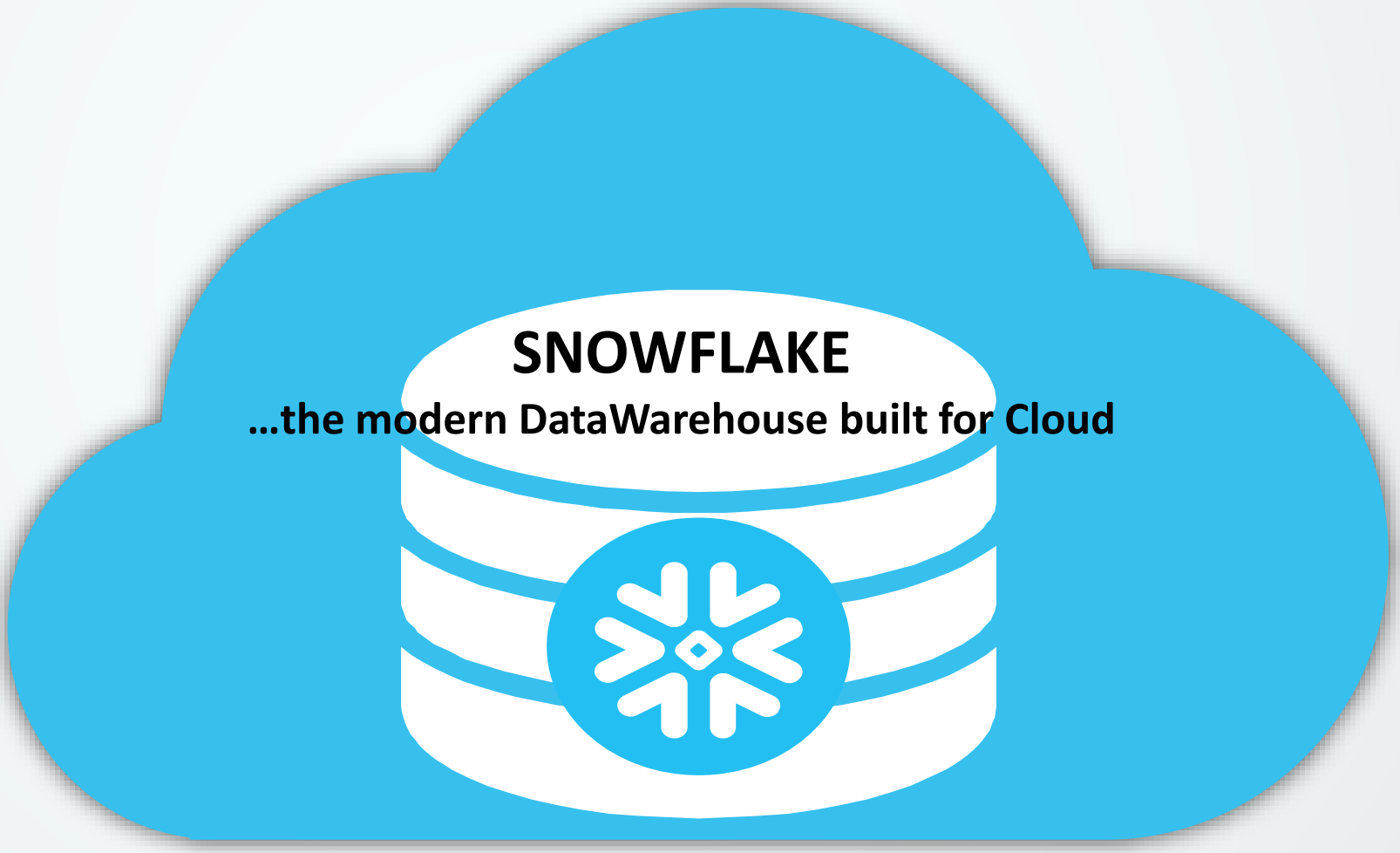


“Simplicity is the ultimate sophistication”

Leonardo da Vinci

Day-3

Snowflake Editions

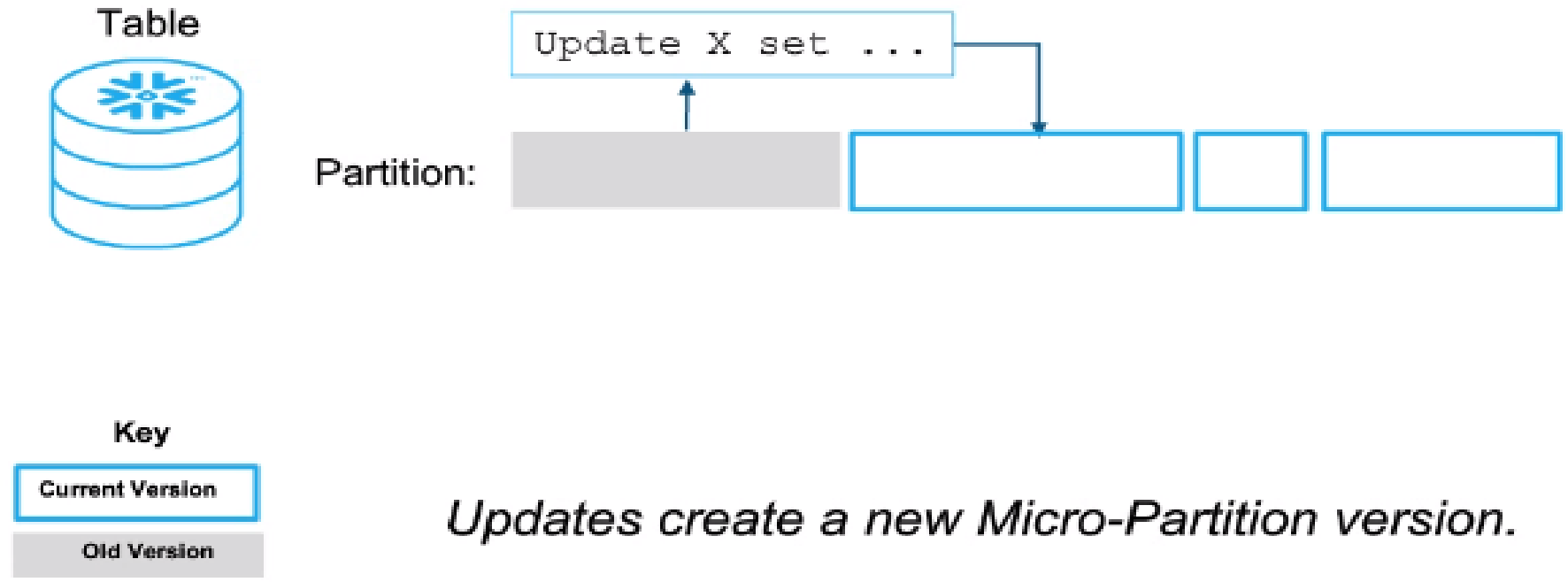
- ❑ AWS Editions and pricing
- ❑ Azure Editions and pricing
- ❑ GCP Editions and pricing
- ❑ Standard Edition
- ❑ Enterprise Edition
- ❑ Business Critical Edition
- ❑ Virtual Private Snowflake(VPS) Edition

Snowflake Editions

STANDARD	ENTERPRISE	BUSINESS CRITICAL	VIRTUAL PRIVATE SNOWFLAKE
Complete SQL Datawarehouse	Standard+	Enterprise+	Business Critical+
Secure Data sharing	Multi-cluster warehouse	HIPPA support	Customer-dedicated virtual servers wherever the encryption key is in memory
Premier Support 24*365	Up to 90 days of time travel	PCI compliance	
1 day of time travel	Annually rekey encrypted data	Data encryption everywhere	Customer-dedicated metadata store
Enterprise-grade encryption	Materialized Views	Tri-Secure secure	
Dedicated Virtual Warehouse	Search Optimization Service	AWS PrivateLink Support	
Federated Authentication	Dynamic Data Masking	AZURE PrivateLink Support	
Database Replication	External Data Tokenization	Database failover and failback	
External Functions		External functions – AWS API	
Snowsight		Gateway Private Endpoints Support	
Create your own data exchange			
Data marketplace access			

Micro-Partition (Falcon De Neigh) & Clustering

DATA IS IMMUTABLE



Updates create a new Micro-Partition version.

Data organized in Micro-partitions in hybrid columnar format.

DATA IN MICRO-PARTITIONS

Date	Store	Cust ID	Amount
01-Jan	Ingatestone	54612	$ 2.75
02-Jan	Ealing	19286	$ 19.00
01-Jan	Ealing	29625	$ 4.99
03-Jan	Ealing	89281	$ 14.99
01-Jan	Ealing	12395	$ 40.25
03-Jan	Ingatestone	19285	$ 18.50
01-Jan	Ealing	39855	$ 50.15
01-Jan	Ingatestone	87252	$ 40.25
02-Jan	Windsor	98261	$ 49.99
03-Jan	Windsor	86542	$ 7.50
03-Jan	Ealing	20202	$ 27.70
02-Jan	Windsor	12662	$ 17.50

Micro Partition 1:

Date	Store	Cust ID	Amount
01-Jan	Ingatestone	54612	$ 2.75
02-Jan	Ealing	19286	$ 19.00
01-Jan	Ealing	29625	$ 4.99
03-Jan	Ealing	89281	$ 14.99
01-Jan	Ealing	12395	$ 40.25

Micro Partition 2:

Date	Store	Cust ID	Amount
03-Jan	Ingatestone	19285	$ 18.50
01-Jan	Ealing	39855	$ 50.15
01-Jan	Ingatestone	87252	$ 40.25

Micro Partition 3:

Date	Store	Cust ID	Amount
02-Jan	Windsor	98261	$ 49.99
03-Jan	Windsor	86542	$ 7.50
03-Jan	Ealing	20202	$ 27.70
02-Jan	Windsor	12662	$ 17.50

```
Where DATE = '03-JAN'
```

Micro-Partition & Clustering

"NATURAL" DATA CLUSTERING

Date	Store	Cust ID	Amount
01-Jan	Ealing	12395	$ 40.25
01-Jan	Ealing	39855	$ 50.15
01-Jan	Ealing	29625	$ 4.99
01-Jan	Ingatestone	87252	$ 40.25
01-Jan	Ingatestone	54612	$ 2.75
02-Jan	Ealing	19286	$ 19.00
02-Jan	Windsor	98261	$ 49.99
02-Jan	Windsor	12662	$ 17.50
03-Jan	Ealing	20202	$ 27.70
03-Jan	Ealing	89281	$ 14.99
03-Jan	Ingatestone	19285	$ 18.50
03-Jan	Windsor	86542	$ 7.50

Micro Partition 1:

Date	Store	Cust ID	Amount
01-Jan	Ealing	12395	$ 40.25
01-Jan	Ealing	39855	$ 50.15
01-Jan	Ealing	29625	$ 4.99
01-Jan	Ingatestone	87252	$ 40.25
01-Jan	Ingatestone	54612	$ 2.75

Micro Partition 2:

Date	Store	Cust ID	Amount
02-Jan	Ealing	19286	$ 19.00
02-Jan	Windsor	98261	$ 49.99
02-Jan	Windsor	12662	$ 17.50

Micro Partition 3:

Date	Store	Cust ID	Amount
03-Jan	Ealing	20202	$ 27.70
03-Jan	Ealing	89281	$ 14.99
03-Jan	Ingatestone	19285	$ 18.50
03-Jan	Windsor	86542	$ 7.50

```
Where DATE = '03-JAN'
```

Micro-Partition & Clustering

CLUSTER BY CUST_ID

Date	Store	Cust ID	Amount
01-Jan	Ealing	12395	$ 40.25
01-Jan	Ealing	39855	$ 50.15
01-Jan	Ealing	29625	$ 4.99
01-Jan	Ingatestone	87252	$ 40.25
01-Jan	Ingatestone	54612	$ 2.75
02-Jan	Ealing	19286	$ 19.00
02-Jan	Windsor	98261	$ 49.99
02-Jan	Windsor	12662	$ 17.50
03-Jan	Ealing	20202	$ 27.70
03-Jan	Ealing	89281	$ 14.99
03-Jan	Ingatestone	19285	$ 18.50
03-Jan	Windsor	86542	$ 7.50

Date	Store	Cust ID	Amount
01-Jan	Ealing	12395	$ 40.25
02-Jan	Windsor	12662	$ 17.50
03-Jan	Ingatestone	19285	$ 18.50
02-Jan	Ealing	19286	$ 19.00
03-Jan	Ealing	20202	$ 27.70

Micro Partition 1:
Range: 12,395 – 20,202

Date	Store	Cust ID	Amount
01-Jan	Ealing	29625	$ 4.99
01-Jan	Ealing	39855	$ 50.15
01-Jan	Ingatestone	54612	$ 2.75

Micro Partition 2:
Range: 29,625 – 54,612

Date	Store	Cust ID	Amount
03-Jan	Windsor	86542	$ 7.50
01-Jan	Ingatestone	87252	$ 40.25
03-Jan	Ealing	89281	$ 14.99
02-Jan	Windsor	98261	$ 49.99

Micro Partition 3:
Range: 86,542 – 98,261

```
Where CUST_ID = 39855
```

PARTITION ELIMINATION

Uses Min/Max Values

Physical Micro-Partitions

	Date	Store	Cust ID	Amount
1	1-Jan	Ealing	12395	$40.25
	1-Jan	Ealing	93855	$50.15
	1-Jan	Ealing	29625	$4.99
	1-Jan	Ingatestone	87252	$40.25
	1-Jan	Ingatestone	54612	$2.75
2	2-Jan	Ealing	19286	$19.00
	2-Jan	Windsor	98262	$49.99
	2-Jan	Windsor	12662	$17.50
3	3-Jan	Ealing	20202	$27.70
	3-Jan	Ealing	89281	$14.99
	3-Jan	Ingatestone	19285	$18.50
	3-Jan	Windsor	86542	$7.50

Metadata Entries (in Memory)

		Date	Store	Cust ID	Amount
1	Min	1-Jan	Ealing	12395	$2.75
	Max	1-Jan	Ingatestone	93855	$50.15
2	Min	2-Jan	Ealing	12662	$17.50
	Max	2-Jan	Windsor	98262	$49.99
3	Min	3-Jan	Ealing	19285	$7.50
	Max	3-Jan	Windsor	86542	$14.99

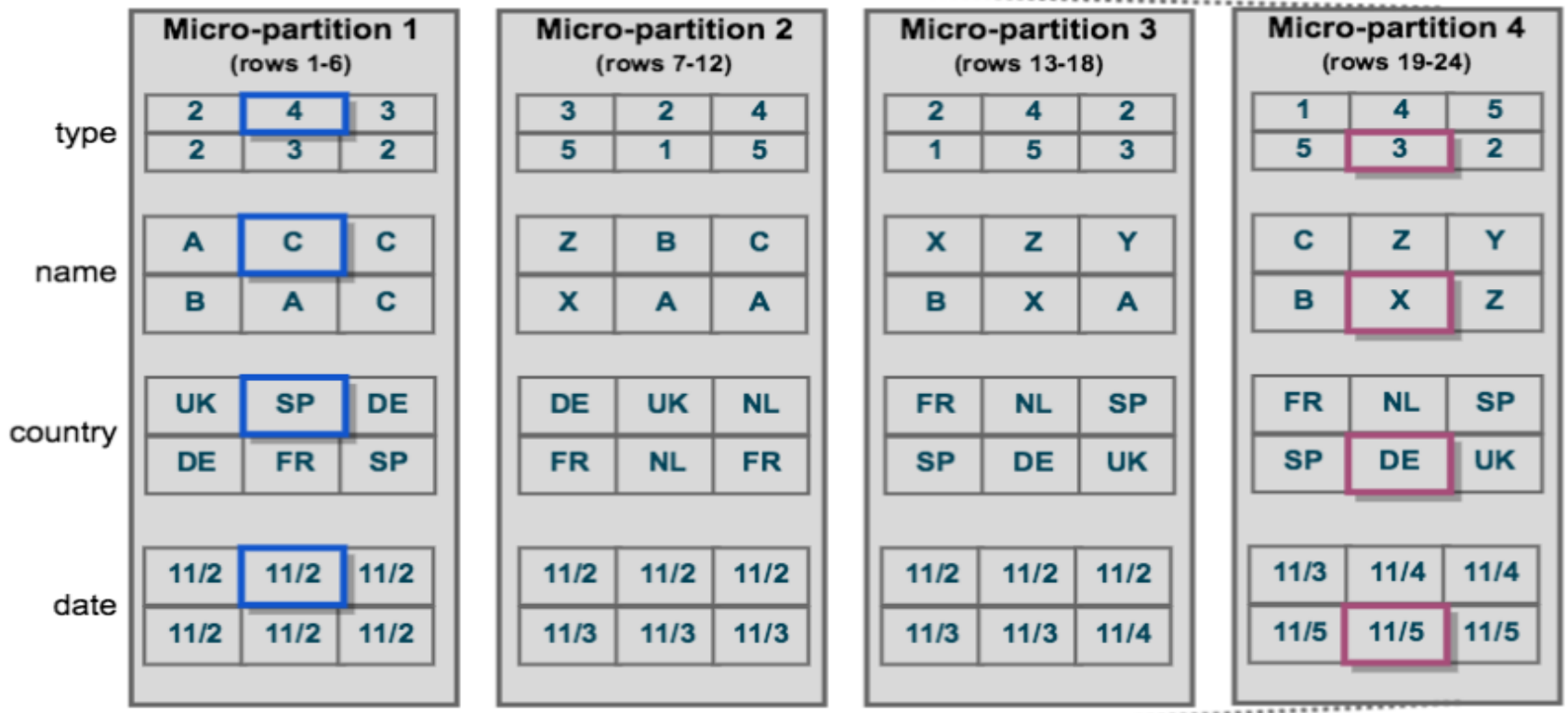
```
Where DATE = 1-JAN
   or AMOUNT > $40.00
   or STORE = 'Windsor'
```

Micro-Partition & Clustering

Logical Structure

type	name	country	date
2	A	UK	11/2
4	C	SP	11/2
3	C	DE	11/2
2	B	DE	11/2
3	A	FR	11/2
2	C	SP	11/2
3	Z	DE	11/2
2	B	UK	11/2
4	C	NL	11/2
5	X	FR	11/3
1	A	NL	11/3
5	A	FR	11/3
2	X	FR	11/2
4	Z	NL	11/2
2	Y	SP	11/2
1	B	SP	11/3
5	X	DE	11/3
3	A	UK	11/4
1	C	FR	11/3
4	Z	NL	11/4
5	Y	SP	11/4
5	B	SP	11/5
3	X	DE	11/5
2	Z	UK	11/5

Physical Structure



Micro-Partition & Clustering

Points to be noted for Micro-Partitions

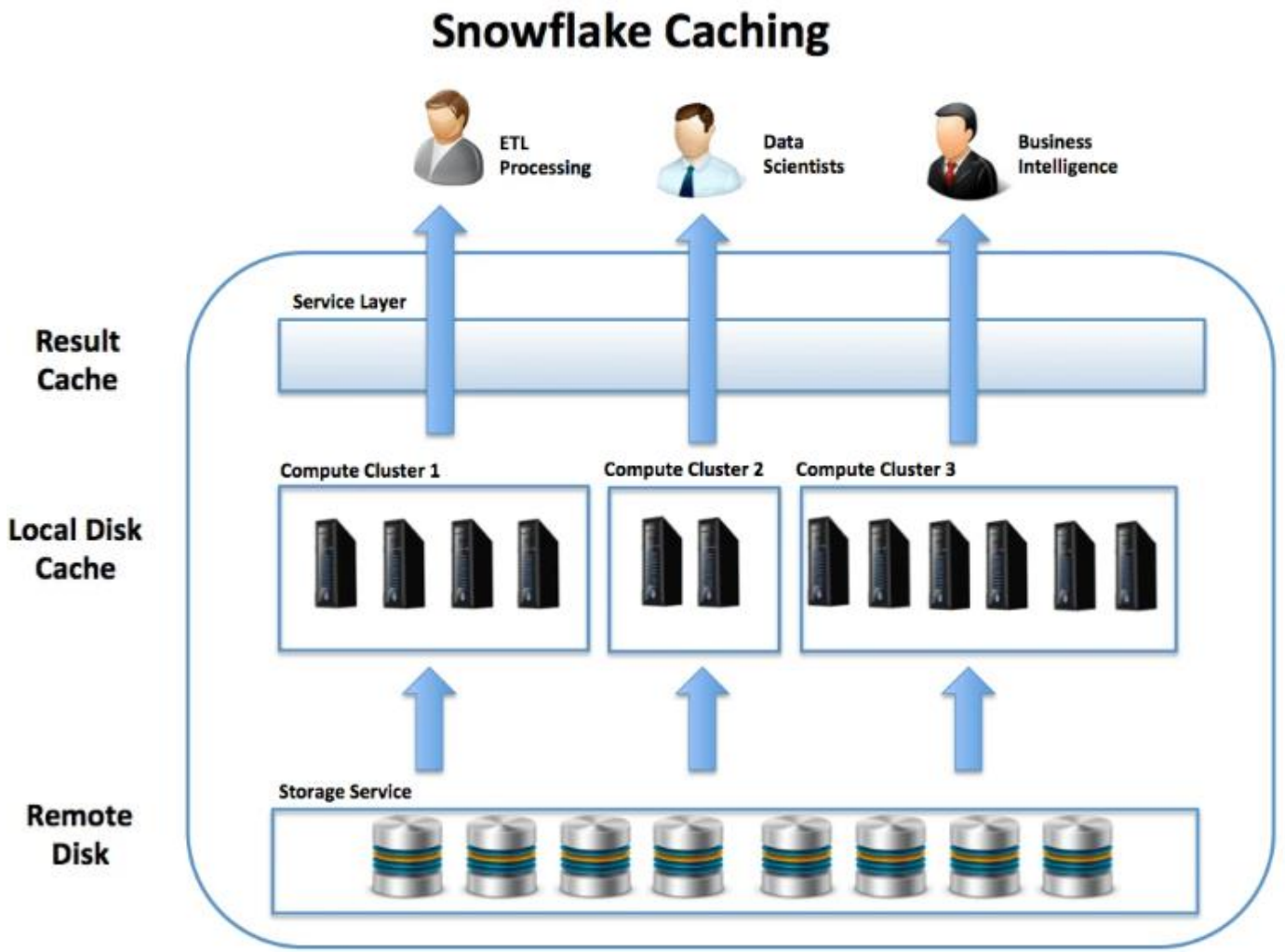
- Micro-partitioning is automatically performed on all Snowflake tables
- Micro-partitions don’t need to be explicitly defined up-front or maintained by users
- Each micro-partition contains between 50 MB and 500 MB of uncompressed data
- Snowflake stores metadata about all rows stored in a micro-partition
  - The range of values for each of the columns in the micro-partition
  - The number of distinct values
  - Additional properties used for both optimization and efficient query processing.
- Columns are also compressed individually within micro-partitions

Micro-Partition & Clustering

Points to be noted for Clustering

- Once Clustering key has been defined on a table, no additional administration is required
- Snowflake recommends a maximum of 3 columns (or expressions) per key for multi-terabyte tables
- Clustering Improves scan efficiency in queries by skipping data that does not match filtering predicates
- Re-clustering in Snowflake is automatic; no maintenance is needed
- Re-clustering also results in storage costs, consumes credits

Snowflake Cache



Snowflake Cache

Result Cache

- All query results stored for 24 hours unless underlying data changes
- Identical queries are returned instantly without requiring compute

Metadata Cache

- Improves compile times for queries against commonly used tables

Virtual Warehouse Cache

- Data loaded into warehouses is stored in local SSD storage
- Cache entries are invalidated if underlying data changes