

HCC 단일 가상화 서비스 플랫폼에서 애플리케이션 시험

우준, 이국화
한국과학기술정보연구원
wjnadia@kisti.re.kr, ghlee@kisti.re.kr

A application testing on HCC single virtualization service platform

Joon Woo, Guohua Li
KISTI

요 약

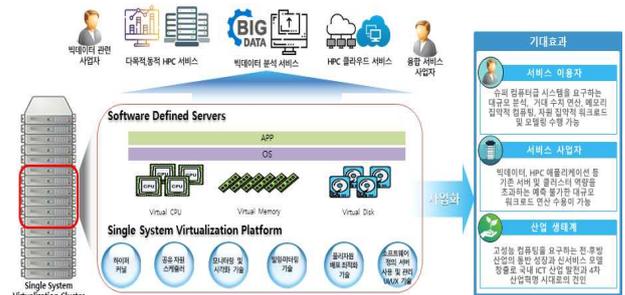
단일 가상화 서비스 플랫폼은 메모리 및 컴퓨팅 집약적 워크로드를 수행하기 위한 고성능 시스템 환경의 신속한 구축을 지원하는 클라우드 기반의 소프트웨어 정의 서버를 위한 핵심 기술이다. 본 연구는 다수의 물리 노드를 통합하여 하나의 고성능 단일가상서버로 구성하기 위해 개발된 HCC 단일 가상화 서비스 플랫폼에서 대용량 데이터 처리 및 대규모 연산이 필요한 NGS 기반 농생명 유전체 조립 프로그램과 이상 기상의 탐지 분석을 위한 GOES 위성자료 전처리 프로그램을 시험하여 활용 적합성을 검증하였다.

1. 서론

단일 가상화 서비스 플랫폼은 지능 정보사회의 신융합분야의 대규모 분석이나 메모리 집약적 컴퓨팅, 컴퓨팅 집약적 워크로드, 모델링을 수행하기 위한 고성능 시스템 환경의 신속한 구축을 지원하는 클라우드 기반의 소프트웨어 정의 서버를 위한 핵심 기술이다. 본 연구에서는 역가상화 기반의 HPC 클라우드 서비스를 제공하기 위해 개발된 HCC(Harmony Cloud Classic) 단일 가상화 서비스 플랫폼에서 바이오 및 기상 분야의 애플리케이션을 시험하여 활용 적합성을 검증하였다.

2. HCC 단일 가상화 서비스 플랫폼

HCC 단일 가상화 서비스 플랫폼[1]은 (그림 1)과 같이 다수의 물리 노드의 하드웨어 자원을 역가상화하여 단일 소프트웨어 정의 서버를 구성한다. 물리 노드의 하드웨어 자원 제약에서 벗어나 사용자가 필요한 자원을 동적으로 구축할 수 있다. 또한, 기존의 Single System Image 기반의 분산 컴퓨팅 서비스와 다르게 사용자마다 독립적인 대용량 계산 수용이 가능한 환경을 신속하게 제공할 수 있다.

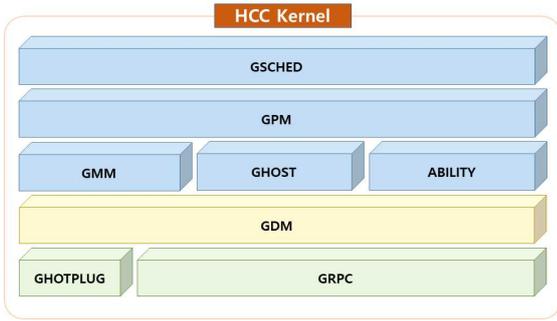


(그림 1) HCC 단일 가상화 서비스 플랫폼 개요

HCC 플랫폼은 하드웨어를 제어하는 운영체제 커널부터 사용자에게 제공되는 포털 서비스까지 종합적인 계층을 아우르는 플랫폼이다. HCC 플랫폼은 크게 HCC 커널 레이어, HCC 미들웨어 레이어, HCC 서비스 레이어로 구성되어 있다.

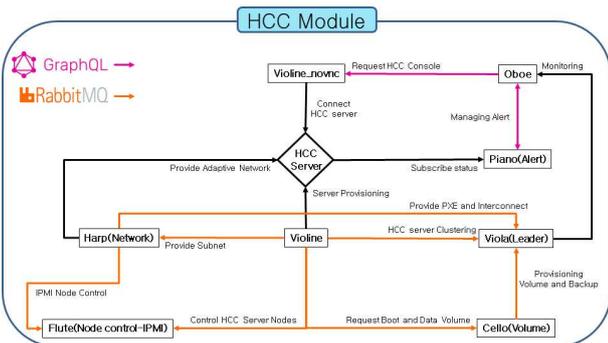
HCC 커널 레이어는 클러스터 기반 멀티 노드 환경에서 SSI(Single System Image)를 구현하는 커널 계층이다. (그림 2)와 같이 HCC 커널 모듈 사이의 원격 프로시저 호출 기능을 제공하는 GRPC와 클러스터 범위의 작업 스케줄링 기능을 제공하는 GSCHED, 클러스터 범위의 메모리 관리 기능을 제공하는 GMM, 클러스터 범위의 파일 관리 기능을 제공하는 GDM, 물리 노드 추가/제거에 대한 자동 인식 기능을 제공하는 GHOTPLUG, 클러스터 범위의 프로세스 관리 기능

을 제공하는 GPM, 메타 데이터 조회 기능을 제공하는 GHOST 등으로 구성되어 있다.



(그림 2) HCC 커널 레이어 구조

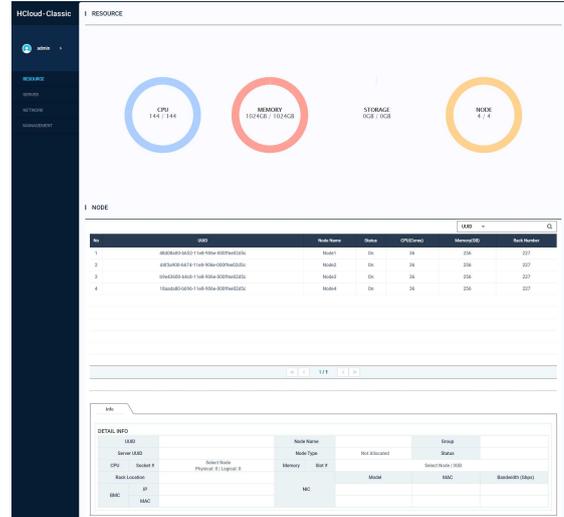
HCC 미들웨어 레이어에서 (그림 3)과 같이 각 모듈들은 마이크로 서비스 형태로 구현되었으며 고유의 기능을 가지고 독립된 데이터베이스를 소유하면서 동작한다. 특정 모듈이 기능을 요청하거나 자원을 요청할 때 해당 기능이나 자원에 대한 접근 권한을 가진 모듈을 통해서 수행된다. 그리고 타 모듈의 데이터베이스에 직접 접근할 수 없으며, 해당 모듈의 기능을 사용하기 위해 메시지 큐 혹은 GraphQL을 통해 API를 호출하여 수행한다. 특정 모듈의 프로시저를 호출하는 메시지들을 보관하는 큐로 RabbitMQ를 사용한다. 그리고, 사용자 역할 관리 및 인증 서비스와 웹 기반 대시보드를 제공하는 oboe, IMPI 제어와 물리 노드를 관리하는 flute, 네트워크를 관리하는 harp, HCC 서버의 상태와 사용량 정보에 대한 모니터링 및 알람 기능을 제공하는 piano 등의 모듈로 구성된다.



(그림 3) HCC 미들웨어의 모듈 개요도

HCC 서비스 레이어는 (그림 4)와 같이 HCC 사용자와 HCC 관리자를 대상으로 HCC 플랫폼 이용 목적의 서비스를 제공하는 계층이다. 대시보드는 사용자가 HCC 서비스를 이용할 수 있는 사용자 친화적인 인터페이스를 제공하는 포털 서비스이다. 단일가상서버를 구성하는 물리 서버들의 구성 자원 현황을 보여주는 Resource, 단일가상서버의 생성 관리 및 모니터링 기

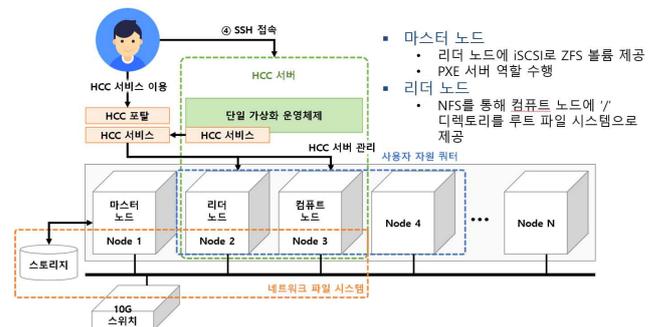
능을 제공하는 Server, 서브넷 등의 네트워크 구성을 지원하는 Network, 사용자 및 쿼터 관리와 과금정보를 제공하는 Management 메뉴로 구성되어 있다.



(그림 4) HCC 서비스 레이어의 대시보드(포털) 화면

3. 애플리케이션 시험

HCC 플랫폼의 활용에 적합한 애플리케이션을 시험하기 위한 테스트베드는 (그림 5) 및 <표 1>과 같이 프로비저닝 서버 역할을 하는 마스터 노드와 마스터 노드에 의해 선택되어 단일가상서버를 형성하는 리더 노드와 컴퓨트 노드들로 구성되어 있다. 리더 노드와 컴퓨트 노드들은 PXE boot에 의해 disk less로 호스트 OS가 부팅된다. 이어서 호스트 OS의 컨테이너 내에서 하이퍼커널이 구동되고 서버 간 상호 연동되어 단일 가상서버가 최종 구성된다. 테스트베드에서 호스트 OS는 HCC 패치 커널을 포함하는 CentOS 6.10을 사용하였다.



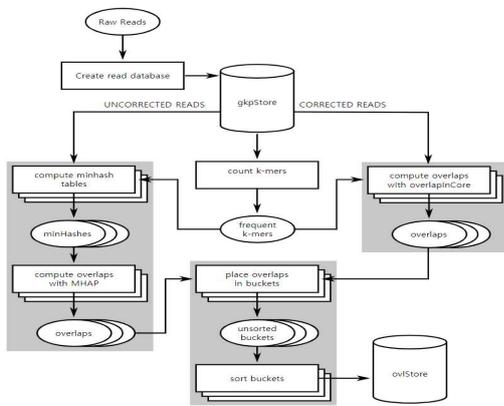
(그림 5) HCC 테스트베드 구성도

<표 1> HCC 테스트베드 사양

구분	사양
서버(4대)	▶ Intel Xeon Gold 6140 2.30GHz 18cores×2ea ▶ DDR4 2,666MHz 256GB
노드 연동네트워크	▶ 10GbE×1ea
Host OS	▶ CentOS 6.10 (Kernel 2.6.32-krq_xeon)
파일시스템	▶ /data(local-xfss,50TB)

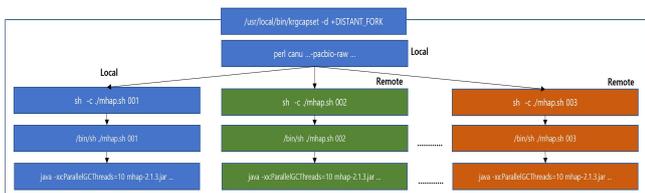
처리해야 하는 데이터 용량 및 계산량이 많아서 대용량 메모리와 다수의 CPU를 장착한 고성능 시스템을 필요로 하여 HCC 활용에 적합한 애플리케이션으로 NGS 기반 농생명 유전체 조립 프로그램 (CANU)과 이상 기상의 탐지 분석을 위한 GOES 위성자료 전처리 프로그램을 선택하여 시험하였다.

첫 번째로, 유전체 조립의 기본 원리는 실험장비를 통해 읽은 대량의 짧은 서열들을 길게 연결하는 작업으로 많은 전산 자원이 필요하다. 본 연구에서는 대표적인 3세대 NGS 장비인 PacBio를 이용하여 단자엽 모델식물인 벼 유전체를 시퀀싱한 중편서열의 조립을 CANU 프로그램[2]을 사용하여 (그림 6)의 워크플로우와 같이 수행하였다.

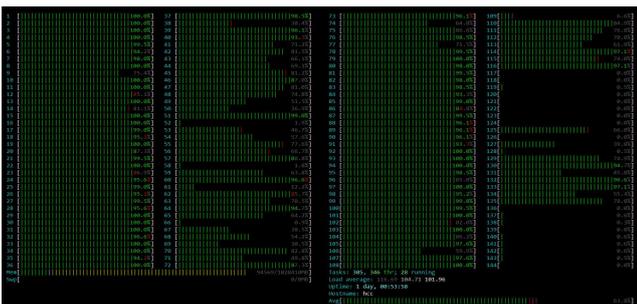


(그림 6) CANU 유전체 조립 과정

(그림 7) 및 (그림 8)과 같이 HCC의 커널 레이어에서 제공하는 주요 기능인 Distant fork[3]를 통해 CANU 실행 시 자식 프로세스가 HCC 단일가상서버를 구성하는 물리 노드 4대에 걸쳐 자동으로 분산 생성되어 정상 동작한다는 것을 확인하였다.



(그림 7) CANU 프로세스 트리 구조도

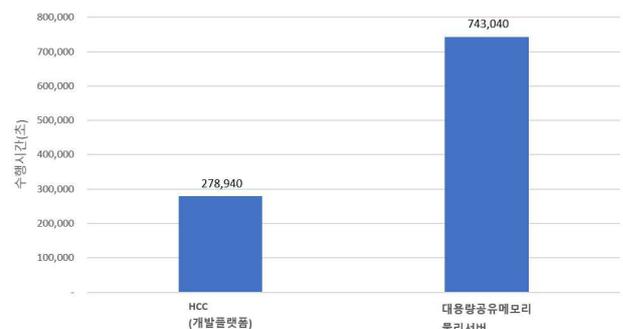


(그림 8) http에서 CPU 사용률 모니터링 화면

144개의 CPU 코어로 구성된 HCC 단일가상서버와 <표 2>와 같이 56개의 CPU 코어와 대용량 공유 메모리를 장착한 물리서버에서 각각 벼 유전체의 조립을 수행한 결과, (그림 9)와 같이 HCC 단일가상서버가 대용량 공유메모리 물리서버 대비 약 2.6배 정도 빠르게 처리하였다. 이와 같이 사용자는 개별 물리 서버의 한계를 넘어 더 많은 CPU 코어와 메모리를 지원하는 HCC 단일가상서버에서 한 대의 고성능 서버를 사용하는 것과 같이 더욱 편리하고 빠르게 벼와 같은 대규모 유전체의 조립을 수행할 수 있다.

<표 2> 대용량 공유메모리 물리서버의 사양

구분	사양
서버	▶ Intel Xeon E7-4830 V4 2.0GHz 14cores×4ea ▶ DDR4 2,666MHz 768GB
노드 연동 네트워크	▶ IB 4x QDR 40bps×1ea
Host OS	▶ CentOS 7.4 (Kernel 3.10.0-693.21.1.el7.x86_64)
파일시스템	▶ /scratch(Lustre, 2.3PB)



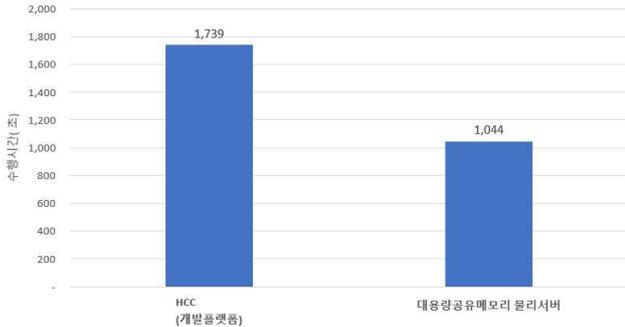
(그림 9) CANU 비교 시험 결과

두 번째로, 대용량의 위성 자료를 기반으로 이상 기상 현상을 탐지 분석하기 위해서는 분석 기간 동안의 위성 자료에서 위성 지정 위치의 경위도와 채널 등을 선택하여 그리드 자료로 변환하는 전처리 작업이 반드시 선행되어야 한다. 미국 1)NOAA는 2)GOES[4]라는 가시 및 적외 채널을 활용한 정지궤도 위성 자료를 공개 제공하고 있다. 2019년에 1년 간의 GOES 위성 자료는 560,640개의 파일과 전체 30.66TB 용량으로 구성되기 때문에, 이를 빠르게 전처리하기 위해서는 다수의 병렬 프로세스를 동시 실행할 수 있는 고성능 계산 시스템이 필요하다.

본 연구에서는 2019년 1~3월까지 3개월간의 GOES 위성자료에 대한 전처리[5][6][7]를 HCC 단일가상서버와 대용량 공유메모리 물리서버에서 수행하였다. 이번 시험에서는 벼 유전체 조립 시험과 다르게 (그림 10)과 같이 대용량 공유메모리 물리서버가 HCC 단일가상

1) National Oceanic and Atmospheric Administration
2) Geostationary Operational Environmental Satellite

서버 대비 약 1.6배 빠르게 처리하였다. 이와 같은 대규모 위성 자료에 대한 전처리에서는 처리할 데이터의 양이 증가하면 파일 I/O 성능이 전처리 성능에 미치는 영향이 더욱 커지기 때문에, <표 3>의 파일 I/O 성능 비교[8]에서 보여주는 바와 같이 소규모 디스크 스토리지로 구성된 HCC 단일가상서버의 공유파일시스템 보다 대규모 디스크 스토리지로 구성된 고성능 병렬 파일시스템인 Lustre를 사용하는 대용량 공유메모리 물리서버에서 더 좋은 성능을 나타낸 것으로 판단된다.



(그림 10) GOES 위성자료 전처리 비교 시험 결과
<표 3> 공유파일시스템 파일 I/O 성능 비교

구분 (파일시스템, 용량)	IOZone 측정 파일 쓰기 성능
HCC 단일가상서버 (XFS, 50TB)	0.98GB/sec
대용량 공유메모리 물리서버 (Lustre, 2.3PB)	4.7GB/sec

4. 결론

본 연구는 다수의 물리 노드의 하드웨어 자원을 통합하여 하나의 단일가상서버로 구성하기 위한 HCC 단일 가상화 서비스 플랫폼에서 대용량 데이터 처리 혹은 대규모 계산이 필요한 NGS 기반 농생명 유전체 조립 프로그램(CANU)과 이상 기상의 탐지 분석을 위한 GOES 위성자료 전처리 프로그램을 시험하였다.

결론적으로, 먼저 지식 프로세스가 HCC 단일가상서버를 구성하는 물리 노드들에 걸쳐 자동으로 분산 생성되고 정상 동작한다는 것을 검증하였다. 또한, 사용자는 개별 물리 서버의 한계를 넘어 더 많은 메모리와 CPU 코어를 지원하는 HCC 단일가상서버에서 한 대의 고성능 서버를 사용하는 것과 같이 더욱 편리하고 빠르게 바와 같은 대규모 유전체의 조립을 수행할 수 있다.

하지만, 대규모 위성 자료에 대한 전처리의 경우, 처리할 데이터 양이 증가하면 파일 I/O 성능이 전처리 성능에 미치는 영향이 더욱 커지기 때문에 HCC 단일가상서버에서 고성능 공유파일시스템의 구성이 필요하다는 것을 알게 되었다.

사사

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.2019-0-00020, 소프트웨어 정의 서버 기반 클라우드 서비스를 위한 단일 가상화 서비스 플랫폼 핵심 기술 개발)

참고문헌

- [1] 우준, 이국화, 최예지, 임상범, HCC 단일 가상화 서비스 플랫폼 시험 대전, KIST, 2020
- [2] CANU, <https://canu.readthedocs.io/en/latest/index.html>
- [3] Kenighed, http://www.kenighed.org/wiki/index.php/Main_Page
- [4] NOAA GOES server, <https://www.goes.noaa.gov/>
- [5] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: Planetary-scale geospatialanalysis for everyone. Remote sensing of Environment 202, 18 - 27.
- [6] Sun, K., Zhu, Y., Pan, P., Hou, Z., Wang, D., Li, W., Song, J., 2019. Geospatial data ontology: the semantic foundation of geospatial data integration and sharing. Big Earth Data 3, 269 - 296.
- [7] Yang, C., Yu, M., Li, Y., Hu, F., Jiang, Y., Liu, Q., Sha, D., Xu, M., Gu, J., 2019. Big earth data analytics: A survey. Big Earth Data 3, 83 - 107.
- [8] IOZone, <http://www.iozone.org>