

Uzaktan Algılama Ürünlerinin Arşivlenmesi ve Ürünlere Erişim

Hüseyin Gökhan Akdağ

Meteoroloji Genel Müdürlüğü

Sistem Yönetimi ve Donanım Şube Müdürlüğü

Ankara

gakdag@mgm.gov.tr

Murat Demir

Meteoroloji Genel Müdürlüğü

Sistem Yönetimi ve Donanım Şube Müdürlüğü

Ankara

mdemir@mgm.gov.tr

ÖZET

Uzaktan algılama sistemlerinin çeşitliliği ve bu sistemlerden elde edilen ürünlerin sayısı gelişen teknolojiyle birlikte her geçen gün artmaktadır. Çoğu zaman, bu sistemlerden elde edilen geçmiş zamanlı ürünlere, çeşitli sebeplerle (akademik, askeri, istatistiki, vb..) ihtiyaç duyulmaktadır. Bu ürünlerin gerek yapısı gerekse boyutları sebebiyle geleneksel veri tabanı teknolojileri üzerinde uzun süreli olarak arşivlenmesi mümkün olmamaktadır. Geneli büyük boyutlu olan uzaktan algılama, sayısal model ve analiz ürünlerinin Meteoroloji Genel Müdürlüğü bünyesinde arşivlenmesine 2007 yılında başlanmış olup, 2018 yılında ise mevcut teknolojik alt yapı değiştirilmiş ve arşivleme kapasitesi arttırılmıştır. Bu sayede hem ürünlerin arşivlenmesinin hem de ürünlere erişimin daha hızlı ve kolay bir şekilde yapılması amaçlanmıştır. Bu bildiride arşivleme sistemi üzerinde arşivlenen verilerin detaylı bilgisi ve hangi yollarla bu ürünlere erişilebileceği konuları üzerinde durulacaktır.

Anahtar Kelimeler — uzaktan algılama ürünleri, ürünlere erişim, arşiv sistemi

1. GİRİŞ

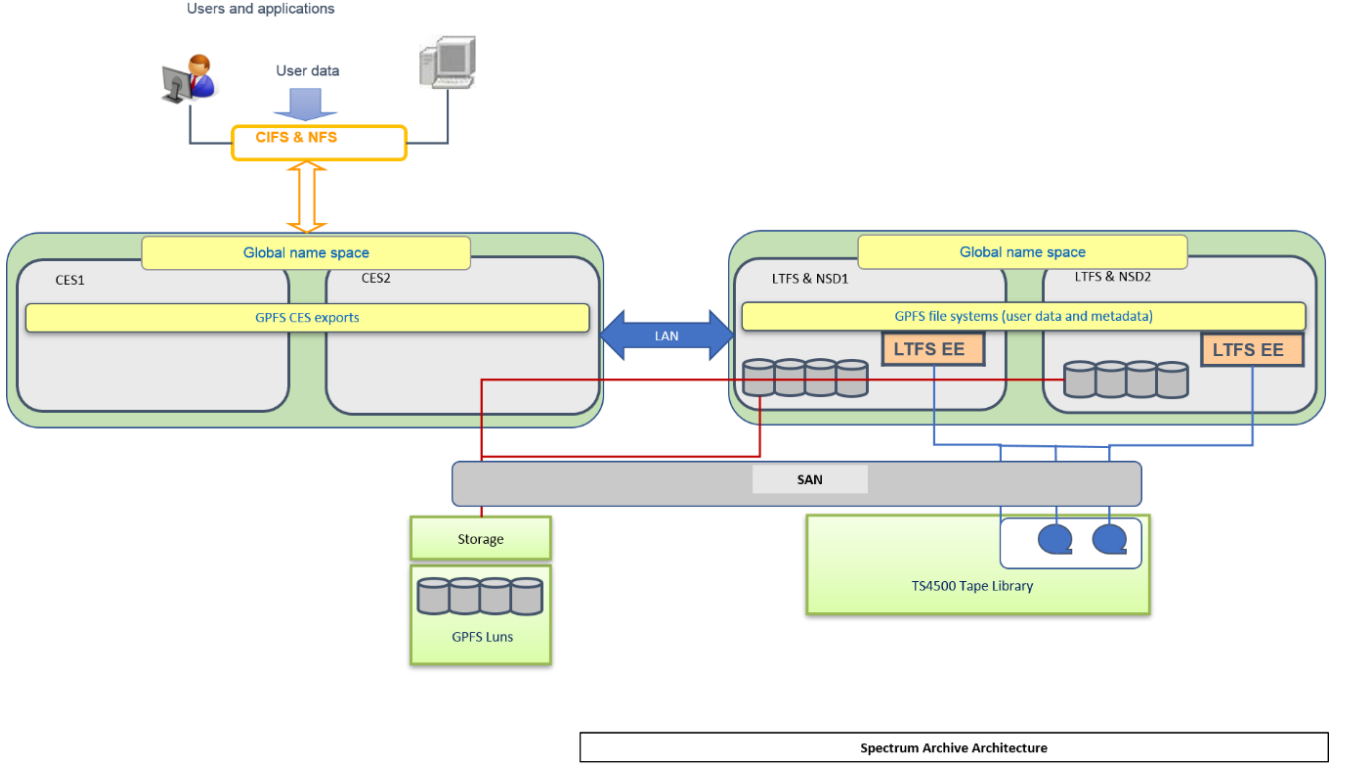
Genel Müdürlüğümüz bünyesinde 2007 yılında kurulan arşivleme sistemi ile Meteorolojik Uydu, Meteorolojik Radar ve Sayısal Model verileri gibi büyük boyutlu veriler arşivlenmeye ve ilgililere sunulmaya başlanmıştır. Veritabanında arşivlemeye uygun gözlem ve ölçüm verileri ise Veritabanı sunucularında saklanmakta ve bu sunucular üzerinden sunumları yapılmaktadır.

İlk kurulan sistemde donanım olarak 22 TB disk ve 290 TB kapasiteli IBM Teyp kütüphanesi, yazılım olarak da IBM Tivoli Storage Manager yazılımı kullanılmıştır. 2012 yılında yapılan güncelleme ile de sunucular ve disk sistemi yenilenmiş, hem disk sisteminin kapasitesi hem de teyp sisteminin kapasitesi de 350 TB'a çıkarılmıştır.

2. YENİ ARŞİV SİSTEMİ

Arşiv sisteminde kullanılan teyp kütüphanesinin kullanım ömrünün sonuna gelmesi ve aynı zamanda arşivlenecek olan verilerin boyutunun artması sebebiyle arşiv sisteminin yenilenmesine

ihtiyaç duyulmuştur. 2018 yılı ikinci yarısında yapılan güncellemeyle birlikte disk sistemleri ile artık destek alınamayan teyp kütüphanesi yenilenmiş, cache olarak kullanılmak üzere 100 TB kapasiteli IBM Storvize V7000, 1 PB kapasiteli IBM TS4500 teyp kütüphanesi ve yazılım olarak IBM Spectrum Protect ve Spectrum Archive yazılımları alınmıştır. Arşiv sisteminin blok diyagramı Şekil 1’de gösterildiği gibidir.



Şekil 1: Arşiv Sistemi Blok Diyagramı.

Verilerin asıl arşivlenme ortamları teyp kütüphanesi olduğu için, veri yedekliliğinin sağlanması için verilerin çift kopya olarak arşivlenmesi planlanmış olup bu nedenle net arşivleme kapasitesi 500 TB olarak gerçekleşmiştir.

3. VERİLERİN TOPLANMASI

Yeni kurulan arşivleme sisteminde veriler ilgili şubelerden (SHT Şb., Uzaktan Algılama Şb, Deniz Met. Şb., Araştırma D.) CES1 ve CES2 sanal sunucularına Linux - scp ile toplanmakta, CES sunucularında çalışan betiklerle LTFS1 ve LTFS2 fiziksel sunucuları altındaki GPFS dosya sistemine aktarılmakta, buradan da kalıcı olarak teyp sistemine çift kopya olarak arşivlenmektedir.

Veriler ilk olarak CES1 ve CES2 sanal sunucularında bulunan ilgili geçici dizinlere gönderilmektedir. 2 Gün süre ile bu dizinlerde tutulan veriler daha sonra kalıcı olarak arşivlenmek üzere arşiv dizinlerine taşınmaktadır. Bu dizinler, daha önce bahsedilen 100 TB boyutundaki disk sistemi üzerinde bulunmaktadır. Disk sistemi 60% doluluk oranına ulaşana kadar veriler hem disk sistemi üzerinde hem de teyp sistemi üzerinde bulundurulmaktadır. Disk sistemi üzerinde bulunan verilere erişimin sadece teyp sisteminde üzerinde olan veriye erişime kıyasla çok daha hızlı olması sebebiyle yakın tarihli yaklaşık 6 aylık bir veri hem disk hem de teyp ortamında bulundurulmaktadır.

4. ARŞİVLENEN ÜRÜNLER

Meteoroloji Genel Müdürlüğü bünyesinde 2007 yılında beri büyük boyutlu verilerin arşivlenmesi yapılmaktadır. Hangi verilerin arşivleneceği, veriyi üreten birimlerin isteği üzerine belirlenmektedir. Bu kapsamda Sayısal Hava Tahmini Şube Müdürlüğü, Uzaktan Algılama Şube Müdürlüğü, Deniz Meteorolojisi Şube Müdürlüğü ve Analiz ve Tahminler Şube Müdürlüğü tarafından üretilen ve arşivlenmesini istedikleri veriler arşiv sisteminde arşivlenmektedir. Şekil 2’de Sayısal Hava Tahmini Şube Müdürlüğü tarafından üretilen arşiv verileri gösterilmiştir. Daha önce üretilen bazı ürünlerin çeşitli sebeplere sonlandırılması sonucunda arşivlenmesi yapılmamaktadır. Ama üretildiği zamanlar için bu verilere erişim mümkündür. Şekil 3’de Uzaktan Algılama Şube Müdürlüğü tarafından üretilen ve arşivlenen uydu ürünlerinin bilgisi gösterilmiştir. Burada verilen aylık veri boyutları yaklaşık olarak verilmiş olup değişen hava koşullarına göre değişiklik göstermektedir. Şekil 4’de Radar ağında bulunan radarlardan ve onlardan üretilen bazı birleştirilmiş verilerle ilgili bilgiler verilmiştir. Bu radarlara ek olarak İstanbul deniz radarı ürünlerinde farklı bir dizin üzerinden arşivlenmektedir.

Model Metadata Bilgileri

Urun	Baş.Tarihi	Bit.Tarihi	Durumu	Model Ürün Sayısı (Çalışma Zamanları)	Dosya Boyutu	Veri Uzantısı	Örnek Ürün Adı
ECB	2006-03-09		Devam Ediyor	Günde 4 Dosya (00,06,12,18)	~4x1.2 GB	GRIB	ECB_201710090000_201710120000.grb
ECD	2007-10-02		Devam Ediyor	Günde 2 Dosya (00,12)	~2x1.92 GB	GRIB	ECD_201710090000_201710190000.grb
ECE	2007-10-02		Devam Ediyor	Günde 2 Dosya (00,12)	~2x28 MB	GRIB	ECE_201710090000_201710240000.grb
ECW	2007-10-02	2011-01-31	Sonlandırıldı	Günde 2 Dosya (00,12)	~2x1.5 MB	GRIB	ECW_201101310000_201102050000.grb
MET	2007-10-03	2018-09-31	Sonlandırıldı	Günde 1 Dosya (00)	~918 MB	NetCDF	MET_2017100900.nc
MD0	2007-11-26	2011-05-10	Sonlandırıldı	Günde 1 Dosya	~1.9 MB	crb	MD0_201105101200_201105131200.crb
DREAM	2010-06-02		Devam Ediyor	Günde 1 Dosya	~161 MB		eta_dream_20171001.tar.gz
IMG	2011-01-19	2018-09-31	Sonlandırıldı	Günde 2 Dosya (00,12)	~2x2.9 GB	jpg	IMG_2017100900.tar
ALR	2011-01-20		Devam Ediyor	Günde 72 Dosya (00..72)	~72x233 MB	fa	ALR_2017100900_00.fa
ARP	2011-07-01		Devam Ediyor	Günde 4 Dosya (00,06,12,18)	~4x813 MB	fa	ARP_2017100912.tar
WRF ?	2013-02-20	2018-09-31	Sonlandırıldı	Günde 73 Dosya (00..72)	~73x891 MB	NetCDF	WRF_2017100900_00.nc
WRF	2018-10-01		Devam Ediyor	Günde 24 Dosya (00..23)	~24*6.65 GB	NetCDF	WRF_2018100101_00.nc
LAEF	2013-09-01		Devam Ediyor	Günde 2 Dosya(00,12)	~2x58 MB	GRIB	LAEF_2017101012.grb
AROME	2018-10-01		Devam Ediyor	Günde 49 Dosya(00,48)	~2x1.4 GB	fa	AROME_2018110600_00
WW3	2018-10-01		Devam Ediyor	Günde 2 Dosya(00,12)	~2x2.1 GB	NetCDF	WW3_2018102100

Şekil 2: Model ürünleri.

Veri Boyutları Aylara Göre Değişiklik Gösterebilir.

Urun	Baş.Tarihi	Bit.Tarihi	Durumu	Aylık Veri Boyutu
NOAA	2008-04-09		Devam Ediyor	~630 MB
MSG	2008-04-09	2009-08-03	Sonlandırıldı	~7.5 GB
SAF	2008-05-01		Devam Ediyor	~362 MB
HSAF	2009-08-05		Devam Ediyor	~2.9 GB
MSG_HAM	2009-08-05	2012-05-31	Sonlandırıldı	~6.8 GB
MPEF	2009-08-05	2012-05-31	Sonlandırıldı	~1 GB
RESIM	2009-08-05		Devam Ediyor	~332 MB
MODIS	2012-11-22		Devam Ediyor	~243 MB

Şekil 3: Uydu ürünleri.

Lokasyon	Baş.Tarihi	Bit.Tarihi	Durumu	Aylık Veri Boyutu
ZNG	2007-03-01		Devam Ediyor	~43 GB
ANK	2007-04-04		Devam Ediyor	~20 GB
BLK	2007-04-04		Devam Ediyor	~45 GB
IST	2007-04-04		Devam Ediyor	~50 GB
MGL	2010-09-01		Devam Ediyor	~22 GB
IZM	2010-09-01		Devam Ediyor	~61 GB
ANT	2011-08-01		Devam Ediyor	~54 GB
HTY	2011-08-01		Devam Ediyor	~57 GB
SMN	2012-07-08		Devam Ediyor	~75 GB
TRZ	2012-07-08		Devam Ediyor	~69 GB
CMP	2014-01-02		Devam Ediyor	~1.3 GB
MOB	2015-05-11		Devam Ediyor	~4.6 GB
BRS	2015-06-03		Devam Ediyor	~46 GB
KRM	2015-06-03		Devam Ediyor	~47 GB
AFY	2015-11-26		Devam Ediyor	~50 GB
SRF	2016-01-18		Devam Ediyor	~100 GB
GZT	2016-01-18		Devam Ediyor	~73 GB
SVS	2016-12-19		Devam Ediyor	~30 GB
ERZ	2016-12-19		Devam Ediyor	~62 GB
CP2	2017-08-17		Devam Ediyor	~20 GB
AYT	2018-10-01		Devam Ediyor	~23 GB

Urun	Aylık Veri Boyutu
PPI	15 GB
CAP	37 GB
RN1	2 GB
RNN	13 GB
VIL	8 GB
MAX	9 GB
WND	12 GB
SRI	6 GB
BAS	1.1 GB
TOP	7.5 GB
RAW	758 GB
JPG	51 GB

Şekil 4: Radar ürünleri.

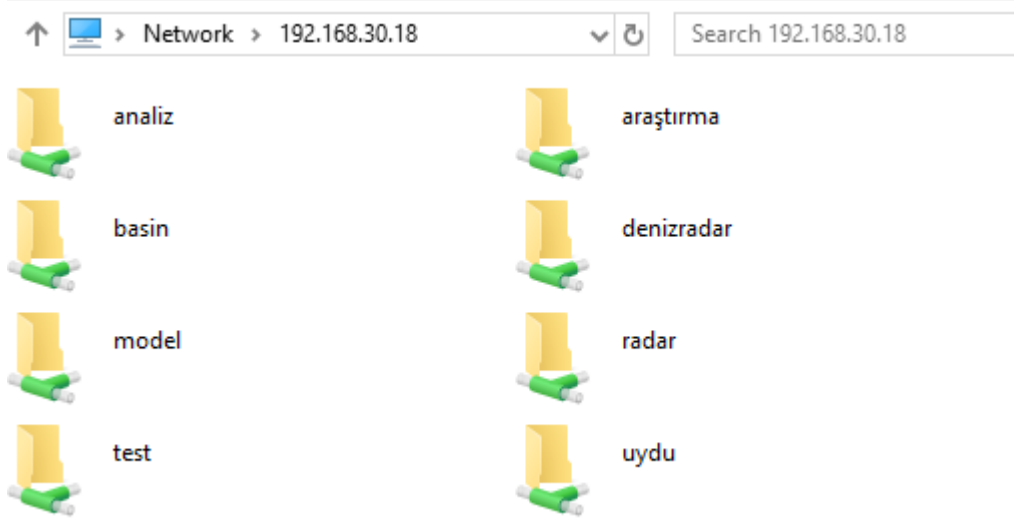
5. VERİYE ERİŞİM YÖNTEMLERİ

Arşiv sisteminde bulunan verilere üç farklı yöntemle erişmek mümkündür. Veriye erişim için öncelikle birim amirinin kullanıcılar için mail aracılığı talepte bulunması gerekmektedir. Bu talep sonrasında kullanıcı Meteoroloji Genel Müdürlüğü Domainde bulunan mgmarsiv kullanıcı grubuna dahil edilir. Bu işlemlerden sonra kullanıcı arşiv sisteminde bulunan bütün verilere erişim sağlayabilir.

5.1. Windows Gezgini Kullanarak Veriye Erişim

Kullanıcılar Windows işletim sistemlerinde bulunan Windows gezgini aracılığı ile [\\192.168.30.18](http://192.168.30.18) veya [\\192.168.30.19](http://192.168.30.19) bağlantısı üzerinden arşiv sistemine erişebilir. Eğer kullanıcılara dizinlere

erişim izni olmadığına dair bir uyarı veriliyor ise bu durumda kullanıcının mgmarsiv grubuna dahil edilmediği anlaşılabilir. Şekil 5’de Windows Gezgini ile ilgili örnek paylaşılmıştır. Windows gezgininde ürünler görüntülenirken bazı ürünlerin sol alt köşesinde x işaretinin olduğu görülmektedir. Bu işaretin anlamı ürünün kalıcı teyp sistemine aktarıldığı ve disk üzerinde bulunmadığı anlaşılmaktadır. Bu durumda veriye erişim yaklaşık 2 dakikalık bir gecikmeyle sağlanmaktadır. Dolayısı ile sadece teypte bulunan verilere erişim sağlanmaya çalışıldığında bir süre beklenmesi gerekmektedir.



Şekil 5: Windows Gezgini ile ürünlere erişim.

5.2. Web Arayüzü Kullanarak Veriye Erişim

Sistem Yönetimi ve Donanım Şube Müdürlüğü tarafından hazırlan web sayfası aracılığı ile arşiv ürünlerine erişim mümkün olmaktadır. Bu web sayfasına <https://arsiv.mgm.gov.tr> adresinden erişim sağlanabilir. Kullanıcılar sisteme giriş yapmamışlar ise sadece arşiv verilerinin metadata bilgilerini ve arşiv sisteminin genel durumunu görüntüleyebilirler. Daha önce mgmarsiv grubuna eklenmiş olan kullanıcılar sisteme kendi domain kullanıcı adı ve parolalarıyla giriş yapabilir ve veri talebinde bulunabilirler.

Buradan yapılacak olan talepler verinin boyutuna, disk sisteminde veya teyp sisteminde oluşuna ve kurumun network hızına göre belirli bir süre içerisinde hazırlanmaktadır. Verilerin hazırlanması aşamasından sonra kullanıcıya elektronik posta aracılığı ile verilerin hazır olduğu

bilgisi iletilmektedir. Eđer verilerin hazırlanması sürecinde herhangi bir hata oluşmuş ise yine kullanıcıya bir bilgilendirme maili gönderilmektedir.

Kullanıcılar hazırlanan verilere 2 süreyle web sayfası üzerinden erişebilirler. 2 günden daha eski veriler web sunucusu üzerinden silinmektedir. Kullanıcılar eđer isterlerse bir Linux makinaya veya bir ftp sunucusuna verilerini otomatik olarak gönderebilirler. Bunun için kullanıcıların Linux makine veya ftp sunucusu için gerekli olan kullanıcı adı ve parolalarını tanımlamaları gerekmektedir. Veri talebi aşamasında kullanıcılar verilere ftp veya scp gönderim şekillerinden birini seçerek sistemin erişebildiđi sunuculara verileri gönderebilirler. Bu aşamada girilen bilgilerin doğruluđu test edilebilir veya daha önceki bilgiler güncellenebilir.

5.3. Ortak Linux Sunucu Üzerinden Veriye Erişim

Linux işletim sistemi kullan kullanıcılar için Linux işletim sistemi olan bir sanal sunucu oluşturulmuştur. Oluşturulan bu sunucu üzerinden arşiv sisteminin model, uydu ve radar dizinlerine NFS ile erişim sağlanabilmektedir. Erişilen veriler ftp veya scp aracılığı ile istenilen başka bir noktaya kopyalanabilir. Yine eđer veriler 6 aydan eskiyse sadece teyp sisteminde olacağı için verilerin kopyalanmasına hemen başlanmaz. Kullanıcıların bir süre beklemesi gerekmektedir.

6. PLANLAN ÇALIŞMALAR

Arşiv sistemi ile ilgili 2020 yılı için bazı güncellemeler planlanmaktadır. Bu güncellemeler sistemin kendisi ile ilgili olup veriye erişim ile ilgili bir deđişiklik olmayacaktır. Buna göre arşiv sistemi teyp kapasitesinin ham olarak 1 PB seviyesine çıkarılması planlanmaktadır. Bu sayede hali hazırda veri boyutları sebebiyle sadece bir kısmını arşivlediđimiz verilerin tamamının arşivlenmesi mümkün olacaktır. Disk kapasitesi 200 TB olacak şekilde bir güncelleme yapılacaktır. Yakın tarihli 1 yıllık verinin hem disk hem teyp sisteminde bulunması veriye erişim hızını arttıracaktır. Arşiv sistemi web sayfası arsiv.mgm.gov.tr üzerinde bazı yazılımsal deđişiklikler yapılması planlanmaktadır.

7. SONUÇLAR

Meteoroloji Genel Müdürlüđu bünyesinde 2007 yılında günümüze büyük boyutlu veya veritabanına arşivlenemeyen ürünlerin arşivlenmesi teyp sistemi üzerine yapılmaktadır. Bu kapsamda arşivlenen toplam veri boyutu yaklaşık olarak 370 TB civarındadır. 4 Farklı birimden 25 den fazla farklı ürünün arşivlenmesi gerçekleştirilmektedir. Günlük arşiv boyutu yaklaşık

olarak 320 GB olarak gerekleřmektedir. Verilere eriřim üç farklı yöntemle gerekleřtirilmektedir. Verilere eriřimde birim amirlerinin bilgisinde kullanıcılara verilere eriřim izni verilmektedir. Yakın zamanda yapılacak olan güncellemeyle birlikte daha fazla verinin arřivlenmesi planlanmakta ve veriye eriřim hızının arttırılması amaçlanmaktadır.