Cloud Services I Compute

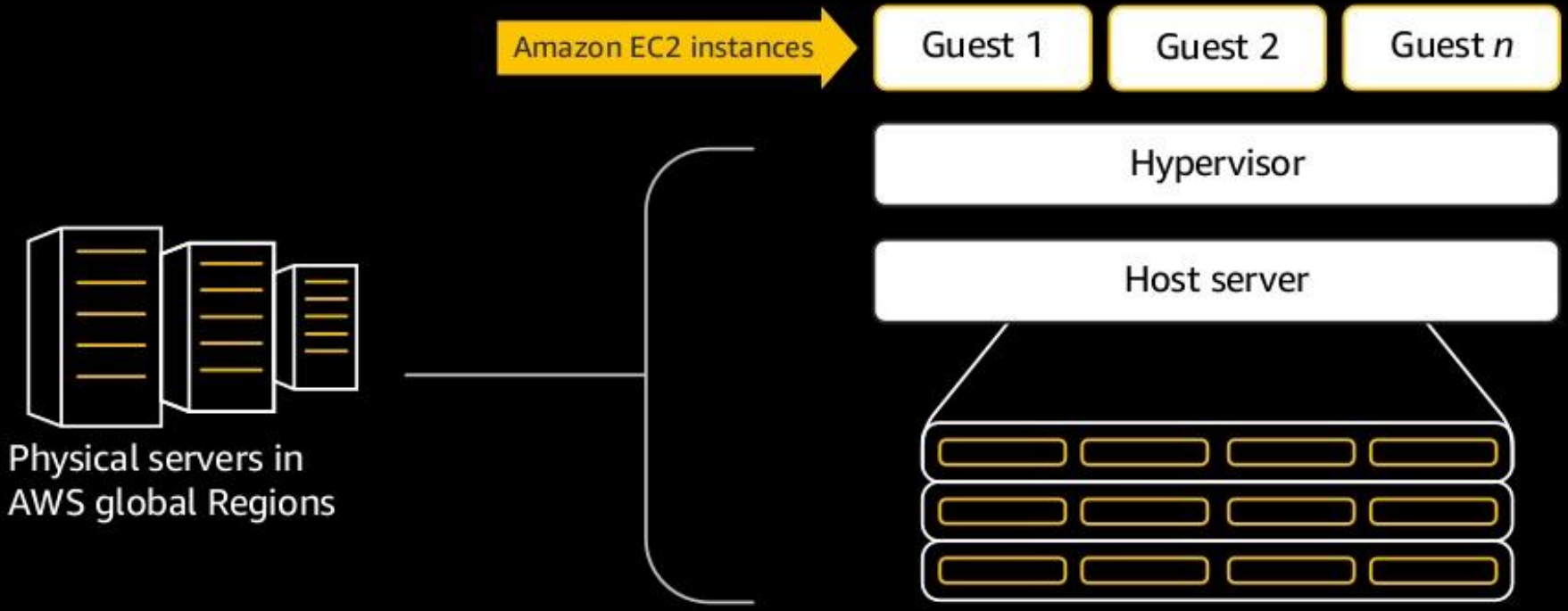
PXL DIGITAL

AWS Cloud Practitioner Essentials (Second Edition): AWS Core Services

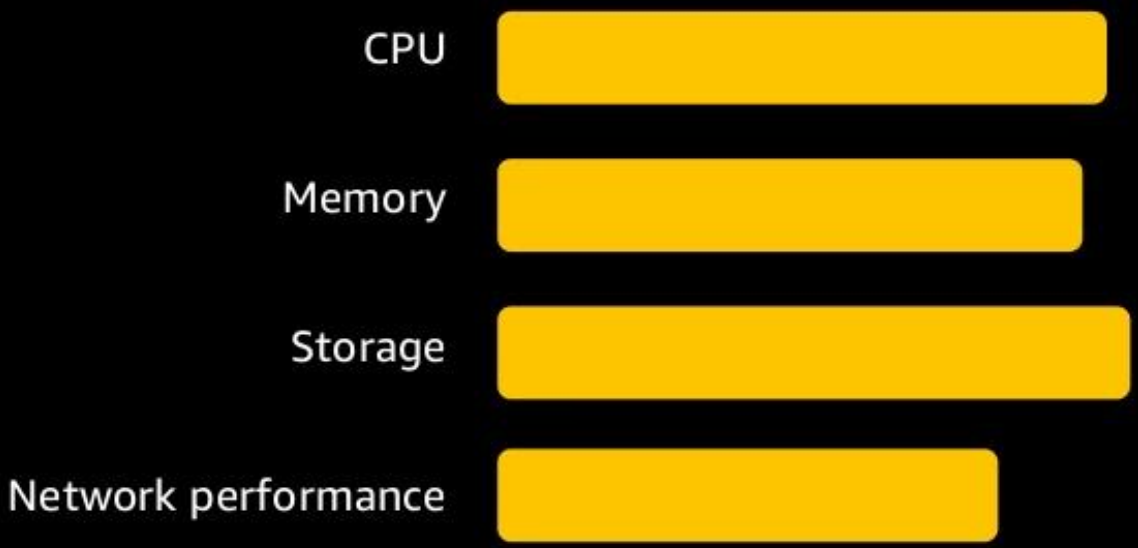
Compute

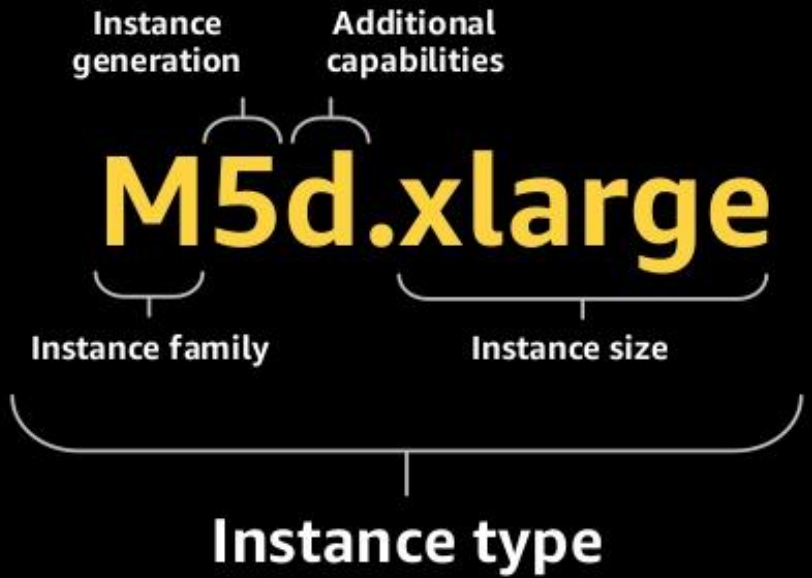
aws SUMMIT

Amazon Elastic Compute Cloud (Amazon EC2)



Amazon EC2 instance characteristics





Amazon machine images (AMIs)

Amazon maintained

Broad set of Linux and Windows images

Kept up-to-date by Amazon in each region

Marketplace maintained

Managed and maintained by AWS Marketplace partners

Your machine images

AMIs you have created from Amazon EC2 instances

Can keep private, share with other accounts, or publish to the community

Choice of processors and architectures

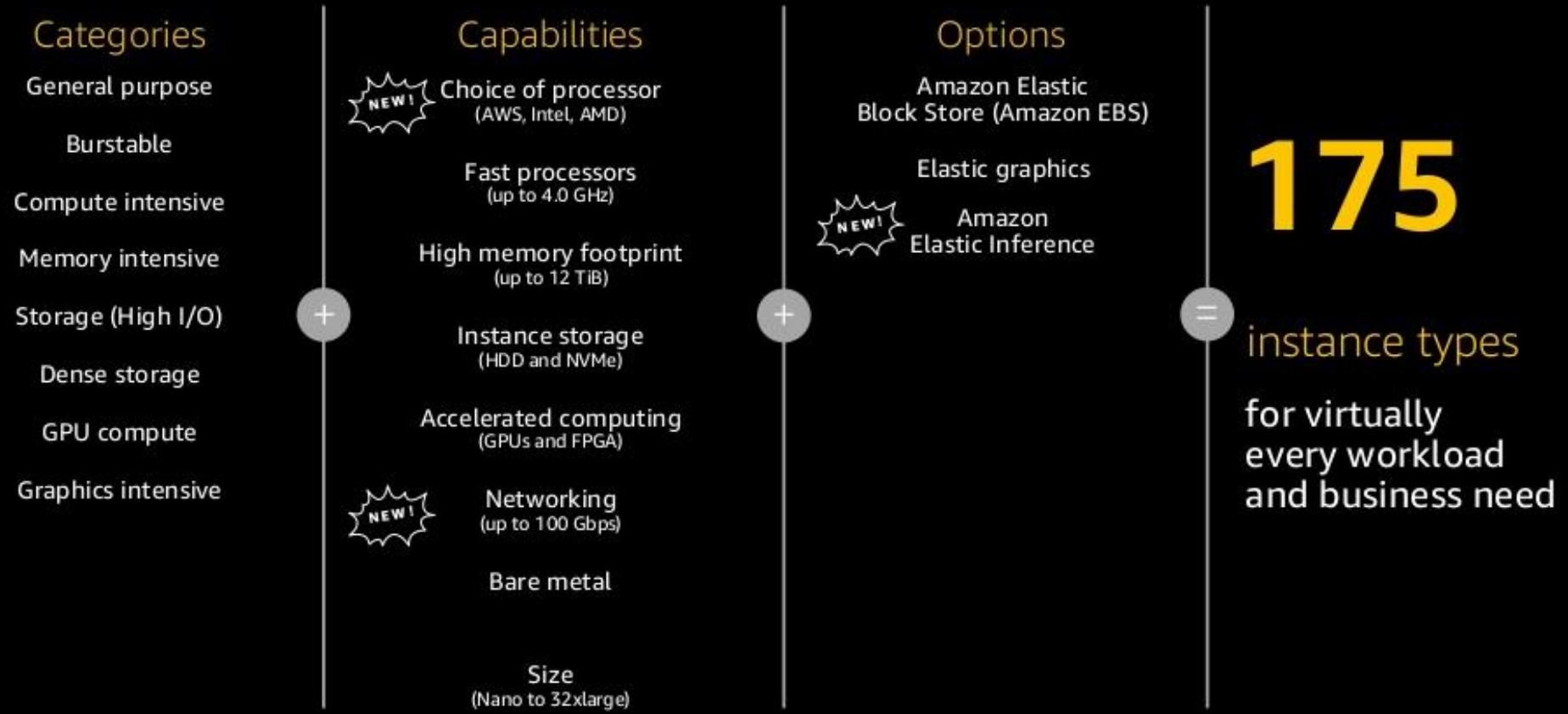
Intel® Xeon® Scalable (Skylake) Processor

AMD EPYC Processor

AWS Graviton Processor based on 64-bit Arm arch

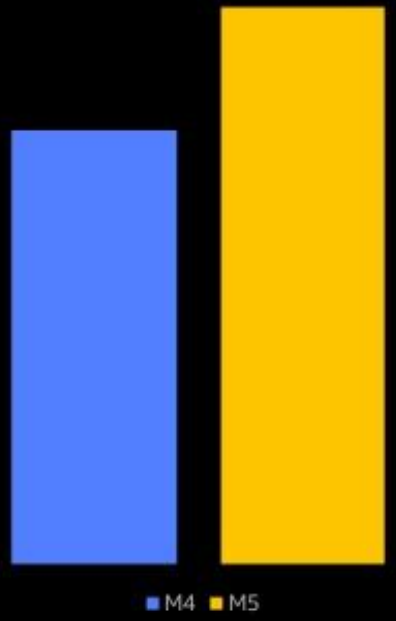
Right compute for each application and workload

Broadest and deepest platform choice



M5: General-purpose instances

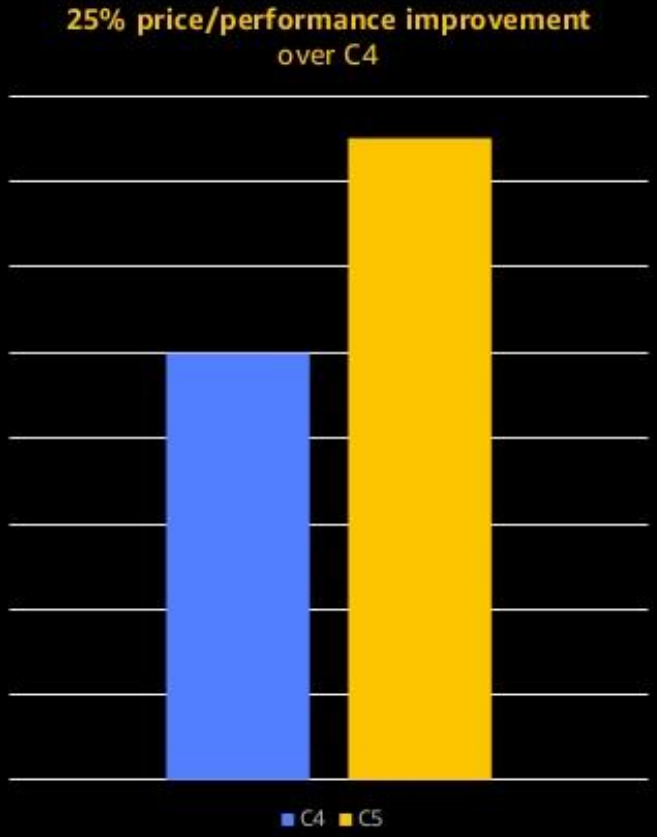
14% price/performance improvement
with M5



- Balance of compute, memory, and networking resources
- Powered by 2.5 GHz Intel Xeon Scalable Processors
- Largest instance size, m5.24xlarge has **96 vCPUs** and **384 GiB of memory**
- Up to 24% instance savings* by migrating from M4 to M5
- Improved network and Amazon EBS performance on smaller sizes
- Support for **Intel AVX-512** offering up to twice the performance for vector and floating point workloads
- M5d variant available with local NVMe-based SSD storage
- M5a variant also available

*Source: https://tsologic.com/wp-content/uploads/2018/11/TSOLogic_Intel_Research_Final_2018-1.pdf

C5: Compute-optimized instances



Custom 3.0 GHz Intel Xeon scalable processors (Skylake)

Up to 72 vCPUs and 144 GiB of memory (2:1 Memory:vCPU ratio)

25 Gbps network bandwidth

Support for Intel AVX-512

C5d with local NVMe-based SSD storage

Up to 50%* AWS instance saving over C4

NETFLIX

"We saw significant performance improvement on Amazon EC2 C5, with up to a 140% performance improvement in industry standard CPU benchmarks over C4."

GRAIL

"We are eager to migrate onto the AVX-512 enabled c5.18xlarge instance size… We expect to decrease the processing time of some of our key workloads by more than 30%."

aws SUMMIT

*Source: https://tsologic.com/wp-content/uploads/2018/11/TSOLogic_Intel_Research_Final_2018-1.pdf

R5: Memory-optimized instances

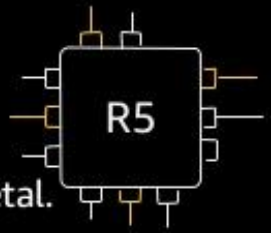
3.1 GHz Intel® Xeon® scalable Processors (Skylake)

Memory-optimized instances with 8:1 GiB to vCPU

Up to 25 Gbps NW bandwidth

R5d instances include up to 3.6 TB of local NVMe SSD

SAP HANA Certified: r5.12xlarge ,r5.24xlarge and r5.metal.



r5.large	7 sizes	r5.24xlarge
16 GiB		768 GiB
2 vCPU		96 vCPU

In-memory caches

High performance databases

Big data analytics

Lower cost for underutilized instances

R5a: Now available with AMD EPYC 7000 processor

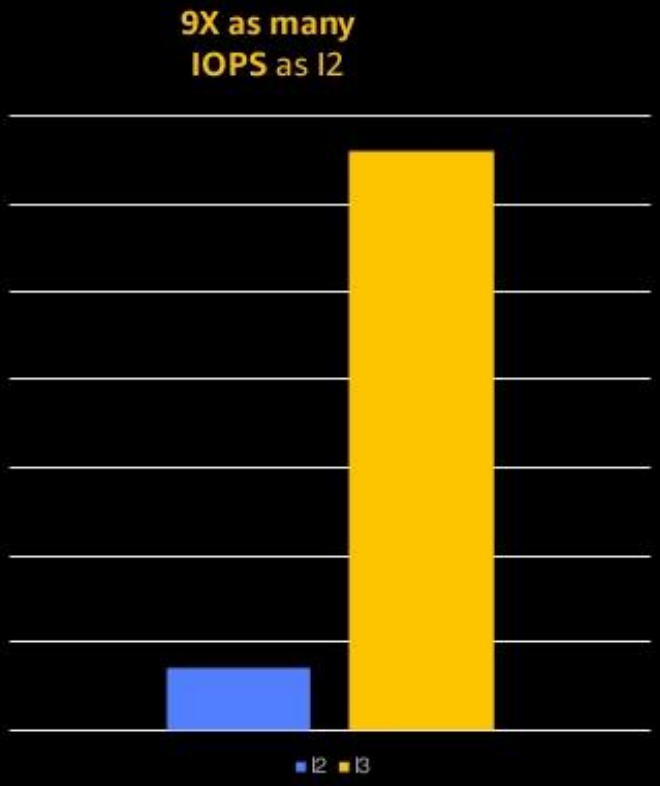
NEW!

R5.metal bare metal instances now available on Intel Xeon scalable processors

Best price to performance

aws SUMMIT

I3: I/O optimized instances



High-perf databases

Real-time analytics

Transactional workloads

No SQL databases

Intel Xeon E5 v4 (Broadwell) processors, with up to 15.2 TB of locally attached NVMe SSD storage, 64 vCPUs, and 488 GiB memory

Lowest cost per IOPS ($/IOPS)

Offers very high random I/O (up to 3.3 million IOPS) and disk throughput (up to 16 GB/s)

Up to 25 Gbps NW bandwidth

Available in bare metal, with i3.metal

aws SUMMIT

T3: Burstable general-purpose instances

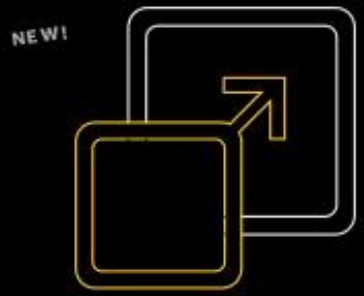
Balance of compute, memory, and network

Baseline level of CPU performance with the ability to burst CPU usage when needed at any time for as long as required

Lowest cost instance at $0.0052 per hour and up to 30% better price performance over T2 using Intel Xeon Scalable Processors

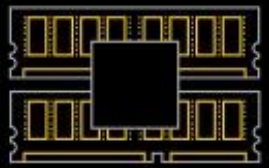
t3.nano
0.5 GiB
2 vCPU
Base perf 5%

t3.2xlarge
32 GiB
8 vCPU
Base perf 40%



With T3 unlimited bursting over baseline is only $0.05 per vCPU-hour, averaged over 24 hours

X1 and X1e: Large-scale memory-optimized

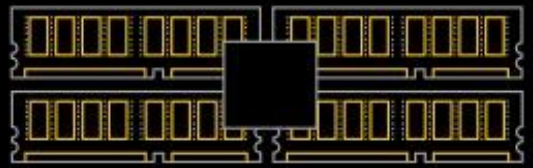


For large in-memory workloads

16:1 GiB to vCPU ratio

In-memory databases (e.g., SAP HANA), big data processing engines (Apache Spark, Presto), in-memory analytics

x1.16xlarge	x1.32xlarge
1 TB	2 TB
64 vCPU	128 vCPU



For memory-intensive workloads and very large in-memory workloads

32:1 GiB to vCPU ratio

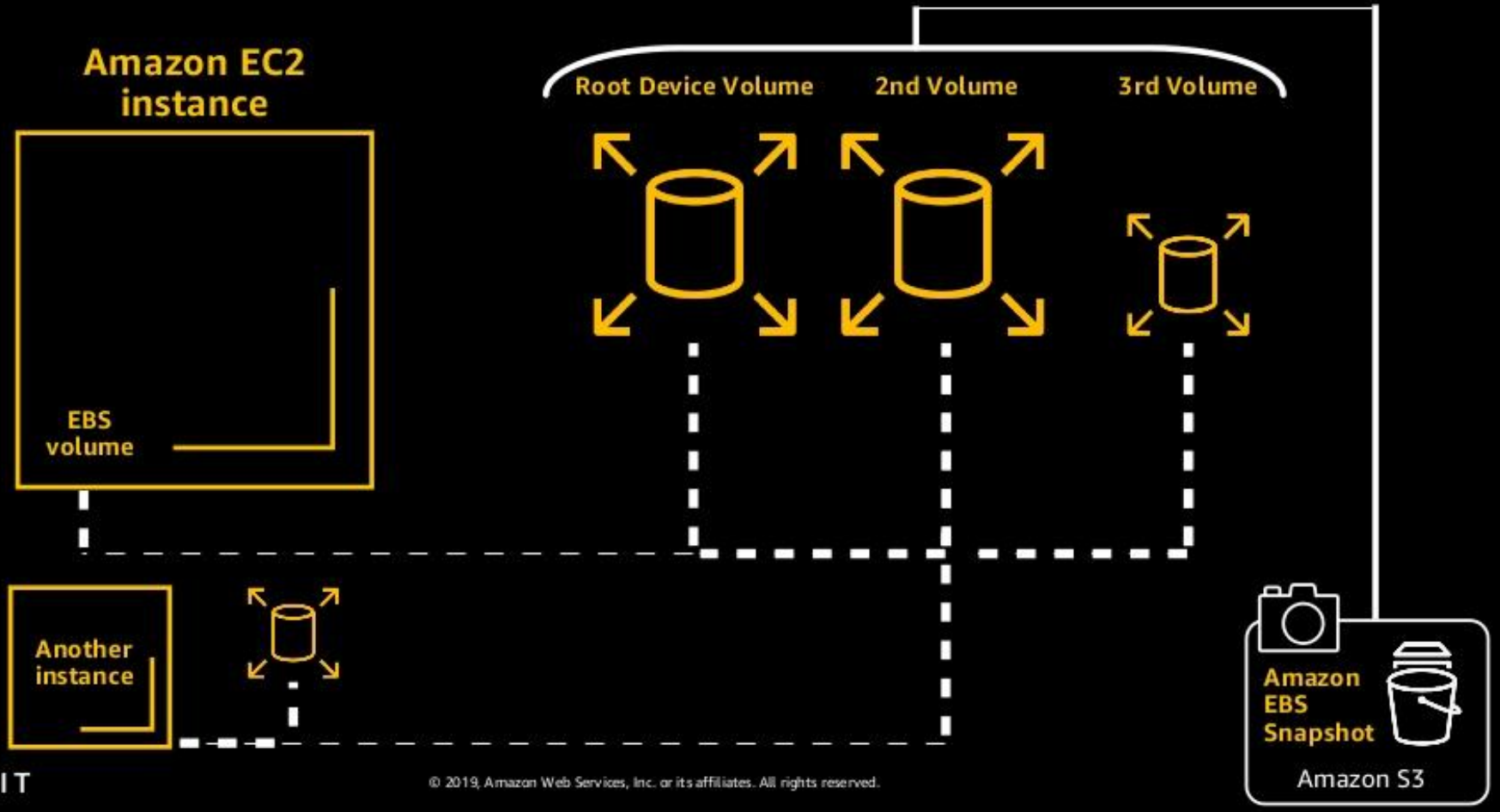
High-performance databases, large in-memory databases (e.g., SAP HANA), and DB workloads with vCPU based licensing (Oracle, SAP)

x1e.xlarge	6 sizes	x1e.32xlarge
122 GiB	● ● ●	4 TB
4 vCPU		128 vCPU

X1e instances are powered by four Intel® Xeon® E7 8880 v3 processors offering up to 128 vCPUs and 3,904 GiB of DRAM-based memory.

aws SUMMIT

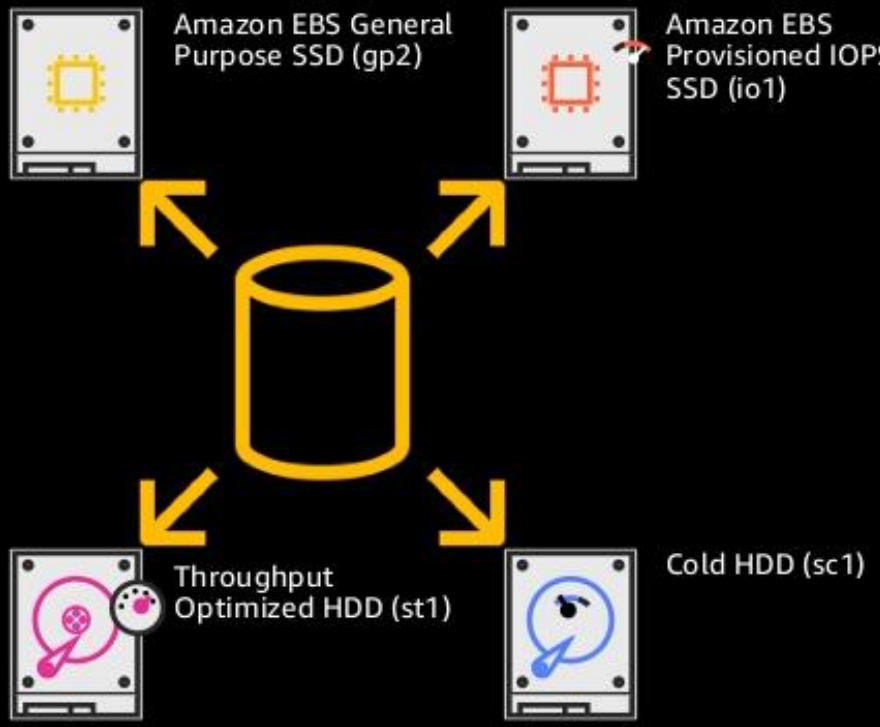
Amazon Elastic Block Store (Amazon EBS)



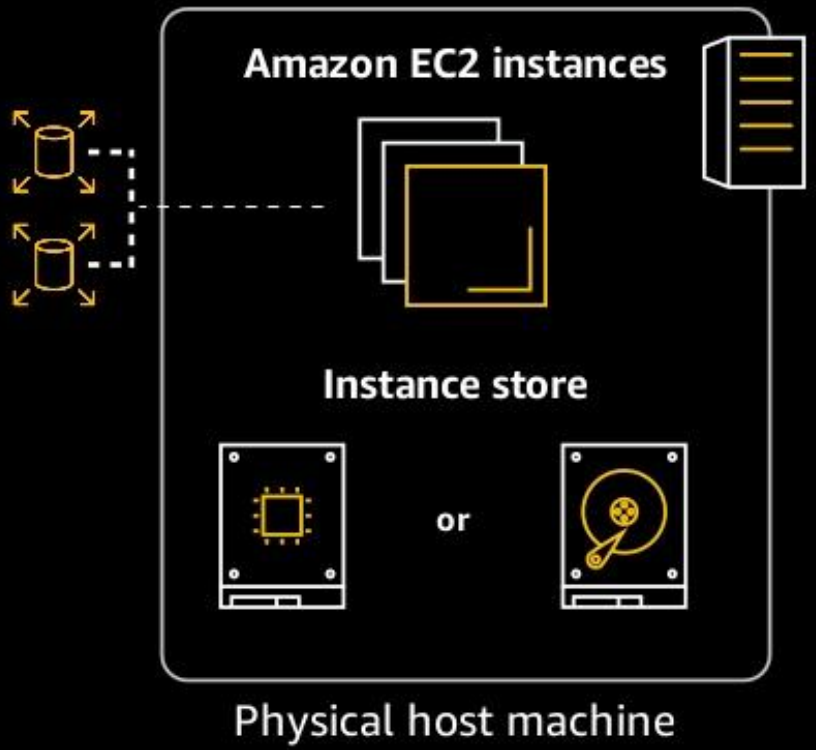
Amazon Elastic Block Store (Amazon EBS)

Amazon EBS SSD-backed volumes

Amazon EBS HDD-backed volumes



Amazon EC2 instance store



Local to instance

Non-persistent data store

Data not replicated (by default)

No snapshot support

SSD or HDD

Compute labs

- Online lab platform
  - AWSGen
  - Launch & configure an EC2 instance with the Wizard
  - Configure security groups on AWS EC2 instances
  - Create an AWS security group for a bastion server to manage EC2 instances