

# 38

## 대규모 딥러닝 고속처리를 위한 분산 딥러닝 플랫폼 기술

### 기술개요

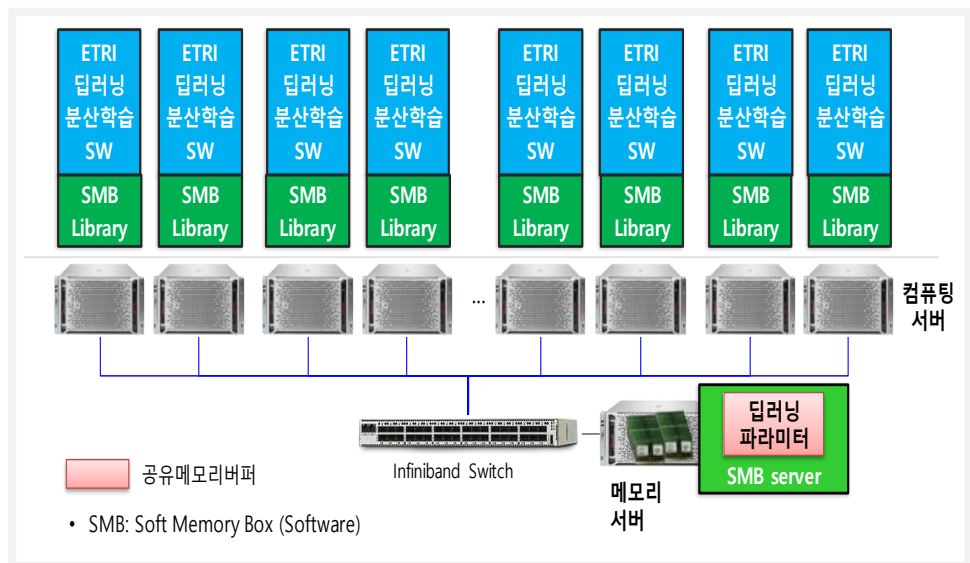
▪ 대규모 딥러닝 고속처리를 위한 분산 딥러닝 플랫폼 기술

- 다수의 고성능 컴퓨팅 서버들을 이용하여 고속으로 대규모 딥러닝 모델의 트레이닝(학습)을 수행하는 분산 딥러닝 플랫폼 기술

### 기술의 특징점

▪ 대규모 딥러닝 학습데이터를 효율적으로 학습시키는 기술

- 높은 정확도를 요구하는 딥러닝 모델 및 학습데이터량이 많은 응용학습 시간을 분산 트레이닝을 통해 획기적으로 개선  
 - Caffe, TensorFlow와 같은 기존 딥러닝 라이브러리를 분산 실행시 발생하는 통신병목을 해소하도록 확장하여 심층 학습속도를 빠르게 향상시키는 분산학습 솔루션



### 적용분야

▪ 클라우드 플랫폼 서비스/온프레미스 AI(딥러닝) 플랫폼 서비스 구축



## 기술완성도 (TRL)

- TRL 6단계 : 파일럿 규모 시작품 제작 및 성능 평가 단계



## 기술이전 내용 및 범위

- 소프트 메모리 박스**
  - 소프트 메모리 박스 클라이언트 기능
  - 소프트 메모리 박스 서버 기능
  - 사용자 레벨 인피니밴드 통신 계층 모듈 기능
- ETRI-Caffe**
  - 분산 프로세스 실행 관리 및 파라미터 공유 메모리 관리 기능
  - 데이터 병렬 트레이닝 및 분산 딥러닝 계산과 파라미터 통신 병렬화 기능
- ETRI-TensorFlow**
  - 딥러닝 트레이닝 태스크 실행 및 소프트 메모리 박스 기반 파라미터 공유 기능
  - 데이터 병렬 트레이닝 및 모델 병렬 트레이닝 기능
  - 비동기식 분산 트레이닝 및 딥러닝 모델 결과 저장 기능
  - 딥러닝 트레이닝 모니터링 기능

## 관련 지재산권 현황

No.	출원번호	특허명	상태
1	2017-0068445	파라미터 서버 및 그것에 의해 수행되는 분산 딥러닝 파라미터 공유 방법	출원
2	2018-0006024	원격 직접 메모리 접근을 통한 분산 처리 장치 및 그 방법	출원
3	2019-0025730	이종 클러스터 기반의 분산 딥러닝 방법 및 이를 위한 장치	출원
4	2017-0060400	파라미터 공유 장치 및 방법	출원

## 기술이전 문의

- 연구성과확산실 (042-860-4946 / hjchoi2@etri.re.kr)