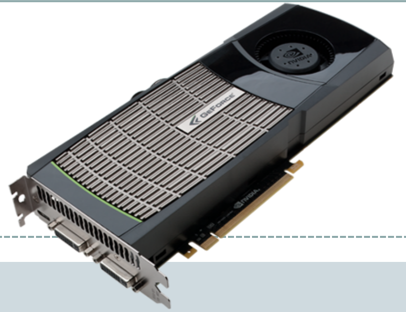
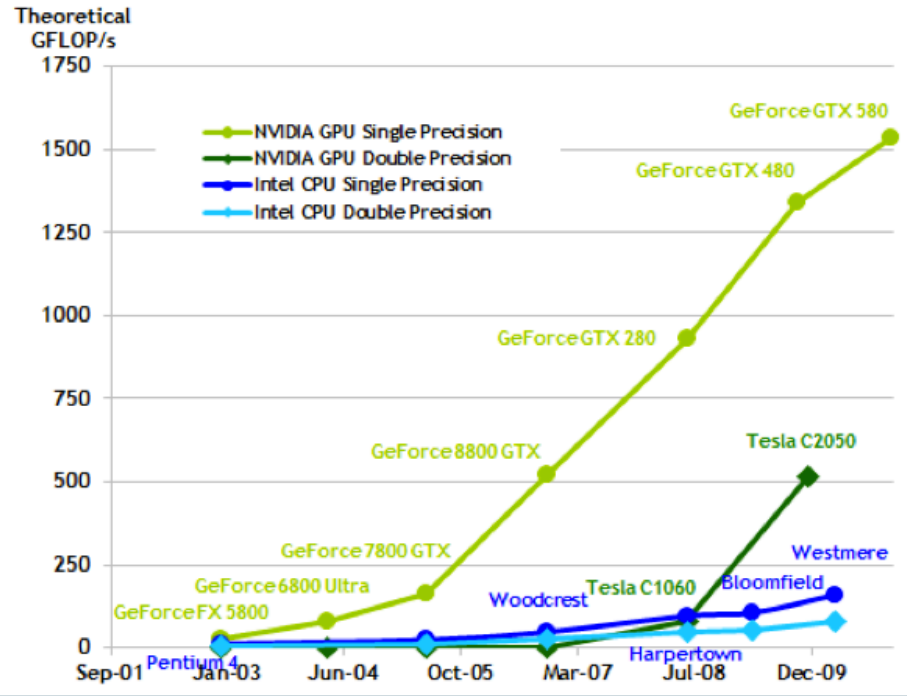
GPU在天文中的应用

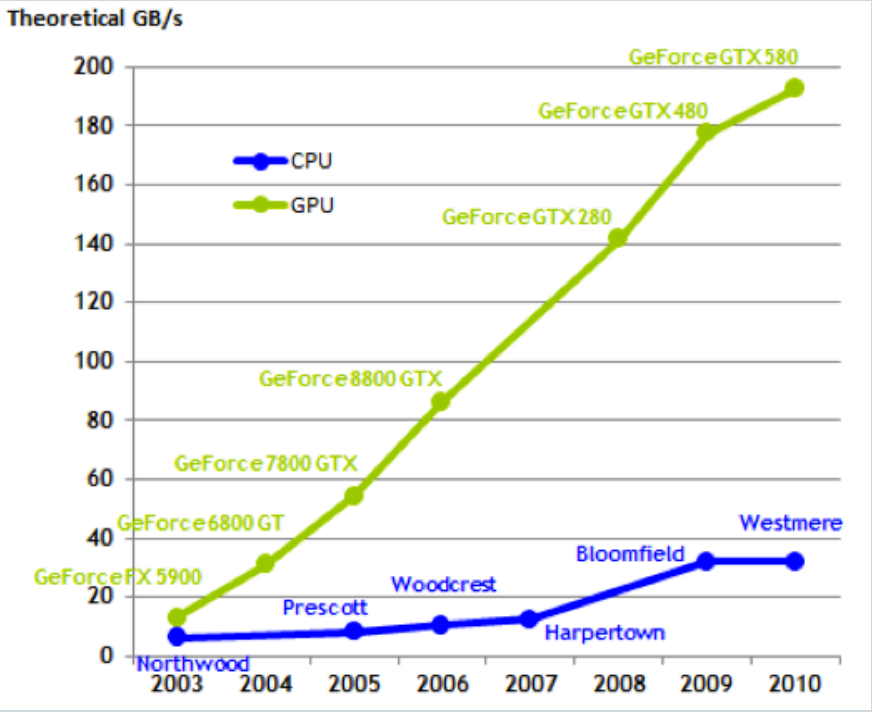
三峡大学/国家天文台
田海俊

2012年11月30日

GPU的发展

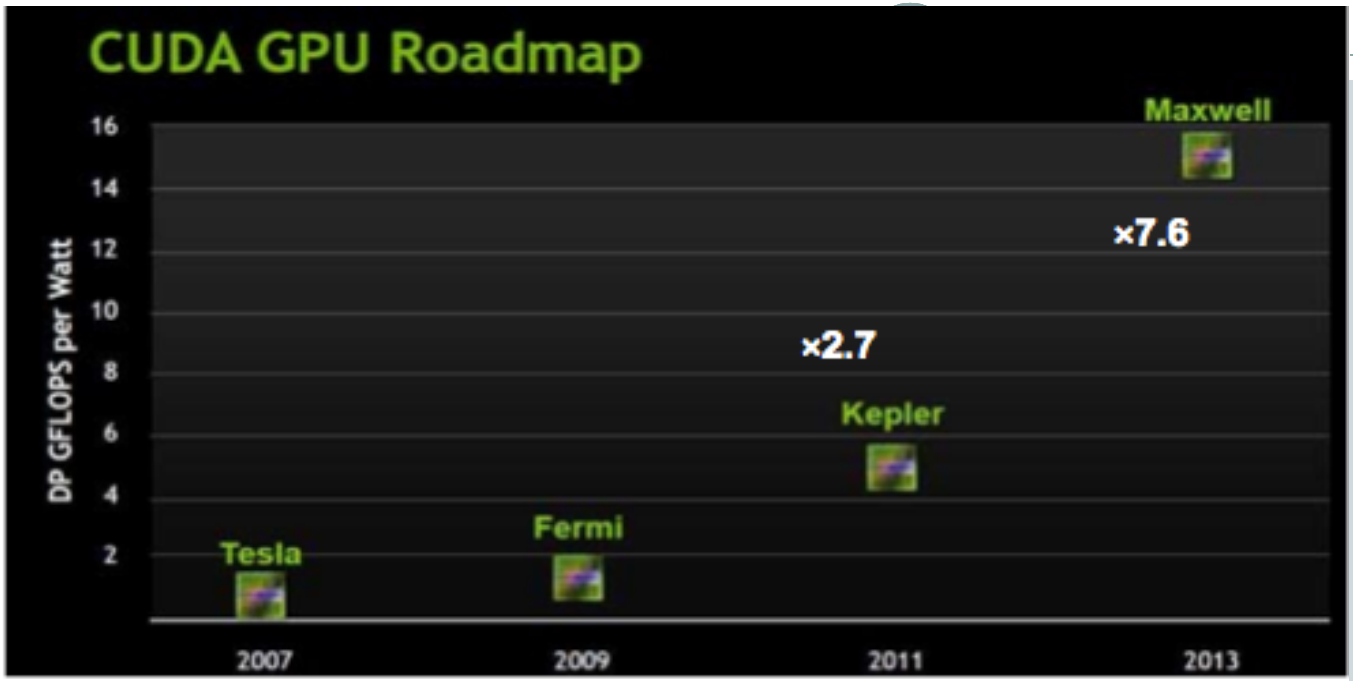






GPU计算能力和数据传输能力

CUDA GPU & PCI-E Roadmap



GPGPU card	Expected release year	Performance
Fermi	2009	1.0×
Kepler	2011	2.7×
Maxwell	2013	7.6×
???	2015	15.2×
???	2017	30.4×

i.e. GPU cards may be expected to deliver 15-30 TFLOP by 2016-2018.

Power dissipation expected to remain constant at 250W per card.

Bus version	Expected release year	Bandwidth
2	2007	8 GB/s
3	2011-2	16 GB/s
4	2015 ??	32 GB/s
5	2019 ??	64 GB/s

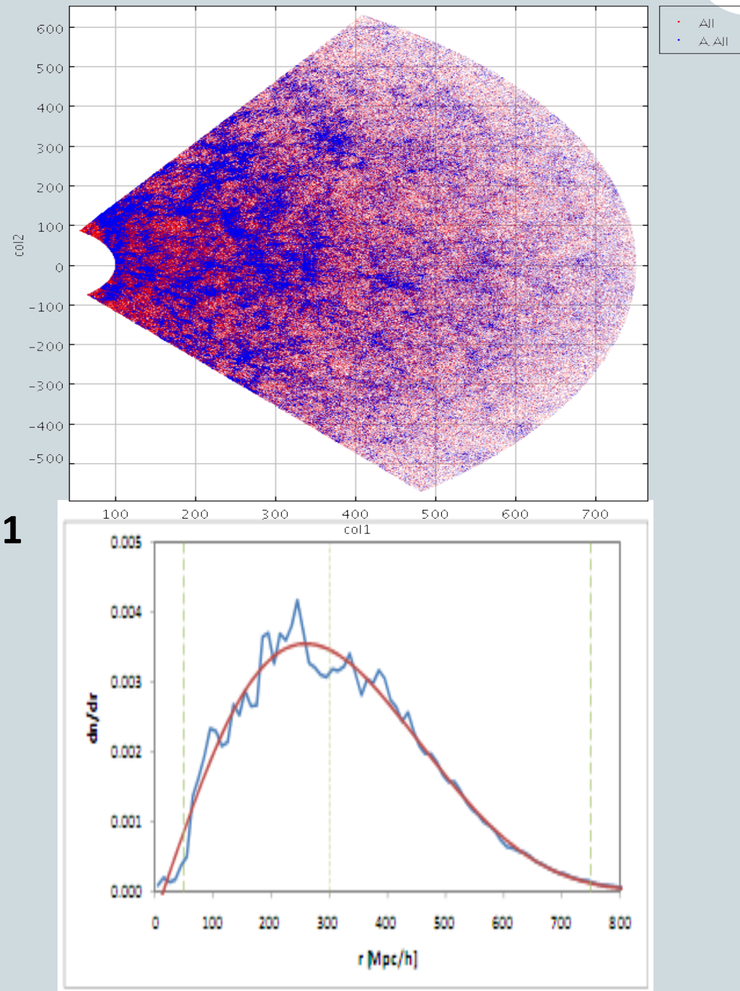
GPU技术在天文中的应用

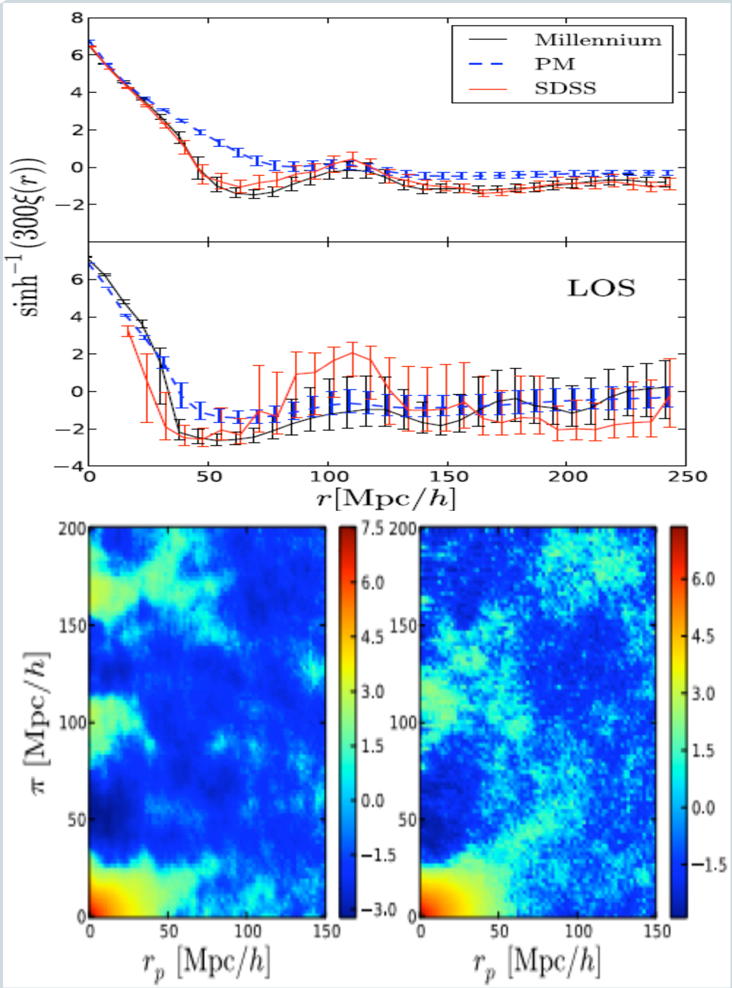
- N体模拟：有助于研究行星系统、恒星凝聚、银河系的演化等（其他领域：分子动力学、基本粒子散射模拟、交通道路模拟等）；
- 射电干涉仪：FX/XF相关器.
- 其他：引力透镜（Thompson2010，Bate2010 ）、地外行星搜寻(Ford2009)、AMR加速(Schive2010)、尘埃温度计算(Jonsson2010)、再电离模拟(Aubert2010)，星系拟合(Barsdell2011)、引力波搜寻（Chung2010)、天文数据的可视化（Hassan2011）、天文数据分类（彭南博2011、裴彤2010）

彭南博,张彦霞《科研信息化技术与应用》2011

GPU应用（一）宇宙大尺度结构

Tian2011





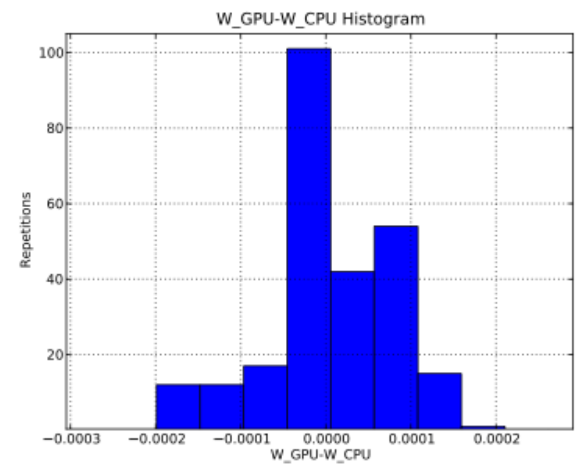
计算由GPU完成，目前将尝试进一步利用GPU完善一套2PCF、nPCF等大尺度结构统计工具

相关函数

Rafael2012(GPU application on 2pcf)

Input file lines	CPU (s)	GTX295 (s)	C1060 (s)	C2050 (s)
$0.43 \cdot 10^6$	$3.60 \cdot 10^4$	$3.01 \cdot 10^2$	$2.91 \cdot 10^2$	$2.19 \cdot 10^2$
$0.86 \cdot 10^6$	$1.44 \cdot 10^5$	$1.20 \cdot 10^3$	$1.16 \cdot 10^3$	$8.76 \cdot 10^2$
$1.00 \cdot 10^6$	$1.98 \cdot 10^5$	$1.61 \cdot 10^3$	$1.56 \cdot 10^3$	$1.17 \cdot 10^3$
$1.29 \cdot 10^6$	$3.24 \cdot 10^5$	$2.68 \cdot 10^3$	$2.59 \cdot 10^3$	$1.97 \cdot 10^3$
$1.72 \cdot 10^6$	$5.76 \cdot 10^5$	— — —	$4.64 \cdot 10^3$	$3.51 \cdot 10^3$
$3.45 \cdot 10^6$	$2.32 \cdot 10^6$	— — —	$1.88 \cdot 10^4$	$1.41 \cdot 10^4$
$6.89 \cdot 10^6$	$9.22 \cdot 10^6$	— — —	$7.45 \cdot 10^4$	$5.61 \cdot 10^4$

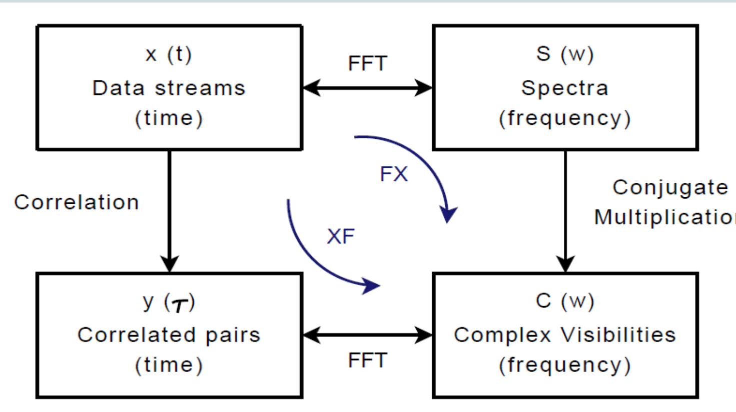
CPU和GPU执行时间对比

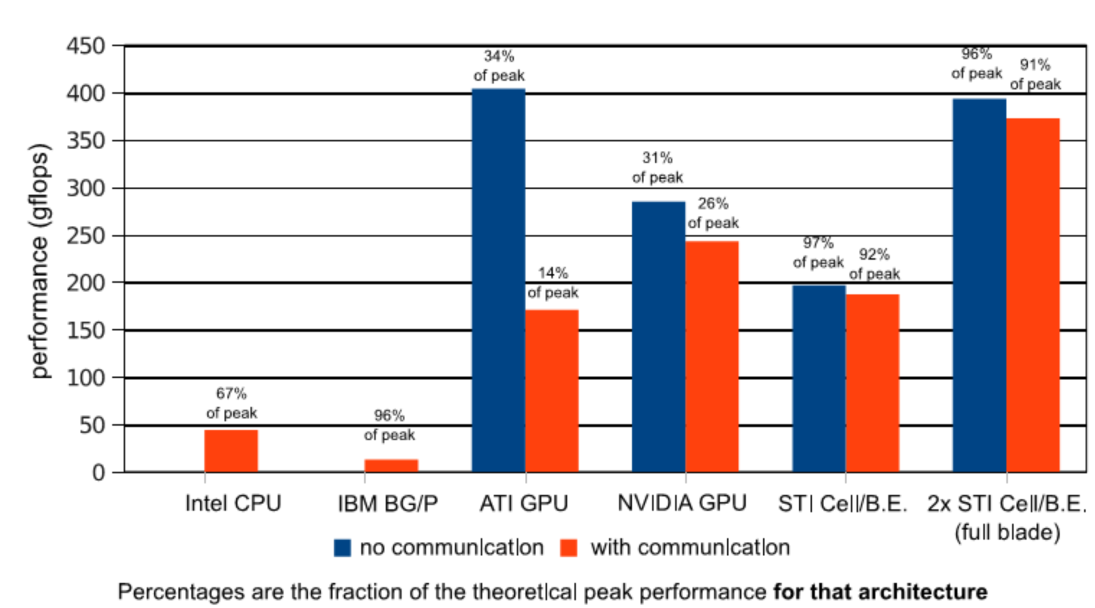


残差

Optimizing the Computation of N-Point Correlations on Large-Scale Astronomical Data (Martch2012)

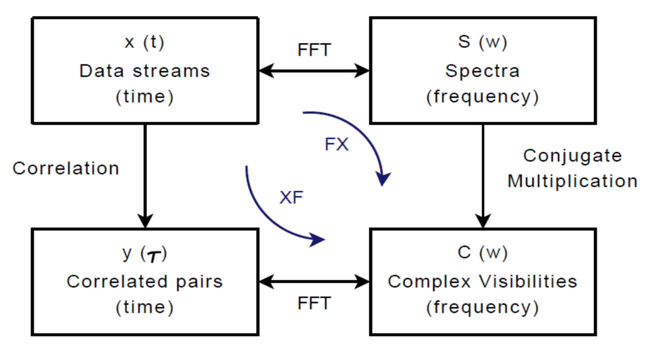
GPU应用（二）射电干涉仪的相关器





Nieuwpoort2009

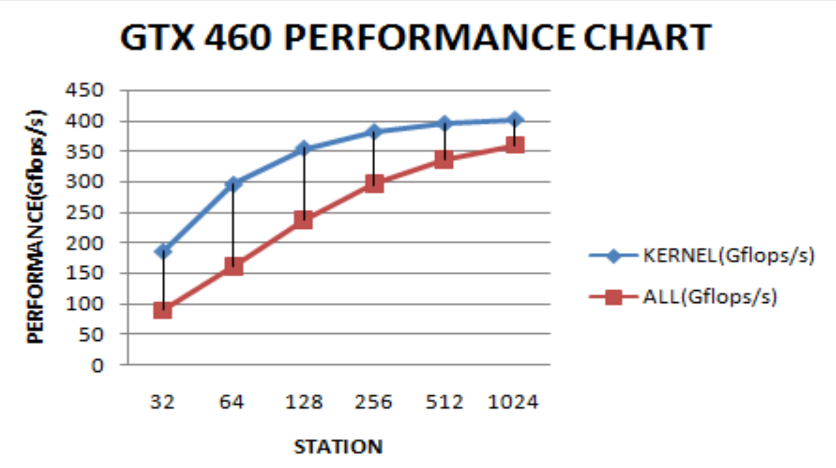
GPU应用（二） 射电干涉仪的相关器

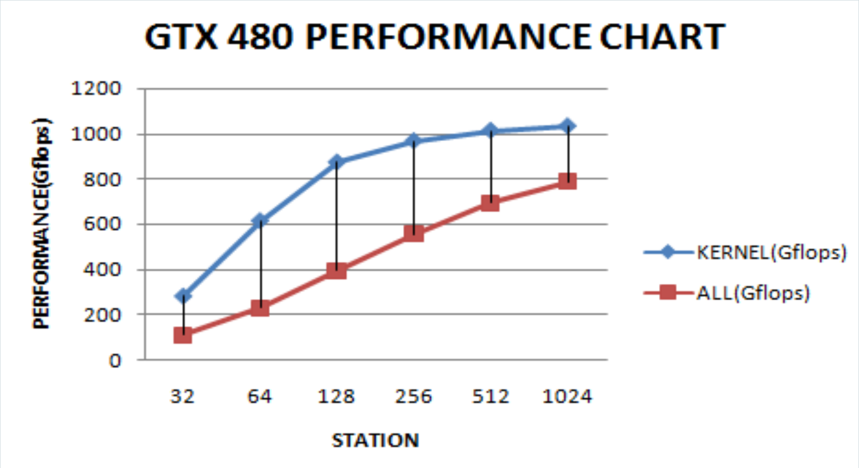


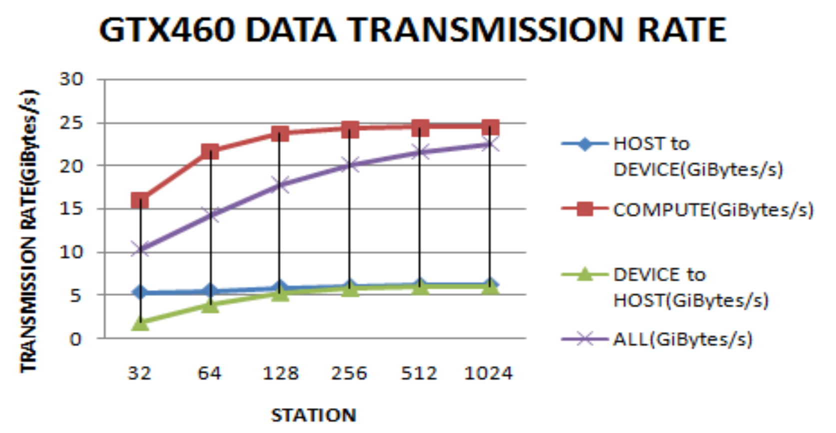
Parameters	Phase I	Phase II
N(antenna)	96	1000
N(pole)	2	2
N(baseline)	4656	500.5K
Frequency	700~800Mhz	700~1500Mhz
N(sampling)	200Mhz	1600Mhz
N(FFT length)	1024	1024
N(bit)	4	4

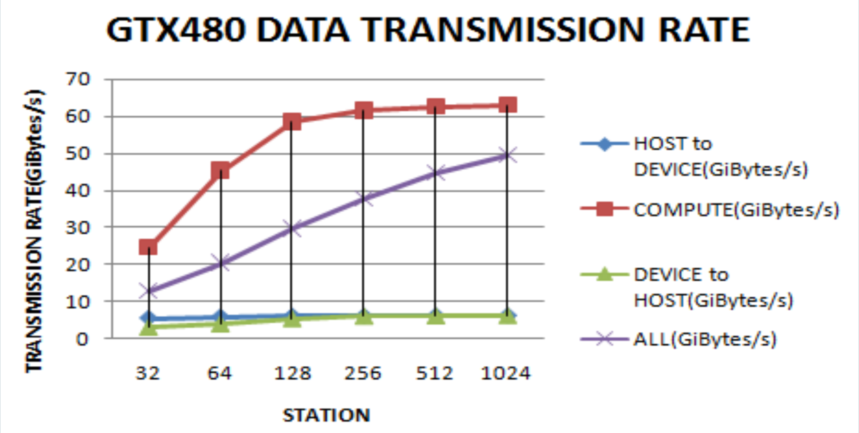
TianLai	F-Step	Corner Turn	X-Step
Phase I 192 data streams ~200MB/s/stream 4.8 GB/s total	**3.072GFLOPS per stream** **~600GPLOPS total**	**0 GFLOPS** **But** **4.8 GB/s**	**Total:** **29.8 TFLOPS**
Phase II 2000 data streams ~1600MB/s/stream 3.2TB/s Total	**24.6 GFLOPS per stream** **~50TPLOPS total**	**0 GFLOPS** **But** **3.2 TB/s**	**Total:** **25.6PFLOPS**

基于单GPU的测试结果





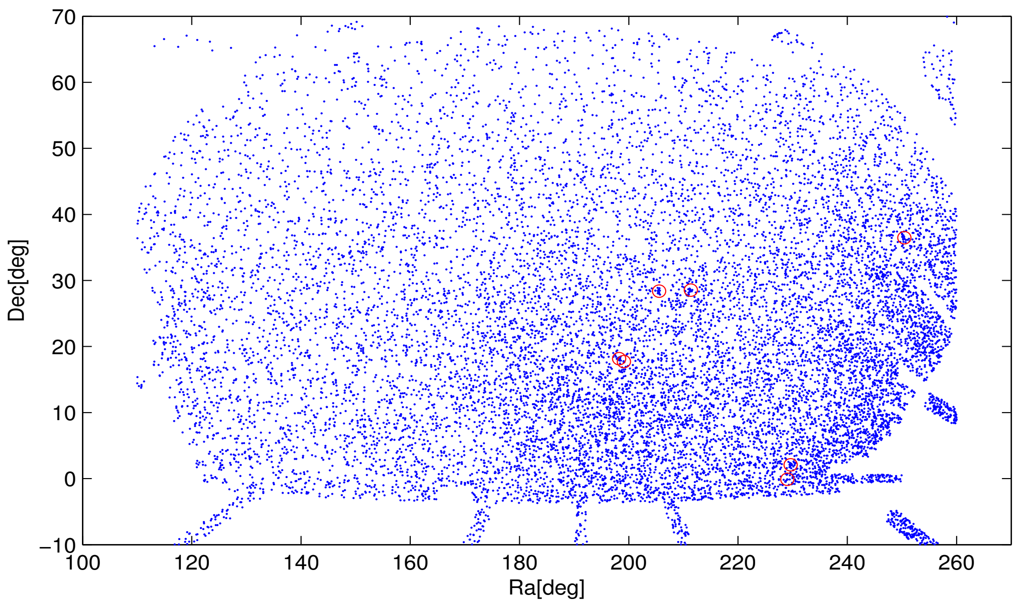


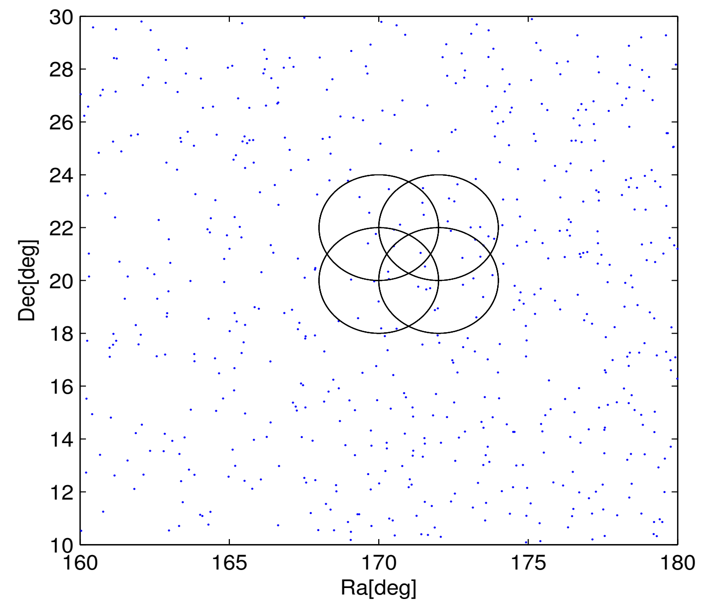


GPU应用（三）:银河系整体消光（基本完成）

- 意义(河外)： 消光对星系相关函数的畸变(Wenjuan Fang 2011)、Ia型超新星观测误差（彭秋和老师质疑）等;
- 红化率：CCM89, Donnell94,Fitzpatrick99等.
- 消光图：SFD98（最流行）
- 问题（SFD98）：某些区域高估40%（Arce&Goodman99, Stanek98, Chen99, Yasuda07,Rowles&Froebrich09）； 高估两倍或更多（|b|<40， Dobashi05）；～14%(SDSS高质量数据Schlafly2011）
- 近两年： Schlafly2010(Blue Tips Method), Jones&West2011 (M dwarf Spectra), 3D消光（LAMOST Jiang, B.W., Liu, X.W.）

GPU应用（三）：银河系整体消光（基本完成）



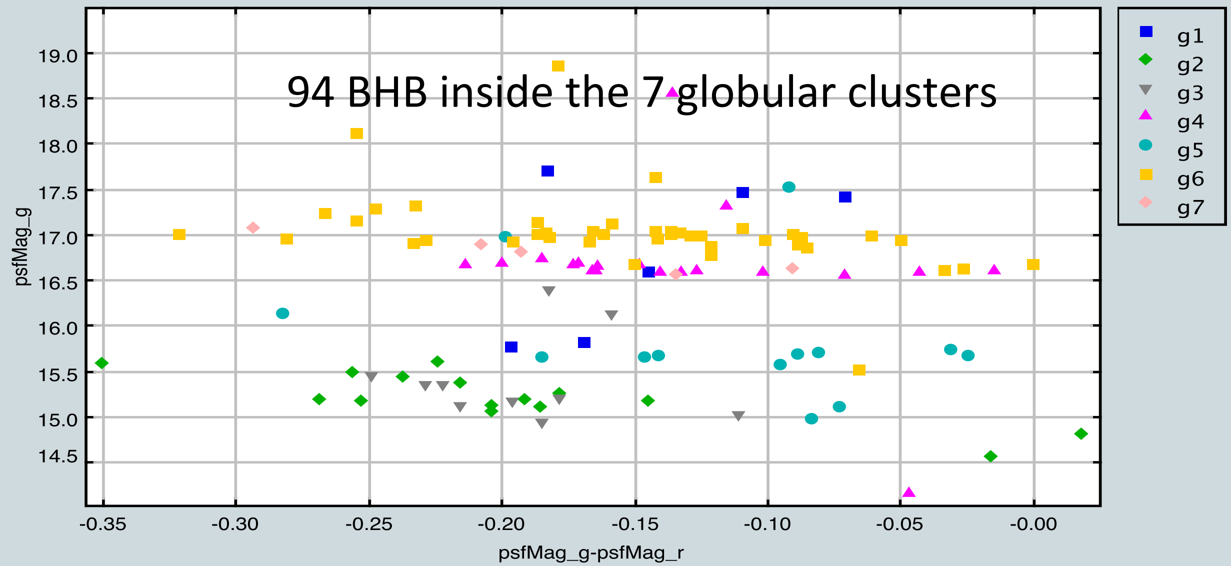


The blue points(left subplot) are 14265 BHB candidates , selected from Smith2010 (SVM trained on the spectroscopic sample from Xue2008), with the following criteria:

1. ra > 110 and ra < 260, dec > −10 and dec < 70
2. g − r > −0.3 (Yanny2000)
3. g<20.0

The red small circles on the left: 7 globular cluster, inside which including 94 BHB stars. On right subplot, the cross black circles are the line-of-sight cells with 2 degree radius

样本选取（Intrinsic BHB sample）



7 GLOBULAR CLUSTERS TAKEN FROM HARRIS 1996 (2010 EDITION)

ID	RA	DEC	l	b	R_{sun}	E(B-V)	V_{BH}	g_{BH} [a]	[Fe/H]
NGC6205	16 41 41.24	+36 27 35.5	59.01	40.91	7.1	0.02	14.9	15.2	-1.53
NGC5024	13 12 55.25	+18 10 05.4	332.96	79.76	17.9	0.02	16.81	17.0	-2.10
NGC5053	13 16 27.09	+17 42 00.9	335.70	78.95	17.4	0.01	16.69	16.8	-2.27
NGC5466	14 05 27.29	+28 32 04.0	42.15	73.59	16.0	0.016	16.52	16.6	-1.98
NGC5272	13 42 11.62	+28 22 38.2	42.22	78.71	10.2	0.01	15.64	15.6	-1.5
NGC5904	15 18 33.22	+02 04 51.7	3.86	46.80	7.5	0.03	15.07	15.2	-1.29
Pal5	15 16 05.25	-00 06 41.8	0.85	45.86	23.2	0.03	17.51	17.5	-1.41

Table 2 lists 94 BHB stars, which are targeted from the above 7 clusters.

方法

Bayes' Theorem (for cell *i*):

$$p_i(\mathbf{E}|D_i) = p(D_i|E_i)P(\mathbf{E}_i) \tag{1}$$

$$p_i(\mathbf{E}|D_i) = \prod_{j=1}^{N_i}\sum_{k=1}^{N_u} p(\hat{\mathbf{c}}_{ij}|(\mathbf{E}_i, \mathbf{c_k}))p(\mathbf{E}_i) \tag{2}$$

Likelihood:

$$p(\hat{\mathbf{c}}_{ij}|(\mathbf{E}_i, c_k)) = \frac{1}{(2\pi|\mathbf{\Sigma}|)^{m/2}} \exp(-\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x}) \tag{3}$$

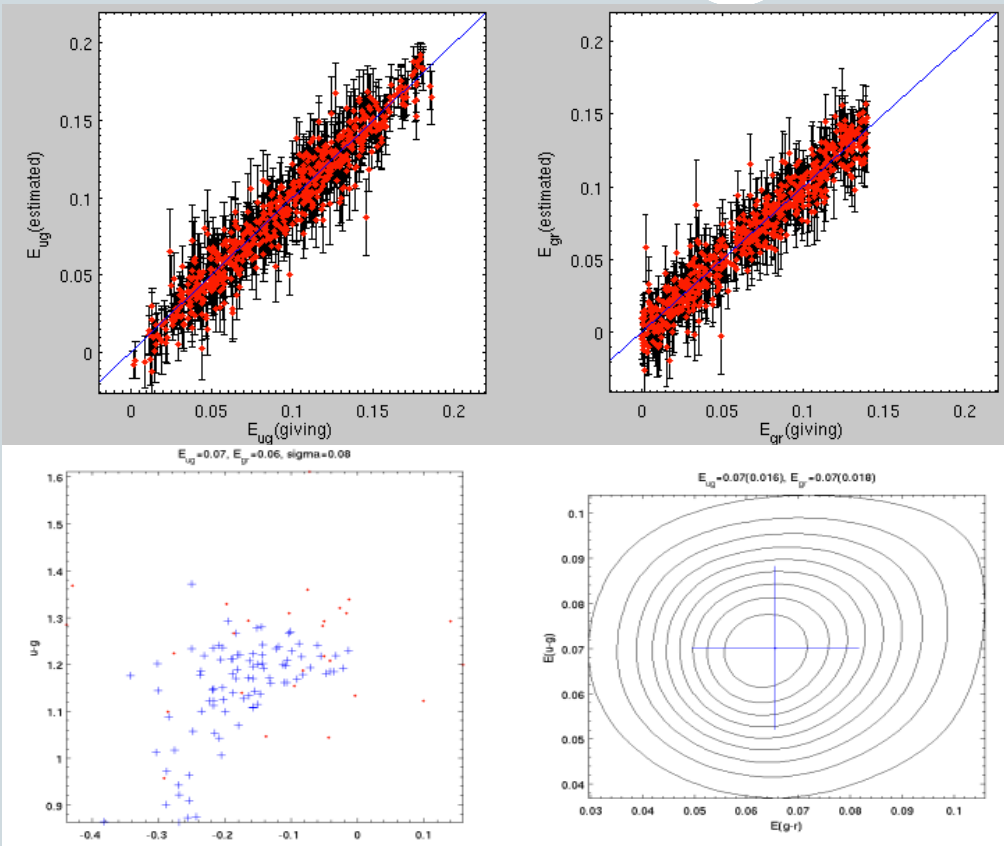
where the $\mathbf{x} = \mathbf{E} + \mathbf{c}_k - \hat{\mathbf{c}}_{ij}$, and the m means rank of $\mathbf{\Sigma}$, while $\mathbf{\Sigma}$ is the covariance matrix of the measurement of the color indexes of the star j, shown in below:

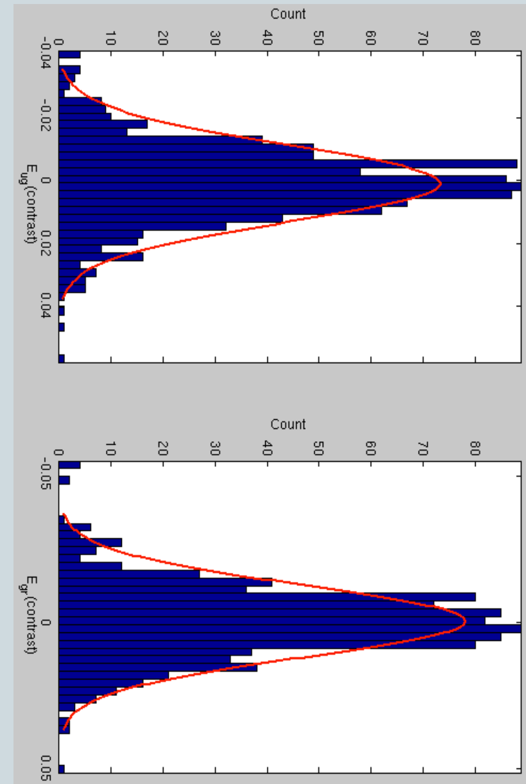
$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_u^2 + \sigma_g^2 & -\sigma_g^2 & 0 & 0 \\ -\sigma_g^2 & \sigma_g^2 + \sigma_r^2 & -\sigma_r^2 & 0 \\ 0 & -\sigma_r^2 & \sigma_r^2 + \sigma_i^2 & -\sigma_i^2 \\ 0 & 0 & -\sigma_i^2 & \sigma_i^2 + \sigma_z^2 \end{bmatrix} \tag{4}$$

here, the σ_u,σ_g, σ_r, σ_i, σ_z are errors of magnitudes.

方法校验

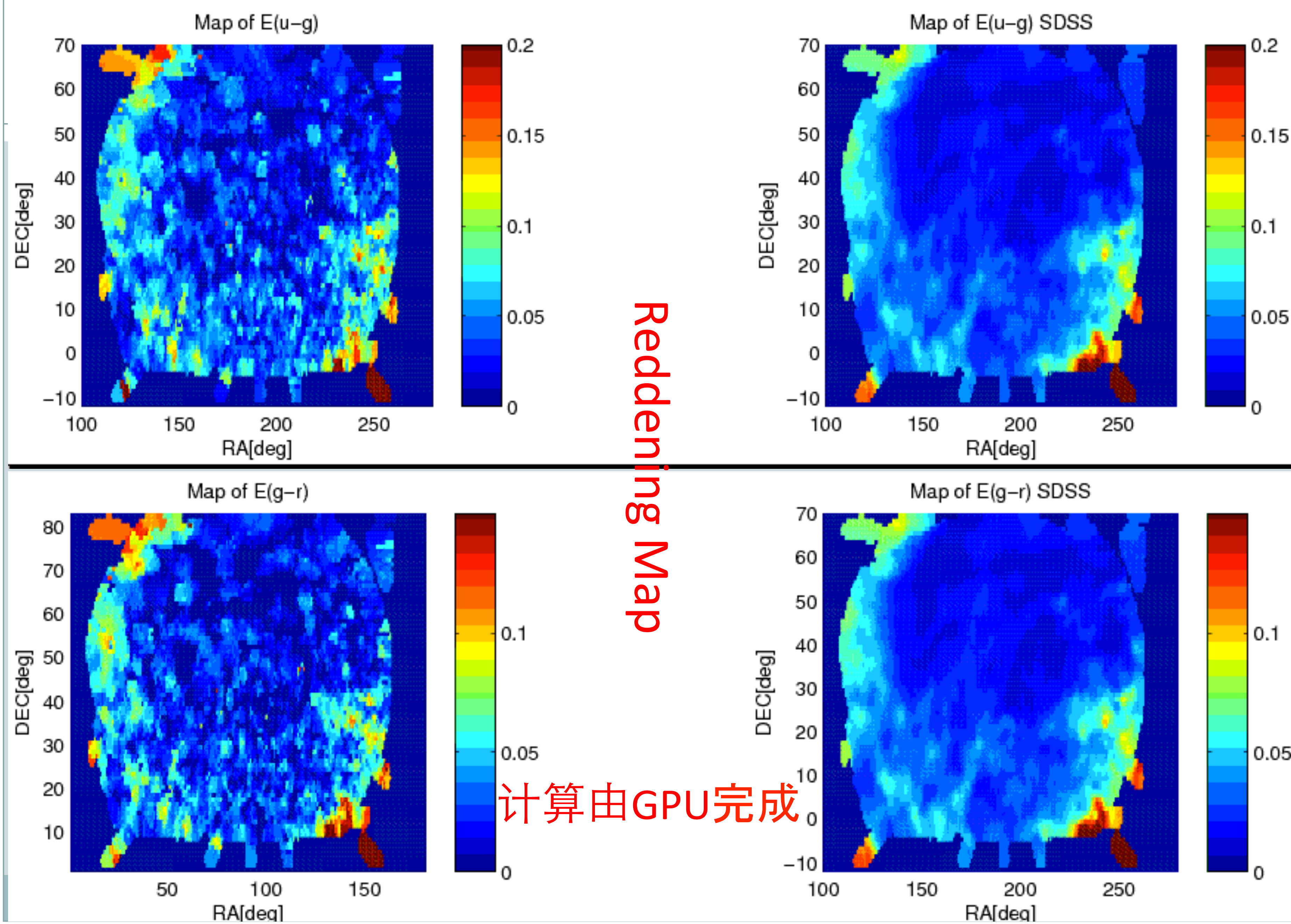
900 simulations



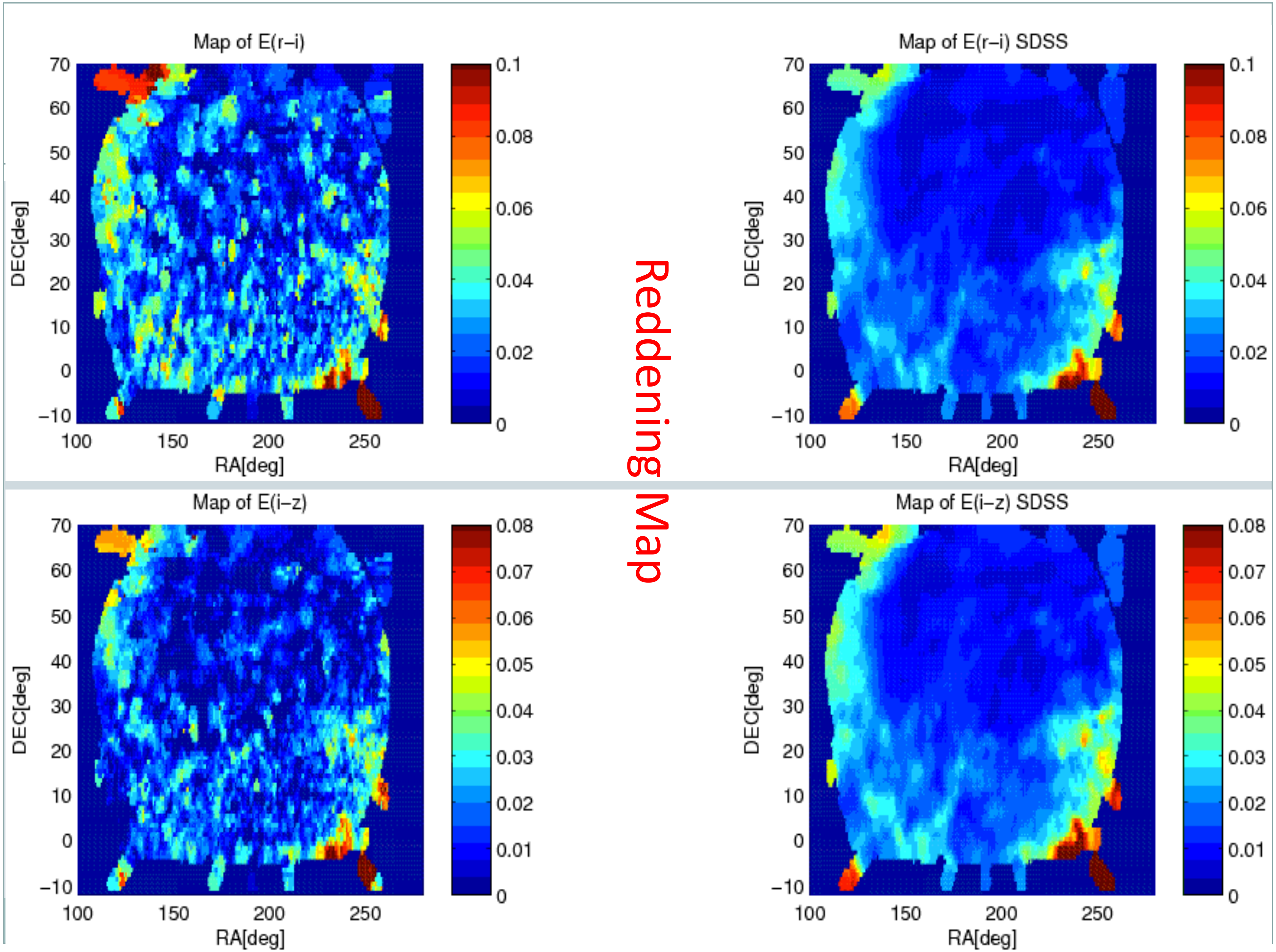


Giving the Reddening with 0~0.2 in u-g, and photometric measurement error in ugr with 0.02~0.08

Reddening Map



计算由GPU完成


Reddening Map
Map of E(r–i)
Map of E(r–i) SDSS
Map of E(i–z)
Map of E(i–z) SDSS
DEC[deg]
RA[deg]

Rv的测量（方法）

CCM89

$$\langle A(\lambda)/A(V)\rangle = a(x) + b(x)/R_V \ .$$

Optical/NIR: $1.1\ \mu m^{-1} \le x \le 3.3\ \mu m^{-1}$ and $y = (x - 1.82)$;

$$a(x) = 1 + 0.17699y - 0.50447y^2 - 0.02427y^3 + 0.72085y^4 + 0.01979y^5 - 0.77530y^6 + 0.32999y^7\ ; \quad (3a)$$

$$b(x) = 1.41338y + 2.28305y^2 + 1.07233y^3 - 5.38434y^4 - 0.62251y^5 + 5.30260y^6 - 2.09002y^7\ . \quad (3b)$$

➔

E(u-g) = Au - Ag

$$E(u-g) = ((a_u + \frac{b_u}{R_V}) - (a_g + \frac{b_g}{R_V})) * A_V \quad (6a)$$

$$E(g-r) = ((a_g + \frac{b_g}{R_V}) - (a_r + \frac{b_r}{R_V})) * A_V \quad (6b)$$

$$E(r-i) = ((a_r + \frac{b_r}{R_V}) - (a_i + \frac{b_i}{R_V})) * A_V \quad (6c)$$

$$E(i-z) = ((a_i + \frac{b_i}{R_V}) - (a_z + \frac{b_z}{R_V})) * A_V \quad (6d)$$

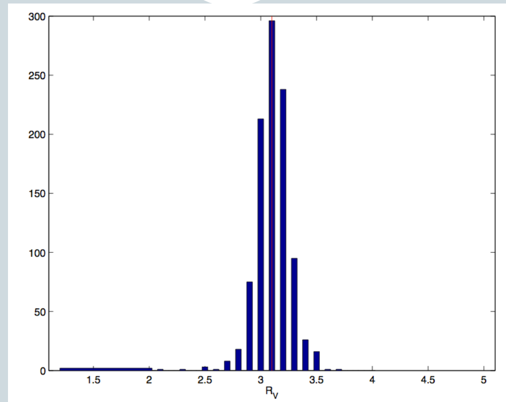
$$\chi^2 = \sum_{k=1}^{N_{obj}} \left(\frac{(E(u_k - g_k) - \bar{E}(u_k - g_k))^2}{\sigma^2_{ug_k}} + \frac{(E(g_k - r_k) - \bar{E}(g_k - r_k))^2}{\sigma^2_{gr_k}} + \frac{(E(r_k - i_k) - \bar{E}(r_k - i_k))^2}{\sigma^2_{ri_k}} + \frac{(E(i_k - z_k) - \bar{E}(i_k - z_k))^2}{\sigma^2_{iz_k}} \right) \quad (7)$$

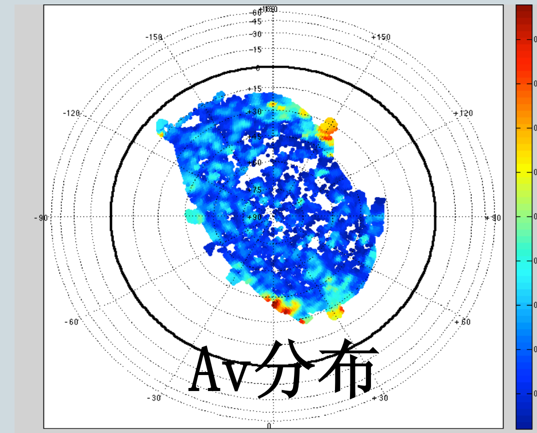
卡方最小确定每个cell的Rv和Av值

Rv的测量（实测和校验）

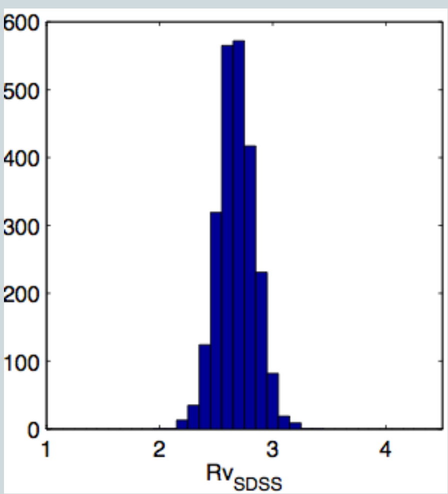
方法检验：

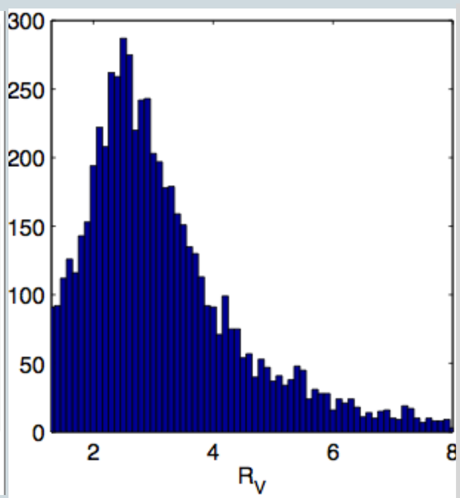
Giving Rv=3.1 and the E(b-v)（SFD98）, and then calculate the E(u-g) and E(g-r) according to the CCM89. Finally, go back to estimate the Rv using our procedure.

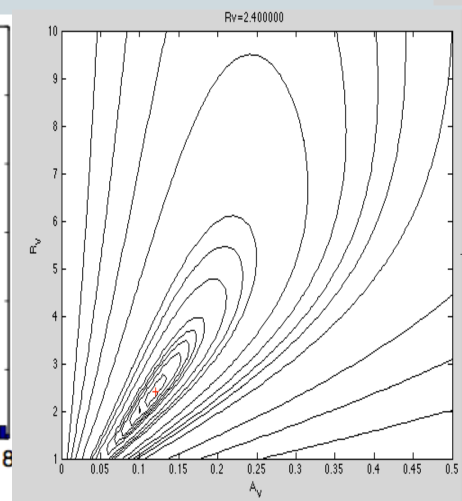


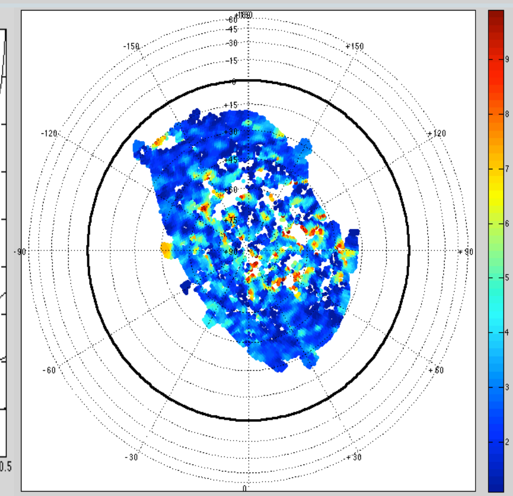


Av分布









Rv的实测分布：

信息与计算平台建设

宗旨

- 服务于天文台的需求
- 为理学院信息与计算科学专业寻求另一特色突破口
- 为三峡大学学生带来更多成长机会

硬件资源

三峡大学理学院数学系集群（60多万）：

- 头节点（IBM System X3650 M2服务器）
- 十个计算节点（Hp z600 workstation）：Intel Xeon E5520四核处理器，2.27GHz主频，1MB二级缓存，8MB三级缓存，8G内存，NVIDIA Quadro FX 3800专业显卡一块

刚为学生配置了一台GPU服务器：

（i7 + 16GB内存 + 2TB硬盘 + 1kw的电源 + 4 PCI-E + GTX480）

资源整合

三峡大学理学院机房：

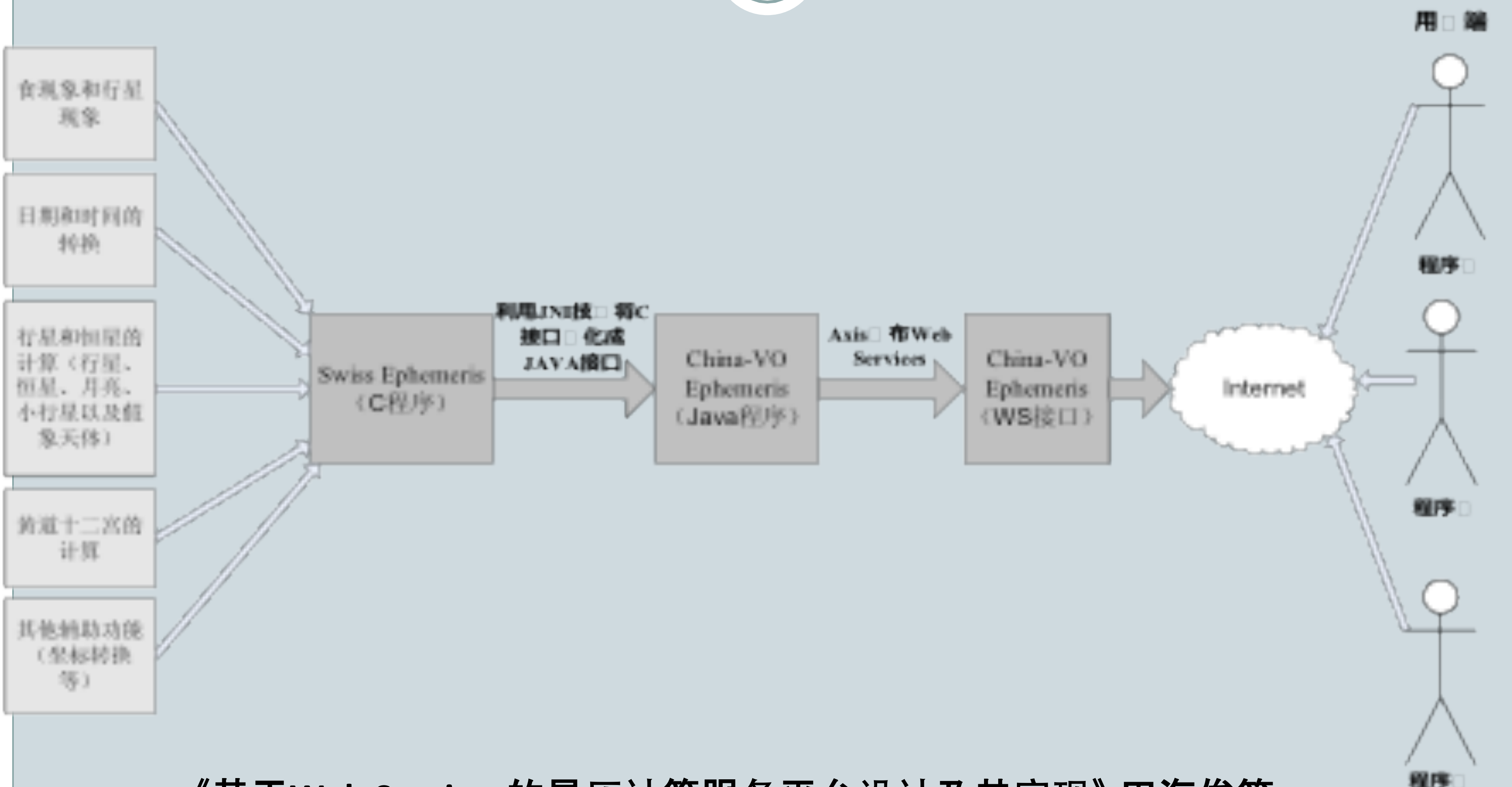
- 新建机房（58机位，信息与计算专业专用）
- 数学建模专业机房（48机位）
- 80多机位的机房
- 20多个机位的机房

平台建设规划：

- 四五十平米的机房整体改造（根据地）：部分用于半开放实验室、部分用于研究生教师办公、小机房。
- 充分利用共享计算资源（学校、国台）
- 学生联合培养模式（大四研一两年专业技能训练+研二研三国台跟踪项目训练）
- 新增一个信息技术联合实验室（目标）

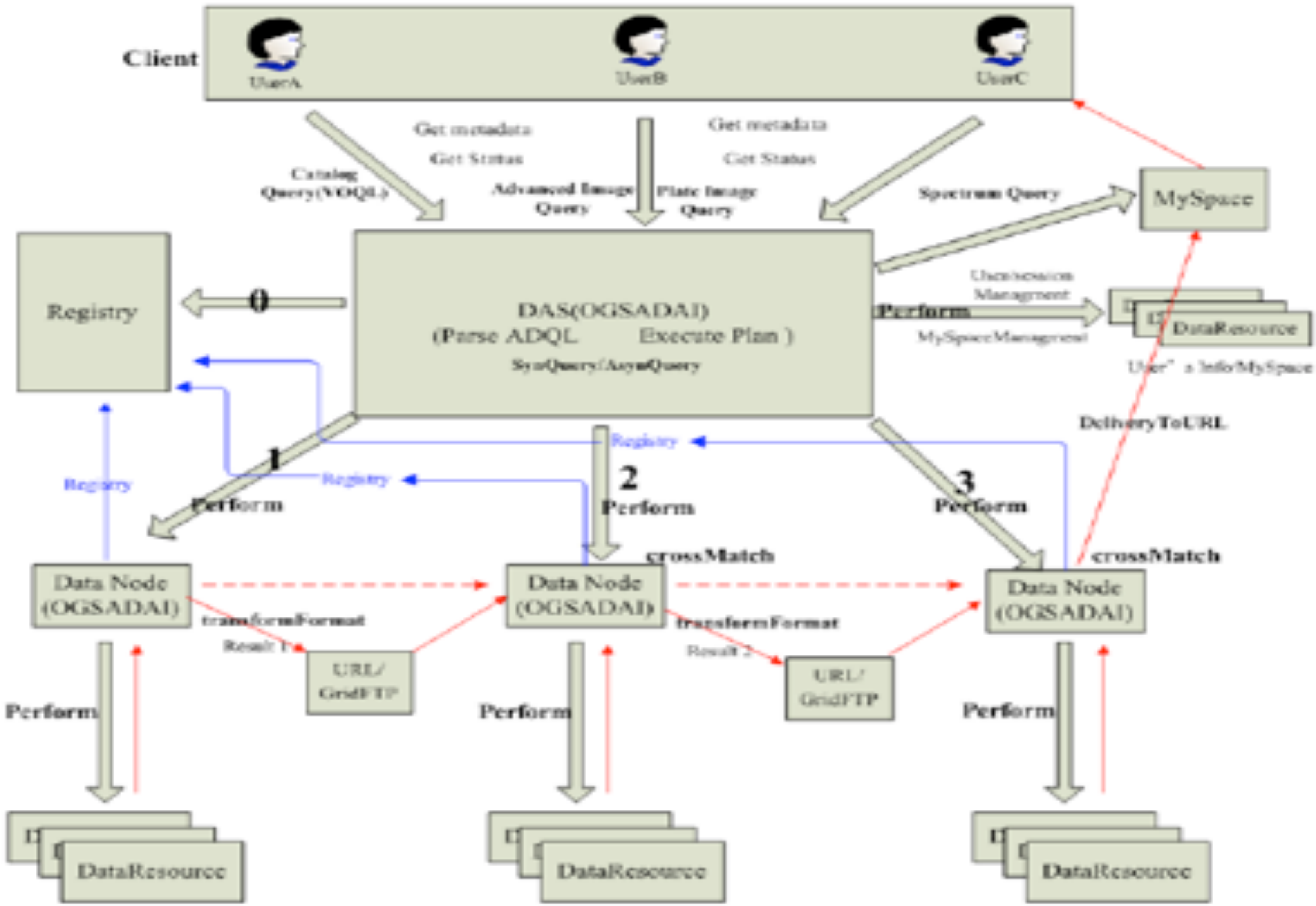
合作：信息与计算机学院、智能与图像研究所、三峡大学软件技术公司

计算服务平台



《基于Web Services的星历计算服务平台设计及其实现》田海俊等

数据节点平台



《虚拟天文台数据访问系统(VO-DAS)任务调度设计与实现》田海俊等

人才培养平台

SciDB

Microsoft® Research
WorldWide Telescope

本科、硕士毕业设计