TeamRoom Plus

IBM TeamRoo

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**13-Oct-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Alan Powell on 14-Oct-10
Last Modified by Alan Powell on 14-Oct-10

OGF DFDL Working Group Call , October 13 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , October 13 2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Suman Kalia (IBM)
Bob McGrath (National Center for Supercomputing Applications)

Apologies
Mike Beckerle (Oco)
Tim Kimber(IBM)

1. Current Actions
Updated Below

2. Draft 43 comments from Steve and Suman
13/10: We went through Steve's many comments and issues.
Issues:
- p32 6.3 Enumerations do not specify a type that an expression should returns. Most are strings but some are
ints. Will state that enums are strings unless otherwise stated.

- p 43 7.3.1 Changed the testValueKind to testKind so it covers attributes as well as value.

- p60 9.2 Will include prefixLength in leftElementFraming

- p62 11 UTF16Width becomes utf16Width and applies when UTF-16, UTF-16BE/LE and their CCSID equivalents.

- p68 12.2 Is is not a schema definition error when emptyValueDelimiterPolicy 'initiator'/'both' and lengthKind 'implicit' for complex elements.

-p85 13.2.1 generateEscape really only applies to escape blocks so could be renamed.

-p111 13.16 nilValue literalCharacter. If not fixed length how many should be output? Steve to check what MRM does.

- p126 17. Remove inputValueCalc/outputValueCalc from simpleType

3. dfdl:roundingIncrement
Problems with rounding and ICU
13/10: not discussed

4. Support for COBOL type PIC 99V99 COMP
The dfdl:binaryDecimalVirtualPoint property should cover 'packed', 'bcd' and 'binary'.
13/10: To support this COBOL type need to allow xs:decimal to have binaryNumberRep 'binary' and apply dfdl:binaryDecimalVirtualPoint

5. AOB - Is support for locales sufficient ?
13/10: not discussed

Meeting closed, 16:45

Next call Wednesday 20 October 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views." All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **129**

Actions raised at this meeting

No	Action
125	-p111 13.16 nilValue literalCharacter. If not fixed length how many should be output? Steve to check what MRM does.
126	-p85 13.2.1 generateEscape really only applies to escape blocks so could be renamed.
127	**3. dfdl:roundingIncrement** Problems with rounding and ICU 13/10: not discussed
128	**5. AOB - Is support for locales sufficient ?** 13/10: not discussed

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing

	17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1' 01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite. 08/09: IBM still progressing 15/09: IBM still progressing, expect tests to be available within a few weeks 22/09: IBM still progressing, expect tests to be available within a few weeks 29/09:Test cases are being prepared. 06/10: Some test cases should be available next week. Steve would like to be able to show the test case information at OGF 30. 13/10: Still progressing
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification. 11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication. 11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'. Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process. 01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and closure of all actions. 08/09: Target for completion September 30. 15/09: as above 22/09: as above 29/09: Draft 43 will be published this week for WG review prior to submitting to OGF 06/10: Draft 43 was distributed on 30 September and includes all outstanding work items. This

	will be the last draft before submission to OGF for the final committee review. All comments should be sent to Alan by Tuesday 12th Oct and final version will be submitted on Friday 15th. 13/10: Thorough review by Steve and comments from Suman and Stephanie. Most updates done. Final review draft will be issued 14th Oct and final version to OGF soon after.
111	**Daffodil DFDL parser** 11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented. We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats. Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil. Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed. It would be helpful if IBM test cases were available to Daffodil. IBM will investigate 25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential. Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list. Bob is planning to make the parser available on the internet to allow testing. His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way. 01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase. Daffodil being placed under formal source control in anticipation of external release. Bob has a start October deadline to create a report on what has been done for his sponsors. It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October. 08/09: Alejandro is marking up Spec draft 42 to indicate which features Daffodil implement. Bob expects Daffodil to be available on the web soon. 15/09: Alejandro had indicated in the specification which functions were implemented in Daffodill. Steve had reviewed and identified which function need to be implemented and which could be considered optional (see action 099). Alejandro is implementing the missing core functions. There was some discussion about the limitations on unordered groups. (stop value and expression not supported). It was agreed that it should be a schema definition error if dfdl:occursCountKind is 'stopValue' on any element within an unordered sequence and a floating element. 22/09: not discussed 29/09: not discussed 06/10: Alejandro has left NCSA. Bob is making the case for continuing and having a replacement. Bob to agree with Steve what can be said at OGF30. 13/10: Bob still progressin project funding and making Daffodil publically available.
112	**DFDL certification process** 25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process. 01/09: In progress, spec needs to state what conformance means, as part of this work 08/09: Discussed what needs to be said in the spec and agreed that details of a conformance test suite should be in another document. Alan to draft conformance section. 15/09: Alan had look at the conformance sections in XML and Schema specifications both of which indicate sections which must be implemented. None just say 'execute the test suite'. They talk in terms of conformance of document, schema and processors.. 22/09: no progress 22/09: Alan has added short Conformance and Optional Features sections to spec which was

		briefly discussed. Discussed naming for processors that don't implement optional features and those that implement all features. 06/10: Need to decide what/how test cases and certification process should occur 13/10: no progress
	123	DFDL tutorial 13/10: Draft of first 3 chapters has been written and will be distributed to WG
	124	DFDL web content on OGF standards pages 13/10: no progress
	125	-p111 13.16 nilValue literalCharacter. If not fixed length how many should be output? Steve to check what MRM does.
	126	-p85 13.2.1 generateEscape really only applies to escape blocks so could be renamed.
	127	3. dfdl:roundingIncrement Problems with rounding and ICU 13/10: not discussed
	128	5. AOB - Is support for locales sufficient ? 13/10: not discussed

Closed actions

No	Action
114	OGF 30 25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session? 09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30. 08/09: Steve to request permission to go 15/09: Travel request has be submitted 22/09: DFDL session is scheduled at 11:00am Monday Oct 25th. 06/10: Steve has approval to travel. 13/10: Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
07	Write DFDL primer		

0			
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**06-Oct-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Alan Powell on 06-Oct-10
Last Modified by Alan Powell on 06-Oct-10

OGF DFDL Working Group Call , October 06 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , October 06 2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Tim Kimber(IBM)
Suman Kalia (IBM)
Bob McGrath (National Center for Supercomputing Applications)

Apologies
Mike Beckerle (Oco)
Alejandro Rodriguez (National Center for Supercomputing Applications)

1. Current Actions
Updated Below

2. truncateSpecifiedLengthString

Suggest rewording as follows:

The position from which data is truncated is determined by the value of the textStringJustification property. If the value of the textStringJustification property is 'left', data is truncated from the right; if the value of the textStringJustification property is 'right', data is truncated from the left. However if the value of the textStringJustification property is 'center' or 'none', truncation does not occur and a processing error occurs if the string is too long.

3. Alignment question - COBOL SYNC RIGHT

Does DFDL v1 need to support COBOL Sync Right? No

3. AOB/ Future actions
- DFDL tutorial
- DFDL web content on OGF standards pages.

Meeting closed, 15:35

Next call Wednesday 13 October 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **125**

Actions raised at this meeting

No	Action
123	DFDL tutorial
124	DFDL web content on OGF standards pages

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1' 01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite. 08/09: IBM still progressing 15/09: IBM still progressing, expect tests to be available within a few weeks 22/09: IBM still progressing, expect tests to be available within a few weeks 29/09:Test cases are being prepared. 06/10: Some test cases should be available next week. Steve would like to be able to show the test case information at OGF 30.
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up

14/07: No response from Joel. Sent email to Greg Newby by no response.
21/07: Still no response.
04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification.
11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication.
11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'.
Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process.
01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and closure of all actions.
08/09: Target for completion September 30.
15/09: as above
22/09: as above
29/09: Draft 43 will be published this week for WG review prior to submitting to OGF
06/10: Draft 43 was distributed on 30 September and includes all outstanding work items. This will be the last draft before submission to OGF for the final committee review. All comments should be sent to Alan by Tuesday 12th Oct and final version will be submitted on Friday 15th.

111	Daffodil DFDL parser

11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented.
We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats.
Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil.
Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed.
It would be helpful if IBM test cases were available to Daffodil. IBM will investigate
25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential.
Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list.
Bob is planning to make the parser available on the internet to allow testing.
His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way.
01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase.
Daffodil being placed under formal source control in anticipation of external release.
Bob has a start October deadline to create a report on what has been done for his sponsors.
It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October.
08/09: Alejandro is marking up Spec draft 42 to indicate which features Daffodil implement. Bob expects Daffodil to be available on the web soon.
15/09: Alejandro had indicated in the specification which functions were implemented in Daffodill. Steve had reviewed and identified which function need to be implemented and which could be considered optional (see action 099). Alejandro is implementing the missing core functions. There was some discussion about the limitations on unordered groups. (stop value and expression not supported). It was agreed that it should be a schema definition error if

		dfdl:occursCountKind is 'stopValue' on any element within an unordered sequence and a floating element. 22/09: not discussed 29/09: not discussed 06/10: Alejandro has left NCSA. Bob is making the case for continuing and having a replacement. Bob to agree with Steve what can be said at OGF30.
	112	**DFDL certification process** 25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process. 01/09: In progress, spec needs to state what conformance means, as part of this work 08/09: Discussed what needs to be said in the spec and agreed that details of a conformance test suite should be in another document. Alan to draft conformance section. 15/09: Alan had look at the conformance sections in XML and Schema specifications both of which indicate sections which must be implemented. None just say 'execute the test suite'. They talk in terms of conformance of document, schema and processors.. 22/09: no progress 22/09: Alan has added short Conformance and Optional Features sections to spec which was briefly discussed. Discussed naming for processors that don't implement optional features and those that implement all features. 06/10: Need to decide what/how test cases and certification process should occur
	114	**OGF 30** 25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session? 09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30. 08/09: Steve to request permission to go 15/09: Travel request has be submitted 22/09: DFDL session is scheduled at 11:00am Monday Oct 25th. 06/10: Steve has approval to travel.
	123	DFDL tutorial
	124	DFDL web content on OGF standards pages

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0	Update dfdl schema with change properties	ongoing	

3 6			
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**29-Sep-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Alan Powell on 30-Sep-10
Last Modified by Alan Powell on 30-Sep-10

OGF DFDL Working Group Call , September 29 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , September 29 2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Tim Kimber(IBM)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)
Bob McGrath (National Center for Supercomputing Applications)
Alejandro Rodriguez (National Center for Supercomputing Applications)

1. Current Actions
Updated Below

2. textNumberRoundingMode .
29/09: The interaction between textNumberRoundingMode and the rounding number in the numberPattern is not clearly described. It was agreed to make textNumberRoundingMode the controlling switch and add 'none' to the enumerations.

3. Syntax of assert /discriminator
29/09: Alan suggested that the value form of assert/discriminator be made the same as the element form of representation properties. Steve felt that they were not the same as format properties (eg they can have defaults) so should not have the same syntax.
Agreed there will be a new property
    dfdl:testKind 'expression' 'pattern'

Meeting closed, 16:35

Next call Wednesday 06 October 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee 's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **123**

Actions raised at this meeting

No	Action
121	**2. textNumberRoundingMode .** 29/09: The interaction between textNumberRoundingMode and the rounding number in the numberPattern is not clearly described. It was agreed to make textNumberRoundingMode the controlling switch and add 'none' to the enumerations.

	Closed
122	Syntax of assert /discriminator
	29/09: Alan suggested that the value form of assert/discriminator be made the same as the element form of representation properties. Steve felt that the assert attributes were not the same as format properties (eg they can have defaults) so should not have the same syntax.
	Agreed there will be a new property
	dfdl:testKind 'expression' 'pattern'
	Closed

Current Actions :

No	Action
066	Investigate format for defining test cases
	25/11:IBM to see if it is possible to publish its test case format.
	04/12: no update
	...
	17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite'
	24/03: No progress
	03/03: Discussions have been taking place on the subset of tests that will be provided.
	10/03: work is progressing
	17/03: work is progressing
	31/03: work is progressing
	14/04: And XML test case format has been defined and is being tested.
	21/04. Schema for TDML defined. Need to define how this and the test cases will be made public
	05/05: Work still progressing
	12/05: Work still progressing
	02/06: Work still progressing on technical and legal considerations
	...
	25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1'
	01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite.
	08/09: IBM still progressing
	15/09: IBM still progressing, expect tests to be available within a few weeks
	22/09: IBM still progressing, expect tests to be available within a few weeks
	29/09:Test cases are being prepared.
085	ALL: publicise Public comments phase to ensure a good review..
	14/04: see minutes
	21/04: Press release, OMG and other standards bodies.
	05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec
	15/05: still no public comments
	02/06: No public comments
	16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment.
	23/06: Still no comments. Alan will contact OGF to understand the rest of the process.
	30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet.
	07/07: No response. Alan will chase up

	14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification. 11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication. 11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'. Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process. 01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and closure of all actions. 08/09: Target for completion September 30. 15/09: as above 22/09: as above 29/09: Draft 43 will be published this week for WG review prior to submitting to OGF
111	**Daffodil DFDL parser** 11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented. We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats. Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil. Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed. It would be helpful if IBM test cases were available to Daffodil. IBM will investigate 25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential. Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list. Bob is planning to make the parser available on the internet to allow testing. His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way. 01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase. Daffodil being placed under formal source control in anticipation of external release. Bob has a start October deadline to create a report on what has been done for his sponsors. It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October. 08/09: Alejandro is marking up Spec draft 42 to indicate which features Daffodil implement. Bob expects Daffodil to be available on the web soon. 15/09: Alejandro had indicated in the specification which functions were implemented in Daffodill. Steve had reviewed and identified which function need to be implemented and which could be considered optional (see action 099). Alejandro is implementing the missing core functions. There was some discussion about the limitations on unordered groups. (stop value and expression not supported). It was agreed that it should be a schema definition error if dfdl:occursCountKind is 'stopValue' on any element within an unordered sequence and a floating element. 22/09: not discussed

112	**DFDL certification process** 25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process. 01/09: In progress, spec needs to state what conformance means, as part of this work 08/09: Discussed what needs to be said in the spec and agreed that details of a conformance test suite should be in another document. Alan to draft conformance section. 15/09: Alan had look at the conformance sections in XML and Schema specifications both of which indicate sections which must be implemented. None just say 'execute the test suite'. They talk in terms of conformance of document, schema and processors.. 22/09: no progress 22/09: Alan has added short Conformance and Optional Features sections to spec which was briefly discussed. Discussed naming for processors that don't implement optional features and those that implement all features.
114	**OGF 30** 25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session? 09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30. 08/09: Steve to request permission to go 15/09: Travel request has be submitted 22/09: DFDL session is scheduled at 11:00am Monday Oct 25th.

Closed actions

No	Action

107	**teston/testoff dfdl expression functions** .
	Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.
	04/08: Not discussed
	11/08: Not discussed
	25/08: Not discussed
	01/09: Steve to progress by Sept 30th
	08/09: Steve to progress by Sept 30th
	15/09: The ISO 8583 standard has existence flags at the beginning that are encoded so cannot be defined as an array of bits. Therefore DFDL needs the ability to set individual bits within an unsigned int. However the functions, particualry SetOn/Off, as currently defined are not correct. SetOn returns a byte with the relevant bit set on. This must then be combined with other bytes which isn't very usable. Steve to circulate example of use and suggested improvements.
	22/09: Steve had documented the why the functions were required to parse ISO 8583 messages. He had suggested the following improvements
	`xs:boolean dfdl:testBit(xs:unsignedByte, xs:unsignedByte)`
	
	Returns Boolean true if the bit number given by arg #2 is set on in the byte given by arg #1, otherwise returns Boolean false.
	
	`xs:unsignedByte dfdl:setBits(xs:boolean+)`
	
	Returns an byte being the value of the bit positions provided by the Boolean arguments, where true=1, false=0. The # of args must 8.
	 Note that the bit numbering goes from left to right, in accordance with section 12.3.7.2 of the spec.
	
	The type was changed from unSignedLong to unSignedByte to avoid problems with padding

	when not enough bits are provided. 29/09: Syntax agreed. Some errors in existing function descriptions need fixing. Closed
118	**2. Rules for 'missing' elements** lengthKind='implicit' and xs:maxLength or xs:length is "0": element is missing lengthKind='implicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='explicit' and length is an expression : element is missing if the expression evaluates to zero. lengthKind='explicit' and length is "0": element is missing lengthKind='explicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='pattern' : element is missing if the length of the pattern match is zero lengthKind='prefixed' : element is missing if the prefixed length region parses as zero lengthKind='delimited' and delimiterPolicy='suppressed' and the end of the group has not been encountered : element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='delimited' and delimiterPolicy='suppressed' or 'suppressedAtEnd' and the end of the group has been encountered : element is missing lengthKind='delimited', all other cases : element is missing if its scanned length is zero lengthKind='endOfParent': element is missing if its scanned length is zero It is a schema definition error to specify emptyValueDelimiterPolicy 'initiator' for lengthKind 'explicit' or 'implicit'. 22/09: We discussed if these rules could be simplified to say an element is missing if it's initiator is missing or its content region is empty. Need further discussion. 22/09: Tim suggested simplifying the above rules with An element is missing 1. if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data 2. else if the content region is empty Also: It is a schema definition error if a sequence has 'initiated content and one of its children has emptyValueDelimiterPolicy or nilValueInitiatorPolicy set to 'none' or 'terminator'. It is a schema definition error to specify emptyValueDelimiterPolicy 'initiator' or 'both' for lengthKind 'explicit' or 'implicit'. Closed
119	In passing we noted that the position of the prefix length relative to the initiator was not defined in the grammar and whether the prefix could have an initiator and terminator. Need further discussion 22/09:

	The prefix length region is between the initiator and the content region (leftFraming prefixLengthRegion simple/complexContent) The simpleType for the length prefix can specify any dfdl property with the exception of lengthKind 'prefixed' and 'endOfParent' Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
109	Add 'message' attribute to dfdl:discriminator 01/09: **Closed**: Conclusion was that this is genuinely useful, and has low implementation cost. Will add a 'message' attribute to dfdl:discriminator.	43	
110	Clarify expression limitations for defineVariable, newVariableInstance and setVariable 01/09: **Closed**: Spec should distinguish newVariableInstance defaultValue from setVariable value. For newVariableInstance defaultValue, disallow downward references and references to self (must be usable from the point of declaration) For setVariable allow downward references and references to self, and always evaluate at end of component. (defineVariable defaultValue should be same as newVariableInstance)	43	
113	Be specific about regular expression syntax	43	
108	Updates to hidden mechanism	43	
99	Updates to reflect subsetting and unparser optionality	43	
1	Define what conformance to spec means	43	

1 2			
1 1 5	Clarify allowed lengths for signed binary integers	43	
1 1 6	**2. xs:minLength** The spec currently states ● When an element declaration specifies a default value, and has type xs:string, then xs:minLength must be specified and must be 1 or greater. It is a schema definition error otherwise. The process for defaults and nils means this restriction is no longer needed. Agreed	43	
1 1 7	**3. Is UTF-16 a fixed width or variable width encoding** Proposal -UCS2 is a fixed length encoding -UTF-16 is a variable width encoding. - A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding 15/09: Closed	43	
1 1 8	**2. Document that an empty sequence that is the content of complex type is ignored even when it has annotations** It is a schema definition error if an empty sequence is the content of a complex type	43	
0 9 9	Splitting the specification in simpler sections. 22/09: We reviewed the proposed list of optional features and approved. These will be documented by adding a section that lists these features rather than making them inline. It will be closely related to the conformance section.. Closed	43	
1 0 1	**Semantics of 'fixed'** Proposal: - xs:fixed will not be used for parsing but only for validation and for providing a default value on unparsing. - A new dfdl function will be defined that applies only to simple element and tests whether the element exists including applying all the schema facets and other constraints. 22/09: Discussed whether dfdl:checkConstaints should included exists function. It isn't obvious what the return code should be for elements that don't exist. checkConstarints will check that element does exist. 'true' means the element exists and is valid, 'false' means doesn't exist or exists but doesn't meet constraints. The parameter is a path to a simple or complex element. If complex and it exists return 'true' Closed	43	
1 0 8	**dfdl:hidden** **Global group approach** Summary: ● Particle to hide can be a local element, element ref, local sequence, local choice or group ref ● Particle is removed from its parent into a dedicated global group of composition sequence and replaced in the parent by a new empty local sequence ● The new empty local sequence carries a dfdl:hiddenGroupRef property, other DFDL properties are not allowed Pros: ● Removal of all DFDL annotations and use of the resultant pure XSD results in same infoset ● Global group can be reused Cons: ● Making something hidden is a refactor operation	43	

	● Global group sequence needs DFDL properties setting correctly Alejandro had implemented extensions to the hidden function. 1. Allow hidden sequence to reference a global element. Decided against as Suman had identified some problems with namesapces. 2. Allow the reference global group to contain a choice in addition to a sequence. It was agreed this was a useful extension. 22/09: Closed	
1 1 3	**Regular Expressions .** 15/09: Agreed that should just say that either JAva or PERL regular expressions can be used and for portability the common subset of functions should be used. 22/09: Closed	43
1 1 3 b	**Regular Expressions for Assert /Discriminator .** Allowed as alternative to expression on dfdl:assert and dfdl:discriminator ● Pattern may be specified as attribute or element value ● Attribute: new testPattern attribute ● Element value: braces ( ) indicate pattern instead of expression 15/09: Do not need the braces as expressions start with '{'. Need to state rules for where the patter matching starts in the data stream. 22/09: Closed	43
1 1 5	**Clarify allowed lengths for signed integer types when rep is binary integer (ie, two's complement)** 01/09: No technical reason to restrict lengths to 2^x bytes, could be odd, could be bits. But rare in practise so if we do relax, limit any core subset to 2^x bytes. 22/09: Agreed that there should be not restrictions on lengths. Closed	43

1
0
7
	teston/testoff dfdl expression functions .
	
	22/09: Steve had documented the why the functions were required to parse ISO 8583 messages. He had suggested the following improvements
	`xs:boolean dfdl:testBit(xs:unsignedByte, xs:unsignedByte)`
	
	Returns Boolean true if the bit number given by arg #2 is set on in the byte given by arg #1, otherwise returns Boolean false.
	
	`xs:unsignedByte dfdl:setBits(xs:boolean+)`
	
	Returns an byte being the value of the bit positions provided by the Boolean arguments, where true=1, false=0. The # of args must 8.
	 Note that the bit numbering goes from left to right, in accordance with section 12.3.7.2 of the spec.
	
	The type was changed from unSignedLong to unSignedByte to avoid problems with padding when not enough bits are provided.
	29/09: Syntax agreed. Some errors in existing function descriptions need fixing. Closed	

1 1 8	**2. Rules for 'missing' elements** 22/09: Tim suggested simplifying the above rules with An element is missing 1. if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data 2. else if the content region is empty Also: It is a schema definition error if a sequence has 'initiated content and one of its children has	

	emptyValueDelimiterPolicy or nilValueInitiatorPolicy set to 'none' or 'terminator'. It is a schema definition error to specify emptyValueDelimiterPolicy 'initiator' or 'both' for lengthKind 'explicit' or 'implicit'. Closed		
1 1 9	In passing we noted that the position of the prefix length relative to the initiator was not defined in the grammar and whether the prefix could have an initiator and terminator. Need further discussion 22/09: The prefix length region is between the initiator and the content region (leftFraming prefixLengthRegion simple/complexContent) The simpleType for the length prefix can specify any dfdl property with the exception of lengthKind 'prefixed' and 'endOfParent' Closed		
1 2 1	2. textNumberRoundingMode . 29/09: The interaction between textNumberRoundingMode and the rounding number in the numberPattern is not clearly described. It was agreed to make textNumberRoundingMode the controlling switch and add 'none' to the enumerations.	43	
1 2 2	Syntax of assert /discriminator 29/09: Alan suggested that the value form of assert/discriminator be made the same as the element form of representation properties. Steve felt that the assert attributes were not the same as format properties (eg they can have defaults) so should not have the same syntax. Agreed there will be a new property dfdl:testKind 'expression' 'pattern'		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**22-Sep-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Alan Powell on 23-Sep-10
Last Modified by Alan Powell on 23-Sep-10

OGF DFDL Working Group Call , September 22 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , September 22 2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Tim Kimber(IBM)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)
Bob McGrath (National Center for Supercomputing Applications)
Alejandro Rodriguez (National Center for Supercomputing Applications)

1. Current Actions
Updated Below

2. Rules for 'missing' elements

lengthKind='implicit' and xs:maxLength or xs:length is "0": element is missing

lengthKind='implicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent)

lengthKind='explicit' and length is an expression : element is missing if the expression evaluates to zero.
lengthKind='explicit' and length is "0": element is missing
lengthKind='explicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent)

lengthKind='pattern' : element is missing if the length of the pattern match is zero
lengthKind='prefixed' : element is missing if the prefixed length region parses as zero

lengthKind='delimited' and delimiterPolicy='suppressed' and the end of the group has not been encountered : element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent)
lengthKind='delimited' and delimiterPolicy='suppressed' or 'suppressedAtEnd' and the end of the group has been encountered : element is missing
lengthKind='delimited', all other cases : element is missing if its scanned length is zero

lengthKind='endOfParent': element is missing if its scanned length is zero

It is a schema definition error to specify emptyValueDelimiterPolicy 'initiator' for lengthKind 'explicit' or 'implicit'.

22/09: We discussed if these rules could be simplified to say an element is missing if it 's initiator is missing or its content region is empty. Need further discussion.

In passing we noted that the position of the prefix length relative to the initiator was not defined in the grammar and whether the prefix could have an initiator and terminator. Need further discussion

Meeting closed, 16:35

Next call Wednesday 29 September 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **120**

Actions raised at this meeting

No	Action
118	**2. Rules for 'missing' elements** lengthKind='implicit' and xs:maxLength or xs:length is "0": element is missing lengthKind='implicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='explicit' and length is an expression : element is missing if the expression evaluates to zero. lengthKind='explicit' and length is "0": element is missing lengthKind='explicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='pattern' : element is missing if the length of the pattern match is zero lengthKind='prefixed' : element is missing if the prefixed length region parses as zero lengthKind='delimited' and delimiterPolicy='suppressed' and the end of the group has not been encountered : element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='delimited' and delimiterPolicy='suppressed' or 'suppressedAtEnd' and the end of the

group has been encountered : element is missing

lengthKind='delimited', all other cases : element is missing if its scanned length is zero

lengthKind='endOfParent': element is missing if its scanned length is zero

It is a schema definition error to specify emptyValueDelimiterPolicy 'initiator' for lengthKind 'explicit' or 'implicit'.

22/09: We discussed if these rules could be simplified to say an element is missing if it's initiator is missing or its content region is empty. Need further discussion.

| 119 | Clarify position and restriction on prefix length element. |
| | In passing we noted that the position of the prefix length relative to the initiator was not defined in the grammar and whether the prefix could have an initiator and terminator. Need further discussion |

Current Actions :

No	Action
066	Investigate format for defining test cases
	25/11:IBM to see if it is possible to publish its test case format.
	04/12: no update
	...
	17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite'
	24/03: No progress
	03/03: Discussions have been taking place on the subset of tests that will be provided.
	10/03: work is progressing
	17/03: work is progressing
	31/03: work is progressing
	14/04: And XML test case format has been defined and is being tested.
	21/04. Schema for TDML defined. Need to define how this and the test cases will be made public
	05/05: Work still progressing
	12/05: Work still progressing
	02/06: Work still progressing on technical and legal considerations
	...
	25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1'
	01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite.
	08/09: IBM still progressing
	15/09: IBM still progressing, expect tests to be available within a few weeks
	22/09: IBM still progressing, expect tests to be available within a few weeks
085	ALL: publicise Public comments phase to ensure a good review..
	14/04: see minutes
	21/04: Press release, OMG and other standards bodies.
	05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec
	15/05: still no public comments
	02/06: No public comments
	16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve

also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment.
23/06: Still no comments. Alan will contact OGF to understand the rest of the process.
30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet.
07/07: No response. Alan will chase up
14/07: No response from Joel. Sent email to Greg Newby by no response.
21/07: Still no response.
04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification.
11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication.
11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'.
Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process.
01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and closure of all actions.
08/09: Target for completion September 30.
15/09: as above
22/09: as above

| 107 | **teston/testoff dfdl expression functions** .
Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.
04/08: Not discussed
11/08: Not discussed
25/08: Not discussed
01/09: Steve to progress by Sept 30th
08/09: Steve to progress by Sept 30th
15/09: The ISO 8583 standard has existence flags at the beginning that are encoded so cannot be defined as an array of bits. Therefore DFDL needs the ability to set individual bits within an unsigned int. However the functions, particualry SetOn/Off, as currently defined are not correct. SetOn returns a byte with the relevant bit set on. This must then be combined with other bytes which isn't very usable. Steve to circulate example of use and suggested improvements.
22/09: Steve had documented the why the functions were required to parse ISO 8583 messages. He had suggested the following improvements
`xs:boolean dfdl:testBit(xs:unsignedByte, xs:unsignedByte)`

     Returns Boolean true if the bit number given by arg #2 is set on in the byte given by arg #1, otherwise returns Boolean false.

`xs:unsignedByte dfdl:setBits(xs:boolean+)`

     Returns an byte being the value of the bit positions provided by the Boolean arguments, where true=1, false=0. The # of args must 8.
 Note that the bit numbering goes from left to right, in accordance with section 12.3.7.2 of the spec.

The type was changed from unSignedLong to unSignedByte to avoid problems with padding |

111	**Daffodil DFDL parser** 11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented. We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats. Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil. Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed. It would be helpful if IBM test cases were available to Daffodil. IBM will investigate 25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential. Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list. Bob is planning to make the parser available on the internet to allow testing. His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way. 01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase. Daffodil being placed under formal source control in anticipation of external release. Bob has a start October deadline to create a report on what has been done for his sponsors. It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October. 08/09: Alejandro is marking up Spec draft 42 to indicate which features Daffodil implement. Bob expects Daffodil to be available on the web soon. 15/09: Alejandro had indicated in the specification which functions were implemented in Daffodill. Steve had reviewed and identified which function need to be implemented and which could be considered optional (see action 099). Alejandro is implementing the missing core functions. There was some discussion about the limitations on unordered groups. (stop value and expression not supported). It was agreed that it should be a schema definition error if dfdl:occursCountKind is 'stopValue' on any element within an unordered sequence and a floating element.
112	**DFDL certification process** 25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process. 01/09: In progress, spec needs to state what conformance means, as part of this work 08/09: Discussed what needs to be said in the spec and agreed that details of a conformance test suite should be in another document. Alan to draft conformance section. 15/09: Alan had look at the conformance sections in XML and Schema specifications both of which indicate sections which must be implemented. None just say 'execute the test suite'. They talk in terms of conformance of document, schema and processors..
114	**OGF 30** 25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session? 09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30. 08/09: Steve to request permission to go 15/09: Travel request has be submitted
115	**Clarify allowed lengths for signed integer types when rep is binary integer (ie, two's complement)**

	01/09: No technical reason to restrict lengths to 2^x bytes, could be odd, could be bits. But rare in practise so if we do relax, limit any core subset to 2^x bytes. 08/09: not discussed 15/09: not discussed 22/09: Agreed that there should be not restrictions on lengths. Closed
118	2. Rules for 'missing' elements lengthKind='implicit' and xs:maxLength or xs:length is "0": element is missing lengthKind='implicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='explicit' and length is an expression : element is missing if the expression evaluates to zero. lengthKind='explicit' and length is "0": element is missing lengthKind='explicit' and length is not "0": element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='pattern' : element is missing if the length of the pattern match is zero lengthKind='prefixed' : element is missing if the prefixed length region parses as zero lengthKind='delimited' and delimiterPolicy='suppressed' and the end of the group has not been encountered : element is missing if it has an initiator AND emptyValueDelimiterPolicy excludes the initiator AND the initiator is not found in the data (regardless of discriminators or initiatedContent) lengthKind='delimited' and delimiterPolicy='suppressed' or 'suppressedAtEnd' and the end of the group has been encountered : element is missing lengthKind='delimited', all other cases : element is missing if its scanned length is zero lengthKind='endOfParent': element is missing if its scanned length is zero It is a schema definition error to specify emptyValueDelimiterPolicy 'initiator' for lengthKind 'explicit' or 'implicit'. 22/09: We discussed if these rules could be simplified to say an element is missing if it 's initiator is missing or its content region is empty. Need further discussion.
119	In passing we noted that the position of the prefix length relative to the initiator was not defined in the grammar and whether the prefix could have an initiator and terminator. Need further discussion

Closed actions

No	Action
117	3. Is UTF-16 a fixed width or variable width encoding
	Appendix A: About UTF-16 and Unicode Character Codes above 0xFFFF When we define UTF-16 to be a fixed-width double-byte wide character set we say that each UTF-16 codepoint is represented by 2 bytes. Notice the careful use of the term 'codepoint' here. Unicode/ISO10646 characters can have character codes as large as 0x10FFFF which requires 3 bytes to store (21 bits actually); however in UTF-16 characters with more than 2 bytes of code are encoded as two codepoints, called a surrogate pair; hence, UTF-16 is fixed-width, 2 bytes

per *codepoint* . It is not 2 bytes per Unicode character. UTF-16 is really a variable-width encoding, but the characters that require the surrogate-pair treatment are so infrequently used that UTF-16 is most often treated like a 16-bit fixed-width character set. It is the acknowledgement of the existence of surrogate pairs that leads to the "codepoint" vs. "character code" distinction.

UTF-32 is a fixed width encoding with a full 4-bytes per character code. It represents all of Unicode with the same width per character.

Hence, when we refer to lengths in character strings we will often refer to length in characters, but we qualify that it means 2-byte *codepoints* when the character set encoding is UTF-16. Hence, when the property lengthUnitKind is 'characters' and the charset is 'UTF-16', then the units are actually 16-bit codepoints, not Unicode characters.

Proposal

-UCS2 is a fixed length encoding

-UTF-16 is a variable width encoding.

- A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding

15/09: Proposal agreed. Closed

099	Splitting the specification in simpler sections . 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call . 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset. 21/07: no progress 04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function . Purpose of subsetting is: 1. Allow simpler implementations . (main purpose) 2. Simplify tooling 3. Simplify specification . Steve to contact previous members of WG to check if we have the correct subsets 11/08: Steve sent an email to previous members of the WG asking for opinions on splitting the specification. Bob McGrath from National Center For Supercomputing responded that they had implemented about 80% of the function. Alejandro will send a description of the function they have implemented. Action will be raised to track the Daffodil implementation 11/08: not discussed 01/09: NCSA implementation description received. Making the unparser optional is a good idea (NCSA do not need one) . Work will progress on the subsets. 08/09: No progress 15/09: Steve proposed making 'obscure' properties optional rather than subsetting parts of the specification. See minutes for proposed list of optional properties. 22/09: We reviewed the proposed list of optional features and approved. These will be documented by adding a section that lists these features rather than making them inline. It will be closely related to the conformance section.. Closed
101	**Semantics of 'fixed'** 21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error . It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration. Suggestions - a parser option that 'converted all validation errors to processing errors '

- a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element.

Stephanie will produce some examples of how this could be used..

04/08: Stephanie had produced examples but they were not discussed due to lack of time

11/08: We started to discuss Stephanie's HIPPA example but ran out of time.

25/08: Not discussed

01/09: Discuss next week

08/09: Stephanie sent an example of an X12 document showing how an element with the same name was defined in different groups with different enumerations.

Proposal:

- xs:fixed will not be used for parsing but only for validation and for providing a default value on unparsing.

- A new dfdl function will be defined that applies only to simple element and tests whether the element exists including applying all the schema facets. (need to check with Tim why he wanted to only apply enumerations)

dfdl:exists(xpath , **true ¦ false**) true means apply facets, false means don'e apply facets.

```
                          <xs:element ref="REF_BillingProviderTaxIdentification_2010AA">
                              <xs:annotation>
                                    <xs:documentation>Discrimination needed to
distinguish REF segments</xs:documentation>
                                    <xs:appinfo source="http://www.ogf.org/dfdl/">
                                        <dfdl:discriminator
test="{dfdl:exists(./REF01__ReferenceIdentificationQualifier, true)}"/>
                                    </xs:appinfo>
                              </xs:annotation>
```

15/09: Decided that a separate dfdl:checkConstraints (or other suitable name) function, that checks all constraints not just facets, was simpler than extending fn:exists. It applies to both simple and complex element and to groups

22/09: Discussed whether dfdl:checkConstaints should included exists function. It isn't obvious what the return code should be for elements that don't exist. checkConstarints will check that element does exist. 'true' means the element exists and is valid, 'false' means doesn't exist or exists but doesn't meet constraints. The parameter is a path to a simple or complex element. If complex and it exists return 'true'

Closed

| 108 | **dfdl:hidden**
There has been some discussion on whether the 'hidden' global group should be indicated in some way.
04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.
25/08: there has been some offline discussions about simplifying how hidden elements are implemented. The proposal is
  • dfdl:hidden property on xs:element only
  • xs:minOccurs and xs:maxOccurs MUST be 0 when hidden
  • dfdl:minOccurs and dfdl:maxOccurs for hidden elements only.
  • An element is 'required' when dfdl:minOccurs >0 and normal default processing occurs.
The schema, without dfdl annotations, must match the infoset so assumption is that non-DFDL tools, such as mappers, will ignore/not show elements with xs:minOccurs and xs:maxOccurs = '0'
01/09: The above proposal is flawed due to use of maxOccurs = 0 (this was identified back in 2008 hence current spec).
Bob confirmed that NCSA models use hidden in a big way, so punting hidden beyond 1.0 is not |

an option.
Two candidates:
- As per spec but with syntactic improvements to make it clear that the two xs:sequences do not take any dfdl:sequence properties
- Place a flag directly on a local element and force minOccurs to be 0. Simpler syntax but the semantic changes, as the element *could* be legally in the infoset, although a DFDL parser would never put it there.
Steve will circulate the two proposals for next week.
Bob to talk to Alejandro as the NCSA implementation is currently more flexible than the spec, allowing the groupref to point to a choice, and an elementref. Are these really needed?
08/09: Discussed the Global Group and Hidden Flag approaches.
Decided to stay with Global Group with dfdl:sequence properties rather than the dfdl:hidden annotation. It was agreed that there would be no extra properties on the 'hidden' global group as the syntax was messy as it should really be on the sequence and there are currently no dfdl properties on global groups.

Global group approach
Summary:
- Particle to hide can be a local element, element ref, local sequence, local choice or group ref
- Particle is removed from its parent into a dedicated global group of composition sequence and replaced in the parent by a new empty local sequence
- The new empty local sequence carries a dfdl:hiddenGroupRef property, other DFDL properties are not allowed

Pros:
- Removal of all DFDL annotations and use of the resultant pure XSD results in same infoset
- Global group can be reused

Cons:
- Making something hidden is a refactor operation
- Global group sequence needs DFDL properties setting correctly

The Daffodil parser allows the hidden annotation to reference global elements in addition to global groups. It was noted that this lost the particle properties but we need to discuss with Alejandro.
15/09: Alan circulated section describing Hidden Sequence Groups. Noted that hidden is now only described in the sequence sections of the specficiation.
Noted some editorial changes.
Alejandro had implemented extensions to the hidden function.
1. Allow hidden sequence to reference a global element. Decided against as Suman had identified some problems with namesapces.
2. Allow the reference global group to contain a choice in addition to a sequence. It was agreed this was a useful extension.
22/09: Closed

113	**Regular Expressions .**

25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?
a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema.
01/09: There are many variations of regexp language, it seems wise to specify one that we know contains functions like lookaround, which makes it easy to say things like 'give me everything up to but not including x'. This rules out XML Schema and POSIX, it needs Perl 5 or Java.
08/09: Agreed that specification should define the regular expression language (if only by referring to other specifications) .
Should allow a common subset of PERL and Java expressions languages. Alan to update regular expression section.
15/09: Agreed that should just say that either JAva or PERL regular expressions can be used and for portability the common subset of functions should be used.

	22/09: Closed	

113b	**Regular Expressions for Assert /Discriminator .**
	25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?
	b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both.
	01/09: Tim to convince Steve (via example) that use of regexp in asserts is needed in 1.0.
	08/09: Agreed that this is a useful function
	Allowed as alternative to expression on dfdl:assert and dfdl:discriminator
	  ● Pattern may be specified as attribute or element value
	  ● Attribute: new testPattern attribute
	  ● Element value: braces ( ) indicate pattern instead of expression
	15/09: Do not need the braces as expressions start with '{'. Need to state rules for where the patter matching starts in the data stream.
	22/09: Closed	

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
109	Add 'message' attribute to dfdl:discriminator 01/09: **Closed**: Conclusion was that this is genuinely useful, and has low implementation cost. Will add a 'message' attribute to dfdl:discriminator.	43	not
110	Clarify expression limitations for defineVariable , newVariableInstance and setVariable 01/09: **Closed**: Spec should distinguish newVariableInstance defaultValue from setVariable value. For newVariableInstance defaultValue, disallow downward references and references to self (must be usable from the point of declaration) For setVariable allow downward references and references to self, and always evaluate at end of component. (defineVariable defaultValue should be same as newVariableInstance)	43	not
113	Be specific about regular expression syntax	43	not

108	Updates to hidden mechanism	43	not
99	Updates to reflect subsetting and unparser optionality	43	not
112	Define what conformance to spec means	43	not
115	Clarify allowed lengths for signed binary integers	43	not
116	2. xs:minLength The spec currently states ● When an element declaration specifies a default value, and has type xs:string, then xs:minLength must be specified and must be 1 or greater. It is a schema definition error otherwise. The process for defaults and nils means this restriction is no longer needed. Agreed		
117	3. Is UTF-16 a fixed width or variable width encoding Proposal -UCS2 is a fixed length encoding -UTF-16 is a variable width encoding. - A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding 15/09: Closed		
118	2. Document that an empty sequence that is the content of complex type is ignored even when it has annotations It is a schema definition error if an empty sequence is the content of a complex type		
099	Splitting the specification in simpler sections. 22/09: We reviewed the proposed list of optional features and approved. These will be documented by adding a section that lists these features rather than making them inline. It will be closely related to the conformance section.. Closed		
101	Semantics of 'fixed' Proposal: - xs:fixed will not be used for parsing but only for validation and for providing a default value on unparsing. - A new dfdl function will be defined that applies only to simple element and tests whether the element exists including applying all the schema facets and other constraints. 22/09: Discussed whether dfdl:checkConstaints should included exists function. It isn't obvious what the return code should be for elements that don't exist. checkConstarints will check that element does exist. 'true' means the element exists and is valid, 'false' means doesn't exist or exists but doesn't meet constraints. The parameter is a path to a simple or complex element. If complex and it exists return 'true' Closed		
108	dfdl:hidden Global group approach Summary: ● Particle to hide can be a local element, element ref, local sequence, local choice or group ref ● Particle is removed from its parent into a dedicated global group of composition sequence and replaced in the parent by a new empty local sequence ● The new empty local sequence carries a dfdl:hiddenGroupRef property, other DFDL properties		

	are not allowed Pros: • Removal of all DFDL annotations and use of the resultant pure XSD results in same infoset • Global group can be reused Cons: • Making something hidden is a refactor operation • Global group sequence needs DFDL properties setting correctly Alejandro had implemented extensions to the hidden function. 1. Allow hidden sequence to reference a global element. Decided against as Suman had identified some problems with namesapces. 2. Allow the reference global group to contain a choice in addition to a sequence. It was agreed this was a useful extension. 22/09: Closed		
1 1 3	**Regular Expressions .** 15/09: Agreed that should just say that either JAva or PERL regular expressions can be used and for portability the common subset of functions should be used. 22/09: Closed		
1 1 3 b	**Regular Expressions for Assert /Discriminator .** Allowed as alternative to expression on dfdl:assert and dfdl:discriminator • Pattern may be specified as attribute or element value • Attribute: new testPattern attribute • Element value: braces ( ) indicate pattern instead of expression 15/09: Do not need the braces as expressions start with '{'. Need to state rules for where the patter matching starts in the data stream. 22/09: Closed		
1 1 5	**Clarify allowed lengths for signed integer types when rep is binary integer (ie, two's complement)** 01/09: No technical reason to restrict lengths to 2^x bytes, could be odd, could be bits. But rare in practise so if we do relax, limit any core subset to 2^x bytes. 22/09: Agreed that there should be not restrictions on lengths. Closed		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**15-Sep-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Alan Powell on 16-Sep-10
Last Modified by Alan Powell on 17-Sep-10

OGF DFDL Working Group Call , September 15 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , September 15-2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Bob McGrath (National Center for Supercomputing Applications)
Alan Powell (IBM)
Tim Kimber(IBM)
Alejandro Rodriguez (National Center for Supercomputing Applications)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)

1. Current Actions
Updated Below

Action 099 Proposed optional properties

Feature	Detection	Pre-reqs	NCSA ?	Comments
Validation	External switch	None, but of limited	No	

		value without Simple type restrictions		
Simple type restrictions	xs:simpleType in xsd	None	No	
Nils	xs:nillable='yes' in xsd	None	No	
Defaults	xs:default or xs:fixed in xsd	None	No	
Bi-di	dfdl:textBiDi='yes'	None	No	
Bits	dfdl:alignmentUnits='bits' or dfdl:lengthUnits='bits'	None	No	
Delimited binary	dfdl:representation='binary' (or implied binary) and dfdl:lengthKind='delimited'	None	No	
Patterns	dfdl:lengthKind='pattern' or dfdl:assert testPattern	None	Yes	
Zoned numbers	dfdl:textNumberRep='zoned'	None	No	z/OS
Packed numbers	dfdl:binaryNumberRep='packed'	None	Yes	z/OS
Packed calendars	dfdl:binaryCalendarRep='packed'	None		z/OS
S/390 floats	dfdl:binaryFloatRep='ibm390Hex'	None	No	z/OS
Unordered	dfdl:sequenceKind='unordered'	None	Not clear	
Floating	dfdl:floating='yes'	None	No	
dfdl functions in expression language	dfdl: functions in expression	None	No	Enables use of off-the-shelf XPath 2.0 package
Hidden	dfdl:hiddenRef <> "	None	Yes	
Calculated values	dfdl:inputValueCalc <> " or dfdl:outputValueCalc <> "	None	Yes	
Escape schemes	dfd;defineEscapeScheme in xsd	None	Yes	
Encodings	Any dfdl:encoding value beyond the core list	None	Not clear	
Asserts	dfdl:assert in xsd	None	Yes	
Discriminators	dfdl:discriminator in xsd	None	Yes	
Prefixed lengths	dfdl:lengthKind='prefixed'	Simple type restrictions	Not clear	Requires simple types
Variables	Variables in expression language	None	Yes	

2. Document that an empty sequence that is the content of complex type is ignored even when it has annotations

One thing to point out is that the authors should avoid

```
<xs:complexType>
  <xs:sequence dfdl:hiddenGroupRef="..."/>
</xs:complexType>
```

(The same applies to other annotations on sequences, long- or short-form.) The schema spec will discard that sequence (see [1] definition of "effective content" clause 2.1.2). The following works:

```
<xs:complexType>
  <xs:sequence>
    <xs:sequence dfdl:hiddenGroupRef="..."/>
  </xs:sequence>
</xs:complexType>
```

[1] http://www.w3.org/TR/2004/REC-xmlschema-1-20041028/#key-exg

It is a schema definition error if an empty sequence is the content of a complex type

Meeting closed, 16:35

Next call Wednesday 22 September 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **118**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 25/08: Will chase to allow Daffodil access to test cases. The WG should define how

	implementation confirm that they 'conform to DFDL v1' 01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite. 08/09: IBM still progressing 15/09: IBM still progressing, expect tests to be available within a few weeks
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification. 11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication. 11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'. Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process. 01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and closure of all actions. 08/09: Target for completion September 30. 15/09: as above
099	Splitting the specification in simpler sections. 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call. 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset. 21/07: no progress 04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function. Purpose of subsetting is: 1. Allow simpler implementations. (main purpose) 2. Simplify tooling 3. Simplify specification. Steve to contact previous members of WG to check if we have the correct subsets

	11/08: Steve sent an email to previous members of the WG asking for opinions on splitting the specification. Bob McGrath from National Center For Supercomputing responded that they had implemented about 80% of the function. Alejandro will send a description of the function they have implemented. Action will be raised to track the Daffodil implementation 11/08: not discussed 01/09: NCSA implementation description received. Making the unparser optional is a good idea (NCSA do not need one) . Work will progress on the subsets. 08/09: No progress 15/09: Steve proposed making 'obscure' properties optional rather than subsetting parts of the specification. See minutes for proposed list of optional properties.

101	**Semantics of 'fixed'**
	21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error.
	It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration.
	Suggestions
	- a parser option that 'converted all validation errors to processing errors'
	- a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element.
	Stephanie will produce some examples of how this could be used..
	04/08: Stephanie had produced examples but they were not discussed due to lack of time
	11/08: We started to discuss Stephanie's HIPPA example but ran out of time.
	25/08: Not discussed
	01/09: Discuss next week
	08/09: Stephanie sent an example of an X12 document showing how an element with the same name was defined in different groups with different enumerations.
	Proposal:
	- xs:fixed will not be used for parsing but only for validation and for providing a default value on unparsing.
	
	- A new dfdl function will be defined that applies only to simple element and tests whether the element exists including applying all the schema facets. (need to check with Tim why he wanted to only apply enumerations)
	dfdl:exists(xpath , **true ¦ false**) true means apply facets, false means don'e apply facets.
	
	`<xs:element ref="REF_BillingProviderTaxIdentification_2010AA">`
	`<xs:annotation>`
	`<xs:documentation>`Discrimination needed to distinguish REF segments`</xs:documentation>`
	`<xs:appinfo source="http://www.ogf.org/dfdl/">`
	`<dfdl:discriminator test="{dfdl:exists(./REF01__ReferenceIdentificationQualifier, `**true**`)}"/>`
	`</xs:appinfo>`
	`</xs:annotation>`
	15/09: Decided that a separate dfdl:checkConstraints (or other suitable name) function, that checks all constraints not just facets, was simpler than extending fn:exists. It applies to both simple and complex element and to groups

107	**teston/testoff dfdl expression functions .** Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required. 04/08: Not discussed 11/08: Not discussed 25/08: Not discussed 01/09: Steve to progress by Sept 30th

108	dfdl:hidden

dfdl:hidden

There has been some discussion on whether the 'hidden' global group should be indicated in some way.

04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.

25/08: there has been some offline discussions about simplifying how hidden elements are implemented. The proposal is

- dfdl:hidden property on xs:element only
- xs:minOccurs and xs:maxOccurs MUST be 0 when hidden
- dfdl:minOccurs and dfdl:maxOccurs for hidden elements only.
- An element is 'required' when dfdl:minOccurs >0 and normal default processing occurs.

The schema, without dfdl annotations, must match the infoset so assumption is that non-DFDL tools, such as mappers, will ignore/not show elements with xs:minOccurs and xs:maxOccurs = '0'

01/09: The above proposal is flawed due to use of maxOccurs = 0 (this was identified back in 2008 hence current spec).

Bob confirmed that NCSA models use hidden in a big way, so punting hidden beyond 1.0 is not an option.

Two candidates:

- As per spec but with syntactic improvements to make it clear that the two xs:sequences do not take any dfdl:sequence properties
- Place a flag directly on a local element and force minOccurs to be 0. Simpler syntax but the semantic changes, as the element *could* be legally in the infoset, although a DFDL parser would never put it there.

Steve will circulate the two proposals for next week.

Bob to talk to Alejandro as the NCSA implementation is currently more flexible than the spec, allowing the groupref to point to a choice, and an elementref. Are these really needed?

08/09: Discussed the Global Group and Hidden Flag approaches.

Decided to stay with Global Group with dfdl:sequence properties rather than the dfdl:hidden annotation. It was agreed that there would be no extra properties on the 'hidden' global group as the syntax was messy as it should really be on the sequence and there are currently no dfdl properties on global groups.

Global group approach

Summary:

- Particle to hide can be a local element, element ref, local sequence, local choice or group ref
- Particle is removed from its parent into a dedicated global group of composition sequence and replaced in the parent by a new empty local sequence
- The new empty local sequence carries a dfdl:hiddenGroupRef property, other DFDL properties are not allowed

Pros:

- Removal of all DFDL annotations and use of the resultant pure XSD results in same infoset
- Global group can be reused

Cons:

- Making something hidden is a refactor operation
- Global group sequence needs DFDL properties setting correctly

The Daffodil parser allows the hidden annotation to reference global elements in addition to

		global groups. It was noted that this lost the particle properties but we need to discuss with Alejandro.
		15/09: Alan circulated section describing Hidden Sequence Groups. Noted that hidden is now only described in the sequence sections of the specficiation.
		Noted some editorial changes.
		Alejandro had implemented extensions to the hidden function.
		1. Allow hidden sequence to reference a global element. Decided against as Suman had identified some problems with namesapces.
		2. Allow the reference global group to contain a choice in addition to a sequence. It was agreed this was a useful extension.
111		**Daffodil DFDL parser**
		11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented.
		We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats.
		Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil.
		Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed.
		It would be helpful if IBM test cases were available to Daffodil. IBM will investigate
		25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential.
		Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list.
		Bob is planning to make the parser available on the internet to allow testing.
		His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way.
		01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase.
		Daffodil being placed under formal source control in anticipation of external release.
		Bob has a start October deadline to create a report on what has been done for his sponsors.
		It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October.
		08/09: Alejandro is marking up Spec draft 42 to indicate which features Daffodil implement. Bob expects Daffodil to be available on the web soon.
		15/09: Alejandro had indicated in the specification which functions were implemented in Daffodill. Steve had reviewed and identified which function need to be implemented and which could be considered optional (see action 099). Alejandro is implementing the missing core functions. There was some discussion about the limitations on unordered groups. (stop value and expression not supported). It was agreed that it should be a schema definition error if dfdl:occursCountKind is 'stopValue' on any element within an unordered sequence and a floating element.
112		**DFDL certification process**
		25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process.
		01/09: In progress, spec needs to state what conformance means, as part of this work
		08/09: Discussed what needs to be said in the spec and agreed that details of a conformance test suite should be in another document.
		Alan to draft conformance section.
		15/09: Alan had look at the conformance sections in XML and Schema specifications both of which indicate sections which must be implemented. None just say 'execute the test suite'. They talk in terms of conformance of document, schema and processors..
113		**Regular Expressions .**

	25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?
	a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema.
	01/09: There are many variations of regexp language, it seems wise to specify one that we know contains functions like lookaround, which makes it easy to say things like 'give me everything up to but not including x'. This rules out XML Schema and POSIX, it needs Perl 5 or Java.
	08/09: Agreed that specification should define the regular expression language (if only by referring to other specifications).
	Should allow a common subset of PERL and Java expressions languages. Alan to update regular expression section.
	15/09: Agreed that should just say that either JAva or PERL regular expressions can be used and for portability the common subset of functions should be used.
113b	**Regular Expressions for Assert /Discriminator .**
	25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?
	b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both.
	01/09: Tim to convince Steve (via example) that use of regexp in asserts is needed in 1.0.
	08/09: Agreed that this is a useful function
	Allowed as alternative to expression on dfdl:assert and dfdl:discriminator
	• Pattern may be specified as attribute or element value
	• Attribute: new testPattern attribute
	• Element value: braces ( ) indicate pattern instead of expression
	15/09: Do not need the braces as expressions start with '{'. Need to state rules for where the patter matching starts in the data stream.
114	**OGF 30**
	25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session?
	09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30.
	08/09: Steve to request permission to go
	15/09: Travel request has be submitted
115	**Clarify allowed lengths for signed integer types when rep is binary integer (ie, two's complement)**
	01/09: No technical reason to restrict lengths to 2^x bytes, could be odd, could be bits. But rare in practise so if we do relax, limit any core subset to 2^x bytes.
	08/09: not discussed
	15/09: not discussed

Closed actions

No	Action
117	**3. Is UTF-16 a fixed width or variable width encoding**
	Appendix A: About UTF-16 and Unicode Character Codes above 0xFFFF
	When we define UTF-16 to be a fixed-width double-byte wide character set we say that each UTF-16 codepoint is represented by 2 bytes. Notice the careful use of the term 'codepoint' here. Unicode/ISO10646 characters can have character codes as large as 0x10FFFF which requires 3 bytes to store (21 bits actually); however in UTF-16 characters with more than 2 bytes of code are encoded as two codepoints, called a surrogate pair; hence, UTF-16 is fixed-width, 2 bytes per *codepoint* . It is not 2 bytes per Unicode character. UTF-16 is really a variable-width encoding, but the characters that require the surrogate-pair treatment are so infrequently used that UTF-16 is most often treated like a 16-bit fixed-width character set. It is the

acknowledgement of the existence of surrogate pairs that leads to the "codepoint" vs. "character code" distinction.

UTF-32 is a fixed width encoding with a full 4-bytes per character code. It represents all of Unicode with the same width per character.

Hence, when we refer to lengths in character strings we will often refer to length in characters, but we qualify that it means 2-byte *codepoints* when the character set encoding is UTF-16.

Hence, when the property lengthUnitKind is 'characters' and the charset is 'UTF-16', then the units are actually 16-bit codepoints, not Unicode characters.

Proposal

-UCS2 is a fixed length encoding

-UTF-16 is a variable width encoding.

- A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding

15/09: Closed

| 118 | **2. Document that an empty sequence that is the content of complex type is ignored even when it has annotations**

One thing to point out is that the authors should avoid

<xs:complexType>
  <xs:sequence dfdl:hiddenGroupRef="..."/>
</xs:complexType>

(The same applies to other annotations on sequences, long- or short-form.) The schema spec will discard that sequence (see [1] definition of "effective content" clause 2.1.2). The following works:

<xs:complexType>
  <xs:sequence>
    <xs:sequence dfdl:hiddenGroupRef="..."/>
  </xs:sequence>
</xs:complexType>

[1] http://www.w3.org/TR/2004/REC-xmlschema-1-20041028/#key-exg

It is a schema definition error if an empty sequence is the content of a complex type |

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not

070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
109	Add 'message' attribute to dfdl:discriminator 01/09: **Closed**: Conclusion was that this is genuinely useful, and has low implementation cost. Will add a 'message' attribute to dfdl:discriminator.	43	not
110	Clarify expression limitations for defineVariable, newVariableInstance and setVariable 01/09: **Closed**: Spec should distinguish newVariableInstance defaultValue from setVariable value. For newVariableInstance defaultValue, disallow downward references and references to self (must be usable from the point of declaration) For setVariable allow downward references and references to self, and always evaluate at end of component. (defineVariable defaultValue should be same as newVariableInstance)	43	not
113	Be specific about regular expression syntax	43	not
1008	Updates to hidden mechanism	43	not
99	Updates to reflect subsetting and unparser optionality	43	not
112	Define what conformance to spec means	43	not
115	Clarify allowed lengths for signed binary integers	43	not
116	**2. xs:minLength** The spec currently states ● When an element declaration specifies a default value, and has type xs:string, then xs:minLength must be specified and must be 1 or greater. It is a schema definition error otherwise. The process for defaults and nils means this restriction is no longer needed. Agreed		
117	**3. Is UTF-16 a fixed width or variable width encoding** Proposal -UCS2 is a fixed length encoding -UTF-16 is a variable width encoding. - A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding 15/09: Closed		
118	**2. Document that an empty sequence that is the content of complex type is ignored even when it has annotations** It is a schema definition error if an empty sequence is the content of a complex type		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**08-Sep-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Alan Powell on 10-Sep-10
Last Modified by Alan Powell on 10-Sep-10

OGF DFDL Working Group Call , September 8 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , September 8-2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Bob McGrath (National Center for Supercomputing Applications)
Alan Powell (IBM)
Suman Kalia (IBM)

Apologies
Mike Beckerle (Oco)
Tim Kimber(IBM)
Alejandro Rodriguez (National Center for Supercomputing Applications)

1. Current Actions
Updated Below

2. xs:minLength
The spec currently states

- When an element declaration specifies a default value, and has type xs:string, then xs:minLength must be specified and must be 1 or greater. It is a schema definition error otherwise.

The process for defaults and nils means this restriction is no longer needed.

Agreed Closed

3. Is UTF-16 a fixed width or variable width encoding

Appendix A: About UTF-16 and Unicode Character Codes above 0xFFFF

When we define UTF-16 to be a fixed-width double-byte wide character set we say that each UTF-16 codepoint is represented by 2 bytes. Notice the careful use of the term 'codepoint' here. Unicode/ISO10646 characters can have character codes as large as 0x10FFFF which requires 3 bytes to store (21 bits actually); however in UTF-16 characters with more than 2 bytes of code are encoded as two codepoints, called a surrogate pair; hence, UTF-16 is fixed-width, 2 bytes per *codepoint*. It is not 2 bytes per Unicode character. UTF-16 is really a variable-width encoding, but the characters that require the surrogate-pair treatment are so infrequently used that UTF-16 is most often treated like a 16-bit fixed-width character set. It is the acknowledgement of the existence of surrogate pairs that leads to the "codepoint" vs. "character code" distinction.

UTF-32 is a fixed width encoding with a full 4-bytes per character code. It represents all of Unicode with the same width per character.

Hence, when we refer to lengths in character strings we will often refer to length in characters, but we qualify that it means 2-byte *codepoints* when the character set encoding is UTF-16. Hence, when the property lengthUnitKind is 'characters' and the charset is 'UTF-16', then the units are actually 16-bit codepoints, not Unicode characters.

Proposal
-UCS2 is   a fixed length encoding
-UTF-16 is a variable width encoding.
- A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding
Action raised

Meeting closed, 16:35

Next call Wednesday 15 September 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **117**

Actions raised at this meeting

No	Action
117	**3. Is UTF-16 a fixed width or variable width encoding** Appendix A: About UTF-16 and Unicode Character Codes above 0xFFFF When we define UTF-16 to be a fixed-width double-byte wide character set we say that each UTF-16 codepoint is represented by 2 bytes. Notice the careful use of the term 'codepoint' here. Unicode/ISO10646 characters can have character codes as large as 0x10FFFF which requires 3 bytes to store (21 bits actually); however in UTF-16 characters with more than 2 bytes of code are encoded as two codepoints, called a surrogate pair; hence, UTF-16 is fixed-width, 2 bytes per *codepoint* . It is not 2 bytes per Unicode character. UTF-16 is really a variable-width encoding, but the characters that require the surrogate-pair treatment are so infrequently used that UTF-16 is most often treated like a 16-bit fixed-width character set. It is the acknowledgement of the existence of surrogate pairs that leads to the "codepoint" vs. "character code" distinction. UTF-32 is a fixed width encoding with a full 4-bytes per character code. It represents all of Unicode with the same width per character. Hence, when we refer to lengths in character strings we will often refer to length in characters, but we qualify that it means 2-byte *codepoints* when the character set encoding is UTF-16. Hence, when the property lengthUnitKind is 'characters' and the charset is 'UTF-16', then the units are actually 16-bit codepoints, not Unicode characters. Proposal -UCS2 is a fixed length encoding -UTF-16 is a variable width encoding. - A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1' 01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite. 08/09: IBM still progressing
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification. 11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication. 11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'. Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process.

	01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and closure of all actions. 08/09: Target for completion September 30.
099	Splitting the specification in simpler sections. 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call. 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset. 21/07: no progress 04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function. Purpose of subsetting is: 1. Allow simpler implementations. (main purpose) 2. Simplify tooling 3. Simplify specification. Steve to contact previous members of WG to check if we have the correct subsets 11/08: Steve sent an email to previous members of the WG asking for opinions on splitting the specification. Bob McGrath from National Center For Supercomputing responded that they had implemented about 80% of the function. Alejandro will send a description of the function they have implemented. Action will be raised to track the Daffodil implementation 11/08: not discussed 01/09: NCSA implementation description received. Making the unparser optional is a good idea (NCSA do not need one). Work will progress on the subsets. 08/09: No progress
101	Semantics of 'fixed' 21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error. It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration. Suggestions - a parser option that 'converted all validation errors to processing errors' - a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element. Stephanie will produce some examples of how this could be used.. 04/08: Stephanie had produced examples but they were not discussed due to lack of time 11/08: We started to discuss Stephanie's HIPPA example but ran out of time. 25/08: Not discussed 01/09: Discuss next week 08/09: Stephanie sent an example of an X12 document showing how an element with the same name was defined in different groups with different enumerations. Proposal: - xs:fixed will not be used for parsing but only for validation and for providing a default value on unparsing. - A new dfdl function will be defined that applies only to simple element and tests whether the element exists including applying all the schema facets. (need to check with Tim why he wanted to only apply enumerations) dfdl:exists(xpath , **true ¦ false**) true means apply facets, false means don'e apply facets. <xs:element ref="REF_BillingProviderTaxIdentification_2010AA"> <xs:annotation>

		<xs:documentation>Discrimination needed to distinguish REF segments</xs:documentation> <xs:appinfo source="http://www.ogf.org/dfdl/"> <dfdl:discriminator test="{dfdl:exists(./REF01__ReferenceIdentificationQualifier, **true**)}"/> </xs:appinfo> </xs:annotation>
107		**teston/testoff dfdl expression functions** . Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required. 04/08: Not discussed 11/08: Not discussed 25/08: Not discussed 01/09: Steve to progress by Sept 30th 08/09: Steve to progress by Sept 30th
108		**dfdl:hidden** There has been some discussion on whether the 'hidden' global group should be indicated in some way. 04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued. 25/08: there has been some offline discussions about simplifying how hidden elements are implemented. The proposal is • dfdl:hidden property on xs:element only • xs:minOccurs and xs:maxOccurs MUST be 0 when hidden • dfdl:minOccurs and dfdl:maxOccurs for hidden elements only. • An element is 'required' when dfdl:minOccurs >0 and normal default processing occurs. The schema, without dfdl annotations, must match the infoset so assumption is that non-DFDL tools, such as mappers, will ignore/not show elements with xs:minOccurs and xs:maxOccurs = '0' 01/09: The above proposal is flawed due to use of maxOccurs = 0 (this was identified back in 2008 hence current spec). Bob confirmed that NCSA models use hidden in a big way, so punting hidden beyond 1.0 is not an option. Two candidates: - As per spec but with syntactic improvements to make it clear that the two xs:sequences do not take any dfdl:sequence properties - Place a flag directly on a local element and force minOccurs to be 0. Simpler syntax but the semantic changes, as the element *could* be legally in the infoset, although a DFDL parser would never put it there. Steve will circulate the two proposals for next week. Bob to talk to Alejandro as the NCSA implementation is currently more flexible than the spec, allowing the groupref to point to a choice, and an elementref. Are these really needed? 08/09: Discussed the Global Group and Hidden Flag approaches. Decided to stay with Global Group with dfdl:sequence properties rather than the dfdl:hidden annotation. It was agreed that there would be no extra properties on the 'hidden' global group as the syntax was messy as it should really be on the sequence and there are currently no dfdl properties on global groups. **Global group approach** Summary: • Particle to hide can be a local element, element ref, local sequence, local choice or group ref • Particle is removed from its parent into a dedicated global group of composition sequence

	and replaced in the parent by a new empty local sequence
	● The new empty local sequence carries a dfdl:hiddenGroupRef property, other DFDL properties are not allowed
	Pros:
	● Removal of all DFDL annotations and use of the resultant pure XSD results in same infoset
	● Global group can be reused
	Cons:
	● Making something hidden is a refactor operation
	● Global group sequence needs DFDL properties setting correctly
	The Daffodil parser allows the hidden annotation to reference global elements in addition to global groups. It was noted that this lost the particle properties but we need to discuss with Alejandro.
111	**Daffodil DFDL parser**
	11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented.
	We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats.
	Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil.
	Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed.
	It would be helpful if IBM test cases were available to Daffodil. IBM will investigate
	25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential.
	Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list.
	Bob is planning to make the parser available on the internet to allow testing.
	His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way.
	01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase.
	Daffodil being placed under formal source control in anticipation of external release.
	Bob has a start October deadline to create a report on what has been done for his sponsors.
	It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October.
	08/09: Alejandro is marking up Spec draft 42 to indicate which features Daffodil implement. Bob expects Daffodil to be available on the web soon.
112	**DFDL certification process**
	25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process.
	01/09: In progress, spec needs to state what conformance means, as part of this work
	08/09: Discussed what needs to be said in the spec and agreed that details of a conformance test suite should be in another document.
	Alan to draft conformance section.
113	**Regular Expressions .**
	25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?
	a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema.
	01/09: There are many variations of regexp language, it seems wise to specify one that we know contains functions like lookaround, which makes it easy to say things like 'give me everything up

	to but not including x'. This rules out XML Schema and POSIX, it needs Perl 5 or Java. 08/09: Agreed that specification should define the regular expression language (if only by referring to other specifications) . Should allow a common subset of PERL and Java expressions languages. Alan to update regular expression section.
113b	**Regular Expressions for Assert /Discriminator .** 25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient? b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both. 01/09: Tim to convince Steve (via example) that use of regexp in asserts is needed in 1.0. 08/09: Agreed that this is a useful function Allowed as alternative to expression on dfdl:assert and dfdl:discriminator - Pattern may be specified as attribute or element value - Attribute: new testPattern attribute - Element value: braces ( ) indicate pattern instead of expression
114	**OGF 30** 25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session? 09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30. 08/09: Steve to request permission to go
115	**Clarify allowed lengths for signed integer types when rep is binary integer (ie, two's complement)** 01/09: No technical reason to restrict lengths to 2^x bytes, could be odd, could be bits. But rare in practise so if we do relax, limit any core subset to 2^x bytes. 08/09: not discussed
117	**3. Is UTF-16 a fixed width or variable width encoding** Appendix A: About UTF-16 and Unicode Character Codes above 0xFFFF When we define UTF-16 to be a fixed-width double-byte wide character set we say that each UTF-16 codepoint is represented by 2 bytes. Notice the careful use of the term 'codepoint' here. Unicode/ISO10646 characters can have character codes as large as 0x10FFFF which requires 3 bytes to store (21 bits actually); however in UTF-16 characters with more than 2 bytes of code are encoded as two codepoints, called a surrogate pair; hence, UTF-16 is fixed-width, 2 bytes per *codepoint* . It is not 2 bytes per Unicode character. UTF-16 is really a variable-width encoding, but the characters that require the surrogate-pair treatment are so infrequently used that UTF-16 is most often treated like a 16-bit fixed-width character set. It is the acknowledgement of the existence of surrogate pairs that leads to the "codepoint" vs. "character code" distinction. UTF-32 is a fixed width encoding with a full 4-bytes per character code. It represents all of Unicode with the same width per character. Hence, when we refer to lengths in character strings we will often refer to length in characters, but we qualify that it means 2-byte *codepoints* when the character set encoding is UTF-16. Hence, when the property lengthUnitKind is 'characters' and the charset is 'UTF-16', then the units are actually 16-bit codepoints, not Unicode characters. Proposal -UCS2 is a fixed length encoding -UTF-16 is a variable width encoding. - A new property dfdl:UTF16Fixed 'yes ¦ no' treat UTF-16 as a fixed width encoding

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
109	Add 'message' attribute to dfdl:discriminator 01/09: **Closed**: Conclusion was that this is genuinely useful, and has low implementation cost. Will add a 'message' attribute to dfdl:discriminator.	43	not
110	Clarify expression limitations for defineVariable, newVariableInstance and setVariable 01/09: **Closed**: Spec should distinguish newVariableInstance defaultValue from setVariable value. For newVariableInstance defaultValue, disallow downward references and references to self (must be usable from the point of declaration) For setVariable allow downward references and references to self, and always evaluate at end of component. (defineVariable defaultValue should be same as newVariableInstance)	43	not
113	Be specific about regular expression syntax	43	not
108	Updates to hidden mechanism	43	not
99	Updates to reflect subsetting and unparser optionality	43	not
112	Define what conformance to spec means	43	not
115	Clarify allowed lengths for signed binary integers	43	not
116	**2. xs:minLength** The spec currently states • When an element declaration specifies a default value, and has type xs:string, then		

xs:minLength must be specified and must be 1 or greater. It is a schema definition error otherwise.

The process for defaults and nils means this restriction is no longer needed.

 Agreed

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF	
Project	**DFDL 1.0**
Meeting Date	**25-Aug-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 26-Aug-10
Last Modified by Alan Powell on 26-Aug-10

OGF DFDL Working Group Call , August 25-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , August 25-2010

Attendees
Alan Powell (IBM)
Suman Kalia (IBM)
Tim Kimber(IBM)
Bob McGrath (National Center for Supercomputing Applications)
Alejandro Rodriguez (National Center for Supercomputing Applications)

Apologies
Mike Beckerle (Oco)
Stephanie Fetzer (IBM)
Steve Hanson (IBM)

1. Current Actions
Updated Below

2. Regular Expressions .
 The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?
Discussed two aspects:
a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema.
b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both.

3. OGF 30
OGF30 takes place on October 25-29 in Brussels
Should we have a WG session?

Meeting closed, 16:30

Next call Wednesday 1 September 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **115**

Actions raised at this meeting

No	Action
112	DFDL certification process
113	**2. Regular Expressions .** The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient? Discussed two aspects: a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema. b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both.
114	**3. OGF 30** OGF30 takes place on October 25-29 in Brussels Should we have a WG session?

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1'
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up

	14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification. 11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication. 11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'. Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process.
099	Splitting the specification in simpler sections . 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call. 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset. 21/07: no progress 04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function . Purpose of subsetting is: 1. Allow simpler implementations . (main purpose) 2. Simplify tooling 3. Simplify specification . Steve to contact previous members of WG to check if we have the correct subsets 11/08: Steve sent an email to previous members of the WG asking for opinions on splitting the specification. Bob McGrath from National Center For Supercomputing responded that they had implemented about 80% of the function. Alejandro will send a description of the function they have implemented. Action will be raised to track the Daffodil implementation 11/08: not discussed
101	**Semantics of 'fixed'** 21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error . It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration. Suggestions - a parser option that 'converted all validation errors to processing errors' - a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element. Stephanie will produce some examples of how this could be used.. 04/08: Stephanie had produced examples but they were not discussed due to lack of time 11/08: We started to discuss Stephanie's HIPPA example but ran out of time. 25/08: Not discussed
107	**teston/testoff dfdl expression functions .** Are these functions still needed. They were introduced to allow individual bits to be set in a byte . Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required. 04/08: Not discussed 11/08: Not discussed

108	**dfdl:hidden**
	There has been some discussion on whether the 'hidden' global group should be indicated in some way.
	04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.
	25/08: there has been some offline discussions about simplifying how hidden elements are implemented. The proposal is
	• dfdl:hidden property on xs:element only
	• xs:minOccurs and xs:maxOccurs MUST be 0 when hidden
	• dfdl:minOccurs and dfdl:maxOccurs for hidden elements only.
	• An element is 'required' when dfdl:minOccurs >0 and normal default processing occurs.
	The schema, without dfdl annotations, must match the infoset so assumption is that non-DFDL tools, such as mappers, will ignore/not show elements with xs:minOccurs and xs:maxOccurs = '0'
109	**dfdl:discriminator : the 'message' attribute**
	From Tim:
	I remembered the reason why I thought this was a good idea.
	Consider the situation where someone is generating their DFDL schema from meta-data. The model is large, and consists of many references to global structures. Each global structure (e.g. an HL7 segment) is identified in a particular way. Sometimes the segment is required, sometimes it is not. Sometimes it occurs as a child of a choice group, and sometimes not. Regardless, it is highly likely that the segment will be identified in the same way wherever it occurs. A natural decision for the modeler would be to create a dfdl:discriminator on all references to the segement, even if the ref is not under a point of uncertainty. It's harmless, and it carries no performance penalty. If we disallow the "message" attribute, it will force the modeler to put in extra logic to work out whether the ref is under a POI, and generate an assert/discriminator as appropriate.
	I'd be interested to know what Steph thinks about this - I think I've heard her say that she sometimes uses discriminators where an assert would have done the job, just to maintain consistency throughout the model.
	04/08: not discussed.
	11/08: Not discussed
	25/08: Not discussed
110	**Semantics of newVariableInstance and setVariable**
	what should a DFDL processor (parser or serializer) do when it cannot evaluate the expression in a newVariableInstance or setVariable annotation?
	Moving the setting of variable values after the element has been parsed just creates other problems A new instance must be available to other expressions on the same component, and to the children of a group/element. So it cannot be left until the end of the element.
	On the other hand, there are clearly some types of setVariable / newVariableInstance annotations which *cannot* be evaluated until the element has been parsed.
	For the parser, it might be OK to
	- evaluate the expression when the component (element or group) is started
	- if it cannot be evaluated, add it to a list of annotations that must be processed at the end of the component
	- if in the mean time any other expressions attempt to access the variable that was being set/created then throw a processing error (because the result will be undefined). This will probably require the variable/instance to be placed into a 'not available' state until its expression

	is resolvable 25/08: There was a brief discussion as IBM needs a resolution soon. Is it possible to restrict newVariableInstance to backward references only so remove the problem? setVariable must obviously be able to access the current value.
111	**Daffodil DFDL parser** Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented. We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats. Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil. Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed. It would be helpful if IBM test cases were available to Daffodil. IBM will investigate 25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential. Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list. Bob is planning to make the parser available on the internet to allow testing. His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way. Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process.
112	DFDL certification process
113	**2. Regular Expressions .** The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient? Discussed two aspects: a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema. b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both.
114	**3. OGF 30** OGF30 takes place on October 25-29 in Brussels Should we have a WG session?

Closed actions

No	Action
104	Expressions Discuss error behaviour when evaluating an expression in various contexts - All properties: wrong type returned : schema definition error exception when evaluating expression : schema definition error referenced variables/paths not available : schema definition error - Properties which allow a forward reference referenced variables/paths not available : no error. DFDL processor continues processing until the

	expression result is available, then acts on the result.
	21/07: Steve stated the current definition that returning the incorrect type was a schema definition error and everything else was a processing error.
	04/08: Not discussed
	25/08: Closed

Work items:

N o	Item	target version	stat
0 0 5	Improvements on property descriptions		not
0 1 2	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0 3 6	Update dfdl schema with change properties	ongoing	
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**01-Sep-10 (Wed)**
Meeting Time	**15:00 - 16:30**

Created by Steve Hanson on 26-Aug-10
Last Modified by Steve Hanson on 01-Sep-10

OGF DFDL Working Group Call , September 1 2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , August 25-2010

Attendees
Stephanie Fetzer (IBM)
Steve Hanson (IBM)
Tim Kimber(IBM)
Bob McGrath (National Center for Supercomputing Applications)

Apologies
Mike Beckerle (Oco)
Alan Powell (IBM)
Suman Kalia (IBM)
Alejandro Rodriguez (National Center for Supercomputing Applications)

1. Current Actions
Updated Below

2. Twos complement integer lengths .
See new action 115 below.

3. Validating a DFDL schema prior to DFDL processing
Not required by spec, so if it happens it must be careful when validating global objects.
This is not a spec issue, but is something for IBM implementation to watch out for.
NCSA implementation does not do this.
Agreed that this was out of scope for WG call.

Meeting closed, 16:05

Next call Wednesday 8 September 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **116**

Actions raised at this meeting

No	Action
115	Clarify allowed lengths for signed integer types when rep is binary integer (ie, two's complement) 01/09: No technical reason to restrict lengths to 2^x bytes, could be odd, could be bits. But rare in practise so if we do relax, limit any core subset to 2^x bytes.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 25/08: Will chase to allow Daffodil access to test cases. The WG should define how implementation confirm that they 'conform to DFDL v1' 01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite.
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification. 11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication. 11/08: Specification is now 'awaiting author changes' before being submitted to the OGF technical committee for approval as a 'proposed specification'. Alan would like to have the updated specification complete by Sept 10th. The WG needs to complete all actions by then or decide that they do not need to be included in this phase of the process. 01/09: Alan and Steve have discussed and propose Sept 30th for completion of draft 43 and

	closure of all actions.
099	Splitting the specification in simpler sections.
	07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call.
	14/07:Discussed Steve's proposal and Suman's and Alan's comments.
	Need to add choice, validation, facets.
	Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset.
	21/07: no progress
	04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function.
	Purpose of subsetting is:
	1. Allow simpler implementations. (main purpose)
	2. Simplify tooling
	3. Simplify specification.
	Steve to contact previous members of WG to check if we have the correct subsets
	11/08: Steve sent an email to previous members of the WG asking for opinions on splitting the specification. Bob McGrath from National Center For Supercomputing responded that they had implemented about 80% of the function. Alejandro will send a description of the function they have implemented.
	Action will be raised to track the Daffodil implementation
	11/08: not discussed
	01/09: NCSA implementation description received. Making the unparser optional is a good idea (NCSA do not need one). Work will progress on the subsets.
101	Semantics of 'fixed'
	21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error.
	It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration.
	Suggestions
	- a parser option that 'converted all validation errors to processing errors'
	- a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element.
	Stephanie will produce some examples of how this could be used..
	04/08: Stephanie had produced examples but they were not discussed due to lack of time
	11/08: We started to discuss Stephanie's HIPPA example but ran out of time.
	25/08: Not discussed
	01/09: Discuss next week
107	teston/testoff dfdl expression functions.
	Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.
	04/08: Not discussed
	11/08: Not discussed
	25/08: Not discussed
	01/09: Steve to progress by Sept 30th
108	dfdl:hidden
	There has been some discussion on whether the 'hidden' global group should be indicated in some way.
	04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can

see they underlying schema. To be continued.

25/08: there has been some offline discussions about simplifying how hidden elements are implemented. The proposal is
- dfdl:hidden property on xs:element only
- xs:minOccurs and xs:maxOccurs MUST be 0 when hidden
- dfdl:minOccurs and dfdl:maxOccurs for hidden elements only.
- An element is 'required' when dfdl:minOccurs >0 and normal default processing occurs.

The schema, without dfdl annotations, must match the infoset so assumption is that non-DFDL tools, such as mappers, will ignore/not show elements with xs:minOccurs and xs:maxOccurs = '0'

01/09: The above proposal is flawed due to use of maxOccurs = 0 (this was identified back in 2008 hence current spec).
Bob confirmed that NCSA models use hidden in a big way, so punting hidden beyond 1.0 is not an option.
Two candidates:
- As per spec but with syntactic improvements to make it clear that the two xs:sequences do not take any dfdl:sequence properties
- Place a flag directly on a local element and force minOccurs to be 0. Simpler syntax but the semantic changes, as the element *could* be legally in the infoset, although a DFDL parser would never put it there.
Steve will circulate the two proposals for next week.
Bob to talk to Alejandro as the NCSA implementation is currently more flexible than the spec, allowing the groupref to point to a choice, and an elementref. Are these really needed?

111	Daffodil DFDL parser
	11/08: Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements approximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented.
	We discussed the scenarios that motivated the development which was to extract data from various sources and transform into canonical formats.
	Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil.
	Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed.
	It would be helpful if IBM test cases were available to Daffodil. IBM will investigate
	25/08: Alejandro had sent a list of the functions that he has implemented and Steve ahd responding indicating the extra functions he thought were essential.
	Since then Alejandro has implemented some of the missing functions, such as escape schemes, pre-defined variables, binary decimal numbers, etc, and will update his list.
	Bob is planning to make the parser available on the internet to allow testing.
	His organisation is being reorganised and he doesn't know what the priority of Daffodill will be so it is essential that we move quickly. It would help if IBM could indicate its support for Daffodil in some semi-formal way.
	01/09: Alejandro updating Daffodil to include escape schemes, unordered sequences and ignoreCase.
	Daffodil being placed under formal source control in anticipation of external release.
	Bob has a start October deadline to create a report on what has been done for his sponsors.
	It would be great if we could get Daffodil on the web and have run some IBM tests so it could be highlighted at OGF 30 at end October.

112	DFDL certification process
	25/08: Discussed how to certify DFDL implementations. Alan to investigate if OGF have a defined process.
	01/09: In progress, spec needs to state what conformance means, as part of this work

113	2. Regular Expressions .
	25/08: The DFDL regular expressions should provide lookahead and backreferences. Is the current regular expression language sufficient?

Discussed two aspects:
a. Is the XML regular expression language the correct one to use. Tim asked if DFDL needs to specify an language at all and should leave it to implementers to pick one. That would inhibit portability of schema.
b. A regular expression property on an assert/discriminator as an alternative to the test expression. Either a DFDL expression or a regular expression could be specified but not both.
01/09: There are many variations of regexp language, it seems wise to specify one that we know contains functions like lookaround, which makes it easy to say things like 'give me everything up to but not including x'.   This rules out XML Schema and POSIX, it needs Perl 5 or Java.
Tim to convince Steve (via example) that use of regexp in asserts is needed in 1.0.

114	3. OGF 30
	25/08: OGF30 takes place on October 25-29 in Brussels. Should we have a WG session?
	09/01: Given emergence of NCSA implementation and spec completion target of 30th Sept it makes sense to host a session at OGF 30.

Closed actions

No	Action
109	**dfdl:discriminator : the 'message' attribute**
	From Tim:
	I remembered the reason why I thought this was a good idea.
	Consider the situation where someone is generating their DFDL schema from meta-data. The model is large, and consists of many references to global structures. Each global structure (e.g. an HL7 segment) is identified in a particular way. Sometimes the segment is required, sometimes it is not. Sometimes it occurs as a child of a choice group, and sometimes not. Regardless, it is highly likely that the segment will be identified in the same way wherever it occurs. A natural decision for the modeler would be to create a dfdl:discriminator on all references to the segement, even if the ref is not under a point of uncertainty. It's harmless, and it carries no performance penalty. If we disallow the "message" attribute, it will force the modeler to put in extra logic to work out whether the ref is under a POI, and generate an assert/discriminator as appropriate.
	I'd be interested to know what Steph thinks about this - I think I've heard her say that she sometimes uses discriminators where an assert would have done the job, just to maintain consistency throughout the model.
	04/08: not discussed.
	11/08: Not discussed
	25/08: Not discussed
	01/09: **Closed:** Conclusion was that this is genuinely useful, and has low implementation cost. Will add a 'message' attribute to dfdl:discriminator.
110	**Semantics of newVariableInstance and setVariable**
	11/08: what should a DFDL processor (parser or serializer) do when it cannot evaluate the expression in a newVariableInstance or setVariable annotation?
	Moving the setting of variable values after the element has been parsed just creates other problems A new instance must be available to other expressions on the same component, and to the children of a group/element. So it cannot be left until the end of the element.
	On the other hand, there are clearly some types of setVariable / newVariableInstance annotations which *cannot* be evaluated until the element has been parsed.
	For the parser, it might be OK to
	- evaluate the expression when the component (element or group) is started
	- if it cannot be evaluated, add it to a list of annotations that must be processed at the end of the component
	- if in the mean time any other expressions attempt to access the variable that was being set/created then throw a processing error (because the result will be undefined). This will probably require the variable/instance to be placed into a 'not available' state until its expression is

resolvable
25/08: There was a brief discussion as IBM needs a resolution soon. Is it possible to restrict newVariableInstance to backward references only so remove the problem? setVariable must obviously be able to access the current value.
01/09: **Closed:** Spec should distinguish newVariableInstance defaultValue from setVariable value. For newVariableInstance defaultValue, disallow downward references and references to self (must be usable from the point of declaration)
For setVariable allow downward references and references to self, and always evaluate at end of component.
(defineVariable defaultValue should be same as newVariableInstance)

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
109	Add 'message' attribute to dfdl:discriminator	43	not
110	Clarify expression limitations for defineVariable, newVariableInstance and setVariable	43	not
113	Be specific about regular expression syntax	43	not
108	Updates to hidden mechanism	43	not
99	Updates to reflect subsetting and unparser optionality	43	not
112	Define what conformance to spec means	43	not
11	Clarify allowed lengths for signed binary integers	43	not

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**11-Aug-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 12-Aug-10
Last Modified by Alan Powell on 12-Aug-10

OGF DFDL Working Group Call , August 11-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , August 11-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)
Bob McGrath (National Center for Supercomputing Applications)
Alejandro Rodriguez (National Center for Supercomputing Applications)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)

1. Current Actions
Updated Below

2 Semantics of newVariableInstance and setVariable

what should a DFDL processor (parser or serializer) do when it cannot evaluate the expression in a newVariableInstance or setVariable annotation?

Moving the setting of variable values after the element has been parsed just creates other problems A new instance must be available to other expressions on the same component, and to the children of a group/element. So it cannot be left until the end of the element.
On the other hand, there are clearly some types of setVariable / newVariableInstance annotations which *cannot* be evaluated until the element has been parsed.

For the parser, it might be OK to
- evaluate the expression when the component (element or group) is started
- if it cannot be evaluated, add it to a list of annotations that must be processed at the end of the component
- if in the mean time any other expressions attempt to access the variable that was being set/created then throw a processing error (because the result will be undefined). This will probably require the variable/instance to be placed into a 'not available' state until its expression is resolvable
11/08: Not discussed. Action raised

3 Daffodil DFDL parser implementation at National Center for Supercomputing Applications
Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements aproximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented.

We discussed the scenarios that motivated the development which was to extract data from various soucres and transform into canonical formats.

Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil.

Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed.

It would be helpful if IBM test cases were available to Daffodil. IBM will investigate

Meeting closed, 16:30

Next call Wednesday 25 August 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee 's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **112**

Actions raised at this meeting

No	Action
110	**Semantics of newVariableInstance and setVariable** what should a DFDL processor (parser or serializer) do when it cannot evaluate the expression in a newVariableInstance or setVariable annotation? Moving the setting of variable values after the element has been parsed just creates other problems A new instance must be available to other expressions on the same component, and to the children of a group/element. So it cannot be left until the end of the element. On the other hand, there are clearly some types of setVariable / newVariableInstance annotations which *cannot* be evaluated until the element has been parsed. For the parser, it might be OK to - evaluate the expression when the component (element or group) is started - if it cannot be evaluated, add it to a list of annotations that must be processed at the end of the component - if in the mean time any other expressions attempt to access the variable that was being set/created then throw a processing error (because the result will be undefined). This will probably require the variable/instance to be placed into a 'not available' state until its expression is resolvable
111	**Daffodil DFDL parser** Bob and Alejandro described the new implementation that they have developed. It is a new code

base and is not based on the Deffudle prototype. It is written in scala and implements aproximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented.

We discussed the scenarios that motivated the development which was to extract data from various soucres and transform into canonical formats.

Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil.

Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed.

It would be helpful if IBM test cases were available to Daffodil. IBM will investigate

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 11/08: work continues
085	ALL: publicise Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over and can we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response.

	21/07: Still no response.
	04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification.
	11/08: Received a response from Joel that the WG can decide if a re- public review is necessary before becoming a 'proposed recommendation'. Alan responded that the WG agreed that a re-review was not necessary. The next stage is for OGF review committee to approve publication.
099	Splitting the specification in simpler sections.
	07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call.
	14/07:Discussed Steve's proposal and Suman's and Alan's comments.
	Need to add choice, validation, facets.
	Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset.
	21/07: no progress
	04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function.
	Purpose of subsetting is:
	1. Allow simpler implementations. (main purpose)
	2. Simplify tooling
	3. Simplify specification.
	Steve to contact previous members of WG to check if we have the correct subsets
	11/08: Steve sent an email to previous members of the WG asking for opinions on splitting the specification. Bob McGrath from National Center For Supercomputing responded that they had implemented about 80% of the function. Alejandro will send a description of the function they have implemented.
	Action will be raised to track the Daffodil implementation
101	**Semantics of 'fixed'**
	21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error.
	It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration.
	Suggestions
	- a parser option that 'converted all validation errors to processing errors'
	- a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element.
	Stephanie will produce some examples of how this could be used..
	04/08: Stephanie had produced examples but they were not discussed due to lack of time
	11/08: We started to discuss Stephanie's HIPPA example but ran out of time.
104	**Expressions**
	Discuss error behaviour when evaluating an expression in various contexts
	- All properties:
	wrong type returned : schema definition error
	exception when evaluating expression : schema definition error
	referenced variables/paths not available : schema definition error
	- Properties which allow a forward reference
	referenced variables/paths not available : no error. DFDL processor continues processing until the expression result is available, then acts on the result.
	21/07: Steve stated the current definition that returning the incorrect type was a schema definition error and everything else was a processing error.
	04/08: Not discussed

107	**teston/testoff dfdl expression functions .**
	Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.
	04/08: Not discussed
108	**dfdl:hidden**
	There has been some discussion on whether the 'hidden' global group should be indicated in some way.
	04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.
109	**dfdl:discriminator : the 'message' attribute**
	From Tim:
	I remembered the reason why I thought this was a good idea.
	Consider the situation where someone is generating their DFDL schema from meta-data. The model is large, and consists of many references to global structures. Each global structure (e.g. an HL7 segment) is identified in a particular way. Sometimes the segment is required, sometimes it is not. Sometimes it occurs as a child of a choice group, and sometimes not. Regardless, it is highly likely that the segment will be identified in the same way wherever it occurs. A natural decision for the modeler would be to create a dfdl:discriminator on all references to the segment, even if the ref is not under a point of uncertainty. It's harmless, and it carries no performance penalty. If we disallow the "message" attribute, it will force the modeler to put in extra logic to work out whether the ref is under a POI, and generate an assert/discriminator as appropriate.
	I'd be interested to know what Steph thinks about this - I think I've heard her say that she sometimes uses discriminators where an assert would have done the job, just to maintain consistency throughout the model.
	04/08: not discussed.
110	**Semantics of newVariableInstance and setVariable**
	what should a DFDL processor (parser or serializer) do when it cannot evaluate the expression in a newVariableInstance or setVariable annotation?
	Moving the setting of variable values after the element has been parsed just creates other problems A new instance must be available to other expressions on the same component, and to the children of a group/element. So it cannot be left until the end of the element.
	On the other hand, there are clearly some types of setVariable / newVariableInstance annotations which *cannot* be evaluated until the element has been parsed.
	For the parser, it might be OK to
	- evaluate the expression when the component (element or group) is started
	- if it cannot be evaluated, add it to a list of annotations that must be processed at the end of the component
	- if in the mean time any other expressions attempt to access the variable that was being

	set/created then throw a processing error (because the result will be undefined). This will probably require the variable/instance to be placed into a 'not available' state until its expression is resolvable
111	**Daffodil DFDL parser** Bob and Alejandro described the new implementation that they have developed. It is a new code base and is not based on the Deffudle prototype. It is written in scala and implements aproximately 80% of the features in the public comments draft of DFDL V1. Alejandro will send a list of the features not implemented. We discussed the scenarios that motivated the development which was to extract data from various soucres and transform into canonical formats. Bob offered to make Daffodil available for the WG to assess the functionality. IBM WG members will get approval the company to allow them to receive Daffodil. Bob raised the question that if Daffodil becomes the public implementation of DFDL then we will need to work out how that would be funded and managed. It would be helpful if IBM test cases were available to Daffodil. IBM will investigate

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		

TeamRoom Plus

IBM TeamRoo

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**04-Aug-10 (Wed)**
Meeting Time	**15:00 - 16:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 05-Aug-10
Last Modified by Alan Powell on 05-Aug-10

OGF DFDL Working Group Call , August 04-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , August 04-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Ella Belisario (IBM)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)
Tim Kimber(IBM)

1. Current Actions
Updated Below

2. dfdl:hidden
There has been some discussion on whether the 'hidden' global group should be indicated in some way.

Option A:
Leave the DFDL specification unchanged, in which case we must expose the internal DFDL constructs to the user of the DFDL editor. That means the following:
- to make a group member hidden, a local group and a global group must be created to represent the hidden item. This is a refactoring operation.
- to make the UI consistent, the global group needs to be exposed as a first-class modelling construct. That means that it appears along with local element/element ref/local group/group ref in any context where the user is dealing with group members.
There are more details of what the editor needs to do in Steve's comments on the first note in this chain.

Option B:
Make one change to the DFDL specification, to make it possible for a DFDL editor to present dfdl:hidden as a simple flag on any group member
The change is to put a DFDL annotation on the global sequence group that contains the hidden items. This would allow the DFDL editor to distinguish these global groups from the real ones that belong to the user's schema. This allows a hidden item to be represented by a pair of *internal* groups, both of which have a dfdl:annotation. One is local, and appears where the hidden item used to appear, The other is global and is referenced by the local one. They are created and deleted as pairs whenever an item is hidden or un-hidden.

A lively discussion. The specification works as currently defined so discussed whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.

3. dfdl:discriminator : the 'message' attribute
From Tim:

I remembered the reason why I thought this was a good idea.

Consider the situation where someone is generating their DFDL schema from meta-data. The model is large, and consists of many references to global structures. Each global structure (e.g. an HL7 segment) is identified in a particular way. Sometimes the segment is required, sometimes it is not. Sometimes it occurs as a child of a choice group, and sometimes not. Regardless, it is highly likely that the segment will be identified in the same way wherever it occurs. A natural decision for the modeler would be to create a dfdl:discriminator on all references to the segement, even if the ref is not under a point of uncertainty. It's harmless, and it carries no performance penalty. If we disallow the "message" attribute, it will force the modeler to put in extra logic to work out whether the ref is under a POI, and generate an assert/discriminator as appropriate.

I'd be interested to know what Steph thinks about this - I think I've heard her say that she sometimes uses discriminators where an assert would have done the job, just to maintain consistency throughout the model.
04/08: not discussed.

Meeting closed, 16:00

Next call Wednesday 11 August 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **110**

Actions raised at this meeting

No	Action
108	**dfdl:hidden** There has been some discussion on whether the 'hidden' global group should be indicated in some way. 04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.
109	**dfdl:discriminator : the 'message' attribute** From Tim: I remembered the reason why I thought this was a good idea. Consider the situation where someone is generating their DFDL schema from meta-data. The model is large, and consists of many references to global structures. Each global structure (e.g. an HL7 segment) is identified in a particular way. Sometimes the segment is required, sometimes it is not. Sometimes it occurs as a child of a choice group, and sometimes not. Regardless, it is highly likely that the segment will be identified in the same way wherever it occurs . A natural decision for the modeler would be to create a dfdl:discriminator on all references to the segement, even if the ref is not under a point of uncertainty. It's harmless, and it carries no performance penalty. If we disallow the "message" attribute, it will force the modeler to put in extra logic to work out whether the ref is under a POI, and generate an assert/discriminator as appropriate. I'd be interested to know what Steph thinks about this - I think I've heard her say that she sometimes uses discriminators where an assert would have done the job, just to maintain consistency throughout the model.

	04/08: not discussed.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 21/07: work continues 04/08: work continues
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over andcan we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response. 04/08: Joel has responded that it is up to the WG to decide if the changes are significant enough to need additional review. Alan to contact David Martin and Erwin Laure for guidance if we split the specification.
099	Splitting the specification in simpler sections. 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call. 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or

profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset.
21/07: no progress
04/08: Steve had updated proposed groups of function. (Subset_proposal_v2.ppt). We discussed whether its is better to have discrete sets of functions or expanding levels of function.
Purpose of subsetting is:
1. Allow simpler implementations. (main purpose)
2. Simplify tooling
3. Simplify specification.
Steve to contact previous members of WG to check if we have the correct subsets.

101	**Semantics of 'fixed'**
	21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error.
	It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration.
	Suggestions
	- a parser option that 'converted all validation errors to porcessing errors'
	- a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element.
	Stephanie will produce some examples of how this could be used..
	04/08: Stephanie had produced examples but they were not discussed due to lack of time
104	**Expressions**
	Discuss error behaviour when evaluating an expression in various contexts
	- All properties:
	wrong type returned : schema definition error
	exception when evaluating expression : schema definition error
	referenced variables/paths not available : schema definition error
	- Properties which allow a forward reference
	referenced variables/paths not available : no error. DFDL processor continues processing until the expression result is available, then acts on the result.
	21/07: Steve stated the current definition that returning the incorrect type was a schema definition error and everything else was a processing error.
	04/08: Not discussed
107	**teston/testoff dfdl expression functions .**
	Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.
	04/08: Not discussed
108	**dfdl:hidden**
	There has been some discussion on whether the 'hidden' global group should be indicated in some way.
	04/08: A lively discussion. The specification is works as currently defined so whether changes need to be made to make tooling easier. There shouldn't be 'conventions' in particular tooling as they must be able to properly deal with schema from other tools that would not obey those conventions. Steve stated that it is often dangerous to hide too much from users when they can see they underlying schema. To be continued.
109	**dfdl:discriminator : the 'message' attribute**
	From Tim:
	I remembered the reason why I thought this was a good idea.
	Consider the situation where someone is generating their DFDL schema from meta-data. The model is large, and consists of many references to global structures. Each global structure (e.g.

an HL7 segment) is identified in a particular way. Sometimes the segment is required, sometimes it is not. Sometimes it occurs as a child of a choice group, and sometimes not. Regardless, it is highly likely that the segment will be identified in the same way wherever it occurs. A natural decision for the modeler would be to create a dfdl:discriminator on all references to the segment, even if the ref is not under a point of uncertainty. It's harmless, and it carries no performance penalty. If we disallow the "message" attribute, it will force the modeler to put in extra logic to work out whether the ref is under a POI, and generate an assert/discriminator as appropriate.

I'd be interested to know what Steph thinks about this - I think I've heard her say that she sometimes uses discriminators where an assert would have done the job, just to maintain consistency throughout the model.
04/08: not discussed.

Closed actions

No	Action
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed. 16/05: went through Steves comments. Steve to update draft. 23/06: Steve's updates to the rules discussed. See minutes. Rest of document needs updating. 30/06: Discussed Alans updates. Some corrections. Alan will send out updated copy for review before next call. 07/07: Discussed Alan updates and Tim and Steve's comments. Still some corrections and updates. 14/07: Discussed Alan updates (v9) Still some corrections and updates. 21/07: Discussed Draft 10. Shouldn't mention input/outputValueCalc in this section. Mention defaulting in calculated values section. Move details of nilValue from nilKind property. 04/08: Draft v11.1 was included in Spec Draft v42. Minor corrections. Closed.
102	Clarify the specification of error reporting from a DFDL processor - section 2.3 needs to be updated 21/07: Section is not clear. Alan will update. 04/08: Sections updated in draft 42. Closed
103	Asserts and discriminators - specify the scope of forward references. Must be downward-only. The expression must be resolvable by the time the component on which it is positioned goes out of scope - otherwise it is a processing error. 21/07: Agreed 04/08: Updated in draft 42. Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		

TeamRoom Plus

IBM TeamRoo

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**21-Jul-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 22-Jul-10
Last Modified by Alan Powell on 22-Jul-10

OGF DFDL Working Group Call , July 21-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , July 21-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)

1. Current Actions
Updated below

2. Alignment of children when forward references .
Expressions with forward references require caching of serialized output which is difficult when children have

alignments that are greater that their parent.
Agreed that 'the alignment of a child must be less than or equal to the alignment of its parent'. This allows serializing parts of the tree without having to redo alignments.

3. teston/testoff dfdl expression functions .
Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.

Meeting closed, 16:30

Next call Wednesday 28 July 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **108**

Actions raised at this meeting

No	Action
106	**2. Alignment of children when forward references .** Expressions with forward references require caching of serialized output which is difficult when

	children have alignments that are greater that their parent. Agreed that 'the alignment of a child must be less than or equal to the alignment of its parent'. This allows serializing parts of the tree without having to redo alignments. Closed
107	3. teston/testoff dfdl expression functions . Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 21/07: work continues
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over andcan we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response. 21/07: Still no response.
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing

	14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed. 16/05: went through Steves comments. Steve to update draft. 23/06: Steve's updates to the rules discussed. See minutes. Rest of document needs updating. 30/06: Discussed Alans updates. Some corrections. Alan will send out updated copy for review before next call. 07/07: Discussed Alan updates and Tim and Steve's comments. Still some corrections and updates. 14/07: Discussed Alan updates (v9) Still some corrections and updates. 21/07: Discussed Draft 10. Shouldn't mention input/outputValueCalc in this section. Mention defaulting in calculated values section. Move details of nilValue from nilKind property.
099	Splitting the specification in simpler sections. 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call. 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset. 21/07: no progress
101	**Semantics of 'fixed'** 21/07: Discussed whether not matching the 'fixed' value should be a validation error or processing error. Decided that for consistency it should be a validation error. It would be useful however to avoid having to duplication of facet information in an assert which could become unwieldy for, say, a large enumeration. Suggestions - a parser option that 'converted all validation errors to porcessing errors' - a dfdl expression function that 'applied all facets' or 'applied specific facet' to a particular element. Stephanie will proce some examples of how this could be used.
102	**Clarify the specification of error reporting from a DFDL processor** - section 2.3 needs to be updated 21/07: Section is not clear. Alan will update.
103	**Asserts and discriminators** - specify the scope of forward references. Must be downward-only. The expression must be resolvable by the time the component on which it is positioned goes out of scope - otherwise it is a processing error. 21/07: Agreed
104	**Expressions** Discuss error behaviour when evaluating an expression in various contexts - All properties:

	wrong type returned : schema definition error
	exception when evaluating expression : schema definition error
	referenced variables/paths not available : schema definition error
	- Properties which allow a forward reference
	referenced variables/paths not available : no error. DFDL processor continues processing until the expression result is available, then acts on the result.
	21/07: Steve stated the current definition that returning the incorrect type was a schema definition error and everything else was a processing error.
107	**teston/testoff dfdl expression functions .** Are these functions still needed. They were introduced to allow individual bits to be set in a byte. Steve to look at TLog and ISO 8583 formats that use existence flags to see if they are still required.

Closed actions

No	Action
105	**Grammar for DFDL expression**
	- currently allows multiple predicates on a single expression. Sounds as if this was unintended. 21/07: Current definition is incorrect: Will delete predicateList from the grammar and replace with predicate. Closed
106	**2. Alignment of children when forward references .** Expressions with forward references require caching of serialized output which is difficult when children have alignments that are greater that their parent. Agreed that 'the alignment of a child must be less than or equal to the alignment of its parent'. This allows serializing parts of the tree without having to redo alignments. Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
08	Implement RFC2116		

3			
1 0 5	AP: Describe trailingSkipBytes for delimited formats. Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.'		
1 0 6	AP: Skip Bytes should allow bits Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.		
1 0 7	Remove timing from dfdl:assert		
1 0 8	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies. 30/06: Decided that defaulting for variable occurrence arrays should always be to minOccurs. If separatorPolicy is 'required' then just the separators will be output up to maxOccurs and unbounded is an error.		
1 0 9	define semantics of choiceKind 'fixedLength' dfdl:choiceKind becomes dfdl:choiceLengthKind with emuns 'implicit' (length of selected branch) and 'explicit' (length specified by dfd;choiceLength) New property dfdl:choiceLength		
1 1 0	nilIndicatorPath and nilIndicatorIndex properties 14/07: As part of nils default discussion the WG decided to drop dfdl:nilIndicator and dfdl:nilIndicatorIndex		
1 1 1	AP: using textStringPadCharacter with charRef '%#r' on multi-byte encoding 14/07: Although there was some merit in having separate textPadKinds it was felt it was too late to make the change. The property descriptions (for numbers, calendars and booleans as well) will be improved. Closed		
1 1 2	Alignment enumerations when dfdl :lengthUnits = 'bits' When alignmentUnits='bits', the rule that alignment must be 1,2,4,8 etc or multiple of 2 looks rather hard to justify. It was agreed to remove the restriction on aligmentUnits for both 'bytes' and 'bits'. Closed		
1 1 3	Grammar for DFDL expression - currently allows multiple predicates on a single expression. Sounds as if this was unintended. 21/07: Current definition is incorrect: Will delete predicateList from the grammar and replace with predicate. Closed		
1 1 4	2. Alignment of children when forward references . Expressions with forward references require caching of serialized output which is difficult when children have alignments that are greater that their parent. Agreed that 'the alignment of a child must be less than or equal to the alignment of its parent'. This allows serializing parts of the tree without having to redo alignments. Closed		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**14-Jul-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 15-Jul-10
Last Modified by Alan Powell on 15-Jul-10

OGF DFDL Working Group Call , July 14-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , July 14-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)
Suman Kalia (IBM)

Apologies
Mike Beckerle (Oco)

1. Current Actions
Updated below

2 Nils and Defaults .
Discussed Alan updates (v9). Updates still needed.

3. using textStringPadCharacter with charRef '%#r' on multi-byte encoding

textStringPadCharacter	DFDL String literal
	The pad~~ding~~ character or byte ~~value~~ that is used when justifying or trimming text elements.
	A pad character can be specified using a literal or DFDL entities.
	A pad byte ~~value~~ must be specified using the %#r entity.
	~~DFDL validation rules~~
	- if a pad character is specified when lengthUnits='bytes' then the pad character must be a single-byte character.
	- if a pad byte ~~value~~ is specified when lengthUnits='characters' then the encoding must be a fixed-width encoding. Padding and trimming is be applied using an array of N pad bytes ~~values~~, where N is the width of a character in the fixed-width encoding.
	Annotation: dfdl:element, dfdl:simpleType

Would adding fillByte to dfdl:textPadKind enumeration make it clearer.
padChar is a pad character (%#r not allowed)
fillByte is a fill byte

Although there was some merit in having separate textPadKinds it was felt it was too late to make the change. The property descriptions (for numbers, calendars and booleans as well) will be improved as above.

4. nilIndicatorPath and nilIndicatorIndex properties

These properties seem a bit of an anomaly. Tim has suggested they can be simplified.

Tim had emaled Mike with a proposal and Steve had responded with 3 alternatives
1) Leave things as they are, perhaps renaming 'nilIndicator' to 'indicator'

2) Collapse nilIndicatorPath and nilIndicatorIndex to a single property. It means the path is repeated, but it makes nils consistent with length and occurs. For all, you can always use a string variable to hold the constant part of the path and concatenate with the index. And if we improve usability in a future DFDL release, we would improve it for nils, lengths and occurs at the same time.

3) Drop the properties for 1.0 altogether on the grounds that it is a rare case. I've never seen such a format, but we know Mike has.
As part of nils default discussion the WG decided to drop dfdl:nilIndicator and dfdl:nilIndicatorIndex as it was felt the complications it introduced on unparsing were not worth the additional function. This may be revisited for DFDL v2.

5. Alignment enumerations when dfdl:lengthUnits = 'bits'
When alignmentUnits='bits', the rule that alignment must be 1,2,4,8 etc or multiple of 2 looks rather hard to justify.
We should either
a) remove the rule entirely and allow any postiive integer > 1 or
b) make the rule apply only when alignmentUnits='bytes'
It was agreed to remove the restriction on aligmentUnits for both 'bytes' and 'bits'.
Also suggested that a note should be added to the default alignment and implicit length tables that 'although the values are specified in bits that does not imply that the dfdl:lengthUnits can be specified as 'bits' for all the simple types

6. Semantics of 'fixed'
Not discussed. Action Raised.

7. Clarify the specification of error reporting from a DFDL processor
- section 2.3 needs to be updated
Not discussed. Action Raised

8 Asserts and discriminators
- specify the scope of forward references. Must be downward-only. The expression must be resolvable by the time the component on which it is positioned goes out of scope - otherwise it is a processing error.
Not discussed. Action Raised

9 Expressions
Discuss error behaviour when evaluating an expression in various contexts
- All properties:
wrong type returned : schema definition error
exception when evaluating expression : schema definition error
referenced variables/paths not available : schema definition error
- Properties which allow a forward reference
referenced variables/paths not available : no error. DFDL processor continues processing until the expression result is available, then acts on the result.
Not discussed. Action Raised

10. Grammar for DFDL expression
- currently allows multiple predicates on a single expression . Sounds as if this was unintended.
Not discussed. Action Raised

Meeting closed, 17:00

Next call Wednesday 21 July 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee 's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **106**

Actions raised at this meeting

No	Action
101	**Semantics of 'fixed'**
102	**Clarify the specification of error reporting from a DFDL processor** - section 2.3 needs to be updated
103	**Asserts and discriminators** - specify the scope of forward references. Must be downward-only. The expression must be resolvable by the time the component on which it is positioned goes out of scope - otherwise it is a processing error.
104	**Expressions** Discuss error behaviour when evaluating an expression in various contexts - All properties: wrong type returned : schema definition error exception when evaluating expression : schema definition error referenced variables/paths not available : schema definition error - Properties which allow a forward reference referenced variables/paths not available : no error. DFDL processor continues processing until the expression result is available, then acts on the result.
105	**Grammar for DFDL expression** - currently allows multiple predicates on a single expression. Sounds as if this was unintended.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations ... 14/07: work continues
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over andcan we update the published version with WG updates. No response yet. 07/07: No response. Alan will chase up 14/07: No response from Joel. Sent email to Greg Newby by no response.
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed. 16/05: went through Steves comments. Steve to update draft. 23/06: Steve's updates to the rules discussed. See minutes. Rest of document needs updating. 30/06: Discussed Alans updates. Some corrections. Alan will send out updated copy for review

	before next call. 07/07: Discussed Alan updates and Tim and Steve's comments. Still some corrections and updates. 14/07: Discussed Alan updates (v9) Still some corrections and updates.
099	Splitting the specification in simpler sections. 07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call. 14/07:Discussed Steve's proposal and Suman's and Alan's comments. Need to add choice, validation, facets. Also how does an implementation declare which subsets it supports. Suggested levels and/or profiles. Steve highlighted a problem when a DFDL schema from an implementation of just the core functions was moved to a full DFDL implementation what should happen about the missing properties. Does the full implementation need to be aware of subsets of functions? Should it raise a schema definition error for use of a function not in the subset.
101	Semantics of 'fixed'
102	Clarify the specification of error reporting from a DFDL processor - section 2.3 needs to be updated
103	Asserts and discriminators - specify the scope of forward references. Must be downward-only. The expression must be resolvable by the time the component on which it is positioned goes out of scope - otherwise it is a processing error.
104	Expressions Discuss error behaviour when evaluating an expression in various contexts - All properties: wrong type returned : schema definition error exception when evaluating expression : schema definition error referenced variables/paths not available : schema definition error - Properties which allow a forward reference referenced variables/paths not available : no error. DFDL processor continues processing until the expression result is available, then acts on the result.
105	Grammar for DFDL expression - currently allows multiple predicates on a single expression. Sounds as if this was unintended.

Closed actions

No	Action
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO 15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example. 02/06: no progress 16/06: no progress 23/06: no progress

	30/06: Alan looked at Tim's description of calculable length and suggested that that real use case may be much simpler. If real use case is COBOL and C importers then it would be cleaner to require the 'fixed length' to be specified on the enclosing complex element and remove choiceKind 'fixed length'. Ask Suman is COBOL and C importers can be enhance to provide length on cpmplex element. 07/07: Steve had proposed 5 possible approaches to supporting main use case. i) Change the importer to add a parent element for each group. Con: This changes the logical model by inserting an extra level. ii) Allow dfdl:length to be carried on embedded xs:choice (along with dfdl:lengthUnits & dfdl:lengthKind). iii) New dfdl:choiceLength property to be carried on embedded xs:choice (always bytes). iv) dfdl:choiceKind as today. Con: length from TD model is lost and must be recalculated. v) dfdl:choiceKind but with limitations to make the calculation easier. Con: length from TD model is lost and must be recalculated. Agreed a modified version of iii) dfdl:choiceKind becomes dfdl:choiceLengthKind with emuns 'implicit' (length of selected branch) and 'explicit' (length specified by dfd;choiceLength) New property dfdl:choiceLength 14/07: Closed
096	AP: using textStringPadCharacter with charRef '%#r' on multi-byte encoding 14/07: Although there was some merit in having separate textPadKinds it was felt it was too late to make the change. The property descriptions (for numbers, calendars and booleans as well) will be improved. Closed
097	nilIndicatorPath and nilIndicatorIndex properties 07/07: Tim had emaled Mike with a proposal and Steve had responded with 3 alternatives 1) Leave things as they are, perhaps renaming 'nilIndicator' to 'indicator' 2) Collapse nilIndicatorPath and nilIndicatorIndex to a single property. It means the path is repeated, but it makes nils consistent with length and occurs. For all, you can always use a string variable to hold the constant part of the path and concatenate with the index. And if we improve usability in a future DFDL release, we would improve it for nils, lengths and occurs at the same time. 3) Drop the properties for 1.0 altogether on the grounds that it is a rare case. I've never seen such a format, but we know Mike has. Not discussed. 14/07: As part of nils default discussion the WG decided to drop dfdl:nilIndicator and dfdl:nilIndicatorIndex as it was felt the complications it introduced on unparsing were not worth the additional function. This may be revisited for DFDL v2.. Closed
100	**Alignment enumerations when dfdl :lengthUnits = 'bits'** When alignmentUnits='bits', the rule that alignment must be 1,2,4,8 etc or multiple of 2 looks rather hard to justify. It was agreed to remove the restriction on aligmentUnits for both 'bytes' and 'bits'. Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0	Update dfdl schema with change properties	ongoing	

3 6			
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		
1 0 5	AP: Describe trailingSkipBytes for delimited formats. Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.'		
1 0 6	AP: Skip Bytes should allow bits Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.		
1 0 7	Remove timing from dfdl:assert		
1 0 8	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies. 30/06: Decided that defaulting for variable occurrence arrays should always be to minOccurs. If separatorPolicy is 'required' then just the separators will be output up to maxOccurs and unbounded is an error.		
1 0 9	define semantics of choiceKind 'fixedLength' dfdl:choiceKind becomes dfdl:choiceLengthKind with emuns 'implicit' (length of selected branch) and 'explicit' (length specified by dfd;choiceLength) New property dfdl:choiceLength		
1 1 0	nilIndicatorPath and nilIndicatorIndex properties 14/07: As part of nils default discussion the WG decided to drop dfdl:nilIndicator and dfdl:nilIndicatorIndex		
1 1 1	AP: using textStringPadCharacter with charRef '%#r' on multi-byte encoding 14/07: Although there was some merit in having separate textPadKinds it was felt it was too late to make the change. The property descriptions (for numbers, calendars and booleans as well) will be improved. Closed		
1 1 2	**Alignment enumerations when dfdl :lengthUnits = 'bits'** When alignmentUnits='bits', the rule that alignment must be 1,2,4,8 etc or multiple of 2 looks rather hard to justify. It was agreed to remove the restriction on aligmentUnits for both 'bytes' and 'bits'. Closed		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**07-Jul-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 08-Jul-10
Last Modified by Alan Powell on 08-Jul-10

OGF DFDL Working Group Call , July 07-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key pla
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , July 07-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)
Suman Kalia (IBM)

Apologies
Mike Beckerle (Oco)

1. Current Actions
Updated below

2 Nils and Defaults .
Discussed Alan updates and Tim and Steve's comments. Still some corrections and updates. Needs to have a introductory
paragraph.

3 DFDL property types and other issues .

Tim has proposed more specific types for some properties. In particular separating the different kinds of entities.

- Modify the meaning and usage of the type 'DFDL String Literal' and modify the remainder of the specification accordingly

- Improve the description of DFDL entities to avoid confusion over the intended usage of raw byte values.

- Clarify the standard sentence about forward references in DFDL expressions - the current text implied that the restrictions o applied to the unparser.

On parsing: No forward references allow for expressions that provide property values. Asserts and discriminators may have downward references (need to be defined)
and expression must be evaluated by the time the element has been processed.
On unparsing: outputValueCalc may have a forward reference.

1.1   DFDL Properties

Properties on DFDL annotations may be one or more of the following types

•   DFDL string literal
The property value is a string that represents a sequence of literal bytes or characters which appear in the data stream.
•   ~~List of DFDL string literals~~
~~The property value is a space-separated list of DFDL string literals. When parsing, if more than one string literal in the list~~
~~matches the portion of the~~
~~data stream being evaluated then the longest matching string literal in the list must be used.~~
~~When unparsing, the first string literal in the list must be used.~~
•   DFDL expression
The property value is an DFDL subset XPath 2.0 expression that returns a value derived from other property values and/or fr the DFDL infoset.
•   DFDL regular expression
The property value is a regular expression that can be used as a pattern to calculate the length of an element by applying tha pattern to the sequence
of literal bytes or characters which appear in the data stream.
•   Enumeration
The property value is one of the allowed values listed in the property description.
•   Logical Value ~~Simple Type~~
The property value is a string that describes a logical value. The type of the logical value is one of the XML Schema simple ty
•   QName
The property value is an XML Qualified Name as specified in "Namespaces in XML ".

Identify which types can be a list or Union
The property value is a space-separated list. When parsing, if more than one string literal in the list matches the portion of the data stream
being evaluated then the longest matching string literal in the list must be used.
When unparsing, the first value in the list must be used.

1.1.1.1   DFDL Entities in String Literals
XML entities should not be listed as they are valid in any property value.

1.1.1.1   **Character classes in DFDL String literals**

1.1.1.1 Raw byte values in DFDL String Literals

4. using textStringPadCharacter with charRef '%#r' on multi-byte encoding

textStringPadCharacter	DFDL String literal
	The padding character or byte value that is used when justifying or trimming text elements.
	A pad character can be specified using DFDL entities.
	A pad byte value must be specified using the %#r entity.
	DFDL validation rules
	- if a pad byte value is specified when lengthUnits='characters' then the encoding must be a fixed-width encoding.
	- if a pad character is specified when lengthUnits='bytes' then the pad character must be a single-byte character.
	If a pad byte value is specified when lengthUnits='characters' then padding and trimming must be applied
	using an array of N pad byte values, where N is the width of a character in the fixed-width encoding.
	Annotation: dfdl:element, dfdl:simpleType

Would adding fillByte to dfdl:padChar enumeration make it clearer.
padChar is a pad character (%#r not allowed)
fillByte is a fill byte

5. nilIndicatorPath and nilIndicatorIndex properties
These properties seem a bit of an anomaly. Tim has suggested they can be simplified.
Tim had emaled Mike with a proposal and Steve had responded with 3 alternatives
1) Leave things as they are, perhaps renaming 'nilIndicator' to 'indicator'

2) Collapse nilIndicatorPath and nilIndicatorIndex to a single property. It means the path is repeated, but it makes nils consist
with length and occurs. For all, you can always use a string variable to hold the constant part of the path and concatenate wit
index. And if we improve usability in a future DFDL release, we would improve it for nils, lengths and occurs at the same time

3) Drop the properties for 1.0 altogether on the grounds that it is a rare case. I've never seen such a format, but we know Mike
has.
Not discussed.

Meeting closed, 17:00

Next call Wednesday 14 July 2010 15:00 UK (10:00 ET)

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **100**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing

	17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations 16/06: work continues 23/06: work continues 30/06: work continues 07/07: work continues
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over andcan we update the published version with WG updates. No response yet. 07/07: No response. Alan will chance up
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed. 16/05: went through Steves comments. Steve to update draft. 23/06: Steve's updates to the rules discussed. See minutes. Rest of document needs updating. 30/06: Discussed Alans updates. Some corrections. Alan will send out updated copy for review before next call. 07/07: Discussed Alan updates and Tim and Steve's comments. Still some corrections and updates.
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL

	importer deals with ODO.
	21/04: Need to define 'calculable length' and WMB importer ODO behaviour.
	05/05: TK: Still need definition of calculable length.
	SKK: WMB COBOL imported behaviour with ODO
	15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example.
	02/06: no progress
	16/06: no progress
	23/06: no progress
	30/06: Alan looked at Tim's description of calculable length and suggested that that real use case may be much simpler. If real use case is COBOL and C importers then it would be cleaner to require the 'fixed length' to be specified on the enclosing complex element and remove choiceKind 'fixed length'. Ask Suman is COBOL and C importers can be enhance to provide length on cpmplex element.
	07/07: Steve had proposed 5 possible approaches to supporting main use case.
	i) Change the importer to add a parent element for each group. Con: This changes the logical model by inserting an extra level.
	ii) Allow dfdl:length to be carried on embedded xs:choice (along with dfdl:lengthUnits & dfdl:lengthKind).
	iii) New dfdl:choiceLength property to be carried on embedded xs:choice (always bytes).
	iv) dfdl:choiceKind as today. Con: length from TD model is lost and must be recalculated.
	v) dfdl:choiceKind but with limitations to make the calculation easier. Con: length from TD model is lost and must be recalculated.
	Agreed a modified version of iii)
	dfdl:choiceKind becomes dfdl:choiceLengthKind with emuns 'implicit' (length of selected branch) and 'explicit' (length specified by dfd;choiceLength)
	New property dfdl:choiceLength
096	AP: using textStringPadCharacter with charRef '%#r' on multi-byte encoding
	07/07: not discussed
097	nilIndicatorPath and nilIndicatorIndex properties
	07/07: Tim had emaled Mike with a proposal and Steve had responded with 3 alternatives
	1) Leave things as they are, perhaps renaming 'nilIndicator' to 'indicator'
	2) Collapse nilIndicatorPath and nilIndicatorIndex to a single property. It means the path is repeated, but it makes nils consistent with length and occurs. For all, you can always use a string variable to hold the constant part of the path and concatenate with the index. And if we improve usability in a future DFDL release, we would improve it for nils, lengths and occurs at the same time.
	3) Drop the properties for 1.0 altogether on the grounds that it is a rare case. I've never seen such a format, but we know Mike has.
	Not discussed.
099	Splitting the specification in simpler sections.
	07/07: Steve sent a proposal but not discussed. Alan will arrange a separate call.

Closed actions

No	Action
092	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies.
	16/06: no progress
	23/06: discussed
	- whether when number of instances doesn't match specified number of occurrences is it an error or should missing instances be defaulted? Decided it is an error.
	- defaulting occurs up to minoccurs unless separator policy is required when default up to maxOccurs and unbounded is an error.
	30/06: Decided that defaulting for variable occurrence arrays should always be to minOccurs. If

	separatorPolicy is 'required' then just the separators will be output up to maxOccurs and unbounded is an error. 07/07: Agreed last call's decision. Closed

Work items:

N o	Item	target version	stat
0 0 5	Improvements on property descriptions		not
0 1 2	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0 3 6	Update dfdl schema with change properties	ongoing	
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		
1 0 5	AP: Describe trailingSkipBytes for delimited formats. Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.'		
1 0 6	AP: Skip Bytes should allow bits Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.		
1 0 7	Remove timing from dfdl:assert		
1 0 8	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies. 30/06: Decided that defaulting for variable occurrence arrays should always be to minOccurs. If separatorPolicy is 'required' then just the separators will be output up to maxOccurs and unbounded is an error.		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**30-Jun-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 01-Jul-10
Last Modified by Alan Powell on 01-Jul-10

OGF DFDL Working Group Call , June 30-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
of your learning.

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , June 30-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Suman Kalia (IBM)
Mike Beckerle (Oco)

1. Current Actions
Updated below

2 Nils and Defaults .
Reviewed Alan's latest draft.
Agreed with definition of 'required parent context'
Preferred the combining of initiated and not initiated rules.
Would still like a separate definition of nillable

3 DFDL property types and other issues .

Tim has proposed more specific types for some properties. In particular separating the different kinds of entities.

- Modify the meaning and usage of the type 'DFDL String Literal' and modify the remainder of the specification accordingly

- Improve the description of DFDL entities to avoid confusion over the intended usage of raw byte values.

- Clarify the standard sentence about forward references in DFDL expressions - the current text implied that the restrictions o

1.1 **DFDL Properties**

Properties on DFDL annotations may be one or more of the following types

- DFDL string literal

The property value is a string that represents ~~describes~~ a sequence of literal bytes ~~and~~ or characters which appear in the data

- List of DFDL string literals

The property value is a space-separated list of DFDL string literals. When parsing, if more than one string literal in the list ma longest matching string literal in the list must be used.

When unparsing, the first string literal in the list must be used.

- DFDL expression

The property value is an XPath 2.0 expression that ~~calculates~~ ~~evaluates to~~ returns a value derived from other property values ~~calculate property values, and to calculate logical values for simple elements.~~

- DFDL regular expression

The property value is a regular expression that can be used as a pattern to calculate the length of an element by ~~comparing a~~ which appear in the data stream.

- Enum

The property value ~~is a string literal that~~ must be one of the allowed values listed in the property description.

- QName

The property value is ~~a string literal that also conforms to~~ an XML Qualified Name as specified in "Namespaces in XML "

- Simple Type

The property value is a string that describes a logical value. The type of the logical value is one of the XML Schema simple ty

1.1.1.1 DFDL Entities in String Literals

1.1.1.1 **Character classes in DFDL String literals**

1.1.1.1 **Raw byte values in DFDL String Literals**

4. using textStringPadCharacter with charRef '%#r' on multi-byte encoding

textStringPadCharacter	DFDL String literal
	The padding character or byte value that is used when justifying or trimming text elements.

	A pad character can be specified using DFDL entities.
	A pad byte value must be specified using the %#r entity.
	DFDL validation rules
	- if a pad byte value is specified when lengthUnits='characters' then the encoding must be a fixed-width encoding.
	- if a pad character is specified when lengthUnits='bytes' then the pad character must be a single-byte character.
	If a pad byte value is specified when lengthUnits='characters' then padding and trimming must be applied
	using an array of N pad byte values, where N is the width of a character in the fixed-width encoding.
	Annotation: dfdl:element, dfdl:simpleType

Would adding fillByte to dfdl:padChar enumeration make it clearer.
padChar is a pad character (%#r not allowed)
fillByte is a fill byte

5. nilIndicatorPath and nilIndicatorIndex properties
These properties seem a bit of an anomaly. Tim has suggested they can be simplified.
Not discussed

6 Remove timing from dfdl:assert.
As timing property has been removed from discriminators it is inconsistent to still have it on assert. Agreed
There was also a discussion about adding a message property to discriminators which was inconclusive so decided not to ch

7. Splitting the specification in simpler sections.
Concern has been raised that a full DFDL implementation is complicated and expensive.
A number of reviewer have said that they might implement only part of the sepcification.
It has been suggested that the WG should look to see if it possible to split the specification into more
manageable optional sections.

Meeting closed, 17:00

Next call Wednesday 07 July 2010 15:00 UK (10:00 ET)

Create Action Items
Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create
Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items
will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **100**

Actions raised at this meeting

No	Action
096	AP: **using textStringPadCharacter with charRef '%#r' on multi-byte encoding**
097	**nilIndicatorPath and nilIndicatorIndex properties**
099	Splitting the specification in simpler sections.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing

	02/06: Work still progressing on technical and legal considerations 16/06: work continues 23/06: work continues 30/06: work continues
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment. 23/06: Still no comments. Alan will contact OGF to understand the rest of the process. 30/06: Alan has emailed Joel asking what the process is now public comment period is over andcan we update the published version with WG updates. No response yet.
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed. 16/05: went through Steves comments. Steve to update draft. 23/06: Steve's updates to the rules discussed. See minutes. Rest of document needs updating. 30/06: Discussed Alans updates. Some corrections. Alan will send out updated copy for review before next call.
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO 15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example. 02/06: no progress 16/06: no progress 23/06: no progress 30/06: Alan looked at Tim's description of calculable length and suggested that that real use

092	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies.
	case may be much simpler. If real use case is COBOL and C importers then it would be cleaner to require the 'fixed length' to be specified on the enclosing complex element and remove choiceKind 'fixed length'. Ask Suman is COBOL and C importers can be enhance to provide length on cpmplex element.
092	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies. 16/06: no progress 23/06: discussed - whether when number of instances doesn't match specified number of occurrences is it an error or should missing instances be defaulted? Decided it is an error. - defaulting occurs up to minoccurs unless separator policy is required when default up to maxOccurs and unbounded is an error. 30/06: Decided that defaulting for variable occurrence arrays should always be to minOccurs. If separatorPolicy is 'required' then just the separators will be output up to maxOccurs and unbounded is an error.
096	AP: using textStringPadCharacter with charRef '%#r' on multi-byte encoding
097	nilIndicatorPath and nilIndicatorIndex properties
099	Splitting the specification in simpler sections.

Closed actions

No	Action
098	Remove timing from dfdl:assert

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
105	AP: Describe trailingSkipBytes for delimited formats. Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.'		
1	AP: Skip Bytes should allow bits		

06	Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.		
107	Remove timing from dfdl:assert		
108			

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**23-Jun-10 (Wed)**
Meeting Time	**13:00 - 14:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 24-Jun-10
Last Modified by Alan Powell on 25-Jun-10

OGF DFDL Working Group Call , June 23-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
of your learning.

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , June 23-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Suman Kalia (IBM)
Mike Beckerle (Oco)

1. Current Actions
Updated below

2 Nils and Defaults .
Reviewed Steve's updates to the parsing and unparsing rules.
- it might be possible to simplify the nil processing in the rules by having a defintion up front .
- 'nil indicator set' need a definition
- Discussion about whether 'usNilForDefaults' could be replaced by some way of defining a literal, out of band value for the d

3 DFDL property types and other issues .

Tim has proposed more specific types for some properties. In particular separating the different kinds of entities.

- Modify the meaning and usage of the type 'DFDL String Literal' and modify the remainder of the specification accordingly

- Improve the description of DFDL entities to avoid confusion over the intended usage of raw byte values.

- Clarify the standard sentence about forward references in DFDL expressions - the current text implied that the restrictions o

1.1   DFDL Properties
Not discussed

Properties on DFDL annotations may be one or more of the following types

- •   DFDL string literal

The property value is a string that represents ~~describes~~ a sequence of literal bytes ~~and~~ or characters which appear in the dat

- •   List of DFDL string literals

The property value is a space-separated list of DFDL string literals. When parsing, if more than string literal in the list matche matching string literal in the list must be used.

When unparsing, the first string literal in the list must be used.

- •   DFDL expression

The property value is an XPath 2.0 expression that ~~calculates~~ evaluates to a value derived from other property values and/or ~~property values, and to calculate logical values for simple elements.~~

- •   DFDL regular expression

The property value is a regular expression that can be used as a pattern to calculate the length of an element by comparing t appear in the data stream.

- •   Enum

The property value is a string literal that must be one of the allowed values listed in the property description.

- •   QName

The property value is a string literal that also conforms to an XML Qualified Name as specified in "Namespaces in XML "

- •   DFDL Simple Type

The property value is a string that describes a logical value. The type of the logical value is one of the XML Schema simple ty

- •   Non-negative integer

The property is a non-negative integer value What does this add?

- ● Integer

     The property is an integer value. What does this add?

1.1.1.1   DFDL Entities in String Literals

1.1.1.1   **Character classes in DFDL String literals**

1.1.1.1   **Raw byte values in DFDL String Literals**

4. using textStringPadCharacter with charRef   '%#r' on multi-byte encoding

textStringPadCharacter	DFDL String literal
	The padding character or byte value that is used when justifying or trimming text elements.
	A pad character can be specified using DFDL entities. A pad byte value must be specified using the %#r entity.
	DFDL validation rules - if a pad byte value is specified when lengthUnits='characters' then the encoding must be a fixed-width encoding. - if a pad character is specified when lengthUnits='bytes' then the pad character must be a single-byte character.
	If a pad byte value is specified when lengthUnits='characters' then padding and trimming must be applied using an array of N pa
	Annotation: dfdl:element, dfdl:simpleType

Would adding fillByte to dfdl:padChar enumeration make it clearer.
padChar is a pad character (%#r not allowed)
fillByte is a fill byte

5. nilIndicatorPath and nilIndicatorIndex properties
These properties seem a bit of an anomaly. Tim has suggested they can be simplified.

Meeting closed, 14:30

Next call Wednesday 30 June 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

Next action: **096**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations 16/06: work continues

	23/06: work continues
085	ALL: publicize Public comments phase to ensure a good review..
14/04: see minutes
21/04: Press release, OMG and other standards bodies.
05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec
15/05: still no public comments
02/06: No public comments
16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment.
23/06: Still no comments. Alan will contact OGF to understand the rest of the process.
086	AP: Nils and Defaults during unparsing - update table
31/03: TK to documetn use cases for parsing
14/04: Investigate new property to control empty string behaviour.
21/04: After investigation a new property is not required. New rules developed and tables updated.
Need examples of complexTypes to confirm tables apply.
Review Nils, defaulting spec section.
05/05: Discussed defaulting complex elements. Tables updated but need to add terminator.
SH; to confirm WMD behaviour when infoset item has no value on unparsing
Need to describe defaulting choices.
15/05: More discussion. Alan updating sections
26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft.
02/06: Alan updated description. Please review.
Discussed Stephanie's example using discriminators. Decided no changes needed.
16/05: went through Steves comments. Steve to update draft.
23/06: Steve's updates to the rules discussed. See minutes. Rest of document needs updating.
088	define semantics of choiceKind 'fixedLength'
31/03: TK to provide definition of calculable length.
Investigate PL/I varchars and Cobol occurs dependingon.
14/04Tim had distributed a document starting the definition of calculable length for the longest choice member.
Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO.
21/04: Need to define 'calculable length' and WMB importer ODO behaviour.
05/05: TK: Still need definition of calculable length.
SKK: WMB COBOL imported behaviour with ODO
15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example.
02/06: no progress
16/06: no porgress
23/06: no progress
092	AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies.
16/06: no progress
23/06: discussed
- whether when number of instances doesn't match specified number of occurrences is it an error or should missing instances be defaulted? Decided it is an error.
- defaulting occurs up to minoccurs unless separator policy is required when default up to maxOccurs and unbounded is an error.

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration	041	
098	occursCountKind="expression" is parsing only	041	
099	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.	041	
100	Need to state in 4.1.2 Infoset that value is optional.	041	
101	When dfdl:textNumberRep is 'zoned' only the pattern for positive numbers is used. Only the following pattern characters may be used: '+' MUST BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is signed '+' MAY BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is unsigned 'V' MAY BE used to indicate the location of an implied decimal point	041	

	'0' indicates the number of required digits (including overpunched). '#' indicates the number optional digits.		
1 0 2	Also textNumberPolicy implies it applies to zoned, but doesn't state what zoned behaviour it covers. I think it should be consistent with binaryNumberCheckPolicy for packed, and control whether positive punched data is accepted or rejected when parsing an unsigned type, and whether unpunched data is accepted or rejected when parsing a signed type.	041	
1 0 3	- textStandardBase - textStandardGroupSeparator - textStandardExponentCharacter - textStandardInfinityRep - textStandardNanRep - textStandardZeroRep - textZonedSignStyle	041	
1 0 4	4 Unsigned decimal Proposed solution: - Allow xs:nonNegativeInteger which enables unsigned unbounded integers to be modelled. The problem to solve is then just for xs:decimal. - Call the new property dfdl:decimalSigned. It only applies to xs:decimal or user defined restrictions thereof. It applies to all physical decimals, as its name implies (not just zoned or packed).	041	
1 0 5	AP: Describe trailingSkipBytes for delimited formats. Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.'		
1 0 6	AP: Skip Bytes should allow bits Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**16-Jun-10 (Wed)**
Meeting Time	**15:00 - 16:30**
Meeting Editors	Alan Powell

Created by Alan Powell on 17-Jun-10
Last Modified by Alan Powell on 18-Jun-10

OGF DFDL Working Group Call , June 16-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key pla
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , June 16-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Suman Kalia (IBM)
Mike Beckerle (Oco)

1 Nils, defaults and unparsing
Steve had done thorough review which we went through.
Main points
- The 4th case of 'has a default' describes nils processing and shouldn't be included in defaulting as it don't
always apply.
- required context not sufficiently described .

Steve will update draft.

2 OGF 29
Chicago, June 20-22
Suman may be able to attend subject to finding reasonable travel cost.

3 Current Actions
Updated below

4 Trailing alignment with lengthKind delimited or endOfParent
Is it sensible to allow Trailing skip bytes with LK delimited

The grammar says that trailing alignment is applied before the separator

So the parser must search for the separator then remove the trailing bytes from the value and length.

@Data1 *Data2*#   - Test1 value=Data1 length=5

@Data1*Data2*#   - Test1 value=Da length=2   (do we require that the trailing data must be the fillByte?)

But bear in mind that the grammar effectively says:

   RightFraming = Terminator TrailingSkipBytes

So the combination of dfdl:lengthKind="delimited" & dfdl:trailingSkipBytes would work when the delimiter is a terminator

Steve was keen that there was consistent behaviour over different lengthKinds so that we shouldn't just disallow trailingSkipBytes for delimited . We discussed allowing TSB when terminator was specified . Alan to propose wording.
Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.' Agreed
It was noted that this emphasised that terminators and postFix separators do not behave the same.

5 More on bit oriented formats
Steve has been looking at a bit oriented format from the space industry

```
Compound Field C1 { //length 10 bits
        Field S1 = 1 byte
        Field S2 = 2 bits
}

Compound Field C2 { //length 2 bytes
        Field S3 = 3 bits
        Field S4 = 10 bits
}
```


The 'unused' fields could be defined as hidden or as leading/trailing skip bytes of a nearby field. But LSB and TSB are specified in bytes only.

Steve proposed that LSB and TSB should be renamed leadingSkip and trailingSkip and can be specified in bits or bytes as specified by alignmentUnits or a new property skipUnits

Ageed that it should be possible to specify bits.
- LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip
- units are specified by dfdl:alignmentUnits.

6 separatorPolicy with optional fields and variable arrays .
In particular:
- separatorPolicy 'required' with variable arrays
- supressAtEnd with required fields that have defaults.
Tim will dig out his original writeup

7 DFDL property types .
Tim has proposed more specific types for some properties. In particular separating the different kinds of entities.
Tim and Alan discussed after the call
Confirmed that only inputValueCalc can have forward reference during parsing.

Meeting closed, 16:30

Next call Wednesday 23 June 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **096**

Actions raised at this meeting

No	Action
094	AP: Improve DFDL property types
095	AP: Skip Bytes should allow bits
	Ageed that it should be possible to specify bits.
	- LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip
	- units are specified by dfdl:alignmentUnits.

Current Actions :

No	Action
066	Investigate format for defining test cases
	25/11:IBM to see if it is possible to publish its test case format.
	04/12: no update
	...
	17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite'
	24/03: No progress
	03/03: Discussions have been taking place on the subset of tests that will be provided.
	10/03: work is progressing
	17/03: work is progressing
	31/03: work is progressing
	14/04: And XML test case format has been defined and is being tested.
	21/04. Schema for TDML defined. Need to define how this and the test cases will be made public

		05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations 16/06: work continues
085		ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments 16/06: Public comments period has ended with no external comments. Alan had posted changes made in draft 041. Steve suggested send a note to the WG highlighting these changes. Steve also suggested requesting an extension as other IBM groups may review. We discussed whether this was necessary as changes will need to be made during the implementation phase anyway. Alan to ask OGF what the process is for changes post public comment.
086		AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed. 16/05: went through Steves comments. Steve to update draft.
088		define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO 15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example. 02/06: no progress 16/06: no porgress
092		AP: Confirm behaviour of defaulting with various occursCountKinds and separator policies. 16/06: no progress
093		AP: Describe trailingSkipBytes for delimited formats. 16/06: wording above agreed. Closed

Closed actions

No	Action
093	AP: Describe trailingSkipBytes for delimited formats. 16/06: wording above agreed. Closed
095	AP: Skip Bytes should allow bits Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration	041	
098	occursCountKind="expression" is parsing only	041	
099	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.	041	
100	Need to state in 4.1.2 Infoset that value is optional.	041	
101	When dfdl:textNumberRep is 'zoned' only the pattern for positive numbers is used. Only the following pattern characters may be used: '+' MUST BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is signed '+' MAY BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is unsigned	041	

			041	
	'V' MAY BE used to indicate the location of an implied decimal point '0' indicates the number of required digits (including overpunched). '#' indicates the number optional digits.			
1 0 2	Also textNumberPolicy implies it applies to zoned, but doesn't state what zoned behaviour it covers. I think it should be consistent with binaryNumberCheckPolicy for packed, and control whether positive punched data is accepted or rejected when parsing an unsigned type, and whether unpunched data is accepted or rejected when parsing a signed type.		041	
1 0 3	- textStandardBase - textStandardGroupSeparator - textStandardExponentCharacter - textStandardInfinityRep - textStandardNanRep - textStandardZeroRep - textZonedSignStyle		041	
1 0 4	4 Unsigned decimal Proposed solution: - Allow xs:nonNegativeInteger which enables unsigned unbounded integers to be modelled. The problem to solve is then just for xs:decimal. - Call the new property dfdl:decimalSigned. It only applies to xs:decimal or user defined restrictions thereof. It applies to all physical decimals, as its name implies (not just zoned or packed).		041	
1 0 5	AP: Describe trailingSkipBytes for delimited formats. Alan suggested 'dfdl:terminator must be specified and not empty if dfdl:lengthKind is delimited or endOfParent.'			
1 0 6	AP: Skip Bytes should allow bits Ageed that it should be possible to specify bits. - LSB and TSB renames to dfdl:leadingSkip, dfdl:trailingSkip - units are specified by dfdl:alignmentUnits.			

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**02-Jun-10 (Wed)**
Meeting Time	**15:00 - 16:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 03-Jun-10
Last Modified by Alan Powell on 03-Jun-10

OGF DFDL Working Group Call , June 02-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plai details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , June 02-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Stephanie Fetzer (IBM)

Apologies
Tim Kimber(IBM)
Suman Kalia (IBM)
Mike Beckerle (Oco)

1 Nils, defaults and unparsing
Continue review proposal emailed by Alan
The proposal was not discussed on the call. Please send comments to Alan by next Monday.

Stephanie sent an email discussing whether using discriminators had any effect on defaulting using a COBOL occursDependingOn example. It seems as though there is not a problem. However the discussion prompted questions about how occursCountKind effected defaulting. In particular if the occureCountKind is expression referring to another field in the data which doesn't match the number of occurences is it an error should defaulting occur?. Action raised to follow up.

2 OGF 29
Chicago, June 20-22
Suman was unable to get authorization to attend.

3 Current Actions
Updated below

4 Trailing alignment with lengthKind delimited or endOfParent
Is it sensible to allow Trailing skip bytes with LK delimited

The grammar says that trailing alignment is applied before the separator

So the parser must search for the separator then remove the trailing bytes from the value and length.

@Data1 *Data2*#   - Test1 value=Data1 length=5

@Data1*Data2*#   - Test1 value=Da length=2   (do we require that the trailing data must be the fillByte?)

But bear in mind that the grammar effectively says:

    RightFraming = Terminator TrailingSkipBytes

So the combination of dfdl:lengthKind="delimited" & dfdl:trailingSkipBytes would work when the delimiter is a terminator

Steve was keen that there was consistent behaviour over different lengthKinds so that we shouldn't just disallow trailingSkipBytes for delimited. We discussed allowing TSB when terminator was specified. Alan to propose wording.

5 More on bit oriented formats
Steve has been looking at a bit oriented format from the space industry

```
Compound Field C1 { //length 10 bits
        Field S1 = 1 byte
        Field S2 = 2 bits
}

Compound Field C2 { //length 2 bytes
        Field S3 = 3 bits
        Field S4 = 10 bits
}
```


The 'unused' fields could be defined as hidden or as leading/trailing skip bytes of a nearby field. But LSB and TSB are specified in bytes only.

Steve proposed that LSB and TSB should be renamed leadingSkip and trailingSkip and can be specified in bits or bytes as specified by alignmentUnits or a new property skipUnits

6 separatorPolicy with optional fields and variable arrays .
In particular:
- separatorPolicy 'required' with variable arrays

- supressAtEnd with required fields that have defaults.

Meeting closed, 16:00

Next call Wednesday 09 June 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee 's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **094**

Actions raised at this meeting

No	Action
092	AP: Confirm behaviour of defaulting with various occursCountKinds
093	AP: Describe trailingSkipBytes for delimited formats.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing 02/06: Work still progressing on technical and legal considerations
085	ALL: publicize Public comments phase to ensure a good review..

	14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments 02/06: No public comments
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft. 02/06: Alan updated description. Please review. Discussed Stephanie's example using discriminators. Decided no changes needed.
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO 15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example. 02/06: no progress
092	AP: Confirm behaviour of defaulting with various occursCountKinds
093	AP: Describe trailingSkipBytes for delimited formats.

Closed actions

No	Action
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc. 14/04: no prorgess 21/04: no progress 05/05: no progress 26/05: Alan has looked at the precedence tables and the look OK. Suggest closing. 02/06: Closed
091	The IBM team had questioned whether timezones should be just UTC+/- offset or should GMT and other short time zones be allowed. Alan to investigate 15/05: Alan to talk to Ruth. 26/05: Ruth confirmed that it was only the offset from UTC or GMT. For DFDL purposes thates are the same. Suggest closing. 02/06: closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration	041	
098	occursCountKind="expression" is parsing only	041	
099	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.	041	
100	Need to state in 4.1.2 Infoset that value is optional.	041	
101	When dfdl:textNumberRep is 'zoned' only the pattern for positive numbers is used. Only the following pattern characters may be used: '+' MUST BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is signed '+' MAY BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is unsigned 'V' MAY BE used to indicate the location of an implied decimal point '0' indicates the number of required digits (including overpunched). '#' indicates the number optional digits.	041	
102	Also textNumberPolicy implies it applies to zoned, but doesn't state what zoned behaviour it covers. I think it should be consistent with binaryNumberCheckPolicy for packed, and control whether positive punched data is accepted or rejected when parsing an unsigned type, and whether unpunched data is accepted or rejected when parsing a signed type.	041	
10	- textStandardBase	041	

3	- textStandardGroupSeparator - textStandardExponentCharacter - textStandardInfinityRep - textStandardNanRep - textStandardZeroRep - textZonedSignStyle		
1 0 4	4 Unsigned decimal Proposed solution: - Allow xs:nonNegativeInteger which enables unsigned unbounded integers to be modelled. The problem to solve is then just for xs:decimal. - Call the new property dfdl:decimalSigned. It only applies to xs:decimal or user defined restrictions thereof. It applies to all physical decimals, as its name implies (not just zoned or packed).	041	

TeamRoom Plus

Clarified default for choices.

Definition:'required'
Clarified that required is in terms of the infoset.

Didn't like the used of 'index' in the definition of variable occurrence elements.

Agreed that ASSERTs do not make an element required. Stephanie to consider examples to confirm this is correct.

Suggest introducing 'required context' to explain defaulting in complex elements.

Discussed the restriction that choice branches may noy have minOccurs = 0 and whether that was sufficient. The was introduced to prevent the empty string being implicitly a valid choice but Mike pointed out that making all the children of a choice branch optional has the same effect. Not resolved.

Discussed whether anonymous sequences and choices were required. Concluded that it wasn't a valid question when required is in terms of the infoset. Thus if anonymous sequences and choices have an initiator or terminator they will always be present in the data stream. If they need to be optional then the seq/choice has to be wrapped in an element.

Suggest removing the definition of success.

Largely agreed on the definition of missing.

2 OGF 29
Chicago, June 20-22
Not discussed

3 Current Actions
Not discussed

4 Trailing alignment with lengthKind delimited or endOfParent
Is it sensible to allow Trailing skip bytes with LK delimited

The grammar says that trailing alignment is applied before the separator

So the parser must search for the separator then remove the trailing bytes from the value and length.

@Data1 *Data2*#   - Test1 value=Data1 length=5

@Data1*Data2*#   - Test1 value=Da length=2   (do we require that the trailing data must be the fillByte?)

But bear in mind that the grammar effectively says:

    RightFraming = Terminator TrailingSkipBytes

So the combination of dfdl:lengthKind="delimited" & dfdl:trailingSkipBytes would work when the delimiter is a terminator.
Not discussed

Meeting closed, 16:30

Next call Wednesday 02 June 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create
Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items
will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **092**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases
	25/11:IBM to see if it is possible to publish its test case format.
	04/12: no update

	... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing 12/05: Work still progressing
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc. 14/04: no prorgess 21/04: no progress 05/05: no progress 26/05: Alan has looked at the precedence tables and the look OK. Suggest closing.
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections 26/05: Discussed draft updates. Stephanie to confirm asserts do not make an element required. Alan will update draft.. All: review rest of draft.
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO 15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example.
091	The IBM team had questioned whether timezones should be just UTC+/- offset or should GMT and other short time zones be allowed. Alan to investigate 15/05: Alan to talk to Ruth. 26/05: Ruth confirmed that it was only the offset from UTC or GMT. For DFDL purposes thates are the same. Suggest closing.

Closed actions

No	Action	

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration		
098	occursCountKind="expression" is parsing only		
099	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.		
100	Need to state in 4.1.2 Infoset that value is optional.		
101	When dfdl:textNumberRep is 'zoned' only the pattern for positive numbers is used. Only the following pattern characters may be used: **'+' MUST BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is signed** **'+' MAY BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is unsigned** 'V' MAY BE used to indicate the location of an implied decimal point		

	'0' indicates the number of required digits (including overpunched). '#' indicates the number optional digits.		
1 0 2	Also textNumberPolicy implies it applies to zoned, but doesn't state what zoned behaviour it covers. I think it should be consistent with binaryNumberCheckPolicy for packed, and control whether positive punched data is accepted or rejected when parsing an unsigned type, and whether unpunched data is accepted or rejected when parsing a signed type.		
1 0 3	- textStandardBase - textStandardGroupSeparator - textStandardExponentCharacter - textStandardInfinityRep - textStandardNanRep - textStandardZeroRep - textZonedSignStyle		
1 0 4	4 Unsigned decimal Proposed solution: - Allow xs:nonNegativeInteger which enables unsigned unbounded integers to be modelled. The problem to solve is then just for xs:decimal. - Call the new property dfdl:decimalSigned. It only applies to xs:decimal or user defined restrictions thereof. It applies to all physical decimals, as its name implies (not just zoned or packed).		

TeamRoom Plus

IBM TeamRoo

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**12-May-10 (Wed)**
Meeting Time	**15:00 - 16:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 13-May-10
Last Modified by Alan Powell on 13-May-10

OGF DFDL Working Group Call , May 12 -2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , May 12 -2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Suman Kalia (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Steve Marting (Progeny)
Mike Beckerle (Oco)

1 Nils, defaults and unparsing

From last calls minutes
- *It was agreed that on unparsing a simple infoset item with no value will be treated the same as missing. Steve*

H to check WMB behaviour. Need to state in 4.1.2 Infoset that value is optional. "

Something to note. We say that for a simple element infoset item **[dataValue]** has a value (either type compliant or nil) and for a complex element infoset item **[dataValue]** has 'no value'. Also we say that for a simple element infoset item **[children]** has 'no value' and for a complex element infoset item **[children]** it **may** have a value (so it could also have 'no value'). The implication therefore is that an element infoset item is deduced to be simple or complex based on the [dataValue] property. If we let a simple item take 'no value' for [dataValue] as proposed, then there is no reliable way to identify an element infoset item as simple or complex.

Is that a problem in practice? After parsing, all simple items will have **[dataValue]** set. For items created by some other means, and presumably destined for unparsing, you could have an item with neither **[dataValue]** nor **[children]** set. We could say (1) that an element infoset item is known to be simple or complex only when it is processed by the unparser, at which point the item is matched to the DFDL schema and its type resolved. This is what broker's MRM parser does. Alternatively (2) we introduce a simple/complex flag.

Note that I tried creating an MRM 'infoset' item for a string element but just gave a name and no value. The MRM serializer looked at the element, matched it to its model, found it was simple, and treated the value as NULL which invoked null processing. The XML serializer also assumed NULL.

Personally I think we should adopt (1) but interpret 'no value' for a simple item as nil and not missing.

Options for simple types:

1) Treat like missing from infoset. Means that an optional element will not be output, and a required element will go through default value processing.

2) Treat like [dataValue] is nil. Means nil processing is invoked if nillable. This is what MRM does today.

3) Give an error - [dataValue] must always be set. It is up to the hosting application to apply a semantic to a missing value from its own 'infoset' and always set something when calling DFDL, if it wants to avoid such errors.

After discussion general agreement that in the absence of any other reason compatibility with existing products meant option 2. Steve H suggested we get Mike's opinion.

Alan has been updating the Nils/Defaults section and raised the following questions

A) Does nil processing apply to optional elements? In particular should a nilValue of empty string apply to optional element.

Nil processing does apply to optional elements including a nilValue of %#ES;

B) Does the nilValueDelimiterPolicy and emptyyValueDelimiterPolicy of a child apply on parsing when the parent is missing. That is should there be different rules for when the just child is missing than when the parent is missing.

After discussion although it was felt that there could be formats that required that support we weren't aware of any so V1 will say that the child policy properties do apply.

2 OGF 29
Chicago, June 20-22

Suman volunteered to attend to present. Note the public comment period ends 15 June.

3 Current Actions
see below

4 zoned numbers clarifications
The specification of the number pattern for zoned numbers is not clear.
A.
Clarification:
When dfdl:textNumberRep is 'zoned' only the pattern for positive numbers is used. Only the following pattern characters may be used:

 '+' MUST BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is signed
 '+' MAY BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is unsigned
 'V' MAY BE used to indicate the location of an implied decimal point
 '0' indicates the number of required digits (including overpunched).
 '#' indicates the number optional digits.
Agreed

B.
Also textNumberPolicy implies it applies to zoned, but doesn't state what zoned behaviour it covers. I think it should be consistent with binaryNumberCheckPolicy for packed, and control whether positive punched data is accepted or rejected when parsing an unsigned type, and whether unpunched data is accepted or rejected when parsing a signed type.
Agreed
C.
Finally, several textNumber properties only apply to standard or zoned. We could distinguish these by name, which is what we have done elsewhere. So:
- textStandardBase
- textStandardGroupSeparator
- textStandardExponentCharacter
- textStandardInfinityRep
- textStandardNanRep
- textStandardZeroRep
- textZonedSignStyle
Stephanie noted that changes to properties caused rework of test cases so, in general, should be resisted
Agreed to change

Meeting closed, 14:30

Next call Wednesday 19 May 2010 15:00 UK (10:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items

will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **092**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public

	05/05: Work still progressing 12/05: Work still progressing
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc. 14/04: no prorgess 21/04: no progress 05/05: no progress
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec 15/05: still no public comments
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices. 15/05: More discussion. Alan updating sections
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO 15/05: Suman sent an expmle of an imported Cobol ODo which suggested that the maximum space was reserved. He will extend the example.
091	The IBM team had questioned whether timezones should be just UTC+/- offset or should GMT and other short time zones be allowed. Alan to investigate 15/05: Alan to talk to Ruth.

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
0	Reordering the properties discussion: move representation earlier, improve flow of topics		

			not
1 2			
0 3 6	Update dfdl schema with change properties	ongoing	
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		
0 9 7	Remove functions that returns duration		
0 9 8	occursCountKind is parsing only		
0 9 9	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.		
1 0 0	Need to state in 4.1.2 Infoset that value is optional.		
1 0 1	When dfdl:textNumberRep is 'zoned' only the pattern for positive numbers is used. Only the following pattern characters may be used: '+' MUST BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is signed '+' MAY BE present at the beginning or end of the pattern to indicate whether the leading or trailing digit carries the overpunched sign, if the logical type is unsigned 'V' MAY BE used to indicate the location of an implied decimal point '0' indicates the number of required digits (including overpunched). '#' indicates the number optional digits.		
1 0 2	Also textNumberPolicy implies it applies to zoned, but doesn't state what zoned behaviour it covers. I think it should be consistent with binaryNumberCheckPolicy for packed, and control whether positive punched data is accepted or rejected when parsing an unsigned type, and whether unpunched data is accepted or rejected when parsing a signed type.		
1 0 3	- textStandardBase - textStandardGroupSeparator - textStandardExponentCharacter - textStandardInfinityRep - textStandardNanRep - textStandardZeroRep - textZonedSignStyle		
1 0	4 Unsigned decimal Proposed solution:		

| 4 | - Allow xs:nonNegativeInteger which enables unsigned unbounded integers to be modelled. The problem to solve is then just for xs:decimal.
- Call the new property dfdl:decimalSigned. It only applies to xs:decimal or user defined restrictions thereof. It applies to all physical decimals, as its name implies (not just zoned or packed). | | |

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**05-May-10 (Wed)**
Meeting Time	**13:00 - 14:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 06-May-10
Last Modified by Alan Powell on 06-May-10

OGF DFDL Working Group Call , May 05-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , May 05-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Suman Kalia (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Steve Marting (Progeny)
Mike Beckerle (Oco)

1. DFDL Specification status

Alan had contacted an automotive standards body (AIAG) to generate interest in getting comments. The request had been forwarded to a number of OMG groups some of whom have at least looked at the specification. He will encourage them to make comments.

Steve H had commented on the OMG MDMI site that the DFDL spec was available

2 Nils, defaults and unparsing

Alan had distributed tables and examples for defaulting complex elements. Steve H had updated the tables.

- emptyValueDelimiterPolicy does apply to complex elements.
- on parsing a complex element is empty if its content region is empty. That is there no child delimiters or content.
- on unparsing defaulting occurs if a complex item is missing or present present but has no children.

Defaulting for complex elements on parsing

initiator region	content region			infoset contains
matching	Not empty	default exists for any missing required children		complex element with child elements (real & default values)
matching	empty	default exists for any required children	*emptyValueDelimiterPolicy* must be **initiator or both**	complex element with child elements (real & default values)
empty	Not empty	default exists for any missing required children	*Initiator* must not be specified	complex element with child elements (real & default values)
empty	empty	default exists for any required children	*emptyValueDelimiterPolicy* must be **none** or *Initiator* must not be specified	complex element with child elements (real & default values)

Defaulting for complex elements on unparsing

infoset contains			initiator region contains	content region contains
not supplied	default exist for all missing required children and final representation of complex element is empty string	*emptyValueDelimiterPolicy* **none** or *Initiator* is not specified	empty	empty
supplied but with no children		*emptyValueDelimiterPolicy* **initiator or both**	initiator string	
supplied with some children	default values exist for all missing required content and final representation of	*Initiator* is not specified	initiator string	representation of complex element and its children
		Initiator is not specified	empty	

complex element is **not** empty string				

- The tables need to add terminator which may make them too cluttered
- It was agreed that on unparsing a simple infoset item with no value will be treated the same as missing. Steve H to check WMB behaviour. Need to state in 4.1.2 Infoset that value is optional.

- Conformed that emptyValueDelimiterPolicy only applies is initiator or terminator is set.

- Need to describe defaulting of choices.

- Confirmed that a separator is inserted/expected when postfix separator and supressedAtEnd.

3. Current Actions :
Updated below

4 AOB

The IBM team had questioned whether timezones should be just UTC+/- offset or should GMT and other short time zones be allowed. Alan to investigate

Tim also questioned the teston/off functions. Alan stated that setting bits in a array was not possible. Will look at ISO8583 to see if needed.

Meeting closed, 14:10

Next call Wednesday 12 May 2010 13:00 UK (8:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

Next action: **092**

Actions raised at this meeting

No	Action
091	The IBM team had questioned whether timezones should be just UTC+/- offset or should GMT and other short time zones be allowed. Alan to investigate

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public 05/05: Work still progressing
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc.

No	Action
	14/04: no progress 21/04: no progress 05/05: no progress
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes 21/04: Press release, OMG and other standards bodies. 05/05: Alan and Steve H have contacted other standards bodies. Will ask them to add comments on spec
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing 14/04: Investigate new property to control empty string behaviour. 21/04: After investigation a new property is not required. New rules developed and tables updated. Need examples of complexTypes to confirm tables apply. Review Nils, defaulting spec section. 05/05: Discussed defaulting complex elements. Tables updated but need to add terminator. SH; to confirm WMD behaviour when infoset item has no value on unparsing Need to describe defaulting choices.
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon. 14/04Tim had distributed a document starting the definition of calculable length for the longest choice member. Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO. 21/04: Need to define 'calculable length' and WMB importer ODO behaviour. 05/05: TK: Still need definition of calculable length. SKK: WMB COBOL imported behaviour with ODO
091	The IBM team had questioned whether timezones should be just UTC+/- offset or should GMT and other short time zones be allowed. Alan to investigate

Closed actions

No	Action
090	Semantics of separatorPolicy 14/04: Discussion on detecting when complex types should be defaulted. Examples needed. 21/04: Complex defaulting examples needed for new rules 05/05: Subsumed into action 086 defaults. CLOSED

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0	Update dfdl schema with change properties	ongoing	

3 6			
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		
0 9 7	Remove functions that returns duration		
0 9 8	occursCountKind is parsing only		
0 9 9	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.		
1 0 0	Need to state in 4.1.2 Infoset that value is optional.		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**21-Apr-10 (Wed)**
Meeting Time	**13:00 - 14:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 22-Apr-10
Last Modified by Alan Powell on 22-Apr-10

OGF DFDL Working Group Call , April 21-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , April 21-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Mike Beckerle (Oco)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Steve Marting (Progeny)
Suman Kalia (IBM)

1. DFDL Specification status

There have been no public comments so far but this is not unexpected. Steve H is pursuing a press release from IBM and OCO. Mike will send updated information to the OMG.
Alan will pursue contacts from the Athena EU project.

2 Nils, defaults and unparsing

Steve H, Tim and Alan have had further discussions and concluded and new property was not required so propose

- Elements that are zero length and elements that are 'missing' from the data stream should be treated the same.
- missingValueInitiatorPolicy should be renamed to emptyValueInitiatorPolicy to distinguish from a missing infoset value on unparsing
- Any 'initiator' policy needs to also include terminators so the enumerations should be extended to "none", "initiator", "terminator", "both". and the properties renamed emptyValueDelimiterPolicy and nilValueDelimiterPolicy .

The tables have been reworked

Nils and Defaults on Parsing

initiator region	content region			Logical Value
matching	content	*content matches nilValue*	*nilValueDelimiterPolicy* must be **initiator or both**	nil
		Content does not match nilValue		parse content to get value
matching	empty	*empty string is a nil value*	*nilValueDelimiterPolicy* must be **initiator or both**	nil
		A default is specified	*emptyValueDelimiterPolicy* must be **initiator or both**	default value (may be nil, ~~empty space~~ or value)
empty	content	*content matches nilValue*	*nilValueDelimiterPolicy* must be **none**	nil
		Content does not match nilValue	*Initiator* must not be specified	parse content to get value
			~~*Initiator* is specified~~	~~ERROR~~
empty	empty	*empty string is a nil value*	*nilValueDelimiterPolicy* must be **none**	nil
		A default is specified	*emptyValueDelimiterPolicy* must be **none**	default value (may be nil, ~~empty space~~ or value)

Note. The default value can be the empty space.

Nils and Defaults on Unparsing

Logical Value			initiator region contains	content region contains
Nil (implies nillable)	*nilValueDelimiterPolicy* **none**		empty	representation of nil
	nilValueDelimiterPolicy **initiator or both**		initiator string	
"" (empty string) (type is xs:string or xs:hexBinary)	*emptyValueDelimiterPolicy* **none**		empty	empty
	emptyValueDelimiterPolicy **initiator or both**		initiator string	
missing	useNilForDefault	*nilValueDelimiterPolicy* **none**	empty	representation of nil
		nilValueDelimiterPolicy **initiator or both**	initiator string	
	default is empty string (type is xs:string or xs:hexBinary)	*emptyValueDelimiterPolicy* **none**	empty	empty
		emptyValueDelimiterPolicy **initiator or both**	initiator string	
	Other default		initiator string	representation of the default value.
a non-nil non-empty-string value			initiator string	representation of the logical value

The above tables (with minor changes) were agreed.
The tables currently apply to simpleTypes, examples are needed to confirm that they are correct for complexTypes as well.
Steve questioned whether complexTypes can be nillable as WTX allow this. Stephanie will investigate.
Subsequently Stephanie confirmed that this was only needed for XML support in WTX so is not needed in DFDL.

The Nils, defaulting section in the spec will be reviewed/rewritten to incorporate the new understanding.

3. Current Actions :
Updated below.

4 Expressions during Unparsing .

In order to reduce complexity and potential performance problems it is proposed to
- nilValue and occursStopValue should NOT be able to take an expression. Agreed
- Very few properties should be allowed to have forward references. Table below agreed

Here's the list of properties that can take an expression, along with some suggested restrictions on forward referencing:

byteOrder	backward reference only
encoding	backward reference only
initiator	backward reference only

terminator	backward reference only
separator	backward reference only
outputNewline	backward reference only
length	~~simple elements ~~: backward reference only. ~~complex elements : forward reference allowed~~ ~~Quote "When unparsing with dfdl:lengthKind="explicit", the calculation of dfdl:representationLength() returns the value of the dfdl:length property." A simple element's length needs to be known in advance, else it cannot be parsed. A complex element's length might get resolved while parsing its content, so a forward reference within the scope of the complex element is possible.~~ ~~Having said that, we might want to impose a restriction in IBM's DFDL processor v1.0 that only backward references are allowed.~~
escapeCharacter	backward reference only
escapeEscapeCharacter	backward reference only
textNumberDecimalSeparator	backward reference only
textNumberGroupSeparator	backward reference only
textNumberExponentCharacter	backward reference only
binaryFloatRep	backward reference only
textBooleanTrueRep	backward reference only
textBooleanFalseRep	backward reference only
~~nilValue~~	~~backward reference only~~ ~~Not clear why this needs to be specified as an expression. (may already be on a list of open issues somewhere)~~
nilIndicatorPath	backward reference only
occursCount	backward reference only
~~occursStopValue~~	~~Not clear why this needs to be specified as an expression.~~
[inputValueCalc]	[Parser only] Backward reference only
outputValueCalc	Forward or backward reference

Meeting closed, 14:10

Next call Wednesday 28 April January 2010 13:00 UK (8:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee 's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **091**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested. 21/04. Schema for TDML defined. Need to define how this and the test cases will be made public
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc.

	14/04: no prorgess
	21/04: no progress
085	ALL: publicize Public comments phase to ensure a good review..
	14/04: see minutes
	21/04: Press release, OMG and other standards bodies.
086	AP: Nils and Defaults during unparsing – update table
	31/03: TK to documetn use cases for parsing
	14/04: Investigate new property to control empty string behaviour.
	21/04: After investigation a new property is not required. New rules developed and tables updated.
	Need examples of complexTypes to confirm tables apply.
	Review Nils, defaulting spec section.
088	define semantics of choiceKind 'fixedLength'
	31/03: TK to provide definition of calculable length.
	Investigate PL/I varchars and Cobol occurs dependingon.
	14/04Tim had distributed a document starting the definition of calculable length for the longest choice member.
	Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO.
	21/04: Need to define 'calculable length' and WMB importer ODO behaviour.
089	Need to be able to define unsigned decimal and integer
	31/03: MB to agree proposed solution
	14/04: MB to agree proposed solution
	21/04: Agreed. Closed
090	Semantics of separatorPolicy
	14/04: Discussion on detecting when complex types should be defaulted. Examples needed.
	21/04: Complex defaulting examples needed for new rules

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
0	Write DFDL primer		

7 0			
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		
0 9 7	Remove functions that returns duration		
0 9 8	occursCountKind is parsing only		
0 9 9	nilValue and OccuresStopValue cannot have an expression. On unparsing only outputValueCalc can have a forward reference.		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**14-Apr-10 (Wed)**
Meeting Time	**13:00 - 14:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 15-Apr-10
Last Modified by Alan Powell on 15-Apr-10

OGF DFDL Working Group Call , April 14-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , April 14-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Stephanie Fetzer (IBM)
Suman Kalia (IBM)
Tim Kimber(IBM)

Apologies
Mike Beckerle (Oco)

1. DFDL Specification status
After a slight delay the DFDL v1 Specification was opened to public comment on April 15th until June 15th.
There was some discussion about publicizing the review. Steve H mentioned he started the process for getting approval for IBM to be included in a press release. Steve M indicated that Progeny may also be willing.

Steve M mentioned that Progeny were working on an OMG mediation framework that requires DFDL and theat Progeny will have at least a partial implementation of DFDL that may satisfy the OGF requirement.

2 Nils, defaults and unparsing

Tim had distributed some uses cases of fixed and delimited formats with initiated and non-initaited element. The was considerable discussion of the initiated, delimited cases

```
ONE:one,THREE:three<END>
ONE:one,,THREE:three<END>
```

The discussion was mainly about whether was any difference in intent between missing and zero length in these examples and whether zero length string or default should be inserted into the infoset, Tim suggested using separatorPolicy required as a trigger. Alan pointed out that the table in the spec used the presence or absence of a default value to decide is the default or empty string should be the result. Eventually concluded that an additional property may be needed and a proposal will be developed.

In addition it was noted that the interaction between dfdl:initiatedContent and dfdl:missingValueInitiatorPoilcy may need clarifying.
It was also question whether XSDL allows the empty space to be specified as a default value.

3 dfdl:separatorPoilicy.

The main discussion was about how to detect when complex types should be defaulted which was similar to the defaulting discussion for simpleTypes. The added complication for complex types is that the content region is the children so should you detect missing by parsing the first child? It was agreed that examples simpler to the ones above would be helpful.

Meeting closed, 14:45

Next call Wednesday 21 April January 2010 13:00 UK (8:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **091**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update ... 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing 14/04: And XML test case format has been defined and is being tested.
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc. 14/04: no progress
085	ALL: publicize Public comments phase to ensure a good review.. 14/04: see minutes

086	AP: Nils and Defaults during unparsing - update table
	31/03: TK to documetn use cases for parsing
	14/04: Investigate new property to control empty string behaviour.
088	define semantics of choiceKind 'fixedLength'
	31/03: TK to provide definition of calculable length.
	Investigate PL/I varchars and Cobol occurs dependingon.
	14/04Tim had distributed a document starting the definition of calculable length for the longest choice member.
	Alan had done some investigation of COBOL occurs depending on and when used in the working section of a program then the maximum storage was reserved but when used in the linkage section the dependent number was uses. We need to understand how the WMB COBOL importer deals with ODO.
089	Need to be able to define unsigned decimal and integer
	31/03: MB to agree proposed solution
	14/04: MB to agree proposed solution
090	Semantics of separatorPolicy
	14/04: Discussion on detecting when complex types should be defaulted. Examples needed.

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration		
0	occursCountKind is parsing only		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**31-Mar-10 (Wed)**
Meeting Time	**13:00 - 14:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 31-Mar-10
Last Modified by Alan Powell on 31-Mar-10

OGF DFDL Working Group Call , March 31-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , March 31-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Stephanie Fetzer (IBM)
Mike Beckerle (Oco)
Tim Kimber(IBM)

Apologies
Suman Kalia (IBM)

0 DFDL specification status

The draft specification was passed by the OGF technical committee on March 30th so will now be made available for public c

1 Nils, defaults and unparsing
The semantics need clarifying

I have reorganized the table

Logical Value			initiator region contains
Nil (implies nillable)	nilValueInitiatorPolicy 'prohibited'		empty
	nilValueInitiatorPolicy 'required'		initiator string
"" (empty string) (type is xs:string or xs:hexBinary)	missingValueInitiatorPolicy 'prohibited'		empty
	missingValueInitiatorPolicy 'required'		initiator string
missing	useNilForDefault	nilValueInitiatorPolicy 'prohibited'	empty
		nilValueInitiatorPolicy 'required'	initiator string
	default is empty string (type is xs:string or xs:hexBinary)	missingValueInitiatorPolicy 'prohibited'	empty
		missingValueInitiatorPolicy 'required'	initiator string
	Other default		initiator string
a non-nil non-empty-string value			initiator string

Note: If nilValue is the empty string then missingValueInitiatorPolicy is not used.

Other changes:

missingValueInitiatorPolicy	Enum Valid values 'required', 'prohibited'
	Specifies whether to expect an initiator when an element is missing. Ignored unless dfdl:initiator is sp string).
	'required'- Indicates that the dfdl:initiator followed by empty content is the required syntax to indicate t
	'prohibited' - Indicates that empty content is the required syntax to indicate that the element is missing initiator implies that real content must follow.
	Use of 'prohibited' implies an ordered sequence. If used on an initiated element of an unordered grou error.
	If the element is required, defaulting occurs as defined above. This property also applies on unparsing, when the data to be written (after ~~nil value and~~ default value content.
	Annotation: dfdl:element (string or hexbinary)

The WG agreed that the unparsing table was correct though there was some discussion and misunderstanding about the con empty string did not use missingValueInitiatorPolicy. The comment was asking fro agreement from reviewers and does not a tavle. The meaning of the columns needs a better explanation.

SH pointed out that missingValueInitiatorPolicy applies to all simple types not just string and hexbinary and to complex types parsing and needs to include unparsing.

There was a lot of discussion about paring behaviour. In particular whether there needs to be a differentiation between 'zero from the data stream'.

Action: Tim will document the parsing use cases so we can verify the parsing table.

2 dfdl:choiceKind 'fixedLength'

The main issues are:
a) The calculation of the length of the longest branch is not obvious.
b) The length units to use - the dfdl:lengthUnits property does not exist on a choice
c) The name could be better

Proposal is therefore to retain the property but to:
i) State the conditions that must apply to use this property, and enforce them in the validator => schema definition error other
ii) Decouple the choice from its parent by calculating the length of each branch based solely on the properties of the branches of any parent dfdl:lengthKind

The use cases are COBOL redefines. C unions and PL/1 ?

Restrictions:
- At least one of the choice branches must be of calculable length. (Tim to provide definition of calculable)

To be calculable
- minOccurs must equal maxOccurs for all children
- All lengths will be calculated in bytes. All encodings must be fixed width.

Actions:
- investigate how PL/1 allocates storage for varchars
- investigate how COBOL allocates storage for 'occurs dependingon' and how these are modelled in DFDL.

3 dfdl:occursCountKind 'expression'

When dfdl:occursCountKind is 'expression' the occursCount should only be used when parsing. On unpasring the number of minOccurs etc

Note this behaviour is different from dfdl:lengthKind.

Agreed that the dfdl:occursCountKind 'expression' is only used on parsing.

4 Unsigned decimal

Whether a logical type is signed or not is used to determine whether a sign is output in some representations. However there or integer so these will always be output with a sign so COBOL declarations such as
```
05  CFCFDN-FLD12              PIC  9(05)V99 COMP-3.
```

cannot be supported.

Proposed solution:

- Allow xs:nonNegativeInteger which enables unsigned unbounded integers to be modelled. The problem to solve is then jus

- Call the new property dfdl:decimalSigned. It only applies to xs:decimal or user defined restrictions thereof. It applies to all p name implies (not just zoned or packed).

We considered added a dfdl:unsignedDecimal type but rejected it as causing confusion between dfdl type restrictions and us

The preferred solution is as proposed. Action Mike to agree.

5 dfdl:separatorPoilicy.

dfdl:separatorPolicy enumeration definitions are not complete as they do not discuss arrays or defaulting.

separatorPolicy ="required"
All the separators must be there.
No group member, except for required members with a default, may be omitted from the data stream.
Every group member must have a separator in the correct position.
Optional members and variable-occurrence arrays will be allowed, separators will be expected for all maxOccurs
** It is a schema definition error if maxOccurs="unbounded" for any member of the group

separatorPolicy ="suppressed"
Any group member (optional or required with a default) can be omitted from the data stream, in which case its separator ~~mu~~

For prefix and infix, if a separator is found then another group member is expected.
For postfix, if enclosing markup is not found immediately after a postfix separator then another group member is expected.

Enclosing markup or end of data can occur where an infix or prefix separator is expected. This terminates the group.
Enclosing markup or end of data can occur immediately after a postfix separator (if dfdl:documentfinalSeparatorCanBeMissir separator may be missing at the end of the datastream). This terminates the group.

If the group terminates before all of its members have been parsed then any required members in the remainder of the group infoset.
The entire group can have a zero-length representation, which is indicated by enclosing markup at the start of the group's co

If another group member is expected and the next group member is a point of uncertainty then the identity of the next group r determined
by resolving the point of uncertainty as described in section XXX.

If another group member is expected but no group member is found in the data stream then it is a processing error.
If another group member is expected but all group members have been parsed then it is a processing error.

separatorPolicy ="suppressedAtEnd "

suppressAtEnd has the same rules as 'required' up to the last required member that does not have a default, then it has the s

Open questions
1. If a group has separatorPolicy='required' and one of its optional members has a zero-length representation, and that mem
zero-length representation, what should the parser do?
A similar question arises for required, defaultable elements with a zero-length representation when they are based on a type
Should the default value be used, or the empty string?
2. How should the occursXXX properties be applied when there are other ways to determine the number of occurences (e.g.
3. What algorithm should the parser use to decide whether to default a complex element or group (when the element/group i
group)?
4. How does unparsing work for separated groups in all of these scenarios?

The changes above were suggested. The unparsing behaviour needs adding.
There was discussion about defaulting of complex elements and when this should occur. More discussion is needed.

6 Current Actions
Updated below

Meeting closed, 15:00

Next call Wednesday 7 April January 2010 13:00 UK (8:00 ET)

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create
Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items
will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **091**

Actions raised at this meeting

No	Action
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon.
089	Need to be able to define unsigned decimal and integer 31/03: MB to agree proposed solution
090	Semantics of separatorPolicy

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still investigating 10/02: IBM is still investigating

	17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressing 31/03: work is progressing
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc.
085	ALL: publicize Public comments phase to ensure a good review..
086	AP: Nils and Defaults during unparsing - update table 31/03: TK to documetn use cases for parsing
088	define semantics of choiceKind 'fixedLength' 31/03: TK to provide definition of calculable length. Investigate PL/I varchars and Cobol occurs dependingon.
089	Need to be able to define unsigned decimal and integer 31/03: MB to agree proposed solution
090	Semantics of separatorPolicy

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration		
0	occursCountKind is parsing only		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**17-Mar-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 18-Mar-10
Last Modified by Alan Powell on 18-Mar-10

OGF DFDL Working Group Call , March 17-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , March 17-2010

Attendees
Suman Kalia (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Stephanie Fetzer (IBM)
Mike Beckerle (Oco)

Apologies
Tim Kimber(IBM)

Steve H reported on the DFDL presentation at OGF 28
Overall OGF registered attendees thought to be around 250.

Number at DFDL session: 7. Apart from Erwin (Data AD), DFDL was new to everyone.

Notes from session (from minutes taken by Erwin Laure based on questions asked):
- Validation on input and output? Means conforming to the schema defined (e.g. integer between 0 and 100). dfdl assert c[...] for asserting correctness of the data. Further complex validation would be via an external step (eg, Schematron invocatio[...]
- Using DFDL to model data structures generated by OO code, particularly use of inheritance? Could be a use case for allo[...] complex type inheritance in DFDL in the future.
- Data often comes with units, infoset would want to reflect this? Could be achieved using expressions, dfdl:inputValueCal[...] dfdl:hidden to apply a scaling factor based on units.
- What's the efficiency of DFDL? Particularly, can tests be turned off for fast read/write? The spec defines the behavior of t[...] not how to implement it. For instance, validation is not mandatory.
- Open source reference implementation would be good to have. (Lots of nodding).
- Scientific floating point data compresses badly. Knowing data structure can allow a more intelligent compression. DFDL i[...] intending to do transformations but that could be done on top.
- Encryption/compression could also be a use case for multi-layers, or for additional functions in the expression language.[...]
- Need to reach out to DAIS-WG and DR-WG for (public) comments and to see whether DFDL will be actively used by thos[...] Open source implementation would obviously help here.
- Comment on the spec: Should be a proposed recommendation (GFD-P-R) not informational (GFD-I).

Draft 040 has been submitted to OGF and will be reviewed by the technical committee on March 30th. It is expect[...] for public comment in early April.

There needs to be some publicity to ensure that enough comments are made.
Mike suggested a press release from his company with IBM comment. Action raised
May be able to get IBM representatives on other standards bodies such as ACCORD. OMG, to interest their memb[...]
Should contact all the contributors to the WG to ask them to review.
Mike will update the information he sent in response to an RFI from OMG

1 Process for dealing with internal issues during Public comment phase   .
The public comments process will be used to make updates

2. Nils and Defaults during unparsing
The table 17 in section 13.16.2 was corrected in draft 40 but here are still some ambiguities, for example when nil is the defa[...]

Logical Value	nilValueInitiatorPolicy	Has default value specified	missingValueInitiatorPolicy	initiator region contains	content region[...]
Nil (implies nillable)	prohibited		don't care	empty	representation of nilKind, nilVa[...]
	required			initiator string	
"" (empty string) Note that this implies that the element type is xs:string	don't care	don't care	prohibited	empty	empty st[...]
			required	initiator string	
a non-nil non-empty-string value	don't care		don't care	initiator string	The representation[...] value[...]
Not supplied	don't care	Yes (non-empty sting)	don't care	Initiator string	The representation[...] value.[...]
		Yes (empty string)	prohibited	empty	empty[...]
			required	initiator string	empty[...]

3 dfdl:choiceKind

The main issues are:
a) The calculation of the length of the longest branch is not obvious.
b) The length units to use - the dfdl:lengthUnits property does not exist on a choice
c) The name could be better

Proposal is therefore to retain the property but to:
i) State the conditions that must apply to use this property, and enforce them in the validator => schema definition error other
ii) Decouple the choice from its parent by calculating the length of each branch based solely on the properties of the branches
irrespective of any parent dfdl:lengthKind

4 DFDL time functions

`fn:timezone-from-dateTime`	Returns the timezone from an `xs:dateTime` value.

`fn:timezone-from-date`	Returns the timezone from an `xs:date` value.
`fn:timezone-from-time`	Returns the timezone from an `xs:time` value.

Function	Meaning
`fn:adjust-dateTime-to-timezone`	Adjusts an `xs:dateTime` value to a specific timezone, or to no all.
`fn:adjust-date-to-timezone`	Adjusts an `xs:date` value to a specific timezone, or to no timez
`fn:adjust-time-to-timezone`	Adjusts an `xs:time` value to a specific timezone, or to no timez

All return an xs:duration. Do we need these functions?

Meeting closed, 14:10

Next call Wednesday 17 March January 2010 13:00 UK (9:00 ET)
Call will be for one hour only
NOTE: East coast is 4 hours behind the UK for the next two weeks

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views . " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **087**

Actions raised at this meeting

No	Action
085	ALL: publicize Public comments phase to ensure a good review..
086	AP: Nils and Defaults during unparsing - update table

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format . 04/12: no update 09/12: no update 16/12: reminded dent to project manager

	23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still investigating 10/02: IBM is still investigating 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided. 10/03: work is progressing 17/03: work is progressin
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc.
085	ALL: publicize Public comments phase to ensure a good review..
086	AP: Nils and Defaults during unparsing - update table

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
097	Remove functions that returns duration		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**10-Mar-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 16-Mar-10
Last Modified by Alan Powell on 16-Mar-10

OGF DFDL Working Group Call , March 10-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , March 10-2010

Attendees
Suman Kalia (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Mike Beckerle (Oco)

1. Spec issues

documentFinalTerminatorCanBeMissing :
Let's try to avoid creating another property for the postfix separator scenario. I think this property provides a
way of modelling the data naturally.

We can recommend use of infix-with-a-terminator rather than 'postfix' if the final terminator can be missing.
WG deferred any changes the property at this stage

dfdl:lengthKind endOfParent

'endOfParent' has almost the same meaning as 'delimited' so should have the same semantics.

- the item's terminator (if specified)
- an enclosing construct's separator or terminator
- the end of an enclosing construct designated by its known length
- the end of the data stream

The effect would be the the element could be ended by the nearest known length parent not just the immediate parent. Also the immediate parent could have lengthKind 'implicit'

Agreed with the definition and that dfdl:lengthKind 'endOfParent should apply to simpleTypes only.

dfdl:lengthUnits 'characaters'

LengthUnits 'characters' is currently for simple representation 'text' only. Should it be allowed for complex?
Agreed dfdl:lengthUnits should be allowed on complex elements. The characters are in the encoding of the complex element. The endocings of the children must be the same as the parent..

Selectors

Should they be moved to an appendix.
Agreed should moved

DFDL Expression Syntax

Value comparison operators are 'eq' 'lt' 'gt' etc not '=' '<' '>' because these were defined poorly in XPATH v1 and replaced in XPATH V2. Unfortunately most examples us '=' etc
The examples need changing and '=' '<' will not be accepted

Syntax of expressions in string literals. Add
- A DFDL string literal which begins with { and does not end with } does not need to have its leading { escaped.

The syntax rules for expressions are
- Must start with a '{' in the first position and end with '}' in the last position.
- '{' in any other position if treated as a literal
- '}' in any position other than the last position is treated as a literal.
- '{{' as the first characters are treated as the literal '{' and not a DFDL expression.

IgnoreCase

Should ignoreCase apply to numberInfinityRep and numberNaNRep
Agreed

Escapes.

Should escapeCharacter, escapeEscapeCharacter etc allow DFDL expressions.
Agreed

OutputNewLine should allow an expression
Agreed

Clarify dfdl : InputValueCalc and outputValueCalc
It is a schema definition error if both inputvalueCalc and outputValueCalc are specified.
InputValueCalc is examined on unparsing so needs to be added to unparsing property precedence.

Action to review inputValueCalc and putputValueCalc

2. Current Actions :
Updated below

3 Status of specification   (for OGF 28)
Draft 40 will be published by tomorrow and submitted to OGF on friday.

Meeting closed, 14:10

Next call Wednesday 10 March January 2010 13:00 UK (9:00 ET)
NOTE: East coast is 4 hours behind the UK for the next two weeks

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **085**

Actions raised at this meeting

No	Action
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc.

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still investigating 10/02: IBM is still investigating 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress

	03/03: Discussions have been taking place on the subset of tests that will be provided.
	10/03: work is progressing
084	Check behaviour of dfdl:inputValueCalc and outputValueCalc.

Closed actions

No	Action
080	AP:Clarify semantics of fn:poisition and fn:count
	17/02: no progress
	24/03: No progress
	03/03: no progress. There are other functions which return a duration which need investgating
	10/03: changed to dfdl:position and dfdl:count. closed

Work items:

N o	Item	target version	stat
0 0 5	Improvements on property descriptions		not
0 1 2	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0 3 6	Update dfdl schema with change properties	ongoing	
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 8 3	Implement RFC2116		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**03-Mar-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 03-Mar-10
Last Modified by Alan Powell on 03-Mar-10

OGF DFDL Working Group Call , March 03-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plai
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or
what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , March 03-2010

Attendees
Suman Kalia (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Peter Lambros (IBM)
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

Apologies
Mike Beckerle (Oco)

1. 16.2 scannablility with lengthKind pattern :

In summary, you can use a data pattern on any element (complex, simple text, simple binary) as long as the
bytes are legal in the stated encoding, which where binary data is involved in practice means an 8-bit ASCII

encoding.

Binary data can be handled using some of the conveniences of text by way of treating it as text with encoding="iso-8859-1". In this case literal text, such as length patterns, is interpreted as in the iso-8859-1 character encoding, and the correspondence of byte values in the data to a string in the DFDL infoset is one to one. That is, byte with value N, produces an infoset character with character code N.

The WG agreed to the solution of recommending the use of "iso-8859-1" for binary fields. The write-up should note that he main use case is for a complex element that contains binary fields.

2. Current Actions :
Updated below

3 Steve H issues with draft 039

1) Name of property dfdl:textNumberRepresentation is not consistent with dfdl:binaryNumber**Rep**, dfdl:binaryFloat**Rep**, etc.
Agreed

2) The dfdl:numberPattern etc properties that have been moved from the defunct dfdl:textNumberFormat annotation to dfdl:element etc should be called dfdl:**text**NumberPattern etc. Otherwise users will think they apply to binary numbers too.
Agreed

3) In section 14.3 on sequences, there are several sub-sections that talk about parsing according to different ways of specifying length (ie, lengthKind). But dfdl:sequence no longer carries dfdl:lengthKind so I think these sub-sections are not in the right place. I think they should be in section 12, under the correct 12.3.x lengthKind sub-section.
MB has moved these sections

4) Section 19 on built-in specifications. Given that we don't have any for public comment phase we should reword this section.
Agreed

4 Tim's (major) issues with draft 039

12.2 Delimiters : Text Markup
- The term 'Delimiters' is not accurate. Most readers will not think of an initiator as a 'delimiter'.
- It's not 'Text' markup any more - especially since v0.39 has allowed lengthKind="delimited" for elements with binary representation.
Title should be 'Markup' and explanation can then deal with what it really is, rather than justifying the innaccurate title :-)
Will be changed to Delimiters - Add delimiters to glossary (DFDLlimits delimiters to terminators and separator when dfdl:lengthKind=delimited)

Syntax for specifying markup :
It's not clear from this description that each item in the space-separated list is a DFDL string literal.
Section 6.3.1 will be revised to separate string literals and dfdl expressions etc.
If possible will add the list of property types (listof DFDL string literals, etc) and update property tables

initiator (and all other space-separated properties)

It is not clear whether the order of the space-separated properties matters. Must the parser test them in the order in which they are specified?

(Q: What if %ES; is the first in the list?)

The parser must look for the longest first. Unparsing will output the first in the list.

The %ES entity should be not be allowed for initiator, separator and terminator. Should probably be limited to use in nilValues but will check whether it is useful in other properties..

occursStopValue should a value in the value space of the base type of the array.

terminator:

is it OK if the final terminator is missing within the scope of a known-length parent? Seems like a reasonable extension of the rule (in all other scenarios, the end of a known-length parent acts like the end of the data stream for items with its scope).

documentFinalTerminatorCanBeMissing :

Let's try to avoid creating another property for the postfix separator scenario. I think this property provides a way of modelling the data naturally.

We can recommend use of infix-with-a-terminator rather than 'postfix' if the final terminator can be missing.

outputNewLine

Should we validate that the 'characterOrCharacters' are all newline characters from the set described by the %NL; mnemonic? Otherwise the DFDL serializer will output data which cannot be parsed by the DFDL parser.

Agreed

dfdl:lengthKind endOfParent

'endOfParent' has almost the same meaning as 'delimited' so should have the same semantics.

- the item's terminator (if specified)
- an enclosing construct's separator or terminator
- the end of an enclosing construct designated by its known length
- the end of the data stream

The effect would be the the element could be ended by the nearest known length parent not just the immediate parent. Also the immediate parent could have lengthKind 'implicit'

Not discussed but emails indicate Agreed

choiceKind 'Fixed'

When lengthKind='implicit' all alternative branches of the choice are padded to the fixed length of the largest one so that overall the entire choice construct is fixed length

There must be a restriction that the length of at least one choice must be statically defined.

Not discussed but emails indicate Agreed

5 DFDL v1 Specification completion .

Draft 40 will be published early next week. Will be sent to OGF next week prior to OGF 28

Meeting closed, 14:10

Next call Wednesday 10 March January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **084**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format .

No	Action
	04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still investigating 10/02: IBM is still investigating 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress 03/03: Discussions have been taking place on the subset of tests that will be provided.
080	AP:Clarify semantics of fn:poisition and fn:count 17/02: no progress 24/03: No progress 03/03: no progress. There are other functions which return a duration which need investgating

Closed actions

No	Action
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress 29/01: No progress. The predefined formats do not need to be available when the spec is published. Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numbers are define is very usable. He will document for next call 03/02: No progress 10/02: No progress 17/03: No progress 24/03: No progress 03/03: The wording of this section will be changed to say that 'implementations may choose to provide predefined formats. The DFDL WG intends to supply a limited set'

	Closed
079	MB:Encoding for binary fields when lenghtkind is pattern 17/02: Discussed but no conclusion 24/03: Mike has found an encoding that matches the first 255 codepoints of iso 10646. Will document its use for binary fields. 03/03: Wording in minutes agreed. Closed
083	MB:To correct syntax diagram for FinalUnused and suggest wording for the Sequence section 03/03: Mike has supplied updates to Sytax, length and Sequence sections. Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
070	Write DFDL primer		
071	Write test cases.		
083	Implement RFC2116		
084	MB:Encoding for binary fields when lenghtkind is pattern		
085	MB:To correct syntax diagram for FinalUnused and suggest wording for the Sequence section		
086	dfdl:textNumberRepresentation to dfdl:textNumberRep		
087	dfdl:NumberPattern etc to dfdl:**text**NumberPattern (not for calendarformat but check individual properties)		
08	Update built-in specifications section.		

8			
0 8 9	12.2 Will be changed to Delimiters - Add delimiters to glossary (DFDLlimits delimiters to terminators and separator when dfdl:lengthKind=delimited)		
0 9 0	Section 6.3.1 will be revised to separate string literals and dfdl expressions etc. If possible will add the list of property types (listof DFDL string literals, etc) and update property tables		
0 9 1	The parser must look for the longest first. Unparsing will output the first in the list. The %ES entity should be not be allowed for initiator, separator and terminator. Should probably be limited to use in nilValues but will check whether it is useful in other properties ..		
0 9 2	occursStopValue should a value in the value space of the base type of the array.		
0 9 3	outputNewLine MUST be one of newline characters from the set described by the %NL; mnemonic		
0 9 4	**dfdl:lengthKind endOfParent** 'endOfParent' has almost the same meaning as 'delimited' so should have the same semantics. • the item's terminator (if specified) • an enclosing construct's separator or terminator • the end of an enclosing construct designated by its known length • the end of the data stream The effect would be the the element could be ended by the nearest known length parent not just the immediate parent. Also the immediate parent could have lengthKind 'implicit'		
0 9 5	ChoiceKinf fixed:There must be a restriction that the length of at least one choice must be statically defined.		
0 9 6	The WG agreed to the solution of recommending the use of "iso-8859-1" for binary fields. The write-up should note that he main use case is for a complex element that contains binary fields.		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**24-Feb-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 24-Feb-10
Last Modified by Alan Powell on 24-Feb-10

OGF DFDL Working Group Call , February 24-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar
details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , February 24-2010

Attendees
Mike Beckerle (Oco)
Suman Kalia (IBM)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Peter Lambros (IBM)
Stephanie Fetzer (IBM)

Apologies
Tim Kimber(IBM)

1. Remaining 037 review issues

A:

16.2 scannablility with lengthKind pattern :

Confirm that this is what we agreed

In summary , you can use a data pattern on any element (complex , simple text , simple binary) as long as the bytes are lega
in the stated encoding , which where binary data is involved in practice means an 8-bit ASCII encoding .

Mike B

I found an official reference which has no "greyed out" codepoints. All 256 values are "mapped".

The following ftp table (see URL below) officially defines the mapping for 8859-1 to unicode/iso10646.

The table includes all 256 codepoints - some are specified as just <control> i.e., have no specific meaning, but thei
8859 codepoint maps one-to-one and onto a unicode/10646 codepoint with the same value.

Note that this property holds for 8859-1. It does not hold for 8859-2 to 8859-16, as these have character codes
substituted into them that map to other places in the iso10646 codepoint space.

Here's the correspondence table:

ftp://ftp.unicode.org/Public/MAPPINGS/ISO8859/8859-1.TXT

If we reference this mapping table in the references of the DFDL spec, then I believe we can say that using
encoding="iso-8859-1", you can treat binary data as textual, use patterns, etc., and the relationship to/from the info
always insures preservation of the values of the bytes (parsing), and creation of bytes whose values exactly match t
string codepoints (unparsing).

This language can be added to the section on lengthKind="pattern" and binary data:

Binary data can be handled using some of the conveniences of text by way of treating it as text with
encoding="iso-8859-1". In this case literal text, such as length patterns, is interpreted as in the iso-8859-1 characte
encoding, and the correspondence of byte values in the data to a string in the DFDL infoset is one to one. That is, b
with value N, produces an infoset character with character code N. [reference to above FTP site].

B:
Glossary

Variable-Occurrence Item - Optional elements have a variable number of occurrences (0 or 1) and arrays also can have a
variable number of occurrences (when minOccurs < maxOccurs). So when we say an item with a variable number of
occurrences, this can mean either an optional element, or an array where minOccurs < maxOccurs. In either array or optiona
elements, we have the additional constraint that the DFDL representation properties do not preclude a variable number of
occurrences. When dfdl:occursCountKind='explicit' and dfdl:occursCount has a literal constant as its value, or an expression
statically evaluates to a constant, then the DFDL properties are specifying exactly the number of occurrences for all instances
and so are said to preclude a variable number of occurrences. If dfdl:occursCount has a formula as its expressed value, then
DFDL properties do not preclude a variable number of occurrences.

MikeB Comment:

This idea that you can have minOccurs < maxOccurs, but dfdl:occurs is equal to a constant and dfdl:occursKind="explicit" is
causing us a bunch of grief in these definitions.

Can we be conservative and just say it is a schema definition error if minOccurs < maxOccurs but the length is static, i.e., an
explicit constant-valued expression?

WG decided that the wording can remain as currently written

C:

DFDL Schema Component Model

Only changes needed are: remove wildcards.

Add following to describe shading:

The shaded boxes have a direct corresponding element syntax and therefore appear in DFDL schema

D:

Sequence Groups

Mike B: To correct syntax diagram for FinalUnused and suggest wording for the Sequence section

E: Check other comments in document .
Please look at the remaining comments in draft 039 and suggest solutions

2. Go through Actions
Updated below

3 DFDL v1 Specification completion .

Draft 039 will be publish today.
WG review and Comments by 3 March
Draft 40 with updates for OGF submission - available 5 March

Meeting closed, 14:00

Next call Wednesday 3 March January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

Next action: **084**

Actions raised at this meeting

No	Action
083	MB:To correct syntax diagram for FinalUnused and suggest wording for the Sequence section

Current Actions :

No	Action
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update

	25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress 29/01: No progress. The predefined formats do not need to be available when the spec is published. Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numbers are define is very usable. He will document for next call 03/02: No progress 10/02: No progress 17/03: No progress 24/03: No progress
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still investigating 10/02: IBM is still investigating 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite' 24/03: No progress
079	MB:Encoding for binary fields when lenghtkind is pattern 17/02: Discussed but no conclusion 24/03: Mike has found an encoding that matches the first 255 codepoints of iso 10646. Will document its use for binary fields.
080	AP:Clarify semantics of fn:poisition and fn:count 17/02: no progress 24/03: No progress
081	AP: Inf and Nan The description is the way ICU behaves but need clarification. It isn't clear how inf and Nan are represented in the infoset. Need to investigate if XML allows these values 17/02: XML allows Nan and inf for float and double Dfdl will do the same. Requires more investigation of ICU. 24/03: Alan send clarification that Inf and Nan are limited to float and Double. Closed.
083	 MB:To correct syntax diagram for FinalUnused and suggest wording for the Sequence section

Closed actions

No	Action
081	AP: Inf and Nan The description is the way ICU behaves but need clarification. It isn't clear how inf and Nan are represented in the infoset. Need to investigate if XML allows these values 17/02: XML allows Nan and inf for float and double Dfdl will do the same. Requires more

investigation of ICU.
24/03: Alan send clarification that Inf and Nan are limited to float and Double. Closed.

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
069	ICU fractional seconds	039	
070	Write DFDL primer		
071	Write test cases.		
072	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix	039	
073	Rename dfdl:separatorPolicy="required" to "always".	039	Def
078	document UPA checks	039	
079	Semantics of length=0, nil handling and defaults. (A071)	039	
080	Tlog: Allow LengthKind delimited for packed/bcd (A074)	039	
081	Update empty sequence section (A075)	039	
08	semantics of minOccurs= 0 on choice branches (A076)	039	

2			
0 8 3	Implement RFC2116		
0 8 4	Length|Kind pattern scanability rules	039	
0 8 5	Invalid character substitution	039	
0 8 6	infoset round tripi: Rephrase sentence 'It is possible to define a schema so that when infoset unparsed and the datastream reparsed, the same infoset will be produced'	039	
0 8 7	Clarify use of relative paths in global components.	039	
0 8 8	'DFDL expression'	039	
0 8 9	Ageed that dfdl:represetnation 'text' is implied for strings and dfdl:represetnation 'binary' is implied for hexbinary	039	
0 9 1	textStringPadCharacter textNumberPadCharacter must be a 1 byte character if the char set encoding is variable width?	039	
0 9 2	finalDocumentTerminatorCanBeMissing and finalDocumentSeparatorCanBeMissing allowed only in 'default' format	039	
0 9 3	remove textNumberFormat and textCalendarFormat.	039	
0 9 4	Alignment should be 1 based	039	
0 9 5	AP: Inf and Nan	039	

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**17-Feb-10 (Wed)**
Meeting Time	**13:00 - 14:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 18-Feb-10
Last Modified by Alan Powell on 18-Feb-10

OGF DFDL Working Group Call , February 17-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , February 17-2010

Attendees
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Peter Lambros (IBM)
Stephanie Fetzer (IBM)

Apologies
Mike Beckerle (Oco)
Suman Kalia (IBM)
Tim Kimber(IBM)

1. Comments of latest discriminators doc v 5.

Steve questioned: It is a processing error if none of the choice branches are known to exist

Steve: I know we agreed on this but I'm concerned that I am now unable to model the situation where I have a header, a footer, and in between either nothing or exactly 1 of a choice of n records

It was agreed that the body could be an optional complex element with the choice as its content.

Also confirmed

dfdl:discriminator can be an annotation on
- a local xs:element declaration

- an xs:element reference

- an xs:group reference (when the top level of a choice branch)

- an xs:sequence (when the top level of a choice branch)

- an xs:choice (when the top level of a choice branch)

Action 045 will be closed.

2. Remaining 037 review issues

16.2 scannablility with lengthKind pattern :
Confirm that this is what we agreed
In summary, you can use a data pattern on any element (complex, simple text, simple binary) as long as the bytes are legal in the stated encoding , which where binary data is involved in practice means an 8-bit ASCII encoding.

By 8-bit ASCII we really mean an encoding where all the codepoints from 0-255 map to the equivalent value. Subsequent investigation indicates that 'all' 8-bit ASCII encodings have gaps so there isn't a valid character.
Mike has suggested
1) for all ascii-based character sets, we say that bytes 0x00 to 0xFF all map to exactly those codepoints in ISO 10646 for the infoset, and vice versa.

2) define dfdl:encoding="bytes" as a special character set name which has the above property.

Briefly discussed but no conclusion.

3. Go through Actions
Updated below

4. Biztalk comparison
Short discussion which noted that DFDL could match Biztalk representation properties but was missing support for attributes and some simple types. Decided that we don't need to do anything in DFDL v1.

5 DFDL v1 Specification completion.
It was agreed that the spec should be in the OGF review process by OGF 28 in March.
Alan stated it needed 2 more drafts.
Draft 39 with all the updates completed for a final WG review - available 24 Feb
Comments by 3 March
Draft 40 with updates for OGF submission - available 5 March

Meeting closed, 14:00

Next call Wednesday 24 February January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **083**

Actions raised at this meeting

No	Action

No	Action

Current Actions :

No	Action
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress 29/01: No progress. The predefined formats do not need to be available when the spec is published. Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numbers are define is very usable. He will document for next call 03/02: No progress 10/02: No progress 17/03: No progress
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still investigating 10/02: IBM is still investigating 17/02: IBM is willing in principle to publish the test case format and some of the test cases. May need some time to build a 'compliance suite'
079	MB:Encoding for binary fields when lenghtkind is pattern 17/02: Discussed but no conclusion
080	AP:Clarify semantics of fn:poisition and fn:count

No	Action
	17/02: no progress
081	AP: Inf and Nan The description is the way ICU behaves but need clarification. It isn't clear how inf and Nan are represented in the infoset. Need to investigate if XML allows these values 17/02: XML allows Nan and inf for float and double Dfdl will do the same. Requires more investigation of ICU.

Closed actions

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way. 29/07: No Progress 05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description. 13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms. 20/01: Mike will write up 27/01: further discussion of discriminators 29/01: Alan had emailed both proposals but not enough time to discuss 02/02: Agreed to adopt 'component exists' semantics for discriminators 10/02: 'component exists' proposal updated. comments by next call. 17/02: reviewed needs minor updates. Closed
082	MB: Should alignment be 0 or 1 based 17/02: Mike has reworded the section to be 1 based and Agreed this was OK Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
069	ICU fractional seconds	039	
070	Write DFDL primer		
071	Write test cases.		
072	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix	039	
073	Rename dfdl:separatorPolicy="required" to "always".	039	Def
078	document UPA checks	039	
079	Semantics of length=0, nil handling and defaults. (A071)	039	
080	Tlog: Allow LengthKind delimited for packed/bcd (A074)	039	
081	Update empty sequence section (A075)	039	
082	semantics of minOccurs= 0 on choice branches (A076)	039	
083	Implement RFC2116		

0 8 4	Length|Kind pattern scanability rules		
0 8 5	Invalid character substitution		
0 8 6	infoset round tripi: Rephrase sentence 'It is possible to define a schema so that when infoset unparsed and the datastream reparsed, the same infoset will be produced'		
0 8 7	Clarify use of relative paths in global components.		
0 8 8	'DFDL expression'		
0 8 9	Ageed that dfdl:represetnation 'text' is implied for strings and dfdl:represetnation 'binary' is implied for hexbinary		
0 9 1	textStringPadCharacter textNumberPadCharacter must be a 1 byte character if the char set encoding is variable width?		
0 9 2	finalDocumentTerminatorCanBeMissing and finalDocumentSeparatorCanBeMissing allowed only in 'default' format		
0 9 3	remove textNumberFormat and textCalendarFormat.		
0 9 4	Alignment should be 1 based		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**10-Feb-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 11-Feb-10
Last Modified by Alan Powell on 11-Feb-10

OGF DFDL Working Group Call , February 10-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , February 10-2010

Attendees
Mike Beckerle (Oco)
Steve Hanson (IBM)
Alan Powell (IBM)
Suman Kalia (IBM)
Peter Lambros (IBM)
Tim Kimber(IBM)
Stephanie Fetzer (IBM)

Apologies
Steve Marting (Progeny)

1. Comments of latest discriminators doc .
Stephanie has sent some editorial comments. Action will be closed next week subject to comments from others.

2. Remaining 037 review issues

16.2 scannablility with lengthKind pattern :
Confirm that this is what we agreed
In summary, you can use a data pattern on any element (complex, simple text, simple binary) as long as the bytes are legal in the stated encoding, which where binary data is involved in practice means an 8-bit ASCII encoding.

By 8-bit ASCII we really mean an encoding where all the codepoints from 0-255 map to the equivalent value. Subsequent investigation indicates that 'all' 8-bit ASCII encodings have gaps so there isn't a valid character.
Mike has suggested
1) for all ascii-based character sets, we say that bytes 0x00 to 0xFF all map to exactly those codepoints in ISO 10646 for the infoset, and vice versa.
2) define dfdl:encoding="bytes" as a special character set name which has the above property.
Action rasied

- Tracker Issue: [schema] is an absolute or relative SCD. Why bother allowing absolute?

Both Abosolute and relative SCDs are allowed. No change needed.

- Tracker Issue: Glossary as the place for centralized definitions, or should they be repeated there, but also introduced at point of first use, or should we put the definitions only at the places where they are discussed, and xref from the glossary?

The glossary will be used for definitions.

- TBD: Issue - semantics of expressions containing relative paths that are inherited via ref to a dfdl:defineFormat. (also section 10.3)

This issue applies to expressions used in any global component not just defineFormats. It was agreed that relative paths would be allowed in global components expect for the dfdl:format annotation on an xs:schema. The relative paths are resolved when the global component is reference, not where it is defined..

- TBD: Issue - XPath term - we are not consistent about using the term XPath, or "expression" when referring to our expression language. I prefer to call it our expression language, and then in the section that defines it state that it is a strict subset of XPath 2.0.

The term 'DFDL expression' will be used.

- TBD: Issue - fn:position is unclear given that we've just said we don't support sequences in the expression language.

Action: clarify semantics of fn:position and fn:count. Relax single sequence restriction.

- TBD: Issue - order of sections. Scoping rules section should come before variables section, which uses these concepts.

Expressions and regular expressions will be moved to the back of the spec as they are 'advanced' features.

- TBD: Issue: Case sensitivity of enum names - did we say whether this is case sensitive or not? I believe it should be case sensitive.

  Already agreed that they are case sensitive

- Issue: dfdl:representation - Strings in binary rep. I see no reason why elements of type xs:string will examine dfdl:representation. They shouldn't' care what it is, they are always "text". I should be able to specify a bunch of inter-mixed binary number and string elements without having to specify dfdl:representation="text' just to avoid an error on the string type elements. I believe xs:string type ignores dfdl:representation (always behaves as if dfdl:representation is 'text').(If we change this then the property precedence section for simpletypes changes slightly as representation="text" is implied if type is string.)
That will make it impossible to introduce a binary representation of text later

What is "a binary representation of text"? Is there a real issue here. This is a primary convenience and clarity issue for me. I do not want to have to change to representation="text" for every string inside a cobol structure, which is ultimately a binary representation object. To me type="string" is enough. I want to put in the file scope level of the schema a representation="binary", and then decorate the elements with the specifics of their types, but I do not expect to have to put representation="text" on anything.

I do not understand what you are trying to achieve by requiring representation="text" for things that are already textual based on the type.

Ageed that dfdl:represetnation 'text' is implied for strings and dfdl:represetnation 'binary' is implied for hexbinary

The rest of the issues below I think we need to discuss on calls.

- textStringPadCharacter textNumberPadCharacter - did we agree that this character must be a "minimum width" character if the char set encoding is variable width? (i.e., the pad char must be 1 byte if the encoding is UTF-8.

  textStringPadCharacter textNumberPadCharacter must be a 1 byte character if the char set encoding is variable width?

- numberInfinityRep numberNanRep - Is this applicable only to xs:double and xs:float? Also, what I've seen requires a distinction of sign. I.e., there are positive and negative infinities often printing

as -inf and +inf.

The description is the way ICU behaves but need clarification. It isn't clear how inf and Nan are represented in the infoset. Need to investigate if XML allows these values. Action raised

- TBD: Issue - \n in regular expressions - clarify relationship of this to entities like NL entity. Also, if I include an entity like WSP* in a regular expression (can I?) does it then match accordingly?
It appears that some of our multi-valued entities like WSP+ create conditional "matching" behavior without having to use regular expressions, e.g., when WSP+ is used as a separator. But can you use entities like WSP+ in a regular expression? It seems you should be able to use regular "single valued" entities in a regular expression, its these multi-valued ones that have tricky semantics.
Added Unicode values to /n, /t,/r. Disallow DFDL general (multi-option) entities in regular expressions

- .
- 14.1 Alignment - TBD: Issue - zero-based thinking here. But all the bits stuff and everything else in DFDL uses 1-based reasoning. Need to revisit to make this sensible for 1 based world.
Added implicit alignment table. TBD zero-based

It was felt that it was more natural to have alignment 0,2,4 etc rather than 1,3,5 etc. MB to rewrite section. Action raised

- finalTerminatorCanBeMissing - spec is not clear. Also is there a finalSeparatorCanBeMissing Changed to finalDocumentTerminatorCanBeMissing and finalDocumentSeparatorCanBeMissing. Not sure where finalDocumentSeparatorCanBeMissing should be specified. Looks odd on 'distinguished root'. These properties operate differently from other properties as they are defined on the 'distinguished root' but affect some lower down element. Effectively they are put in scope by a different mechanism

We discussed if the propoerties should be on the distinguished root and it's sequence but deciced that because these were really global properies that would only be allowed in the 'default' format block and be 'scoped' over the whole schema.

3. Go through Actions

Meeting closed, 14:40

Next call Wednesday 17 February January 2010 13:00 UK

Create Action Items

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **083**

Actions raised at this meeting

No	Action
079	AP:Encoding for binary fields when lenghtkind is pattern
080	AP:Clarify semantics of fn:poisition and fn:count
081	AP: Inf and Nan The description is the way ICU behaves but need clarification. It isn't clear how inf and Nan are represented in the infoset. Need to investigate if XML allows these values
082	MB: Should alignment be 0 or 1 based

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way. 29/07: No Progress

	05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description. 13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms. 20/01: Mike will write up 27/01: further discussion of discriminators 29/01: Alan had emailed both proposals but not enough time to discuss 02/02: Agreed to adopt 'component exists' semantics for discriminators 10/02: 'component exists' proposal updated. comments by next call.
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress 29/01: No progress. The predefined formats do not need to be available when the spec is published. Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numbers are define is very usable. He will document for next call

	03/02: No progress 10/02: No progress
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still invetsigating 103/02: IBM is still invetsigating
079	Encoding for binary fields when lenghtkind is pattern
079	AP:Encoding for binary fields when lenghtkind is pattern
080	AP:Clarify semantics of fn:poisition and fn:count
081	AP: Inf and Nan The description is the way ICU behaves but need clarification. It isn't clear how inf and Nan are represented in the infoset. Need to investigate if XML allows these values
082	MB: Should alignment be 0 or 1 based

Closed actions

No	Action
077	SKK: mapping of COBOL numbers to textNumberFormats. 03/02: Suman documented the problem. Agreed to remove textNumberFormat and textCalendarFormat. 10/02: closed
078	MB: Reword section 2.3.1 incorporating markup order rules. 10/02:closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
0	ICU fractional seconds	039	

069			
070	Write DFDL primer		
071	Write test cases.		
072	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix	039	
073	Rename dfdl:separatorPolicy="required" to "always".	039	Def
078	document UPA checks	039	
079	Semantics of length=0, nil handling and defaults. (A071)	039	
080	Tlog: Allow LengthKind delimited for packed/bcd (A074)	039	
081	Update empty sequence section (A075)	039	
082	semantics of minOccurs= 0 on choice branches (A076)	039	
083	Implement RFC2116		
084	Length|Kind pattern scanability rules		
085	Invalid character substitution		
086	infoset round tripi: Rephrase sentence 'It is possible to define a schema so that when infoset unparsed and the datastream reparsed, the same infoset will be produced'		
087	Clarify use of relative paths in global components.		
088	'DFDL expression'		
089	Ageed that dfdl:represetnation 'text' is implied for strings and dfdl:represetnation 'binary' is implied for hexbinary		
091	textStringPadCharacter textNumberPadCharacter must be a 1 byte character if the char set encoding is variable width?		
0			

9 2	finalDocumentTerminatorCanBeMissing and finalDocumentSeparatorCanBeMissing allowed only in 'default' format		
0 9 3	remove textNumberFormat and textCalendarFormat.		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**03-Feb-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 04-Feb-10
Last Modified by Alan Powell on 04-Feb-10

OGF DFDL Working Group Call , February 02 & 03-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plar details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , February 02 & 03-2010

Attendees
Mike Beckerle (Oco)
Steve Hanson (IBM)
Alan Powell (IBM)
Suman Kalia (IBM)
Peter Lambros (IBM)
Tim Kimber(IBM)

Apologies
Stephanie Fetzer (IBM)
Steve Marting (Progeny)

1. Discriminators

Discussed the 'parent exists' and 'component exists' options for dfdl:discriminator semantics. The WG agreed that DFDL would adopt the 'component exists' semantics where the discriminator indicates that the component

it is on exists and does not say anything about the components parent.

Examples need minor changes and full syntax. Arrays would be the same as optional components.

Discriminators are not allowed on
- Global groups and the top level sequence or choice of a global group.
- Global element decalrations
- The top level group of a complex type.
- Anonymous groups other than when it is the top level of a choice branch.

The discriminator timing property was also discussed. It was agreed that a parser should be able to tell when it is possible to evaluate the expression so the timing property will be removed.
The following will be added instead 'The expression will be evaluated when the referenced elements are known to exist or known not to exist.'

Timing on asserts was also mention but deferred for the time being.

Alan will update the 'component exists' proposal.

1. Action 077 Cobol and numberFormats
Suman highlighted a problem using textNumberFormat for COBOL numbers where a separate textNumberFormat is required for each length. This is because a different numberPattern is required for each one. It was suggested that numberPattern was move from textNumberFormat to the standard properties but it was decided to get rid of textNumberFormat altogether and move all the properties. Tim pointed out that this meant that is was now not possible to just vary the number properties so there may be more definedFormats but this was not felt to be a problem.

For consistency it was agreed that textCalendarFormat would also be removed. EscapeScheme is different so will be retained.

2. Remaining 037 review issues

2. I agree with the existing comment that the RFC2119 key words should be upper case. Agreed that the RFC2119 keywords should be in upper case and wherever possible the spec should be reworded to use the key words.

16.2 scannablility with lengthKind pattern : One use case it to find the end of of a character element by looking for 'binary' bytes of the following element.
Discussed whether dfdl:lengthKind pattern should be allowed for binary elements. Pattern scanning inherently treats the data as characters but for single byte encodings a binary byte can be specified by its character codepoint.   Agreed that dfdk:lengthKind pattern will be allowed for binary elements when the encoding is US-ASCII. (why didn't we say any single byte character set?)
Section 16.2 says that that the children on a complex type must not change the encoding. This will be relaxed when encoding is US-ASCII.

DFDL raw entities will not be allowed in a pattern.

Tracker Issue: illegal character encodings for parsing and unparsing.
What should DFDL do when it finds illegal bytes when converting strings during parsing and unparsing?
Discussed adding a new property to declare a substitution character. Discussed what products do and ICU. Decided to follow ICU.
Conversions to Unicode (during parsing) will substitue characters that cannot be converted with the unicode substitute character (U+FFFD). When converting from Unicode (during unparsing) they will be substituted with the substitute character for that encoding, for example 0x1A (Control-Z) for ASCII.

Tracker Issue : Processing-time Schema Definition Errors
Mike will reword Section 2.3.1

Tim will supply rules for the order that terminators and separators are look for in the data stream

Tracker Issue : "round trip" for infoset. Should we omit the whole point ?
Rephrase sentence 'It is possible to define a schema so that when infoset unparsed and the datastream reparsed, the same infoset will be produced'

3. Go through Actions

Meeting closed, 15:00

Next call Tuesday 10 February January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **079**

Actions raised at this meeting

No	Action
078	MB: Reword section 2.3.1 incorporating markup order rules.

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way. 29/07: No Progress 05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description. 13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms. 20/01: Mike will write up 27/01: further discussion of discriminators 29/01: Alan had emailed both proposals but not enough time to discuss 02/02: Agreed to adopt 'component exists' semantics for discriminators
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat

	annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress 29/01: No progress. The predefined formats do not need to be available when the spec is published. Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numbers are define is very usable. He will document for next call 03/03: No progress
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress 03/02: IBM is still invetsigating
077	SKK: mapping of COBOL numbers to textNumberFormats. 03/02: Suman documented the problem. Agreed to remove textNumberFormat and textCalendarFormat.
078	MB: Reword section 2.3.1 incorporating markup order rules.

Closed actions

No	Action

Work items:

No	Item	target version	stat
0	Improvements on property descriptions		not

ID	Description		
05			
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
069	ICU fractional seconds	039	
070	Write DFDL primer		
071	Write test cases.		
072	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix	039	
073	Rename dfdl:separatorPolicy="required" to "always".	039	Def
078	document UPA checks	039	
079	Semantics of length=0, nil handling and defaults. (A071)	039	
080	Tlog: Allow LengthKind delimited for packed/bcd (A074)	039	
081	Update empty sequence section (A075)	039	
082	semantics of minOccurs= 0 on choice branches (A076)	039	
083	Implement RFC2116		
084	Length|Kind pattern scanability rules		
085	Invalid character substitution		

TeamRoom Plus

IBM TeamRoo

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**29-Jan-10 (Fri)**
Meeting Time	**13:00 - 14:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 01-Feb-10
Last Modified by Alan Powell on 01-Feb-10

OGF DFDL Working Group Call , January-29-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , January-29-2010

Attendees
Mike Beckerle (Oco)
Steve Hanson (IBM)
Alan Powell (IBM)
Suman Kalia (IBM)
Peter Lambros (IBM)
Tim Kimber(IBM)

Apologies
Stephanie Fetzer (IBM)
Steve Marting (Progeny)

1 Review Schedule

The WG acknowledged that we are not going to meet the schedule to be available for public comment by OGF 28. Agreed to
specification as soon as possible.

OGF prereview is confirmed to take about 4 weeks assuming no document updates are required. We are behind schedule to March.

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	– 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec – 08 Jan 2010	WG
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	
Grid recommendation process		1 Jan - 1 April 2011	

2. Go through Actions
Updated below

Action 077: Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numl will document for next call.

3. TLog
TLOG
The individual fields are a mixture of ASCII strings, proprietary packed decimals, and the occasional pure binary data. All fiel Fields of all types can be fixed length or variable length with a maximum. Pure binary data is preceded by a field giving the ac

Packed decimals. Like a packed decimal in the IBM sense. These can carry negative numbers but use a leading xD sign nibb unsigned. Odd number of digits (including sign if present) are padded with xF nibble. This is best illustrated using examples.

```
1234      =>      x12x34
123       =>      xF1x23
-1234     =>      xFDx12x34
-123      =>      xD1x23
```

Proposal

1) The 'variable length with a maximum' will be handled using a post-timing assertion. Note this only applies on parsing. **

2) dfdl:lengthKind 'delimited' is permitted for numbers when dfdl:representation is 'binary' and dfdl:binaryNumberRep is 'pack
know in advance the range of bytes being used, and therefore to choose suitable delimiters.

3) Core DFDL 1.0 will **not** be enhanced to handle the TLOG packed decimal type. A future version of DFDL will provide an e
user-defined types to be handled. In the 1.0 timeframe IBM may implement its own proprietary extension to handle this type.

** While this can result in output from a DFDL unparser that can not be re-parsed, that is a problem general to the use of assertions, and a future version of
enhancements to the assertion annotation.

4 Action 071 Semantics of length=0, nil handling and defaults.
Changed unparsing behaviour - we must honour the property - the existing behaviour of always writing the initiator means we
writing empty content and enum is 'suppress'. When reading, assume that section *15.13 has been updated to include compl*

No change to enums.

missingValueInitiatorPolicy	Enum
	Valid values 'required', 'prohibited'
	Specifies whether to expect an initiator when an element is missing. Ignored unless dfdl:initiator is specified and is not "" (em
	'required' - Indicates that the dfdl:initiator followed by empty content is the required syntax to indicate that the element is mis
	'prohibited' - Indicates that empty content is the required syntax to indicate that the element is missing. The presence of an in
	Use of 'prohibited' implies an ordered sequence. If used on an initiated element of an unordered group it is a schema definitio
	If the element is required, defaulting occurs as defined above.
	This property also applies on unparsing, when the data to be written (after nil value and default value processing) is empty co
	Annotation: dfdl:element

Unparsing. *The branch of a choice output when a complex element is required but missing from the infoset is the first branch
a processing error*.

5. Go through Actions

6. Discriminators
Not discussed)

7. Draft 037 review issues

not discussed

- Case of enumerations. We should follow the XSDL convention which is that enumerations are case sensitive

- dfdl:lengthKind='Pattern scannability: A complex element with lengthKind=Pattern will use its dfdl:encoding property as the children irrespective of the child's encoding property.

Go through unanswered issues in Mike's comments document

Meeting closed, 14:10

Next call Tuesday 02 February January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **078**

Actions raised at this meeting

No	Action

077	SKK: mapping of COBOL numbers to textNumberFormats.

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way. 29/07: No Progress 05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description. 13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms. 20/01: Mike will write up 27/01: further discussion of discriminators 29/01: Alan had emailed bot proposals but not enough time to discuss
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value

	is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress 29/01: No progress. The predefined formats do not need to be available when the spec is published. Suman said that he had been mapping COBOL structures to DFDL and it didn't look as though the way text numbers are define is very usable. He will document for next call
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update 27/01: no progress 29/01: no progress
077	SKK: mapping of COBOL numbers to textNumberFormats.

Closed actions

No	Action
064	MB/SH Request WG presentation at OGF 28 25/11: Session requested 04/12: no update 09/12: no update 16/12: SH has changed request to a general session rather tha WG in the hope of attracting more people. 23/12: no update 06/01: not heard anything yet 13/01: no update 20/01: no update 27/01: Session confirmed Closed
071	Semantics of length=0, nil handling and defaults. 23/12:SH no update 06/01: SH has started 13/01: SH proposal review. Minor updates to be made 20/01: Reviewed updated proposal. Need to agree on unparsing empty choices. 27/01: Steve H had sent update but not discussed due to lack of time 29/01: See minutes. Update 15.3 for complex. document missingValueInitiatorPolicy. Unparsing. *The branch of a choice output when a complex element is required but missing from the infoset is the first branch of the choice that does not result in a processing error*.

	Closed
074	SH: Proposal for parsing TLog 27/01: Proposal discussed and agreed to allow delimited for binary packed/bcd fields 29/01: See minutes. Confirmed delimited for packed/bcd Closed
075	SH: rewrite empty sequences section 27/01: Steve provide written section 29/01: scetion rewritten Closed
076	SH semantics of minOccurs= 0 on choice branches 29/01: Steve confirmed that XSDL allows minOccurs=0 for branches of a choice which means that the empty sequence in a valid result. WG decided that DFDL will not allow minOccurs =0 on branches of a choice. Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
069	ICU fractional seconds	039	
070	Write DFDL primer		
071	Write test cases.		
072	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix	039	
073	Rename dfdl:separatorPolicy="required" to "always".	039	Def
078	document UPA checks	039	
079	Semantics of length=0, nil handling and defaults. (A071)	039	
0	Tlog: Allow LengthKind delimited for packed/bcd (A074)	039	

8 0			
0 8 1	Update empty sequence section (A075)	039	
0 8 2	semantics of minOccurs= 0 on choice branches (A076)	039	

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**27-Jan-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 28-Jan-10
Last Modified by Alan Powell on 28-Jan-10

OGF DFDL Working Group Call , January-26 & 27-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you
next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , January-26 & 27-2010

Attendees
Mike Beckerle (Oco) (27)
Steve Hanson (IBM) (26 & 27)
Alan Powell (IBM)   (26 & 27)
Steve Marting (Progeny) (27)
Suman Kalia (IBM) (27)
Peter Lambros (IBM) (26)
Tim Kimber(IBM)   (26 & 27)

Apologies
Stephanie Fetzer (IBM)

1. Discriminators
There was a long discussion of Alan's document which merged the 'Resolving Uncertainty' and 'Discriminators' documents.
It proposes that discriminators 'confirm the existence of their parent sequence' and did not allow discriminators on simple ele
confirm the existence of that element based on its contents or other data. Mike pointed out that this did not allow the uses cas
previous field in the data stream (eg a flag) indicating the existence of a field without wrapping that field in a extra sequence .

The uses cases are:
- A sequence is 'known to exist' by testing one or more of its children
- A sequence is 'known to exist' by testing some previous data
- A simple element is 'known to exist' by testing some previous data
- A simple element is 'known to exist' by testing its content

An alternative proposal was discussed where a discriminator where allowed on particles and   'confirmed the existence of the component'

Alan will write up both proposals.

Also mentioned was the meaning of minOccurs= 0 on a choice branch. Steve will ask Sandy Gao

2. Unparsing lengthKind =Pattern

The table below was modified to make the text unparing length the minLength/outputMinLength.
Mike suggested that lengtKind pattern should be limited to text fields but need to understand Tlog requirements.

	Representation		
	text	*binary*	
String	minLength	Not applicable	
Float	outputMinLength	32	
Double	outputMinLength	64	
Decimal/Integer	outputMinLength	Minimum number of bytes to represent significant di... sign	
Long, UnsignedLong	outputMinLength	packed/bcd : as decimal	binary: 64
Int, Unsignedint	outputMinLength	packed/bcd : as decimal	binary: 32
Short, Unsignedshort	outputMinLength	packed/bcd : as decimal	binary: 16
Byte, Unsignedbyte	outputMinLength	packed/bcd : as decimal	binary: 8
DateTime	outputMinLength	packed/bcd : as decimal	binarySeconds: 32, binaryMilliseconds:64
Date	outputMinLength	packed/bcd : as decimal	binarySeconds: 32, binaryMilliseconds:64
Time	outputMinLength	packed/bcd : as decimal	binarySeconds: 32, binaryMilliseconds:64
Boolean	outputMinLength	32	
HexBinary	Not applicable	Length of infoset value	

3. TLog
Reviewed Steve's proposal

Proposal
1) The 'variable length with a maximum' can not be handled using a post-timing assertion, because assertions only apply on Given that IBM MRM allows you to model this case, I think we should allow dfdl:length & dfdl:lengthUnits to be specified whe dfdl:lengthKind is 'delimited' or 'pattern'. Extraction during parse is still by scanning. If the physical representation exceeds tha after extraction, it is a processing error. Similarly when unparsing, if the physical representation exceeds that length prior to c processing error.

2) dfdl:binaryNumberRep is extended with another value 'tlog'. Associated property dfdl:binaryDecimalVirtualPoint is applicab Associated property dfdl:binaryPackedSignCodes is not applicable, there being only xD to indicate negative numbers, and no nibbles for unsigned or zero.

3) dfdl:lengthKind 'delimited' is permitted for numbers when dfdl:representation is 'binary' and dfdl:binaryNumberRep is 'pack or 'tlog' because it is possible to know in advance the range of bytes being used, and therefore to choose suitable delimiters.

1) can be achieved on parsing with an assert but cannot be enforced on unparsing
2) Discussed the desirability of properties/enumerations specific particular formats versus providing more extensibility. Decid were needed for DFDL v1
3) Steve H felt that this was needed. It was agreed to add to spec (I need to confirm this)

4 Action 071 Semantics of length=0, nil handling and defaults.
Not discussed

5. Go through Actions

6. Draft 037 review issues

- Case of enumerations. We should follow the XSDL convention which is that enumerations are case sensitive

- dfdl:lengthKind='Pattern scannability: A complex element with lengthKind=Pattern will use its dfdl:encoding property as the when scanning its children irrespective of the child's encoding property.

Go through unanswered issues in Mike's comments document

7 Review Schedule
OGF prereview is confirmed to take about 4 weeks assuming no document updates are required. We are behind schedule to for public review by March. Draft 038 will be available at the end of this week.

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	– 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec – 08 Jan 2010	WG

	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	
Grid recommendation process		1 Jan - 1 April 2011	

Meeting closed, 15:10

Next call Friday 29 January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

Next action: **077**

Actions raised at this meeting

No	Action
076	SH semantics of minOccurs= 0 on choice branches

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way. 29/07: No Progress 05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update

	and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description. 13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms. 20/01: Mike will write up 27/01: further discussion of disciminators
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progress 20/01: no progress 27/01: no progress
064	MB/SH Request WG presentation at OGF 28 25/11: Session requested 04/12: no update 09/12: no update 16/12: SH has changed request to a general session rather tha WG in the hope of attracting more people. 23/12: no update 06/01: not heard anything yet 13/01: no update 20/01: no update 27/01: Session confirmed
066	Investigate format for defining test cases

	25/11:IBM to see if it is possible to publish its test case format.
	04/12: no update
	09/12: no update
	16/12: reminded dent to project manager
	23/12: SH will send another reminder.
	06/01: Another reminder will be sent
	13/01: no update
	20/01: no update
	27/01: no progress
071	Semantics of length=0, nil handling and defaults.
	23/12:SH no update
	06/01: SH has started
	13/01: SH proposal review. Minor updates to be made
	20/01: Reviewed updated proposal. Need to agree on unparsing empty choices.
	27/01: Steve H had sent update but not discussed due to lack of time
074	SH: Proposal for parsing TLog
	27/01: Proposal discussed and agreed to allow delimited for binary packed/bcd fields
075	SH: rewrite empty sequences section
	27/01: Steve provide written section
076	SH semantics of minOccurs= 0 on choice branches

Closed actions

No	Action

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
042	Mapping of the DFDL infoset to XDM	none	not
069	ICU fractional seconds	039	
070	Write DFDL primer		
071	Write test cases.		
0	it is a processing error if the number of occurrences in the data does not match the value of the	039	

7 2	expression or prefix		
0 7 3	Rename dfdl:separatorPolicy="required" to "always".	039	Def
0 7 8	document UPA checks	039	

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**20-Jan-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 21-Jan-10
Last Modified by Alan Powell on 21-Jan-10

OGF DFDL Working Group Call , January-20-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , January-20-2010

Attendees
Mike Beckerle (Oco)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Suman Kalia (IBM)
Peter Lambros (IBM)

Apologies
Stephanie Fetzer (IBM)
Tim Kimber(IBM)

1. **045 - Disciminators**
 Mike has started to write proposal.

2 **Unparsing lengthKind = 'pattern'**

The Specification currently says that on unparsing dfdl:lengthKind='pattern' behaves like dfdl:lenghtKind='implict' but since we limited 'implict' to certain logical/representation combinations this is no longer correct.

Proposals
1. When unparsing complex elements with dfdl:lengthKind='pattern', the length is the length of the children (same as 'implicit')

When unparsing simple elements with dfdl:lengthKind='pattern' the length is the implicit length for those that have one and length of the data supplied in the infoset converted to the representation with no padding/filling for the rest.

For string/text and hexbinary/binary that is reasonable
For number/text/standard and zoned it is governed by the numberPattern
For number/binary/ binary use 'implicit' lengths
For number/binary/packed or bcd use minimum number of bytes
For Calendar/text it is governed by the calendarpattern
For Calendar/binary/packed or bcd use minimum number of bytes
For Calendar/binary/binarymilliseconds or binaryseconds use implicit lengths
For Boolean/text use the length of the true/false rep.
For Boolean/binary use implicit length (32)

2. A new property dfdl:patternOutputLengthKind

3. We could limit dfdl:lenghtKind 'pattern' to complex elements so that there is always a lengthKind for child simple elements but this would introduce 'unnecessary' complex items in the infoset when there is only one child.

The WG decided that option 1 was preferable but it needs to be written in terms of a set of 'base' unparsing rules which are augmented depending on the dfdl:lengthKind.

There was some discussion about whether the pattern should be reapplied to the data stream produced to ensure that it could be parsed by the same schema. Perhaps this should be part of validation. There was no agreement.

MB suggested that dfdl:lengthKind 'pattern' should only be allowed for elements with representation 'text' and complex elements with children all 'text'. Steve H felt that it was needed for complex elements containing binary data.

Alan will write up 'base' unparsing length rules.

3. TLOG
Steve H took us through his further investigation of TLog format

Finally had some time to look at the TLOG stuff again (the format emitted by IBM 4680 & 4690 POS controllers). This time I've looked at the MRM code, & spoken with domain expert David Bennett, to establish the full behaviour. Unfortunately there are some issues which mean DFDL 1.0 is not capable of handling it.

1. The individual fields are a mixture of ASCII strings, proprietary packed strings (rare), proprietary packed

decimals, binary integers (rare). All fields are delimited by a separator.

2. The fields are all defined with a length in bytes, but most of the string and decimal fields are actually variable length. If the data exceeds the length in bytes when parsing or unparsing the MRM throws an error. However, talking to David B, the length is really intended to be used if there was ever a fixed length equivalent format, so is really for validation only. To parse the current TLOG formats it is sufficient to use the delimiter.
Note to WG: Validation using the specified length only works for strings . Should we allow dfdl:length to be specified when dfdl:lengthKind=delimited or pattern , and use it as an extra constraint when parsing/unparsing?

3. Scanning for the separator, or maybe use of a data pattern, is needed to extract the variable length data, including packed decimal data (which we would consider a <u>binary</u> type).
Note to WG: Should we allow binary scanning when the users says it is safe to do so ?

4. Packed strings. This is a packed data type used when the range of possible chars in the string is limited to 0-9, A-F. You can view this as a BCD that can also carry A-F. MRM does not try to turn this into an integer, instead it treats it as character data. That is x12x34 would result in string '1234'. Odd numbers of digits are padded with a x0 nibble. On unparsing the MRM throws an error if a character other than 0-9, A-F is encountered. In practice, these packed strings are rare and invariably only carry 0-9 anyway. I'm still trying to establish why this data type is needed, as it could be treated as a logical integer and modelled as BCD.

5. Packed decimals. Like a packed decimal in the IBM sense. These can carry negative numbers but use a leading xD sign nibble. No sign nibble if positive. Odd number of digits (including sign if present) are padded with xF nibble. This is best illustrated using examples.
```
1234       =>       x12x34
123        =>       xF1x23
-1234      =>       xFDx12x34
-123       =>       xD1x23
```
Note to WG: Should we support this data type natively ?

6. Some of the packed decimals are interpreted as an array of bit flags. I assume that DFDL would model these using dfdl:hidden.

Steve H is going to make a proposal about how DFDL can parse TLog

4, Empty Sequences
Is the current section correct

1.1   **Empty Sequences**

A sequence having no children is syntactically legal in DFDL; however, a sequence having no children must have content length zero. It can still have non-zero length Prefix and Suffix regions, but the SequenceContent region in between must be of length zero. It is a schema definition error if the SequenceContent region of an empty sequence is not length zero.

Steve H will re-word.

5. Mike/Steve review issues
Mike B had to leave but a number a issues were discussed

- Case of enumerations. We should follow the XSDL convention which is that enumerations are case sensitive

- dfdl:lengthKind='Pattern: dfdl entities will be allowed in the lengthPattern regular expression except for generic entities such as WSP and NL

- dfdl:lengthKind='Pattern scannability: A complex element with lengthKind=Pattern will use its dfdl:encoding property as the encoding when scanning its children irrespective of the child's encoding property.

Mike to comment on remaining review issues.

6. Go through remaining actions
Updated below

Action 071 Semantics of length=0, nil handling and defaults.
1) Steve H to propose new name for dfdl :defaultValueInitiatorPolicy

Looking more at this, we should be consistent with the similar property for missing separators. So I've changed the enums too - now 'require' and 'suppress', Also changed unparsing behaviour - we must honour the property - the existing behaviour of always writing the initiator means we can not successfully re-parse if writing empty content and enum is 'suppress'. When reading, assume that section 15.13 has been updated to include complex as well as simple elements.

missingValueInitiatorPolicy	Enum
	Valid values 'require', 'suppress'
	Specifies whether to expect an initiator when an element is missing. Ignored unless dfdl:initiator is specified and is not "" (empty string).
	'require' - Indicates that the dfdl:initiator followed by empty content is the required syntax to indicate that the element is missing.
	'suppress' - Indicates that empty content is the required syntax to indicate that the element is missing. The presence of an initiator implies that real content must follow.
	Use of 'suppress' implies an ordered sequence. If used on an initiated element of an unordered group it is a schema definition error.
	If the element is required, defaulting occurs as defined above.
	This property also applies on unparsing, when the data to be written (after nil value and default value processing) is empty content.
	Annotation: dfdl:element

We should similarly change the enums for nilValueInitiatorPolicy to 'require' and 'suppress'.

2) Not recorded in minutes , but there was a discussion around my bullet on choices   .

Worth noting that the concept of 'required' for the elements of a choice does not apply. Even if minOccurs > 0.

The issue was on unparsing. Which branch of a choice do we output when a complex element is required but missing from the infoset? I think it should be the first branch of the choice that does not result in a processing error.

NOTE: Work Item 071 changes dfdl:separatorPolicy enumeration from require to always. Need to make sure that is consistent with this proposal.

7 Review Schedule
OGF prereview is confirmed to take about 4 weeks assuming no document updates are required. We are behind schedule to be available for public review by March. Draft 038 will be available at the end of this week.

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	– 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec – 08 Jan 2010	WG
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	
Grid recommendation process		1 Jan - 1 April 2011	

Meeting closed, 14:30

Next call Tuesday 26 January 2010 13:00 UK and Wednesday 27th January 2010 13:00 UK.

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

Next action: **076**

Actions raised at this meeting

No	Action
074	SH: Proposal for parsing TLog
075	SH: rewrite empty sequences section

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing
	27/05: Psuedo code has been circulated. Review for next call
	03/06: Comments received and will be incorporated
	09/06: Progress but not discussed
	17/06: Discussed briefly
	24/06: No Progress
	01/07: No Progress
	15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way.
	29/07: No Progress
	05/08: No Progress. Will document behaviour as a set of rules.
	12/08: No Progress
	...
	16/09: no progress
	30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update

	and reissue
	07/10: Updated proposal was discussed.Comments will be incorporated into the next version.
	14/10: Alan to update proposal to include array scenario where minOccurs > 0
	21/10: Updated proposal reviewed
	28/10: Updated proposal reviewed see minutes
	04/11: Discussed semantics of disciminators on arrays. MB to produce examples
	11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples.
	18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules.
	25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc.
	04/11: Further discussion about arrays.
	09/12: Reviewed proposed discriminator semantic.
	16/12: Reviewed discriminator examples and WTX semantic.
	23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call
	06/01:B Connolly not available. SF to provide more complete description.
	13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms.
	20/01: Mike will write up
049	20/05 AP Built-in specification description and schemas
	03/06: not discussed
	24/06: No Progress
	24/06: No Progress (hope to get these from test cases)
	15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide.
	...
	14/10: no progress
	21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web.
	28/10: no progress
	04/11: no progress
	11/11: no update
	18/11: no update
	25/11: Agreed to try to produce for CSV and fixed formats
	04/12: no update
	09/12: no update
	16/12: no update
	23/12: no update
	06/01: no progress. If there is no resource to complete this action it can be deferred
	13/01:no progess
	20/01: no progess
064	MB/SH Request WG presentation at OGF 28
	25/11: Session requested
	04/12: no update
	09/12: no update
	16/12: SH has changed request to a general session rather tha WG in the hope of attracting more people.
	23/12: no update
	06/01: not heard anything yet
	13/01: no update
	20/01: no update
066	Investigate format for defining test cases
	25/11:IBM to see if it is possible to publish its test case format.
	04/12: no update
	09/12: no update

	16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent 13/01: no update 20/01: no update
071	Semantics of length=0, nil handling and defaults. 23/12:SH no update 06/01: SH has started 13/01: SH proposal review. Minor updates to be made 20/01: Reviewed updated proposal. Need to agree on unparsing empty choices.
074	SH: Proposal for parsing TLog
075	SH: rewrite empty sequences section

Closed actions

No	Action
056	MB Resolve lengthUnits=bits including fillbytes 12/08: No Progress ... 28/10: no progress 04/11: MB to look at lengthUnits = bits 11/11: no update 18/11: no update 25/11: no update 04/12: no update. ALan will set up a separate call to progress this action. 09/12: no update. ALan will set up a separate call to progress this action. 16/12: MB, SH and AP had a separate call. MB to distribute proposal 23/12: Discussed proposal. MB will updated 06/01: V4 discussed and approved 13/01: Mike updated proposal. Closed
073	SH: Control of overpunching zoned positive sign 13/01: no update 20/01: Proposal agreed. Closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
011	How speculative parsing works (combining choice and variable-occurence - currently these are separate) (from action 045)		awa
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
038	Improve length section including bit handling	038	son
0	Mapping of the DFDL infoset to XDM	none	not

042			
069	ICU fractional seconds	038	
070	Write DFDL primer		
071	Write test cases.		
072	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix	038	
073	Rename dfdl:separatorPolicy="required" to "always".	038	
074	- Last 'postFix' separator is not optional - Terminators are mandatory. - dfdl:documentFinalTerminatorCanBeMissing - dfdl:documentFinalSeparatorCanBeMissing (Action (70))	038	
075	Remove occursCountKind="useAvailableSpace".	038	
076	dfdl:documentRoot, will be defined that can only be on global elements. The DFDL spec does not have to define the format of parameters to the DFDL processor but will indicate that it must be possible to adresss any element. Agreed that ANY element within the schema cane be the starting point for parsing or unparsing. dfdl:documentRoot no longer required	038	
077	*'delimited' means the item is delimited by the item's terminator (if specified) or an enclosing construct's separator or end of the enclosing construct designated by its known length or its terminator.* The definition of EndOfParent also needs improving.	038	
078	document UPA checks	038	
079	Restrictions on use of 'special' entities in regular expressions	038	
080	LengthUnit=bits (A056)	038	
081	Case of enumerations. We should follow the XSDL convention which is that enumerations are case sensitive	038	
082	dfdl:lengthKind='Pattern: dfdl entities will be allowed in the lengthPattern regular expression except for generic entities such as WSP and NL	038	
083	dfdl:lengthKind='Pattern scannability: A complex element with lengthKind=Pattern will use its dfdl:encoding property as the encoding when scanning its children irrespective of the child's encoding property.	038	
0	Control of overpunching zoned positive sign	038	

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**13-Jan-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 12-Jan-10
Last Modified by Alan Powell on 14-Jan-10

OGF DFDL Working Group Call , January-13-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

1. **045 - Disciminators**
Discussion with Stephanie

2. **Zero length elements**
Review Steve H email

3. **Difference between dfdl :lenghtKind = Delimited and endOfParent**

'delimited' means the item is delimited by the item's terminator (if specified) or an enclosing construct's separator or the end of the enclos construct designated by its known length or its terminator.
the only difference with dfdl:lentghKind='endOfParent' is that the latter includes the 'end of the data stream' and applies to bir fields.
We should either
- Add 'end of data stream' to delimited and remove 'endOfParent'
- Make 'endOfParent' be specifically for only 'end of data stream'

4. **Go through remaining actions**

5 **Draft 037 review**

From comments:

a **DFDL Subset of XML Schema**   (TBD: need means for an implementation to indicate it is using non-standard extensions?

b. Question whether infoset MUST be in schema order. Request for 'bitstream order'

c. Dealing with 'Grammar ambiguity' errors

6 **Review Schedule**

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	– 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec – 08 Jan 2010	WG
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	
Grid recommendation process		1 Jan - 1 April 2011	

Minutes

Meeting Minutes

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , January-13-2010

Attendees
Mike Beckerle (Oco)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Stephanie Fetzer (IBM)
Suman Kalia (IBM)
Peter Lambros (IBM)
Tim Kimber(IBM)

Apologies

1. **045 - Disciminators**
Stephanie took us through her email Subject: **[DFDL-WG] Bob & Steph's WTX 'Discriminators ' write-up**

 WTX Identifiers are similar to DFDL discriminators

-Discriminators may only be placed on the physical representation of a group. That is why we see them on partition groups and sequence groups but not on choice groups (or unordered groups – covered below).

In partitioned groups we have a subtype of each possible group - so each possible group may have a discriminator.

When WTX expresses choice groups it expresses them as a group containing all of the possible child groups – so at the top level 'choice group' there is no component of the actual group content- so no use for a discriminator. But each choice which may itself be a group may have a discriminator. Choice groups are special in that the choice model construct simply lists the components and only one may occur...at this level a discriminator on one of the choices may not be very useful. Inside of each choice's components a discriminator could be used to indicate the existence of that choice.

-The WTX UI does not allow discriminators on the components of Unordered Groups. This may be due to the fact that the position of the discriminator has significance (all rules at or above the discriminator must evaluate to true). If the group is unordered it would be difficult to enforce. Will need to discuss for DFDL.

-A group may have either zero or one discriminators. No group may have more than one discriminator.

-The discriminator may have two significant parts
o   it's location (mandatory). The discriminator is placed on a component of a group and makes all of the cardinality and rules at that point and above become part of it's concept.
o   it's rule (optional)
A group with a component which has a discriminator should have some 'rule' associated with it. In WTX if there is no explicit rule then the implicit rule is 'PRESENT ($)'. We will need to decide if such implied rules will be allowed in DFDL.

-A group may only have a discriminator on a mandatory component. Once again, this impacts a choice group where by definition all components are optional - which will not have a discriminator .

This has been an issue of debate in WTX. We could have implemented checking on optional elements quite

easily Over the years this has been questioned (as our UI allows them to be placed on optional elements) but once we explained the way the engine worked no customers perceived this as a deficiency. In DFDL we will need to determine if this is needed.

-In WTX we do allow a discriminator to be placed on a mandatory fixed size array (a repeating mandatory component with n:n cardinality). It's component rule can either refer to the entirety of the array (PRESENT($) meaning the whole of the array is present) or can call out a specific rule against one if the iterations. This is not done often in practice.

-In WTX it is common to have multiple levels of discriminators when we are working with nested groups.

We discussed whether DFDL should not allow discriminators on unordered groups or groups with floating elements. Agree that discriminators should be allowed

Also discussed whether timing 'before/after' was required are WTX only has after. Decided to keep timing property.

Suggested should not be allowed on variable length arrays to be consistent with not being allowed on optional elements.

Mike agreed to write up rules in dfdl terms and extent to cover other points of uncertainty besides choices.

2. Zero length elements
Steve H took us through his email subject: **[DFDL] zero length (was Re: Fw: TDS length reference) ** updated **
This proposes that zero length fields should not be a processing error

Proposal:

1. Parsing

Simple elements

1) It is **not** a schema definition error **nor** a processing error if a length is being used to extract data and it is zero. This covers dfdl:lengthKind implicit, explicit, prefixed and endOfParent (when parent length is known). The result is 'empty content'. (Note that for implicit, XSDL allows maxLength/length facet to be 0, so disallowing it for others is not consistent).

2) It is **not** a processing error if scanning for data and the length of the returned bytes is zero. This applies to dfdl:lengthKind delimited, pattern and endOfParent (when parent length is not known). The result is 'empty content'. (This is just stating the obvious).

(The above two rules ensure that it is possible to apply empty content to trigger optional, nil value or default value processing regardless of data type and dfdl:lengthKind).

3) Optional, nil and default processing are applied as per spec.

4) If the element is required, and nil value or default value is not used, and empty string is not in the lexical space of the element's type, then it is a processing error.

The two initiator related properties dfdl:nilValueInitiatorPolicy and dfdl:defaultValueInitiatorPolicy define whether nils and defaults are applied when initiated empty content is found, they don't affect the definition of empty content or what it means for the type.

[Note: If you recall, this discussion was triggered by a customer that was using an expression to calculate the

length of a standard text decimal. He wanted 0 length to mean 0 ended up in the infoset. He can achieve this by making the element required with a default value of 0.]

Complex elements

It is possible to get returned empty content for a complex element for cases 1) and 2) above.

1) If the complex element is optional then it is not added to the infoset.

2) If the complex element does not have an initiator specified & is required then it is added to the infoset.

3) If the element has an initiator specified then dfdl:defaultValueInitiatorPolicy applies
        - required => element is added to infoset only if initiator is present (processing error if no initiator & empty content)
        - prohibited => element is added to infoset only if initiator is not present (initiator implies real content follows so processing error if initiator & empty content)

4) If the complex element is added to the infoset, then the parser processes the child content of the complex type. This may or may not cause a processing error.
<tk>I presume a processing error would be caused by
- any group having an initiator or terminator (same as 5. below)
- any group having a prefix or postfix delimiter
- any group with more than one member having an infix delimiter
- any required element within the complex element having an initiator and dfdl:defaultValueInitiatorPolicy="required"
- any required element within the complex element having a terminator
- any required element which does not have a default value specified, and for which a zero-length representation is illegal
- other error scenarios?
</tk>
<smh>Correct. Basically you are going through the element's content (model group plus children) and attempting to parse. When you extract the data you get back empty content. This may or not cause a processing error. This was agreed on the call as the correct behaviour. In summary, for empty content to be valid for the complex element then it must also be valid for at least one content model</smh>
 If it doesn't then default value processing applies for required child elements. If we don't do this then we will not create default values for all missing required simple elements, and that would be wrong.

5) If the contained sequence or choice has an initiator or terminator then it is a processing error.
<tk>
So it's OK to have a choice among the children of the complex element? If so, the specification should define the rules for picking a branch of the choice. The DFDL processer *could* always pick the first branch, but what if the first branch triggers a processing error and a different branch would not have done?
</tk>
<smh>I think it's the same as with real content. Parser will start against the first branch of the choice and see where it gets. Usual speculative parsing rules apply. If it has not discriminated successfully and a processing error occurs it will cause backtracking and the next branch will be tried. If it finds a valid content model for the empty content we are ok. If it doesn't it's a processing error.</smh>

2. Unparsing

Simple elements

Data in the infoset can result in empty content being added to the bit stream (ie, nothing), with an accompanying 0 value in any length prefix or length expression field, if appropiate to the dfdl:lengthKind.

Complex elements

The absence from the infoset of a required complex element will cause any specified initiator to be output, plus if there are required children then default values will be output for those children. If we don't do this then we will not create default values for nested missing required simple elements, and that would be wrong. This enables creation of a sparse infoset containing just the elements with explicit values, with the rest defaulting regardless of nesting.

3. Choices

Worth noting that the concept of 'required' for the elements of a choice does not apply. Even if minOccurs > 0.

4. Outstanding Issues

Is it ok to reuse dfdl:defaultValueInitiatorPolicy for complex elements? Should it be renamed? Should we add a separate property for complex elements?

Steve H to propose new name for dfdl:defaultValueInitiatorPolicy

3. Difference between dfdl :lenghtKind= Delimited and endOfParent

'delimited' means the item is delimited by the item's terminator (if specified) or an enclosing construct's separator or the end of the enclosing construct designated by its known length or its terminator.
the only difference with dfdl:lentghKind='endOfParent' is that the latter includes the 'end of the data stream' and applies to binary fields.
We should either
- Add 'end of data stream' to delimited and remove 'endOfParent'
- Make 'endOfParent' be specifically for only 'end of data stream'

Short discussion. Alan agreed to try to write up description of endOfParent for review

4. Go through remaining actions
No enough time

5 Draft 037 review

From comments:

a DFDL Subset of XML Schema (TBD: need means for an implementation to indicate it is using non-standard extensions?)
Believe that this was to allow users to indicate they are using unsupported schema components. Agreed to defer fron DFDL v1

b. Question whether infoset MUST be in schema order. Request for 'bitstream order'
Short discussion. Main reason for schema order is allow the infoset to be validated against a schema. Agree to leave as schema order

c. Dealing with 'Grammar ambiguity' errors
Not discussed

6 Review Schedule

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	- 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec - 08 Jan 2010	WG
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	
Grid recommendation process		1 Jan - 1 April 2011	

Meeting closed, 15:20

Next call 20 January 2010 13:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **074**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way. 29/07: No Progress 05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update

		and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description. 13/01: Stephaine took us through a description of WTX identifiers. Mike agreed to write up in DFDL terms.
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deferred 13/01:no progess	
064	MB/SH Request WG presentation at OGF 28 25/11: Session requested 04/12: no update 09/12: no update 16/12: SH has changed request to a general session rather tha WG in the hope of attracting more people. 23/12: no update 06/01: not heard anything yet 13/01: no update	
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent	

	13/01: no update
068	Should the roots of messages be designated.? 09/12: Yes. New dfdl:documentRoot property Closed 16/12: reopened and decided to drop property subject to agreement from SKK and SF 23/12: SKK review decision to drop dfdl:documentRoot 13/01: closed
071	Semantics of length=0, nil handling and defaults. 23/12:SH no update 06/01: SH has started 13/01: SH proposal review. Minor updates to be made
073	SH: Control of overpunching zoned positive sign 13/01: no update

Closed actions

No	Action
056	MB Resolve lengthUnits=bits including fillbytes 12/08: No Progress ... 28/10: no progress 04/11: MB to look at lengthUnits = bits 11/11: no update 18/11: no update 25/11: no update 04/12: no update. ALan will set up a separate call to progress this action. 09/12: no update. ALan will set up a separate call to progress this action. 16/12: MB, SH and AP had a separate call. MB to distribute proposal 23/12: Discussed proposal. MB will updated 06/01: V4 discussed and approved 13/01: Mike updated proposal. Closed
068	Should the roots of messages be designated.? 09/12: Yes. New dfdl:documentRoot property Closed 16/12: reopened and decided to drop property subject to agreement from SKK and SF 23/12: SKK review decision to drop dfdl:documentRoot 13/01: closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
011	How speculative parsing works (combining choice and variable-occurence - currently these are separate) (from action 045)		awa
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
0	Update dfdl schema with change properties	ongoing	

3 6			
0 3 8	Improve length section including bit handling		son
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 6 9	ICU fractional seconds		
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 7 2	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix		
0 7 3	Rename dfdl:separatorPolicy="required" to "always".		
0 7 4	- Last 'postFix' separator is not optional - Terminators are mandatory. - dfdl:documentFinalTerminatorCanBeMissing - dfdl:documentFinalSeparatorCanBeMissing (Action (70))		
0 7 5	Remove occursCountKind="useAvailableSpace".		
0 7 6	dfdl:documentRoot, will be defined that can only be on global elements. The DFDL spec does not have to define the format of parameters to the DFDL processor but will indicate that it must be possible to adresss any element. Agreed that ANY element within the schema cane be the starting point for parsing or unparsing. dfdl:documentRoot no longer required		
0 7 7	*'delimited' means the item is delimited by the item's terminator (if specified) or an enclosing construct's separator or end of the enclosing construct designated by its known length or its terminator.* The definition of EndOfParent also needs improving.		
0 7 8	document UPA checks		
0 7 9	Restrictions on use of 'special' entities in regular expressions		
0 8 0	LengthUnit=bits (A056)		

TeamRoom Plus

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	**DFDL 1.0**
Meeting Date	**06-Jan-10 (Wed)**
Meeting Time	**13:00 - 15:00**
Meeting Editors	Alan Powell

Created by Alan Powell on 07-Jan-10
Last Modified by Alan Powell on 12-Jan-10

OGF DFDL Working Group Call , January-06-2010

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

1. **045 - Disciminators**
Continue discussion of WTX semantics.

2. **Does the parser/serializer have to start on a global element ?**
Agree removal of dfdl:documentRoot property
Need Stephanie and Suman to agree.

3. **LengthUnits = 'bits'**
Review Mike's proposal

4. **Semantics of length =0**
Need to go through each of the lengthKinds to understand the interaction of length=0, nil handling and defaults.
ACTION

5. **Byte Order Mark and Unicode signature**
Tim has investigated ICU behavior and it looks OK

Before we conclude that the spec is OK as it stands, we should consider whether it is correct to treat a BOM as a character. T Unicode standard makes a clear distinction between characters and BOMs:
http://www.unicode.org/versions/Unicode5.0.0/ch02.pdf section 2.13 (don't skip the first couple of paragraphs)
http://unicode.org/faq/utf_bom.html#bom1

6. **Go through remaining actions**

7. **Draft 037 review**
Please submit comments by Monday 21st Dec

Agree terminology

Agreed that:
-A schema has **components** which have dfdl **annotations**
.- Dfdl **annotations** have **properties** (same as XSDL) but need to distinguish between format/representation properties and o
properties such as name, ref, selector.

Not discussed.
-The ~~data/data stream /message~~ has ~~elements/fields~~
-The infoset has **items**

8. **Plan to finish DFDL v 1 Schedule**

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	– 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec – 08 Jan 2010	WG
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations. What went well, or what would you do differently next time? Document those so others can take advantage of your learning .

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , January-06-2010

Attendees
Mike Beckerle (Oco)
Steve Hanson (IBM)
Alan Powell (IBM)
Steve Marting (Progeny)
Stephanie Fetzer (IBM)
Suman Kalia (IBM)

Apologies
Peter Lambros (IBM)

1. **045 - Disciminators**
Bob Connolly was not available to discuss WTX.   Stephanie discussed some of the WTX behaviour and agreed to provide more for the next call. We need a definition of the restricted positions where discriminators can be used and the semantics

2. **LengthUnits = 'bits'**
Reviewed v4 of Mike's proposal and suggested editorial changes. If possible would like some words to describe how 'little-endian' is dealt with. Approved.

3. **Selector support in V 1**
Implementation of selectors is proving more difficult than expected so IBM requested that selectors were deferred. In order to give reviewers and indiction of how the function might look in a future release and the opportunity to say if it is a higher priority the description will be move to an appendix named 'Features under consideration for a future version'

4. **Go through remaining actions**
updated below

5 **Draft 037 review**
Mike has reviewed and raised many comments

-Significant issues:
The spec is not consistent about how different types of errors are dealt with.
Schema definition errors can occur at design and runtime and stop processing
'Grammar ambiguity' errors may not be statically detectable so are processing errors.

- Use of entities in regular expressions. 'Special' entities such as WSP+ will not be allowed.

5A. AOB

Steve H raised an issue with zoned number support. Customers have requested the ability control whether the sign is overpunched for positive numbers.
Action raised

6 Review Schedule

Activity		Schedule	Who
Complete Action items		- 18 Dec 2009	WG
Complete Spec	Write up work items	– 23 Dec 2009	AP
	Restructure and complete specification	- 23 Dec 2009	AP
	Issue Draft 038	23 Dec 2009	
WG review	WG review	7 Dec – 08 Jan 2010	WG
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 039	15 Jan 2010	
	Incorporate review comments	4 Jan - 29 Jan 2010	AP +
	Issue Draft 040	29 Jan 2010	
Initial OGF Editor Review	Initial Editor review	1 Feb - 1 Mar 2010	OGF
	Initial GFSG review	1 Feb - 1 Mar 2010	
	Issue Draft 041	1 Mar 2010	
OGF Public Comment period (60 days)		1 Mar - 30 Apr 2010	OGF
OGF 28 Munich		15-19 March 2010	
Incorporate comments	Incorporate comments	28 May 2010	
	Issue Draft 042	28 May 2010	
Final OGF Editor Review	Final Editor review	June 2010	OGF
	final GFSG review	June 2010	
	Issue Final specification	30 June 2010	
Publish proposed recommendation		1 July 2010	
Grid recommendation process		1 Jan - 1 April 2011	

Meeting closed, 15:00

Next call 13 January 2010 13:00 UK

Record the to-do's and individuals assigned by entering the appropriate information in the form below . Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab .

Action Items and Other Meeting Documents

View: ResultDocs

Next action: **074**

Actions raised at this meeting

No	Action
073	SH: Control of overpunching zoned positive sign

Current Actions :

No	Action
045	20/05 AP: Speculative Parsing 27/05: Psuedo code has been circulated. Review for next call 03/06: Comments received and will be incorporated 09/06: Progress but not discussed 17/06: Discussed briefly 24/06: No Progress 01/07: No Progress 15/07: No progress. MB not happy with the way the algorithm is documented, need to find a better way.

	29/07: No Progress 05/08: No Progress. Will document behaviour as a set of rules. 12/08: No Progress ... 16/09: no progress 30/09: AP distributed proposal and others commented. Brief discussion AP to incorporate update and reissue 07/10: Updated proposal was discussed.Comments will be incorporated into the next version. 14/10: Alan to update proposal to include array scenario where minOccurs > 0 21/10: Updated proposal reviewed 28/10: Updated proposal reviewed see minutes 04/11: Discussed semantics of disciminators on arrays. MB to produce examples 11/11: Absorbing action 033 into 045. Maybe decorated discrminator kinds are needed after all. MB and SF to continue with examples. 18/11: Went through WTX implementation of example. SF to gather more documentation about WTX discriminator rules. 25/11: Further discussion. Will get more WTX documentation. Need to confirm that no changes need to Resolving Uncertainty doc. 04/11: Further discussion about arrays. 09/12: Reviewed proposed discriminator semantic. 16/12: Reviewed discriminator examples and WTX semantic. 23/12: SF to provide better description of WTX behaviour and invite B Connolley to next call 06/01:B Connolly not available. SF to provide more complete description.
049	20/05 AP Built-in specification description and schemas 03/06: not discussed 24/06: No Progress 24/06: No Progress (hope to get these from test cases) 15/07: No progress. Once available, the examples in the spec should use the dfdl:defineFormat annotations they provide. ... 14/10: no progress 21/10: Discussed the real need for this being in the specification. It seemed that the main value is it define a schema location for downloading 'known' defaults from the web. 28/10: no progress 04/11: no progress 11/11: no update 18/11: no update 25/11: Agreed to try to produce for CSV and fixed formats 04/12: no update 09/12: no update 16/12: no update 23/12: no update 06/01: no progress. If there is no resource to complete this action it can be deffered
056	MB Resolve lengthUnits=bits including fillbytes 12/08: No Progress ... 28/10: no progress 04/11: MB to look at lengthUnits = bits 11/11: no update 18/11: no update 25/11: no update 04/12: no update. ALan will set up a separate call to progress this action. 09/12: no update. ALan will set up a separate call to progress this action. 16/12: MB, SH and AP had a separate call. MB to distribute proposal 23/12: Discussed proposal. MB will updated 06/01: V4 discussed and approved

No	Item	
064	MB/SH Request WG presentation at OGF 28 25/11: Session requested 04/12: no update 09/12: no update 16/12: SH has changed request to a general session rather tha WG in the hope of attracting more people. 23/12: no update 06/01: not heard anything yet	
066	Investigate format for defining test cases 25/11:IBM to see if it is possible to publish its test case format. 04/12: no update 09/12: no update 16/12: reminded dent to project manager 23/12: SH will send another reminder. 06/01: Another reminder will be sent	
068	Should the roots of messages be designated.? 09/12: Yes. New dfdl:documentRoot property Closed 16/12: reopened and decided to drop property subject to agreement from SKK and SF 23/12: SKK review decision to drop dfdl:documentRoot	
071	Semantics of length=0, nil handling and defaults. 23/12:SH no update 06/01: SH has started	
073	SH: Control of overpunching zoned positive sign	

Closed actions

No	Action
072	Byte Order Mark and Unicode signature 23/12: Confirm BOM will be treated and part of the content. 06/01: MB v37 comments suggest additional description. closed

Work items:

No	Item	target version	stat
005	Improvements on property descriptions		not
011	How speculative parsing works (combining choice and variable-occurence - currently these are separate) (from action 045)		awa
012	Reordering the properties discussion: move representation earlier, improve flow of topics		not
036	Update dfdl schema with change properties	ongoing	
03	Improve length section including bit handling		son

8			
0 4 2	Mapping of the DFDL infoset to XDM	none	not
0 6 9	ICU fractional seconds		
0 7 0	Write DFDL primer		
0 7 1	Write test cases.		
0 7 2	it is a processing error if the number of occurrences in the data does not match the value of the expression or prefix		
0 7 3	Rename dfdl:separatorPolicy="required" to "always".		
0 7 4	- Last 'postFix' separator is not optional - Terminators are mandatory. - dfdl:documentFinalTerminatorCanBeMissing - dfdl:documentFinalSeparatorCanBeMissing (Action (70))		
0 7 5	Remove occursCountKind="useAvailableSpace".		
0 7 6	dfdl:documentRoot, will be defined that can only be on global elements. The DFDL spec does not have to define the format of parameters to the DFDL processor but will indicate that it must be possible to adresss any element. Agreed that ANY element within the schema cane be the starting point for parsing or unparsing.		
0 7 7	*'delimited' means the item is delimited by the item's terminator (if specified) or an enclosing construct's separator or end of the enclosing construct designated by its known length or its terminator.* The definition of EndOfParent also needs improving.		
0 7 8	document UPA checks		
0 7 9	Restrictions on use of 'special' entities in regular expressions		