

แนวทางการทำวิจัย KU Digital and AI Platform

ศ.ดร. จันทนา จันทราพรชัย
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์



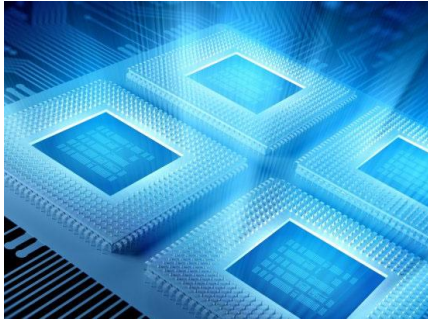


HPC = High Performance Computing

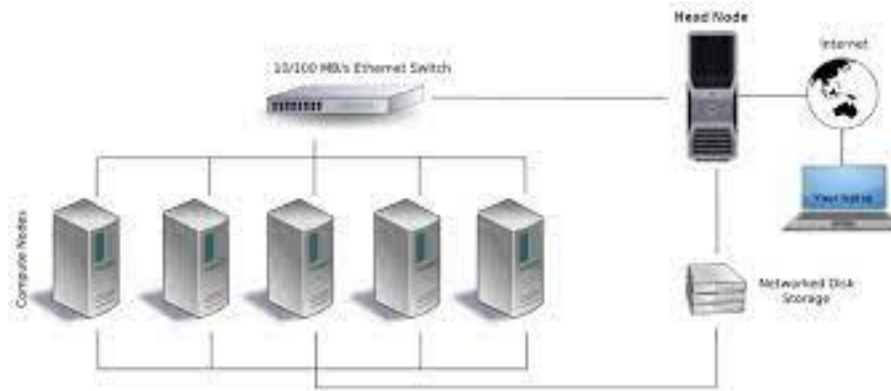
AI = Artificial Intelligence

All about Performance!!! Motivation: Long computing -> algorithm, large data

Throughputs : IPS, Images per sec, Operations per sec, Requests per sec etc.



multicore



Clusters



Multi-GPUs

Application Areas: HPC + API

Deep Learning and AI applications

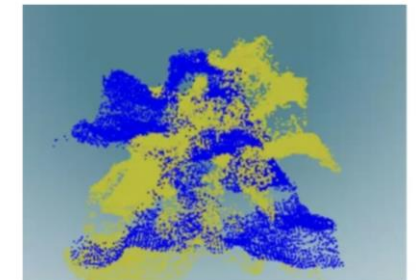
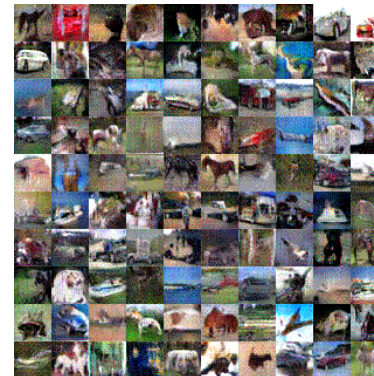
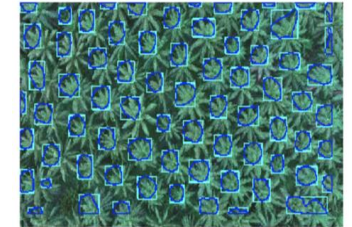
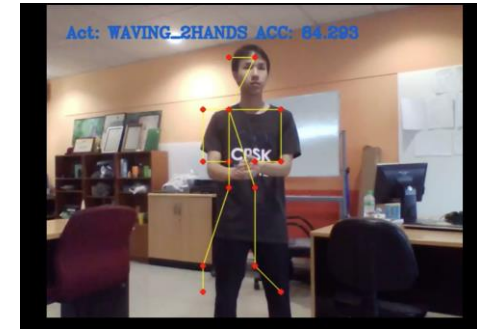
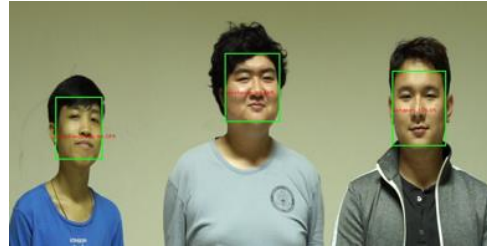
- Model training: new model, new dataset, new algorithm etc.
- Model deployment: new services, new sw architecture, etc.

High Performance Computing

Parallel & Distributed Computing

Cloud Computing: VM

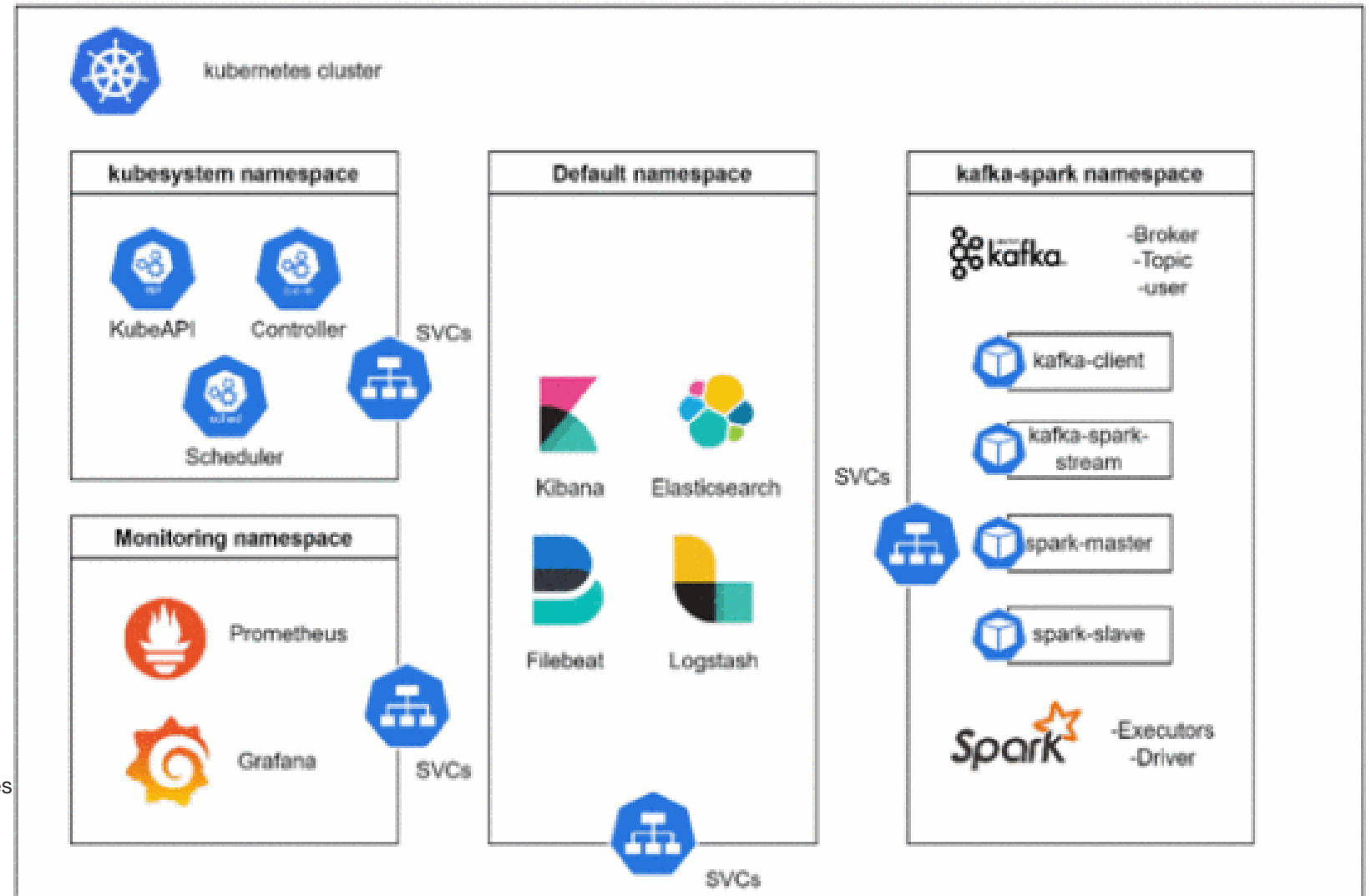
Accelerated GPUs



Example Researches (1)

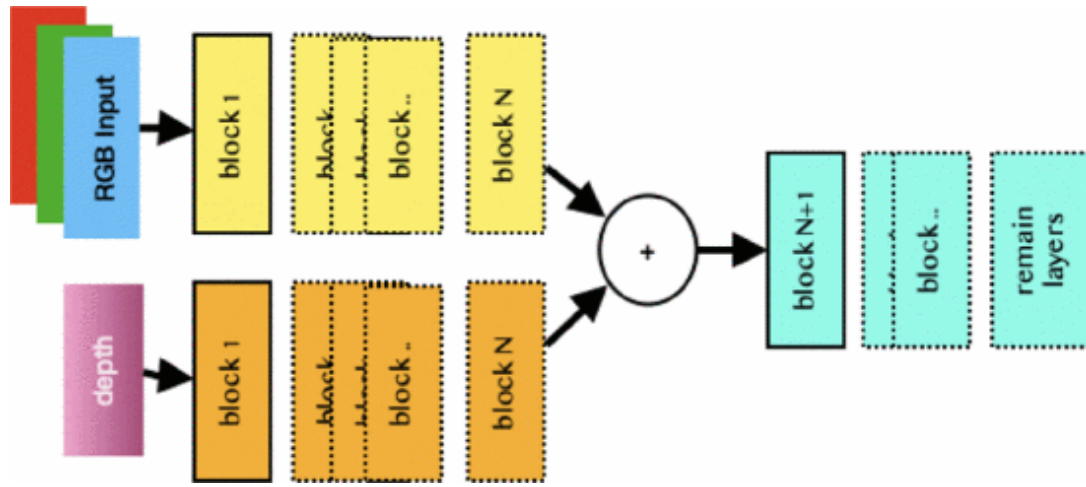
Big data autoscaling Platform

Autoscaling Data
Streaming Microservice on Kubernetes Cluster

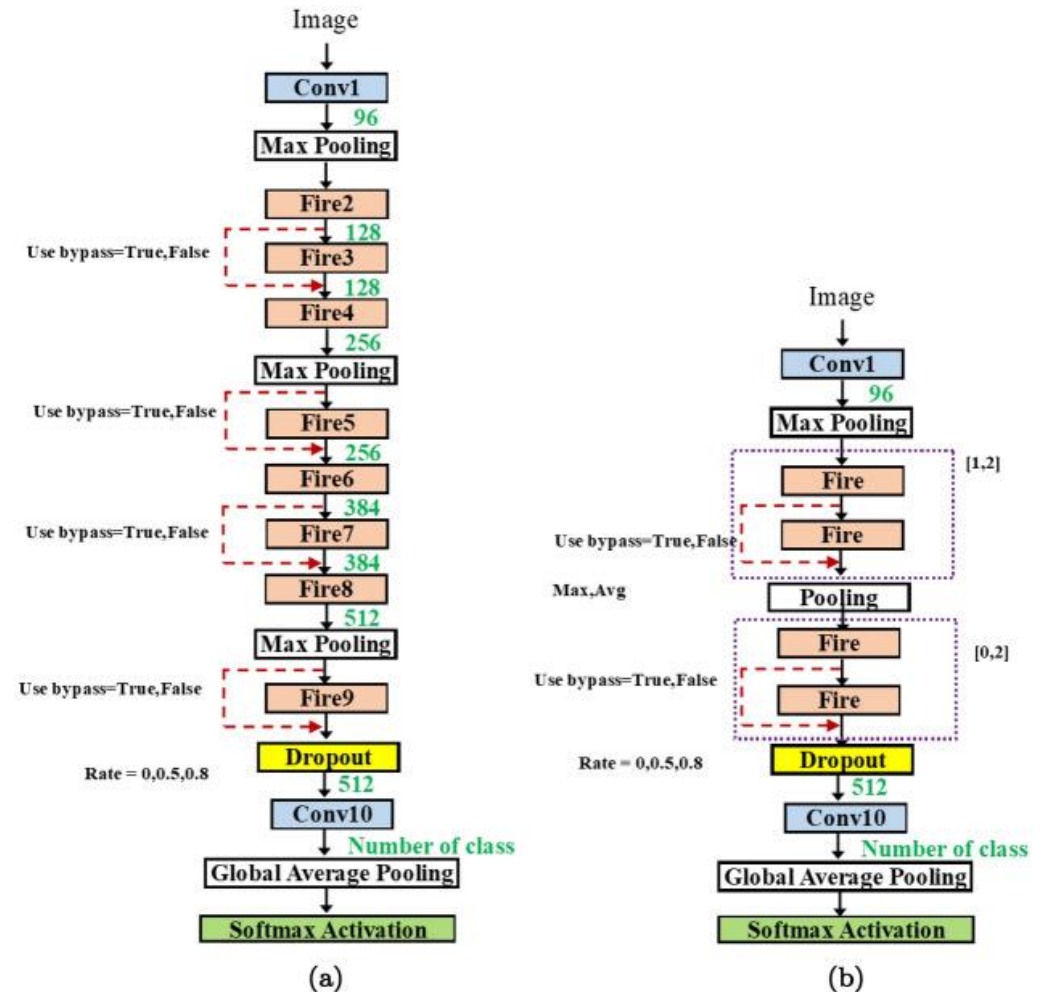


Example Researches (2)

Architecture Exploration



K. Soongswang, P. Romphet and C. Chantrapornchai, "Enhancing MobileNetV2 Performance with Layer Replication and Splitting for 3D Face Recognition Task Using Distributed Training," 2023 *International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC)*, Jeju, Korea, Republic of, 2023, pp. 1-6, doi: 10.1109/ITC-CSCC58803.2023.10212828.



Chantrapornchai C, Kajkamhaeng S, Romphet P. Micro-architecture design exploration template for AutoML case study on SqueezeSEMAuto. *Sci Rep.* 2023 Jun 30;13(1):10642. doi: 10.1038/s41598-023-37682-0. PMID: 37391458; PMCID: PMC10313661.



Q & A