**4️⃣ Splunk Architecture Explained: Key Components and Design Principles**

**Overview**

Splunk architecture is designed to **collect, index, search, analyze, and visualize machine data** in real-time.

It follows a **distributed architecture**, meaning components can run on a single machine (small setup) or across multiple machines (enterprise scale).

**Key Components of Splunk Architecture**

**1. Forwarders (Data Collectors)**

* Installed on **source systems** (servers, applications, devices).
* Responsible for **collecting logs/events** and sending them to the Indexer.
* Two types:
  + **Universal Forwarder (UF):** Lightweight, forwards raw data.
  + **Heavy Forwarder (HF):** Can parse/filter data before sending.

**2. Indexers (Data Storage & Processing)**

* The **core component** where data is stored and processed.
* Functions:
  + Receives data from forwarders.
  + **Indexes (organizes)** data for fast searching.
  + Stores data in **time-series format** (making historical searches faster).
* Can scale horizontally (multiple indexers in a cluster).

**3. Search Head**

* The **user interface** for searching and analyzing data.
* Provides:
  + **Search Processing Language (SPL)** for queries.
  + Dashboards and reports.
  + Role-based access control for users.
* Multiple Search Heads can be clustered for **high availability**.

**4. Deployment Server**

* Used to **manage configurations** of multiple forwarders and Splunk components.
* Helps in **centralized management** for large-scale deployments.

**5. Cluster Master (Indexing Cluster Management)**

* Manages **indexer clustering** for scalability and fault tolerance.
* Ensures data is **replicated** across multiple indexers.

**6. License Master**

* Manages Splunk licenses.
* Ensures compliance with daily data ingestion limits.

**Data Flow in Splunk**

1. **Data Collection:**

Logs are collected using **Forwarders**.

1. **Data Processing & Indexing:**

Indexers parse, process, and store the data.

1. **Searching & Visualization:**

Users interact with the **Search Head** to query, analyze, and visualize data.

**Design Principles**

1. **Scalability**
   * Splunk can handle massive amounts of data by adding more indexers/search heads.
2. **Reliability & Fault Tolerance**
   * Indexer clustering ensures data replication and availability.
3. **Real-Time Processing**
   * Provides near real-time monitoring and alerting.
4. **Security & Access Control**
   * Role-based access and integration with enterprise authentication systems.
5. **Extensibility**
   * Can integrate with **cloud platforms, third-party tools, and apps** via APIs.

**Summary**

Splunk’s architecture is built around **Forwarders, Indexers, and Search Heads**.

* **Forwarders** collect the data.
* **Indexers** store and process the data.
* **Search Heads** allow users to search, analyze, and visualize the data.

Supporting components like **Deployment Server, Cluster Master, and License Master** make Splunk scalable, manageable, and enterprise-ready.