

# Best Practices for Developing With Data Flow on Sandbox

5-April-2023

# FLOW DEV - **NEVER** DO THIS IN THE **SHARED** SANDBOX



Don't upload your private or corporate data



Don't use real usernames, passwords and the like.



All participants have access to your flows and data



Limit resource usage

# FLOW DEVELOPMENT BEST PRACTICES



Name your processors/connections



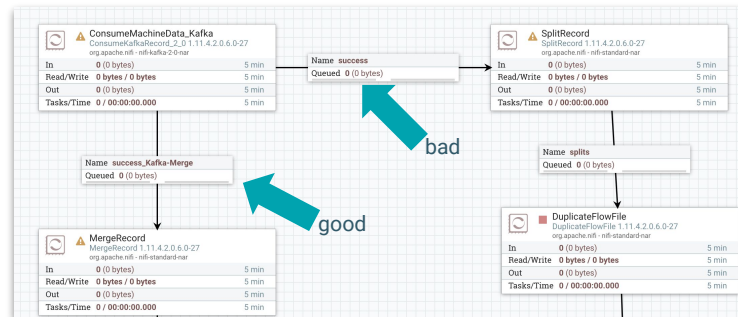
Parameterize connection information



Tag sensitive properties as "sensitive"



Define controller services on process group level (except *Default NiFi SSL Context Service*)



### Configure Processor

Invalid

SETTINGS SCHEDULING PROPERTIES COMMENTS

Required field

Property	Value
Kafka Brokers	#{Kafka Brokers}
Topic Name(s)	#{Machine Data Topic}
Topic Name Format	names

GENERAL CONTROLLER SERVICES

Name	Type	Bundle	State	Scope
AWSCredentialsProviderC...	AWSCredentialsProviderC...	org.apache.nifi - nifi-aws-n...	Disabled	NIFI Flow
ActionHandlerLookup	ActionHandlerLookup 1.1...	org.apache.nifi - nifi-rules...	Invalid	MachineDataToWarehouse
CDPSSLContextServiceLo...	StandardRestrictedSSLCo...	org.apache.nifi - nifi-ssl-co...	Invalid	MachineDataToWarehouse
CSVReader CDP Schema	CSVReader 1.11.4.2.0.6.0...	org.apache.nifi - nifi-record...	Invalid	NIFI Flow

# FLOW DEVELOPMENT BEST PRACTICES

## CDPEnvironment Parameter & Default SSLContextService

### CDP Environment Parameter

- Use whenever Hadoop configuration files are needed
- CDF detects parameter usage, obtains the Hadoop configuration files from SDX, makes them available to NiFi pods and replaces parameter value accordingly
- No copying of config files required anymore

### Default NiFi SSL Context Service

- Use whenever SSL Context Service is required to interact with CDP service in target environment
- CDF detects reference, creates a key and truststore for the target environment and configures a *Default NiFi SSL Context Service* accordingly
- No more manual creation of truststores and moving around of certificates to interact with CDP services
- *Default NiFi SSL Context Service* must be an *external controller service* i.e. defined **outside** of the process group that's exported

### Configure Controller Service

SETTINGS

PROPERTIES

COMMENTS

Required field



Property	Value	
Configuration Resources	#(CDPEnvironment)	→
Username	#(CDP Workload User)	→
Password	No value set	↑

### Configure Processor

Invalid

SETTINGS

SCHEDULING

PROPERTIES

COMMENTS

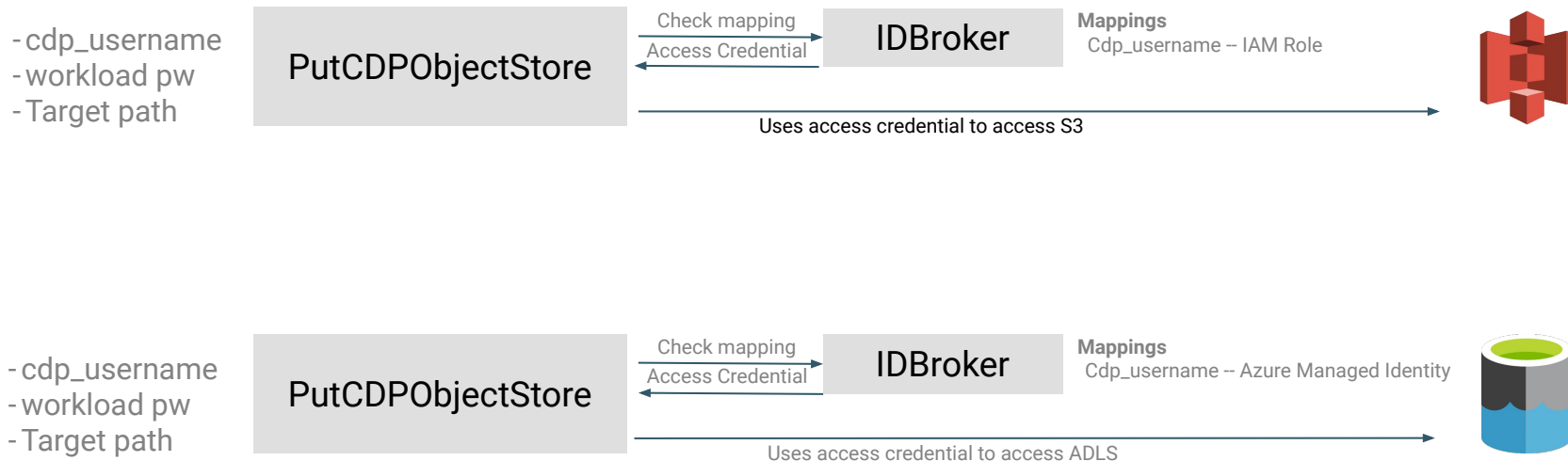
Required field



Property	Value	
Kerberos Keytab	No value set	↑
Username	#(CDP Workload User)	→
Password	No value set	↑
Token Auth	false	↑
SSL Context Service	Default NiFi SSL Context Service	→
Group ID	#(Kafka Consumer Group Id)	→

# FLOW MANAGEMENT CLUSTERS - CDP DATA ACCESS

## Object Store Access Governed by SDX



[https://docs.cloudera.com/management-console/cloud/security-overview/topics/security\\_how\\_identity\\_federation\\_works\\_in\\_cdp.html](https://docs.cloudera.com/management-console/cloud/security-overview/topics/security_how_identity_federation_works_in_cdp.html)