Linux Distributions and Applications Analysis During Linux Standard Base Development

Denis Silakov
Institute for System Programming, RAS
silakov@isprras.ru
World of Linux

- Distributions
  - Hundreds
- Applications
  - Thousands
- Users
  - Millions
Application Development

Distributions

- Suse
- Ubuntu
- Red Hat
- Debian

Application
Application Portability

- **Source level portability**
  - Application Program Interface (API) standards
  - Programming languages standards

- **Binary level portability**
  - Application Binary Interface (ABI) standards
  - Standards on executable files, shared libraries, etc.
Linux Standard Base – Binary Standard

Distributions

- Suse
- Ubuntu
- Red Hat
- Debian

LSB

Application
Linux Ecosystem

Distribution

- Libraries
- C++ Classes
- Binary Interfaces
- Commands
- Modules of Interpreted Languages

Application
Standardization Criteria

- Presence in distributions
  - Only «major» distributions are taken into account

- Applications needs

- Maturity
How Much Data Do We Have?

- Typical Linux Distribution (1 DVD)
  - 1-3 thousands of libraries
  - 200-700 thousand of binary interfaces

- LSB 3.2
  - 50 libraries
  - ~39,000 binary interfaces

- Linux Applications
  - 2-30 libraries
  - 100 – 10,000 binary interfaces
Collect Data to a Storage

Distributions

Database

Applications
What Should Be Analysed

Distribution

Library1

Library2

Dynamic loader

LibraryN

Application

Exec1

Library1

Exec2

Library2
Executable and Linking Format (ELF)

- Shared object (library)
  - Soname
  - Exported symbols

- Executable
  - “Needed” libraries (sonames)
  - Required symbols
Executable and Linking Format (ELF)

- **Shared object (library)**
  - Soname
  - Exported symbols

- **Executable**
  - “Needed” libraries (sonames)
  - Required symbols
  - !!! No mapping between symbols and libraries
“Needed” Libraries Problems

- No symbols are used from needed library
  - Usually due to tools like `pkg-config`

- No needed libraries provide used symbols
  - Necessary libs are loaded as dependencies of other needed libs

- Loader «workarounds»
  - `syscall`, `system`, ...
  - `libdl` (`dlsym`, `dladdr`, ...)
Data Extraction

- binutils
  - readelf
  - objdump
  - nm
  - ...

- Analogues from elfutils
  - libdl
  - ...

Interpreted Languages

- Perl and Python modules only

- Data extraction
  - Perl: `perldeps.pl` from rpm
  - Python: `modulefinder` module
Applications Analysis Problems

- Different dependencies in different environments
  - “import” statements under «if» or «try»
  - Import of «all submodules» or «everything from a directory»
Applications Analysis Problems

- Different dependencies in different environments
  - "import" statements under «if» or «try»
  - Import of «all submodules» or «everything from a directory»

More ambiguous then binary files
Data Analysis

Distributions

Applications

Data collection tools (based on readelf, perldeps, ...)

Database

?
Analytical Tools

- **LSB Navigator**
  - Web tool upon the central database

- **LSB ATK Manager**
  - Standalone tool shipping database information
Can We Include These Symbols?
Can We Include These Symbols?

Statistics for LSB Candidates

Note: Different versions of the same application are considered as single application.
Total number of distributions we are looking at: 21

<table>
<thead>
<tr>
<th>Interface</th>
<th>Distro</th>
<th>Total</th>
<th>Proprietary</th>
<th>Mixed License</th>
<th>Open Source</th>
<th>Header</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>winc</td>
<td>21 (list)</td>
<td>481 (list)</td>
<td>64</td>
<td>3</td>
<td>41:4</td>
<td>unistd.h</td>
<td>Yes</td>
</tr>
<tr>
<td>__clpsa_toupper_loc</td>
<td>21 (list)</td>
<td>270 (list)</td>
<td>33</td>
<td>4</td>
<td>233</td>
<td>sysv.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmode</td>
<td>21 (list)</td>
<td>57 (list)</td>
<td>27</td>
<td>0</td>
<td>30</td>
<td>sysv.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmem</td>
<td>21 (list)</td>
<td>24 (list)</td>
<td>17</td>
<td>0</td>
<td>7</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemnode</td>
<td>21 (list)</td>
<td>32 (list)</td>
<td>10</td>
<td>0</td>
<td>22</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>17 (list)</td>
<td>10</td>
<td>0</td>
<td>7</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>14 (list)</td>
<td>10</td>
<td>0</td>
<td>4</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemmove</td>
<td>21 (list)</td>
<td>12 (list)</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>11 (list)</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>11 (list)</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemsetcomp</td>
<td>21 (list)</td>
<td>7 (list)</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>5 (list)</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>5 (list)</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>2 (list)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>2 (list)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>1 (list)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>1 (list)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>1 (list)</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>stdlib.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>wmemset</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>file</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>file</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>file</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>file</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
<tr>
<td>file</td>
<td>21 (list)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>wchar.h</td>
<td>Yes</td>
</tr>
</tbody>
</table>
## Standalone Application Analysis

This page contains analytical information about the application you have tested. It contains the following data:

1. List of the external libraries needed by the application (based on DT_NEEDED ELF section).
2. List of the external interfaces actually required by the application to run.

Please send this information to the Linux Foundation for publishing in the Linux Applications section of the LSB Navigator.

### List of Required Libraries (12 entries total; 12 entries shown)

<table>
<thead>
<tr>
<th>Library</th>
<th>Interfaces Used</th>
<th>In LSB?</th>
<th>Application Modules</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>libICE.so.6</td>
<td>0</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libSM.so.6</td>
<td>0</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libX11.so.6</td>
<td>93</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libXt.so.6</td>
<td>6</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libxc.so.6</td>
<td>145</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libXi.so.2</td>
<td>4</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libgcc_s.so.1</td>
<td>1</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libln.so.6</td>
<td>22</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libm.so.6</td>
<td>27</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libpng.so.0</td>
<td>817</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libstdc++.so.6</td>
<td>1</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>libz.so.1</td>
<td>9</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
</tbody>
</table>

### List of Required Interfaces (1125 entries total; 1125 entries shown)

<table>
<thead>
<tr>
<th>Interface</th>
<th>Library</th>
<th>In LSB?</th>
<th>Application Modules</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>XChangeProperty</td>
<td>libX11</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
<tr>
<td>XChangeWindowAttributes</td>
<td>libX11</td>
<td>Yes</td>
<td>X11lib/ice0.6-26-20080218.8opera</td>
<td>No information</td>
</tr>
</tbody>
</table>
Future Directions

“On which distributions the application can be launched?”

- Shell scripts analysis
  - Commands usage
  - Non-standard features used
Collected Data

- 38 distributions
- \(\sim\) 1100 applications
  - 220 proprietary
- \(\sim\) 25 millions of records in 17 tables
URLs

- linuxbase.org
- isprases.linux-foundation.org
- linuxxtesting.org
Thank You!