

Monitoring and test suite for the SPD Computing system

Alexey Konak
JINR MLIT
konak@jinr.ru
Tomsk, Russia
XI SPD collaboration meeting

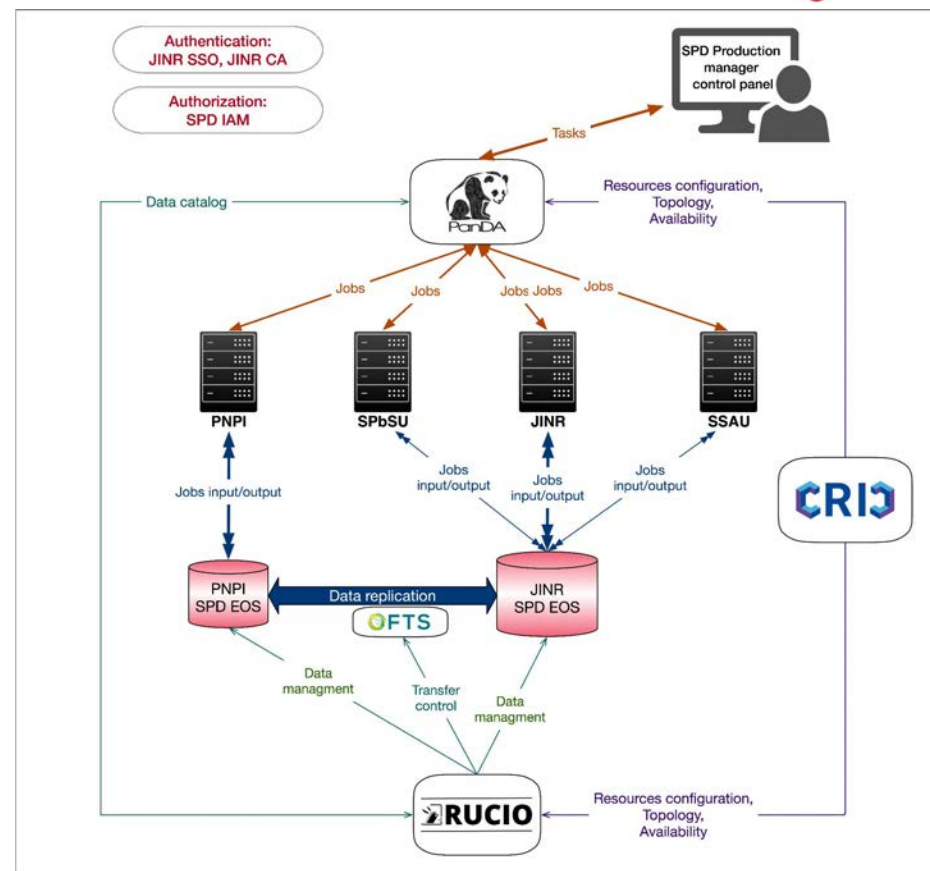
20.05.2026

Infrastructure

We already have a fairly extensive infrastructure.

- Authentication system: JINR SSO
- Authorization system: IAM
- Information system: CRIC
- Data management system: Rucio
- Data transfer service: FTS
- Workload management system: PanDA
- Computing Elements at JINR, PNPI, etc.
- Storage Elements at JINR and PNPI

Disabling any of these components may cause the entire system to shut down.



Informed means prepared

The goal is to increase awareness of the system status in real time. We need to test (to verify that all should work as expected) and monitor the statuses of the systems

To do that we are developing:

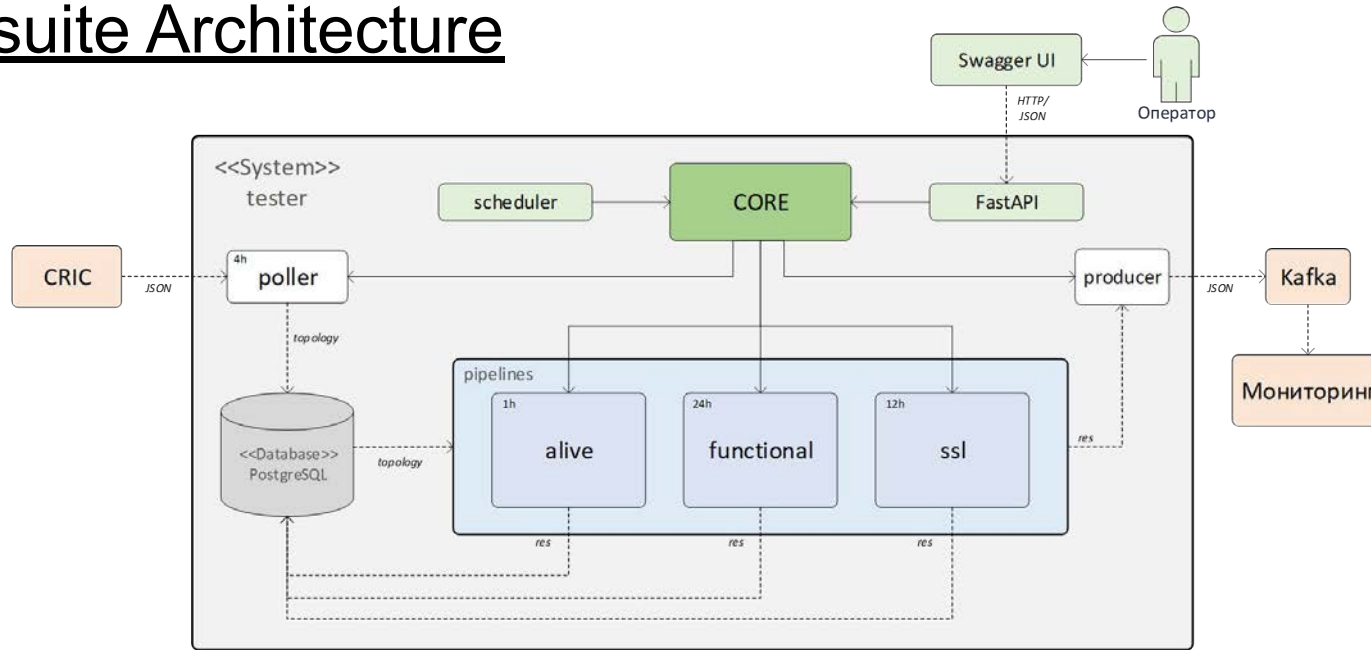
- A software platform for testing components of a geographically distributed data processing system for the SPD experiment
- Monitoring of the components of the geographically distributed data processing system for the SPD experiment.

Test suite Requirements

The requirements for the testing system being developed are as follows:

- Testing – support for various types of testing modules (Availability, Functional, SSL).
- Save information – store service topologies, results history in the database, logging the progress of testing.
- Integration and monitoring – importing the list of services from CRIC, Transfer of results to the monitoring system in real time.
- Management and schedule – scheduled periodic startup, manual tests.

Test suite Architecture



- The Core manages the testing lifecycle. It starts on a timer, forming test pipelines, performing the testing process, sending the results to the monitoring system.
- Pipeline is a chain of tests. Each pipeline runs independently. The tests within the same pipeline follow one after the other.

About tests

Alive

Services and systems must be accessible over the network.

Executed every 30 minutes.

Each type of object has its own test:

- Rucio – rucio ping
- PanDA – request “isAlive”
- Harvester – request “/icons”
- FTS – fts-rest-server-status
- CE – arcinfo -c <endpoint>
- SE – getinfo of / with webdav and xrootd protocols

SSL

Host certificates of services must be valid

Executed every day.

- Requesting a host certificate.
- Verification of the certificate.
- Checking the certificate's lifetime.

LifeTime > 7 days – SUCCESS

LifeTime ≤ 7 days – WARNING

LifeTime ≤ 0 – CRITICAL

Functional

Services and systems must correctly perform operations.

Executed every day.

Special tests for each service.
Not everything is ready yet.

Rucio – upload/download
FTS – delegate+transfer

Testing process



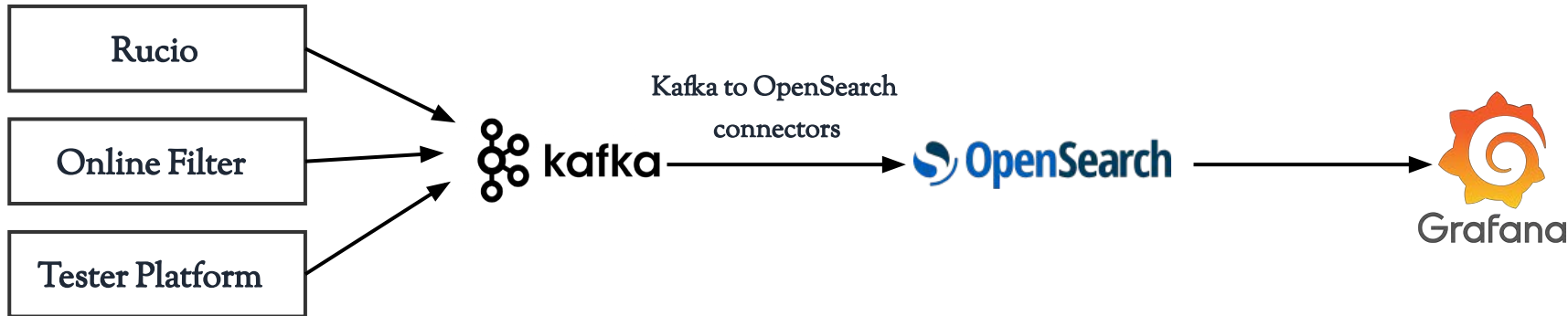
The system core performs the following actions:

- gets a list of testing services from CRIC;
- generates test pipelines (chains);
- runs test chains;
- saves the test results and sends them to the monitoring system.

Monitoring System

The main purpose of creating such monitoring is:

- Collect information about the health and operation of systems in one place
- Visualization of this data for analysis by system administrators



Data sources: collecting traces and metrics from critical experiment nodes and the testing system.

Apache Kafka: provides fault-tolerant reception and buffering of large data streams in real time.

OpenSearch: indexing and long-term storage of monitoring data with the possibility of fast full-text search.

Grafana: building dynamic dashboards and visualizations for operational monitoring of the distributed environment.

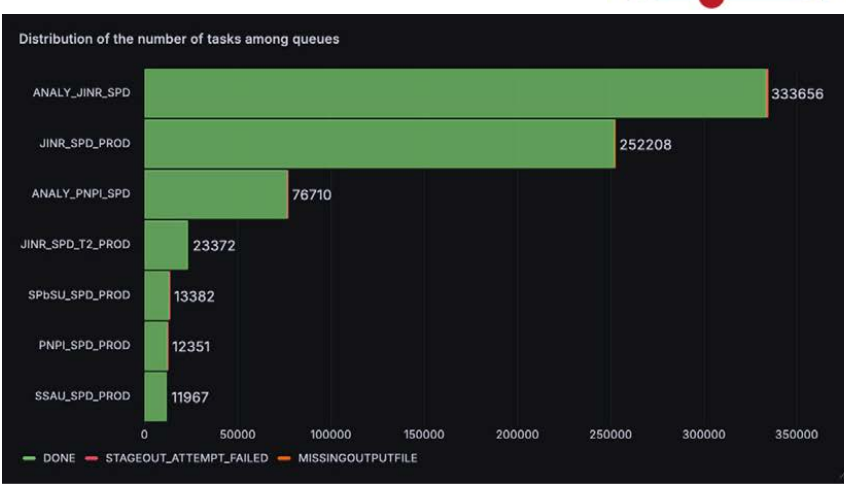
Rucio monitoring

- Collects messages and traces in OpenSearch
- Visualization this metrics

We can monitor:

- Deletion process (how many files deleted, where, the amount of free space, the time of the operation, errors that occur during the deletion process, etc.)
- Replication process (states of finished transfers, errors, amount of transferred data, transition states before transfer start, etc.)
- Users and Production system interactions with Rucio (upload/download files, created datasets, errors during single transaction, etc.)

Rucio dashboard examples



Online Filter monitoring

Collects info from

- Workload management system (about jobs and active pilots)
- Workflow management system (about tasks)
- Data management system (info about storage capacity, datasets)

We can track the process of completing tasks, errors that occur during processing, the fullness of input and output storage, etc.

Online Filter dashboard examples [1]

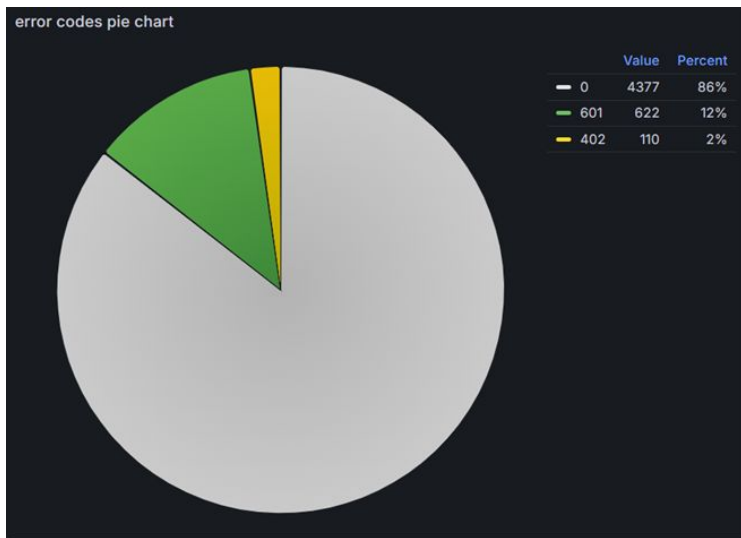
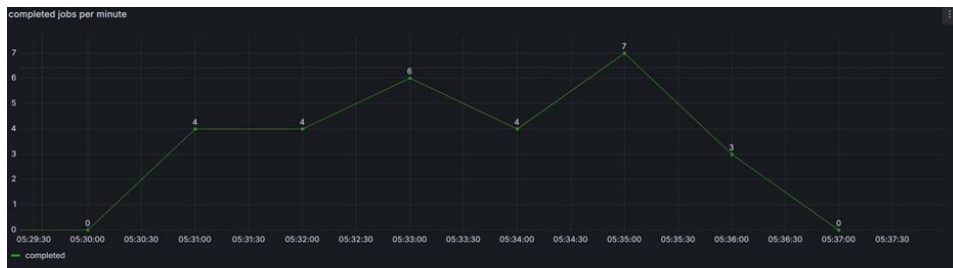
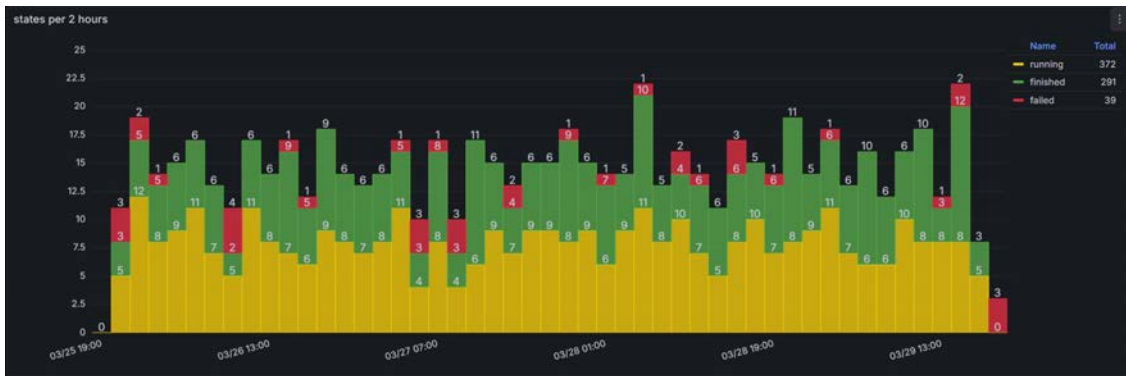
Info about Tasks

tasks info				
datetime	task_id	state	error_code	workflow_id
06.05.26 14:33:30	26398	finished	0	10490
06.05.26 14:33:29	26393	finished	0	10488
06.05.26 14:31:55	26398	running	0	10490
06.05.26 14:31:50	26393	running	0	10488
06.05.26 14:31:42	26384	finished	0	10485
06.05.26 14:31:41	26381	finished	0	10484
06.05.26 14:31:41	26391	finished	0	10488
06.05.26 14:31:40	26397	finished	0	10490
06.05.26 14:30:47	26390	running	0	10487
06.05.26 14:30:39	26388	finished	0	10487
06.05.26 14:30:07	26397	running	0	10490
06.05.26 14:30:06	26399	cancelled	0	10490
06.05.26 14:29:50	26384	running	0	10485
06.05.26 14:29:45	26381	running	0	10484
06.05.26 14:29:37	26382	finished	0	10485
06.05.26 14:29:36	26374	finished	0	10482
COUNT	48	48	48	48



Online Filter dashboard examples [2]

Info about Jobs



Online Filter dashboard examples [3]

job info						
datetime ▾	job_id ▾	state ▾	pilot_id ▾	task_id ▾	error_code ▾	exit_code ▾
29.03.26 15:14:38	160046	completed	65	20362	0	0
29.03.26 15:14:36	160046	stage_out	65	20362	0	0
29.03.26 15:14:34	160046	post_processing	65	20362	0	0
29.03.26 15:14:32	160046	finished	65	20362	0	0
29.03.26 15:14:30	160046	running	65	20362	0	0
29.03.26 15:14:28	160046	stage_in	65	20362	0	0
29.03.26 15:14:26	160046	pre_processing	65	20362	0	0
29.03.26 15:14:25	160046	obtained	65	20362	0	0
29.03.26 15:13:55	160043	failed	62	20362	402	0
29.03.26 15:13:55	160041	completed	64	20362	0	0
29.03.26 15:13:53	160041	stage_out	64	20362	0	0
29.03.26 15:13:52	160043	failed	62	20362	402	0
29.03.26 15:13:51	160041	post_processing	64	20362	0	0
29.03.26 15:13:50	160043	running	62	20362	0	0
COUNT	364	364	364	364	364	364

Info about pilots workload

Avg Memory	Avg IOPS
9.29%	9.75M io/s
Active pilots	Avg CPU
5	7.93%

completion		
all jobs	completed jobs	completion
30	28	93.3%
avg time in state		
state	avg time	
obtained	5.62 s	
pre_processing	6.30 s	
stage_in	2.58 s	
running	2.73 s	
finished	2.00 s	
post_processing	2.11 s	
stage_out	2.19 s	

Distributed computing infrastructure monitoring

Visualization of data obtained after tests of the SPD distributed computing infrastructure:

- availability of systems and services
- info about host-certificates
- info about functional tests (dashboard not implemented yet)

Distributed computing infrastructure dashboard example [1]



test results									
timestamp ↑	site	service...	service_name	result	error_type	mes...	endpoint	check_type	response...
05.05.26 12:24	PNPI	SE-XROOTD	PNPI_EOS_xrootd	success		Alive	root://mss3.pnpi.nw.ru:1094	XRootD Stat	0.726
05.05.26 12:24	JINR	SE-XROOTD	JINR_SPD_EOS-XROOTD	success		Alive	root://eos-spd.jinr.ru:1094	XRootD Stat	0.326
05.05.26 12:24	PNPI	SE-WEBDAV	PNPI_EOS_webdav	success		Alive	davs://mss3.pnpi.nw.ru:8000	WebDAV Stat	0.379
05.05.26 12:24	JINR	SE-WEBDAV	JINR_SPD_EOS-WEBDAV	success		Alive	davs://eos-spd.jinr.ru:8000	WebDAV Stat	0.224
05.05.26 12:24	INPBSU	SE-WEBDAV	INPBSU-SPD-STORM	fail	ServiceNotFound	Failed afte	davs://grid02.hep.by:8443	WebDAV Stat	0.437
05.05.26 12:24	JINR	Rucio	Rucio-prod	success		Alive (Ruci	https://spd-rucio.jinr.ru:443	Rucio Client	0.0217
05.05.26 12:24	JINR	PanDA	PanDA-prod	success		Alive	https://vm221-129.jinr.ru:25443	HTTP	0.0200
05.05.26 12:24	JINR	Harvester	SPD-221-126	success		Alive	https://vm221-126.jinr.ru:8080	HTTP	0.0274
05.05.26 12:24	JINR	FTS	SPD-FTS	success		Alive	https://spd-fts.jinr.ru:8446	HTTP	0.0554
05.05.26 12:24	SSAU	CE	SSAU-CE	fail	Network	Failed afte	https://spd.ssau.ru:443	ARC CE	7.20
05.05.26 12:24	SPbSU	CE	SPbSU-CE	fail	Network	Failed afte	https://alice30.grid.pu.ru:443	ARC CE	6.96
05.05.26 12:24	PNPI	CE	PNPI-CE	success		Alive	https://gt2.pnpi.nw.ru:443	ARC CE	0.387
05.05.26 12:24	JINR	CE	JINR-CE-T2	success		Alive	https://lcgce01.jinr.ru:443	ARC CE	0.267
05.05.26 12:24	JINR	CE	JINR-CE-T1	fail	SSL	Failed afte	https://ce03.jinr-t1.ru:443	ARC CE	2.61
05.05.26 12:24	INPBSU	CE	INPBSU-CE	fail	Network	Failed afte	https://grid04.hep.by:443	ARC CE	7.13

Distributed computing infrastructure dashboard example [2]

certificates info							
timestamp	issuer	cert not after	service name	status	days left	service type	endpoint
05.05.26 15:19	JINR	27.10.26 9:56	spd-requests.jinr.ru	VALID	174	manual	https://spd-requests.jinr.ru:443
05.05.26 15:19	JINR	17.04.27 6:05	spd-cric.jinr.ru	VALID	346	manual	https://spd-cric.jinr.ru:443
05.05.26 15:19	JINR	08.10.26 12:26	spd-iam.jinr.ru	VALID	156	manual	https://spd-iam.jinr.ru:443
05.05.26 15:19	RDIG	08.02.27 0:05	PNPI_EOS_webdav	VALID	278	SE-WEBDAV	davs://mss3.pnpi.nw.ru:8000
05.05.26 15:19	RDIG	08.02.27 0:05	PNPI-CE	VALID	278	CE	https://gt2.pnpi.nw.ru:443
05.05.26 15:19	JINR	20.02.27 10:12	JINR_SPD_EOS-WEBDAV	VALID	290	SE-WEBDAV	davs://eos-spd.jinr.ru:8000
05.05.26 15:19	RDIG	26.03.27 9:07	JINR-CE-T2	VALID	324	CE	https://lcgce01.jinr.ru:443
05.05.26 15:19	RDIG	08.08.26 9:32	JINR-CE-T1	VALID	94	CE	https://ce03.jinr-t1.ru:443
05.05.26 15:19	JINR	24.10.26 14:33	INPBSU-SPD-STORM	VALID	172	SE-WEBDAV	davs://grid02.hep.by:8443
05.05.26 15:19	JINR	24.10.26 14:35	INPBSU-CE	VALID	172	CE	https://grid04.hep.by:443
05.05.26 15:19	JINR	17.12.26 12:16	PanDA-prod	VALID	225	PanDA	https://vm221-129.jinr.ru:25443
05.05.26 15:19	JINR	17.04.27 6:11	SPD-221-126	VALID	346	Harvester	https://vm221-126.jinr.ru:8080
05.05.26 15:19	JINR	27.01.27 10:17	SSAU-CE	VALID	266	CE	https://spd.ssau.ru:443
05.05.26 15:19	JINR	16.07.26 21:46	SPbSU-CE	VALID	72	CE	https://alice30.grid.pu.ru:443
05.05.26 15:19	JINR	12.01.27 8:27	SPD-FTS	VALID	251	FTS	https://spd-fts.jinr.ru:8446
05.05.26 15:19	JINR	17.11.26 8:00	Rucio-prod	VALID	195	Rucio	https://spd-rucio.jinr.ru:443

Thank you for your attention!