

List all Instance Type.

1. General Purpose Instances

These instances provide a balance of compute, memory, and networking resources and can be used for a variety of diversified workloads.

- **T Series (Burstable Performance Instances):**
 - **t3, t3a, t2** - Cost-effective instances that provide a baseline level of CPU performance with the ability to burst to higher levels when needed.
- **M Series (Balanced Memory and Compute):**
 - **m6g, m5, m5a, m5n, m5zn, m4** - Instances with a balance of compute, memory, and networking resources for a broad range of applications, including web servers, application servers, and databases.

2. Compute Optimized Instances

These instances are ideal for compute-intensive applications.

- **C Series (Compute Optimized):**
 - **c7g, c6g, c6gd, c6i, c5, c5a, c5n, c4** - Designed for compute-heavy workloads like high-performance web servers, scientific modeling, batch processing, and video encoding.

3. Memory Optimized Instances

These instances are designed for memory-intensive applications, such as high-performance databases and real-time big data analytics.

- **R Series (Memory Optimized):**
 - **r6g, r5, r5a, r5n, r4** - High memory-to-CPU ratio, ideal for workloads such as high-performance databases, in-memory caches, and data analysis.
- **X Series (Extreme Memory Optimized):**
 - **x1e, x1** - Designed for large-scale, in-memory, and SAP HANA applications.
- **u Series (High Memory Instances):**
 - **u-6tb1.metal, u-9tb1.metal, u-12tb1.metal** - Bare metal instances with massive memory (up to 12 TB) for memory-bound applications.

4. Storage Optimized Instances

These instances are designed for workloads that require high, sequential read and write access to very large datasets on local storage.

- **I Series (Storage Optimized):**
 - **i3, i3en, i4i** - High-performance local storage optimized for applications such as NoSQL databases and transactional workloads.
- **D Series (Dense Storage Instances):**
 - **d2** - Storage-intensive workloads, such as big data analytics, data warehousing, and log processing.
- **H Series (High Storage):**
 - **h1** - Optimized for high-throughput, low-latency storage with instances suited for data-intensive applications.

5. GPU Instances

GPU instances are used for applications requiring graphical processing power, such as machine learning, deep learning, 3D rendering, and video transcoding.

- **P Series (GPU Optimized):**
 - **p4, p3** - Instances designed for machine learning, high-performance computing, and deep learning applications.
- **G Series (Graphics Optimized):**
 - **g5, g4ad, g4dn** - Instances designed for graphics-intensive applications such as video transcoding, 3D rendering, and gaming.
- **Inf1 (Inference Optimized):**
 - **inf1** - Designed specifically for machine learning inference workloads.

6. High Performance Computing (HPC) Instances

HPC instances are optimized for running highly parallel workloads like simulations, financial modeling, and scientific research.

- **H Series (High Performance Computing):**
 - **hpc6id** - Instances designed for high-performance computing workloads with dense storage for data-intensive calculations.

7. Bare Metal Instances

Bare metal instances provide direct access to the underlying hardware, allowing for workloads that require full control over the hardware.

- **i3.metal**
- **m5.metal**
- **c5.metal**
- **r5.metal**

These instances allow users to run applications that need to access hardware features directly, such as specialized workloads that require virtualization-free environments.

8. Networking Optimized Instances

These instances are designed to handle high throughput and low latency networking.

- **n2 instances (NVIDIA-powered network-optimized):** Used for workloads requiring high-performance networking and GPU resources, such as machine learning and high-performance computing.

Summary of Instance Families and Their Key Use Cases:

- **General Purpose:** Balanced compute, memory, and networking (e.g., t3, m5).
- **Compute Optimized:** CPU-intensive workloads (e.g., c6g, c5).
- **Memory Optimized:** Memory-intensive workloads (e.g., r5, x1).
- **Storage Optimized:** Storage-heavy applications (e.g., i3, d2).
- **GPU Instances:** GPU-based workloads (e.g., p3, g4dn).
- **Bare Metal Instances:** Full control over the hardware (e.g., i3.metal, m5.metal).
- **High Performance Computing:** HPC workloads (e.g., hpc6id).