

HPC 응용의 런타임 실행 패턴에 따른 소프트웨어 정의 계산 구조 분석

오 지 선, 이 효 정, °김 윤 희

An Analysis of Software-defined Compute Architecture for HPC Application Runtime Execution Patterns

Ji Sun Oh, Hyojeong Lee, °Yoonhee Kim

요 약

최근 기기들의 고성능화에 따라 데이터양의 폭증 및 트래픽 동적 변화가 심화되고 더불어 클라우드 컴퓨팅의 확산으로 데이터 서비스의 다양성을 고려한 소프트웨어 중심의 자원 관리가 중요해지고 있다. 대용량 데이터를 고속으로 처리해야 하는 데이터 집약적인 HPC 응용은 실행 요구사항 중 실시간 분석 및 예측을 요구하는 경우 적시적 데이터 처리 및 분석과 결과 검증에 시간제한 요소를 고려하여 소프트웨어 정의 중심의 계산 및 전송구조를 정의하고 실행해야 한다. 본 논문에서는 고성능 데이터 전송 환경인 ScienceDMZ에 대한 국내외 동향과 소프트웨어 정의 기술을 바탕으로 실시간 처리 및 결과 검증 문제를 해결을 요구하는 응용에 대한 실행 패턴을 분석한다. 이를 위하여 기상 예보 계산 모델의 실행 패턴 분석과 시간적인 한계를 가진 고성능 연구를 위한 적시적 데이터 저장과 계산 실험의 통합적 연동의 필요성을 제안한다. 또한 응용 프로그램의 이해를 기반으로 자원 사용 사례 패턴에 따라 실행 자원을 동적으로 할당 및 반환하는 것이 중요하다. 본 논문은 데이터 집약적인 HPC 응용 사례로 WRF 모형을 선정하여 수집된 입력 데이터의 변화에 따른 분석의 런타임 특성을 분석한다. 분석 결과를 통해 적시적으로 데이터를 할당하는 HPC 응용을 위한 소프트웨어 정의 계산 구조를 정의하고 중요성을 제고하였다.

Key Words : HPC, ScienceDMZ, SDN, SDC, WRF

ABSTRACT

In the advent of high performing and sophisticated devices, explosion of large-scale data and dynamic variation of traffic change are intensifying. Cloud computing enables software-based resource management considering diversity of data services over on-demand computing infrastructure. As real-time analysis of large-scale data and forecasting of data-intensive HPC application are required, software definition technology should be applied to temporal data processing, analysis, and result verification. In this paper, we have investigated trends and software definition technologies of ScienceDMZ, which provides high speed data transmission environment for scientific big data. In order to solve the problem of real-time processing and result verification, we propose a temporal application and integration of data storage and calculation experiment for high performance research with time limit through weather forecast infrastructure. It is also important to identify resource use case patterns based on the understanding of application programs and assign them appropriately. This paper analyzes the runtime characteristics of input data by selecting WRF model as a data - intensive HPC application. We defined the execution structure of the HPC application that allocates the data in a timely manner through the analysis results and raised the importance.

※이 논문은 2017 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017R1A2B4005681)

° Corresponding Author : Dept. of Computer Science, Sookmyung Women's University, yulan@sookmyung.ac.kr

논문번호 : KNOM2017-01-04, Received June 13, 2017; Revised July 20, 2017; Accepted August 10, 2017

I. 서론

소프트웨어 정의 기술은 IT 인프라를 동적으로 관리하여 응용 실행의 최적화를 위한 적시적 자원 관리 기술을 제공한다. 최근 서비스 특성에 따라 효율적인 네트워크 구조 개선을 위한 OpenFlow^[1]와 같은 오픈소스 기반의 SDN(Software-defined Network) 기술을 적용한 동적 네트워크 관리 연구가 활발하다. 이와 함께 기기들의 고성능화에 따른 데이터의 폭증, 트래픽 동적 변이 심화와 더불어 클라우드 컴퓨팅의 확산과 서비스의 다양성을 고려한 소프트웨어 중심의 관리가 중요해지고 있다. 이러한 대규모의 데이터를 처리하는 과학 계산 응용의 요구사항을 충족시키기 위해 SDN을 시작으로 스토리지, 인프라 그리고 데이터센터 전체 영역으로 확산되었다. 고성능 환경을 활용한 SDC(Software-defined Compute), SDS(Software-defined Storage) 기술이 새롭게 정의되고, 관련 연구가 활발히 진행 중이다.

기존의 HPC연구는 필요 자원이 독점적으로 활용 가능한 시점에 분석을 진행하는 일괄 배치 작업 형태를 활용하고 있다. 그러나 기상 재난 예측과 같은 과학계산 응용은 일괄 배치 처리로 끝나기 어렵고, 연속적인 여러 단계를 거쳐 장시간 처리 되는 경우가 많다. 이러한 응용은 실시간에 가까운 분석 및 결과 활용을 요구하기 때문에 런타임 특성에 따라 필요로 하기 때문에 자원의 동적 처리를 요구한다. 그러나 과학계산 응용의 자원 및 실행 인프라의 실시간 성능 처리와 결과 검증 문제에 대한 연구가 부족한 상황이다. 따라서 전송, 보안, 인프라의 실시간 성능 처리 및 결과 검증 문제를 해결하여 시간적 제한을 가지고 결과를 도출해야 하는 고성능 연구 환경이 필요하다.

본 논문은 국내외에서 진행되고 있는 고성능 데이터 전송이 가능한 ScienceDMZ 환경 연구에 대한 조사와 관련 소프트웨어 정의 기술(SDN, SDC)기술에 대해 설명하였다. 실시간 예측 분석의 일례로 기상 재난 예측 분석 시나리오를 소개하고 실시간 침수 예보를 가능토록 하는 소프트웨어 정의 인프라 구조를 제안하였다. 시나리오를 바탕으로 대기 연구 및 운영에

측 응용을 활용하여 실시간 예측 분석 성능에 영향을 미치는 요소를 파악하기 위한 실험을 진행하였다. 이를 바탕으로 과학응용 데이터 집중형 응용을 효율적으로 수행하는 인프라를 구축하기 위하여 소프트웨어 사용자 정의 기술인 SDC, SDN을 적용하여 대용량 데이터 처리가 필요한 HPC 응용 실행 구조를 제안한다.

본 논문의 나머지 부분은 다음과 같이 다섯 개의 순서로 구성되어 있다. 제 2장에서는 관련 연구에 대해 소개하고 제 3장에서는 기상 재난 예측 시나리오를 통해 실시간 예보 인프라 구조를 설명한다. 제 4장에서는 기상 응용의 실시간 예측 성능에 영향을 미치는 요소 분석 실험과 그 결과를 보인다. 제 5장에서는 실시간 과학 응용을 위한 소프트웨어 정의 HPC응용 실행 구조를 제안한다. 마지막 결론으로 향후 구축할 인프라에 대한 계획을 포함한다.

II. 관련 연구

2.1. ScienceDMZ

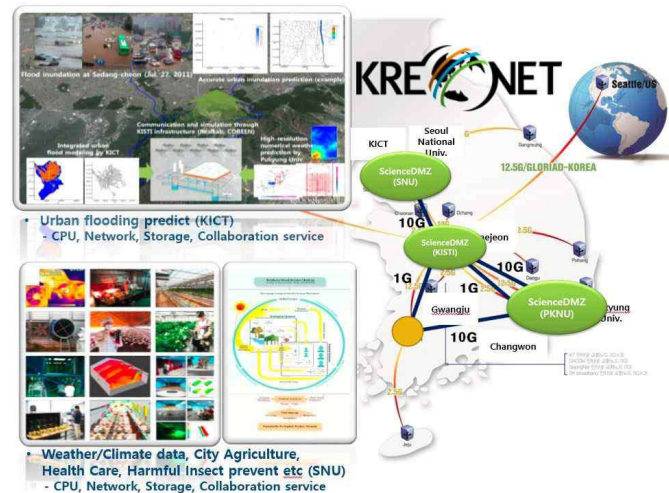


그림 1. 국내 대기과학분야 ScienceDMZ기반의 고성능 인프라^[2]

Fig. 1. High-performance infrastructure based on ScienceDMZ in Korea's atmospheric science field^[2]

미국 ESnet에서 제안된 ScienceDMZ는 데이터 집중형 과학을 위한 통합적 전송체제이며 개념적 용어이다. TCP환경에서의 최적화된 전송 성능의 보장을 위하여 트래픽 손실 없는 네트워크 상태를 추구한다. 망 분리는 초고속

전송 환경 구성을 위해 일반적인 데이터 트래픽과 ScienceDMZ 데이터 트래픽에 의해 이루어진다. 트래픽 손실이 없는 경우 LAN과 international 구간에 대해 Line Rate의 90% 이상의 전송 성능을 나타낸다. 또한 고성능 데이터 전송을 위해 저장, 캐싱, 보안 기능까지 감안한 전용 노드(DTN)가 필요하다. 데이터 전용 전송 노드 구축을 통해 과학데이터 전송 환경을 구축한다. 또한 전송 성능을 보장하기 위해 ScienceDMZ간 연결된 DTN의 성능을 실시간으로 검증하여 성능을 보장한다^{2)[3]}. 이를 기반으로 고성능 데이터 전송 서비스를 확충하여 고성능 연구 환경을 제안하였다. 그림 1은 국내 ScienceDMZ기반의 고성능 인프라이다.

2.2. SDN, SDC

SDN 기술은 이중의 스위치와 라우터를 통해 손쉽게 네트워크를 제어할 수 있는 OpenFlow에 기반하여 출현하였다. 소프트웨어 중심의 네트워킹 기술 개발을 위해 네트워크를 컴퓨터 시스템으로, OpenFlow는 하드웨어와 네트워크 운영체제 사이를 연결하는 인터페이스로 하여 네트워킹 시스템 모델을 정의한다. SDN 기술은 네트워크 사용자가 하드웨어 기반 연구에서 벗어나 다양한 응용들을 손쉽게 개발할 수 있는 환경을 제공할 것이다⁴⁾. SDC 기술은 SDN의 개념에서 확장되었다. 데이터센터의 새로운 트렌드 기술로 데이터와 자원을 작업에 올바르게 배치하는데 있어서 작업의 동적인 수행을 고려한다⁵⁾. 응용 실행의 최적화를 위한 자원관리 기술인 SDC를 통해 데이터센터의 자원을 효율적으로 사용할 수 있다.

III. 기상 재난 예제 시나리오

고성능 전송, 보안, 인프라의 실시간 성능 처리 및 결과 검증 문제가 해결되지 않는 기존 환경에서 시간적인 한계를 가지고 결과를 도출해야 하는 고성능 연구는 ScienceDMZ와 같은 인프라 구축이 필요하다. 응용의 예로는 국지성 태풍 예보와 같은 기상 재난 안전 관련 분석, 예측의 경우이다.

기상청은 정기적인 예보를 위해 과거 데이터

를 바탕으로 다년간 개발된 기상모델을 이용하여 슈퍼컴퓨터와 같은 고성능 자원을 활용한 단기, 중기, 장기 예보를 실시하고 있다. 그러나 갑작스러운 기상 변화에 대한 분석 및 예측 서비스는 시간적 제약 하에 어려움이 있다. 그림 2는 실시간 침수 예보를 가상화 하여 제공하는 인프라 구조를 나타낸다. 과거의 재난, 지형, 기상정보와 변동하는 기상 데이터를 위성 및 레이더로 측정한다. 데이터 고속전송 및 분석을 위한 DMZ 기술을 통한 ICT인프라를 구축하여 지형 및 광역, 국지 구름 모델 데이터를 전송한다. 데이터를 통해 기상 재난 상황의 원인 분석 및 모델을 구현하며 분석 결과를 통해 강우 예상 데이터를 전처리하고 기상 모델에 적용한다. 과학적 모델 기반 기상재난을 예측하여 개선방안을 제시하고 재발 방지를 위한 정보 인프라를 제공한다.

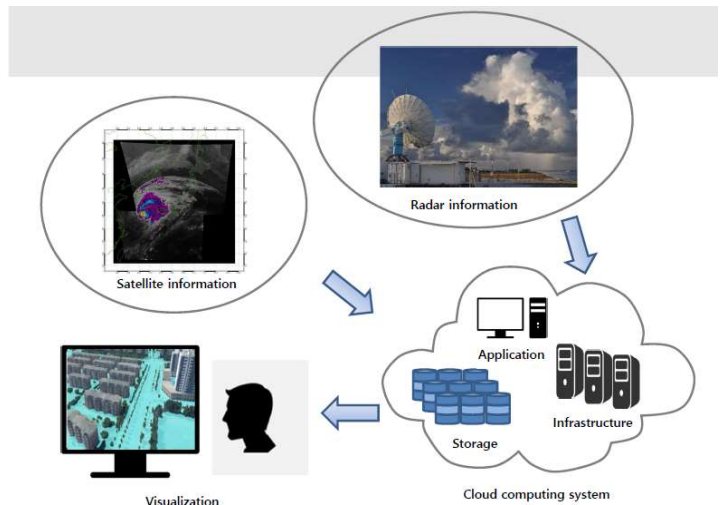


그림 2 . 실시간 침수 예보 인프라^[6]
Fig. 2. Real-time flood forecasting infrastructure^[6]

이처럼 국지적 태풍 예보, 도로별 침수, 홍수 지역 예측 서비스 등과 같은 실시간 분석 및 적용 연구는 대용량 데이터의 이동, 충분한 고성능 자원, 새로운 분석 모델 개발 등을 필요로 한다. 대용량의 기상, 기후 데이터의 전송, 저장, 처리가 필요하며 데이터 분석 적용을 위한 데이터 흐름 연구가 필요하다. 고성능 데이터 전송이 가능한 ScienceDMZ 환경에서 데이터 저장과 계산 실험의 적시적 적용 및 통합적 연동이 가능한 상황을 제공해야 한다. 또한 응용

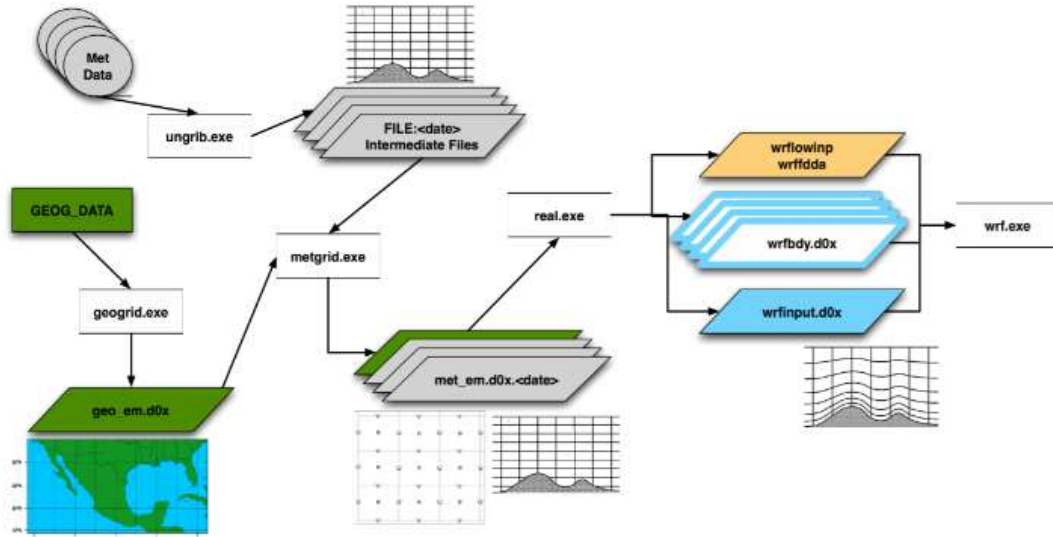


그림 3 . WRF 응용 실행 구조도^[8]
 Fig. 3. Execution structure of WRF application^[8]

프로그램의 이해를 기반한 지능적 데이터 배치와 계산 배치가 우선적으로 필요하다.

IV. 대기 예측 모델 실행 부하 실험 및 분석

대기 연구 및 운영 예측 응용인 WRF(Weather Research and Forecasting)^[7]는 전처리 단계인 WPS(WRF Pre-processing System)와 모형 본체인 WRF로 구성되어 이상실험(ideal case) 또는 실제 기상 환경에 대한 실험과 예보(real case)를 실시한다.

데이터 집약적인 응용 실험에 앞서 WRF 응용의 특성을 분석하였다. 그림 3은 WRF 응용 실행 구조도이며 이에 따른 실행 과정은 다음과 같다. WPS에서는 WRF 본체에 사용될 도메인과 변수를 보간 한다. 그림 3의 geogrid.exe, ungrrib.exe와 metgrid.exe 모듈이 이에 해당된다. geogrid.exe 모듈은 지구 데이터 GEOG_DATA를 입력 받아 예보를 수행할 도메인을 정의한 geo_em_d0x 파일을 출력한다. ungrrib.exe 모듈은 해당 지역의 기상 데이터 Met Data에서 WRF 모델 구동에 사용되는 변수를 추출하고 이를 이진 파일로 표현한다. metgrid.exe 모듈은 geogrid.exe 모듈이 정의한 도메인에 ungrrib.exe 모듈이 추출한 변수를 보간 한다. 이러한 전처리 과정이 완료되면 WRF 본체의 모듈 중 real.exe가 모델 구

동에 사용될 초기장 wrfinput 파일과 경계장 wrfbdy 파일을 생성한다. 초기장과 경계장을 입력으로 wrf.exe 모듈이 적분으로 이루어진 실제 예보 연산을 수행한다.

본 실험에서는 AWIP(All Weather Insulated Panels)에서 제공하는 미국 동부 해안의 겨울 폭풍 데이터 재분석 자료를 사용하였다. 적분기간은 2000년 1월 24일 12시부터 1월 25일 12시까지 총 24시간이며, 측정 도메인 크기는 2190×1800km²이다. 실험환경은 Intel(R)Core(TM)i7 CPU950@3.07GHz, 8core, RAM26GB 이며, WRF version은 3.6.1을 사용하였다.

표 1. 해상도 변수
 Table 1. Resolution variable

Contents	Description
E_WE	동서 격자의 수
E_SN	남북 격자의 수
DX	동서 격자의 크기(m)
DY	남북 격자의 크기(m)

위 AWIP 재분석 자료 기준으로 실제 연산이 이루어지는 WRF 모형 본체의 수행 시간이 WRF 모델 총 수행 시간의 약 94퍼센트를 차지한다. 따라서 본 연구의 데이터 집약적인 응

용 실험을 위해 WRF모델 본체에 사용되는 도메인의 해상도를 변경하였다. 표 1은 도메인의 해상도를 조절하기 위한 변수를 나타낸다.

표 2. 연구에 사용된 실험 디자인
Table 2. Experiment design used in this study

Experiment	E_WE, E_SN	DX, DY (m)
30km	74, 61	30000, 30000
20km	111, 92	20000, 20000
10km	222, 183	10000, 10000
5km	444, 366	5000, 5000
3km	740, 610	3000, 3000

도메인의 동서 길이는 $(E_WE-1)*DX$, 남북 길이는 $(E_SN-1)*DY$ 이다. 도메인의 기존 크기를 유지하면서 고해상도로 조절하기 위해 DX, DY의 값을 감소시키고, E_WE와 E_SN 변수 값을 증가시켜 실험하였다. 실험은 총 5가

지로 진행하였으며 각 실험의 변수 값은 표 2와 같다.

WRF 모형 본체의 각 모듈의 수행 시간 및 메모리 점유율 사용량은 glances 툴을 사용하여 측정하였다. 그림 4는 real.exe와 wrf.exe의 수행 시간 및 최대 메모리 사용량을 보여준다. 격자 크기에 따른 각 모듈의 수행 시간은 격자 크기가 작아져 해상도가 높아질수록 증가한다. 격자 크기 30km, 20km, 10km의 경우 real.exe 모듈에서 최대 5.7초의 수행 시간이 소요되고 메모리 사용량이 감지되지 않을 정도로 미미하여 수행 시간에 영향을 미치지 않는다. 반면 5km의 경우에 92.3%의 메모리 사용량을 보이고 수행 시간은 142초로 급격히 증가한다. wrf.exe 모듈의 경우에 격자 크기 30km, 20km, 10km의 각각 최대 메모리 사용량은 4.6%, 8.8%, 30%이며, 5km의 경우 최대 92.7%이다.

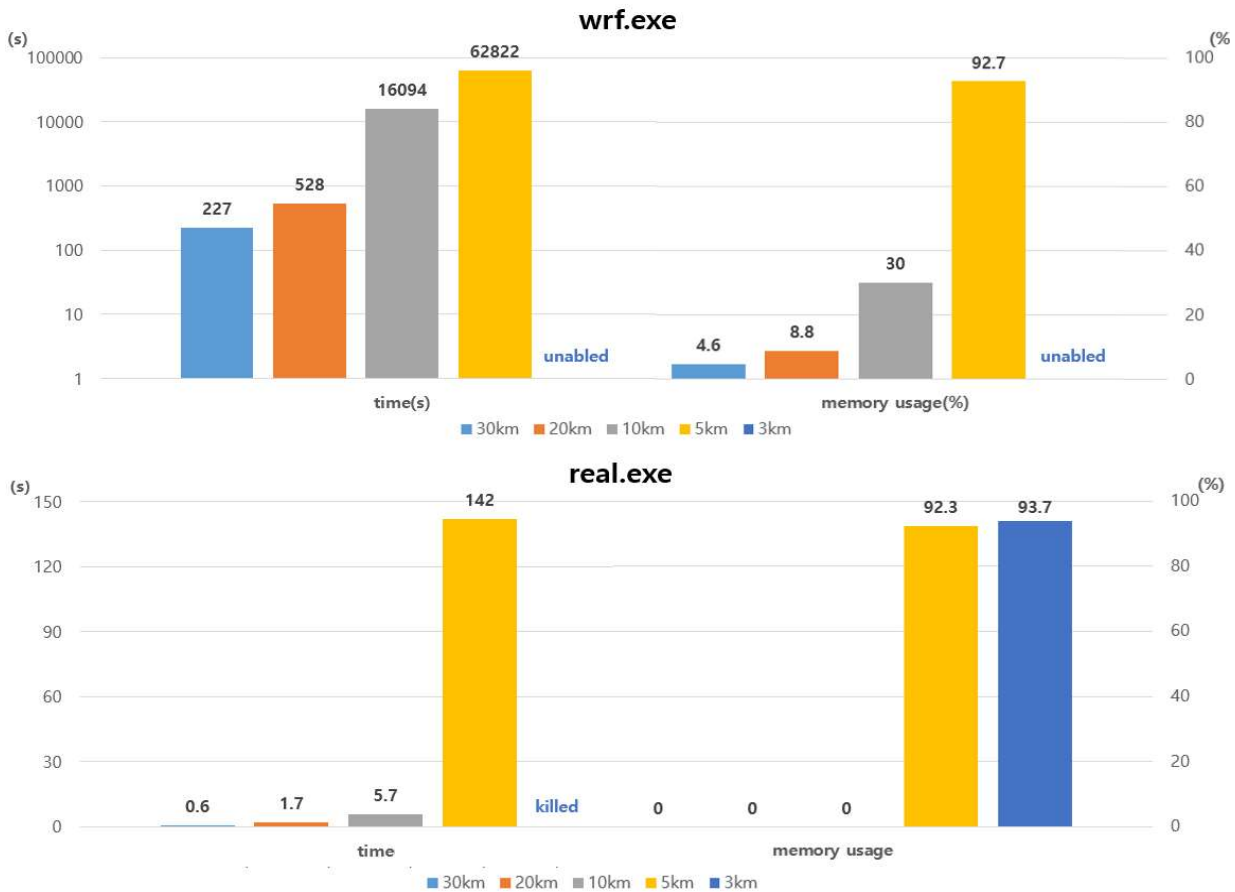


그림 4 . real.exe와 wrf.exe의 시간 및 최대 메모리 사용량
Fig. 4. Time and maximum memory usage of real.exe and wrf.exe

수행 시간 또한 1074분 2초(62822초)로 real.exe 모듈의 경우와 같이 급격히 증가함을 확인하였다. 3km의 경우에는 real.exe 모듈 단계에서 메모리를 최대 93.7%까지 사용하면서 수행 도중 비정상적으로 종료되었다.

본 실험 결과에 따르면 WRF 모형 본체에 사용되는 입력 데이터의 부하가 커질수록 수행 시간과 메모리 사용량이 급격히 증가함을 알 수 있다. 따라서 데이터 집약적 예보 수행을 위해서는 메모리 자원을 적시적으로 활용하는 인프라 구축 필요하다.

V. HPC 응용 실행을 위한 소프트웨어 정의 컴퓨트(SDC) 구조

본 논문에서는 앞 장의 응용 시나리오에 적합한 소프트웨어 정의 기술을 활용하여 응용의 특성을 고려한 실시간 HPC응용 실행 구조를 정의하였다.

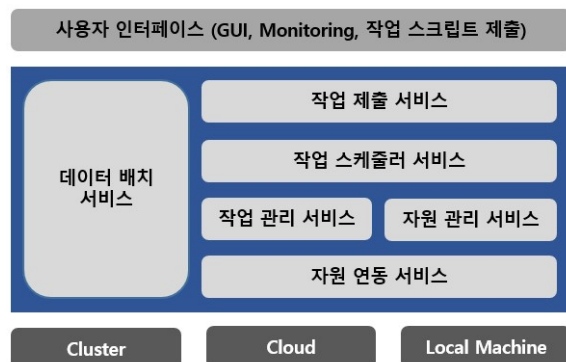


그림 5 . 소프트웨어 정의 컴퓨트(SDC) 구조
Fig. 5. Software defined compute (SDC) structure

그림 5는 본 논문에서 제안하는 기능 구조도이다. 이 구조는 차세대 고성능 인프라에서 발생하는 데이터 전송 및 계산 배치의 문제를 효과적으로 해결할 수 있도록 데이터 집약적인 응용을 위한 데이터 배치 및 작업 스케줄러를 포함한다. 소프트웨어 정의 컴퓨트(SDC) 구조는 다음 단계와 같은 작업으로 진행된다. 사용자는 사용자가 지정한 작업 제출 내용이 명세 되어 있는 파일을 사용자 인터페이스를 통해 제출한다. 작업 제출 모듈에서 제출된 파일의 구문을 분석하고 응용 프로그램의 유형, 실행 파일 경로, 매개 변수, 종속성 등과 같은 정보를 추출한다. 파일에 지정된 매개 변수 범위에 따라 작업을 생성한 다음 해당 작업의 정보와 함께 작

업 관리 모듈로 전달한다. 작업 관리 모듈은 스케줄링 정책을 기반으로 작업들을 스케줄링하고 자원 관리 모듈로 제출한다. 자원 관리 모듈은 할당된 여러 유형의 자원에 따라 사전에 설정된 계획을 통해 가상머신을 실행한다.^[9]

플랫폼의 작업 제출 및 작업 스케줄러 서비스는 응용의 특성 분석을 통해 찾아낸 각각의 특징 정보를 이용하여 세밀한(fine-grained) 자원 요구를 만족시키는 SDC 기술 기반의 작업 스케줄링 기법을 제공한다. 예를 들어 4장의 실험을 통해 기상 재난 예측 모델의 HPC 응용은 메모리 자원 집중 사용 패턴을 보여주었다. 이러한 응용의 메모리 자원 사용 패턴 분석을 기반으로 메모리 자원 재구성 기법을 수행할 경우, 필요 시점에 동적으로 메모리 할당과 반환이 가능하다^[10]. 이를 바탕으로 응용의 “bursty” 자원 요구가 있는 실행 단계에서 메모리 유휴 자원 사용을 최대화하여 수행할 수 있도록 한다. 또한 메모리 자원의 조기 반환 시에는 대기 작업에게 유휴 상태의 메모리를 동적으로 조기에 할당할 수 있으므로 대기 작업의 수 및 시간을 줄이고 시스템 작업 자원 사용 효율을 극대화할 수 있다.

기상 재난과 같은 HPC 응용은 대규모의 입력 데이터, 중간 데이터와 출력 데이터를 생성한다. 데이터의 생성 시점 및 종류에 따라 사용되는 용도와 사용 빈도가 다르다. 따라서 데이터의 생성 시점과 종류를 분류하고 그 결과에 따라 관리, 사용한다. 이러한 데이터 배치 기법을 통해 응용에 사용되는 데이터 특성 이해를 기반으로 하여 지능적으로 데이터를 분배, 배치해야 한다. 데이터 집약적인 HPC 응용의 데이터와 작업 간의 관계 정보를 통해 데이터 이동을 최소화하는 데이터 배치 서비스를 제공하는 것이 한 예이다. 또한 데이터 배치 전략에서 성능 하락의 큰 요인인 네트워크 병목현상으로 인한 지연을 최소화 한다. SDN환경에서의 데이터 움직임을 최소화하기 위해 네트워크 전송 속도 및 대역폭을 고려한 서비스를 제공한다.

VI. 결 론

본 논문에서는 HPC 응용 실행을 위한 소프트웨어 정의 컴퓨트 구조 대해 정의하였다. 응

용의 특성과 데이터의 이동을 고려하면서 효과적인 스케줄링을 제공하여 자원의 효율성 및 활용성을 극대화 할 수 있다. 또한 ScienceDMZ와 같은 국가적 차원의 고성능 데이터 전송체계에서 데이터배치 기법은 데이터 집약적인 HPC 응용 과학 연구를 위한 기술 발전으로 이어진다. 현재 그리고 향후 국가적으로 추진되고 있는 다양한 과학 응용 개발 사업에도 원천기술로써 적극적으로 활용이 가능하다.

References

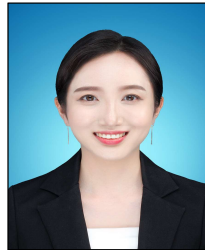
- [1] <http://www.openflow.org/videos>
- [2] 문정훈, and 이민선. "데이터 집중형 (Data-Intensive Science) 과학을 위한 ScienceDMZ 기반 빅데이터 전송 연구." 한국정보과학회 학술발표논문집 : 1359-1361, 2016
- [3] Dart, Eli, et al. "The science dmz: A network design pattern for data-intensive science." Scientific Programming 22.2: 173-185, 2014
- [4] ETRI, "미래 네트워킹 기술 SDN," 2012 Electronics and Telecommunications Trends.
- [5] Nam, Yoonsung, et al. "Workload-aware resource management for software-defined compute." Cluster Computing 19.3 : 1555-1570, 2016
- [6] Seungsoo Lee, "Development and Application of 1D-2D Coupled Urban Inundation Model for Real-time forecasting based on Remote Sensing Technique", KNOM tutorial 2016
- [7] WRF, <http://www2.mmm.ucar.edu/wrf/users/>
- [8] <http://www2.mmm.ucar.edu/wrf/users/>
- [9] 안주림, et al. "과학 응용 실험을 위한 하이브리드 자원 활용 소프트웨어정의 컴퓨팅 (SDC) 플랫폼 설계." 한국정보과학회 학술발표논문집 : 1499-1501, 2017
- [10] Adufu, Theodora, Yoonhee Kim. "Dynamic Memory Allocation for Scientific Workflows in Containers." 정보과학회논문지 44.5 : 439-448, 2017

오 지 선 (Ji sun Oh)



2017년 숙명여자대학교, 컴퓨터과학과 학사
 2017년 ~ 현재 숙명여자대학교, 컴퓨터과학과 석사과정
 <관심분야> 분산시스템, 클라우드 컴퓨팅

이 효 정 (Hyojeong Lee)



2014년 ~ 현재 숙명여자대학교, 컴퓨터과학과 학사과정
 <관심분야> SDN, 빅데이터 처리 기술, 기계학습

김 윤 희 (Yoonhee Kim)



1991년 숙명여자대학교 학사
 1996년 Syracuse University 전산학과 석사
 2000년 Syracuse University 전산학과 박사
 1991년 ~ 1994년 한국전자통신연구소(ETRI) 연구원
 2000년 ~ 2001년 Rochester Institute of Technology 컴퓨터공학과 조교수
 2001년 ~ 2004년 숙명여자대학교 컴퓨터과학과 조교수
 2004년 ~ 2009년 숙명여자대학교 컴퓨터과학과 부교수
 2009년 ~ 현재 숙명여자대학교 소프트웨어학부 교수
 <관심분야> 클라우드 컴퓨팅, 그리드/클라우드 관리