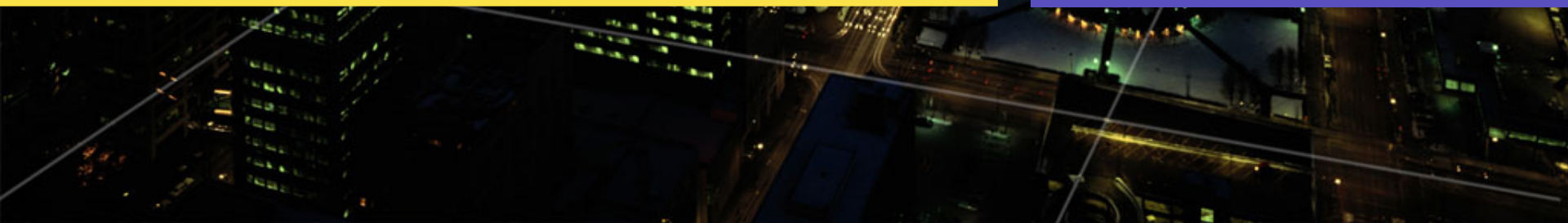
Solaris 10

Fischer Erik

Zsemlye Tamás

Módly Zoltán

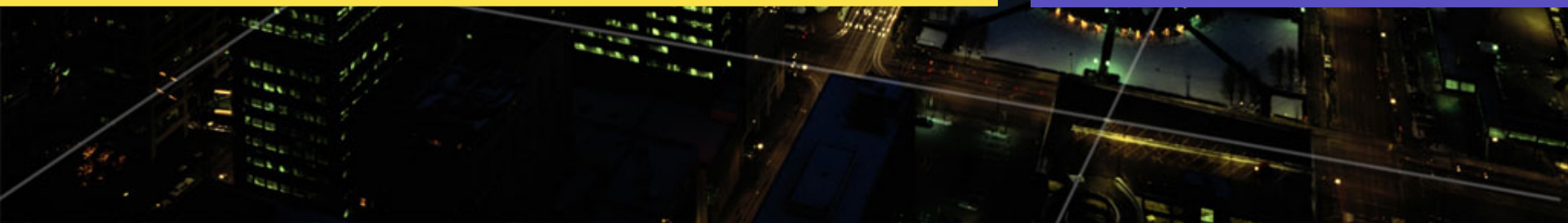
(A filmben való megjelenés sorrendjében.)



Menü

- Bevezető
- ZFS file rendszer
- DTrace
- Zónák
- Privilege rendszer
- BART
- Egyéb security gyöngyszemek
- SMF

Bevezető



Solaris 10

- Minden idők egyik legkomplexebb operációs rendszere és projektje a Sun-nál
- A feladat mérete meghaladja a Solaris 7-et (az első 64-bites kernel)
- 700+ alprojekt
- Ingyenes!!! (Nincs támogatás)
- OpenSolaris

Nagyobb projektek

- JDS 3 integráció
- 64-bites AMD64/EM64T kernel
- Linux emulációs alrendszer
- Új TCP/IP stack - FireEngine (ez nem túl aprócska)
- 64k page base size

És kisebbek...

- IP6to4 router
- USB szalagmeghajtók
- Anoním memória lapméret kontroll
- Aszinkron socket-ek
- CMT ütemező
- Koherens konzol
- Új CPU performancia számláló projekt

És kisebbek...

- Dinamikus erőforrás pool-ok
- Felhasználó szintű mount
- Oracle DHCP modul
- File event monitoring
- FS-VM szeparálás
- Teljes fizikai memória HAT mapping
- GNU Zebra (OSPF, RIP, BGP)
- Interrupt targetting

És kisebbek...

- iSCSI hooks
- Kernel cage megszüntetése
- Kernel crypto framework
- SVM javítások (multiTB metadevice, resource configuration)
- Virtualizált Layer 2 networking

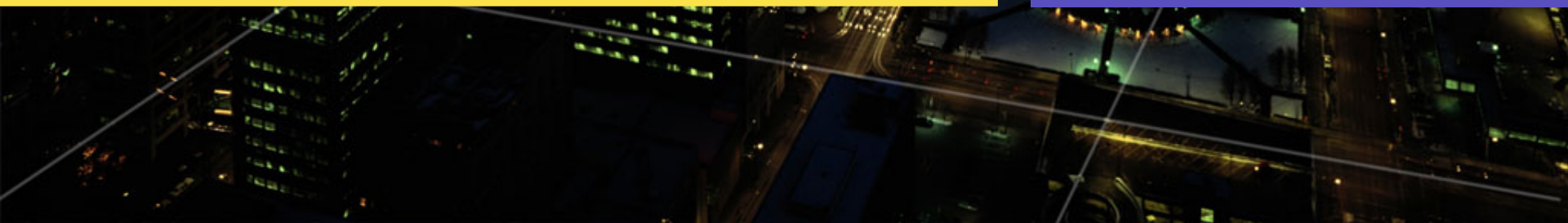
És kisebbek...

- Memory DR
- MPxIO integráció
- MultiTB UFS és diszk
- NCA multiszerver
- NFSv4
- NIS/NIS+2LDAP
- NUMA optimalizáció

És kisebbek...

- PAM bővítések
- Packet filter hook
- rcapd – fizikai memória menedzsment
- Wheel mouse
- USB mass storage
- WAN boot

ZFS



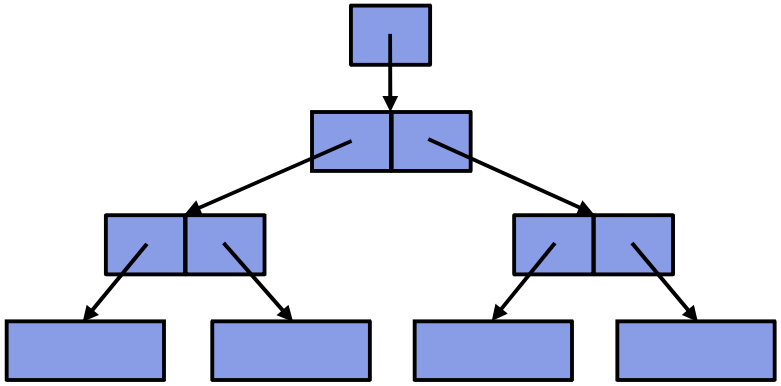
Sun microsystems

ZFS

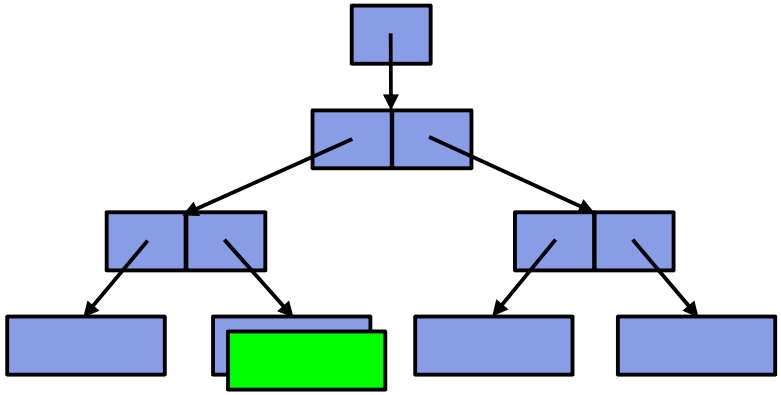
- A diszk VM-je, nincs szükség kötet kezelőre
- Tranzakcionális
- Teljesen integritás védett
- 128-bites (256 quadrillió zettaB)
- Dinamikus metaadatok
- Beépített tömörítés, titkosítás
- Diszk scrubbing
- Replikáció
- Minden művelet copy-on-write
- Gyors snapshot, verziózás

COW

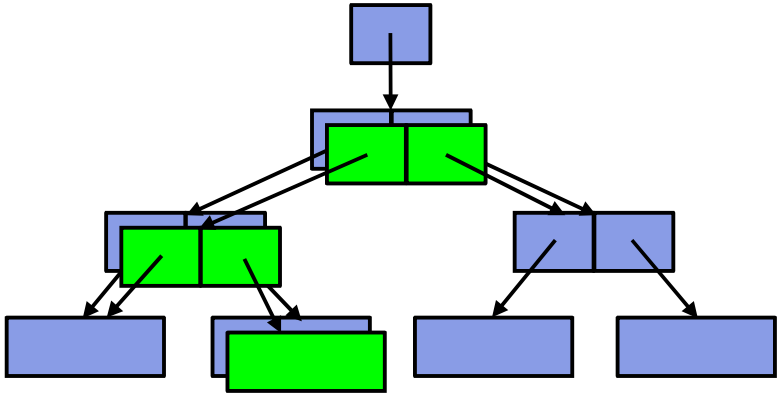
1. Az eredeti



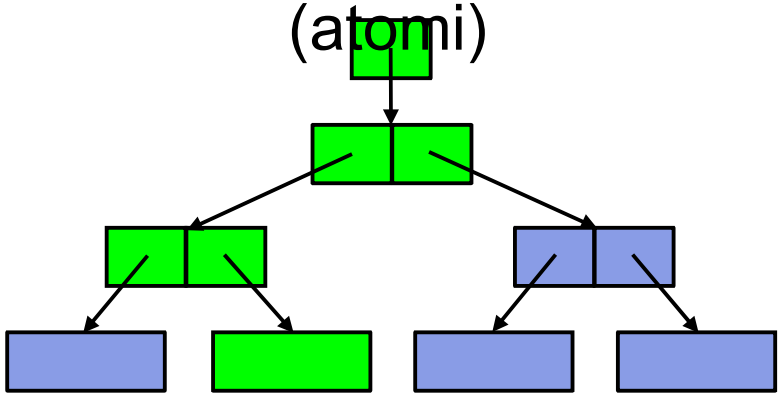
2. Adat blokk COW



3. Indirekt blokk COW

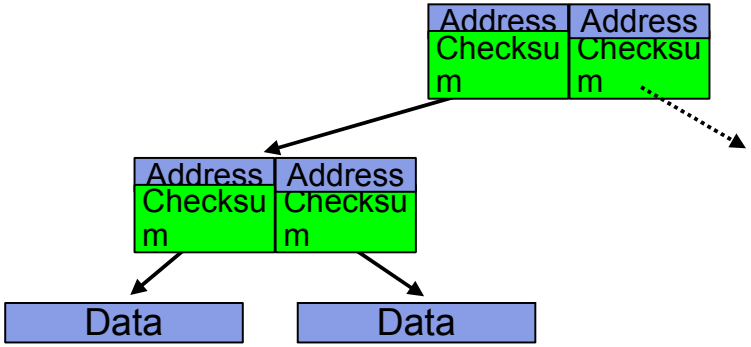


4. Überblock módosítás (atomi)

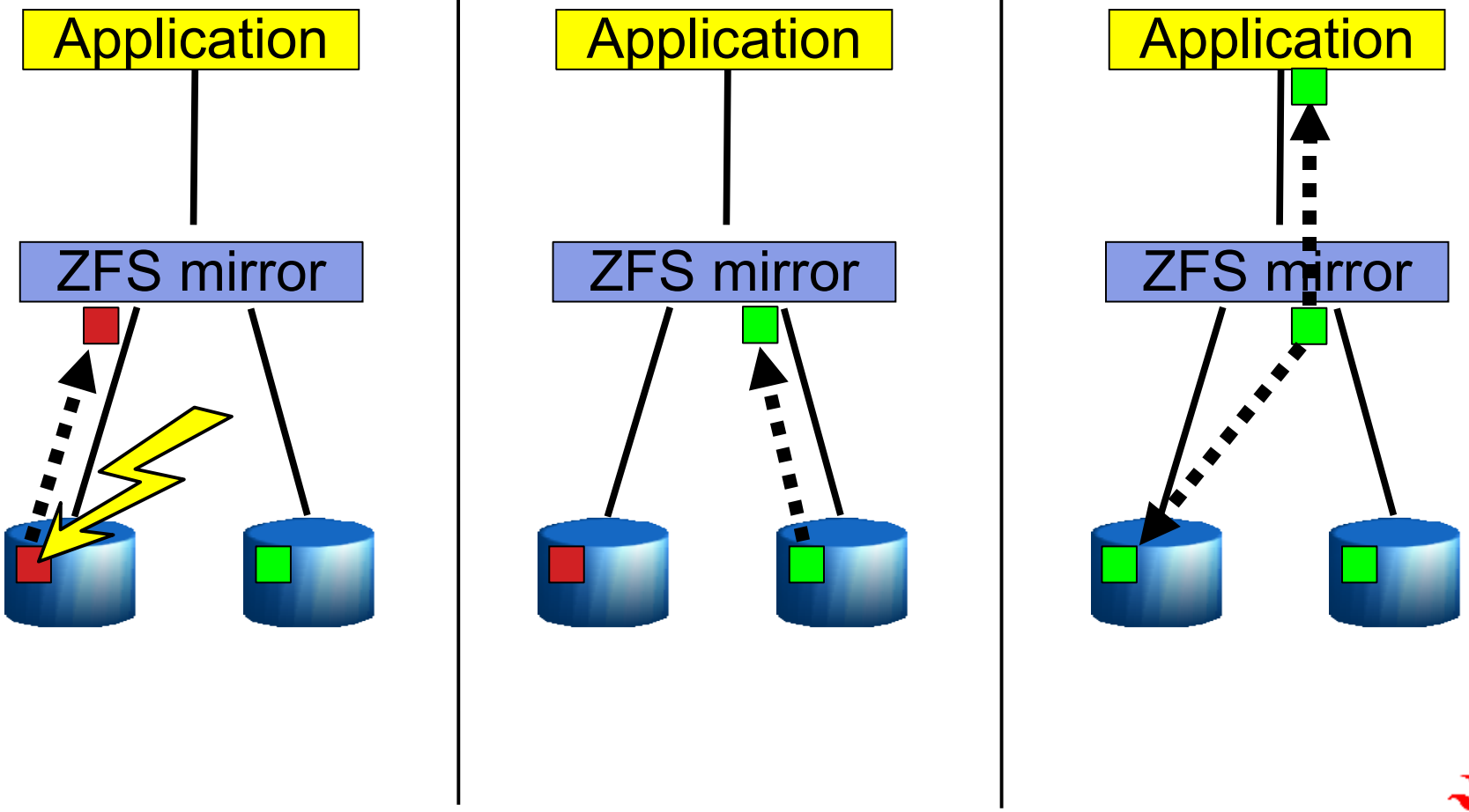


Checksum

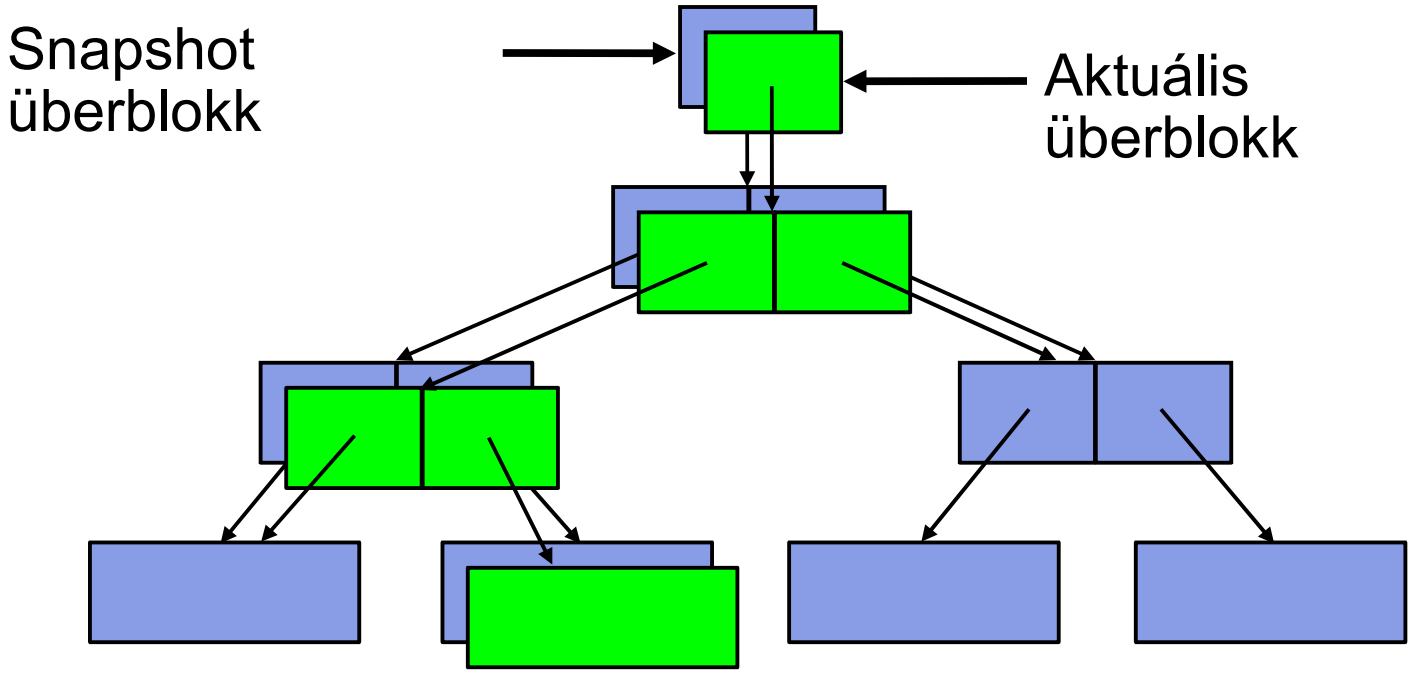
- Bit rot
- Phantom writes
- Misdirected reads and writes
- DMA parity errors
- Driver bugs
- Accidental overwrite



Önjavítás


Application
ZFS mirror
Application
ZFS mirror
Application
ZFS mirror

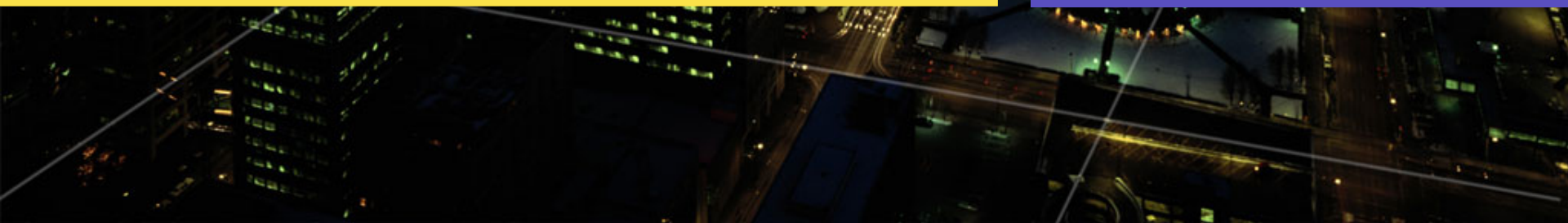
Snapshot



Parancsok

- zpool
  - A kötetkezelő
  - create, destroy, add, import, export, df, iostat, vdevs, devices, config
- zfs
  - A fájlrendszer kezelő
  - create, destroy, rollback, ls, mount, unmount
- zvol
  - Emulációs kötetkezelő
  - create, destroy, ls

DTrace



Hibakeresés – in vitro

- Végzetes, nem reprodukálható hibák
  - Core dump
  - Postmortem mdb(1), dbx(1)
- Tranziens hibák
  - Programozási hibák vagy percepcionális teljesítmény problémák
    - Ad hoc technikák vagy szinte semmi
    - truss(1), mdb(1), *stat(1M), sar(1), prstat(1)
    - Sun Studio Performance Analyzer

Hibakeresés – in vivo

- Invaziv technikák
  - Bináris instrumentálás
  - Forrás szintű instrumentálás
  - Interposer library-k
  - Debug library-k
  - Debug kernel
  - Általában durva beavatkozások
  - Általában lassú
  - Általában nagy az additív hibák injekciójának esélye

Elvárások

- Az ideális dinamikus hibakereső rendszer
  - Képes információ gyűjtésére
  - Az adatszerkezetek módosítása
  - Mindezt éles rendszereken is
  - Teljesen biztonságosan
  - Teljesítmény veszteség nélkül

DTrace

- Interpretált probe alapú nyelv, prédikátumokkal és akciókkal
- Dinamikus függvény be- és kilépési pont instrumentációs rendszer
- Jellemzők:
  - User és Kernel rétegben is működik
  - 40000+ probe egy átlagos Solaris 10 rendszeren
  - Alapállapotban csak root-ként működik
  - 3 privilégiumot igényel: dtrace_kernel, dtrace_proc és dtrace_user
  - 410 oldalas dokumentáció

Probe

- Az instrumentáció pontos helye egy hierarchiában (leírója egy 'n'-es)
- Egy provider bocsájtja a rendelkezésünkre
- Minden provider modulokra és egy függvényekre tagolódik
- A probe-nak van neve (általában entry és return)
- dtrace -l

<provider, module, function, probe_name>

Provider-ek

- Szép számban akadnak
  - fbt – szintem minden entry és return a kernelben (~39000 db)
  - syscall – syscall tábla (~450 db)
  - profile
  - lockstat
  - proc
  - sysinfo
  - vminfo
  - sched
  - io
  - fpuinfo
  - sdt (~190 db)
  - mib – TCP/IP-hez kapcsolódó függvények (~430 db)

DTrace fogyasztók

- Gyakorlatilag nincs felső korlátja, a DTrace multiplexel
- dtrace(1M) a command line fogyasztó
- Vannak programozott fogyasztók is

A D nyelv

- C és awk keveréke
- Teljes hozzáférés a kernel C típusaihoz
- Teljes hozzáférés a statisztikákhoz és globális információkhoz
- String támogatás

Szkriptek

- dtrace(1M) támogatja a szkripteket

```
#! /usr/sbin/dtrace –s
```

- Szkript szintaktika

```
provider:module:function:name[,
provider:module:function:name]*
/predicate/
{
 action;
 [action;]*
}
```

Beépített változók

- pid – aktuális processz ID
- execname – az aktuális programnév
- timestamp – a bootolás óta eltelt idő [ns]
- probemod, probefunc, probename – a probe adatai

...

Változók

- Skalár
  - Szokásos típusok (char, short, int, long, long long, float, double, long double)
  - **NINCS FLOAT ARITHMETIKA!**
- Asszociatív array-ek
  - Továbbfejlesztett Perl hash
  - A kulcs egy 'n'-es, de mindig azonos szignatúrájú (!!)
  - Pl. array[execname, timestamp]

Akciók

- trace() – az argumentum értékét a trace pufferbe helyezi
- tracemem() – az argumentumban megadott memóriacímtől adott hosszig ment értékét a trace pufferbe
- stack() – a kernel stack trace-t a trace pufferbe helyezi
- ustack() – a user stack trace-t a trace pufferbe helyezi
- exit() – a DTrace fogyasztót kilépésre kényszeríti

...

Destruktív akciók

- A –w flag-gel engedélyezni kell
- stop() – megállítja az aktuális processzt
- raise() – szignált küld az adott processznek
- panic() – kernel pánik
- chill() – adott időre leállítja a processzt
- copyout() és copyoutstr() – adott változóból adott számú adatot egy memória helyre másol

Intelligens megjelentés

- printf() – trace() a printf(3C) jótulajdonságaival
- printa() – a printf(3C), de aggregációkra

Thread lokális változók

- Azonos név, de per thread adat
- Jelölése: self->
- A nem definiált változóknak nulla az értékük
- Ha nullázunk egy thread lokális változót, avval deallokáljuk (!!)

Aggregációk

- Natív DTrace típus
- Egy aggregáció olyan f(x) függvény, ahol x egy tetszőleges n elemű adathalmaz és:

$$f(f(x_0) \cup f(x_1) \cup \ldots \cup f(x_n)) = f(x_0 \cup x_1 \cup \ldots \cup x_n)$$

- Ilyenek a count, a sum, az avg, a max és a min (pl. a median nem az éppen ezért nincs is ilyen)
- Speciális aggregáció a quantize és az lquantize

Aggregációk 1.

```
syscall::write:entry
{
  @counts[execname] = count();
}
```

```
dtrace: script './cnt.d' matched 1 probe
^C

dtrace                            1
csh                               1
xscreensaver                      2
dtwm                             12
ls                               14
sdtperfmeter                     22
dtterm                           67
gnome-cd                        279
```

Aggregációk 2.

```
syscall::write:entry
{
  self->start=timestamp;
}
syscall::write:return
/self->start/
{
  @counts[execname] = quantize(timestamp-self-
    >start);
  self->start=0;
}
```

Aggregációk 2.

```
dtrace: script './cnt2.d' matched 2 probes
^C

  sdtperfmeter
           value  ------------- Distribution ------------- count
            8192 |                                         0
           16384 |@@@@@@@@@@@@@@@@@@@                      8
           32768 |@@@@@@@@@@@@@@@@@@@@@@@@@                10
           65536 |                                         0

  dtwm
           value  ------------- Distribution ------------- count
            8192 |                                         0
           16384 |@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@    21
           32768 |@@@@@@@@                                 5
           65536 |                                         0

  dtterm
           value  ------------- Distribution ------------- count
            4096 |                                         0
            8192 |@                                        3
           16384 |@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@    61
           32768 |@@@@@@@@@@@@@                            25
           65536 |@@@                                      6
          131072 |                                         0

  gnome-cal
           value  ------------- Distribution ------------- count
            4096 |                                         0
            8192 |                                         1
           16384 |@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ 212
           32768 |@@@                                      15
           65536 |                                         0
```

Normalizálás

- Az aggregációk konstanssal normálhatóak
- Ez az adatokat nem változtatja meg

```
dtrace:::BEGIN
{
 start=timestamp;
}
syscall::write:entry
{
  @counts[execname] = count();
}
dtrace:::END
{
 normalize(@counts, (timestamp - start) / 1000000000);
}
```

Normalizálás

```
dtrace: script './cnt3.d' matched 3 probes
^C

  sdtprocess                                                    0
  sort                                                          0
  dtrace                                                        0
  ps                                                            0
  init                                                          0
  csh                                                           0
  xscreensaver                                                  0
  picld                                                         0
  ls                                                            1
  more                                                          1
  tail                                                          1
  sdtperfmeter                                                  2
  dtwm                                                          2
  dtterm                                                        9
  gnome-cd                                                     27
```

Műveletek normákkal

- clear(@aggr) – adott norma törlése
- trunc(@aggr, n) – adott normált aggregáció első ‘n’ elemének “megtartása”

Pragmák

#pragma D option <name>[=<value>]

- Alapvetően hangolható paraméter beállítások
- Speciális kapcsolók beállítása

#pragma D option destructive

#pragma D option quiet

Spekulatív traceing

- speculation() – Spekulatív nyomkövetés indítása, visszatérési érték egy ID
- speculate(ID) – Minden további akció kimenete a spekulatív pufferbe megy
- commit(ID) – A spekulatív puffer tartalma a fő pufferbe kerül
- discard(ID) – A spekulatív puffer tartalma eldobódik

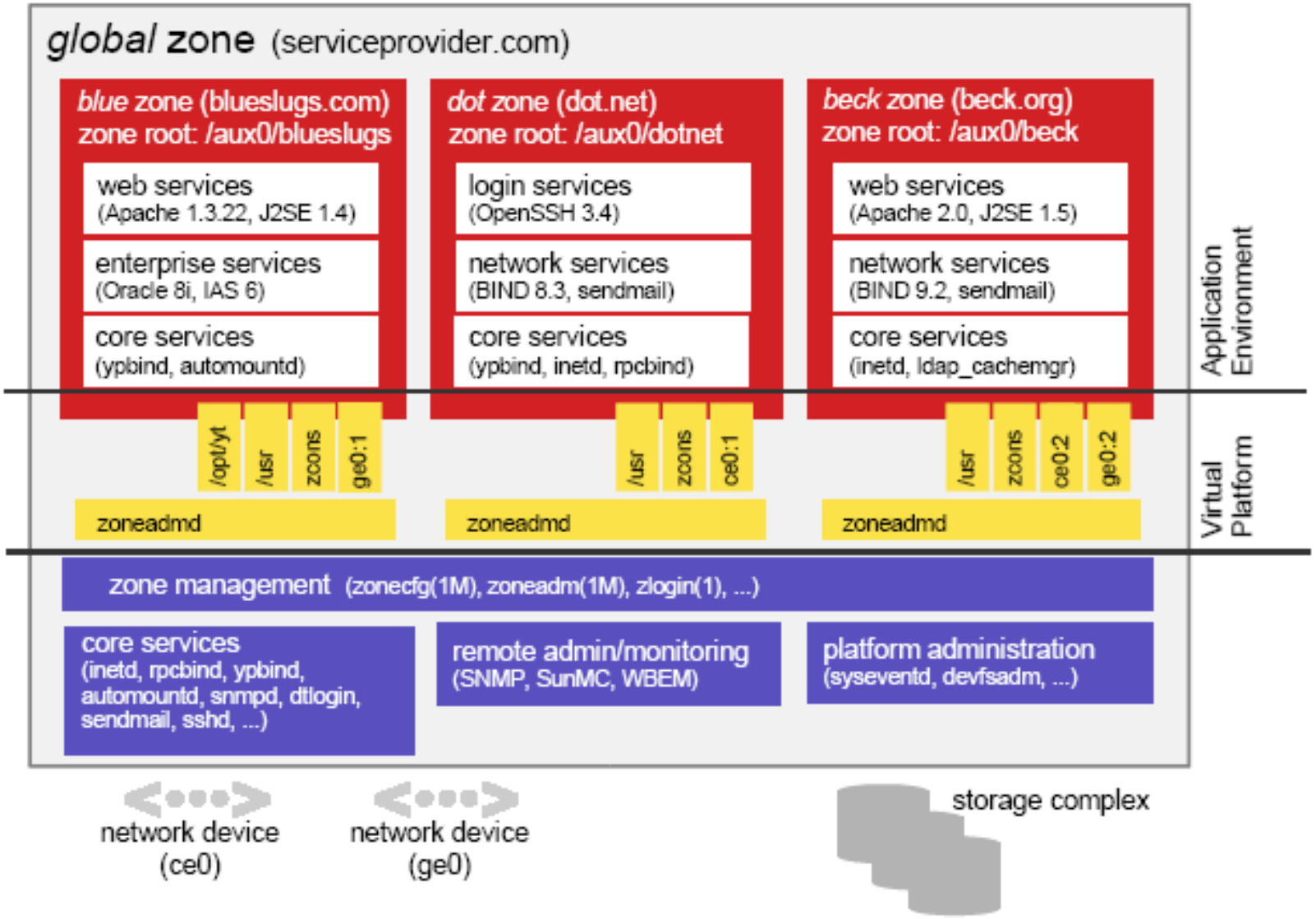
Zónák

Sun microsystems

Zónák

- Egyetlen kernel megosztása több virtualizált applikációs konténer között
- Processz “bedobozolás”
  - erőforrás és biztonsági izoláció
  - láthatóság és kontrol (globális, nem globális)
- Virtualizált hardver (!), de nem virtuális gép
- Kintről önálló operációs rendszernek tűnik

Zónák – blokk diagramm



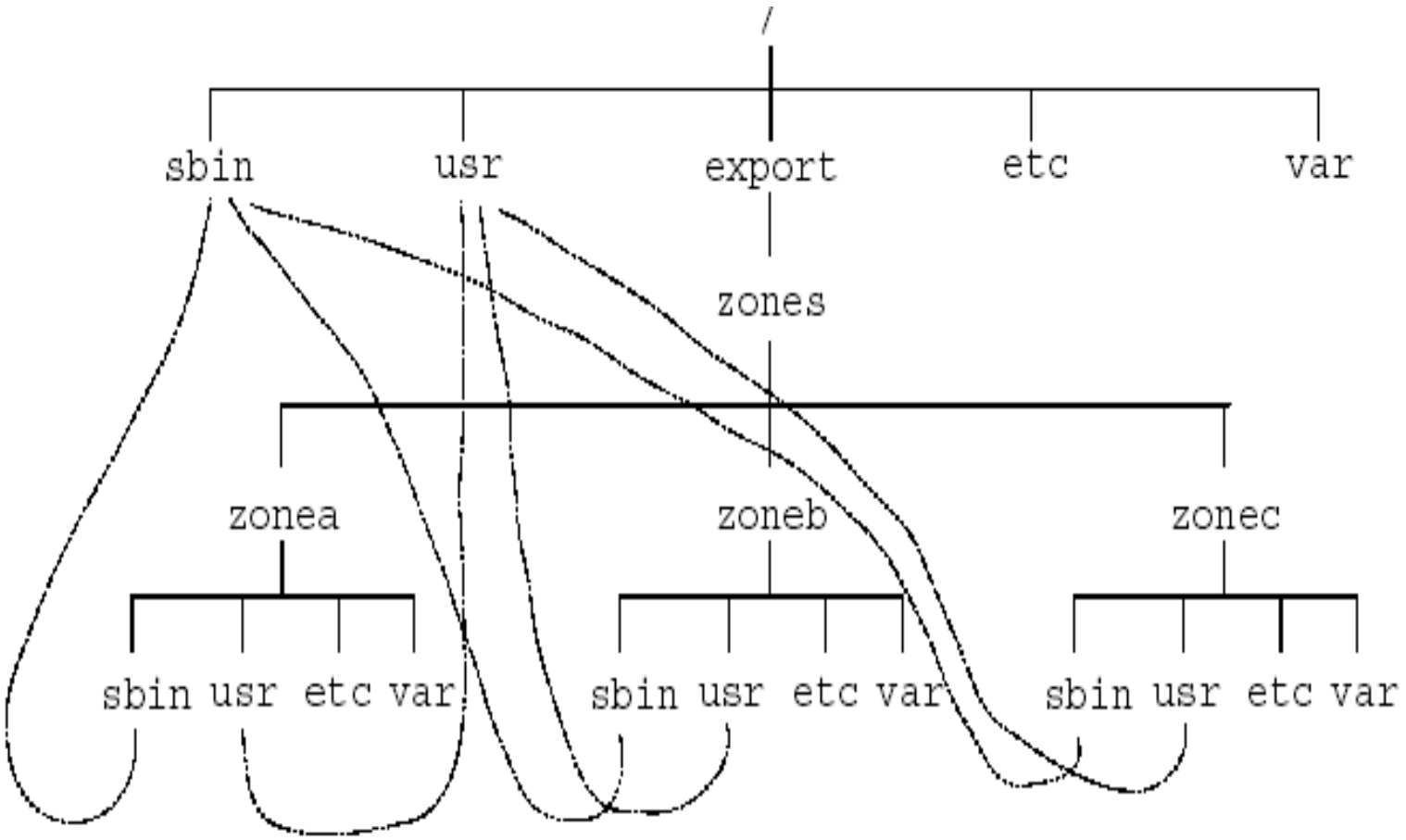
Zónák – virtualizáció

- Rejtett fizikai szint
  - /devices, mknod csak globális zónában
  - /dev, loopback és konstruált (zero, log, ..)
- Önálló IP cím(ek), multiplexelt kommunikáció
  - hme0, hme0:1 (globális zóna által definiálva, plumb csak globális zónában)
  - Snoop csak globális zónában
  - Globális zóna mindent, helyi zóna csak csak saját IP forgalmát látja
- Önálló címtér és címtár
  - Lokálisan megadott DNS/LDAP/NIS v. file
- Önálló IPC tér
  - Zóna között hálózati (sw loopback), v. megosztott filerendszerrel
- Kettős syslog
- Kettős pkg, patch adminisztráció

Zónák – virtualizáció

- Virtualizált /proc, /etc/mnttab
  - ps, mount csak a zóna adatait mutatja
- Önálló fájlrendszer
- Privát filerendszer pontok (/, /var, /etc)
- Bizonyos fájlrendszerek read-only módon beszármazhatnak egy zónába (pl. /usr)
- Korlátozott mount/umount lehetőségek
  - NFS mount zónánként
- Korlátozott privilégiumok (globális, nem globális)
  - kill, pkill
  - chmod, chown, setuid
  - link, unlink nem engedélyezett

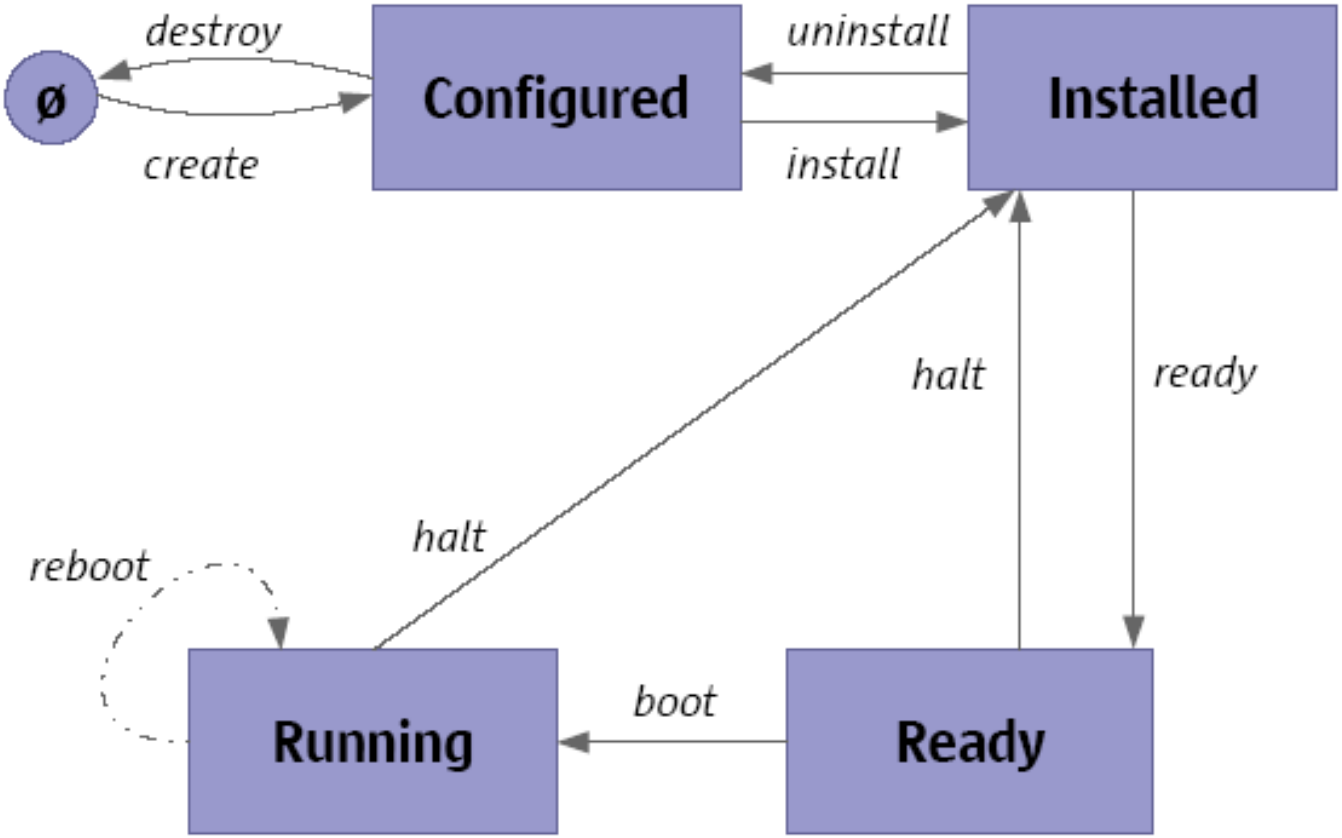
Zónák – file rendszer virtualizáció



Zóna – állapotok

- Konfigurált (Configured)
  - Teljes és elmentett konfiguráció (verifikált)
  - /etc/zones/{SUNWdefault.xml, XYZ.xml}
  - /etc/zones/index
- Telepített (Installed)
  - A zóna gyökerében telepített operációs rendszer csomagok
  - Nem teljes: sysidcfg
- Kész (Ready)
  - A hálózat, a fájl rendszer és a virtualizált eszközök rendelkezésre állnak (/dev)
- Működő (Running)
  - Felhasználói processzek futnak a zónában
- Incomplete, Shutting Down, Down
  - Átmeneti állapotok

Zóna – állapotok



Zóna alap parancsok

- **zonecfg:**
  - Zóna konfiguráció létrehozása, módosítása.
- **zoneadm**
  - Zóna adminisztrálása
- **zlogin**
  - Zónába lépés, nem hálózat!
  - zlogin –C zona => mintha hw konzol, „suninstall” első belépéskor
  - zlogin –l user zone parancssor
- **zonename**
  - Zónanév lekérdezés (zlogin után)

Zóna létrehozás

- Globális zónából hozhatók létre új zónák
- minták:
  - /etc/zones/{SUNWdefault.xml, SUNWblank.xml}
- zonecfg –z highzone –f config.txt
- Kulcs szavak:
  - zonepath, autoboot, pool
  - inherit-pkg-dir, fs
  - net, device, rctl
  - attr
  - verify, commit, export
  - delete

Zóna minta, DEFAULT

```
<?xml version="1.0"?>
<!--
    Copyright 2003 Sun Microsystems, Inc.  All rights reserved.
    Use is subject to license terms.
    ident "@(#)SUNWdefault.xml   1.1  03/12/09 SMI"
    DO NOT EDIT THIS FILE.  Use zonecfg(1M) instead.
-->
<!DOCTYPE zone PUBLIC "-//Sun Microsystems Inc//DTD Zones//EN"
    "file:///usr/share/lib/xml/dtd/zonecfg.dtd.1">
<zone name="default" zonepath="" autoboot="false">
  <inherited-pkg-dir directory="/lib"/>
  <inherited-pkg-dir directory="/platform"/>
  <inherited-pkg-dir directory="/sbin"/>
  <inherited-pkg-dir directory="/usr"/>
</zone>
```

Zóna minta, config.txt

```
create -F
set zonepath=/opt/test_fs/zones/highzone
set autoboot=false
add inherit-pkg-dir
set dir=/opt/XXX
end
add net
set address=129.159.190.204
set physical=hme0
end
add rctl
set name=zone.cpu-shares
add value
     (priv=privileged,limit=75,action=none)
end
```

```
add fs
set dir=/mnt
set special=/dev/dsk/c1t2d0s0
set raw=/dev/rdsk/c1t2d0s0
set type=ufs
end
add fs
set dir=/opt/local
set special=/usr/local
set type=lofs
end
add device
set match=/dev/dsk/c1t2*
end
verify
info
commit
```

Zóna minta, parancssor

```
global# zonecfg -z highzone
highzone: No such zone configured
Use 'create' to begin configuring a new zone.
zonecfg:highzone> create
zonecfg:highzone> set zonepath=/opt/test_fs/zones/highzone
zonecfg:highzone> add net
zonecfg:highzone:net> set physical=hme0
zonecfg:highzone:net> set address=192.9.200.55/24
zonecfg:highzone:net> end
zonecfg:highzone> verify
zonecfg:highzone> commit
zonecfg:highzone> ^D
global# zonecfg -z highzone info zonepath
zonepath: /export/home/highzone
global#
```

Zóna minta, parancssor

```
global# mkdir /usr/local
global# zonecfg -z highzone
zonecfg:highzone> add fs
zonecfg:highzone:fs> set dir=/usr/local
zonecfg:highzone:fs> set special=/data/local/highzone
zonecfg:highzone:fs> set type=lofs
zonecfg:highzone:fs> end
zonecfg:highzone> ^D
global# zonecfg -z highzone info fs
fs:
        dir: /usr/local
        special: /data/local/highzone
        type: lofs
        options: []
global#
```

Zóna létrehozás

- IP cím
  - IPv4 automatikusan lehet, IPv6 manuálisan
  - Netmask automatikusan a globális zónából
- Öröklés
  - SUNWdefault template alapján inherit-pkg-dir
  - /usr, /sbin, /platform, /lib
  - /opt is javasolt
- Device
  - Fizikai eszköz is örököltethető, körültekintéssel
  - Globális root definiálja, locális csak használ

Zóna adminisztráció

- zoneadm –z zona ...
  - install, uninstall
  - ready
  - boot, reboot, halt
  - list (-c, -cv)
- zlogin –C zona
  - alternativa:
    - ${zonapath}/root/etc/sysidcfg

Zóna installáció

```
global# zoneadm -z highzone install
Preparing to install zone <highzone>.
Creating list of files to copy from the global zone.
Copying <2319> files to the zone.
Initializing zone product registry.
Determining zone package initialization order.
Preparing to initialize <1361> packages on the zone.
Initialized <1361> packages on zone.
Successfully initialized zone <highzone>.
```

Zóna listázás, ellenőrzés

- Listázás

```
global% zoneadm list -cv
  ID NAME             STATUS         PATH
   0 global           running        /
   - drop             installed      /drop-dir
   8 fracture         running        /zones/fracture
   - highzone         installed      /export/home/highzone
```

- Ellenőrzés (Verify)

```
global# zoneadm -z drop verify
/drop-dir must not be group readable.
/drop-dir must not be group executable.
/drop-dir must not be world readable.
/drop-dir must not be world executable.
could not verify zonepath /drop-dir because of the above errors.
zoneadm: zone drop failed to verify
```

Interaktív első boot

- `sysidtool(1M)`

```
[NOTICE: zone booting up]

SunOS Release 5.10 Version s10_52 32-bit
Copyright 1983-2004 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
Hostname: twilight
The system is coming up.  Please wait.

Select a Language

0. English
1. French
2. Japanese
3. Simplified Chinese
4. Traditional Chinese

Please make a choice (0 - 4), or press h or ? for help:
```

Inicializálás más módon

- `sysidcfg(4)` file editálás, alternatíva

```
global# cat /export/home/highzone/root/etc/sysidcfg
system_locale=C
terminal=xterm
network_interface=primary {
        hostname=highhost
}
security_policy=NONE
name_service=NIS {
        domain_name=engineering
}
timezone=CET
root_password=MWQxNuo89EjI2
```

Zóna adminisztráció

- Minden pkg: globális => lokális zóna
- Minden patch: globális => lokális zóna
- Globális root gyakorlatilag bármit módosíthat, globális user csak olvashat
- Editálható csomagok: pkgmap, pkginfo
  - Nem editálhatóak az inherit-pkg-dir alatti csomagok (csak metadata másolt)
- Hagyományos sysadmin tevékenység

Zóna adminisztráció

- -z zonename
  - ifconfig, ps, prstat, pkill, pgrep
- -Z
  - ps, prstat, df
- netstat: zóna kapcsolatait mutatja
- Lokális zónában nem működik:
  - Prtdiag, prtconf, eeprom, snoop, sysdef
  - NFS szerver, zónaközi NFS mount
  - DHCP discover
  - Defaultrouter változtatás
  - Kernel modul betöltés, kitörlés
  - Hálózati interfész csatolás

Izoláció, példa

```
# zonename
highzone
# ls /proc
18332  18359  18389  18427  18437  18444  18459  18491  18546  19777
18335  18386  18398  18433  18443  18456  18464  18545  18549  20243

# zonename
smallzone
# ls  /proc
19959  19986  20018  20056  20066  20073  20088  20120  20177  20180
19962  20015  20027  20062  20068  20085  20093  20176  20178  20240
```

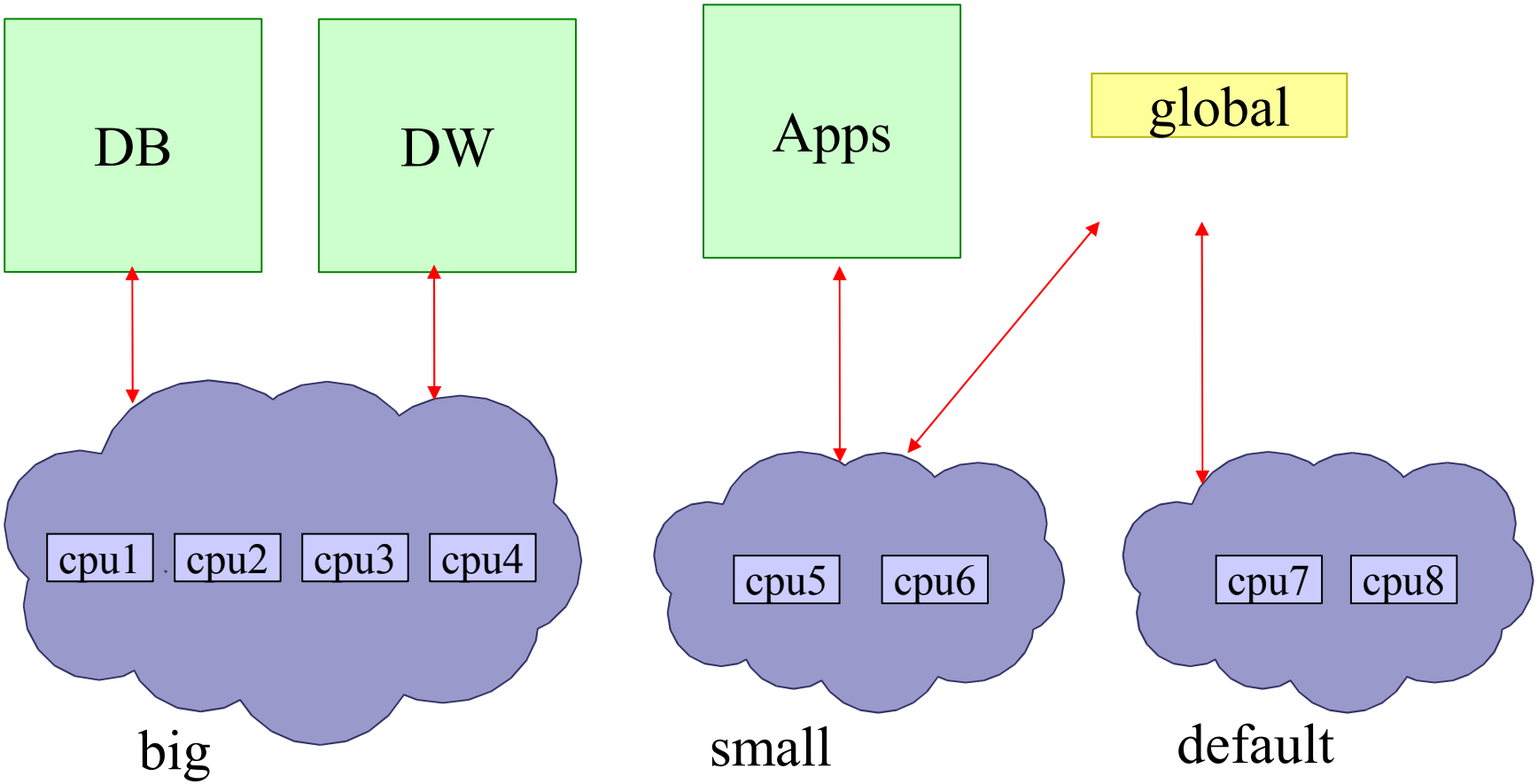
Zónák – erőforrás menedzsment

- A Solaris 9 továbbfejlesztett erőforrás menedzsmentje
- Erőforrás limitek definiálhatók a zónákhoz
  - Globális dinamikusan változtatható
- Egy erőforrás pool-ra több zóna is multiplexelhető
  - Automatikusan, zóna konfigurációban
  - Lokális zóna virtuális információt lát a pool-ról
  - Lokális zóna csak a hozzárendelt pool-t adminisztrálhatja
- A projekt adatbázis virtualizált, zónánként
- Minden zóna önálló rcapd-t futtat
  - Globális zóna beállítás csak a globálisra érvényes
- Két szintű FSS, két szintű model
  - Global => share2zone, local => share2project

Erőforrás menedzsment

- Új elemek
  - IPC erőforrások:
    - project.max-shm-ids, ..max-msg-ids, ..max-sem-ids
    - project.max-shm-memory, process.max-sem-nsems
    - process.max-sem-ops, process.max-msg-qbytes
  - Port, device:
    - project.max-device-locked-memory
    - project.max-port-ids, process.max-port-events
  - Titkosítás (cryptographic resource control):
    - project.max-crypto-memory
  - Egyéb:
    - project.max-lwps, project.max-tasks
    - project.max-contracts

Zónák, Pool-ok

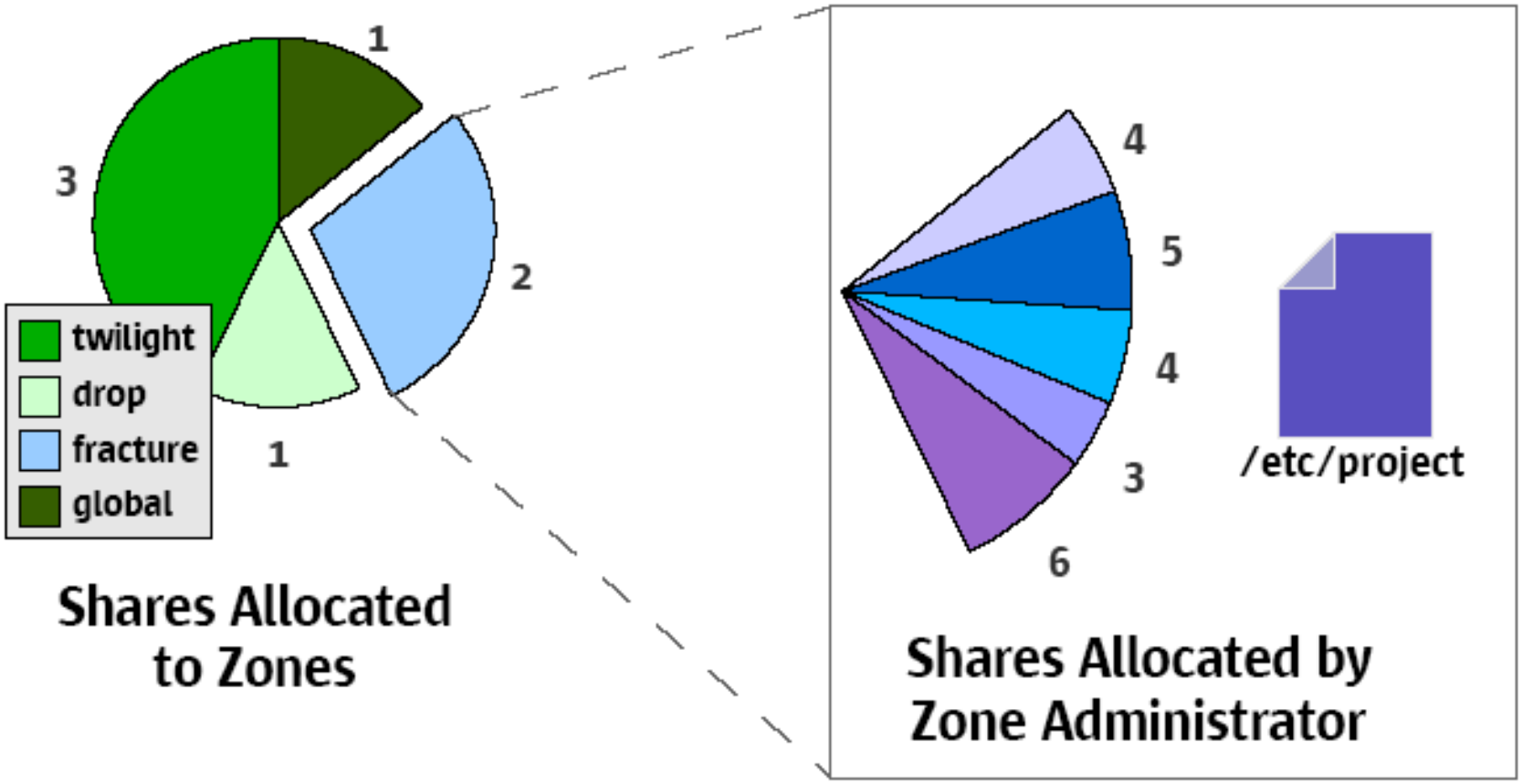


Resource Poolok

```
global% poolstat
                                   pset
   id pool                     size used load
    4 tide                        2 0.00 0.09
    0 pool_default                2 0.00 0.09
    3 whirl                       4 0.00 0.00

highzone% poolstat
                                   pset
   id pool                     size used load
    4 tide                        2 0.00 0.09
```

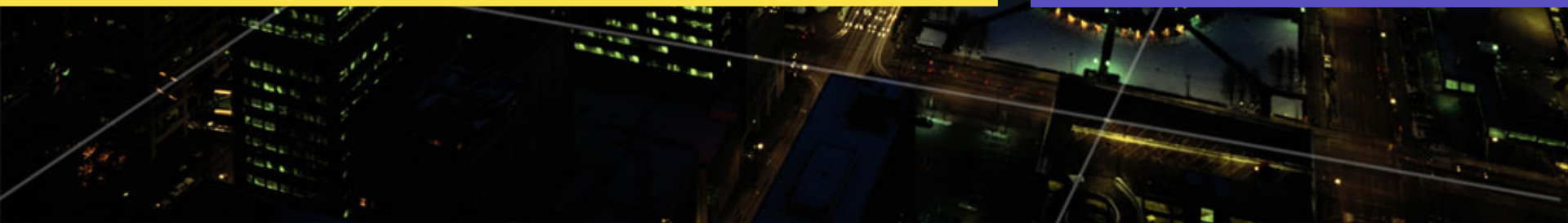
Zónák – erőforrás menedzsment



Kérdés!

Sorolja fel a
Zóna állapotait.....

ipfilter



ipfilter

- Nyílt forráskódú csomagszűrő
- 1.0 1993-ban
- Darren Reed fejleszti
- http://coombs.anu.edu.au/~avalon/ip-filter.html

ipfilter jellemzők

- Text fájl konfiguráció
- IPv6 (egyelőre kikapcsolt)
- Pici és csak pár csomag
- Nincs titkosítás
- Nincs lopakodó üzemmód
- NAT és port transzláció
- Nincs I18N
- Nincs CMG vagy HA

ipfilter jellemzők

- Stateful és stateless csomag vizsgálat
- Kernel alapú szűrés
- Protokol proxik (TCP, UDP , FTP, rcmds etc.)
- Naplózás

ipfilter szabályok

- Csomag blokkolás, átrengedés vagy naplózás
- Interfész, irány, IP cím, protokol, port, IP opció alapján
- Mindig az utolsó szabály érvényes
- IP cím csoportok: ippools
- Különálló fájlok a csomagszűrésre és a NAT-ra

Tipikus szabály fájl

```
block in all
pass in log quick on ce0 proto tcp from 129.159.190.0/8 to any port = 80
block out all
pass out log quick on ce0 proto tcp from any to 129.159.190.0/8 to any port = 80
```

ipfilter komponensek

- Kernel modulok
  - ipf – csomagszűrő
  - pfil – STREAMS driver
- Adminisztratív eszközök
  - ipf(1M) – csomagszűrő konfiguráció
  - ipnat(1M) – NAT konfiguráció
  - ipmon(1M) – Napló nézegető
  - ipfstat(1M) – Statisztika
  - ippool(1M) – ippools konfigurátor
- Szkriptek
  - /etc/init.d/pfil –pfil modul
  - /etc/init.d/ipfboot – csomagszűrő, NAT és ippools indító

Csomagok

- SUNWipfr - IP Filter Utilities, (Root)
- SUNWipfu - IP Filter Utilities

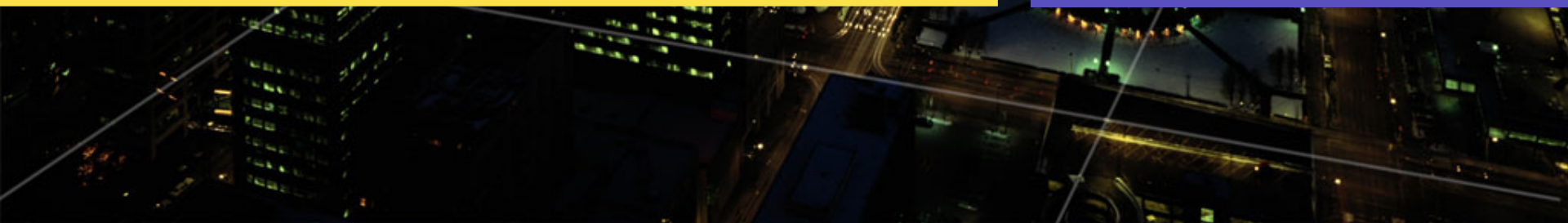
Teljesítmény

- ~11% sávszélesség veszteség (netperf2 mérés)
- Nincs boot idő hatás, ha nem konfigurált (default)
- < 1s boot idő növekedés, ha konfigurált

Jövő

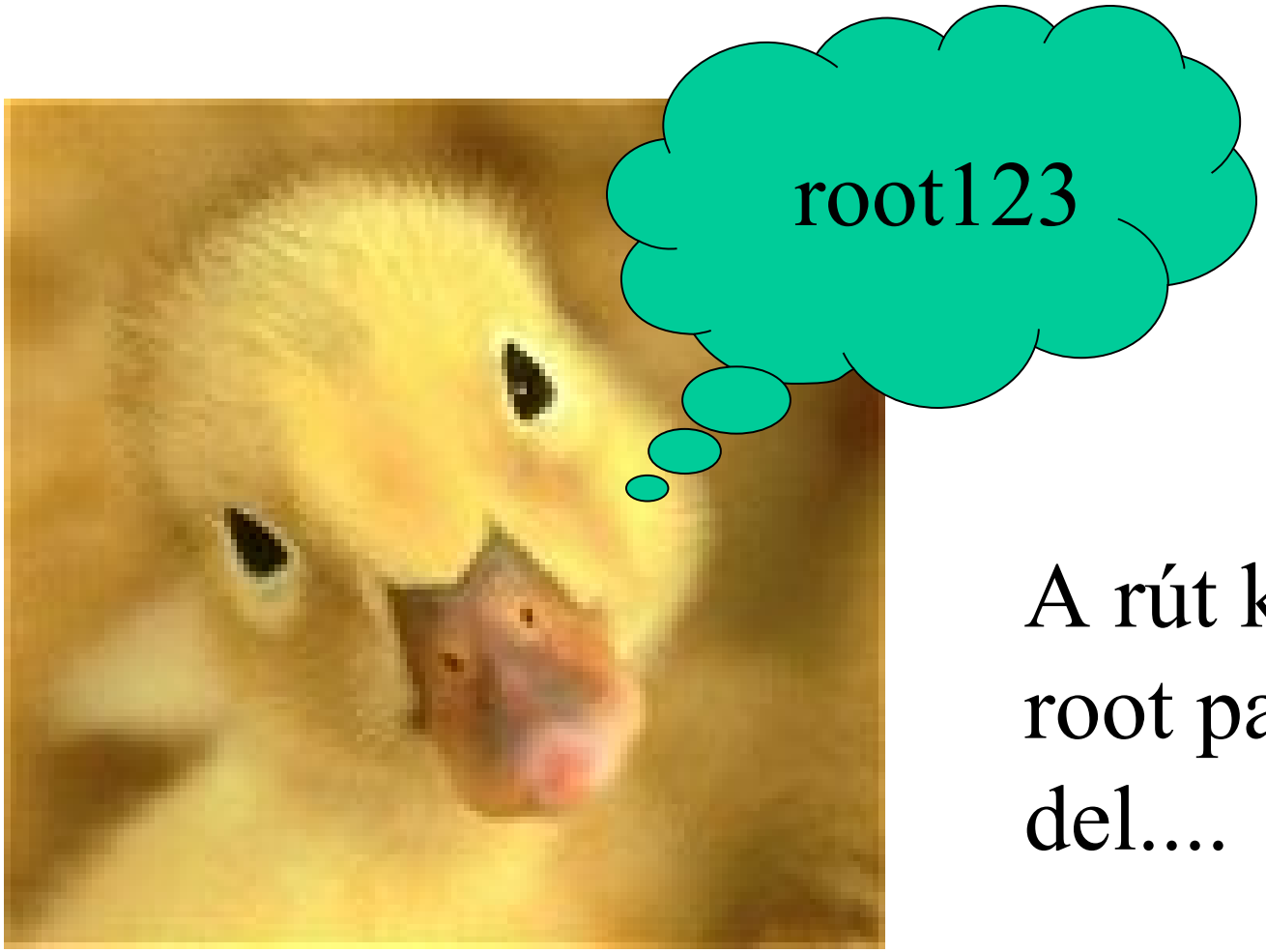
- IPv6
- IPMP
- Loopback csomagszűrés (zonák, Trusted Solaris)
- Jobb FireEngine integráció

Privilege rendszer



Kicsi, sárga, csúnya és nagyon veszélyes. Mi az?

Kicsi, sárga, csúnya és nagyon veszélyes. Mi az?



A rút kiskacsa a root password-del....

RBAC

- Solaris 8 óta az operációs rendszer része
- Átlagos felhasználó, de csak su-val közelíthető meg
- A szerepek 4 fájlban definiálódnak és parancs szinten konfigurálhatóak
- A Solaris 10-től privilégium kiegészítést is kaptak

Privileges - alapgondolat

- Least privilege
- A mindenható root jogainak szétszedése apró darabokra
  - Egyenként elvenni vagy odaadni
- Sok minden fut root-ként, aminek nem kellene
  - Pl. alacsony portokra bindoló web szerver

Megvalósítás

- A kernel nem UID==0-ra vizsgál, hanem az éppen szükséges privilege-re
- 47 egyedileg ki-bekapcsolható priv
- Bővíthető
- RBAC integráció
- Kompatibilitás: root minden priv-vel rendelkezik
- Privilege aware programozás

47 Privilege

"cpc_cpu"	Access to per-CPU perf counters
"dtrace_kernel"	DTrace kernel tracing
"dtrace_proc"	DTrace process-level tracing
"dtrace_user"	DTrace user-level tracing
"file_chown"	Change file's owner/group IDs
"file_chown_self„	Give away (chown) files
"file_dac_execute„	Override file's execute perms
"file_dac_read"	Override file's read perms
"file_dac_search„	Override dir's search perms
"file_dac_write"	Override (non-root) file's write perms
"file_link_any"	Create hard links to diff uid files
"file_owner"	Non-owner can do misc owner ops
"file_setdac"	Non-owner can set file perms (no seuid)
"file_setid"	Set uid/gid (non-root) to diff id
"ipc_dac_read"	Override read on IPC, Shared Mem perms
"ipc_dac_write"	Override write on IPC, Shared Mem perms
"ipc_owner"	Override set perms/owner on IPC
"net_icmpaccess„	Send/Receive ICMP packets
"net_privaddr"	Bind to privilege port (<1023+extras)
"net_rawaccess"	Raw access to IP
"proc_audit"	Generate audit records
"proc_chroot"	Change root (chroot)
"proc_clock_highres"	Allow use of hi-res timers
"proc_exec"	Allow use of execve()
"proc_fork"	Allow use of fork*() calls
"proc_info"	Examine /proc of other processes
"proc_lock_memory"	Lock pages in physical memory
"proc_owner"	See/modify other process states
"proc_priocntl"	Increase priority/sched class
"proc_session"	Signal/trace other session process
"proc_setid"	Set process UID
"proc_taskid"	Assign new task ID
"proc_zone"	Signal/trace processes in other zones
"sys_acct"	Manage accounting system (acct)
"sys_admin	System admin tasks (node/domain name)
"sys_audit"	Control audit system
"sys_config"	Manage swap
"sys_devices"	Override device restricts (exclusive)
"sys_ipc_config"	Increase IPC queue
"sys_linkdir"	Link/unlink directories
"sys_mount"	Filesystem admin (mount,quota)
"sys_net_config"	Config net interfaces,routes,stack
"sys_nfs"	Bind NFS ports and use syscalls
"sys_res_config"	Admin processor sets, res pools
"sys_resource"	Modify res limits (rlimit)
"sys_suser_compat"	3rd party modules use of suser
"sys_time"	Change system time

Interesting — Some interesting privileges

Basic — Non-root privileges

Privilege sets

- Effective (E) set
  - Ami éppen számít
  - P-ből bővíthető
- Permitted (P) set
  - Ezt csak csökkenteni lehet
  - Amit innen elveszek, az E-ből is elmegy
- Inheritable (I) set
  - Örökölhető exec-en keresztül
- Limit (L) set
  - I felső korlátja (exec)
  - Amit innen elveszek, az nem módosítja P-t és E-t

Exec szabályok

- Limit nem változik
- I’=I∩L azaz L limitálja I-t
- E’=I’; P’=I’ , azaz örökli az L-lel szűkített I-t

Alapbeállítások

- Basic set
  - Nem módosítható
  - file_link_any, proc_exec, proc_fork, proc_info, proc_session
- E: basic
- I: basic
- P: basic
- L: all

Parancsok

ppriv -l [-v]	az összes priv [részletes] listázása
ppriv [-v] <pid>	az adott process priv set-jeinek [részletes] listázása
ppriv -eD <parancs>	miért nem működik az adott parancs????
ppriv –s <priv> <pid>	process priv beállítása

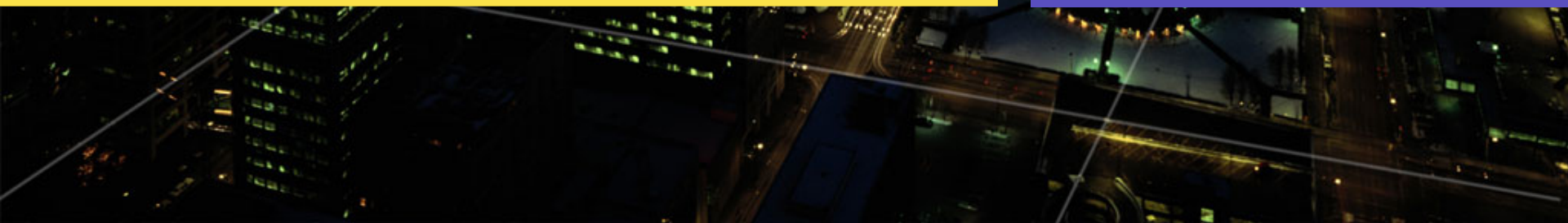
Kivételek – privilege escalation ellen

- Bizonyos esetekben nem az érintett priv-t várja el a rendszer, hanem az „All” priv készletet (hacsak nem vagyok root)
  - root által tulajdonolt file módosítása
  - root által tulajdonolt process módosítása
  - stb., pl:
    - file_dac_write: file írása jogoktól függetlenül

Adminisztráció

- RBAC összefonódás
  - /etc/security: exec_attr, prof_attr,
  - /etc: user_attr
- useradd, usermod
- roleadd, rolemod, /usr/sadm/bin/smrole
- /usr/sadm/bin/smprofile

BART: Basic Audit Reporting Tool



Sun microsystems

Mire való?

- Mi változott a rendszeren?
- Mi az eltérés két rendszer között?
- (Ugye nem törték fel a gépet...)

Mit tehettünk eddig?

- sidekick.sh
  - md5 signatures
- sfpDB
  - Keresés a Sun adatbázisában
  - MINDEN benne van, amit a Sun valaha is szállított
  - Verzióspecifikus adatok
  - Bármely Solaris verzióra használható

File integritás vizsgálat

- MD5: Letölthető toolkit
- 16 byte-os hash generálása
- File módosításakor változik a hash
- Ellenőrzés:
  - Saját adatbázis (install után, read-only médiára)
  - Solaris Fingerprint Database (http://sunsolve.sun.com)
- Kérdés: hogyan generáljuk?

MD5 signatures

MD5 (/usr/bin/at) = 3e34ddd03c87125902c824e0c58b5a68
MD5 (/usr/bin/awk) = d6451529b2172c6de71032d0de2ee3dc
MD5 (/usr/bin/banner) = 250e4b9590499246b14997990ca45bdf
MD5 (/usr/bin/bash) = c784b19d0034235fbf6de2accc6e86b6
MD5 (/usr/bin/cal) = 7f5f841d2ab9d3a0e0263bd66f403442
MD5 (/usr/bin/cancel) = b8dbf22d06f08f9938c270e73c371bda
MD5 (/usr/bin/cat) = 30f26ab47fd473a1ac0b63f8e62d609b
MD5 (/usr/bin/chgrp) = 8d0234c0770a255c52158df887c1d941
MD5 (/usr/bin/chkey) = 34c38ea44768f50ba3b7ec4b203b624b
MD5 (/usr/bin/chmod) = dbba839836ac12f1b19c2397493fbd84
MD5 (/usr/bin/chown) = 9f857e3b38d457c176575afc437dedcc
MD5 (/usr/bin/cp) = 420a9823e777812cad0c56b40d0a1524
MD5 (/usr/bin/su) = 8b98fb9c314bd5b378d9436b1617d014

Solaris Fingerprint Database

8b98fb9c314bd5b378d9436b1617d014 -
(/usr/bin/su) - 1 match(es)
canonical-path: /usr/bin/su
package: SUNWcsu
version: 11.8.0,REV=2000.01.08.18.12
architecture: sparc
source: Solaris 8/SPARC

BART: Basic Audit and Reporting Tool

- Baseline rögzítés
- Komparálás (vs. baseline vagy másik gép)
- File rendszeren belüli szűrés
- Wildcard alapú exclude/include
- Spec karakterek
- Távoli mount-okat nem követ

Első lépés: control manifest

- File rendszer állapotának rögzítése
  - `acl`
  - `contents` : md5
  - `dest` : symlink hova mutat
  - `devnode` : device node, csak char és block device
  - `dirmtime, lnmtime, mtime` : modification time
  - `uid, gid` : ownership
  - `mode` : file permissions
  - `size`
  - `type` : típus, pl. file, block device, socket, pipe stb.
  - all : a fentiek mindegyike
- Eltérő filerendszert nem követ, pl NFS mount

Manifest formátum

- ! Version 1.0
- ! Thursday, December 04, 2003 (16:17:39)
- # Format:
- #fname D size mode acl dirmtime uid gid
- #fname P size mode acl mtime uid gid
- #fname S size mode acl mtime uid gid
- #fname F size mode acl mtime uid gid contents
- #fname L size mode acl lnmtime uid gid dest
- #fname B size mode acl mtime uid gid devnode
- #fname C size mode acl mtime uid gid devnode
- / D 1024 40755 user::rwx,group::r-x,mask:r-x,other:r-x 3fd9ea47 0 0
- /.java D 512 40755 user::rwx,group::r-x,mask:r-x,other:r-x 3f8dc04d 0 10
- /.java/.userPrefs D 512 40700 user::rwx,group::---,mask:---
- other:--- 3f8dc06b 010
- /.java/.userPrefs/.user.lock.root F 0 100600 user::rwgroup::---,
- mask:---,other:--- 3f8dc06b 0 10 -
- /.java/.userPrefs/.userRootModFile.root F 0 100600 user::rw-,

Második lépés: test manifest és riporting

- Két manifest összehasonlítása, eltérések riportolása
  - Ugyanaz a rendszer, időben eltérő manifest
  - Két rendszer
  - stb.

Második lépés: test manifest és riporting

/su:

gid control:3 test:1

/ypcat:

mtime control:3fd72511 test:3fd9eb23

/vfstab:

mode control:100644 test:100777

acl

control:user::rw-,group::r--,mask:r--,other:r--

test:user::rwx,

Rule file

- Mely könyvtárakban, mely file-okra mit kell nézni
- Global

```
CHECK all
IGNORE acl
```

- További blokkok

```
<path name> [filters]
CHECK/IGNORE <attribs>
```

Parancsok

- bart create <options>

  -R : induló directory, ha nem /

  -I <files>: file lista specifikálása

  -r <rulefile> : rule file megadása

  -n : md5 checking kikapcsolása (gyorsít)

- bart compare <options> <control m.> <test m.>

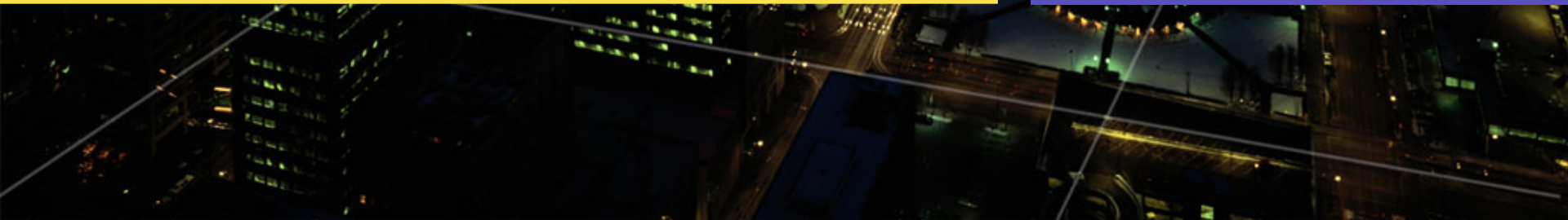
  -r <rulefile> : rule file megadása

  -p : programmatic output

Amire figyelni kell

- Jogosultságok
  - Csak ha van jogom olvasni a file-t, ill. keresni a directory-ban
  - Vagy ha root vagyok
  - Vagy ha van file_dac_read és file_dac_search
- De az eredmény biztonsági szempontból érzékeny!
- BART rendszer integritása????

Egyéb biztonsági kiegészítések



Kriptográfiai keretrendszer

- User és kernel egyaránt
- digest, mac, encrypt, decrypt funkciók
- Algoritmusok
  - User: DES, 3DES, AES, RC4, RSA, DSA, D-H, SHA-1, MD5
  - Kernel: DES, 3DES, AES, Blowfish, SHA-1, MD5
- Signózott elf (!)
- Hardver RNG támogatás
- Hardver accelerator támogatás

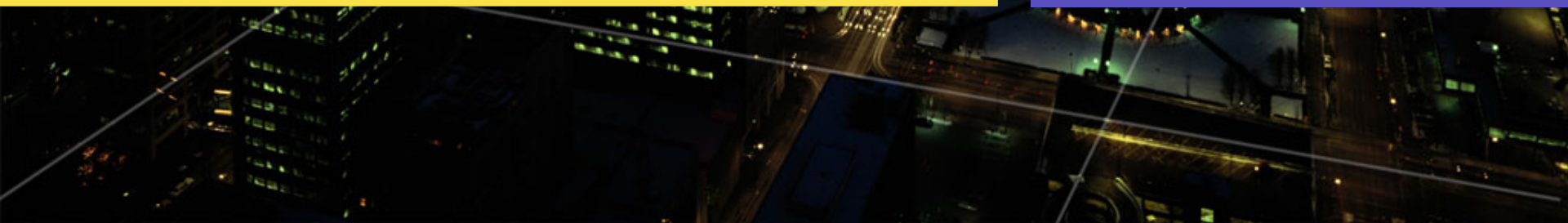
Password kiegészítések

- Password history
- Password ellenőrzés (cracklib) és integrált komplexitás ellenőrzés
- Tiltott password lista
- Cserélhető crypt(3c)
- Account zárolás

Egyéb fontosabb biztonsági apróságok

- Minimalizált operációs rendszer, 167MB, 81 csomag, 28 suid, 11 guid
- C2 audit syslog integráció
- Kerberos 1.3.2
- OpenSSH 3.6p2

SMF
Service Management Facility



Sun microsystems

Solaris 10 SMF

- A teljes OS infrastruktúra egységesítése
- /etc/rc*.d, /etc/inetd.conf és millió más szolgáltatás inditó/konfiguráló hely halála
- Automatikus, függőségeket is figyelembe vevő szolgáltatás indítás
- Új szolgáltatás azonosítás:

FMRI: Fault Management Resource Identifier

  - svc://gép/alrendszer/szolgáltatás:

Solaris 10 SMF

- XML repository
- Egyszerű és könnyen bővíthető
- Alapvetően négy parancs
  - svcs, svcadm, svccfg, inetadm
- Szolgáltatás ki és bekapcsolás
  - svcadm disable system/cron:default
  - svcadm enable network/ssh:default
- Profilok

Solaris 10 SMF

- Lényegesen jobb diagnosztika

  # svcs -x

  svc:/network/ntp:default (Network Time Protocol (NTP).)

  State: maintenance since Mon Oct 18 13:58:42 2004

  Reason: Start method exited with $SMF_EXIT_ERR_CONFIG.

  See: http://sun.com/msg/SMF-8000-KS

  See: ntpq(1M)

  See: ntpdate(1M)

  See: xntpd(1M)

  Impact: 0 services are not running.

Solaris 10 SMF

- Lényegesen jobb diagnosztika

```
# svcs -x -v
svc:/application/print/server:default (LP Print Service)
State: disabled since Mon Oct 18 16:17:27 2004
Reason: Disabled by an administrator.
See: http://sun.com/msg/SMF-8000-05
See: man -M /usr/share/man -s 1M lpsched
Impact: 1 service is not running:
                svc:/application/print/rfc1179:default
```

Service állapotok

- degraded
- disabled
- legacy_run
- maintenance
- offline
- online
- uninitialized

Solaris 10 SMF

- Lényegesen jobb diagnosztika

```
% svcs -d network/smtp:sendmail
STATE          STIME    FMRI
online         18:20:14 svc:/system/identity:domain
online         18:20:26 svc:/network/service:default
online         18:20:27 svc:/system/filesystem/local:default
online         18:20:27 svc:/milestone/name-services:default
online         18:20:27 svc:/system/system-log:default
online         18:20:30 svc:/system/filesystem/autofs:default
% svcs -D network/smtp:sendmail
STATE          STIME    FMRI
online         18:20:32 svc:/milestone/multi-user:default
```

Contract alrendszer

- A SMF egyik háttere
- Garantálja az “szerződött” alkalmazások “szerződésének” betartását
- A szerződés az elérhetőségre és a működés alapvető paramétereire vonatkozik

Contract alrendszer

- Tetszőleges felhasználói program szerződhető
  - ctrun(1)
  - Milyen eseményekre kell újraindítás: core, exit, hwerr
  - Mi legyen a gyermek processzekkel
  - Hány újraindítási probálkozás történjen
  - Pl.

```
ctrun –r 0 –t –f core,exit,hwerr httpd
```

Contract alrendszer

- ctstat(1) “szerződés” statisztika

```
ctstat -a
CTID    ZONEID TYPE    STATE    HOLDER      EVENTS QTIME    NTIME
1       0      process owned    0      0    -      -
4       0      process owned    1      0    -      -
41      0      process owned    7      0    -      -
51      0      process owned    7      0    -      -
71      0      process orphan   -      0    -      -
73      0      process orphan   -      0    -      -
75      0      process orphan   -      0    -      -
78      0      process orphan   -      0    -      -
81      0      process owned    260    0    -      -
83      0      process owned    260    0    -      -
87      0      process dead     -      0    -      -
99      0      process dead     -      0    -      -
```

Contract alrendszer

- Részletes statisztika

```
ctstat -av
CTID    ZONEID TYPE     STATE    HOLDER       EVENTS QTIME    NTIME
75      0      process  orphan   -            0      -        -
        cookie:              0
        informative event set: core signal
        critical event set:    hwerr empty
        fatal event set:       hwerr
        parameter set:         none
        member processes:      221 234 235 242 477 478 479 480 498 515 521
      567 569 571 631 632 639 640 641 642 643 645 653 654 722 735 736
      738 739 741 746 747 749 750 752 757 758 760 761 763 771 772 774
      775 777 782 783 785 786 788 793 794 796 797 799 804 805 807 808
      810 2526 18355 18357
        inherited contracts:   none
```

Köszönjük a figyelmet.

erik.fischer
tamas.zsemlye
zoltan.modly
@sun.com

Sun microsystems