Subject: Bugfix: XmlParser in endless loop
Posted by zsolt on Mon, 02 Oct 2023 16:07:03 GMT
View Forum Message <> Reply to Message

Try this simple code:
ParseXML("");

This function goes into an endless loop, because at the end of sReadXmlNode() function, the line
p.ReadText(); // skip empty text
doesn't do anything.

My first idea was to change it to
p.Skip();// skip empty text
It seems to be better, but the error message will not be too useful.

My proposed change:

```
@@ -1005,16 +1005,15 @@ static XmlNode sReadXmlNode(XmlParser& p, ParseXmlFilter
*filter, dword style)
   return m;
  }
  if(p.IsText()) {
   m.CreateText(p.ReadText());
   m.Shrink();
   return m;
  }
- p.ReadText(); // skip empty text
- return m;
+ throw XmlError("Unexpected text");
 }

 void ParseXmlFilter::EndTag() {}

 XmlNode ParseXML(XmlParser& p, dword style, ParseXmlFilter *filter)
 {
  XmlNode r;
```

Subject: Re: Bugfix: XmlParser in endless loop
Posted by mirek on Mon, 02 Oct 2023 22:55:05 GMT
View Forum Message <> Reply to Message

zsolt wrote on Mon, 02 October 2023 18:07Try this simple code:
ParseXML("");

This function goes into an endless loop, because at the end of sReadXmlNode() function, the line
p.ReadText(); // skip empty text

doesn't do anything.

My first idea was to change it to
p.Skip();// skip empty text
It seems to be better, but the error message will not be too useful.

My proposed change:

```
@@ -1005,16 +1005,15 @@ static XmlNode sReadXmlNode(XmlParser& p, ParseXmlFilter
*filter, dword style)
   return m;
  }
  if(p.IsText()) {
   m.CreateText(p.ReadText());
   m.Shrink();
   return m;
  }
- p.ReadText(); // skip empty text
- return m;
+ throw XmlError("Unexpected text");
 }

 void ParseXmlFilter::EndTag() {}

 XmlNode ParseXML(XmlParser& p, dword style, ParseXmlFilter *filter)
 {
  XmlNode r;
```

I am not 100% sure about removing ReadText to skip empty text, I think there are cornercases that require that (I bet it is actually a fix).

But this definitely should work:

```
 if(p.ReadText().GetCount() == 0) // skip empty text
  throw XmlError("Unexpected text");
```

(in master now)

Subject: Re: Bugfix: XmlParser in endless loop
Posted by zsolt on Wed, 04 Oct 2023 01:10:36 GMT
View Forum Message <> Reply to Message

Yes, it is a good idea. And seems to be working with my user provided test "XML" as well

Thank you.

Subject: Re: Bugfix: XmlParser in endless loop
Posted by mirek on Mon, 09 Oct 2023 08:32:48 GMT
View Forum Message <> Reply to Message

So that fix proved wrong, breaking autotests, so I have reverted and provided proper fix by moving the check / throw one level up:

```
static XmlNode sReadXmlNode(XmlParser& p, ParseXmlFilter *filter, dword style)
{
 XmlNode m;
 if(p.IsTag()) {
  String tag = p.ReadTag();
  if(!filter || filter->DoTag(tag)) {
   m.CreateTag(tag);
   m.SetAttrs(p.PickAttrs());
   while(!p.End())
    if(!Ignore(p, style)) {
     XmlNode n = sReadXmlNode(p, filter, style);
     if(n.GetType() != XML_DOC) // tag was ignored
      m.Add() = pick(n);
    }
   if(filter)
    filter->EndTag();
  }
  else
   p.SkipEnd();
  return m;
 }
 if(p.IsPI()) {
  m.CreatePI(p.ReadPI());
  return m;
 }
 if(p.IsDecl()) {
  m.CreateDecl(p.ReadDecl());
  return m;
 }
 if(p.IsComment()) {
  m.CreateComment(p.ReadComment());
  return m;
 }
 if(p.IsText()) {
  m.CreateText(p.ReadText());
  m.Shrink();
  return m;
```

```
 }
 p.ReadText(); // skip empty text
 return m;
}

void ParseXmlFilter::EndTag() {}

XmlNode ParseXML(XmlParser& p, dword style, ParseXmlFilter *filter)
{
 XmlNode r;
 while(!p.IsEof())
  if(!Ignore(p, style)) {
   XmlNode n = sReadXmlNode(p, filter, style);
   if(n.GetType() != XML_DOC) // tag was ignored
    r.Add() = pick(n);
   else {
    if(p.IsRelaxed())
     p.Skip();
    else
     throw XmlError("Unexpected text");
   }
  }
 return r;
}
```

Subject: Re: Bugfix: XmlParser in endless loop
Posted by zsolt on Thu, 12 Oct 2023 03:15:42 GMT
View Forum Message <> Reply to Message

Thank you.
Working well.
