Safa Onur Macit

data deduplıcatıon

**1) Projenin Amacı**

Günümüzde büyük verilerin kullanılmaya başlaması ile disk teknolojileri de kendisini bu sisteme adapte etti. Yalnız büyük verinin hızlı işlenmesi, yedeklenmesi, geri dönülmesi gibi noktalarda hız gereksinimi ortaya çıktı. Burada da data deduplication devreye giriyor. Mantık olarak birebir aynı kopyası bulunan verilerin tek kopyasını tutup, tek kopyayı kullanarak gerçek verilere linkleme yapan bir mekanizmadır.

**2) Data Deduplication Hakkında Bilgiler**

Data deduplication bir Windows System 2016 rolüdür. Bu servis depo alanındaki çift dataları bularak siler. Daha çok datayı daha az fiziksel disk alanı kullanarak depolamak için yapar.

Herhangi bir server’a data deduplication eklenip çalıştırıldığında;

1- Data segmentlerini daha küçük parçalara ayırır.

2- Birbirleriyle aynı olan parça varsa tanımlar.

3- Kullanılmayan kopyaları referansla değiştirir.

4- Parçaları sıkıştırır.

Windows Server 2016 ile eklenen özellikler;

1- 64 TB’lı size. Önceden 10 TB idi.

2- 1 TB’a kadar dosya boyutunu destekler.

3- Basitleştirilmiş konfigürasyon ( sanallaştırılmış backup uygulamaları için)

Data deduplication nerelerde kullanılmaz

1- Volume’ler sistem ya da boot volume ise kullanılmaz.

2- Volume’ler ya MTR ya da GPT olarak konfigüre edilmelidir. NTFS ya da ReFS ile formatlanmalıdır

3- Volume’ler Windows Server’a bağlı olması gerekmektedir ve çıkarılabilir disklere yapılamaz.

4- Volume’ler Fiber Channel, iSCSI SAN ya da SAS kullanılarak açılabilir.

5-32 KB’ın altındaki dosyalar data deduplication tarafından parçalanamaz.

6- Windows Client işletim sistemleri için uygun değildir.

Data deduplication işlemleri;

Bir volume’de data deduplication başlatıldığı zaman;

a) Bir algoritma kullanılarak 32 ile 128 KB arası datalar değişken boyutta parçalara ayrılır.

b) Aynı olan parçalar belirlenir.

c) Aynı olan parçalar bir parça kılıfına konulur.

d) Tekrarı olan parçalardan 2.sini ya referans ya da tek kopya ile değiştirilir.

e) Parçalar sıkıştırılır.

f) Datayı düzenler.

**3) İşlem Basamakları**

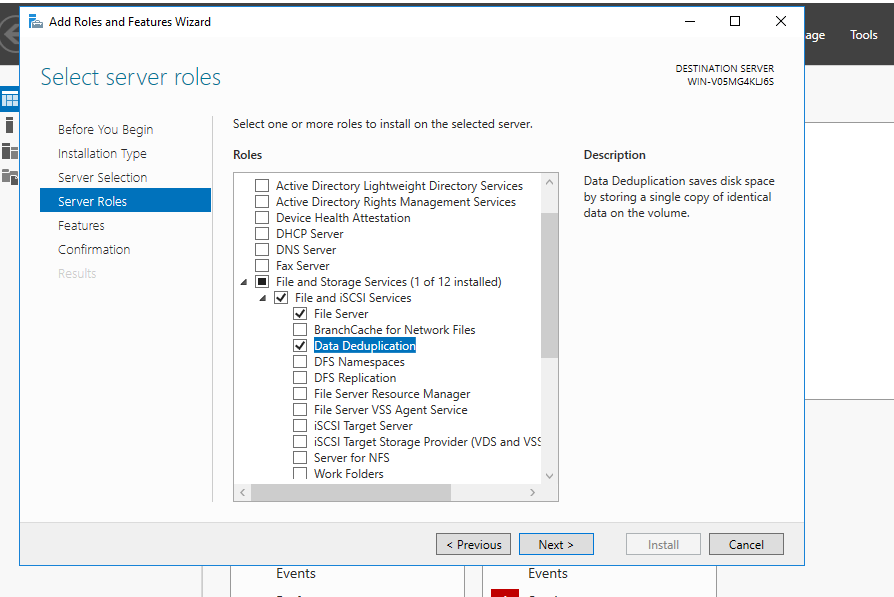
1) Server’a gerekli roller eklenir

2) Server’a disk eklenir.

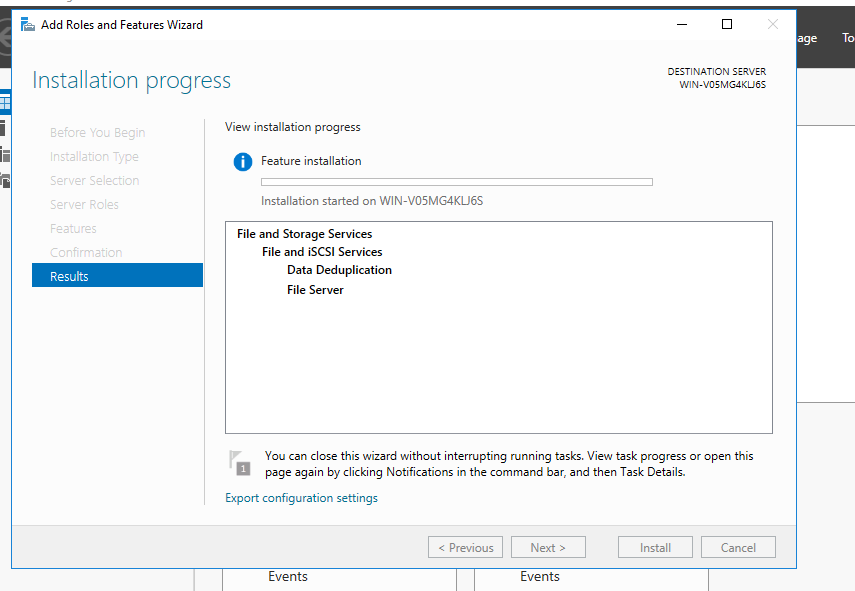
3) Data deduplication konfigüre edilir.

**4) Proje Adımları**

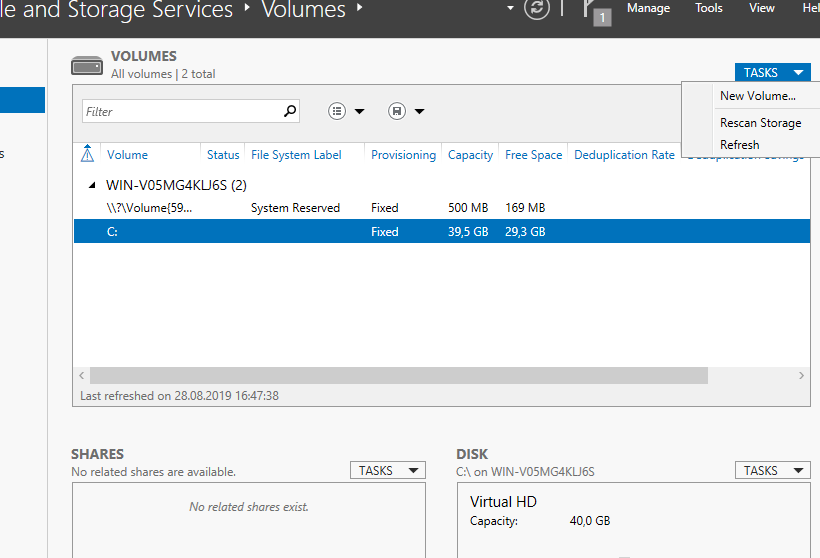
Add roles and features’dan eklenecek rol olarak Data Deduplication seçilir.



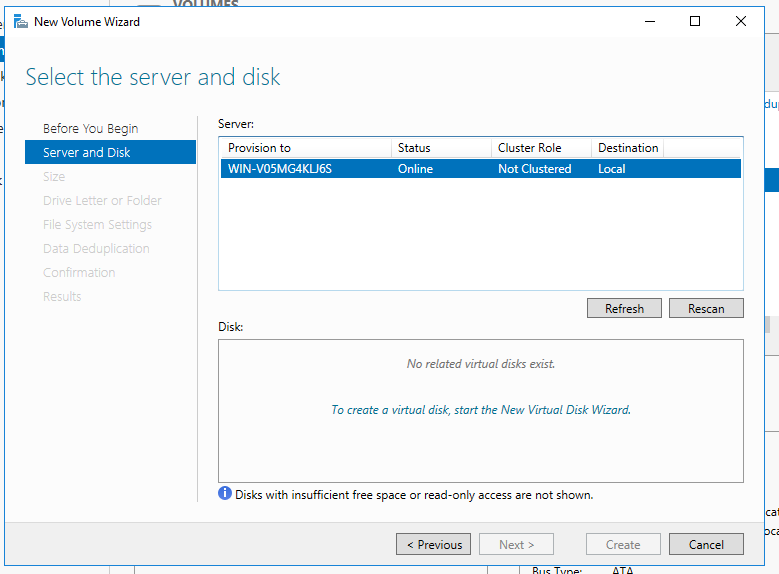
iSCSI data deduplication ile beraber yükleniyor



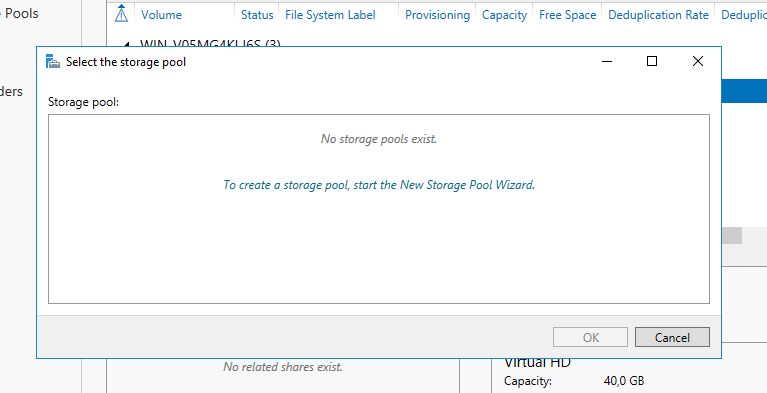
Eklenen yeni disk için new volume denir.



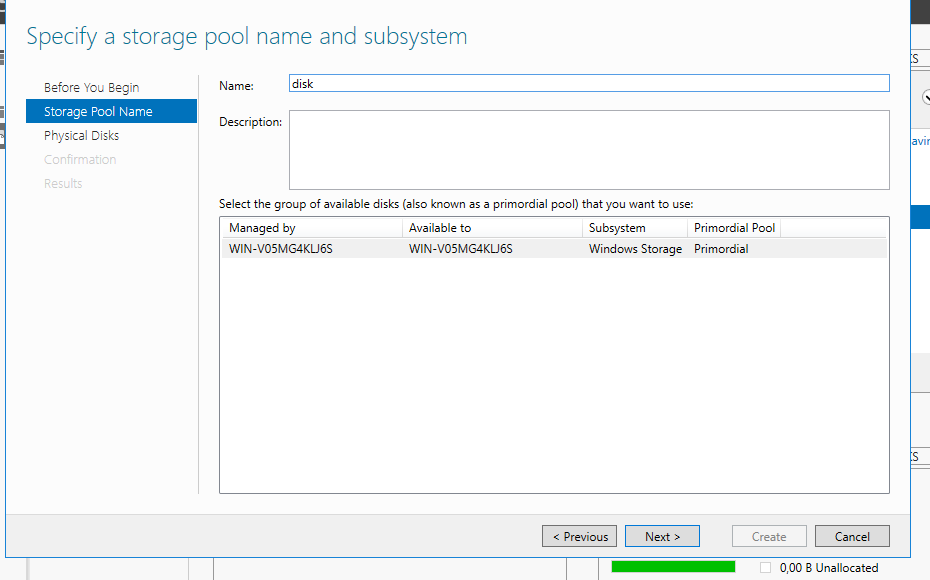
Disk için wizard açılır.



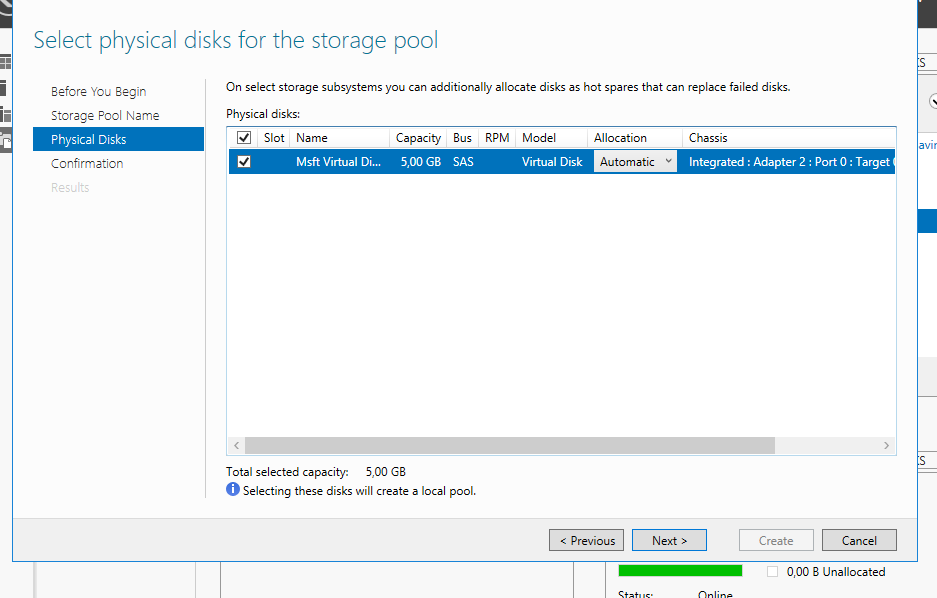
Yeni storage pool wizard açılır



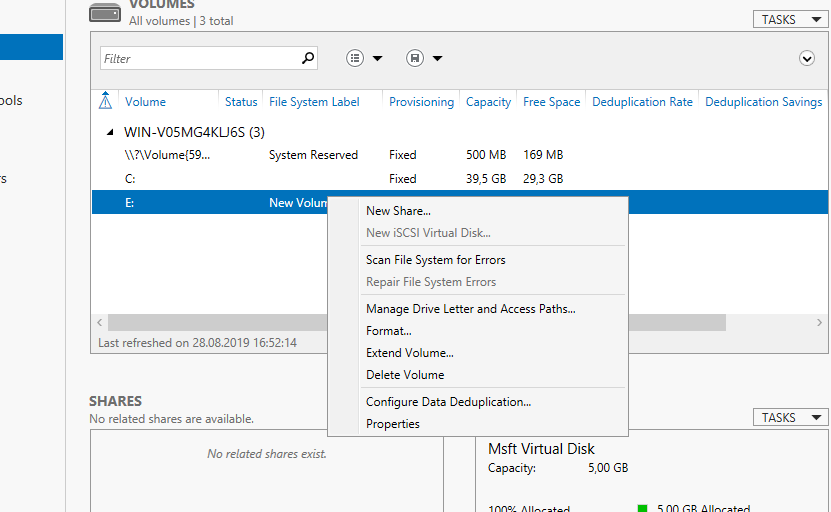
Disk adı verilir.



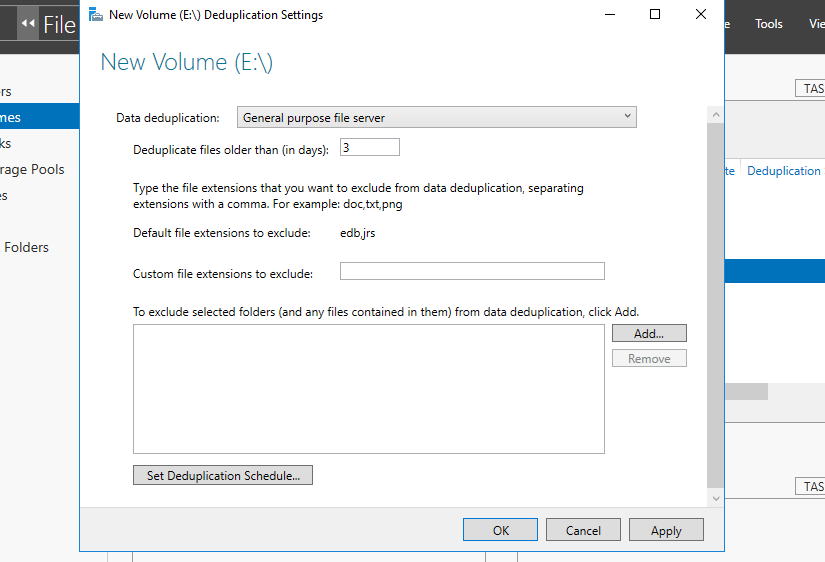
Dışardan eklenen disk göründü.



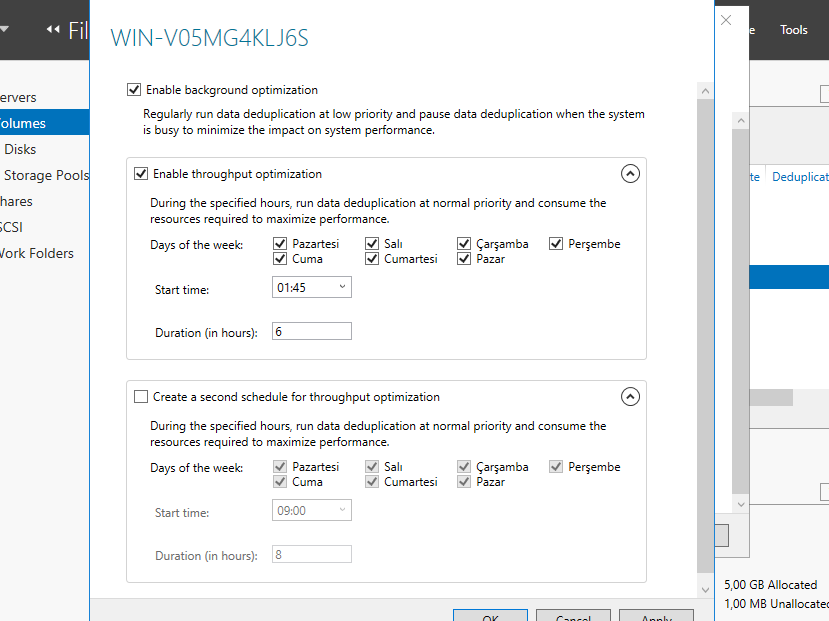
Data deduplication’u konfigüre etmeye başlanır.



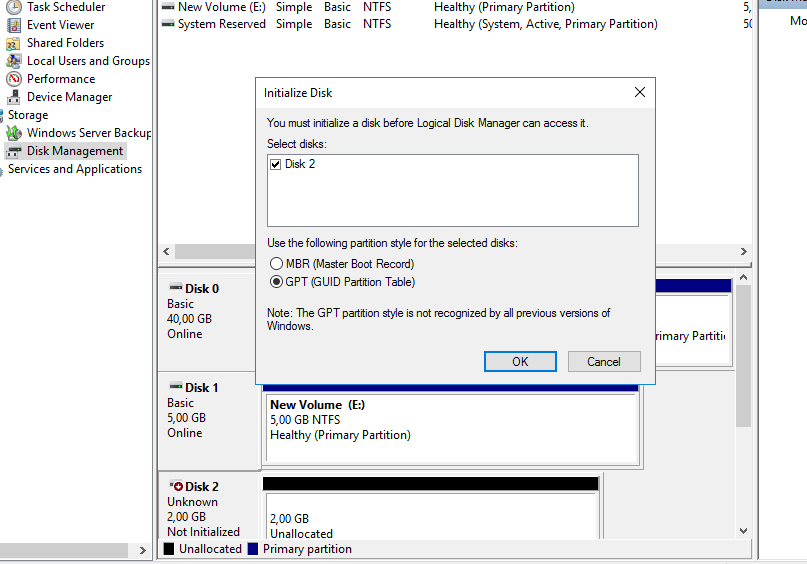
General purpose file server seçilir ve gün ayarı yapılır.



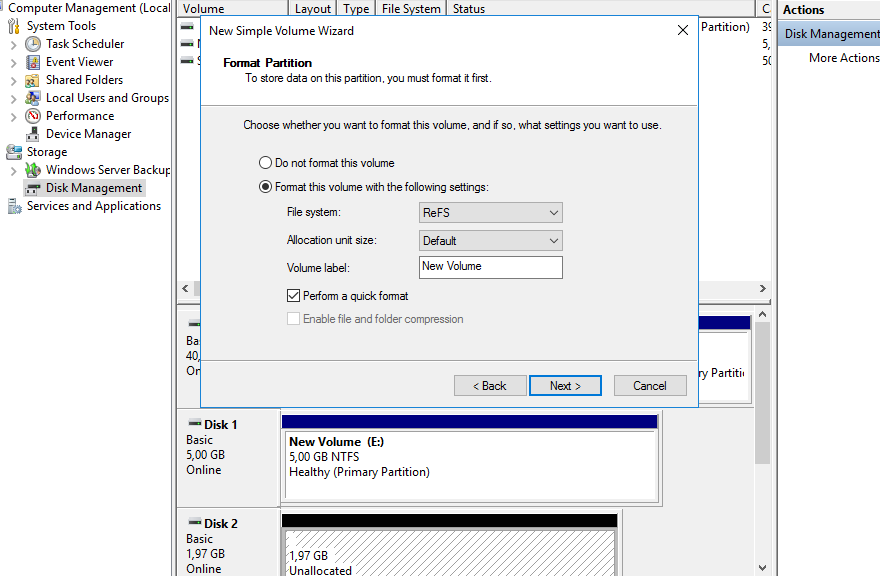
Enable throughput optimization seçilir ve ayar tamamlanır.



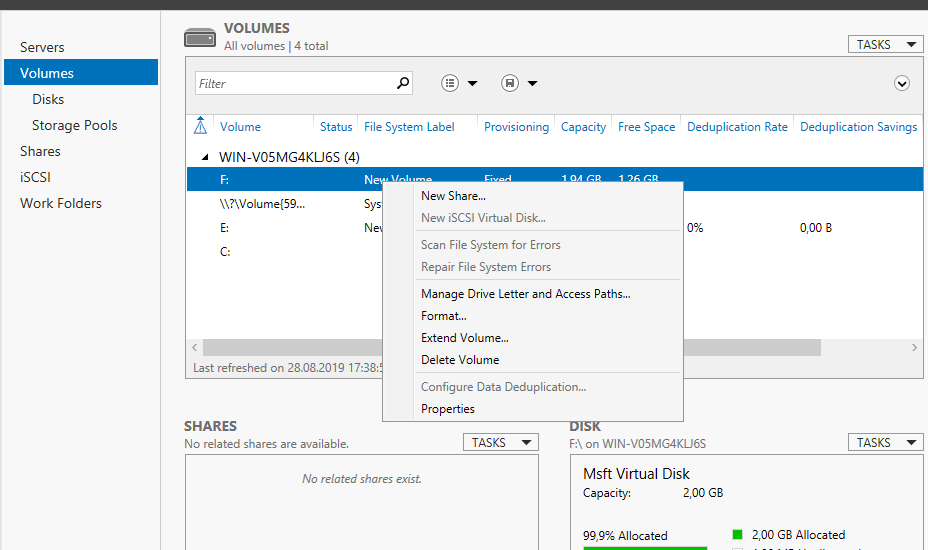
2.Disk eklenir



File sistem olarak ReFS seçilir.

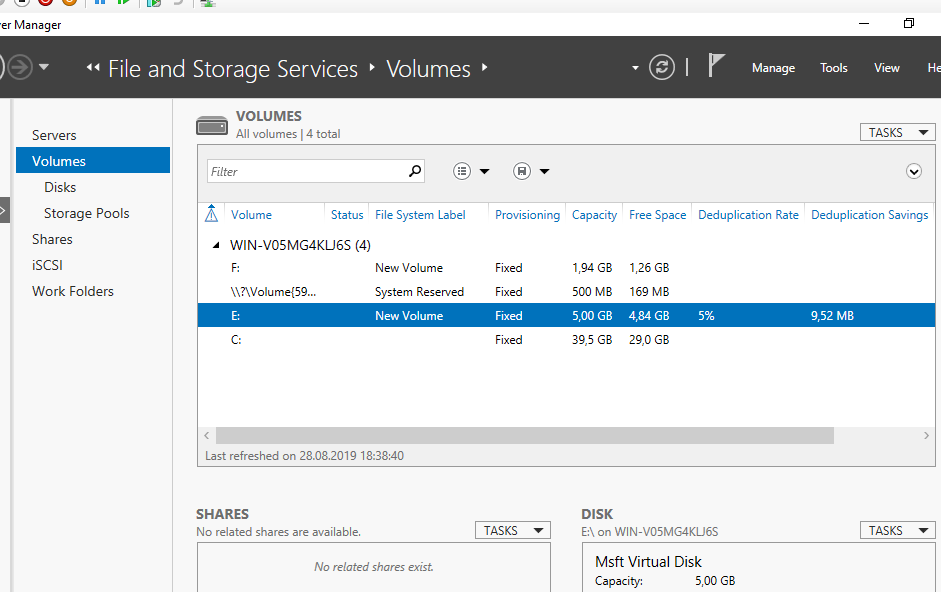


Görüldüğü gibi ReFS deduplica olmuyor.



**5) Sonuç**

Data Deduplication başladı.



Data deduplication’un olduğu PowerShell’den kanıtlandı.

