Day 2 – ELK Stack

INTRODUZIONE

Stack ELK

Lo **stack ELK** è una potente soluzione open-source utilizzata per raccogliere, elaborare, archiviare e visualizzare dati provenienti da fonti diverse, in particolare log e eventi di sistema. È composto da tre componenti principali che lavorano insieme per creare una piattaforma completa di analisi dei dati che sono: **Elasticsearch** , **Logstash** e **Kibana**.

Grazie alla sua flessibilità e scalabilità, lo stack ELK è ampiamente utilizzato in contesti come il monitoraggio delle infrastrutture IT, l'analisi delle prestazioni e la sicurezza informatica, diventando una scelta popolare per costruire un **SIEM** (Security Information and Event Management).

In un'architettura SIEM, lo stack ELK permette di centralizzare e analizzare grandi quantità di log in tempo reale, fornendo strumenti per monitorare, rilevare anomalie, e rispondere a potenziali minacce di sicurezza. La sua architettura modulare consente di adattarlo a diverse esigenze, rendendolo ideale sia per ambienti aziendali che per piccole installazioni.

Elasticsearch

- **E’ un database che archivia log**

Elasticsearch è utilizzato per memorizzare e cercare grandi quantità di log, inclusi syslog, firewall logs e altri eventi di sistema. Grazie alla sua capacità di gestire dati in tempo reale, è ideale per monitorare infrastrutture IT e sistemi di sicurezza.

- **Usa un linguaggio di query:**

Elasticsearch Query Language (ESQL) è il linguaggio utilizzato per interrogare i dati. Permette di filtrare, ordinare e aggregare informazioni con query complesse. ESQL consente agli utenti di ottenere rapidamente insight dettagliati dai dati archiviati, semplificando la ricerca e l'analisi.

- **Usa RESTful API per interagire**

Le interazioni con Elasticsearch avvengono principalmente tramite una RESTful API. Questo consente di effettuare operazioni di ricerca, recupero e gestione dei dati utilizzando richieste HTTP (GET, POST, PUT, DELETE). L'API facilita l'integrazione con altri strumenti e piattaforme per una gestione efficiente dei dati.

Logstash

- **Raccoglie dati da varie sorgenti**

Logstash è responsabile della raccolta di dati (telemetria) provenienti da diverse fonti, come file di log, dati di rete, database e altri endpoint, permettendo di centralizzare le informazioni per una successiva analisi.

- **Trasforma e filtra i dati prima di trasmetterli a Elasticsearch**

Logstash permette di trasformare i dati, applicare filtri, e pulirli prima di inviarli a Elasticsearch per l'archiviazione e l'analisi. Questo passaggio aiuta a organizzare i dati in modo strutturato e ad estrarre solo le informazioni rilevanti.

- **Parsing dei dati**

Logstash permette di fare il parsing dei dati, ovvero mappare le keyword all'interno di un log per trasformarle in campi strutturati. Non tutti i log hanno un parser predefinito, ma con Logstash è possibile creare uno schema personalizzato per estrarre i valori necessari.

Logstash

- **Due metodi principali per raccogliere i dati: Beats ed Elastic Agents**

Esistono vari metodi per raccogliere i dati dagli endpoint:

1. **Beats:** Sono agent leggeri installati sugli endpoint per raccogliere dati specifici, ad esempio:
    a) Filebeat per i file di log
    b) Packetbeat per i dati di reteIn base alla tipologia di dati da raccogliere, installerai un determinato Beat sull'endpoint.
2. **Elastic Agent:** Un singolo agente capace di raccogliere più tipologie di dati in modo centralizzato e gestibile.

Kibana

- **Console web per cercare nei log memorizzati in Elasticsearch**

Kibana è un'interfaccia web che permette di esplorare e analizzare i dati memorizzati in Elasticsearch. Consente agli utenti di cercare e visualizzare rapidamente i log e i dati di sistema, migliorando la capacità di indagine.

- **Feature di visualizzazione: Lens e dashboard**

Kibana include strumenti come Lens, che permettono di importare e visualizzare i dati in dashboard interattive e personalizzabili. È possibile creare visualizzazioni dinamiche per monitorare le prestazioni e analizzare i dati con facilità.

- **Discover tab per cercare dati tramite query ESQL**

Tramite la scheda Discover, gli utenti possono eseguire ricerche avanzate sui dati utilizzando il linguaggio di query ESQL, filtrando i log e ottenendo insights precisi in tempo reale.

- **Funzionalità avanzate: Machine Learning e Metrics Alerting**

Kibana offre funzioni avanzate come il Machine Learning per rilevare anomalie nei dati e generare avvisi automatici.

Perché elk stack

- **Logging centralizzato**

ELK permette di centralizzare tutti i log in un unico sistema, essenziale per la compliance e per facilitare la ricerca dei dati. Con un unico punto di accesso, è più semplice gestire e analizzare le informazioni provenienti da diverse fonti.

- **Flessibilità nell'ingestione dei dati**

Grazie alla sua architettura modulare, lo stack ELK consente di customizzare il processo di ingestione dei dati in base alle esigenze specifiche. Puoi adattare il sistema a diversi formati di log e fonti di dati, utilizzando plugin e pipeline personalizzate.

- **Visualizzazione immediata dei dati**

Le funzionalità di visualizzazione offerte da Kibana rendono le informazioni immediatamente accessibili tramite grafici, dashboard e report. Questo permette di interpretare i dati velocemente e prendere decisioni basate su visualizzazioni chiare e comprensibili.

Benefici dello stack ELK

- **Scalabilità**

Lo stack ELK è progettato per essere scalabile, il che significa che può gestire grandi volumi di dati, supportando ambienti aziendali di grandi dimensioni e in continua crescita, senza compromettere le performance.

- **Ecosistema ricco e integrabile**

Lo stack ELK fa parte di un ampio ecosistema con numerose integrazioni disponibili. Puoi estendere le funzionalità collegandolo ad altre soluzioni o piattaforme, e adattarlo alle esigenze specifiche dell'infrastruttura aziendale.

- **Costituisce la base di molti SIEM**

Molti sistemi di monitoraggio e sicurezza (SIEM) sono costruiti sullo stack ELK grazie alla sua robustezza e flessibilità, rendendolo una scelta popolare per la sicurezza IT e la gestione delle operazioni.