LUSciD Cosmology: The Cosmic Simulator

Michael Norman, UCSD

Robert Harkness, SDSC

Frank Graziani, LLNL

Overarching Goals

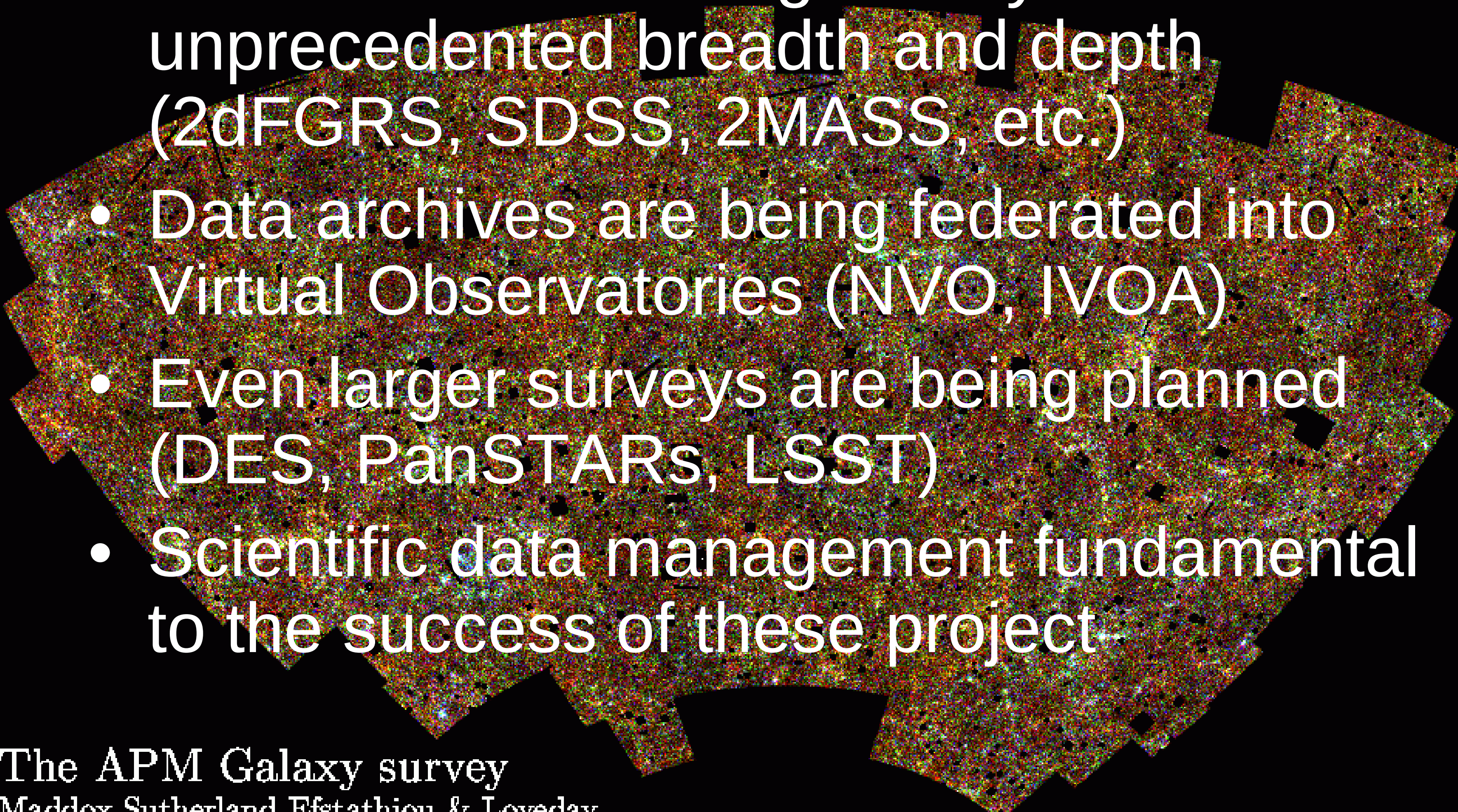
- To advance the state of the art of multiphysics cosmological simulations in both scale and complexity
- To create a public archive of numerical simulation data of high scientific value relevant to current cosmological research
- To foster collaborations with LLNL

Outline

- Background: cosmology and “big data”
- Science application: LightCone simulations
- Year 1 deliverables and status
- Technical details
- Year 2 plans

Cosmology and “big data”

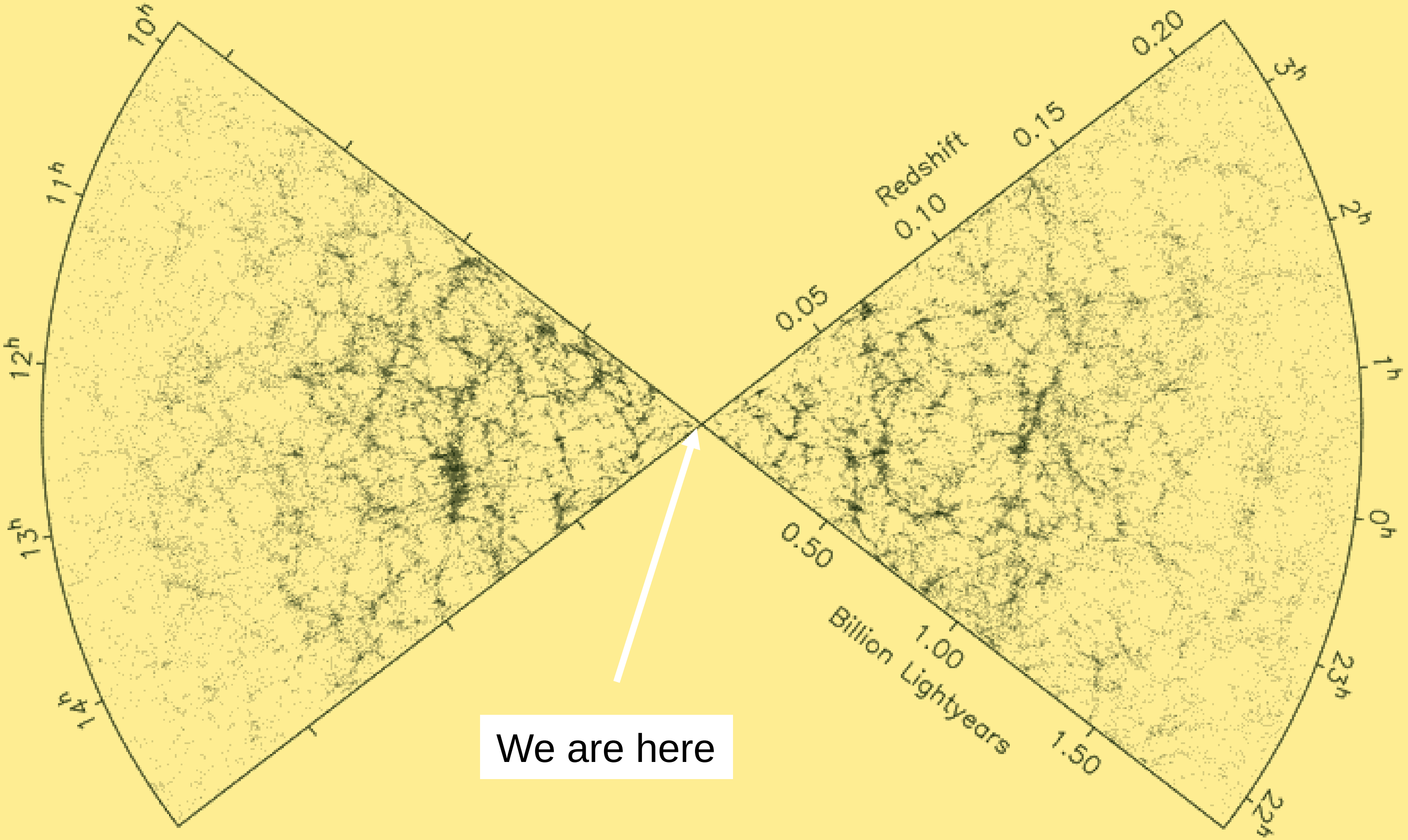
- The universe is being surveyed to unprecedented breadth and depth (2dFGRS, SDSS, 2MASS, etc.)
- Data archives are being federated into Virtual Observatories (NVO, IVOA)
- Even larger surveys are being planned (DES, PanSTARs, LSST)
- Scientific data management fundamental to the success of these project



The APM Galaxy survey
Maddox Sutherland Efstathiou & Loveday

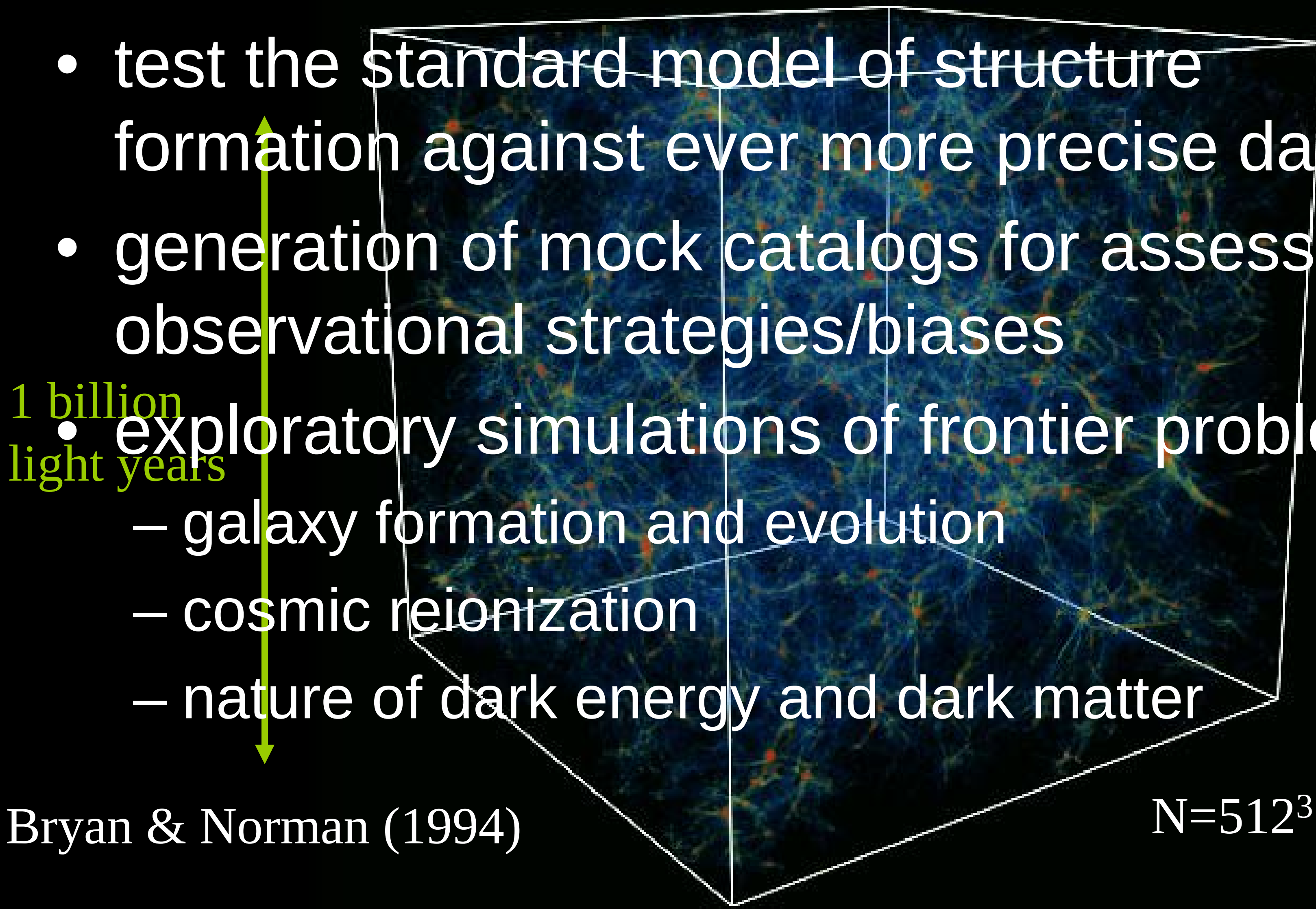
The 2dF Galaxy Redshift Survey

Final Data Release - 30 June 2003



Role of Cosmological Simulations

- test the standard model of structure formation against ever more precise data
- generation of mock catalogs for assessing observational strategies/biases
- exploratory simulations of frontier problems
  - galaxy formation and evolution
  - cosmic reionization
  - nature of dark energy and dark matter



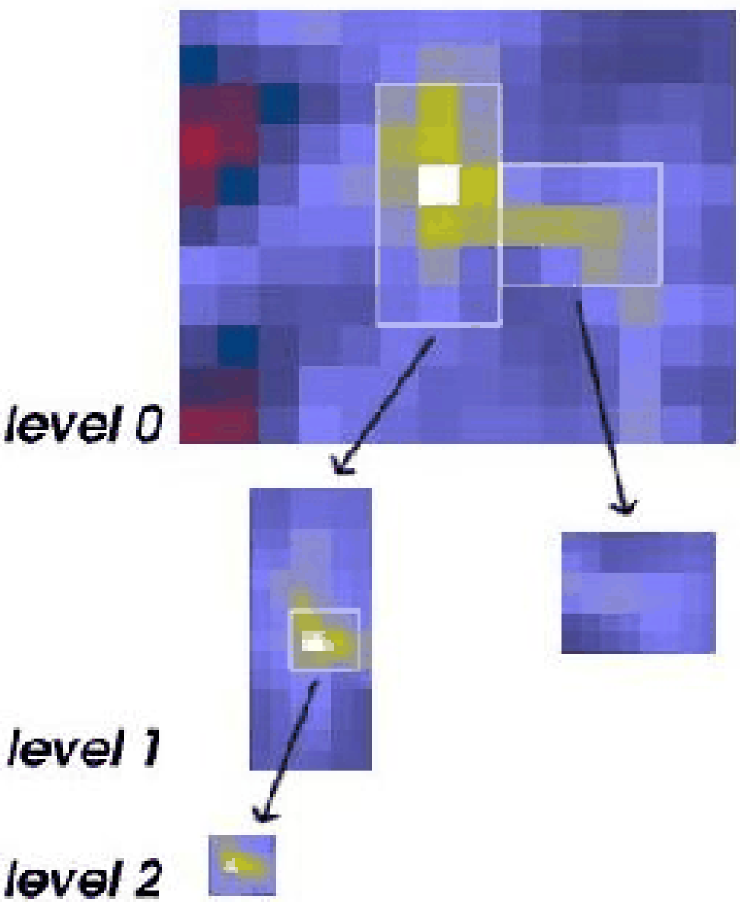
Bryan & Norman (1994)

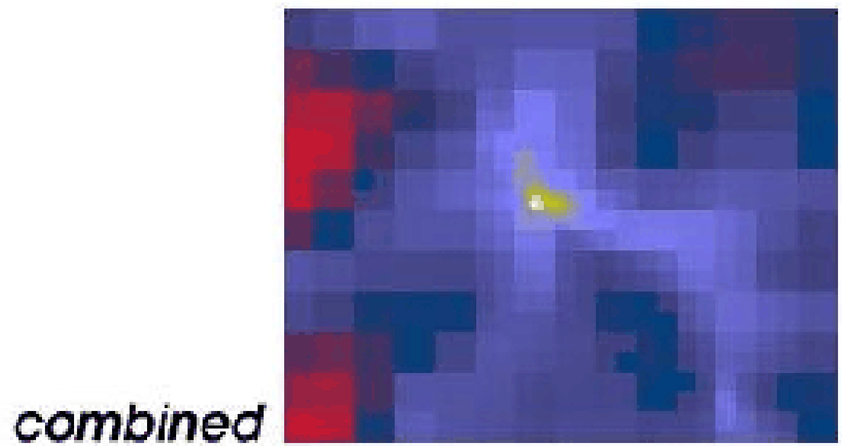
Ultrascale ENZO Simulations

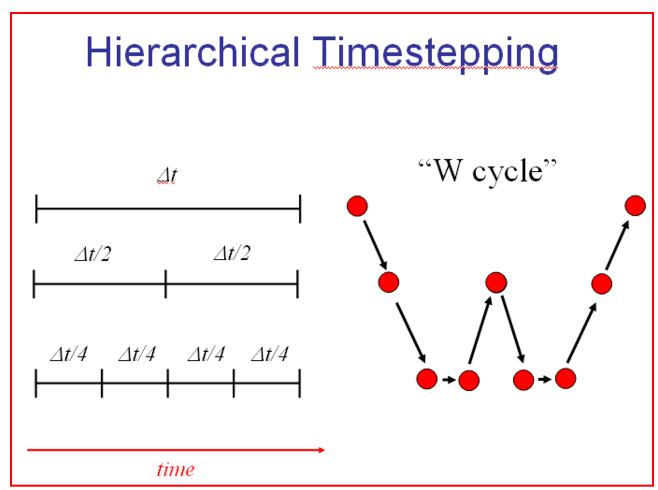
- To remain relevant to cosmological research, cosmological simulations must increase in size and complexity
- Size: survey volumes to large redshift depths exceeding 1 Gpc^3
- Complexity: self-consistent treatment of dark matter, self-gravity, cosmic expansion, gas dynamics, ionization, energy balance, star formation, radiative transfer
- Managing and publishing the TB of output from such simulations is the principal goal of this project
- Incorporating radiative transfer into ENZO is a secondary goal

ENZO: 3D Hybrid Cosmological Adaptive Mesh Refinement Code

(Bryan & Norman 1997, 1999)

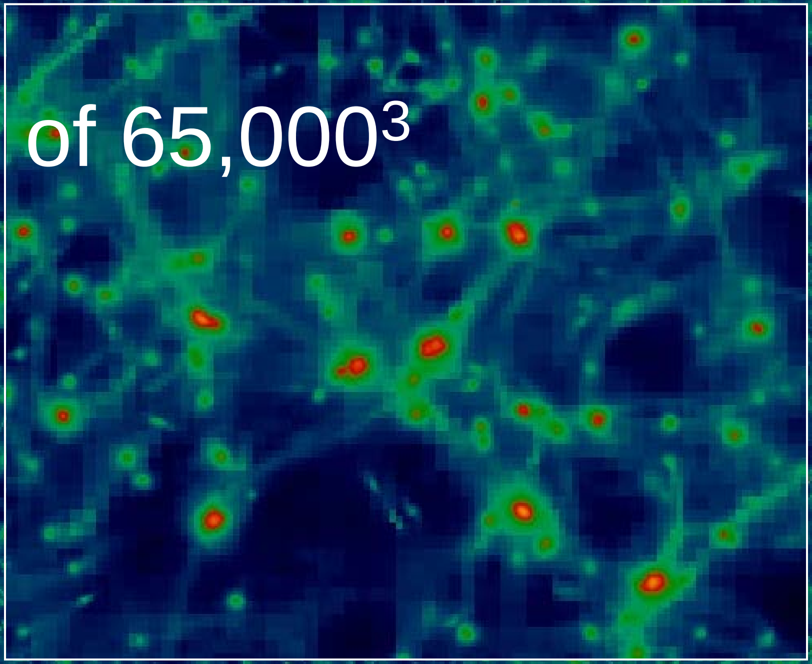






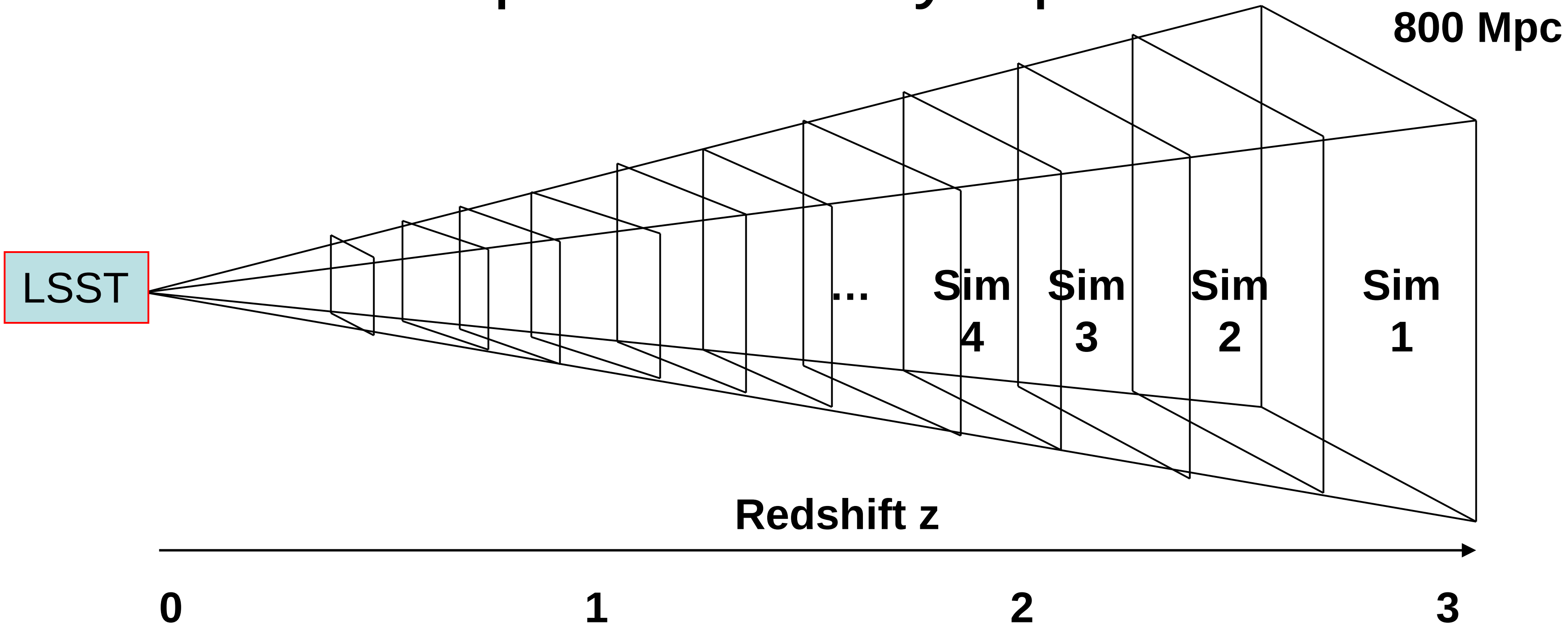
State of the Art

- Galaxy clusters in $(0.7\ \text{Gpc})^3$
- Hybrid simulation: dark matter and gas
- >350,000 subgrids at 7 levels of refinement
- Effective resolution of $65{,}000^3$
- 6 TB output
- done at SDSC



Science Goal: LightCone simulations of galaxy clusters

- 100 square degree survey area
- Constant angular resolution (gigapixel)
- redshift depth: 3➔many Gpc^3



Year 1 deliverables

- Port ENZO to LLNL compute resources and perform scaling studies
- Develop ENZO metadata schema and enhanced file I/O
- Develop Cosmic Simulator (ENZO) analysis pipeline and archive
- Produce first simulated LSST sky map

Yr 1 deliverables and status

- Port ENZO to LLNL compute resources and scaling studies
- Status: DONE
  - Massive AMR simulations are a brave new world
  - Many high performance mods to ENZO
  - 3 of a planned 16 very large AMR simulations of the “lightcone” have been completed on Thunder
  - 4 TB of data transferred from LLNL to SDSC
  - Production runs generate performance data which are being analyzed

Yr 1 deliverables and status

- Develop ENZO metadata schema and enhanced file I/O
- Status: DONE
  - Draft metadata schema posted on project website open for comment
  - Packed AMR file I/O based on HDF5 groups in production
  - *NEW!* Memory-resident I/O yields 100x speedup and reduced exposure to file system instability

Yr 1 deliverables and status

- Develop Cosmic Simulator analysis pipeline and archive
- Status: Excellent Progress
  - Massively parallel versions of ENZO analysis tools developed and are undergoing testing
  - We are running our own SRB server instance to manage ENZO data and metadata
  - We are on track to publish a previously computed 1.5 TB data collection by 9/1/06

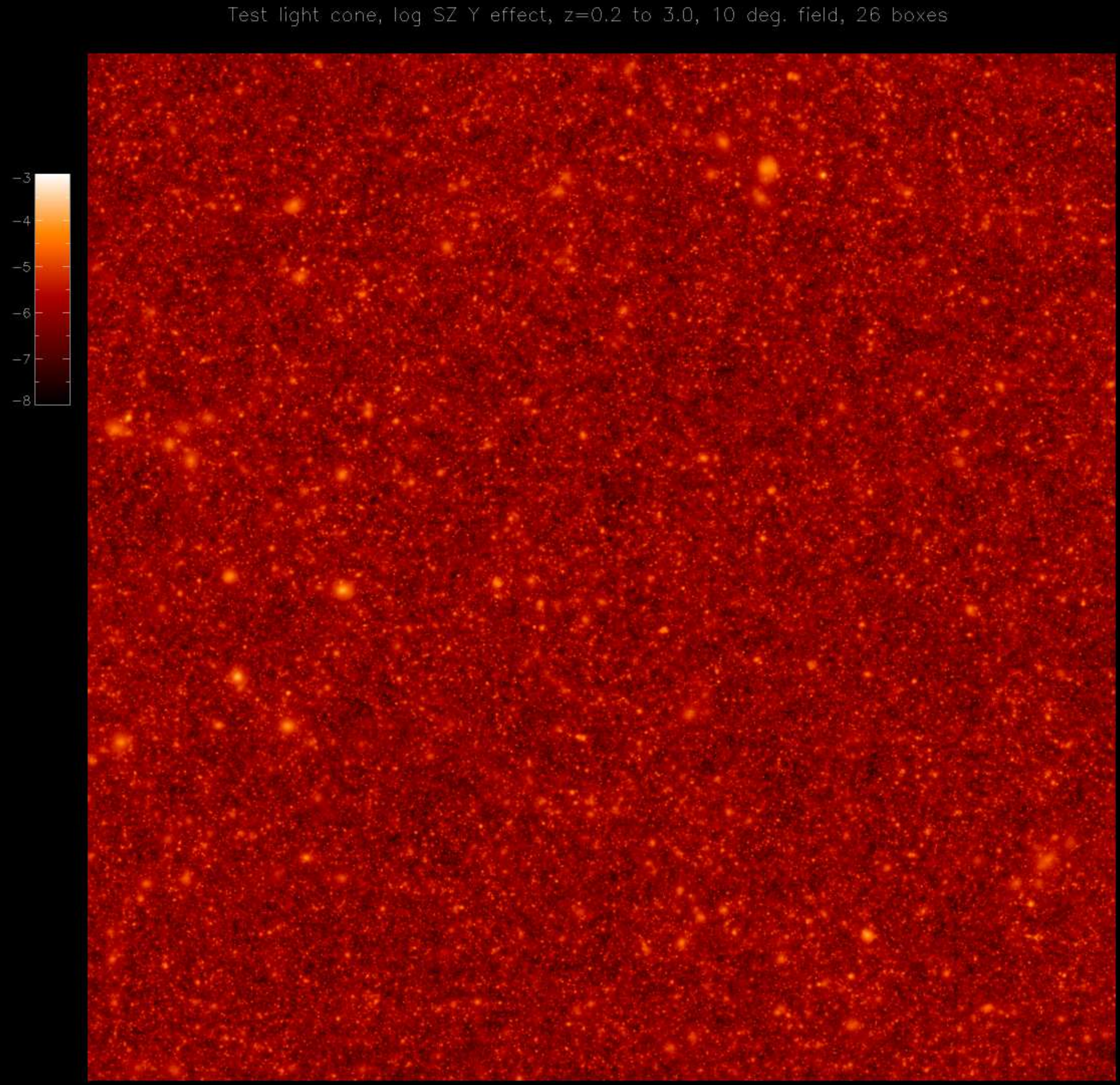
Yr 1 deliverables and status

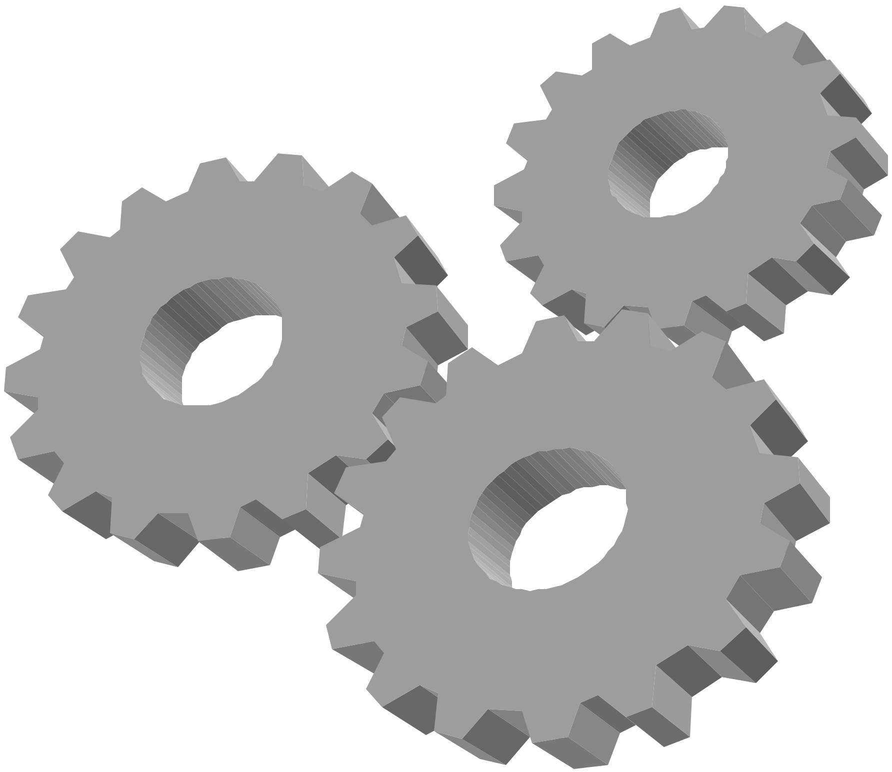
- First simulated LSST sky map
- Status: Good Progress
  - Have produced a "mock" lightcone projection at 8192^2=65 megapixel resolution
  - Have the ability to produce a lightcone projection image at $65{,}000^2$ = 4 gigapixel resolution (LSST will have 3 gigapixel)
  - How to store? display?
  - Time to engage LLNL LSST team

Projected image
8192^2

100 sq. degrees

SZ effect





Technical Details

Organization

Team	**Description**	**Team Leader**	**Members**
Production runs	Running production jobs on Thunder	Robert Harkness	James Bordner
Data management	Design and implementation of Cosmic Simulator data archive @ SDSC	Rick Wagner	Robert Harkness Jake Streeter
Code development	Enhancements and support of Enzo software and tools	Dan Reynolds	James Bordner John Hayes Robert Harkness Alexei Kritsuk Pascal Paschos Rick Wagner
Visibility	Project webpage	Jake Streeter	James Bordner Rick Wagner

Main ENZO Code Developments

Major modifications in Year 1 (Harkness)

- Fast sibling grid locator
- Parallel IC generator, OOC* IC generator
- OOC boundary conditions on top mesh
- Nested rectangular ICs
- Task-to-processor/node mapping
- Memory-resident I/O

*out-of-core

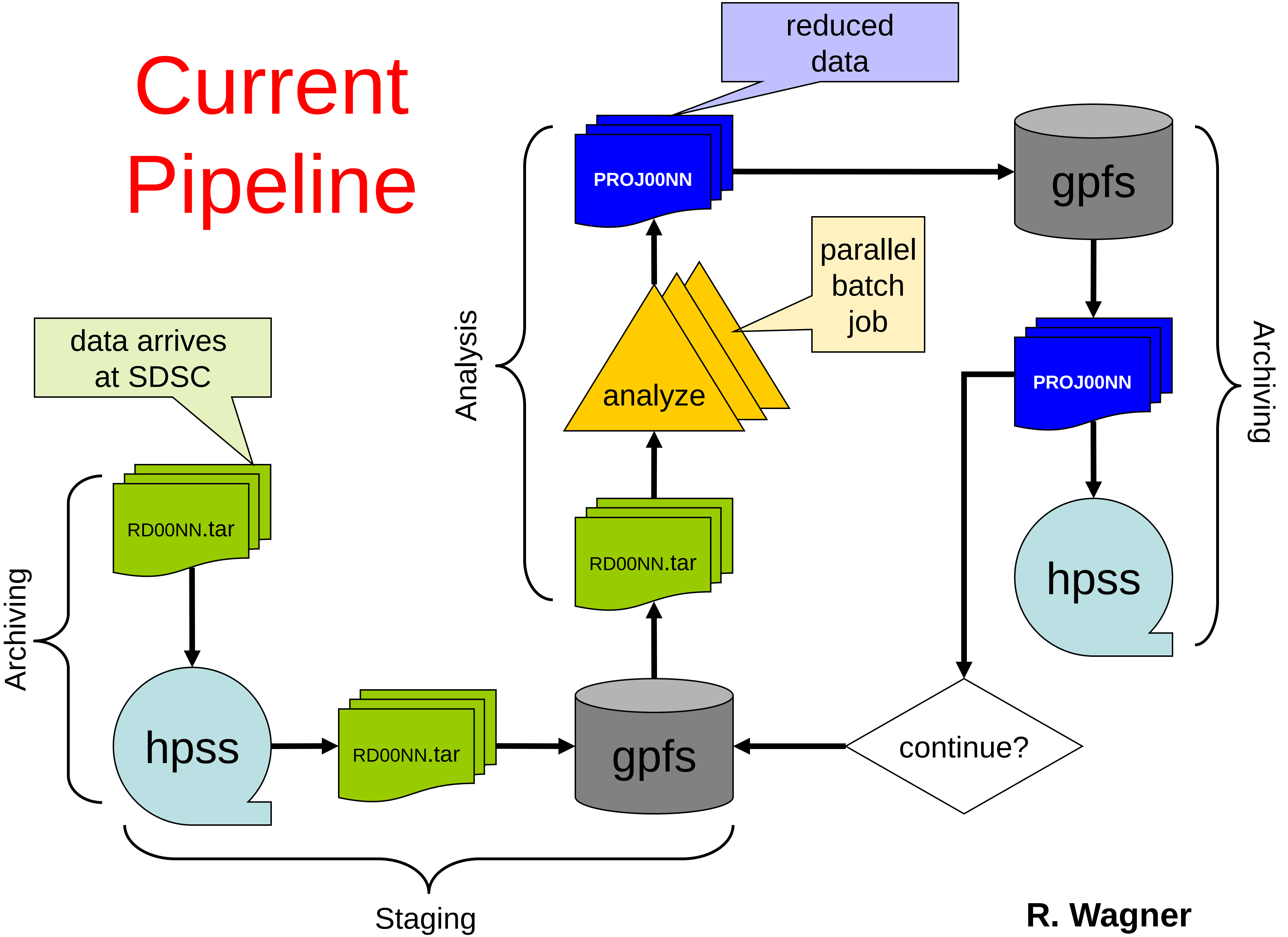
In-core HDF5 File Assembly

- For extreme AMR small I/O chunk size and frequent file extensions can become prohibitive on some file systems and also expose reliability issues (e.g. LLNL Thunder).
  - HDF5 memory driver used to assemble entire HDF5 files in core with single write on closure
  - Performance gains up to 100x, or more
  - Eliminates I/O as a limiting factor in deep AMR

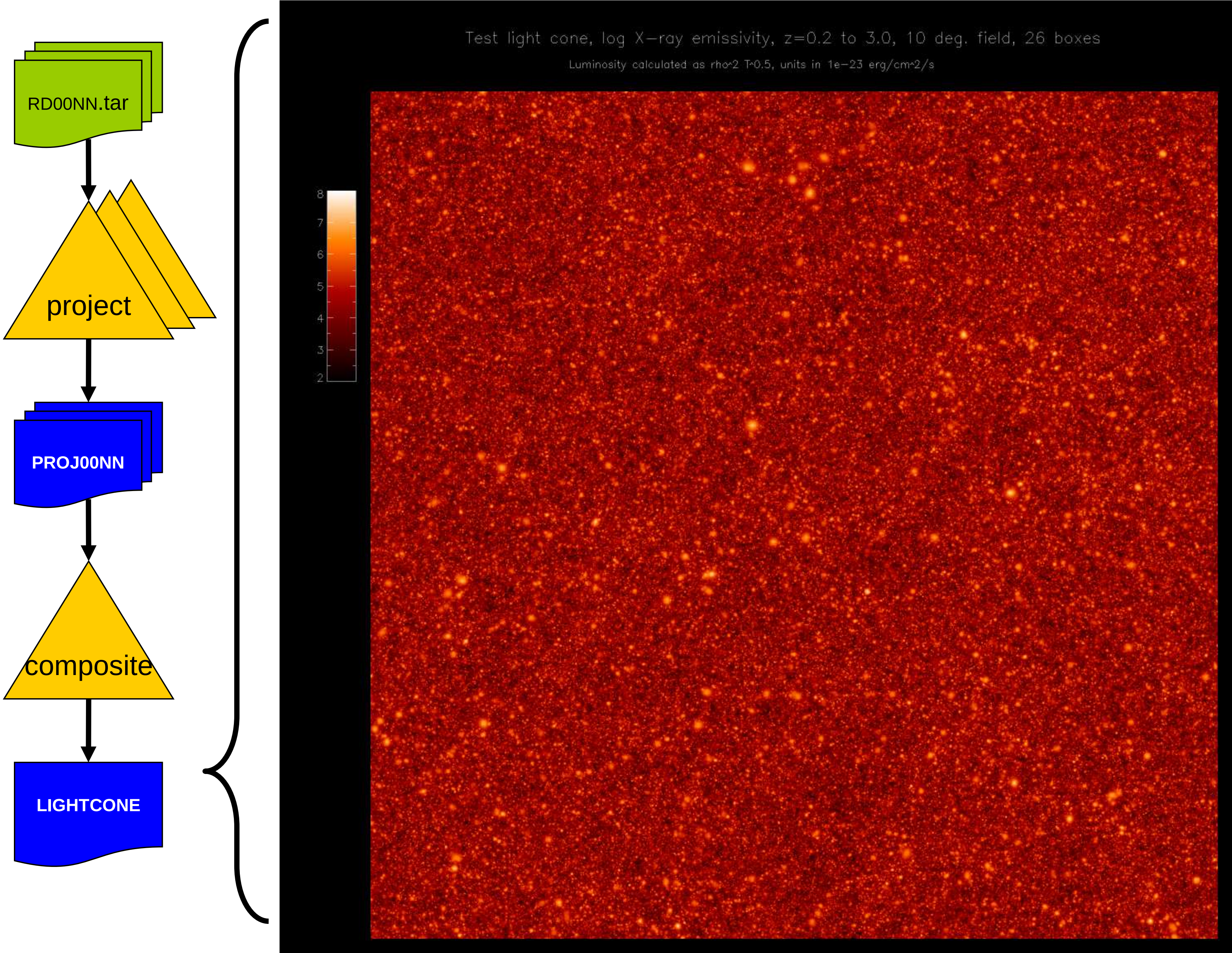
LightCone Runs to Date

- Tile 1 completed at LLNL
- Tiles 6 & 7 started at LLNL, completed at SDSC
- Tile 16 in progress at SDSC
- With HDF5 memory-driver mods Thunder should be able to accommodate the whole sequence
- Data transfer to SDSC is fastest (and most reliable) using BBFTP
- Data transfer to SDSC is most convenient with HPSS hsi or SRB

Current Pipeline



R. Wagner



RD00NN.tar
project
PROJ00NN
composite
LIGHTCONE
Test light cone, log X−ray emissivity, z=0.2 to 3.0, 10 deg. field, 26 boxes
Luminosity calculated as rho^2 T^0.5, units in 1e−23 erg/cm^2/s
8
7
6
5
4
3
2

Archive Design Principles

- Target Audiences:
  - Internal users via shared archive
  - Astronomers via VOs
- Internal Users:
  - Primarily developers
  - Need to manage data, not a workflow GUI
- Astronomers:
  - Need standardized data products
  - Shouldn’t have to deal with our proprietary format

R. Wagner

Schema Design

At a minimum, the schema must be able to answer two questions:

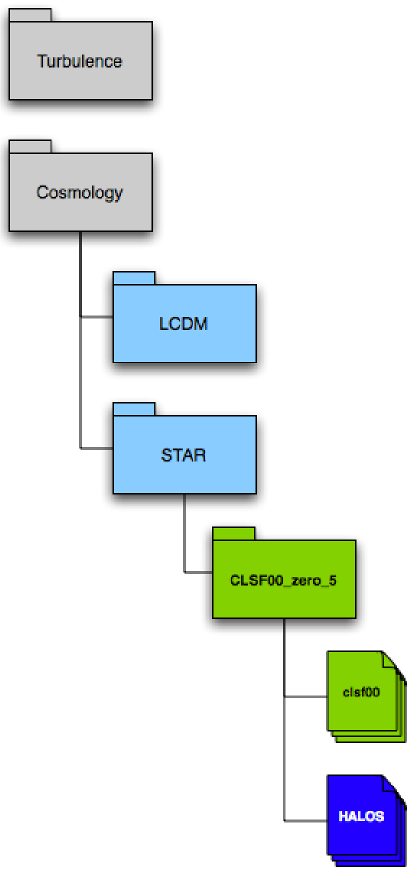
1) How was the object created?
2) What science does it address?

- Considerations:
  - Simple designs promote use
  - Can always be extended

R. Wagner

Archive Components

SRB hierarchy



SRB metadata

```
cable:~ rpwagner$ Sls
/Cosmology/STAR/CLSF00_zero_5:
  clsf00_0000.tar
  clsf00_0004.tar
  clsf00_0008.tar
```

```
UDMS0="RunID"
UDMS1="349"

UDMS0="Program"
UDMS1="Enzo"

UDMS0="ProblemType"
UDMS1="30"

UDMS0="TopGridDims"
UDMS1="128 128 128"

UDMS0="ComovingBoxSize"
UDMS1="256"

UDMS0="CurrentRedshift"
UDMS1="0.08"
```

database tables

t_machines

id	name	site
1	Cobalt	NCSA
2	DataStar	SDSC

t_runs

id	mach	user	procs	nodes
349	1	chen	32	16
350	4	ux453784	64	8

t_users

mach	user	person
2	u4672	5
1	chen	5

t_persons

id	last_name	first_name	email
5	Chen	Johnnie	jdoe@foo.edu
9	Streeter	Jake	cool@goo.com

R. Wagner

SRB Hierarchy "Rules"

Topic	Top-level collection, broad scope ("Cosmology", "Software", Turbulence") May contain: Research Areas
Research Area	Intermediate collection, narrowing scope ("AMR", "Star Formation") May contain: Research Areas, Simulations, Derived Data
Simulation / Sim	Simulation collections and datasets ("GC512_L7", "LCDM_h071") May contain: Simulations, Derived Data
Derived / Derived	Derived collections and datasets ("GC512_L7.projections", "Spectra") May contain: Derived Data

basic taxonomy (Topic, Research Area)

basic provenance (Simulation, Derived)

R. Wagner

SRB Metadata

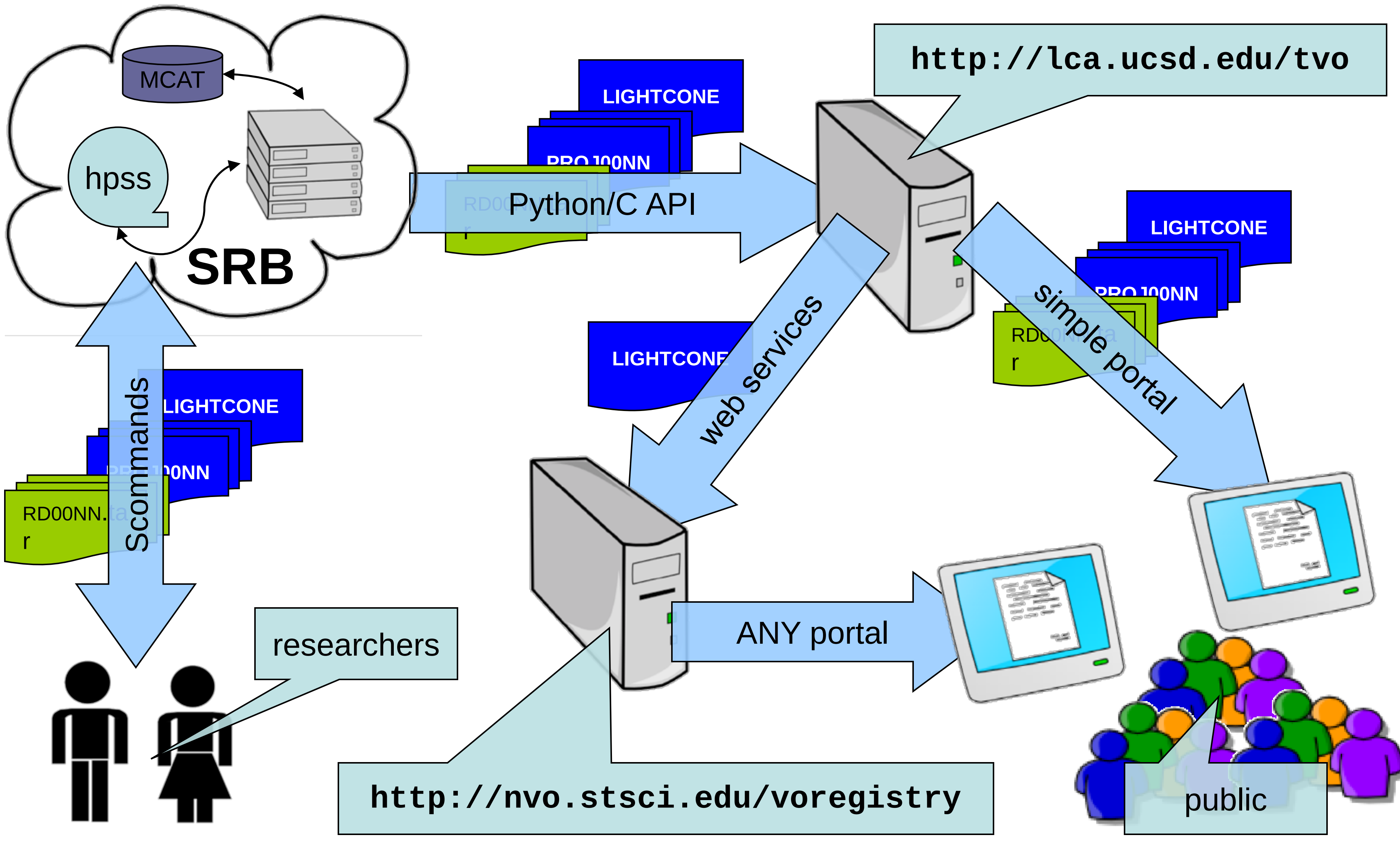
- General & Class
- Inherits Dublin Core

- Link objects to database
- Info for SRB clients

`red`: maps to database

R. Wagner

Basic Archive Design



Relational Database & SRB Extensible Schema

- SRB Extensible Schema:
  - Access separate database via Scommands
  - Eliminates additional database clients
  - Easier to update metadata from various platforms
- Relational Database:
  - Handles abstraction
  - Connects file metadata to other things
  - Can be accessed various ways
  - Well suited to driving web servers

R. Wagner

Year 2 Goals

Production Runs	1) Complete LightCone runs and data analysis 2) First application of radiation transport to cosmic reionization
Data management	1) Provide VO-friendly data products 2) Publish LightCone to the NVO via web services 3) Increase SRB use internally 4) Automate data transport and analysis pipelines
Code Development	1) Implement radiative transfer Step 1 2) Optimize grid hierarchy management *3) amrSolve* linear system solver
Collaboration	1) Radiation transport verification tests (Graziani) 2) LSST data pipeline (Abdulla) 3) LightCone visualizations (SDSC)

Publishing to NVO

- IVOA Theory Interest Group
  - Developing standards for publishing simulation data to VO
  - Nothing finalized
  - Describes some use cases
- Rich metadata will ease any transformation
- References:
  - Theory in the VO
    - http://www.ivoa.net/pub/papers/TheoryInTheVO.pdf
  - IVOA Theory Interest Group
    - http://www.ivoa.net/twiki/bin/view/IVOA/IvoaTheory

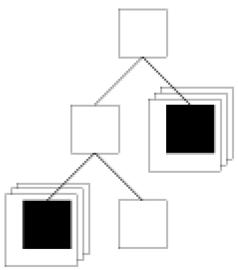
Incorporating Radiation Transport into ENZO: Roadmap

(details in Reynolds et al.)

- Step I: Single group FLD on unigrid, implicitly coupled to ionization and energy equations
- Step II: same as above for AMR
- Step III: Multigroup FLD for AMR, implicitly coupled to ionization and energy equations
- Step IV: Multigroup VTEF for AMR

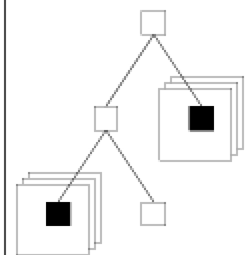
Refactoring Datastructures to Improve Parallel Scalability

Current status



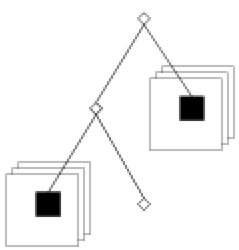
- Enzo uses a large "Swiss army knife" `grid` class
- Only "local" grids store data fields
- But all processors still store all grids
- Total storage is $O(NP)$ $(N(P|G|+|F|)$

Phase 1: Reduce `grid` class size



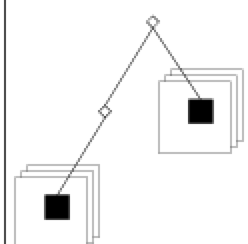
- `grid` class variables may be:
  - Redundant or not needed
  - Constant within level or always
  - Bigger than needed
- Still $O(NP)$, but much smaller constant

Phase 2: Split `grid` class



- Introduce `grid_local` and `grid_remote` subclasses
- `grid_remote` can be much smaller than `grid_local`
- Most grids ($(P-1)/P$) are `grid_remote`
- *Still* $O(NP)$, but even smaller constant

Phase 3: Localize



- Each processor only stores grids it needs
- Truly distributed instead of quasi-distributed
- Finally scalable: $O(N)$
- More difficult to implement

J. Bordner