CDE: Run Any Linux Application On-Demand Without Installation

Philip Guo
pg@cs.stanford.edu
http://www.pgbovine.net/cde.html
Problem

It’s hard to package up your software so that other people can reliably run it ...
It’s hard to package up your software so that other people can reliably run it ...

... and it can be hard to install and run other people’s software ...

[Numpy-discussion] problem installing numpy on Fedora

Charles R Harris charlesr.harris@gmail.com
Wed Apr 13 09:17:15 CDT 2011

On Wed, Apr 13, 2011 at 7:30 AM, c cook <caecook@gmail.com> wrote:

> Hello,
> > I am trying to install numpy 1.6.0b2 version on Fedora and I am getting the following error:
> > In file included from numpy/core/src/private/npy_config.h:4,
> > from numpy/core/src/multiarray/common.o:8,
> > from numpy/core/src/multiarray/multiarraymodule_onefile.c:8:
> > build/src/linux-x86_64-2.4/numpy/core/include/numpy/config.h:34:1: warning: `HAVE_HYPOT` redefined
> > In file included from /home/cce/Apps/Python-2.4/Include/Python.h:8,
> > from numpy/core/src/multiarray/common.o:2,
Problem

It’s hard to package up your software so that other people can reliably run it ... 

... and it can be hard to install and run other people’s software ... 

... especially on Linux
Example: Scientists and Sysadmins

Alice’s computer

```
$ cd /home/alice/cool-experiment/
$ python weather_sim.py tokyo.dat
```

Bob’s computer

```
$ cd /home/bob/cool-experiment/
$ python weather_sim.py tokyo.dat
ERROR: Cannot find py-weather.so
```
Sysadmin woes

Difficult to install custom versions of Python and 3rd-party extension libs on university machines

```bash
cd /home/bob/cool-experiment/
python weather_sim.py tokyo.dat
```
CDE: Automatic packaging of Code, Data, and Environment

1. Create package on your Linux computer
   Prepend any set of commands with `cde`, and CDE runs them and automatically packages up their dependencies

2. Transfer package
   A package is simply a directory of files (~10MB – 500MB), so it can be transferred via USB stick, DVD, or over the Internet

3. Execute software from within package on any modern Linux computer
   Prepend those same commands with `cde-exec`, and CDE runs them natively without any installation
1. CDE overview
2. Seamless execution
3. Run any app on-demand
Alice’s computer

cd /home/alice/cool-experiment/
cde python weather_sim.py tokyo.dat
Bob’s computer

Bob’s computer

cde-exec python weather_sim.py tokyo.dat

Alice’s computer

Alice’s computer

[cde-package]

[cde-root]

[usr]

[lib]

[py-weather.so]

[libc-2.10.so]

[bin]

[python]

[alice]

[cool-experiment]

[weather_sim.py]

[tokyo.dat]
Bob’s computer

Bob no longer needs to bother his university sysadmins!

cde-exec python weather_sim.py tokyo.dat
Creating a package with *cde*

- **cde**\[ptrace\] monitored process
- open("/lib/libc.so.6")

**Timeline**
- copy /lib/libc.so.6 into cde-package/
Executing a package with `cde-exec`

Timeline

```
open("/home/bob/cde-package/cde-root/lib/libc.so.6")
```

```
open("/lib/libc.so.6")
```
16 diverse CDE packages can execute on popular distros with **no installation or configuration**.

Requires user-kernel ABI to be compatible
Run-time slowdowns

Proportional to $\# \text{syscalls} / \text{sec}$.

- SPEC CPU2006 benchmarks: 0% - 4%
- Batch file processing apps: 2% - 28%
- GUI apps: slower start-up
1. CDE overview
2. Seamless execution
3. Run any app on-demand
Seamless execution

cde-exec python <Alice’s scripts> /var/log/httpd/access_log

/home/bob/cde-package/cde-root/usr/bin/python
/home/bob/cde-package/cde-root/usr/lib/logutils.so
/home/bob/cde-package/cde-root/var/log/httpd/access_log 😞
/var/log/httpd/access_log
1. CDE overview
2. Seamless execution
3. Run any app on-demand
Run any app on-demand

Sysadmins maintain a cloud distro farm

Users stream selected apps on-demand
Alice wants to run Eclipse without installing it on her machine

```
mount cloud distro containing Eclipse
```

```
cde-exec -s eclipse
```
CDE streaming mode

Instead of delivering a package, have users connect to your server and stream your apps on-demand.

Solves the package incompleteness problem!
End Linux distro holy wars!

Distro farm houses all versions of all package management systems.

Users have **convenience** and **freedom** to choose from the best of ALL packages on ALL distros!
CDE: Automatic packaging of Code, Data, and Environment

~4000 downloads so far
(Google for “cde linux”)

**Real-world use cases:**
Sharing prototype software
Deploying custom web app stacks
Reproducing research experiments
Running new software on older distros
Deploying computations to cluster/cloud