



Tips for Getting Started with the Azure Data Platform



Practical Tips and Lessons Learned

April 17, 2018

Melissa Coates

Solution Architect, BlueGranite

Microsoft Data Platform MVP

Blog: www.sqlchick.com

Twitter: @sqlchick

Agenda

Goals for this session:

1. Key concepts for using Azure as a data platform
2. Share suggestions and lessons learned through a series of tips and demos

Lots of important things are out of scope though!

- Database migration techniques
- Security
- Azure Stack & private cloud deployments
- Networking & connectivity
- Troubleshooting, performance tuning & growth
- Installation and configuration details
- Monitoring & alerting



Azure services & features change very frequently, so be sure to verify details

Tip #1

Know the What & Why
for Moving to the Cloud

Azure

A public cloud computing platform and infrastructure for building, deploying, and managing MSFT-specific and third party software and services through a global network of Microsoft-managed datacenters

Source: <http://azureplatform.azurewebsites.net/>

Compute		Web & Mobile		Internet of Things & Enterprise Integration	
Virtual Machines	Virtual Machine Scale Sets	Web Apps	Mobile Apps	Azure IoT Hub	Event Hubs
Azure Container Service	Azure Container Registry	Logic Apps	API Apps	Stream Analytics	Notification Hubs
Functions	Batch	Content Delivery Network	Media Services	BizTalk Services	Service Bus
Service Fabric	Cloud Services	Search		Data Catalog	
Networking		Databases		Security + Identity	
Virtual Network	Load Balancer	SQL Database	SQL Data Warehouse	Security Center	Key Vault
Application Gateway	VPN Gateway	SQL Server Stretch Database	DocumentDB	Azure Active Directory	B2C
Azure DNS	Traffic Manager	Redis Cache	Data Factory	Domain Services	Multi-Factor Authentication
ExpressRoute	Network Watcher	Intelligence & Analytics		Developer Services	
Storage		HDInsight	Machine Learning	Visual Studio Team Services	Azure DevTest Labs
Storage: Blobs, Tables, Queues, Files, Disks	Data Lake Store	Cognitive Services	Azure Bot Service*	VS Application Insights	API Management
StorSimple	Azure Backup	Data Lake Analytics	Power BI Embedded	HockeyApp	Developer Tools
Site Recovery		Azure Analysis Services		Service Profiler*	
Monitoring & Management					
Azure Portal	Azure Resource Manager	Azure Advisor	Azure Monitor	Log Analytics	Automation
				Scheduler	

Why the Cloud

Common Pros

- ✓ Eliminate/reduce data center management
- ✓ Elasticity (scaling up/down)
- ✓ Self-service provisioning of services
- ✓ Ease of experimentation (agility)
- ✓ Faster time-to-market
- ✓ Easier high availability and disaster recovery
- ✓ Subscription-based operating expenses (rather than capital expenses with large up-front investment)

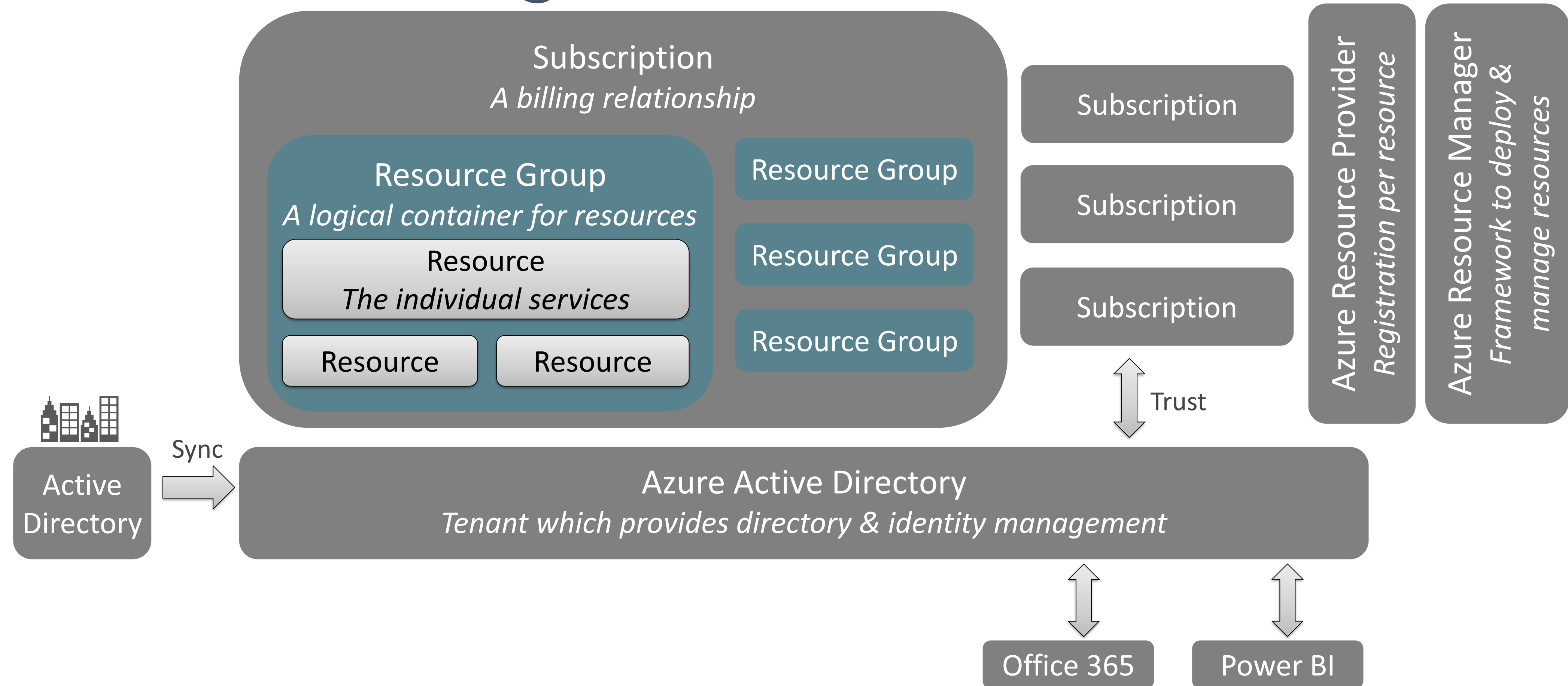


Your goals for the cloud affect the tradeoffs & decisions you will make on control, complexity, cost, performance & security

Common Concerns

- ✓ Uptime guarantees
- ✓ Performance
- ✓ Security
- ✓ Compliance, regulations, legal
- ✓ Sharing of resources (multi-tenancy; noisy neighbors)
- ✓ Data and intellectual property privacy
- ✓ Vendor lock-in/dependency
- ✓ Connecting legacy systems (hybrid/on-prem)
- ✓ Sprawl of self-provisioned services
- ✓ Lack of cloud expertise
- ✓ Complexity ← *The cloud is not *always* easier*
- ✓ Cost ← *The cloud is not *always* cheaper*
- ✓ Difficult to estimate cost up-front

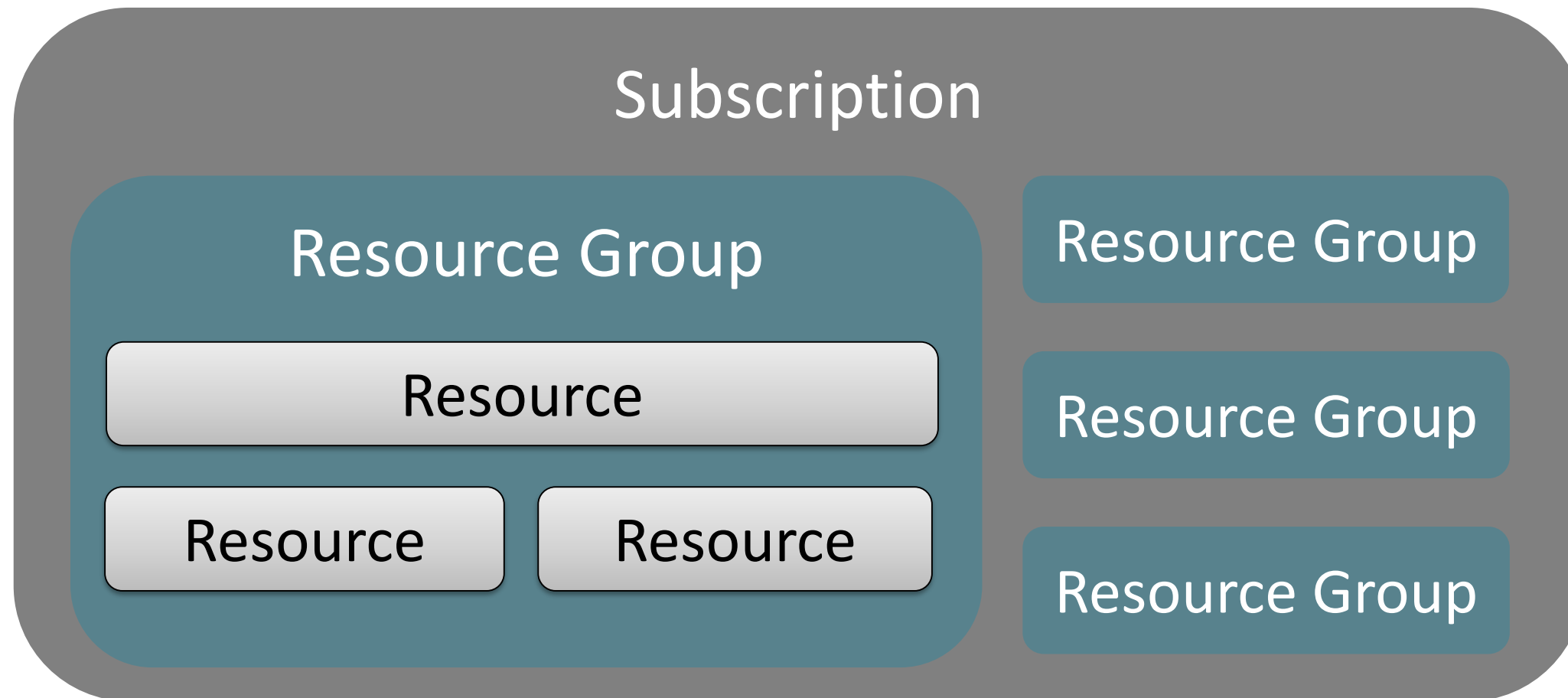
The Azure Lingo



Tip #2

Plan for Resource Groups
Very Carefully

Resource Groups - Considerations



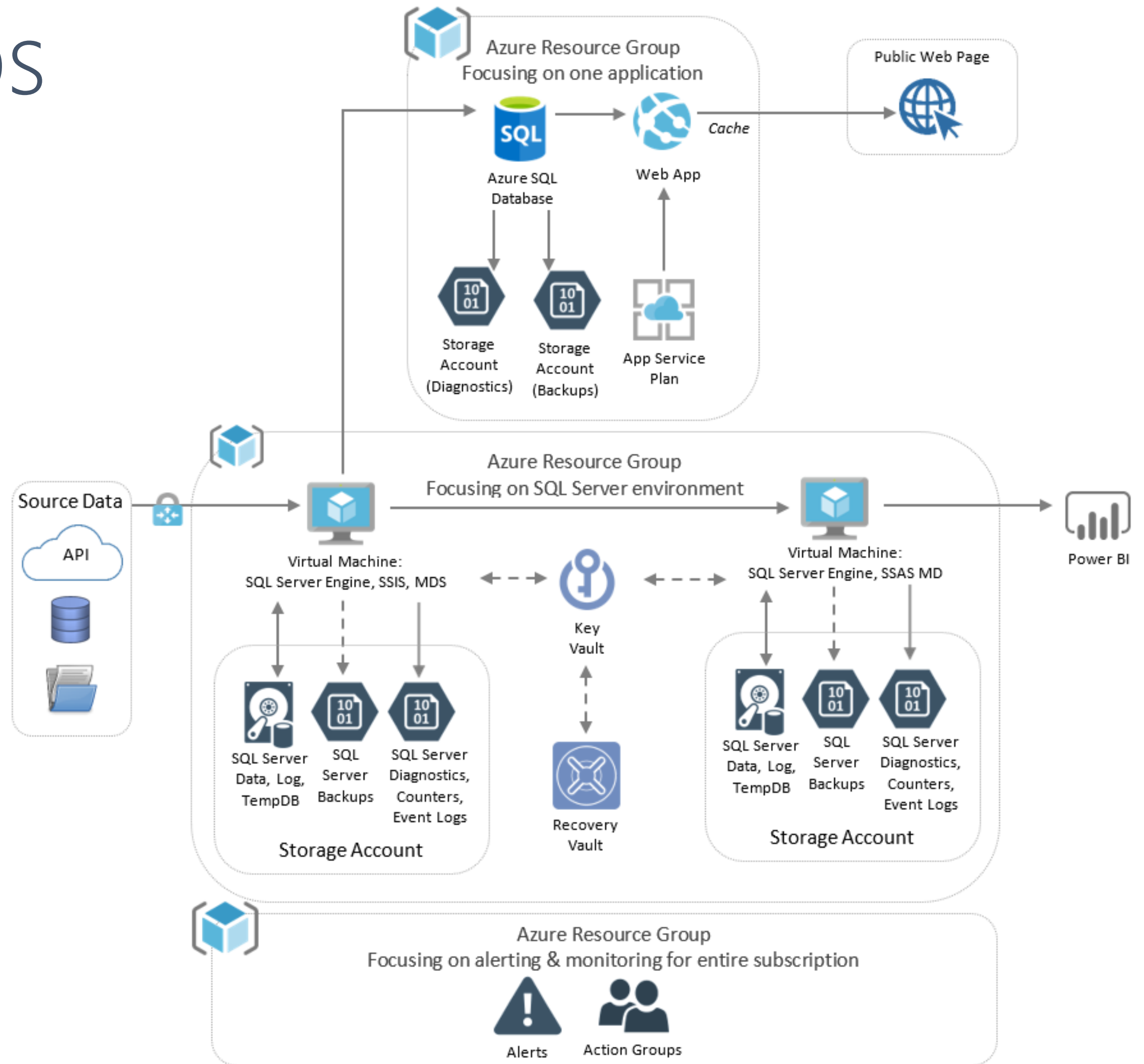
Considerations:

- ✓ Logical organization by purpose
- ✓ Permissions/security boundaries
- ✓ Policies & management of resources as a group
- ✓ Deployment lifecycle -- RGs are the default scope for declarative ARM deployment scripts
- ✓ Location of services
 - Geographic location of data
 - Proximity to business users
 - Co-location of related resources
 - Minimizing latency
 - Minimizing data egress charges
 - Not all resources/services available in every region
 - Backup/recovery considerations

Resource Groups

Suggestions:

- ✓ Plan very carefully for resource groups
- ✓ Keep resource groups more narrow than broad
- ✓ Select the region (location) carefully



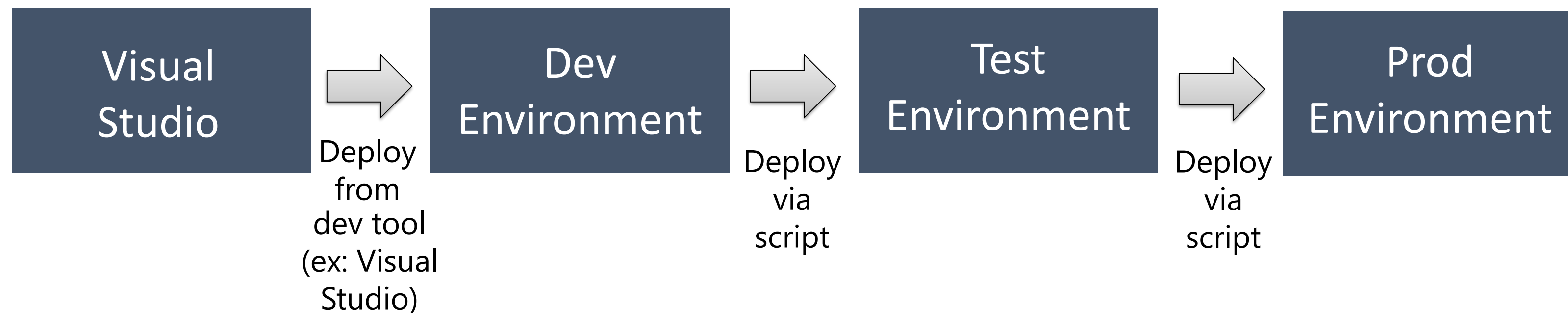
Tip #3

Plan Deliberately for Dev,
Test, Prod Environments

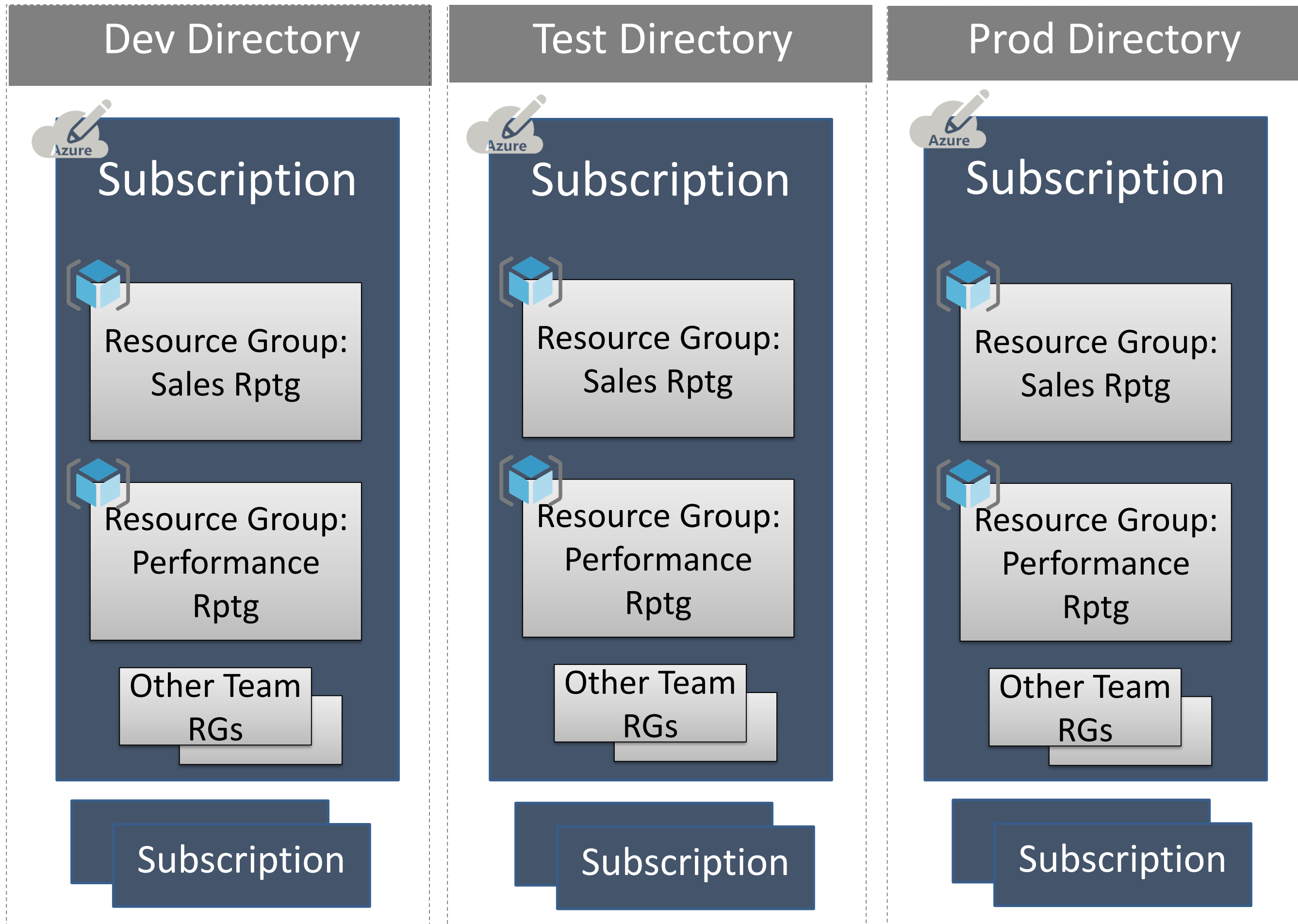
Separation of Dev, Test, Prod Environments

Most commonly environments are segregated by:

1. Directory, or
2. Subscription, or
3. Resource Groups,
4. A combination of 1 and 3, or 2 and 3



Option: Separate By Directory



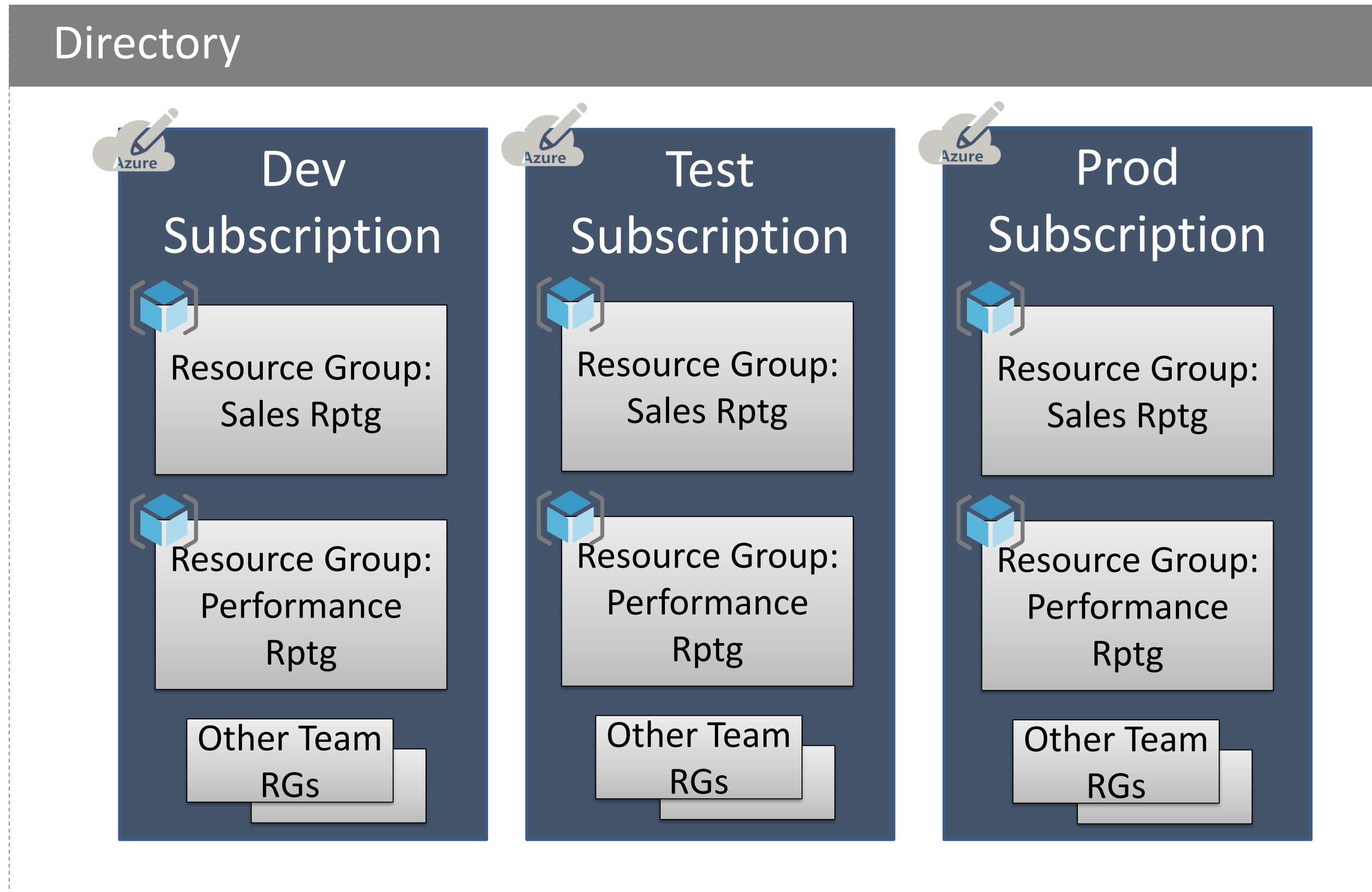
Pros:

- ✓ Clear boundary
- ✓ Offers the most scalability

Cons:

- ✓ More infrastructure to manage
- ✓ A lot of objects intermixed in a subscription - need clear resource group names and/or tags to tell what belongs to which team
- ✓ More complex AAD configuration
- ✓ Separate virtual networks & VPN set up across subscriptions

Option: Separate By Subscription



Pros:

- ✓ Clear boundary

Cons:

- ✓ Cannot always provision new resources if you're just an owner at the resource group level
- ✓ A lot of objects intermixed in a subscription - need clear resource group names and/or tags to tell what belongs to which team
- ✓ Separate virtual networks & VPN set up across subscriptions

Option: Separate by Resource Group

Directory



Analytics Team Subscription



Resource Group:
Sales Rptg Dev



Resource Group:
Sales Rptg Test



Resource Group:
Sales Rptg Prod



Resource Group:
Performance
Rptg Dev



Resource Group:
Performance
Rptg Test



Resource Group:
Performance
Rptg Prod

Other Team Subscriptions

Pros:

- ✓ Isolation of subscriptions by the team who owns/manages
- ✓ Co-admin privileges easier to delegate at the subscription level for each team
- ✓ Billing segregation
- ✓ More subscriptions: might be less likely to hit [Azure subscription limits](#)

Cons:

- ✓ More risk because Prod is mixed in with Dev & Test (mitigate w/ clear names & diff IDs)
- ✓ Separate virtual networks & VPN set up across subscriptions

Demo

Resource Groups, Resources, and
Resource Providers

Tip #4

Use Clear Naming Conventions

Naming Conventions

Purpose → Type of Service → Environment

- ✓ **Type of service** in the name helps with logging/metrics in monitoring scenarios
- ✓ **Environment as the suffix** makes any concatenations easier within scripts
- ✓ **Prod is enumerated** because we separate Dev/Test/Prod within one subscription
- ✓ **No dashes** since all services don't allow them
- ✓ **Camel case** if the service allows it; otherwise lower case

Resource Groups

InternalReporting**RG**Dev
InternalReporting**RG**Test
InternalReporting**RG**Prod

Virtual Machines

BISQL**VM**1Dev
BISQL**VM**1Test
BISQL**VM**1Prod

Storage Accounts

BISQLVM1**DataStd****Strg**Dev
BISQLVM1**BckStd****Strg**Dev
BISQLVM1**DiagStd****Strg**Dev

Tip #5

Use Tags to Categorize Billing
Charges & Other Metadata

Standard Tags

Tags are very helpful for:

- ✓ Billing or cost center categories
- ✓ Environment names
- ✓ Project or system
- ✓ Purpose or application
- ✓ Team, department, business unit
- ✓ Who owns or supports a resource
- ✓ Release or version #s (ex: testing infrastructure)
- ✓ Archival date (ex: if needed only temporarily)
- ✓ Patching or maintenance window or SLA
- ✓ Which customer it applies to (ex: if an ISV)
- ✓ etc...

Meter Name	Consumed Service	Resource Group	Tags
"Standard IO - Block Blob Read Operation Units (in 10,000s)"	"Microsoft.Storage"	"internalreportingrgdev"	"{"billingCategory":"Internal Analytics","environmentType":"Dev","supportContact":"Analytics Team"}"
"Standard IO - Table Write Operation Units (in 10,000s)"	"Microsoft.Storage"	"internalreportingrgdev"	"{"billingCategory":"Internal Analytics","environmentType":"Dev","supportContact":"Analytics Team"}"
"Standard IO - Page Blob/Disk (GB)"	"Microsoft.Storage"	"internalreportingrgdev"	"{"billingCategory":"Internal Analytics","environmentType":"Dev","supportContact":"Analytics Team"}"
"Standard IO - Table (GB)"	"Microsoft.Storage"	"internalreportingrgdev"	"{"billingCategory":"Internal Analytics","environmentType":"Dev","supportContact":"Analytics Team"}"

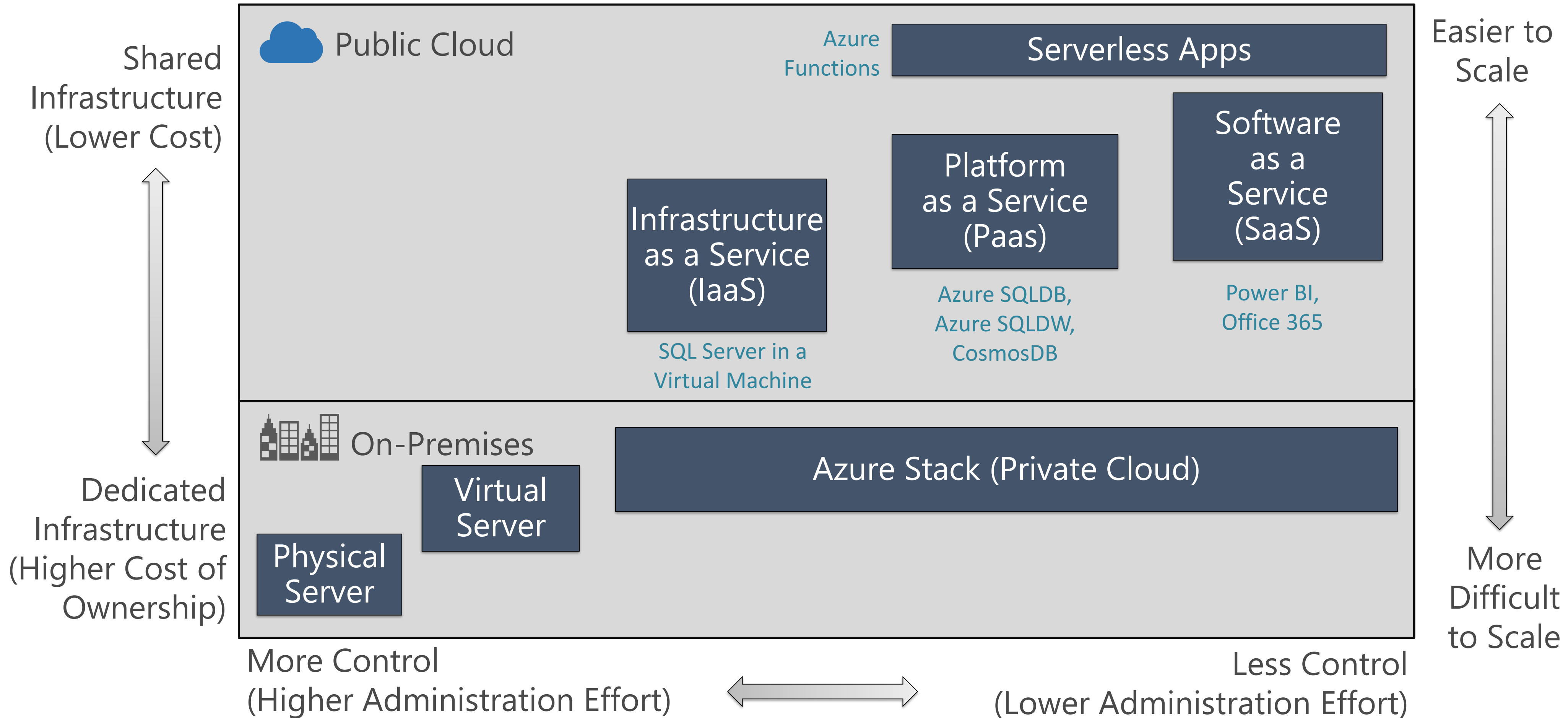
Demo

Find Resources Based on Tags

Tip #6

Determine if IaaS vs PaaS vs SaaS
Best Meets Your Needs

IaaS vs. PaaS vs. SaaS



Comparing the SQL Offerings in Azure

(1/2)



SQL Server in a Virtual Machine
(IaaS)

Run full workload within an Azure virtual machine, including SQL Server, SSIS, SSAS, SSRS, etc



Azure SQL Database
(PaaS)

A relational database-as-a-service (DBaaS)

Non-Managed

A traditional Azure SQLDB deployment (isolated DB)

Managed Instance

Newer - closer feature parity to SQL Server (instance level features)

Elastic Pools

Unpredictable multi-tenant OLTP workloads

(NOT for data warehousing workloads)



Azure SQL Data Warehouse
(PaaS)

An data warehouse-as-a-service (DWaaS) optimized for performance of large-scale analytical workloads

MPP architecture (massively parallel processing)

Comparing the SQL Offerings in Azure

(2/2)



SQL Server in a Virtual Machine (IaaS)

Best for:

- ✓ Migrating existing solutions
- ✓ Running all SQL Server features (ex: SSIS, SSRS, SSAS, MDS)
- ✓ Running 3rd party software (ex: Oracle)
- ✓ Bring your own license (Software Assurance)
- ✓ Isolated dev/test environments
- ✓ Administering all aspects (highest level of control)



Azure SQL Database (PaaS)

Best for:

- ✓ New database solutions
- ✓ Migrating existing database solutions as-is (managed instances which are in preview)
- ✓ OLTP with scaling & pooling needs (unpredictable workloads)
- ✓ DW workloads which are small to medium
- ✓ Reduced administration



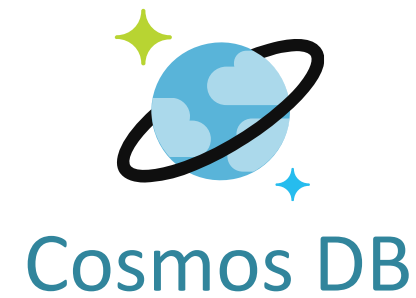
Azure SQL Data Warehouse (PaaS)

Best for:

- ✓ DW with larger data volumes (ideally 1-4TB+ bare min. of 250GB+)
- ✓ Dimensional models which can take advantage of CCI
- ✓ Ability to scale compute up/down, or pause (elasticity)
- ✓ Data Lake Store integration (relational + nonrelational data)
- ✓ Reduced administration

More info: <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-paas-vs-sql-server-iaas>
<https://www.blue-granite.com/blog/is-azure-sql-data-warehouse-a-good-fit>

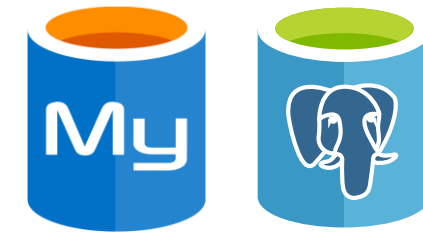
Other PaaS Database Offerings in Azure



Cosmos DB

Best for:

- ✓ Schema-agnostic data model that is write-optimized
 - NoSQL JSON documents
 - Key-value pairs
 - Column-family data
 - Graph data
- ✓ Scalable, globally distributed apps
- ✓ Reduced administration



Azure Database
for MySQL and PostgreSQL

Best for:

- ✓ Open source DB (community editions)
- ✓ Focus is mostly on a front-end application
- ✓ Familiar with MySQL or PostgreSQL
- ✓ Reduced administration



Azure Analysis
Services

Best for:

- ✓ Semantic layer for reporting
- ✓ Alternative to SQL Server Analysis Tabular Model (otherwise run in an IaaS VM)
- ✓ Upsizing a data model from Power BI
- ✓ Sharing some workload with Azure SQLDW or Azure SQLDB
- ✓ Reduced administration

Azure SQL DB: Feature Parity with SQL Server

Many features go first to Azure SQLDB ("cloud first").

However, some features are *not available* in Azure SQL Database such as:

- PolyBase (alternative: elastic queries)
- SQL Server Machine Learning & R Services
- Change data capture
- CLR
- DB snapshots
- Some T-SQL syntax
- Profiler
- Non-primary filegroups

Feature	SQL Server	Azure SQL Database
Active geo-replication	Not supported - see Always On Availability Groups	Supported
Always Encrypted	Supported	Supported - see Cert store and Key vault
AlwaysOn Availability Groups	Supported	Not supported - See active geo-replication
Attach a database	Supported	Not supported
Application roles	Supported	Supported
Auto scale	Not supported	Supported - see Service tiers
Azure Active Directory	Not supported	Supported
Azure Data Factory	Supported	Supported
Auditing	Supported	Supported

Full list:

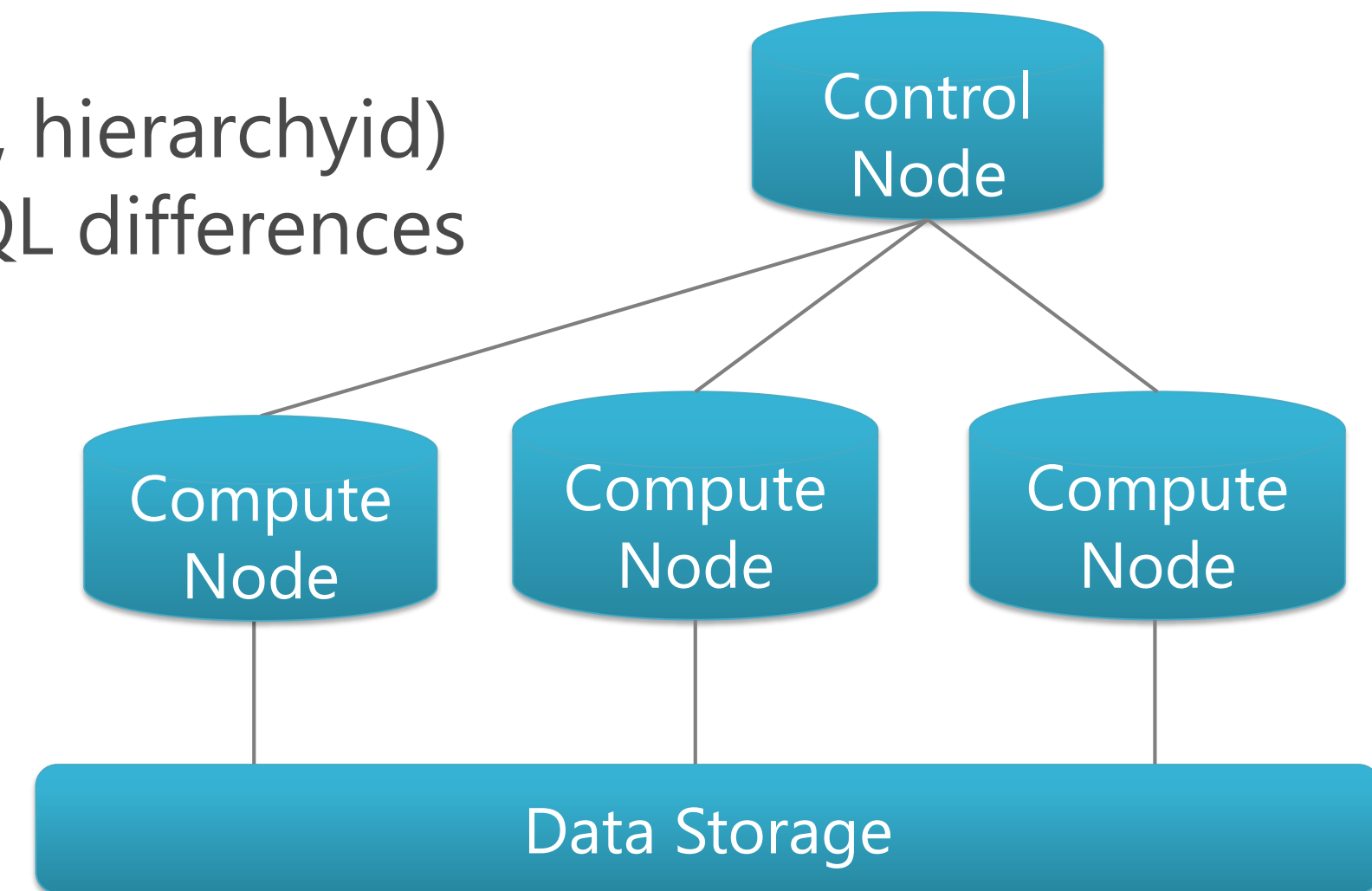
<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-features>

Azure SQL DW: Feature Parity with SQL Server

Some features are *not available* in Azure SQL Data Warehouse:

- Primary key, foreign key, unique key
- Computed columns, default values
- Merge statements
- Some data types (ex: xml, json, geography, geometry, hierarchyid)
- Certain update and delete techniques and other T-SQL differences
- Statistics creation and management
- SSDT database project
- Cross-database queries

The recommended data load technique is via PolyBase.



More info:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-troubleshoot#differences-from-sql-database>

<https://www.blue-granite.com/blog/is-azure-sql-data-warehouse-a-good-fit>

Tip #7

Get an Azure SLA if Possible

How to Get a Service Level Agreement (SLA)



SQL Server in a Virtual Machine (IaaS)

SLA on the VM if:

1. Include the VM in an Availability Set -
Guarantee: 99.95%
or
2. Use Premium Storage for all Data and O/S disks -
Guarantee: 99.9%

Premium Storage is how you get an SLA on a 'single instance' VM. You might consider Standard storage in Dev/Test, and Premium in Prod.



Azure SQL Database (PaaS)

SLA on the database:

- ✓ Guarantee: 99.99% for Basic, Standard, and Premium tiers that customers will have connectivity



Azure SQL Data Warehouse (PaaS)

SLA on the database:

- ✓ Guarantee: 99.99% for Basic, Standard, and Premium tiers that client operations will succeed

More info: <https://azure.microsoft.com/en-us/support/legal/sla/summary/>
https://azure.microsoft.com/en-us/support/legal/sla/virtual-machines/v1_6/

Tip #8

Take Advantage of Efficiencies
Offered by Cloud Services

Starting With a VM Image

(1/2)

The screenshot displays the Azure portal interface for configuring a virtual machine. On the left, the 'Virtual machines' directory is shown with a search filter for 'sql server 2016 SP1'. The search results list several SQL Server images, with 'SQL Server 2016 SP1 Enterprise on Windows Server 2016' selected. On the right, the 'SQL Server settings' and 'Storage configuration' panels are visible. In the 'SQL Server settings' panel, the 'Storage configuration' dropdown is highlighted with a red circle, and the 'Data warehousing' option is selected. Below this, a 'DS12_V2 Standard' VM size is shown with specifications: 4 Cores, 28 GB RAM, 8 Data disks, 12800 Max IOPS, 56 GB Local SSD, Load balancing, and Premium disk support. The price is listed as 485.09 USD/MONTH (ESTIMATED). In the 'Storage configuration' panel, the 'Data warehousing' option is also highlighted with a red circle. A callout bubble points to this selection, stating: 'The DW selection sets storage disk stripe size to 256KB and trace flags 610 & 1117'.

Virtual machines Directory

Compute

Virtual machines and Virtual machines (classic) can now be managed together in the combined list below.

Subscriptions: 1 of 2 selected – Don't see a subscription? Switch directories

Filter by name...

2 items

NAME	PUBLISHER
SQL Server 2016 SP1 Standard on Windows Server 2016	Microsoft
Free License: SQL Server 2016 SP1 Developer on Windows Server 2016	Microsoft
SQL Server 2016 SP1 Enterprise on Windows Server 2016	Microsoft
Free License: SQL Server 2016 SP1 Express on Windows Server 2016	Microsoft
SQL Server 2016 SP1 Web on Windows Server 2016	Microsoft
{BYOL} SQL Server 2016 SP1 Enterprise on Windows Server 2016	Microsoft
{BYOL} SQL Server 2016 SP1 Standard on Windows Server 2016	Microsoft

SQL Server settings

SQL connectivity: Private (within Virtual Network)

Port: 1433

SQL Authentication: Disable Enable

Storage configuration: Data warehousing

DS12_V2 Standard

- 4 Cores
- 28 GB RAM
- 8 Data disks
- 12800 Max IOPS
- 56 GB Local SSD
- Load balancing
- Premium disk support

485.09 USD/MONTH (ESTIMATED)

Storage configuration

Select your desired performance, storage size, and workload to optimize the storage on your virtual machine.

Storage size (TB): 1

1 data disks will be added to the virtual machine. This value was computed based on the value of IOPS, throughput, and storage size.

Storage optimization: Data warehousing

General

Transactional processing

Data warehousing

The DW selection sets storage disk stripe size to 256KB and trace flags 610 & 1117



Some recommendations in the "Performance Best Practices for SQL Server in Azure Virtual Machines" article you'll need to implement yourself. <https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sql/virtual-machines-windows-sql-performance>

Starting With a VM Image

(2/2)

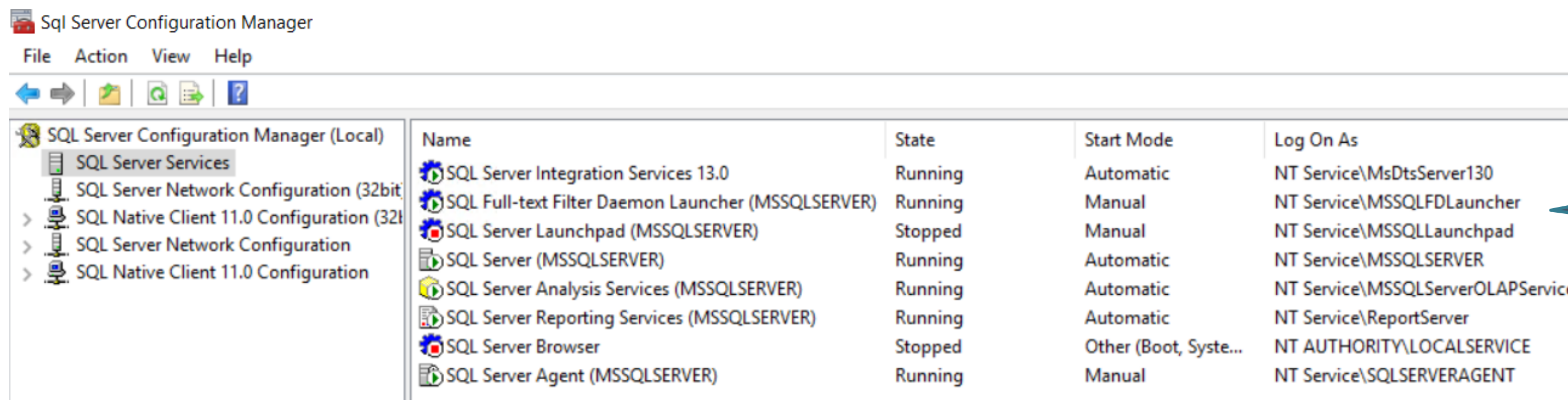
Expect to make changes and customizations to the VM image, such as:

- Domain service accounts for each service
- Disable unused services
- Disk structure + permissions
 - G: Data
 - L: Logs
 - T: TempDB
- Enable disk encryption

By default, there is one data drive on F: provisioned at 1TB. Stay away from D:\ - it's the temp drive.

SQL Server changes to VM image:

- Enable the DAC
- Default DB file locations
- Fill factor
- Optimize for ad hoc workloads
- Max degree of parallelism
- Cost threshold for parallelism
- Startup parameters; trace flags
- Expand # of TempDB files
- Time of SSIS maintenance job etc...



Name	State	Start Mode	Log On As
SQL Server Integration Services 13.0	Running	Automatic	NT Service\MsDtsServer130
SQL Full-text Filter Daemon Launcher (MSSQLSERVER)	Running	Manual	NT Service\MSSQLFDLauncher
SQL Server Launchpad (MSSQLSERVER)	Stopped	Manual	NT Service\MSSQLLaunchpad
SQL Server (MSSQLSERVER)	Running	Automatic	NT Service\MSSQLSERVER
SQL Server Analysis Services (MSSQLSERVER)	Running	Automatic	NT Service\MSSQLServerOLAPService
SQL Server Reporting Services (MSSQLSERVER)	Running	Automatic	NT Service\ReportServer
SQL Server Browser	Stopped	Other (Boot, System...)	NT AUTHORITY\LOCALSERVICE
SQL Server Agent (MSSQLSERVER)	Running	Manual	NT Service\SQLSERVERAGENT

As-is before any customizations

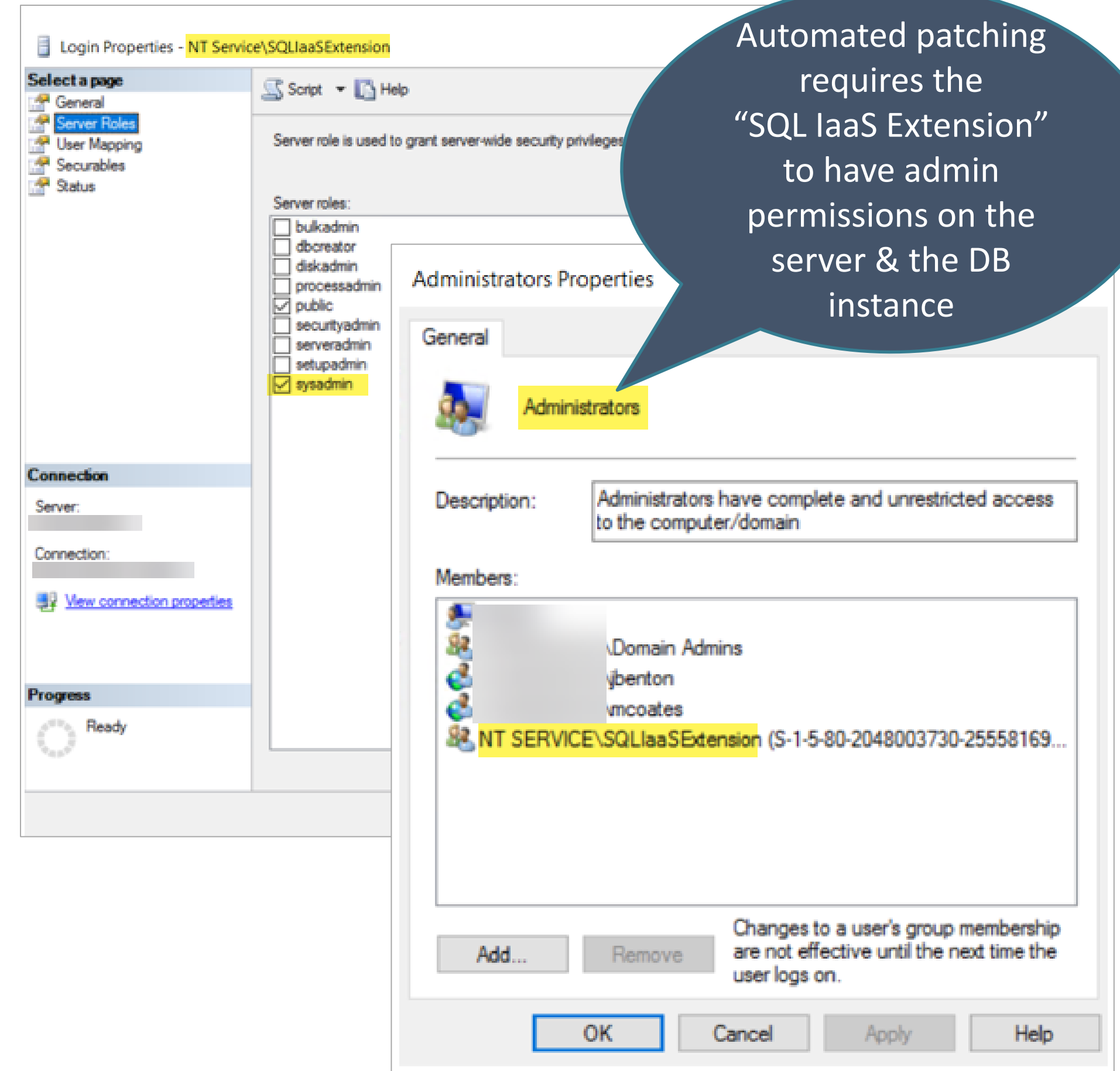
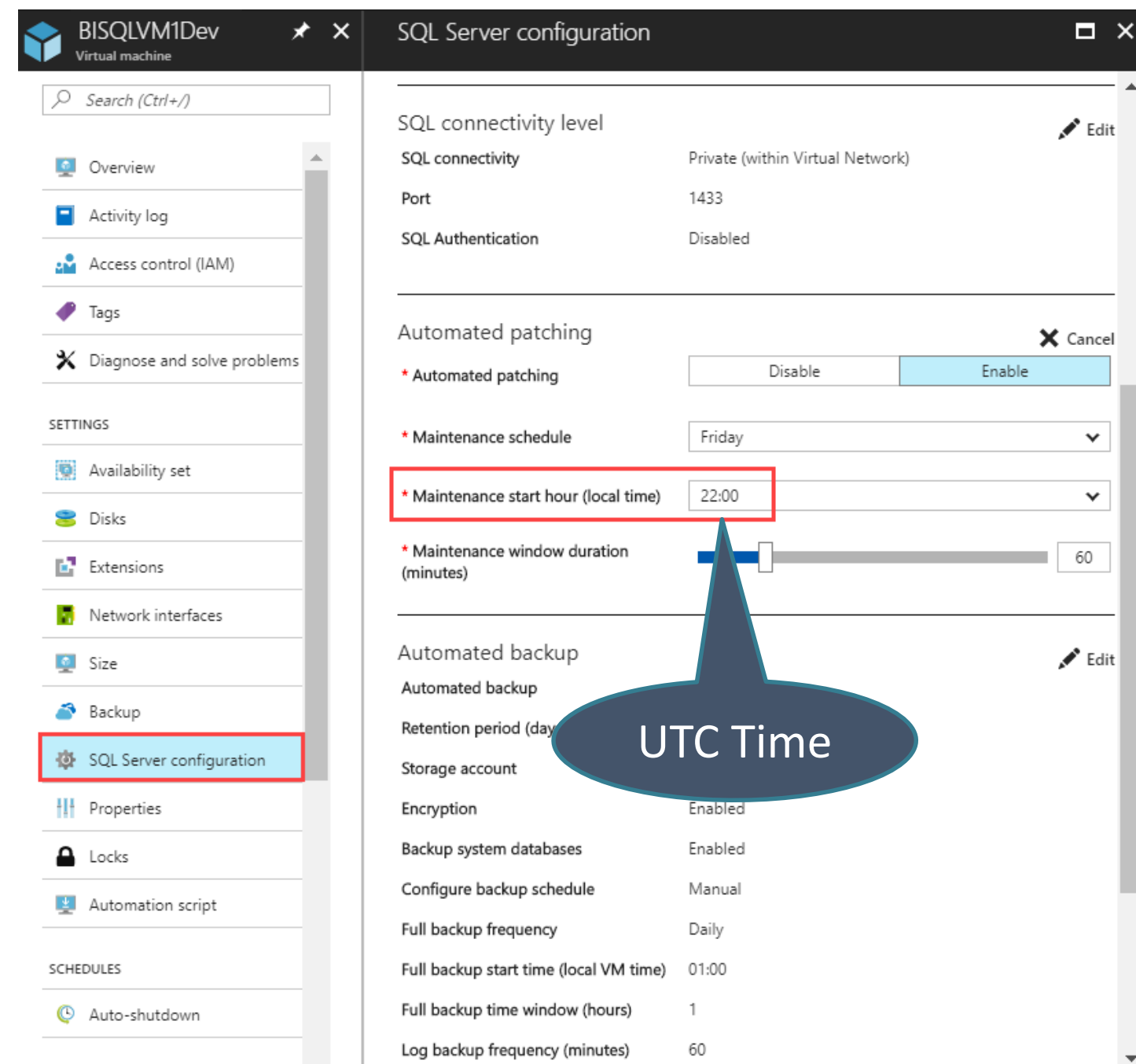
Automatic Patching

Windows and SQL Server patching

PaaS: Done for you (no opt-out or delay).

IaaS: You can optionally allow MSFT to handle.

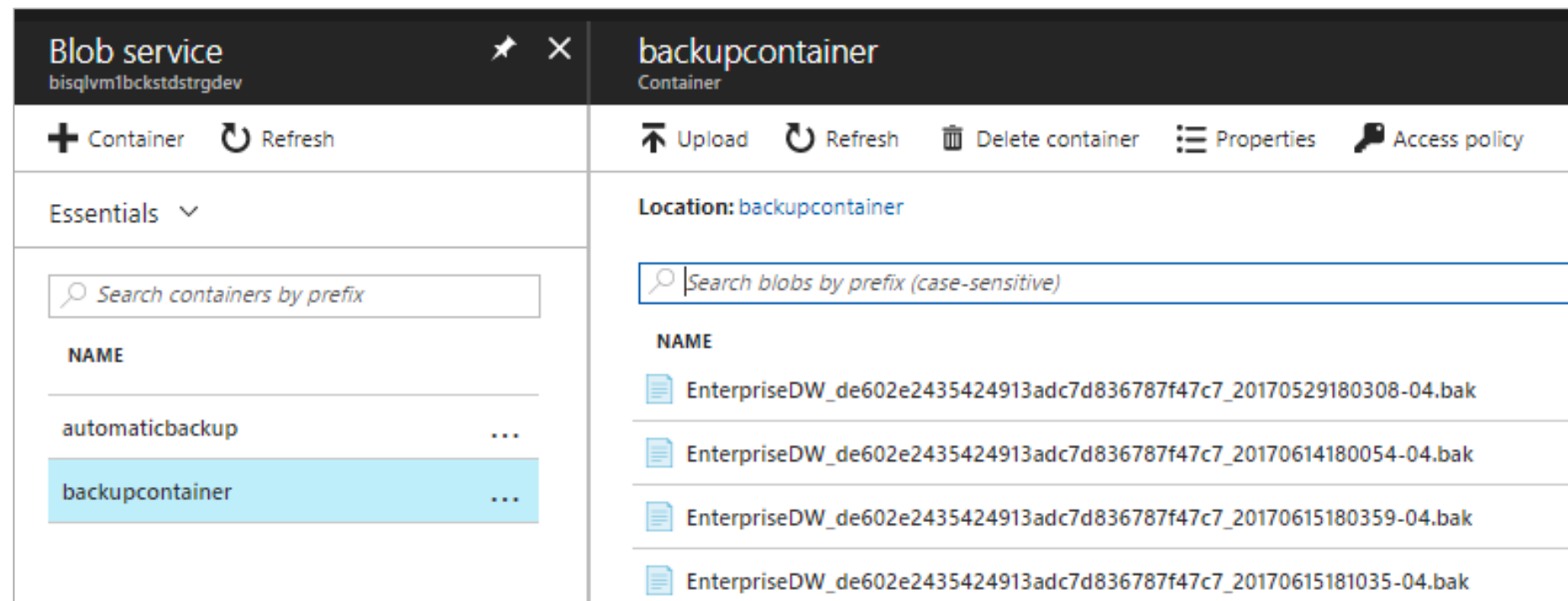
There's also a new VM "Update Management" feature.



More info: <https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sql/virtual-machines-windows-sql-server-agent-extension>

SQL Backups: IaaS

IaaS: You can optionally set up scheduled SQL Server backups.

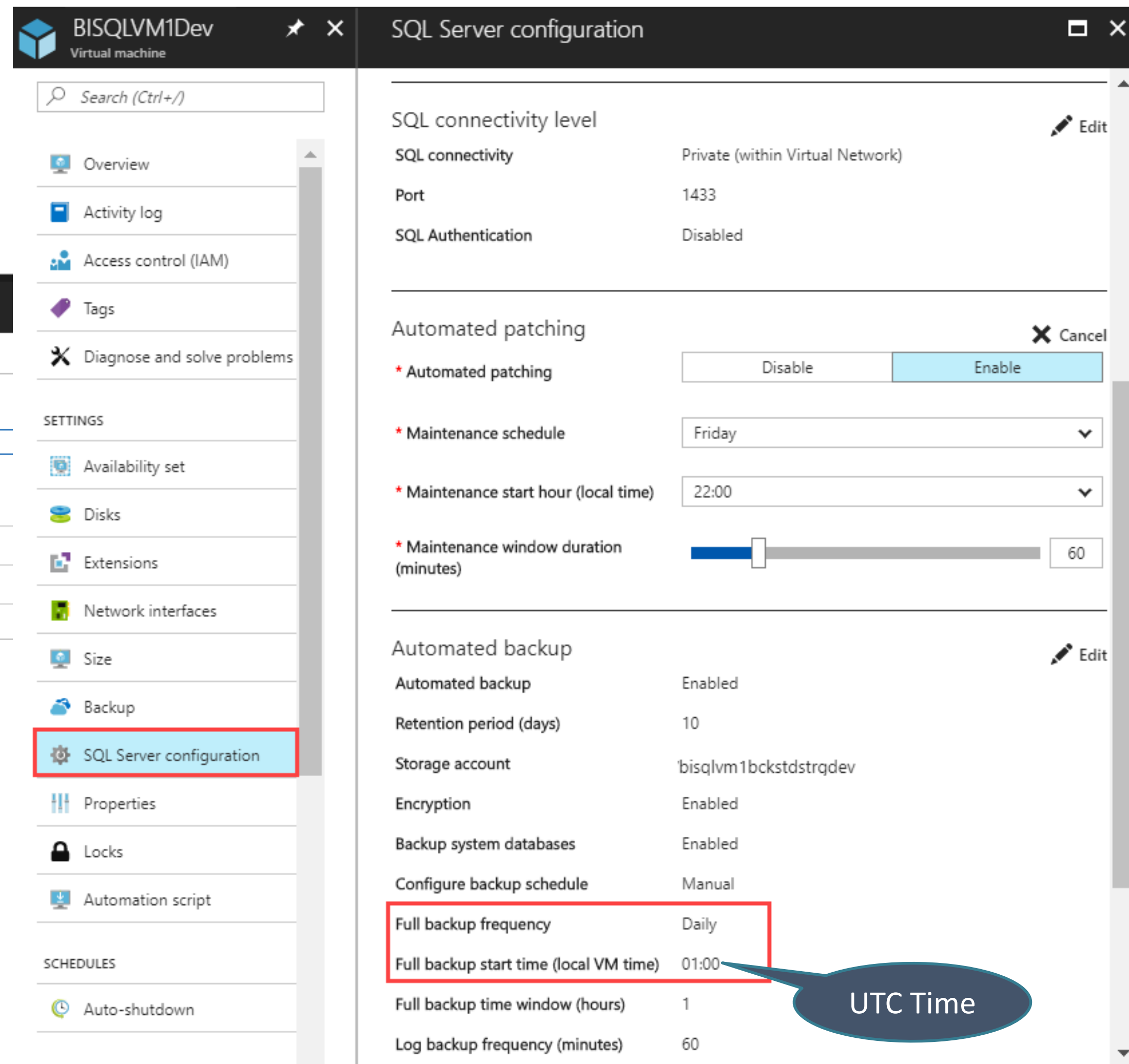


The screenshot shows the Azure Storage Explorer interface for a Blob service named 'bisqlvm1bckstdstrgdev'. The selected container is 'backupcontainer'. The 'Essentials' pane on the left shows a search bar and a list of containers, with 'backupcontainer' selected. The main pane displays a list of backup files with columns for 'NAME' and 'LOCATION'. The files are:

NAME	LOCATION
EnterpriseDW_de602e2435424913adc7d836787f47c7_20170529180308-04.bak	backupcontainer
EnterpriseDW_de602e2435424913adc7d836787f47c7_20170614180054-04.bak	backupcontainer
EnterpriseDW_de602e2435424913adc7d836787f47c7_20170615180359-04.bak	backupcontainer
EnterpriseDW_de602e2435424913adc7d836787f47c7_20170615181035-04.bak	backupcontainer

Suggestions:

- ✓ Be sure that this Azure storage account is set to be encrypted since the "automaticbackup" container contains certificates & keys
- ✓ Provision the storage account with backups in another region, or with geo-redundant storage



The screenshot shows the 'SQL Server configuration' window for a virtual machine named 'BISQLVM1Dev'. The 'Automated backup' section is expanded, showing the following settings:

Setting	Value
Automated backup	Enabled
Retention period (days)	10
Storage account	bisqlvm1bckstdstrgdev
Encryption	Enabled
Backup system databases	Enabled
Configure backup schedule	Manual
Full backup frequency	Daily
Full backup start time (local VM time)	01:00
Full backup time window (hours)	1
Log backup frequency (minutes)	60

A red box highlights the 'Full backup frequency' and 'Full backup start time (local VM time)' settings. A blue callout bubble points to the '01:00' value, containing the text 'UTC Time'.

SQL Backups: PaaS



Azure SQL Database (PaaS)

Basic service: retained 7 days
Standard service: retained 35 days
Premium service: retained 35 days

Full backups: weekly
Differential: every few hours
T-log: every 5-10 minutes

Longer-term backup retention:

- Set up a process to back up the .bacpac file (schema + data), or
- Recovery Services vault (still in Preview)



Azure SQL Data Warehouse (PaaS)

Local snapshots: every 4-8 hours & retained 7 days
Geo-redundant snapshot: retained 24 hours
(Backups occur when DW is online -- not paused)

Longer-term backup retention:

- Not currently an option



*The tradeoff here: less administration,
but also less control*

VM Backups

VM backups are managed in the Recovery Services Vault.

To ensure the VM backups are also encrypted, Azure disk encryption is required.

The screenshot displays the Azure portal interface for a backup item named 'bisqlvm1dev'. The left pane shows the 'Essentials' section with details for the Recovery services vault (BIRecoverServicesVaultEastDev), subscription name, subscription ID, item type (Azure virtual machine), and backup policy (BIBackupPolicyWeekly). It also shows the last backup status as 'Success' and the latest restore point as '2/7/2017, 10:16:41 AM (1 hour(s) ago)'. The right pane shows the 'Restore points' section, filtered for the last 30 days. It includes a table with columns for TIME, CONSISTENCY, and ENCRYPTION. The ENCRYPTION column is highlighted with a red box, showing the value 'Encrypted' for the restore point on 2/7/2017 at 10:16:41 AM. Below the table, there is a summary of restore points for the last 30 days and last 7 days, both showing 1 restore point.

TIME	CONSISTENCY	ENCRYPTION
2/7/2017, 10:16:41 AM	Application Consistent	Encrypted

Restore points	Count
Last 30 days	1
Last 7 days	1

Basic Monitoring

Diagnostic Logs

bankingadsdev - Diagnostics logs
Data Lake Store

Search (Ctrl+)

Refresh

* Subscription Resource group Resource type
Microsoft Azure Sponsorship BankingRGDev Data Lake Store

Microsoft Azure Sponsorship > BankingRGDev > bankingadsdev

Turn on diagnostics to collect the following data.

- Audit
- Requests
- AllMetrics

Metrics

bankingadsdev - Activity log
Data Lake Store

Search (Ctrl+)

Columns Export Log Analytics Operation log (classic)

Select query ...

* Subscription Resource group Resource Event severity
Microsoft Azure Sp... BankingRGDev All resources 4 selected

Timespan Event category Event initiated by
Last month All categories Email or name or s...

Apply Reset

Query returned 22 items. Click here to download all the items as csv.

OPERATION NAME	STATUS	TIME	TIME STAMP	SUBSCRIPTION
RunFinished	Succeeded	4 d ago	Sun Oct 08 2...	Microsoft Azure Sponsorship
Extensiondata	Succeeded	4 d ago	Sun Oct 08 2...	Microsoft Azure Sponsorship
Write Datapipelines	Accepted	4 d ago	Sun Oct 08 2...	Microsoft Azure Sponsorship

SHARED SERVICES

- Alerts
- Metrics
- Metrics (preview)
- Log Analytics
- Activity log
- Service Health

SOLUTIONS

- Application Insights
- Network watcher
- Management solutions

Available metrics

Filter metrics...

You can only select metrics of the same unit (bytes)

- Data Read
- Data Written
- Read Requests
- Total Storage
- Write Requests

Metrics chart

BANKINGADLSDEV

Chart type: Line Time range: Past week

600kB
500kB
400kB
300kB
200kB
100kB
0kB

Oct 6 Oct 7 Oct 8 Oct 9 Oct 10 Oct 11 Oct 12

DATA WRITTEN
549.71 kB

Activity Log

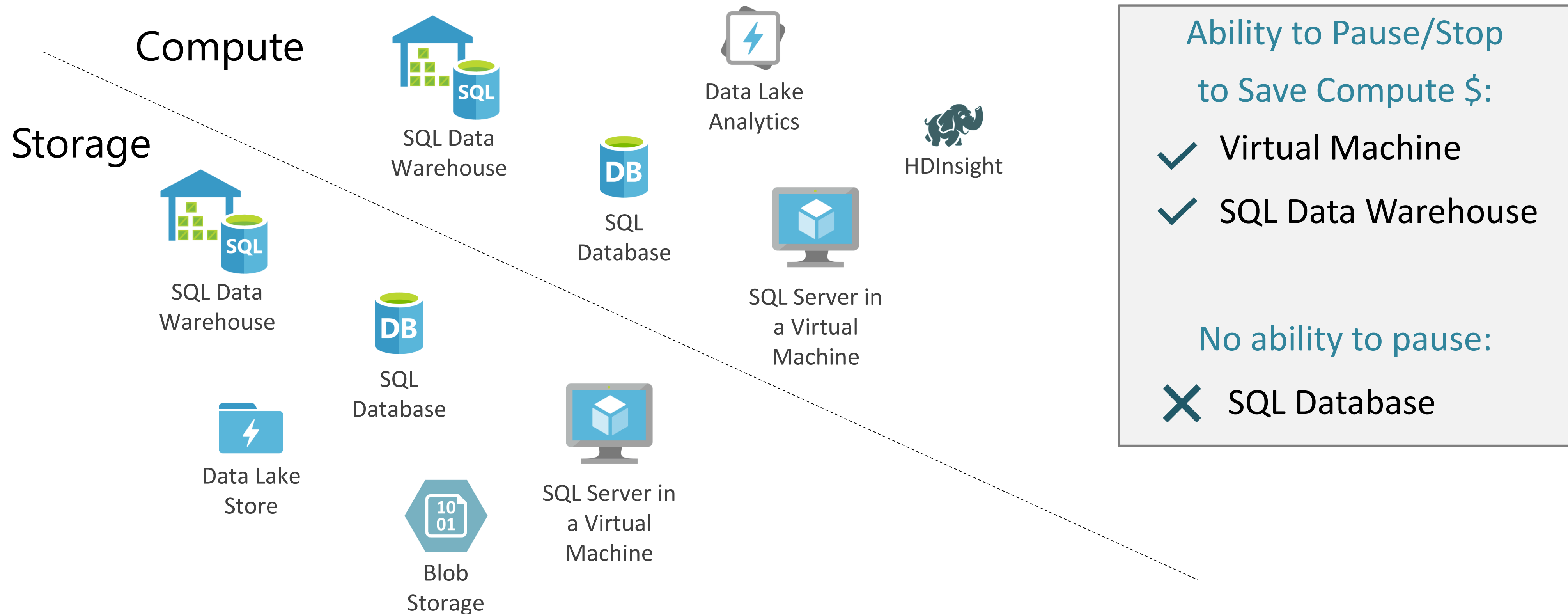
Tip #9

Be Aware of Cost Savings
Techniques

Scaling and Pausing Compute Resources

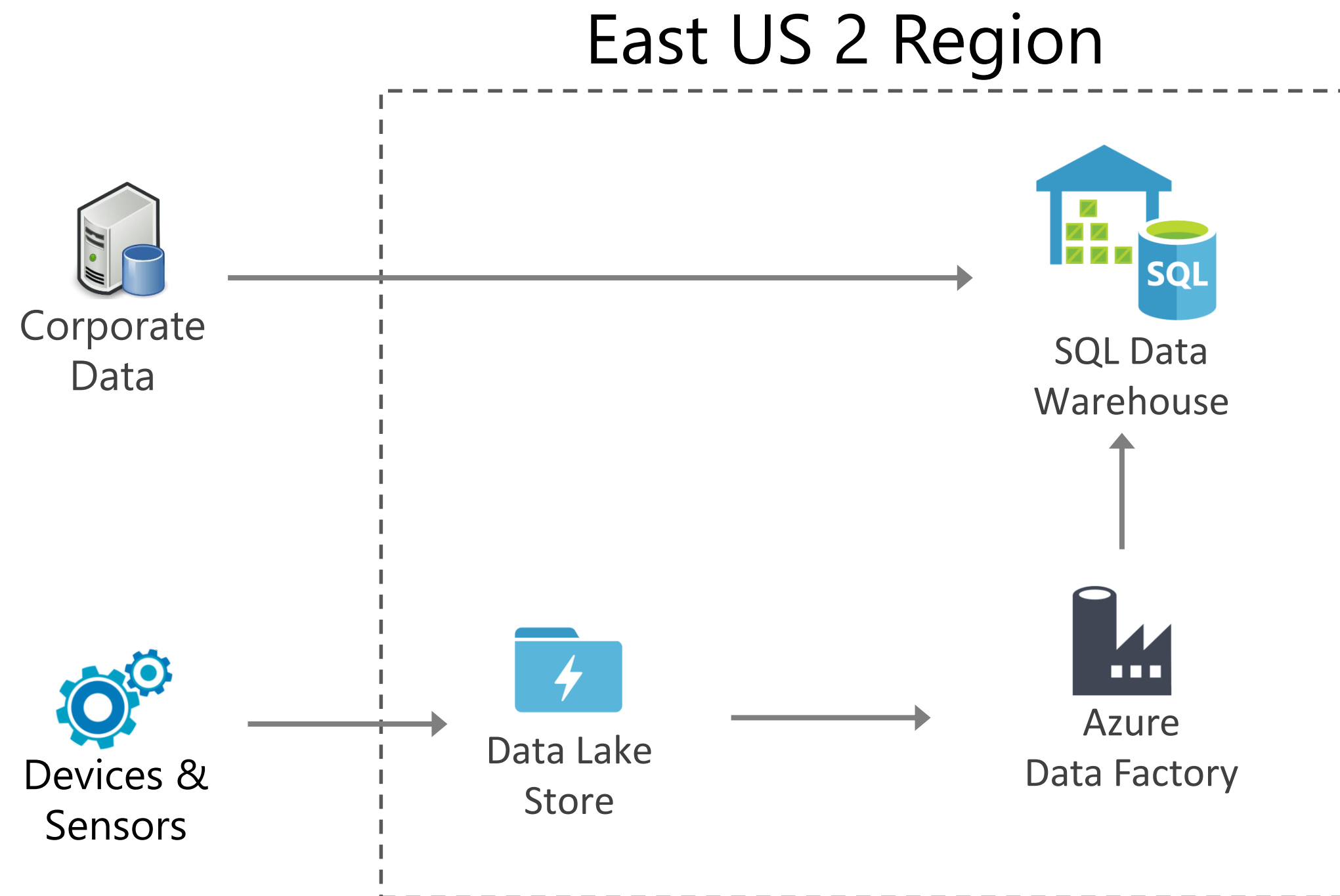
Some services support pausing or stopping compute resources when it is idle.

Some services support scaling compute resources up & down to support varying demands.



Co-Location to Avoid Data Egress Charges

Data egress charges are incurred when data leaves an Azure region. No charges to ingest data into Azure. Data egress charges are very cheap – but they can easily add up at scale.



Scale Lower in Dev/Test Environments

SQL databases


+ Add Columns Refresh

Subscriptions: Analytics – Don't see a subscription? [Switch directories](#)

Filter by name...

All locations

2 items

NAME	STATUS	REPLICATION ROLE	SERVER	PRICING TIER
 sqlwaitssqldatabase	Online	None	sqlwaitsserverdev	Standard: S0
 sqlwaitssqldatabase	Online	None	sqlwaitsserverprod	Standard: S2



In a subscription model, developers and admins can have a huge influence on cost

Automatic Shutdown of VMs

BISQLVM1Dev - Auto-shutdown
Virtual machine

Search (Ctrl+/)

Save Discard Feedback

Like auto-shutdown? DevTest Labs has more features to control costs and streamline your development workflows. [Learn more.](#)

Enabled
 On Off

Scheduled shutdown
9:00:00 PM

Time zone
(UTC-05:00) Eastern Time (US & Canada)

Send notification 15 minutes before auto-shutdown?
 Yes No

* Webhook URL

Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems

SETTINGS

Availability set
Disks
Extensions
Network interfaces
Size
Backup
SQL Server configuration
Properties
Locks
Automation script

SCHEDULES

Auto-shutdown

MONITORING

Metrics
Alert rules

Shutting the VM down within Windows does ***NOT*** stop compute charges. It needs to be de-allocated in Azure to reduce cost.

Dev & Test

Shutdown:

- ✓ Auto-shutdown feature within the VM properties

Start:

- ✓ VM is manually restarted when needed by one of the developers

Production

Shutdown & Start:

- ✓ A machine which needs to be restarted by a certain time (ex: DB engine+SSIS) can be done with:
 - Azure Automation (PowerShell)
 - Microsoft Flow
 - Azure Functions
 - etc...

Saving Money Running Virtual Machines

Low Priority VMs on Azure Batch

Runs on a pool of VMs; takes advantage of compute surplus in Azure

Useful when:

- ✓ Job completion time is flexible
- ✓ Batch job can be parallelized & distributed across VMs

<https://docs.microsoft.com/en-us/azure/batch/batch-low-pri-vms>

https://blogs.msdn.microsoft.com/uk_faculty_connection/2017/05/14/microsoft-azure-makes-vms-affordable-for-everybody/

Prepaid Reserved VM

Upfront commitment for compute capacity at a large discount

Useful when:

- ✓ Predictable, long-term workloads

<https://azure.microsoft.com/en-us/pricing/reserved-vm-instances/>

“Burstable” B-Series VMs

Accumulates ‘saved’ credits for unused CPU cycles—uses the credits for bursting

Useful when:

- ✓ Workload is small but a dedicated VM is still desirable
- ✓ Bursts only occasionally (ex: once a day)

<https://buildazure.com/2017/09/11/more-affordable-azure-vms-with-burstable-b-series/>

Dev/Test Labs

Offers lower rates for certain Azure services.

Useful for:

- ✓ Learning
- ✓ Experimenting
- ✓ Proof of concept
- ✓ Throwaway work
- ✓ Non-production

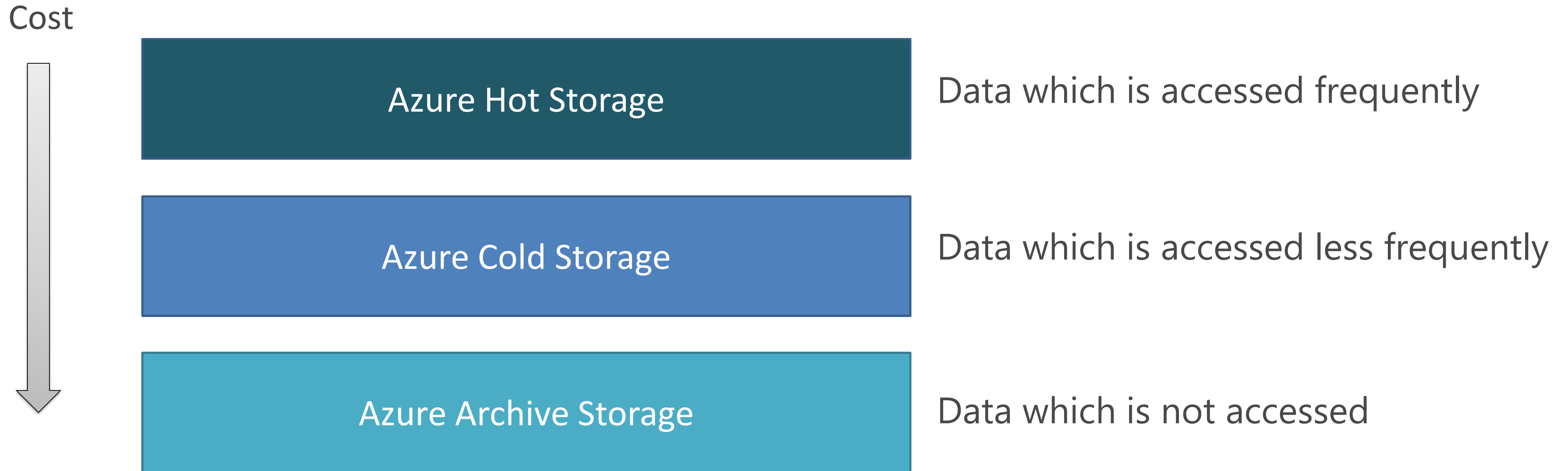
Specify policies and thresholds to control cost.

Requires active Visual Studio subscription.

	FOR INDIVIDUALS	FOR TEAMS	FOR TEAMS
Dev/Test pricing options	Monthly Azure credits for Visual Studio subscribers	Enterprise Dev/Test	Pay-As-You-Go Dev/Test
Number of instances allowed	One per Visual Studio subscriber	Unlimited	Unlimited
Monthly credit included*	\$50, \$100, or \$150	None	None
Billing options	No credit card required Credit card billing for overages (invoicing available)	Always billed to the Enterprise Agreement Consumes prepaid Azure Monetary Commitment funds, if available	Credit card billing (invoicing available)
Billed usage	Overages are billed, after the monthly credit is consumed	All usage is billed	All usage is billed
Allowed to run Windows 10 Virtual Machines	✓	✓	✓
Service Level Agreement	Only for Visual Studio Team Services , Application Insights , and HockeyApp	Only for Visual Studio Team Services , Application Insights , and HockeyApp	Only for Visual Studio Team Services , Application Insights , and HockeyApp
Licensing terms	Restricted to dev/test use only For use by active Visual Studio subscribers, and by end users providing feedback and performing acceptance tests	Restricted to dev/test use only For use by active Visual Studio subscribers, and by end users providing feedback and performing acceptance tests	Restricted to dev/test use only For use by active Visual Studio subscribers, and by end users providing feedback and performing acceptance tests

<https://azure.microsoft.com/en-us/pricing/dev-test/>
<https://azure.microsoft.com/en-us/services/devtest-lab/>

Efficient Use of Storage Tiers



Virtual Machine Storage

	Premium Disk	Standard Disk
Managed Disk	Billed for disk size reservation	Billed for disk size reservation
Unmanaged Disk	Billed for disk size reservation	Billed for space used

Premium (SSD)
IOPS & throughput
are associated with
disk size

	P4	P6	P10	P20	P30	P40	P50
Disk Size	32 GB	64 GB	128 GB	512 GB	1 TB	2 TB	4 TB
Price per month	\$4.81	\$9.29	\$17.92	\$66.56	\$122.88	\$235.52	\$450.56
IOPs per disk	120	240	500	2,300	5,000	7,500	7,500
Throughput per disk	25 MB/second	50 MB/second	100 MB/second	150 MB/second	200 MB/second	250 MB/second	250 MB/second

Standard (HDD)
IOPS & throughput
are based on VM size
+ tier + # of disks

	S4	S6	S10	S15	S20	S30	S40	S50
Disk Size	32 GB	64 GB	128 GB	256 GB	512 GB	1 TB	2 TB	4 TB
Price per month (promotional)	\$0.77	\$1.51	\$2.95	\$5.67	\$10.88			
						VM Tier	Basic Tier VM	Standard Tier VM
						Max Disk size	4095 GB	4095 GB
						Max 8 KB IOPS per disk	Up to 300	Up to 500
						Max Bandwidth per disk	Up to 60 MB/s	Up to 60 MB/s

Use Scripts to Override Default Storage Sizes

Azure SQL DW

- ✓ Requires Premium storage
- ✓ Provisions 1 TB by default if initially created in the portal (or if maxsize is unspecified in a script)

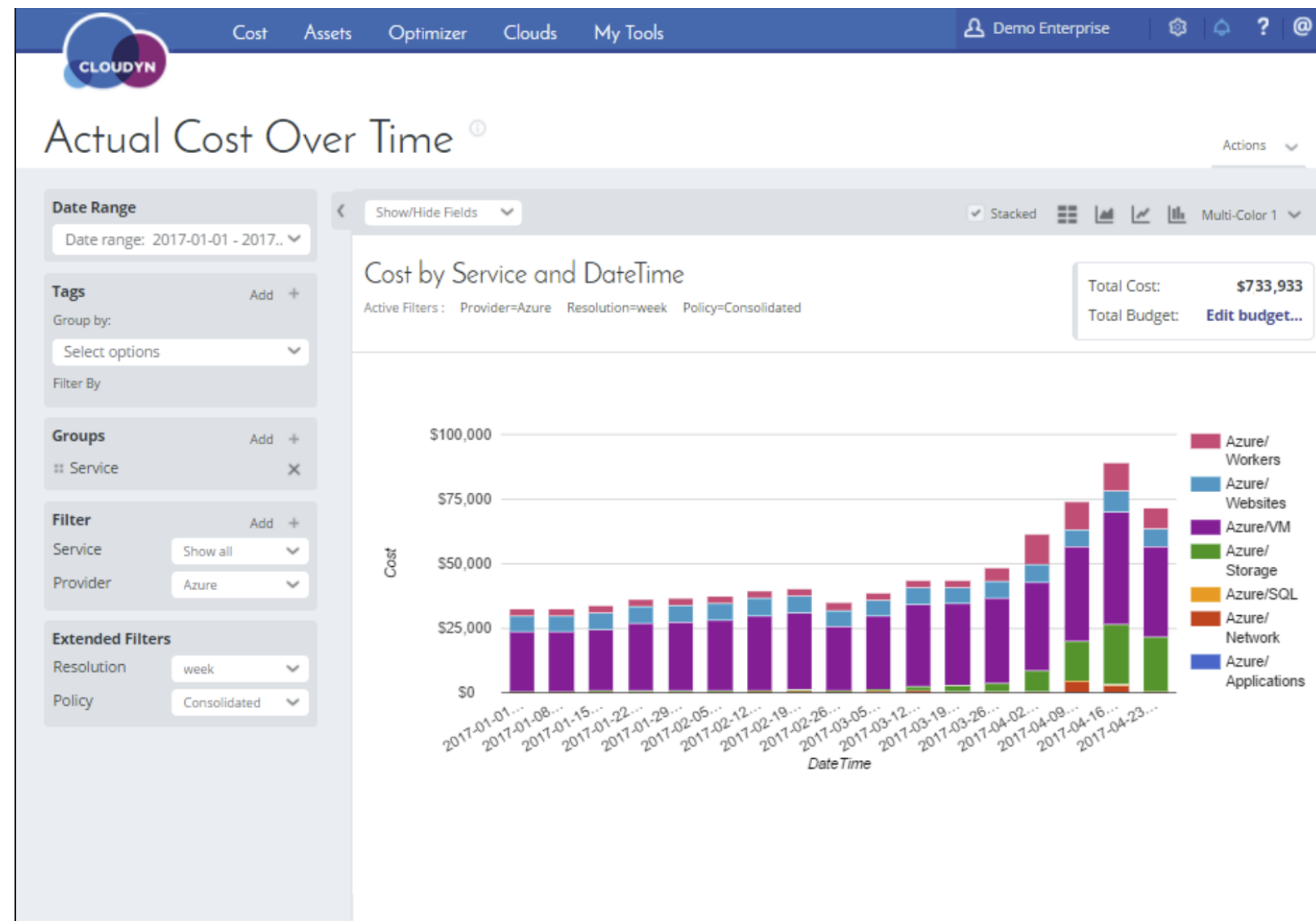
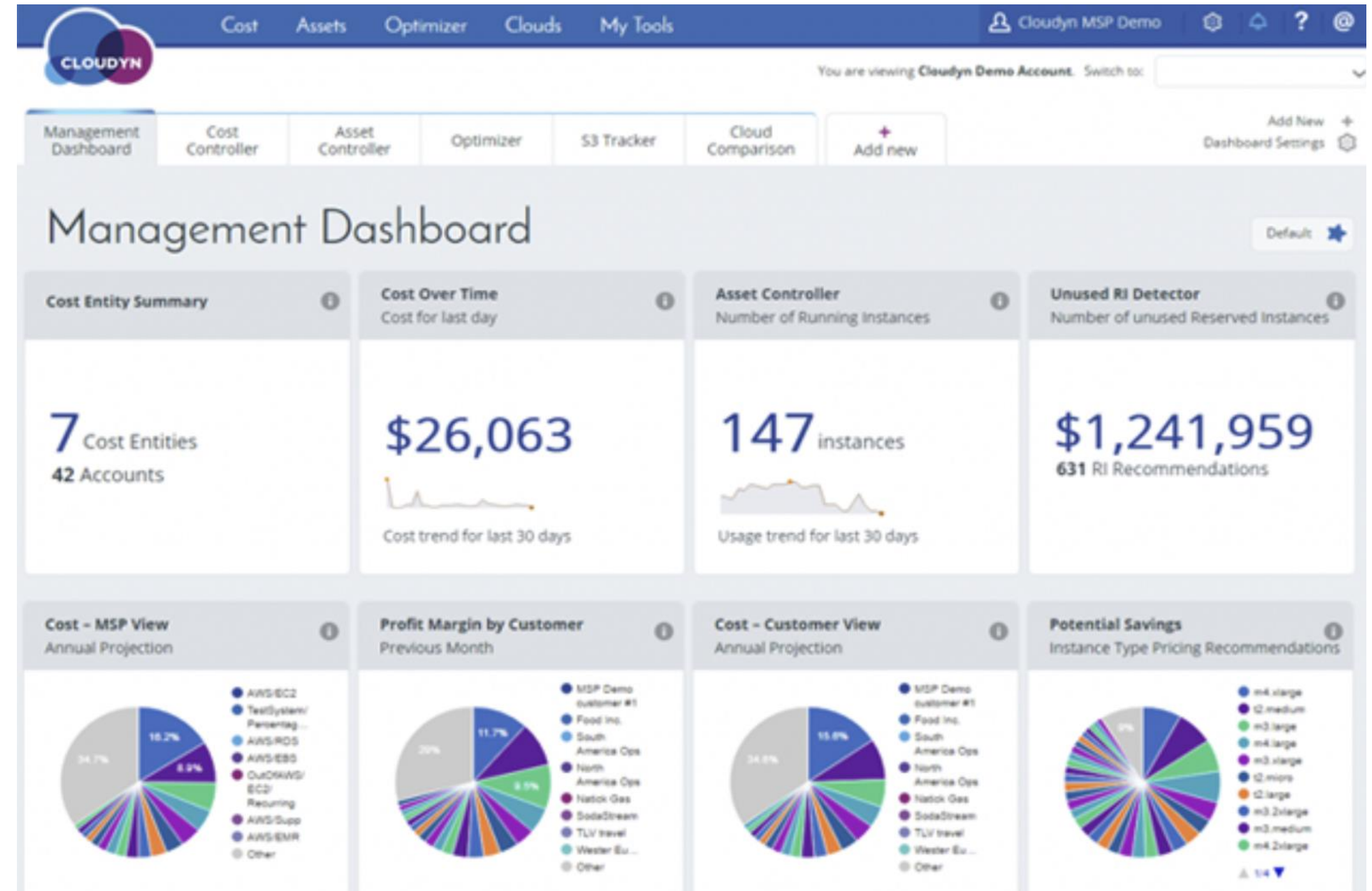
```
CREATE DATABASE BankingSQLSQLDW
COLLATE SQL_Latin1_General_CP1_CI_AS
(
    MAXSIZE = 250 GB
    , EDITION = 'datawarehouse'
    , SERVICE_OBJECTIVE = 'DW100'
);
```

Virtual Machine

- ✓ Pre-built images provision 1 TB by default

Azure Cost Management

A service offering in Azure from Clouduyn to monitor & manage cloud spend. Basic tier is free.



Azure Advisor

A 'best practices' advisor to optimize Azure deployments, including cost optimizations.

Microsoft Azure Advisor recommendations

Download as CSV Download as PDF Configure

Subscriptions: 2 of 24 selected – Don't see a subscription? Switch directories

2 subscriptions All types Active No grouping

Overview High Availability (8) Security (21) Performance (1) Cost (2) All (32)

Category	Recommendations	High impact	Medium impact	Low impact	Impacted resources
High Availability	8	1	7	0	25
Security	21	21	0	0	63
Performance	1	1	0	0	1
Cost	2	2	0	0	11

2,876 USD savings/mo *

Tips & tricks

- You can customize Advisor to process recommendations for resources that matter to you the most.
- You can optimize underutilized virtual machines to reduce your monthly Azure spend.
- You can improve the performance of your SQL Azure databases.
- You can enable virtual machine backup to protect your data from corruption or accidental deletion.

Download recommendations as PDF

Download recommendations as CSV

Demo

Auto-VM Shutdown +
SQLDW Pause +
SQLDB Scale

Tip #10

Understand Implications of
Storage Decisions for IaaS VMs

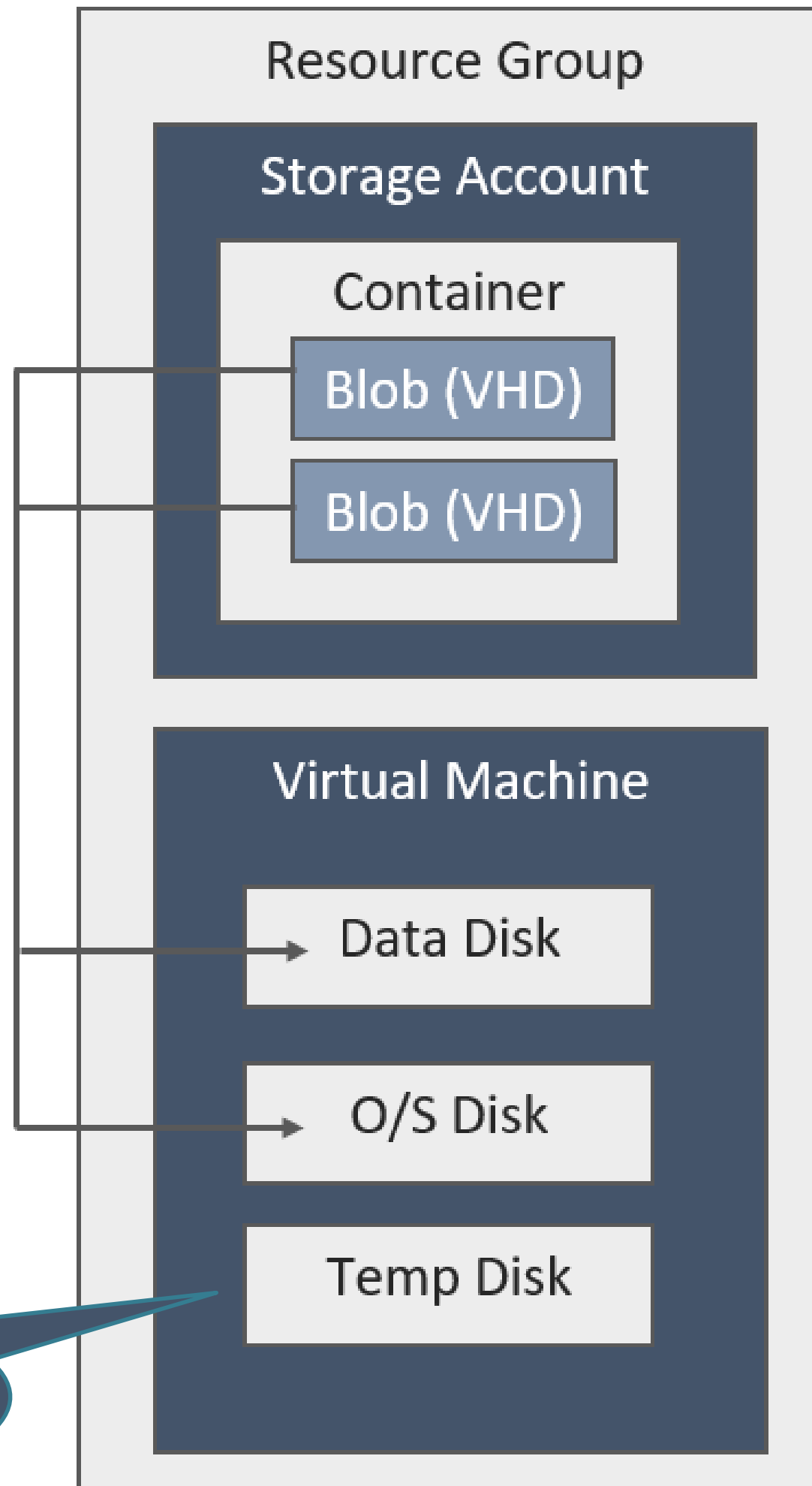
Storage for VMs

The VHDs (virtual hard disks) behind a VM are in Azure Storage



If you delete a VM in Azure, don't forget to also go delete the related storage (if appropriate)

Don't use the temp disk for anything persistent



DS12_V2 Standard	
4	Cores
28	GB
8	Data disks
12800	Max IOPS
56 GB	Local SSD
	Load balancing
	Premium disk support
485.09 USD/MONTH (ESTIMATED)	

of disks allowed depends on the VM size

The VM estimate doesn't include the O/S disk or data disk(s)

Storage Decisions Related to VMs

- 1 Managed or Unmanaged
- 2 Premium (SSD) or Standard (HDD)
- 3 Redundancy (Local, Zone, or Geo-Redundant)
- 4 Encryption (Storage Service Encryption and/or Disk Encryption)






Local is recommended per the Best Practices article

	Premium Disk	Standard Disk
Managed Disk	Billed for disk size reservation	Billed for disk size reservation
Unmanaged Disk	Billed for disk size reservation	Billed for space used

Can't currently reduce a disk size reservation – though it can be increased easily

Organizing Unmanaged Storage for VMs

All data for a single VM is in the same storage account – it is a unit of recovery in case of failure

NAME	RESOURCE GROUP	LOCATION	SKU	BILLINGCATEGORY (TAG)	ENVIRONMENTTYPE (TAG)	SUPPORTCONTACT (TAG)
INTERNALREPORTINGRGDEV						
 bisqlvm1 bck stdstrgdev	InternalReportingRGDev	East US	Standard_RAGRS	Internal Analytics	Dev	BI & Analytics Team
 bisqlvm1 data stdstrgdev	InternalReportingRGDev	East US	Standard_LRS	Internal Analytics	Dev	BI & Analytics Team
 bisqlvm1 diag stdstrgdev	InternalReportingRGDev	East US	Standard_LRS	Internal Analytics	Dev	BI & Analytics Team
SQLSKILLSWAITSLIBRARYRGDEV						
 sqlwaits bck stdstrgdev	SQLSkillsWaitsLibraryRGDev	East US	Standard_RAGRS	SQLSkills Waits Library	Dev	BI & Analytics Team
 sqlwaits diag stdstrgdev	SQLSkillsWaitsLibraryRGDev	East US	Standard_LRS	SQLSkills Waits Library	Dev	BI & Analytics Team

Backups (“bck”), Data, and Diagnostic logging (“diag”) are separated

Read Access Geo-Redundancy (RAGRS) for backup files

VM Disk Performance

Premium (SSD) – IOPS & Throughput are Associated with Disk Size:

	P4	P6	P10	P20	P30	P40	P50
Disk Size	32 GB	64 GB	128 GB	512 GB	1 TB	2 TB	4 TB
Price per month	\$4.81	\$9.29	\$17.92	\$66.56	\$122.88	\$235.52	\$450.56
IOPs per disk	120	240	500	2,300	5,000	7,500	7,500
Throughput per disk	25 MB/second	50 MB/second	100 MB/second	150 MB/second	200 MB/second	250 MB/second	250 MB/second

Standard (HDD) – IOPS & Throughput are Based on VM Size + Tier + # of Disks:

	S4	S6	S10	S15	S20	S30	S40	S50
Disk Size	32 GB	64 GB	128 GB	256 GB	512 GB	1 TB	2 TB	4 TB
Price per month (promotional)	\$0.77	\$1.51	\$2.95	\$5.67	\$10.88	\$20.48	\$40.96	\$81.92

VM Tier	Basic Tier VM	Standard Tier VM
Max Disk size	4095 GB	4095 GB
Max 8 KB IOPS per disk	Up to 300	Up to 500
Max Bandwidth per disk	Up to 60 MB/s	Up to 60 MB/s

<https://docs.microsoft.com/en-us/azure/storage/common/storage-premium-storage>

<https://docs.microsoft.com/en-us/azure/storage/common/storage-standard-storage>

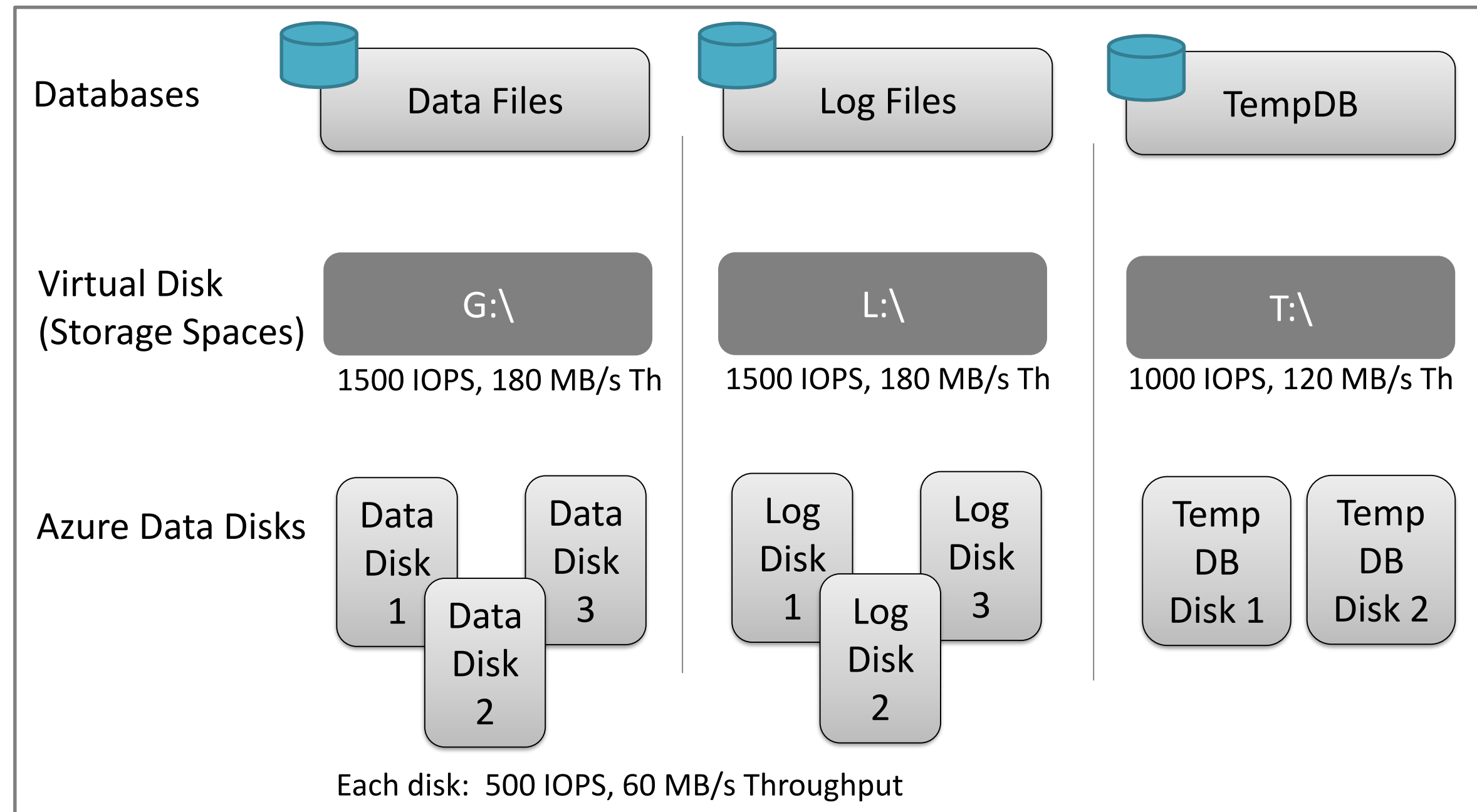
Increasing Performance of Standard Disks

(1/3)

Storage Spaces (aka Virtual Disks) are multiple Azure data disks in a storage pool to collectively share throughput and IOPS limits.

This aggregation of disks is helpful due to lower Standard limits on IOPS & throughput.

(Another alternative to increase disk performance is to spread separate database files on different disks.)



Adapted from:

<https://blogs.msdn.microsoft.com/sqlcat/2013/06/17/performance-guidance-for-sql-server-in-windows-azure-virtual-machines/>

IOPS = # of requests per second (OLTP workloads)

Throughput (Bandwidth) = IOPS x I/O size per specified interval (DW workloads)

Increasing Performance of Standard Disks

(2/3)

Virtual machine: BISQLVM1Dev - Disks

OS disk: BISQLVM1Dev

LUN	NAME	SIZE	STORAGE ACCOUNT TYPE	ENCRYPTION	HOST CACHING
1	bisqlvm1datadisk1dev	550 GIB	Standard_LRS	Not enabled	Read-only
2	bisqlvm1datadisk2dev	550 GIB	Standard_LRS	Not enabled	Read-only
3	bisqlvm1datadisk3dev	550 GIB	Standard_LRS	Not enabled	Read-only
4	bisqlvm1logdisk1dev	550 GIB	Standard_LRS	Not enabled	None
5	bisqlvm1logdisk2dev	550 GIB	Standard_LRS	Not enabled	None
6	bisqlvm1logdisk3dev	550 GIB	Standard_LRS	Not enabled	None
7	bisqlvm1tempdbdisk1dev	550 GIB	Standard_LRS	Not enabled	Read-only
8	bisqlvm1tempdbdisk2dev	550 GIB	Standard_LRS	Not enabled	Read-only

Disks are attached to the VM. A red box highlights the '+ Add data disk' button in the bottom right of the VM configuration pane.

Before disk encryption has been configured

Virtual machine: bisqlvm1datadisk1dev (unmanaged)

NAME: bisqlvm1datadisk1dev

LOGICAL UNIT NUMBER (LUN): 1

STORAGE ACCOUNT TYPE: Standard_LRS

Size (GiB): 550

ESTIMATED PERFORMANCE

IOPS limit	500
Throughput limit (MB/s)	60

SOURCE IMAGE: MicrosoftSQLServer / SQL2016SP1-WS2016 / SQLDEV / latest

Host caching: Read-only

VHD URI: https://bisqlvm1datastdstgdev.blob.core.windows.net/vhds/bisqlvm1datadisk1dev.vhd

Note the single disk performance before it's collectively shared

Microsoft Azure Storage accounts > bisqlvm1datastdstgdev > Blob service > vhds

Container: vhds

Essentials

NAME: vhds

Location: vhds

Search blobs by prefix (case-sensitive)

- bisqlvm1datadisk1dev.vhd
- bisqlvm1datadisk2dev.vhd
- bisqlvm1datadisk3dev.vhd
- BISQLVM1Dev20170521150516.vhd
- bisqlvm1logdisk1dev.vhd
- bisqlvm1logdisk2dev.vhd
- bisqlvm1logdisk3dev.vhd
- bisqlvm1tempdbdisk1dev.vhd
- bisqlvm1tempdbdisk2dev.vhd

Each blob has a lease which associates it with the VM

+ Add data disk

Disks attached to the VM ↑

← Blob storage container

Increasing Performance of Standard Disks

(3/3)

This is a tradeoff: more complexity for greater performance at lower cost

The screenshot shows the Server Manager interface for Storage Pools. The 'Storage Pools' section is active, displaying a table of storage pools. A red box highlights the 'Windows Storage (3)' group, which includes 'DataDisksStoragePool', 'LogDisksStoragePool', and 'TempDBDisksStoragePool'. Below this, the 'VIRTUAL DISKS' section for 'DataDisksStoragePool on BISQLVM1Dev' is shown, with a table listing 'DataVirtualDisks' with a status of 'Simple' and 'Fixed', and a capacity of 1.61 TB. To the left, a 'DataVirtualDisk' details pane shows the 'Property' set to 'Interleave' and the 'Value' set to '262144'. A blue callout bubble points to the 'Value' field with the text: 'Stripe size should be 256KB for a data warehousing workload, or 64KB for OLTP workload'. To the right, the 'PHYSICAL DISKS' section is visible, showing three 'Msft Virtual Disk (BISQLVM1Dev)' entries with a capacity of 550 GB each.

Name	Type	Managed by	Available to	Read-Write Server	Capacity	Free Space	Percent Allocated
Windows Storage (3)							
DataDisksStoragePool	Storage Pool	BISQLVM1Dev	BISQLVM1Dev	BISQLVM1Dev	1.61 TB	0.00 B	
LogDisksStoragePool	Storage Pool	BISQLVM1Dev	BISQLVM1Dev	BISQLVM1Dev	1.61 TB	0.00 B	
TempDBDisksStoragePool	Storage Pool	BISQLVM1Dev	BISQLVM1Dev	BISQLVM1Dev	1.07 TB	0.00 B	

Name	Status	Layout	Provisioning	Capacity	Allocated	Volume
DataVirtualDisks	Simple	Fixed		1.61 TB	1.61 TB	G:

Slot	Name	Status	Capacity	Bus	Usage	Char
	Msft Virtual Disk (BISQLVM1Dev)		550 GB	SAS	Automatic	Inte
	Msft Virtual Disk (BISQLVM1Dev)		550 GB	SAS	Automatic	Inte
	Msft Virtual Disk (BISQLVM1Dev)		550 GB	SAS	Automatic	Inte

Stripe size should be 256KB for a data warehousing workload, or 64KB for OLTP workload

Demo

Azure Storage Account

Tip #11

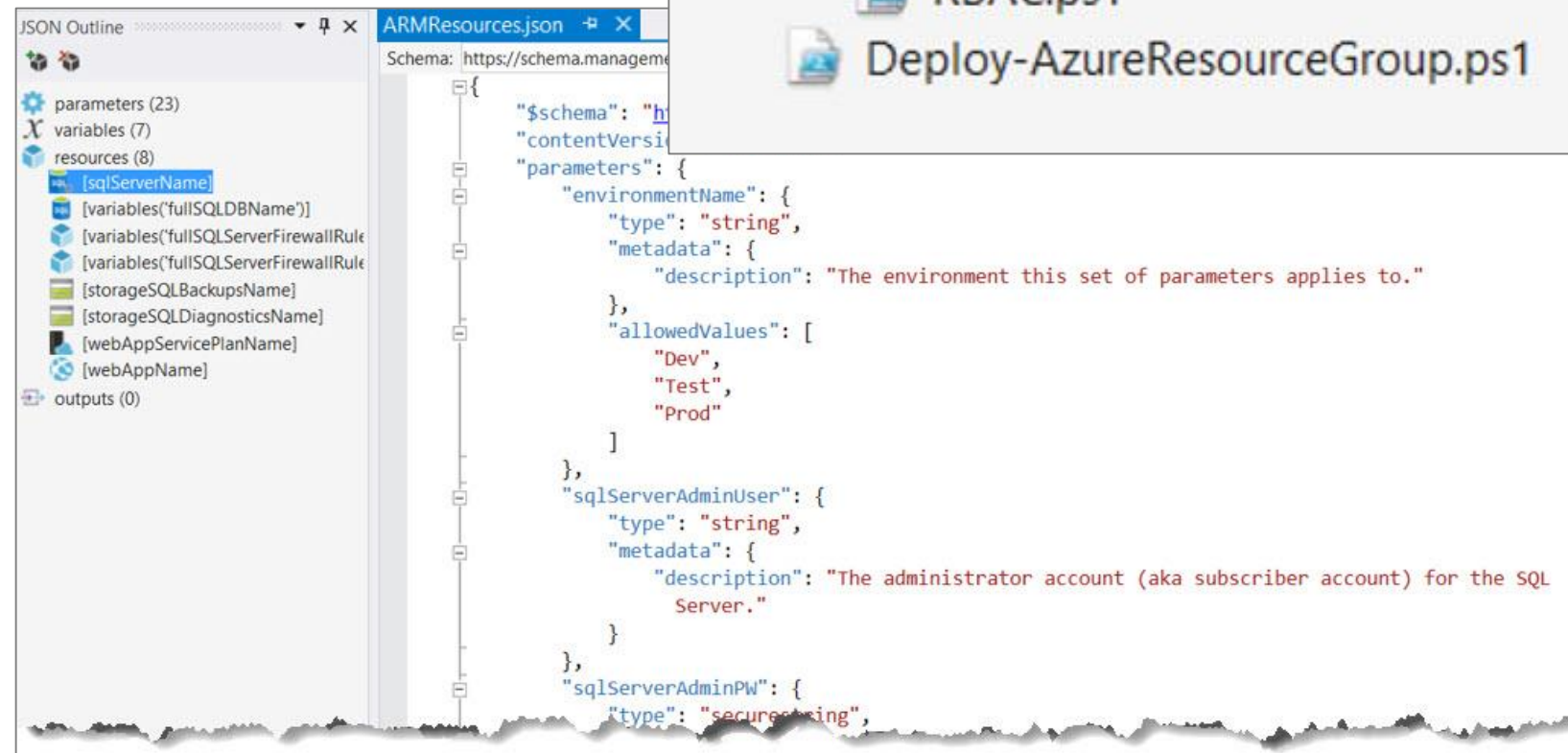
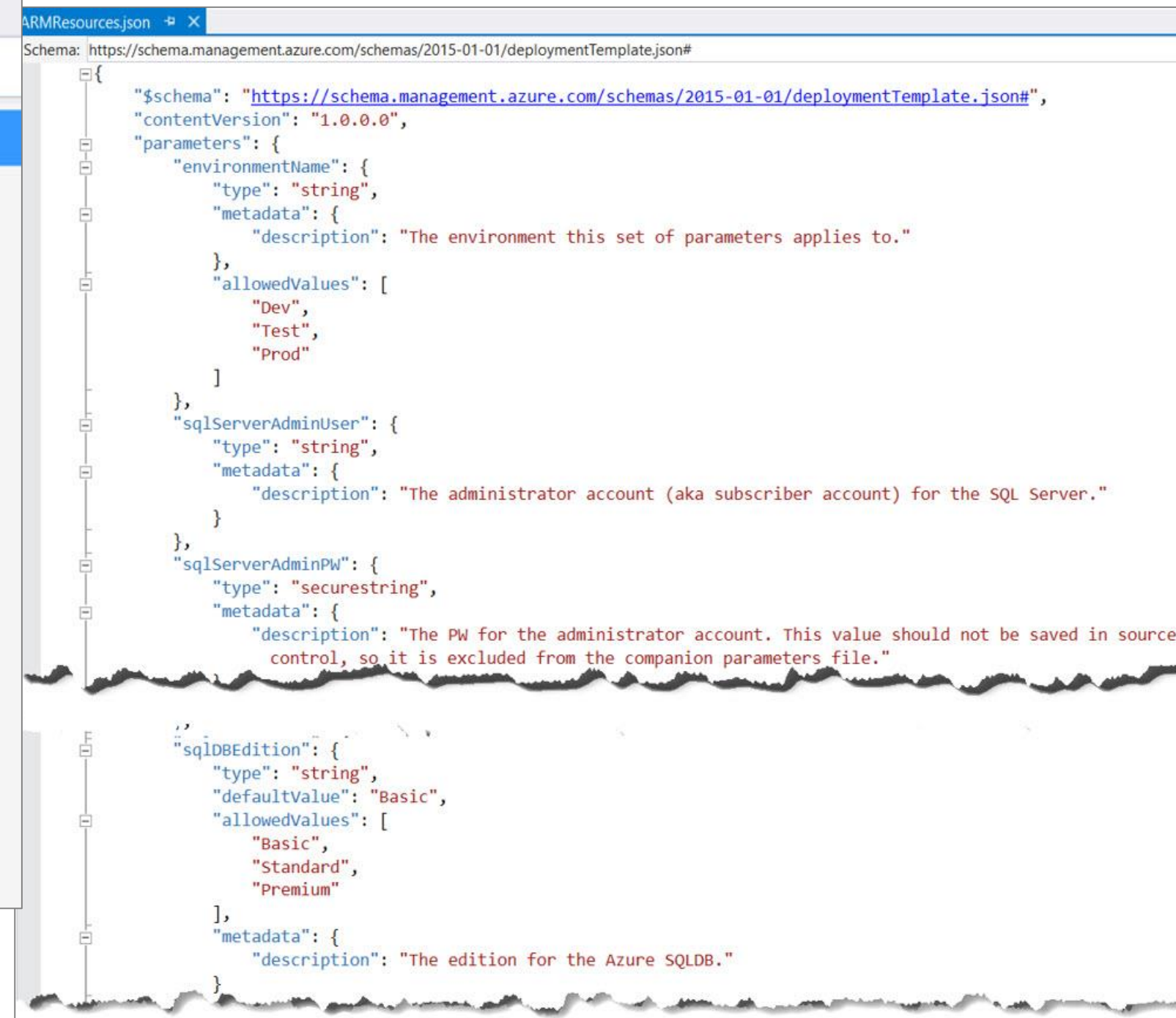
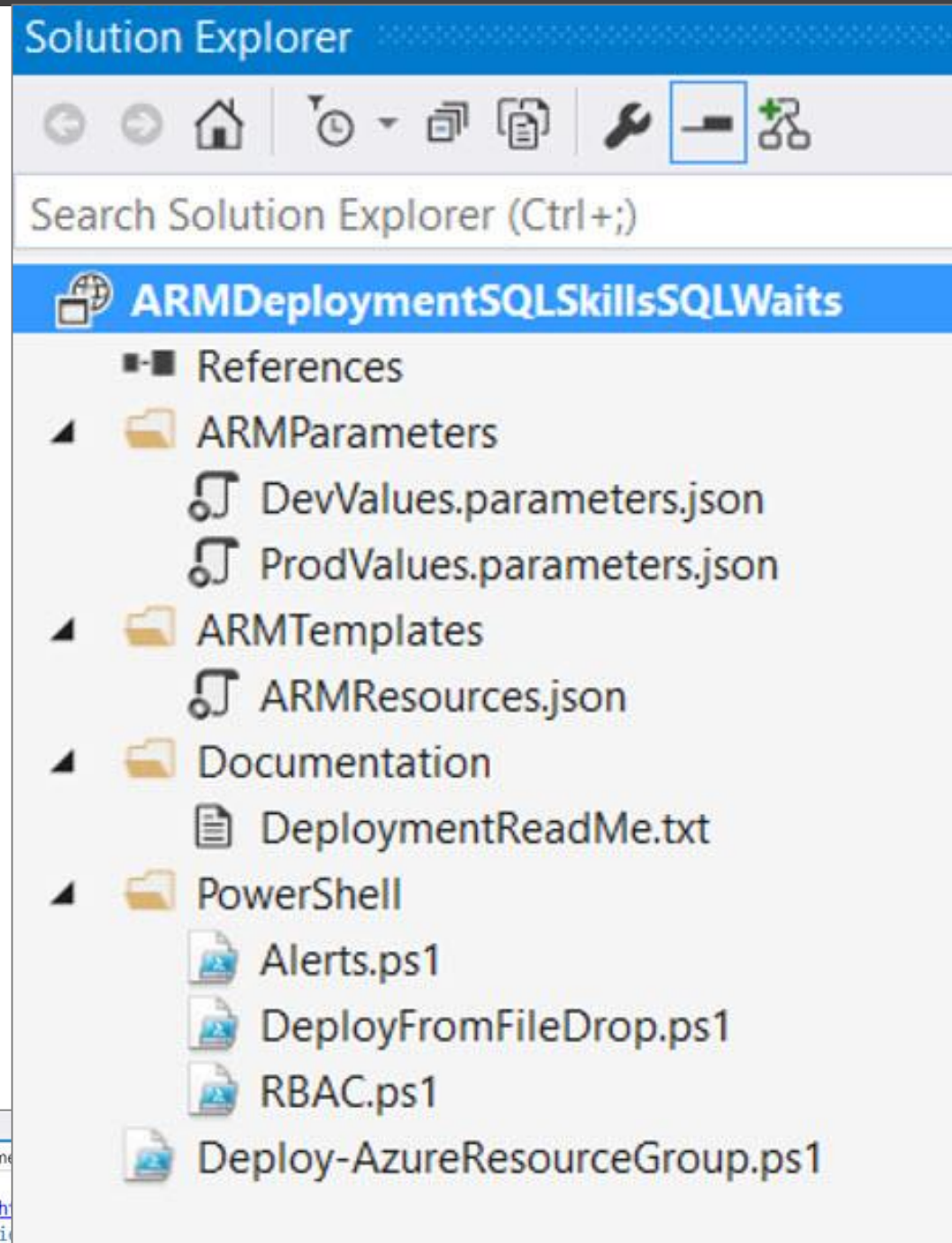
Know When to Leverage ARM
and/or PowerShell

ARM Templates

Goal is for deployments to be modular & repeatable.

Invest time in learning how to do ARM templates so you can make a good decision when to use them.

This is known as “**infrastructure as code**” which can be checked into source control from a Visual Studio project.

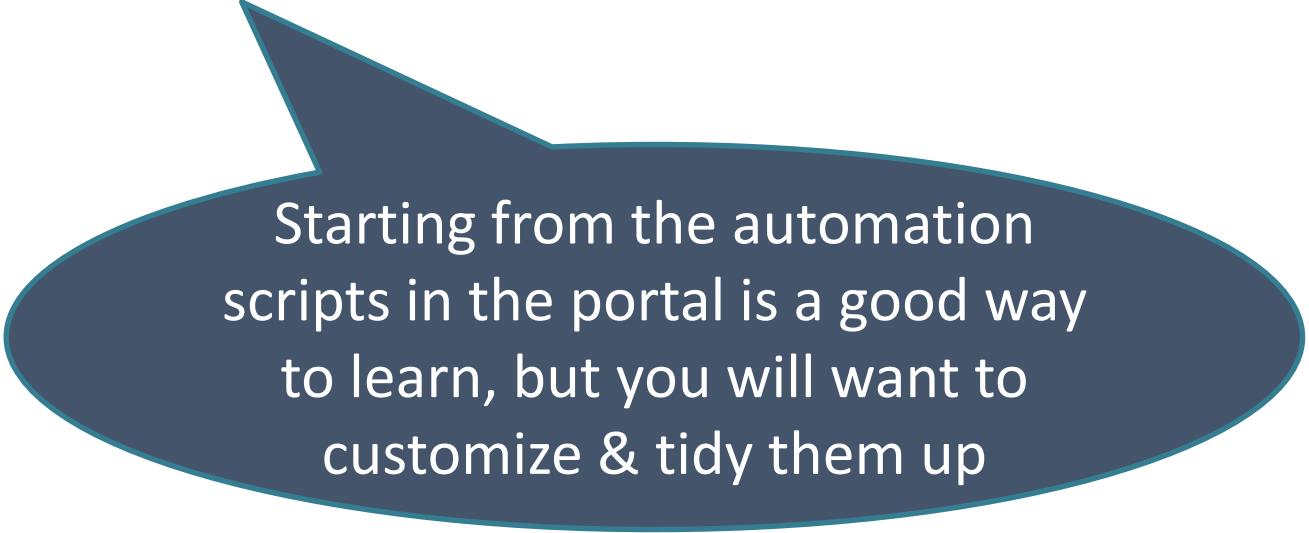


Blog post on ARM deployments:
<https://blogs.sentryone.com/melissacoates/getting-started-with-automated-arm-deployments-in-azure/>

ARM Templates

Use ARM if You Intend To...

- ✓ Include the configuration of Azure resources in **source control** ("Infrastructure as Code"), and/or
- ✓ **Repeat the deployment process** numerous times, and/or
- ✓ **Automate** deployments, and/or
- ✓ Employ **continuous integration** techniques, and/or
- ✓ Utilize **DevOps** principles and practices, and/or
- ✓ Repeatedly provision & de-provision **testing infrastructure**
- ✓ Do something **unsupported** by the portal interface



Starting from the automation scripts in the portal is a good way to learn, but you will want to customize & tidy them up

PowerShell



Automation is a balance of:

Wanting infrastructure in source code vs. little need for deployment repeatability

PowerShell scripts can be useful for:

- Deploying ARM templates
- Tasks unsupported in the portal
- Deploying specific items
 - Azure Data Factory
 - AAD service principal
 - etc...
- Managing resources
 - Assigning tags
 - Enabling disk encryption
 - Provisioning key vault
 - Defining alerts
 - Start/stop a virtual machine

```
#Input Area
$subscriptionName = 'InsertSubscriptionName'
$resourceGroupName = 'SQLSkillsWaitsLibraryRGDev'
$deploymentName = 'SQLSkillsSQLWaitsDeployment'
$templateFilePath = 'O:\Deployments\SQLWaits\2017-06-07\ARMResources.json'
$parameterFilePath = 'O:\Deployments\SQLWaits\2017-06-07\DevValues.parameters.json'

#ARM template deployment
New-AzureRmResourceGroupDeployment `
  -Name $deploymentName `
  -ResourceGroupName $resourceGroupName `
  -TemplateFile $templateFilePath `
  -TemplateParameterFile $parameterFilePath
```

From:

<https://blogs.sentryone.com/melissacoates/getting-started-with-automated-arm-deployments-in-azure/>

Recommendations for ARM and PowerShell

Use ARM for Provisioning Azure Resources When...

- Recognition of **inter-dependencies** is important
- Speed of deployment is important - ARM is **faster** than PowerShell or CLI (due to **parallel deployment**)

Use PowerShell For...

- Execution of **ARM template**
- Execution of **misc. or piecemeal deployments**
- **Management** of resources (ex: start/stop a VM) often in conjunction with **Azure Automation**
- Execution of action, resource deployment, or configuration **not supported by ARM**

Demo

Automation Script in Azure Portal
+
Visual Studio ARM Project

Key Takeaways

The **pace of change** in Azure is frequent—plan to keep up with announcements. New features can change a previous decision really fast. New services and features are constant.

Look at **using PaaS** unless you really need an IaaS solution – i.e., the simplest solution that gets the job done. And expect complications, delays, and a learning curve along the way.

Don't shortcut the initial planning phase – do a **POC** if you can to learn. Don't assume a feature is definitely available or works as you expect it to. For unfamiliar features or services, consider creating a **sandbox** area to **test & learn** on first.

Plan to invest some time learning **ARM templates & PowerShell**, if you haven't already.

Pay attention to **opportunities for cost savings**. Developers & administrators have a huge impact on cost in the cloud.

Resources

Conceptual documentation: <https://azure.microsoft.com/>

Reference documentation: <http://msdn.microsoft.com/>

Documentation: <https://docs.microsoft.com/en-us/azure/>

Samples: <http://github.com/azure-samples>

Templates: <https://github.com/Azure/azure-quickstart-templates>

Azure Architecture Center: <https://docs.microsoft.com/en-us/azure/architecture/>