 ,	x	
	NiFi Settings			
0	GENERAL REPORTING TASK CONTROLLER SERVICES REPORTING TA	Add Registry Client		
4		Name	+	
C	Name 🔺	DataEngineeringRegistry	n	
		URL http://localhost:18080		
		Description		
		Registry for Data Engineering with Python		
_		CANCEL ADD		
	2 Last updated: 13:56:56 MDT			
N	liFi Flow			





	toppoloje andistrikularina andraja list			🚽 III\ 🗉 📽
C'û i localhos	t:18080/nifi-registry/#/explorer/grid-list		⊠ ☆	⊻ II\ 🗉 📽
NiFi Registry / All 🔫				anonymous 💡 🔒
			T Sort by	r. Name (a - z) 👻
SCf - DataEngineeringPython Flow		VERSIONS 1		^
BUCKET IDENTIFIER	CHANGE LOG			ACTIONS -
c8e69fca-7f08-44a1-967c-32a1a5709404	Version 1 - 13 minutes ago by anonymous			
FLOW IDENTIFIER				
7aa48258-ce5c-44d8-ae7a-6e67d9dbaa4c	First commit May-15-2020 at 2:10 PM			
DESCRIPTION				
Chapter 6: SeeClickFix API to Elasticsearch				

🖌 Apache NiFi Registry Syster 🗙 🏾 🎄 NiFi Flow	× +			
$\leftarrow \rightarrow$ C \triangle \bigcirc \bigcirc \bigcirc localhost:93	00/nifi/?processGroupId=9466c3ca-4c6d-388	4-ac72-af4a27919fb0∁	oonentids=1a1026fd-0172- 🚥 🗵 🛧	⊻ II\ 🗊 📽 🗏
nifi 🖉 🖓 🖙 📴	o: 🚏 🍃 🗾			≡
🗱 0 🗮 81 / 882.05 KB 💿 0 💿 0	🕨 0 📕 46 🔺 11 🤺 0	√ 0 * 1	💿 0 🕚 0 ? 0 🎜 14:39:46 MDT	Q 🖸
Read/Write 0 bytes / 0 bytes	5 min 📲 Name success		SplitJson 1.11.3	
Out 0 (0 bytes)	5 min Queued 2 (528.77 KB)		org.apache.nifi - nifi-standard-nar In 0 (0 bytes) 5 min	
Navigate Tasks/Time 0/==(0:00.000	5 min		Read/Write 0 bytes 5 min	
			Out 0 (0 bytes) 5 min	
@ Q []!!			Tasks/Time 0 / 00:00:00.000 5 min	
			Name split	
			Queued 36 (65.75 KB)	
	EvaluateJsonPath EvaluateJsonPath 1.11.3 org.apache.nifi - nifi-standard-nar		Coords ExecuteScript 1.11.3	
	In 0 (0 bytes)	Name success	org.apache.nifi - nifi-scripting-nar 0 (0 bytes) 5 min	
	Read/Write 0 bytes / 0 bytes	Queued 1 (2.3 KB)		
	Out 0 (0 bytes)	5 min	Out 0 (0 bytes) 5 min	
	Tasks/Time 0 / 00:00:00.000	5 min	Tasks/Time 0 / 00:00:00.000 5 min	
	Name matched			
	Oueued 0 (0 bytes)	Name matched		
🖕 Operate 📃		Queued 0 (0 bytes)		
Multiple components selected	ElasticSCF PutElasticsearchHttp 1.11.3 org.apache.nifi - nifi-elasticsearch-nar			
5 min	In 0 (0 bytes)	5 min	New DataWarehouse PutElasticsearchHttp 1.11.3	
5 min	Read/Write 0 bytes / 0 bytes	5 min	org.apache.nifi - nifi-elasticsearch-nar	
* * * • =	Out 0 (0 bytes)		In 0 (0 bytes) 5 min	
	Tasks/Time 0 / 00:00:00.000		Read/Write 0 bytes / 0 bytes 5 min	
🖆 🖪 🔟 🥒 🏛 DELETE			Out 0 (0 bytes) 5 min	
			Tasks/Time 0 / 00:00:00.000 5 min	
			₩	
3)				

NiFi Flow » * SCF

	🗞 NiFi Flow	× +					
← → ୯ ŵ	🗊 🛈 loca	lhost:9300/nifi/			… ⊠ ☆	± III\ ₪	
nifil 🖸 🖓		i iei 🚏 🏅	= //				≡
🇱 0 🔳 81 / 882.05 KB	0 0	0 🕨 0 📕 46	🛕 11 🌾 0 🛹 0 🗰 1	○ 0 ○ 0	? 0 😂 14:45:53 MDT		Q
Ø Navigate	8						
Q Q []!							
		\star SCF					
			ion 1 in "DataEngineeringRegistry - DataEngine	eeringPython'			
		Local changes have b					
A		Queued	44 (864.15 KB)				
🖞 Operate 💧 NiFi Flow		In	0 (0 bytes) → 0		5 r	min	
Process Group a1fe8006-0170-1000-ae13-cf025be97ef5		Read/Write	0 bytes / 0 bytes		5 r	min	
* 1 * 🕨 🖬 🗗 🖬	8	Out	$0 \rightarrow 0 \ (0 \ \text{bytes})$		5 r	min	
🔁 🖪 🚺 🖌 🛍 DELETE							
			0 0 0 ? 0				

📝 Apache NiFi Registry Systen 🗙 🏾 🎄	NiFi Flow	v × +							
← → ℃ ŵ	0	localhost:9300/nifi/?processGroupId=roo	t&componentIds=9466c3ca-4c6d-3884	-ac72-af4a27919fb0 •••		☆ ⊻	\ ⊡	۲	≡
nifi 🖸 Ð	⊡ >	וסו 🗳 א	<i>≡∥,</i>					E	≡
🗱 0 🗮 81 / 882.05 KB	• 0	Show Local Changes				4:50:27 MDT		Q	D
 Navigate Q Q [3] [:] 		The following changes have been made Displaying 3 of 3 Filter	to SCf (Version 1).				1		
		Component Name 🔺	Change Type	Difference					
	-11	ElasticSCF	Auto-Terminated Relationships Changed	From '[failure, retry]' to '[success, failur	→				
		New DataWarehouse	Component Added	Processor was added	→				
			Component Added	Connection was added	→				
	_								
Operate						E main			
SCF Process Group						5 min	_		
9466c3ca-4c6d-3884-ac72-af4a27919fb0						5 min			
🌣 🤊 🦎 🕨 🖬 📑 🚠						5 min			
🙆 🖪 🗐 🥒 🛍 DELETE									
				CLO	SE				
NiFi Flow									

🕴 Apache NiFi Registry Systen 🗙 🌾	NiFi Flow × +					
(←) → C [*] ⁽¹⁾	Icalhost:9300/nifi/		🗟 .	☆ <u>⊀</u>	∭\ ⊡	© ≡
nifi 🖸 🕹		• <i>=</i> //				≡
🇱 0 🗮 81 / 882.05 KB	◎ 0	▲ 11 🕺 0 🛩 0 🗰 0	○ 1 ● 0 ? 0 2 1	15:00:37 MDT		Q 🛛
Ø Navigate		0.005				
ତ୍ର୍ 🖸 🗄 🛙		SCF				
			9 🗛 3 🄆 0			
		Add Process Group	64.15 KB)			
		Process Group Name	oytes) → 0	5 min		
			es / 0 bytes	5 min		
		Import	(0 bytes)	5 min		
😋 Operate			0 ? 0			
NiFi Flow Process Group						
a1fe8006-0170-1000-ae13-cf025be97ef5		CANCEL ADD				
	3					
NiFi Flow						

🖊 Apache NiFi Registry Systen 🗙	🎄 NiFi Flow	× +						
← → ♂ ☆	0 i la	ocalhost:9300/nifi/			***	⊎ ☆	⊻ II\ 🗊	
nifi 🖸 🖓		احة احة	-	= //,		_		Ξ
👪 0 🔳 81 / 882.05 KB	0	Import Version				5:01:07 MDT		Q 🛛
		Registry						
Navigate		DataEngineeringRegistry	r		•	·		
QQ []!		Bucket						
		DataEngineeringPython			·	·		
		Flow Name SCf				• 5 min		
			_			5 min		
		Version - Created 2 05/15/2020	14:51:22.133	Comments Added the New Data Warehouse		5 min		
😋 Operate			14:10:52.067	First commit				
NiFi Flow Process Group								
a1fe8006-0170-1000-ae13-cf025be97ef5	5							
	* -							
					CANCEL			
NiFi Flow								
O Create a New Repository ×	+							
(←) → ♂ ŵ	🔽 🔒 h	ttps:// github.com /new				⊌ ☆	⊻ II\ 🗈	© ≡
<u> </u>		Create a new	reposito	ory				
		A repository contains al Import a repository.	l project files,	, including the revision history. Already hav	ve a project repository elsewhere?			
								- 1
		Owner	Renosito	ory name *				
		PaulCrickard -	/ NifiRegi					
			are short and	d memorable. Need inspiration? How abou	ut cuddly-eureka?			
		Description (optional)						
		Description (optional) Flows from Nifi			-			
					-			
		Flows from Nifi	his repository. Ye	You choose who can commit.	-			
		Flows from Nifi Flows from Nifi Flows from Nifi Anyone can see the private			-			
		Flows from Nifi Flows from Nifi Flows from Nifi Anyone can see the private		You choose who can commit. mmit to this repository.	-			
		Flows from Nifi Flows from Nifi Flows from Nifi Anyone can see the private	can see and com	mmit to this repository.				
		Flows from Nifi Flows from Nifi Flows from Nifi Public Anyone can see tt Private Vou choose who of Skip this step if you're in Vitialize this repos	an see and com mporting an ex itory with a R	mmit to this repository. existing repository. README	-			
		Flows from Nifi Flows from Nifi Flows from Nifi Public Anyone can see tt Private Vou choose who of Skip this step if you're in Vitialize this repos	an see and com mporting an ex itory with a R ately clone the re	mmit to this repository. existing repository.				

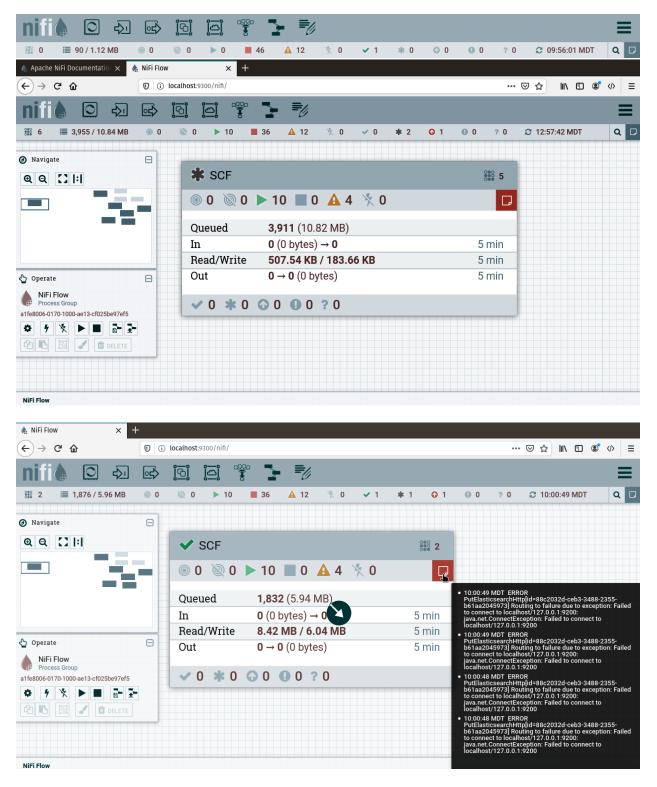
Personal Access Tokens	s x +				
) → ሮ û	🗊 🔒 https://gith	ub.com/settings/tokens		🖂 ☆	⊻ II\ 🗊 📽
Search or jump		Pull requests Issues Marketplac	ce Explore		ب+ ب
	Settings / Developer setting:	S			Signed in as PaulCrickard
	GitHub Apps	Personal access tok	kens	Generate new token	Set status
	OAuth Apps	Need an API token for script	ts or testing? Generate a personal access token	for quick access to the GitHub API.	Your profile
	Personal access tokens		like ordinary OAuth access tokens. They can be used in		Your repositories Your projects
					Your stars Your gists
	© 2020 GitHub, Inc. Terms Priv	vacy Security Status Help	Contact C	GitHub Pricing API Training Blog Abou	Feature preview Help
					Settings
					Sign out
ew personal access to	oken × +				
lew personal access to → C' 企		ub.com/settings/tokens/new		⊚ ☆	<u>↓</u> II\ 🗊 ®
				··· 🖂 🎝	⊻ II\ 🗊 ©
	🛛 🖨 https://gith			··· ତ ជ	<u>↓</u> II/ E ©
	🛛 🖨 https://gith		cess token	··· ତ ជ	<u>↓</u> II\ 🗊 ©́
	Settings / Developer settings	New personal acc			<u>≁</u> II/ Œ ©
	Image: Settings Developer settings GitHub Apps	New personal acc	CESS TOKEN tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentic	an be used instead of a password for Git	<u>⊼</u> II/ ⊡ ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal acc	tion like ordinary OAuth access tokens. They ca	an be used instead of a password for Git	<u>⊼</u> III/ ED ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used	tion like ordinary OAuth access tokens. They ca	an be used instead of a password for Git	<u>⊼</u> III/ ⊡ @
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used Note	tion like ordinary OAuth access tokens. They ca	an be used instead of a password for Git	<u>⊼</u> IIV Œ ®
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used Note	tion like ordinary OAuth access tokens. They ca	an be used instead of a password for Git	<u>⊼</u> II/ ⊑ €
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used to Note Nifi What's this token for? Select scopes	tion like ordinary OAuth access tokens. They ca	In be used instead of a password for Git ation.	<u>≁</u> III\ CD ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used to Note Nifi What's this token for? Select scopes	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentic	In be used instead of a password for Git ation.	<u>⊼</u> III/ ED ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal acc Personal access tokens funct over HTTPS, or can be used in Note Nifi What's this token for? Select scopes Scopes define the access for repo repo:status	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentica personal tokens. Read more about OAuth scop	In be used instead of a password for Git ation.	<u>⊼</u> III/ ED ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used to Note Nifi What's this token for? Select scopes Scopes define the access for repo repo repo repo_deployment	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentica personal tokens. Read more about OAuth scop Full control of private repositories Access commit status Access deployment status	In be used instead of a password for Git ation.	<u>⊼</u> II/ Œ ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used to Note Nifi What's this token for? Select scopes Scopes define the access for repostatus repo_deployment repo_deployment repo_deployment repo_deployment repo_deployment	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authenticat personal tokens. Read more about OAuth scop Full control of private repositories Access commit status Access deployment status Access public repositories	In be used instead of a password for Git ation.	<u>⊼</u> II/ Œ ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used to Note Nifi What's this token for? Select scopes Scopes define the access for repostatus repo_deployment repo_deployment repoinvite	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentica personal tokens. Read more about OAuth scop Full control of private repositories Access commit status Access deployment status Access public repositories Access repository invitations	In be used instead of a password for Git ation.	<u>⊼</u> III/ (⊡ (s)
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used in Note Nifi What's this token for? Select scopes Scopes define the access for repo repo_teployment public_repo repo_teployment public_repo repo_trivite security_events	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentica personal tokens. Read more about OAuth scop Full control of private repositories Access commit status Access deployment status Access public repositories Access repository invitations Read and write security events	In be used instead of a password for Git ation.	<u>⊼</u> III/ E ©
New personal access to) → C ^a û	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal acc Personal access tokens funct over HTTPS, or can be used Note Nifi What's this token for? Select scopes Scopes define the access for erpo_ erpo_deployment erp	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentica personal tokens. Read more about OAuth scop Full control of private repositories Access commit status Access deployment status Access repositories Access repository invitations Read and write security events Upload packages to github package registry	an be used instead of a password for Git ation.	<u>⊼</u> III/ ⊑ ©
	Image: Settings Developer settings GitHub Apps OAuth Apps	New personal access tokens funct over HTTPS, or can be used in Note Nifi What's this token for? Select scopes Scopes define the access for repo repo_teployment public_repo repo_teployment public_repo repo_trivite security_events	tion like ordinary OAuth access tokens. They ca to authenticate to the API over Basic Authentica personal tokens. Read more about OAuth scop Full control of private repositories Access commit status Access deployment status Access public repositories Access repository invitations Read and write security events	an be used instead of a password for Git ation.	<u>⊀</u> III (()

parterickard@pop-os:~\$ git clone https://github.com/Pauterickard/Nilikegistry.git
Cloning into 'NifiRegistry'
remote: Enumerating objects: 3, done.
remote: Counting objects: 100% (3/3), done.
Unpacking objects: 100% (3/3), done.
remote: Total 3 (delta 0), reused 0 (delta 0), pack-reused 0

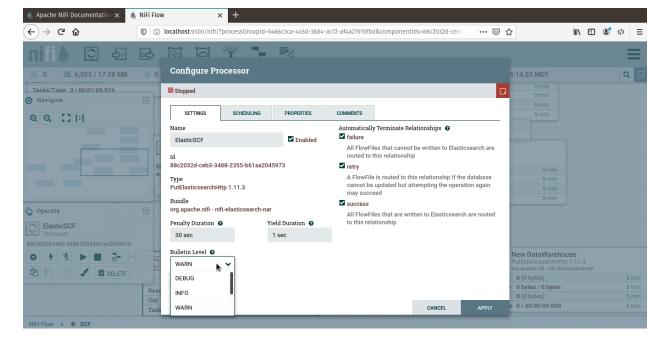
Open - A limitations under the Lice	providers.xml -/nifi-registry-0.6.0/conf		Save = 😣
> <providers></providers>			
<pre></pre>	in this file must be listed in the order nce Provider (Must occur once and only or oviders (May occur 0 or more times) tence Provider (Must occur once and only	nce)	ch is the following:
	fi.registry.provider.flow.FileSystemFlow/ Storage Directory">./flow_storage <td></td> <td></td>		
<property name="Flow
<property name=" remo<br=""><property name="Remo</td><td><pre>fi.registry.provider.flow.git.GitFlowPers storage Directory">/home/paulcrickard/N: te To Push">origin</property> te Access User">paulcrickard</property> te Access Password">YOUR TOKEN&/property>	ifiRegistry		
<br <flowpersistenceprovider <class>org.apache.ni ></class></flowpersistenceprovider 	fi.registry.provider.flow.DatabaseFlowPe	rsistenceProvider	
<property name="Scri</td><td>fi.registry.provider.hook.ScriptEventHoo
pt Path"></property> ing Directory">	<provider< class=""></provider<>		
		XML 🔻 Tab Width: 8 🕶	Ln 35, Col 59 🔻 INS
M Mill Davidson			
$ \begin{array}{c} & & \\ & & \\ \hline \\ \hline$	calhost:18080/nifi-registry/#/explorer/grid-list	··· 🛛 🕁	<u>↓</u> II\ 🗊 📽 😑
NiFi Registry / All 🔻			anonymous 😯 🎤
		T	Sort by: Name (a - z) 🔻
SCf - DataEngineeringPyth Flow	ON VERSION 3	٩S	^
BUCKET IDENTIFIER	CHANGE LOG 2		ACTIONS -
c8e69fca-7f08-44a1-967c-32a1a57094	Version 3 - a minute ago		
FLOW IDENTIFIER	by anonymous		
7aa48258-ce5c-44d8-ae7a-6e67d9dba	a4c Added a third warehouse May-15-2020 at 4:25 PM		
DESCRIPTION			
Chapter 6: SeeClickFix API to Elasticse	arch Version 2 - 2 hours ago by anonymous		
	Version 1 - 2 hours ago by anonymous		

) NifiRegistry/DataEngineerin \leftrightarrow \rightarrow C' \textcircled{a}		NifiRegistry/tree/master/DataEngineeringPython	Q Recommendation ⊙ ☆	<u>↓</u> II\ 🗊 📽 =
		Read the guide		
Ļ	PaulCrickard / NifiRegistry		O Unwatch → 1 ★ Star 0 % Fork 0	
	Code () Issues 0 () Pull requests 0	O Actions III Projects O III Wiki II S	ecurity 0 🛄 Insights 🗘 Settings	
E	Branch: master - NifiRegistry / DataEnginee	ringPython /	Create new file Upload files Find file History	
	root Added a third warehouse		Latest commit 50f8faf 1 minute ago	
	SCf.snapshot	Added a third warehouse	1 minute ago	
	bucket.yml	Added a third warehouse	1 minute ago	
© :	2020 GitHub, Inc. Terms Privacy Security Status	Help 🔘	Contact GitHub Pricing API Training Blog About	

Chapter 9: Monitoring and Logging Data Pipelines

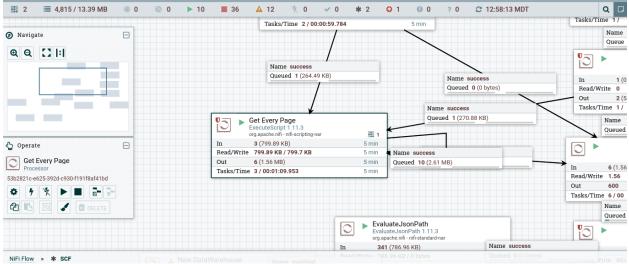


🎄 NiFi Flow 🛛 🗙 🕂					
↔ → ♂ ŵ	🗊 🛈 localhost:9	300/nifi/?processGroupId=9466c3ca-4c6d-3884-ac72-	af4a27919fb0&compon	entids=88c2032d-ceb3	
nifi 🖸 🖓	ि हिं	iai 🚏 🍃 🗾			≡
888 2 🗮 3,064 / 8.64 MB	0 Ø 0	🕨 10 📕 36 🔺 12 🧏 0	✓ 0 ¥ 2	O 1 O ? O	😂 10:01:52 MDT 🛛 🔍 🗖
Ø Navigate	8	EvaluateJsonPath EvaluateJsonPath 1.11.3 org.apache.nifi - nifi-standard-nar	Name su	1	Coords ExecuteScript 1.11.3 org.apache.nlfi - nlfi-scripting-nar
Q Q [] :	e matched	In 1,272 (2.86 MB) Read/Write 2.86 MB / 0 bytes		0 (0 bytes)	1,271 (2.21 MB) Vrite 2.21 MB / 2.86 MB
	ed 1,272 (2.8	,816 (8.57 MB) ,272 / 00:00:12.951	5 min 5 min		ut 1,271 (2.86 MB) asks/Time 1,271 / 00:03:44.866
	5 min 5 min				
	5 min 5 min	Name matched Queued 0 (0 bytes)		Name matched Queued 1,272 (2.86	6 MB)
🖞 Operate		► ElasticSCF PutElasticsearchHttp 1.11.3	R		New Data PutElasticsea
C ElasticSCF Processor		org.apache.nifi - nifi-elasticsearch-nar	• 1	10:02:05 MDT ERROR	org.apach8.nifi- 2022d-cob2-2499-2255-b61-22045972] Bouting to() by (85)
88c2032d-ceb3-3488-2355-b61aa2045973		In 1,409 (3.17 MP	5 min	PutElasticsearchHttp[id=88c2 failure due to exception: Faile	2032d-ceb3-3488-2355-b61aa2045973] Routing to UDI es) ed to connect to localhost/127.0.0.1:9200; the Dibytes / Failed to connect to localhost/127.0.0.1:9200
		Read/Write 3.17 MB / 0 by		java.net.ConnectException: F 10:02:05 MDT ERROR	alled to connect to localhost/127.0.0.1:9200 Out 0 (0 bytes)
		Tasks/Time 1,409 / 00:00:21.399	5 min	PutElasticsearchHttp[id=88c2 failure due to exception: Faile	2032d-ceb3-3488-2355-b61aa2045973] Routing to 00:0 ed to connect to localhost/127.0.0.1:9200: Failed to connect to localhost/127.0.0.1:9200
				10:02:05 MDT ERROR PutElasticsearchHttp[id=88c2 failure due to exception: Faile	2032d-ceb3-3488-2355-b61aa2045973] Routing to ed to connect to localhost/127.0.0.1.9200: alleld to connect to localhost/127.0.0.1.9200
NiFi Flow » * SCF				10:02:04 MDT ERROR PutElasticsearchHttp[id=88c2 failure due to exception: Faile	2032d-ceb3-3488-2355-b61aa2045973] Routing to ed to connect to localhost/127.0.0.1:9200:



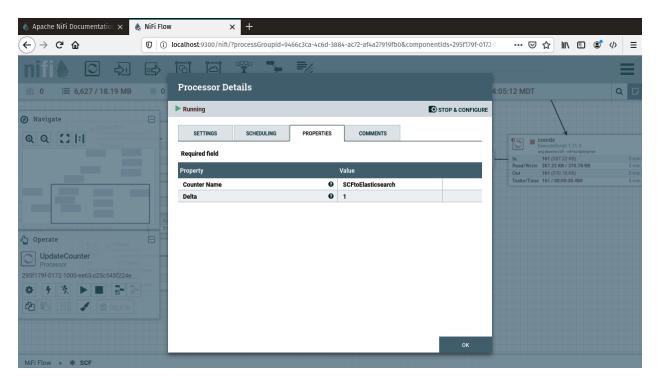
NiFi Bulletin Board

Filter by message
13:12:21 MDT ERROR 88c2032d-ceb3-3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3-3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1.9200: java.net.ConnectException: Failed to connect to localhost/127.0.0.1.9200
13:12:21 MDT ERROR 88c2032d-ceb3-3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3-3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1:9200; java.net.ConnectException: Failed to connect to localhost/127.0.0.1:9200
13:12:24 MDT ERROR 88c2032d-ceb3-3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3-3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1:9200: java.net.ConnectException: Failed to connect to localhost/127.0.0.1:9200
13:12:24 MDT ERROR 88c2032d-ceb3-3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3-3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1:9200: java.net.ConnectException: Failed to connect to localhost/127.0.0.1:9200
13:12:24 MDT ERROR 88c2032d-ceb3-3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3-3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1:9200; java.net.ConnectException: Failed to connect to localhost/127.0.0.1:9200
13:12:25 MDT ERROR 88c2032d-ceb3-3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3-3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1:9200; java.net.ConnectException: Failed to connect to localhost/127.0.0.1:9200
13:12:25 MDT ERROR 88c2032d-ceb3:3488-2355-b61aa2045973 PutElasticsearchHttp[id=88c2032d-ceb3:3488-2355-b61aa2045973] Routing to failure due to exception: Failed to connect to localhost/127.0.0.1:9200; java.net.ConnectException: Failed to connect to localhost/127.0.0.1:9200
Auto-refresh Clear
🗼 Apache Nifi Documentatio 🗙 🍓 Nifi Flow 🗙 🕂
$(\leftarrow) \rightarrow \mathbb{C} \textcircled{0} \textcircled{0} \texttt{localhost.9300/nifl/?processGroupId=9466c3ca-4c6d-3884-ac72-af4a27919fb0&componentIds=53b2821c-e625} \cdots $
= · · · · · · · · · · · · · · · · · · ·
ﷺ 2



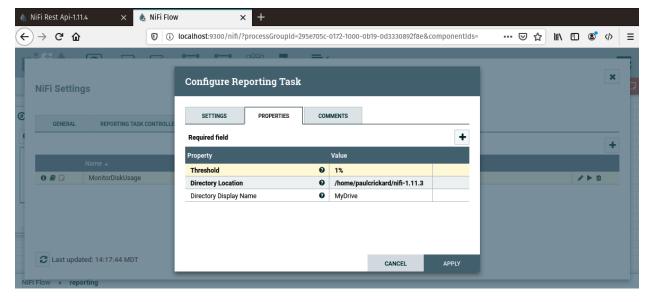
×

EvaluateJsonPath EvaluateJsonPath 1.11.3 org.apache.nifi - nifi-standard-nar	
In 162 (373.16 KB)	5 min
Read/Write 373.16 KB / 0 bytes	5 min
Out 486 (1.09 MB)	5 min
Tasks/Time 162 / 00:00:00.725	5 min
Name matched	
Queued 0 (0 bytes)	
UpdateCounter UpdateCounter 1.11.3 org.apache.nifi - nifi-standard-nar	
In 162 (373.16 KB)	5 min
Read/Write 0 bytes / 0 bytes	5 min
Out 162 (373.16 KB)	5 min
Tasks/Time 162 / 00:00:01.296	5 min
Name success	
Queued 162 (373.16 KB)	
ElasticSCF PutElasticsearchHttp 1.11.3 org.apache.nlfi - nlfi-elasticsearch-nar	
In 0 (0 bytes)	5 min
	5 min
Read/Write 0 bytes / 0 bytes	
Out 0 (0 bytes)	5 min

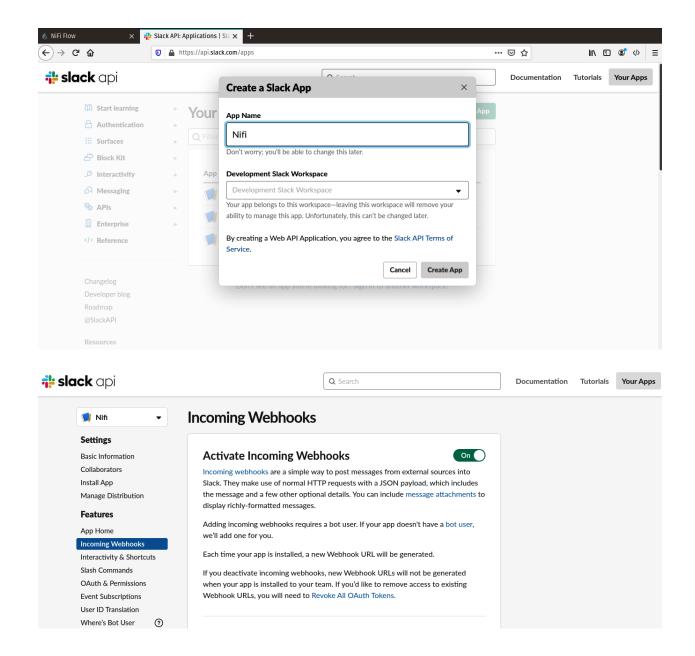


		ē—ē	n00n 🗮 🚍 /			2 ×
NiFi Counters						
Displaying 2 of 2						
Filter	by name	~				
Context 🔺			Name	Value		
All UpdateCounter's			SCFtoElasticsearch	162	C	
UpdateCounter (295f179f-017	2-1000-ee63-c25c545f224e)		SCFtoElasticsearch	162	ຽ	

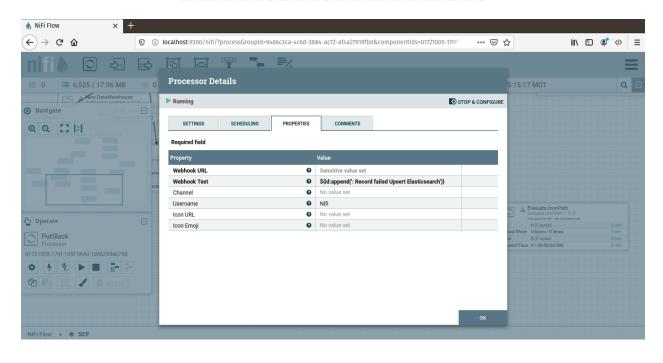
		Add Reporting Task				
iFi Settings		Source	Displaying 15 of 15		Filter]
0515041		all groups 🗸 🗸	Туре 🔺	Version	Tags	
GENERAL	REPORTING TASK CONTRO		AmbariReportingTask	1.11.3	ambari, metrics, reporting	
		ambari azure datadog disk	AzureLogAnalyticsProvena	1.11.3	provenace, log analytics, repor	
	_	garbage collection gc heap	AzureLogAnalyticsReportin	1.11.3	metrics, log analytics, reportin	
	me 🔺	jvm log log analytics	ControllerStatusReportingT	1.11.3	stats, log	
M D M	onitorDiskUsage	memory metrics	DataDogReportingTask	1.11.3	datadog, metrics, reporting	J → 1
		monitor monitoring	MetricsReportingTask	1.11.3	metrics, reporting	
		prometheus provenace repo	MonitorDiskUsage	1.11.3	disk, repo, warning, storage, m	
		reporting	MonitorMemory	1.11.3	jvm, memory, warning, monitor,	
		restricted site	PrometheusReportingTask	1.11.3	prometheus, metrics, time seri	
		site to site stats	ScriptedReportingTask	1.11.3	luaj, python, jython, js, execute,	
		status storage warning	SiteToSiteBulletinReporting	1.11.3	site, restricted, bulletin, site to	
C Last updated:	14:17:44 MDT		Service works, this reporting tas will send the metrics from the p	Ambari Metrics Service (AMS) sk should be scheduled to run revious iteration, and calculat	mbari-nar . Due to how the Ambari Metrics every 60 seconds. Each iteration it e the current metrics to be sent on ther than 60 seconds may produce	

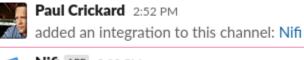


	NiFi Flov	v	×	+											
(←) → ୯ ŵ	00	localhost	:9300/nifi/?pi	rocessGroupId	=295e705c-01	172-1000-0b19	-0d3330892	f8e∁	onentids=			⊌ ☆	lii\	•	<⇒ Ξ
nifil 🖸 🖓	₀=\$>	<u>ស្រុ</u> រ		¥ 🟅	. <i>≡</i> ⁄/										≡
🗱 0 🔳 6,627 / 18.19 MB	o (0 @	0 • 0	48	A 12	Ϋ́ Ο	✓ 0	* 2	<mark>0</mark> 1	0	? 0	C 14:17:08 MDT			Q [
⊘ Navigate QQ [] :									14:1 Mor exce and	7:07 MDT itorDiskUsa reds config 434.82 GB	WARNING ge[id=017; ured thresh (96.82%) fr	21003-179f-195f-9cbe- old of 1%, having 14.27 ee	27f0f068b38 GB / 449.09	a) MyDrive GB (3.18%) used
♦ Operate Image: Propositing Proposition 295e705c-0172-1000-0b19-0d3330892f8e Image: Proposition Image: Proposition Proposition Image: Proposition Image: Proposition Proposition Image: Proposition Image: Proposition Proposition Image: Proposition Proposition Image: Proposition Proposition Proposition Image: Proposition Proposition Image: Proposition Proposition Proposition Proposition Image: Proposition Proposition Image: Proposition Proposi	Slack AF		ions Str 🗙										III	F1 (8)	
i slack api		nups://a	DI.Stack.com/	apps		Q Searc	-b					ত দ্ব Documentation	III Tutorials		Apps
🕅 Start learning	Þ					L									
 Authentication Surfaces Block Kit ✓ Interactivity Messaging 	> > >		App Name	DS y name or wo	rkspace	Workspace DA2ND	ce	Distribu Not distri	ution State	reate New	Арр				
III Surfaces 슈 Block Kit グ Interactivity 슈 Messaging 아 APIs			Filter apps by	y name or wo	rkspace		ce		ution State		Арр				
iii Surfaces 合 Block Kit グ Interactivity 介 Messaging)))))		Filter apps by	y name or wo	rkspace	DA2ND		Not dist	ition State		(App				



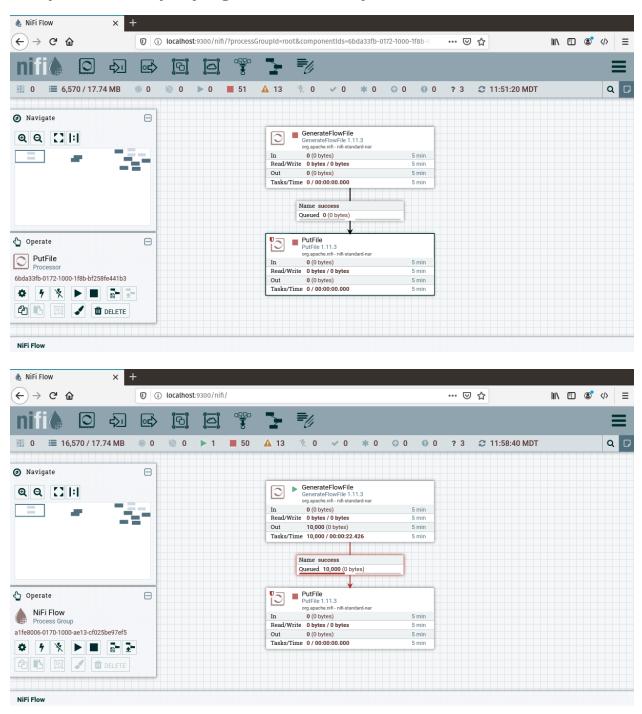
0 P	asticSCF ItElasticsearchHttp 1.11.3 g.apache.nlfi - nlfi-elasticsearch-nar	D
In	101 (233.94 KB)	5 min
Read/Write	233.94 KB / 0 bytes	5 min
Out	101 (233.94 KB)	5 min
Tasks/Time	2/00:00:00.556	5 min
Qu	eued 1 (2.22 KB)	
	PutSlack PutSlack 1.11.3 irg.apache.nlfi - nlfi-slack-nar	
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min



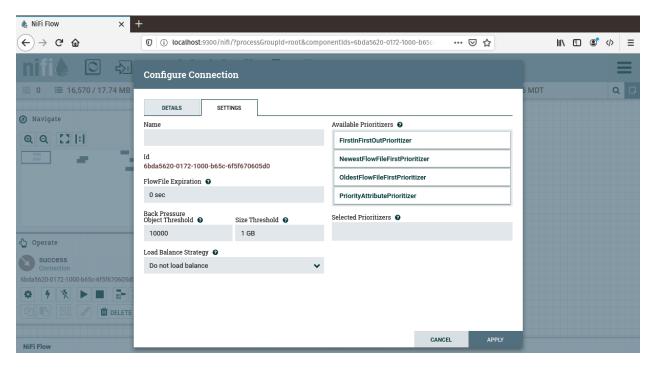


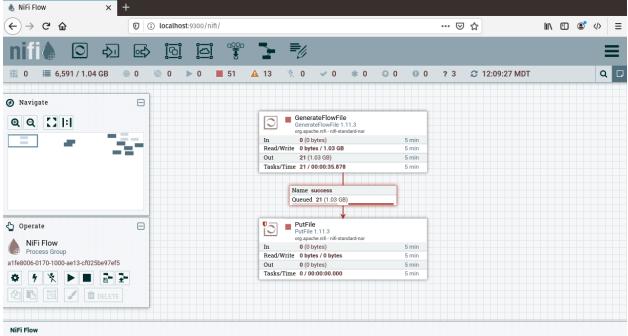
New

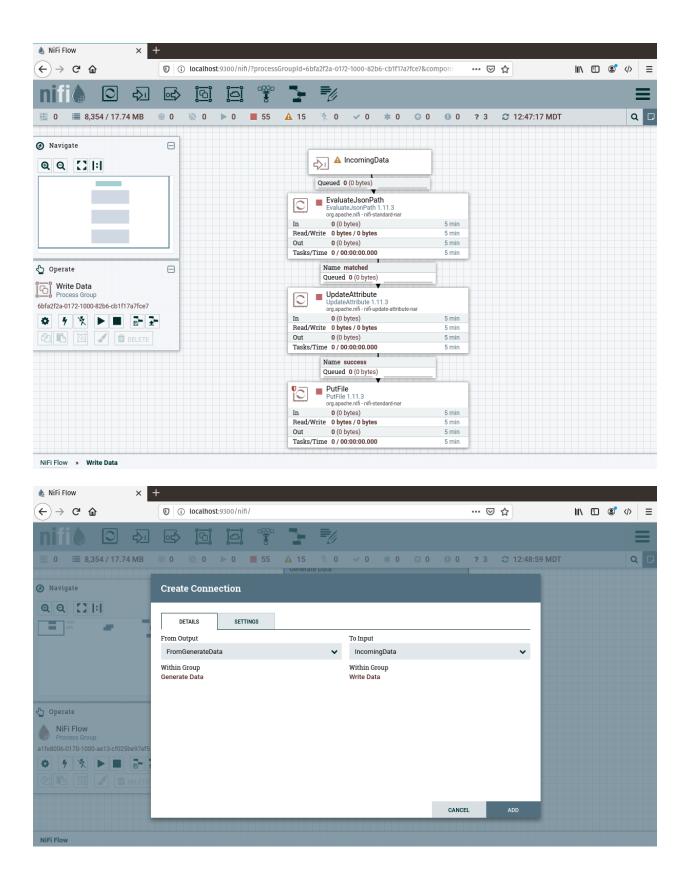
Nifi APP 3:03 PM 7781528: Record failed Upsert Elasticsearch

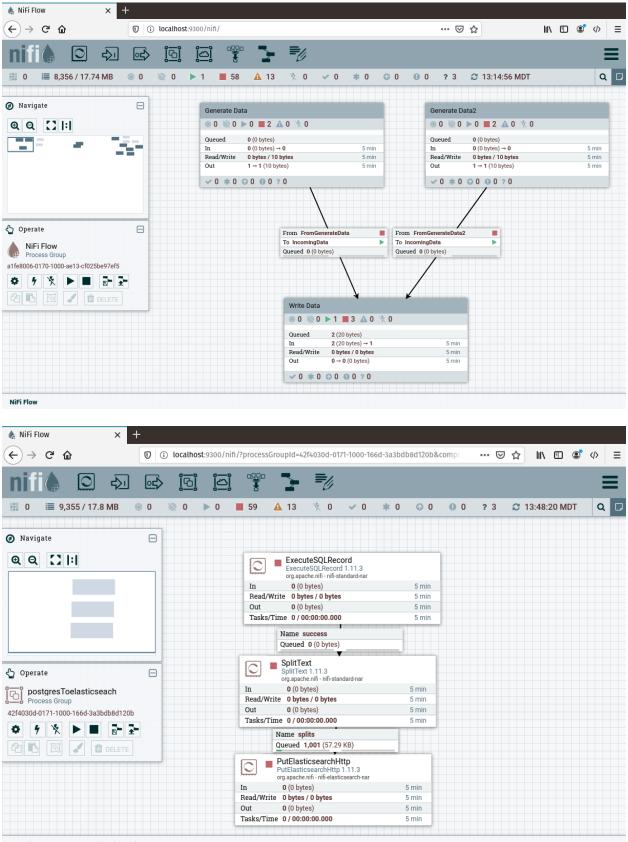


Chapter 10: Deploying Your Data Pipelines











🌲 NiFi Flow	× +				
← → ♂ û	i localhost	:9300/nifi/?processGroupI	d=root&componentIds=42	f4030d-0	0171-1000-166d-3 … 🖂 🏠 💷 🔿
nifi 🖸					≡
👬 0 🔳 9,355 / 17.8	MB 💿 0 📎 0	▶ 0 🔳 59 🔺 1	3 🌾 0 🖌 0	* 0	💿 0 🔍 0 ? 3 😂 13:53:12 MDT 🛛 Q 🗖
5 min Navigate	Variables				
ଷ୍ ଷ୍ଟ : : : : : : : : : : : : : : : : : : :	Process Group postgresToelasticseach			+	Variables elastic
	Scope	Name 🔺	Value		Referencing Processors 😧
	postgresToelasticseach	elastic	http://localhost:9200	Û	Referencing Controller Services 📀
	NiFi Flow postgresToelasticseach	elastic index	http://127.0.0.1:9200 nifivariable	→ @	None
		Index	Inivaliable		Unauthorized Referencing Components 😨
🖒 Operate					
-					
Process Group					
42f4030d-0171-1000-166d-3a3b	c				
* 1 × •					
26.					
	Parameters are replacing v	ariables. Learn more about	the greater power and secu	rity of na	rameters. CANCEL APPLY
	T drameters are replacing v		and greater power and seed	nty or pu	CANCEL APPEr
NiFi Flow					
🎄 NiFi Flow	× 🥃 loo	alhost:9200/_cat/in	dices × +		
$\overleftarrow{\bullet}$ > C $\widehat{\bullet}$	Ū	i localhost:9200)/_cat/indices		
yellow open frompo yellow open nifiva yellow open fromni yellow open test green open .kibana yellow open scf green open .apm-a green open .kiban yellow open users	riable fi a_task_manager_1 _sample_data_ecomme gent-configuration	TRi0TFoKSoCd NBAL_aLdQ1y4 VWacjWq_S9a0 InLCVInfTa6X erce -6sdGU13T12G 7xSytRRETcyD 51toZCyRS2yW gSgxKq0xR0y7	PtrufvSNkA 1 1 22 _2HFnf9HAQ 1 1 10 iK90kDFauw 1 1 40 fW6l-YbcIA 1 1 mmifoGntXA 1 0 IjE2rU_ZlQ 1 0 40 Kt3wXRjaYg 1 1 23 lnLJZWF5-A 1 0 IFCG0rKiMw 1 0 xh2TqBybKw 1 1 10	001 0 004 0 2 0 575 0 365 0 0 0 92 25) 7.1kb 7.1kb) 4.9mb 4.9mb) 3.6mb 3.6mb) 283b 283b ; 979kb 979kb

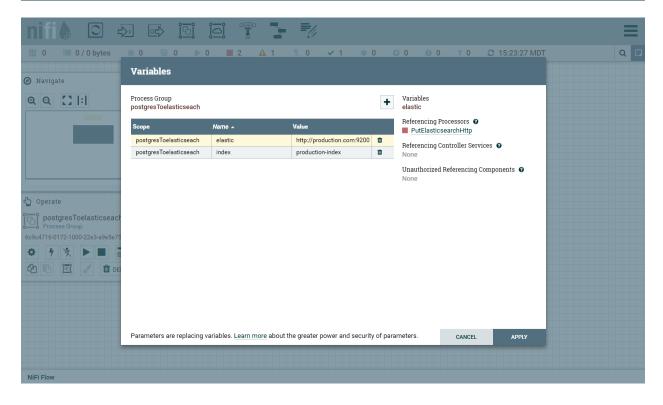
♠ NiFi Flow × + ← → ⊂ ω □ □	localboot:0000/nifi/2process	Groupid=root&componentids=0172100c-f40c-1c05-b 🗸 🚥 🛛 🟠	<u>↓</u> III\ 🗊	© <> =
				• • =
······································	الم الم	2 -0		_
🗱 0 🗮 8,354 / 17.74 MB 💿 0	፩ 0 ▶ 0 ■ 59	▲ 16 🏌 0 🗸 3 🗰 1 💽 3 🚇 0 ? 0	C 15:57:47 MDT	Q []
🔊 Navigate 📃	DEV			
-		O Dev		
ଷ୍ ପ୍ 🚺 🔛		◎ 0 ◎ 0 ▶ 0 ■ 0 ▲ 0 ½ 0		
		Queued 0 (0 bytes)		
	-	In 0 (0 bytes) → 0 5 min		
		Read/Write 0 bytes / 0 bytes 5 min Out 0 → 0 (0 bytes) 5 min		
		✓ 0 * 0 © 0 ● 0 ? 0		
	-			
	1			
🖞 Operate 🖂	TEST			
C Dev		O Dev		
Process Group	-			
0172100c-f40c-1c05-b7c0-04a92c04be89		◎ 0 📎 0 🕨 0 🔳 0 🛕 1 🌾 0		
◆ ∮ 头 ▶ ■ 📑 🚠		Queued 0 (0 bytes)		
🖄 🖪 🔝 🥒 🏛 DELETE	-	In 0 (0 bytes) → 0 5 min		
		Read/Write 0 bytes / 0 bytes 5 min Out 0 → 0 (0 bytes) 5 min		
		✓ 0 * 0 © 0 ● 0 ? 0		
	PROD			
	PROD			
		✓ Dev		
		◎ 0 🗟 0 ▶ 0 🔳 0 🔺 2 🧏 0		
		Queued 0 (0 bytes)		
		In 0 (0 bytes) → 0 5 min		
		Read/Write 0 bytes / 0 bytes 5 min Out 0 → 0 (0 bytes) 5 min		
		✓ 0 * 0 ⊙ 0 ● 0 ? 0		

NiFi Settings

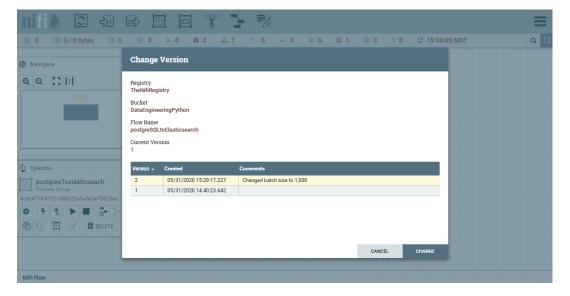
NiFi Settings				_	×	:
GENERAL	REPORTING TASK CONTROLLER SERVICES	S REPORTING TASKS	REGISTRY CLIENTS			
			,		+	-
Name 🔺	l	ocation	De	scription		
TheNifiRegistry		http://10.0.0.148:18080	P	ull data pipelines to production	10	

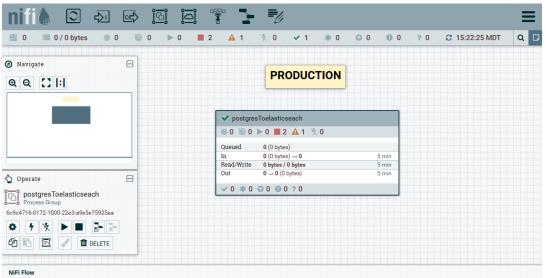
C Last updated: 15:25:24 MDT

nifi 🖸 🗗				. ≡ ⁄/								Ξ
🎎 0 🔳 0 / 0 bytes 💿 0	۵ 🖉	▶ 0 ■ 0	Δ 0	∛ 0	✓ 0	* 0	0	0	? 0	€ 15:18:5) MDT	Q 🛛
O Navigate	Import V	rsion										
ର୍ଷ୍ 🖸 🗄	Registry TheNifiReg	jistry								~		
	Bucket DataEngine	eeringPython								~		
	Flow Name postgreSQ	LtoElasticsearch								*		
🖒 Operate	Version 🗸	Created	C	omments								
NiFi Flow Process Group	1	05/31/2020 14:40:2	3.642									
6c95dfc9-0172-1000-62ff-8deb24eae1ac												
								CANCEL		IMPORT		
NiFi Flow												



3 0 ■ 0 / 0 bytes ③ 0 ◎ 0 ▶ 0		1 🕕 0 ? 0 🎜 15:29:35 MDT	Q [
🔊 Navigate 🖂			
Q Q []!	PRODUCTION		
	O postgresToelasticseach		
	◎ 0 🔌 0 ▶ 0 📕 2 🔺 1 🦎 0		
	Queued 0 (0 bytes)		
	In 0 (0 bytes) → 0 Read/Write 0 bytes / 0 bytes	5 min 5 min	
🖞 Operate 📃	Out $0 \rightarrow 0$ (0 bytes)	5 min	
NIFI Flow Process Group	✓ 0 * 0 ⊙ 0 ⊕ 0 ? 0		
6c95dfc9-0172-1000-62ff-8deb24eae1ac			
* 1 % 🕨 🖬 🚠			

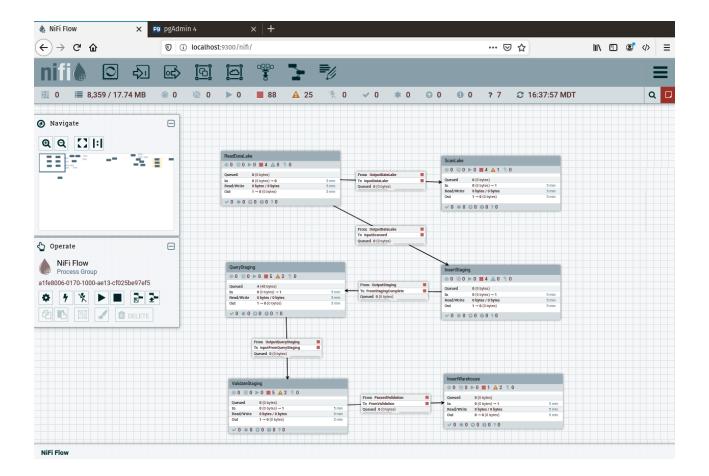


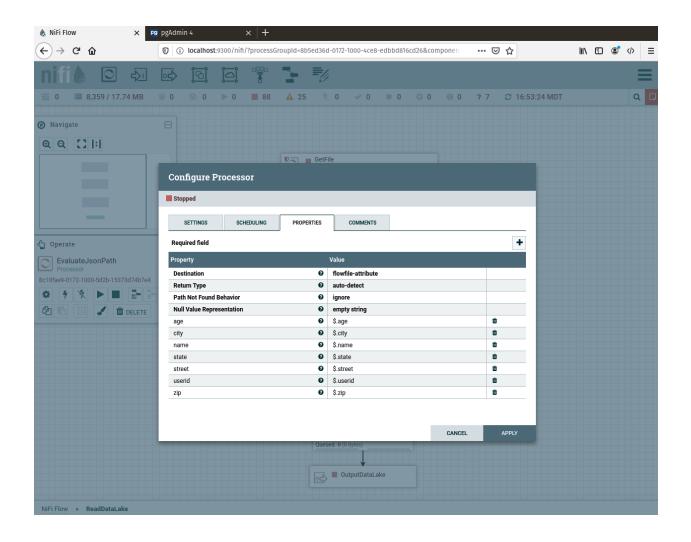


Chapter 11: Building a Production Data Pipeline

\rightarrow C $\textcircled{0}$	(i) loc	alhost/pgadmi	n4/browser/					⊠ ☆		lii\ C	9	
Admin File - Object - Tool	s ∨ H	lelp 🗸								🕌 pos	stgres@l	ocalho
rowser		5	🗉 🚡 Dashboar	d Pr	operties SQL St	atistics Dep	pendencies	Dependents	🗊 test/postgre	s@MyPost	greSQL	< 3
> 🍔 postgres			e B	~ [■ Q - 4 ·	• 6 0	♂ ~ T	✓ No limit	-	• &		200
> 📑 prod												
▼ 🥮 test	📑 sta	aging						×		0		
> 🐼 Casts	Gener	ral Columns	Constraints A	dvance	ed Parameters S	Security SQI	_			Scratch P	ad	
> @ Catalogs			-									
C Event Triggers S S Extensions	Inher	ited from table	(s) Select to in	herit fr	om							
 Extensions Foreign Data Wrappers 	Colu											
> Clanguages	Colu	mns						+				
 Schemas (1) 		Name	Data type		Length/Precision	Scale	Not NULL?	Primary key?				
✓ ♦ public												
> ^A _B ↓ Collations	C t	t userid	bigint	*			Yes	Yes				
> 🏠 Domains	CZ ff	name					No	No				
FTS Configurations		Indiric	text	Ŧ								
FTS Dictionaries	C 🛍	jage	integer	*			No	No				
> Aa FTS Parsers	_											
 Image: Second Sec	U T	street	text	*			No	No				
> (a) Functions	C n	t city	text				No	No				
Materialized Views		,	lexi	*								
> () Procedures	C f	t state	text				No	No				
> 1.3Sequences	~ ~											
✓ ☐ Tables (1)	ľ	zip	text	*			No	No				
✓												
✓												
🚦 userid	i	?				× Cancel	🗘 Rese	t 🖹 Save				
ame 🖥							-					
age street												
city												
state												
> ► Constraints												
> 🚴 Indexes												
> 🥅 Rules												
> 🛟 Triggers												
 Initial Content 												
> 🛅 Types												
> 🧐 Views > 🚓 Login/Group Roles												

	pgAdmin 4	× +										
\leftrightarrow > C \textcircled{a}	(i) localhost/pgad	min4/browser/						… ⊠ ☆	lii/	•	\$ `\$	/>
g Admin File 🗸 Object 🗸	• Tools • Help •								🕌 P	ostgr	es@loca	alhost
Browser	s m 1	Dashboard	Properties	SQL	Statistics	Dependencies	Dependents					
 Servers (1) 												
✓		Туре					Name					
 Databases (4) dataengineering 		🧢 Primary	Key				public.staging_pk	ey			auto	
> Spostgres												
> sprod												
✓ ≡ test												
> 🐼 Casts												
> 💖 Catalogs												
Event Triggers	Create >											
> 領 Extensions > ≝ Foreign Data Wrap	Refresh											
> Canguages	Count Rows											
 ✓ Schemas (1) 	Delete/Drop											
✓ ♦ public	Drop Cascade											
➤ ^A _B ↓ Collations	Reset Statistics											
> 🏠 Domains	Import/Export											
> 🕞 FTS Config	Maintenance											
 IN FTS Diction Aa FTS Parser 	Scripts >	CREATE Sci	int.									
> 💽 FTS Temple												
> 📑 Foreign Tat		DELETE Scr										
> (iii) Functions	Backup	INSERT Scr										
> 📴 Materialize	Restore	SELECT Scr										
> (() Procedures	View/Edit Data >	UPDATE Sc	ript									
 1.3 Sequences Tables (1) 	Query Tool											
 ✓ Image: Tables (1) ✓ Image: Staging 	Properties											
> 📋 Colum	าทร											
> > Const												
> 🚴 Indexe												
> 🧰 Rules												
> 🛟 Trigge												
 > (i) Trigger Funct > Types 	tions											
> log Views												
> 🐴 Login/Group Roles												
> 🔁 Tablespaces												





🎄 NiFi Flow 🛛 🗙 🚺	9 pgAdmii	n 4		× 🖻	NiFi Regis	try	×	+								
(←) → C' ŵ	00	localhost	:9300/nifi	1/?process	sGroupId=	8c4f488e-01	72-1000-eaa	f-805ad2ae0t	53&compo	onentids=		•	⊠ ☆	III\ 🗊	3 <	⊳ ≡
nifi 🖸 🕹	⊡ >	<u>[6]</u>				₽IJ										≡
8,359 / 17.74 MB	⊚ 0	Ø 0		0	88	A 25	Υ Ο	✓ 3	* 1	0 3	0	? 0	C 17:20:06 MDT			۹ 🛛
Ø Navigate	Θ															
@ Q ∷ : 													딝 📕 InputDataLake			
													Queued 0 (0 bytes)			
		0	Evalu	uateJsonf iateJsonPat	th 1.11.3							0	ScanContent ScanContent 1.11.3 org.apache.nifi - nifi-standard-nar			
C Operate		In Road/B	0 (0	oache.nifi - nifi D bytes) ytes / 0 byt e			5 min	Name ma Queued (In Read/V	0 (0 bytes) Vrite 0 bytes / 0 bytes		min min	
CanLake		Out	0 (0	0 bytes) 0 bytes) 00:00:00.00			5 min 5 min	-				Out	0 (0 bytes) Time 0 / 00:00:00.000		min min	
8c4f488e-0172-1000-eaaf-805ad2ae0b53																
				e matched			1									
			Queu	ied 0 (0 byt	tes)								Name matched, unmatched Queued 0 (0 bytes)			
			Put	tSlack	¥											
			Put: org.a	Slack 1.11.3 apache.nifi - n (0 bytes)	3 iifi-slack-nar		5 min						\			
			/Write 0	bytes / 0 by (0 bytes)	ytes		5 min 5 min						▲ OutputScanned			
		Tasks	;/Time 0	/ 00:00:00.	.000		5 min									
NIFI Flow » ScanLake																
_													_			
<u>+</u>					pa	aulcric	kard@p	op-os: ~	/stagi	ng			С	ર ≡		×
What data woul 1. Files o 2. Relatio : 2	n a	file	esys	tem	(fo								Spark)			
Which database 1. MySQL 2. Postgre 3. Redshif 4. Snowfla 5. other - : 2	s t ke							Alche	my c	onne	ectio	on s	string?			
Give your new [my_postgres_						t na	me.									
Next, we will `section		Ũ													tab	ole
of this config	fil	e: g	rea	t_e>	крес	tati	ons/1	uncom	mitt	ed/o	conf:	ig_v	variables.y	nl:		
What is the ho What is the po What is the us What is the pa	rt f erna	or t me f	he or	post the	tgre pos	s co tgre	nnect s cor	tion? nnect	54] ion?	:[32] pc]						

🎄 NiFi Flow	× P	pgAdmin 4 X 🔯 Data documentation complex 🧧 edit_staging.validation -	+ × Iul		
÷) → ℃	ŵ	0 localhost:8888/notebooks/edit_staging.validation.ipynb	··· 🖂 🕅 🗓	℗ ⟨>	
	💭 jupyter 🦸	dit_staging.validation (unsaved changes)	Logout		
	File Edit Vi	ew Insert Cell Kernel Help	Not Trusted 🖋 Kernel O		
	₽ ₩ ∞				
		<pre>context = ge.data_context.DataContext()</pre>			
		<pre># Feel free to change the name of your suite here. Renaming this will n # remove the other one. expectation_suite_name = "staging.validation" suite = context.get expectation_suite(expectation_suite_name) suite.expectations = []</pre>	ot		
		<pre>batch_kwargs = { "datasource": "tagingtable", "limit": 1000, "schema": "public", "table": "staging", } batch = context.get batch(batch kwargs, suite)</pre>			
		Create & Edit Expectations Add expectations by calling specific expectation methods on the batch object. They all begin v tab. You can see all the available expectations in the <u>expectation glossary</u> .	with <code>.expect_</code> which makes autocompleting easy using		
		Table Expectation(s)]		
	In []:	<pre>batch.expect_table_row_count_to_be_between(max_value=0, min_value=0)</pre>			
	In []:	<pre>batch.expect_table_column_count_to_equal(value=7)</pre>			
	In []:	<pre>batch.expect_table_columns_to_match_ordered_list(column_list=["userid", "name", "age", "street", "city", "state", "z)</pre>	ip"]		
		Column Expectation(s)			
		No column level expectations are in this suite. Feel free to add some here. They all begin with b	atch.expect_column		
		Save & Review Your Expectations			

great_expectations	Home / sta	aging.validation / 20200608T01493	4.687170Z / 7be1a20654d	51dfb571fb4353a703c	lef	
Expectation Validation Result Evaluates whether a batch of data matches expectations.	Status:	ew tion Suite: staging.validation Succeeded Expectations	2			
Actions		ul Expectations	2			
Validation Filter:	Unsucces Success F	esful Expectations	0 100%			
Show Failed All Only	Show more	e info				
How to Edit This Suite	Table-	Level Expectations				
Show Walkthrough	Status	Expectation			Observed Value	
	•	Must have exactly 7 columns.			7	
Table of Contents Overview	⊘	Must have these columns in this	order: userid, name, age, st	reet city, state zip	['userid', 'name', 'age', 's	treet', 'city', 'state', 'zip']
Table-Level Expect						

Chapter 12: Building an Apache Kafka Cluster

AFACHE **fka**. A distributed streaming platform

HOME **PUBLISH & SUBSCRIBE** PROCESS STORE LATEST NEWS INTRODUCTION KAFKA SUMMIT 2020 Read and write streams of Write scalable stream Store streams of data safely AUG 24 - AUG 25, 2020 QUICKSTART data like a messaging processing applications that in a distributed, replicated, AK RELEASE 2.5.0 system. react to events in real-time. fault-tolerant cluster. APRIL 15, 2020 USE CASES AK RELEASE 2.4.1 Learn more » Learn more » Learn more » MARCH 12, 2019 DOCUMENTATION AK RELEASE 2.2.2 PERFORMANCE DECEMBER 1 2019 AK RELEASE 2.3.1 POWERED BY OCTOBER 24, 2019 Stream Processor PROJECT INFO Арр ECOSYSTEM Stream CLIENTS Арр Processo **EVENTS** CONTACT US APACHE App DB App DB ... 💓 @apachekafka

Apache ZooKeeper™ Project + Documentation + Developers + ASF +

Welcome to Apache ZooKeeper?

Apache ZooKeeper is an effort to develop and maintain an open-source server which enables highly reliable distributed coordination

What is ZooKeeper?

ZooKeeper is a centralized service for maintaining configuration information, naming, providing distributed synchronization, and providing group services. All of these kinds of services are used in some form or another by distributed applications. Each time they are implemented there is a lot of work that goes into fixing the bugs and race conditions that are inevitable. Because of the difficulty of implementing these kinds of services, applications initially usually skimp on them, which make them brittle in the presence of change and difficult to manage. Even when done correctly, different implementations of these services lead to management complexity when the applications are deployed.

Learn more about ZooKeeper on the ZooKeeper Wiki.

Getting Started

Start by installing ZooKeeper on a single machine or a very small cluster.

1. Learn about ZooKeeper by reading the documentation. 2. Download ZooKeeper from the release page.

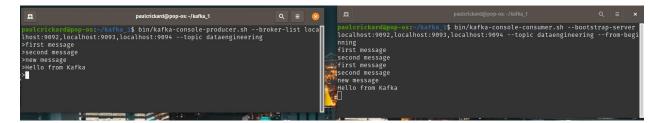
Getting Involved

Apache ZooKeeper is an open source volunteer project under the Apache Software Foundation. We encourage you to learn about the project and contribute your expertise. Here are some starter links:

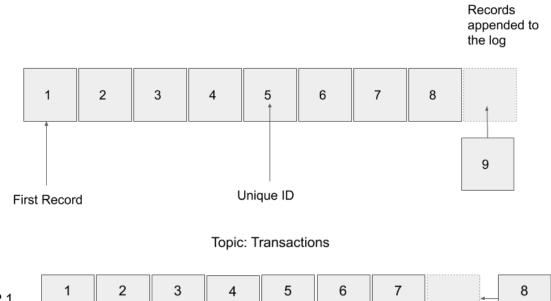
- 1. See our How to Contribute to ZooKeeper page
- 2. Give us feedback: What can we do better? 3. Join the mailing list: Meet the community.
- Copyright @ 2010-2020 The Apache Software Foundation, Licensed under the Apache License, Version 2.0.

Apache ZooKeeper, ZooKeeper, Apache, the Apache feather logo, and the Apache ZooKeeper project logo are trademarks of The Apache Software Foundation.





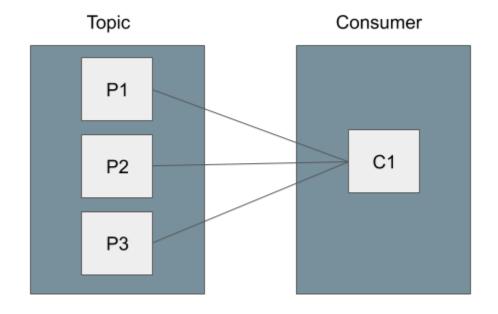
Chapter 13: Streaming Data with Kafka

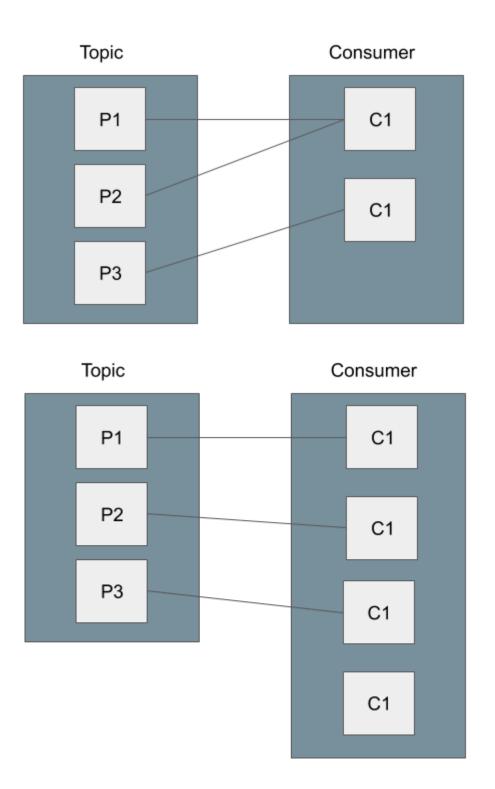


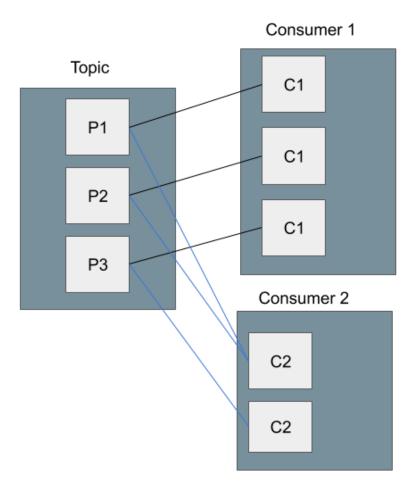
P 1	K:A	 K:A							
P 2	1 K:B	2 K:B	3 K:B	4 K:B		5 K:B			
P 3	1 K:C	2 K:C	3 K:C	4 K:C	5 K:C		6 K:C		

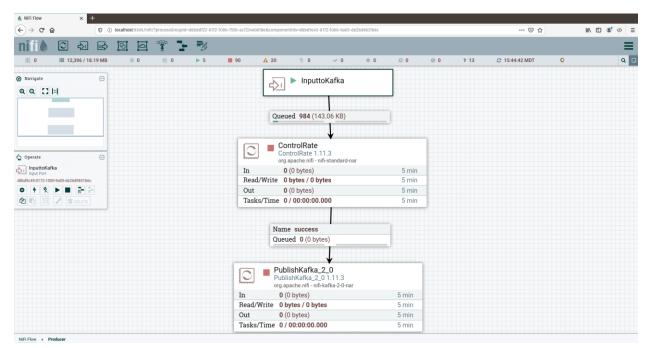
Older

Newer

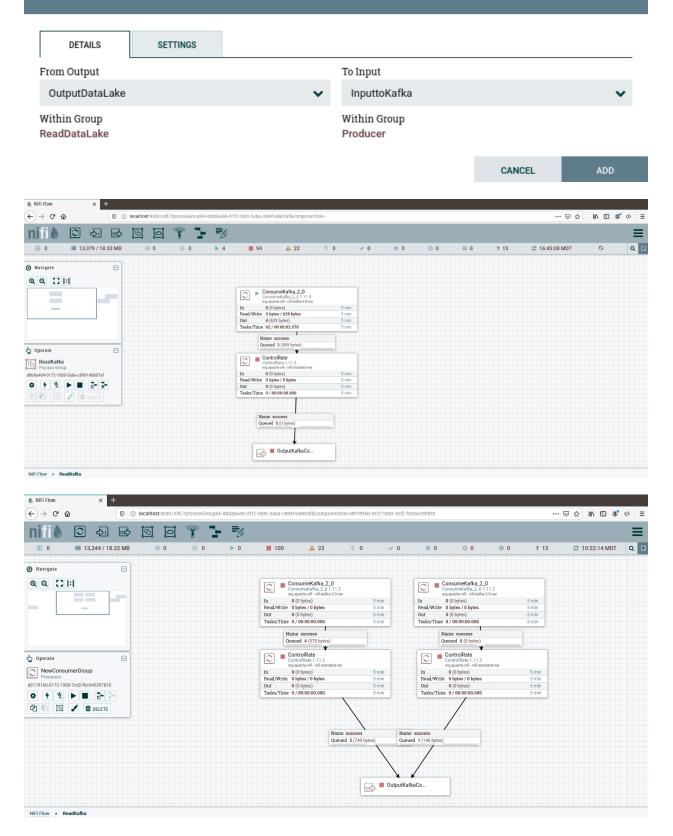


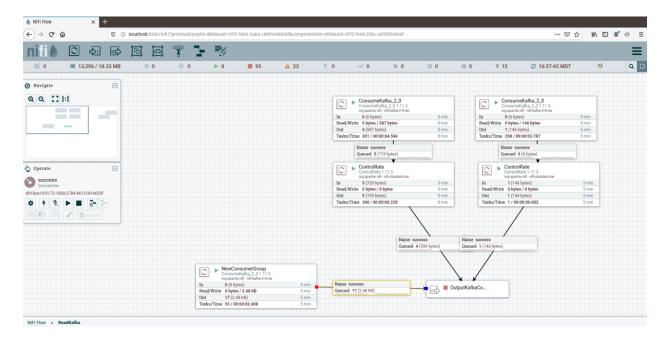


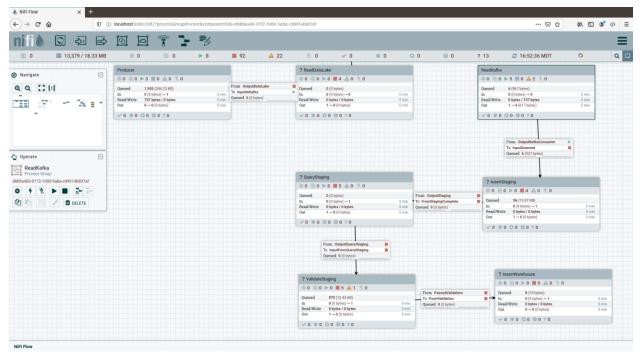


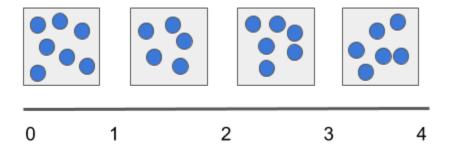


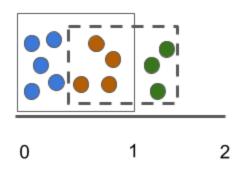
Create Connection



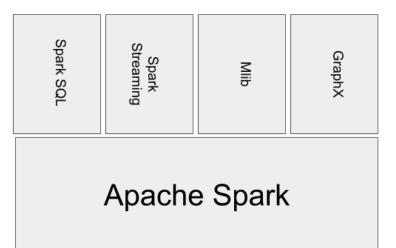


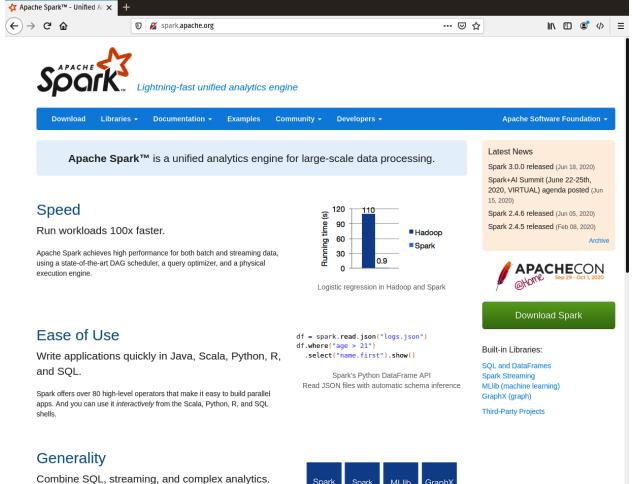






Chapter 14: Data Processing with Apache Spark





Spark Spark MLlib GraphX SQL Streaming (machine (graph) learning)

	raries 👻 Do	ocumentation ·	- Examples	Community -	Developers 👻			Apache	Software F	oundation
Download A	∆nache	Snark	M					Latest News	6	
	-	-						Spark 3.0.0 r		
1. Choose a Spark re	elease: 3.0.0	(Jun 18 2020)	~					Spark+AI Sui 2020, VIRTU	•	
2. Choose a package	e type: Pre-b	uilt for Apache I	Hadoop 2.7	~				15, 2020)	, ,	•
3. Download Spark:	spark-3.0.0-bin	-hadoop2.7.tgz						Spark 2.4.6 r	`	
	using the 3.0.0) signatures, che	ecksums and proje	ct release KEYS.				Spark 2.4.5 r	eleased (Fel	b 08, 2020)
12. ark Master at spark://po × → C û and and and and and and and and and and	+ () () localho Master at s	st:8080	os.localdoma		Cald 2.12. Spark 0.0	+ is pre-built v		6	ז לב על אין	I\ [] (\$ 42
bet that, Spark 2.x is print ark Master at spark://poi \times \rightarrow C $\hat{\omega}$ \therefore spark/hop-os.localdomain: ε workers: 1 es in use: 2.7 foital, 0.0 used nory in use: 2.7 foital, 0.0 used: nory in use: 2.7 foital, 0.0 used: hory in used: hory in used: hory in use: 2.7 foital, 0.0 used: hory in used: hory in used: hory in used: hory in used:	+ The contract of the contrac	st:8080				+ is pre-duiit v		6	7☆⊻∥	N 🗉 🕸 4:
te that, Spark 2.x is pr 12. → C ŵ → C ŵ 2.0.0 Spark I L: spark//pop-os.localdomain: re Workers: 1 res in use: 2 Total, 0 Used mory in use: 2.7 GIB Total, 0.Comp vers: 0 Running, 0 Completed tus: ALIVE	+ () () locatho Master at s 7077 D B Used oleted	st:8080				+ is pre-duiit v		6	J ☆ ⊻ ∥	N 🖸 🐠 ()
ote that, Spark 2.x is pr 12. → C ☆ → C ☆ 2.0.0 Spark I L: spark//pop-os.localdomain: ve Workers: 1 res in use: 2 Tolk 0 Used mory in use: 2.7 GIB Total, 0.05 sources in use: 0 Running, 0 Completed tus: ALIVE Workers (1)	+ () () localho Master at s 7077 D B Used oleted	st:8080		in:7077				6		
tick that, Spark 2.x is price that, Spark 2.x is price that, Spark 2.x is price that a spark.//poi × → C ☆ ↓ → C ☆ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓	+ Master at s 7077 D B Used Deted	st:8080			State	+ IS pre-Duilt V	Memory 2.7 GB (0.0 B 1			N 🗊 😰 🕫
the that, Spark 2.x is print bark Master at spark://poix → C û 3.0.0 Spark I L: spark.//pop-os.localdomain: we Workers: 1 es in use: 2 Total, 0 Used mory in use: 2.7 GIB Total, 0.0 bilications: 0 Running, 0 Completed tus: ALIVE Workers (1) orker Id	+ () () locatho Master at s 7077 D B Used Dieted 148-9911	st:8080		in:7077 Address	State	Cores	Memory			

Welcome to /__/___/___//__//__/ _\ \/ _ \/ _ `/ __/ '_/ /__ / .__/_,_/_/ /_/_\ version 3.0.0 /_/

Using Python version 3.7.5 (default, Nov 20 2019 09:21:52) SparkSession available as 'spark'.

) Get Started with PySpark a	× 😄 Ċ Home Pag	ge - Select or creat 🗙	릗 Spark -	Jupyter Notebook	× 褖 E	kamples	Apache Spark 🗙	+							
-)→ ൙ û		localhost:8888/noteboo	ks/Spark.ip	ynb								• ⊠ ☆	<u>↓</u> III\	•	° <>
		🗂 Jup	yter s	park Last Checkp	ioint: an hour	ago (un	saved changes)				2 Logout				
		File	- Edit Vie	w Insert C	ell Kerne	el Hel	n			Trusted	Python 3 O				
			≫ दी ।		Run 🔳 🕻					Indated	Python's O				
			e 40	•) • • •	Kull 📕 🤇	-	Code 🗾 🖻								
			In [1]:	<pre>import finds findspark.in</pre>	park it()										
			In [18]:	import pyspa from pyspark	.sql impo	rt Spar	kSession								
			Te [21].	spark=SparkS	ession.bu	ilder.m	aster("spark:/,	/pop-os	.localdomain:	077").get0rC	reate()				
			10 [21]:	<pre>import random NUM_SAMPLES=</pre>	1										
				def inside(p x, y = ra return x		dom(), < 1	random.random()								
				<pre>count = spar print("Pi is</pre>	k.sparkCo roughly	ntext.p {}".for	arallelize(rang mat(4.0 * count	ge(0,NU	M_SAMPLES)).fi _SAMPLES))	lter(inside).	.count()				
				Pi is roughly	y 4.0										
			In [22]:	spark.stop()											
Get Started with PySpark at 🗙	🔵 Home Page - S	Select or creat 🗙 🛛 🥭 Spa	ark - Jupyter I	Notebook 🗙 🙀	Examples A	pache Spa	ark 🗙 📾 Spark M	laster at sp	park://por 🗙 🛛 🍪 p	yspark.sql module ·	– Pysp 🗙 🕂				
-)→ ሮ û	🗊 🛈 local	host:8080										⊌	☆ ⊻ ∥		3°
M		spark://pop-o													
status: ALIVE															
Worker Id worker-20200629125219-10.0.	0 1/8-9911				ddress 0.0.0.148:991	1	State	Cores 2 (2 Use	d) 2.7 GiB (1	024.0 MiB Used)			Re	ources	
				1	0.0.0.140.001	*		2 (2 050	u) 2.1 UD (3	524.0 Milb 0300)					
 Running Applications Application ID 	(1)	Name		Cores Me	emory per Ex	recutor	Resources Per Exe	cutor	Submitted	Time	User		State		Duration
app-20200629195112-0002		(kill) Pi-Estimati	on		24.0 MiB	coutor	TRESOURCES I CI EXC	outor	2020/06/29		paulcrickar	ł	RUNNING		4 s
Completed Application	ıs (2)					Beer			Culoulus d Thurs		11			-	
Application ID app-20200629194952-0001		Name pyspark-shell	Cores	Memory pe 1024.0 MiB	r Executor	Reso	urces Per Executor		Submitted Time 2020/06/29 19:49	:52	Daulcrickard		State FINISHED		Duration
app-20200629180816-0000		pyspark-shell	2	1024.0 MiB					2020/06/29 18:08	:16	paulcrickard		INISHED	3	84 min
+	+ .	+			+-			+		+		4	+		
1	_c0 _			-	c2		_c3	3	_c4	_c5		_c6			_c7
, I	name a	•		stre	+- 0-+1		city	·+ /	state	++		1 0 0			lat
 Patrick Her			lonat												
		36 2502 S													
Arthur Ga							manhaver				-39.84				
Gary Valer											-30.304				
+															
only showing															
root															
c0: s1	tring (nullable :	= tru	ie)											
c1: st	tring (nullable :	= tru	ie)											
c2: st	tring (nullable :	= tru	ie)											
c3: s1															
c4: st	tring (م ا ا د ا ا س	- + -	<u>م</u> ا											

|-- _c4: string (nullable = true)
|-- _c5: string (nullable = true)
|-- _c6: string (nullable = true)
|-- _c7: string (nullable = true)

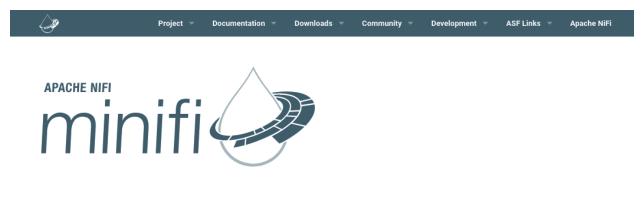
	street	city	state zip	lng	lat
Patrick Hendrix 23	5755 Jonathan Ranch	New Sheriland W	isconsin 60519	103.914462 -	59.0094375
Grace Jackson 36	2502 Stewart Plaz	Ramirezville	Arizona 91946	170.503858	58.1631665
Arthur Garcia 61	627 Liu Brooks	Freemanhaven	Kansas 97783	-39.845646	38.689889
Gary Valentine 29	9682 Theresa Vist	Allenborough	Oregon 81537	-30.304522	81.2722995
Erin Mclean 23	9349 Williams Lan	East Markmouth	Ohio 4300	-110.860085	11.476733

only showing top 5 rows

root

|-- name: string (nullable = true)
|-- age: integer (nullable = true)
|-- street: string (nullable = true)
|-- city: string (nullable = true)
|-- state: string (nullable = true)
|-- zip: integer (nullable = true)
|-- lng: double (nullable = true)
|-- lat: double (nullable = true)

Chapter 15: Real-Time Edge Data – Kafka, Spark, and MiNiFi



A subproject of Apache NiFi to collect data where it originates.

About

MiNiFi—a subproject of Apache NiFi—is a complementary data collection approach that supplements the core tenets of NiFi in dataflow management, focusing on the collection of data at the source of its creation.

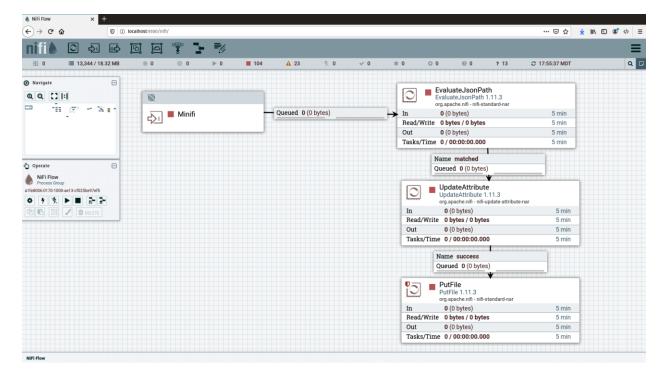
Specific goals for the initial thrust of the MiNiFi effort comprise:

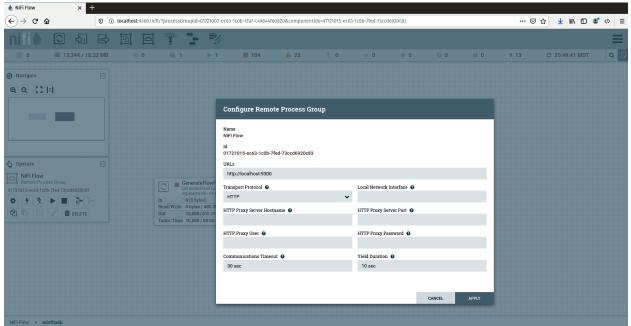
- Small size and low resource consumption
- · Central management of agents
- Generation of data provenance (full chain of custody of information)
- · Integration with NiFi for follow-on dataflow management

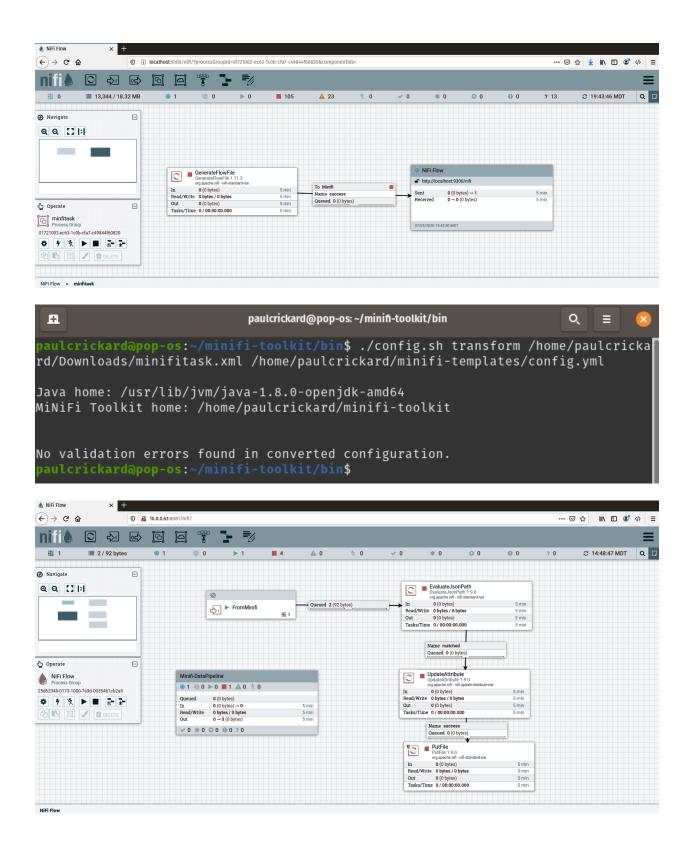
Perspectives of the role of MiNiFi should be from the perspective of the agent acting immediately at, or directly adjacent to, source sensors, systems, or servers.

```
if ! shopt -oq posix; then
    if [ -f /usr/share/bash-completion/bash_completion ]; then
        . /usr/share/bash-completion/bash_completion
    elif [ -f /etc/bash_completion ]; then
        . /etc/bash_completion
    fi
    fi
    fi
    export JAVA_HOM=/usr/lib/jvm/java-1.11.0-openjdk.amd64
    export SPARK_HOME=/home/paulcrickard/spark3
    export SPARK_HOME=/home/paulcrickard/spark3
    export PATH=$SPARK_HOME/bin:$PATH
    export MINIFI_HOME=/home/paulcrickard/minifi
    export PATH=$MINIFI_HOME/bin:$PATH
```

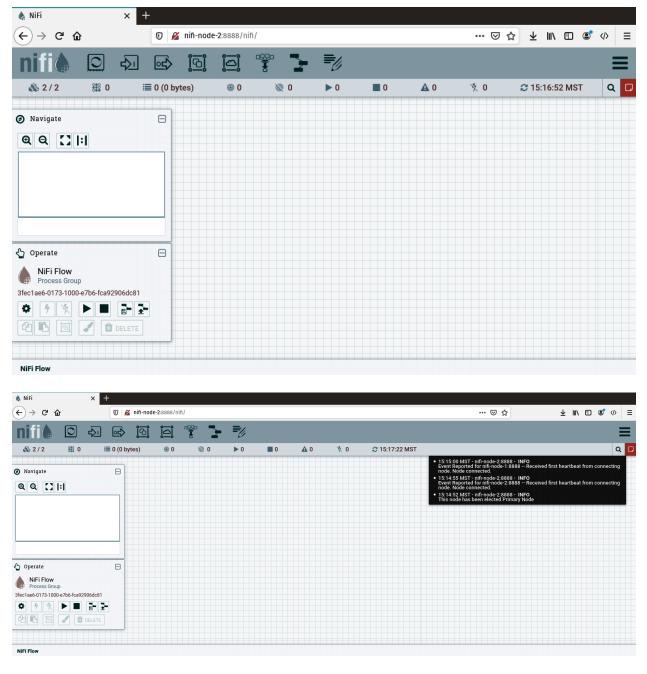
Site to Site properties nifi.remote.input.host= nifi.remote.input.secure=false nifi.remote.input.socket.port=1026 nifi.remote.input.http.enabled=true nifi.remote.input.http.transaction.ttl=30 sec nifi.remote.contents.cache.expiration=30 secs





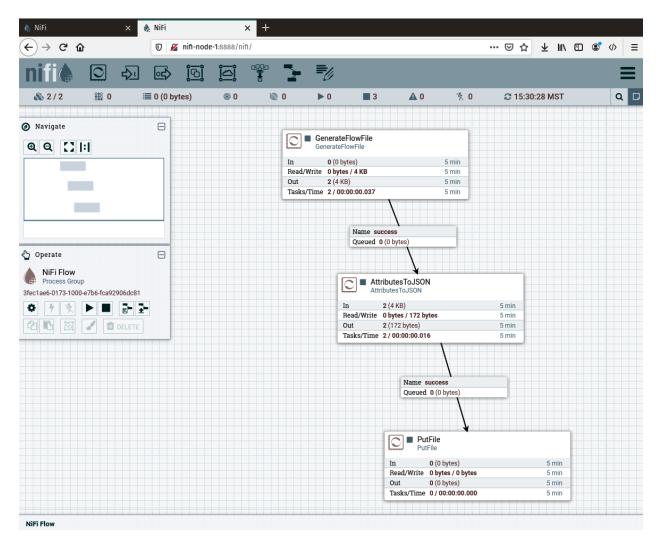


Appendix



NiFi Cluster Displaying 2 of 2

Filter		by address 🗸 🗸						
	Node Address 🔺		Active Thread Count	Queue / Size	Status	Uptime	Last Heartbeat	
0	nifi-node-1:8888		0	0 / 0 bytes	CONNECTED	07/11/2020 16:11:47 MDT	07/11/2020 16:18:51 MDT	Ф
0	nifi-node-2:8888		0	0 / 0 bytes	CONNECTED, PRIMARY, COORDINATOR	07/11/2020 16:14:48 MDT	07/11/2020 16:18:47 MDT	Ф



£		paulcrickard@pop-os	s: ~/output	Q = 8
paulcrickard@po	<pre>op-os:~/output\$</pre>	ls		Ī
4274490619179	4314534837026	4354551784292	4394578909080	data.txt
4284523756564	4324537815176	4364556613360	4404582691395	
4294527247783	4334542190236	4374560985786	4414585033417	
4304531894134	4344546106426	4384567644867	4424589700344	
	<pre>op-os:~/output\$</pre>			
			d":"556872b1-f4	34-47f1-a610-5893f319
2441"}paulcric	card@pop-os:~/ou	utput\$		

🎄 Nifi 🛛 🗙	🍖 Nifi × +	
↔ → ♂ ŵ	🔽 🔏 niñ-node-2:8888/niñ/	
nifi 🖸 🕹		≡
land 2 / 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	≡ 0 (0 bytes)	© 15:31:03 MST Q
Navigate		
@ Q []	C = GenerateFlowFile In Configure Processor	_
	Read/Wi Out	
	Tasks/T SETTINGS SCHEDULING PROPERTIES	COMMENTS
	Scheduling Strategy	Run Duration 0ms 25ms 50ms 100ms 250ms 500ms 1s 2s
🖒 Operate	Timer driven V Timer driven V Run Schedule e	Lower latency Higher throughput
GenerateFlowFile Processor	On primary node 10 sec	
3ff6813b-0173-1000-ffff-ffffcdd7bb81	CRON driven	
* * * • • •		
		CANCEL APPLY
	PutFile	
	In 0 (0 bytes) Read/Write 0 bytes / 0 bytes	
	Out 0 (0 bytes) Tasks/Time 0 / 00:00:000	
	Tasks/Time 0/00:00:00	
NiFi Flow		

NiFi Cluster

	Cluster ing 2 of 2							
Filter		by address	~					
	Node Address 🔺		Active Thread Count	Queue / Size	Status	Uptime	Last Heartbeat	
0	nifi-node-1:8888		0	4 / 2.25 KB	CONNECTED	07/11/2020 16:11:47 MDT	07/11/2020 16:38:55 MDT	ڻ
0	nifi-node-2:8888		0	4 / 2.25 KB	CONNECTED, PRIMARY, COORDINATOR	07/11/2020 16:14:48 MDT	07/11/2020 16:38:54 MDT	Φ

C Last updated: 16:38:55 MDT

NiFi Cluster

|--|

Node Address 🔺	Active Thread Count	Queue / Size	Status	Uptime	Last Heartbeat	
Inifi-node-1:8888	0	8 / 10.25 KB	CONNECTED	07/11/2020 17:00:55 MDT	07/11/2020 17:06:41 MDT	¢
o nifi-node-2:8888	0	5 / 4.25 KB	CONNECTED, PRIMARY, COORDINATOR	07/11/2020 16:54:18 MDT	07/11/2020 17:06:41 MDT	Φ

×

C Last updated: 17:06:41 MDT

NiFi Cluster

Display	ing 2 of 2							
Filter		by address 🗸	•					
	Node Address 🔺		Active Thread Count	Queue / Size	Status	Uptime	Last Heartbeat	
0	nifi-node-1:8888				DISCONNECTED	No value previously set	No value previously set	¢û
0	nifi-node-2:8888		0	11 / 16.25 KB	CONNECTED, PRIMARY, COORDINATOR	07/11/2020 16:54:18 MDT	07/11/2020 17:07:36 MDT	Φ

C Last updated: 17:07:38 MDT

NiFi Cluster

Display Filter	ing 2 of 2	by address	~					
	Node Address 🔺		Active Thread Count	Queue / Size	Status	Uptime	Last Heartbeat	
0	nifi-node-1:8888		0	13 / 20.25 KB	CONNECTED	07/11/2020 17:00:55 MDT	07/11/2020 17:08:32 MDT	Φ
0	nifi-node-2:8888		0	13 / 20.25 KB	CONNECTED, PRIMARY, COORDINATOR	07/11/2020 16:54:18 MDT	07/11/2020 17:08:32 MDT	Q

C Last updated: 17:08:35 MDT