libferris

Dr Martin
Talk overview
Why

declare variable $docurl := "customers.db"
declare variable $customerid := "131";

<resultdata>
{
  for $c in ferris-doc( $docurl )
  /customers/*[@id=$customerid]
  return <person cid="{ $c/@id }"
    surname="{ $c/@familyname }" />
}
</resultdata>
Extended Attribute (EA) interface for all

```bash
$ ferrisls -l \n  --show-ea=name,size,width a.png
a.png 44624  488

$ date >df
$ fls --show-ea=name,content df
df Thu Apr 30 21:37:36 EST 2009

$ fls --show-ea=name,md5 df
df c473631d731726eed41ec8f1de2d0178
```
EA is first rate

$ fls --ferris-filter='(name=~.*db)'
hash.db  play.db

$ fls -lh --show-ea=name,size \
   --ferris-sort=':!#:size'
play.db  8192
hash.db   12288

$ fls -lh --ferris-sort=\n   '(name-extension)(:#!:size)'
Cute..but - Why?

IOZone num diff
Files are FS too!

$ ferrisls -lh \n--show-ea=name,kb,reclen,...
  read,write,reread \n  iozone-64kb-chunks-ext3.iozone

0  64  4 4207076 357529  4274062
1  64  8 3588436 394870  3363612
10 128 128 3867787 444501  3445772
100 131072 128 3602456 437633  3224094
FS are tables too!

.load libferrissqlitevtable.so
create virtual table fs using libferris(
  './iozone-64kb-chunks-ext3.iozone',
  '',
  name text, size int, md5 text,
  kb int, reclen int, write int, rewrite int, read int, reread int,
  ranread int, ranwrite int, bkwdread int, ...
  fwrite int, frewrite int, fread int, freread int,
  content text);

select kb,reclen,read,write from fs limit 3;
FS are tables too!

$ sqlite3 -init iozone-select1.sql

... 
64|4|4207076|357529
64|8|3588436|394870
128|128|3867787|444501
rm -f /tmp/tt.db;
touch /tmp/tt.db
export TSCHEMA="name text, kb int, reclen int…"
sqlite3 /tmp/tt.db <<EOF

create virtual table a using libferris(
   '$FILENAME1', 'recursive=false,', $TSHEMA );

create virtual table b using libferris(
   '$FILENAME2', 'recursive=false,', $TSHEMA );

EOF
IOZone diff...

.separator ''
select a.kb, a.reclen,
   a.write - b.write ,
   a.rewrite - b.rewrite ,
   a.read – b.read ...
from a, b
where a.kb = b.kb
   and a.reclen = b.reclen;

EOF
IOZone diff...

$ ./iozone-diff.sh \
iozone-64kb-chunks-ext3.iozone \
iozone-64kb-chunks-ext3align.iozone

...

64 4 3769 324697 1314631 1611163 1159812 401119 287352 -146933 0 -101911 -99518 1293718 1177091
64 8 -32027 -112288 -1400542 -2026041 -1542059 -416132 0 109205 -2620671 -157787 -268980 -1070993 -1507633
128 128 23362 -319365 -114766 -1488444 -1174766 -342633 239031 5119 102434 5193 -255690 1888844 1654505
131072 128 -57525 25540 93447 -15465 -70771 -23478 129855 214708 479410 38341 146806 -198559 -156104
Setup...

$ cat ./setup-iozone.sh
#!/bin/bash

for fn in *.iozone
do
echo -n recordfile | \
ferris-redirect -T --ea=ferris-type $fn
  echo -n iozone | \
ferris-redirect -T -x --ea=ferris-recordfile-metadata-url $fn
done
Extensible?

$ ferrisls -0 ~/.ferris/recordfile-types.db/iozone
...
<context
ferris-recordfile-field-names="junk, kb, reclen, write, rewrite, read, reread, ...
" ferris-recordfile-fs="[ \t]+"
ferris-recordfile-grep="^[ ]+[0-9]+[ ]+[0-9]+"
name="iozone"/>
ferris-recordfile-rs
   RS from awk, rec sep rex

ferris-recordfile-fsplit
   Regex (capture)s

ferris-recordfile-grep
   Negate lines

ferris-recordfile-attribute-that-names-file

ferris-recordfile-field-timeparsers
   date=%d/%b/%Y:%H:%M:%S %z|other=%...
The Forest

Web Photos
Web Videos
Social Nets

These are All Filesystems

XML  ISAM

Firefox
amaroK
PostgreSQL powered
mythTV
(C)Lucene
PostgreSQL
RDF
Custom inverted
Xapian, **Federations**
Beagle, Strigi
Web
Boost memory map,
with URL regex prefilter
findexadd, feaindexadd
findexquery, feaindexquery
Search is a filesystem....

$ echo 'drop database findexaudio' | psql
$ fcreate `pwd` \
   --create-type=eaindexpostgresql

$ find /Audio  | feaindexadd -P `pwd` -1
$ feaindexcompact -P `pwd`
$ feaindexquery -P `pwd` '(url=~River)'
file:///Audio/willie_nelson/.../Whiskey River.flac
...

$ fls --ea-index-path=`pwd` \\
'eaq://(url=~Dinah)'
file:///Audio/bing_crosby/... ...Dinah.flac
Web is an FS

ferriscp dpic.jpg 23hq://monkeyiq/upload

gfcp test.avi vimeo://upload/

echo whatever \
   | ferris-redirect facebook://status

   gfcp ~/fuse/my-phone/latest.mp4 youtube://upload/
Web Photos as FS

$ gfcp small-test-image.jpg flickr://me/upload
$ fls -lh --show-ea=name,size,webphoto-service,webphoto-photo-id small-test-image.jpg
small-test-image.jpg  683  flickr3857255129

$ fls -0 flickr://me
by-id
contacts
defavs
not-in-any-photosets
photosets
recent
upload
Web Photos as FS

$ fcat flickr://me/by-id/3831809948 >| /tmp/out.jpg
$ feh /tmp/out.jpg

$ fls -0 flickr://me/not-in-any-photosets/443571910_comments
7213593- ... monkeyiq 09 Aug 4 11:48 four more
7213593- ... monkeyiq 09 Aug 4 13:06
7213593- ... monkeyiq 09 Aug 4 13:13 test9
7213593- ... monkeyiq 09 Jul 30 12:00 water dragon

$ echo what you looking at | ferris-redirect -T \ 
    flickr://me/not-in-any-photosets/443571910_comments/new

$ echo test10 | ferris-redirect -T \ 
    flickr://me/not-in-any-photosets/...\ 
    443571910_comments/7213...571910-721576218147771773
Web is an FS

$ ftouch google://docs/appender.txt
$ date | ferris-redirect –trunc \ google://docs/appender.txt
$ fcat google://docs/appender.txt
Fri Aug 21 11:54:25 EST 2009
$ date | ferris-redirect –append \ google://docs/appender.txt
$ fcat google://docs/appender.txt
Web is an FS

$ export SPREAD=\google://spreadsheets/smalltest1/Sheet1
$ echo 5  | ferris-redirect --ea b $SPREAD/9
$ echo 23 | ferris-redirect --ea c $SPREAD/9

$ echo -n "=sum(B9,C9)"
 | ferris-redirect --ea d $SPREAD/9
$ fcat -a d $SPREAD/9
28
Google?

Whatever
SQLite on Google

.load libferrissqlitevtable.so
create virtual table fs using libferris(
  'google://spreadsheets/smalltest1/Sheet1',
  '',
  name text,
  a text, b int, c int, d int,
  e text, f text, size int
);
SQLite on Google

.headers on
select a,b,c,d,b+c as result
from fs
where a = 'foo';

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>result</th>
</tr>
</thead>
<tbody>
<tr>
<td>foo</td>
<td>21</td>
<td>25</td>
<td>46</td>
<td>46</td>
</tr>
</tbody>
</table>
select max(b),max(c),
    avg(d) as mean,
    sum(d) as total
from fs
where not (b = "" or c = ")
group by b,c;

max(b)|max(c)|mean|total
LA|barry|77.0|77
LA|john|12.0|12
LA|peter|54.5|109
NY|barry|77.66666666666667|233
NY|peter|56.0|56
Tricksey FSeses

$ fls -0 pg://localhost/play/cars
bmw  cz 100 bmw  count-name-place
bmw  de 1000 bmw  count-name-place

$ fls --xml
pg://localhost/play/cars
...
<context count="100"
    name="bmw" place=""

$ fls --rdf pg://localhost/play/cars
<pg://localhost/play/cars/10000-skoda-czech%20rep.>
<http://...rdf/ferris-attr/primary-key>
"count-name-place"^^<http://w..chema#string> .
...

$ fcat 'pg://localhost/play/atan(0.5)'
0.4636476090000806
$ echo 1 | ferris-redirect \namarok://control/toggle-play-pause

$ fcat 'dbus://localhost/session/
org.kde.amarok/Player/
org.freedesktopMediaPlayer/
PlayPause()'

$ fls -0 amarok://playlist
0 3:27 Bing Crosby ...
The Classic Crooners Collection   Muddy Water ...

$ ferriscp "amarok://playlist/" .
```bash
$ fcat xwin://localhost/clipboard/1 about

$ ferriscp emacs://localhost/ben/sample xwin://localhost/clipboard/0

$ fcreate . --create-type=db4 --rdn=new.db
$ echo hi there | ferris-redirect -T new.db/embedded-file
$ db_dump -p new.db

... embedded-file
hi there
```
$ fedit new.db/embedded-file
dd
boo
ZZ
$ db_dump -p new.db embedded-file
  boo
  \0a
$ fcat new.db/embedded-file
  boo
$
$ xmllint --format new.xml
<?xml version="1.0"?>
<rootelement>
  <child1>hi there</child1>
  <child2>something else</child2>
</rootelement>

$ fcat new.xml/rootelement/child1
hi there

$ fedit new.xml/rootelement/child1

and again
xsltfs://

Forward
Reverse

xsltfs://context/file/tmp/foo?
stYLESHEET=a.xsl&
reverse-stylesheet=b.xsl
$ fls -0 --show-headings \ 
pg://localhost/play/msgs

id msgfoo namelprimary-key
1 msg1value foo1value 1 id
ferris-filesystem-to-xsltfs-sheets \ 
  --plugin excel2003 --fuse pg-play-msgs \ 
  pg://localhost/play/msgs

$ cat ~/ferrisfuse/mount-pg-play-msgs.sh
#!/bin/bash
ferrisfs -u 'single-file-filesystem:xsltfs? implicit-sheets=
3://context/pg/localhost/play/msgs.xml'
-F '.*.xml$'  pg-play-msgs

ooffice ~/ferrisfuse/pg-play-msgs/msgs.xml
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>name</td>
<td>id</td>
<td>msg</td>
<td>foo</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td>1 msg1value</td>
<td>foo1value</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td></td>
<td>2 msg2value</td>
<td>foo2val</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
<td>5 newXXX1</td>
<td>new2</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sheet1
What is Real?

user-overlay-links.xml

  <user-overlay-links>
    <link-by-regex>
      <vlink>
        <match>^gphoto:[/]+$</match>
        <target>Canon PowerShot G7 (PTP mode)</target>
        <link-name>G7</link-name>
      </vlink>
    </link-by-regex>
  </user-overlay-links>
More?

Linux Journal
Index and Search

LDN / linux.com
SQLite series
Apache KML series

xml.com
XQuery

lwn.net
Intro
Consulting

Customization
Optimization
Custom builds
Regression testing

... lets talk
Questions .... Tomatoes?