

다중 컴퓨터 CC-NUMA 시스템의 진단 프레임워크 설계

◦ 김주만, 정낙주, 정태일
운영체제연구팀, 한국전자통신연구원
e-mail : jmkim@computer.etri.re.kr

The Design of Diagnostic Framework for Multi-Computer CC-NUMA Systems

Joo Man Kim, Nak Ju Jeong, Tae Il Jeong
Operating System Team, ETRI

요약

본 논문은 CC-NUMA 시스템과 같은 다중 컴퓨터 시스템에서 분산된 자원에 대한 효율적인 관리 및 진단을 위한 프레임워크의 설계에 대해 기술한다. CC-NUMA 시스템이란 고성능 상호 연결 망에 결합되어 하드웨어 방식으로 메모리 일관성을 제공하는 다중 컴퓨터이다. 이러한 시스템은 주로 엔터프라이즈 서버용으로 사용되기 때문에 TCO(Total Cost of Ownership)에 근거한 관리 및 진단 기능이 필수적으로 요구된다. 본 연구는 고성능 멀티미디어 서버(MX-Server)의 진단 및 관리 기능으로 제공되는 SDC(Server Diagnostic Card)의 하드웨어 및 소프트웨어 구조 설계 시 고려 사항에 대하여 논하였다.

1. 서론

하드웨어 기술의 발전으로 단일 보드 SMP가 보편화되었고 응용 환경의 변화 없이 확장된 처리 성능을 갖는 고도의 시스템에 대한 연구 개발이 이루어졌다. 그러나 SMP 기술은 성능 향상을 위한 확장성에 있어서 일정 처리기 수로 제한되며 따라서 고속 상호 연결 망에 노드를 결합하여 확장성을 제고한 CC-NUMA 시스템에 대한 연구가 90년대 중반 이후 학계 및 산업계에서 폭 넓게 진행되고 있다.

MX-Server는 16-way CC-NUMA 구조를 갖는 고성능 서버로서 4개의 인텔 Merced 프로세서를 갖는 SHV가 상호 연결 망 브리지인 CCA(Cache Coherent Adapter)를 통해 결합되어 하드웨어에 의한 메모리 일관성을 제공한다. 즉, 분산된 각 노드의 메모리는 일정 시스템 초기화 단계를 통해 하나의 거대한 공유 메모리 구조를 형성하여 사용자로부터 하여금 대형 다중 처리기 시스템에서와 같은 응용 환경을 제공한다.

이러한 구조에서 단일 콘솔 문제가 제기되며, 여러 개의 구성요소가 복잡하게 결합된 대형 시스템으로서 진단

및 관리 문제가 대두된다. 진단 프레임워크 설계는 이들 문제점과 향후 시스템 유지 보수에 따른 TCO 관점에서 하나의 해결책으로 제안되었다.

2. MX-Server의 진단 프레임워크

MX-Server의 각 노드에는 진단 카드인 SDC가 장착되고, SDC는 사설망(Private Network)로 연결되어 콘솔 PC에 연결된다. 또한 운영체제 수행 시 MX-Server 관리자는 공공망(public network)을 통하여 콘솔 PC 또는 원격 워크스테이션에 연결된다.

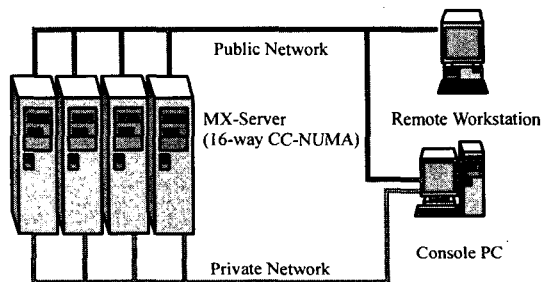


그림 1. MX-Server 진단 구조도

그림 1.은 MX-Server 의 진단 구조를 보여주고 있으며, 서버의 원격지 제어 및 관리를 위해 사설 망과 공공 망에 연결된 클라이언트를 두고있다. 클라이언트인 콘솔 PC 또는 원격 워크스테이션은 서버를 다음과 같은 상태에서 GUI 기반으로 진단 및 관리 할 수 있는 관리 소프트웨어를 탑재하고 있다.

- 온라인(On-Line) 상태 : 운영체제가 수행중인 상태로 서버상에 진단 및 관리 에이전트가 운영체제의 서비스를 콘솔 PC 및 원격 워크스테이션에 제공하는 상태.
- 오프라인(Off-Line) 상태 : 운영체제가 동작하지 않은 상태로 콘솔 PC 는 사설 네트워크에 연결된 SDC 를 통하여 서버를 진단할 수 있는 상태이며, 공공 네트워크는 접근 불가 상태이다.
- 전원차단(Power-Off) 상태 : 서버에 전원이 공급되지 않는 상태. SDC 는 배터리 백업에 의해 동작하며 SDC 내부 진단 또는 서버 전원 관리가 가능하다.

PC 에 연결된다. SHV 와는 PCI 슬롯과 I2C Serial Bus, Feature Connector 를 통하여 SHV 와 인터페이스하며 데이터 및 진단 정보를 교환한다.[그림 2]

다음은 SDC 의 설계 시 고려하여야 할 사항들이다.

- 서버 및 콘솔 PC 와의 인터페이스는 산업 표준을 따르는 접속 방법을 채택하고 소프트웨어 제어가 가능하여야 한다.

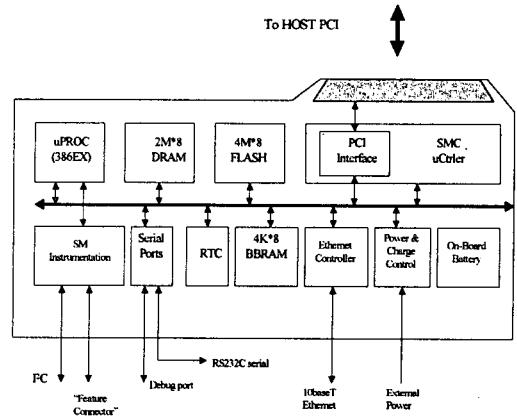


그림 2 SDC 내부 구조도

3. 동작 특성

시스템 BIOS 를 기본 BIOS 와 확장 BIOS 로 분할하여 노드의 POST 과정을 ROM 에 내장된 기본 BIOS 에서 수행하고, 각 노드를 통합하고 CC-NUMA 전역 공유 메모리 구조를 구축하는 확장 BIOS 는 SDC 카드의 플래시 메모리 파일 시스템에 저장되어 기본 BIOS 요청에 따라 SHV 메모리로 적재되어 수행된다.

SDC 는 초기화와 함께 자신의 IP 번지를 콘솔 PC 에서 수행되는 관리자에게 등록하여 연결을 설정한다. 콘솔의 그래픽 메모리는 PCI 번지로 매핑되어 실제 SDC 메모리로 경로가 재 지정된다. 따라서 SDC 의 관련 모듈에 의하여 콘솔 및 키보드 경로 재 지정을 하므로서 단일 콘솔을 구현한다.

SDC 는 충전기에 의해 전원이 항상 공급되므로 시스템 전원 또는 리셋을 원격지 제어로 가능하게 한다.

4. 진단 카드구조 및 설계

SDC 는 MX-Server 의 각 SHV 노드마다 장착되며, 각 SDC 들은 사설 네트워크를 통하여 클라이언트 콘솔

- Embedded 시스템용 실시간 운영체제가 탑재되어 수행될 수 있는 환경을 제공하고 내부 자원의 확장성을 고려한다.
- 서버 진단을 위해 I2C 버스 및 Feature Connector 를 제공하고, 자료의 송수신을 위하여 64 비트 슬롯 호환성을 갖는 PCI 규격을 갖는다.
- 내부 배터리의 수명을 가능한 길게 하기 위하여 저전력 소모, 저 발열의 Embedded 시스템용 프로세서를 채택한다.

5. 소프트웨어 구조 및 설계

SDC 의 자원 제어 및 응용 프로그램에 대한 타스크 생성과 관리를 위해 실시간 커널(RTOS)을 사용한다. 커널의 기능으로 타스크 관리와 인터럽트 서비스, 파일 입출력 관리 및 TCP/IP 통신 프로토콜을 제공한다. 하부에 여러 장치 제어를 위해 장치 구동기인 BSP(Board Specific Package) 모듈을 둔다. 콘솔 PC 와는 Ethernet 으로 연결되며 상호 TCP/IP 프로토콜을 따른다.

진단 프레임워크를 구성하는 소프트웨어 모듈의 설계 사항은 다음과 같다.

- 서버 모듈
CC-NUMA 시스템의 운영체제인 COSMIX 에 인텔 BMC 장치 구동기를 두고 상위에 On-Line 진단 관리를 위한 에이전트 소프트웨어를 개발한다.
- SDC 실시간 커널
Embedded 시스템용 실시간 커널을 채용한다. 제한된 자원 관리와 인터럽트 서비스를 수행하는 요건을 만족하여야 한다. 커널 사이즈는 충분히 작아야 하며 선점 가능한(Preemptible) 스케줄링 정책이 제공되어야 한다. 또한 멀티 프로그래밍 환경을 위한 태스크 관리 기법을 제공하여야 하고 백터 방식 인터럽트 서비스가 제공되어야 한다. 응용 프로그램 인터페이스를 제공하여야 하고 커널 기능 추가를 위한 프리미티브를 제공하여야 한다. 또한 파일 시스템 관리 기능과 TCP/IP 지원 모듈을 포함하여야 한다.

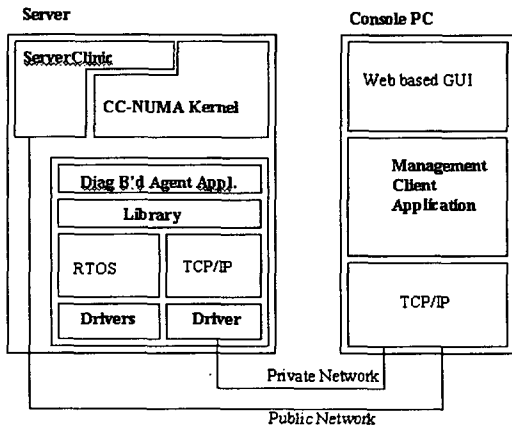


그림 3. 진단 프레임워크 소프트웨어 구조

- SDC 진단 관리 모듈
SDC 내부 자원뿐만 아니라 서버 BMC 기반의 진단 대상에 대한 관리 모듈을 포함한다. 다중 노드의 단일 콘솔 입출력 기능을 제공하고, 서버에서 발생한 이벤트를 수시로 보고 받아 관리할 수 있고, 전원의 인가 및 단절에 대한 제어가 가능할 것이다.
- 클라이언트 콘솔 PC 소프트웨어

사설망 또는 공공망에 연결된 원격 콘솔은 그래픽 지원이 가능한 PC 운영체제를 채용하여 서버 및 SDC 를 진단/관리할 수 있는 소프트웨어 모듈을 둔다. 표준 소켓 프로그래밍에 의해 각각의 망에 인터페이스하며 가상 입출력 처리 모듈을 두어 터미널 및 키보드 에뮬레이션, 플로피, CD 롬을 서버 장치로 지원한다. 사용자 인터페이스는 웹 기반의 GUI를 지원한다.

- CC-NUMA 패키지
CC-NUMA 시스템의 초기화 절차를 제어하는 모듈을 가진다. 이러한 초기화 모듈은 ROM 에 내장 되거나 원격 콘솔 환경에서 필요 시 개발되어 동적인 적재가 가능할 것이다. 이 패키지에는 인텔의 다중 컴퓨터 사양을 만족하는 형상 관리 기법을 포함한다.

6. 결론

본 논문에서는 엔터프라이즈 서버에서 TCO 관점의 비용 절감과 시스템 관리의 효율성을 제고하는 고성능 컴퓨터 서버인 MX-Server 의 진단 프레임워크에 대한 설계 내용을 기술 하였다. SDC 는 PCI 카드로 제작하여 보통의 서버 시스템에서 진단 및 관리가 용이하도록 설계되며 내장형 실시간 커널과 진단 에이전트 소프트웨어 및 원격 콘솔의 관리 소프트웨어에 대한 설계 개념 및 방향에 대하여 기술하였다.

본 연구에서 제안된 SDC 및 관련 소프트웨어는 어떤 특정 시스템에 특화 되지 않은 일반적인 클러스터링 시스템의 단일 콘솔 지원과 네트워크 기반의 진단 시스템에 적용할 수 있도록 확대 개발될 것이다.

7. 참고문헌

- [1] L. Beaudet, F. Eshragh, "System-level diagnostics troubleshoot multiprocessors," Computer Design, pp.77-82, June 1987
- [2] J. Bartlett, R. Carr, et al, "Fault Tolerance in Tandem computer Systems," in Reliable Computer Systems: Design and Evaluation, Bedford, MA: Digital Press, 1992
- [3] 이해동, 김주만, 김용연, "그래픽 기반 SPAX 진단 시스템 개발," '97 한국정보과학회 총칭지부 학술발표회, 12, 1997