**IDRP-SCM**

**Optimization Details**

SCPO-DAY

**SCPO-DAY->**

**DP-00-IE-SEARS-VNDPACK-LOC-SDAY -> DP-00-IE-SEARS\_VEND\_PACK\_LOC**

Date:- 6/2/2019

Developed By :- Heena Salim Shaikh

**JOB DETAILS**

I have worked on the following scripts:

|  |  |  |  |
| --- | --- | --- | --- |
| **SN** | **Job name** | **Average Run Time** | **Status** |
| 1 | DP-00-IE-SEARS\_VEND\_PACK\_LOC | 02:15:00 | Did the processing as soon as possible so as to achieve the performance benefit. |

This job have following four scripts those are in critical path :-

|  |  |  |
| --- | --- | --- |
| **SN** | **Scripts Names** | **ExecTime(min)** |
| 1 | perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sears\_vendor\_package\_location.pig | 30 |
| 2 | perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sourced\_sears\_location.pig | 20 |
| 3 | perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_candidate\_sears\_warehouse.pig | 14 |
| 4 | perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sears\_vendor\_package\_exploding\_assortment\_location.pig | 15 |
| 5 | perform\_item\_eligibility\_srsvndrpklocn\_smith\_\_idrp\_eligible\_sears\_vendor\_package\_location\_current.pig | 10 |
| 6. | perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sourced\_sears\_warehouse.pig | 7 |

* **Waiting for jobs:-**
* DPG-DP-00-IE-BLD-MSTR-DATA-P2-SDAY-OK :- 7:50 AM ET
* DPG-DP-01-IE-TME-PHSD-STS-P1-SDAY-OK :- 6:57 AM ET
* DPG-DP-02-IE-TME-PHSD-STS-P1-SDAY-OK :- 7:06 AM ET
* DPG-DP-03-IE-TME-PHSD-STS-P1-SDAY-OK :- 7:02 AM ET
* DPG-DP-04-IE-TME-PHSD-STS-P1-SDAY-OK :- 7:06 AM ET
* DPG-DP-00-IDRP-I2K-REBUY-P1-SDAY-OK :- 7:50 AM ET
* DIJ-Daily\_Inv\_DC-Oks :- 7:53 AM ET

**DP-00-IE-SEARS\_VEND\_PACK\_LOC**

**1: Sears Vend Pack Script:-**

* **Pig Name:**

perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sears\_vendor\_package\_location.pig

* **Shell Name:-**

perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sears\_vendor\_package\_location.sh

* **Job execution Time:-**

1. minutes

* **Script Start Time:-**

4/2/2019 9:33:17 AM

* **Input Files:-**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr No** | **Input Files** | **Arrival Time of the File (ET AM)** | **File Count** |
| 1 | "/smith/idrp/vend\_pack\_combined" | 2/4/2019 7:51:00 | 11024804 |
| 2 | "/smith/idrp/eligible\_loc" | 2/3/2019 11:20:00 | 165406 |
| 3 | /smith/idrp/ksn\_attribute/current | 2/4/2019 7:59 | 3034640 |
| 4 | "/smith/idrp/vend\_pack\_dc\_combined" | 2/4/2019 8:24:00 \*\* | 19506142 |
| 5 | /smith/idrp/shc\_item\_combined | 2/4/2019 8:05 | 8145691 |
| 6 | "/work/idrp/sourced\_sears\_location" | 2/4/2019 9:32:00 | 51717427 |
| 7 | /smith/idrp/scan\_based\_trading\_sears\_location/current | 2/5/2019 8:45:00 AM | 7517468 |

* **Output Files:-**

|  |  |  |
| --- | --- | --- |
| **Sr No** | **Output Files** | **File Count** |
| 1 | "/work/idrp/sears\_vendor\_package\_location" | 51569770 |

* **Solution for optimization :-**

From all of above files we can process following 5 files at 8:24 pm only , So that only two file will be process at 9:33 ET and it will execute within 10 minute, so we will get 20 minutes of performance benefit here.

* **"/smith/idrp/vend\_pack\_combined"**
* **"/smith/idrp/eligible\_loc"**
* **"/smith/idrp/vend\_pack\_dc\_combined"**
* **“/smith/idrp/shc\_item\_combined”**
* **"/smith/idrp/vend\_pack\_dc\_combined**"

**2:Sourced Sears location**

* **Pig Name:-**

perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sourced\_sears\_location.pig

* **Shell Name:-**

perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sourced\_sears\_location.sh

* **Job execution Time:**-

20 minutes

* **Script Start Time:-**

04-02 09:13:53 AM

* **Input Files:-**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr No** | **Input Files** | **Arrival Time of the File (ET AM)** | **File Count** |
| 1 | "/work/idrp/sourced\_sears\_store" | 2/4/2019 9:04 | 47424127 |
| 2 | "/work/idrp/sourced\_sears\_import\_center" | 2/4/2019 9:13 | 66771 |
| 3 | "/work/idrp/sourced\_sears\_warehouse" | 2/4/2019 9:02 | 4191256 |
|  | "/work/idrp/amz" | 7/3/2018 7:18 | 11 |
| 5 | "/smith/idrp/ksn\_attribute/current" | 2/4/2019 7:59 \*\*\* | 3034640 |

* **Output Files:-**

|  |  |  |
| --- | --- | --- |
| **Sr No** | **Output Files** | **File Count** |
| 1 | /work/idrp/sourced\_sears\_location | 51717427 |

* **Solution for optimization :-**

From all of above files we can process following 2 files at 8:00 pm only ,So that only rest

three file will be process at 9:13 ET and it will execute within 15 minute, so we will get 15 minutes of performance benefit here.

* "/**work/idrp/amz"**
* **"/smith/idrp/ksn\_attribute/current"**

**3:Eligible Script:-**

* **Pig Name:-**perform\_item\_eligibility\_srsvndrpklocn\_smith\_\_idrp\_eligible\_sears\_vendor\_package\_location\_current.pig
* **Shell Name:-**  
  perform\_item\_eligibility\_srsvndrpklocn\_smith\_\_idrp\_eligible\_sears\_vendor\_package\_location\_current.sh
* **Job execution Time:-**

10 minutes

* **Script Start Time:-**

04-02 10:19:34 AM

* **Input Files:-**

|  |  |  |  |
| --- | --- | --- | --- |
| **SN** | **Input Files** | **File Arrival Time (ET AM)** | **File Count** |
| 1 | /smith/idrp/eligible\_loc | 2/3/2019 11:20 | 165406 |
| 3 | /work/idrp/sears\_vendor\_package\_vendor\_location | 2/4/2019 10:19 | 570576 |
| 4 | /work/idrp/sears\_vendor\_package\_exploding\_assortment\_location | 2/4/2019 10:14 | 1385690 |
| 5 | /work/idrp/sears\_vendor\_package\_location | 2/4/2019 9:58 | 51569770 |
| 6 | /smith/idrp/ksn\_attribute/current | 2/4/2019 7:59\*\* | 3034640 |

|  |  |  |
| --- | --- | --- |
| **SN** | **Output Files** | **File Count** |
| 1 | "/work/idrp/netapp\_ext\_srs\_vp\_loc" | 35796537 |
| 2 | "/smith/idrp/eligible\_sears\_vendor\_package\_location/current" | 53526036 |
| 3 | /work/idrp/netapp\_ext\_srs\_vp\_str" | 16647119 |

* **Solution for optimization :-**

From all of above files we can process following 2 files at 8:00 pm only ,

So that only rest three file will be process at 10:19 ET and it will execute within 7 minutes, so we will get 3 minutes of performance benefit here.

* **/work/idrp/sears\_vendor\_package\_vendor\_location**
* **/work/idrp/sears\_vendor\_package\_exploding\_assortment\_location"**
* **/work/idrp/sears\_vendor\_package\_location**

**4: Candidate**

* **Pig Name:-**

perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_candidate\_sears\_warehouse.pig

* **Shell Name:-**  
  perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_candidate\_sears\_warehouse.sh
* **Job execution Time:-**

15 minutes

* **Script Start Time:-**

04-02 08:30:57 AM

* **Input Files:-**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr No** | **Input Files** | **Arrival Time of the File (ET AM)** | **File Count** |
| 1 | /gold/inventory/dc/sears\_item\_facility/current | 2/3/2019 11:02:00 PM previous day ET | 7798228 |
| 2 | /gold/item/aprk/current | 2/4/2019 5:15 | 377177 |
| 3 | /gold/inventory/srim/daily\_inventory/ | 2/3/2019 21:20 | 7199199 |
| 4 | /work/idrp/sears\_location\_xref | 2/4/2019 8:06 | 19871 |
| 5 | /gold/inventory/dc/sears\_item\_owner\_current | 2/4/2019 07:51 | 8197794 |
| 6 | /smith/idrp/ksn\_attribute/current | 2/5/2019 7:39 \*\*\*\* | 3034640 |

|  |  |  |
| --- | --- | --- |
| **Sr No** | **Output Files** | **File Count** |
| 1 | "/work/idrp/candidate\_sears\_warehouse" | 4194940 |
| 2 | /work/idrp/candidate\_sears\_warehouse\_missing\_source\_error | 81012 |

* **Solution for optimization :-**

From all of above files we can process following 5 files at 7:30 pm only ,

So that only rest two file will be process at 9:13 ET and it will execute within 5 minute, so we will get 10 minutes of performance benefit here.

* **/gold/inventory/dc/sears\_item\_facility/current**
* **/gold/item/aprk/current**
* **/gold/inventory/srim/daily\_inventory/**
* **/work/idrp/sears\_location\_xref**
* **/gold/inventory/dc/sears\_item\_owner\_current**

**5.Assortment**

* **Pig Name:-**

perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sears\_vendor\_package\_exploding\_assortment\_location.pig

* **Shell Name:-**  
  perform\_item\_eligibility\_srsvndrpklocn\_work\_\_idrp\_sears\_vendor\_package\_exploding\_assortment\_location.sh
* **Job execution Time:-**

15 minutes

* **Script Start Time:-**

04-02 08:30:57 AM

* **Input Files:-**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr No** | **Input Files** | **Arrival Time of the File (ET AM)** | | **File Count** |
| 1 | /smith/idrp/core\_ima\_assortments\_prepacks\_exploding\_assortment/current | 2019-02-04 07:54 AM ET | 481402 | |
| 2 | /work/idrp/sears\_vendor\_package\_location | 2019-02-03 09:34 AM ET | 51569770 | |
| 3 | /smith/idrp/vend\_pack\_combined | 2019-02-04 07:30 AM ET | 11024804 | |
| 4 | /smith/idrp/shc\_item\_combined | 2019-02-04 07:43 AM ET \*\* | 8145691 | |
| 5 | /smith/idrp/ksn\_attribute/current | 2019-02-04 07:39 AM ET | 3034640 | |
| 6 | /work/idrp/dummy\_vend\_whse\_ref | 2015-06-04 15:37 AM ET | 27 | |
| 7 | /work/idrp/sears\_location\_xref | 2019-02-04 08:06 AM ET | 19871 | |
| 8 | /smith/idrp/eligible\_loc | 2019-02-03 11:20 AM ET | 165406 | |

* **Output Files:-**

|  |  |  |
| --- | --- | --- |
| **Sr No** | **Output Files** | **File Count** |
| 1 | "/work/idrp/sears\_vendor\_package\_exploding\_assortment\_location" |  |

* **Solution for optimization :-**

We can process the files till 7:45 AM ET , so we will process only 2 files after.

* **/smith/idrp/core\_ima\_assortments\_prepacks\_exploding\_assortment/current**
* **/work/idrp/sears\_vendor\_package\_location**
* **/smith/idrp/vend\_pack\_combined**
* **/smith/idrp/shc\_item\_combined**
* **/smith/idrp/ksn\_attribute/current**
* **/work/idrp/dummy\_vend\_whse\_ref**
* **/work/idrp/sears\_vendor\_package\_location**
* **/work/idrp/sears\_location\_xref**
* **/smith/idrp/eligible\_loc**

**THANK YOU ☺!!**