

IBM “Big Data”

Emre Uzuncakara

emre@tr.ibm.com

Big Data Sales

“Büyük Veri” Nedir?

Hacim

12 terabyte
Günlük Tweet verisi

Ürün Analizi

Hız

5 milyon
Ticari hareket - saniyede

Potansiyel suistimal hareketleri

Çeşitlilik

100's video
Gözetleme kameraları

Olay inceleme

350 Milyar
Yıllık akıllı sayaç

Enerji kullanımını tahminleme

500 milyar
Günlük CDR kaydı

Churn analizi

80% Veri büyüme oranı
Resim, video, doküman ...

Müşteri memnuniyetini geliştirme

Hangi Veriler?

Uygulama Verileri



- Hacim
- Biçimsel

Makine Verileri



- Akışkan / Sürekli
- Yarı biçimsel

Sosya Medya



- Çeşitlilik
- Biçimsel olmayan
- Tutarlılık sorunları

Kurumsal İçerik



- Çeşitlilik
- Biçimsel olmayan
- Hacim

Neler Yapılabilir?

Finans

- Suistimal Tespiti
- 360° Müşteri Analizi



Enerji

- Hava durumu analizi
- Enerji yönlendirme

Lojistik

- Optimizasyon
- Trafik önleme



IT

- Log Analizi
- Siber güvenlik



Sağlık

- Erken uyarı sistemleri
- Hasta kayıt analizi



Perakende

- 360° Müşteri Analizi
- Gerçek zamanlı kampanya



Telco

- Pazarlama
- Network analizi



Güvenlik

- Çok yönlü gözetleme
- Siber güvenlik
- Erken suç uyarı sistemleri

Büyük Veri Yaklaşımı

Geleneksel Yaklaşım

Biçimsel Veri ve Tekrarlayan Raporlar

İş Ünitelerinin yapacağı analizler hemen hemen belirlidir



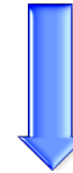
BT sorulacak soruları ve sorguları en hızlı yanıtlayacak platformu oluşturur



Big Data Yaklaşımı

İteratif ve Keşfe Yönelik

BT Yaratıcı keşiflere yönelik bir platform ve veri zenginliği sunar



İş Ünite'leri neler keşfedebileceğini araştırır, tarihsel veriye özgürce erişirler



Büyük Veri problemleri her kurumda karşımıza çıkabiliyor

Bir kamu kurumu ses kayıtlarının analizini saatlerden **70 Milisaniye'**ye düşürdü



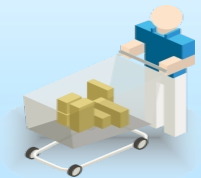
Bir enerji üreticisi güç iletimi problemlerini engellemek için **10 PB** veriyi dakikalar mertebesinde analiz ediyor



Bir hastane akışkan sağlık indikatörlerini sürekli izleyip, **24 saat önceden** hayati problemleri görebiliyor



Bir perakende şirketi sorgu süreleri **%80** oranında düşürdü ve stoklarını optimize etti



Büyük bir borsa **2 PB veri** üzerinde 26 saat süren analizleri **2 dakikaya** kadar indirdi



Bir telekom şirketi network indikatörlerini artık akışkan veri üzerinden takip ediyor ve donanım maliyetlerini **%90** azalttı.



Ham Veriyi İnceleyin --biçimi ne olursa olsun

(IBM InfoSphere BigInsights)

Veri Ambarınızı hızlandırın

(IBM PureData System for Analytics)

Veri Ambarı maliyetlerinizi düşürün

(IBM InfoSphere BigInsights)

Akışkan Veriyi İşleyin

(IBM Infosphere Streams)

Verinizi Keşfedin

(IBM Infosphere Data Explorer)

Analitik Uygulamalar

BI / Raporlama	Analiz/ Görselleştirme	Fonksiyonel Uyg.	Endüstri Uyg.	Veri Madenciliği	İçerik Analitiği
-------------------	---------------------------	---------------------	------------------	---------------------	---------------------

IBM Big Data Platformu

Görselleştirme
ve Keşif

Uygulama
Geliştirme

Sistem
Yönetimi



Hızlandırıcılar

Hadoop
Sistemi



Akışkan
veri işleme



Veri
Ambarı



Veri Entegrasyonu ve Sahipliği



“Big Data” Platformu – Hadoop Sistemi

InfoSphere BigInsights

- Çok çeşitli ve yüksek hacimli verileri yönetir
- Açık kaynak Hadoop sistemine kurumsal yaklaşımlar getirdi
 - Performans Optimizasyonu
 - Geliştirme Platformu
 - Kurumsal Uygulama Entegrasyonu
 - Analitik Hızlandırıcılar
 - Uygulama Hızlandıcıları
 - Görsellik
 - Güvenlik



InfoSphere Streams – “Big Data” için Akan Veri Analizi

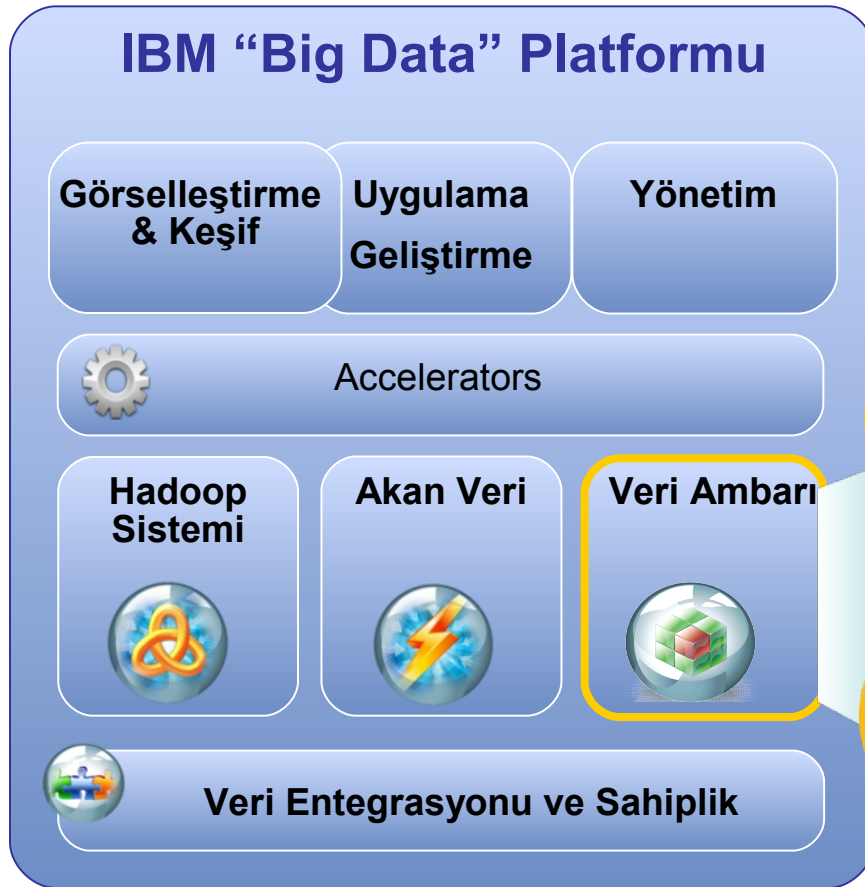
- Hareket halindeki verinin analizi için tasarlandı
 - Çoklu veri işleme
 - Esnek Ölçeklenebilirlik
- Çeşitli verilerin işlenmesi ve analizi
 - Biçimsel, biçimsel olmayan, video, ses, vb.
 - Gelişmiş analitikler
- Gerçek zamanlı entegrasyon
 - Veri ambarı
 - Hadoop Sistemi



The IBM Big Data Platformu

Farklı Analitikler için Farklı Çözümler

PureData



Kompleks Analitik

Analitik ve Raporlama

Operasyonel Analitik

Operasyonel Sistemler



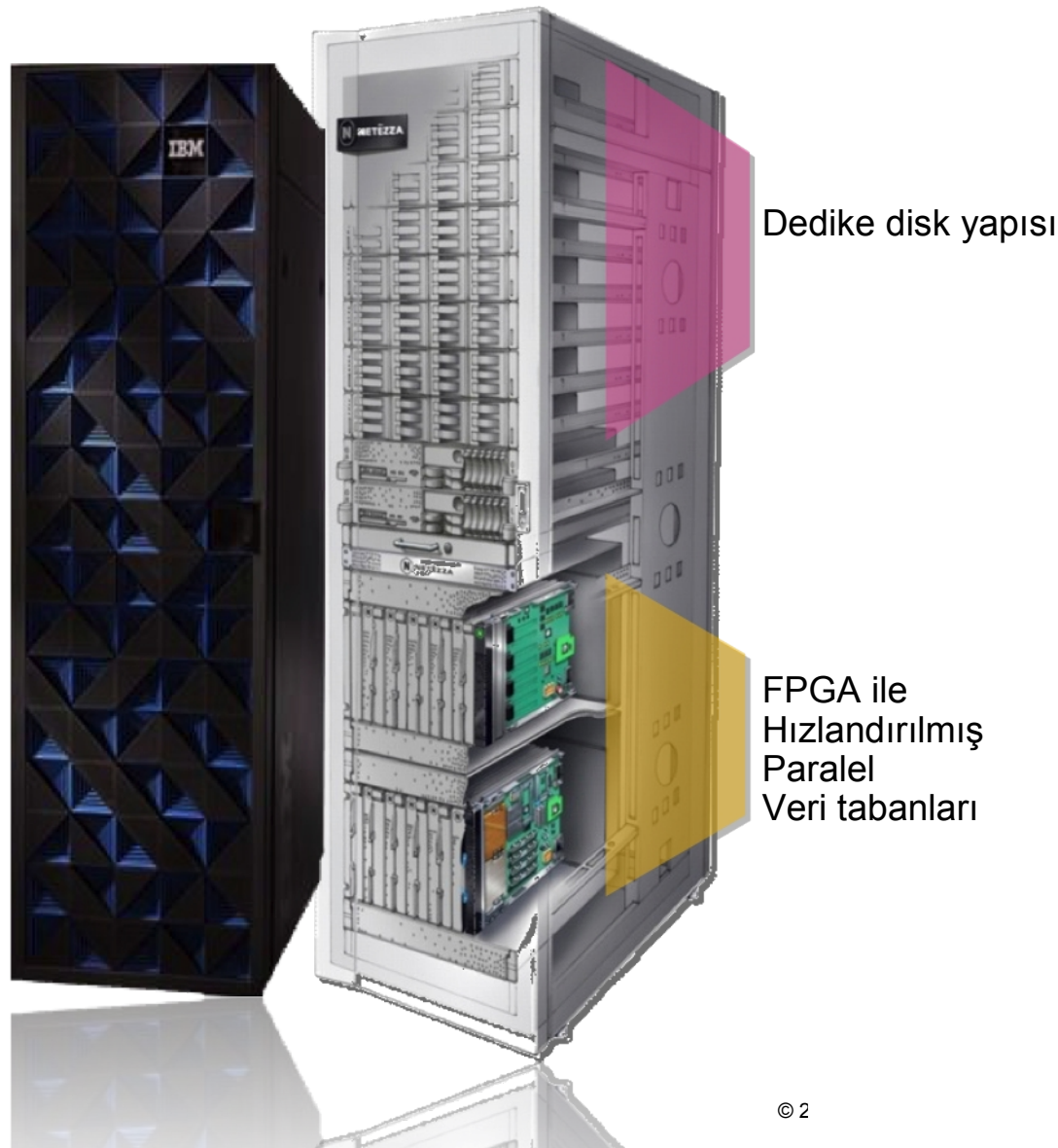
IBM Pure Data System for Analytics

Amaç Odaklı Yaklaşım

Hız: Geleneksel veri ambarlarına göre 10-100 kat hız

Basitlik: Minimum yönetim gereksinimi

Ölçeklenebilirlik: Veri hacmine göre ölçeklenebilirlik



Dedike disk yapısı

FPGA ile Hızlandırılmış Paralel Veri tabanları

IBM PureData System for Operational Analytics



İnteraktif ve analitik sorgular için karma sistem

- Kullanıma hazır optimize edilmiş
- Operasyonel veri analizi için geliştirilmiş
- Otomatik iş yükü dağılımı
- 10 kata kadar sıkıştırma oranı
- Kolay yönetim



Vestas : Büyük Veri ile rüzgar güllerinin konumlarının belirlenmesi

- 2.8 petabyte hava durumu verisi
- Hava durumu tahmin süresi:
Hafta → Saat
- Enerji aktarma optimizasyonu

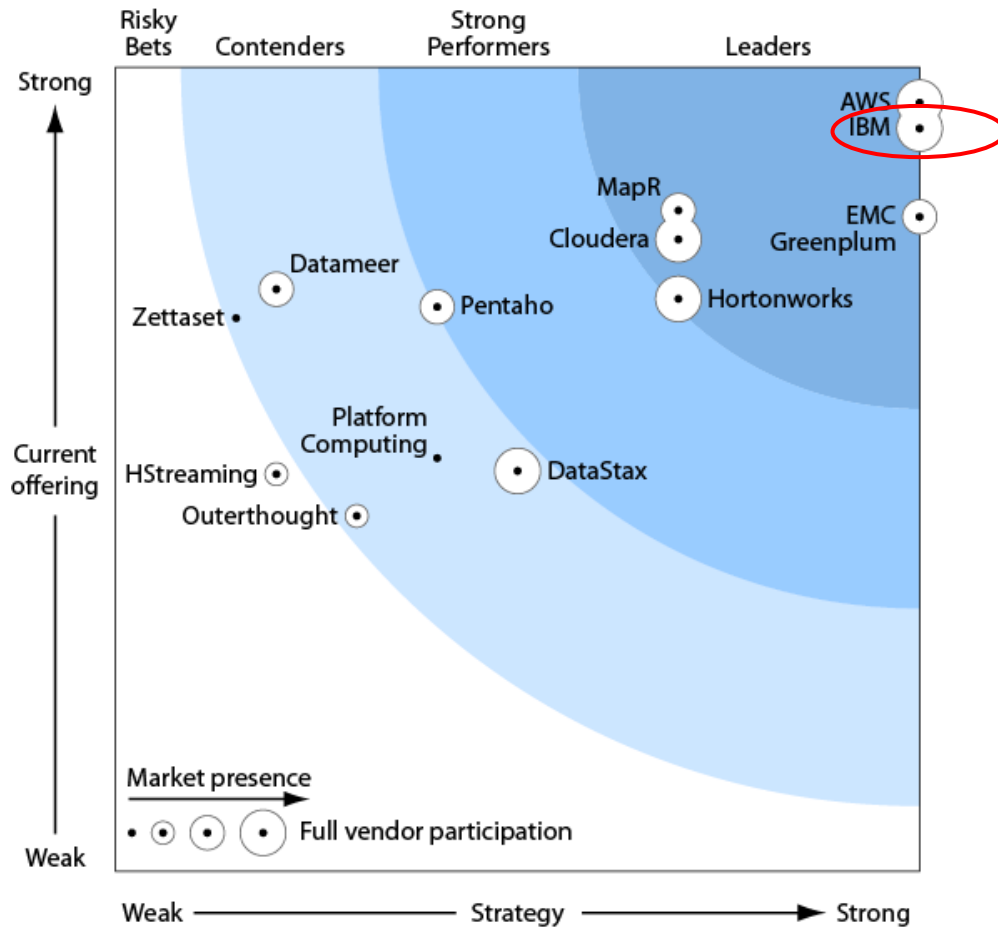


Telco (Asya): Faturalama maliyetleri düřtü ve müşteri memnuniyeti arttı.

- **Günlük 6 milyar CDR kaydı analizi**
- **Gerçek zamanlı analitik ve veri entegrasyonu**
 - **Veri işleme süresi:**
12 saat → 1 saniye
 - **Donanım maliyeti %83 azalma**
- **Network sorunlarına proaktif yaklaşımlar**

IBM 'in yeri

“IBM Hadoop platformu ve uygulamalarında en büyük portföye sahip”



February 2012 “The Forrester Wave™: Enterprise Hadoop Solutions, Q1 2012”

Teşekkürler

