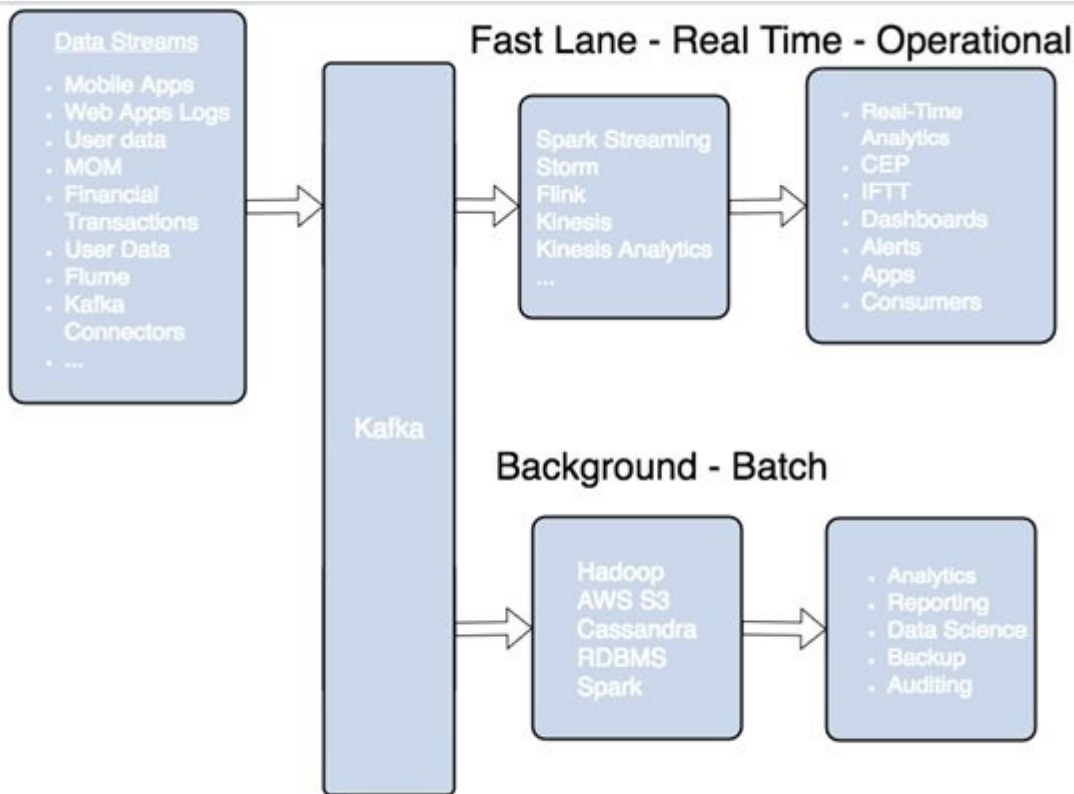


کافکا معماری پردازش جریان داده‌ها است که اولین بار در سال 2011 از سوی مهندسان لینکدین به منظور اداره کردن حجم انبوهی از داده‌هایی که به صورت بلادرنگ تولید می‌شوند ابداع گردید. کافکا اغلب در معماری‌های مربوط به استریم کردن بلادرنگ داده‌ها مورد استفاده قرار می‌گیرد تا یک تحلیل بلادرنگ را ارائه کند. در شش سال گذشته رشد انفجاری کافکا تعجب کارشناسان را برانگیخته است.

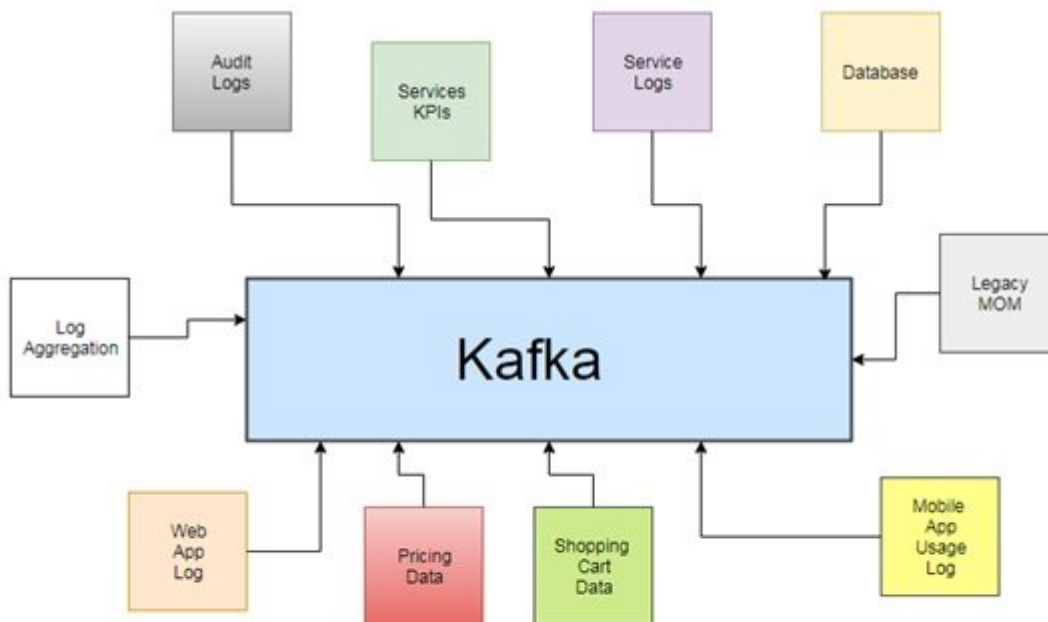
رشد کافکا واقعا جالب توجه بوده است. بیش از یک سوم از 500 شرکت مطرح جهان از کافکا استفاده می‌کنند. ده آژانس برتر مسافرتی، هفت بانک برتر جهانی، هشت شرکت بیمه‌گر بزرگ و نه شرکت بزرگ فعال در حوزه مخابرات تنها گوشه‌ای از سازمان‌های بزرگی هستند که از کافکا استفاده می‌کنند. کافکا برای استریم کردن بلادرنگ داده‌ها، جمع‌آوری بزرگ داده‌ها یا انجام تحلیل‌های بلادرنگ مورد استفاده قرار می‌گیرد. کافکا توسط میکروسرویس‌های درون حافظه‌ای مورد استفاده قرار گرفته تا پایداری بیشتری را فراهم کند. کافکا می‌تواند برای تغذیه رویدادهای مربوط به CEP (سامانه‌های پردازش رویدادهای پیچیده) و سامانه‌های اتوماسیون IoT/IFTTT به کار گرفته شود.



Kafka streaming architecture diagram

چرا باید از کافکا استفاده کنیم؟

از آن جایی که کافکا سریع، گسترش‌پذیر و بادوام بوده و همچنین در مقابل مشکلات و خرابی‌ها پایداری خوبی دارد در نتیجه گزینه ایده‌آلی است تا در مکان‌هایی که JMS، RabbitMQ و AMQP ممکن است مورد توجه قرار نگیرند، به کار گرفته شود. درست در سامانه‌های تبادل پیام که با حجم بالایی از داده‌ها و پاسخگویی روبرو هستیم و به جای ActiveMQ یا نمونه‌های مشابه در نظر داریم از سامانه‌های مدرن‌تر استفاده کنیم، کافکا و Pub/Sub ایده‌آل هستند. قابلیت اطمینان و توانایی بالا از ویژگی‌های اصلی کافکا به شمار می‌روند. فاکتورهایی که باعث شده‌اند کافکا در زمینه کارهایی همچون پیگیری تماس‌ها یا پیگیری داده‌های مربوط به حس‌گرهای اینترنت اشیا جایی که MOM سنتی ممکن است چندان مورد توجه قرار نگیرد، قابل استفاده باشد.



Kafka decoupling data streams

کافکا می‌تواند بدون مشکل خاصی با اسپارک، HBase، Flink، Storm، Streaming، و Flume برای ارائه تحلیل‌های بلادرنگ و پردازش جریان داده‌ای به کار گرفته شود. کارگزاران کافکا از استریم‌های حجیم پیام‌ها در ارتباط با ارائه تحلیل‌های بلادرنگ و قابل پیگیری در هادوپ یا اسپارک پشتیبانی می‌کنند. همچنین Kafka Streaming می‌تواند برای تحلیل‌های بلادرنگ نیز مورد استفاده قرار گیرد. به طو خلاصه، کافکا برای پردازش جریانی، پیگیری فعالیت‌های سایت‌ها، جمع‌آوری و نظارت بر معیارها (متریک‌ها)، جمع‌آوری گزارش‌ها، تحلیل بلادرنگ، CEP، به کارگیری داده‌ها در اسپارک و هادوپ، CQRS، جواب‌گویی مجدد به پیام‌ها و بازیابی خطا به کار گرفته می‌شود.

چه سازمان‌هایی از کافکا استفاده می‌کنند؟

بسیاری از شرکت‌های بزرگ که نیاز دارند حجم بالایی از داده‌ها را مدیریت کنند از کافکا استفاده می‌کنند. لینکدین از کافکا برای ردیابی اطلاعات مربوط به فعالیت‌ها و معیارهای عملیاتی استفاده می‌کند. تویتر از کافکا در قالب بخشی از Storm و در ارتباط با زیرساخت پردازشی استریم‌های خود استفاده می‌کند. Square از کافکا برای انتقال همه رویدادها به مراکز داده Square (گزارش‌ها، رویدادهای سفارشی، معیارها و...)، ارسال خروجی به Splunk، Graphite و پیاده‌سازی سامانه‌های هشدار دهنده Esper-like/CEP استفاده می‌کند. همچنین شرکت‌های بزرگ دیگری همچون سیسکو، اوبر، پی‌پال، CloudFlare و نت‌فلیکس از کافکا استفاده می‌کنند.

تاریخ انتشار:

16 شهریور 1396

نشانی منبع:

<https://www.shabakeh-mag.com/news/world/9554/%DA%A9%D8%A7%D9%81%DA%A9%D8%A7-%D9%85%D8%B9%D9%85%D8%A7%D8%B1%DB%8C-%D9%BE%D8%B1%D8%AF%D8%A7%D8%B2%D8%B4-%D8%AC%D8%B1%DB%8C%D8%A7%D9%86-%D8%AF%D8%A7%D8%AF%D9%87%E2%80%8C%D8%A7%DB%8C-%D9%84%DB%8C%D9%86%DA%A9%D8%AF%DB%8C%D9%86%D8%8C->

%D8%A8%D8%AF%D9%88%D9%86-%D8%B1%D9%82%DB%8C%D8%A8