

Amazon Redshift

Table of Contents



- 1. 2 Versions or Mode
- 2. What is Amazon Redshift?
- 3. Amazon Redshift Serverless
- 4. Amazon Redshift provisioned clusters overview
- 5. Amazon Redshift Provisioned Cluster Database
- 6. Amazon Redshift Serverless Key Concepts
- 7. Understanding Amazon Redshift Serverless Capacity
- 8. Serverless vs. Provisioned

2 Versions or Mode

Redshift Provisioned Mode

Redshift Serverless

1. 🏭 Petabyte-scale data warehouse service

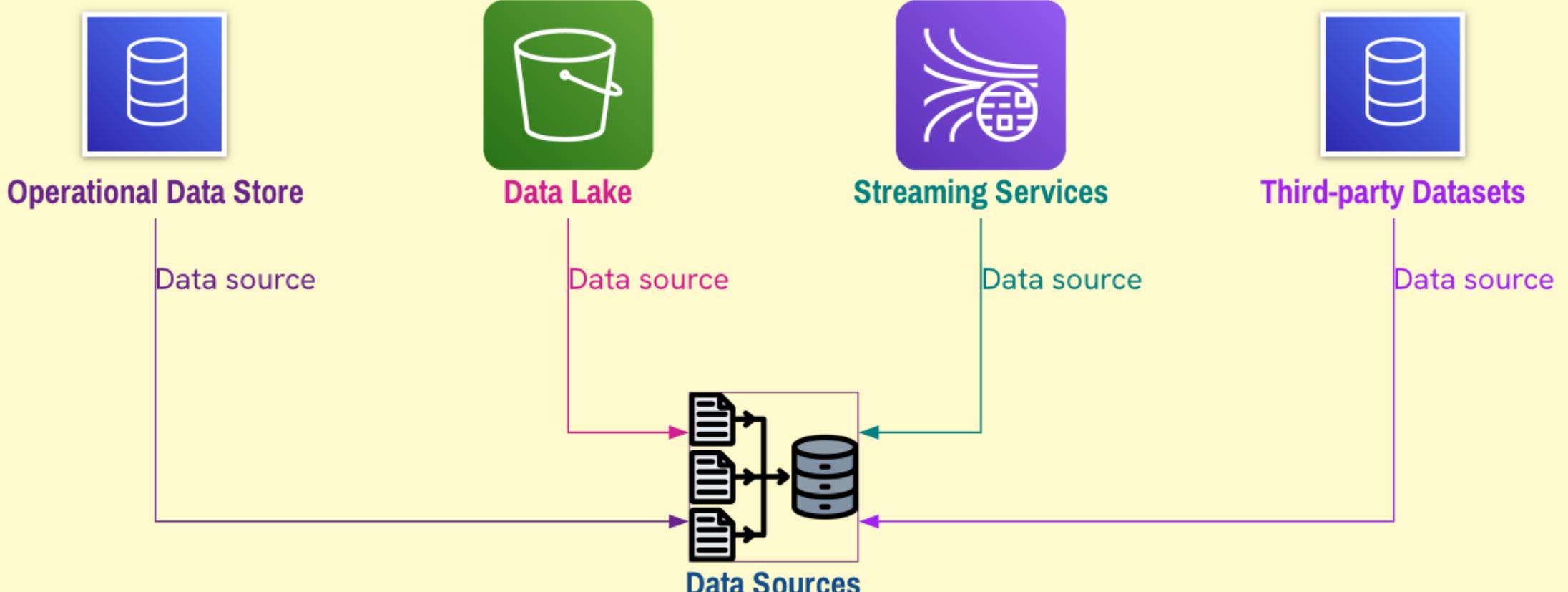
Cloud-based, fully managed

Large-scale data processing and analysis

What is Amazon Redshift?

Amazon Redshift

Data Sources



5. Provides data from



Amazon Redshift

1. Offers



Redshift Serverless

2. Uses



MPP Engine

3. Implements



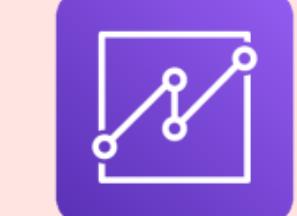
AutoMaterialized Views

4. Provides



Query Editor v2

6. Integrates with



QuickSight

7. Integrates with



Athena

Integrated AWS Services



SageMaker

8. Supports ML with



EMR (Apache Spark)

9. Enables analytics with



Glue Data Catalog

What is Amazon Redshift?

3. 💰 Pay-per-use pricing model

charges only for active usage

No upfront costs

Cost-effective for varying workloads

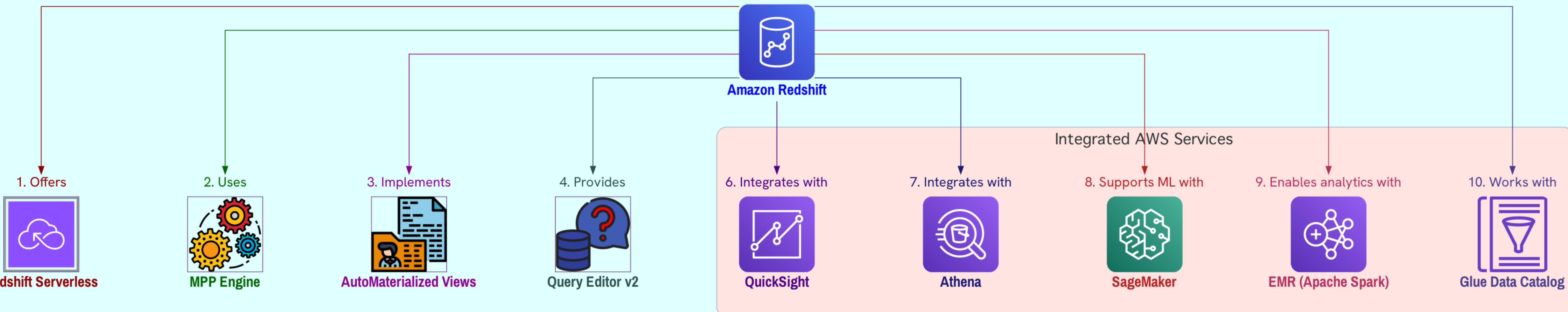
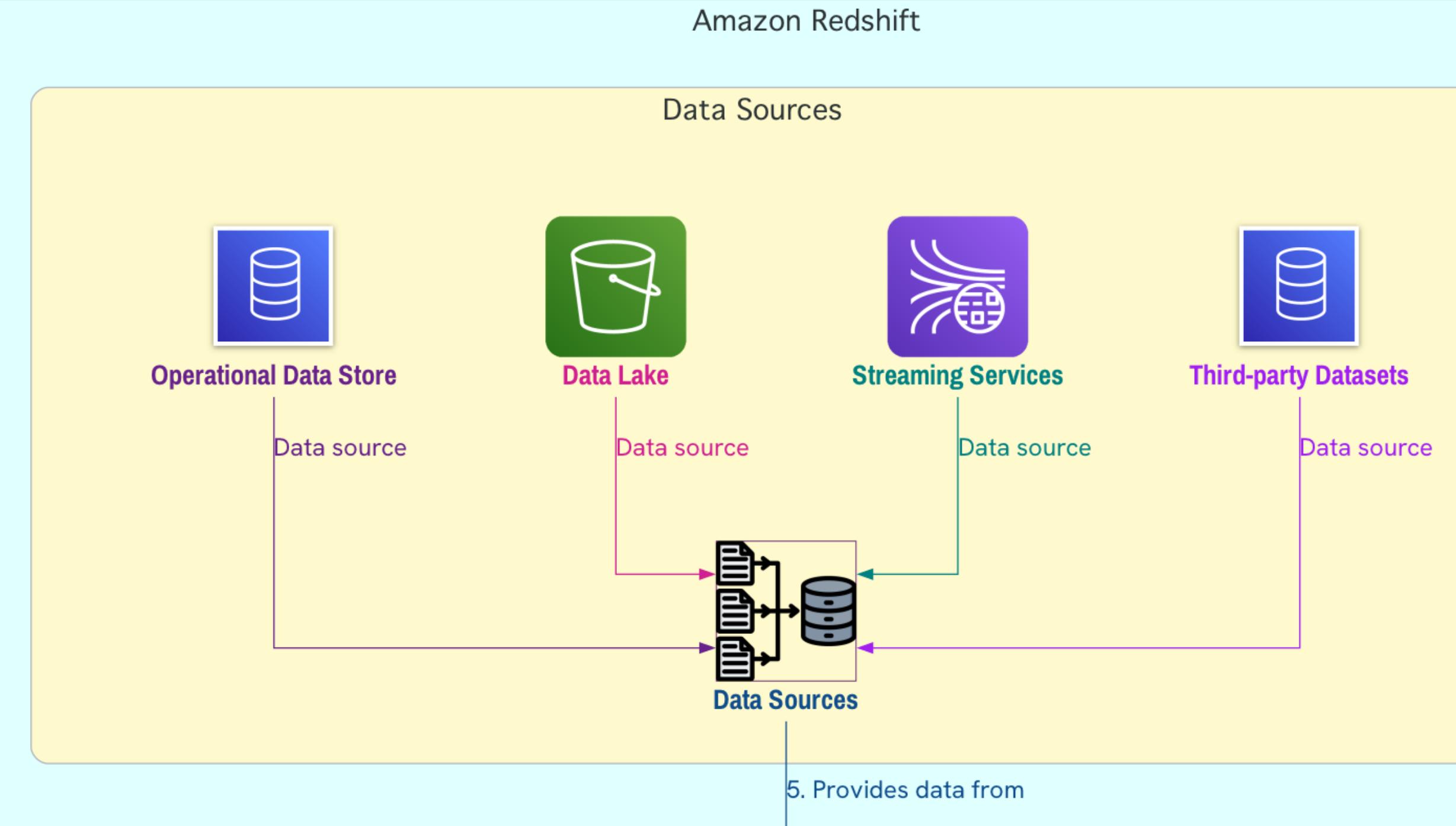
4. 🔎 Fast query performance with SQL tools

⚡ Immediate data loading and querying

💻 Amazon Redshift query editor v2

📊 Compatible business intelligence tools

📈 Fast performance across all dataset sizes



What is Amazon Redshift?

5. 🔒 Secure data access from multiple sources

Secure access, combine, and share data

Operational data stores

Data lakes

Streaming services

Third-party datasets

Comprehensive data analysis

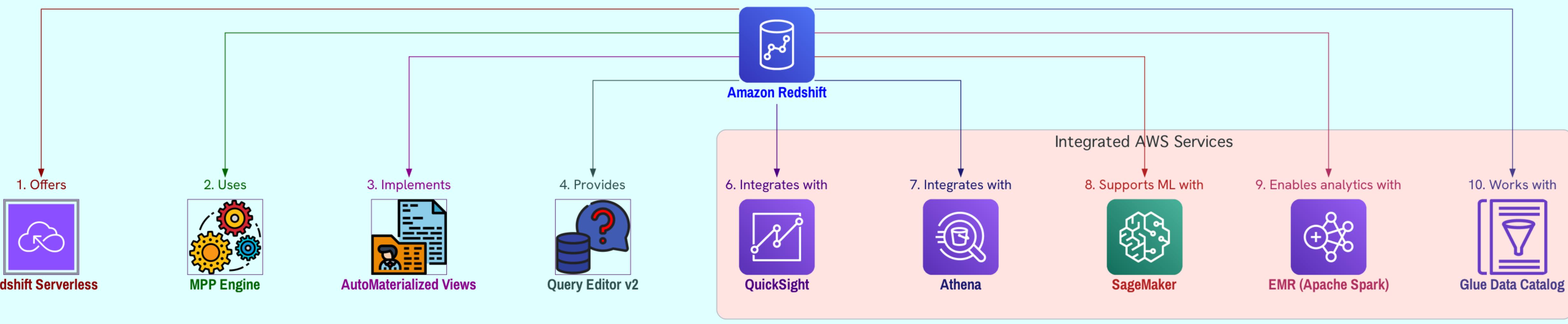
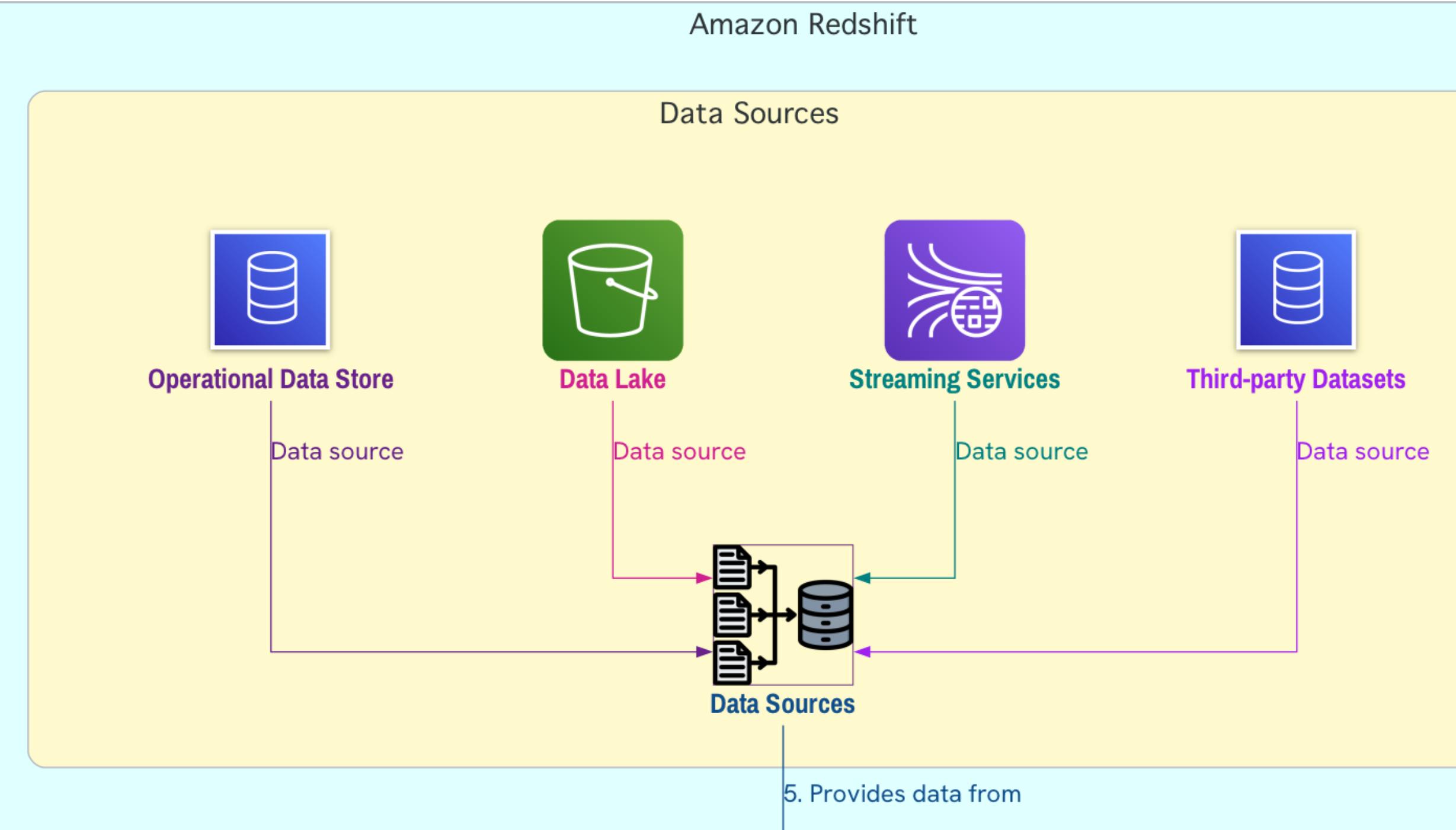
6. 🤝 Deep integration with AWS services

Zero-ETL approaches

Near real-time analytics

SQL-based machine learning

Apache Spark analytics



What is Amazon Redshift?

7. 🧑 Zero-administration environment

Easily scalable

🛠 Reduces operational overhead

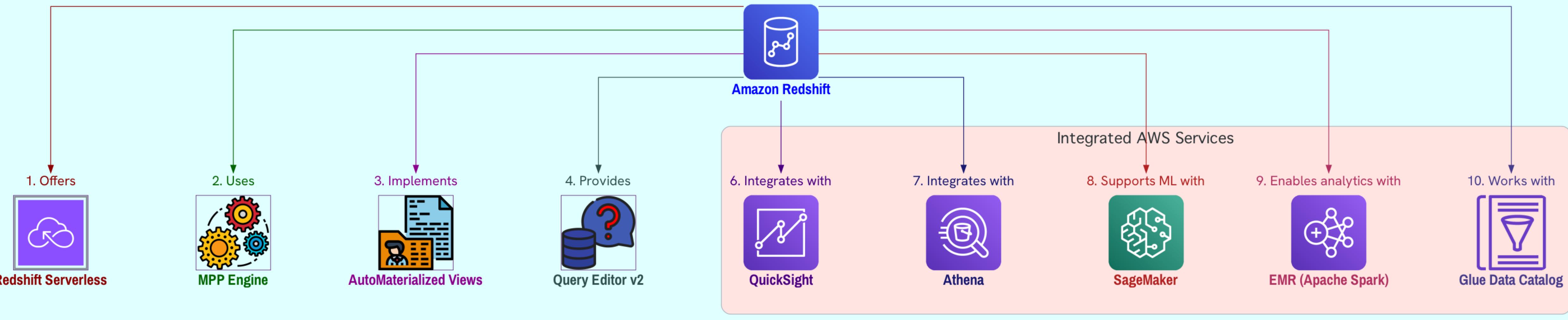
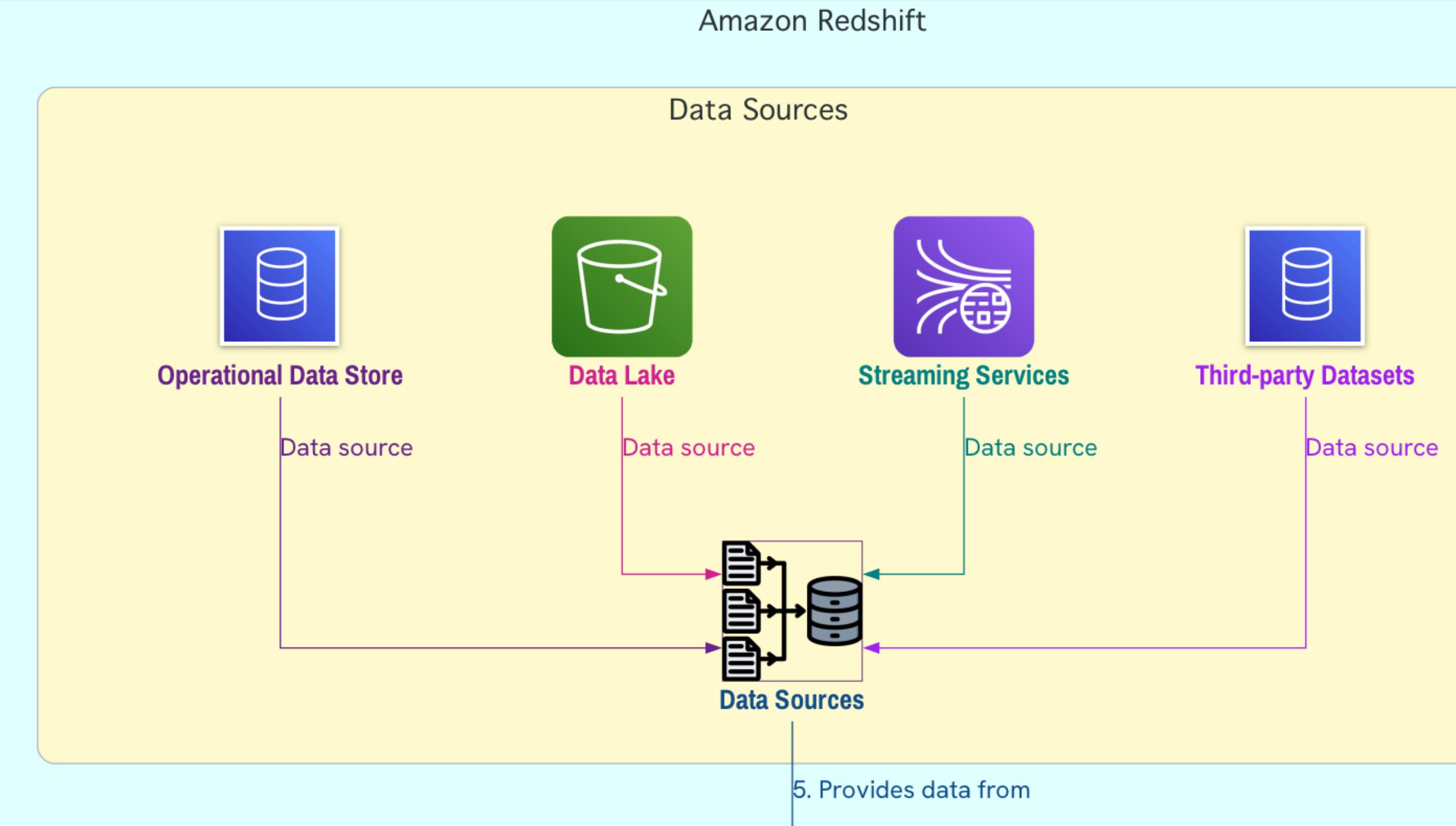
🎯 Focus on data analysis

8. ⚡ MPP engine for efficient scaling

12 34 Massively Parallel Processing

💽 Separate compute and storage

🚀 High performance and scalability



What is Amazon Redshift?

9. 🧠 ML-driven performance innovations

🔮 AutoMaterialized Views

🔧 Automatic query performance optimization

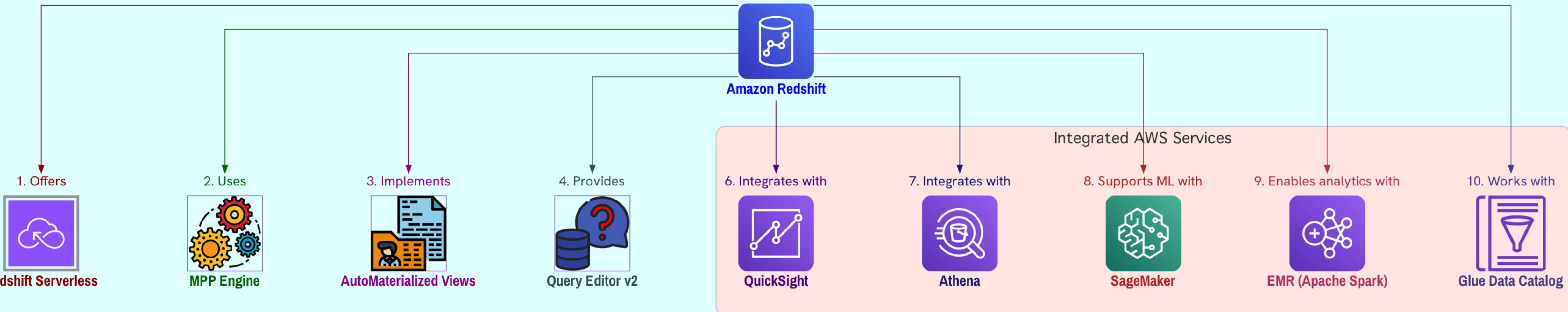
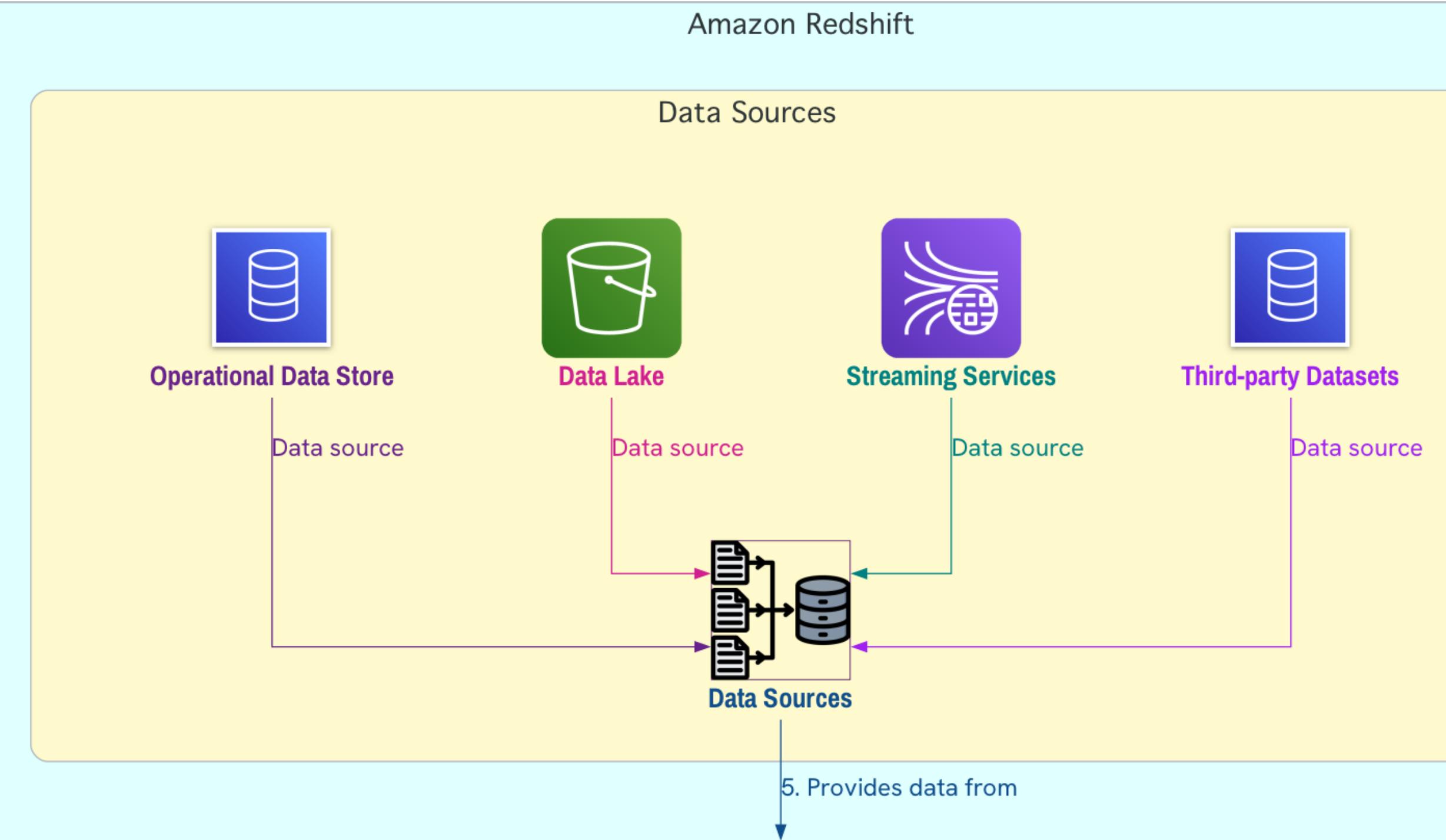
📈 Improved system efficiency

10. 💼 5x better price performance vs competitors

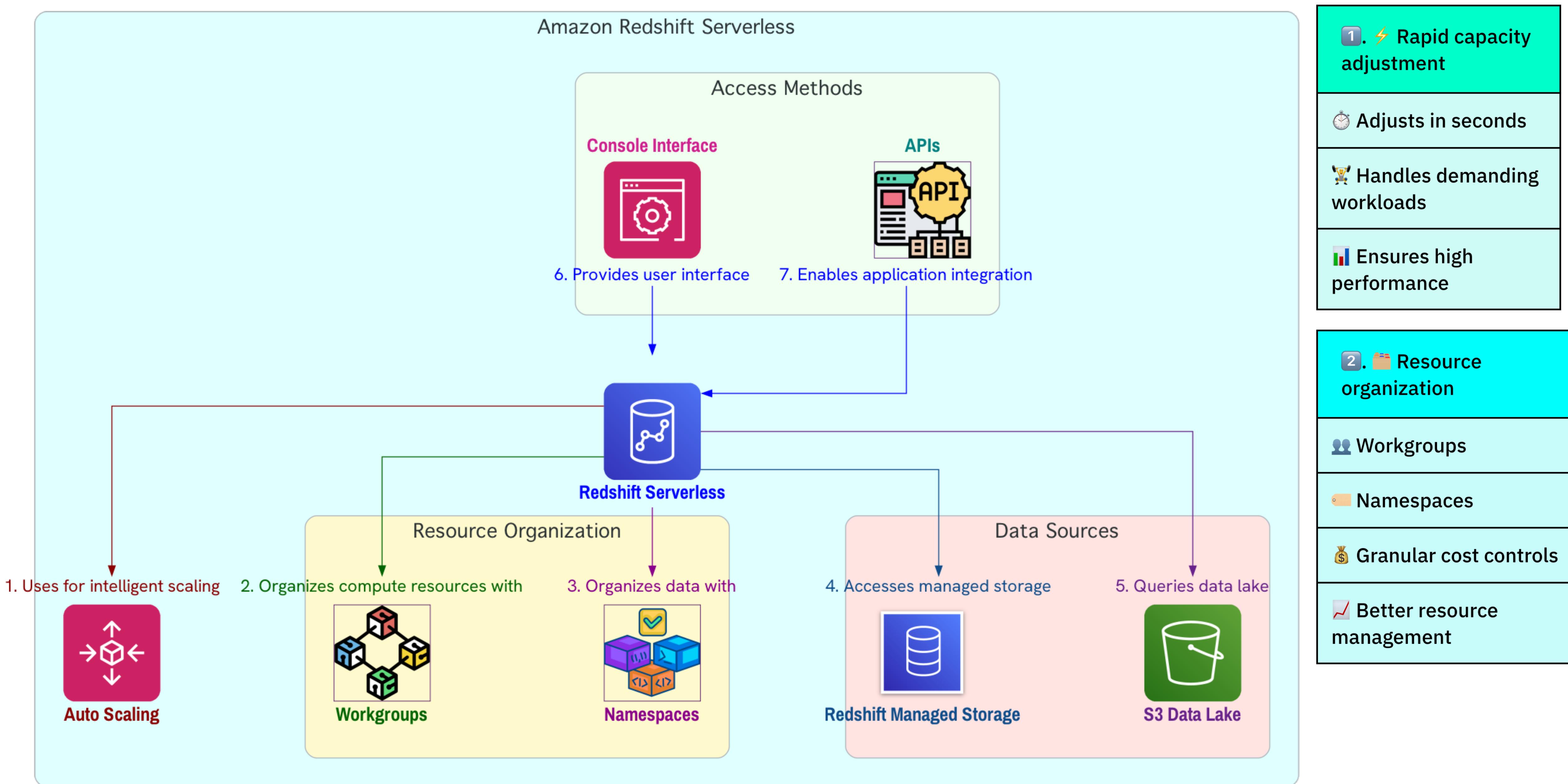
💰 Cost-effective solution

📊 Maximize data analytics capabilities

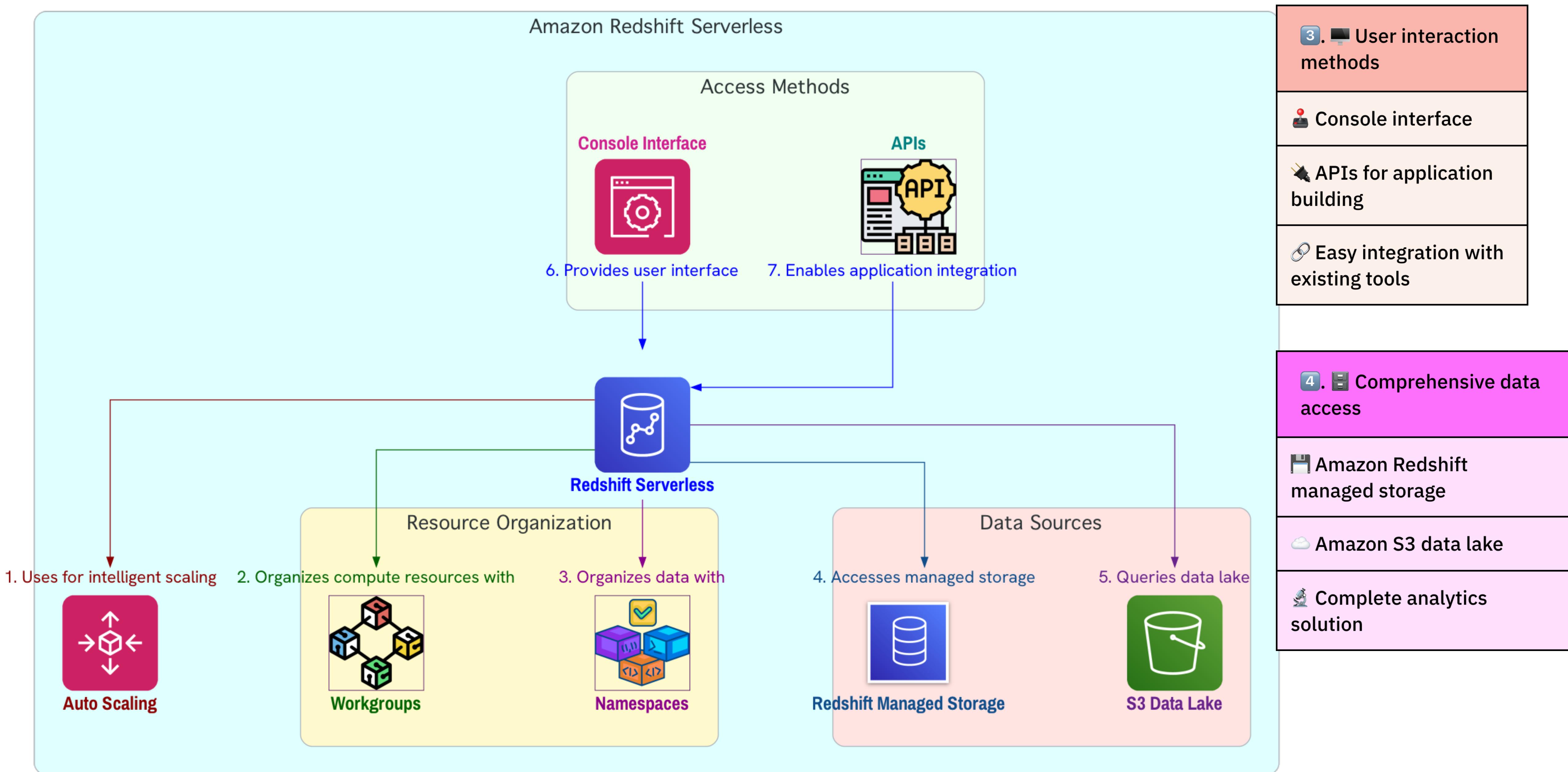
📉 Minimize costs



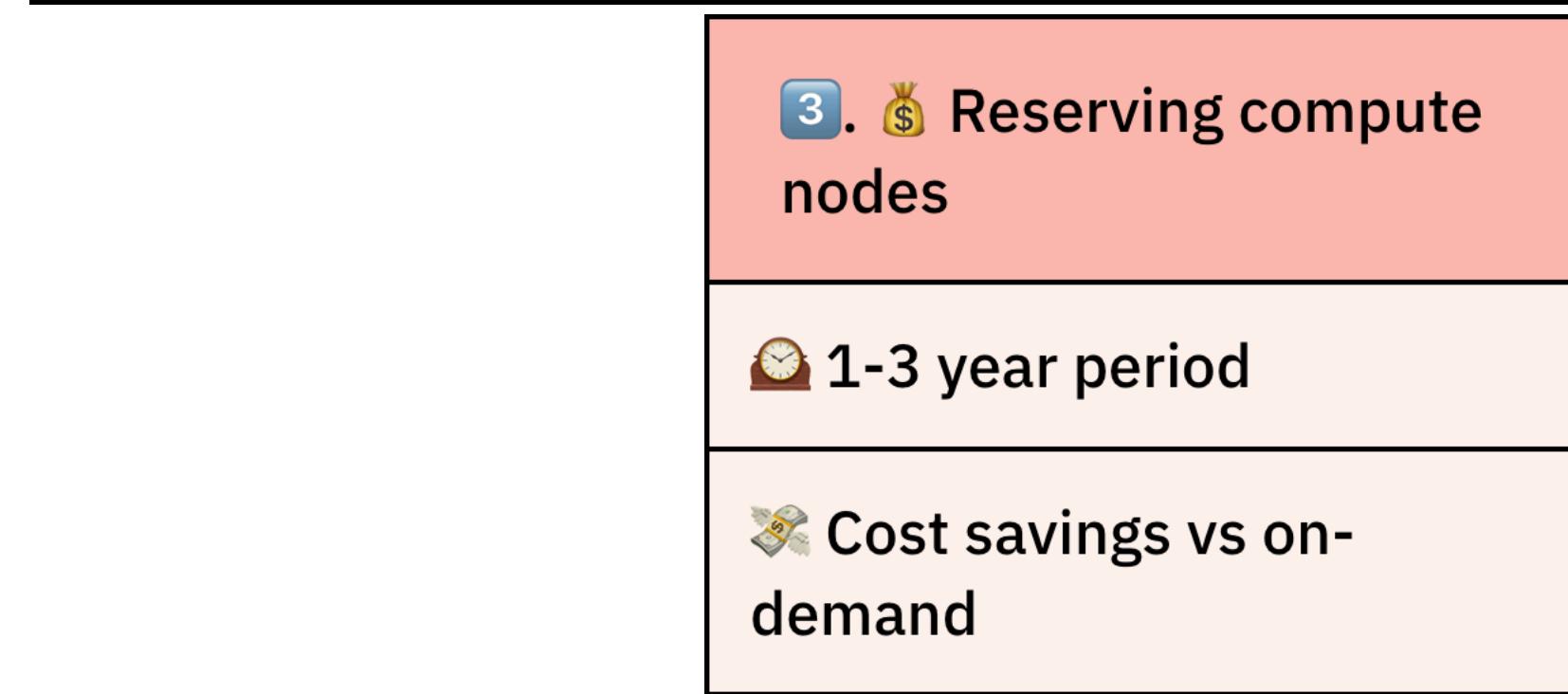
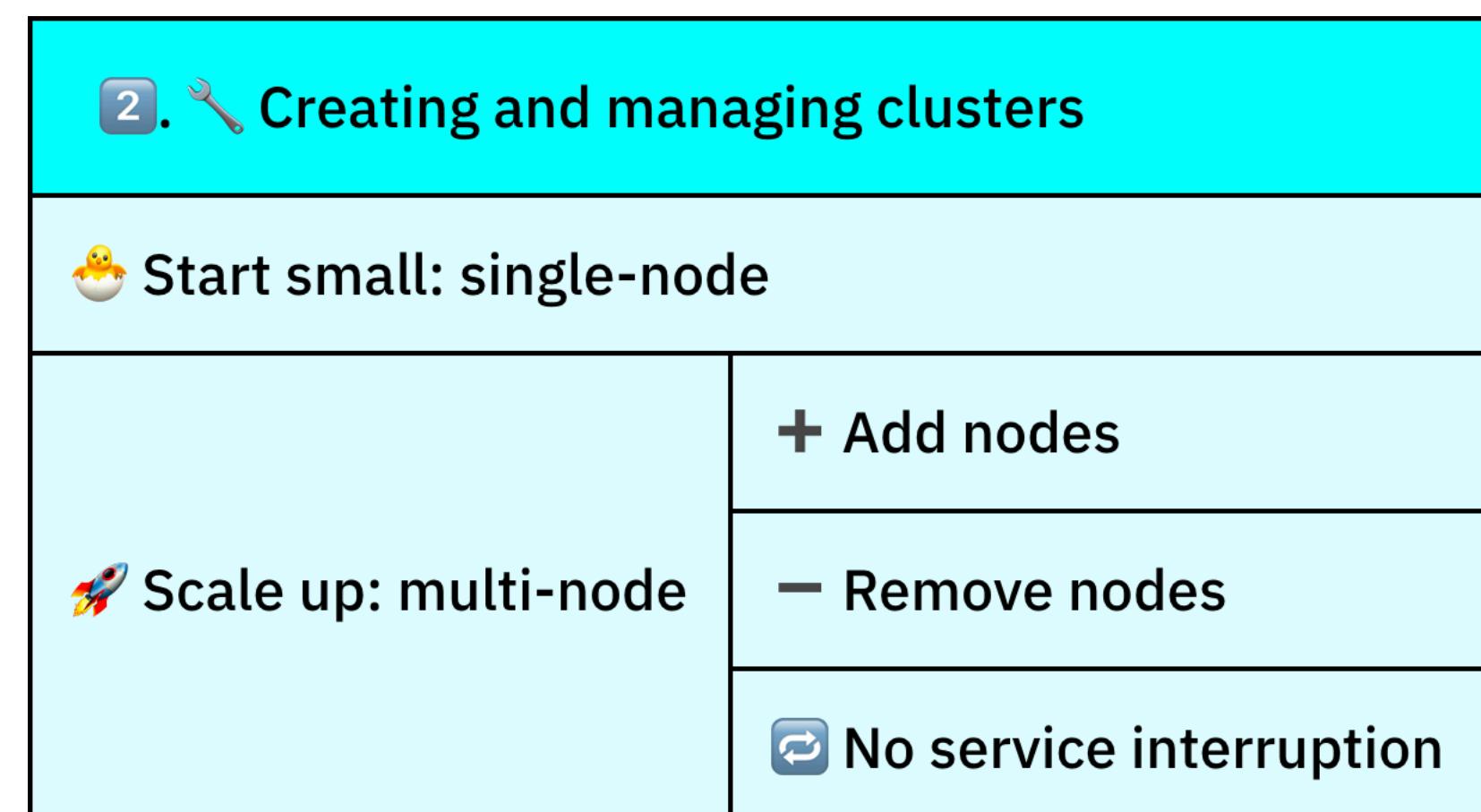
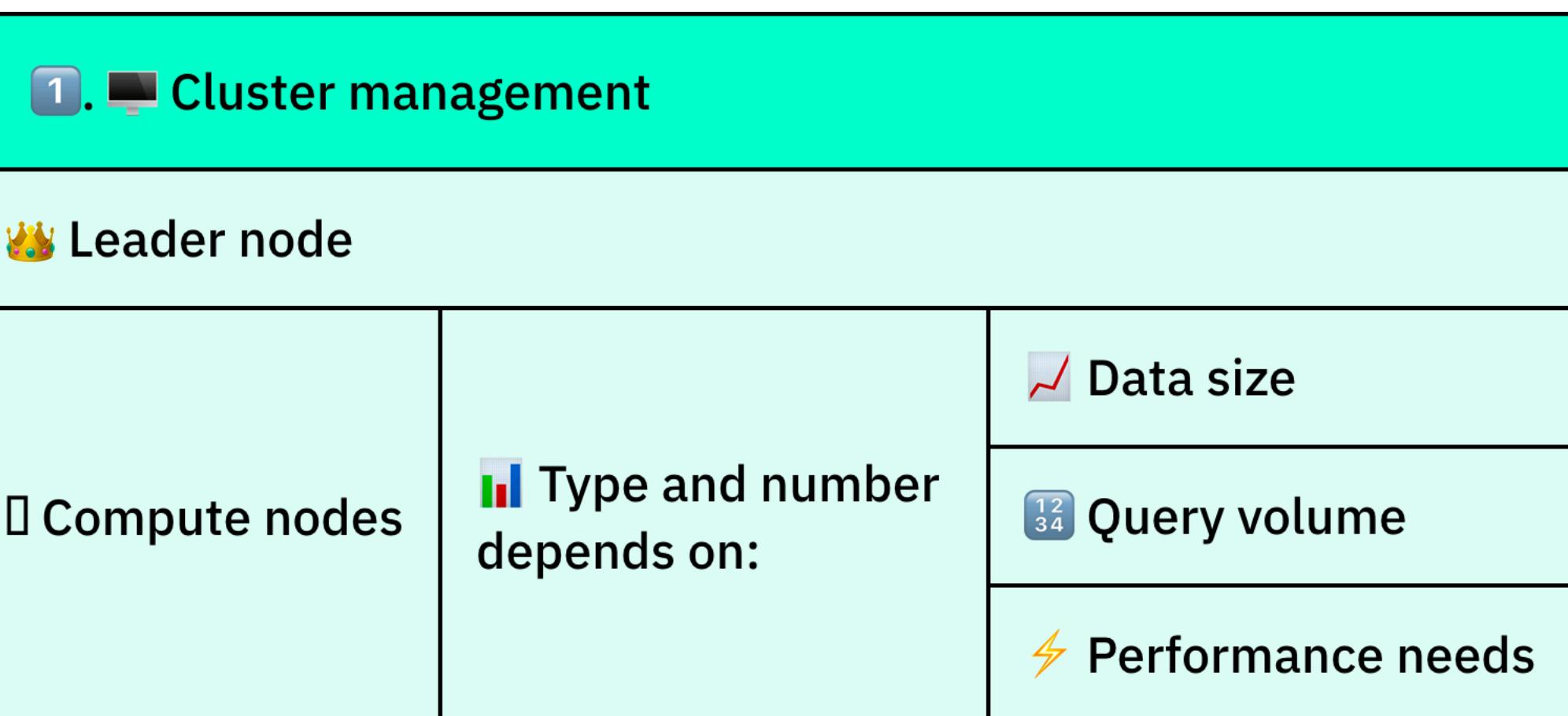
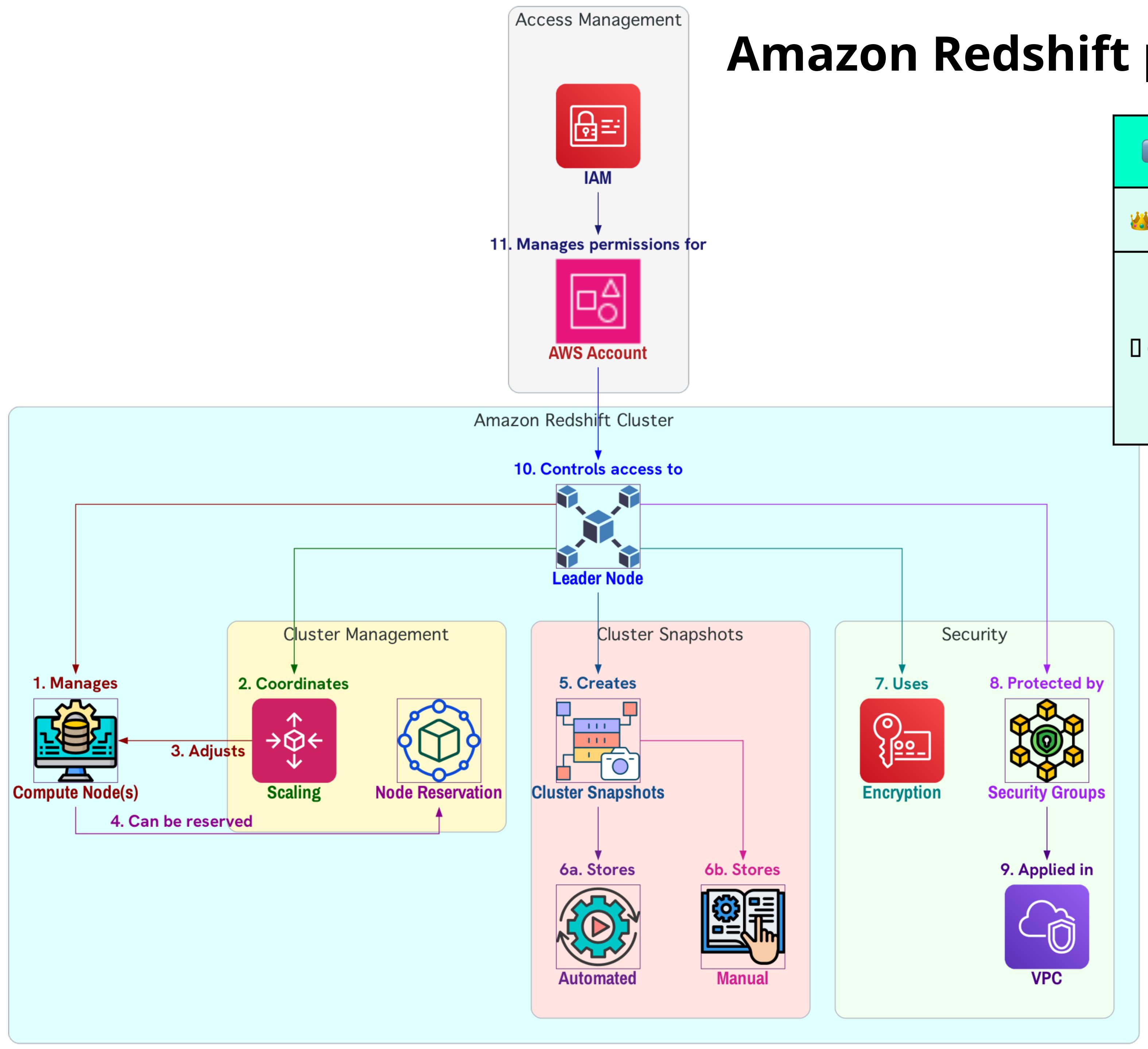
Amazon Redshift Serverless



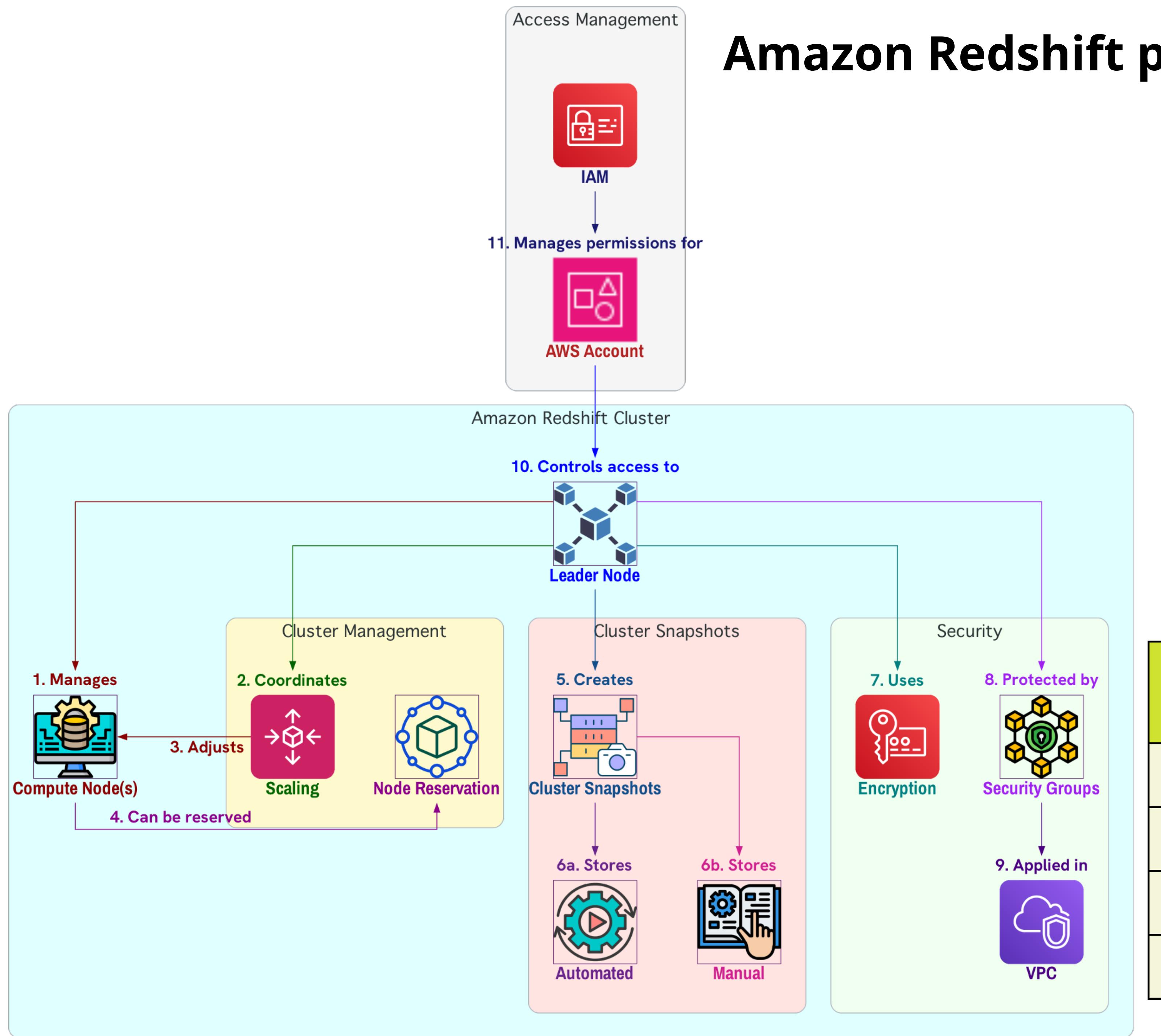
Amazon Redshift Serverless



Amazon Redshift provisioned clusters overview

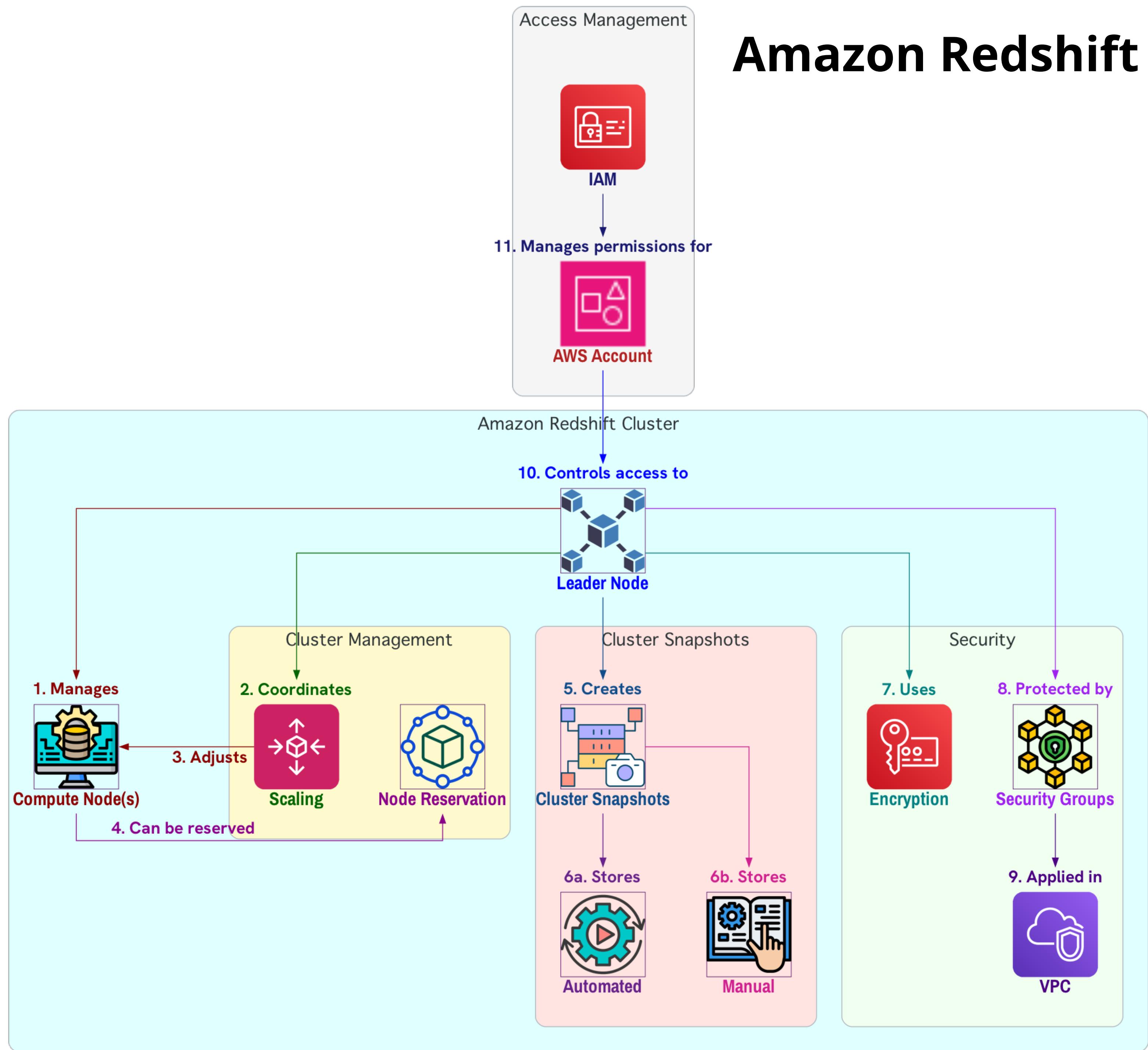


Amazon Redshift provisioned clusters overview



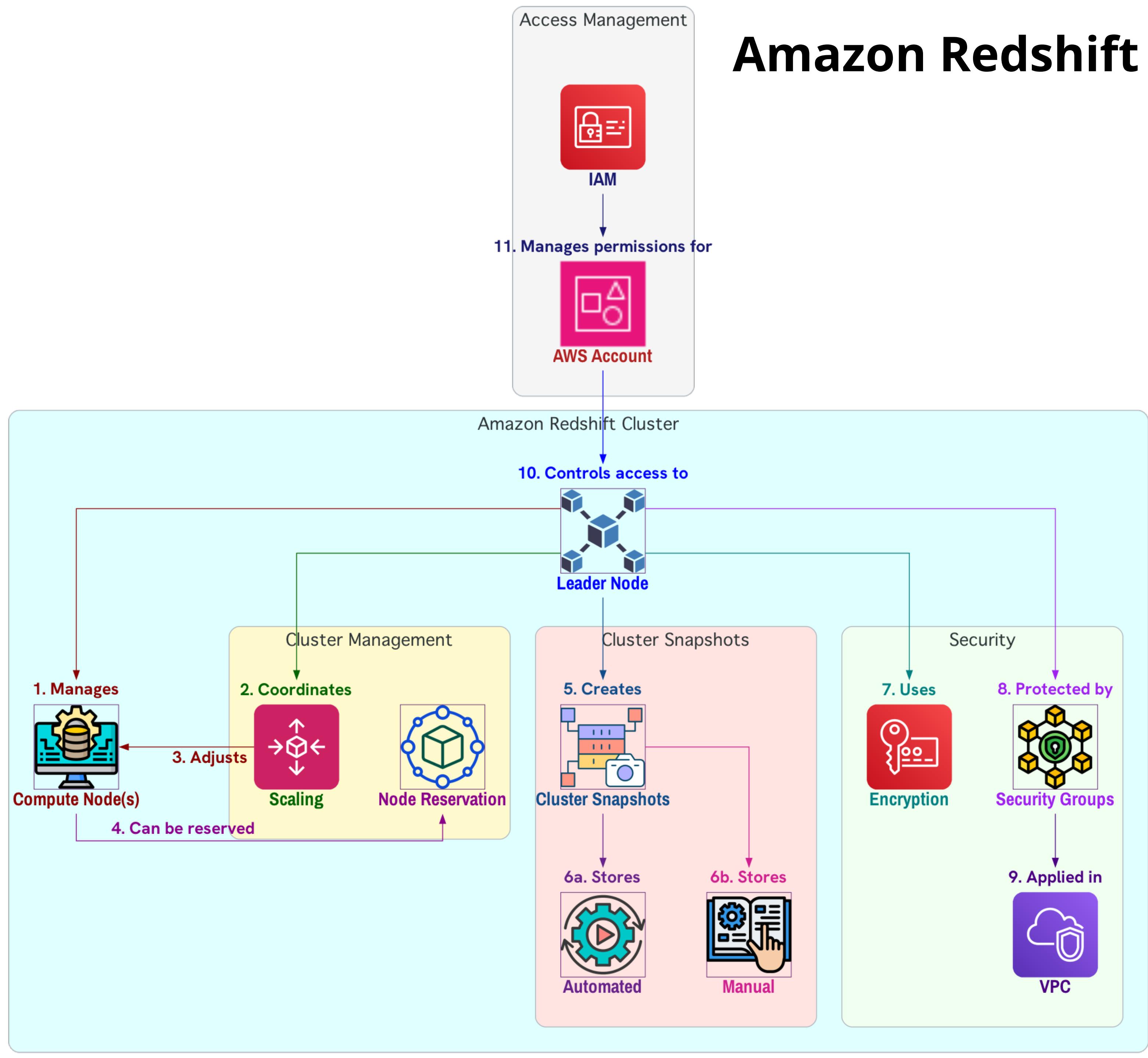
| | |
|---------------------------------------|---------------------------------------|
| 4. 📸 Creating cluster snapshots | |
| ⌚ Point-in-time backups | |
| 🤖 Automated snapshots | |
| 👤 Manual snapshots | |
| Stored in Amazon S3 | 🔒 Encrypted SSL connection |
| Creates new cluster | ⌚ Point-in-time backups |
| Restoration process | 🤖 Automated snapshots |
| Imports data from snapshot | 👤 Manual snapshots |
| 5. 🔒 Cluster access and security | 5. 🔒 Cluster access and security |
| 🚫 Access control | 🚫 Access control |
| 🌐 Connectivity rules | 🌐 Connectivity rules |
| 🔒 Data encryption | 🔒 Data encryption |
| 🔗 Connection encryption | 🔗 Connection encryption |
| 6. 🔑 AWS accounts and IAM credentials | 6. 🔑 AWS accounts and IAM credentials |
| 👤 AWS account access only | 👤 AWS account access only |
| 🔒 Cluster locked down | 🔒 Cluster locked down |
| 👥 IAM user accounts | 👥 IAM user accounts |
| 🛡️ Permission management | 🛡️ Permission management |

Amazon Redshift provisioned clusters overview



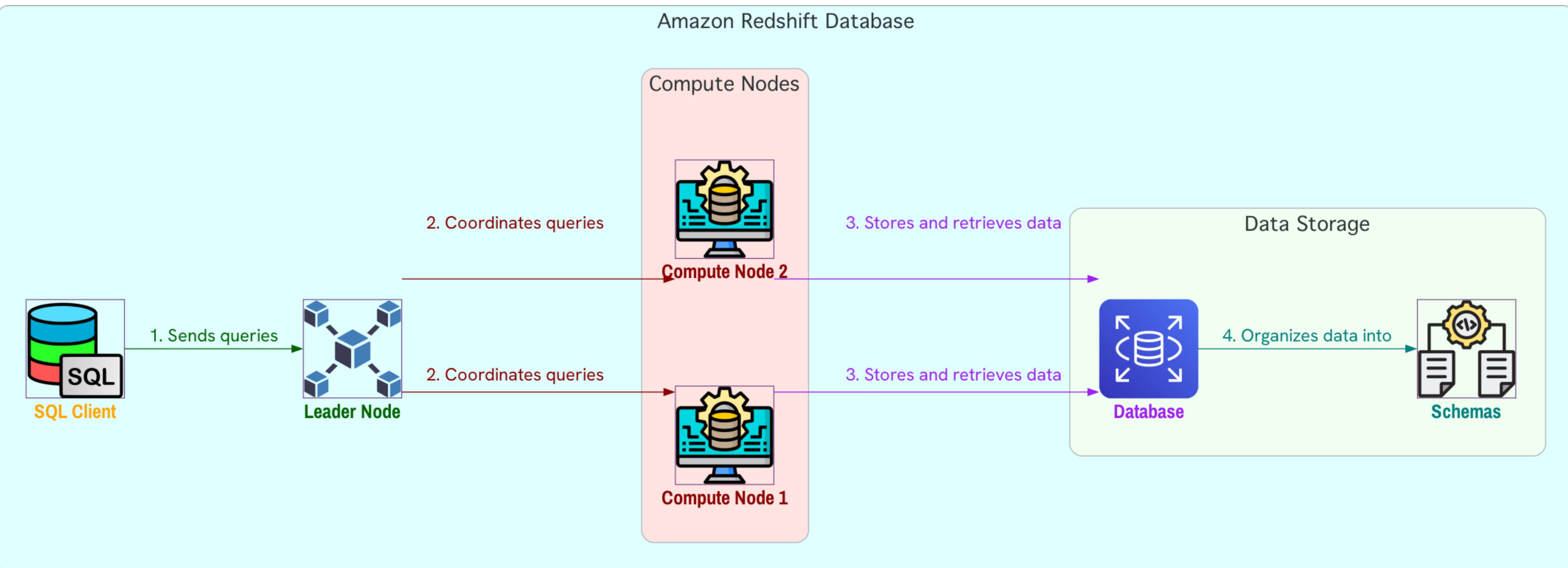
| | |
|-------------------------|------------------------------|
| 7. 🛡️ Security groups | |
| 🔒 Closed by default | |
| 🔑 IAM for API resources | Console CLI API SDK |
| 🔒 SQL client access | JDBC ODBC |
| ☁️ EC2-VPC platform | VPC security groups |

Amazon Redshift provisioned clusters overview



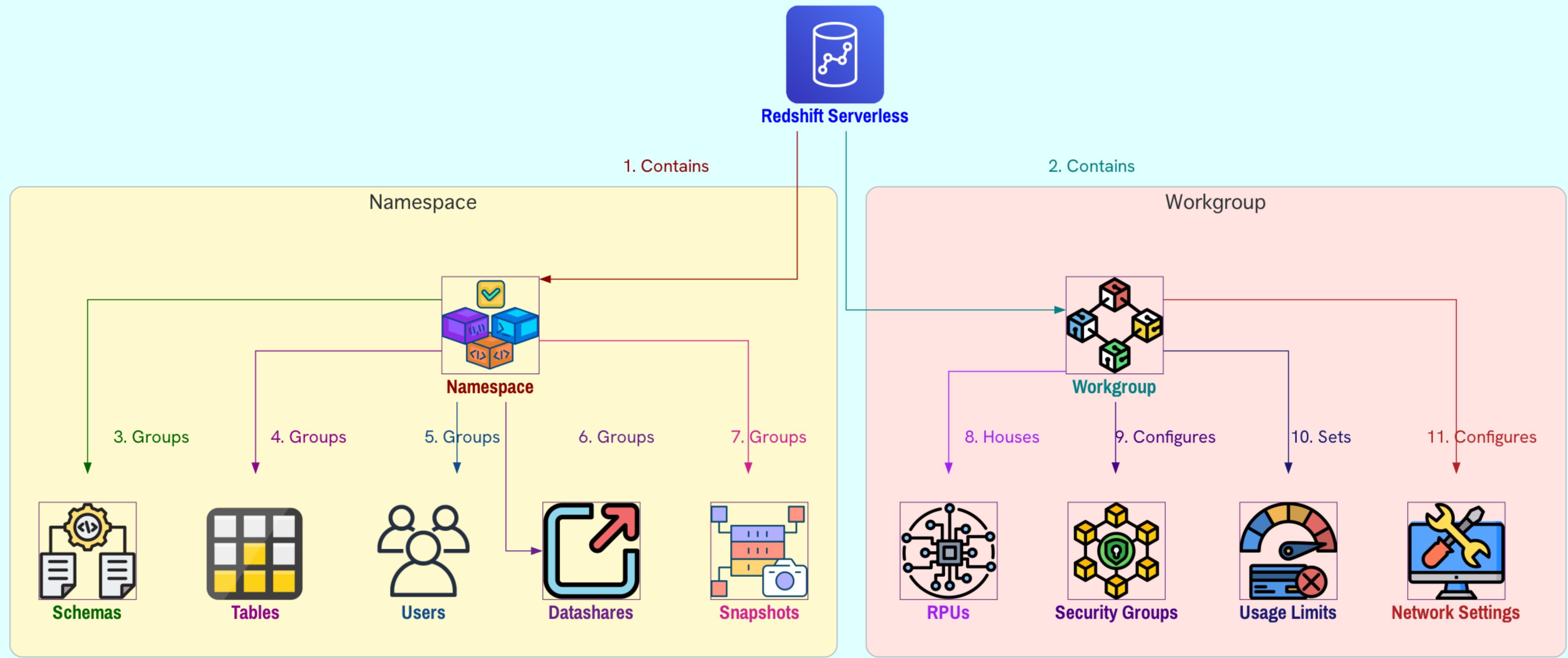
| | |
|-----------------------------|-----------------------|
| 8. 🔒 Encryption | Unload data |
| Optional cluster encryption | Reload to new cluster |
| AWS KMS for key management | Cluster |
| Immutable property | Backups |
| Switching encryption status | Restored clusters |
| Applies to: | |

Amazon Redshift Provisioned Cluster Database



Amazon Redshift Serverless Key Concepts

Amazon Redshift Serverless



Understanding Amazon Redshift Serverless Capacity

RPUs

Amazon Redshift Serverless measures data warehouse capacity in **Redshift Processing Units (RPUs)**. RPUs are resources used to handle workloads.

Understanding Amazon Redshift Serverless Capacity

RPUs

Amazon Redshift Serverless measures data warehouse capacity in **Redshift Processing Units (RPUs)**. RPUs are resources used to handle workloads.

Base Capacity

This setting specifies the base data warehouse capacity Amazon Redshift uses to serve queries. Base capacity is specified in RPUs.

- One RPU provides 16 GB of memory
- Default base capacity: 128 RPUs
- Adjustable range: 8 RPUs to 512 RPUs (in units of 8)
- Higher base capacity improves query performance, especially for resource-intensive data processing jobs

Flexibility for Different Workloads

- **Minimum capacity:** 8 RPU
- **8, 16, and 24 RPU base capacities:** Targeted for workloads requiring less than 128TB of data
- **32 RPU or more:** Required for data requirements greater than 128 TB
- **32 or more RPU:** Recommended for workloads with tables having large number of columns and higher concurrency

Flexibility for Different Workloads

- **Minimum capacity:** 8 RPU
- **8, 16, and 24 RPU base capacities:** Targeted for workloads requiring less than 128TB of data
- **32 RPU or more:** Required for data requirements greater than 128 TB
- **32 or more RPU:** Recommended for workloads with tables having large number of columns and higher concurrency

Considerations and Limitations

1. Storage Capacity Limits:

- **8 or 16 RPU configurations:** Support up to 128 TB of Redshift managed storage capacity
- **More than 128 TB of managed storage:** Minimum 32 RPU required

2. Query Impact:

- Editing workgroup's base capacity might cancel some running queries

| Feature | Serverless | Provisioned |
|--------------------------|---|--|
| Workgroup and Namespace | Uses namespaces for database objects/users and workgroups for compute resources | Uses clusters with compute nodes and a leader node |
| Node types | Automatically managed, can specify base capacity | User chooses node types and count |
| Workload management | Automatically scales based on workloads | Concurrency scaling enabled manually |
| Port | 5431–5455 or 8191–8215 | Any port |
| Resizing | Not applicable, can change base RPU capacity | Manual cluster resize |
| Pausing and resuming | Not needed, pay only when queries run | Manual pause/resume |
| Spectrum queries | Billed when compute resources process workloads | Separate Spectrum capacity |
| Compute-resource billing | Per-second, 60-second minimum | Per-second when not paused |
| Maintenance window | No window, updates handled seamlessly | User-chosen recurring time |
| Encryption | Always encrypted with AWS KMS | Optional encryption |
| Storage billing | GB per month | Separate from compute for RA3 nodes |
| User management | IAM or Redshift users | IAM or Redshift users |
| JDBC/ODBC compatibility | Compatible | Compatible |
| Credentials on sign-in | Not required every instance | Required for each sign-in |
| Data API | Supports (uses workgroup-name parameter) | Supports (uses cluster-identity parameter) |

Serverless vs. Provisioned

| Feature | Serverless | Provisioned |
|-------------------------|--|--|
| Snapshots | Supports snapshots and recovery points | Supports snapshots |
| Data Sharing | Supports all features, including cross-type | Supports cross-database, account, region sharing |
| Tracks | No concept of tracks | Supports current and trailing tracks |
| System tables and views | Supports new tables and views | Supports existing set |
| Parameter groups | No concept of parameter groups | Supports parameter groups |
| Query monitoring | Requires database connection, in sync with system tables | May not show all data in system tables |
| Audit logging | CloudWatch destination, no S3 support | S3 and CloudWatch support |
| Event notifications | Uses Amazon EventBridge | Managed through Redshift console |



**Thanks
for
Watching**