

System Calls Compatibility

Status of syscall support in Gramine

Background

- Paper “[A Study of Modern Linux API Usage and Compatibility: What to Support When You’re Supporting](#)”
 - Study of Linux API usage across all applications and libraries in the Ubuntu Linux distribution
 - Propose metrics for evaluating the relative maturity of a prototype system or compatibility layer
 - Use a combination of **static analysis** to understand API usage and survey data to weight the relative importance of applications to end users
- Paper “[Loupe: Driving the Development of OS Compatibility Layers](#)”
 - Use **dynamic analysis** and determine the OS features that need to be implemented to bring support for a target set of applications and workloads
 - Linux API usage measurements for set of applications results are on website <https://github.com/unikraft/loupedb>
 - It detects if a syscall can be **stubbed** (return -ENOSYS) or **faked** (dummy supported: return 0)
 - “only as few as 20% of (used) system calls reported by static analysis”

Metrics

- API Importance (API usage): Which APIs should be implemented first?
 - For a given API, the probability that an installation includes at least one application requiring the given API.
 - How disruptive its absence would be to applications and end users
- System's API compatibility: What is the progress of API support in a system?
 - % of a user's installed apps supported by the system
 - Weighted Completeness for ***static analysis***
Taking Popularity into Consideration (<https://popcon.debian.org/>)
 - Unweighted Completeness for ***dynamic analysis***

Static analysis

Note that the data only identifies all possible APIs could be invoked in a package. It doesn't matter if application can work with **stubbed** (`return -ENOSYS`) or **faked** (dummy supported: `return 0`) system call

Gramine v1.8

Top 20 not yet supported syscalls ordered by (API Importance)

setresgid	used in 8306 package
setresuid	used in 8306 package
mremap	used in 6808 package
link	used in 2309 package
symlink	used in 1805 package
getresuid	used in 3008 package
getresgid	used in 3004 package
utimes	used in 1474 package
inotify_add_watch	used in 1947 package
inotify_init	used in 1956 package
lchown	used in 1409 package
setxattr	used in 1355 package
inotify_rm_watch	used in 1899 package
inotify_init1	used in 1883 package
shmctl	used in 2671 package
prctl	used in 1539 package
shmat	used in 2668 package
splice	used in 1333 package
shmdt	used in 2847 package
shmget	used in 1938 package

Excluding `setresgid setresuid getresuid setresgid setreuid setregid getrusage prctl shmget shmat shmctl shmdt semget semop semtimedop semctl`

Top 20 not yet supported syscalls ordered by (API Importance)

mremap	used in 6808 package
link	used in 2309 package
symlink	used in 1805 package
utimes	used in 1474 package
inotify_add_watch	used in 1947 package
inotify_init	used in 1956 package
lchown	used in 1409 package
setxattr	used in 1355 package
inotify_rm_watch	used in 1899 package
inotify_init1	used in 1883 package
splice	used in 1333 package
utime	used in 718 package
sync	used in 518 package
times	used in 1381 package
utimensat	used in 125 package
capget	used in 96 package
name_to_handle_at	used in 80 package
timerfd_create	used in 151 package
timerfd_settime	used in 143 package
mount	used in 87 package

Dynamic analysis

Implemented syscalls: 177

Unweighted Completeness = 74.380 %

Top 20 not yet supported syscalls ordered by (API Importance)

<code>prctl</code>	used in 8 datasets, works faked in 2 datasets
<code>utimensat</code>	used in 4 datasets, works faked in 1 datasets
<code>timerfd_create</code>	used in 4 datasets
<code>link</code>	used in 4 datasets, works faked in 1 datasets
<code>restart_syscall</code>	used in 3 datasets
<code>io_setup</code>	used in 3 datasets
<code>io_getevents</code>	used in 3 datasets
<code>io_submit</code>	used in 3 datasets
<code>times</code>	used in 3 datasets
<code>shmget</code>	used in 3 datasets
<code>shmat</code>	used in 3 datasets
<code>getrusage</code>	used in 3 datasets
<code>mremap</code>	used in 2 datasets
<code>timerfd_settime</code>	used in 2 datasets, works faked in 1 datasets
<code>symlink</code>	used in 2 datasets
<code>getresuid</code>	used in 1 datasets, works faked in 1 datasets
<code>timer_settime</code>	used in 1 datasets, works faked in 1 datasets
<code>timer_delete</code>	used in 1 datasets, works faked in 1 datasets
<code>sync</code>	used in 1 datasets
<code>set_mempolicy</code>	used in 1 datasets
<code>sync_file_range</code>	used in 1 datasets
<code>signalfd4</code>	used in 1 datasets
<code>memfd_create</code>	used in 1 datasets
<code>...</code>	

`prctl`

Faked works in wrks:

```
['envoy-benchmark-envoy-test',  
'mongodb-benchmark-mongodb']
```

Required by wrks:

```
['elixir-benchmark-elixir-test',  
'gegl-benchmark-gegl-test',  
'kafka-benchmark-kafka-test',  
'mysql-benchmark-mysql-test',  
'libreoffice-benchmark-libreoffice-test',  
'jython-benchmark-jython-test']
```

`utimensat`

Faked works in wrks:

```
['flask-suite']
```

Required by wrks:

```
['golang-pie-example-benchmark-golang-test',  
'django-suite',  
'golang-httpserver-benchmark-golang-test']
```

`link`

Faked works in wrks:

```
['haproxy-suite-regression-testsuite']
```

Required by wrks:

```
['ocrmypdf-benchmark-ocrmypdf-test',  
'django-suite', 'gnupg-benchmark-gnupg-test']
```

`timerfd_create`

Required by wrks:

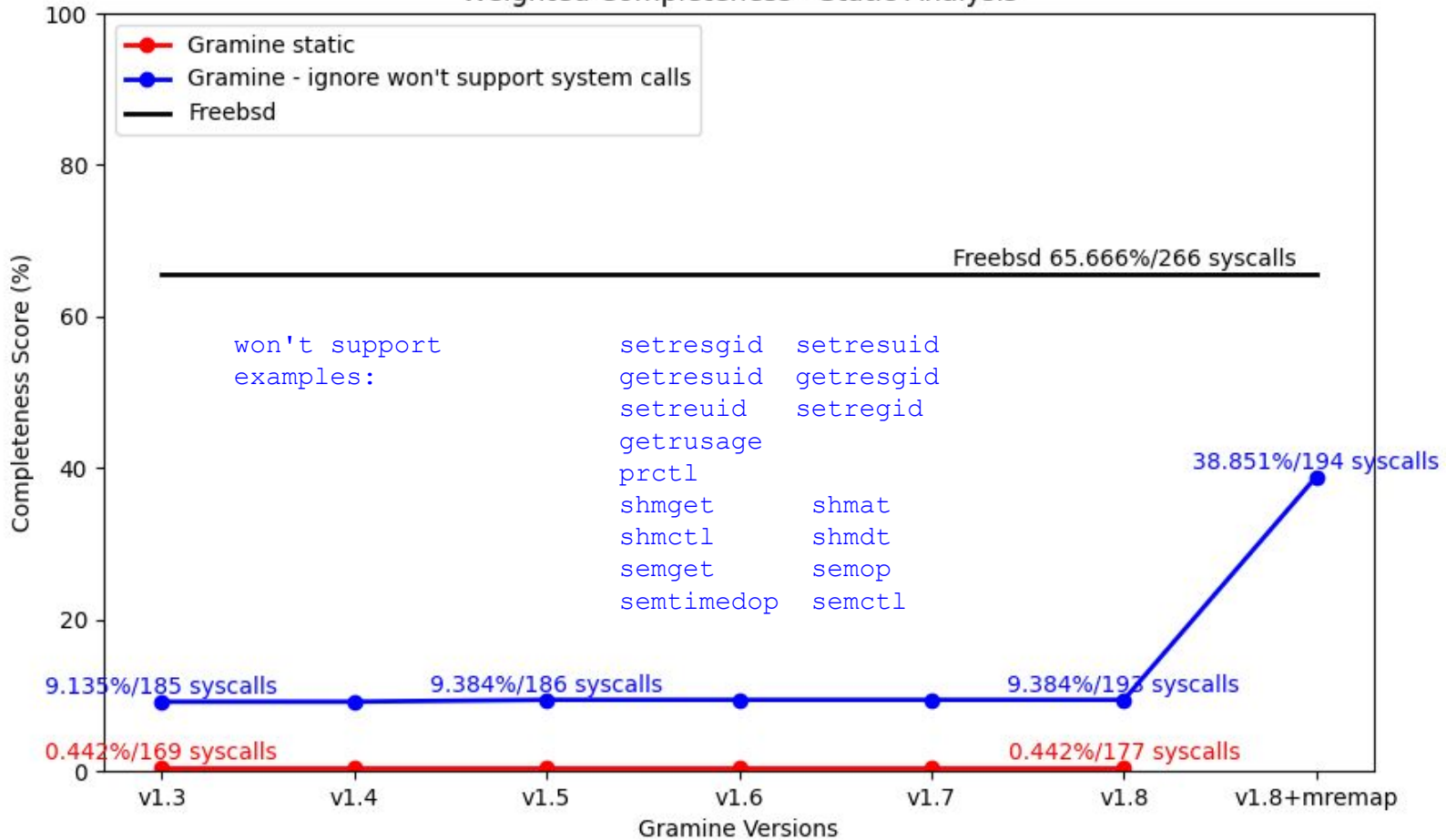
```
['minecraft-server-benchmark-minecraft-test',  
'fio-benchmark-fio-test',  
'elixir-benchmark-elixir-test',  
'mongodb-benchmark-mongodb']
```

`mremap`

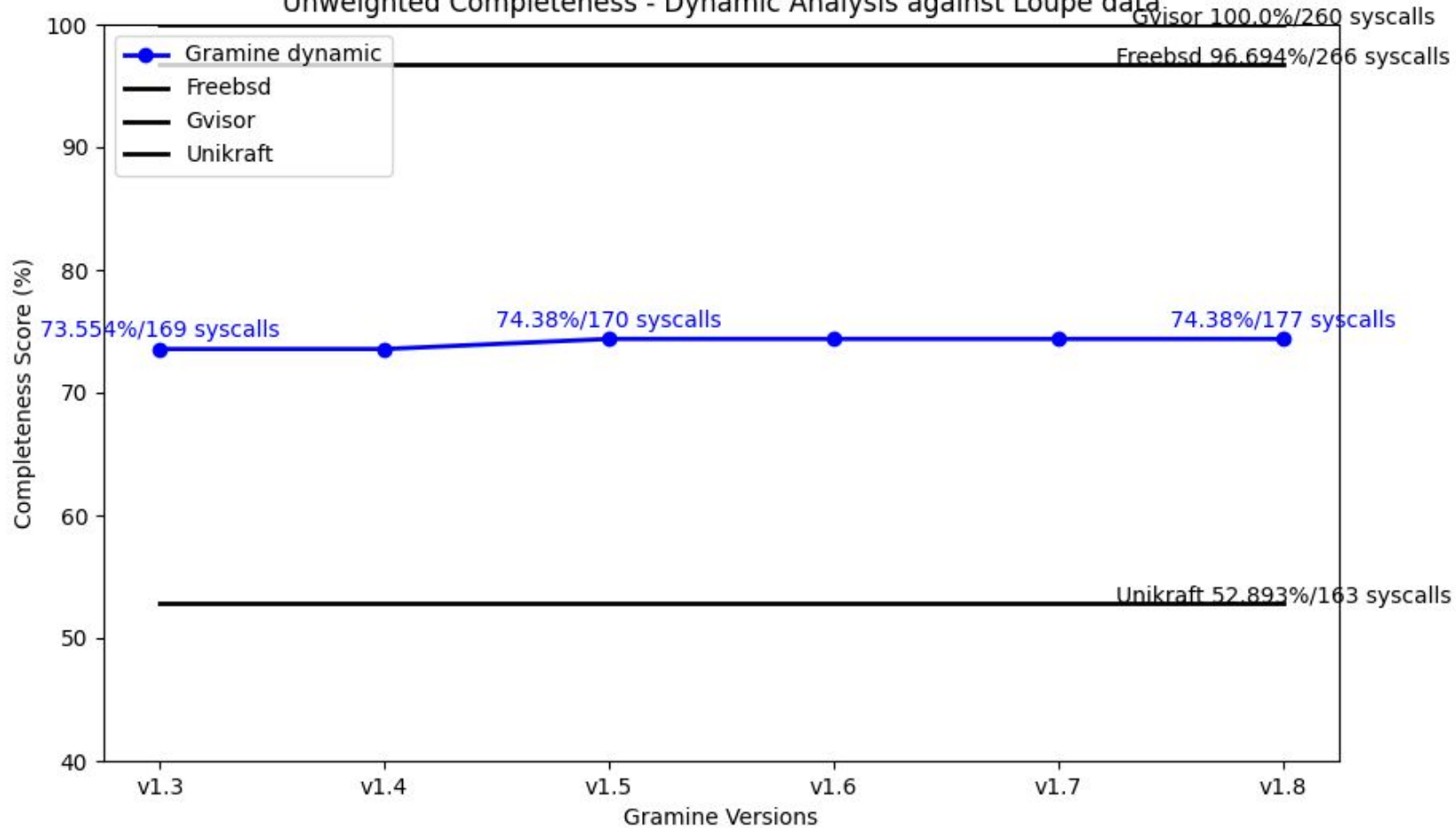
Required by wrks:

```
['golang-pie-example-benchmark-golang-test', 'gimp-benchmark-gimp-test']
```

Weighted Completeness - Static Analysis



Unweighted Completeness - Dynamic Analysis against Loupe data



Support plan example produced by Loupe tool

Step	Implement	Fake	to support
1	times		openssl
2	sync		linpack, blogbench
3	prctl		libreoffice
4	getrusage		kafka, jython
5	link		gnupg
6	utimensat		golang-httpserver
7	restart_syscall		golang-example
8	mremap		golang-pie-example
9		shmat	gimp
10	set_mempolicy	shmdt, shmget, shmctl	amg
11	timerfd_create		mongodb
12			minecraft-server
13	symlink		ocrmypdf
14	timerfd_settime, memfd_create		elixir
15	shmdt, shmget, shmat		gnuradio
16	semget, semctl, semtimedop		dbench
17	io_getevents, io_submit, io_setup		aio-stress
18	get_mempolicy		mysql
19	signalfd4, sync_file_range		fio