



ceph



**Ceph Türkiye Meetup**

# **Ceph Tasarımında Dikkat Edilmesi Gereken Hususlar**

Ramazan ÖZTEMUR

22.11.2017

Ankara

# Gündem

- Gereksinimlerin Belirlenmesi
- Replika vs Erasure Coding
- Performans Maliyet Kapasite
- İş Yüküne Özel tasarım
- Donanım Seçimi
- Önerilen Donanım Listesi
- Örnek Tasarım
- Soru/Cevap

# Gereksinimlerin Belirlenmesi

- Ticari Limit ve Gereksinimler
  - Bütçe
  - Amaç (Günlük iş, HPC, Media Stream, Arşiv, özel amaç vb)
  - Tür (Block, Object, FS)

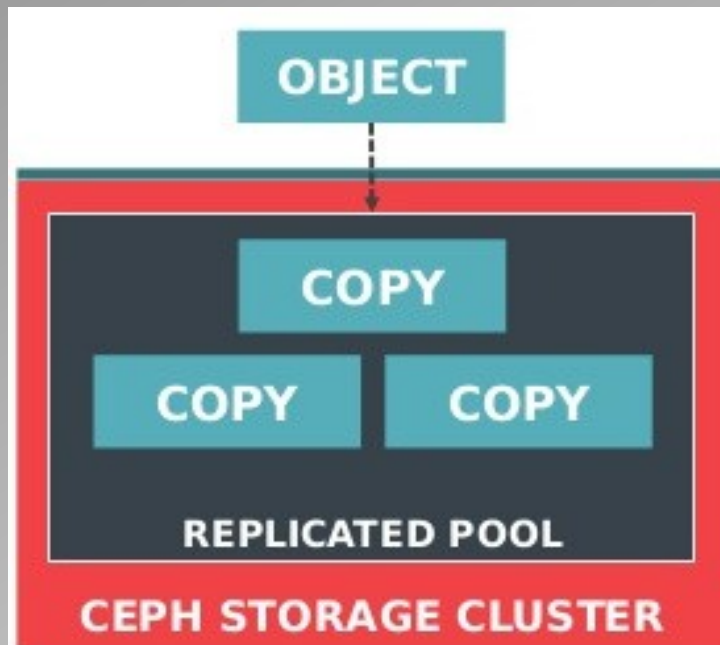
# Gereksinimlerin Belirlenmesi

- Teknik Gereksinimler
  - Net Kapasite?
  - Veri Büyüme Hızı ?
  - IOPS ?
  - Kümeden Beklenen Transfer Hızı, MB/s ?
  - Replikasyon ?
  - Hangi Uygulama ?
  - Okuma(%70) / yazma(%30) oranı
  - Yazılan Veri Deseni

# Gereksinimlerin Belirlenmesi

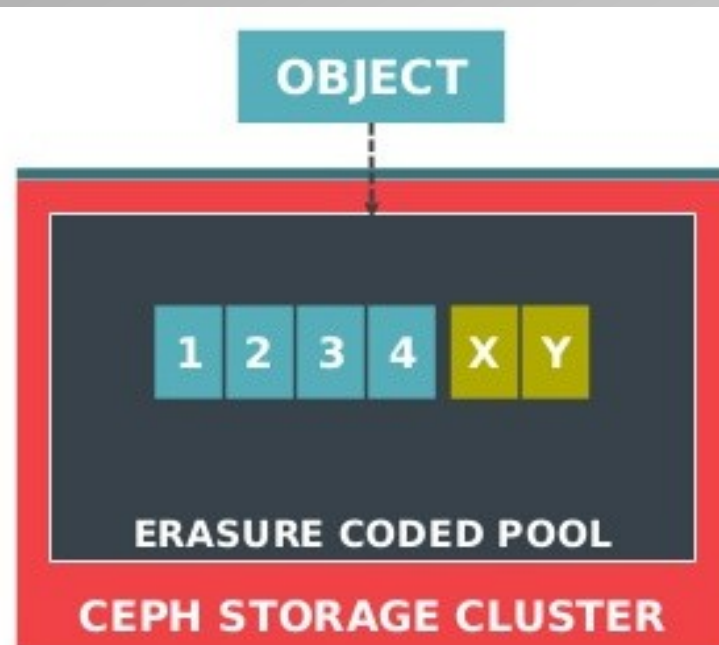
- Teknik Gereksinimler
  - Performans / Maliyet önceliği
  - Kaç Disk / Host Kaybı Tolere Edilmeli ?
  - Replika Farklı Rack Kabin/DC/Region de Olacak mı?
  - Küme, Ağ Trafiği? Kabin?

# Replika vs Erasure Coding



Full copies of stored objects

- Very high durability
- 3x (200% overhead)
- Quicker recovery



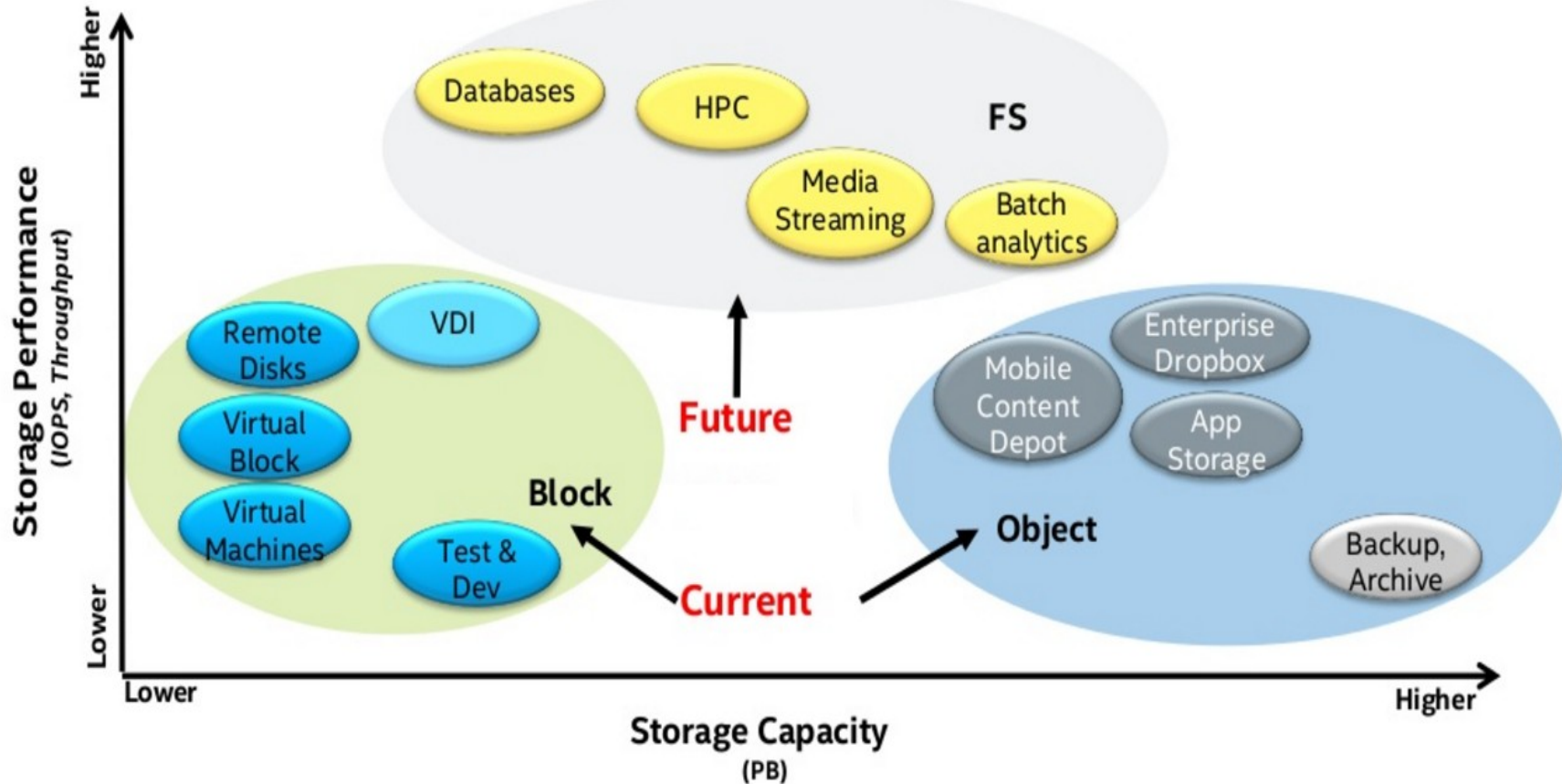
One copy plus parity

- Cost-effective durability
- 1.5x (50% overhead)
- Expensive recovery

# Performans Maliyet Kapasite

- Performans / Maliyet / Kapasite
  - Replika – Erasure Coding
  - Küme Disk Tipleri
    - Full Flash
    - Cache + Mekanik Disk
    - Çeşitli Disk Havuzları
  - Ağ
  - Küme/Sunucu/Disk Sayısı

# İş Yüküne Özel tasarım





# Donanım Seçimi

- Küme Sunucuları
- Ağ Katmanı
- Disk
- Journal Disk

# Önerilen Donanım Listesi

- Küme Sunucuları
  - Diskler
    - 10 GbE kart / 8 – 10 SAS HDD
    - 10 GbE kart / ~ 12 SATA HDD
  - Journal
    - 4 - 6 Disk / 1 SSD journal

# Önerilen Donanım Listesi

- Küme Sunucuları
  - RAM
    - 1TB / 1 GB RAM
  - CPU
    - OSD / 0.5 Core veya 1 GHz CPU
    - SSD / 2 Core
  - Ağ
    - Genel Ağ – 1 / 10 Gbps
    - Küme Ağ – 1 / 10 / 40 Gbps

# Önerilen Donanım Listesi

- Küme Monitör Sunucusu
  - En az 3 (3, 5)
  - Çalışan Ceph Servis / 1 GB RAM
  - Çalışan Ceph Servis / 10 GB Disk
  - Genel Ağ - 1 / 10 GbE

# Önerilen Donanım Listesi

- Küme Ceph MetaData Sunucu (MDS)
  - Yedekli önerilir, Monitor Sunucular
  - Çalışan Servis / 1 GB RAM
  - Çalışan Servis / 10 GB Disk
  - Genel Ağ - 1 / 10 GbE

# Önerilen Donanım Listesi

- Küme Rados GW
  - Object Storage API sunucu
  - Genelde Mon Sunucuda

# Önerilen Donanım Listesi

- Tahmin Edilemeyen Yüklerde Genel Kabul
  - %70 Read / %30 Write IOPS oranı
  - ~ 4-8 KB Random Read Paterni
- Küme IOPS verimliliği
  - 4-8 K Random Read 0.88
  - 4-8 K Random Write 0.64

# Örnek Tasarım

- Gereksinimler
  - 1 PB Hacim
  - 500 VM
  - 100 IOPS / VM
  - Total 50.000 IOPS



# Örnek Tasarım

- Performans Odaklı
  - 2620v3 CPU (6 çekirdek)
  - 64 GB RAM
  - 2 x 500 GB SATA HDD (OS)
  - 20 x 1.8 TB SAS 10K RPM HDD
  - 4 x 200 GB Journal
  - 1 x 1 Gbps + 4 x 10 GbE (2 bond)
  - $1000 / (20 * 1.8 * 0.85) * 3 = 98$  sunucu

# Örnek Tasarım

- Performans Odaklı Tasarım Başarımı
  - 10K SAS disk : 250 IOPs
  - Cluster okuma oranı:
    - $250 \times 20 \times 98 \times .88 = 431200$  IOPs
    - Yaklaşık 800 IOPs (VM başına)
  - Cluster yazma oranı:
    - $250 \times 20 \times 98 \times .88 = 313600$  IOPs
    - Yaklaşık 600 IOPs (VM başına)

# Örnek Tasarım

- Düşük Maliyet
  - CPU: 2 x 2630v3 (16 cores)
  - RAM: 192 GB
  - HDD (OS): 2 x SATA 500 Gb
  - HDD (OSD): 28 x SATA 4 Tb (7.2k,3.5")
  - SSD (write journal): 6 x SSD 200 GB
  - NIC: 1 x 1 GbE, 4 x 10 GbE (2 bond)
  - $1000 / (28 * 4 * 0.85) * 3 = 32$  sunucu



# Örnek Tasarım

- Maliyet Odaklı Tasarım Başarımı
  - SATA Disk : 150 IOPs
  - Cluster okuma rating:
    - $150 \times 32 \times 28 \times 0.88 = 118,272$  IOPs
    - Yaklaşık 200 IOPs (VM başına)
  - Cluster yazma rating:
    - $150 \times 32 \times 28 \times 0.64 = 86,016$  IOPs
    - Yaklaşık 150 IOPs (VM Başına)

# Yaklaşan Etkinlikler



- Ceph Türkiye 4. Meetup  
29.11.2017 Saat:19:00  
İstanbul YTE Davutpaşa  
Kampüsü Teknoparkı, A1 Blok
- Ceph Türkiye 5. Meetup  
07.12.2017 Saat:19:00  
Ankara Ataköşk Hotel
- Ceph Türkiye 6. Meetup  
14.12.2017 Saat:19:00  
İstanbul İTÜ Ayazağa Kampüsü  
SC3 Ulusal Yüksek Başarımlı Hesaplama M.

# Soru / Cevap

- Teşekkürler
- Meetup davet / sunum / salon
- Sorular

