Atmospheric Science Data Center Update

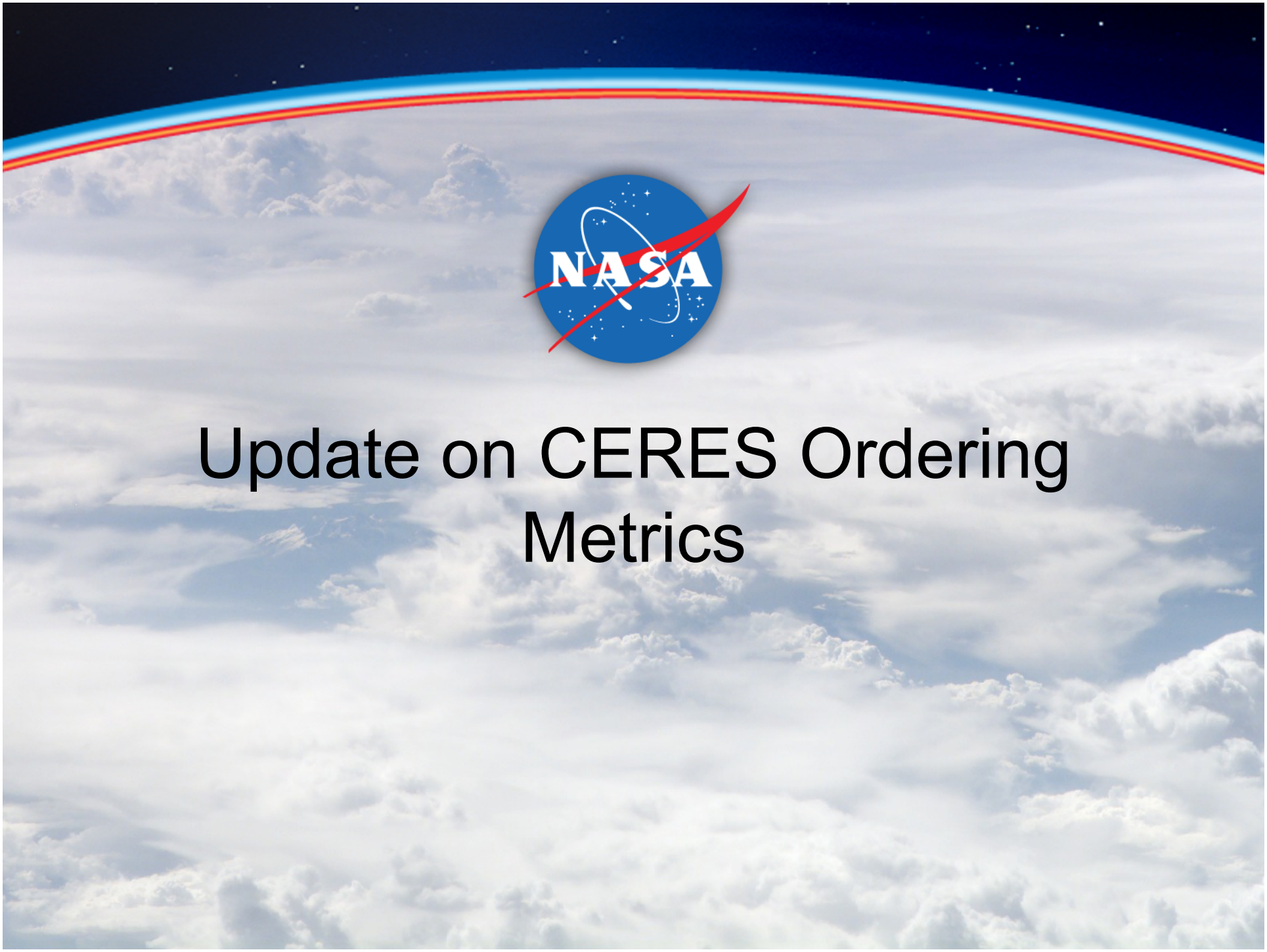
CERES Science Team Meeting
26 April 2011

John M Kusterer

Head, ASDC

Agenda

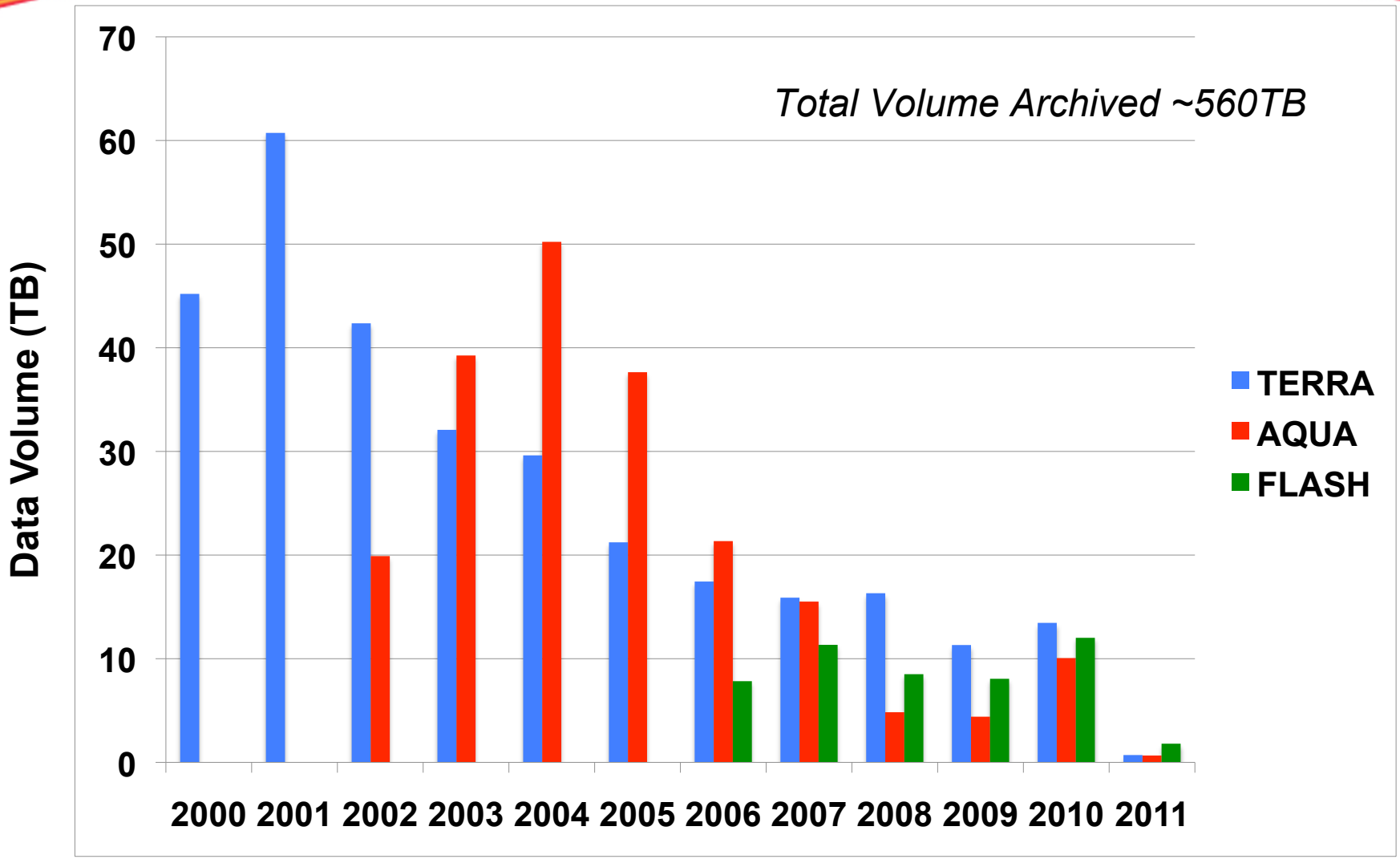
- Update on CERES Ordering Metrics
- Status of CERES Production Environment
- CERES Data Access
  - Subsetting
  - Web pages
  - Future vision of data access



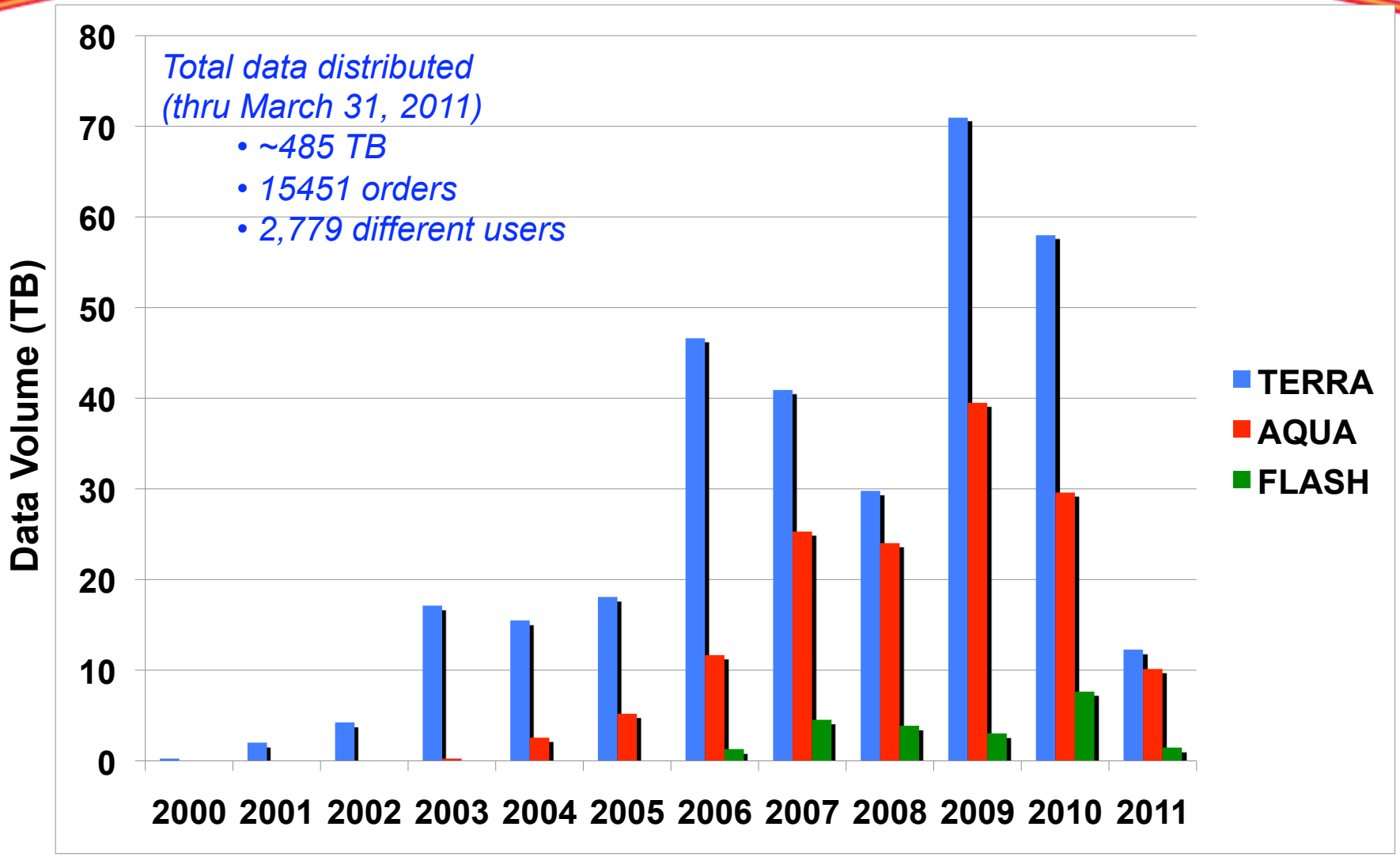
Update on CERES Ordering Metrics

CERES and FLASHFlux Archive Volume

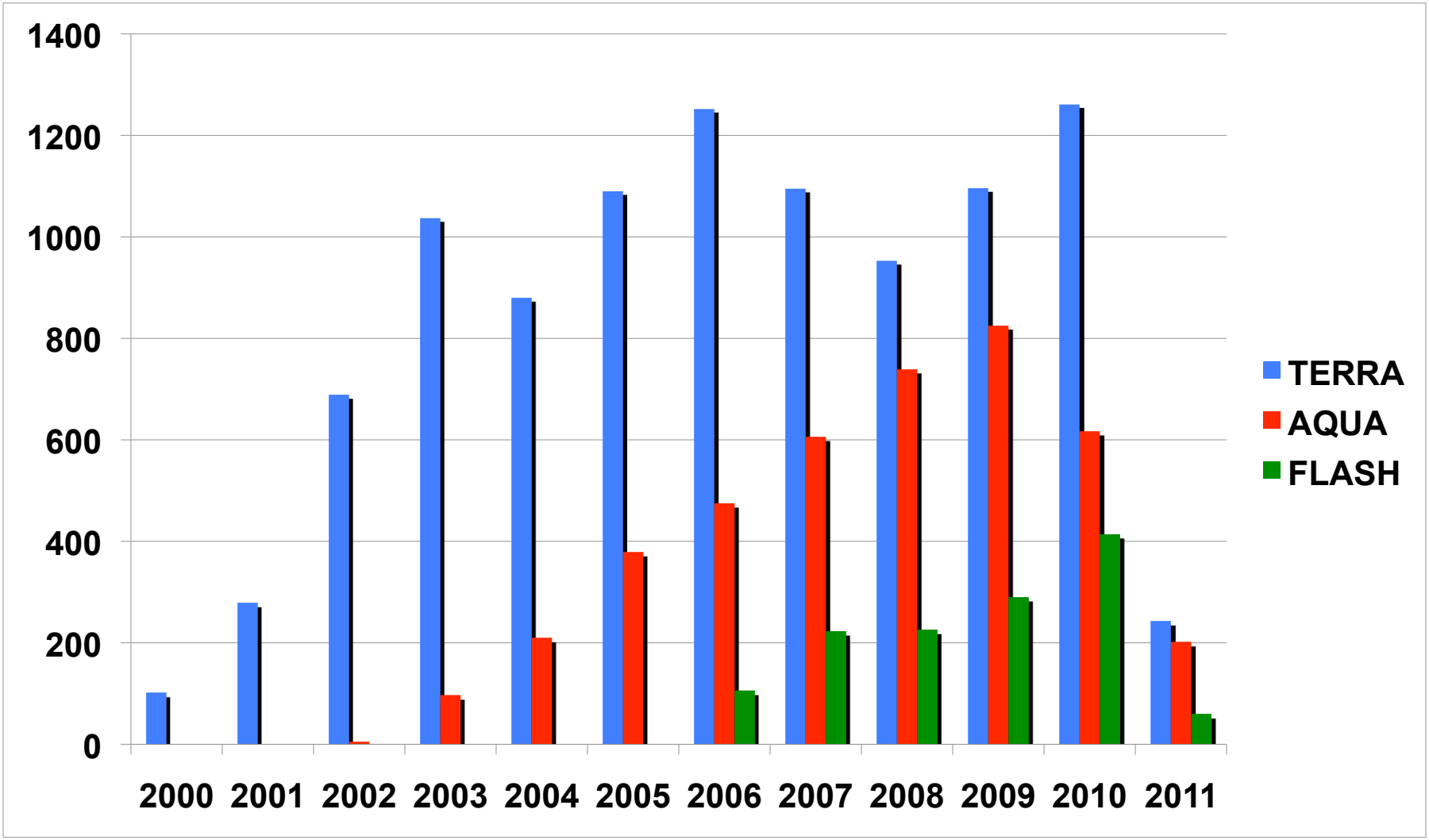
By Data Year



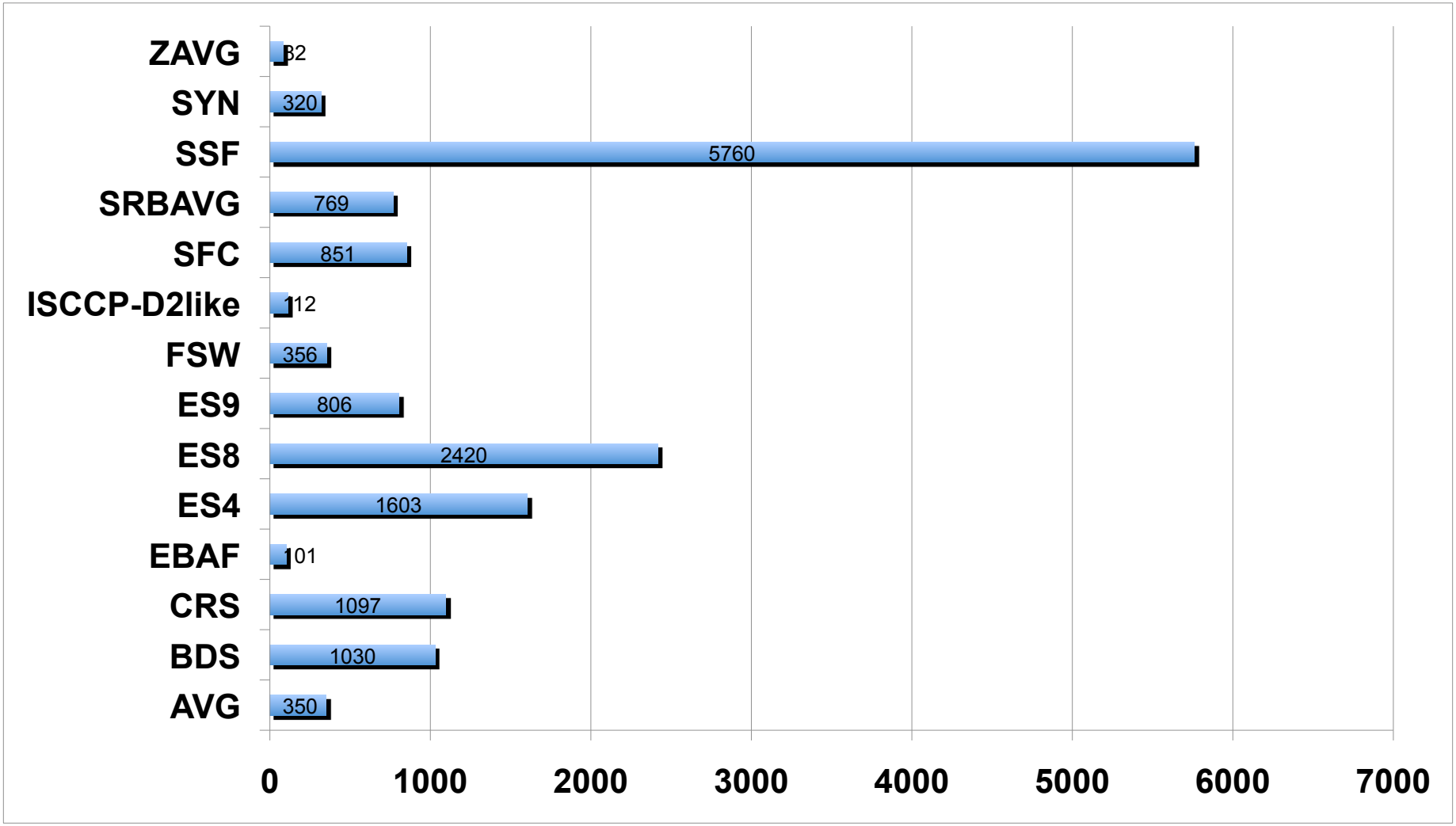
CERES and FLASHFlux Data Distribution



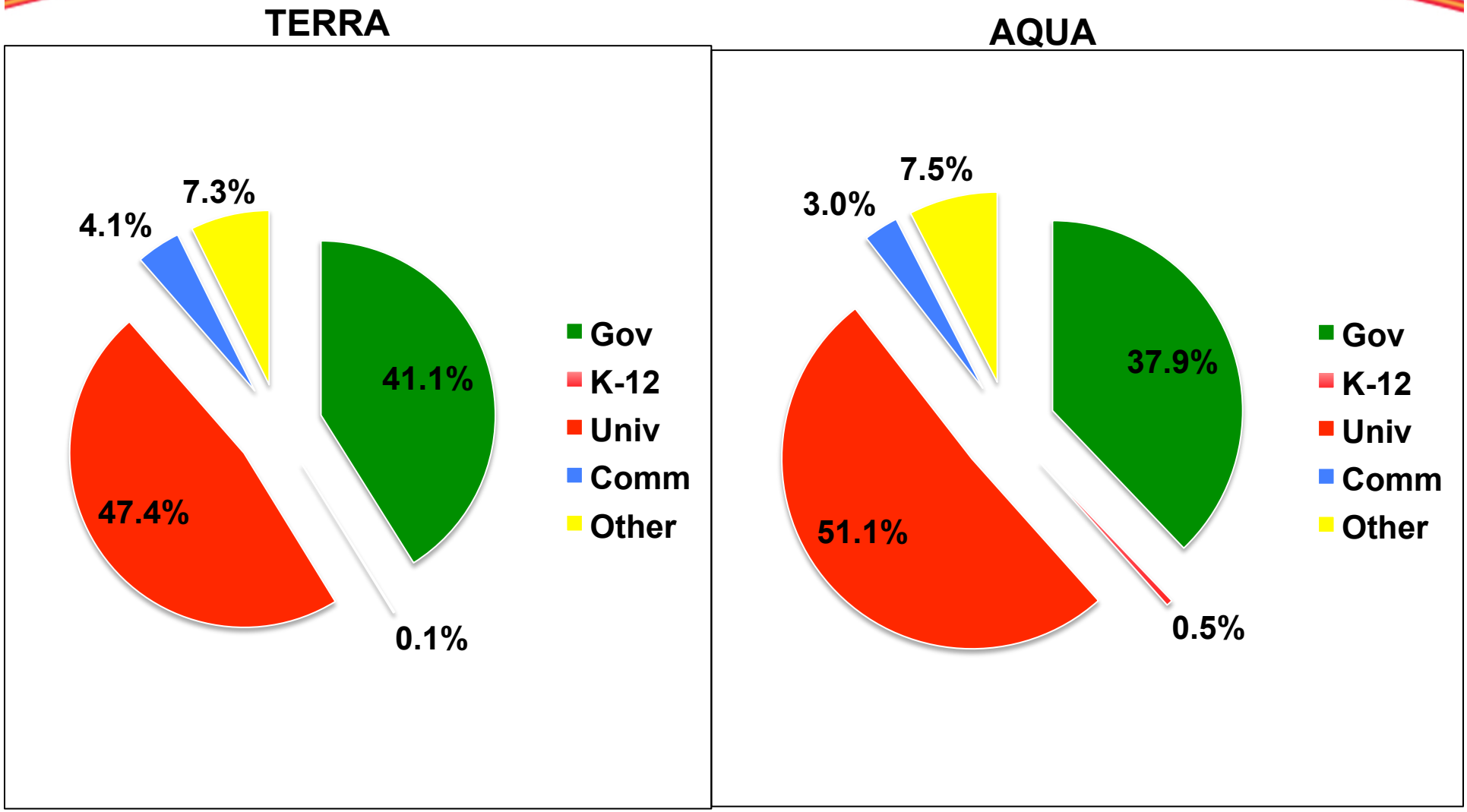
CERES and FLASHFlux Data Orders



CERES Orders by Product (Mar 2000- March 2011)

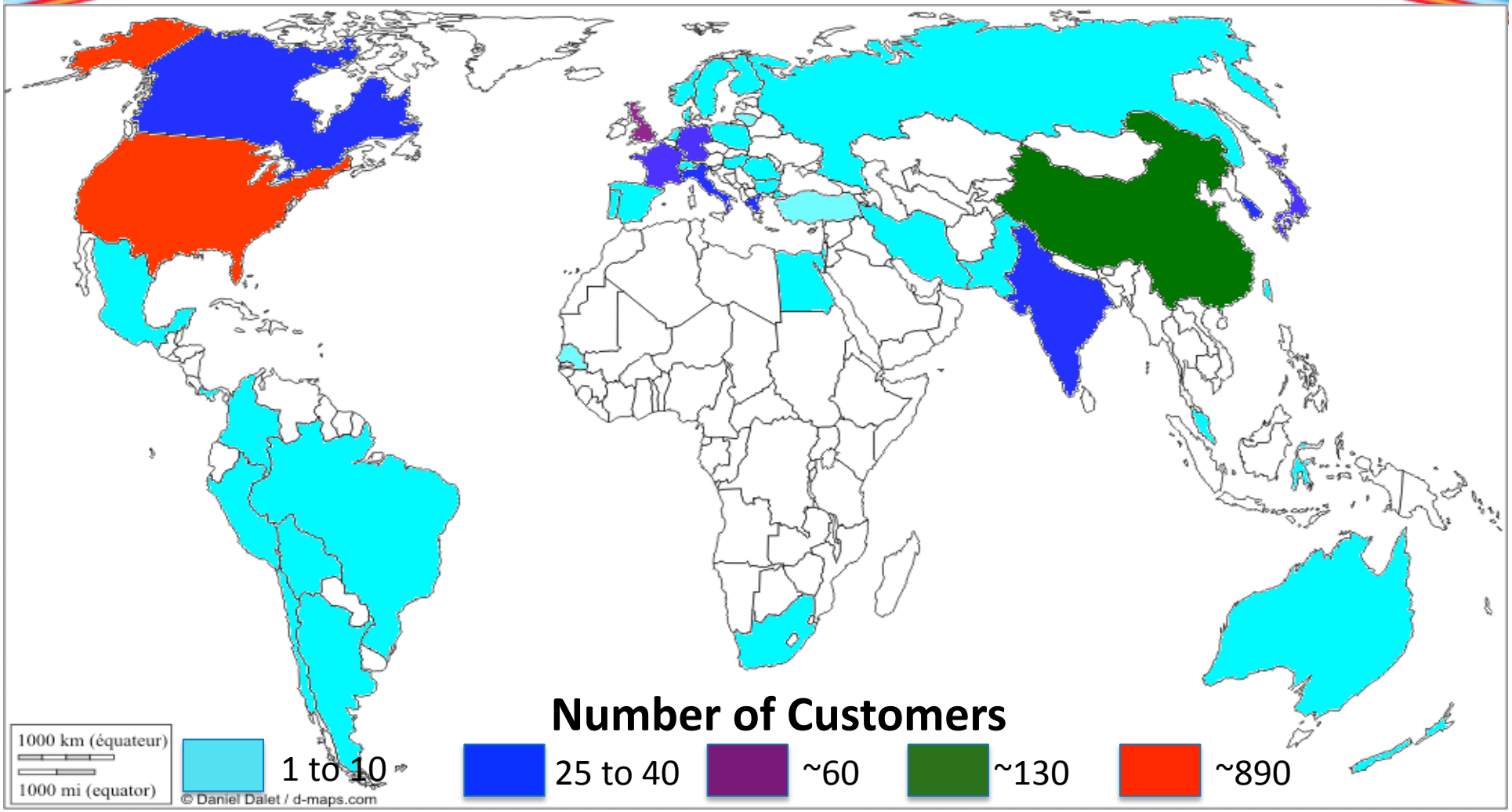


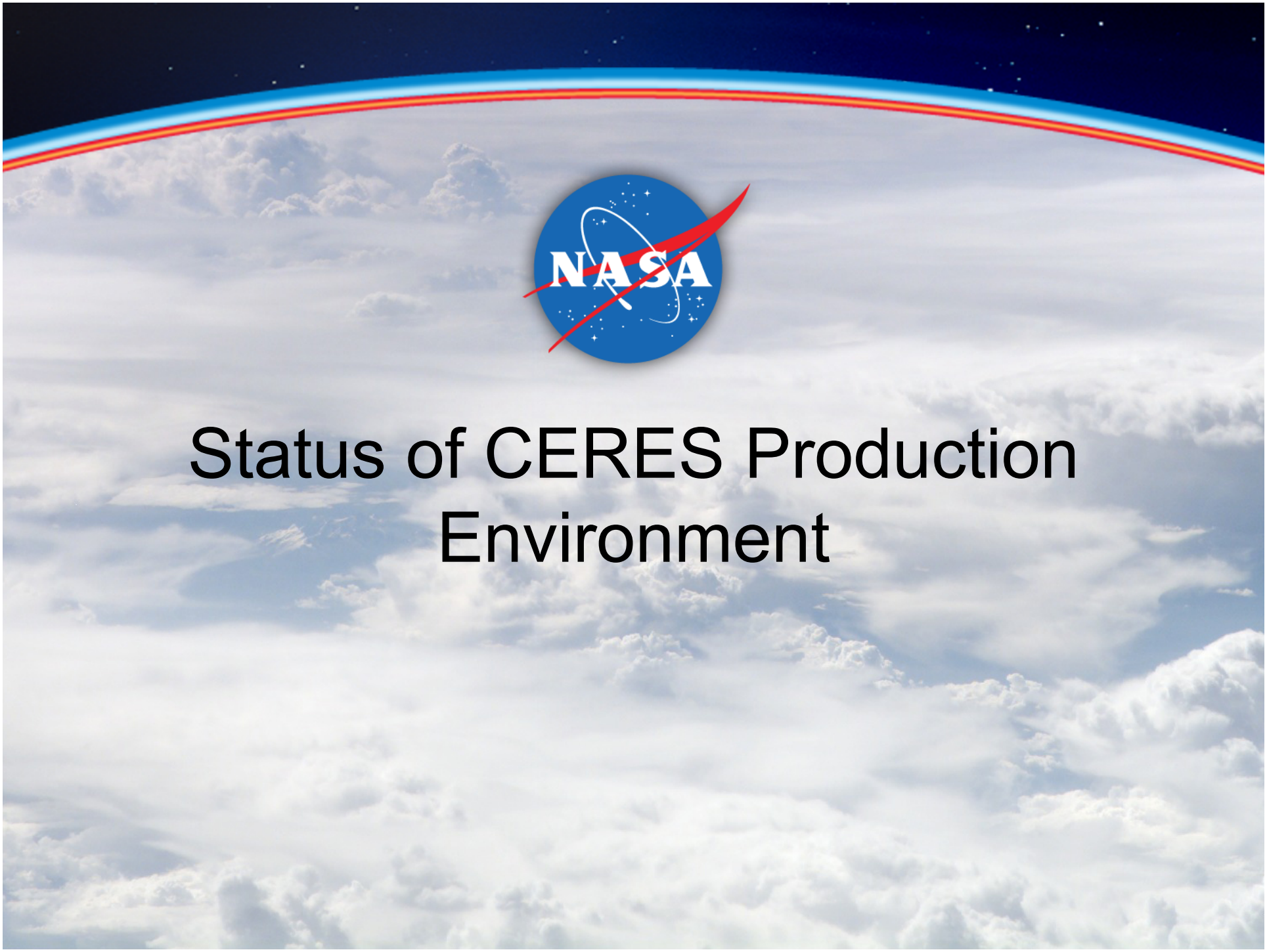
CERES Customers by Affiliation



ASDC CERES Data Users

Number of Customers



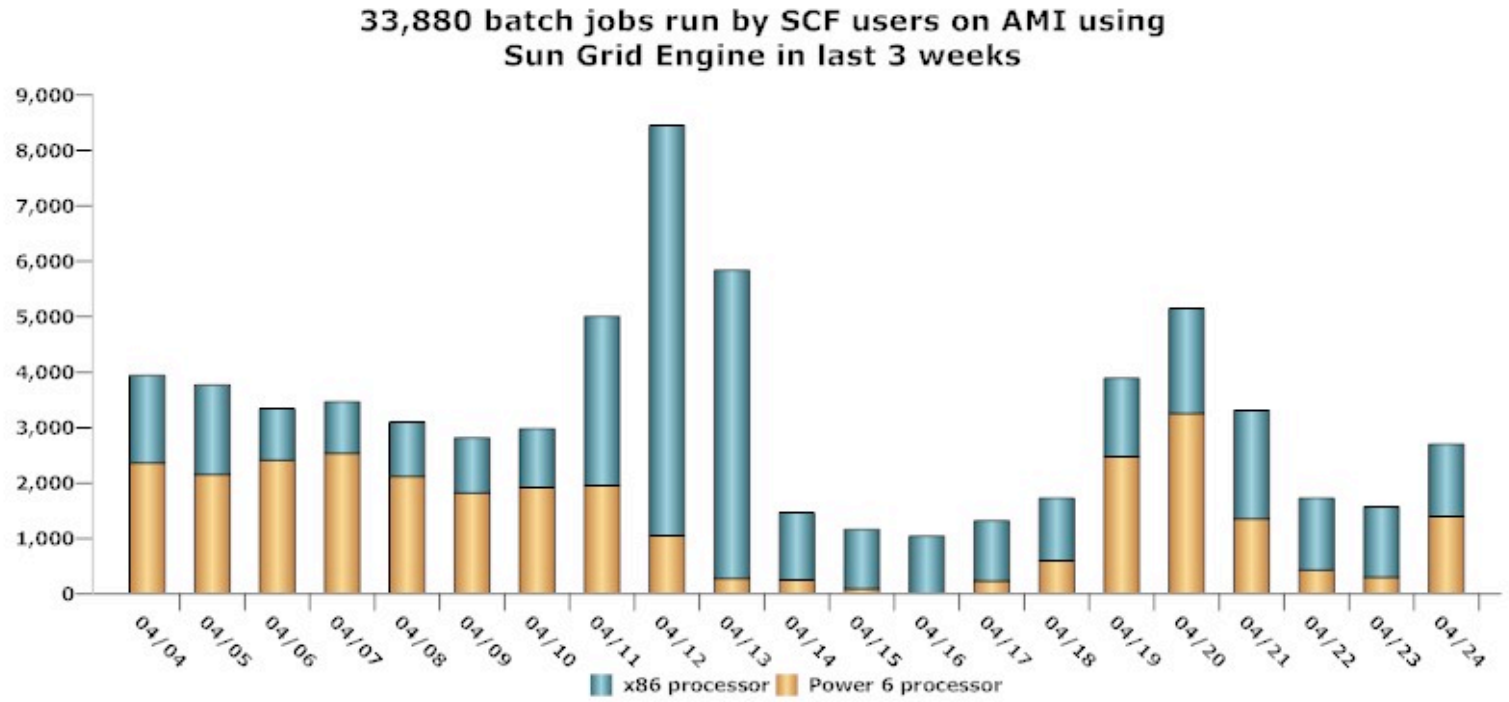


Status of CERES Production Environment

AMI

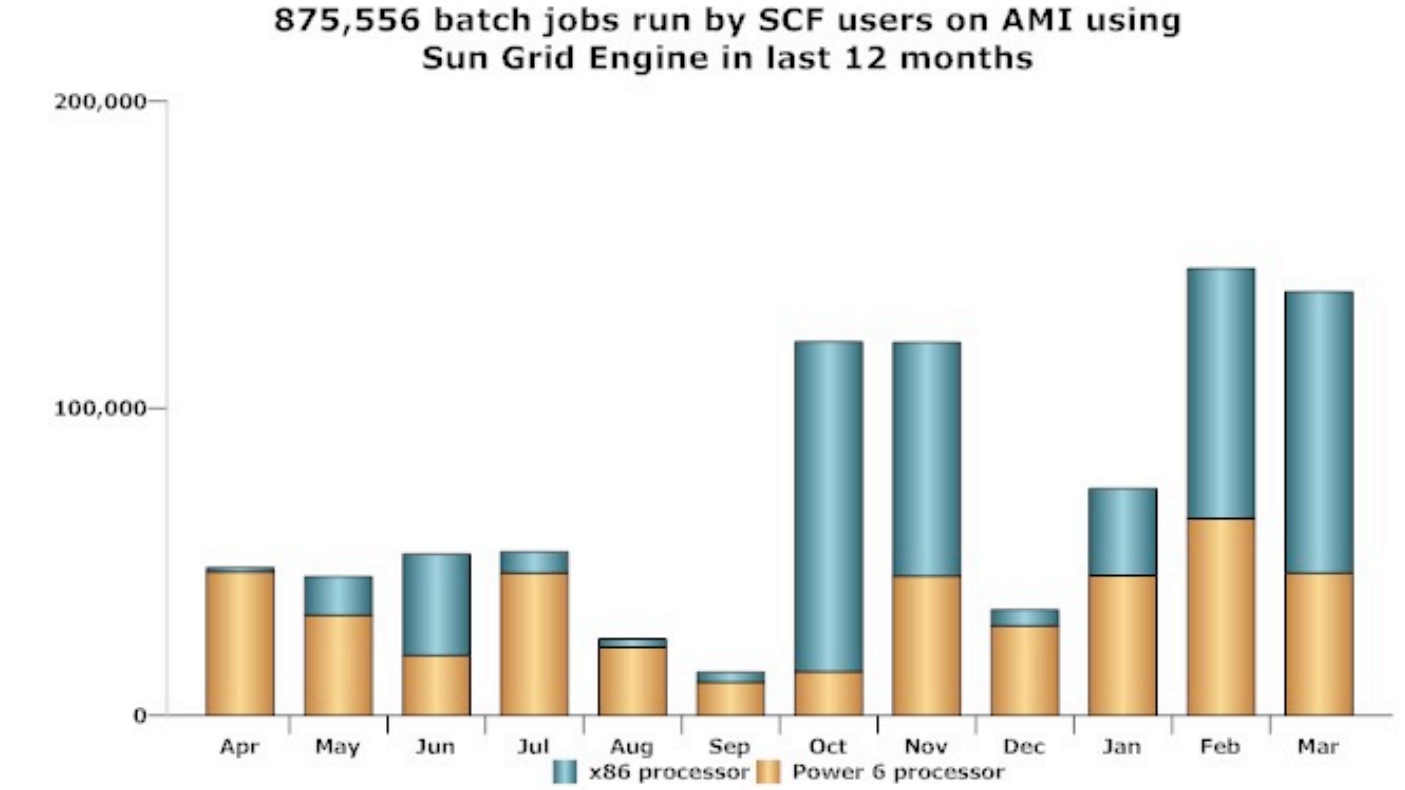
- ASDC Modernization through Integration
- New (relatively) processing and storage based on IBM technologies
- Challenges with initial setup and configuration
  - No one person has all the right answers
  - Consultants, vendors experts, in-house expertise
    - Convergence was difficult
- Stability issues primary concern
- Orchestrated a redesign effort within constraints
- Technical review early April
  - Extremely valuable in validating design
- AMI has been very stable recently

Status of CERES Processing Environment

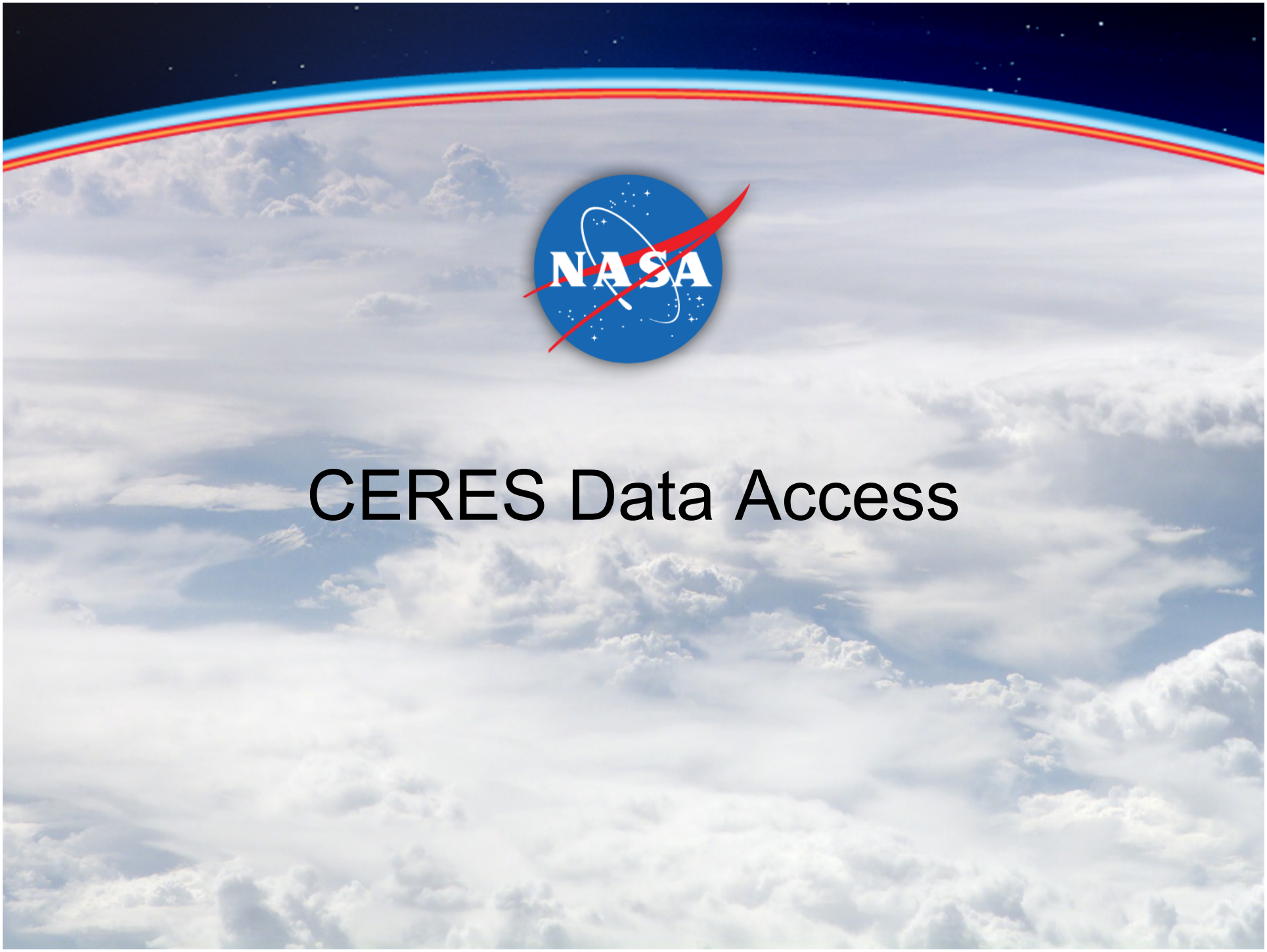


	04/04	04/05	04/06	04/07	04/08	04/09	04/10	04/11	04/12	04/13	04/14	04/15	04/16	04/17	04/18	04/19	04/20	04/21	04/22	04/23	04/24
Power 6 processor	2,359	2,155	2,415	2,531	2,117	1,823	1,918	1,951	1,053	270	245	92	0	226	590	2,469	3,255	1,355	427	296	1,391
x86 processor	1,580	1,616	924	930	987	987	1,059	3,056	7,400	5,567	1,219	1,071	1,040	1,089	1,133	1,417	1,895	1,954	1,294	1,275	1,309

Status of CERES Processing Environment



	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Power 6 processor	46,789	32,752	19,680	46,487	22,186	10,853	14,340	45,399	29,328	45,595	64,119	46,385
x86 processor	1,390	12,586	32,899	6,877	2,873	3,320	107,379	76,037	5,114	28,335	81,566	91,642



CERES Data Access

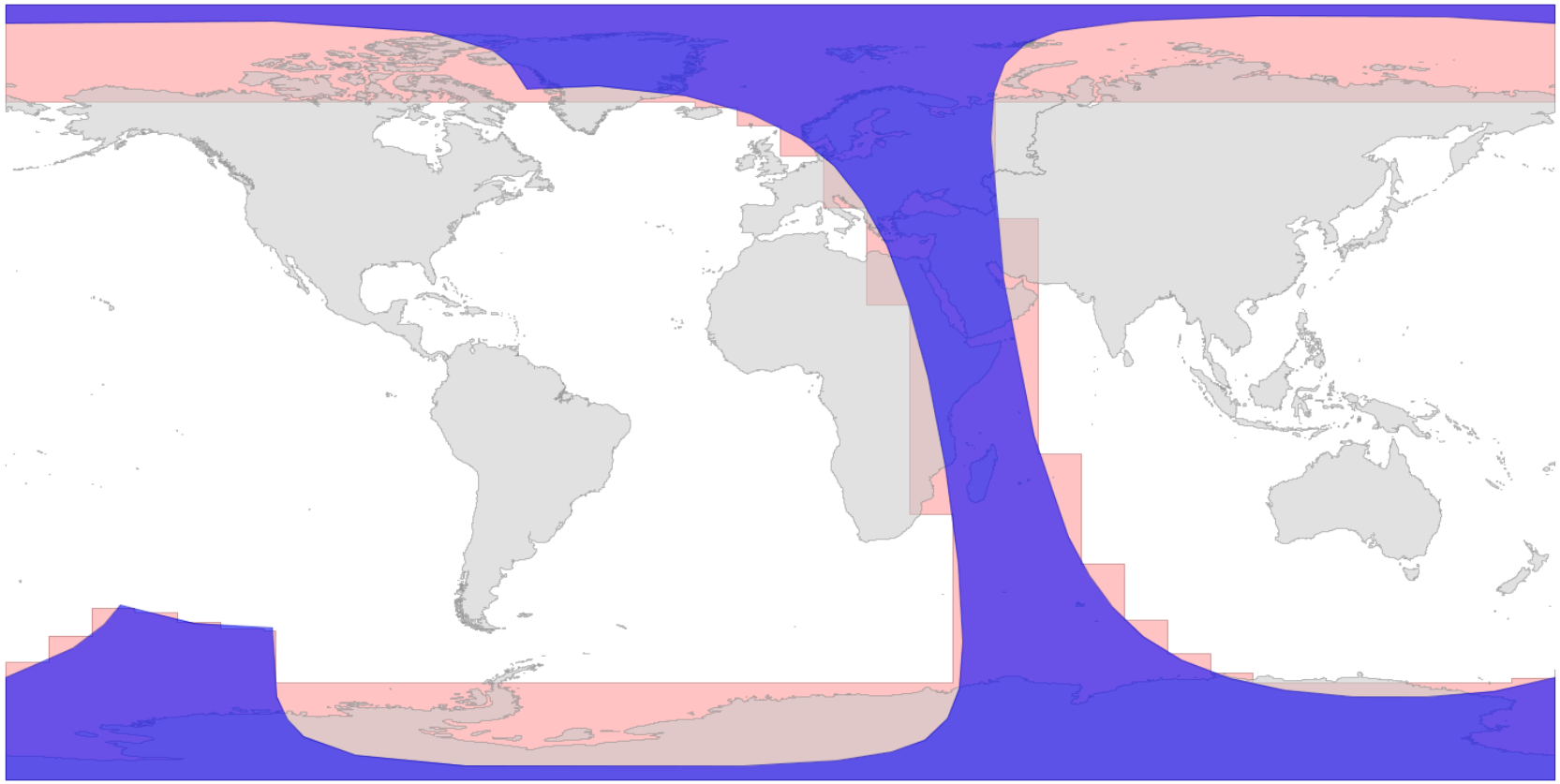
CERES Subsetting

- ASDC hosting current CERES Level 3 subsetter
  - Well received by the science community
- Current collaborative effort underway to develop CERES Level 2 subsetting capability
  - Development model has been rapid prototyping (iterative)
  - Progress to date
    - Subsetter developed for SSF Level 2 Edition 3A data
    - Subsetter framework (facilitates interaction between user interface and data subsetter), data subsetter, and final distribution is complete
    - Subsetter's spatial metadata database contains more complete and accurate footprints of each CERES granule than anything previously available
    - Subsetted files can be delivered in either HDF4 or NETCDF
    - Multiple NETCDF subset results can be bundled into a single file

CERES Data Comparison Plots

New spatial metadata mined directly from the CERES hdf granules to use in new CERES subsetter

Default spatial metadata supplied with the CERES granules in the .met files

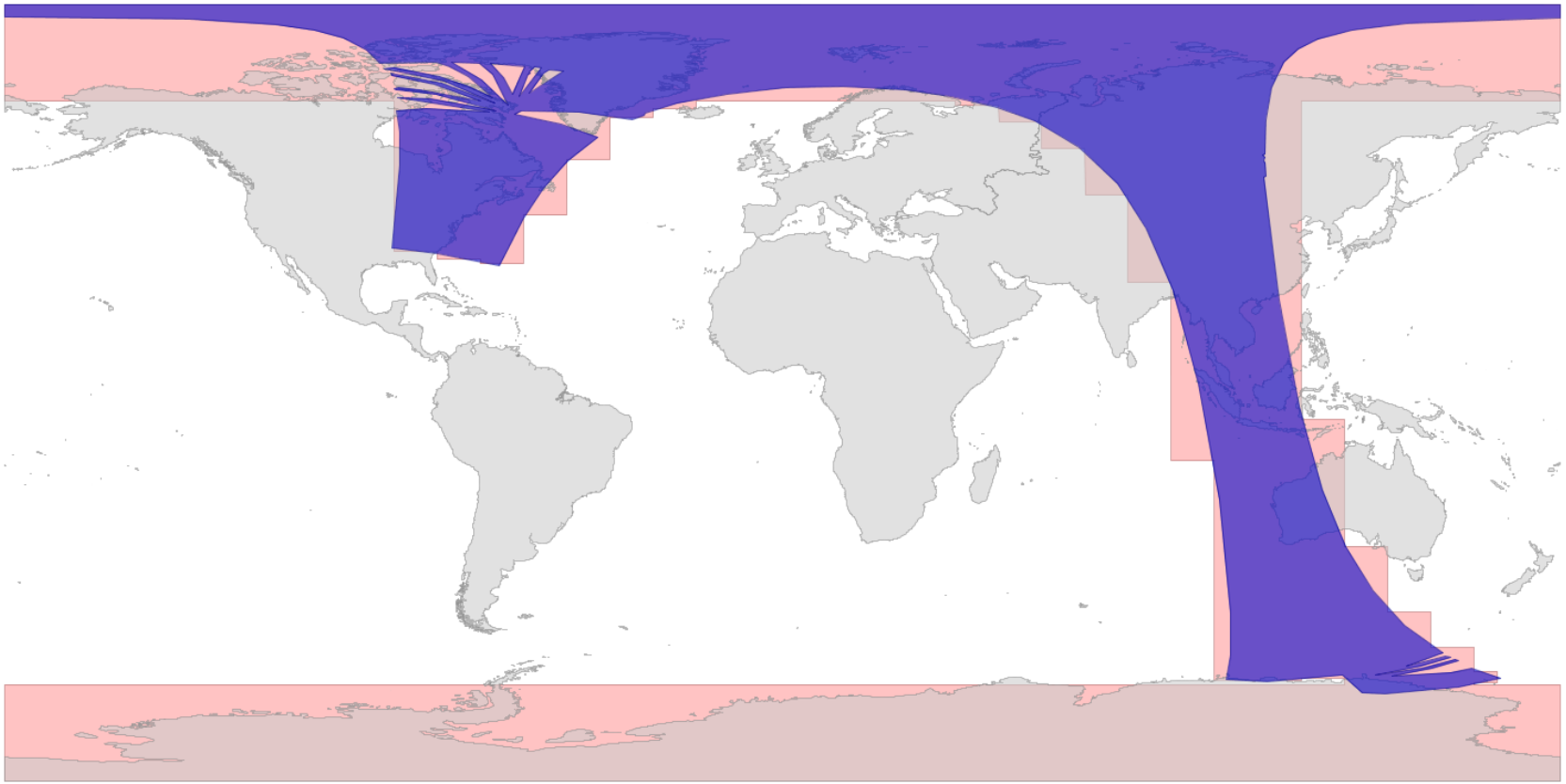


CERES SSF Aqua-FM3-MODIS Edition3A

CERES Data Comparison Plots

New spatial metadata mined directly from the CERES hdf granules to use in new CERES subsetter

Default spatial metadata supplied with the CERES granules in the .met files

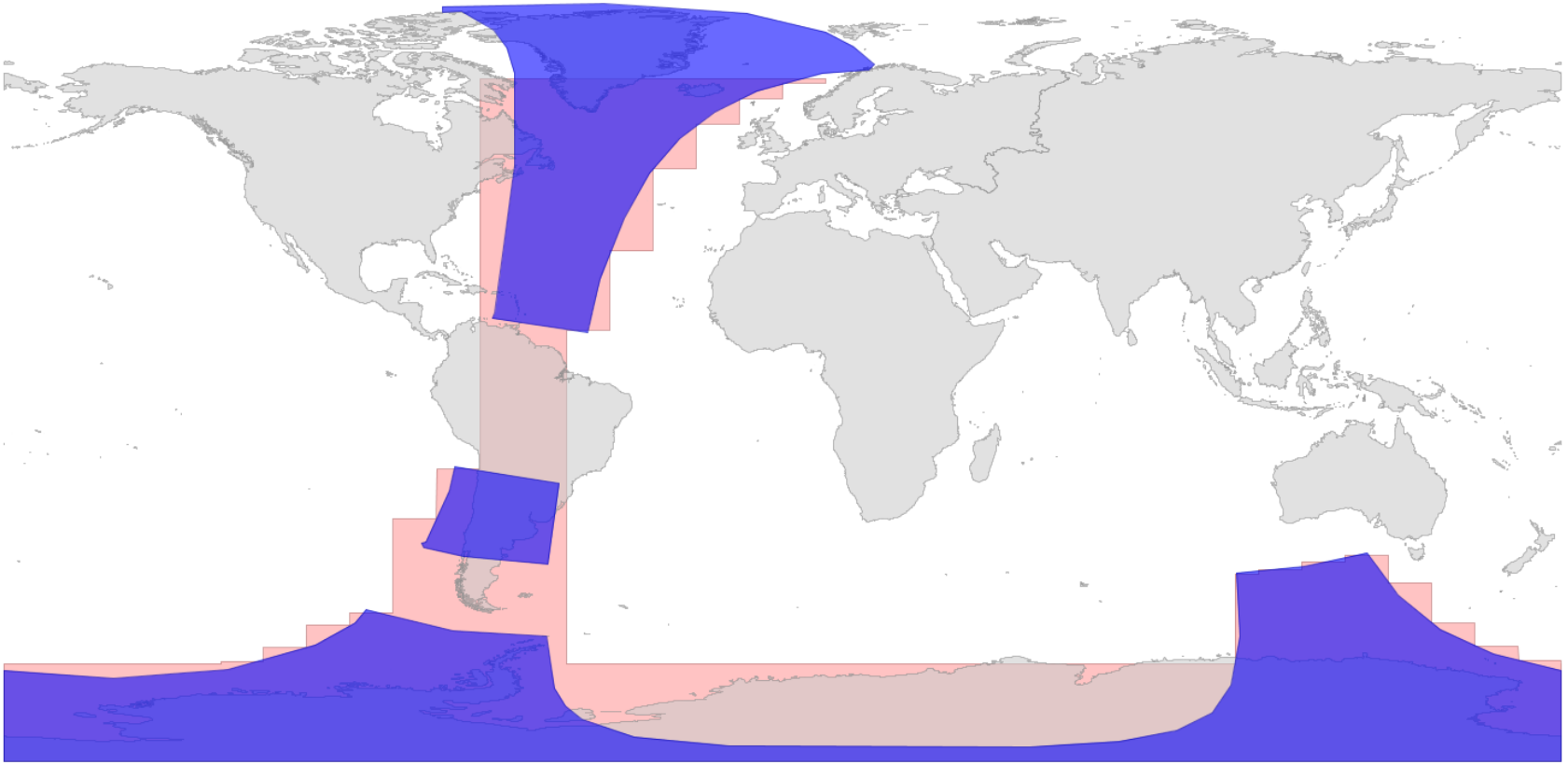


CERES SSF Terra-FM1-MODIS Edition3A

CERES Plots

New spatial metadata mined directly from the CERES hdf granules to use in new CERES subsetter

Default spatial metadata supplied with the CERES granules in the .met files



CERES SSF Terra-FM1-MODIS Edition3A

Improving CERES Ordering

- Current ASDC CERES landing pages and order tools were developed years ago
  - Largely static HTML with simple pull downs
- Working with CERES science team to improve the user experience
  - Plan to leverage web page design supporting CERES subsetters
    - Developed pages have been customer driven
    - Need to ensure that all customer needs are addressed
  - Planning long-term development and support models
    - Agreements on interfaces between user interfaces and subsetting/ordering
    - Availability, maintenance, user support

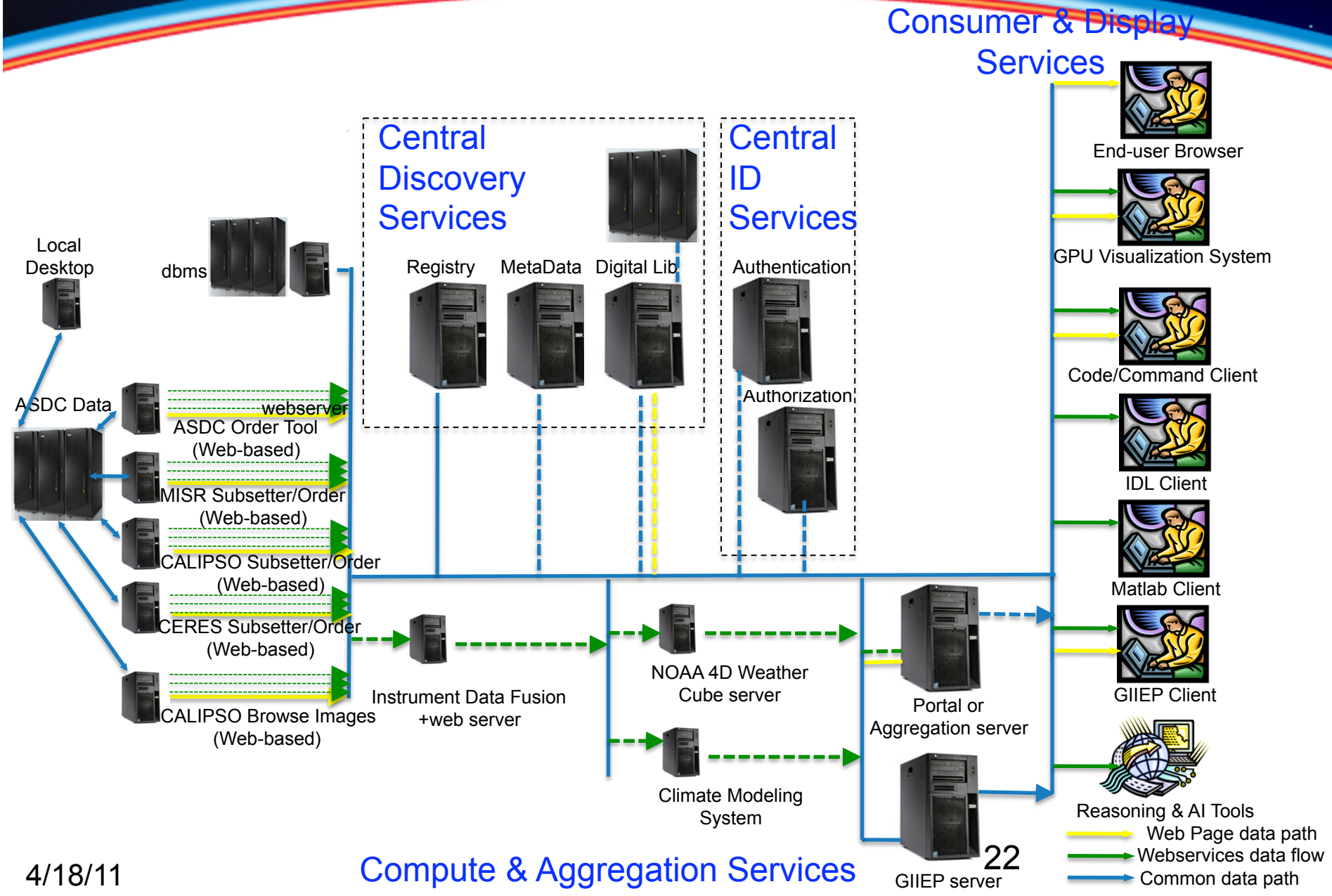
Future: Understanding Benefits of our Data

- Instrument Science Teams create the data and use it themselves
  - Quality checks
  - Algorithm improvements
  - Inter-calibration
- Other Instrument Science Teams
  - Data Fusion with their own instrument output
  - Inter-calibration of similar measurements
  - Evaluation/Comparison of Alternate Algorithms
- Climate Modeling (Global, regional and thematic studies)
- Climatology and Assimilation Models
- AI tools can help make unexpected correlations of measurements
    - Non-obvious relationship analysis
    - Edge Detection algorithms
- Justification for future instrument missions
  - Based on volume/diversity of use
- Data Stewardship investments
  - Based on volume of access and “value” of the data to the community

Future: Net-centric Ops (NCO) Applied to ASDC

- NCO = Information sharing and re-use
  - Involves discovery, accessibility and understanding
  - Many paths to the same data
    - Keep existing ordering capabilities and add more pathways
    - Instrument Team Defined Storage Model based on instrument output
    - Access model does not have to be dependent upon the storage model
  - Data must be convenient to the researcher-users
- Machine-to-machine interface is critical to practical use of “big data”
  - Minimize manual handling of the data to use it
  - Must convey restrictions, boundary conditions, error bands
  - ASDC MUST make Recipients comfortable that ASDC will guarantee long-term availability for future review
    - Data Stewardship Means Never HAVING to say You’re Sorry

Notional Webservices View



Conclusion

- Metrics indicate the ASDC is healthy in supporting its mission
- AMI has become more stable, more effective and customers are utilizing its capability
- Subsetting efforts are moving forward and are responsive to user needs
- Effort to improve the user experience for those visiting ASDC data pages is underway
- Strategic planning is underway to make data available to customers in new paradigms