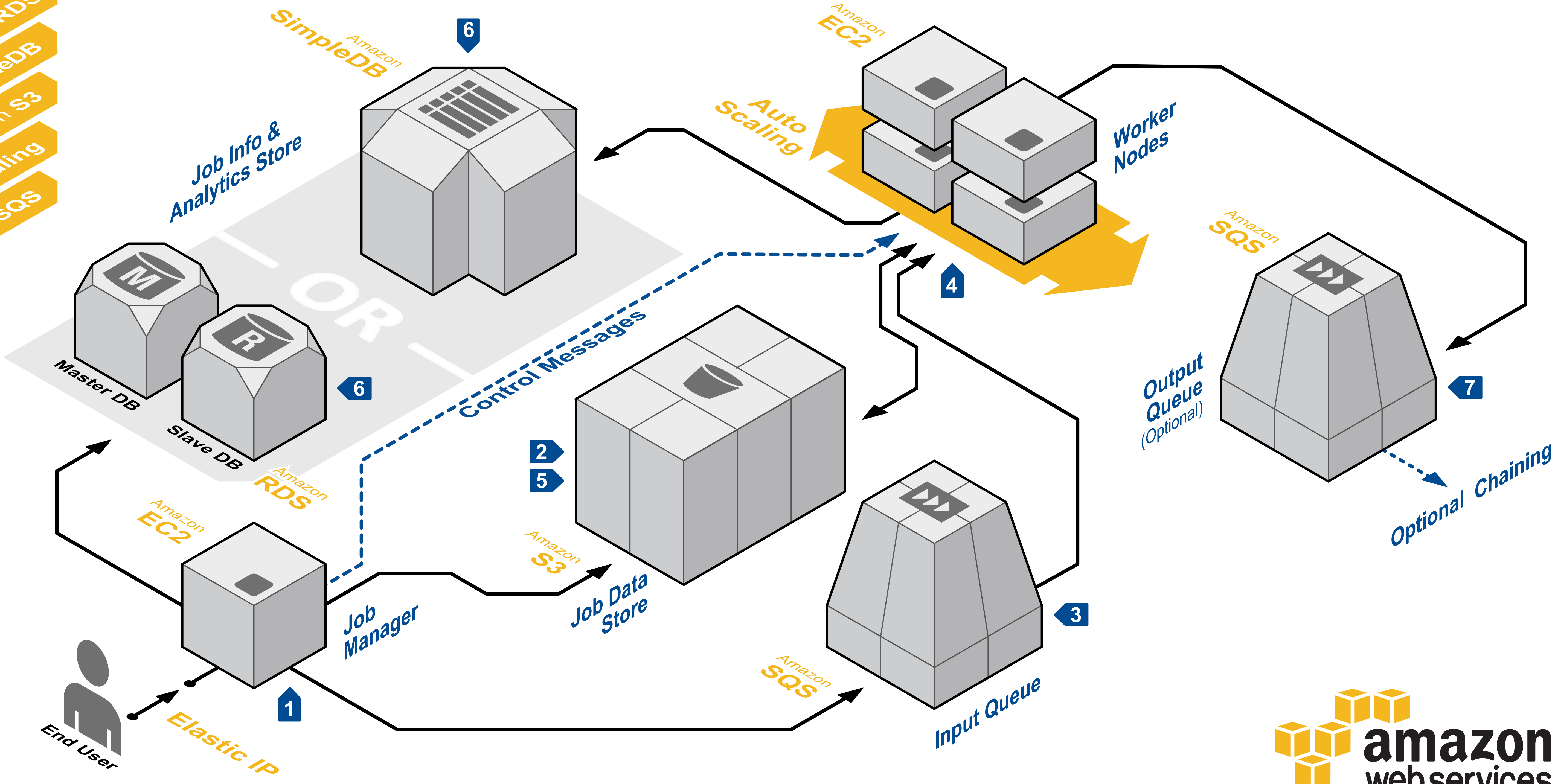
AWS Reference Architectures

- Amazon EC2
- Amazon RDS
- Amazon SimpleDB
- Amazon S3
- Auto Scaling
- Amazon SQS

BATCH PROCESSING

Batch processing on AWS allows for the on-demand provisioning of a multi-part job processing architecture that can be used for instantaneous or delayed deployment of a heterogeneous, scalable "grid" of worker nodes that can quickly crunch through large batch processing tasks in parallel. There are numerous batch oriented applications in place today that can leverage this style of on-demand processing, including claims processing, large scale transformation, media transcoding and multi-part data processing work.

Batch processing architectures are often synonymous with highly variable usage patterns that have significant usage peaks (e.g., month-end processing) followed by significant periods of underutilization.
There are numerous approaches to building a batch processing architecture. This document outlines a basic batch processing architecture that supports job scheduling, job status inspection, uploading raw data, outputting job results, grid management, and reporting job performance data.



System Overview

1 Users interact with the Job Manager application which is deployed on an **Amazon Elastic Compute Cloud (EC2)** instance. This component controls the process of accepting, scheduling, starting, managing, and completing batch jobs. It also provides access to the final results, job and worker statistics, and job progress information.

2 Raw job data is uploaded to **Amazon Simple Storage Service (S3)**, a highly-available and persistent data store.

3 Individual job tasks are inserted by the Job Manager in an **Amazon Simple Queue Service (SQS)** input queue on the user's behalf.

4 Worker nodes are **Amazon EC2** instances deployed on an **Auto Scaling** group. This group is a container that ensures health and scalability of worker nodes. Worker nodes pick up job parts from the input queue automatically and perform single tasks that are part of the list of batch processing steps.

5 Interim results from worker nodes are stored in **Amazon S3**.

6 Progress information and statistics are stored on the analytics store. This component can be either an **Amazon SimpleDB** domain or a relational database such as an **Amazon Relational Database Service (RDS)** instance.

7 Optionaly, completed tasks can be inserted in an **Amazon SQS** queue for chaining to a second processing stage.