

# FreeBSD Jails Lightweight Virtualization?

Αλέξανδρος Κοσιάρης alex@noc.grnet.gr

Πανεπιστήμιο Πειραιά - Software Libre Society

17 Σεπτεμβρίου 2010

Το κλασσικό μοντέλο ασφαλείας (DAC) του UNIX κάνει δύσκολη τη ζωή του διαχειριστή

Το κλασσικό μοντέλο ασφαλείας (DAC) του UNIX κάνει δύσκολη τη ζωή του διαχειριστή

- Ανθρώποι/ομάδες που δεν έχουν σχέσεις εμπιστοσύνης μεταξύ τους (ή ακόμη και με τον ίδιο)

Το κλασσικό μοντέλο ασφαλείας (DAC) του UNIX κάνει δύσκολη τη ζωή του διαχειριστή

- Ανθρώποι/ομάδες που δεν έχουν σχέσεις εμπιστοσύνης μεταξύ τους (ή ακόμη και με τον ίδιο)
- Συνεχής ανάγκη αναπροσαρμογής των κανόνων από τον διαχειριστή

Το κλασσικό μοντέλο ασφαλείας (DAC) του UNIX κάνει δύσκολη τη ζωή του διαχειριστή

- Ανθρώποι/ομάδες που δεν έχουν σχέσεις εμπιστοσύνης μεταξύ τους (ή ακόμη και με τον ίδιο)
- Συνεχής ανάγκη αναπροσαρμογής των κανόνων από τον διαχειριστή
- Λιγοστές δυνατότητες ανάθεσης δικαιωμάτων σε άλλους χρήστες

Το κλασσικό μοντέλο ασφαλείας (DAC) του UNIX κάνει δύσκολη τη ζωή του διαχειριστή

- Ανθρώποι/ομάδες που δεν έχουν σχέσεις εμπιστοσύνης μεταξύ τους (ή ακόμη και με τον ίδιο)
- Συνεχής ανάγκη αναπροσαρμογής των κανόνων από τον διαχειριστή
- Λιγοστές δυνατότητες ανάθεσης δικαιωμάτων σε άλλους χρήστες
- Πρακτικά δύο είδη χρηστών: ο root, και... οι υπόλοιποι

Έχουν προταθεί και υλοποιηθεί διάφορες λύσεις

- `sudo`
- `chroot` (2)
- Filesystem ACLs
- Mandatory Access Control (SELinux, AppArmor, TrustedBSD, RSBAC, Solaris roles,...)
- Virtualization...

Έχουν προταθεί και υλοποιηθεί διάφορες λύσεις

- sudo
- chroot (2)
- Filesystem ACLs
- Mandatory Access Control (SELinux, AppArmor, TrustedBSD, RSBAC, Solaris roles,...)
- Virtualization...

Κάποια ασχολούνται με να κλειδώσουν κάπου(και σε κάποια επίπεδα) τους χρήστες δίνοντας τους μόνο τα απαραίτητα, άλλα προσθέτουν δυνατότητες στο DAC ώστε να είναι πιο εύκολος ο έλεγχος και κάποια επιλέγουν να τους πετάξουν σε ένα άλλο παράλληλο και μακρινό σύμπαν



## Ακόμη μία λύση!

### Enter FreeBSD Jails

Η ιδέα προτάθηκε από τον Poul-Henning Kamp και τον Robert Watson το 2000, και εμφανίστηκε για πρώτη φορά(υλοποιημένη από τον rhk) στο FreeBSD 4.0

# Τι είναι τα Jails;

## FreeBSD jail είναι:

Μία εξέλιξη του chroot(2) syscall ώστε να υπάρχει διαχωρισμός μίας διεργασίας και των παιδιών της όχι μόνο στο επίπεδο του filesystem αλλά και στα επίπεδα των διεργασιών και δικτύου

Μία διεργασία μέσα σε ένα jailed περιβάλλον δεν μπορεί:

- Να δει ή αλληλεπιδράσει με διεργασίες εκτός του jail
- Να μεταβάλλει με οποιονδήποτε τρόπο τον πυρήνα
- Να προσαρτήσει(mount) άλλα filesystems
- Να δει άλλο μέρος του filesystem πέρα του δικού του
- Να δημιουργήσει device nodes(aka mknod)
- Να αλλάξει/χρησιμοποιήσει δικτυακούς πόρους

Ένα jail χρειάζεται:

- Ένα FreeBSD filesystem hierarchy
- Μία IP διεύθυνση(μοιρασμένη με τον host)(optional)
- Μία διεργασία που θα τρέξει στον jailed χώρο

Εάν το αρχικό πρόγραμμα που θα τρέξει είναι το /etc/rc μπορούμε να σηκώσουμε ένα ολόκληρο Virtual FreeBSD σύστημα

- Εξαιρετικά φτηνό virtualization - Ελάχιστη χρήση σε δίσκο και κυρίως μνήμη

- Εξαιρετικά φτηνό virtualization - Ελάχιστη χρήση σε δίσκο και κυρίως μνήμη
- Πλήρης διαχωρισμός των διαφόρων jails - χρηστών/ομάδων

- Εξαιρετικά φτηνό virtualization - Ελάχιστη χρήση σε δίσκο και κυρίως μνήμη
- Πλήρης διαχωρισμός των διαφόρων jails - χρηστών/ομάδων
- Προσβάσιμο από τον host το guest σύστημα
  - Διαχείριση
  - Αναβάθμιση
  - Παρακολούθηση

- Κοινό networking stack για όλα τα jails (και τον host)

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota



- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)
  - Δικτύου

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)
  - Δικτύου
- Λειψό networking stack

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)
  - Δικτύου
- Λειψό networking stack
  - Μία IPv4 διεύθυνση per jail ( για <FreeBSD 8.0)

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)
  - Δικτύου
- Λειψό networking stack
  - Μία IPv4 διεύθυνση per jail ( για <FreeBSD 8.0)
  - No IPv6 ( για < FreeBSD 8.0)

- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)
  - Δικτύου
- Λειψό networking stack
  - Μία IPv4 διεύθυνση per jail ( για <FreeBSD 8.0)
  - No IPv6 ( για < FreeBSD 8.0)
  - Το localhost είναι βραχυκυκλωμένο. Προβλήματα με TCP wrappers, Application Level ACLs(πχ apache)



- Κοινό networking stack για όλα τα jails (και τον host)
- Έλλειψη εύκολα διαχειρίσιμων μηχανισμών quota
  - Δίσκου
  - Μνήμης
  - CPU (Τα CPU Sets διορθώνουν αυτό μερικώς στο FreeBSD 7.1)
  - Δικτύου
- Λειψό networking stack
  - Μία IPv4 διεύθυνση per jail ( για <FreeBSD 8.0)
  - No IPv6 ( για < FreeBSD 8.0)
  - Το localhost είναι βραχυκυκλωμένο. Προβλήματα με TCP wrappers, Application Level ACLs(πχ apache)
  - No firewall support

- Bind όλες οι υπηρεσίες στην μία IP,IPv6 του host
  - sendmail\_enable="NO"
  - inetd\_flags="-wW -a 10.10.10.10"
  - rpcbind\_enable="NO"
  - syslogd\_flags="-s -b 10.10.10.10"
  - sshd\_config ListenAddress

# Προετοιμασία Host (2)

```
1 D=/here/is/the/jail
2 cd /usr/src
3 mkdir -p $D
4 make world DESTDIR=$D
5 make distribution DESTDIR=$D
6 mount -t devfs devfs $D/dev
7 jail $D myhostname 10.10.10.20 /etc/rc
```

## Εργαλεία

- **jls** - Δείχνει τα running jails
- **jexec** - Με argument το jailid (από την jls) και ένα command(πχ tcsh) μπαίνεις στο jail

Ένα νέο jail είναι σαν να έχει γίνει μόλις installed.  
Χρειάζεται:

- Ένα άδειο fstab
- Disabled τον portmapper (rpcbind\_enable="NO" στο /etc/rc.conf")
- /etc/resolv.conf
- newaliases για να μην γκρινιάζει το sendmail
- network\_interfaces="" στο /etc/rc.conf
- Root password
- Timezone

Και είμαστε έτοιμοι.

Για να ξεκινάνε τα jails on host boot βάζουμε περίπου τα εξής στο `/etc/rc.conf`

```
1 jail_set_hostname_allow=""NO
2 jail_enable="YES"
3 jail_list="webserver ldapserver"
4 jail_interface="rl0"
5 jail_devfs_enable="YES"
6 jail_procfs_enable="NO"
7 jail_devfs_ruleset="devfsrules_jail"
8 jail_webserver_rootdir="/usr/jails/webserver"
9 jail_webserver_hostname="webserver"
10 jail_webserver_ip="10.0.0.20"
11 jail_ldapserver_rootdir="/usr/jails/ldapserver"
12 jail_ldapserver_hostname="ldapserver"
13 jail_ldapserver_ip="10.0.0.30"
```

Σχετικά παλιό εργαλείο. Επιτρέπει εύκολα:

- Δημιουργία/καταστροφή
- Εκκίνηση/τερματισμό
- Αναβάθμιση/backup/restore

Μειονεκτήματα:

- Το πρώτο jail φτιάχνεται χειροκίνητα
- Σχετικά υψηλή χρήση δίσκου(σε σχέση με άλλα εργαλεία)
- Non-maintained

Νεώτερο εργαλείο. Ακόμη αναπτύσσεται. Επιτρέπει:

- Binary εγκατάσταση/αναβάθμιση/καταστροφή
- Πολύ φτηνά jails σε χώρο(2MB)
- Εύκολη αναβάθμιση
- Flavours(πχ FAMP, LDAP κτλ)
- ZFS jails
- Jail specific Routing Tables ( μέσω του setfib(2) )
- CPUSets

Ένα μειονέκτημα: Δεν ξεκινά αυτόματα τα jails με δικτύωση. Υπάρχουν trivial patches αλλά δεν γίνονται δεκτά upstream. Ευτυχώς το patch management των FreeBSD ports δουλεύει υπέροχα.

## VIMAGE

VIMAGE(Experimental): Ένας νέος τύπος container. Πρόκειται για ένα jail με virtualized networking stack

Virtualized:

- IPv4,IPv6
- NFS
- IPFW / PF firewalls,
- BPF, raw / routing sockets...



Χρησιμοποιείται από:

- PostgreSQL
- eUKhost
- RootBSD
- InetServices
- NTUA NOC