

Minutes of the Bellevue meeting of ISO/IEC JTC 1/SC 34/WG4, 2013-06-17/20

Rex Jaeschke (rex@RexJaeschke.com)

2013-06-26

1. Opening remarks

The meeting started at 13:50 on 2013-06-17. The convener, Murata-san, welcomed everyone to the 17th face-to-face meeting of WG4.

2. Roll call of delegates

The following members were present during part or all of the meeting:

Name	Affiliation	Employer/Sponsor
Sam Oh	SC 34 Chair	Sungkyunkwan University
Makoto Murata	WG4 Convener, JP	International University of Japan
Dennis Hamilton	Invited expert	Consultant
Xia HOU	CN	Beijing Information Science and Technology University
Florian Reuter	Invited expert	Consultant
Rex Jaeschke	Ecma, Project Editor	Consultant
Jim Thatcher	Ecma, US	Microsoft
John Haug	Ecma, US	Microsoft
Francis Cave	GB	Francis Cave Digital Publishing
Alex Brown	GB	Griffin Brown Digital Publishing Ltd.

Name	Affiliation	Employer/Sponsor
Gareth Horton	GB	Datawatch
Jaeho Lee	KR	University of Seoul

Present were 12 people representing 5 NBs, and 1 liaison.

3. Adoption of the agenda

The agenda ([SC 34 N 1908](#)) was adopted, with the following change: All occurrences of this meeting's dates should be changed from 2013-06-18/21 to 2013-06-17/20.

4. Administration

Approval of Previous Meeting Minutes [[WG4 N 0257](#)]

The draft minutes were approved as circulated.

Outstanding Action Items

1. John H. to report re the addition of Digital Signature support for OPC. **Pending**
2. Jesper to write a Wiki page on the Assembla site ([https://www.assembla.com/spaces/IS29500/wiki/Package-level_encryption_\(OPC\)](https://www.assembla.com/spaces/IS29500/wiki/Package-level_encryption_(OPC))) proposing how to move this idea forward. **Pending**
3. Chris to update his application-defined MCE elements example (sent to the list some months ago) to also incorporate output documents. **Done**
4. Murata-san will produce and circulate a set of proposed solutions for all OPC-related DRs in advance of the next F2F meeting. **Done**
5. Murata-san will revise his proposed resolution to DR 11-0010 to include multi-column and vertical writing. **Pending** (depends on the resolution of DR 11-0008)
6. Re DR 13-0004, "SML: Which of ext and extLst is an application-defined extension element?", Murata-san to test ignorable attributes of extLst elements. **Done**
7. John H. will update the MCE spec based on the feedback from the telcon. **Done**
8. Rex will distribute the latest MCE spec as a committee document, along with an editorial report identifying the changes. **Done** [see [WG4 N 0260](#)]
9. WG4 members will review the MCE spec draft in advance of a detailed walk-through in June. **Done**
10. Murata-san to invite Florian to call-in to the Seattle meeting w.r.t the MCE spec walk-through. **Done**

11. Chris and John H. will produce a set of comments and questions on WG4 N 0207, “Improving Part 2 in reply to DRs”, and to send them to the email list. **Done**
12. Rex will distribute a new base document for Part 2, which complies with the ISO Style Guidelines, along with an editorial report identifying the changes. **Done** [see WG4 N [0259](#)]

Report from the WG4 Secretariat

The following NBs and liaisons have registered delegates to WG4: BR, CA, CH, CI, CN, CZ, DE, DK, Ecma, FI, FR, GB, IN, IT, JP, KR, NL, NO, OASIS, PL, US, W3C, XML Guild, and ZA. All requests for additions, deletions, and changes to the delegate list should be sent to the WG4 Secretariat (rex@RexJaeschke.com).

The WG4 email list is e-SC34-WG4@ecma-international.org. The document repository is at http://lucia.itscj.ipsj.or.jp/itscj/servlets/ScmDoc10?Com_Id=w4.

5. Defect Reports

The public online DR log is at <http://www.29500sc34comments.org/>. Access individual DRs via the hyperlinks contained within the spreadsheet’s left-most column.

DR 10-0015, “OPC: Relationship Markup”

Murata-san presented a slide containing the following:

- OPC
 - Part names and zip item names differ (%HH)
 - Interleaving
 - pack URI scheme
- ODF Package UTF-8
- EPUB OCF UTF-8
- EPUB CFI
 - Any resource?
 - Temporal/spatial/textRange

- Pack URI scheme
 - any part in a package is referenceable
 - fragments in parts have to be referenced by frag identifiers that follow pack URIs
 - ODF uses fragment identifiers

Separately, we discussed those relative references beginning with "/" and those not beginning with "/". Are both allowed? How are they interpreted? Do OOXML and XPS correctly use them?

From Part 2, §9.3.2:

The Target attribute of the Relationship element holds a URI that points to a target resource. Where the URI is expressed as a relative reference, it is resolved against the base URI of the Relationships source part. The xml:base attribute shall not be used to specify a base URI for relationship XML content.

But what is a relative reference? We reviewed RFC 3986. We found that both those beginning with "/" and those not beginning with "/" are allowed, as shown in RFC 3986, section 4.2. Those beginning with "/" are called *absolute-path references*, while those not beginning with "/" are called *relative-path references*. Confusingly, both are relative references.

However, absolute-path references and relative-path references are resolved differently, as shown in the algorithm in RFC 3986, section 5.2.2. When an absolute-path reference is resolved relative to a base URI, the scheme and authority in the base URI are used, but the path in it is not used. Meanwhile, when a relative-path reference is resolved relative to a base URI, not only the scheme and authority but also the path in the base URI are used. In other words, the difference is that the path in the base URI is used for resolving relative-path references but is not used for resolving absolute-path references.

But what is a base URI when a relative reference appears as the value of the Target attribute of the Relationship element?

From 9.2.1:

Relative references from a part are interpreted relative to the base URI of that part. By default, the base URI of a part is derived from the name of the part, as defined in §B.3 where B.3 is the definition of Pack URI

From Terms and Definitions:

pack URI — A URI scheme that allows URIs to be used as a uniform mechanism for addressing parts within a package. Pack URIs are used as Base URIs for resolving relative references among parts in a package.

Thus, OPC defines the base URI of a part for a relationship target as the pack URI of the source part.

What is a path in a pack URI then? We find that it is a path in an OPC package (i.e., a path from the root of the OPC package). It is not the path of an OPC package, since the path of the OPC package is rather embedded within the authority of the pack URI.

Therefore, we conclude that relative references beginning with "/" are relative to the root of the OPC package and that those not beginning with "/" are relative to the source part. Some XPS files, which are generated by Windows (print to XPS), contain relative references beginning with "/" and they are indeed relative to the root of the OPC package. Meanwhile, other XPS files, which are generated by Office, and OOXML files contain relative references not beginning with "/" and they are indeed relative to the source part.

It took a lot of time for us to reach this conclusion. To help others to reach the same conclusion easily, we ought to improve §9.3.2.

We also discussed "///". Part 2, §B.3 step #5 could leave you with "///" if the path has a leading "/"? Does RFC 3986 have an algorithm for combining a base URI and a relative reference (that has a leading "/"?)? We think that RFC 3986 does not, but some HTTP servers do.

Re “resolution of absolute path reference and relative path reference”, Murata-san provided the following examples:

We assume that an OPC package is available at <http://www.example.com/dir1/example.opc>

Consider a relative path reference "foo1/foo2" in relation to a part "/part1/part11"

The pack URI for the part "/part1/part11" is:

pack://http%3c,,www.example.com,dir1,example.opc/part1/part11

From this URI, we construct:

scheme: pack

authority: http%3c,,www.example.com,dir1,example.opc

path: /part1/part11

From these three and a relative reference "foo1/foo2", we will construct

pack://http%3c,,www.example.com,dir1,example.opc/part1/part11/foo1/foo2

Then, consider an absolute path reference "/foo1/foo2" (which is also a relative reference) in relative to a part "/part1/part11".

This time, the part of the pack URI is not used. Thus, we will construct

pack://http%3c,,www.example.com,dir1,example.opc/foo1/foo2

In other words, "/foo1/foo2" is relative to the root directory within the OPC package.

6. Revising Part 3 (Semantics of Markup Compatibility and Extensions)

(See WG4 N 0260 and https://www.assembla.com/spaces/IS29500/wiki/Semantics_of_MCE.)

See Murata-san's two postings on 2013-06-09.

Rex provided an overview of his Editor's Report provided as part of WG4 N 0260.

We walked through WDO reviewing issues raised by various members. Rex made comments and tracked changes to the draft, and assigned owners for issue resolution, as follows (all editorial issues will be handled by Rex):

1. §1, "Scope": Why is this OOXML-specific? Wasn't the original intent to allow Part 3 (and Part 2) for non-OOXML applications? Yes.
2. §2, "Conformance"
 - We agreed to strike all of §2.1 and to correct §2.2.
 - §2.3, "Application Conformance": Rewrite this subclause to remove the notion of producer and consumer, and specify requirements on markup processors. These requirements are NOT purely syntactic. Remove all mention of conformance class MCE.

Owner: Murata-san; Done

Result: An MCE processor is conformant if it satisfies the requirements specified in §11.

Doing this impacts Part 1, whose §2.1 mentions “conformance class MCE”. Will need to reword that.

Owner: Francis

- **Eventually, we agreed to get rid of the conformance clause entirely. However, we still need to reword Part 1’s §2.1.**

3. §4, “Normative References”:

- Move NVDL entry to Bibliography as only alluded to in an informative annex.
- Add an entry to Bibliography for RELAX NG.
- Add 29500-1 and -4 to Bibliography.
- Remove URI entry, as this spec no longer uses that term URI.
- Remove Unicode entry, as it is not referenced by this spec.
- Update reference to XML Base to latest edition. **[Are we sure this won’t break anything?]**

Owner: Murata-san. **Done**

Result: <http://www.w3.org/TR/2009/REC-xmlbase-20090128/>

- Move 3 x XML Schema entries to Bibliography as only alluded to in an informative annex.
- Add a reference to XML Information Set.

4. §4, “Terms and Definitions”:

- Review the newly added text elsewhere to make sure the terms in this clause get updated to reflect the new text.
- Strike all but first sentence in opening paragraphs.
- Remove the following terms, as they are no longer used: *byte* **and** *compatibility-rule attribute*.
- Global change of “configuration” to “MCE configuration”.
- Review the need for markup consumer, producer, and document. Current definitions don’t seem quite right anyway. Objections were raised as they refer to markup document, which in turn refers to a file format; the conformance of a consumer isn’t relevant.

Owner: Murata-san

- Change “markup preprocessor” to “MCE processor” and consolidate all variants to that term.
- Re “namespace, ignorable” and “namespace, understood”, should we simply define the adjectives in the context of a namespace? (A namespace can be both ignorable and understood.) Do we need to define

“Non-understood”?

Owner: Francis

Result: Create 4.12 entry “namespace” with two subordinate entries, as follows:

4.12.1 ignorable namespace

XML namespace, identified in markup, whose elements and attributes are ignored if the namespace is not an understood namespace

4.12.2 understood namespace

XML namespace that is included in an MCE configuration

Also, add a normative reference to XML Information Set (Second Edition), which is where it is defined.

- Can “recognize” be removed or reformulated?

Owner: Murata-san

5. §6, “General Description”: Consider moving the final paragraph with bullet list to the previous clause, and making §6 informative.
6. §8, “Markup Compatibility Fundamentals”:
 - Get rid of §8.1, “Core Concepts”, as proposed.
 - Removed proposed new text from §8.1, “Markup Compatibility Namespace” and moved the remainder to §9.1, “Introduction”.
 - Consider moving §8.2, “Error Handling” to the processing model clause thereby eliminated §8 altogether.
7. §A (now B), “(informative) Validation Using NVDL”

- Change B.2’s title and make minor correction to narrative.
- Change B.3’s title and after the first paragraph, add a note regarding the example’s being transitional and how a strict example would compare.

Owner: Murata-san; **Done**

Result: Note: This NVDL script handles the conformance class "Transitional". A similar NVDL script for the conformance class "Strict" can be created by replacing each transitional namespace by a corresponding strict namespace and removing <namespace> elements for transitional-only features such as VML.

On Thursday morning, Chuck Jazdzewski joined us by phone for an hour of Q&A. (Chuck was one of the original authors of MCE.) We had a very productive discussion. Below, is a summary of some of that discussion:

1. Declaring application-defined extension elements

WG4 asked why there are no mechanisms for XML documents to declare application-defined extension elements. Chuck provided a very clear answer: Application-defined extension elements have to be chosen when original markup consumers are created, and thus we must not allow new producers to declare new application-defined extension elements, which cannot be handled by existing consumers.

Note: WG4 should reconsider its decision to introduce the ExtesionElement attribute.

2. Dropping PreserveAttributes and PreserveElements

Chuck agreed on the WG4 decision to drop them and further said that they were incorporated into the first edition of ISO/IEC 29500-3 by mistake.

3. Dropping namespace subsumption

At first, Chuck was surprised to hear that WG4 has decided to drop namespace subsumption. It has been used for XAML, for example.

However, he agreed that namespace subsumption could be done by some technologies for namespace renaming. Such renaming works as a thin layer between the MCE processor and application programs. He continues to believe that his implementation, which does namespace subsumption as well as MCE processing, is conformant. WG4 agreed.

4. Non-MCE attributes and child elements of AlternateContent elements

The current design allows them only when they are in an ignorable namespace. Chuck said that this restriction is intended to simplify the design/implementation, if not mandating use of infosets. He did not oppose allowing any foreign attributes and child elements.

5. Inherited attributes

WG4 asked why MCE mentions inherited attributes in the XML namespace and no other inherited attributes. Chuck said that this is intentional. When some elements are unwrapped, consumers do not know how to handle

these elements and thus their attributes should be ignored together. Chuck also said that all inherited attributes in the XML namespace should be mentioned in 29500-3.

6. Copying attributes

The current draft mentions copying of attributes from ancestors to subordinates and mentions the use of infoset for such copying. Chuck thinks that the use of infosets would make MCE simpler. He further pointed out that text chunks directly under an unwrapped element cannot hold inherited attributes and that some nicely-defined infoset can allow text chunks to have attributes.

7. Multiple invocation of MCE

The current working draft does not allow MCE resumption within application-defined extension elements. Rather, it is assumed that subordinate elements of such application-defined extension elements may be handled by another invocation of the MCE processor. Chuck agreed on such multiple-invocation and said that the original MCE concentrated on a single invocation.

We agreed to produce another WD for distribution and review to WG4 as soon as possible. Then after the Delft meeting, we hope to distribute a further WD to the public for comment.

7. Revising Part 2 (Open Packaging Conventions)

(See WG4 N [0259](#).)

The following mail from Chris and Murata-san is relevant:

2013-05-20 MURATA Makoto:

Chris and John,

2013/5/21 Chris Rae <Chris.Rae@microsoft.com>:

> Hi all - many thanks for reposting this, Murata-san. As we mentioned on the WG4 call last week, John and I have been looking through this document and have a couple of questions:

> 1. From item 2: As we discussed on the conference call, we think that ISO/IEC 295000 does not intent permitting a part name with a trailing slash, but the non-terminal ipath-absolute in RFC 3987 does. To this end, we think that we should define part name as 1*("/" isegment-nz).

Agreed. There are no reasons to allow "/" at the end or "//".

> 2. From item 2: The list of constraints is a subset of the original constraints in 9.1.1. Can you remember why the other constraints weren't included?

1) A part URI shall not be empty. [M1.1]

is retained.

2) A part URI shall not have empty segments. [M1.3]

This is required and my wording "An isegment shall be non-empty" and your definition (i.e., 1*("/") isegment-nz)) capture this.

3) A part URI shall start with a forward slash ("/") character. [M1.4]

This is captured by both your definition (i.e., 1*("/") isegment-nz)) and mine (ipath-absolute).

4) A part URI shall not have a forward slash as the last character. [M1.5]

My wording "An isegment shall be non-empty" implies this constraint.

5) A segment shall not hold any characters other than pchar characters. [M1.6]

This has to be changed for allowing non-ASCII characters in part names. We have to allow ipchar. We only have to use isegment and isegment-nz.

> 3. From item 2: Are asterisk and colon characters to be prohibited entirely inside part names, or are we prohibiting only part names that are an asterisk or colon in their entirety?

Anywhere in part names.

> 4. From item 12: Please can you specify which exact five steps are removed from the conversion procedure?

They are shown in the word document attached to the mail available at:

<http://mailman.vse.cz/pipermail/sc34wg4/2011-June/002305.html>

Digital Signature: We discussed the document [MS-OFFCRYPTO]: Office Document Cryptography Structure ([http://msdn.microsoft.com/en-us/library/cc313071\(v=office.12\).aspx](http://msdn.microsoft.com/en-us/library/cc313071(v=office.12).aspx)) and ODF w.r.t. packaging.

We spent considerable time discussing DR 10-0015, “OPC: Relationship Markup”. (See the DR section above.)

8. OOXML Extensions

Nothing to report

9. Other Business

We thanked Microsoft and John Haug for hosting the meeting and the Boeing tour and dinner.

10. Future meetings

Face-to-Face Meetings:

- 2013-09-09/13, Delft, NL (with other WGs and Plenary, as follows:
09 Mon: 10:00 SC 34 opening plenary; 14:30-17:00 WG1
10 Tue: 09:00-12:00 WG5; 13:30-17:00 WG4
11 Wed: 09:00-17:00 WG4
12 Thu: 09:00-17:00 WG4
13 Fri: 09:00-12:00 WG4 and WG3 in parallel; 13:00 SC 34 closing plenary)
- 2014-03-03/07, Berlin, DE (with other WGs; specific days/times TBD)
- 2014-06-16/20, Prague, CZ (with other WGs; specific days/times TBD)

Teleconferences:

- 2013-07-16, 13:00 GMT (US/PT 06:00, GB 14:00, DE/DK/FR/CZ 15:00, JP 22:00)
- 2013-08-13, 13:00 GMT (US/PT 06:00, GB 14:00, DE/DK/FR/CZ 15:00, JP 22:00)

11. Adjournment

Adjourned by unanimous consent at 17:00 on 2013-06-20.