



Dataflow Management From Edge to Core with Apache NiFi

Andy LoPresto | @yolopey

Sr. Member of Technical Staff at Hortonworks, Apache NiFi PMC & Committer

06 February 2019 Dataworks Summit Melbourne

Acknowledgement of Country

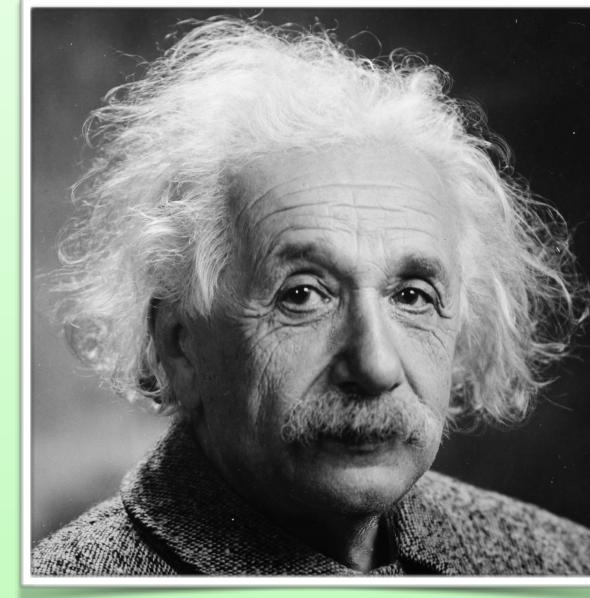
I acknowledge the Traditional Owners of the land on which we are meeting. I pay my respects to their Elders, past and present, and the Aboriginal Elders of other communities who may be here today.

Gauging Audience Familiarity With NiFi



“What’s a NeeFee?”

No experience with dataflow
No experience with NiFi



“I can pick this up pretty quickly”

Some experience with dataflow
Some experience with NiFi



“I refactored the Ambari integration endpoint to allow for mutual authentication TLS during my coffee break”

Forgotten more about NiFi than most of us will ever know

Agenda

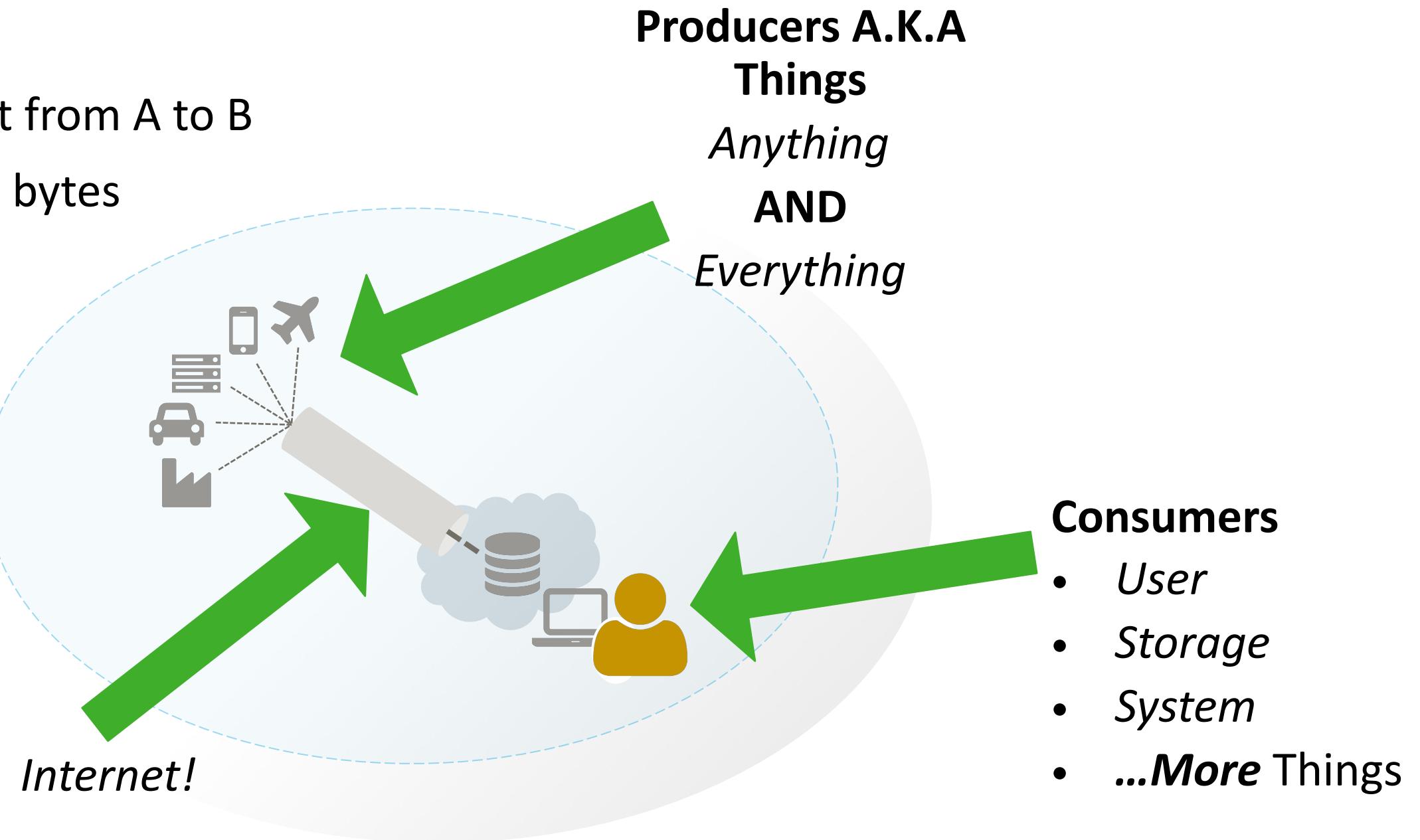
- *What is dataflow and what are the challenges?*
 - Apache NiFi
 - Apache MiNiFi
 - Apache NiFi Registry
 - Complementary Tools
 - Community
-
- All slides provided online, so no need to transcribe



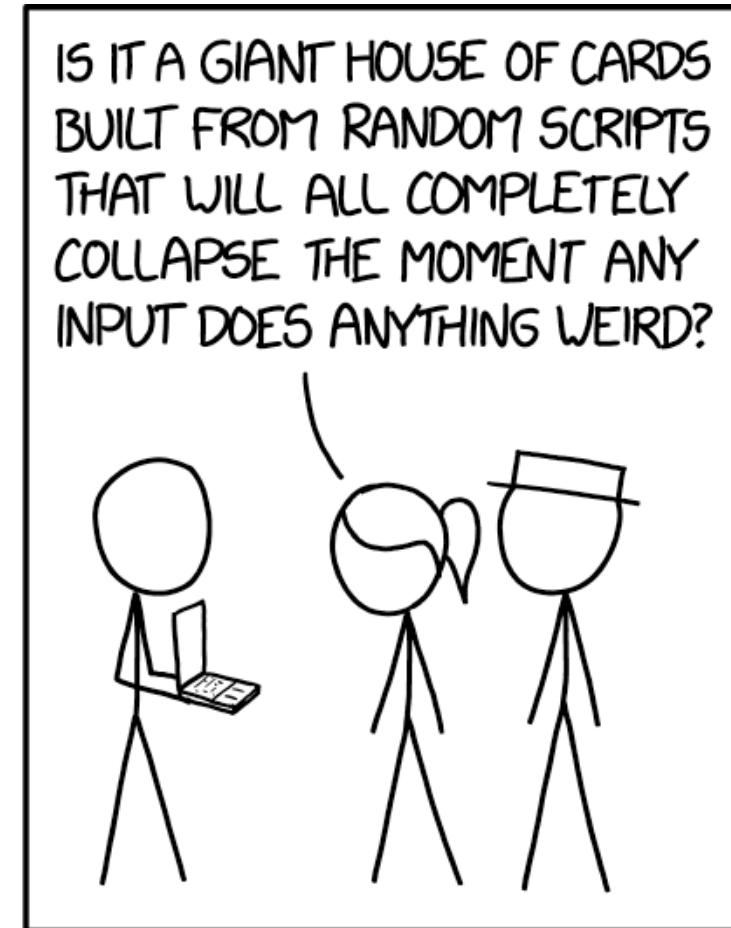
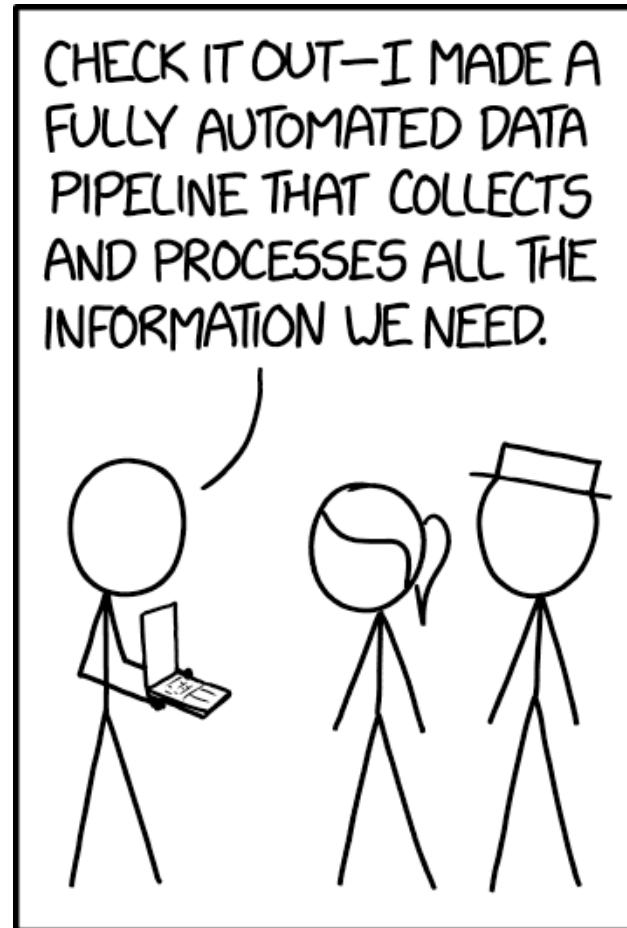
What is dataflow?

What is dataflow?

- Moving some content from A to B
- Content could be any bytes
 - Logs
 - HTTP
 - XML
 - CSV
 - Images
 - Video
 - Telemetry



Moving data effectively is hard



"Data Pipeline" <https://xkcd.com/2054/>

Dataflow Challenges In 3 Categories

Data

- Standards
- **Formats**
- Protocols
- Veracity
- Validity
- Schemas
- Partitioning/
Bundling

Infrastructure

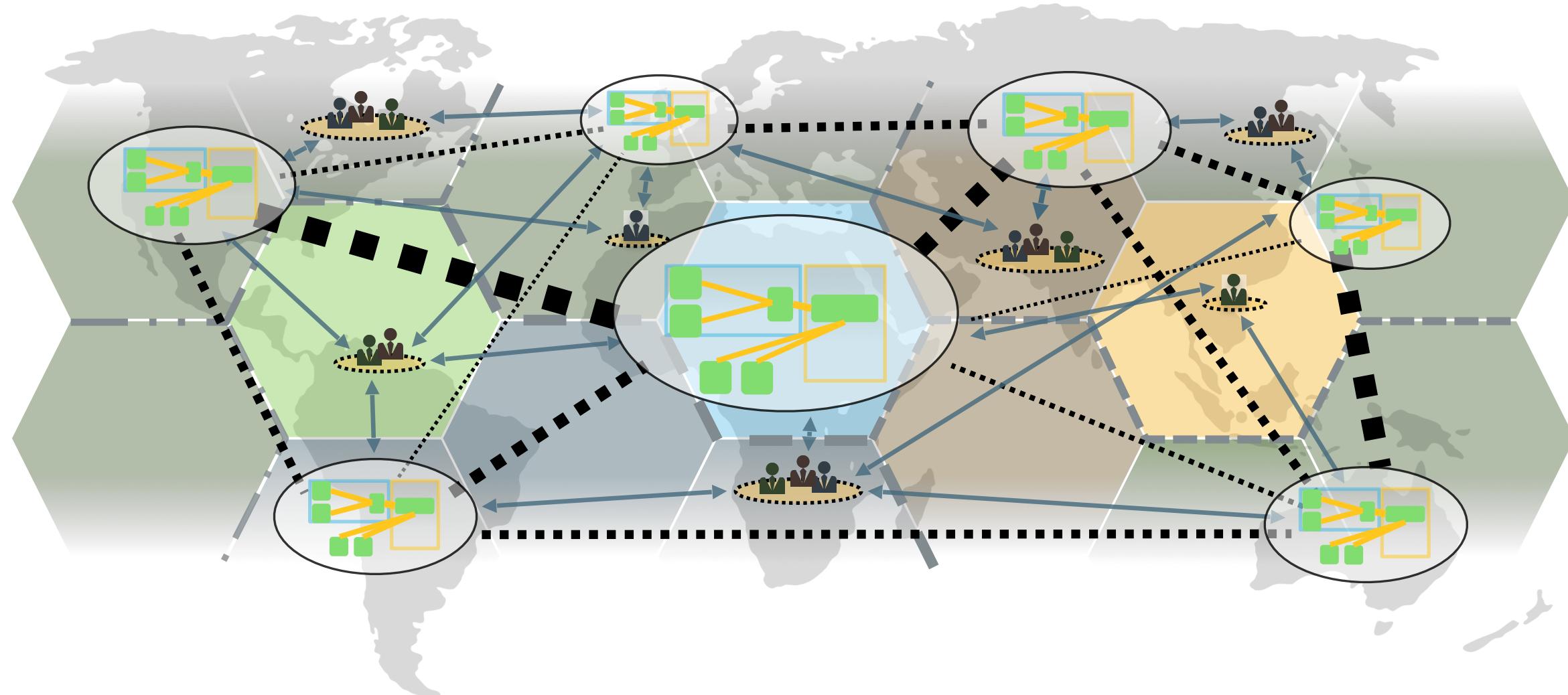
- “Exactly Once”
Delivery
- Ensuring
Security
- **Overcoming**
Security
- Credential
Management
- Network

People

- Compliance
- “**That** [person |
team | group]”
- **Consumers
Change**
- **Requirements
Change**
- “Exactly Once”
Delivery

Let's Connect Lots of As to Bs to As to Cs to Bs to Δ s to Cs to φ s

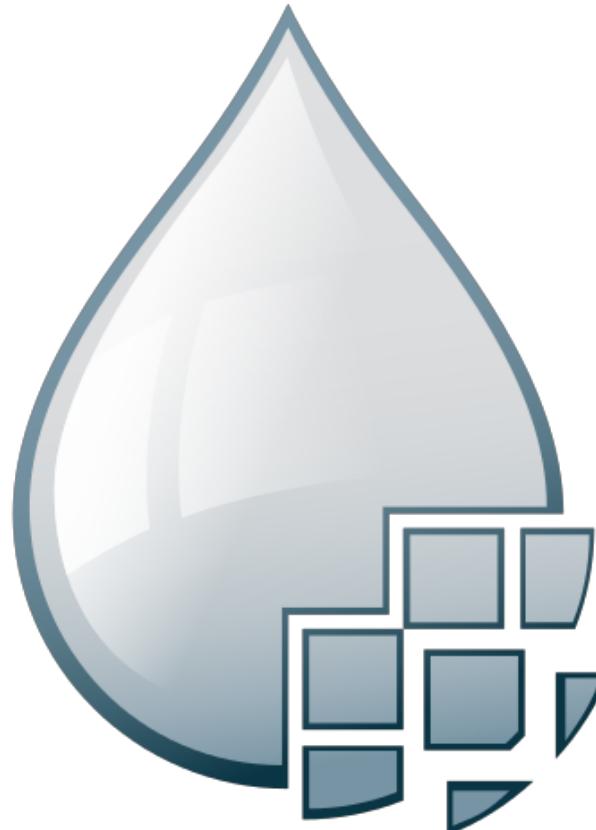
Raise your hand if you want to maintain Python scripts for the rest of your life



Apache NiFi

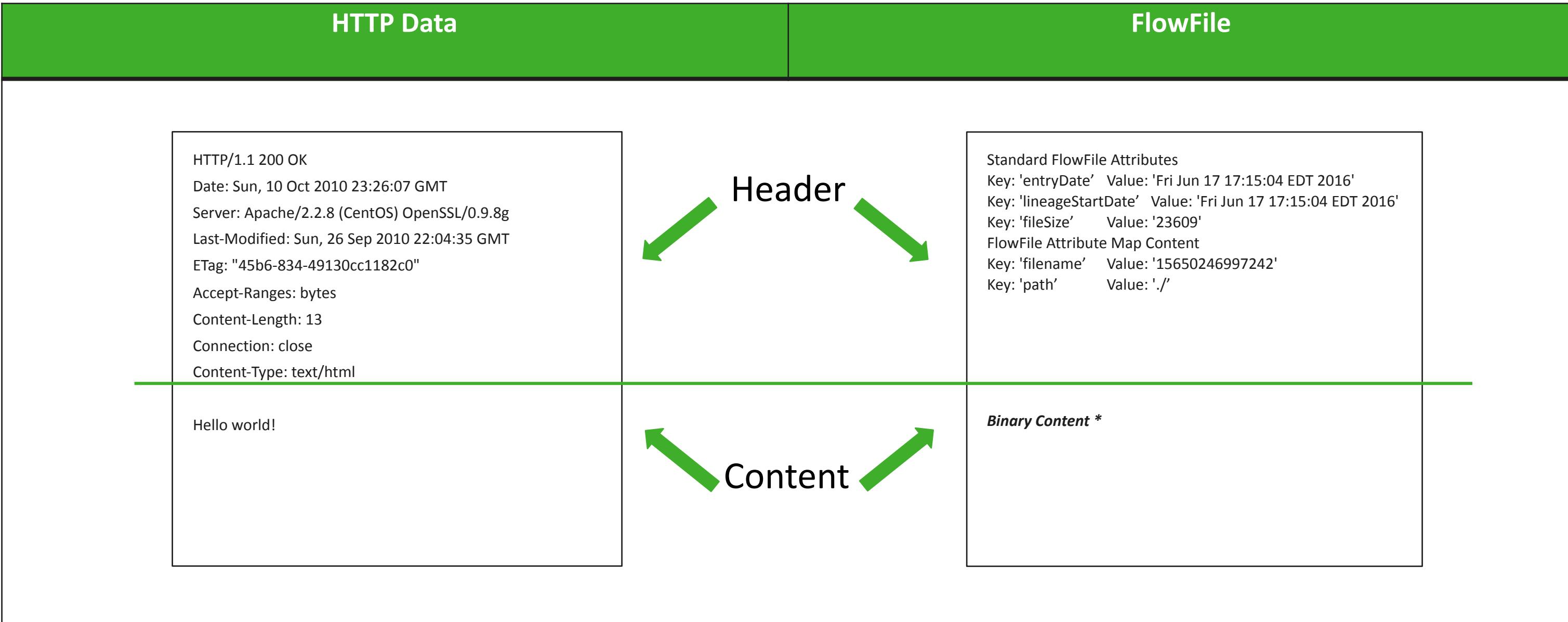
Apache NiFi

Key Features



- Guaranteed delivery
 - Data buffering
 - Backpressure
- Pressure release
- Prioritized queuing
- Flow specific QoS
 - Latency vs. throughput
 - Loss tolerance
- Data provenance
- Supports push and pull models
- Recovery/recording a rolling log of fine-grained history
- Visual command and control
- Flow templates
- Pluggable, multi-tenant security
- Designed for extension
- Clustering

Flowfiles Are Like HTTP Data



User Interface

Less of this...

```
1. scratch/release_verification (bash)
nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT (bash)
nifi.flowfile.repository.always.sync=false
nifi.swap.manager.implementation=org.apache.nifi.controller.FileSystemSwapManager
nifi.queue.swap.threshold=20000
nifi.swap.in.period=5 sec
nifi.swap.in.thresholds=1
nifi.swap.out.period=5 sec
nifi.swap.out.thresholds=4
nifi.swap.out.threads=4

# Content Repository
nifi.content.repository.implementation=org.apache.nifi.controller.repository.FileSystemRepository
hw12203:.../space/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT (NIFI-3313) alopresto
⌚ 157s @ 17:40:39 $ nfk1

Java home: /Users/alopresto/.jenv/versions/1.8
NiFi home: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT

Bootstrap Config File: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/conf/bootstrap.conf

2017-01-30 17:40:43,812 INFO [main] org.apache.nifi.bootstrap.Command Apache NiFi has accepted the Shutdown Command and is shutting down now
Name: SUCCESS
2017-01-30 17:40:43,854 INFO [main] org.apache.nifi.bootstrap.Command Waiting for Apache NiFi to finish shutting down...
2017-01-30 17:40:45,867 INFO [main] org.apache.nifi.bootstrap.Command NiFi has finished shutting down.

hw12203:.../space/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT (NIFI-3313) alopresto
⌚ 9s @ 17:40:49 $ subl conf/nifi.properties
hw12203:.../space/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT (NIFI-3313) alopresto
⌚ 5s @ 17:40:55 $ nfgo

Java home: /Users/alopresto/.jenv/versions/1.8
NiFi home: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT

Bootstrap Config File: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/conf/bootstrap.conf

2017-01-30 17:41:13,874 INFO [main] org.apache.nifi.bootstrap.Command Starting Apache NiFi...
2017-01-30 17:41:13,874 INFO [main] org.apache.nifi.bootstrap.Command Working Directory: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT
2017-01-30 17:41:13,874 INFO [main] org.apache.nifi.bootstrap.Command Command: /Users/alopresto/.jenv/versions/1.8/bin/java -classpath /Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.conf:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/jcl-over-slf4j-1.7.12.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/jul-to-slf4j-1.7.12.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/logback-classic-1.1.3.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/logback-core-1.1.3.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-api-1.2.0-SNAPSHOT.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-documentation-1.2.0-SNAPSHOT.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-framework-api-1.2.0-SNAPSHOT.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-nar-utils-1.2.0-SNAPSHOT.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-properties-1.2.0-SNAPSHOT.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-runtime-1.2.0-SNAPSHOT.jar:/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/.lib/nifi-slf4j-api-1.7.12.jar -Dorg.apache.jasper.compiler.disablejsr199=true -Xmx512m -Xms12m -Dsun.net.http.allowRestrictedHeaders=true -Djava.net.preferIPv4Stack=true -Djava.awt.headless=true -XX:+UseG1C -Djava.protocol.handler.pkgs=sun.net.www.protocol -Dnifi.properties.path=/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT/.conf/nifi.properties -Dnifi.bootstrap.listen.port=5972 -Dapp=Nifi -Dorg.apache.nifi.bootstrap.config.log.dir=/Users/Alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT/logs org.apache.nifi.NiFi

hw12203:.../space/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT (NIFI-3313) alopresto
⌚ 21s @ 17:41:17 $ nfk1

Java home: /Users/alopresto/.jenv/versions/1.8
NiFi home: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT

Bootstrap Config File: /Users/alopresto/Workspace/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT/conf/bootstrap.conf

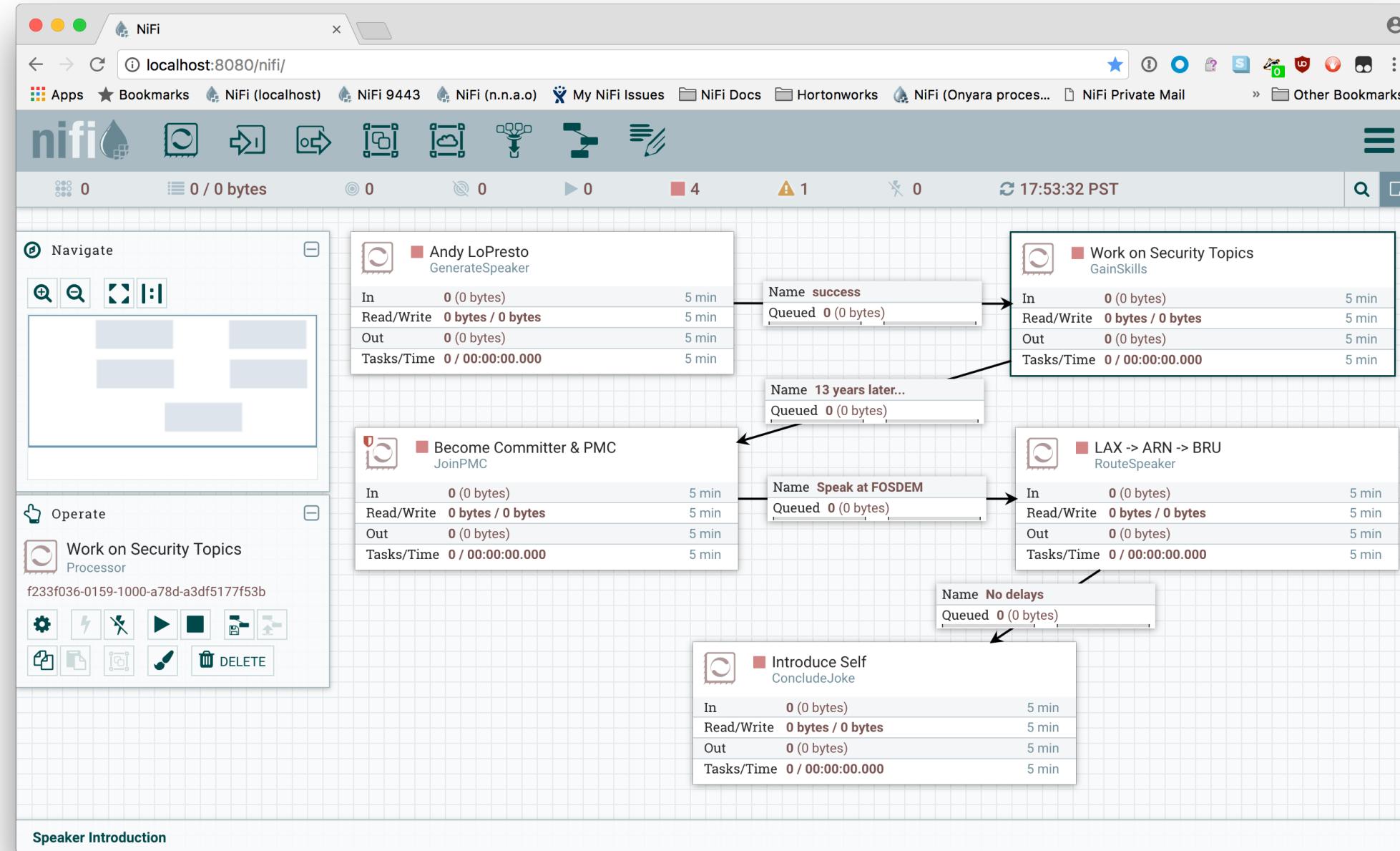
2017-01-30 20:54:26,971 INFO [main] org.apache.nifi.bootstrap.Command Apache NiFi has accepted the Shutdown Command and is shutting down now
2017-01-30 20:54:27,075 INFO [main] org.apache.nifi.bootstrap.Command Waiting for Apache NiFi to finish shutting down...
2017-01-30 20:54:29,085 INFO [main] org.apache.nifi.bootstrap.Command NiFi has finished shutting down.

hw12203:.../space/nifi/nifi-assembly/target/nifi-1.2.0-SNAPSHOT-bin/nifi-1.2.0-SNAPSHOT (NIFI-3313) alopresto
⌚ 1159s @ 20:54:33 $ 
```

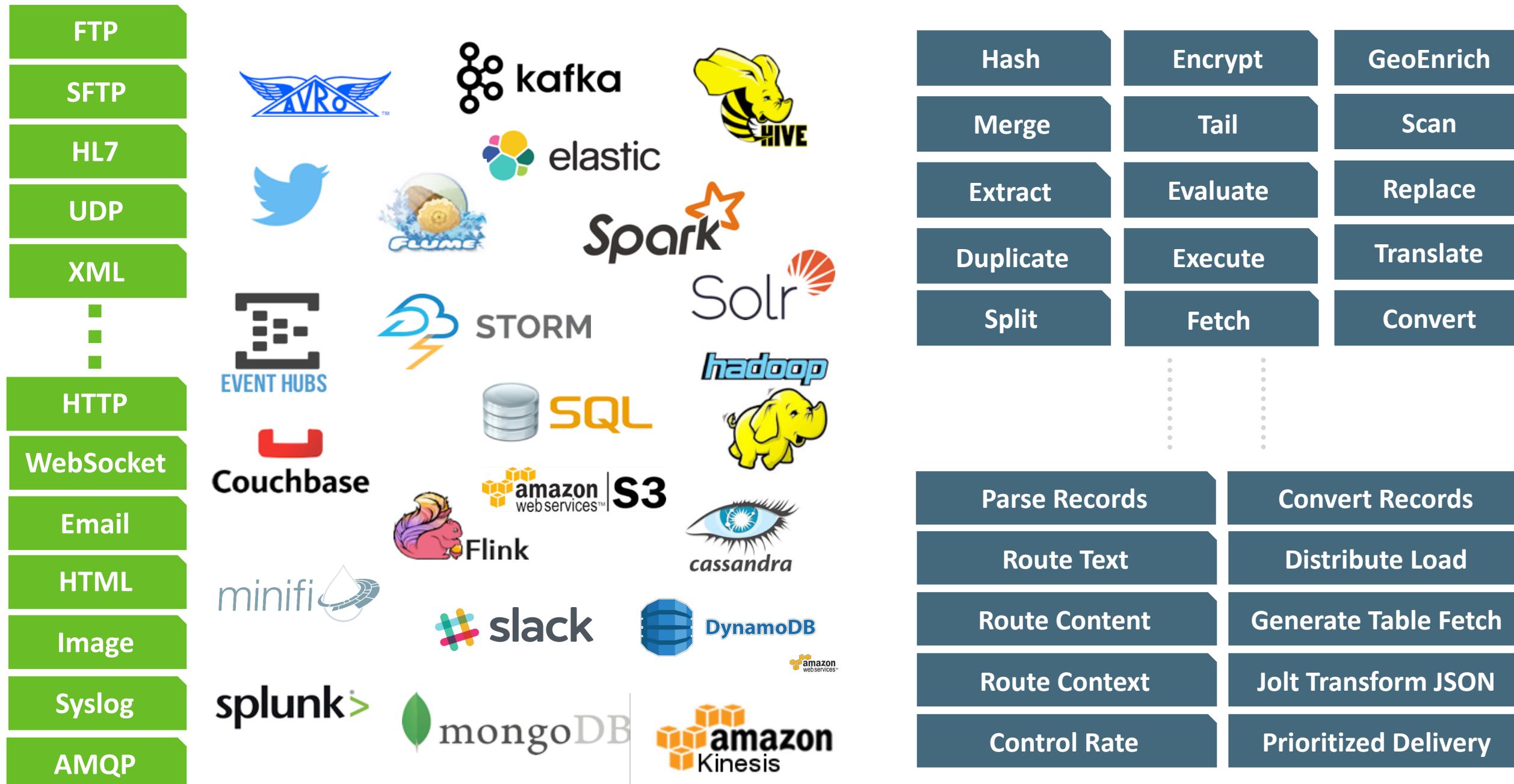


User Interface

Less of this..... more of this



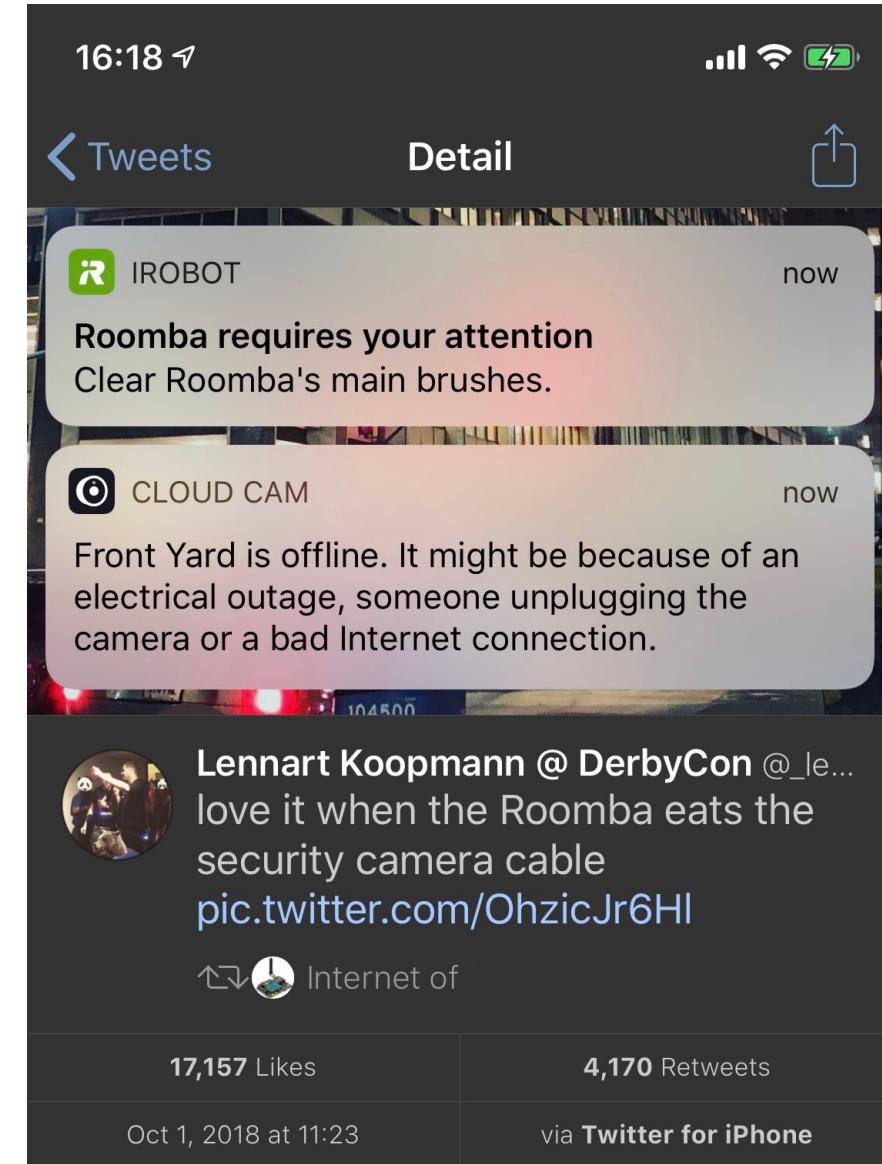
Deeper Ecosystem Integration: 286+ Processors, 61 Controller Services



Apache MiNiFi

IoT Challenges

- Limited computing capability
- Limited power/network
- **Restricted software library/platform availability**
- **No UI**
- Physically inaccessible
- Not frequently updated
- **Competing standards/protocols**
- Scalability
- **Privacy & Security**



@_lennart

So Why Do We Need A Different Solution?

- NiFi is designed to “own the box”
- NiFi 0.7.x started up in about 10-15 minutes on RP3 (593 MB)
- NiFi 1.x started up in about 30 minutes on RP3 (760 MB)
 - 33 new processors
 - Rewrite for multi tenant authorization
 - Complete UI overhaul

```
▶hw12203:/Users/alopresto/Workspace/scratch/rp3b-demo (master) alopresto
└─ 113s @ 17:09:05 $ ssh pi@my-raspberry-pi
^C
▶hw12203:/Users/alopresto/Workspace/scratch/rp3b-demo (master) alopresto
└─ 145s @ 17:09:37 $ █
```

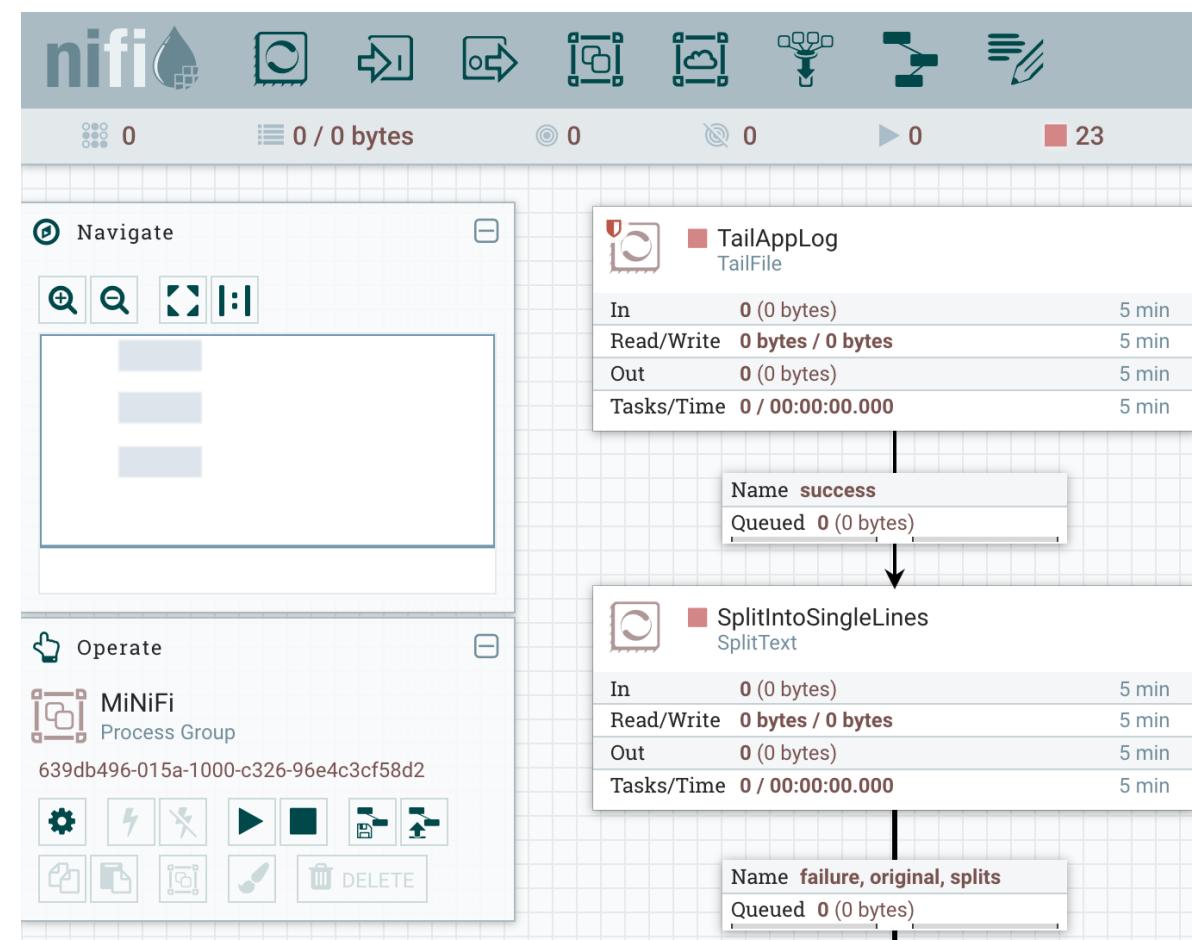
Apache NiFi Subproject: MiNiFi

- Get the key parts of NiFi close to where data begins and provide bidirectional communication
- NiFi lives in the data center — give it an enterprise server or a cluster of them
- MiNiFi lives as close to where data is born and is a guest on that device or system
 - IoT
 - Connected car
 - Legacy hardware

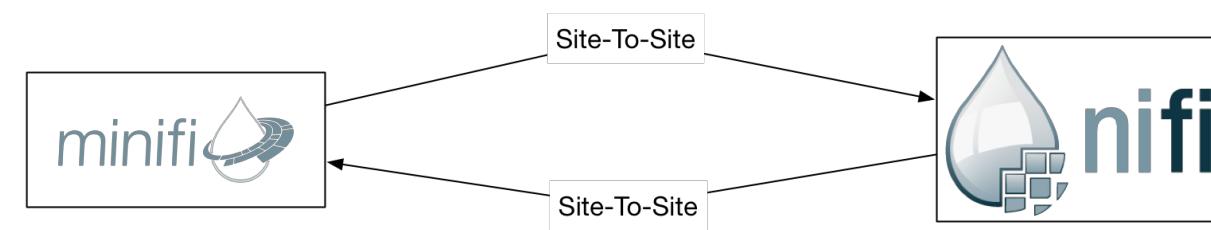


Flavors of MiNiFi

- MiNiFi Java (v0.5.0)
 - Modified version of NiFi
 - No UI
 - YAML configuration
 - Reduced processor count
 - 63+ by default, more available with additional NARs
- MiNiFi C++ (v0.5.0)
 - Written from scratch
 - 33 processors by default
 - Bi-directional site-to-site & provenance data

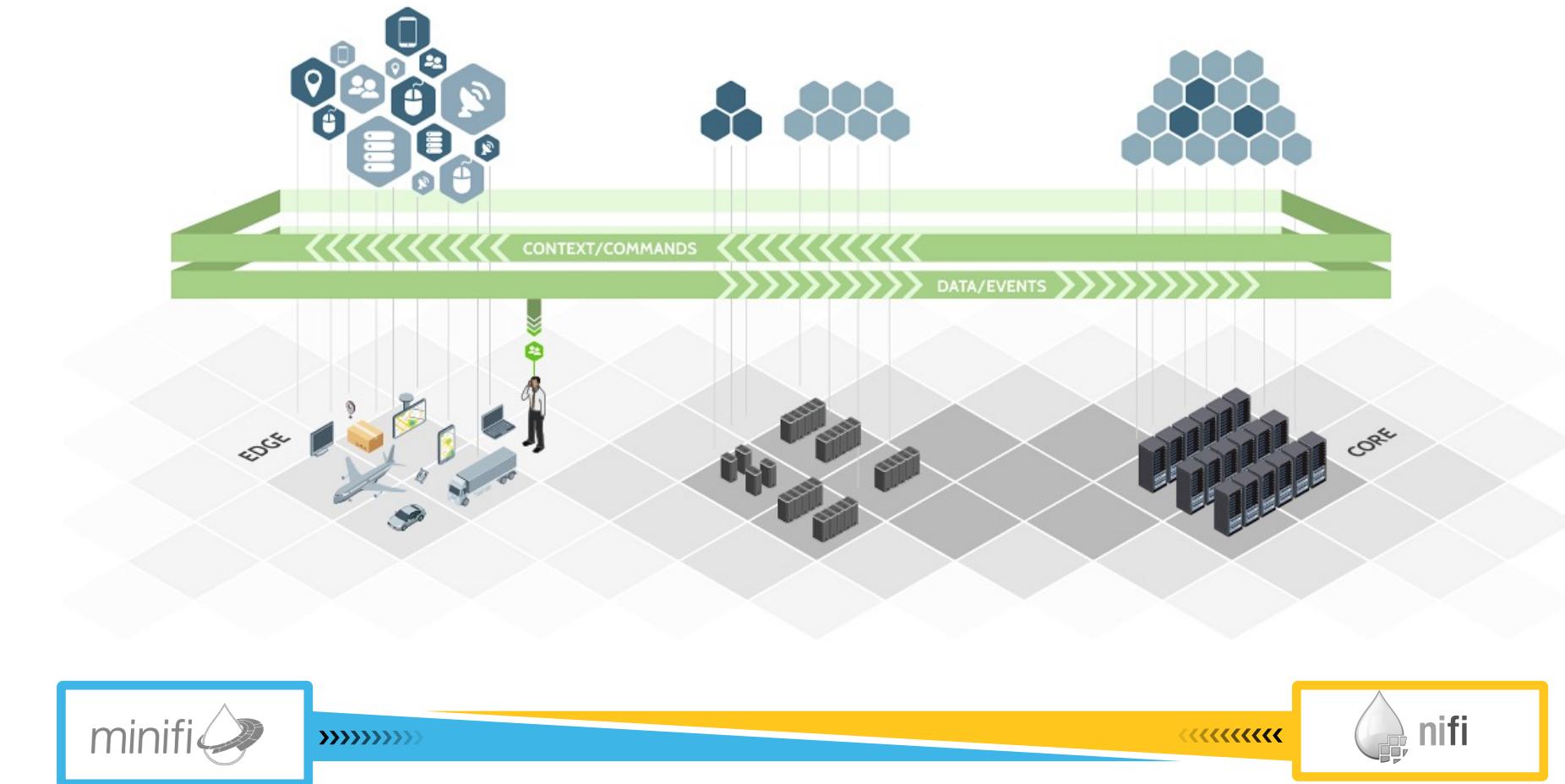


```
Security Properties:  
keystore: /tmp/ssl/localhost-ks.jks  
keystore type: JKS  
keystore password: localtest  
key password: localtest  
truststore: /tmp/ssl/localhost-ts.jks  
truststore type: JKS  
truststore password: localtest  
ssl protocol: TLS  
Sensitive Props:  
key:  
algorithm: PBWEWITHMD5AND256BITAES-CBC-OPENSSL  
provider: BC  
  
Processors:  
- name: TailAppLog  
  class: org.apache.nifi.processors.standard.TailFile  
  max concurrent tasks: 1  
  scheduling strategy: TIMER_DRIVEN  
  scheduling period: 10 sec  
  penalization period: 30 sec  
  yield period: 1 sec  
  run duration nanos: 0  
  auto-terminated relationships list:  
    Properties:  
      File to Tail: logs/minifi-app.log  
      Rolling Filename Pattern: minifi-app*  
      Initial Start Position: Beginning of File  
- name: SplitIntoSingleLines  
  class: org.apache.nifi.processors.standard.SplitText  
  max concurrent tasks: 1  
  scheduling strategy: TIMER_DRIVEN  
  scheduling period: 0 sec  
  penalization period: 30 sec  
  yield period: 1 sec  
  run duration nanos: 0  
  auto-terminated relationships list:  
    - failure  
    - original  
  Properties:  
    Line Split Count: 1  
    Header Line Count: 0  
    Remove Trailing Newlines: true
```



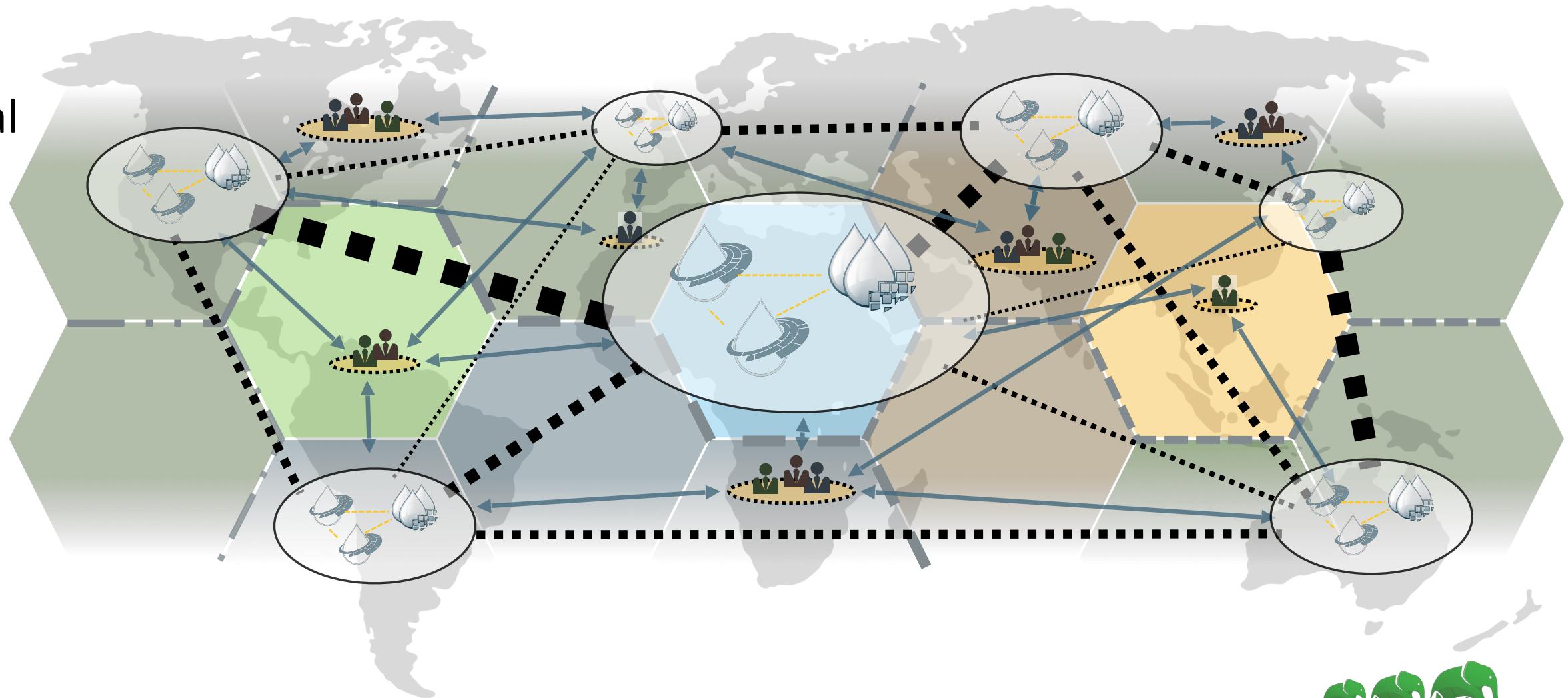
How Does MiNiFi Interact With NiFi?

- NiFi
 - Design flows
 - Aggregate data from many sources
 - Perform routing/analysis/SEP
- MiNiFi
 - Receive flows
 - Collect data
 - Send for processing



Let's Add Dimensionality

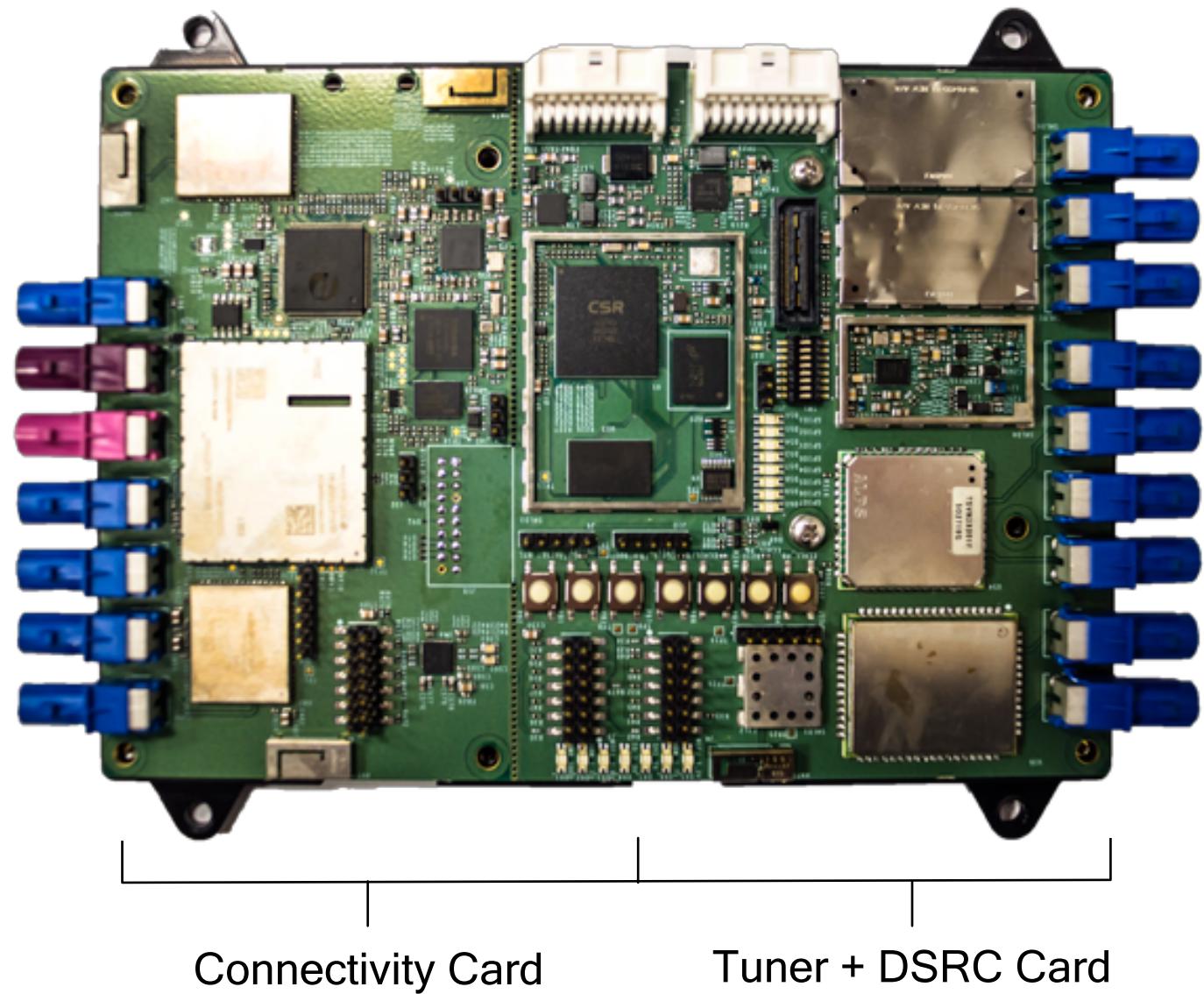
- We've been imagining EDGE to CORE as a bi-directional linear system
- Let's expand that to the real world



What does MiNiFi provide?

- Data tagging/provenance
- Governance from edge (geopolitical restrictions)
- Security (encryption, certificate-based authentication)
- Low latency (immediate reactions & decision-making)

Connected Car Reference Platform Box



MiNiFi Exfil

- Site-to-Site
 - NiFi protocol
 - Two implementations
 - Raw socket
 - HTTP(S)
- Secured with mutual authentication TLS
 - HTTP(S), (S)FTP, JMS, Syslog, File, Email, Process

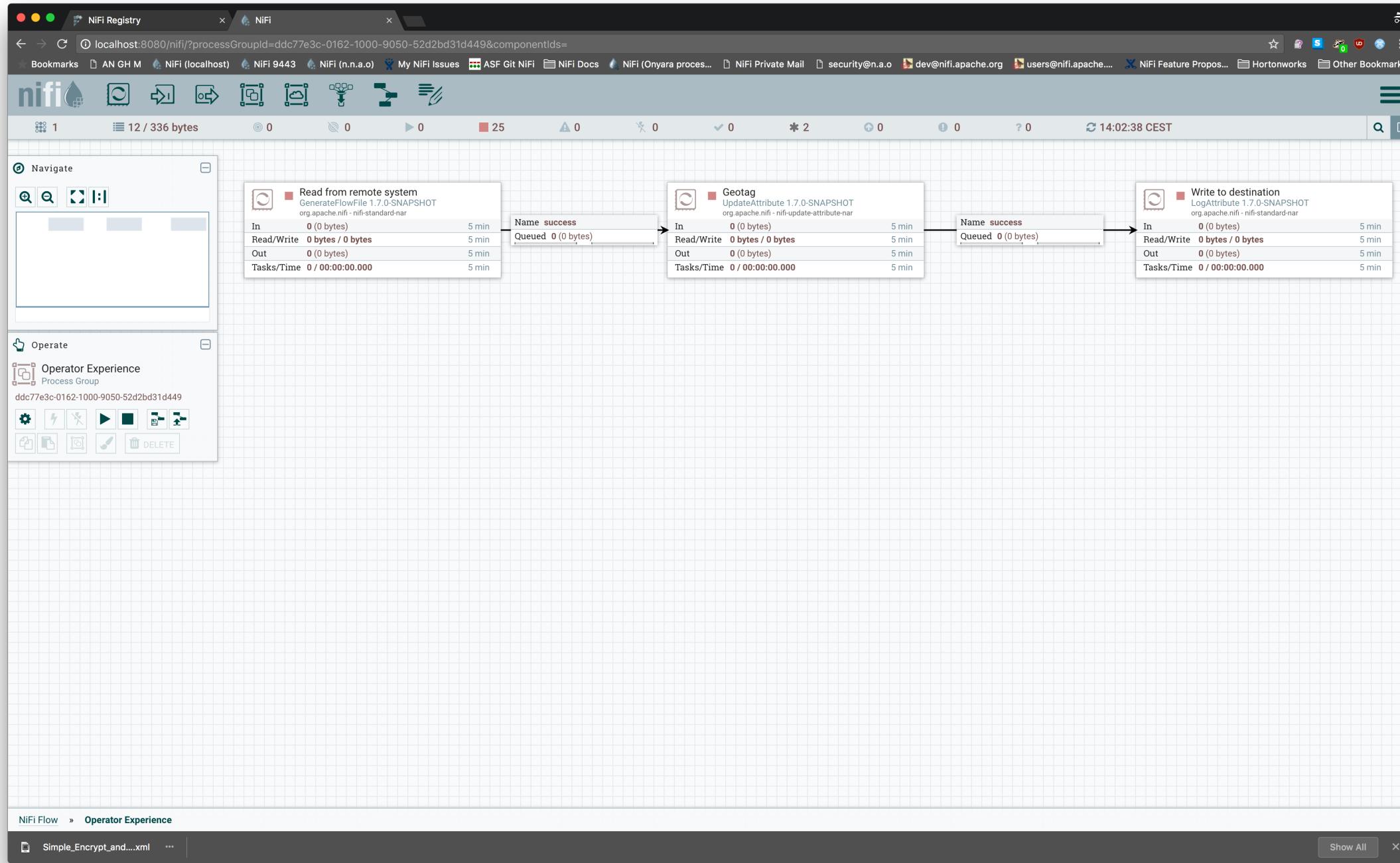
Apache NiFi Registry

Flow Development Lifecycle (FDLC)

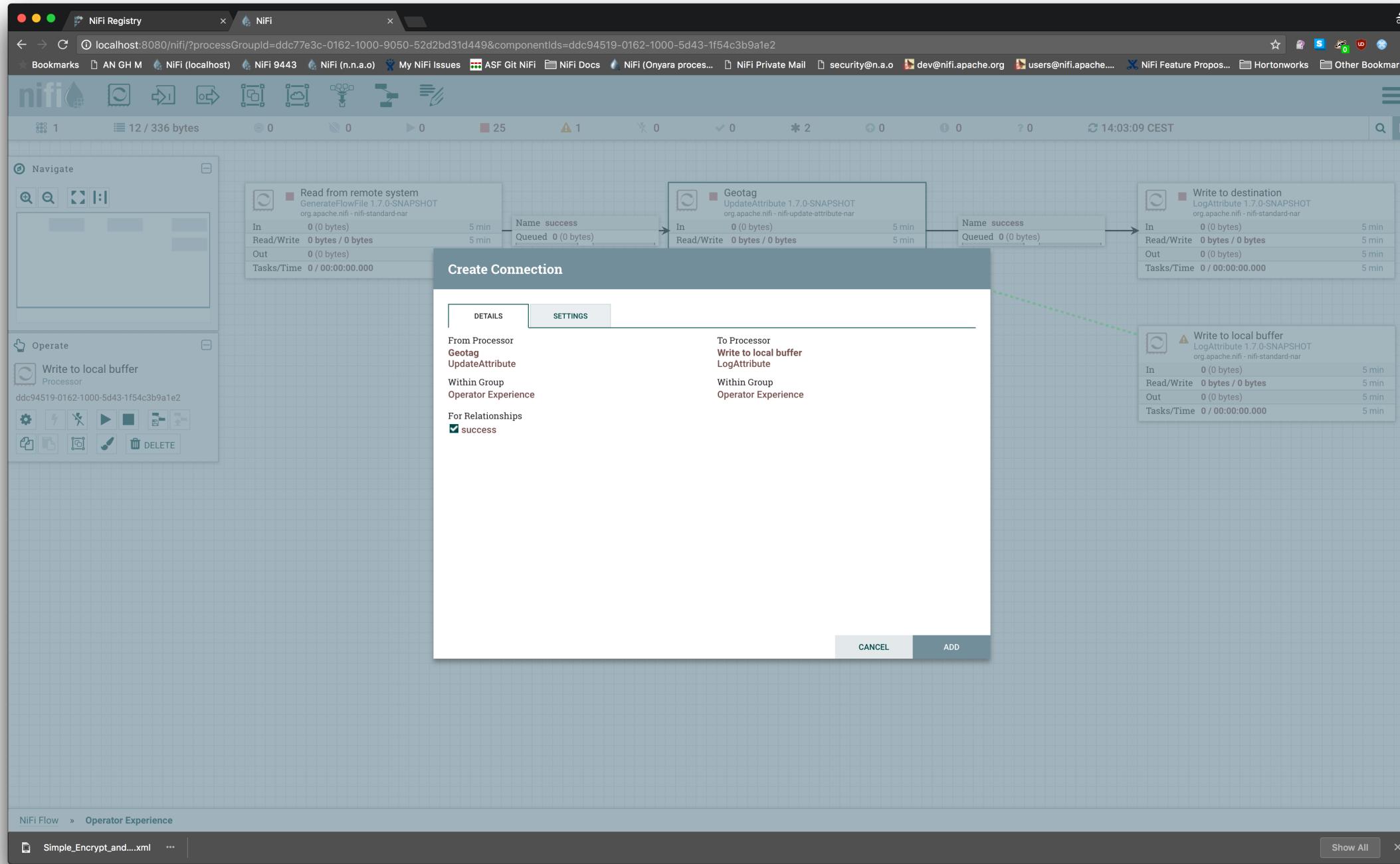
- Origins of NiFi
- Operator Experience
 - MC data, don't drop, mitigate temporarily
- Version Control
- Environment Promotion



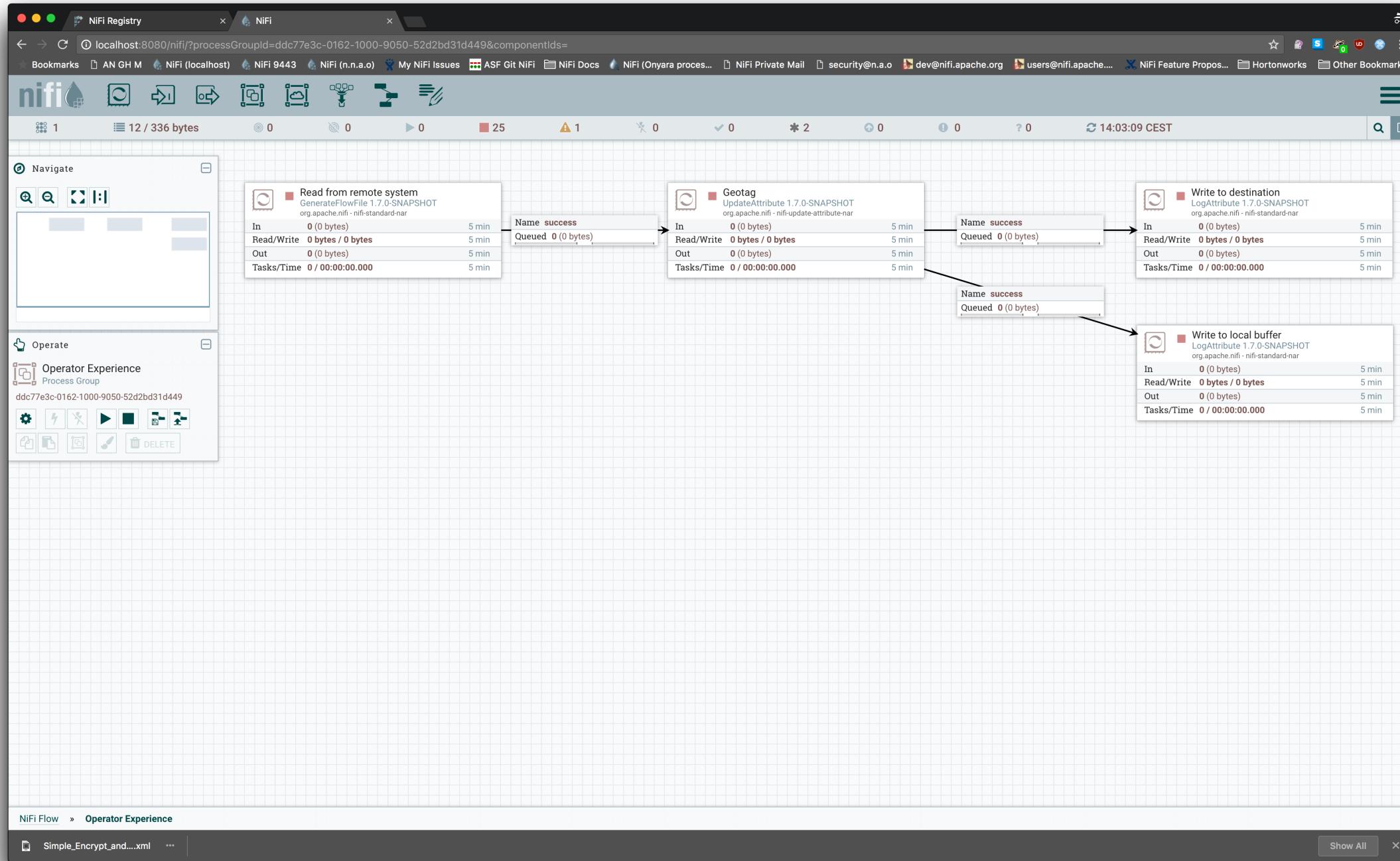
Operator Experience



Operator Experience



Operator Experience



Component Property History

- Shows previous values (user, time changed)
- Sensitive values are always encrypted at rest and never returned via the API

Configure Processor

SETTINGS SCHEDULING PROPERTIES COMMENTS

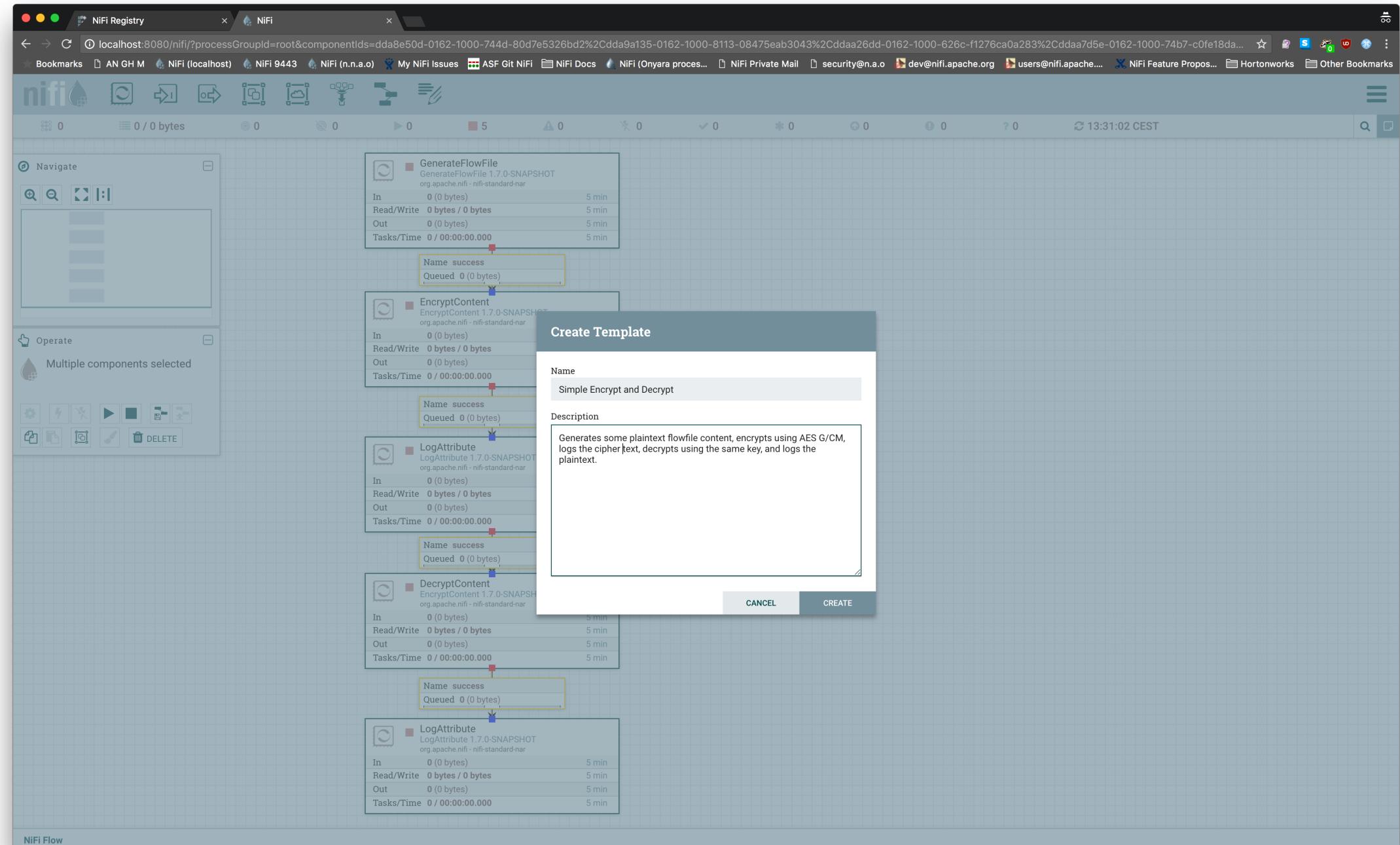
Required field +

| Property | Value |
|------------------------------------|--|
| Mode | Encrypt |
| Key Derivation Function | None |
| Encryption Algorithm | AES_GCM |
| Allow insecure cryptographic modes | Not Allowed |
| Password | No value set |
| Raw Key (hexadecimal) | Sensitive value set |
| Public Keyring File | In keyed encryption, this is the raw key, encoded in hexadecimal |
| Public Key User Id | Supports expression language: false |
| Private Keyring File | History: • ***** - 02/21/2017 20:30:46 PST (CN=alopresto, OU=Apache NiFi) |
| Private Keyring Passphrase | No value set No value set |

CANCEL APPLY

Exporting Flows

- XML templates
- Copying flow.xml.gz between systems



Challenges

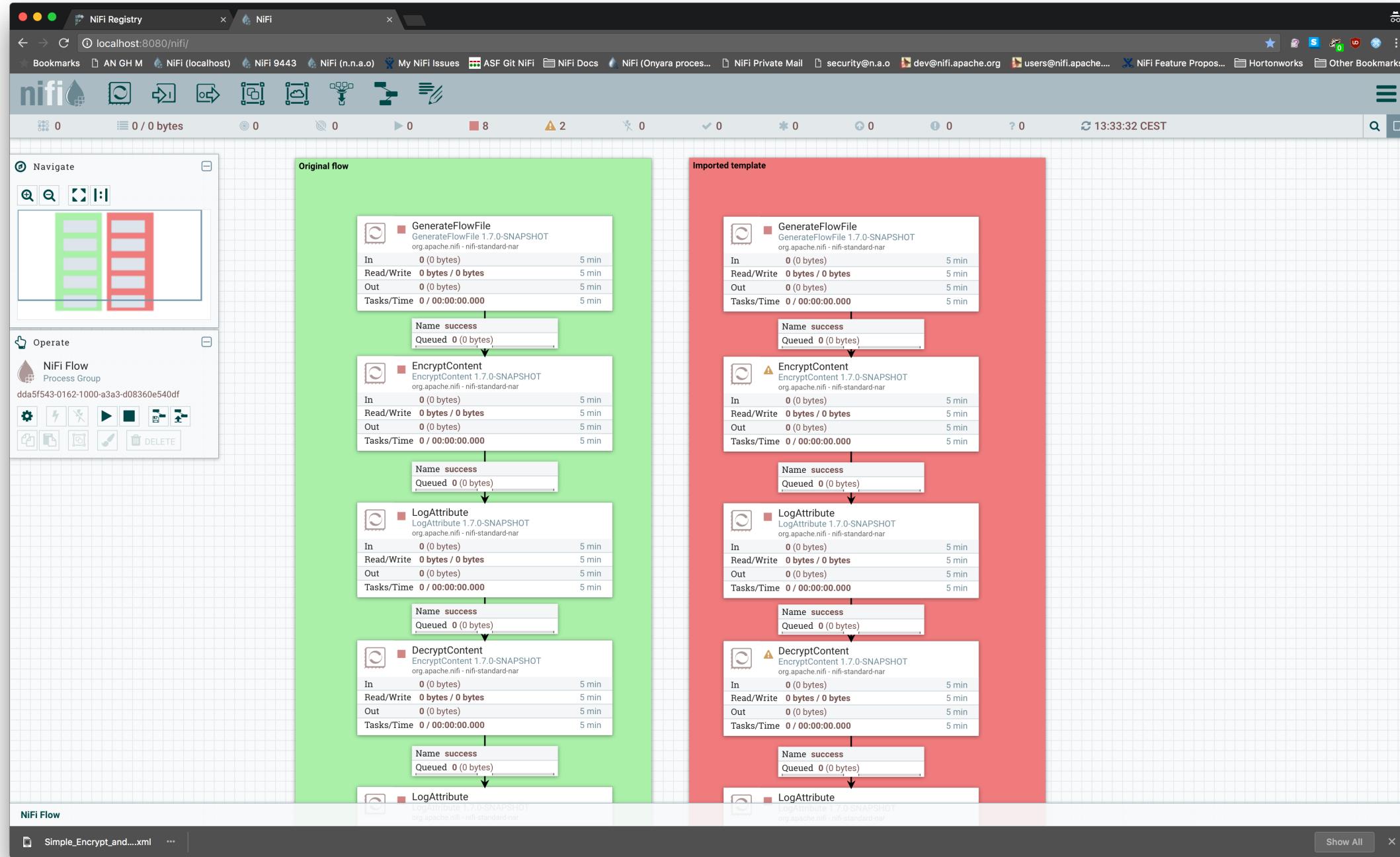
- Templates
 - Updates/replacement
 - Sensitive property replacement
- Flow.xml.gz migration
 - Key synchronization
- Environment promotion
 - Approval processes
 - Verifiability

Template Replacement

- Export a new version of template
- Transfer (somehow)
- Verify?
- Import onto canvas side-by-side existing flow
- Stop processors
- Empty queues
- Reconnect queues
- Start
- Pray?

```
-rw-r--r-- 1 alopresto staff 381B Apr 19 15:02 flow_20170626-171416_label_description.xml  
-rw-r--r-- 1 alopresto staff 1.4K Apr 19 15:02 flow_20170628-111620_xxe_transform_xml.xml  
-rw-r--r-- 1 alopresto staff 1.4K Apr 19 15:02 flow_20170628-155452_xxe_with_patch.xml  
-rw-r--r-- 1 alopresto staff 1.4K Apr 19 15:02 flow_20170629-103436_xxe_password_leak.xml  
-rw-r--r-- 1 alopresto staff 1.9K Apr 19 15:02 flow_20170717-192339_trusted_hostname.xml  
-rw-r--r-- 1 alopresto staff 2.7K Apr 19 15:02 flow_20170718-101751_evaluate_json_path.xml  
-rw-r--r-- 1 alopresto staff 5.2K Apr 19 15:02 flow_20170721-175745_site_to_site_secure.xml  
-rw-r--r-- 1 alopresto staff 5.2K Apr 19 15:02 flow_20170724-161539.xml  
-rw-r--r-- 1 alopresto staff 5.7K Apr 19 15:02 flow_20170724-181056_listen_http_sslv3.xml  
-rw-r--r-- 1 alopresto staff 2.3K Apr 19 15:02 flow_20170726-155905_encrypted_spk.xml  
-rw-r--r-- 1 alopresto staff 5.2K Apr 19 15:02 flow_20170802-172619_simple_gen_and_log.xml  
-rw-r--r-- 1 alopresto staff 5.2K Apr 19 15:02 flow_20170804-164011_simple_gen_and_log.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20170807-154118_groovy_json_uppercase.xml  
-rw-r--r-- 1 alopresto staff 1.8K Apr 19 15:02 flow_20170807-220536_put_mongo_record.xml  
-rw-r--r-- 1 alopresto staff 6.1K Apr 19 15:02 flow_20170809-105304_date_difference.xml  
-rw-r--r-- 1 alopresto staff 6.4K Apr 19 15:02 flow_20170809-174400_json_record_path.xml  
-rw-r--r-- 1 alopresto staff 1.3K Apr 19 15:02 flow_20170809-210516_validate_xml.xml  
-rw-r--r-- 1 alopresto staff 1.4K Apr 19 15:02 flow_20170815-102522_json_to_sql.xml  
-rw-r--r-- 1 alopresto staff 1.5K Apr 19 15:02 flow_20170822-104407_json_formatter.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20170823-123729_evaluatexpath.xml  
-rw-r--r-- 1 alopresto staff 2.5K Apr 19 15:02 flow_20170825-115758_encrypt_and_decrypt.xml  
-rw-r--r-- 1 alopresto staff 2.0K Apr 19 15:02 flow_20170825-121351_encrypt_decrypt_clean.xml  
-rw-r--r-- 1 alopresto staff 2.0K Apr 19 15:02 flow_20170825-121656_encrypt_decrypt_aligned.xml  
-rw-r--r-- 1 alopresto staff 2.3K Apr 19 15:02 flow_20170825-190225_listen_http_and_invoke_http_with_tls1_2.xml  
-rw-r--r-- 1 alopresto staff 5.0K Apr 19 15:02 flow_20170831-153220_psaltis_meetup_template_local.xml  
-rw-r--r-- 1 alopresto staff 4.0K Apr 19 15:02 flow_20170905-175343_xxe_template.xml  
-rw-r--r-- 1 alopresto staff 4.3K Apr 19 15:02 flow_20170905-185907_funnel_bus.xml  
-rw-r--r-- 1 alopresto staff 8.1K Apr 19 15:02 flow_20170906-111801_chained_get_html_element_NIFI-4356.xml  
-rw-r--r-- 1 alopresto staff 2.5K Apr 19 15:02 flow_20170909-121848_sydney_demo.xml  
-rw-r--r-- 1 alopresto staff 3.2K Apr 19 15:02 flow_20170919-202215_meetup_sydney.xml  
-rw-r--r-- 1 alopresto staff 3.8K Apr 19 15:02 flow_20170920-190328_crash_course.xml  
-rw-r--r-- 1 alopresto staff 2.9K Apr 19 15:02 flow_20170926-180548_handle_http_with_tls1_2.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20170929-172333_encrypt_decrypt_1_4_0_verification.xml  
-rw-r--r-- 1 alopresto staff 6.3K Apr 19 15:02 flow_20171003-192928_update_record_and_replace_text.xml  
-rw-r--r-- 1 alopresto staff 8.9K Apr 19 15:02 flow_20171004-180529_web_server_and_json.xml  
-rw-r--r-- 1 alopresto staff 3.8K Apr 19 15:02 flow_20171103-183334_groovy_json_to_sql_srsw.xml  
-rw-r--r-- 1 alopresto staff 1.5K Apr 19 15:02 flow_20171127-190651_local_site_to_site.xml  
-rw-r--r-- 1 alopresto staff 2.1K Apr 19 15:02 flow_20180101-145059_registry_rc_verification.xml  
-rw-r--r-- 1 alopresto staff 1.4K Apr 19 15:02 flow_20180104-104211_test_count_text.xml  
-rw-r--r-- 1 alopresto staff 2.0K Apr 19 15:02 flow_20180104-112138_test_count_text_remote.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20180202-100218_consume_jms_test.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20180202-105514.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20180202-131629_consume_jms_10_threads.xml  
-rw-r--r-- 1 alopresto staff 1.7K Apr 19 15:02 flow_20180202-160620_consume_jms_multithread.xml  
-rw-r--r-- 1 alopresto staff 1.6K Apr 19 15:02 flow_20180208-200731_simple_demo_route_on_time_parity.xml  
-rw-r--r-- 1 alopresto staff 5.2K Apr 19 15:02 flow_20180214-182110.jms_deserialization_test.xml  
-rw-r--r-- 1 alopresto staff 2.6K Apr 19 15:02 flow_20180216-205336_attribute_evaluator.xml  
-rw-r--r-- 1 alopresto staff 5.7K Apr 19 15:02 flow_20180228-130457_attribute_soq.xml  
-rw-r--r-- 1 alopresto staff 8.3K Apr 19 15:02 flow_20180307-171740_NIFI-4928.xml  
-rw-r--r-- 1 alopresto staff 1.5K Apr 19 15:02 flow_20180313-171310_python_process_attribute.xml  
-rw-r--r-- 1 alopresto staff 2.4K Apr 19 15:02 flow_20180315-113634_NIFI-4246-0Auth.xml  
-rw-r--r-- 1 alopresto staff 1.9K Apr 19 15:02 flow_20180328-115056.xml  
-rw-r--r-- 1 alopresto staff 2.5K Apr 19 15:02 flow_20180405-111419_prioritizers.xml
```

Template Replacement



NiFi Registry for Dataflows

Introducing Apache NiFi Registry 0.3.0

- Previously, flows were exported via XML templates
 - Didn't contain sensitive values
 - Couldn't be updated in-place
 - No tracking system
- NiFi Registry brings asset management as first-class citizen to NiFi
- Flows can be versioned

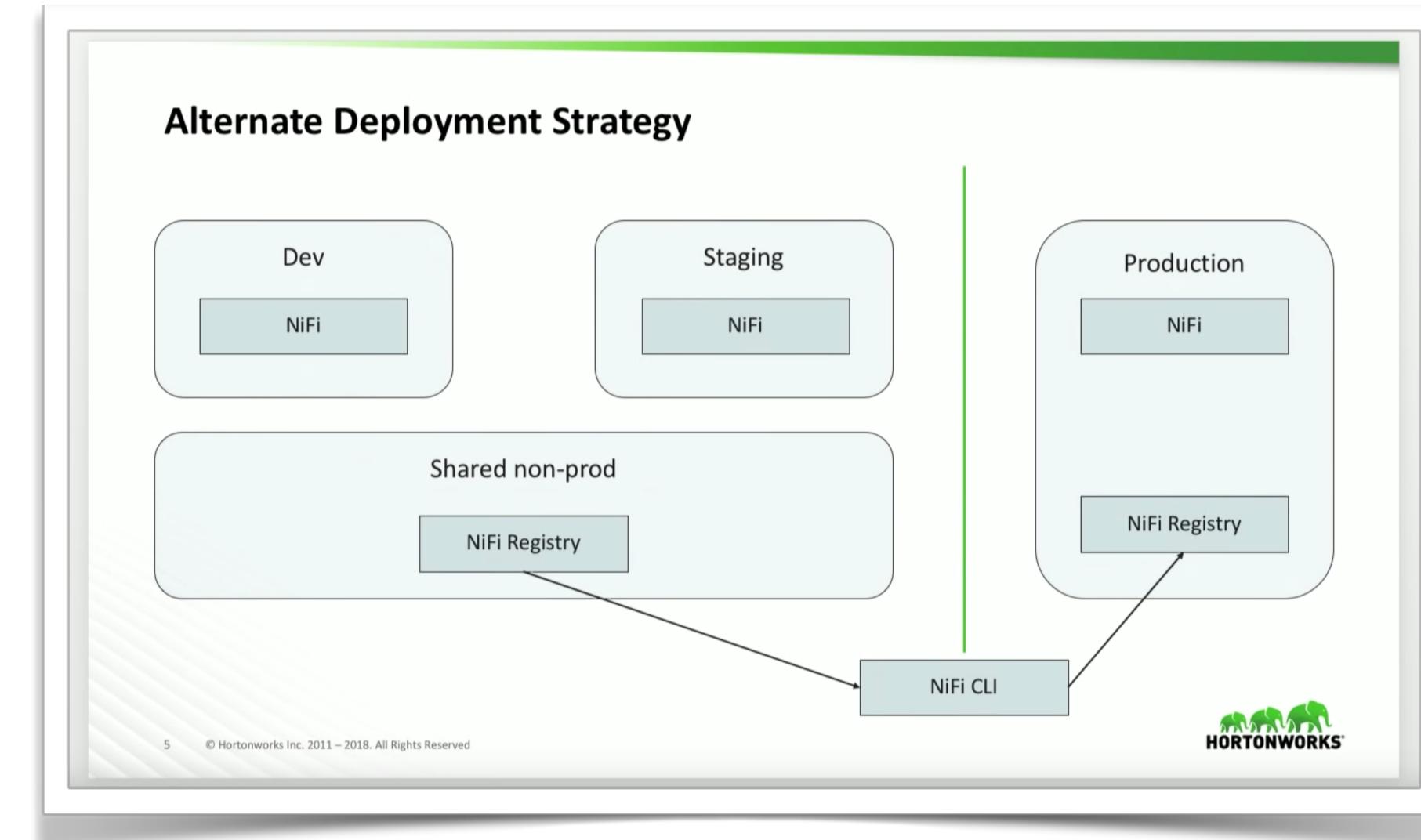
The screenshot shows the Apache NiFi Registry interface. At the top, there's a navigation bar with the title "NiFi Registry / All" and user information "registry_user LOGOUT". Below the navigation bar, there's a search bar and a sorting dropdown set to "Sort by: Name (a - z)". The main content area displays two flows: "Flow 1 - Bucket 1" (1 version) and "Flow 2 - Bucket 2" (2 versions). For "Flow 2 - Bucket 2", a detailed view is shown, including its description "Description 2" and a "CHANGE LOG" section. The change log lists three entries:

- Version 2 - 40 minutes ago by registry_user (highlighted)
- Add processors Dec-26-2017 at 11:23 PM
- Version 1 - 41 minutes ago by registry_user

A "DESCRIPTION" section is also visible, containing the text "Description 2". On the right side of the screen, there's a "ACTIONS" dropdown menu.

Flows can be promoted between environments

- Connect multiple NiFi instances to a NiFi Registry instance
- Communicate between multiple NiFi Registry instances
 - via multiple Registry Clients
 - via *NiFi CLI*



Extensibility

- Git-backed persistence
 - Share flows via GitHub, etc.
- Commit hooks
 - Register a hook & action
 - “When a new version of the flow is committed to QA Registry, email the QA team and post in the QA Deploy Slack channel”
- Pluggable DB implementations

Event Hooks

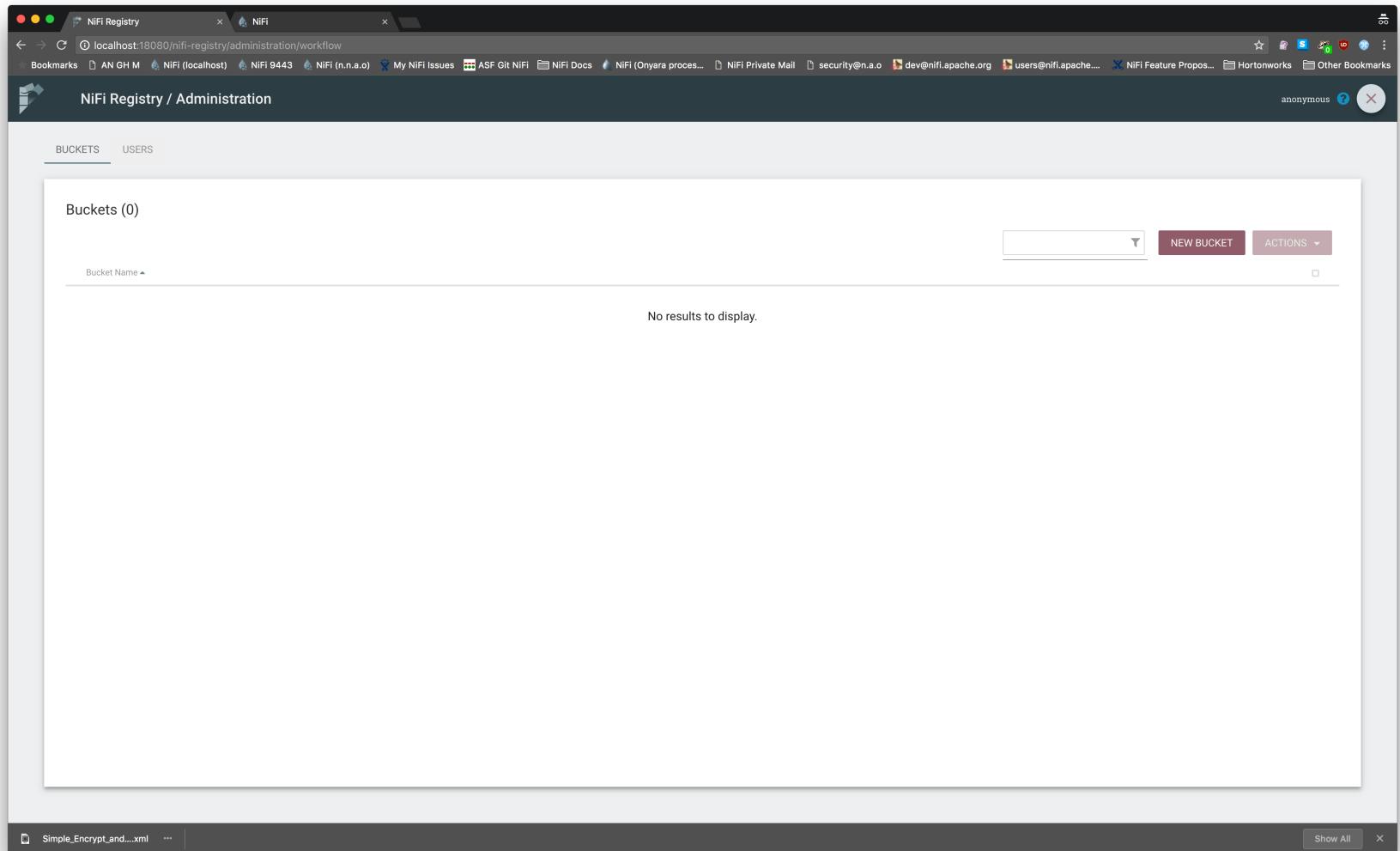
Event hooks are an integration point that allows for custom code to be triggered when NiFi Registry application events occur.

| Event Name | Description |
|---------------------|---|
| CREATE_BUCKET | A new registry bucket is created. |
| CREATE_FLOW | A new flow is created in a specified bucket. Only triggered on first time creation of a flow with a given name. |
| CREATE_FLOW_VERSION | A new version for a flow has been saved in the registry. |
| UPDATE_BUCKET | A bucket has been updated. |
| UPDATE_FLOW | A flow that exists in a bucket has been updated. |
| DELETE_BUCKET | An existing bucket in the registry is deleted. |
| DELETE_FLOW | An existing flow in the registry is deleted. |
| REGISTRY_START | Invoked once the NiFi Registry application has been successfully started. This is only invoked after a complete and successful start. |

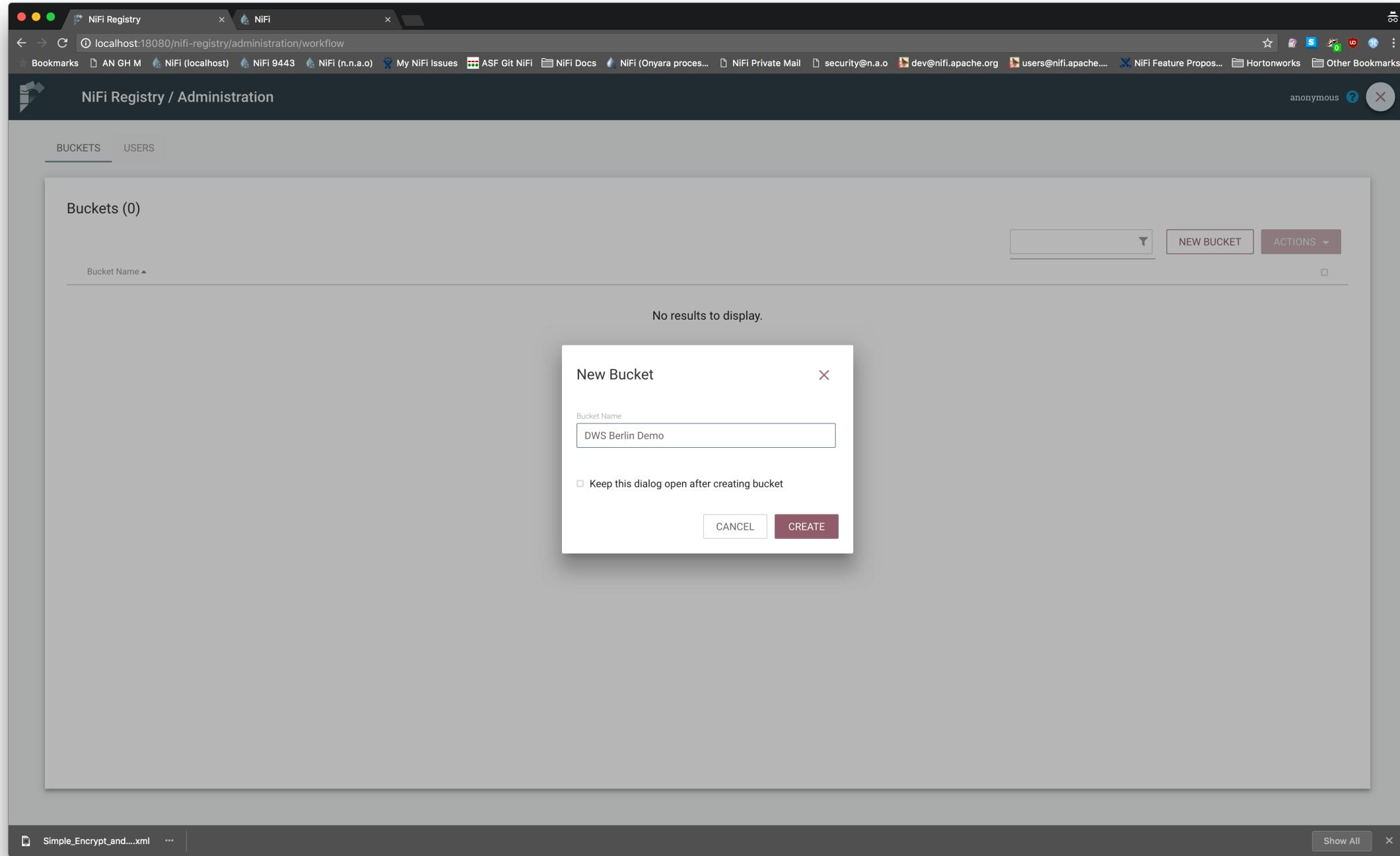
Demo

Create Registry

- Install nifi-registry
- \$ mvn clean install
- \$./bin/nifi-registry.sh start
- Browse to <http://localhost:18080>



Create Bucket



Connect to NiFi

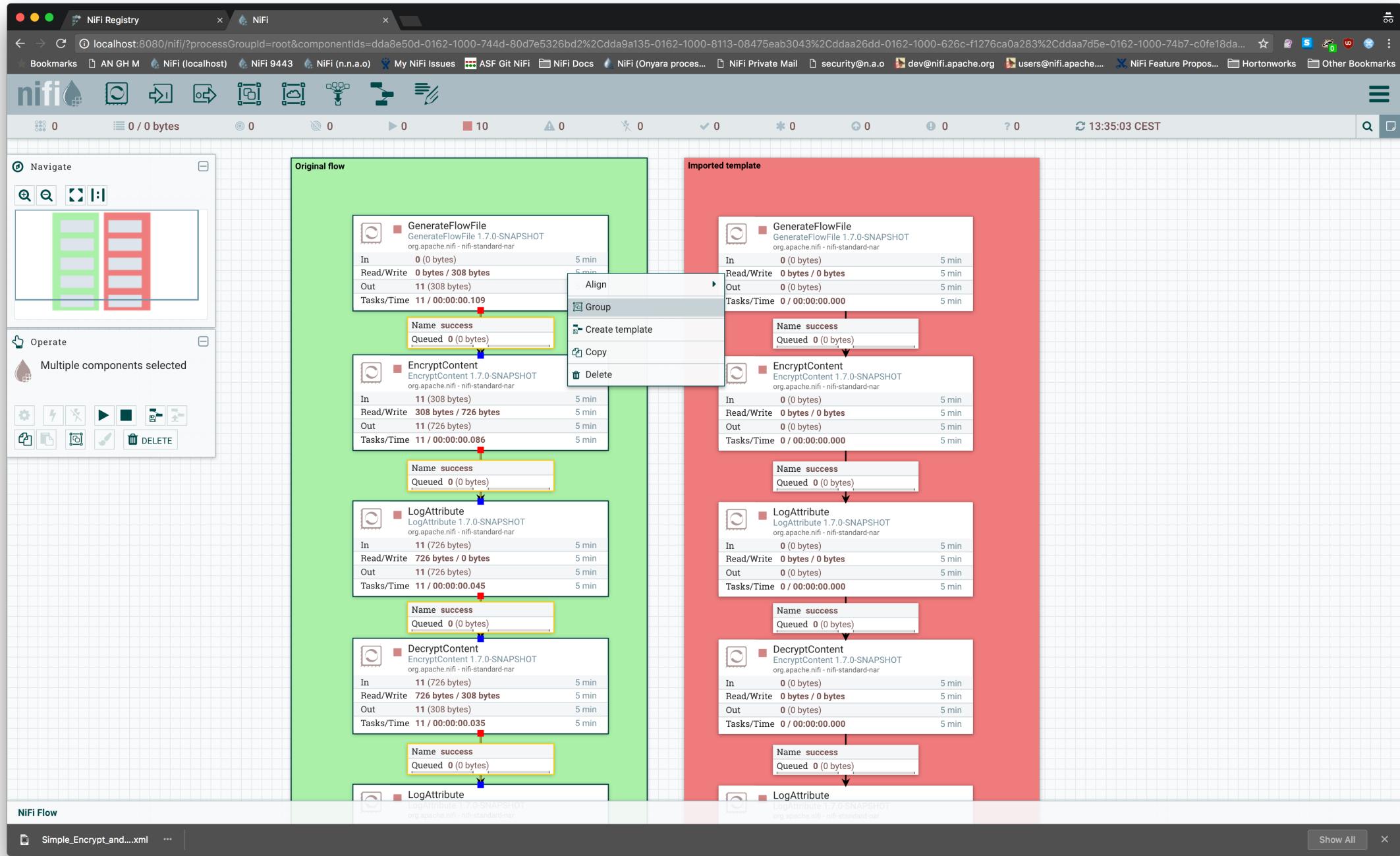
The screenshot shows the NiFi Registry interface in a web browser. The title bar indicates the window is titled 'NiFi Registry' and 'NiFi'. The address bar shows the URL 'localhost:8080/nifi/'. The main content area is titled 'NiFi Settings' and has tabs for 'GENERAL', 'REPORTING TASK CONTROLLER SERVICES', 'REPORTING TASKS', and 'REGISTRY CLIENTS'. The 'REGISTRY CLIENTS' tab is selected. A modal dialog box titled 'Add Registry Client' is open in the center. It contains fields for 'Name' (set to 'DWS Berlin Registry'), 'URL' (set to 'http://localhost:18080'), and 'Description' (set to 'A NiFi Registry instance used during the DWS Berlin FDLC demo.'). At the bottom of the dialog are 'CANCEL' and 'ADD' buttons. In the background, there is a table with columns 'Name', 'Location', and 'Description'. At the bottom of the page, there is a footer with links like 'NIIFI FLOW', 'Simple_Encrypt_and...xml', 'Show All', and a close button.

Connect to NiFi

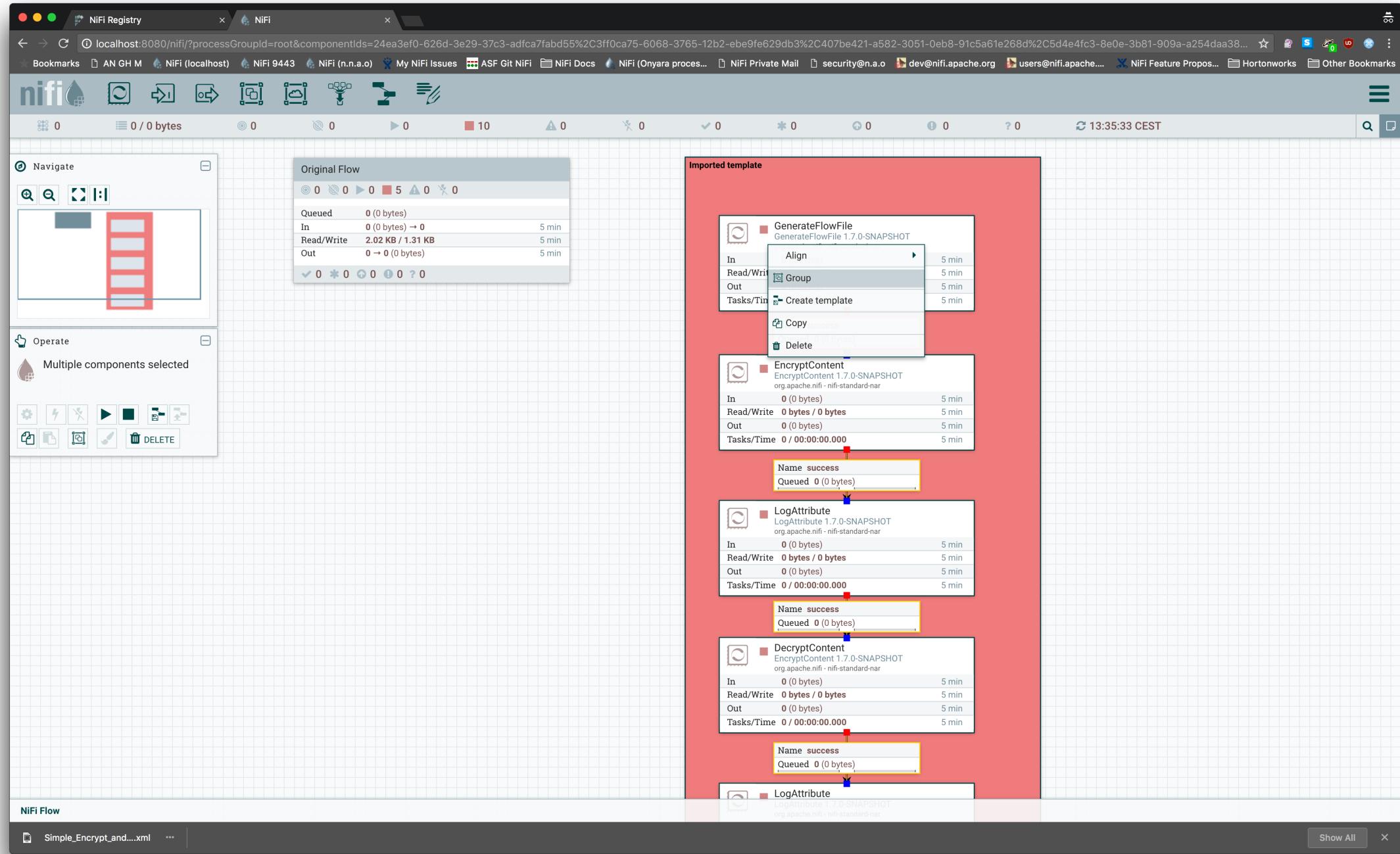
The screenshot shows the NiFi Registry interface in a web browser. The title bar indicates the window is titled "NiFi Registry" and the address bar shows "localhost:8080/nifi/". The main content area is titled "NiFi Settings" and displays a table of "REGISTRY CLIENTS". The table has three columns: "Name", "Location", and "Description". There is one entry: "DWS Berlin Registry" with "http://localhost:18080" in the Location column and a descriptive note in the Description column. The bottom of the screen shows a "NIFI FLOW" tab with a file named "Simple_Encrypt_and...xml". A status message at the bottom left says "Last updated: 13:36:00 CEST".

| Name | Location | Description |
|---------------------|------------------------|--|
| DWS Berlin Registry | http://localhost:18080 | A NiFi Registry instance used during the DWS Berlin FDLC demo. |

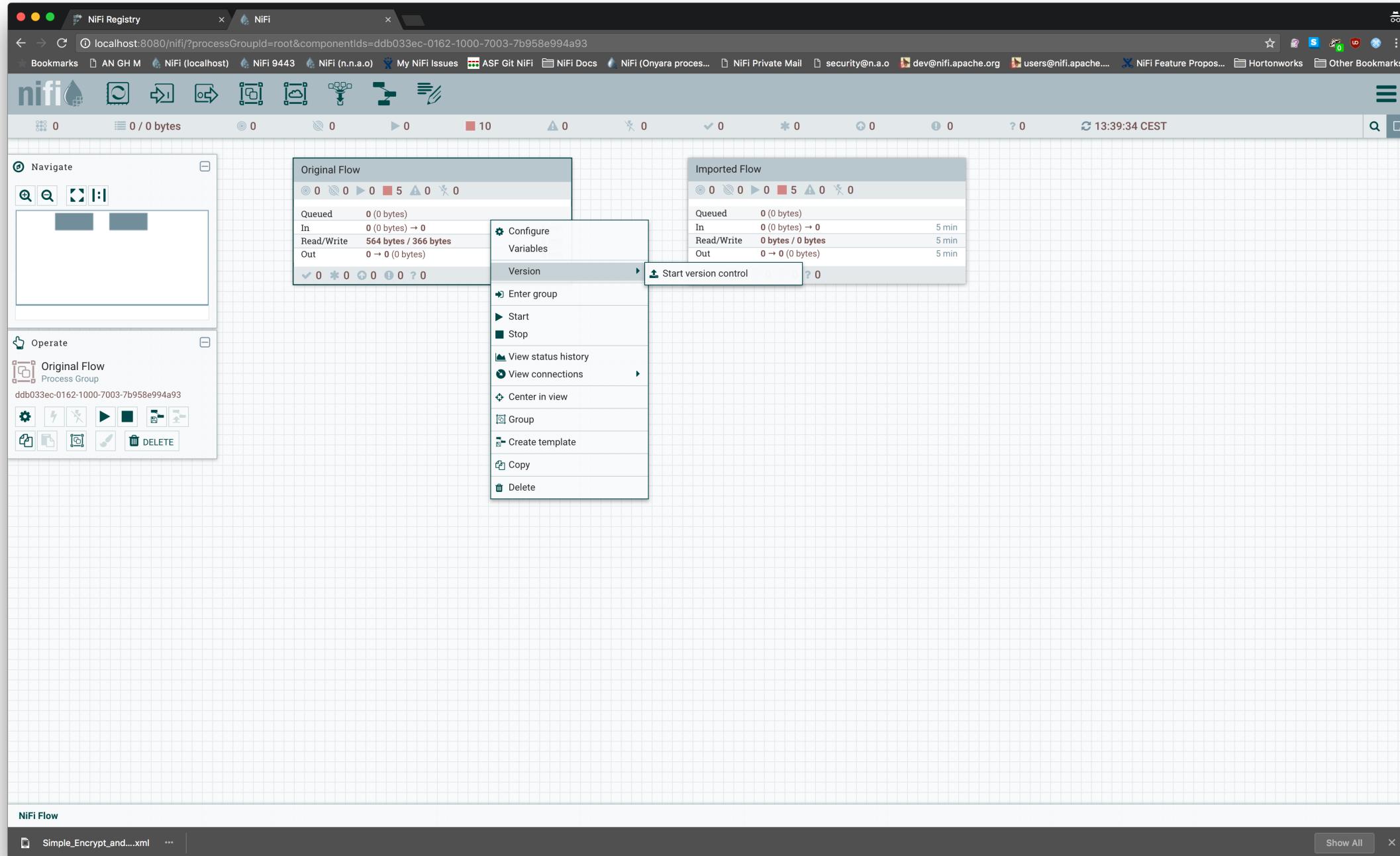
Create Process Group



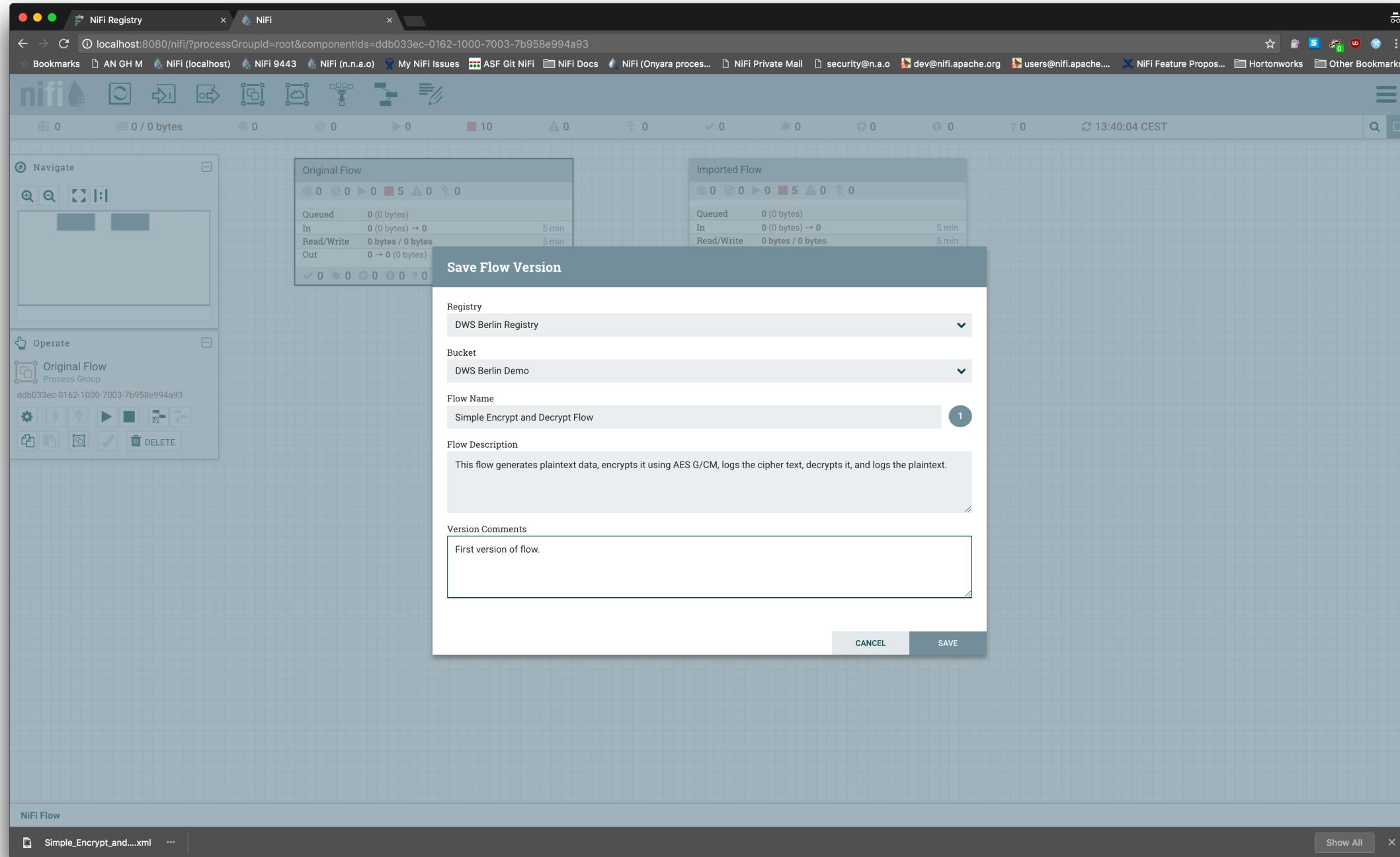
Create Process Group



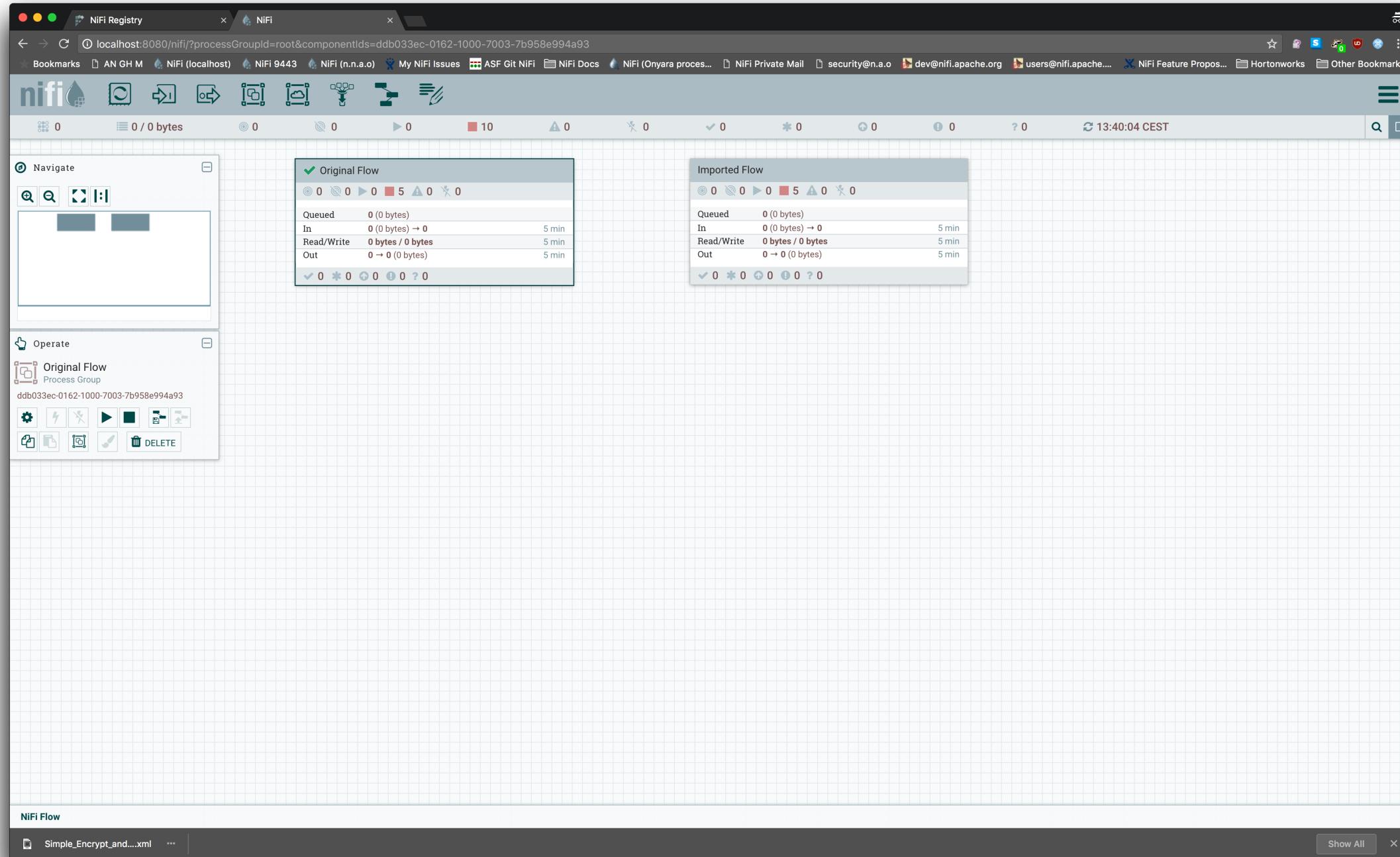
Commit Version



Commit Version



Commit Version



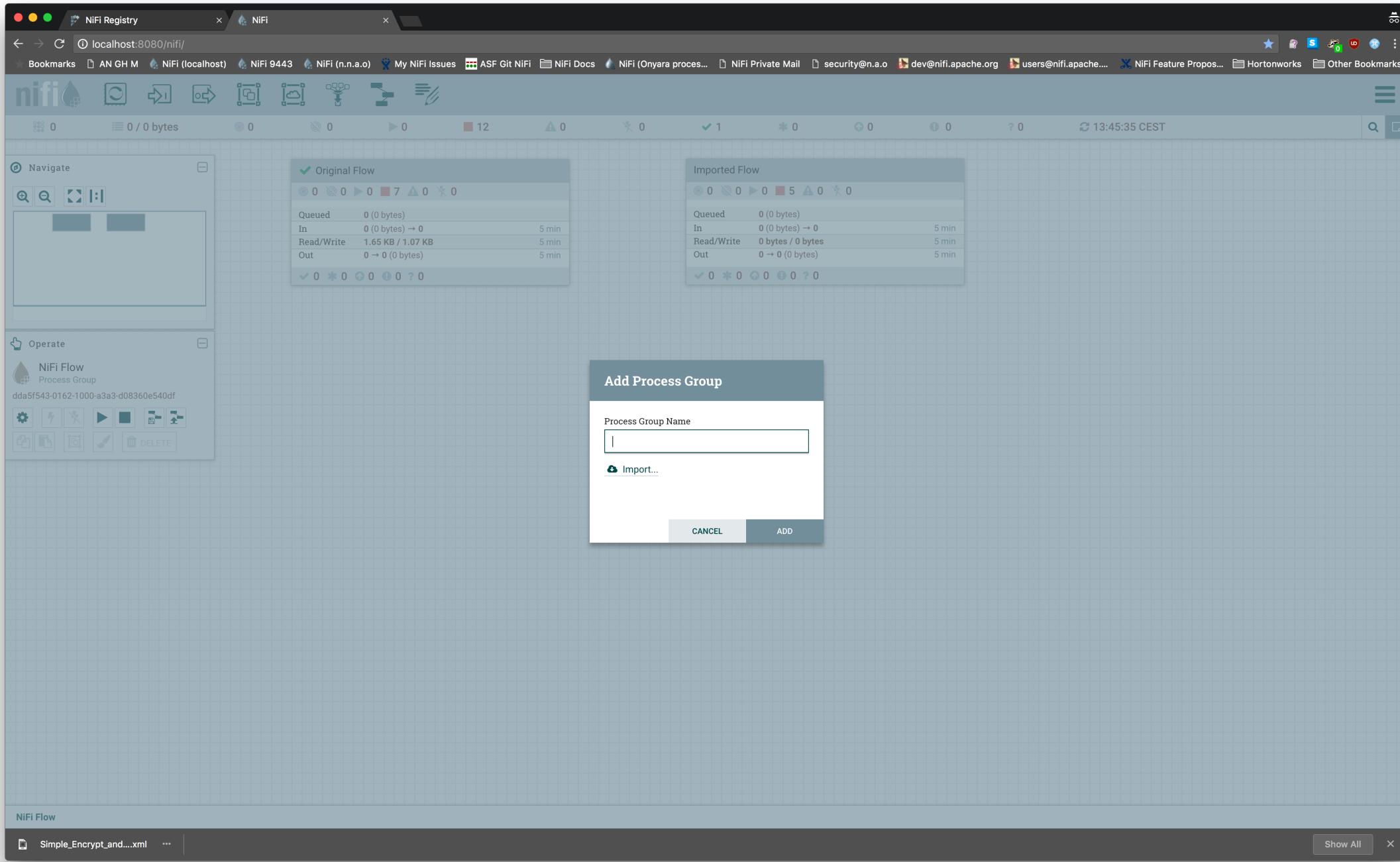
View flow in Registry

The screenshot shows a web browser window for the NiFi Registry at localhost:18080/nifi-registry/explorer/grid-list. The title bar says "NIFI Registry" and "NIFI". The address bar shows the URL. The page header includes a search bar, a sort dropdown set to "Name (a-z)", and a user status "anonymous". The main content area displays a flow named "Simple Encrypt And Decrypt Flow - DWS Berlin Demo". To the right of the flow name is a "VERSIONS" section showing "2". Below this, there are sections for "BUCKET IDENTIFIER" (d7df1fc1-d14b-4bb5-a8a6-ea57a60c5c98), "FLOW IDENTIFIER" (b288f8eb-8adf-455a-b887-785239e97d33), and "DESCRIPTION" (This flow generates plaintext data, encrypts it using AES GCM, logs the cipher text, decrypts it, and logs the plaintext.). On the right side of the flow details, there is a "CHANGE LOG" section with two entries:

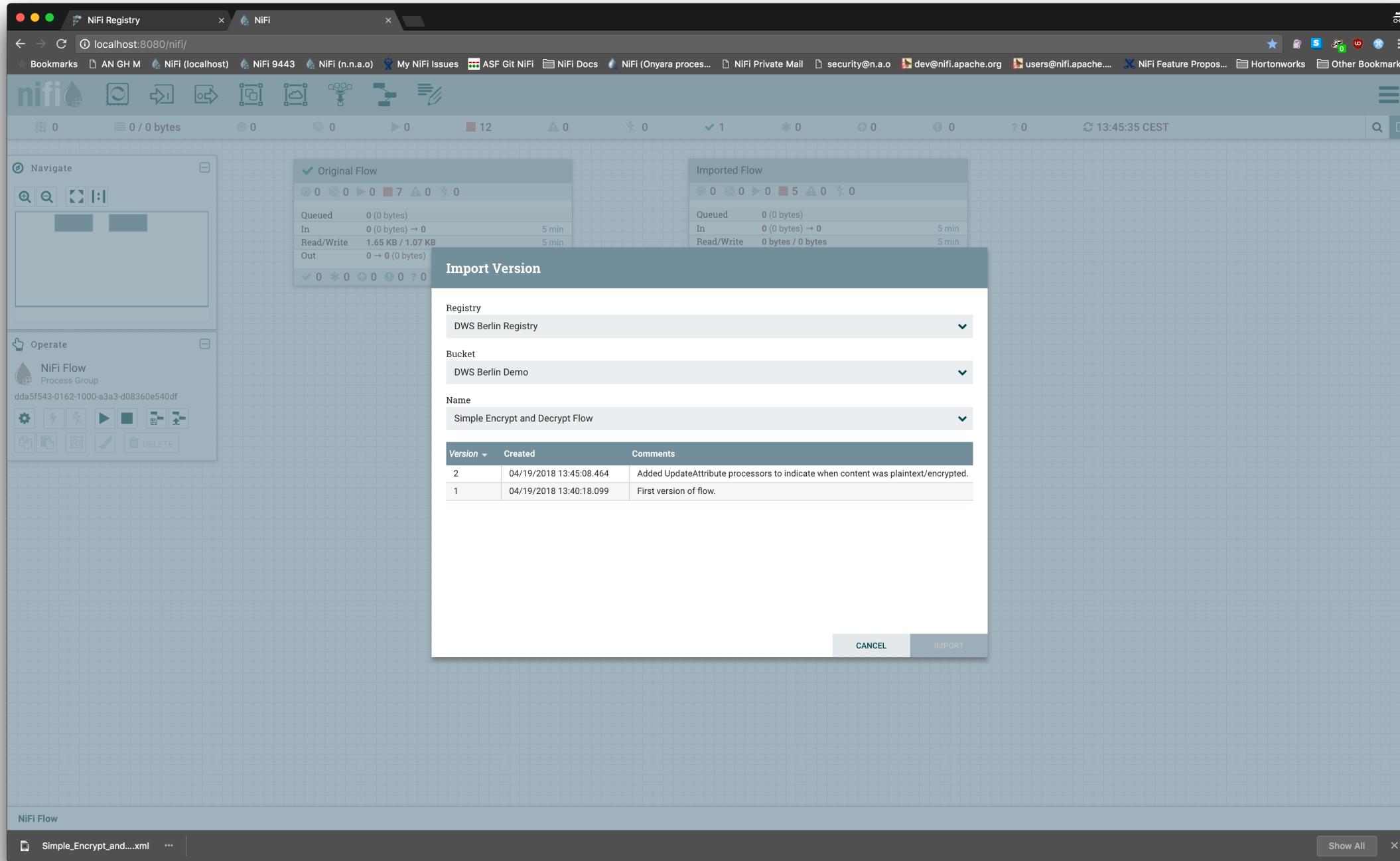
- Version 2 - a few seconds ago by anonymous: Added UpdateAttribute processors to indicate when content was plaintext/encrypted. (Apr-19-2018 at 1:45 PM)
- Version 1 - 5 minutes ago by anonymous: (no details provided)

At the bottom left of the main content area, there is a file icon followed by "Simple_Encrypt_and....xml" and an ellipsis "...". At the bottom right, there are "Show All" and "X" buttons.

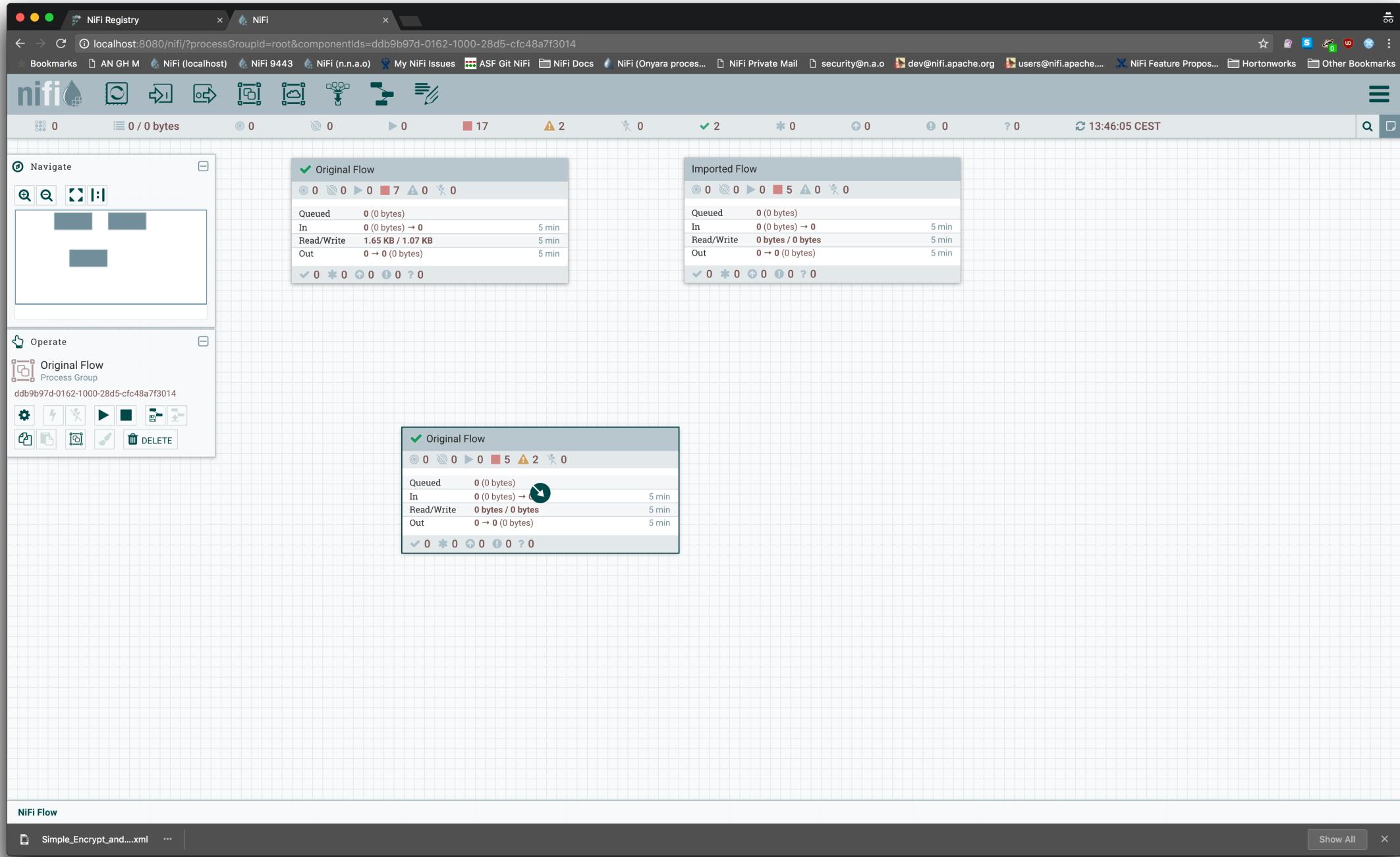
Import new instance into NiFi



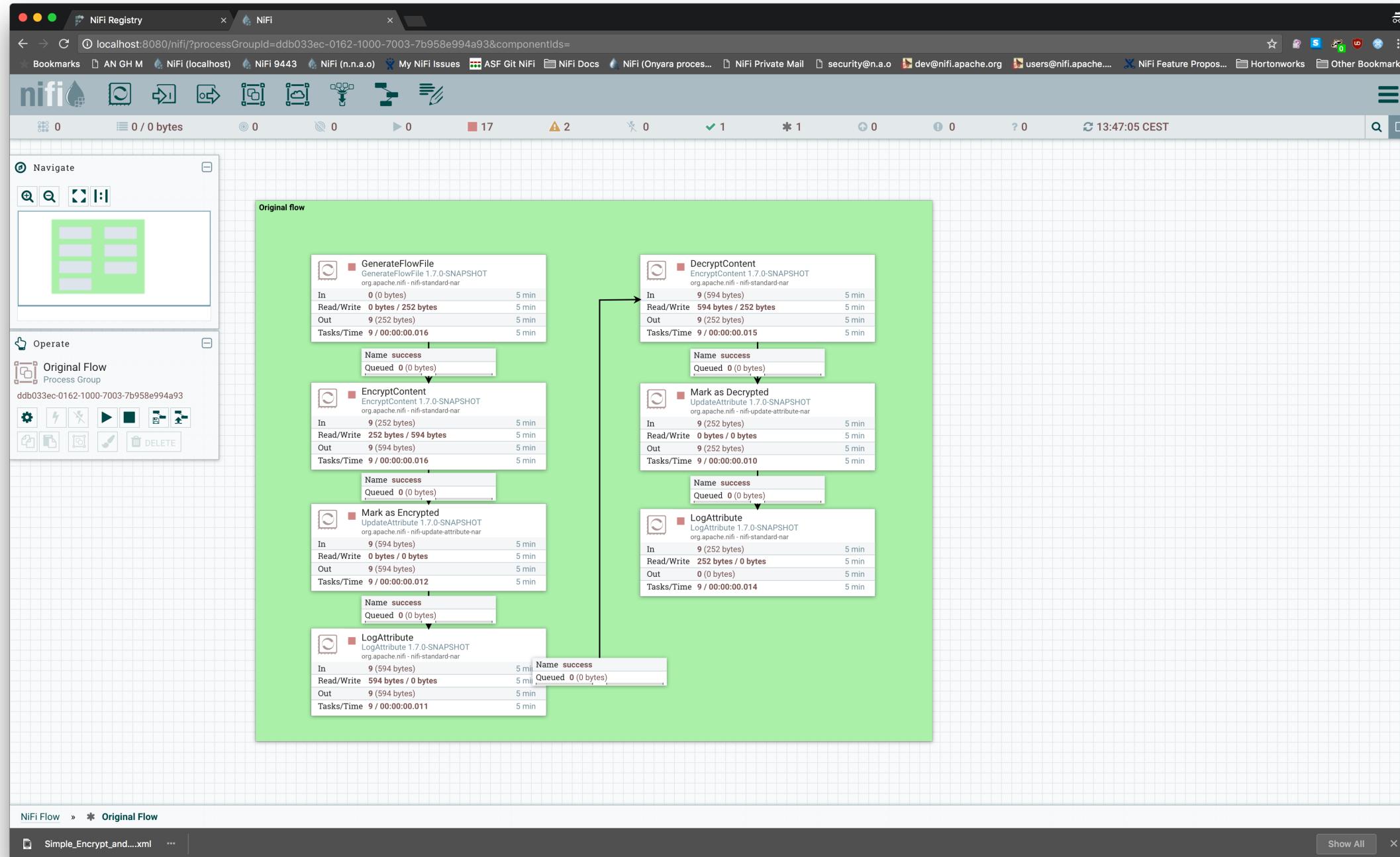
Import new instance into NiFi



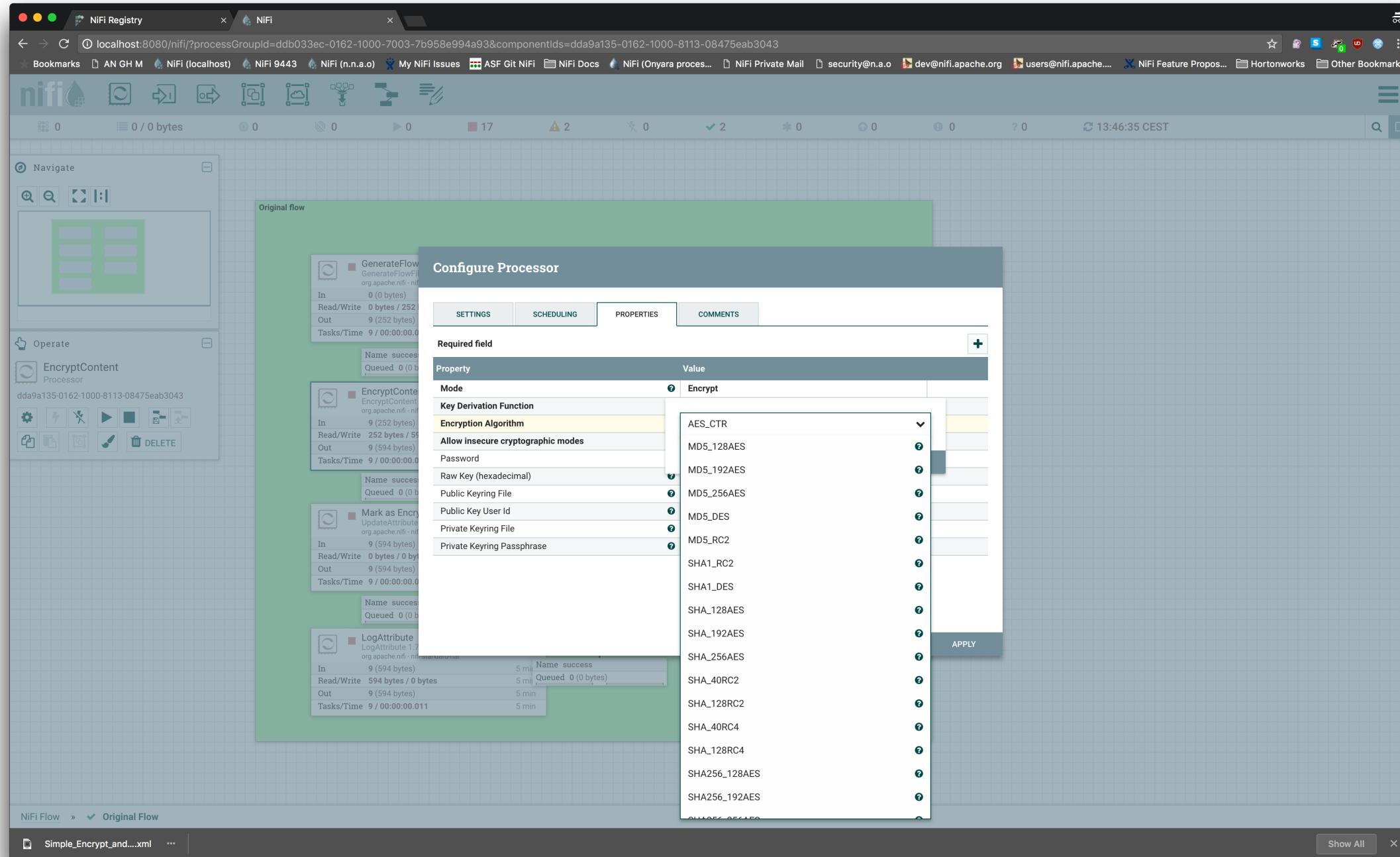
Import new instance into NiFi



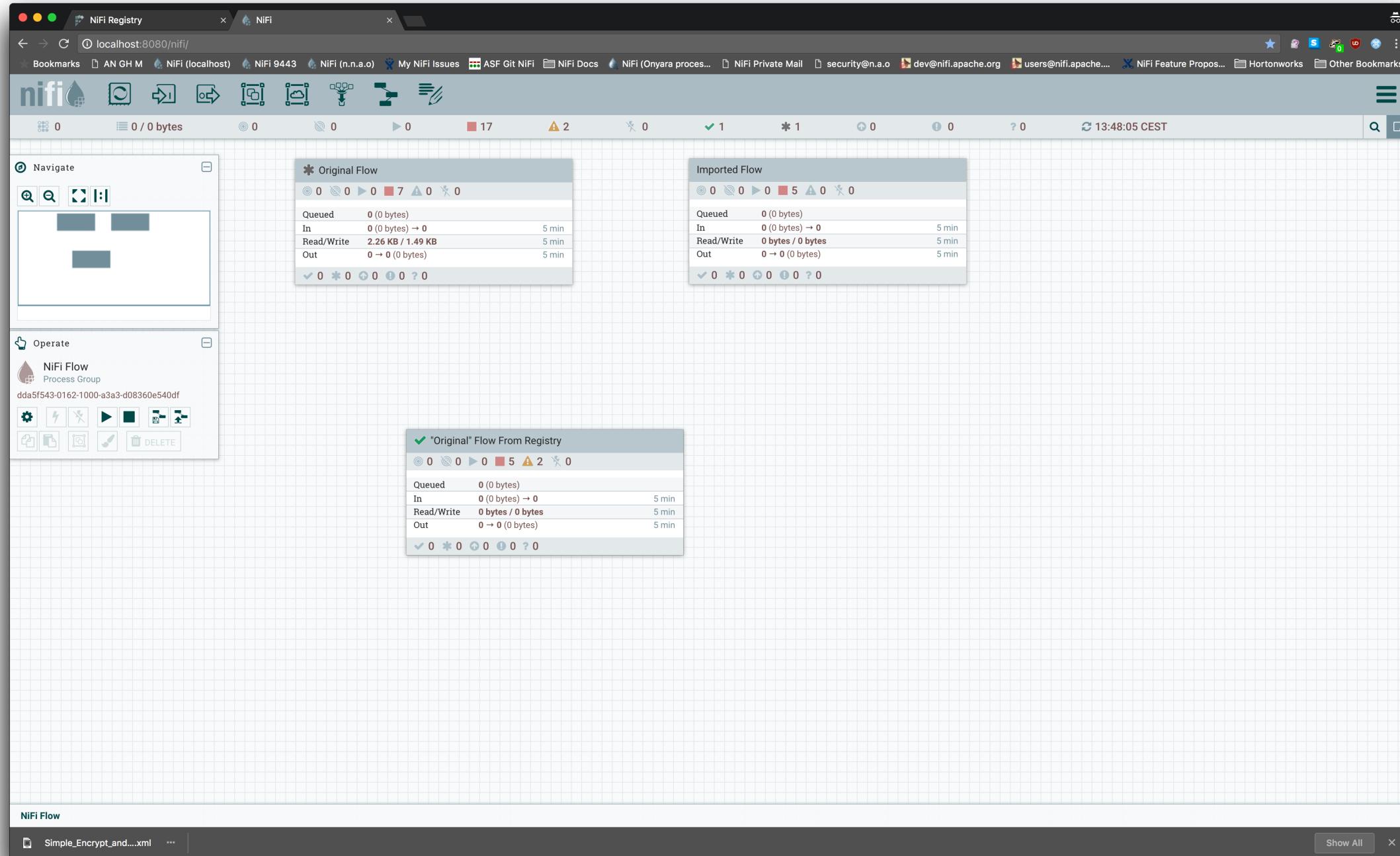
Modify the original flow



Modify the original flow

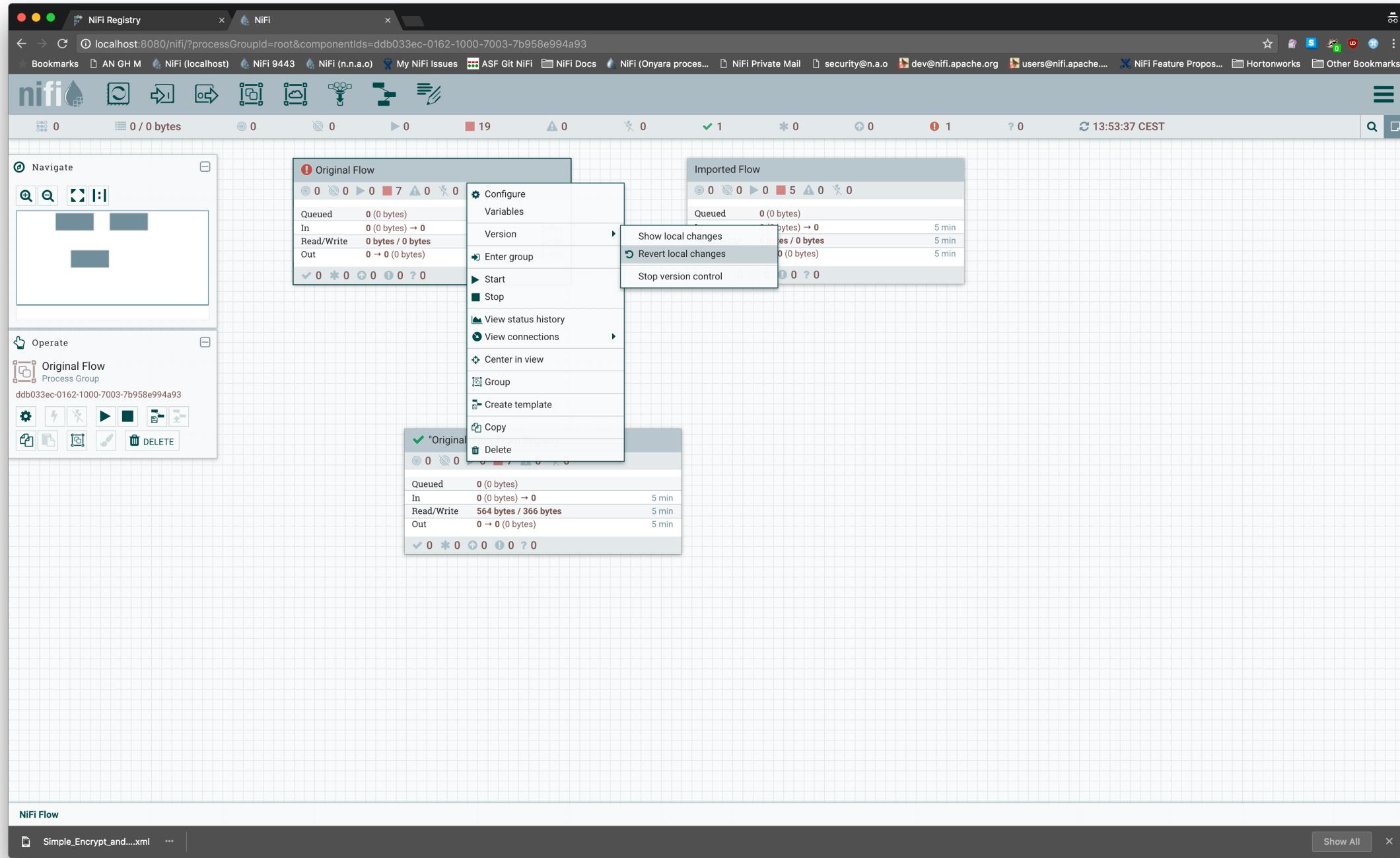


Modify the original flow

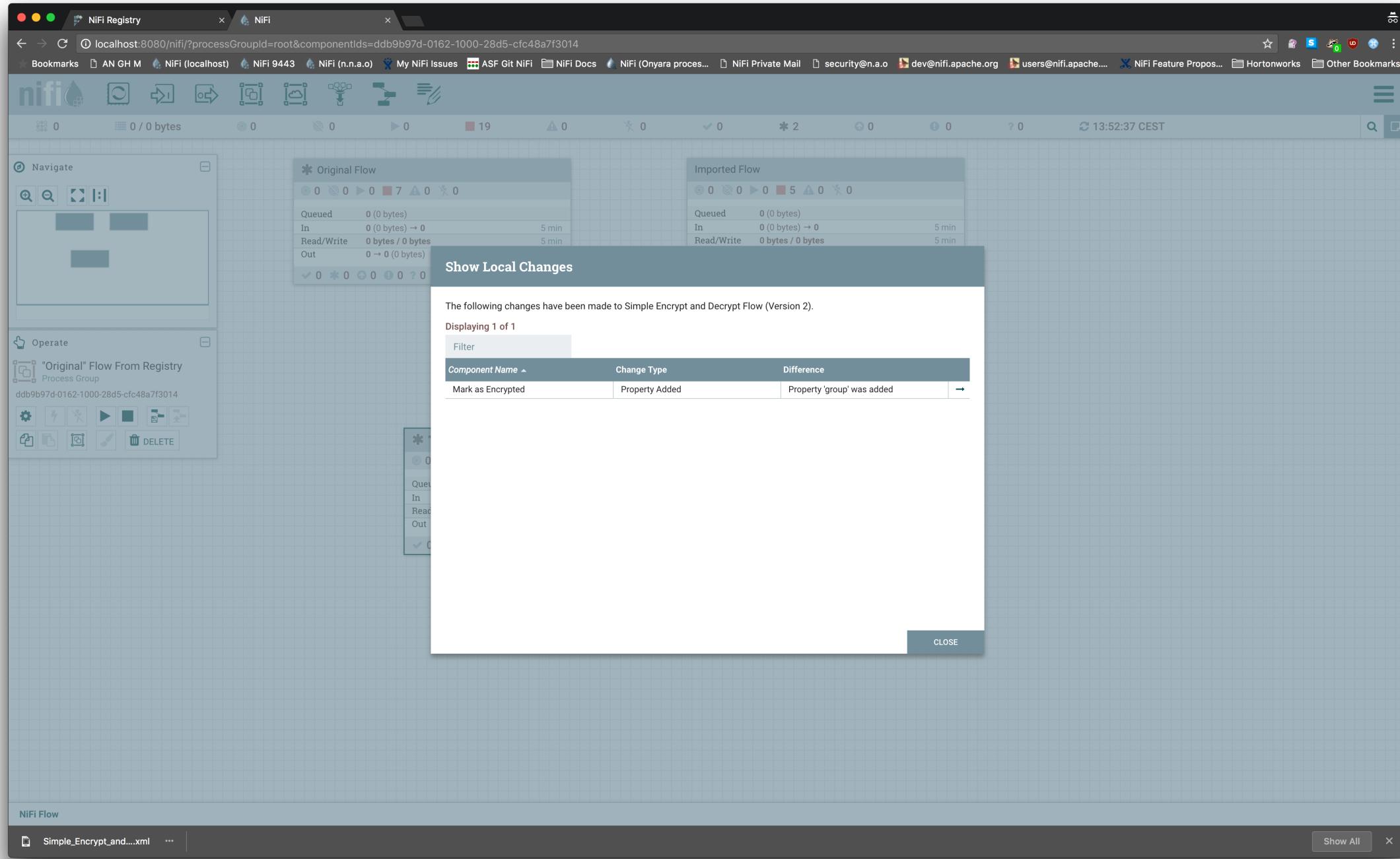


See local changes before committing

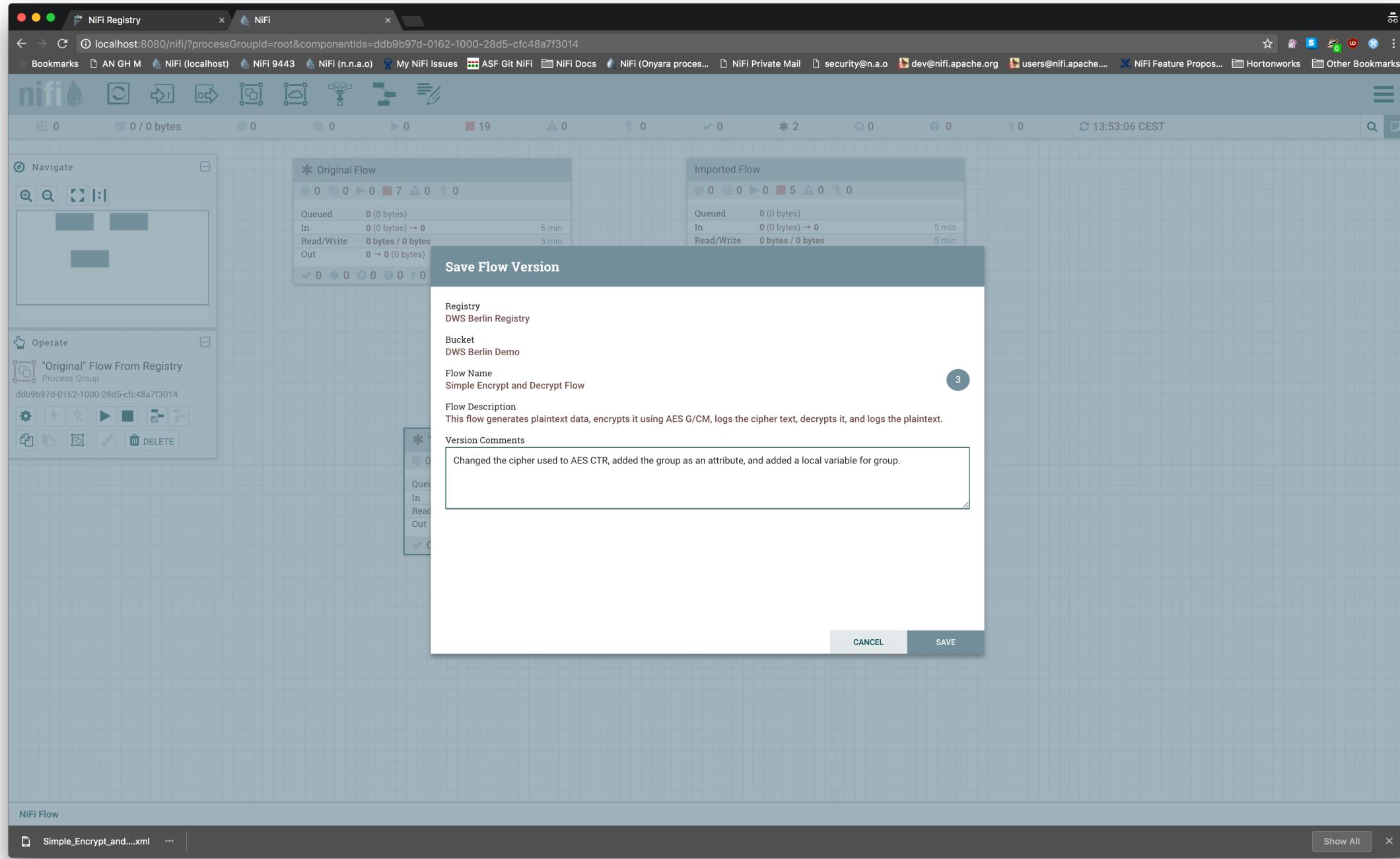
See local changes before committing



See local changes before committing



Commit



Commit

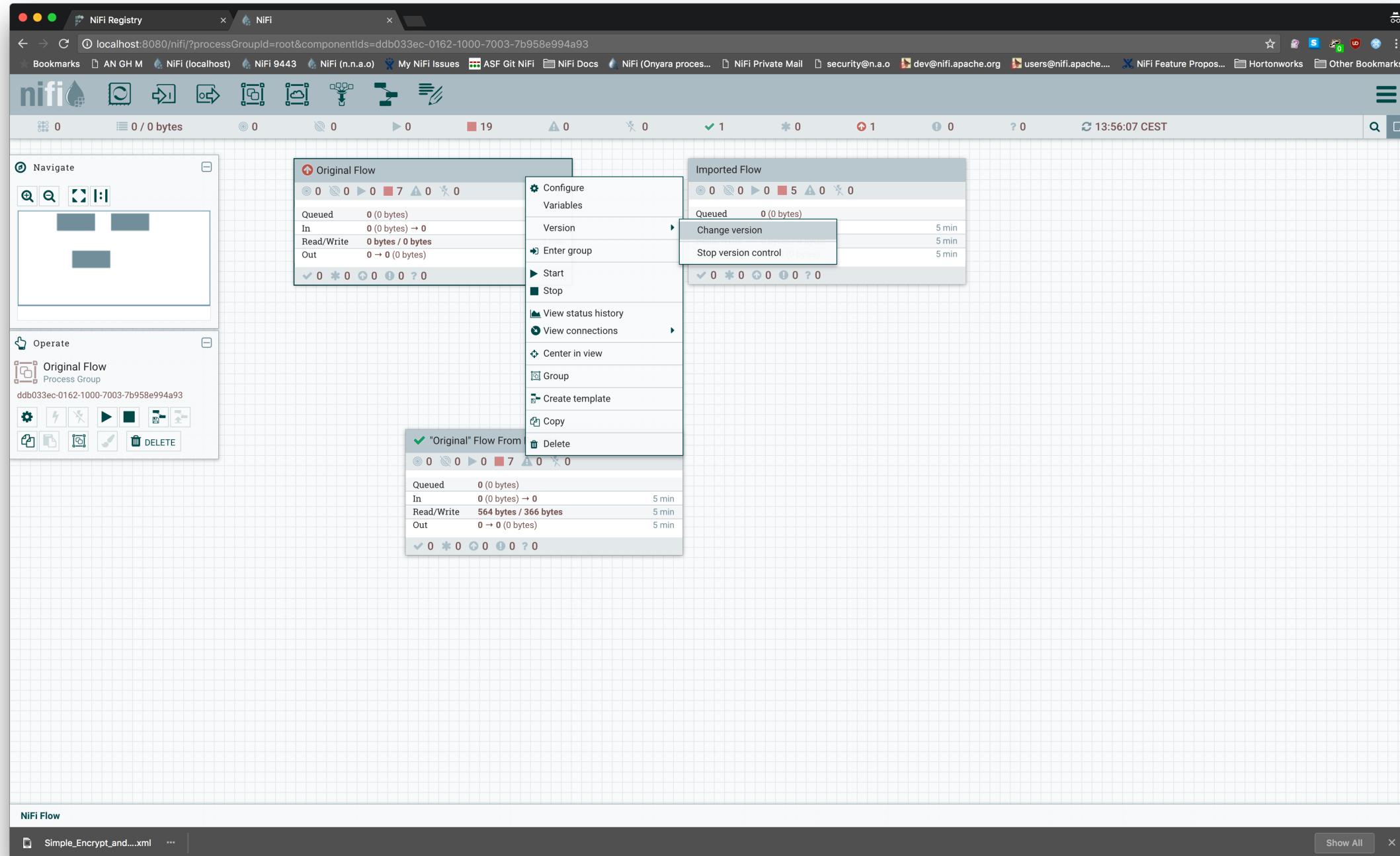
The screenshot shows a web browser window for the NiFi Registry at localhost:18080/nifi-registry/explorer/grid-list. The title bar says "NiFi Registry" and "NiFi". The address bar shows the URL. The top navigation bar includes links for Bookmarks, AN GH M, NiFi (localhost), NiFi 9443, NiFi (n.n.a.o), My NiFi Issues, ASF Git NiFi, NiFi Docs, NiFi (Onyara proces..., NiFi Private Mail, security@n.a.o, dev@nifi.apache.org, users@nifi.apache.org, NiFi Feature Propos..., Hortonworks, and Other Bookmarks. The user is logged in as "anonymous".

The main content area displays a flow named "Simple Encrypt And Decrypt Flow - DWS Berlin Demo". It shows "2 VERSIONS". On the left, there are sections for "BUCKET IDENTIFIER" (d7df1fc1-d14b-4bb5-a8a6-ea57a60c5c98) and "FLOW IDENTIFIER" (b288f8eb-8adf-455a-b887-785239e97d33). The "DESCRIPTION" section states: "This flow generates plaintext data, encrypts it using AES GCM, logs the cipher text, decrypts it, and logs the plaintext." To the right, under "CHANGE LOG", three versions are listed:

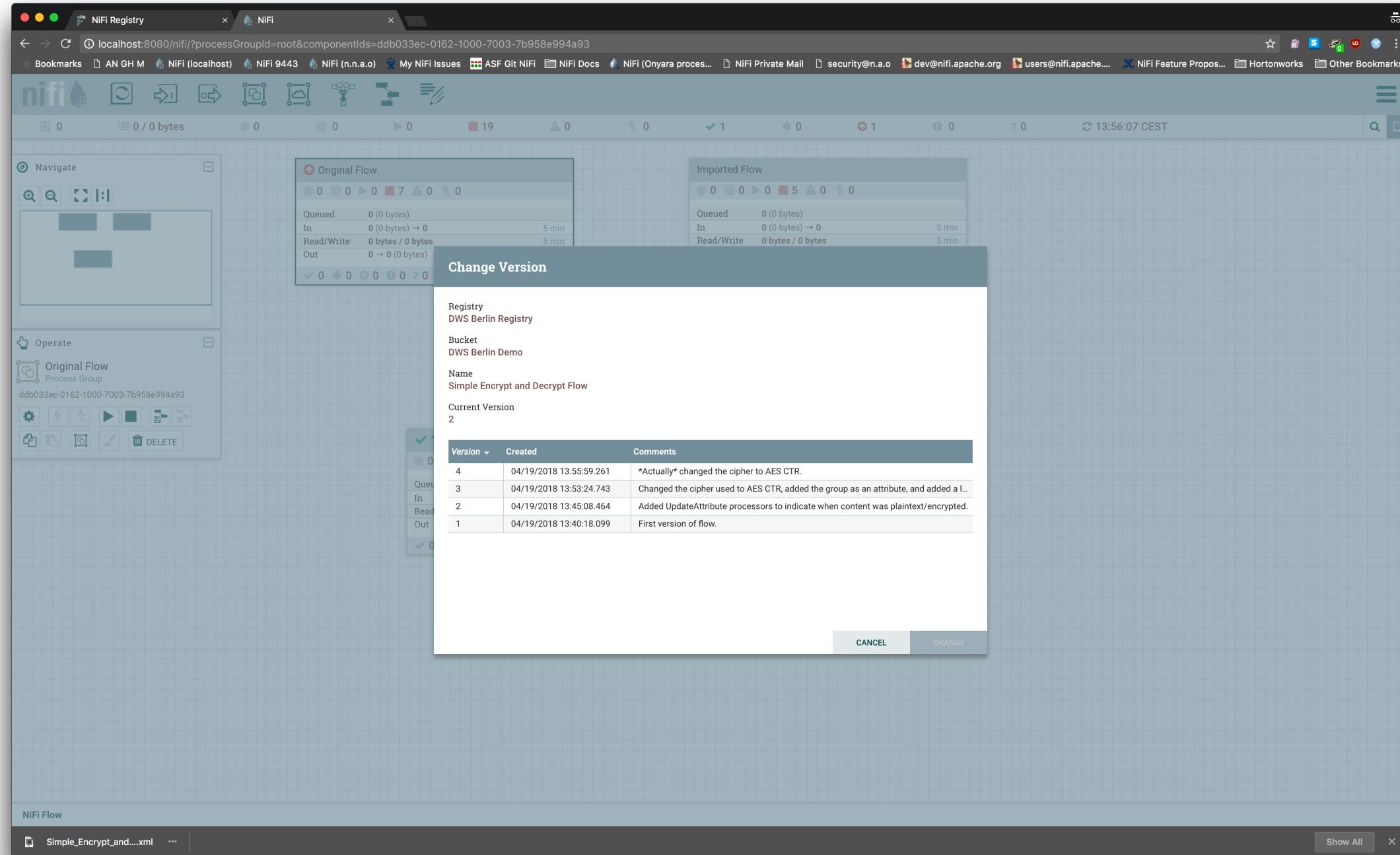
- Version 3 - a few seconds ago (by anonymous): "Changed the cipher used to AES CTR, added the group as an attribute, and added a local variable for group." (Apr-19-2018 at 1:53 PM)
- Version 2 - 8 minutes ago (by anonymous): (no details provided)
- Version 1 - 13 minutes ago (by anonymous): (no details provided)

At the bottom of the browser window, the tab bar shows "Simple_Encrypt_and...xml" and "Show All".

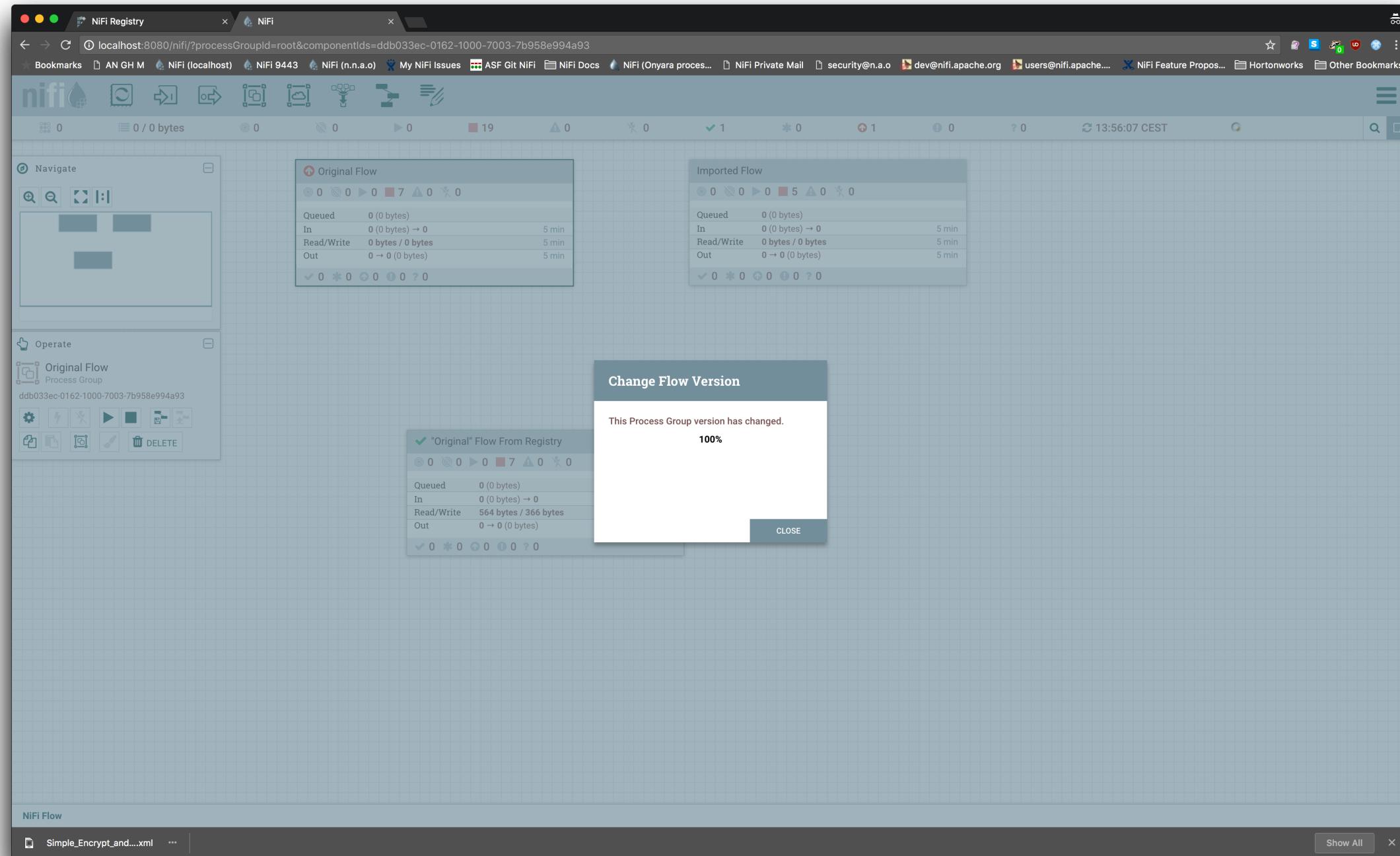
Update new instance from Registry



Update new instance from Registry



Update new instance from Registry



Complementary Tools

Complementary Tools

- NiFi Toolkit
- NiPyAPI
- MiNiFi Converter Toolkit

NiFi Toolkit

- TLS Toolkit
 - Generates, signs, and packages keys and certificates for NiFi services (node/cluster, clients)
- Encrypt Config
 - Protects sensitive configuration values like passwords
- CLI
 - Interacts with NiFi & NiFi Registry to operate on flows

The default output type in interactive mode is `simple`, and the default output type in standalone mode is `json`.

Example of simple output for list-buckets:

```
#> registry list-buckets -ot simple  
  
My Bucket - 3c7b7467-0012-4d8f-a918-6aa42b6b9d39
```

Example of json output for list-buckets:

```
#> registry list-buckets -ot json  
[ {  
  "identifier" : "3c7b7467-0012-4d8f-a918-6aa42b6b9d39",  
  "name" : "My Bucket",  
  "createdTimestamp" : 1516718733854,  
  "permissions" : {  
    "canRead" : true,  
    "canWrite" : true,  
    "canDelete" : true  
  },  
  "link" : {  
    "params" : {  
      "rel" : "self"  
    },  
    "href" : "buckets/3c7b7467-0012-4d8f-a918-6aa42b6b9d39"  
  }  
} ]
```

NiPyAPI

- Python wrapper around NiFi REST API
- Community-provided by Daniel Chaffelson
- Exposes common operations for automation, batch processing, recursion, etc.

```
dev_bucket = nipyapi.versioning.get_registry_bucket(dev_bucket_name)
dev_ver_flow = nipyapi.versioning.get_flow_in_bucket(
    dev_bucket.identifier,
    identifier=dev_ver_flow_name
)
dev_export = nipyapi.versioning.export_flow_version(
    bucket_id=dev_bucket.identifier,
    flow_id=dev_ver_flow.identifier,
    mode='yaml'
)
```

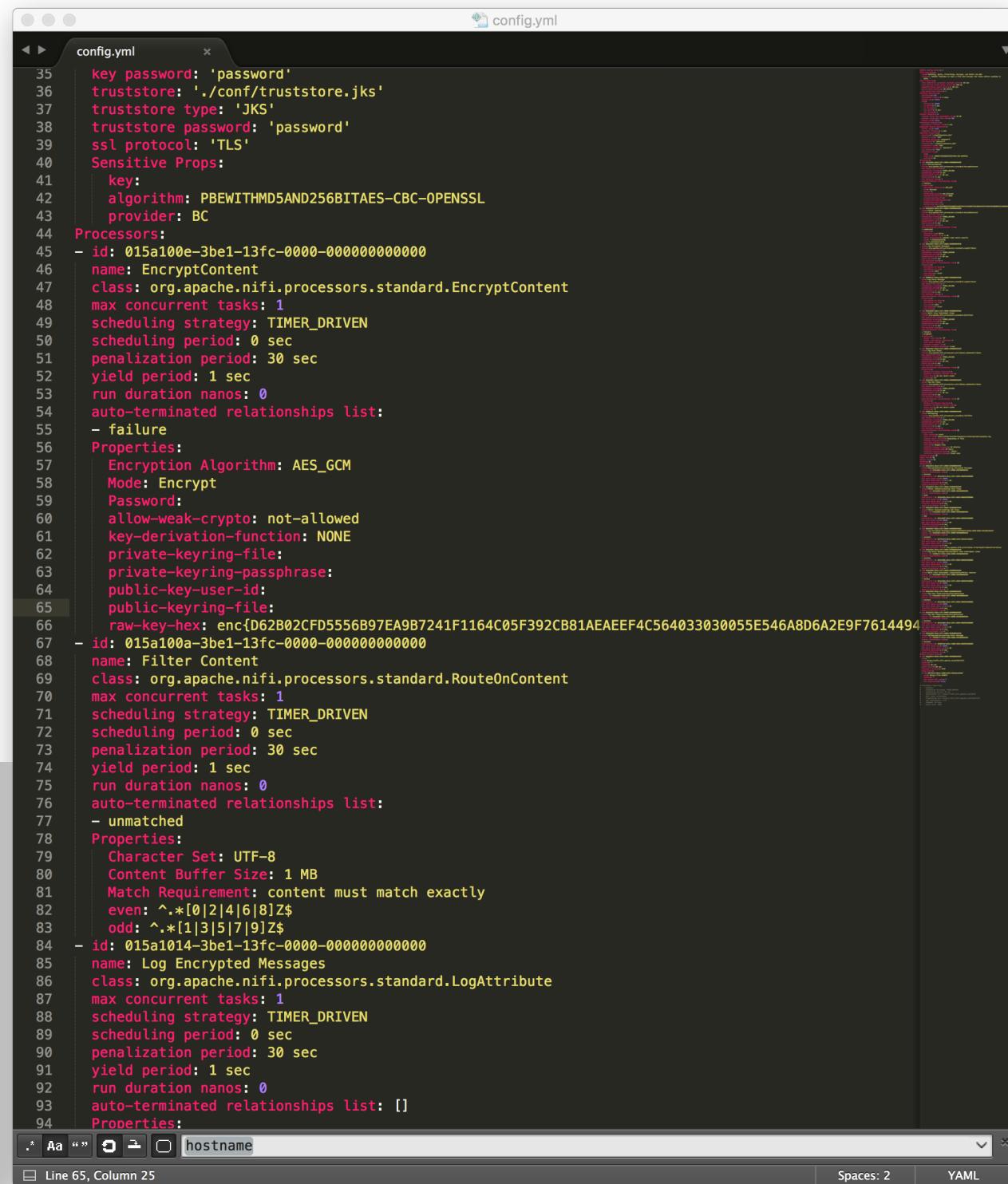
MiNiFi Converter Toolkit

- Save as template from NiFi
- Run \$./bin/config.sh transform template.xml config.yml
- MiNiFi flow ready to run

```
▶ hw12203:...inifi-toolkit-assembly/target/minifi-toolkit-0.1.0-bin/minifi-toolkit-0.1.0 (master) alopresto
2s @ 21:54:55 $ ./bin/config.sh transform ~/Downloads/TailFile\,_Split\,_Prioritize\,_Encrypt\,_and_Exfil_via_S2S.xml config.yml

Java home: /Users/alopresto/.jenv/versions/1.8
MiNiFi Toolkit home: /Users/alopresto/Workspace/scratch/release_verification/minifi-java-0.1.0/minifi-0.1.0/minifi-toolkit/minifi-toolkit-assembly/target/minifi-toolkit-0.1.0-bin/minifi-toolkit-0.1.0

No validation errors found in converted configuration.
▶ hw12203:...inifi-toolkit-assembly/target/minifi-toolkit-0.1.0-bin/minifi-toolkit-0.1.0 (master) alopresto
41s @ 21:55:37 $ ll
total 104
drwxr-xr-x  8 alopresto  staff  272B Feb 21 21:55 .
drwxr-xr-x  3 alopresto  staff  102B Dec  1 11:13 ..
-rw-r--r--  1 alopresto  staff   18K Nov 30 22:55 LICENSE
-rw-r--r--  1 alopresto  staff   11K Nov 30 22:55 NOTICE
-rw-r--r--  1 alopresto  staff   4.6K Nov 30 22:55 README
drwxr-xr-x  4 alopresto  staff  136B Dec  1 11:13 bin/
-rw-r--r--  1 alopresto  staff   8.9K Feb 21 21:55 config.yml
drwxr-xr-x 83 alopresto  staff   2.8K Dec  1 11:13 lib/
▶ hw12203:...inifi-toolkit-assembly/target/minifi-toolkit-0.1.0-bin/minifi-toolkit-0.1.0 (master) alopresto
10s @ 21:55:47 $
```

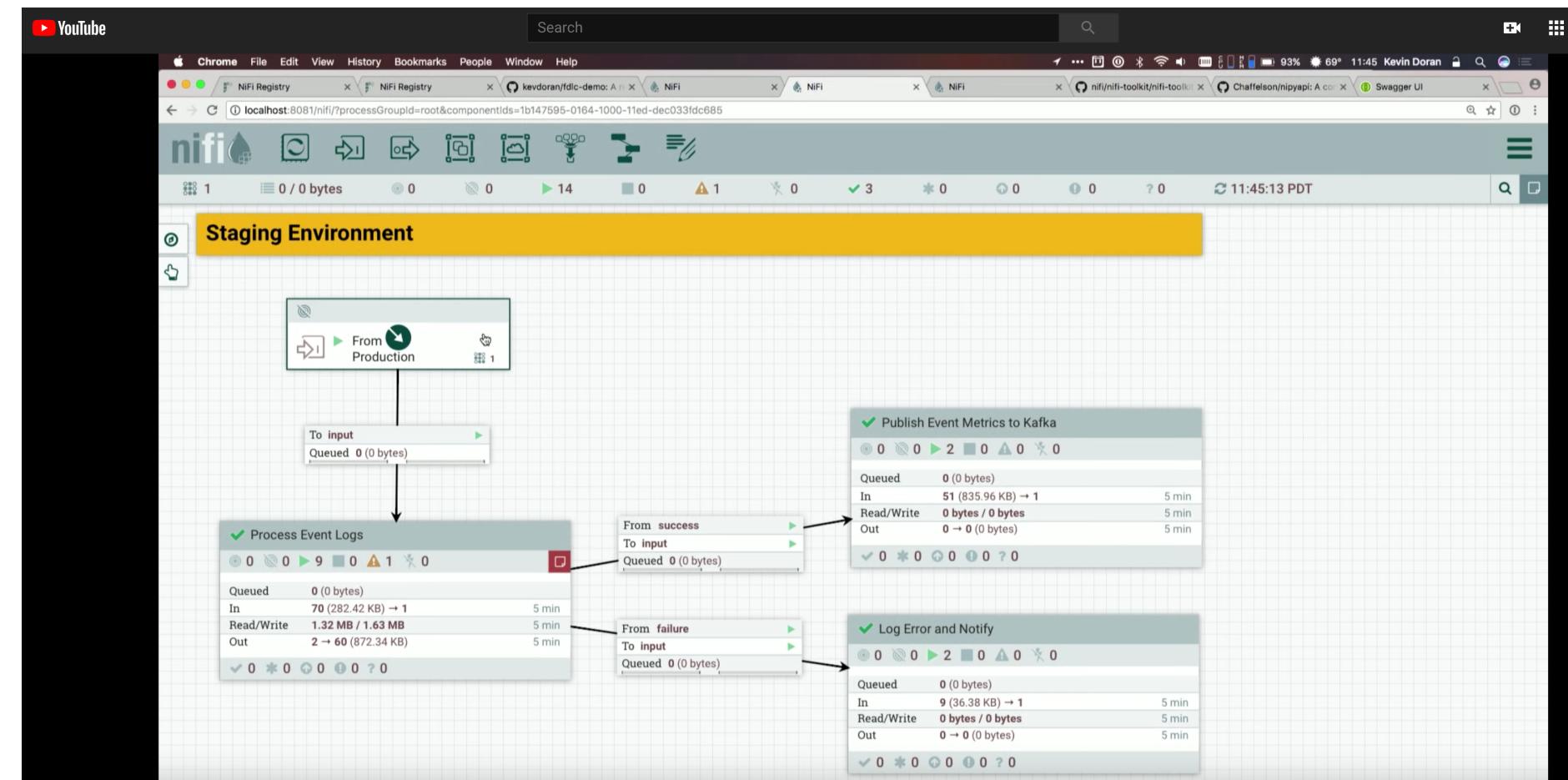


```
config.yml
35 key password: 'password'
36 truststore: './conf/truststore.jks'
37 truststore type: 'JKS'
38 truststore password: 'password'
39 ssl protocol: 'TLS'
40 Sensitive Props:
41   key:
42     algorithm: PBEWITHMD5AND256BITAES-CBC-OPENSSL
43     provider: BC
44 Processors:
45 - id: 015a100e-3be1-13fc-0000-000000000000
46   name: EncryptContent
47   class: org.apache.nifi.processors.standard.EncryptContent
48   max concurrent tasks: 1
49   scheduling strategy: TIMER_DRIVEN
50   scheduling period: 0 sec
51   penalization period: 30 sec
52   yield period: 1 sec
53   run duration nanos: 0
54   auto-terminated relationships list:
55 - failure
56 Properties:
57   Encryption Algorithm: AES_GCM
58   Mode: Encrypt
59   Password:
60   allow-weak-crypto: not-allowed
61   key-derivation-function: NONE
62   private-keyring-file:
63   private-keyring-passphrase:
64   public-key-user-id:
65   public-keyring-file:
66   raw-key-hex: enc{D62B02CFD5556897EA9B7241F1164C05F392CB81AEAEF4C564033030055E546A8D6A2E9F7614494
67 - id: 015a100a-3be1-13fc-0000-000000000000
68   name: Filter Content
69   class: org.apache.nifi.processors.standard.RouteOnContent
70   max concurrent tasks: 1
71   scheduling strategy: TIMER_DRIVEN
72   scheduling period: 0 sec
73   penalization period: 30 sec
74   yield period: 1 sec
75   run duration nanos: 0
76   auto-terminated relationships list:
77 - unmatched
78 Properties:
79   Character Set: UTF-8
80   Content Buffer Size: 1 MB
81   Match Requirement: content must match exactly
82   even: ^.*[0|2|4|6|8]Z$
83   odd: ^.*[1|3|5|7|9]Z$
84 - id: 015a1014-3be1-13fc-0000-000000000000
85   name: Log Encrypted Messages
86   class: org.apache.nifi.processors.standard.LogAttribute
87   max concurrent tasks: 1
88   scheduling strategy: TIMER_DRIVEN
89   scheduling period: 0 sec
90   penalization period: 30 sec
91   yield period: 1 sec
92   run duration nanos: 0
93   auto-terminated relationships list: []
94 Properties:
```

Community

More Resources

- [FDLC with Apache NiFi](#), Kevin Doran
- [NiPyAPI Docs](#), Daniel Chaffelson
- [DevOps Tips](#), Tim Spann
- [Automate Workflow](#), Pierre Villard



New Announcements

- NiFi 1.8.0 – 26 Oct 2018 (212+ Jiras)
 - Jetty, DB improvements
 - Auto load-balancing queues
 - TLS Toolkit w/ external CA
 - Record processor improvements
- MiNiFi C++ 0.5.0 – 6 June 2018
- MiNiFi Java 0.5.0 – 7 July 2018
- NiFi Registry 0.3.0 – 25 Sept 2018



Community Health

apache / nifi

Unwatch ▾ 154 Unstar 1,103 Fork 1,041

Code Pull requests 145 Projects 0 Insights

Mirror of Apache NiFi

4,720 commits

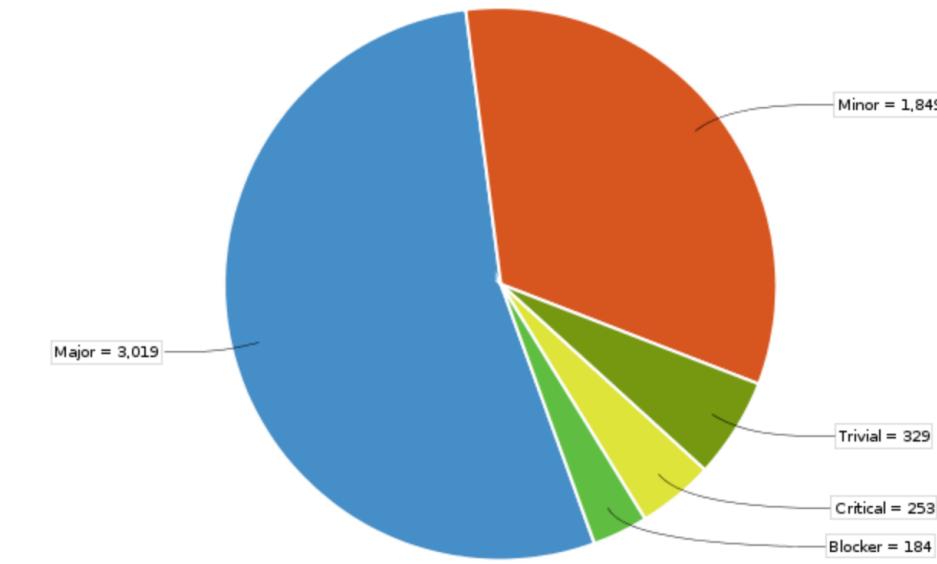
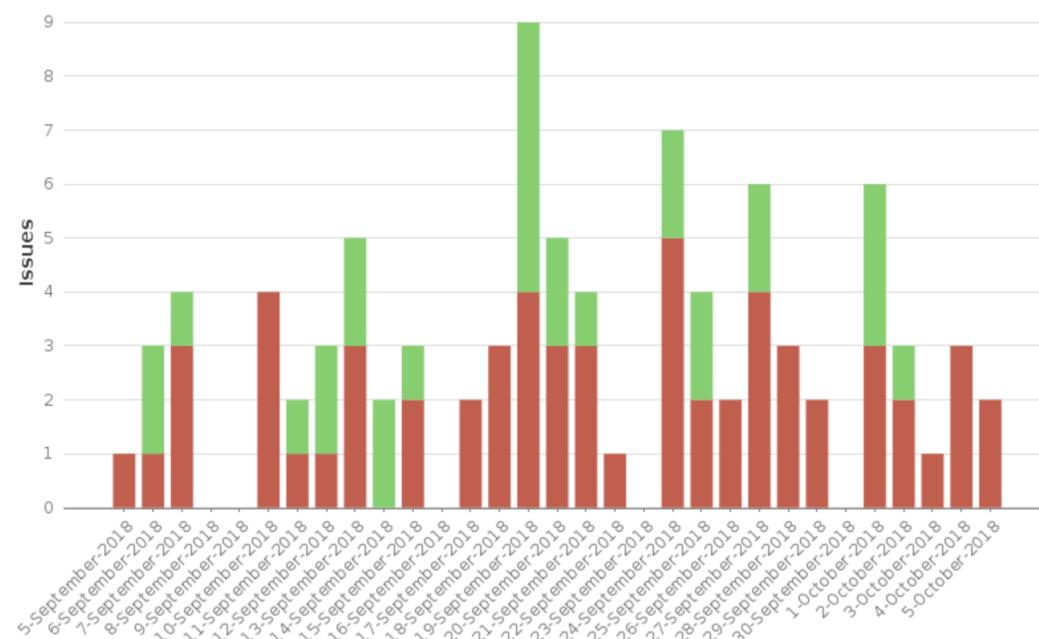
52 branches

63 releases

200 contributors

Apache-2.0

This chart shows the issues created in the last 30 days



Learn more and join us

Apache NiFi site

<https://nifi.apache.org>

Subproject MiNiFi site

<https://nifi.apache.org/minifi/>

Subscribe to and collaborate at

dev@nifi.apache.org

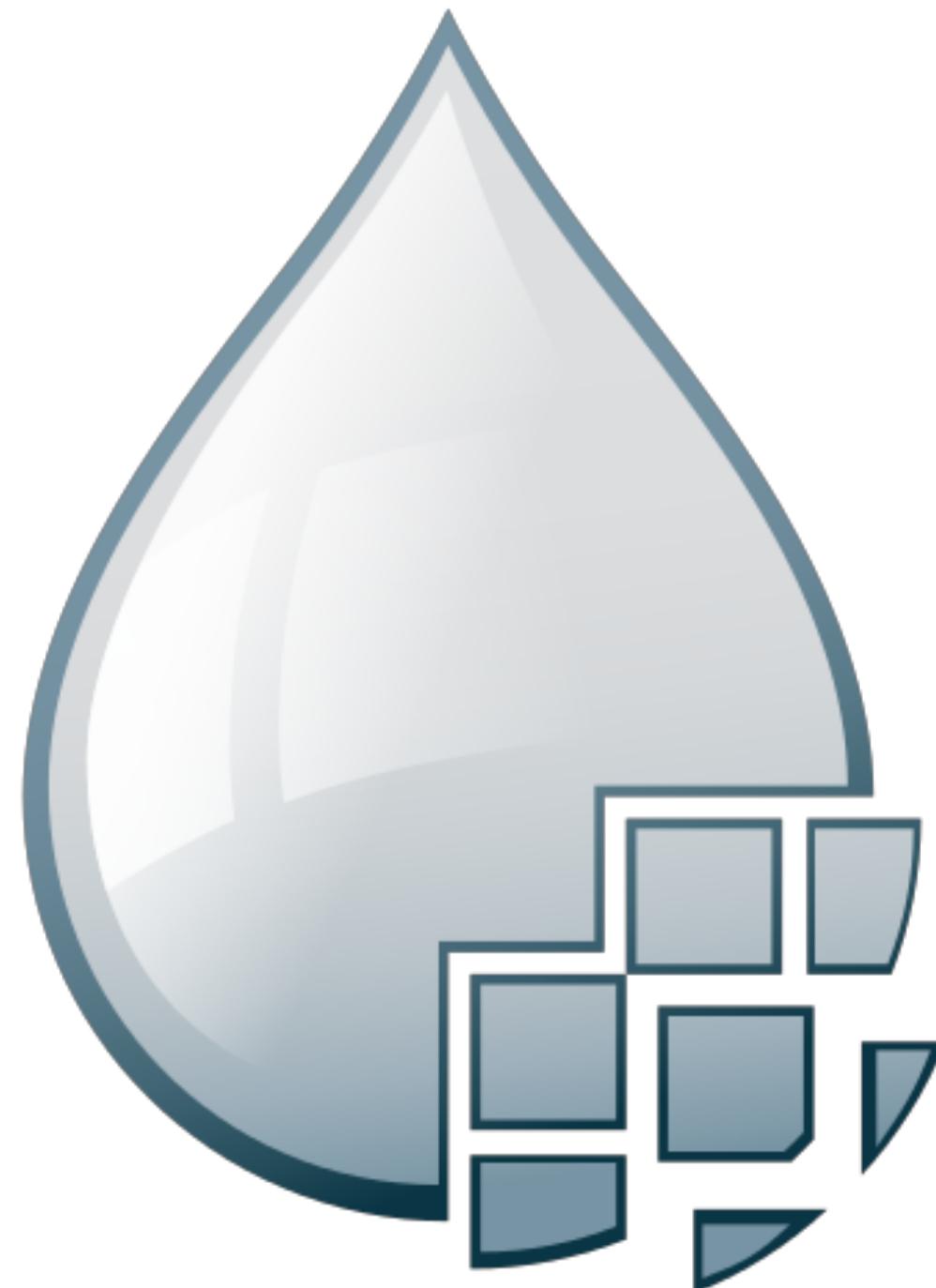
users@nifi.apache.org

Submit Ideas or Issues

<https://issues.apache.org/jira/browse/NIFI>

Follow us on Twitter

[@apachenifi](https://twitter.com/apachenifi)



More NiFi Today

| Title | Time | Room |
|---|-------------|----------|
| The First Mile – Edge and IoT Data Collection with Apache NiFi and MiNiFi | 1100 - 1140 | Room 103 |
| Apache NiFi Crash Course | 1400 - 1600 | Room 109 |
| Dataflow Management From Edge to Core with Apache NiFi | 1650 - 1730 | Room 112 |
| Using Spark Streaming and NiFi for the Next Generation of ETL in the Enterprise | 1650 - 1730 | Room 103 |



Thank you

alopresto@hortonworks.com | alopresto@apache.org | [@yolopey](https://twitter.com/yolopey)
github.com/alopresto/slides