

OMÜ, BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ, VERİTABANI LABORATUVARI, 6.DENEY
HAZIRLIĞI
2024, BAHAR

Öğr_Numarası, İsim	20060358 Zeynep Sıla KAYMAK
Deney Sorusu 1' i cevaplandırınız. (En fazla 150 kelime)	
<p>Bu sorunun çözümü, geçici bellek yerine kalıcı bir depolama alanı kullanmaktır. Önerim, smsAtılacakListe gibi verileri bir veritabanında saklamaktır. Veritabanı, verilerin diskte kalıcı olarak saklanması sağlar ve her uygulama yeniden başlatıldığında veriler korunur. Ayrıca, veritabanları genellikle yüksek performans sunar ve saniyede onbinlerce erişime uygun olabilirler. Veritabanı işlemleri için ORM (Object-Relational Mapping) araçları kullanarak veritabanı etkileşimini kolaylaştırabiliriz. Bu şekilde, her yeni versiyonda veriler kalıcı olarak saklanır ve uygulamanın kapatılmasıyla kaybolmaz. Sonuç olarak, güncelleme yapıldıktan sonra bile randevu bildirimlerinin sorunsuz bir şekilde devam etmesi sağlanır.</p>	
Deney Sorusu 2' yi cevaplandırınız. (En fazla 500 kelime)	

Türkiye'nin ormanlık bölgelerindeki IoT cihazlarından gelen büyük veri hacimlerini yönetmek için modern dağıtık veri işleme ve depolama çözümleri gerekmektedir. Verilerin hem yüksek frekansta toplanması hem de gerçek zamanlı olarak işlenmesi, sistemin ölçeklenebilir ve yüksek performanslı olmasını zorunlu kılar. Bu gereksinimleri karşılamak için Apache Kafka ve Apache Flink gibi araçların entegrasyonu oldukça etkili olabilir.

Veri Toplama ve İletme: Apache Kafka

Apache Kafka, büyük veri akışlarını yüksek verimlilikle işleyebilen dağıtık bir akış işleme platformudur. IoT cihazlarından gelen verileri toplamak için ideal bir çözümdür. Kafka'nın temel özellikleri şunlardır:

- **Yüksek Performans ve Düşük Gecikme Süresi:** Kafka, büyük miktarda veriyi düşük gecikme süresiyle işleyebilir. Bu, 50.000 IoT cihazından gelen sürekli veri akışını yönetmek için kritiktir.
- **Yatay Ölçeklenebilirlik:** Kafka, yatay olarak ölçeklenebilir, yani veri hacmi arttıkça yeni düğümler ekleyerek sistemin kapasitesini artırabilirsiniz.
- **Güvenilirlik ve Dayanıklılık:** Kafka, verileri güvenilir bir şekilde depolar ve veri kaybını önler.

Kafka, her IoT cihazından gelen veriyi belirli bir konuya (topic) yazarak toplar ve bu verileri kuyruğa alır ve tüketicilere iletir.

Gerçek Zamanlı Veri İşleme: Apache Flink

Veriler Kafka'da toplandıktan sonra, gerçek zamanlı olarak işlenmesi gerekir. Bu noktada, Apache Flink devreye girer. Flink, yüksek verimlilikle gerçek zamanlı veri akışı işleyebilen bir platformdur. Flink'in özellikleri şunlardır:

- **Durumlu Veri İşleme:** Flink, durumlu veri işleme yapabilir, bu da geçmiş verilerle ilişkili hesaplamaların yapılmasını sağlar.
- **Yüksek Performans:** Flink, düşük gecikmeli ve yüksek verimlilikle veri işleyebilir, bu da büyük veri hacimlerinin anında analiz edilmesini sağlar.
- **Esneklik:** Flink, çeşitli veri kaynaklarından veri alabilir ve işleyebilir, bu da Kafka ile entegrasyonunu kolaylaştırır.

Flink, Kafka'dan gelen verileri alır ve gerçek zamanlı analizler yaparak, anomalileri tespit edebilir, alarm üretebilir veya öngörülerde bulunabilir. Örneğin, belirli bir bölgede anormal duman miktarı algılandığında, hemen bir alarm oluşturulabilir.

Yük Dengeleme ve Ölçeklenebilirlik

Böylesine büyük bir sistemde, yük dengeleyiciler kritik bir rol oynar. Gelen istekleri ve veri akışlarını dengeli bir şekilde dağıtarak, her bir sunucunun aşırı yüklenmesini önlerler. Bu amaçla, Nginx veya HAProxy gibi yük dengeleyici çözümler kullanılabilir.

Yük dengeleyiciler, Kafka ve Flink düğümleri arasında yükü dengeler. Bu sayede, her bir düğüm optimal performansta çalışabilir. Ayrıca, sistemin ölçeklenebilirliği sağlanmış olur; veri hacmi arttıkça yeni Kafka ve Flink düğümleri eklenerek sistem kapasitesi artırılabilir.