Three Things to Consider:
Microsoft Azure Data Integration Tools
Three Things to Consider:
Microsoft Azure Data Integration Tools

"88% of organizations using cloud today, have a cloud-first strategy."

"Hype Cycle for Cloud Computing," Gartner

According to Forrester, as of the beginning of 2019, close to 60% of North American enterprises rely on public cloud platforms, which is five times the percentage five years prior. On top of this, organizations use their internal data centers as private clouds.

At 76% year-over-year growth as of January 2019, Microsoft’s Azure is not only one of the largest but also one of the fastest growing public cloud platforms. However, as with any technology platform, it isn’t a given that organizations can simply sign up for an account and reap immediate rewards without a careful approach.

Cloud data integration can apply to a variety of use cases: whether it be from various sources into an ADLS data lake, migrating on-premises to the Azure cloud, running real-time analytics in the cloud, or integrating into cloud systems. Organizations interested in leveraging the power of Azure should consider the following three tips when architecting data integration solutions:

1. Secure Your Data

For years security, or rather perceived lack thereof, was one of the main concerns holding back cloud adoption. However many data breaches are caused by delayed patching of known problems or poor password management, and given that data breaches can break the cloud providers’ business, there has been a strong focus on cloud data security features.

As you adopt the Azure cloud data security is your responsibility. Following are some tips to consider:

- Use Azure’s virtual network and avoid the use of public IP addresses if possible.
- Encrypt data, especially on the network between public IP addresses, using an AES256 or stronger encryption method.
- Limit access to the on-premises data stores by keeping the firewall closed if at all possible. Take advantage of a proxy in a DMZ (de-militarized zone) if a port must be open from the cloud into the on-premises system.
- Use strong authentication when connecting from the outside in, e.g. by deploying two-factor authentication. Once in Azure use the Azure Active Directory.
- Take advantage of the Azure-native Storage Service Encryption using Microsoft-managed keys, or by storing your own in the Azure Key Vault.
2 You Want to Optimize Your Network Utilization Because:

The best way to get the most use out of network resources is to avoid unnecessary use, and optimize the unavoidable use. Look for data integration solutions that:

- Compress data before sending it across the wire. Balance the compression ratio with the cost of compression/decompression to satisfy your business needs for latency and other resources’ utilization.

- Use a Change Data Capture (CDC) mechanism, ideally by identifying relevant changes as close to the data source as possible.

- Use optimization techniques to leverage available bandwidth even if latency is relatively high.

3 Validate Your Data

Irrespective of whether you (gradually) migrate on-premises databases into Azure, you implement a data warehouse using, for example, Azure SQL Data Warehouse, or you build a data lake in the Azure Data Lake Store (ADLS), data accuracy should be top of mind. Minimize risk during data migrations by validating data for accuracy before switching over to the database in Azure. Build end users’ trust in the data warehouse or data lake by regularly validating the data and share accuracy reports.

Regardless of the use case, these three tips will ensure your cloud initiative is a success:

- SECURE YOUR DATA
- OPTIMIZE NETWORK UTILIZATION
- VALIDATE YOUR DATA

Would you like to discuss your cloud data integration needs?

Contact us: info@hvr-software.com