FreeBSD Jailing, A Secure Virtual Machine

Sept. 1, 2004, Isaac Levy Presenter

NYCBUG

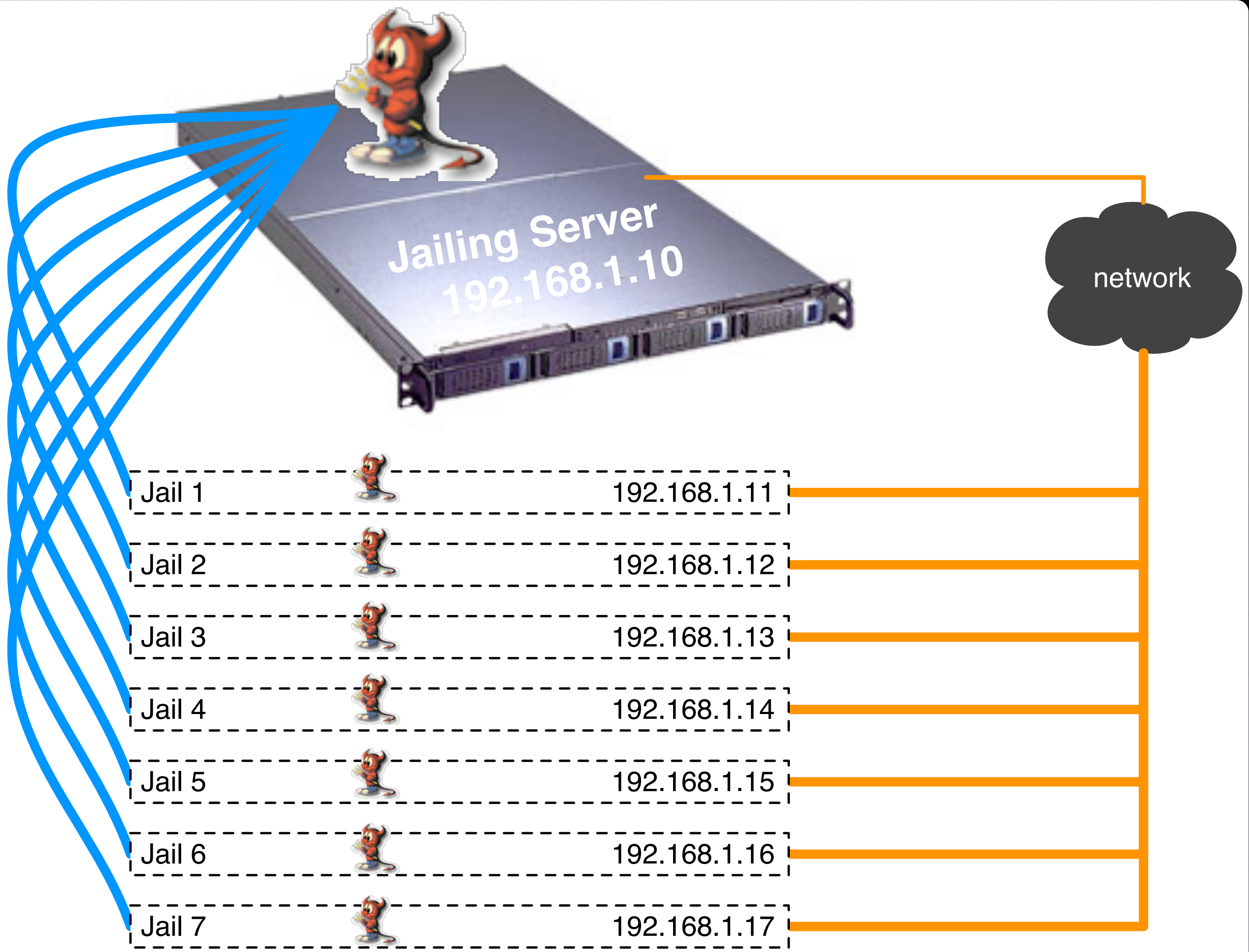
NEW YORK CITY *BSD USER GROUP

Definitions of Jailing

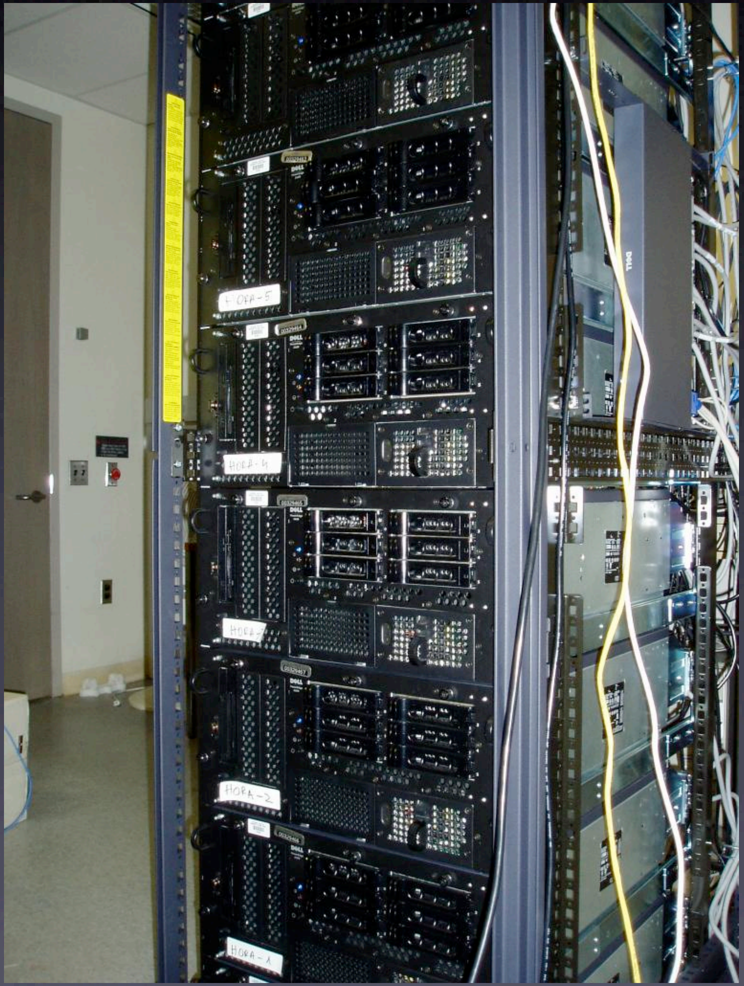
- what is **jail(8)**:
  - a system call in FreeBSD
  - virtual system image
  - process tree based
- what jail **jail(8)** is not:
  - chroot (ala OpenBSD vocabulary)
  - it is not a classical machine emulator

Process Tree:

```
JailingServer:
  \_init
    \_daemon/process etc...
    \_daemon/process etc...
    \_daemon/process etc...
    \_daemon/process etc...
    \_jail (Jail 1)
      \_daemon/process etc...
      \_daemon/process etc...
      \_daemon/process etc...
    \_jail (Jail 2)
      \_daemon/process etc...
      \_daemon/process etc...
      \_daemon/process etc...
    \_jail (Jail 3)
      \_daemon/process etc...
      \_daemon/process etc...
      \_daemon/process etc...
    \_jail (Jail 4)
      \_daemon/process etc...
      \_daemon/process etc...
      \_daemon/process etc...
```



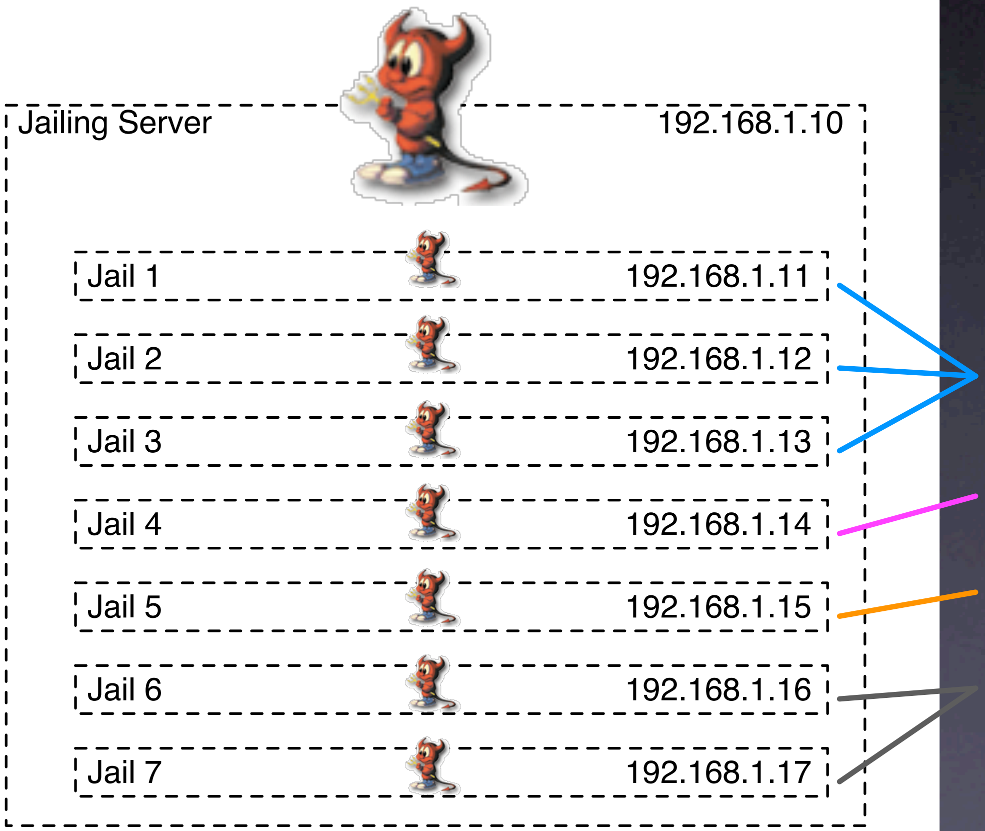
maintaining old junk?



- 3 webservers
- 1 local-use dns cache
- fileserver (for 2 people)
- 2 dev servers

jail(8)!

(special thanks to jail(2)!)



Jailing Server	192.168.1.10
Jail 1	192.168.1.11
Jail 2	192.168.1.12
Jail 3	192.168.1.13
Jail 4	192.168.1.14
Jail 5	192.168.1.15
Jail 6	192.168.1.16
Jail 7	192.168.1.17

3 webservers

1 local-use dns cache

fileserver (for 2 people)

2 dev servers

Some Uses for **jail(8)**

- when entire OS is needed for a small service
- separating specific services securely
- Learning/Development Servers
- Hardware Resource Sharing
- High Availability / Low Cost

Poor Uses for **jail(8)**

- Kernel Access
- Network Interface Access
- Device Driver Access
- Some Applications require particular low-level system calls:
  - Notably, PostgreSQL doesn’t run securely until FreeBSD 5.x stable
- when chroot(8) will do the job **simpler**

How To **jail(8)**

- GREAT instructions in **man jail**, but nutshell,
    - build a FreeBSD userland from source somewhere on host machine, minor tweaks.
    - create an IP alias interface
    - run the jail call with the IP, and userland, the jail ‘boots’, so to speak.

Best Pracitces

- ssh into jails to manage their processes!
- You always can see straight filesystem/userland from master jail
- Design your jailing system carefully, be creative (Hans note about readonly mounts)
- Use highest secure practices possible for master server...

Process Tree:

```
JailingServer
  \_init
      \_daemon/process etc...
      \_daemon/process etc...
      \_daemon/process etc...
      \_daemon/process etc...
      \_jail (Jail 1)
          \_daemon/process etc...
          \_daemon/process etc...
          \_daemon/process etc...
      \_jail (Jail 2)
          \_daemon/process etc...
          \_daemon/process etc...
          \_daemon/process etc...
      \_jail (Jail 3)
          \_daemon/process etc...
          \_daemon/process etc...
          \_daemon/process etc...
      \_jail (Jail 4)
          \_daemon/process etc...
          \_daemon/process etc...
          \_daemon/process etc...
```

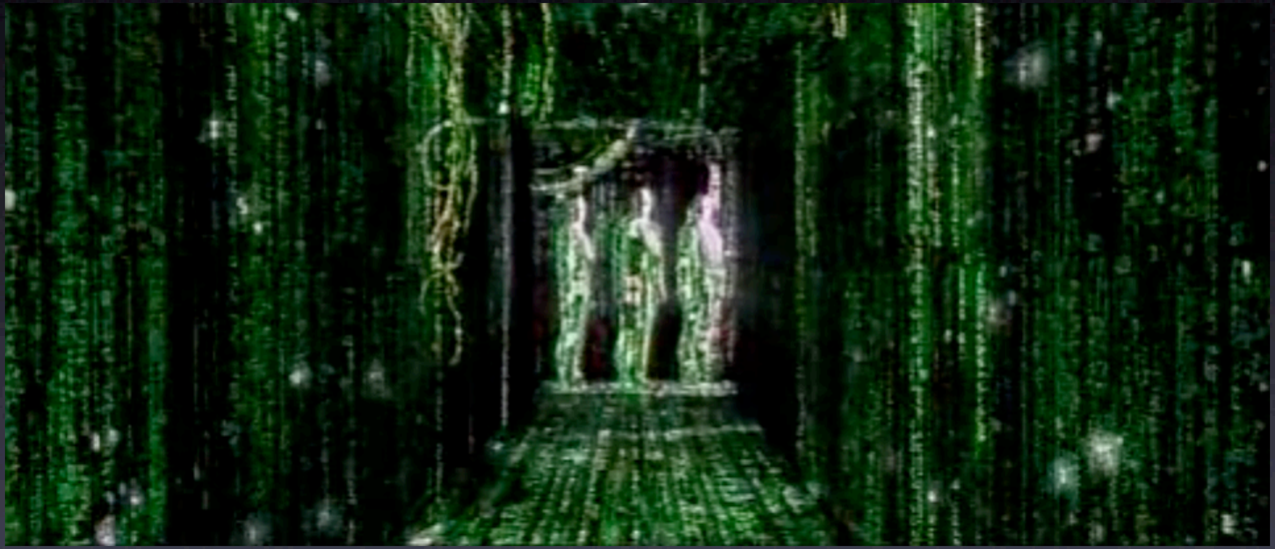
http://www.samag.com/documents/s=1151/sam0105d/0105d.htm

OpenRoot Project, fork-bombs, FreeBSD SecureLevels, reality, truth, and process control

PostgreSQL *sigh*

- Time At iMeme, people want PostgreSQL
- if you give jails sysvipc,
  - /sbin/sysctl -w jail.sysvipc_allowed=1
    - EXCEPT this *fix* lets jails access main system memory directly, therefore shooting security in the foot...
- This kind of issue *can* arise elsewhere.

Questions/Discussion



(the following urls will hit nycbug-talk late tonight)

http://www.freebsd.org/doc/en_US.ISO8859-1/books/arch-handbook/jail.html

http://www.samag.com/documents/s=1151/sam0105d/0105d.htm

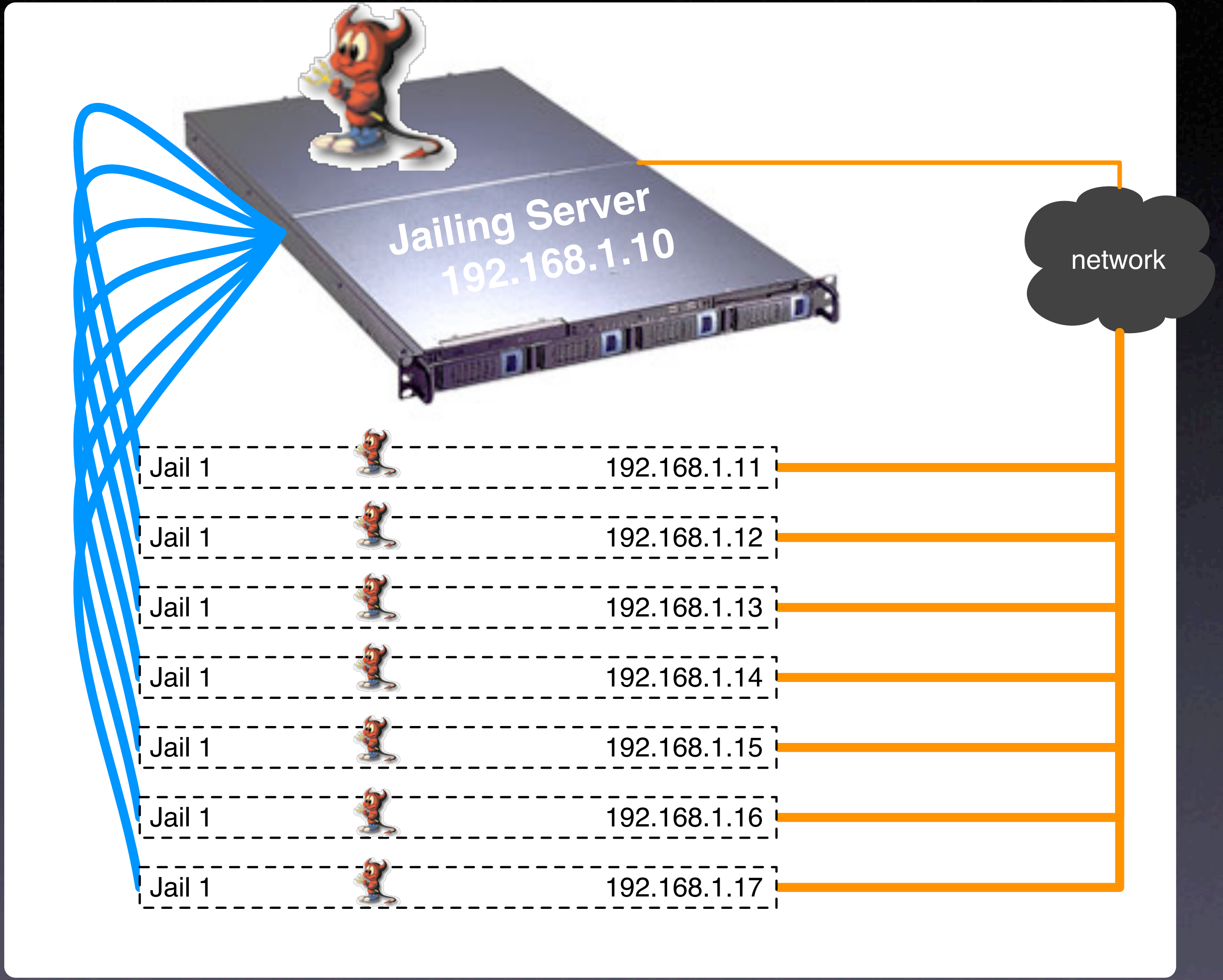
http://www.the-labs.com/FreeBSD/JailTools/

http://www.bpfh.net/simes/computing/chroot-break.html

isaac@diversaform.com

Special Thanks:

- **Jon Ringuette** of iMeme, taught me to jail.
- The jail feature was written by **Poul-Henning Kamp** for R&D Associates http://www.rndassociates.com/ who contributed it to FreeBSD.
- **Robert Watson** wrote the extended documentation, found a few bugs, added a few new features, and cleaned up the userland jail environment.



Jailing Server
192.168.1.10

network

Jail 1 192.168.1.11

Jail 1 192.168.1.12

Jail 1 192.168.1.13

Jail 1 192.168.1.14

Jail 1 192.168.1.15

Jail 1 192.168.1.16

Jail 1 192.168.1.17