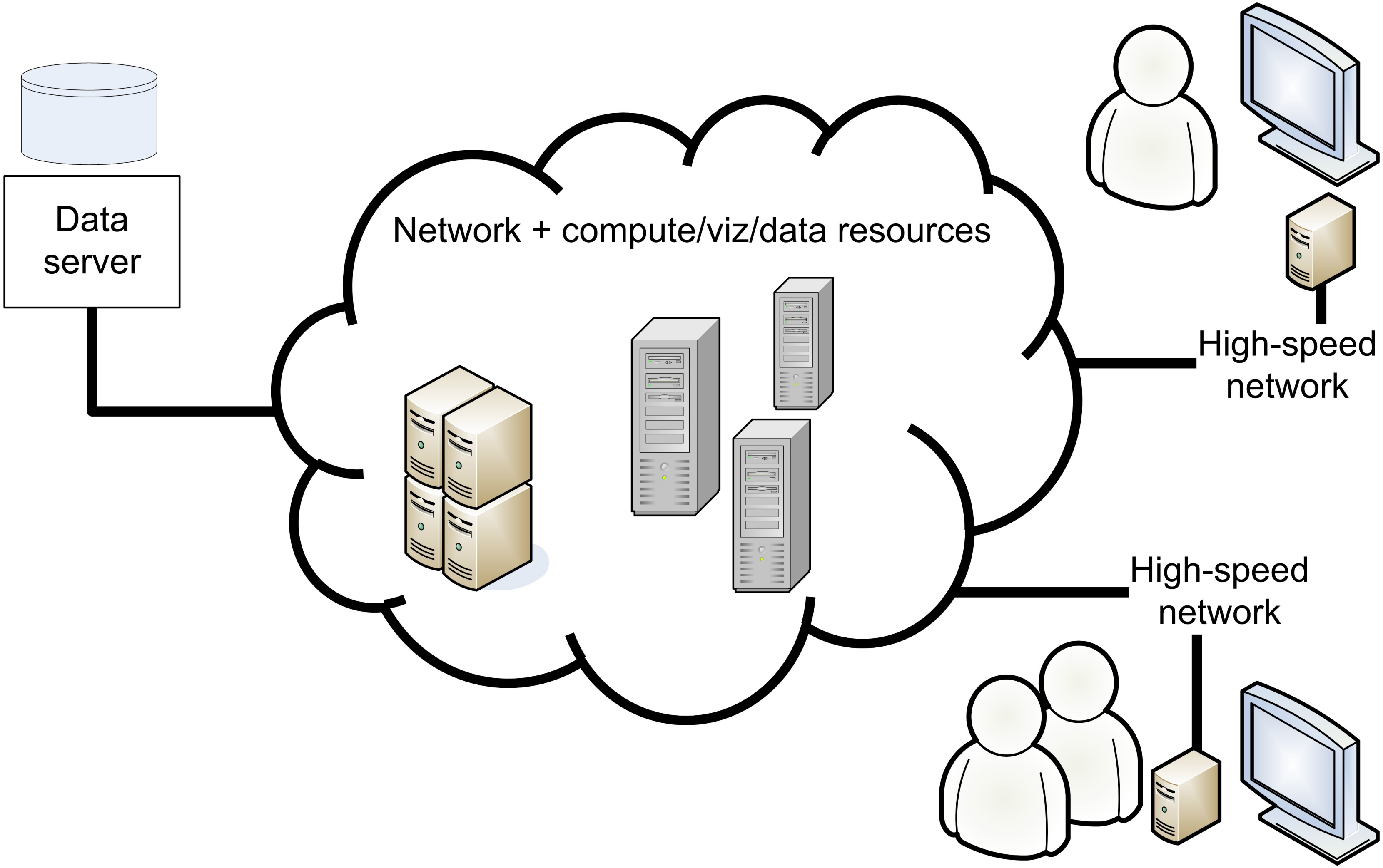
CSC 7700: Scientific Computing Module B: Networks and Data Lecture 5: Grid-based Visualization

Dr. Andrei Hutanu

Distributed Visualization

- Visualization pipeline
- Existing distributed visualization systems
- eaviv

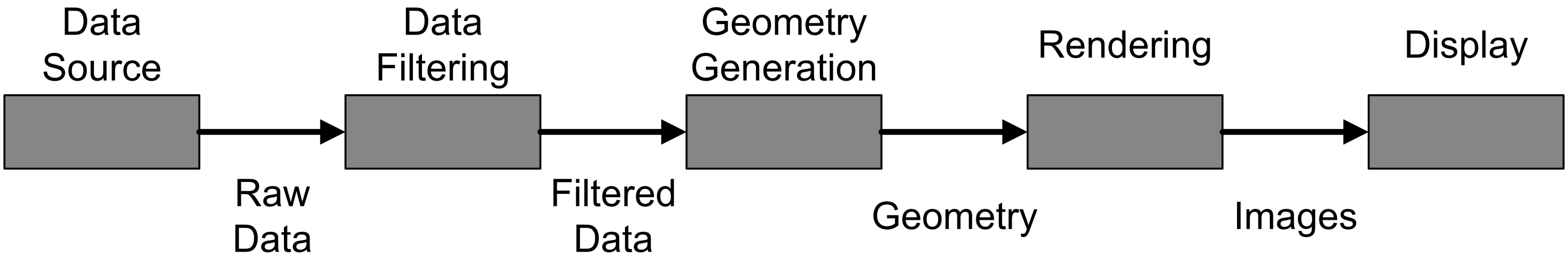
Motivating Scenario



Two approaches

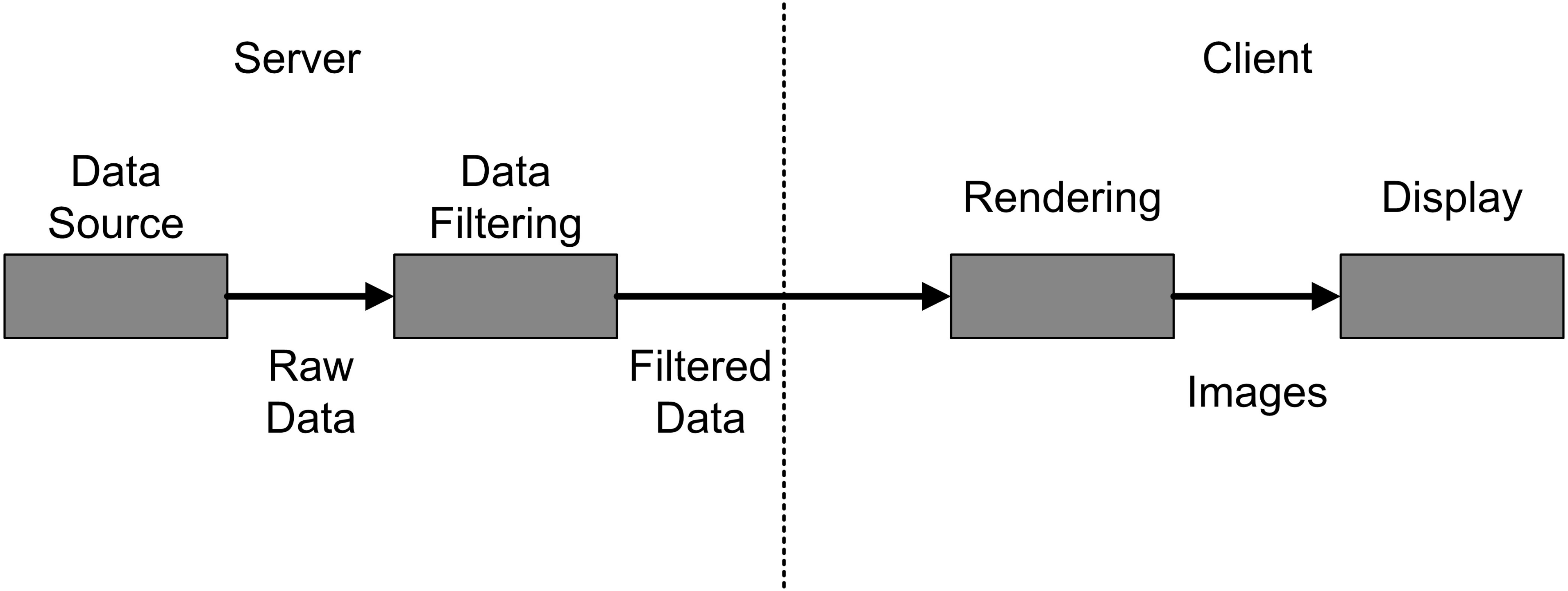
- Copy the data to visualize at the user location: staging
  - If network bandwidth & time permit
  - If powerful visualization resources are available locally
- Visualize the data from where it is located
  - Remote/distributed visualization
  - Useful when resource availability (network or time, local graphics) is limited or unable to deal with the data size(think Terabytes of data)
  - Focus of this lecture

Visualization Pipeline

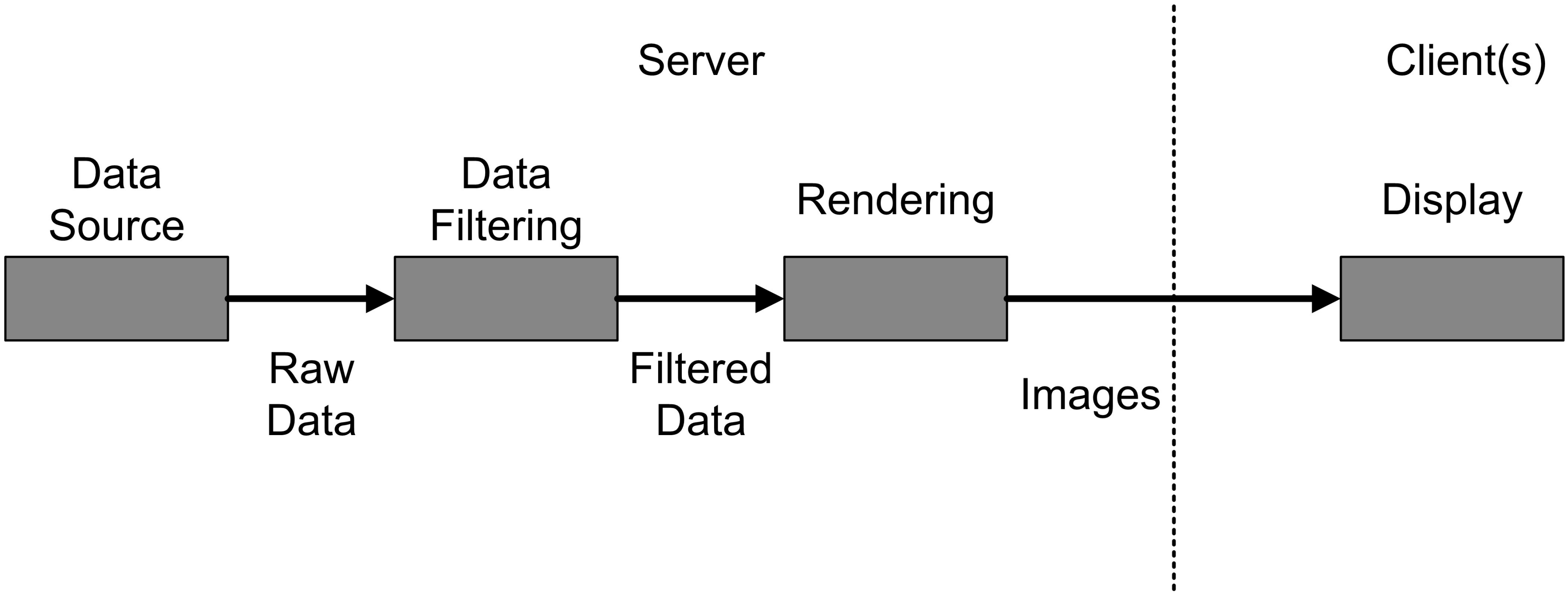


- Data source: disk or memory
  Data filter: selecting the data of interest
  Geometry generation: generate triangles/points
  Rendering: transform visualization primitives into images
  Images display: on one or more screens
- Interaction controls all components

Visualization of Remote Data (volume rendering)



Video Streaming



Server

Client(s)

Data Source

Data Filtering

Rendering

Display

Raw Data

Filtered Data

Images

VNC

- Tool for remote desktop viewing
- VNC server captures the screen and transmits it to the client (that has mouse and keyboard control over the remote desktop)
- Widely used for remote visualization
- Method of choice for remote visualization in TeraGrid
- Regular vncserver starts non-GPU accelerated X
    - For GPU-accelerated remote desktops use x0vncserver (X needs to be running and allowing connections from the user running the x0vncserver)

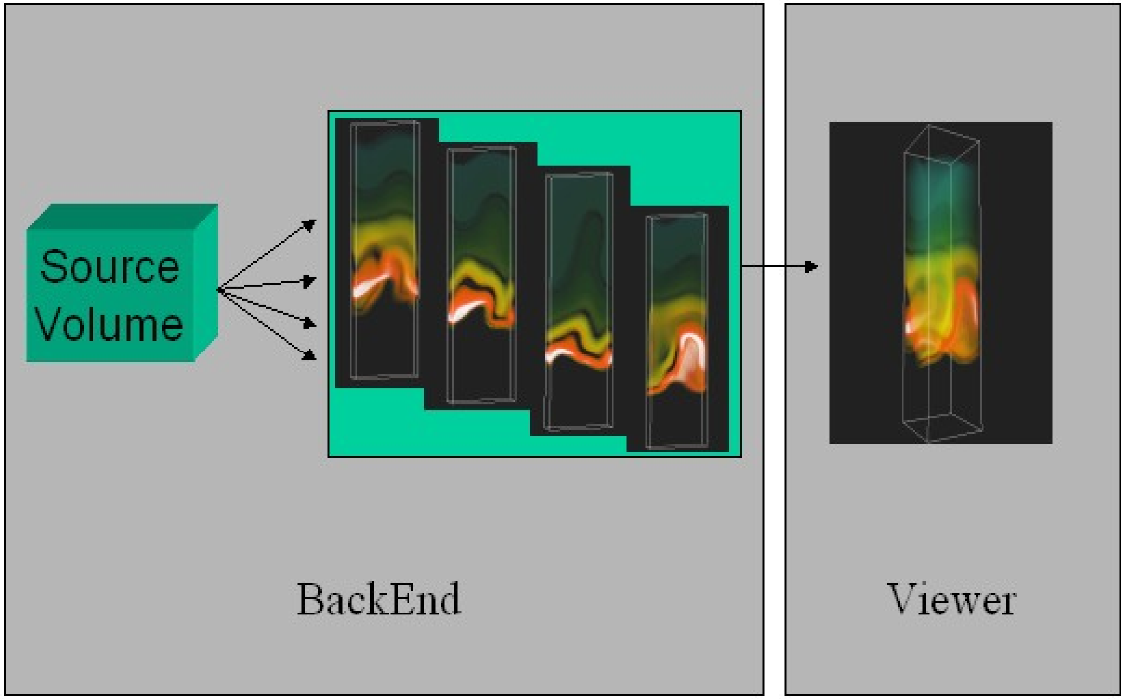
ParaView

- Client - server
  - Client connects to a parallel renderer running on a server
  - Rendering happening on the server if data is large, or on the client after transferring over the network
  - Decision made based on configurable parameters and video (for display) or data (for rendering) is transmitted over the network
- Client - rendering server - data server
  - Server separated into two: data server responsible for filtering, located near the data, and a rendering server that does only rendering (these are usually on the same LAN)

VisIt

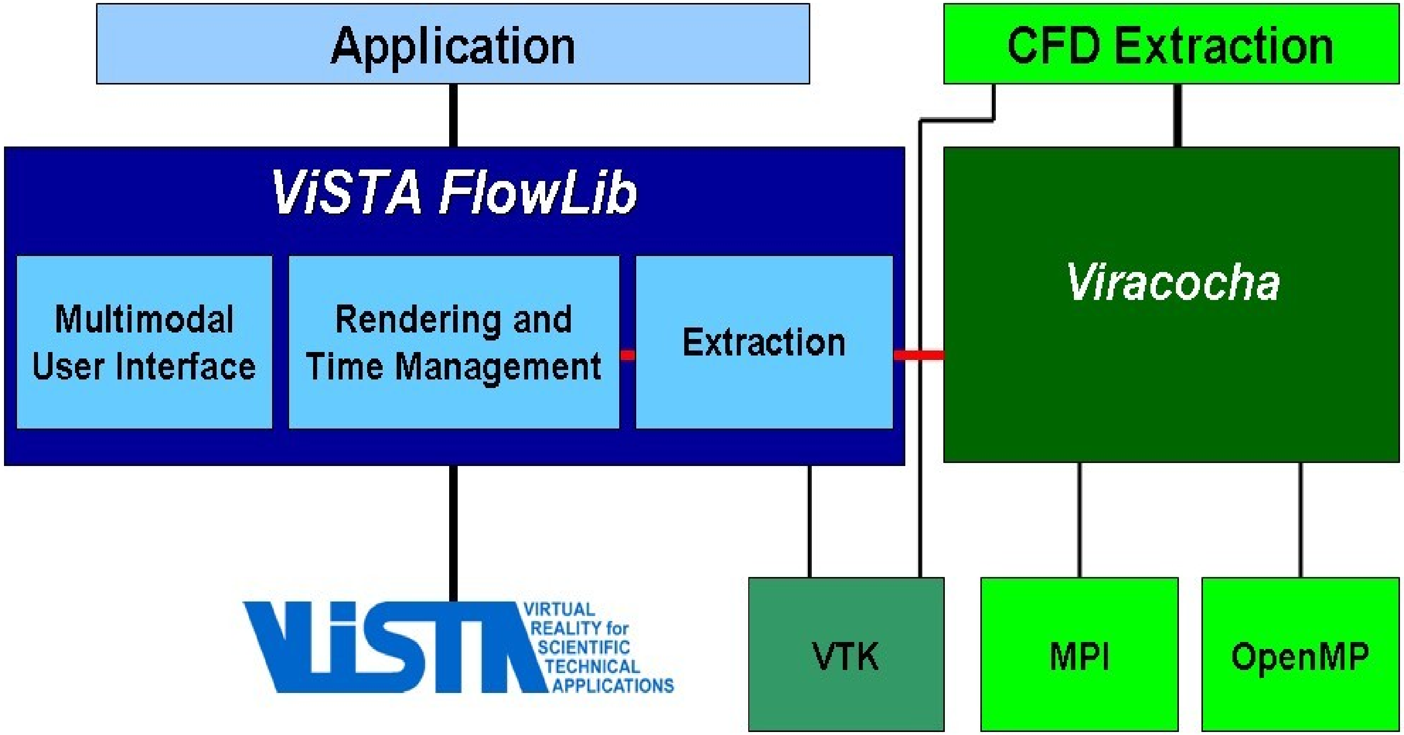
- Client - server distribution
    - Parallel rendering server
        - Local viewer & interaction client
        - Image streaming
    - Data processing and filtering server
        - Data transfer to client for interactive rendering
- Commonly, server reads the data from file
- Have the option to have simulation deliver the data to visit (separate server in two components, but on the same cluster) – visualization of live simulation data

Distributed Visualization Research



- Visapult
  - Data server -> Paralel back-end generates images
  - Images combined on viewer (image-based rendering)

Viracocha


Application
CFD Extraction
ViSTA FlowLib
Multimodal User Interface
Rendering and Time Management
Extraction
Viracocha
VISTA
VIRTUAL REALITY for SCIENTIFIC TECHNICAL APPLICATIONS
VTK
MPI
OpenMP

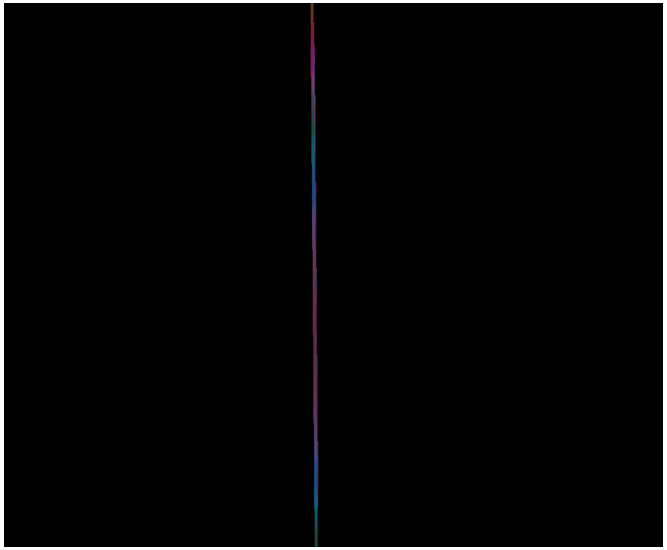
- Parallel data processing (filters)
- Parallel visualization
- Connecting over network

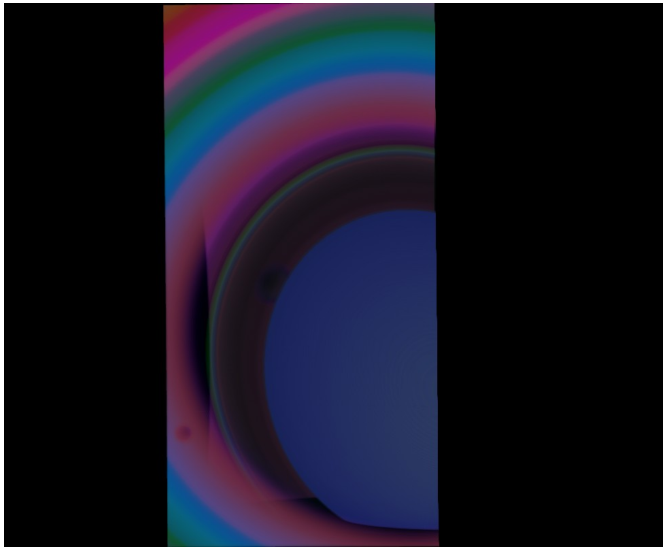
eaviv Project

- Interactively visualize large datasets generated by application scientists (simulations, image data)
  - High frame rate (>= 5 fps)
  - High data rate for I/O (>= 5 Gbps)
  - Large data (>= 10 Gigabytes of visualized data)
  - High resolution (>= 1 megapixel)
  - Good quality (no artifacts, good interaction response)
  - Enable collaborative visualization
  - Fast updates (<1 second)

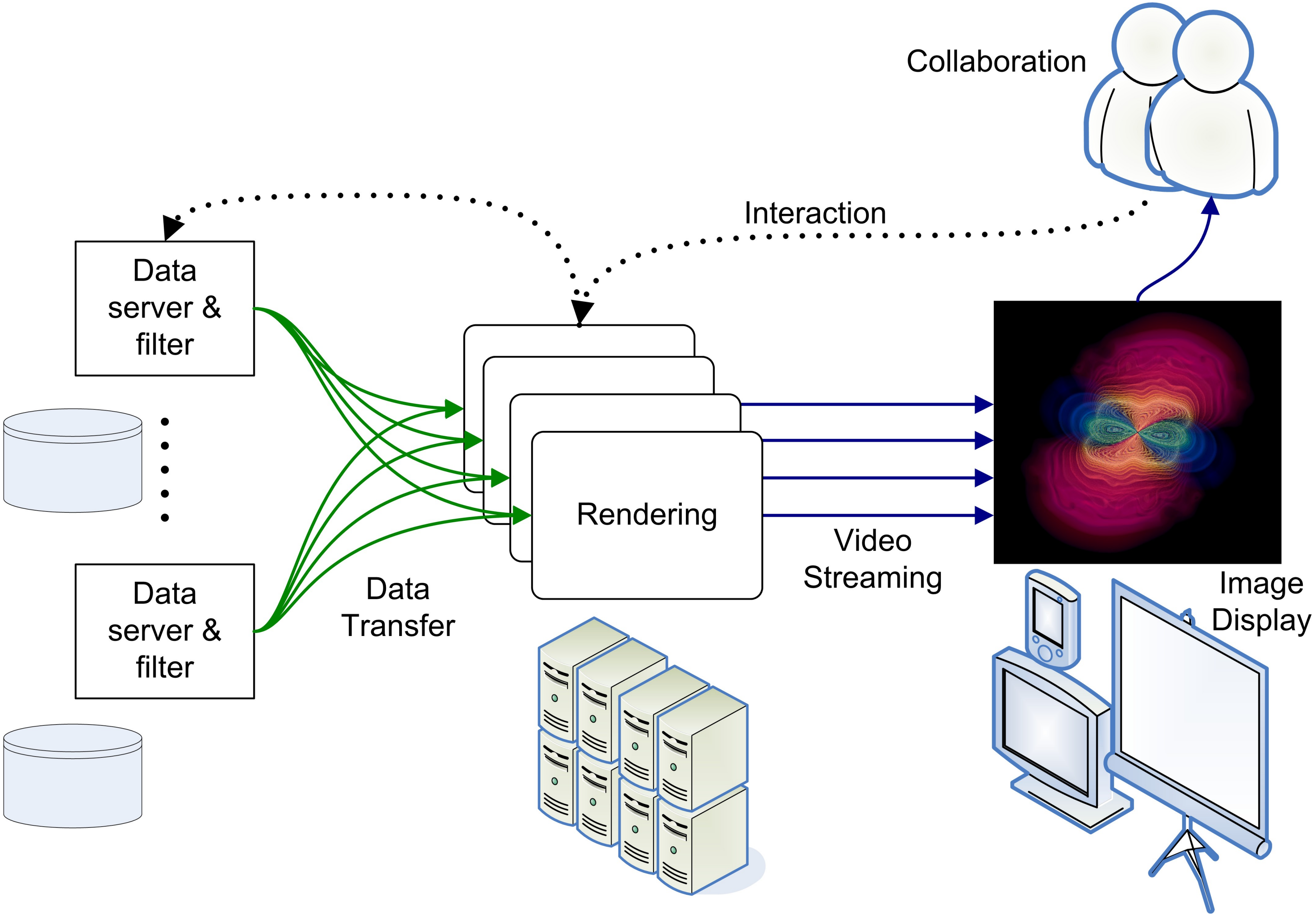
Benefits of distribution

- Increased I/O rate (see movies)
- Increase data size
(top image: laptop only visualization,
bottom image: distributed visualization on laptop using remote cluster)
- Collaborative visualization capabilities





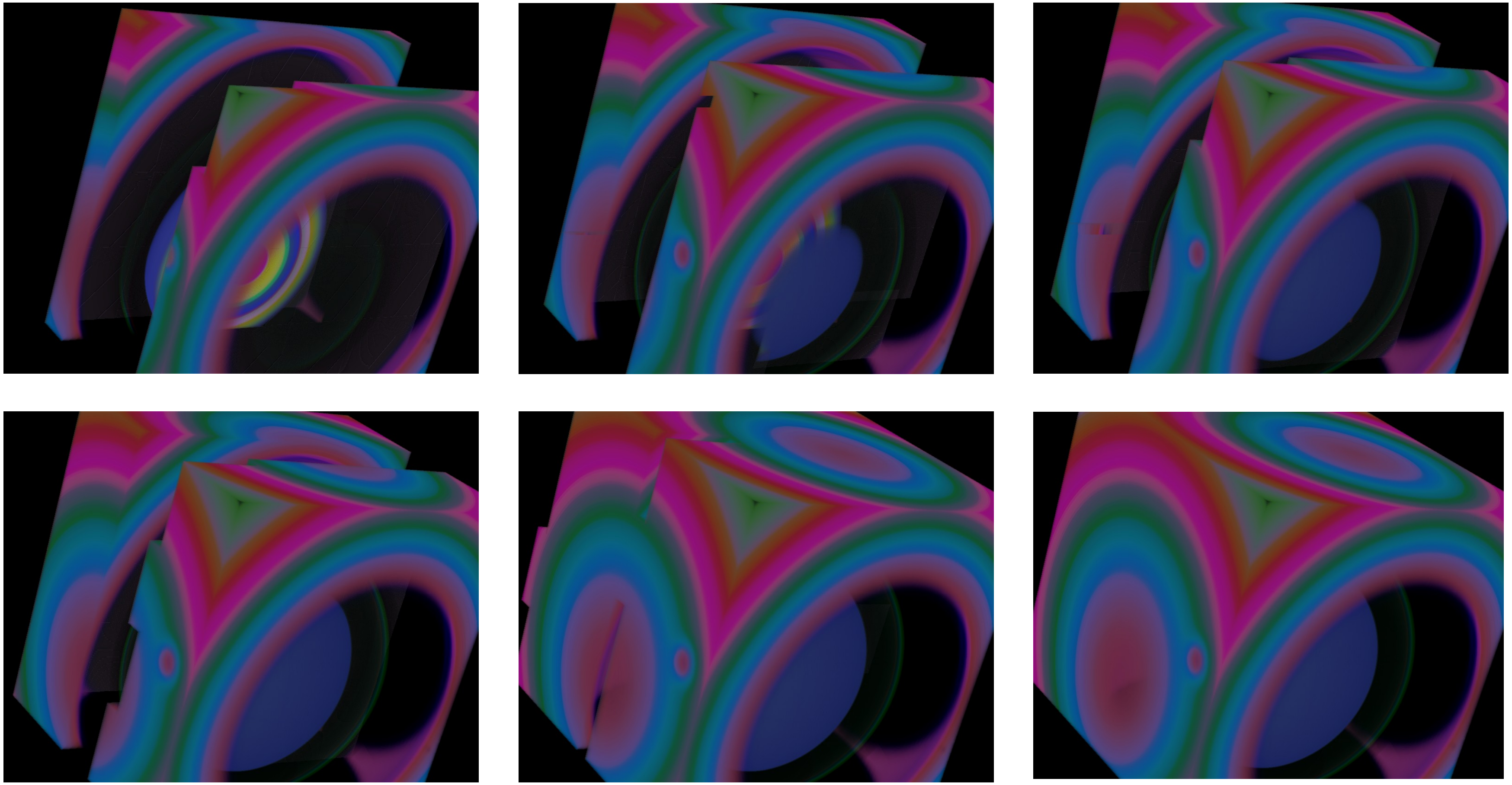
System Architecture



eaviv Components: I/O

- Networks can be faster than local storage
- Distributed data servers
- Use main memory to cache data
- High operation throughput, fast protocols

Components: Rendering



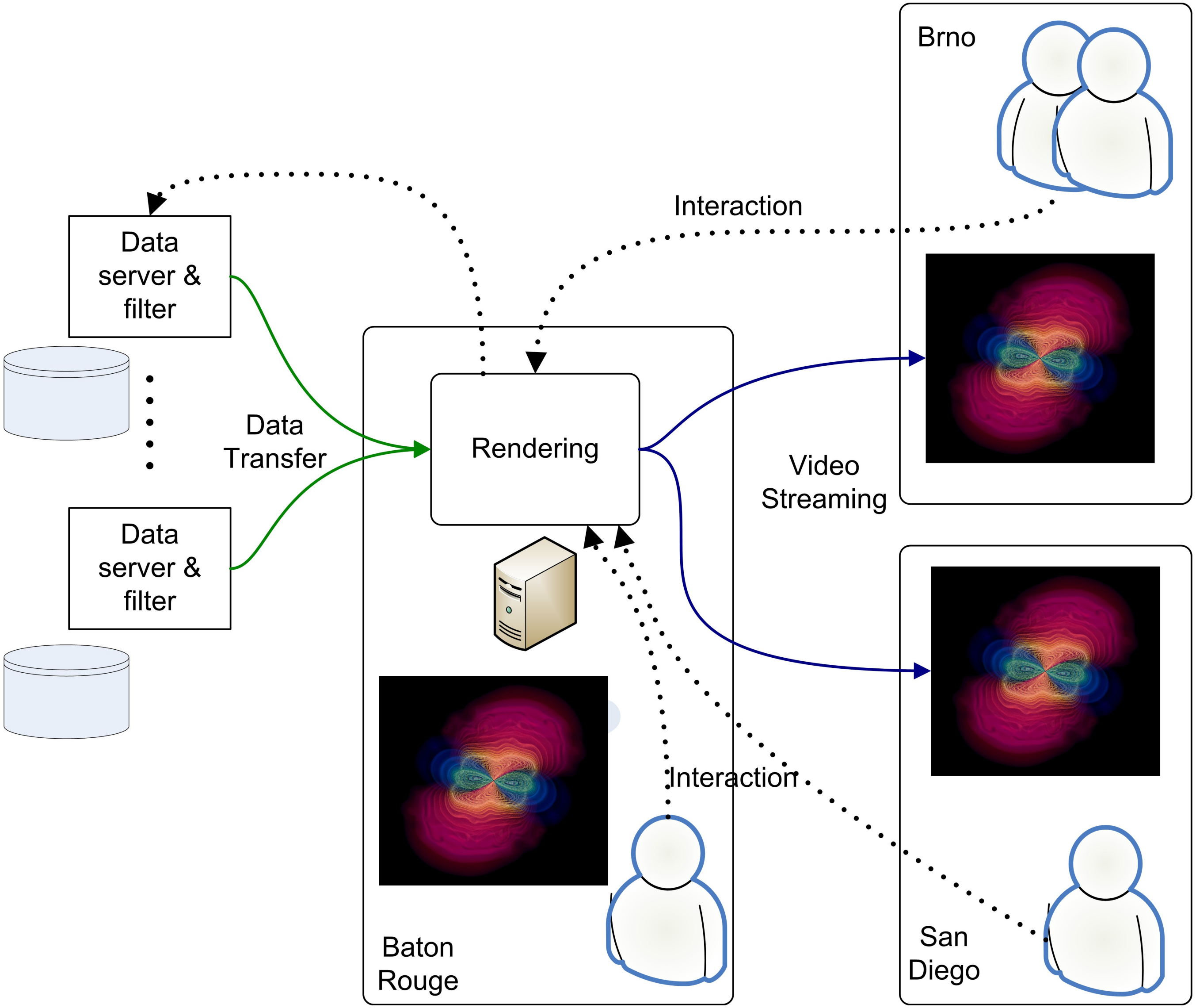
- Only data sections; Load data progressively. GPU

Components: Streaming, Interaction

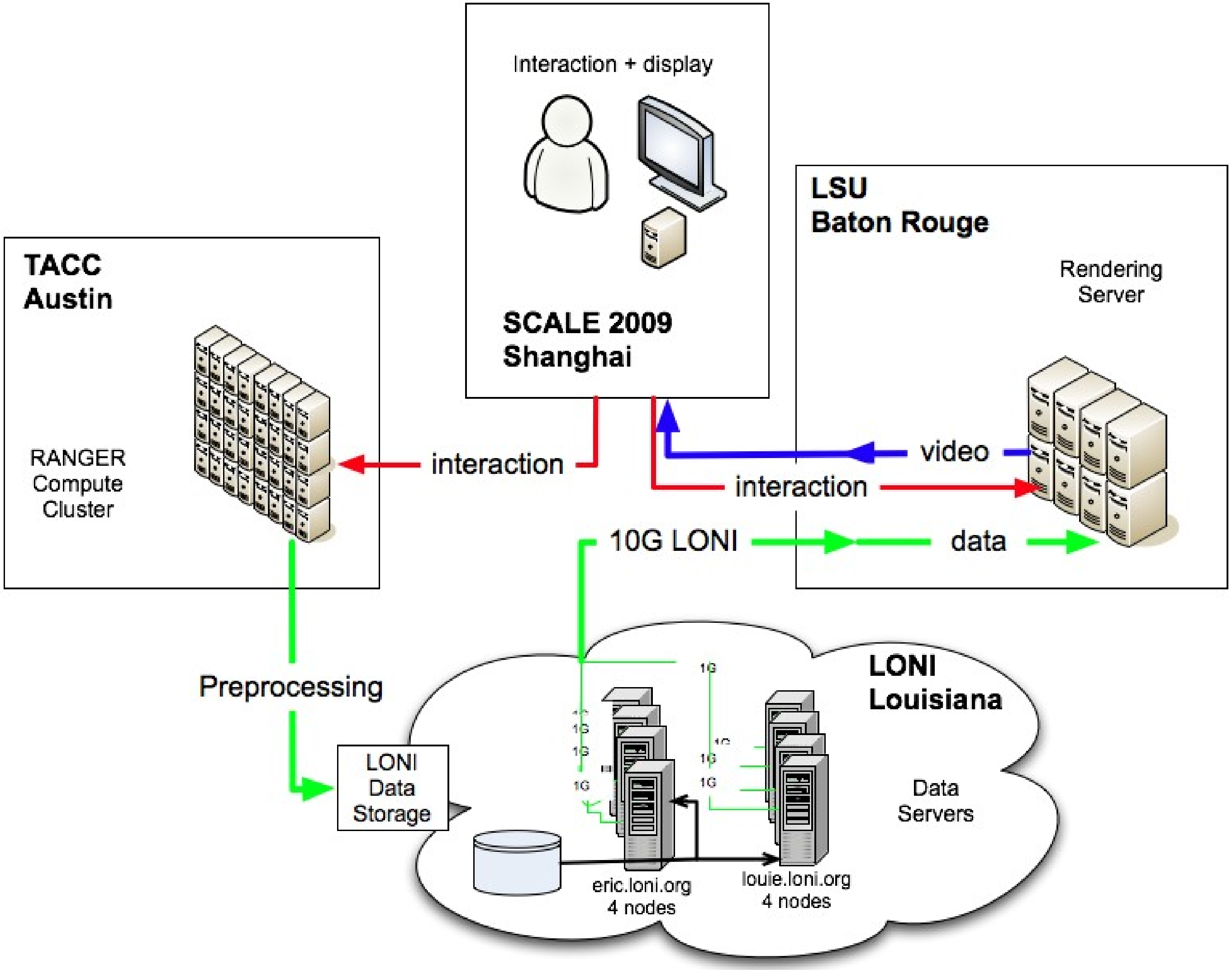
- Streaming
  - Images from remote renderer; collaboration
  - Using high-speed networks
    - High resolution
    - High frame rate
    - No compression (low latency)
- Interaction
  - Modify parameters (zoom, pos)
  - Mouse not really usable
  - Using “viz tangibles”; collaboration

iGrid 2005

- Remote data transfer 3 times faster than local
- Collaborative visualization (three-way)
- Uncompressed HD (hardware)
- Issues: data size, data transfer speed



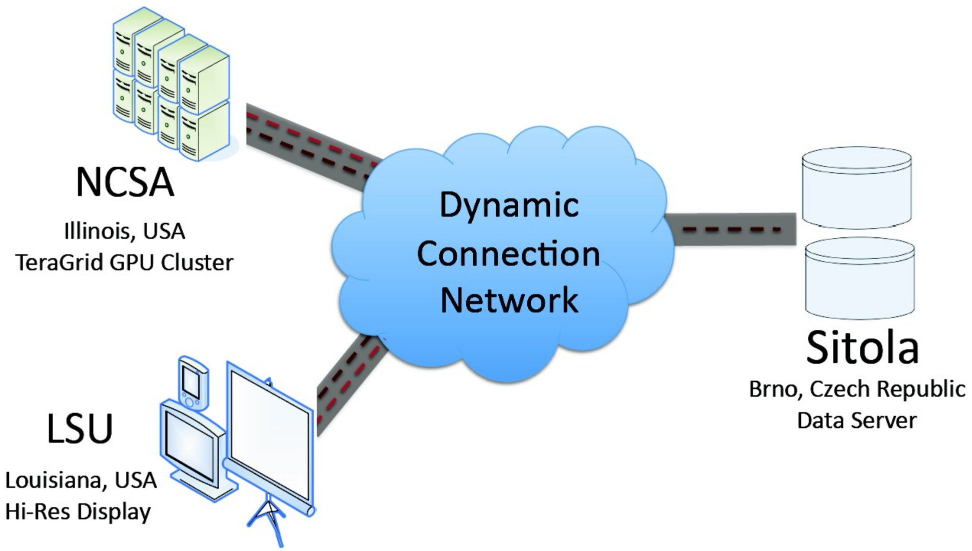
IEEE SCALE 2009



- 2048 core simulation code
- 8 node renderer at LSU
- Remote LONI nodes as data servers
- 20 timesteps cached remotely
- Scalable Approach!
- First Prize Winner
- http://www.youtube.com/watch?v=12VrgSAPIWw

Testbed for distributed app development

- Connecting NCSA, LSU and Sitola (Czech Republic) using dynamic circuit networks by Internet2
- PSNC (Poland) to be added soon
- GPU clusters
- Storage resources
- Large displays



Example utilization

Parallel rendering on Lincoln (NCSA)

Allocate ION network between NCSA and LSU

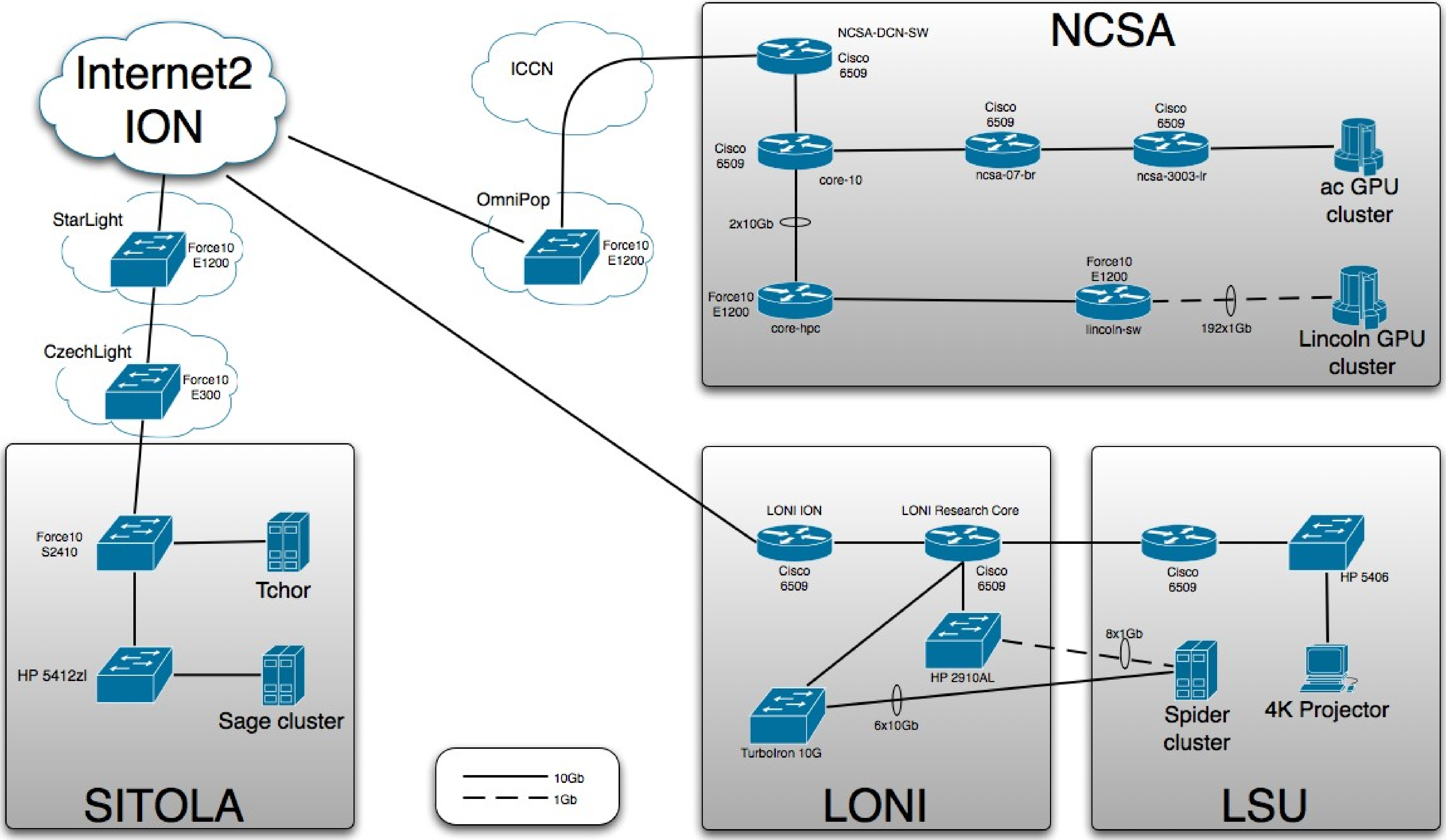
Video stream to LSU 4k projector (about 500 Mbps bandwidth utilization)

Interaction with visualization from LSU

Can have multiple viewer + interaction points (various locations): Collaboration

http://www.youtube.com/watch?v=jg5HQ1wc9_A

Network testbed



Full circle

- Integrated application development
- Take the lectures in reverse
    - Going from top issue (visualization of remote data) into lower-level components
    - Middleware (to actually distribute the application)
    - Remote I/O and data management (to deal with the remote data)
    - Video streaming
    - Data transport protocols

Final remarks

- Suggest to read the research papers listed on the wiki
  - https://wiki.cct.lsu.edu/sci-comp/Networks_and_Data#Research_Papers
- 10% of final grade from the report
  - Unless you decide to do the optional second assignment for 5% (then report will count for 5%)
- 10% in the final exam
  - Five questions (can have multiple parts) covering one each lecture
  - You will be able to select four out of five (equal grade for each, no bonus for answering all five)