

Parsing xml to and replacing specific tags shell script

For the below xml ,I need to replace `<studentStatus>` for `<studentName>CLASSA</studentName>` to `<studentStatus>failed</studentStatus>` .

```
<studentFile>

  <student>
    <studentName>CLASSA</studentName>
    <studentStatus>Success</studentStatus>
    <studentActions>
      <studentAction>
        <studentType>Juniour</studentType>
        <studentStatus>Completed</studentStatus>
        <studentMsg/>
      </studentAction>
      <studentAction>
        <studentType>HighSchool</studentType>
        <studentStatus>Completed</studentStatus>
        <studentMsg/>
      </studentAction>
    </studentActions>
  </student>
  <student>
    <studentName>CLASSB</studentName>
    <studentStatus>Success</studentStatus>
    <studentActions>
      <studentAction>
        <studentType>Senior</studentType>
        <studentStatus>Completed</studentStatus>
      </studentAction>
      <studentAction>
        <studentType>Middle</studentType>
        <studentStatus>Completed</studentStatus>
      </studentAction>
    </studentActions>
  </student>
</studentFile>
```

What I got so far,

```
xmllint -xpath "/studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus" myxml.xml
```

now i got the status of the student as Completed , now this value should be changed to Failed . Only for `<studentType>Juniour</studentType>` . How should I edit the xml inorder to get it as ,

```
<studentFile>

  <student>
    <studentName>CLASSA</studentName>
    <studentStatus>Success</studentStatus>
    <studentActions>
      <studentAction>
        <studentType>Juniour</studentType>
        <studentStatus>Failed</studentStatus>
        <studentMsg/>
      </studentAction>
      <studentAction>
        <studentType>HighSchool</studentType>
        <studentStatus>Completed</studentStatus>
        <studentMsg/>
      </studentAction>
    </studentActions>
  </student>
  <student>
    <studentName>CLASSB</studentName>
    <studentStatus>Success</studentStatus>
    <studentActions>
      <studentAction>
        <studentType>Senior</studentType>
        <studentStatus>Completed</studentStatus>
      </studentAction>
      <studentAction>
        <studentType>Middle</studentType>
        <studentStatus>Completed</studentStatus>
      </studentAction>
    </studentActions>
  </student>
</studentFile>
```

Can this be done using sed. I know there are tools like xsltproc but not sure if this is installed in all nodes in our cluster .

Any help will be appreciated. Thanks in advance!

bash


shell

xml-parsing

xmlstarlet

xmllint

edited May 15 '17 at 4:50

Cyrus


37k

3

29

63

asked May 14 '17 at 18:23

Neethu

1,384

1

10

29

1 "Can this be done using sed." See: stackoverflow.com/a/1732454/3016153 (applies to XML just as well). – michael.hor257k May 14 '17 at 18:29

Do you have access to `xmlstarlet` ? – Cyrus May 14 '17 at 18:30

Pretty sure python comes with a xml module on most systems, could try that. – 123 May 14 '17 at 18:39

Your file is no valid XML. At least one `</student>` is missing. – Cyrus May 14 '17 at 18:40

@Cyrus edited now . I checked for xmlstarlet. `xmlstarlet` -bash: xmlstarlet: command not found – Neethu May 14 '17 at 18:43

Update value with `xmllint` in `file.xml` :

```
xmllint --shell file.xml << EOF
cd /studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus
set failed
save
EOF
```

or without [here document](#):

```
echo -e "cd /studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus\nset failed\nsave" | xmllint --shell file.xml
```

Update: With bash and XML in a variable:


```
xml=$(xmllint --shell <(echo "$xml") << EOF
cd /studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus
set failed
save -
EOF
)
```

or without [here document](#):

```
xml=$(echo -e "cd /studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus\nset failed\nsave -" | xmllint --shell <(echo "$xml"))
```

edited May 15 '17 at 19:18

answered May 14 '17 at 19:47



Cyrus

37k 3 29 63

if my xml is in a variable , how can this command be modified? – Neethu May 15 '17 at 3:43

1 I've updated my answer. – Cyrus May 15 '17 at 4:47

It works . Thanks @Cyrus . – Neethu May 15 '17 at 13:26

Hi @Cyrus , I tried adding `xml=$(xmllint --shell <(echo "$xml") << EOF cd /studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus set failed save - EOF)` to my script but it says: unexpected EOF while looking for matching `)' , should i enclose them in quotes ? – Neethu May 15 '17 at 18:56

I've updated my answer. – Cyrus May 15 '17 at 19:19

In case if ***xmlstarlet*** (a command line toolkit to query/edit/check/transform XML documents) is accessible:

```
xmlstarlet ed -u "//studentAction/studentStatus[preceding-sibling::studentType[1][text() = 'Juniour'] \
and ancestor::student/studentName[text() = 'CLASSA']]\" -v failed students.xml
```

The above will output the initial XML document with needed replacement

The command details:


`ed -u` - edit/update mode

`//studentAction/studentStatus` - xpath expression to select `studentStatus` element which has:

- `preceding-sibling::studentType[1][text() = 'Juniour']` - preceding sibling element `studentType` with value `Juniour`
- `ancestor::student/studentName[text() = 'CLASSA']` - nearest element `studentName` with value `CLASSA`

edited May 14 '17 at 19:26

answered May 14 '17 at 19:20



RomanPerekhrest

45.4k 3 19 48

Not sure the point of this answer, they say in the comments it isn't... – 123 May 14 '17 at 19:25

1 @123, that's pretty sad. Missed that comment while preparing the anser – RomanPerekhrest May 14 '17 at 19:28

The comment they just added makes it seem like it probably will be a useful answer, so not too sad – 123 May 14 '17 at 19:33

1/23/2018


bash - Parsing xml to and replacing specific tags shell script - Stack Overflow

@123, the sad fact is that someone's unable to install such useful tools – RomanPerekhrest May 14 '17 at 19:38

xlmlint , as the name implies, is for parsing and validating XML, not editing it. If you can install `xmlstarlet` on your cluster, you can do the following:

`xmlstarlet ed --update "/studentFile/student[studentName='CLASSA']/studentActions/studentAction[studentType='Juniour']/studentStatus" --value "Failed" *file*`

answered May 14 '17 at 19:20



Brian from QuantRocket

1,952 12 13