# Data Format Description Language (DFDL) v1.0

**Experience Document 8**

**Experimental Feature for supporting indirection**

Status of This Document

Grid Working Document (GWD)

Copyright Notice

Copyright © Open Grid Forum, (2021). Some Rights Reserved. Distribution is unlimited.

Abstract

This document provides experience information to the OGF community on the Data Format Description Language (DFDL) 1.0 specification (GFD-P-R.207).

It documents an experimental feature that supports indirection such as offsets and pointers which may be used in various data formats.

**Contents**

[1. Description 2](#_Toc64536648)

[2. Examples 3](#_Toc64536649)

[3. DFDL Data syntax grammar update 5](#_Toc64536650)

[4. SECURITY Considerations 7](#_Toc64536651)

[5. Contributors 7](#_Toc64536652)

[6. Intellectual Property Statement 7](#_Toc64536653)

[7. DIsclaimer 7](#_Toc64536654)

[8. Full Copyright Notice 7](#_Toc64536655)

[9. References 8](#_Toc64536656)

# Description

The following experimental properties allow for data to be accessed in a non-sequential or non-contiguous format by using indirection values to locate the data contents. When parsing, the indirection value is used to access the contents in the data stream to populate the infoset. When unparsing, an indirection value is created to store the contents in the data stream when creating the data stream from the infoset. The indirection value does not exist in the infoset, but only in the data stream.

The LeadingAlignment and TrailingAlignment regions apply to the indirection value in the data stream. The alignmentUnits must be bytes or it is a schema definition error. The initiator and terminator as well as all regions between the LeftFraming and RightFraming regions are applied to the data content accessed through the indirection value.

|  |  |
| --- | --- |
| Property | Description |
| dfdlx:indirectionKind | Enum  Specifies the type of indirection value used to access the contents of the data stream.  Valid values are ‘pointer’, ‘offset’, ‘none’.  ‘pointer’ means a numeric value gives an implementation-dependent starting byte position of the content.  ‘offset’ means a numeric value that when added to indirectionBase gives the starting byte position of the content.  ‘none’ means no indirection is used and is the default value.  Annotation: dfdl:choice, dfdl:element, dfdl:group, dfdl:simpleType, dfdl:sequence |
| dfdlx:indirectionType | QName  Name of a simple type derived from xs:nonNegativeInteger or any subtype of it.  This type, with its DFDL annotations, specifies the representation of the indirection value.  It is a schema definition error if the xs:simpleType specifies any of:   * dfdl:lengthKind ‘delimited’, ‘endofParent’, or ‘pattern’ * dfdl:lengthKind ‘explicit’ where length is an expression * dfdl:outputValueCalc * dfdl:initiator or dfdl:terminator other than empty string * dfdl:alignment other than ‘1’ * dfdl:leadingSkip or dfdl:trailingSkip other than ‘0’ * dfdl:lengthUnits other than ‘bytes’ * dfdl:alignmentUnits other than ‘bytes’   Annotation: dfdl:choice, dfdl:element, dfdl:group, dfdl:simpleType, dfdl:sequence |
| dfdlx:indirectionBase | String  Provides the absolute or relative path to an element upon which the offset is based. This path must not be a forward reference to an element that has not yet been processed. The offset value is added to the starting position of this element.  Required only when dfdl:indirectionKind is ‘offset’.  Annotation: dfdl:choice, dfdl:element, dfdl:group, dfdl:simpleType, dfdl:sequence |
| dfdlx:indirectionUnusedValue | Integer  Specifies the indirection value that indicates when the data content is missing. If this property is specified, all child elements must contain a default value specification or it can result in a parse error.  Annotation: dfdl:choice, dfdl:element, dfdl:group, dfdl:simpleType, dfdl:sequence |

1. Examples

The following example illustrates how to describe a pointer to a null-terminated string (common in languages like C).

<xs:element name="myString" type="xs:string"  
 dfdl:lengthKind="delimited" dfdl:encoding="UTF-8"  
 dfdl:terminator="%NUL;" dfdlx:indirectionKind="pointer"  
 dfdlx:indirectionType="ptr32\_t"  
 dfdlx:indirectionUnusedValue="0" default=""/>

<xs:simpleType name="ptr32\_t" dfdl:representation="binary"  
 dfdl:lengthKind="explicit" dfdl:length="4"  
 dfdl:byteOrder="bigEndian">  
 <xs:restriction base="integer"/>  
</xs:simpleType>

The data stream may look like the following for element ‘myString’ with a string value of “test”.

Location Hex values  
00000000 0012A000  
  
0012A000 7465737400  
 t e s t

The following example defines an array of 3 pointers to complex elements defined by ns0:myStruct.

<xs:element name="myArray" type="ns0:myStruct"  
 dfdl:lengthKind="implicit" dfdlx:indirectionKind="pointer"  
 dfdlx:indirectionType="ptr32\_t"  
 dfdlx:indirectionUnusedValue="0"  
 minOccurs="3" maxOccurs="3" dfdl:occursCountKind="fixed"/>

<xs:complexType name="myStruct">  
 <xs:sequence>  
 <xs:element name="subItem1" type="xs:int"   
 dfdl:lengthKind="explicit" dfdl:length="4"   
 default="0"/>  
 <xs:element name="subItem2" type="xs:string"   
 dfdl:lengthKind="explicit" dfdl:length="12"   
 default=""/>  
 </xs:sequence>  
</xs:complexType>

<xs:simpleType name="ptr32\_t" dfdl:representation="binary"  
 dfdl:lengthKind="explicit" dfdl:length="4"  
 dfdl:byteOrder="bigEndian">  
 <xs:restriction base="integer"/>  
</xs:simpleType>

The data stream may look like the following for element ‘myArray’ with the contents of each occurrence in a different location and the contents of the 2nd occurrence being empty.

Location Hex values  
00000000 00147000 00000000 00146000

00146000 ...

00147000 ...

The following example defines an offset to a 6 byte hexBinary value with the offset being calculated from the start of the current element.

<xs:element name="myData" type="xs:hexBinary"  
 dfdl:lengthKind="explicit" dfdl:length="6"  
 dfdlx:indirectionKind="offset"  
 dfdlx:indirectionType="uint"  
 dfdlx:indirectionBase="."/>

<xs:simpleType name="uint" dfdl:representation="binary"  
 dfdl:lengthKind="explicit" dfdl:length="4"  
 dfdl:byteOrder="bigEndian">  
 <xs:restriction base="xs:unsignedInt"/>  
</xs:simpleType>

The data stream may look like the following for element ‘myData’ with a hexBinary value of ‘123456789ABC’.

Location Hex values  
00000100 00000108

00000208 123456789ABC

1. DFDL Data syntax grammar update

The following outlines changes that can be made to the grammar for indirection support with changes indicated in blue:

|  |
| --- |
| Document = DocumentElement IndirectionElement\*  DocumentElement = SimpleElement | ComplexElement  SimpleElement = SimpleLiteralNilElementRep | SimpleEmptyElementRep |  SimpleNormalRep | IndirectionRep  ~~SimpleEnclosedElement = SimpleElement | AbsentElementRep~~  ComplexElement = ComplexLiteralNilElementRep | ComplexNormalRep |  ComplexEmptyElementRep | IndirectionRep  ~~ComplexEnclosedElement = ComplexElement | AbsentElementRep~~  EnclosedElement = SimpleElement | ComplexElement | AbsentElementRep  IndirectionElement = SimpleElement | ComplexElement |
| AbsentElementRep = ***Absent*** |
| IndirectionRep = IndirectionLeftFraming PrefixLength IndirectionContent  IndirectionRightFraming  IndirectionLeftFraming = LeadingAlignment  IndirectionRightFraming = TrailingAlignment  IndirectionContent = ***LeftPadding*** [ IndirectionValue ***| IndirectionUnusedValue*** ] RightPadOrFill  IndirectionValue = ***SimpleValue*** |
| SimpleEmptyElementRep = EmptyElementLeftFraming EmptyElementRightFraming  ComplexEmptyElementRep = EmptyElementLeftFraming EmptyElementRightFraming  EmptyElementLeftFraming = LeadingAlignment ***EmptyElementInitiator*** PrefixLength  EmptyElementRightFraming = ***EmptyElementTerminator*** TrailingAlignment |

1. SECURITY Considerations

No security issues have been raised.

1. Contributors

Bradd Kadlecik,

IBM Software Group,

Poughkeepsie, NY, USA

[braddk@uk.ibm.com](mailto:smh@uk.ibm.com)

1. Intellectual Property Statement

The OGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the OGF Secretariat.

The OGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the OGF Executive Director.

1. DIsclaimer

This document and the information contained herein is provided on an “As Is” basis and the OGF disclaims all warranties, express or implied, including but not limited to any warranty that the use of the information herein will not infringe any rights or any implied warranties of merchantability or fitness for a particular purpose.

1. Full Copyright Notice

Copyright (C) Open Grid Forum (2021). Some Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included as references to the derived portions on all such copies and derivative works. The published OGF document from which such works are derived, however, may not be modified in any way, such as by removing the copyright notice or references to the OGF or other organizations, except as needed for the purpose of developing new or updated OGF documents in conformance with the procedures defined in the OGF Document Process, or as required to translate it into languages other than English. OGF, with the approval of its board, may remove this restriction for inclusion of OGF document content for the purpose of producing standards in cooperation with other international standards bodies.

The limited permissions granted above are perpetual and will not be revoked by the OGF or its successors or assignees.

1. References

[DFDL] OGF DFDL 1.0 specification: <http://www.ogf.org/documents/GFD.207.pdf/>