DataStage Health Check Assessment for

UWV, Netherlands

Srikrishna Uppuluri
IBM Expert Labs
Email: sriuppul@in.ibm.com

Version 1.0

Abstract: This document reports the findings of a health check of UWV's DataStage implementation conducted on-site in the week of 4th to 15th December 2023. Because of the complex scope of the DataStage implementation, only critical components experiencing recent issues were reviewed. Recommendations for improving the DataStage implementation in terms of resilience, performance and functionality were made in conformance with best practices

Date	Version	Description	Author
18/12/23	1.0	Initial version	Srikrishna Uppuluri

Table of Contents

Purpose of this document

The primary scope of this assessment is to review existing IBM Infosphere Information Server 11.7.1.x (aka IIS) deployment in Production and provide recommendations with a view on “Best Practices” on IIS 11.7.1 deployment

This review covers all aspects of IIS and DataStage which includes Architecture, Configurations and deployment of DIM application

Methodology

The information gathered is based on discussions with UWV DataStage team , documentation provided, and technical evaluations of system configurations. In all areas of review, the facts observed in running system configuration take precedence over facts obtained via documentation and interviews.

Finally, observations, findings and recommendations were formatted into this report

Environment findings and recommendations

IIS 11.7. Architecture

- IIS 11.7.1.3 is deployed on 4-tier architecture with dedicated servers for each tier
- Tier deployed
    - Services
    - Engine
    - Repository
    - Microservices
- Primary component used is DataStage to do DIM application processing
- IGC is used for data lineage
- Governance and Metadata management components have less utilization.

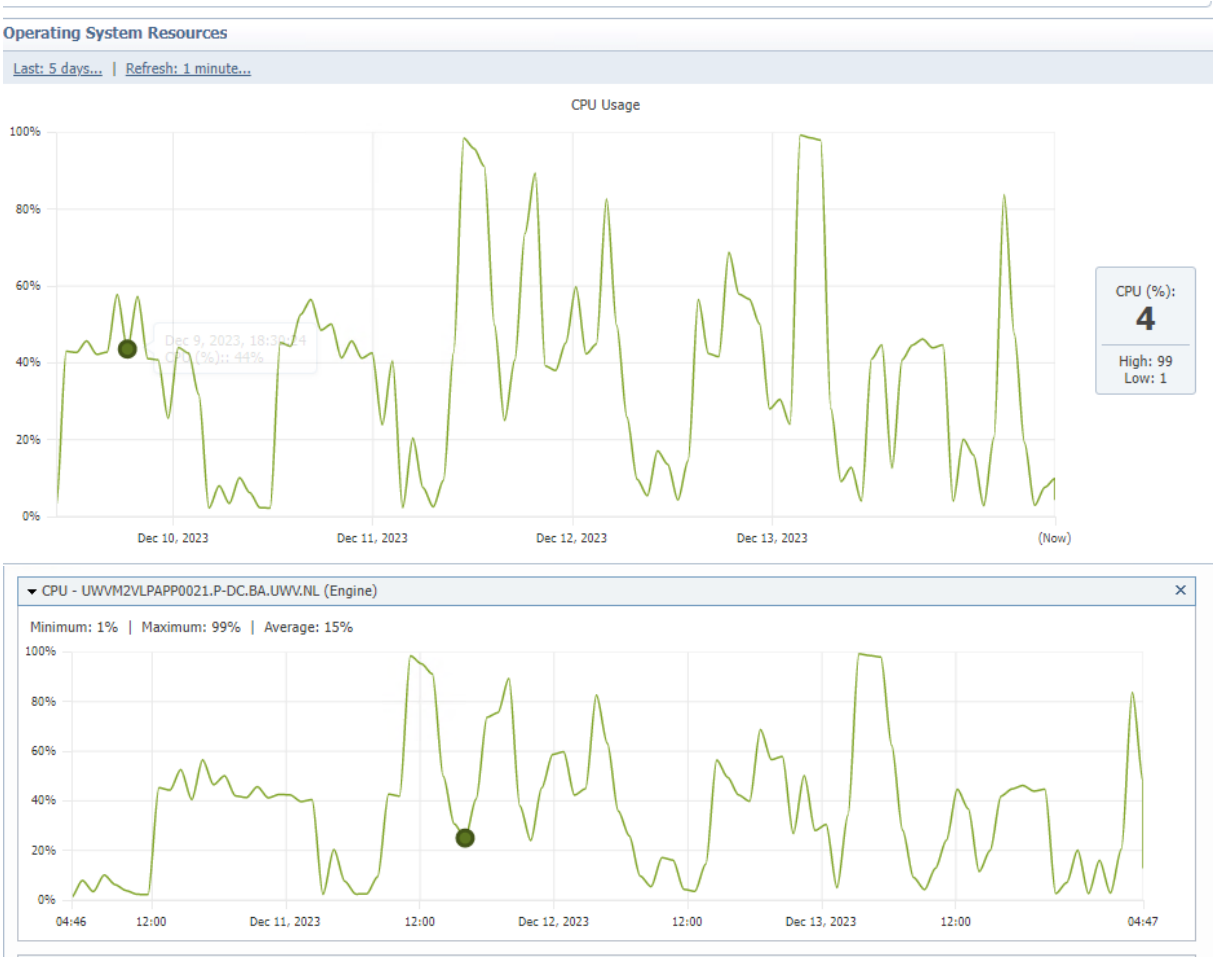
Environment and Workload

- Approximately ~40,000 jobs run in a 24 hour period, indicating an extremely high workload environment

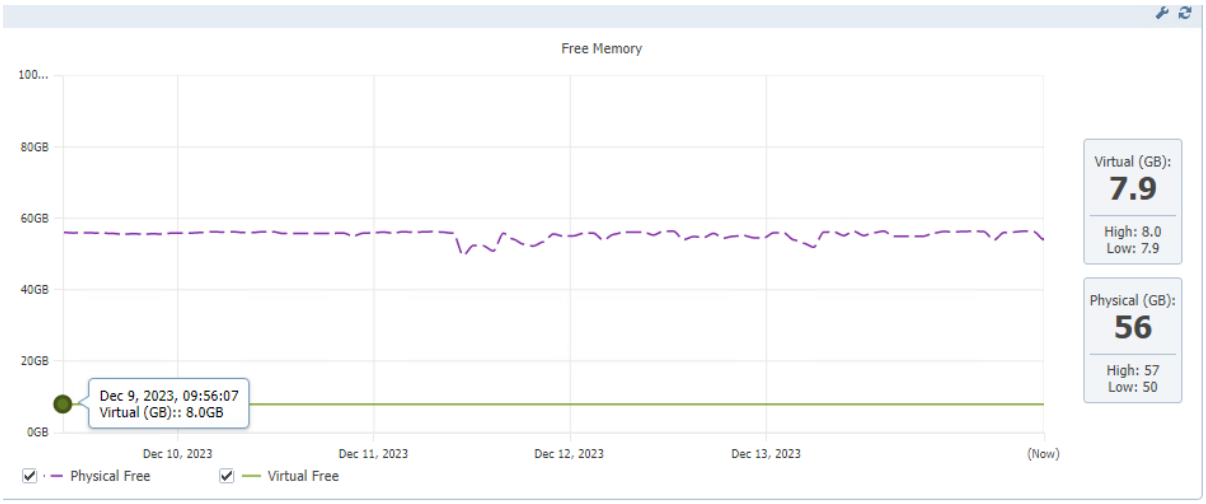
- The job failures(0.035%) are low and primarily due the data/file availability.
- A high concurrency of ~60 jobs at certain times. Most of the time, the average concurrency is ~30 jobs.

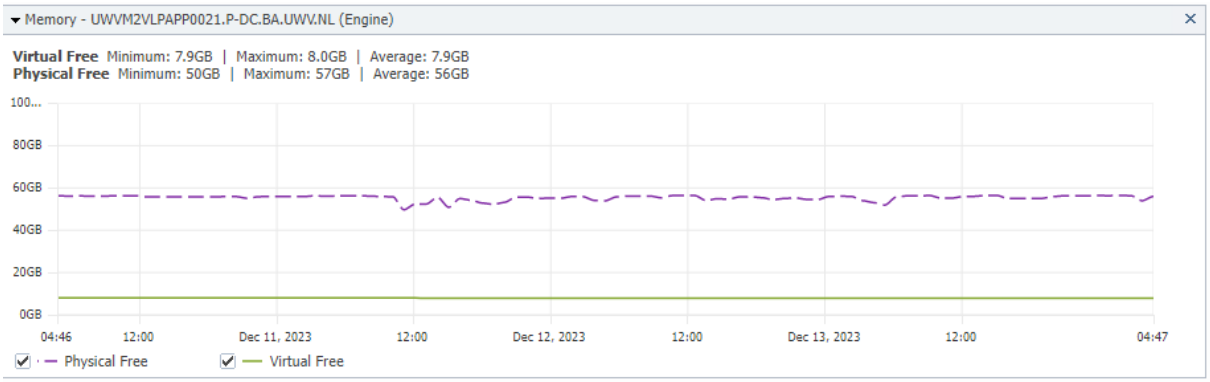
For any change in the workload, uvconfig tunables and WLM settings must be re-evaluated

- CPU utilization is 15% on average during a 5-day period. There are occasions when the CPU is completely utilized. The highest utilization corresponds to the number of jobs submitted at that moment. If possible, plan jobs before or after high usage times to ensure that all processes have consistent CPU cycle availability.

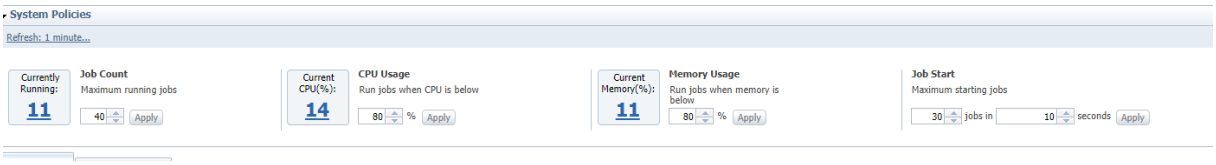


- Memory utilization is low. There were no problems observed.





- WLM settings.



ISALite General Health Check Results

Following are the findings from the ISALite reports. Majority of the fundamental checks are successful.

- shared_hot directory 87% full

```
/dev/mapper/vg_shared-lv_grid          515924224      73808  489619840   1% /w001/shared_grid
/dev/mapper/vg_shared-lv_scratch      1056754100      90432 1002960400   1% /w001/shared_scratch
uwvm2vlpapp0024:/w001/shared_hot      2113642496 1745155072  261097472  87% /w001/shared_hot      <===  << Warning: Disk usage is higher than 80% >>
uwvm2vlpapp0023:/w001/shared_cold    10652618752 3055389696 7060341760  31% /w001/shared_cold
tmpfs                                    6578952          0    6578952   0% /run/user/180805478
```

  A 90% full disk can cause a variety of concerns and potential problems such as performance degradation, fragmentation, insufficient space causing errors in job executions, and so on

- Duplicate terms in a unique category path. Glossary author should review the listed terms in the report and remove any duplicates.

- uvconfig tunables (discussed in the specific section)

Usage on multiple Information Server/DataStage interfaces

The list provides a view on the usage of different interfaces. Information is derived from the savailable logs(~ 15 day period) on the server

```
    685  Client="Command client"
    184  Client="DataStage Administrator"
     63  Client="DataStage Designer"
    144  Client="DataStage Director"
    786  Client="DataStage Operations Console"
     24  Client="DataStage Server"
      6  Client="DataStage Server CLI"
     98  Client="IGC"
      3  Client="Information Analyzer Validation"
     55  Client="Information Analyzer Web Interface"
     64  Client="InfoSphere Metadata Asset Manager Command-Line Interface"
     71  Client="InfoSphere Metadata Asset Manager Web Interface"
     24  Client="ISALite"
 122152  Client="Server client"
    169  Client="Web Console"
```

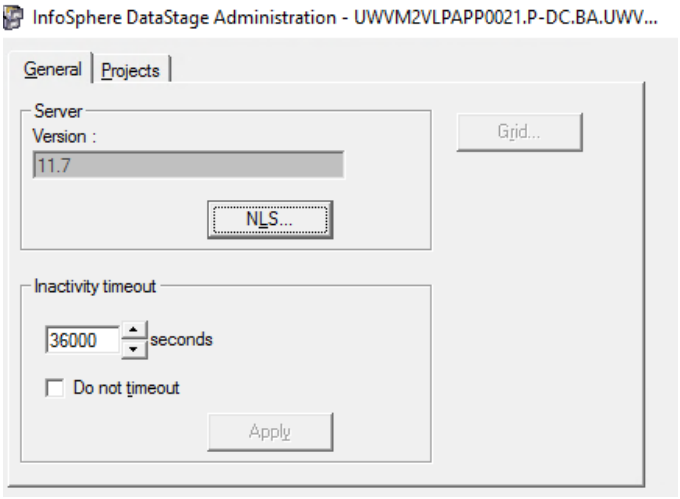
The operations console is used the most to monitor jobs, followed by DataStage Client interfaces.

All of these sessions will be added to the global sessios. Having an inactive timeout, even a large value, in the global session settings and in DataStage settings will help maintain the system health by removing the unused/leftover sessions. If there are too many orphan processes left, DataStage may become unstable and some of the processes may start to utilize a lot of CPU

Recommendation implemented

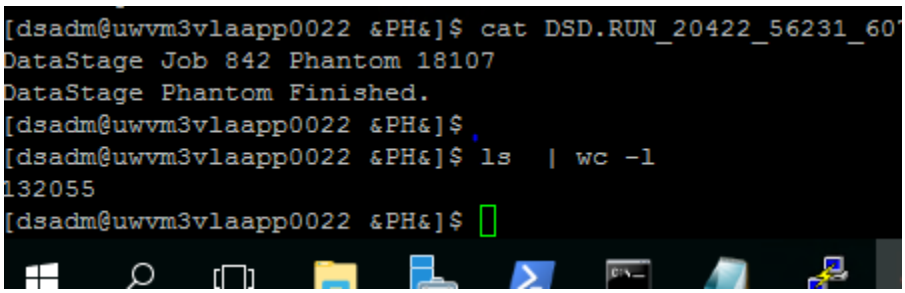
- Session Timeout

  The “Do not timeout” option has been changed to 36000 seconds. It can be increased to 86400 or higher for jobs running more than a day



- Clean-up of &PH& directory

  Around 130,000 files accumulated in the &PH& folder of DIM projects on ACC and PROD servers. If files are not automatically removed by the engine process, this directory must be cleaned on a regular basis.



- uvconfig tunables

  In the weeks leading up to the review, few jobs started failing with “code=-14 [Timed out while waiting for an event]” error. Recurrence of the issue was reduced with disabling of the vmotion. However, the jobs occasionally failed with same error.

  Following the below technote

https://www.ibm.com/support/pages/datastage-job-fails-code-14-timed-out-while-waiting-event/

- DSWaitResetStartup and DSWaitStartup has been increased to 600 at instance level
- Below uvconfig paramters are tuned
  MFILES 400
  T30FILE 1024
  RLTABSZ 600
  MAXRLOCK 599
  DMEMOFF 0xb0000000
  PMEMOFF 0xb31ee000
  CMEMOFF 0xb31f1000
  NMEMOFF 0xb33f2000

  GLTABSZ is still having a value of 75.It can be increased to 600 to make it same as RLTABSZ.

  Following technote will provide a view on calculation if the tunables

  https://www.ibm.com/support/pages/node/484687

- DataStage projects on NFS mount point
  DataStage project folders were created on the NFS shared location in all of the environments. For non-clustered setups, it is always preferable to have the project folders and installation binaries on the local disk.

DIM Application Review

Staging overview:
A total of 4 staging frameworks (source to staging)
- PlatBestandharnas (File format)
  Location in Designer: 10 Staging/ 50 Jobs/DL_Bestand_Verwerking

- Dblink harnas (Dblink format)
  Location in Designer: 10 Staging/ 50 Jobs/DL_DatabaseLink_Verwerking

- Datapump harnas (Datapump format)
  Location in Designer: 10 Staging/ 50 Jobs/DL_Datapump_Verwerking

- Platbestandharnas Resafasa (File format – only for source Resafasa)
  Location in Designer: 10 Staging/ 50 Jobs/DL_Resafasa_Verwerking

ODS Overview

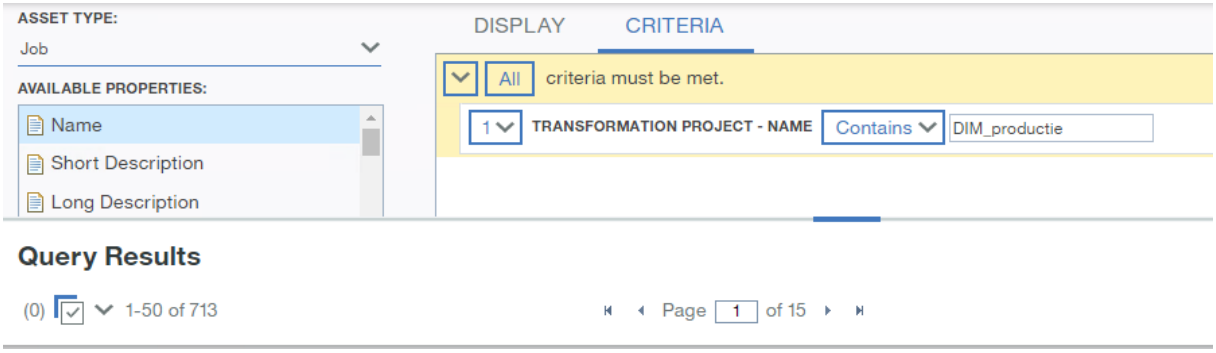
bronzone (ODS) harnas: Staging to ODS
- Bronzoneharnas (for all type of deliveries, incremental, full load, stackable)
  Location in Designer: 20 Bronzone / 50 Jobs/

- BronzoneStandharnas (can only be used for full load)
  Location in Designer: 20 Bronzone /50 Jobs/B Alternatieve Bronzone Verwerking

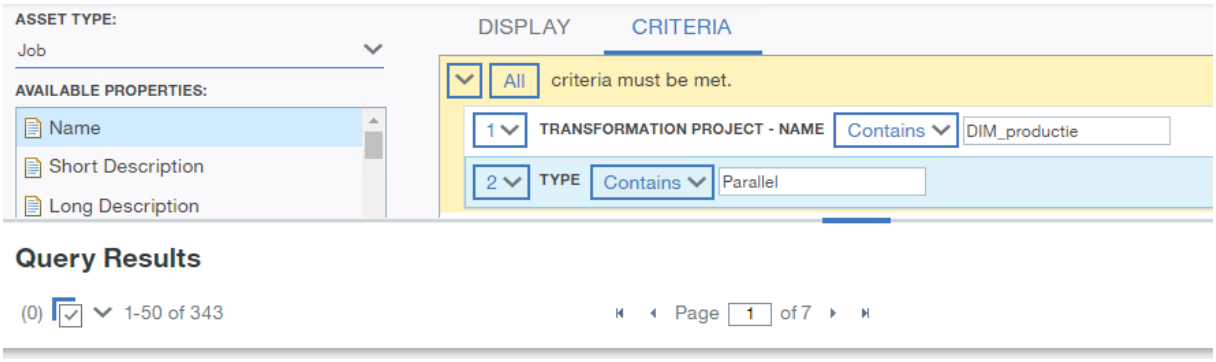
Production project and job statistics

Project	DIM_productie
Total Jobs	713
Parallel Jobs	343
Sequence Jobs	280
Server Jobs	90
UWV Routines	4

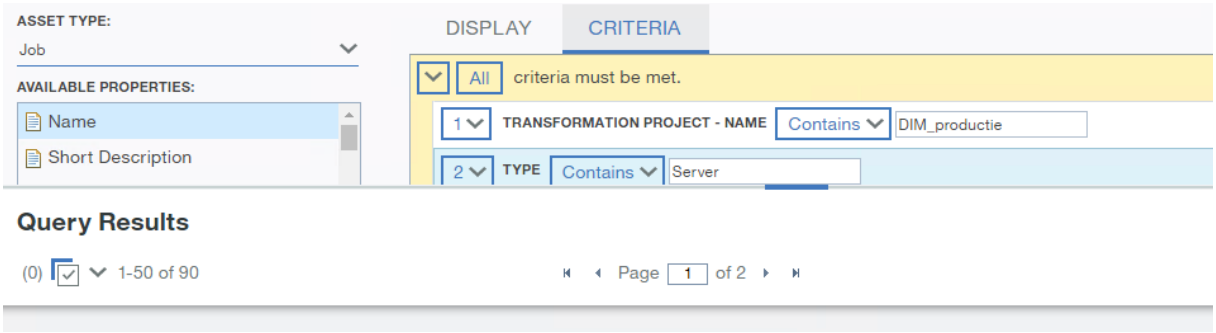
Total Jobs:



Parallel Jobs:

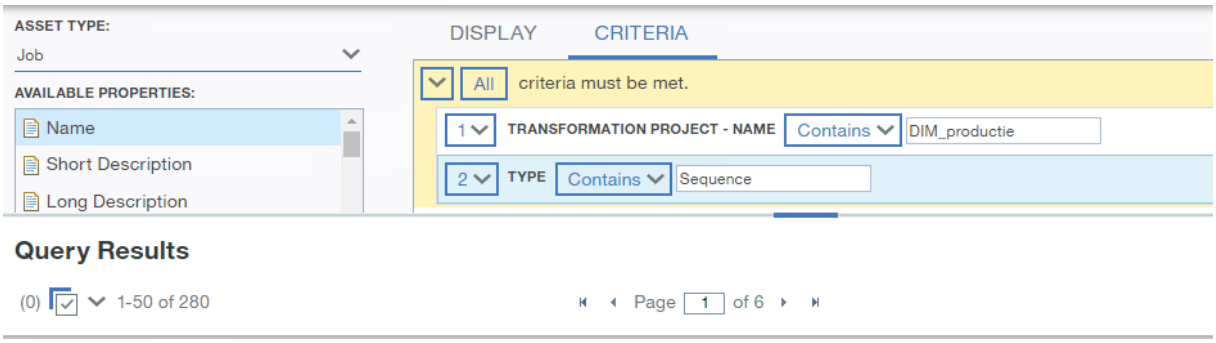


Server Jobs:

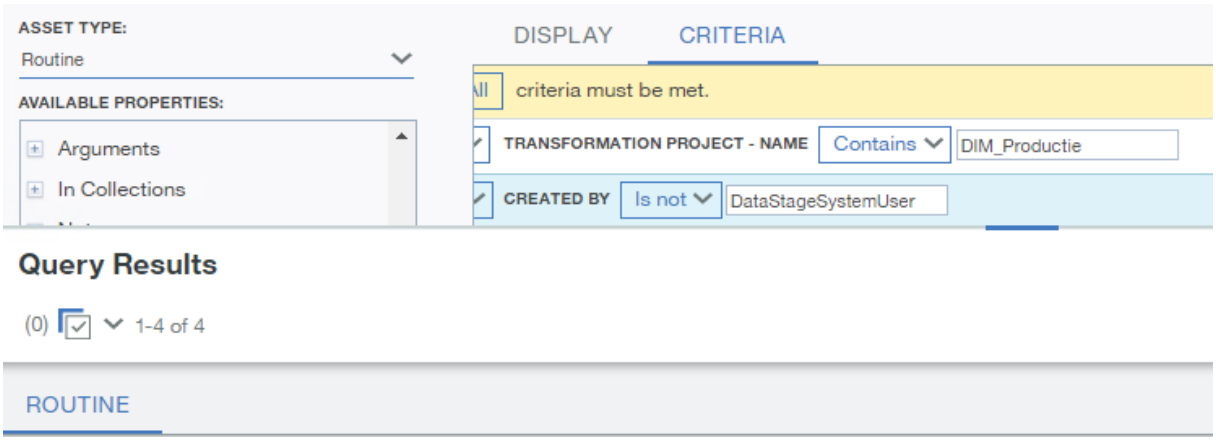


Sequence Jobs

Version 1.0



Routines developed by UWV:



DIM Observations

Due to the complexity of the DIM Application's implementation, only Staging and a partial review of ODS capability were possible during the assessment.

- Majority of the jobs in staging and ODS extensively use RCP allowing the jobs to run for multiple sources

  For the majority of staging flows, the source definitions, schema files are created/generated on each execution, even if they are executed on the same day. Staging tables are dropped and re-created every day based on the schema files. Recommend to re-design/optimize to generate new schemas only when source file structure changes

- Oracle statements from files are extensively used in staging and ODS areas. In the Oracle connector, many jobs refer to the sql statements file as before or after sql. Job executions (details in the following section) reveal that some of the long-running jobs are the ones invoking the sql statement. Some of the transformation logic can be migrated to DataStage to take use of the partitioning and parallelism capabilities.

- Most of the jobs are not configured use the Oracle Connector parallel read capability. Using parallel read in Oracle improves the read performance , which is often considered as a bottleneck (Details in the sample job tuning section)

- Order by clauses are used in several select queries in Oracle connector stages.. The downstream stages are configured to 'Don't Sort' which is good design practice. But performing the sort in parallel outside of the database removes the sequential bottleneck of sorting within the database

- DataStage director only supports time based triggers. If the flow needs to maintain dependency across several systems, it is recommend to continue using IWS scheduler.

- The project contains approximately 80 server jobs. It is recommended to convert server jobs to parallel jobs in order to take advantage of DataStage parallel features

A more thorough assessment of jobs is advised in order to optimize batch design , identify areas of performance improvement.

Operations Database insights

The operations database's data can offer valuable information into the behaviour and execution of jobs. A good place to start identifying long-running jobs and patterns.

The following findings are based on job run data from 2023-Nov-1 to 2023-Dec-15. All of the results are based on parallel and server job executions. Sequence jobs were not considered since the elapsed time is roughly equal to the total of the elapsed times of the individual jobs.

1) The table shows the elapsed times classified by duration.
   ~97% of the times, jobs complete in <1 min
   ~1100 times, jobs ran more than 30 nin

1	LessThan-1MIN	461641
2	LessThan-2MIN	6580
3	LessThan-5MIN	3227
4	LessThan-10MIN	1250
5	LessThan-30MIN	708
6	LessThan-60MIN	257
7	Greaterthan-2HR	84
8	LessThan-2HR	75

2) List of jobs and the number of times the execution time exceeds 2 hours

	JOBNAME	COUN...
1	job_DL_RunSQLStatement	48
2	job_DL_LaadStgHub	12
3	job_DL_RunStatement	5
4	job_DL_LaadStgSrcStand	4
5	job_DL_RunImport	4
6	job_RunDeelBackupBrz	4
7	job_DL_LaadBronzoneTabel	3
8	job_DL_Laad_DeltaIncrGeenBronStatus_Opt	1
9	job_DL_LaadResafasaPakbon	1
10	job_URS_Bepaling_DatumRecht	1

3) List of jobs and the number of times the execution time exceeds 30 minutes

	JOBNAME	COUN...
1	job_DL_RunSQLStatement	96
2	job_DL_LaadStgSrcStand	49
3	job_INT_PEER_ANTWOORD_HT	40
4	job_DL_LaadBronzoneTabelStatus	34
5	job_BDR_PEER_ANTWOORD	34
6	job_DL_LaadStgHub	28
7	job_DL_RunStatement	21
8	job_DL_LaadBronzoneTabel	18
9	job_RunDeelBackupBrz	17
10	job_DL_RunImport	13
11	job_DL_Laad_DeltaIncrGeenBronStatus_Opt	10
12	job_Deelnemer_Laden_Eerdere_Levering	7

4) List of jobs, invocation id and the number of times the execution time exceeds 30 minutes

	JOBNAME	INVOCATIONID	COUN...
1	job_DL_LaadBronzoneTabelStatus	117_wwf0118d	29
2	job_INT_PEER_ANTWOORD_HT	PEER_PO	20
3	job_INT_PEER_ANTWOORD_HT	PEER_PM	20
4	job_BDR_PEER_ANTWOORD	PEER_PM	17
5	job_BDR_PEER_ANTWOORD	PEER_PO	17
6	job_Deelnemer_Laden_Eerdere_Levering	-	7
7	job_DL_LaadStgSrcStand	r0067_vast_10	6
8	job_DL_RunStatement	100_wwo_303	6
9	job_DL_LaadStgSrcStand	r0098_cat_vast_10	6
10	job_DL_LaadStgSrcStand	r0059_res_vast_10	6
11	job_DL_RunImport	109_uzs	6
12	job_DL_LaadStgSrcStand	r0015_bt_vast_10	6
13	job_DL_LaadStgSrcStand	r0066_uitk_vast_10	6
14	job DL RunSQLStatement	uzz toeknnngn 6	5

The elapsed time, invocation id, and occrence can be used to identify likely candidates for performance tuning. The initial list might include the top ten jobs in each category.

5) Usage of configuration file in project. Majority of the jobs execute with default configuration of 2 nodes. There are no jobs that use average.apt(4 node) configuration file

CONFIGFILENAME	COUNT(1)
/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	200
/w001/app/ibm/as1/InformationServer/Server/Configurations/high.apt	86
/w001/app/ibm/as1/InformationServer/Server/Configurations/single.apt	9

6) Jobs with elapsed time for more than 30min and configuration file usage

	JOBNAME	CONFIGFILENAME	COUN...
1	job_DL_RunSQLStatement	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	96
2	job_DL_LaadStgSrcStand	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	49
3	job_INT_PEER_ANTWOORD_HT	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	40
4	job_BDR_PEER_ANTWOORD	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	36
5	job_DL_LaadBronzoneTabelStatus	/w001/app/ibm/as1/InformationServer/Server/Configurations/single.apt	34
6	job_DL_LaadStgHub	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	28
7	job_DL_RunStatement	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	21
8	job_RunDeelBackupBrz	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	18
9	job_DL_LaadBronzoneTabel	/w001/app/ibm/as1/InformationServer/Server/Configurations/high.apt	18
10	job_DL_RunImport	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	13
11	job_DL_Laad_DeltaIncrGeenBronStatus_Opt	/w001/app/ibm/as1/InformationServer/Server/Configurations/default.apt	10

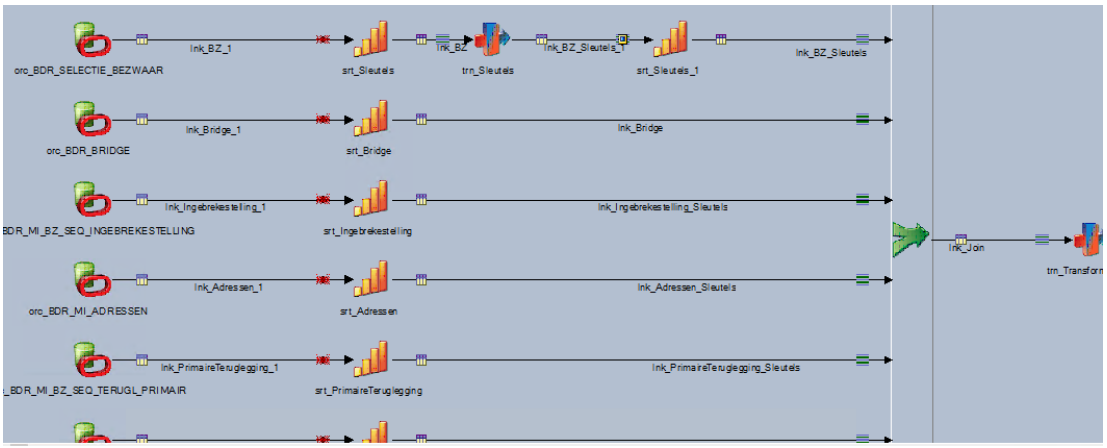
Jobs using defaults configuration file and having higher elapsed time can be run with average.apt (4 node) to improve performance.

Only the jobs having transformation logic in DataStage can be tried for performance improvement using configuration changes. Some of the jobs directly run the sql statements from a file as before and after sql statements . To improve performance of those jobs, either the queries in the sql statements needs to be tuned or sql logic needs to be migrated to DataStage if applicable

Sample Performance Tuning of long running job in Acceptance

One of the observation in Acceptance environment is the job job_Bdr_BezwaarSelectie running for more than 50 hours. The job was doing a select of 54 months of data and joining with multiple tables.

Iteration-1



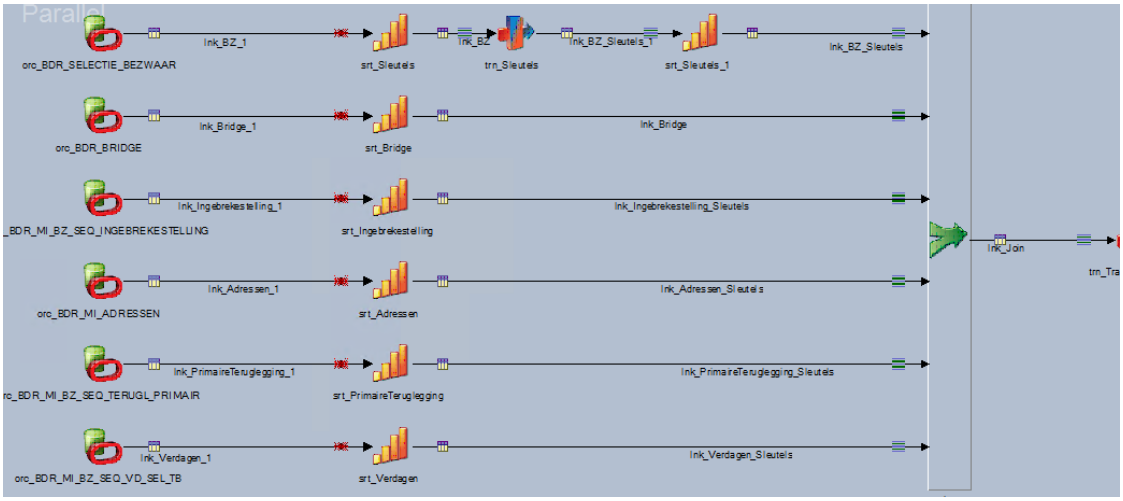
The query in the oracle connector ora_BDR_SELECTIE_BEZWAAR is having an order by clause on 1.3B rows for 54 months. Though the subsequent sort stages are re-using the sort order, the query took long time to produce the complete result.

For tuning , the job design is copied to fjob_Bdr_BezwaarSelectie_No_Target without the target stage and ran for one month of data to benchmark the elapsed time. Elapsed time is ~24-25 min (in multiple runs)

11:37:48	14-12-2023	Info	orc_BDR_SELECTIE_BEZWAAR,0: Oracle client version: 12.2.0.1.0
11:37:48	14-12-2023	Info	orc_BDR_MI_BZ_SEQ_INGEBREKESTELLING,0: Oracle client version: 12....
11:37:48	14-12-2023	Info	orc_BDR_MI_ADRESSEN,0: The connector connected to Oracle server A_...
11:37:49	14-12-2023	Info	orc_BDR_MI_ADRESSEN,0: Oracle client version: 12.2.0.1.0
11:37:54	14-12-2023	Info	orc_BDR_MI_BZ_SEQ_INGEBREKESTELLING,0: Number of rows fetched ...
11:38:55	14-12-2023	Info	orc_BDR_MI_BZ_SEQ_TERUGL_PRIMAIR,0: Number of rows fetched on th...
11:40:42	14-12-2023	Info	orc_BDR_MI_BZ_SEQ_VD_SEL_TB,0: Number of rows fetched on the curre...
12:01:45	14-12-2023	Info	orc_BDR_MI_ADRESSEN,0: Number of rows fetched on the current node: 1...
12:01:48	14-12-2023	Info	orc_BDR_BRIDGE,0: Number of rows fetched on the current node: 28879600.
12:01:49	14-12-2023	Info	orc_BDR_SELECTIE_BEZWAAR,0: Number of rows fetched on the current ...
12:01:49	14-12-2023	Info	main_program: Step execution finished with status = OK.
12:01:49	14-12-2023	Info	main_program: Startup time, 0:02; production run time, 24:00.
12:01:49	14-12-2023	Info	Parallel job reports successful completion
12:01:49	14-12-2023	Control	Finished Job fjob_Bdr_M_BezwaarSelectie_NoTarget.

>Job name	Status	Started	On date	Last ran	On date	Elapsed time
fjob_Bdr_M_BezwaarSelectie_NoTarget	Finished	11:37	14-12-2023	12:01	14-12-2023	00:24:05

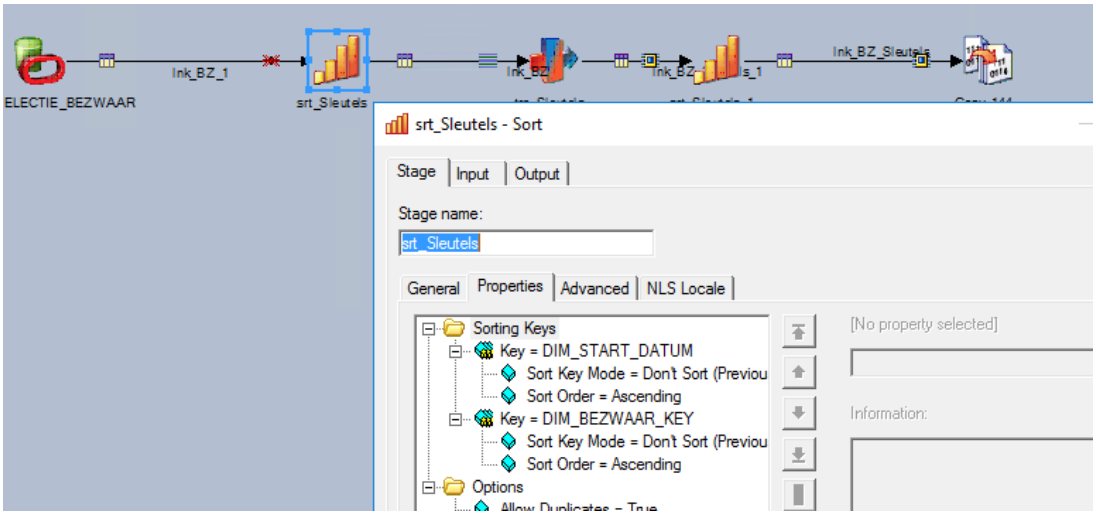
Created a new job fjob_Bdr_BezwaarSelectie_No_Target_Tuned with tuning in Oracle connector and sort stages. Elapsed time is ~13-14 minutes. 40% improvement



Job name	Status	<Started	<On date	<Last ran	<On date	Elapsed time
fjob_Bdr_M_BezwaarSelectie_NoTarget_Tuned	Finished	08:26	15-12-2023	08:40	15-12-2023	00:14:09

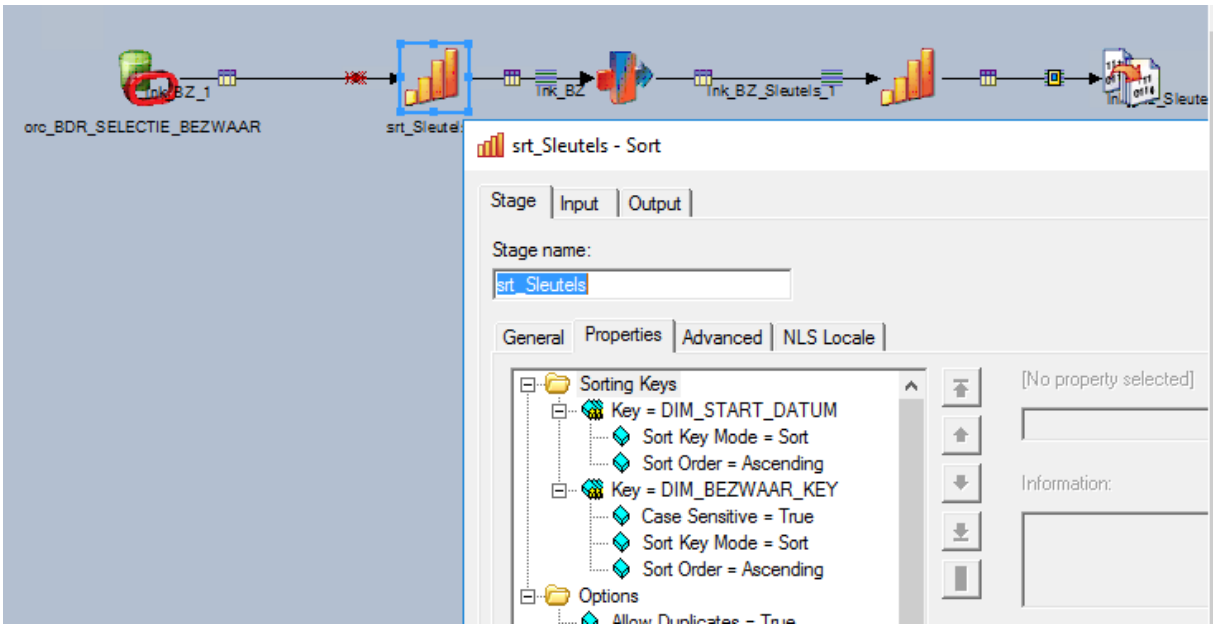
Iteration-2

Sample job with oracle order by clause and sequential read : Elapsed time of ~16-17min



fjob_Bdr_M_BezwaarSelectie_NoTarget_Src	Finished	08:45	15-12-2023	09:02	15-12-2023	00:16:40	B&

Modified job with Oracle parallel read and DataStage sort : Elapsed time of ~9-10 min



Job name	Status	<Started	<On date	<Last ran	<On date	Elapsed time	Description
fjob_Bdr_M_BezwaarSelectie_NoTarget_Tuned_Src	Finished	09:02	15-12-2023	09:12	15-12-2023	00:09:43	B&B bezwaa

Reference links for design best practices

The URL provided below can help comprehend the best practices in job design

https://www.ibm.com/docs/en/iis/11.7?topic=reference-job-design-tips

https://www.ibm.com/docs/en/iis/11.7?topic=tips-designing-good-performance

https://www.ibm.com/docs/en/iis/11.7?topic=reference-improving-performance