TeamRoom Plus

IBM TeamRo

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

Meeting about Meetings\OGF

Project	DFDL 1.0
Meeting Date	28-Jun-11 (Tues)
Meeting Time	15:00 - 16:00

Created by Steve Hanson on 09-Mar-11
Last Modified by Steve Hanson on 28-Jun-11

OGF DFDL Working Group Call , 28 June 2011

Agenda

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

Open Grid Forum : Data Format Description Language Working Group

OGF DFDL Working Group Call , 28 June 2011

1. Spec issues from IBM
As implementation has progressed, a number of issues have been spotted in the DFDL 1.0 spec
by the IBM implementation team. List so far for resolution.

6.3. Bullet for logical value. State that the string must obey the lexical representation of the type.

6.3.1. Explicit statement of white space rules is missing. Just white space for a DFDL string literal property is an error.
When a property can be a list of string literals, white space between values is collapsed.

7.7. What is the scope of a variable with respect to included/imported xsds?
Can I refer from xsd A to a variable defined in parent xsd B?

9.2. Grammar. Does not allow for a prefix length itself having a prefixed length, so does not match property description
(ie, the change to permit this is draft 045 omitted to update the grammar).

13. Be explicit on the DFDL literals allowed for text properties. See separate e-mail.

13.6. When textNumberRep is zoned, property description should state that base is assumed to be 10.

13.6. State that dfdl:ignoreCase not honoured by textStandardExponentCharacter, textStandardInfinityRep,
textStandardNanRep as ICU does not offer that flexibility.

13.6 Allow multiple characters for textStandardExponentCharacter to handle reps like 1.23x10^4 as ICU allows that.
Note that property name will have to change, eg, to textStandardExponent..

13.6. Change meaning of textNumberCheckPolicy 'lax' to align more closely with ICU:
"If 'lax' and dfdl:textNumberRep is 'standard' then grouping separators are ignored, leading and trailing whitespace
is ignored, leading zeros are ignored."

13.6. Disallow the use of empty string for textStandardExponentCharacter.
State property must be set if the pattern contains an 'E'

13.6. Disallow the use of empty string for textStandardDecimalSeparator,
State property must be set if the pattern contains a '.'

13.6 Disallow the use of empty string for textStandardGroupingSeparator,
State property must be set if the pattern contains a ','

13.6. Allow decimal separator to be a List of DFDL String Literals or a DFDL expression.
This allows modelling of EDIFACT where user can choose decimal separator in ISA header but '.' is also allowed.

13.6. When textNumberPadCharacter is '0' which it commonly is, a value of say '00000' will get trimmed to the empty string, whereas the intent would surely be to trim to '0'. Need a rule for this.

13.6.Should textStandardDecimalSeparator ignored when the logical type is not decimal/float/double?

13.6.1.1. Erroneously mentions sigDigits in reference to the BNF for textNumberPattern.

13.6.1.1. Formatting. Uses terms 'minimum/maximum integer/fraction digits' but does not define them.

13.9. Should state that textBooleanTrue/FalseRep properties are used after trimming when parsing,
and before padding when unparsing. If lengthKind is explicit or implicit and either textPadKind or textTrimKind is none then the properties must have the same length else schema definition error.

13.16. Nil literal character. When representation is binary, the encoding property is not used, which implies that a nil value must only contain DFDL hex literals (%rXX;)

13.16. Clarify nilKind 'literalCharacter' and 'literalValue' and whether nil test is applied before or after trimming.

15. Using a discriminator to resolve each branch of a choice with a very large number of branches has a performance impact as each branch must be evaluated in turn. In cases where the name of the branch element is known or derivable in advance (ie, each discriminator expression is of the form {..\tag eq "elem1"} or similar) then consider adding a property to dfdl:choice that provides an expression that returns a QName, being the name of the branch element.
Note: IBM MRM has something like this.

16.1. Revise entire section.

16.2. State it is a processing error if the stop value is missing when parsing.

2. Spec issues from Mike Beckerle
A number of issues have been spotted by Mike. Some have been resolved by action 139 below, but others remain.
See separate e-mail.

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 , 28 June 2011

Attendees
Steve Hanson (IBM)
Steph Fetzer (IBM)
Mike Beckerle (Deloitte)

Apologies
Suman Kalia (IBM)

1. Spec issues from IBM
6.3. Bullet for logical value. State that the string must obey the lexical representation of the type.
Agreed

6.3.1. Explicit statement of white space rules is missing. Just white space for a DFDL string literal property is an error. When a property can be a list of string literals, white space between values is collapsed.
Decided to be conservative so literal white space only ever used as list token separator.
Use entities if literal white space is needed as part of the property value.

7.7. What is the scope of a variable with respect to included/imported xsds?
Can I refer from xsd A to a variable defined in parent xsd B?
Existing spec statement actually covers this, so if A includes/imports B then B can refer to a variable defined in A. Clarify 7.7 to make it clearer that this is allowed and reference is via QName in the usual way..

9.2. Grammar. Does not allow for a prefix length itself having a prefixed length, so does not match property description (ie, the change to permit this is draft 045 omitted to update the grammar).
Allow a prefix length type to have a prefix length, but not beyond. Types can't be the same.
Action 142 raised for Mike to supply grammar change.

13. Be explicit on the DFDL literals allowed for text properties. See separate e-mail.
Agreed with a couple of modifications that IBM team need to verify. Action 143 raised.

13.6. When textNumberRep is zoned, property description should state that base is assumed to be 10.
Agreed

13.6. State that dfdl:ignoreCase not honoured by textStandardExponentCharacter, textStandardInfinityRep, textStandardNanRep as ICU does not offer that flexibility.
Action 141 raised to investigate use of a list.

Remaining issues to be discussed on next call.

2. Spec issues from Mike Beckerle
A number of issues have been spotted by Mike. Some have been resolved by action 139 below, but others remain. To be discussed on next call.

Meeting closed
16:05 UK

Next call
Tues 5th 15:00 UK (10:00 EST)

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: **144**

Actions raised at this meeting

No	Action
141	Should text number exponent, infinity and Nan rep properties be lists? 28/6: There is certainly a requirement for DFDL to handle multiple reps for these properties. If ignoreCase is not an option due to ICU then should these allow a list?
142	Update grammar for PrefixLengthType 28/6: Needs to reflect that a prefix length type can itself have a prefix length, but not beyond that. Grammar must reflect this.
143	Clarify any restrictions that apply to DFDL properties that can take a DFDL string literal. 28/6. WG made some updates to IBM's proposal, IBM will verify updates.

Current Actions:

No	Action
111	**Daffodil DFDL parser (Bob, Joe)** 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 responded 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 progressing project funding and making Daffodil publically available.
10/11: NCSA internal & sponsor (US National Archive in Washington DC - Electronic Records Administration) reviews passed. NCSA have new resource allocated - Joe Futrelle. Bob has started open source paperwork. ETA end December.
17/11: Joe has started coming up to speed with Daffodil. Bob is waiting for signoff from the university to open source the code.
24/11: not discussed
01/12: Joe is becoming familiar with the code. Waiting for university to give permission to make available.
08/12: Disclosure process in progressing. Bob has permission to make the web page available so hopefully that will be soon.
The US National Archive have an article on their website about DFDL. Unfortunately it is a little out of date.
15/12: Hope to get permission to make the source code available within days. Web demo page will be available when resources have been identified.
22/12: An internal demonstration web page has been made available. Initial testing highlight some problems. Joe investigating.
12/01: Web site has been move to a public URL http://daffodil.ncsa.uiuc.edu/. Joe thinks there are multiple problems that will show up when he can run the test cases.
19/01: Joe has had little time to work on Daffodil. Without significant research into the code it is difficult to know how compliant daffodil is. Discussed if there are ways the WG could help. Suggestions
- Help review the code
- Review the Daffodil test cases for correctness
- Simplify the IBM test cases to be 'unit tests'
Alan to ask for permission to review test cases and possibly simplifying the IBM test cases
26/01: Joe has ot had any time to work on Daffodil. We briefly discuessed if it was possible to use XSLT to transform test cases so they didn't use defaulting but it didn't seem possible and effort would be better targetted at the Daffodil code.
Alan had asked for permission to view test cases and subsequently heard that it has been given.
02/02: Joe will send the Daffodil test cases for Alan to review.
09/02: Alan has received the test cases and will feed back comments. Joe is looking at support for the default format
16/02: Alan has corrected some of the daffodil test cases and identified common problems
23/02: No update
02/03: No progress

	09/03: No progress 30/03: From Joe: "The current status of my work is that, having identified issues with the schema parser, I am now going to attempt to independently assess the processors. That is: I am going to hand-configure the processors in a series of tests so that I can see what they can parse and what they can't. In the process I may be able to refactor them so that they are less closely coupled to the schema parser, setting the stage for a new schema parser implementation." 13/04: Joe has analysed the schema parser part of Daffodil - it is looking for annotations in wrong place, scoping are rules broken, it is using wrong property names. Needs a rewrite. But the parser itself is in better shape. Unfortunately Joe is leaving NCSA in 4-5 weeks time. Net is that any Daffodil on the web will be a partial, non-conformant implementation. Must make it very clear that this is the case. Also unlikely that Joe will be replaced. Joe will see whether there is any chance to work on DFDL in his new job. 04/05: Bob will create some white papers on what to do when implementing a DFDL parser, on the state of the current Daffodil implementation, and on lessons learned. The former is of particular interest as the IBM implementation is throwing up some interesting design points and a couple of difficulties in the area of empty/missing/nil/default processing. 18/05: No update 01/06: No update 08/06: No update 15/06: No change in status. Bob wants to spread the new wider to find contributors to Daffodil, as NCSA does not have the qualified staff available, and no new hires are likely. US NSA have shown an interest in contributing. Steve has invited them to future DFDL WG calls. 28/06: No update
123	**DFDL tutorial (Steve)** 13/10: Draft of first 3 chapters has been written and will be distributed to WG 10/11: Posted to grid forge here (http://forge.gridforum.org/sf/go/doc16106?nav=1), work continuing at IBM to define a standard example-based chapter framework and to author additional chapters. Contributors welcome! 17/11: Steve, Stephanie and Alan had a meeting to discuss the best structure for the tutorial and decide which examples to use throughout. The meeting raised more questions. Further discussions will be held. 24:11: The list of topics to be covered in the remaining lessons has been produced and a lesson template. Alan will write lesson 4 01/12: Alan has started lesson 4 which covers fixed and variable fields and arrays. 08/12: Alan has almost completed lesson 4. Will send out for review. 15/12: First draft of lesson 4 is available for review. Alan to send to Bob and Joe. 22/12: Alan has distributed drafts for tutorials on Basic Structure and Optional/Repeating elements. Please review 12/01: Alan distributed a tutorial for choices and updated the others. Alan and Steve reviewed them and updated versions will be sent soon. Should start on the 'representation' tutorials soon. 19/01: The tutorials for basic structure, optional/arrays and choices have be updated. Please review. The tutorial for text elements should be available soon. 26/01: No comments received about 3 tuorials distributed last week. Alan is still working on Text representation. 02/02: Steve has sent comments on three tutorials. Alan to send updated versions by the end of the week. Alan has also distributed the first part of the tutorial on text representation and would like feedback. 09/02: Steve had reviewed tutorials 3,4,5 and updated versions have been distributed. Joe reviewed lesson on text elements. Main points. Using 'represented as text' is confusing. Examples are too cluttered. Suggest simple targeted examples but still build up to final complete schema 23/02: New versions distributed and Steve has commented. 02/03: Alan has published the final versions of tutorials 4,5,6 and is working on text respresentations. There was some discussion about the detail that needs to be covered. Should limit it to 'common usage' and refer to the spec for details of edge cases. 09/03: Alan distributed an update to the text tutorial. Please review. 30/03: Steve has spent half a day tidying up lessons 1 to 6 and has uploaded them as pdfs to gridforge. They are now more coherent, and many inconsistencies and errors fixed.

	Ownership of draft lessons (text properties, binary properties, advanced features) has been passed to Steve. Also need to make a schema available for the examples. 13/04: Steve is working on the text properties tutorial. 04/05: No progress 18/05: No progress 01/06: No progress 08/06: No progress 15/06: This is on hold until Steve clears up spec issues and other workload. Steph has looked at the later lessons, and noted that they are more direct compared to the more wordy earlier lessons. 28/06: On hold.
124	**DFDL web content on OGF standards pages (Steve, Bob)** 13/10: no progress 10/11: no progress 17/11: Alan has looked at the OGF web pages and there aren't many standards listed. Some of the links point to very short primers rather than the specification ... 08/12: Alan to produce some information to be ready for when spec is approved. Still no word about is it was discussed/approved at OGF meeting 15/12: no progress 22/12: Steve has developed a summary web page for DFDL which will be sent to OGF when spec is approved. 12/01: Not heard from Joel about updated OGF pages Alan to chase. Will also track other site updates: Wikipedia, IBM developerworks etc. 19/01: Still no response from Joel. Other web site that need updating - IBM virtual XML - Defuddle - Wikipedia - Need google trawl for others Also need to make spec and tutorials more accessible on the web, eg in pdf and/or html format. 26/01: Still not heard from Joel about OGF web pages. PDF versions of the Specification and tutorials have been uploaded to gridforge. 02/02: A DFDL web page is available at www.ogf.org/dfdl. We need to update the IBM virtual XML and MCSA Defuddle pages. Will ask Mike Beckerle to update his DFDL page. 09/02: Wikipedia DFDL page is available. 23/02: All sites except Defuddle have been updated. 02/03: NCSA web pages have been updated. The DFDL WG home page needs updating and should provide links to spec and tutorials. 09/03: Steve has updated the DFDL WG home page and is in contact with Edinburgh University to update an old DFDL presentation. Would like to have a separate DFDL tutorial page to link to the individual lessons. 30/03: Bob will chase the update of the old Defuddle web pages. IBM investigating conversion of spec and tutorials from pdf to html for usability from browsers. 13/04: Bob still sorting out the Defuddle updates. IBM work to convert the spec to html has started. 04/05: Jim Myers has updated the sourceforge download page but not the Defuddle home/overview pages, Bob will chase. IBM has converted the spec to html pages, needs some tidying before being made public. 18/05: IBM aims to publish the web version by end June. Steve checking with OGF whether the web spec is a 'derivative work' in terms of the copyright notice, or can be considered an actual copy of the spec. 01/06: Still waiting on status of web spec from OGF 08/06: OGF have come back to us and said that it is ok to create a web version of the spec. If it is unchanged then it is a copy, otherwise it is a derivative work. Either way, the copyright covers this so there is not a problem. The web version is looking good, some minor tidy-up changes needed where formatting is not quite right. When IBM is happy with it, Steve will circulate to the WG for review. 15/06: Ongoing 28/06: Steve needs to take a final look at all the pages, get any problems fixed, then

	distribute to the WG
132	**Publishing DFDL xsd (Suman)** 08/12: Agreed that it should be made available. Suman has started the approval process in IBM 15/12: no progress 22/12: no update 12/01: Suman is getting approval from IBM to publish. 19/01: Waiting to get IBM approval to publish 26/01: no update 02/02: No update 09/02: no update 23/02: no update 02/03: Suman is working through the IBM process to permit publication. There was discussion about what licence the XSD would be published under and how that would effect use in products. Suman to investigate 09/03: No update 30/03: Suman has sent information to IBM legal. Reminded him about the license issue. 13/04: No update. 04/05: IBM has permission to release the DFDL model xsds to WG members only, Suman has a couple more changes to make and will send to Steve for review. License clarification needed. 18/05: Awaiting response from IBM legal. Suman will send Steve the model xsd for review. 01/06: 08/06. Awaiting response from IBM legal. Suman will send Steve the model xsd for review. 15/06: Steve has received the xsds (there are three of them) and will review. 28/06: Not reviewed yet
133	**Make a set of default formats available (Suman)** 19/01: Suman expects some default formats to be ready by Feb 9th. Will need approval to publish 26/01: Stephanie sent the defaults used by test cases to Suman 02/02: no update 09/02: no update 23/02: no update 02/03: Same as 132 09/03: No update. Same license issue apply though. 13/04: No update. 04/05: IBM will make one default format available. Suman is working through the IBM process to permit publication. 18/05: Awaiting response from IBM legal. 01/06: 08/06: Awaiting response from IBM legal. IBM also want to prove that the default format has the properties sensibly defined to plans to include in internal testing. 15/06: No change. 28/06: No change
136	**Arrays with missing elements (Steve)** There is a problem when there are empty/missing array elements with an index greater than minOccurs. For example: xs:element name"array" minOccurs=0, maxOccurs=10 lengthKind='delimited' Datastream: ,,,value3,value4 Infoset will contain: array[1] = value3 array[2] = value4 This is because elements with an index greater than minOccurs are optional and so do not get defaulted.

	Unparsing this infoset will produce: Datastream: value3,value4 You could make the empty space (%ES;) the nil value which will work for simple elements but not for complex. Infoset will contain: array[1] = nil array[2] = nil array[3] = value3 array[4] = value4 23/02: Discussed options 1. Changed definition of required for arrays to be 'required up to the last instance in the data stream of the array' 2. add index to the element info item Steve to investigate if XDM uses an index. 02/03: No progress 09/03: No update 30/03: Stephanie recognised the issue from IBM's WTX. Here, the solution was to provide an option so that the user explicitly chose whether the position in the array was significant, as it is not always and on output some users do not want defaults or blanks to appear. Not yet resolved. 13/04: Steve has verified that XDM does not carry index information, but will check with IBM's W3C rep to see if that has ever been expressed as a requirement. Steve also said that the area of defaulting missing required elements on parsing (especially complex elements) is one that the IBM implementation team has raised some concerns about, so Steve and the team are looking at this area again. It is possible that the spec will change to clarify behaviour, and so this action should be used to cover this work. 04/05: In progress. 18/05: In progress, at minimum some clarification to the spec is needed 01/06: Discussed to bring Mike up to speed. Solution will depend on other spec discussions. 08/06: Still under discussion 15/06: Will come back to this after action 140 resolved 28/06: As above
139	**Spec issue: is specified length extraction sensitive to in -scope delimiters ?** 01/06: See minutes 08/06: Still under discussion. Useful input from Steph via e-mail which shows that some formats require early termination is detected and reported. 15/06: Discussed in depth. Conclusions so far: a) In-scope delimiters - Using lengthKind 'explicit' means that scanning has been turned off and in-scope delimiters are not looked for within or between elements - Same deal for 'implicit' for both simple and complex elements. For complex, we are working bottom up and length is defined by the children only. - Same deal for lengthKinds 'prefixed' and 'pattern'. - This implies that when lengthKind is not 'delimited' or 'endOfParent' it resets the in-scope markup 'stack' b) The technique of using a dfdl:assert to police a length constraint when lengthKind is 'delimited' was discussed. - Several industry formats such as HL7 and TLOG are delimited but impose such physical length constraints - The only issue is that specifying a dfdl:assert for every element is cumbersome - Validation can be used to do this - but only works reliably for xs:string because the xs:min/maxLength facets do not apply to number/calendar/boolean types - Tim suggested using the (unused) dfdl:length property to do this - Should this be an actual length or a maximum length? It was noted that the industry formats mentioned above typically specify a maximum length. - There is an analogous situation today with padding to a minimum length - xs:minLength facet is used for strings, but dfdl:textOutputMinLength is used for other types

	- To be concluded 28/06: An extra call was held between Mike, Steve, Tim and Steph. a) Confirmed the rules above b) Use of dfdl:length or a separate property proved problematic. Proposal is to have a new failure type called 'Recoverable Error' after which the parser will continue. Importantly, it does not cause backtracking to take place when speculating. It can be raised via a new enum attribute on dfdl:assert called 'failureType'. Full proposal will be mailed to the WG in due course. Also noted that the combination of lengthUnits=bytes and the use of min/maxLength facets was allowed but words missing from the spec to explain this. Also noted that section 5.2.2 is misleading and needs revising. Steph noted that the new failure type should be used with care as it stops backtracking.
140	**Spec issue: Parsing: 'missing' v 'empty', role of initiators , default values** 01/06: See minutes. 08/06: Still under discussion. Tim has sent Mike a selection of data formats to guide the discussion. 15/06: Not discussed - an extra call has been scheduled to go through this. 20/06: A series of extra calls are being held between Mike, Steve, Tim and Steph.

Closed actions

No	Action

Deferred actions

No	Action
129	**Press release to publicise DFDL (Steve)** Steve is pulling together a press release at IBM. Want to include as many contributors and interested parties as possible.NCSA are keen to be included. Also likely that US National Archive will want to be included. Mike has indicated OCO are too. 17/11: no progress ... 08/12: Still no response from IBM press office 15/12: no progress .... 09/03: No progress 30/03: Making this action deferred until IBM is in a position to say something more concrete about any implementation.
131	**Transformation of DFDL properties to a canonical form (Joe)** 08/12: Joe has produced a XSLT to transform a DFDL schema to a canonical element form. When tested it should be made available on the WG gidforge site. 15/12: Alan tested against test dfdl schema which worked correctly (after fixing some errors in the schema) 22/12: no update 12/01: Joe has some defects to fix before making available on gridforge. 19/01: There is a difficult problem to solve before Joe make the style sheet public 26/01: Working on problems 02/02: no progress 09/02: As it wasn't a simple as exoected this will be treated as a low priority action 23/02: Low prioity 09/03: Low priority 30/03: Deferring for now
066	**Investigate format for defining test cases (All)** 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 10/11: Legal issues cleared, IBM in process of collecting 100 example test cases, ideally ones that fit the 'extended conformance' of NCSA Daffodil 17/11: Work is progressing on verifying the test cases. It should be possible to distribute to the WG in 2 weeks. 24/11: About half the test cases have been completed and are being reviewed internally. 01/12: Test cases should be available shortly 08/12: The test cases are in internal IBM review. Probably need a bit of reorganising before publication Stephanie gave a brief overview of the format of the test cases. 15/12: Ruth joined the call to provide the latest status. The test cases have been updated and a draft read.me produced. Although not ready for public distribution Ruth will send them to Joe for feedback. 22/12: Test cases were sent to Joe for initial testing which found some problems in the Daffodil parser 12/01: All current tests use a default format whih Daffodil doesn't currently support. Joe suggested that there should be test that defined the same function using different definition forms. Also suggested that default formats should be provided by the WG. This had always been the intention. Action 133 raised to track. 19/01: There is currently no resource available in IBM to make more tests available. IBM to discuss how/if it can make a 'minimal compliance test suite' available. 26/01: Action kicked off within IBM. There was a brief discussion abot naming and organisation of test cases but no preferences were expressed 02/02: IBM will not have the resources to develop a full test suite in the near future. Steve suggested that we produce a list of required test cases so that anyone could supply them. 09/02: Steve had previously sent a list of areas to be tested. Please review. 23/02: Please review Steve's list of areas to be tested 02/03: Alan had reviewed Steve's list and we went through his comments. Agreed there is no need for separate tests for the infoset or for dfdl: property lists, unions etc but comment will be added that these should be exercised during property testing. 09/03: Alan updated the test document. Need more introduction and perhaps adopting the OGF template. 30/03. Ownership of test document passed to Steve. This action is merged with 112 and will cover all aspects of compliance suite. 13/04: IBM will not have time to create a compliance suite in the near future. Probably best to make this action deferred for now.

Work items:

No	Item	Owner	Target	Status

043	Track errata list for 1.0 of the spec:		Steve	N/A	
	1	The ref property needs to state that circular paths are a schema definition error.			
	2	Clarify what packed and BCD calendars mean. - No need to use a separate VDP property. The only place where a decimal point can occur is for fractional seconds. This is detectable from the pattern at the boundary of 's' and 'S', ie sS - Property calendarPatternKind = 'explicit' must be used with binary calendar representations, as the defaults for 'implicit' use non-numeric characters. Schema defn error otherwise. - Property binaryCalendarRep should restate the rule from property calendarPattern. - Examples to be provided.			
	3	It looks like the DFDL spec does not fully state the time zone symbol behaviour, as it quotes from http://icu-project.org/apiref/icu4c/classSimpleDateFormat.html instead of from http://userguide.icu-project.org/formatparse/datetime. It should say: z Time Zone: specific non-location (z, zz, or zzz) PDT zzzz Time Zone: specific non-location Pacific Daylight Time Z Time Zone: RFC 822 (Z, ZZ, ZZZ) -0800 ZZZZ Time Zone: localized GMT GMT-08:00 v Time Zone: generic non-location PT vvvv Time Zone: generic non-location Pacific Time V Time Zone: generic non-location PT VVVV Time Zone: generic location United States (Los Angeles) Note that both Z and ZZZZ can be followed by 'U' which is a DFDL extension.			
	4	Property precedence corrections: - 22.1.1 & 22.2.1. Binary representations can have delimited lengthKind .			

		- 22.1.2 & 22.2.2. Complex elements can have endOfParent lengthKind.			
		5 Spec talks about 'content region' but it should be more specific in terms of the grammar, and use 'SimpleContent region' or 'ComplexContent region', or both.			
		6 Properties hiddenGroupRef, inputValueCalc and outputValueCalc change to behave like dfdl:ref. That is, they can not be placed in scope, and are only set at their point of use. Empty string is not an allowed value. Note that this removes the need for errata 6.			
		7 On parsing, state that nilLiteralCharacter test takes place on the untrimmed rep value.			
		8 Clarify that when lengthUnits is bytes, using lengthKind implicit for a string interprets the min/maxLength facets as byte values and not characters when parsing/unparsing.			
		9 6.3. Bullet for logical value. State that the string must obey the lexical representation of the type.			
		10 6.3. Literal white space only ever used as list token separator. Use entities if literal white space is needed as part of the property value.			
		11 7.7. Clarify to make clearer that if A.xsd includes B.xsd then B can refer to a variable in A and reference is via QName in the usual way..			
		12 9.2 and 12.3.4. Allow a prefix length type to have a prefix length, but not beyond. Types can't be the same.			
		13 13.6. When textNumberRep is zoned, property description should state that base is assumed to be 10.			