**IR types & networking · Linked service auth (Managed ID / KeyVault) · Dataset vs linked service · Copy activity options · Fault-tolerance & retries · Monitor hub · Alert rules · Cost view**

**1. Integration Runtime (IR) Types & Networking**

**Types of Integration Runtime:**

| **Type** | **Description** | **Use Case** |
| --- | --- | --- |
| **Azure IR** | Fully managed by Microsoft | Cloud data movement and transformation |
| **Self-hosted IR** | Installed on-prem or in VM | On-premises data sources (SQL, Oracle, etc.) |
| **Azure SSIS IR** | Runs SSIS packages in ADF | Lift & shift SSIS-based ETL |

**Networking Options:**

* **Private Endpoint**: Connect securely to ADF via Azure Private Link.
* **Managed VNet**: ADF creates an isolated VNet to run activities securely.
* **Bring Your Own VNet (BYOVNet)**: Use your own VNet with Self-hosted IR.

**2. Linked Service Authentication (Managed Identity / Key Vault)**

**Linked Service Authentication Options:**

* **Azure Managed Identity (MSI)**:
  + ADF authenticates to resources like Blob, SQL DB using its system-assigned identity.
  + Recommended for production — no hardcoded secrets.
* **Azure Key Vault**:
  + Secrets (keys, connection strings) are stored securely.
  + Linked services fetch credentials at runtime.
  + Rotate secrets without redeploying ADF.

**Best Practice:**

* Use **Managed Identity + Key Vault** for secure, scalable authentication.

**3. Dataset vs Linked Service**

| **Feature** | **Dataset** | **Linked Service** |
| --- | --- | --- |
| Purpose | Describes the **shape and structure** of data (e.g., file format, table) | Defines **connection details** to data source |
| Example | CSV dataset with delimiter , | Azure Blob Storage account connection |
| Reusability | Specific to pipeline activities | Shared across multiple datasets/pipelines |
| Coupling | Depends on linked service | Independent |

**4. Copy Activity Options**

**Source Options:**

* Full load, incremental load via watermark
* File-based: Wildcards, modified datetime filters
* Database: Query, table, stored procedure

**Sink Options:**

* File partitioning by date, size, or column
* Merge, upsert (with data flow), append

**Performance Tuning:**

* Parallel copy (based on partitions)
* Compression & staging
* Pre-copy and post-copy scripts

**5. Fault Tolerance & Retries**

**Key Features:**

* **Retry Policy**: Set retry count and interval on failed activity
* **Skip Failed Rows**: In Copy activity, skip corrupt data rows
* **Log Bad Records**: Store in error directory
* **Timeout Settings**: Activity-level timeout to avoid long-running jobs

**Use Cases:**

* Data integrity issues
* Transient network failures
* Source system throttling

**6. Monitor Hub (Monitoring Pane in ADF)**

**Capabilities:**

* Visualize pipeline runs, trigger runs, activity failures
* Drill down into activity duration and logs
* Filter by status (succeeded, failed, in progress)

**Options:**

* Resubmit pipeline run
* Cancel in-progress runs
* Set alerts (linked to Azure Monitor)

**7. Alert Rules (via Azure Monitor)**

**Alert Targets:**

* Pipeline failure
* Activity timeout
* IR going offline

**Actions:**

* Email
* Teams via Webhook
* Trigger Logic Apps / Azure Functions

**Best Practice:**

* Define severity-based alerts (Critical, Warning)
* Combine with log analytics queries for custom metrics

**8. Cost View**

**Visibility:**

* Azure Cost Management + ADF monitoring metrics
* Monitor:
  + Data movement cost
  + Pipeline execution time
  + IR activity cost (vCPU hours)

**Optimization Tips:**

* **Avoid idle SSIS IRs** (auto-pause)
* **Consolidate pipelines** where possible
* Use **data partitioning** to speed up movement and reduce vCPU cost
* Prefer **Azure IR** unless on-prem access is mandatory

**Summary Tips:**

| **Topic** | **Key Tip** |
| --- | --- |
| IR | Use Azure IR where possible; only use SHIR for on-prem |
| Auth | Prefer Managed Identity + Key Vault |
| Linked Service vs Dataset | Use LS for connection, DS for metadata |
| Copy | Enable fault-tolerance and test with small batches |
| Monitoring | Automate alerts and use diagnostics logs |
| Cost | Review ADF usage in Cost Management monthly |