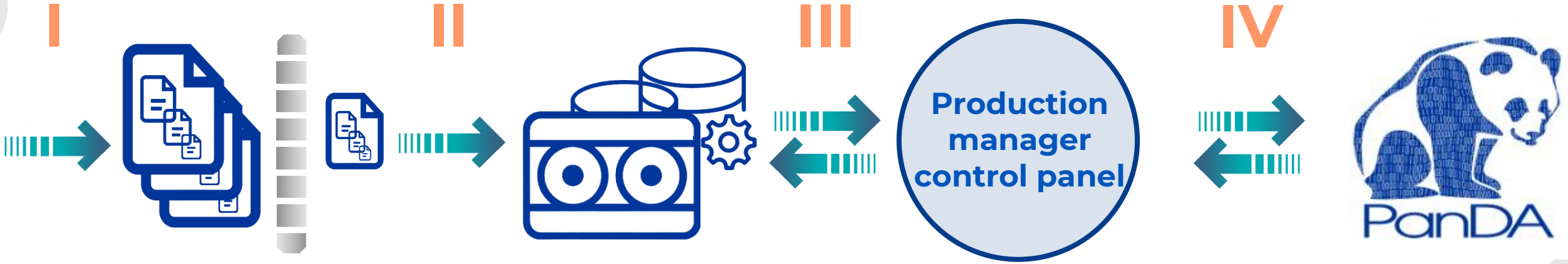
XII SPD collaboration meeting

Production system status update

Speaker: Monakov Nikita Glebovich
software engineer , MLIT JINR

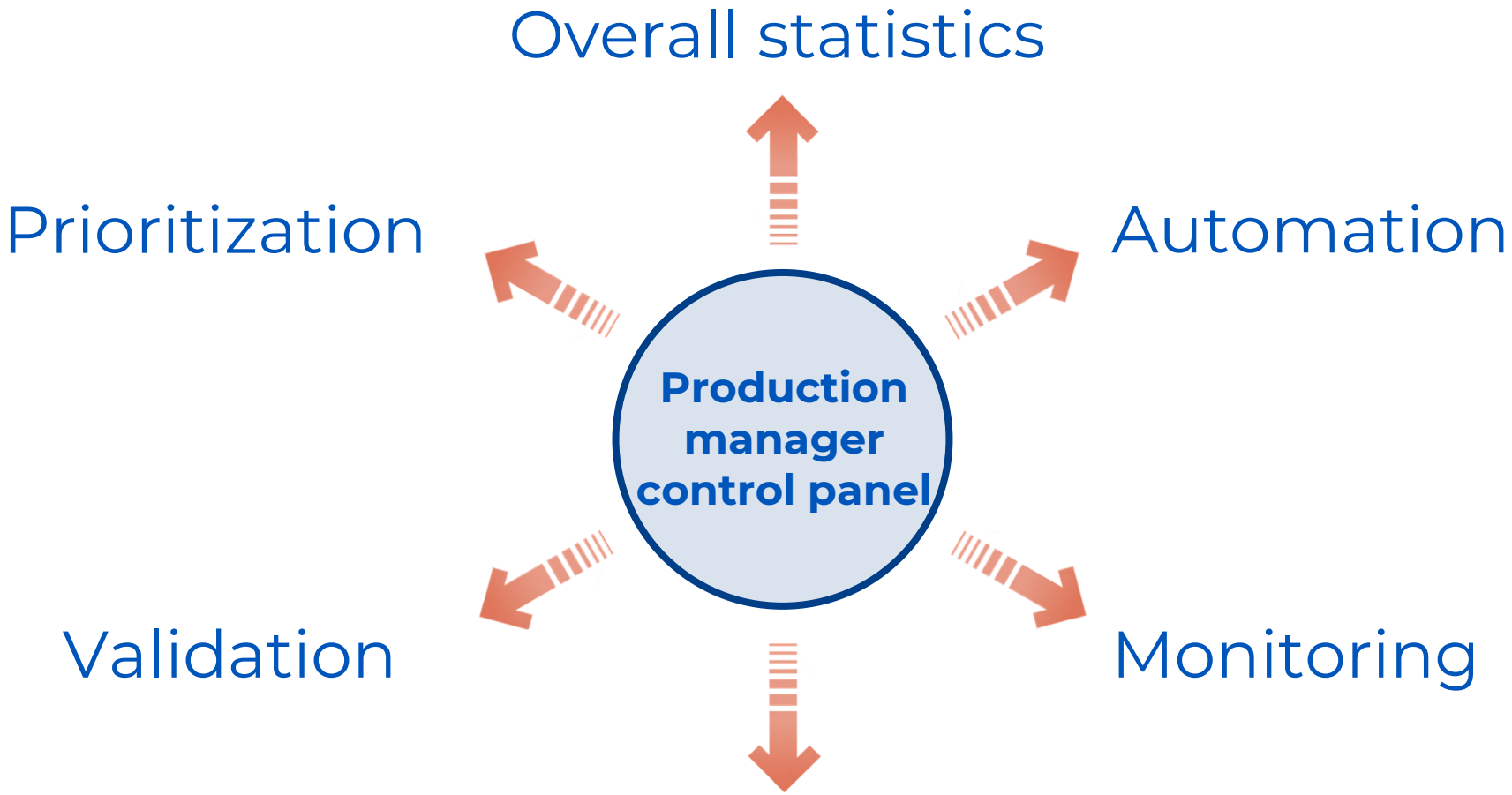
Motivation for developing the system



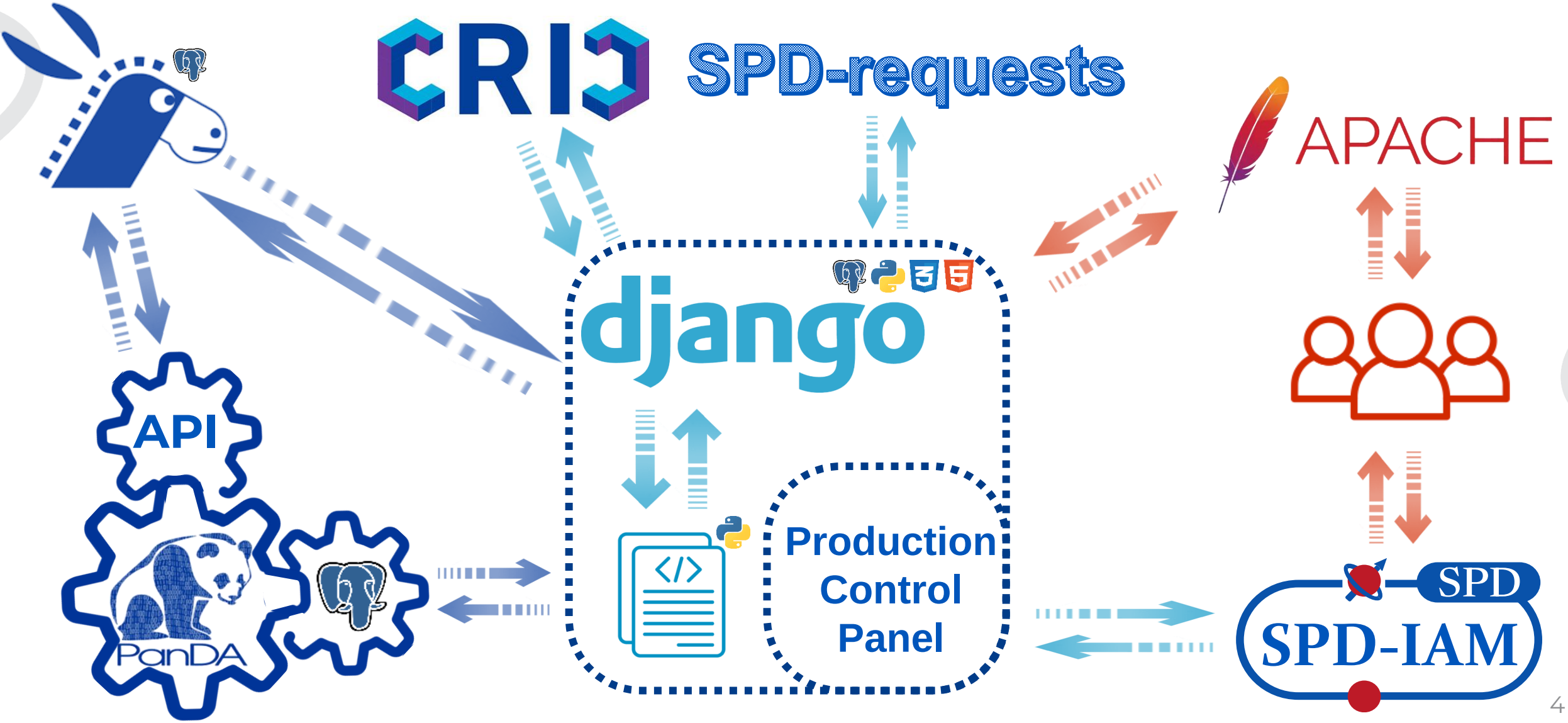
The system is designed to:

- Centralize data processing and enhance its scalability
- Simplify task submission workflows
- Efficiently allocate computing resources to physics groups based on quotas and priorities
- Reduce overhead from multi-stage processing of invalid tasks

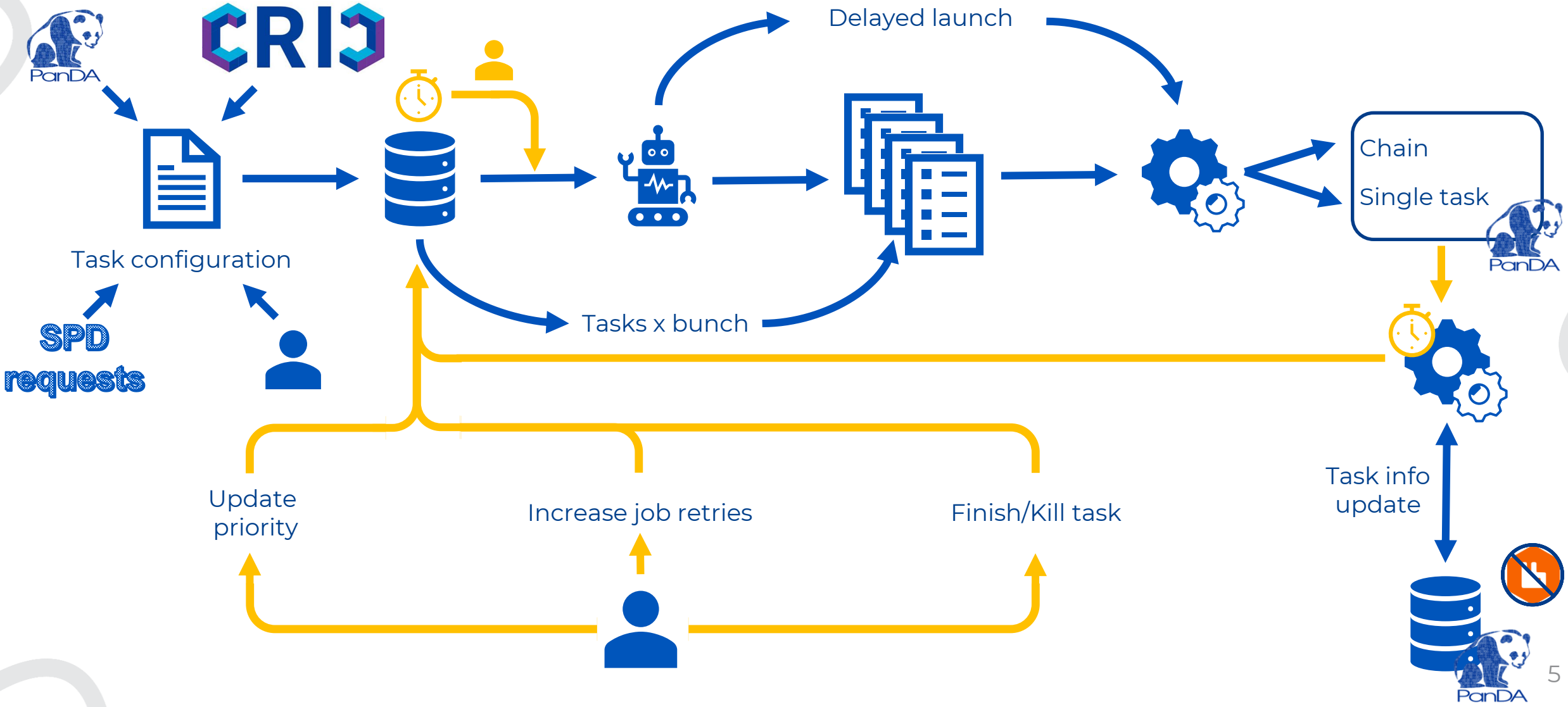
Current service concept



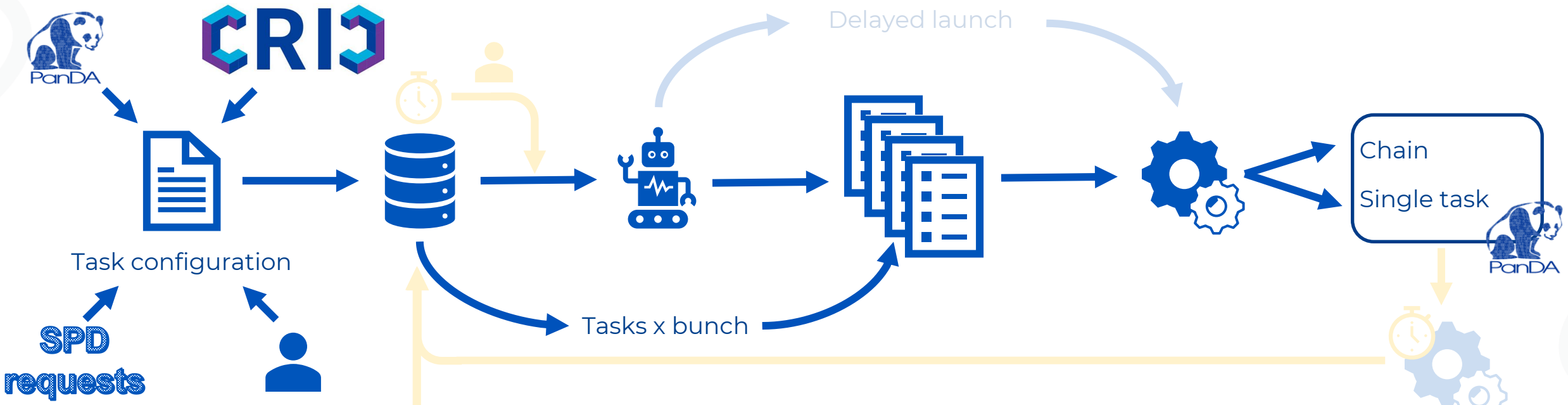
Technology stack



Work done 1/5


PanDA
CRIC
Delayed launch
Chain
Single task
PanDA
Task configuration
SPD
requests
Tasks x bunch
Update
priority
Increase job retries
Finish/Kill task
Task info
update
PanDA

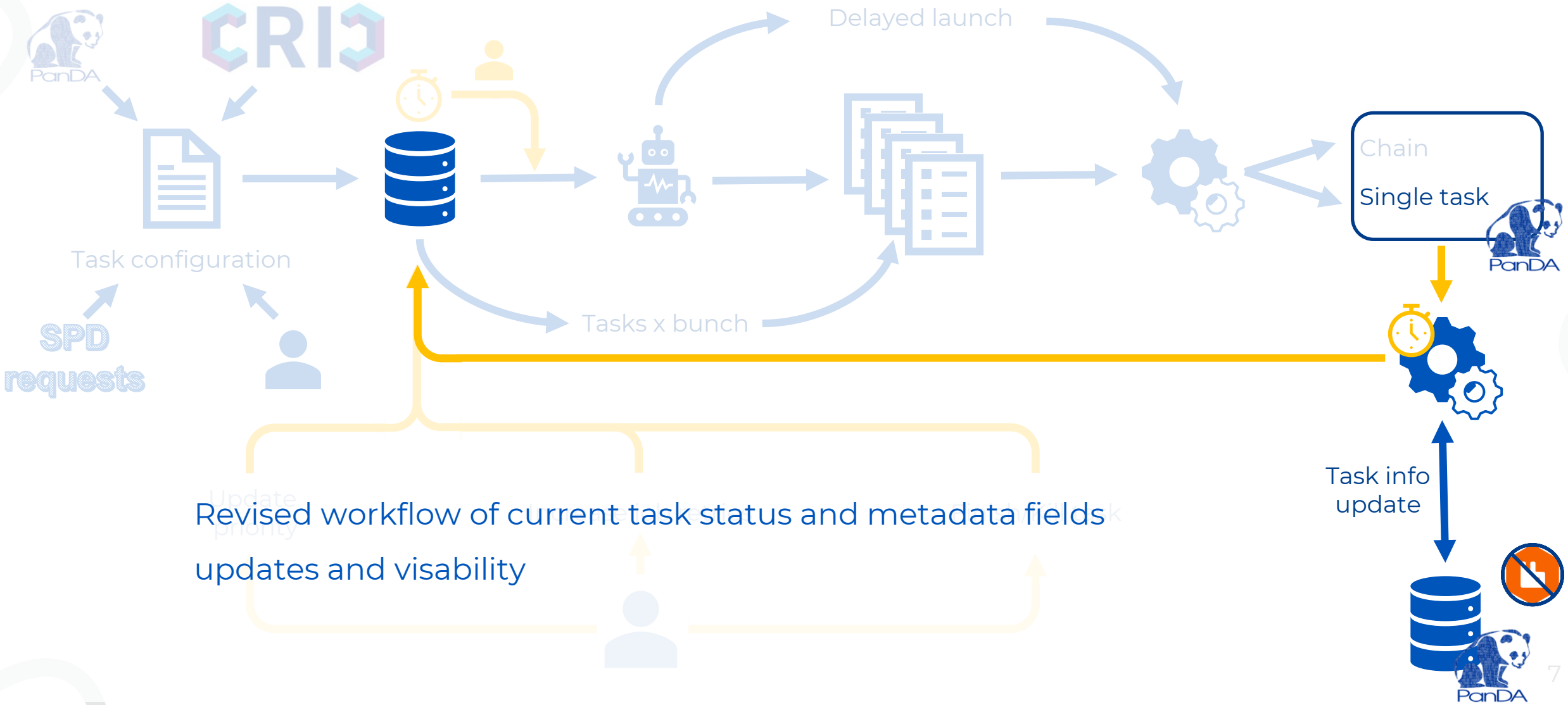
Work done 2/5



Updated the internal task registration mechanism (user form and internal DB schema) for subsequent task submission to an external experiment service (PanDA WMS)

Task info update

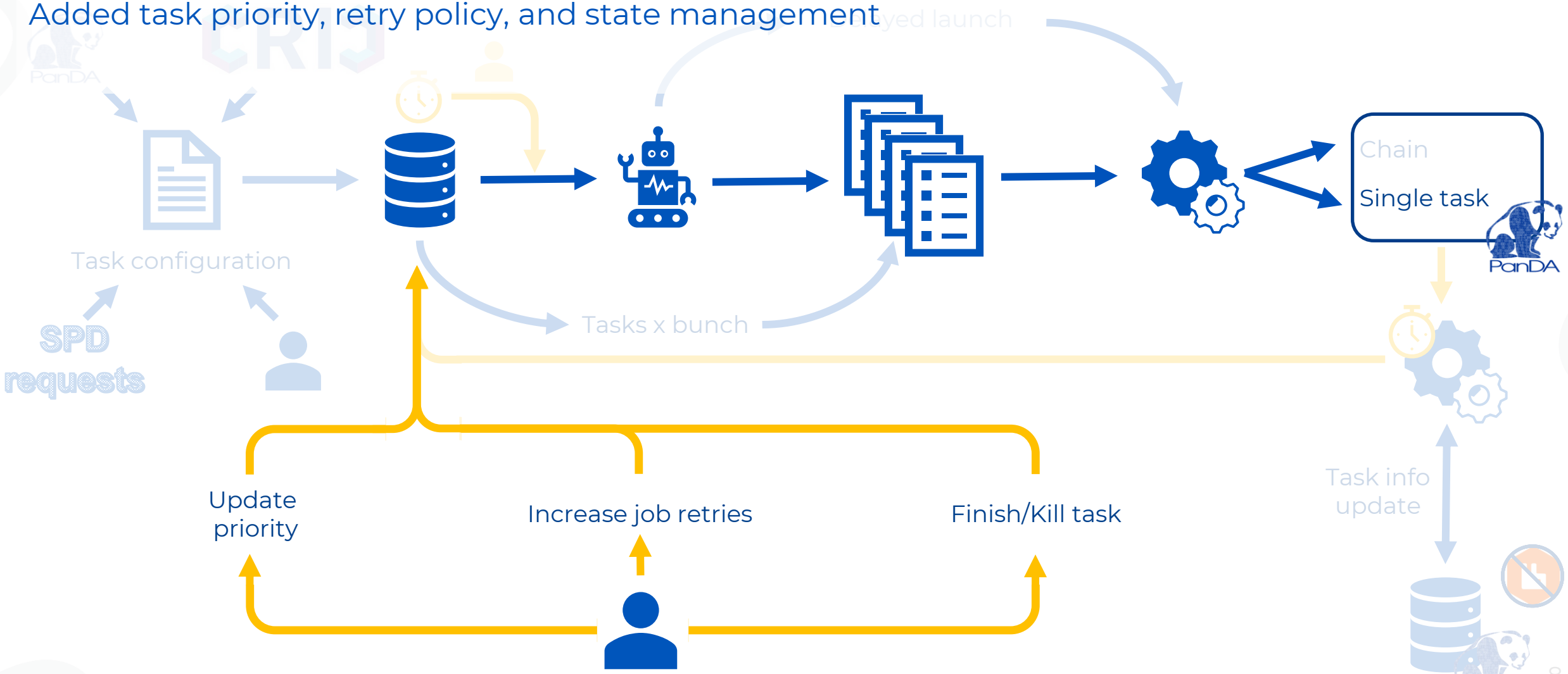
Work done 3/5



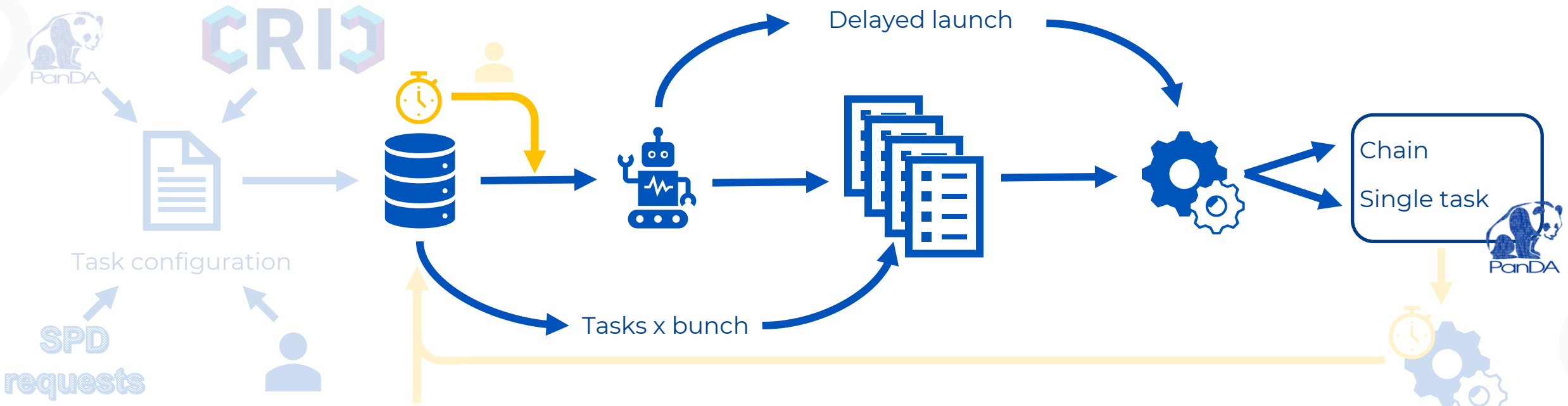
Revised workflow of current task status and metadata fields updates and visability

Work done 4/5

Added task priority, retry policy, and state management



Work done 5/5



- Implemented workflow for chains of arbitrary length
- Auto priority increase *

  new launch – update previous

  complete 90% - update self

Task creation forms

Simulation

Task Creation

Processing Type: MC

Data Type: SIMUL

Task name:

Request: PROD2026-018

Tasks to launch: 1

Output dataset name: MC2026_S2.

Naming convention here

Total events per task: 0

Events per job: 0

Cloud: RU

Data disk: SPDDATA

Skip scout:

Initial offset: 0

Job retry policy: 5

Path to execution files:

smth like -> /cvmfs/spd.jinr.ru/production/MC/minbias-P8-spdroot417-dev.10GeV.V01

Path to container:

smth like -> /cvmfs/spd.jinr.ru/images/spdroot-dev-4.1.7.sif

Create task

Chains

Create Chain

Processing Type: MC

Chain name:

Chain step #1

Step type: Select step type

Select step type
SIMUL
RECO

Step type: Select step type

Remove

Add step

Create chain

Monitoring



Select field | Enter value | Filter

Task ID	Task name ↑ ↓	Parent ID	Creator	Status	Substatus	Done jobs	Default/Current priority	Total events	Submit time ↑ ↓	Start time ↑ ↓	End time ↑ ↓
799	PROD2026-032.RECO.2	799	Elena Zemlyanichkina	running	pending	9472	900/910	0	16:14, 13 Nov 2026	16:14, 13 Nov 2026	None
798	PROD2026-035.RECO.1	798	Elena Zemlyanichkina	finished	---	9728	900/910	0	06:50, 11 Nov 2026	06:50, 11 Nov 2026	06:09, 18 Nov 2026
797	PROD2026-034.RECO.1	797	Elena Zemlyanichkina	finished	---	9787	900/910	0	06:49, 11 Nov 2026	06:49, 11 Nov 2026	15:08, 16 Nov 2026
796	PROD2026-033.RECO.1	796	Elena Zemlyanichkina	finished	---	9935	900/910	0	06:41, 11 Nov 2026	06:41, 11 Nov 2026	14:13, 15 Nov 2026
795	PROD2026-025.RECO.2	795	Elena Zemlyanichkina	finished	---	9300	900/910	0	12:21, 10 Nov 2026	12:22, 10 Nov 2026	15:05, 16 Nov 2026
794	PROD2026-024.RECO.2	794	Elena Zemlyanichkina	running	finishing	9991	900/910	0	06:14, 10 Nov 2026	06:14, 10 Nov 2026	None
793	PROD2026-022.RECO.2	793	Elena Zemlyanichkina	finished	---	8703	900/900	0	14:54, 09 Nov 2026	14:54, 09 Nov 2026	16:49, 16 Nov 2026
792	TEST_RETRIES_3	792	Nikita Monakov	failed	---	0	900/900	0	11:49, 09 Nov 2026	11:49, 09 Nov 2026	06:59, 12 Nov 2026
791	PROD2026-032.RECO.1	791	Elena Zemlyanichkina	finished	---	7749	900/900	0			
790	PROD2026-021.RECO.2	790	Elena Zemlyanichkina	finished	---	8789	900/900	0			

Page: 1 2 3 ... 78 79 80

All tasks brief overview*

*Filtering & sorting functionality available

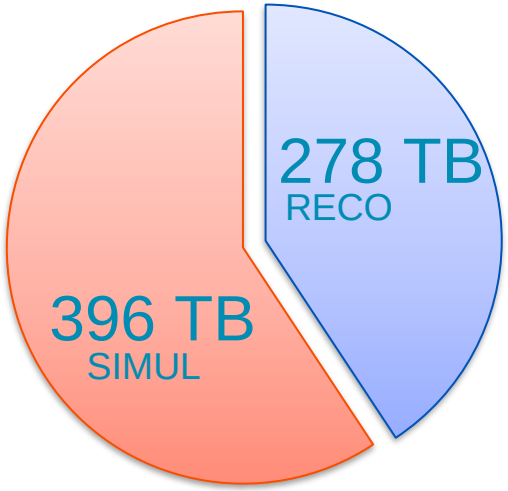
Exact task information and manipulation

Up priority (+5) | Increase retries | Finish task | Kill task :(

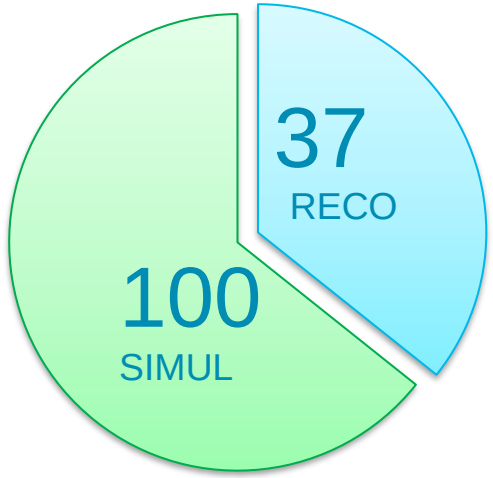
Task ID	799
Parent task ID	799
Chain task ID	None
Name	PROD2026-032.RECO.2
Type	managed
VO	spd.nica.jinr
Creator	Elena Zemlyanichkina
Task status	running
Task substatus	pending
Total events	0
Total jobs done	9472
Total requested jobs	184627
Priority	900
Current priority	910
Submit time	16:14 November 13, 2026
Start Time	16:14 November 13, 2026
Time stamp	8:15 November 18, 2026
End Time	
JEDI parameters	{"taskName": "PROD2026-032.RECO.2", "nFilesPerJob": 1, "noInput": false, "userName": "Elena Zemlyanichkina", "vo": "spd.nica.jinr", "taskPriority": 900, "architecture": "x86_64", "cloud": "RU", "workingGroup": "spd.nica.jinr", "skipScout": false, "reqID": 54, "maxAttempt": 70, "transUses": "A", "transHome": null, "transPath": "https://vm221-126.jinr.ru:8080/spd_simu_VA_transform.sh", "processingType": "step2", "prodSourceLabel": "managed", "taskType": "test", "ramCount": 1900, "log": {"dataset": "MC2026_S2.minbias-P8-spdrootdev-240.27GeV-UU.PROD2026-032.RECO.2.log", "type": "template", "param_type": "log", "token": "ddd:*DATADISK", "destination": "(type=DATADISK)", "offset": 210001, "value": "MC2026_S2.minbias-P8-spdrootdev-240.27GeV-UU.PROD2026-032.RECO.2.${SN}.log.tgz"}, "jobParameters": [{"type": "constant", "value": "singularity run --bind /cvmfs/spd.jinr.ru/production/MC/2026-06-23-test-prod:/prod -H ./:/WORKDIR /cvmfs/spd.jinr.ru/images/spdroot-4.2.0 spdroot.py -b -q '/prod/reco.C(4000, '), {"type": "constant", "value": "\\\""}, {"type": "template", "param_type": "input", "value": "${INS}", "dataset": "MC2026_S2:MC2026_S2.minbias-P8-spdrootdev-240.27GeV-UU.PROD2026-032.SIM.1.S"}, {"type": "constant", "value": "\\\","}, {"type": "constant", "value": "\\\""}, {"type": "template", "param_type": "input", "value": "${INP}", "dataset": "MC2026_S2:MC2026_S2.minbias-P8-spdrootdev-240.27GeV-UU.PROD2026-032.SIM.1.P"}, {"type": "constant", "value": "\\\","}, {"type": "constant", "value": "\\\""}, {"type": "template", "param_type": "output", "token": "ddd:*DATADISK", "destination": "(type=DATADISK)", "value": "r.MC2026_S2.minbias-P8-spdrootdev-240.27GeV-UU.PROD2026-032.RECO.2.${SN/P}.root", "dataset": "MC2026_S2.minbias-P8-spdrootdev-240.27GeV-UU.PROD2026-032.RECO.2.R", "offset": 210001}, {"type": "constant", "value": "\\\","}, {"type": "template", "value": "${RNDMSEED}", "param_type": "number", "offset": 210001}, {"type": "constant", "value": ")'"}]}

Statistics on XI meeting

- Successfully processed
  about 440k jobs
  across 137 tasks



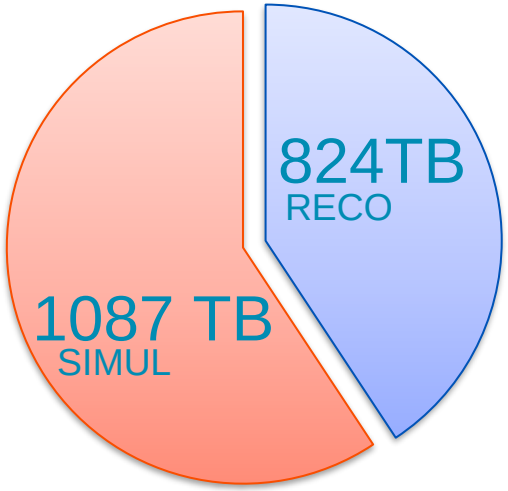




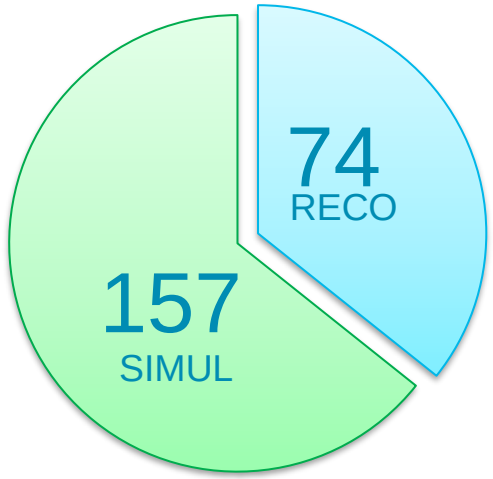
- Total output datasets volume –
  more than 680 TB

Current statistics

- Successfully processed
  about 1.1 million jobs
  across 231 tasks







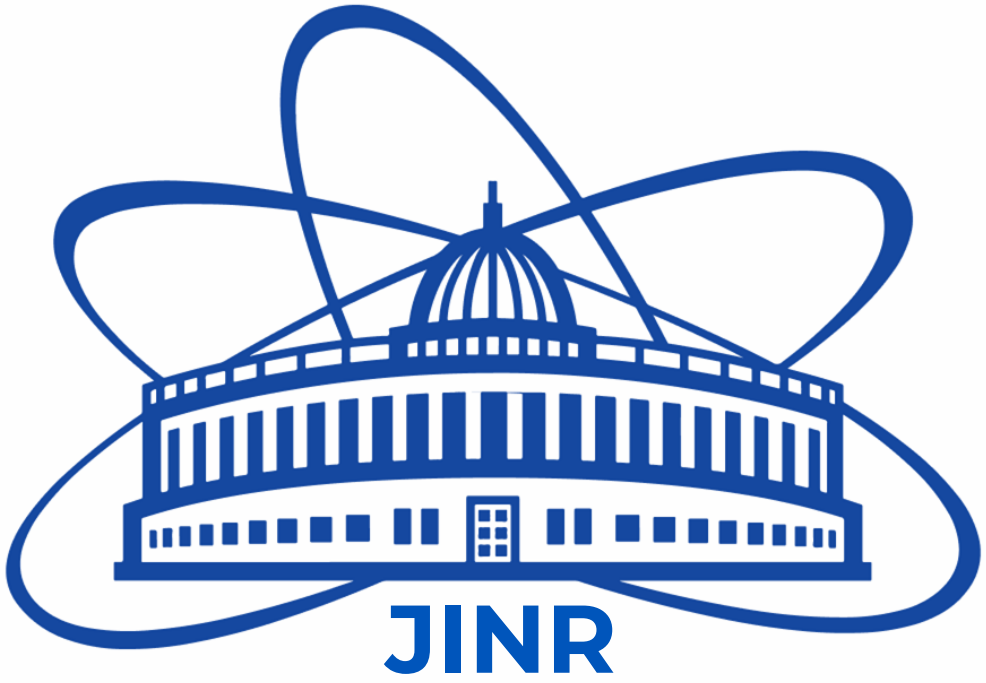
- Total output datasets volume –
  more than 1.9 PB

Current tasks and upcoming plans



- Delayed launch for users
  - ✓ Token storage
  - ✓ Token refreshing
  - ✓ Task modifying
  - ❑ FE and DB schema update

- Further automation of task workflow
  - ❑ Auto complete container and exe paths for exact request
  - ❑ Auto chain generation for full request
  - ❑ Dataset postprocessing



Thanks for attention

Work done previous time



- A web application prototype has been developed
  - User authentication via SPD-IAM
  - Task creation following naming conventions
  - Job status tracking with filtering and sorting capabilities
  - Configured decorators and middleware
  - First version of cloning mechanism
- Implementation of automation scripts
- Integration with PanDA WMS using JWT
- Logging interactions with PanDA WMS
- Task submitting process was highly refactored
- Internal DB structure for task chains, task bunches and single tasks
- FE & BE for chain processing
- Second version of cloning mechanism
- Integration with spd-requests.jinr.ru