

# NASA EOSDIS\* Evolution in the BigData Era

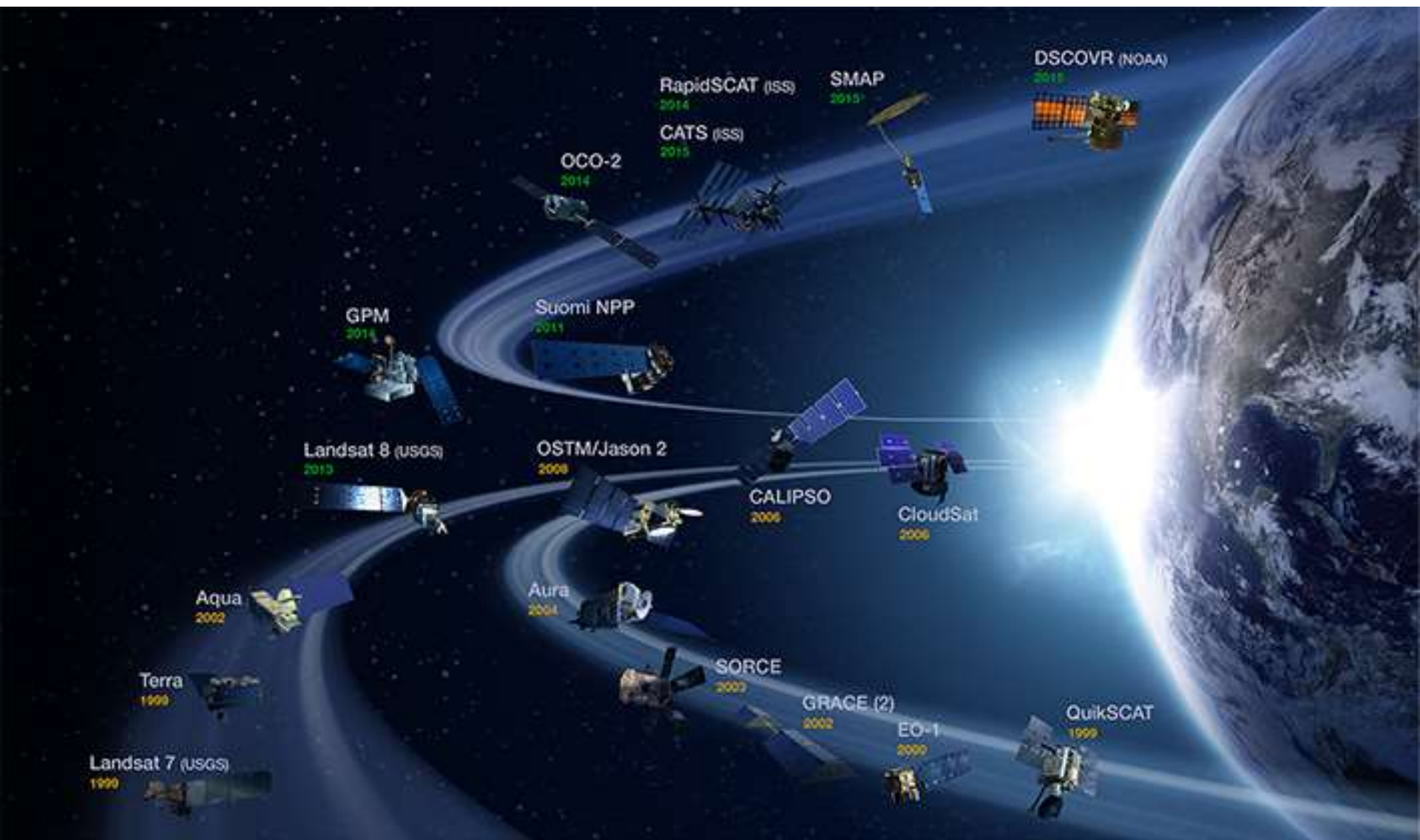
\*Earth Observing System Data and Information System



Christopher Lynnes  
EOSDIS Systems Architect  
Code 586, NASA/GSFC  
HPC Forum 2015



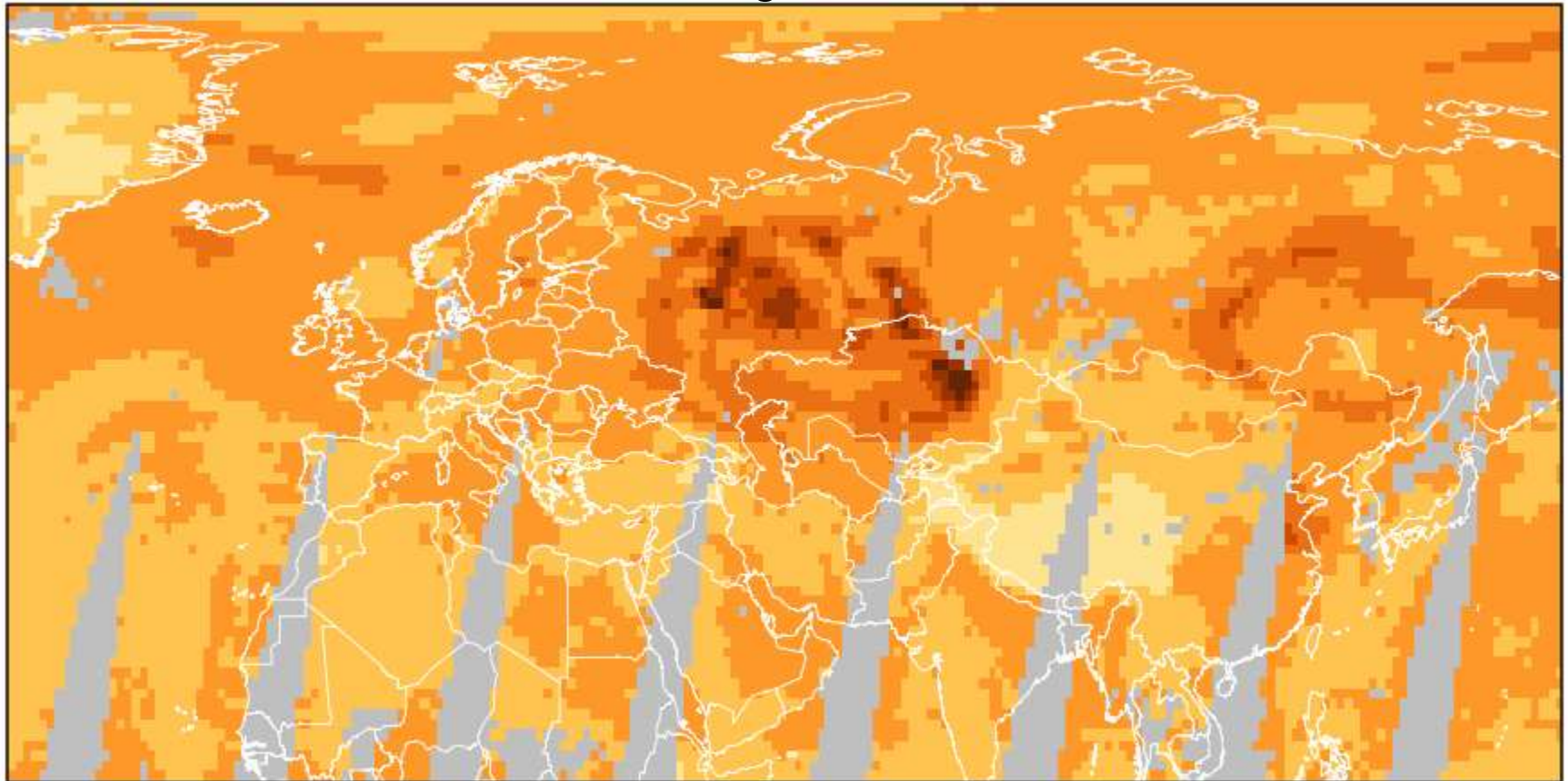
# *EOSDIS processes, archives and distributes data from Earth observing satellites*



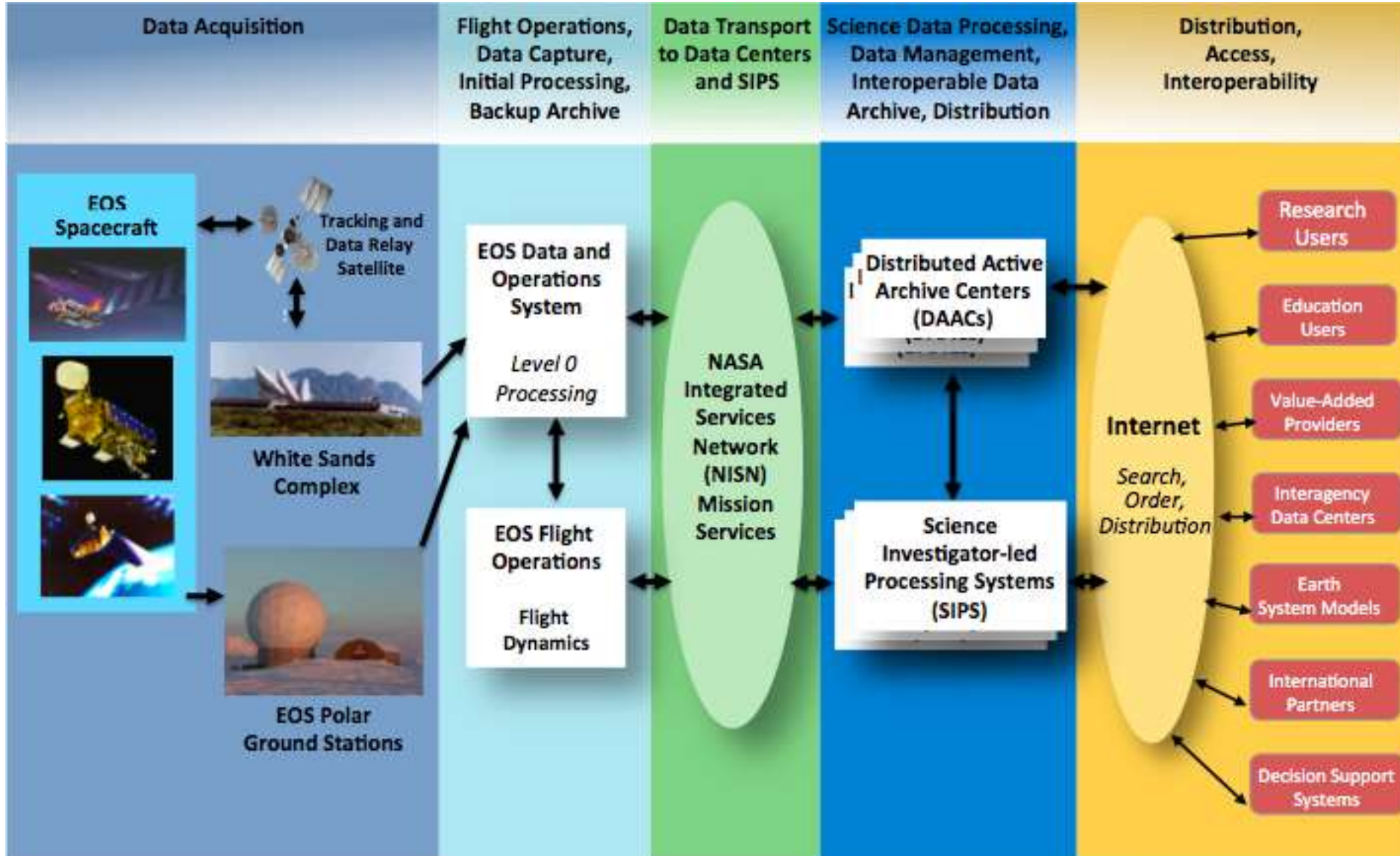
# Example: Atmospheric Infrared Sounder (Aqua Satellite)



Total Column Carbon Monoxide, Night  
11 Aug 2011



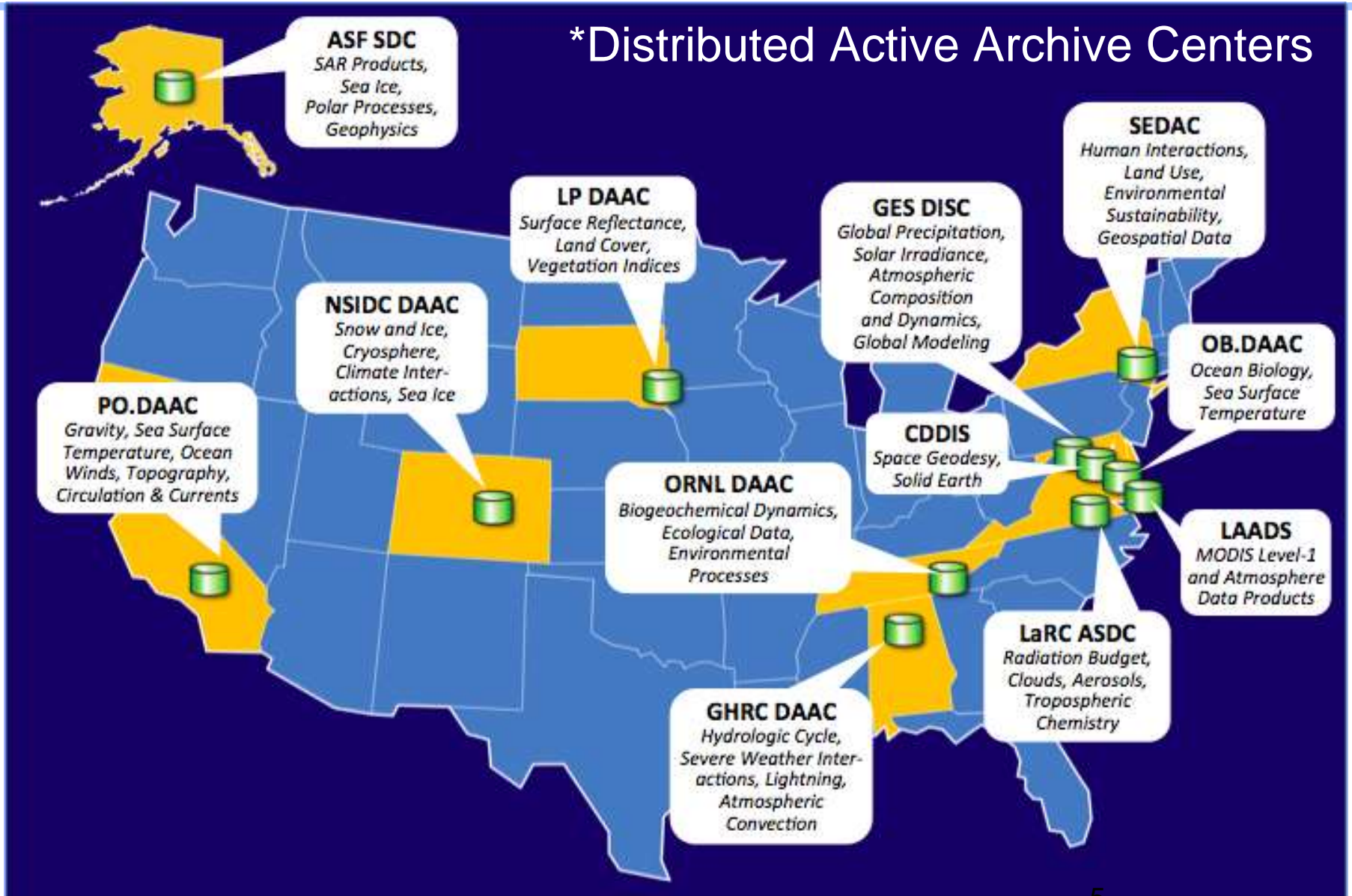
# EOSDIS manages data from downlink to distribution



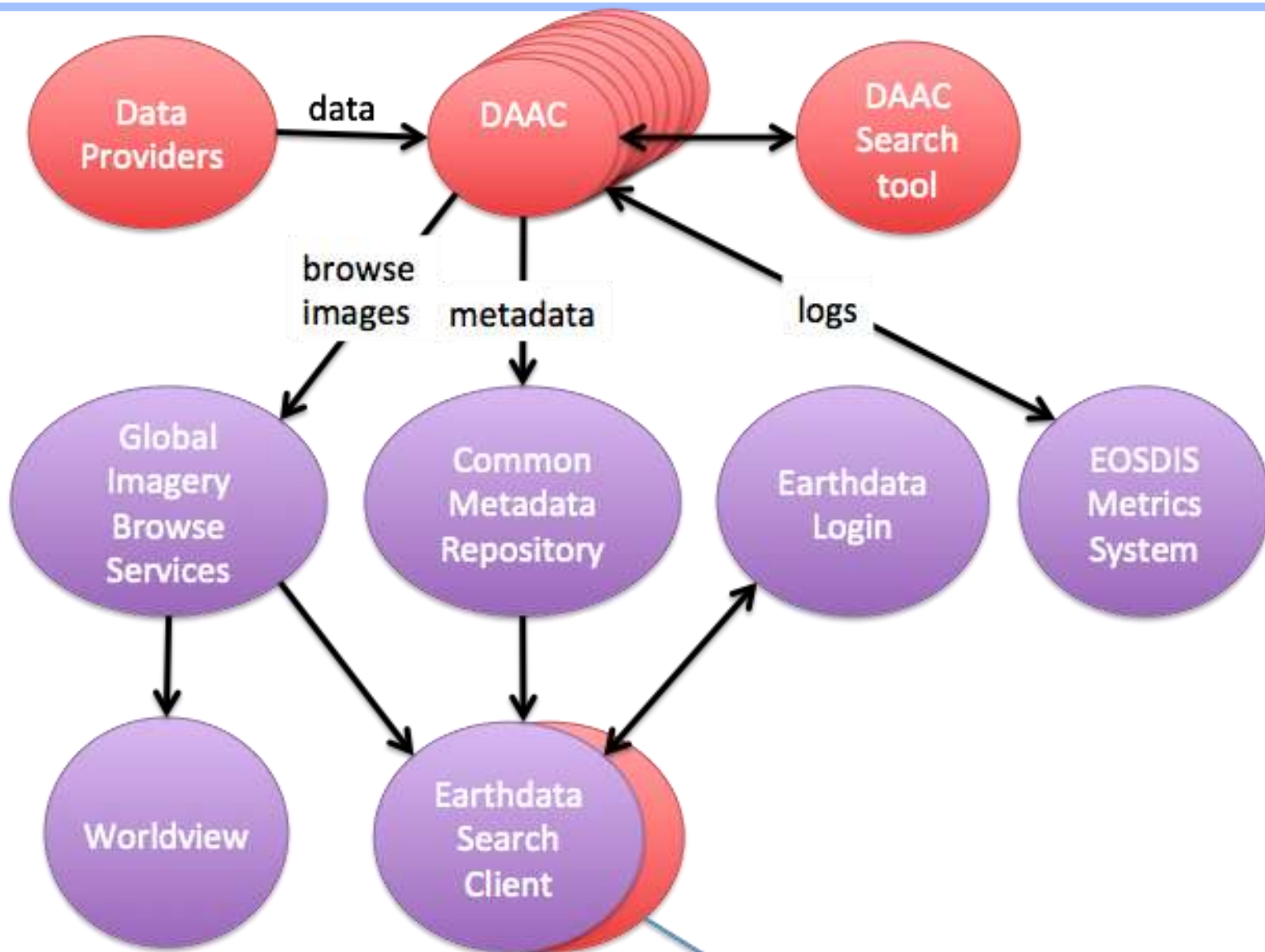
# Data are archived and distributed by DAACs oriented around science disciplines



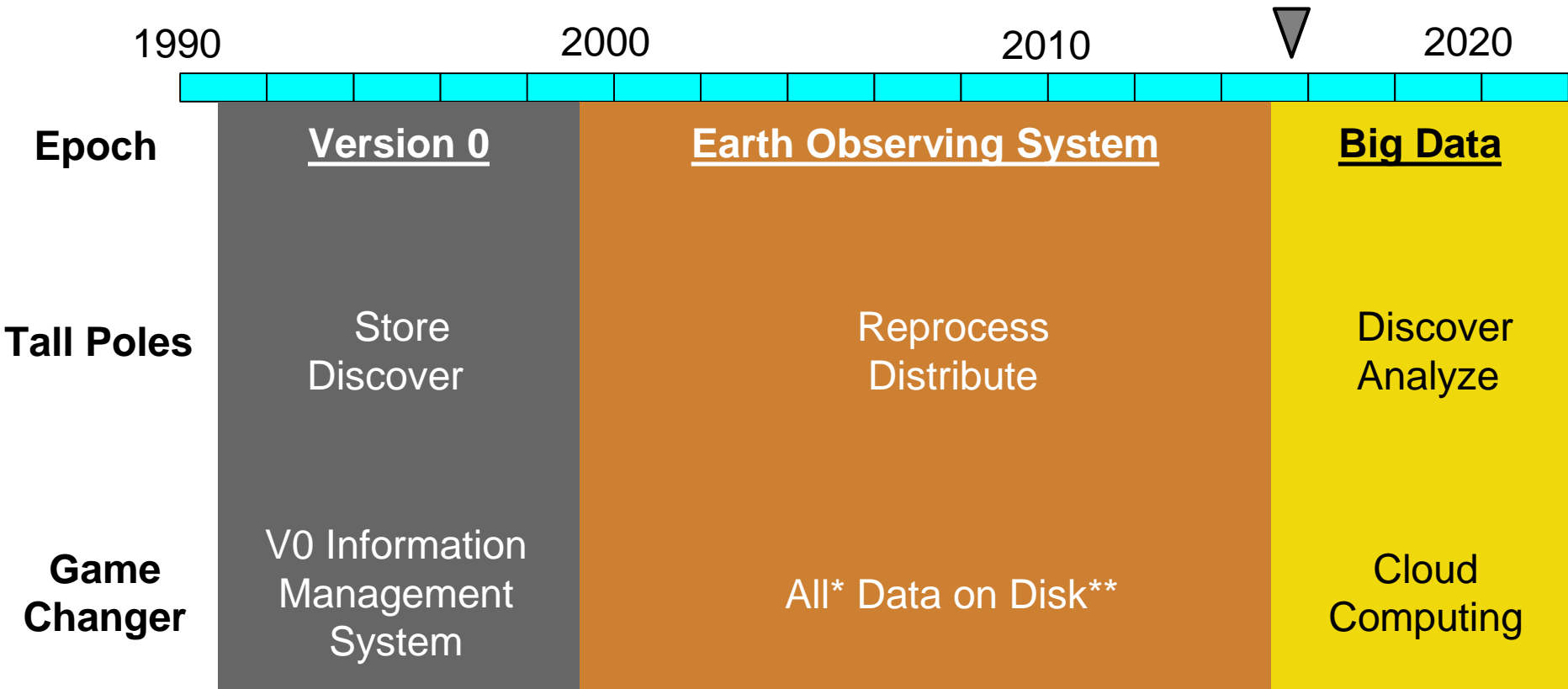
## \*Distributed Active Archive Centers



# DAACs and users are supported by EOSDIS Common Services



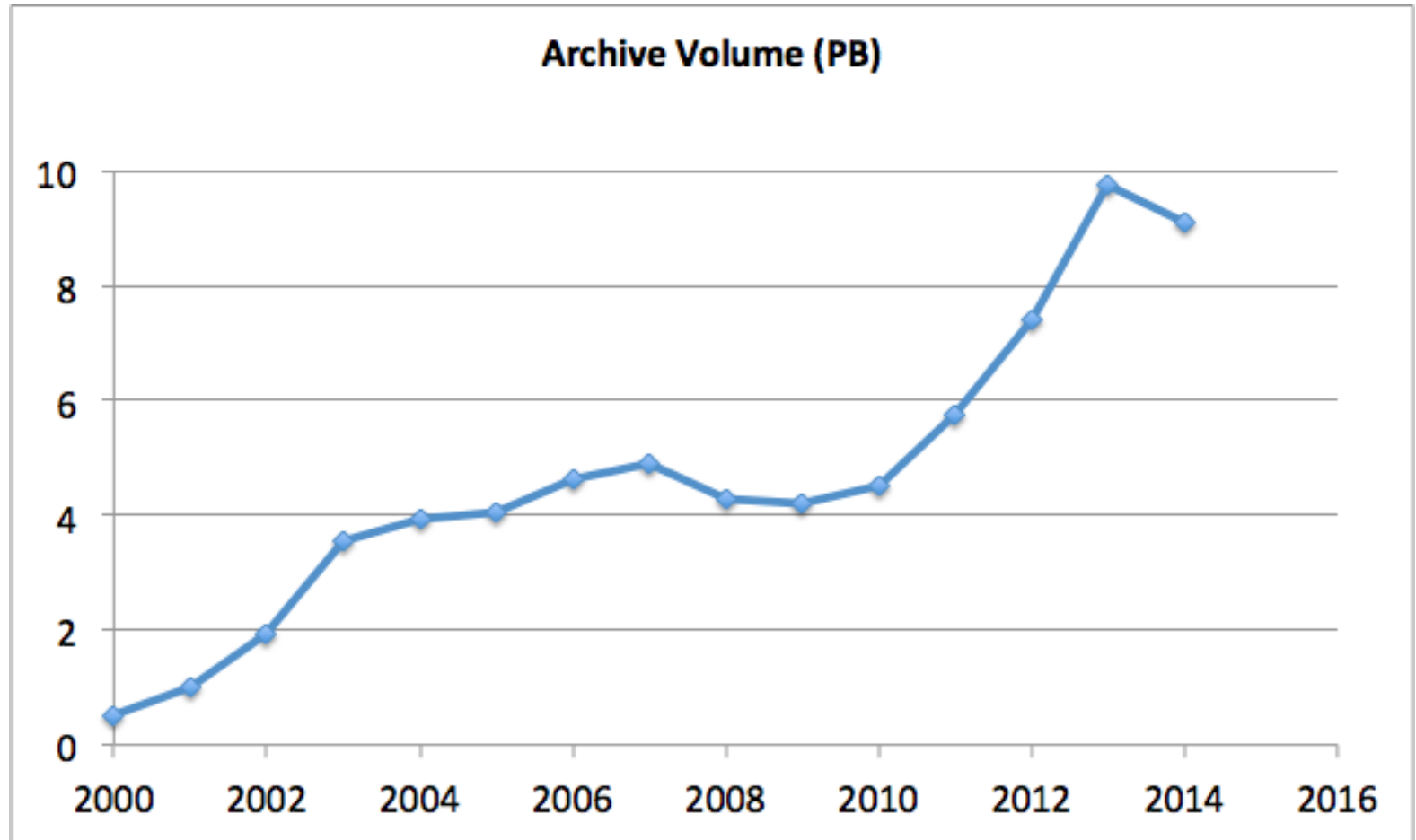
# EOSDIS Evolves Continually



\*Almost

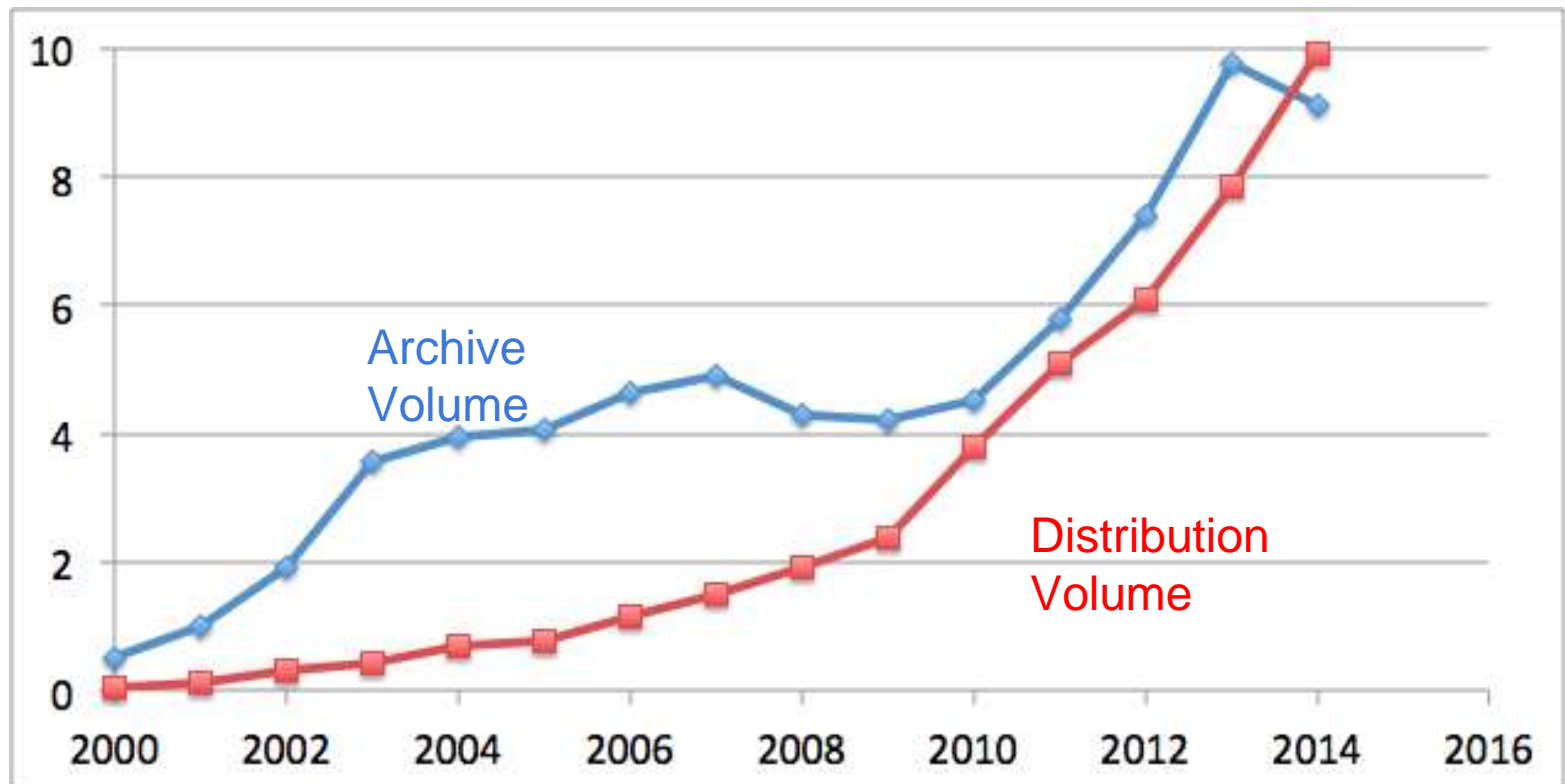
\*\*Thank you, HDF internal compression!

# Big Data Volume Growth

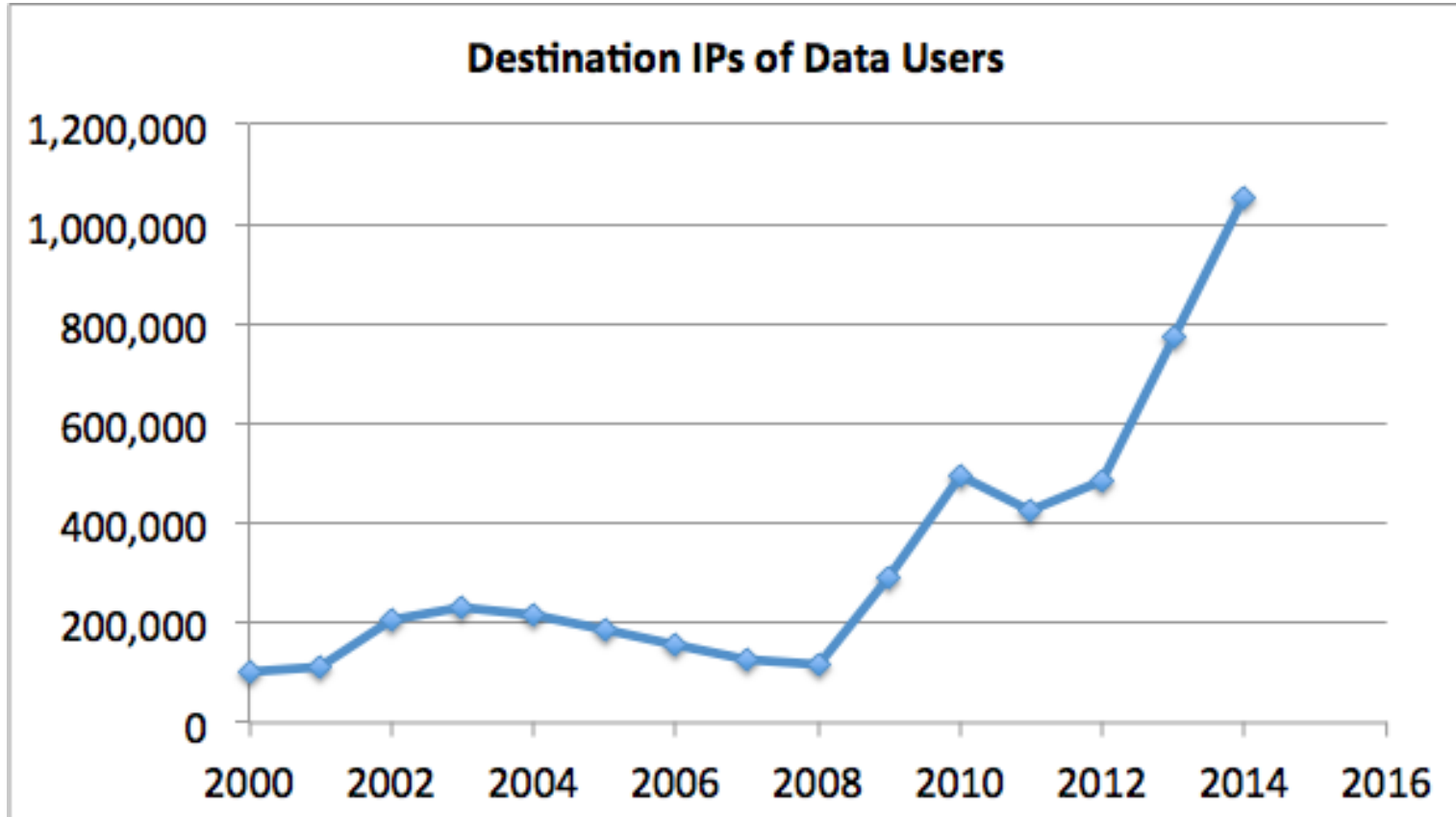




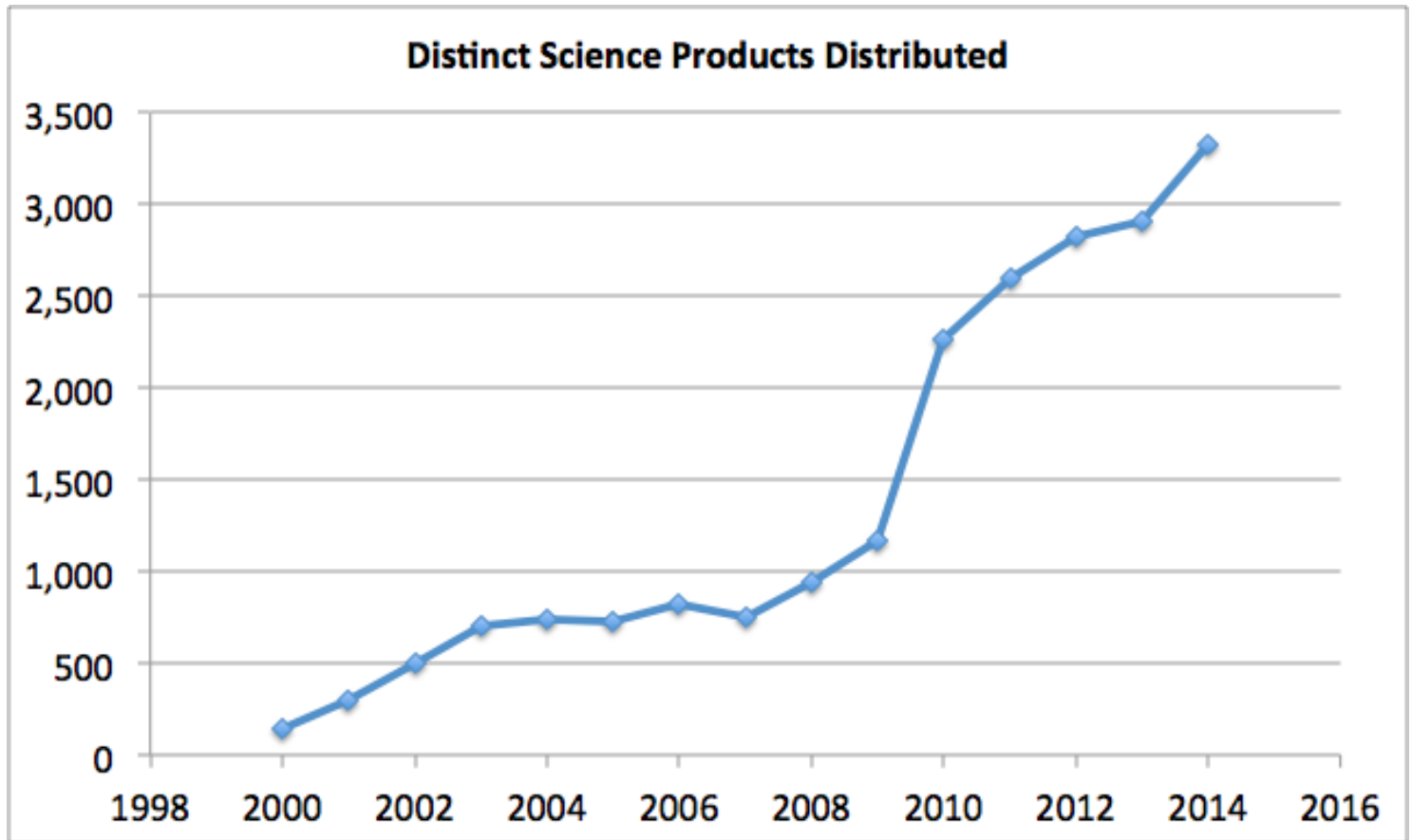
# Big Data Distribution Growth



# Big Data User Growth



# Big Data Variety Growth





EOSDIS in the Big Data epoch will  
*enable* more analysis closer to the data.

# Let's Break that Down...

---



**“more analysis** closer to the data”

# “More Analysis”



*More Complexity*



Subset

**Data Variable**      **Spatial Area**      **Quality Filter**

Transform

**Reprojection**      **Mosaicking**

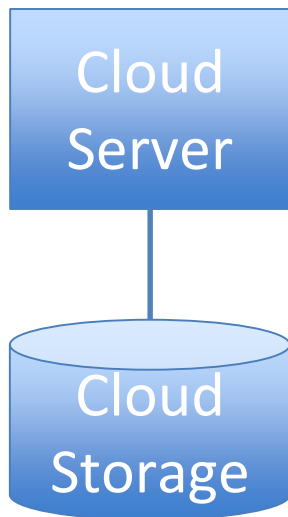
Analyze

**Simple Stats**      **Complex Stats**      **End User's Algorithm**

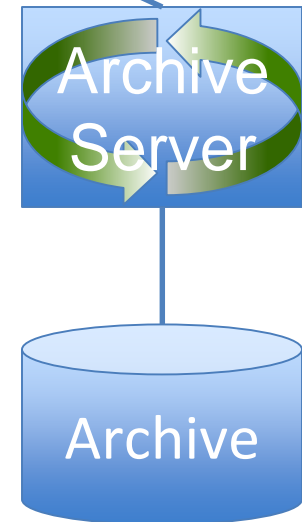


“more analysis **closer to** the data”

# “Close To” = At Archive

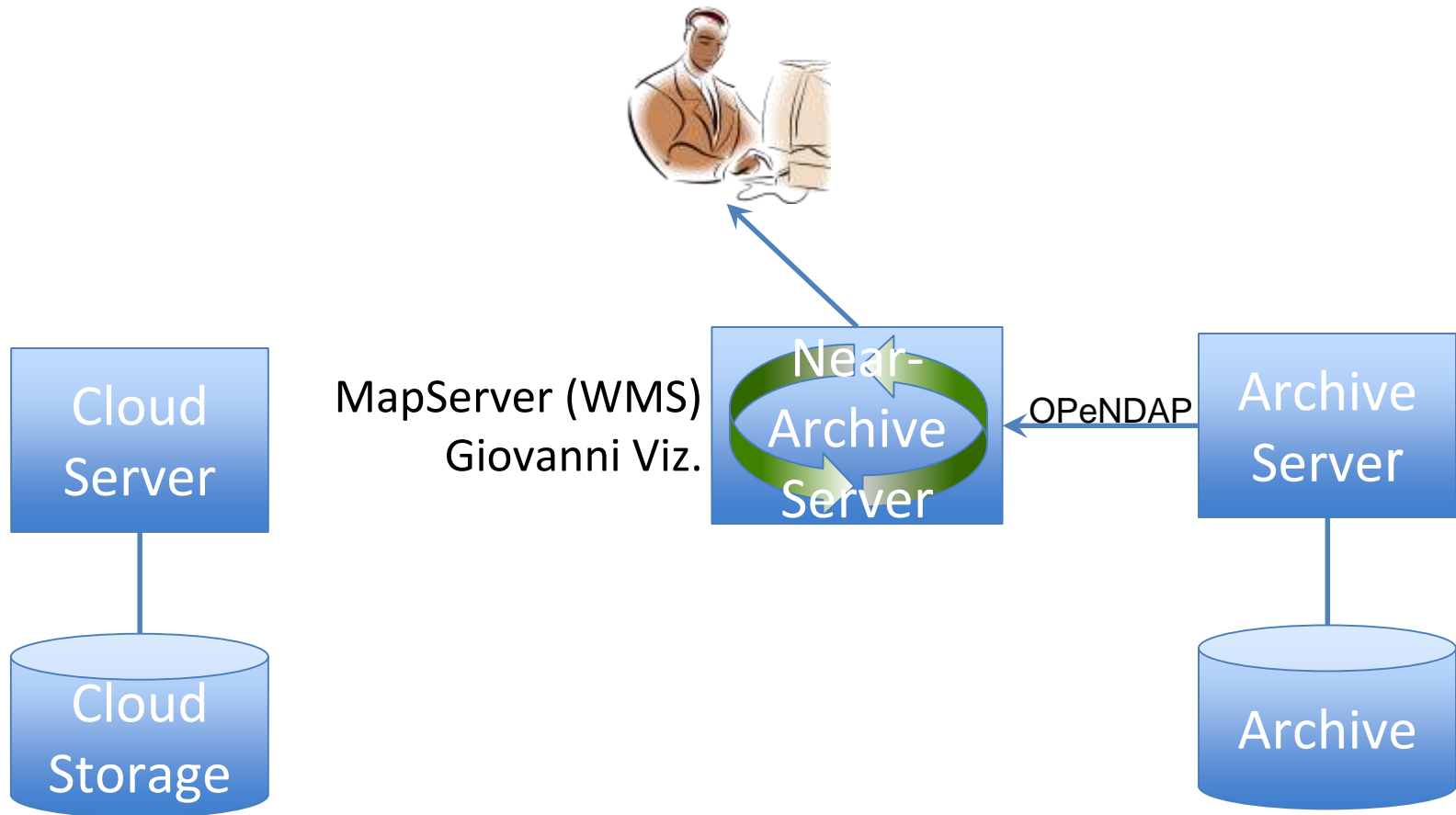


GrADS Data Server  
ArcGIS for Server

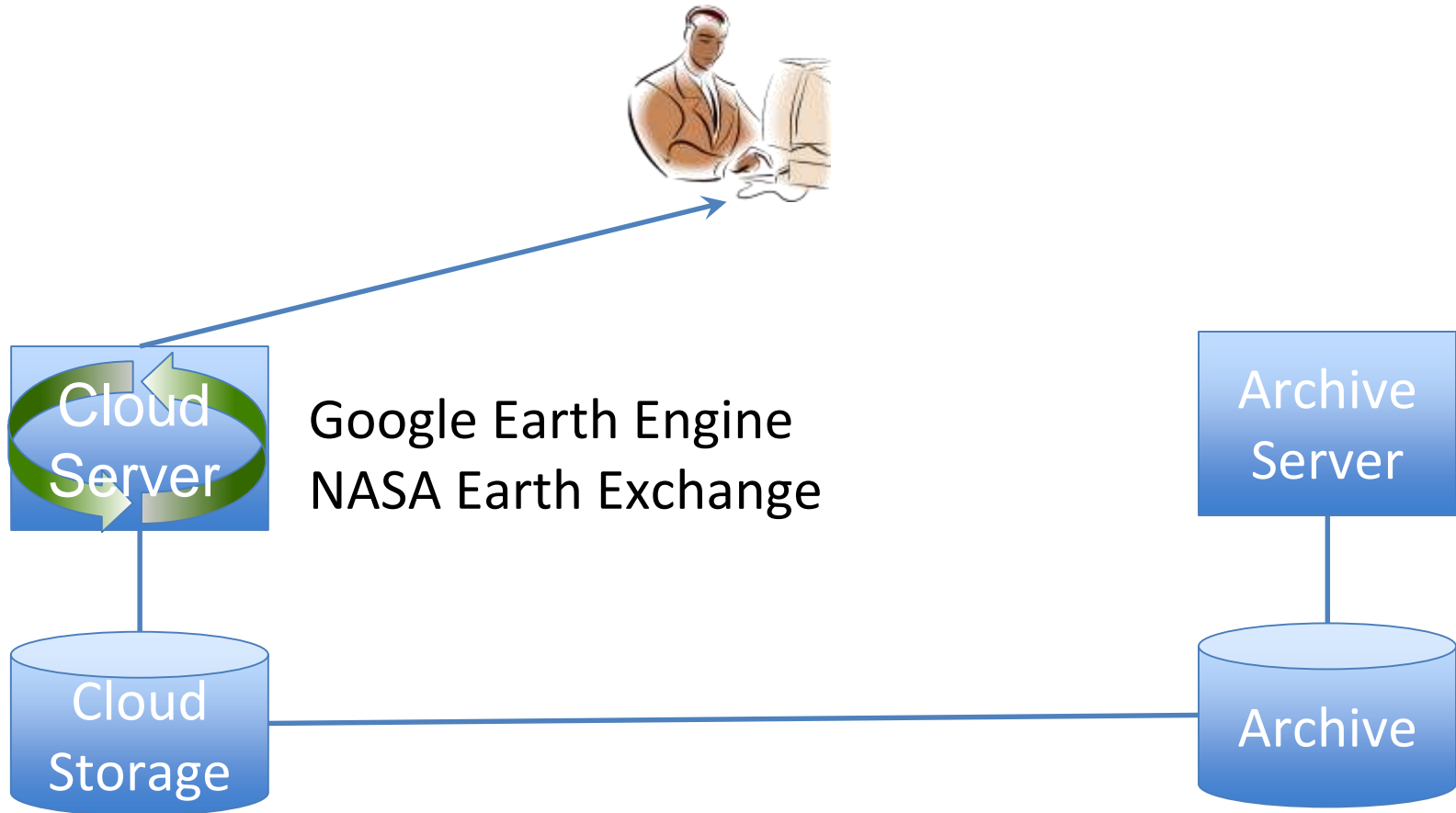




# “Close To” = Near Archive



# “Close To” = Near Processing





“more analysis closer to **the data**”

# “The Data”



Suitable Data Form	Processing Technologies
Original Data as Archived	GrADS Data Server ArclIMS for Server
Groomed Data (reformatted, annotated)	Earth System Grid Giovanni
Reorganized Data	Google Earth Engine SciDB MapReduce

# What's Next?

