

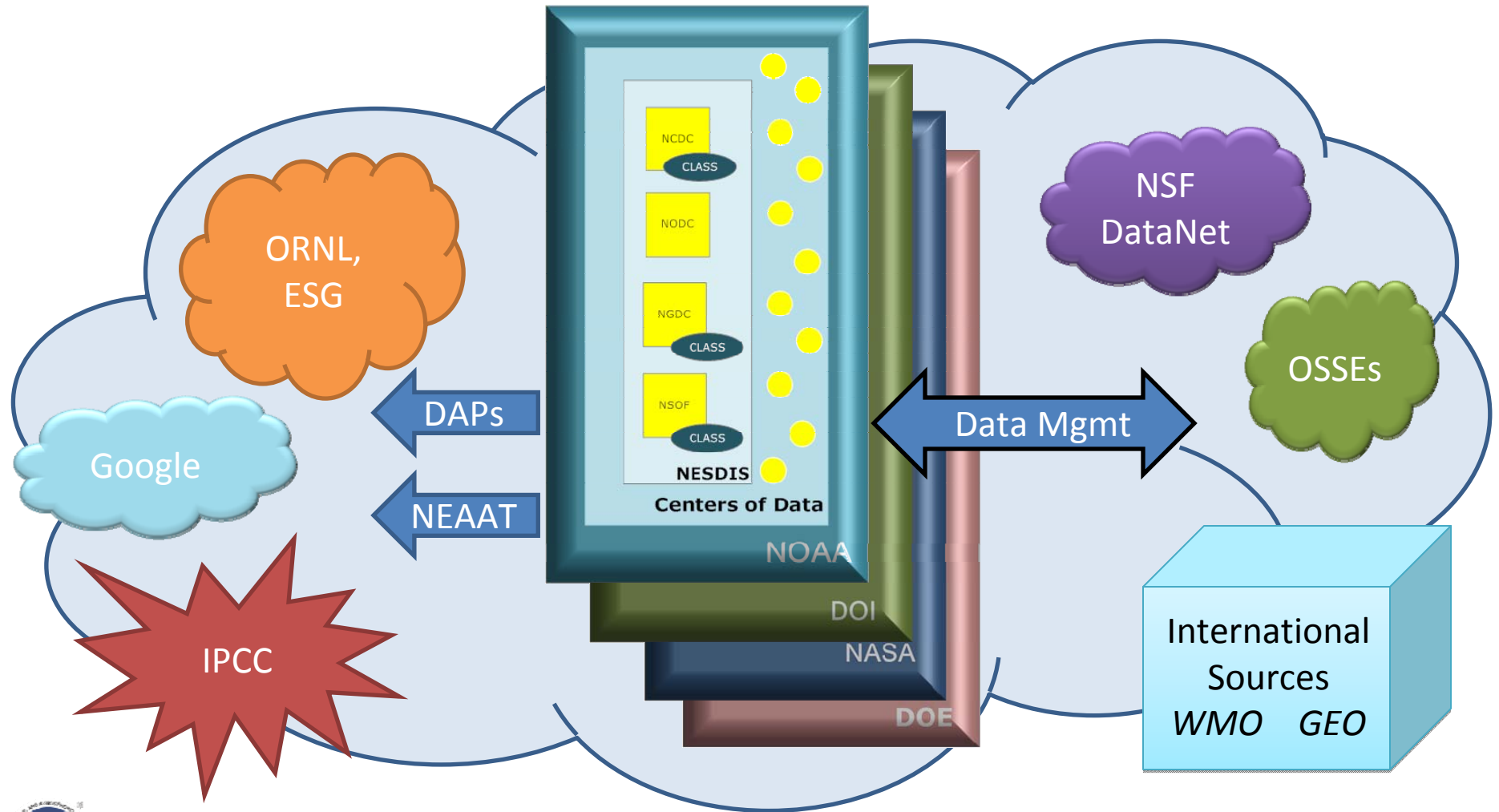
# Vision of Shared Services within NOAA's Data Centers

Nancy Ritchey  
Archive Branch Chief  
National Climatic Data Center

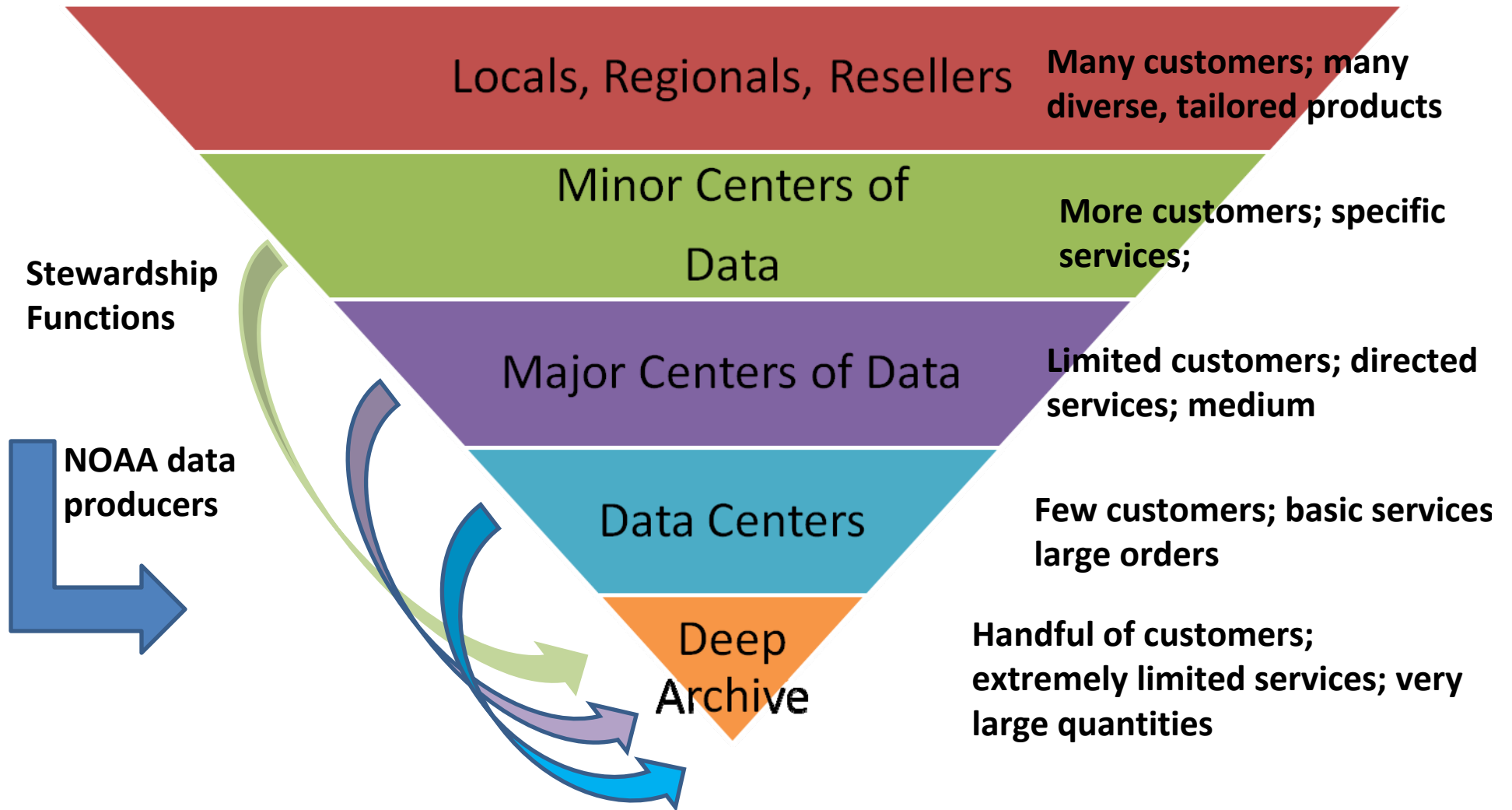
- Federation of Data Centers with NOAA's many Centers of Data, as well as other Agency/Department's data grids
- Utilize rules-based federation to enable:
  - Data conversions or production on the fly
  - Reduced bandwidth and number of datasets to store
  - Full utilization of NOAA and/or commodity computing resources
- Use of workflows to move data in/out of archive, to users, etc.



# NOAA's Data Systems Function in a Wide Information Landscape



# Tiered Services: Data



# What's *THE* problem?

- Data volume doubling every 2 years?
  - Moore's Law: processor power doubles each 1.5 years
  - Data storage density similarly doubling every 1.5 years
  - Not *THE*, maybe *A* problem: flat or flatter real costs over time?
- Data Heterogeneity? Expansive, but a one-time effort?
- Scientific Data Stewardship
  - Labor intensive - expensive
  - Recurring: doesn't decrease exponentially with time?
  - Required for each dataset archived
  - THE* problem for the next decade+



# What's a "Federation"?

- Standards- and services-based framework which exploits multiple authoritative data sources that are separately administered
- Each element can securely access data and metadata throughout the federation
- Often no need to move data around, just move the metadata
- Define workflows to move/distribute/update data and utilize federated resources



# How does a Federation help with the problems?

- Can leave the “active” and most easily accessible copy of the data with the Data Steward, in a specific service Tier, or out on the Cloud
  - Simplifies stewardship functions
  - Reduces storage and bandwidth requirements
- Reduces load on Tier 0 resources (deep archive)
- Workflows can be defined to process or otherwise change the data on-the-fly to meet specific member needs



# How to employ federation in NOAA

- Deploy rules-based distributed data management software to the authoritative data sources
- Provide tiered services, adopt common standards
- If secure storage is desired, enable coordinated ingest and archive of data from the data steward to lower tier resources
- Establish workflows to/from data sources
  - Plug into cloud processing and storage resources
- Provide virtual data collections to users for search+



# NOAA's Data Systems Function in a Wide Information Landscape

