Grid technologies for large-scale projects

N. S. Astakhov, A. S. Baginyan, S. D. Belov, A. G. Dolbilov, A. O. Golunov, I. N. Gorbunov, N. I. Gromova, I. A. Kashunin, V. V. Korenkov, V. V. Mitsyn, V. V. Palichik, S. V. Shmatov, T. A. Strizh, E. A. Tikhonenko, V. V. Trofimov, <u>N. N. Voitishin</u>, V. E. Zhiltsov

Laboratory of Information Technologies

JINR

Alushta, Crimea 2016

Grid technologies - a way to success

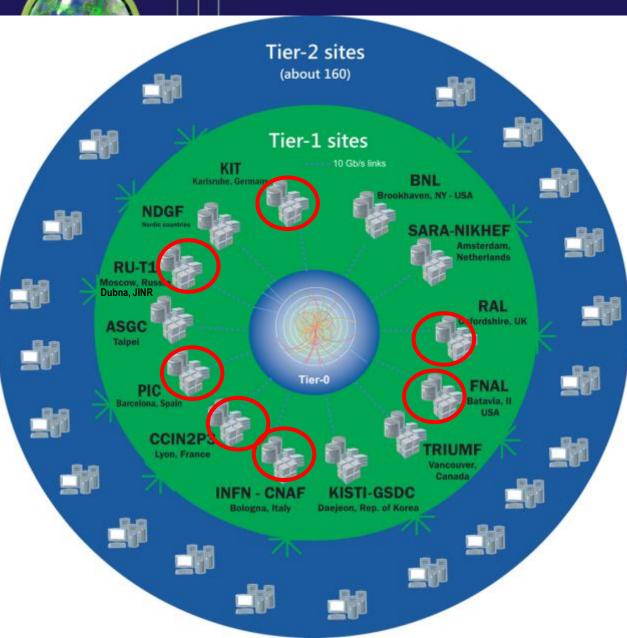
On a festivity dedicated to receiving the Nobel Prize for discovery of Higgs boson, CERN Director professor Rolf Dieter Heuer directly called the **gridtechnologies one of three pillars of success** (alongside with the LHC accelerator and physical installations).



Without implementation of the grid-infrastructure on LHC it would be impossible to process and store enormous data coming from the collider and therefore to make discoveries.

Nowadays, every large-scale project will fail without using a distributed infrastructure for data processing.

LHC Computing Model



Tier-0 (CERN): • Data recording • Initial data reconstruction • Data distribution

Tier-1 (11 \rightarrow 14 centres):

- Permanent storage
- Re-processing

Analysis

Simulation

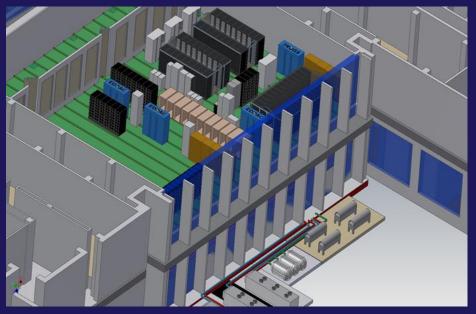
Tier-2 (>200 centres):

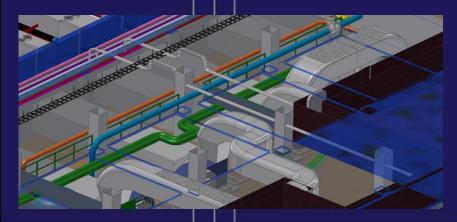
- Simulation
- End-user analysis

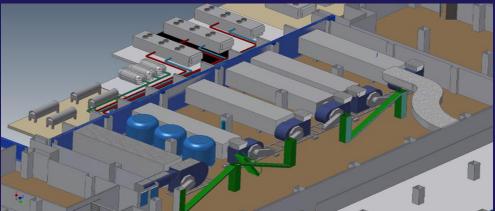
Creation of CMS Tier1 in JINR

Engineering infrastructure (a system of uninterrupted power supply, climate - control);

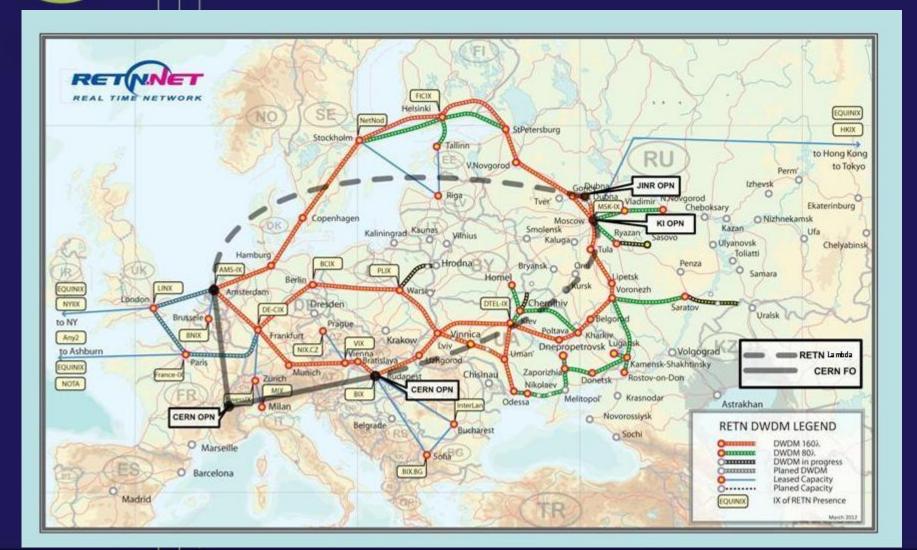
- High-speed reliable network infrastructure with a dedicated reserved data link to CERN (LHCOPN);
- Computing system and storage system on the basis of disk arrays and tape libraries of high capacity;
 100% reliability and availability.

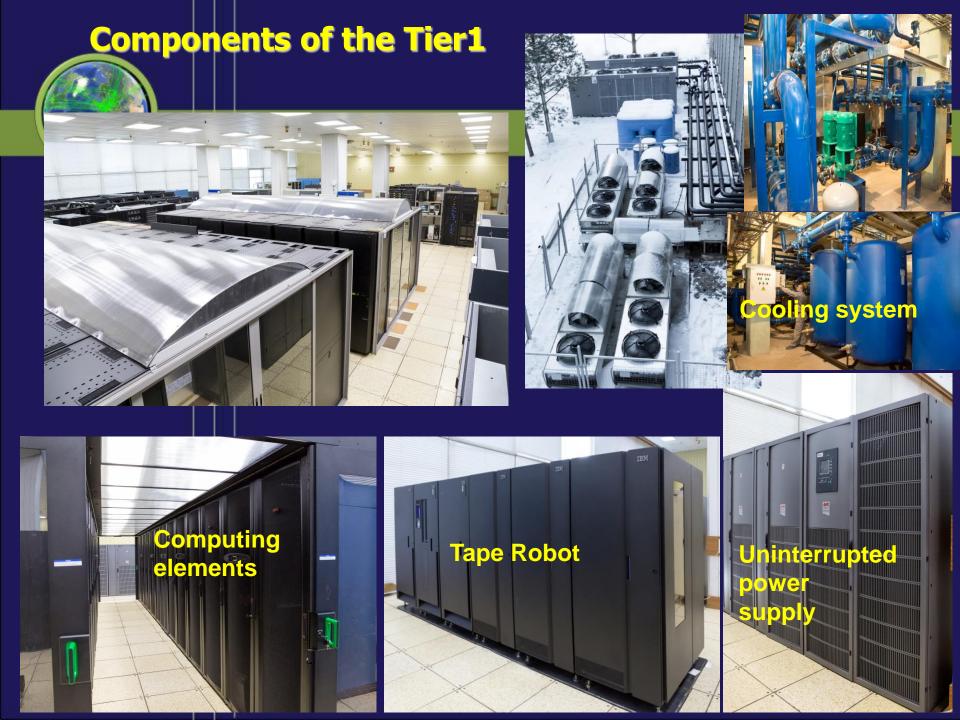






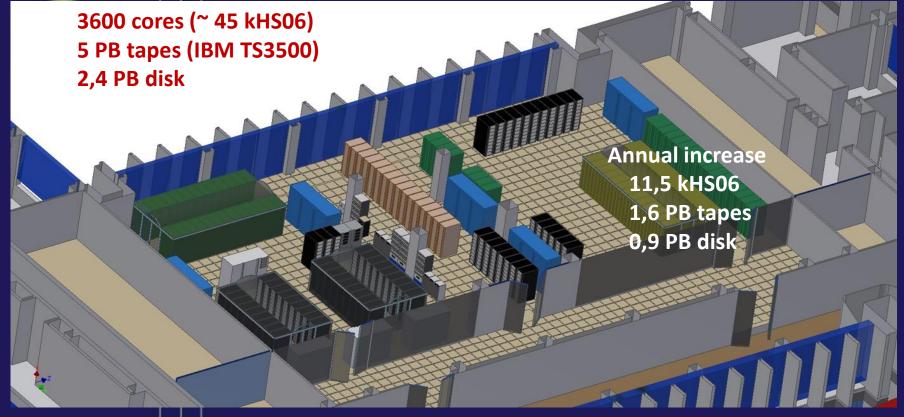
JINR Tier-1 Connectivity Scheme







Current configuration and plans

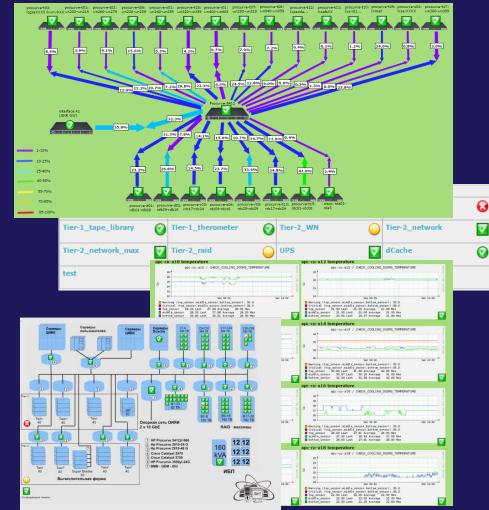


JINR Tier-1 monitoring system

JINR Tier-1 monitoring system provides real-time information about:

- * work nodes;
- * disk servers;
- network equipment;
- uninterruptible power supply elements;
- cooling system.

It also can be used for creating network maps and network equipment load maps, for drawing state tables and different plots.

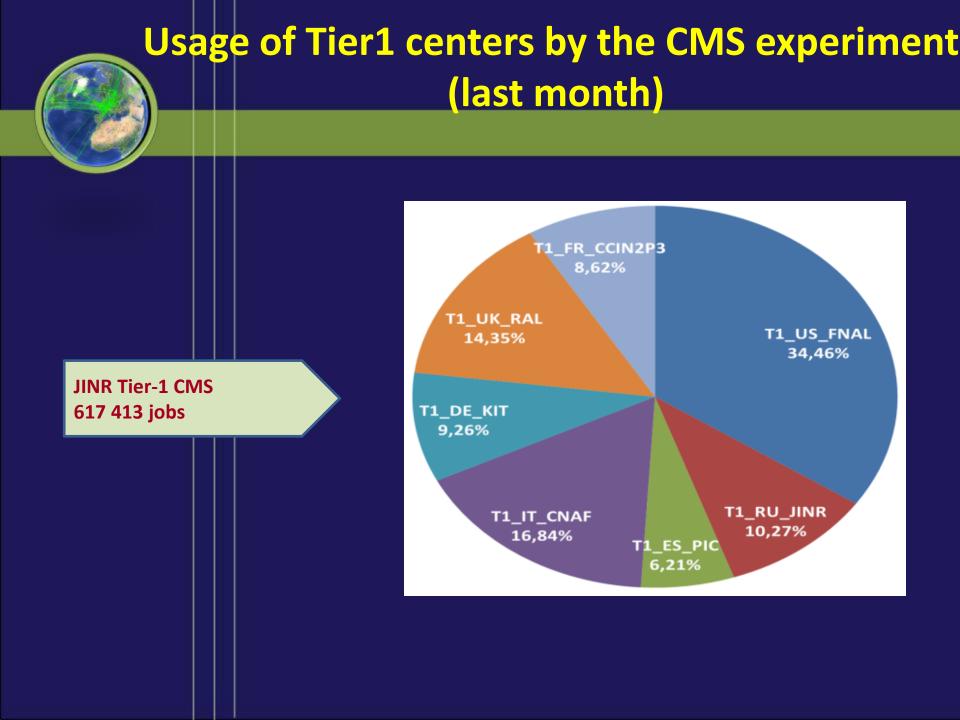


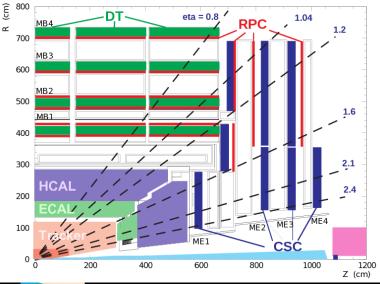
JINR Tier's Dashboard



	KNODES	WORK NODE	WORK NODES		JINR 12 JOBS	
wn040-044	KNODES	work NODE: wn280-284 7 7 7	WORK NODES	UPS	lxbsrv01.jinr.ru / O4EOK_308S	
wn045-049		wn285-289	vmi005-009 v v v	UPS-01 UPS-02 UPS-06 UPS-07	244	
		wn290-294 V 0 V	wni010-014		and the manufacture	
wn050-054	<u> </u>			c 100 c 100 c 100 c 100		
wn055-059	<u> </u>	wn295-299 💟 💟 💟	vni015-019 V V V V	2 T 29 T 30 T 33 ⁴² T 27 1 29 P 4 P 5 P 5 P 5 P 30	the formation of the	
wn100-104	22222	wn300-304 💟 💟 💟	<u>v</u> <u>v</u>		L'AND CAN THE CAR AND	
wn104-109	7 7 7 7 7	wn304-309 💟 💟 💟	🛛 🗹 🛛 RAIDS	UPS-03 UPS-04 UPS-08 UPS-11	0.4 k 06:00 08:00 10:00 12:00 14:00	
wn110-114	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	wn310-314 💟 💟 💟	🟹 🟹 rda01-04 🛛 🟹 🟹 🟹		The number os jobs in queue	
wn115-119	7 7 7 7 7	wn315-319 🔽 🟹 🟹	🟹 🔽 rda05-09 🛛 🟹 🟹 🟹	C100 C100 C100 C100	The number of current jobs	
wn120-124	77777	wn320-324 💟 💟 🔽	🛛 🔽 rda10-13 📝 🟹 🟹	7 35 T 35 37 T 31 31 T 34 T 34		
wn125-129	7 7 7 7 7	wn325-329 💟 💟 💟	🛛 🔽 rdb01-04 🛛 🟹 🟹 🟹		2015 OK	
wn200-204	V V V 0 V	wn330-334 👿 🗹 🗹	🟹 🟹 rdb05-09 🔽 🟹 🟹 🟹	PP – power packs T – themperature		
wn204-209	7 7 7 7 7 7	wn335-339 💟 💟 🟹	🛛 🟹 rdb10-14 🔍 🟹 🟹 🟹		2 NETWORK	
wn210-214	30000	wn340-344 💟 💟 💟	🗸 🟹 rdb15-19 🖉 🖉 🖉 🖉	103:kgsexxX t01:wn200- t02:wn260- t06:wn220- s01:wn240- t18:wn320- r01:w	curve procurve procup	
wn215-219	33333	wn345-349 🔍 🔍 🔍	👽 👽 rdb20-24 🔍 👽 👽 👽		1493 wn319 wn059 ks-r05 kpanda kgoqd10 02-bse-XXX wn	
wn220-224	00000	wn350-354 🔍 🔍 🟹	7 7 rdc01-04 7 7 7 7		~ ~ ~ ~ ~ ~ ~ ~ ~	
wn225-229	77777	wn355-359 👽 👽 🐽	7 7 rdc05-09 7 7 7 7 7			
wn230-234	33333	wn360-364 👽 👽 👽	7 7 rdc10-14 7 7 7 7 7	0.1% 15.8% 0.1% 0.1% 0.1%		
wn235-239	77777	wn365-369 👽 👽 👽	🗸 🗸 rdc15-19 🗸 🗸 🗸 🗸	0.4%	1% 0.1% 0.1% 0.0% 0.0% 0.0%	
wn240-244		wn370-374 V V	V V rdc20-24 V V V V V	LIN	0.0%	
wn245-249		wn375-379 7 7	🗸 🟹 rdb25-29 🕘 🗸 🗸 🗸			
wn250-254		wn380-384 V V V		18.4% 0.2% 9.0% 2.7%	0.3% 2.6% 3.4% 0.6% 0.6%	
wn255-259		wn385-389 7 7 7		25.4%	0.0% 0.0% 0.0%	
wn260-264		wn390-394 7 7 7	JINR GW	Proti	urvesal2	
wn265-269		wn395-399 7 7 7		0.8%	2.6%	
wn270-274		wn400-403 7 7 7			Procurve-5412 / CHECK_TRAFFIC_SNMP_PROCURVE_A1	
wn275-279	3 3 3 3 3 3	W19009403 V V	<u>21</u>	0.0% 0.1% 0.0% 2.1% 0.1% 0.0% 0.2% 0.0		
	MPERATURE			0.346 (0.046) 2.146 (0.146) 00.146 (0.246) 000		
		apc-cu-a17 🔽 2			g zeek II. A . A II.	
apc-cu-a01 apc-cu-a02			10-25%	7.1% 0.0% 12.3% 7.8% 5.0% 23.3% 1.0%	when the day of the state of the	
apc-cu-a03				1.975 A.4.976	- tother we with a to be achieved with	
apc-cu-a04				<u> </u>	06.00 00.00 10.00 12.00 14.00	
apc-cu-a05	23 apc-cu-a14	🔽 30 apo-cu-a22 🟹 2	70-85% procurve	procurve procurve procurve procurve procurve procurve cisco.rda01		
apc-cu-a06	25 apc-cu-a15			d02xdb09- d03xdb17- s04xdc09- r03xdc25- s13xdc17- s15xdc01- rdb16 rdb24 rdc16 rdc29 rdc24 rdc08	Input trafic	
apc-cu-a07	21 apc-cu-a16	27 apc-cu-a25 👽 2				

			UPS	JINR T1 USED CORES				
WORK NODES	WORK NODES	RAIDS	npc-gi-01	bfsrv.jinr-tl.ru / CHECK_USED_CORES				
wna000-004 🔽 🟹 🟹 🟹	wna130-134 🔽 🟹 🟹 🟹	rda000-004 🔽 🟹 🟹 🟹	1012					
wna005-009 🔽 🟹 🟹 🟹	wna135-139 🔽 🔽 🟹 🟹	rda005-009 🔽 🟹 🟹 🟹						
wna010-014 🔽 🟹 🟹 🟹	wna140-144 🔽 🗹 🟹 🟹	rda010-014 🛛 🗹 🟹 🟹 🟹	54Z - T 23	80 00 01:00 02:00 02:00 04:00 05:00 04:00 07:00 09:00 09:00				
wna015-019 🔽 🟹 🟹 🟹	wna145-149 🔽 🔽 🔽 🔽	rda015-019 🔽 🔽 🔽 🟹						
wna020-024 🔽 🟹 🟹 🟹	wna150-154 🔽 🟹 🟹 🟹	rda020-024 🔽 🟹 🟹 🟹	GZ <u>13</u> ▼ ups_loas	94.00% OK JINR T1 JOBS				
wna025-029 🔽 🟹 🟹 🟹	wna155-159 🔽 🔽 🔽 🔽	rda025-029 🔽 🔽 🔽 🔽	apo-gl-02	bfsrv.jinr-tl.ru / OHECK_2005				
wna030-034 🔽 🟹 🟹 🟹	lxvh000-004 🔽 🗹 🔽 🔀	rdt000-004 🔽 🔽 🔽 🔽	1002	****				
wna035-039 🔻 🟹 🟹 🟹 🗸	lxvh005-009	rdt005-007 📅 📅 📅	6100	101				
wna040-044 👽 👽 👽 👽	Hast Name	Host (Last state refresh: 2016-02-16 09-21/ lovR005 Sm-11.ev (lovR005)	64)	0.58				
wna045-049 🗸 🗸 🗸 🗸	State	DOWN (HARD - 1/10) (Host Check Timed Out)		00.00 02:00 82:00 03:00 04:00 85:00 06:00 07:00 68:00 09:00				
	Last Check	2016-02-16 06: 19:59		The number of mulicore jobs 233 The number of current jobs 1298 The number os jobs in queue 911				
wna050-054 🔽 🗹 🟹 🟹	Next Check Last State Charge	2016-02-16 06:25:38 2016-02-05 17:02:27		OV				
wna055-059 🔽 🟹 🟹 🟹	Summary State	50WN		onternettorat				
wna060-064 🔽 🟹 🟹 🟹	Summary Output	The Host is DOWN. There are 1 CRITICAL, 7 UNK	NOWN Services.	W-11 T1-SW-12 T1-SW-17 T1-SW-18				
wna065-069 🔽 🟹 🟹 🟹	CHECK_TEMP	Carrie Check Time						
wna070-074 🔽 🟹 🟹 🟹	CHECK_NRPE_NETWO	RK_INTERFACE UNKNOWN CHECK_NEPE: Souks	Conecul after 45	0.0% 0.0% 0.1% 0.1% 0.0% 0.0% 0.0% 0.0%				
wna075-079 🔽 🏹 🟹 🟹	CHECK_DHART	UNKNOWN CHECK_NRPE: Seeke	t timeout after 45					
wna080-084 🔽 🟹 🟹 🟹	CHECK_NRFE_CPU_R	UNKNOWN CHECK_NRPE: Seeke	t timeout after 45					
wna085-089 🔽 🟹 🟹 🟹	CHECK_HCELOG	UNKNOWN CHECK_NEPE: Seeks	Conecut after 45	178,226 8.85				
wna090-094 🔽 🟹 🟹 🟹	CHECK_HOD_TEMP	UNKNOWN CHECK_NRPE: Socks	t timeout after 45	0.10 1.10 1.10 1.10 1.10 1.10 1.10 1.10				
wna095-099 🔽 🟹 🟹 🟹	CHECK_LOAD	UNKNOWN CHECK_NRPE: Seeks	t timeout after 45	0.1%				
wna100-104 🔽 🟹 🟹 🟹	CHECK_DIR_FILLED	UNKNOWN CHECK_NRPE: Seeks	t timeout after 45	0.3%				
wna105-109 🔽 🟹 🟹 🟹			JINR GW	1.0% BROCADE FACTORY				
wna110-114 🔽 🟹 🟹 🟹				0.1%				
wna115-119 🔽 🟹 🟹 🟹	TAPE LIBRARY	Tier-1 dCache		RAIN BROCADE LACTORY / CHECK TRAFFIC SMP. REOCADE TendinabitEthernet48				
wna120-124 🔽 🟹 🟹 🟹	drives01-06 🔽 🟹 🟹 🟹 🟹	se-hd02 🔽		6.1% 0.0% 0.0% 6.0 M				
wna125-129 🔽 🟹 🟹 🟹	drives07-12 🔽 🟹 🟹 🟹 🟹	Tier-1 blades CMM						
		r14cmm0-2 🔽 🔽 🔽						
apc-rc-10 217.2	c-rc-10 🕱 17.2 THEMPERATURE r24cmm1-4 😨 😨 😨							
apc-rc-13 2 17,4	t1-temper01.lit-tech.jinr.ru	CHECK_THERMOMETER	25-40%					
aporc-19 216.8								
apc-rc-24 217.1 12 14 apc-rc-27 217.1 12			70-85%	- Colportance (Molit) 130:00				
apc-rc-2/ 1/.1			85-100	2% T1-SW-3 T1-SW-4 Input trafic (Mbit) 100,32				

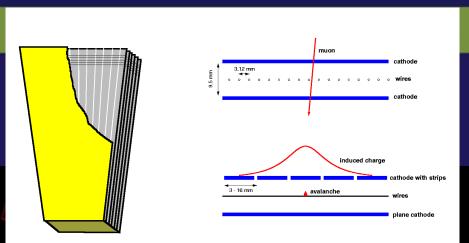




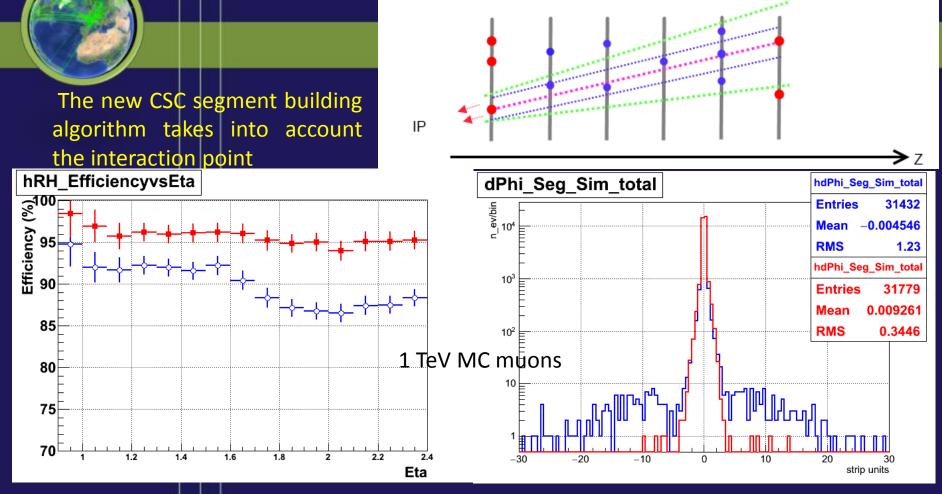
CMS Experiment at the LHC, CERN

Data recorded: 2011-Jun-05 09:01:21.346043 GMT(04:01:21 CDT) Run / Event: 166512 / 337493970

JINR LIT CMS group Cathode Strip Chambers (CSC)



JINR LIT CMS group Development of a new CSC segment building algorithm

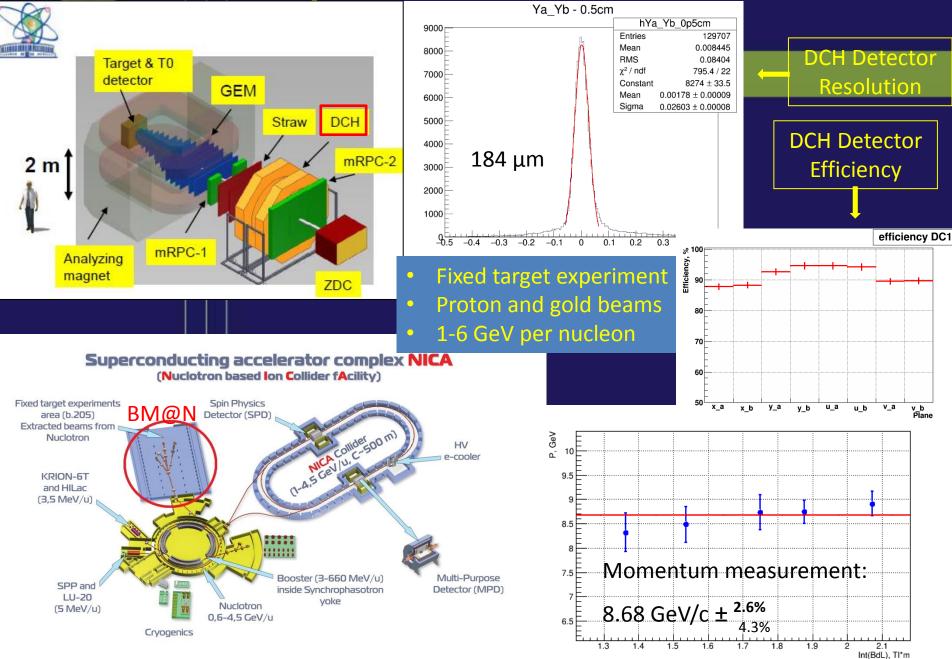


Difference in angles of the reconstructed and simulated segments (strip_unit ~ 3mrad)

Reconstruction efficiency vs pseudorapidity

~5000 CRAB jobs ~5TB skimmed datasets

Baryonic Matter at Nuclotron



NICA Project

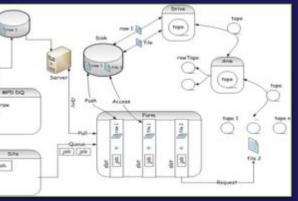


NICA project data amount estimation:

- High frequency of registered events (up to 6 KHz);
- in Au-Au collisions at the expected energies more than 1000 charged particles per event will be produced;
- overall number of events 19 billions;
- overall amount of data produced per year –
 30 PB and 8.4 PB more after analysis.

The model of a distributed computer infrastructure

The model for detailed investigation and estimation: ✓Tape robot, ✓Disk array, ✓CPU Cluster.





Conclusions

- The concept of grid perfectly fits the LHC project, making possible tremendous amounts of data transfers and job processing.
- * The JINR Tier-1 site along with the Russia Tier-2 sites gives the possibility for physicists from JINR, member states to fully participate in processing and analysis of the LHC experimental data.
- * An enormous experience in building and maintaining big data processing and storing center was obtained that can be very useful for the development of large scale projects at JINR and other member states.

Thank you for your attention!

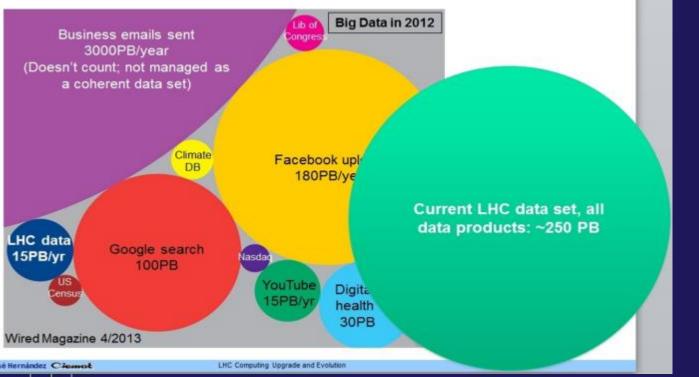


Back-up slides



Stepping into Big Data era

Where is LHC in Big Data Terms?



The comparison plot of the worldwide processed data shows that the amount of data that comes from the LHC fits the term of Big Data.

The expected amount of data received from LHC should become 2.5 times bigger in Run2, that will require the increase in resources and the optimization of their usage.