

بسم الله الرحمن الرحيم

گزارش استقرار آزمایشگاه خط داده

جزئیات عملیاتی

- سیستم عامل: LTS22.04.5 Ubuntu
- استقرار: Docker
- سرویس‌های راه‌حل:
 - پایگاه داده NoSQL: Elasticsearch 9.1.5
 - ابزار دیداری‌سازی داده‌ها: Kibana 9.1.5
 - پایگاه داده SQL: PostgreSQL 16.10-alpine3.22
 - ارکستراسیون: Airflow 2.9.2
 - ارکستراسیون: NiFi 1.27.0
- نسخه گزارش: ۰.۱

صورت مسئله

راه‌کار خط داده‌ها معمولاً شامل حداقل سه مؤلفه معماری منطقی است که در نمودار زیر شرح داده شده است:



در باره این شکل گفتنی است، گره‌ها به‌طور منطقی نمایانگر پیاده‌سازی هستند، اما هر گره می‌تواند حاوی جزئیات و اجزای پیچیده‌ای باشد. در این تمرین، هدف ما ایجاد یک آزمایشگاه ساده است که شامل تنها یک گره بوده و به‌عنوان محیطی برای تمرین و کار با داده‌ها عمل کند.

راه حل

1.1 نوشتن Docker Compose

برای کانتینر Elasticsearch و Kibana، [در تارنمای رسمی elastic پست وبلاگی نوشته شده است](#) که نحوه نوشتن Docker Compose مناسب را توضیح می‌دهد. برای اهداف تجاری و محیط‌های عملیاتی، از گواهی‌نامه SSL استفاده می‌شود، اما در این آزمایشگاه نیازی به آن نیست. برای سایر سرویس‌ها، جزئیات نوشته شده در Docker Compose توسط ChatGPT نوشته شده و همان‌ها نیازمندی‌های موجود را برطرف می‌کند، بنابراین با اندکی تغییر آمده است. از آن‌جا که نوشتن کد و موارد مشابه در نرم‌افزار Microsoft Word دشواری‌هایی دارد، بنابراین نشانی مخزن گیت‌هاب برای بررسی اسناد مربوط به استقرار صورت گرفته در این‌جا [الصاق](#) شده است.

1.2 نکات استقرار

در ابتدا، برای ثبت و persist کردن داده‌های elastic search من از مسیرهای مشخص شده مانند data/esdata/. استفاده کردم، اما کار نکرد و نتواستم دلیل آن را با جستجوهای کوتاهی که انجام دادم پیدا کنم. شرکت ما یک مخزن تصاویر داکر Nexus دارد که به عنوان پروکسی برای اهداف تحریمی عمل می‌کند. اما می‌توان از تصاویر با راه‌حل‌های ضد تحریم دیگر نیز استفاده کرد.

1.3 بررسی وضعیت آزمایشگاه

برای بررسی صحت عمل‌کرد، لاگ سرویس‌ها بررسی شد تا با عدم نمایش exception نسبت به کارکرد صحیح استقرار اطمینان حاصل شود، همین‌طور سرویس‌هایی که مانند کیبانا بخش وب داشتند، در مرورگر بررسی شدند.

پایان

علی جمعه‌ئی

۲۹ مهرماه ۱۴۰۴