

# Myrinet 상에서 VI-GM 을 이용한 CFS 의 효율적인 통신기법

\*장기성, \*이주열, \*유관중, \*\*박의수, \*\*최현호, \*\*\*유원경  
\*충남대학교, \*\*보건대학교, \*\*\*성신여대  
\*{zephyr17, esil7153, kjyoo}@micrlab.cnu.ac.kr, \*\*{uspark, hyuno}@hit.ac.kr,  
\*\*\*wyoo@sungshin.ac.kr

## An Implementation of Communication Modules on CFS using of VI-GM On Myrinet

\*Gi-Sung Jang, \*Joo-Youl Lee, \*Kwan-Jong Yoo,  
\*\*Ui-Su Park, \*\*Hyun-Ho Choi, \*\*\*Weon-Kyung Yoo  
\*Chungnam Univ., \*\*Daejeon Health Sciences College, \*\*\*Sungshin Women' s Univ.

### 요 약

클러스터 파일 시스템의 성능은 노드 내부 연산의 성능 못지않게 노드간의 통신 성능이 전체 시스템의 성능에 큰 영향을 미친다. 노드간의 통신 성능을 높이는 방법으로 사용자 수준 통신 프로토콜이 있는데 이 중에서 M-VIA 를 선정하여 노드간 통신모듈을 구현함으로써 운영체제의 간섭으로 인한 네트워크 계층간의 데이터 복사에 의한 병목현상을 줄이고자 하였다. 이를 통해 약간의 성능 우위를 보였다. 하지만 하드웨어의 급속한 발전으로 최근의 클러스터링 파일 시스템에는 Myrinet 이나 ServerNet, QNet, SCI(Scalable Coherent Interface) 등의 고속 인터페이스를 통해 연결하는 것이 일반화 되어 있다. 본 논문에서는 노드간의 초고속 통신을 지원해 주는 Myrinet 환경에서 제공해 주는 사용자 수준의 통신 프로토콜인 VI-GM (Virtual Interface Architecture over GM)을 사용하여 클러스터 파일 시스템의 노드간의 통신모듈을 구현하였다. 그리고 TCP, M-VIA, VI-GM 으로 통신모듈을 구현한 각각의 클러스터 파일 시스템의 성능을 비교 평가하였다.

### 1. 서론

인터넷의 급속한 성장으로 대용량의 데이터를 효과적으로 관리하기 위해 가장 많이 사용되는 방법이 클러스터 파일 시스템을 활용하는 것이다. 클러스터 파일 시스템은 파일 시스템에 클러스터링 기법이 적용된 것으로써 대용량 멀티미디어 파일의 입출력 시간을 단축하고 결함 허용성(Fault Tolerance) 및 고가용성(High Availability)을 제공하는 기술로 주목 받고 있다[1]. 클러스터 파일 시스템이 안정적이고 효율적인 방법으로 대용량의 멀티미디어 데이터를 빠르게 분산 저장하기 위하여 필요한 교환 메커니즘인 통신 모듈은 클러스터 파일 시스템을 구성하는 각 노드의 Daemon 이 다른 노드에서 실행되고 있는 Daemon 과 자료를 주고받을 수 있도록 하여야 한다[1].

기존의 TCP/IP 를 이용한 통신 모듈은 여러 네트워크 계층을 거치면서 생기는 오버헤드와 데이터를 주고받을 때 커널이 직접 개입함으로써 System Buffer 로의 잦은 데이터 복사나 시스템 콜 등으로 클러스터 파일 시스템 전체 성능의 저하를 가져왔다. 이를 위해 본 논문에서는 네트워크 계층을 거치면서 생기는 병목현상과 커널의 관여로 생기는 오버헤드로 인해 생기는 클러스터 파일 시스템의 성능저하를 해결하기 위한 새로운 통신 모듈이

필요하게 되었다[2]. 리눅스 클러스터 파일 시스템의 성능에 큰 영향을 미치는 노드간의 통신모듈을 Fast Ethernet 상에서 Zero-Copy 를 지원해 주는 사용자 수준 통신 프로토콜인 M-VIA 를 이용하여 노드간의 통신 모듈을 설계 및 구현하였다. 하지만 최근의 클러스터 파일 시스템은 Myrinet 이나 ServerNet, QNet, SCI(Scalable Coherent Interface)처럼 고속의 인터페이스 장치를 이용하여 노드간의 통신모듈을 구축하는 것이 일반화 되는 추세이다. 본 논문에서는 Myrinet 상에서 제공해 주는 사용자 수준 프로토콜인 VI-GM (Virtual Interface Architecture over GM)을 사용하여 클러스터 파일 시스템의 노드간의 통신모듈을 설계 및 구현하였다. 그리고 최종적으로 기존의 TCP/IP 통신모듈과 M-VIA 통신모듈 VI-GM 통신모듈을 비교 평가한다[3][5][6][7].

본 논문의 구성은 다음과 같다. 2 장에서는 관련연구를 살펴보고 3 장에서는 기존 시스템과 제안 시스템을 살펴보고 Myrinet 기반의 VI-GM 통신의 설계 및 구현을 하였다. 4 장에서는 구현된 VI-GM 통신 모듈을 실험을 통해 성능을 비교 평가하였다. 5 장에서는 본 논문의 결론 및 향후 연구 과제를 기술한다.