#### 7/16/2012

Tamás Budavári (Johns Hopkins University)

## SPATIAL SEARCHES IN ASTRONOMY DATABASES MULTI-DIMENSIONAL INDEXING FOR

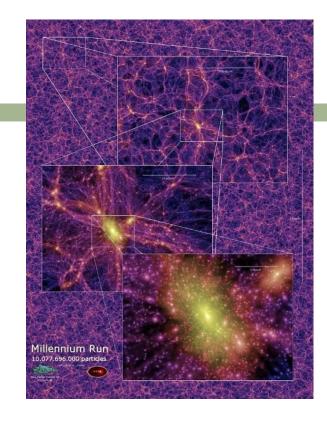
SIMULATIONS AND OBSERVATIONS

Tamás Budavári (Johns Hopkins University)

7/16/2012

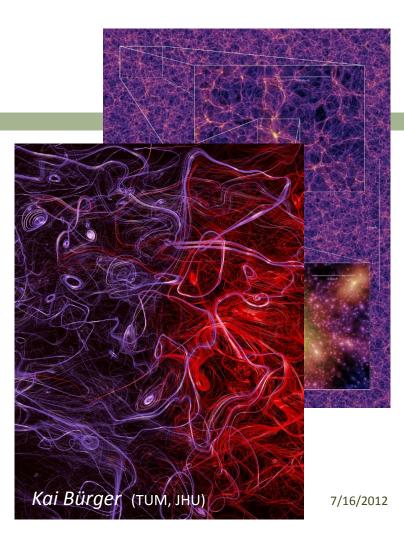
## **Storing Simulations**

- Millennium Run (MPA)
  - 10 billion particles, 64 snapshots
  - FoF groups and merger trees
- Millennium XXL
  - 300 billion particles
- MultiDark Bolshoi
- Turbulence simulations (JHU)
   1024<sup>4</sup> grid, 27TB



## **Storing Simulations**

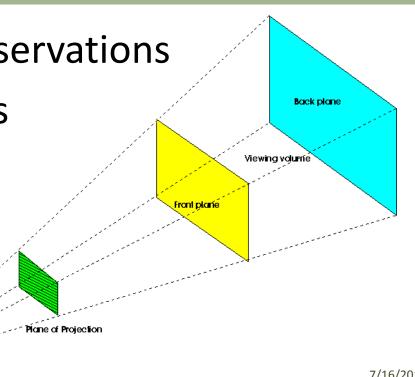
- Millennium Run (MPA)
  - 10 billion particles, 64 snapshots
  - FoF groups and merger trees
- Millennium XXL
  - 300 billion particles
- MultiDark Bolshoi
- Turbulence simulations (JHU)
   1024<sup>4</sup> grid, 27TB



### **Observing Simulations**

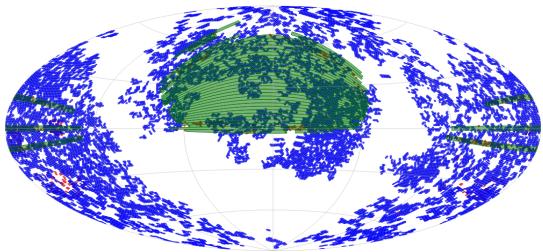
Comparison to real observations
Lots of spatial searches
In the database?

Center of Projection





# For precise window function Virtual surveys



- Query shapes in SQL
   Indexing with space-filling curve
   Combine for spatial searches
   Periodic boxes
  - Celestial sphere

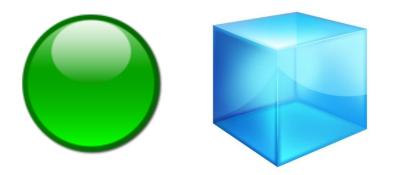


#### Databases

- Which one to use depends on the task
  - □ Sqlite, MySQL, PostGRES, DB2, Oracle, SQL Server
    - Free "express versions" of the big ones, too
- Customization is a must
  - There is always something missing
    - Extend by loading your libraries



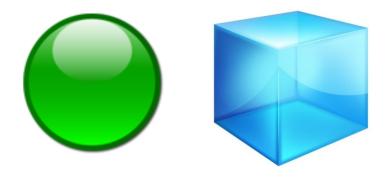
# Geometric primitivesSphere, Box, Cone...



#### IShape interface

TopoPoint Contains(Point p);
TopoShape GetTopo(Box b);
Box GetBoundingBox();

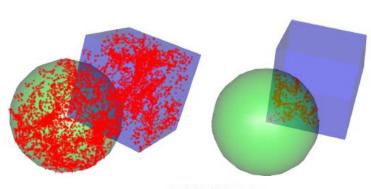
# Geometric primitivesSphere, Box, Cone...

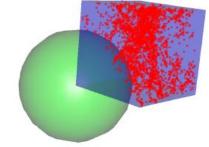


#### IShape interface

TopoPoint Contains(Point p);
TopoShape GetTopo(Box b);
Box GetBoundingBox();

# Composites Intersect, Union, Difference...





In SQL



```
/* Sphere */
Declare @s Sphere = Sphere::New(1,2,3,10);
-- Check if a point is inside
select @s.ContainsPoint(1,2,3), @s.ContainsPoint(99,0,0);
go
```

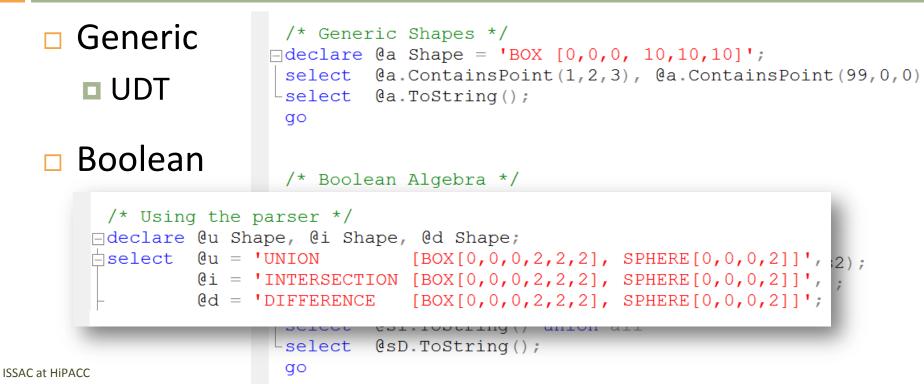
```
/* Box */
Deleclare @b Box = Box::New(0,0,0,10,10,10);
select @b.ContainsPoint(1,2,3), @b.ContainsPoint(99,0,0);
select @b.ToString(); -- string representation
go
```

```
/* String Representation */

declare @x Box = 'BOX [0,0,0, 10,10,10]'

select @x.ContainsPoint(1,2,3), @x.ContainsPoint(99,0,0);
go
```

#### Generic /\* Generic Shapes \*/ $\neg$ declare @a Shape = 'BOX [0,0,0, 10,10,10]'; select @a.ContainsPoint(1,2,3), @a.ContainsPoint(99,0,0) **UDT** Lselect @a.ToString(); qo Boolean /\* Boolean Algebra \*/ $\neg$ declare @s1 Shape = 'BOX [0,0,0, 10,10,10]'; Methods declare $(s_2 \text{ Sphere } [0, 0, 0, 5]';$ declare @sU Shape = Shape::NewUnion(@s1,@s2); declare @sI Shape = Shape::NewIntersection(@s1,@s2); declare @sD Shape = Shape::NewDifference(@s1,@s2); ⊨select @sU.ToString() union all select @sI.ToString() union all Lselect @sD.ToString();



## **Indexing Tables**

- Better performance of queries
  - Instantaneous range searches
  - Fast JOINs
- □ Syntax

CREATE INDEX ix\_Name ON Table
 (X ASC, ...) INCUDE (V, ...)

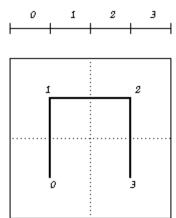
#### Multi-Dimensional

- Map the space to a simple index
- Different kinds of Space-Filling Curves
  - Morton's Z-curve
  - Peano-Hilbert Curve

Tamás Budavári

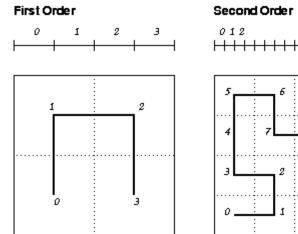
#### Hierarchical space filling

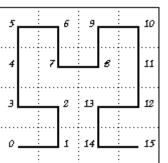
#### First Order



Tamás Budavári

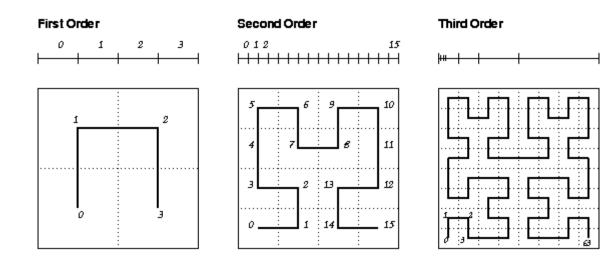
#### Hierarchical space filling



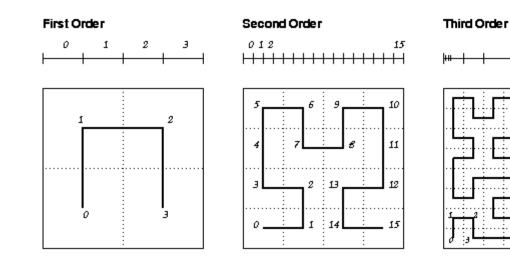


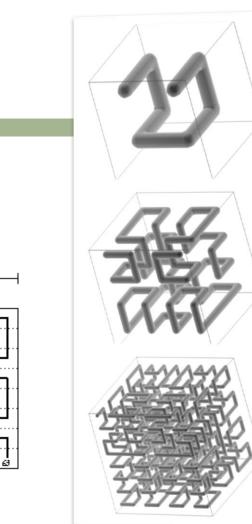
15

#### Hierarchical space filling



#### Hierarchical space filling

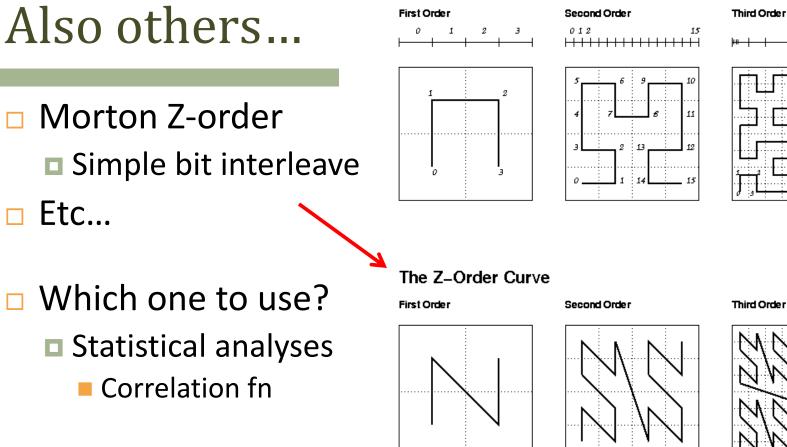




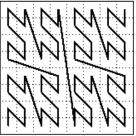
2012

vári

#### The Hilbert Curve

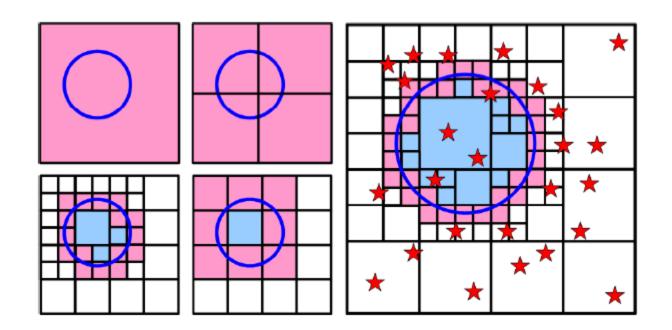




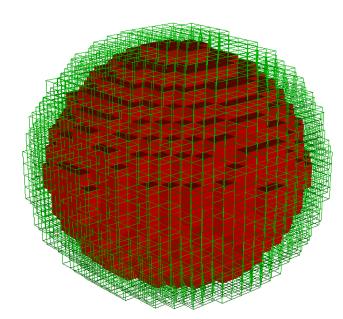


ISSAC at HiPACC

### Divide and Conquer



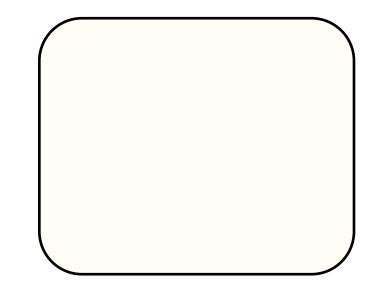
Inside approximationOutside overshoot



Inside approximation
 Outside overshoot

#### They are Key ranges

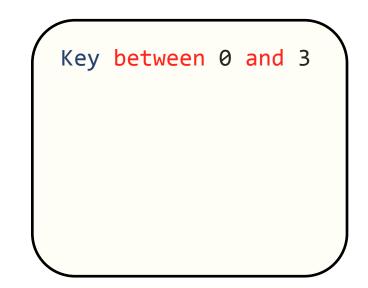
Level 0	Leve	el 1		Lev	vel 2	
	2	2	10	11	14	15
0	2	3	8	9	12	13
0	0	1	2	3	6	7
	0	1	0	1	4	5



Inside approximation
 Outside overshoot

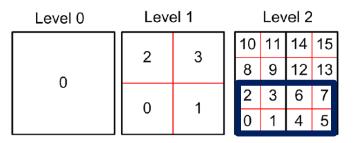
#### They are Key ranges

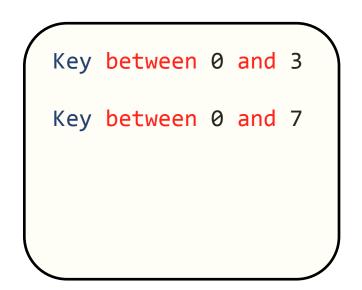
Level 0	Leve	el 1		Lev	el 2	
	2	3	10	11	14	15
0	2	3	8	9	12	13
0	0	1	2	3	6	7
	0	I	0	1	4	5



Inside approximation
 Outside overshoot

#### They are Key ranges





Inside approximation
 Outside overshoot

They are Key ranges

Level 0	Leve	el 1	_	Lev	vel 2	
	2	3	10	11	14	15
0		3	8	9	12	13
0	0	1	2	3	6	7
	0	1	0	1	4	5

Tamás Budavári

Key between 0 and 3

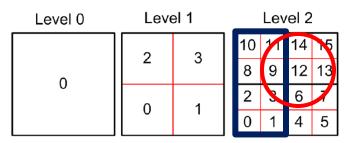
Key between 0 and 7

Key between 0 and 3 or

Key between 8 and 11

Inside approximation
 Outside overshoot

They are Key ranges



Tamás Budavári

Key between 0 and 3

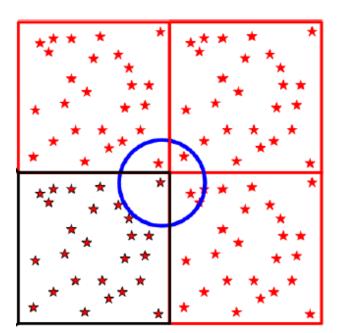
Key between 0 and 7

Key between 0 and 3 or

Key between 8 and 11

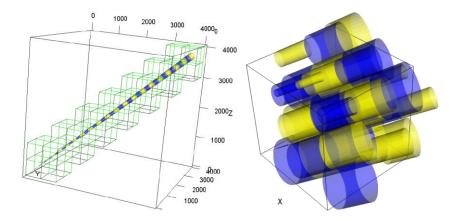
#### **Periodic Boundaries**

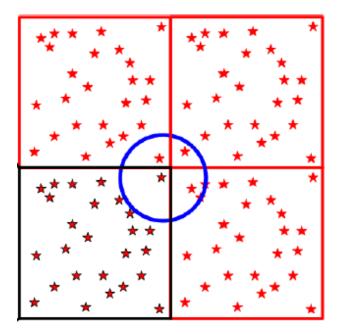
Infinite with periodicity
 Have to search all boxes



#### **Periodic Boundaries**

# Infinite with periodicityHave to search all boxes





### Searching in SQL

Key filterBy Cover

```
ShiftX,-Y,-Z
Where?
```

ISSAC at HiPACC

```
/* Peano-Hilbert Cover */
Description for the select * from for for the select * from for the select * select * from for the select * select * from for the select * from f
```

```
Real!
```

🗆 E.g.,

```
/* Multiple searches around POI */
@with QueryShapes (FoFID, Shape) as
  (
    select top 10 FoFID, Shape::NewSphere(X,Y,Z,10)
    from MilliMil..FoF
    where SnapNum=63 order by M_TopHat200 desc
  )
  select distinct s.FoFID, g.GalaxyID, g.X+c.ShiftX as X,
    g.Y+c.ShiftY as Y,
```

```
g.Z+c.ShiftZ as Z
```

```
from QueryShapes s
```

```
g.Z+c.ShiftZ))
```

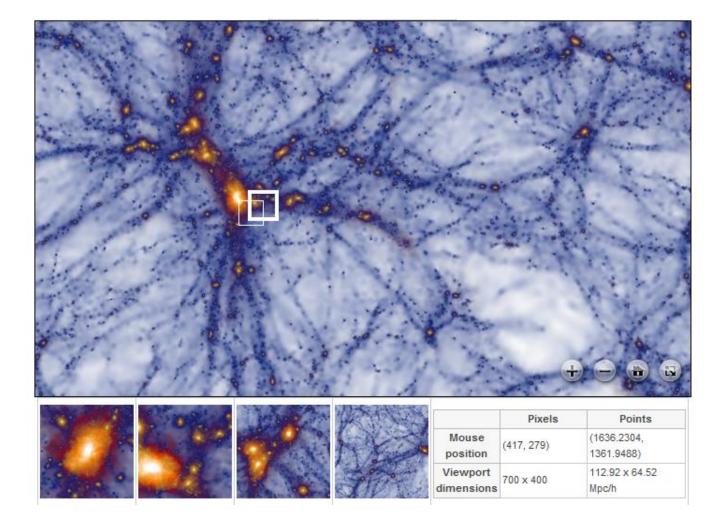
#### **Online Interfaces**

Largest simulations Search and visualize 10 billion+ objects and growing... Indra 512 simulations Coming soon at JHU

	se - Windows Internet Explorer		- 0 >
	p://www.mutidark.org/MultiDark/MyDB	🔹 🔛 🔧 🗙 🔁 Bing	ρ.
Ele Edit View Fa	vorites Iools Help is simulation • Search • · · · · · · · · · · · · · · · · · ·		Sign In •
Favorites	🛅 demos 👻 🛅 tools 👻 🛅 varia 👻 🤔 The General Theory of Empl 😢 ch6 👘 MUCN - Manag		
SquirrelMail 1.4.21	MutDark Database 🗴 😘 Milennium Simulation 😘 Documentation	👌 🕶 Bage	<ul> <li>Tgols -</li> </ul>
A State of the	MultiDark Database		and the second s
A CALL		<b>白体与自生性的学校</b> 会	2
Home			
Query Form	Query the MultiDark Database		
Credits	Welcome Gerard Lemson!		
Documentation	*** All data provided here are currently only in a pre-release state and MUST NOT (yet) be used Place your SQL statement directly in the text area below and submit your request by pressing one of the	for SCIENTIFIC purposes. ***	
Databases	Please note, that there is a timeout and row limit for each query. Streaming oueries, return unlimited number of rows in CSV format and are cancelled after 420 seconds.		
Bolshoi MDR1	Browser queries: return a maximum of 1000 rows in HTML format and are cancelled after 30 seconds.		
miniMDR1 Spatial3D_db		2	
Private (MvDB)			
Databases			
	abases - Windows Internet Explorer	+2 X Bing P -	
1		* X 🖸 Bing 🖉 -	
dit Vew Favorites To	ols Help	sian In •	
orites 🕜 SquirreiMal 1	.4.21 😘 Milennium Simula 🗴 😘 Job Runner - Main 🔗 VirtualTelescopeConfi	🔄 💌 Bage 🕶 Tools 🕶	
Virgo - Mil	lennium Database		
ntation	ennium Database	<u> </u>	
entation S/Acknowledgments	Watome Grand Lanzon. Bravaning averains tutkon unlimited number of revisi in CBV format and are cancelled after 420 seconds. Brevisar queries retirum maximum of 10000 overs in HTML, format and are cancelled after 30 seconds.		
ntation 5/Acknowledgments	Watome Grand Lamon. Benannig avera et utilinité number of reve in CBV format and are cancelled after 420 seconds. Brosser quelles return mainume (1000 over in HTLL, brinat and are cancelled after 38 seconds. #46erc #mail:conversit(slapics, justaver)) as a fortaire diseasement infondisor	VA: 354	
ntation 5/Acknowledgments	Witcome Canadia         Second parameters           Basering approximation unlimited number of trave in CDV formal and are cancelled after 420 seconds.         Biosear queries return mainmum of 10000 rove in HTML formal and are cancelled after 20 seconds.           Extrast and cancel and the control of	<u> </u>	
ntation S/Acknowledgments tion	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode		2
ntation S/Acknowledgments dion Patabases axies orrees	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode	Guery (Meant)	2006 👻
ntation S/Acknowledgments tion Natabases axies oxtes 010a id	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode		2 2016 •
ntation SlAcknowledgments dion Vatabases axies otrees 010a di nniumIII nii	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode	Guery (Meant)	2006 ¥
ntation SAcknowledgments tion States	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode	Cuery (stream) Cuery (stream)	- 20% •
ntation sidebases sides sides sides sides sidebases sides sidebase	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode	Ouer (stream) Deer (stream) Esplan	
ntation /Acknowledgments ion atobases xixies xixies trees of trees trees of trees of trees of trees of trees trees of trees of trees of trees of trees of trees of trees of trees of trees of trees of trees of trees of trees of trees trees trees trees of trees of trees	Viscome Gerard Lamson. Bearing queries return unlimited number of rows in CBV format and are canceled after 420 seconds. Browser queries return maximum of 10000 rows in rNLK format and are canceled after 50 seconds. From millionance/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode Miner discussion/destartgoode	Ouer (stream) Deer (stream) Esplan	20%
ntation /Acknowledgments tion atabases xixies xixies trees 0100 mini mi mi mi mi mi mi mi mi mi mi mi mi mi	Wetome Grand Lamon. Benaming years in them unlimited number of rows in CBV format and are cancelled after 420 seconds. Benaming years and unlimited number of rows in CBV. Lomat and are cancelled after 320 seconds. Frieder standards (Standards (Standa	Ouer (stream) Deer (stream) Esplan	
entation SAcknowledgments dion witabases strik	Watchine Grand Lamson.           Bharanning queries intern unlimited number of rows in CBV format and are cancelled affer 429 seconds.           Entrance queries intern unlimited number of rows in CBV format and are cancelled affer 429 seconds.           Facility Tameters (Tameters) are international are cancelled affer 429 seconds.           Facility Tameters (Tameters) are international are cancelled affer 429 seconds.           Facility Tameters (Tameters) are international are cancelled affer 429 seconds.           Material cameters (Tameters) are international are cancelled affer 429 seconds.           Material cameters (Tameters) are international are cancelled affer 429 seconds.           Material cameters (Tameters) are international are cancelled affer 429 seconds.           Material cameters (Tameters) are international are cancelled affer 429 seconds.           Material cameters > 0           Material cameters of rows to relation to be query form.	Ouer (stream) Deer (stream) Esplan	2006 •
ntation Addinoviedgments ion etabases trites trites 010a dommit dommit di With apphotos salatices salatices salatices (4/vDB) Dotabases (4/vDB) Dotab	Watcome Grand Lamasol.           Brearing gaves that multimited number of reves in CBV format and are cancelled after 420 seconds.           If a second secon	Ouer (stream) Deer (stream) Esplan	-
witchowiedgments SAcknowiedgments Sicher witchoses avides avides d d ummant all ummants al	Vestions Grand Lamosi.         Barance grance.         Bara	Ouer (stream) Deer (stream) Esplan	
entation SiAcknowledgments don valabases axies of too of t	Wateries Grand Lamson.         Braaming operation that multimed number of rows in CDV format and are cancelled after 20 seconds.         In the control operation of the control operation of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operation.         Interfer attendence (Depted Lamson) is an information of the control operattendence (Depted Lamson) is an information of	Ouer (stream) Deer (stream) Esplan	2 2016 ¥
entation SAcknowledgments don variabases axies axies otrees otrees otrees otrees don minimum all diffi mapshots claiores claiores claiores claiores tatores ta	Vestions Grand Lamosi.         Barance grance.         Bara	Ouer (stram) Deer (stram) Expan Hep	3 30% *
Virgo - Mil vitation Stacksnowledgments dion Databases arises ar	Wateries Gerard Lamson.         Baraming operation that individe outputs of rows in CDV formation date canceled after 20 seconds.         Individe Transmission of the CDV operation of the CDV formation date canceled after 20 seconds.         Individe Transmission of the CDV operation of the CDV ope	Ouer (stram) Deer (stram) Expan Hep	2 2006 *
Antalion SJAcksoveldgments tion Valabases axies of too within anii ali Milli M	Waters Grand Lamon.           Branning operation that multiplied number of rows in CDV formating and are cancelled after 20 seconds.           Interfer amount in the number of rows in CDV formating are cancelled after 20 seconds.           Interfer amount interference Displicit. Lamonare of the composition of	Ouer (stram) Deer (stram) Expan Hep	- - - 
retation station stations retabases strice	Wateries Gerard Lamson.         Baraming operation that individe outputs of rows in CDV formation date canceled after 20 seconds.         Individe Transmission of the CDV operation of the CDV formation date canceled after 20 seconds.         Individe Transmission of the CDV operation of the CDV ope	Ouer (stram) Deer (stram) Expan Hep	- - -

*i*ári



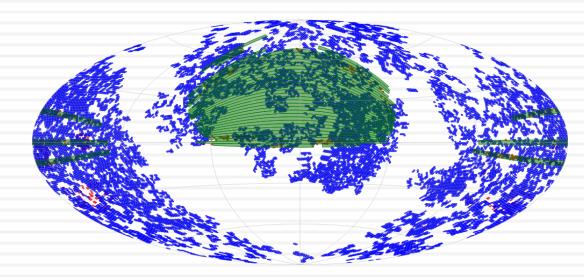


#### Web Services

Programming interfaces
 Execute SQL queries
 Most flexible
 Inject probes in simulations

- Turbulence
- Cosmology





### No Sky Coverage?

2 M. F. Pedbost et al.

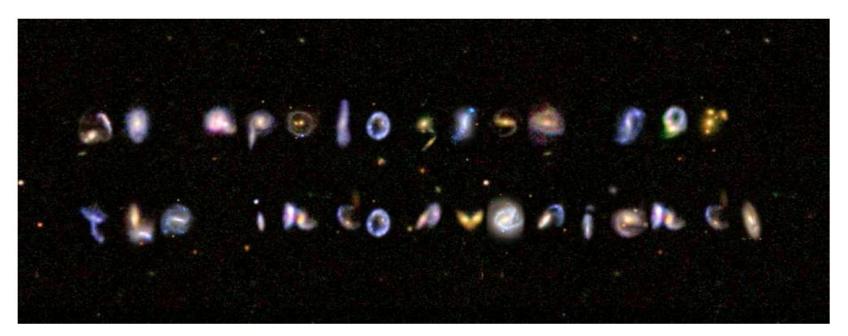
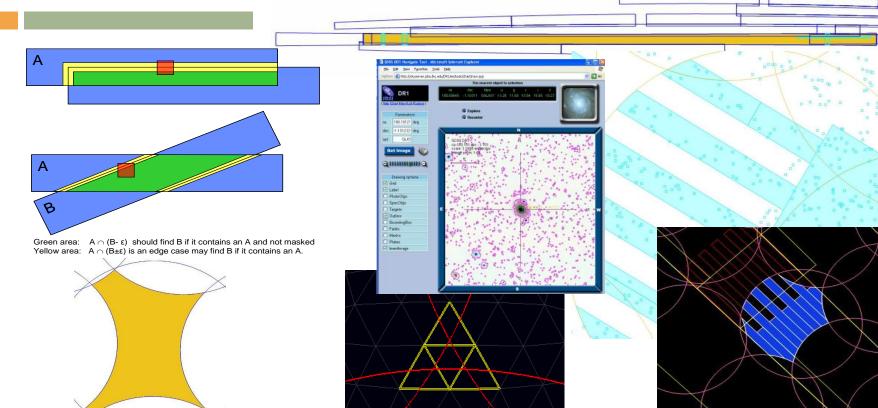


Figure 1. SDSS colour composite image (vri) for our prototype unusual galaxy cluster, at  $RA = 16^{h}23^{m}76^{s}$ ,  $Dec = +97^{\circ}62'12''$ , identified by Galaxy Zoo participants. North is at the top, East is to the left.

### **Spherical Geometry**



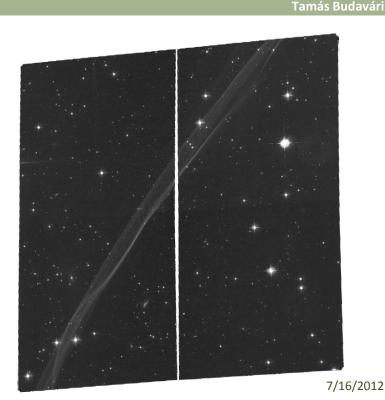
## Approaches to Consider

- Pixel mapsSensitivity, etc...
- Equations of shapes
   Spherical "vector graphics"

□ And beyond...

#### An Observation

- FITS header with WCS
   Image dimensions map to the geometry
- More exposures?
  - No common pixel coordinate-system
  - Overlapping areas



#### **Common Pixels**

- Pre-defined pages of an atlas
   Standard in cartography
- Image pyramids of hierarchical pixels
   Including HTM, Igloo, HEALPix, SDSSPix, etc...



Always approximate!

## **Practical Implementation**

 Looking at Terapixels
 We know how to work with images
 Now have commodity Internet
 We have cheap hard-drives
 WorldWideTelescope.org Sky in G••gle Earth

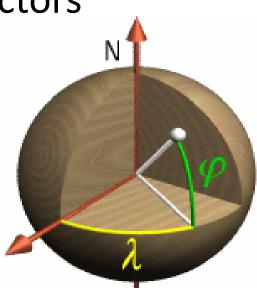
Integrated catalogs for efficiency
 How about more surveys?



# Drawing with Equations

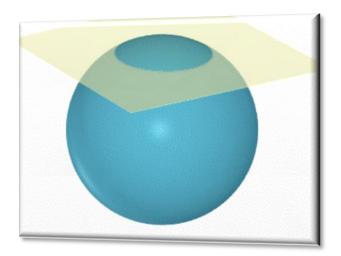
#### Working with 3D normal vectors

- Benefits include
  - No wraparound
  - No projections
  - No singularities



# Drawing with Equations

- Direct 3D approach
   Halfspace → Circle/Cap
   Convex → Simple shapes
- Region
   Unions of convexes
  - Patches on the sphere



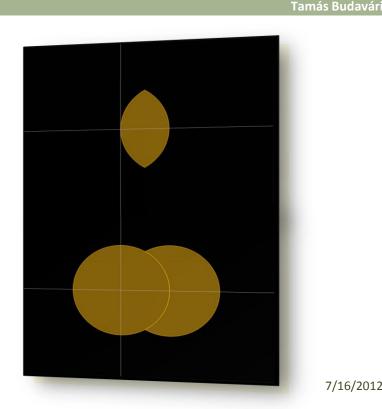
## Point in Region Test

- *Halfspace:* one side of a plane  $(\vec{n}, c)$ ■ Inside, when  $\vec{n} \cdot \vec{x} > c$
- *Convex:* a collection of halfspaces
   Inside, when inside all halfspaces
- *Region:* a collection of convexes
   Inside, when inside any convex

# Shape Operations

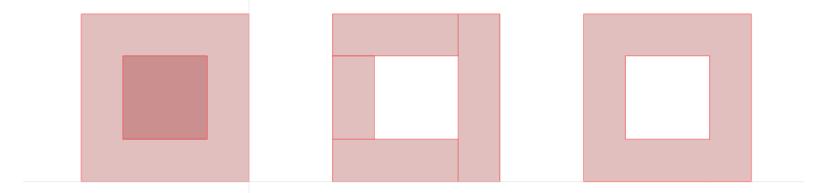
- Intersection Concat halfspace lists
- Union
  - Concat convex lists
  - Unique coverage
  - Analytic area

#### Boolean algebra



## Difference of Convexes is a Region

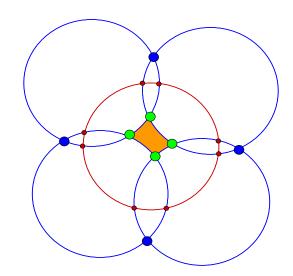
#### □ The set of Regions is closed for the Boolean ops



## Simplification

#### Eliminate redundant halfspaces

- First handle trivial combinations of constraints
- Then solve geometry on the surface
  - Derive Roots, Arcs, Patches
- Eliminate redundant convexes
  - Some trivial cases, but...
- Make convexes disjoint
  - Unique coverage, area, etc.
- Stitch together convexes
  - When possible



### SphericalLib .NET

- □ C# code ~ 10k lines
  - OS independent (Windows, Un\*x w/ Mono)
  - Documentation via Sandcastle
- Great performance!
   Sloan Digital Sky Survey in 10s (13× larger than USA in area)

## Numerical Imprecision

- Double precision calculations
   IFFE 754 standard
- Degeneracy
  - When are two vectors the same?
- Spatial resolution limit
  - Roughly 30 cm on Earth
- Lots of tricks from Graphics Gems



Sky coverage of the Sloan Digital Sky Survey's 5<sup>th</sup> Data Release and the Galaxy Evolution Explorer's 2<sup>nd</sup> Public Release

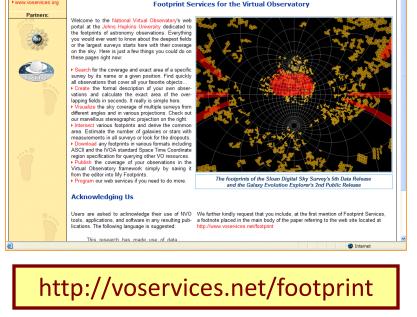
# Region in SQL

```
DECLARE @s VARCHAR (MAX), @r VARBINARY (MAX),
        @z VARCHAR (MAX), @u VARBINARY (MAX)
SELECT Qs = 'REGION CIRCLE J2000 180 0 60',
       @z = 'POLY J2000 180 0 182 0 182 2 180 2',
       Qr = sph.fSimplifyString(Qs),
       @u = sph.fUnion(@r, sph.fSimplifyString(@z))
SELECT sph.fGetArea(@r), sph.fGetArea(@u)
-- 3.14151290574491 6.35572804450646
/*
 SQL Server Execution Times:
   CPU time = 0 ms, elapsed time = 1 ms.
*/
```

## **Footprint Services**

#### □ All about coverage

- Editor and calculator
- Online public repository
- On-the-fly visualization
- STC translator, etc...
- Web services
  - Simple programming



**Footprint Service** 

NVØ

VO Services

Tamás Budavári

login | registe

#### ASTRONOMICAL DATA ANALYSIS SOFTWARE AND SYSTEMS XVI

#### The Westin La Paloma Resort & Spa Tucson, AZ, USA 15–18 October 2006

This volume contains papers presented at the 16th annual conference on Astronomical Data Analysis Software and Systems (ADASS XVI). The meeting themes addressed challenges and solutions for very large, compute-intensive systems. Major news astronomical facilities such as ALMA, Gasis, LSST, and the Square-Klönneter Array, to name but a few, are in various strages of planning and design. Their rouporting data systems will exceed the current state of the art in many dimensions by multiple orders of magnitude, and will demand a major portion of the facility construction and operating costs. A key challenge will be to extract their full science potential through distributed data systems that make new, innovative use of Grid, database, web service, and visualization technologies.

The 13 invited and 36 contributed talks, 116 posters, seven fieor demonstration booths, three focus demo sessions, and seven Birds-of-a-Feather sessions combined to provide a thorough overview of the latest developments in astronomical software, applications, data facilities, and algorithms. The key topics for this meeting were Challenges & Solutions for Large Data, Advances in Imaging & Caliboriton Algorithms, Quality Management in Artsnomical Data Management Systems, Medern Grid Computing in Astronomy, Architectures for Large Astronomy. Software Systems, and Solar Neighborhood & Planetary Astronomy. While a large number of contributed papers followed these themes, the full range of traditional topics included planning for legacy architers such as that for HST, the continuing development.

The ADASS conference series has a strong tradition of participant-lead topical sessions within the meeting. This year a record number of "Birds of a Feather" sessions were held on a mix of established and emerging topics, including FTTS, Artenomoulo Data Processing and the VO. The Emerging Infrastructure of Autonomous Astronomy, Building Observatory Leggery Archives, Pitpeline Processing of Spectroscopic Observations, IRAP Users and Developers, and Next Generation of Visualization Toels for Astrophysics. Summaries of these sessions are included in these proceedings. This book is suitable for software developers, designers of astronomical software systems, astronomere who use astronomical software, and students of any of these fields.





Available online at www.aspbooks.org



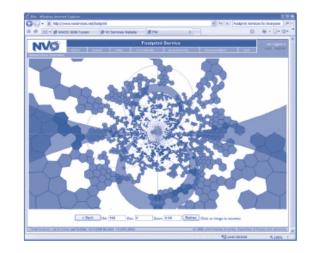
ASTRONOMICAL DATA ANALYSIS SOFTWARE AND SYSTEMS XVI

Edited by Richard A. Shaw, Frank Hill and David J. Bell



#### VOLUME 376

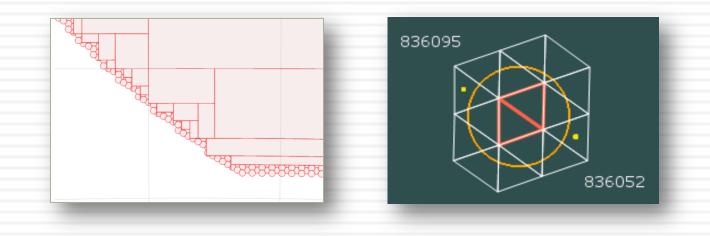
#### ASTRONOMICAL DATA ANALYSIS SOFTWARE AND SYSTEMS XVI



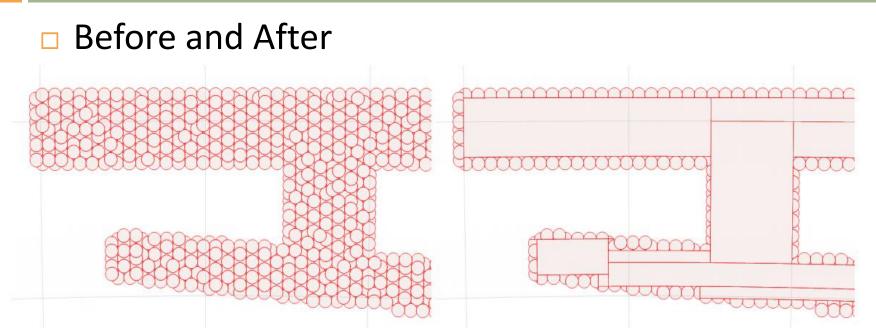
Edited by Richard A. Shaw, Frank Hill and David J. Bell

ASPCS

# <sup>55</sup> Hybrid Solutions

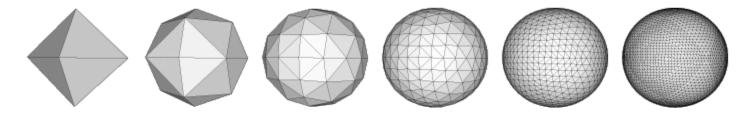


### Heuristic Simplification

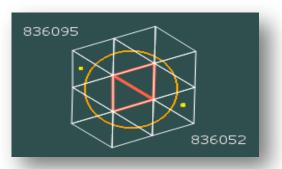


Indexing the Sky

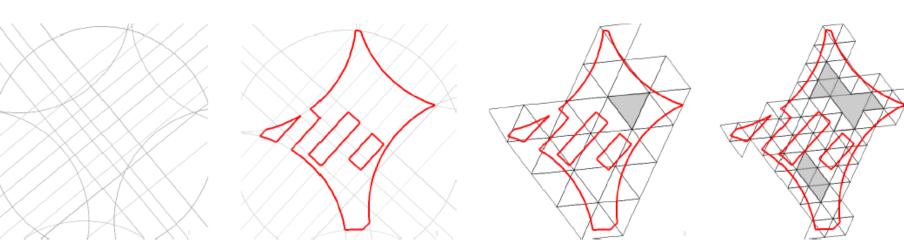
#### Hierarchical Triangular Mesh



 Region approximation
 Fast filtering using HTM ID ranges



# Anatomy of an SDSS Region



# HTM Filtering

```
WITH Cover AS
(
    SELECT * FROM dbo.fHtmCoverRegion
        ('REGION CIRCLE J2000 180 0 10')|
)
SELECT 0.ObjID
FROM PhotoObj AS 0 INNER JOIN Cover AS c
    ON 0.HtmID BETWEEN c.HtmIDStart AND c.HtmIDEnd
```



- Store simulations, e.g., the reference Millennium
  - Simulations take 10x longer than analysis
- Databases enable fast searches
  - Custom routines
  - Space-filling curves

Direct comparison of observed universe to sims