

ATLAS Deneyinde Veri İşleme

Nurcan Öztürk

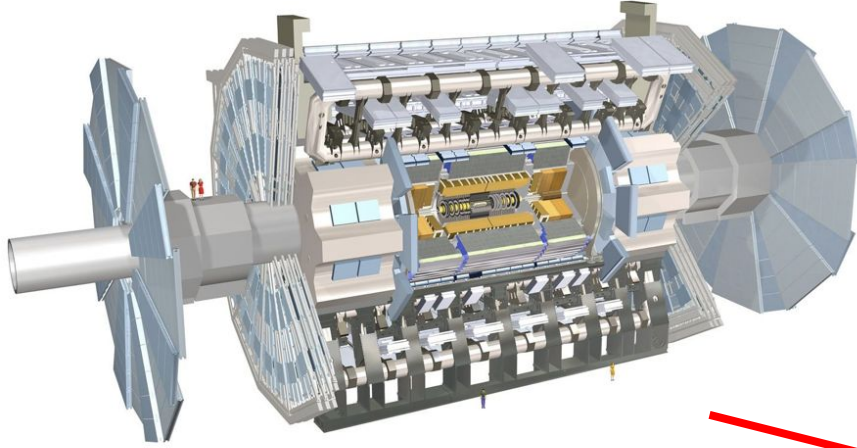
University of Texas at Arlington

10. Türk Öğretmen Çalıştayı

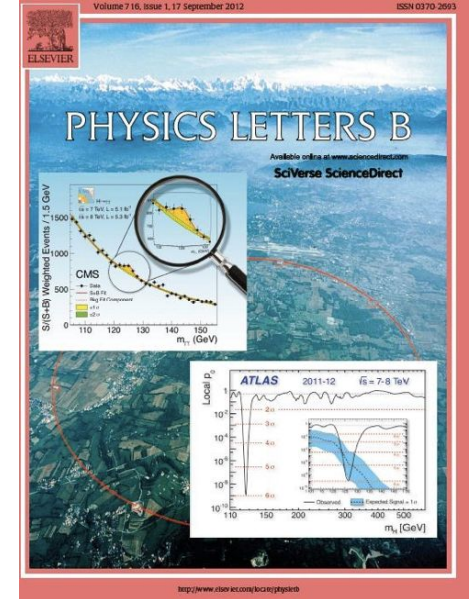
Ocak 27-31, 2020 CERN



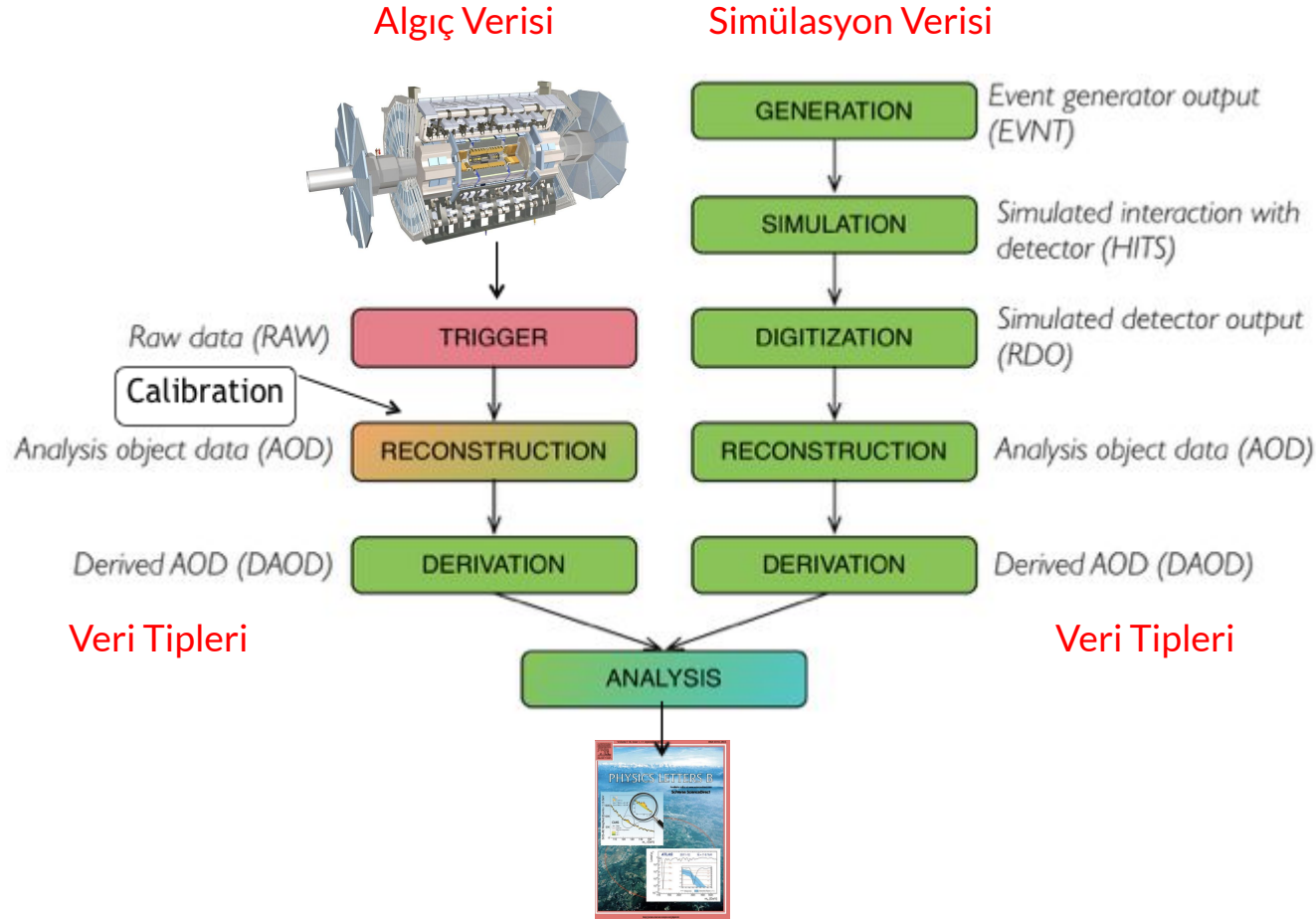
Algıçtan makaleye nasıl gidiyoruz?



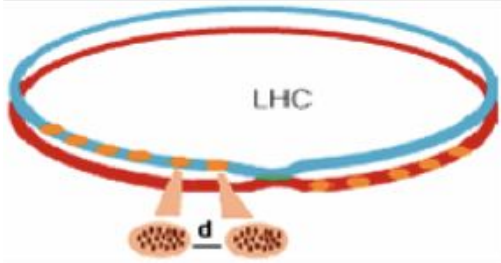
Veri Alımı
Tetikleme Sistemi
Veri Hazırlama
Veri İşleme
Veri Analizi
Fizik



Algıçtan makaleye nasıl gidiyoruz?

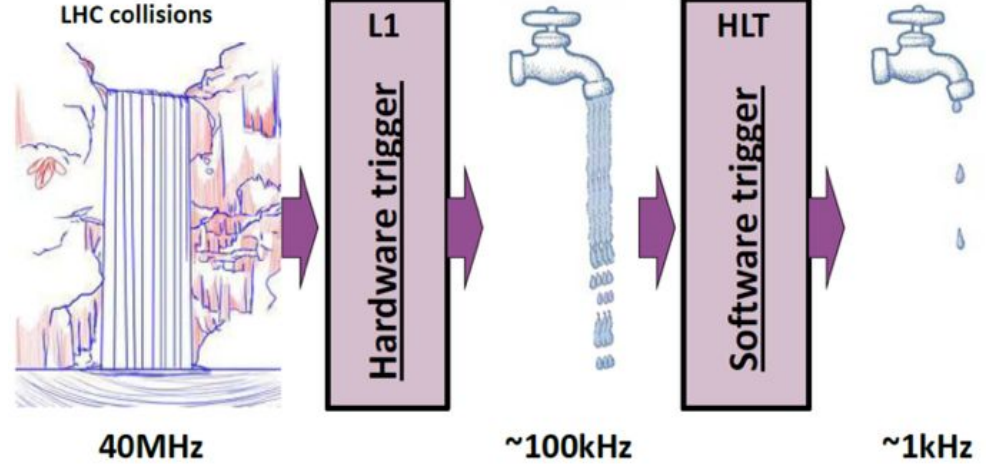


Tetikleme (Trigger) Sistemi



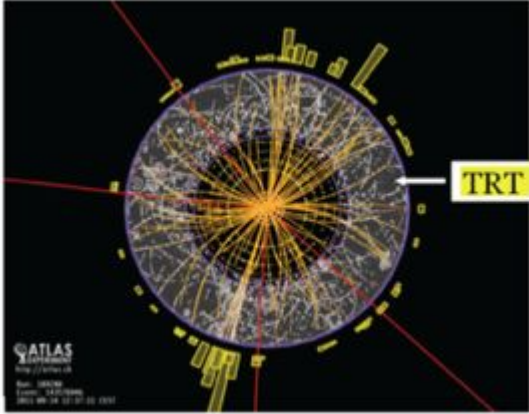
- Her bir proton demetinde 10^{11} tane proton var
- Proton demetleri arası 25ns
- Çarpışma frekansı 40 MHz

25 tane çarpışma noktası olan bir olay (event)

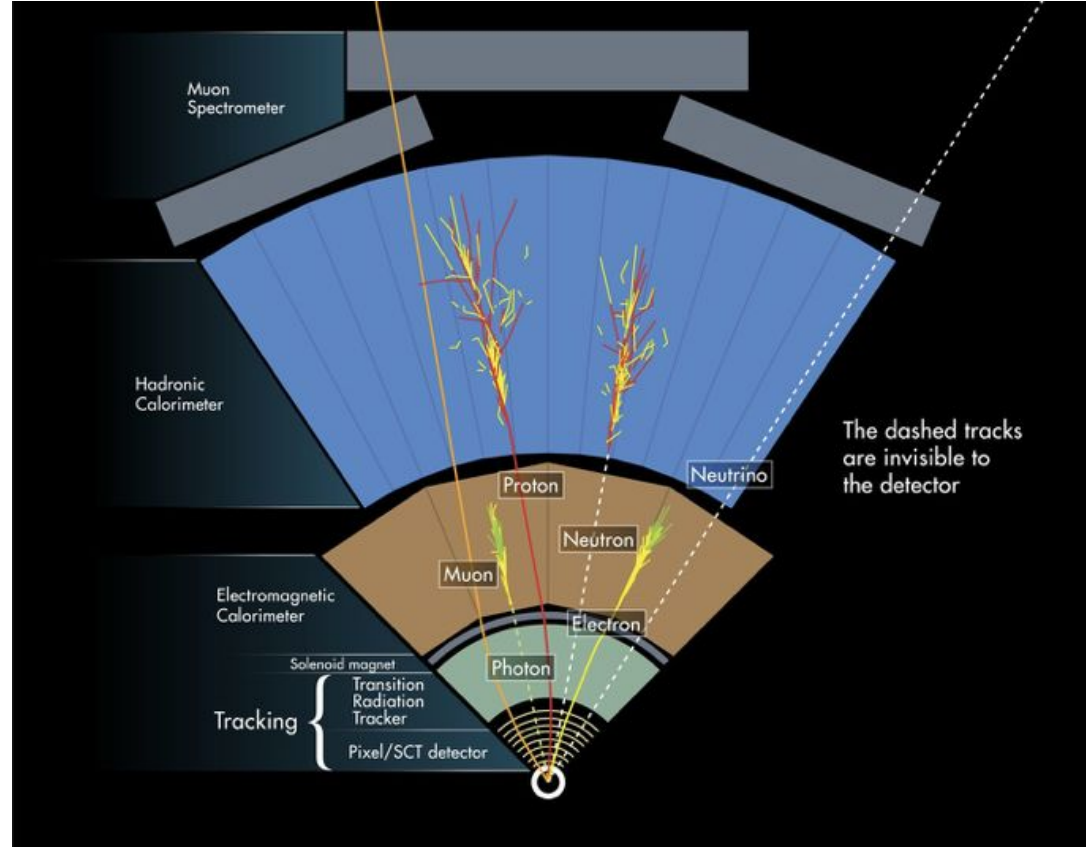


Tüm çarpışma olayların kaydetmek mümkün değil, tetikleme sistemi ile ancak ilginç olaylar kaydedilir

Veri İşleme (Reconstruction)



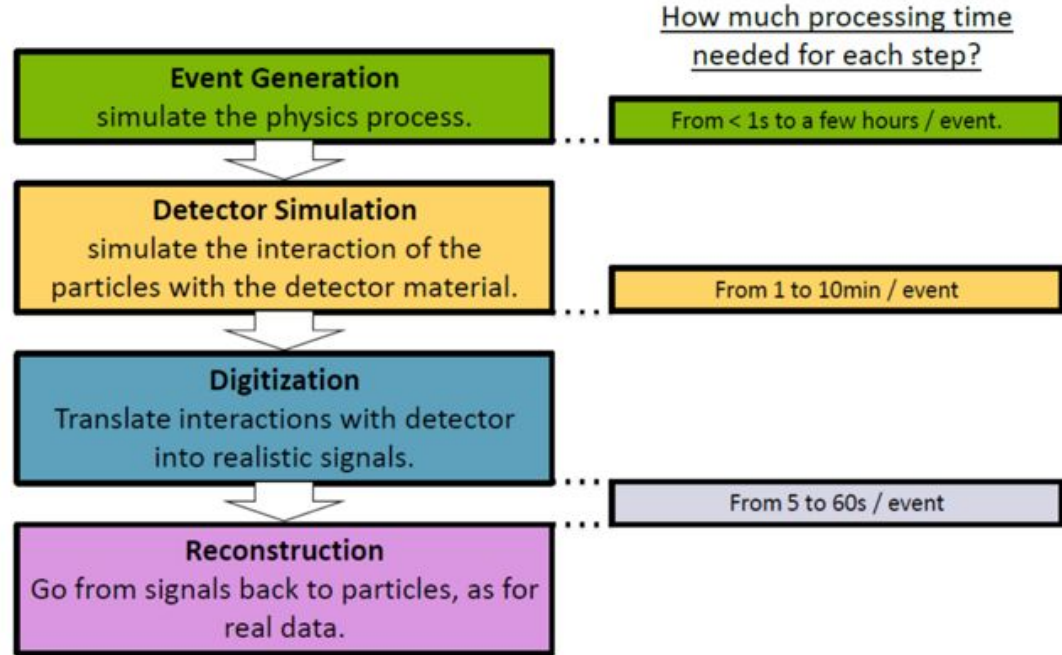
- Tetikleme sisteminden alınan ham verinin (RAW) kalibrasyonu yapılır ve fizik analizinde kullanılacak objelere (elektron, muon, jet,..) dönüştürülür.
- Kalibrasyon: Algıçın konumu, proton demetinin konumu, algıçın gürültülü kanalları belirlenir. Bu bilgi ile veri işlenir.



Simülasyon Verisi (Monte-Carlo)

Simülasyon verisi neden gerekli:

- İlginç fizik süreçlerinin modellenmesi
- Bu süreçler için beklenen arka plan olaylarının modellenmesi
- Algıçın performansının anlaşılmasının test edilmesi
- Ham verinin işlenip fizik objelerinin oluşturulmasındaki etkinliğin çalışılması
- Sistemik belirsizliklerin çalışılması
- ...



Veri Setleri ve İsimleri

DATA

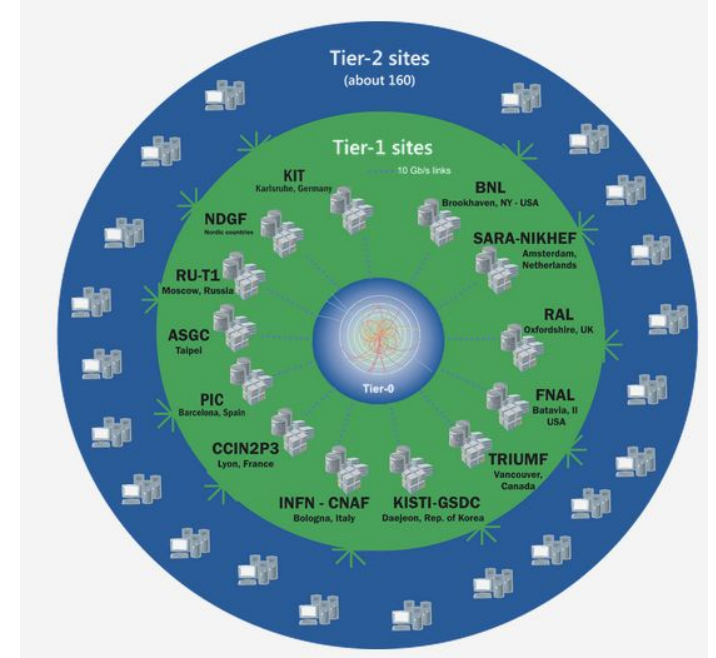
Project Stream Format
data18_13TeV.00349051.physics_Main.deriv.DAOD_EGAM1.f926_m1955_p3544
Run number Step AMI tag

MONTE CARLO

Project Physics description Format
mc16_13TeV.364250.Sherpa_222_NNPDF30NNLO_1111.deriv.DAOD_EGAM1.e5894_s3126_r9364_p3526
Dataset ID Step AMI tag

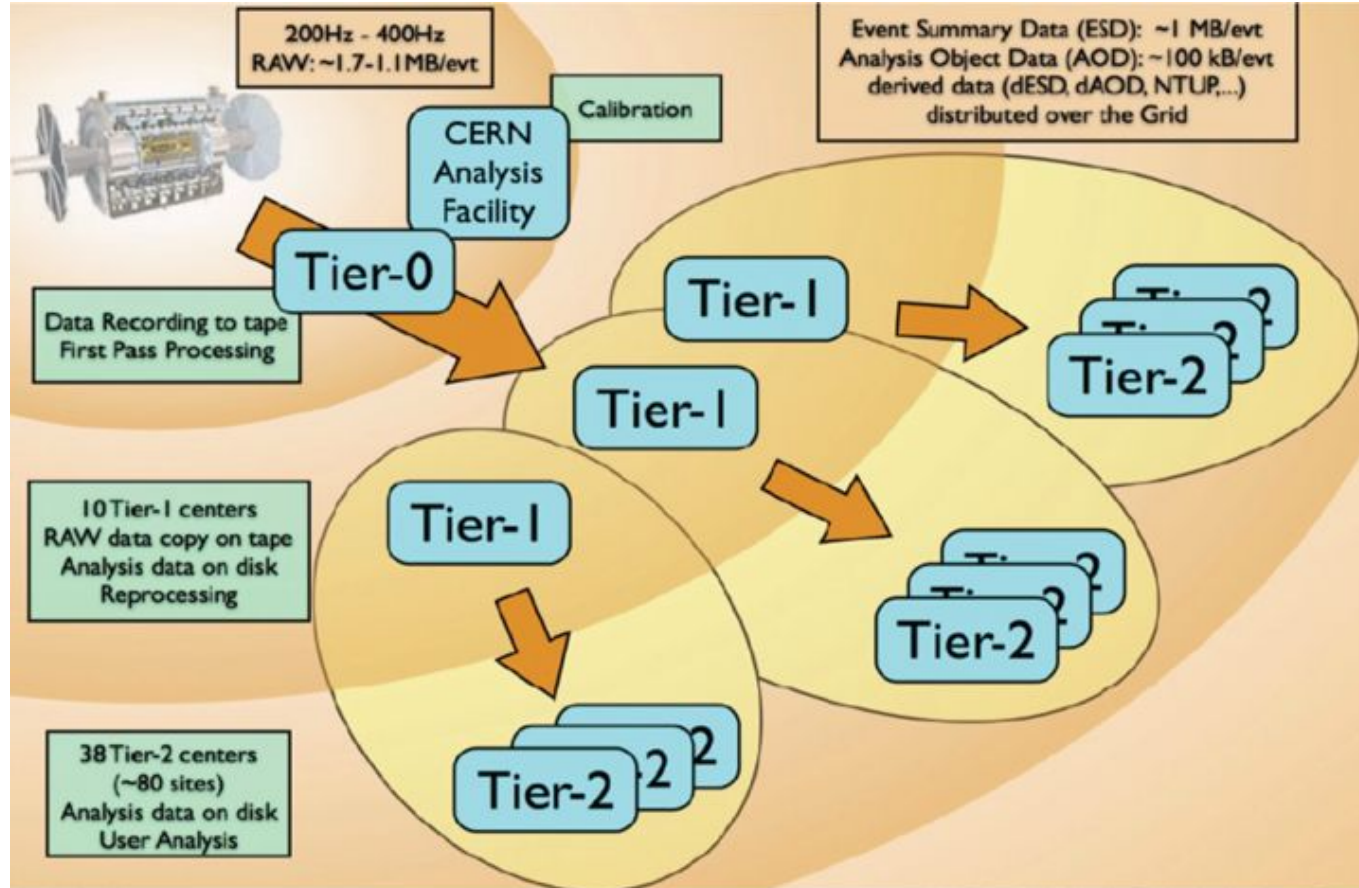
Veri Miktarı - Grid Sistemi

- LHC'de üretilen veri miktarı 50+ PetaByte/yıl
- Veri analizi yaklaşık 500 bin tane CPU (işlemci ünitesi) gerektiriyor
- Analizciler dünyanın birçok ülkesinde bulunuyor (ATLAS için 38 ülke)
- CERN ancak veri depolama ve işleme kapasitesinin % 20-30'unu karşılıyor
- Verilerin dağıtılması ve kullanıcıların kolaylıkla ulaşması gerekiyor -> GRID sistemi



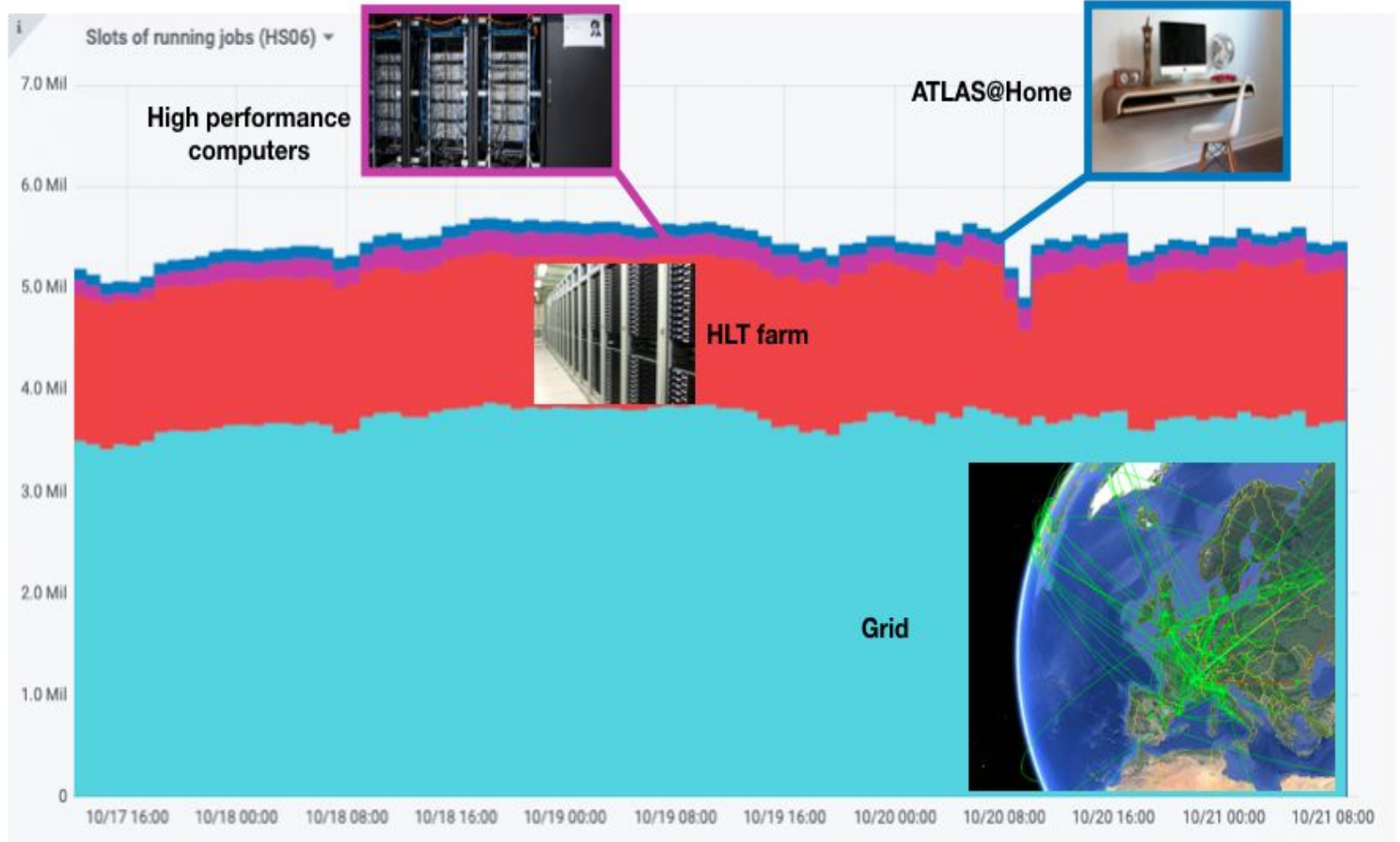
WLCG (Worldwide LHC Computing Grid) Sistemi, 160+ bilgisayar merkezi, 40+ ülke, 13 Tier-1 merkezi, 160 Tier2 merkezi

Veriler Nerede Tutuluyor? - Grid Sistemi



Veri İşleme Nerede Yapılıyor?

- Grid sistemi
- CERN (HLT farm)
- Yüksek performanslı bilgisayar merkezleri
- Evdeki bilgisayarlar



Türkiye Grid Sisteminde mi?

<https://wlcg-public.web.cern.ch/>

- Evet, Avrupa Grid sistemine bağlı, EGI (European Grid Infrastructure)
- Tier-2 merkezleri var
- ATLAS deneyi için TÜBİTAK-ULAKBİM'de. Hollanda Tier-1 merkezine bağlı
- CMS deneyi için Boğaziçi-ODTÜ-Çukurova konsorsiyomu

