Hardening for Blue Teams

Edição 04 - Janeiro/2023

Security by Design

Azure Data Factory

Criado por: Felipe Silvany , Yves Peixoto e Antonio Luciano

SUMÁRIO

CONFIDENCIALIDADE *3*
SOBRE O AUTOR *4*
SOBRE ESTE DOCUMENTO *5*
AGRADECIMENTOS *6*
1 - INTRODUÇÃO *7*
2 – CONCEITOS BÁSICOS *8*
3 – HARDENING *12*
4 – LEITURA ADICIONAL *14*

CONFIDENCIALIDADE

Este documento contém informações confidenciais ou privilegiadas, sendo seu sigilo protegido por lei, não sendo autorizado o uso, cópia ou divulgação das informações ou tomar qualquer ação baseada nessas informações.
Desta forma, os destinatários destes documentos comprometem-se em:

1. Não utilizar as informações confidenciais a que tiver acesso, para gerar benefício próprio exclusivo e/ou unilateral, presente ou futuro, ou para o uso de terceiros.
2. Não efetuar nenhuma gravação ou cópia da documentação confidencial a que tiver acesso.
3. Não apropriar-se de material confidencial e/ou sigiloso da tecnologia que venha a ser disponível.
4. Não repassar o conhecimento das informações confidenciais, responsabilizando-se por todas as pessoas que vierem a ter acesso às informações, por seu intermédio, e obrigando-se, assim, a ressarcir a ocorrência de qualquer dano e/ou prejuízo oriundo de uma eventual quebra de sigilo das informações fornecidas.

SOBRE O AUTOR

Felipe Silvany é coordenador de Cybersecurity, com 20 anos de experiência na área de tecnologia, com sólidos conhecimentos em resposta à incidentes de segurança, SOC, hardering, análise de vulnerabilidades, pentest e grandes projetos de cybersecurity em ambientes onpremises, Cloud e mobile (AppSec).
Sólida experiência na implementação de soluções técnicas e estruturação de equipes de segurança de alta performance.
Responsável pela implementação de serviços, ferramentas e processos como: SOC/CSIRT 24x7, playbooks, Plano de Resposta à Incidentes de Segurança (RI), SIEM com base no MITRE ATT&CK, hardening e baselines de segurança, exercícios de RedTeam como Pentest, análise de vulnerabilidades DAST e SAST na infraestrutura onpremises/cloud (OpSec), mobiles Android e IOS (AppSec) e definição de arquitetura segura com conceitos de Zero Trust. Implementação de ferramentas de BlueTeam como: NGAv/EDR, NGFw, MTLS, NAC, DLP, IDS, IPS, WAF, AppControl, Anti-DDoS e outras.
Graduado em "Sistema de Informação", Pós Graduado em "Projetos e Gerência de Redes" e com MBA em "Gestão Empresarial Estratégica". Membro da ISO/ABNT/CE-021 000.027 (Comissão Especial de Estudo de Segurança da Informação, Segurança Cibernética e Proteção da Privacidade de dados), membro da ANPPD (Associação Nacional dos Profissionais de Privacidade de Dados) e certificado nos mais conceituados frameworks internacionais, como: ISO27001, ISO20000, ITIL, COBIT, Scrum, Kanban e outros. Classificado entre os TOP 04 Ethical Hackers na maior plataforma de BugBounty do Brasil e TOP 05 na maior plataforma de BugBounty da Europa.
Linkedin: https://www.linkedin.com/in/felipe-silvany-69241365/
Github: https://github.com/FelipeSilvany

Yves Peixoto é Especialista em Cybersecurity, com mais de 12 anos de experiência na área de tecnologia. Bacharel em análise de sistemas com pós-graduação em segurança da informação pela UFRJ e algumas certificações de mercado, como por exemplo: CEH, CNDA, ISO, LPI, Azure, entre outras.
Dentro da área de S.I. possui sólidos conhecimentos tanto em áreas ofensivas (Pentest, Appsec, DAST, SAST, IAST, MAST, análise de vulnerabilidades quanto em áreas defensivas (SOC, CSIRT, hardening, firewalls, WAF, IDS, IPS, EDR, infraestrutura onpremises e cloud). Possui histórico de atuação em empresas de diversos seguimentos, privados e públicos.
Linkedin: https://www.linkedin.com/in/yvespeixoto/

SOBRE ESTE DOCUMENTO

Este não é um e-book comum.
Você está prestes a obter um conhecimento privilegiado em segurança da informação e hardening.
Esta obra utiliza uma metodologia focada em *Security by Design* e implementação de boas práticas de segurança desde a concepção de projetos corporativos até sua passagem para produção.
A segurança deve ser sempre preventiva e pró ativa, nunca reativa.
Nesta quarta edição veremos boas práticas na criação e configuração do Azure Data Factory.

Edição 01 – Criação do Documento
Janeiro/2023 – Antônio Luciano, Felipe Silvany e Yves Peixoto

AGRADECIMENTOS

“Ao meu filho, Thor, que me dá toda força para enfrentar todas as dificuldades e tribulações da vida com força e ânimo. À minha esposa, Fabiana, por sempre me apoiar em todos os desafios profissionais e perder várias noites de sono comigo. À minha mãe e meu pai, Maria Aparecida e Adilton, por terem me dado a maior herança que um homem por ter, o estudo. À Deus, que por intermédio do meu Senhor Jesus Cristo conduz a minha vida, me guia e me guarda.” **– Felipe Silvany**

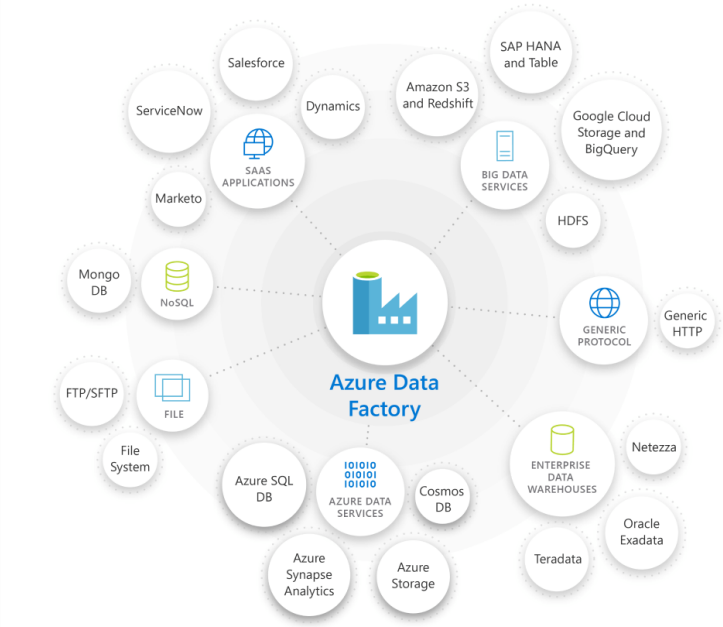
“Primeiramente à Deus e a minha família por todo apoio em minha trajetória, principalmente nos momentos mais difíceis.” **– Yves Peixoto**

1 - INTRODUÇÃO

No mundo de Big Data, os dados brutos e não organizados são, muitas vezes, armazenados em sistemas relacionais, não relacionais e outros sistemas de armazenamento. No entanto, os dados brutos em si não possuem o contexto ou significado apropriados para fornecer uma visão adequada para os analistas, cientistas de dados ou responsáveis por decisões de negócio.

O Big Data requer um serviço que possa orquestrar e operacionalizar processos para refinar esses enormes repositórios de dados brutos em insights de negócio acionáveis. O Azure Data Factory é um serviço de nuvem gerenciado que foi criado para esses projetos híbridos complexos para extrair, transformar e carregar (ETL) e de integração de dados.

O Data Factory foi certificado pela HIPAA e a HITECH, a ISO/IEC 27001, a ISO/IEC 27018 e a CSA STAR.



Este ebook descreve a infraestrutura de segurança básica que os serviços de movimentação de dados no Azure Data Factory usam para ajudar a proteger seus dados. Os recursos de gerenciamento do Data Factory são criados na infraestrutura de segurança do Azure e usam todas as medidas de segurança possíveis oferecidas pelo Azure.

2 – CONCEITOS BÁSICOS

CRIPTOGRAFIA DE CREDENCIAIS

Você deve criptografar e armazenar credenciais para qualquer um de seus armazenamentos de dados locais (serviços vinculados com informações confidenciais) em uma máquina com tempo de execução de integração self-hosted.
Para criptografar os dados confidenciais da carga JSON em um tempo de execução de integração auto-hospedado, execute New-AzDataFactoryV2LinkedServiceEncryptedCredential e transmita a carga JSON. Esse cmdlet garante que as credenciais sejam criptografadas usando DPAPI e armazenadas no nó de tempo de execução de integração auto-hospedado localmente. Ele pode ser executado em qualquer máquina, desde que a opção Remote Access esteja habilitada no tempo de execução de integração auto-hospedado de destino e o PowerShell 7.0 ou superior seja usado para executá-lo.

PROTEGENDO CREDENCIAIS DE ARMAZENAMENTO

Você pode armazenar a credencial do armazenamento de dados no Azure Key Vault, com rotação e periodicidade de expiração. O Data Factory recupera a credencial durante a execução de uma atividade.

CRIPTOGRAFIA DE DADOS EM TRÂNSITO

Se o armazenamento de dados na nuvem oferecer suporte a HTTPS ou TLS, todas as transferências de dados entre os serviços de movimentação de dados no Data Factory e um armazenamento de dados na nuvem deverão ser realizados por meio de um canal seguro HTTPS ou TLS.

> ⓘ Observação
>
> Todas as conexões com o Banco de Dados SQL do Azure e o Azure Synapse Analytics exigem criptografia (SSL/TLS) enquanto os dados estão em trânsito de e para o banco de dados. Ao criar um pipeline usando JSON, adicione a propriedade de criptografia e defina-a como **true** na cadeia de conexão. Para o Armazenamento do Azure, você pode usar **HTTPS** na cadeia de conexão.

CRIPTOGRAFIA DE DADOS EM REPOUSO

Alguns armazenamentos de dados oferecem suporte à criptografia de dados em repouso. Recomendamos que você habilite o mecanismo de criptografia de dados para esses armazenamentos de dados.
A Criptografia de Dados Transparente (TDE) no Azure Synapse Analytics ajuda a proteger contra a ameaça de atividade maliciosa executando criptografia e descriptografia em tempo real de seus dados em repouso. Esse comportamento é transparente para o cliente.

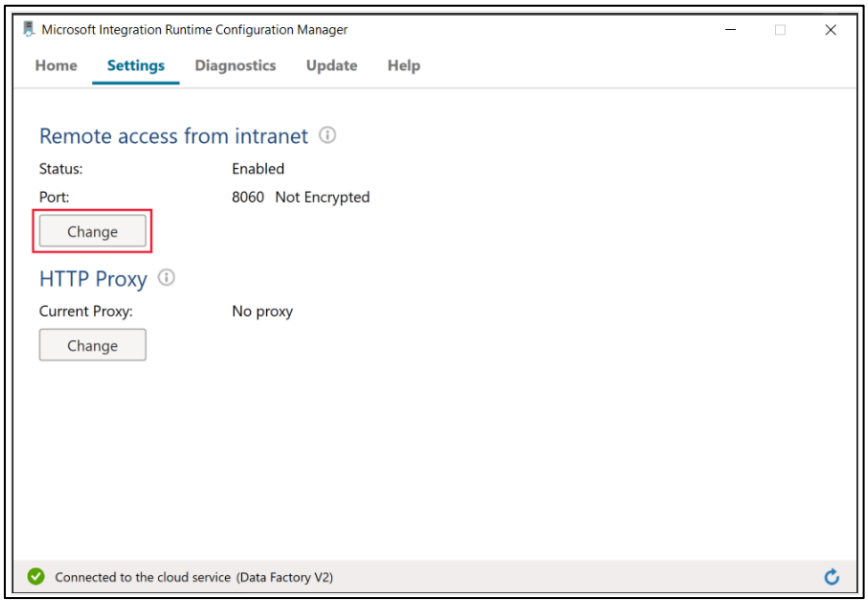
Para Banco de Dados SQL do Azure também dá suporte à criptografia de dados transparente (TDE), que ajuda a proteger contra a ameaça de atividade maliciosa executando criptografia e descriptografia em tempo real dos dados, sem exigir alterações no aplicativo. Esse comportamento é transparente para o cliente.

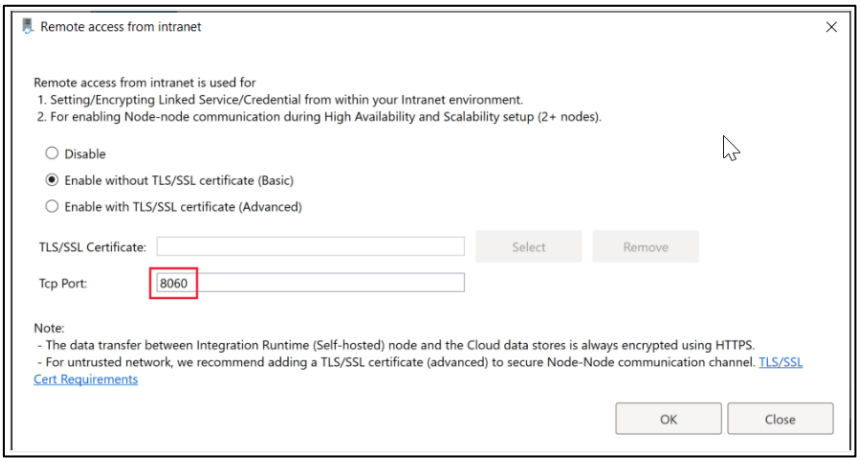
ARMAZENAMENTO DE BLOBS E TABELAS DO AZURE

O armazenamento de BLOBs do Azure e o armazenamento de tabelas do Azure dão suporte à criptografia do serviço de armazenamento (Storage Service Encryption - SSE), que criptografa automaticamente seus dados antes de persistir no armazenamento e descriptografia antes da recuperação.

PORTAS USADAS PARA CRIPTOGRAFAR O SERVIÇO VINCULADO NO TEMPO DE EXECUÇÃO DE INTEGRAÇÃO SELF-HOSTED

Por padrão, quando o acesso remoto da intranet está habilitado, o PowerShell usa a porta 8060 na máquina com tempo de execução de integração auto-hospedado para comunicação segura. Altere esta porta no Integration Runtime Configuration Manager na guia Configurações:



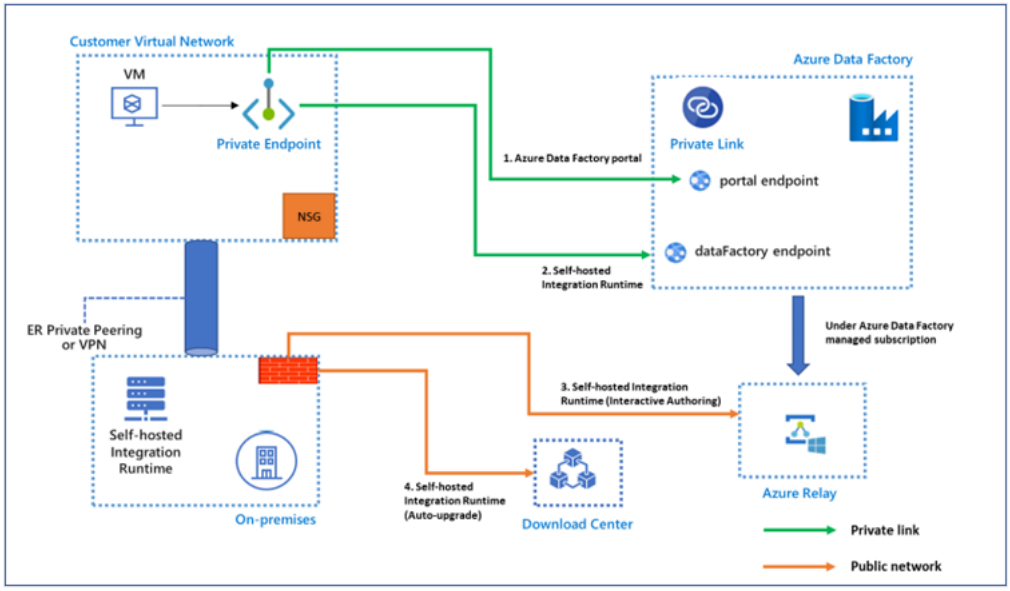


MANAGED VIRTUAL NETWORKING

Utilize endpoints privados para se conectar com segurança aos armazenamentos de dados suportados.

A criação de um integration runtime em uma rede virtual gerenciada garante que o processo de integração de dados seja isolado e seguro.

Com uma rede virtual gerenciada, você pode descarregar a carga de gerenciamento da rede virtual para o Data Factory. Você não precisa criar uma sub-rede para um runtime de integração que poderia eventualmente usar muitos IPs privados de sua rede virtual e exigiria planejamento prévio da infraestrutura de rede.

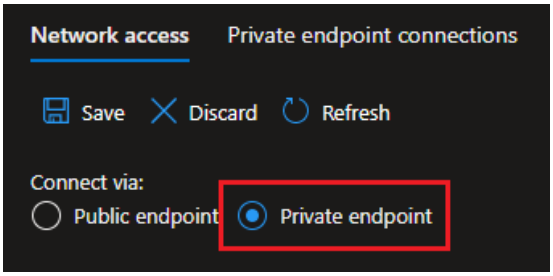


MANAGED PRIVATE ENDPOINTS

Pontos de extremidade privados gerenciados são pontos de extremidade privados criados na rede virtual gerenciada do Data Factory que estabelece um link privado para recursos do Azure. O Data Factory gerencia esses endpoints privados em seu nome.

Quando você usa um link privado, o tráfego entre seus armazenamentos de dados e a rede virtual gerenciada passa inteiramente pela rede de backbone da Microsoft. O link privado protege contra riscos de exfiltração de dados. Você estabelece um link privado para um recurso criando um endpoint privado.

Um ponto de extremidade privado usa um endereço IP privado na rede virtual gerenciada para trazer efetivamente o serviço para ele. Os pontos de extremidade privados são mapeados para um recurso específico no Azure e não para todo o serviço.



ZERO TRUST

O principal conceito trazido pelo Zero Trust é “nunca confie, sempre verifique” ou para autores mais conservadores “confiar, mas sempre verificar”, o que significa que dispositivos e serviços não devem ser confiáveis (confiados) por padrão. Redes, como WAN, LAN corporativa, VPNs são hostis e inseguras por padrão, mesmo que estejam dentro do perímetro de segurança, por isso, todo acesso deve ser validado, mesmo que já tenha sido verificado anteriormente. O Zero Trust aborda princípios chaves como: Autenticação mútua (MFA, 2FA), autenticação de máquina, autenticação de usuário, monitoramento contínuo, registro de todos os logs de acessos, entre outros. Permitir um acesso anônimo no ambiente ou disponibilizado para qualquer origem (ALL) é contra todos os conceitos trazidos pelo Zero Trust.

PRINCÍPIO DO MENOR PRIVILÉGIO

O princípio do menor privilégio é aquele que preza por delegar somente os privilégios necessários para que um determinado elemento (sistema ou pessoa) possa realizar sua função.

MONITORAMENTO

O Azure Monitor fornece logs e métricas de infraestrutura de nível básico para a maioria dos serviços do Azure. Os logs de diagnóstico do Azure são emitidos por um recurso e fornecem dados avançados e frequentes sobre a operação do recurso. O ADF (Azure Data Factory) pode gravar logs de diagnóstico no Azure Monitor.

O Data Factory armazena dados de execução de pipeline por apenas 45 dias. Use o Azure Monitor se desejar manter esses dados por mais tempo. Com o Monitor, você pode rotear os logs de diagnóstico para análise em vários destinos diferentes:

- **Conta de armazenamento**: salve os logs de diagnóstico em uma conta de armazenamento para auditoria ou inspeção manual. Você pode usar as configurações de diagnóstico para especificar o tempo de retenção em dias.
- **Hub de eventos**: transmita os logs para os Hubs de Eventos do Azure. Os logs tornam-se entrada para uma solução de serviço de parceiro/análise personalizada, como o Power BI.
- **Log Analytics**: analise os logs com o Log Analytics. A integração do Data Factory ao Azure Monitor é útil nos seguintes cenários:
    - Você quer escrever consultas complexas em um conjunto avançado de métricas publicadas pelo Data Factory no Monitor. Você pode criar alertas personalizados nessas consultas por meio do Monitor.
    - Você quer monitorar os data factories. Você pode rotear dados de vários alocadores de dados para um só workspace de monitoramento.

BACKUP

Realizar cópias de segurança de ambientes, aplicações ou dados em um determinado momento. Fazer cópias dos softwares, arquivos e outros dados em diferentes dispositivos de armazenamento para a recuperação do sistema em caso de falhas.

No Uso de IRs self-hosted em VMs para o ambiente de Data Factory, habilite o backup do Azure e configure a VM, o período de retenção e a frequência necessária para backups automáticos. Para fazer backup de todo o código no Azure Data Factory, use a funcionalidade de controle do código-fonte no Data Factory.

3 – HARDENING

Garanta que as contramedidas abaixo estejam implementadas em seu Ambiente. Responda "*sim*" caso as contramedidas de segurança estejam implementadas e forneça maiores detalhes técnicos junto à resposta. Responda "*não*" caso as contramedidas de segurança não estejam implementadas no ambiente e considere como um gap de segurança.

1. Certifique-se que os usuários que possuem acesso de gerência ao Data Factory tenham a autenticação MFA/2FA habilitada;
2. Habilite a Microsoft Azure Defender para detecção de ameaças dos recursos do Azure (Azure Virtual Machines, Azure SQL Database, Azure Blob storage e Azure Table storage) que estejam conectados ao Data Factory.
3. Habilite soluções de gerenciamento de vulnerabilidades (Microsoft/Qualys) no caso de utilização de Virtual Machines;
4. Padronize o Azure AD como padrão de identidade e autenticação e elimine credenciais locais;
5. Criptografe e Armazene as credenciais para qualquer um de seus armazenamentos de dados locais;
6. Utilize o Azure Key Vault para armazenamento da credencial de armazenamento de dados, quando o armazenamento não tiver suporte pelo Data Factory ao Managed Service Identity (MSI);
7. Utilize rotação e expiração periódica das credenciais no Azure Key Vault;
8. Utilize criptografia para dados em trânsito HTTPS/TLS 1.2;
9. Utilize criptografia de dados em repouso. A Criptografia de Dados Transparente (TDE) no Azure Synapse Analytics e Azure SQL Database ajudam a proteger contra a ameaça de atividade maliciosa executando criptografia e descriptografia em tempo real de seus dados em repouso;
10. Utilize o Azure SSE (Storage Service Encryption) para criptografar automaticamente o armazenamento de Blobs do Azure e de tabelas;
11. Altere a porta padrão (8060) de acesso da Intranet no Integration Runtime Configuration Manager por outra diferente do padrão conhecido;
12. Crie uma rede virtual gerenciada (managed virtual networking) para descarregar a carga de gerenciamento da rede virtual para o Data Factory, garantindo o processo de integração de dados seja isolado e seguro;
13. Utilize um ponto de extremidade privado (Managed private endpoints), garantindo que o tráfego entre seus armazenamentos de dados e o Data Factory passem por uma VNET exclusiva.
14. Configure retenção de armazenamento de log para ao menos 90 dias;
15. Garanta backups e testes de restore realizados regularmente;
16. Garantir que o Data Factory seja criado em regiões onde há replicação de dados em regiões pares, a fim de proteger contra perda de dados, no caso de desastres. As regiões "Brazil South" e "Southeast Asia" não possuem esta proteção.

17. Garanta junto à equipe de Cybersecurity que o servidor será submetido à testes de segurança, invasão (Pentest) ou vulnerabilidade.
18. Garanta que o *Princípio do Menor Privilégio* seja respeitado e credenciais de serviço ou sistema tenham apenas o mínimo de privilégios necessários para sua autenticação.
19. Garanta que o conceito de Zero Trust seja minimamente respeitando, realizando segmentações de rede restritas, autenticando e armazenando logs de todos os acessos e evitando a publicação de interfaces públicas para comunicações privadas.

4 – LEITURA ADICIONAL

- https://docs.microsoft.com/en-us/azure/data-factory/encrypt-credentials-self-hosted-integration-runtime
- https://docs.microsoft.com/en-us/azure/data-factory/data-movement-security-considerations
- https://docs.microsoft.com/en-us/azure/storage/common/storage-service-encryption
- https://learn.microsoft.com/en-us/azure/data-factory/data-factory-private-link
- https://docs.microsoft.com/en-us/security/benchmark/azure/baselines/data-factory-security-baseline
- https://learn.microsoft.com/en-us/azure/data-factory/
- https://learn.microsoft.com/en-us/azure/data-factory/monitor-using-azure-monitor
- https://learn.microsoft.com/en-us/azure/data-factory/concepts-data-redundancy