

# Вычислительный кластер CL120

Вычислительный кластер проектировался и создавался параллельно с проектным, строительным и монтажно-наладочным процессами компьютерного класса в аудитории №120 в 2024 году. В январе 2025 введен в эксплуатацию.

Было начато реальное использование кластерных параллельных вычислений, освоены классическое и нейросетевое машинное обучение, вошли в арсенал современные языки программирования.

Комплекс получил наименование CL120.

## Он включает в себя

- 35 вычислительных узлов, в качестве которых используются обычные персональные компьютеры;
- 2 коммутатора 24 / 1 Gbit;
- 1 сервер 16 ядер / 128 Gb.

## Программное обеспечение

- Операционная система GNU / Linux Debian 12;
- Фреймворк для параллельной обработки данных и машинного обучения Apache Spark 3.5.1;
- СУБД PostgreSQL 17.

## Языки программирования

- Python 3.11.2;
- PL / pgSQL;
- Другие программные средства.

Все оборудование класса используется в учебное время, согласно расписанию, как обычный учебный компьютерный класс, кроме тех случаев, когда занятие предполагает использование параллельных вычислений.

Во внеурочное время персональные компьютеры перезагружаются в операционную систему GNU / Linux, после чего кластер готов к работе. Впрочем, и в этом случае компьютеры класса могут использоваться, как обычные персоналки - все кластерные вычисления используют ресурсы ПК в фоновом режиме.

Другими словами под Linux класс может быть использован в учебных целях, как с использованием вычислительного кластера студентами, так и с использованием традиционных, однопользовательских программных средств.

## Функции сервера

- распараллеливание выполняемой задачи;
- хранение данных;
- распределение вычислительной нагрузки и данных между узлами;
- объединение результатов полученных от узлов;
- самостоятельная работа над задачей в период, когда узлы (персональные компьютеры) заняты или выключены.

## Вывод

Университет впервые получил вычислительный кластер, способный в режиме параллельных вычислений решать статистические и иные задачи на больших данных, с использованием как классического машинного обучения, так и глубокого обучения на нейросетях.

## Годовой баланс времени

- Учебное время = 8 месяцев \* 20 дней \* 8 часов = 1280 часов в год.
- Полный годовой баланс = 365 \* 7 \* 24 = 61320 часов в год.
- Ресурс вычислителя 61320 - 1280 = 60040 часов в год или 98% годового баланса.

Разумеется, это теоретические / умозрительные выкладки. Машину нужно обслуживать, готовить к каждому большому проекту, небольшие и учебные проекты могут крутиться параллельно, сразу группами, в учебное и не учебное время, поэтому реальный ресурс CL120 можно будет оценить не менее чем через несколько месяцев или лет эксплуатации.

В презентации есть несколько примеров использования вычислительного кластера CL120.

Приглашаем всех, кому нужны исследования больших данных воспользоваться услугами Лаборатории ЭВТ им.Чепоского.

**Мы помогаем произвести:**

- декомпозицию вашей задачи;
- собрать и пред-обработать данные;
- создать и обучить вашу модель;
- внедрить ее в приложение.