Fault-Tolerance 기능을 갖는 HA NAS 시스템의 설계 및 구현

김주영° 박준희* 권혁빈* 서희정 정영준
강원대학교 컴퓨터 과학과, (주)나래시스템*
{kimjy°, hjseo, ychung}@kangwon.ac.kr, {jh.park*, hb.kwon*}@naraesys.co.kr

Design and Implementation of HA NAS with Fault-Tolerance

JooYoung Kim° JoonHee Park*, HyuckBin Kwon*, HeeJeong Seo, YoungJun Chung
Dept. of Computer Science, Kangwon National University, Narae Systems Co., Ltd*

요 약

최근 업무에 컴퓨터를 이용하는 사례가 늘어나고, 각종 컨텐츠 서비스업계가 발전하면서 독립적인 파일 서버 기능만을 처리하도록 만든 네트워크 저장장치인 NAS(Network Attached Storage)의 이용이 점차 증가하는 추세에 있다. NAS는 기존 파일 서버의 문제점을 보완하면서 이기종 플랫폼간의 파일 공유, 스토리지 확장성, 관리 용이성 등을 특징으로 한다. 그러나 NAS 시스템에 장애가 발생할 경우에는 막대한 경제적인 손실이 발생하게 된다.

따라서 본 논문에서는 NAS 시스템 장애가 발생하였을 때, 효율적으로 장애를 복구할 수 있는 HA(High Availability) NAS 시스템을 설계 및 구현하고, 다양한 장애 상황에서 NAS를 이용하는 서비스가 중단없이 제공되는 것을 확인하였다.

1. 서 론

최근 디지털 기기의 보급 확산으로 인한 정지화상, 동영상, 오디오 파일 등 멀티미디어 데이터 이용률이 증가하고 기업이나 관공서 등의 각종 기록 보관용 파일 데이터가 누적되면서 방대한 데이터 처리 및 보관에 대한 관심이 증폭되고 있다. 수년간의 스토리지 업계 동향을 보면 이러한 멀티미디어 데이터와 기록 데이터 등은 일반 파일 서버에 저장하여 공유하는 형태가 보편적이었으나, 이기종 데이터 저장, 서버 추가 비용, 시스템 최적화 등의 문제로 인해 점차 저장 기능만을 전담하는 네트워크 저장 장치인 NAS를 이용하는 추세이다.

NAS(Network Attached Storage)는 독립적으로 파일 서비스 기능만을 처리하도록 만든 네트워크 저장장치로서, 기존 파일서버의 자원 병목 현상 해소, 이기종 플랫폼간의 파일 공유 지원, 스토리지의 확장성 및 관리 용이성 등을 특징으로 하고 있다[1].

컨텐츠 서버 또는 클라이언트에서 NAS를 사용하는 경우에 장애가 발생하게 되면 그와 연관된 서비스는 당연히 치명적인 영향을 받게 된다. 예를 들어, 웹 서비스 회사의 컨텐츠 서버가 NAS를 데이터 저장 스토리지로 이용한다고 가정해 볼 때, 시스템 장애로 인한 서비스 중지는 고객의 이탈 및 각종 컨텐츠 서비스와 관련된 막대한 경제적인 손실로 이어질 것이다. 이와 같이 시스템의 성능이 높아지고 기능이 다양해질수록 그 동작은 복잡해지고 장애의 가능성은 높아진다. 시스템 장애는 계획적인 시스템 중지에 기인하기도 하지만, 하드웨어 장애, 소프트웨어 장애, 인적 착오, 환경 장애 등의 임의적인 요인에 의해 발생할 수도 있다.

이러한 시스템 장애를 극복하기 위한 기존 방법으로 NAS 2대를 이중화하는 방법이 있다. 이 시스템은 NAS1, NAS2로 구성하여, NAS1은 서비스를 수행하고 NAS2는 NAS1의 데이터를 Sync하는 시스템이다. 이와 같은 시스템에서는 NAS1이 시스템 장애가 발생하더라도, NAS2가 똑같은 복사본을 가지고 있기 때문에 계속적인 서비스 지원이 가능하다. 그러나 스토리지 비용이 NAS의 대부분을 차지하기 때문에 경제적인 측면에서 오버헤드가 매우 크다.

따라서 본 논문에서는 NAS 시스템 장애가 발생하였을 때 HA 기능을 제공해주기 위한 고려 사항을 제시하고, 안정적인 서비스 제공을 위한 HA NAS 시스템을 설계 및 구현하였다. 구현된 시스템은 장애 상황에서 NAS를 이용하는 서비스는 거의 영향을 받지 않는 것을 확인할 수 있었다.

본 논문의 2장에서는 HA의 개념에 대해서 소개하고, 3장에서는 제안 시스템 고려사항 및 구현에 대해서 설명하고, 4장에서 결론을 맺는다.

2. HA

과거의 중앙 집중식 컴퓨팅 환경에서는 메인 프레임을 이용하여 연산, 저장, 출력 등을 모두 처리하며, 장애에 대비하여 무정지 시스템을 이용하였다. 이제는 보편적인 컴퓨팅 환경과 사용자의 마인드가 중앙집중식에서 다운사이징 또는 라이트사이징을 통한 분산처리 환경 또는 클라이언트/서버 환경으로 옮겨왔다. 시스템의 성능이 높아지고 기능이 다양해질수록, 동작은 복잡해지면서 장애의 가능성은 높아진다. 시스템의 연속성을 저해하는 장애 요인은 계획된 다운타임(planned downtime)과 계획되지 않은 다운타임(unplanned downtime)으로 구분할 수 있다. 계획된 다운타임은 시스템 변경, 신규 시스템 도입, 데이터 백업, 소프트웨어 추가, 어플리케이션 마이그레이션 등으로, 시스템을 운영하는데 필수 불가결한 요소로

분류할 수 있다. 계획되지 않은 다운타임은 정전 및 UPS 장애, 하드웨어 장애, 소프트웨어 장애, 네트워크 장애 등으로 서비스 시스템에 갑자기 발생하는 장애이다.

HA(High Availability)는 고가용성을 의미하는 것으로서, 위와 같은 다양한 서비스 장애 요인으로부터의 어플리케이션 보호, 미션 크리티컬 어플리케이션 및 운영환경의 지속적인 서비스 보장, 서비스 중단시 복구때까지의 다운타임 최소화 등이 포함된다[2].

HA 시스템은 일반적으로 프라이머리 시스템에 고장이 발생할 때까지 대기하고 있는 예비 하드웨어와 특화된 소프트웨어로 구성된다. 가용성의 허용 수준을 보장하는데 얼마나 많은 비용을 소비할 것인가를 결정하는 것은 개별적인 일이다. HA 소프트웨어는 이러한 내, 외부적 환경으로부터 시스템의 다운 시간을 최소화하는 것을 목적으로 한다. 시스템을 이중화된 소프트웨어로 구축하면 데이터의 연속성이 보장되고 어플리케이션의 서비스 보호, 시스템과 네트워크의 고장 감지 및 복구, 클라이언트/서버 환경에서 시스템의 증설을 용이하게 한다. 또한 HA 소프트웨어로 이중화된 시스템을 구축하면 생산적이고 경제적인 환경을 구축할 수 있다.

HA는 장애를 판단하고 대기 서버를 구동하기 위해서, 하드웨어, 네트워크, 프로세스 등의 장애 포인트를 지속적으로 모니터링해야 한다. 만약에 HA가 상대방 시스템의 상태를 모니터링하고 있다가 이와 같은 장애가 발생할 경우 자동으로 조치하게 된다. 이를 통해 서비스를 지속적으로 제공할 수 있도록 해준다.

HA 시스템의 구성은 양쪽 서버(Active, Standby)에 각각 HA를 설치하게 되고, 바로 아래 어플리케이션, 그 하단에 OS가 있다. HA는 상단에서 시스템과 어플리케이션을 모니터링한다. 이때 서로의 상태를 모니터링하는 연결라인을 Heartbeat라 하는데, 만약에 Active 서버에 장애가 발생하게 되면 모든 서비스는 Standby 서버에서 구동되어 서비스가 이루어진다. HA의 핵심 기능은 Active 서버의 장애를 자동으로 감지해 Standby 서버로 넘기는 failover 기능이 될 것이다[3].

3. 제안 시스템의 설계 및 구현

3.1 HA NAS 설계의 고려사항

NAS에서 HA 기능을 구현하기 위해서는 다음과 같은 사항들에 대해서 고려해야 한다.

첫째, 네트워크의 투명성을 보장해주는 네트워크단의 failover이다. 이것은 NAS를 사용하고 있는 네트워크단의 서비스 제공 서버가 Active 서버의 장애를 인지못하도록 하는 것이다. 네트워크단의 failover를 구현하기 위해서는 Standby 서버가 Active 서버의 IP를 사용하도록 해주고, MAC 어드레스 브로드캐스팅을 해주면, 서비스 제공 서버들은 기존 마운트하고 있는 NAS와의 연결점을 잃어버리지 않게 된다.

둘째, 서비스 장애가 발생한 Active서버가 공유하고 있는 스토리지를 Standby 서버가 다운타임을 최소화하면서 사용할 수 있도록 해주는 스토리지단의 failover이다. 장애가 발생한 Active 서버로 인해서 파일시스템은 비정상적인 상태이므로, 이에 대한 복구를 수행하여 Standby 서버에서 마운트하는데 지장이 없도록 해주어야 한다[5]. 본 논문에서 구현한 시스템의 파일시스템은 백그라운드 복구가 가능하므로 장애가 발생한 파일 시스템을 바로 마운트할 수 있게 된다[6]. 또한 스토리지단의 failover는 하드웨어 자체의 물리적인 장애가 발생할 수 있으므로 이에 대한 대비를 해주어야 한다. 이것은 스토리지 컨트롤러의 예비 포트로 가능해진다. 즉, 스토리지의 컨트롤러에는 2개의 포트를 가진 2개의 컨트롤러가 있기 때문에 컨트롤러의 물리적인 장애가 발생할 때 우회가 가능하다[4].

지금까지 NAS 시스템 장애를 대비하기 위한 HA 기능 구현의 고려사항에 대해서 살펴보았다. 다음 절에서는 위의 고려사항을 구현한 제안 시스템에 대해서 설명한다.

3.2 시스템 구현

본 논문에서 제안하고 있는 시스템 구성은 (그림 1)과 같이 Active 서버와 Standby 서버가 시리얼 통신 라인을 통하여 제어 정보인 Heartbeat를 교환하고, 이들 두 서버가 데이터 정보를 공유할 수 있는 Shared Disk Array 형태를 취하고 있다.

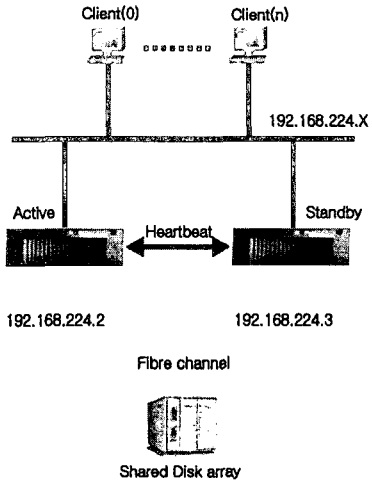

그림.1

그리고 서버(Engine)와 Shared Disk Array(Storage)의 사양은 다음과 같다.

- Engine
  - Hardware : Dual intel Xeon 2.3G CPU
    2G Fibre channel HBA
    1G Ethernet NIC
  - OS : power NRS(FreeBSD 5.1 Base)
  - Protocol : NFS, CIFS

⊃ Storage
- FC disk array(RAID5), 2TB, 1G FC

NAS에서 HA 기능을 구현할 때 중요한 부분 중의 하나가 시스템 모니터링이다. Standby 서버는 일정 시간 간격으로 시리얼 라인을 통하여 Active 서버의 정상동작 여부를 확인하게 된다. Active 서버는 시스템의 동작과 관련된 상태를 Shared memory에 저장하여 자동검출 프로그램에 의해 주기적으로 갱신하도록 되어 있다. 만일 장애가 발생하게 되면 자동검출 프로그램에 의해 상태가 변경되며 이는 시리얼 라인을 통해 Standby 서버에 의해 감지된다. 장애를 감지한 Standby 서버는 Active 서버와 동일한 동작을 수행하기 위한 준비작업을 통하여 클라이언트에 서비스를 제공하게 된다.

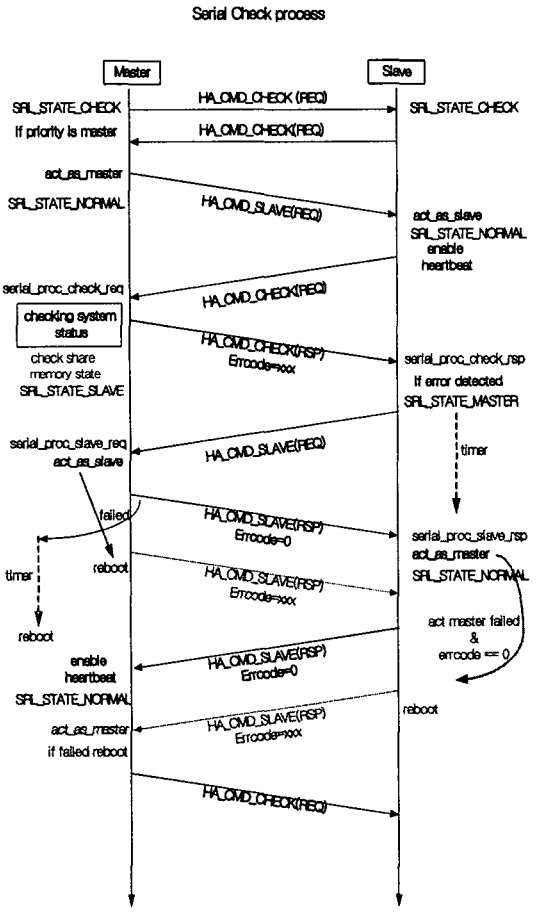

그림 2

위 (그림 2)는 Active 서버와 Standby 서버간에 연결된 시리얼 라인을 통해 모니터링 하기 위한 흐름도이다. Active 서버(master)와 Standby 서버(slave)의 역할이 정해지면(act_as_master, asc_as_slave), master에서는 네트워크 설정, 파일시스템 마운트, NAS 서비스(NFS, CIFS) 등을 수행하고, slave는 master의 상태를 정해진 주기로 체크하게 된다(serial_proc_check_req). 만약에 시스템 장애가 검출되면(if error detected), master는 slave의 역할(act_as_slave)과 reboot을 수행하고, slave는 master의 역할(act_as_master)을 수행하게 된다.

위 시스템을 이용하여 성능을 테스트 해 본 결과, 다양한 장애 상황에서 NAS를 이용하는 서비스가 거의 영향을 받지 않는 것을 확인하였다. 이것은 장애 상황이 발생하였을때, HA 기능을 통하여 Active 서버가 제공하던 서비스를 Standby 서버에서 그대로 제공하고 있음을 의미한다.

4. 결론

NAS(Network Attached Storage)는 독립적으로 파일 서비스 기능만을 처리하도록 만든 네트워크 저장장치로서, 기존 파일서버의 자원 병목 현상 해소, 이기종 플랫폼간의 파일 공유 지원, 스토리지의 확장성 및 관리 용이성 등을 특징으로 하고 있다. 그러나 NAS 서버에 장애가 발생하게 되면 막대한 손해가 발생하게 된다.

본 논문에서는 NAS 시스템 장애가 발생하였을 때 HA 기능을 제공해주기 위한 고려 사항을 제시하고, 안정적인 서비스 제공을 위한 HA NAS 시스템을 설계 및 구현하였다. 본 논문에서 제안한 시스템은 NAS header(Active 서버, Standby 서버)와 Shared Disk Array로 구성되어, Active 서버에 장애가 발생한 경우, Standby 서버가 그 역할을 대신하도록 설계되었다. 본 논문에서 구현한 시스템은 다양한 장애가 발생했을때 NAS를 이용하는 서비스가 거의 영향을 받지 않는 것을 확인할 수 있었다. 향후 연구에서는 장애 발생시 파일시스템의 완전한 복구문제, 메모리에 존재하는 데이터 손실문제 등에 대해서 연구가 이루어져야 할 것이다.

5. 참고문헌

[1] Garth A. Gibson and Rodney Van Meter, "Network Attached Storage Architecture," Communications of the ACM, Vol.43, No.II, pp.37-45, 2000.

[2] Evan Marcus, Hal Stern, "Blueprints for High Availability Second Edition," Wiley Publishing, Inc, 2003

[3] http://www.clusterkorea.org

[4] "WHITEPAPER, QLogic Fibre Channel Software: Enhancing Disk Adapter Performance, Function, and Control," qlogic corp, 2000

[5] Greg Lehey, "The Vinum Volume Manager," USENIX Annual Technical Conference, 1999

[6] Marshall Kirk McKusick, "Running "Fsck" in the Background," BSDCon Conference, 2002