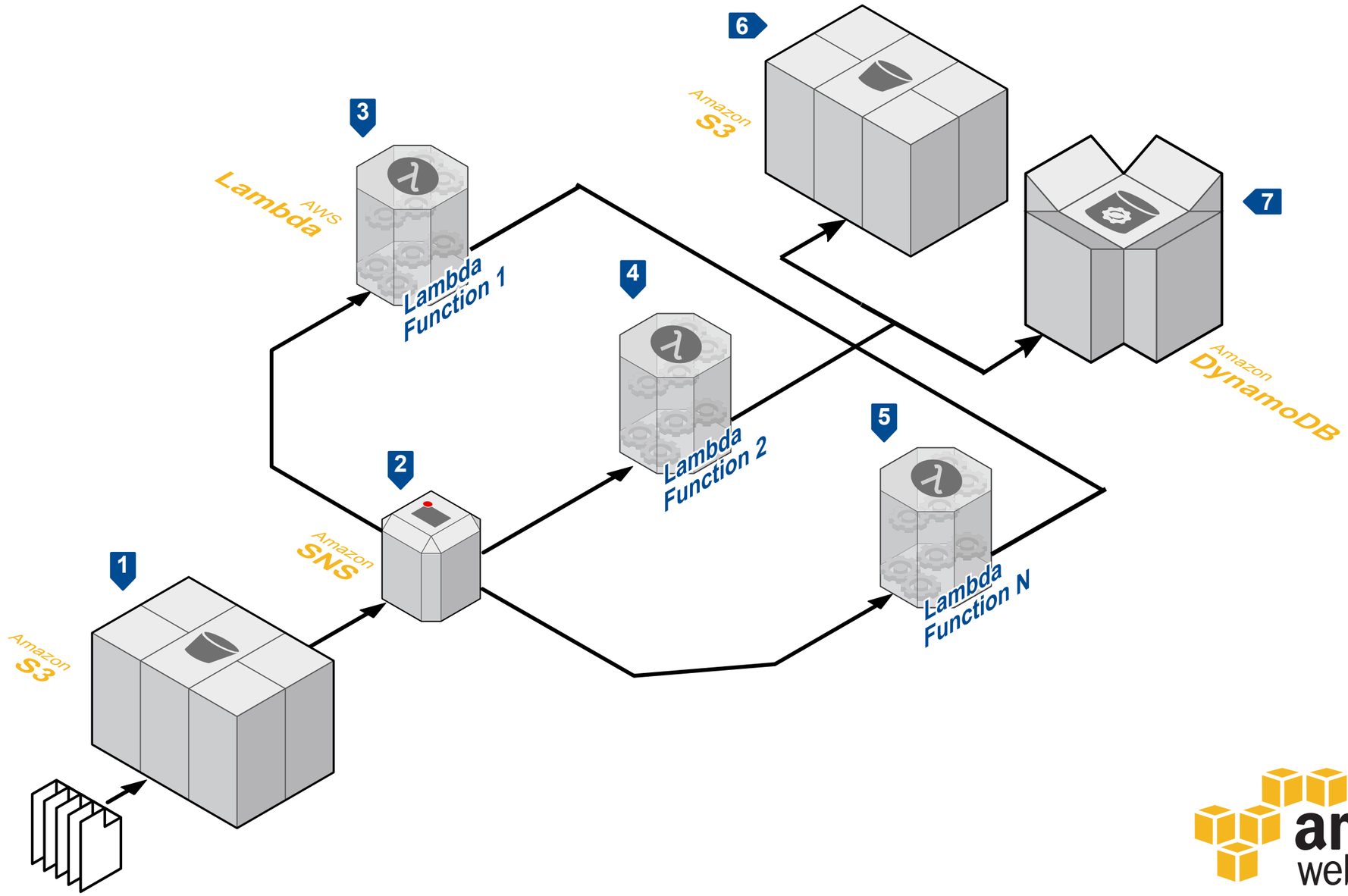
AWS Reference Architectures

- Amazon S3
- Amazon SNS
- AWS Lambda
- Amazon DynamoDB

AWS LAMBDA: REAL-TIME FILE PROCESSING

Building real-time file processing systems can be complex and expensive. Variable workloads can result in low utilization rates of compute infrastructure. AWS Lambda provides a serverless event-driven compute platform ideal for processing multiple workloads in parallel without the need to provision or manage servers. Lambda continuously scales applications by running code in parallel in response to triggers from AWS services such as Amazon S3.



amazon web services

System Overview

1. Objects are uploaded into **Amazon Simple Storage Service (S3)**, a highly available and persistent data store. The S3 bucket publishes an event notification to an **Amazon Simple Notification Service (Amazon SNS)** topic.
2. **Amazon SNS** is a fully managed push messaging system that can fan-out messages to multiple subscribers.
3. **AWS Lambda** is used to create a layer of processing for a wide variety of data sets and subsequently sends the results to a post-processing storage layer. For example, Lambda Function 1 can process objects put into the S3 bucket and create a data derivative.
4. Lambda Function 2 can, in parallel, process and create another data derivative from the same objects in S3.
5. N number of Lamba functions can be created to process data, all without the need to provision or manage servers.
6. A post-processing storage layer such as S3 can be added to store the results in a cost effective and durable fashion.
7. Alternatively, results can be sent to **Amazon DynamoDB** for further processing and/or querying with low-latency.