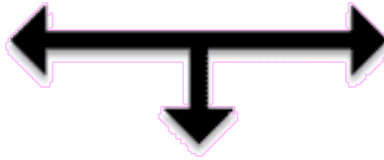
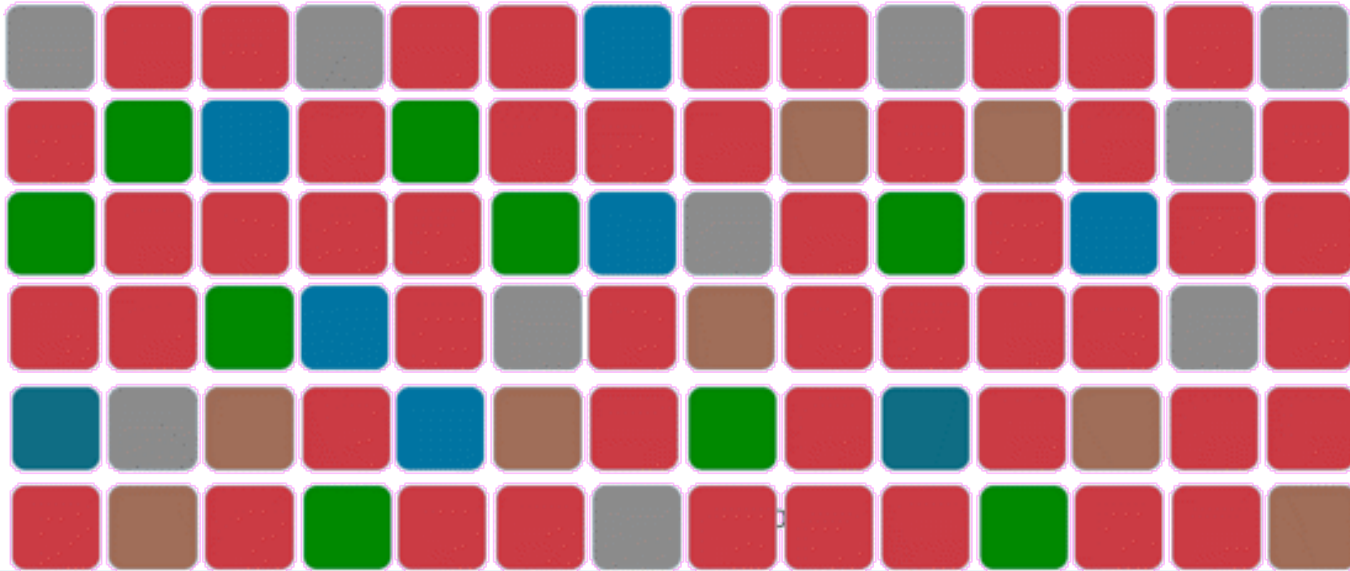


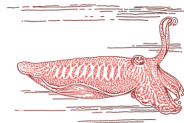
CRUSH



ZEBRABULUT



Özhan KARAMAN
ozhan@zebrabulut.com
RHCE, CLP, CLA, CNI



CUTTLEFISH
CUTTLEFISH
CUTTLEFISH



HAMMER

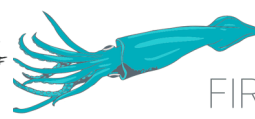


JEWEL

Linked in 'den beni takip
edebilirsiniz



ARGONAUT

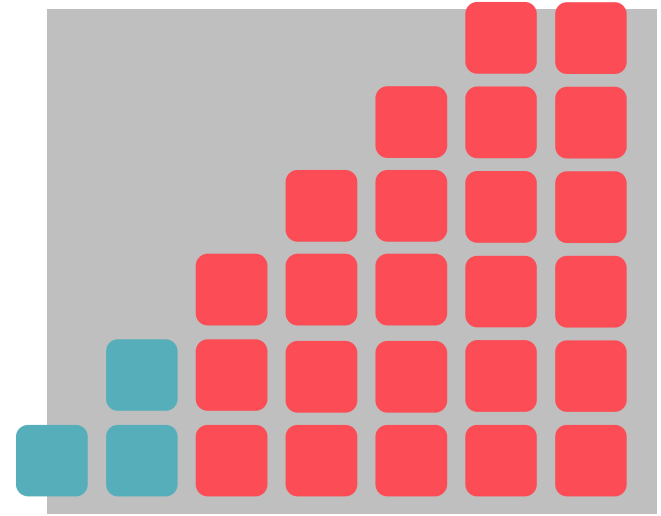


FIREFLY



Genel Durum

2020 yılına kadar **39 ZB** saklama alanının kullanılmasını bekleniyor. 2016 yılı itibari ile **1.5 ZB** alan kullanılmaktadır.



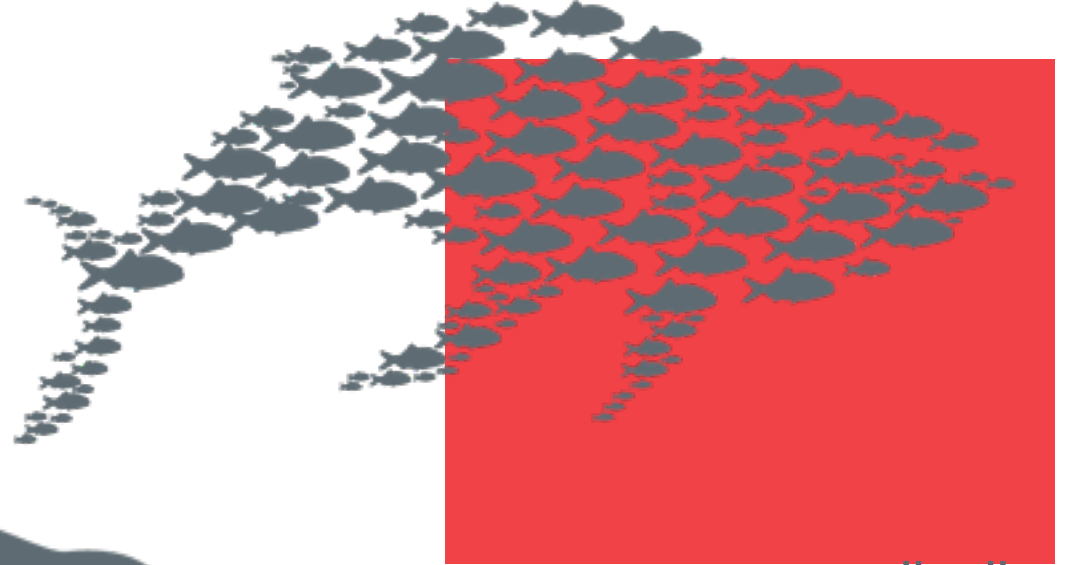
Mevcut Problem



- Mevcut Sistemler ölçeklenemiyor, büyüyemiyor.
- Maliyet ve karmaşıklık artıyor
- Önceden yeni platformlara teknolojilere yatırım yapmak gerekiyor.

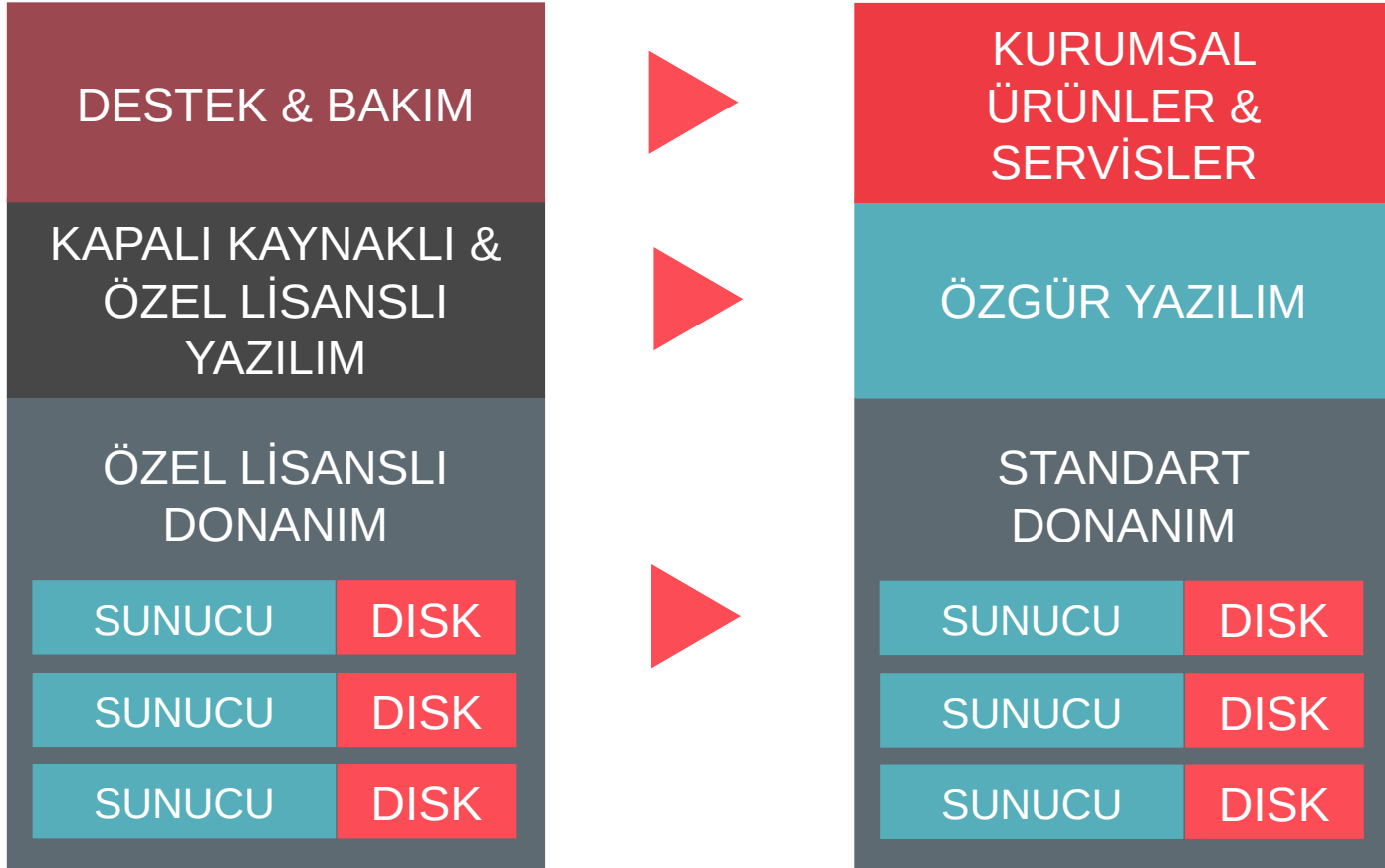
Çözüm

GEÇMİŞ: DİKEY BÜYÜME



GELECEK: YATAY BÜYÜME

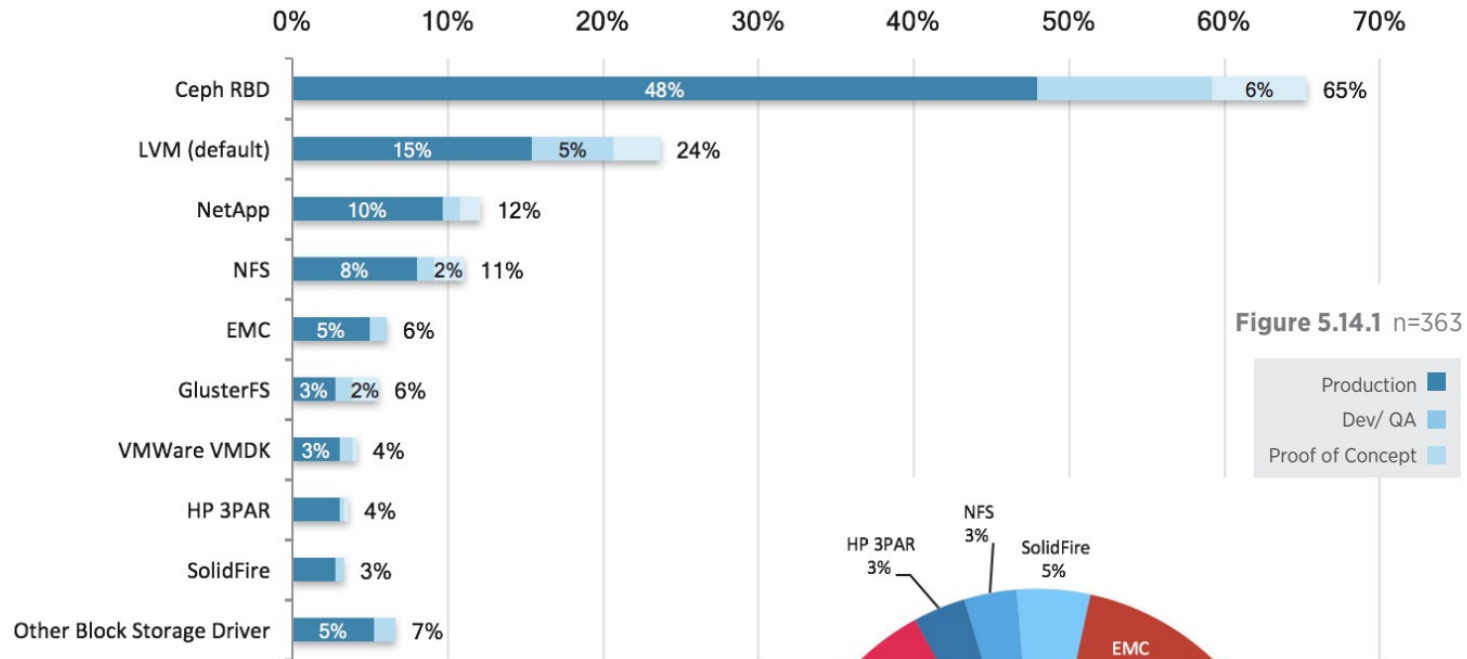
Saklama Alanında Devrim



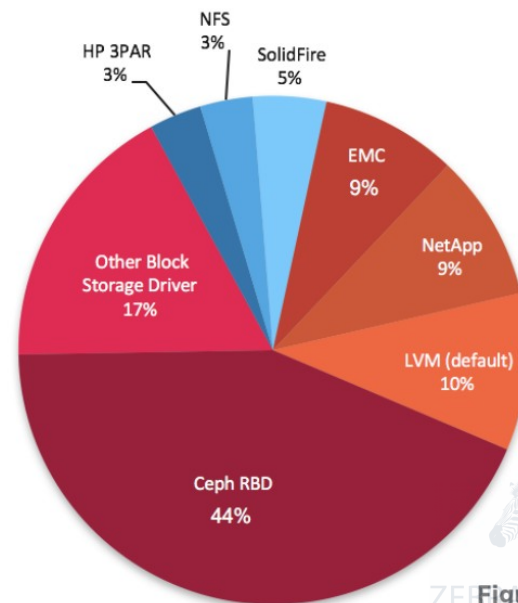
Genel Olarak Ceph

- Dağıtık Dosya Sistemi
- Enine Genişleyebilir, Ölçeklenebilir
- Tek Nokta Hatası İçermez, Tüm Yapı Yüksek Erişimlidir
- Kendini İyileştirme ve Yönetme
- Standart x86 Sunucu Donanımı Üstünde Çalışır
- Obje, Blok ve Posix Dosya Sistemi Desteği Sunar
- GPLv2 Kamu Lisansı Kullanmaktadır
- Topluluk Tarafından Yönetilen ve Kurumsal Desteği Olan
- OpenStack için en çok tercih edilen saklama alanı teknolojisidir

OpenStack Kullanıcı Anketi 04/2017

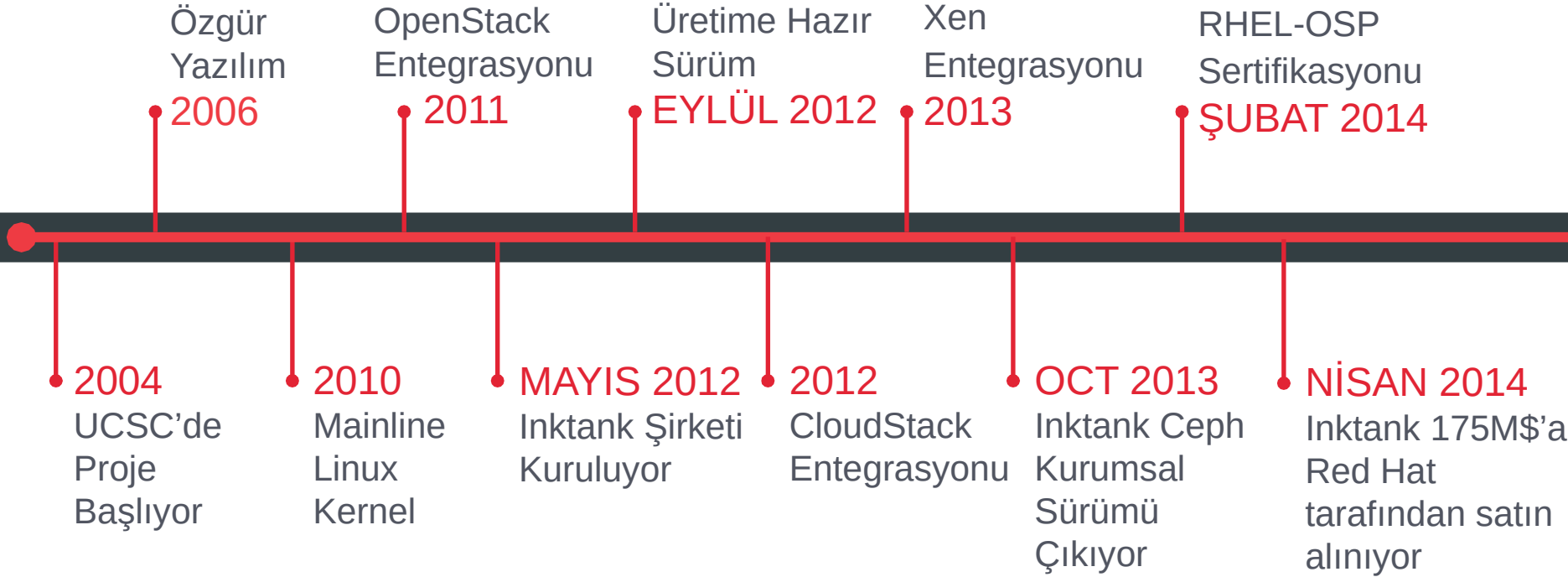


Among the largest clouds with 1,000 or more cores, Ceph RBD is still dominant, but not used by the majority, while other block storage drivers were also popular.



Ceph Kronolojisi

Ceph'in 10 senelik ilerleyiři



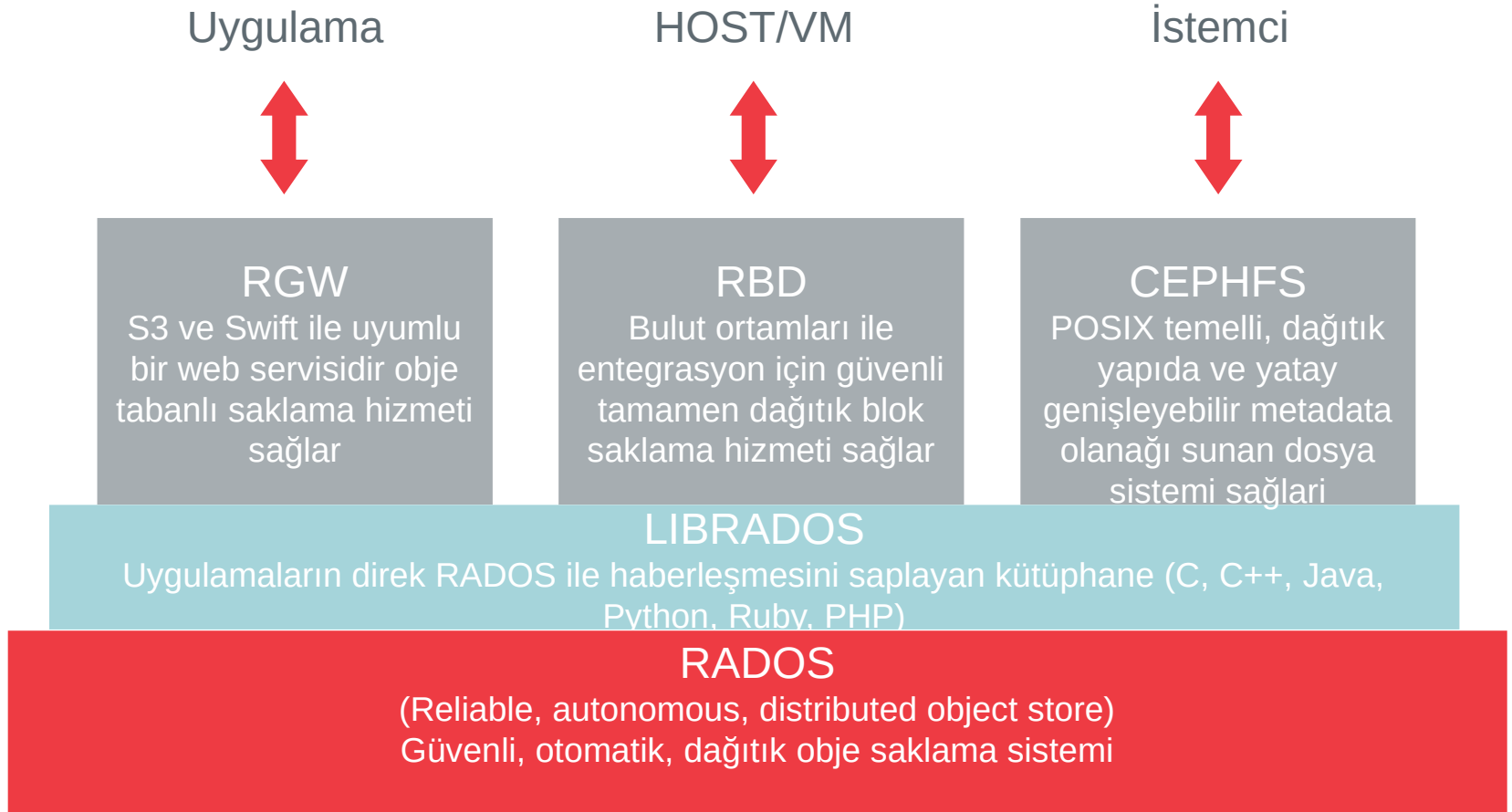
Türkiye'de Hangi Şirketler Kullanıyor ?

- Cloudeos www.cloudeos.com
- UyumSoft
- Tübitak Ulakbim
- Türksat
- Bilkent Üniversitesi Bilgisayar Merkezi
- Gözen Holding – FreeBird Airlines

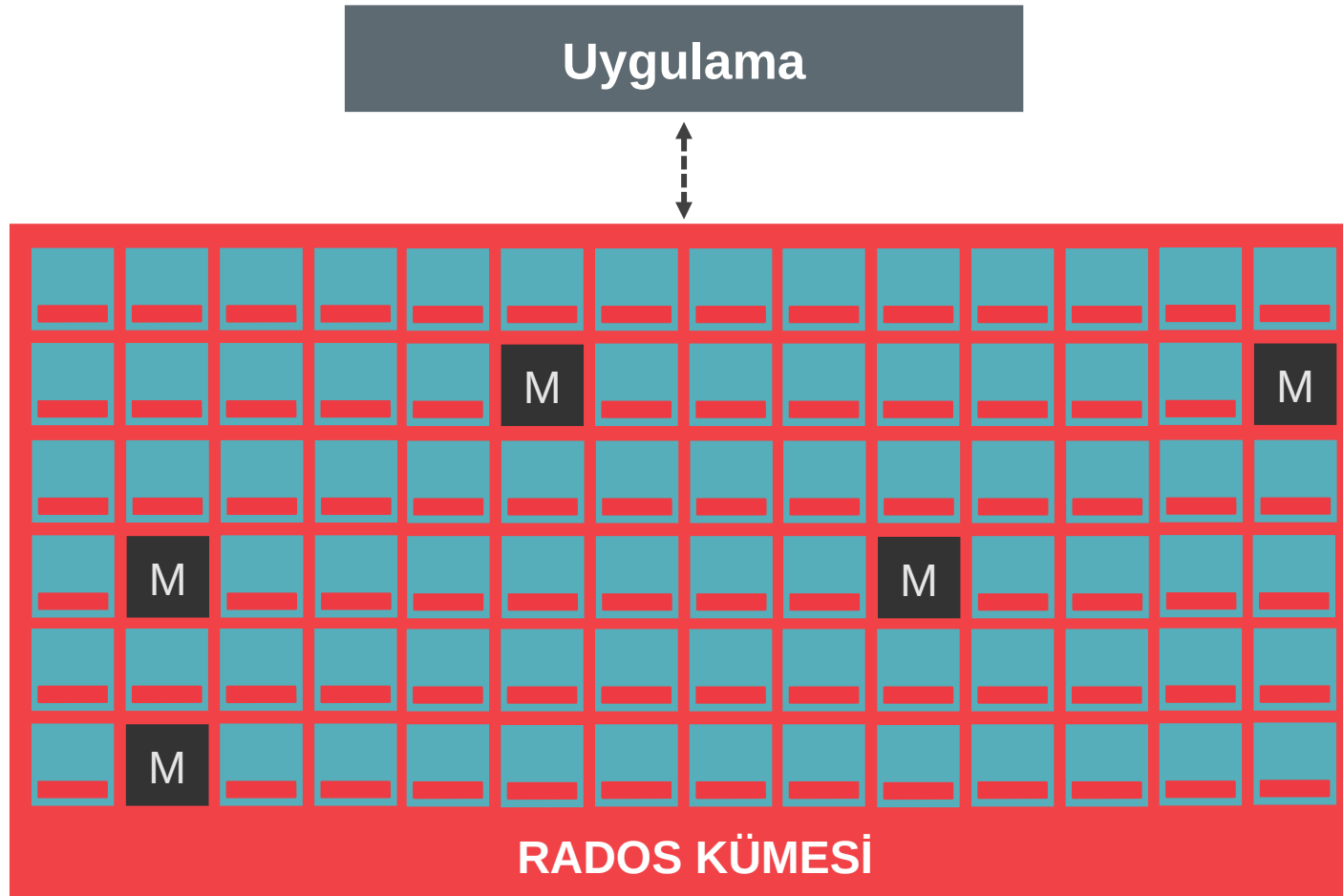


MIMARI

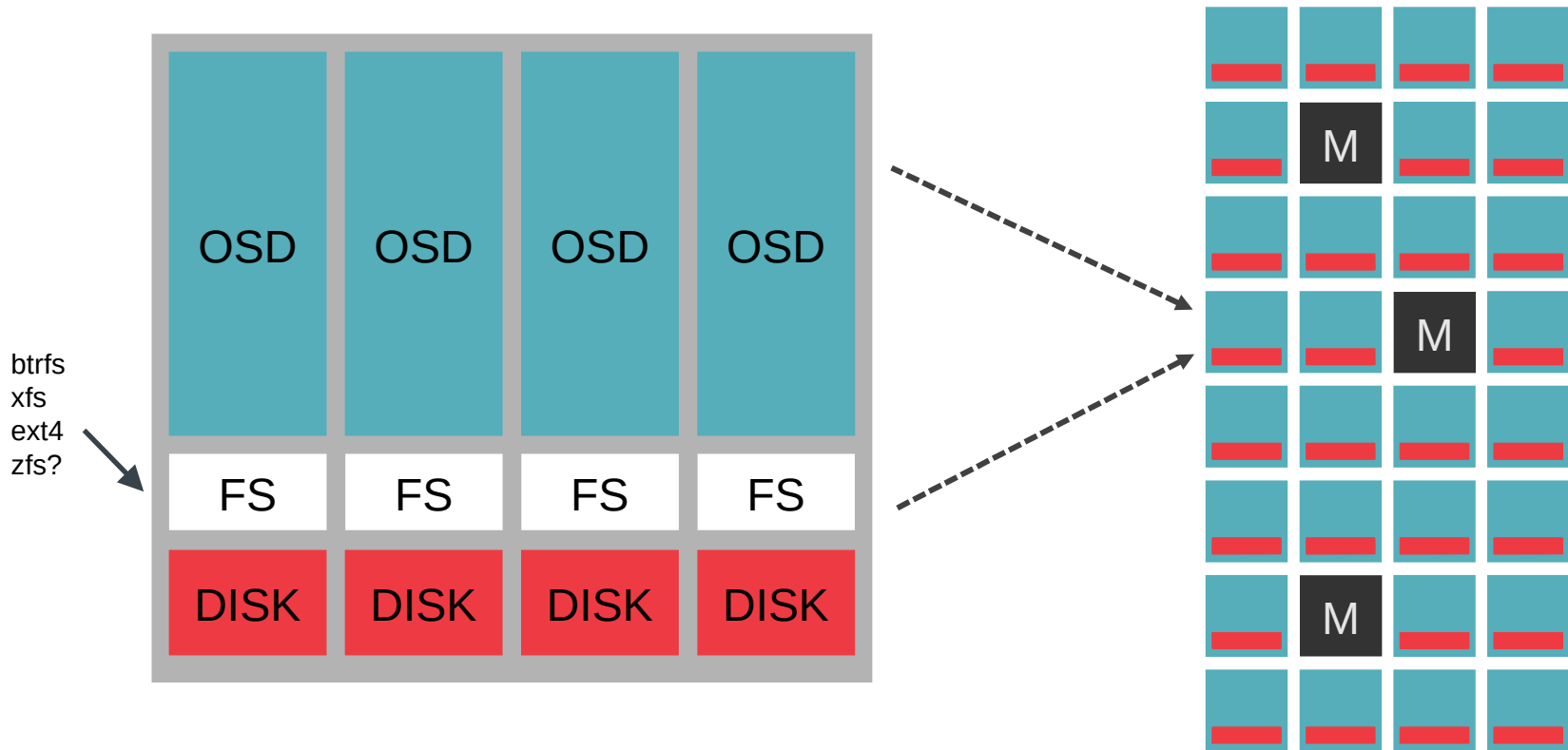
Genel Mimari



RADOS Kumesi



Objekt Saklama Uygulamaları - OSDs



RADOS Bileşenleri



OSD Daemons:

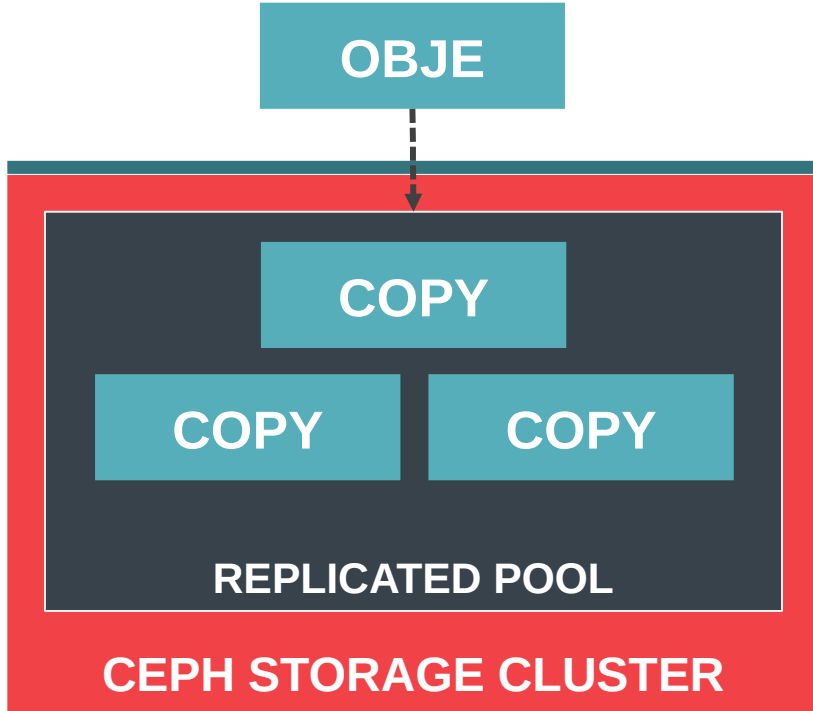
- Her bir kümeden ihtiyaca göre 10k-100k OSD olabilir
- Her bir fiziksel disk için 1 adet çalışır
- Objeleri saklar ve istemcilere gönderir
- Replikasyon ve eşleme için akıllı yöntemler kullanır



Monitör Daemons:

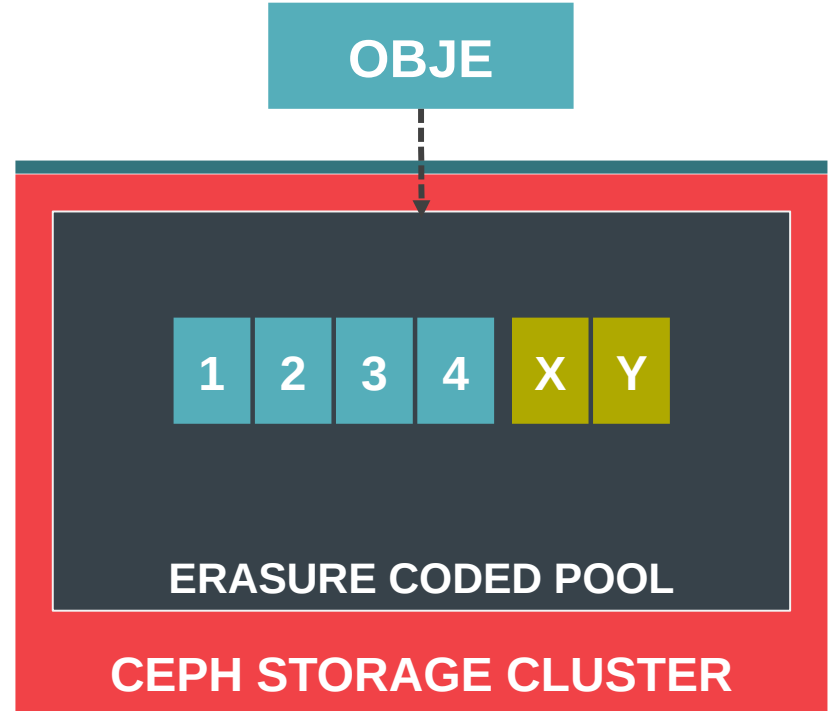
- Küme'nin genel yöneticisidir. Tekil ve dağıtık tüm kararları o verir ve yönetir.
- Tek sayılı ihtiyaca göre 3 veya 5 adettir
- İstemcilere data göndermez sadece sistemin sağlıklı çalışması için gereken haritaları tutar

Veri Yerleşim Teknikleri



Saklanacak dosyanın n kopyası

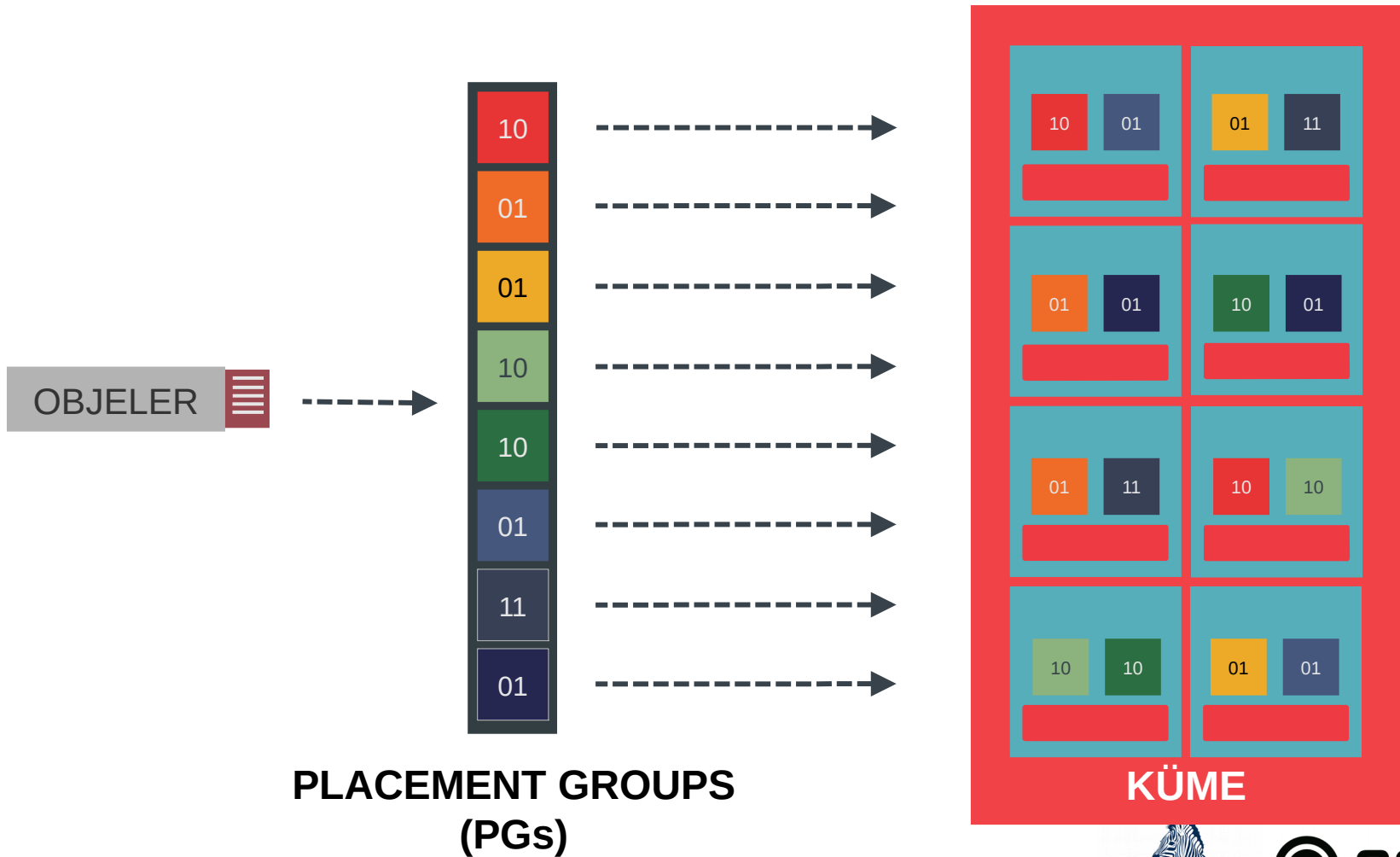
- Çok Yüksek Dayanıklılık
- Hızlı Kurtarma



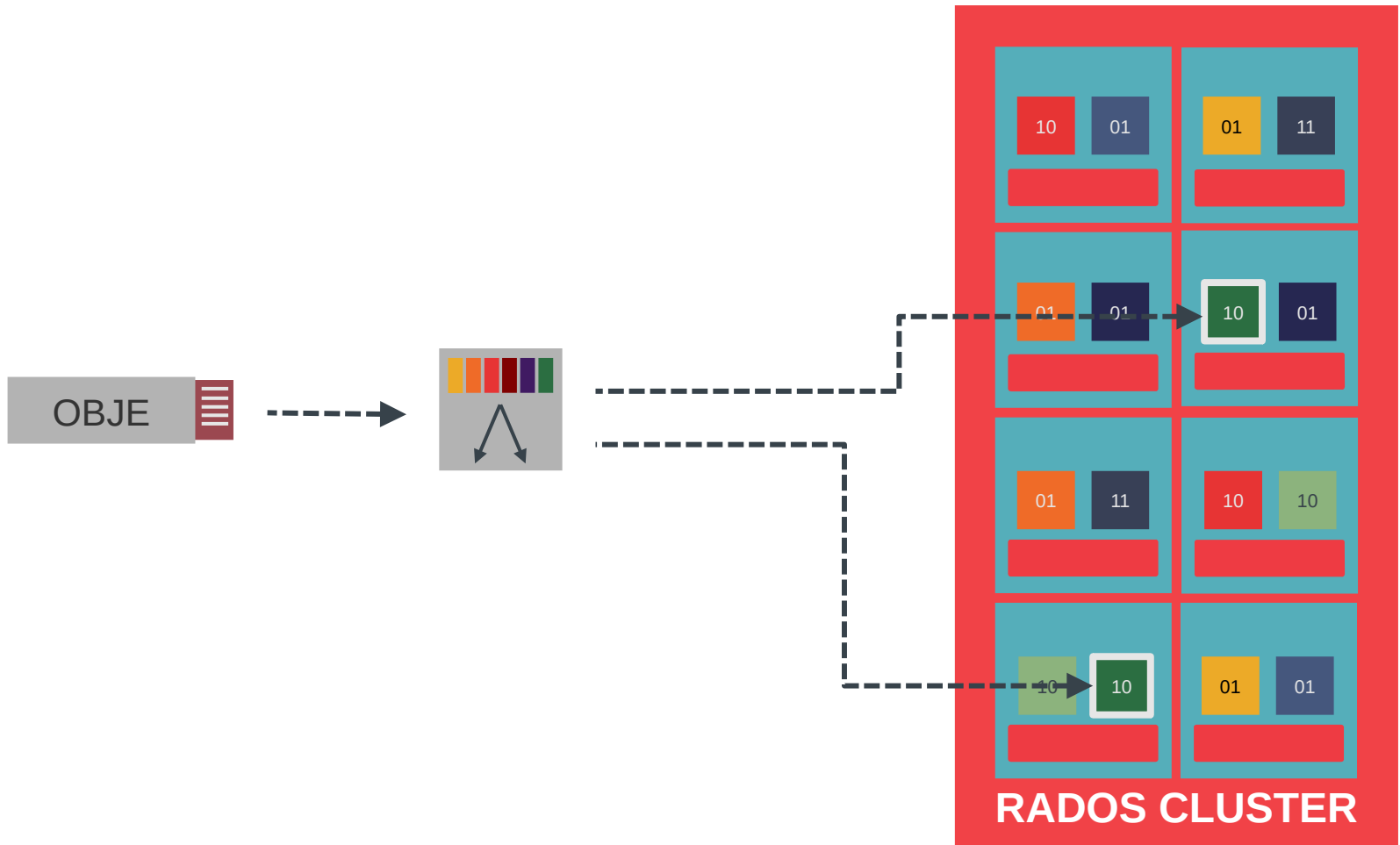
Tek kopya ve partiy dosyası

- Maliyet Etkin Dayanıklılık
- Pahalı, Zorlu Kurtarma

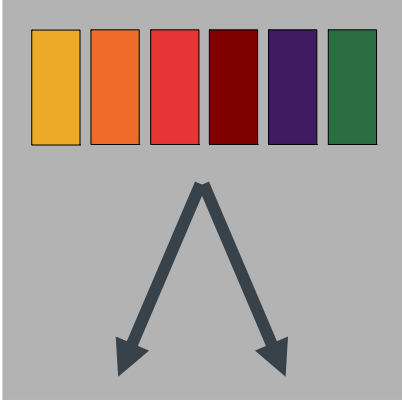
Genel CRUSH Yapısı



Hesaplama Çok Kolaydır



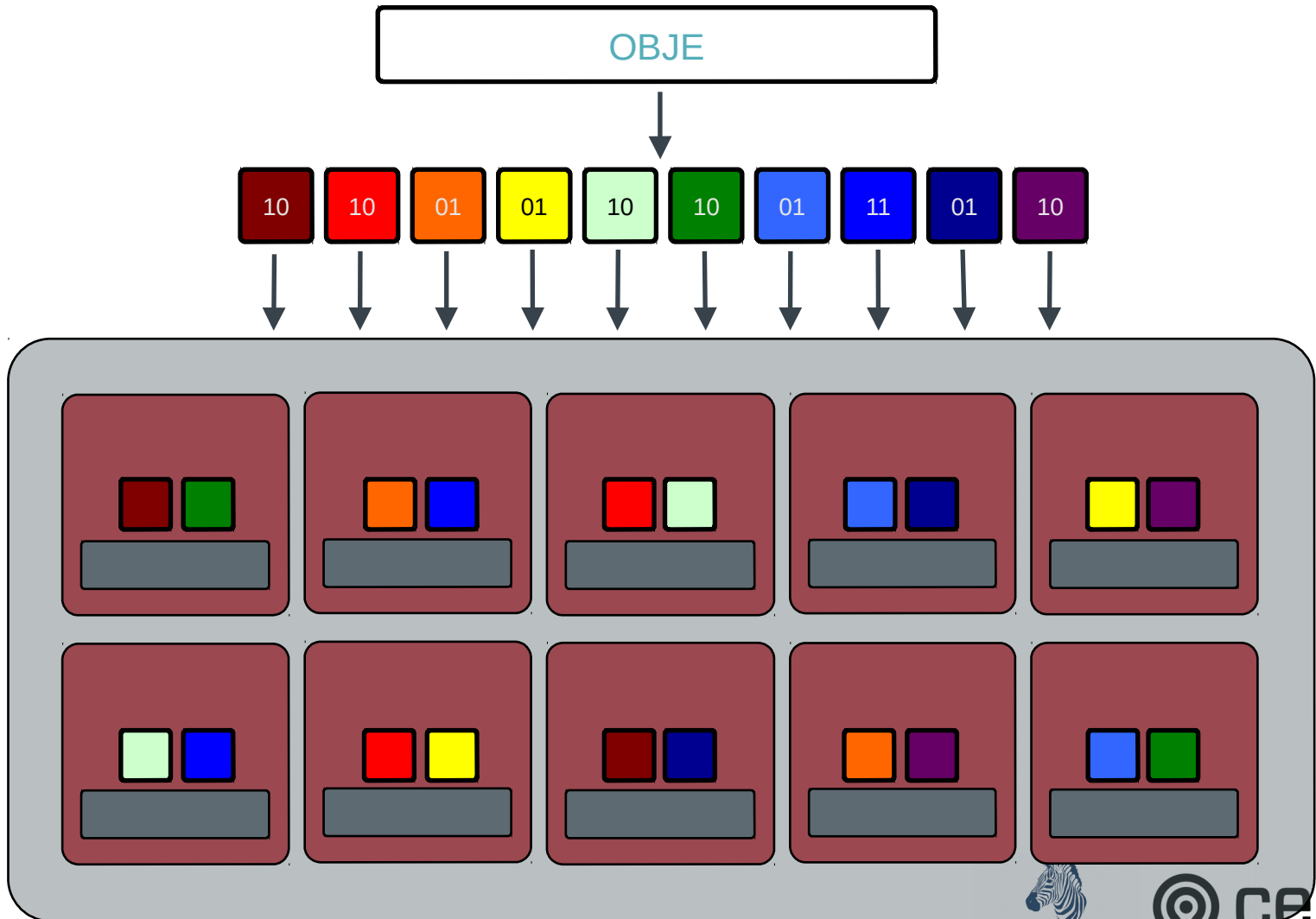
CRUSH: Dinamik Obje Konumu Belirleme



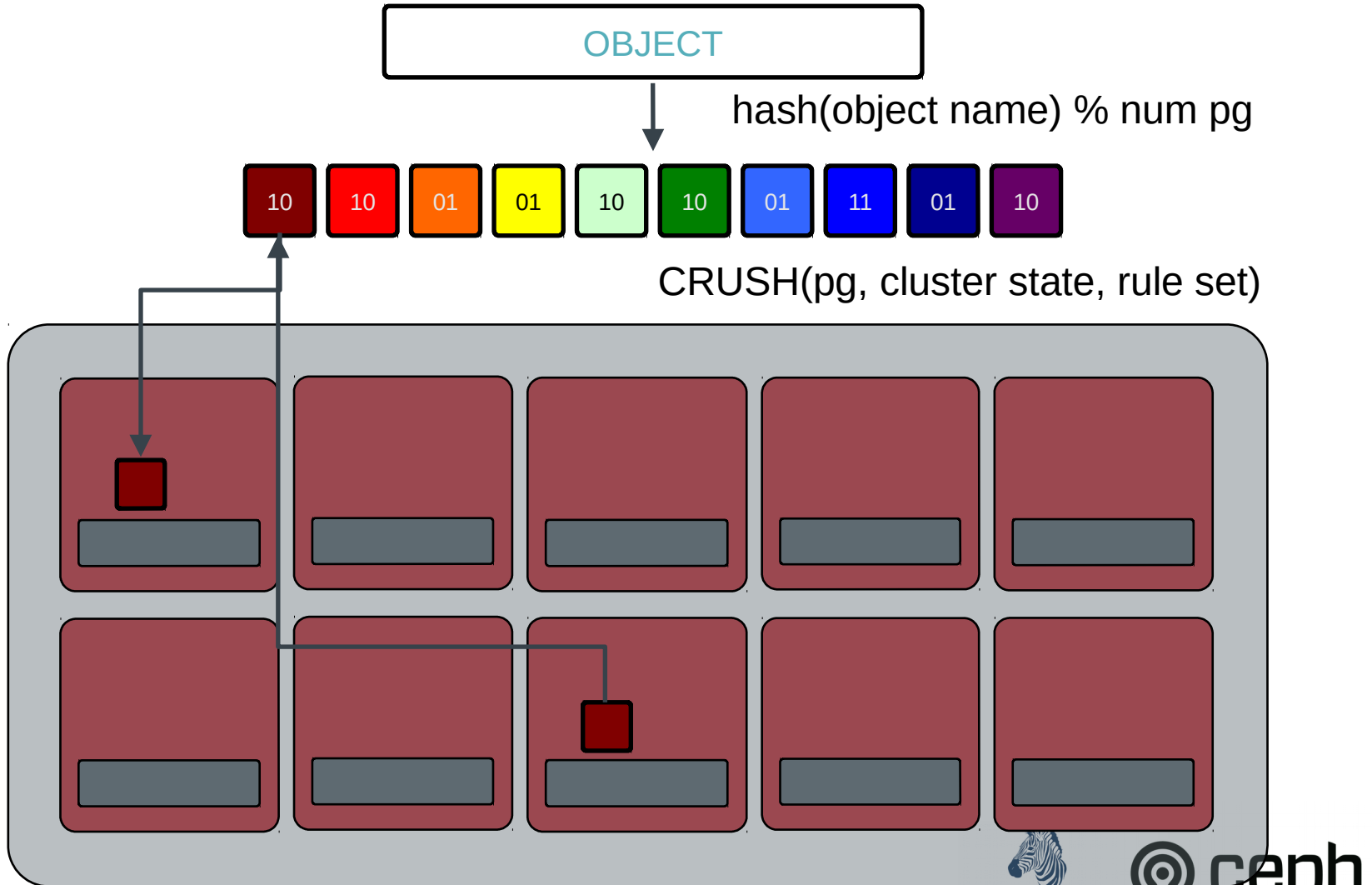
CRUSH:

- Kurallı Rastgele data yerleştirme algoritması
 - Hızlı hesaplama
 - Tekrarlanan, deterministik
- İstatistiksel olarak düzgün(uniform) dağılım
- Kararlı haritalama
 - Değişimlerde minimum veri göçü
- Kural tabanlı konfigürasyon
 - Mevcut altyapıdan haberdar
 - Ayarlanabilir Replikasyon
 - Ağırlık verme

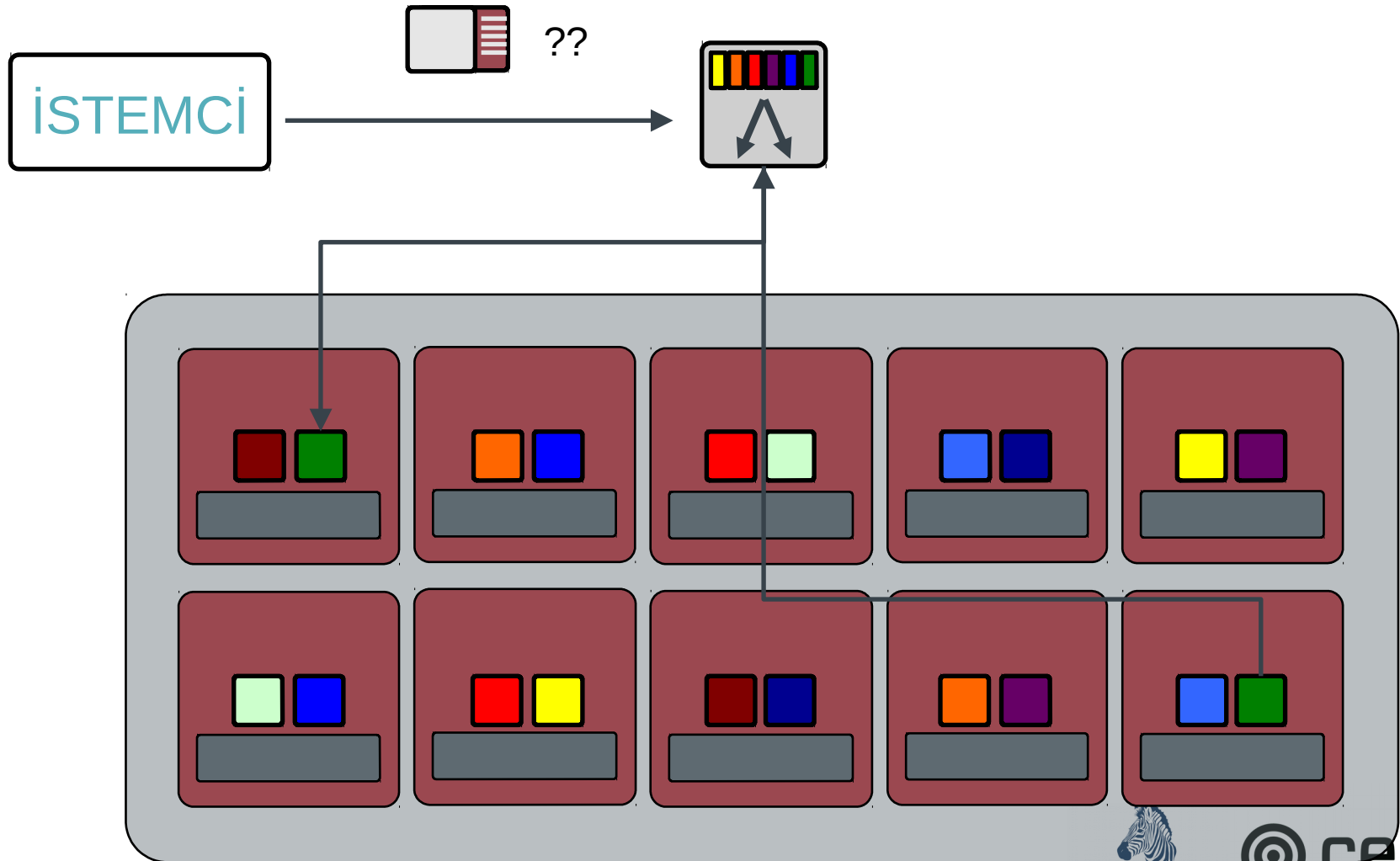
CRUSH



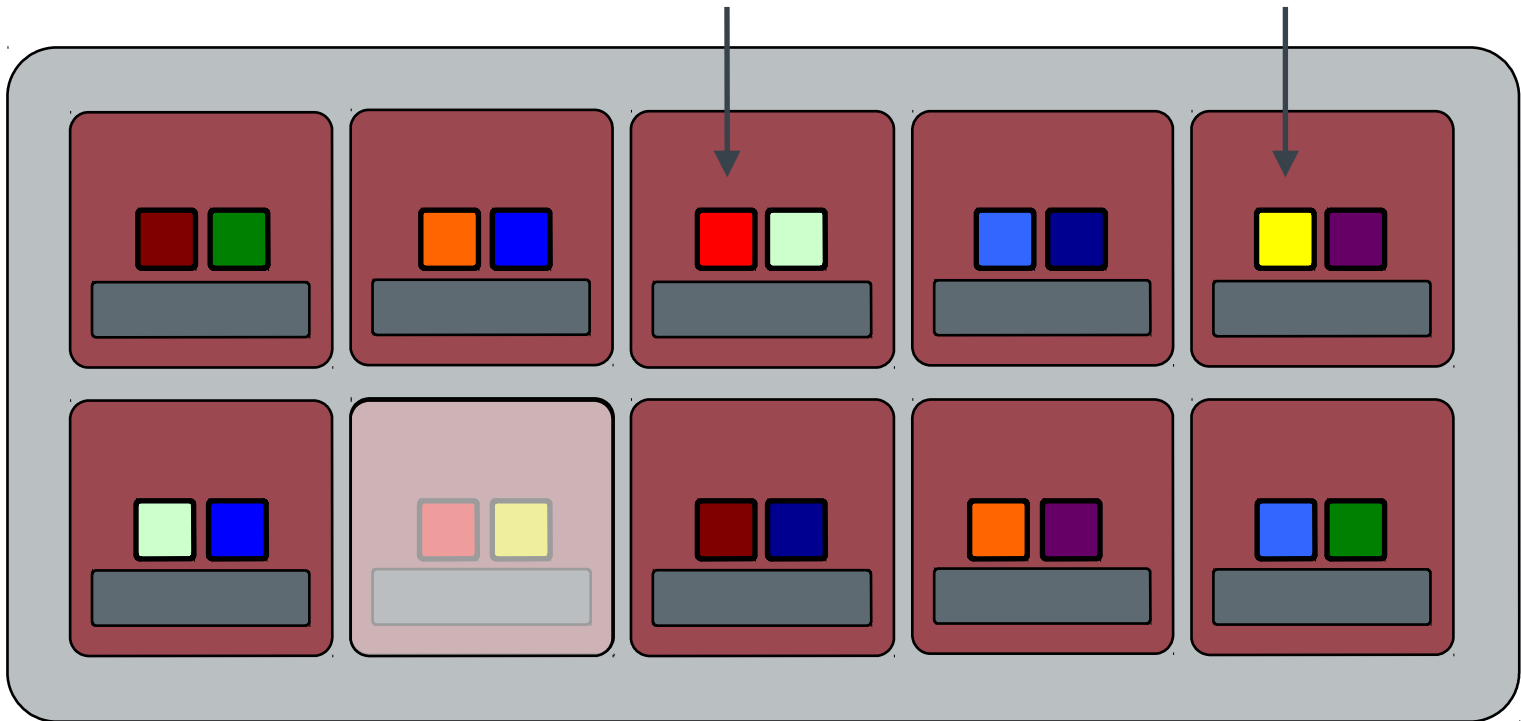
CRUSH: Verinin Yerleşimi



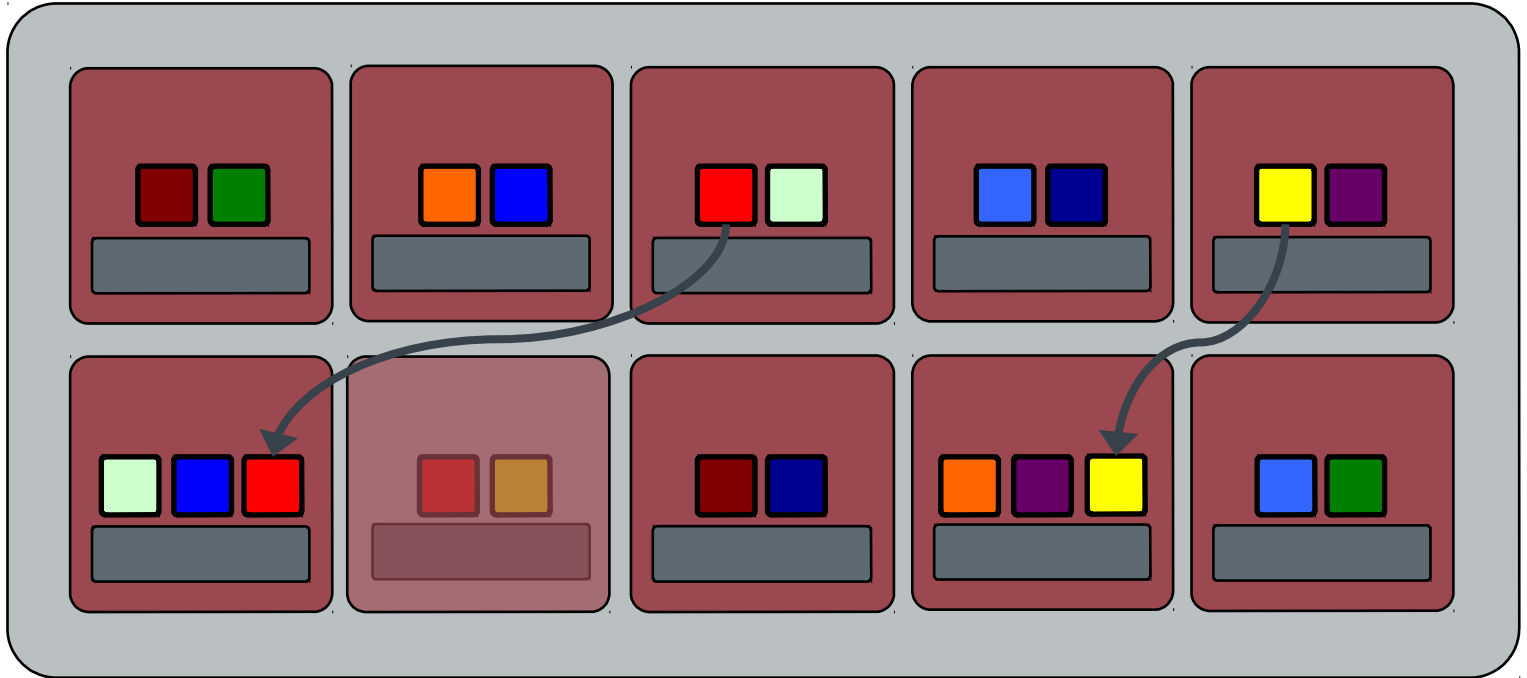
CRUSH: İstemci Erişimi



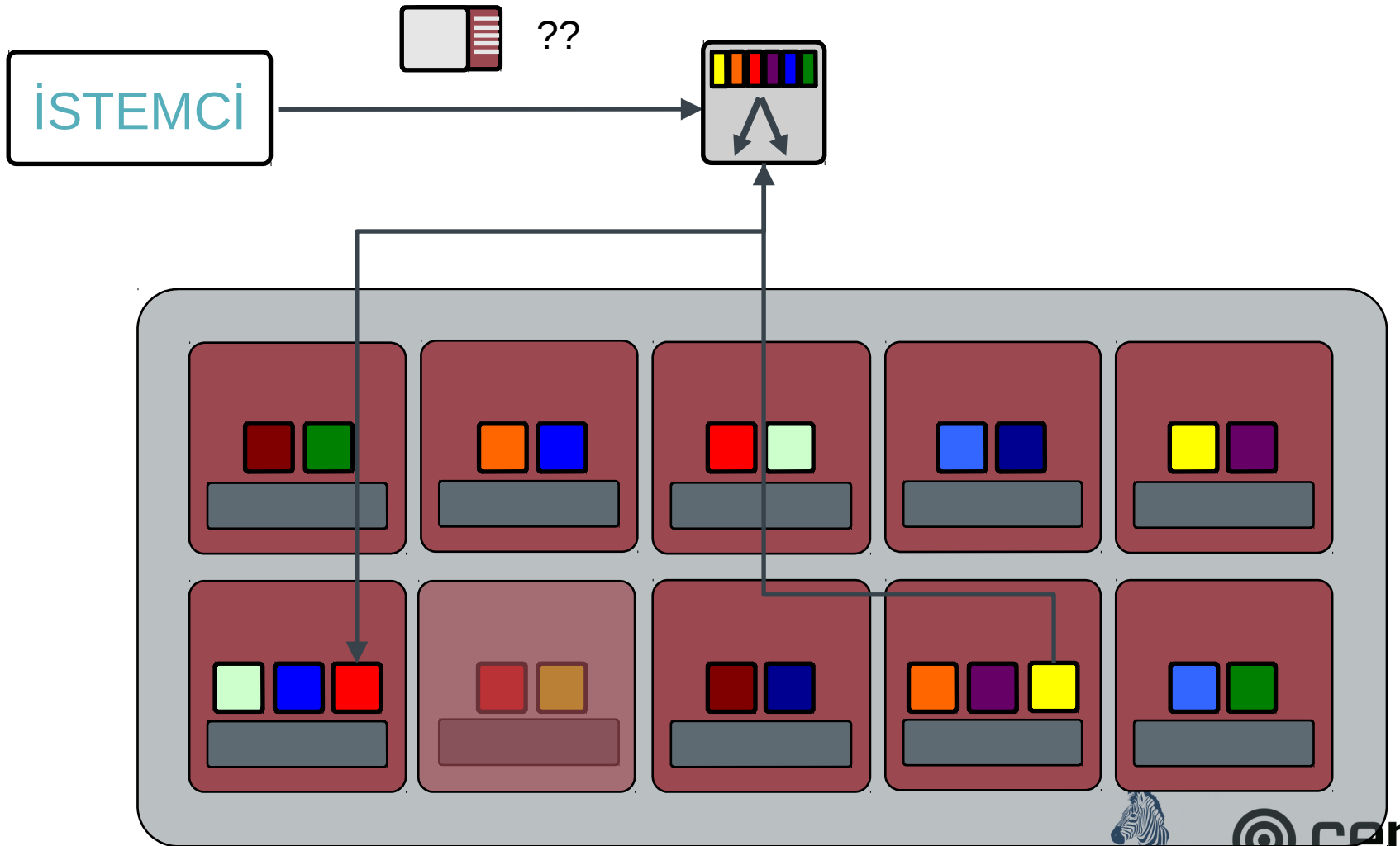
CRUSH: OSD'nin Gitmesi Durumu



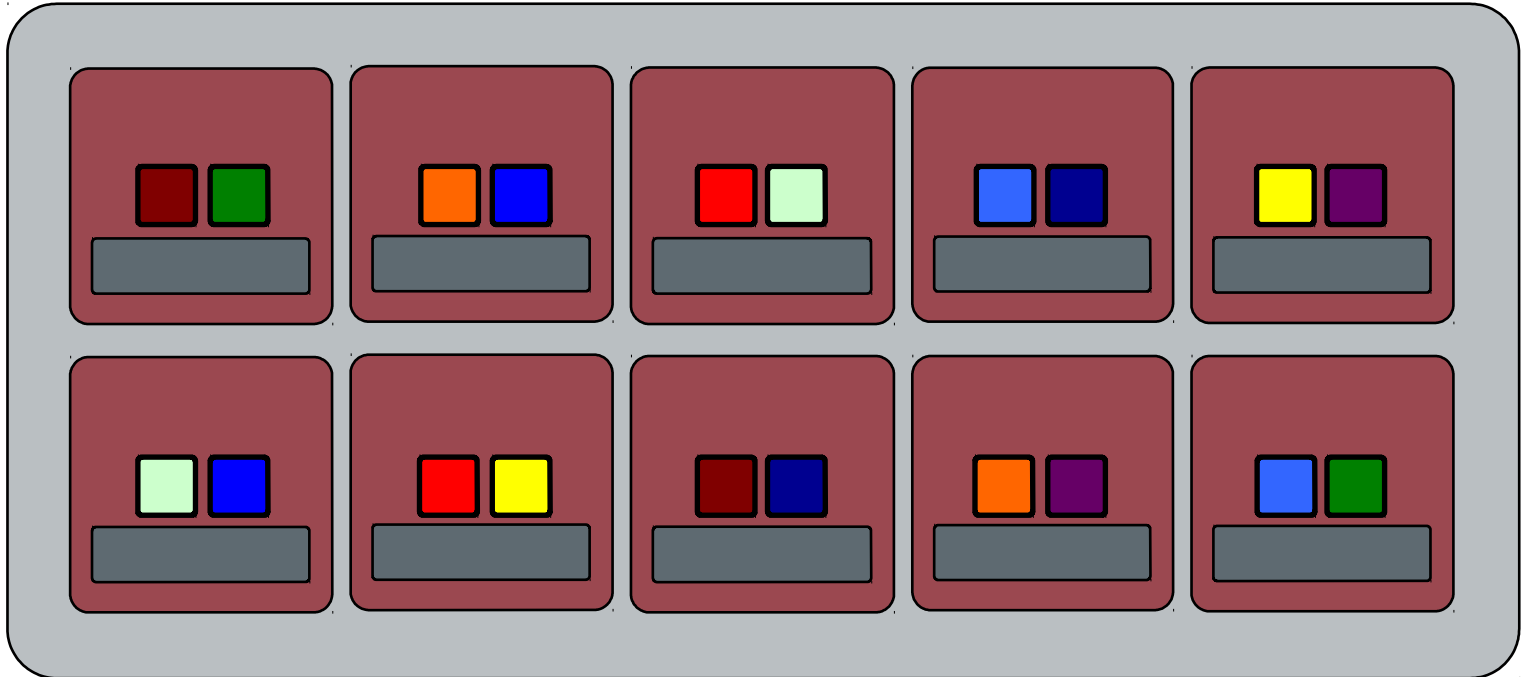
CRUSH: Arızalanan OSD'deki verilerin kurtarılması



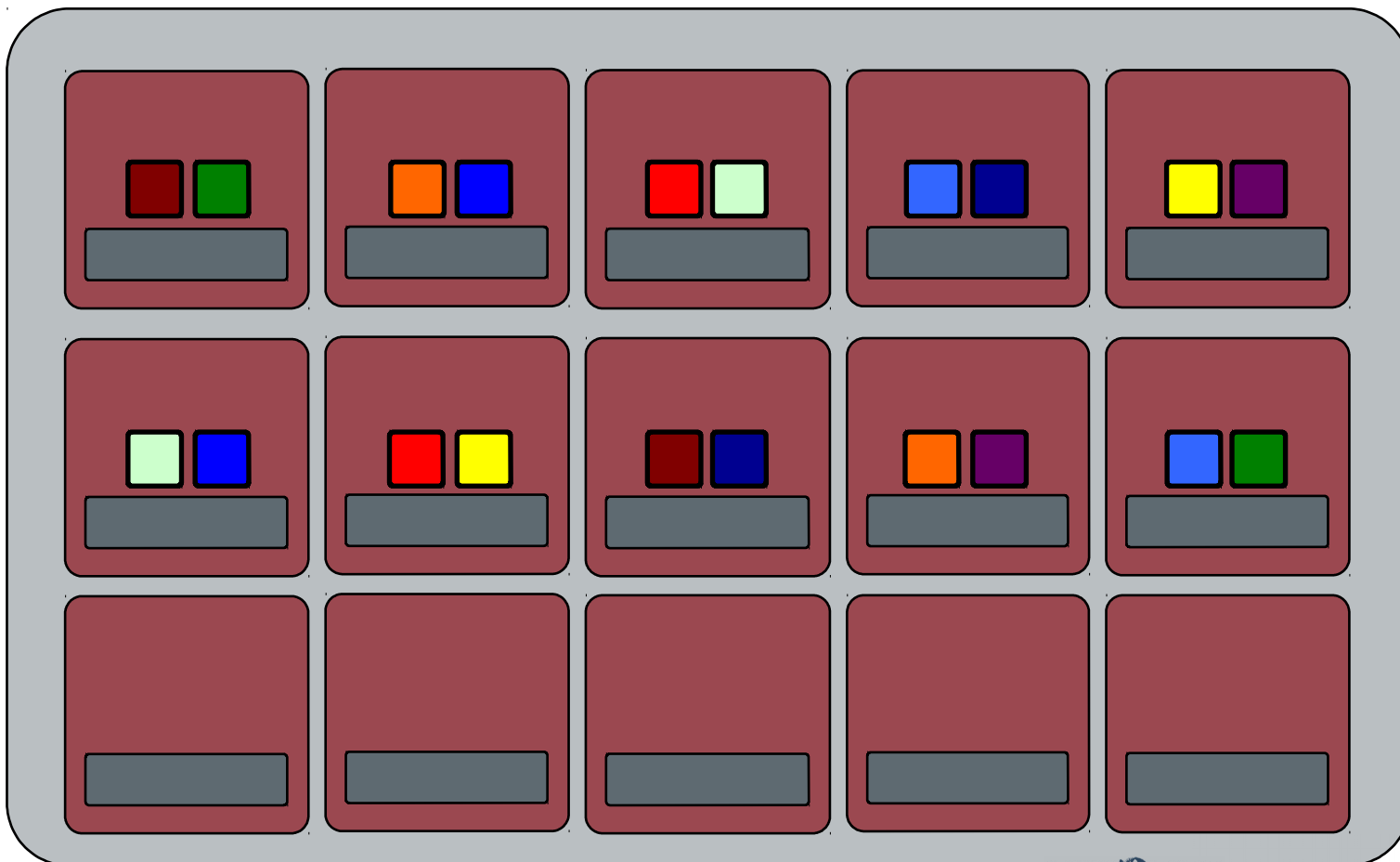
CRUSH: Yeni Dağılım



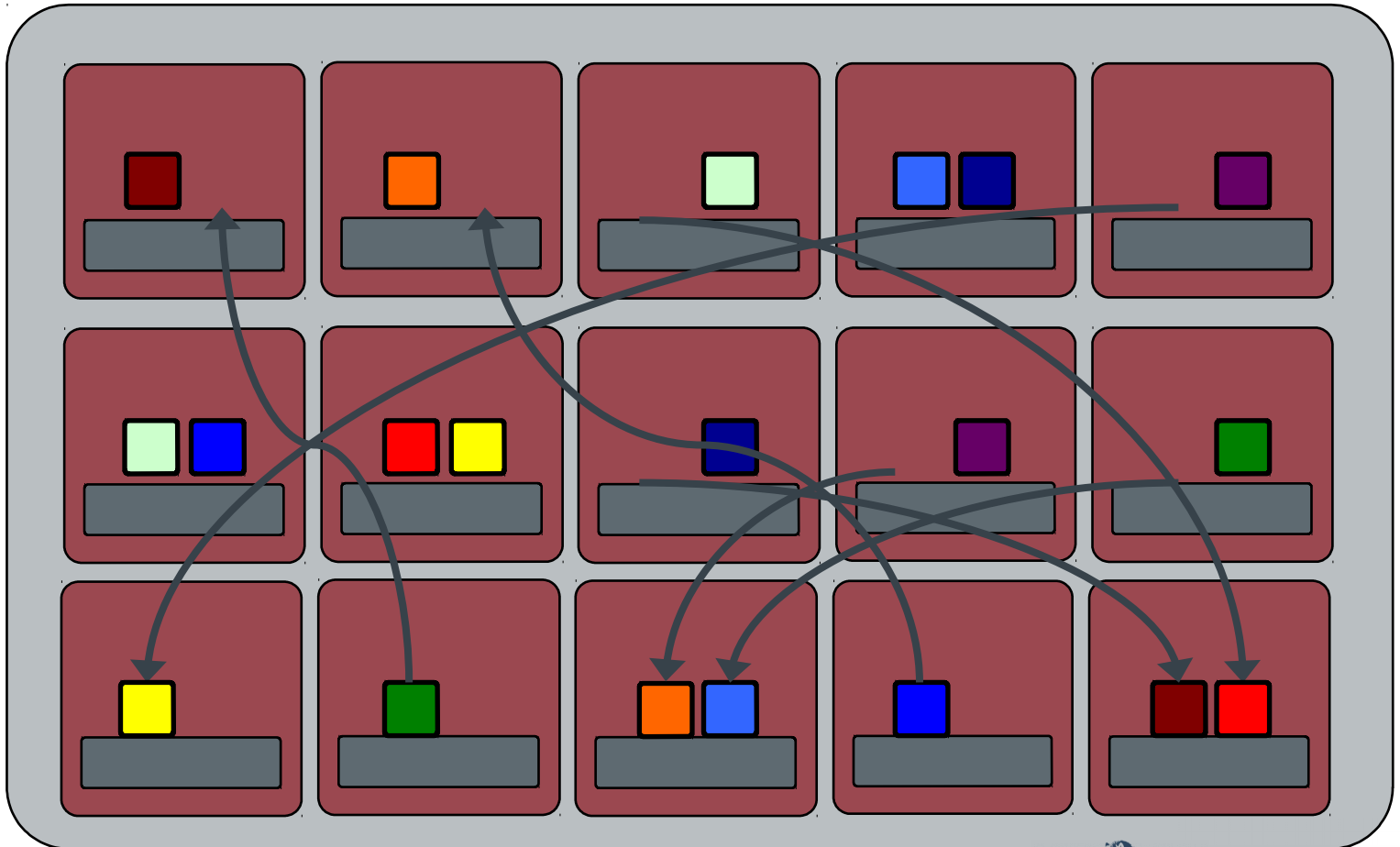
CRUSH:



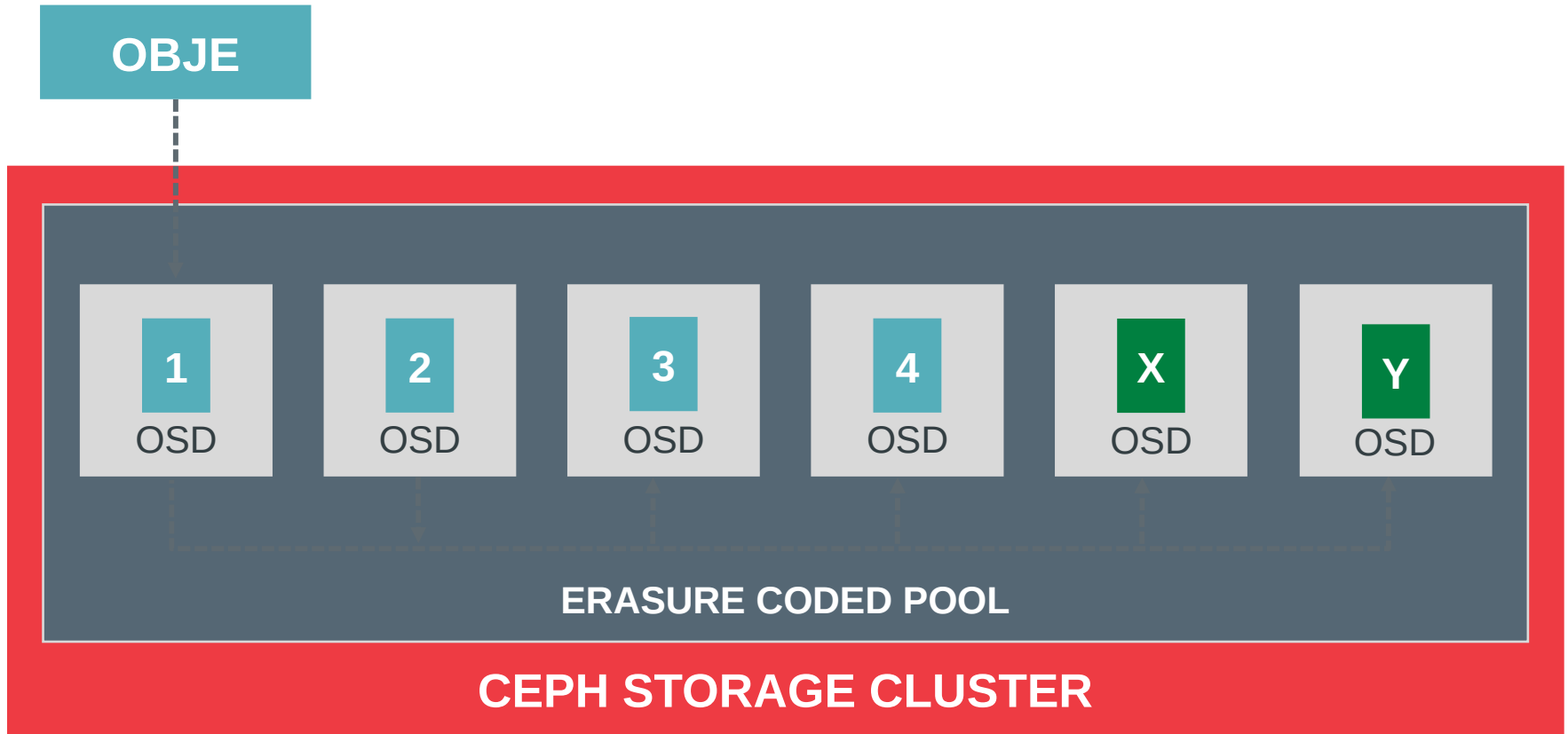
CRUSH



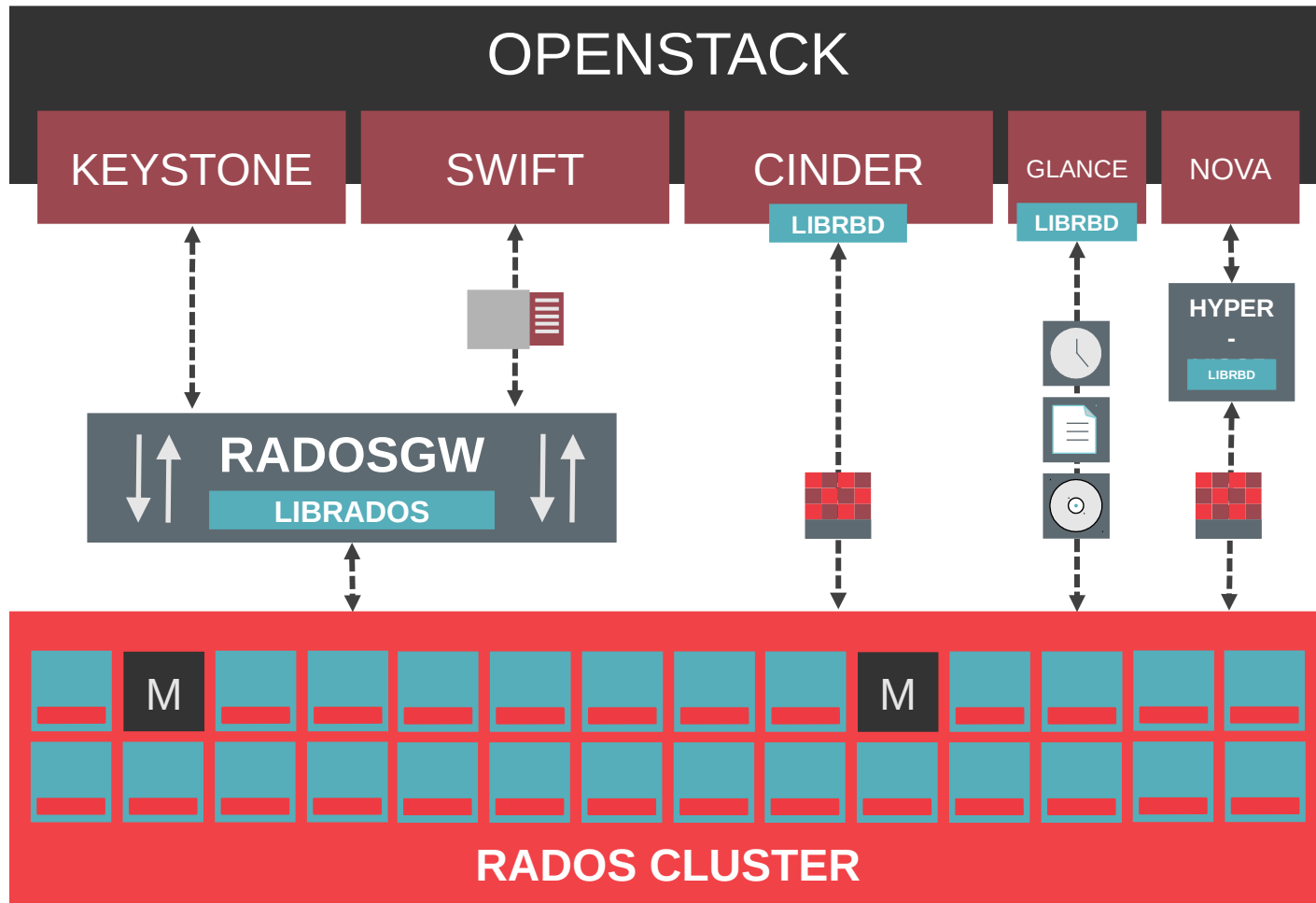
CRUSH



ERASURE CODING: Nasıl Çalışır



CEPH ve OPENSTACK Entegrasyonu





Sorular ve Cevaplar

Teşekkürler!

