제36회 한국정보처리학회 추계학술발표대회 논문집 제18권 제2호 (2011. 11)

IPMI와 PSMI 연동을 통한 서버 전력 소비량 모니터링 시스템

이상훈, 최린
고려대학교 전기전자전파공학부
e-mail:{smile97, lchoi}@korea.ac.kr

An Interlocking Interface between IPMI and PSMI for Monitoring the Power Consumption of Server Systems

Sang Hoon Lee, Lynn Choi
School of Electrical Engineering, Korea University

요 약

스마트폰을 비롯하여 스마트 TV, 클라우드 서비스 등과 같은 네트워크 기반 서비스의 증가로 인해 IDC의 서비스 부하 증가 속도가 더욱 증가하고 있다. 이로 인해 IDC의 운영 및 관리 비용이 급격히 증가하고 있다. 특히 IDC의 전력 소비량은 IDC 관리 비용 중 상당한 부분을 차지하고 있다. 이에 따라 본 설계서에서는 IDC의 에너지 효율적인 관리를 위한 기본 시스템으로 적용 가능한 서버 모니터링 시스템을 소개한다. 본 설계서를 통해 제안하는 서버 모니터링 시스템은 현재 산업 표준으로 받아들여지고 있는 IPMI (Intelligent Platform Management Interface)와 PSMI (Power Supply Monitoring Interface)의 연동을 통해 서버의 하드웨어 상태 및 전력 소비량을 측정할 수 있도록 하였다.

1. 서론

스마트폰을 필두로 스마트 TV, 클라우드 시스템 등과 같이 최근 급격한 성장을 보이고 있는 네트워크 기반 응용 서비스들에서는 기존의 PC 환경에서 제공되던 응용들에 비해 서비스를 제공함에 있어서 서버의 역할이 증가하고 있다. 이러한 흐름으로 인해 IDC (Internet Data Center)의 규모는 더욱 빠르게 증가하고 있으며, 이러한 IDC의 운영 및 관리에 소요되는 비용도 급격히 증가하고 있다. 특히 IDC의 전력 소비량은 서버 및 제공되는 서비스가 증가함에 따라 꾸준히 증가하고 있으며 이에 따른 비용이 IDC 관리 비용 중 상당한 부분을 차지하고 있다. 또한, 높은 전력 소비로 인해 발생되는 서버실의 온도 증가는 과열로 인한 장비 이상을 야기할 수도 있다. 이에 따라 IDC의 에너지 효율적인 관리를 위한 다양한 기술들에 대한 필요성이 높아지고 있다[1].

최근 Enterprise급 고성능 서버에서 서버 관리를 위해 사용되고 있는 IPMI (Intelligent Platform Management Interface)[2]는 사실상 산업 표준으로 받아들여지고 있다. IPMI는 software agent가 필요 없는 원격 서버 관리 표준 규격으로 센서를 통한 시스템 장치 감시, 컴포넌트 제어 및 시스템 이벤트 기록 방법을 정의하고 있다. 이와 더불어 CPU에 독립적인 BMC (Baseboard Management Controller)를 이용한 내장형 관리 시스템과 별도의 회선을 통한 통신 방법을 정의하고 있다. 이로 인해 IPMI는 서버 시스템의 운영체제 및 주변 장치들의 구성이 전혀 다른 서버들에서도 동일한 인터페이스를 통한 원격 제어 기능을 제공할 수 있다. 특히 IPMI 기반 하드웨어 자원 모니터링 시스템은 out-of-band 방식의 처리를 이용하여 관리 대상 서버에 대기 전원만 들어온 상태에서도 동작 가능하며, 하드웨어 모니터링을 위한 에이전트를 운영체제에서 수행하지 않으므로 모니터링으로 인한 CPU 자원 사용을 방지하는 장점을 제공한다.

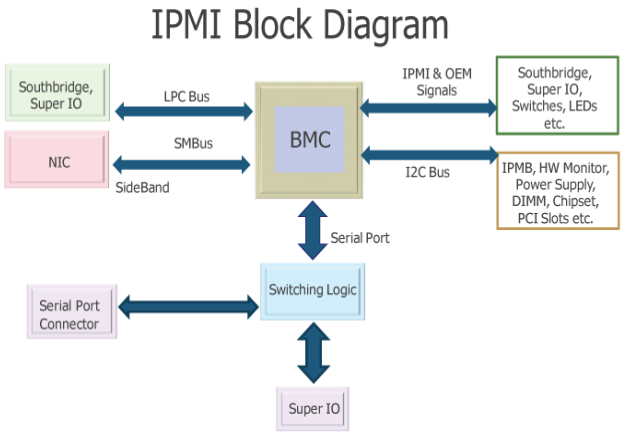
IPMI에서 제공하는 시스템 모니터링 기능들은 서버의 메인보드에 부착된 각종 센서 장치들을 이용하여 구현된다. 즉, IPMI의 모니터링 기능은 제공되는 센서의 종류에 따라 달라질 수 있는데, 일반적으로 CPU 전압, 팬 스피드, 온도 등과 관련한 센서들이 설치되어 있다[3]. 그런데, 일반적인 서버의 메인 보드에서는 전력 소비량을 측정할 수 있는 센서가 부착되어 있지 않으며 이로 인해 IPMI만을 이용해서는 서버의 전력 소비량을 측정하기 어렵다.

서버에서 사용되는 총 전력 소비량은 메인 보드가 아닌 전원 공급 장치에서 제공하는 PSMI (Power Supply Monitoring Interface)[4]를 통해 측정할 수 있다. PSMI에서 측정된 전력 소비량은 I2C bus[5]를 통해 EEPROM에 기록되는데, PSMI와 같이 IPMI의 BMC도 I2C bus를 통해 EEPROM에 연결되어 있으므로 BMC은 이를 통해 PSMI 측정값을 읽어올 수 있다. 본 설계서를 통해 제안하는 서버 모니터링 시스템은 EEPROM을 통한 IPMI와 PSMI의 연동을 이용하여 서버의 메인 보드에 부착된 센서의 측정값뿐만 아니라 전원 공급 장치에서 측정된 전력 소비량을 확인할 수 있도록 하였다.

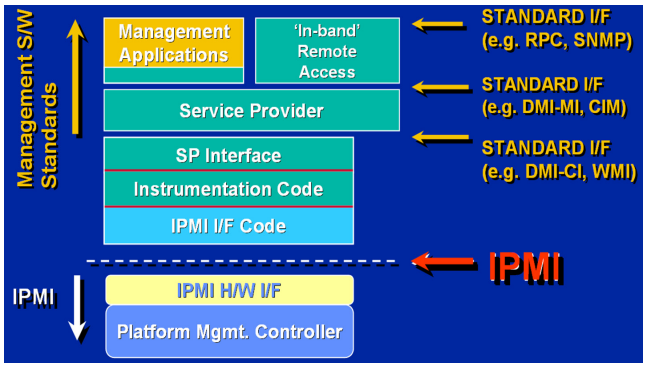
2. 관련 이론 및 연구

2.1 Intelligent Platform Management System (IPMI)

IPMI 표준은 서버 하드웨어의 관리에 집중된 구조로서 1998년 Intel과 HP를 중심으로 약 70여개의 회사가 모여서 만든 IPMI 포럼에 의해 제정되었다[6]. IPMI 표준 기반 관리 구조는 기존의 소프트웨어 agent 기반 관리가 아닌 agentless 원격 서버 관리 표준 규격으로 시스템 하드웨어 및 센서를 감시하여 시스템 컴포넌트를 제어하고 시스템 이벤트를 기록하는 방법을 정의한 확장성 표준이다. IPMI 표준 지원을 위한 BMC (Baseboard Management Controller)는 핵심 마이크로 제어 기능이 내장되어 있으며 이를 통해 서버의 물리적 특성, 예를 들어 온도, 사용 전압, 쿨링 팬의 회전 속도, 본체의 개폐여부 등에 대한 모니터링이 가능하다. 2002년에 소개된 IPMI 1.5 표준 규격에서는 IPMI over LAN 규격이 추가되었으며 이는 원격 시스템 제어 및 이벤트 로그 원격 접속, 오류 발생에 대한 알람과 같은 원격 관리를 가능하게 해준다. 즉, IPMI를 이용하여 서버 관리자는 로컬 서버에서는 물론, 네트워크 상의 다른 서버로부터도 모니터링 정보를 획득할 수 있다. 그림 1은 IPMI의 BMC interface를 포함하는 block diagram을 보여주며, 그림 2는 IPMI 지원 프로그램 인터페이스 구조 및 동작을 나타낸다.



(그림 1) IPMI의 block diagram



(그림 2) IPMI의 software stack

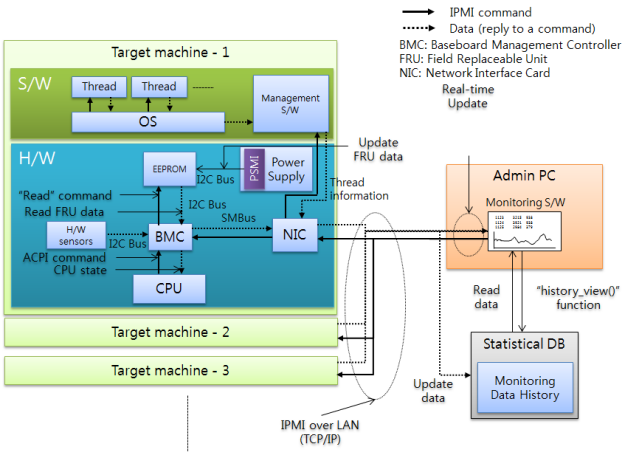
2.2 Power Supply Monitoring System (PSMI)

PSMI [4]는 SSI forum에서 발표한 전원 공급 장치 관리 인터페이스 표준으로 전원 공급 장치의 모니터링, 관리, 진단, 제어 및 이벤트 기록 등의 기능을 제공한다. 특히 출력 전압 및 전류를 측정하여 서버가 소비하는 전력량을 산출할 수 있도록 되어 있다. PSMI에서 제공하는 전력 소비량은 실제 전원 공급 장치를 통해 공급되는 전력량을 측정하는 것이므로 서버의 메인 보드에 별도의 센서를 부착하여 측정하는 것에 비해 더 정확한 값을 얻을 수 있는 장점이 있다. PSMI 표준안에서는 총 8개의 출력 전압 (3.3V, 5V, 12V1, 12V2, 12V3, 12V4, -12V, 및 5VSB)를 제공하는 전원 공급 장치의 표준 모델을 정의하고 있다. 특히 해당 모델에서 모니터링 및 제어를 위한 관련 인터페이스에 대한 정의도 포함하고 있다.

3. IPMI와 PSMI 연동 시스템 설계

서버 구성에 따라 PSMI를 지원하지 않고 PMBus만 지원하는 경우도 있는데, 이 때, IPMI는 PMBus를 통해 전원 공급 장치에 설치된 전력 센서의 측정값을 읽어올 수 있다. 다만, PMBus를 사용할 경우 PSMI와 달리 전원 공급 장치에 설치된 전력 센서 측정값만을 읽어 올수 있다. PSMI가 설치된 경우에는 전력 센서 측정값뿐만 아니라 전원 공급 장치의 팬 회전 속도, 공급되는 전류량, 전원 공급 장치 제조 정보 등 다양한 정보를 읽어올 수 있다. 본 논문에서는 PSMI를 지원하는 전원 공급 장치가 설치된 환경에서 사용 가능한 IPMI와 PSMI의 연동 시스템을 제안한다.

앞서 다루었듯이 IPMI와 PSMI는 메인보드상의 EEPROM을 통해 산호 연동될 수 있다. 그림 3에서와 같이 전원 공급 장치는 I2C 버스를 통해 전력 센서의 측정값을 FRU (Field Replaceable Unit) data 형태로 EEPROM에 기록한다. FRU data format은 IPMI 표준안에 함께 정의되어 있다. 표 1, 2 와 3은 전원 공급 장치에 사용되는 FRU data format을 보여준다[7]. FRU data



(그림 3) 서버 모니터링 시스템

update와 read는 I2C bus를 통해 이루어지도록 설계되었다. 서버 시스템에 따라 PSMI와 EEPROM의 통신 회선으로 PMBus를 사용하는 경우도 있는데, 이러한 환경에서도 BMC가 EEPROM에 저장된 PSMI 측정값을 읽어오기 위한 회선은 I2C bus를 사용하도록 되어 있다.

그림 3 및 표 1, 2, 3에서 알 수 있듯이 IPMI에서 PSMI 측정값을 읽어오기 위해서는 IPMI에서 FRU data read 함수를 이용해야 한다. 이를 위해 본 설계서에서 제안하는 모니터링 시스템은 IPMItool 1.8.11[8]을 이용하여 관련 함수를 구현하였다. IPMItool은 서버의 IPMI 인터페이스를 통해 모니터링 및 제어 기능을 제공하는 공개 소프트웨어로 이 외에도 IPMIUtil[9], FreeIPMI[10], OpenIPMI[11] 등이 있다.

<표 1> FRU device Information

Area Type	Description
Common Header	As defined by the FRU document
Internal Use Area	Not required, do not reserve
Chassis Info Area	Not applicable, do not reserve
Board Info Area	Not applicable, do not reserve

<표 2> FRU Device Product Information Area

Field Name	Field Description
Manufacturer Name	{Formal name of manufacturer}
Product Name	{Manufacturer's model number}
Product part/model number	Customer part number
Product Version	Customer current revision
Product Serial Number	{Defined at time of manufacture}
Asset Tag	{Not used, code is zero length byte}
FRU File ID	{Not required}
PAD Bytes	{Added as necessary to allow for 8-byte offset to next area}

<표 3> Multi-record Information Area

Field Name (PS info)	Field Information Definition
Overall Capacity (watts)	550 {Low power version would be 450}
Peak VA	610 {Low power version would be 490}
Inrush current (A)	50
Inrush interval (ms)	5
Low end input voltage range 1	90
High end input voltage range 1	140
Low end input voltage range 2	180
High end input voltage range 2	264
A/C dropout tol. (ms)	20
Binary flags	Set for: hot Swap support, Autoswitch, and PFC
Peak Wattage	Set for: 10s, 610W
Combined wattage	Set for: 5V & 3.3V combined wattage of 115W
Predictive fail tach support	Not supported
Field Name (Output)	**Field Description**
Output Information	Set for: Stanby on +5VSB
All other output fields	Using parameters in the EPS12V spec.

IPMI와 PSMI 연동과 관련한 기본적인 기능들은 IPMItool 1.8.11에 포함된 ipmi-chassis.c와 ipmi-fru.c에 다음과 같은 함수들로 구현되어 있다.

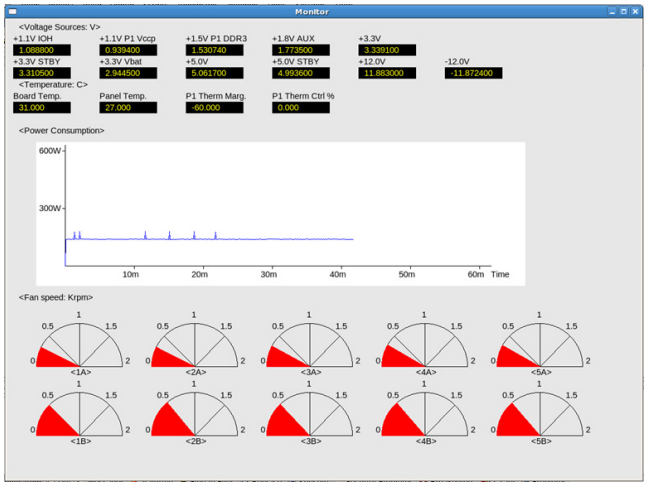
int ipmi_sensor_monitor(struct ipmi_intf *intf)

int ipmi_fru_print_all(struct ipmi_intf * intf)

void fru_area_print_multirec_bloc(struct ipmi_intf * intf, struct fru_info * fru, uint8_t id, uint32_t offset);

위 함수들 외에도 FRU data 중 전력 소비량 데이터만을 읽기위한 별도의 함수를 추가로 구현하였고, PSMI 측정 데이터와 보드에 부착된 센서의 측정 데이터를 한 번에 확인할 수 있도록 관련 함수들을 통합 및 수정하였다.

아래는 IPMItool에 구현된 int ipmi_fru_print()를 이용하여 작성된 FRU data read 함수를 보여준다. FRU data read 함수는 IPMItool 내에 기본적으로 구현되어 있는 함수이다. 이 함수를 통해 EEPROM에 저장된 PSMI의 FRU data를 읽어오도록 함수를 수정하였다. PSMI와 IPMI의 연동을 통해 서버 모니터링 시스템에서는 메인보드의 전압, 팬 스피드, 온도 및 전력 소비량을 측정할 수 있고, 이를 표출하기 위한 사용자 인터페이스는 그림 4와 같다.



(그림 4) IPMI와 PSMI 연동을 통한 서버 모니터링 시스템

```
1:  int ipmi_fru_monitor_print(struct ipmi_intf * intf,
struct sdr_record_fru_locator * fru){
2:     char desc[17];
3:     uint8_t save_addr;
4:     int rc = 0;
5:     if (fru == NULL)
6:       return ipmi_fru_monitor_print_(intf, 0);
7:     if (fru->dev_type != 0x10 &&
8:        (fru->dev_type_modifier != 0x02 ||
9:        fru->dev_type < 0x08 || fru->dev_type > 0x0f))
10:       return -1;
```

```
11:    if (fru->dev_slave_addr == IPMI_BMC_SLAVE_ADDR
&& fru->device_id == 0)
12:     return 0;
13:    memset(desc, 0, sizeof(desc));
14:    memcpy(desc,  fru->id_string,  fru->id_code  &
0x01f);
15:    desc[fru->id_code & 0x01f] = 0;
16:    printf("FRU Device Description : %s (ID %d)\n",
desc, fru->device_id);
17:    switch (fru->dev_type_modifier) {
18:      case 0x00:
19:      case 0x02:
20:        intf->target_addr = fru->dev_slave_addr;
21:     rc = ipmi_fru_monitor_print_(intf, fru->device_id);
22:        intf->target_addr = save_addr;
23:        break;
24:     case 0x01:
25:        rc = ipmi_spd_print_fru(intf, fru->device_id);
26:        break;
27:     default:
28:        if (verbose)
29:          printf(" Unsupported device 0x%02x "
30:          "type 0x%02x with modifier 0x%02x\n",
31:          fru->device_id, fru->dev_type,
32:          fru->dev_type_modifier);
33:        else
34:          printf(" Unsupported device\n");
35:    }
36:    return rc;
37: }
```

(그림 5) FRU data read 함수 구현

그림 5는 FRU data read를 위한 ipmi_fru_monitor_print를 보여준다. 이 함수의 주요 기능은 라인 17번의 switch문으로, EEPROM에 저장된 FRU data의 device type을 인지하여 power supply에 해당하는 정보만을 읽어오도록 한다. 라인 21이 power supply의 FRU data를 읽는 부분으로 이는 ipmi_fru_monitor_print_ 라는 별도의 함수를 통해 구현되었다.

4. 결론

서버 관리 기술 분야에서 산업 표준으로 여겨지는 IPMI와 PSMI의 연동은 각 기술들이 제공하지 못하는 값들을 단일 관리 에이전트를 통해 확인할 수 있도록 한다. 이는 서버 모니터링의 효율성을 증가시키고, 나아가 서버 제어 정책을 수립함에 있어서 보다 정확한 데이터를 통해 에너지 효율적인 제어 정책을 적용할 수 있도록 한다. 특히 표준안에 기반을 두어 시스템을 연동함으로써 대상 서버의 사양, 운영체제, 제공 서비스 등에 영향을 받지 않는 서버 모니터링 시스템을 구축할 수 있다.

서버 관리 기술 분야에서 산업 표준으로 여겨지는 IPMI와 PSMI의 연동은 각 기술들이 제공하지 못하는 값들을 단일 관리 에이전트를 통해 확인할 수 있도록 한다. 이는 서버 모니터링의 효율성을 증가시키고, 나아가 서버 제어 정책을 수립함에 있어서 보다 정확한 데이터를 통해 에너지 효율적인 제어 정책을 적용할 수 있도록 한다. 특히 표준안에 기반을 두어 시스템을 연동함으로써 대상 서버의 사양, 운영체제, 제공 서비스 등에 영향을 받지 않는 서버 모니터링 시스템을 구축할 수 있다.

참고문헌

[1] Roger S. Pressman "Software Engineering A Practltiners' Approach" 3rd Ed. McGraw Hill

[1] K. Kant, P. Mohapatra, "Guest Editors' Introduction: Internet Data Centers," IEEE Computer, Vol. 37, Issue 11. 2004, pp. 35~37

[2] Intelligent Platform Management Interface (IPMI), http://www.intel.com/design/servers/ipmi/

[3] Intel Server Board Products http://www.intel.com/ko_KR/products/server/motherboard/index.htm

[4] Power Supply Monitoring Interface (PSMI), https://ssiforum.org/

[5] I2C bus (Inter-Integrated Circuit bus), http://en.wikipedia.org/wiki/I%C2%B2C

[6] H. Jordan, "An Introduction to the Intelligent Platform Management Interface," Magazine of Dell Power Solutions, 2004

[7] EPS12V Power Supply Design Guide, V2.91

[8] IPMItool, http://ipmitool.sourceforge.net

[9] IPMIUtil, http://ipmiutil.sourceforge.net

[10] FreeIPMI , http://www.gnu.org/software/freeipmi

[11] OpenIPMI, http://openipmi.sourceforge.net