
전공소개서

박주원

Contents 1

전공학업내용



연구 분야

Big Data

- 빅데이터 플랫폼인 Hadoop Eco-system의 최적화 연구
- 대용량 데이터를 효율적으로 처리하기 위한 알고리즘 및 프레임워크 연구
- 대량의 스몰파일을 빅데이터 플랫폼에서 효율적으로 처리하기 위한 연구

Deep Learning

- 초거대규모 CPU 기반 환경에서 Deep Learning 모델의 병렬 확장성 연구
- GPU 환경에서 분산 Deep Learning 최적화 연구
- MR-IoT융합 기반의 재단 대응 인공지능 응용 기술 연구

Parallel and Distributed Computing

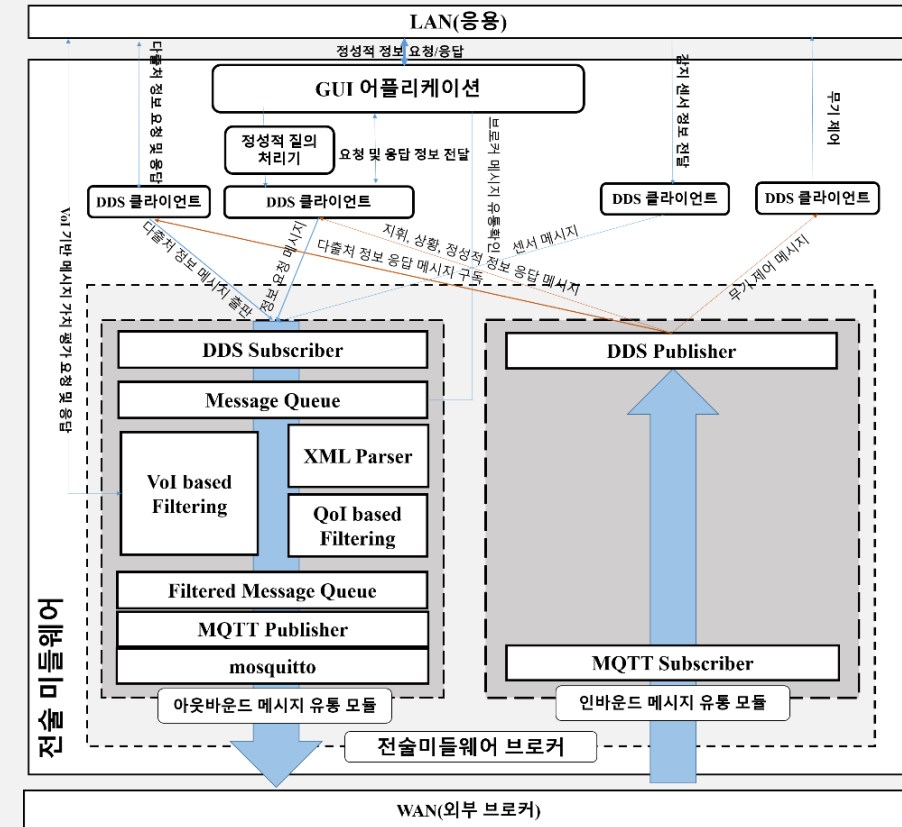
- 미래 전술네트워크 상호운용성을 위한 아키텍처 및 협업 모델 기술 연구
- 전술 네트워킹 미들웨어에서 적응형 정보교환 메시징 서비스 및 가치기반 정보평가 기술 연구
- Pub/Sub 메시징 시스템 최적화 연구

프로젝트 소개

- 계약기관 : 국방과학연구소
- 연구기간 : 2019-01 ~ 2020-05

프로젝트 요약

- 지역 DDS 네트워크 간 효율적인 메시지 유통을 위한 OpenDDS 기반의 하이브리드 전술 미들웨어 구현
 - MQTT+DDS 구현체를 활용한 하이브리드 전술 미들웨어
 - 전술 리파지토리의 질의변환 과정을 지원하는 정성적 정보 요구 표현 정의 및 유통
 - 네트워크 적응형 인코딩 기법을 활용한 XML 기반의 C2 메시지 포맷
 - 효율적인 요소 색인 및 QoI/VoI 가치기반 메시지 필터링
 - IaaS에서 시험된 오픈소스 미들웨어 기반의 프로토타입

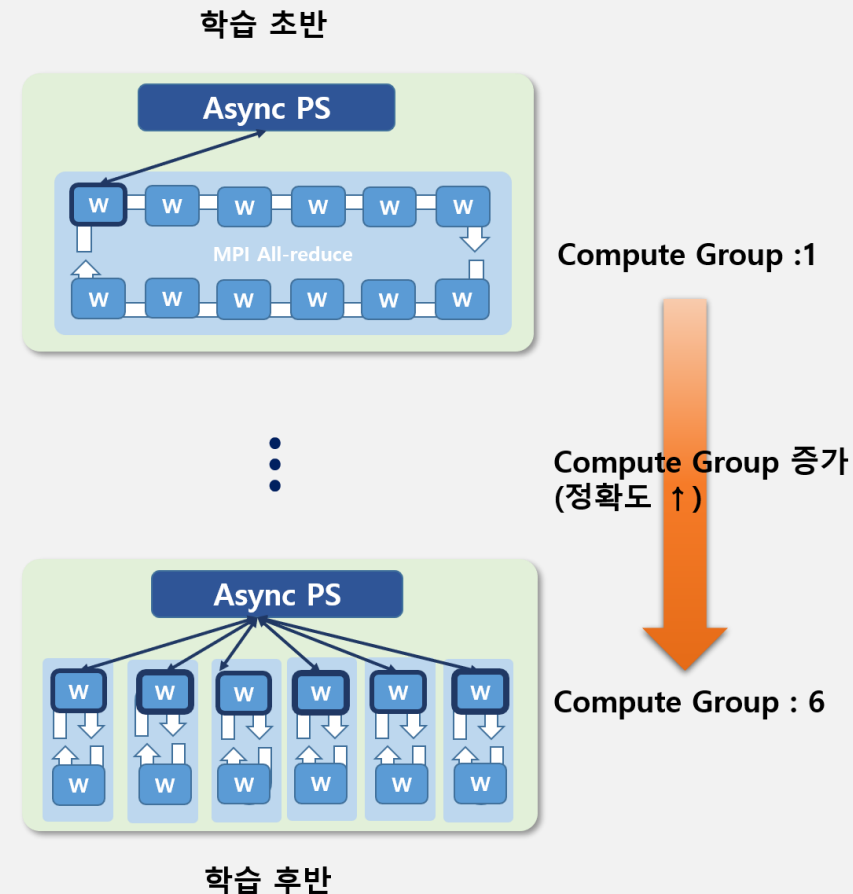


프로젝트 소개

- 계약기관 : 한국과학기술정보연구원 (KISTI)
- 연구기간 : 2019-01 ~ 2019-10

프로젝트 요약

- 초거대규모 CPU 기반 환경에서 딥러닝 응용의 병렬확장성 실험
 - 슈퍼컴퓨터인 누리온에서 딥러닝 응용의 특성자료 수집을 위한 거대규모 데이터 벤치마크 실험 수행
 - 초거대규모 CPU 기반 환경에서 수행 시 제한요소 및 성능저하요소 식별
- 초거대규모 CPU 기반 환경에서 딥러닝 플랫폼 운용을 위한 최적화 연구
 - 딥러닝 학습 속도 및 정확도 향상을 위한 캐쉬 활용 기법, 비동기통신기반 SGD 알고리즘 등 최적화 기법 연구
 - 거대규모 데이터 벤치마크를 통해 메모리와 네트워크에서 이점을 가지는 CPU 기반 환경에서 적합한 특징 선별

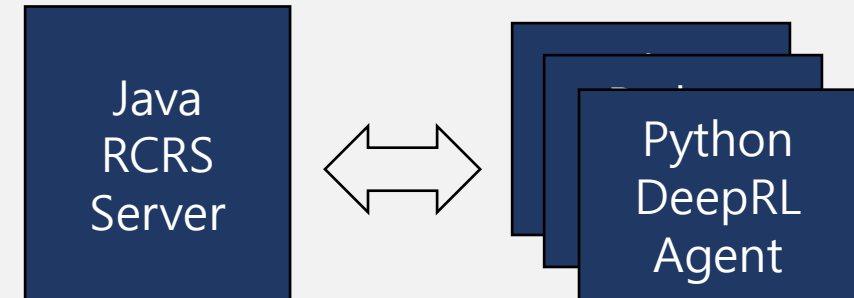
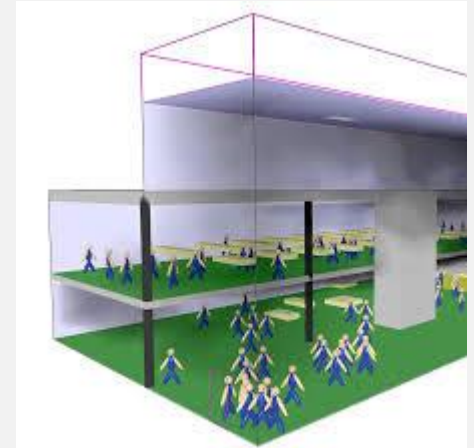
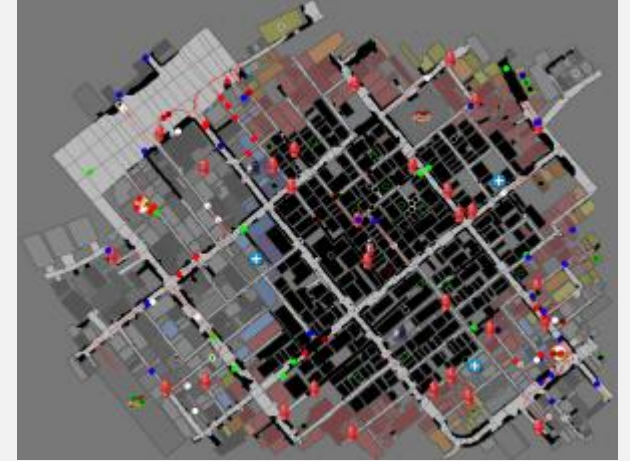


프로젝트 소개

- 계약기관 : 아주대학교 ITRC 센터
- 연구기간 : 2020-01 ~ 2020-12

프로젝트 요약

- 재난대응 콘텐츠 기반의 심화 기술적응으로 MR-IoT의 저전력 지능형 제어시스템 개발 및 이에 따른 재난대비대응 응용플랫폼 개발
 - 화재에 대한 구조 시뮬레이션 중 하나인 RCRS(RoboCup Rescue Simulation) 환경을 기반으로 하는 강화학습 환경 구축
 - 시뮬레이션 서버와 통신을 할 수 있는 파이썬 기반 강화학습 환경 클라이언트 구현

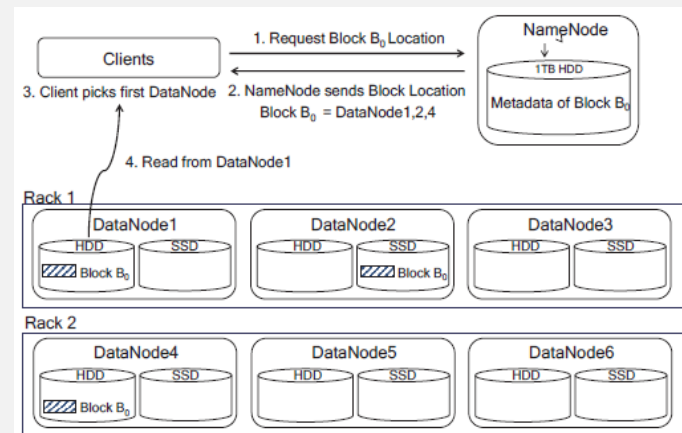


Contents 3 논문 I

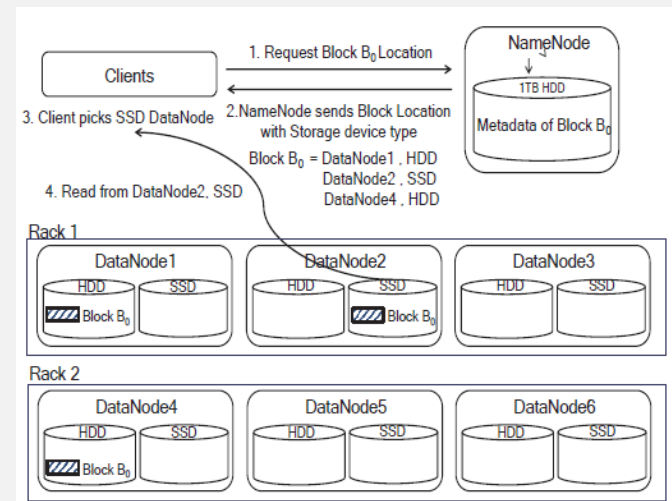
Data Temperature-aware Block Management Method for HDFS Heterogeneous Storage

논문 요약

- HDFS (Hadoop Distributed File System)에서 Heterogeneous Storage 전략을 효율적으로 활용하기 위한 기법 제안
- HDFS는 기본적으로 DataNode의 모든 Storage를 Homogeneous라 여기고 작동
 - 따라서 데이터를 읽을 때 Data Locality만을 고려하여 가장 가까운 노드에 접근
 - 즉, Storage Type을 고려하지 않음
- 본 논문에서는 액세스 빈도에 따라 데이터 온도를 분류하고, 분류한 온도에 맞는 Storage 정책을 적용
 - 액세스 빈도가 높은 데이터의 경우 높은 온도로 분류되고 성능이 좋은 Storage(ex. SSD, RAM Disk)에 저장하는 정책 적용
- 데이터 블록에 접근할 때 Data Locality뿐만 아니라 Storage Type도 고려
 - Data Locality와 Storage Type을 고려하여 우선순위 지정



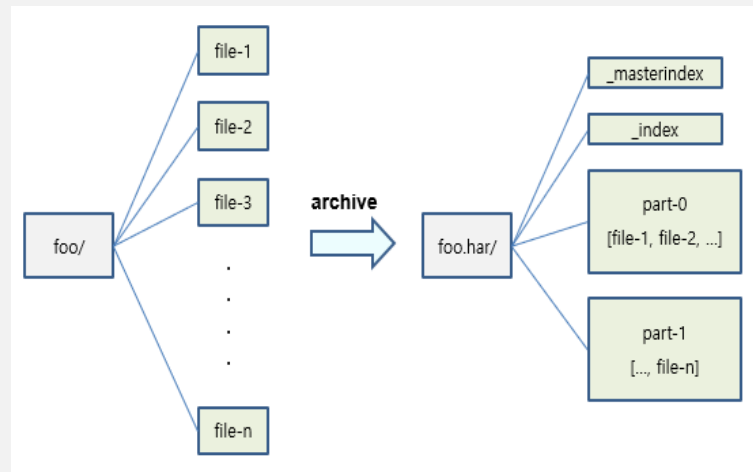
기존 기법



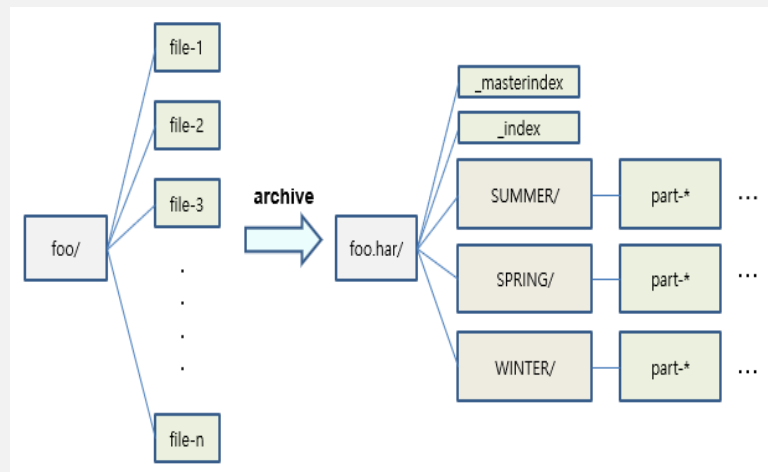
제안 기법

논문 요약

- 스몰파일에 Heterogeneous Storage 전략을 효과적으로 수행시키기 위한 기법 제안
- 기존 아카이빙 기법들은 Heterogeneous 환경을 고려하지 않아 Heterogeneous Storage 전략 수행이 어려움
- 스몰파일의 액세스 패턴에 따라 분류
 - 데이터 저장 시간, 액세스 빈도에 따라 온도를 분류
 - 분류한 온도에 맞는 Storage 정책 설정
- Hadoop에서 제공하는 아카이빙 기법인 HAR의 아카이브 구조를 변형하여 구현
 - 분류한 온도에 맞는 HDFS Storage 정책을 설정하기 위해 아카이브 내에 각 온도의 디렉토리를 구성
- 기존 기법 대비 아카이브 생성 시간면에서 약 28%의 성능 향상



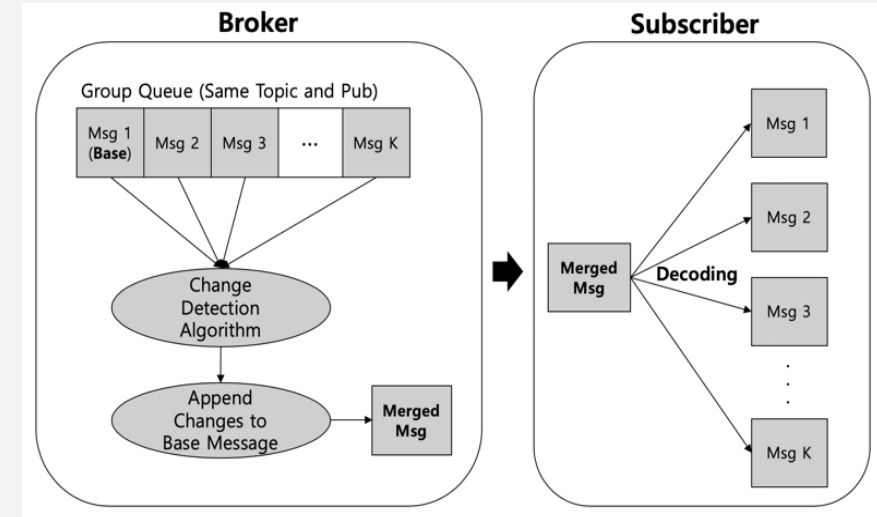
기존 기법



제안 기법

논문 요약

- 전송 네트워크 환경은 대역폭, 지연율, 신뢰성, 안정성, 연결성과 관련된 문제가 발생
- 이러한 환경에 대한 출판/구독 모델 기반 메시징 시스템의 문제는 전달 시간 지연 문제를 가짐
- K-way 메시지 통합을 통해 메시지에 대한 헤더 추가, 제거 및 페이로드 직렬화, 직렬화 해제와 같은 작업이 한번만 이루어지도록 제안
- 메시지 통합을 통해 전체 메시지 전송량이 감소하기 때문에 전달 지연시간(Delivery Latency)을 줄일 수 있음
- K-way 메시지 통합은 네트워크 속도와 통합 과정의 오버헤드를 고려하여 k 값을 설정
- 같은 토픽, 같은 출판자로부터 발급된 메시지들 중 먼저 발급된 순서로 k개의 메시지를 그룹 큐에 삽입
- 그룹 큐에 있는 K개의 메시지를 KF-Diff+ 알고리즘을 통해 차이 탐지 후 하나의 메시지로 통합



Contents 4

특허 출원

대규모 네트워크 환경에서의 효율적인 메시지 전송을 위한 요소기반 XML 표현 기법 및 필터링 시스템

발명 특허 소개

- 발명자 : 오상윤, 배민호, 여상호, 박주원, 민경준, 주현수(아주대), 박규동, 류동국(국과연)
- 한정되고 불안정한 통신망 환경에서의 신속/효율/효과적인 XML 기반 자료 교환

