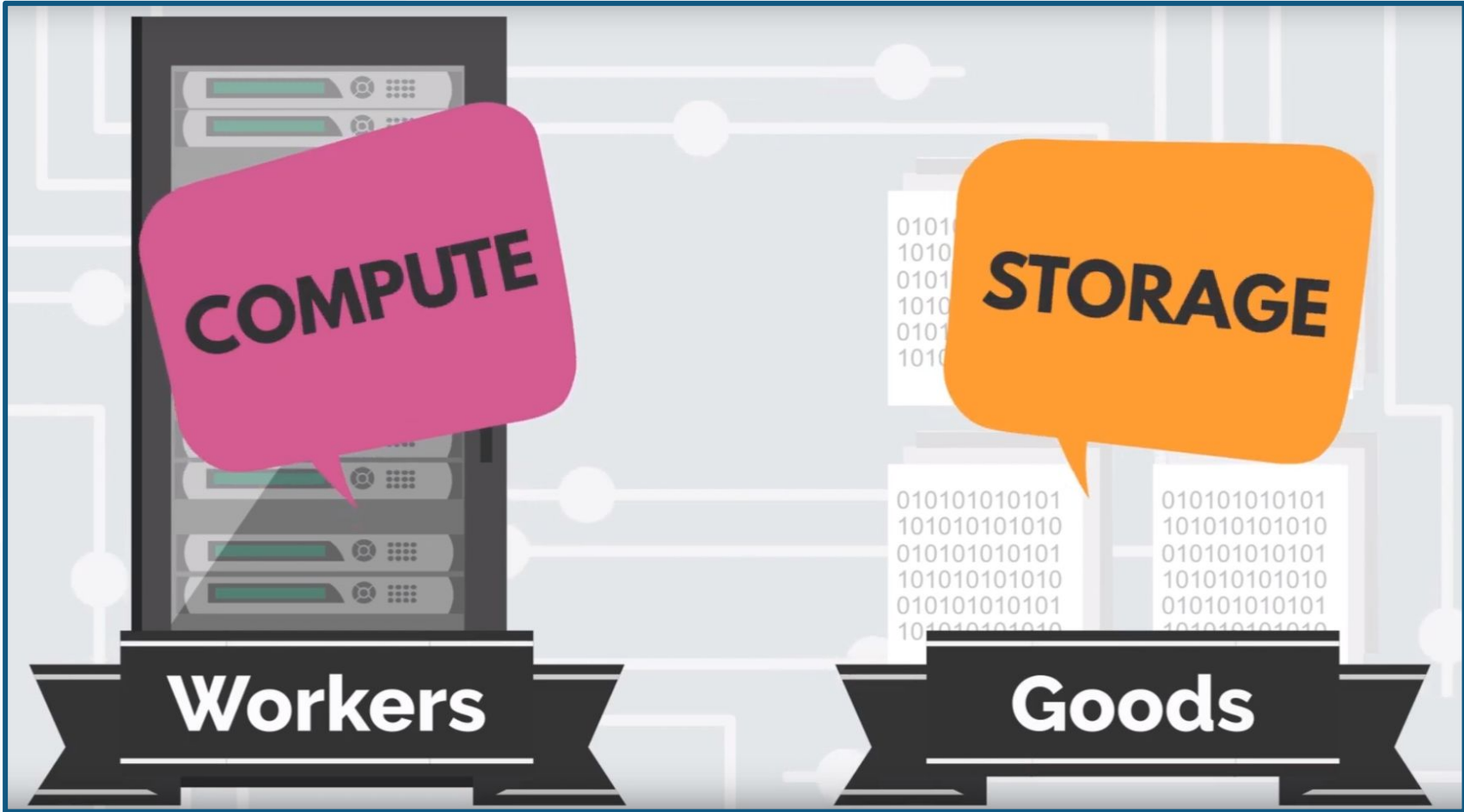


Lesson 7: Warehousing Concepts

Review Notes



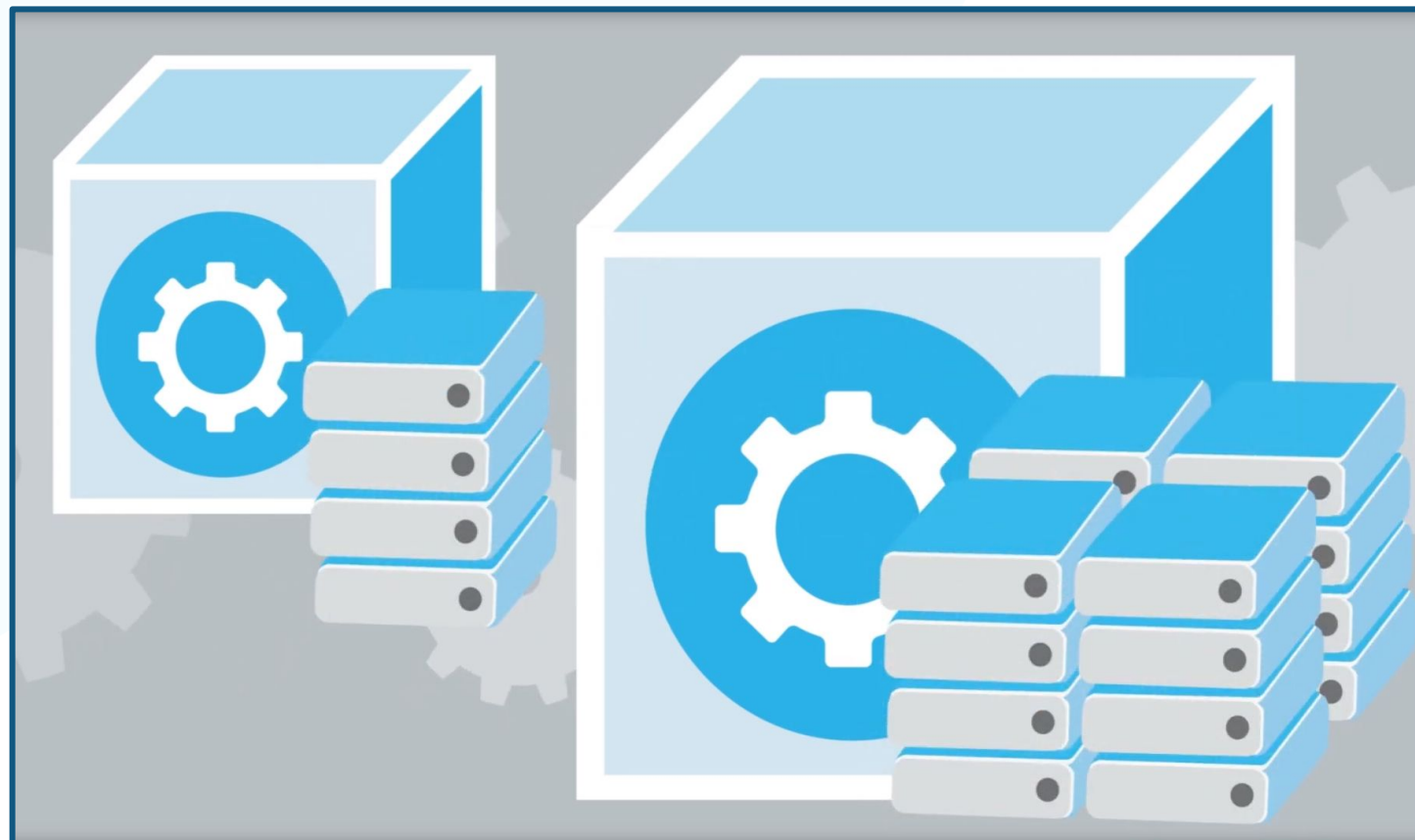
Real World Warehouse Term	Traditional Data Warehouse Term	Snowflake Alternate Term
Goods	Data	Storage
Workers	Servers	Compute

In Snowflake, when speaking about a Warehouse, we are not speaking of a place for storing data, we are always speaking about compute power that is brought to bear for data processing.

- Warehouse power is used for loading data.
 - Warehouse power is used for unloading data.
 - Warehouse power is used for querying data.
- Warehouses do NOT contain data.
 - Warehouses do NOT store data.
 - Warehouses are NOT “places” - they are collections of computing power.

Lesson 7: Warehousing Concepts

Review Notes



In both real-world warehouses and in Snowflake Warehouses, the bigger the warehouse, the bigger the team of workers doing the work

In Snowflake, each warehouse has a single cluster of servers. Think of this as being like a real-world warehouse that has a single workforce or team.

Larger warehouses have a more workers in their workforce. Smaller warehouses have fewer workers in their workforce. But each can still be thought of as having a single workforce.

In Snowflake you should think of a warehouse as having a single cluster. The number of servers in that cluster vary based on the warehouse's size designation.

This distinction is important for understanding multi-cluster warehouses later in this lesson.

Lesson 7: Warehousing Concepts

Review Notes



When creating a warehouse, you choose a size. The size of the warehouse you choose is the same as choosing the number of servers in the cluster.

1 Cluster of 4 Servers

1 Cluster of 64 Servers

Create Warehouse

Name *

NEW_WAREHOUSE

Size

X-Large (16 credits / hour)

X-Small (1 credit / hour)

Small (2 credits / hour)

Medium (4 credits / hour)

Large (8 credits / hour)

X-Large (16 credits / hour)

2X-Large (32 credits / hour)

3X-Large (64 credits / hour)

4X-Large (128 credits / hour)

Auto Suspend

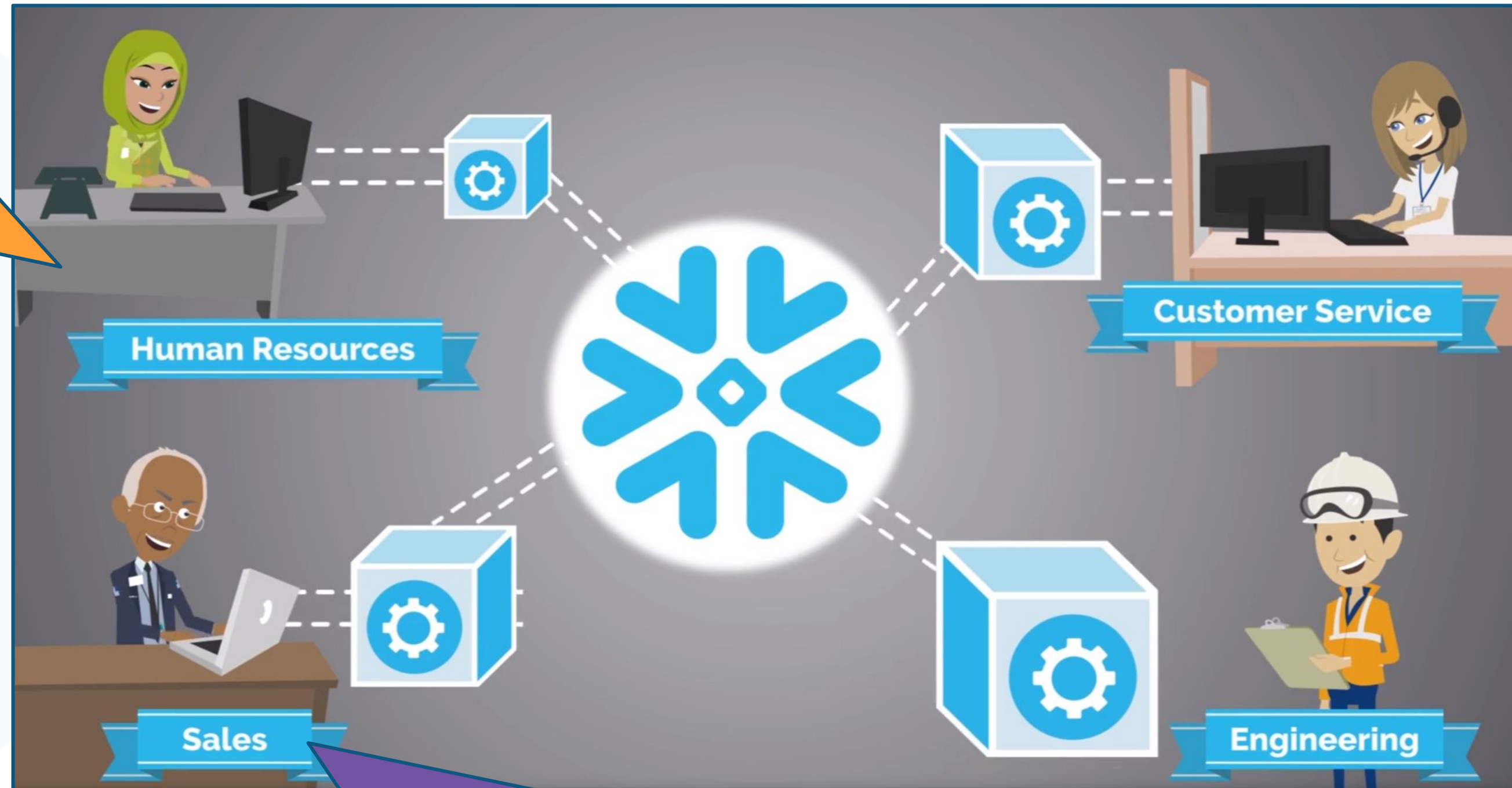
Comment

Show SQL

Lesson 7: Warehousing Concepts

Review Notes

HR might need a small warehouse that runs for long hours every day.



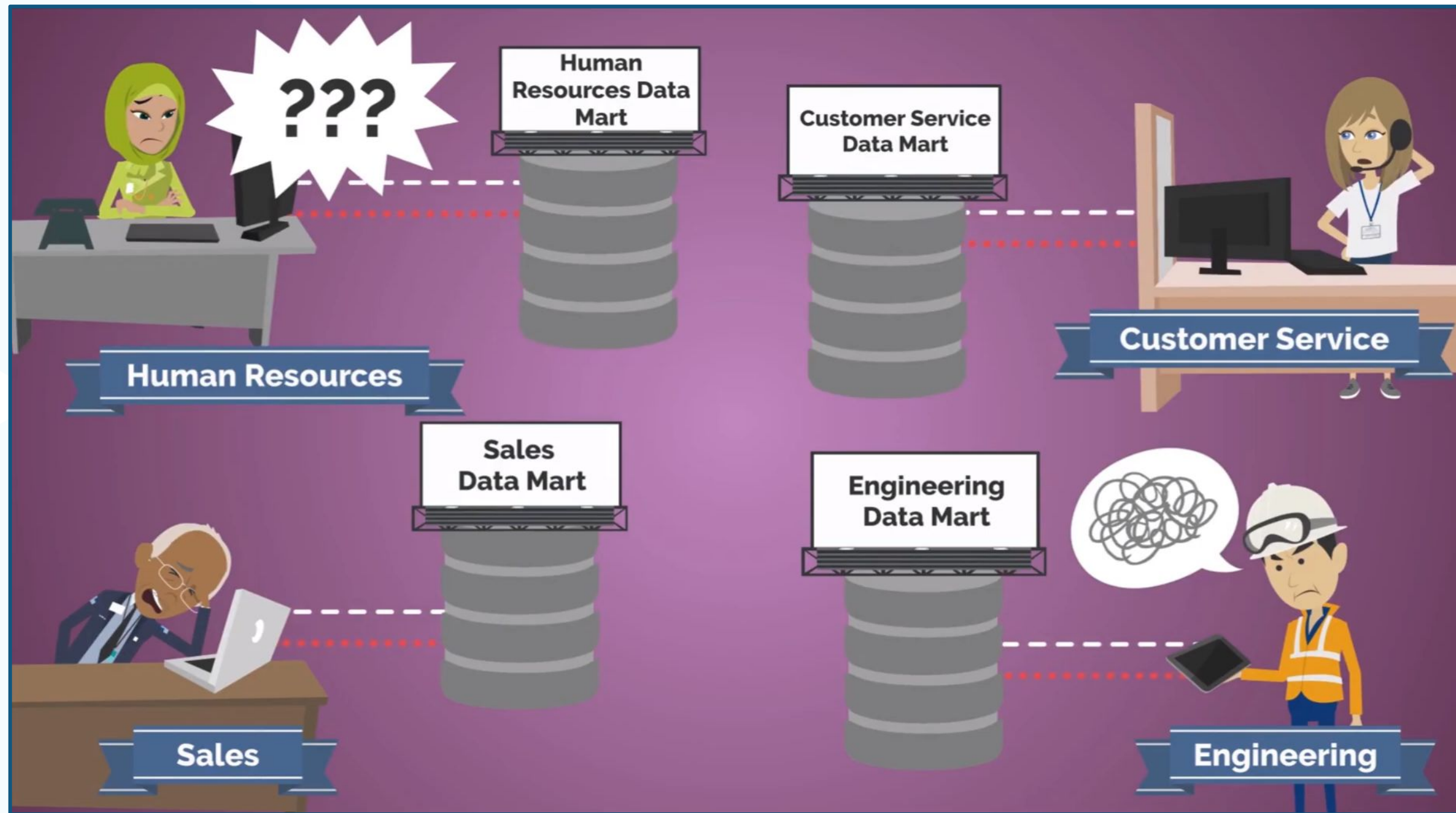
Sales might need several warehouses. A small warehouse for simple daily reports and a large warehouse for complex Month-End and Quarter-End reports.

Engineering might need a very large warehouse that rarely runs but processes very complex queries in short bursts of activity.



Lesson 7: Warehousing Concepts

Review Notes



Giving different departments their own warehouses is not the same as giving Data Marts to different departments.

Data Marts were generally silos of data, sectioned off to improve performance.

Siloed data often had to be replicated to multiple data marts when more than one department wanted the same data.

All Snowflake Warehouses have access to all the data, all the time.

Snowflake Warehouses are not storage locations, they are collections of workforce, or compute power so separate warehouses are nothing like storage silos.

Lesson 7: Warehousing Concepts

Review Notes



- You can change the size of a warehouse at any time.
- This process is manual.
- Moving to a larger warehouse is called “Scaling Up”.
- Moving to a smaller warehouse is called “Scaling Down.”

Create Warehouse

Name*

Size

Auto Suspend ☐

Comment

Show SQL

X-Small (1 credit / hour)

Small (2 credits / hour)

Medium (4 credits / hour)

Large (8 credits / hour)

X-Large (16 credits / hour)

2X-Large (32 credits / hour)

3X-Large (64 credits / hour)

4X-Large (128 credits / hour)

- Smaller warehouses cost less while running for the same amount of time as a larger warehouse.
- Another way to potentially control costs is to have warehouses set to “Auto-Suspend” after a period of time and only resume when a query is run (Auto Resume).



Lesson 7: Warehousing Concepts

Review Notes





- Snowflake offers another, more advanced option for optimizing warehouse usage. This option is called “Elastic Data Warehousing”
- Elastic Data Warehousing has two additional components or facets (that non-elastic warehouses do not have):
 - Multi-Cluster options
 - Scaling Policy options
- Snowflake Standard and Premier Editions do NOT include Elastic Data Warehousing.




Lesson 7: Warehousing Concepts

Review Notes







Databases




Shares




Warehouses



Worksheets



History





Account


Warehouses


Last re


Manage your warehouses from this page. To operate on your data, you need to create one or more warehouses.


 Create...

 Configure...

 Suspend...

 Resume...

 Drop...

 Transfer Ownership

Status	Warehouse Name	Size	Clusters	Scaling Policy	Run...	Que...
Suspended	COMPUTE_WH	X-Large	min: 1, max: 2	Standard	0	0

With Elastic Warehousing you have to choose cluster minimum and maximum for each warehouse.

In this example, an X-Large Warehouse has a minimum cluster count of 1 and maximum cluster count of 2.

Since an X-Large Cluster is made up of 16 servers, the warehouse shown will use 16 servers at a minimum and 32 at its maximum.

Lesson 7: Warehousing Concepts

Review Notes

Create Warehouse

Name *

Size

X-Large (16 credits / hour)

Learn more about virtual warehouse sizes

here

Maximum Clusters

2

Multi-cluster warehouses improve the query throughput for high concurrency workloads.

Minimum Clusters

1

The number of active clusters will vary between the specified minimum and maximum values, based on number of concurrent users/queries.

Scaling Policy

Standard

The policy used to automatically start up and shut down clusters.

Auto Suspend

10 minutes

The maximum idle time before the warehouse will be automatically suspended.

☒ Auto Resume

?

Comment

Show SQL

Cancel

Finish

With Elastic Warehousing you will also need to choose a Scaling Policy.

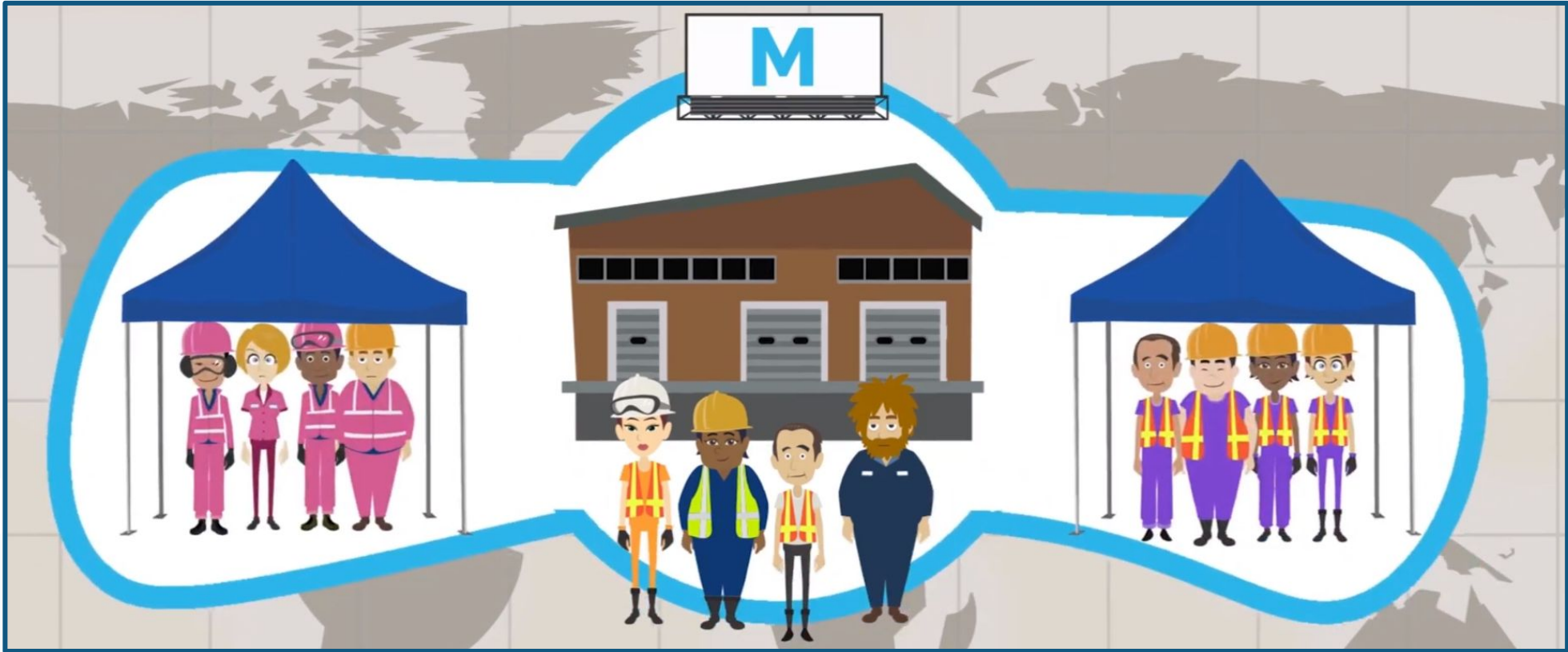
A Scaling Policy helps Snowflake decide when to move from a single cluster, to more than one and when to snap back to a single cluster.

Read more about Scaling Policies in the Snowflake online documentation [here](#).



Lesson 7: Warehousing Concepts

Review Notes



Remember that we defined each warehouse as having one cluster with multiple servers. With Multi-Cluster Elastic Warehousing, that definition fails.

Once we begin to speak of Elastic Warehousing, think of the warehouse not as having “one and only one” cluster, but as having a single “base cluster” that defines the number of servers in all other clusters.

In this scenario, when clusters are added, each one has the same number of servers within it. So a medium-sized warehouse adds or removes servers in groups of four every time because the base cluster is four servers.

Warehouse Name	Warehouse Size	Minimum Cluster Setting	Maximum Cluster Setting	Number of Servers Used when Min Clusters Activated	Number of Servers Used when MAX Clusters Activated
MY_LITTLE_OBSTINATE_WAREHOUSE	X-SMALL (1)	1	1	1 server used when 1 cluster is active	1 server used when 1 cluster is active
MY_MIDDLING_SUPER_ELASTIC_WH	MEDIUM (4)	1	10	4 servers used when 1 cluster is active	40 servers used when 10 clusters are active
MY_REALLY_BIG_SOMEWHAT_FLEXIBLE_WH	2X-LARGE (32)	1	3	32 servers used when 1 cluster is active	96 servers used when 3 clusters are active