

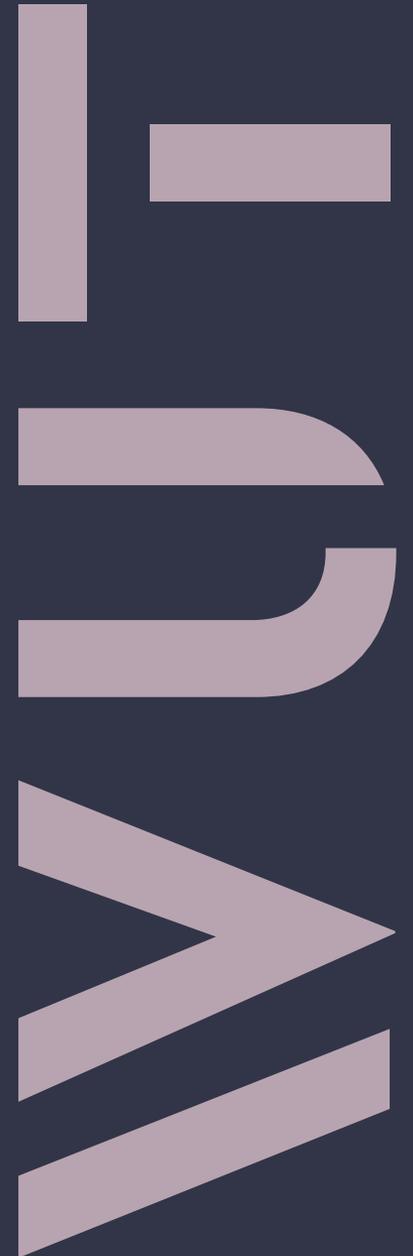
Managing large scale computing for MPD experiment.

Krzysztof Dygnarowicz
Supervisor:
Oleg Rogachevsky



Faculty of Electronics
and Information
Technology

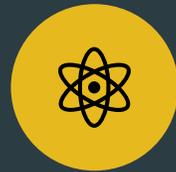
WARSAW UNIVERSITY OF TECHNOLOGY



What we will talk about



What is
MPDroot



Why we
Simulate



Resources



Tasks

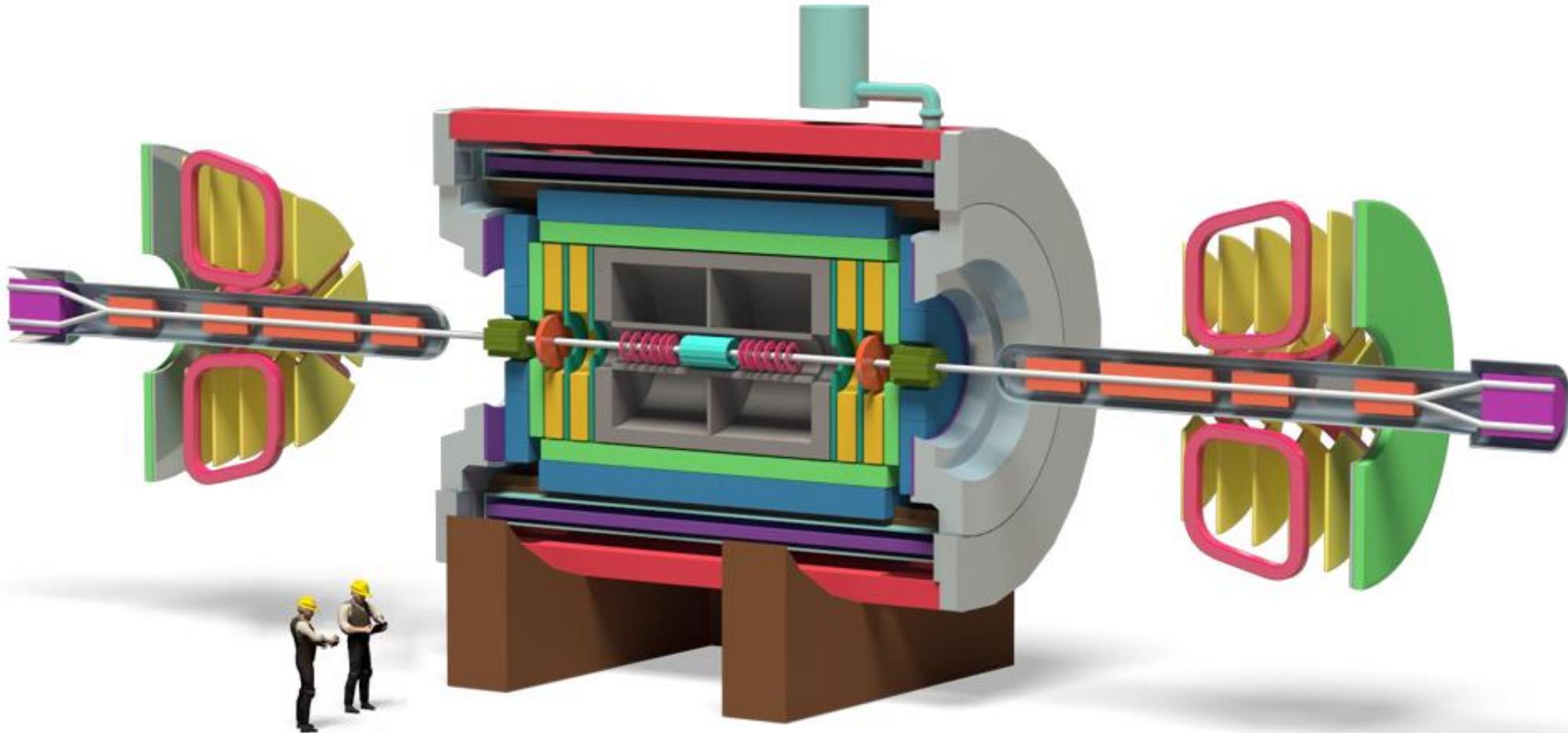


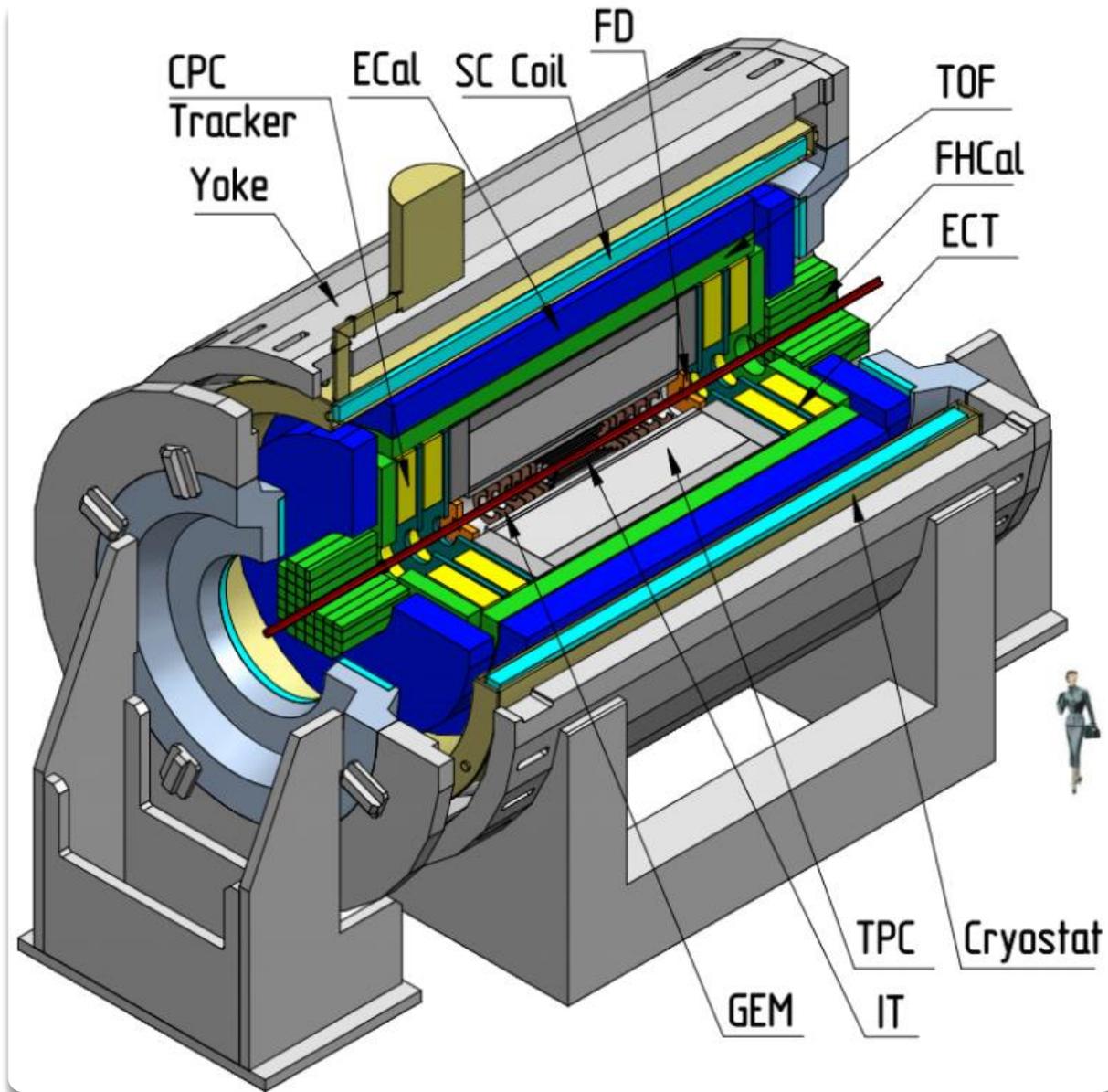
Results



Future
plans

MPD Simulations





MPDroot

Why we doing simulations now?

To check if our model makes sense

To make reconstruction programs

To know what to implement in FPGA

NICA subsystem	Technical data rate (GB/s)	Event rate (kHz)	Event size (MB)	Full event size (GB)	Mean data transfer rate (Gb/s)	Data volume (TB/24 hours)
Accelerators						
2019-2020	0.5				0.1	4
>2020	1.5				0.3	10
BM@N						
2019-2020		30	0.5	15	20	100
>2020		50	0.7	35	100	300
MPD						
2021-2022		0.1	1	0.1	10	200
>2022		6	2	12	100	600
SPD						
>2023		50	0.5	25	100	1000

Large data problem

Resources

- ▶ NICA Cluster
- ▶ Heterogeneous platform “HybriLIT”
 - ▶ SUPERCOMPUTER “Govorun”
 - ▶ HybriLIT
- ▶ Other clusters working in Russian Data Intensive Grid (RDIG)

HYBRILIT HETEROGENEOUS COMPUTING PLATFORM

Unified software and information environment

HybriLIT education and testing cluster

SUPERCOMPUTER «GOVORUN»

Intel Xeon

Intel Xeon Phi

Nvidia Tesla K20

Nvidia Tesla K40

Nvidia Tesla K80

10 computation nodes

40 nodes

CPU-component
Intel Skylake



21 nodes

CPU-component
Intel Xeon Phi (KNL)



5 nodes

GPU-component
GPU DGX-1 Volta
(NVIDIA Tesla V100)



Total Peak Performance
Double precision **500 Tflops**
Single precision **1000 Tflops**

HYBRILIT

Software

SYSTEM SOFTWARE (LINUX, SLURM, MODULES)



Scientific Linux



slurm workload manager

...



PARALLEL COMPUTING SOFTWARE



Enabling HPC since 1997



NVIDIA CUDA

...

APPLICATION PACKAGES



...

LICENSED SOFTWARE PACKAGES



...

FRAMEWORKS AND LIBRARIES FOR MACHINE LEARNING AND BIG DATA ANALYTICS



LIBRARIES



...

3	n01p003	288	203.7	1.5 %	43 %	0 %	0 %	55.5 %	30.24 GB	62.06 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s
4	n01p004	288	206.41	1.1 %	43.4 %	0 %	0 %	55.5 %	29.66 GB	56.98 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s
5	n01p005	288	206.79	1.3 %	43.2 %	0 %	0 %	55.5 %	30.62 GB	61.60 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s
6	n01p006	288	198.53	3.7 %	85.2 %	0 %	0 %	11.1 %	43.45 GB	50.08 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s

			CPU				Memory			Disks		Network	
name	cores	load	sys	user	used	cached	total	Write	Read	In	Out		
blade01	48	49.93	5.5 %	94.5 %	23.37 GB	101.97 GB	125.88 GB	0 B/s	0 B/s	2.7 KB/s	8.13 KB/s		

14	n01p014	288	208.69	1.3 %	43.2 %	0 %	0 %	55.5 %	30.38 GB	60.96 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s
15	n01p015	288	197.67	3.6 %	85.3 %	0 %	0 %	11.1 %	47.70 GB	11.26 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s
16	n01p016	288	216.23	1.3 %	43.3 %	0 %	0 %	55.5 %	35.49 GB	55.90 GB	94.12 GB	0 B/s	0 B/s	66 B/s	2.48 KB/s
Total		4608	3262.19	-	-	-	-	-	624.26 GB	647.16 GB	1.47 TB	0 B/s	0 B/s	1.89 KB/s	40.8 KB/s
Average		288	203.89	2.4 %	64.3 %	0 %	0 %	33.3 %	39.02 GB	40.45 GB	94.12 GB	0 B/s	0 B/s	120 B/s	2.55 KB/s

Monitoring of resources

Job Monitor

Updated: 2019-08-21 09:18 [UTC](0 00:00) Displaying topics 1 - 500 of 40442

JobId	Status	MinorStatus	ApplicationStatus	Site	JobName	LastUpdate[UTC]	LastSignOfLife[UTC]	SubmissionTime[UTC]	Owner
<input type="checkbox"/> 152973	Done	Execution Complete	Unknown	CLOUD.JINR.ru	BulkSubmission_sim_1k_187	2019-08-21 08:33:48	2019-08-21 08:33:48	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152972	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_186	2019-08-21 08:37:44	2019-08-21 08:37:44	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152971	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_185	2019-08-21 09:11:41	2019-08-21 09:11:41	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152970	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_184	2019-08-21 08:52:16	2019-08-21 08:52:16	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152969	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_183	2019-08-21 09:08:06	2019-08-21 09:08:06	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152968	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_182	2019-08-21 08:38:17	2019-08-21 08:38:17	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152967	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_181	2019-08-21 08:14:36	2019-08-21 08:14:36	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152966	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_180	2019-08-21 08:36:39	2019-08-21 08:36:39	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152964	Done	Execution Complete	Unknown	CLOUD.JINR.ru	BulkSubmission_sim_1k_178	2019-08-21 08:28:14	2019-08-21 08:28:14	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152961	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_175	2019-08-21 08:36:00	2019-08-21 08:36:00	2019-08-21 05:52:55	nkutovskiy

<input type="checkbox"/> 169212	Waiting	Pilot Agent Submission	CLOUD.NU.kz	SimpleLsTest
<input type="checkbox"/> 169211	Done	Execution Complete	CLOUD.GREN...	SimpleLsTest
<input type="checkbox"/> 169210	Done	Execution Complete	CLOUD.GREN...	SimpleLsTest

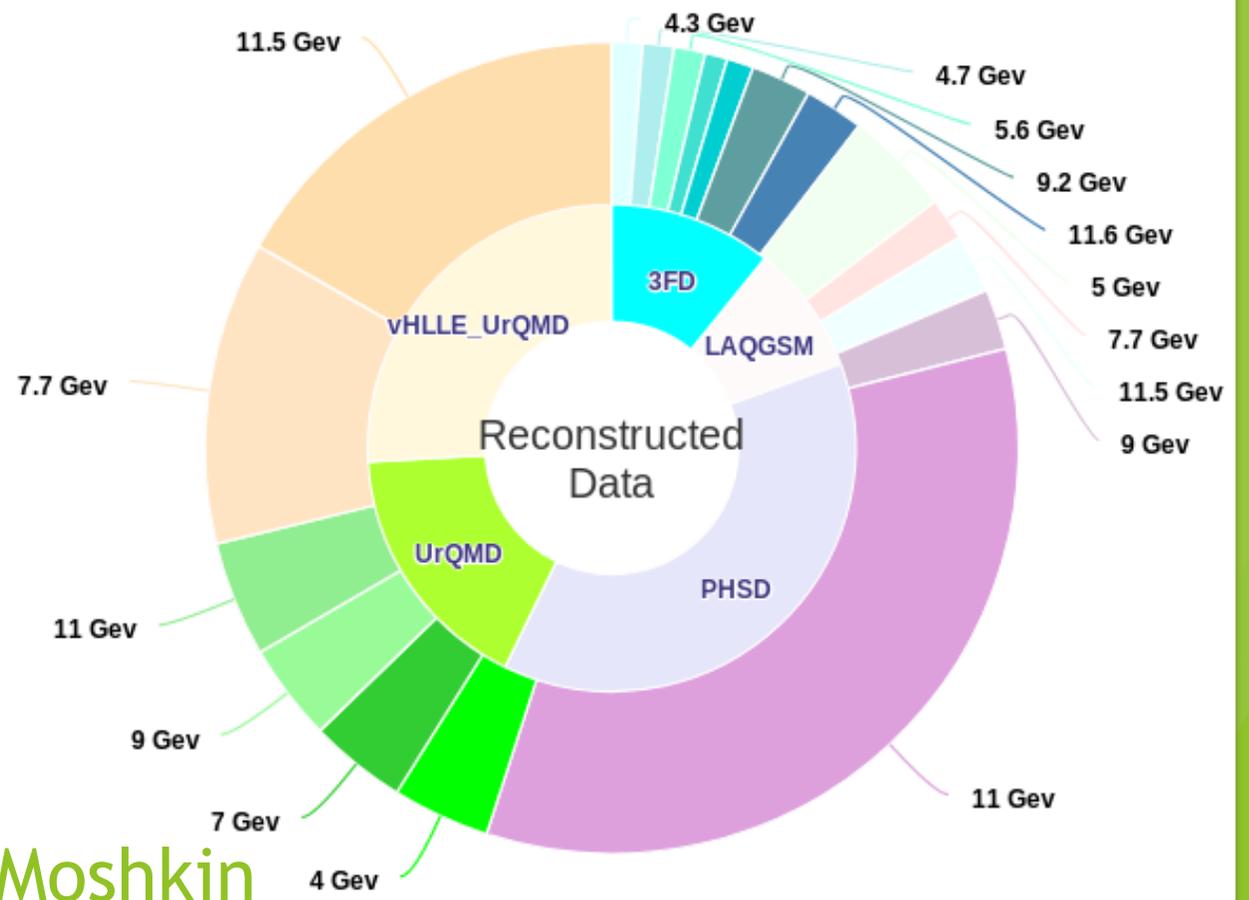
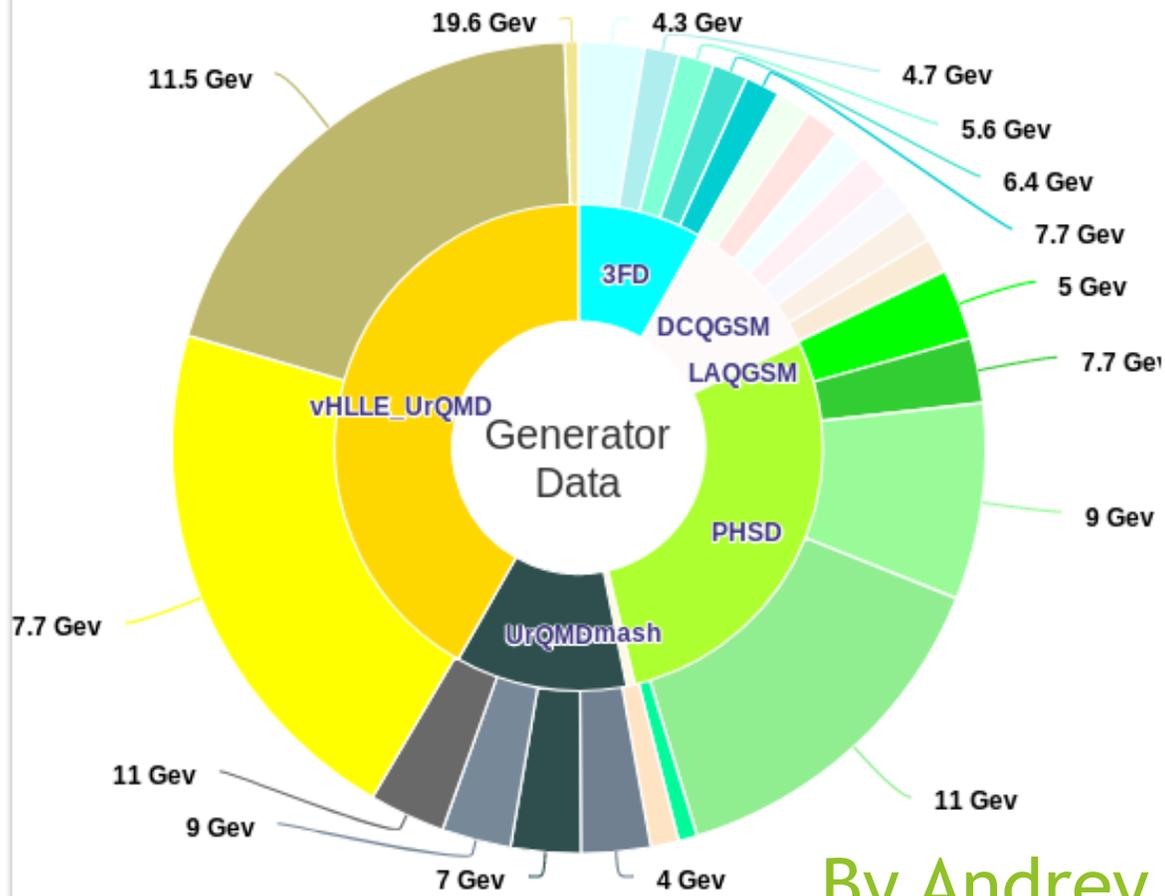
<input type="checkbox"/> 152944	Done	Execution Complete	Unknown	CLOUD.GREN...	BulkSubmission_sim_1k_158	2019-08-21 08:11:59	2019-08-21 08:11:59	2019-08-21 05:52:55	nkutovskiy
<input type="checkbox"/> 152942	Done	Execution Complete	Unknown	CLOUD.JINR.ru	BulkSubmission_sim_1k_156	2019-08-21 08:18:11	2019-08-21 08:18:11	2019-08-21 05:52:55	nkutovskiy

Monitoring jobs

Show jobs | Items per page: 100 | Page 1 of 1 | Updated: -

<input type="checkbox"/>	Name	Tier	GridType	Flag	MaskStatus	Status	Efficiency (%)	Staging	Running	Completed	Done	Stalled
Country: Unknown												
<input type="checkbox"/>	Multiple	Tier-2	Unknown		NoMask	Good	100.0	0	0	0	2	0
Country: Russia												
<input type="checkbox"/>	DIRAC.JINR-TIER.ru	Tier-2	DIRAC		Active	Good	100.0	0	442	0	3723	0
<input type="checkbox"/>	DIRAC.JINR-CREAM.ru	Tier-2	DIRAC		Active	Good	100.0	0	412	0	2693	0
<input type="checkbox"/>	CLOUD.PRUE.ru	Tier-2	CLOUD		Active	Good	100.0	0	0	0	9	0
<input type="checkbox"/>	CLOUD.JINR.ru	Tier-2	CLOUD		Active	Good	100.0	0	0	0	397	1
Country: Georgia												
<input type="checkbox"/>	CLOUD.GRENA.ge	Tier-2	CLOUD		Active	Good	99.0	0	43	129	572	0
Country: Belarus												
<input type="checkbox"/>	CLOUD.INP.by	Tier-2	CLOUD		Active	Good	100.0	0	0	0	4	0
Country: Azerbaijan												
<input type="checkbox"/>	CLOUD.IPANAS.az	Tier-2	CLOUD		Active	Good	100.0	0	0	0	16	0

Other clusters in RDIG



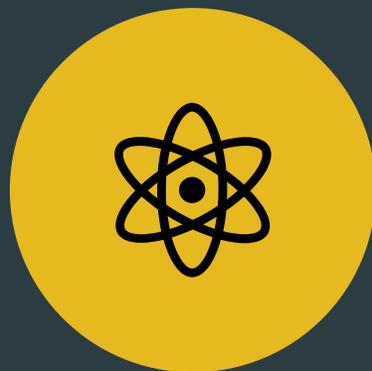
By Andrey Moshkin

Data of simulations

My tasks



HOW TO GET
ACCESS?



HOW TO WORK WITH
DIRAC AND LIT
CLUSTER?



HOW TO USE DIRAC
ON NICA CLUSTER?

What has been done?

- ▶ Scripts to product events and reconstruction working on LIT cluster.
- ▶ DIRAC job manager scripts - **in progress.**
- ▶ Instructions how to get access and run your jobs on LIT and by DIRAC job manager.

Outline of scripts

1. Get information of JOBID
2. Link generator and MpdRoot
3. Generate data (runqmd.bash for URQMD)
4. Run them by detector (runMC.c)
5. Reconstruct event from detector data (reco.c)
6. Copy data to safe storage
7. Erase unnecessary data

Instruction

- ▶ How to get access
- ▶ How to start Jobs
- ▶ Where to store data

system.

2.1 Getting access

First you have to fill form on website^[2]. As Project funding source during my registration I chose 02-0-1065-2007/2019. Mostly during working with MPDroot you will use CPU. Next you have to get a signers and get it to place in LIT mentioned in website.

2.2 Starting work

First you need to chose which resources you will be using. Mostly you should add module of Goverun.

```
$ module add GVR/v1.0-1
$ sinfo
```

After this you should see available nodes .

PARTITION	AVAIL	TIMELIMIT	NODES	STATE	NODELIST
kn1	up	30-00:00:0	1	mix	n01p016
kn1	up	30-00:00:0	15	alloc	n01p
skylake*	up	30-00:00:0	1	comp	n02p022
skylake*	up	30-00:00:0	12	mix	n02p
skylake*	up	30-00:00:0	20	alloc	n02p
dgx	up	30-00:00:0	1	drain	dgx05
dgx	up	30-00:00:0	3	mix	dgx
dgx	up	30-00:00:0	1	alloc	dgx01

To manage platform and jobs cluster uses SLURM system. It provides three key functions:

Future plans

Improve	Improve scripts and file storage system
Extend	Extend instruction of NICA cluster and future improvements of DIRAC system
Generate	Generate events

Sources

- ▶ <https://home-hlit.jinr.ru/#/>
- ▶ http://hlit.jinr.ru/en/installed_software_eng/
- ▶ http://hlit.jinr.ru/en/for_users_eng/hardware-and-software-environment_eng/
- ▶ <http://db-nica.jinr.ru/mpdmc/stat.php>
- ▶ <https://home.cern/science/computing/storage>
- ▶ http://mpd.jinr.ru/wp-content/uploads/2018/06/NICA_computing_TDR_1.03.pdf
- ▶ <http://nica.jinr.ru/projects/mpd.php>



Thanks for your attention
kdygnaro@protonmail.com