University of Wisconsin SSEC Data Center

May 7, 2012

SSEC Data Center

- The SSEC Data Center mission is to create and maintain the facilities, human expertise and technology necessary to provide SSEC scientists and collaborators with the highest quality geophysical data in a timely fashion and to provide real-time data access, archive and retrieval services as necessary to support SSEC's scientific programs.
- The Data Center is treated as a project and is expected to breakeven.
- The Data Center is supported through data services. ALL users pay for services, both internal to SSEC and the outside community.

SSEC Data Center - Activities

- Assist Satellite Operations Control Center and other agencies in satellite checkout and troubleshooting of related problems
- Support to field experiments
  - Special archiving
  - Extended staffing, either on-call or on-site
- Provide large dataset backups for users
  - Read, write, and copy tapes
  - Provide specialized archives of user data
- Provide Help Desk support to users of the SSEC SDI, and assist in the generation of SDI user documentation
- Assist McIDAS User Services team with McIDAS testing for system upgrades
- Provide archive data to in-house and external users
- Top level provider of GOES data to the Unidata UNIWISC IDD feed, and hosts Unidata server for IDD relaying and GOES ADDE serving

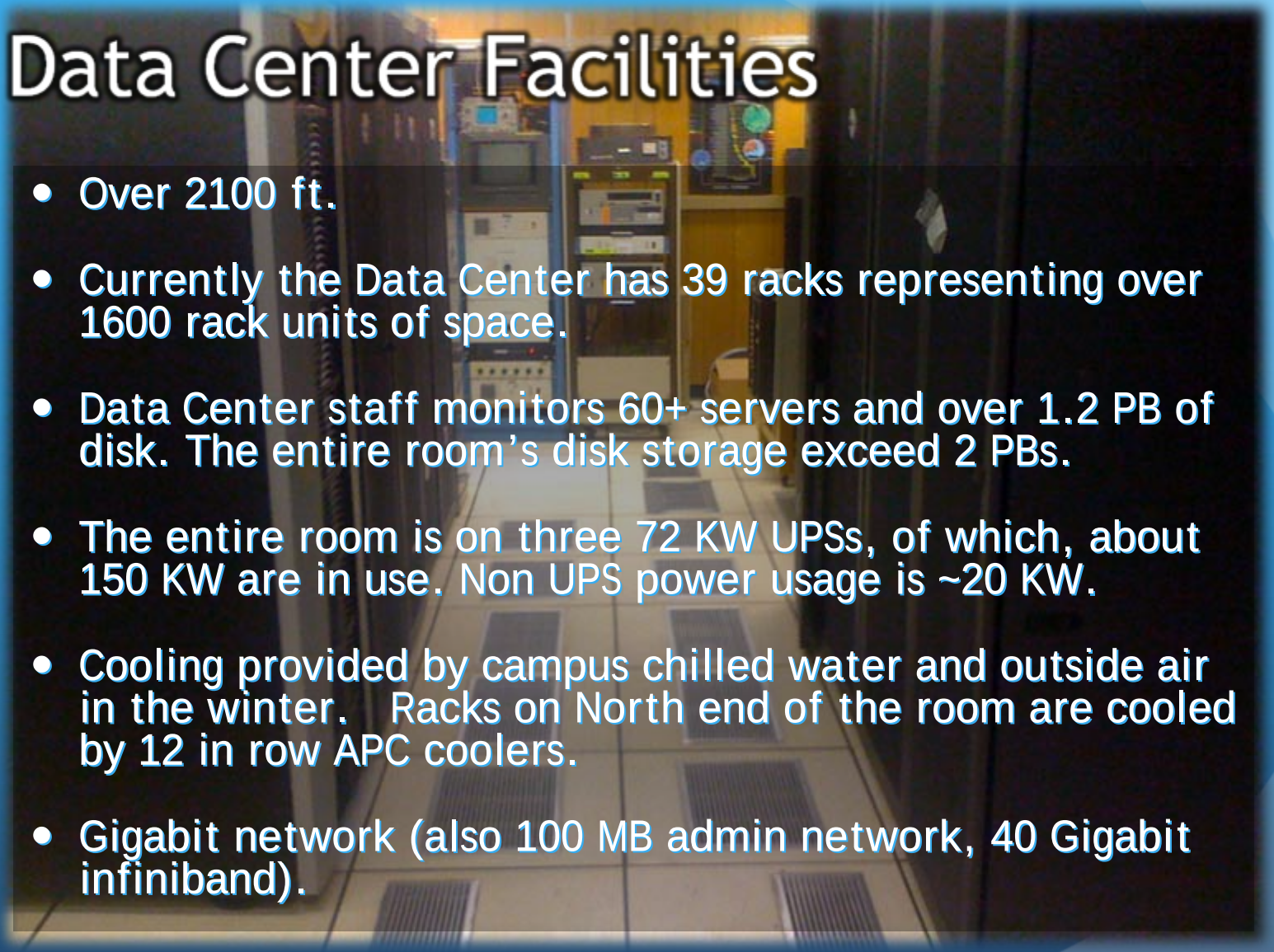
SSEC Data Center - Staff

Staffed M-F , 7:30 AM - 11:00 pm Central time.

- 3 FTE ~100% time
  - Archivist & Computer Operator (1st shift)
  - Computer Operator (1st shift)
  - Computer Operator (2nd shift)

- 5 FTE ~portions of their time
  - Program Manager
  - System Programmer
  - Data Base Programmer
  - Research Specialist (PM assistant)
  - Antenna/Communication technician
- 1 Student programmer
- 1 Student archive assistant

Data Center Facilities

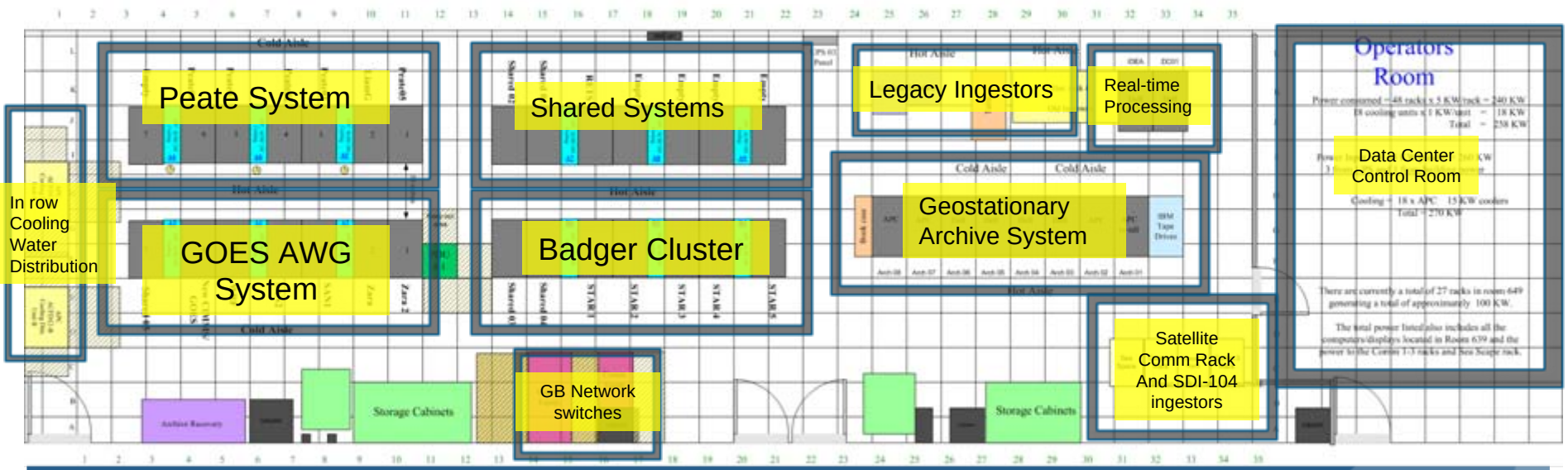


- Over 2100 ft.
- Currently the Data Center has 39 racks representing over 1600 rack units of space.
- Data Center staff monitors 60+ servers and over 1.2 PB of disk. The entire room's disk storage exceed 2 PBs.
- The entire room is on three 72 KW UPSs, of which, about 150 KW are in use. Non UPS power usage is ~20 KW.
- Cooling provided by campus chilled water and outside air in the winter. Racks on North end of the room are cooled by 12 in row APC coolers.
- Gigabit network (also 100 MB admin network, 40 Gigabit infiniband).

Facilities layout

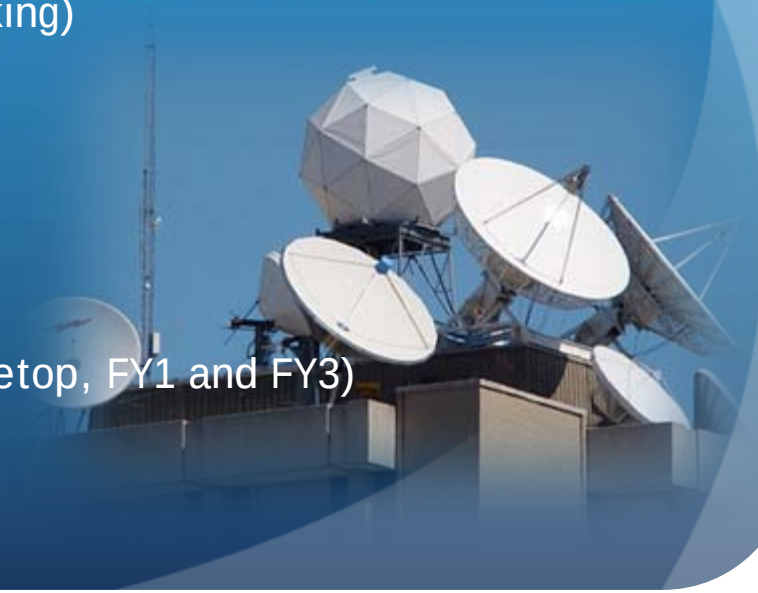
Existing Configuration

Early 2012



Data Center Antennas

- C-Band
  - 11 meter (87° West - POES Wallops Relay, MSG)
  - 7.3 meter (101° West – POES Fairbanks Relay, MTSAT, Noaaport)
  - 6.3 meter heated backup (101° West – POES Fairbanks Relay, MTSAT, Noaaport)
- L-Band
  - 7.3 meter (75° West –GOES-East Primary)
  - 4.6 meter (135° West –GOES-West Primary)
  - 4.5 meter (60° West –GOES-SA auto tracking)
  - 4.5 meter (105° West –GOES-test/spare)
  - 3.7 meter (offline spare)
- X-Band
  - 4.4 meter (Tracking – EOS)
- X/L Band
  - 2.4 meter (Tracking – Suomi NPP, EOS,metop, FY1 and FY3)



SSEC Data Center Incoming Data

February 2012

170+ GB/day
via Satellite
(C-band. L-band, X-band)

GOES satellites	~96 GB/day
International Geo Satellites	~47 GB/day
NOAA Polar	~27 GB/day
Miscellaneous Polar and Non satellite	~85 GB/day
MODIS polar from NASA archive	~150 GB/day
NPP (VIIRS CrIS ATMS)	~700 GB/day

935+ GB/day
via Internet
(ftp, LDM, ADDE, http)

SSEC Data Center Outgoing Data

Four primary methods of Data delivery

1.ADDE
2.HTTP
3.FTP
4.LDM (Unidata local data manager)

Outgoing Data

February 2012

- On average over 1 million ADDE transactions per day
- Over 1 TB data distributed per day

Real-time Data

The SSEC Data Center receives data from 9 different geostationary satellites and 10 different polar orbiting satellites. Most data are available in near real-time via ADDE. Other methods of data access are available upon request.

Geostationary Satellites

GOES East/West/ S. America

FY-2

MTSAT

Kalpana

Meteosat *(access restricted)*

Polar Satellites

- NOAA AVHRR
- NPP
- METOP
- EOS (AQUA/TERRA)
- FY1,FY3

NOAAPORT/Conventional Data

- Model Output (GFS, RAP, etc)
- NEXRAD
- NWS Text output

Current Real-Time Satellite Data availability (ADDE and/or ftp)

Geostationary

US		Eumetsat		Japan		India		China	
Dataset	Delay	Dataset	Delay	Dataset	Delay	Dataset	Delay	Dataset	Delay
GOES-15 (WEST) GOES-13 (EAST) GOES-12 (SA)	< 5 min	Meteosat-9 (Prime) Meteosat-7 (IODC)	<15 min	MTSAT-2	<5 min	Kalpana1	.75 – 2+ hours	FY-2D FY-2E	15-30 min

Polar

NOAA POES (Wallops and Gilmore Creek DOMSAT Relays, and DDS)		EOS		Suomi NPP		Other	
Dataset	Delay	Dataset	Delay	Dataset		Dataset	Delay
NOAA-15 NOAA-16 NOAA-18 NOAA-19	GAC 1.5 hours – 6+hours HRPT <20 minutes	AQUA TERRA	<30 min	VIIRS CRIS ATM	<30 min	METOP FY1, FY3	1.5 hours – 6+hours < 10 minutes after pass

Archive Data

Over 685 TBs online.

Geostationary Satellites

- SMS-1&2 *(1978-1981)*
- GOES-1 through GOES-7 *(1978-1996)*
- GOES-8 through GOES-15 *(1994-Present)* (East, West and South America)

International Geo Satellites

- GMS/MTSAT *(1998-Present)*
- Meteosat/Meteosat IODC *(1998-Present)*
- FY2 (2004-Present)
- Kalpana *(2005-Present)*

NOAAPORT/Conventional Data

- Model Output *(1996-Present)**
- Insitu Point Observations *(1976-Present)*

**available via ADDE and THREDDS*

Recent Data Center Changes

Increased Computer Room Security

Update to Control room

New Programmer

- Kevin Hallock
- Data Center Support
- New Archive metadata Database and Inventory interface
- XCD replacement

New Dual X/L Band antenna

S4 Supercomputing Cluster

•Funded by the National Oceanic and Atmospheric Administration (NOAA), S4 is used by NOAA and UW researchers to run data assimilation experiments with the goal of improving the NOAA operational weather models used to generate weather forecasts for the United States.

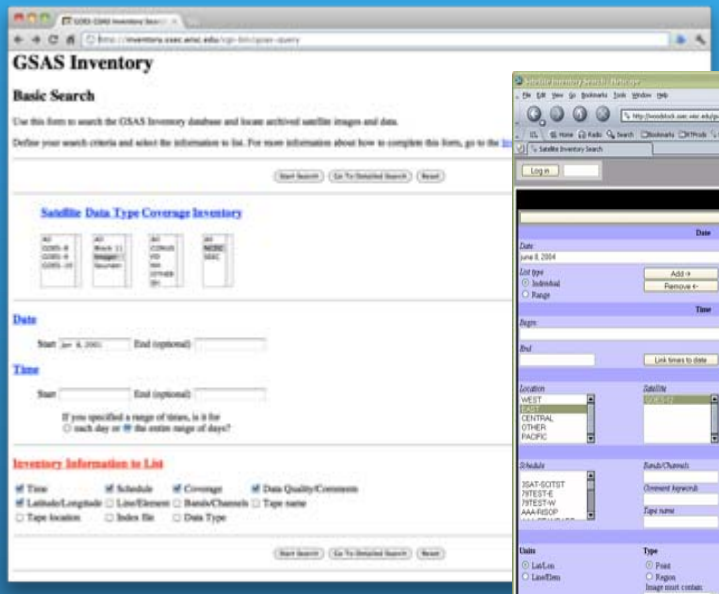
•The system was designed, installed, and is maintained by the UW SSEC Technical Computing Group.

S4 Hardware

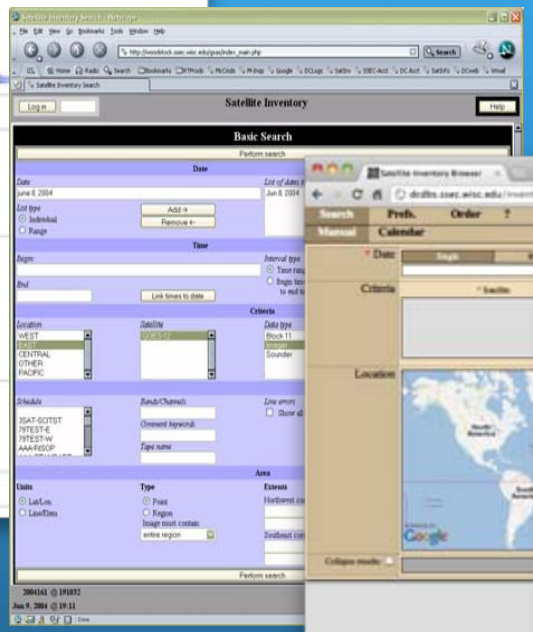
- Largest High Performance Compute Cluster at UW Madison
- Total compute: 3072 cores, 8TB memory
- Total storage: 456 TB
- 40 Gigabit/s Infiniband Network Interconnect

Current Data Center Projects

Satellite Data Archive Inventory Interface rewrite



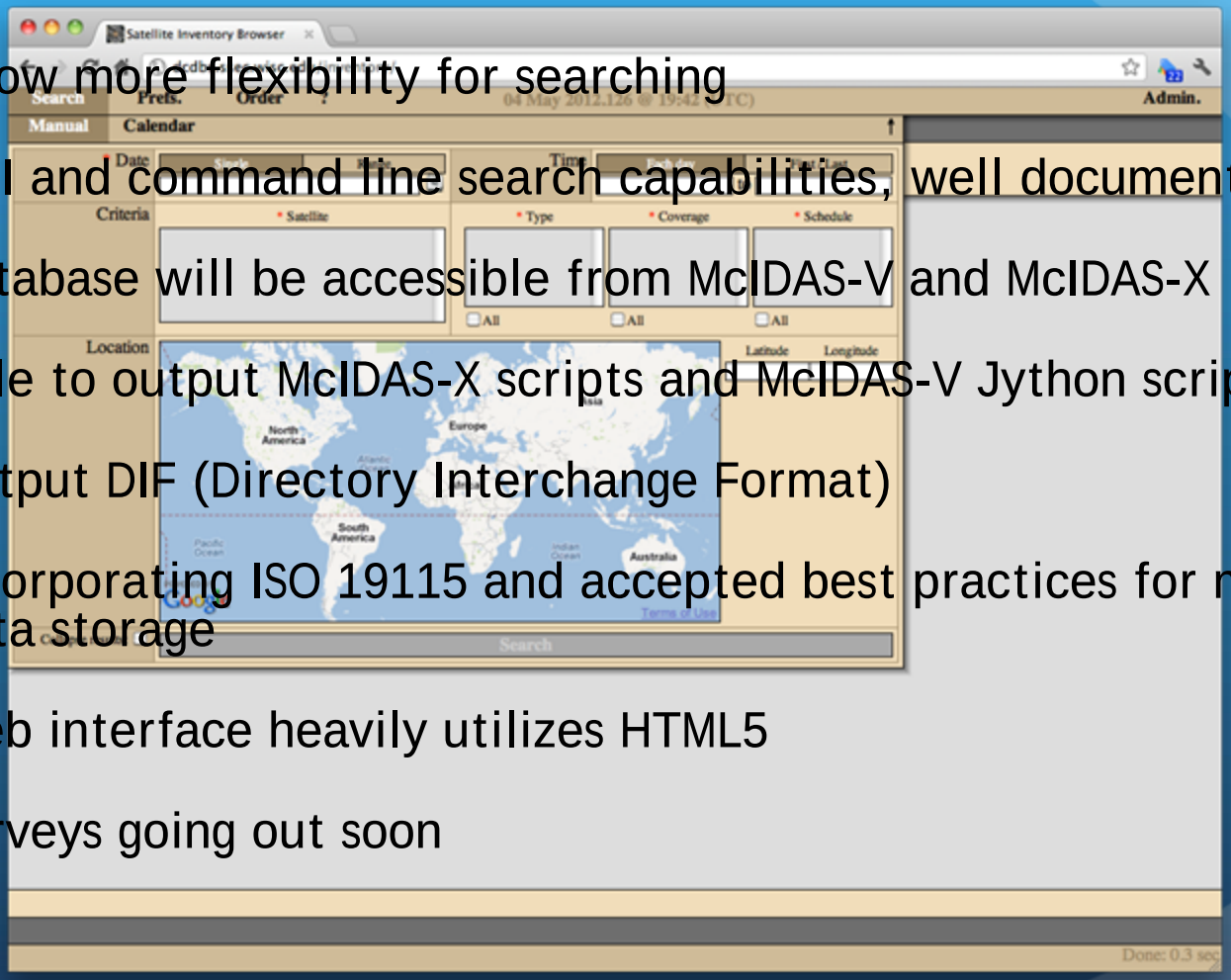
(1996-2003)



(2003-2007)

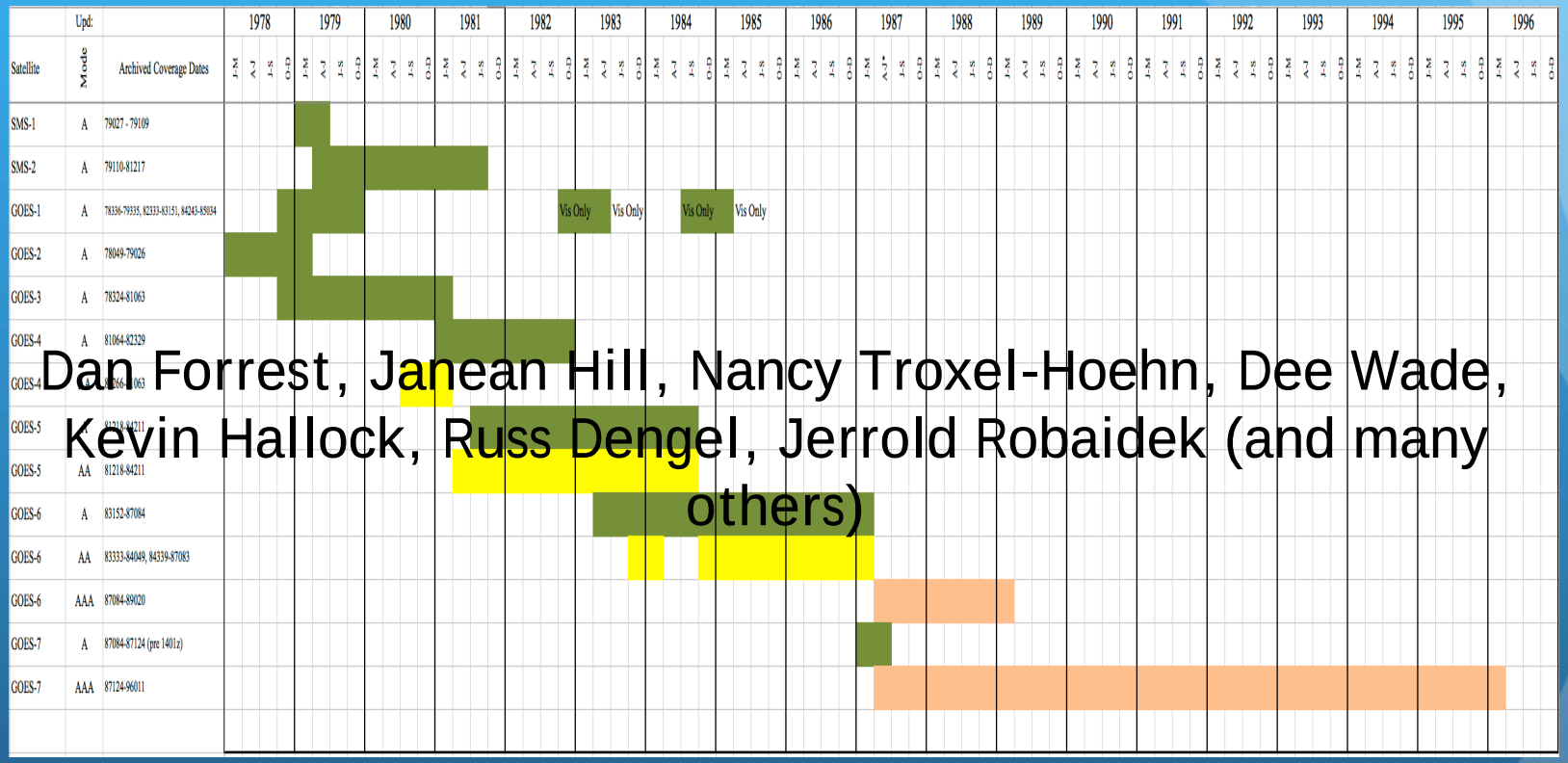
(2007-Present)

Satellite Data Archive Inventory Interface rewrite



- Allow more flexibility for searching
- GUI and command line search capabilities, well documented API
- Database will be accessible from McIDAS-V and McIDAS-X
- Able to output McIDAS-X scripts and McIDAS-V Jython scripts
- Output DIF (Directory Interchange Format)
- Incorporating ISO 19115 and accepted best practices for meta data storage
- Web interface heavily utilizes HTML5
- Surveys going out soon

Early GOES Archive - Data Interrogation and Reconciliation



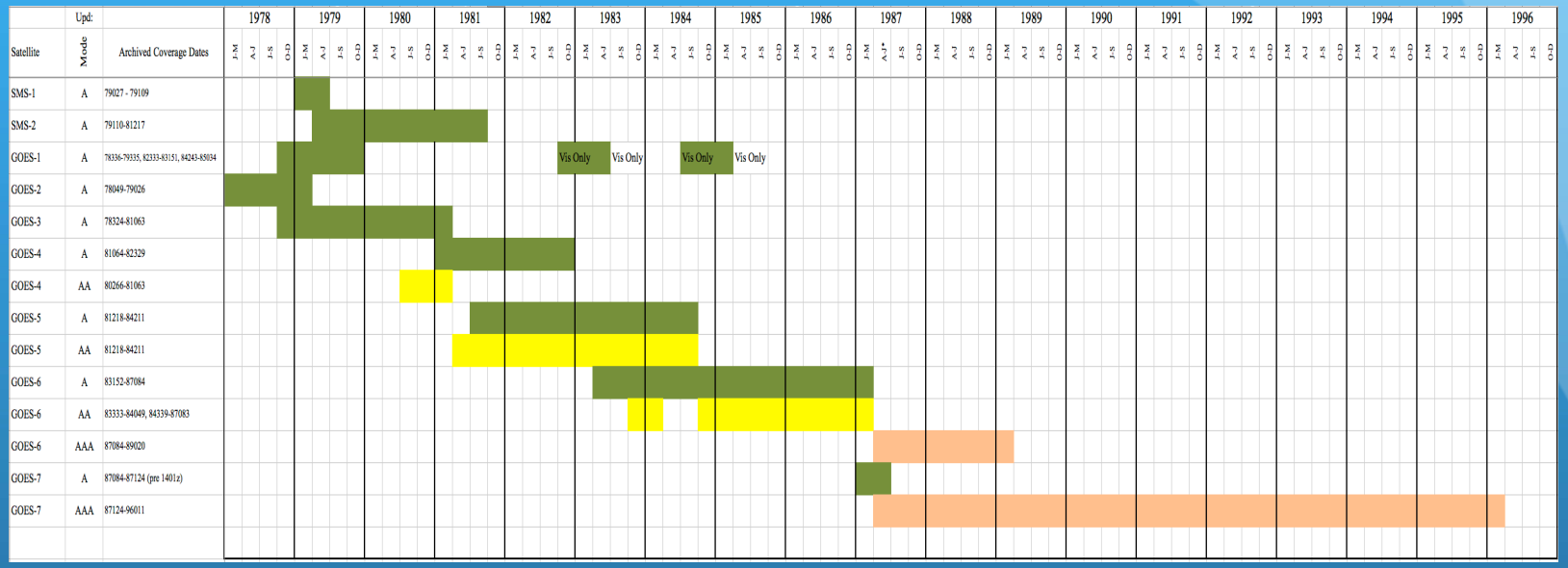
Dan Forrest, Janean Hill, Nancy Troxel-Hoehn, Dee Wade, Kevin Hallock, Russ Dengel, Jerrold Robaidek (and many others)

GOES Archive History

- Mid-1970s: SSEC began recording data to U-matic tape
- 1978: SSEC started keeping the data
- 1979: SSEC became the official archive of GOES data (ended in 2004)

Dr. Vern Suomi

Early GOES



- May 17, 1974 (SMS-1) – Jan 11, 1996 (GOES-7)
- SMS-1, SMS-2, GOES-1 through GOES-7
- Three Modes
  - MODE-A
  - MODE-AA
  - MODE-AAA

Archive recording 1978-1997

U-matic Recorder Archive Hardware (1989)

U-matic
8 hours of data per tape 1,095 tapes per year
(1978-1997)

1997: Switched to 3590

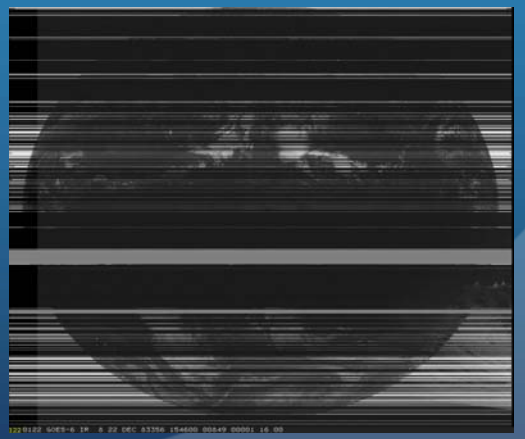
IBM 3590
~2 days of data per tape 182 tapes per year
(1997-2005)

Data Rescue

- 1990s: U-matic tapes were deteriorating
- 1997: Data rescue project began
- Reingest all data from (Early GOES)1978 to 1996 and first few years of GVAR (1994-1997) and put on 3590 tape.
- One copy sent to NCDC in Asheville, NC; one copy kept at SSEC

Data Rescue project completed in 2005

- All data copied to 3590 tapes
- Copy of all tapes sent to NCDC
- Known issues
  - Multiple copies of some images (sometimes more than 8 copies)
  - Significant errors and data dropouts
    - tape degradation
    - recording problems
    - noise in the original signal



On-line archive (2005)

(a little side track)

- 61 TB of disk (RAID-5) enough space for 6 years of GOES-East and West.
- All new GOES data archived on-line (Sep 2005)
- Backed up to lto-2
- Data service to users, went 27 minutes to pull an IR full disk from U-matic – 15 seconds via ADDE, no operator needed!

Wow!

GOES Archive Storage Media History

U-matic
8 hours of data per tape
1,095 tapes per year
(1978-1997)

IBM 3590
~2 days of data per tape
182 tapes per year
(1997-2005)

All data now on-line.
Early GOES data: 1978 to 1996 >100 TB.
All other data: 1994 to present almost 600 TB.
(2005 - Present day)

LTO-5
~135 days of data per tape less than 3 tapes per year
(2010-Present day)

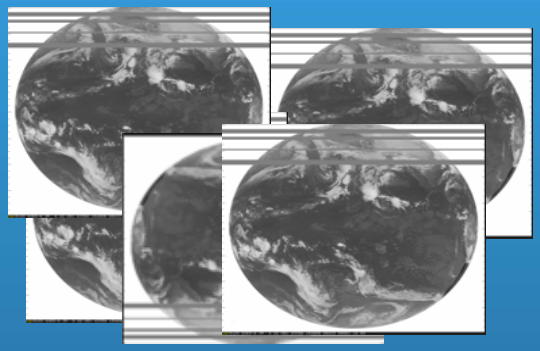
Entire GVAR and International GEO Archive put online (2008)

- Large request by a private company (late 2007)
- User was first sent to NCDC, came back to SSEC
- Funded expansion of the system at SSEC for archiving Geostationary Satellite Data online.

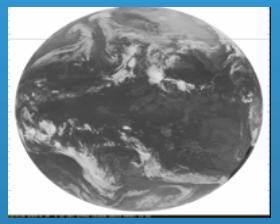
All GOES data online 2009

- By mid 2009 all GOES data from rescue was put on line.
- Data had many issues:
  - Noisy
  - Multiple images
  - Missing images
- Plan was to merge multiple images
- Replay some U-matic tapes, if needed to fix noisy data or look for missing data

Processing Plan



Objectively merge redundant data

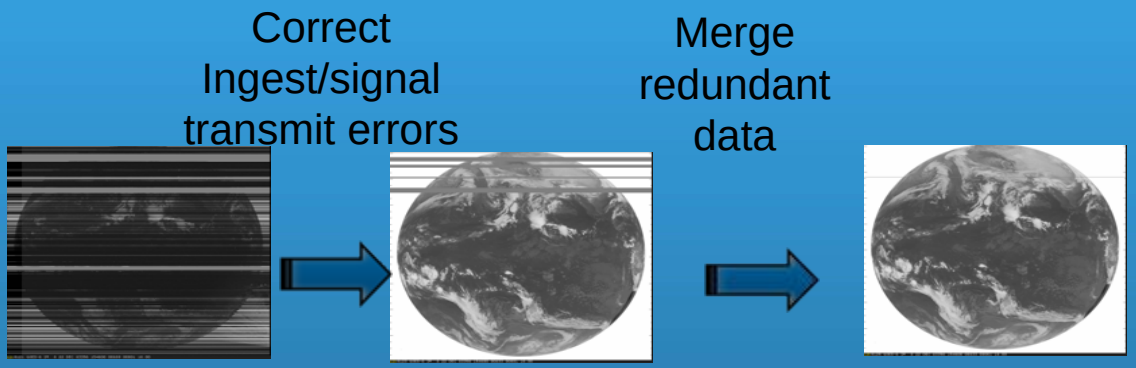


***Get rid of multiple images and merge into a "best" image*.**

Types of Correctable Errors Encountered in Mode-A data

- **Lines deleted/added**: Caused by multiple/bad sync or ground station send errors, sometimes resulted in multiple images in one index.
- **Bit slip**: Bits inserted between sync and type blocks and/or data, probably from when U-matic was first recorded, or may have happened during playback.
- **Fixed type:** IR (0) or VIS (1-8) incorrect, sometimes due to bit slip or multiple/bad sync.
- **Fixed line number:** Line number was changed, caused by bit error.
- **Framing Error**: IR data block size was incorrect.

Interrogation and Reconciliation Process

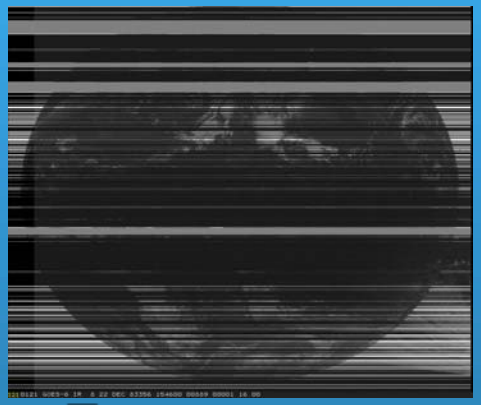


NO data are changed, averaged or otherwise manipulated. All data are uncovered by correcting errors made by earlier software deficiencies.

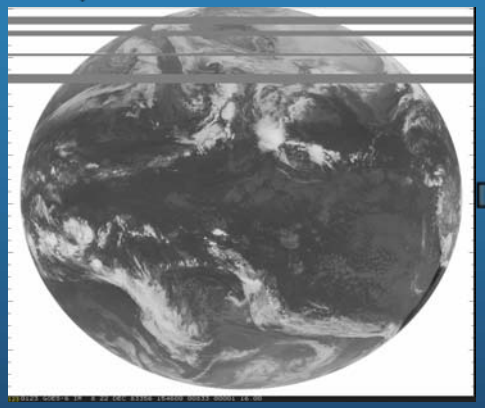
Processing Results

GOES-6 Dec 22, 1983 15:46 UTC

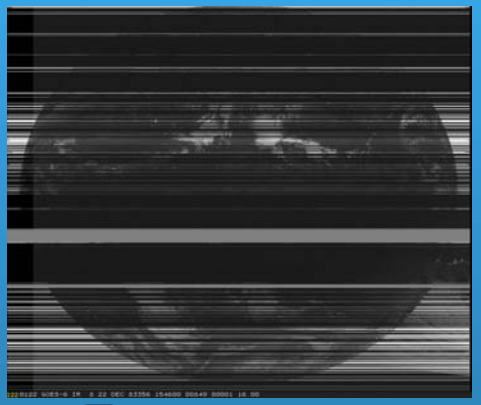
Original playback from U-matic tape



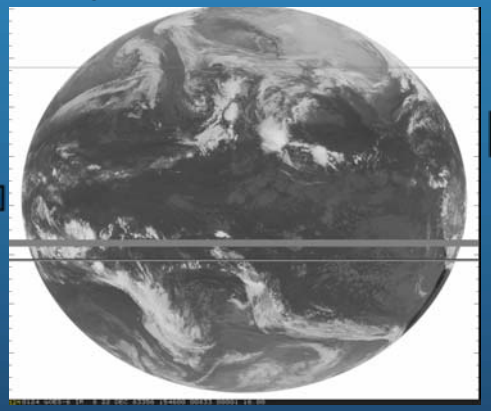
Processed

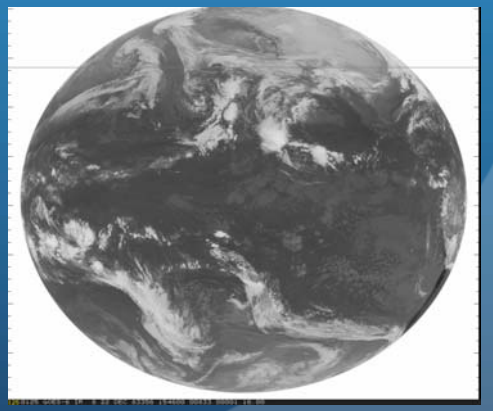


2nd attempt playback from U-matic tape



Processed

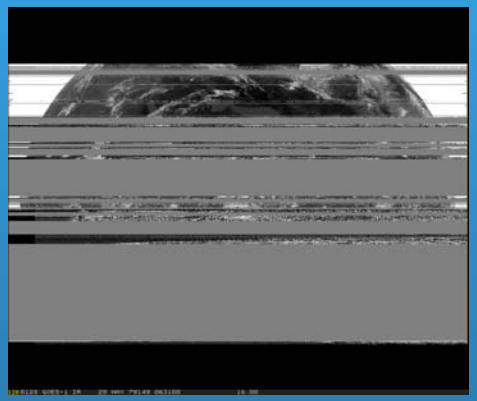




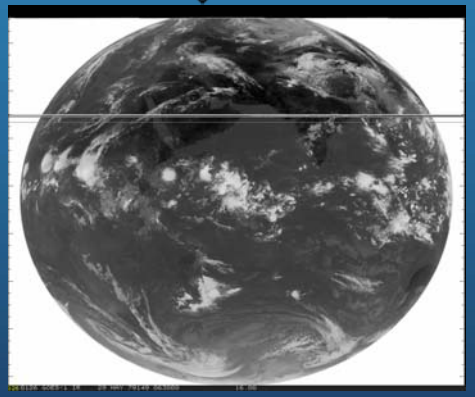
Processing results

GOES-1 May 29, 1979 06:30 UTC

Original playback from U-matic Tape



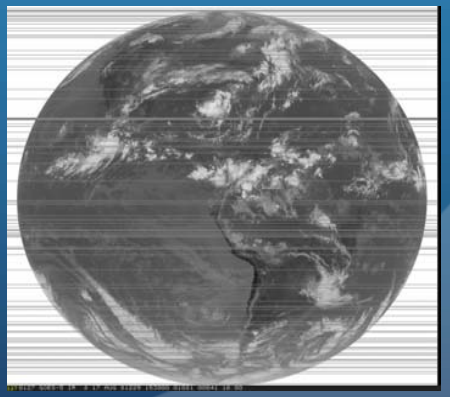
Processed



GOES-5 Aug 17, 1981 15:30 UTC

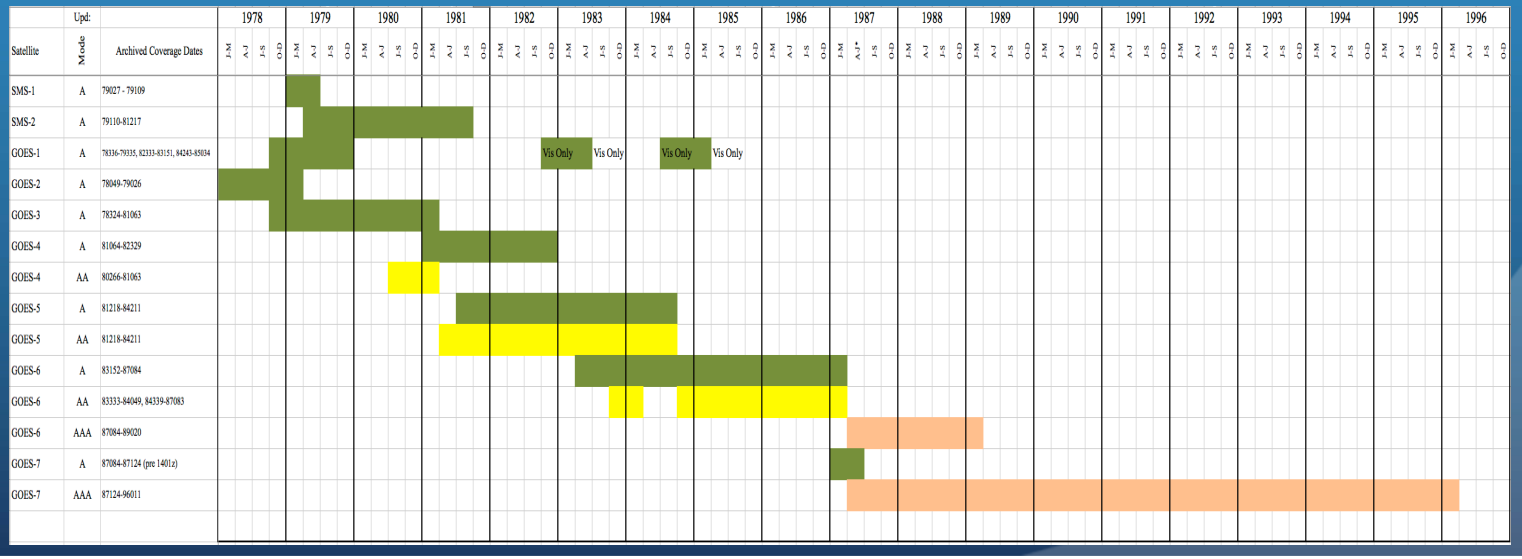
Original Playback from U-matic

Processed



Processing Statistics

- There are over **717,000** images archived from the operational GOES satellites during the period 1978 to 1996.
- All Mode-A data (over **306,470** images) have now been processed.



Processing Statistics

- **~2,800** Mode-A images thought to be completely lost were recovered
- **~8,100** images had corrections to at least 95% of their image lines
- **~25,500** images had framing errors corrected that affected every visible scan in an image
- *Totals above represent nearly 2 full years of data!*

Archive work in progress

- Complete MODE-A work
- Complete MODE-AA and MODE-AAA
- Improve ADDE Server Performance
- Update metadata database interface

End