Saturday, December 11, 2010

Saturday, December 11, 2010

FreeBSD

Berkeley Software Distribution

Jails

Francisco Cabrita aka @include

Porto, 11 Dezembro 2010

Rápida Introdução às FreeBSD Jails

Francisco Alves Cabrita *tcc include*
francisco@nortenet.pt
sufixo.com

Coimbra, 2 de Setembro de 2006

Disclaimer

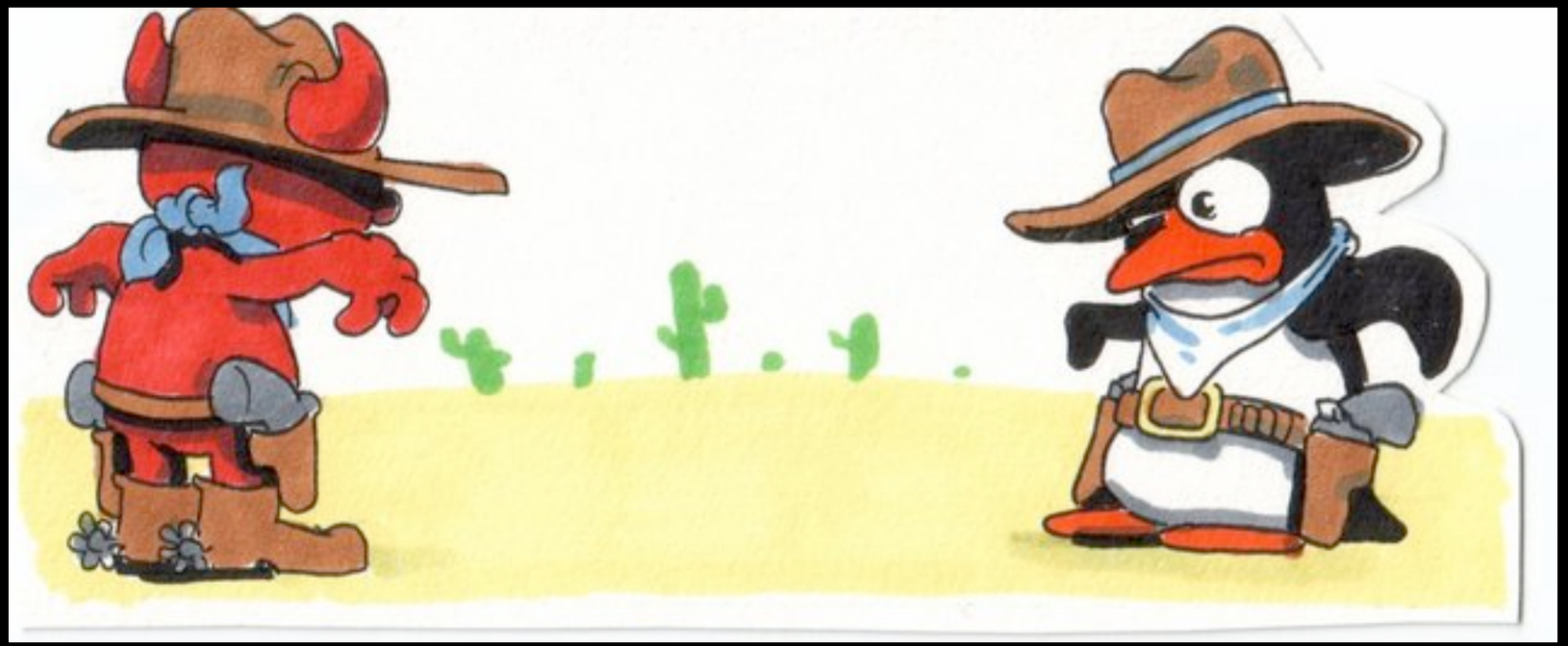
- Provoca dores intensas
- Arruma dois tipos de cada vez

Agenda

- $ whoami
- Um pouco de FreeBSD
- Um pouco de Jails

Apresentação

- Viana Castelo, 31
- Lisboa, 9
- Inf. Gestão, 15d
- Runner, 4
- Brandia 4 anos, bofh
- log 4 anos, bofh
- Vodafone 2 anos, sysadmin
- SAPO 6 meses, *



*BSD is dying

- Anonymous Coward, Slashdot

1993, 1994, 1995, 1996, 1997, 1998, 1999, 2001, 2002,
2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010

FreeBSD


hp
invent


f5
NETWORKS


Adobe


NASA


NOKIA
Connecting People


Google


SIEMENS


nyi


Juniper
NETWORKS


YAHOO!


IBM


McAfee
Proven Security


NORTEL
NORTHERN TELECOM


ORACLE


NetApp

- Unix like - Open Source de Lic. BSD
- Origem 386BSD, reescrito 4.4BSD-Lite ’95
- Monolítico, Multi-processing/threaded
- Kernelland / Userland

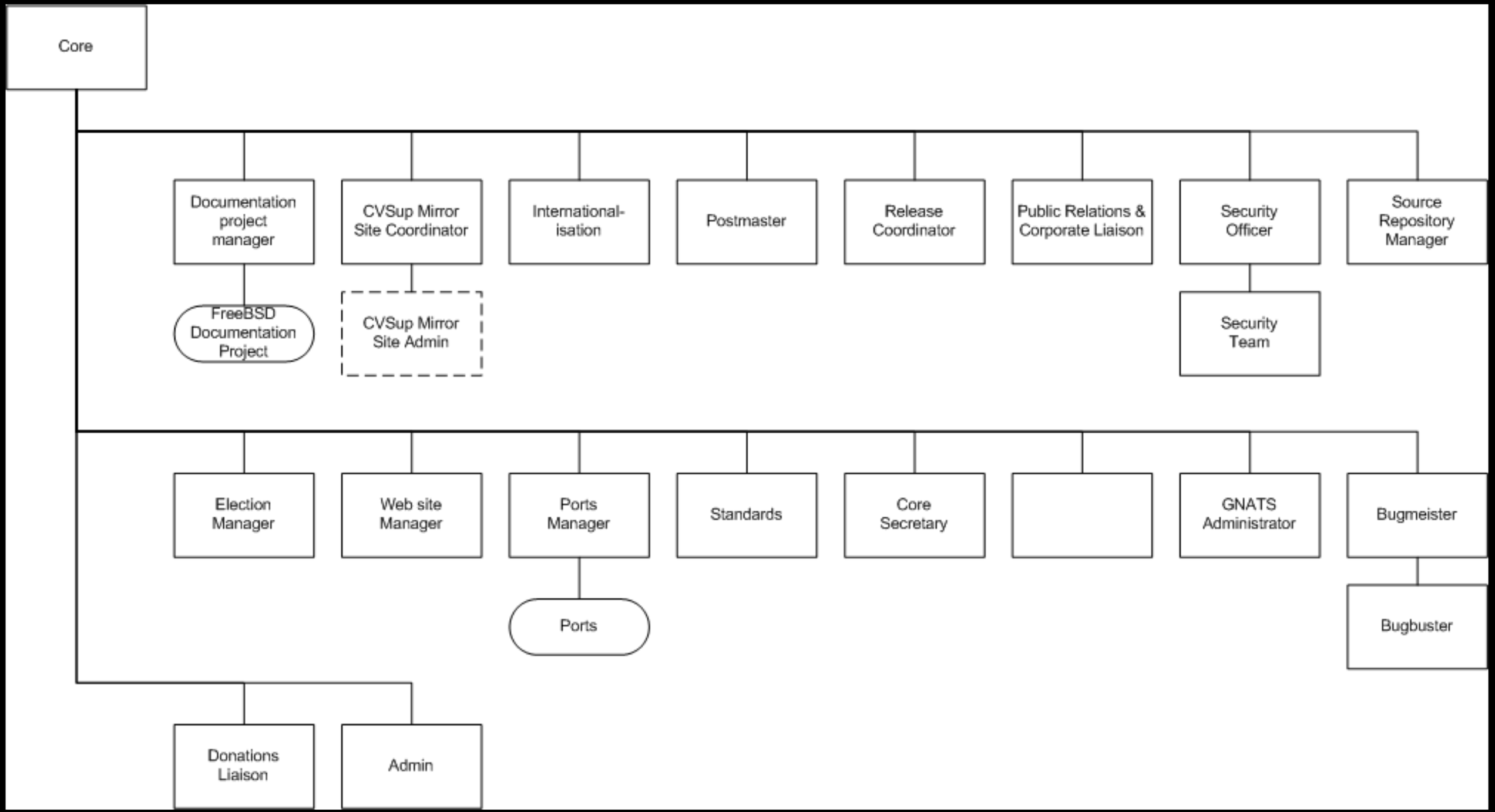
- FFS << UFS
- SCM [http://wiki.freebsd.org/VersionControl]

The FreeBSD Copyright

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

- FreeBSD Kernel, user space
- Security officer, release engineering
- Ports, Packages
- FreeBSD releases
- Manual, Handbook, web pages, marketing
- Tech support, debugguing, etc
- Eventos
- Mailing lists (~100)
- Clusters



http://www.freebsd.org/doc/en/books/dev-model/book.html
http://7n4q.sl.pt

- Source developers
- Core team
- Release Eng Team
- Release Eng Build Team
- Security Officer
- Security Team
- Ports Team
- Ports Managers
- Doceng Team
- Documentation Team
- Vendor Relations Team
- Foundation Board of Directors
- Foundation Operations Manager
- FreeBSD.org admins@
- FreeBSD.org webmaster
- Cluster admins
- Mirrors Team
- Donations Team
- Marketing Teams
- Perforce Admins
- CVSUP Mirrors Team
- Postmaster

- CVS Admins
- Perforce Contributers
- FreeBSD GNOME Project
- FreeBSD KDE Project
- Mono on FreeBSD
- OO on FreeBSD
- Java on FreeBSD
- FreeBSD Standards
- FreeBSD Tinderbox
- SoC Mentors
- Stress Testing
- Monthly Status Reports
- Coverity Team
- KAME Project
- TrustedBSD Project
- PC-BSD
- DesktopBSD
- FreeSBIE
- FreeNAS
- PfSense
- DrangonflyBSD

Events

- USENIX ATC
- BSDcan
- BSDcon
- EuroBSDcon
- AsiaBSDcon
- NYCBSDcon
- MeetBSD
- BSDconTR

s/PortoLinux/PortoBSD

Flavours & Distros

- FreeBSD
- NetBSD
- OpenBSD
- DragonflyBSD

- FreeNAS
- pfSense
- PC-BSD
- DesktopBSD
- FreeSBIE
- PicoBSD

Jails

Poul-Henning Kamp

- http://docs.freebsd.org/44doc/papers/jail/jail.html
- http://www.freebsd.org/doc/en_US.ISO8859-1/books/handbook/jails.html
- http://www.freebsd.org/cgi/man.cgi?query=jail&sektion=8
- http://en.wikipedia.org/wiki/FreeBSD_Jail
- http://wiki.freebsd.org/Jails
- http://sufixo.com/articles/jails_barcamp06.pdf
- http://sufixo.com/articles/jails.pdf

Tipos de Virtualização

HVM/FVM/PVM

Containers

FreeBSD jails?
O que são?

- ‘Máquinas Virtuais’ auto-contidas
- Enclausuramento da Userland
- Prisões / *Containers*

Objectivos

- manter a semântica dos mecanismos de controlo de acesso o que permite manter a compatibilidade com demais aplicações.
- permitir que cada jail tenha o seu root e que as suas actividades estivessem limitadas ao processo, ficheiros e network associadas a ele.

Jails

- ~190Mb
- perspectiva do sysadmin:
  - apenas um host. login via SSH
- perspectiva do host:
  - directório com FreeBSD minimalista
  - processos marcados com J

chroot on steroids

- está restrita ao seu processo e filhos
- Não pode aceder a processos no host
- Hardware não é emulado
- Não há um kernel na Jail
- Partilha recursos com o host

steroids e mais

- Multiplos IPs (IPv4/v6/no-IP) por Jail
- Firewalling por Jail
- Jails hierarquicas
- Correr Linux numa Jail ‘autch’

Para que servem I

- Poucos recursos físicos
- Labs / Ambientes de desenvolvimento
- Hosting
- Package building
- Sistemas homogéneos

Para que servem 2

- Reutilização de Hardware
- Flexibilidade e rapidez na implementação de servidores
- Administração contextualizada
- Segurança e Privacidade

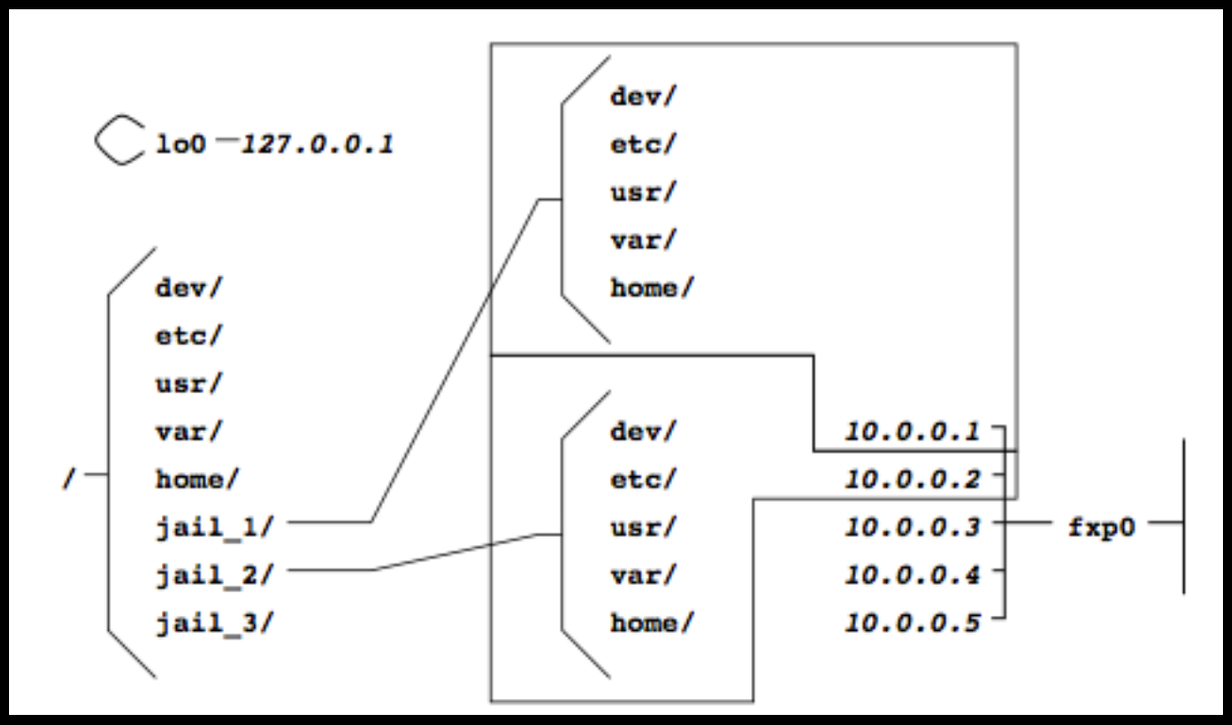
Como?

Simples

```
$ ifconfig NOME_INTERFACE inet alias IP_DA_JAIL netmask 0xffffffff
$ mkdir -p /usr/jails/NOME_DA_JAIL/tmp/build
$ cd /usr/src
$ make -j4 world DESTDIR=/usr/jails/NOME_DA_JAIL
$ cd etc
$ make distribution DESTDIR=/usr/jails/NOME_DA_JAIL

$ cd /usr/jails/NOME_DA_JAIL
$ touch etc/fstab ; echo "hostname=\"NOME\"" >> etc/rc.conf
$ echo "WRKDIRPREFIX=/tmp/BUILD" >> etc/make.conf
$ echo "sshd_enable="YES" >> etc/rc.conf
$ echo "ListenAddress IP_DA_JAIL" >> etc/ssh/sshd_config
```

Start


lo0 — 127.0.0.1
/
dev/
etc/
usr/
var/
home/
jail_1/
jail_2/
jail_3/
dev/
etc/
usr/
var/
home/
dev/
etc/
usr/
var/
home/
10.0.0.1
10.0.0.2
10.0.0.3
10.0.0.4
10.0.0.5
fxp0

config

```
$ vi /etc/rc.conf.local

jail_enable="YES"
jail_set_hostname_allow="NO"
jail_socket_unixiproute_only="YES"
jail_sysvipc_allow="YES"
jail_stop_jailer="NO"
jail_list="WEBSERVER1 NOME DA JAIL MAIS OUTRA JAIL"
jail_WEBSERVER1_rootdir="/usr/jails/WEBSERVER1"
jail_WEBSERVER1_hostname="webserver1"
jail_WEBSERVER1_ip="10.1.1.50"
jail_WEBSERVER1_exec="/bin/sh /etc/rc"
jail_WEBSERVER1_devfs_enable="YES"
```

Start

Manualmente

$ jail /usr/jails/<PATH> <NOME_JAIL> <IP_JAIL> /bin/sh

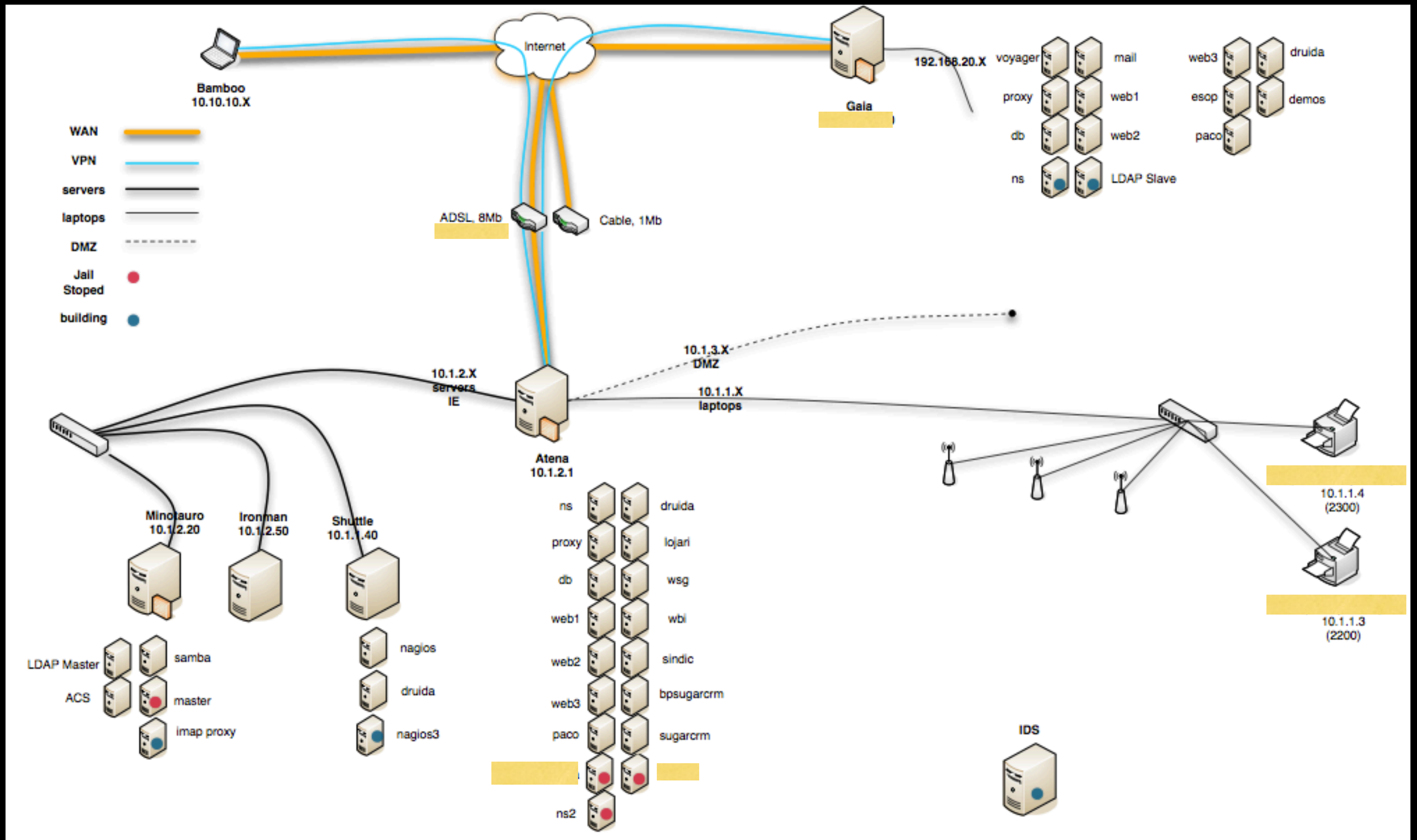
Via rc.d

$ /etc/rc.d/jails start <NOME_JAIL>

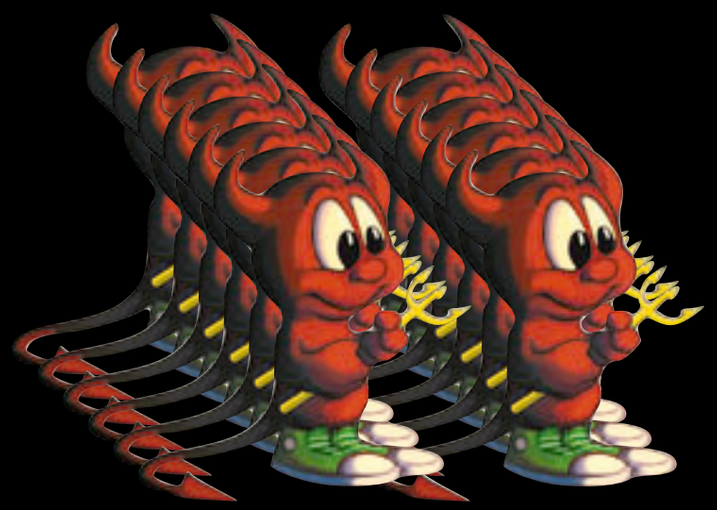
FreeBSD81 [Running]

```
[root@freebsd /etc/rc.d]# uname -a
FreeBSD freebsd 8.1-RELEASE FreeBSD 8.1-RELEASE #0: Mon Jul 19 02:55:53 UTC 2010
    root@almeida.cse.buffalo.edu:/usr/obj/usr/src/sys/GENERIC  i386
[root@freebsd /etc/rc.d]# jls
   JID  IP Address      Hostname                      Path
     5  127.0.0.2       web01                         /usr/jails/web01
[root@freebsd /etc/rc.d]# ps aux | grep 'J'
root  41601  0.0  0.2  3348  1016  ??  IsJ  10:39AM   0:00.00 /usr/sbin/syslogd
root  41722  0.0  0.6  6704  3128  ??  IsJ  10:39AM   0:00.00 /usr/sbin/sshd
root  41733  0.0  0.2  3376  1188  ??  IsJ  10:39AM   0:00.01 /usr/sbin/cron -J
root  41780  0.0  0.2  1820   824  v0  R+   10:41AM   0:00.01 grep J
[root@freebsd /etc/rc.d]# jexec 5 uname -a
FreeBSD web01 8.1-RELEASE FreeBSD 8.1-RELEASE #0: Mon Jul 19 02:55:53 UTC 2010
    root@almeida.cse.buffalo.edu:/usr/obj/usr/src/sys/GENERIC  i386
[root@freebsd /etc/rc.d]# jexec 5 ps aux
USER    PID %CPU %MEM    VSZ   RSS  TT  STAT STARTED      TIME COMMAND
root  41601  0.0  0.2   3348  1016  ??  IsJ  10:39AM   0:00.00 /usr/sbin/syslogd
root  41722  0.0  0.6   6704  3128  ??  IsJ  10:39AM   0:00.00 /usr/sbin/sshd
root  41733  0.0  0.2   3376  1188  ??  SsJ  10:39AM   0:00.01 /usr/sbin/cron -J
root  41782  0.0  0.2   3428   956  v0  R+J  10:42AM   0:00.02 ps aux
[root@freebsd /etc/rc.d]#
```

Left ⌘


Internet
Bamboo
10.10.10.X
Gala
192.168.20.X
voyager
mail
proxy
web1
db
web2
ns
LDAP Slave
web3
druida
esop
demos
paco
WAN
VPN
servers
laptops
DMZ
Jail
Stoped
building
ADSL, 8Mb
Cable, 1Mb
10.1.3.X
DMZ
10.1.2.X
servers
IE
10.1.1.X
laptops
Atena
10.1.2.1
10.1.1.4
(2300)
10.1.1.3
(2200)
Minotauro
10.1.2.20
Ironman
10.1.2.50
Shuttle
10.1.1.40
LDAP Master
samba
ACS
master
imap proxy
nagios
druida
nagios3
10.1.2.1
ns
druida
proxy
lojari
db
wsg
web1
wbi
web2
sindic
web3
bpsugarcrm
paco
sugarcrm
ns2
IDS

Jails++



- children.max (default 0)
- children.cur
- parent (0 para top)
- cpuset.id
- ipv6
- persist

Porreiro pá

- Memory Devices
- 0..N IPs p/ Jail
- unionfs
- ZFS
- VIMAGE
- Limite número processos p/ Jail

Austin Chuck

Internals I

- jail(2) system call
  - struct prison é alocada e populada com argumentos
  - é linkada à struct proc executada
  - struct prison ref counter++
  - chroot define root path

http://fxr.watson.org/fxr/source/kern/kern_jail.c#L96

Internals 2

- struct prison não pode ser alterado assim que criado
- Utilizados hooks no código da criação/remoção de processos que mantêm ref counters da struct prison; são libertados no fim
- Cada novo processo na jail, herda a ref para a struct prison da jail criadora. Isto põe o processo na mesma jail.

Internals 3

- Alterações ao Kernel
  - interfaces de report de processos
  - sinais entre processos
  - networking e shared mem

locks I

- kldload
- modify sysctls
- u/mount
- criar mkdev
- alterar securelevels
- alterar confs de networking

unlocked, mas...

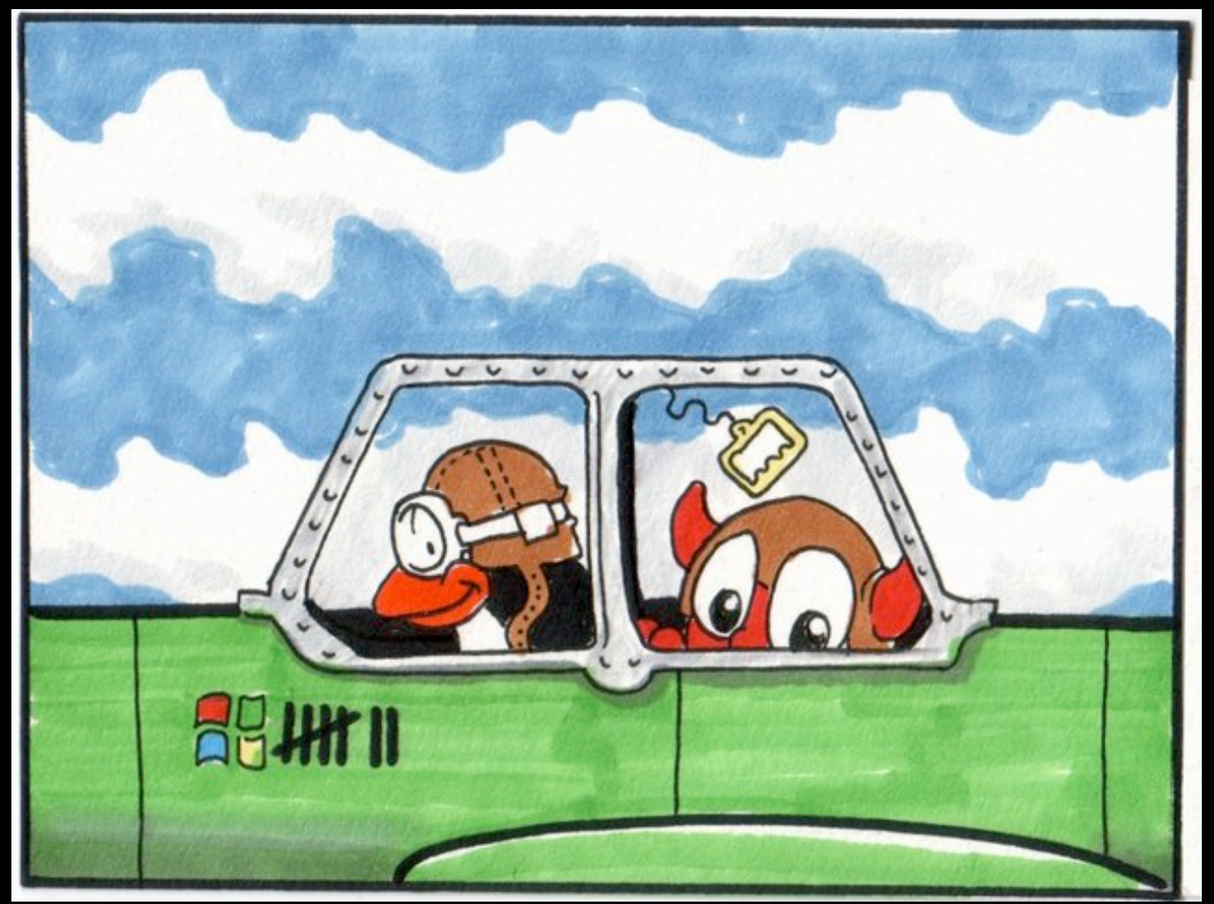
- signals entre processos da mesma jail
- chown/chmod files na jail
- binding

Limitações

- scheduling resources

Hoje não :)

http://www.debian.org/ports/kfreebsd-gnu/



http://phaq.phunsites.net/2007/01/06/debian-gnukfreebsd-inside-native-freebsd-jail/

Obrigado!!