Configuration and Optimization

– Azure Synapse Service

Eshant Garg
Data Engineer, Architect, Advisor
eshant.garg@gmail.com

Azure Synapse Analytics

Introduction

Configuration settings

Common tasks – Big data analytics

Backup and Restore DW

Price optimization

Managing Workload

Monitoring

Support and Troubleshooting

Delete all resources

Azure Synapse Analytics

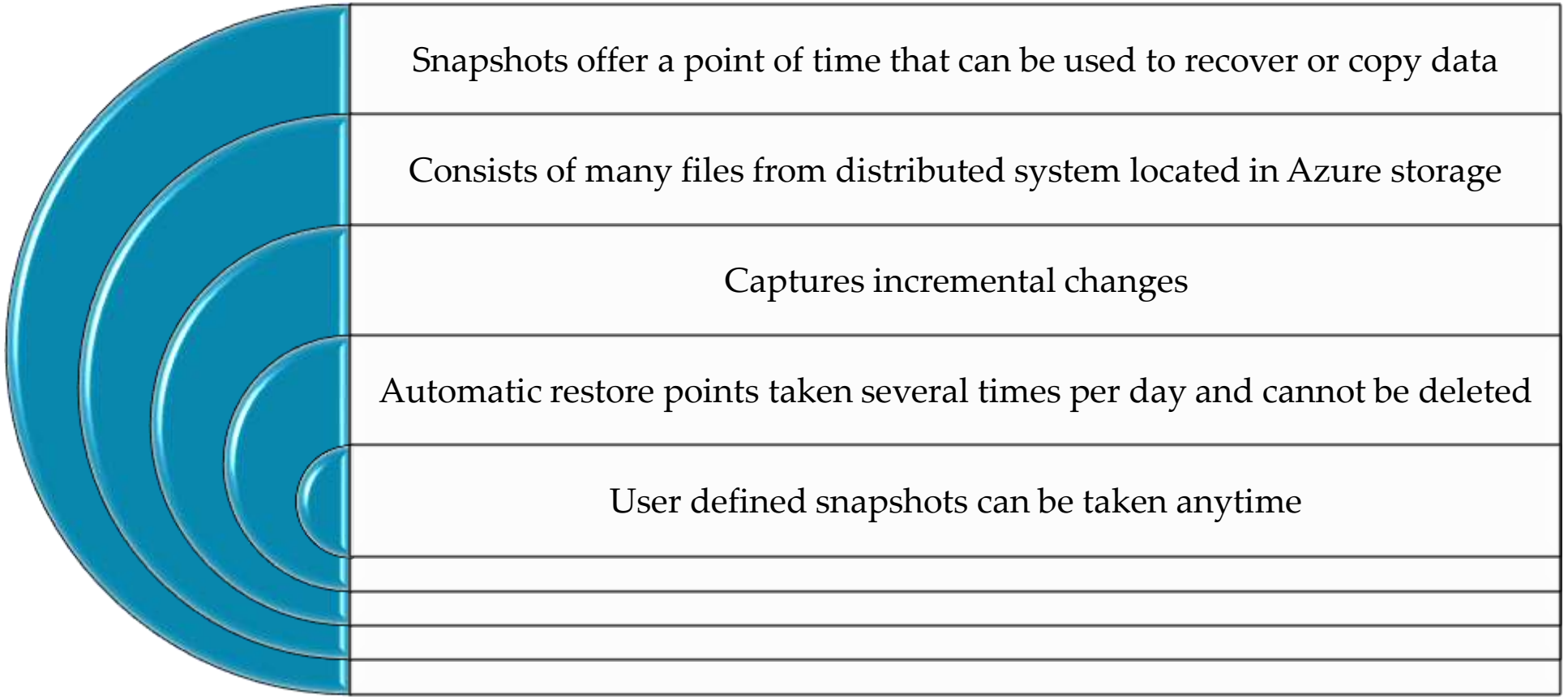
Examining Configuration options for an SQL pool in Azure Synapse Analytics

- Overview of Data Warehouse interface
- How to Scale
- How to view Activities
- Apply Tags and Locks
- Various other settings

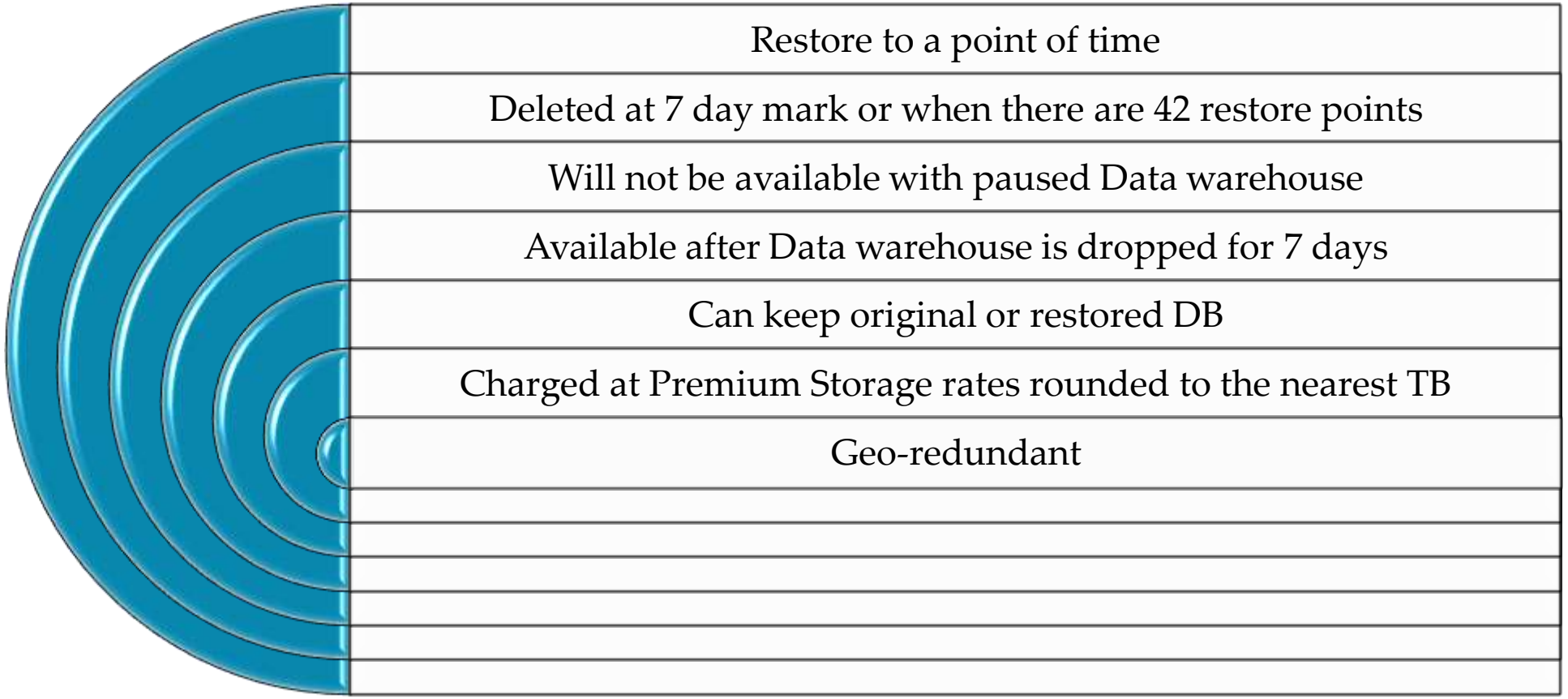
Performing Common Tasks with Azure Synapse Analytics in the Portal

- Streaming Analytics
- Load Data
- Use the Query Editor
- Build Reports
- Perform Analytics

Backup


Snapshots offer a point of time that can be used to recover or copy data
Consists of many files from distributed system located in Azure storage
Captures incremental changes
Automatic restore points taken several times per day and cannot be deleted
User defined snapshots can be taken anytime

Restore


Restore to a point of time
Deleted at 7 day mark or when there are 42 restore points
Will not be available with paused Data warehouse
Available after Data warehouse is dropped for 7 days
Can keep original or restored DB
Charged at Premium Storage rates rounded to the nearest TB
Geo-redundant

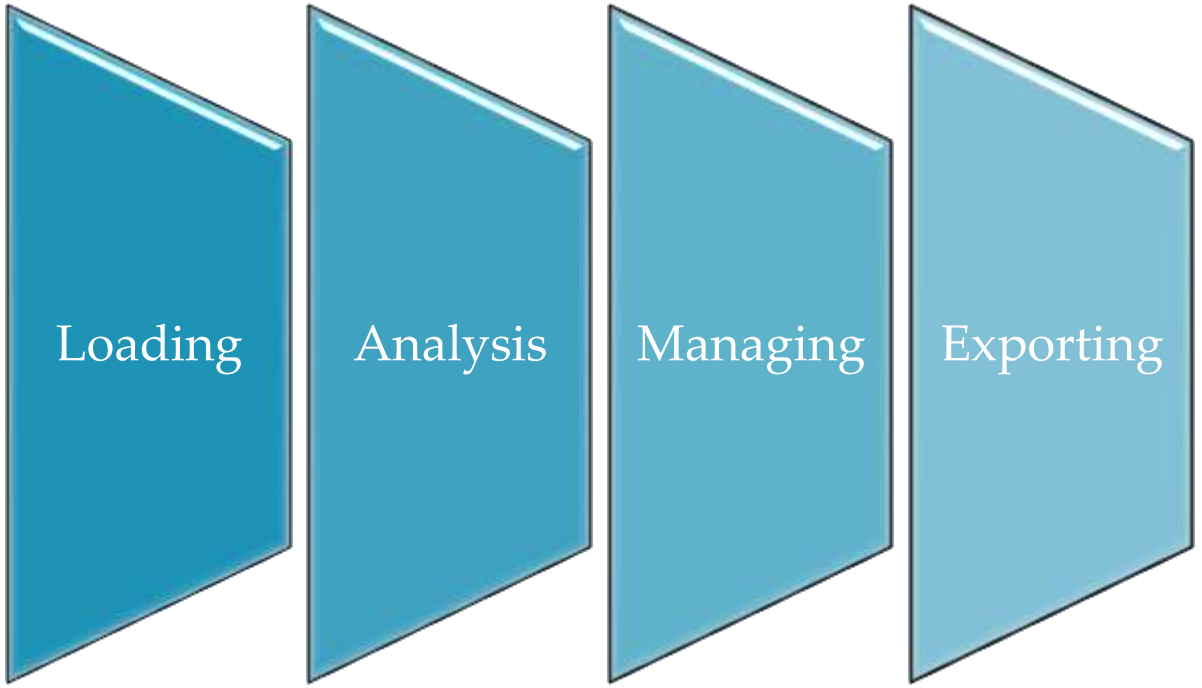
Price Optimization

Managing workloads

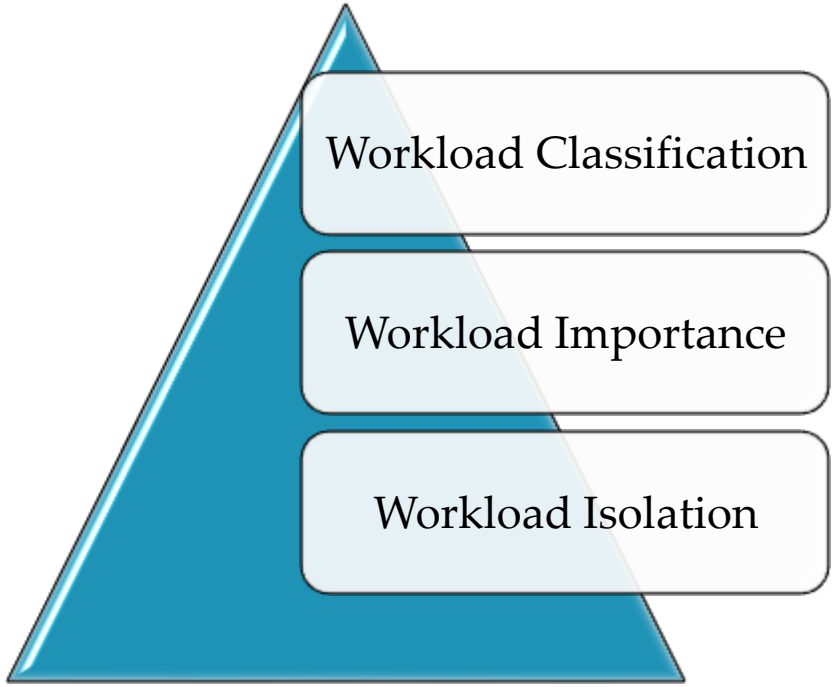
Workload Management ensures that enough resources exist for data warehousing activities like loading, transforming or querying data.

A successful workload management scheme effectively manages resources, ensures highly efficient resource utilization, and maximizes return on investment (ROI).

Data warehouse workloads



Data warehouse workloads



Workload Classification

- Assign users to a role that has a corresponding resource class
- CREATE WORKLOAD CLASSIFIER
- A classification that is based on a database user will take precedence over a role membership

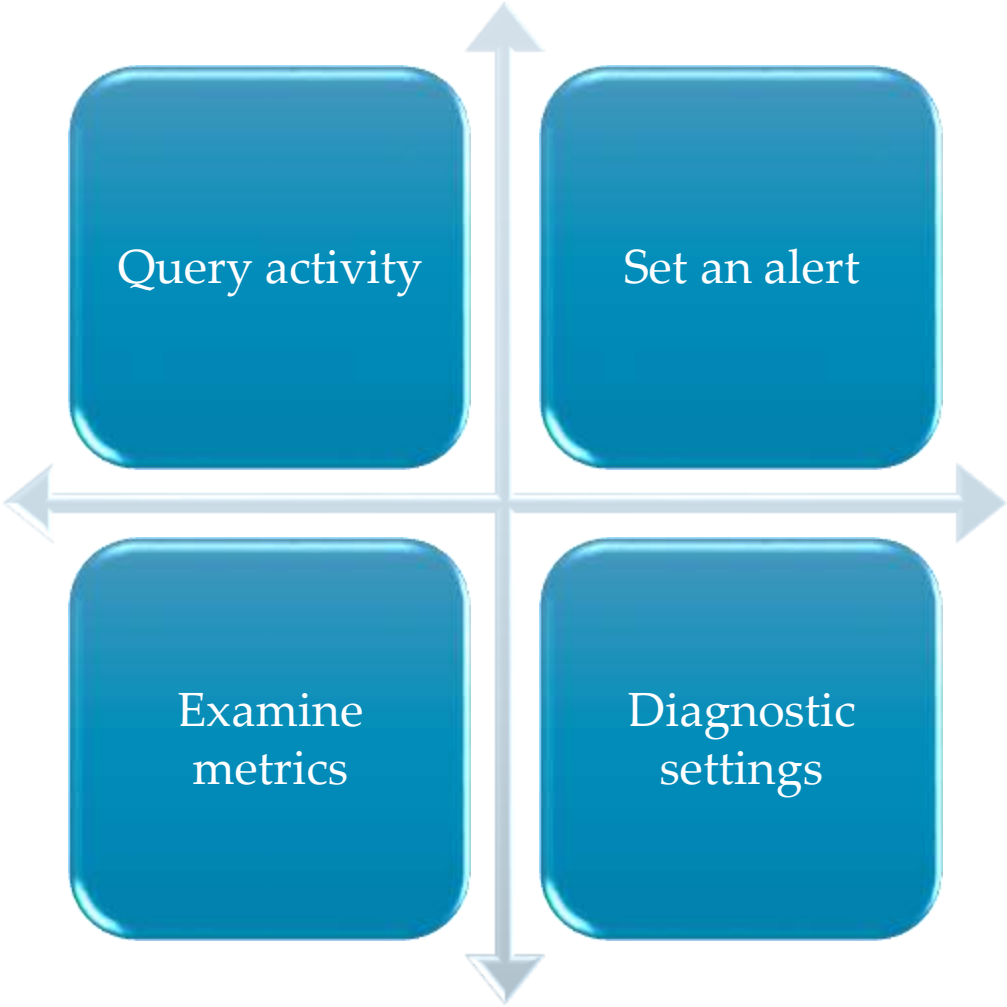
Workload Importance

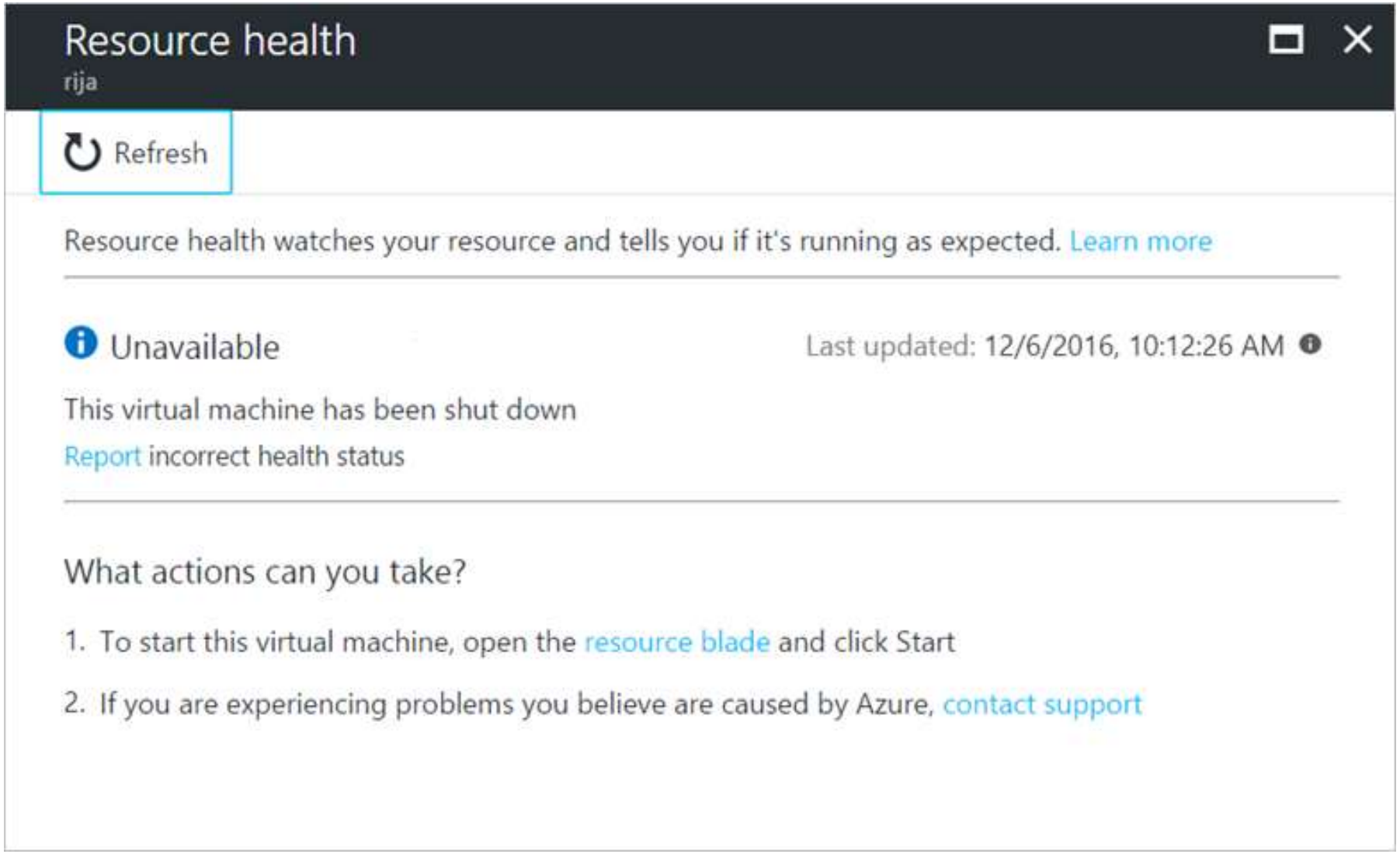
- Workload importance influences the order in which a request gets access to resources.
- 5 level of importance – Low, below_normal, normal, above_normal and high
- A request with higher importance will be run before a request with lower importance
- Can prevent lock contention

Workload Isolation

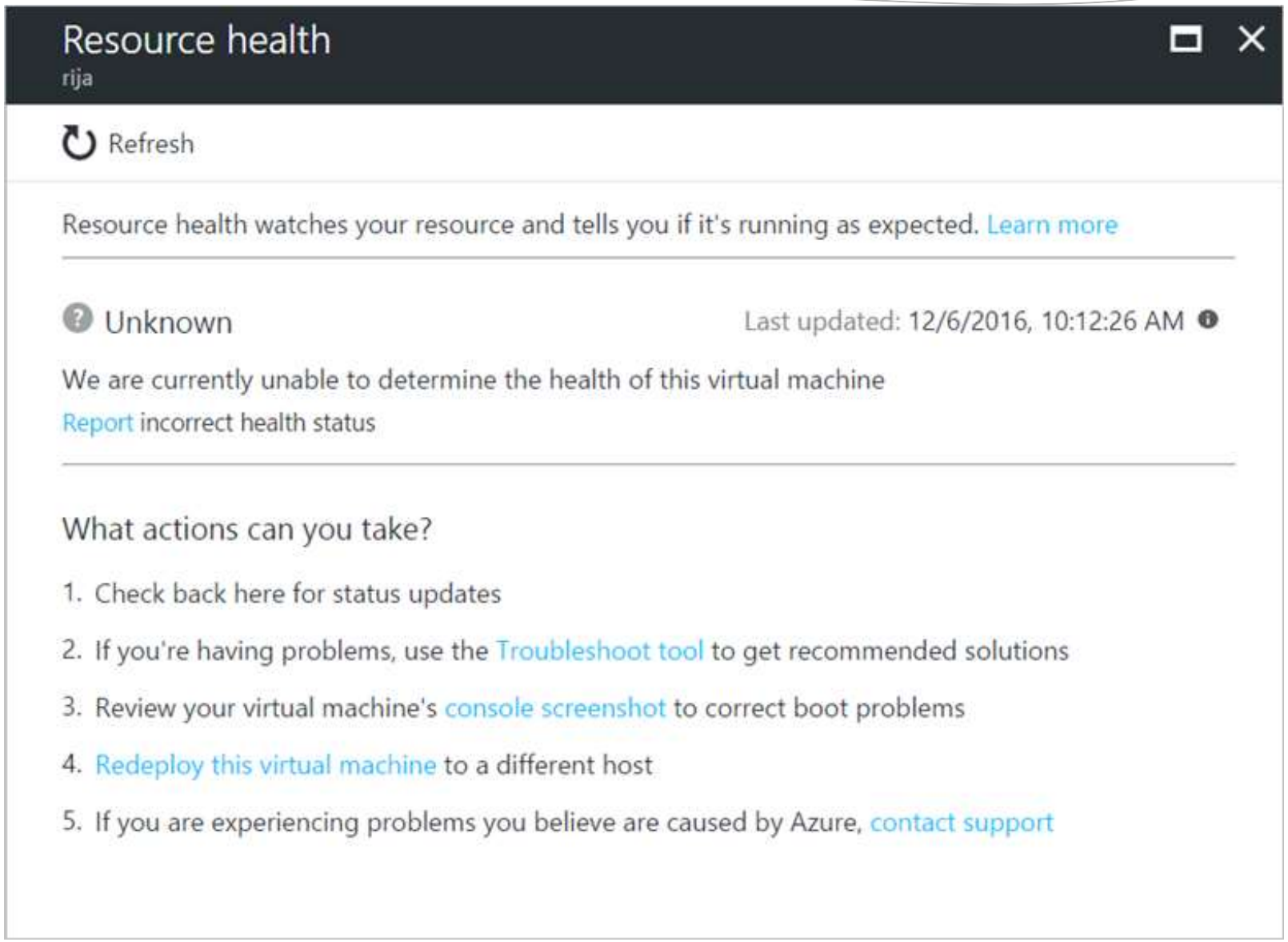
- Workload isolation reserves resources for a workload group and use workload groups to define isolation
- Guarantees implicit level of concurrency with the MIN_PERCENTAGE_RESOURCE parameter
- Define a maximum level of concurrency with CAP_PERCENTAGE_RESOURCE parameter

Monitoring

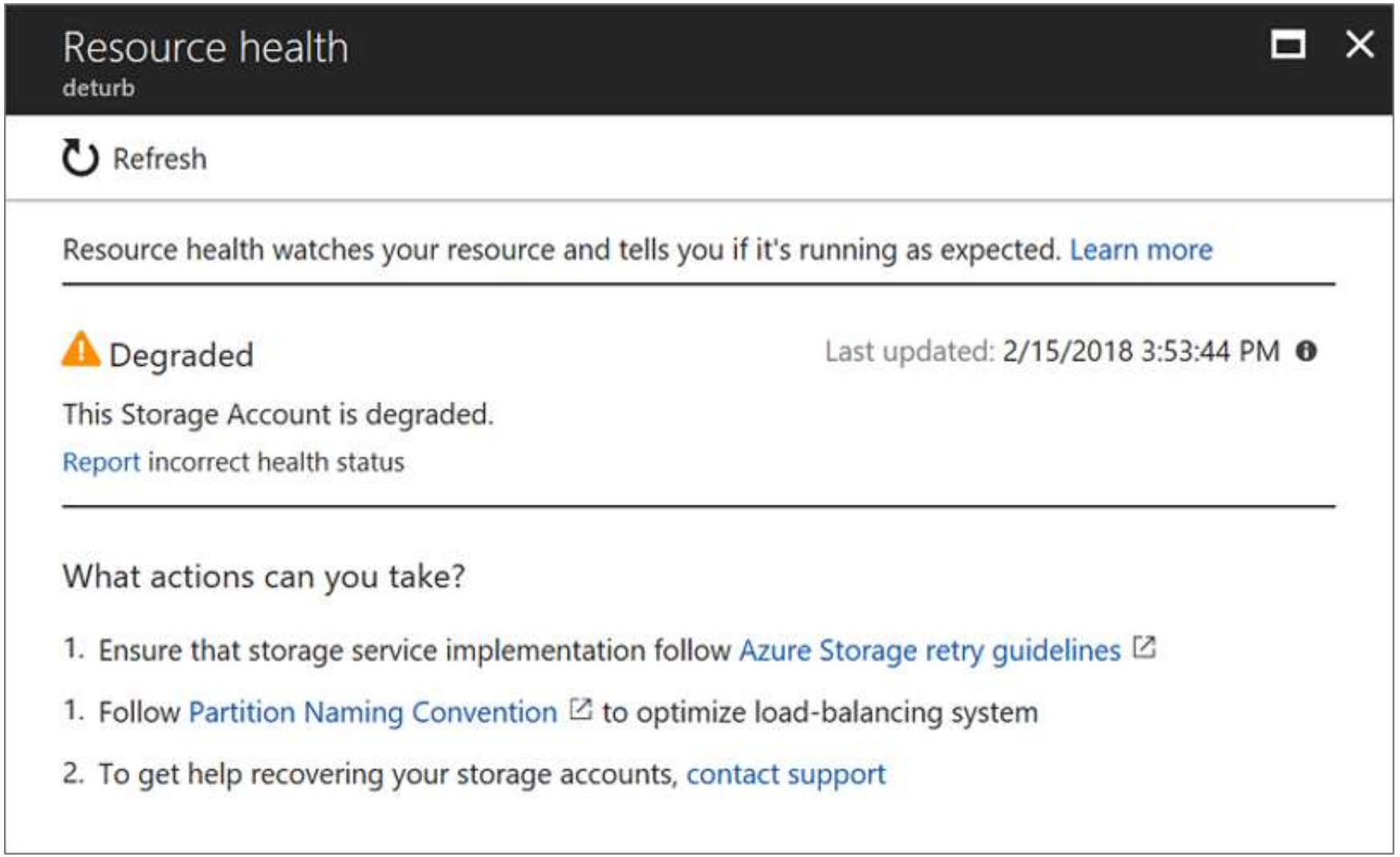

Query activity
Set an alert
Examine metrics
Diagnostic settings



Source: Microsoft



Source: Microsoft



Source: Microsoft

Summary

Configuration settings

Common tasks – Big data analytics

Backup and Restore DW

Workload management

Monitoring

Support and Troubleshooting

Delete all resources