



SAKARYA ÜNİVERSİTESİ
Bilgisayar ve Bilişim Bilimleri Fakültesi
Bilgisayar Mühendisliği Bölümü

BSM 313
NESNELERİN İNTERNETİ VE UYGULAMALARI

(Internet of Things (IoT) and Applications)

BÜYÜK VERİ VE BULUT BİLİŞİM

Prof. Dr. Cüneyt BAYILMIŞ



Büyük Veri ve Nesnelerin İnterneti

(Big Data and IoT)

- ❑ Birbirine bağlı çok sayıda fiziksel cihazın veya sensörün internet üzerinden ürettiği veri büyük veri (**big data**) olarak adlandırılır.
 - Büyük veri, kavramsal olarak verinin hacmini ifade etmektedir.
 - Büyük veri, büyük hacimli, yüksek hız ve çok çeşitli verilerin yönetimi ve analizini ifade etmektedir
- ❑ IOT Verileri saklanmalı, işlenebilmeli ve anlamlandırılmalıdır.
- ❑ Saklama için verimli ve efektif disk alanına ihtiyaç vardır.
- ❑ Big Data kabul edilebilir bir sürede işlenebilir olmalıdır.
- ❑ Bulut sistemler uzaktan yönetilebilir, istenildiği zaman boyutu ölçeklenebilir yapıdadır.
- ❑ Bunlardan dolayı IOT cihazlarından toplanan veriler Bulut sistemler üzerinde hesaplanabilir ve anlamlandırılabilir.

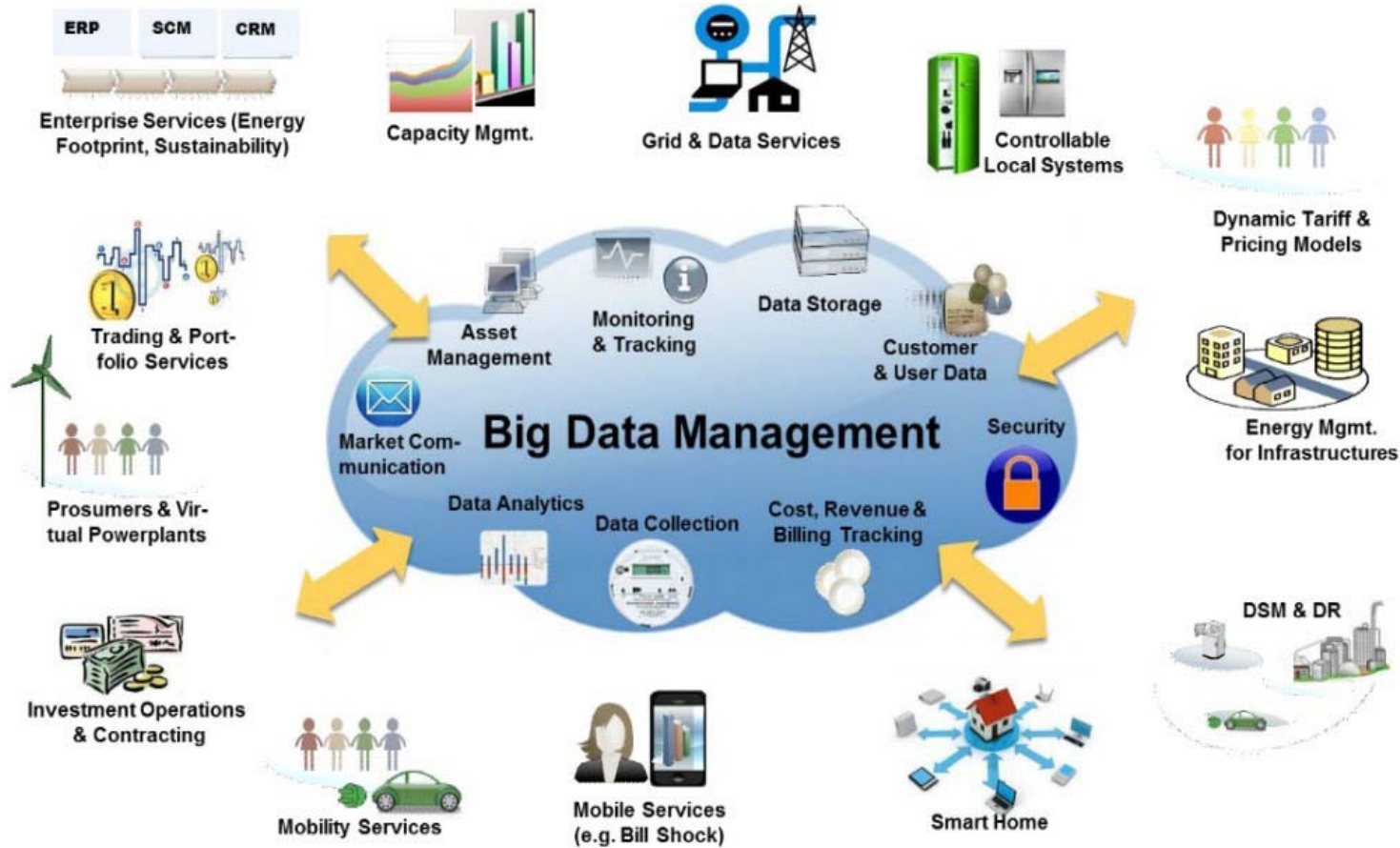


Büyük Veri Artma Nedenleri

- ❑ Sosyal ağ (facebook, twitter, linkedin vb.) kullanımının artması,
- ❑ Akıllı sensörler sayesinde lokasyon duyarlı cihazların artması,
- ❑ Fiziksel dünya hakkında bilgi yakalayan ve ileten akıllı sensörlerin sayısındaki artış
- ❑ Özet olarak sebeplerden biri olarak;
nesnelerin interneti diyebiliriz.

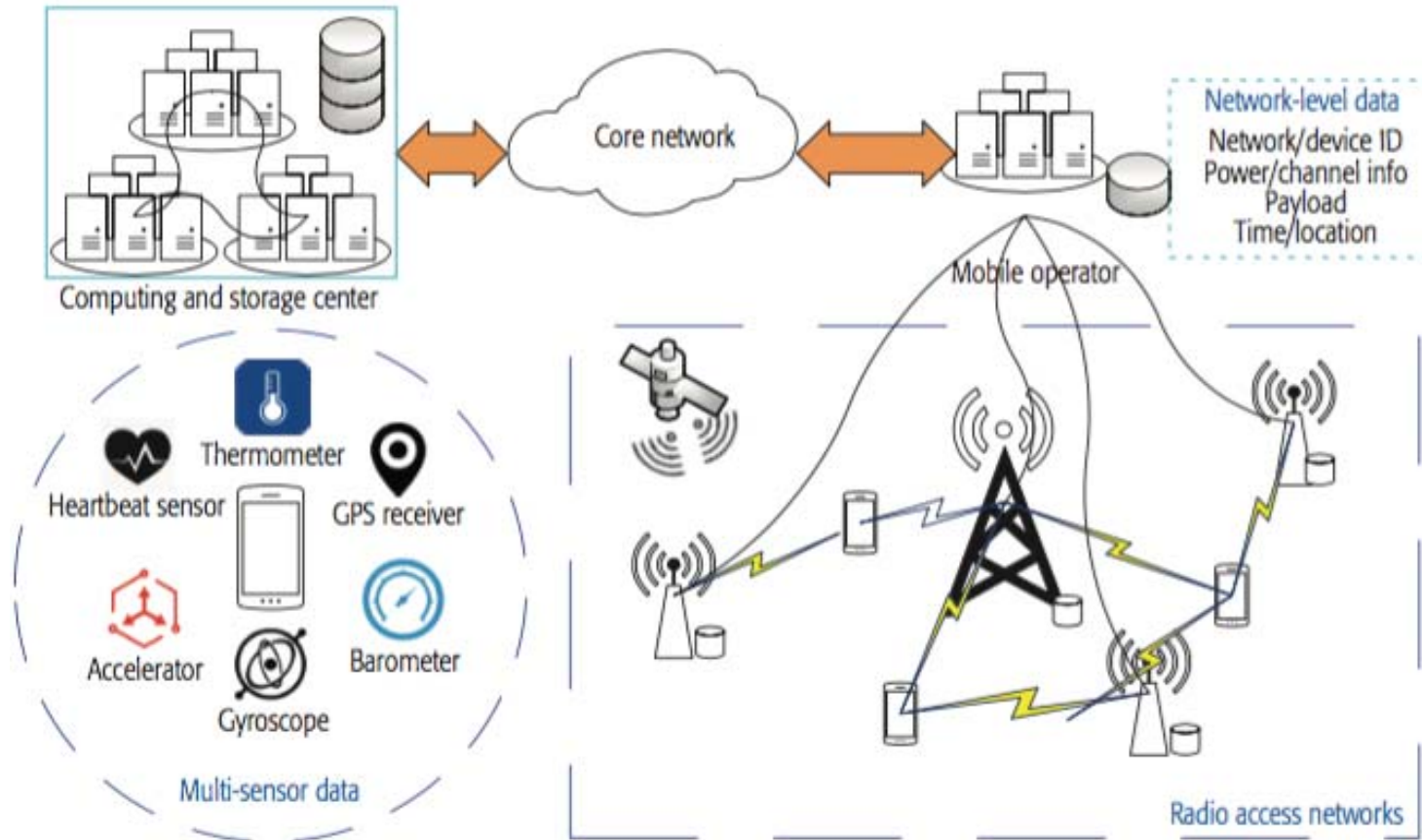


Büyük Veri Yönetimi

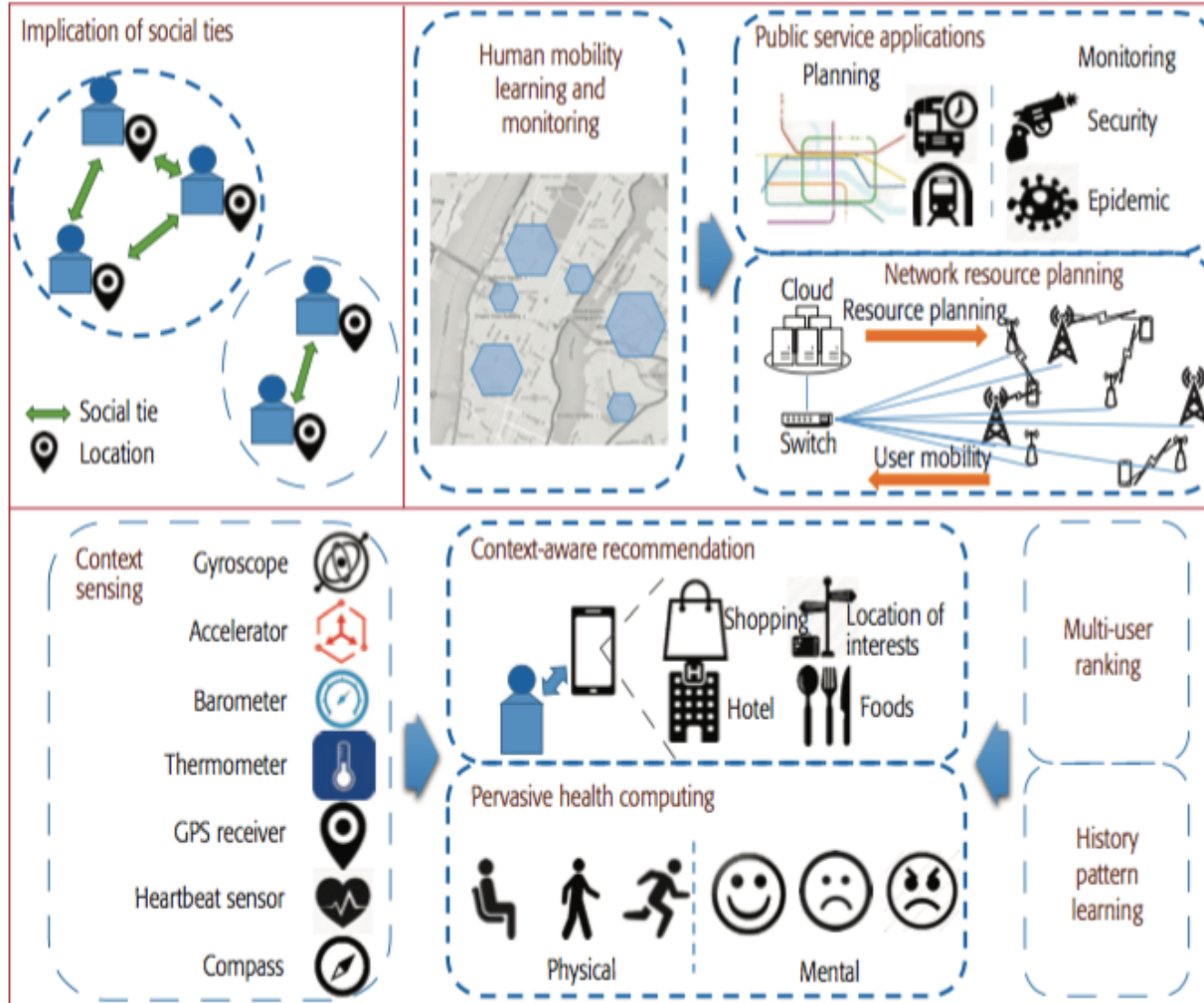


Mobil Büyük Veri Kaynakları

- ❑ Mobil veri, gezgin kullanıcılar (akıllı telefon) ya da mobil ağlardan (gateway, baz istasyonu) elde edilir.
- ❑ Mobil veri iki sınıfa ayrılır.
 - Uygulama Verisi: Akıllı telefon ya da gömülü sistem sensörlerinden (GPS, ivmeölçer vb.) toplanan veriler.
 - Ağ Verisi: Ağ operatörleri tarafından toplanan veriler (Kullanıcı ID, konum, cihaz tipi, servis tipi vb.)



Mobil Büyük Veri Uygulamaları



❑ Mobil veri uygulamaları kişisel servisler geliştirmek için bireysel veri ve topluluk aktivitelerinin analizine dayalı topluluk verisi olarak sınıflandırılabilir.

- Bireysel servisler, içerik farkında öneriler, aktivite tanıma vb.
- Topluluk servisleri ise kamu servisleri, akıllı şehir uygulamaları, trafik yönetimi vb.

Kaynak: X. Cheng, L. Fang, X. Hong, and L. Yang, "Exploiting Mobile Big Data: Sources, Features, and Applications," IEEE Network, vol. 16, no. 1, pp. 72-79, 2017

Büyük Veri ve Nesnelerin İnterneti

(Big Data and IoT)

Büyük veri için zorluklar

❑ Hacim (Volume)

- Büyük verinin, veri boyutu, veri kümesi ile ilgilenir.

❑ Hız (Velocity)

- Verinin üretilme hızı ile ilgilenir. Verinin işlenmesi için çeşitli özniletiklerin hesaplanması için gerekli algoritmaların analizini ele alır.

❑ Çeşitlilik (Variety)

- Veri kümesindeki yapısal çeşitliliği ifade eder.

❑ Doğruluk (Veracity)

- Verilerin doğrulunu ele alır.

❑ Değer (Value)

- Verilerin belirleyici özelliği olarak ifade edilen değerdir.



Büyük Veri ve Nesnelerin İnterneti

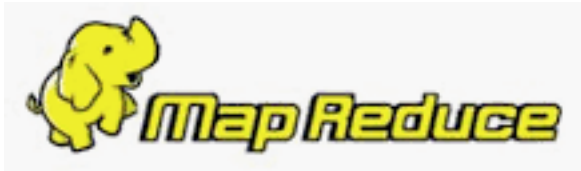
(Big Data and IoT)

Özellik	Zorluk	Teknik (Çözüm Yöntemi)
Hacim (Volume)	Depolama / Ölçek	Dağıtık Dosya Sistemleri
Hız (Velocity)	Hızlı İşleme	Paralel Programlama
Çeşitlilik (Variety)	Heterojenlik	NoSQL Veritabanları
Değer (Value)	Bilgi Keşfi (Knowledge Discovery)	Veri Madenciliği
Doğruluk (Veracity)	Anlambilim (Semantics) Analitik (Analytics)	Algoritmaları



Nesnelerin İnterneti İçin Büyük Veri Analizi

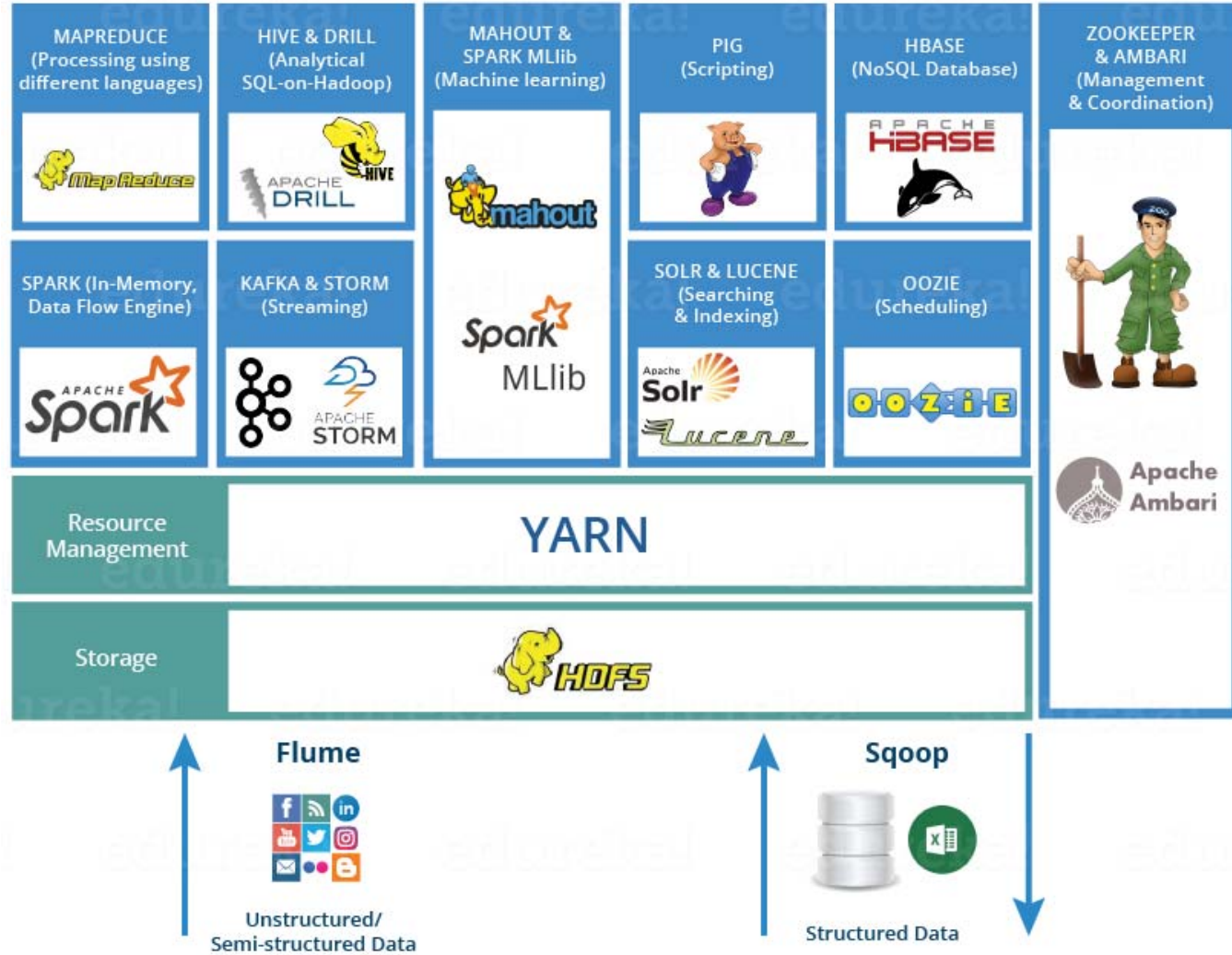
- ❑ Büyük Veri için analiz platformları



- ❑ Bu platformlar nesnelerin interneti gereksinimlerini karşılamaktadır.
- ❑ Gerçek zamanlı (real-time) çalışmayı desteklerler.
- ❑ Özel olarak geliştirilmiş analiz programları yerine nesnelerin interneti için yaygın olarak kullanılan platformlardan yararlanmak daha iyi bir çözümdür.

IoT Büyük Veri Araçları

❑ Hadoop Ekosistemi

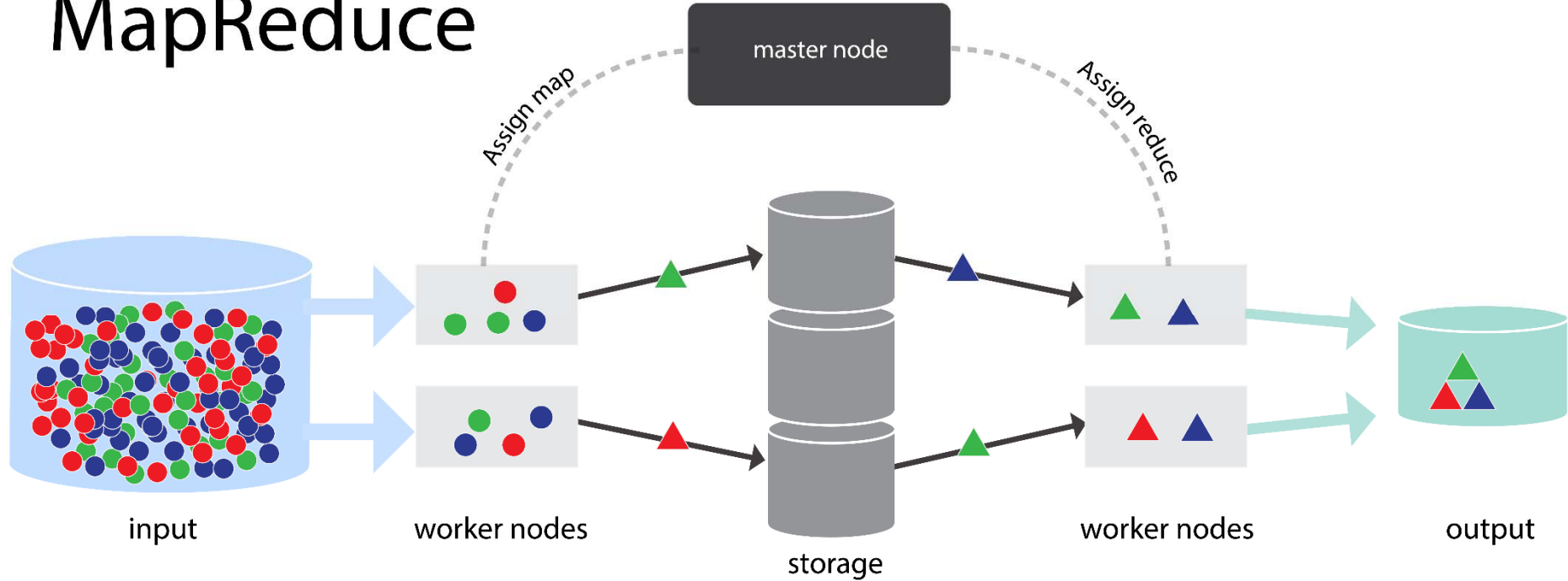




- ❑ Hadoop, açık kaynak kodlu Java için geliştirilmiş büyük veri işleme aracıdır.
- ❑ Depolama bölümü olarak dağıtık veri sistemi olarak bilinen HDFS (Hadoop File System), işleme modülü olarak ta MapReduce kullanmaktadır.
- ❑ Hadoop, HDFS ve MapReduce bileşenlerinden oluşan bir yazılımdır.
- ❑ Hadoop tarafından analiz edilecek veri, HDFS üzerinde tutulmaktadır.
Genel olarak Hadoop birden fazla bilgisayarın oluşturduğu kümeler (**cluster**) üzerinde koşar. Bu durum hem verinin hem de işlerin dağıtılmasına imkan verir.
- ❑ Hadoop'un sunduğu özellikler:
 - İşlem gücü, hata toleransı, esneklik, düşük maliyet, ölçeklenebilirlik.
- ❑ En yaygın kullanım örneği Facebook,

- ❑ MapReduce, 2004 yılında Google tarafından geliştirilmiş, verilerin harita ve azaltma fonksiyonları kullanır.
- ❑ Büyük veriyi, büyük kümelerle işlemeyi amaçlayan bir dağıtık programlama modelidir.
- ❑ MapReduce iki kısımdan oluşur:
 - **Map**: Bir yığının tüm üyelerini sahip olduğu fonksiyon ile işleyerek bir sonuç listesi döndürür. Çok büyük veri kümelerini bucket olarak adlandırılan iki veya daha fazla alana/kovaya böler.
 - **Reduce**: Paralel bir şekilde çalışan iki ya da daha fazla map fonksiyonundan dönen sonuçları harmanlar ve işler/çözer.
- ❑ MapReduce programı farklı dillerde geliştirilebilir.
- ❑ Google, web sayfalarını indekslemek için MapReduce kullanır.

MapReduce



Map Phase

- ❑ Ana (master) düğüm verileri alıp, daha ufak parçalara ayırıp işçi düğümlere dağıtır.
- ❑ İşçi düğümler bu işleri tamamladıkça sonucunu ana düğüme geri gönderir.
- ❑ Özetle, analiz edilen veri içerisinde almak istenilen veriler çekilir.

Reduce Phase

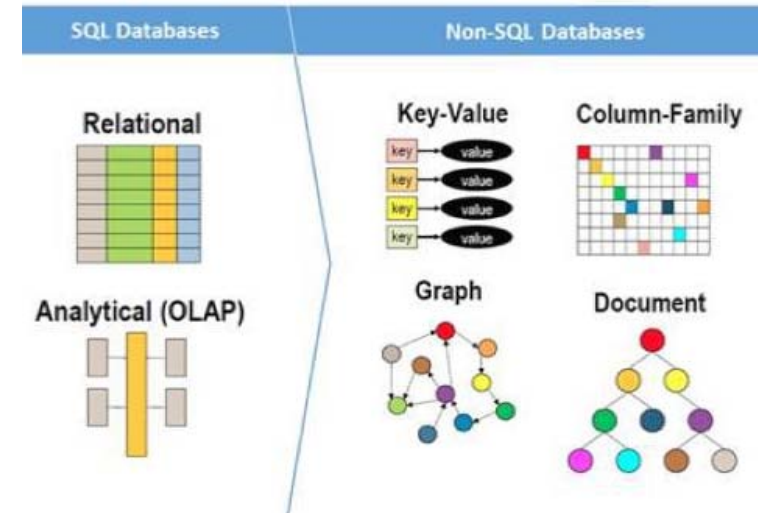
- ❑ Tamamlanan işler için mantığına göre birleştirilerek sonuç elde edilir.
- ❑ Özetle, çekilen veri üzerinde istenilen analiz gerçekleştirilir.

Nesnelerin İnterneti Büyük Veri Araçları

NoSQL



- ❑ SQL dili kullanılmadığından 'Not Only SQL' olarak ta adlandırılır.
- ❑ NOSQL kullanan sistemlere örnek olarak Google ve Amazon tarafından kullanılan Big Table, DynamoDB verilebilir.
- ❑ **Veri için bir kullanım ömrü belirlenebilir** (bir anahtara bağlanabilir), çok uzun tutulmayı gerektirmeyen veriler NoSQL'de tutulabilir.
- ❑ Çok veri trafiğine sahip sitelerde veritabanından yazma ve okuma işlemlerini hızlandırmak için **veriler sütun halinde tutulabilir**.
- ❑ Veriler doküman gibi saklanabilir. (MongoDB)
- ❑ Çizge kuramlı (Graph) veritabanları, kullanıcılar arasındaki bağları/ilişkileri yakalayabilmeyi kolaylaştırır (sosyal ağlar).
- ❑ Hiyerarşik veritabanları, coğrafik veritabanı türleri.

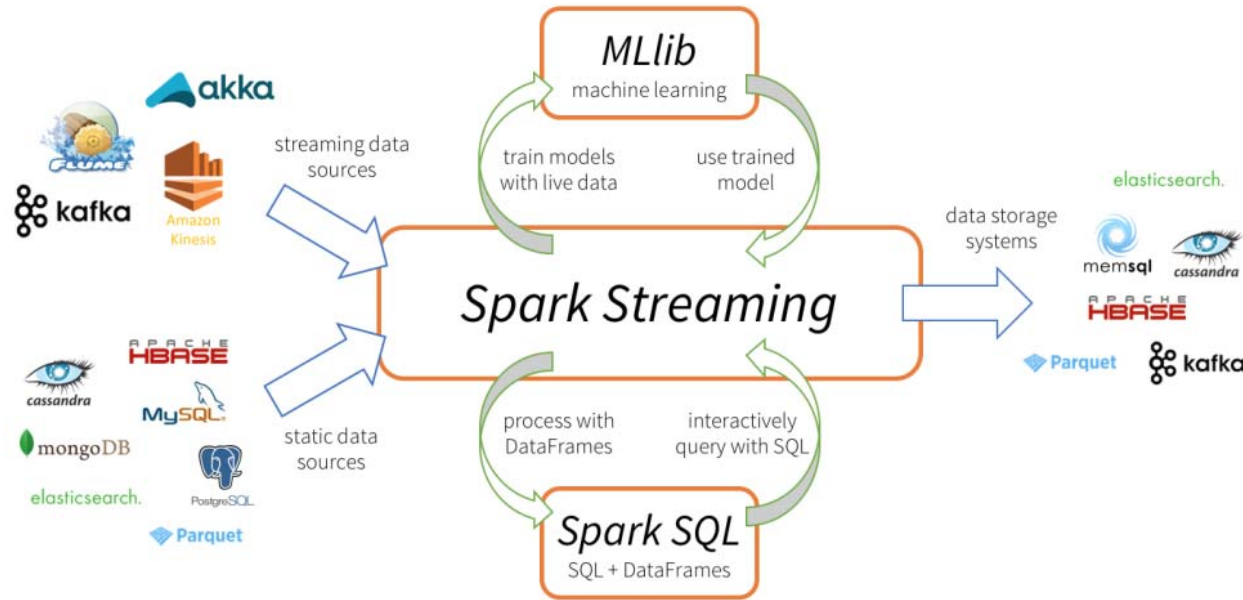


- ❑ Doküman odaklı bir veri modeli kullanan açık kaynaklı NoSQL veritabanıdır.
- ❑ İlişkisel veritabanlarındaki gibi tablolar ve satırlar yerine koleksiyon ve belgeler mimarisi üzerine inşa edilmiştir.
- ❑ NoSQL veritabanlarına benzer olarak, bir koleksiyondaki belgelerin farklı alanlara ve yapılar sahip olmasına imkan tanıyan dinamik şema tasarımını destekler.
- ❑ Belge saklama ve veri değişimi formatı olarak BSON kullanır. İkili temsilini JSON benzeri belgeler.
- ❑ Önemli Özellikleri
 - Master, slave çoğaltmasını destekler.
 - Boyut farketmeksizin dosyaları kolayca depolar.
 - Birden fazla sunucu üzerinde çalışabilir.
 - Parçalara yerleştirilen veriler sebebiyle otomatik bir yük dengeleme konfigürasyonuna sahiptir.

IoT Büyük Veri Araçları

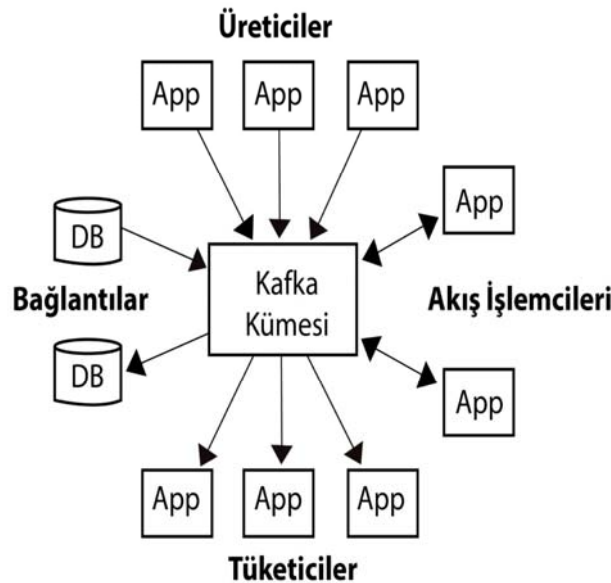


- ❑ Apache Spark, genel amaçlı, hızlı, yüksek seviye API desteği ve zengin büyük veri araçları sunan küme hesaplama platformudur.
- ❑ Hadoop'un MapReduce yapısına alternatif olarak geliştirilmiştir.
- ❑ Farklı dilleri (R, Python, Scala, Java) destekleyen açık kaynaklı bir platformdur.
- ❑ **Desteklediği Birimler/Teknolojiler**, **Spark SQL** (veri analizi için sorgu işlemleri), **Spark Streaming** (akış ile alınan veri işlemleri), **MLlib** (Makine öğrenmesi), **GraphX** (Grafik algoritmaları)



- ❑ Apache Storm, gerçek zamanlı olarak akışlı (stream) verileri işlemek için kullanılan bir sistemdir.
- ❑ Yüksek hacimli verilerin işlenmesi için dağıtılmış bir gerçek zamanlı hesaplama sistemidir.
- ❑ **Özellikleri**
 - **Yüksek hızlı çalışma** (Düğüm başına saniyede bir milyondan fazla kayıt yapma)
 - **Kolay kullanım** (Standart konfigürasyonlar)
 - **Güvenilir** (Her bir veri biriminin (tuple) en az bir kez işlenmesini garanti eder. Sadece arızalar oluştuğunda mesajlar tekrarlanır)
 - **Ölçeklenebilirlik** (Makinelerin bir kümesinde çalışan paralel hesaplama imkanı)
 - **Hata toleransı** (İşçiler öldüğünde, storm onları otomatik olarak yeniden başlatır. Bir düğüm ölürse, işçi bir başka düğümde yeniden başlatılır)

- ❑ Apache Kafka, dağıtık veri akış platformudur.
- ❑ Sistemler ya da uygulamalar arasında güvenilir şekilde gerçek zamanlı veri akış hatları oluşturur.
- ❑ Küme olarak adlandırılan bir ya da birden fazla sunucuda çalıştırılabilir.
- ❑ Kafka kümesi, konu adında kategorilerde kayıt akışları saklar. Her kayıt, bir anahtar, bir değer ve bir zaman damgasından oluşur.
- ❑ Üretici, Tüketici, Akış ve Bağlayıcı olmak üzere 4 temel API desteği sunar.



❑ 2010 yılında geliştirilmeye başlanmış, dağıtık, ölçeklenebilir, açık kaynak, kurumsal dereceli arama motorudur.

❑ Bir NoSQL'dir.

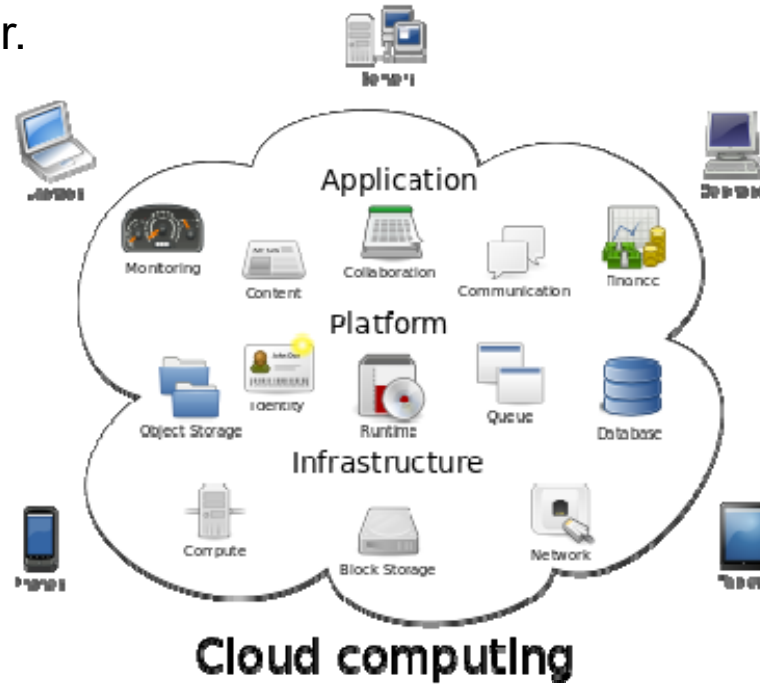
❑ Özellikleri

- Geniş hacimli verilere karşı hızlı ve duyarlı arama
- Belgelerin depoya indekslenmesi
- Normalleştirilmemiş belge depolama, verilere hızlı ve doğrudan erişim
- Geniş dağıtılabirlik ve yüksek ölçeklenebilirlik

Bulut Biliřim

(Cloud Computing)

- ❑ Bulut kavram olarak, ok sayıda cihaz/bilgisayar kaynaklı veri merkezlerini ve kullanıcıların bu veri kaynaklarına internet zerinden erişim imkanı veren ađ sistemini ifade etmektedir.
- ❑ Bulut Biliřim ise, kullanıcıların bilgi ve kaynaklara (depolama alanları, cihazlar vb.) internet zerinden erişimi, paylařımı ve bilgi zerinde ortak iřlem yapmalarını sađlayan bir teknolojidir.



Bulut Bilişim Kullanımının Faydaları

(Benefits of Cloud Computing)

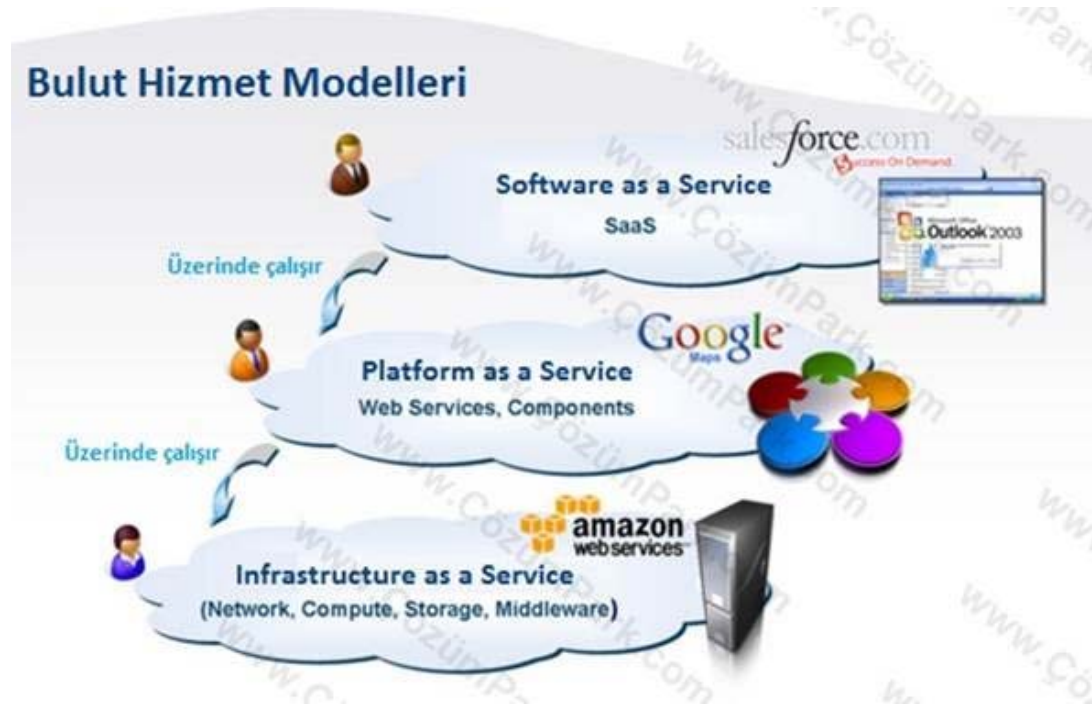
- ❑ Esneklik ve Verimlilik: Kapasite artırma ve azaltma işlemlerinin kolaylığı,
- ❑ İnternet erişimi olan her yerden kullanım imkanı,
- ❑ Güvenlik,
- ❑ Yazılım güncelleme,
- ❑ Maliyet kazancı (düşüklüğü),



Bulut Bilişim

(Cloud Computing)

- ❑ Bulut teknolojisi, yazılım, platform ve altyapı servislerini sağlar.
 - Altyapı olarak fiziksel donanım (depolama alanları) temini, sanallaştırma vb.
 - Uygulama geliştirme amaçlı veri alanı temini, erişilebilirlik vb. (Windows Azure, Google Apps vb.),
 - Yazılım olarak ise, sağlayıcıların buluta uygulama yazılımlarını yükleme ve erişim hizmeti sunması,



Bulut Bilişim

(Cloud Computing)

❑ Servis olarak Yazılım (Software as a Service, SaaS)

- Geleneksel sistemlerdeki yazılımın kurulumu, lisanslama, destek gibi iş yüklerini ortadan kaldırır.
- Kullanıcılara gerekli altyapı ve platform üstünde belirli uygulamaları kullanmaya imkan sağlar.

❑ SaaS modelindeki yazılım ürünleri

- Kurumsal Kaynak Planlaması (Enterprise Resource Planning, ERP)
- Müşteri İlişkileri Yönetimi (Customer Relationship Management, CRM)
- İnsan Kaynakları (İK),

❑ Avantajları

- Kullanıcının kullanılan teknolojiler ile ilişkisinin olmaması,
- Teknik destek kolaylığı (On-line),
- İnternet üzerinden her yerden kullanılabilmesi,
- Maliyet (kurulum, lisanslama, destek vb.)

❑ Dezavantajları

- IaaS ve PaaS ile uyum problemleri,
- Üründe sınırlı değişiklik yapabilme,



Nesnelerin İnterneti ve Bulut Bilişim

(Big Data and Cloud Computing)

Karşılaşılan Zorluklar

- ❑ Senkronizasyon : Bulut Sistem tarafında cihazların verileri eş zamanlı alma zorluğu
- ❑ Standardizasyon: IOT tarafında farklı üreticilerin cihazları arasındaki iletişim zorluğu
- ❑ Dengeleme : Farklı altyapıların bulut bilişimde dengelenmiş servis sunabilmesi
- ❑ Güvenlik : Bulut ve cihazlar arasındaki güvenlik
- ❑ Yönetim : IoT cihazlarının, Bulut Bilişimin kaynaklarının ve bileşenlerinin ayrı olması
- ❑ Güvenilirlik : Kullanıcıların istediği sonuçların doğrulanması



Nesnelerin İnterneti ve Bulut Bilişim

(Big Data and Cloud Computing)

Platform Karşılaştırmaları

- ❑ IOT sensörleri ve uç noktaları için bir çok bulut bilişim uygulaması vardır.
- ❑ Genel olarak bulut sistemlere bağlanmak için API sunarlar.
- ❑ Sensör ağlarının bağlanabileceği protokolleri destekleyebilirler. Hemen hemen hepsinde IOT cihazları için RESTful servisleri vardır.
- ❑ Sonuçları XML, CSV olarak dışa aktarılabilirdiği gibi grafiksel arayüzleri de bulunabilir.
- ❑ IOT verilerine göre uyarı veya alarm verilebilir.
- ❑ IOT verilerine göre uç cihazlar kontrol edilebilir.



Nesnelerin İnterneti ve Bulut Bilişim

(Big Data and Cloud Computing)

Platform Karşılaştırmaları

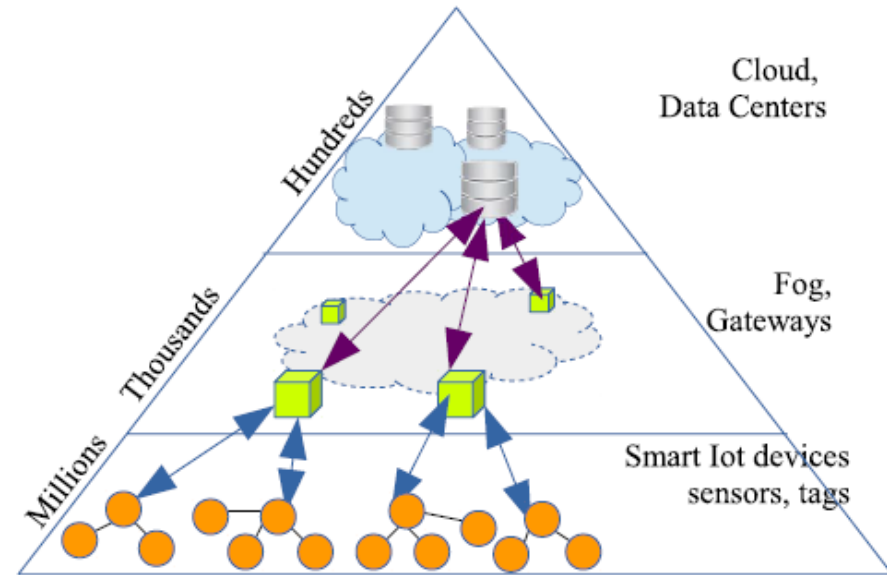
Platform	Gateway	Provision	Assurance	Billing	Application Protocol			
					REST	CoAP	XMPP	MQTT
Arkessa	-	+	+	-	+	-	-	+
Axeda	+	+	+	+	+	-	-	-
Etherios	+	+	+	-	+	-	-	-
LittleBits	-	-	-	-	+	-	-	-
NanoService	+	+	+	-	+	+	-	-
Nimbits	-	-	-	-	+	-	+	-
Ninja Blocks	+	-	-	-	+	-	-	-
OnePlatform	+	+	+	-	+	+	+	-
RealTime.io	+	+	-	-	+	-	-	-
SensorCloud	+	+	-	-	+	-	-	-
SmartThings	+	+	-	-	+	-	-	-
TempoDB	-	-	-	-	+	-	-	-
Thingworx	-	+	+	-	+	-	-	+
Xively	+	+	+	+	+	-	-	+



Nesnelerin İnterneti için Sis Bilişim

(Fog Computing for IoT)

- ❑ IoT cihazların üretmiş oldukları veriyi işlenmek üzere merkezi bir sunucuya göndermek yerine önce yerel bir noktada analiz edilmesini ve analiz sonucu ihtiyaç olunan kadarının merkezi sunucuya (buluta) gönderilmesini temel alan mimaridir.
- ❑ Sis Bilişim, Bulut Bilişim ve IOT cihazları arasında yer alır. Köprü görevi görür.
- ❑ IOT için cevap zamanını azaltıp daha iyi performans sağlar.
- ❑ Bant genişliğinin kullanımını azaltır.
- ❑ Güvenlik ve gizliliği artırır.



Nesnelerin İnterneti için Sis Bilişim

(Fog Computing for IoT)

Sis Bilişimin Nesnelerin İnternetine Sağladığı Avantajlar

- ❑ Konum: Sis Bilişim IOT Cihazlara daha yakındır.
- ❑ Dağıtık: Küçük parçalar ayrılıp mikro merkezler oluşturulabiliyor.
- ❑ Ölçeklenebilirlik : Bulut sistemlere göre kurulumu daha az maliyetli olduğu için limite ulaşımlarda yeni mikro merkezler oluşturulabilir.
- ❑ Cihaz Yoğunluğu
- ❑ Mobil Destek
- ❑ Gerçek Zamanlı İletişim: Bulut sisteme göre cevap zamanı daha hızlıdır.
- ❑ Ön Analiz: Veri boyutu anlamlandırılarak azaltılabilir.



KAYNAKLAR

❖ Temel Kaynaklar

- Prof. Dr. Cüneyt BAYILMIŞ ve Prof. Dr. Kerem KÜÇÜK, “Nesnelerin İnternet’i: Teori ve Uygulamaları”, Papatya Yayınevi, 2019.

❖ Diğer Kaynaklar

- A. Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, “*Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications*”, IEEE Communication Survey&Tutorials, vol. 17 (4), 2347-2376 ,2015.
- C. Tsai, C. Lai, M. Chiang, and L. T. Yang, “Data mining for Internet of Things: A survey,” IEEE Commun. Surveys Tuts., vol. 16, no. 1, pp. 77– 97, 1st Quart. 2014
- TÜBİTAK Bulut Bilişim ve Büyük Veri Araştırma Laboratuvarı, b3lab.org
- X. Cheng, L. Fang, X. Hong, and L. Yang, “Exploiting Mobile Big Data: Sources, Features, and Applications,” IEEE Network, vol. 16, no. 1, pp. 72-79, 2017