WHITE PAPER

OPTIMIZING SPLUNK LOG INGESTION WITH CLOUDERA DATAFLOW





Government agencies and commercial entities must retain data for several years and commonly experience IT challenges due to increased data volumes and new sources coming online. Due to these factors, they are starting to undergo degradation in the performance of Security Information & Event Management (SIEM's) tools like Splunk. To continue to meet mission needs, address the increase in data sources that require protection, and manage costs, they have started research strategies that complement their Splunk investments while looking for solutions that meet or exceed their organization's policies.

The following requirements must be met to increase the performance of Splunk and maximize IT investment:

- · Route data to utilize cost-effective data destinations
- · Preprocess incoming data by filtering and removing extraneous fields
- Aggregate logs
- · Replay any data
- · Easily move workloads to Splunk Cloud
- · Standardize the onboarding of additional data sources
- Provide observability of data pipelines

This white paper will focus on how agencies use DataFlow for universal data distribution as a solution for Splunk optimization with the technical details required to re-create this work. This white paper does not provide details on how to deploy CDF, instead it includes sample data files and a template for a CDF data flow, that can be found here: https://github.com/Brookslan/UDDS-SIEM-Optimization

Also included is a data flow template demonstrating how to route, filter, and aggregate data from Windows Application Event or IP Stream logs. This approach is designed to be hybrid in nature, meaning it can be used on-premises, in the cloud, or as a combination of both. Due to the flexibility of this architecture, it can be slowly phased into an existing Splunk deployment without interrupting the service.

Routing Data To Utilize Cost-Effective Data Destinations

There are many reasons agencies are seeking to leverage more cost-effective data storage locations to continue to scale and meet business SLAs. They also need flexibility in terms of supporting future architectures and applications. Currently, object storage has become the industry standard due to its ability to:

- · Grow to their needs as data volumes increase
- · Continue to perform analytics at scale
- Allow for organizations to leverage data tiers (hot, medium, or cold) for storage optimization
- Facilitate a hybrid architecture

There are many different software and hardware vendors that support their object storage products, and many of them provide an API that is compliant with the AWS S3 API. For this white paper, we will be using the native capabilities of DataFlow to demonstrate writing a file as an object into a specific bucket location. To complete this task, the NiF processor PutS3Object is required.

In this image, we used the QueryRecords NiFi processor to query and route the corresponding results to a file, which can be set to a different file format (CSV, JSON). Once the output file has been created, we can write this to a pre-existing object storage bucket. This bucket can be hosted on-premise or in AWS cloud infrastructure.



In this set of images, we are showing the configuration values of the PutS3Object NiFi processor to communicate with the object storage instance. This example is using an onpremise deployment of object storage using Apache Ozone. The IP address host containing Ozone Manager is required to complete this. The user will also need to create a bucket in this object storage instance.

If you're using S3 in AWS's cloud, you will need to provide the proper Region of the bucket, and you can leave the Endpoint Override URL value empty. In your environment, the user list may need to be set from your environment, but we are using the user "hadoop" for these values.

Please note that we have security enforcement on the AWS S3 protocol turned off. If security enforcement is turned on, the user must provide the values for Access Key ID and Secret Access Key that match their AWS credentials.



The configuration fields that are important to enter correctly are:

The configuration fields that are important to enter correctly are:
Object Key:
Bucket:
Storage Class:
Endpoint Override URL:
http://<Object Store Host I

- ReducedRedundancy http://<Object Store Host IP>:<Object Store Port> true testUser
- Use Path Style Access:
- Access Key:
 Secret Access Key:
 - 123456

Stopped SCHEDULING PROPERTIES COMMENTS Required field	Configure Pro	ocessor			
SETTINGS SCHEDULING PROPERTIES COMMENTS Required field Property Value Object Key Ø \$(EventFileName) Bucket Ø \$(EventFileName) Content Type Ø \$(EventFileName) Content Type Ø \$(EventFileName) Content Disposition Ø \$(EventFileName) Content Disposition Ø \$(EventFileName) Cache Control Ø \$(EventFileName) Caches Control Ø \$(EventFileName) Access Key ID Ø \$(EventFileName) Secret Access Key ID Sensitive value set Secret Access Key ID Sensitive value set Aves Credentials File No value set AWS Credentials Provider service No value set Object Tags Prefix No value set Remove Tag Prefix Palse Storage Class ReducedRedundancy	Stopped				
Required field Property Value Object Key Ø \$(EventFileName) Image: Colspan="2">Colspan="2">Content Key Bucket Imformationevent Image: Colspan="2">Content Type Ø application/json Content Disposition Itachment Image: Colspan="2">Content Disposition Content Disposition Itachment Image: Colspan="2">Content Disposition Cache Control No value set Image: Colspan="2">Content Colspan="2">Content Disposition Cache Control No value set Image: Colspan="2">Content Colspan="2">Content Disposition Cache Control No value set Image: Colspan="2">Content Colspan="2">Content Disposition Secret Access Key ID Sensitive value set Image: Colspan="2">Content Colspan="2">Content Secret Access Key MS Credentials File No value set Image: Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Content Colspan="2">Colspan="2">Content Colspan="2">Content Colspan="2">Colspan="2">Content Colspan="2">Colspan="2">Content Colspan="2">Colspan="2"	SETTINGS	SCHEDULING	PROPERTIES	COMMENTS	
Value Object Key S(EventFileName) Bucket informationevent Bucket informationevent Content Type application/json Content Disposition attachment Cache Control No value set Access Key ID Sensitive value set Secret Access Key No value set Object Tags Prefix No value set Object Tags Prefix No value set Remove Tag Prefix False	Required field				+
Object Keyا(EventFileName)Image: Sector of the sector of	Property		<u>0</u>	Value	
BucketinformationeventContent Typeapplication/jsonContent DispositionattachmentCache ControlNo value setCache ControlSensitive value setAccess Key IDSensitive value setSecret Access KeySensitive value setCredentials FileNo value setAWS Credentials Provider serviceNo value setObject Tags PrefixNo value setRemove Tag PrefixFalseStorage ClassReducedRedundancy	Object Key		0	\${EventFileName}	1
Content Typeoapplication/jsonoContent DispositionattachmentoCache ControlNo value setoAccess Key IDSensitive value setoSecret Access KeySensitive value setoCredentials FileNo value setoAWS Credentials Provider serviceNo value setoObject Tags PrefixNo value setoRemove Tag PrefixFalseoStorage ClassReducedRedundancyo	Bucket		0	informationevent	
Content DispositionattachmentattachmentCache ControlNo value setContentAccess Key IDSensitive value setContentSecret Access KeySensitive value setContentCredentials FileNo value setContentAWS Credentials Provider serviceNo value setContentObject Tags PrefixNo value setContentRemove Tag PrefixFalseContentStorage ClassReducedRedundancyContent	Content Type		0	application/json	
Cache ControlNo value setImage: ControlAccess Key IDSensitive value setImage: ControlSecret Access KeySensitive value setImage: ControlCredentials FileNo value setImage: ControlAWS Credentials Provider serviceNo value setImage: ControlObject Tags PrefixNo value setImage: ControlRemove Tag PrefixFalseImage: ControlStorage ClassReducedRedundancyImage: Control	Content Disposition	n	0	attachment	
Access Key IDSensitive value setSecret Access KeySensitive value setCredentials FileNo value setAWS Credentials Provider serviceNo value setObject Tags PrefixNo value setRemove Tag PrefixFalseStorage ClassReducedRedundancy	Cache Control		0	No value set	
Secret Access Key Sensitive value set Credentials File No value set AWS Credentials Provider service No value set Object Tags Prefix No value set Remove Tag Prefix False Storage Class ReducedRedundancy	Access Key ID		0	Sensitive value set	
Credentials File No value set AWS Credentials Provider service No value set Object Tags Prefix No value set Remove Tag Prefix False Storage Class ReducedRedundancy	Secret Access Key		0	Sensitive value set	
AWS Credentials Provider service No value set Object Tags Prefix No value set Remove Tag Prefix False Storage Class ReducedRedundancy	Credentials File		0	No value set	
Object Tags Prefix No value set Remove Tag Prefix False Storage Class ReducedRedundancy	AWS Credentials Pr	rovider service	0	No value set	
Remove Tag Prefix Ø False Storage Class Ø ReducedRedundancy	Object Tags Prefix		0	No value set	
Storage Class Ø ReducedRedundancy	Remove Tag Prefix		0	False	
	Storage Class		0	ReducedRedundancy	

Stopped				
SETTINGS	SCHEDULING	PROPERTIES	COMMENTS	
Required field				-
Property			Value	
Region		0	US East (N. Virginia)	
Communications	Timeout	0	30 secs	
Expiration Time Ru	ıle	0	No value set	
FullControl User L	ist	0	hadoop	
Read Permission I	Jser List	0	hadoop	
Write Permission	User List	0	hadoop	
Read ACL User Lis	it	0	hadoop	
Write ACL User Lis	st	0	hadoop	
Owner		0	hadoop	
Canned ACL		0	\${s3.permissions.cannedacl}	
SSL Context Servi	ce	0	No value set	
Endpoint Override	URL	0	http://172.19.0.11:9878	

Stopped			
SETTINGS SCHEDULING PROPERTIES	s	COMMENTS	
Required field			-
Property		Value	
Signer Override	0	Default Signature	
Multipart Threshold	0	5 GB	
Multipart Part Size	0	5 GB	
Multipart Upload AgeOff Interval	0	60 min	
Multipart Upload Max Age Threshold	0	7 days	
Server Side Encryption	0	None	
Encryption Service	0	No value set	
Use Chunked Encoding	0	true	
Use Path Style Access	0	true	
Proxy Configuration Service	0	No value set	
Proxy Host	0	No value set	
Proxy Host Port	0	No value set	

 $Optimizing \ Splunk \ Log \ Ingestion \ With \ Cloudera \ Data Flow$

Pre-Process Incoming Data by Filtering and Removing Extraneous Fields

To improve the performance of Splunk or other SIEMs, it is critical to be selective about the data sent to be indexed and searched in the future. There are many advantages to limiting the amount of data an analyst has to sift through later. With less data to be searched through, it allows for queries to be completed quicker, and critical events can be given more attention or immediate action.

Concerning filtering data, CDF can complete this task using native processors. Data can also be filtered based on any data or metadata associated with the flow file. In the following example, we will be using the event severity level as a filter to determine if data should be routed to our SIEM or object storage. This example will pull the severity level, which is a metadata value, and only send events that are non-information to our SIEM (i.e., Splunk). Information events will be routed to our object storage bucket.

To complete this task, we will use the QueryRecord NiFi processor to query incoming records for severity level and event codes. This acts as a filter and routes the results based on the query. In this example, the following SQL statements are used for severity and event codes. Please note, the RPATH call is used to traverse through nested JSON elements in the input file.

- Severity Information SELECT * FROM FLOWFILE WHERE RPATH ("result", '/severity') = 'Information'
- Severity Warning SELECT * FROM FLOWFILE WHERE RPATH ("result", '/severity') = 'Warning'
- Event Code 102 SELECT * FROM FLOWFILE WHERE RPATH ("result", '/EventCode') = '102'
- Event Code 1001 SELECT * FROM FLOWFILE WHERE RPATH ("result", '/EventCode') = '1001'



Processor Details			
Running		0	STOP & CONFIGU
SETTINGS SCHEDULING PROPERTIES	31	COMMENTS	22
Required field			
Property		Value	
Record Reader	0	WInEventAppJsonTreeReader	-
Record Writer	0	WinEventAppJsonRecordSetWriter	-
Include Zero Record FlowFiles	0	false	
Cache Schema	0	true	
EventCode1001	0	SELECT * FROM FLOWFILE WHERE RPATH ('result', '/EventCode'	
EventCode102	0	SELECT * FROM FLOWFILE WHERE RPATH ("result", '/EventCode'	
Severity - Informational	0	SELECT * FROM FLOWFILE WHERE RPATH ('result', '/severity') = '	
Severity - Non-Informational	0	SELECT * FROM FLOWFILE WHERE RPATH ("result", '/severity') = '	

Removing unrequired or extra fields can be accomplished using the built-in SQL capabilities in the QueryRecords NiFi processor. Once those attributes are present in the query, they will be added to the output file that would be sent to a SIEM or an object storage instance.

Aggregate Logs

Log aggregation can enable organizations with greater control over the flow of data through their infrastructure as well as control over how data is written to its destination. In many ways, this can be considered transformation in an ETL process. By taking advantage of the native capabilities of the QueryRecords NiFi processor, the results from the queries are merged together out of the box. This allows developers to leverage all of the results of the query into a single output file. The following images show the results of the query merged into a single output file.



Replay Any Data

All NiFi processors can replay any data that has been processed through them. This capability is a valuable tool for development or debugging purposes and is very easy to use. In the following images, a user must right-click on the processor of interest and select the option to view data provenance. A list of events will be displayed to the user, and one of those events can be selected. Under the Content tab, the replay button is available, which can be clicked to replay that data.

WHITE PAPER

CLOUDERA

	e 🕆 🤰 💋	
🗞 3/3 🗄 0 🔳 0 (0 bytes) 💿 0	0 ► 0 ■ 7 ▲ 9 3 0 ✓ 0 ≠ 0	⊖ 0 0 0 7 0 C 16:17:01 UTC
© Norigate 😑	Name success Queued 0 (0 bytes)	
Operate Opera	Query & Route Records QueryRecord 1.16.0.2.2.5.1-1 org apache.nif - nif-standard-nar In 0 (0 bytes) Read/Write 0 bytes / 0 bytes Out 0 (0 bytes) Tasks/Time 0 (0 bytes)	Configure Start Start
rentCode 1001 0 (0 bytes)		View connections View connections View connections Center in view Change access policies Center in view Change color Craste template Copy Delive

Filter	by compo	oent name 🗸						
	Date/Time -	Type	Paretile blad	tire:	Component Name	Component Type	Nada	
0	06/16/2022 20 15:37 897 UTC	CONTENT, MODIFIED	160cfe11-18fe-425f0c02-773a/b03	239 bytes	AttributesTp.JSON	AttributesToJSON	ec42507902e5 8090	8-
0	06/16/2022 20 15:37.897 UTC	CLONE	160cfs11-18fs-425f-0c62-773a/b63.	239 bytes	AttributesTo.1904	AltributesTo. ISON	m:42503902e5 8080	A
0	06/16/2022 20 15 37 #94 UTC	CONTENT, MODIFIED	18v18927-ac/74-4530-9242-cb2189.	-461 bytes	AttributesTp.3504	AttributesToJSON	di:42503902e5 8080	5-
0	06/16/2022 20.15.37 #V0 UTC	CLONE	18e18927-ac74-4530-92d2-cb2189	461 bytes	AttributesToJSDN	AttributesTo.150N	dic42503902x5.8080	A -
0	06/16/2022 20 15 37 #92 UTC	CONTENT, MODIFIED	\$329660e-036d-4314-x792-767148	217 bytes	AttributesTo.JSON	AttributesTp.JSDN	dc42503902e5.8000	A-
0	06/16/2022 20 19:37 886 UTC	CLONE	\$329660e-0366-4314-a792-767148.	217 bytes	AttributesTo.JSQN	AttributesTs.JSON	6c42503902x5 8080	5-
0	06/16/2022 20 15:37 884 UTC	CONTENT_MODIFIED	149ce5be-9cb8-4a33-8d57-19bc98.	453 bytes	Attributes7s.ISON	AttributesTo.ISON	dc42503902e5 8080	5-
0	06/16/2022 20:15:37 884 UTC	CLONE	¥48ce5be9cb84a33-8d57-19bc98.	A63 bytes	AttributesTaJSON	AttributesTo.JDON	dt/42503902e5.8080	5-
0	06/16/2022 20 15:37 882 UTC	CONTENT, MODIFIED	1e9928a14c4F4e41-8066-6131e746	515 bytes	AttributesTa, JSON	AttributesTa. ISON	dc42507902e5.8080	A-
0	06/16/2022 20:10:37 ##1 UTC	CLONE	8e9928al-to4F-te41-0066-6131e748.	515 bytes	AttributesTp.190N	AttributesTa.ISON	di;42507902x5 8080	A-
0	06/16/2022 20 15 37 880 UTC	CONTENT_MODIFIED	83410745-9108-4465-8585-962043.	231 bytes	AttributesTe.JDON	AttributesTaUSON	dc42507902x5 8080	3-
0	06/16/2022 20:15:37 880 UTC	CLONE	#3a1d746-9158-4a65-8586-3625a3	231 bytes	AttributesTa.tION	AttributesTe, ISON	dc42507902e5 8080	A
0	06/16/2022 2015:27.877 UTC	CONTENT_MODIFIED	e121/2994 70ae-8100-05a5-ac50030.	463 3ytes	ArtifibutesTo.190N	ActributesTa.JSON	6:42503902+5 8090	A-4
0	06/16/2222 20.15.37.877 UTC	CLONE	#121/2994 70ae-4100-55a5-ac50030.	463 bytes	AttributesTo.150N	AttributesTa.JSDN	0(42503902+5 8080	4-4
0	06/16/2022 20:15:37 #75 UTC	CONTENT_MODIFIED	838c48cd #5e0 4679 8000 70/k23	463 bytes	AthibutesTeJSON	AttributesTeJ30N	0142507902x5 8080	A-4
0	06/16/2022 20 15:57 874 UTC	CLONE	038:48:d afeb 4679 8000 70/823.	463 bytes	AttributesTo, ISON	AttributesTo.350N	ds;42503902e5 8080	4-
0	06/16/2022 20 15:37 871 UTC	CONTENT, MODIFIED	#6434654519456e.6fs7149235	463 bytes	ArteiloutesTo.1504	AttributesTo.JSON	dc42503902e5 8080	A-4
0	06/16/2022 20 15:37:871 UTC	CLONE	a6d3ebd5#319-4b6e-bfa71d9235.	453 bytes	AttributesTp.150N	AltributesTo.JSON	dc42503902e5 8080	A-
8	06/16/2222 20:15:37:867 UTC	CONTENT_MODIFIED	##3625bc c55# 4e6F6719 #258c95	305 bytes	AttributesTp.JSON	AttributesTeJSON	dc42303902e5 8080	5-
0	06/16/2022 20 15:37 865 UTC	CLONE	##1625bc c29e 4et/ 6718 #258cb1.	309 bytes	AttributesTp.JSON	AttributesTp.JS04	do-42503902e5.8080	A-

C Last updated 20.22:17 UTC

DETAILS	ATTRIBUTES	CONTENT	
Container			Container
default			default
Section			Section
98			36
Identifier			Identifier
1655308450265-98			1655308450196-36
Offset			Offset
0			3969
Size			Size
2 KB			239 bytes
LOWNLOAD	VIEW		L DOWNLOAD OVIEW
Replay			
Connection Id			
5af1cd6d-3edf-39e0-	-b1c6-33ae82e0ff4f		
~			

ОК

Easily Move Workloads to Splunk Cloud with Built-In Connectors

Moving data into a cloud-based Splunk instance can be accomplished using the native NiFi built-in processor called PutSplunk. This processor easily allows organizations to push data from their on-premise or cloud environment to Splunk instances regardless of location. By taking advantage of this capability, SIEM workloads can be straightforwardly moved to the cloud in parallel as depicted below.



The configuration of this processor must be set with the hostname of the Splunk instance. The following images display the processor and configuration screen. For this white paper, the specific Splunk values have not been included.

CANCEL

CLOUDERA

Invalid				
SETTINGS SCHEDULING	PROPER	IES COMMENTS		
Required field				
roperty		/alue		
Hostname	0	localhost		
Port	0	No value set	alloset	
Max Size of Socket Send Buffer	0	1 MB	ainost	
Idle Connection Expiration	0	5 seconds		
Timeout	0	10 seconds		
Character Set	0	UTF-8		
Protocol	0	TCP		
Message Delimiter	0	No value set		
SSL Context Service	0	No value set		

Standardize the Onboarding of Additional Data Sources

One of the significant advantages of using CDF as a UDDS is the ability to standardize how new data sources get collected and moved throughout an enterprise - build once, use many times.

By giving teams a standard approach to ingesting new data sources, organizations can streamline their process to acquire new datasets and use data in their missions or feed downstream applications. This approach decouples data sources from their destinations and brings immediate business value.



Provide Observability of Data Pipelines

Organizations require observability of their data pipelines. Observability is critical in validating chains of custody for audits and validation, but it's also essential for performance. CDF keeps a very granular level of detail about each piece of data that it ingests. According to the Apache NiFi documentation, "As the data is processed hrough the system, transformed, routed, split, aggregated, and distributed to other endpoints, an audit trail is created and stored within Apache NiFi's Provenance Repository." This implies that any and all steps that are used to process data can be stored and tracked.

We can select Data Provenance from the Global Menu. All provenance events will be listed, and the complete data provenance can be viewed by selecting the icon on the right-hand side of the event. The out-of-the-box observability capabilities of CDF allow for a comprehensive view of data pipelines, which is critical for organizations today.

The following images detail how the complete data provenance listing can be selected for all events. The last image displays all of the steps used for a particular event from the moment the data was acquired and ingested into CDF.



NiFi Data Provenance

Ø ×

Filte	b	y component name 🗸 🗸						
	Date/Time =	Туре	FlowFile Unid	Size	Component Name	Component Type	Node	
0	06/16/2022 20:49:58.3	DOWNLOAD	8efdf604-c3cc-4af2-bc	2.1 KB	No value set	NiFi Flow	dc42503902e5.8080	A
0	06/16/2022 20:49:56.6	DOWNLOAD	Sefdf604-c3cc-4af2-bc	2.1 KB	No value set	NIFI Flow	dc42503902e5.8080	a
0	06/16/2022 20:49:48.5	ATTRIBUTES_MODIFIED	5134982c-4a6f-4a30-a	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	å →
0	06/16/2022 20:49.48.5	ATTRIBUTES_MODIFIED	4a26742a-21ba-42bf-a	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	s →
0	06/16/2022 20:49:48.5	ATTRIBUTES_MODIFIED	0d248749-5ac5-4e49-a_	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	3-
-	- Octain /2022 20:49.48.5_	ATTRIBUTES_MODIFIED	6c38c412-98f5-4ecd-ab.	239 bytes	MergeContent	MergeContent	dc42503902e5:8080	3.→
0	06/16/2022 20:49:48.5	ATTRIBUTES_MODIFIED	39c521c3-9694-4598-b	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	& →
0	06/16/2022 20:49:48.5	ATTRIBUTES_MODIFIED	72b0e688-b64f-46cf-be_	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	s →
0	06/16/2022 20:49:48.5	ATTRIBUTES_MODIFIED	bcdb0b8d-c416-42dc-b	239 bytes	MergeContent	MergeContent	dc42503902e5:8080	_& →
0	06/16/2022 20:49:48.5	ATTRIBUTES_MODIFIED	b398e2ed-5ba0-437c-b	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	2-4
0	06/16/2022 20:49:48.5_	ATTRIBUTES_MODIFIED	2b71ebc4-14bd-4638-9_	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	3-
0	06/16/2022 20:49:48.5	DROP	39c521c3-9694-4598-b	239 bytes	MergeContent	MergeContent	dc42503902e5.8080	& →
0	06/16/2022 20:40:48 5	DROP	775666688.5641.46-1.54	239 butes	MerneContent	MemelContent	4+42503002+5 8080	8-

About Cloudera

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to Al. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.

Learn more at cloudera.com

Connect with Cloudera

About Cloudera: cloudera.com/more/about.html

Read our blog: blog.cloudera.com

Follow us on Twitter: twitter.com/cloudera

Visit us on Facebook: facebook.com/cloudera

See us on YouTube: youtube.com/cloudera

Join the Cloudera Community: community.cloudera.com

Read about our customers' successes:

cloudera.com/more/customers

KIFI Data Provenance

Conclusion

Government agencies and commercial entities will need to continue to address growing requirements and IT challenges. They will need solutions that are flexible enough to enable hybrid deployments, possess the ability to scale to the growing volume of data, and complement existing IT investments like Splunk. In addition to these IT challenges, they will need an universal solution that can ingest data from new sources as they come online, such as IoT devices, and deliver data to destinations, such as cloud based applications or future storage devices. Due to these factors, using CDF as an UDDS allows organizations to address the degradation in the performance of their SIEMs, and allow them to continue to meet and exceed the future needs of the mission.

Get Started Today

Wherever you are on your hybrid cloud journey, a first-class data distribution service is critical for successfully adopting a modern hybrid data stack. Cloudera DataFlow for the Public Cloud (CDF-PC) provides a universal, hybrid, and streaming-first data distribution service that enables customers to gain control of their data flows.

Take our interactive product tour: https://www.cloudera.com/products/dataflow/cdp-tour-dataflow.html

Sign up for a free trial: https://www.cloudera.com/campaign/try-cdp-public-cloud.html

CLOUDERA

Cloudera, Inc. 5470 Great America Pkwy, Santa Clara, CA 95054 USA cloudera.com

© 2022 Cloudera, Inc. All rights reserved. Cloudera and the Cloudera logo are trademarks or registered trademarks of Cloudera Inc. in the USA and other countries. All other trademarks are the property of their respective companies. cldr_title_wp_101