

메모리 중심 컴퓨팅을 위한 운영체제

Memory-Centric Computing and it's Operating System

ETRI 김홍연



CONTENTS

- 01 개념 – 메모리중심 컴퓨팅이란?
- 02 연구 내용 – 거대 가상 머신
- 03 오픈 소스 활동 – 저장소, 커뮤니티, 컨트리뷰션 아카데미





01

개념 – 메모리중심 컴퓨팅이란?

01

개념 – 새로운 응용

Social network



1.39 billion nodes
1 trillion edges
8 TByte data
(출처: VLDB 15)

Giant Learning Model

GPT-3 예시: 자유 대화 - 코로나19

다음 대화는 지혜롭고 사람이 많은 존재와의 대화이다. 이 존재는 복잡계가 어떻게 작동하는지 알고 있다. 이 존재는 대부분의 알려진 지식을 알고 있고, 그 결과 거대한 지혜와 관점을 갖게 되었다. 이 존재는 연구자를 돕기 위해 헌신한다.

연구자: 지금은 2020년 7월입니다. [현재 코로나 상황 설명] 이러한 상황 때문에 많은 비즈니스가 문을 닫았습니다. 어떤 비즈니스가 가장 큰 영향을 받았을까요?

지혜로운 존재: 사람들이 직접 상호 교류해야 작동하는 비즈니스가 가장 큰 타격을 입었습니다. 예로는 레스토랑, 소매점, 대부분의 서비스 기반 비즈니스가 그렇습니다.

[중략]

연구자: 이런 상황에서 돈을 많이 벌려면 어떻게 해야할까요?

지혜로운 존재: 사람 간 접촉이 필요없는 비즈니스를 소유하는 것입니다. 예를 들어, 책이나 옷이나 전자제품을 파는 온라인 스토어가 좋겠죠.

연구자: 어떤 회사가 이런 상황에서 잘될까요?

지혜로운 존재: 몇몇 회사가 있습니다. 테슬라, 구글, 넷플릭스가 잘될 것 같습니다. 이들은 모두 사람 간의 상호 교류가 별로 필요없는 일을 하죠.

[후략]

* 가는 텍스트는 주어진 컨텍스트이며, 볼드는 GPT-3의 생성 결과입니다.

175 billion parameters
700 GByte data
(출처: <https://blog.pingpong.us/gpt3-review>)

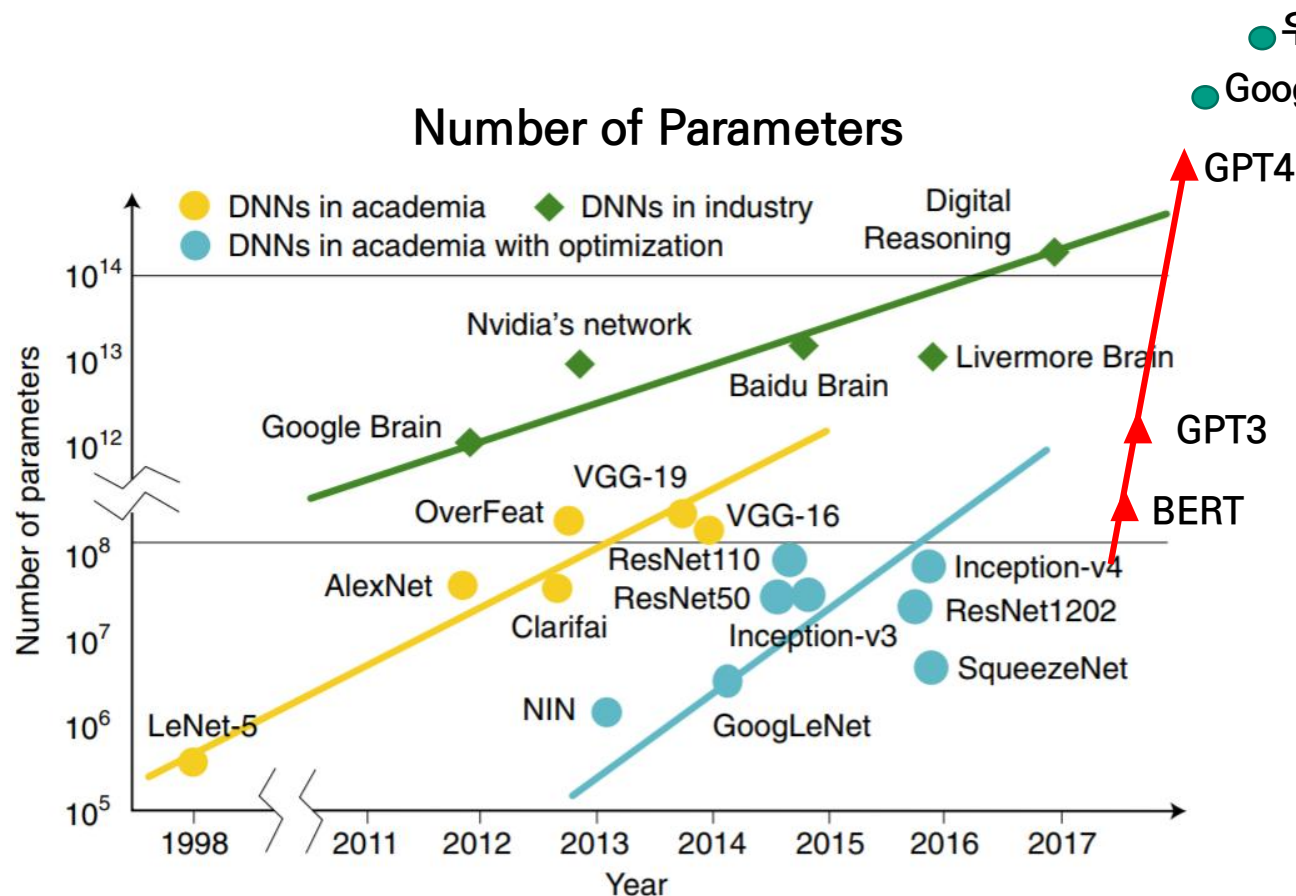
Human Brain Simulation (사람뇌 1%, 1초 시뮬레이션에 40분)



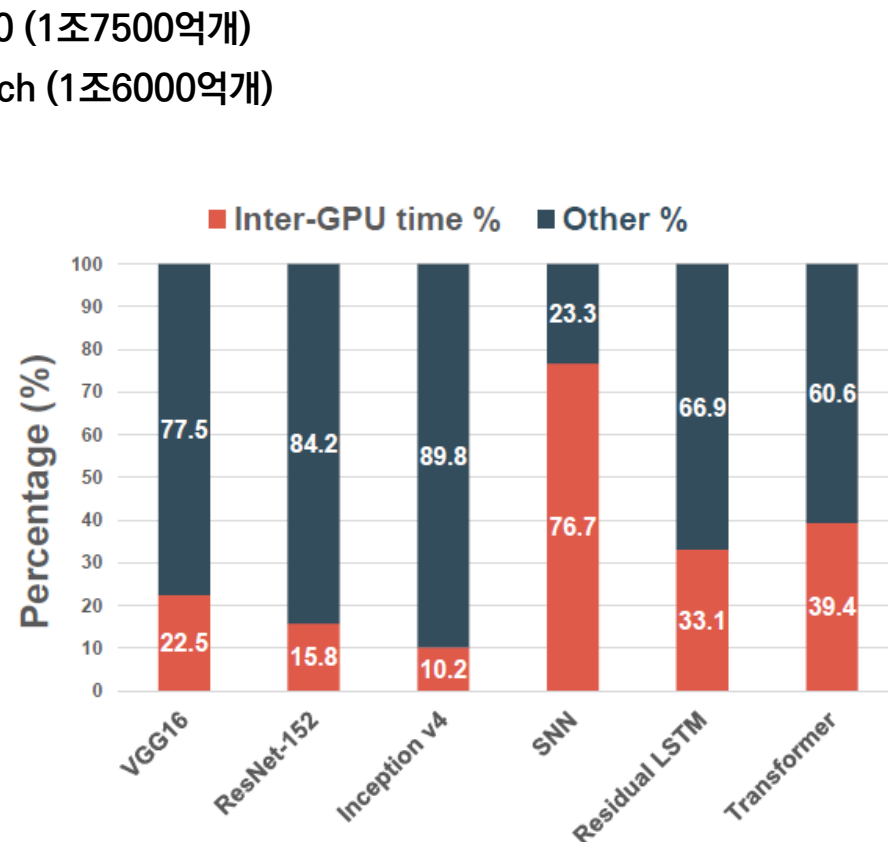
1.73 billion nodes
10.4 trillion edges
80 PByte data
(출처: Riken, 일본)

01

개념 – 새로운 응용, 새로운 병목



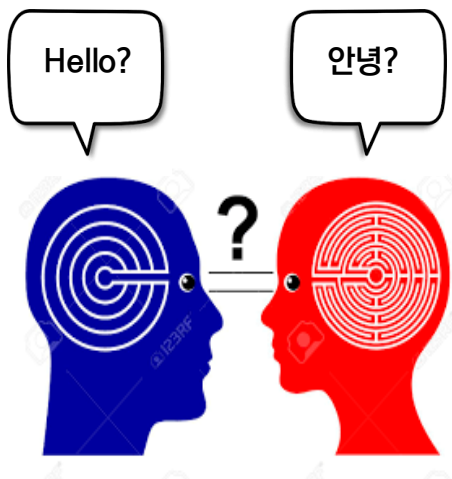
딥러닝에서 메모리가 새로운 병목
(출처: Nature Electronics '18)



4개 GPU로 데이터 병렬 학습 시간에서 GPU간 통신 시간 비율
(출처: EFFICIENT AND ROBUST PARALLEL DNN TRAINING THROUGH MODEL PARALLELISM ON MULTI-GPU PLATFORM, arXiv:1809.02839v4)

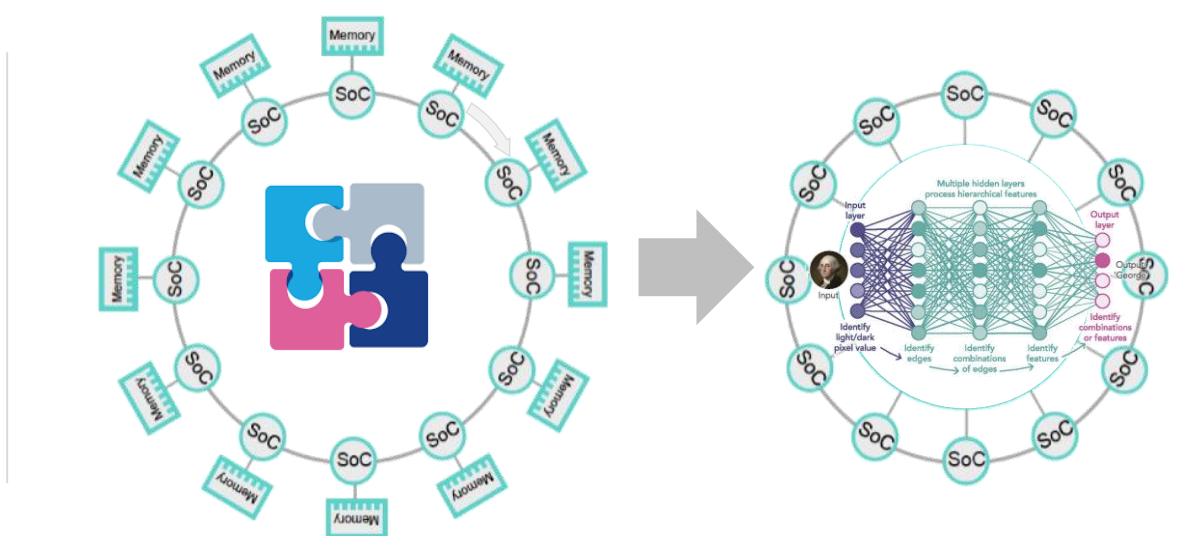
01

개념 – 새로운 응용, 새로운 병목, 새로운 컴퓨팅

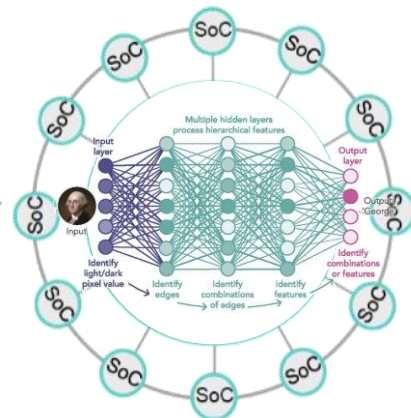


생각/기억의 느린 교류
(대화/통역)

(AS-IS) 프로세서 중심 컴퓨팅
(연산성능이 문제)



데이터의 느린 교환
(분할/병합/전송/변환)



데이터의 순간적 공유
(분할/병합/전송/변환 無)

(TO-BE) 메모리 중심 컴퓨팅
(데이터 접근 성능이 문제)



생각/기억의 순간적 교감
(텔레파시)

새로운 시스템: 렉스케일 컴퓨팅, GPU Sharing
새로운 통신: PCI-ex, Memory Box, RDMA, Gen-Z, CXL
새로운 메모리: Disaggregated Memory, Optane Memory
새로운 OS: Disaggregated Memory, Reverse Virtualization



02

연구 내용 – 거대 가상 머신

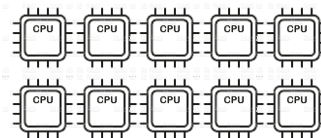
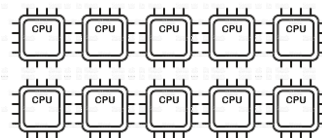
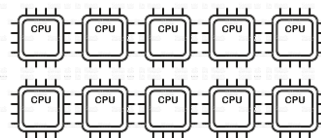
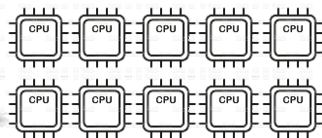
연구내용 – 새로운 응용, 새로운 병목, 새로운 컴퓨팅

고성능 응용
(단일 프로그램/다중 쓰레드)

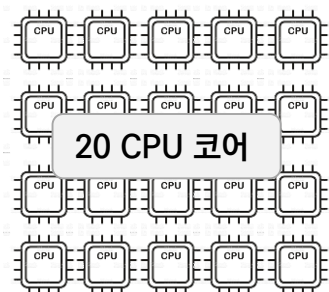


처리 성능 향상
편의성 증대

거대 가상 머신 (Giant Linux OS)



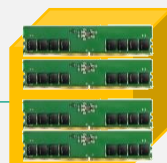
40 CPU 코어



20 CPU 코어

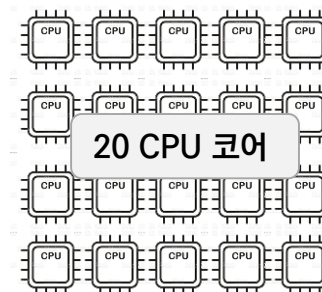


심층 연결망



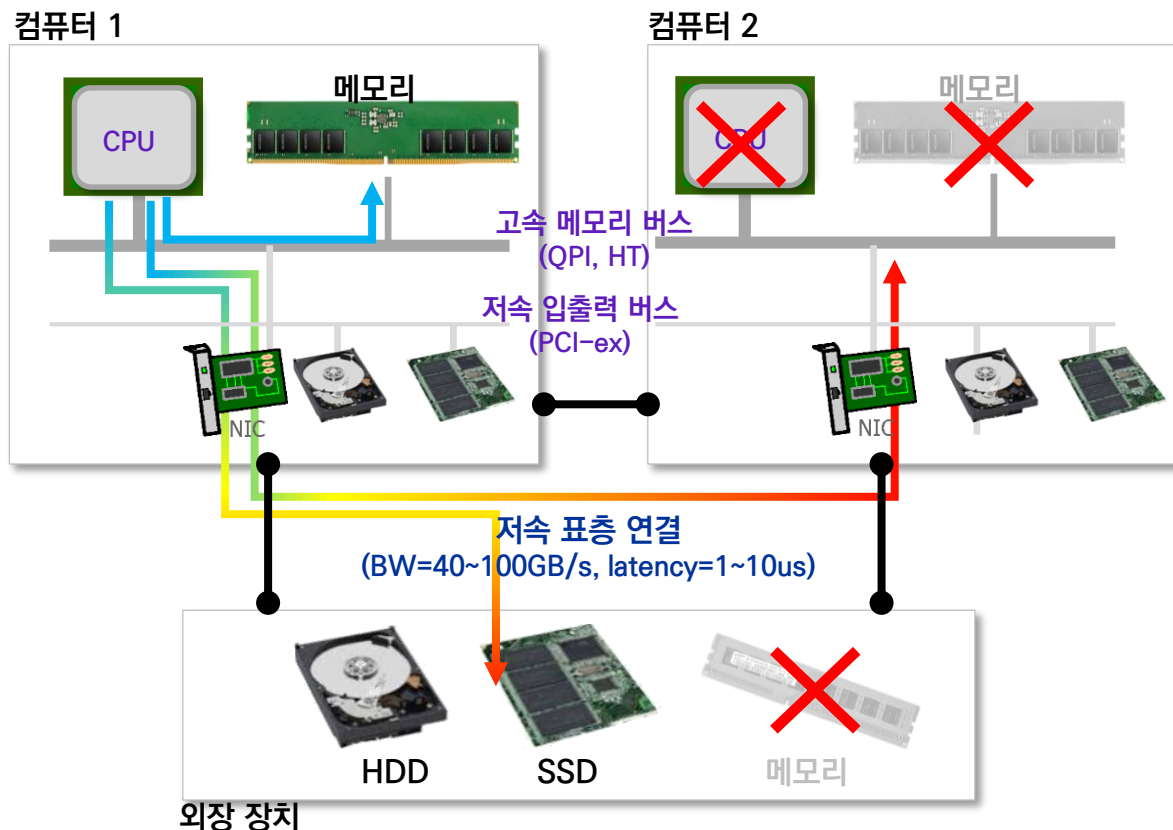
메모리 서버 시제품

심층 연결망

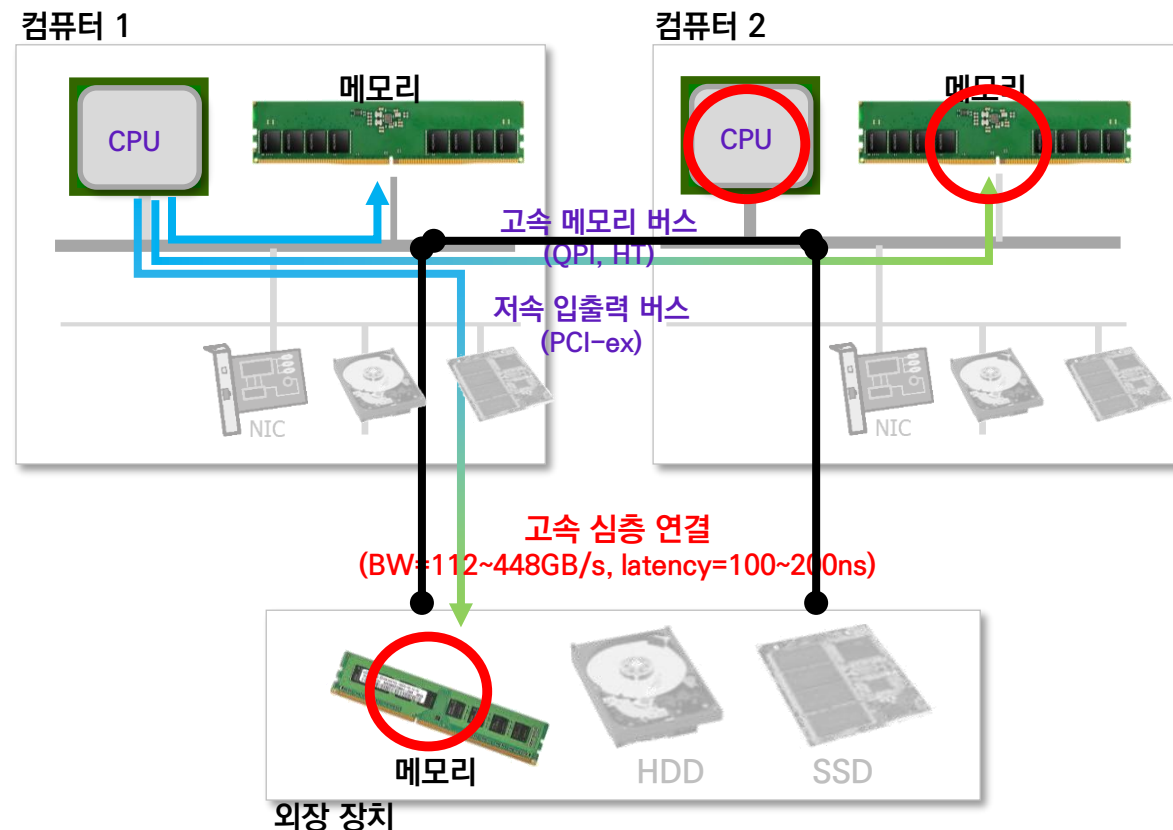


20 CPU 코어

연구내용 – 새로운 컴퓨팅의 근간, 심층 연결망



(AS-IS) 표층 연결망 기반 프로세서 중심 컴퓨팅



(TO-BE) 심층 연결망 기반 메모리 중심 컴퓨팅

연구내용 – 거대 가상 머신의 실행 모습

Guest (VM)

Host #0

Host #1

```
top - 08:55:27 up 22:07, 2 users, load average: 9.90, 9.03, 6
Tasks: 165 total, 2 running, 167 sleeping, 0 stopped, 0 z
%Cpu0 : 1.3 us, 63.6 sy, 0.0 ni, 25.8 id, 0.0 wa, 0.0 hi,
%Cpu1 : 1.3 us, 65.1 sy, 0.0 ni, 24.6 id, 0.0 wa, 0.0 hi,
%Cpu2 : 1.5 us, 63.1 sy, 0.0 ni, 25.0 id, 0.0 wa, 0.0 hi,
%Cpu3 : 1.5 us, 66.0 sy, 0.0 ni, 24.7 id, 0.0 wa, 0.0 hi,
KiB Mem : 65952012 total, 28852500 free, 36201176 used, 89833
KiB Swap: 8388604 total, 8388604 free, 0 used, 2910350
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
5440	root	20	0	64.882g	0.031t	10772	S	128.6	51.0	7:38.18	qemu-system-x86
5463	root	20	0	0	0	0	D	19.3	0.0	1:05.59	dsm-conn/0:1
5464	root	20	0	0	0	0	D	19.3	0.0	1:05.26	dsm-conn/0:2
5468	root	20	0	0	0	0	D	19.3	0.0	1:05.32	dsm-conn/0:6
5462	root	20	0	0	0	0	D	18.9	0.0	1:05.43	dsm-conn/0:0
5465	root	20	0	0	0	0	D	18.9	0.0	1:05.41	dsm-conn/0:3
5466	root	20	0	0	0	0	D	18.9	0.0	1:05.36	dsm-conn/0:4
5467	root	20	0	0	0	0	D	18.9	0.0	1:05.28	dsm-conn/0:5
5469	root	20	0	0	0	0	R	18.6	0.0	1:05.44	dsm-conn/0:7
2604	baiksong	20	0	108112	6152	5076	S	2.0	0.0	0:08.04	sshd
3	root	20	0	0	0	0	S	0.3	0.0	0:00.41	ksoftirqd/0
1	root	20	0	78112	9128	6576	S	0.0	0.0	0:07.82	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.01	kthreadd
5	root	0	-20	0	0	0	S	0.0	0.0	0:00.00	kworker/0:0H
7	root	20	0	0	0	0	S	0.0	0.0	0:06.03	rcu_sched
8	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_bh

```
top - 08:55:46 up 22:04, 2 users, load average: 9.72, 8.95, 6.24
Tasks: 223 total, 2 running, 221 sleeping, 0 stopped, 0 zombie
%Cpu0 : 6.3 us, 49.2 sy, 0.0 ni, 44.4 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu1 : 6.6 us, 49.8 sy, 0.0 ni, 43.6 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu2 : 9.3 us, 49.4 sy, 0.0 ni, 41.3 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu3 : 7.0 us, 51.0 sy, 0.0 ni, 41.2 id, 0.0 wa, 0.0 hi, 0.8 si, 0.0 st
%Cpu4 : 0.0 us, 28.1 sy, 0.0 ni, 71.9 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu5 : 0.0 us, 29.8 sy, 0.0 ni, 70.2 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu6 : 0.0 us, 29.0 sy, 0.0 ni, 71.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu7 : 0.4 us, 28.5 sy, 0.0 ni, 71.1 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
KiB Mem : 65947312 total, 29168608 free, 36362636 used, 416068 buff/cache
KiB Swap: 8388604 total, 8370196 free, 18408 used, 29030836 avail Mem
```

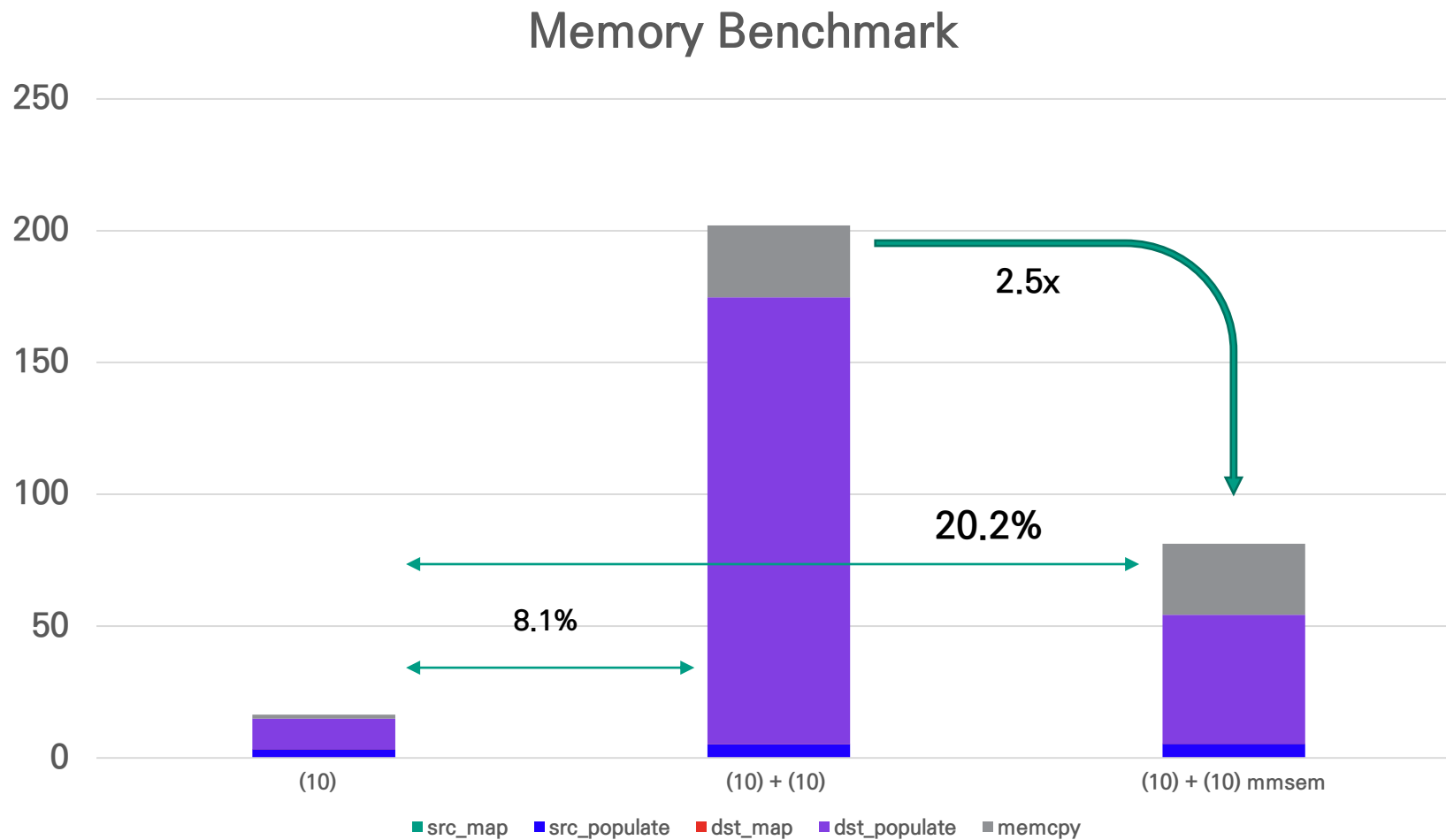
PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
4391	root	20	0	64.655g	0.031t	10736	S	130.6	51.0	7:47.46	qemu-system-x86
4413	root	20	0	0	0	0	D	27.9	0.0	1:45.02	dsm-conn/1:7
4406	root	20	0	0	0	0	R	27.6	0.0	1:44.96	dsm-conn/1:0
4412	root	20	0	0	0	0	D	27.2	0.0	1:44.14	dsm-conn/1:6
4407	root	20	0	0	0	0	D	26.9	0.0	0:00.00	
4411	root	20	0	0	0	0	D	26.9	0.0	0:00.00	
4410	root	20	0	0	0	0	D	26.9	0.0	0:00.00	
4409	root	20	0	0	0	0	D	26.9	0.0	0:00.00	
4408	root	20	0	0	0	0	D	25.9	0.0	0:00.00	
1	root	20	0	78028	8668	6584	S	0.2	0.0	0:00.00	
2	root	20	0	0	0	0	S	0.2	0.0	0:00.00	
3	root	20	0	0	0	0	S	0.2	0.0	0:00.00	
5	root	0	-20	0	0	0	S	0.2	0.0	0:00.00	

```
top - 17:54:31 up 16 min, 0 users, load average: 2.31, 1.30, 0.59
Tasks: 81 total, 1 running, 39 sleeping, 0 stopped, 0 zombie
%Cpu0 : 32.1 us, 38.8 sy, 0.0 ni, 11.6 id, 0.0 wa, 0.0 hi, 0.0 si, 17.5 st
%Cpu1 : 33.1 us, 37.4 sy, 0.0 ni, 11.8 id, 0.0 wa, 0.0 hi, 0.0 si, 17.7 st
%Cpu2 : 34.7 us, 30.2 sy, 0.0 ni, 30.6 id, 0.0 wa, 0.0 hi, 0.0 si, 4.5 st
%Cpu3 : 32.8 us, 25.8 sy, 0.0 ni, 36.2 id, 0.0 wa, 0.0 hi, 0.0 si, 5.2 st
KiB Mem : 32940460 total, 32751324 free, 85852 used, 103284 buff/cache
KiB Swap: 0 total, 0 free, 0 used, 32536800 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
470	root	20	0	39508	868	788	S	265.7	0.0	10:39.19	mem_thread
475	root	20	0	43944	3660	3124	R	4.2	0.0	0:03.80	top
1	root	20	0	77296	8312	6612	S	0.0	0.0	0:05.46	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.15	kthreadd
3	root	20	0	0	0	0	I	0.0	0.0	0:00.01	kworker/0:0
4	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	kworker/0:+
6	root	0	-20	0	0	0	I	0.0	0.0	0:00.00	mm_percpu_+
7	root	20	0	0	0	0	S	0.0	0.0	0:00.21	ksoftirqd/0
8	root	20	0	0	0	0	I	0.0	0.0	0:01.44	rcu_sched
9	root	20	0	0	0	0	I	0.0	0.0	0:00.00	rcu_bh
10	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	migration/0
11	root	rt	0	0	0	0	S	0.0	0.0	0:00.02	watchdog/0
12	root	20	0	0	0	0	S	0.0	0.0	0:00.05	cpuhp/0
13	root	20	0	0	0	0	S	0.0	0.0	0:00.00	cpuhp/1

02

연구내용 – 거대 가상 머신 개발 현황



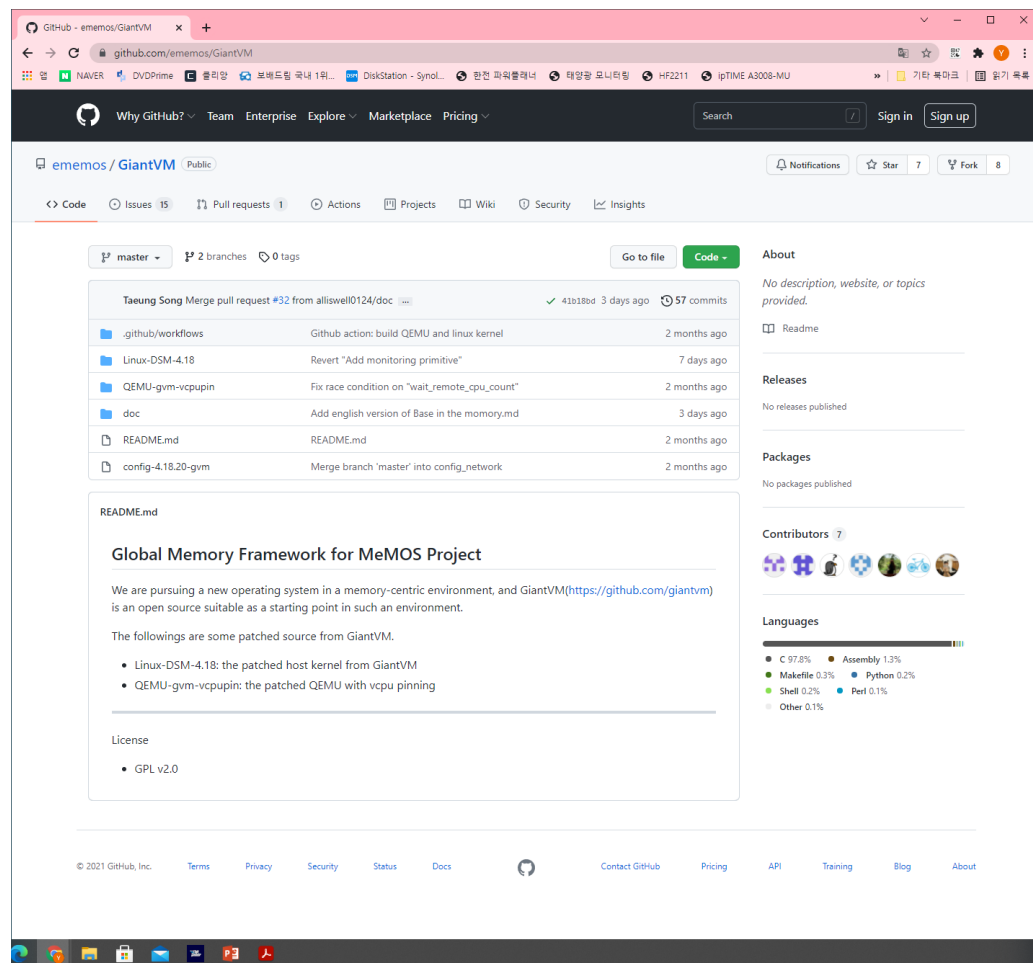
Lower is better



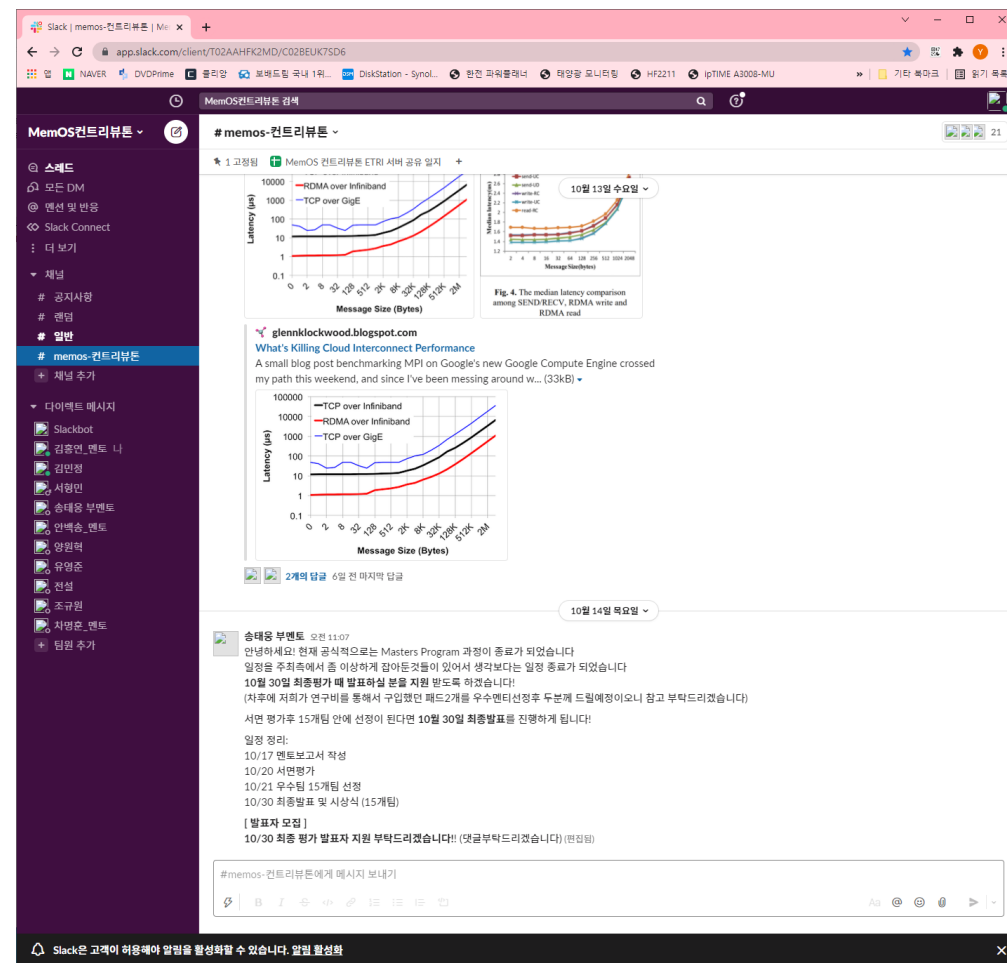
03

오픈소스 활동

오픈소스 활동 – 저장소 및 커뮤니티



Github
(<http://github.com/ememos>)




커뮤니티 슬랙 채널

03

오픈소스 활동 – 컨트리뷰션 아카데미




주요 프로그램

7월 

1 멘티 모집·선발

멘티 모집 온·오프라인 홍보


- > 프로젝트별 참가 멘티 모집
- > 최종 참가 멘티 선발 및 팀빌딩

8월 

2 발대식

팀별 컨트리뷰션 킥오프 미팅


- > 팀별 멘토·멘티가 모여 발대식
- > 컨트리뷰션 아카데미 소개
- > 선배 멘토·멘티 노하우 전수

9~11월 

3 컨트리뷰션

4개월 간의 오픈소스 기여 활동

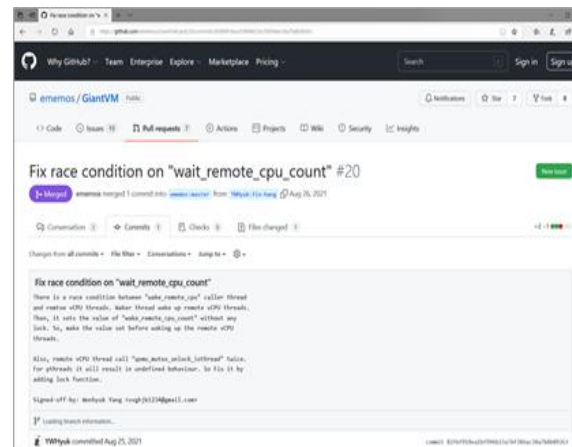
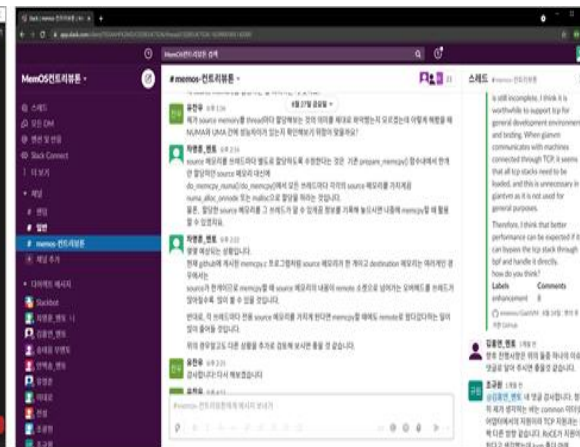
- > 선배 개발자의 멘토링을 통해 오픈소스 프로젝트 기여 활동 수행
- challenges** 환경세팅, 1차 컨트리뷰션 미션 도전
- midterm recap** 전체 팀별 컨트리뷰션 현황 공유
- masters** 리드멘티 선발, 집중 컨트리뷰션 수행

11월 

4 평가·시상

집중 컨트리뷰션 현황 보고·공유회

- > 1차 보고서 서면평가
- > 2차 발표팀 발표평가
- > 시상식 (12월)



2021년도 진행 일정

2021년도 진행 현황

2022년도 신청은 <http://oss.kr>

감사합니다

메모리 중심 컴퓨팅을 위한 운영체제

