

Bootcamp IGTI: Engenheiro de Dados**Desafio**

Módulo 4	Infraestrutura de Dados e Arquitetura Escalável
-----------------	--

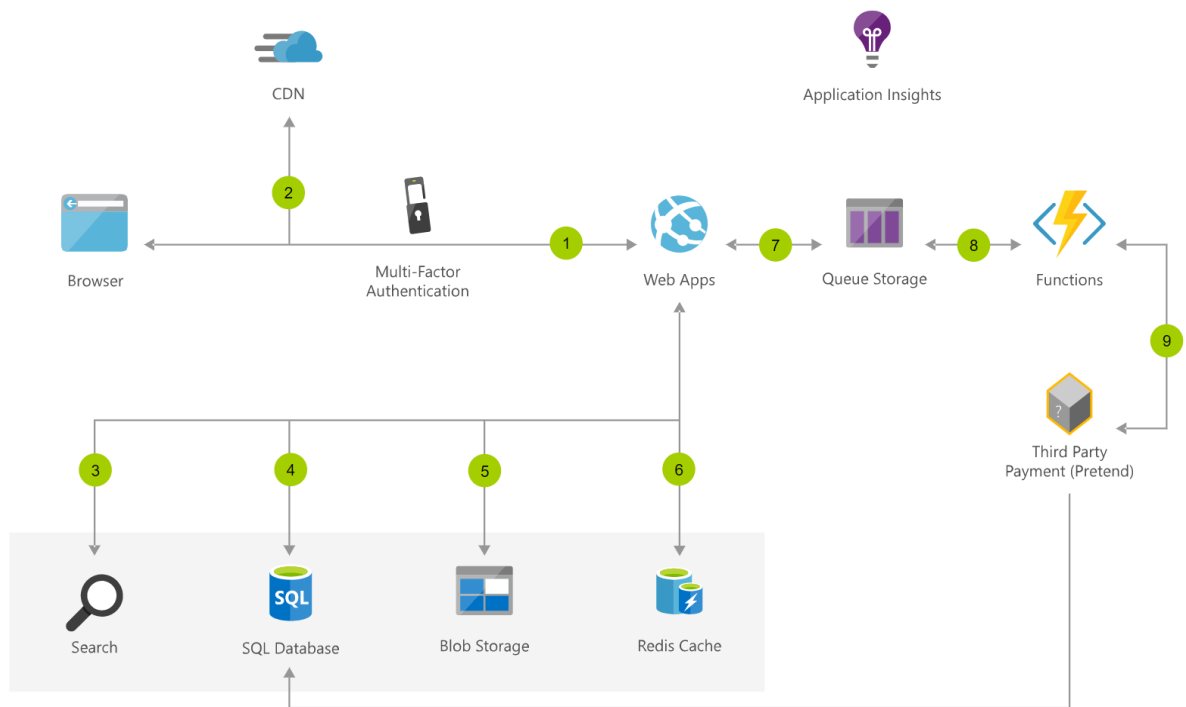
Objetivos

Exercitar os seguintes conceitos trabalhados no módulo:

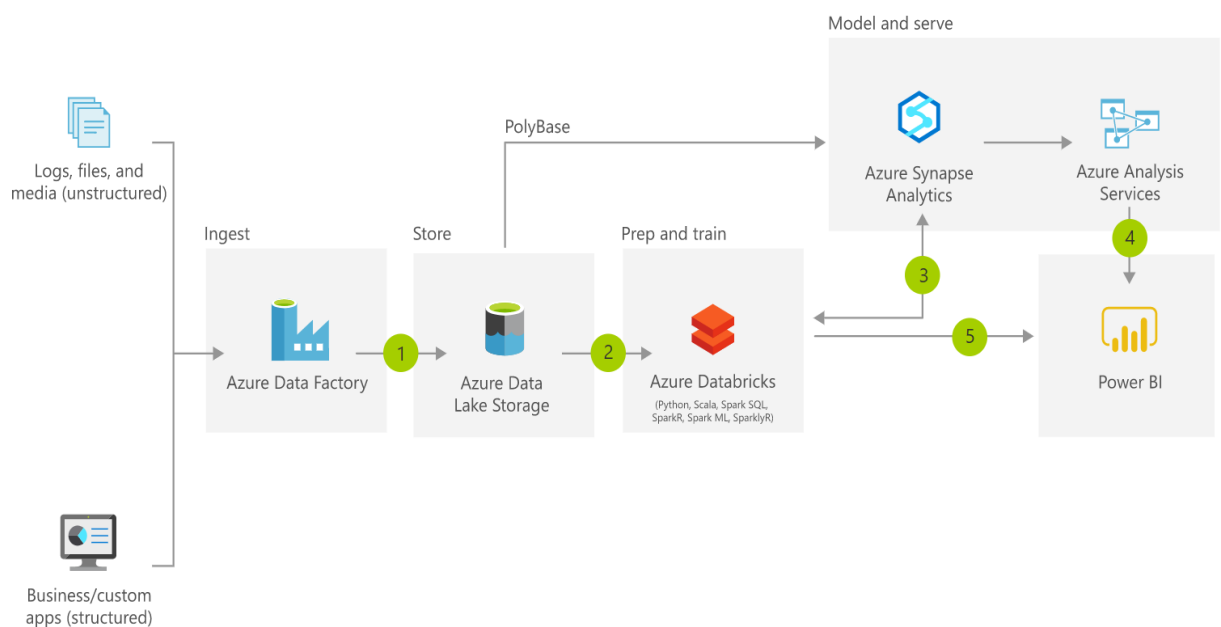
- ✓ Infraestrutura como serviço na nuvem para armazenamento de dados;
- ✓ Arquiteturas escaláveis de dados em infraestrutura como serviço na nuvem;
- ✓ Plataforma como serviço na nuvem para armazenamento de dados;
- ✓ Arquiteturas escaláveis de dados em plataformas como serviço na nuvem.

Enunciado

Você foi contratado para fazer a implantação e configuração, no Azure, das arquiteturas escaláveis para o armazenamento de dados de uma solução Web de comércio eletrônico. Essa solução possui a topologia mostrada a seguir.



Além disso, também foi feito o desenho da solução do data warehouse, que será implementado sobre a topologia da solução principal, mostrado abaixo:



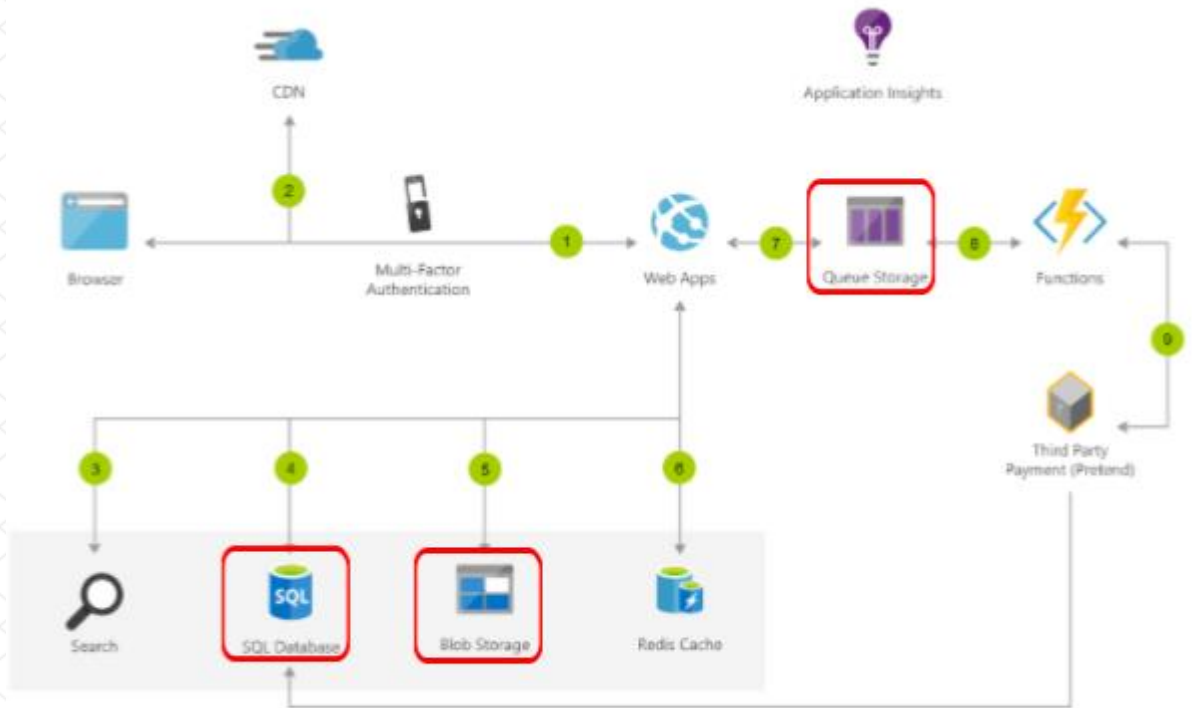
Como pré-requisitos para o provisionamento dos recursos, foram colocados os seguintes pontos:

- Uma conta de armazenamento para cada uma das duas soluções (Web/Data warehouse), usando discos magnéticos;
- Para essa camada, além da escalabilidade vertical, também deverá ser possível escalar horizontalmente o workload de leitura, bem como possuir um ambiente de DR com failover automático. Para os dados não estruturados, armazenados no respectivo recurso do Azure Storage, deverá ser permitida a leitura direta em outra região/zona para escalar o workload de leitura;
- Na solução de data warehouse, o Data Factory fará a ingestão de dados não estruturados (logs de navegação, textos e mídias das redes sociais da empresa), além de dados estruturados da solução Web;
- O Azure Synapse Analytics, da solução de data warehouse, deverá permitir armazenar e trabalhar com dados relacionais e dados não estruturados;
- O cluster do Azure DataBricks deverá ter a capacidade de executar workloads escritos na linguagem Scala, e ter no mínimo 3 worker nodes e no máximo 6 de propósito geral, cada um com 32 GB de RAM e 8 vCores.

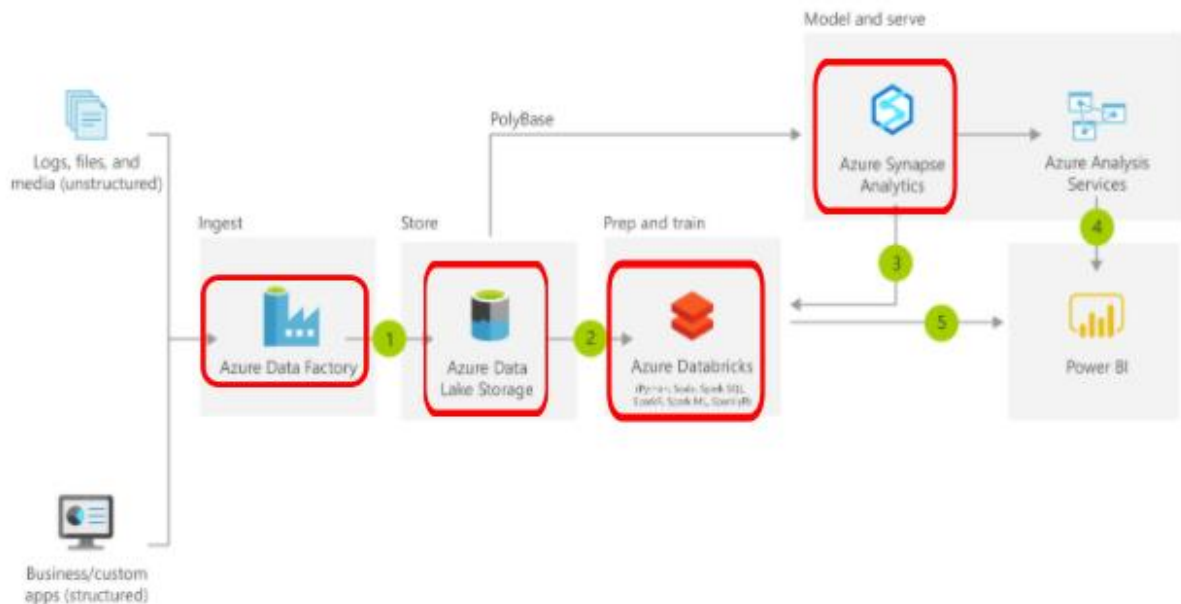
Atividades

Os alunos deverão desempenhar as seguintes atividades:

1. No Azure, usando a free tier, provisionar os recursos da solução Web a seguir, destacados em vermelho;



2. No Azure, usando a free tier, aprovisionar os recursos da solução de Data Warehouse a seguir, destacados em vermelho;



3. Responder às 15 questões objetivas do desafio, considerando os pré-requisitos relacionados anteriormente.