

open



USE



IMPROVE



EVANGELIZE

Using Dtrace for Gnome Performance Analysis

Krishnan Parthasarathi
Desktop Sustaining Engineer
Sun Microsystems, India

開
放
的
열린
مفتوح
libre
मुक्त
ಮುಕ್ತ
livre
libero
ముక్త
开放的
açık
open
nyílt
⋮⋮⋮
πππ
オープン
livre
ανοικτό
offen
otevřený
öppen
открытый
வெளிப்படை



Agenda

Dtrace Intro

Problem Statement

Existing approaches

What Dtrace can offer?

Pid provider

USDT probes – How to

Real World USDT probes eg.

Where are the probes?

Future work



DTrace Intro

Dynamic tracing framework

'Sensors'/Probes are placed all over Solaris system

These 'sensors' are programmable – D Language

Can gather information on probe firing

See <http://wikis.sun.com/display/DTrace/Introduction>



Dtrace probes

No of probes in system

```
krish@trantor:~$ pfexec dtrace -l | wc -l  
64974
```

No of system calls

```
krish@trantor:~$ pfexec dtrace -n 'syscall:::entry @[probefunc] = count()'  
dtrace: description 'syscall:::entry ' matched 236 probes
```

```
^C
```

fstat	1
mmap	1
recv	1
schedctl	1

```
[...]
```



Dtrace Oneliners

Syscall count by program,

```
dtrace -n 'syscall:::entry { @num[execname] = count(); }'
```

```
dtrace: description 'syscall:::entry ' matched 236 probes
```

```
^C
```

updatemanagernot	2
ntpd	5
tracker-indexer	5
trackerd	6
[....]	
Xorg	557
soffice.bin	719
dtrace	805
firefox-bin	1922



Dtrace onliners contd.

Syscall count by syscall,

```
dtrace -n 'syscall:::entry { @num[probefunc] = count(); }'
```

```
dtrace: description 'syscall:::entry ' matched 236 probes
```

```
^C
```

fstat	1
mmap	1
schedctl	1
sigpending	1
[..]	
read	453
pollsys	784
ioctl	1324

See http://www.brendangregg.com/DTrace/dtrace_oneliners.txt



Problem Statement

To collect performance metrics of Gnome applications in a non-intrusive manner

Desirable properties

- Non-intrusive

- In production



Existing approaches

Adding debug statements of varying importance -
debug levels

- violates the non-intrusiveness constraint

- Production and debug binaries are different

In Linux, Using systemtap to profile gnome -
<http://live.gnome.org/GnomePerformance/SystemTap>
time spent in syscalls can be recorded

- some tracing features like ustack are developer only

- lacks probes specific to apps



What Dtrace can offer?

Pid provider to get surface details

USDT probes exposes app specific metrics

USDT probes can be present in production code

Zero disabled probe effect – non-intrusive



Pid provider

Pid provider can trace any instruction as specified by an absolute address or function offset.

Pid provider has no probe effect when probes are not enabled.

Probes only induce probe effect on those processes that are traced.



Using pid provider..

```
krish@trantor:~/D-scripts# dtrace -n 'pid5596::g_local_file_read*:entry {}'  
dtrace: description 'pid5596::g_local_file_read*:entry ' matched 1 probe  
CPU   ID          FUNCTION:NAME  
  1   18      g_local_file_read:entry  
  0   18      g_local_file_read:entry  
  0   18      g_local_file_read:entry  
  0   18      g_local_file_read:entry  
^C
```

The above oneliner was run against the nautilus process.



Limitation of pid provider

To retrieve objects other than primitive data types is non-trivial

Tracing places other than entry/return is difficult



USDT probes

User Level Statically Defined Tracing
Framework for placing custom probes in
application code

Can be used in,

- Performance analysis

- Debugging



USDT probes – How To

helloworld.c – simple C program with dtrace probe

```
#include <stdio.h>
#include <sys/sdt.h>
int main() {
    int i;
    for(i=0;i<5;i++) {
        DTRACE_PROBE(world, sample);
        printf("hello\n");
    }
    return 0;
}
```

world.d – describes the probes published by the provider

```
provider world {
    probe sample();
};
```



USDT probes – How To

Steps to compile app with probes

```
cc -c helloworld.c
```

```
dtrace -32 -G -s world.d helloworld.o
```

```
cc -o helloworld world.o helloworld.o
```



USDT probes – How To

Sample output

```
krish@trantor:/tmp$ pfexec dtrace -n 'world$target:::sample' -c ./helloworld
```

```
dtrace: description 'world$target:::sample' matched 1 probe
```

```
hello
```

```
hello
```

```
[..]
```

```
dtrace: pid 1195 has exited
```

CPU	ID	FUNCTION:NAME
1	64974	main:sample
1	64974	main:sample
1	64974	main:sample
1	64974	main:sample
1	64974	main:sample



Real World USDT eg.

Apache

```
int apache_receive_request(request_rec *r)
{
    DTRACE_PROBE1(apache,
                  receive__request,
                  r);
    return DECLINED;
}
```



Real World USDT eg.

Mysql

```
#include <mysqlprovider.h>
mysql_parse {
...
bool mysql_execute_command(THD *thd)
{
    MYSQL_QUERY_EXECUTE_START(thd->thread_id);
```

where,

```
#define MYSQL_QUERY_EXECUTE_FINISH(arg0) \
    __dtrace_mysql__query_execute__finish(arg0)
```



Where are the probes?

Probes are placed in core gnome libs

Glib

Provides the core object system used in GNOME
Utility functions for strings and common
data structures.

Gio

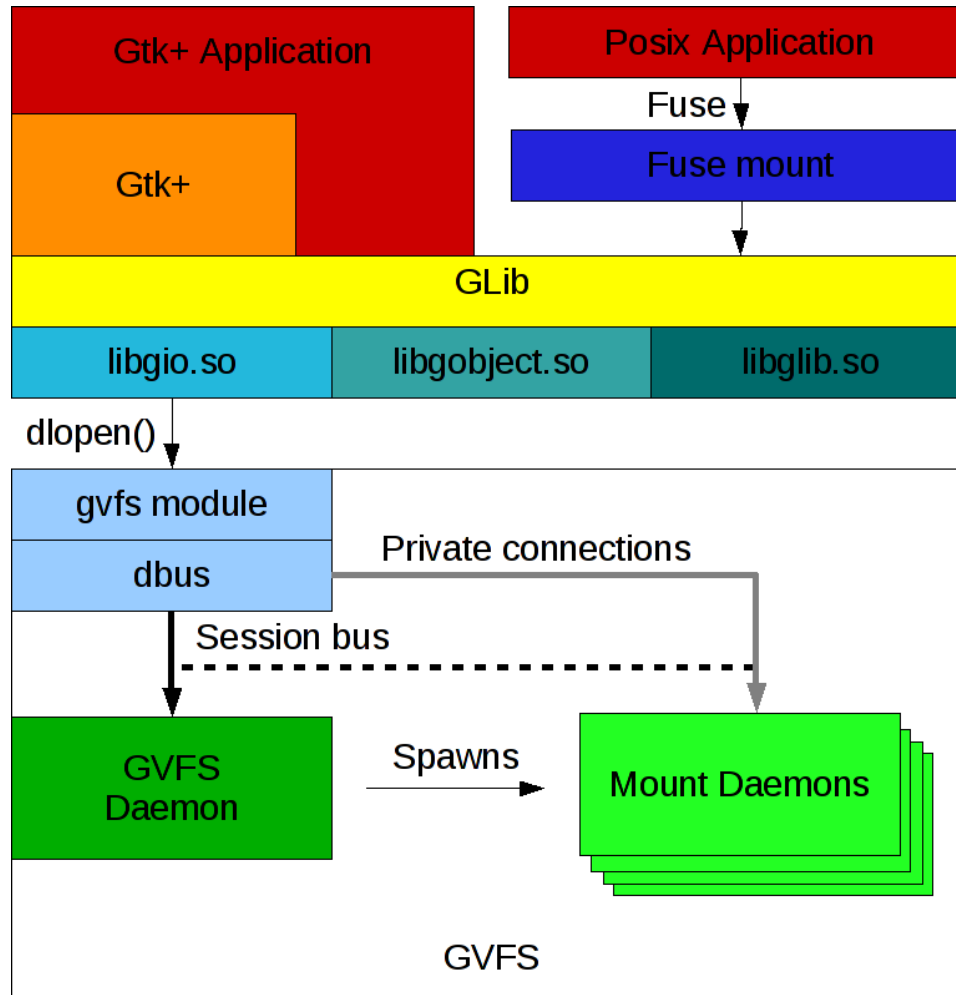
Easy-to-use VFS API
Provides higher-level, document-centric interfaces

Gtk

Library for creating graphical user interfaces

Gnome stack

Component view



Probes in Glib

Probe matrix

Probe name	Description	Arg0	Arg1
g-malloc	Fires on entering g_malloc()	No. of bytes to be allocated	
g-free	Fires on entering g_free()	Address of the memory to be freed	
gslice-alloc	Fires on entering g_slice_alloc()	No. of bytes to be allocated	
gslice-free	Fires on entering g_slice_free1()	No. of bytes to be freed	Address of the memory to be freed



Custom script – Mean Life Time

```
mean_life_time.d
```

```
#!/usr/sbin/dtrace -s
```

```
/*  
This script prints the average time for which a particular memory address, allocated through the  
g_malloc(), is alive  
*/
```

```
long long addr_to_ts[long long]; /* stores the timestamp at which 'addr' was allocated */
```

```
pid$$1::g_malloc:return
```

```
{  
addr_to_ts[arg1] = vtimestamp;
```

```
}  
glib_prov$$1::g-free
```

```
/addr_to_ts[arg0] > 0/
```

```
{  
@[execname] = quantize(vtimestamp - addr_to_ts[arg0]);
```

```
addr_to_ts[arg0] = 0;
```

```
}
```

Sample Output – MLT

MLT for Nautilus objects

```
krish@trantor:~/D-scripts# ./mean_life_time.d `pgrep nau`
dtrace: script './mean_life_time.d' matched 2 probes
^C
```

```
nautilus
```

value	Distribution	count
1024		0
2048	@@@@@@@@@@@@@@@@@@@@@@	28
4096	@@@@@@@@@@	14
8192		0
16384		0
32768		0
65536		0
131072		0
262144		0
524288	@@@@@@@@	12
1048576	@@	3
2097152		0

```
krish@trantor:~/D-scripts# █
```



Custom script – Distribution of sizes

```
memory_distribution.d
```

```
#!/usr/sbin/dtrace -s
```

```
/*
```

```
This script prints the distribution of sizes of objects allocated by an application.
```

```
*/
```

```
glib_prov$$$1:::g-malloc
```

```
{
```

```
    @[execname] = quantize(arg0);
```

```
}
```


Sample output – Distribution of sizes

Distribution of sizes of Nautilus objects

```
krish@trantor:~/D-scripts# ./memory_distribution.d `pgrep nau`
dtrace: script './memory_distribution.d' matched 1 probe
^C

nautilus
value  ----- Distribution ----- count
-1    |                                     0
 0    | @@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ 255
 1    |                                     0
 2    |                                     0
 4    |                                     0
 8    | @@@@@@                               68
16    | @@@@@@@@                             102
32    | @@@@@@@@@@@@                         153
64    |                                     0
128   | @                                       17
256   |                                     0

krish@trantor:~/D-scripts# █
```

Accidental finding

Zero size memory requests

```
libglib-2.0.so.0.2105.0`g_malloc+0x1b  
libgobject-2.0.so.0.2105.0`g_object_newv+0xc5  
libgobject-2.0.so.0.2105.0`g_object_new_valist+0x3f3  
libgobject-2.0.so.0.2105.0`g_object_new+0x6a  
libgdk-x11-2.0.so.0.1709.0`_gdk_pixmap_new+0x141  
libgdk-x11-2.0.so.0.1709.0`gdk_pixmap_new+0x3c  
libgdk-x11-2.0.so.0.1709.0`gdk_x11_draw_pixbuf+0x983  
libgdk-x11-2.0.so.0.1709.0`gdk_window_draw_pixbuf+0xc4  
libgdk-x11-2.0.so.0.1709.0`gdk_draw_pixbuf+0x10a  
nautilus`nautilus_icon_canvas_item_draw+0x698  
nautilus`eel_canvas_group_draw+0x167  
nautilus`eel_canvas_expose+0x181  
nautilus`expose_event+0x38
```

All the stacks indicate drawing of icons.



Probes in Gtk

Probe matrix

Probe name	Description
file-chooser-dialog-init-begin	Fires on entering <code>gtk_file_chooser_dialog_init</code>
file-chooser-dialog-init-end	Fires just before leaving <code>gtk_file_chooser_dialog_init</code>



Custom script – file-chooser-dialog init

```
#!/usr/sbin/dtrace -s
/* This script prints the time taken to instantiate a file-chooser-dialog object */
gtk_prov$$$1:::file-chooser-dialog-init-begin
{
self->ts = vtimestamp;
}
gtk_prov$$$1:::file-chooser-dialog-init-begin
{
}
gtk_prov$$$1:::file-chooser-dialog-init-end
/self->ts != 0/
{
@[execname] = avg(vtimestamp - self->ts);
}
gtk_prov$$$1:::file-chooser-dialog-init-end
/self->ts != 0/
{
}
```



Sample output – file-chooser-dialog init

```
krish@trantor:~/D-scripts# ./gtk-file-chooser.d 1450
```

```
dtrace: script './gtk-file-chooser.d' matched 4 probes
```

```
CPU   ID           FUNCTION:NAME
  0  65095  gtk_file_chooser_dialog_init:file-chooser-dialog-init-begin
  0  65096  gtk_file_chooser_dialog_init:file-chooser-dialog-init-end
  1  65095  gtk_file_chooser_dialog_init:file-chooser-dialog-init-begin
  1  65096  gtk_file_chooser_dialog_init:file-chooser-dialog-init-end
```

```
^C
```

```
gedit                61097
```

Average time spent in initializing file-chooser-dialog object is 61097 ns



Probes in Gio

Probe Matrix

Probe name	Description
localfile-open-begin	Fires on entering <code>g_local_file_open_readwrite()</code>
localfile-open-end	Fires before leaving <code>g_local_file_open_readwrite()</code>
localfile-read-begin	Fires on entering <code>g_local_file_read()</code>
localfile-read-end	Fires before <code>g_local_file_read()</code>

Probes in Gvfs

Probe Matrix

Probe Name	Description
ftp-open-begin	Fires on entering do_open_for_read() implementation of the ftp backend.
ftp-open-end	Fires before leaving do_open_for_read() implementation of the ftp backend.
ftp-read-begin	Fires on entering do_read() implementation of the ftp backend.
ftp-read-end	Fires before leaving do_read() implementation of the ftp backend.



Measure File I/O Performance

Probes in Gio measure the localfile i/o latencies

Probes in Gvfs measure file i/o latencies associated with respective backends - ftp, http, webdav, etc.

Using USDT probes added..

```
krish@trantor:~/D-scripts# dtrace -n 'gio_prov5596:::localfile-read-begin {trace(copyinstr(arg0));
ustack()}'
```

```
dtrace: description 'gio_prov5596:::localfile-read-begin ' matched 1 probe
```

```
CPU  ID          FUNCTION:NAME
```

```
0 64996 g_local_file_read:localfile-read-begin
/export/home/krish/.thumbnails/normal/83a3725f4a3df0fc493e4472fe9adf80.png
libgio-2.0.so.0.2105.0`g_local_file_read+0x34
libgio-2.0.so.0.2105.0`open_read_async_thread+0x5e
libgio-2.0.so.0.2105.0`run_in_thread+0x48
libgio-2.0.so.0.2105.0`io_job_thread+0x38
libglib-2.0.so.0.2105.0`g_thread_pool_thread_proxy+0xcb
libglib-2.0.so.0.2105.0`g_thread_create_proxy+0x131
libc.so.1`_thrp_setup+0x7e
libc.so.1`_lwp_start
```

Filename is 'extracted' from Gfile object



Dtrace Patch for Glib

```
krish@trantor:~/packages/BUILD$ diff -N SUNWglib2-2.21.6{,.dtrace}/i386/glib-2.21.6/glib/glib_probes.d
```

```
0a1,6
```

```
> provider glib_prov {  
>     probe g__malloc(int);  
>     probe g__free(void*);  
>     probe gslice__alloc(int);  
>     probe gslice__free(int, void*);  
> };
```

```
krish@trantor:~/packages/BUILD$ diff SUNWglib2-2.21.6{,.dtrace}/i386/glib-2.21.6/glib/gmem.c
```

```
32a33
```

```
> #include <sys/sdt.h>  
  
124a126  
>     DTRACE_PROBE1(glib_prov, g__malloc, n_bytes);
```

```
186a189
```

```
>     DTRACE_PROBE1(glib_prov, g__free, (void*) mem);
```

```
krish@trantor:~/packages/BUILD$ diff SUNWglib2-2.21.6{,.dtrace}/i386/glib-2.21.6/glib/gslice.c
```

```
29a30
```

```
> #include <sys/sdt.h>
```

```
802a804,806
```

```
>  
>     DTRACE_PROBE1(glib_prov, gslice__alloc, mem_size);  
>
```

```
852a857
```

```
>     DTRACE_PROBE2(glib_prov, gslice__free, mem_size, (void*) mem_block);
```



Dtrace Patch for Gio

```
krish@trantor:~/packages/BUILD$ diff -N SUNWglib2-2.21.6{,.dtrace}/i386/glib-2.21.6/gio/gio_probes.d
```

```
0a1,7
```

```
> provider gio_prov {  
>     probe localfile__open__begin(char*);  
>     probe localfile__open__end();  
>     probe localfile__read__begin(char*);  
>     probe localfile__read__end();  
> };  
>
```

```
krish@trantor:~/packages/BUILD$ diff -N SUNWglib2-2.21.6{,.dtrace}/i386/glib-2.21.6/gio/glocalfile.c
```

```
24a25
```

```
> #include <sys/sdt.h>  
1296a1298  
>     DTRACE_PROBE1(gio_prov, localfile__read__begin, local->filename);  
1299d1300  
<  
1320a1322  
>     DTRACE_PROBE(gio_prov, localfile__read__end);  
1365a1368  
>     DTRACE_PROBE1(gio_prov, localfile__open__begin, G_LOCAL_FILE (file)->filename);  
1374a1378  
>     DTRACE_PROBE(gio_prov, localfile__open__end);
```



Dtrace Patch for Gtk

```
krish@trantor:~/packages/BUILD$ diff
  SUNWgtk2-2.17.11/i386/gtk+-2.17.11/gtk/gtkfilechooserdialog.c{.orig,}
34a35
> #include <sys/sdt.h>
84a86
>     DTRACE_PROBE(gtk_prov, file__chooser__dialog__init__begin);
104a107
>     DTRACE_PROBE(gtk_prov, file__chooser__dialog__init__end);
krish@trantor:~/packages/BUILD$ diff SUNWgtk2-2.17.11/i386/gtk+-2.17.11/gtk/gtk_probes.d{.orig,}
0a1,6
> provider gtk_prov {
>     probe label__class__init__begin();
>     probe file__chooser__dialog__init__begin();
>     probe label__class__init__end();
>     probe file__chooser__dialog__init__end();
> };
```



Dtrace Patch for Gvfs

```
krish@trantor:~/packages/BUILD$ diff SUNWgnome-gvfs-2.27.91/gvfs-1.3.6/gvfs_probes.d{.orig,}
```

```
0a1,6
```

```
> provider gvfs_prov {
>     probe ftp__read__begin(char*);
>     probe ftp__read__end();
>     probe ftp__open__begin(char*);
>     probe ftp__open__end();
> };
```

```
krish@trantor:~/packages/BUILD$ diff SUNWgnome-gvfs-2.27.91/gvfs-1.3.6/daemon/gvfsbackendftp.c{.orig,}
```

```
26a27
```

```
> #include <sys/sdt.h>
```

```
640a642
```

```
>     DTRACE_PROBE1(gvfs_prov, ftp__open__begin, filename);
```

```
670a673
```

```
>     DTRACE_PROBE(gvfs_prov, ftp__open__end);
```

```
695a699
```

```
>     DTRACE_PROBE1(gvfs_prov, ftp__read__begin, G_VFS_BACKEND_FTP (backend)->host_display_name );
```

```
712a717
```

```
>     DTRACE_PROBE(gvfs_prov, ftp__read__end);
```



Future Work

To ship Gnome on OpenSolaris, with Dtrace enabled by default

D-Bus, the IPC mechanism Gnome uses is also a good candidate for adding probes

open



USE



IMPROVE



EVANGELIZE

Thank you!

Krishnan Parthasarathi
Desktop Sustaining Engineer
krishnan.parthasarathi@sun.com

“open” artwork and icons by chandan:
<http://blogs.sun.com/chandan>

開
放
的
열린
مفتوح
libre
मुक्त
ಮುಕ್ತ
livre
libero
ముక్త
开放的
açık
open
nyílt
•••••
πικρ
オープン
livre
ανοικτό
offen
otevřený
öppen
ОТКРЫТЫЙ
வெளிப்படை