

# 병렬 디스크 시스템의 모델링 및 모의 실험

백승훈, 김경호, 박규호

한국과학기술원, 전기및전자공학과, 컴퓨터공학연구소

## Modeling and Simulation of a Parallel Disk System

Sung-Hoon Baek, Kyung Ho Kim, Kyu Ho Park

Dept. of Electrical Engineering, KAIST

### 요 약

본 논문에서는 형식론적 모델링 및 모의 실험 방법을 사용하여 병렬 디스크의 성능 예측 및 영상 데이터용 비선형 편집기의 개발을 위한 여러 형태의 구조에 대해서 모의 실험한 결과 및 성능에 영향을 미치는 요소들에 대해 다루고 있다. 컴퓨터가 발전하기 전에는 테이프를 이용한 선형 편집기를 사용하였으나, 컴퓨터가 발전하고 또한 입출력 장치의 성능과 버스의 성능이 향상되면서 RAID(Redundant Array of Inexpensive Disks)가 개발되었고, 이것을 이용한 비선형 편집기가 개발되었다. 병렬디스크는 비선형 편집기뿐만 아니라 VOD(Video On Demand)나 데이터 서버등 응용 분야가 많고, 요구되는 성능에 따라서 다양한 구조의 병렬디스크가 사용된다. 매우 큰 대역폭이 요구되는 병렬 디스크의 경우에는 다양한 버스 및 버퍼를 사용한 복잡한 계층적 구조가 요구된다. 그러나 이렇게 복잡한 병렬디스크를 개발하기에 적당한 개발 도구가 없어 필요한 성능에 꼭 맞는 시스템을 개발하기가 어렵다. 특히 디스크의 성능을 수학적으로 쉽게 구할 수 없다. 또한 시스템이 실제로 완성되기 전에는 디스크의 운영 알고리즘의 개발 및 파일 시스템의 개발이 쉽지 않다. 그래서 쉬운 소프트웨어 개발과 디스크의 성능을 쉽게 측정하기 위해서 객체 지향 성능 예측 시스템(OOPPS, Object Oriented Performance Prediction System)을 개발하였고, 이것을 이용하여 HDTV용 비선형 편집기의 성능을 예측하고, 필요한 성능을 만족하는 구조를 선택하여 실제로 개발하였다.

## 1. 개 요

현대 마이크로 프로세서 기술이 매년 40에서 60퍼센트의 놀라운 속도로 발전하고 있다. 또한 디스크의 용량이 마이크로 프로세서 기술의 발전 속도로 비슷하게 커져가고 있다. 그러나 디스크의 성능은 매년 7에서 10퍼센트의 성장밖에 하지 못 하였다. 그래서 새로운 입출력 장치로서 병렬 디스크가 발명되었다. 이것은 데이터를 여러 개의 디스크에 동시에 접근함으로써 하나의 빠르고 큰 디스크로 보이게 하는 것이다. 병렬 디스크는 여러 가지 용도로 사용되면서 많은 연구가 이루어졌다. 실시간 운영을 위한 디스크 운영 알고리즘, 많은 사용자를 수용할 수 있는 VOD의 다양한 구조, 입출력 서버를 위한 캐쉬 정책 등이 이러한 연구들이다.[1,2]

큰 대역폭이 요구되거나 실시간 운영이 요구되는 병렬 디스크 시스템은 평면 구조의 병렬 디스크는 필요한 대역폭을 만족시킬 수 없어 계층적 구조의 형태를 가지는 경우가 많다. 이런 시스템의 구조는 선택 사항이 많을 뿐만 아니라 그 성능을 예측하기가 쉽지 않아서 제품 개발 시간(time to market)과 더불어 가격 및 판매 성능 또는 파소 성능 없는 시스템을 찾기가 어렵다.

병렬 디스크 시스템의 성능을 쉽게 구하기 위하여 시스템 수

준의 설계 방법과 형식론적인 모의 실험 방법으로 빠르고 쉽게 성능을 예측할 수 있는 방법론(객체 지향 성능 예측 시스템)을 개발하였다.[3] 이것은 UML과[5] DEVS[5] 형식론을 이용하여 시스템 모델링 및 모의 실험 방법을 통한 성능 예측 시스템에 관한 방법론이다. 형식론은 쉬운 접근과 기술의 이해능력뿐만 아니라 쉽게 검증될 수 있는 장점이 있다. UML은 추상 수준에서의 시스템 모델링과 형식론적 시스템 기술방법과 그래픽 사용자 인터페이스를 제공한다. DEVS는 실제적 시뮬레이터를 담당한다. UML을 이용하여 그래픽 환경에서 병렬 디스크 시스템을 모델링하고 성능을 분석할 수 있고, 모의 실험 모델을 이용하여 실제 하드웨어 없이도 소프트웨어 개발과 검증을 수행 할 수 있다.

이 방법론을 이용한 모의 실험을 통하여 요구되는 성능에 필요한 병렬 디스크 시스템의 구조를 구하고 병렬 디스크 시스템의 성능에 크게 좌우되는 요소인 데이터 요청 크기 및 디스크의 수와 버스 계층 구조에 대한 결과를 얻었다.

## 2. 배경 지식

앞에서는 UML과 DEVS형식론을 이용한 객체 지향 성능 예측 시스템을 언급하였다. 전설에서 청사진을 만드는 것만큼이나 시스템을

구체적으로 설계하기에 앞서 시스템을 추상 수준에서 전체적으로 모델링하는 것이 중요하다. 좋은 모델은 프로젝트 팀원들과 의사교환을 하는데 큰 도움이 되고 큰 시스템을 이해하는 데에 매우 도움이 된다. 이러한 목적으로 만들어진 것이 UML이다. 이것은 실시간 시스템, 소프트웨어, 비즈니스 모델 등을 기술하고 문서화하고 가시화 하는데 사용되는 가시적 객체 지향 모델링 언어이다. UML은 크고 복잡한 시스템을 표현하는데 성공적임이 입증되어져 있다.

DEVS는 이산 사건 시스템을 기술하는 수학적 형식론으로서 객체 지향적이고 계층적으로 모델을 표현할 수 있다. DEVS는 Atomic-model과 Coupled-model로 구성되어 있다. Atomic-model은 외부 사건 또는 내부 시간 초과에 의해 상태가 천이하는 동적 모델에 해당하며, 출력은 내부 시간 초과에 의해 상태가 천이 될 때에 발생된다. 이러한 출력은 다른 모델의 외부 사건을 발생시켜 다른 모델의 상태 천이를 발생시킨다. Coupled-model은 그 인터페이스는 Atomic-model과 같지만 Atomic-model 또는 Coupled-model의 집합으로 구성되어 있다. 각 모델들은 입력과 출력으로 연결이 되며, 유한한 상태를 가지고, 스스로 발생시키는 사건과 외부에서 발생하는 수동적인 사건들로 상태는 천이한다.

객체 지향 성능 예측 시스템에서의 모의 실험 순서는 다음과 같다. UML로써 시스템의 구조를 GUI 환경에서 모델링하고, 이것으로부터 자동적으로 DEVS 모의 실험 모델을 생성하고, 모의 실험을 수행하여 결과를 얻는다. 디스크, 버스, 버퍼와 같은 널리 사용되는 모델들은 라이브러리로 되어 있으며, 상속의 방법을 통하여 주어진 라이브러리를 수정하거나 새로운 모델을 만들어 라이브러리에 추가 할 수 있는 기능을 갖고 있다. DEVS의 시뮬레이터로서는 DEVS++를[6] 사용하였다.

### 3. 예제 연구

#### 3.1. 시스템 구조 및 모델링

본 예제는 HDTV를 위한 비선형 편집기 시스템(NESH : Nonlinear Editing System for HDTV)의 제작에 관한 것이다. 이 비선형 편집기는 RAID(Redundant Array of Inexpensive Disks)를 이용하여 실시간으로 비디오 프레임 단위로 임의의 위치를 편집하여 재생, 저장하는 시스템이다. 이 시스템의 핵심 기술은 고용량의 HDTV 데이터를 전송할 수 있는 대역폭과 데이터들이 최악의 조건으로 흩어져 있어도 실시간으로 임의 접근하여 재생, 저장할 수 있는 병렬 디스크이다. 이것을 구현하는 데에는 많은 선택 사항이 있다. 디스크는 몇 개가 필요하며 디스크에 사용될 버스는 무엇으로, 디스크 컨트롤러는 무엇을 사용하며, 컨트롤러를 몇 개 사용하는 등의 많은 선택사항이 있다. 가격이나 성능에서 처지거나 앞서지 않는 시스템을 개발하기 위해서 앞서 언급한 객체 지향 성능 예측 시스템을 사용하였다

비선형 편집기는 매우 높은 대역폭이 필요하기 때문에 성능을 만족시키기 위해서 고성능의 디스크와 버스를 사용해야만 하였다. 두 개의 Ultra Wide SCSI 버스에 각각 용량이 9 GBytes인 Barracuda 9 Disc Drive(ST19171) 네 개씩 설치되어 있고 SCSI controller로서 SYM53C875를 사용하였다. SYM53C875는 두 개의 SCSI bus 인터페이스와 한 개의 PCI bus를 지원한다. 비디오 스트림과 디스크 사이의 데이터 완충 버퍼 컨트롤러로서 SBC(Striping

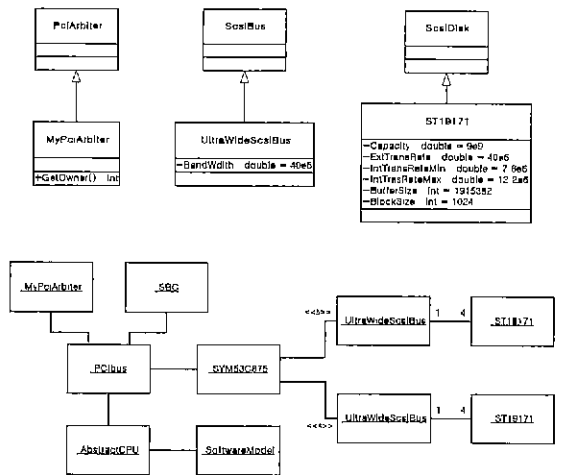


그림 1 객체 지향 성능 예측 시스템으로 설계한 병렬 디스크 시스템의 구조

Buffer Controller)를 FPGA를 사용하여 설계하였다. 실제 시스템에서는 SBC를 통하여 비디오 컨트롤 보드로 연결되는 버스가 연결되어 있다

그림 1은 객체 지향 성능 예측 시스템을 사용하여 디스크 시스템을 모델링한 것이다. 객체 지향 성능 예측 시스템은 UML 도구를 이용하여 그래픽환경에서 시스템을 모델링한다. 그림 1에서 윗부분의 다이어그램은 주어진 라이브러리에서 원하는 새로운 클래스를 정의하는 것을 보여 주고 있다 PCI Arbiter의 버스 획득자 알고리즘을 바꾸고, 일반 목적용 SCSI 버스에서 Ultra Wide SCSI 버스를 만들고, 일반 목적용 디스크 클래스에서 ST19171에 맞는 디스크 값을 정의하는 것을 보이고 있다 그림의 아랫부분은 새로 만든 클래스나 라이브러리에 있는 클래스를 실제화(Instantiation) 한 객체를 이용하여 객체간의 연결관계를 설계한 것이다 SBC는 새로운 ASIC이므로 라이브러리에 존재하지 않는다. 이런 것들은 그래픽 환경에서 DEVS 모델을 구현할 수 있는 DEVShop이나[6] C++ 수준에서 DEVS++ 모의 실험 모델을 사용자가 설계해야한다. 소프트웨어와 DEVS++의 인터페이스로서 AbstractCpu를 두었다. SoftwareModel에서는 디스크를 효율적으로 동작시키기 위한 알고리즘을 구현하였다. 이 소프트웨어 모델에서 나온 데이터는 AbstractCpu를 거치고 PCI bus를 거쳐 SBC나 SCSI 컨트롤러로 전달되도록 하였다. 주요 데이터 패스는 SBC와 디스크를 잇는 길이다. SBC는 데이터를 발생시키거나 소모하도록 설계되어져 있다.

이렇게 설계한 UML 다이어그램으로부터 DEVS 모의 실험 모델이 만들어진다. 이것은 C++로 구현되어져 있으며, 이것을 컴파일하면 실행 가능하고 결과 값을 구할 수 있다.

#### 3.2. 모의 실험

그림 2~4는 병렬디스크 시스템의 모의 실험 결과이다. 그림 2는 SCSI 버스 한 개에 디스크 구동 방법을 파이프라인 모드 없이 평범한 방법으로 실험한 성능 지표이고, 그림 3은 SCSI 버스 한 개에 디스크를 파이프라인 방법으로 구동하였을 때의 성능 지표이며, 그림 4

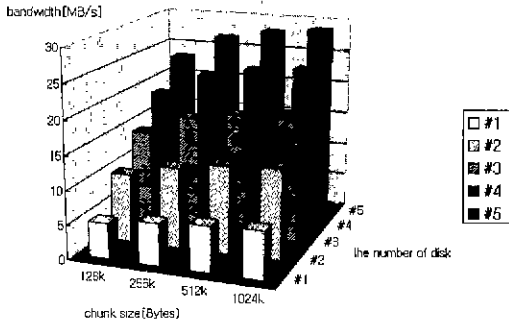


그림 2 SCSI 버스가 한 개일 때의 병렬 디스크의 성능

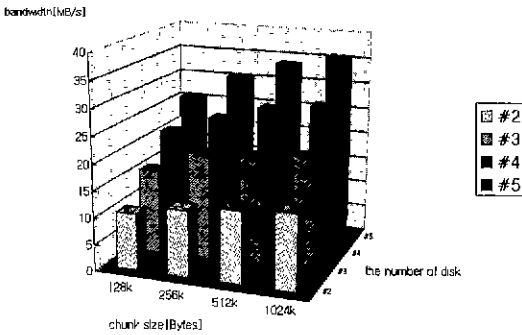


그림 3 SCSI버스가 두 개이고 파이프라인 모드로 디스크를 구동하였을 때의 성능

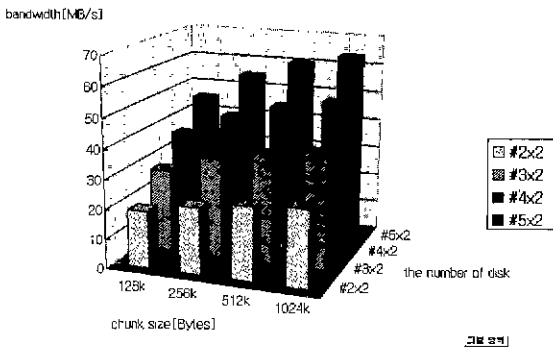


그림 4 SCSI 버스가 두 개이고 파이프라인 모드로 디스크를 구동하였을 때의 성능

는 SCSI 버스 두 개에 디스크를 파이프라인 방법을 사용하여 구동하였을 때의 성능 값이다. 각 실험은 30회 반복하여 평균값의 1%이내의 편차가 나왔다. 그림 3은 그림 2보다 약 1.2배 가량의 성능 향상을 보였다. 그림 4는 SCSI 버스를 두 개 사용하였기 때문에 약 1.8배의 성능 향상을 보였다. 또한 한 번에 요청하는 데이터의 크기(chunk size)가 성능에 큰 영향을 보이는 것을 알 수 있었다. 이 실험에서 Software-to-DEVS 인터페이스를 이용하여 병렬 디스크를 파이프라인으로 구동하는 알고리즘을 개발할 수 있었다. 또한 디스크 모델

이 실제 데이터를 저장할 수 있도록 하여 병렬 디스크에서의 파일 시스템을 개발하는데 이 시뮬레이션 모델을 사용하였다.

### 3.3. 동시 하드웨어/소프트웨어 개발

이제까지의 시스템 개발은 하드웨어가 실제로 제작이 끝난 뒤에 그 시스템에 동작하는 소프트웨어를 만들거나, 하드웨어와 독립적인 소프트웨어만 하드웨어가 구현되기 전에 개발이 가능하였다. 시스템이 복잡해질수록 개발과정의 시스템 오동작이 소프트웨어 때문인지 하드웨어 때문인지 판단하기가 힘들어진다

객체 지향 성능 예측 시스템은 동시 하드웨어/소프트웨어 개발에 다소 도움이 준다. DEVS 모델들은 성능 측정뿐만 아니라 시스템의 동작을 모방하여 동작하기 때문에 이 것이 가능하다. 소프트웨어의 SEND/RECEIVE와 같은 통신 API들은 소프트웨어-DEVS 인터페이스(그림 1의 AbstractCpu 모델에 해당한다)를 통하여 DEVS의 메시지로 변환되어 DEVS 시스템 모델로 전달되어 각 모델이 동작되도록 한다.

## 4. 결 론

큰 대역폭이 필요한 비선형 편집기를 설계하기 위하여 먼저 여러 가지 구조에 대한 모의 실험이 필요하였다. 객체 지향 성능 예측 시스템을 이용하여 그래픽환경에서 쉽게 원하는 구조를 설계하고 변경이 가능하였다. 또한 DEVS 모의 실험 모델은 디스크 및 버스의 복잡한 동작의 표현이 가능하여 정확한 디스크 시스템의 성능을 얻을 수 있었다. 한번에 요청하는 데이터의 크기(chunk size)가 디스크 성능에 매우 큰 영향을 끼치고 성능이 디스크의 수가 늘어도 성능에는 기대만큼의 성능 향상은 없었다 그러나 디스크를 파이프라인 방법으로 구동하게 되면 디스크 수를 늘리는 것보다 훨씬 성능에 좋은 효과를 보이는 것도 알았다. SCSI 버스를 증가 시켰을 때 성능 향상의 정도를 구체적으로 알 수 있었다. 이 모의 실험 결과로써 요구되는 시스템에 가장 적합한 구조를 얻을 수 있었다.

## 참 고 문 헌

- [1] G. R. Ganger, B. L. Worthington, R. Y. Hou, and Y. N. Patt, "Disk arrays", IEEE Computer, pp.30-36, March 1994
- [2] D. Gemmell, H. M. Vin, D. D. Kandlur, P. V. Rangan, and L. A. Rowe, "Multimedia storage servers - A tutorial". IEEE Computer, pp.40-49, May, 1995.
- [3] S. H. Baek, J. S. Hong, K. H. Park, "An Object Oriented Performance Prediction System", submitted to Korea Information Science Society Fall Conference, 1998
- [4] Douglass, Bruce Powel, "Real-time UML : developing efficient objects for embedded systems", Addison Wesley Press, 1998.
- [5] B. P. Zeigler, "Object-Oriented Simulation with Hierarchical, Modular Models". Academic Press, 1990.
- [6] J. S. Hong, "A Formal Method for Object-Oriented Real-Time Software Development", Ph.D Thesis, KAIST, 1998