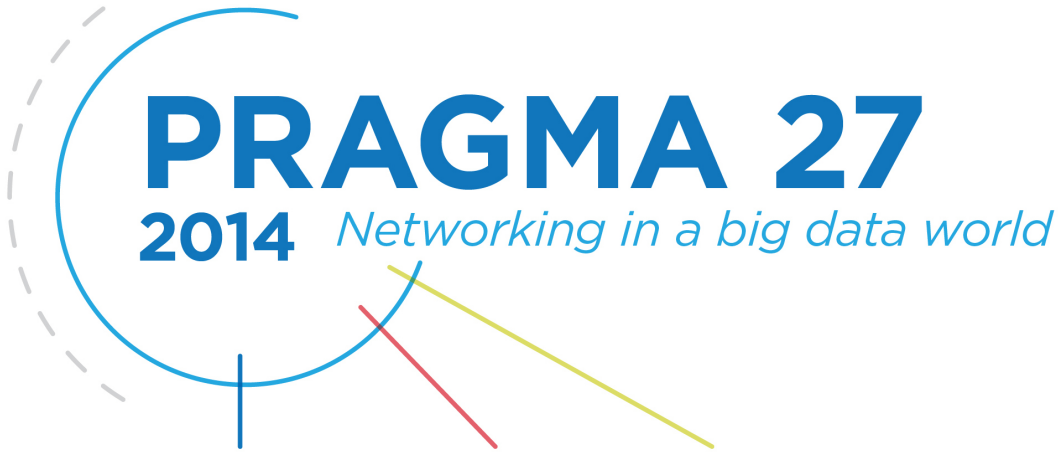
Fingerprint

Luca Clementi, Philip Papadopoulos



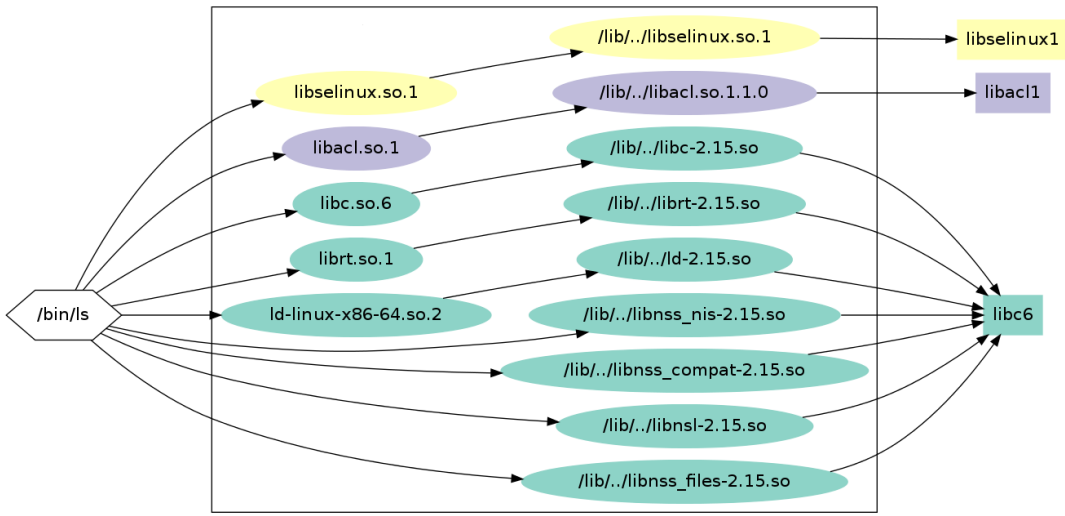
The problem

- Modern scientific research is dependent on reliable software environments
- Every applications relay on several shared libraries and external files

```
# mpirun
mpirun: error while loading shared libraries: libxml2.so.2:
      cannot open shared object file: No such file or directory
```

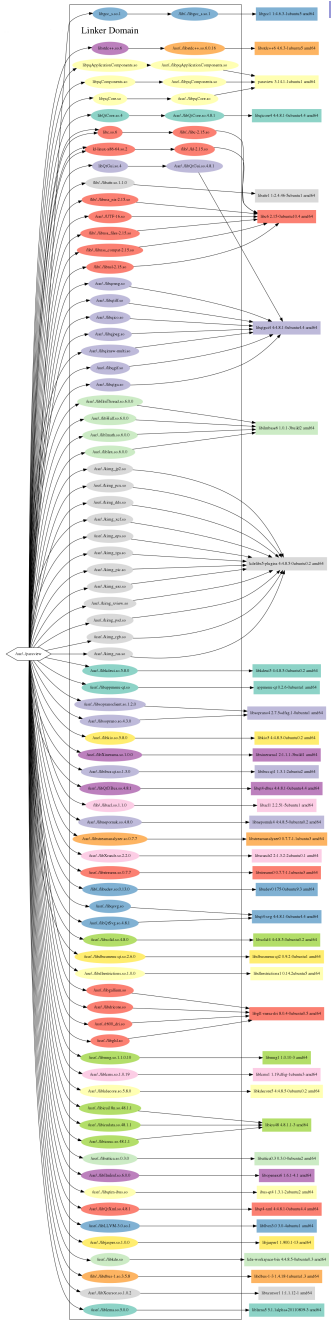
ls -l

- 9 files
- 12 shared libraries
  - 4 dynamically loaded
- 4 OS packages



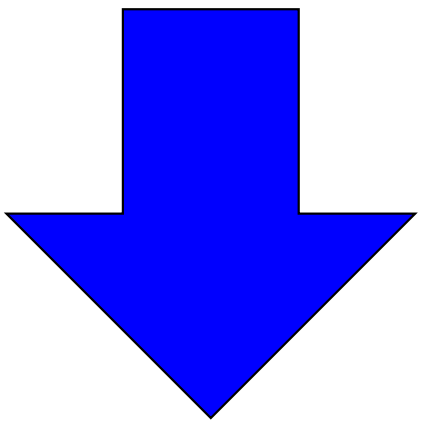
Paraview

- 139 files
- 382 shared libraries
  - 68 dynamically loaded
- 122 OS packages



Our approach

Fingerprinting application dependencies



Command line tool

What is it...

Fingerprint is a software tool which can do:

- Dependencies **discovery**
- Dependencies **display**
- **Verify** dependencies
- **Compose** a new system based on a list of dependencies

Static Fingerprinting

Collect requirements from a list of binary files.

- Parse ELF (Executable Linkable Format) headers.

Static Fingerprinting

```
# fingerprint -c /bin/ls
```

Output written to a file: Swirl

But...

Linux supports dynamic loading of shared libraries...

- Explicit dependency: a dependency which is declared in the binary header
- Implicit dependency: a dependency which is loaded dynamically at runtime (`dlopen()`)

Dynamic Fingerprinting

Dynamic Fingerprinting

```
# fingerprint -c -x "ls -l"
```

- run the process and trace it
- detect each invocation of `open()`
- read `/proc/ID/mmaps`

Swirl File

The `Swirl` file stores info regarding all the traced files.

For each file it saves its:

- Full path and symbolic links
- Explicit dependencies (soname)
- Implicit dependencies (pointers to SwirlFile)
- Major and minor version of symbols provided
- Opened files (pointers to SwirlFile)
- Environment variables
- Hash, architecture, package, rpath, executable

Dependencies Display

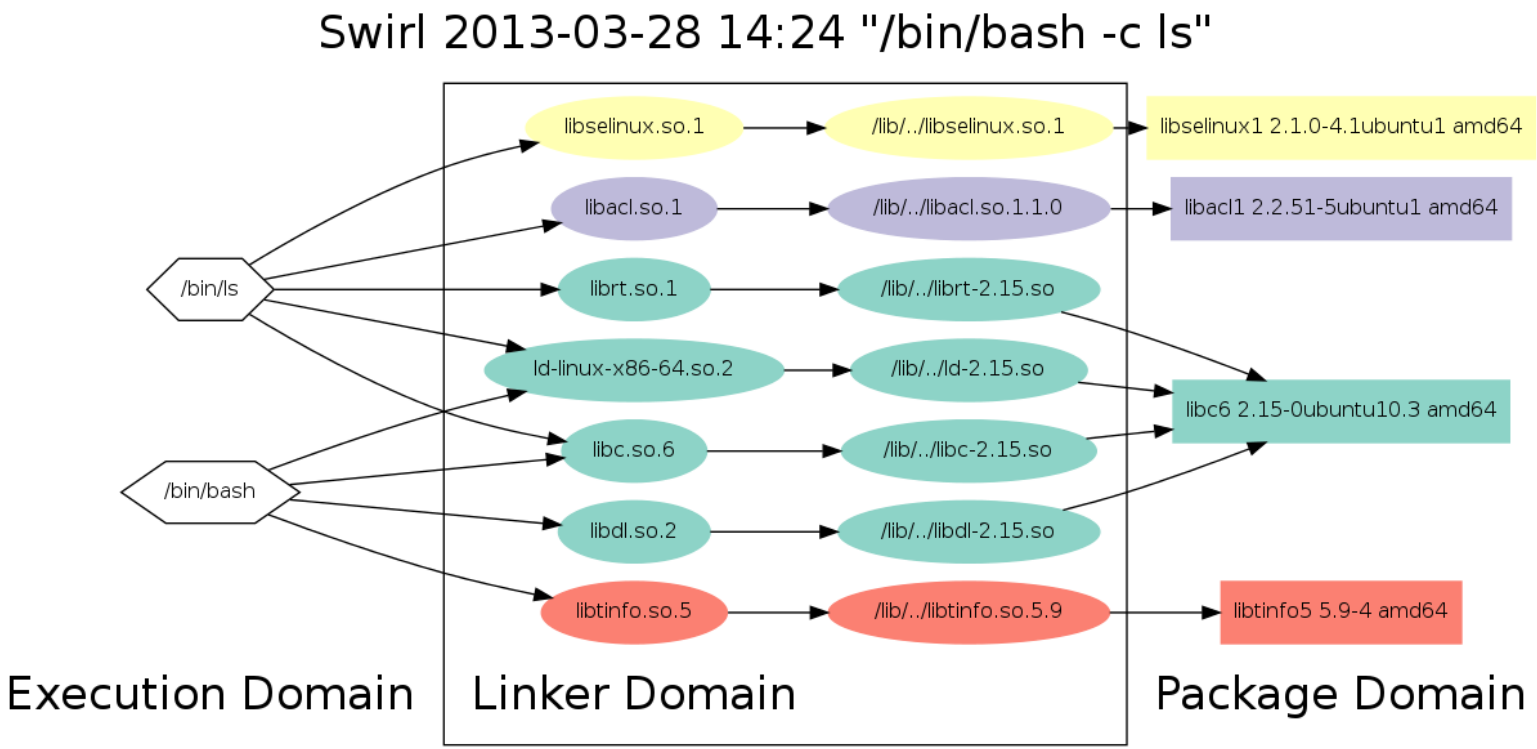
A Swirl can be displayed with text

```
# fingerprint -d
  /bin/ls
    /lib/x86_64-linux-gnu/ld-2.15.so
    /lib/x86_64-linux-gnu/libacl.so.1.1.0
    /lib/x86_64-linux-gnu/libc-2.15.so
    /lib/x86_64-linux-gnu/librt-2.15.so
    /lib/x86_64-linux-gnu/libselinux.so.1
    /lib/x86_64-linux-gnu/libattr.so.1.1.0
    /lib/x86_64-linux-gnu/libpthread-2.15.so
    /lib/x86_64-linux-gnu/libdl-2.15.so
    /lib/x86_64-linux-gnu/libnss_compat-2.15.so --(Dyn)--
    /lib/x86_64-linux-gnu/libnsl-2.15.so --(Dyn)--
    /lib/x86_64-linux-gnu/libnss_nis-2.15.so --(Dyn)--
    /lib/x86_64-linux-gnu/libnss_files-2.15.so --(Dyn)--
      Opened files:
        /etc/localtime
        /etc/group
        /etc/passwd
        /usr/lib/locale/locale-archive
        /usr/lib/x86_64-linux-gnu/gconv/gconv-modules.cache
        /usr/share/locale-langpack/en/LC_MESSAGES/coreutils.mo
        /etc/locale.alias
        /etc/nsswitch.conf
        /proc/filesystems
```

Dependencies Display

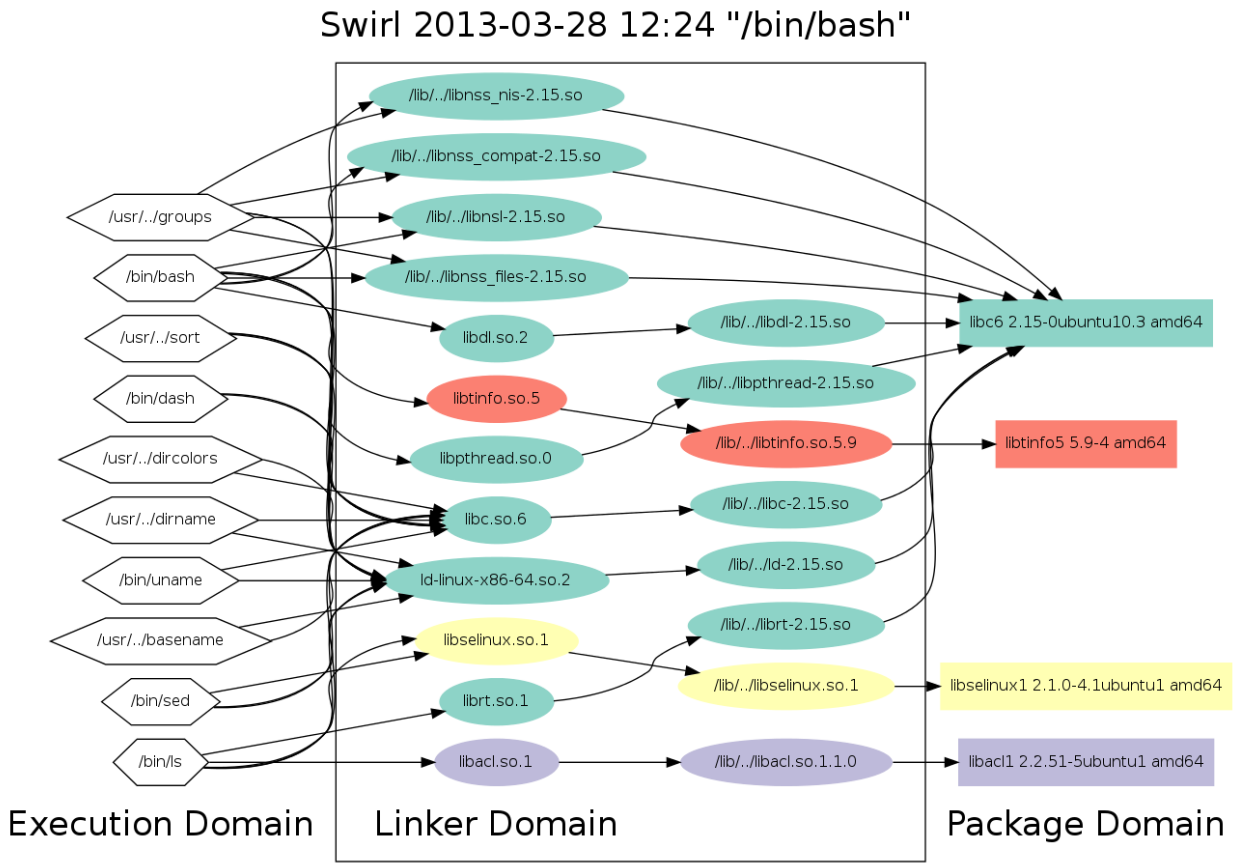
A Swirl can be displayed with Graphviz.

Non-interactive bash `bash -c "ls"`



Dependencies Display

Interactive bash ls



System Validation

Does this system satisfy this set of requirements?

```
# fingerprint -yv
The file output.swirl failed.

Missing Dependecies:
  libunwind-x86_64.so.8()(64bit)
  libunwind-ptrace.so.0()(64bit)
```

Software Stack Integrity

Did any of the requirements change?

```
# fingerprint -yvi
The file output.swirl failed.

Modified Dependecies:
      libunwind-ptrace.so.0()(x86_64)
                        wrong hash
```

Compose a Cluster from a Swirl

- Given a Swirl file we can compose a software environments which provides all the requirement specified in a given Swirl
- It supports only Rocks Clusters

Definition

Rocks is an open source Linux distribution based on Centos that enables users to easily build computational clusters

Compose a Roll

- Create a Swirl (with dynamic tracing)

```
# fingerprint -cx "ls -l"
# ls -lh output.swirl
-rw-rw-r-- 1 root root 26K Oct 16 09:10 output.swirl
```

- Create a Swirl archive

```
# fingerprint -cr
# ls -lh output.tar.gz
-rw-rw-r-- 1 root root 13M Oct 16 09:16 output.tar.gz
```

Compose a Roll

- Copy output.tar.gz on a Rocks Cluster and create a Roll.

  To way to port an application:

  - Native Mode

    ```
    # fingerprint -cm
    ```
  - Sandbox Mode

    ```
    # fingerprint -cmz
    ```

Definition

Roll is the name for a software package in Rocks

Port Dock6 on RedHat 6

Dock6 is a virtual machine used in Pragma 25 and 26 for docking (NAIST).

- It is based on Centos 5.9
- On a 32 bit system
- Uses mpich 3.0.4

Port binaries on on a Rocks 6.1 (Based on Centos 6.3) 64 bit.

Port Dock6 on RedHat 6

- Origin: Dock6 (Centos 5.9 32 bit)
  `@rocks-213`
- Destination: Rocks 6.1 (Centos 6.3 64 bit)
  `@cluster`

Questions?

Software available at:
https://github.com/rocksclusters/FingerPrint

For info:

- lclementi at ucsd.edu