Exiv2 - Bug #1112

lost timezone information in XMP dates

28 Aug 2015 16:46 - Jakub Wilk

Status:	Closed	**Start date:**	28 Aug 2015
Priority:	Normal	**Due date:**	
Assignee:	Robin Mills	**% Done:**	100%
Category:	metadata	**Estimated time:**	15.00 hours
Target version:	0.26		

Description

Exiv2 loses timezone information in some XMP dates:

```
$ cp test.xmp test.xmp.bak
$ exiv2 -M 'del Xmp.dc.title' test.xmp
$ diff test.xmp.bak test.xmp
6c6
<     xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
---
>     xmp:CreateDate="2012-02-01T15:28:00"/>
```

Tested with svn r3911. AFAICT, it didn't happen in 0.24.

Related issues:

Related to Exiv2 - Feature #1050: Extend exif <> iptc <> xmp conversions	**Assigned**	**04 Apr 2015**
Related to Exiv2 - Bug #864: Mapping of Exif DateTime fields to XMP changed i...	**Closed**	**30 Oct 2012**

Associated revisions

Revision 3923 - 03 Sep 2015 20:21 - Robin Mills

#1112. Fix submitted. Also added typedefs to datasets.hpp for Exiv2::Dictionary, Exiv2::StringSet, Exiv2::StringVector

Revision 3924 - 03 Sep 2015 20:35 - Robin Mills

#1112. Regression detector added to test/bugfixes-test.sh

History

#1 - 29 Aug 2015 15:23 - Alan Pater

- Category set to metadata

- Assignee set to Alan Pater

This is a feature, not a bug. ;-)

exiv2 0.25 was updated to follow MWG guidelines(1) and attempts to map the appropriate values between exif and xmp. The catch is that exif does not contain timezone data.

```
~$ exiv2 -pa test.xmp
Exif.Photo.DateTimeDigitized                 Ascii      20  2012:02:01 09:28:00
Iptc.Application2.DigitizationDate           Date        8  2012-02-01
Iptc.Envelope.CharacterSet                   String      3
Xmp.xmp.CreateDate                           XmpText    25  2012-02-01T16:28:00+02:00
~$  cp test.xmp test.xmp.bak
~$  exiv2 -M 'del Xmp.dc.title' test.xmp
~$  diff test.xmp.bak test.xmp
6c6
<     xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
---
>     xmp:CreateDate="2012-02-01T09:28:00"/>
```

The result you see comes from Exif.Photo.DateTimeDigitized.

(1) http://www.metadataworkinggroup.org/pdf/mwg_guidance.pdf#page=37

#2 - 29 Aug 2015 17:16 - Robin Mills

- Status changed from New to Resolved

- Target version set to 0.26

- % Done changed from 0 to 100

Thanks for looking at this, Alan. And thanks to Jakub for such a clear issue report and test case. Thank You, Gentlemen.

I'm going to assign this to Alan to ensure that he's credited in the release notes. I'm also going to mark it 100% Resolved. If Jakub doesn't contest Alan's reply, we'll mark this "Closed" before we ship v0.26.

#3 - 30 Aug 2015 20:31 - Alan Pater

- Status changed from Resolved to Assigned

On second thought, I am not sure that correct behaviour is being shown here.

The sample xmp file only has the Xmp.xmp.CreateDate property

```
~$ cat test.xmp
<?xml version="1.0" encoding="UTF-8"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="XMP Core 4.4.0-Exiv2">
 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about=""
    xmlns:xmp="http://ns.adobe.com/xap/1.0/"
   xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
 </rdf:RDF>
</x:xmpmeta>
```

Exif.Photo.DateTimeDigitized and the IPTC tags are being generated by exiv2 0.25 and adjusted to my local timezone. I don't see why the conversion is being triggered here. As Jakub said, this does not happen with 0.24

#4 - 30 Aug 2015 22:28 - Alan Pater

Digging in further, on exiv2 0.24 and earlier, Exif.Photo.DateTimeDigitized was mapped to Xmp.exif.DateTimeDigitized rather then Xmp.xmp.CreateDate.

So, with an xmp file containing Xmp.exif.DateTimeDigitized:

```
$ cat exif.xmp
<?xml version="1.0" encoding="UTF-8"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="XMP Core 4.4.0-Exiv2">
 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about=""
    xmlns:exif="http://ns.adobe.com/exif/1.0/"
   exif:DateTimeDigitized="2012-02-01T16:28:00+02:00"/>
 </rdf:RDF>
</x:xmpmeta>
```

Results in:

```
$ exiv2 -pa exif.xmp
Exif.Photo.DateTimeDigitized                 Ascii      20  2012:02:01 09:28:00
Xmp.exif.DateTimeDigitized                   XmpText    25  2012-02-01T16:28:00+02:00
$ cp exif.xmp exif.xmp.bak
$ exiv2 -M 'del Xmp.dc.title' exif.xmp
$ diff exif.xmp.bak exif.xmp
6c6
<    exif:DateTimeDigitized="2012-02-01T16:28:00+02:00"/>
---
>    exif:DateTimeDigitized="2012-02-01T09:28:00"/>
```

So 0.24 DOES shows the same behaviour as 0.25 but with a xmp file containing Xmp.exif.DateTimeDigitized.

#5 - 30 Aug 2015 22:51 - Robin Mills

Going back to Jakub's issue report, his test file in pure XMP. So why are we loosing TZ data when the user removed *the Title*. I think it's worth finding (and reconsidering) the change which introduced this.

Can it be as simple as mapping one Exiv2 key into XMP. Looking at my Stonehenge photo, the Exif Dates don't have TZ info - that's being stored in the Nikon Makernote Exif.NikonWt.Timezone:

```
 $ exiv2 -pa -g Date -g Time -g description -g Caption http://dev.exiv2.org/attachments/download/805/DSC_7154.
jpg
```

```
Exif.Image.DateTime                           Ascii      20  2015:07:16 20:25:28
Exif.Photo.ExposureTime                       Rational    1  1/400 s
Exif.Photo.DateTimeOriginal                   Ascii      20  2015:07:16 15:38:54
Exif.Photo.DateTimeDigitized                  Ascii      20  2015:07:16 15:38:54
Exif.NikonWt.Timezone                         SShort      1  UTC +00:00
Exif.NikonWt.DateDisplayFormat                Byte        1  Y/M/D
Exif.Photo.SubSecTime                         Ascii       3  00
Exif.Photo.SubSecTimeOriginal                 Ascii       3  00
Exif.Photo.SubSecTimeDigitized                Ascii       3  00
Exif.GPSInfo.GPSTimeStamp                     Rational    3  14:38:55.9
Exif.GPSInfo.GPSDateStamp                     Ascii      11  2015:07:16
Iptc.Application2.Caption                     String     12  Classic View
Xmp.xmp.ModifyDate                            XmpText    25  2015-07-16T20:25:28+01:00
Xmp.dc.description                            LangAlt     1  lang="x-default" Classic View
$
```

The photo was taken at 3:38pm. The Xmp.xmp.ModifyDate was added by Picasa when I added the Caption "Classic View" about 8:25pm

#6 - 30 Aug 2015 23:31 - Alan Pater

Well, convert.cpp dates to March 2008.

http://dev.exiv2.org/projects/exiv2/repository/changes/trunk/src/convert.cpp

But I am wondering the same. Why are the conversions getting called and applied when they are not being explicitly asked for?

#7 - 30 Aug 2015 23:33 - Alan Pater

Also note that test.xmp does not contain a title. The test case is trying to remove a non-existent value.

#8 - 31 Aug 2015 10:10 - Robin Mills

I've built a "vanilla" v0.24 this morning to investigate something on the forum from Mikayel. I've rerun Jakub's test. Here's the output:

```
603 rmills@rmillsmbp:~/Downloads $ cp test.xmp test.xmp.bak
604 rmills@rmillsmbp:~/Downloads $ exiv2 -M 'del Xmp.dc.title' test.xmp
605 rmills@rmillsmbp:~/Downloads $ diff test.xmp test.xmp.bak
1c1
< <?xml version="1.0" encoding="UTF-8"?>
---
> <?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
8a9
> <?xpacket end="w"?>
\ No newline at end of file
606 rmills@rmillsmbp:~/Downloads $
```

As you can see, XMPsdk has made xml changes, however the CreateDate retains the TZ information in v0.24.

```
613 rmills@rmillsmbp:~/Downloads $ grep CreateDate test.*
test.xmp:   xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
test.xmp.bak:   xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
614 rmills@rmillsmbp:~/Downloads $
```

Just for information, here's test.xmp.bak

```
614 rmills@rmillsmbp:~/Downloads $ cat test.xmp.bak
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="XMP Core 4.4.0-Exiv2">
 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about=""
    xmlns:xmp="http://ns.adobe.com/xap/1.0/"
   xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
 </rdf:RDF>
</x:xmpmeta>
<?xpacket end="w"?>615 rmills@rmillsmbp:~/Downloads $
```

It's a public holiday in England today and therefore pouring with rain. I'll build different SVN revisions and see when/why this changes has taken place.

#9 - 31 Aug 2015 10:35 - Robin Mills

No doubt about it. This arrived with r3659:

```
633 rmills@rmillsmbp:~/Downloads $ cp test.xmp.bak test.xmp ; exiv2 -M 'del Xmp.dc.title' test.xmp ; diff test
.xmp.bak test.xmp
```

```
1c1
< <?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
---
> <?xml version="1.0" encoding="UTF-8"?>
4a5
>     xmlns:exif="http://ns.adobe.com/exif/1.0/"
6c7,8
<    xmp:CreateDate="2012-02-01T16:28:00+02:00"/>
---
>    exif:DateTimeDigitized="2012-02-01T14:28:00"
>    xmp:CreateDate="2012-02-01T14:28:00"/>
9d10
< <?xpacket end="w"?>
\ No newline at end of file
634 rmills@rmillsmbp:~/Downloads $ exiv2 -vV -g svn -g version
exiv2 0.24 001800 (64 bit build)
version=4.2.1 Compatible Apple LLVM 6.1.0 (clang-602.0.53)
svn=3658
id=$Id: version.cpp 3639 2015-03-25 19:55:02Z robinwmills $
635 rmills@rmillsmbp:~/Downloads $
```

Here are the same commands on r3658:

```
620 rmills@rmillsmbp:~/Downloads $ cp test.xmp.bak test.xmp ; exiv2 -M 'del Xmp.dc.title' test.xmp
621 rmills@rmillsmbp:~/Downloads $ diff test.xmp.bak test.xmp
1c1
< <?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
---
> <?xml version="1.0" encoding="UTF-8"?>
9d8
< <?xpacket end="w"?>
\ No newline at end of file
622 rmills@rmillsmbp:~/Downloads $ exiv2 -vV -g svn -g version
exiv2 0.24 001800 (64 bit build)
version=4.2.1 Compatible Apple LLVM 6.1.0 (clang-602.0.53)
svn=3658
id=$Id: version.cpp 3639 2015-03-25 19:55:02Z robinwmills $
623 rmills@rmillsmbp:~/Downloads $
```

Good News: it's still raining! So I'll dig into the code this morning. Hoping for a sunny afternoon with friends coming for BBQ/dinner.

#10 - 31 Aug 2015 12:13 - Alan Pater

- File exif.xmp added

Test file that shows issue existed pre r3659.

#11 - 31 Aug 2015 12:36 - Alan Pater

Here I show that the issue existed even earlier. Exiv2 0.23 was released in April 2013.
This is with the exif.xmp test file rather then test.xmp.

```
$ exiv2 -vV
exiv2 0.23 001700 (64 bit build)
```

Let's look directly inside the exif.xmp file. It only contains a value for Xmp.exif.DateTimeDigitized.

```
$ cat exif.xmp
<?xml version="1.0" encoding="UTF-8"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="XMP Core 4.4.0-Exiv2">
 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about=""
    xmlns:exif="http://ns.adobe.com/exif/1.0/"
   exif:DateTimeDigitized="2012-02-01T16:28:00+02:00"/>
 </rdf:RDF>
</x:xmpmeta>
```

Now let's see how exiv2 interprets that file.

```
$ exiv2 -pa exif.xmp
Exif.Photo.DateTimeDigitized                 Ascii      20  2012:02:01 09:28:00
Xmp.exif.DateTimeDigitized                   XmpText    25  2012-02-01T16:28:00+02:00
```

exiv2 displays a value for Exif.Photo.DateTimeDigitized that does not exist in the exif.xmp file. It has not yet written any new information to the exif.xmp file.

```
$ cp exif.xmp exif.xmp.bak

$ exiv2 -M 'del Xmp.dc.title' exif.xmp
```

That was the command that wrote changes to the exif.xmp file. Exiv2 takes the value it has generated for Exif.Photo.DateTimeDigitized and writes it to Xmp.exif.DateTimeDigitized.

```
$ diff exif.xmp exif.xmp.bak
6c6
<    exif:DateTimeDigitized="2012-02-01T09:28:00"/>
---
>    exif:DateTimeDigitized="2012-02-01T16:28:00+02:00"/>
```

#12 - 31 Aug 2015 12:58 - Alan Pater

Robin, the change to the xmp header is, I think, unrelated to the lost timezone issue. The xmp header change happened in r3662 via issue #774.

#13 - 31 Aug 2015 13:05 - Robin Mills

Thanks, Alan. I've reproduced your findings with a "vanilla" v0.24 (built from our published/archived v0.24 http://www.exiv2.org/archive.html).

I think we're both right. There's something causing us to loose TZ data. This loss is new in r3659 with xmp.CreateDate ***AND*** has been forever in exif.DateTimeDigitized.

I haven't looked at r3662 yet. Hold your guns, buster. Gimme a chance to understand what's going on here. It's still pouring here, so I'll dig about a bit more. I've never visited this part of the code before. In fact, I didn't even know until earlier this year that exiv2 had conversion code.

#14 - 31 Aug 2015 15:03 - Robin Mills

There's no doubt that the issue with r3659 is coming from src/convert.cpp:346:

```
{mdExif, "Exif.Photo.DateTimeDigitized", "Xmp.xmp.CreateDate", &Converter::cnvExifDate,&Converter::cnvXmpDate
},
```

That code is converting an Exif date (which doesn't have TZ information) into Xmp.xmp.CreateDate.

I don't know why it's being called for a "pure xmp" file. It feels as though something has manufactured Exif.Photo.DateTimeDigitized from the original xmp.CreateDate and removed the TZ information. Then we convert the Exif date into an XML date which can (but doesn't in this case) store TZ. And here's the evidence:

```
$ exiv2 -pv -g Exif ~/Downloads/test.xmp.bak
cnvXmpDate  erase_= 0 from,to = Xmp.xmp.CreateDate,Exif.Photo.DateTimeDigitized ...
... value= 2012-02-01T16:28:00+02:00 result = 2012-02-01T16:28:00+02:00
0x9004 Photo        DateTimeDigitized           Ascii      20  2012:02:01 14:28:00
$
```

The output cvnXmpDate erase_ cvnXmpDate is coming from debug code I have added locally *change listed below*.

cvnXmpDate created a string of 26 bytes which includes the TZ. However something chopped it to 20 bytes as specified http://www.exiv2.org/Exif2-2.PDF p24 Section 4.6.5. The long-standing conversion Exif.Photo.DateTimeDigitized -> Xmp.exif.DateTimeDigitized has the same behaviour, as you have shown with the file exiv.xmp

I believe r3659 is correct when operating on "real" image files (jpeg etc). And I believe r3662 is good as it's an XML formatting/presentation change.

I think the issue here is that we should not use those conversion tables for .xmp files.

My local change to convert.cpp is:

```
    void Converter::cnvXmpDate(const char* from, const char* to)
    {
        Exiv2::XmpData::iterator pos = xmpData_->findKey(XmpKey(from));
        if (pos == xmpData_->end()) return;
        if (!prepareExifTarget(to)) return;
#ifdef EXV_HAVE_XMP_TOOLKIT
        std::string value = pos->toString();
        std::cout << "cnvXmpDate  erase_= " << erase_ << " from,to = " << from << "," << to << " value= " << v
alue ;
        if (!pos->value().ok()) {
... unchanged ...
        }

        std::cout << " result = " << (*xmpData_)[from].value().toString() << std::endl;

        if (erase_) xmpData_->erase(pos);
#else
```

```
# ifndef SUPPRESS_WARNINGS
        EXV_WARNING << "Failed to convert " << from << " to " << to << "\n";
# endif
#endif // !EXV_HAVE_XMP_TOOLKIT
    }
```

#15 - 31 Aug 2015 15:47 - Alan Pater

I believe the conversion functions only apply to xmp files. I tested exiv2 0.23 on a jpeg file with the same xmp value set:

```
$ exiv2 -pa exif.jpg
Xmp.exif.DateTimeDigitized                   XmpText    25  2012-02-01T16:28:00+02:00

$ exiv2 -M 'del Xmp.dc.title' exif.jpg

$ exiv2 -pa exif.jpg
Xmp.exif.DateTimeDigitized                   XmpText    25  2012-02-01T16:28:00+02:00
```

No change, no loss of data. The conversion functions did not get triggered.

Let's try something else:

-eX creates an xmp sidecar file.

```
$ exiv2 -eX exif.jpg
```

-iX imports and converts data from the xmp sidecar file.

```
$ exiv2 -iX exif.jpg
```

Exif.Photo.DateTimeDigitized has been added with a value derived from Xmp.exif.DateTimeDigitized and adjusted to the timezone of this computer.

```
$ exiv2 -pa exif.jpg
Exif.Image.ExifTag                           Long        1  26
Exif.Photo.DateTimeDigitized                 Ascii      20  2012:02:01 09:28:00
Xmp.exif.DateTimeDigitized                   XmpText    25  2012-02-01T16:28:00+02:00
```

#16 - 31 Aug 2015 16:12 - Robin Mills

Yuck, this is horrible. I don't think this is a sensible feature - however it's been there a long time and we're stuck with it. We shouldn't massage/manufacture anything. Never. We should report/update what's there. I think there's a feedback loop in the code. xmpsidecar.cpp reads the XMP, then runs the convertors to *"manufacture fake"* Exif and Iptc data.

```
        copyXmpToIptc(xmpData_, iptcData_);
        copyXmpToExif(xmpData_, exifData_);
    } // XmpSidecar::readMetadata
```

I think, later on, we push the *"fake manufactured"* Exif data back to XML and loose the TZ information.

Sadly, it's still pouring and there will be no BBQ today. I might as well keep working on this, it's more interesting than the rain.

#17 - 31 Aug 2015 20:56 - Robin Mills

- Assignee changed from Alan Pater to Robin Mills

- % Done changed from 100 to 60

- Estimated time set to 30.00 h

There's no doubt that this puzzle is limited to sidecar XMP->Exif->XMP Date tags loosing their time-zone data. The Exif Data is a 20 byte string without TZ data.

Alan is quite correct. This has always occurred when the sidecar contains:

```
exif:DateTimeDigitized="2012-02-01T16:28:00+02:00
```

causes a new tag to appear:

```
xmp:CreateDate="2012-02-01T16:20:00"
```

Jakub is quite correct that r3696 has changed this tag:

```
xmp:CreateDate="2012-02-01T16:28:00+02:00"
```

to be rewritten as

```
xmp:CreateDate="2012-02-01T16:20:00"
```

The previous behaviour was to ignore xmp:CreateDate and it would pass unchanged.

Do we really need to fix this, Jakub? I've spent about 12 hours on this and I'm wondering if this deserves more time.

#18 - 01 Sep 2015 02:56 - Alan Pater

The changes that occur when using the -eX and -iX flags are by design so they should stay as is. With issue #1050 I am hoping to make those conversions both more robust and fully compliant with MWG guidelines, including ensuring that data is never lost.

But I remain unsure that it is correct that the convert functions get triggered in this case. I don't think that is by design. It looks like an oversight.

#19 - 03 Sep 2015 20:24 - Robin Mills

- % Done changed from 60 to 90

Fix submitted: r3923. No side effects! Passes test suite without changing anything in the test suite.

TODO:
Add regression detector with Jakub's to test suite.

#20 - 03 Sep 2015 20:37 - Robin Mills

- Status changed from Assigned to Resolved

- % Done changed from 90 to 100

- Estimated time changed from 30.00 h to 15.00 h

r3924. Submitted regression detector.

I'm going to mark this resolved/100% for now and, if nothing else arises from this, this issue will be closed prior to shipping v0.26.

#21 - 13 Oct 2015 19:54 - Robin Mills

- Status changed from Resolved to Closed

Files

test.xmp	358 Bytes	28 Aug 2015	Jakub Wilk
exif.xmp	334 Bytes	31 Aug 2015	Alan Pater