



Национальный исследовательский ядерный университет «МИФИ»
Институт ядерной физики и технологий
Кафедра физики элементарных частиц №40

Отчет о научно-исследовательской работе студента
на тему:

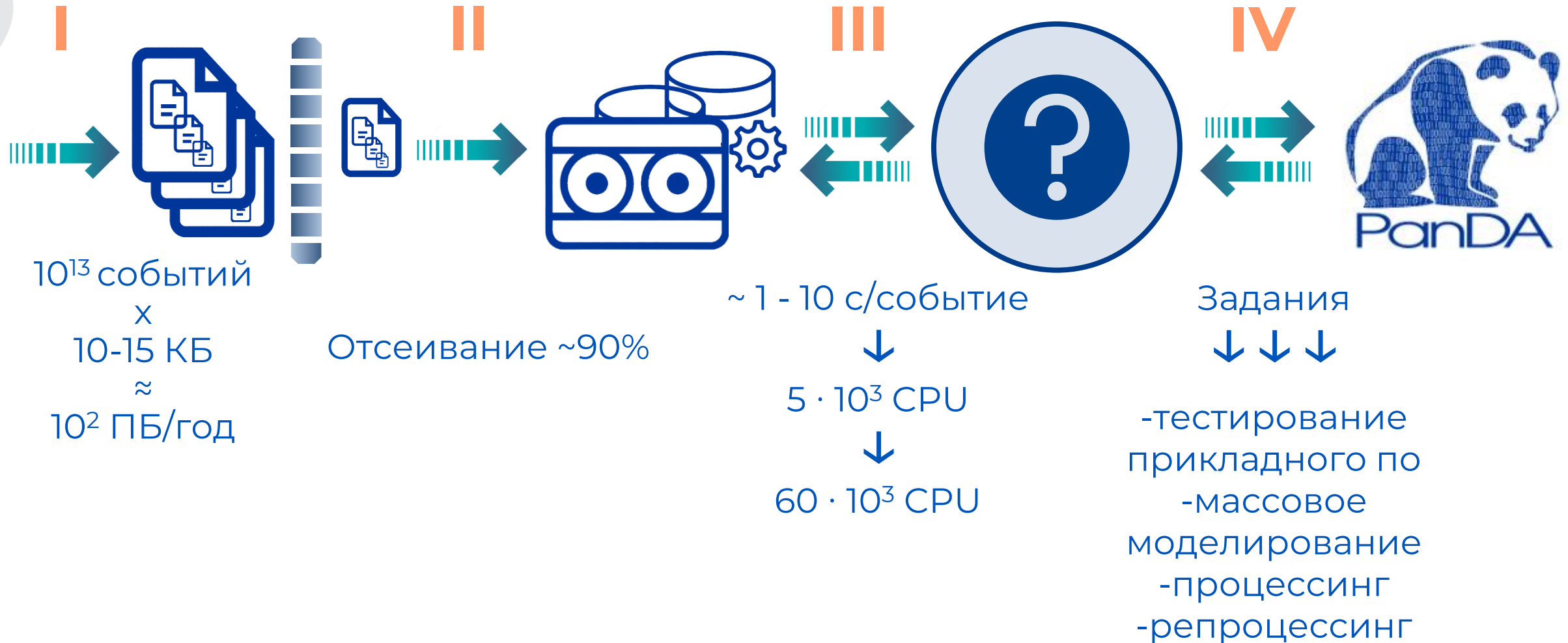
СИСТЕМА УПРАВЛЕНИЯ ПРОЦЕССАМИ ОБРАБОТКИ В ГЕОГРАФИЧЕСКИ РАСПРЕДЕЛЕННОЙ СИСТЕМЕ ОБРАБОТКИ ДАННЫХ SPD

Москва 2025



- Мотивация создания и текущая концепция сервиса
- Проделанная работа
- Текущие задачи
- Дальнейшее расширение функционала и поддержка стабильной работы

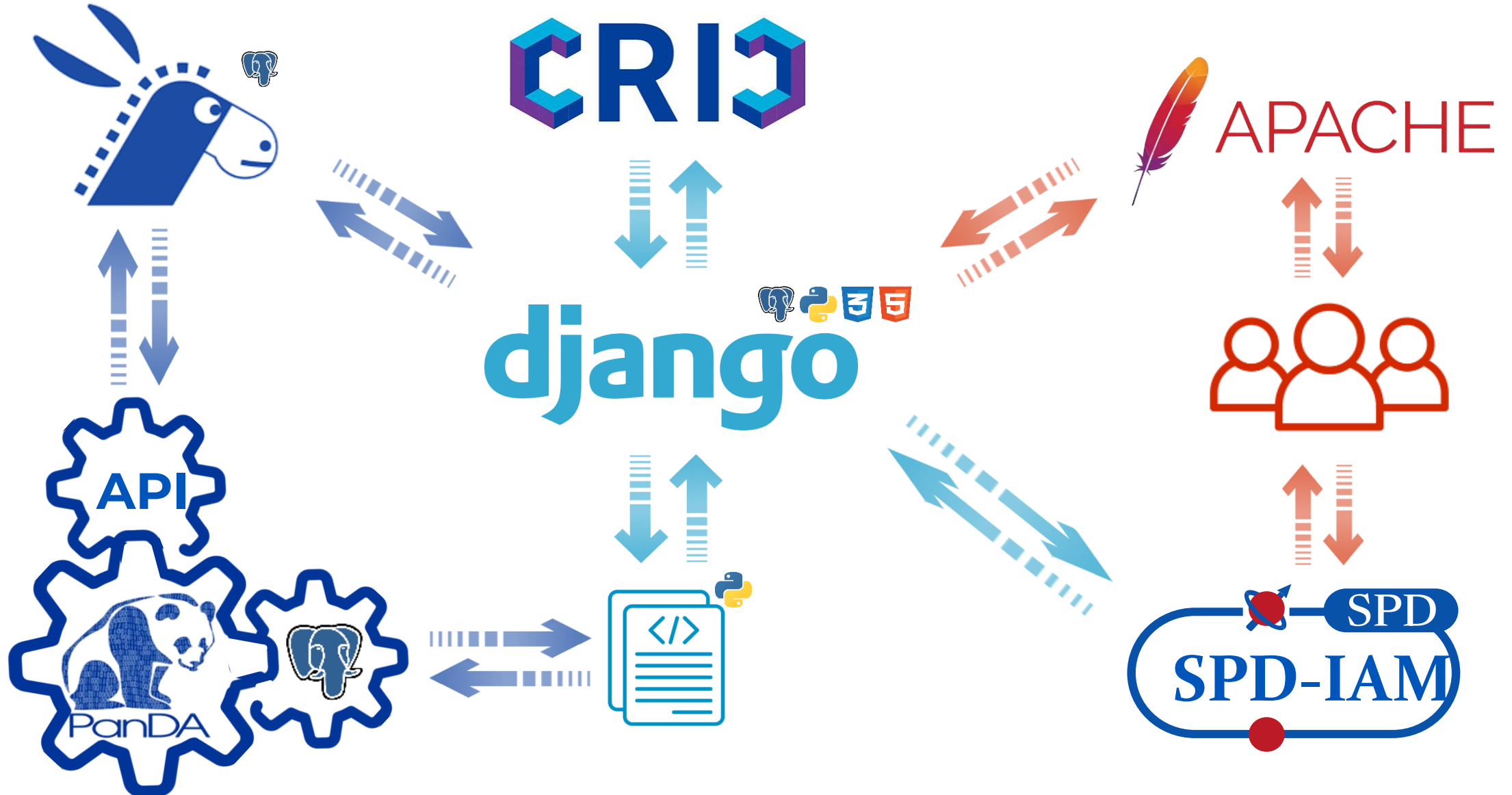
Мотивация создания сервиса



Текущая концепция сервиса



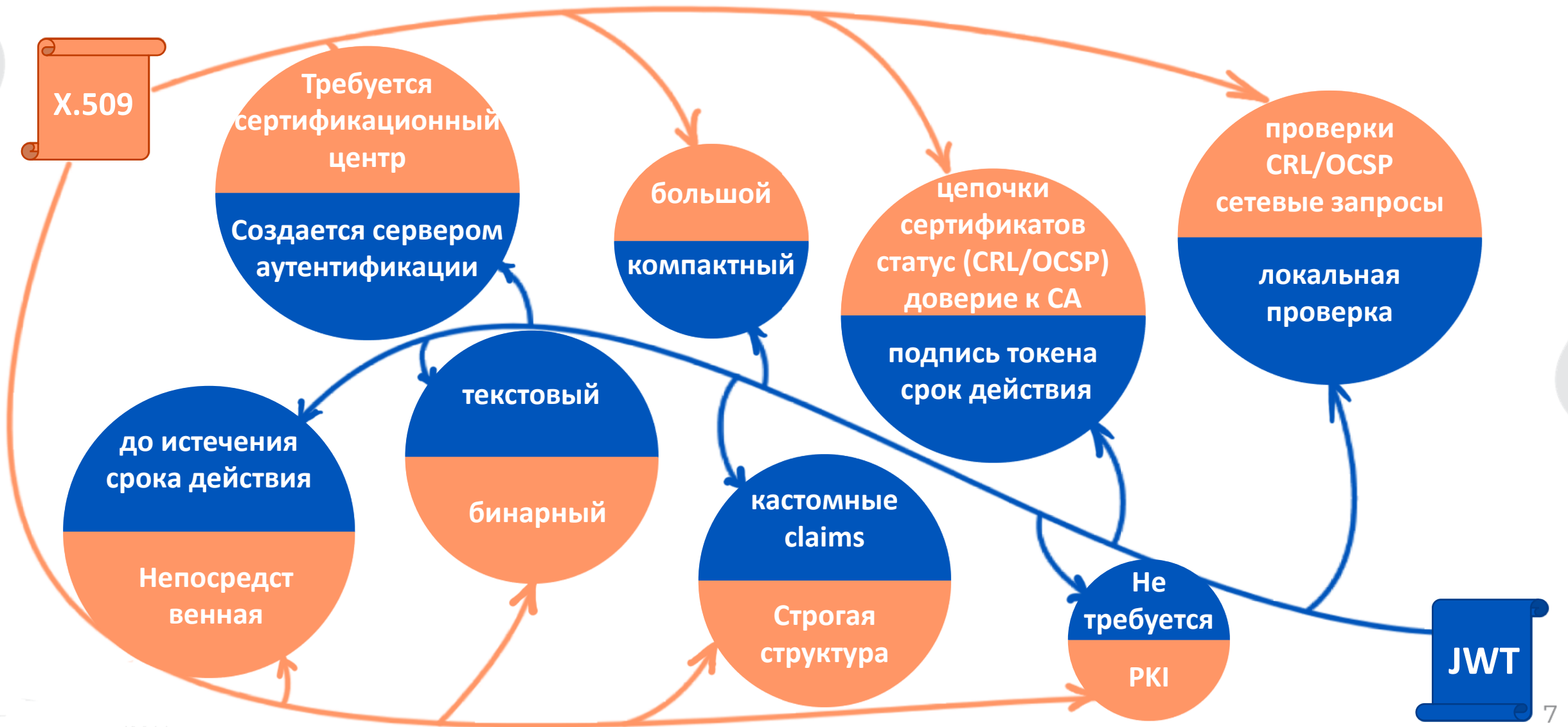
Technology stack



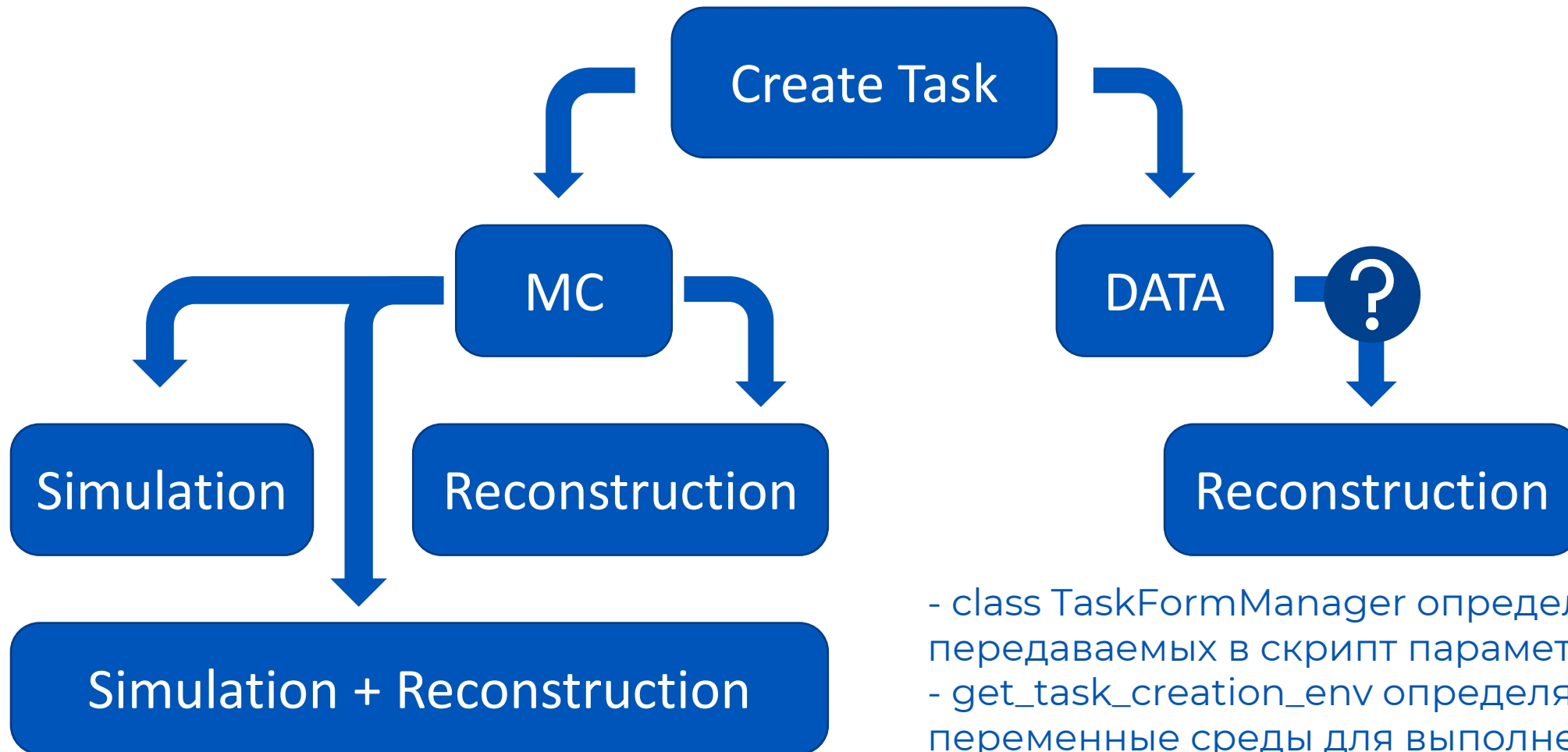
Проделанная работа

- Создан прототип веб-приложения
 - аутентификации пользователя с использованием `spd-iam`
 - создание заданий в соответствии с принятыми конвенциями
 - отслеживание статуса задания
 - настроены декораторы и `middleware`
- Начата реализация автоматизирующих скриптов
- Налажено взаимодействие с PanDA WMS с использованием JWT
- Логирование взаимодействий с PanDA WMS

Проделанная работа



Проделанная работа



- class TaskFormManager определяет набор передаваемых в скрипт параметров
- get_task_creation_env определяет переменные среды для выполнения
- run_send_task_script инициирует выполнение скрипта, ведет логгирование, передает данные, необходимые для цепочек

Текущие задачи и дальнейшие планы

- Реализация механизма клонирования заданий*
- Расширение функционала создания заданий
- Ведение общей статистики
- Мониторинг заданий
- Расширение валидации
- Расширение автоматизации

Заключение

На данном этапе научной работы начато и продолжается освоения ПО, необходимого для реализации системы контроля данных высокого уровня. Были изучены соответствующие системы, реализованные в рамках других мега-сайнс экспериментов, произведено полное развертывание и настройка приложения в рамках облачного сервиса ОИЯИ, имплементированы механизмы коммуникации с прочими сервисами в рамках эксперимента, начат процесс масштабной унификации кода приложения и расширены его внутренние возможности.



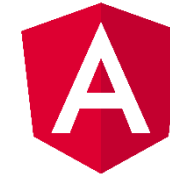
Благодарю за внимание



Анализ технологий для построения системы



django



MVT-архитектура (Model View Template)

- разделение обязанностей
- повторное использование кода
- масштабируемость



ORM (Object Relational Mapper)

- оптимизация часто используемых операций с БД
- экранирование параметров запросов
- автоматическая обработка подключений к БД и управление соединениями
- поддержка миграций



SPA (Single Page Application)

- маршрутизация и навигация между компонентами
- поддержка Lazy Loading для оптимизации загрузки приложений



Dependency Injection



Реализованный admin-интерфейс

Анализ технологий для построения системы



- **Встроенные функции обеспечения безопасности**
 - защита от CSRF (Cross-site Request Forgery)
 - защита от XSS (Cross-site Scripting)
 - HTTPS (HyperText Transfer Protocol Secure)

- **Встроенные функции обеспечения безопасности**
 - аутентификация и управление сессиями пользователей
 - автоматическая обработка подключений к базе данных и управление соединениями
 - защита от SQL-инъекций
 - управление паролями
 - ограничение доступа на уровне представлений

- **Формы и валидация**
 - модуль **ReactiveForms** для создания и валидации форм
 - встроенные механизмы валидации данных на стороне клиента
- **Тестирование**
 - модульное тестирования компонентов и сервисов
 - инструменты для автоматизации тестирования

Анализ технологий для построения системы



Созданные в рамках веб-приложения задания должны быть отправлены на выполнение. PanDA client реализует соответствующий функционал, используемый для взаимодействия с PanDA - системой управления рабочими нагрузками.



Отправка заданий; автоматизация процесса подготовки и запуска заданий на распределённых ресурсах [prun ...]

- inDS: Входной набор данных
- outDS: Выходной набор данных
- exes: Команда для выполнения
- bexes: Создание бинарных файлов
- nFiles: Количество файлов для обработки
- writeInputToTxt: Запись входных файлов в текстовый файл
- outputs: Спецификация выходных файлов
- containerImage: Образ контейнера для выполнения
- architecture: Запуск на GPU
- match: Фильтр входных данных
- nJobs: количество задач в задании
- nEventsPerJob: количество событий в задаче



Интеграции с Athena [pathena ...]

- inDS: Входной набор данных
- outDS: Выходной набор данных
- trf: Трансформационный скрипт
- containerImage: Образ контейнера для выполнения
- nJobs: количество задач в задании
- nEventsPerJob: количество событий в задаче



Интерфейс для управления заданиями и мониторинга их статуса [pbook ...]

- kill
- finish
- retry
- show
- showl

Introduction to JWT

JWT is a compact and secure way to transmit claims between two parties, it consists of Header, Payload, and Signature.

```
eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.  
eyJzdWIiOiIxMjM0NTY3ODkwIiwibmFtZSI6IkpvaG4gRG9lIiwiaWF0IjoiOnRydWV9.  
TjVA95OrM7E2cBab30RMHrHDcEfxjoYZgeFONFh7HgQ
```



Base64
decode



gibberish?

Structure of a JSON Web Token (JWT)

- **JWT** standardize **certain claims** that are useful in the context of some common operations.
- Another key aspect of JWT is **security**:
- **Signing** them, using JSON Web Signatures (JWS, RFC 7515)
- **Encrypting** them, using JSON Web Encryption (JWE, RFC 7516)



```
{  
  "alg": "HS256",  
  "typ": "JWT"  
}
```

```
{  
  "sub": "1234567890",  
  "name": "George White",  
  "admin": true,  
  "iat": 1516239022  
}
```

```
Base64URLSafe(  
  HMACSHA256(<header>.  
  <payload>, <secret key>  
  )  
)
```

Practical Applications of JWT

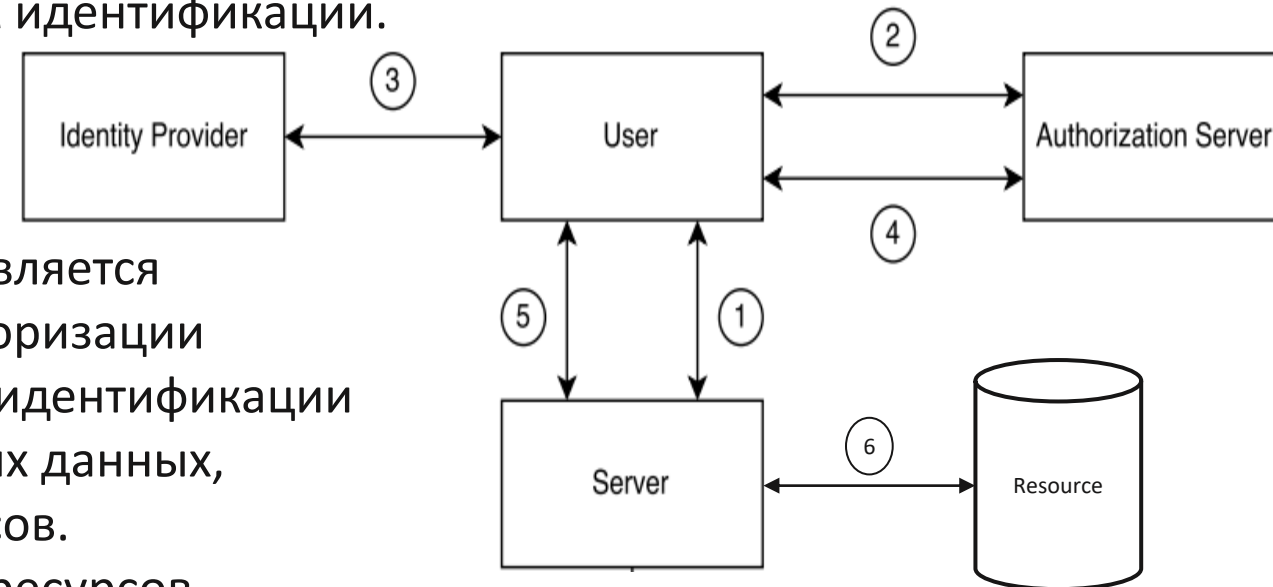
Пользователь пытается получить доступ к ресурсу, который контролируется сервером. У пользователя нет подходящих учетных данных для доступа к ресурсу, поэтому сервер перенаправляет его на сервер авторизации. Сервер авторизации настроен таким образом, чтобы позволить пользователю войти, используя учетные данные, которые управляются провайдером идентификации.

Пользователь перенаправляется сервером авторизации на страницу входа провайдера идентификации.

Пользователь успешно выполняет вход и перенаправляется обратно на сервер авторизации. Сервер авторизации использует предоставленные провайдером идентификации данные для получения необходимых учетных данных, чтобы предоставить доступ к серверу ресурсов.

Пользователь перенаправляется обратно на сервер ресурсов через сервер авторизации. Запрос теперь содержит необходимые учетные данные для доступа к ресурсу.

Пользователь получает доступ к ресурсу.



RSA (Rivest-Shamir-Adleman) — один из наиболее известных алгоритмов, где публичный и приватный ключи связаны через числа.

1. Генерация ключей:

- Выбираются два больших простых числа p и q .
- Вычисляется их произведение $n = p \cdot q$.
- Вычисляется $\phi(n) = (p - 1)(q - 1)$ (функция Эйлера).

2. Создание публичного ключа:

- Выбирается число e (показатель шифрования), которое взаимно просто с $\phi(n)$.
Обычно $e = 65537$.

3. Создание приватного ключа:

- Вычисляется d (показатель расшифровки), которое является мультипликативно обратным к e по модулю $\phi(n)$:

$$d \cdot e \pmod{\phi(n)} = 1$$

$$L(n) = \exp \left(\left(\frac{64}{9} \right)^{1/3} (\ln n)^{1/3} (\ln \ln n)^{2/3} \right)$$

4. Полученная пара ключей:

- Публичный ключ: (n, e)
- Приватный ключ: (n, d)

2.2 Как они работают вместе?

- Сообщение M шифруется с помощью публичного ключа:

$$C = M^e \pmod{n}$$

Здесь C — зашифрованное сообщение.

- Зашифрованное сообщение C расшифровывается приватным ключом:

$$M = C^d \pmod{n}$$

Аббревиатуры, назначения и всякие полезности

ADC - ATLAS Distributed Computing

AGIS - ATLAS Grid Information System

AMI - ATLAS Meta-data Interface

ARC - Advanced Resource Connector (Nordugrid middleware)

ATLAS - a toroidal apparatus for LHC

CE - Computing Element

DB - database

DDM - Distributed Data Management

DEFT - Database Engine for Tasks - the second layer of ProdSys2

EGI/EGEE - European Grid Initiative (one of three LHC grid flavours)

JEDI - Jobs Execution and Definition Interface - the third layer of ProdSys2

HPC - High-Performance Computing (=supercomputers)

HTC - High-Throughput Computing (=Grid)

NDGF, Nordugrid - one of three LHC grid flavors

OSG - Open Science Grid (one of three LHC grid flavours)

PanDA - Production and Distributed Analysis WMS - the bottom layer of ProdSys2

ProdSys2 - the second generation of ATLAS Production System

pyAMI - python AMI (I/F to access AMI database(s))

Rucio - The second generation of ATLAS DDM

SE - Storage Element

TID - Task ID, unique task identifier

WFM - Workflow Management

WMS - Workload Management System

Аббревиатуры, назначения и всякие полезности

ProdSys -production system

Rucio – data management system, каталогизация, хранение метаданных о файлах

Задание->задачи-> события логи (STDERR STDOUT задач) -----PanDA и Rucio работает с распараллеливанием рассылкой по пилотам(вычислительным кластерам) а потом сборкой

JEDI(job execution and definition interface) – позволяет пользователю перейти от работы с задачами к работе с заданиями

AGIS-> CRIC описание всех ресурсов, полная топология системы

Grouping tier	Field	Description	Example
0	[YEAR]	Main Scope - the year of data production	2050
1	[MC DATA]	Real data or simulated data	DATA
2	[energy][polarization]		250LT
3	[desc]	Short name of physics aim	minbias
4	[RunNumber]	Run number for DATA, ID for MC	27189
5	[data type]	EVGEN, SIMUL, RECO....	RAW
6	[<u>DatasetUID</u>]	unique ID of the dataset	636763fd78df7d
7	[Version]	for reprocessing	0

с какими центрами дружим: хранение /ЦП

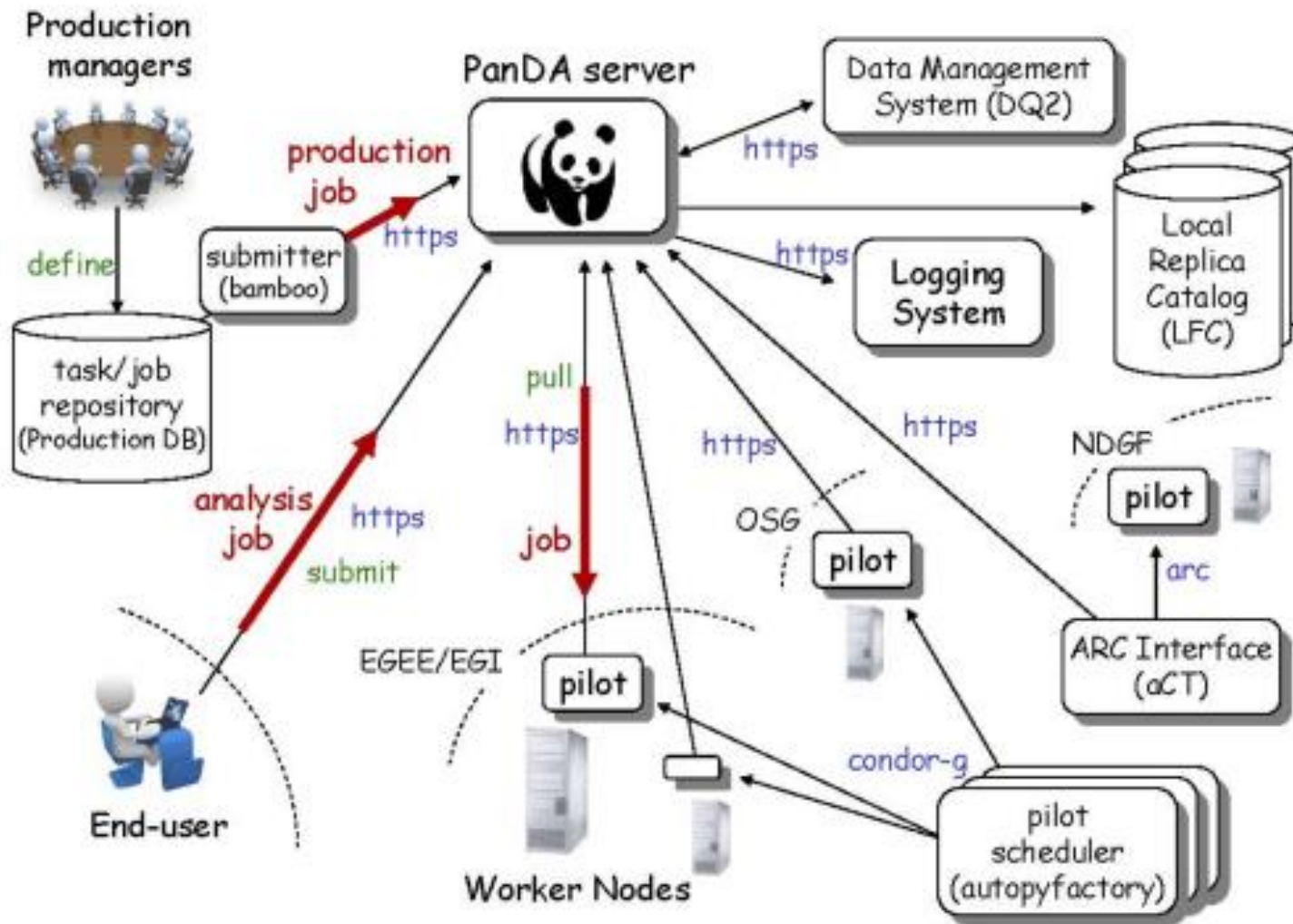
- Первый МИФИческий
- ПИЯФ Гатчина 200 ТБ/6 к
- НИИЯП БГУ / 400
- СПбГУ
- рассчитываем на НИЦ Курчатовский Институт
- Локально >5ПБ/1.5 к

Аббревиатуры, назначения и всякие ПОЛЕЗНОСТИ

Common task definition errors

- “The task is rejected because of inconsistency. XXX”
 - nEventsPerJob of parent is not equal to specified nEventsPerInputFile
 - To fix: nEventsPerInputFile should be changed
 - nEventsPerJob is not divisible by nEventsPerInputFile without remainder AND nEventsPerInputFile is not divisible by nEventsPerJob without remainder
 - To fix: nEventsPerInputFile or nEventsPerJob should be changed
- “Input data list is empty”
 - No inputs or JO are provided
 - To fix: check step/request parameters
- “Invalid request parameter: DSID”, “Invalid request parameter: Energy”, “Suitable XXX candidate not found in evgeninputfiles.csv”
 - Wrong energy provided for Evgen step of there is no necessary line in the file ‘share/evgeninputfiles.csv’ on cvmfs
 - To fix: check request parameter or update/fix ‘evgeninputfiles.csv’ file
- “Output data are missing”, “These requested outputs are not defined properly: ZZZ”
 - The task cannot be defined without output but TRF does not support for some of specified output formats (“XXX.YYY.ZZZ”)
 - To fix: check step parameters (output formats), check TRF (asetup ..., *_tf.py -dumpgargs)
- “Number of events to be processed is mandatory when task has no input”
 - The task input is not properly defined. Missing dataset/container or JO
 - To fix: check input parameters of step (“Dataset”)
- “[Check duplicates] The task is rejected”, “No more input files”
 - All available events are already processed with given configuration
 - To fix: change step configuration (project, tags, formats, etc.)

PanDA



Task ID	Task name ↑ ↓	Parent ID	Creator	Status	Done jobs	Default/Current priority	Total events	Submit time ↑ ↓	Start time ↑ ↓
269	2024.27GeV.test-MB.2st.DSSD.reco	269	Artem Petrosyan	finished	4764	900/900	0	12:34, 13 Jan 2025	12:35, 13 Jan 2025
268	2024.27GeV.dev-MB.2st.DSSD.00001.SIMUL.1	268	Artem Petrosyan	finished	4825	900/900	0	11:42, 26 Dec 2024	11:49, 26 Dec 2024
266	2024.27GeV.test-MB.2st.DSSD.reco	266	Artem Petrosyan	finished	4838	900/900	0	12:26, 13 Dec 2024	12:27, 13 Dec 2024
265	2024.27GeV.test-MB.2st.DSSD.reco	265	Artem Petrosyan	finished	31363	900/900	0	12:31, 06 Dec 2024	12:33, 06 Dec 2024
263	2024.27GeV.test-MB.2st.DSSD.00007.SIMUL.1	263	Artem Petrosyan	finished	34598	900/900	0	08:19, 03 Dec 2024	08:27, 03 Dec 2024
262	2024.27GeV.test-MB.2st.DSSD.00006.SIMUL.1	262	Artem Petrosyan	finished	37677	900/900	0	12:53, 28 Nov 2024	13:01, 28 Nov 2024
261	2024.27GeV.test-MB.2st.DSSD.00005.SIMUL.1	261	Artem Petrosyan	finished	4860	900/900	0	08:38, 28 Nov 2024	08:45, 28 Nov 2024
259	2024.27GeV.test-MB.2st.DSSD.00003.SIMUL.1	259	Artem Petrosyan	finished	4940	900/900	0	19:29, 27 Nov 2024	19:37, 27 Nov 2024
249	2024.27GeV.test-MB.2st.DSSD.reco	249	Artem Petrosyan	finished	4977	900/900	0	09:12, 21 Nov 2024	09:13, 21 Nov 2024
230	2024.27GeV.test-MB.2st.DSSD.reco	230	Artem Petrosyan	finished	4934	900/900	0	15:48, 12 Nov 2024	15:49, 12 Nov 2024

Task ID	269
Parent task ID	269
Chain task ID	None
Name	2024.27GeV.test-MB.2st.DSSD.reco
Type	managed
VO	spd.nica.jinr
Creator	Artem Petrosyan
State status	finished
Total events	0
Total jobs done	4764
Total requested jobs	6430
Priority	900
Current priority	900
Submit time	12:34 January 13, 2025
Start Time	12:35 January 13, 2025
Time stamp	22:54 January 17, 2025
JEDI parameters	<pre> {"nFilesPerJob": 1, "nEventsPerJob": 4000, "noInput": false, "taskName": "2024.27GeV.test-MB.2st.DSSD.reco", "userName": "Artem Petrosyan", "vo": "spd.nica.jinr", "taskPriority": 900, "architecture": "x86_64", "transUses": "A", "transHome": null, "transPath": "http://159.93.221.125:8080/spd_simu_VA_transform.sh", "processingType": "step2", "prodSourceLabel": "managed", "taskType": "test", "workingGroup": "spd.nica.jinr", "cloud": "JINR", "ramCount": 1900, "log": {"dataset": "test.2024.MC.27GeV.dev- minbias.00001.RECO.1.log", "type": "template", "param_type": "log", "token": "DATADISK", "value": "test.2024.MC.27GeV.dev-minbias.00001.RECO.1.\${SN}.log.tgz"}, "jobParameters": [{"type": "constant", "value": "singularity run --bind /cvmfs/spd.jinr.ru/production/MC/2024.27GeV.dev-MB.2st.DSSD:/prod -H ./:WORKDIR /cvmfs/spd.jinr.ru/images/spdroot-dev-4.1.7.sif spdroot.py -b -q '/prod/reco.C(4000, ", {"type": "constant", "value": "\\\"", {"type": "template", "param_type": "input", "value": "\${INS}", "dataset": "test:test.2024.MC.27GeV.dev-minbias.00001.SIMUL.1.S"}, {"type": "constant", "value": "\\\"", {"type": "constant", "value": "\\\"", {"type": "template", "param_type": "input", "value": "\${INP}", "dataset": "test:test.2024.MC.27GeV.dev-minbias.00001.SIMUL.1.P"}, {"type": "constant", "value": "\\\"", {"type": "constant", "value": "\\\"", {"type": "template", "param_type": "output", "token": "DATADISK", "value": "r.test.2024.MC.27GeV.dev-minbias.00001.RECO.1.\${SN/P}.root", "dataset": "test.2024.MC.27GeV.dev-minbias.00001.RECO.1.R"}, {"type": "constant", "value": "\\\"", {"type": "template", "value": "\${RNDMSEED}", "param_type": "number"}, {"type": "constant", "value": ")"}]} </pre>

```
● [monakov@10-220-18-146 clown]$ ll /var/log/highLevelService/taskSender/
total 4
-rw-rw-rw--+ 1 apache apache 3278 Sep 24 15:25 create_send_2024-09-24.log
-rw-rw-rw--+ 1 apache apache    0 Sep 30 02:54 create_send_2024-09-30.log
● [monakov@10-220-18-146 clown]$ cat /var/log/highLevelService/taskSender/create_send_2024-09-24.log
2024-09-24 02:58:06,812 Sending task named '995.Data.9584.polarizationType1.5984895.48954.EVGEN.84948.94498'
  dataset name : '84948'
  event(s) per job : 15
  no input flag is : True
  task creator : Monakov Nikita Glebovich
  task VO : spd.nica.jinr
  task priority : 900
  architecture : x86_64
  cloud : RU
  site : JINR_SPD_TEST
  workingGroup : spd.nica.jinr

2024-09-24 02:58:06,813 /home/monakov/highLevelService/toma/sendTask.py
2024-09-24 02:59:05,485 Sending task named '0.Data.0.polarizationType1.00.0.EVGEN.0.0'
  dataset name : '0'
  event(s) per job : 123
  no input flag is : True
  task creator : Monakov Nikita Glebovich
  task VO : spd.nica.jinr
  task priority : 900
  architecture : x86_64
  cloud : RU
  site : JINR_SPD_TEST
  workingGroup : spd.nica.jinr
```


Task Creation

Year:	<input type="text" value="2024"/>
Data source:	<input type="text" value="MC"/>
Energy [GeV]:	<input type="text" value="1000"/>
Polarization:	<input type="text" value="polarizationType1"/>
Description:	<input type="text" value="Some description"/>
Run number:	<input type="text" value="59"/>
Data type:	<input type="text" value="SIMULATION"/>
Dataset name:	<input type="text" value="myDataset.123456.name"/>
Version:	<input type="text" value="1"/>

Create task

Task Creation

Task name:	<input type="text" value="2024.MC.1000.polarizationType1.Some description.59.SIMULATION.m"/>
Total events:	<input type="text" value="1000000"/>
Total jobs:	<input type="text" value="1000"/>

Create task