MODIS Data Processing System

MODAPS

Curt Tilmes
NASA Goddard Space Flight Center
Code 922, Greenbelt MD 20771
Phone: +1-301-614-5534 Fax: +1-301-614-5269
Curt.Tilmes@gsfc.nasa.gov

Presented at the THIC Meeting at the Sheraton Barcelo in Annapolis MD 21401-3094 on May 9, 2001

THIC Inc. The Premier Advanced Recording Technology Forum

MODIS Data Processing System
MODAPS

Curt Tilmes
NASA Goddard Space Flight Center
Code 922, Greenbelt MD 20771
Phone: +1-301-614-5534 Fax: +1-301-614-5269
Curt.Tilmes@gsfc.nasa.gov

Presented at the THIC Meeting at the Sheraton Barcelo Annapolis Hotel
173 Jennifer Road, Annapolis MD 21401
on May 9, 2001

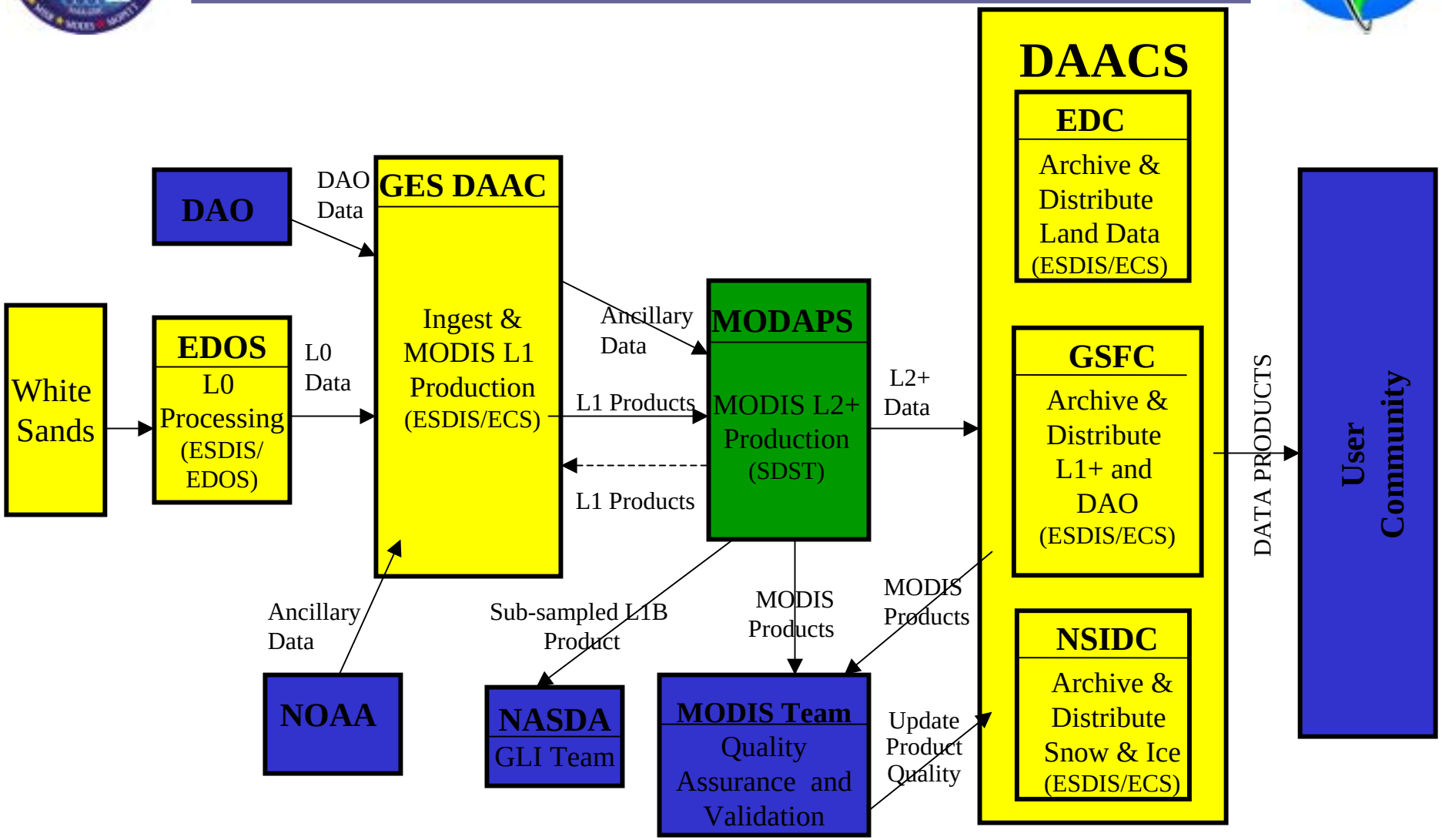
MODIS

- MODerate resolution Imaging Spectroradiometer
- Key instrument on EOS-Terra and EOS-Aqua Missions
- 36 spectral bands from the visible through thermal infrared
- Spatial resolution: 250m(2 bands), 500m(5 bands), 1km (29 bands)
- Nearly complete global coverage in 1 day, all of Earth in 2 days
- More than 40 earth science products including:
  - Sea Surface Temperature and Ocean Color, Cloud Properties and Aerosols, Land Surface Temperature, Vegetation Indices, Land Cover Change and Measures of Global Productivity

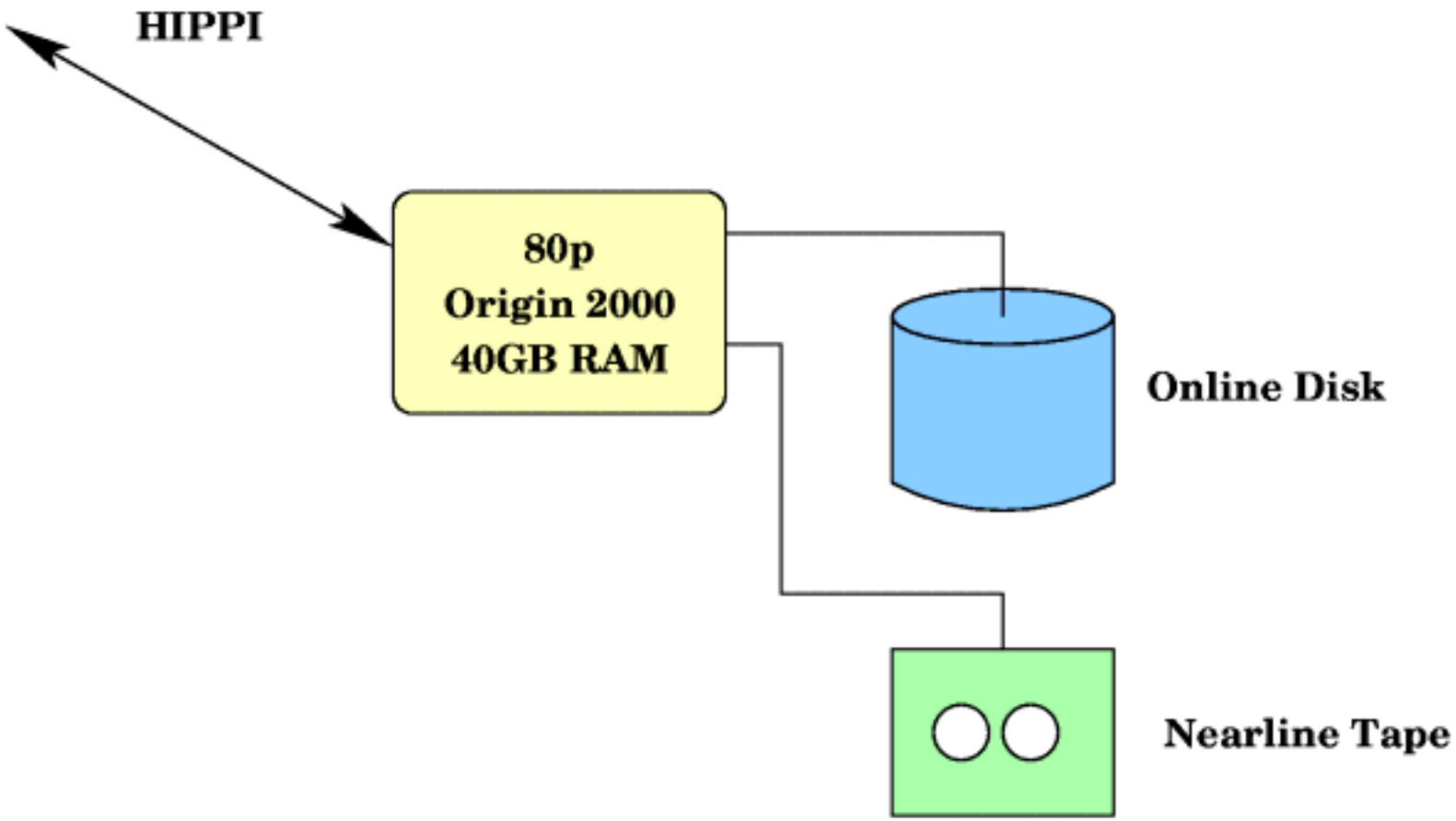
Product Volumes

- Each MODIS instrument produces **70GB** of raw data per day
- Goddard Earth Sciences Distributed Active Archive Center (GES DAAC) produces calibrated radiance, cloud mask and earth location products (**295GB/day**) and ships them to MODAPS
- MODAPS (MODIS Data Production System) makes the MODIS higher level science products (**400GB/day**) and ships them to the MODIS Science Team and to 3 DAACs for public distribution
- In 2003, the MODIS Team has a goal of producing up to **8TB** of products per day and shipping **15TB** per day to the community
  - Current requirements are 3TB/day produced and 3TB shipped
  - USGS EROS Data Center ships 55GB of Landsat 7 ETM+ scenes per day

Data Product Flow

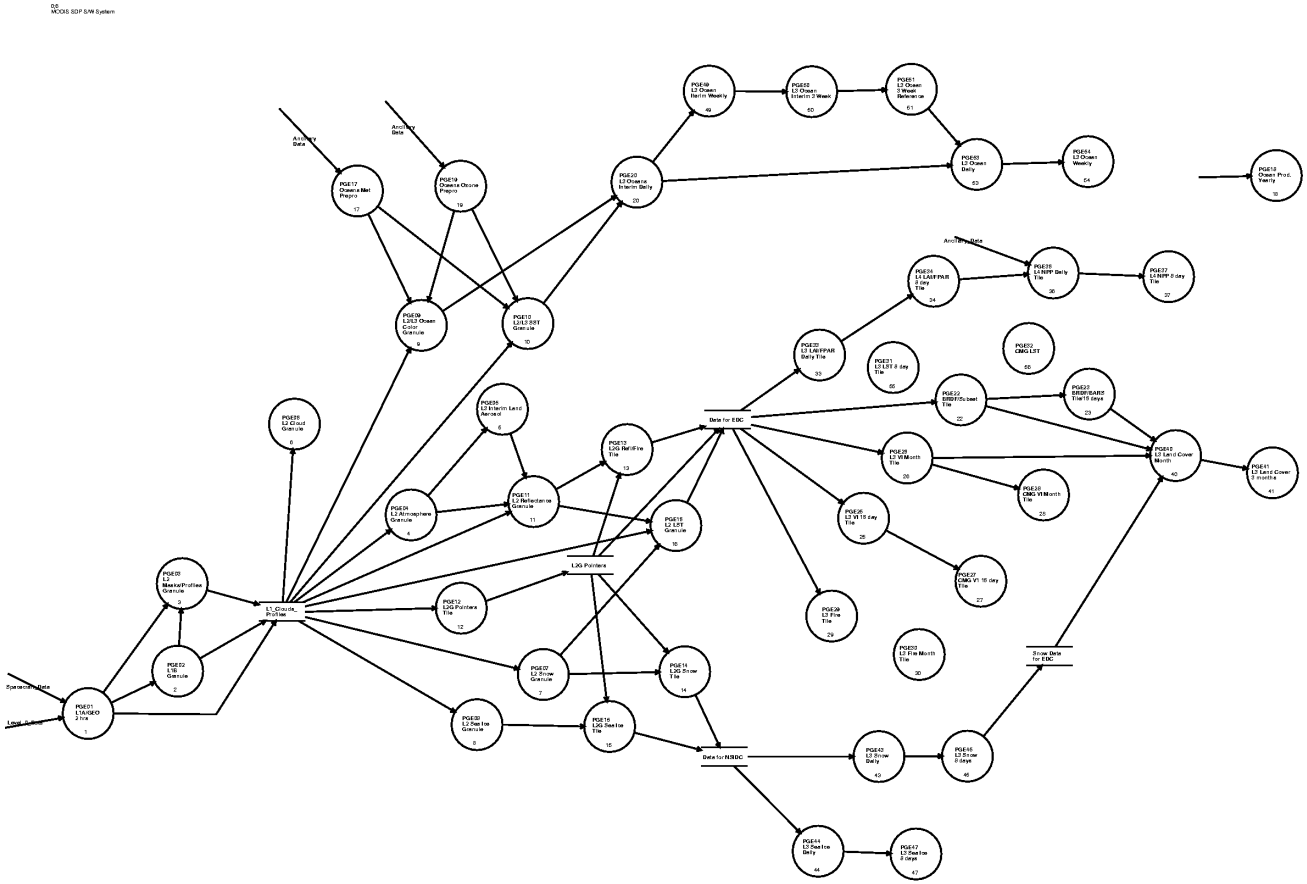


At Launch System

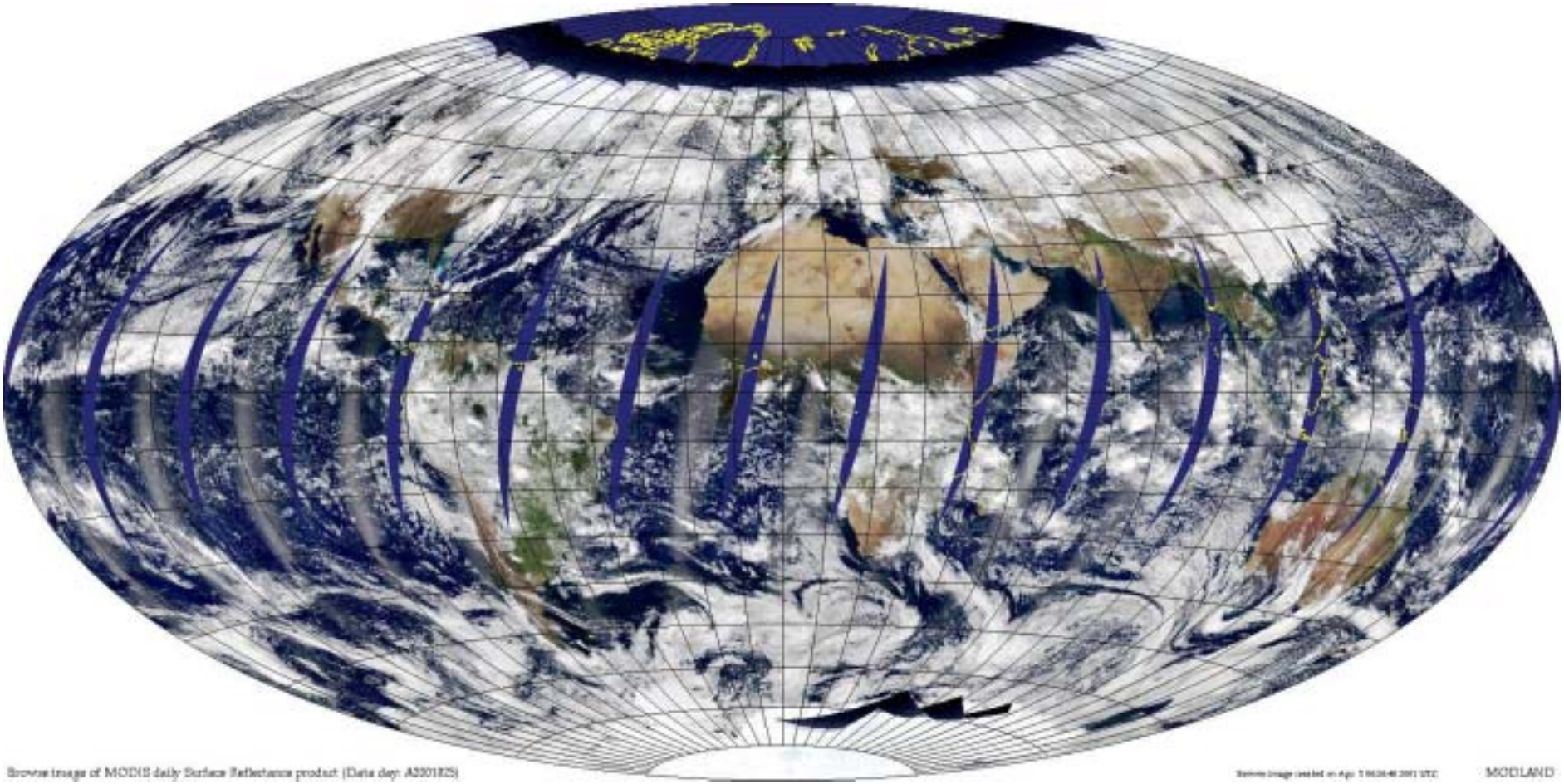


Process Flow

MODIS SDP S/W System



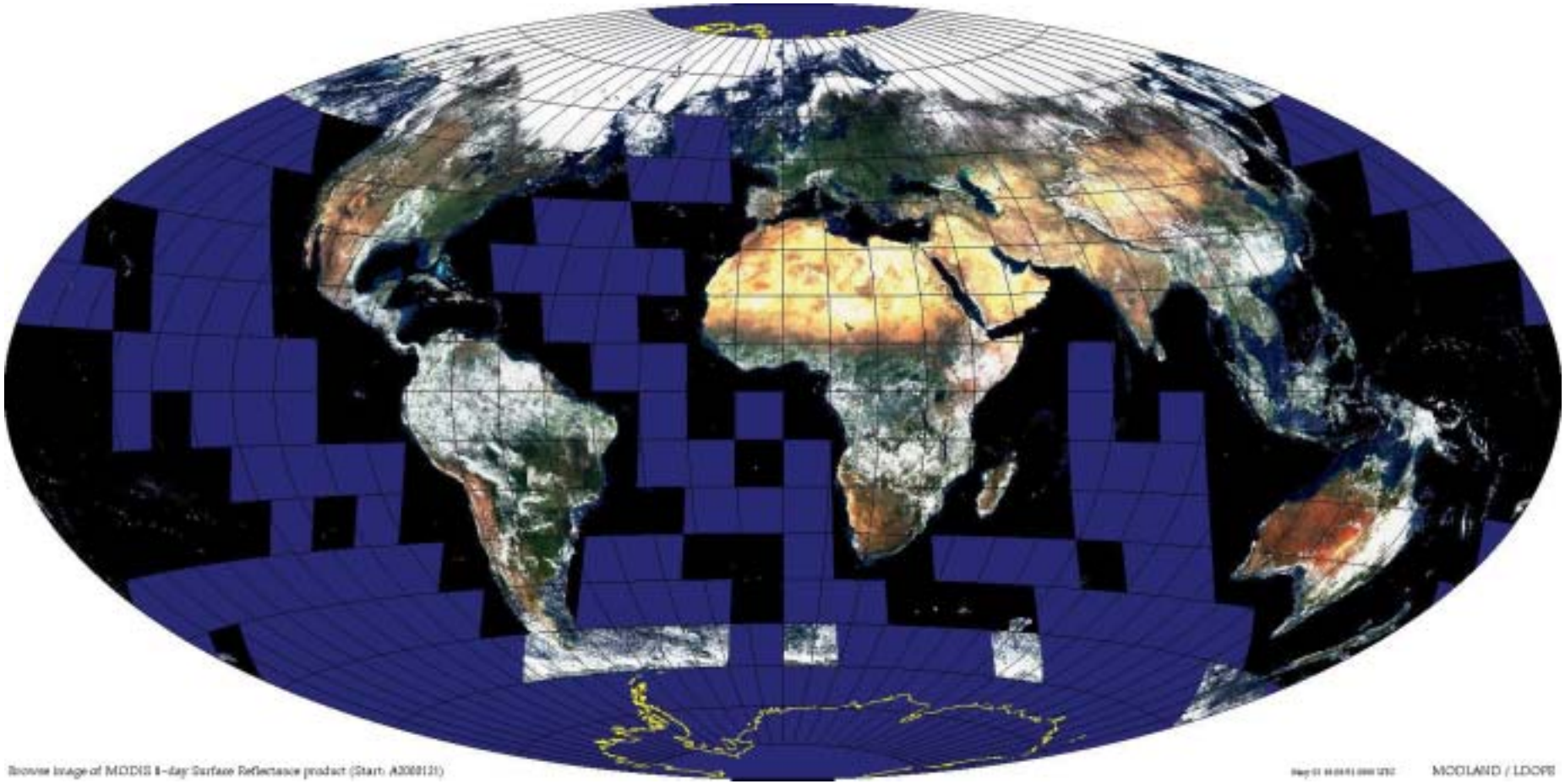
Level 2


Browse image of MODIS daily Surface Reflectance product (Data day: A2001125)
MODLAND

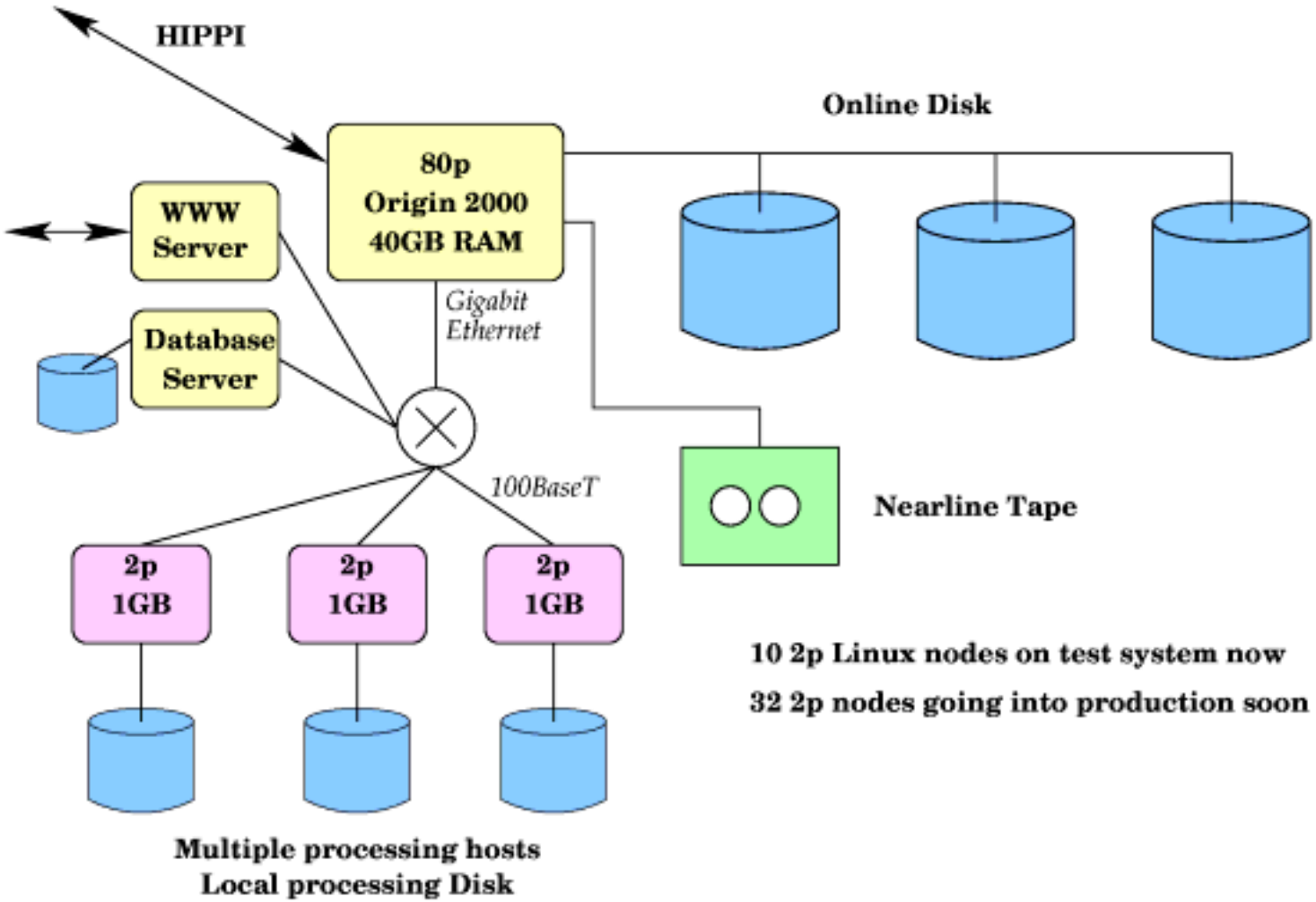
Level 2 Production

- 7 Level 2 PGEs
- Input 5 minute granules, output 5 minute granules
- 288 x 5 minute granules per day
- 144 day-mode, 144 night-mode executions
- Input ~730MB per 5 minutes
- Output ~580MB day-mode and 175MB night mode
- Around an hour of processing time for 5 minutes of data
- Very suited to distributed processing

Land Gridded Production



Distributed Architecture



Distributed Architecture

- Much cheaper per processor cost for 2p nodes than central node
- Local disk space on processing nodes
- Copy input files down through Multiple Gigabit Ethernet to switched 100BaseT
  - ~10MB/s, rcp in less than 2 minutes
- Read input files once on central node, then read repeatedly locally
- Process locally for a while, then copy outputs back to central node.
- 2GB RAM per node, caching benefit

Online Disk

- At launch 10TB of LSI Logic Metastor (1 Filesystem!)
- Added Data Direct San Data Director
- Connected through 8 Host Fibre Channel Interfaces
- 20 Drive Fibre Channel Interfaces
- Dual 8+1 RAID, plus 2 spares channels
- Currently 35TB on primary host
- Soon another 28TB on reprocessing host
- ~70TB total online
- Split into 1TB per filesystem.
- Database tracks location of each file.

Nearline Tape

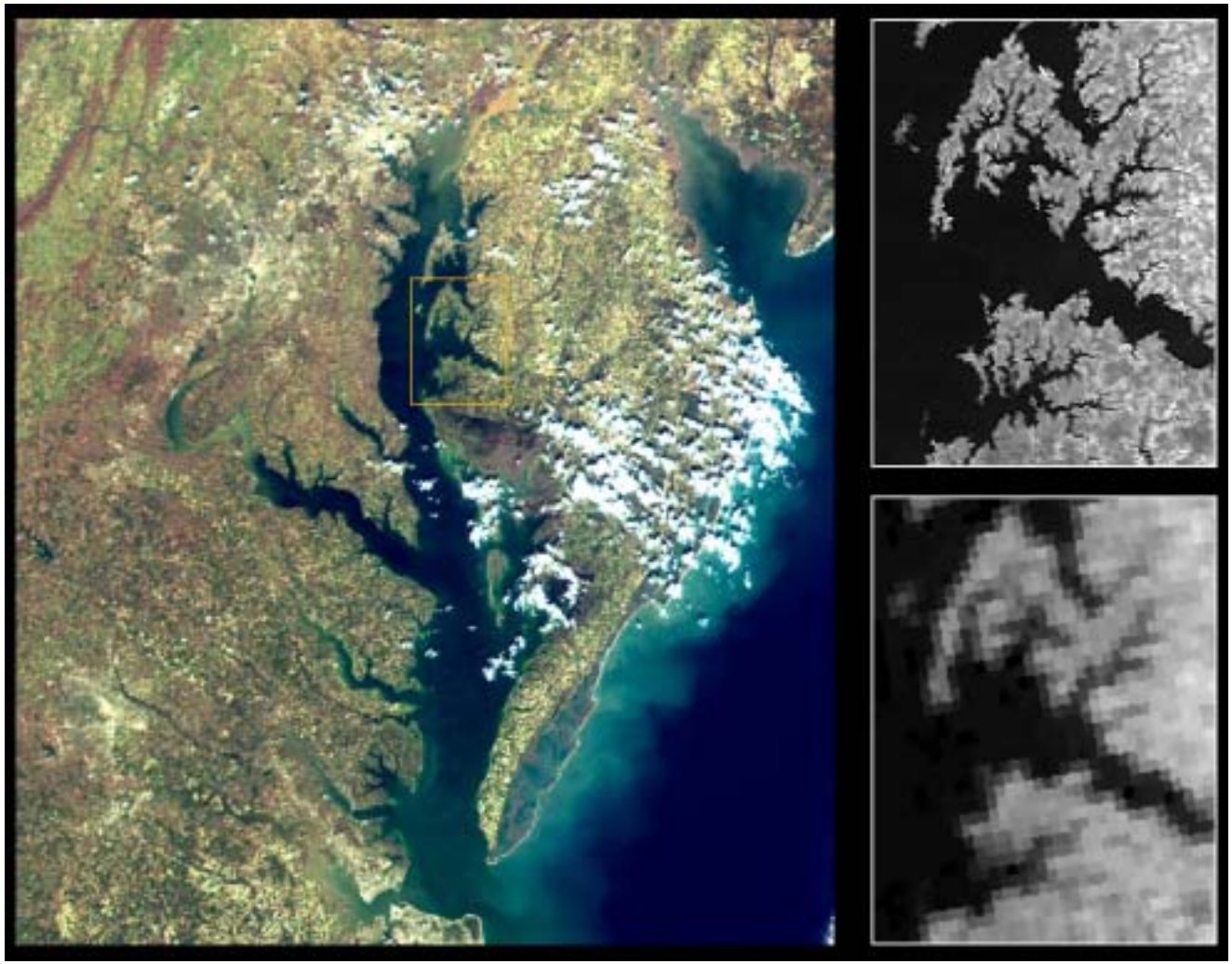
- ADIC Scalar 1000 with 12 AIT-2 drives on both primary and reprocessing system. 1182 tapes per jukebox.
- Connected via Chapparal Fibre Channel to SCSI bridges
- 50GB native, ~100GB with 2/1 compression
- 6MB/s native, we see 5-10MB/s
- About 2TB/day transfer to/from tape possible
- Retrieves take time mounting tapes, searching
- Not enough – adding another 12 drives to each jukebox.
- Driven by homegrown database driven scripts calling Legato Networker Archive Option

Summary

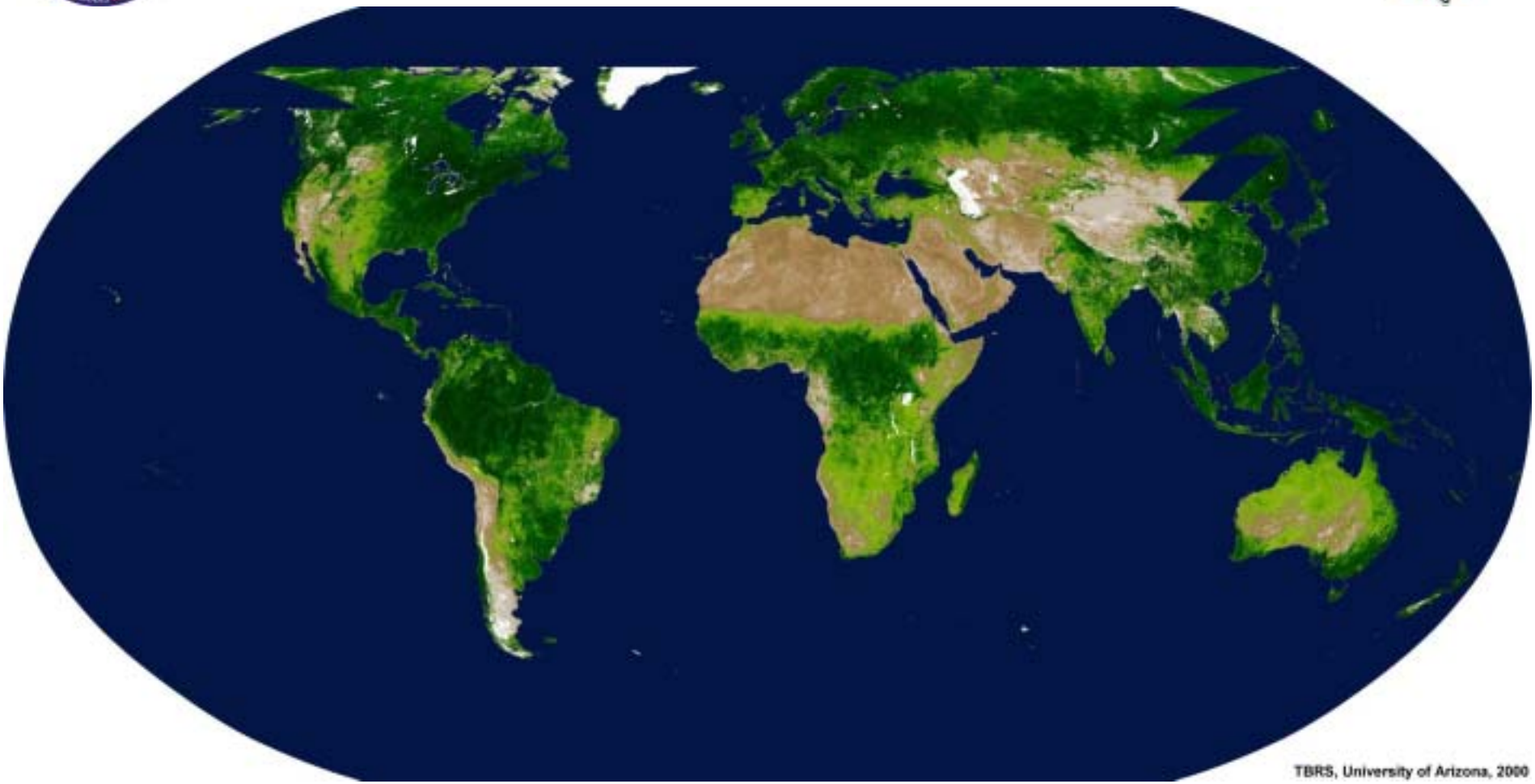
- With Aqua launching soon, and reprocessing needs growing, we need to continue scaling the whole system.
- Commodity Linux processors have good bang/$
- All Level 2 and 2G processes ported to Linux/Intel
- HIPPI link to DAACs ~75MB/s = 6TB/day per system, Two systems (processing and reprocessing).
- DAACs have a big job scaling enough to meet long term archive and distribution needs
- http://modis.gsfc.nasa.gov

MODIS Sample Images

250M vs. 1KM



16day 500M NDVI Composite



Fire Detection