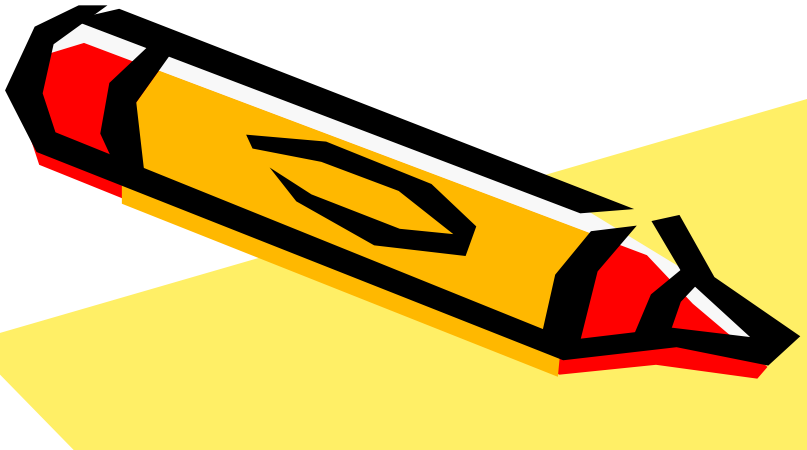

AB control system
Front-End computers

System aspects

Introduction

- Different kinds of front-end computers
- Boot mechanism
- Remote terminal access
- Remote reset
- NFS
- Transfer.ref files
- Management Makefile

front-end computers

different types

Front-ends are usually diskless computers but may have various forms:

- VME 32 crates (mainly WES crates)
- VME 64 Wiener crates
- VXI crates
- 2U PCs
- Compact-PCI crates (mainly Wiener crates using the same FAN units than WME ones)
- Normal PCs (but with many restrictions on supported network or video cards).

Different Operating systems

- LynxOS (http://www.LynuxWorks.com):
  - Mainly version 4.0 on VME or VXI PowerPC (ppc4) or PCs (x86),
  - but still a few more ancient: 3.1.0a for PowerPC (ppc), or 2.5 for MC680x0 (68k) or PCs (pc2).
- Linux
  - Diskless version derived from SLC3 (Scientific Linux version 3) used on consoles or servers (L86).
  - A few standalone PCs are just running SLC3 for WorldFIP validation (Linux).
  - A Linux version for CES RIO8064 exists but is not yet used.
- OS/9
  - A few antique 68k VME or G64 systems are still running OS/9 (TIS or AB/BI on the SPS).

Front-end boot servers

From 2006 we are using 2 new Linux servers as boot servers installed in CCR (Building 874 – Prevessin control room building):

- Cs-ccr-feop for technical network
- Cs-ccr-felab for General Purpose network

These are running SLC4, acting as bootp, tftp and nfs servers for the diskless front-ends.

As these are HP Proliant, their ILO can be accessed using http://csm-ccr-feop (or http://csm-ccr-felab). This permits remote power off/on and console access across the web.

Dsm is used for incremental backup/restore.

Note: if these are down, most other servers and consoles do not boot any more as these are also used to furnish their IP addresses.

Diskless boot mechanism

- *After BIOS self-test, diskless client issues a network BOOTP request, using Bootp protocol (or built-in PXE).*
- *This request is forwarded to bootp servers due to special programming of network routers.*
- *When server finds (in /etc/dhcpd.conf file) corresponding MAC address, it sends a reply with IP address of client, server, and default router, domain name server informations, boot file name, accelerator name and first mount info.*
- *Client downloads corresponding file using tftp protocol. (PXE downloads a preboot to download the same image).*

Reference file is /acc/sys/adm/dhcpd.conf propagated using /acc/sys/adm/update_dhcpd script

Boot (continued)

- Then BIOS jumps into the downloaded system image which embeds an initial RAM-disk,
- After basic initialization, initial script (/etc/rc for LynxOS or /linuxrc for Linux) repeats bootp request (using bootpc or dhclient) to find its name, IP address, default router, DNS info and from where to mount root filesystem.
- Then it does issue the first first NFS mount (/acc/sys).
- Then rc.S script is then executed (chained by /etc/rc from LynxOS systems or started by init for Linux systems).

Note: For PCs, PXE do not work fine on 10Mb/s ethernet and is not present on ancient PCs. An EtherBoot floppy is then used, but this change the required declaration in dhcpd.conf to be the same as for CPCI cards.

Remote terminal access

VME or CPCI cards have an RS232C serial line used for setup and diagnostics (usually 9600 bauds, 8bits no parity).

Some PCs BIOS can also be configured to use an RS232C serial line to print-out all diagnostics.

On all front-ends we try to get this line connected either to a terminal concentrator or to new Wiener VME FAN units that provides remote terminal access facility.

After boot, diskless systems are configured to use this serial line as console

Note: Some servers like HP Proliant also embed a system (e.g. ILO) which permits a remote access to the console.

Different remote-terminal servers

Terminal concentrators are from different types:

- *IT/CS managed DECserver 90/TL named …-..-…-TDE90-.. (installed in network start points on Meyrin site)*
- *IT/CS managed Annex named …-…-…-TXY8…-.. (found around SPS or TS places)*
- *AB/CO managed Radlynx named CCT-…-RAD… (found in LHC PCs installations).*
- *AB managed Wiener FAN units.*

Warning: each kind of terminal concentrator implies a different wiring in the RJ45/DB25 adaptor !!!

DSC terminal service

As remote terminal service cannot be shared by different users, a small task dsc_term runs on boot servers to:

- Connect permanently to the terminal server,
- Re-export a telnet service to multiple clients (using the dsc_terminal script)
- Keep a log file (limited to 1Mb): /nfs/cs-ccr-feop/local/syslog/<dsc_name>

Log files are kept for 1 week

Reference file is /acc/sys/adm/ds90tl

Remote Reset

In order to permit a reboot without going in front of equipment, different remote-reset systems are hidden behind /mcr/reset/rem_reset program:

- *Legacy PS using EM DIGIO:*
  - *pulse from a VMOD-DOR or an ICV96 VME card.*
  - *On AD: remote-reset box driven by DIGIO-A*
- *Legacy SL using SL-equip calls*
- *Wiener crates embedding remote-reset across SNMP (cfv-xxxx -> cfvm-xxxx or dleixxxx -> dleixxxx-fan)*
- *Twido PLCs used mainly for LHC.*

After reset, the remote reset programs chains a call to the dsc_terminal service in order to ease the reboot follow-up.

VME SAC module

System Access Controller is a CERN developed VME32 card acting as VME bus controller and providing VME SYSRESET to the CPU and other cards in the crate.

For PS, this is used using a blocking level (not TTL/) with modified versions to resist to spikes in the Linac.

For SPS this was used using a TTL signal. A "user program" is programmed in the BIOS to read ethernet MAC address from the SAC (this permits a CPU exchange without changing the dhcpd.conf file).

Note: SAC modules cannot be used any more with new VME64 Wiener crates.

Note also: Only a single VME card can act VME bus master (Slot 1 functionality).

CPU switches

On all CES RIO boards, at least one of the switches block has to be checked:

- Slot 1: this represent if the CPU is master of the VME bus and should be off if a SAC module is present
- VME SYSRESET signal should be programmed as an input if reset is taken from a SAC module of from the Wiener FAN unit
- Front-Panel reset: should be programmed as an output if unused.

Network File System

AFS is not available on LynxOS and not widely used. Front-end systems are therefore using only the NFS protocol to share files across the network.

NFS initially designed by SUN for diskless systems, currently have 3 versions:

- *Version 2 based only on UDP*
- *Version 3 introducing a TCP transport*
- *Version 4 introducing ACL capabilities.*

LynxOS only implements NFS version 2 and has no auto-mounter.

As HP-UX version 10.20 (still used by samoa) also only implements version 2, most of our usage of NFS is still based only on this version.

Without NFS caching, block size is limited to 512bytes/UDP datagram while caching permits to use up to 8192bytes/UDP datagram.

Note: A side effect of NFS caching is Directory caching (usually setup to 60s). Files removed by a client may be still reached by other clients that are still accessing them...

Automount

On normal Linux computers, we widely use autofs.

This permits to have filesystems mounted only when required, but a side effect, ls reports only mounted directories in automount points.

```
cd /acc/dsc
ls
ls /acc/dsc/oper
ls
```

Reference files are in /etc:

- *auto.master*
- *auto.nfs + auto.nfs.sub to mount /nfs/xxx/yyy*
- *auto.ps to mount /ps/xxx*
- *auto.acc + auto.dsc to mount /acc/xxxx and /acc/dsc/xxx*

NFS protection

NFS protection is weak, based only on IP address and UIDs.

On our systems, this is driven by a file /user/sysadm/hosts_equiv/hosts.equiv located on samoa, used by different scripts that rebuild /etc/netgroup with an exhaustive list of AB trusted systems and re-exports local filesystems:

- /user/sysadm/tools/Protect_NFS on HPUX 10.20
- /user/sysman/tools/Protect_NFS on HPUS 11.10
- /afs/cern.ch/group/pz/sue/Protect_NFS on SLC3
- /ps/sys/slc4/adm/Protect_NFS on SLC4.

Ask to a system manager to add missing systems to the list and to re-run the right scripts…

Main file-systems

- /acc/sys holds the reference OS for diskless systems. This is duplicated from cs-ccr-feop to cs-ccr-felab.
  /root -> /acc/sys/<cpu>
- /acc/dsc is split between felab (lab, tst) and feop (oper, mcr, lin, psb, cps, ade, sps, lhc, ln3, lei, ctf).
  /usr/local -> /acc/dsc/oper/<cpu>
  /dsc/local -> /acc/dsc/<acc>/<dsc_name>
  /dsc/data -> /acc/dsc/<acc>/data
- /ps/local/<cpu> holds include files and libraries used for development.
  /usr/local/lib -> /ps/local/<cpu>/lib
  /usr/local/include -> /ps/local/<cpu>/include.
- /ps/src/dsc/<acc>/<dsc_name> is used for DSC management

DSC management Makefile

In /ps/src/dsc/<mach>/<dsc_name>, we have a Makefile:

- *"make transfer.ref" regenerate the specific applications startup file*
- *"make new_dtab" extracts all GM and FESA data, rebuilds applications and install in /acc/dsc/<mach>/<dsc_name>, i.e. in /dsc/local as seen from the DSC.*
- *"make clean" remove all binaries*
- *"make clobber" remove almost everything but Makefile.*
- *"make" just recompile*
- *"make install" do not re-extract database info but (after a clean) permits to re-link using latest version of libraries.*

Transfer.ref

To launch specific applications, wreboot is used on front-ends, with a transfer.ref file automatically generated from the database in order to keep information used by ioconfig library

Wreboot permits to restart applications detected as not running by clic. See "wreboot –h"

Front-end's generated transfer.ref is delivered to /dsc/local/etc (while it is still found - manually edited - in /user/pca/<host_name> for servers and consoles managed by AB/CO/IN).

Transfer.ref format

We first find lines starting with #-# information used by ioconfig library:

- Description of all VME modules
- List of CAMAC and G64 crates associated with the DSC,
- Definition of logical events used by application (e.g. for PPMAQI/PPMCVI)

We then find lines starting with #% commands executed before starting upfiles and wreboot: These are usually expressed as negative startup sequences in the database and cannot be surveyed by clic.

Other lines starting with # are just comments.

We then find normal upfiles/wreboot commands.

Upfiles/Wreboot command lines

Generic format is:

- Source directory
- Source file name
- Destination directory
- Destination file name (clic name)
- File owner
- Files group
- File chmod parameters (e.g. 555)
- Type of command: command, server, slink, dir, file or "-"
- Delay before next command or priority if negative value,
- % command %

Upfiles/Wreboot cont'd

- Upfiles execute a copy or re-create a symbolic link of the specified source file/dir to the specified destinations and then adjust owner/group/mode. If destination is equal to source, no action is done.
- Wreboot executes a change of current user/group to the specified ones, then change current working directory to the destination directory and then executes command between ending % % if transfer.ref line type is command or server.
- If the transfer.ref line type is server, destination file name is passed to clic for survey (in /etc/server.init)

Clic/Xcluc

- *Clic is used to survey different critical parameters like critical processes or local file-systems on our systems.*
- *Application server names are transmitted by wreboot to clic across /etc/server.init file.*
- *Clic is reporting to a Central logger using sysReporter*
- *Then the central logger info may be displayed using the xcluc or the legacy PS alarm applicatiion.*

Note: for automatic integration of a computer into the PS alarm, at least 1 equipment (DSCPM) must exist and be declared as belonging to this accelerator.

Front-End users

The passwd and group files are derived from HP samoa every night in /acc/sys/adm.

Normal user's home directory is on slnfs1 which is mounted only after reboot.

Special user's home directories or extra file-systems maybe mounted by transfer.ref specific commands.

As ssh is available only once reboot is finished, in case of problems, telnet is available for diagnostic login (as root user) – except on Linux.

GM Data-table save

At least in the past, there wa no crontab under LynxOS. Therefore the GM data-table save is triggered from the boot server's cron table at regular intervals, sending a directed UDP packet to the front-end.

For the data-table save to work, a DSCPM equipment is required and SAVE property set to 1.

Reference data-table save cron table is in /user/pca/cs-ccr-feop/dtsav

CNIC implications

CNIC has been introduced to block exchanges between General Purpose Network (GPN 137.138.xx.xx) and Technical Network (TN 172.18.xx.xx).

Only computers within a trusted list can communicate with computers.

Computers connected using Dynamically allocated IP addresses, e.g. Wi-Fi (128.142) will never have a direct access the technical network.

In case of communication problems, first verify rights using http://network

Unknown MAC addresses will be blocked on the TN. This implies to maintain correct MAC information in http://network for each connected device.

How to work from non-trusted systems

Most of our Linux installations are in the CNIC trusted list, especially abcopl1 is now on the GPN with the right rights.

For Windows XP systems, you may always use Windows Terminals Services and connect to CernTsxxx computers. Please ask Pierre Charrue for more information.

Note that Only computers in trusted lists may access our file servers…