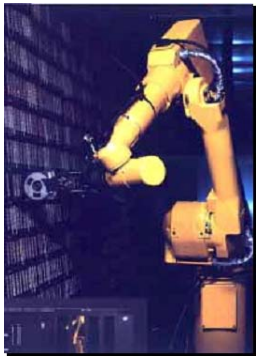
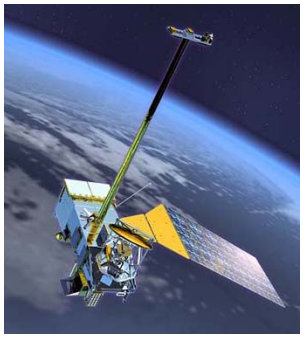
The Grid: an IT Infrastructure for NOAA in the 21st Century

Mark Govett
NOAA
Forecast Systems Laboratory

Data Systems

- Massive increase in the type and volume of data
  - Observations
    - Global: Satellite, RAOBS,
    - National: Radar, Aircraft, Ship, Buoy, Surface, Precipitation,
    - Regional: Profiler, observers
    - Local: Transportation, Utilities
  - Modeling Systems
    - Global
    - Regional
    - Local
  - Archival
    - NESDIS Data Centers
    - Local Data Archive

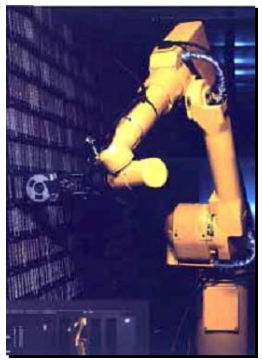


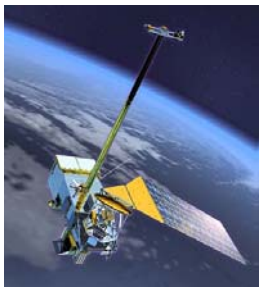




Data Systems

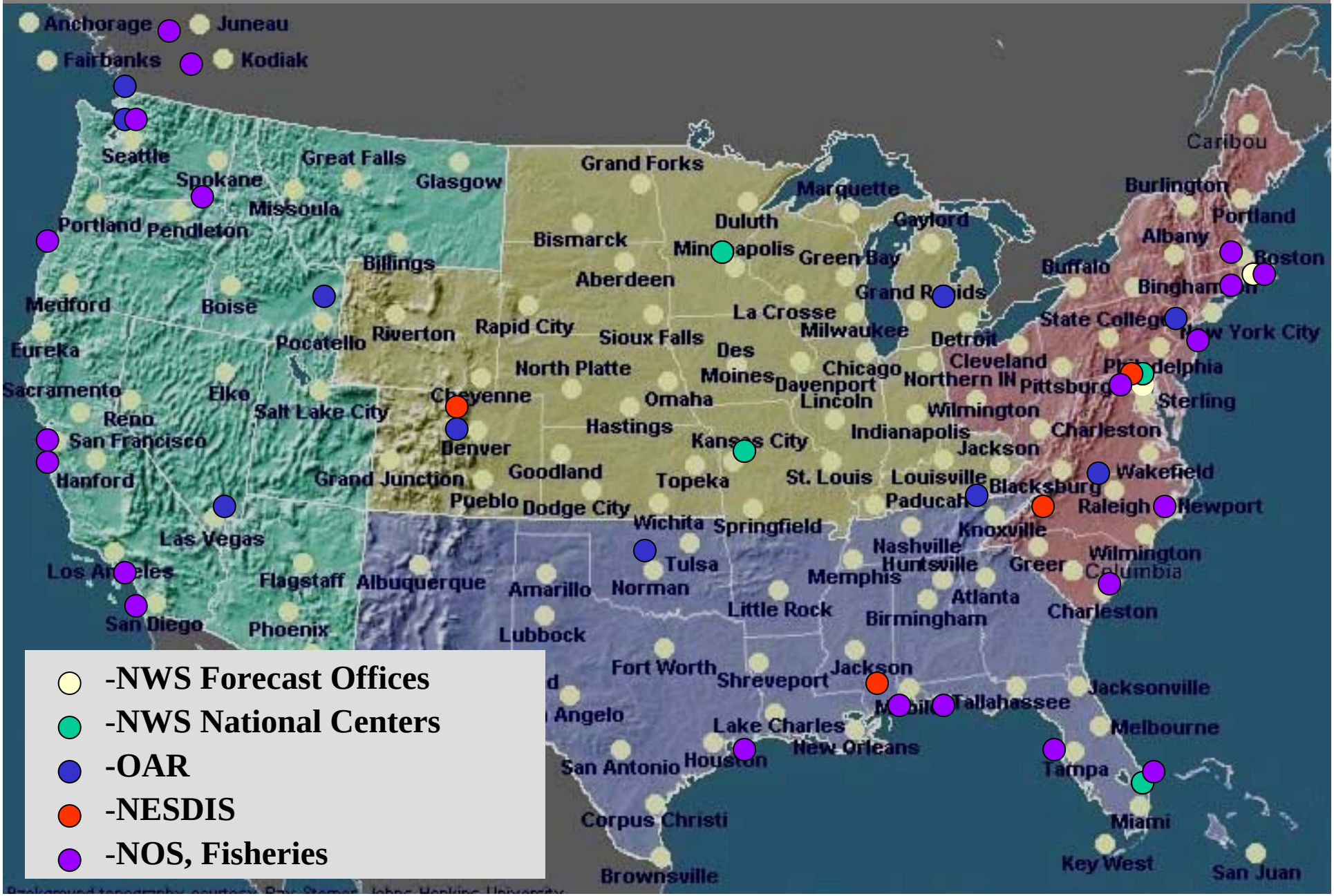
- Dissemination to:
  - NOAA
  - Government Agencies
  - Commercial & Universities
  - General Public
- By 2004, NOAA will ingest and process more new data in one year than was contained in the total digital archive in 1998







NOAA Facilities

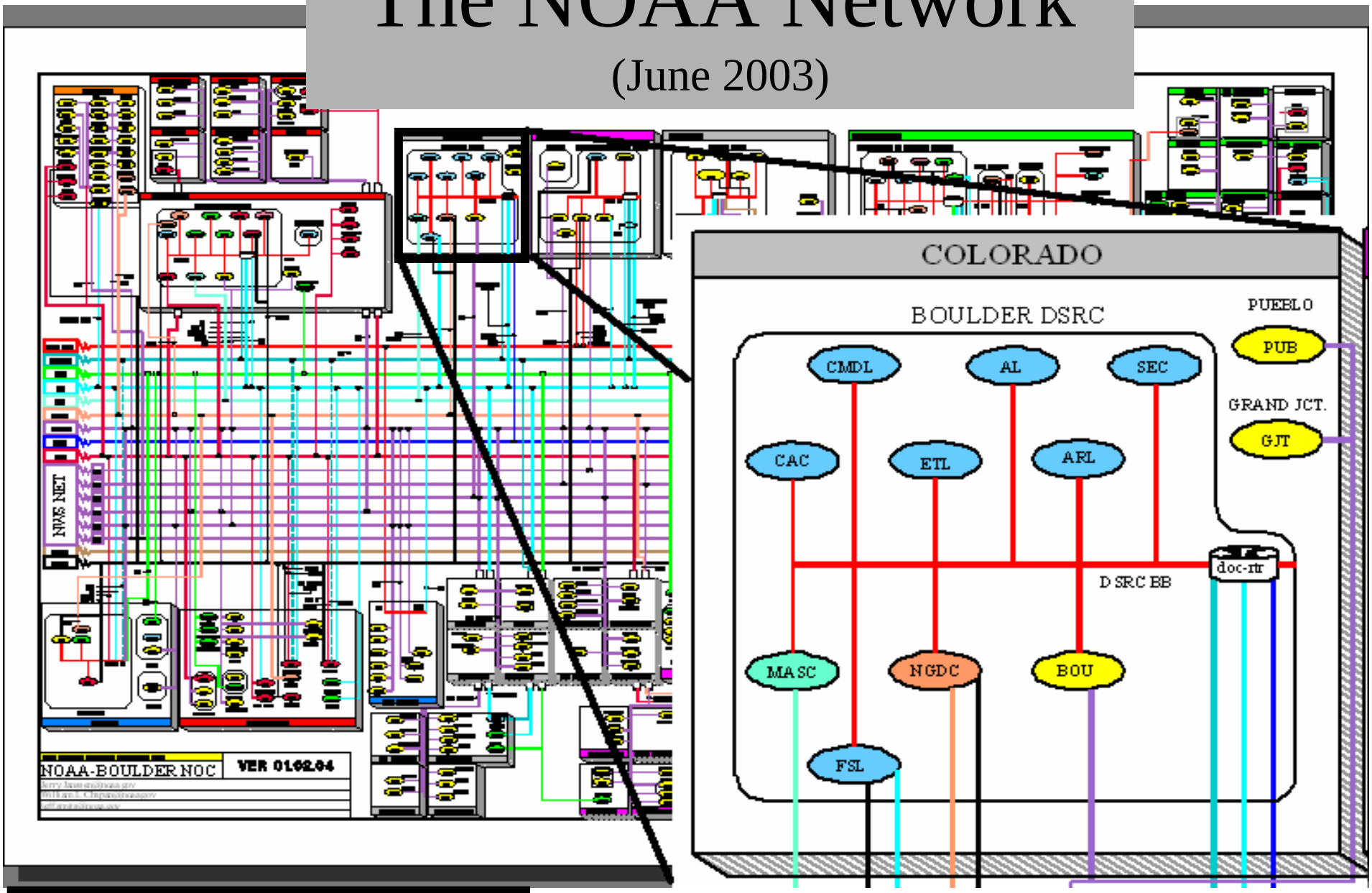


Data Distribution via (NOAAPORT)

- Benefits
  - Broadband distribution of data
  - Cost effective
- Limitations
  - Scalability
  - Forecasters need more data
- Led to increased use of the Internet

The NOAA Network

(June 2003)



Courtesy of Jerry Janssen, NOC

NOAA's IT Infrastructure

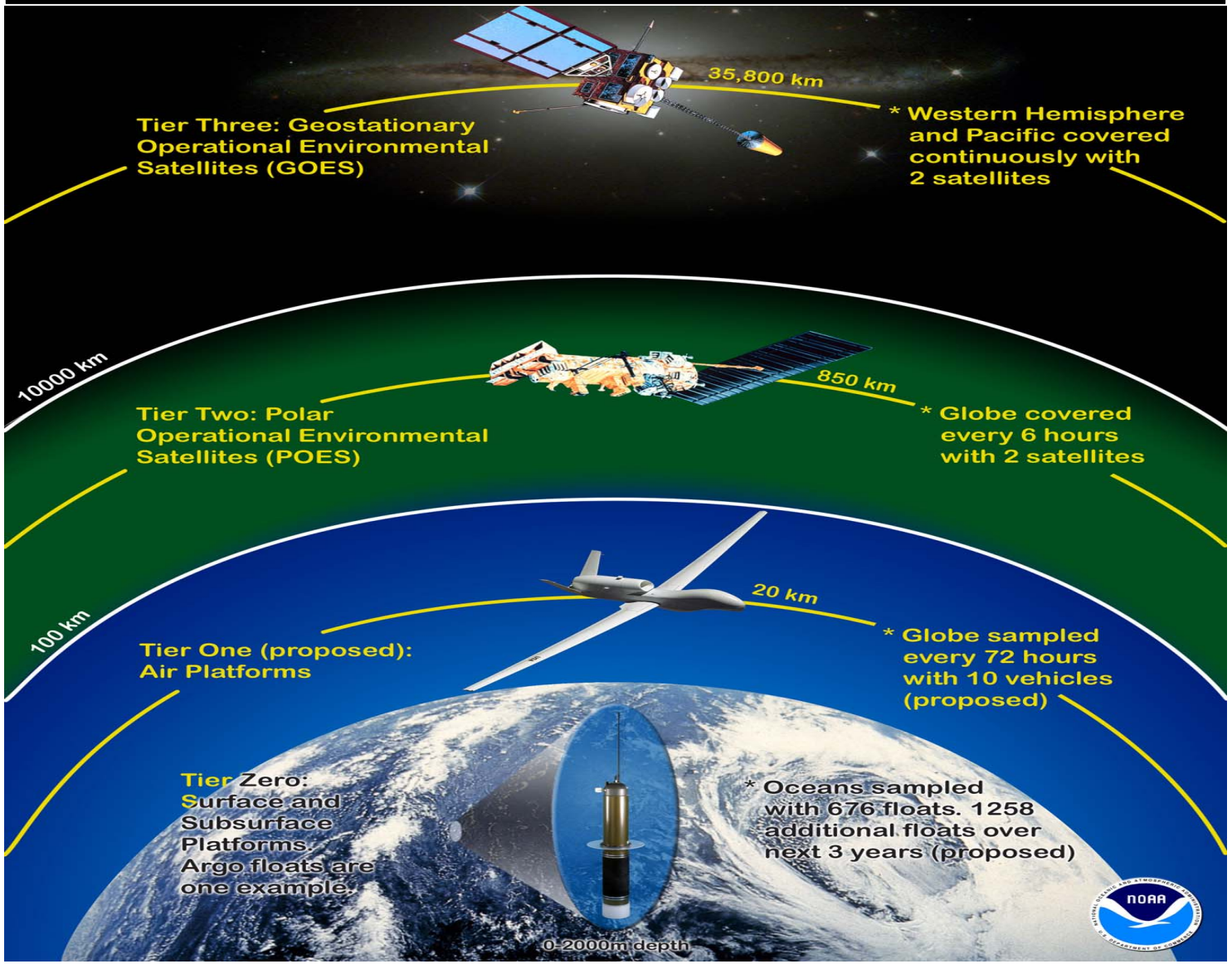
- Built up over time to support programs and requirements
- Independent systems
  - Inefficient
  - Costly
- cannot not sustain future growth without incurring huge costs

An IT Infrastructure for the next decade and beyond

- "*End goal - A Fully wired, networked and integrated system that provides for data processing, distribution, & archiving*"

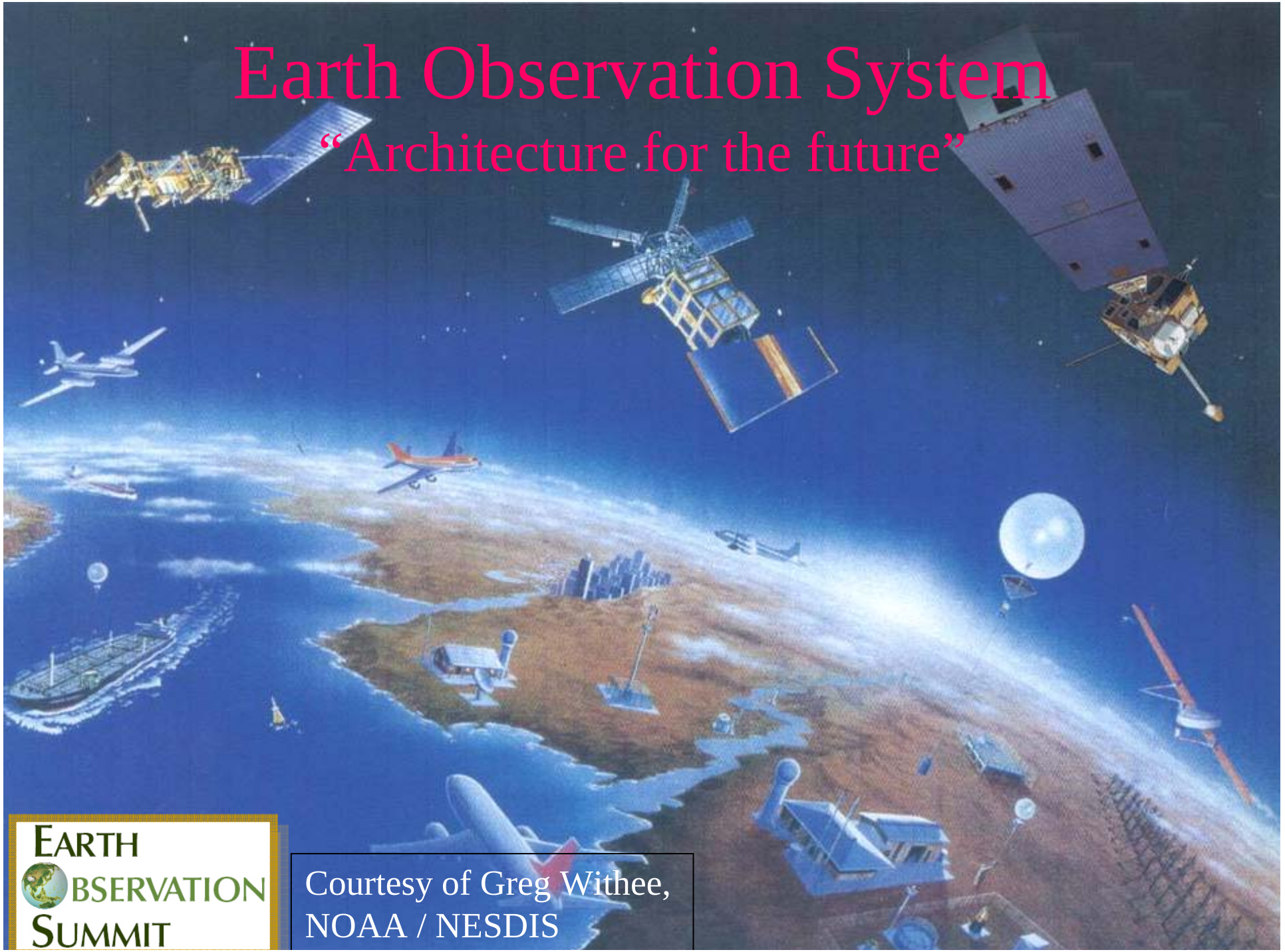
  - C. Lautenbacher, NOAA Administrator

NOAA Observing System Architecture



Earth Observation System

"Architecture for the future"



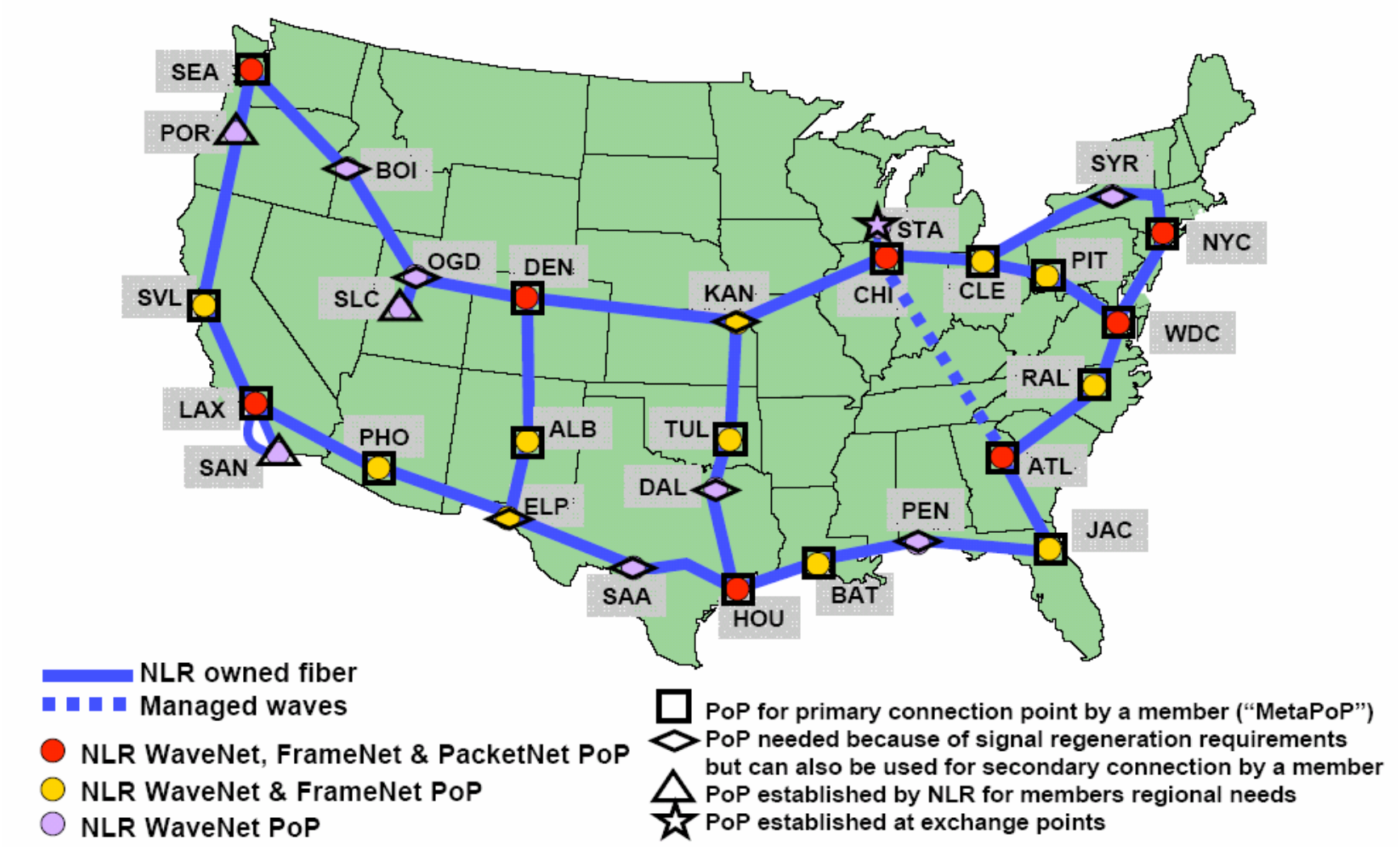
EARTH OBSERVATION SUMMIT

Courtesy of Greg Withee,
NOAA / NESDIS

Grids at NOAA: A Vision

- 21st Century IT Infrastructure for NOAA
  - scalable, adaptable, efficient, secure
- NOAA is already moving toward building a more cost effective IT Infrastructure
  - Address the Network
  - Distributed remote data access
    - NOMADS, CLASS
  - Distributed computing environment

National Lambda Rail



Data Grids

- Continue to utilize NOAAPORT for general distribution but increase internet use
- Use grid mechanisms to provide and control access to data
  - If users can’t locate, access or utilize data it’s value is diminished
- Build demand-based data delivery systems
- Exploit grid mechanisms to discover, access, transfer, and utilize data

Compute Grids

- View independent resources as a collective NOAA resource
- Meta-scheduling across NOAA HPC systems
- Create grids to serve virtual communities

Possible NOAA Grids
(serving Virtual Organizations)

Data Grids

Data Archives
- **NGDC**
- **NCDC**
- **NODC**

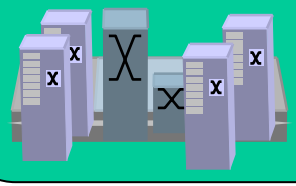
Satellite Data
- **NASA**
- **DoD**
- **OSO**

Compute Grids

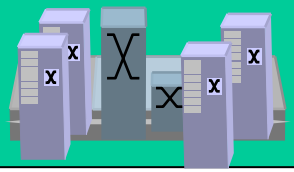
Meso-scale Modeling
- NCEP
- NCAR
- FSL

NWS Central Region
- NCEP
- Forecast Offices

Hardware Infrastructure



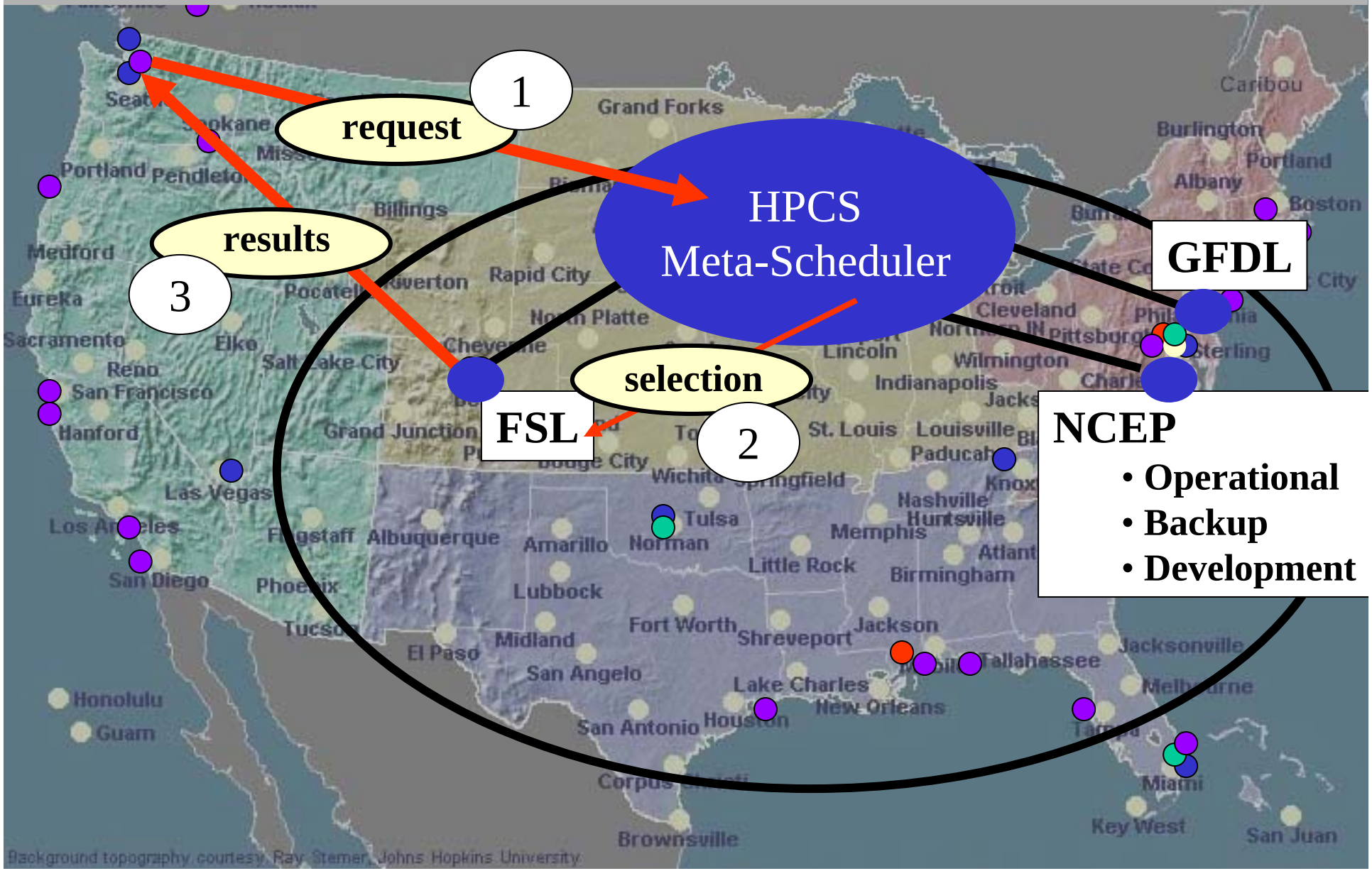
network



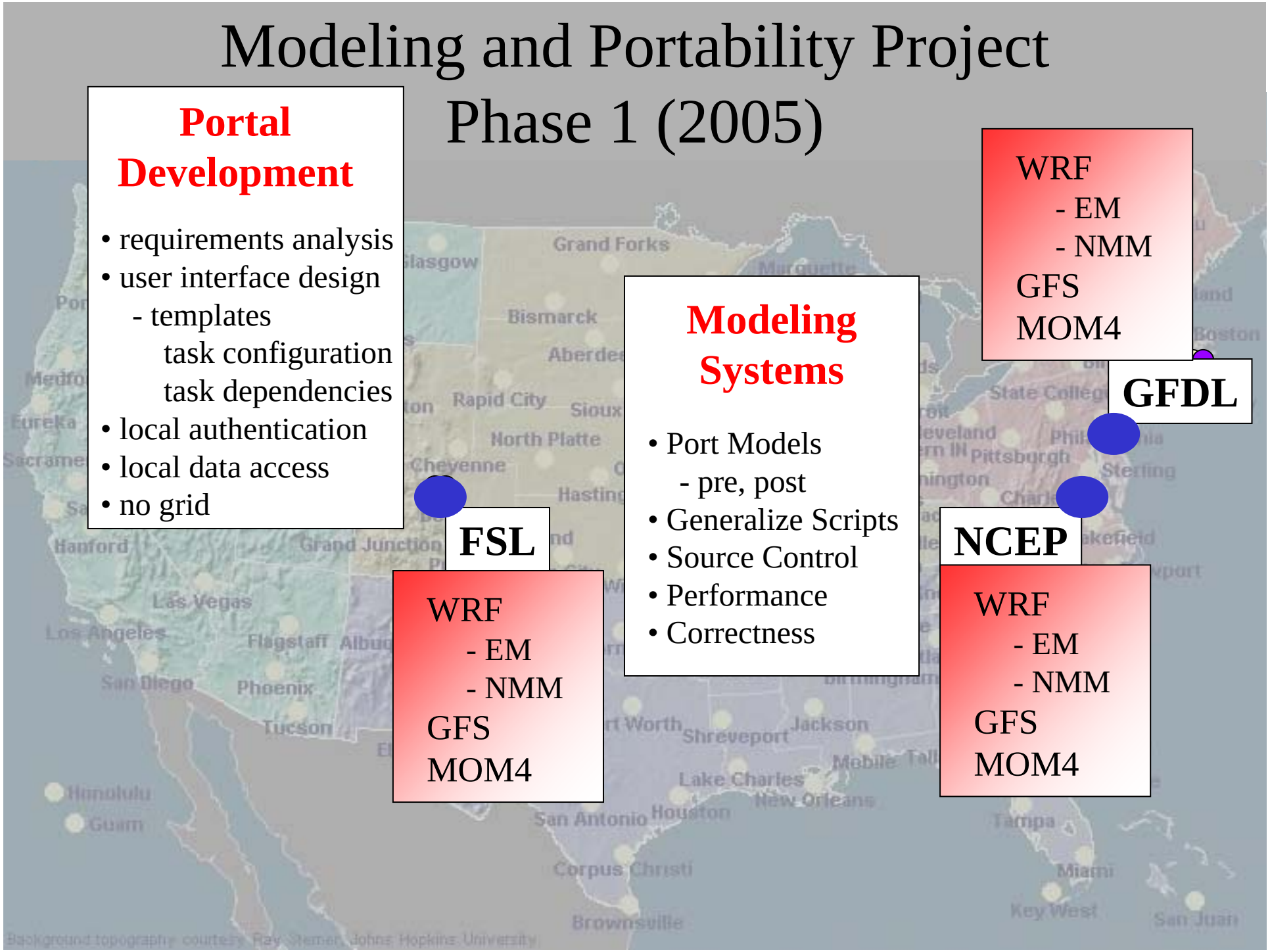
Moving from Vision to Reality

- Current Grid Projects
  - Data on demand
    - NOMADS & FSL
  - Building a NOAA Meta-scheduler
    - GFDL, PMEL, FSL
  - Web-based grid-enabled portal
    - FSL, NSF

A NOAA Meta-Scheduler



Modeling and Portability Project Phase 1 (2005)



Portal Development

- requirements analysis
- user interface design
  - templates
    - task configuration
    - task dependencies
- local authentication
- local data access
- no grid

Modeling Systems

- Port Models
  - pre, post
- Generalize Scripts
- Source Control
- Performance
- Correctness

GFDL

WRF
- EM
- NMM

GFS
MOM4

FSL

WRF
- EM
- NMM

GFS
MOM4

NCEP

WRF
- EM
- NMM

GFS
MOM4

Modeling and Portability Project Phase 2 (2006)