



## ❖ SIMPLE QUEUE SERVICES (SQS):

- SQS is a reliable, scalable, fully-managed **message queuing service**.
- It enables you to decouple and scale microservices, distributed systems, and serverless applications.
- SQS **send, store, and receive messages** between software components at any volume, without losing messages.
- Amazon SQS allows producers to send messages to a queue. Messages are then stored in an SQS Queue.
- When consumers are ready to process new messages, they poll them from the queue.
- Applications, microservices, and multiple AWS services can take the role of producers or consumers.

## ➤ BENEFITS AND FEATURES:

- Highly scalable Standard and FIFO queues
- Durability
- Availability
- Security
- Batching

## ➤ USE CASES:

- Increase application reliability and scale
- Decouple microservices and process event-driven applications
- Ensure work is completed cost-effectively and on time
- Maintain message ordering with deduplication

## ➤ MESSAGE QUEUE TYPES:

- Amazon SQS supports two types of queues.
  - standard queues
  - FIFO queues.

### **STANDARD QUEUES:**

- It offers maximum throughput, best-effort ordering, and at-least-once delivery.
- UNLIMITED THROUGHPUT:** Standard queues support a nearly unlimited number of API calls per second, per API action (SendMessage, ReceiveMessage, or DeleteMessage).

**AT-LEAST-ONCE DELIVERY:** A message is delivered at least once, but occasionally more than one copy of a message is delivered.

**BEST-EFFORT ORDERING:** Occasionally, messages are delivered in an order different from which they were sent.

### **FIFO QUEUES:**

- These are designed to guarantee that messages are processed exactly once, in the exact order that are sent.

**HIGH THROUGHPUT:** If you use batching, FIFO queues support up to 3,000 messages per second, per API method (SendMessageBatch, ReceiveMessage, or DeleteMessageBatch). The 3,000 messages per second represent 300 API calls, each with a batch of 10 messages. To request a quota increase, submit a support request. Without batching, FIFO queues support up to 300 API calls per second, per API method (SendMessage, ReceiveMessage, or DeleteMessage).

**EXACTLY-ONCE PROCESSING:** A message is delivered once and remains available until a consumer processes and deletes it. Duplicates aren't introduced into the queue.

**FIRST-IN-FIRST-OUT DELIVERY:** The order in which messages are sent and received is strictly preserved.

## ➤ STANDARD QUEUE Vs FIFO QUEUE:

- Amazon SQS supports two types of queues – standard queues and FIFO queues.

Standard Queue	FIFO Queue
<b>Unlimited Throughput</b> – Standard queues support a nearly unlimited number of transactions per second (TPS) per API action.	<b>High Throughput</b> – By default, FIFO queues support up to 3,000 messages per second, per API action ( <code>SendMessage</code> , <code>ReceiveMessage</code> , or <code>DeleteMessage</code> ), with <b>batching</b> . To request a limit increase, <a href="#">file a support request</a> . FIFO queues support up to 300 messages per second, per API action ( <code>SendMessage</code> , <code>ReceiveMessage</code> , or <code>DeleteMessage</code> ) without batching.
<b>At-Least-Once Delivery</b> – A message is delivered at least once, but occasionally more than one copy of a message is delivered.	<b>Exactly-Once Processing</b> – A message is delivered once and remains available until a consumer processes and deletes it. Duplicates aren't introduced into the queue.
<b>Best-Effort Ordering</b> – Occasionally, messages might be delivered in an order different from which they were sent.	<b>First-In-First-Out Delivery</b> – The order in which messages are sent and received is strictly preserved.
<b>Durability and redundancy</b> – Standard queues ensure high durability by storing multiple copies of each message across multiple AWS Availability Zones. This ensures that messages are not lost, even in the event of infrastructure failures.	
<b>Visibility timeout</b> – Amazon SQS allows you to configure a visibility timeout to control how long a message stays hidden after being received, ensuring that other consumers do not process the message until it has been fully handled or the timeout expires.	