홍수예보시스템을 위한 분포형 유역유출모형 기반 HPC 프레임워크 설계

이영준*, 이정용*, 고덕구*
*동부엔지니어링(주)
e-mail: dudwnsbug@dbeng.co.kr

Design of HPC Framework based on Grid based Rainfall-runoff Model for Flood Forecasting System

Young-Jun Lee*, Jeong-Yong Lee*, Deuk-Koo Koh*
*Water Resources Dept. GIS Team, Dongbu Engineering Co., Ltd.

요 약

오늘날 우리나라의 강수량은 대부분 하절기에 집중되고 있으며 국지성 집중호우에 의한 돌발홍수 발생으로 인명피해가 크게 발생하고 있는 실정이다. 이를 대비하기 위해서는 빠른 시간 안에 좁은 지역에 대한 기상정보 분석이 필요하다. 본 논문에서는 보다 효과적인 기상예보를 위하여 홍수예보모델을 기반으로 모의테스트를 진행 후 결과 분석을 바탕으로 고성능 컴퓨팅(HPC) 기반 홍수예보시스템을 설계하였다.

1. 서론

세계적으로 기후변화로 인한 자연재해 발생빈도가 증가하고 있으며 우리나라 또한 이에 대한 대책마련이 필요한 실정이다. 특히 우리나라의 경우는 연강수량이 대부분 하절기에 집중되고 있으며 이에 따른 국지성 집중호우로 돌발홍수 발생하여 인명피해가 발생하고 있다. 이러한 집중호우 및 홍수는 빠른 예보가 필수적이며 이를 위한 시스템 구축이 필요하다.

본 논문에서는 보다 빠르고 정확한 홍수예보를 위하여 고성능 컴퓨팅(HPC) 기반 홍수예보시스템을 설계하고자 한다. 이를 위하여 단일 클러스터 환경에서 홍수예보모델에 대하여 모의테스트를 실시하고 수행시간과 기상예보를 위한 최소 소요시간을 비교분석하여 홍수예보시스템에 적합한 HPC를 설계하고 적합한 병렬처리 개발방안을 제시하였다.

2. 홍수예보모델

많은 홍수예보모델 중 한국건설기술연구원에서 개발한 분포형 유역유출모형(Grid based Rainfall-runoff Model)을 이용하였다. GRM은 물리적 분포형 강우-유출 모형으로 격자형으로 공간분포 된 수문지형학적 입력자료와 강우자료를 이용하여 실시간 유출해석을 수행한다. 돌발홍수를 대비한 실시간 홍수예보를 위해서는 최대한 빠르고 좁은 간격범위의 유출해석 결과가 나올수록 효용성이 높다.

* 본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음(과제번호 18AWMP-B127555-02)

본 논문에서 사용되는 분포형 유역유출모형은 .NET 기반, Windows 환경으로 병렬코딩 기법를 사용하여 개발되었다. 입력자료는 낙동강권역의 100m 격자 모델과 낙동강권역 중권역 중 하나인 금호강유역의 500m 격자 모델을 사용하였다. 또한, 실시간 홍수예보를 위해서 각 모델별 24시간 기준으로 2분 이내에 홍수예보모델 수행이 완료됨을 기준으로 삼았다.

3. 모의테스트 환경구성 및 수행방안

모의테스트를 위하여 Windows 운영체제 기반의 테스트 베드를 구축하였으며, 단일 클러스터 상 최적의 CPU 구성을 확인하기 위하여 2.4Ghz, 20 Core의 CPU 2개를 장착하였다.

<표 1> 테스트 베드 구축 환경

OS	CPU	Memory	I/O
Windows	20Core * 2CPU	128GB	10 Gigabit

모의테스트는 낙동강권역의 500m 격자 모델과 중권역 금호강유역의 100m 격자 모델을 대상으로 GRM이 실행되었을 경우를 산정하였다.

Core 수 사용을 제한하여 10, 20, 40 Core 환경에서 낙동강권역의 500m 격자 모델에 대한 모의테스트 소요시간을 측정하고 금호강유역의 100m 격자 모델도 같은 방법으로 테스트하여 소요시간을 측정값을 비교하였다.

4. 모의테스트 결과

낙동강권역 24시간/ 500m 격자 모델을 대상으로 CPU의 수행 시 사용되는 Core 수를 다르게 설정하여 모의테스트를 수행하여 수행시간을 측정하고 전체의 평균값을 계산하였다.

<표 2> GRM 수행시간 (낙동강권역, 24시간 기준)

구 분	1차	2차	…	평균
10 Core	45초	45초		44.2초
20 Core	38초	37초		37.8초
40 Core	39초	38초		38.2초

테스트 결과 수행시간은 모두 2분 이내로 조건을 충족하였으며, 20 Core에서 37.8초로 가장 빠르게 측정되었다. 한편으로 2 CPU를 사용하는 40 Core 환경에서는 20 Core에서보다 오히려 속도가 느리게 측정되었다.

금호강유역의 24시간/ 100m 격자 모델을 대상으로 같은 방법으로 모의테스트를 수행하여 수행시간을 측정하고 결과를 낙동강권역과 비교하였다. 그 결과 계산된 소요시간은 낙동강권역과 마찬가지로 20 Core에서 7분 48초로 가장 빠르게 측정되었다.

40 Core에서 오히려 수행시간이 길어지는 이유로는 프로그램에서 수치계산시 필요한 쓰레드보다 많은 쓰레드가 할당되어 대기시간이 길어지고 쓰레드의 시작과 종료 등의 비용이 높아져 전체 성능의 저하를 가져온 것으로 판단되며, 이는 GRM의 수치계산이 20개의 쓰레드에서 최적화되어 있음을 의미한다.

<표 3> GRM 수행시간 (금호강유역, 24시간 기준)

구 분	1차	2차	…	평균
10 Core	8분 36초	8분 32초		8분 34초
20 Core	7분 48초	7분 49초		7분 48초
40 Core	8분 33초	8분 32초		8분 33초

GRM을 수행 시 24시간 기준으로 소요시간은 2분 이내가 이상적이나, 계산결과 금호강유역 100m 격자 모델의 경우 2분을 초과하였다.

단일 병렬화 환경에서는 CPU 성능과 수를 늘리는 방법으로는 500m 격자보다 높은 해상도에서 2분 이내의 소요시간을 기대할 수 없으며, HPC 클러스터를 구성하여 금호강유역의 소유역들에 대한 유출해석 계산을 병렬로 수행한다면 계산시간 단축이 가능할 것으로 판단된다.

5. HPC 프레임워크 설계

HPC 클러스터는 일반적으로 여러 대의 노드로 구성되어 단일 클러스터 환경에 비해 계산 및 처리속도가 뛰어나지만 계산 자원 및 네트워크 속도를 고려한 효율적인 설계가 필요하다.

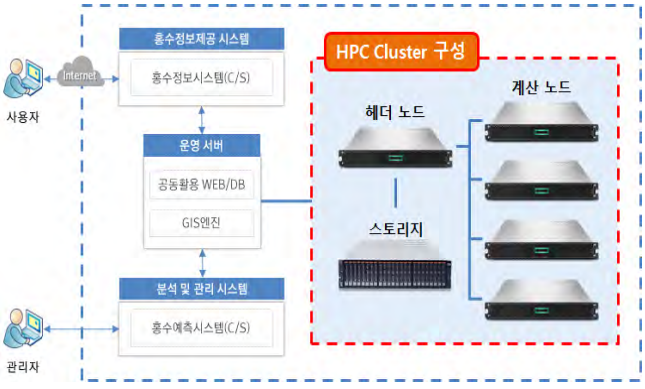
본 논문에서는 HPC 클러스터를 관리하기 위한 운영서버를 구성하였고, 클라이언트 시스템은 운영서버를 통하여 HPC 자원을 사용할 수 있도록 설계하였다.

HPC 클러스터는 약 6분정도의 시간단축이 필요하며, 이는 단일 클러스터 대비 약 4배의 계산속도 향상이 필요하다. 이를 위하여 전체 소유역을 4개 단위로 나누어 계산한다 가정하여 4대의 계산노드를 구성하였다. 더불어 계산노드 관리를 위한 헤더노드 1대와 기상자료 저장을 위한 스토리지노드 1대를 추가로 구성하였다.

<표 4> HPC 클러스터 노드 구성

구 분	사 양	수 량
헤더 노드	- CPU : intel Xeon 2.4Ghz 20 Core - MEM : DDR4 128GB 16s/T	1
계산 노드	- CPU : intel Xeon 2.4Ghz 20 Core - MEM : DDR4 128GB 16s/T	4
스토리지 노드	- CPU : intel Xeon 2.4Ghz 20 Core - HDD : SAS 12Gb 84TB	1

이렇게 구성된 홍수예보시스템의 사용자는 클라이언트 시스템에서 제공하는 UI를 통하여 기상자료를 입력하고 명령을 내리게 되며, 운영서버는 공동활용 WEB/DB을 통하여 전달받은 명령을 작업스케줄러에 따라 HPC 클러스터를 통하여 수행하고 이를 모니터링 한다. 각 노드의 상태, 성능, 에러 등의 모니터링 관련 히스토리와 작업 관련 정보들은 공동 DB에 저장되고 기상예측 및 홍수정보 관련 정보들은 HPC 스토리지 노드에 저장된다.



<그림 1> 홍수예보시스템 HPC 설계(안)

6. 병렬화 시 고려사항

HPC 환경에서 병렬화 개발을 위해서 MPI(Message Passing Interface), OpenMP(Open Multi-Processing) 방식을 가장 많이 사용하고 있다. MPI와 OpenMP를 효과적으로 사용하기 위해서는 HPC의 장비 구성을 파악하여야 효과적인 프로세스 및 쓰레드의 구성 및 할당이 가능하다.

<표 5> HPC 설계에 따른 병렬처리 설계방안

구분	개발방안
Open MP	○ 최대 20개의 Thread를 활용한 코드 병렬화 지원 설계 ○ 128GB 메모리 내에서의 해석범위 설계
MPI	○ 최대 80개의 분산코어의 Scale-out 구조 MPI 코드 병렬화 설계 ○ 512GB의 분산메모리 구조를 고려한 Scale-out 구조의 계산모형 설계

7. 결론

본 논문에서는 홍수예보를 위하여 분포형 유역유출모형 모의테스트를 기반으로 홍수예보시스템의 HPC 프레임워크를 설계하여 기상시스템 구축 시 참고자료가 될 수 있도록 하였다. 앞으로 제시된 설계를 바탕으로 실제 HPC 클러스터를 실제적으로 구축하고 모의테스트를 수행함으로써 홍수예보시스템 구성을 계속 진행할 예정이다.

참고문헌

[1] 최윤석, 김경탁, 김주훈 “기상레이더와 분포형 모형을 이용한 실시간 유출해석 시스템 개발 및 평가” 한국습지학회지 제14권 제3호, 2012년 8월

[2] 김길호, 최윤석, 원영진, 김경탁 “분포형 모형과 클라우드 서비스를 이용한 낙동강 실시간 유출해석시스템 개발 및 성능평가“ 한국지리정보학회지 20(3), 2017년 9월

[3] 정성영, 박진혁, 허영택, 정관수 “분포형 강우유출모형 병렬화 처리기법 적용” 한국수자원학회 논문집 Vol. 43 No. 8, 2010년 8월