



# New Data Reduction Tool

**Josh Wisniewski**  
Tools Architect

**IBM z/TPF**  
April 11, 2016

10 Minutes

# New Data Reduction Tool

5 Minutes

## Examples

## Q&A

# Disclaimer

Any reference to future plans are for planning purposes only. IBM reserves the right to change those plans at its discretion. Any reliance on such a disclosure is solely at your own risk. IBM makes no commitment to provide additional information in the future.

# New Data Reduction Tool

## Problem statement

- The existing data reduction tool is one of the few remaining tools that requires a z/OS system and PL/I compiler in support of z/TPF. If the existing data reduction tool was rewritten to run Linux on z, like most other z/TPF tools, tooling on z/OS systems and PL/I compiler licenses could be eliminated.
- It can be cumbersome to feed the existing data reduction results into analytics or other tooling that aid in understanding the data.

## New data reduction tool

- This new tool will be written on Linux on z in C++.

## Benefits

- May be able to eliminate PL/I compiler license.
- z/OS interaction is not required to run new data reduction tool.
- Simple modern interface (no JCL training required).
- New data format in CSV for ease of import into analytics tooling, interpreting the data with a spreadsheet program (ie sort by column, graphing, and etc. functions) and etc.
- New z/TPFDF statistics report which is not available in the existing data reduction tool.

## As is scenario

1. Data collection starts on z/TPF.
2. Results are written to tape on z/TPF.
3. Data collection stops.
4. Data collection tape is mounted on z/OS.
5. Modify JCL to run existing data reduction against the tape.
6. Textual reports are produced as output.
7. For results to be piped into analytics tooling, customer must write a program to screen scrape the results from textual reports.

## To be scenario

1. Data collection starts on z/TPF.
2. Results are written to tape on z/TPF.
3. Data collection stops.
4. Data collection tape is mounted on **Linux on z**.
5. Issue **pptapetofile** command on Linux on z to convert the tape to a binary file.
6. Modify new data reduction configuration file as needed.
7. Issue new data reduction command, **tpfdataredux**, on Linux on z.
8. Output in CSV (comma separated file format) and/or textual reports.
9. CSV output files are easily consumed into analytics tooling and spreadsheet software.

## First deliverable of New Data Reduction Tool

- Environment Summary Report in CSV format.
- z/TPFDF Statistics Report in CSV format.
- Targeting beta in early 2Q2016 with Sponsor Users.
- Incorporate Sponsor User feedback.
- Targeting GA at end 2Q2016.
- APAR number: PJ42460



## Future plan for New Data Reduction Tool

- Efforts will continue to leverage Design Thinking iterations:
  - Plan content of next deliverable based upon Sponsor User input:
    - New CSV reports (ie System Summary, ECB Heap, and etc.)
    - Textual version of a delivered report (ie Environment or z/TPFDF)
  - Coding/testing efforts.
  - Beta and incorporate Sponsor User input.
  - GA deliverable.

## Future Direction Intent

- The new data reduction tool will only convert the valued reports, options and etc. as directed by Sponsor User input.
- Our intent is to eventually deprecate the existing data reduction tool that runs on z/OS.
- At some point new features and maintenance will only be delivered in the new data reduction tool.

# Example

New tape conversion utility

- General purpose utility that is used to copy and format data from tape into binary file format on the Linux on z file system. This utility can be used for processing tapes for other purposes.

- Usage:

```
pptapetofile -t tapedevice[volser] -o outputFile
```

- Sample:

```
pptapetofile -t /dev/tape[T1G100] -o dacol_20160311.bin
```

## New data reduction tool configuration file

- The configuration file is used to specify input and output options. It will also be used to specify which reports should be generated and any associated options (for existing data reduction these options are specified as JCL cards).
- Sample file: opt.cfg

```
1 @general
2 inputfile=dacol_20160311.bin
3 outputdir=.
4 reports=(ALL)
```

## New data reduction tool

- This command is used to run the new data reduction tool to create CSV and textual reports from the binary file on Linux on z that was created from the data collection tape.

- Usage:

```
tpfdataredux CONFIGURATIONFILE.cfg
```

- Sample:

```
tpfdataredux opt.cfg
```

- Environment Report Sample CSV output

```
1 Date Reduced,06/04/2016
2 Time Reduced,12:12:33
3 Input File Reduced,/projects/tpfdataredux/current/tapes/dataCollection24by7.bin
4
5
6 TPF ENVIRONMENT SUMMARY REPORT SYSTEM WIDE
7 DATE COLLECTED,25 MAR
8 DATE REDUCED,06/04/2016
9 DATA COLLECTION TAPE VOLUME SERIAL NUMBER,T1G476
10 ENVIRONMENT,PR/SM-SH
11 PARTITION,TPFP1
12 CAP INDICATOR,N
13 WAIT COMP,N
14 IBM MOD,2827-750
15 SERIAL NUMBER,0005A5B6
16 TPF CPUID,B
17 TPF RELEASE AND VERSION,z/TPF EE
18 PUT LEVEL,010112
19 VIRTUAL MODE,VEQV
20 STATIC POWER SAVE MODE,temp placeholder
21
22 OPERATOR INPUT REQUEST,ZMEAS V/SFPM/6060/08080208
23 MODE OF COLLECTION,SAMPLING
24 INTERVAL SYSTEM,8 SEC
25 FILE,8 SEC
26 PROGRAM,2 SEC
27 MESSAGE,8 SEC
28 PERIOD,60 SEC
29 FILE COLLECTION INTERCEPT SKIP FACTOR,9999
30 PROGRAM COLLECTION INTERCEPT SKIP FACTOR,9999
31 START TIME,11:29:56
32 END TIME,12:29:24
33
34 ACTIVE UTILITIES,START,END
35 DATA COLLECTION ,YES,YES
36
```

**Example**

- Environment Report Sample CSV output continued

```

37 ACTIVE TRACES,START,END
38     ECB TRACE                ,YES,YES
39     ENTER/BACK TRACE        ,YES,YES
40     SYSTEM LOG              ,YES,YES
41     IO TRACE                ,YES,YES
42     BRANCH TRACE           ,YES,YES
43     SOCKET TRACE           ,YES,YES
44     BLOCK CHECK MODE       ,YES,YES
45     I/O CCW TRACE          ,YES,YES
46     EXTENDED FUNCTION TRACE,YES,YES
47     FUNCTION TRACE         ,YES,YES
48     IDLE TIMER             ,YES,YES
49     ECB HEAP TRACE         ,YES,YES
50
51 COLLECTOR OPTIONS,OPTION,SUBOPTION
52     NONE,,
53
54
55 CONFIGURATION SUMMARY
56 CPU CONFIGURATION
57 MACH TYPE,SERIAL NR,CPU ID
58     2827-750,005A5B6,B,DATA COLLECTED ON THIS CPU
59     ,0015A5B6,C,
60     ,00BA5B6,D,
61     ,00CA5B6,M,
62
63
64 MDBF CONFIGURATION
65 SS NAME,SS ORD,SS STATE,SSU NAME,SSU ORD,
66 BSS ,0,NORM
67   ,,HPN ,0
68 WP  ,1,NORM
69   ,,WP1 ,1
70   ,,WP2 ,2
71   ,,WP3 ,3

```

# Sponsor User Interest

- We have a few Sponsor Users signed up (one beta level) already.
- If you're interested in being a Sponsor User, let us know.



# New Data Reduction Tool

What is it:

- This new tool will be written on Linux on z in C++ to replace the existing data reduction tool that currently runs on z/OS.

Benefits:

- May be able to eliminate PL/I compiler license.
- Simplified modern interface without a need for z/OS interaction.
- New data output format as CSV for ease of import into analytics tooling, spreadsheet programs and etc.
- New z/TPFDF statistics report which is not available in the existing data reduction tool.

**Thank you!**

Questions or comments?

# Trademarks

- IBM, the IBM logo, ibm.com and Rational are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “[Copyright and trademark information](#)” at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

## Notes

- Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.
- All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
- This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
- All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
- Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
- Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.
- This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.