

ОПИСАНИЕ И УСЛОВИЯ ПРЕДОСТАВЛЕНИЯ УСЛУГ

УСЛУГА GPU ACCELERATED SERVERS

1. НАИМЕНОВАНИЕ УСЛУГИ

- 1.1. Наименование Услуги: «GPU Accelerated Servers».
- 1.2. Настоящий документ содержит описание состава Услуги, ее базовой функциональности, возможных сопутствующих и дополнительных услуг, общего порядка подключения, изменения и отключения Услуги, условий предоставления и ограничений.

2. ИНФОРМАЦИЯ ОБ УСЛУГЕ

- 2.1. Краткое описание Услуги.
GPU Accelerated Servers — облачные GPU-серверы на базе графических процессоров NVIDIA® Tesla® V100. Они подходят для сценариев, требующих одновременных высокопроизводительных вычислений в реальном времени, таких как задачи искусственного интеллекта и глубокого обучения, автоматизация научных и инженерных расчетов, 3D-анимация.
- 2.2. Функциональные особенности:
 - Высокопроизводительные вычисления (High Performance Computing, HPC).
 - Поддержка параллельных вычислений NVIDIA CUDA и распространенных фреймворков глубокого обучения, таких как TensorFlow, Caffe, PyTorch и MXNet.
 - Высокопроизводительное хранилище данных.
 - Высокопроизводительная сеть.
 - Поддержка GPU Direct.
- 2.3. Обеспечение защиты инфраструктуры:
Защита инфраструктуры облачной платформы обеспечивается на следующих уровнях:
 - на физическом уровне;
 - на сетевом уровне;
 - на инфраструктурном.

3. БАЗОВАЯ ФУНКЦИОНАЛЬНОСТЬ И МЕТРИКИ УСЛУГИ

- 3.1. Метрики и базовая функциональность GPU Accelerated Servers описана в Приложении 7.40.Н к Договору.

4. ТАРИФИКАЦИЯ УСЛУГИ

- 4.1. Для данной Услуги используется Динамическая тарификация (Pay as you go). Динамическая тарификация предполагает оплату пула ресурсов по факту их потребления Заказчиком в течение Отчетного периода.
- 4.2. Окончательная стоимость Услуги в Отчётном периоде формируется на основании объёма потребленных ресурсов и в соответствии с тарифом, установленным в Приложении 7.40.Н к Договору.

- 4.3. Динамическая тарификация, осуществляющаяся в почасовом порядке (из расчета стоимости 1 (одного) часа), начиная с первой минуты использования. Неполные часы использования Услуги, начиная с первой минуты, округляются до полного часа пользования Услугой.
- 4.4. Данные (отчёты об использовании) по Динамической тарификации Заказчик может запросить у менеджера по Договору.

5. ИНЫЕ УСЛОВИЯ, ПРИМЕНИМЫЕ К УСЛУГЕ

- 5.1. Подключение/изменение/удаление Услуги возможно посредством совершения действий в Консоли Управления SberCloud.Advanced.
- 5.2. Использование Услуги подразумевает пост оплатный порядок расчетов.
- 5.3. Оплата за использование услуги осуществляется в безналичном формате на основании выставленного Исполнителем счета.