Case Study 1: General Purpose Instances T & M Series

1. Overview of Recommended Instance

General purpose instances offer a balanced ratio of **compute, memory, and networking** resources, making them ideal for applications that require a mix of these. In this case, **T3, T4g, M5, and M6i** instances are recommended.

- **T-Series (T3, T4g):** Burstable performance instances with a baseline CPU and the ability to burst when needed. Cost-efficient and suited for moderate workloads.

- **M-Series (M5, M6i):** Steady performance instances for general-purpose workloads requiring consistent compute and memory capacity.

These instance families are ideal for **startups** and small to medium-scale web applications needing reliable and affordable cloud infrastructure.

2. Characteristics

Instance Type	vCPUs	Memory (GiB)	Network Performance	Processor Type	Key Feature
T3	2–8	1–32	Up to 5 Gbps	Intel Xeon	Burstable, cost-effective
T4g	2–8	1–32	Up to 5 Gbps	AWS Graviton2	Arm-based, 40% better price/perf
M5	2–96	8–384	Up to 25 Gbps	Intel Xeon	Balanced and consistent
M6i	2–128	8–512	Up to 50 Gbps	Intel Xeon (Ice Lake)	New-gen, better perf/price

- **T-Series:** Ideal for spiky workloads. Credits system allows saving costs on idle time.

- **M-Series:** Ideal for steady-state workloads with reliable performance.

3. Why They Are Suitable

Scenario A – SaaS Project Management Tool:

- **T3/T4g** instances are ideal for collaborative tools with **bursty, interactive usage**, where compute load can fluctuate based on user activity.

- **T4g** offers lower costs with good performance using Graviton2.

Scenario B – Small E-commerce Website:

- **M5/M6i** provides a good base for **consistent performance**, especially during predictable traffic patterns.

- **Auto Scaling** can be applied on top for dynamic traffic loads without over-provisioning.

4. Consideration Detailing the Instance Series

- **T3/T4g:**

  - Use **CPU credit system**; not ideal for sustained high-CPU applications.

  - T4g supports **ARM architecture**, which requires compatible software packages.

  - Ideal for dev/test, microservices, small web servers.

- **M5/M6i:**

  - Provide **fixed performance**, suitable for backend services and larger web apps.

  - **M6i** offers better performance at a lower cost than M5 due to newer generation hardware.

  - Broader OS support and compatibility.

5. Comparison and Selection

Criteria	T3/T4g	M5/M6i
Performance Consistency	Moderate (burstable)	High (steady)
Cost Efficiency	High for low-moderate	Better for consistent
Scalability	Good with Auto Scaling	Excellent with Auto Scaling
Architecture	Intel (T3), ARM (T4g)	Intel Xeon (latest gen)

Selection Guidance:

- **T4g** for cost-sensitive, compatible, and bursty workloads (Scenario A).

- **M5/M6i** for predictable and moderate traffic (Scenario B), or when compatibility with x86 is required.

6. Key Considerations Supporting the Business Case

- **Startup-focused Strategy:**

  o Start with **T4g** for development and early deployment for reduced cost.

  o As traffic grows or stability is critical, migrate to **M5 or M6i**.

- **Cost Control:**

  o T-Series allows for **lower operational costs** with occasional CPU spikes.

  o M-Series offers **predictable billing** with Reserved Instances or Savings Plans.

- **Future Readiness:**

  o M6i allows for horizontal scaling and handles enterprise-grade loads.

  o T4g introduces ARM-based optimization, preparing for long-term cloud-native architecture.

7. Conclusion

The selection of **T3, T4g, M5, and M6i** instances meets the startup's needs by offering **balanced compute resources** for both dynamic and steady workloads.

- **For early development or low traffic,** T4g provides excellent cost-effectiveness.

- **As user load grows**, migrating to **M5 or M6i** ensures performance and scalability.

- This hybrid approach supports the startup's **agility, budget constraints, and growth roadmap**, making general-purpose instances the best fit for both given scenarios.