Multi-Instrument Intercalibration Framework (MIIC) Update

CLARREO Science Team Meeting
Jan. 7-9, 2014

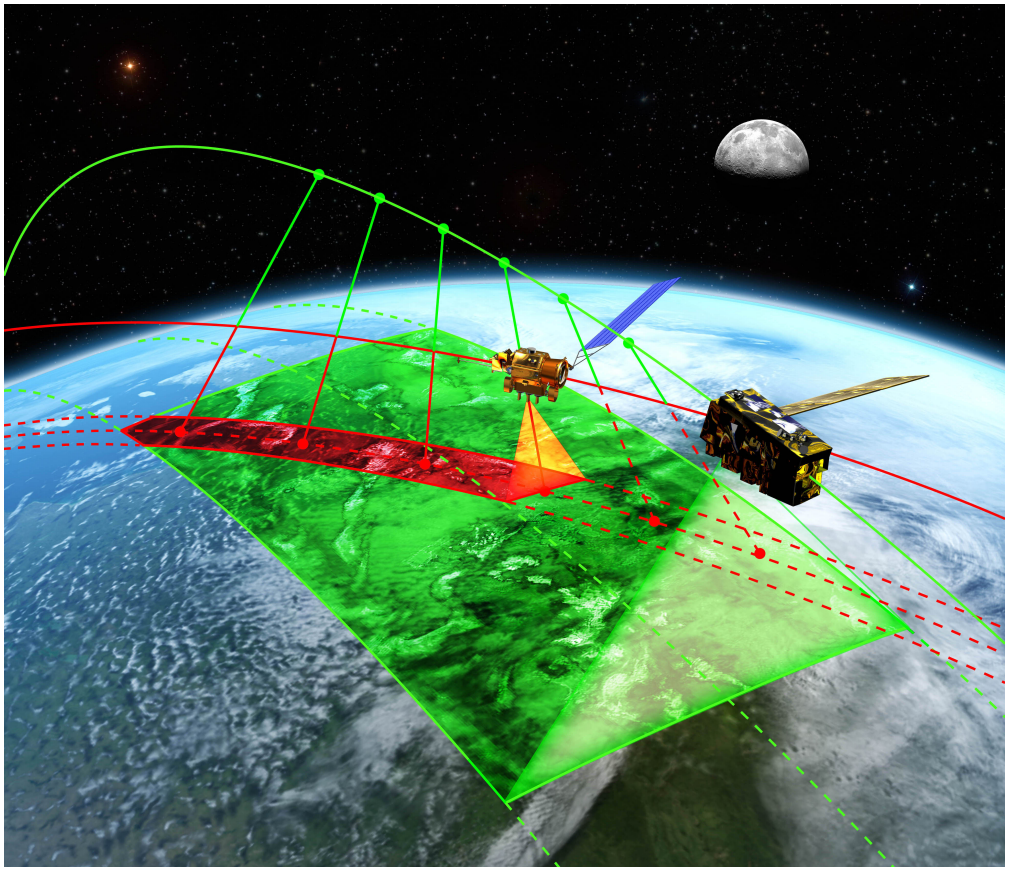
Chris Currey[1], Constantine Lukashin[1], Dave Doelling[1],
Carlos Roithmayr[1], Aron Bartle[2], Thomas Thompson[2],
Dave Johnson[4], James Gallagher[3]

(1) NASA, (2) Mechdyne, (3) OPeNDAP Inc., (4) SSAI

What is the MIIC Framework?

- Collection of software designed to work in a *distributed collaborative environment* to support LEO-GEO and LEO-LEO intercalibration
- Provides event prediction, remote data acquisition, and analysis *web services*

How MIIC Benefits CLARREO?



The MIIC software framework provides efficient access to distributed data for intercalibration – finds and acquires matched samples for instruments on separate spacecraft

MIIC ROSES-2011 ACCESS Proposal Objectives

The primary objective of the Advancing Collaborative Connections for Earth System Science (ACCESS) program is to **enhance, extend, and improve existing components of NASA's distributed and heterogeneous data and information systems infrastructure**. ACCESS projects increase the interconnectedness and reuse of key information technology software and techniques underpinning the advancement of Earth science research.

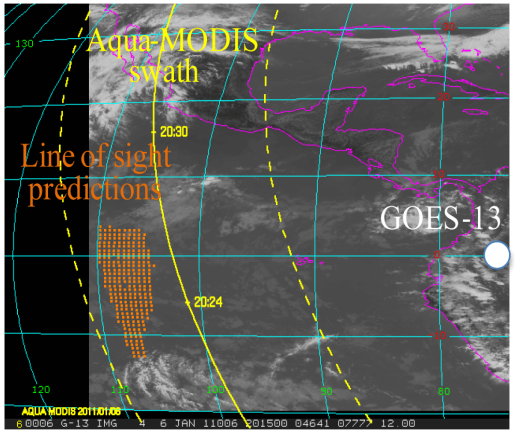
Specific MIIC Framework objectives:

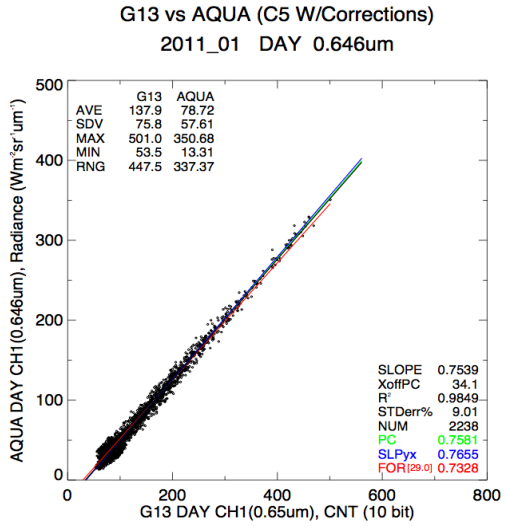
- **Develop of a common framework to improve access to co-located measurements with matched viewing geometries from multiple instruments to advance the intercalibration of current sensors**
- **Advance collaborative connections between NASA's Earth science production systems, instrument Cal-Val data systems, and Distributed Active Archive Center (DAAC) online archives**
- Improve access to multi-instrument multi-satellite datasets contained within the LaRC ASDC online depot
- Support the international intercalibration goals of WMO **GSICS** and CEOS WGCV
- Support NPOESS Preparatory Project (NPP) and Joint Polar Satellite System (JPSS) CERES calibration and validation activities
- Support future missions developed through the NASA's Earth Systematic Missions Office and existing NASA Earth science data systems with a distributed processing infrastructure
- Improve the calibration accuracy and consistency of GEO and LEO operational instrument data products

The **goal of this proposal** is to demonstrate an extendable distributed processing framework that allows instrument teams to **perform significantly more processing at data source centers versus transmitting large volumes of data to a science data system for processing**.

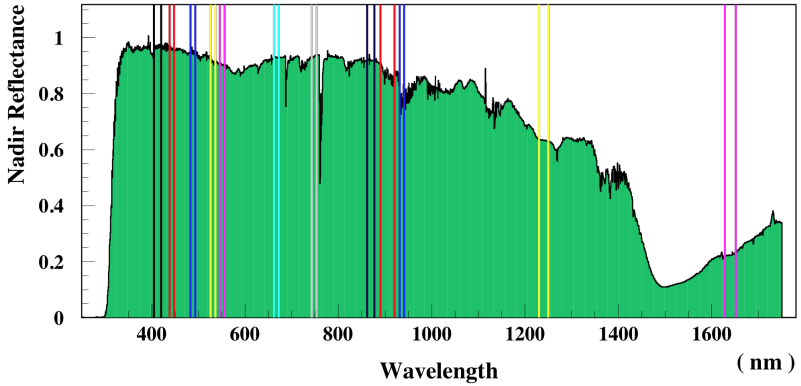
Proposal #1 - Demonstrate **feasibility and benefit of intercalibration services** in DAAC environment
Proposal #2 - Deployment in NASA ASDC and NOAA NCDC facilities

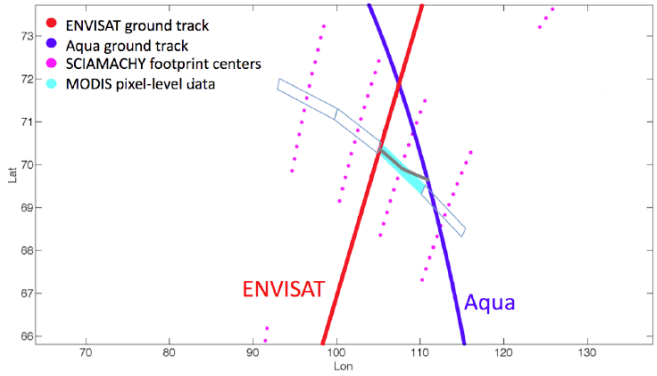
Requirements - Reflected Solar IC Use Cases





Build 1 LEO-GEO Reflected Solar IC Use Case: GOES-13 and Aqua MODIS; 0.5° equal angle grid cells, client and server-side filtering (*Doelling et al.*)

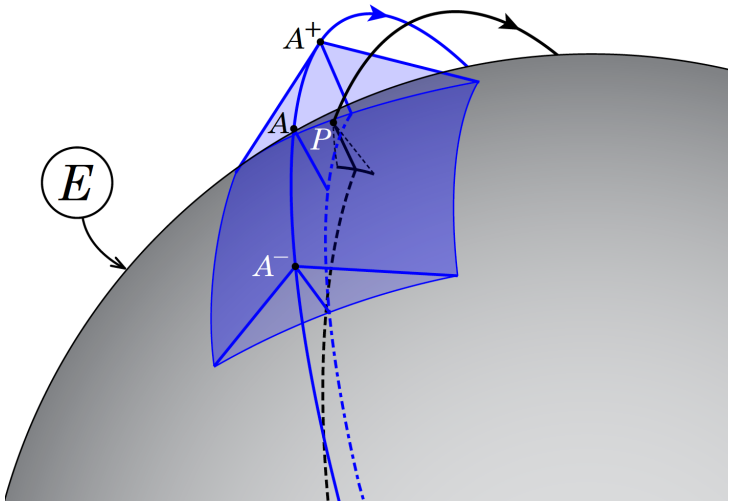


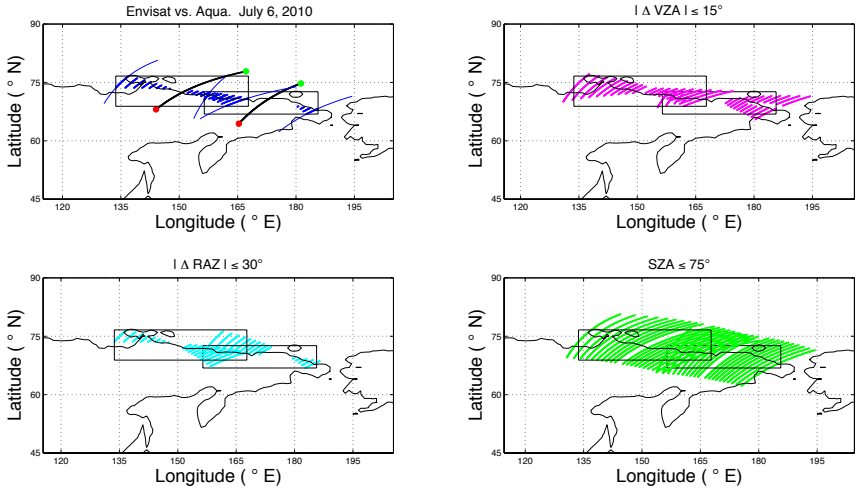


Build 2 LEO-LEO Reflected Solar IC Use Case: Envisat SCIAMACHY and Aqua MODIS; spectrally convolve SCIAMACHY data with MODIS RSRs and spatially convolve MODIS 1 km pixels w/ SCIAMACHY iFOVs (*Doelling et al.*)

Build 2 LEO-LEO Event Prediction

(*Roithmayr*)

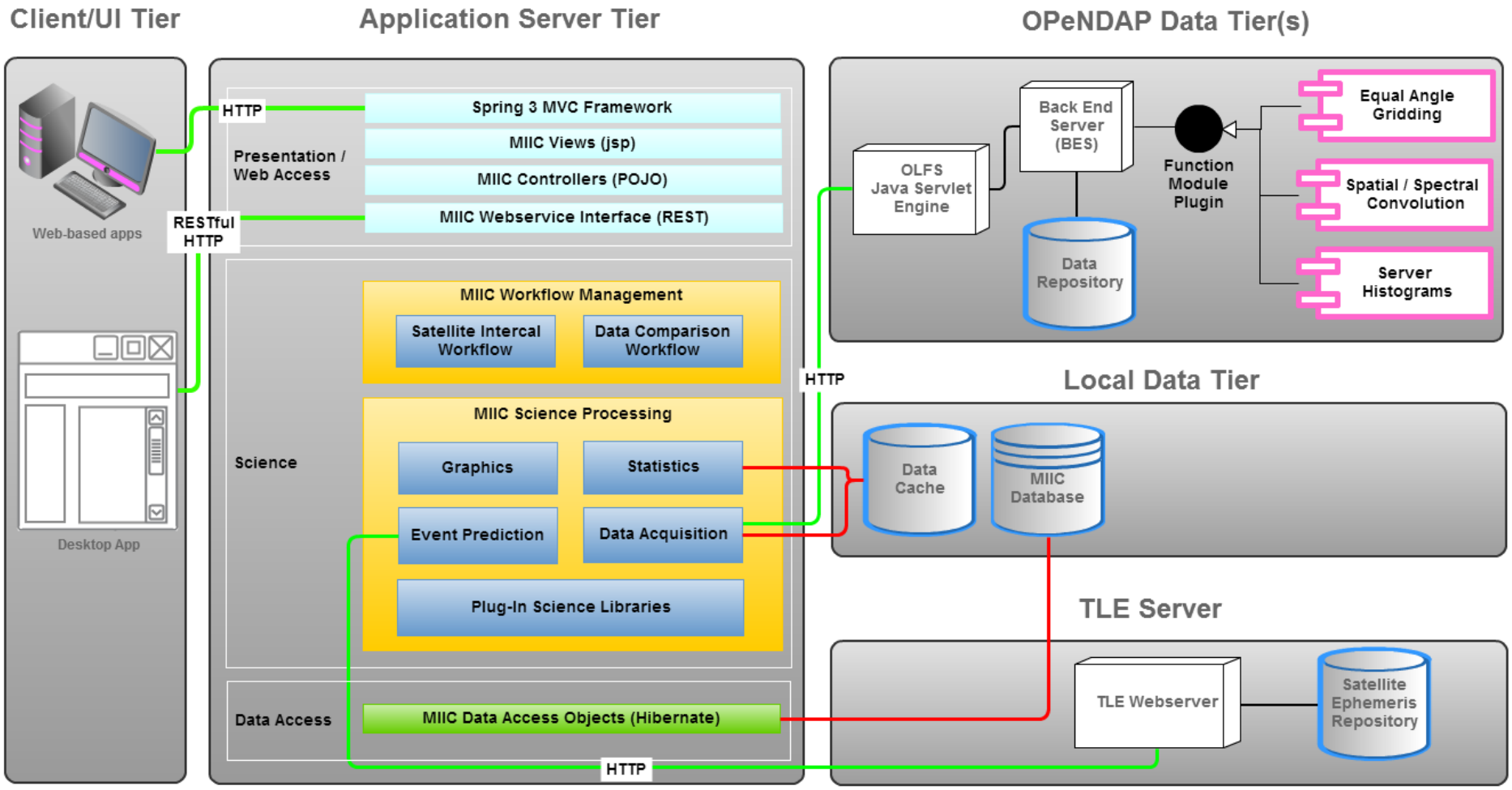




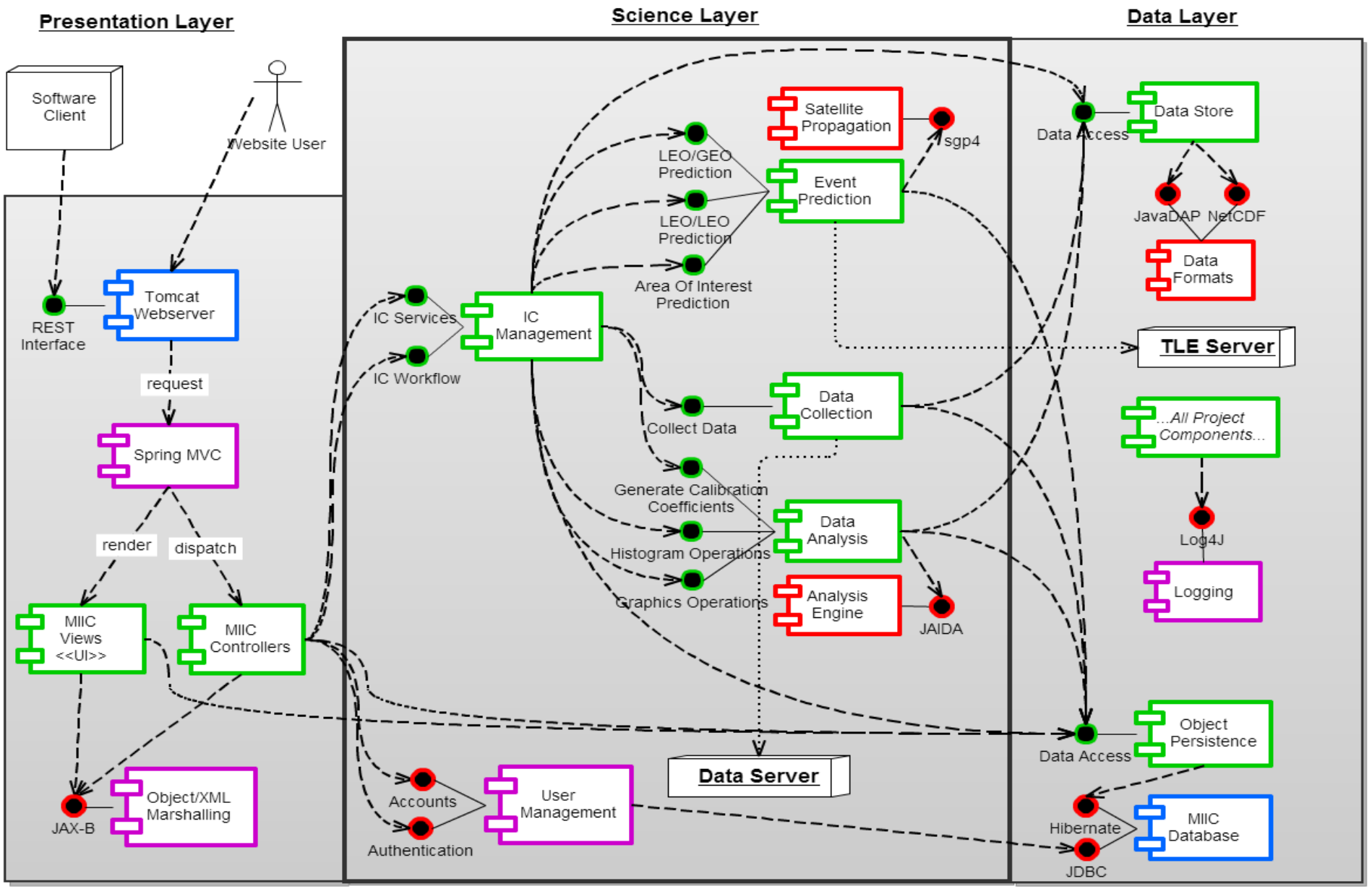
LEO-LEO Event Prediction; rectangle defines lat-lon box that meets all criteria (top left), SZA ≤ 75°, |Δ RAZ| ≤ 30°, |Δ VZA| ≤ 15°, |Δ T| ≤ 2.5 min. Matching Aqua MODIS visible data with Envisat SCIAMACHY.

The LEO-LEO Event Prediction algorithm uses orbital modeling to find matches in viewing geometry for measurements taken from instruments on separate spacecraft. We have modified the CLARREO algorithm to find matches on existing LEO scanners. First create a tent structure that moves with one spacecraft, then predict the view conditions when the other instrument is inside the tent. *C. Roithmayr et al., CLARREO Approach for Reference Inter-Calibration of Reflected Solar Sensors: On-Orbit Data Matching and Sampling, IEEE Tran. Geo. Rem. Sensing, 2013.*

MIIC N-Tier Architecture



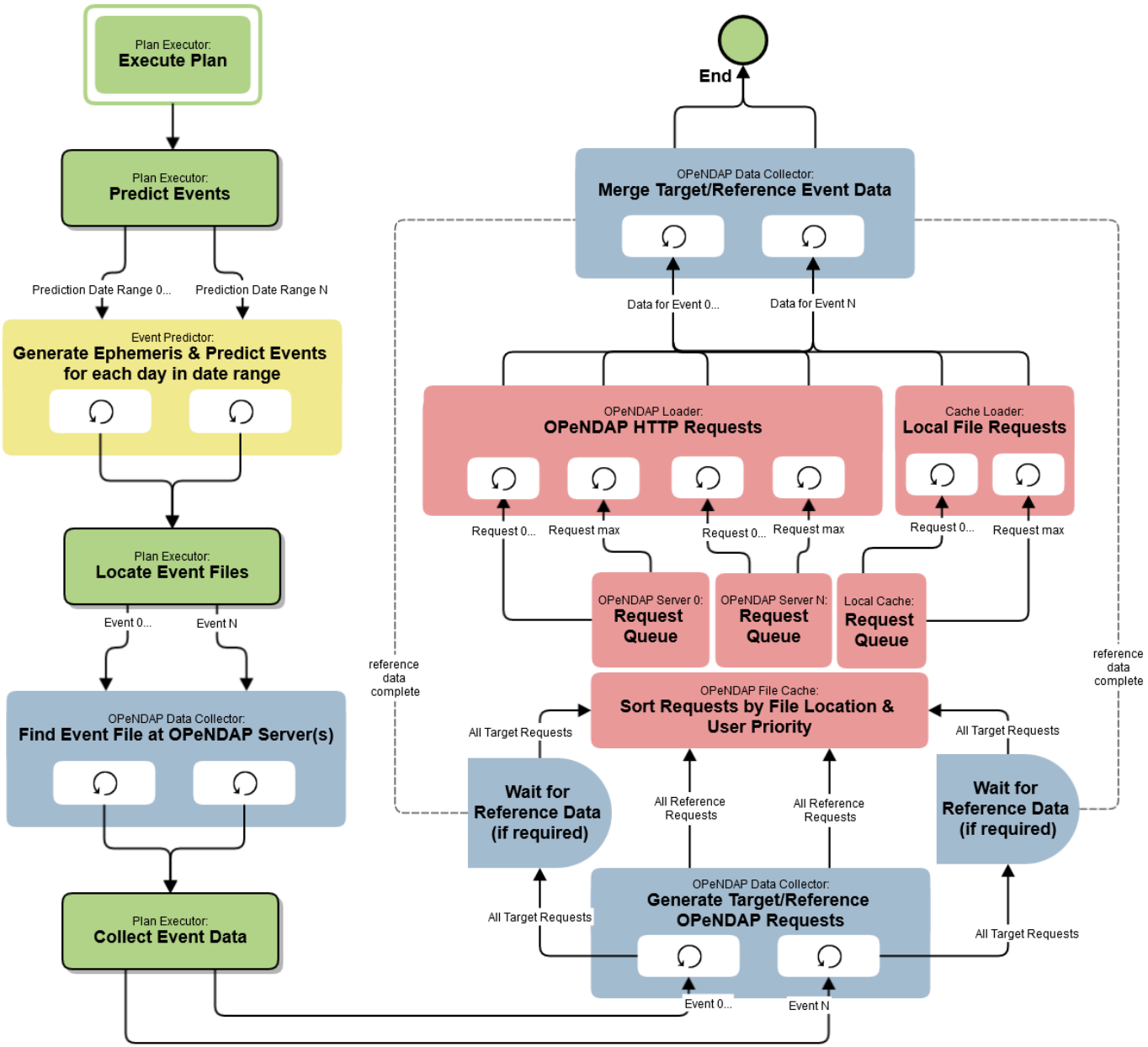
Application Tier Component Architecture



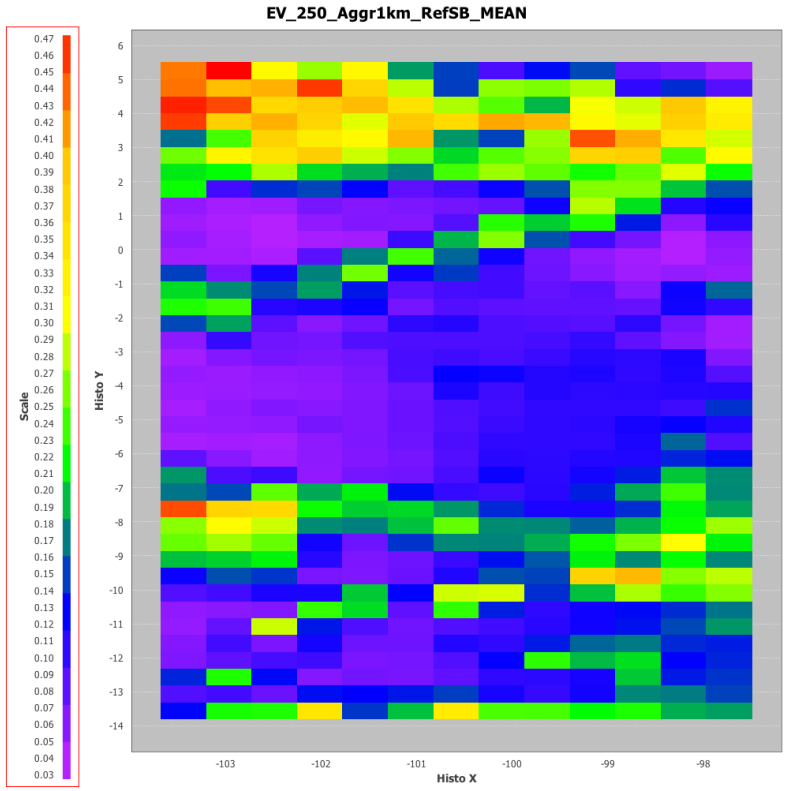
MIIC Framework Services Implemented

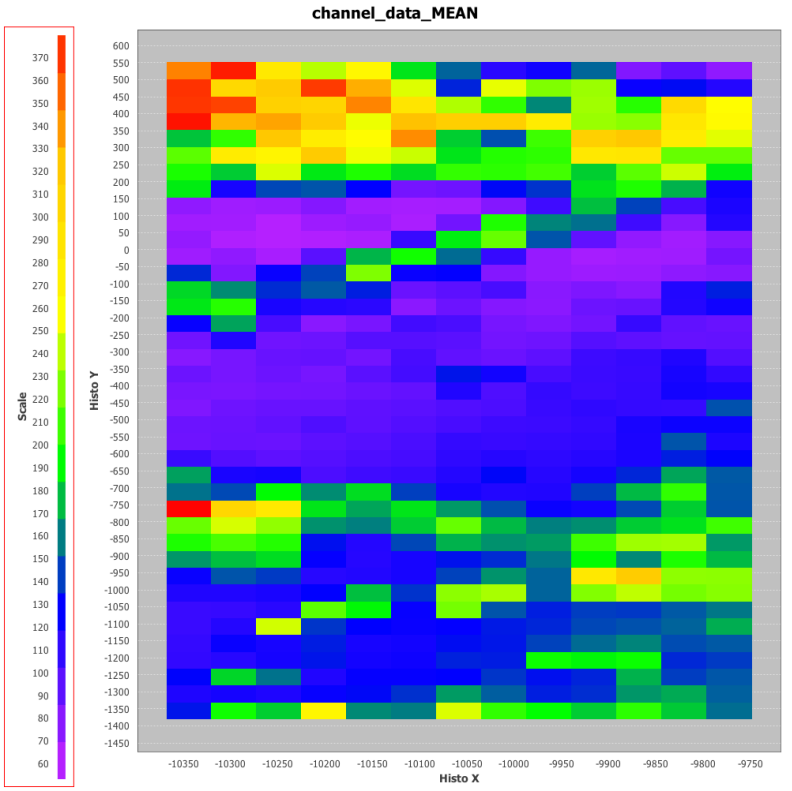
- Event Prediction
  - Users define ICPlans (time period and instruments)
  - MIIC client calculates Lat/Lon extents and times for matched data
  - Support both LEO and GEO satellite instruments if spacecraft TLEs available and spacecraft altitudes different
  - Using CLARREO-like IC sampling algorithms; parameters tunable
- Remote Data Access
  - Efficiently access data from remote data centers using OPeNDAP
  - Only access files that contain data in Event boundaries
  - Execute user developed functions, eg., spatial convolution, gridding, histograms, within remote OPeNDAP servers
  - Works with data products in hdf or netCDF
  - Deployed on new ASDC web server and offsite at Mechdyne
- Regression Analysis

Parallelism within the Application Tier



2DHistogram Server-side Function Processed Event





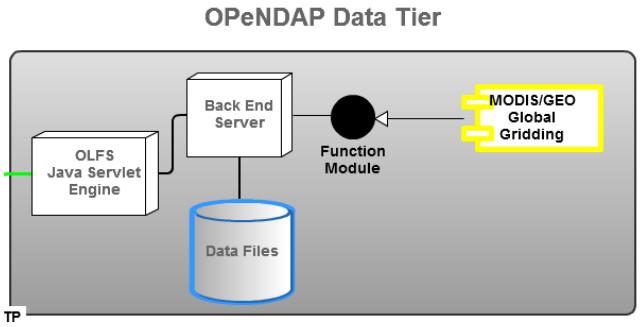
Event (State: DATA_LOCATED)
from: 2011.01.01 AD at 19:53:10 GMT
to: 2011.01.01 AD at 20:18:10 GMT
box: -13.788960361722237,5.507211389678214,-103.67675765569243,-97.49095771216804
ref files: [http://projects.mechdyne.com:8080/opendap/hyrax/MYD021KM.A2011001.2015.005.2011002204747.hdf,
http://projects.mechdyne.com:8080/opendap/hyrax/MYD021KM.A2011001.2005.005.2011002205625.hdf,
http://projects.mechdyne.com:8080/opendap/hyrax/MYD021KM.A2011001.1950.005.2011002203851.hdf,
http://projects.mechdyne.com:8080/opendap/hyrax/MYD021KM.A2011001.2000.005.2011002205114.hdf,
http://projects.mechdyne.com:8080/opendap/hyrax/MYD021KM.A2011001.1955.005.2011002204824.hdf,
http://projects.mechdyne.com:8080/opendap/hyrax/MYD021KM.A2011001.2010.005.2011002204113.hdf]
tgt files: [http://projects.mechdyne.com:8080/opendap/hyrax/MCIDAS.G-13.2011.01.01.1945.08K.bin]

MIIC Design Features

- Operates in a distributed collaborative enterprise environment (secure, scalable, reliable)
- Software is extendable
- Provides access to L1 & L2 data
  - Build 1: MODIS L1B (hdfeos), GOES-13 (bin -> nc)
  - Build 2: MODIS L1B (hdfeos), SCIAMACHY (bin -> nc)
  - Lesson learned: stick with hdf or netCDF products!
  - New products: SCARAB Flux (h5), CERES SSF (hdf), NPP VIIRS Subset (hdf), GEO products from CLASS (nc), MODIS L2 (hdfeos)
- Web services follow the MVC design pattern (Spring Framework)
- Use open source software
  - SGP4, JFreeChart, Apache Commons, JavaDAP,
  - Hibernate ORM + Postgres database
- Deploy on servers with access to large data repositories (ASDC)
- Developers can extend code on client-side (Java) or server-side (C++)

OPeNDAP Data Tier: MIIC Server-side functions

Server-side functions are called using HTTP commands – MIIC client builds OPeNDAP constraint expressions



http://…/MCIDAS.G-13.2011.01.01.2145.08K.bin.dods ← data URL
?eageogrid(.5, ← equal angle GEO grid, ½ degree
GRID_VAR,TIME, ← scan time included in grid
GRID_VAR, latitude, GRID_VAR, longitude,
GRID_VAR, channel_data,"0", -99, ← channel 0, -99 is missing value
FILTER_VAR, latitude,-35,10, ← include only lat from -35 to 10
FILTER_VAR, longitude, -133,-104,
MIN_GRID_PTS,100) ← min number of points per grid cell or fill

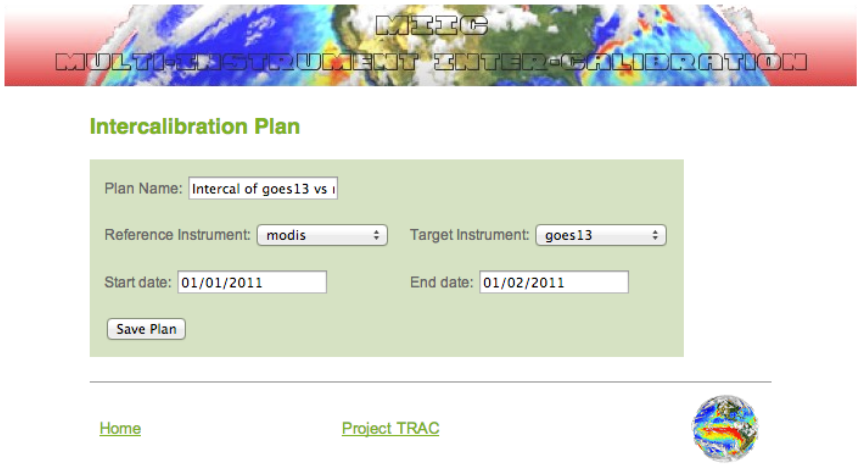
Output: Structure containing MEAN, STD, COUNT for each Grid variable[#lon][#lat]

MIIC Server-side Functions (SSFs)

- Developed Equal Angle Gridding, Spatial and Spectral Convolution, 1D and 2D Histograms SSFs
- Newest 2DHistogram function provides generic access to parameters and will replace earlier eagridding()
- Use OPeNDAP C++ library of server-side classes (libdap) to access variables (DDS, DAS, DataDDS) from low level handlers (hdf, netcdf)
- Complex development environment!
- Received support from OPeNDAP and The HDFGroup
  - Added HTTP POST to allow clients to specify larger constraint expressions to include large Arrays (eg., RSR values for spectral convolution)
  - Users provide functions in separate linkable modules (eg., MIICmodule)
  - Multiple mods to HDF4 handler (memory leaks, L1B MODIS QC flags)
- Biggest issue – memory leaks within nested server-side functions (histo2D) and OPeNDAP HDF4 Handler
- Future – extend our DAP2 SSF “subsetting/filtering” or use future (unfunded) DAP4 capabilities for single instrument use cases?

Presentation/User Interface Layer

JavaServer Pages



RESTful API

http://projects.mechdyne.com:8080/opendap/hyrax/july2010/MODIS/MYD021KM.A2010183.0215.005.2010183172615.hdf.ascii?histogram_2d_avg(x_axis(Longitude,100.,160.,60),y_axis(Latitude,68.,78.,20),define_var(fixed_dim(EV_1KM_RefSB,"Band_1KM_RefSB",0)),"exclude_flow")

http://projects.mechdyne.com:8080:/miic/icplan/create

- Simple template web pages delivered with software
- Need to finalize web pages during deployment
- RESTful interface provides client programs access to MIIC services
- A. Bartle's fall CS class, William & Mary, class projects built MVC controller and web pages for MIIC (5 teams)

Summary

- **Met objectives of MIIC #1 ROSES-2011 proposal**
- **MIIC Framework software will significantly reduce CLARREO science data system development costs**
- Numerous opportunities for collaboration
  - IRAD 2014 Data Fusion Engine (Lukashin)
  - ECHO EVI (Pilewskie)
  - MIIC follow-on proposal ROSES-2013 with NOAA NCDC
  - CERES Order/Visualization Tool (focus L1 & L2 data)
  - CLARREO IR intercalibration algorithms?
- Distributed computing environments are complex, requires software staff familiar with the newer programming languages and development tools
- Invited to present at the GSICS Annual meeting, Mar. 2014
- Need to resolve issues with code delivery to GSFC ESDS
  - How to handle future updates to code?
  - Need to verify software licensing for distribution