

# **Рабочая группа по анализу и обеспечению качества работы СКЦ**

Июнь 2021

# Скрипты для автоматической реакции

## Прологи/эпilogи блокируют узел

- Не прошла проверка
  - Отвалился GPU
  - Отвалилась память
  - Провисла производительность теста
- Не получилось загрузить драйвер
- Не получилось выполнить чистку / сменить режим

# Скрипты для автоматической реакции

## Нестандартные квоты — блокирование пользователей

- Panasas
- Узло-часы

# Скрипты для автоматической реакции

## Реакции (оповещения) повешения графаны:

- Высокая загрузка служебного сервера
- Мало места на дисках
- Много сбойных узлов
- Большой скачок в статусах задач
- Температура растёт

# Скрипты для автоматической реакции

## Проверка состояния сервисов и т. п. (оповещение)

- Тома люстры
- Opensm
- Slurmctld / slurmdbd
- Серверы лицензий
- Трапы от оборудования

**Оповещения — пока «глупые», о смене состояния или событии (trap, etc)**

# Скрипты для автоматической реакции

## Правила для оценки эффективности приложений и СК в целом:

- Задача зависла
- Очень странное поведение задачи
- ...

**Пока только простые повешения**

# Почему не используете больше?

Требуется **вручную проверить**, верно ли сработал критерий (пример: правила определения аномальных задач)

Понятно, что проблема есть, но неясно, как ее **автоматом решить** (пример: аномальные запуски пакетов);

# Почему не используете больше?

Все понятно, **нужно реализовать** (пример: «умный» перезапуск агентов мониторинга)

Есть условия, при которых **автомат не должен работать** (ручной режим, подозрение на другие причины проблемы и т. п.), и их не всегда получается формализовать, но над этим работаем



# Что хотелось бы автоматизировать в первую очередь?

- Оповещение через мессенджеры пользователей, что их текущая **задача зависла**/работает не корректно (почти все есть для этого)
- **Перезапуск** служебного оборудования (свичи, сервера и т.д.) или узлов, если недоступны с последующим **прогоном тестов**
- Перезапуск **сервера люстры** при некоторых проблемах
- **Автоотключение** при резком перегреве, нехватке энергии (нет острой актуальности)

# Какого ПО вам больше всего не хватает?

- ПО для аккуратной **оценки эффективности** приложения (более сложными методами, чем сейчас) и для сравнения с похожими по смыслу приложениями
- ПО для **анализа логов**:
  - **автоматизированной** оценки первопричин сбоев
  - **ручного** анализа логов (категории, подстроки и т.п.)
- ПО для проведения **тестов** (тестируем nhs)
- Работающие аналоги **Easybuild, Spack!**

# Что из этого можно было бы сделать отчуждаемым?

- Что реализуемо в принципе, то и можно (нужно)
  - Скрипты оповещения и блокировки
  - Тесты
  - Система для поиска аномалий
  - ...