

Windows NT 운영체제 커널에서 Disk mirroring 기능 설계

김성관*, 장승주*, 지동해**, 김학영**, 이정배***, 노영욱****

(sjjang@hyomin.dongueui.ac.kr)

동의대학교 컴퓨터공학과*

한국전자통신연구원 프로그래밍 환경 연구팀**

부산외국어대학교 컴퓨터공학과***

신라대학교 컴퓨터교육과****

The design of disk mirroring function in Windows NT kernel

Kim Sung-Kwan*, Jang Seung-Ju*, Gi Dong-Hea**, Kim Hak-Young**,

Lee Jung-Bea***, Lho Young-UK****

Dongueui Univ., Dept. of Computer Engineering*

ETRI, Programming Language Section**

PUF, Dept. of Computer Engineering***

Silla Univ., Dept. of Computer Education****

요 약

Disk mirroring 시스템은 안정성과 고 가용성(High Availability)을 제공하기 때문에 고성능 시스템에서 많이 사용되고 있다. 본 논문에서는 Windows NT 커널에 remote node 에 대한 disk mirroring 기능을 추가하여 시스템 자체에서 이러한 기능을 제공할 수 있는 모델을 설계하고자 한다. 설계 시스템은 Windows NT TDI 계층을 사용하여 network 으로 연결된 두 대의 시스템이 master node 와 slave node 의 pair 로 구성되며, slave node 가 master node 의 디스크를 mirror 하게 된다. 또한 fault tolerance 기능을 제공하여 node 의 장애 발생 시 log 를 생성하고, 이 정보를 바탕으로 장애가 발생한 node 에 대하여 데이터 복구 기능을 제공하도록 한다.

1. 서론

디스크 저장 기술의 발전으로 최근에는 고용량의 데이터를 저장해야 하는 서버가 급격히 증가하고 있다. 또한 network 기술과 환경의 발전으로 network 을 통한 데이터 교환이 보편적으로 사용되고 있다. 이러한 고용량의 데이터를 저장함에 있어 데이터의 안정성과 디스크 매체의 고 효율성이 요구되고 있는 상황이다. 디스크 입출력의 성능을 향상시키기 위한 기술로서 다수의 디스크를 사용하여 디스크 입출력을 서비스하는 여러 종류의 디스크 배열(disk array) 시스템과 각 디스크에 저장할 데이터를 할당하는 정책 등이 활발히 연구, 개발되어 사용되고 있다[6,7].

최근 PC 성능의 발전으로 UNIX 계열의 운영체제가 주로 사용되던 서버시장에서 Windows NT 의 사용이 점차 증대되고 있다. Windows NT 는 micro kernel 을 기반으로 한 내부구조를 가지고 있으며, system service 기능은 각 기능을 담당하는 manager 들이 수행한다[1]. Windows NT 의 커널 부분은 잘 정의된 계층구조를 가지고 있다. 대신 Kernel source 가 공개된 linux 등과는 다르게 kernel 에 접근하기 위해서는 kernel mode driver 의 형태로 계층구조에 포함시켜야 한다[10].

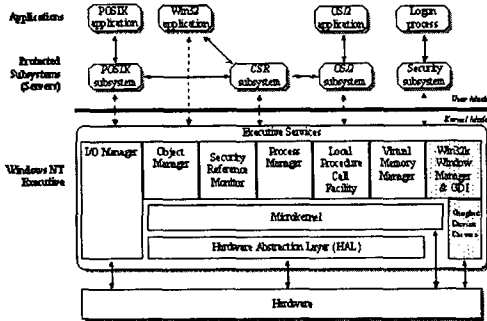
디스크 저장 기술에서 mirroring 기법은 높은 신뢰도와 고 효율성을 제공하기 때문에 고성능 디스크 시스템에서 많이 사용되고 있다. 이러한 disk mirroring 기능과 fault tolerance 기능을 Windows NT

커널 모드 자체에 추가하여 network 을 통하여 pair 로 구성된 mirror set 시스템을 설계하여 안정성과 고 가용성을 보장할 수 있는 모델을 제시한다[8].

본 논문의 구성은 2 장에서 관련연구 사항을 언급 하고, 3 장에서 설계 모델을 정의한다. 4 장에서 결론 과 고려사항의 언급으로 마무리한다.

2. 관련연구

2.1 Windows NT Microkernel 구조



[그림 1] Windows NT 4.0 Architecture

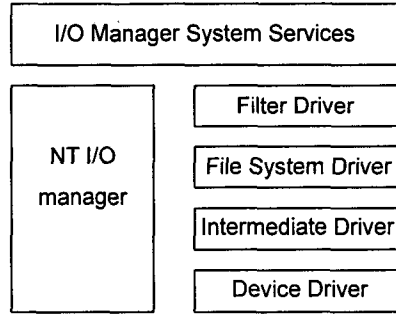
Windows NT 의 커널 모드 구조는 [그림 1]에서 보 는 바와 같이 크게 3 부분으로 나누어진다.

- Executive
- Microkernel
- Hardware Abstraction Layer(HAL)

Windows NT Executive 는 프로세스간 통신, I/O, 프 로세스/스레드, 메모리 관리와 같은 기본적인 운영체 제 서비스를 담당하며, 커널은 다중 프로세스 동기 화, 인터럽트, 스케줄링과 같은 저수준의 운영 시스 템 작동을 수행한다. HAL 은 상위 계층과 하드웨어 에 대한 추상화 기능을 담당한다. Executive 는 Microkernel 과 밀접한 관계를 유지하며 시스템의 전 반적인 기능을 수행한다. 이 중에서 Executive 에 속 한 I/O Manager 는 시스템 I/O 를 관리하며, 커널 모 드 드라이버에 대한 framework 을 제공한다[3,4,10].

2.2 Kernel-mode Driver 구성

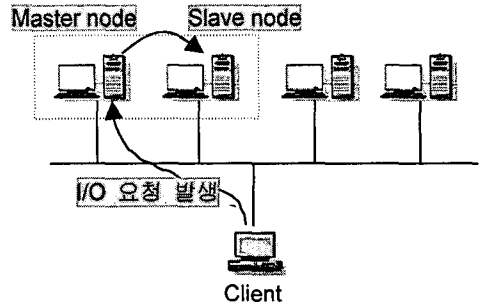
다양한 종류의 드라이버들은 계층적인 구조를 가 지며 I/O Manager 를 통하여 전달되는 요청(IRP)을 처리하게 된다. Windows NT 커널 모드 내에 disk mirroring 기능을 구현하기 위해서는 File System Driver 에 mirroring 기능을 추가할 수 있는 filter driver 의 형태로 개발하여야 한다. Filter driver 의 특 징은 기존에 존재하는 driver 의 상/하부에 위치하여 특정한 드라이버에 전달되는 I/O 요청을 가로채어 새로운 기능을 추가할 수 있는 것이다[2,3,4].



[그림 2] Kernel-mode driver 구성

3. Mirroring system 설계

3.1 시스템 구성

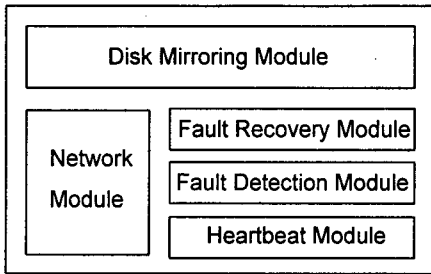


[그림 3] Mirroring 시스템 구성도

시스템을 구성하는 mirroring set 은 두 대의 컴퓨 터가 pair 로 구성되며, network 을 통하여 연결되어 있다. 사용자의 요청이 발생하면 이 요청은 master node 로 전달되고, mirroring 모듈에 의해서 두 node 의 디스크에 동일한 데이터 이미지가 저장된다[5].

만약 두 node 중 하나의 node 에 장애가 발생하면 사용자의 요청은 상대편 node 로 전달되어 사용자의 요청에 항상 응답할 수 있는 안정성을 제공한다. 이 러한 기능을 제공하기 위해서 각 pair 의 node 들은 상대 node 가 active 상태를 인식하기 위해 heartbeat 을 체크하여 node 의 상태를 파악한다. 즉, alive 메시지를 주기적으로 교환하여 timeout 내에 alive 메시지가 없으면 상대 node 에 failure 가 발생 한 것으로 간주하고, 이 시점부터의 사용자 요청을 log 에 기록기 위해 fault detection 모듈에 해당 node 에 대한 failure 를 통보한다. 이 log 는 장애가 발생 한 node 가 서비스를 재실행 할 때까지의 정보와 해당 node 에 대한 정보를 포함하고 있다. Recovery 기능은 master node 와 slave node 에서 양방향으로 수행할 수 있도록 symmetric failover 기법을 사용한 다.

3.2 Driver module 구성



[그림 4] Mirroring driver module 구성

Mirroring 기능을 수행할 driver 는 [그림 4]에서 보
는바 와 같이 5 개의 모듈로 구성된다.

Disk Mirroring 모듈은 짝이 되는 node 에게 disk 의
내용을 mirroring 하는 기능을 담당하는 핵심 모듈이
며, Network 모듈은 Windows 의 TDI 계층을 사용하
여 network 기능을 수행한다[9,10]. 사용자의 요청이
나 다른 node 로의 mirroring 이미지 전송 그리고
alive check 등은 이 network 모듈을 통하여 전달된
다. Heartbeat 모듈은 상대 node 의 active 상태 파악
과 fault tolerance 기능을 제어한다. 주기적인 메시지
교환을 통하여 alive 여부를 파악하고, 상대 node 의
장애가 파악되면 Fault detection 모듈과 Fault
recovery 모듈이 실행할 수 있도록 제어권을 넘기게
된다. Fault detection 모듈은 장애가 발생한 상대
node 에 대한 정보와 복구되기 전까지의 사용자 요
청에 대한 log 를 기록하는 기능을 수행한다. Fault
recovery 모듈은 생성된 log 를 분석하여 상대 node
에 대하여 동일한 데이터를 유지할 수 있는 복구 기
능을 수행한다.

이러한 모듈들로 구성된 driver 는 실행중인
Windows NT Filesystem Driver 에 attach 되어 동작하
는 filter driver 로 존재한다.

4. 결론

이상으로 Windows NT 커널내부에 신뢰성과 고
가용성을 지닌 disk mirroring 기능을 수행할 수 있도
록 구성된 모델의 설계를 제시하였다. Windows NT
와 같은 Microkernel 기반의 계층화된 운영체제에서
는 기존의 코드를 변경하지 않고도 커널 내부에 추
가적인 기능을 추가할 수 있는 장점이 있음을 알 수
있다.

Disk Mirroring 시스템은 높은 신뢰도를 제공하기
때문에 많은 분야에서 활용 될 수 있다. 최근엔 고
비용의 하드웨어 장비를 사용하지 않고도 소프트웨
어적으로도 우수한 성능의 mirroring 기능을 지원할
수 있어 더욱 활용 가치가 높다.

기존의 Windows NT 에서 제공하는 mirroring 기
능은 disk failure 발생시 읽기 요청에 대한 수행 성능
이 우수하지만, local 시스템에 한정 될 뿐만 아니라
저장 매체의 종류나 controller 에 대한 제약을 가지
고 있다.

본 논문에서 제시하는 모델은 network 을 통한
mirroring 기능을 제공하고, fault tolerance 기능을 보
장하여 시스템의 가용성을 증대 시킬 수 있다. 또한
이러한 시스템 구조에서는 하드웨어에 대한 제약이
줄어들며, 사용자에 대해 유연한 서비스를 제공할
수 있다.

향후 시스템에서 고려되어야 할 사항으로는 node
의 active 상태를 파악하기 위한 heartbeat check 매
시지 교환이 시스템 성능을 저하시키지 않으면서 서
비스의 중지를 가져오지 않도록 설정하는 frequency
문제와 적절한 프로토콜의 사용 여부이다. 또한 각
pair set 을 확장하여 여러 시스템으로 구성하였을 때
각 pair set 에 대한 정보를 파악하여 사용자 요청에
대한 정확한 서비스 제공할 수 있도록 하는 것이다.

참고문헌

- [1] David A. Solomon, Inside Windows NT Second Edition, Microsoft Press, 1998.
- [2] Rajeev Nagar, Windows NT File System Internals, O'Reilly, 1997.
- [3] Art Baker, The Windows NT Device Driver Book, Prentice Hall, 1997.
- [4] Peter G. Viscarola, W. Anthony Mason, Windows NT Device Driver Development, Macmillan, 1999.
- [5] Shenze Chen, don Towsly, "Performance of a Mirrored disk in a Real-Time Transaction System", ACM SIGMET -RISCM Performance Evaluation, Vol5, 198-207, 1991
- [6] A. L. Reddy, P. Banerjee, "An Evaluation on Multiple-Disk I/O Systems", IEEE Trans. On Computers, Vol38, No.12, pp.1680-1690, 1989
- [7] D. Patterson, G. Gibson, and R. Katsz, "A Case for Redundant Arrays of Inexpensive Disk(RAID)", in Proc ACM SIGMOD, pp.109-116, 1988
- [8] Microsoft High Availability <http://www.microsoft.com/technet/avail/default.htm>
- [9] Open Systems Resources <http://www.osr.com>
- [10] Windows NT 4.0 DDK Documentation