

# Status Updates on STAR Computing at KISTI



November 16, 2012

Seo-Young Noh  
Haengjin Jang



{rsyoung, hjjang}@kisti.re.kr

# Contents

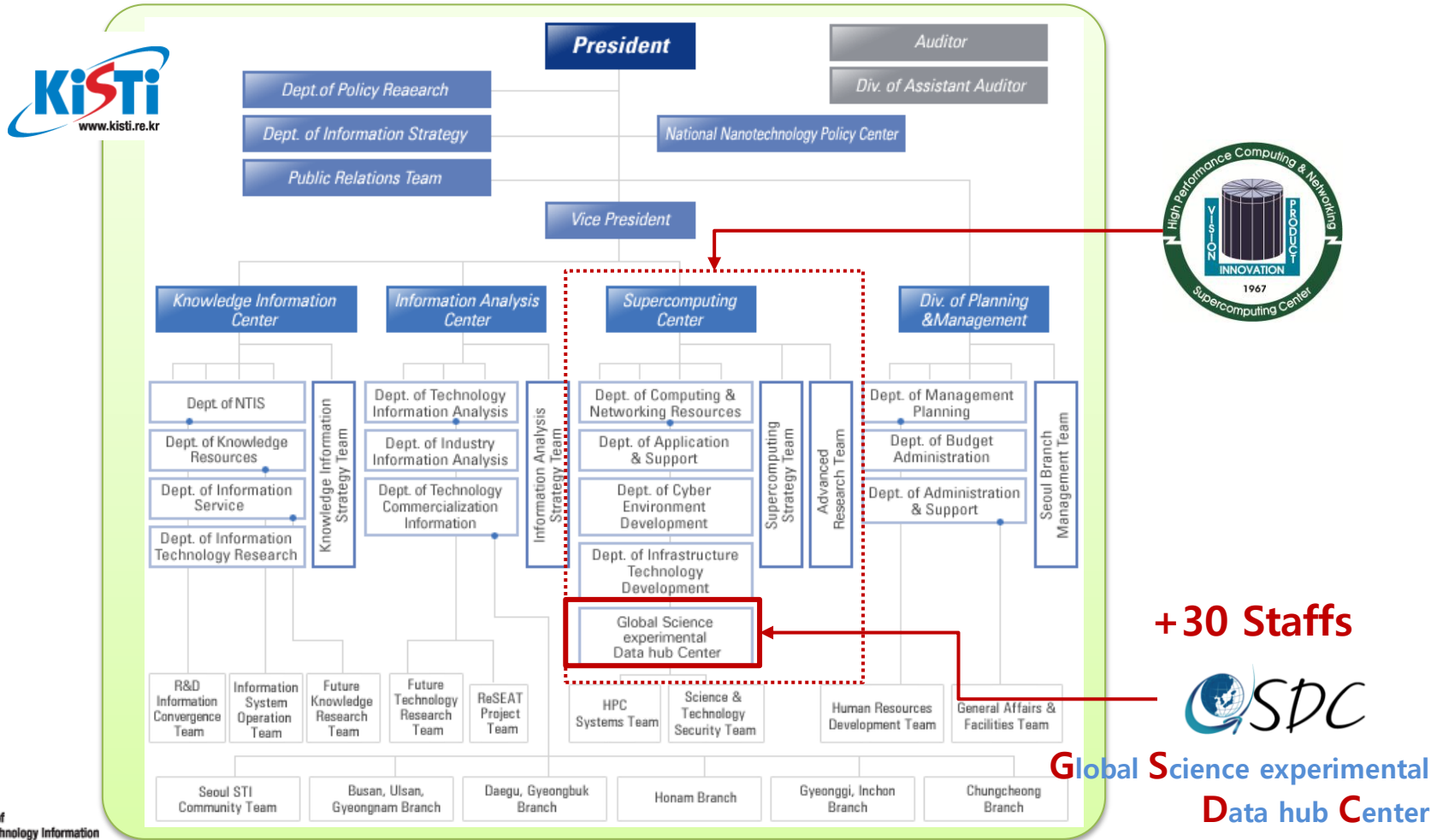
- 1. Introduction to GSDC**
- 2. Computing Resources**
- 3. STAR Computing**
- 4. MoU Summary & Conclusions**

---

# Introduction to GSDC

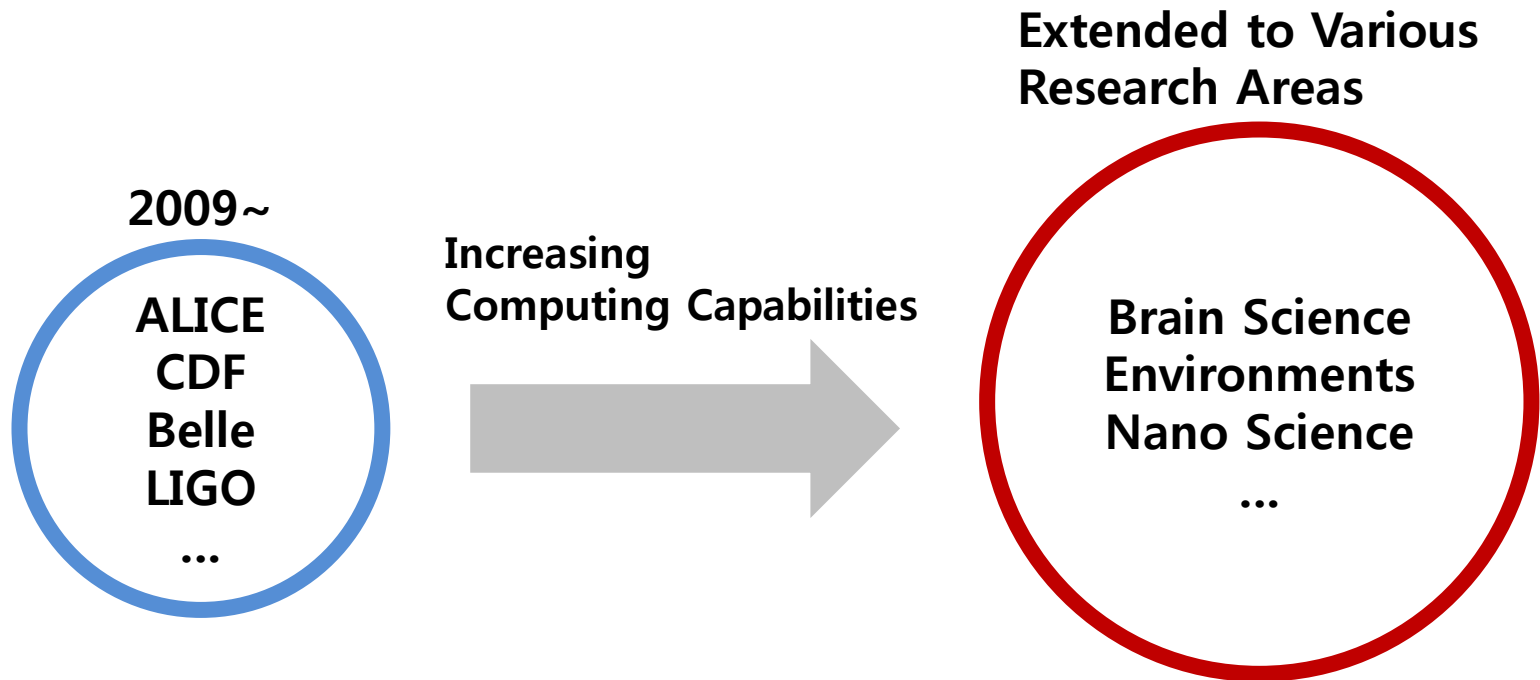
# Organization

**GSDC is under Supercomputing Center,  
focusing on Data Intensive Computing**



# Mission

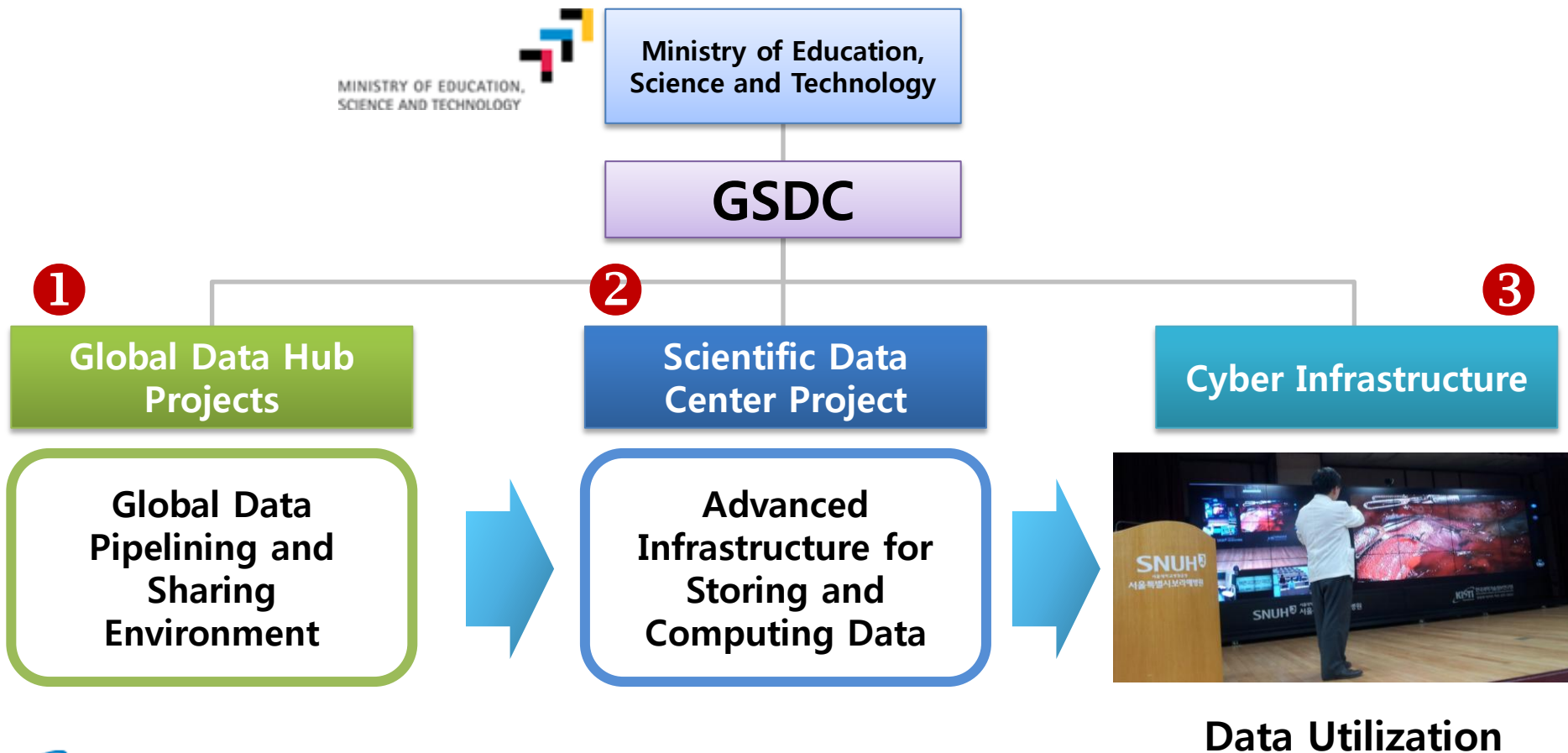
**Supporting Data Intensive Researches through Advanced Data  
Computing Environment sharing Global Experimental Data**



# Major Projects

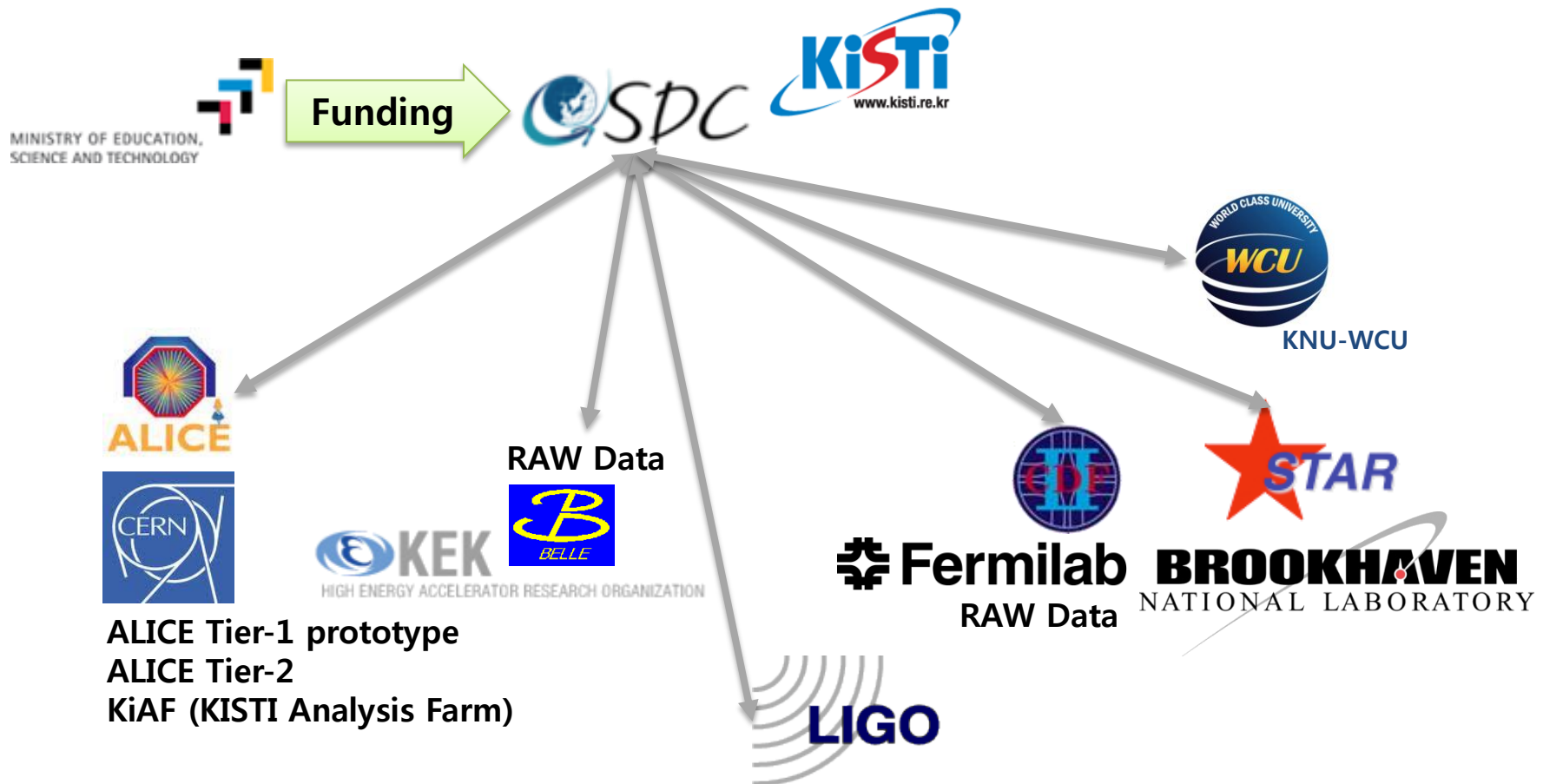
3 major projects seamlessly connected each other.

Data is everything! Data is being stored, shared, and utilized.



# Major Experiments

Supporting data intensive computing research areas with tight collaboration with research communities



---

# Computing Resources



# Resources – by 2011

**+2,000 Computing Cores, +2 PB Storage**

**Doubling computing resources every year!**

## ■ Server

Year	Spec	Mem.	Node	Core
2008~ 2009	Dell, Intel Xeon E5405x2(2.0GHz 4Core)	16GB	6	48
	HP, Intel Xeon E5420x2(2.5GHz 4Core)	16GB	16	128
	IBM, Intel Xeon E5450x2(3.0GHz 4Core)	16GB	38	304
2010	IBM, Intel Xeon X5650x2(2.66GHz 6Core),	24GB	36	432
2011	Intel Xeon X5650 x2(2.66GHz 6Core)	24GB	75	900
	Intel Xeon X5640 x2(2.4GHz 6Core)	24GB	35	420
<b>Total</b>			<b>206</b>	<b>2,232</b>

## ■ Storage

Year	Model	Type	Usable Size	Physical Size
2008 ~ 2009	NetApp FAS2050(SAN only, RAID6)	SAN	30 TB	48 TB
	NetApp FAS6080(SAN & NAS, RAID6)	SAN & NAS	200 TB	334 TB
2010	Hitachi USP-V (SAN & NAS, RAID6)	SAN & NAS	600 TB	960 TB
2011	EMC CX4-960	SAN	1,240 TB	1,800 TB
<b>Total</b>			<b>2,070 TB</b>	<b>3,142 TB</b>

# Resources - New Server & Storage in 2012

New +1,500 cores, +1.2 PB storage, 700 TB tape systems with high specification were introduced.

## ■ Server

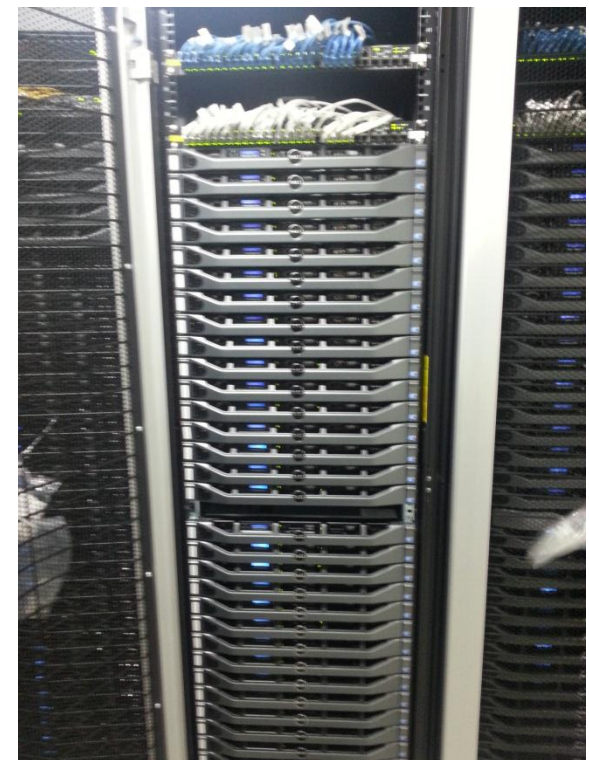
Year	Spec	Mem.	Node	Core
2012	Intel Sandy Bridge x2 (2.6GHz 8Core)	<b>96GB</b>	<b>94</b>	<b>1500</b>

## ■ Storage

Year	Model	Type	Usable Size	Physical Size
2012	EMC NL108	NAS	<b>1,200 TB</b>	<b>1,920 TB</b>

## ■ Tape

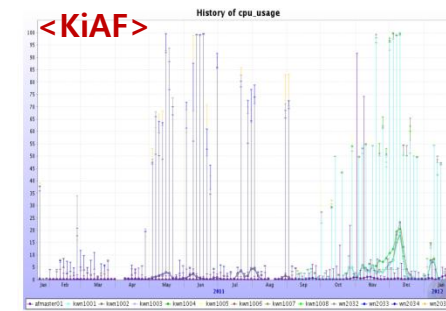
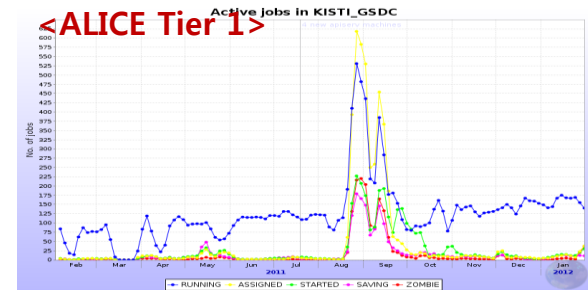
Year	Model	Type	Usable Size
2012	IBM	Tape	<b>700TB</b>



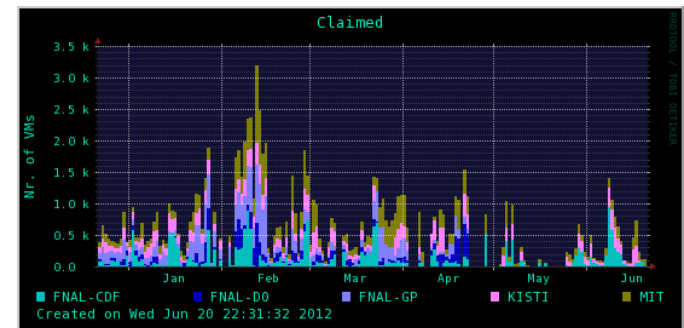
# Resource Allocation

Resource allocations and support ranges are defined through consultation with user communities.

Experiments		Computing Cores	Storage (TB)
ALICE	Tier 1	900	1000
	KiAF	96	
CDF	CDF Farm	432	180
STAR	STAR Farm	1024	50
KNU	HCP Farm	(200)	(50)
Belle	Belle Farm	126	170
LIGO	LIGO Farm	120	100
Neuro	Gbrain Farm	120	52
Earth Sci.	Storage Only	-	50
Etc.	DMRC, Nano,	-	310 TB



<CDF Farm>



---

# STAR Computing

# Background

## History

- ➔ Many discussions on KISTI's contribution to STAR collaboration since 2007
- ➔ Vision to Asia regional center for STAR computing, but not fruitful results

### Plan of STAR Collaboration at KISTI

#### • KISTI Resources Plans for STAR

Items	2008	2009	2010	2011
KISTI's SUN Cluster System	4th Supercom 1st phase AMD Opteron 2GHz (Barcelona)	4th Supercom 1st phase system AMD Opteron 2GHz (Barcelona)	4th Supercom 2nd phase system Intel Xeon 3.3GHz+ (Gainestown)	
KISTI Resources Plans for STAR				
minimum #of CPUs	32-150	150-250	about 500	about 700
Storage	Disk(TB)	20	100	200
	Tape(TB)	50	200	500
Network	GLORIAD	~10G	~20G	



**Not much progressed!**

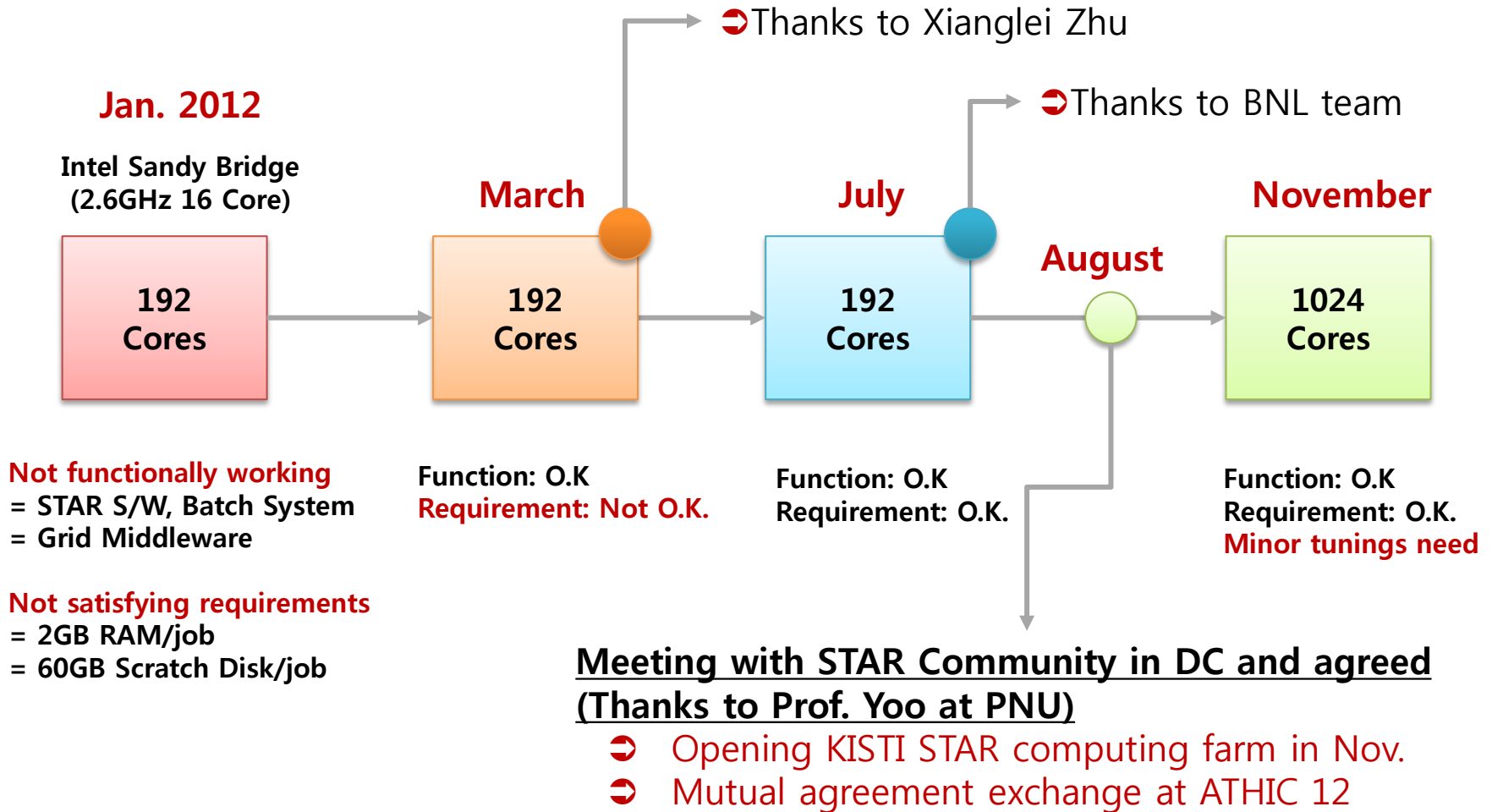
#### • KISTI Manpower for STAR

Network Specialist	Dr. Dong Kyun KIM, Computer Scientist
System Engineer	Mr. Jun WOO, Computer Engineering (Master Degree)
Storage Manager	Mr. Kelvin CHOI, Computer Science (Master Degree)
Codes Optimization	
Physics Research	Dr. Hyun Woo KIM, Physicist (Particle Physics)

\*Courtesy from Prof. Yoo's final report in 2009

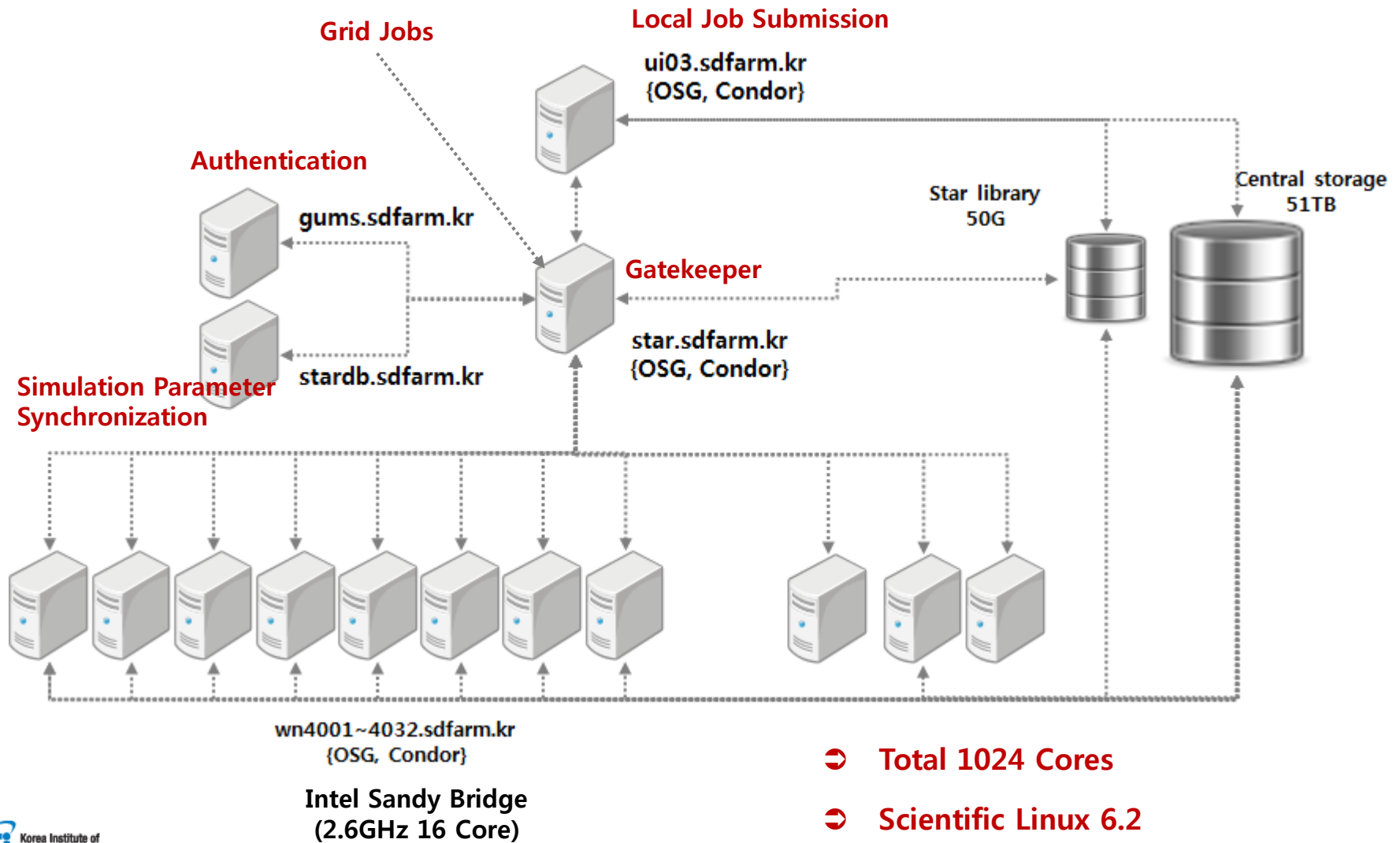
**In January 2012, set up a master plan for STAR computing farm at KISTI which keeps compromised resources at our best.**

# Transitions from Old to New



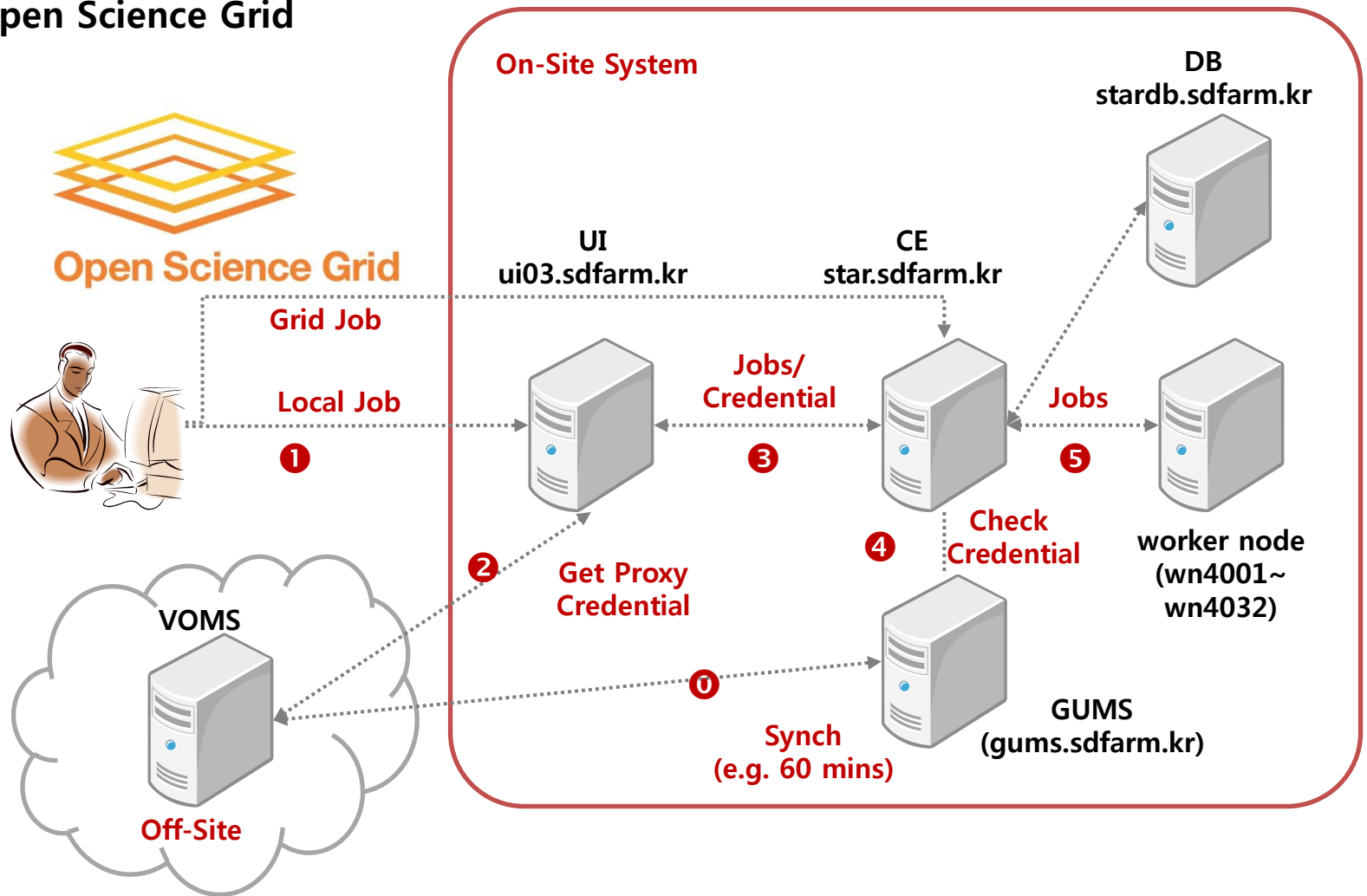
# STAR Computing Farm @ KISTI

## ■ System Configuration



# STAR Computing Farm @ KISTI

## Open Science Grid





# Local Computing Test

## ■ Computing Pool

```
root@star:~  
File Edit View Search Terminal Help  
slot23@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:41  
slot24@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:34  
slot25@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:35  
slot26@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:36  
slot27@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:37  
slot28@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:38  
slot29@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:39  
slot2@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:36  
slot30@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:40  
slot31@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:41  
slot32@wn4094.sdfa LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:34  
slot3@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:37  
slot4@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:38  
slot5@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:39  
slot6@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:40  
slot7@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:41  
slot8@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:07  
slot9@wn4094.sdfar LINUX X86_64 Unclaimed Idle 0.000 3020 0+23:30:35  
Total Owner Claimed Unclaimed Matched Preempting Backfill  
X86_64/LINUX 1024 0 0 1024 0 0 0  
Total 1024 0 0 1024 0 0 0  
[root@star ~]#
```

⇒ Current: 1024 Cores

# Local Computing Test

## Job Submission Test

```
[star@ui03 00.test]$ condor_status
```

Name	OpSys	Arch	State	Activity	LoadAv	Mem	ActvtyTime
slot10@wn4001.sdfa	LINUX	X86_64	Unclaimed	Idle	0.000	3020	4+02:56:48
slot11@wn4001.sdfa	LINUX	X86_64	Unclaimed	Idle	0.000	3020	4+02:56:49
slot12@wn4001.sdfa	LINUX	X86_64	Unclaimed	Idle	0.000	3020	4+02:56:50
slot13@wn4001.sdfa	LINUX	X86_64	Unclaimed	Idle	0.000	3020	4+02:56:51
slot14@wn4001.sdfa	LINUX	X86_64	Unclaimed	Idle	0.000	3020	4+02:56:52

----- lines are omitted -----

slot5@wn4094.sdfar	LINUX	X86_64	Unclaimed	Idle	0.000	3020	3+07:06:30
slot6@wn4094.sdfar	LINUX	X86_64	Unclaimed	Idle	0.000	3020	3+07:06:31
slot7@wn4094.sdfar	LINUX	X86_64	Unclaimed	Idle	0.000	3020	3+07:06:32
slot8@wn4094.sdfar	LINUX	X86_64	Unclaimed	Idle	0.000	3020	3+07:06:25
slot9@wn4094.sdfar	LINUX	X86_64	Unclaimed	Idle	0.000	3020	3+07:06:26

```
Machines Owner Claimed
X86_64/LINUX 1024 0
Total 1024 0
[star@ui03 00.test]$
```

① ui03 machine

Local users, not grid users,  
can log on to ui03.sdfarm.kr  
to submit jobs!

② 1024 c

```
[star@ui03 00.test]$
[star@ui03 00.test]$ condor submit multiple.ids
Submitting
job(s).....
.....
.....
----- lines are omitted -----
.....
.....
.....
5000 job(s) submitted to cluster 8909.
[star@ui03 00.test]$
```

③ job submission

④ 5000 jobs submitted

# Local Computing Test

## ■ Queue Filling Test

```
[star@ui03 00.test]$ condor_q
```

### ① Queue status checking

```
-- Submitter: ui03.sdfarm.kr : <134.75.123.38:10205> : ui03.sdfarm.kr
ID      OWNER      SUBMITTED      RUN_TIME ST PRI  SIZE CMD
8909.0  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.1  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.2  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.3  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.4  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.5  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.6  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.7  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.8  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.9  star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
8909.10 star       11/11 08:24    0+00:00:26 R  0   0.0 count.sh 1 10
```

1024 slots are filled by 5000 jobs and working properly returning correct results.

```
----- lines are omitted -----
8909.4993 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
8909.4994 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
8909.4995 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
8909.4996 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
8909.4997 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
8909.4998 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
8909.4999 star       11/11 08:24    0+00:00:00 I  0   0.0 count.sh 1 10
```

```
4913 jobs; 1 completed, 0 removed, 3888 idle, 1024 running, 0 held, 0 suspended
[star@ui03 00.test]$
[star@ui03 00.test]$
```

### ② All slots are filled!

# Grid Middleware Test

## ■ Authentication Test

```
[star@ui03 ~]$ globusrun -a -r star.sdfarm.kr
```

```
GRAM Authentication test successful
```

```
[star@ui03 ~]$
```

⇒ Working with GUMS

## ■ Managed Fork Test

⇒ Asking hostname through fork

```
[star@ui03 ~]$ globus-job-run star.sdfarm.kr:2119/jobmanager-fork /bin/hostname  
star.sdfarm.kr
```

⇒ Correctly return the hostname

```
[star@ui03 ~]$
```

## ■ Condor Job Manager Test

⇒ Asking hostname through Condor

```
[star@ui03 ~]$ globus-job-run star.sdfarm.kr/jobmanager-condor /bin/hostname  
wn3089.sdfarm.kr
```

⇒ A job assigned to wn3089  
returns its hostname

```
[star@ui03 ~]$
```

# Grid Middleware Test

## ■ Submitting multiple grid jobs

↻ ui03.sdfarm.kr

↻ 10 grid jobs submitted to star.sdfarm.kr

```
[star@ui03 condor_based]$ ./batch-job-submit.sh
1 : https://star.sdfarm.kr:40199/16217745584214884961/10812896844644919431/
2 : https://star.sdfarm.kr:40199/16217745582034350181/10812896844644919431/
3 : https://star.sdfarm.kr:40199/16217745581520385896/10812896844644919431/
4 : https://star.sdfarm.kr:40199/16217745582035273836/10812896844644919431/
5 : https://star.sdfarm.kr:40199/16217745582538169456/10812896844644919431/
6 : https://star.sdfarm.kr:40199/16217745581160514931/10812896844644919431/
7 : https://star.sdfarm.kr:40199/16217745584205947511/10812896844644919431/
8 : https://star.sdfarm.kr:40199/16217745582774283386/10812896844644919431/
9 : https://star.sdfarm.kr:40199/16217745582998781566/10812896844644919431/
10 : https://star.sdfarm.kr:40199/16217745580627009666/10812896844644919431/
[star@ui03 condor_based]$
```

↻ root@star.sdfarm.kr

```
[root@star ~]# condor_q
```

```
-- Submitter: star.sdfarm.kr : <134.75.123.61:9619> : star.sdfarm.kr
ID      OWNER   SUBMITTED  RUN_TIME ST PRI  SIZE  CMD
147.0   star    6/4 13:55   0+00:00:07 R  0    0.0  data
148.0   star    6/4 13:55   0+00:00:03 R  0    0.0  data
149.0   star    6/4 13:55   0+00:00:03 R  0    0.0  data
```

```
3 jobs; 0 completed, 0 removed, 0 idle, 3 running, 0 held, 0 suspended
```

```
[root@star ~]#
```

↻ Filling the queue gradually

# Grid Middleware Test

## ■ Results of multiple grid jobs

➡ ui03.sdfarm.kr

```
[star@ui03 condor_based]$ ./getstatus.sh
```

➡ Asking the status of 10 grid jobs submitted to star.sdfarm.kr

```
1 https://star.sdfarm.kr:40199/16217745584214884961/10812896844644919431/ DONE
2 https://star.sdfarm.kr:40199/16217745582034350181/10812896844644919431/ DONE
3 https://star.sdfarm.kr:40199/16217745581520385896/10812896844644919431/ DONE
4 https://star.sdfarm.kr:40199/16217745582035273836/10812896844644919431/ DONE
5 https://star.sdfarm.kr:40199/16217745582538169456/10812896844644919431/ DONE
6 https://star.sdfarm.kr:40199/16217745581160514931/10812896844644919431/ DONE
7 https://star.sdfarm.kr:40199/16217745584205947511/10812896844644919431/ DONE
8 https://star.sdfarm.kr:40199/16217745582774283386/10812896844644919431/ DONE
9 https://star.sdfarm.kr:40199/16217745582998781566/10812896844644919431/ DONE
10 https://star.sdfarm.kr:40199/16217745580627009666/10812896844644919431/ DONE
[star@ui03 condor_based]$
```

➡ Successfully done

# Grid Middleware Test

## ■ Copying a file from CE using GridFTP

➔ Copy star.sdfarm.kr/proc/cpuinfo to ui03.sdfarm.kr/tmp/cpuinfo

➔ ui03.sdfarm.kr

```
[star@ui03 ~]$ globus-url-copy gsiftp://star.sdfarm.kr/proc/cpuinfo \
file:///tmp/cpuinfo
[star@ui03 ~]$
```

➔ cat /tmp/cpuinfo

```
[star@ui03 ~]$ cat /tmp/cpuinfo
processor       : 0
vendor_id     : GenuineIntel
cpu family    : 6
model        : 23
-----  종  략  -----

processor      : 3
vendor_id    : GenuineIntel
cpu family   : 6
model       : 23
model name  : Intel(R) Xeon(R) CPU           X3320 @ 2.50GHz
stepping    : 7
cpu MHz     : 2000.000
cache size  : 3072 KB
physical id : 0
siblings    : 4
core id     : 3
cpu cores   : 4
apicid      : 3
fpu         : yes
fpu_exception : yes
cpuid level : 10
wp          : yes
flags       : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat
pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm syscall nx lm constant_tsc
pni monitor ds_cpl vmx smx est tm2 ssse3 cx16 xtpr sse4_1 lahf_lm
bogomips    : 5000.02
clflush size : 64
cache_alignment : 64
address sizes : 36 bits physical, 48 bits virtual
power management:
```

proc/cpuinfo file at star.sdfarm.kr has been successfully copied to ui03.sdfarm.kr/tmp/cpuinfo

# Grid Middleware Test

## ■ Copying a file to CE using GridFTP

➔ ui03.sdfarm.kr

```
[star@ui03 ~]$ globus-url-copy file:///proc/cpuinfo \
gsiftp://star.sdfarm.kr/tmp/cpuinfo-deleteme
```

➔ star.sdfarm.kr

```
[star@ui03 ~]$
```

```
[star@star ~]$ ls -al /tmp/cpuinfo-deleteme
-rw-r--r-- 1 cdf osgcdf 5164 Jun 11 10:34 /tmp/cpuinfo-deleteme
[star@star ~]$
[star@star ~]$ cat /tmp/cpuinfo-deleteme
processor      : 0
vendor_id     : GenuineIntel
cpu family    : 6
model         : 23
model name    : Intel(R) Xeon(R) CPU           X5450  @ 3.00GHz
stepping      : 6
cpu MHz       : 1992.000
cache size    : 6144 KB

-----  중  략  -----

bogomips      : 5985.00
clflush size  : 64
cache_alignm  : 64
address sizes : 38 bits physical, 48 bits virtual
power managem:
```

proc/cpuinfo file at ui03.sdfarm.kr has been successfully copied to the computing element



# SUMS Test

## ■ SUMS: STAR Unified Meta Scheduler

```
[ui03] ~/sums_test/> cat test.xml ① Test file  
<?xml version="1.0" encoding="utf-8" ?>  
<job name="Test" mail="false" nProcesses="3">
```

```
<command>  
echo $USER  
/bin/date  
sleep 60  
</command>
```

➔ Getting \$USER and Date

```
[ui03] ~/sums_test/> star-submit test.xml ② Submitting test jobs  
STAR Unified Meta Scheduler 1.10.7  
<stdout URL="file" >  
<stderr URL="file" >  
Your Log file can be found at: ../star.log  
</job>  
Reading request description file : test.xml  
Analyzing XML...XML OK  
Testing queue (kisti_condor_Queue@STAR - Local Access Method) ...Done (passed)  
Writing files for process 6BC7B31FDE60DB7F2C097194058754DA_2..done.  
Writing files for process 6BC7B31FDE60DB7F2C097194058754DA_1..done.
```

➔ 3 jobs to kisti\_condor\_Queue

```
[ui03] ~/sums_test/> condor_q ③ Checking submitted jobs  
  
-- Submitter: ui03.sdfarm.kr : <134.75.123.38:14994> : ui03.sdfarm.kr  
ID      OWNER   SUBMITTED   RUN_TIME ST PRI  SIZE  CMD  
186.0   star    6/10 18:11   0+00:00:27 R  10   0.0  sched6BC7B31FDE60D  
187.0   star    6/10 18:11   0+00:00:27 R  10   0.0  sched6BC7B31FDE60D  
188.0   star    6/10 18:11   0+00:00:27 R  10   0.0  sched6BC7B31FDE60D  
  
3 jobs; 0 completed, 0 removed, 0 idle, 3 running, 0 held, 0 suspended  
[ui03] ~/sums_test/>
```

➔ 3 jobs running

# SUMS Test

```
[ui03] ~/sums_test/> ls
6BC7B31FDE60DB7F2C097194058754DA.session.xml
sched6BC7B31FDE60DB7F2C097194058754DA_2.list
_condor_stderr
sched6BC7B31FDE60DB7F2C097194058754DA.report
_condor_stdout
shed6BC7B31FDE60DB7F2C097194058754DA_0.err.out
```

① Log file

② Result file

```
[ui03] ~/sums_test/> ls
6BC7B31FDE60DB7F2C097194058754DA.session.xml
sched6BC7B31FDE60DB7F2C097194058754DA_2.list
_condor_stderr
sched6BC7B31FDE60DB7F2C097194058754DA.report
_condor_stdout
```

----- 중 략 -----

```
sched6BC7B31FDE60DB7F2C097194058754DA_2.csh*
```

```
[ui03] ~/sums_test/>
```

```
[ui03] ~/sums_test/> cat shed6BC7B31FDE60DB7F2C097194058754DA_0.out
```

```
We are starting on node: wn3086.sdfarm.kr
```

```
HOME is not defined
```

```
Using getent method
```

```
HOME is now /star/u/star
```

```
chkread :: File or directory /afs/rhic.bnl.gov/ is not readable
```

```
STAR Unified Meta Scheduler 1.10.7 we are starting in $SCRATCH :
/star/scratch/star/6BC7B31FDE60DB7F2C097194058754DA_0
```

```
total 0
```

```
star
```

```
Sun Jun 10 18:11:23 KST 2012
```

```
[ui03] ~/sums_test/>
```

⇒ Successfully print  
\$USER & Date

⇒ A job has been  
successfully done

```
--
.94058754DA_2.csh*
shed6BC7B31FDE60DB7F2C097194058754DA_0
.94058754DA_0_2.condor
.94058754DA_0.csh*
.94058754DA_0.condor.log
.94058754DA_0.list
shed6BC7B31FDE60DB7F2C097194058754DA_0.condor.log
11:06 Job submitted from host: <134.75.123.38:14994>
11:22 Job executing on host: <134.75.124.87:9678>
11:31 Image size of job updated: 15500
e of job (KB)
12:23 Job terminated.
(return value 0)
$ 0 00:00:00 - Run Remote Usage
$ 0 00:00:00 - Run Local Usage
$ 0 00:00:00 - Total Remote Usage
$ 0 00:00:00 - Total Local Usage
By Job
ived By Job
t By Job
eceived By Job
: Usage Request
: : 1
Disk (KB) : 12 12
Memory (MB) : 16
...
[ui03] ~/sums_test/>
```

---

# MoU Summary & Conclusions

# MoU Summary

## ■ Computing Resources

Resource	Y. 2012	Y. 2013	Y. 2014*
CPU	1024	1024	1024
Disk Storage (TB)	50	50	50

\* Resource allocation will be discussed with STAR community

## ■ Primary Contact Points in Communication

Service	Name	Organization
GSDC STAR System Management	Seo-Young Noh	GSDC
STAR Software	Xianglei Zhu	STAR Community
STAR Computing General	Jerome Lauret	BNL
STAR Collaboration General	In-Kwon Yoo	Pusan National University

# Conclusions

- ➔ **Big achievement this year, been tried since 2007!**
- ➔ **Many thanks to Xiaglei Zhu, BNL team, and Prof. Yoo for STAR S/W installation, testing, and communication**
- ➔ **Playing an important role in STAR computing as a Asia STAR computing center.**
- ➔ **Official MoU will lay the groundwork for two parties' tight collaborations**

**Thank You!**