

Be My Azure DBA

DBA or DSA?



Paul Andrew | Senior Consultant

07/09/2018

 @MrPaulAndrew



Gold Data Analytics
Gold Data Platform
Gold Cloud Platform





<https://github.com/mrpaulandrew>

CommunityEvents

Demo code, content and slides from various community events.

● C++

[{Event}-{Location}-{Year}](#)

Disclaimer:

Blunt Englishmen with a dry
sense of humour.

Agenda

Context

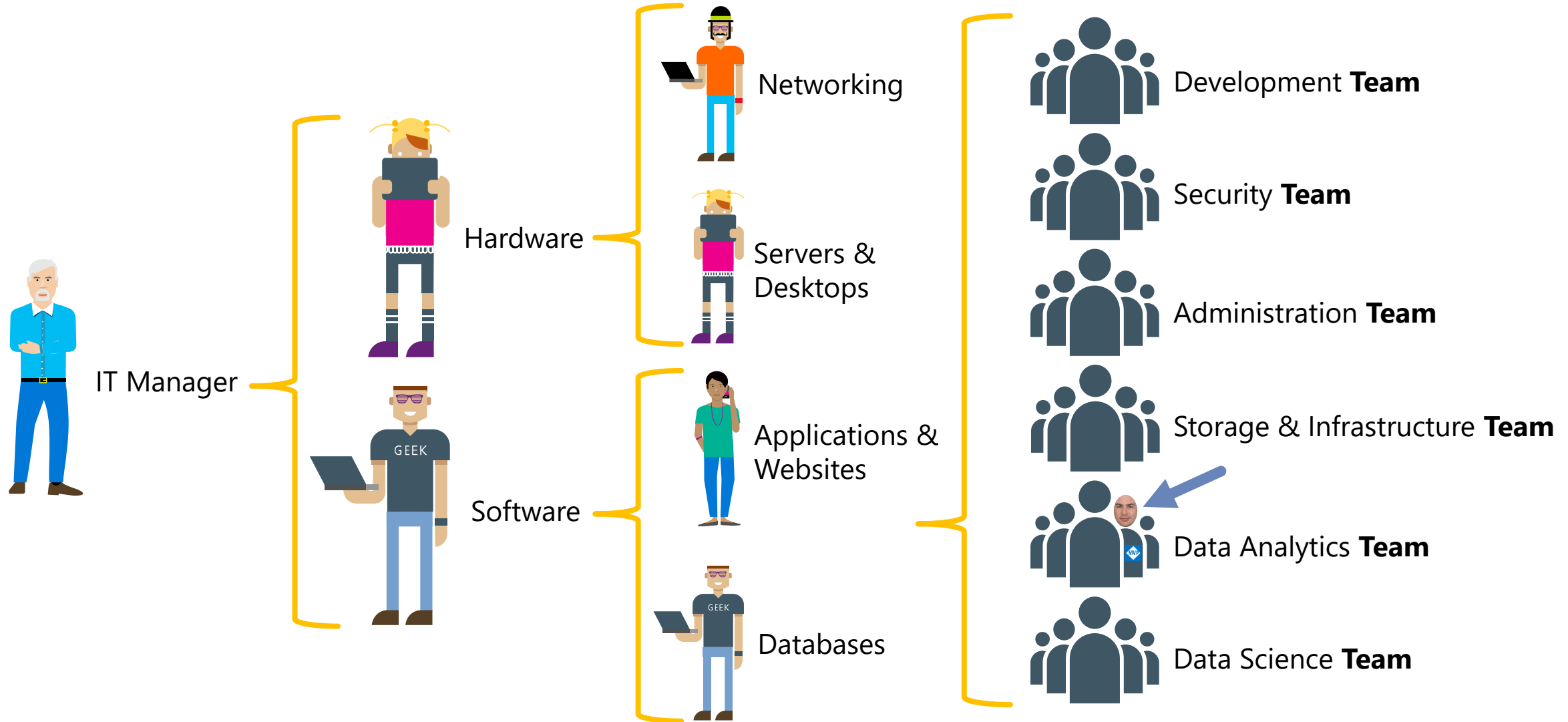
IT Industry Roles
Knowledge

The Modern Data Warehouse

Components
Considerations

Conclusions

IT Industry Role Evolution

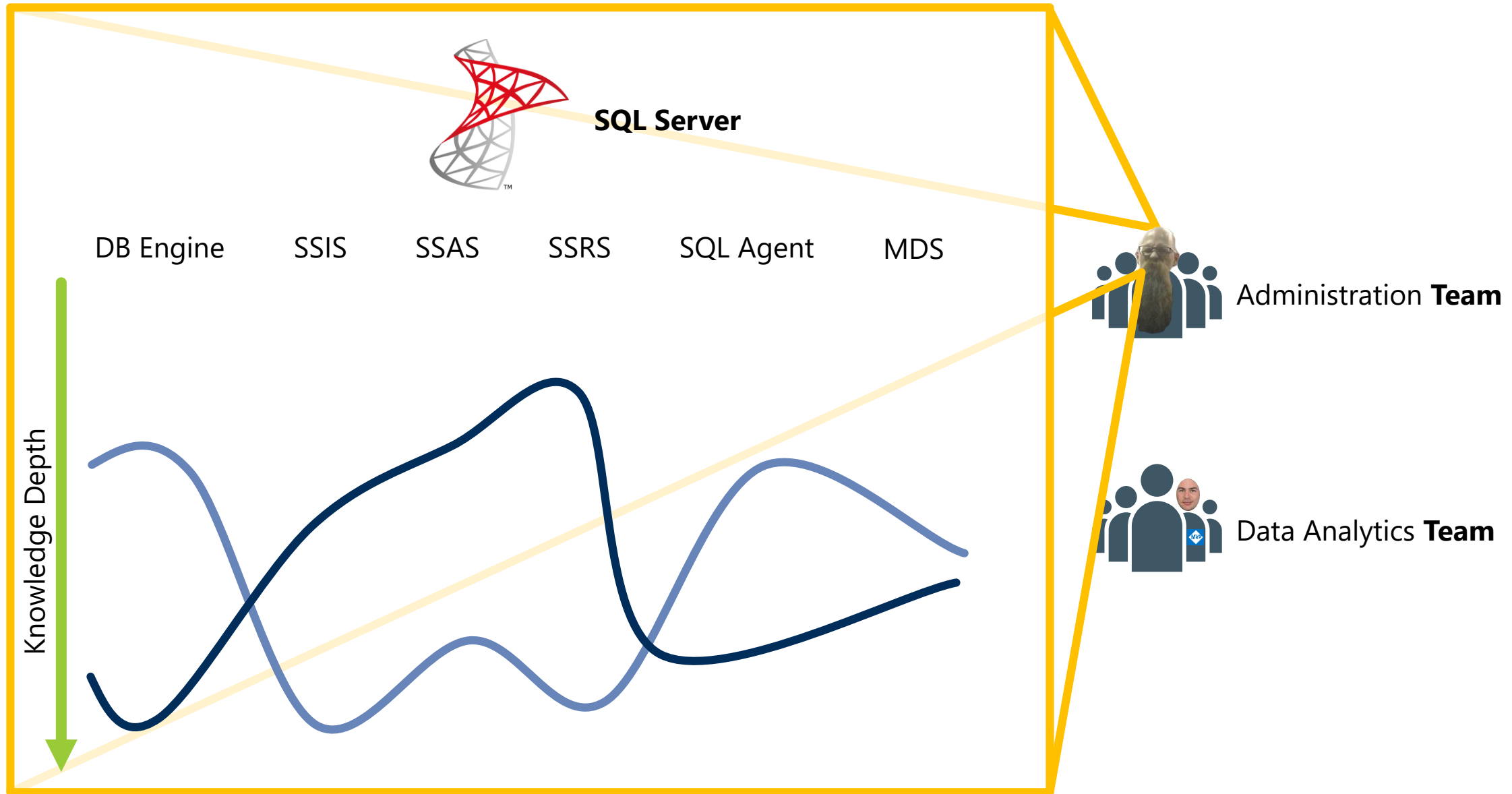


Time

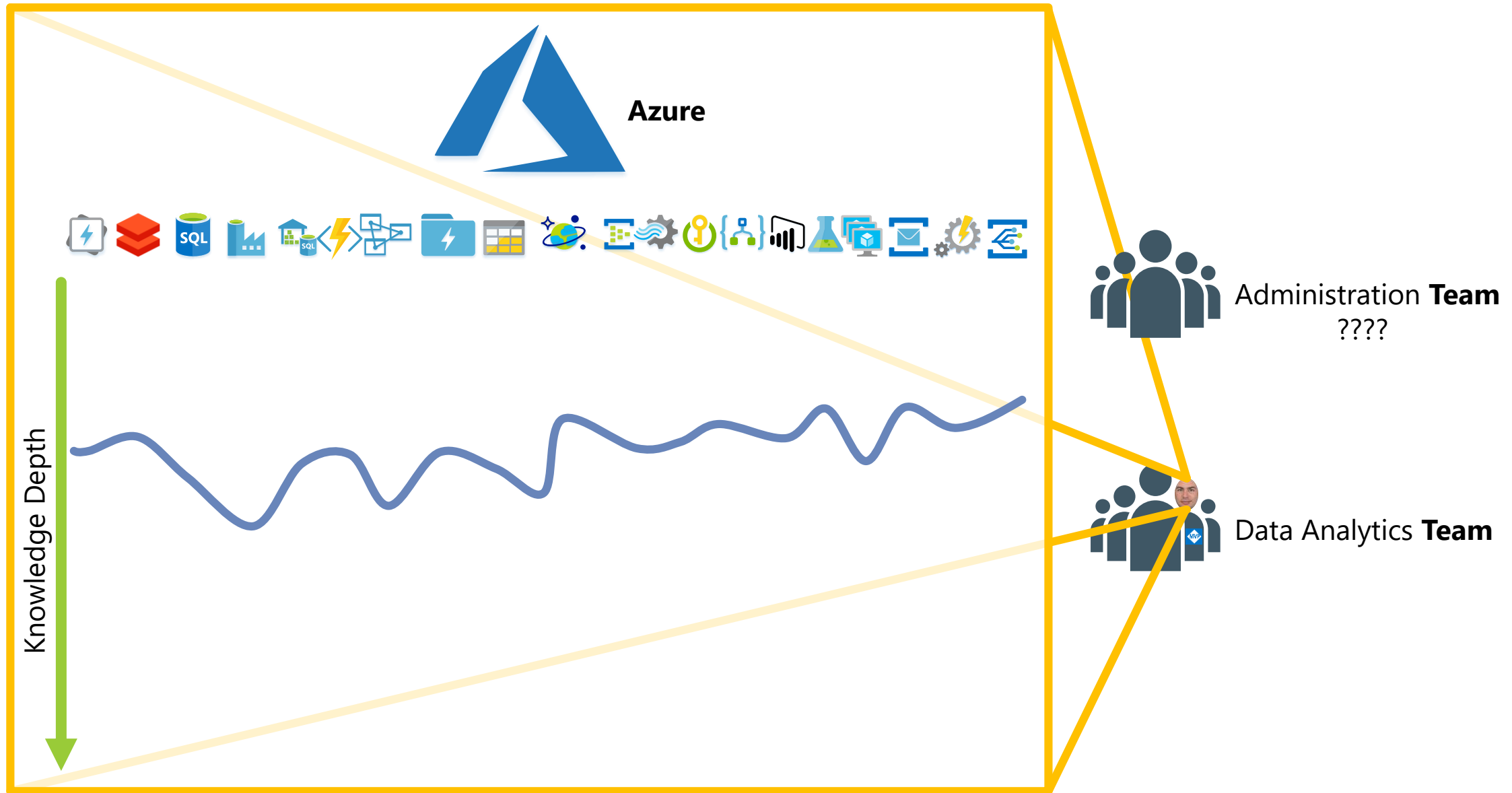
On Premises Solution



On Premises Solution



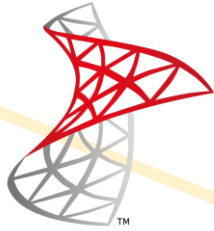
Cloud PaaS Solution



Hybrid Solution



Azure



SQL Server



**Azure
SQL Server**

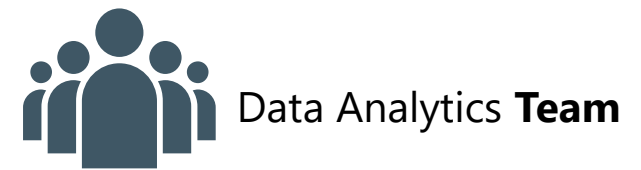


Data Analytics Team



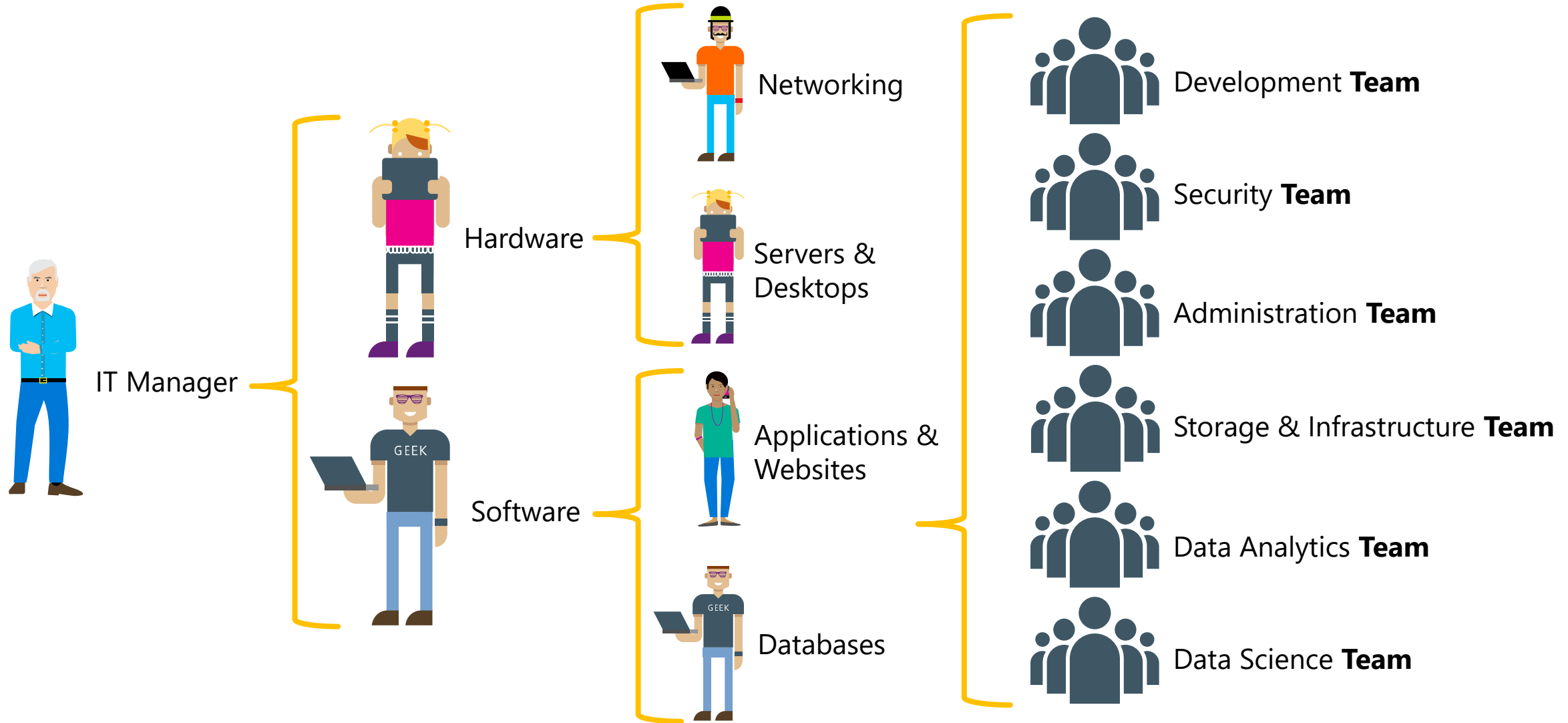
Data Analytics **Team**

IT Industry Role Evolution



Time

IT Industry Role Evolution



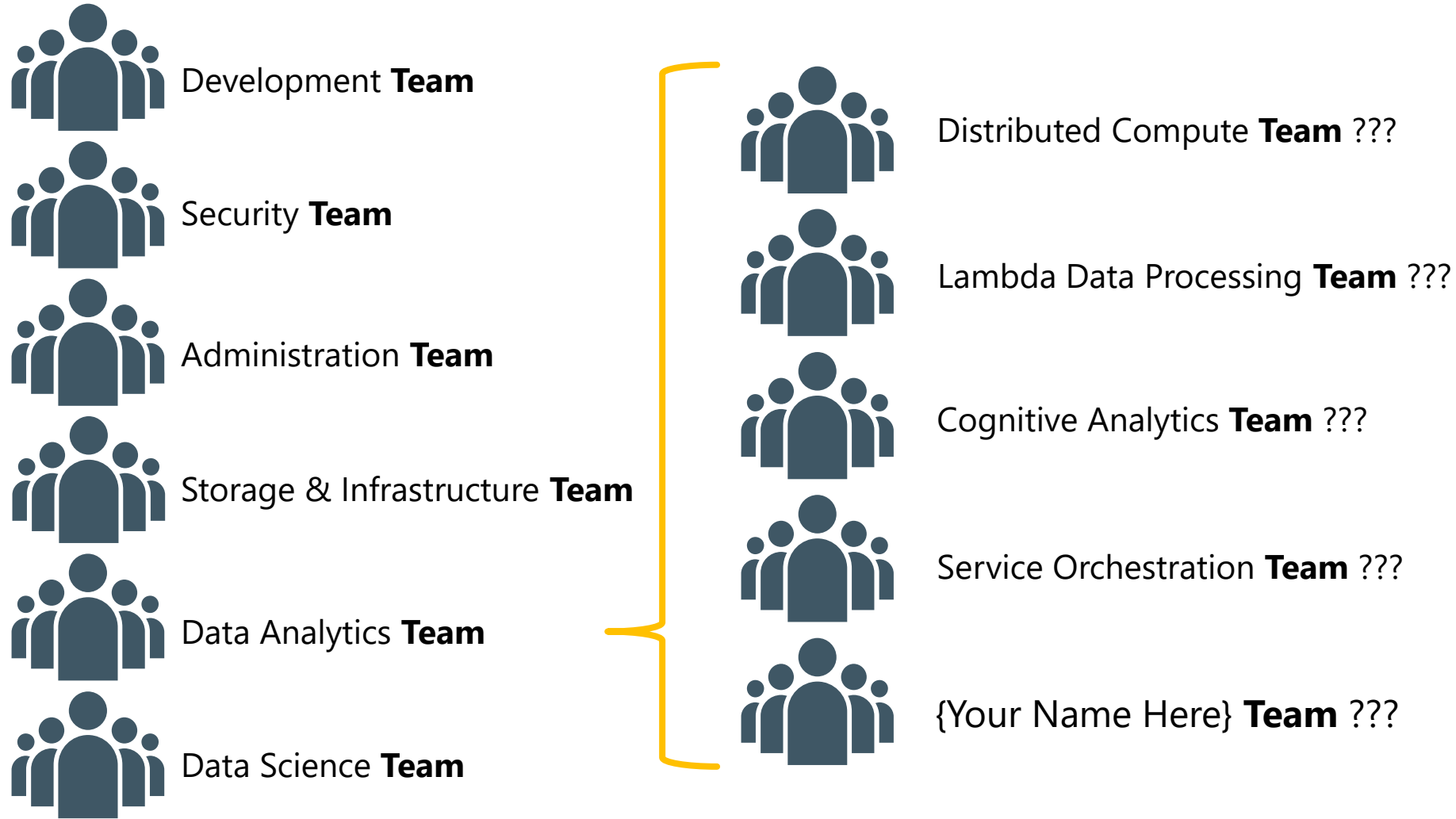
Time

IT Industry Role Evolution

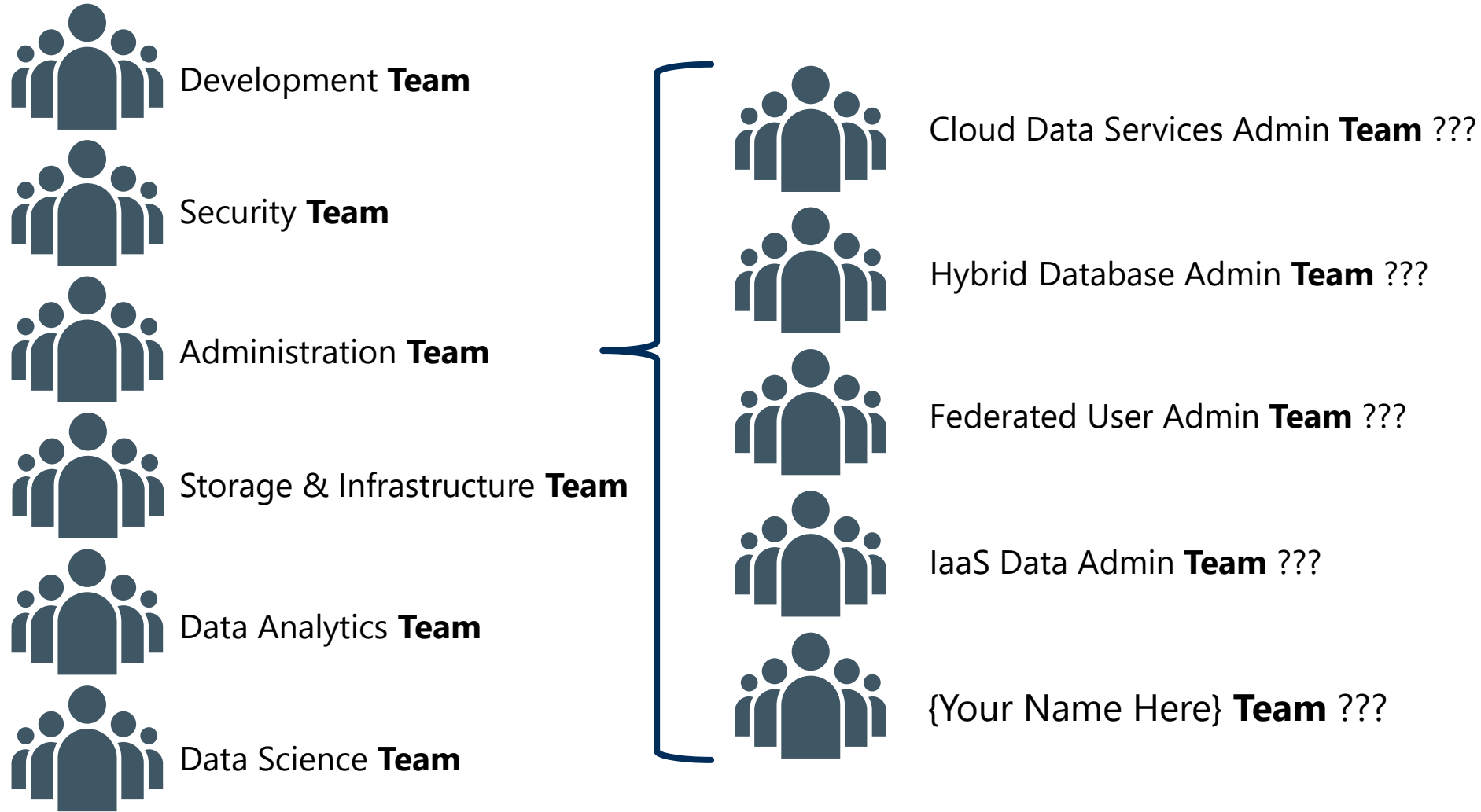


Time

IT Industry Role Evolution



IT Industry Role Evolution





Agenda

Context

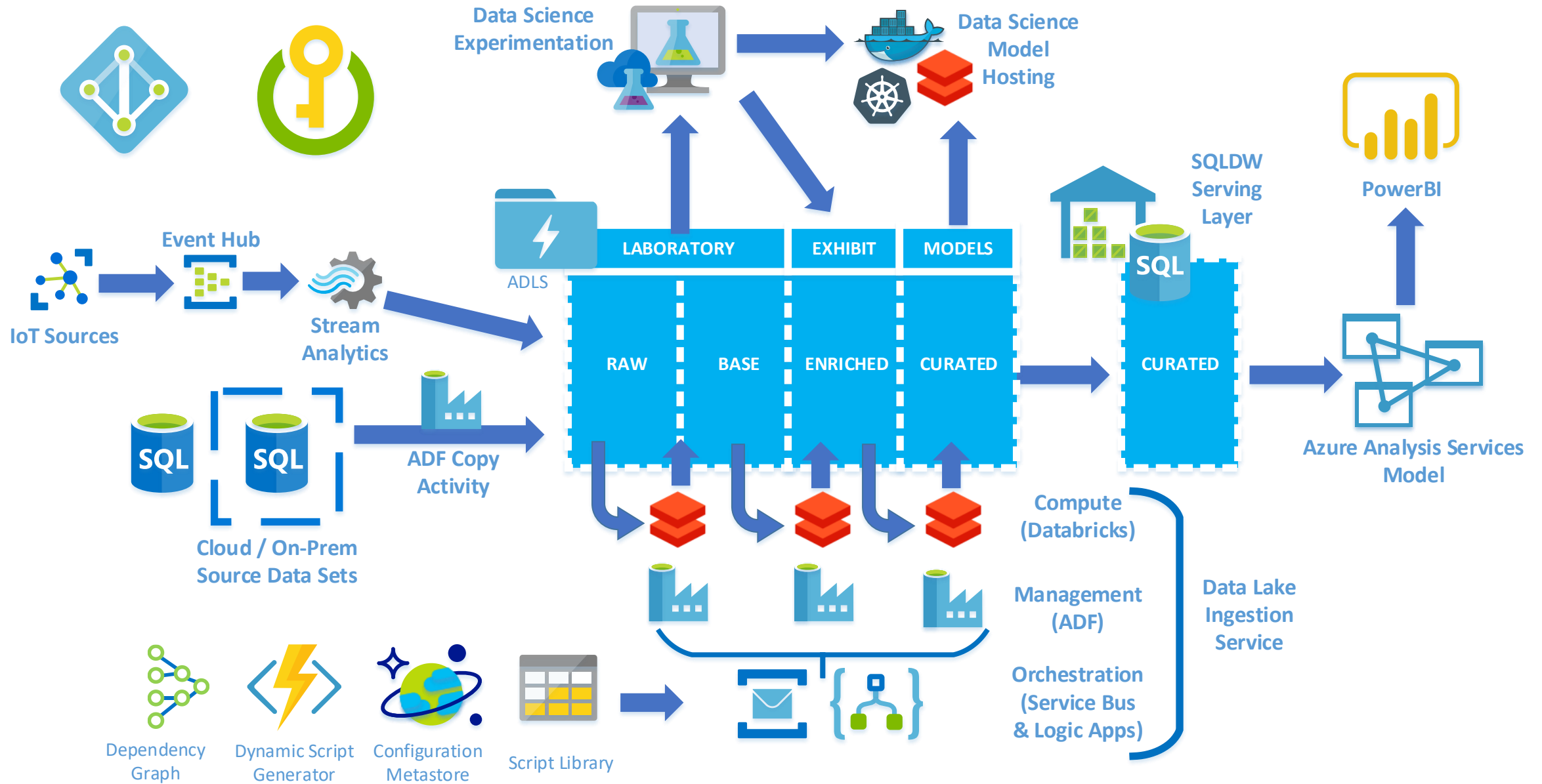
IT Industry Roles
Knowledge

The Modern Data Warehouse

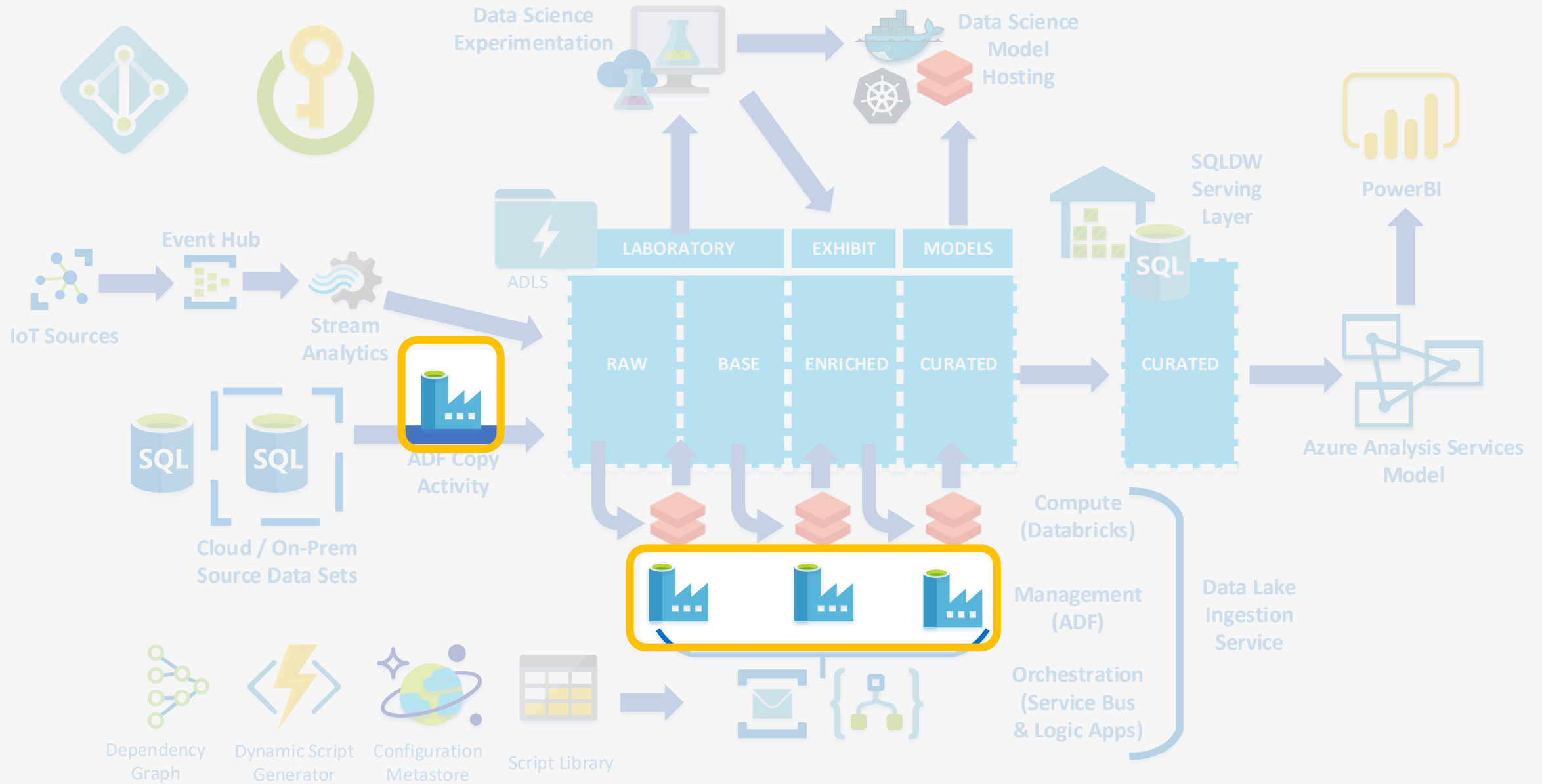
Components
Considerations

Conclusions

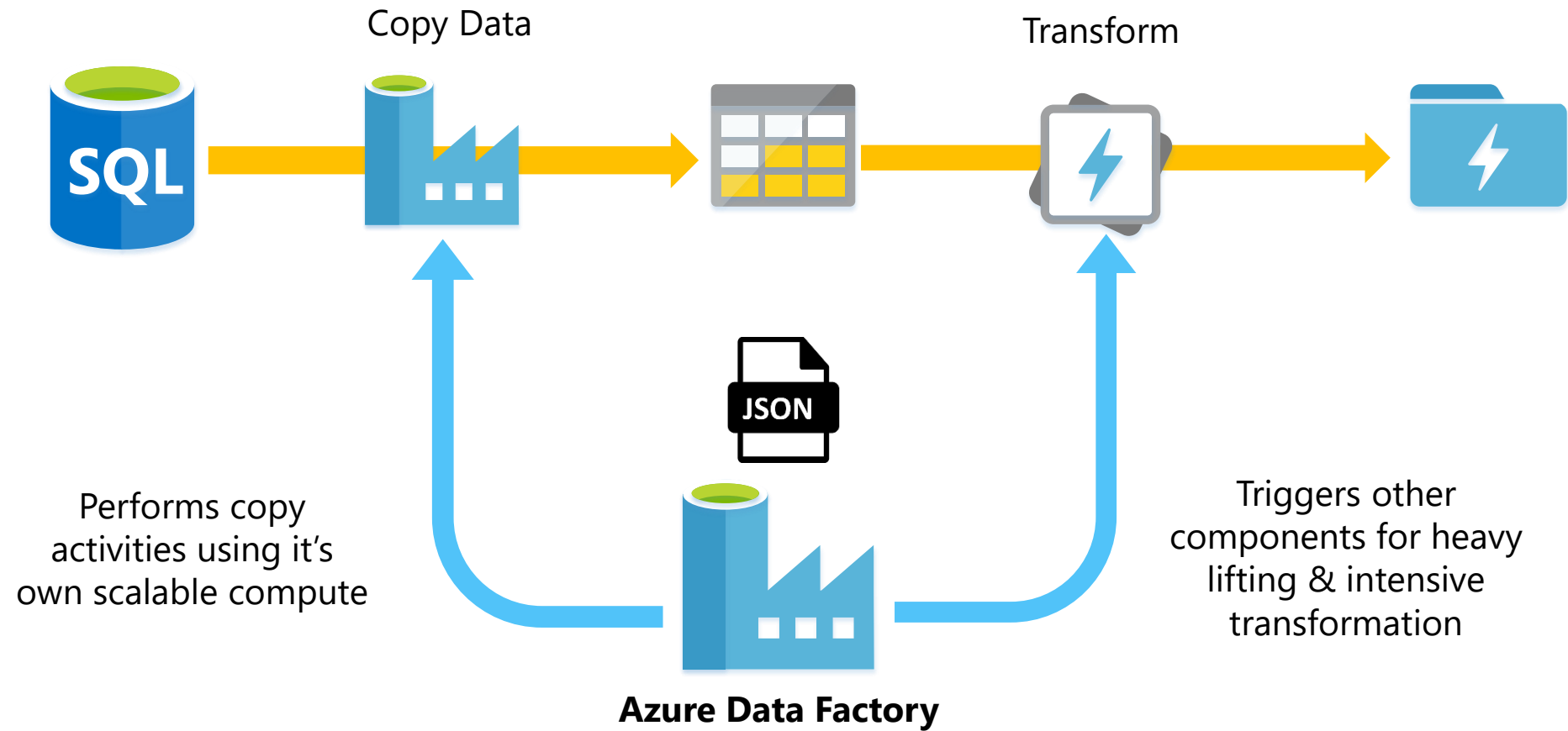
The Modern Data Warehouse



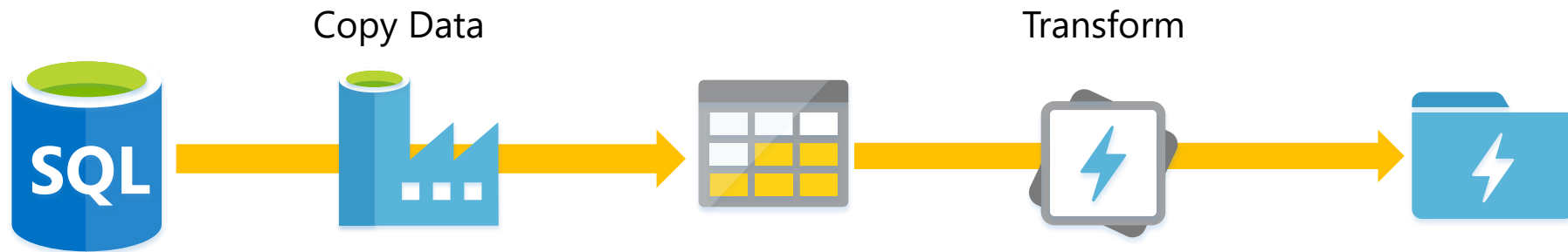
The Modern Data Warehouse



What is Azure Data Factory?

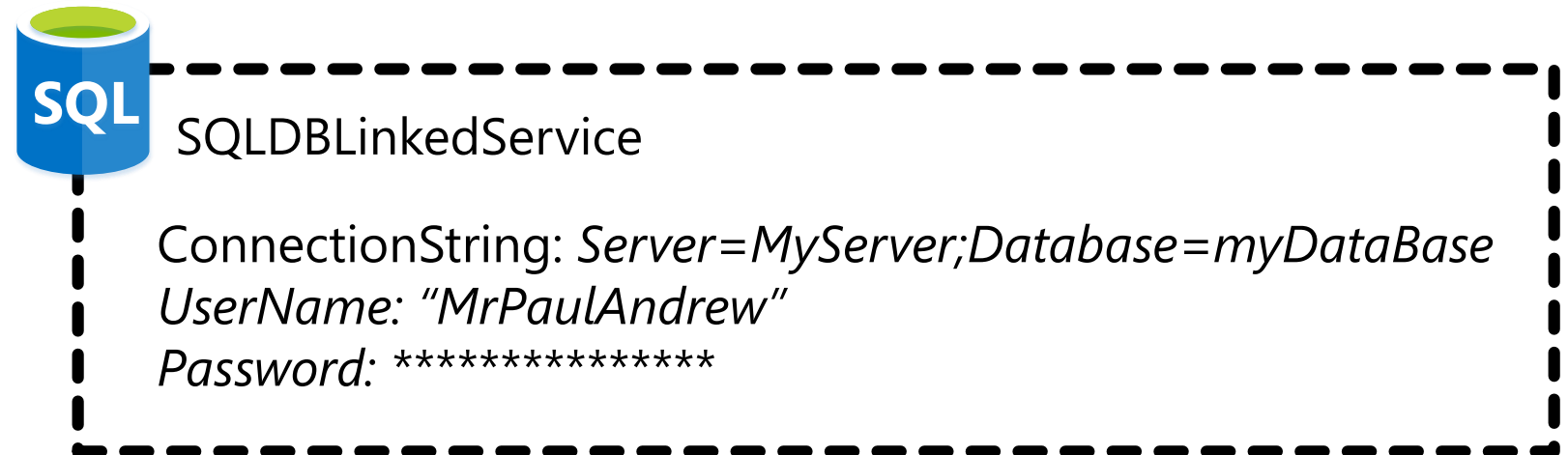
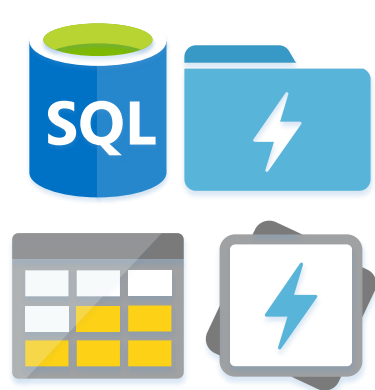


Data Factory Components

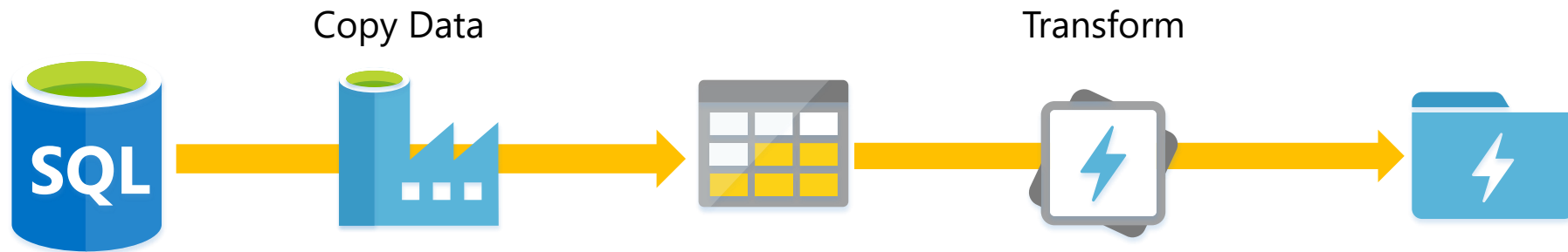


1 Linked Services – How do I connect?

Like the SSIS Connection Manager!



Data Factory Components



1

Linked Services

2

Data Sets – What slices/partitions does my data have?

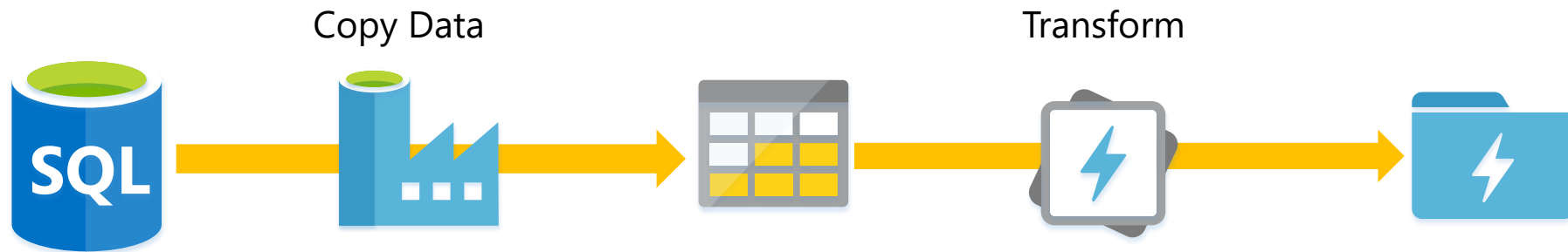


dbo.DimCustomer



/RAW/Orders/2018/01/01/Orders.csv

Data Factory Components



1

Linked Services

2

Data Sets

3

Activities – What do we want to happen?
With what conditions?



U-SQL Activity

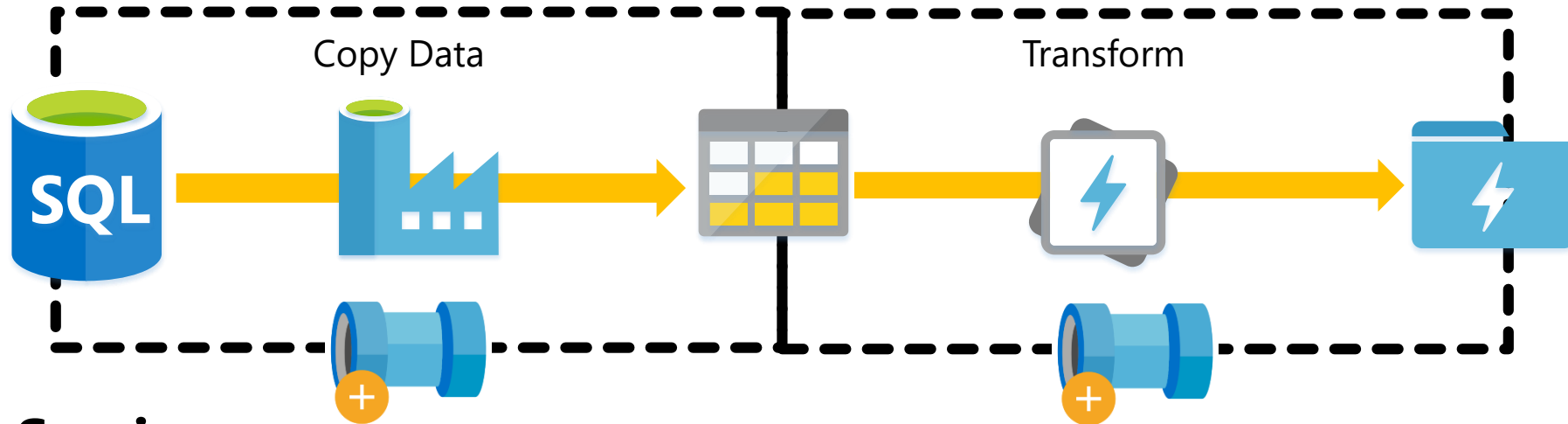
Script: *wasb://:myscripts/ProcessOrders.usql*

AUs: *5 units*

Priority: *1000*

Parameters: *@Output = "RAW/Orders/..."*

Data Factory Components



1

Linked Services

2

Data Sets

3

Activities

4

Pipelines – What groups of work do I want to do?

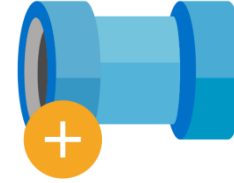
Integration Runtimes



1

Azure
Integration Runtime

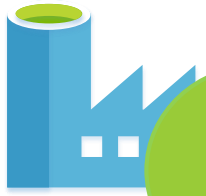
Movement Hours



Activity
Orchestration



Flexible Region



2

SSIS
Integration Runtime

SSIS Package
Execution



Specified Region



3

Self Hosted
Integration Runtime

Local Compute



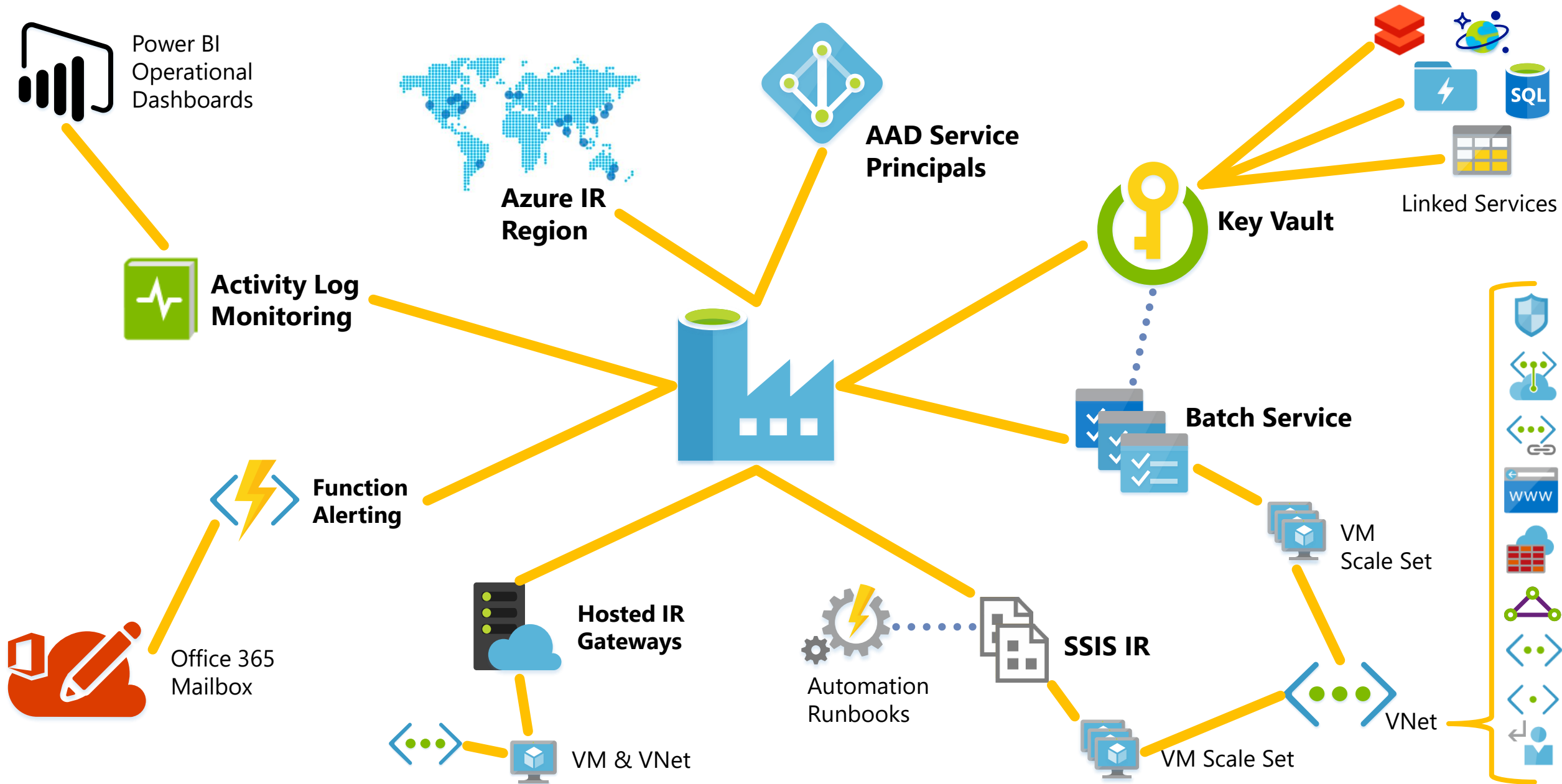
Activity
Orchestration



On-Prem Server

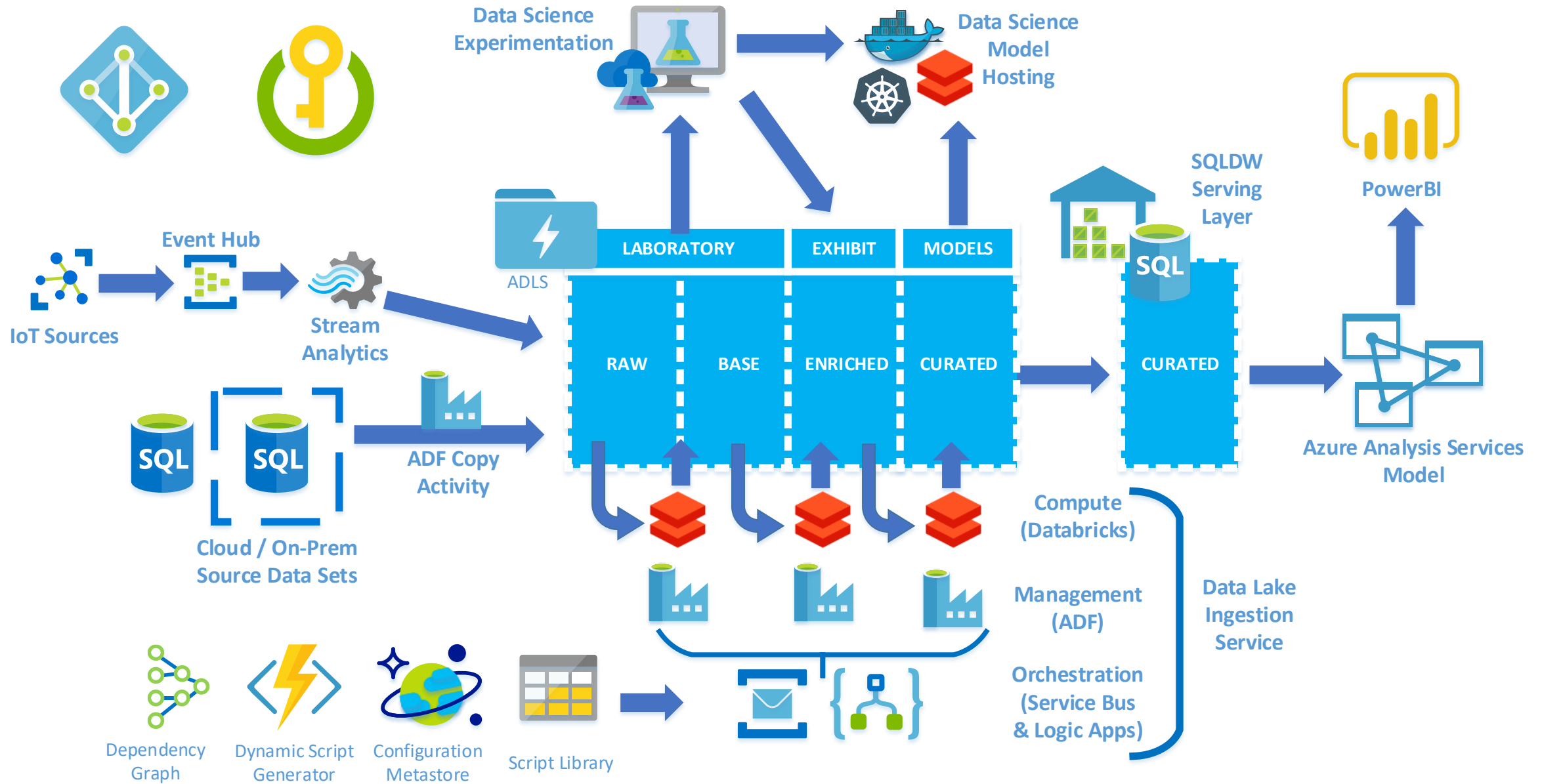


Administration Considerations – Azure Data Factory

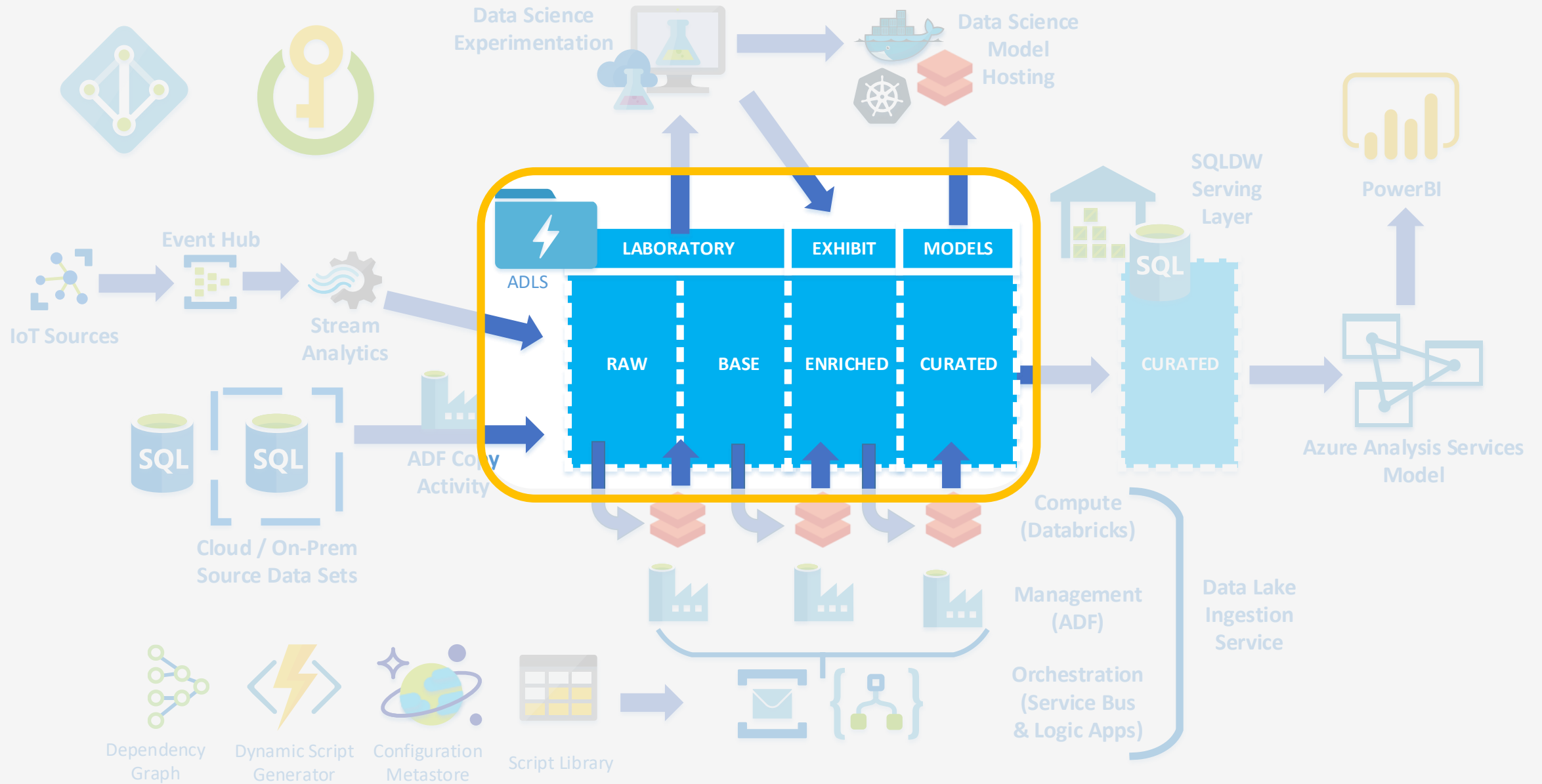




The Modern Data Warehouse



The Modern Data Warehouse



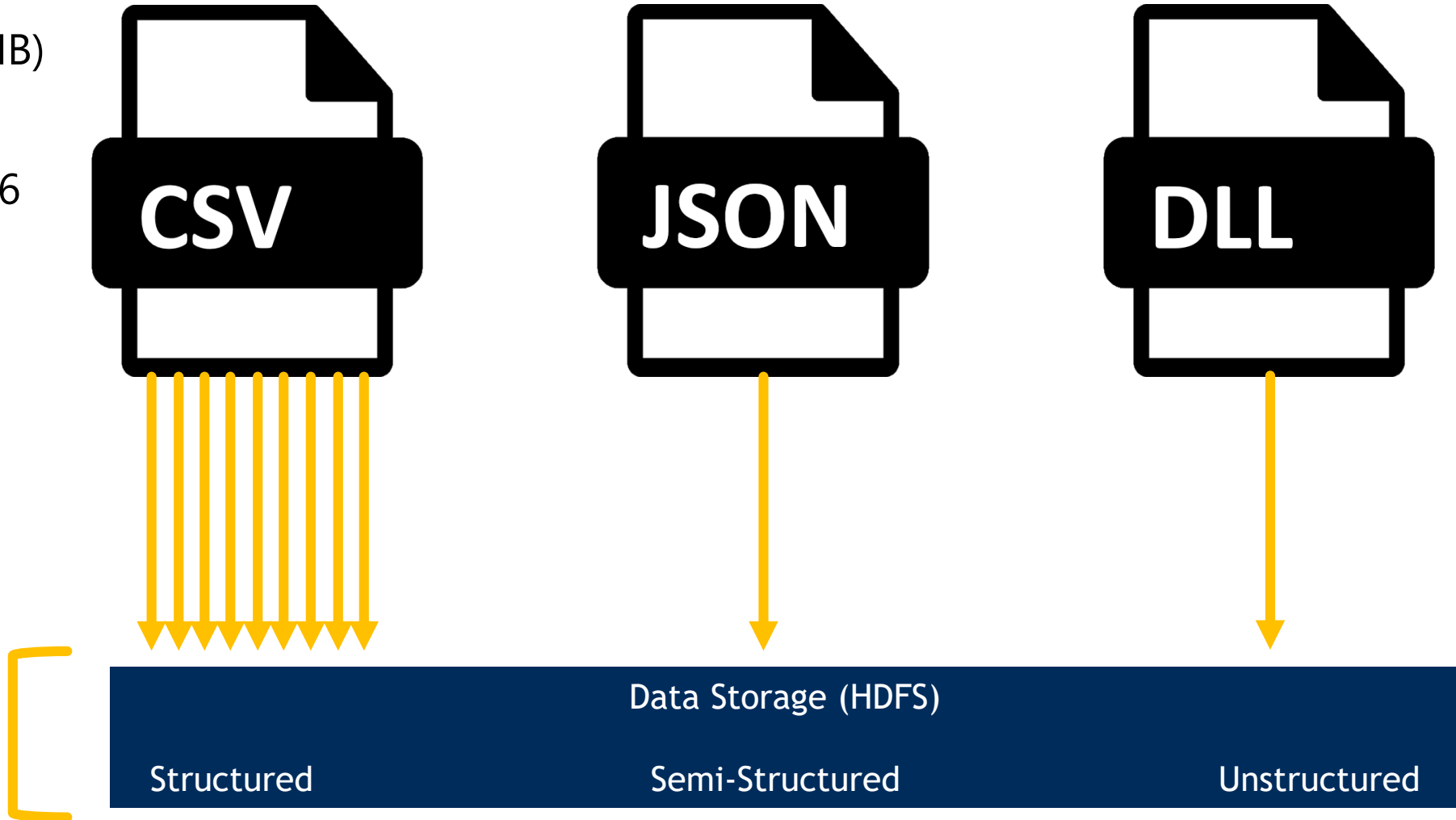
What is Azure Data Lake Storage?

HDFS Extents (MB)

Default: 128
Variable: 4 to 256

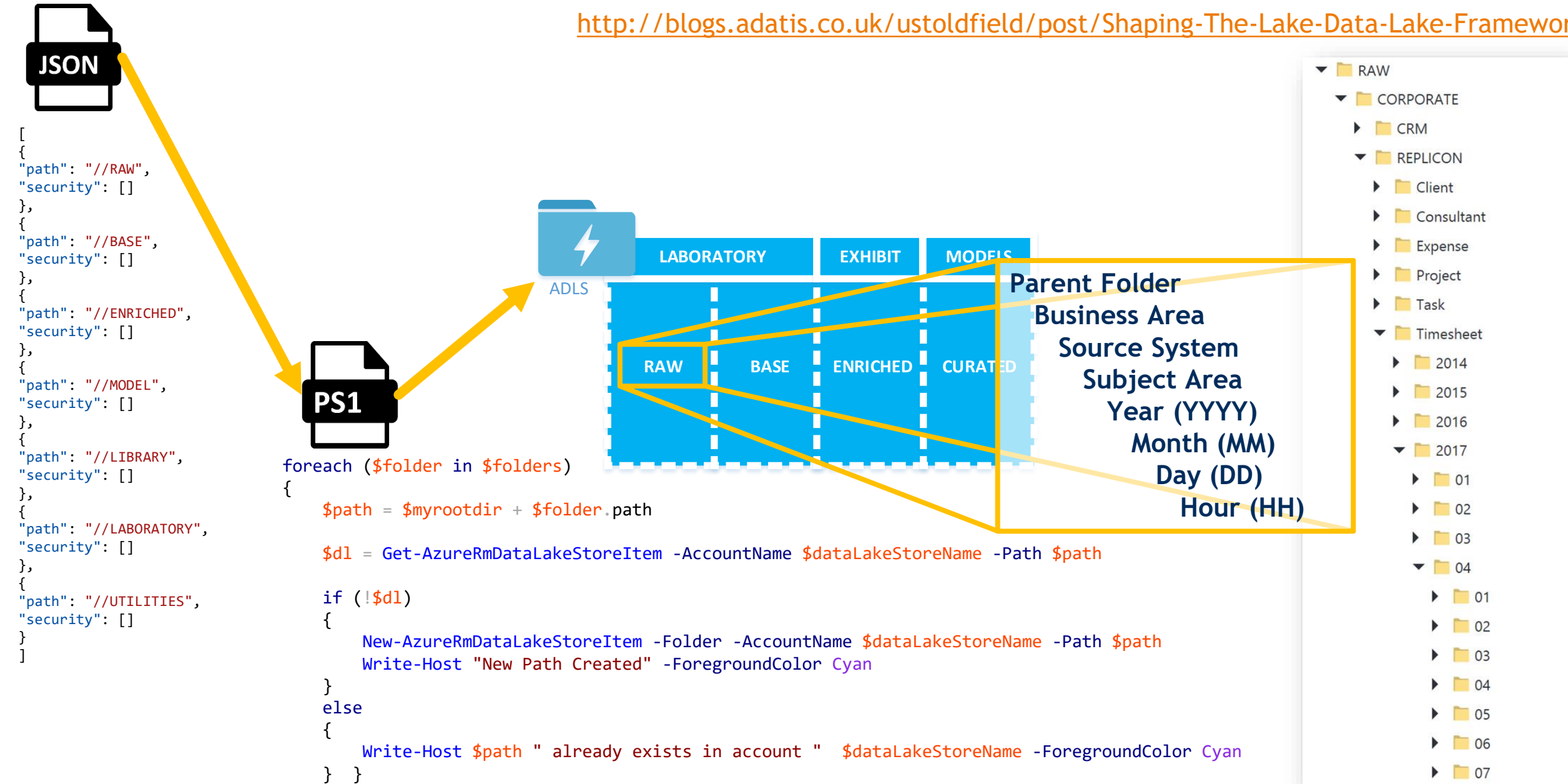


Storage

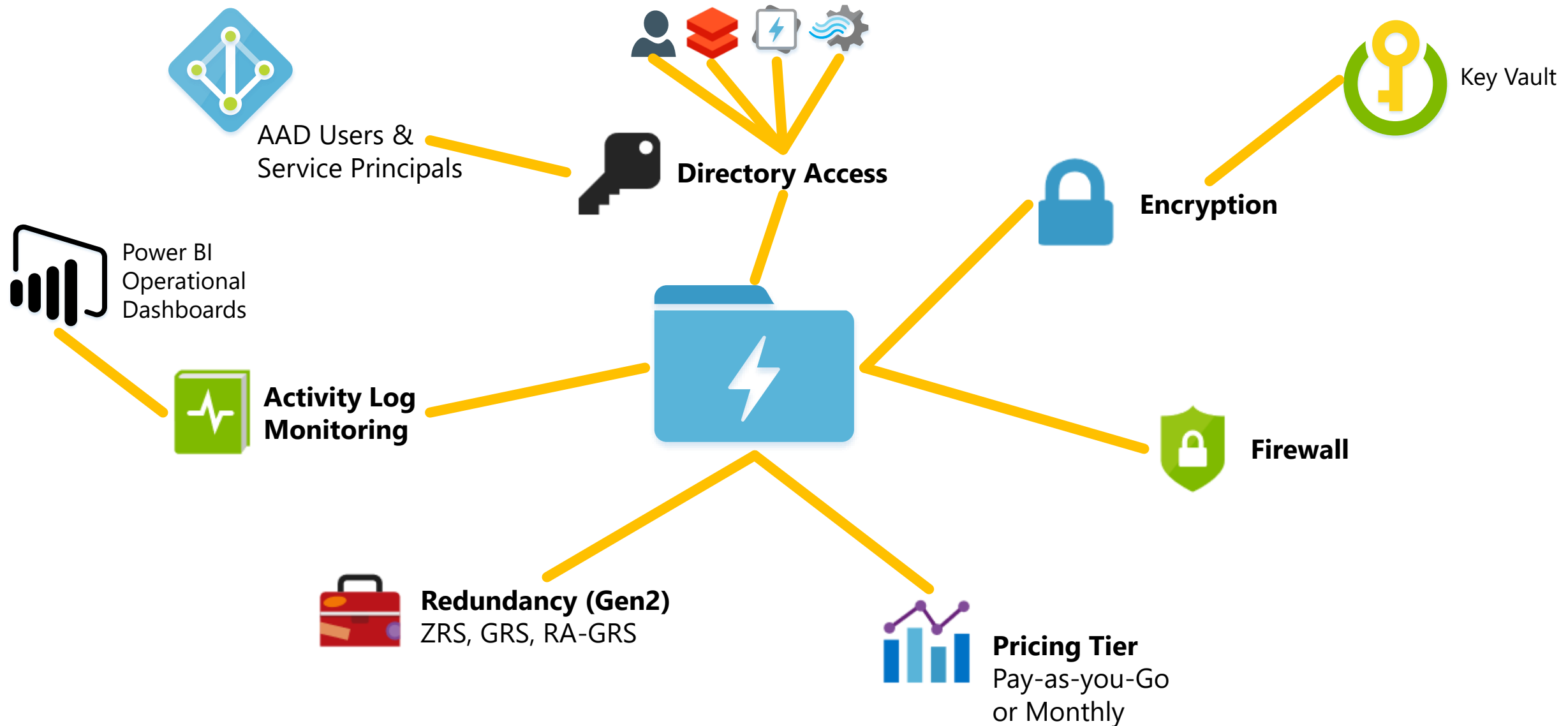


Azure Data Lake Storage

<http://blogs.adatis.co.uk/ustoldfield/post/Shaping-The-Lake-Data-Lake-Framework>

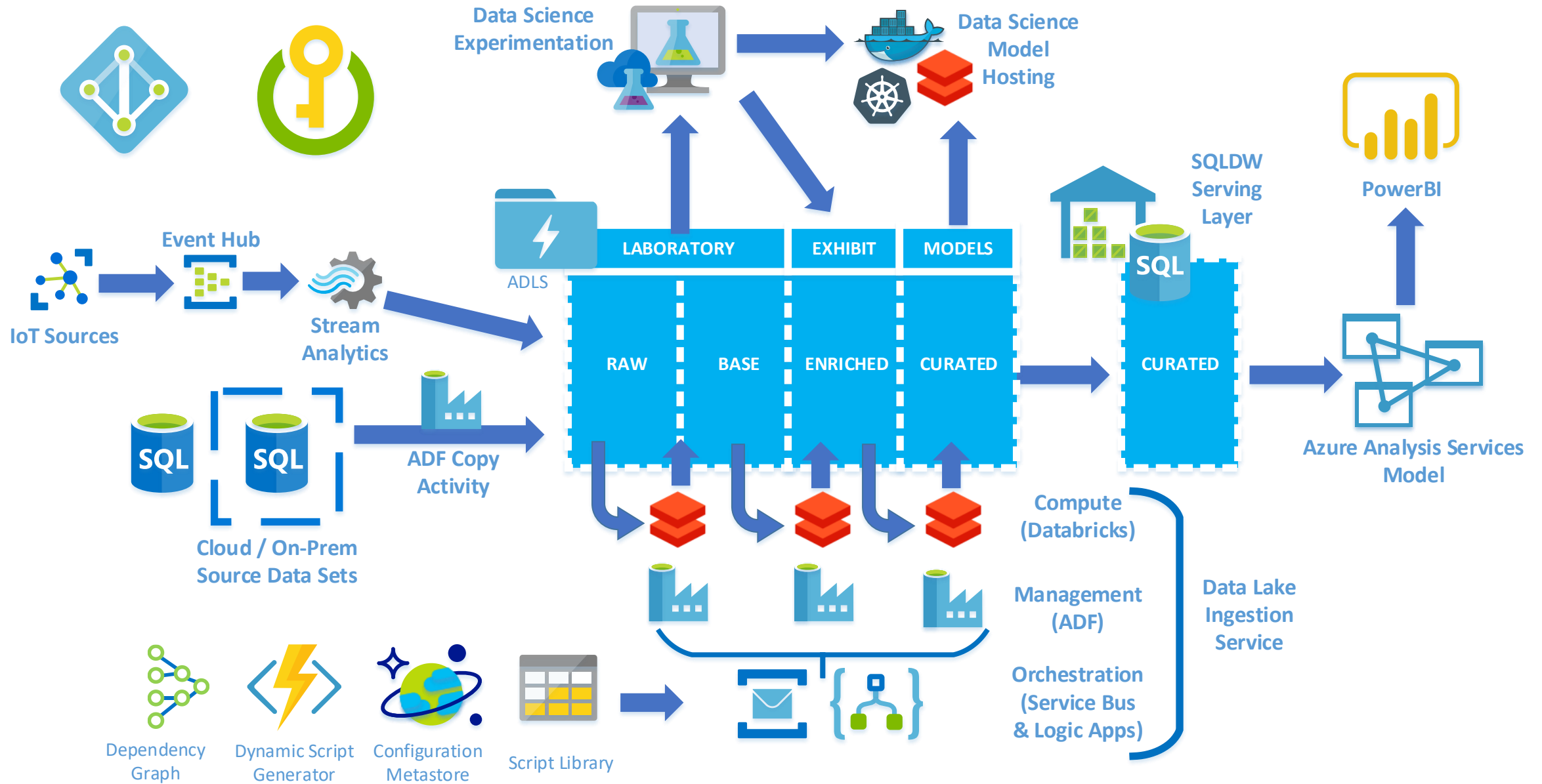


Administration Considerations – Azure Data Lake

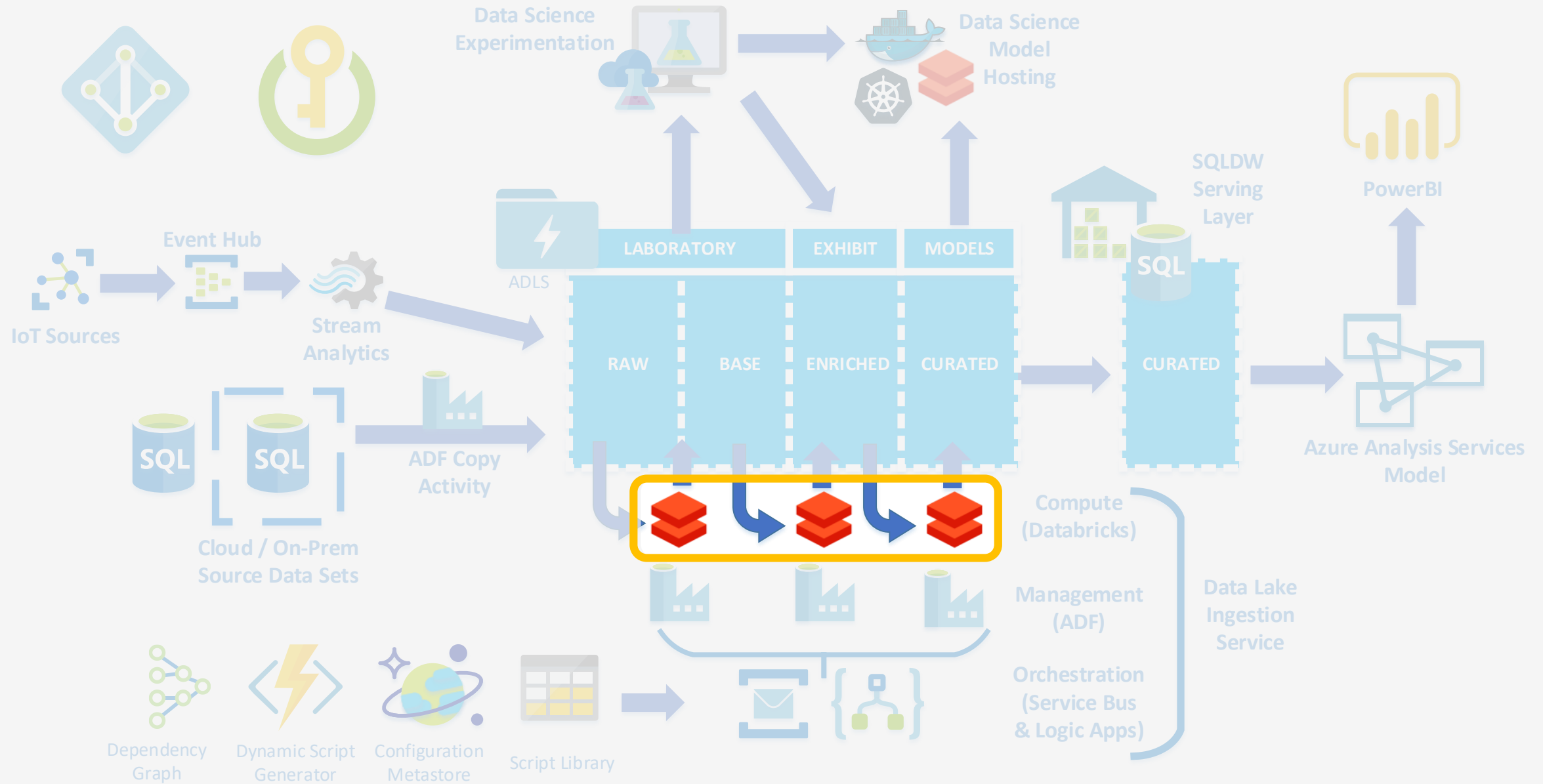




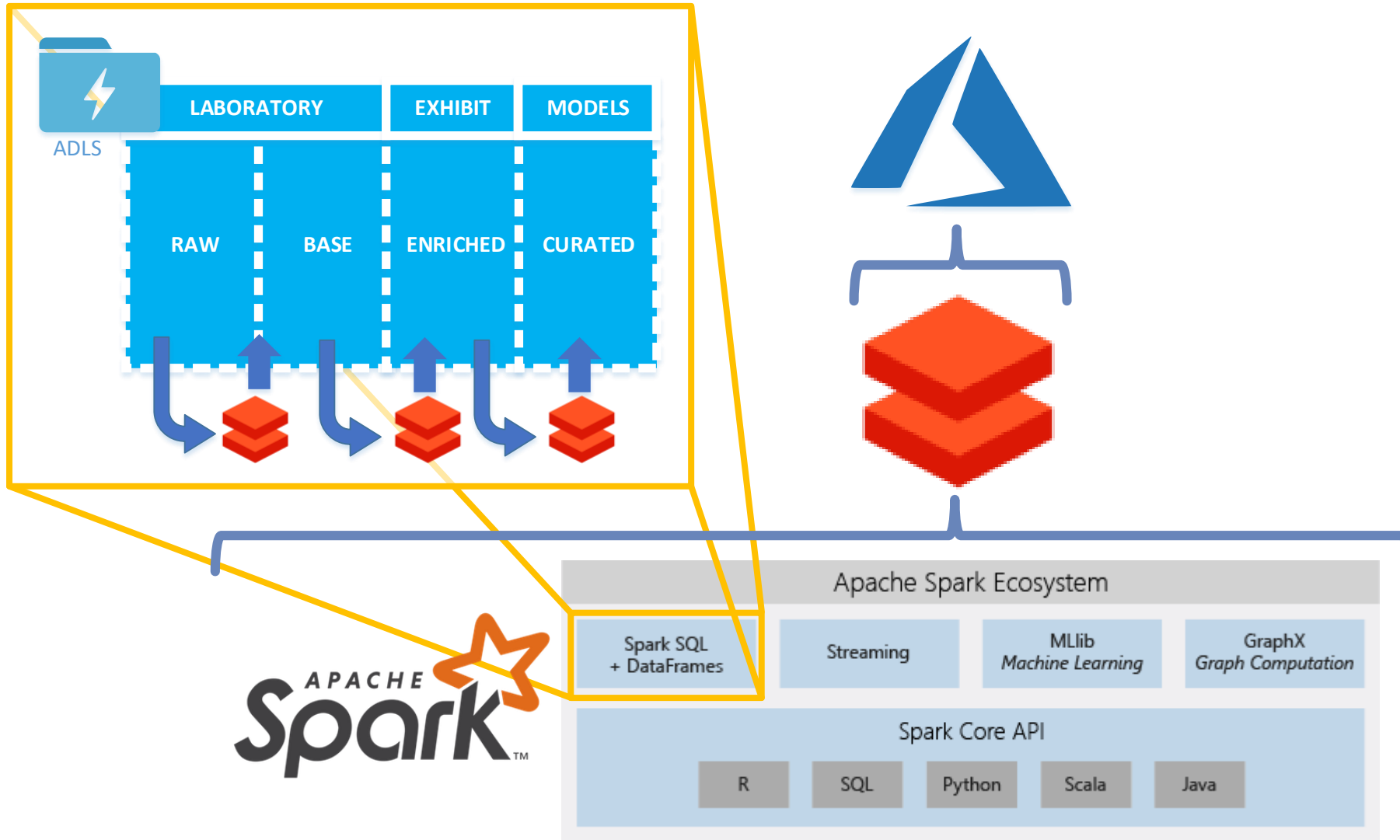
The Modern Data Warehouse



The Modern Data Warehouse



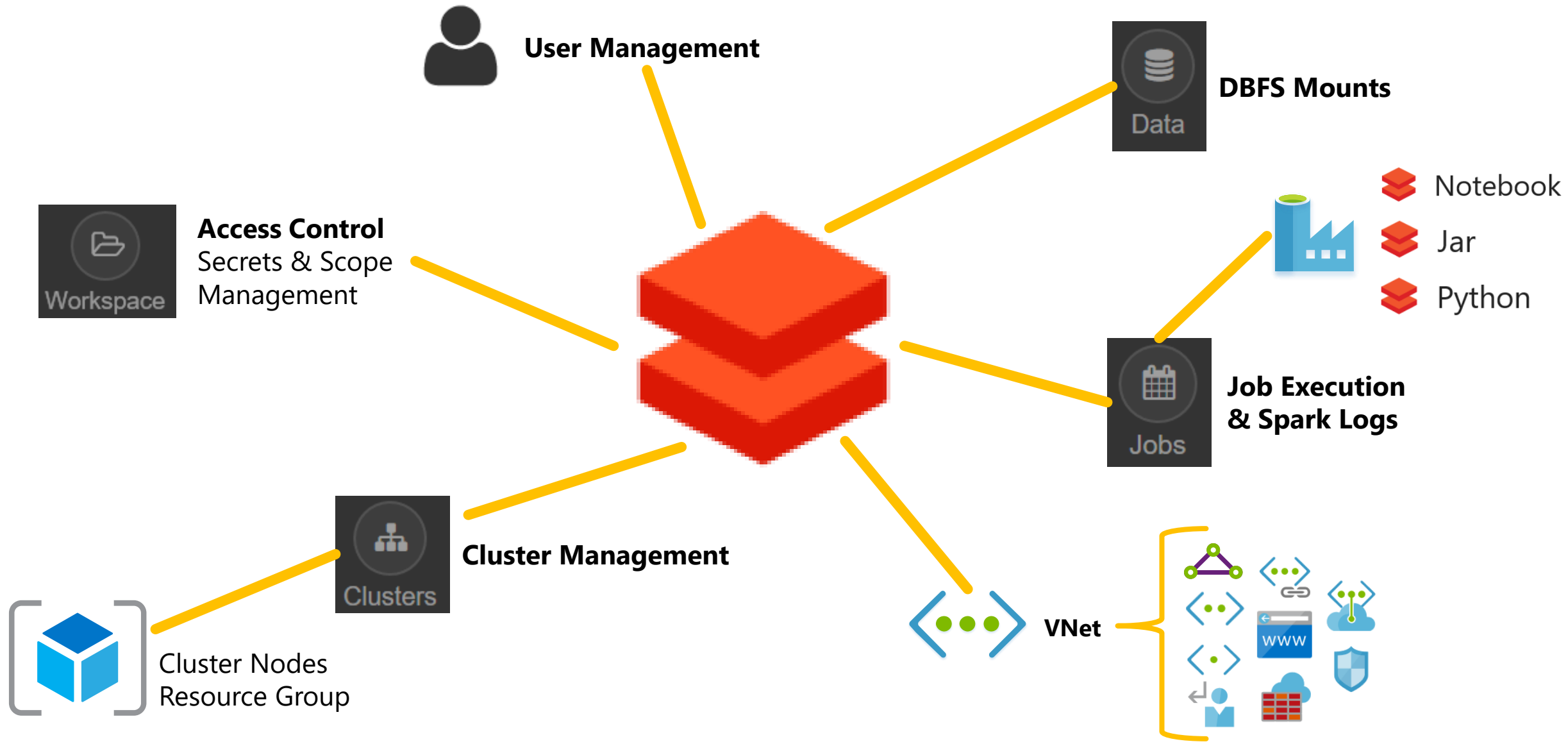
What is Azure Databricks?



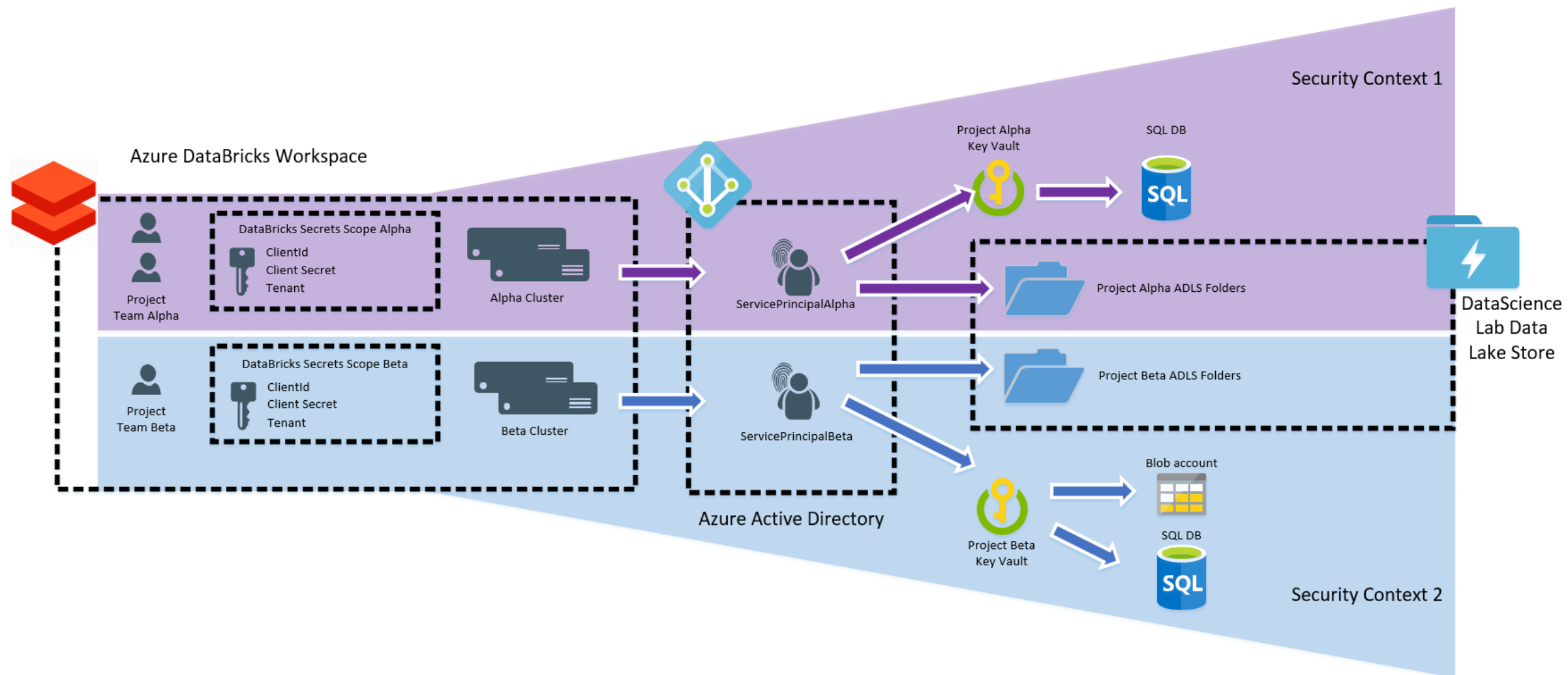
Databricks founded by
Matei Zaharia

2013

Administration Considerations – Azure Databricks

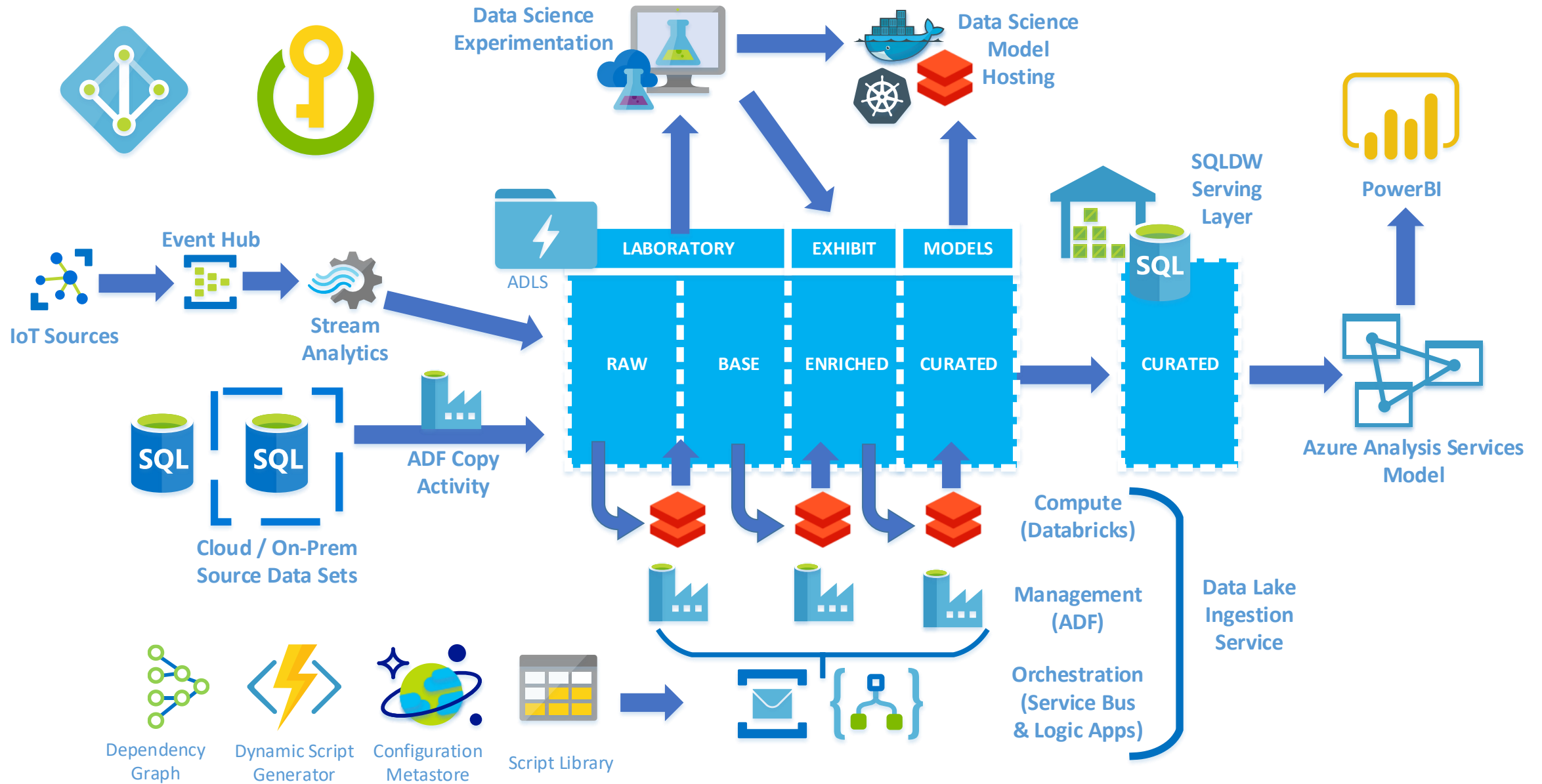


Databricks Workspace Scopes

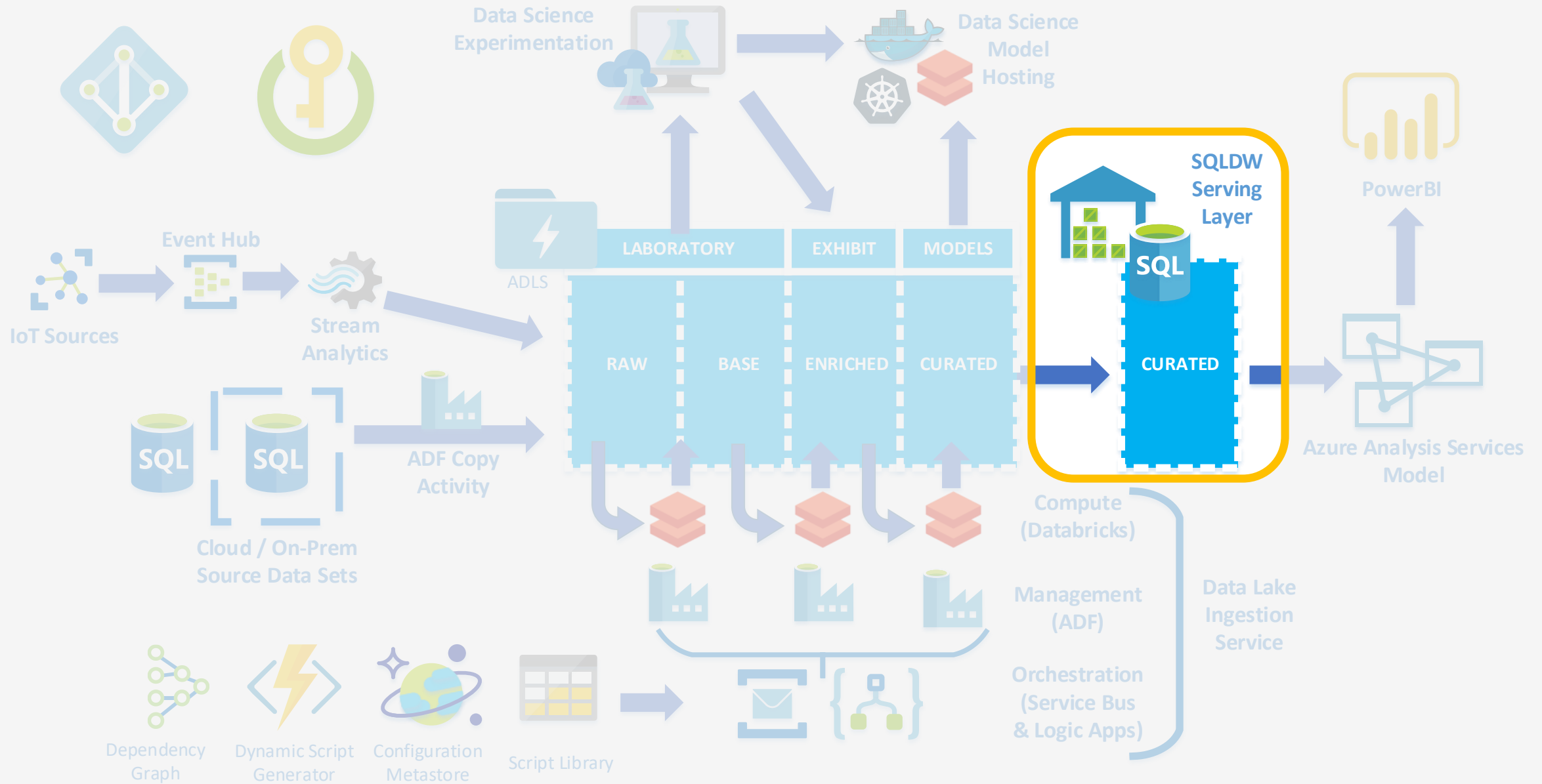




The Modern Data Warehouse



The Modern Data Warehouse



What is Azure SQL Data Warehouse?

Azure SQL Database (SQLDB)



T-SQL

Compute Node

Database Storage Engine

Database Transaction Units (DTU's) –
Scale Up Compute

Azure SQL Data Warehouse (SQLDW)



D-SQL

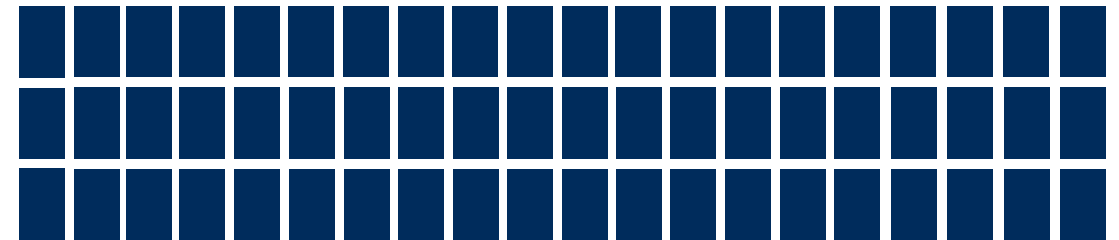
Control Node

Compute Node

Compute Node

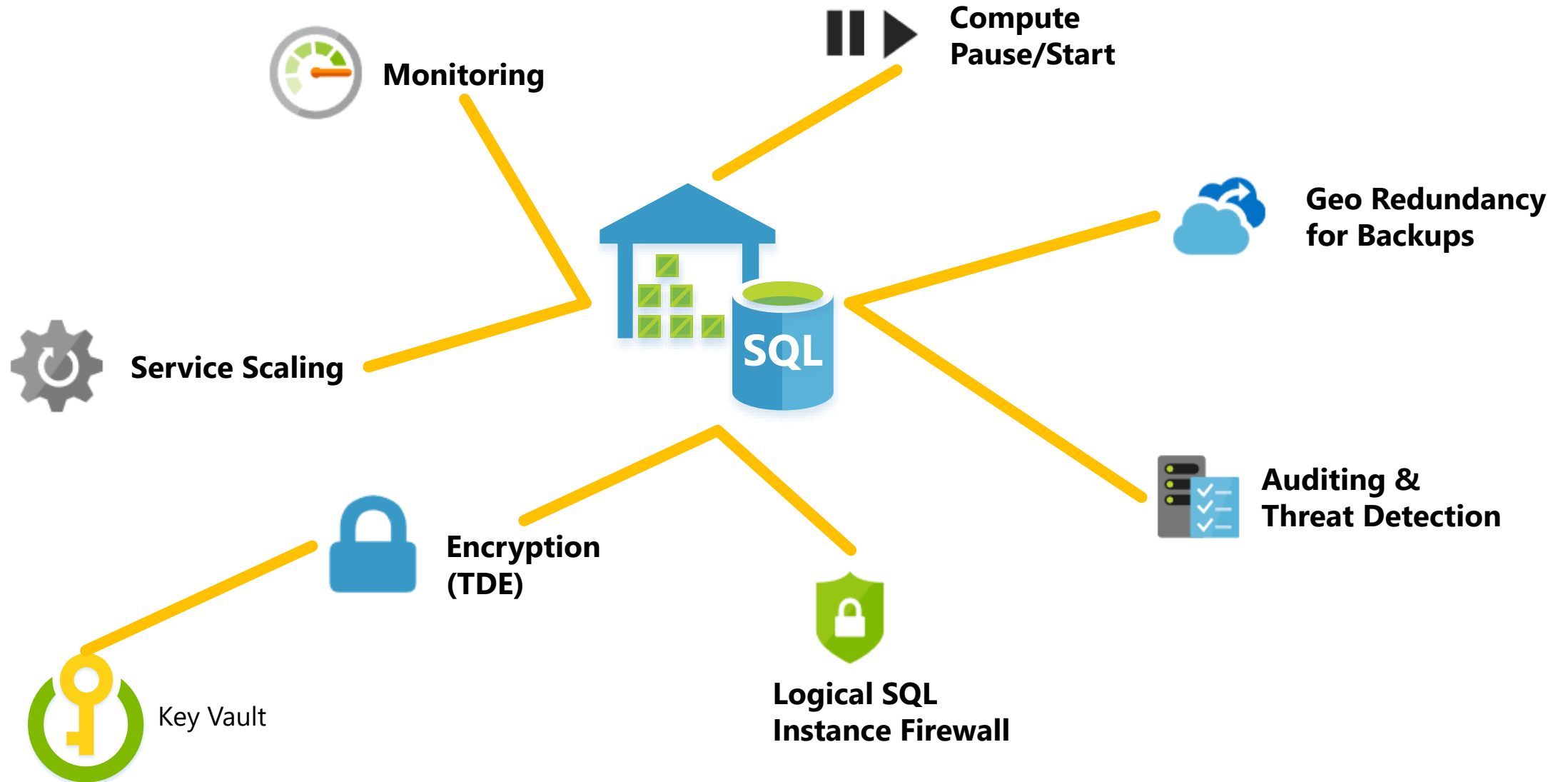
Compute Node

Compute Node



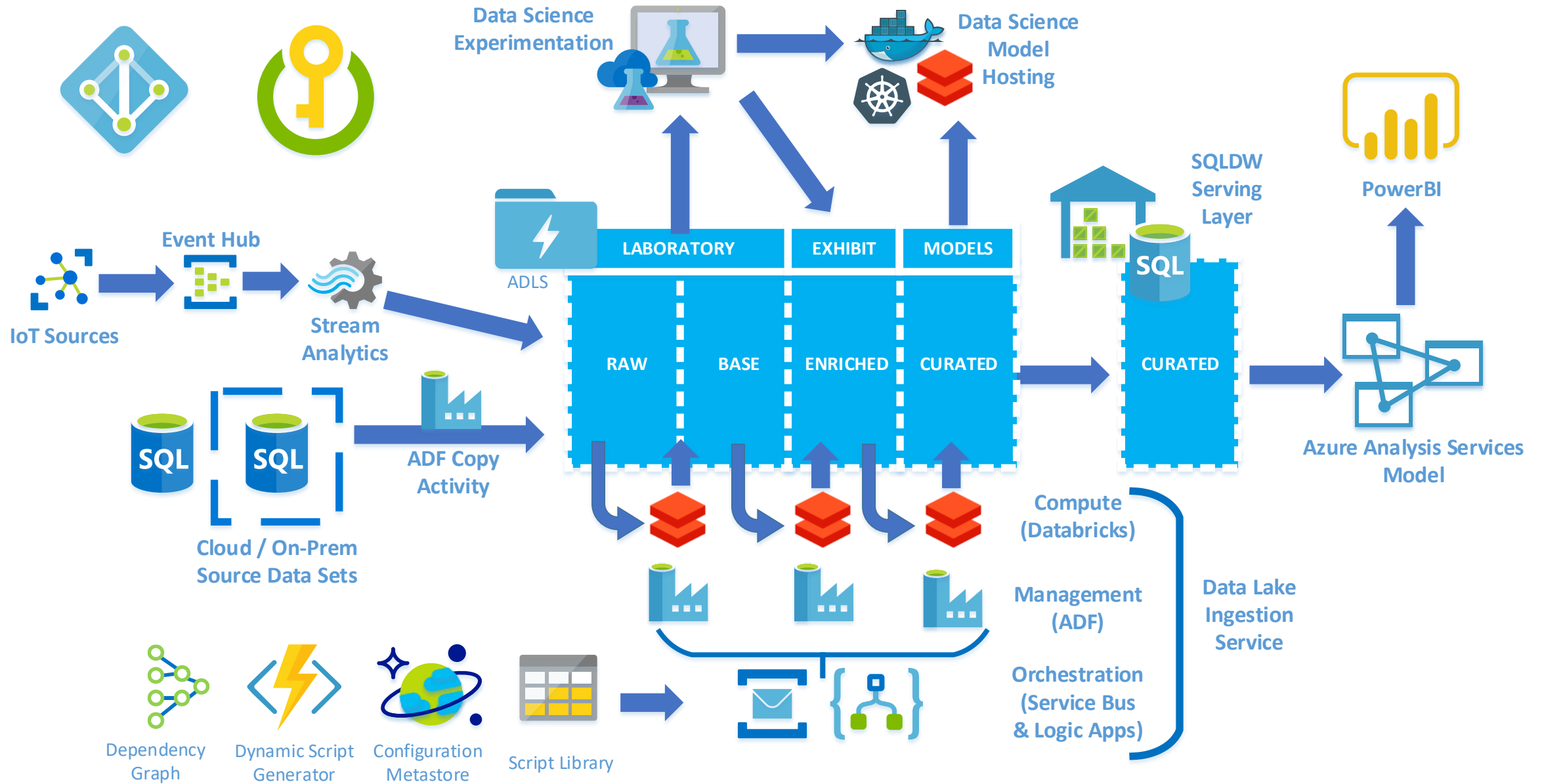
Database Warehouse Units (DWU's) –
Scale Out Compute

Administration Considerations – Azure SQL Data Warehouse

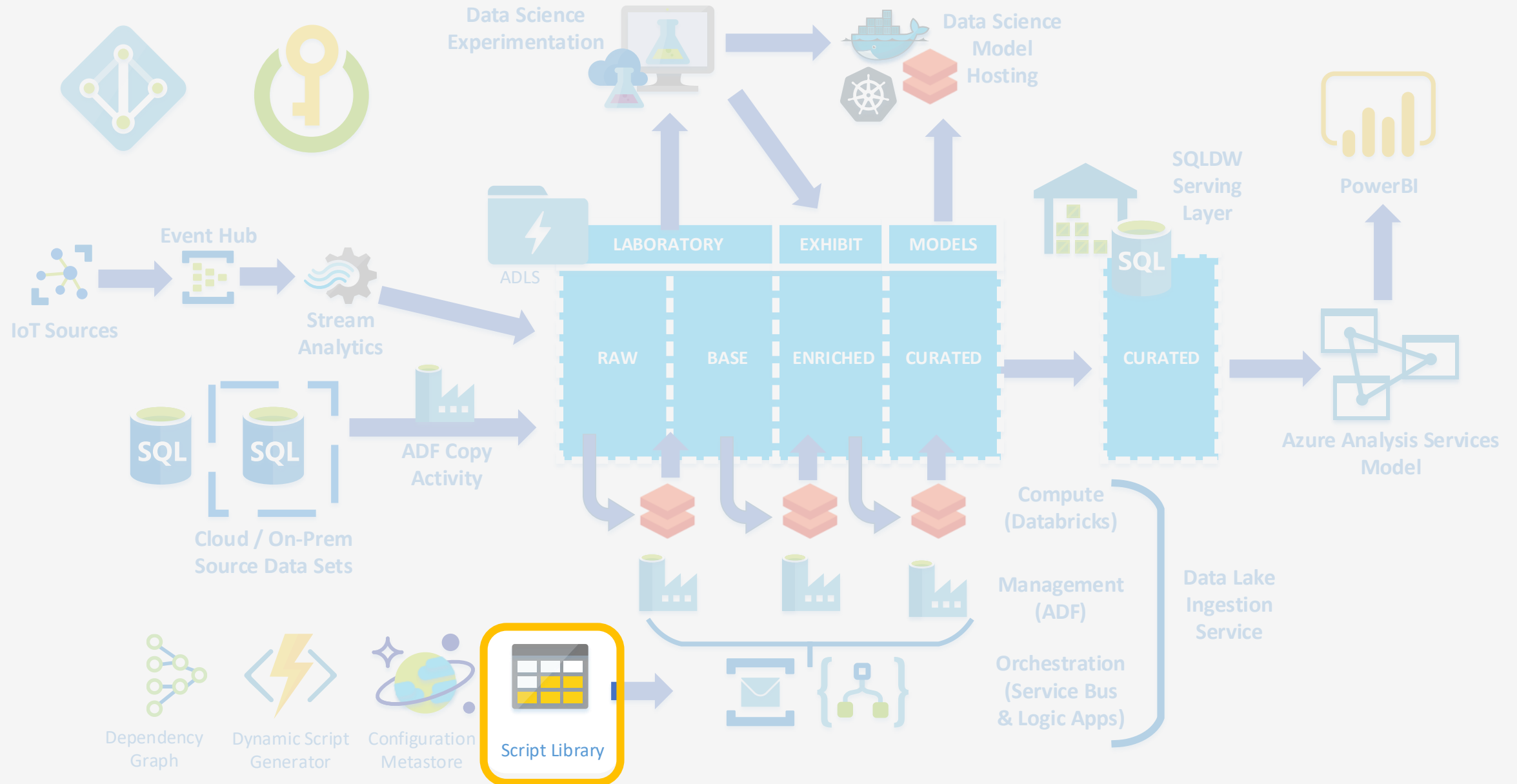




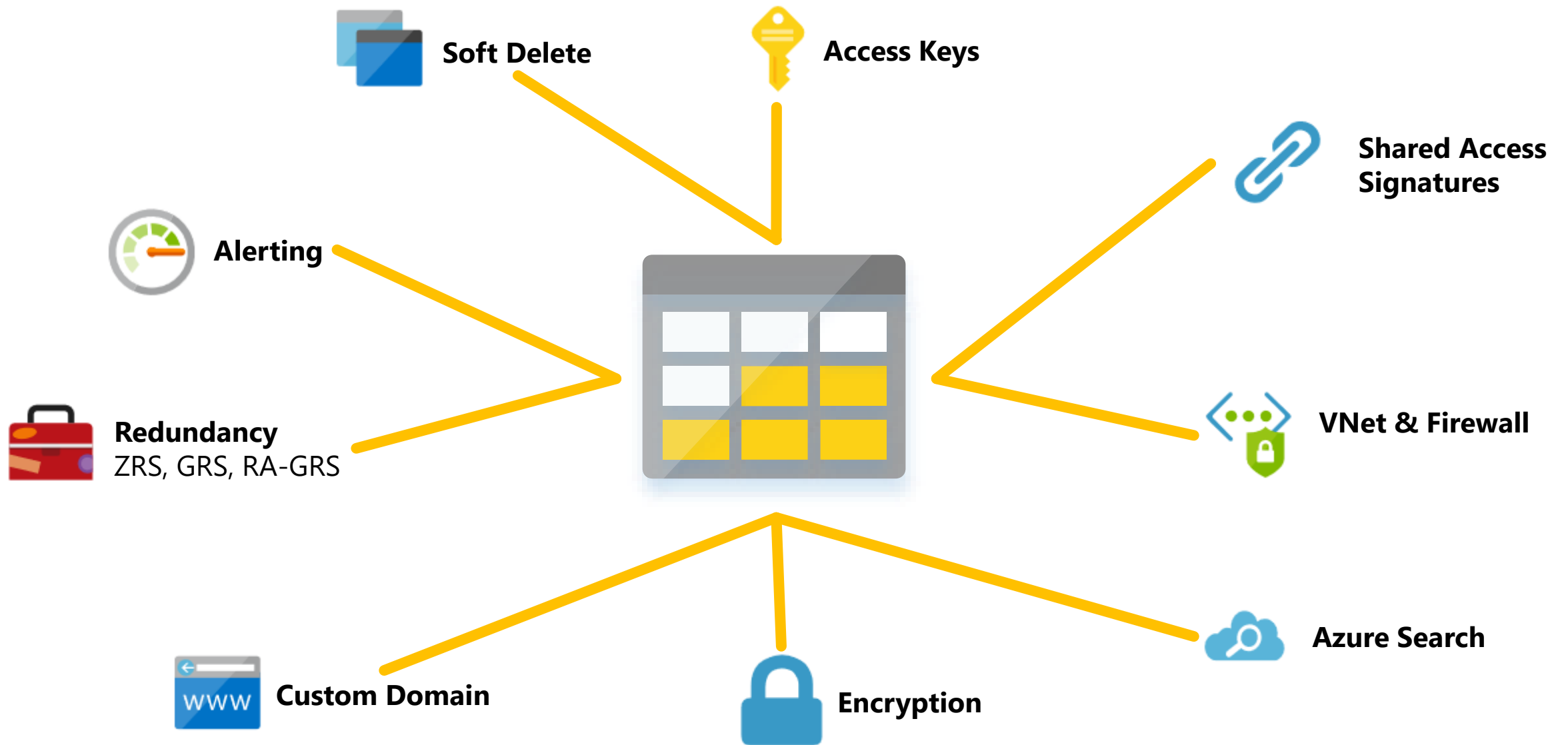
The Modern Data Warehouse



The Modern Data Warehouse

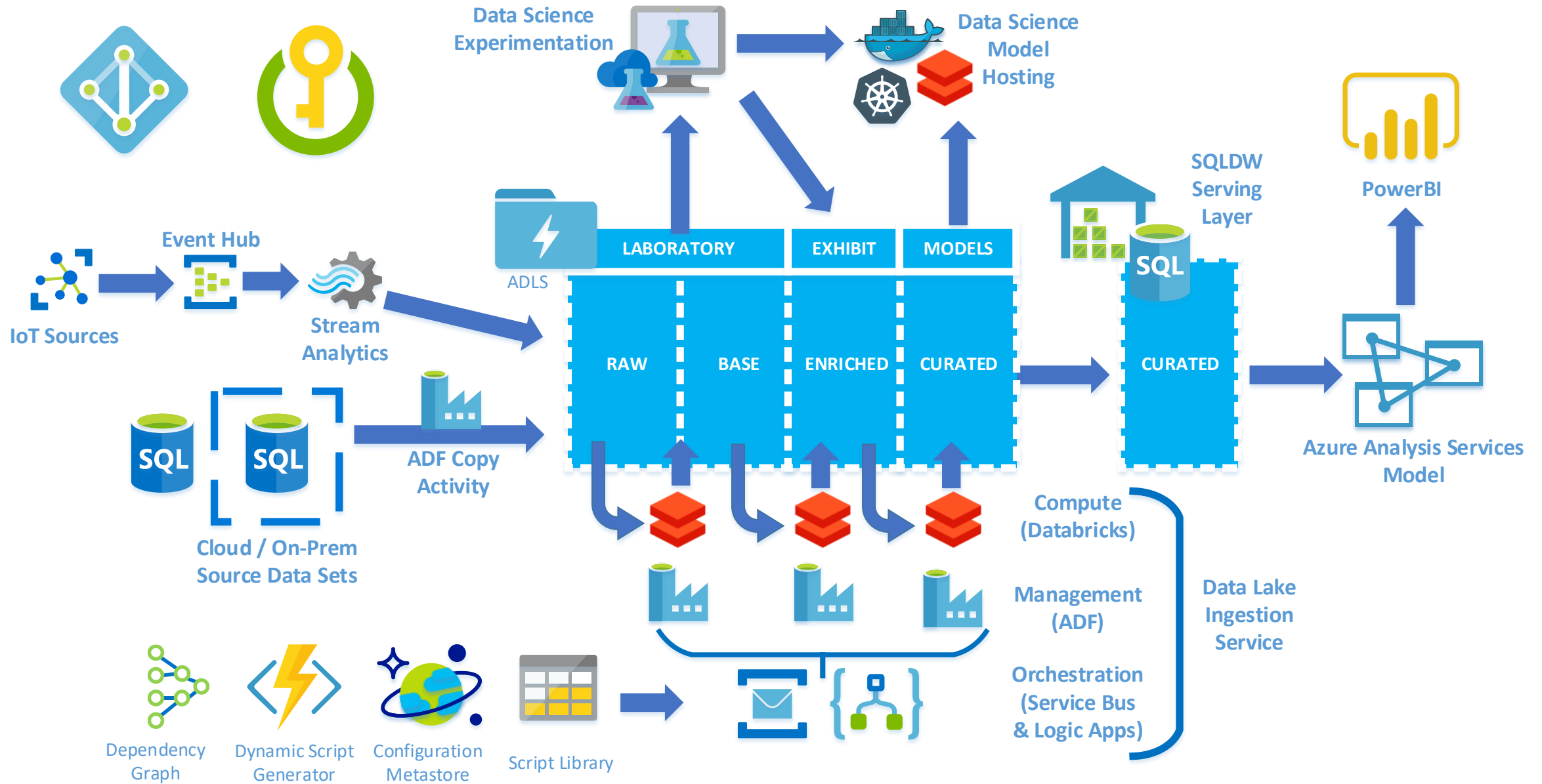


Administration Considerations – Azure Storage Account

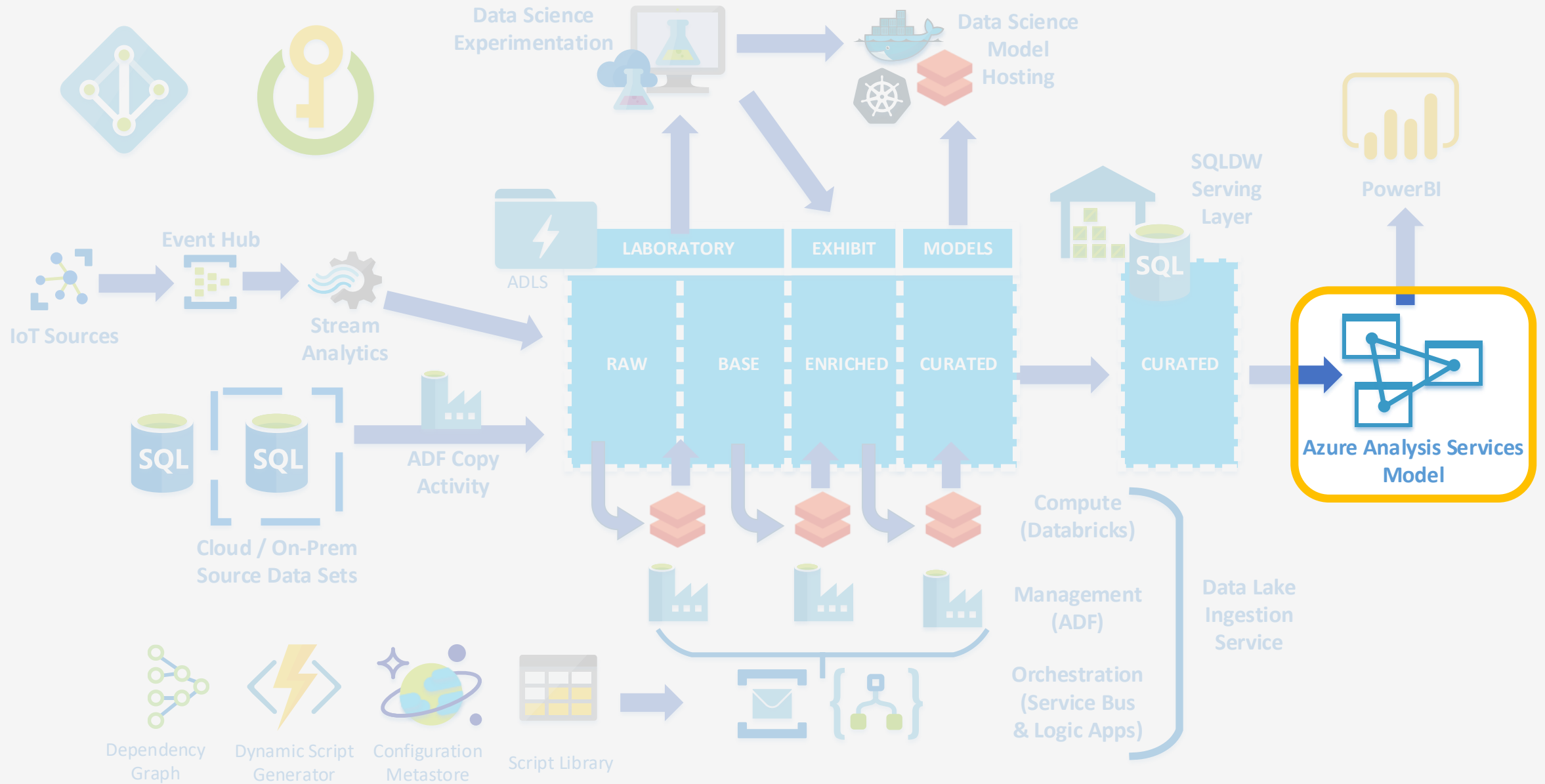




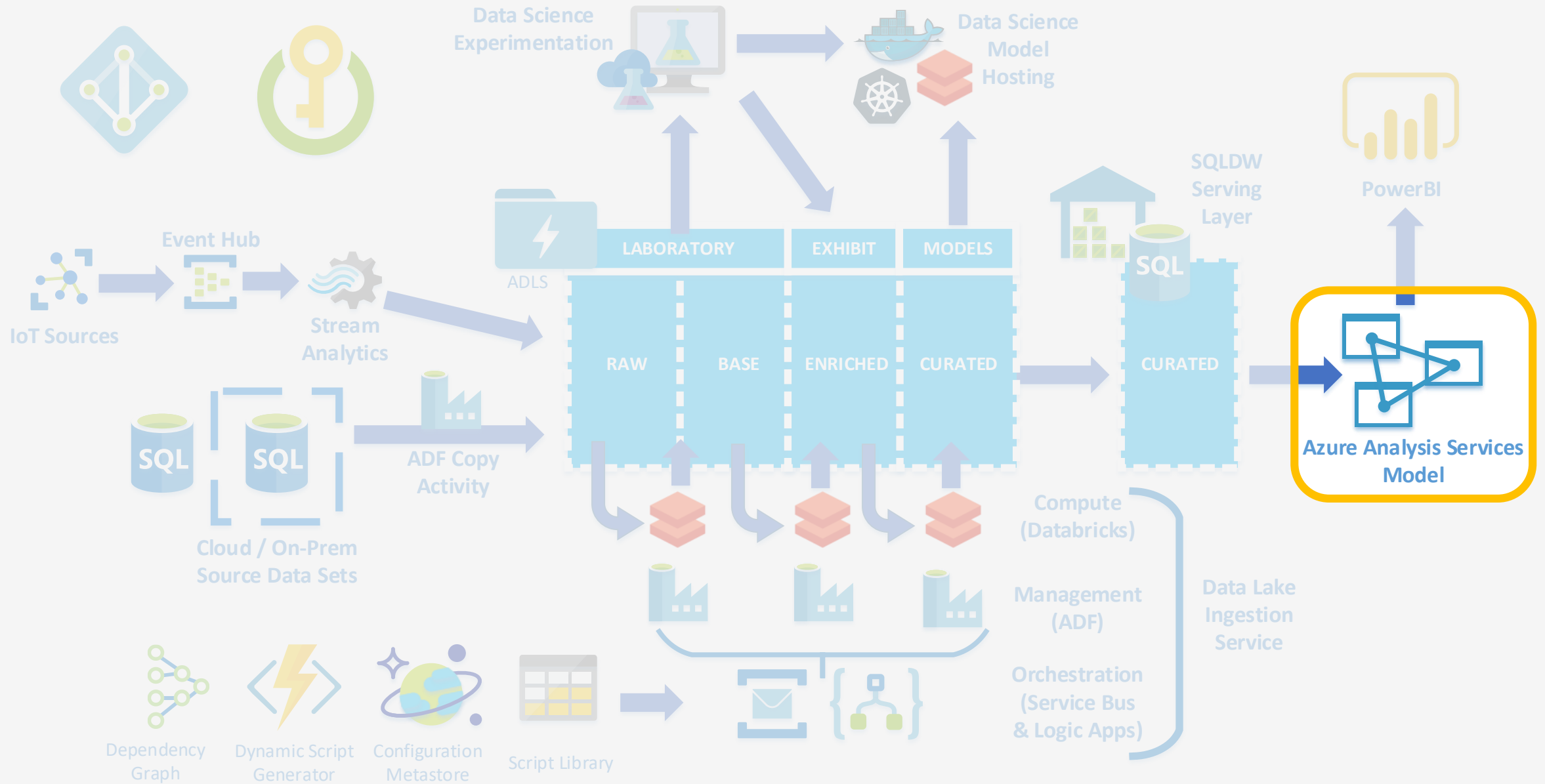
The Modern Data Warehouse



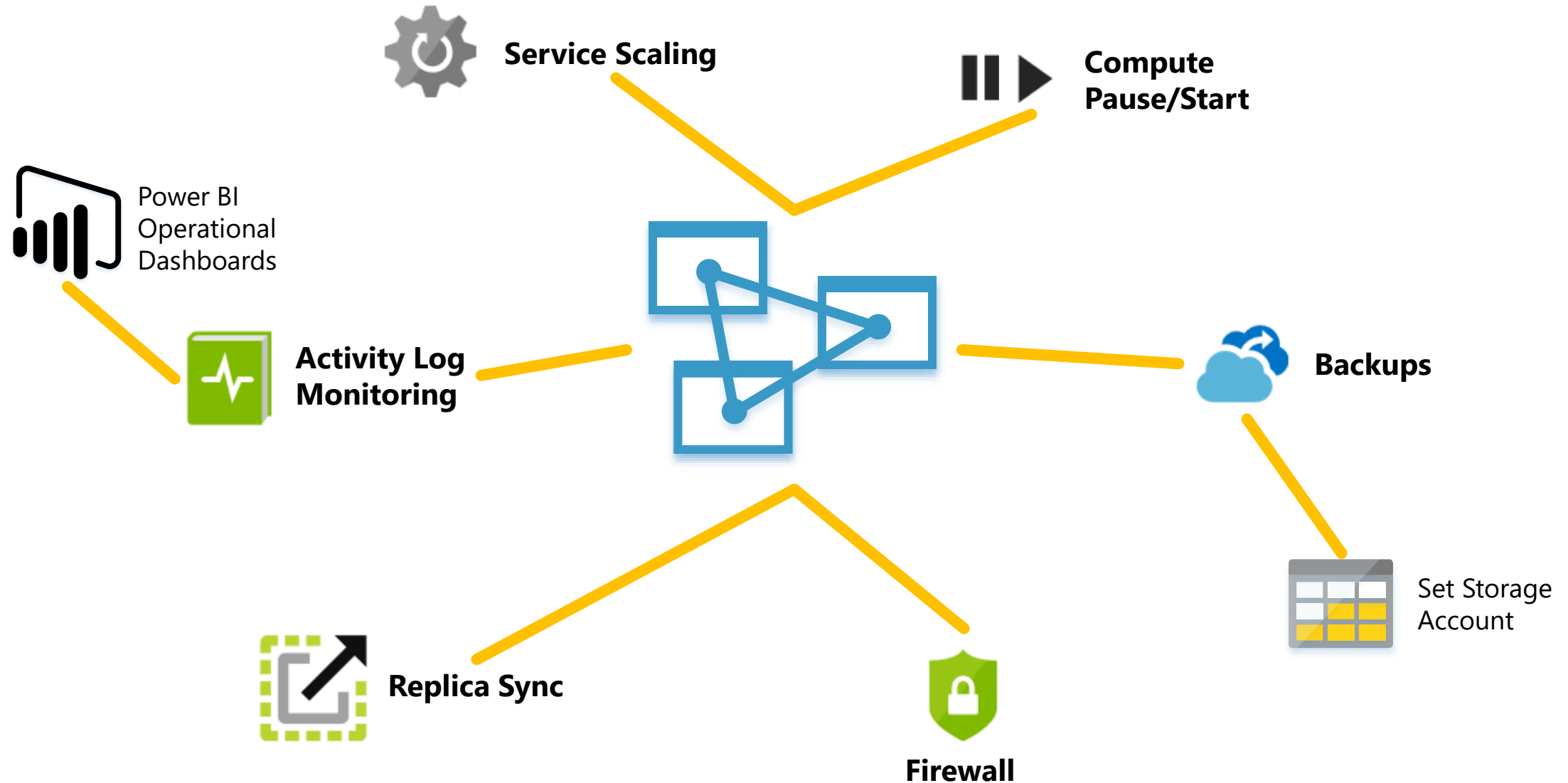
The Modern Data Warehouse



The Modern Data Warehouse



Administration Considerations – Azure Analysis Services





Agenda

Context

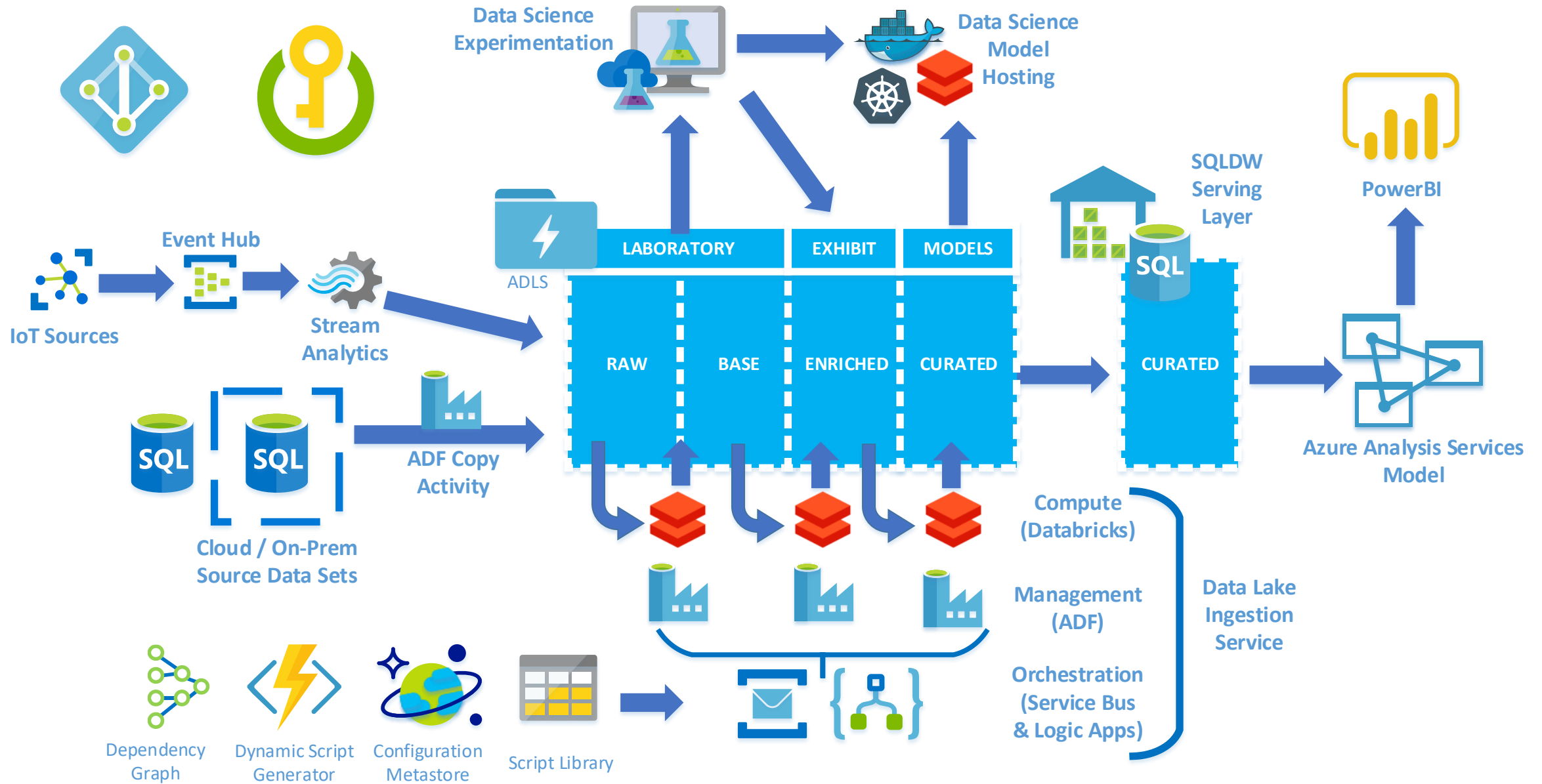
IT Industry Roles
Knowledge

The Modern Data Warehouse

Components
Considerations



Conclusions

The Modern Data Warehouse



Conclusion

DBA to DSA – a role Paul really wants to evolve to support cloud tech.

Task	 On Premises	 Azure
User & Service Account Management	<ul style="list-style-type: none">• Server login• DB user• DB roles	<ul style="list-style-type: none">• AAD service principal• Key Vault• Secret scopes
Job Scheduling	<ul style="list-style-type: none">• SQL Agent• ETA Window	<ul style="list-style-type: none">• Service orchestration• Un-pause service• Scale up service
Increase Compute/RAM	<ul style="list-style-type: none">• Allocate more cores• Increase instance RAM	<ul style="list-style-type: none">• VM Scale Set Cores Quota• Manage offline services during scaling.
SSAS Database Backup	<ul style="list-style-type: none">• Maintenance plan• SQL Agent• XMLA	<ul style="list-style-type: none">• Enable• Setup storage

Paul thinks we need a
DSA, do you agree?

Thanks for Listening

Paul Andrew



@MrPaulAndrew



Blog: <http://mrpaulandrew.com>

Email: paul@mrpaulandrew.com