AIX Commands

2009. 07.

UNIX TEAM
SSO/ITD/GTS
IBM KOREA

목 차

1. SMIT 6

1.1. 실행 환경 6
1.2. 사용 명령어 6

2. 시스템 기동과 정지 7

2.1. 시스템 Startup 7
2.2. 시스템 Shutdown 9

3. 시스템 사용자 및 그룹 관리 10

3.1. 사용자 관리 10
3.2. 그룹 관리 12

4. Logical Volume Manager 14

4.1. Volume Group 14
4.2. Physical Volume 16
4.3. Logical Volume 17

5. 파일 시스템 20

5.1. 파일 시스템 생성 20
5.2. 파일 시스템 관리 22

6. Paging Space 와 Dump Device 25

6.1. Paging Space 25
6.2. Dump Device 26

7. 백업과 복구 29

7.1. 백업 29
7.2. 복구 33
7.3. 상황 별 디스크 교체 절차 35

8. 소프트웨어 관리 39

8.1. 설치 39
8.2. 유지 관리 40

9. 네트워크 관리 43

9.1. hostname 설정 43
9.2. ip 설정 43
9.3. 기본적인 tcp/ip 설정 44
9.4. EtherChannel 구성 44
9.5. 라우팅 테이블 추가 삭제 45
10. 시스템 모니터링 48
10.1. nmon 48
10.2. sar 50
10.3. vmstat 51
10.4. ps 51
10.5. iostat 51
10.6. netstat 51
10.7. svmon 51
10.8. 기타 모니터링 58
11. 시스템 장애 처리 59
11.1. HACMP 장애 59
11.2. GPFS 장애 61
11.3. Filesystem umount 장애 63
11.4. 기타 장애 처리 66
12. Problem Determination 71
12.1. 기본적인 PD 절차 71
12.2. 문제 해결을 위한 기본적인 데이터와 기본적인 H/W PD 절차 71
12.3. Dump Handling Process 72
12.4. SNAP 76
12.5. TRACE 77
12.6. alog 77
13. AIX Parameter 78
13.1. CPU & Memory Parameter 78
13.2. Network Parameter 79
14. HMC (Hardware Management Console) 81
14.1. HMC 개요 81
14.2. LPAR & DLPAR 를 통한 자원 재분배 81
14.3. HMC 백업과 복구 92

15. HACMP 94

15.1. HACMP 구성 방식 94
15.2. HACMP Config Guidelines 96
15.3. HACMP 구성 절차 102

16. GPFS 117

16.1. GPFS Software Install & 환경설정 117
16.2. GPFS Cluster Creation 118

17. OS Migration 125

17.1. 시스템 OS backup 수행 125
17.2. migration 사전 절차 125
17.3. migration 작업 절차 125
17.4. migration 사후 절차 127

18. NTP (Network Time Protocol) Server 128

18.1. SETTING UP AN NTP SERVER 128
18.2. SETTING UP AN NTP CLIENT 128

19. Openssh 130

19.1. openssl 설치 (from AIX Toolbox for Linux) 130
19.2. openssl 설치 (from AIX Expansion Pack) 130
19.3. Creating the key on each node 130
19.4. Using a shell script for automatic login 131

20. TCP_Wrapper 133

20.1. TCP_Wrapper 다운로드 133
20.2. TCP_Wrapper 구성 133

21. WAS 135

21.1. Stand-alone Application Server 시작과 종료 135
21.2. Network Deployment Cell 시작과 종료 135

22. Apache 137

22.1. Apache rpm 설치 방법 137
22.2. Apache 시작과 중지 137

23. DBMS 139

23.1. DB2 시작과 종료 139

23.2. 오라클 시작과 종료 140

24. Storage 141

24.1. 사전 참고 사항 141
24.2. IBM vpath, RDAC 명령어 141
24.3. EMC powerpath 명령어 142
24.4. Hitachi HDML 명령어 144

25. 소프트웨어 라이프 사이클 145

25.1. AIX 145
25.2. HACMP 146
25.3. JAVA 146

1. SMIT

AIX 에서는 시스템 관리의 편의성을 위해 “System Management Interface Tools (SMIT)”라는 대화형 툴을 제공한다. SMIT 의 구성은 메뉴와 다이얼로그로 이루어져 있으며 Menu-Driven 방식으로 원하는 관리영역을 선택하고 실행하게 된다.

1.1. 실행 환경

사용자 인터페이스는 ASCII 및 X-Window 를 지원하며, 일반적으로 ASCII 모드를 주로 이용

1.1.1. ASCII Mode

- Non-Graphical Interface, Curses 기반의 화면 구조

1.1.2. X-Windows Mode

- Graphical Interface, Motif 기반의 화면 구조

1.2. 사용 명령어.

1.2.1. smit

- Motif/X-Windows 기반의 SMIT 명령어로 터미널 유형을 참조하여 적절한(ASCII 또는 X-Windows) 모드로 명령을 수행
- smit –M: 강제적인 X-Windows 모드 수행

1.2.2. smitty

- ASCII 기반의 SMIT 명령으로 Motif 에 비해 처리가 빠름
- smit –C: 강제적인 ASCII 모드 수행

1.2.3. 단축 경로

- SMIT 명령 실행 시 특정단어를 입력하여 해당 메뉴로 직접 이동

단축경로	내 용
smit system	시스템 환경변수 설정 및 관리
smit jfs2	JFS2 파일시스템 구성 및 관리
smit lvm	LVM 구성 및 관리
smit installp	소프트웨어 설치 및 관리
smit tcpip	TCP/IP 구성 및 관리
smit user	사용자 생성 및 관리

.

2. 시스템 기동과 정지

2.1. 시스템 Startup

플러그인 된 모든 장치들의 연결 여부 및 전원 상태를 확인한 후 서버의 전원을 올리면, 전면부에 장착된 “LED”를 통해 부팅 과정이 표시된다. “OK”가 표시되어 있으면 Shutdown 상태이며 정상적으로 시스템이 기동된 상태라면 “LED”에 아무것도 표시되지 않아야 한다

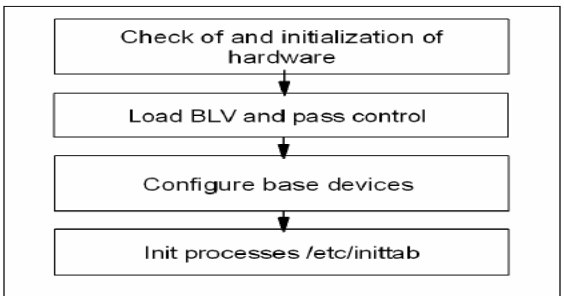
2.1.1. 부트 과정

- 시스템 부트 단계에서는 서버 자체 및 주변 장치에 대한 하드웨어 테스트, 운영체제 커널 로드 및 실행, 주변장치들에 대한 구성을 검증한다.

부트 과정	내 용
Normal Boot	일반적인 시스템 부트로 운영체제가 설치된 디스크를 통한 부팅
Service Boot	운영체제 설치 및 복구를 위해 부팅 가능한 미디어로 부팅
Network Boot	네트워크로 연결된 타 서버로부터 부트이미지를 받아서 부팅

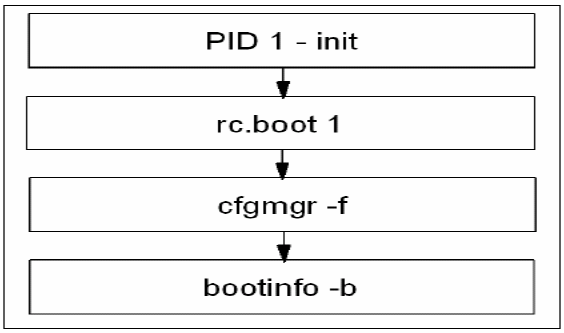
- 일반적인 부트 과정

일련의 하드웨어 점검 절차가 선행되며, 지금은 생산되지 않는 MCA 버스 구조의 서버에서는 “BIST(Built-in Self Test)-POST(Power-on Self Test)” 2 단계, 현재의 PCI 버스 구조의 서버에서는 “POST” 1 단계로 통합되어 부팅이 진행된다.



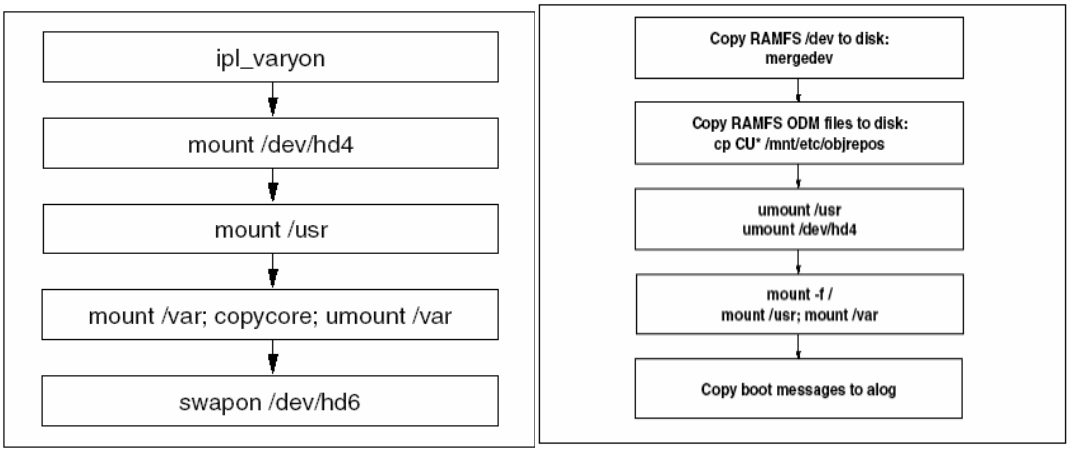
- 부트 단계 1

“rc.boot 1” 실행 및 cfgmgr 명령을 통해 rootvg 억세스를 위한 기본 장치들에 대한 점검을 수행하고 “bootinfo –b” 명령을 통해 최종 부트 장치를 확인한다.



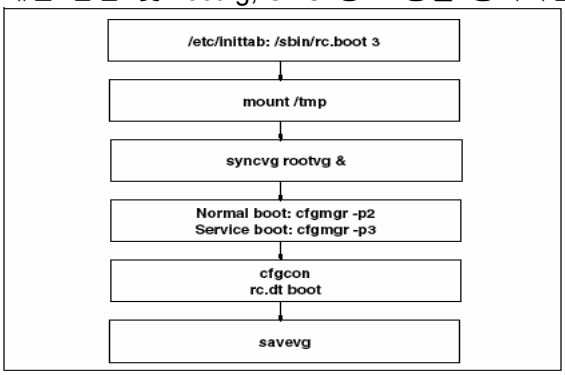
- 부트 단계 2

정상적으로 rootvg 이 varyon 되면, 운영체제 파일시스템들이 마운트된다. 최종적으로 부트 메시지가 alog 명령을 통해 콘솔로 출력되거나 에러 로그에 기록이 된다.



- 부트 단계 3

최종적인 부트 과정으로 "/etc/inittab"에 등록된 내용들을 순차적으로 실행하고, 주변장치들에 대한 점검 및 rootvg, ODS 정보 등을 동기화한다.



2.1.2. 일반적인 LED 코드 및 조치 방법

- LED 551, 555 또는 557

구 분	내 용
장애 내용	- damaged filesystem - A damaged JFS log device - A failing disk in the machine that is a member of the rootvg
조치 방법	① Diag를 통한 하드웨어 점검이나 문제를 해결하기 위해서는 부팅 가능한 미디어를 통해 Service Boot 과정을 실행한다. ② "maintenance menu"에서 "Access a Volume Group and start a shell before mounting file systems"를 선택한다. ③ 파일시스템 복구를 위해 fsck 명령을 실행한다. # fsck –y /dev/hd1 # fsck –y /dev/hd2

- LED 552, 554 및 556

구 분	내 용
장애 내용	- A damaged filesystem - A damaged JFS log device - A damaged IPL-device record or a damaged IPL-device magic number (the magic number indicates the device type) - A damaged copy of the Object Data Manager (ODM) database on the boot logical volume - A hard disk in the inactive state in the rootvg - A damaged superblock

조치 방법	① Diag를 통한 하드웨어 점검이나 문제를 해결하기 위해서는 부팅 가능한 미디어를 통해 Service Boot 과정을 실행한다. ② “maintenance menu”에서 “Access a Volume Group and start a shell before mounting file systems”를 선택한다. ③ “LED 551,555 또는 557”에 대한 조치방법으로 복구가 되지 않는 경우, 백업데이터를 이용한 복구를 수행한다. 만일, 운영체제 파일시스템 중 “/dev/hd4 (root filesystem)”이 복구가 되지 않는 경우는 시스템 재설치 외에는 방법이 없다. ④ “Maintenance Mode”로 부팅 시 “not an AIX filesystem”, “not a recognized file system type” 등의 에러메시지 출력 시, 파일시스템 superblock 복구를 위해 다음을 실행한다. # dd count=1 bs=4k skip=31 seek=1 if=/dev/hd4 of=/dev/hd4 ⑤ BLV의 ODM을 복구하기 위해 다음을 실행한다. # /usr/sbin/mount /dev/hd4 /mnt # /usr/sbin/mount /dev/hd2 /usr

- LED 518, 553

구 분	내 용
장애 내용	- LED 518: “/usr, /var” 파일시스템 마운트 장애 - LED 553: “/tmp, /” 파일시스템 용량 부족, inittab 파일 비정상
조치 방법	① rootvg 및 파일시스템 장애 조치 방법에 따라 복구한다. ② “/tmp, /” 파일시스템의 용량 및 “/etc/inittab” 파일의 문법을 점검한다.

2.2. 시스템 Shutdown

2.2.1. Shutdown 과정

- 시스템의 정상적인 종료 절차로 현재 수행중인 모든 프로세스의 종료 및 운영체제의 정상 종료 과정을 수행한다. 각 시스템 별로 운영체제 외의 서비스 프로세스들(Database, WAS 등)은 운영체제 종료 이전에 각 서비스들의 절차에 맞는 정상 종료 절차를 반드시 선행해야 한다. 만일, 서비스들의 정상 종료가 선행되지 않으면 운영체제는 이와 관련된 모든 프로세스에게 Kill 신호(TERM 후 실패 시 KILL)을 보내서 종료하게 되므로 데이터의 유실을 초래할 수 있다. 관련 명령어는 “shutdown, halt”를 사용한다.

2.2.2. Shutdown 명령

- 서버 리부팅 시 반드시 bootlist 를 확인 하고, boot image 생성 후 리부팅 하는 것이 바람직함.

bootlist –m normal –o (bootlist 확인)

bosboot –ad /dev/disk (boot image 생성)

구 분	내 용	
명령어	/usr/sbin/shutdown	
옵션	-F -r -l	fast shutdown shutdown 실행 후 시스템 재시작 shutdown log 작성, /etc/shutdown.log
관련 파일	/etc/rc.shutdown	

3. 시스템 사용자 및 그룹 관리

운영체제는 시스템 접근 권한을 부여하기 위한 기본적인 사항으로 Login ID 를 생성하여 사용자에게 부여하고 모든 Login ID 는 최소한 하나 이상의 그룹에 반드시 포함되어야 한다. 시스템 관리의 모든 권한을 가진 사용자를 superuser 라고 하며, “root”라는 Login ID 를 사용한다.
그룹은 사용자들의 집합이며, 동일 그룹 내 사용자들간의 자원 공유를 목적으로 한다. 그룹의 속성은 superuser 또는 superuser 가 관리역할을 부여한 사용자에 의해 변경이 가능하다

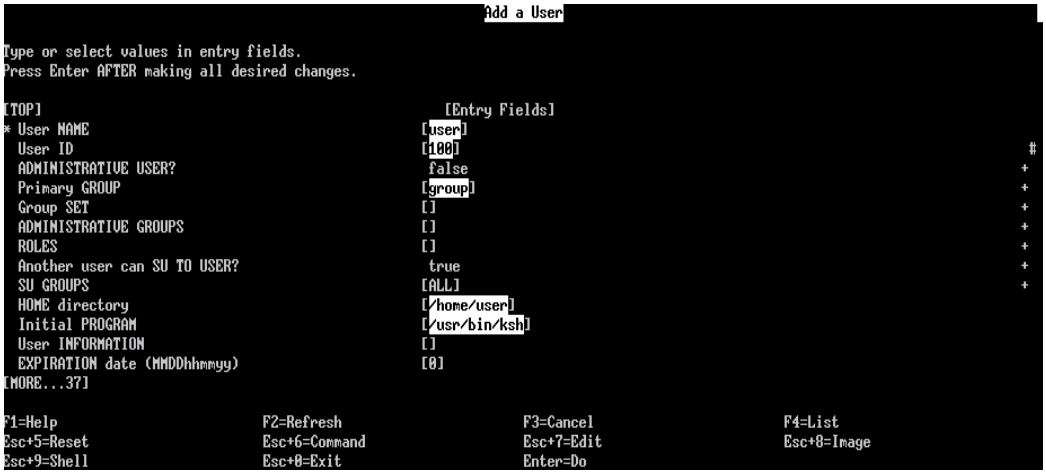
3.1. 사용자 관리

3.1.1. 사용자 추가

- smitty mkuser

사용자 생성을 위한 최소 입력 사항은 “사용자명”이다. 그 외의 값들은 미입력시 시스템이 정한 기본값들이 입력된다. 그러나, 실제로 운영체제가 사용자를 구분하는 것은 사용자 이름이 아닌 사용자 ID (UID)이며 고유한 값이어야 한다.

- 사용 명령어: mkuser home=/testuser testuser



- 사용자 환경 초기화

구 분	내 용	
/etc	시스템 전체 유저에게 적용되는 환경 파일이 있는 경로	
	profile	사용자 로그인 환경 초기화
	environment	프로세스/쉘 환경 초기화
$HOME	개별 사용자에게 적용되는 환경 파일이 있는 경로, 사용자 홈 디렉토리	
	.profile	사용자 로그인 환경 초기화 (자동 생성)
	.kshrc	프로세스/쉘 환경 초기화 (사용자 생성) “.profile”에서 ENV=$HOME/filename 변수를 통해 지정
적용 순서	/etc/profile → /etc/environment → $HOME/.profile → $HOME/.kshrc	

- 사용자 패스워드

사용자의 초기 패스워드는 관리자가 사용자 생성시 지정하며, 해당 사용자가 최초 로그인시 패스워드 변경 절차가 기본적으로 진행된다. 패스워드 변경 절차를 수행하지 않은 사용자는 패스워드 파일의 2 번째 필드가 “*”로 구분된다. 정상적인 경우, 해당 필드는 “!”로 표시된다.

- 사용 명령어: passwd testuser

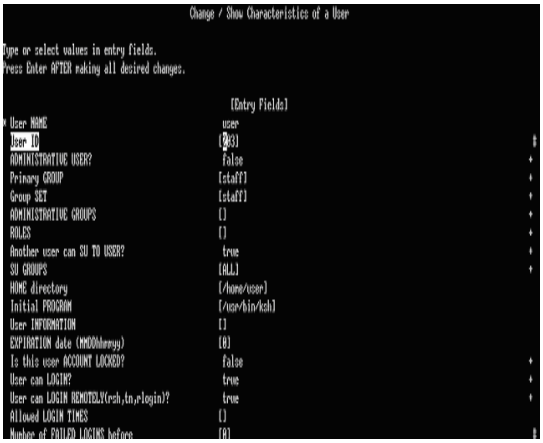
구 분	내용	
관련 명령어	/usr/bin/passwd	
관련 파일	/etc/passwd	사용자명, UID, 홈 디렉토리, 쉘 등에 대한 정보 기록
	/etc/security/passwd	사용자 실제 패스워드가 암호화되어서 기록

TIP> 스크립트를 이용해 사용자 패스워드 초기화 하기
echo "user-name:password" | chpasswd
pwdadm -f NOCHECK $username

3.1.2. 사용자 속성 변경

- smitty chuser

첫 화면에서 변경을 원하는 사용자명을 입력 후, 원하는 항목의 값을 변경한다.



- 사용자 제한 변경

사용자의 시스템 자원 사용 제한을 변경하는 것은 성능에 민감한 요소로 작용하므로 사용자의 업무 성격을 파악하고 자원의 효율적 할당을 위해 적절한 제한 값을 설정해야 한다. 설정의 변경은 위 화면의 붉은색 테두리 안의 값을 변경하면 된다. 적용을 위해서 사용자는 재 접속을 해야 하며, 변경 전 실행된 프로세스는 반드시 재시작 해야 한다.

내 용		
관련 파일	/etc/security/limits	사용자의 시스템 자원 사용 제한 값을 기록

- UID 변경 시 주의사항

유닉스 운영체제는 사용자 구분을 UID 라는 숫자를 통해 식별한다. UID 가 변경되는 경우, 실제 사용자가 생성했던 파일들의 소유권은 자동으로 변경되지 않는다. 따라서, 관리자가 직접 소유권을 변경된 UID 에 맞게 변경해 주어야 한다.

3.1.3. 사용자 삭제

- smitty rmuser

사용자명 입력 및 인증정보 삭제 여부 선택 후 진행한다

- 사용 명령어: rmuser -p testuser

- 파일의 삭제

사용자를 삭제해도 다음과 같이 실제 사용자가 생성했던 파일들은 삭제되지 않는다. 따라서, 필요 없는 파일들은 관리자가 직접 삭제해야 한다.

3.1.4. 사용자 관리 관련 Command

- lsuser: 이 명령은 AIX 에서 User 를 확인하는 명령

lsuser root: 이 명령은 root user 에 대한 모든 속성을 볼 수 있다.
root id=0 pgrp=system groups=system,bin,sys,security,cron,audit,lp home=/ shell=/usr/bin/ksh
#

#more /etc/passwd
"/etc/passwd" 35 lines, 1587 characters
root:!:0:0::/:/usr/bin/ksh
daemon:!:1:1::/etc:
...
생략 ~~~
...
snapp:*:210:12:snapp login user:/usr/sbin/snapp:/usr/sbin/snappd
picokid:!:241:205:HyunJin/Sys Admin:/home/picokid:/usr/bin/ksh
#

위 정보 중 맨 아래 line 을 확인해 보겠다.
picokid:!:241:205:HyunJin/Sys Admin:/home/picokid:/usr/bin/ksh

id : pidokid
UID : 241
GID : 205
사용자 정보: HyunJin/Sys ADmin
Home Directory: /home/picokid
Default Shell : /usr/bin/ksh (Korn Shell)

** ! 와 * 로 표시되는 부분이 있다. 이 부분은 암호 필드이며, 사용자 ID 에 암호가 있는 경우 암호 필드에는 ! (느낌표)가 있다. 사용자 ID 에 암호가 없는 경우 암호 필드에는 *(별표)가 있다. 암호화된 암호는 /etc/security/passwd 파일에 저장된다.

more /etc/security/passwd
root:
        password = pCPKl/RbIIRHM
        lastupdate = 1172112666
        flags =

daemon:
        password = *
…
생략 ~~~~
...
picokid:
        password = sfwBsnpMI5UZU
        lastupdate = 1171240485
        flags =

#
위에 설명에 있는 것처럼 passwd 가 암호화되어서 나온다.

3.2. 그룹 관리

3.2.1. 그룹 추가

- smitty mkgroup

그룹 생성을 위한 최소 입력 사항은 “그룹명”이다. UID 와 마찬가지로 운영체제가 그룹을 구분하는 것은 그룹 이름이 아닌 그룹 ID (GID)이며 고유한 값이어야 한다.

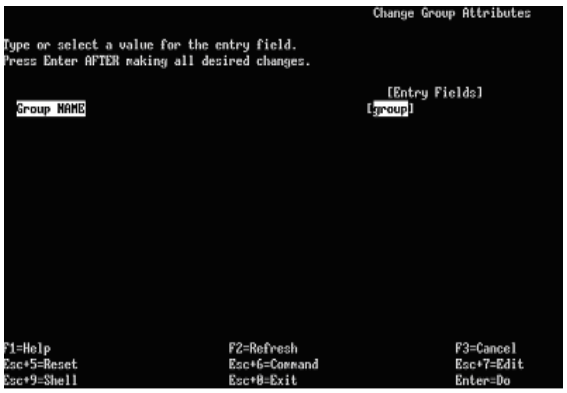
- 사용 명령어: mkgroup -'A' id='900' users='testuser' testgrp
- 그룹 파일

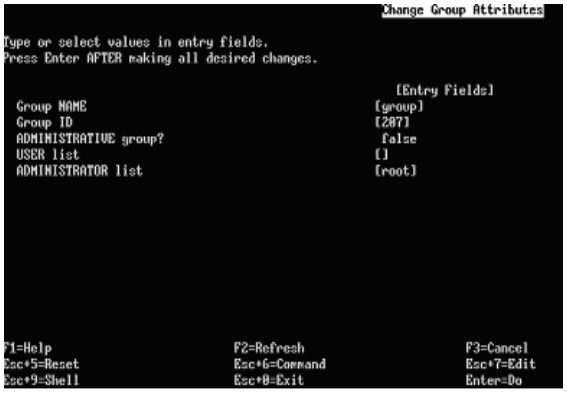
구 분	내 용	
관련 파일	/etc/group	그룹명, GID, 그룹원 목록 등이 기록
	/etc/security/group	관리자 그룹 여부, 그룹 관리자 정보를 기록

3.2.2. 그룹 속성 변경

- smitty chgroup

첫 화면에서 변경을 원하는 그룹명을 입력 후, 원하는 항목의 값을 변경한다.





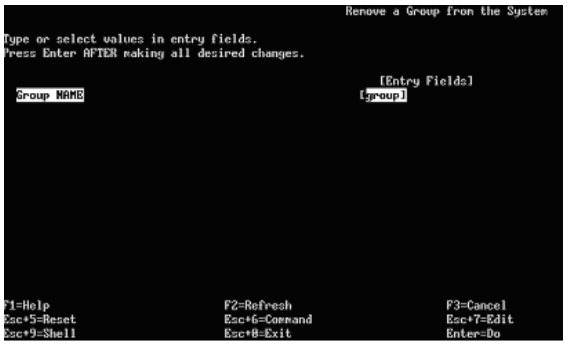
- GID 변경 시 주의사항

UID 변경 시와 마찬가지로 실제 사용자가 생성했던 파일들의 그룹에 대한 소유권은 자동으로 변경되지 않는다. 따라서, 관리자가 직접 소유권을 변경된 GID 에 맞게 조정해 주어야 한다.

3.2.3. 그룹 삭제

- smitty rmgroup
- 사용 명령어: rmgroup testgrp
- 그룹 삭제 시 주의사항

삭제하려는 그룹이 특정 사용자(들)의 “Primary Group”인 경우에는 오류가 발생하며, 관련 사용자를 모두 삭제하거나 해당 사용자들의 “Primary Group”을 다른 그룹으로 변경해야만 한다.



- 파일의 소유권 변경

그룹을 삭제해도 다음과 같이 실제 사용자가 생성했던 파일들의 그룹 소유권은 변경되지 않는다. 따라서, 관리자가 직접 소유권을 변경하거나 필요 없는 파일들은 직접 삭제해야 한다.

4. Logical Volume Manager

LVM 은 Logical Volume Device Driver (LVDD)와 LVM subroutine interface library 로 구성된다. LVDD 는 가상의 Device Driver 로 논리적 주소를 물리적 주소로 변환하여 I/O 요청을 해당 Device Driver 로 전달하는 등의 모든 I/O 를 관리한다. LVM subroutine interface library 는 논리적, 물리적 디스크의 작업을 수행하는 시스템 관리 명령어들이 사용하는 루틴들을 위한 라이브러리로 이루어져 있다.

구 분	내 용
가용성	운영체제 기준의 미러링 기능 제공으로 안정성을 향상시킬 수 있다. RAID 어댑터를 통한 미러링은 아니다.
	온라인 상태에서 미러링 해제 및 파일시스템의 증가가 자유롭다.
성능	스트라이핑을 통해 I/O를 분산할 수 있다

4.1. Volume Group

Physical Volume(PV)의 집합으로 최대 15 자까지 이름을 지정할 수 있으며, 고유의 Volume Group ID (VGID)를 가진다. 시스템에 데이터를 기록하기 위한 저장공간을 제공하기 위해서 반드시 PV 은 하나의 VG 에 포함되어야 한다. 특별히 운영체제가 포함된 VG 은 "rootvg"이라고 하며, 변경할 수 없다.

4.1.1. Volume Group Descriptor Area (VGDA)

- VG 정보를 가지고 있는 영역으로 VG 내의 PV 에 기록된다.
- 디스크의 동적인 할당, 제거에 이용된다.
- VG 의 PV 및 LV 정보를 포함한다.
- LVM 관련 명령에 의해 정보가 갱신된다.

4.1.2. Volume Group Status Area (VGSA)

- PP 및 PV 와 관련된 시스템 설정 정보를 포함한다.
- LV device driver 에 의해 관리된다.
- VGDA 와 VGSA 에는 time stamp 가 기록된다.

4.1.3. 관련 명령어

구 분	내 용	
lsvg	VG 관련 정보를 출력	
	-l vg_name	VG에 포함된 LV 정보 표시
	-p vg_name	VG에 포함된 PV 정보 표시
varyonvg	VG을 사용 가능한 상태로 전환	
varyoffvg	VG을 사용 불가능한 상태로 전환	
importvg	다른 시스템의 VG 정보를 현재 시스템에 추가	
	-y vg_name pv_name	지정한 VG명을 사용하여 추가
exportvg	현재 시스템에서 VG 정보를 제거 (다른 시스템에 import를 하기 위해 수행)	
extendvg	VG에 PV를 추가	
reducevg	VG에서 PV를 제거	

4.1.4. vg sync

- fast path: smitty syncvg_lv
- 사용 명령어: syncvg –l testlv

4.1.5. vg type

VG Type	Maximum PVs	Maximum LVs	Maximum User definable LVs	Maximum PPs per VG	Maximum PP size
Normal VG	32	256	255	32512 (1016*32)	1 GB
Big VG	128	512	511	130048 (1016*128)	1 GB
Scalable VG	1024	4096	4095	2097152	128 GB

- Normal VG 예시

```
rsks1:/etc/tunables>lsvg rootvg
VOLUME GROUP:       rootvg                  VG IDENTIFIER:  0001708700004c000000011f
VG STATE:           active                  PP SIZE:        32 megabyte(s)
VG PERMISSION:      read/write              TOTAL PPs:      542 (17344 megabytes)
MAX LVs:            256                     FREE PPs:       222 (7104 megabytes)
LVs:                13                      USED PPs:       320 (10240 megabytes)
ACTIVE PVs:         1                       AUTO ON:        no
MAX PPs per VG:     32512
MAX PPs per PV:     1016                    MAX PVs:        32
LTG size (Dynamic): 256 kilobyte(s)         AUTO SYNC:      no
HOT SPARE:          no                      BB POLICY:      relocatable
```

- Big VG 예시

```
rsks1:/etc/tunables>lsvg Big_VG
VOLUME GROUP:       Big_VG                  VG IDENTIFIER:  000afdad00004c000000011f
VG STATE:           active                  PP SIZE:        16 megabyte(s)
VG PERMISSION:      read/write              TOTAL PPs:      541 (8656 megabytes)
MAX LVs:            512                     FREE PPs:       541 (8656 megabytes)
LVs:                0                       USED PPs:       0 (0 megabytes)
ACTIVE PVs:         1                       AUTO ON:        yes
MAX PPs per VG:     130048
MAX PPs per PV:     1016                    MAX PVs:        128
LTG size (Dynamic): 256 kilobyte(s)         AUTO SYNC:      no
HOT SPARE:          no                      BB POLICY:      relocatable
```

- Scalable VG 예시

a rsks1:/etc/tunables>

rsks1:/etc/tunables>lsvg Scal_VG

VOLUME GROUP: Scal_VG VG IDENTIFIER: 000afdad00004c000000011f

VG STATE: active PP SIZE: 8 megabyte(s)

VG PERMISSION: read/write TOTAL PPs: 1076 (8608 megabytes)

MAX LVs: 256 FREE PPs: 1076 (8608 megabytes)

LVs: 0 USED PPs: 0 (0 megabytes)

ACTIVE PVs: 1 AUTO ON: yes

MAX PPs per VG: 32768 MAX PVs: 1024

LTG size (Dynamic): 256 kilobyte(s) AUTO SYNC: no

HOT SPARE: no BB POLICY: relocatable

4.2. Physical Volume

4.2.1. Physical Volume (PV)

- 하나의 물리적 디스크를 지칭하는 단위
- 시스템 상에서 고유의 Physical Volume ID (PVID)를 가진다.

4.2.2. Physical Partition (PP)

- PV 를 분할하는 물리적인 기본 단위
- PV 당 1,016 개의 개수 제약이 있으나, VG 생성 후"-t factor"에 의해 변경 가능하다
- LV 를 이루는 Logical Partition (LP)과 매핑된다.

4.2.3. 관련 명령어

구 분	내 용	
lspv	PV 관련 정보를 출력	
	-l pv_name	PV에 포함된 LV 정보 표시
	-p pv_name	PV에 포함된 LV 배치 정보 표시
migratepv	PV에 포함된 PP를 하나 또는 그 이상의 지정 PV로 이동	
	-l lv_name pv1 pv2	지정한 LV의 PP를 PV1에서 PV2로 이동
getlvcb	-TA lv_name	LVCB에 기록된 모든 정보를 출력

4.2.4. PVID 관련 작업

- PVID 생성

```
# chdev –l hdiskX –a pv=yes
```

- PVID 삭제

```
# chdev –l hdiskX –a pv=clear
```

- PVID 가 변경되었을 때 복구하는 방법

```
# vi restore_pvid.sh
  pvid=$1
  disk=$2
  set -A a `echo $pvid|\
  awk ' {
```

```
        for (f=1; f <= length($0); f=f+2) {
        print "ibase=16\nobase=8\n"toupper(substr($0,f,2))
        }
            }'|\
bc 2>/dev/null`
/usr/bin/echo "\0"${a[0]}"\0"${a[1]}"\0"${a[2]}"\0"${a[3]}"\0"\
${a[4]}"\0"${a[5]}"\0"${a[6]}"\0"${a[7]}"\0\0\0\0\0\0\0\0\c"|\
dd bs=1 seek=128 of=/dev/$disk
```

이전의 PVID 확인 (# lqueryvg -p hdiskX -At)

```
# restore_pvid.sh old_pvid hdiskX
# rmdev –dl hdiskX
# cfgmgr
# recreatevg –L / -Y NA –y vg00 hdiskX
```

=> vg00 에 속해있는 모든 디스크를 명시해야 한다.

=> "-L / -Y NA" option 을 사용하지 않으면 lv 와 filesystem 이름이 변경된다.

4.3. Logical Volume

4.3.1. Logical Volume (LV)

- PP 와 매핑된 LP 의 집합으로 PV 를 임의로 분할한 영역
- 물리적인 제약(PV)없이 공간 할당이 가능하다.
- LV 단위의 미러링 기능을 제공한다.

4.3.2. Logical Partition (LP)

- LV 를 이루는 논리적인 기본 단위

4.3.3. Logical Volume Control Block (LVCB)

- 모든 LV 의 시작 부분에 위치하며, LV 의 정보를 저장

4.3.4. 관련 명령어

구 분	내 용	
lslv	LV 관련 정보를 출력	
	-l lv_name	LV의 PV 배치 정보 표시
synclvodm	VG의 LV 관련 정보를 업데이트 또는 rebuid	
getlvcb	-TA lv_name	LVCB에 기록된 모든 정보를 출력

4.3.5. lv 추가

- fast path: smitty mklv
- 사용 명령어: mklv -y'testlv' khan_data01vg 10 dlmfdrv27

4.3.6. lv 삭제

- fast path: smitty rmlv
- 사용 명령어: rmlv 'testlv'

4.3.7. lv size 늘리기

- fast path: smitty extendlv
- 사용 명령어: extendlv testlv 10 dlmfdrv27

4.3.8. lv 속성 변경

- fast path: smitty chjfs2
- 사용 명령어: chlv -x'1024' testlv

4.3.9. lv size 줄이기

- lv 는 줄일 수 없으므로 lv 를 줄여야 할 경우에는 해당 lv 를 백업 받은 후 lv 를 삭제하고 다시 lv 를 만들어 준다.

4.3.10. lv mirror copy

- fast path: smitty mklvcopy
- 사용 명령어: mklvcopy testlv 2 hdisk1 or mklvcopy testlv 2 hdisk1 -k (sync 해주는 옵션)

4.3.11. jfs log 관련 작업

< Tips on JFS log device >

1. JFS log device name 은 Volume Group 별로 고유의 이름으로 변경하거나 새로 생성한다.
2. heavy IO filesystem 의 경우에는 별도의 JFS log device 를 생성하여 사용하도록 한다.
3. JFS2 inline log - 관리상 편리, 성능상 큰 차이는 없음.

<Creating a new JFS log device>

1. Create the logical volume to be used as the JFS log device with the jfslog or the jfs2log type, the desirable logname, volume group and number of physical partitions:

   mklv -t jfslog -y <loglvxx> vgname 1

loglvxx is the name of the JFS log device.

NOTE: Substitute jfs2log instead of jfslog if you are creating a jfslog for a JFS2 file system.

2. Format the new logical volume as a log device using following command (answer yes to destroy):

   logform /dev/<loglvxx>

loglvxx is the name of the JFS log device.

3. Unmount any filesystems wanting to use the new JFS log device:

   unmount /dev/<lvname>

4. Update /etc/filesystems and the logical volume control block (lvcb) of the filesystems needing to use the new JFS log device by using the following command:

   chfs -a log=/dev/<loglvxx> /mountpoint

   loglvxx is the name of the JFS log device.

5. Mount filesystems using new JFS log device by using the following:

mount /dev/<lvname>

There is no general rule for JFS2 log size. You need to increase one by one.

LABEL: JFS_LOG_WAIT IDENTIFIER: CF71B5B3 /

LABEL: JFS_LOG_WRAP IDENTIFIER: 061675CF

Determine which log device to increase. This can be determined by its Device Major/Minor Number in the error log, enter:

Device Major/Minor Number

000A 0003

The preceding numbers are hexadecimal numbers and must be converted to decimal values. In this example, hexadecimal 000A 0003 equals decimal numbers 10 and 3.

Determine which device corresponds with these Device Major/Minor Numbers, enter:

ls -al /dev | grep "10, 3"

<Increase the size of /dev/hd8 >

1. extendlv hd8 1
2. Boot the machine into Service Mode (also known as Maintenance Mode).
3. Format /dev/hd8 to use all of the space in the logical volume and answer y, to destroy the log, after entering the following command:

logform /dev/hd8

<Increasing the size of a user created JFS log device>

1. Extend the logical volume where the JFS log device resides on with 1 physical partition, enter:

extendlv <loglvxx> 1

loglvxx is the name of the JFS log device.

2. Determine which of the mounted filesystems use this log device, enter:

mount

The log device is indicated in the last column under options.

3. Unmount these filesystems that use the JFS log device specified in step 2, enter:

unmount /dev/<lvname>

4. Format the JFS log device to use all of the space in the logical volume and answer y, to destroy the log, after entering the following command:

logform /dev/<loglvxx>

loglvxx is the name of the JFS log device.

5. Mount all filesystems you unmounted in step 3, enter:

mount <lvname>

5. 파일 시스템

5.1. 파일 시스템 생성

5.1.1. 파일 시스템 생성

#smitty lvm (or #smitty mklv)

```
------------------------------------------------------------------------
                                Logical Volume Manager
Move cursor to desired item and press Enter.
  Volume Groups
  Logical Volumes
  Physical Volumes
  Paging Space
------------------------------------------------------------------------
                                   Logical Volumes
Move cursor to desired item and press Enter.
  List All Logical Volumes by Volume Group
  Add a Logical Volume
  Set Characteristic of a Logical Volume
  Show Characteristics of a Logical Volume
  Remove a Logical Volume
  Copy a Logical Volume
------------------------------------------------------------------------
                                Add a Logical Volume
Type or select a value for the entry field.
Press Enter AFTER making all desired changes.

                                                [Entry Fields]
* VOLUME GROUP name                            [rootvg]                  +
------------------------------------------------------------------------
                                Add a Logical Volume
Type or select values in entry fields.
Press Enter AFTER making all desired changes.
[TOP]                                           [Entry Fields]
  Logical volume NAME                             [test_lv]
* VOLUME GROUP name                               rootvg
* Number of LOGICAL PARTITIONS                   [32]
  Logical volume TYPE                            [jfs2]
------------------------------------------------------------------------
# lsvg -l rootvg|grep test
test_lv             jfs2       32     32     1     closed/syncd   N/A
#smitty jfs2
------------------------------------------------------------------------
                          Enhanced Journaled File Systems

Move cursor to desired item and press Enter.

  Add an Enhanced Journaled File System
  Add an Enhanced Journaled File System on a Previously Defined Logical Volume
  Change / Show Characteristics of an Enhanced Journaled File System
  …
  생략~~~
  …
  Unmount Snapshot for an Enhanced Journaled File System
  Change Snapshot for an Enhanced Journaled File System
  Rollback an Enhanced Journaled File System to a Snapshot
------------------------------------------------------------------------
```

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

```
                                                [Entry Fields]
* LOGICAL VOLUME name                             test_lv                                +
* MOUNT POINT                                       [/test]
(이하 생략)
```

```
# lsvg -l rootvg
rootvg:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT

loglv00             jfs2log    1       1       1    closed/syncd  N/A
test_lv             jfs2       32      32      1    closed/syncd  /test
```

- 사용 명령어: mklv -y'testlv' -t'jfs2' rootvg 10 hdisk0

5.1.2. lv mirror copy

- fast path: smitty mklvcopy
- 사용 명령어: mklvcopy testlv 2 hdisk1 or mklvcopy testlv 2 hdisk1 –k (sync 해주는 옵션)

5.1.3. vg sync

- fast path: smitty syncvg_lv
- 사용 명령어: syncvg –l testlv

5.1.4. filesystem add

- fast path: smitty crjfs2lvstd
- 사용 명령어: crfs -v jfs2 -d'testlv' -m'/testfs' -A

5.1.5. 파일 시스템 mount/umount

#mount
현재 mount 된 File System 내용을 보여준다. (참고 #more /etc/filesystems)

```
# lsvg -l logvg
logvg:
LV NAME         TYPE          LPs     PPs     PVs  LV STATE        MOUNT POINT
backuplv        jfs2          306     306     1    closed/syncd    /backup
loglv00         jfs2log       1       1       1    open/syncd      N/A
varlogs         jfs2          149     149     1    open/syncd      /var/logs
www_a_lv        jfs2          46      46      1    open/syncd      /www/a
#
```

현재 backuplv 는 closed 된 상태로 /backup File System 이 mount 되어 있지 않다.

#mount /backup -> F/S 을 mount 하는 방법

```
#df -g|grep /backup
/dev/backuplv      19.12      15.51     19%        22      1% /backup
# lsvg -l logvg|grep /backup
backuplv            jfs2        306     306     1    open/syncd      /backup
```

 # umount /backup -> F/S 을 umount 하는 방법

```
#df -g|grep backup
#lsvg -l logvg|grep /backup
backuplv            jfs2        306     306     1    closed/syncd    /backup
#
```

TIP> Filesystem 이 umount 되지 않는 경우 조치방안
- fuser 명령어를 이용하여 조치

5.2. 파일 시스템 관리

5.2.1. df

file System 의 사용현황을 파악할 수 있다.
df 는 기본적으로는 512 blocks size 로 보여진다.
옵션 -k (1024-blocks)

```
#df
Filesystem         512-blocks  Free    %Used     Iused %Iused  Mounted on
/dev/hd4              262144       158448    40%       2527       4%       /
/dev/hd2            10092544     1871424    82%      41612       4%       /usr
...
생략~~~
...
/dev/www_a_lv     6029312      3021976    50%       1821      1%         /www/a
/dev/perfmgr        131072       123056      7%         43       1%         /var/adm/perfmgr
#
```

옵션 -m (MB blocks), -g(GB blocks): 5.2 부터 지원

5.2.2. du

이 명령어는 directory(File System)의 Disk 사용률을 알아보기 위한 명령어.

-s 옵션은 합계이다.

```
# du /var/logs
648       /var/logs/adsm
713416   /var/logs/.tivoli/syslog
785344   /var/logs/.tivoli
65448    /var/logs/alerts050718
…
생략~~~
…
2245556 /var/logs/today
5470356 /var/logs
#du -sk /var/logs
5470496 /var/logs
#
```

5.2.3. du와 df 차이 발생시 해결 방법

du 와 df 명령의 결과 차이가 상당히 벌어지는 경우가 종종 생긴다면 어떤 process 가 사용하는 어떤 파일이 문제를 일으키는지 확인하실 필요가 있습니다.
그런 경우 다음을 참고하시기 바랍니다.

1. du 와 df 로 나오는 파일 시스템의 크기에 차이가 난다면 다음과 같은 명령으로 이상 파일을 확인할 수 있습니다.

fuser -dV /dev/hd3

위에서 '/dev/hd3' 는 문제되는 파일 시스템의 lv 이름으로 변경하시기 바랍니다

그러면 다음과 같은 결과가 나오게 됩니다.

/dev/hd3:

inode=284 size=4446 fd=63 34574

inode=284 size=4446 fd=63 38408

inode=318 size=6814 fd=63 49230

각 행의 마지막에 나오는 숫자가 PID 입니다. 즉, 해당 process 에서 사용하던 파일을 온전히 closing 을 하지 않아 발생되는 현상입니다.

2. fuser 결과에 나오는 PID 에 대해 process name 을 ps 명령의 결과에서 찾습니다.

ps -ef|grep 34574

그러면 어떤 process 인지 확인됩니다.

온전히 closing 이 되지 않는 파일의 경우 inode 만 확인할 수 있을 뿐, 파일명을 알 수는 없습니다. 따라서 아래 절차를 거쳐 어떤 파일이 문제인지 모니터링해야 합니다.

3. 이제는 du 결과와 df 결과가 같아지도록 fuser 결과에 나오는 process 를 kill 로 내리거나 재부팅을 합니다.

4. lsof 명령과 ncheck 명령으로 process 가 사용하는 파일명을 확인해둡니다.

"lsof -a -p <PID>" 를 수행하여 해당 process 의 파일정보를 확인합니다.

COMMAND	PID	USER	FD	TYPE	DEVICE	SIZE/OFF	NODE NAME
java	344210	root	cwd	VDIR	10,4	12288	4128 / (/dev/hd4)
java	344210	root	0r	VCHR	2,2	0t0	9337 /dev/null
java	344210	root	1w	VREG	10,9	3234	33435 /opt (/dev/hd10opt)
java	344210	root	2w	VREG	10,9	3234	33435 /opt (/dev/hd10opt)
java	344210	root	3r	VREG	10,5	236996	213208 /usr (/dev/hd2) <==
java	344210	root	4r	VREG	10,5	6842243	213088 /usr (/dev/hd2)

.. 생략…

위는 "lsof -a -p 344210"의 일부를 발췌한 것으로 '<=='표시된 행을 보면 해당 'VREG' type 이 일반 파일의 의미하고 /usr(/dev/hd2)의 inode 가 '213208' 인 파일입니다.

그 파일의 이름을 확인하려면 다음 명령을 수행합니다.

#ncheck -i 213208 /dev/hd2

이렇게 하여 process 가 사용하는 파일의 inode 와 파일명에 대한 mapping table 을 작성해둡니다

5.2.4. filesystem size 늘리기

- fast path: smitty chjfs2
- 사용 명령어: chfs –a size=+1G testfs

5.2.5. filesystem size 줄이기

- fast path: smitty chjfs2
- 사용 명령어: chfs –a size=-1G testfs

5.2.6. filesystem 삭제

- fast path: smitty rmjfs2
- 사용 명령어: rmfs –r /testfs

6. Paging Space 와 Dump Device

유닉스 운영체제에는 Real Memory (Physical Memory) 이외에도 메모리 관리자에 의해 디스크에 할당되어 실제 메모리의 확장된 영역처럼 사용되는 "Paging Space (Swap Space)"가 있다. 중단된 프로세스 및 관련 데이터 등을 저장하는 영역으로 이용된다. "Dump Device"는 시스템이 이상이 있을 경우 당시의 시스템 상태 정보(메모리 상에 존재하는 모든 정보)를 스냅샷 형태로 저장하는 공간이다. 별도로 지정하지 않는 경우 1 차 덤프 장치는 "Paging Space"가 지정된다.

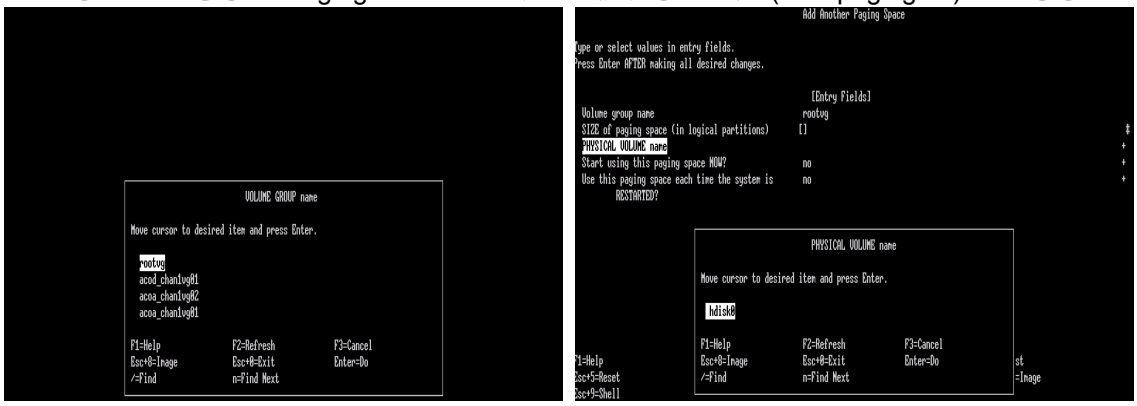
6.1. Paging Space

6.1.1. Paging Space 부족 관련 증상

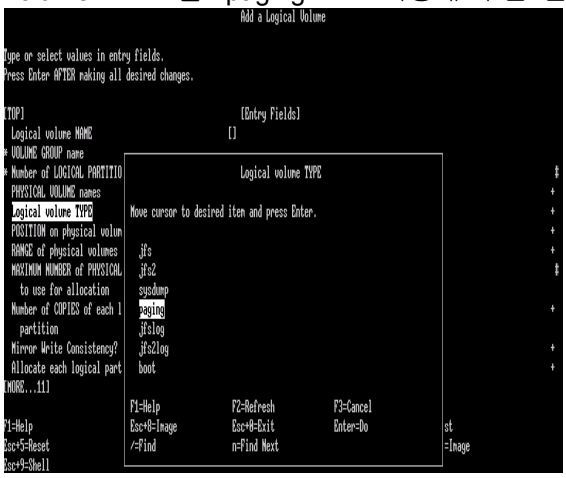
Paging Space 가 부족하면 응용프로그램이 시작이 안되거나 수행중인 프로세스가 강제 종료될 수 있으며, 고갈되면 시스템이 다운되거나 Hang-up 증상이 나타난다. 페이징 스페이스 부족의 징후가 발견되면 상황에 따라 영구적으로 추가하거나 튜닝을 통해 해결해야 한다.

6.1.2. smitty mkps (추가)

VG 선택 메뉴에서 원하는 VG 를 선택하고 크기 및 디스크, 시작 옵션 등을 설정한다. 이와 같은 방법으로 생성된 Paging Device 는 시스템에서 정한 이름(/dev/pagingXX)으로 생성된다.



사용자가 원하는 이름으로 Paging Device 를 생성하고자 하는 경우, LV 생성 메뉴에서 "Logical volume TYPE"을 "paging"으로 지정해 주면 된다.

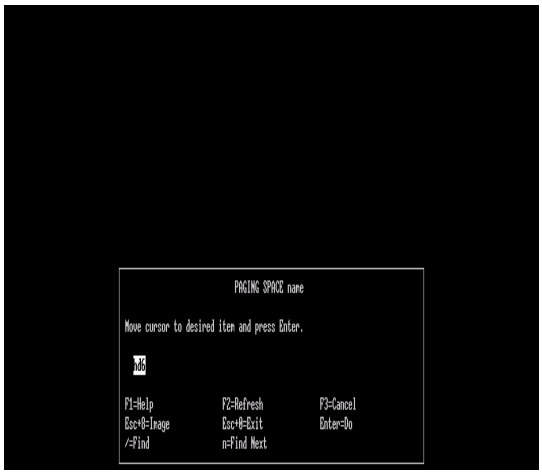


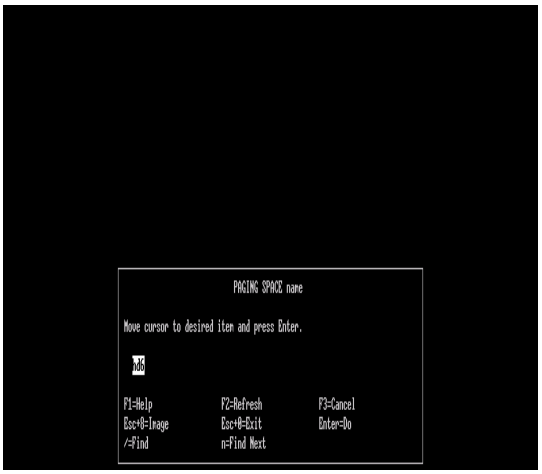
6.1.3. 관련 명령어

구 분	내 용
명령어	/usr/sbin/swapon, swapoff, lsps
Paging Device On	swapon –a | DeviceName …

Paging Device Off	swapoff DeviceName …	
Paging Device 상태	lsps –s	Paging Space 사용률 표시
	lsps -a	Paging Device 별 정보 표시
관련 파일	/etc/swapspaces	

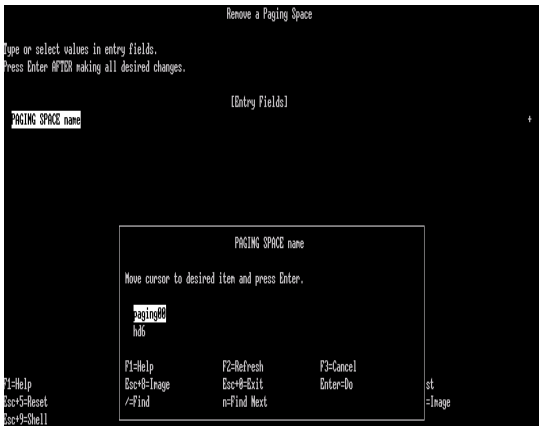
6.1.4. smitty chps (변경)





6.1.5. smitty rmps (삭제)

Paging Device 를 삭제하고자 하는 경우, 활성화 상태에서는 에러가 발생한다. swapoff 명령을 통해 비활성화 상태로 변경한 후 삭제가 가능하며, 최소한 하나 이상의 Paging Space 는 활성화 상태여야 한다.



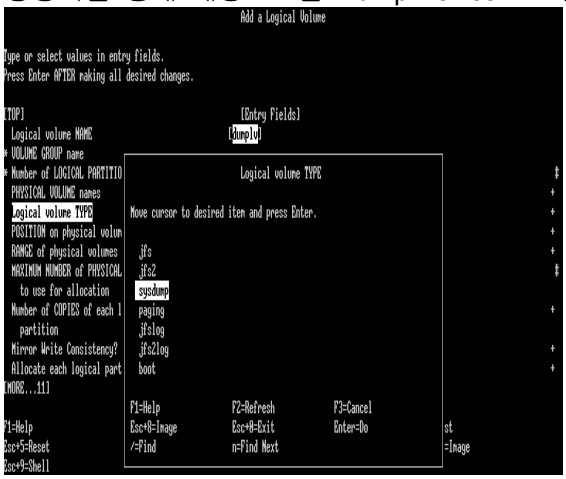
TIP> paging space 의 추가 및 삭제는 VG 의 pp size 를 반드시 확인하여 작업을 해야 하고, 작업시 여유 공간여부를 확인해야함.
만일 공간이 부족하여 분리해야 하는 경우 Paging space 사이즈가 동일해야 한다.
Hd6 1GB 이면 paging 1GB 으로 맞춰서 해야 한다

6.2. Dump Device

시스템 덤프는 예기치 않은 시스템의 정지 또는 특정 장애시에 발생한다. 덤프는 발생한 시점의 시스템 상태를 스냅샷 형태로 “Primary Dump Device”에 기록하며, 덤프 발생 이전 시점의 데이터들은 기록하지 않는다.
AIX 4.3.3 이하에서는 Paging Space 를 dump device 로 사용하였으나, AIX 5L 5.1 에서는 Real Memory 가 4GB 이상인 경우에는 install 시 dedicated dump device (/dev/lg_dumplv)를 자동으로 생성한다. 만일 Dump Device 가 모자랄 경우에는 Paging Space (/dev/hd6)를 사용하며 Real Memory 가 4GB 이하인 경우에는 default dump device 로서 Paging Space 를 사용한다.

6.2.1. smitty mklv (추가)

LV 생성시 “Logical volume TYPE”을 “sysdump”로 지정하여 LV 를 생성한 후 sysdumpdev 명령어를 통해 해당 LV 를 Dump Device 로 지정한다



6.2.2. 관련 명령어

구 분	내 용
-P	시스템 재시작 이후에도 적용
-p	Primary Dump Device로 지정
-s	Secondary Dump Device로 지정
-l	현재 Dump Device 정보 표시
-L	최종 Dump 기록 표시
-e	현재 시스템의 예상 덤프 크기 표시
-K	reset키를 통한 강제 덤프가 가능하도록 설정 (TRUE)
명령어	/usr/bin/sysdumpdev

6.2.3. dump option 변경

AIX 를 사용하는 시스템에 Hang 이 발생하면 강제 Dump 를 통한 분석이 필요한데, 강제 Dump 자체가 가 Fail 되어 고객의 큰 불만을 야기하는 경우가 아직도 많이 발생하고 있다. 문제 발생 가능성을 사전에 차단하기 위하여 아래의 정보를 공유 하오니 담당하는 고객사 시스템 환경을 확인하고 조치한다.

1. 대상 시스템
   - AIX 5.3 TL05 이상
2. 문제 현상 및 발생 가능 문제
   - 강제 Dump 시 Dump failure
3. 권고사항: 아래의 dump 관련 hidden option 들이 제대로 설정이 되어 있는지 확인

- 현재 Dump Option 확인하는 방법

```
# odmget SWservAt | grep -Ep "dump_compress|parallel_dump|dump_hang_prevent"
SWservAt:
        attribute = "dump_compress"
        deflt = "ON"
        value = "ON"

SWservAt:
        attribute = "dump_compress_factor"
        deflt = "5"
        value = "5"
```

SWservAt:
attribute = "dump_hang_prevent"
deflt = "OFF"
value = "OFF"

SWservAt:
attribute = "parallel_dump"
deflt = "ON"
value = "ON"

- Dump Option 변경 방법

아래와 같이 세가지 option 들을 변경하도록 한다.

- sysdumpdev -c (turn off dump compression) <== compression 을 disable 하는 것이 보다 바람직하긴 하나, 설치된 memory 량이 커서 실제 compression 없이 dump 를 쏟을 경우 공간이 부족한 환경인 경우 이를 그냥 enable 하여 사용하도록 할 것
- sysdumpdev -x (turn off parallel dump) <== Disabling parallel dump would take into effect after the reboot
- sysdumpdev -H (turn on dump_hang_prevent)

* 주의: 변경 시 dump_hang_prevent 부터 절대 enable 시키면 안된다. 순서는 반드시 "sysdumpdev -x;sysdumdev -H" 이다. <== 순서가 바뀌면 error 가 발생할 수 있다.
* 참고로 조만간 release 될 AIX 5.3 TL09 와 AIX 6.1 최신에서는 dump_hang_prevent option 은 default 로 enable 시키도록 변경을 하게 되어 더 이상 지금처럼 configure 를 할 필요가 없어질 예정이다.

7. 백업과 복구

7.1. 백업

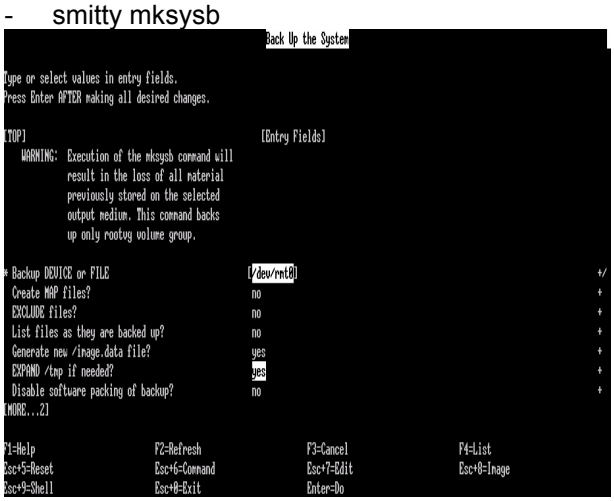
7.1.1. 시스템 백업 & Disk 복제 (alt disk)

현재 운영중인 시스템과 동일한 이미지를 테이프나 CD 로 백업하며, 시스템 장애시 백업 시점과 동일한 환경으로 운영체제를 복구할 수 있다. 대상은 오직 rootvg 만 해당되며, 마운트된 파일시스템을 대상으로 한다.

- 관련 명령어

구 분	내 용	
명령어	/usr/bin/mksysb (shell script)	
옵션	-i	mkszfile 명령어를 수행 (rootvg 정보를 수집하여 /image.data 파일 생성/갱신)
	-X	/tmp 공간 부족 시 자동으로 크기 조정
관련 파일	/image.data, /bosinst.data	

- smitty mksysb



- 그냥 O/S 를 alt 뜰 때는 다음과 같은 옵션들을 주로 사용.

기본: # alt_disk_install -C hdisk1

image.data 수정 후 받을 때 (mirror 된 Disk 중 1 copy 만 받는 게 목적)

alt_disk_install -i /image.data -C hdisk1

application 살아있는 상태에서 log 같이 file 변환이 심한 부분만 빼고 받을 때 (뺄 부분은 /etc/exclude.rootvg 에다 정의)

alt_disk_install -e /etc/exclude.rootvg -C hdisk1

7.1.2. 일반 데이터 백업

사용자가 생성한 데이터를 대상으로 시스템에서 제공하는 백업 명령어들을 이용하여 저장 가능한 매체를 통해 백업을 수행한다. 가장 일반적인 명령어로 tar, backup 등이 있다.

- tar

파일명을 기준으로 백업하는 가장 일반적인 명령어로 사용이 쉽다. 단일파일 크기가 8GB 까지 백업 가능하다.

구 분	내 용
-c	지정된 백업 대상을 archive
-x	백업된 파일을 지정된 곳으로 extract
-t	백업된 파일 목록을 확인 (화면으로 표시)
-v	진행 상황을 화면으로 출력
-f	백업 장치를 지정
명령어	/usr/bin/tar

- backup
tar 에 비해 기능과 유연성이 뛰어나지만, 사용이 어렵다

구 분	내 용
-i	표준입력을 통해 받은 파일 정보를 이용하여 백업
-u	/etc/dumpdates에 백업에 대한 정보를 기록/갱신 (Incremental Backup)
-Level	백업 레벨을 지정. default=9 (0~9 레벨 / Incremental Backup)
-v	진행 상황을 화면으로 출력
-f	백업 장치를 지정
명령어	/usr/sbin/backup

- backup 예제
[/]# cd /UMI/DW/ascentech/NHN -> 백업할 디렉토리로 이동...
[/UMI/DW/ascentech/NHN]#ls -alt -> 현재 디렉토리 내의 파일들에 대해 알아 보았다.
total 65409016
drwxrwxrwx 24 root system 4096 Apr 9 11:51 ..
-rw-r----- 1 portal db2grp1 15126561318 Mar 29 23:50 SYSDISKDS.exp
-rw-r----- 1 portal db2grp1 367 Mar 29 23:50 exportmsg.log
drwxr-xr-x 2 portal db2grp1 256 Mar 29 23:23 .
-rw-r--r-- 1 portal db2grp1 501 Mar 29 23:16 1.sql
-rw-r----- 1 portal db2grp1 18353761462 Mar 29 23:08 SYSCPUDS.exp
-rw-r----- 1 portal db2grp1 2680228 Mar 29 21:30 IPADDRESST.exp
-rw-r----- 1 portal db2grp1 1428869 Mar 29 21:30 SERVERMASTERT.exp
-rw-r----- 1 portal db2grp1 4958428 Mar 29 21:30 NODEMASTERT.exp
[/UMI/DW/ascentech/NHN]#find . -print -> 이 명령은 현재 디렉토리에 있는 것들을 Print 하는 내용이다. 아래 backup 명령을 수행하기 위해 내용을 확인해 보았다.
.
./NODEMASTERT.exp
./SERVERMASTERT.exp
./IPADDRESST.exp
./1.sql
./exportmsg.log
./SYSCPUDS.exp
./SYSDISKDS.exp
[/UMI/DW/ascentech/NHN]#tctl -f /dev/rmt0 rewind -> tctl 이란 명령은 단순히 tape Device 를 작동하기 위한 명령이다. rewind 옵션은 되감기

[/UMI/DW/ascentech/NHN] **#find . -print|backup -ivf /dev/rmt0**
Please mount volume 1 on /dev/rmt0
... and press Enter to continue

Backing up to /dev/rmt0

(이하생략)
Done at Thu Apr 12 13:35:01 2007; 65409000 blocks on 1 volume(s)
[/UMI/DW/ascentech/NHN]#

- find . -print|backup -ivf /dev/rmt0 설명
현재 디렉토리의 파일들을 나열(find . -print)해서 해당 내용을 backup 명령을 통해서 Tape(/dev/rmt0)으로 백업을 받도록 하는 명령이다.

7.1.3. Copy 명령어 모음

- cpio 명령어

cpio [option] [device_name]
cpio [option] [file_name]

디렉토리 내의 파일들을 백업
ls | cpio -ocv > backup.cpio // 현재 디렉토리의 파일들을 backup.cpio 파일로 백업
ls backup.cpio // backup.cpio 파일이 정상적으로 생성되었는지 확인
cpio -it < backup.cpio // backup.cpio 파일에 기존 파일들이 존재하는지 확인

backup.cpio 파일을 현 디렉토리에 복원
mkdir a // 복구를 위한 새 디렉토리 생성
cd a // 새 디렉토리로 이동
cpio -ivc < ../backup.cpio // 현재 디렉토리의 파일들을 backup.cpio 란 파일로 백업

특정 파일만 복원
cpio -ivc "in* < ../backup.cpio // in 으로 시작하는 파일만 복구

- dd 명령어
지정한 블록 크기만큼 파일을 복사할 때 사용하는 명령어
dd [option] [변수=값]
- 시스템 전체를 테이프 장치로 백업: # dd if=/ of=/dev/st0
- 파티션을 테이프 장치로 백업: # dd if=/dev/sda3 of=/dev/st0
- 하드 to 하드: # dd if=/dev/sda of=/dev/sdb

- gzip/gunzip 명령어
gzip: 압축하는 명령어 / gunzip(=gzip -d) : 압축 푸는 명령어
gzip [option] [file_name]
gzip 에는 파일을 묶는 기능이 없어 -r 옵션으로 디렉토리를 압축하면 디렉토리 내의 파일들이 각각 gz 파일로 압축된다.
gunzip [option] [file_name]

- bzip2/bunzip2 명령어
gzip 보다 압축률이 좋음 (-r 옵션은 없음)
bunzip2 (=bzip2 -d)

- compress / uncompress

압축률이 가장 낮음

uncompress(=compress -d)

- tar 명령어

여러 개의 파일을 하나로 묶는 명령어

tar [option] [생성할 tar 파일] [묶을 파일]

tar [option] [tar 파일] -C [풀어줄 위치]

tar 를 이용한 증분 백업

mkdir /data

mkdir /backup

cp -f /bin/* /data // /bin 의 파일들을 모두 /data 에 복사

cp -f /bin/* /home // /bin 의 파일들을 모두 /home 에 복사

>> step1: Full Backup : /home 과 /data 의 파일들을 /backup 에 백업

tar zcfp /backup/home_bak_full.tgz -g /backup/home_backup /home

tar zcfp /backup/data_bak_full.tgz --listed-incremental /backup/data_backup /data

>> step2: 1 차 증분 백업: /data 와 /home 에 파일을 추가한 후 /backup 에 백업

cp /sbin/* /data/

cp /sbin/* /home/

tar zcfp /backup/home_bak_1.tgz -g /backup/home_backup /home

tar zcfp /backup/data_bak_1.tgz -g /backup/data_backup /data

>> step3: 2 차 증분 백업: /data 에만 파일을 추가한 후 /backup 에 백업

cp -f /boot/* /data/

tar zcfp /backup/home_bak_2.tgz -g /backup/home_backup /home

tar zcfp /backup/data_bak_2.tgz -g /backup/data_backup /data

증분 백업 복구

rm -rf /data // /data 삭제

rm -rf /home // /home 삭제

>> step1: Full Backup 파일 복구

cd backup

tar zxf home_bak_full.tgz -C /

tar zxf data_bak_full.tgz -C /

ls -l /home | more

ls -l /data | more

>> step2: 1 차 증분 백업 파일 복구

tar zxf home_bak_1.tgz -g ./home_backup -C /

tar zxf data_bak_1.tgz -g ./data_backup -C /

ls -l /home | more

ls -l /data | more

>> step3: 2 차 증분 백업 파일 복구

tar zxf home_bak_2.tgz -g ./home_backup -C /

tar zxf data_bak_2.tgz -g ./data_backup -C /

ls -l /home | more

ls -l /data | more

7.2. 복구

7.2.1. 시스템 복구

시스템 백업 미디어를 통해 운영체제를 복구하며, 백업 시점 이후의 변경 데이터들은 복구가 되지 않으므로 부가적인 데이터 복구가 필요하다.

- 복구 사전 절차

구 분	내 용
부트 장치 확인	테이프 부팅 가능 여부 확인 (1: 가능 / 0: 불가능) # bootnfo -e
	서비스로 부팅 가능한 장치들이 내림차순으로 표시된다. # bootlist –m service –o
서비스 부트 장치 등록	서비스 부팅이 가능하도록 테이프 장치를 등록한다. # bootlist –m service rmt0

- 복구 절차

테이프를 이용한 시스템 부팅은 가능하나 특성상 I/O 시간이 느리므로 부팅 가능한 CD 와 시스템백업 미디어를 같이 사용하여 부팅을 진행한다. 운영체제 설치 시와 마찬가지로 System Management Services (SMS) 메뉴를 이용하여 복구를 진행한다.

구 분	내 용
SMS 메뉴 1단계	- 운영체제 CD 및 mksysb 테이프를 넣고 시스템 재시작 - memory, keyboard, network, scsi … 화면 출력 시 “1” 입력 - “7. Select Boot Options” 선택

```
Main Menu

 1   Select Language
 2   Change Password Options NOT available in LPAR mode
 3   View Error Log
 4   Setup Remote IPL (Initial Program Load)
 5   Change SCSI Settings
 6   Select Console NOT available in LPAR mode
 7   Select Boot Options


--------------------------------------------------------------------------------
Navigation keys:

                                                X = eXit System Management Services
--------------------------------------------------------------------------------
Type the number of the menu item and press Enter or Select a Navigation key: _
```

SMS 메뉴 2단계	“1. Select Install or Boot a Devices” 선택

```
 1. Select Install or Boot a Device
 2. Select Boot Devices
 3. Multiboot Startup




--------------------------------------------------------------------------------
Navigation keys:
M = return to main menu
ESC key = return to previous screen                    X = eXit System Management Services
--------------------------------------------------------------------------------
Type the number of the menu item and press Enter or Select a Navigation key: _
```

SMS 메뉴 3단계	“3. CD/DVD 선택”

```
  Select Device Type
 1.  Diskette
 2.  Tape
 3.  CD/DVD
 4.  IDE
 5.  Hard Drive
 6.  Network
 7.  None
 8.  List All Devices




--------------------------------------------------------------------------------
Navigation keys:
M = return to main menu
ESC key = return to previous screen                    X = eXit System Management Services
--------------------------------------------------------------------------------
Type the number of the menu item and press Enter or Select a Navigation key: _
```

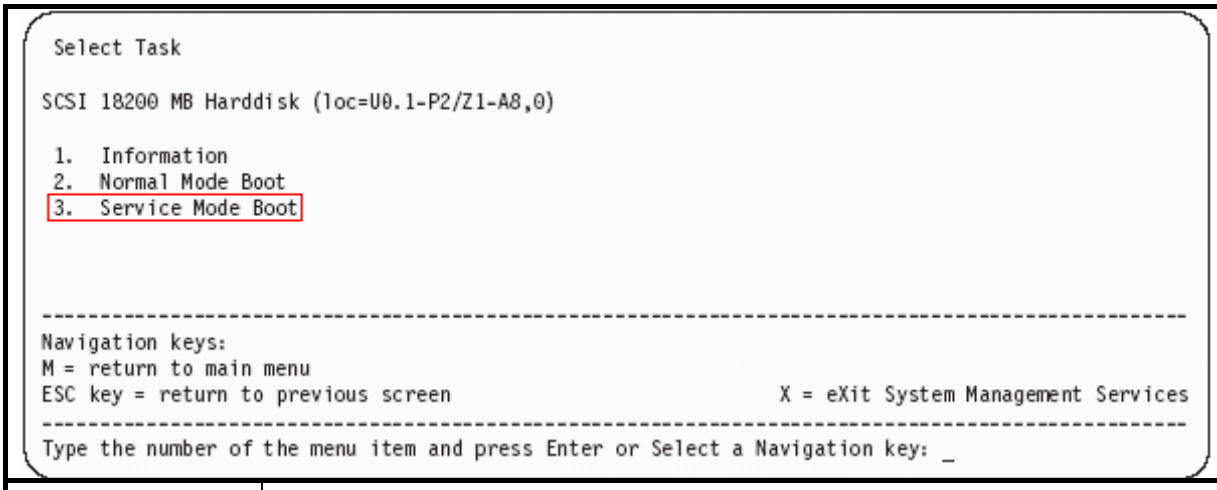
SMS 메뉴 4단계	부팅할 장치의 Position 값을 “1”로 설정 (하드디스크 예)

```
  Select Device
Device  Current  Device
Number  Position  Name
 1.        1      SCSI 18200 MB Harddisk (loc=U0.1-P2/Z1-A8,0)
 2.        -      SCSI 18200 MB Harddisk (loc=U0.1-P2/Z1-A9,0)
 3.        -      SCSI 18200 MB Harddisk (loc=U0.1-P2/Z1-Aa,0)
 4.    None
 5.   List all devices




--------------------------------------------------------------------------------
Navigation keys:
M = return to main menu
ESC key = return to previous screen                    X = eXit System Management Services
--------------------------------------------------------------------------------
Type the number of the menu item and press Enter or Select a Navigation key: _
```

SMS 메뉴 5단계	“Service Mode Boot” 선택



Service 부트 단계	시스템 복구 메뉴 선택 후 테이프 장치를 선택한다. 확인 절차가 끝나고 Restore가 진행되며, 종료 후 자동으로 재 시작한다

7.2.2. Restore

tar 및 pax 를 통해 백업 받은 대상은 동일 명령어를 이용하여 복구하면 된다. backup 명령어를 통해 백업 받은 대상은 restore 명령어를 통해 복구가 가능하다.

- 사용 명령어

구 분	내 용
-T	백업 목록을 화면으로 출력
-x	백업된 대상을 파일 단위로 복구
-r	백업된 대상을 파일시스템 단위로 복구
-v	진행 상황을 화면으로 출력
-f	백업 장치를 지정
명령어	/usr/sbin/restore

- restore 예제

backup tape 를 DAT Device 에 넣고 아래 명령을 실행하시기 바란다.

#lsdev -Cc tape --> rmt# available 확인
#tctl -f /dev/rmt0 rewind
#cd restore 할 Directory
#restore -xvf /dev/rmt0

7.3. 상황 별 디스크 교체 절차

사전 요건: 모든 작업에 앞서서 system image backup 과 data backup 이 반드시 수행되어야 하며 rootvg 가 아닌 volume group 에서의 disk 교체 상황

7.3.1. Case 1

상황: physical volume 하나로 구성된 volume group 에서 disk 의 장애가 발생하였고, data access 가 되지 않는 상황

*조치방법: 새로운 디스크를 추가하여 volume group 을 재생성하고, file system 을 재 작성하여 미리 backup 된 backup media 로부터 data 를 복구해야 한다

*작업순서:

- 기존의 volume group 을 삭제 또는 export 한 후, 시스템에서 디스크 정보를 삭제한다

#reducevg uservg hdisk1
#rmdev -dl hdisk1
hdisk1 deleted
또는

#varyoffvg uservg
#exportvg uservg
#rmdev -dl hdisk1
hdisk1 deleted
- 시스템을 shutdown 한 후, 기존의 장애 디스크를 제거하고, 새로운 디스크를 추가한다..
- 시스템을 rebooting 한 후, 시스템에 구성된 disk 를 확인한다.
#lsdev -Cc disk
hdisk0 Available 00-00-0S-0,0 16 Bit SCSI Disk Drive
hdisk1 Available 00-00-0S-1,0 16 Bit SCSI Disk Drive
#lspv
hdisk0 004308010610eac7 rootvg
hdisk1 none none
- volume group 을 새로 생성한 후 기존의 data 를 위한 file system 을 생성한다.
(예: PP size 가 32MB 이고, hdisk1 하나로 구성된 volume group 의 생성)
#mkvg -s 32 uservg hdisk1 or
#smitty mkvg
#smitty jfs
- data backup media 로부터 data 를 restore 한다.
- data 가 올바르게 복구되었는지 확인한다

7.3.2. Case 2

상황: physical volume 한 개로 구성된 volume group 에서 disk 장애 발생 (data 는 access 가능한 상태임)

*조치 방법: 시스템 shutdown 후 새로운 disk 를 추가하고, 시스템을 재 부팅 시킨다. (hot-swap 지원 시는 shutdown 절차 없이 디스크 장착 후, cfgmgr 명령을 실행한다.)
디스크 구성을 확인하고, 기존의 volume group 에 추가한다. 이어서 data 의 migration 을 수행한다. 작업이 완료되면 volume group 으로부터 장애 디스크를 제거하고, 시스템으로부터 장애 디스크를 삭제한다.
*작업 순서:
- 디스크를 추가한다.
#lspv
hdisk0 004308010610eac7 rootvg
hdisk1 00430801dd899e6b uservg
hdisk2 none none
- 기존 volume group 에 새로 장착된 디스크를 추가시킨다.
#extendvg -f uservg hdisk2
#lspv
hdisk0 004308010610eac7 rootvg
hdisk1 00430801dd899e6b uservg
hdisk2 004308010c27279e uservg
- 기존 hdisk1 에 있는 data 를 hdisk2 로 옮긴다.
#migratepv hdisk1 hdisk2
- data 가 정상적으로 옮겨졌는지 확인한다.
#lspv -l hdisk1
#lspv -l hdisk2
hdisk1:

LV NAME	LPs	PPs	DISTRIBUTION	MOUNT POINT
lv00	85	85	01..40..28..16..00	/oracle
lv03	39	39	39..00..00..00..00	/srcdata
lv02	39	39	39..00..00..00..00	/work
loglv00	1	1	00..01..00..00..00	N/A
lv01	100	100	00..39..00..00..00	/ora_data
paging00	2	2	00..00..02..00..00	N/A

- 기존 장애 disk 를 volume group 으로부터 제거한다.
#reducevg uservg hdisk1

- 장애 디스크를 시스템으로부터 제거한다.
#rmdev -dl hdisk1

7.3.3. Case 3

상황: OS mirror 된 volume group 의 장애 disk 교체

*조치방법: mirror 된 volume group 을 unmirror 시킨 후, 장애디스크를 제거한다.
제거 방법은 위의 2 번 경우와 동일하다. 제거 후에 새로운 디스크를 volume group 에 추가한 후, 다시 미러를 구성한다.
*작업순서:
- 장애 디스크 확인
#errpt
613E5F38 1212232603 P H LVDD I/O ERROR DETECTED BY LVM
A668F553 1212092003 P H hdisk1 DISK OPERATION ERROR
- volume group 의 미러 해제.
#unmirrorvg uservg hdisk1
- volume group 으로부터 disk 제거
#reducevg uservg hdisk1
- 시스템으로부터 disk 제거
#rmdev -dl hdisk1
- 디스크를 교체 후 volume group 에 새로운 디스크 추가
#extendvg uservg hdisk1
- 다시 미러 재구성.
#mirrorvg uservg

7.3.4. Case 4

상황: hdisk1 과 hdisk2 그리고, hdisk3 로 구성된 volume group 에서의 hdisk3 장애시 디스크 교체. (data access 불가)

* 조치방법: 3 개의 디스크로 구성된 volume group 의 경우, 한 개의 disk 가 failure 되더라도 quorum 이 유지 됨으로 인해 volume group 전체가 깨지진 않는다. 위의 경우와 동일하게 디스크를 추가 후, 손상된 disk 에서 가지고 있던 file system 에 대해서 backup 으로부터 data 를 복구해야 한다.
* 작업순서:
- 장애 디스크 확인
#errpt
CD546B25 1212234003 I O SYSPFS FILE SYSTEM RECOVERY REQUIRED
D2A1B43E 1212234003 P U SYSPFS FILE SYSTEM CORRUPTION
F7DDA124 1212233603 U H LVDD PHYSICAL VOLUME DECLARED MISSING
52715FA5 1212233603 U H LVDD FAILED TO WRITE VOLUME GROUP STATUS ARE
A668F553 1212092003 P H hdisk3 DISK OPERATION ERROR
- 장애디스크에 포함된 filesystem 의 제거
#lspv -l hdisk3
hdisk3:
LV NAME LPs PPs DISTRIBUTION MOUNT POINT
samplelv 3 3 00..03..00..00..00 /sample/data
tivoli 20 20 00..06..00..00..00 /tivoli
#umount /sample/data
#umount /tivoli
#rmfs /sample/data
#rmfs /tivoli
- 디스크를 volume group 으로부터 제거한다. file system 이 여러 disk 에 걸쳐 있는 경우 위 명령은 실패할 수도 있지만 다음 명령으로 제거 가능하다.
#reducevg -d uservg hdisk3
- 장애 디스크를 system 으로부터 제거한다.

#rmdev -dl hdisk3
- 새로운 디스크 장착 후, volume group 에 추가한다.
#extendvg uservg hdisk3
- 새로운 disk 에 기존의 file system 을 재 생성한다.
#smitty jfs
- 생성한 file system 을 mount 하고, 미리 backup 된 media 로부터 data 를 복구한다.

7.3.5. Case 5

상황: hdisk1 과 hdisk2 그리고, hdisk3 로 구성된 volume group 에서의 hdisk3 장애시 디스크 교체 (data access 가능)

*조치방법: data access 가 가능한 상태이기 때문에 새로운 디스크 volume group 에 추가한 후, 기존의 file system 과 logical volume 을 새로 추가된 disk 로 이동시킬 수 있다. 때로는 data 의 변동이 거의 없는 경우, backup 으로부터 복구도 좋은 선택일 수 있다.
*작업순서:
- 장애디스크 확인
#errpt -a
- 시스템에 새로운 disk 장착
#lspv
hdisk0 00191169d980d2db rootvg
hdisk1 001911696ded263c uservg
hdisk2 001911696ded2d99 uservg
hdisk3 001911696e164352 uservg
hdisk4 none none
- 새로 추가된 disk 를 uservg 에 추가
#extendvg uservg hdisk4
- 장애디스크(hdisk3)로부터 새로운 디스크(hdisk4)로 모든 logical volume 의 이동
#migratepv hdisk3 hdisk4
- volume group 으로부터 장애 디스크 제거
#reducevg uservg hdisk3
- 시스템으로부터 장애디스크 제거
#rmdev -dl hdisk3

7.3.6. Case 6

상황: hdisk1 과 hdisk2 그리고, hdisk3 로 구성된 volume group 에서의 hdisk2 와 hdisk3 디스크 동시 장애 (access 불가)

조치방법: 이 경우는 Quorum 이 구성되지 못한 상태이기 때문에, volume group 은 자동적으로 vary off 되고, 그러므로 이 volume group 은 사용할 수 없는 상태가 된다 (단, SSA 에서는 예외). 이 경우에는 반드시 data 의 backup 이 있어야 하고, 그것으로부터 data 를 복구하여야 한다. (이와 같은 장애의 경우 volume group 이 활성화 되어 있는 paging space 를 가지고 있다면 시스템 crash 또는 dump 를 발생시킬 수 있다.)
*작업순서:
- error disk 확인.
#errpt or errpt -a
- corrupt 된 volume group 제거
#exportvg uservg
- 시스템으로부터 디스크 제거
#rmdev -dl hdisk2
#rmdev -dl hdisk3
- 새로운 디스크 장착 후 volume group 재생성.
#smitty vg 이용.
- lv 및 file system 재생성.
- data restore

8. 소프트웨어 관리

AIX 에서의 소프트웨어들은 보통 SMIT 에 의해 관리된다. 이는 설치/제거시의 복잡한 선택적 옵션들을 관리자가 알아보기 쉽도록 구성되어 있기 때문이며, 실제 수행되는 명령은 installp, instfix, lslpp 등에 의해 처리된다. 또한, Licensed Program Products (LPP)라는 소프트웨어 패키지 개념을 사용하며, Backup File Format (BFF)이라는 자체적인 설치이미지 포맷을 가지고 있다.

8.1. 설치

8.1.1. 설치 메뉴 구성

SMIT 실행 시 단축메뉴인 “installp”를 입력하면 다음과 같은 하위 메뉴가 표시된다.

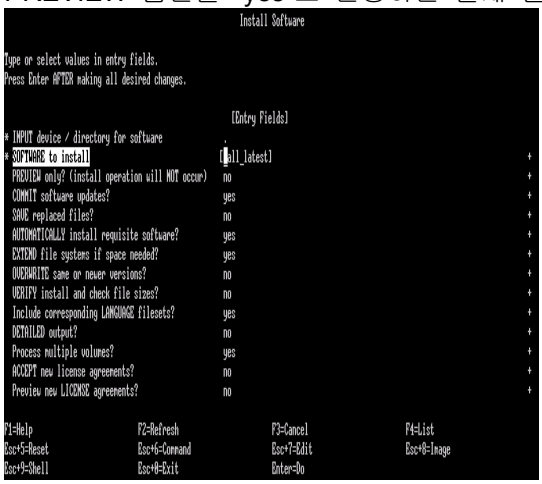
구 분	내 용
Install Software	최종적인 버전의 패키지를 기준으로 표시되며 원하는 최종 버전설치 시 선택
Update Installed Software to Latest Level (Update All)	현재 설치된 패키지의 업그레이드 시 선택
Install Software Bundle	시스템에서 지정한 번들에 해당하는 패키지 설치 시 선택
Update Software by Fix (APAR)	Fix 패키지 설치 시 선택
Install and Update from ALL Available Software	모든 버전의 패키지가 표시되며, 원하는 버전을 선택하여 설치 시 선택

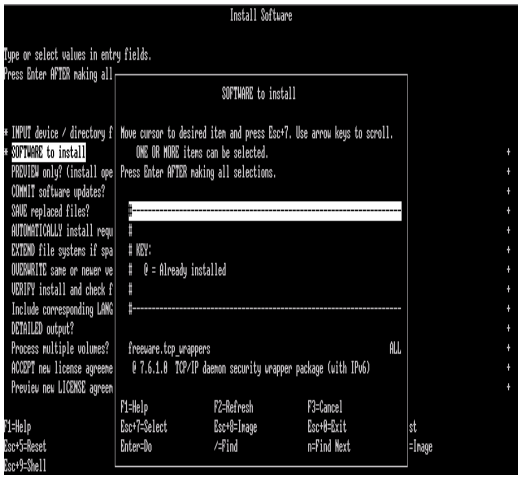
8.1.2. Install Software 메뉴

초기 메뉴에서 인스톨 미디어 또는 설치 이미지 파일이 있는 디렉토리를 지정한다

설치할 패키지들의 목록을 확인한 후 원하는 파일셋을 선택하여 설치한다.

PREVIEW 옵션을 “yes”로 변경하면 실제 설치는 진행하지 않고 성공여부를 점검할 수 있다.

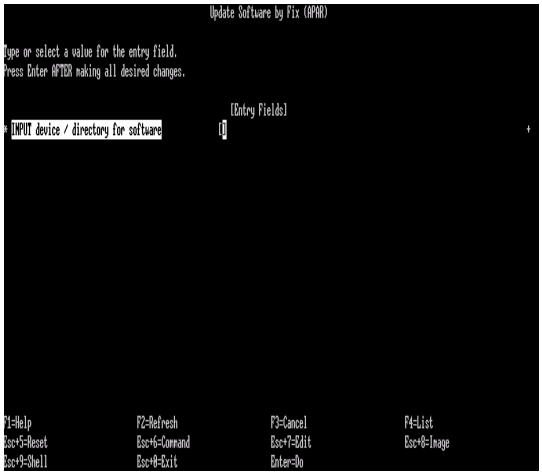


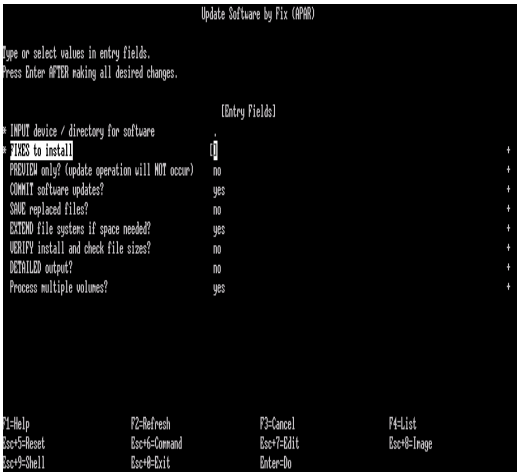


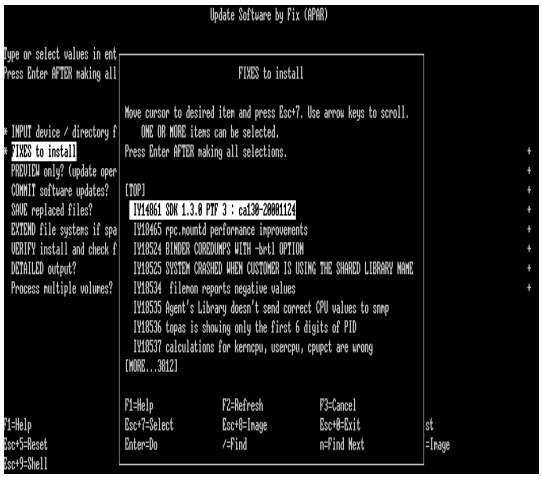
COMMIT 옵션 변경에 따라 파일셋의 설치 상태가 “Commited / Applied”로 구분된다. Applied 상태의 파일셋은 이전 버전으로의 복귀 시 Reject 과정을 거치면 가능하지만 Commited 상태의 파일셋은 불가능하다. 보통의 경우, PTF 를 적용 시 Applied 로 설치 후 안정성이 확인되면 Commited 로 전환하는 것이 일반적인 방법이다.

8.1.3. Update Software by Fix (APAR)

초기 메뉴에서 업데이트 미디어 또는 PTF/APAR 이미지 파일이 있는 디렉토리를 지정한다 설치할 PTF/APAR 들의 목록을 확인한 후 원하는 PTF/APAR 선택하여 설치한다.







- Command 로 통한 iFIX 설치

```
instfix -ik   IZ1589                ; 패치 설치여부 확인
emgr -e clvm_10_2.epkg.Z –p         ; Preview, Rebooting 필요성 여부 확인
emgr -e clvm_10_2.epkg.Z –X         ; Interim Fix 설치
emgr –l                             ; 설치결과 확인
```

- Command 로 통한 eFIX 제거

```
emgr –l                             ; iFix List
emgr -lv3 -L lavel                  ; iFix List (Level 3)
emgr -r -L lavel                    ; iFix Remove
```

8.2. 유지 관리

시스템에 설치된 소프트웨어를 유지 관리하는 것은 설치보다 많은 주의를 요구한다. 운영 중에 발견된 소프트웨어적인 결함이나 기타 장애에 대한 대응, 상호 연관성이 있는 소프트웨어에 대한 보호 등 많은 고려사항이 존재한다. 현재 설치된 소프트웨어에 대한 확인은 “lslpp” 명령을 통해 이루어진다.

8.2.1. 설치 소프트웨어 확인

- lslpp 명령

구 분	내 용	
명령어	/usr/bin/lslpp	
옵션	-l	파일셋 명, 최종 레벨, 설치 상태, 파일셋에 대한 설명을 표시
	-f	파일셋에 포함된 파일들의 설치 경로를 표시
	-h	설치 및 업데이트 히스토리 정보를 표시

8.2.2. 패치 다운로드

IBM 소프트웨어와 관련된 Patch 는 몇 가지 유형이 있다. efixes, Program Temporary Fix (PTF), Authorized Program Analysis Report (APAR) 등이 있으며 지정된 웹사이트에서 선택적으로 다운로드 받아서 설치할 수 있다

.

https://techsupport.services.ibm.com/server/aix.fdc

위의 URL 을 따라서 접속하면 다음과 같은 페이지가 표시되며 서버, 버전 및 원하는 Fix 를 확인하여 메뉴를 선택하고 다운로드를 진행한다

8.2.3. 패치 관련 용어

1) PTF (Program Temporary Fix)

PTF 는 보고된 결함 부분에 대한 픽스이다. 이 픽스는 일시적이고 PTF 의 다음 버전에 포함될 경우 없어진다. PTF 는 단일 픽스를 포함할 수도 있지만 일반적으로 여러 개의 픽스를 포함하며 단일 파일세트와 연관되어 있다.
예를 들어, PTF U476294 는 파일세트 PEX_PHIGS.graPHIGS.rte.base 용이다. 특히 U476294 는 파일세트 PEX_PHIGS.graPHIGS.rte.base 에서 5.1.0.10 으로 업데이트한다. 5.1.0.10 은 Version.Release.Maintenance/Modification.Fix (버전.릴리스.유지/변경.픽스) (V.R.M.F)으로 알려져 있다. PTF 는 11 개의 픽스(밑에 정의된 APAR 로 알려짐)가 포함된다.

2) PMR (Problem Management Record)

PMR 은 고객 보고 문제에 관한 추적 레코드이다.

3) APAR (Authorized Program Analysis Report)

APAR 는 픽스와 문제 관리 레코드를 연계시킨다. APAR 번호를 사용해 요구 픽스를 얻는다. 소프트웨어 요구사항을 문서화하는 경우 PTF/PMR 번호보다는 APAR 번호를 사용하는 것이 좋다. instfix -ivk APAR_NUMBER 명령을 사용해 항상 시스템 상의 APAR 설치 여부를 결정한다. 반면 설치된 PTF 는 추적기능이 없다.
위의 U476294 와 관련된 예에서 PTF 는 IY18782, IY18936, IY18950, IY19534, IY19690 (2 개의 결점), IY19765, IY20521, IY20877, IY20919 및 IY20921 등의 APAR 를 포함한다. 임의의 APAR 는 다른 APAR 을 PTF 에 포함시켜 파일세트 연관성을 얻는다. PTF 는 여러 개의 APAR 픽스를 포함한다는 점에서 APAR 및 PTF 는 밀접하게 연계되어 있다. APAR 는 PTF 패키징을 이용해 전달되는 단일 픽스이다.

4) ML (Maintenance Level)

ML 은 Base Operating System(BOS) 또는 옵션 소프트웨어 제품을 현 릴리스 레벨로 업그레이드하는데 필요한 서비스 업데이트이다. 신규 AIX 5L 서비스 전략의 일환으로 2006 년 시작된 ML 은 TL 로 대체되었다. TL 을 다음과 같이 정의한다.

5) TL (Technology Level)

TL 은 반년에 한 번 출시되는 AIX 5L 릴리스에 관한 신규 용어이다. 이 릴리스는 하드웨어, 소프트웨어 기능 및 서비스 업데이트 기능을 포함한다. 첫 번째 TL 은 소프트웨어 서비스, 하드웨어 기능 및 설정기능만 있다. 두 번째 TL 은 하드웨어 기능 및 설정기능, 소프트웨어 서비스 및 신규 소프트웨어 기능을 포함한다. TL 을 설치하는 작업은 "모 아니면 도" 방식으로 봐야 한다. 즉 필수 요건들을 추가해 부분이 아닌 전 TL 을 설치해야 한다. TL 설치 전 시스템을 백업한다.

6) SP (Service Pack)

SP 는 용이한 사용자 식별을 위해 한 데 모아진 TL 간에 출시되는 서비스 전용 업데이트(PTF 로도 알려져 있음)들로 이루어져 있다. 이와 같은 픽스들은 전반적이고 중요한

보안 관련 사항들을 나타낸다. 각 릴리스에 대한 최신 TL(예, 5300-04 및 5200-08)상의 N, N-1 릴리스(예, V5.3 및 V5.2)의 경우, SP 가 제공된다.

7) CSP (Concluding Service Pack)

CSP 은 TL 에 관한 최근의 서비스 팩이다. CSP 는 서비스 팩과 같이 전반적이고 주요한 보안 관련 사항들에 관한 픽스들을 포함하지만 이와 같은 범주에 들어가는 신규 출시 기술레벨에서 나오는 픽스들을 포함하기도 한다. 따라서 CSP 는 신규 기술레벨의 일환으로 출시되었던 서비스 가운데 상당히 작은 부분집합을 포함한다. 임시 픽스의 활용을 통해 CSP 는 TL 상의 서비스를 연장시킨다.

8) 임시 픽스

임시 픽스라는 용어는 " 비상 픽스" 또는 "efix"를 대신하는 용어로 사용된다. 여전히 비상 픽스라는 용어가 적용되는 상황(최소한의 테스트로 밤중에 픽스가 이루어지는 상황)이 있지만 포괄적인 테스트를 거친 업데이트를 적용할 때까지는 임시 상태를 의미한다는 점에서 임시 픽스라는 용어가 더 적합하다. 각 릴리스에 대한 최근의 두 개의 TL 상의 기술지원 릴리스(예, V5.3 및 V5.2)의 경우 비-보안 관련 사항을 나타내는 임시 픽스가 제공된다.

9. 네트워크 관리

9.1. hostname 설정

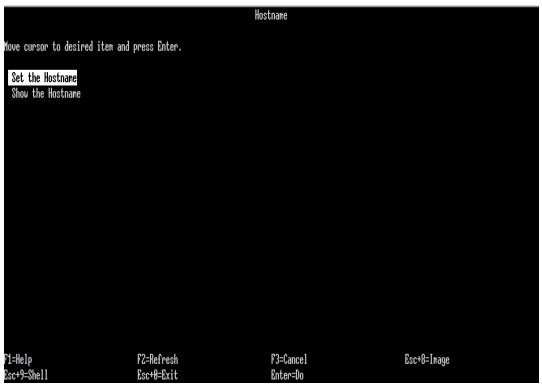
호스트 명은 시스템에 부여하는 명칭으로 반드시 고유한 이름일 필요는 없으나, 네트워크로 연결된 시스템들 간의 원활한 호스트 관리를 위해서 고유한 이름을 부여하는 것이 일반적이다.

9.1.1. 사용 명령어

구 분	내 용
관련 파일	/etc/hosts
명령어	/usr/bin/hostname

9.1.2. smitty hostname

호스트 명 설정을 위해 “Set the Hostname”을 선택 후 원하는 호스트 명을 입력한다.



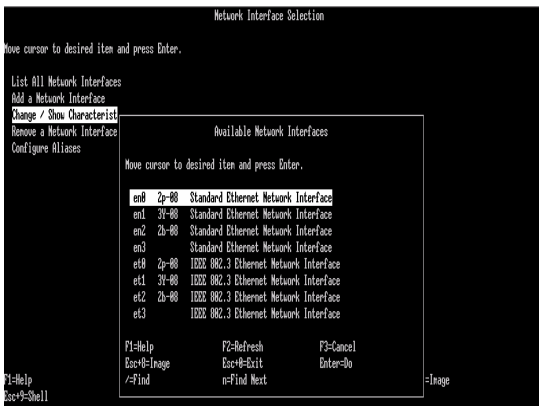
9.2. ip 설정

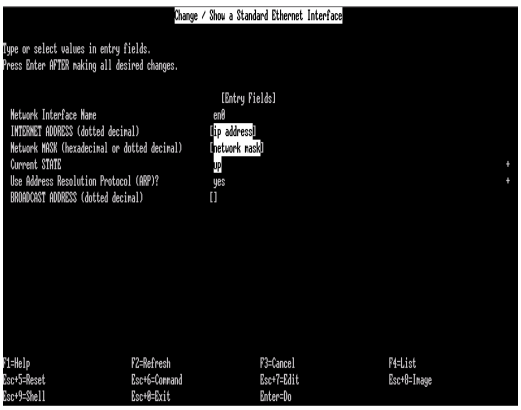
네트워크 서비스를 위해 각 시스템에는 고유한 통신주소가 필요한데 가장 일반적인 통신 방법인 TCP/IP 를 사용하기 위해 IP 주소를 할당한다. 네트워크 구성 환경에 따라 고유의 공인(일반적인 Internet 환경)/사설(일반적인 Intranet 환경) IP 주소를 각 시스템에 부여한다. IP 주소를 할당하기 위해선 시스템에 통신 어댑터가 존재해야 하며, 가장 범용적인 Ethernet 어댑터를 사용한다. AIX 에서는 장치관리자(cfgmgr)를 통해 손쉽게 어댑터(예. ent0, ent1)를 추가할 수 있으며, 어댑터와 연결된 논리적인 인터페이스(예. en0, en1, …)도 자동으로 구성된다.

9.2.1. 관련 명령어

구 분	내 용
-a	시스템의 모든 인터페이스를 출력
명령어	/usr/sbin/ifconfig

9.2.2. smitty inet

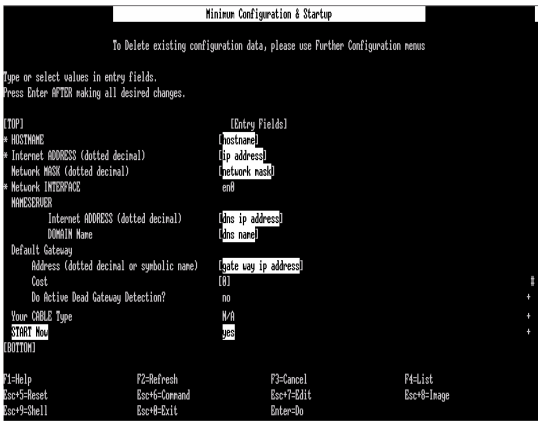




9.3. 기본적인 tcp/ip 설정

IP 주소를 부여 받은 시스템은 서로 다른 네트워크 상의 시스템들과 통신하기 위해 라우터 또는 게이트웨이를 이용하여 통신한다. 또한, 도메인 명을 사용하는 네트워크 환경에서는 IP 주소를 도메인 명으로, 도메인 명을 IP 주소로 변환하여 시스템에 알려주는 DNS 의 IP 주소 역시 알아야 한다.

9.3.1. smitty mktcpip



9.4. EtherChannel 구성

EtherChannel 은 원래 Cisco 사의 trademark 로서, 일반적으로는 network interface aggregation 또는 trunking 이라고도 한다. 즉, EtherChannel 은 특정 hardware 나 software 제품을 지칭하는 것은 아니고, Cisco 에서 만든 일종의 기술 표준이다.
이 기술을 이용하면, 2 개 또는 4 개의 Ethernet interface 를 모아서, 하나의 새로운 virtual Ethernet interface 를 만들어낼 수 있다. 이렇게 하여, 4 개의 Ethernet interface 에 하나의 공통 IP address 를 줄 수 있으므로, bandwidth 를 크게 늘릴 수 있고, network load balancing 및 high availability 도 구현할 수도 있다. 이 기술을 사용하기 위해서는 먼저 network switch 가 EtherChannel 표준을 지원하는 것이어야 하고, 또 OS 에서도 EtherChannel 을 지원해야 한다.

9.4.1. EtherChannel 방식

1) Standard Mode
외부로 나가는 traffic 이 어느 interface 를 이용하는지를, 목적지의 IP address 를 hashing 하여 그 마지막 bit 에 따라 결정한다. 가령, 마지막 bit 가 0 이면 첫 번째 interface 를 사용하고, 1 이면 두 번째 interface 를 사용하는 방식이다. 이 방식에 따르면 외부로 나가는 packet 이 어떤 순서로 어떤 interface 를 이용할지 보장이 되지만, 전체적인 bandwidth 에 대해 load balancing 이 된다고는 보장할 수 없다.

2) Round-robin Mode
글자 그대로, 각 interface 마다 한 개씩의 packet 을 분배해서 던져주는 방식이다. 이 방식으로는 전체 bandwidth 의 load balancing 이 비교적 잘 되지만, packet 이 밖으로 보내지는 순서는, 전체 channel 에 대해 원래 보낸 순서와는 다소 다를 수도 있다.

3) Network IP backup Mode (netif_backup)
이 방식은 비교적 새로 추가된 방식으로 AIX 433 maintenance level 8 을 적용하면 이 mode 가 사용 가능해진다. 이 mode 에서는 한번에 하나의 adapter 만을 active 하게 사용한다. 만약 이 연결에 문제가 발생하면, channel 은 자동적으로 현재 active 한 Ethernet 과 backup 상태였던 Ethernet 을 swapping 한다. 즉, 이 mode 에서는 load balancing 효과는 없고, availability 만을 보장하게 된다. 그렇다면 다른 방식에 비해서 기능이 떨어지는 편인데 왜 굳이 이 mode 가 사용될까? 이 mode 는 switch 장비가 EtherChannel 을 지원하지 않아도 사용이 가능하기 때문이다. 일반 dummy hub 에서 사용할 때도 구성이 가능하다.

9.4.2. smitty EtherChannel

구 분	내 용
Number of Retries	ping 실패 시, retry 횟수 지정
Retry Timeout (sec)	각각의 retry마다 몇 초씩 timeout을 줄 것인지 지정
Internet Address to Ping	입력된 IP 주소로 ping을 하면서 이상유무 점검

TIP> EtherChannel 구성 방법: Active-Active, Active-Backup

9.5. 라우팅 테이블 추가 삭제

9.5.1. permanent routing table add/delete - smitty route => on inet0 ODM

```
blue:/>smitty mkroute

Add Static Route
==> chdev -l inet0 $arg7 -a route=$1,$arg,$arg2,$arg3,$arg4,$arg5 (...)
==>
inet0 changed

blue:/>netstat -rn
(...)
10.1/16          10.0.0.1          UG        0          0 en0      -      -
<== added

blue:/>lsattr -El inet0
authm          65536                        Authentication Methods            True
bootup_option no                            Use BSD-style Network Configuration True
gateway                                     Gateway                           True
hostname       blue                         Host Name                         True
rout6                                       IPv6 Route                        True
route          net,-hopcount,0,,,,,,10.1.0.0,10.0.0.1 Route                  True
route          net,-hopcount,0,,0,9.187.213.1          Route                  True
blue:/>odmget -q attribute=route CuAt

CuAt:
        name = "inet0"
        attribute = "route"
        value = "net,-hopcount,0,,0,9.187.213.1"
        type = "R"
        generic = "DU"
        rep = "s"
        nls_index = 0
```

```
CuAt:
        name = "inet0"
        attribute = "route"
        value = "net,-hopcount,0,,,,,,10.1.0.0,10.0.0.1"
        type = "R"
        generic = "DU"
        rep = "s"
        nls_index = 0

smitty rmroute deleted the static route both ODM and current routing table.
blue:/>netstat -rn | grep 10.1
10.1/16          10.0.0.1          UG        0          0 en0      -      -
blue:/>
blue:/>smitty rmroute
Remove Static Route
==>  chdev -l inet0 -a delroute=$1,$arg2,$2,$3
==>
inet0 changed
blue:/>lsattr -El inet0 | grep 10.1
blue:/>
blue:/>netstat -rn | grep 10.1
blue:/>
```

9.5.2. Temporary routing table add/delete - route command

```
blue:/>route add -net 10.2.0.0 -netmask 255.255.0.0 10.0.0.1
10.0.0.1 net 10.2.0.0: gateway 10.0.0.1

blue:/>netstat -rn | grep 10.2
10.2/16          10.0.0.1          UG        0          0 en0      -      -
```

Tried to remove this temporary route but failed in smitty route menu because it's not static route. But we can remove it by route delete command

```
blue:/>smitty rmroute
Remove Static Route
==>  chdev -l inet0 -a delroute=$1,$arg2,$2,$3
==>
Method error (/usr/lib/methods/chginet):
        0514-009 Cannot delete an object from the device
                 configuration database.
0821-216 chginet: Cannot delete route (net,10.2.0.0,10.0.0.1) from CuAt.

blue:/>netstat -rn | grep 10.2
10.2/16          10.0.0.1          UG        0          0 en0      -      -
blue:/>route delete -net 10.2.0.0 -netmask 255.255.0.0 10.0.0.1
10.0.0.1 net 10.2.0.0: gateway 10.0.0.1
blue:/>netstat -rn | grep 10.2
blue:/>
```

9.5.3. 10.0.x.x net addr 로의 routing entry 추가 하기

```
blue@/>ifconfig en2
en2:
flags=5e080863,c0<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,6
4BIT,CHECKSUM_OFFLOAD(ACTIVE),PSEG,LARGESEND,CHAIN
>
        inet 9.187.213.222 netmask 0xffffff00 broadcast 9.187.213.255
```

```
        tcp_sendspace 131072 tcp_recvspace 65536
blue@/>ifconfig en0
en0:
flags=5e080863,c0<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,6
4BIT,CHECKSUM_OFFLOAD(ACTIVE),PSEG,LARGESEND,CHAIN
>
       inet 10.0.0.4 netmask 0xffffff00 broadcast 10.0.0.255
        tcp_sendspace 131072 tcp_recvspace 65536
blue@/>
blue@/>netstat -rn
Routing tables
Destination      Gateway           Flags    Refs     Use  If    Exp  Groups

Route Tree for Protocol Family 2 (Internet):
default          9.187.213.1       UG       3    601138 en2      -     -
9.187.213.0      9.187.213.222     UHSb     0         0 en2      -     -   =>
9.187.213/24     9.187.213.222     U        4    246046 en2      -     -
9.187.213.222    127.0.0.1         UGHS     13       325 lo0      -     -
9.187.213.255    9.187.213.222     UHSb     0        15 en2      -     -
10.0.0.0         10.0.0.4          UHSb     0         0 en0      -     -   =>
10/24            10.0.0.4          U        0         0 en0      -     -
10.0.0.4         127.0.0.1         UGHS     0         2 lo0      -     -
10.0.0.255       10.0.0.4          UHSb     0         0 en0      -     -
127/8            127.0.0.1         U        7      1603 lo0      -     -

blue@/>

blue@/>route add -net 10.0.0.0 -netmask 255.255.0.0 10.0.0.1
10.0.0.1 net 10.0.0.0: gateway 10.0.0.1
blue@/>netstat -rn
Routing tables
Destination      Gateway           Flags    Refs     Use  If    Exp  Groups

Route Tree for Protocol Family 2 (Internet):
default          9.187.213.1       UG       3    601101 en2      -     -
9.187.213.0      9.187.213.222     UHSb     0         0 en2      -     -   =>
9.187.213/24     9.187.213.222     U        4    246018 en2      -     -
9.187.213.222    127.0.0.1         UGHS     13       325 lo0      -     -
9.187.213.255    9.187.213.222     UHSb     0        15 en2      -     -
10.0.0.0         10.0.0.4          UHSb     0         0 en0      -     -   =>
10/24            10.0.0.4          U        1         0 en0      -     -   =>
10/16            10.0.0.1          UG       0         0 en0      -     -
10.0.0.4         127.0.0.1         UGHS     0         2 lo0      -     -
10.0.0.255       10.0.0.4          UHSb     0         0 en0      -     -
127/8            127.0.0.1         U        7      1603 lo0      -     -

Route Tree for Protocol Family 24 (Internet v6):
::1              ::1               UH       0         0 lo0      -     -
blue@/>
<== Added entry is the belows:
10/16            10.0.0.1          UG       0         0 en0      -     -
```

10. 시스템 모니터링

10.1. nmon

10.1.1. nmon Download

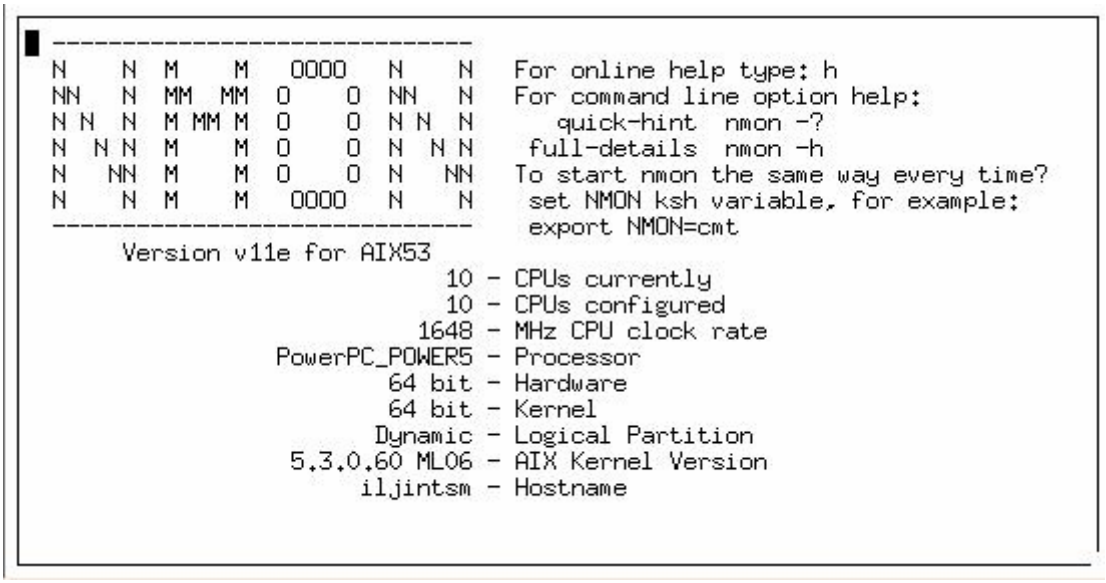
- nmon download link
- http://www-941.ibm.com/collaboration/wiki/display/WikiPtype/nmon
- nmon analyser download link
- http://www-941.ibm.com/collaboration/wiki/display/WikiPtype/nmonanalyser
- nmon 성능: AIX 와 리눅스 성능 분석을 위한 무료 툴 (한글) Link
- http://www.ibm.com/developerworks/kr/library/au-analyze_aix/

10.1.2. file upload to Server

- hostname:/home/osh
#/home/osh #>ls -lrt
-rw-r----- 1 osh staff 597126 10 월 1 일 10:56 nmon4aix_11e.tar.gz (현재 최신 버전)
#/home/osh #>mkdir /ibm/nmon11 (nmon 을 설치할 Directory)
#/ibm #>chmod 755 /ibm/nmon11
#/ibm #>cd /home/osh
#/home/osh #>ls -lrt
-rw-r----- 1 osh staff 597126 10 월 1 일 10:56 nmon4aix_11e.tar.gz
#/home/osh #>cp ./nmon4aix_11e.tar.gz /ibm/nmon11/ (설치할 Directory 에 Copy)
#/home/osh #>cd /ibm/nmon11
#/ibm/nmon11 #>ls -lrt
-rw-r----- 1 root system 597126 10 월 1 일 10:58 nmon4aix_11e.tar.gz
#/ibm/nmon11 #>gunzip nmon4aix_11e.tar.gz (gz 확장자를 푸는 방법)
#/ibm/nmon11 #>tar -xvf nmon4aix_11e.tar (tar 확장자 푸는 방법)
x nmon, 1371 바이트, 3 개의 미디어 블록.
x nmon_aix51, 335989 바이트, 657 개의 미디어 블록.
x nmon_aix52ml2, 386667 바이트, 756 개의 미디어 블록.
x nmon_aix52ml5, 381241 바이트, 745 개의 미디어 블록.
x nmon_aix53, 403428 바이트, 788 개의 미디어 블록.
x README_v11.txt, 2527 바이트, 5 개의 미디어 블록.
#/ibm/nmon11 #>ls -lrt
총계 5992
-rwxr-xr-x 1 210 ipsec 1371 1 월 29 일 2006 nmon
-rwxr-xr-x 1 210 staff 335989 7 월 5 일 2006 nmon_aix51
-rwxr-xr-x 1 210 214 381241 7 월 5 일 2006 nmon_aix52ml5
-rwxr-xr-x 1 210 staff 386667 7 월 5 일 2006 nmon_aix52ml2
-rw-r--r-- 1 210 ipsec 2527 7 월 5 일 2006 README_v11.txt
-rwxr-xr-x 1 210 staff 403428 7 월 6 일 2006 nmon_aix53 (현재 OS 와 맞는 버전을 가지고 사용)
-rw-r----- 1 root system 1525760 10 월 1 일 10:58 nmon4aix_11e.tar
#/ibm/nmon11 #>chown root:system * (ownership 변경)

10.1.3. nmon 실행 Test

#/ibm/nmon11 #>./nmon_aix53 (topas 처럼 사용 가능하며 동작하는 지 체크를 위해 실행)



10.1.4. nmon Data gather

#/ibm/nmon11 #>vi nmon1.sh -> (성능 Data 를 gather 할 때 script 를 작성해서 동작하도록 한다)
cd /ibm/nmon11/nmonout #Data File 이 쌓이는 곳을 지정한다.
/ibm/nmon11/nmon_aix53 -f -d -t -s 60 -c 1440
위 옵션식으로 사용한다. 두 가지 옵션만 알면 된다
-s 60 (60 초 단위로 gather)
-c 1440 (count 옵션으로 1440 번 수행
60 *1440 = 60 초 단위로 1440 번 수행한다는 얘기 -> 60(sec)*60(min)*24(hour) -> 하루 단위로 Data Gather
#/ibm/nmon11 #>mkdir nmonout -> (nmon Data 가 쌓이는 Directory)
#/ibm/nmon11 #>chmod 755 nmonout
#/ibm/nmon11 #>ls -lrt
총계 6208
-rwxr-xr-x 1 root system 1371 1 월 29 일 2006 nmon
-rwxr-xr-x 1 root system 335989 7 월 5 일 2006 nmon_aix51
-rwxr-xr-x 1 root system 381241 7 월 5 일 2006 nmon_aix52ml5
-rwxr-xr-x 1 root system 386667 7 월 5 일 2006 nmon_aix52ml2
-rw-r--r-- 1 root system 2527 7 월 5 일 2006 README_v11.txt
-rwxr-xr-x 1 root system 403428 7 월 6 일 2006 nmon_aix53
-rw-r----- 1 root system 1525760 10 월 1 일 10:58 nmon4aix_11e.tar
-rw-r--r-- 1 root system 70 10 월 1 일 11:17 nmon1.sh
drwxr-xr-x 2 root system 256 10 월 1 일 11:17 nmonout
#/ibm/nmon11 #>
#/ibm/nmon11 #>vi nmon1.sh (입맛에 맞게 수정)
cd /ibm/nmon11/nmonout

#하루 Schedule 단위로 할 때는 아래와 같이 한다.
#1. Detailed
#/ibm/nmon11/nmon_aix53 -f -d -t -s 60 -c 1440
#2. Normal
#/ibm/nmon11/nmon_aix53 -f -d -t -s 300 -c 288

/ibm/nmon11/nmon_aix53 -f -d -t -s 300 -c 84 #5 분 단위로 84 번 수행(7 시간 수행)
"nmon1.sh" 4 행, 117 자

```
#/ibm/nmon11 #>./nmon1.sh
ksh: ./nmon1.sh: 0403-006 실행 권한이 거부되었다.
#/ibm/nmon11 #>chmod 755 ./nmon1.sh
#/ibm/nmon11 #>./nmon1.sh -> nmon data gather script 수행
#/ibm/nmon11 #>ps -ef|grep nmon
    root 2031820       1   0  11:21:00      -  0:00 /ibm/nmon11/nmon_aix53 -f -d -t -s 300 -c 84
#/ibm/nmon11 #>cd nmonout
#/ibm/nmon11/nmonout #>ls -lrt
-rw-r--r--    1 root     system        90969 10 월  1 일  11:21 #_071001_1121.nmon
#/ibm/nmon11/nmonout #>cd ..
#/ibm/nmon11 #>
```

위와 같이 설정하면 nmon 을 사용 가능하게 된다. 필요할 때마다 nmon1.sh 의 내용을 수정해서 사용한다.

10.1.5. nmon Schedule Gather

만약 매일 Data 를 수집하려면 아래와 같이 crontab 에 등록하면 된다.

```
#crontab -e (작성)
#------------------------------
#nmon datagather
0 0 * * * /ibm/nmon11/nmon1.sh  # 매일 0 시에 수행되게 한다.
#------------------------------
```

10.1.6. 기타

nmon 은 AIX 5.3 TL09 이상부터 OS 에 포함되어 별도 설치하지 않아도 된다.

10.2. sar

10.2.1. 사용법과 옵션

- sar [-ubdycwaqvmAMS] [-o file] t [n]
: 't'초 동안 'n'회의 sar 결과를 'file'명의 binary 로 만듦
- sar [-ubdycwaqvmAMS] [-s time] [-e time] [-i sec] [-f file]
: 'file'의 내용 중 '-s time' 부터 '-e time'까지 '-I sec'를 주기로 하여 Data 를 추출함
-> 위 결과를 파일로 저장하고자 할 경우

```
sar -d -s 01:00 -e 11:00 -i 1 -f sa20 >> new_file_name
```

- 옵션
  - -.u : CPU Data
  - -.b : Buffer 정보
  - -.d : Disk or Tape 정보
  - -.y : TTY 정보
  - -.c : System Call 정보
  - -.s : System Swapping or Switching 정보
  - -.q : Queue Length 정보
  - -.v : Text, Process, Inode, File table 정보
  - -.m : Message and Semaphore 정보
  - -.A : 모든 Data(udqbwcayvm)
  - -.M : 각 Processor 별 sar 정보

10.2.2. 예제

- sar #1 #2
=> #1 초를 주기로 #2 회 동안의 sar 정보를 모음
-. 0,5,10,15,20,25,30,35,40,45,50,55 * * * 0-6 /usr/lib/sa/sa1
=> 매 5 분마다 sa 정보 수집
-. 55 13 * * * /usr/lib/sa/sa2 -s 1:00 -e 23:40 -i 300 -A
=> 매일 13:55 분에 01:00 ~ 23:40 시간 동안 300 초를 주기로 모든 정보를 모음

10.3. vmstat

10.3.1. 사용법과 옵션

- CPU 사용 정보에 대한 전반적인 특징을 체크 할 수 있다.
- 사용 명령어: vmstat 2 10 (2 초 간격으로 10 번 수행)
- CPU 관리 기준: swap 에서 대기중인 kernel thread 수의 양이 2 이상 일 경우 병목으로 판단하며, 이 경우 CPU 증설이나 상위 모델로의 증설을 권고하는 것이 바람직하다.

10.4. ps

10.4.1. 사용법과 옵션

- CPU 를 많이 사용하는 프로세서를 체크 할 수 있다
- 사용 명령어: ps aux | head -10

10.5. iostat

10.5.1. 사용법과 옵션

- Disk 사용 정보에 대한 전반적인 특징을 체크 할 수 있다.
- 사용 명령어: iostat –d hdisk0 hdisk1 5 (5 초 간격으로 5 번 수행)
- Disk 관리 기준: 지속적으로 % iowait 가 40% 이상이거나 사용자 application 에 의해 특정 disk 의 %tm_act 가 70%를 초과하는 경우 Disk I/O 의 병목으로 판단하며, 이 경우 시스템의 구성 환경 및 파일 시스템 재구성을 통한 tuning 작업을 수행해야 한다. Tuning 이후에도 지속적으로 Disk I/O 의 병목이 발생하면 물리적 disk 증설이나 교체를 권고하는 것이 바람직하다.

10.6. netstat

10.6.1. 사용법과 옵션

- Network 사용 정보에 대한 전반적인 특징을 체크 할 수 있다.
- 사용 명령어: netstat –v
- Network 관리 기준: 지속적으로 Network bandwidth 의 30% 이상을 사용하는 경우 병목으로 판단하며, 이 경우 환경변수 조정에 의한 tuning 작업을 수행해야 한다.
  Tuning 이후에도 지속적으로 Network 병목이 발생하면 업그레이드를 권고하는 것이 바람직하다.

10.7. svmon

10.7.1. 개요

- Physical memory, Paging space 정보를 가장 정확하게 보여주는 도구이다.
- svmon 명령을 수행하는 시점의 메모리 정보를 보여준다.
- svmon 명령이 수행 되지 않는 경우에는 다음 fileset 이 설치되어 있는지 확인한다.
  : bos.perf.tools

10.7.2. 명령어 옵션

-G : Global Report
-C: 명령어 별 메모리 사용 현황을 보여줌.
-D: 세그먼트 별 메모리 사용 현황을 보여줌. (Frame 정보 포함)

-S: 세그먼트 별 메모리 사용 현황을 보여줌. (Segment 별로…)

-F: 특정 세그먼트에 대해 Frame 단위의 메모리 사용 현황을 보여줌.

-P: 프로세스 별 메모리 사용 현황을 보여줌.

-U: 사용자(login id)별로 메모리 사용 현황을 보여줌.

-W: WLM 의 class 별로 메모리 사용 현황을 보여줌.

-m: Source segment 와 Mapping segment 별로 보여줌.

-d: 사용하는 모든 프로세스 보여줌. (-U 또는-C 옵션과 함께 사용)

-I: 해당 세그먼트를 사용하는 process id 까지 보여줌.

-v: virtual space 순으로 정렬.

-g: paging space 사용량 순으로 정렬.

-p: pinned 영역 사용량 순으로 정렬.

-u: real memory 사용량 순으로 정렬.

-c: client 세그먼트만 보여줌.

-f: persistent 세그먼트만 보여줌.

-w: working 세그먼트만 보여줌.

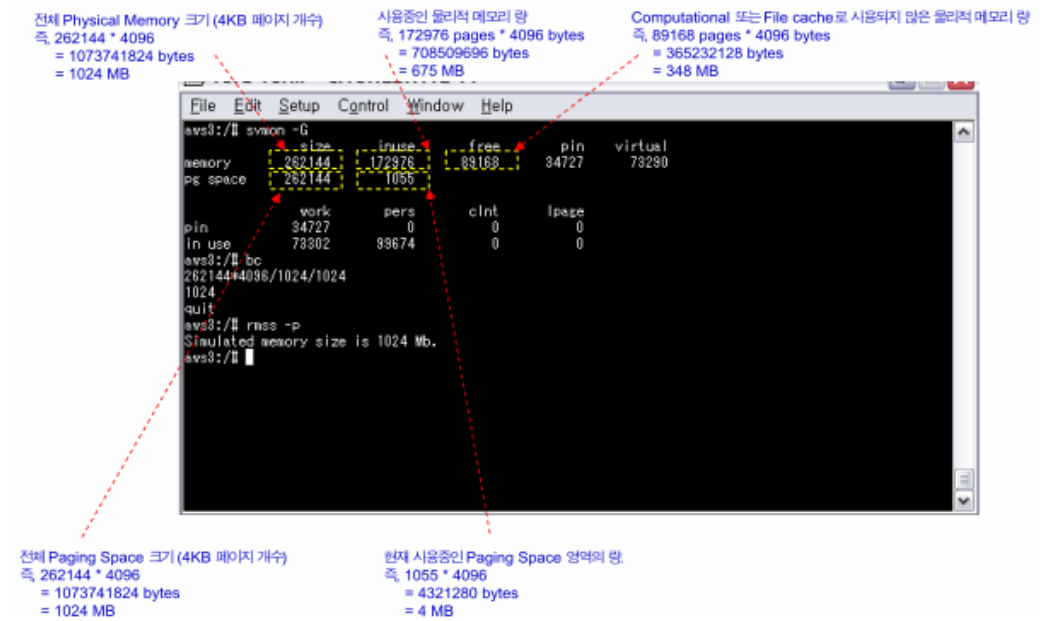
-s: system 세그먼트만 보여줌.

-n: non-system 세그먼트만 보여줌.

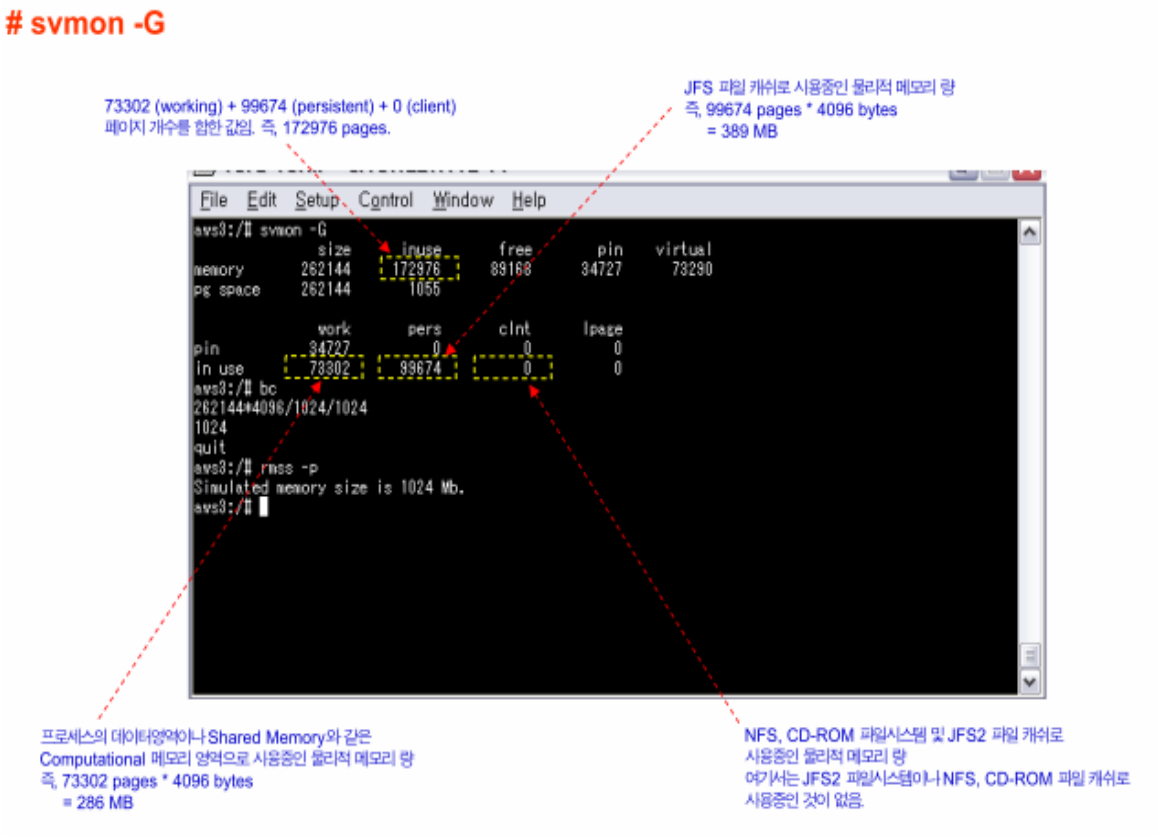
-r: 세그먼트내의 allocated page range 정보도 보여줌.

10.7.3. 현재 시점의 메모리 사용 현황Summary (I)

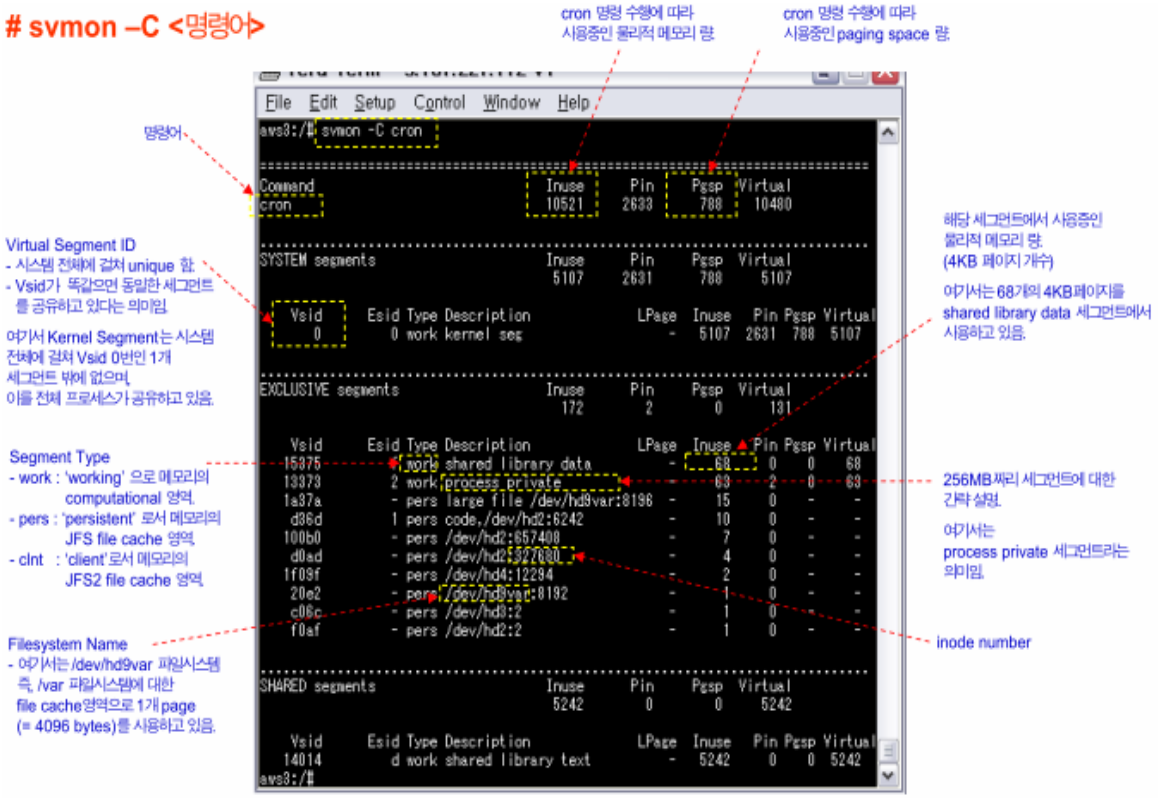
 # svmon -G



10.7.4. 현재 시점의 메모리 사용 현황 Summary (II)

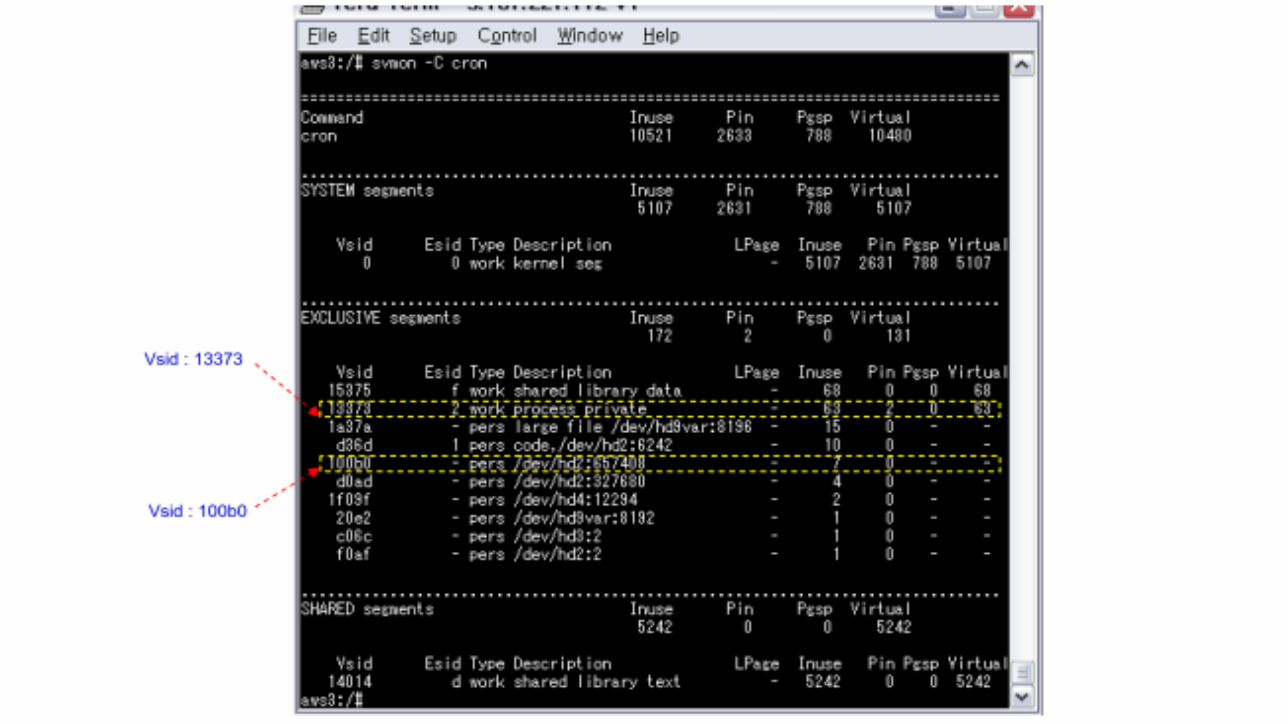


10.7.5. 특정 명령어가 사용중인 메모리 사용 현황



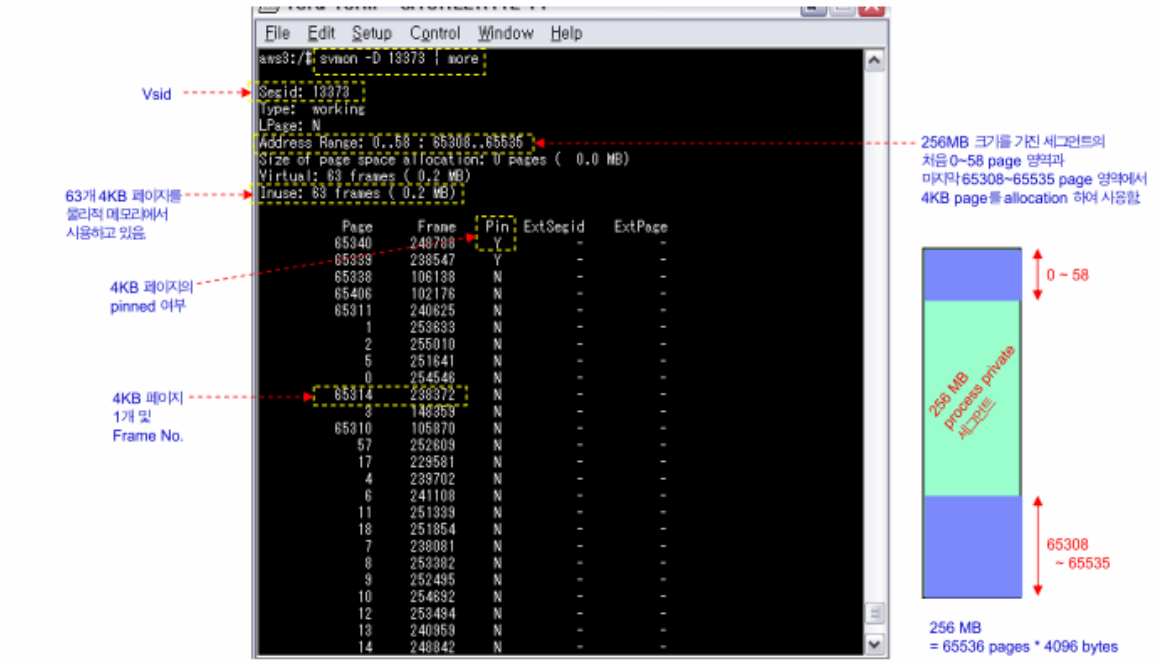
10.7.6. 특정 세그먼트에 대한 메모리 사용 현황

앞 例 에서 본 **cron** 명령어가 사용하는 메모리 현황 중에서 다음 **2**개 세그먼트를 선택하여 살펴보도록 합니다.



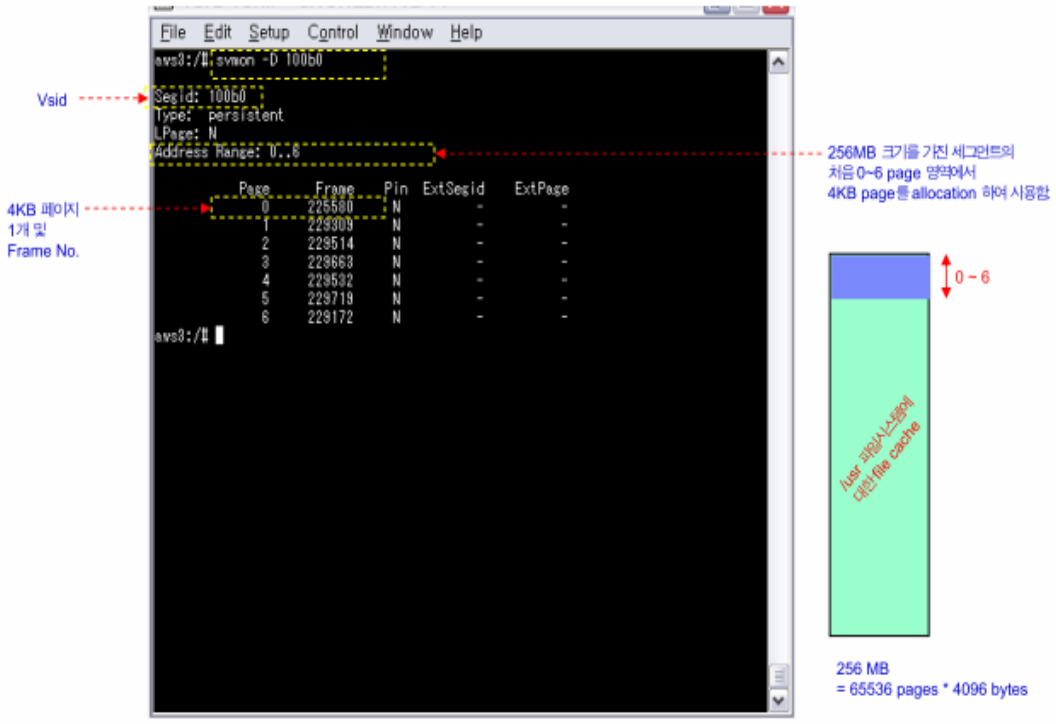
10.7.7. 특정 세그먼트에 대한 메모리 사용 현황(Frame 정보를 볼 경우) working segment 예

 # svmon –D <Vsid>



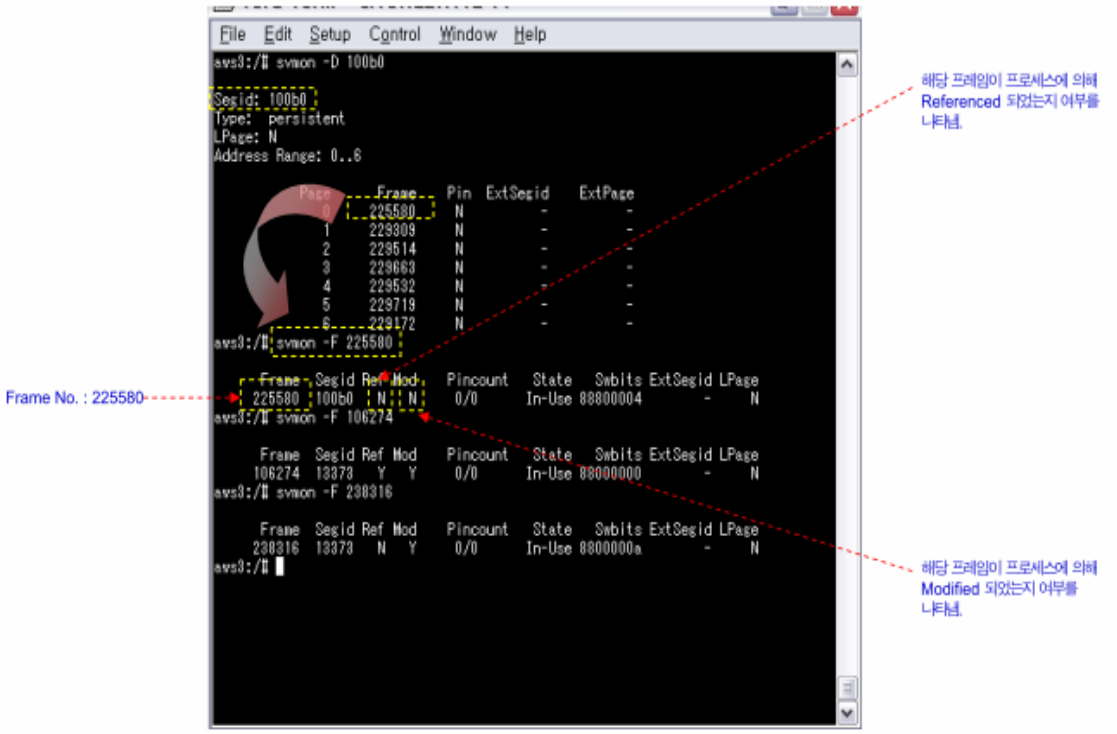
10.7.8. 특정 세그먼트에 대한 메모리 사용 현황(Frame 정보를 볼 경우) persistent segment 예

 # svmon –D <Vsid>



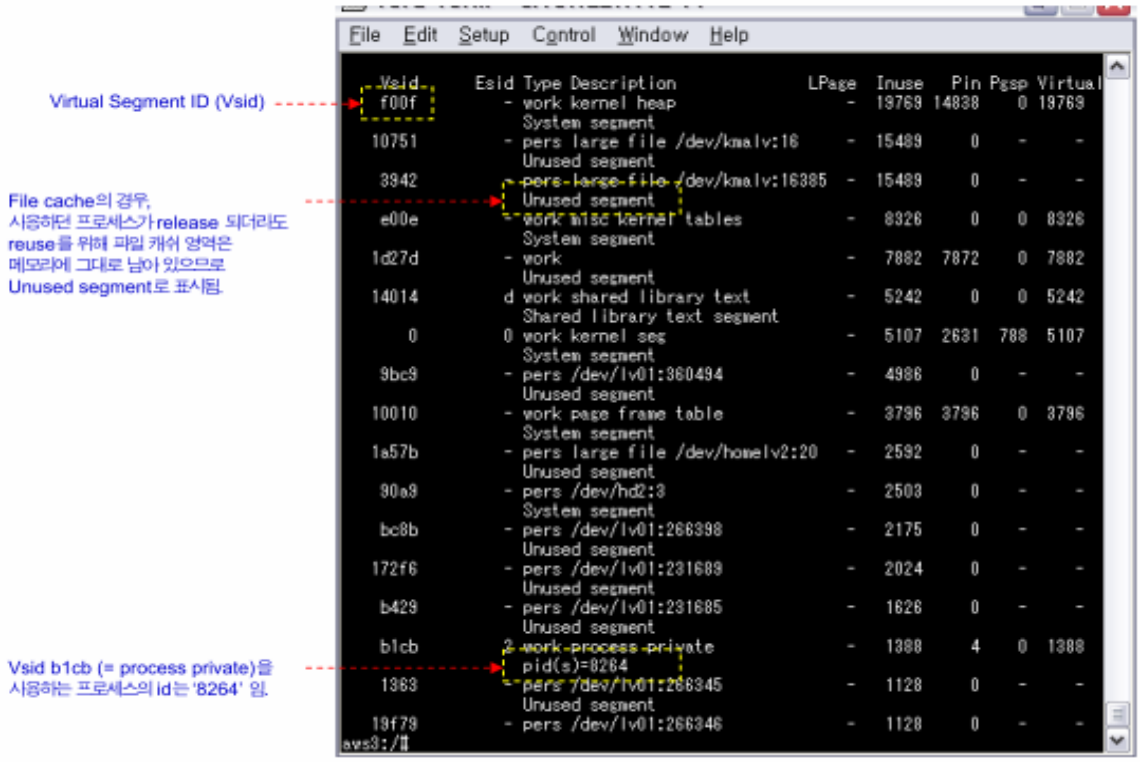
10.7.9. 특정 메모리 프레임(= 4KB 메모리페이지)에 대한 정보 조회

 # svmon –F <Frame ID>



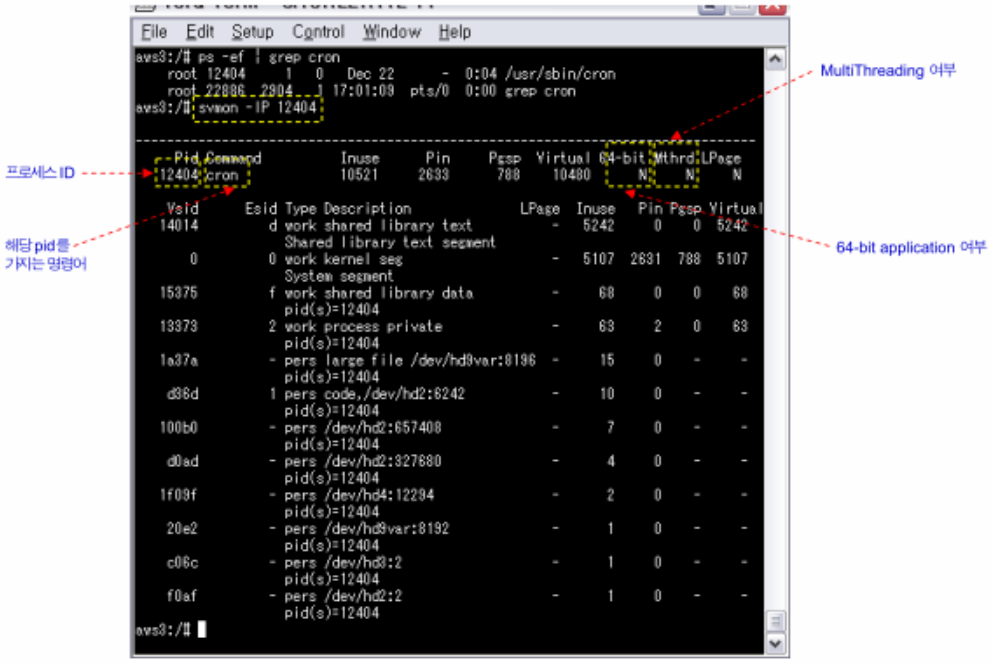
10.7.10. 세그먼트 별 메모리 사용 현황 조회

그림은 'svmon –IS | more' 명령 수행 결과입니다.



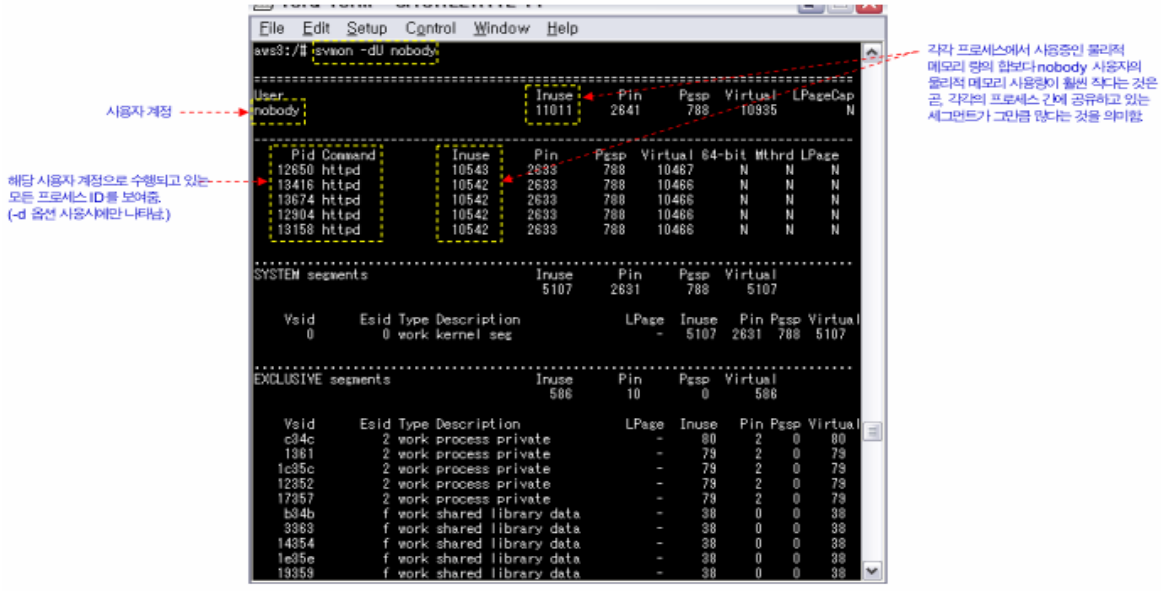
10.7.11. 특정 프로세스가 사용하는 메모리 현황 조회

svmon –P <Process ID>



10.7.12. 특정 사용자 계정이 사용하는 메모리 현황 조회

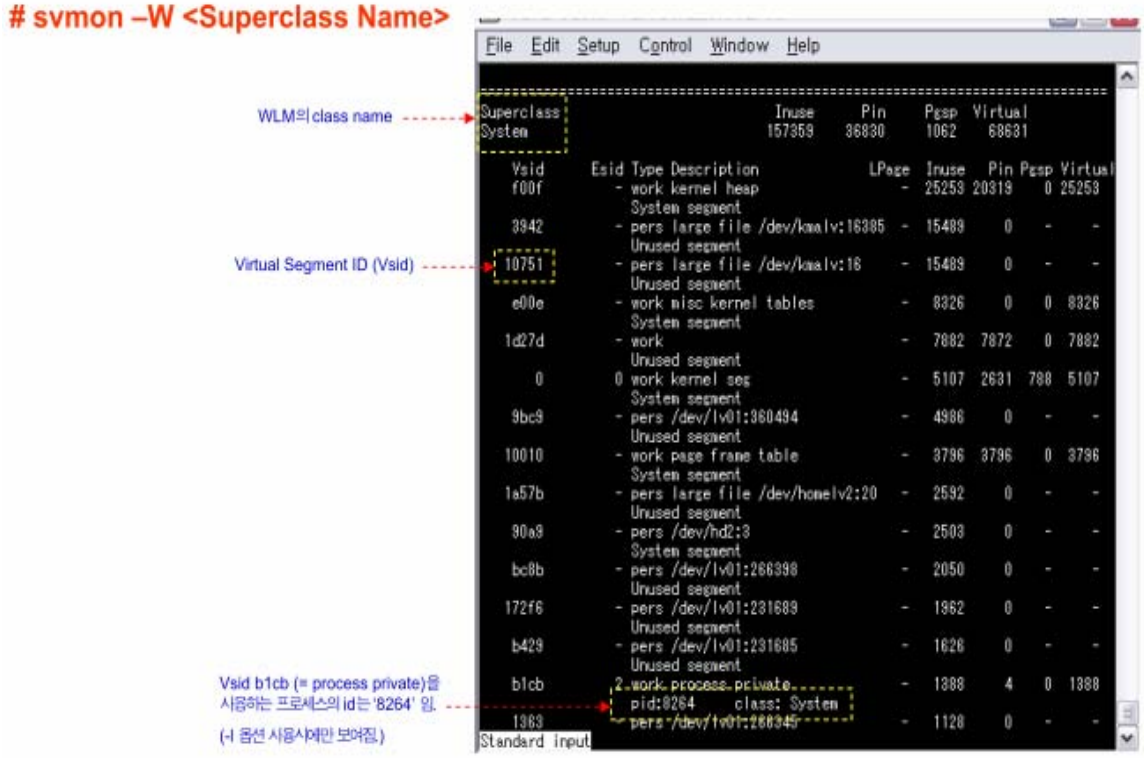
 # svmon –U <User ID>



10.7.13. WLM (Work Load Manager)의Class별 메모리 사용 현황 조회

반드시 WLM을 먼저 startup 시켜야 합니다.

 # svmon –W <Superclass Name>



10.8. 기타 모니터링

10.8.1. 모니터링 command와 종료 방법

구분	용도	종료
iptrace	- 인터넷 프로토콜에 대한 인터페이스 레벨 패킷 추적	kill 명령어로 종료
trace	- 선택된 이벤트 기록	trcstop 명령어로 종료
filemon	- 파일시스템의 성능을 모니터하고 논리적 파일, 가상 메모리 세그먼트, 논리적 볼륨(LV) 및 물리적 볼륨(PV)를 위해 I/O 활동을 보고 (backgroud로 수행)	trcstop 명령어로 종료
netmon	- 네트워크I/O와 네트워크 관련 CPU 사용에 관하여 모니터하고 통계를 보고(backgroud로 수행)	trcstop 명령어로 종료
perfpmr.sh	- performance 및 시스템 정보 gathering (backgroud로 수행)	trcstop 명령어로 종료

10.8.2. iptrace script 예

```
#!/bin/ksh
IPTRACE_LOG=/tmp/tcptest/iptrace_${1}.log
IPREPORT_OUT=/tmp/tcptest/ipreport_${1}.out
D_HOST=node1

startsrc -s iptrace -a " -a -P tcp -d $D_HOST -b $IPTRACE_LOG"
명령어/Application 수행
stopsrc -s iptrace
ipreport -ns $IPTRACE_LOG > $IPREPORT_OUT
```

10.8.3. tcpdump 사용 예

```
tcpdump -i en6 host 128.2.201.71  and port 80
```

11. 시스템 장애 처리

11.1. HACMP 장애

11.1.1. 로그 파일 확인

- 로그파일 위치

버전 5.3 이하: /tmp/hacmp.out, /usr/es/sbin/cluster/history/cluster.MMDDYYYY

버전 5.4 이상: /var/hacmp/log/hacmp.out, /usr/es/sbin/cluster/history/cluster.MMDDYYYY

- 문제가 발생하였을 경우 1차적으로는 hacmp.out 파일을 주의 깊게 확인해봐야 한다. HACMP에서 발생하는 대부분의 문제는 log/hacmp.out 파일안에서 그 원인을 찾을 수 있다. hacmp.out 파일은 hacmp event script의 모든 수행 결과를 보여주므로 매우 자세한 정보를 포함하고 있다. 너무 많은 내용이 포함되어 있어 찾고자 하는 정보를 찾기가 쉽지 않으므로 cluster.MMDDYYYY 파일을 먼저 확인하여 문제가 발생한 event와 시각을 확인한 후 hacmp.out을 확인하면 좀 더 쉽게 필요한 정보를 찾을 수 있다.

또 errpt 를 이용하여 필요한 정보를 얻을 수 있다.

11.1.2. HACMP 기동 시 한쪽 노드에서 halt가 발생하는 경우

이러한 경우는 대부분 여러 가지 이유로 인해서 두 노드간에 HACMP ODM 정보가 일치하지 않기 때문에 발생한다. 이럴 경우에는 양쪽 노드에서 모두 HACMP를 내린 후 위에서 Sync를 수행한 다음 HACMP를 기동하면 해결될 수 있다.

11.1.3. HACMP가 기동되지 않을 경우

- clcomES 데몬이 살이 있는지 확인한다. (# lssrc -s clcomES) 또는 clcomES 데몬이 사용하는 6191 포트가 보안 툴 등에 의해 막혀있지 않은지 확인하다.

- sync를 수행해 본다. sync는 모든 노드에서 hacmp가 내려가 있으면 아무 노드에서 수행하여도 되지만 online node가 있으면 online node에서 sync를 수행하여야 한다.

- clstat 명령어를 사용하여 cluster의 상태를 확인한다. 만약에 SubStatus가 “Stable” 상태가 아니면 현재 hacmp가 활성화되어 있는 노드에서 /var/hacmp/log/hacmp.out 을 확인한다. 에러가 발생한 원인을 확인하여 수작업으로 조치를 한 후 hacmp recovery를 수행하여 SubStatus가 “Stable” 상태가 되도록 한 후 hacmp를 기동하여야 한다.

※ hacmp recovery 절차

smit hacmp > Problem Determination Tools > Recover From HACMP Script Failure

=> (현재 에러가 발생한 노드를 선택한 후 Enter)

11.1.4. HACMP가 기동되었으나 Resource가 올라오지 않는 경우

- /var/hacmp/log/hacmp.out을 검색하여 기동 중에 어느 부분에서 에러가 발생하였는지 확인한다. 가능한 원인으로는 Volume Group을 varyonvg 과정에서 에러가 발생하거나, File System을 mount 하는 과정에서 에러가 발생하거나, Application Server에 등록되어 있는 스크립트를 수행하는 과정에서 에러가 발생할 수 있다. 원인을 파악한 후 조치한 후 hacmp를 종료시키고 재기동한다.

- clRGinfo 명령어를 수행하였을 때 특정 Resource Group의 State가 “Error”로 되어있으면 /var/hacmp/log/hacmp.out을 검색하여 원인을 찾아서 수동으로 조치한 후 해당 Resource Group만 Online 시켜준다.

smit cl_admin > HACMP Resource Group and Application Management > Bring a Resource Group Online > (해당 Resource Group을 선택) > (해당 노드를 지정한 후 Enter)

11.1.5. Sync 수행 시 에러가 발생하는 경우

- clcomES 데몬이 살이 있는지 확인한다. (# lssrc -s clcomES) 또는 clcomES 데몬이 사용하는 6191 포트가 보안 툴 등에 의해 막혀있지 않은지 확인하다.

- /usr/es/sbin/cluster/etc/rhosts 파일안에 모든 노드의 정확한 ip address가 들어있는지 확인한다..

11.1.6. HACMP가 정상적으로 종료되지 않을 경우

- /var/hacmp/log/hacmp.out을 검색하여 기동 중에 어느 부분에서 에러가 발생하였는지 확인한다. 가능한 원인으로는 Volume Group을 varyoffvg 과정에서 에러가 발생하거나, File System을 umount 하는 과정에서 에러가 발생하거나, Application Server에 등록되어 있는 스크립트를 수행하는 과정에서 에러가 발생할 수 있다. 원인을 파악한 후 수동으로 조치(varyoffvg VolumeGroup , umount Filesystems, 해당 프로세스 강제 종료)한 후 hacmp recovery를 수행한다.

11.1.7. Takeover가 정상적으로 수행되지 않을 경우

- /var/hacmp/log/hacmp.out을 검색하여 기동 중에 어느 부분에서 에러가 발생하였는지 확인한다. 가능한 원인으로는 Volume Group을 varyonvg 과정에서 에러가 발생하거나, File System을 mount 하는 과정에서 에러가 발생하거나, Application Server에 등록되어 있는 스크립트를 수행하는 과정에서 에러가 발생할 수 있다. 원인을 파악한 후 수동으로 조치(varyonvg VolumeGroup , mount Filesystems)한 후 hacmp recovery를 수행한다. 또는 clfindres 명령어를 수행하였을 때 특정 Resource Group의 State가 “Error”로 되어있으면 해당 Resource Group만 Online 시킨다.

11.1.8. Config_too_long 발생 시 해결 절차.

- ps 를 통해 해당 process 확인 / 관련 Parent PID 확인. 실제 Cluster event 와 상관없는 것이라면 kill 명령어로 강제 종료

- Recover From HACMP Script Failure 를 통해 해당 Script 를 Reset 시킴

- 일정 시간이 지나도 해결이 안될 경우 해당 서버를 rebooting 합니다.

- device busy 로 인한 umount fail 시 아래절차로 확인합니다.

genld-l |grep -p FS_name

(umount 가 안되는 filesystem 상에 load 된 object 가 있는 지 확인하고 해당 Process number 를 확인 kill -9 process_id)

11.2. GPFS 장애

11.2.1. 로그파일 확인 (/var/mmfs/log)

gpfs 는 운용상의 메시지나 에러 데이터를 각 노드의 /var/adm/ras 디렉토리 밑에 보관하며 mmfs.log.date.nodename 의 이름으로 생성한다. 가장 최근의 파일은 mmfs.log.lastest 이름으로 symbolic link 되어있다.

장애가 발생시에는 일차적으로 mmfs 로그 파일을 분석하여 원인을 확인하여야 한다..

다음은 mmfs 로그의 예이다.

Tue Feb 14 16:36:42 EST 2006 runmmfs starting

Removing old /var/adm/ras/mmfs.log.* files:

mv: cannot rename /var/adm/ras/mmfs.log.previous to /var/adm/ras/mmfs.log.previous.save:

No such file or directory Loading kernel extension from /usr/lpp/mmfs/bin . . . /usr/lpp/mmfs/bin/mmfskxload: /usr/lpp/mmfs/bin/aix64/mmfs64 is already loaded at 64532224. Tue Feb 14 16:36:43 2006: mmfsd64 initializing. {Version: 3.1.0.0 Built: Feb 13 2006 22:15:25} ... Tue Feb 14 16:36:44 2006: OpenSSL library loaded

Tue Feb 14 16:36:48 2006: Accepted and connected to 89.116.94.86 k155n06

Tue Feb 14 16:36:51 2006: mmfsd ready Tue Feb 14 16:36:52 EST 2006: mmcommon mmfsup invoked

Tue Feb 14 16:36:54 2006: Accepted and connected to 89.116.94.88 k155n08

Tue Feb 14 16:36:54 2006: Accepted and connected to 89.116.94.85 k155n05

Tue Feb 14 16:36:54 2006: Accepted and connected to 89.116.94.87 k155n07

Tue Feb 14 16:36:54 2006: Accepted and connected to 89.116.94.89 k155n09

Tue Feb 14 16:36:55 2006: Accepted and connected to 89.116.94.2 k156lnx02

Tue Feb 14 16:42:05 2006: Command: mount fs1 684072

Tue Feb 14 16:42:05 2006: Node 89.116.94.70 (k154n06) appointed as manager for fs1.

Tue Feb 14 16:42:30 2006: Node 89.116.94.70 (k154n06) completed take over for fs1.

Tue Feb 14 16:42:30 2006: Loaded policy 'for stripe group fs1': parsed 4 Placement Rules, 0 Migrate/Delete/Exclude Rules

Tue Feb 14 16:42:30 2006: Command: err 0: mount fs1 684072

11.2.2. GPFS daemon이 올라오지 않는 경우

GPFS daemon 의 확인은 다음과 같이 수행한다.

```
 # ps   -ef | grep mmfs
root 307398 352476    0    Mar 16      -   1:43 /usr/lpp/mmfs/bin/aix64/mmfsd64
root 352476      1    0    Mar 16      -   0:00 /bin/ksh /usr/lpp/mmfs/bin/runmmfs
```

현재의 구성은 GPFS daemon 이 시스템 부팅 시 자동적으로 올라오게 되어있으며, 자동적으로 올라오지 않을 경우에는 /etc/inittab 에 다음과 같은 항목이 들어있는지 확인한다.

mmfs:2:once:/usr/lpp/mmfs/bin/mmautoload >/dev/console 2>&1

GPFS daemon 이 올라오지 않았을 경우에는 해당 노드에서 다음과 같이 수행한다.

mmstartup

위 명령어로 gpfs daemon 이 올라오지 않을 경우에는 다음 사항을 확인하다.

1. 각 노드로 ping 테스트 등을 이용하여 GPFS 네트워크 연결이 정상적인지 확인한다. 이때 주의할 점은 서비스 네트워크가 아닌 GPFS 전용 네트워크의 연결을 확인하여야 한다.

ping nodeA_gpfs

2. 각 노드로 remote command가 정상적으로 수행되는지 확인한다. 현재의 구성은 remote command로 /bin/ssh, /bin/scp를 사용하도록 되어있다.

ssh nodeA_gpfs

3. mmfs 로그 파일의 내용을 확인하여 원인을 찾는다. 정확한 원인을 찾을 수 없다면 유지보수 업체에 연락하여 지원을 받도록 한다.

11.2.3. GPFS 파일시스템이 올라오지 않는 경우

파일시스템이 mount 되지 않을 경우에는 다음 사항을 확인한다.

1. mmlsmgr 명령어를 사용하여 현재 해당 파일시스템의 매니저 노드가 어떤 노드인지 확인한 후 매니저 노드에서 GPFS daemon이 정상적인지 확인한다

```
pfcta01@root:/>mmlsmgr
file system         manager node
---------------- ------------------
gp_pfctap_cfep1    128.5.93.143 (pfcta02_gpfs)
gp_pfctap_cfep2    128.5.93.143 (pfcta02_gpfs)
gp_pfctap_clog1    128.5.93.143 (pfcta02_gpfs)
gp_pfctap_clog2    128.5.93.143 (pfcta02_gpfs)

Cluster manager node: 128.5.93.143 (pfcta02_gpfs)
```

2. Quorum을 만족하는지 점검한다. 2 노드 클러스터 환경에서는 하나 이상의 노드에서 GPFS daemon이 올라와있고, 하나 이상의 tiebreaker disk가 available 상태로 있어야 Quorum을 만족한다.

tiebreaker disk 는 mmlsconfig 명령어로 확인한다.

```
pfcta01@root:/>mmlsconfig
Configuration data for cluster pfcta_cluster.pfcta01_gpfs:
---------------------------------------------------------
clusterName pfcta_cluster.pfcta01_gpfs
clusterId 9224882277421251870
clusterType lc
autoload yes
```

minReleaseLevel 3.2.1.3

dmapiFileHandleSize 32

pagepool 512M

maxMBpS 1600

failureDetectionTime 10

tiebreakerDisks tiebreaker_nsd01;tiebreaker_nsd02;tiebreaker_nsd03

File systems in cluster pfcta_cluster.pfcta01_gpfs:

/dev/gp_pfctap_cfep1

/dev/gp_pfctap_cfep2

/dev/gp_pfctap_clog1

/dev/gp_pfctap_clog2

3. 각 노드로 ping 테스트 등을 이용하여 GPFS 네트워크 연결이 정상적인지 확인한다. 이때 주의할 점은 서비스 네트워크가 아닌 GPFS 전용 네트워크의 연결을 확인하여야 한다.

ping nodeA_gpfs

11.2.4. GPFS 명령어가 수행되지 않는 경우

GPFS 명령어가 수행되지 않을 경우에는 다음 사항을 확인한다.

1. 해당 노드에 GPFS daemon이 정상적으로 올라와있는지 확인한다.

2. 각 노드로 ping 테스트 등을 이용하여 GPFS 네트워크 연결이 정상적인지 확인한다. 이때 주의할 점은 서비스 네트워크가 아닌 GPFS 전용 네트워크의 연결을 확인하여야 한다.

ping nodeA_gpfs

3. 각 노드로 remote command가 정상적으로 수행되는지 확인한다. 현재의 구성은 remote command로 /bin/ssh, /bin/scp를 사용하도록 되어있다.

ssh nodeA_gpfs

11.3. Filesystem umount 장애

11.3.1. fuser를 이용해서 파일시스템에 속한 파일을 사용중인 PID 찾기

fuser -cux /ggg

/ggg: 311314(root)

(파일시스템에 속한 파일을 사용중인 PID 가 리스트 된다.)

ps -fT 311314

UID PID PPID C STIME TTY TIME CMD

root 311314 291068 0 01:10:46 pts/0 0:00 ./fopen

(ps 명령으로 해당 PID 들을 확인.)

kill 311314

11.3.2. 사용중인 라이브러리가 파일시스템에 있는지 확인

genld -l | grep "/ggg"

사용중인 PID 가 있으면 kill 한다.

11.3.3. 사용중인 shared 라이브러리가 파일시스템에 있는지 확인

genkld | grep "/ggg"

해당 파일시스템에 속한 라이브러리가 현재 사용 중이면 리스트 된다.

여기에 있으면 통상 genld -l 명령으로도 확인된다.

slibclean

사용되지 않고 있는 커널과 라이브러리 모듈들을 삭제처리 한다.

11.3.4. svmon 명령을 이용해서 찾아 보는 방법

svmon -Sl | grep "/dev/fslv00"

19589 - clnt /dev/fslv00:6 s 2 0 - -

19569 - clnt /dev/fslv00:4 s 1 0 - -

6596 - clnt /dev/fslv00:5 s 0 0 - -

해당 파일시스템 관련하여 메모리에 올라와 있는 Vsid 를 확인한다.

svmon -S 19589 19569 6596 -l

Vsid Esid Type Description PSize Inuse Pin Pgsp Virtual

19589 - clnt /dev/fslv00:6 s 2 0 - -

Unused segment

19569 - clnt /dev/fslv00:4 s 1 0 - -

Unused segment

6596 - clnt /dev/fslv00:5 s 0 0 - -

pid(s)=311314

해당 Vsid 를 누가 사용 중인지 확인할 수 있다.

이 경우는 PID 311314 가 /dev/fslv00 의 inode 5 번 파일을 사용 중.

(Unused segment 는 원칙적으로 umount 를 방해하지 못한다.

11.3.5. shared memory 와 관련되어 있는 파일이 있는지 찾기

ipcs -mP

IPC status from /dev/mem as of Wed May 30 01:50:59 KORST 2007

T ID KEY MODE OWNER GROUP

Shared Memory:

m 0 0x58000844 --rw-rw-rw- root system

SID:0x1b20b PINSIZE:0 LGPG: -

m 1048577 0x0d0016f1 --rw-rw---- root system

SID:0x13523 PINSIZE:0 LGPG: -

m 1048578 0xffffffff --rw-rw---- root system

ServiceSuite 21/23 RS070531-T83

SID:0x84d8 PINSIZE:0 LGPG: -

m 1048579 0x78002117 --rw-rw-rw- root system

SID:0x13403 PINSIZE:0 LGPG: -

(shared memory 에 대한 SID 를 확인합니다.

kdb

(0)> scb 2 0x1b20b | grep gnode

(0)> scb 2 0x13523 | grep gnode

(0)> scb 2 0x84d8 | grep gnode

(0)> scb 2 0x13403 | grep gnode

(kdb 로 들어가서 해당 SID 가 혹시 gnode 포인터를 갖고 있는지 확인한다.

(이 예제에서는 gnode 포인터를 갖고 있는 것이 존재하지 않습니다.

(0)> gnode [gnode 포인터] | grep gn_rdev

gn_vnode..... F10001001783B7F8 gn_rdev..... 8000000A0000000E

(만일 gnode 포인터가 있다고 가정하면 위와 같이 해당 LV 를 확인한다..

(8000000A0000000E 는 major, minor 번호가 (0xA,0xE) 즉, (10,14) 인 LV 이다..

(0)> vfs

GFS DATA TYPE FLAGS

[생략]

9 F1000100143B4A90 013ECD20 F100010017912080 JFS2 DEVMOUNT

... /dev/fslv00 mounted over /ggg

(0)> vfs 9 | grep vmt_fsid

vmt_fsid..... 8000000A0000000E 0000000000000000

위의 gn_rdev 와 vmt_fsid 가 동일하면 아래와 같이 svmon 명령으로 해당 SID 를 사용중인 PID 를 찾아낸다.

물론 이 경우는 gn_rdev 와 vmt_fsid 가 다르므로 더 이상 진행할 필요는 없다.

svmon -S 0x1b20b -l

Vsid Esid Type Description PSize Inuse Pin Pgsp Virtual

1b20b 3 work shared memory segment s 4 2 0 4

pid(s)=122950

11.3.6. 여전히 umount가 안 되는지 최종 확인

umount /ggg

umount: 0506-349 Cannot unmount /dev/fslv00: The requested resource is busy.

ServiceSuite 22/23 RS070531-T83

kdb

(0)> vfs

GFS DATA TYPE FLAGS

[생략]

9 F1000100143B4A90 013ECD20 F100010017912080 JFS2 DEVMOUNT

... /dev/fslv00 mounted over /ggg

(0)> vfs 9 | grep -v " 0 " | egrep "F1000100143B4A9|COUNT"

COUNT GNODE VFSP TYPE FLAGS

4 F1000100179253F8 1 F1000100179250B0 F1000100143B4A90 DIR ROOT

- 여기까지 진행했는데도 여전히 umount 가 안 된다면 위와 같이 아직도 참조 COUNT 가 0 이 아닌 파일들이 DIR ROOT 이외에도 존재

- 위의 1 번~5 번까지 방법으로 다 했는데도 여전히 umount 가 안 된다면 강제덤프를 받아서 IBM 기술지원 센터의 도움을 받아야 한다.

11.4. 기타 장애 처리

11.4.1. 특정 포트를 사용하고 있는 프로세스 찾는 방법

1) netstat -Aan | grep port_number (note the PCB address).

2) lsof | grep pcb_address (note the PID).

3) ps -ef | grep pid.

그러면 port 를 사용중인 process 를 알 수 있습니다.

[실습]

root@v5:/> netstat -Aan | head -2; netstat -Aan | grep "\.23 "

Active Internet connections (including servers)

PCB/ADDR	Proto	Recv-Q	Send-Q	Local Address	Foreign Address	(state)
f1000089c00f9b58	tcp	0	0	*.23	*.*	LISTEN
f1000089c04dd358	tcp4	0	0	10.10.11.210.23	10.10.11.233.36466	ESTABLISHED
f1000089c0484b58	tcp4	0	0	10.10.11.210.23	9.187.195.77.1725	ESTABLISHED
f1000089c04e0b58	tcp4	0	0	10.10.11.210.23	9.187.211.163.1513	ESTABLISHED
f1000089c04df358	tcp4	0	309	10.10.11.210.23	9.187.213.162.1400	ESTABLISHED

```
f1000089c0424b58      tcp4       0          0          10.10.11.210.23
9.187.211.49.2787     ESTABLISHED
f1000089c04e0358      tcp4       0          0          10.10.11.210.23
9.187.213.48.3966     ESTABLISHED
f1000089c054c358      tcp4       0          0          10.10.11.210.23
9.187.213.226.5375    ESTABLISHED

root@v5:/> lsof | head -1; lsof | grep f1000089c04dd358
COMMAND     PID   USER   FD   TYPE             DEVICE
SIZE/OFF   NODE NAME
(unknown)     0   root  cwd   VDIR             10,4
2560          2 /  (/dev/)
telnetd  229532   root   0u   IPv6  0xf1000089c04dd358
0t232         TCP  v5:telne)
telnetd  229532   root   1u   IPv6  0xf1000089c04dd358
0t232         TCP  v5:telne)
telnetd  229532   root   2u   IPv6  0xf1000089c04dd358
0t232         TCP  v5:telne)

root@v5:/> ps -ef | grep 229532
    root  73860  229532   0  11:17:59  pts/2  0:00  -ksh
    root 131322  336082   1  14:30:47  pts/3  0:00  grep 229532
    root 229532   77902   0  11:17:58      -  0:00  telnetd -ag
```

[유의사항]

Long-pending 된 socket structure 를 delete 하거나, socket 을 잡고 있는 process id 를 알아내기 위하여 rmsock utility 는 가능하면 사용하지 마시기 바랍니다. AIX 를 crash 로 몰고 갈 수 있습니다.

따라서 비정상적인 socket problem 에 대해서는 그 상태에서 IBM 에 문제 분석을 요청하시던지, 급한 경우에는 해당 socket application 을 재기동 하시기 바랍니다.

(물론 경우에 따라서는 application 이 stop 되어도 socket 이 netstat output 에 남을 수 있습니다.)

해당 socket 을 사용하는 pid 를 알고 싶으신 경우에는 netstat -aAn 과 kdb 를 사용하거나 lsof 를 설치하여 lsof -i command 를 실행하여 알아낼 수 있습니다.

```
#  netstat  -Aan
.....
f100060007f44b98   tcp4    0    0    10.0.1.11.60000    10.0.1.247.54521
ESTABLISHED
```

```
#  kdb
The specified kernel file is a 64-bit kernel.
Preserving 1393063 bytes of symbol table
First symbol __mulh
.....
(0)>  sockinfo  f100060007f44b98  tcpcb
---- TCPCB ----(@ F100060007F44B98)----
     seg_next......@F100060007F44B98  seg_prev......@F100060007F44B98
     t_softerror... 00000000 t_state..... 00000004 (ESTABLISHED)
     t_timer..... 00000000 (TCPT_REXMT)
.....
proc/fd:  588/5
```

위의 예에서는 pid 588 번에 file descriptor # 5 번입니다.

11.4.2. Defunct Process 처리 방법

- Defunct Process 의 정의

AIX 에서는 각 프로세스들이 계층을 형성하고 있어 parent 프로세스가 존재합니다. 정상적으로 프로세스가 종료되면 parent 프로세스에게 그 정보가 전달이 되고 parent 프로세스는 child 프로세스에 대한 정보를 수정하게 됩니다. 만약 parent 프로세스가 child 프로세스에 대한 정보를 얻지 못하게 되면 그 child 프로세스는 defunct 로 빠지게 됩니다.

defunct 프로세스는 이미 멈춰 zombie 프로세스라고도 하는데 시스템에 그냥 떠 있는 상태로 kill 명령이 적용되지 않습니다. defunct 프로세스는 상태 정보나 자원 사용정보를 유지하기 위해 아주 적은 양의 메모리를 차지하고 있기 때문에 메모리 리소스에 대한 염려는 크지 않습니다만 간혹 max # of process 에 걸려 당황하게 되는 경우가 있습니다.

한 프로세스가 종료되었을 때 그것을 제거하는 것은 parent 프로세스의 책임입니다. 따라서 defunct 프로세스가 된 원인을 찾기 위해서는 먼저 parent 프로세스에 대한 점검이 필요합니다. defunct 프로세스 문제는 PPID 가 1 인 경우와 1 이 아닌 경우로 나뉘어집니다. PPID 가 1 인 경우는 /etc/inittab file 에 문제가 있는 경우입니다. 이 때는 inittab file 에 설정되어 있는 프로세스들을 하나하나 점검하면서 문제가 되는 프로세스를 찾아 나가야 합니다.

- PPID 가 1 인 경우 해결방법

1) 먼저 다음의 두 가지를 확인합니다.

① /etc/inittab 파일을 열어서 "install_assit" 프로세스가 있는지 보고 있다면 제거합니다. install_assis 는 OS 를 처음 설치하면 보이는 configuration tool 입니다. 설치 후에는 사용할 필요가 없으므로 제거하시면 됩니다.

② dce 의 설치여부를 확인해서 제거합니다.

2) “/etc/inittab” vs. “who .d”

/etc/inittab file 의 내용과 who -d 에서 나타나는 프로세스를 비교해 보면서 defunct 프로세스를 생성하는 원인을 찾아낼 수 있습니다. who -d 에서 나타나지 않는 프로세스를 모두 comment(:) 처리 했다가 하나씩 살리면서 어느 것이 defunct 프로세스를 만들어 내는지를 확인하여 해당 프로세스를 제거합니다.

who 명령어는 현재 login 한 사용자를 보여주는데 그 중 -d flag 는 init 에 의해 재실행되지 않고 만기된 모든 프로세스를 표시해 줍니다. 무효 프로세스에 대해 표시되는 종료 필드는 무효 프로세스의 종결 및 그 값(대기하여 리턴됨)을 포함하고 있습니다. 이 플래그는 응용 프로그램이 리턴한 오류 번호를 보고 프로세스가 종료된 이유를 판별하는 데 유용합니다.

두 output 에서 보이는 process 를 비교하여 제거하는 순서는 다음과 같습니다.

root@smactr:/> vi /etc/inittab

/etc/inittab file 을 열어서 세 번째 칼럼이 "wait" 이거나 “once”인 프로세스를 찾습니다.

root@smactr:/> more /etc/inittab

cron:2:respawn:/usr/sbin/cron

piobe:2:wait:/usr/lib/lpd/pio/etc/pioinit >/dev/null 2>&1 # pb cleanup

qdaemon:2:wait:/usr/bin/startsrc -sqdaemon

uprintfd:2:respawn:/usr/sbin/uprintfd

diagd:2:once:/usr/lpp/diagnostics/bin/diagd >/dev/console 2>&1

logsymp:2:once:/usr/lib/ras/logsymptom # for system dumps

root@smactr:/> who -d

나타나는 프로세스를 확인해서 inittab 과 비교해 보고 inittab 에만 나타나는 프로세스들을 찾습니다.

root@smactr:/> who -d

. . Feb 15 15:24 old 43520 id=SvcAgen term=15 exit=0

. . Feb 18 17:14 old 38448 id=hc term=0 exit=1

root@smactr:/> vi /etc/inittab

who -d 에서 나타나지 않은 프로세스는 모두 comment(:) 처리하여 실행하지 않습니다.

root@smactr:/> telinit q

inittab 을 재실행합니다.

3 분 정도 기다린 후에 defunct process 가 있는지 확인합니다.

root@smactr:/> ps -ef |grep defunct

모든 defunct 프로세스가 사라졌으면 /etc/inittab 에서 comment 로 막았던 프로세스를 하나씩 재실행시킵니다. 한 번에 한 프로세스씩 살리고 inittab 을 telinit 명령어로 refresh 합니다.

root@smactr:/> telinit q

- PPID 가 1 이 아닌 경우 해결 방법

만약 parent process ID 가 1 이 아닌 경우는 해당 PPID 가 PID process 를 제거하는 책임을 가진다.

이러한 프로그램이 종료된 child process 의 제거를 담당하도록 coding 이 되어야 함은 당연하지만 이러한 루틴이 포함되지 않아 지속적으로 defunct process 가 발생되면 해당 software vendor 에게 잠재적 결함이 있음을 알려야 한다.

제거되지 않고 defunct 로 남아 있는 process 의 한가지 원인은 shell 에서 발생하는데 이는 shell 이 종료되면 함께 사라지게 된다.

따라서 shell 로 기인한 경우 지속적으로 defunct 가 남아 있게 되면, 해당 shell program 이 looping 상태가 아닌지 등을 점검해 보아야 한다.

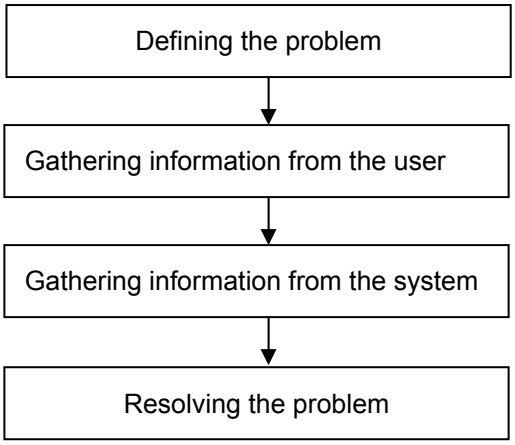
11.4.3. PSDB url

https://techdev.austin.ibm.com/psdb/systemp/searchAll.do

12. Problem Determination

12.1. 기본적인 PD 절차

문제해결(PD)을 위한 기본적인 절차는 아래와 같이 크게 4 단계로 볼 수 있다.



※ Ref. IBM redbook “Problem Solving and Troubleshooting in AIX 5L”
Chapter1. Problem determination introduction

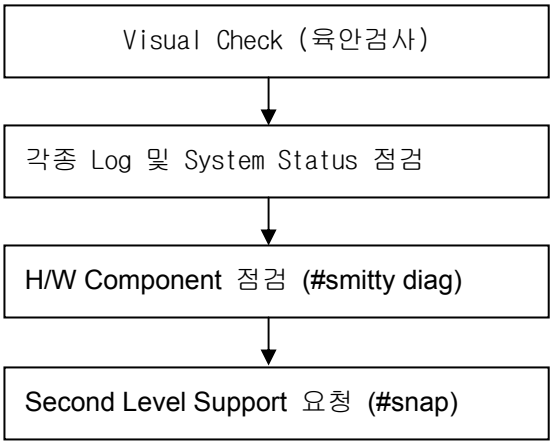
12.2. 문제 해결을 위한 기본적인 데이터와 기본적인 H/W PD 절차

12.2.1. PD 시작 전에 반드시 알아야 할 정보

- M/T, Model & S/N
- Environment (OS Version, PTF level, F/W level, Application, Sub-system)
- 문제 발생 시점과 빈도
- 증상
- 장애에 따른 영향도
- 작업가능 시간과 허용시간

12.2.2. 기본적인 H/W PD 절차

- 기본적인 H/W PD 의 절차는 다음과 같다.



- snap 명령어는 System Configuration Information 의 수집하고, 해당 정보를 압축된 tar File 로 생성한다. 생성된 정보는 Disk, Tape 또는 Remote System 으로 전송할 수 있다.
- Attention/Alarm Lamp 의 의미와 Error code 의 의미 등은 각 기종마다 다르므로 해당 기종의 Service guide 를 참조하여 H/W 문제에 대한 PD 를 진행한다.

12.2.3. H/W PD Step별 Check 대상 및 분석 방법

Step	Check 대상	분석 방법 및 참고 자료
Visual Check	Attention/Alarm Lamp	System Service Guide 참조
	OP Panel	System Service Guide 참조
	Cable 연결 상태	System Installation Guide 참조
Check Logs	Error report	#errpt | pg or #errpt -a | pg
	Service event in FSP	HMC Operation Guide
	Service Processor log	System Service Guide 참조
	Root user mail	#mail
	alog	#smitty
	Diag report	#diagrept (Diagnostic Guide 참조)
	smit log	#vi /smit.log
General Check	H/W Resource	#lsdev –Cs scsi #bindprocessor –q #lsattr –El mem0 #lscfg –v
	System Firmware Level	System Service Guide 참조
	Adapter Firmware Level	Web Site
	Device Firmware Level	Web Site
	EC Level	Web Site

12.3. Dump Handling Process

12.3.1. Overview

올바른 Dump 환경설정과 Data Gathering 방법을 알림으로써, Dump 분석 시간을 최소화하여, 시스템의 안정적인 운용을 위한 목적이다

12.3.2. Dump 환경 설정

올바른 Dump 환경설정은 Dump 발생시 Dump 의 원인을 분석하는데 꼭 필요한 Dump 자료를 얻기 위해서 아무리 강조해도 지나치지 않은 매우 중요한 부분이다.

Dump 환경은 #smitty Dump 메뉴를 이용하거나, #sysdumpdev 명령어를 이용하여 설정할 수 있다.

1) Dump 환경 설정

#sysdumpdev -l 명령으로 Dump 환경을 확인하도록 한다.

<#sysdumpdev -l 명령어 실행 예>

```
# sysdumpdev –l
primary                 /dev/hd6
secondary               /dev/sysdumpnull
copy directory          /var/adm/ras
forced copy flag        TRUE
always allow dump       TRUE
dump compression        ON
```

- Primary dump device
  시스템은 1 차로 Primary dump device 인 /dev/hd6 에 Dump 를 받으려고 시도한다.
- Secondary dump device
  실패할 경우는 2 차로 Secondary dump device 인 /dev/sysdumpnull 에 시도한다. (이 경우는 /dev/sysdumpnull 이므로 Dump 를 받지 않고 버리게 된다.)
- Forced copy flag
  Rebooting 되어 올라올 때 Forced copy flag 가 TRUE 이면, Copy Directory 인 /var/adm/ras 에 Dump device 에 받은 Dump 를 파일로 복사하게 된다. 만일 Copy Directory 공간이 부족하여 Dump 를 복사할 수 없다면, 일반적으로 복사하는 것을

포기하고 Rebooting 된다. 하지만 Dump device 가 Paging device 인 /dev/hd6 일 경우는 Tape 으로 Dump 를 복사할 지 유무를 시스템 콘솔을 통해 묻게 되어 있다.

- Always allow dump
  Always allow dump 가 TRUE 이면, 시스템 콘솔에서 Ctrl-Alt-NumPad1 을 눌러서 강제로 Dump 를 Initiate 할 수 있다. 일반적으로 강제 Dump 가 필요하면, Reset Button 을 이용해서 받는 방법을 권장한다.
- Dump compression
  Dump compression 이 ON 이면, Dump 는 압축되어 Dump device 에 저장된다. 압축된 Dump 가 어떤 경우 Corrupt 되어 읽혀지지 않는 경우가 있으므로, 저장공간이 가능하다면 Dump compression 은 OFF 로 설정 하는 것을 권장한다.

2) Dump Data 저장 공간
Primary dump device 는 발생할 Dump 를 받을 수 있을 만큼 충분한 공간을 확보하고 있어야 하며, #lslv Logical Volume 명령으로 Dump device 의 크기를 확인할 수 있다.
아래의 경우 LPs*PP SIZE = 32*16 megabyte(s) = 512 MB 이다.

<#lslv Logical Volume 명령 실행 예>

```
# lslv hd6
LOGICAL VOLUME:     hd6                    VOLUME GROUP:   rootvg
LV IDENTIFIER:      000cf5ad00004c0000000101b6b2a614.2 PERMISSION:
read/write
VG STATE:           active/complete        LV STATE:       opened/syncd
TYPE:               paging                 WRITE VERIFY:   off
MAX LPs:            512                    PP SIZE:        16 megabyte(s)
COPIES:             2                      SCHED POLICY:   parallel
LPs:                32                     PPs:            64
STALE PPs:          0                      BB POLICY:      non-relocatable
INTER-POLICY:       minimum                RELOCATABLE:    yes
INTRA-POLICY:       middle                 UPPER BOUND:    32
MOUNT POINT:        N/A                    LABEL:          None
MIRROR WRITE CONSISTENCY: off
EACH LP COPY ON A SEPARATE PV ?: yes
Serialize IO ?:     NO
```

발생할 Dump 의 크기를 예측하는 명령은 #sysdumpdev -e 이다.

<#sysdumpdev –e 명령어 실행 예>

```
# sysdumpdev –e
0453-041 Estimated dump size in bytes: 19503513
```

위의 경우 19503513 bytes = 약 19 MB 정도이며 따라서 현재 Dump 가 발생한다면 512 MB 크기의 hd6 가 충분히 Dump 를 받을 수 있다.
유의할 점은 #sysdumpdev -e 의 결과는 현재 시점의 예상 Dump 크기이기 때문에, Memory 사용이 많을 때 값을 기준으로 해야 의미가 있다.

참고로, Dump 는 Memory 전체를 복사하는 것이 아니라, Selective copy 이기 때문에 Memory 보다 항상 작으며 아울러 Paging Device 에 있는 내용은 Dump 에 Selective copy 대상이 되지 않는다. 기타 Dump 환경설정은 #smitty dump 메뉴에서 쉽게 하실 수 있다.

12.3.3. Dump 받기 및 확인

Dump 는 Dump 를 받으라는 지시를 누가 내렸느냐에 따라, System initiated dump 와 User initiated dump (강제 Dump)로 나눌 수 있다.

1) Dump 생성 방법
강제 Dump 는 일반적으로 Reset Button 을 눌러서 받는다. 특히 강제 Dump 는 시스템 사용자가 판단하여, Dump 를 받기 때문에, 무슨 작업을 했고, 왜 Dump 를 받게 되었는지 충분한 설명이 반드시 Dump 분석자에게 전달되어야 한다.

시스템 Hang 으로 판단한 경우는, 시스템에 접근 할 수 있는 모든 시도를 다 해보아야 하며 통상 Telnet, FTP, Ping, 콘솔을 통한 로그인 등을 시도해보고 시도한 결과를 기록한 후에 강제 Dump 를 시도한다.

보통 #sysdumpstart 명령이나, #smit Menu 를 통해서 강제 Dump 를 받는 경우가 있는데, 이는 권장사항이 아니다.
강제 Dump 는 Reset 버튼을 눌러서 받기를 권장한다.

2) Dump Status Code

일단 Dump 가 시작되면, OP panel 의 LED 를 통해 볼 수 있는 Dump Status code 를 통해 Dump 의 진행상태를 확인 할 수 있으며, 아래에 몇 가지 Dump Status code 를 참조하기 바란다.

<Dump Status Code 예>

```
0c0   Dump completed successfully
0c2   User-initiated dump in progress.
0c6   User-initiated dump in progress to secondary dump device.
0c9   system initiated dump in progress, dump progress by # of bytes
0cc   Dump process switched to secondary dump device.
```

- 정상적인 경우, System initiated dump 는 0c9 -> 0c0 으로, User initiated dump 는 0c2 -> 0c0 으로 LED 가 변하게 된다.
- 0c0 이 나오기 전에 Dump 받는 것을 중단하면, Partial dump 가 생성된다. 이 경우는 Dump 분석이 안 되는 경우가 많으니 주의해야 한다.
- Dump Status code 0c0 은 Dump 가 Dump Device 에 성공적으로 생성되었음을 의미한다.
- Dump 가 생성되고, 시스템이 Rebooting 되면, #sysdumpdev -L 명령으로 최근에 생성된 Dump 정보를 확인할 수 있다.
- 아래의 예는 #dumplv 라는 Dump device 에 압축된 Dump 가 239 MB 정도 받혔고, 압축을 풀면 1.6 GB 정도 된다.
- Dump 발생시점은 Mon Jan 17 19:37:29 2005 이고, 성공적으로 생성되었다.

3) Dump 생성 정보 확인

- Dump 가 성공적으로 생성되었는지의 최종 확인은 #kdb 명령으로 직접 Dump 가 읽혀지는지 확인해 보는 것으로 가능하며, #kdb /dev/dumplv 실행결과 (0)> 와 유사한 형태의 kdb 프롬프트를 받으면 Dump 가성공적으로 생성된 것을 의미한다.
- 하지만 위와 같이 Dump 가 압축된 경우는 #snap -Dd/Directory/ 명령으로 Dump File 을 내려서 #kdb Dump File 명령으로 확인해야 한다.

```
# sysdumpdev –L
Device name:              /dev/dumplv
Major device number: 10
Minor device number: 10
Size:                     239520256 bytes
Uncompressed Size:    1600872622 bytes
Date/Time:               Mon Jan 17 19:37:29 2005
Dump status:             0
dump completed successfully
```

12.3.4. Dump 자료 수집 및 분석의뢰

Dump 를 포함한 분석에 필요한 자료는 #snap -acd /Directory 명령으로 받는 것이 좋다. 그러나 외장 디스크가 많아서 LVM 정보를 수집하는 데 너무 많은 시간이 소요된다면, #snap -gfkDcd /Directory 명령을 사용해도 무방하다.

1) Dump Data 수집 방법

- #snap -acd /Directory 명령을 사용하면 /Directory 밑에 분석에 필요한 Dump 와 시스템 환경 정보가 모두 수집된다. Dump Directory 밑에 UNIX Kernel 과 Dump File 이 들어 있는지 반드시 확인해야 한다.

- Dump File 을 Uncompress 해서 #kdb Dump File 명령으로 (0)> 와 유사한 형태의 kdb 프롬프트가 뜨는지 확인한다.
- 모두 문제가 없다면 /Directory 밑에 Dump 와 다른 Snap 정보를 압축한 snap.pax.Z File 이 생성되며, 이 snap.pax.Z File 을 Second Level Supporter 에게 전달한다.
- Second Level Supporter 에게 자료를 전달하는 방식은 아래 방법 중 한 가지 방법을 이용한다.

2) Dump Data 분석 의뢰 방법

① FTP 전송 방법

ftp aix.kr.ibm.com
id : aix
pw : aix
위의 방법으로 로그인하면 /home/ftp/pub/dump/ Directory 에 접속이 되며,
여기에 원하는 이름으로 Directory 를 생성하고, snap.pax.Z 파일을 올려놓는다.

② Tape 으로 전달하는 방법

#pax 명령을 이용해서 snap.pax.Z 파일을 Tape 에 넣으면 되며, Tape 에 넣을 때는 절대경로로 묶지 말고, 반드시 상대경로로 묶도록 한다.
Tape 을 본사 13 층 pSeries S/W 기술지원센터로 보내주시기 바랍니다.
(☎ 1588-5801)

③ Second Level Supporter 에게 직접 전달하는 방법

snap.pax.Z File 을 xxxxx.yyy.zzz.pax.Z (xxxxx.yyy.zzz 는 PMR 번호)로 이름을 바꾸어 아래 testcase Server 에 FTP 로 전송한다.
ftp testcase.software.ibm.com
id : anonymous
pw : mail_id@mail_address
cd aix/toibm/
이 Directory 로 자료를 전송한다.

① 마지막으로, 아래 덤프분석의뢰양식을 작성해서 메일로 aixcc@kr.ibm.com 에게 보낸다.

<덤프분석의뢰양식 작성 예>

Dump Analysis Request Form
.
Entitlement
Machine Type & Serial : 7040 2255555
Customer Name : 한국은행
고객담당자
- Name : 김철수 과장,
- Phone 011-222-3333,
- E-Mail cskim@koreabank.com

기술대표
- Name : 홍길동 (삼주)
- Phone 011-898-7777,
- E-Mail gdhong@kr.ibm.com

.
Environment
OS : AIX 4.3.3 ML11
Software : HACMP 4.4.1, Oracle 8.1.7

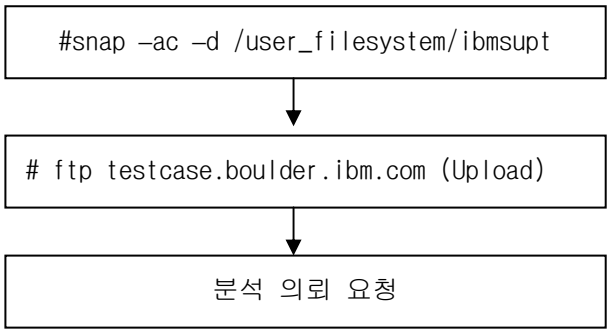
Hardware : ESS disk, LPAR
.
Testcase : aix.kr.ibm.com: ./k-bank/snap.pax.Z
.
Problem Description
I ran rm command. It didn't give me any prompt.
Tried to open new telnet session from my Windows PC to this host.
It showed only "Trying..." message but didn't show any login prompt.
Ping to this host from my Windows PC was okay.
Went to the system console.
Pressed Enter key several times to get login prompt but showed nothing but black screen.
Pressed reset button on the system to get forced dump.
The system showed 888-102-000-0C0 on LED and rebooted automatically.
System works okay at this moment.
Gathered "snap -a" and uploaded this testcase.
Would you analyze this dump why the system was hung?

12.3.5. 주의 사항

① Dump Reason Code 가 207 의 경우는 Dump File 이 필요 없다. Error Report 에 있는 CHRP error log 를 분석하면 된다.
② 자료를 #pax 로 묶을 때 꼭 상대패스로 묶도록 한다.
③ Dump 환경 설정 시 가능하다면, Compression option 은 NO 로 한다.
④ 강제 Dump 의 경우는 반드시 받기 직전 문제상황을 상세히 기술하도록 한다.
⑤ Dump 자료를 보내기 전에 Dump File 과 UNIX File 이 들어있는지 꼭 확인한다.
⑥ Dump File 외에 Snap 자료를 함께 보내도록 한다..

12.4. SNAP

시스템 운영도중 문제발생에 대한 분석을 위해 기본적으로 필요한 데이터를 수집하기 위해 Snap 을 받도록 한다.
유형별 문제에 따라 Snap 은 여러 가지 Flag 로 받을 수 있다.
사용법은 #man page 를 참고하기 바라며, 아래는 일반적으로 모든 Snap 정보(Dump 포함)를 받아 분석의뢰까지 진행하는 일반적인 방법이다.



※ pax 로 묶인 Snap 정보를 Uncompress 하는 방법: #pax –rvf filename.pax.Z
TIP 1> S/W 문제해결을 위해 기본적으로 필요한 데이터
1. 일반적인 이슈 : snap -gc
2. TCIPIP : snap -gtc
3. LVM : snap -gfLc
4. HACMP : snap -ec
5. DUMP : snap -ac
6. GPFS : gpfs.snap (예외적으로 /tmp/gpfs.snapOut/all.xxxx.tar 형태로 생성됨)

TIP 2> H/W 문제해결을 위해 기본적으로 필요한 데이터
1. snap -gc
2. HMC iqyylog(or pedbg) - IBM 엔지니어가 gathering
3. HMC FSPdump - IBM 엔지니어가 gathering

12.5. TRACE

Trace System 은 시스템 활동이나 이벤트들의 연속적인 플로우를 캡처하기 때문에 특정 시점의 문제를 디버킹 하는데 아주 좋은 Tool 이다.

Trace Facility 는 대량의 데이터를 모으기 때문에 분석을 위해 필요한 Hook ID 를 미리 지정함으로써 원하는 데이터를 모을 수도 있다. 이를 위해서는 문제에 적합하게 적용할 요구되는 Hook ID 를 선택해야 하는 어려움이 있는 것이 사실이다. 정확한 Hook ID 를 위해서는 Second Level Supporter 의 도움을 받아, Trace 정보를 받는 것이 좋다.

1) Trace Data 수집
Trace 수집은 Buffer 사용방법에 따라 달라지며 아래 3 가지 방법이 있다.
① Alternate Buffer mode (Default)
Buffer 에 쓰여진 모든 Event 들이 Trace Log File 에 쓰여짐
② Circular Buffer mode (-l)
Trace event 들이 Buffer 에 계속 쌓이다가 문제가 발생한 시점에 Trace 를 멈추게 되는 시점에 Buffer 에 있는 것이 Log File 에 쓰여진다.
③ Single Buffer mode (-f)
하나의 Buffer 를 가지고 이 Trace Buffer 가 차게 되면 Event Trace 는 멈추게 된다. 즉, Buffer 가 Full 이 되거나 Trace 가 Stop 될 때 까지는 어떠한 것도 Trace Log File 에 기록되지 않는다. AIX 5.1 부터 최대 Trace Buffer 크기는 536,870,368 Bytes 이다.

2) Circular 방식의 Trace Data 수집 예
Trace 를 받는 방법에는 위와 같이 여러 가지 방법이 있다.
예를 들어 Circular 방식의 Background Mode 로 아래와 같이 Trace 를 걸 수 있다.
trace –a –l –L16000000 –T8000000 –o trace.raw
위와 같은 방법으로 받은 Trace 를 Second Level Supporter 에게 분석요청을 하도록 한다.

3) 참고 자료
Trace 사용방법 및 분석에 대한 자세한 내용은 Redbook : Problem Solving and Troubleshooting in AIX5L,SG24-5496-01 의 Chapter 11 Event tracing 부분을 참고한다.

12.6. alog

alog 는 시스템에 문제가 발생하거나 boot process 를 trace 해야 할 경우 사용한다.

1) alog flag
-f LogFile : 로그 파일의 이름을 지정한다. 만약 파일이 없으면 하나를 만든다. 로그 파일이 쓰기 가능하지 않을 경우, /dev/null 에다가 쓴다
-L : alog db 에 기 정의된 로그 타입의 리스트를 보여 준다. -L flag 를 -t flag 와 함께 쓰면, 로그 타입의 속성과 함께 나열한다.
-o : 로그 파일의 내용 보기
-q : stdin 을 로그 파일에 복사한다. 쓰지 못하면 stdout 으로 보낸다
-t logtype : alog db 에 정의된 로그를 확인한다.

2) alog -L 로 리스트 보기
```
[root@ifis:/]alog -L
  boot
  bosinst
  nim
  console
  dumpsymp
```

3) alog -o -t LogType 로 로그 내용 보기
```
[root@ifis:/]alog -o -t dumpsymp
```

13. AIX Parameter

13.1. CPU & Memory Parameter

```
[khan1:/] vmo -a
         cpu_scale_memp = 8
 data_stagger_interval = 161
                 defps = 1
   force_relalias_lite = 0
             framesets = 2
             htabscale = n/a
     kernel_heap_psize = 4096
          kernel_psize = 16777216
  large_page_heap_size = 0
          lgpg_regions = 0
             lgpg_size = 0
       low_ps_handling = 1
       lru_file_repage = 1
     lru_poll_interval = 10
             lrubucket = 131072
          maxclient% = 30
             maxfree = 1088
             maxperm = 2848458
          maxperm% = 30
            maxpin = 8045802
            maxpin% = 80
       mbuf_heap_psize = 65536
       memory_affinity = 1
         memory_frames = 9961472
         memplace_data = 2
  memplace_mapped_file = 2
memplace_shm_anonymous = 2
    memplace_shm_named = 2
        memplace_stack = 2
         memplace_text = 2
memplace_unmapped_file = 2
                mempools = 4
                 minfree = 960
                 minperm = 1424228
                minperm% = 15
              nokilluid = 0
                 npskill = 32768
               npsrpgmax = 262144
               npsrpgmin = 196608
            npsscrubmax = 262144
            npsscrubmin = 196608
                 npswarn = 131072
          num_spec_dataseg = 0
                numpsblks = 4194304
      page_steal_method = 0
           pagecoloring = n/a
        pinnable_frames = 8899564
pta_balance_threshold = n/a
   relalias_percentage = 0
               rpgclean = 0
            rpgcontrol = 2
                 scrub = 0
            scrubclean = 0
soft_min_lgpgs_vmpool = 0
       spec_dataseg_int = 512
       strict_maxclient = 1
         strict_maxperm = 0
              v_pinshm = 0
 vm_modlist_threshold = -1
       vmm_fork_policy = 1
```

- ✓ maxperm%: 파일캐쉬로 사용할 수 있는 최대 메모리의 백분율
- ✓ minperm%: 파일 캐쉬로 사용할 수 있는 최소 메모리의 백분율
- ✓ maxclient%: jfs2, nfs등이 사용할 수 있는 최대 메모리의 백분율
- ✓ minfree: free real memory를 유지하기 위한 최소값. VMM(virtual memory mansge)의 page stealer는 minfree값 이하로 free list에 있는 page 수가 감소하게 되면 수행된다.
- ✓ maxfree: free real memory를 유지하기 위한 최대값. VMM(virtual memory mansge)의 page stealer는 maxfree값 이상으로 free list에 있는 page 수가 증가하면 정지한다.
- ✓ strict_maxperm: maxperm% 값 이상의 메모리 사용에 대한 lock(값1) and unlock(값0)

권고 값

<Memory Parameter>

minperm%=3

maxperm%=90

maxclient%=90

page_steal_method=1

lru_file_repage=0

<I/O Parameter>

j2_nBufferPerPagerDevice=1024

pv_min_pbuf=1024

numfsbufs=2048

lvm_bufcnt=12

13.2. Network Parameter

```
[khan1:/] no -a
                 arpqsize = 12
                arpt_killc = 20
               arptab_bsiz = 7
                arptab_nb = 149
                 bcastping = 0
       clean_partial_conns = 0
                 delayack = 0
              delayackports = {}
          dgd_packets_lost = 3
             dgd_ping_time = 5
            dgd_retry_time = 5
         directed_broadcast = 0
           extendednetstats = 0
                  fasttimo = 200
         icmp6_errmsg_rate = 10
           icmpaddressmask = 0
ie5_old_multicast_mapping = 0
                   ifsize = 256
          inet_stack_size = 16
               ip6_defttl = 64
                ip6_prune = 1
            ip6forwarding = 0
        ip6srcrouteforward = 1
        ip_ifdelete_notify = 0
                  ip_nfrag = 200
             ipforwarding = 0
                ipfragttl = 2
         ipignoreredirects = 0
                ipqmaxlen = 512
           ipsendredirects = 1
         ipsrcrouteforward = 1
            ipsrcrouterecv = 1
            ipsrcroutesend = 1
           llsleep_timeout = 3
                  lo_perf = 1
               lowthresh = 90
                main_if6 = 0
               main_site6 = 0
                 maxnip6q = 20
                   maxttl = 255
                medthresh = 95
              mpr_policy = 1
             multi_homed = 1
               nbc_limit = 16934502
           nbc_max_cache = 131072
           nbc_min_cache = 1
        nbc_ofile_hashsz = 12841
               nbc_pseg = 0
         nbc_pseg_limit = 1992294
         ndd_event_name = {all}
      ndd_event_tracing = 0
          ndp_mmaxtries = 3
          ndp_umaxtries = 3
               ndpqsize = 50
              ndpt_down = 3
              ndpt_keep = 120
             ndpt_probe = 5
         ndpt_reachable = 30
           ndpt_retrans = 1
          net_buf_size = {all}
          net_buf_type = {all}
       net_malloc_police = 0
          nonlocsrcroute = 1
                nstrpush = 8
             passive_dgd = 0
       pmtu_default_age = 10
             pmtu_expire = 10
pmtu_rediscover_interval = 30
           psebufcalls = 20
              psecache = 1
          pseintrstack = 24576
             psetimers = 20
        rfc1122addrchk = 0
               rfc1323 = 1
               rfc2414 = 1
          route_expire = 1
       routerevalidate = 1
              rto_high = 64
            rto_length = 13
             rto_limit = 7
               rto_low = 1
                  sack = 0
                sb_max = 1310720
send_file_duration = 300
            site6_index = 0
           sockthresh = 85
              sodebug = 0
          sodebug_env = 0
            somaxconn = 1024
             strctlsz = 1024
             strmsgsz = 0
            strthresh = 85
            strturncnt = 15
         subnetsarelocal = 1
      tcp_bad_port_limit = 0
                 tcp_ecn = 0
      tcp_ephemeral_high = 65535
       tcp_ephemeral_low = 32768
            tcp_finwait2 = 1200
         tcp_icmpsecure = 0
       tcp_init_window = 0
tcp_inpcb_hashtab_siz = 24499
            tcp_keepcnt = 8
           tcp_keepidle = 14400
           tcp_keepinit = 150
          tcp_keepintvl = 150
   tcp_limited_transmit = 1
            tcp_low_rto = 0
           tcp_maxburst = 0
            tcp_mssdflt = 1460
       tcp_nagle_limit = 65535
     tcp_nagleoverride = 0
            tcp_ndebug = 100
          tcp_newreno = 1
       tcp_nodelayack = 0
     tcp_pmtu_discover = 1
        tcp_recvspace = 65536
        tcp_sendspace = 65536
        tcp_tcpsecure = 0
         tcp_timewait = 1
              tcp_ttl = 60
      tcprexmtthresh = 3
             thewall = 19922944
      timer_wheel_tick = 0
    udp_bad_port_limit = 0
    udp_ephemeral_high = 65535
     udp_ephemeral_low = 32768
udp_inpcb_hashtab_siz = 24499
     udp_pmtu_discover = 1
         udp_recvspace = 655360
         udp_sendspace = 65536
               udp_ttl = 30
            udpcksum = 1
               use_isno = 1
        use_sndbufpool = 1
```

- ✓ thewall: Network에서 mbuf 관리를 위해 할당된 메모리로 AIX 4.3.2 이후에는 default로 실 메모리의 1/2이나 1048576(1GB)로 설정되는 runtime 값
- ✓ sb_max: 소켓에 허용된 최대 버퍼사이즈 이며, sb_max는 최소한 tcp_sendspace, tcp_recvspace 또는 udp_recvspace 값 중 제일 큰 값의 두 배 이상 이여야 하며, buffer 효율성이 50% 이상이면 tcp 및 udp byte limit까지 사용할 수 있게 된다.
- ✓ tcp_sendspace : 데이터 송신을 위한 시스템의 기본 소켓 버퍼 크기. 반드시 sb_max값 이하로 지정해야 한다.

- ✓ udp_sendspace: UDP 데이터 송신을 위한 시스템의 기본 소켓 버퍼 크기 ifconfig 명령으로 인터페이스마다 설정 가능하다. 반드시 sb_max값 이하로 지정해야 한다.
- ✓ udp_recvspace: UDP 데이터 수신을 위한 시스템의 기본 소켓 버퍼 크기 ifconfig 명령으로 인터페이스 마다 설정 가능하다. 반드시 sb_max값 이하로 지정해야 한다. (권장 값:262144)
- ✓ rfc1323: 고성능 TCP확장에 의해 지정된 TCP 성능 향상을 가능하게 함. '1'인 경우 TCP window size를 최대 32bit로 사용할 수 있어 tcp_sendspace와 tcp_recvspace 값을 64KB 이상으로 설정 할 수 있다. 그렇지 않은 경우 값을 증가 시켜도 아무런 의미가 없다.
- ✓ tcp_nodelayack: 특정 TCP패킷에 대한 ACK를 지연하고 대신 다음에 전송되는 패킷과 함께 보낸다. 성능은 좋아지나 시스템 오버헤드가 발생한다. (값1:On, 값0:Off)
- ✓ ipqmaxlen: IP input queue의 길이를 관리하는 값
- ✓ use_isno: 사용자가 변경한 인터페이스 설정 값을 우선 사용하도록 지정. (값1:On, 값0:Off)

권고 값

< AIX 4.3.3~ >

tcp_nodelayack=1 (Connection type, rebooting 불필요)

ipignoreredirects=1 (Dynamic type, rebooting 불필요)

tcp_nagle_limit=0 (Dynamic type, rebooting 불필요)

rfc1323=1

sb_max=1310720

tcp_sendspace=262144

tcp_recvspace=262144

udp_sendspace=65536

udp_recvspace=655360

tcp_pmtu_discover=0

udp_pmtu_discover=0

ipqmaxlen=200

< AIX 5.2~, 해당 network option이 no -a 에서 보일 때에만 해당 >

lo_perf=0

< AIX 5.3 TL05~ >

timer_wheel_tick=10 (Rebooting 필요)

tcp_low_rto=200 (timer_wheel_tick을 변경한 이후, rebooting하여야만 적용됨)

단, AIX 5.3 TL06 Base에서 tcp_low_rto가 100까지밖에 정의되지 않는 bug가 있으므로, 만일 해당 버전에서 200으로 설정할 때에 error가 발생하면, 100으로 입력해 주시면 됩니다. TL06 SP1부터는 fix가 적용되어 있으므로, 정상적으로 200으로 설정할 수 있습니다.

참고: AIX 6.1 최적화와 성능 튜닝 (한글)

http://www.ibm.com/developerworks/kr/library/au-aix6tuning/index.html#N100FA

14. HMC (Hardware Management Console)

14.1. HMC 개요

HMC (Hardware Management Console)는 IBM System p5 를 관리하기 위한 콘솔로서,
IBM Linux 상에서 HMC 솔루션이 운영됩니다.

HMC 를 통해 다음 과 같은 업무를 수행할 수 있습니다.

- LPAR (Logical Partition) 생성 및 변경 관리
- CoD (Capacity on Demand) 활성화
- System Firmware Update
- ASMI 를 통한 서비스 프로세서 메뉴 접속
- Service Agent 를 통한 원격 감시 기능

14.2. LPAR & DLPAR를 통한 자원 재분배

14.2.1. LPAR 개요

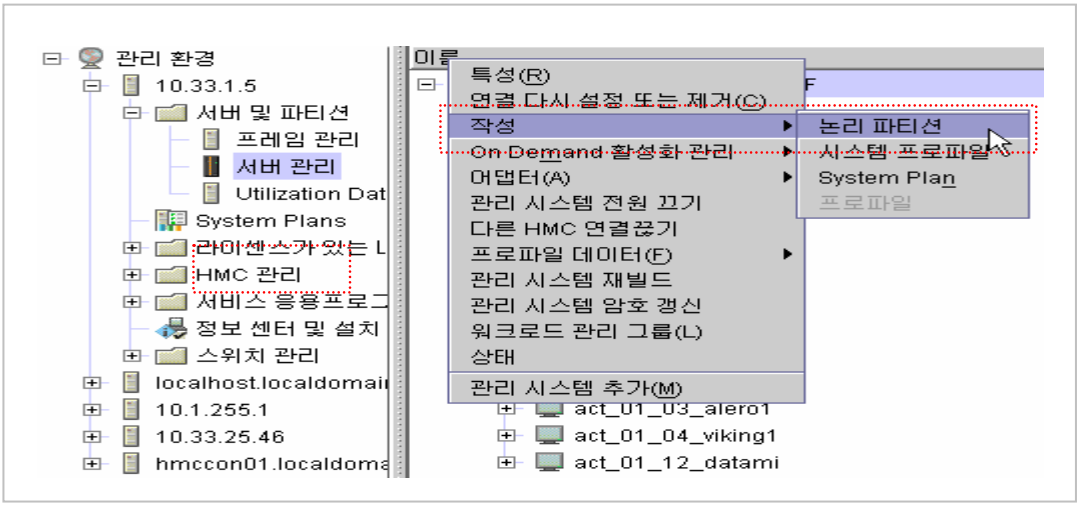
LPAR (Logical Partition) 는 하나의 물리적인 서버를 논리적으로 나누어서 쓰는 것을 의미합니다.
즉 물리적인 서버를 논리적인 partition 으로 나누고 각 논리적 partition 에 CPU 와 Memory 등의 리소스를 나누어 할당해 줌으로써, 하나의 물리적인 서버 위에 여러 대의 논리적 서버를 구성하여 다양한 업무를 독립적으로 운영할 수 있게 해줍니다.
LPAR 를 통해 다음과 같은 업무를 수행할 수 있습니다.

- CPU 추가 및 제거 (동적 재분배 가능)
- Memory 추가 및 제거 (동적 재분배 가능)
- Adapter 추가 및 제거 (동적 재분배 가능)
- 공유 I/O 장치 추가 및 제거 (동적 재분배 가능)

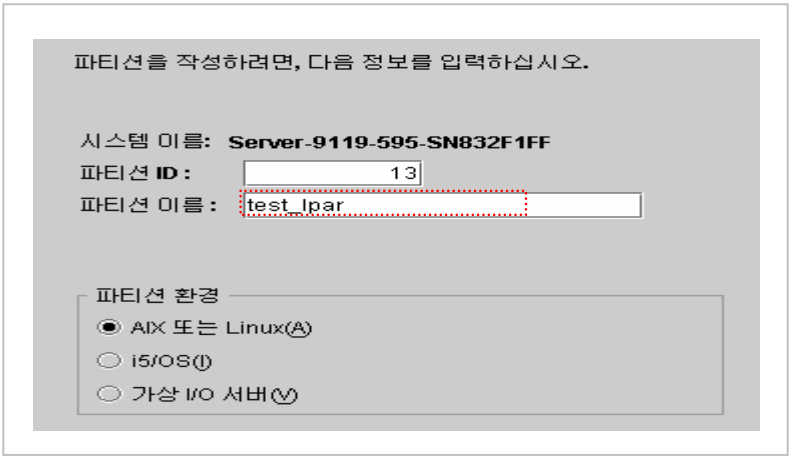
14.2.2. LPAR 구성

- LPAR 구성 화면 선택

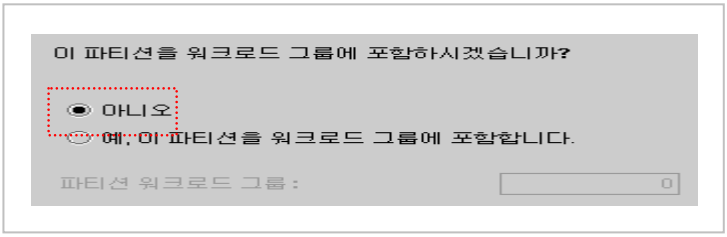
► 서버관리 클릭 후 시스템에 마우스를 가져다 놓고 오른쪽 마우스를 클릭한다.
► 아래 화면처럼 “작성” 에서 “논리 파티션”을 선택한다.



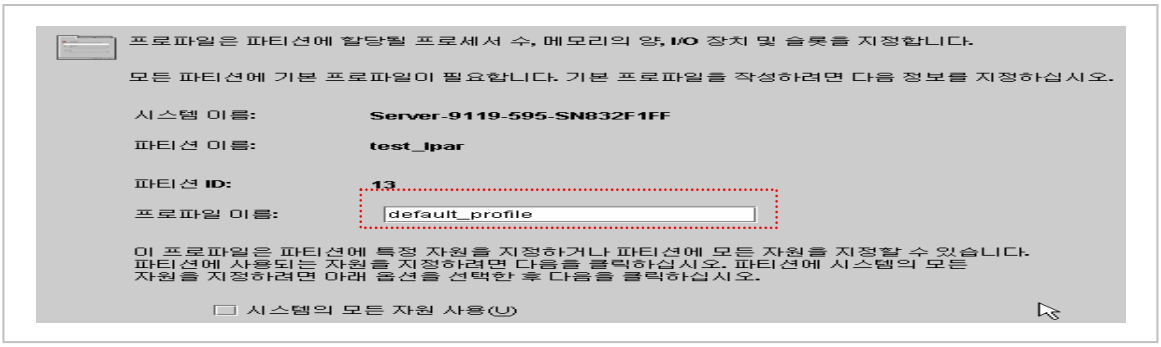
▶ 파티션 이름을 작성한다.



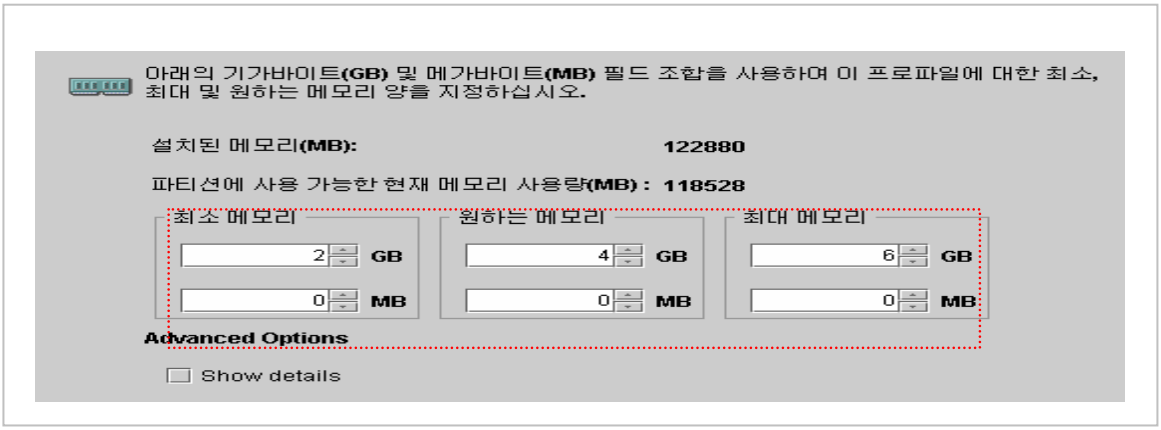
▶ 아니오 를 선택한다.



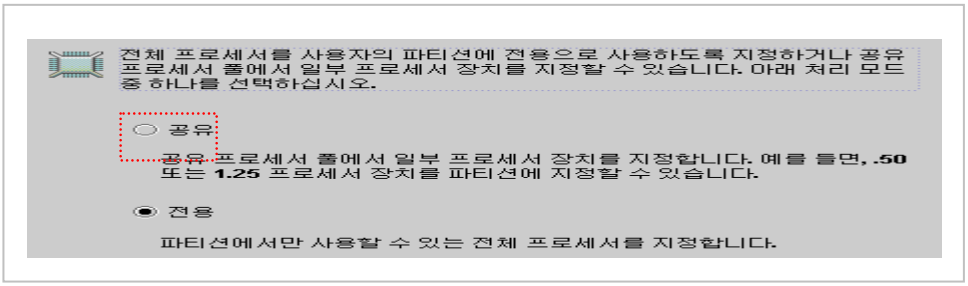
▶ 프로파일 이름을 작성한다.



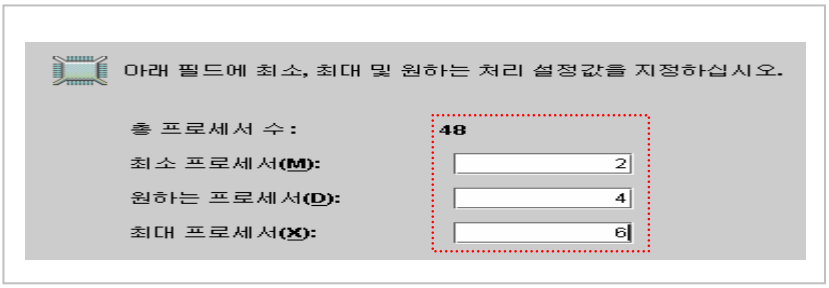
► 사용할 Memory 를 지정한다. (동적 Memory 재구성을 위한 최소, 최대 Memory 를 지정한다.)



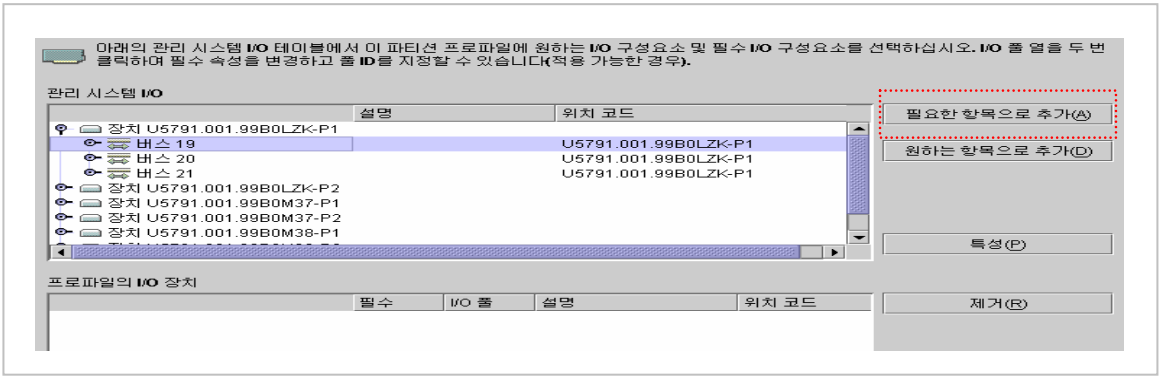
► CPU 사용 방법 선택



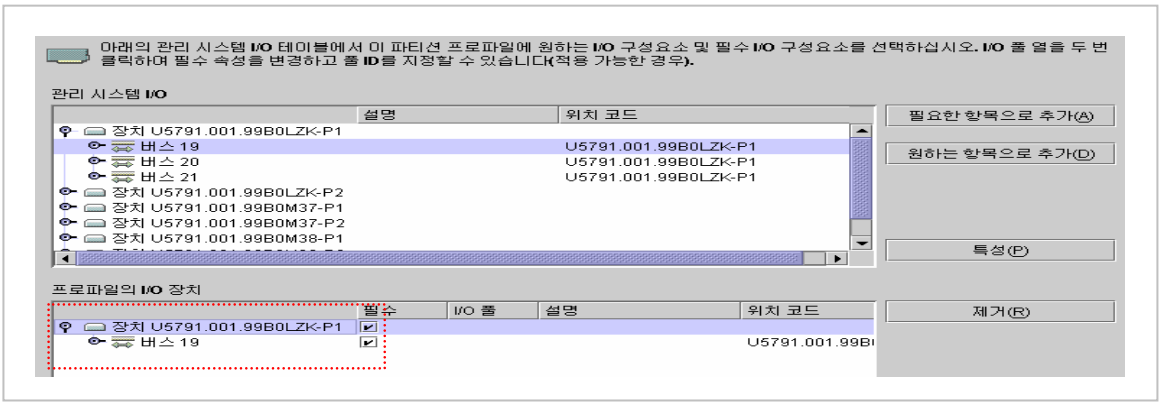
► 사용할 CPU 를 지정한다. (동적 CPU 재구성을 위한 최소, 최대 CPU 를 지정한다.)



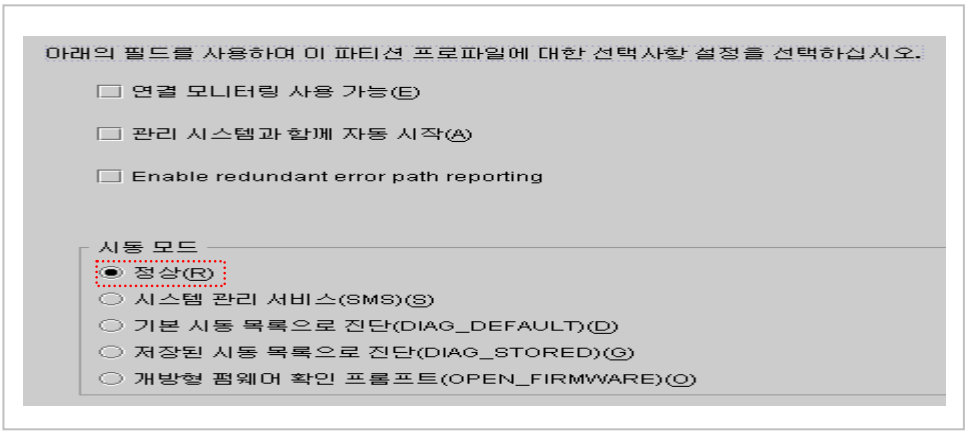
► 사용할 I/O 를 장치들을 지정한다.



▶ 사용할 I/O 를 장치를 “필요한 항목으로 추가”하면 프로파일의 I/O 장치에 추가가 된다.



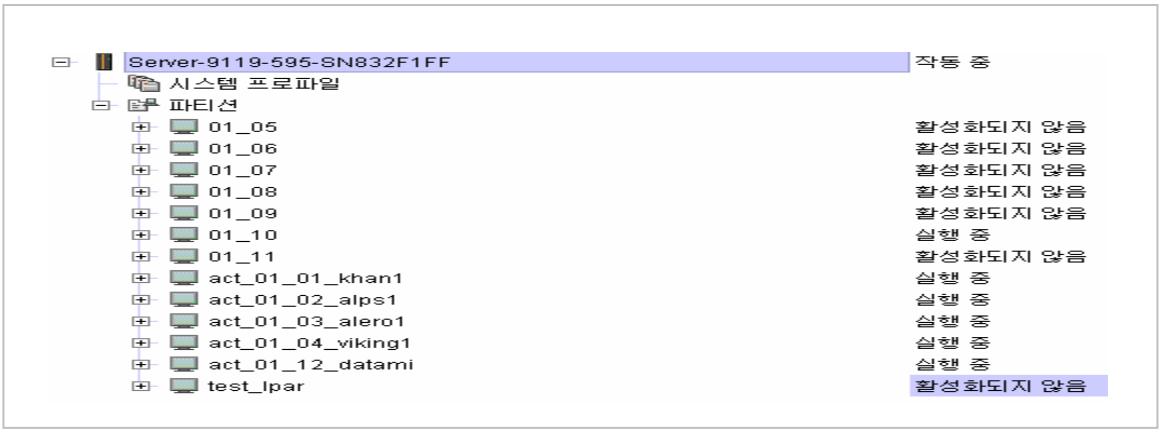
▶ 시동 모드를 선택 한다.



▶ LPAR 구성 요약 확인

▶ LPAR 구성 요약 확인 후 “확인”을 선택하면 LPAR 가 생성된다.

▶ 새로 생성된 LPAR 확인

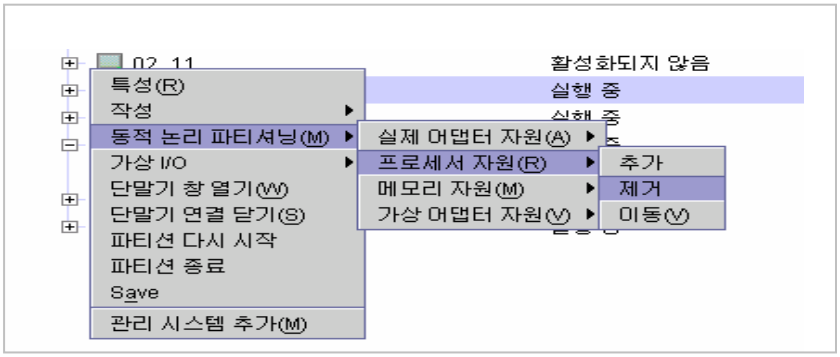


14.2.3. CPU DLAR (CPU 자원 변경)

- HMC 통해 해당 파티션의 Processor 를 제거

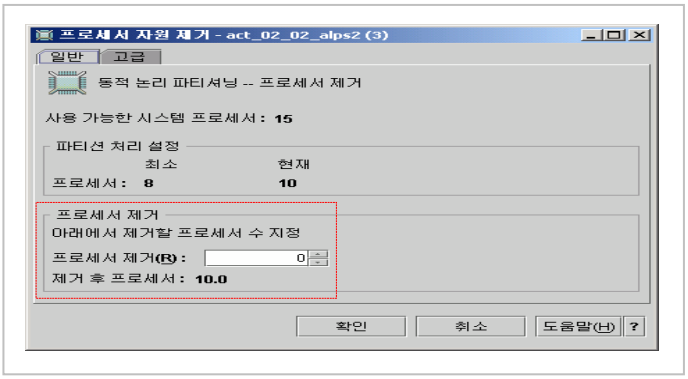
1) 제거 메뉴 선택

► HMC Console 전체 화면 → 해당 파티션을 선택하고 오른쪽 마우스 클릭 → 동적 논리적 파티셔닝 → 프로세서 자원 → 제거 선택



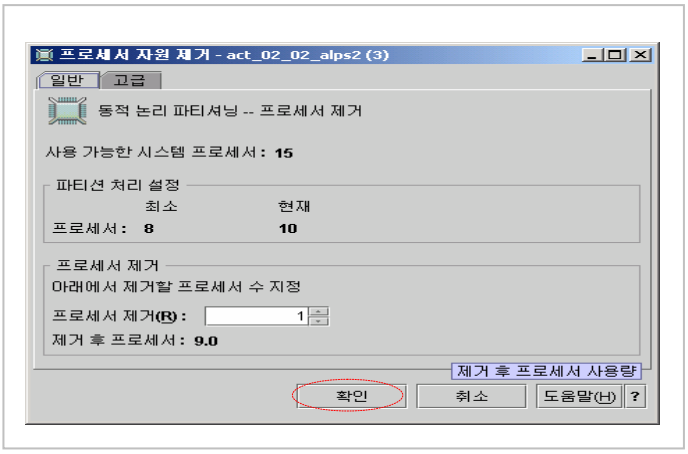
2) 제거할 Processor 개수 지정

► Processor 자원 제거 화면



3) Resource 제거

► 확인 버튼 클릭 후 Resource 제거

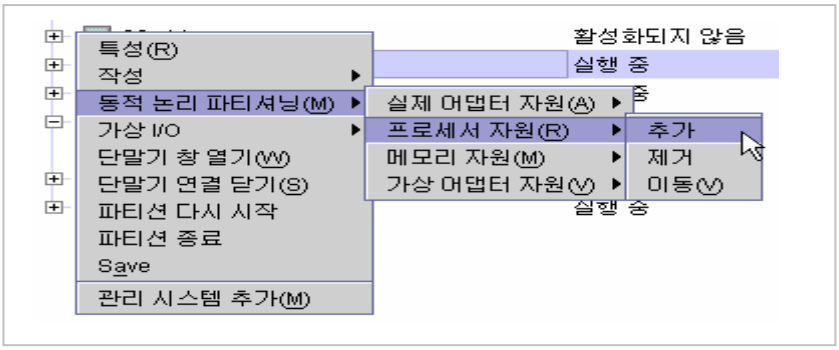


► Resource 를 제거시키면 진행 화면이 나오고 완료되면 사라짐.

- HMC 통해 해당 파티션의 Processor 를 추가

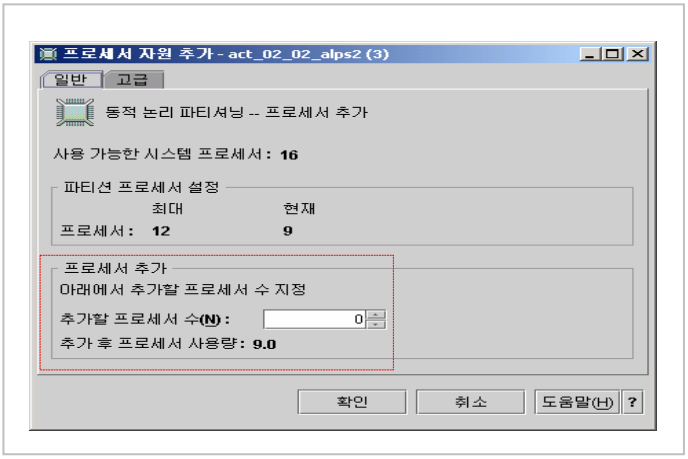
1) 추가 메뉴 선택

► HMC Console 전체 화면 → 해당 파티션을 선택하고 오른쪽 마우스 클릭 → 동적 논리적 파티셔닝 → 프로세서 자원 → 추가 선택



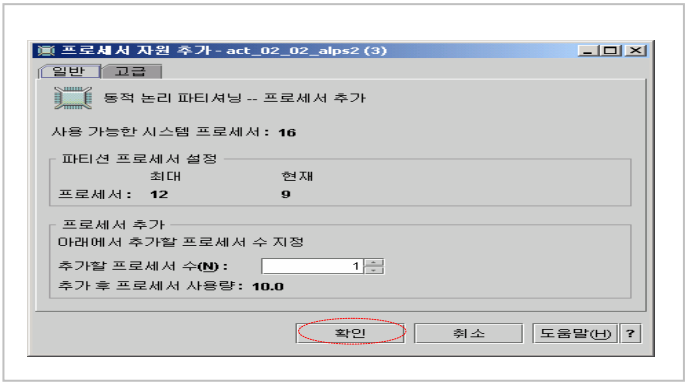
2) 추가할 Processor 개수 지정

► Processor 자원 추가 화면



3) Resource 추가

► 확인 버튼 클릭 후 Resource 추가

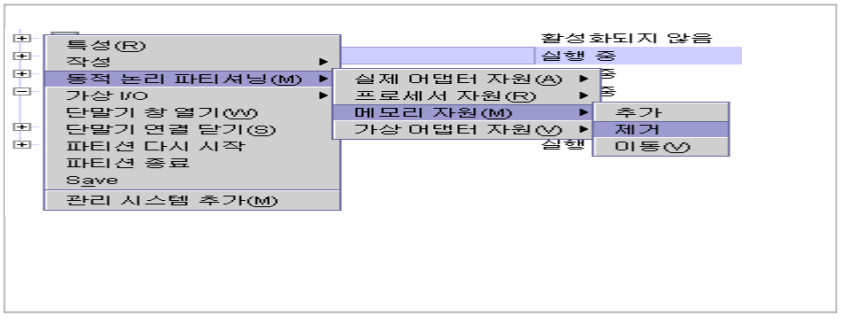


14.2.4. Memory DLPAR (Memory 자원 변경)

- HMC 통해 해당 파티션의 Memory 를 제거

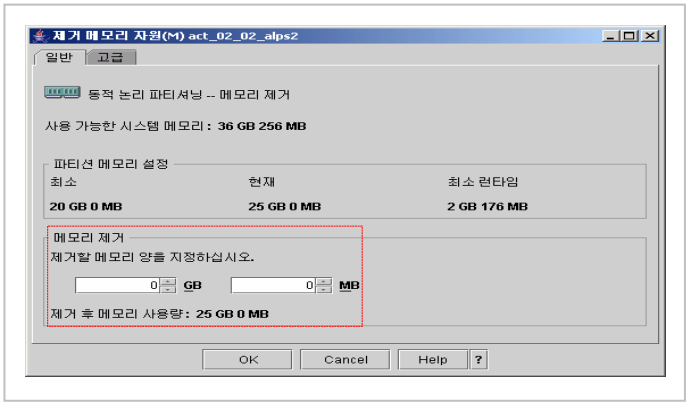
1) 제거 메뉴 선택

► HMC Console 전체 화면 → 해당 파티션을 선택하고 오른쪽 마우스 클릭 → 동적 논리적 파티셔닝 → 프로세서 자원 → 제거 선택



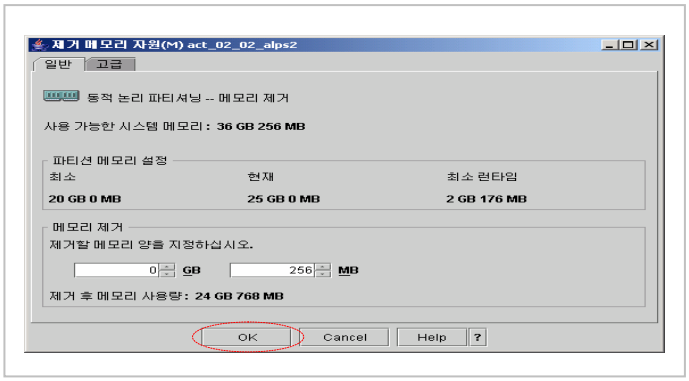
2) 제거할 Memory size 지정

► Memory 자원 제거 화면



3) Resource 제거

► 확인 버튼 클릭 후 Resource 제거

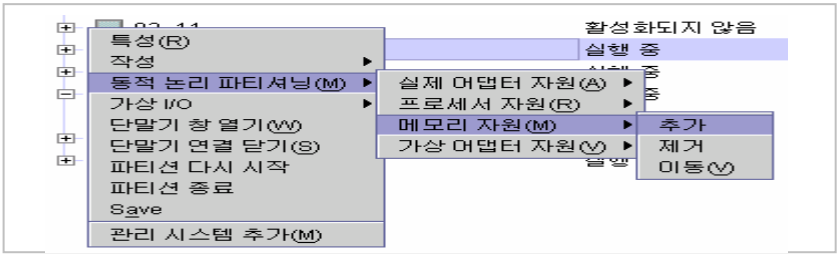


► Resource 를 제거시키면 다음과 같은 진행 화면이 나오고 완료되면 사라짐.

- HMC 통해 해당 파티션의 Memory 를 추가

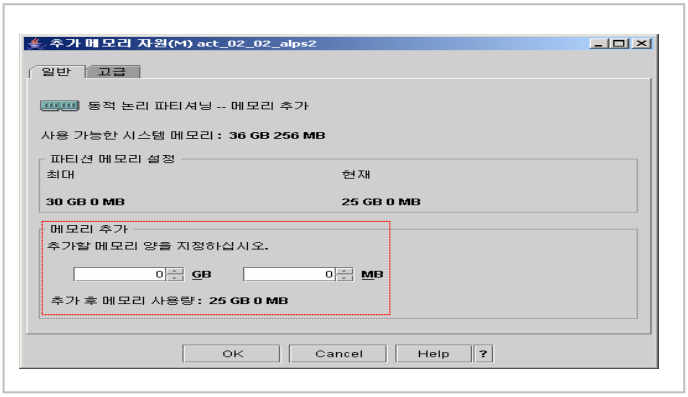
1) 추가 메뉴 선택

► HMC Console 전체 화면 → 해당 파티션을 선택하고 오른쪽 마우스 클릭 → 동적 논리적 파티셔닝 → 프로세서 자원 → 추가 선택



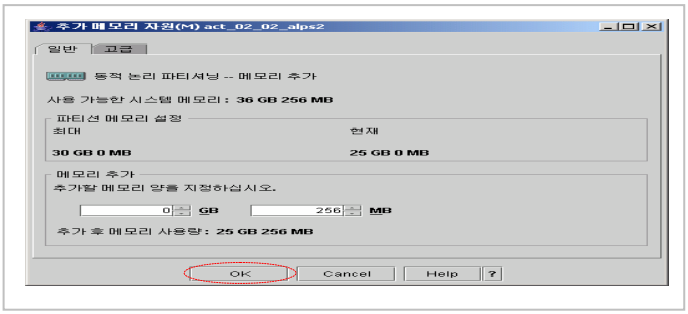
2) 추가할 Memory size 지정

► Memory 자원 추가 화면



3) Resource 추가

► 확인 버튼 클릭 후 Resource 추가

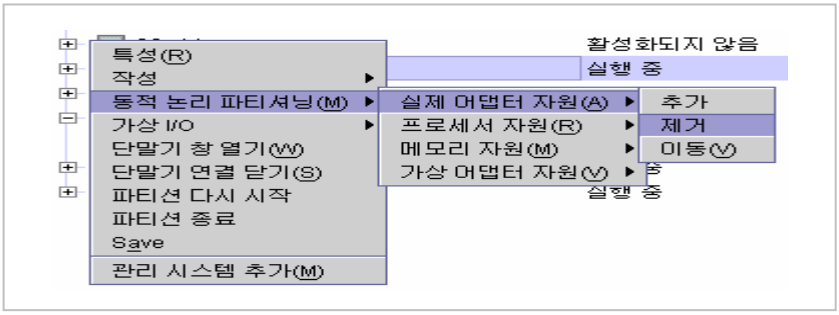


14.2.5. Adapter DLPAR (Adapter 자원 변경)

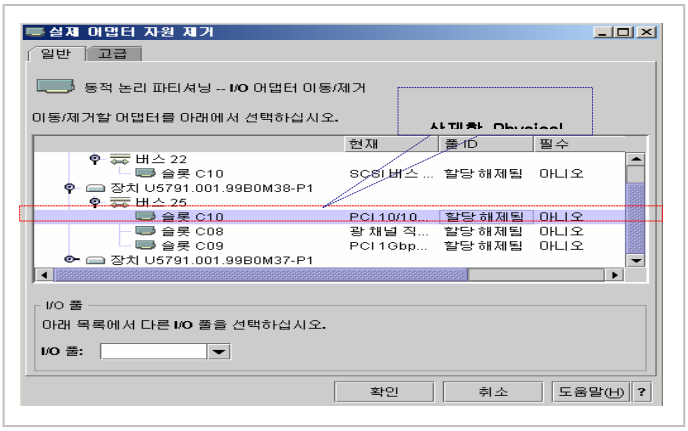
- HMC 통해 해당 파티션의 Adapter 를 제거

1) 제거 메뉴 선택

► HMC Console 전체 화면 → 해당 파티션을 선택하고 오른쪽 마우스 클릭 → 동적 논리적 파티셔닝 → 프로세서 자원 → 제거 선택



2) 제거할 Adapter 지정



3) Resource 제거

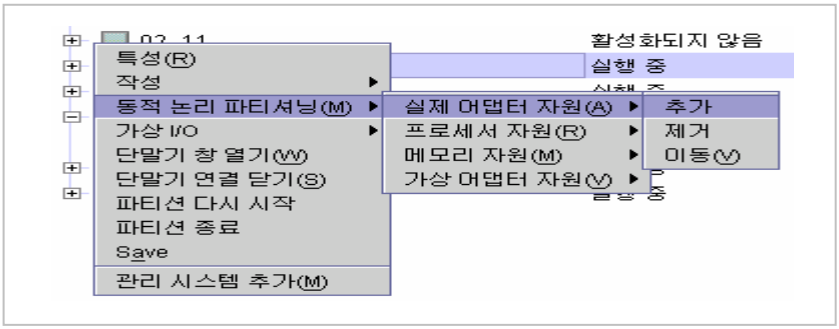
► 확인 버튼 클릭 후 Resource 제거

► Resource 를 제거시키면 다음과 같은 진행 화면이 나오고 완료되면 사라짐.

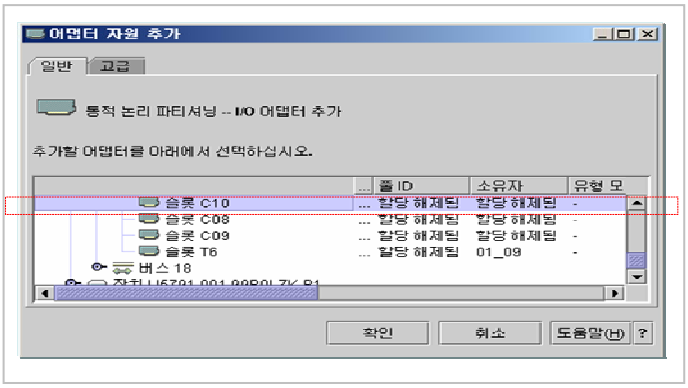
- HMC 통해 해당 파티션의 Adapter 를 추가

1) 추가 메뉴 선택

► HMC Console 전체 화면 → 해당 파티션을 선택하고 오른쪽 마우스 클릭 → 동적 논리적 파티셔닝 → 프로세서 자원 → 추가 선택



2) 추가할 Adapter 지정



3) Resource 추가

► 확인 버튼 클릭 후 Resource 추가

14.2.6. 공유 I/O 장치 DLPAR (I/O 장치 자원 변경)

- 4mm DAT 장치 및 CDROM 이동

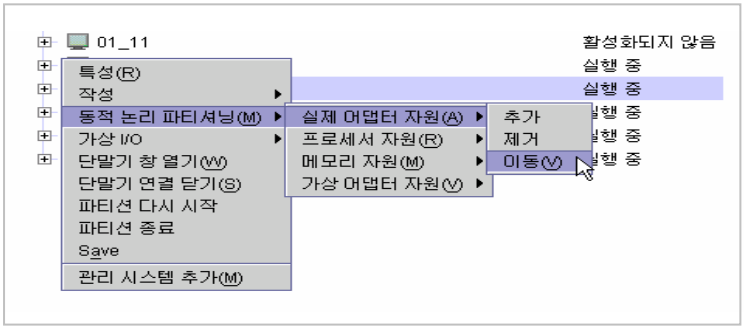
1) 제거할 장치 확인 및 삭제

► 서버에서 제거할 장치 확인 및 삭제

```
[khan1:/] lsdev -C |grep rmt0
rmt0             Available 0D-08-00-2,0   LVD SCSI 4mm Tape Drive
[khan1:/] lsdev -C |grep cd0
cd0              Available 0D-08-01-5,0   16 Bit LVD SCSI DVD-RAM Drive
[khan1:/] lsdev -C |grep 0D-08
cd0              Available 0D-08-01-5,0   16 Bit LVD SCSI DVD-RAM Drive
rmt0             Available 0D-08-00-2,0   LVD SCSI 4mm Tape Drive
rmt1             Available 0D-08-01-6,0   LVD SCSI 4mm Tape Drive
scsi1            Available 0D-08-00       PCI-X Dual Channel Ultra320 SCSI
scsi2            Available 0D-08-01       PCI-X Dual Channel Ultra320 SCSI
sisscsia0        Available 0D-08          PCI-XDDR Dual Channel Ultra320 S
[khan1:/] lsdev -Cl sisscsia0 -F parent
pci13
[khan1:/] rmdev -Rl pci13
```

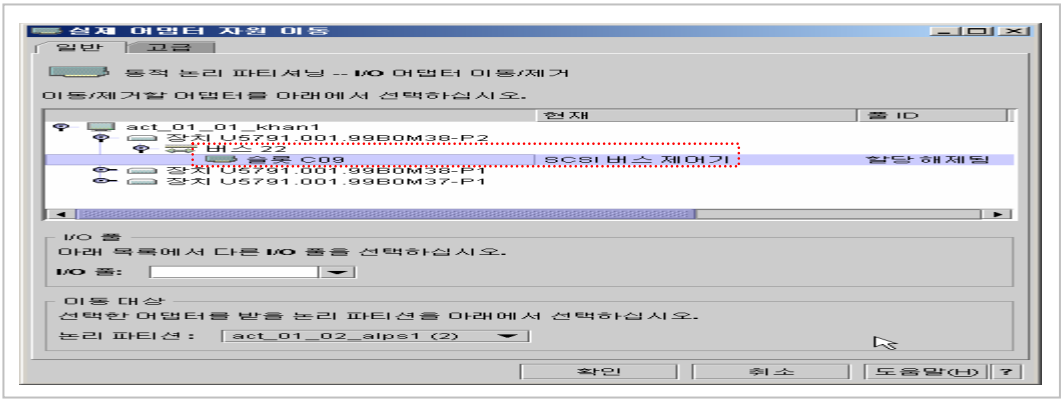
2) 이동 메뉴 선택

► 동적 논리 파티셔닝 → 실제 어댑터 자원 → 이동 선택



3) 이동할 장치 선택

► 이동할 장치 선택 후 이동 대상에 “논리 파티션” 선택



4) 다른 서버에서 장치 추가

► “cfgmgr –v” 명령어를 사용하여 장치 추가

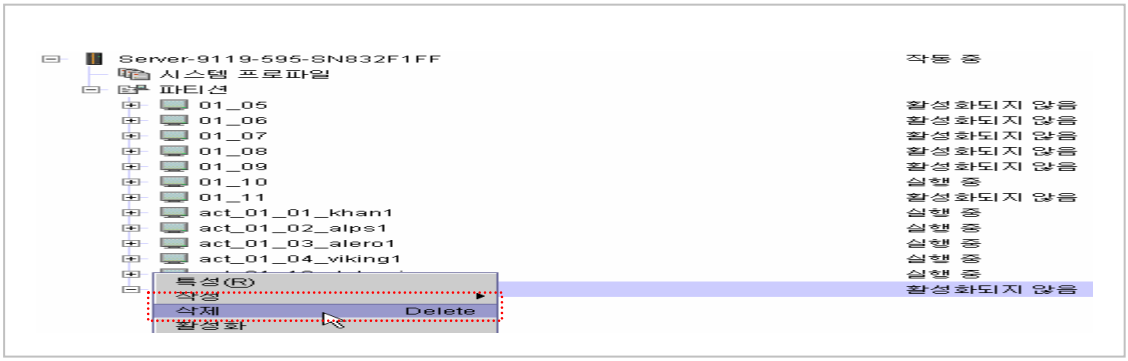
► 추가된 장치 확인

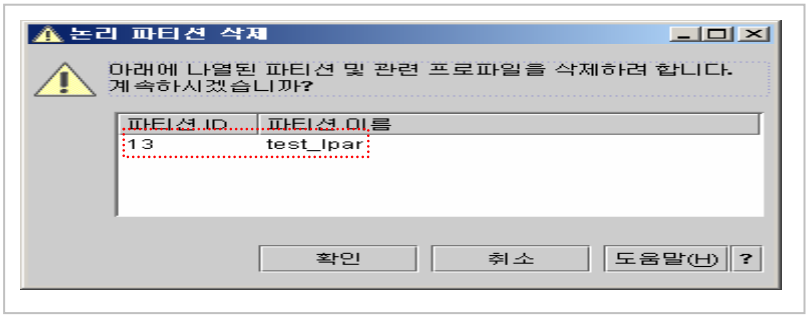
```
[alps1:/] lsdev -C |grep rmt0
rmt0             Available 0D-08-00-2,0   LVD SCSI 4mm Tape Drive
[alps1:/] lsdev -C |grep cd0
cd0              Available 0D-08-01-5,0   16 Bit LVD SCSI DVD-RAM Drive
```

14.2.7. LPAR 삭제

- LPAR 삭제

► 삭제할 partition 선택하여 오른쪽 마우스를 눌러 삭제





14.3. HMC 백업과 복구

14.3.1. Backing up critical HMC data

Using the HMC, you can back up all important data, such as the following:
User-preference files
User information
HMC platform-configuration files
HMC log files
HMC updates through Install Corrective Service

The Backup function saves the HMC data stored on the HMC hard disk to the following:
DVD media
USB Flash Memory Device
Remote system mounted to the HMC file system (such as NFS)
Remote site through FTP

Back up the HMC after you have made changes to the HMC or to the information associated with logical partitions.
Note: Before data can be saved to removable media, the media must be formatted. To format media, click HMC Management > Format Media and follow the steps.

To back up the HMC, you must be a member of one of the following roles:
super administrator
operator
service representative

To back up the HMC critical data, do the following:
In the Navigation area, click HMC Management.
Select Back up HMC Data.
Select an archive option. You can back up to media on the local system, back up to a mounted remote system, or send backup data to a remote site.
Follow the instructions on the window to back up the data.

14.3.2. Restoring critical HMC data

To restore the HMC data, you must be a member of one of the following roles:
super administrator
operator
service representative

Select the data-restoration procedure based on the data archiving method used:

Restoring from DVD

Restore data that was archived to DVD.

If the critical console data has been archived on a DVD, do the following:

Select 1 - Restore Critical Console Data from the menu displayed at the end of the HMC reinstallation.

Insert the DVD containing the archived console data. On the first boot of the newly installed HMC, the data is automatically restored.

Restoring from a remote server

Restore data that was archived to a remote FTP or NFS sever.

If the critical console data has been archived remotely, do the following:

Manually reconfigure network settings to enable access to the remote server after the HMC is newly installed. For information about configuring network settings, see Configuring the HMC.

In the Navigation area, click the Licensed Internal Code Maintenance icon.

In the Contents area, click the HMC Code Update icon.

Select Restore Remote Console Data.

Select the type of remote restoration.

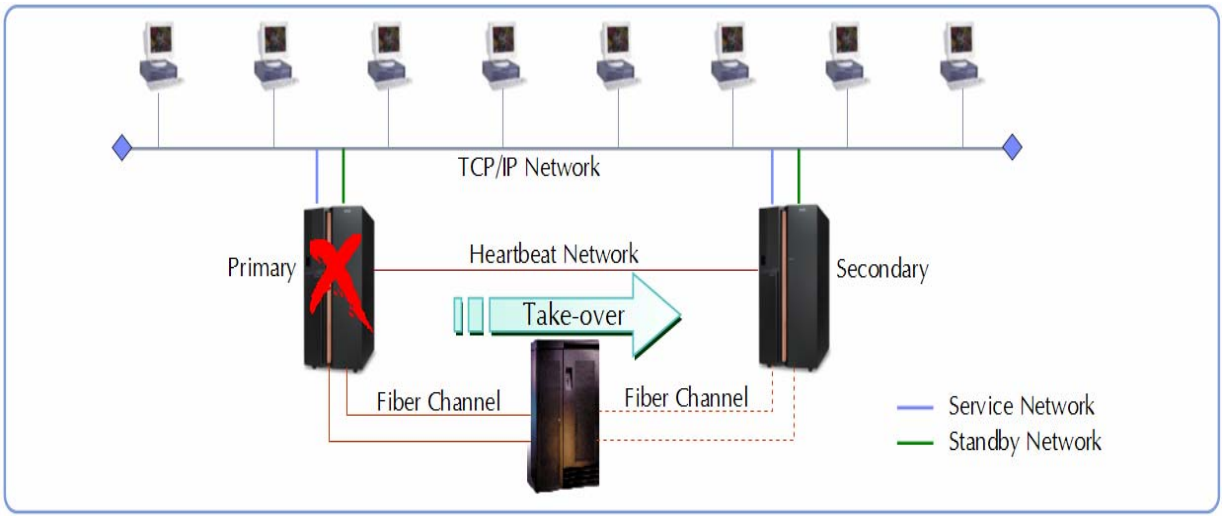
Follow the directions on the window to restore the critical console data. The data automatically restores from the remote server when the system is rebooted.

15. HACMP

15.1. HACMP 구성 방식

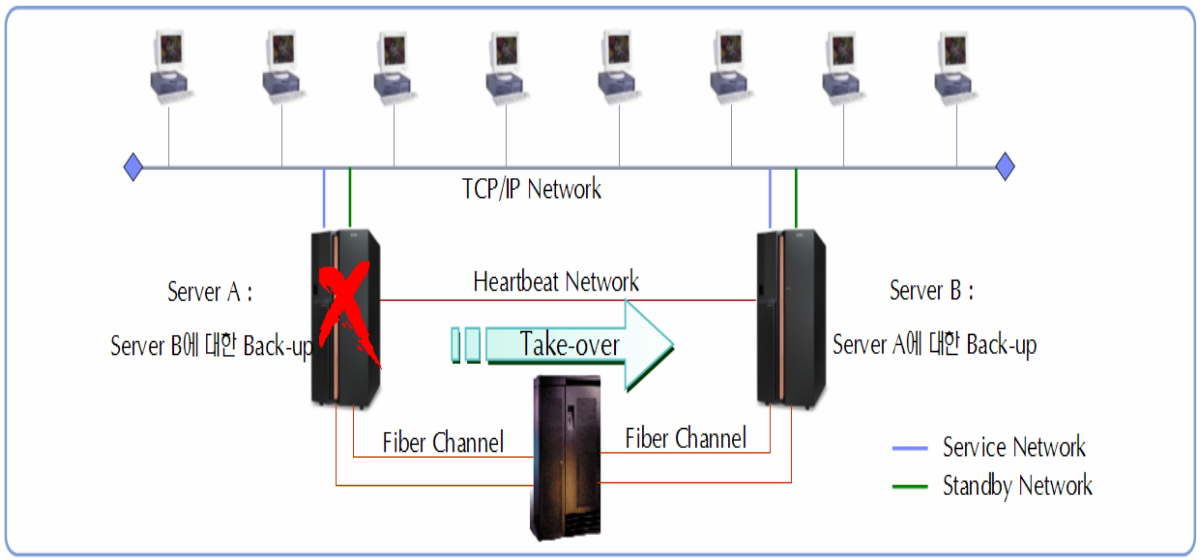
15.1.1. Active-Standby(hot – Standby) 방식

Back-up 시스템은 Primary 시스템이 장애가 발생할 때까지 기다리고 있다가 Primary 에 장애가 발생하면 공유자원(Shared Resource)의 관리를 계속하며, 운영하고 있는 Application 은 재수행합니다. 이때 서버에서 운용하고 있는 프로그램(예, DBMS)은 코드의 변환이 불필요합니다. Back-up 시스템은 Primary 시스템이 클러스터상에 다시 Join 할 때까지 서비스를 계속합니다. Primary 와 Standby 의 역할은 고정되어 있는 환경입니다. Primary 시스템이 복구가 완료되면, Standby 로 Take-over 되었던 공유자원과 Application 이 Fall-back 되게 됩니다.



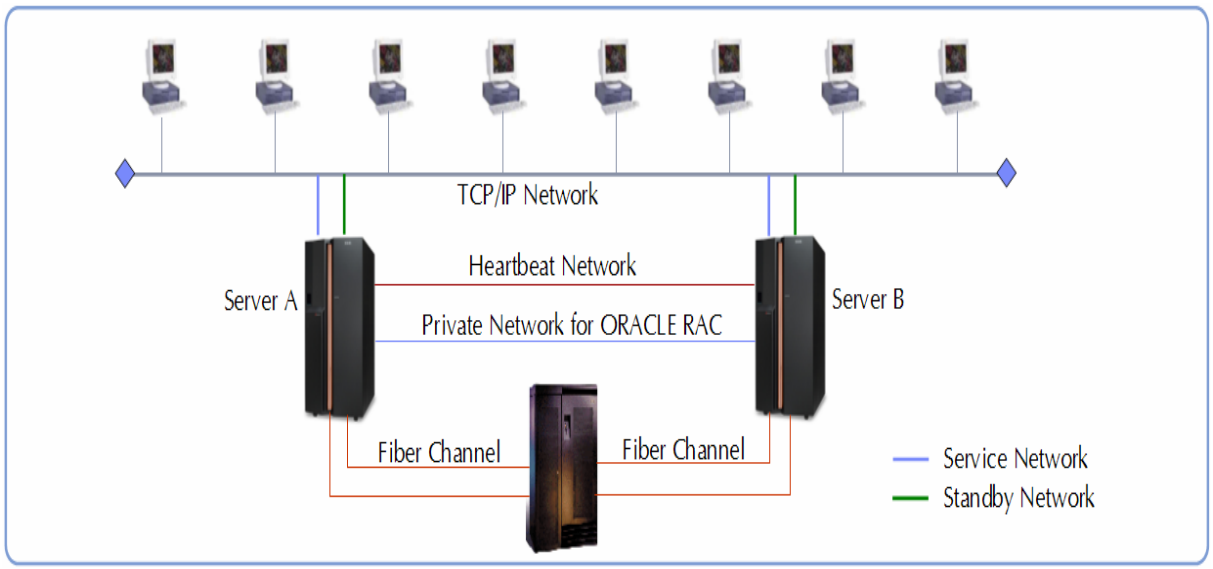
15.1.2. Mutual Take Over 방식

One-Sided Take-over Mode 와 거의 유사하지만, 어는 한쪽 시스템에 장애가 발생하면 다른 쪽 시스템이 공유자원 및 운영 중인 Application 을 Take-over 합니다. 이러한 구성은 클러스터상의 두 시스템 모두가 중요한 Application 을 수행하거나 양쪽 시스템의 부하를 동일하게 감당할 수 있을 때 사용합니다.



15.1.3. Concurrent VG Access

운용 중인 Application 이 양쪽 Server 가 동시에 접근이 가능하며 이는 HA 측면도 있지만, 서버 용량확장 및 load balancing 에 의미가 있습니다. 하나의 데이터베이스에 두 개의 시스템이 동시에 데이터를 요구하여 사용할 수 있으며, 이를 지원하는 데이터베이스로는 ORACLE 9i 의 RAC 같은 Shared Architecture 을 지원하는 데이터베이스를 사용할 수 있습니다. 이때 시스템 측면의 성능향상은 Linear 하지는 않지만 ORACLE RAC 와 운영 시 2 Node 구성 시에는 약 1.8 배, 4 Node 구성 시에는 약 3.3 배, 6 Node 구성 시에는 약 4.7 배의 성능향상을 기대할 수가 있습니다.



15.2. HACMP Config Guidelines

15.2.1. HACMP 구성 시 고려 사항

- HACMP 5.1 이후는 IP Alias 방식으로 가급적 구성한다.

HACMP 4.5 version 까지는 주로 IP replacement 방식을 사용하였습니다. HACMP Service 가 시작되면 Boot IP 가 Service IP 로 바뀌고 adapter failure 시에 Standby adapter 가 IP 를 take-over 받게 됩니다. Primary Node failure 가 발생하면 Standby Node 로 Service IP 가 take-over 됩니다.

HACMP 5.1 Version 부터 IP Aliases 방식을 주로 사용하며 System startup 시에 persistent ip (관리용 IP)가 boot ip 에 alias 방식으로 추가되며 HACMP Service 가 시작되면 Service IP 도 역시 alias 방식으로 추가됩니다.

Adapter failure 나 Node down 시에도 마찬가지로 alias 방식으로 기존 boot ip 에 추가적으로 생성되게 됩니다. 장점은 IP Replacement 방식에 비해 take-over 속도가 빠릅니다.

- Service IP, Boot1 IP, Boot2 IP 등의 IP 는 Different Subnet 으로 구성한다.

Service IP 와 Boot IP 는 반드시 서로 다른 subnet 으로 구성해야 합니다.

planning 단계에서 Service IP, Boot 1 IP, Boot2 IP 의 network 을 반드시 분리해서 설정해야만 Cluster verification 시 error 를 방지할 수 있습니다.

Persistent IP 는 Service IP 와 같은 대역의 IP 를 사용합니다.

아래는 file 은 /etc/hosts 파일의 예제입니다. (oracle rac 환경이며 interconnect IP 도 분리합니다.)

```
170.10.183.108   persist                    #Persist
170.10.183.110   server1          #Service
50.10.1.1                boot             #Boot
10.10.1.1                rac1             #Interconnect1
10.10.2.1                rac2                      #Interconnect2

170.10.183.109   persist                    #Persist
170.10.183.111   server2          #Service
50.10.1.2                boot             #Boot
10.10.1.2                rac1             #Interconnect1
10.10.1.2                rac2             #Interconnect2
```

- Remote 관리를 위해 Persistent IP 를 정의 한다.

Persistent IP 의 사용 목적은 HACMP Service 관리를 위해 추가적으로 할당되는 alias 방식의 IP 입니다.

HACMP Service IP 는 HACMP Service 가 Startup 되어야만 IP 가 살아나기 때문에 HACMP Service 를 Stop 시키면 사설 IP 인 Boot IP 만 있는 경우는 Remote 에서 telnet 접속을 할 수 없습니다. 물론, 콘솔에서 직접 Login 해서 HACMP Service 를 Startup 시킬 수는 있지만 관리상의 불편함 때문에 Persistent IP 를 사용할 것을 권장합니다.
또한 AIX Booting 시에 Persistent IP 가 binding 되어있어야만 Default Gateway 가 자동으로 설정되게 됩니다.

- “Discover HACMP-related Information from Configured Nodes” 가 정상적으로 수행되는지 확인할 것.
Configuration 정보 변경 시 반드시 재수행 해줍니다.
아래 메뉴에서 수행해주며 정상 수행완료 시 Command : OK message 가 나옵니다.
Cluster Node 구성 후 이 메뉴를 실행하면 기본적인 Network 및 Volume Group 구성을 조사해서 /usr/es/sbin/cluster/etc/config/cl*_config (clip_config, clvg_config) 파일에 해당 정보를 기록합니다.

smitty hacmp > Extended Configuration > Discover HACMP-related Information from Configured Nodes

COMMAND STATUS

Command: OK stdout: yes stderr: no

Before command completion, additional instructions may appear below.

[TOP]

Discovering IP Network Connectivity

Retrieving data from available cluster nodes. This could take a few minutes...

Discovered [2] interfaces

IP Network Discovery completed normally

Discovering Volume Group Configuration

Initializing..

Gathering cluster information, which may take a few minutes...

Processing...

Storing the following information in file

/usr/es/sbin/cluster/etc/config/clvg_config

[MORE...219]

- Cluster ID 는 default 값인 1 을 사용하지 않는다.
Cluster ID 번호는 Network 내에서 유일해야 하며 특히 하나 이상의 Cluster 가 Network 내에 존재할 때 겹치지 않도록 주의하여야 합니다. Cluster ID 의 default 값이 1 이므로 이러한 위험

성을 줄이기 위해 될 수 있는 한 Cluster ID 는 1 을 사용하지 않는 것이 좋습니다.
아래 예제처럼 /usr/es/sbin/cluster /clstat –a 명령을 실행하면 cluster Name, Cluster ID, State 등을 확인할 수 있습니다.

- Config_too_long 이 발생하지 않도록 Cluster 를 사전 검증한다.
config_too_long 메시지는 Cluster event 가 Cluster config 에 정의된 time-out 기간보다 실행시간이 더 오래 걸릴 경우 보여지게 된다.
Default 로 정의된 time-out 시간은 360 초이다.
HACMP Event 가 360 초 이상 실행시간이 걸리면 HACMP 는 config_too_long Warning message 를 /tmp/hacmp.out file 에 뿌리게 된다. 매 30 초마다 logging 됩니다.
아래 메시지는 rg_move Event 가 정상적으로 수행되지 않을 때 보여지는 config_too_long 메시지입니다.

Jan 18 20:51:58 EVENT START: config_too_long 720 /usr/es/sbin/cluster/events/rg_move.rp
:config_too_long[64] [[high = high]]
:config_too_long[64] version=1.11
:config_too_long[65] :config_too_long[65] cl_get_path
HA_DIR=es
:config_too_long[67] NUM_SECS=720
:config_too_long[68] EVENT=/usr/es/sbin/cluster/events/rg_move.rp
::config_too_long[88] grep -v cname
CLUSTER=hkmc_cluster
:config_too_long[89] TIME=720
:config_too_long[90] sleep_cntr=0
:config_too_long[95] [-x /usr/lpp/ssp/bin/spget_syspar]
WARNING: Cluster hkmc_cluster has been running recovery program '/usr/es/sbin/cluster/events/rg_move.rp' for 720 seconds. Please check cluster status.
WARNING: Cluster hkmc_cluster has been running recovery program '/usr/es/sbin/cluster/events/rg_move.rp' for 750 seconds. Please check cluster status.
WARNING: Cluster hkmc_cluster has been running recovery program '/usr/es/sbin/cluster/events/rg_move.rp' for 780 seconds. Please check cluster status.
WARNING: Cluster hkmc_cluster has been running recovery program '/usr/es/sbin/cluster/events/rg_move.rp' for 810 seconds. Please check cluster status.

config_too_long 문제를 해결하기 위한 첫 번째 Solution 으로는 문제가 있는 Event Script 를 확인 후 수정하는 것이며 HACMP 에서 Event duration time 을 증가시키는 것도 한 방법입니다.
Event duration time 은 아래 메뉴에서 설정할 수 있습니다.
smitty hacmp > Extended Configuration > Extended Event Configuration > Change/Show Time

Until Warning

Max. Event-Only Duration (in seconds)	Default 값은 180 초이며 Acquiring or releasing Resource Group 과 연관되지 않는 Event 의 Maximum duration time 값을 설정한다.
Max Resource Group Processing Time	Default 값은 180 초이며 Acquiring or releasing Resource Group 과 관련된 Event 의 Maximum duration time 값을 설정한다.
Total time to process a Resource Group Event Before a warning is displayed	config_too_long script 구동 전에 Cluster manager 가 기다리는 Total time 을 의미하며 위 두 개 field 값의 합으로 Edit 가 불가능하다.

- Serial Network 을 반드시 설정한다.

Node 의 IP Subsystem Failure 와 Node 자체의 Failure 를 구분하기 위해 모든 Cluster 는 IP 가 없어도 Node 간에 통신이 가능하도록 Serial Network 와 같은 Non-IP Network 의 구축을 권고합니다.

Non-IP heartbeat 네트워크는 cluster node 를 직접적으로 연결하는 serial network 로 node/network failure 를 구분해준다.

Cluster 내의 각 node 가 정상적으로 동작하는지를 확인하기 위하여 주기적으로 heartbeat 을 주고 받음으로써, 불필요한 takeover 를 방지합니다.

Serial 네트워크는 8-port serial adapter 의 RS232 라인으로 구성됩니다.

Heartbeat 는 HACMP 는 cluster manager 에 정의된 모든 IP interface 를 통해 교환된다.

Serial network 에 failure 가 발생하면, error log 에 message 가 더해지나, failover action 은 이뤄지지 않는다.

Minimum 2 개의 Line 이 권장되며 Adapter 까지 이중화하는 게 좋습니다.

Non-IP heartbeat 는 SSA 와 SCSI 공유 디스크를 통해서도 가능합니다.

tty 설정 값 중 'Enable LOGIN'은 disable 로 설정할 것을 권장합니다.

- Resource 에 참여하는 Volume Group 에 대한 auto_varyon Option 을 disable 한다.

Shared Volume Group 은 각 Node 에서 import 될 때 동일한 Volume Group Major Number 를 가져야 하며 Cluster Service 가 실행되지 않은 상황에서 Varyon 이 되지 않도록 autovaryon Option 을 no 로 설정해야 합니다.

또한 Shared Volume Group 내의 모든 filesystem 은 Resource Group 에 정의되어 있어야 하며 auto mount 와 disk accounting option 은 no 로 설정해야 합니다.

사용 가능한 Volume Group Major Number 를 아래 명령으로 확인합니다.

[root:/] lvlstmajor

55,57...61,62

- Primary Node 에서 Volume Group 생성 후에 chvg -an “volume group name” 명령으로 autovaryon Option 을 no 로 설정합니다.
- Primary Node 에서 varyoffvg 명령을 수행합니다.
- Standby Node 에서 importvg -V “설정한 Major number” -y “volume group name” “diskname” 으로 import 해줍니다.
- Standby Node 에서 chvg -an “volume group name” 으로 마찬가지로 autovaryon Option 을 no 로 설정합니다.
- Standby Node 에서 varyoffvg 명령을 수행합니다.
- lsattr -El “Volume Group name” | grep auto_on 명령으로 현 설정상태를 확인할 수 있다.

- OEM disk 를 공유하는 경우 해당 Vendor 에서 제공하는 reset 방식을 custom disk method 에 등록한다.

EMC 또는 Hitachi Storage 를 Shared Disk 로 사용할 경우 Storage Vendor 에서 제공하는 custom disk method 를 등록해야 합니다.

- SNMP 서비스를 반드시 사용하도록 설정하며 보안을 위해 community 설정을 public 이 아닌 다른 값으로 변경한다.

SNMP 는 단순히 clstat 등의 cluster info 를 확인하는 목적만으로 사용되는 것은 아닙니다.
clsmuxpdES 는 SMUX (snmp multiplexing) protocol 을 통해 cluster object 들을 감시하고 update 받고 합니다.
SNMP 가 비정상적으로 동작하여 /usr/es/adm/cluster.log 에 invalid error 등을 만나게 되면 cluster 가 비정상적으로 동작할 수도 있습니다.
따라서, SNMP 는 반드시 사용해야 하며 보안이슈가 있으면 Public 이 아닌 다른 community name 을 지정해주면 됩니다.
HACMP 환경하에서는 SNMP v1 을 사용할 것을 권장합니다.
AIX 5.2 부터는 Default 가 SNMP v3 이기 때문에 다음 command 로 변경해주며 설정파일은 /etc/snmpd.conf 입니다.
#snmpv3_ssw -1

- 가급적 Cluster 는 동일 fix version 을 사용해야 하며, AIX patch Level 도 동일하게 설정한다.

Cluster 간에는 동일한 AIX Version 과 HACMP Version 을 유지해야 하며 “**lppchk –v**”로 점검했을 때 빠진 파일셋이 없도록 OS 를 유지해야 합니다.
Node 간의 정합성을 위해 생성되는 User ID, Group ID 및 Password 등은 모든 Node 에서 일치시키는 것이 권장됩니다.

- sync 작업 완료 후에 odmget HACMPtopsvcs 를 통해 instance id 가 동일한지 확인한다. HACMP instance number 가 불일치 할 경우 HACMP sync 가 error 가 나게 되며, 또한 HACMP Service 가 정상적으로 startup 되지 않을 수도 있고 take-over 도 정상적으로 수행하지 못하게 됩니다.

Instance number 는 sync 작업시 수치가 증가하게 됩니다.

현재 instance number 는 아래 두 개의 명령으로 확인가능하며 모든 노드에서 동일해야 합니다.

 # odmget HACMPtopsvcs

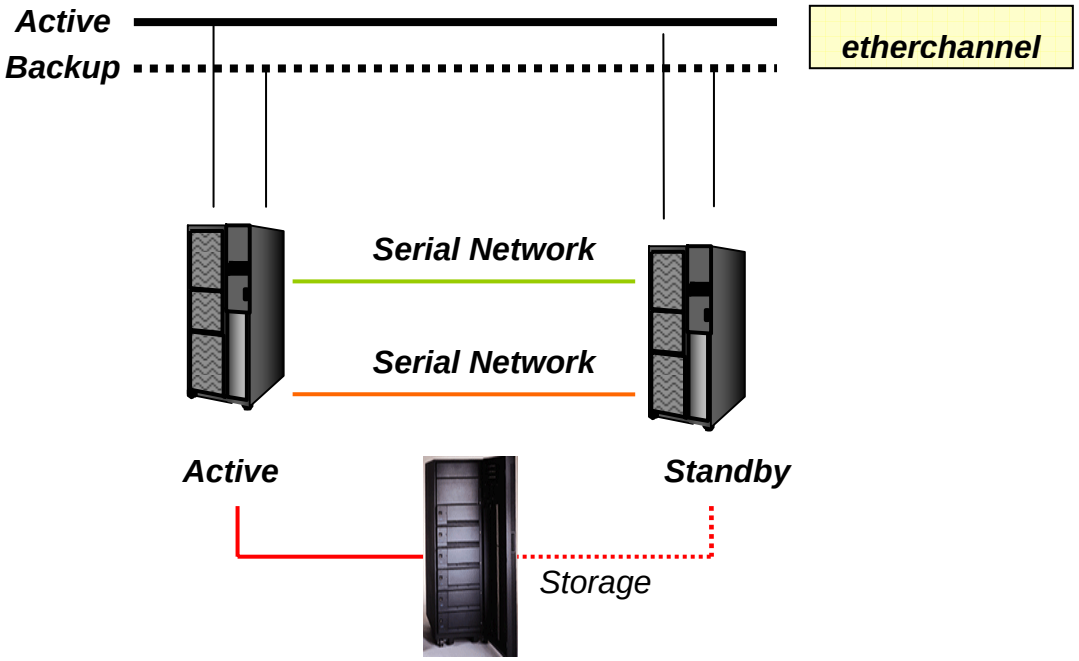
 # cat /var/ha/run/top*/machines.*.lst

- Dead Man Switch 관련 항목을 설정한다.

Dead Man Switch 란 Clean takeover 를 수행하기 위한 도구로서 각 Node 에 존재합니다. 주어진 시간 내에 이 Dead Man Switch 가 reset 되지 않으면 해당 Node 를 다운시키고 takeover 를 수행하게 되는데 이로 인해 work load 가 크게 걸리는 Node 의 상태를 잘못 판단하여 일어나지 말아야 할 takeover 가 일어나는 경우가 있습니다. 이러한 경우를 Dead Man Switch Timeout 문제라 하는데 일반적으로 Cluster 구성 시에 다음과 같은 점을 고려해서 Dead Man Switch Timeout 문제를 방지 합니다.

- I/O Pacing enable 시키는데 high water mark 를 33, low water mark 를 24 로 설정한다.
- sync daemon 이 수행되는 간격을 조정한다. 보통 10 초 간격을 권장.
- Ethernet adapter 의 failure detection rate 을 slow 로 조정한다.
- Etherchannel Network 을 구성하여 Network topology 를 단순화한다.

AIX 5.1 이후 버전과 HACMP 4.5 버전 이후부터 HACMP 상에서 Etherchannel 을 지원합니다. Etherchannel 의 장점은 Default 로 OS 에서 제공되며 HACMP 에서 사용되는 topology 를 단순화 해줍니다. 또한 IP Subnet 의 수도 줄여주는 장점이 있습니다.



15.3. HACMP 구성 절차

15.3.1. Etherchannel 설정

- Etherchannel 에서 사용할 device 를 rmdev -l 명령으로 define 시킨다.
- smitty etherchannel > Add an etherchannel / Link Aggregation 메뉴에서 한 개의 network interface 를 선택하고 Backup adapter 를 한 개 지정해준다.
- 10/100 NIC card 로 Etherchannel 구성 시에는 link polling 기능을 enable 시켜줘야 합니다. #chdev -l ent# -a poll_link=yes 명령으로 설정할 수 있습니다.
- 실제 어떤 Interface 에서 서비스되고 있는지 확인하는 명령은 #entstat -d ent# 이며, primary Adapter 항목을 보면 확인할 수 있습니다. (여기서 ent#는 etherchannel 로 생성된 가상 device 이다.)

아래 내용은 etherchannel setup 절차입니다. ent0 를 primary adapter, ent1 을 backup adapter 로 설정하는 예제입니다.

실행이 완료되면 추가로 가상의 Adapter 가 생성되며 primary adapter 의 MAC address 를 가져갑니다.

Add an EtherChannel / Link Aggregation

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]
EtherChannel / Link Aggregation Adapters	ent0
Enable Alternate Address	no
Alternate Address	[]
Enable Gigabit Ethernet Jumbo Frames	no
Mode	standard
Hash Mode	default
Backup Adapter	ent1
Automatically Recover to Main Channel	yes
Perform Lossless Failover After Ping Failure	yes
Internet Address to Ping	[]
Number of Retries	[]
Retry Timeout (sec)	[]

15.3.2. Network 설정

- 아래 절차로 IP 를 설정합니다.

smitty mktcpip 실행 후 etherchannel 로 설정한 가상 adapter 를 선택합니다. Persistent ip 및 gateway 를 입력합니다.

Minimum Configuration & Startup

To Delete existing configuration data, please use Further Configuration menus

Type or select values in entry fields.

Press Enter AFTER making all desired changes.

	[Entry Fields]	
* HOSTNAME	[server1]	
* Internet ADDRESS (dotted decimal)	[156.147.184.11]	
Network MASK (dotted decimal)	[255.255.255.0]	
* Network INTERFACE	en10	
NAMESERVER		
Internet ADDRESS (dotted decimal)	[ ]	
DOMAIN Name	[ ]	
Default Gateway		
Address (dotted decimal or symbolic name)	[156.147.184.1]	
Cost	[0]	#
Do Active Dead Gateway Detection?	no	+
Your CABLE Type	N/A	+
START TCP/IP daemons Now	no	

netstat -rn 명령으로 gateway 가 정상적으로 입력되었는지 확인합니다. 또한 rebooting 시 적용될 수 있도록 inet0 에 항목이 입력되어있는지 확인합니다.

#lsattr -El inet0 |grep route 명령으로 확인 가능합니다.

smitty chinet 명령으로 Persistent IP 를 Boot IP 로 변경해줍니다.

Change / Show a Standard Ethernet Interface

Type or select values in entry fields.

Press Enter AFTER making all desired changes.

	[Entry Fields]	
Network Interface Name	en10	
INTERNET ADDRESS (dotted decimal)	[10.10.10.1]	
Network MASK (hexadecimal or dotted decimal)	[255.255.255.0]	
Current STATE	up	+
Use Address Resolution Protocol (ARP)?	yes	+
BROADCAST ADDRESS (dotted decimal)	[]	
Interface Specific Network Options		
('NULL' will unset the option)		

rfc1323 []
tcp_mssdflt []
tcp_nodelay []
tcp_recvspace []

모든 노드에 Base IP setting 후에 서로간에 통신여부를 ping test 를 통해 확인합니다.

A. /etc/hosts 파일 구성

```
server1:/> vi /etc/hosts
127.0.0.1          loopback localhost      # loopback (lo0) name/address
10,10,10,1         server1_boot            # boot address
10.10.10.2         server2_boot            # boot address
156.147.184.10     server1_svc          # service address
156.147.184.11     server1_per             # persistent address
156.147.184.12     server2_per             # persistent address
```

먼저, HACMP 구성에 사용될 각종 address 를 각 node 의 /etc/hosts file 에 미리 입력을 해놓아야 합니다.
이 구성에서와 같이, ethernet adapter 1 장만을 사용하는 경우, node 당 boot address 1 개, service address 1 개, 그리고 필요 시 persistent address 1 개를 입력해둡니다.

Persistent address 는 HACMP version 5.1 부터 생긴 개념입니다. 즉, HACMP version 4 에서는, boot 및 service address 의 subnet 이 같았지만, version 5 부터는 boot 및 service address 의 subnet 이 다르므로, HACMP 를 기동하기 전에는 PC 등으로부터 system 에 접속하기가 번거롭습니다. 이를 극복하기 위해, service address 와 같은 subnet 상에 존재하는 "persistent address"라는 개념을 만들어, 이를 IP alias 형태로 구성시킵니다. 이 address 는 service address 와 같은 subnet 상에 위치하며, alias 형태로 지원된다는 점에서 service address 와 같지만, HACMP daemon 이 구동되지 않은 상태에서도 사용할 수 있고, 또 장애시에도 failover 되지는 않는다는 점에서는 service address 와 다릅니다.
Boot address 와 service address 는 서로 다른 subnet 상에 있어야 한다는 것에 완료 후에.
Ethernet adapter 가 2 개 이상 있는 system 에서 HACMP 를 구성할 때는, 두 번째 ethernet adapter 에 설정하는 2 번째 boot address 는 첫 번째 ethernet adapter 에 설정된 1 번째 boot address 와도 다른 subnet 에 있어야 한다는 것에도 유의하십시오.

15.3.3. Volume Group 구성

Shared Volume Group 을 구성합니다. 먼저 Primary Node 에서 Volume Group 생성 합니다.
Shared Volume Group 의 VG Major number 는 모든 노드에서 일치해야 합니다.

[root@server1:]/>lvlstmajor # 44 번부터 사용 가능한 경우의 예제입니다.

44...

smitty mkvg 명령으로 shared volume group 을 생성합니다.

Add an Original Volume Group

Type or select values in entry fields.

Press Enter AFTER making all desired changes.

	[Entry Fields]	
VOLUME GROUP name	[]	
Physical partition SIZE in megabytes		+
* PHYSICAL VOLUME names	[]	+
FORCE the creation of volume group?	no	+
Activate volume group AUTOMATICALLY at system restart?	yes	+
Volume group MAJOR NUMBER	[ ]	+#
Create VG Concurrent Capable?	no	+

생성된 Volume Group 에 파일시스템을 생성합니다.

파일시스템 속성 중 automount option 을 no 로 설정합니다.

Add an Enhanced Journaled File System

Type or select values in entry fields.

Press Enter AFTER making all desired changes.

	[Entry Fields]	
Volume group name	datavg	
SIZE of file system		
Unit Size	Megabytes	+
* Number of units	[]	#
* MOUNT POINT	[]	
Mount AUTOMATICALLY at system restart?	no	+
PERMISSIONS	read/write	+
Mount OPTIONS	[]	+
Block Size (bytes)	4096	+
Logical Volume for Log		+
Inline Log size (MBytes)	[]	#
Extended Attribute Format	Version 1	+

Enable Quota Management? no

+

Volume Group 의 속성 중 autovaryon 속성을 disable 해줍니다.

#chvg -an datavg 명령으로 변경.

아래 명령으로 반영 여부를 확인합니다.

```
[root@server1:]/>lsattr -El datavg
auto_on           n                                   N/A True
conc_auto_on      n                                   N/A True
conc_capable      n                                   N/A True
gbl_pbufs_ppv 0                                       N/A True
gbl_pbufs_pvg 0                                       N/A True
timestamp         45f025dc0ad2d9c5                    N/A True
vg_pbufs_ppv      0                                   N/A True
vgserial_id       00cfa93f00004c000000010fcb0d831f N/A False
```

Primary Node 에서 varyoffvg 후 다음절차로 Backup Node 에 importvg 해줍니다.

```
[root@server1:]/>varyoffvg datavg
[root@server1:]/>importvg –V 44 –y datavg hdisk100
[root@server1:]/>chvg –an datavg
[root@server1:]/>varyoffvg datavg
[root@server1]/>ls -al /dev | grep datavg                    #VG Major number 확인
crw-rw----    1 root         system           44,    0 Dec 29 06:53 datavg
```

15.3.4. HACMP Config

아래절차로 cluster 를 추가해줍니다. 반드시 한쪽 Node 에서 작업합니다.

smttiy hacmp > Extended Configuration > Extended Topology Configuration > Configure an HACMP Cluster > Add/Change/Show an HACMP Cluster

Add/Change/Show an HACMP Cluster

Type or select values in entry fields.

Press Enter AFTER making all desired changes.

[Entry Fields]

* Cluster Name [server_cluster]

NOTE: HACMP must be RESTARTED

on all nodes in order for change to take effect

아래 절차로 각 Cluster Node 를 추가해주며 communication path 로 /etc/hosts 파일에 지정한 boot adapter 를 지정해줍니다. 이 IP 는 HACMP sync 용도로 사용됩니다.

smttiy hacmp > Extended Configuration > Extended Topology Configuration > Configure an HACMP Node > Add a Node to the HACMP Cluster

Add a Node to the HACMP Cluster

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

	[Entry Fields]
* Node Name	[server1]
Communication Path to Node	[server1_boot]

아래 절차로 각 Node 에서 Network 정보 및 Volume 정보를 gathering 합니다.
Gathering 된 정보는 /usr/es/sbin/cluster/etc/config/cl*_config (clip_config, clvg_config) 파일에 저장됩니다.

smttiy hacmp > Extended Configuration > Discover HACMP-related Information from Configured Nodes

COMMAND STATUS

Command: OK stdout: yes stderr: no

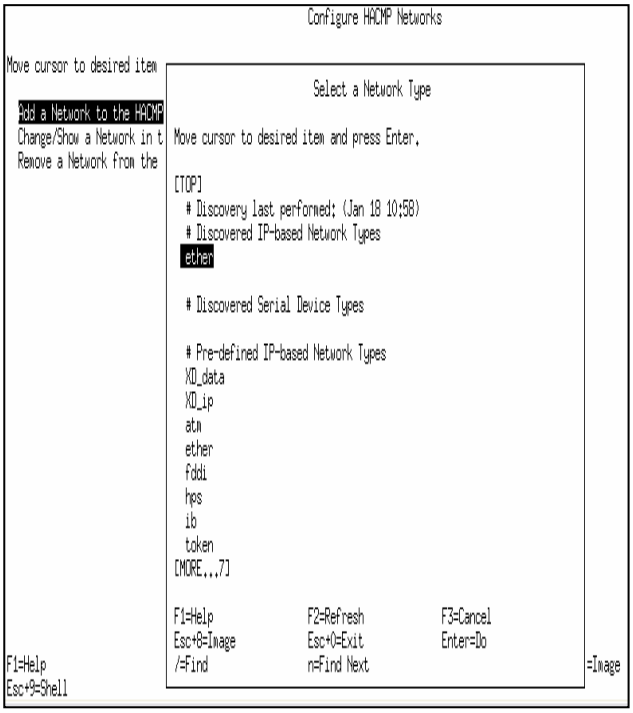
Before command completion, additional instructions may appear below.
[TOP]
Discovering IP Network Connectivity
Retrieving data from available cluster nodes. This could take a few minutes...
Discovered [2] interfaces
IP Network Discovery completed normally
Discovering Volume Group Configuration
Initializing..
Gathering cluster information, which may take a few minutes...
Processing...
Storing the following information in file
/usr/es/sbin/cluster/etc/config/clvg_config
[MORE...219]

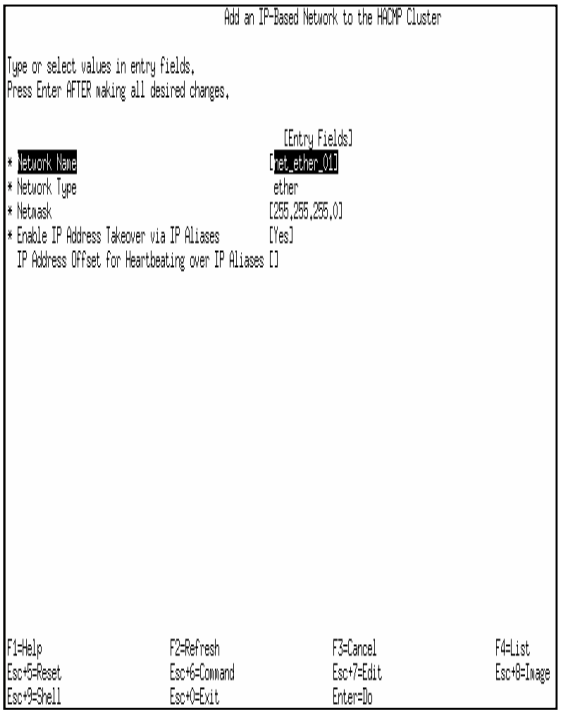
아래 절차로 Network 을 구성합니다. Discovered IP-Based Network Types 인 ether 를 선택합니다.

Network Name 은 그냥 사용해도 되면 변경도 가능합니다.

Enable IP Address Takeover via IP Aliases 항목을 IP Aliasing 방식을 사용하면 YES 로 설정하고 예전방식인 IP Replacement 방식을 사용하려면 No 로 설정합니다.

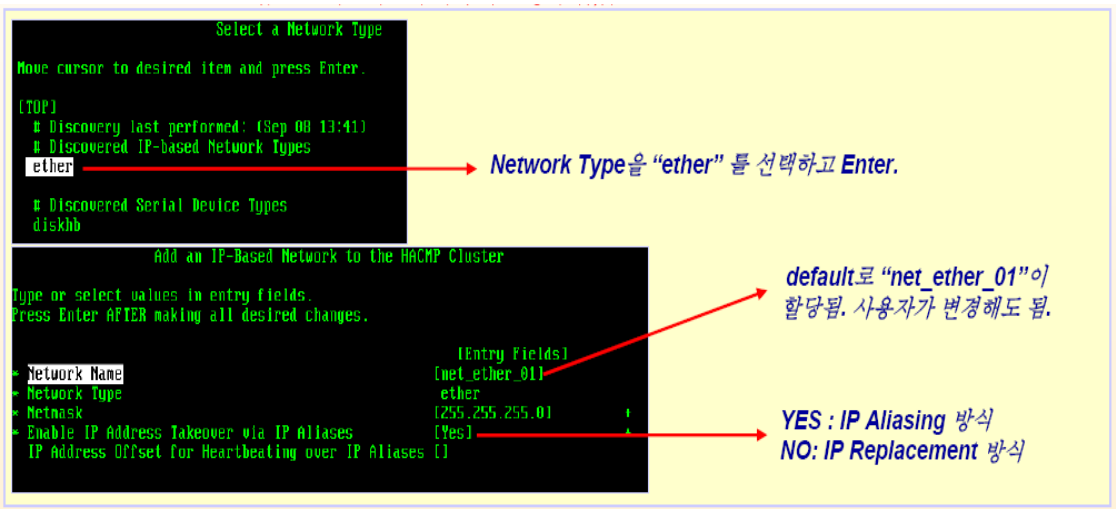
smttiy hacmp > Extended Configuration > Extended Topology Configuration > Configure an HACMP Networks > Add a Network to the HACMP Cluster





아래 절차로 Network Communication Interface 를 설정합니다. Discover 단계에서 가져온 항목을 F7 로 선택해줍니다.

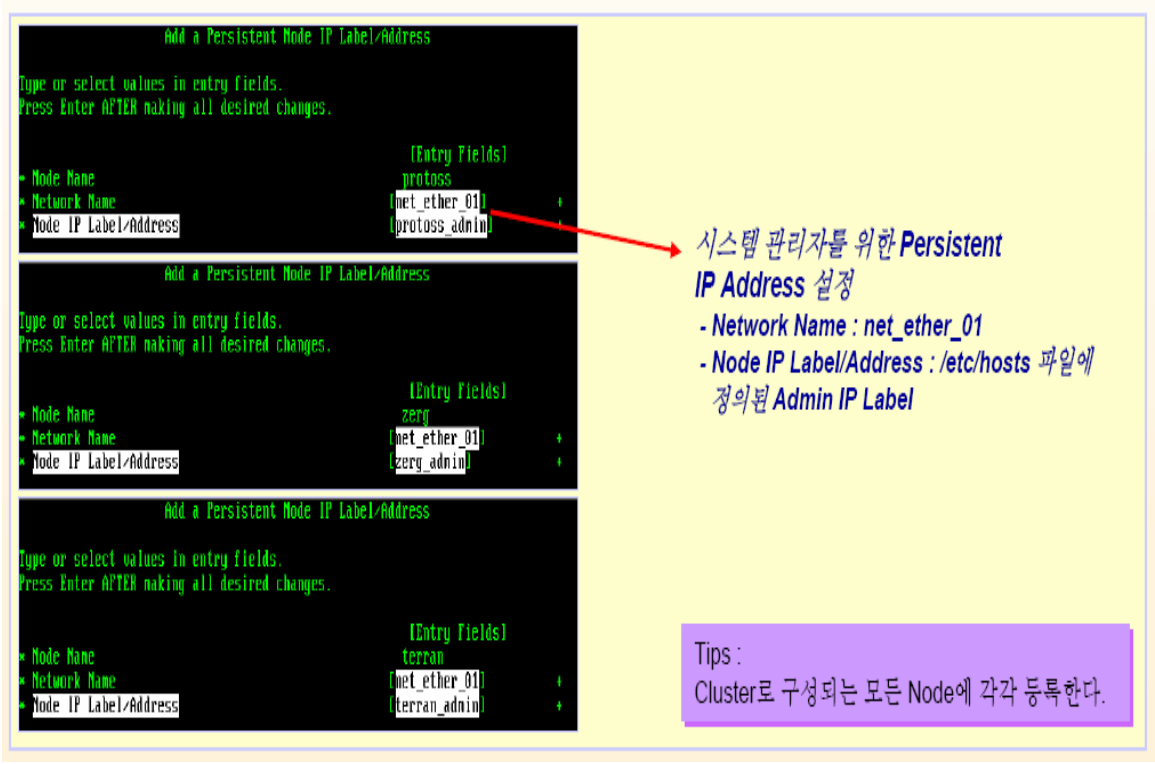
smttiy hacmp > Extended Configuration > Extended Topology Configuration > Configure a HACMP Communication Interfaces and Devices > Add a Communication Interfaces and Devices



아래 절차로 Persistent IP 를 설정해줍니다. Persistent IP 는 sync 이후에 alias 형태로 binding 됩니다.

smttiy hacmp > Extended Configuration > Extended Topology Configuration > Configure a HACMP Persistent Node IP Label / Addresses > Add a Persistent Node IP Label / Address

해당 Node 를 선택해줍니다. 모든 노드에 대해 한차례씩 Persistent IP 를 등록해줍니다.



해당 Field 에서 F4 를 누르면 해당 항목을 선택할 수 있습니다.

protoss node 에는 protoss_admin 을 선택, zerg node 에서는 zerg_admin 선택, terrab 에서는 terran_admin 을 선택해줍니다.

아래 절차로 Service IP 를 등록해줍니다.

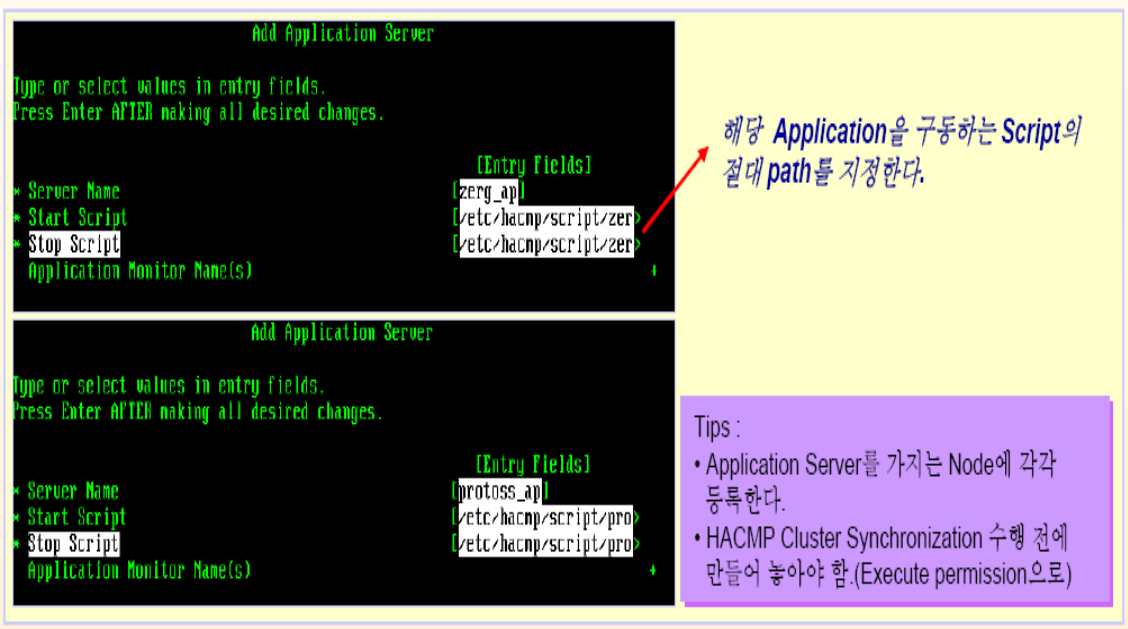
smttiy hacmp > Extended Configuration > Extended Resource Configuration > HACMP Extended Resource Configuration > Configure HACMP Service IP Labels / Addresses > Add a Service IP Label / Address

Configurable on Multiple Nodes 항목을 선택한다. (take-over 되는 항목)

아래 절차로 Application Server 를 등록해줍니다.

사전에 HACMP 모든 노드에 Application start script 및 stop script 를 생성해 놓는다.

smttiy hacmp > Extended Configuration > Extended Resource Configuration > HACMP Extended Resource Configuration > Configure HACMP Applications Servers > Configure HACMP Application Servers > Add an Application Server

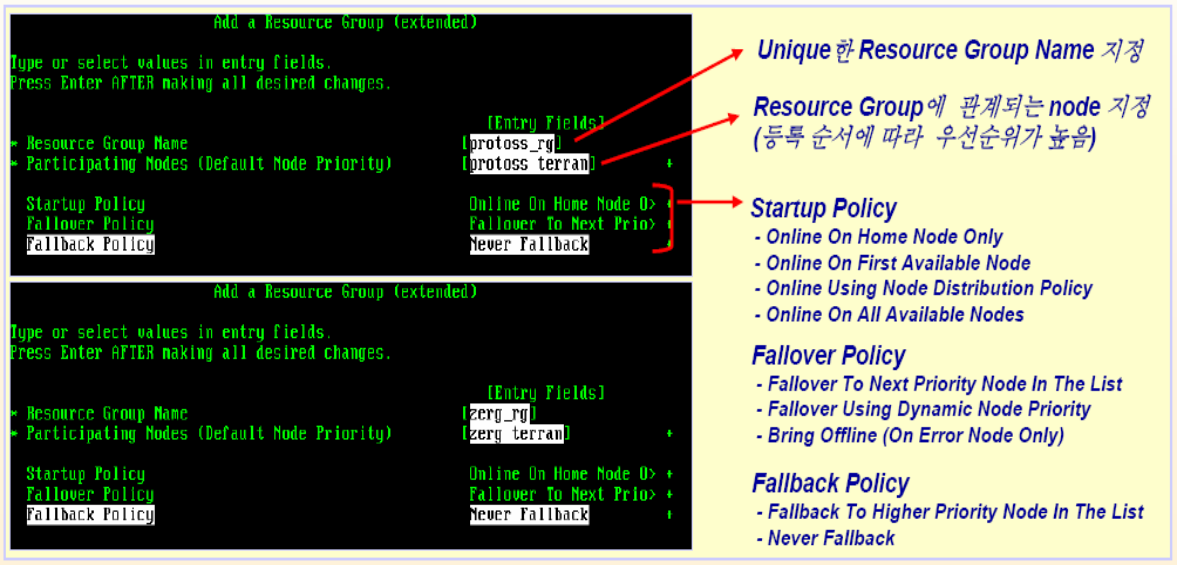


아래 절차로 Resource Group 을 생성합니다.

아래 예는 Active-Standby 방식의 예이며, Concurrent 구성 시에는 Online On Home Node Only 항목을 Online on All Available Node 로 변경해줘야 한다.

Fallback Policy 를 Never Fallback 으로 설정하면 Active Node 가 Standby Node 로 Take-Over 된 이후에 Active Node 를 다시 살려도 Resource Group 을 자동으로 가져오지 않으며 rg_move 옵션으로 수동으로 가져와야 한다.

smttiy hacmp > Extended Configuration > Extended Resource Configuration > HACMP Extended Resource Group Configuration > Add a Resource Group

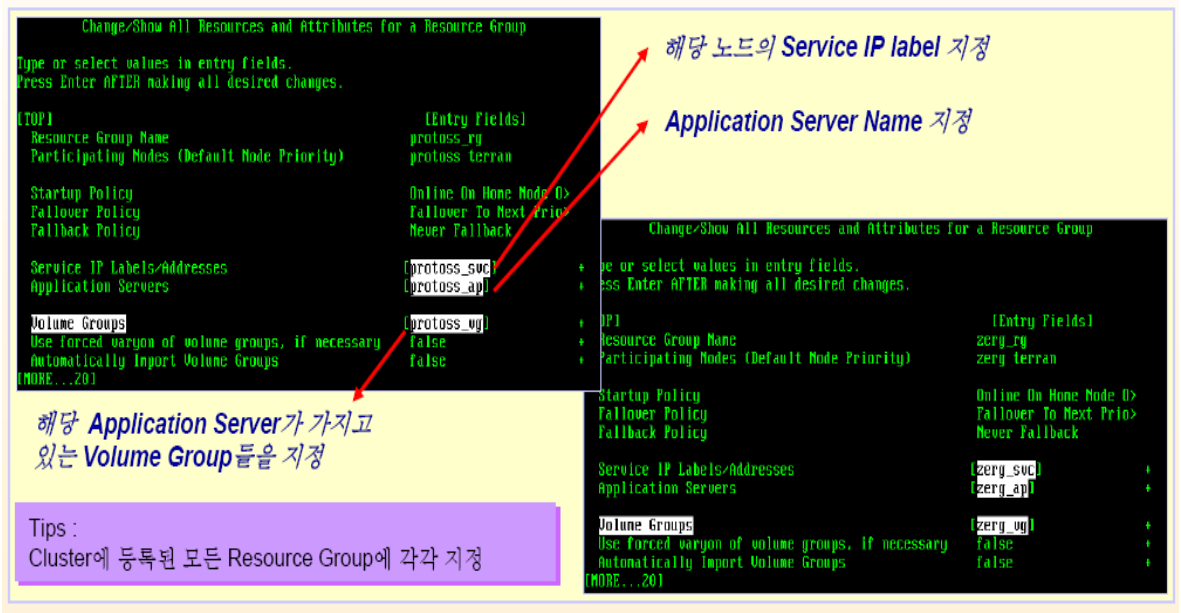


Participating Nodes 항목에 들어간 Node 의 순서가 Resource Group 에 대한 Ownership 순서입니다.

아래 절차로 Resource Group 속성을 변경합니다.

smttiy hacmp > Extended Configuration > Extended Resource Configuration > HACMP Extended Resource Group Configuration > Change / Show Resources and Attributes for a Resource Group

이전에 생성한 Resource Group 을 선택합니다.

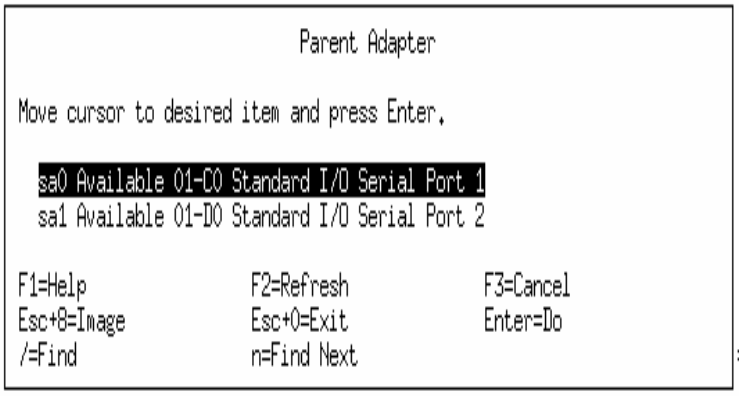


Service IP, Application Server, Volume Group 항목을 입력

아래 절차로 Serial Heartbeat Network 을 구성합니다.

양쪽 Node 에서 연결한 tty 를 AIX 구성합니다.

smitty mktty > Add a tty > tty rs232 Asynchronous Terminal 선택



```
                                    Add a TTY

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

[TOP]                                          [Entry Fields]
  TTY type                                     tty
  TTY interface                                rs232
  Description                                  Asynchronous Terminal
  Parent adapter                               sa0
* PORT number                                  [1]                                                    +
  Enable LOGIN                                 disable                                                +
  BAUD rate                                    [9600]                                                 +
  PARITY                                       [none]                                                 +
  BITS per character                           [8]                                                    +
  Number of STOP BITS                          [1]                                                    +
  TIME before advancing to next port setting   [0]                                                    +#
  TERMINAL type                                [dumb]
  FLOW CONTROL to be used                      [xon]                                                  +
  OPEN DISCIPLINE to be used                   [dtropen]                                              +
  STTY attributes for RUN time                 [hupcl,cread,brkint,icrnl,opost,tab3,onlcr,isig,icanon,echo,echoe,echok,ec> +
  STTY attributes for LOGIN                    [hupcl,cread,echoe,cs8]
  LOGGER name                                  []
  STATUS of device at BOOT time                [available]                                            +
  REMOTE Reboot ENABLE                         no                                                     +
  REMOTE Reboot STRING                         [#@reb@#]
  TRANSMIT buffer count                        [16]                                                   +#
```

Port Number 를 입력합니다.

Enable LOGIN 항목이 disable 인지 확인

Serial connection 이 정상적인지 확인하는 절차는 아래와 같습니다. (Cluster start 하기 전 확인할 경우)

#[root@server1:] cat < /dev/tty0

#[root@server2:] cat /etc/hosts > /dev/tty0

#[root@server1:] stty > /dev/tty0

#[root@server2:] stty < /dev/tty0

---

server1 node 쪽에서 hosts file 의 내용이 정상적으로 전송되는지 확인합니다.

*** cluster 가 운영 중일 때는 위의 명령을 실행치 말고 아래의 명령어로 확인

#[root@server1:]lssrc –ls topsvcs

```
Subsystem         Group            PID     Status
 topsvcs          topsvcs          336086  active
Network Name   Indx Defd Mbrs St Adapter ID      Group ID
net_ether_02_0 [ 0]    2    2  S 12.1.10.1       12.1.10.2
net_ether_02_0 [ 0] en1          0x403e9c55      0x403e9c67
HB Interval = 1 secs. Sensitivity = 10 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 1310 ICMP 0 Errors: 0 No mbuf: 0
Packets received: 1922 ICMP 0 Dropped: 0
NIM's PID: 536760
diskhb_0       [ 1]    2    2  S 255.255.10.1    255.255.10.3
diskhb_0       [ 1] rhdisk33    0x803ea244      0x803ea256
HB Interval = 2 secs. Sensitivity = 4 missed beats
Missed HBs: Total: 0 Current group: 0
Packets sent    : 643 ICMP 0 Errors: 0 No mbuf: 0
Packets received: 643 ICMP 0 Dropped: 0
NIM's PID: 540860
diskhb_1       [ 2]    2    2  S 255.255.10.0    255.255.10.2
diskhb_1       [ 2] rhdisk41    0x803ea245      0x803ea256
HB Interval = 2 secs. Sensitivity = 4 missed beats
```

- Disk Heartbeat 으로 구성된 hdisk? 를 통한 Packet sent/Received 상태를 모니터링 함.

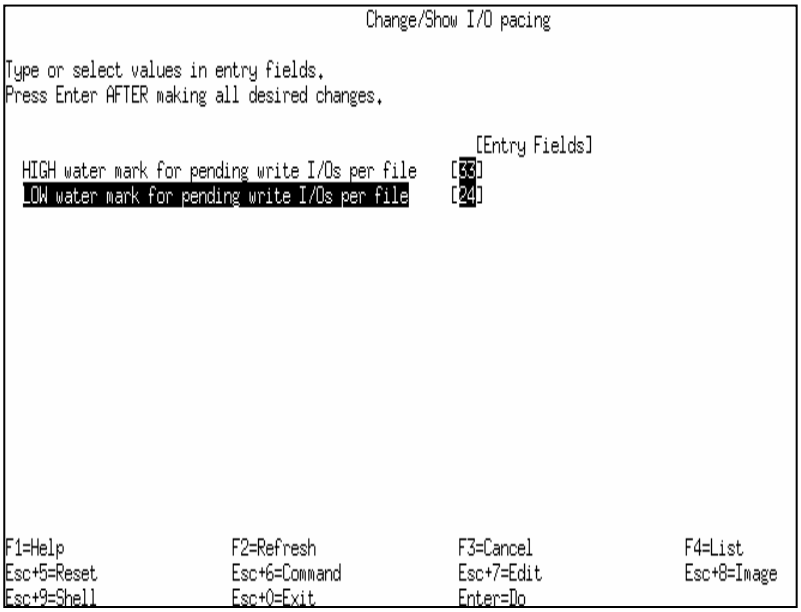
- 약 10초 간격으로 명령을 수행하여 Packet 증가를 확인하면 됨.

아래 절차로 HACMP Serial Heartbeat Network 을 구성합니다.

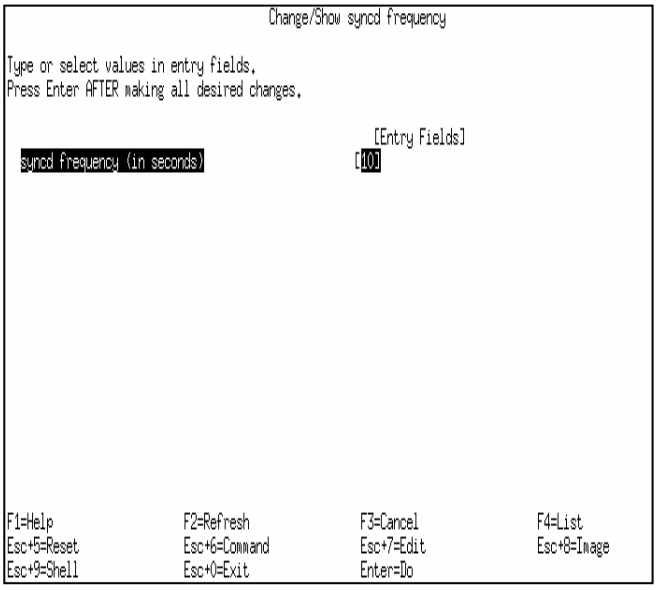
smitty hacmp > Extended Configuration > Extended Topology Configuration > Configure HACMP Communication Interfaces/Devices > Add Communication Interface and Devices > Add Discovered Communication Interface and Devices > Communication Devices

아래 절차로 HACMP tuning parameter 항목을 설정합니다.

smitty hacmp > Extended Configuration > Extended Performance Tuning Parameters Configuration > Change / Show I/O pacing



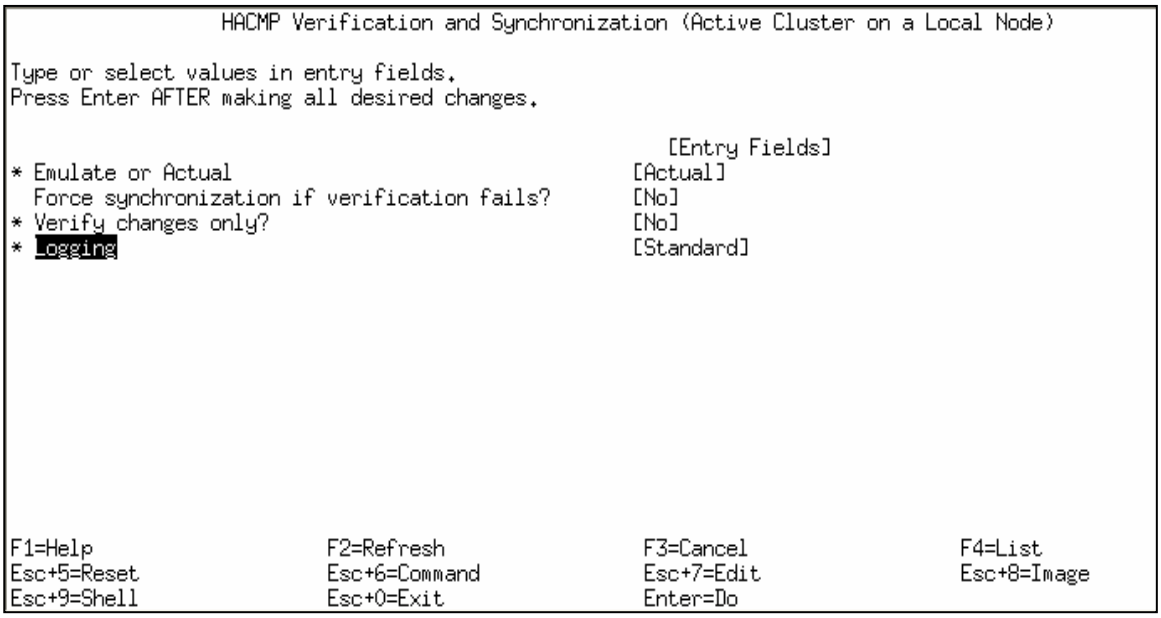
High water mark : 33, Low water mark : 24 로 설정합니다.



Syncd frequency 를 10 으로 변경합니다.

아래 절차로 HACMP 설정을 sync 시킵니다.

smitty hacmp > Extended Configuration > Extended Verification and Synchronization



Default 항목을 유지한 후 엔터를 치면 verify & sync 작업이 진행됩니다.

결과가 OK 가 떨어지면 구성이 정상적으로 완료된 것입니다.

아래 절차로 SNMP community 를 public 이 아닌 다른 값으로 변경합니다.

stopsrc -s snmpd

stopsrc -s clsmuxpdES (if hacmp service is active)

stopsrc -s clinfoES (if hacmp service is active)

- modify /etc/snmp.conf file to add the new community name

vi /etc/snmpd.conf ◊ community hkmc_cluster

- change the default community name used by the clinfoES SRC

/usr/es/sbin/cluster/utilities> ./cl_community_name

<= cluster 에서 사용할 수 있는 community name 확인

chssys -s clinfoES -a "-c hkmc_cluster"

odmget SRCsubsys|grep -p clinfoES <= ODM 에 변경되었는지 확인

startsrc -s snmpd

startsrc -s clsmuxpdES

startsrc -s clinfoES

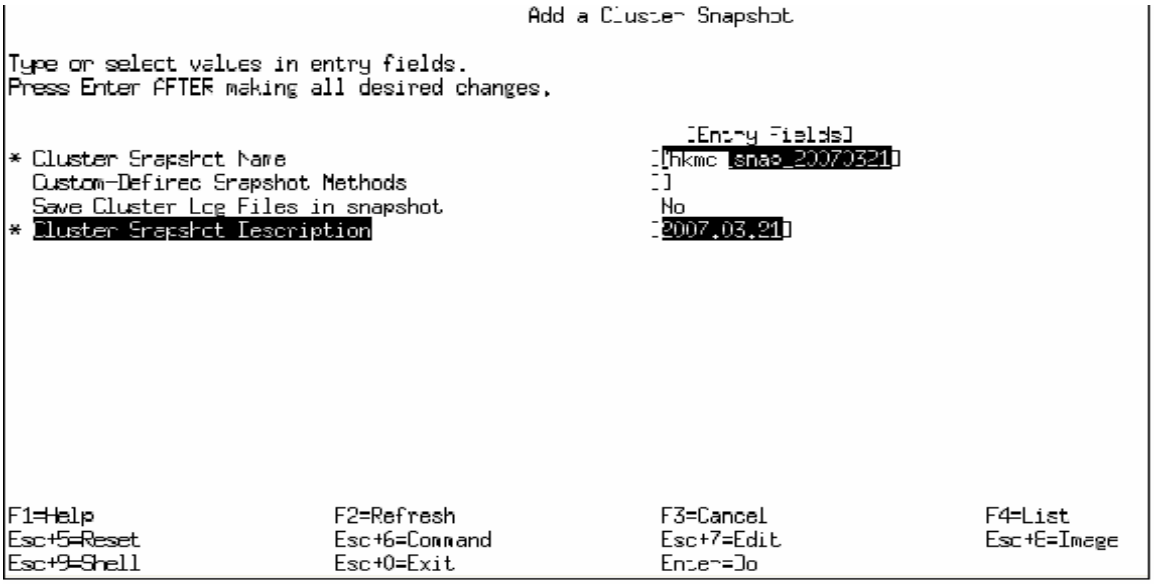
lssrc -Ss clinfoES | awk –F '{print $3}' <= 변경여부 확인

chssys -s hostmibd -a "-c hkmc_cluster“

- 각 서버에서 tail -f /usr/tmp/snmpd.log 파일에서 error 가 없는지 확인하며 /etc/rc.tcpip 파일을 열어서 start /usr/sbin/hostmibd "$src_running“ 항목을 주석 처리하여 hostmibd service 는 사용하지 않는다. hostmibd 는 default 로 public 으로 설정되어 있다.

마지막으로 구성정보를 snapshot 기능을 통해 Backup 합니다. 이로써 HACMP 구성이 완료됩니다. 전체 system 을 한차례 rebooting 시켜줍니다. (booting 이후 persistent ip 및 default gateway 정상여부 확인)

smitty hacmp > Extended Configuration > Snapshot Configuration > Add a Cluster Snapshot



Snapshot name 및 description 을 입력합니다.

/usr/es/sbin/cluster/snapshots/ 디렉토리에 odm, info 확장자로 파일이 생성된다.

16. GPFS

16.1. GPFS Software Install & 환경설정

- GPFS 구성을 위한 환경 설정

export PATH=$PATH:/usr/lpp/mmfs/bin

setclock server1; date ; rsh server2 date

rsh server1 ls ; rsh server2 ls

- GPFS 클러스터 노드의 시간동기화가 절대적으로 요구된다.

setclock server1; data ; rsh server2 date

- /tmp/mmfs: 문제진단 데이터가 수집되는 디렉토리로 200MB 이상이어야 한다.

- # GPFS 네트워크는 IPAT(IP Address Takeover) 불가능, Alias 주소 불가능

- /etc/netsvc.conf 파일

hosts=local, bind

- /etc/hosts 파일을 편집해 준다

GPFS 용 network 는 Private IP 를 이용해 adapter 사이에 cross cable 을 사용해서 연결한다.

GPFS 용 network 는 etherchannel 을 사용해 구성하는 것을 권장한다.

OPS 용 network 도 private IP 를 이용해 direct 연결 후 etherchannel 을 사용해 구성한다.

cat /etc/hosts

SVC IP

12.26.5.138 server1

12.26.5.141 server2

For GPFS network (etherchannel)

7.1.2.1 server1_gpfs

7.1.2.2 server2_gpfs

- .rhost

양쪽의 node 가 rsh & rcp 가 작동하도록 rhosts 를 수정해준다.

cat .rhosts

server1

server2

server1_gpfs

server2_gpfs

- mmfs subsystem

lssrc –s mmfs 로 보여지는 상태는 GPFS v2.3 에서부터는 inoperative 로 보인다.

v2.3 이전 version 에서는 active 상태가 정상이다.

16.2. GPFS Cluster Creation

16.2.1. GPFS Node 정보

- Node

클러스터 내 개별 운영 시스템 이미지

- Node descriptor

GPFS 에서 node 가 어떻게 사용될 지에 대한 설명으로 다음과 같이 표현된다

NodeName:NodeDesignations

NodeName ◊ hostname 또는 IP address

designation ◊ Mananger|Client 및 quorum|nonquorum 설정

quorum: node 가 quorum member 에 추가된다.

manager: node 가 filesystem 을 전담하는 file system manager 로 선택된다.

small system 에는 일반적으로 필요 없으며 large system 에서는 권장된다.

note: filesystem manager 의 지정상태는 mmlsmgr 로 확인할 수 있다.

- Node 등록 예

node name 과 node descriptor 가 포함된 gpfs.allnodes 라는 임의의 file 을 만들어 준다.

#cat gpfs.allnodes

server1_gpfs:quorum

server2_gpfs:quorum

- Node 등록 후 유의사항

node name 은 등록 후에는 변경할 수 없다.

node name 을 변경하기 위해서는 cluster 를 제거하고 다시 구성해야 한다.

/etc/hosts file 에서 node name 을 변경하지 않는다.

/etc/hosts file 에서 node name 을 변경 시에는 node 간의 통신이 이루어지지 않아 cluster 가 동작하지 않는다. 따라서 gpfs filesystem 도 daemon 은 동작하지만 filesystem 은 작동하지 않는다.

node name 의 변경 시에는 exportfs 를 이용한다.

1. mmdelnode 로 모든 node 를 삭제한다.
2. mmexportfs –a –o exportfile_name -> 모든 filesystem 을 export 한다.
3. mmcrcluster 로 cluster 와 node 를 생성한다.
4. mmimportfs –a –I exportfile_name -> 모든 filesystem 을 import 한다.

16.2.2. GPFS Cluster 생성

- GPFS Cluster 생성 및 확인

Server1_gpfs 을 primary 로, server2_gpfs 를 secondary server 로 설정

```
# mmcrcluster -n /home/gpfs/gpfs.allnodes -p server1_gpfs -s server2_gpfs -C gpfs_bmt
# mmlscluster
```

option

-p primary

-s secondary

-C cluster name

GPFS 용 디스크 등록을 위한 description file 생성

```
# cat /home/gpfs/disk.desc
hdisk2:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:1:
hdisk3:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:1:
hdisk4:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:1:
hdisk5:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:2
hdisk6:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:2
hdisk7:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:2
hdisk8:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:3
hdisk9:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:3
hdisk10:server1_gpfs:ibmbmt2_gpfs:dataAndMetadata:3:
```

- GPFS 용 디스크 등록을 위한 description file 생성

디스크 등록을 위한 description file 은 vpath 의 경우에도 동일하다

ex) hdisk2:server1_gpfs:server2_gpfs:dataAndMetadata:1:

vpath 가 생성되지 않은 시점에서 hdisk 로 작업을 한 경우에도 일반 VG 와 동일하게 cfgmgr –v 를 실행했을 때 hdisk 의 정보가 vpath 로 이관된다.

vpath 를 사용하는 disk 에서 hdisk 를 사용하여 작업할 때는 여러 개의 path 중에 오직 한 path 만을 잡아서(예: cfgmgr –vl fcs0) 작업한 후 cfgmgr –v 로 vpath 를 만든다.

note : hdisk2:server1_gpfs:server2_gpfs:dataAndMetadata:1:
hdisk3:server1_gpfs:server2_gpfs:dataAndMetadata:1:
hdisk4:server1_gpfs:server2_gpfs:dataAndMetadata:1:

위의 description file 에서 1 이 의미하는 것은 failure group 이다.

따라서 disk 개수를 quorum 이 만족되어야 한다.

- GPFS 용 디스크인 NDS Disk 생성

```
# mmcrnsd -F /home/gpfs/disk.desc
mmcrnsd: Processing disk hdisk2
mmcrnsd: Processing disk hdisk3
mmcrnsd: Processing disk hdisk4
. . .
```

mmcrnsd: Propagating the changes to all affected nodes.

This is an asynchronous process.

- 생성 NSD Disk 확인

mmlsnsd

File system	Disk name	Primary node	Backup node
(free disk)	gpfs1nsd	server1_gpfs	server2_gpfs
(free disk)	gpfs2nsd	server1_gpfs	server2_gpfs
(free disk)	gpfs3nsd	server1_gpfs	server2_gpfs
(free disk)	gpfs4nsd	server1_gpfs	server2_gpfs
(free disk)	gpfs5nsd	server1_gpfs	server2_gpfs
(free disk)	gpfs6nsd	server1_gpfs	server2_gpfs
(free disk)	gpfs7nsd	server1_gpfs	server2_gpfs

….

- 구성한 GPFS 정보 확인

mmlsconfig

Configuration data for cluster gpfs_test.gp1:

---

clusterName gpfs_bmt.server1_gpfs

clusterId 72342372957413275

clusterType lc

multinode yes

autoload no

useDiskLease yes

maxFeatureLevelAllowed 806

File systems in cluster gpfs_bmt.server1_gpfs:

(none)

- Quorum 을 tiebreakerDisks 로 설정 (3 개의 디스크 할당)

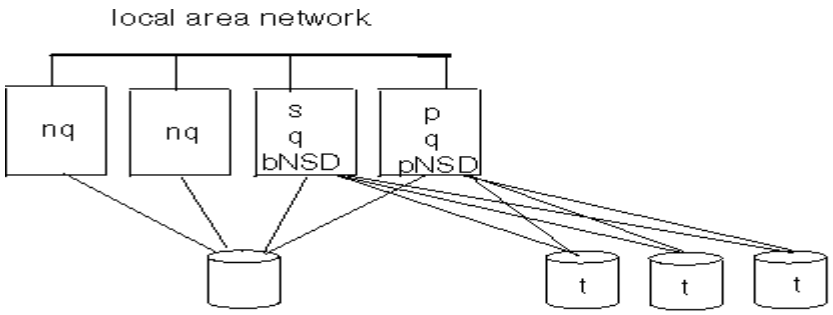
mmchconfig tiebreakerDisks="gpfs1nsd;gpfs4nsd;gpfs7nsd“

Verifying GPFS is stopped on all nodes ...

mmchconfig: Command successfully completed

mmchconfig: Propagating the changes to all affected nodes.

This is an asynchronous process.



p - primary cluster configuration server
s - secondary cluster configuration server
q - quorum node
nq - non-quorum node
pNSD - primary NSD server
bNSD - backup NSD server
t - tiebreaker disk

- 변경된 정보 확인

```
# mmlsconfig
Configuration data for cluster gpfs_bmt.server1_gpfs:
---------------------------------------------
clusterName gpfs_bmt.server1_gpfs
clusterId 72342372957413275
clusterType lc
multinode yes
autoload no
useDiskLease yes
maxFeatureLevelAllowed 806
tiebreakerDisks gpfs1nsd;gpfs4nsd;gpfs7nsd

File systems in cluster gpfs_bmt.server1_gpfs:
-------------------------------------
(none)
```

- GPFS Daemon Startup

```
# mmstartup -a
Fri Apr  8 10:57:40 KORST 2009: mmstartup: Starting GPFS ...

# ps -ef | grep mmfs ; rsh server2_gpfs ps -ef | grep mmfs
    root 270406      1   0 10:50:54   pts/1  0:00 /usr/bin/perl -w /usr/lpp/mmfs/bin/mmgetobjd
6669 0:4:
    root 323592      1   0 11:11:23      -   0:00 /bin/ksh /usr/lpp/mmfs/bin/runmmfs
    root 344140 323592   0 11:11:23      -   0:00 /usr/lpp/mmfs/bin/aix64/mmfsd64
    root 245906 278708   0 11:11:23      -   0:00 /usr/lpp/mmfs/bin/aix64/mmfsd64
```

root 278708 1 0 11:11:22 - 0:00 /bin/ksh /usr/lpp/mmfs/bin/runmmfs

root 282814 1 0 10:50:54 - 0:00 /usr/bin/perl -w /usr/lpp/mmfs/bin/mmgetobjd 6669 0:4:

- GPFS 파일 시스템 구성

각 파일시스템을 구성할 디스크들에 대한 description file 생성

```
# cat disk.desc1
# hdisk3:server1_gpfs:server2_gpfs:dataAndMetadata::gpfs2nsd:::dataAndMetadata:1
# hdisk4:server1_gpfs:server2_gpfs:dataAndMetadata::gpfs3nsd:::dataAndMetadata:1

# cat disk.desc2
# hdisk6:server1_gpfs:server2_gpfs:dataAndMetadata::gpfs5nsd:::dataAndMetadata:2
# hdisk7:server1_gpfs:server2_gpfs:dataAndMetadata::gpfs6nsd:::dataAndMetadata:2

# cat disk.desc3
# hdisk9:server1_gpfs:server2_gpfs:dataAndMetadata::gpfs8nsd:::dataAndMetadata:3
# hdisk10:server1_gpfs:server2_gpfs:dataAndMetadata::gpfs9nsd:::dataAndMetadata:3
```

note : NDS 를 구성할 때 사용했던 description file 이 위와 같은 형식으로 자동으로 수정된다.
이것을 filesystem 을 만들기 위해 필요한 용량만큼 NSD 개수를 조합해 준다.

- 각 Description File 을 이용하여 Filesystem 생성 (본 예에서는 3 개 생성)

```
# mmcrfs /gpfs1 /dev/gpfs1 -F /home/gpfs/disk.desc1 -B 256k -m 1 -M 2 -r 1 -R 2

The following disks of gpfs1 will be formatted on node server1_gpfs:
    gpfs2nsd: size 35507200 KB
    gpfs3nsd: size 35507200 KB
Formatting file system ...
Creating Inode File
Creating Allocation Maps
Clearing Inode Allocation Map
Clearing Block Allocation Map
Flushing Allocation Maps
Disks up to size 118 GB can be added to this file system.
Completed creation of file system /dev/gpfs1.
mmcrfs: Propagating the changes to all affected nodes.
This is an asynchronous process.
```

```
# mmcrfs /gpfs2 /dev/gpfs2 -F /home/gpfs/disk.desc2 -B 256k -m 1 -M 2 -r 1 -R 2
# mmcrfs /gpfs3 /dev/gpfs3 -F /home/gpfs/disk.desc3 -B 256k -m 1 -M 2 -r 1 -R 2
```

note : /gpfs1 과 /dev/gpfs1 은 mount point 가 된다.

- GPFS Cluster Filesystem Mount
- 생성한 GPFS Filesystem 확인 및 mount

```
# lsfs
Name                Nodename   Mount Pt               VFS   Size       Options      Auto
Accounting
/dev/hd4            --         /                      jfs   131072     --           yes  no
/dev/hd1            --         /home                  jfs   131072     --           yes  no
/proc               --         /proc                  procfs --        --           yes  no
/dev/hd10opt        --         /opt                   jfs   262144     --           yes  no
/dev/lv00           --         /source                jfs   10485760 rw             yes  no
/dev/gpfs1          -          /gpfs1                 mmfs  --       rw,mtime,atime,dev=gpfs1 no
no
/dev/gpfs2          -          /gpfs2                 mmfs  --       rw,mtime,atime,dev=gpfs2 no
no
/dev/gpfs3          -          /gpfs3                 mmfs  --       rw,mtime,atime,dev=gpfs3 no
no
```

```
# mount /gpfs1; mount /gpfs2; mount /gpfs3
```

```
# mount
  node          mounted          mounted over     vfs          date           options
-------- ---------------  ---------------  ------ ------------ ---------------
           /dev/hd4          /                    jfs      Apr 07 16:45 rw,log=/dev/hd8
           . . .
           /dev/lv00         /source              jfs      Apr 07 16:46 rw,log=/dev/hd8
           /dev/gpfs1        /gpfs1               mmfs     Apr 08 14:04 rw,mtime,atime,dev=gpfs1
           /dev/gpfs2        /gpfs2               mmfs     Apr 08 14:04 rw,mtime,atime,dev=gpfs2
           /dev/gpfs3        /gpfs3               mmfs     Apr 08 14:04 rw,mtime,atime,dev=gpfs3
```

2.2.11. GPFS Shutdown

- mmshutdown

mmshutdown [-t unmountTimeout] [-a | -W NodeFilename | [-w NodeName[,NodeName...]]

[-n NodeNumber[,NodeNumber...]]]

#mmshutdown -a

Thu Oct 6 14:09:56 KORST 2005: 6027-1341 mmshutdown: Starting force unmount of GPFS file systems

server1_boot1: forced unmount of /gpfs1

server2_boot1: forced unmount of /gpfs1

Thu Oct 6 14:10:01 KORST 2005: 6027-1344 mmshutdown: Shutting down GPFS daemons

server1_boot1: Shutting down!

server1_boot1: 'shutdown' command about to kill process 438418

server2_boot1: Shutting down!

server2_boot1: 'shutdown' command about to kill process 831506

server2_boot1: Master did not clean up; attempting cleanup now

server2_boot1: Thu Oct 6 14:11:01 2005: GPFS: 6027-311 mmfsd64 is shutting down.

server2_boot1: Thu Oct 6 14:11:01 2005: Reason for shutdown: mmfsadm shutdown command timed out

server2_boot1: /var/mmfs/etc/mmfsdown.scr: Subsystem: mmfs Status: inoperative

server2_boot1: /var/mmfs/etc/mmfsdown.scr: /usr/sbin/umount -f -t mmfs

Thu Oct 6 14:11:05 KORST 2005: 6027-1345 mmshutdown: Finished

#mmlsfs gpfs1

mmcommon: 6027-1562 mmlsfs command cannot be executed. Either none of the nodes in the cluster are reachable, or GPFS is down on all of the nodes.

- GPFS Filesystem 유의 사항

- 처음 Storage 에서 host 로 disk 를 할당할 때 size 계산에 유의한다.

- NSD (vpath) 하나가 filesystem 을 만들 수 있는 최소 단위가 된다.

filesystem 의 증가와 축소는 NSD 단위로 이뤄진다.

filesystem 의 dynamic 한 증가작업시에는 filesystem 에 NSD 를 추가시켜 준 후에 작업해야 한다.

filesystem 의 축소 시에도 NSD 단위만큼만 축소시킬 수 있다.

ex) NSD size 가 50G 라면 filesystem 의 최소 size 는 50G 가 된다.

- 작은 file 이 많은 경우 inode 값을 증가시켜 준다.

inode 의 max value 는 103424 이다.

변경된 값은 mmdf 로 확인할 수 있다.

ex) # mmchfs gpfs1 –F 103424

17. OS Migration

17.1. 시스템 OS backup 수행

가) 4mm tape backup

smitty mksysb

나) 시스템 정보 backup

① user 정보: /etc/passwd, /etc/security/passwd

② device 정보: lsdev –Cc adapter, lsdev –Cc disk,

③ vg, pv, lv 정보: lsvg, lspv, lsvg –o, lsvg –o|lsvg –il, lsvg –o|lsvg –ip

④ Filesystem 정보: df –k, /etc/filesystems

⑤ Network 정보: netstat –in, netstat –rn, /etc/hosts

17.2. migration 사전 절차

가) 작업 user 가 root 인지 확인

id ; whoami

uid=0(root) gid=0(system) groups=2(bin),3(sys),7(security),8(cron)

root

나) OS 의 Fileset 정상유무 확인

lppchk –v (output 이 return 안되면 정상임)

다) Rootvg 에 Migration 에 필요한 space 확인

① /tmp : 32MB 의 free space 필요

② / : 10MB 의 free space 필요

③ /usr : 80MB 의 free space

df –kt

lsvg rootvg

라) paging space 크기 확인

① paging space 크기 : 최소 64MB 이상

마) 현재 동작하는 job 확인

① migration 작업 동안 수행될 작업 있는지 확인 (cron)

/var/spool/cron/crontabs 에서 확인

cronadm cron –l

17.3. migration 작업 절차

가) bootlist 변경 및 OS CD 삽입

① bootlist 를 CD rom 으로 변경

bootlist –m normal cd0

bootlist –m normal –o

② OS CD 1 번을 CD-rom 에 삽입

나) System Shutdown & reboot

sync; sync; sync

shutdown –Fr

다) Welcome Screen 은 pass

라) Console 선택

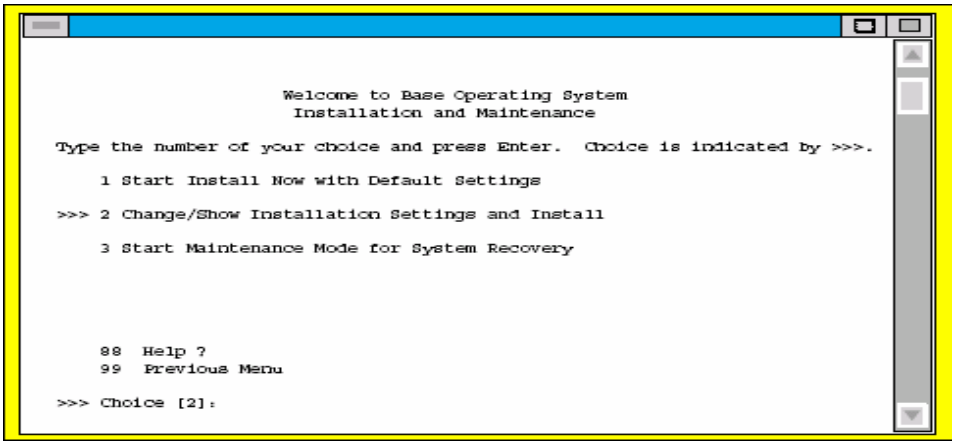
① Console 이 1 개일 경우에는 direct 로 console 에 접속 됨

만약, 여러 개의 Console 이 있을 경우 LED 가 C31 일 때 사용하고자 하는 console 의 숫자를 입력

마) Install 을 위한 language 선택

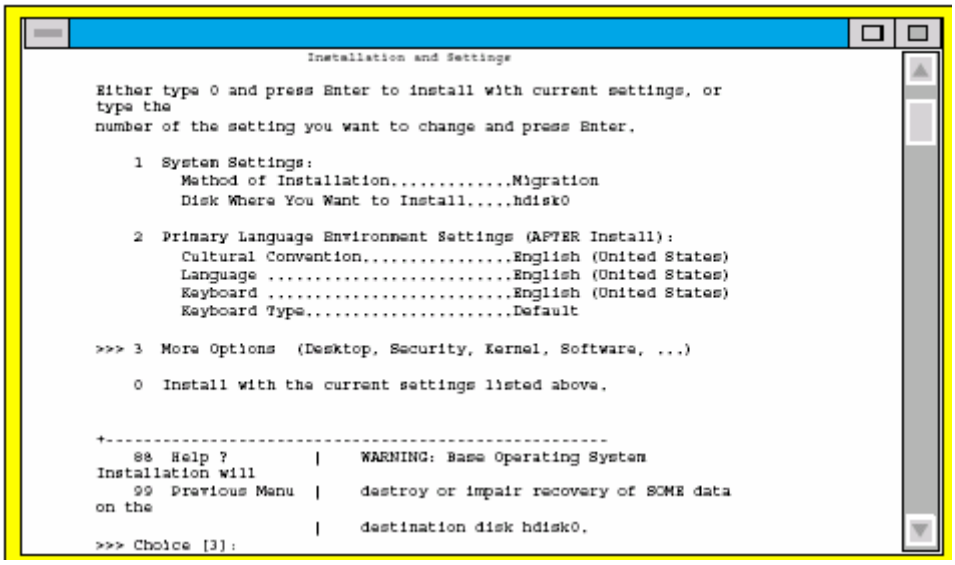
바) Installation and Maintenance 메뉴에서 migration 을 위해서

2 번 Change/Show Installation Settings and Install 을 선택



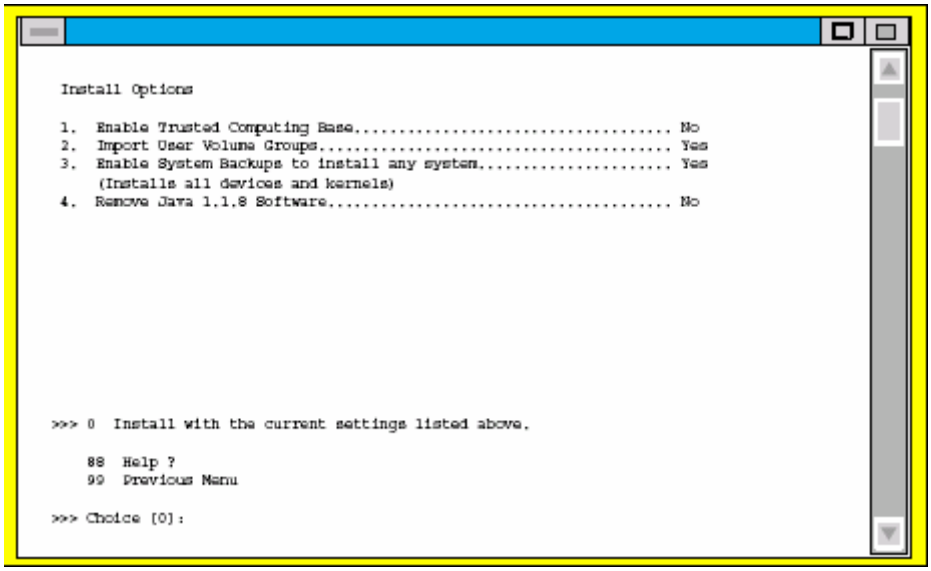
사) Installation and Settings Menu

① Installation and Settings Menu 에서 **3 번 More Options** 를 선택



② Migration 의 “ More Options “ 에서 Option 확인

(각각 작업에 따라 필요 항목 점검 후 Option 결정)



17.4. migration 사후 절차

가) 신규 설치 된 Licensed Program Products 확인

```
# lppchk -l
```

나) OS level 확인

```
# oslevel
```

다) Maintenance Level 확인

```
# instfix –i |grep ML
```

라) 시스템 Fileset 이상 유무 확인

```
# lslpp –l |grep BROKEN
```

18. NTP (Network Time Protocol) Server

18.1. SETTING UP AN NTP SERVER

18.1.1. Edit the /etc/ntp.conf file

File should read:

server 127.127.1.0 prefer
driftfile /etc/ntp.drift
tracefile /etc/ntp.trace

save file and exit.

18.1.2. smitty xntpd-->start at both system restart and now

Let run for approximately 10 minutes before moving to step 4. Otherwise, stratum will show 16.

18.1.3. lssrc -ls xntpd

Stratum should show between 3-4 but ok as long as it doesn't show 16.
Server is ok, now to configure client.

18.2. SETTING UP AN NTP CLIENT

18.2.1. Edit the /etc/ntp.conf file

File should read:
server <ipaddress of ntp server> (example: server 1.2.3.4 if 1.2.3.4 is the NTP server)
driftfile /etc/ntp.drift
tracefile /etc/ntp.trace
save file and exit

18.2.2. At command line enter

ntpdate <ipaddress of ntp server>
If it doesn't say "no server suitable for synchronization found", go to next step.

18.2.3. smitty xntpd-->start at both system restart and now

Let run for approximately 10 minutes or so before going on to step 4.
Otherwise stratum may show 16.

18.2.4. lssrc -ls xntpd

Stratum will probably show 4-5. If it doesn't its ok as long as it doesn't show 16.

Clocks should now be in sync. Repeat the client steps for each of your clients that you need to setup.

19. Openssh

19.1. openssl 설치 (from AIX Toolbox for Linux)

```
# rpm -i openssl-0.9.7d-2.aix5.1.ppc.rpm
# rpm -i openssl-devel-0.9.7d-2.aix5.1.ppc.rpm
# rpm -i openssl-doc-0.9.7d-2.aix5.1.ppc.rpm
    -- or --
# smitty installp
```

19.2. openssl 설치 (from AIX Expansion Pack)

```
# smitty installp
# lslpp -l | grep ssh
  openssh.base.client       4.1.0.5300   COMMITTED
       Open Secure Shell Commands
  openssh.base.server       4.1.0.5300   COMMITTED
       Open Secure Shell Server
  openssh.license           4.1.0.5300   COMMITTED
       Open Secure Shell License
  openssh.base.client       4.1.0.5300   COMMITTED
       Open Secure Shell Commands
  openssh.base.server       4.1.0.5300   COMMITTED
       Open Secure Shell Server
```

19.3. Creating the key on each node

```
# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (//.ssh/id_rsa):   Enter
Created directory '//.ssh'.
Enter passphrase (empty for no passphrase):   Enter
Enter same passphrase again:                  Enter
Your identification has been saved in //.ssh/id_rsa.
Your public key has been saved in //.ssh/id_rsa.pub.
The key fingerprint is:
18:d3:70:5c:f5:8e:1a:41:15:33:c8:35:93:4f:ae:b8 root@m80
# cd /.ssh
# ls -l
total 16
```

```
-rw-------   1 root     system          887 Feb 16 12:19 id_rsa
-rw-r--r--   1 root     system          218 Feb 16 12:19 id_rsa.pub
```

19.4. Using a shell script for automatic login

```
# vi setremotekey
OLDDIR=`pwd`;
if [ -z "$1" ]; then
      echo Need user@host info;
      exit;
fi;
cd $HOME;
if [ -e "./.ssh/id_rsa.pub" ]; then
      cat ./.ssh/id_rsa.pub | ssh $1 'cat >> .ssh/authorized_keys';
else
      ssh-keygen -t rsa;
      cat ./.ssh/id_rsa.pub | ssh $1 'cat >> .ssh/authorized_keys';
fi;
cd $OLDDIR
```

```
$ setremotekey root@p610
The authenticity of host 'p610 (192.168.1.111)' can't be established.
RSA key fingerprint is 9f:e9:6b:d2:5a:d2:d9:ef:f7:83:d9:bf:7f:62:d0:ee.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'p610,192.168.1.111' (RSA) to the list of known hosts.
root@p610's password:   <p610의 root password를 입력>
```

```
=> 반대편 노드(p610)에서 m80 노드에 대해서 수행
p610:/.ssh#   setremotekey root@m80
The authenticity of host 'm80 (192.168.1.112)' can't be established.
RSA key fingerprint is 9f:e9:6b:d2:5a:d2:d9:ef:f7:83:d9:bf:7f:62:d0:ee.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'm80,192.168.1.112' (RSA) to the list of known hosts.
root@m80's password: <m80의 root password를 입력>
```

*** GPFS 구성 시 유의사항**

```
 # mmcrcluster   "   -r /bin/ssh -R /bin/scp "
```

=> setremotekey 명령어 사용시 GPFS interface name 을 hostname 으로 사용한다.

(# setremotekey root@m80_gpfs)

=> setremotekey 명령어는 자기 자신에게도 수행한다.

(on m80 # setremotekey root@m80_gpfs)

20. TCP_Wrapper

20.1. TCP_Wrapper 다운로드

http://www.bullfreeware.com/사이트에서 버전에 맞는 Tcp_wrapper 를 다운로드 합니다.
이 사이트에 AIX 에 많은 파일셋과 freeware 를 다운 받을 수 있습니다.
다운 받은 파일은 tcp_wrappers-7.6.1.0.exe 입니다.
그리고 서버로 전송하여 ./tcp_wrappers-7.6.1.0 .exe 파일을 실행시키면, tcp_wrappers-7.6.1.0.bff 파일이 생겨납니다.
.toc 파일은 지워주셔야 정상적인 설치가 가능합니다.
다음, 그 디렉토리에서 smitty installp 실행 후 [사용 가능한 소프트웨어에서 설치 및 갱신] 선택
옵션은 License 를 YES 로 바꿔 줍니다.

20.2. TCP_Wrapper 구성

일단 telnet, ftp 를 참조하는 파일인 /etc/inetd.conf 파일을 수정합니다.

```
# vi  inetd.conf

  #ftp     stream  tcp6    nowait  root    /usr/sbin/ftpd          ftpd
  #telnet  stream  tcp6    nowait  root    /usr/sbin/telnetd       telnetd -a
  ftp      stream  tcp6    nowait  root    /usr/local/bin/tcpd     ftpd
  telnet   stream  tcp6    nowait  root    /usr/local/bin/tcpd     telnetd
```

수정이 끝났으면, /etc 밑에 hosts.allow hosts.deny 파일을 만들어 접근 허용할 포트와 거부할 포트를 지정합니다.

```
# touch hosts.allow
# touch hosts.deny
```

일단 고객이 모든 포트에 대하여 막기를 원하기 때문에 hosts.deny 파일에는 모든 포트를 거부합니다.

```
# vi hosts.deny
 ALL:ALL
```

ex)
그리고 hosts.allow 파일에서는 고객이 원하는 111.1. 으로 시작하는 포트는 열어 주어야 하며, 또한 나머지 서너 군데를 더 열어 두어야 합니다.

```
 # vi hosts.allow
 ALL:LOCAL: 111.1.
 ALL:10.60.*.*
```

ALL:10.60.*.*

ALL:10.60.*.*

기본 설정이 끝나고, inetd 데몬을 refresh 시켜주면 됩니다.

lssrc -s inetd

refresh -s inetd <- /etc/inetd.conf 파일 수정 시 항상 refresh 시켜야 적용이 된다.

하지만 hosts.allow hosts.deny 수정 시에는 refresh 시키지 않아도 적용이 됩니다.

마지막으로 확인: 다른 IP 대에서 이 서버로 접속할 경우, telnet 이 막혀있는지 확인

21. WAS

21.1. Stand-alone Application Server 시작과 종료

시작

- <WAS_HOME>/profiles/Node_home/bin 디렉토리로 이동, startServer.sh 명령을 치면, 중지된 Application Server 를 시작할 수 있다.
- 만약, 보안설정이 되어 있는 경우에는 <WAS_HOME>/profiles/Node_home/bin/startServer.sh <server_name> -username <username> -password <user's passwod>

종료

- <WAS_HOME>/profiles/Node_home/bin 디렉토리로 이동, stopServer.sh 명령을 치면, Application Server 를 중지시킬 수 있다.
- 보안설정이 되어 있는 경우에는 <WAS_HOME>/profiles/Node_home/bin/stopServer.sh <server_name> -username <username> -password <user's passwod

상태 확인

- <WAS_HOME>/profiles/Node_home/bin 디렉토리로 이동, serverStatus.sh 명령을 치면, 현재 Process 의 상태를 확인 할 수 있다.
- 보안설정이 되어 있는 경우에는 <WAS_HOME>/profiles/Node_home/bin/serverStatus.sh <server_name> -username <username> -password <user's passwod>

21.2. Network Deployment Cell 시작과 종료

Deployment Manager (DMgr) Start/Stop Command

- <WAS_HOME>/profiles/DMgr_home/bin 디렉토리로 이동, startManager.sh OR stopManager.sh 명령을 치면, 중지된 DMgr process 를 시작, 정지 할 수 있다.
- 보안설정이 되어 있는 경우에는 <WAS_HOME>/profiles/<DMGR_HOME>/bin/startManager.sh OR stopManager –username <username> -password <user's passwod>

Node Agent Start/Stop Command

- <WAS_HOME>/profiles/Node_home/bin 디렉토리로 이동, startNode.sh OR stopNode.sh 명령을 치면, 중지된 nodeagent process 를 시작, 정지 할 수 있다.
- 보안설정이 되어 있는 경우에는 <WAS_HOME>/profiles/<NODE_HOME>/bin/startNode.sh OR stopNode -username <username> -password <user's passwod>
- [중요] 반드시 Application Server process 를 시작하기 전에 Node Agent (nodeagent) process 가 시작 되어 있어야 한다. Deployment Manager (DMgr) process 는 상관 없음.

Application Server Start/Stop Command

- <WAS_HOME>/profiles/<NODE_HOME>/bin 디렉토리로 이, startServer.sh OR stopServer.sh 명령을 치면, 중지된 Application Server process 를 시작, 정지 할 수 있다.

- 보안설정이 되어 있는 경우에는 <WAS_HOME>/profiles/<DMGR_HOME>/bin/startServer.sh <Server_Name> OR

stopServer.sh <Server_Name> -username <username> -password <user's passwod>

22. Apache

22.1. Apache rpm 설치 방법

Apache를 다운 받을 수 있는 곳...(AIX Toolbox for Linux Applications)

http://www-03.ibm.com/servers/aix/products/aixos/linux/download.html (AIX Toolbox for Linux Applications CD에도 있습니다. 설치 file들)

This software is incompatible with some earlier versions. If you have rpm.rte installed at a level less than 3.0.5.20, then you will need to reinstall that software in order to continue. See the AIX Toolbox for Linux Applications ReadMe for further information.

- 이 내용은 OS에 rpm.rte라는 File Set이 존재해야 rpm 명령어를 사용할 수 있다는 내용입니다.

다운로드 대상 (아래 순서대로 설치하세요)

1. rpm.rte

2. expat-1.95.7-4.aix5.1.ppc.rpm (libexpat.a(libexpat.so.0) 관련 설치 파일)

3. apache-1.3.31-1.aix5.1.ppc.rpm

**참고로 smitty installp로도 rpm 설치가 가능합니다.

#lslpp -l|grep rpm (먼저 **rpm File을 설치되어 있는 지 확인**합니다.)

```
  rpm.rte                  3.0.5.39  COMMITTED  RPM Package Manager
  rpm.rte                  3.0.5.39  COMMITTED  RPM Package Manager
# ls -lrt
total 6808
-rw-r-----   1 hjcidc   ibmadmin      629604 Mar 28 09:03 apache-1.3.31-1.aix5.1.ppc.rpm
-rw-r-----   1 hjcidc   ibmadmin     2210816 Mar 28 09:03 rpm.rte
-rw-r-----   1 hjcidc   ibmadmin      459599 Mar 28 09:16 expat-1.95.7-4.aix5.1.ppc.rpm
```

#rpm -Uhv ./apache-1.3.31-1.aix5.1.ppc.rpm

```
error: failed dependencies:
        libexpat.a(libexpat.so.0) is needed by apache-1.3.31-1
```

--> libexpat.a(libexpat.so.0) 먼저 설치 필요

#rpm -i ./expat-1.95.7-4.aix5.1.ppc.rpm

#rpm -Uhv apache-1.3.31-1.aix5.1.ppc.rpm -> 그리고 **Apache 설치**

```
apache                  ##################################################
```

22.2. Apache 시작과 중지

서버 Root

/opt/freeware/apache

데몬 위치

/opt/freeware/apache/sbin/apachectl

-> 여기서 apache를 올리고 내리고 할 수 있습니다.

/opt/freeware/apache/sbin/**apachectl stop (시작)**

/opt/freeware/apache/sbin/**apachectl stop (중지)**

/opt/freeware/apache/sbin/**apachectl restart (재시작)**

Configuration file 위치

(기존 설정 내용 중 contents를 아래 httpd.conf 파일에 설정을 해주면 Apache 서버가 동작합니다.)

/etc/opt/freeware/apache/httpd.conf

23. DBMS

23.1. DB2 시작과 종료

DB2 Start

- DB2 Instance User-ID로 Login 한다.

login db2inst1

- CLP에서 Start Command 실행한다.

/home/db2inst1 >> db2start

DB2 Stop

- CLP에서 Stop Command 실행한다.

/home/db2inst1 >> db2stop

- DBMS Stop시 처리사항

1. Application 및 DB2에서 Database에 Connect시 마다 Agent 하나씩 기동되는바,
2. 처리중인 Agent가 존재할 경우 db2stop 명령이 실행되지 않음.
3. 이때에는 기동중인 Agent를 강제 중지 시킨 후 다시 db2stop 명령으로 Stop 시킴
   => /home/db2inst1 >> db2 force application all 또는
   => /home/db2inst1 >> db2 force application (Agent-ID)
   => /home/db2inst1 >> db2stop
4. Force 명령 실행 시 주의사항
   가. Force 명령으로 Application Server의 Agent를 종료했을 경우
   CICS측면에서의 Application Server는 계속 살아있으면서 Database와 Connect 상태로 간주하고 있으나 연결을 차단된 상태임
   나. 이때에는 CICS도 정상종료 후 다시 Start 해야 함.

DB2 실행 확인

- DB2 UDB가 실행 중인지 확인하려면, 아래의 프로세스들이 수행 중인지 확인한다.

ps -ef | grep db2

```
db2inst1  8866 14458   0  21:27:45      -  0:00 db2ipccm
    root 13788     1   0  21:27:45      -  0:00 db2wdog
db2inst1 14458 13788   0  21:27:45      -  0:00 db2sysc
db2inst1 15486 15742   0  21:27:46      -  0:00 db2resyn
db2inst1 15742 14458   0  21:27:45      -  0:00 db2gds
db2inst1 16256 14458   0  21:27:45      -  0:00 db2tcpcm
db2inst1 16514 14458   0  21:27:45      -  0:00 db2tcpcm
db2inst1 17030 15742   0  21:27:46      -  0:00 db2spmlw
db2inst1 17288 14458   0  21:27:46      -  0:00 db2spmrm
```

db2inst1 17546 15742 0 21:27:46 - 0:00 db2srvlst

db2inst1 18672 18084 0 21:30:46 pts/1 0:00 grep db2

23.2. 오라클 시작과 종료

기동중인 리스너 STOP 및 오라클 SHUTDOWN

- 먼저 리스너를 STOP 시키고 DATABASE를 shutdown 해 야함.

- oracle 계정으로 로그인

#su - oracle

Passwd:oracle

- 리스너 중지

$lsnrctl stop

- 오라클 접속

$sqlplus /nolog

- sysdba 권한으로 connect

SQL>connect /as sysdba

- DATABASE SHUTDOWN

SQL>shutdown

DATABASE STARTUP 및 리스너 START

중지절차의 역순. 먼저 DATABASE를 기동하고 리스너를 START 한다.

- 오라클 접속

$sqlplus /nolog

- sysdba 권한으로 connect

SQL>connect /as sysdba

- DATABASE STARTUP

SQL>startup

- 리스너 START

SQL>exit

$lsnrctl start

24. Storage

24.1. 사전 참고 사항

- 일반적으로 외장스토리지는 서버의 FC카드를 통해 SAN 스위치를 거쳐 외장 스토리지와 연결된다 즉, 서버의 FC카드의 개수에 따라 AIX에서 보는 디스크개수는 FC카드개수 X 디스크개수로 보인다. 쉽게 말하면, FC카드가 2장이고, 스토리지에서 할당해준 LUN [디스크,(ex) hdisk, hdiskpower, dlmfdrv...)이 4개면, hdisk는 8개로 보인다.
- 보통 IBM, EMC, Hitachi 디스크는 Load Balancing 을 해주는 s/w를 제공하며 IBM 대형급은 vpath, IBM DS스토리지군은 RDAC, EMC는 powerpath , Hitachi는 HDLM(Hitachi Dynamic Link Manager)을 제공한다.

24.2. IBM vpath, RDAC 명령어

lsvpcfg vpath로 구성된 디스크가 어떤 hdisk와 매핑되어있는지 확인

[결과예제]

vpath0 (Avail pv XXXvg) 75890016000 = hdisk3 (Avail) hdisk19 (Avail) hdisk35 (Avail) hdisk51 (Avail)

위 예제를 보면 FC카드가 4장 있다는 것을 알 수 있다.

datapath query adapter -> FC어댑터 별 LUN을 볼 수 있음

[결과예제]

datapath query adapter

Active Adapters :4

Adpt#	Name	State	Mode	Select	Errors	Paths	Active
0	fscsi0	NORMAL	ACTIVE	70834	0	18	0
1	fscsi2	NORMAL	ACTIVE	71555	0	18	0
2	fscsi3	NORMAL	ACTIVE	70753	0	18	0
3	fscsi1	NORMAL	ACTIVE	70432	0	18	0

위 예제는 FC카드가 4장이고, 각 카드당 18개씩 vpath 디스크가 있다는 것을 알 수 있다.

장애가 있으면 Errors에 카운트됨.

datapath query device -> single SDD vpath device 또는 all SDD vpath devices에 대한 정보 제공

(참고1) lslpp -l devices.sdd*

```
  Fileset                      Level  State      Description
 ----------------------------------------------------------------------------
  Path: /usr/lib/objrepos
  devices.sdd.53.rte           1.6.1.0  COMMITTED  IBM Subsystem Device Driver for AIX V53
  Path: /etc/objrepos
  devices.sdd.53.rte           1.6.1.0  COMMITTED  IBM Subsystem Device Driver for AIX V53
```

fget_config -l dar0 -> 컨트롤러별 디스크할당 보여줌. dac0 , dac1 ...

24.3. EMC powerpath 명령어

- powermt display -> FC당 disk 정보 보기(EMC disk정보를 표시)
- powermt display dev=all -> 모든 EMC hdisk와 hdiskpower의 mapping정보를 표시

  사용 예) powermt display dev=hdiskpower2
- sympd list -> 물리적인 LUN 구성 정보를 보여줌

아래 명령어 실행금지(EMC엔지니어가 수행하는 명령임/상황 별)

- powercf -q -> EMC disk 정보를 갱신
- powermt config -> EMC disk 환경을 전환 (예: hdisk pvid를 hdiskpower pvid로 전환)

[결과예제]

```
# powermt display
Symmetrix logical device count=16
CLARiiON logical device count=0
==============================================================================
----- Host Bus Adapters ---------  ------ I/O Paths -----  ------ Stats ------
###  HW Path                           Summary   Total   Dead  IO/Sec Q-IOs Errors
==============================================================================
    0 fscsi0                         optimal      13      0      -      0      0
    1 fscsi2                         optimal      13      0      -      0     13

[LGDAMRT1] root:/> powermt display dev=all
Pseudo name=hdiskpower0
Symmetrix ID=000292600013
Logical device ID=0668
state=alive; policy=SymmOpt; priority=0; queued-IOs=0
==============================================================================
---------------- Host ----------------   - Stor -   -- I/O Path -  -- Stats ---
###  HW Path                  I/O Paths    Interf.   Mode    State  Q-IOs Errors
==============================================================================
    1 fscsi2                  hdisk17     FA  9hB   active  alive      0      1
    0 fscsi0                  hdisk4      FA  7hB   active  alive      0      0

Pseudo name=hdiskpower1
Symmetrix ID=000292600013
Logical device ID=0669
state=alive; policy=SymmOpt; priority=0; queued-IOs=0
```

```
==============================================================================
---------- Host ----------     - Stor -    -- I/O Path -   -- Stats ---
###   HW Path                  I/O Paths     Interf.   Mode     State   Q-IOs Errors
==============================================================================
    1 fscsi2                   hdisk18     FA  9hB    active   alive       0       1
    0 fscsi0                   hdisk5      FA  7hB    active   alive       0       0
…생략…

[결과예제]
#sympd list
Symmetrix ID: 000292600013
          Device Name             Directors                    Device
-------------------------- ------------ ------------------------------------
                                                                          Cap
Physical                   Sym   SA :P DA :IT   Config         Attribute    Sts   (MB)
-------------------------- ------------ ------------------------------------
/dev/rhdiskpower0          0668 07H:1 07C:D1   RAID-5         N/Grp'd      RW    15000
/dev/rhdiskpower1          0669 07H:1 08C:D0   RAID-5         N/Grp'd      RW    15000
/dev/rhdiskpower2          066A 07H:1 07C:C0   RAID-5         N/Grp'd      RW    15000
/dev/rhdiskpower3          066B 07H:1 08C:C3   RAID-5         N/Grp'd      RW    15000
/dev/rhdiskpower4          066C 07H:1 07C:D3   RAID-5         N/Grp'd      RW    15000
…생략…

(참고2)
# lslpp -l|grep EMC
EMC.Symmetrix.aix.rte        5.3.0.3   COMMITTED   EMC Symmetrix AIX Support
EMC.Symmetrix.fcp.rte        5.3.0.3   COMMITTED   EMC Symmetrix FCP Support
EMCpower.base                5.3.0.0   COMMITTED   PowerPath Base Driver and
EMCpower.consistency_grp     5.3.0.0   COMMITTED   PowerPath Consistency Group
EMCpower.encryption          5.3.0.0   COMMITTED   PowerPath Encryption with RSA
EMCpower.migration_enabler
EMCpower.mpx                 5.3.0.0   COMMITTED   PowerPath Multi_Pathing
SYMCLI.STAR_PERL.rte         7.0.0.0   COMMITTED   EMC Solutions Enabler Star
SYMCLI.SYMRECOVER.rte        7.0.0.0   COMMITTED   EMC Solutions Enabler
EMC.Symmetrix.aix.rte        5.3.0.3   COMMITTED   EMC Symmetrix AIX Support
EMC.Symmetrix.fcp.rte        5.3.0.3   COMMITTED   EMC Symmetrix FCP Support
#
```

24.4. Hitachi HDML 명령어

- dlnkmgr view –path -> path 정보 보여줌

[결과예제]

Paths:000020 OnlinePaths:000020

PathStatus IO-Count IO-Errors

Online 453194683 0

PathID PathName DskName iLU

ChaPort Status Type IO-Count IO-Errors DNum HDevName

000000 08.08.0000000000011F00.0000 HITACHI .DF600F .77010291 0006

1D Online Non 0 0 0 dlmfdrv0

000001 08.0C.0000000000021F00.0000 HITACHI .DF600F .77010291 0006

0D Online Own 84472518 0 0 dlmfdrv0

.....

- dlnkmgr view –drv -> hdisk와의 매핑정보를 볼 수 있다.

[결과 예제]

PathID HDevName Device LDEV

000000 dlmfdrv0 hdisk2 AMS.XXX.0006

000001 dlmfdrv0 hdisk22 AMS.XXX.0006

000002 dlmfdrv1 hdisk3 AMS.XXX.0007

000003 dlmfdrv1 hdisk23 AMS.XXX.0007

.....

- dlnkmgr view –lu -> hdisk과의 매핑정보 및 디스크 상태를 확인할 수 있다.

Product : AMS

SerialNumber : XXXXXX

LUs : 10

iLU HDevName Device PathID Status

0006 dlmfdrv0 hdisk2 000000 Online

hdisk22 000001 Online

0007 dlmfdrv1 hdisk3 000002 Online

hdisk23 000003 Online

.....

25. 소프트웨어 라이프 사이클

25.1. AIX

Release	Original Release	Withdraw from Marketing	Withdraw from Service	End of Fee-based Service Extension	End of Web-based Support (End of Life)	Last Updated
AIX 7.1	TBA Plan 10/2010	TBA assume 9/30/2015	**TBA** assume 9/30/2016	Not announced will require latest TL (TBD) assume 9/30/2016 - 9/30/2019	Not announced assume 09/30/2020	
AIX 6.1 **5765-G62**	11/9/2007	TBA assume 9/30/2012	**TBA** assume 9/30/2013	Not announced will require latest TL (TBD) assume 9/30/2013 - 9/30/2016	Not announced assume 09/30/2017	12-Aug-2008
AIX 5.3 **5765-G03**	8/13/2004	TBA assume 4/30/2011	**TBA** assume 4/30/2012	Not announced will require latest TL (assume 5300-12) assume 5/01/2012 - 9/30/2014	Not announced assume 09/30/2015	23-Dec-2008
AIX 5.2 **5765-E62**	10/18/2002	7/8/2008 withdraw from SPO	**04/30/2009**	05/01/2009-04/30/2012 Will require 5200-10 with latest Service Pack	04/30/2013	12-Aug-2008
AIX 5.1 **5765-E61**	05/04/2001	04/30/2005	**04/01/2006**	04/01/2006 - 04/01/2009 requires 51 ML9	04/01/2010	25-Mar-2008
AIX 4.3	10/31/1997	6/30/2003	**12/31/2003**	12/31/2004 - 12/31/2006 Extended -- 12/31/2008 with new defect for limited customers Requires 433 ML11 DST 2007 NOT available as interim fix	12/31/2008	

25.2. HACMP

Release	Original Release	Withdraw from Marketing	Withdraw from Service	Last Updated
HACMP 5.5 5765-F62	14/11/2008			05-Dec-2008
HACMP 5.4 5765-F62	09/11/2007			05-Dec-2008
HACMP 5.3 5765-F62	12/8/2005	30/9/2008	30/9/2009	25-Mar-2008
HACMP 5.2 5765-F62	13/7/2004	30/4/2006	30/9/2007	25-Mar-2008
HACMP 5.1 5765-F62	11/7/2003	31/3/2005	01/9/2006	06-Dec-2006
HACMP 4.5 5765-E54	12/7/2002	31/12/2003	01/9/2005	25-Mar-2008
HACMP 4.4 5765-E54	23/6/2000	20/6/2003	31/12/2003	06-Dec-2006

25.3. JAVA

Release	Withdraw from Service
Java 6 64-bit	**30-Nov-17**
Java 6 32-bit	**30-Nov-17**
Java 5 64-bit	**30-Sep-12**
Java 5 32-bit	**30-Sep-12**
Java 1.4.2 64-bit	**30-Sep-11**
Java 1.4.2 32-bit	**30-Sep-11**
Java 1.3.1 64-bit	**30-Sep-07**
Java 1.3.1 32-bit	**30-Sep-07**