

İstanbul Teknik Üniversitesi
Fizik Mühendisliği
Fen ve Edebiyat Fakültesi



Fizik Mühendisliği Tasarımı Raporu
Gerçek Zamanlı Sensör Verilerinin İşlenmesi
ve Analizi

Nazif ÇELİK

090200712

Danışman

Prof. Dr. Mehmet Kerim Ramazanoğlu

İçindekiler

1	Giriş	4
1.1	Tasarımın Tanımı ve Amacı	4
1.2	Tasarımın Kapsamı ve Kullanım Alanları	4
2	İş paketleri	5
2.1	Proje ihtiyaçlarının belirlenmesi	5
2.2	Geliştirme ortamı hazırlığı	5
2.3	Literatür Taraması	5
2.4	Kullanılacak teknolojilerin belirlenmesi	5
2.5	Prototip oluşturma	6
2.6	Prototipin gerçek durumlar için testi	6
2.7	Test sonrası revizesi ve optimizasyon süreci	6
2.8	Proje dokümantasyonu	6
3	Uygulama	7
3.1	Kod	7
3.2	Tanım	7
3.3	Mimari Bileşenler	7
3.3.1	Eclipse Mosquitto:	7
3.3.2	Apache Kafka	7
3.3.3	Apache Kafka Connect	8
3.3.4	Apache Kafka Streams	8
3.3.5	Prometheus	8
3.3.6	MongoDB	8
3.3.7	Grafana	8
3.4	Mimari Süreç	8
3.4.1	Veri Toplama	8
3.4.2	Veri Aktarımı	8
3.4.3	Veri İşleme	9
3.4.4	Veri Depolama	9
3.4.5	Veri Görselleştirme	9

4	Sonuçlar ve Tartışma	11
5	Referanslar ve Kullanılmış Teknolojiler	12

Bölüm 1

Giriş

1.1 Tasarımın Tanımı ve Amacı

Nesnelerin İnterneti (IoT) uygulamalarında sensörlerden sürekli olarak büyük miktarda veri toplanmakta ve bu verilerin gerçek zamanlı olarak işlenmesi gerekmektedir. Apache Kafka ve Apache Spark kullanılarak gerçek zamanlı sensör verilerinin işlenmesi ve analiz edilmesi sağlanılacaktır. İşlenen ve analiz edilen verilerin uygun formatta kullanıcıya sunulacaktır. Proje gerçek zamanlı akan veri trafiğini en kısa tepki süresiyle analiz ve sonrasında raporlanmasını hedefliyor.

1.2 Tasarımın Kapsamı ve Kullanım Alanları

Savunma Sanayii sektöründe IoMT (Internet of Military Things) alanında elde edilen radar, sonar, kamera, GPS, LIDAR gibi büyük miktarda veriyi , projenin gerçek zamanlı veri işleme kapasitesi sayesinde askeri operasyonlar için daha hızlı ve az maliyetli karar alma mekanizmasını sağlayabilir . Ayrıca bu verilere dayalı olarak potansiyel tehditlerin tahmin edilmesi, risklerin öngörülmesi ve erken müdahale imkanı sağlar. Sınır bölgelerindeki hareketlilik gerçek zamanlı olarak analiz edilerek, şüpheli aktiviteler tespit edilebilir ve güvenlik güçleri anında bilgilendirilebilir bu sayede geniş bölgelerdeki aktivitelerin daha etkin bir şekilde izlenmesine olanak tanır. Askeri araçlar arasında güvenilir ve hızlı veri paylaşımı mümkün kılınarak birlikte operasyon yürütme kapasitesi artırılabilir. Çevresel değişiklikler, ekipman tabiki, kritik bölgeler gibi hususlarda özellikle meskun mahaldeki başta olmak üzere bir çok operasyonlarda komuta merkezine daha hızlı, güvenli ve etkili karar kabiliyetleri sunar. Kullanılacak olan veri niteliği Unstructured (yapılandırılmamış) tipdeki JSON formatındaki verileri içeriyor.

Bölüm 2

İş paketleri

2.1 Proje ihtiyaçlarının belirlenmesi

Kullanıcı gereksinimleri, sistem performans hedefleri, veri güvenliği ihtiyaçları ve nihai çıktılar gibi detaylar göz önüne alınarak proje kapsamında işlenecek sensör verilerinin türü, veri hacmi, hız, ve işleme gereklilikleri gibi kriterler belirlenir.

2.2 Geliştirme ortamı hazırlığı

Geliştirme sürecinde kullanılacak donanım, yazılım ve altyapı bileşenleri hazırlanır.

2.3 Literatür Taraması

IoT, gerçek zamanlı veri işleme, Apache Kafka ve Spark ile ilgili mevcut araştırmalar ve uygulama örnekleri incelenir. Muhtemel teknik zorluklar üzerine bilgi toplanır.

2.4 Kullanılacak teknolojilerin belirlenmesi

Docker için uygun veri toplama, işleme, depolama, ve analiz teknolojileri seçilir. Ayrıca, veri görselleştirme ve analizi için kullanılacak araçlar ve entegrasyon yöntemleri belirlenir.

2.5 Prototip oluřturma

Seilen teknolojiler kullanılarak temel bir prototip geliřtirilir. Sensörün sanal ortamda simüle edilerek veri akıřı yaratılması Kafka aracılıęıyla Spark’a iletildikten sonra iřlenmesi ve analiz edilmesi adımları gerekleřtirilir.

2.6 Prototipin gerek durumlar iin testi

Prototip, gerek zamanlı sensör verileriyle ve hedef kullanıcı senaryolarına göre test edilir.

2.7 Test sonrası revizesi ve optimizasyon süreci

Sistemin kararlılıęı ve performansı deęerlendirilir, tespit edilen eksiklikler ve hatalar düzeltilir. Performansı artırmak iin Apache Kafka ve Spark konfigürasyonları ince ayarlanır. Verilerin iřleme süresi ve ıktı kalitesi optimize edilir.

2.8 Proje dokümantasyonu

Proje ierięinde kullanılan araçlar ve yöntemler detaylı bir řekilde dokümente edilir. Teknik belgeler, kullanıcı kılavuzları, ve gelecekteki geliřtirmelere yönelik öneriler hazırlanır.

Bölüm 3

Uygulama

3.1 Kod

Github Repo

3.2 Tanım

Bu Uygulama, IoT cihazlarından (genellikle sensörler) mesaj almak ve işlemek için basit bir gerçek zamanlı veri işleme ve yutma boru hattı (ETL Pipeline) önerisi oluşturur. IoT cihazları ve sensörleri binlerce ve milyonlarca mesaj (veya olay) üretir. Bu mesajları yakalayarak ve analiz ederek değerli bilgileri deşifre edebilir ve uygun önlemleri alabilir.

3.3 Mimari Bileşenler

3.3.1 Eclipse Mosquitto:

MQTT protokolünü uygulayan açık kaynaklı bir mesaj aracısıdır. Hafif yapısı sayesinde düşük güçlü cihazlardan sunuculara kadar geniş bir yelpazede kullanılabilir.

3.3.2 Apache Kafka

Gerçek zamanlı veri akışlarını yayınlama, depolama ve işleme imkanı sunan dağıtık bir veri iletim platformudur. Farklı kaynaklardan gelen veri akışlarını işleyerek çeşitli kullanıcılara dağıtabilir.

3.3.3 Apache Kafka Connect

Kafka'yı harici sistemlerle (veritabanları, dosya sistemleri vb.) entegre etmek için kullanılan bir çerçevedir. Verilerin Kafka'ya ve Kafka'dan dış sistemlere taşınmasını kolaylaştırır.

3.3.4 Apache Kafka Streams

Kafka üzerinde çalışan uygulamalar ve mikro hizmetler geliştirmek için kullanılan bir istemci kütüphanesidir. Gelen veri akışlarını işleyerek çıktıları yine Kafka kümelerine yazabilir.

3.3.5 Prometheus

Gerçek zamanlı olay izleme ve uyarı sistemi sağlayan ücretsiz bir yazılımdır. Zaman serisi verilerini toplayarak esnek sorgulama ve gerçek zamanlı uyarı özellikleri sunar.

3.3.6 MongoDB

Açık kaynaklı, belge odaklı bir NoSQL veritabanıdır. Verileri dinamik şemalarla BSON formatında depolar, bu da veri entegrasyonunu kolaylaştırır.

3.3.7 Grafana

Açık kaynaklı bir yazılım olan Grafana, metriklerin görselleştirilmesi ve izlenmesi için kullanılır. Farklı veri kaynaklarından gelen verileri panolar üzerinde görselleştirir.

3.4 Mimari Süreç

3.4.1 Veri Toplama

IoT sensörleri, her saniye sıcaklık, nem, basınç ve ışık verilerini üretir. Bu veriler, Eclipse Paho MQTT kütüphanesi kullanılarak Eclipse Mosquitto aracısına iletilir.

3.4.2 Veri Aktarımı

Kafka Connect, Mosquitto'dan gelen verileri alarak Apache Kafka'ya iletir. Bu işlem için Confluent'ın MQTT kaynağı bağlayıcısı kullanılır.

3.4.3 Veri İşleme

Apache Kafka Streams, Kafka'daki verileri işleyerek zaman pencereleri üzerinden ortalama sıcaklık, nem gibi hesaplamalar yapar.

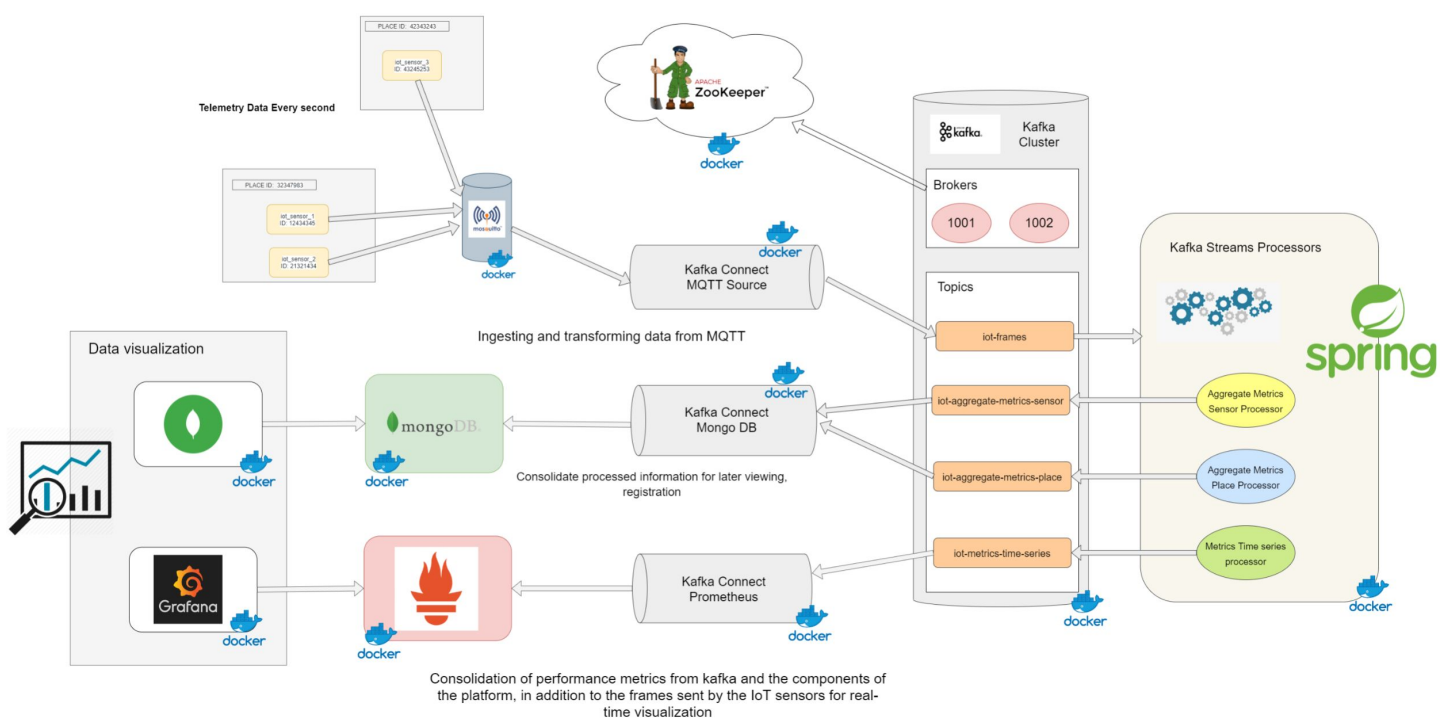
3.4.4 Veri Depolama

İşlenen veriler, MongoDB'de depolanır. MongoDB, sık yazma işlemleri ve homojen belge boyutları için uygun bir NoSQL veritabanıdır.

3.4.5 Veri Görselleştirme

Grafana, MongoDB'deki verileri kullanarak sensör metriklerini ve mimari bileşenlerin performansını gerçek zamanlı olarak görselleştirir.

Internet of Things (IoT) and Event Streaming at Scale with Apache Kafka and MQTT



Bölüm 4

Sonuçlar ve Tartışma

Bu çalışma kapsamında, Savunma Sanayii sektöründe IoMT (Internet of Military Things) alanında elde edilen radar, sonar, kamera, GPS, LIDAR gibi büyük miktarda veriyi , projenin gerçek zamanlı veri işleme kapasitesi sayesinde askeri operasyonlar için daha hızlı ve az maliyetli karar alma mekanizmasını sağlayabilir . Ayrıca bu verilere dayalı olarak potansiyel tehditlerin tahmin edilmesi, risklerin öngörülmesi ve erken müdahale imkanı sağlar. Sınır bölgelerindeki hareketlilik gerçek zamanlı olarak analiz edilerek, şüpheli aktiviteler tespit edilebilir ve güvenlik güçleri anında bilgilendirilebilir bu sayede geniş bölgelerdeki aktivitelerin daha etkin bir şekilde izlenmesine olanak tanır. Askeri araçlar arasında güvenilir ve hızlı veri paylaşımı mümkün kılınarak birlikte operasyon yürütme kapasitesi artırılabilir. Çevresel değişiklikler, ekipman takibi, kritik bölgeler gibi hususlarda özellikle meskun mahaldeki-ler başta olmak üzere bir çok operasyonlarda komuta merkezine daha hızlı, güvenli ve etkili karar kabiliyetleri sunar.

Bölüm 5

Referanslar ve Kullanılmış Teknolojiler

- OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]
- Spring Boot 2.3.3 / Apache Maven 3.6.3
- Kafka Streams
- Micrometer Registry Prometheus
- Eclipse Paho MQTT Client
- Kafka Connect
- Kafka Rest Proxy
- Jackson
- NodeExporter (Exporter for machine metrics)
- Prometheus
- Grafana
- Eclipse Mosquitto
- Mongo DB Express (Web-based MongoDB admin interface, written with Node.js and express)
- Cadvisor (Analyzes resource usage and performance characteristics of running containers)
- kafka-exporter (Kafka exporter for Prometheus)