

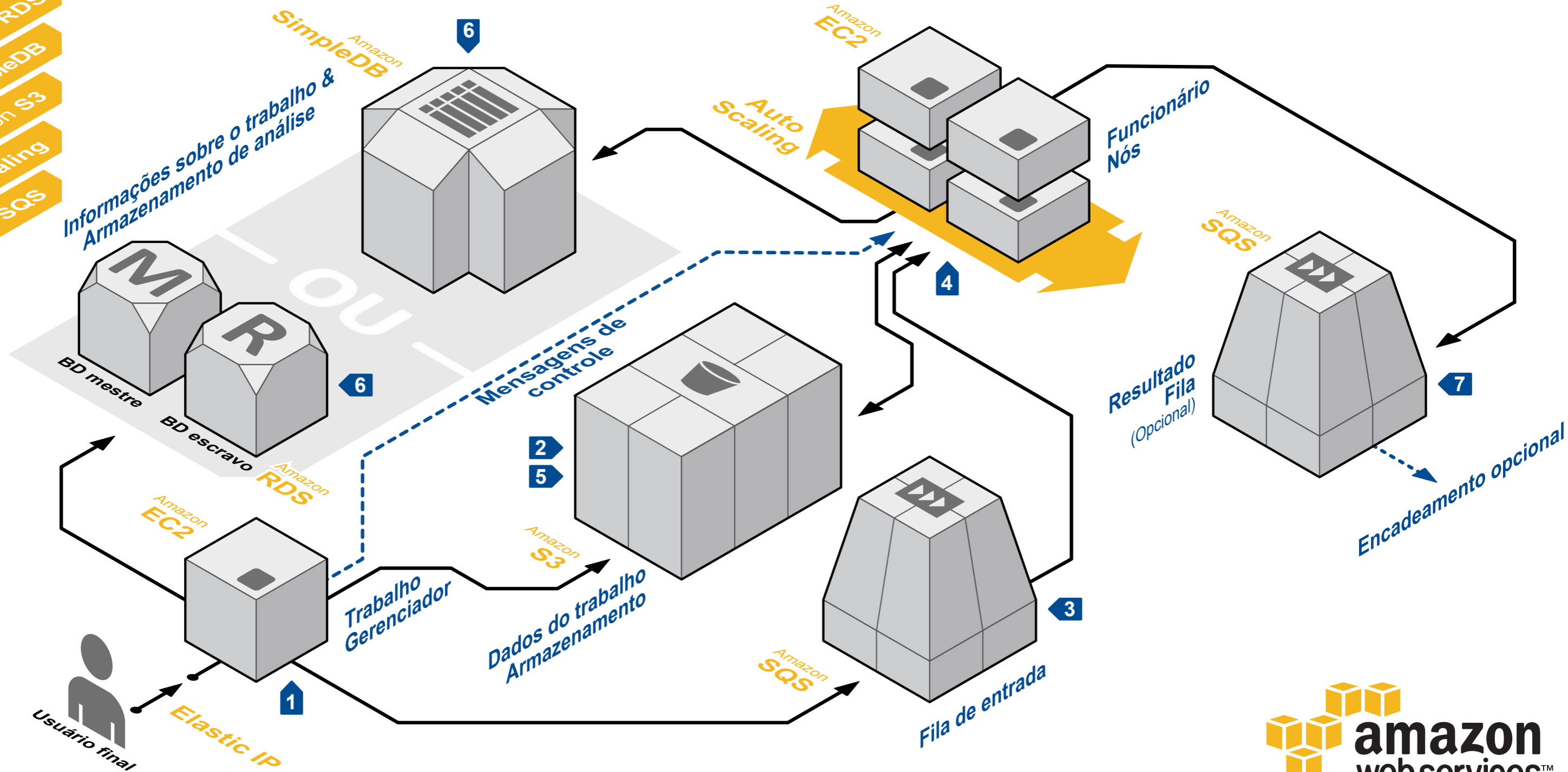
# LOTE EM PROCESSAMENTO

O processamento em lote na AWS permite o provisionamento sob demanda de uma arquitetura de processamento de trabalho de várias partes que pode ser usada para a implementação imediata ou postergada de uma "grade" heterogênea e escalonável de nós de funcionários que podem se agrupar rapidamente por meio de tarefas grandes de processamento em lote em paralelo. Há inúmeros aplicativos voltados para lote em uso atualmente que podem alavancar esse estilo de processamento sob demanda, incluindo o processamento de reclamações, transformação em larga escala, transcodificação de mídia e trabalho de processamento de dados de várias partes.

Com frequência as arquiteturas de processamento em lote são sinônimos de padrões de uso altamente variáveis que têm picos de utilização significativos (por ex., processamento no final do mês) seguidos por períodos significativos de sub-utilização. Há inúmeras abordagens para criar uma arquitetura de processamento em lotes. Esse documento descreve uma arquitetura básica de processamento em lote que oferece suporte à programação de trabalhos, inspeção do status de trabalho, carregamento de dados brutos, exibição de resultados do trabalho, gerenciamento da grade e relato de dados de desempenho do trabalho.

**AWS**  
**Referência**  
**Arquiteturas**

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



## Sistema Visão geral

**1** Os usuários interagem com o aplicativo Gerenciador de Trabalho que está implementado em uma instância do **Amazon Elastic Compute Cloud (EC2)**. Esse componente controla o processo de aceitação, programação, inicialização, gerenciamento e conclusão de trabalhos em lote. Ele também fornece acesso aos resultados finais, às estatísticas de trabalho e dos funcionários, e a informações sobre o andamento do trabalho.

**2** Os dados brutos de trabalho são carregados no **Amazon Simple Storage Service (S3)**, um armazenamento de dados altamente disponível e persistente.

**3** As tarefas de trabalho individuais são inseridas pelo Gerenciador de Trabalho em uma fila de entrada do **Amazon Simple Queue Service (SQS)** em nome do usuário.

**4** Os nós de funcionários são instâncias do **Amazon EC2** implementadas em um grupo de **Auto Scaling**. Esse grupo é um contêiner que assegura o status e a escalabilidade de nós de funcionários. Os nós de funcionários coletam partes do trabalho com base na fila de entrada automaticamente e desempenham tarefas únicas que fazem parte da lista de etapas de processamento em lote.

**5** Os resultados intermediários dos nós de funcionários são armazenados no **Amazon S3**.

**6** As informações sobre o andamento e as estatísticas são armazenadas no armazenamento de análise. Esse componente pode ser um domínio do **Amazon SimpleDB** ou um banco de dados relacional, como uma instância do **Amazon Relational Database Service (RDS)**.

**7** As tarefas concluídas também podem ser inseridas em uma fila do **Amazon SQS** para o encadeamento para uma segunda etapa de processamento.

