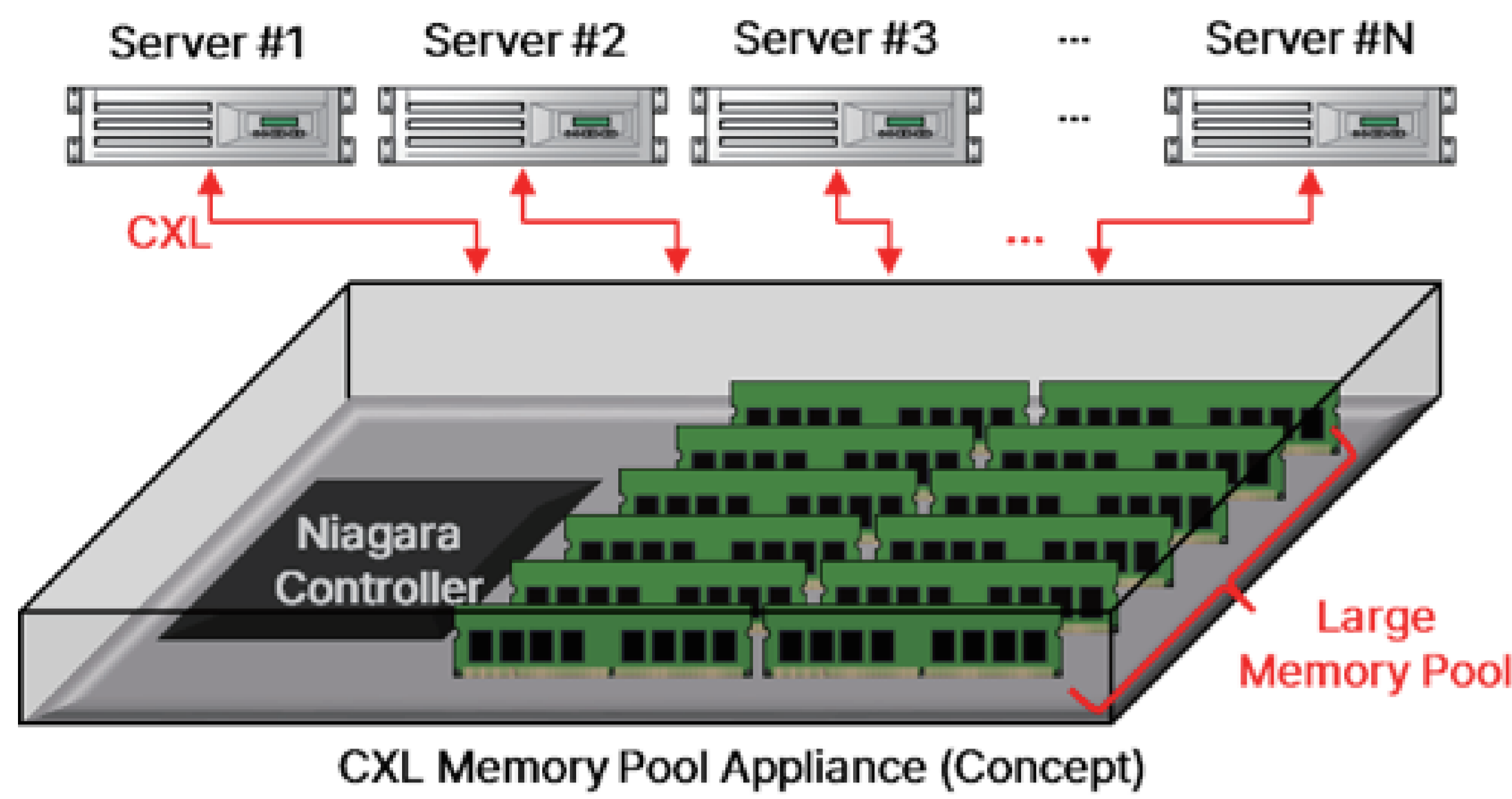


Niagara : Multi-Headed CXL Memory Pool Appliance



Specification (FPGA Prototype 기준)

CXL Interface	CXL 2.0, Gen4x8 최대 8-port
Memory	4CH DDR4 DIMM 최대 1TB
Performance	Latency : 600ns BW : Port당 11GB/s

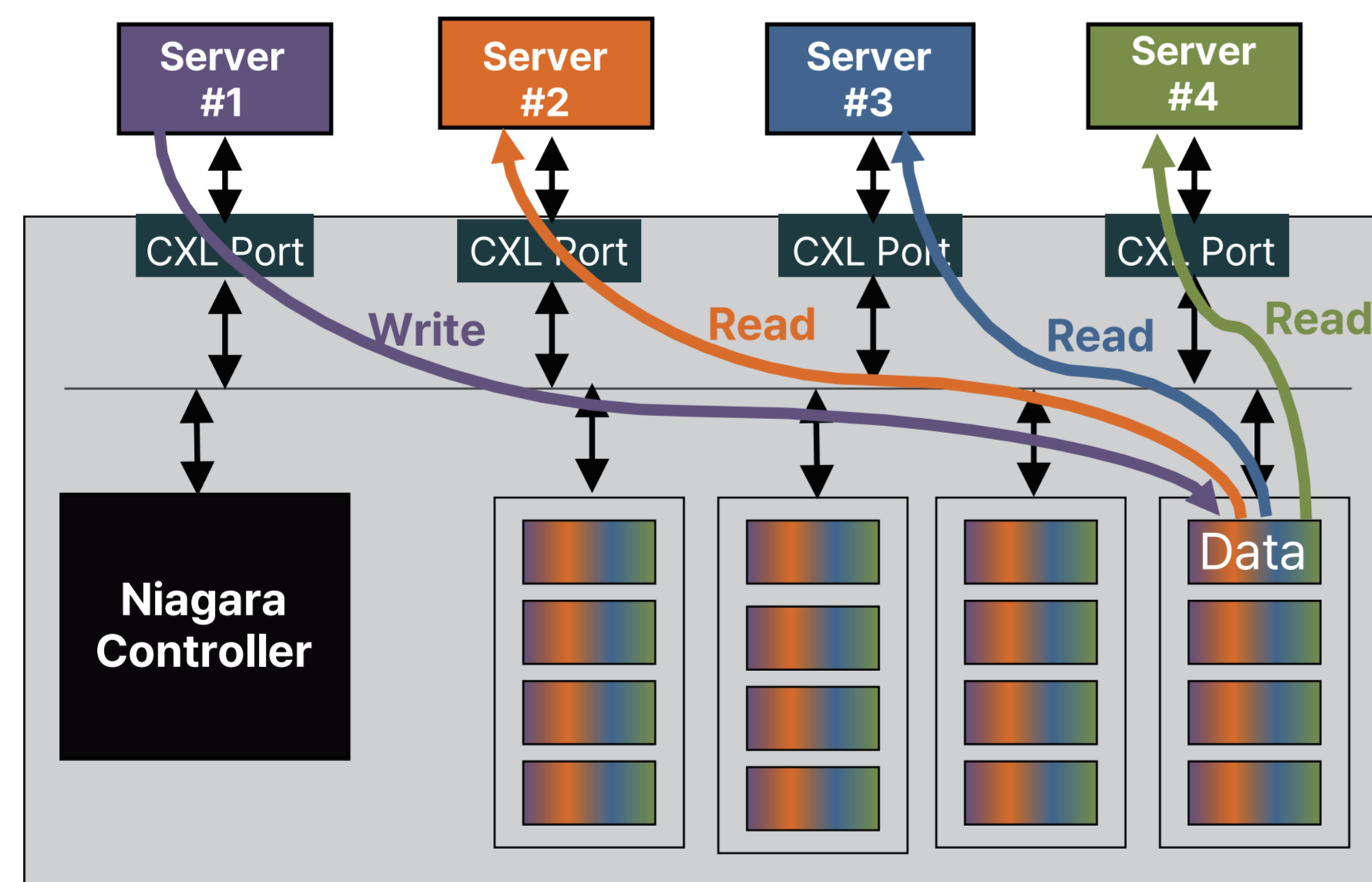
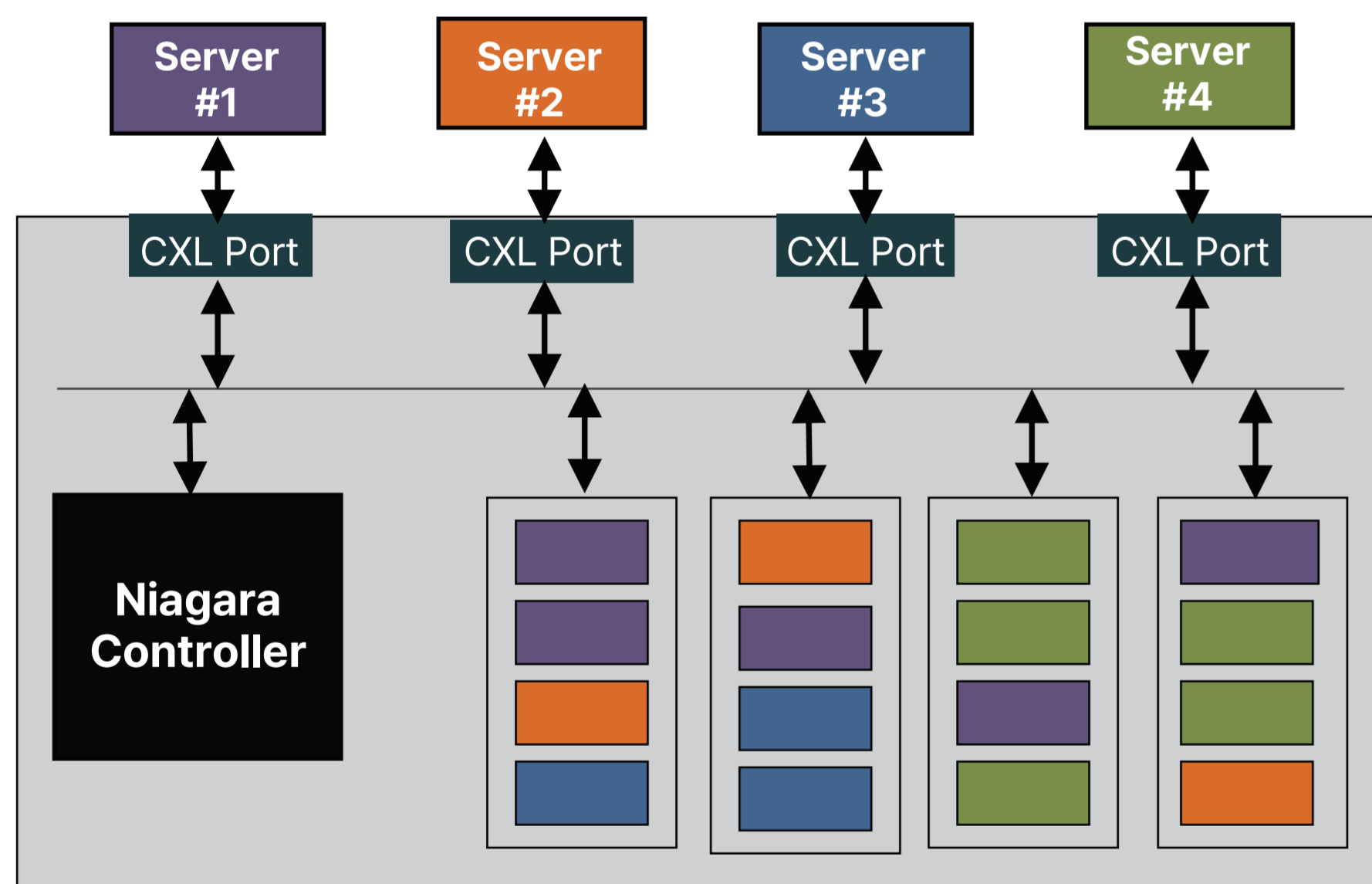
Use Cases

Resource Sharing

- 각 서버의 Memory 요구량 변화에 따라 동적으로 Memory Pool 용량 할당/해제
- 이점 : Memory Utilization 개선으로 전체 Memory 용량 절감
- 활용 : Memory 사용량 예측이 어려운 VM/Container 기반 Cloud Datacenter

Data Sharing

- 각 서버들 간에 Memory Pool 내 동일 주소 기반으로 Data 공유
- 이점 : Data 이동 오버헤드 및 중복 Data 제거로 성능 개선, Memory 용량 절감
- 활용 : Data 공유가 필요한 AI/ML, Big Data Analytics 분산 컴퓨팅 환경



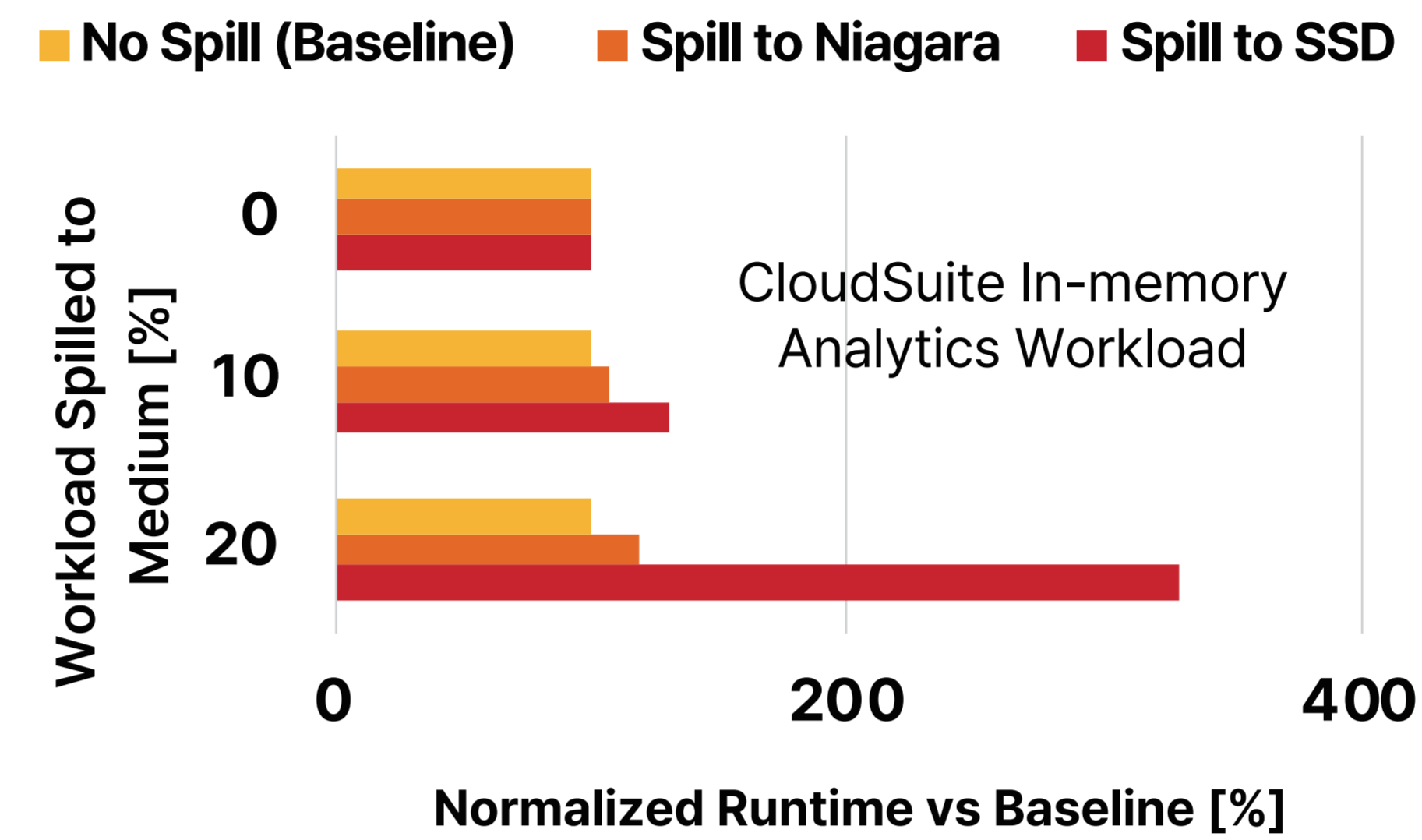
Hardware-Assisted Features

- Block Data Management** : Memory Pool 내 Internal Data Copy/Move 기능
- Snapshot Offloading** : 현재 System 상태의 Persistent Storage 저장 기능 Offloading
- Memory Failure Prediction** : Memory Failure 발생 예측, System 신뢰성 개선

Experimental Results

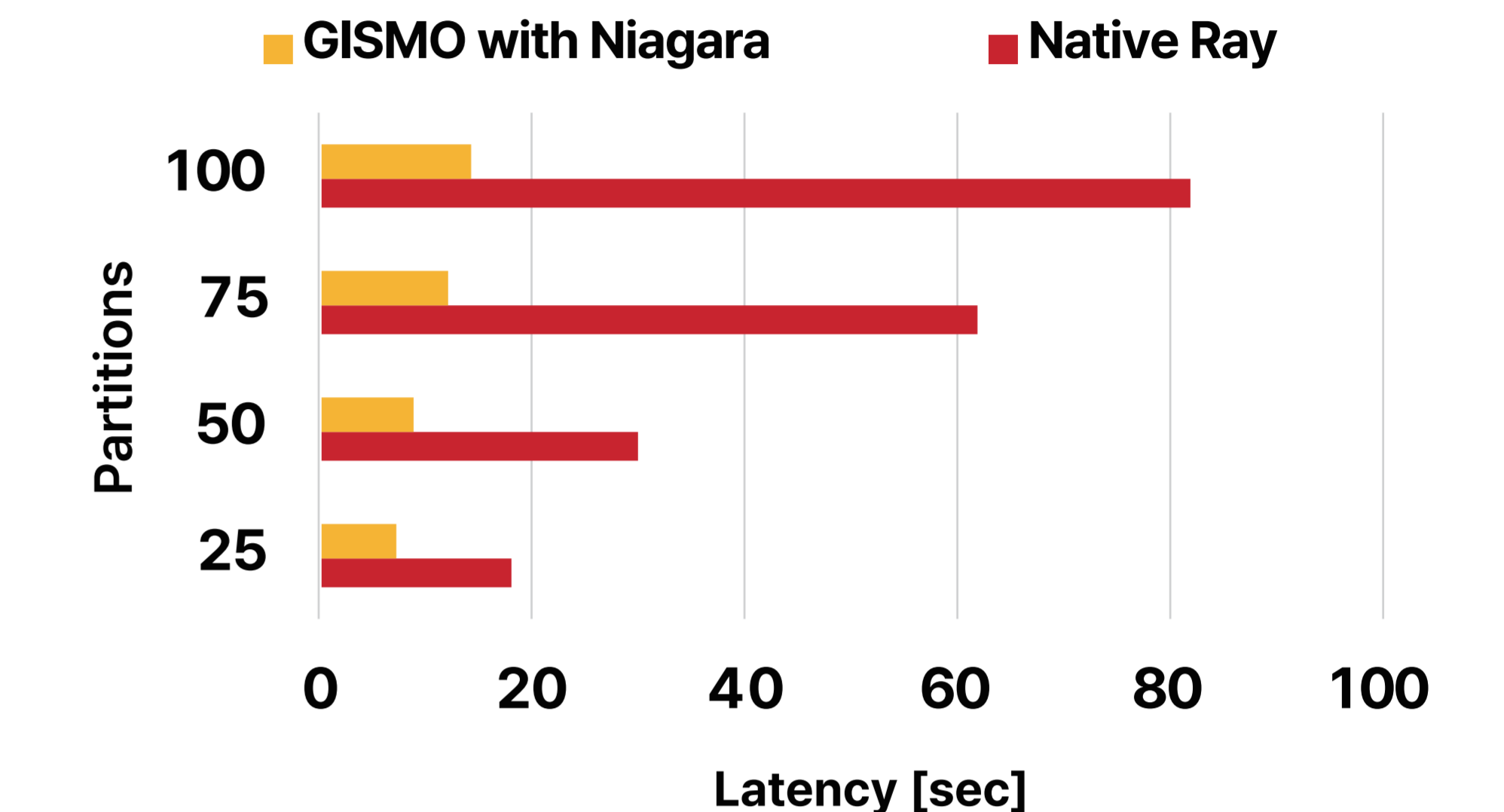
Resource Sharing

- SSD로의 Spill-Over 대신 동적 Memory Pool 용량 추가로 최대 2.5배 성능 개선



Data Sharing

- Memory Pool 상의 Data 공유로 (Gismo*), Native Ray 대비 최대 5.9배 성능 개선



* Gismo: Memory Sharing 지원 분산 컴퓨팅 환경 (MemVerge 공동 연구 개발)

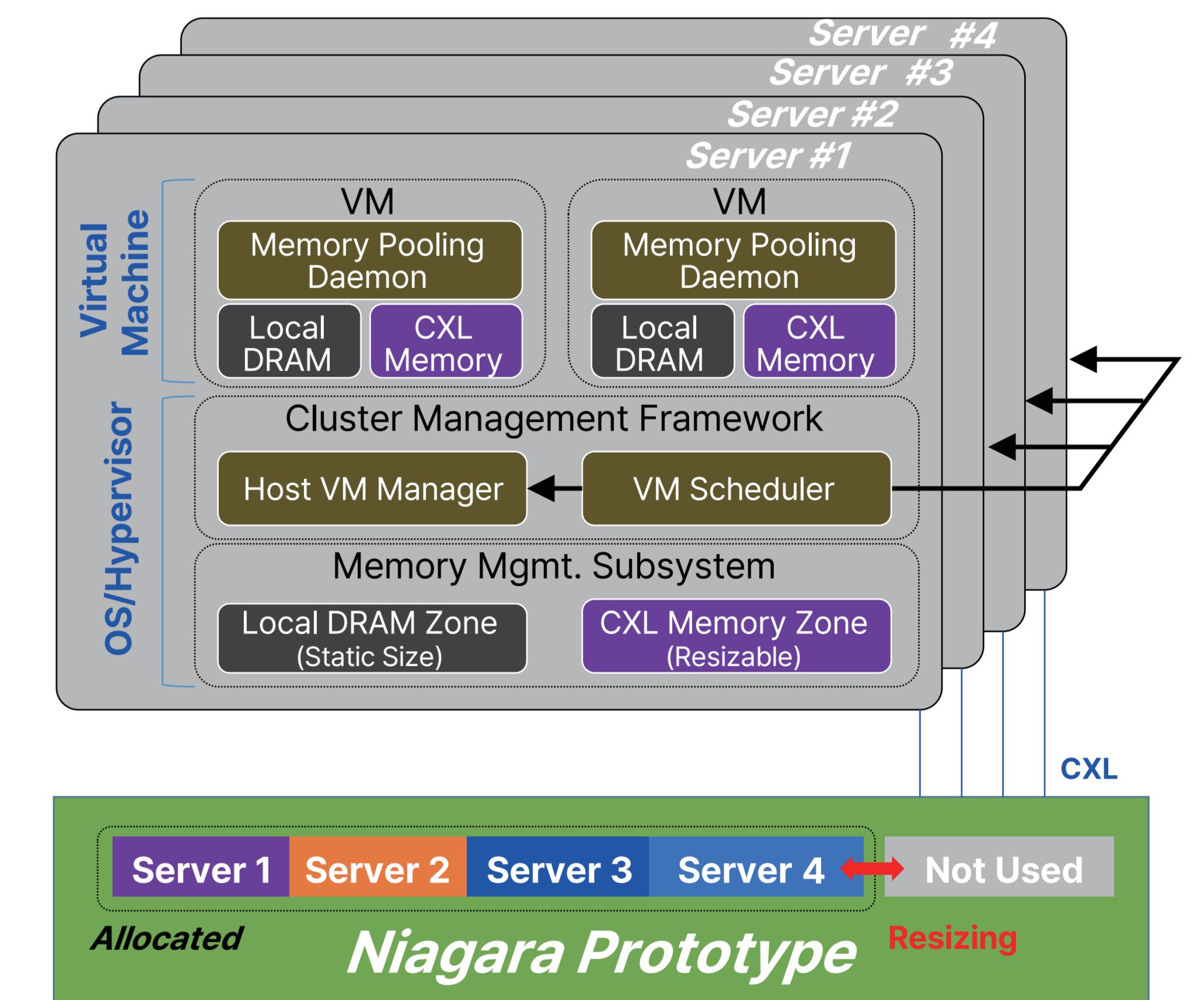
[Live Demo] Multi-Server 환경에서의 Resource Sharing 시연

Hardware 환경

- Memory Pool: SK hynix Niagara 1대
- 서버 : Intel Sapphire Rapids 서버 4대

Software 환경

- Virtual Machine (VM) : Synthetic Workload 구동 및 Memory 사용량 기반으로 용량 조절 기능을 담당하는 Daemon으로 구성
- Cluster Management Framework : VM 실행 서버를 관리하는 VM Scheduler 및 Memory Pool 자원의 요청/반환 기능을 담당하는 VM Manager로 구성



동적 Memory 할당/해제 지원 VM Software Stack

향후 연구 계획

- CXL Pooled Memory Usage 추가 발굴
- CXL Pooled Memory System Research Platform 고도화
- 고객, Partner사와의 협업을 통한 HW/SW Ecosystem 구축, 차별화 기능 발굴 및 개발