

Table of Contents

Troubleshooting:

Collecting Data for Troubleshooting.....	2
Accessing MDE column data taking time.....	4
End-of-day CPD health check.....	5
MDE Publish job appearing hang.....	7
Action	8

Collecting Data for Troubleshooting

CPD platform related problems:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI

Connectivity related issues:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI (CCS)
Application logs from pod	oc cp <wdp-connect-connection pod>/logs ./<wdp-connect-connection pod>-logs oc cp <wdp-connect-connector pod>/logs ./<wdp-connect-connector pod>-logs

Data Class related issues:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI (CCS, WKC, AE)
Application logs from swagger	https://<CPD_URL>/v2/data_profiles/api/explorer/#/Hummingbird%20tasks/getHbTaskLogs

Global Search related issues:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI (CCS, WKC)
Pod logs	catalog-api
Application logs from swagger	https://<CPD_URL>/v2/cams/explorer/#/Assets/retrieveAssets

Metadata Enrichment related issues:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI (CCS, WKC, AE)
Application logs from pod	oc cp <wdp-profiling pod>/logs ./<wdp-profiling pod>-logs oc cp <spark-hb-control-plane pod>/logs ./<spark-hb-control-plane pod>-logs
Job log from UI	Job log for MDE job & Screenshot of the job start time
Application logs from swagger	https://<CPD_URL>/v2/data_profiles/api/explorer/#/Hummingbird%20tasks/getHbTaskLogs

Metadata Enrichment related issues:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI (CCS, WKC)

WKC Reporting related issues:

Data Needs to Capture	
Diagnostic job	Gather diagnostics information from CPD UI (CCS, WKC)
Reporting status from API	<pre>curl -i -k -H "content-type: application/json" -H "Authorization: bearer \$Bearer_TOKEN" -X GET "https://\$HOSTNAME/v3/reporting/heartbeat"</pre> <pre>curl -i -k -H "content-type: application/json" -H "Authorization: bearer \$Bearer_TOKEN" -X GET "https://\$HOSTNAME/v3/reporting/999/register"</pre> <pre>curl -i -k -H "content-type: application/json" -H "Authorization: bearer \$Bearer_TOKEN" -X GET "https://\$HOSTNAME/v3/reporting/bistatus?tenant_id=999&table_name=all"</pre>

OpenShift related issues:

Data Needs to Capture	
Must-gather from CLI	cc adm must-gather
Collect data for projects from CLI	<pre>oc adm inspect/namespace <CPD namespace></pre> <pre>oc adm inspect/namespace <IBM CPD operator namespace></pre>

OpenShift Data Fusion related issues:

Data Needs to Capture	
ODF must-gather from CLI	<pre>oc adm must-gather --image=registry.redhat.io/odf4/ocs-must-gather-rhel8:v4.12 --dest-dir=/tmp/odf-must-gather</pre> <p>or</p> <pre>oc adm must-gather --image=docker-virtual.oneartifactoryci.verizon.com/odf4/ocs-must-gather-rhel8:v4.12</pre>

IBM Storage Fusion related issues:

Data Needs to Capture	
Collect logs from Fusion console	Fusion Web Console > “Help icon” > Support logs > Collect logs.

Accessing MDE column data taking time

24122_DR_MDE_V1_PP2

Columns (13,724)	Asset	Business terms	Data class	Data quality	Review status
CHANGEDATE	XI_EDVERSION GFWV_24122_DisasterRecovery_C				
CHANGEDATE	XI_REPVERSION GFWV_24122_DisasterRecovery_C				
CHANGEDATE	XI_RUNVERSION GFWV_24122_DisasterRecovery_C				

CHANGEDATE

Columns from XI_RUNVERSION

Go to asset in project

Details Governance

Description

Source
GFWV_24122_DisasterRecovery_C / SAPS...

Troubleshooting:

1) For any CPD UI issue a starting point is the browser developer tool.

Name	Status	Type	Initiator	Size	Time	Waterfall
?size=24&scaleFactor=1x&showFallbackMonogram=&page...ps%3A%2F%2Fchrome.google...	200	png	chrome://res...	994 B	2 ms	
db925b85-ed34-49e3-a5f4-8cc8fa9e6719?project_id=36...225-4190-b60c-a072fed121ab&co...	500	xhr	datacatalog.v...	712 B	853 ms	
db925b85-ed34-49e3-a5f4-8cc8fa9e6719?project_id=36...225-4190-b60c-a072fed121ab&co...	304	xhr	datacatalog.v...	316 B	402 ms	
db925b85-ed34-49e3-a5f4-8cc8fa9e6719?project_id=36...225-4190-b60c-a072fed121ab&co...	500	xhr	datacatalog.v...	620 B	6.03 s	
3d87ea75f01484b48e97.svg	200	svg+xml	datacatalog.v... (disk cache)		2 ms	
tokenexpiry	200	xhr	datacatalog.v...	613 B	132 ms	

- Mainly the “Network” and “Console” panels are commonly used for investigating network activities, API call failure and error message returned by a URL.
- Focused on the HTTP return/status code and elapsed time to find the problematic API. Compare elapsed time with previous successful run.

Name	Headers	Payload	Preview	Response	Initiator	Timing
?size=24&scaleFactor=1x&showFallbackMonogram=&...	<p>Request URL: https://datacatalog.verizon.com/gov/metadata-enrichments/api/data-class-assignment/db925b85-ed34-49e3-a5f4-8cc8fa9e6719?project_id=36d882a5-f225-4190-b60c-a072fed121ab&column_name=access_amt</p> <p>Request Method: GET</p> <p>Status Code: 500 Internal Server Error</p>					

- The URL in the “Network > Headers” tab can help to identify the CPD service or component having problem.

Request URL: /gov/metadata-enrichments/api/data-class-assignment/

- Service: gov/WKC/IKC
- Component: metadata-enrichments/data-class-assignment

2) Check [pod reference table](#)^[A6] to find pods associated to the Service/Component.

- Are all pods healthy? ^[A1]
- Any recent warning/fail event associated to the pod? ^[A2]
- Has pod restarted recently? ^[A3]
- Any error in the pod log during the problem? ^[A4]
- What is memory, CPU consumption of pod? ^[A5]

End-of-day CPD health check

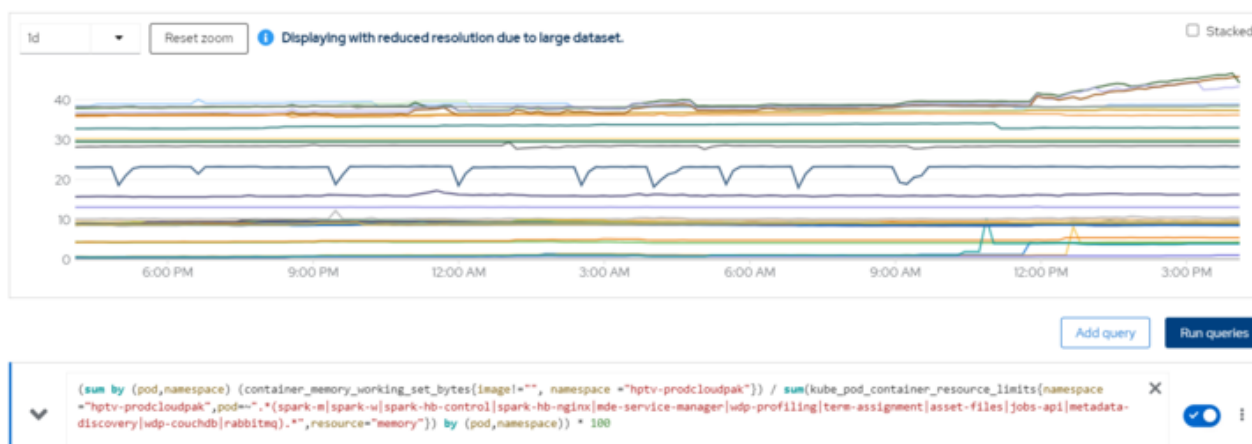
Process of identifying potential issue:

- 1) What are you going to run over night?
 - a. MDI/MDE/Publishing jobs
- 2) What are pods associated to these operations?
 - a. Check [pod reference table](#)^[A6] to find pods associated to the Service/Component.
- 3) Find status of pods using command line or CPD monitor console.
 - a. Are all pods healthy? ^[A1]
 - b. Any pods restarted recently within last 24 hours? If yes, investigate why in restart and take corrective action. ^[A3]

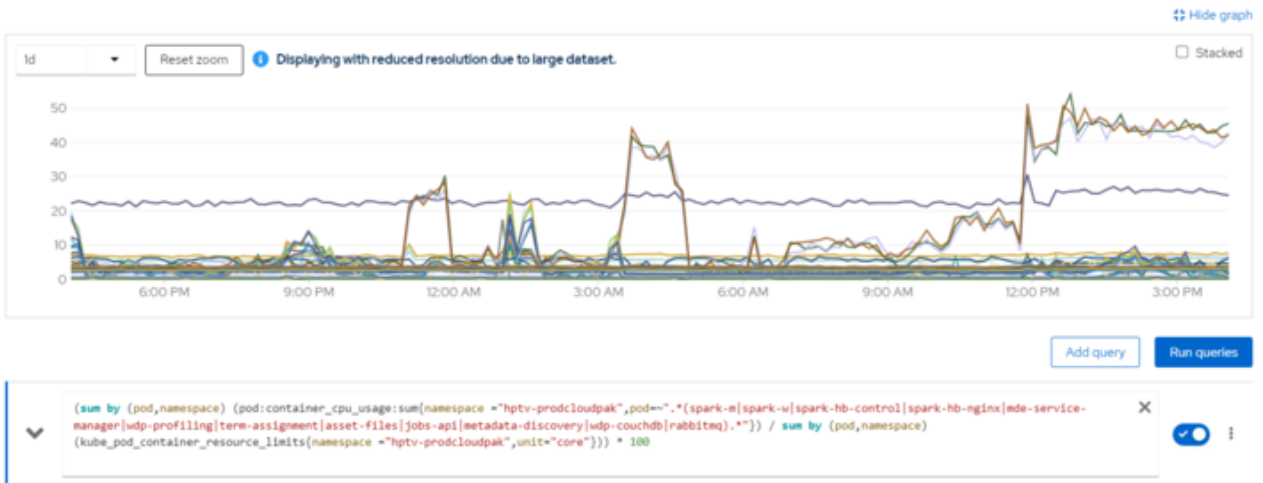
```
# oc get pods | egrep 'spark-m|spark-w|spark-hb-control|spark-hb-nginx|mde-service-manager|wdp-
profiling|term-assignment|asset-files|jobs-api|metadata-discovery|wdp-couchdb|rabbitmq'
```

asset-files-api-7b74f7d4cf-82hzb	1/1	Running	1 (4d4h ago)	4d23h
asset-files-api-7b74f7d4cf-b4xjj	1/1	Running	2 (3d10h ago)	4d23h
asset-files-api-7b74f7d4cf-dgbmz	1/1	Running	0	4d23h
asset-files-api-7b74f7d4cf-hxtpq	1/1	Running	2 (38h ago)	4d23h
asset-files-api-7b74f7d4cf-lr2t8	1/1	Running	1 (4d6h ago)	4d23h
asset-files-api-7b74f7d4cf-z5nqc	1/1	Running	2 (27h ago)	4d23h
jobs-api-6484667d79-5xjv8	1/1	Running	0	5d23h
metadata-discovery-68bfc986dd-gd8zr	1/1	Running	0	5d19h
metadata-discovery-68bfc986dd-rxt7d	1/1	Running	0	5d19h
metadata-discovery-68bfc986dd-w67fn	1/1	Running	0	5d19h
rabbitmq-ha-0	1/1	Running	0	5d3h
rabbitmq-ha-1	1/1	Running	0	5d3h
rabbitmq-ha-2	1/1	Running	0	5d3h
rabbitmq-ha-3	1/1	Running	0	5d3h
spark-hb-control-plane-59759fcd6-89678	2/2	Running	0	8d
spark-hb-nginx-85f7c6995d-6hbjb	1/1	Running	0	5d18h
wdp-couchdb-0	2/2	Running	0	5d21h
wdp-couchdb-1	2/2	Running	0	5d21h
wdp-couchdb-2	2/2	Running	0	5d21h
wdp-profiling-745bd5897b-4vhs4	1/1	Running	0	4d23h
wdp-profiling-745bd5897b-bj262	1/1	Running	0	4d23h

- c. What is memory, CPU consumption of pod over last 24 hours? Measure the pods resource usage from OpenShift console using metrics query. Current resource usage below 40% will be ideal.



CPD/IKC – Knowledge Sharing



MDE Publish job appearing hang

1. Check the job log for any error.
2. Monitor the job log for any progress in statistical values.
3. Check [pod reference table](#)^[A6] to find pods associated to publishing
 - a. Are all pods healthy? ^[A1]
 - b. Any recent warning/fail event associated to the pod? ^[A2]
 - c. Has pod restarted recently? ^[A3]
 - d. Any error in the pod log during the problem? ^[A4]
 - e. What is memory, CPU consumption of pod? ^[A5]

Action

A1. Are all pods healthy? If not:

- [Investigate pod issues](#)
- Try restart the pod

A2. Any recent warning/fail event associated to the pod? If yes:

- [Get an understanding of the event](#) and take corrective action.

A3. Has pod restarted recently? If yes:

- Need to investigate if pod restarted multiple times.
- Check pod description to find the reason of restart.
- Capture the previous pod log. Look for any error message at the end of log to figure out the cause of restart.
- Search exiting Jira for similar error message.
- If need help open support ticket with IBM along with all necessary diagnostics data.

A4. Any error in the pod log during the problem? If yes:

- Search exiting Jira for similar error message.
- If need help open support ticket with IBM along with all necessary diagnostics data.

A5. What is memory, CPU consumption of pod?

- Check the pod description for current resource configured.
- Check the pod description for restart due to lack of resource.
- Using OpenShift console monitor pod resource usage for a period.
- Check with IBM support if pod restarted due to lack of resource.

A6. Pod reference table

- https://github.com/sanjitc/Cloud-Pak-for-Data/blob/main/wkc/troubleshooting/Pods_Making_CPD.md