Scaling of Memory Performance and Capacity with CXL Memory Expander

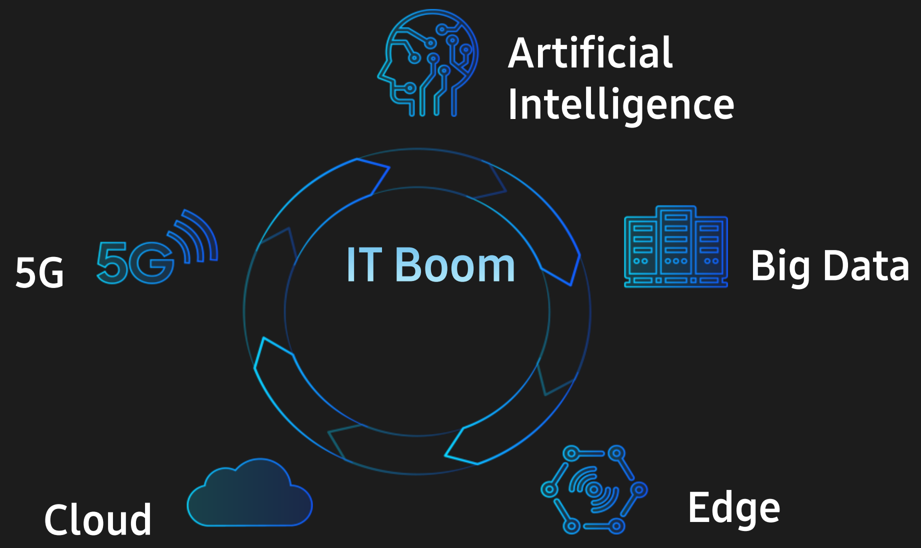
August, 2022 | Samsung Electronics Co., Ltd.

S. J. Park, H. Kim, K.-S. Kim, J. So, J. Ahn, S. Ryu, W.-J. Lee, D. Kim, Y.-J. Kim, J. Seok, J.-G. Lee, H.-Y. Ryu, C. Y. Lee, J. Prout, K.-C. Ryoo, S.-J. Han, M.-K. Kook, J.S. Choi, J. Gim, Y. S. Ki, S. Ryu, C. Park, D.-G. Lee, J. Cho, H. Song, and J. Y. Lee

Agenda

- Industry Trends and Challenges
- Introduction of CXL (Compute Express Link)
- CXL Memory Expander Features
- SMDK: Unified Software Solution for CXL
- Application Benchmark Test Results
- Summary and Future Plan

Industry Trends and Challenges



Artificial Intelligence

IT Boom

5G

Big Data

Cloud

Edge

Massive demand for data-centric technologies and applications

Memory bandwidth and density not keeping up with increasing CPU core count

Need a next gen interconnect for heterogeneous computing and server disaggregation

The Fast-Growing Computing Workloads

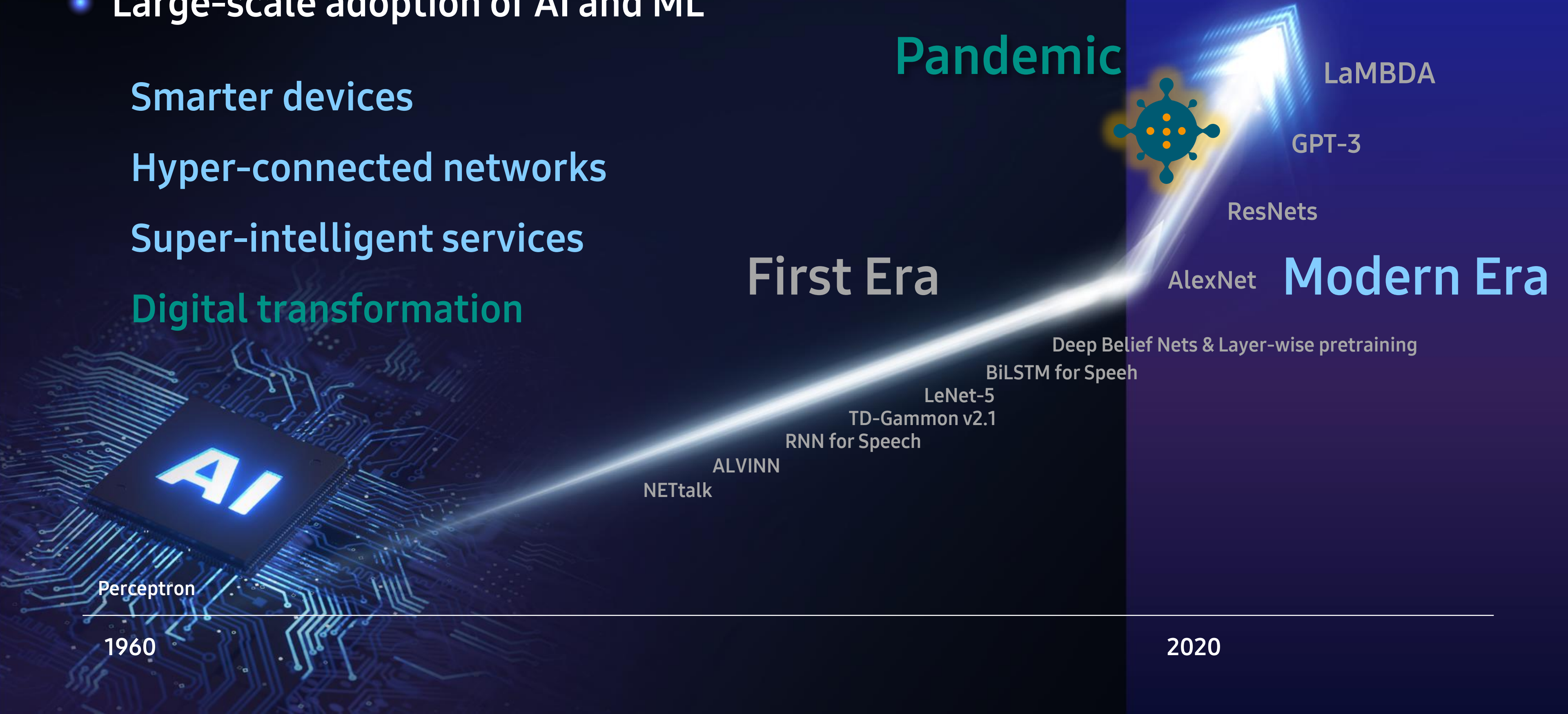
- Large-scale adoption of AI and ML

Smarter devices

Hyper-connected networks

Super-intelligent services

Digital transformation



Pandemic

First Era

Modern Era

Perceptron

NETtalk

ALVINN

RNN for Speech

TD-Gammon v2.1

LeNet-5

BiLSTM for Speeh

Deep Belief Nets & Layer-wise pretraining

AlexNet

ResNets

GPT-3

LaMBDA

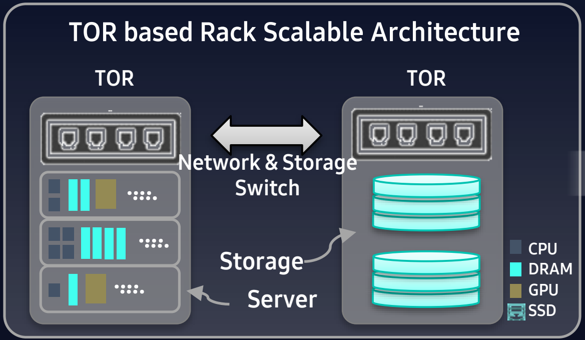
1960

2020

Evolution of Hyperscale Computing Environment

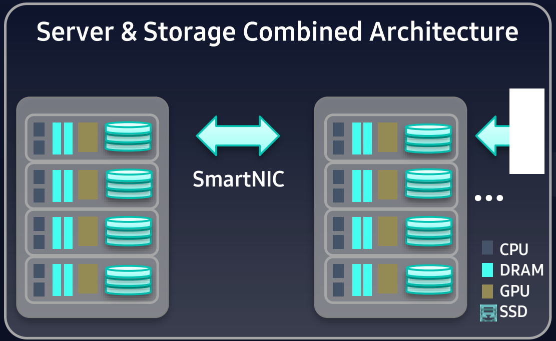
- From Converged to Disaggregated and Composable Architecture

Converged Architecture



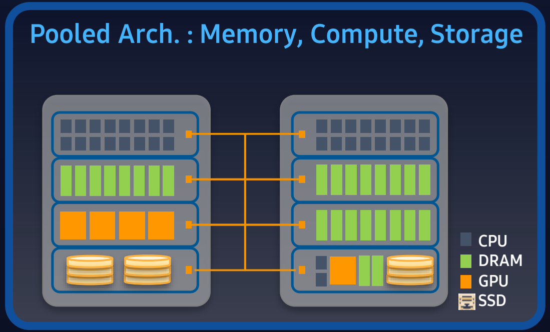
Network Challenge

Hyper-Converged Architecture



Divergence Challenge

Disaggregated / Composable Architecture



Interconnect Challenge

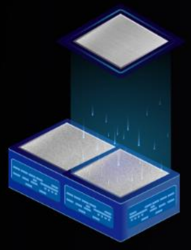
SAMSUNG

The Rising Need for Better Connectivity

- A universal interconnect solution needed for connectivity

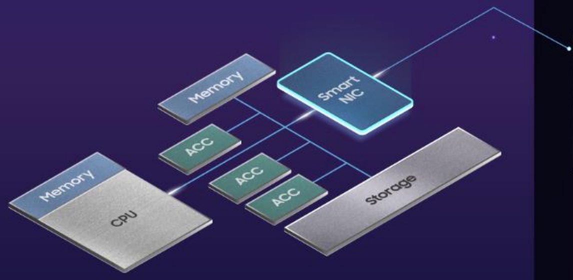
SoC Interconnect

DIE / PACKAGE



Processor Interconnect

NODE



CXL Compute Express Link™

A new class of interconnect
for device connectivity in the era of AI

Data Center Interconnect

DATA CENTER

Customer Interconnect

MOBILE / BROADBAND

CXL: Solution for the Era of HPC

- CXL as the core of composble computing infrastructure

Key Features of CXL Interface

Cache Coherence

Connectivity

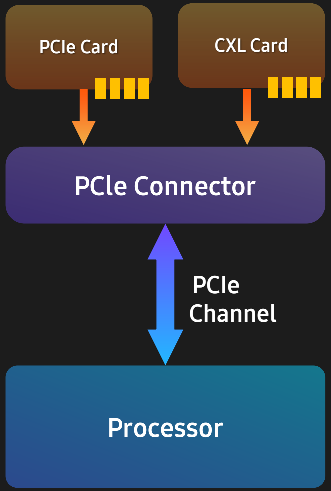
Byte Addressable

Low Latency

CXL Features

SAMSUNG

CXL is a high-performance, low-latency protocol that leverages PCIe physical layer

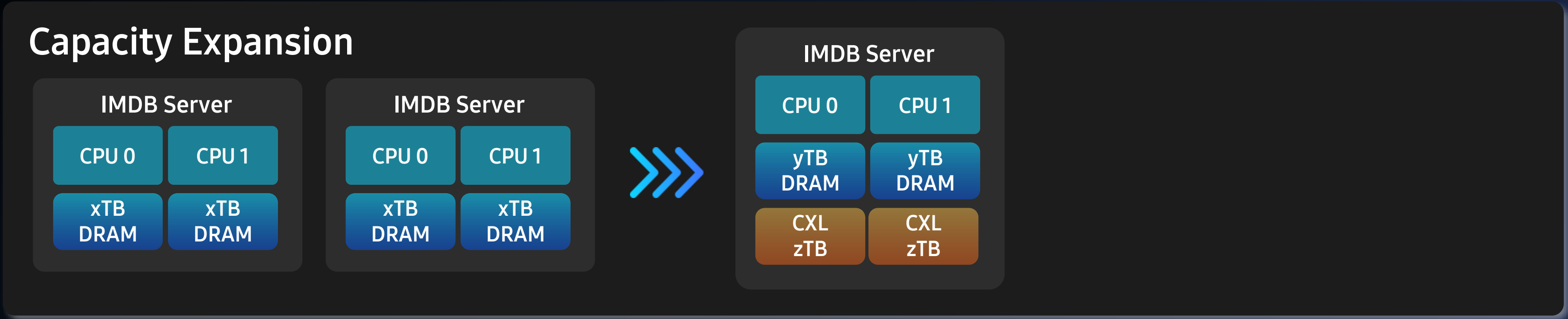


- High-speed and low-latency interconnect
- Leverages PCIe Physical layer (PCIe 5.0, PCIe 6.0)
- Supports various types of memories (volatile, non-volatile)
- CPU and CXL device memory coherency
- Supports switching and memory pooling
- Supports link level integrity and data encryption
- Open standard (non-proprietary)
- Broad industry support in CXL consortium
- Regular specification updates (CXL 1.1, CXL 2.0, CXL 3.0)

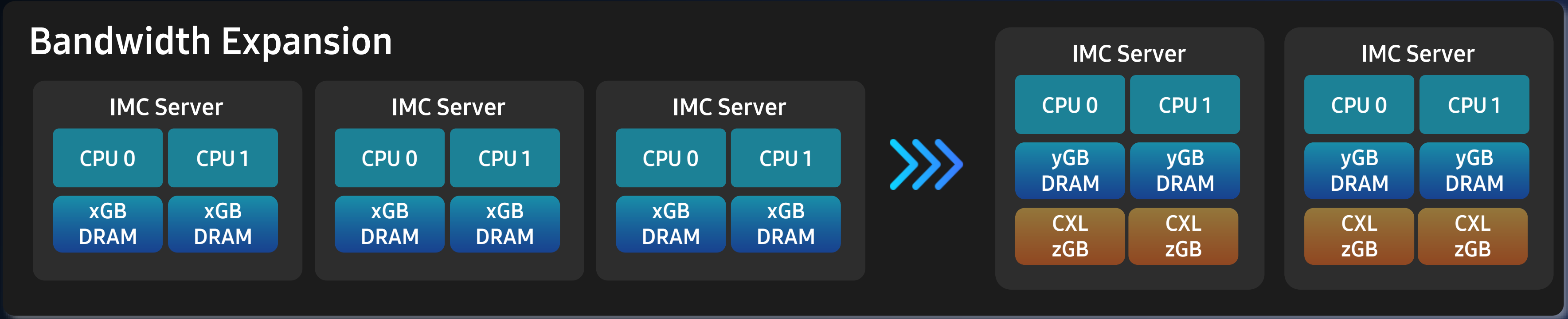
SAMSUNG

CXL Use Cases (1/2)

Capacity and Bandwidth Expansion



Capacity Expansion

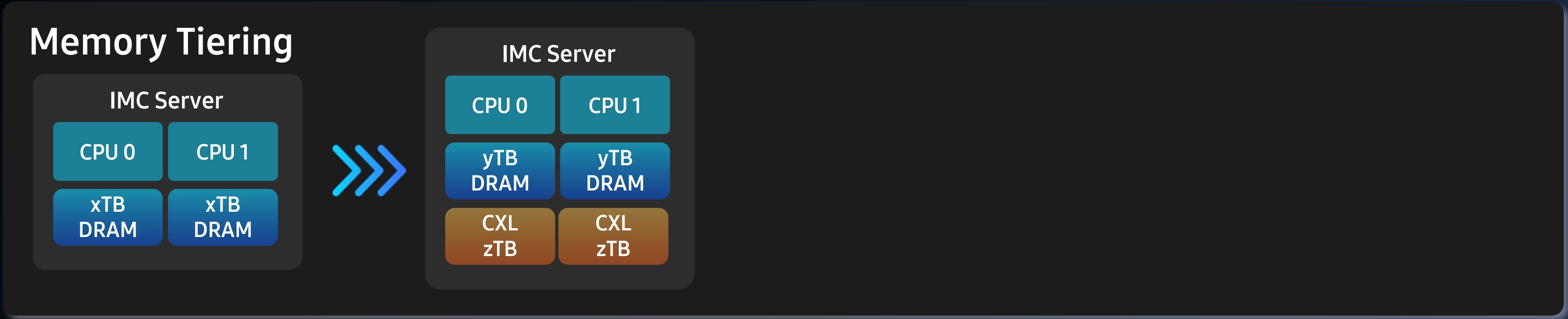


Bandwidth Expansion

CXL Use Cases (2/2)

- Tiering and Pooling

Memory Tiering



IMC Server

CPU 0 | CPU 1

xTB DRAM | xTB DRAM

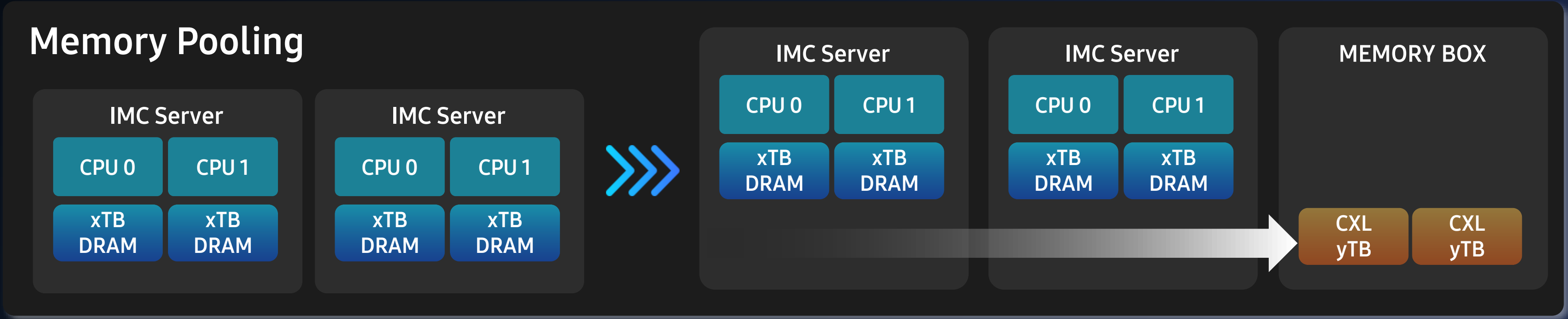
IMC Server

CPU 0 | CPU 1

yTB DRAM | yTB DRAM

CXL zTB | CXL zTB

Memory Pooling



IMC Server

CPU 0 | CPU 1

xTB DRAM | xTB DRAM

IMC Server

CPU 0 | CPU 1

xTB DRAM | xTB DRAM

IMC Server

CPU 0 | CPU 1

xTB DRAM | xTB DRAM

IMC Server

CPU 0 | CPU 1

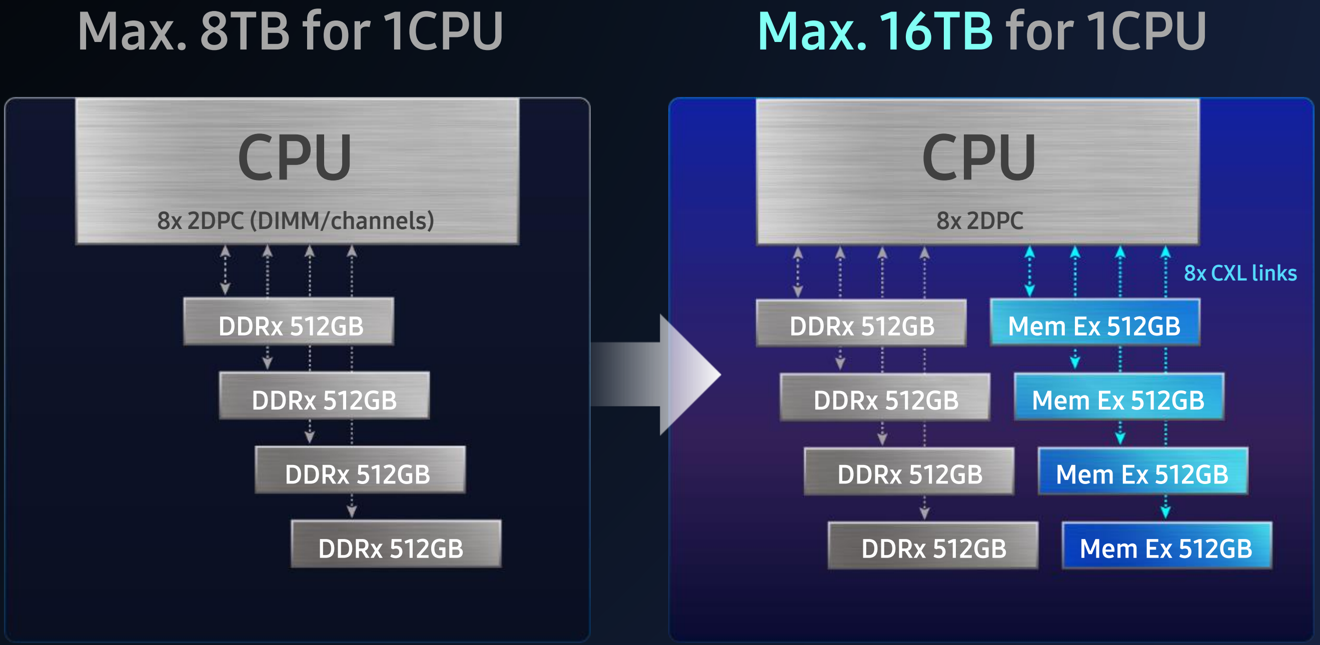
xTB DRAM | xTB DRAM

MEMORY BOX

CXL yTB | CXL yTB

CXL Memory Expansion Solution

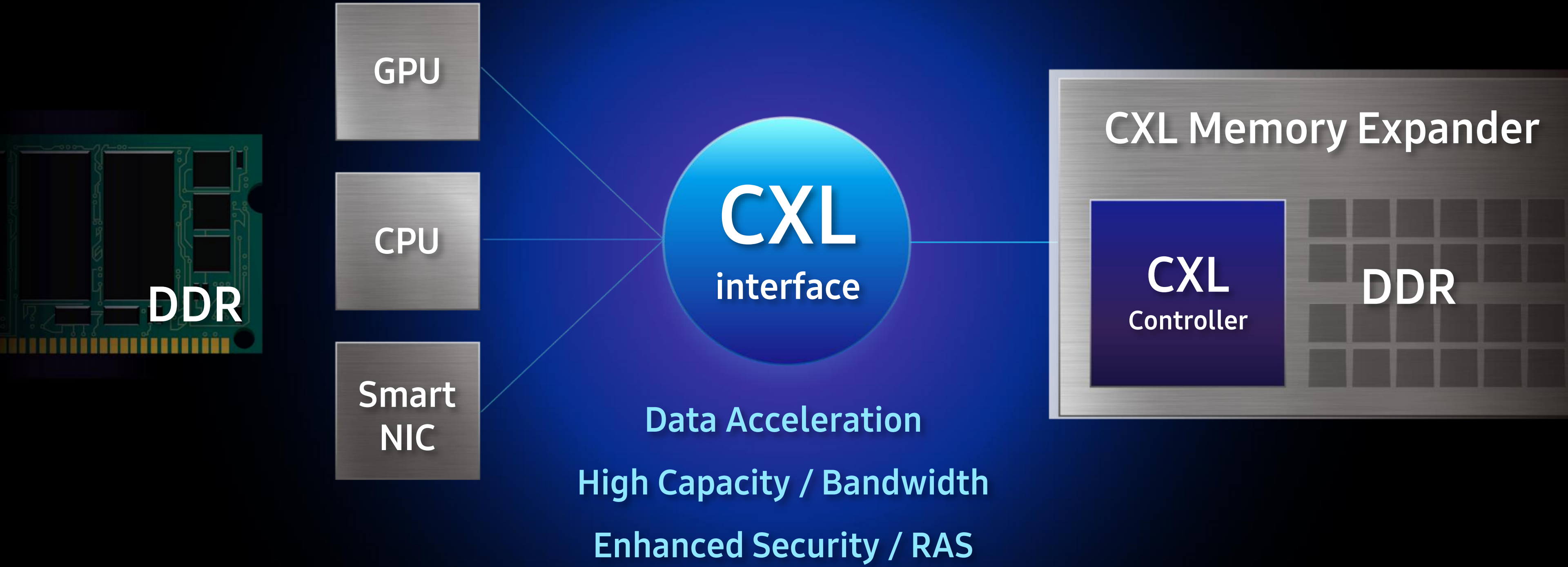
- Doubled Capacity than Conventional Memory



Note: Max capacity varies with system configurations

CXL Memory Expander

- New Solution for Memory Dominant Applications



CXL Memory Expander Line-up

- Built with FPGA and ASIC Controller

’21

’22

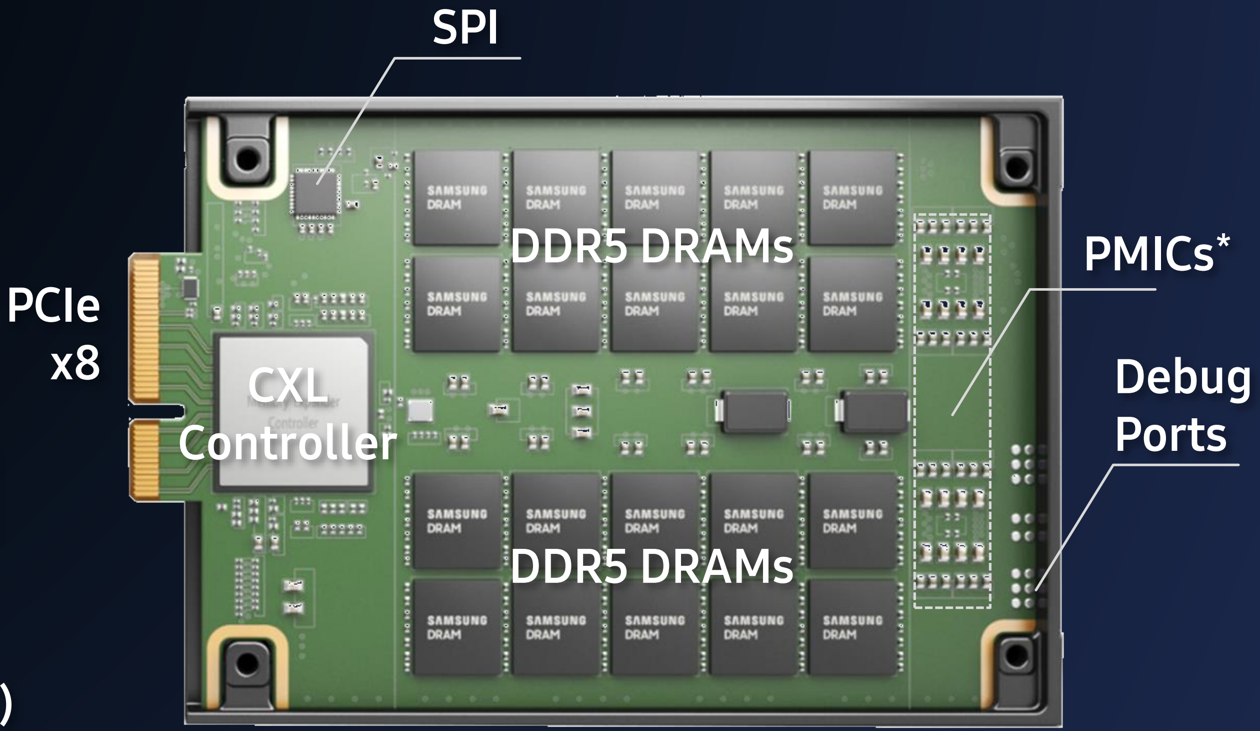
FPGA PCIe 3.0 (x16)	Host (Controller)	ASIC PCIe 5.0 (x8)
DDR4 3200, 128GB	Media (DRAM)	DDR5 4800+, 512GB

As of August, 2022

CXL Memory Expander (1/3)

- Solution Overview

Enclosure (2T)



E3.S Form Factor

* Bottom-side

CXL Memory Expander (2/3)

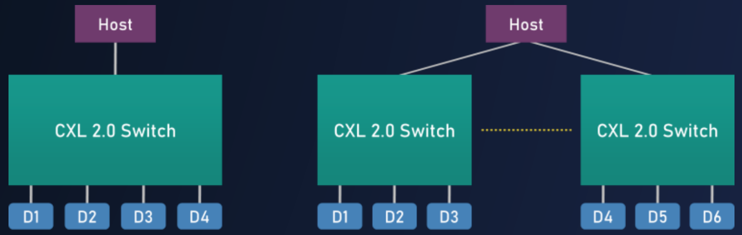
- Product Features

- Form Factor - EDSFF (E3.S)
- Media - DDR5 4800
- Module Capacity - Max 512 GB
- CXL Link Width - x8
- Maximum CXL Bandwidth - 32GB/s (PCIe 5.0)
- Other Features - RAS, Interleaving, Diagnostics etc.
- Availability - Q3'22 for evaluation/testing

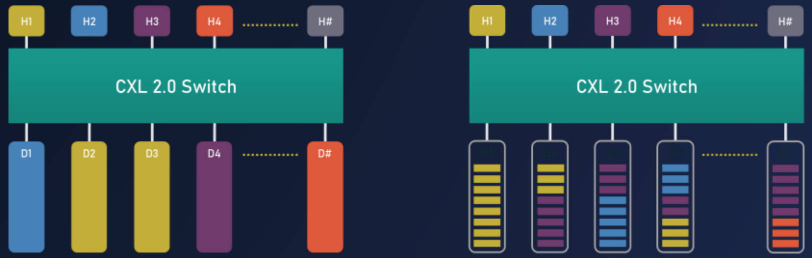
CXL Memory Expander (3/3)

Supported Features

- CXL 2.0
- Device Type: Type 3
- Support viral and data poisoning
- Memory error injection
- Multi-symbol ECC
- Media scrubbing
- Post package repairs (hard/soft)



CXL Memory Pool

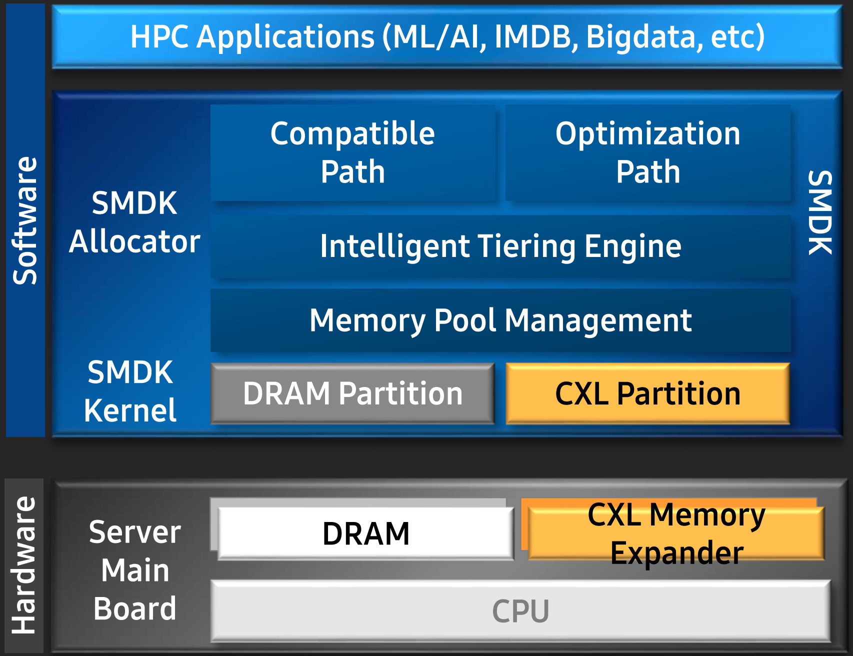


* Image Source: CXL Consortium

SMDK*, Unified Interface for Memory

* Scalable Memory Development Kit

SW development kit to enable Software-Define Memory system on heterogeneous memories



Plugin

Two selectable paths, Compatible and Optimization Path, without or with modification of application SW

Intelligent Tiering Engine supports memory tiering scenarios with prioritiy, capacity, bandwidth, and so on

Memory Pool Management supports scalability reflecting memory request status and system resource

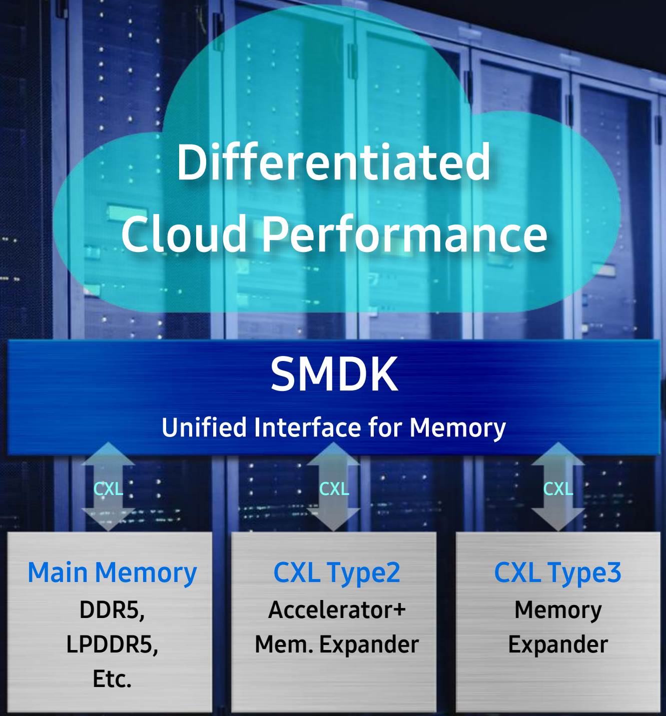
Kernel

Memory Partitioning allows logical memory views for heterogeneous physical DRAM and CXL memory

Benefits of SMDK

SAMSUNG

- **Unified SW Solution**
  Full-stack SW all about heterogenous memory system
- **Client Experience**
  Transparent as well as Optimized Memory uses
- **CXL Ecosystem**
  OSS for CXL Industry and Research field



Differentiated Cloud Performance

SMDK

Unified Interface for Memory

Main Memory
DDR5,
LPDDR5,
Etc.

CXL Type2
Accelerator+
Mem. Expander

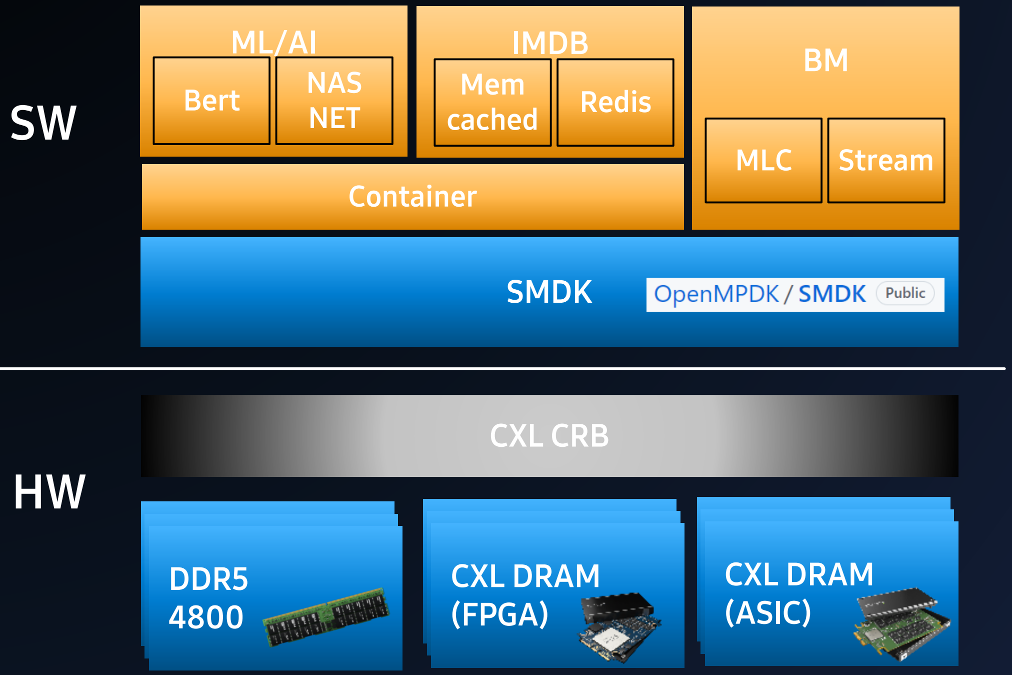
CXL Type3
Memory
Expander

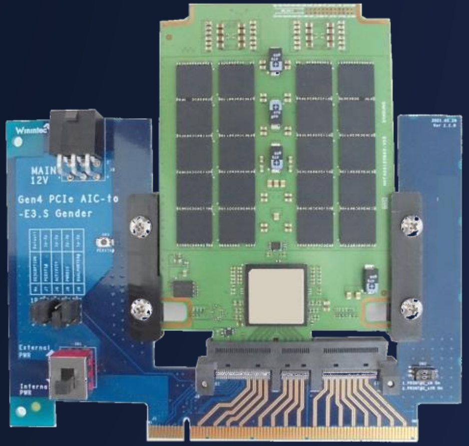
SMDK is available as open source on GitHub

➢ https://github.com/OpenMPDK/SMDK/releases/tag/smdk_v1.1

Experimental Setup

- Configuration of Test Bed



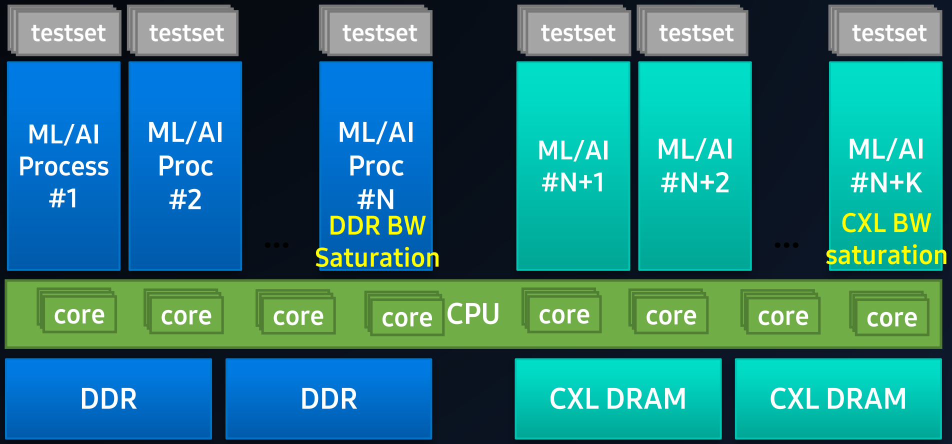


Memory Expander
w/ EDSFF Riser Card

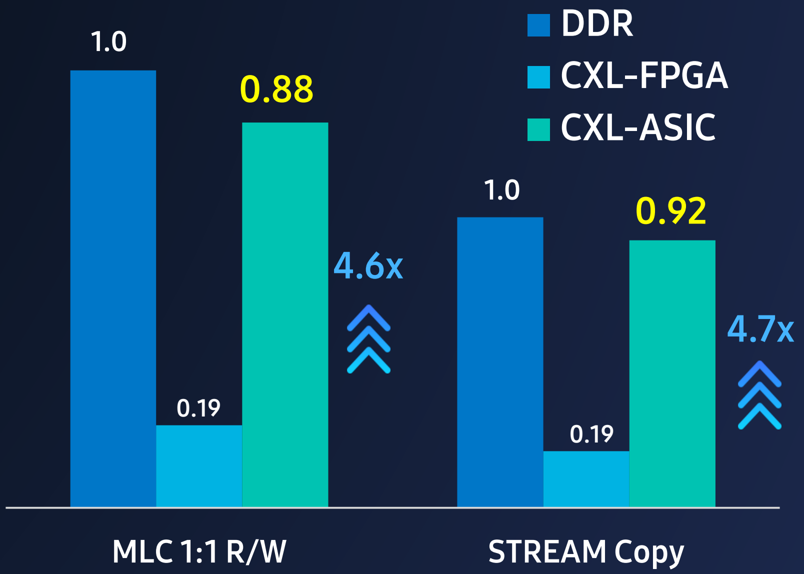
SAMSUNG

Memory Benchmark Test Results

- Comparable Performance with DDR Memory



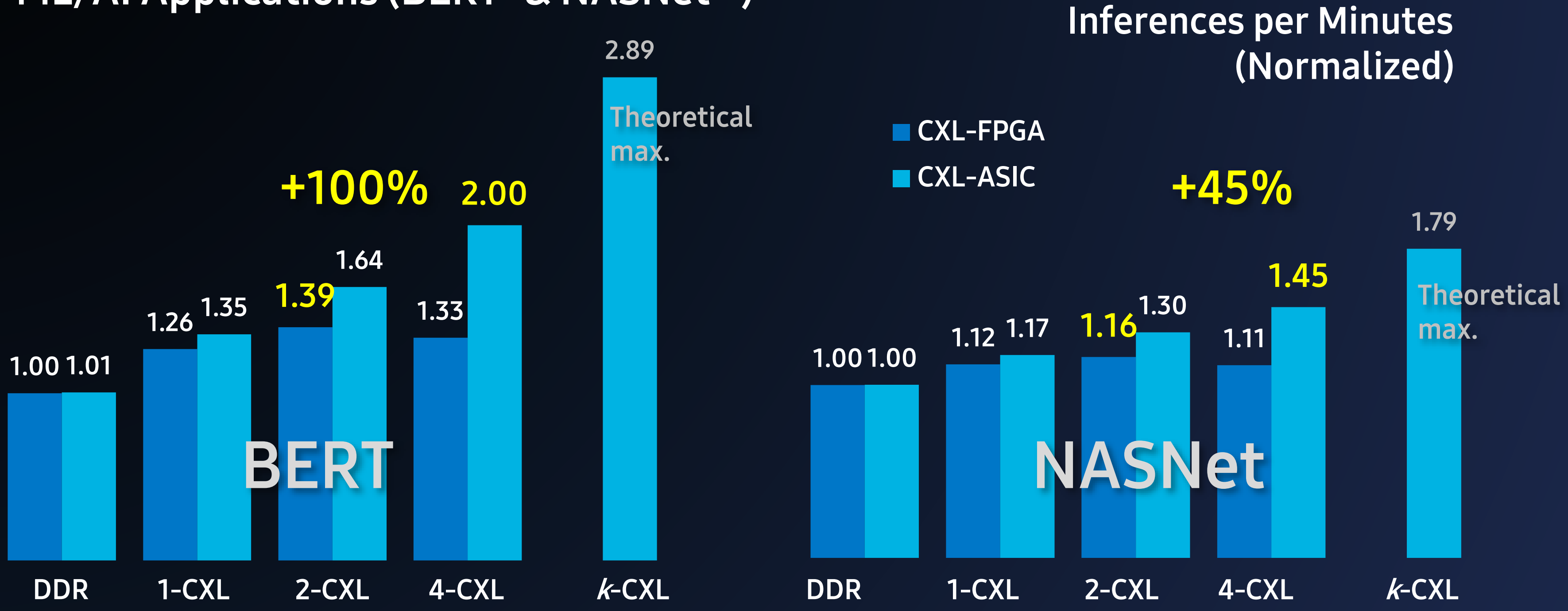
Normalized Bandwidth



SAMSUNG

System Test Results (ML/AI)

- ML/AI Applications (BERT* & NASNet**)



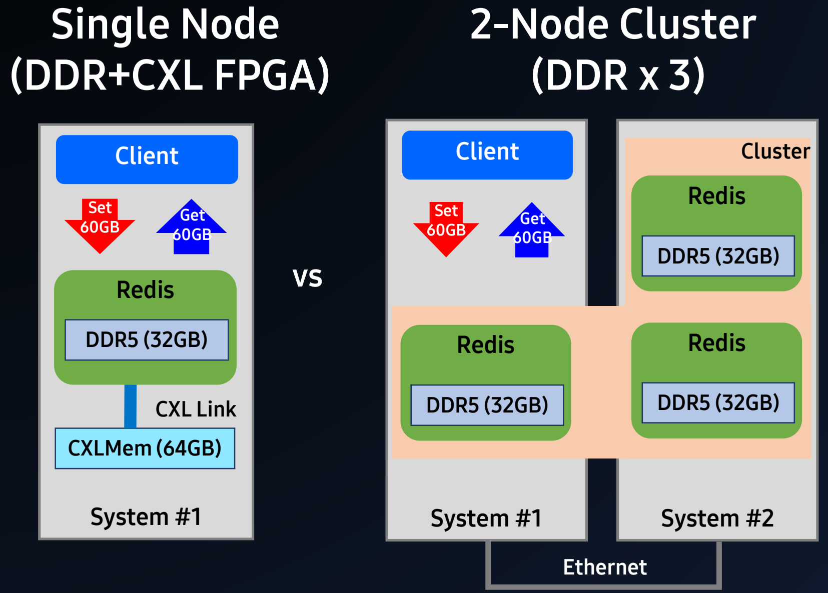
* Bidirectional Encoder Representations from Transformers

** Neural Architecture Search Network

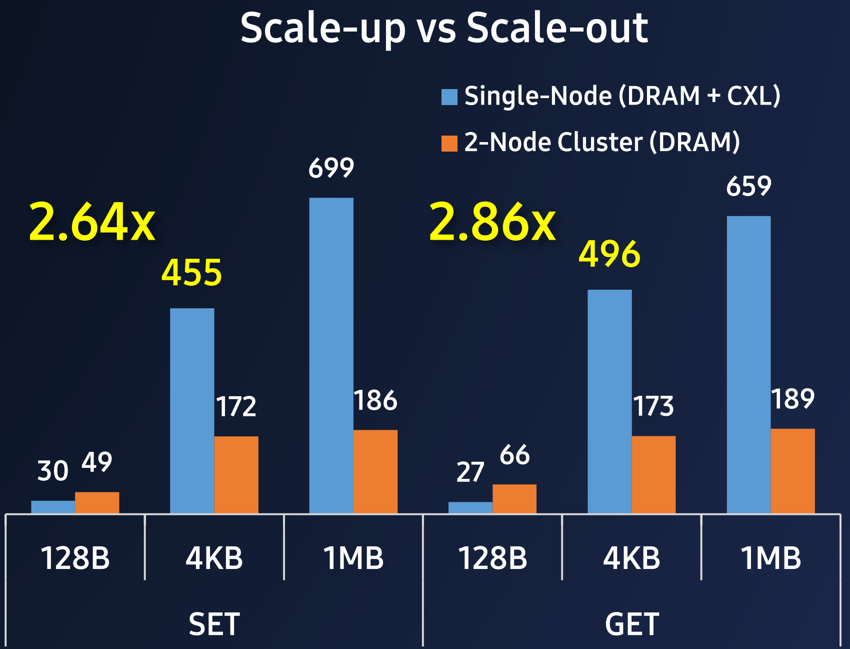
(See appendix for detail test condition)

System Test Results (IMDB)

- IMDB Redis* Memory Usage (Scale-up vs Scale-out)



Performance [MB/s]



* Remote dictionary server

(See appendix for detail test condition)

A Proven Memory Expansion Solution

SAMSUNG

Increasing System Memory Capacity

2X

Increase

Widening Memory Bandwidth

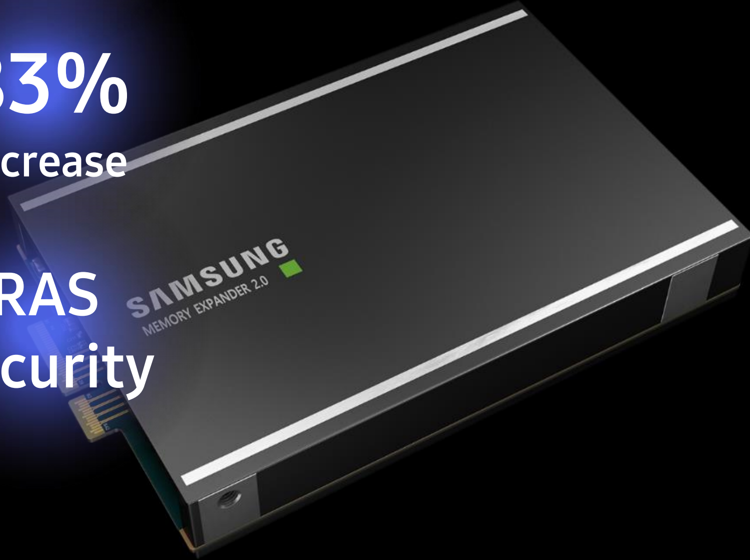
83%

Increase

Supporting RAS/Security based on Memory Controller

RAS

Security



SAMSUNG

Summary and Future Plan

- AI and pandemic drive demand for memory bandwidth and capacity, and new interconnect standard CXL allows expansion of memory
- Samsung developed the industry's first ASIC-based 512GB CXL memory expander, which will be available for early evaluation this quarter
- Memory intensive applications such as IMDB and AI/ML have been tested on CXL memory expander with an open-source software, SMDK
- Samsung to cooperate further on CXL 3.0 and beyond, and provide next-gen memory solutions like memory disaggregation, SDM*, and more

* Software-defined memory

Enhanced Data Service

AI/ML
NLP, Recommendation

Edge Computing

Industry First CXL™ Memory Expanders

SAMSUNG

Appendix

Test Condition (ML/AI and IMDB)

ML/AI

For BERT and Nasnet
TensorFlow (CPU) >= 1.11.0+, Python ~ 3.7, Numpy < 1.20.0

For BERT
Multi-process, 3 cores/process, batch-size:128,
max_seq_num:256, num-test-data/process: 512
dataset=CoLA
do_train=true, do_eval=true, data_dir=$GLUE_DIR/CoLA
vocab_file=vocab.txt
init_checkpoint=$BERT_BASE_DIR/bert_model.ckpt
max_seq_length=128, train_batch_size=32
learning_rate=2e-5, num_train_epochs=3.0

For NASNet
Multi-process, 3 cores/process, batch-size: 100,
eval_image_size:236, num-test-data/process: 200
dataset_name=imagenet, num_preprocessing_threads=4
labels_offset=0, model_name=inception_v3
preprocessing_name=inception_v3
moving_average_decay=None, quantize=False, use_grayscale=False

IMDB

For scale-up vs scale-out

```
Redis-server :master
cluster-enabled yes
cluster-node-timeout 300000
save ""
stop-writes-on-bgsave-error yes
rdbcompression yes
rdbchecksum yes
rdb-del-sync-files no
repl-diskless-sync no
rdb-del-sync-files no
replica-serve-stale-data yes
replica-read-only yes
repl-diskless-sync-delay 5
repl-diskless-load disabled
repl-disable-tcp-nodelay no
replica-priority 100
client-output-buffer-limit replica 0 0 0
maxclients 1000000
maxmemory-policy noeviction
maxmemory-samples 10
maxmemory-eviction-tenacity 100
repl-diskless-sync no  #master-replica간 disk-based sync
rdb-del-sync-files no
replica-serve-stale-data yes
replica-read-only yes
replica-priority 100
client-output-buffer-limit replica 0 0 0
io-threads 4
io-threads-do-reads yes
```

```
Redis-server :replica
save ""
port 6380
replicaof 127.0.0.1 6379
replica-read-only yes
stop-writes-on-bgsave-error yes
rdbcompression yes
rdbchecksum yes
rdb-del-sync-files no
repl-diskless-sync no
rdb-del-sync-files no
replica-serve-stale-data yes
replica-read-only yes
repl-diskless-sync-delay 5
repl-diskless-load disabled
repl-disable-tcp-nodelay no
replica-priority 100
client-output-buffer-limit replica 0 0 0
maxclients 1000000
maxmemory-policy noeviction
maxmemory-samples 10
maxmemory-eviction-tenacity 100
replica-lazy-flush no
lazyfree-lazy-user-del no
lazyfree-lazy-user-flush no
oom-score-adj no
oom-score-adj-values 0 200 800
disable-thp yes
```