

# **Linux Distributions and Applications Analysis During Linux Standard Base Development**

**Denis Silakov**

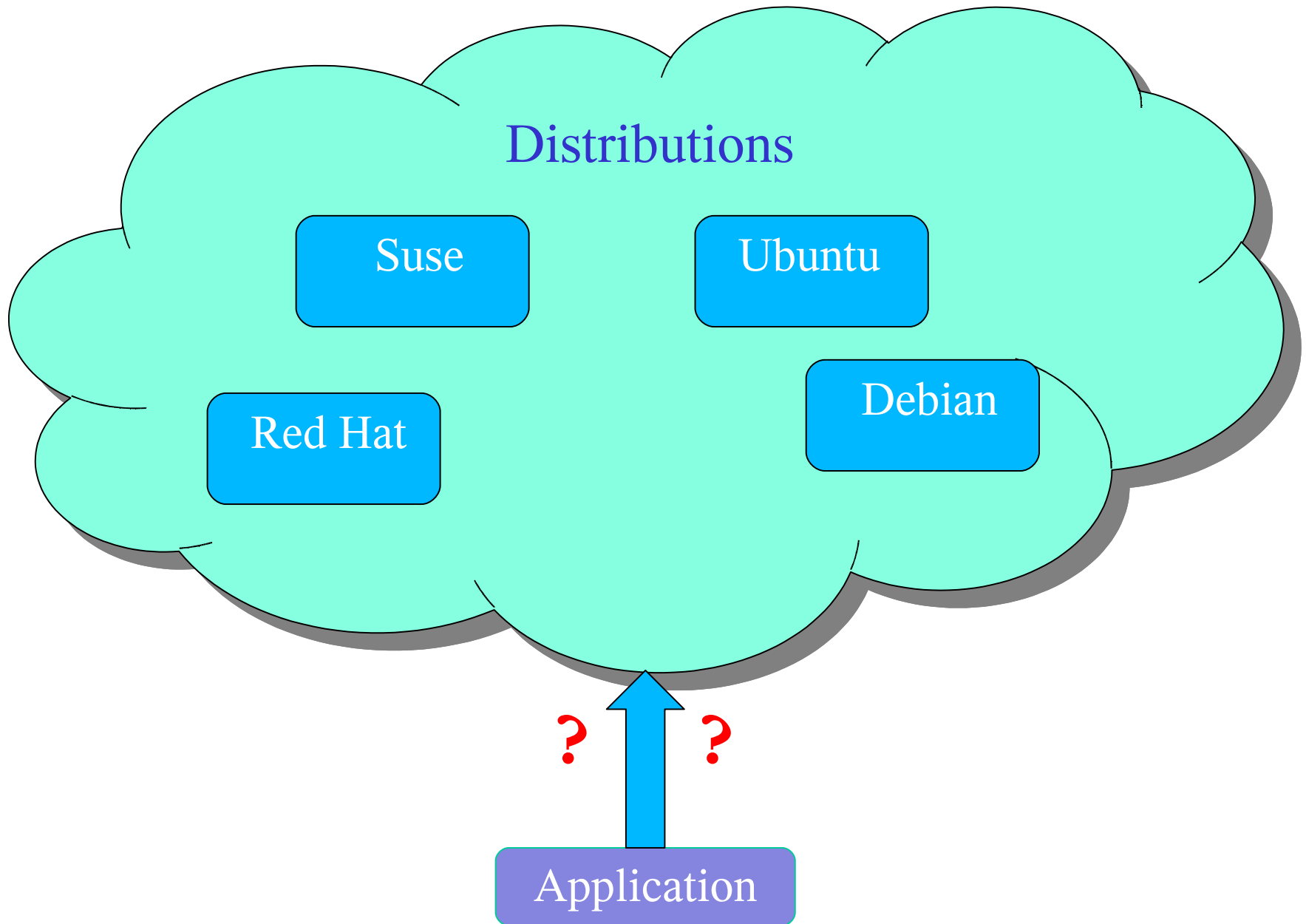
Institute for System Programming, RAS

*silakov@ispras.ru*

# World of Linux

- Distributions
  - ◆ Hundreds
- Applications
  - ◆ Thousands
- Users
  - ◆ Millions

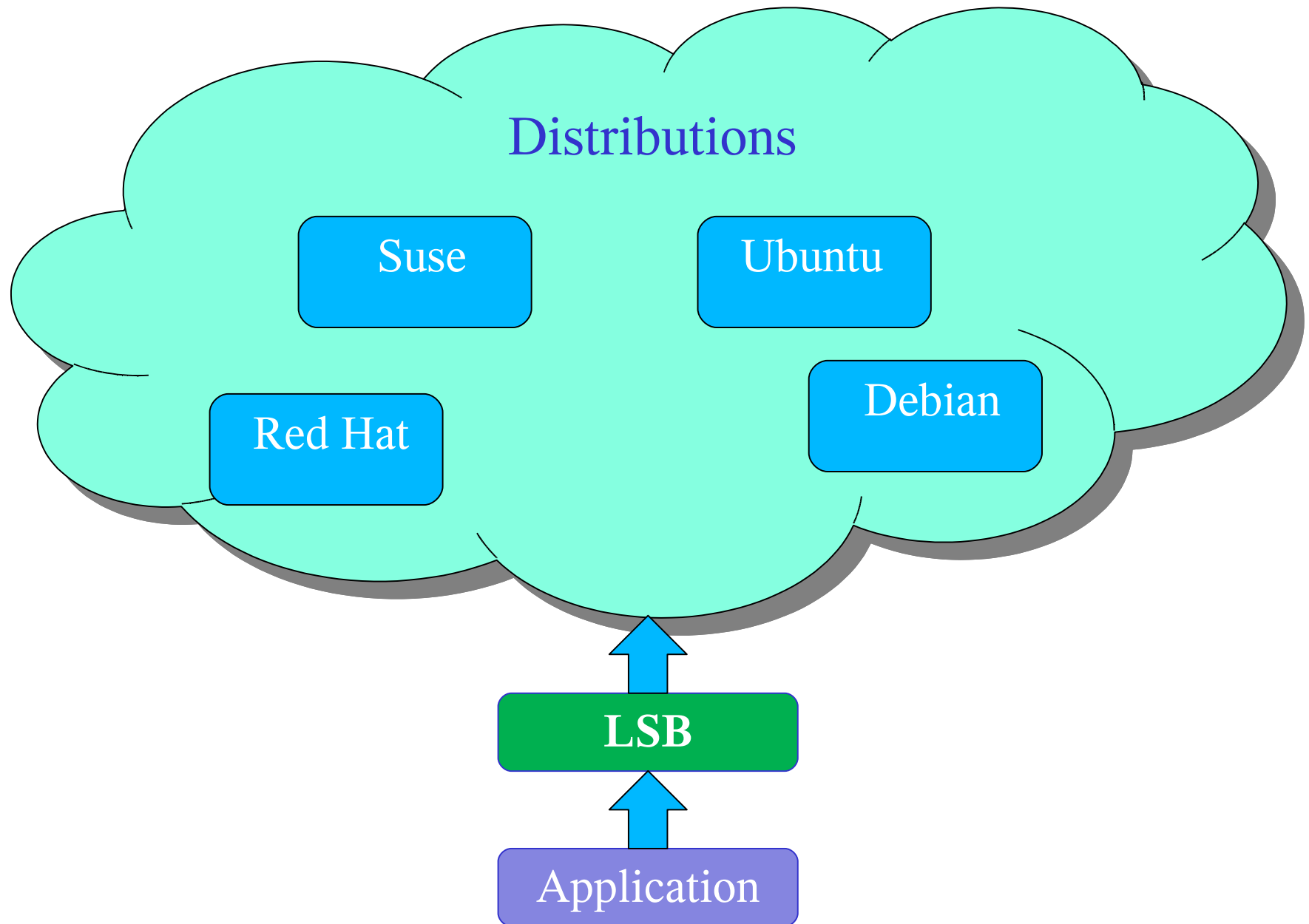
# Application Development



# 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.

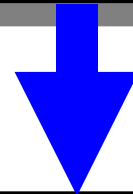
# Linux Standard Base – Binary Standard



# Linux Ecosystem

## Distribution

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



**Application**

# Standardization Criteria

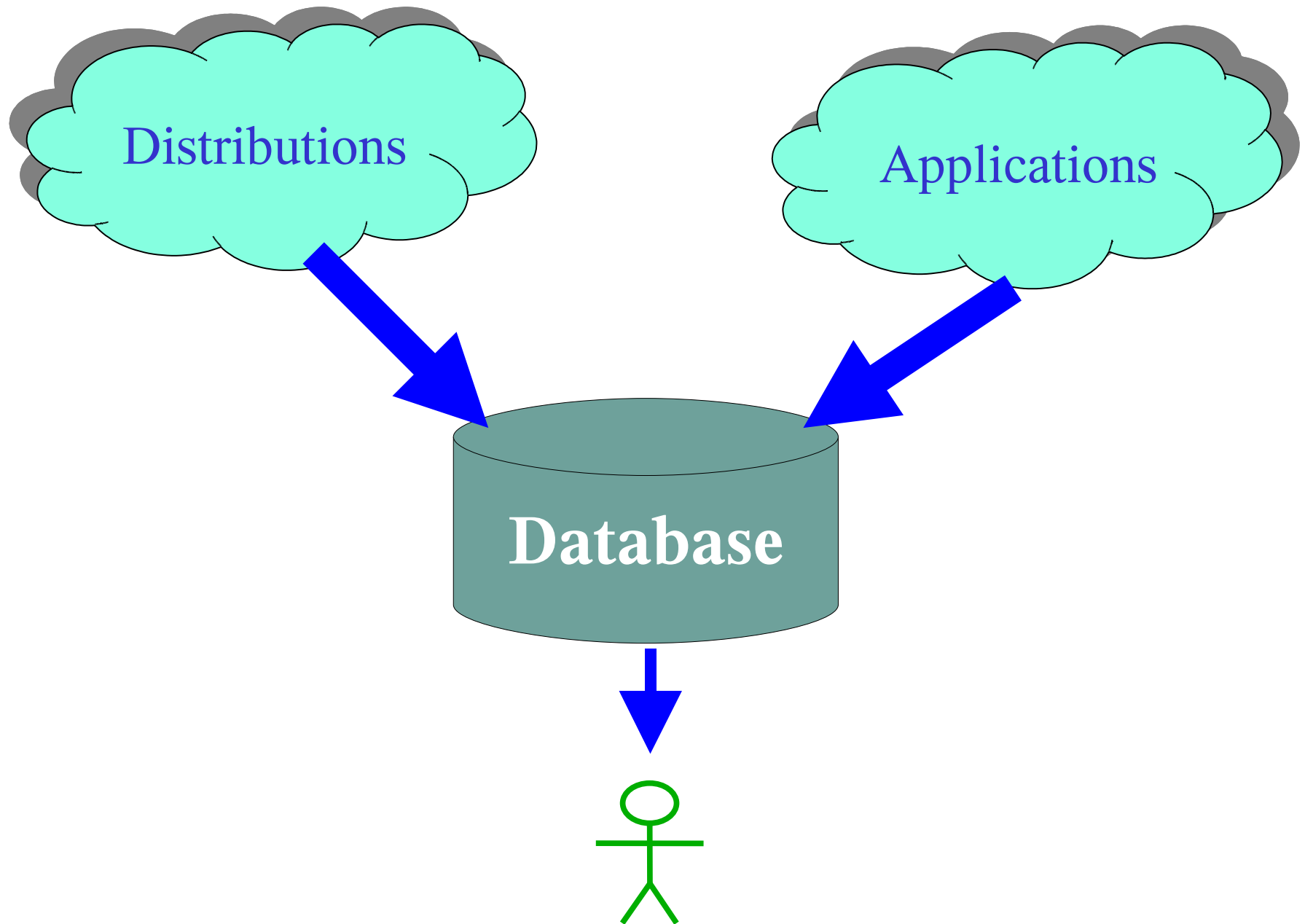
- Presence in distributions
  - ◆ Only «major» distributions are taken into account
- Applications needs
- Maturity

# How Much Data Do We Have?

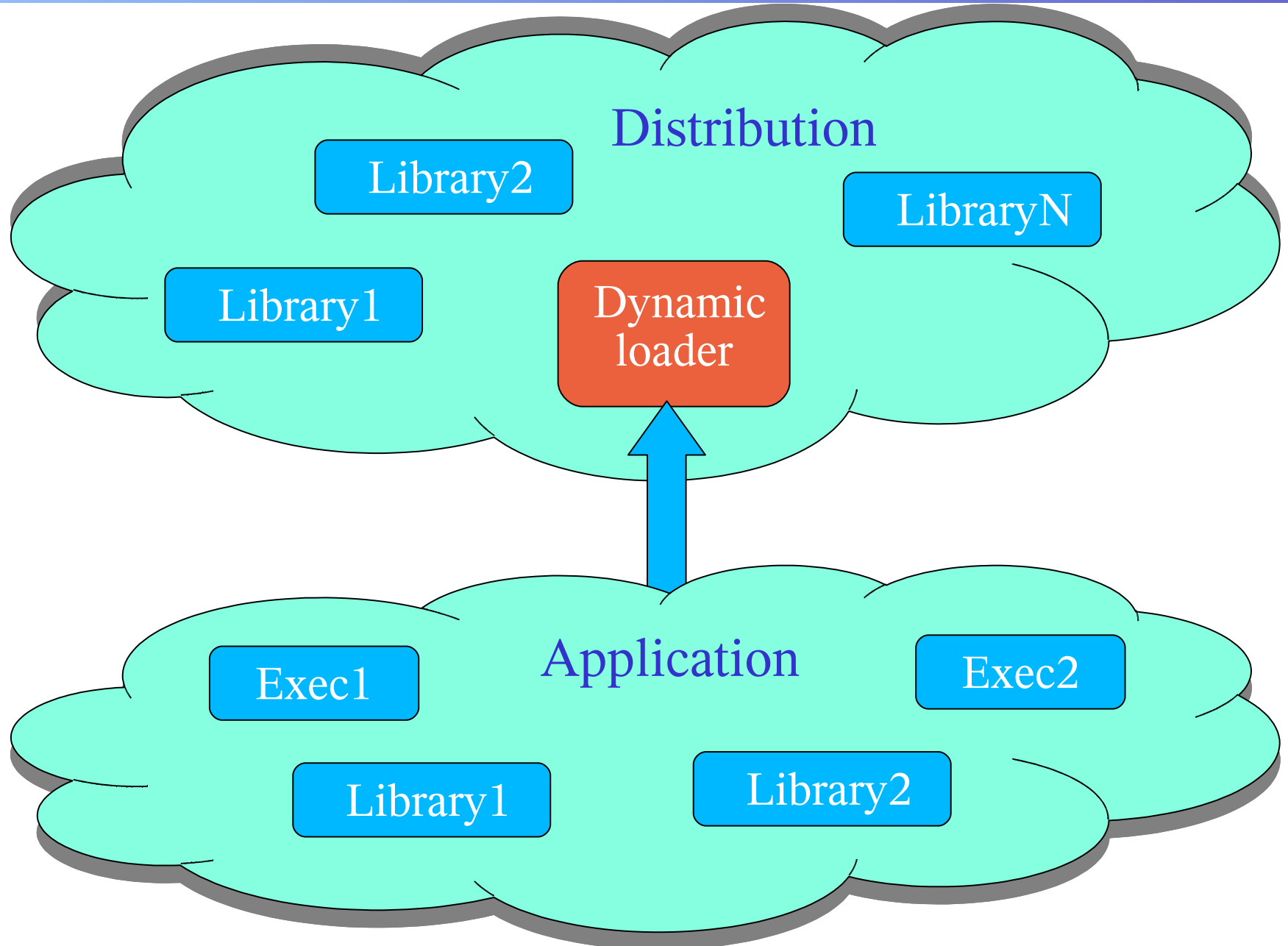
- Typical Linux Distribution (1 DVD)
  - ◆ 1-3 thousands of libraries
  - ◆ 200-700 thousand of binary interfaces
- LSB 3.2
  - ◆ 50 libraries
  - ◆ ~39.000 binary interfaces
- Linux Applications
  - ◆ 2-30 libraries
  - ◆ 100 – 10.000 binary interfaces



# Collect Data to a Storage



# What Should Be Analysed



# 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»

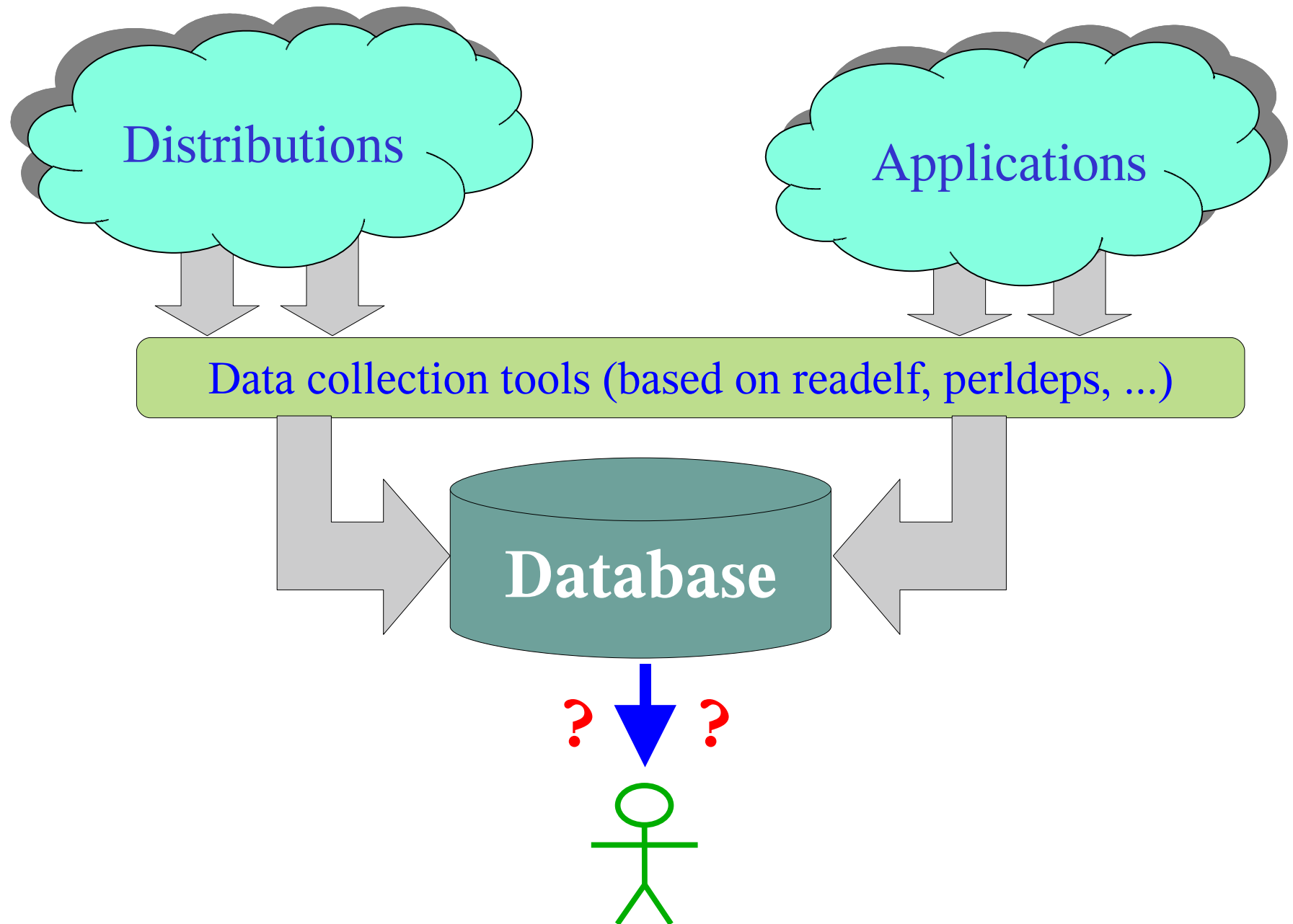


# Applications Analysis Problems

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

**More ambiguous than binary files**

# Data Analysis



# Analytical Tools

- LSB Navigator
  - ◆ Web tool upon the central database
- LSB ATK Manager
  - ◆ Standalone tool shipping database information

# Can We Include These Symbols?

The screenshot shows a web browser window displaying the Linux Foundation Database Navigator interface. The page is titled "Statistics for LSB Candidates" and is in "Administration mode". The "Set Library" dropdown is set to "\*\*None\*\*" and the manual entry field contains "libc".

**List of interfaces:**

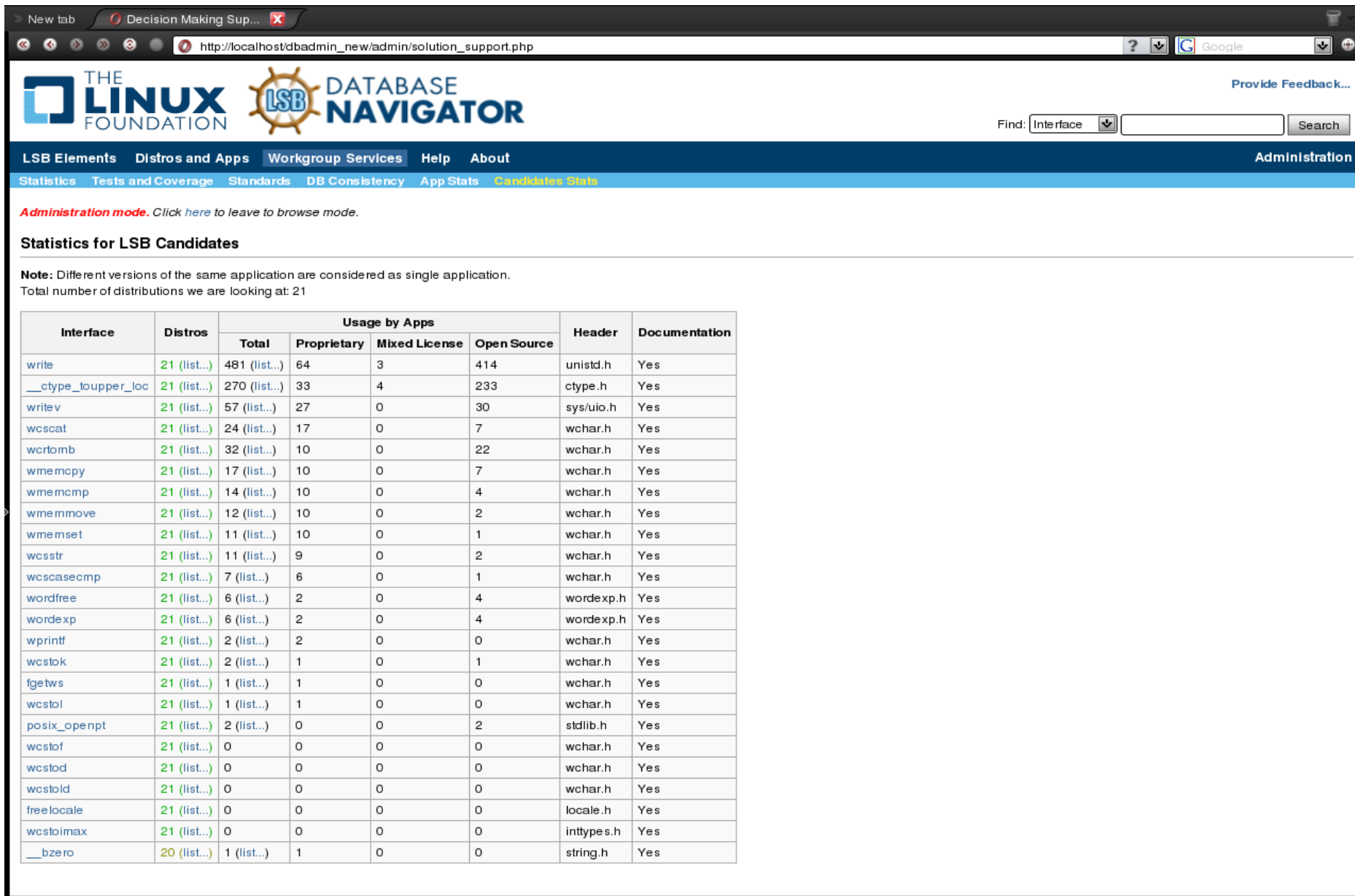
```
_bzero
// comment
garbage entry
freelocale
posix_openpt
__ctype_toupper_loc
fgetws
wrtomb
wscasecmp
wscat
wcastr
wcstod
wcstof
wcstoimax
wcstok
wcstol
wcstold
wmemcmp
wmemcpy
wmemmove
wmemset
wordexp
wordfree
wprintf
write
writev
```

**Distributions we are interested in:**

- Asianux 2.0 for x86
- Asianux Server 3.0 for x86
- BOSS 2.0 for x86
- BOSS 1.0 for x86
- Debian 4.0 for x86
- Debian 3.1r2 for x86
- Debian 4.0 for PPC32
- Fedora 8 for x86
- Fedora 7 for x86
- Gentoo 2007.0 for x86
- Mandriva 2008 for x86
- Mandriva 2007\_Spring for x86
- Mandriva 2006 for x86
- Mandriva Corporate Server 4 for x86
- Oracle Enterprise Linux 5 for x86
- Oracle Enterprise Linux 5 for x86-64
- RHEL 5 for x86
- RHEL 4 for x86
- RHEL 4 for x86-64
- SLES 10 for x86
- SLES 10 SP1 for IA64
- SLES 10 for IA64
- SLES 10 SP1 for x86-64
- SLES 10 for S390X
- Ubuntu 7.10 for x86
- Ubuntu 7.04 for x86
- Ubuntu 6.06 for x86
- Ubuntu 7.10 for x86-64
- Ubuntu 7.04 for x86-64
- Xandros Desktop 4.1 for x86
- Xandros Server 1.0 for x86
- openSUSE 10.3 for x86-64
- openSUSE 10.2 for x86-64

Get Statistics

# Can We Include These Symbols?



The screenshot shows a web browser window displaying the Linux Foundation Database Navigator website. The page title is "Statistics for LSB Candidates". A note indicates that different versions of the same application are considered as single applications, and the total number of distributions being looked at is 21. A table titled "Usage by Apps" lists various interfaces and their usage across different license types (Total, Proprietary, Mixed License, Open Source) and documentation status.

**Administration mode.** Click [here](#) to leave to browse mode.

### 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

Interface	Distros	Usage by Apps				Header	Documentation
		Total	Proprietary	Mixed License	Open Source		
write	21 (list...)	481 (list...)	64	3	414	unistd.h	Yes
__ctype_toupper_loc	21 (list...)	270 (list...)	33	4	233	ctype.h	Yes
writev	21 (list...)	57 (list...)	27	0	30	sys/uio.h	Yes
wcscat	21 (list...)	24 (list...)	17	0	7	wchar.h	Yes
wcrtomb	21 (list...)	32 (list...)	10	0	22	wchar.h	Yes
wmemcpy	21 (list...)	17 (list...)	10	0	7	wchar.h	Yes
wmemcmp	21 (list...)	14 (list...)	10	0	4	wchar.h	Yes
wmemmove	21 (list...)	12 (list...)	10	0	2	wchar.h	Yes
wmemset	21 (list...)	11 (list...)	10	0	1	wchar.h	Yes
wcsstr	21 (list...)	11 (list...)	9	0	2	wchar.h	Yes
wcscasecmp	21 (list...)	7 (list...)	6	0	1	wchar.h	Yes
wordfree	21 (list...)	6 (list...)	2	0	4	wordexp.h	Yes
wordexp	21 (list...)	6 (list...)	2	0	4	wordexp.h	Yes
wprintf	21 (list...)	2 (list...)	2	0	0	wchar.h	Yes
wcstok	21 (list...)	2 (list...)	1	0	1	wchar.h	Yes
fgetws	21 (list...)	1 (list...)	1	0	0	wchar.h	Yes
wcstol	21 (list...)	1 (list...)	1	0	0	wchar.h	Yes
posix_openpt	21 (list...)	2 (list...)	0	0	2	stdlib.h	Yes
wcstof	21 (list...)	0	0	0	0	wchar.h	Yes
wcstod	21 (list...)	0	0	0	0	wchar.h	Yes
wcstold	21 (list...)	0	0	0	0	wchar.h	Yes
freelocale	21 (list...)	0	0	0	0	locale.h	Yes
wcstoimax	21 (list...)	0	0	0	0	inttypes.h	Yes
__bzero	20 (list...)	1 (list...)	1	0	0	string.h	Yes

# Standalone Application Analysis

The screenshot displays the 'Test Results' page for 'opera on x86' in the Linux Foundation Application Testkit Manager. The page includes a navigation menu with 'Application Check', 'Results', 'Help', 'About', and 'Administration'. The main content area shows the analysis results, including a list of required libraries and interfaces. The 'List of Required Libraries' table is filtered by 'Presence in LSB: \*No filter\*' and shows 12 entries. The 'List of Required Interfaces' table is also filtered by 'Presence in LSB: \*No filter\*' and 'Library: \*No filter\*', showing 1125 entries.

## Analysis Results for opera on x86

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

- List of the external **libraries** needed by the application (based on DT\_NEEDED ELF section).
- 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)

Filter by: Presence in LSB: \*No filter\*

Library	Interfaces Used	In LSB?	Application Modules	Comments
<a href="#">libICE.so.6</a>	0	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libSM.so.6</a>	0	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libX11.so.6</a>	93	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libXext.so.6</a>	6	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libc.so.6</a>	145	Yes	/usr/lib/opera/9.26-20080218.6/missingsyms.so /usr/lib/opera/9.26-20080218.6/opera /usr/lib/opera/9.26-20080218.6/spellcheck.so /usr/lib/opera/9.26-20080218.6/works	No information
<a href="#">libdl.so.2</a>	4	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libgcc_s.so.1</a>	1	Yes	/usr/lib/opera/9.26-20080218.6/spellcheck.so	No information
<a href="#">libm.so.6</a>	22	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libpthread.so.0</a>	27	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libqt-mt.so.3</a>	817	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libstdc++.so.6</a>	1	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">libz.so.1</a>	9	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information

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

Filter by: Presence in LSB: \*No filter\* Library: \*No filter\*

Interface	Library	In LSB?	Application Modules	Comments
<a href="#">XChangeProperty</a>	<a href="#">libX11</a>	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information
<a href="#">XChangeWindowAttributes</a>	<a href="#">libX11</a>	Yes	/usr/lib/opera/9.26-20080218.6/opera	No information

# Future Directions

- “On which distributions the application can be launched?”
- Shell scripts analysis
  - ◆ Commands usage
  - ◆ Non-standard features used

# Collected Data

- 38 distributions
- ~ 1100 applications
  - ◆ 220 proprietary
- ~ 25 millions of records in 17 tables



# URLs

- [linuxbase.org](http://linuxbase.org)
- [ispras.linux-foundation.org](http://ispras.linux-foundation.org)
- [linuxtesting.org](http://linuxtesting.org)



**Thank You!**