

Displaying Records from Flat Files Using awk



Andrew Mallett

@theurbanpenguin | www.theurbanpenguin.com

XML and awk

Display tagged data

Virtual Hosts

Catalogs move to XML

Display tagged data

Normally an awk record will match lines in the file. Using a little imagination we can use the RS variable to match a record to a blank line: **RS="\n\n"**. We can then search for the correct record and all associated lines will be shown.

```
<VirtualHost *:80>
DocumentRoot /www/example
ServerName www.example.org
# Other directives here
</VirtualHost>
```

```
$ cat xml.awk
BEGIN { RS="\n\n" }
$0 ~ search { print }
$ awk -f xml.awk \
search=example vh.conf
```

The idea came from displaying Apache Virtual Host detail

Imagine many virtual host definitions in the one file, each separated by a blank line

We could extract the record for example.org using awk

Use sed to remove all blank lines and then add blank lines between records if needed



Demo Time: Displaying Virtual Host data

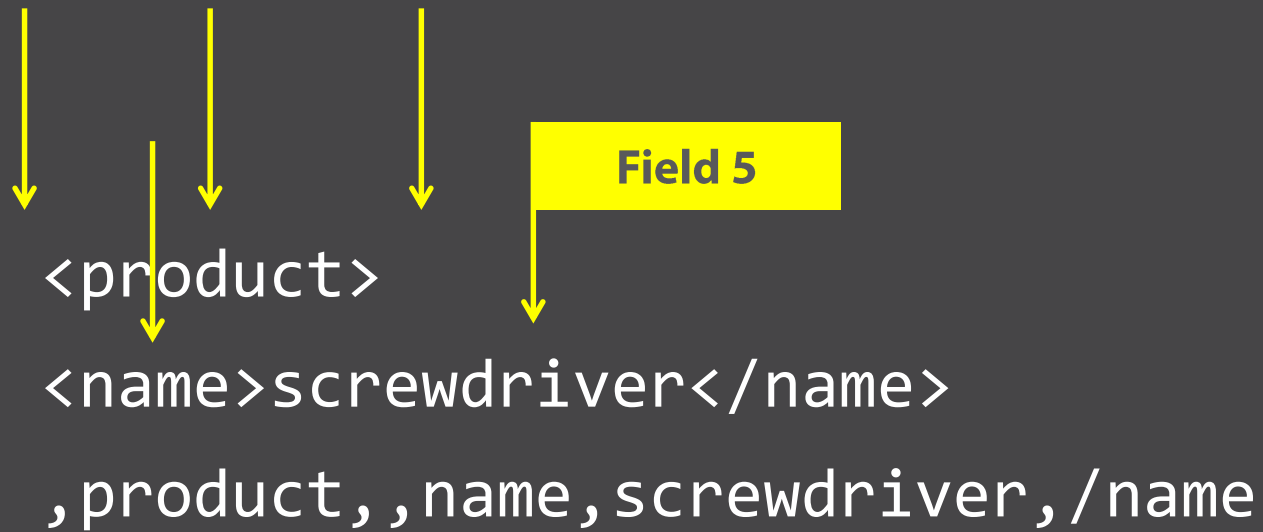
New XML Product Data

- We saw before the using **grep -A** we could display the correct amount of catalog data
- DIY'r'US have now moved to XML based data files
- This makes searching the data much easier for Danny and his team





Demo Time: DIY'r'Us Catalogs



`<product>`
`<name>screwdriver</name>`
`,product,,name,screwdriver,/name`

Break XML Data Further

Using square brackets [] we can define multiple delimiters

'[><]' Define both the opening and closing angle as delimiters

Field numbering: you need to take care, \$5 is screwdriver in this case

Summary



- Reporting from flat files
- Parsing search strings from the CLI
- Display Virtual Host data
- Implementing a new catalog search
- Multiple field delimiters
- Next up.... Analyze logs