Suponha que você trabalha em uma organização que atua com parceiros, e precisa disponibilizar acesso a determinado dado para um dos seus parceiros, sendo parte desses dados, considerados sensíveis pelo jurídico. Como você resolveria esse problema de segurança de dados, nesse caso?

Para resolver esse problema de segurança de dados, eu seguiria os seguintes passos:

- Identificar os dados sensíveis. O primeiro passo é identificar quais dados são considerados sensíveis pelo jurídico. Isso pode ser feito por meio de uma análise dos dados da organização, levando em consideração fatores como o seu valor comercial, o impacto que sua divulgação teria na organização ou nos seus clientes, e as leis e regulamentos aplicáveis.
- Definir as regras de acesso. Uma vez identificados os dados sensíveis, é necessário definir as regras de acesso a esses dados. Essas regras devem ser claras e concisas, e devem especificar quem tem permissão para acessar os dados, quando e por que.
- Implementar as medidas de segurança. As regras de acesso devem ser implementadas por meio de medidas de segurança adequadas. Essas medidas podem incluir autenticação, autorização, criptografia, firewalls e monitoramento.
- No caso específico de dados sensíveis compartilhados com parceiros, é importante garantir que o parceiro também tenha implementado medidas de segurança adequadas para proteger esses dados. Para isso, é possível exigir que o parceiro assine um acordo de confidencialidade (NDA) ou um acordo de nível de serviço (SLA) que contenha cláusulas específicas sobre a segurança dos dados.

Para algumas medidas específicas, eu implementaria as seguintes soluções, para proteger os dados sensíveis:

- Autenticação forte. A autenticação forte deve ser usada para garantir que apenas usuários autorizados tenham acesso aos dados. Isso pode ser feito por meio de métodos como autenticação multifatorial ou autenticação baseada em identidade.
- Autorização granular. A autorização deve ser granular para garantir que os usuários autorizados tenham apenas o acesso necessário aos dados. Isso pode ser feito por meio de mecanismos de controle de acesso baseado em papéis ou funções.
- Criptografia. A criptografia deve ser usada para proteger os dados em trânsito e em repouso. Isso pode ser feito por meio de algoritmos de criptografia fortes e chaves de criptografia seguras.
- Firewalls. Os firewalls devem ser usados para proteger os dados de acessos não autorizados. Isso pode ser feito por meio de regras de firewall que bloqueiem o acesso a portas e serviços específicos.
- Monitoramento. Os dados devem ser monitorados para detectar atividades suspeitas. Isso pode ser feito por meio de ferramentas de monitoramento de segurança que rastreiam eventos e atividades nos sistemas e redes.

A implementação dessas medidas de segurança ajudará a proteger os dados sensíveis compartilhados com parceiros, e a mitigar o risco de vazamento ou comprometimento desses dados.

A solução dada anteriormente é ótima e você já se livrou do problema de dados sensíveis. Então agora seu desafio é permitir que os dados não sensíveis sejam acessados pelo parceiro sem comprometer a segurança dos dados. Suponha que o volume de dado que o parceiro precisa obter semanalmente, tem 10 GB, qual lógica, solução ou tecnologia você utilizaria para possibilitar essa integração?

Para permitir que os dados não sensíveis sejam acessados pelo parceiro sem comprometer a segurança dos dados, eu utilizaria uma solução de integração de dados baseada em APIs. Essa solução permitiria que o parceiro acessasse os dados por meio de uma interface segura e controlada.

A solução específica que eu utilizaria seria baseada nos seguintes princípios:

- Autenticação e autorização fortes. A autenticação e autorização devem ser fortes para garantir que apenas usuários autorizados tenham acesso aos dados. Isso pode ser feito por meio de métodos como autenticação multifatorial ou autenticação baseada em identidade.
- Criptografia. A criptografia deve ser usada para proteger os dados em trânsito e em repouso. Isso pode ser feito por meio de algoritmos de criptografia fortes e chaves de criptografia seguras.
- Controle de acesso baseado em papéis ou funções. O controle de acesso deve ser baseado em papéis ou funções para garantir que os usuários autorizados tenham apenas o acesso necessário aos dados.
- Monitoramento. Os dados devem ser monitorados para detectar atividades suspeitas. Isso pode ser feito por meio de ferramentas de monitoramento de segurança que rastreiam eventos e atividades nos sistemas e redes.

Baseadas nessas soluções, poderíamos implementar as seguintes soluções para mitigar os problemas:

- O parceiro criaria uma conta na API da organização.
- A organização concederia ao parceiro as permissões necessárias para acessar os dados.
- O parceiro usaria a API para solicitar acesso aos dados.
- A organização criptografaria os dados antes de enviá-los ao parceiro.
- O parceiro descriptografaria os dados após recebê-los.

Essa solução permitiria que o parceiro acessasse os dados sem comprometer a segurança dos dados. O volume de dados de 10 GB não seria um problema para essa solução, pois as APIs modernas são capazes de lidar com grandes volumes de dados.

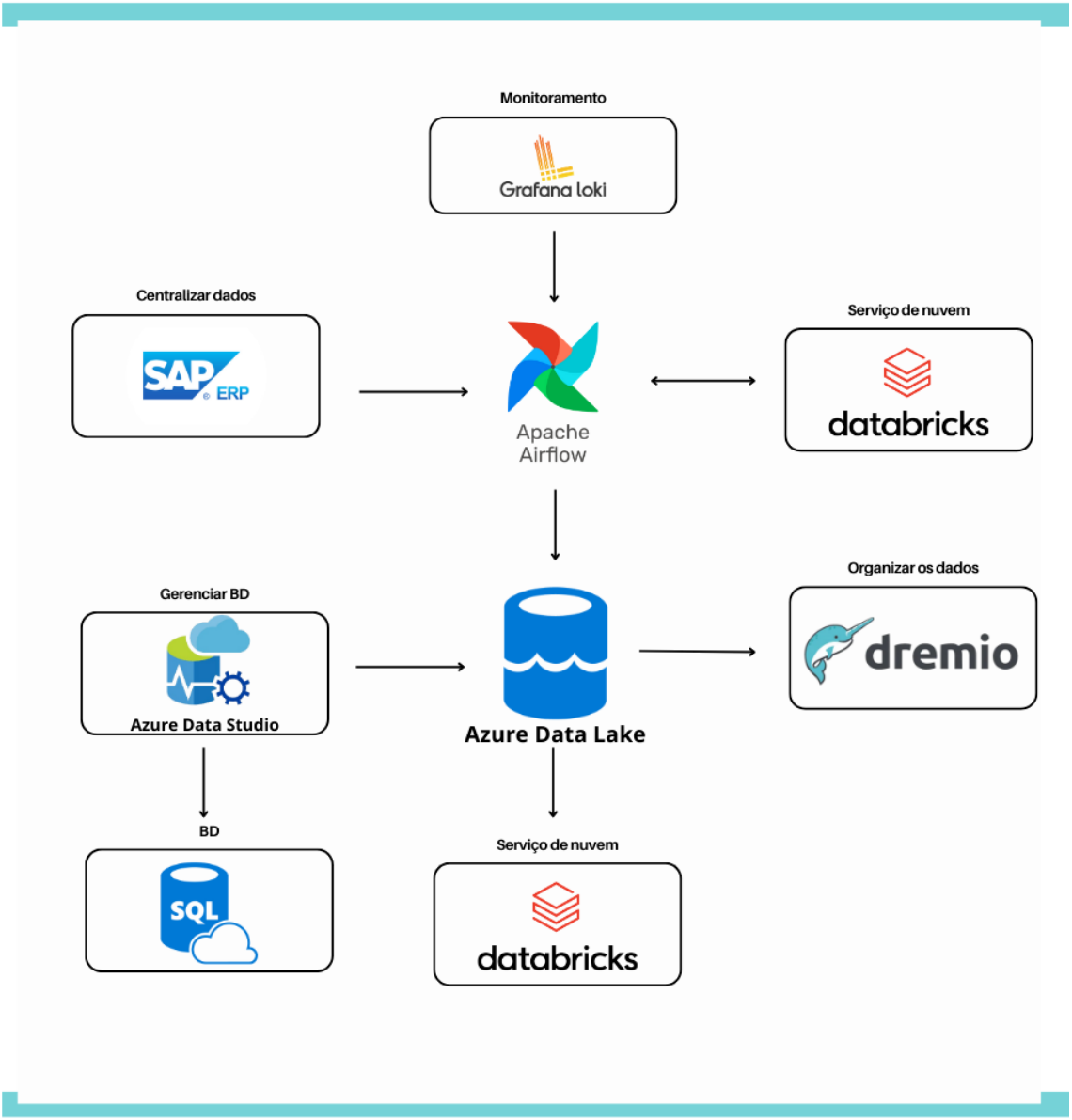
Aqui estão algumas tecnologias específicas que poderiam ser usadas para implementar as soluções:

- API Gateway. Um API Gateway pode ser usado para gerenciar o acesso às APIs.
- Criptografia. Algoritmos de criptografia fortes, como AES-256, podem ser usados para proteger os dados. Controle de acesso baseado em papéis ou funções. Ferramentas de gerenciamento de identidades e acesso (IAM) podem ser usadas para implementar o controle de acesso baseado em papéis ou funções.

- Monitoramento. Ferramentas de monitoramento de segurança podem ser usadas para monitorar os dados e detectar atividades suspeitas.

Ao combinar essas práticas, poderíamos criar uma solução robusta que permite ao parceiro acessar dados não sensíveis de maneira eficiente e segura, sem comprometer a integridade e a segurança dos dados.

Construa a arquitetura de Dados ideal para você, e explique por que você usaria os elementos que escolheu.



Dividi a arquitetura da seguinte forma, usei uma ferramenta para monitoramento (na arquitetura usei a grafana loki, mas existem muitas outras ferramentas), para armazenar e consultar os logs da aplicação. Como centralizador de dados, utilizei o SAP, que ajudaria a empresa nas integrações e na gestão completa dos diferentes setores da empresa. O Apache Airflow seria responsável por programar e monitorar as pipelines de trabalho. Já o Databricks ajudaria na melhoria da capacidade de coletar, gerenciar, processar e analisar dados da empresa. O Azura Data Lake seria utilizado para organizar todos os objetos e

arquivos em uma hierarquia de diretórios e subdiretórios aninhados. O Dremio ficaria responsável por organizar os dados, auxiliando a ferramenta que faria as consultas no banco de dados, que nesta arquitetura usaria um banco de dados SQL.