

Hi Team.

Issue - Parallel execution(with number of parallel execution 2) took 338 milli seconds where as non parallel execution took 49 seconds to get the document name of 15 PDF files(just file name without reading file content) in a directory.

Use Case - Get document name of files in a directory parallely

Talend Tool - Talend Data Integration 6.5.1

Here are the Main Job Design Details:

Using tFileList to get the list of all files in a directory

Using iterate to connect to tRunJob and using the Enable Parallel Execution to process files parallely. Number of parallel executions specified were 2.

Passing the FileName as a context parameter as below to subjob:

```
(String)globalMap.get("tFileList_1_CURRENT_FILEPATH")
```

Sub Job Details:

Getting the FileName as context parameter as below:

```
context.FileName
```

Using tFixedFlowInput to generates a fixed flow from internal variables

using tLogRow to print the PDF file name on console

My System Configuration:

i7-4790 processor with 16 GB RAM

CPU Cores 4

Test Data:

Tried to get the file name of 15 PDF files stored in a directory parallely:

Test Results:

Parallel execution(number of parallel execution were 2) took 338 milli seconds where as non parallel execution took 49 seconds to get the 15 PDF files name

Test Results Details:

Test Results without enabling parallel execution:

Starting job ReadFiles3 at 14:16 11/10/2019.

[statistics] connecting to socket on port 3857

[statistics] connected

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\28997.pdf|  
|-----|
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\34684.pdf|  
|-----|
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\34686.pdf|  
|-----|
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\35495.pdf|  
|-----|
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\36491.pdf|  
|-----|
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\36651.pdf|  
|-----|
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\37561.pdf|  
-----
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\37905.pdf|  
-----
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\39296.pdf|  
-----
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\40822.pdf|  
-----
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\56645.pdf|  
-----
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\56647.pdf|  
-----
```

```
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\57195.pdf|  
-----
```

```
-----  
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\57197.pdf|  
-----
```

```
-----  
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\Geodatafy.pdf|  
-----
```

49 milliseconds

[statistics] disconnected

Job ReadFiles3 ended at 14:16 11/10/2019. [exit code=0]

Test Results with enabling parallel execution:

Starting job ReadFiles3 at 14:15 11/10/2019.

[statistics] connecting to socket on port 4069

[statistics] connected

```
-----  
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\28997.pdf|  
-----
```

```
-----  
-----  
| tLogRow_1 |  
|=-----=  
|FileName |  
|=-----=  
|D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\34684.pdf|  
-----
```

```
-----  
-----  
| tLogRow_1 |  
|=-----=  
|FileName |
```

=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\35495.pdf

| tLogRow_1 |
=====|
|FileName |
=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\34686.pdf

| tLogRow_1 |
=====|
|FileName |
=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\36651.pdf

| tLogRow_1 |
=====|
|FileName |
=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\36491.pdf

| tLogRow_1 |
=====|
|FileName |
=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\37905.pdf

| tLogRow_1 |
=====|
|FileName |
=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\37561.pdf

| tLogRow_1 |
=====|
|FileName |
=====|
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\39296.pdf

| tLogRow_1 |

```

=====
|FileName |
=====
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\40822.pdf
-----

tLogRow_1 |
=====
|FileName |
=====
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\56645.pdf
-----

tLogRow_1 |
=====
|FileName |
=====
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\56647.pdf
-----

tLogRow_1 |
=====
|FileName |
=====
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\57197.pdf
-----

tLogRow_1 |
=====
|FileName |
=====
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\Geodatafy.pdf
-----

tLogRow_1 |
=====
|FileName |
=====
D:\Data\GeoDatafy\Lakshmi\25092019\IndexTest\Temp\57195.pdf
-----

```

338 milliseconds

[statistics] disconnected

Job ReadFiles3 ended at 14:15 11/10/2019. [exit code=0]

Can you please let us know what could be the issue if we enable Parallel execution.

Thanks in advance for your help!

