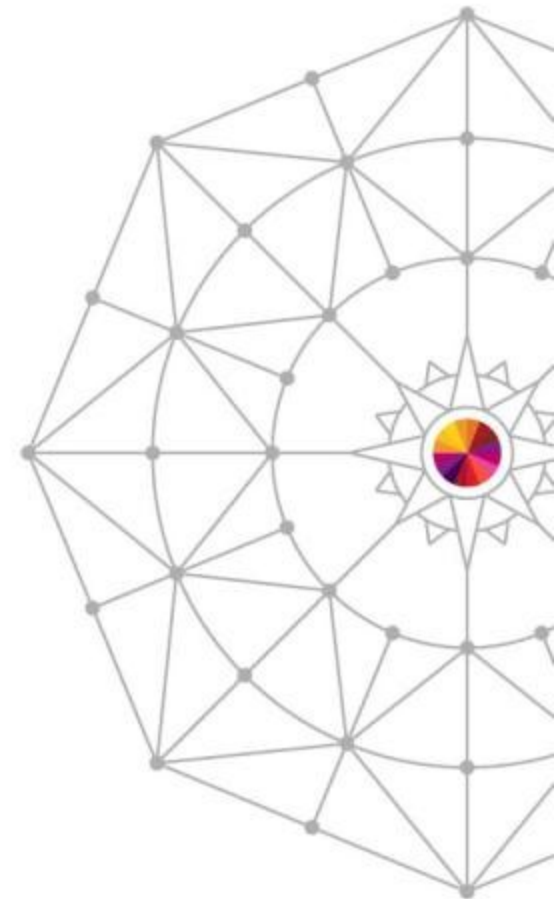


Unlocking Data on z/TPF: Data Format Description Language (DFDL)

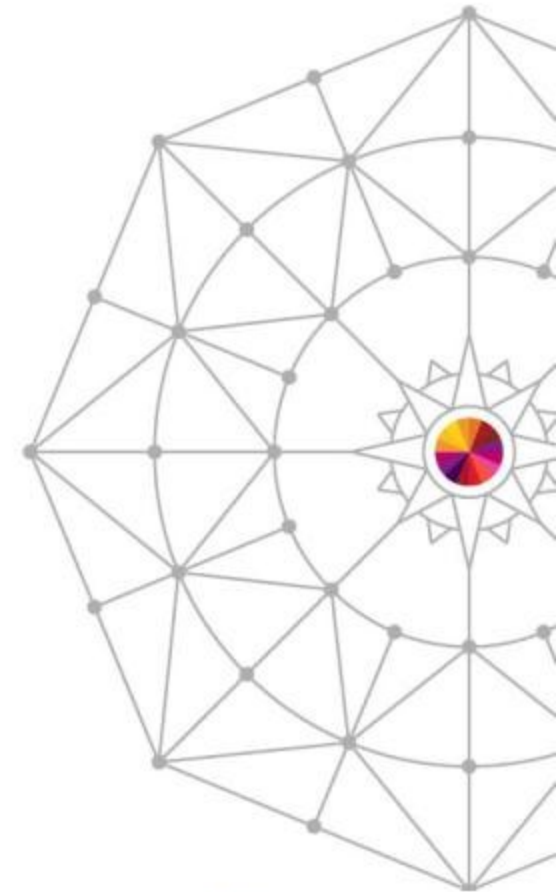
Bob Dryfoos
TPF Development Lab

Main Tent
March 10, 2014



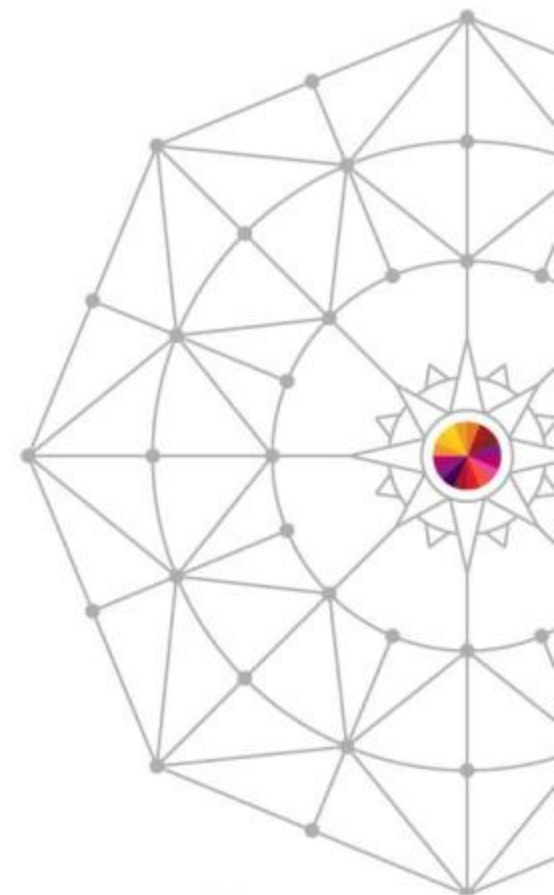
Disclaimer

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.



How will TPF data be formatted for other systems to consume?

Data Format Description Language.



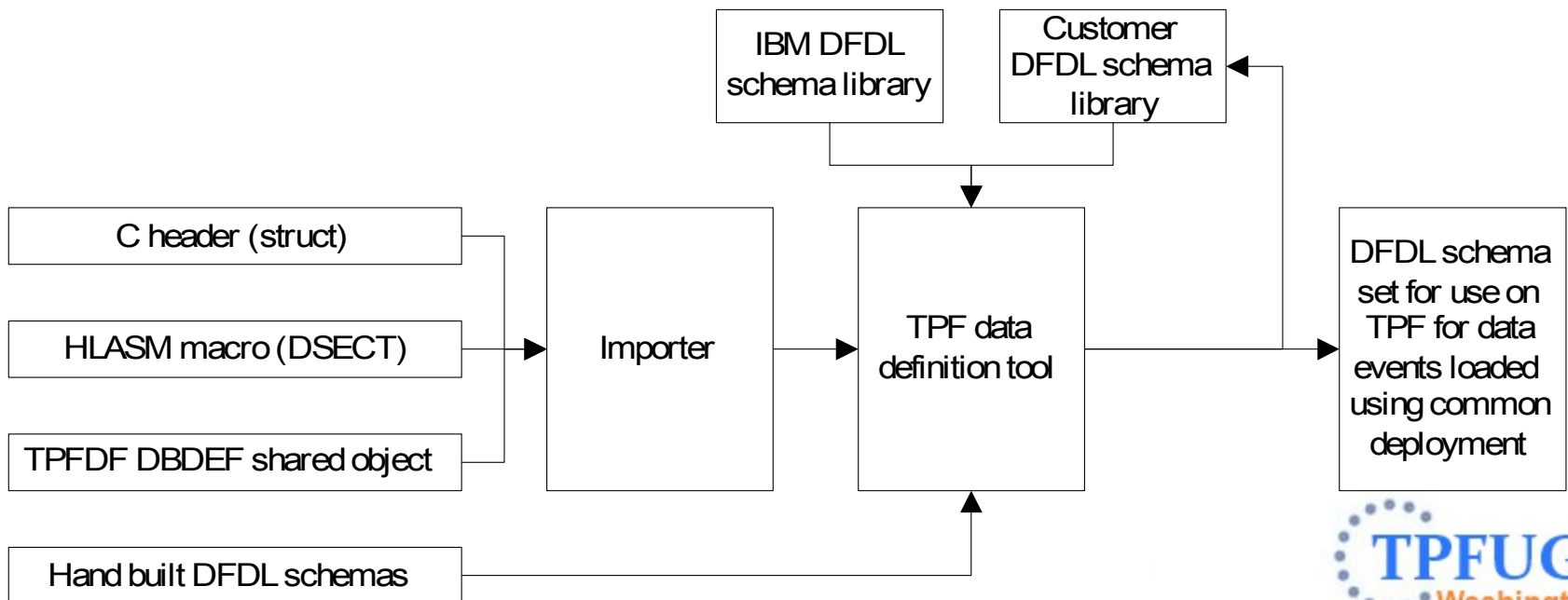
Why Data Format Description Language (DFDL)

- DFDL will be used to:
 - Describe binary data from TPF for use on other platforms.
 - Convert TPFDF, find and file, and other data into XML, JSON, and other consumable formats.
 - Embed application intelligence directly into the DFDL schema.
- DFDL is a standardized consumable format available for use on a variety of platforms.
 - Active open source development community
<http://www.ogf.org/dfdl/>.
 - IBM is adopting DFDL across products
 - Webspere Message Broker.
 - Integration Bus.
 - Rational Developer for System z.
 - InfoSphere Master Data Management.



TPF Data Definition Tool

- Importer provides a means to start the creation of DFDL schemas from existing formats.
- TPF data definition tool allows the user to
 - Specify application intelligence in the DFDL schema.
 - Specify attributes in the DFDL schema such as blanking a field in the data events output message (ie credit card number).



TPF Data Definition Tool

- IBM DFDL tooling shown below as an example implementation.
 - Project explorer, outline, field layout, and field properties views.
- TPF tooling will be Eclipse based and pluggable into TPF Toolkit.

The screenshot displays the IBM DFDL tooling interface, which is Eclipse-based. It features several views:

- Data Project Explorer:** Shows a project named 'myProject' with subfolders for 'Data Diagrams', 'Data Models', and 'Other Files'. The 'Example.xsd' file is selected.
- Outline:** Displays the structure of the data model, including 'Imports', 'Messages', 'Elements', and 'Types'. The 'Complex Types' section is expanded, showing 'IR00_keyA0_History', 'IR00_key90_Address', and 'IR00_key80_Name'.
- Field Layout:** A table showing the structure of the data model. The 'key' element is selected, and its properties are displayed in the 'Representation Properties' view.
- Representation Properties:** Shows the properties of the selected 'key' element, including 'Comment', 'General' (Data Format Reference, Encoding, Byte Order, Ignore Case, Fill Byte), and 'Content' (Representation).

Name	Type
DF_file	
sequence	
length	unsignedShort
key	unsignedByte
choice	
nameRecord	IR00_key80_Name
sequence	
firstName	string
lastName	string
addressRecord	IR00_key90_Address
sequence	
Street	string
State	string
historyRecord	IR00_keyA0_History
sequence	
Text	string

Property	Value
Comment	
General	
Data Format Reference	<default format>
Encoding (code page)	IBM1047
Byte Order	bigEndian
Ignore Case	no
Fill Byte	0
Content	unsignedByte
Representation	text

TPF Data Definition Tool

- DFDL schemas are highly extensible and optimized for reuse:
 - Ability to define data formats, encoding, and etc.
 - Ability to define reusable complex types.
 - Ability to define conditional data layouts.
 - Ability to import other DFDL schemas to facilitate reuse.

Name	Type
[-] [e] DF_file	
[-] ... sequence	
⋮ [e] length	unsignedShort
⋮ [e] key	unsignedByte
⋮ [-] choice	
⋮ [-] [e] nameRecord	IR00_key80_Name
[-] ... sequence	
⋮ [e] FirstName	string
⋮ [e] LastName	string
⋮ [-] [e] addressRecord	IR00_key90_Address
[-] ... sequence	
⋮ [e] Street	string
⋮ [e] State	string
⋮ [-] [e] historyRecord	IR00_keyA0_History
[-] ... sequence	
⋮ [e] Text	string

Advantages of Data Format Description Language (DFDL)

- Designed to describe textual and binary data including a focus on legacy and record formatted data.
- DFDL is implemented based on standard W3C XML Schema Definition Language (XSDL) 1.0 .
 - DFDL tagging is accomplished using annotations.
- The TPF DFDL schemas can be provided to other systems to allow them to consume the TPF data.
- DFDL parsers process data and a DFDL schema to produce XML, JSON and other formats for ease of consumption.
 - DFDL parsers are designed to be highly efficient.

Trademarks

- IBM, the IBM logo, and ibm.com 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.

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.

