

Текстовая расшифровка видео:

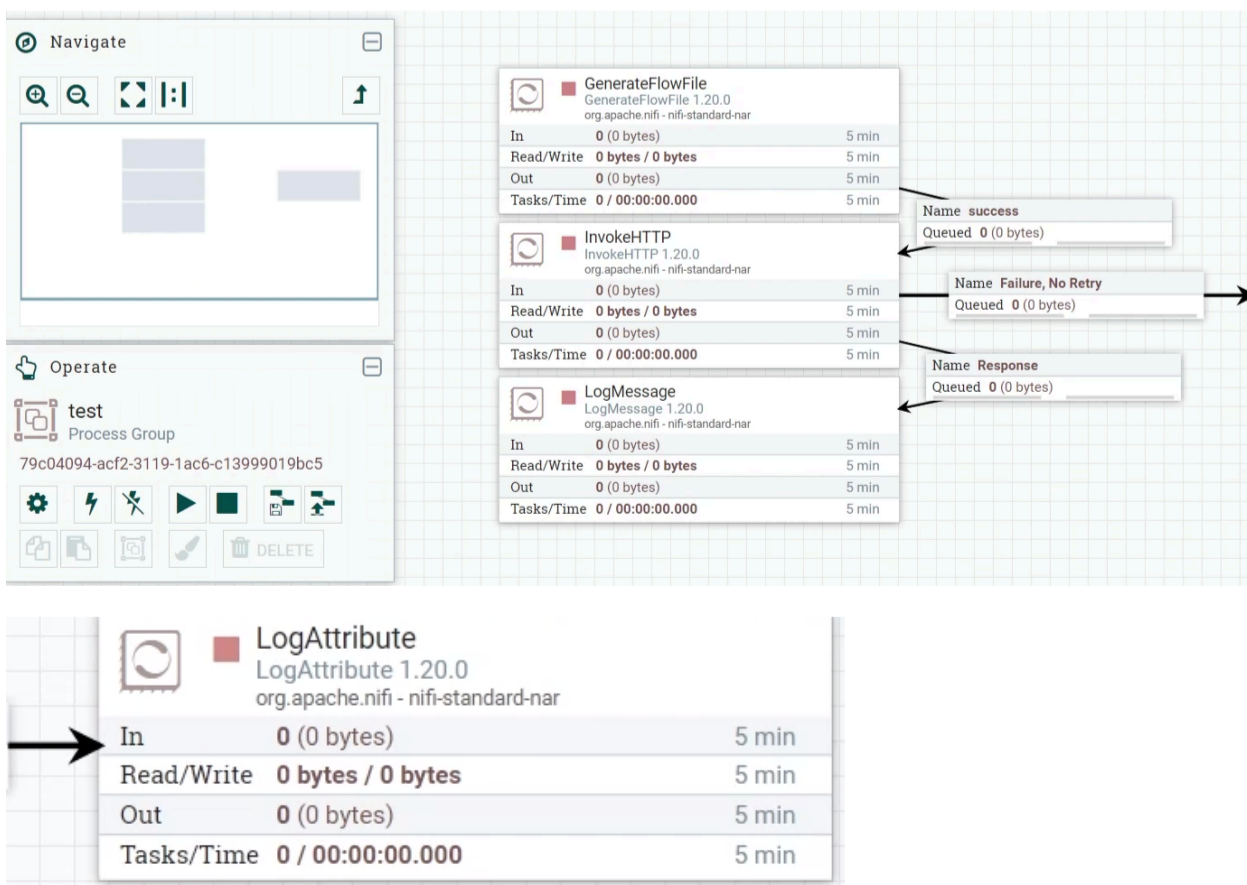
DATA PROVENANCE

План:

- Data Provenance.

Data Provenance

Рассмотрим Data Provenance в консоли NiFi:

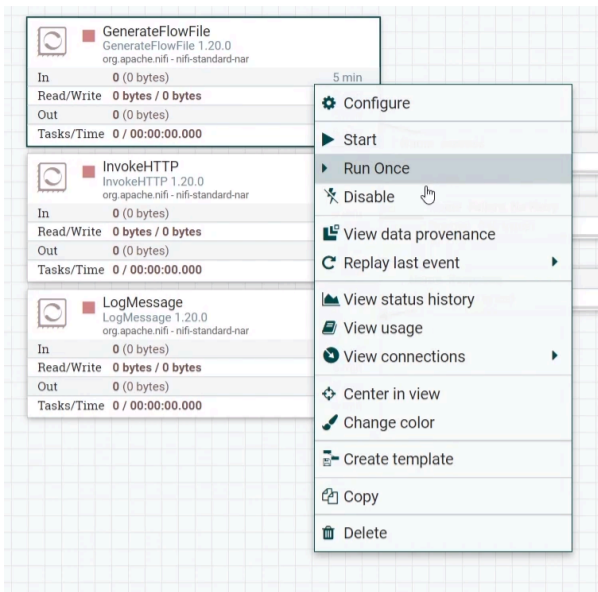


У нас есть простой поток, который будет забирать данные по HTTP.

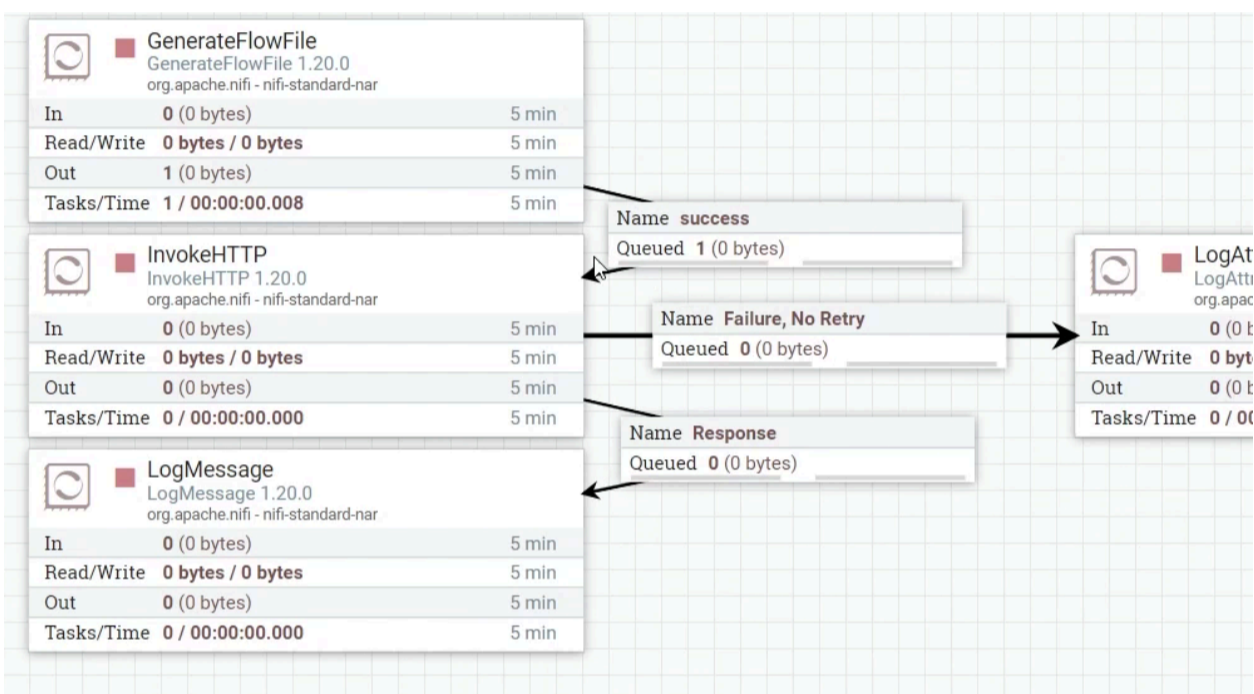


Рассмотрим случай, когда мы используем debug-отладку. Data Provenance довольно обширная и позволяет извлекать данные в любой момент, если позволяет размер репозитория извлекать старые данные, изучать изменения атрибутов, изменения содержимого файла, она может даже делать элементарное логирование.

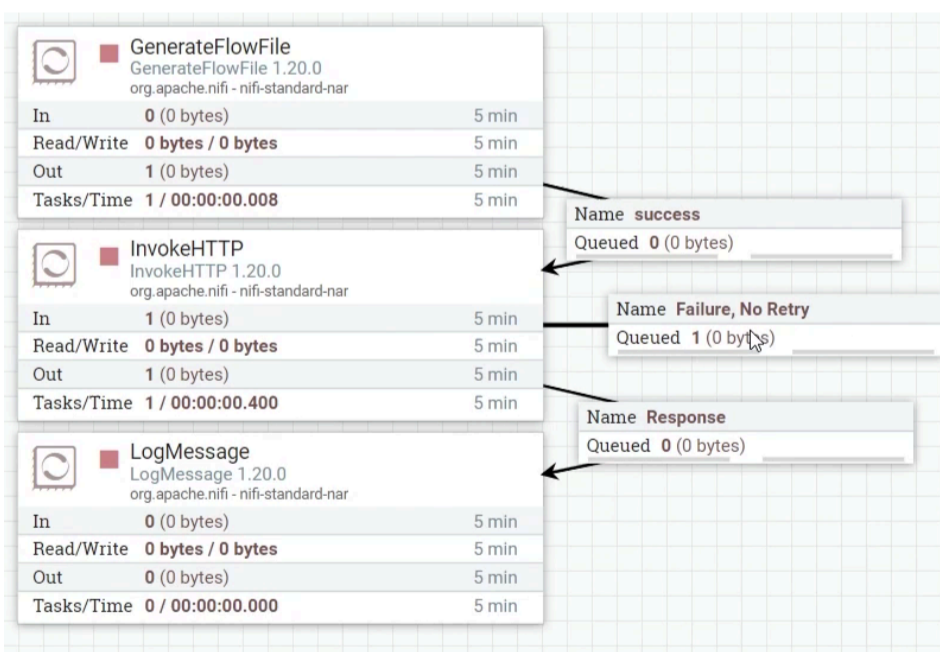
Заострим внимание именно на отладке:



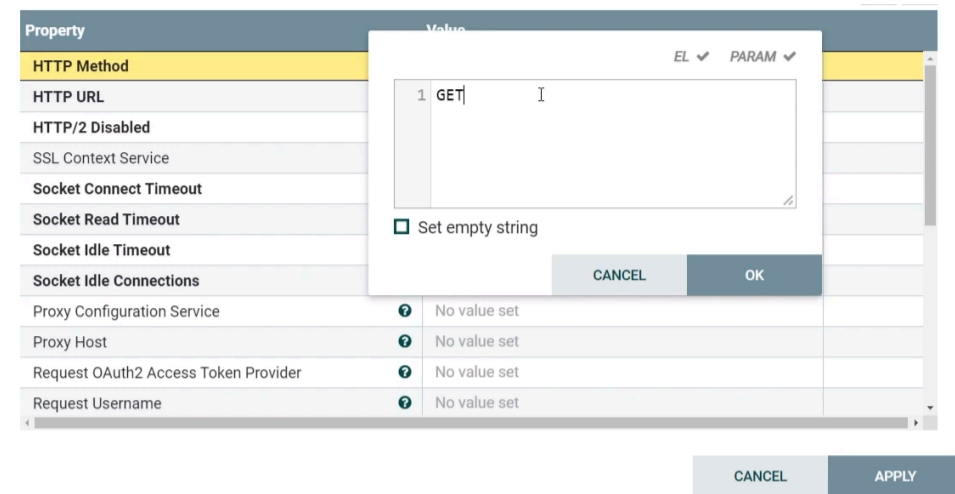
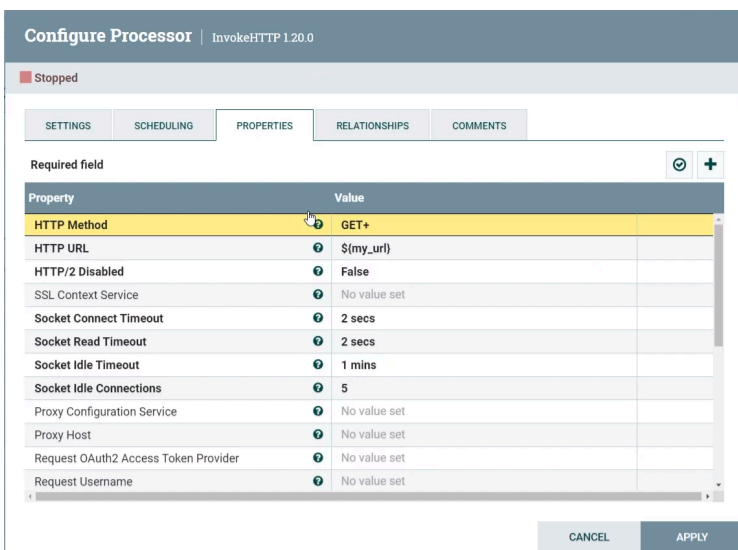
Итак, мы однократно запускаем первый процессор, в следствие чего появляется FlowFile:



Следующий процессор в конфигурации содержит ошибку, при старте данные идут в фейлы:

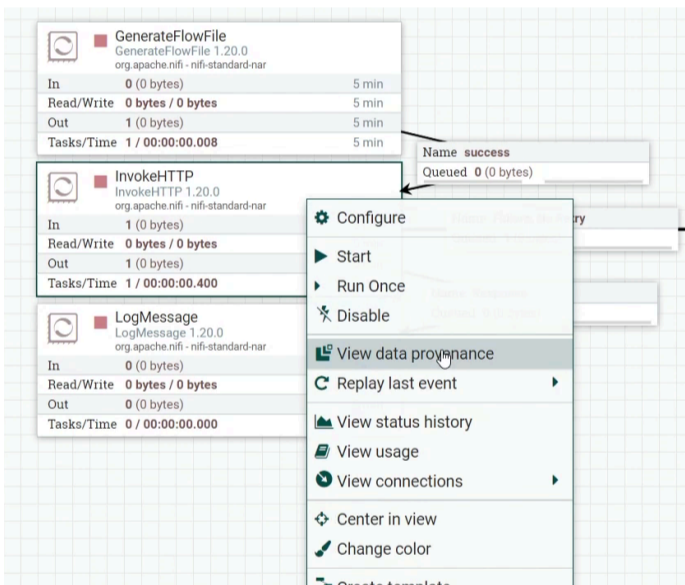


Устраним ошибку (сделана искусственно):

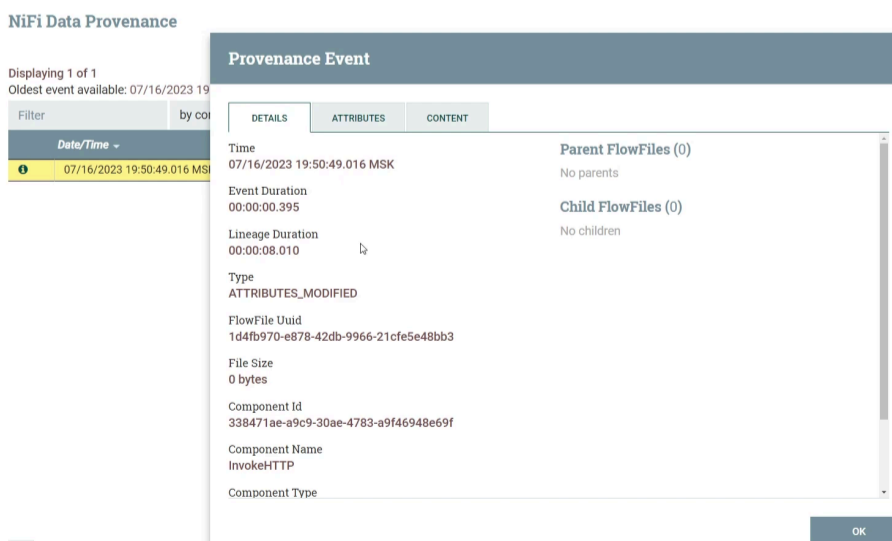


Вопрос: как повторить ту работу, которую делал этот процессор?

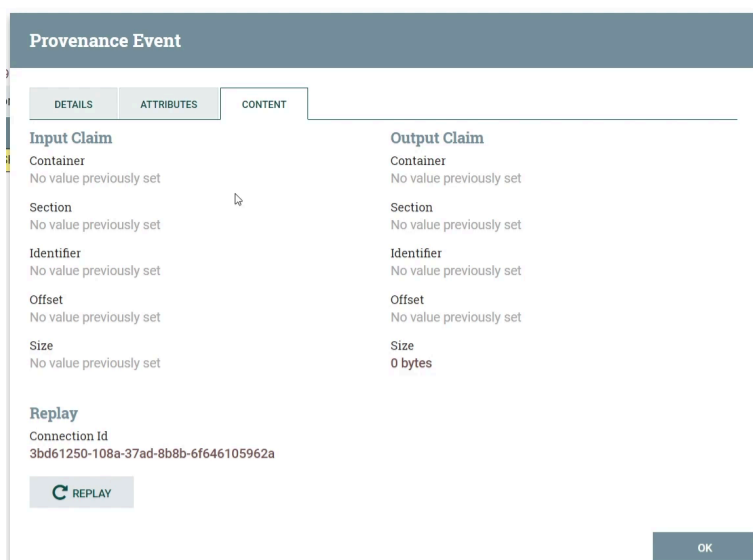
Ответ: для этого можно использовать Data Provenance:



Вызываем Data Provenance через контекстное меню, видим здесь соответствующую запись:



Мы можем изучить содержимое атрибутов и контента FlowFile, которые были здесь ранее в обработке:



В данном случае нас интересует кнопка «Replay»:

Replay

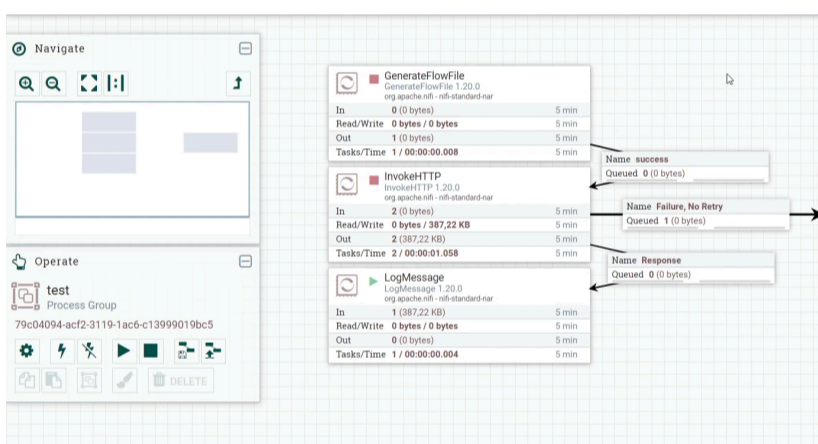
Connection Id

3bd61250-108a-37ad-8b8b-6f646105962a



Нажимая её, мы восстанавливаем на входе перед процессором тот файл, который здесь был ранее.

Запускаем исправленный процесс и видим, что результат появился:



Последний процессор логирует успех обработки, и файл исчезает, то есть файла сейчас в потоке нет, но через Provenance-репозиторий мы можем восстановить его в любой точке нашего потока.

Как вам урок?



Изучил, далее >

Слёрм ©

+7 (495) 248-05-80

[Лицензия №ДЛ-1368 от 22.08.2019](#)

[Политика конфиденциальности](#)

[Публичная оферта](#)