

Версия: 11Ф21 от 11.02.2021г.

Условия оказания услуги «Managed Kubernetes»

Оглавление

Общие положения	3
Состав услуги	3
Порядок внесения изменений в Условия оказания услуги	4
Типы узлов и кластеров	4
Тарифные планы	6
Общая информация	6
Поддержка расширенных компонентов	7
Рабочее время Исполнителя	7
Техническая поддержка	7
Поддерживаемые варианты размещения	9
Обновление кластеров	9
Типы обновлений	10
Каналы обновлений	10
Параметры обновлений кластера	12
Соглашение об уровне сервиса (SLA)	13
Типы доступности	14
Synthetic Availability	14
Nginx Availability	15
Node Group Availability	15
Control Plane Availability	15
M&A Availability	16
Extensions Availability	16
Общие условия SLA для всех типов доступности	16
Гарантируемый уровень сервиса	17
Штрафы за нарушение уровня сервиса	18
Стоимость и порядок оплаты	18
Стоимость обслуживания кластера	18
Стоимость обслуживания узлов в кластере	19
Расчет стоимости и порядок оплаты	20
Дополнительные услуги	20

1. Общие положения

1.1. Состав услуги

В настоящих Условиях оказания услуги «Managed Kubernetes» (Условиях оказания услуги), утвержденных приказом АО «Флант» №01-мк от «13» октября 2020 г. и опубликованных в сети Интернет по адресу <https://flant.ru/services/managed-kubernetes-as-a-service/terms.pdf>, приведена информация, раскрывающая по существу назначение и состав Услуги Managed Kubernetes (далее – Услуга), представляющей собой совокупность технических услуг, оказываемых Исполнителем непрерывно в формате абонентского обслуживания, включающих в том числе следующие:

- Инсталляцию кластера(ов) Kubernetes (<https://kubernetes.io/>) в соответствии с заданными в Заказе параметрами на вычислительной инфраструктуре, предоставляемой Заказчиком и соответствующей требованиям настоящих Условий оказания услуги.
- Инсталляцию вспомогательных программных компонентов для кластеров Kubernetes, имеющих глубокую интеграцию с устанавливаемыми кластерами, с целью обеспечения полностью подготовленной для эксплуатации приложений и различного программного обеспечения Заказчика (далее — ПО) инфраструктурной среды.
- Постинсталляционное техническое сопровождение кластеров и вспомогательных программных компонентов в режиме 24x7x365, включая регулярное обновление до актуальных версий кластеров и вспомогательных компонентов, разрабатываемых Исполнителем, включая компоненты систем мониторинга и агрегации параметров жизнедеятельности кластера.
- Круглосуточный мониторинг работоспособности кластеров Kubernetes и вспомогательных компонентов на базе Prometheus и Grafana, а также реакция в соответствии с правилами, определенными в Регламенте, на возникающие сбои и ошибки в их работе.
- Регулярное резервное копирование конфигурации управляющих компонентов кластера в систему хранения Заказчика или Исполнителя (по выбору Заказчика).
- Консультационную поддержку представителей Заказчика — разработчиков программного обеспечения, системных администраторов, инженеров и прочих технических специалистов, взаимодействующих с Исполнителем по поручению Заказчика (далее — Контактные лица).
- Контактных лиц Заказчика в объеме и в соответствии с правилами, предусмотренными настоящими Условиями оказания услуги.
- Своевременное исправление ошибок, разработку, модернизацию и адаптацию вспомогательных компонентов кластеров Kubernetes, проводимые инженерами Исполнителя в соответствии с собственными производственными планами и внутренними техническими регламентами или по взаимному согласованию Сторон.

Исполнитель не гарантирует, что Услуга будет соответствовать требованиям и ожиданиям Заказчика. Заказчик соглашается с тем, что оказание Услуги осуществляется с использованием

программных средств Исполнителя, которые работают по принципу «как есть» («as is»).

Заказчик обязуется учитывать рекомендации Исполнителя по правилам эксплуатации кластеров Kubernetes, обслуживаемых Исполнителем, которые могут направляться Заказчику с целью предотвращения возникновения сбоев в работе обслуживаемых кластеров Kubernetes, при этом если Заказчик систематически игнорирует такие рекомендации, то все возможные риски, связанные с последствиями игнорирования рекомендаций, несет Заказчик.

В любом случае, ответственность Исполнителя при оказании настоящей Услуги ограничена размерами штрафов, установленных настоящими Условиями оказания услуги, Исполнитель не компенсирует убытки Заказчика, возникшие в связи с оказанием Услуги.

Порядок внесения изменений в Условия оказания услуги

Исполнитель вправе вносить изменения в Условия оказания услуги в одностороннем порядке. Обновленные Условия оказания услуги вступают в силу через 30 (тридцать) календарных дней с момента уведомления Заказчика об изменениях.

Для уведомления Исполнитель направляет письмо с информацией о предстоящем изменении посредством электронной почты на email-адреса Полномочных представителей Заказчика, указанных в ССИ, а также на email-адреса всех представителей Заказчика, с установленной в Личном кабинете (<https://my.flant.com/contacts>) Заказчика пометкой «Контакт для связи по операционным вопросам».

Риск неознакомления с новой редакцией Условий оказания услуги несет Заказчик. Потребление Заказчиком Услуги после вступления в силу изменений в вышеуказанном документе считается согласием с его новой редакцией. Заказчик вправе отказаться от принятия изменений, а равно от пользования Услугой, в установленном Договором порядке.

1.2. Типы узлов и кластеров

Поддерживаются следующие **типы узлов кластера**:

- **Cloud** — виртуальная машина (или выделенный сервер), которая создается компонентами Kubernetes в полностью автоматическом режиме. Такие узлы при необходимости внесения изменений могут быть заново созданы (по стратегии RollingUpdate).
- **VM** — виртуальная машина, созданная ранее или создаваемая вручную из-за отсутствия технической возможности автоматизации (отсутствие API, нестандартный/неподдерживаемый API, платформа виртуализации или провайдер согласно п.1.6). Исполнителю должна быть доступна возможность управления (посмотреть состояние виртуальной машины, осуществить перезагрузку, осуществить переустановку ОС) через веб-браузер, по SSH или через обращение к службе поддержки.

- **Hard** — выделенный сервер. Исполнителю должна быть доступна возможность управления (посмотреть состояние сервера, осуществить перезагрузку, осуществить переустановку ОС) через веб-браузер или обращение к службе поддержки.
- **Special** — узел любого типа, который требует особого обращения и не может быть выведен из кластера и/или переустановлен (пересоздан) без подготовительных работ и/или особого согласования. К таким узлам могут, например, относиться stateful (на которых данные сохраняются локально), well-known (адрес которых используется для подключения), session-termination (на которых установлены постоянные соединения с пользователями, что не могут быть корректно прерваны и переустановлены), узлы с поддержкой вычислений на GPU или узлы с выполнением вычислений, которые нельзя прерывать.

В зависимости от архитектуры кластера, окружения, где он эксплуатируется, и типа используемых узлов поддерживаются три **типа кластеров**:

№	Тип кластера	Описание	Требования и ограничения	Возможные типы узлов
1.2.1.	Cloud	Кластер Kubernetes размещается в одном из поддерживаемых облачных провайдеров или частном облаке.	- все узлы кластера (рабочие и управляющие) расположены в одном регионе облачного провайдера.	- Cloud - Special
1.2.2.	NonCloud	Кластер Kubernetes размещается на предварительно созданных виртуальных машинах и/или на выделенных серверах.	- все узлы кластера (рабочие и управляющие) расположены в одном дата-центре; - между всеми узлами кластера есть единая и стабильная L2-сеть; - работа подсистемы хранения (CSI) и возможность создания сервисов с типом LoadBalancer может быть ограничена или отсутствовать.	- VM - Hard - Special
1.2.3.	Hybrid	Кластер Kubernetes размещается на предварительно созданных виртуальных машинах и/или на выделенных серверах, но часть узлов создается автоматически в одном из поддерживаемых облачных провайдеров или частном облаке.	- все узлы кластера (рабочие и управляющие) расположены в одном дата-центре; - между всеми узлами кластера есть единая и стабильная L2-сеть; - работа подсистемы хранения (CSI) и возможность создания сервисов с типом LoadBalancer может быть ограничена или отсутствовать.	- Cloud - VM - Hard - Special

1.2.4.	Backup	Кластер Kubernetes, аналогичный существующему кластеру, служащий горячим или холодным резервом для основного кластера	-	-
1.2.5.	Migrate	Кластер Kubernetes, аналогичный существующему кластеру, созданный в другом ЦОДе или на других мощностях в текущем ЦОДе с целью переезда в него.	-	-

Максимальное количество узлов любого типа (включая мастер-узлы) в одном кластере любого типа — **150** (включая до 3-х мастер-узлов). Это ограничение, прежде всего, обусловлено невозможностью обеспечить стабильный сбор и обработку метрик с необходимым интервалом в кластере с более чем 150 узлами из-за ограничений Prometheus. При необходимости эксплуатации большего количества узлов рекомендуется создать несколько кластеров.

1.3. Тарифные планы

1.3.1. Общая информация

№		Standard	Silver	Gold	Platinum
1.3.1.1.	Назначение	Среды разработки и некритичные продуктивные среды	Продуктивные среды	Важные продуктивные среды	Важные продуктивные среды и премиальная поддержка ваших специалистов
1.3.1.2.	Сравнительный общий уровень доступности	Низкий	Средний	Высокий	
1.3.1.3.	Минимальный срок подписки	Нет			6 месяцев
1.3.1.4.	Доступные версии Kubernetes	Две последние 1.18, 1.19	Четыре последние 1.16, 1.17, 1.18, 1.19		
1.3.1.5.	Максимальное	15 Hard / 15 Special	150		

	количество типов узлов в кластере	/ 30 VM / 60 Cloud	
--	--	--------------------	--

1.3.2. Поддержка расширенных компонентов

№	Услуга	Standard	Silver	Gold	Platinum
1.3.2.1.	Поддержка GPU Специальная версия Docker	Нет	Да	Да	Да
1.3.2.2.	Поддержка Istio	Нет ¹	Нет ¹	Да ²	Да ²

¹ При использовании Istio в продуктивном кластере, находящемся на одном из тарифов, для которых использование Istio доступно, Istio в кластере для разработки становится доступным вне зависимости от его тарифа.

² Доступность услуги ожидается в IV квартале 2020 г. Будет взиматься дополнительная оплата.

1.4. Рабочее время Исполнителя

Рабочее время Исполнителя — время, в которое по умолчанию проводятся:

- плановые работы;
- автоматические обновления;
- значительные обновления.

Рабочим временем Исполнителя является период с 10:00 до 18:00 (GMT+3) в рабочие дни согласно производственному календарю РФ.

Нерабочим временем Исполнителя является все время, которое не является рабочим временем.

1.5. Техническая поддержка

№	Услуга	Standard	Silver	Gold	Platinum
1.5.1.	Время реакции ¹ на инциденты	до 15 минут	до 10 минут	до 5 минут	до 5 минут
1.5.2.	Первоначальная установка и настройка кластера	Да			
1.5.3.	Мониторинг и обработка инцидентов, связанных с работой кластера, 24x7x365	Да			
1.5.4.	Резервное копирование ключевых компонентов платформы (etcd)	Да			
1.5.5.	Базовые консультации по вопросам архитектуры кластера и приложений(в рабочее время)	Да			

1.5.6.	Регулярные конференц-коллы Для синхронизации совместных действий, прогресса по проекту, выявления потребностей, задач и проблем. Проводятся инженером внедрения.	Нет	Да, один раз в месяц	Да, один раз в неделю
1.5.7.	Возможность голосовой консультации с ведущими инженерами в рабочее время	Нет	Да	Да
1.5.8.	Выделенный технический контакт (ведущий инженер)	Нет		Да
1.5.9.	Оказание Услуги при отсутствии прямого доступа Исполнителя по SSH к кластеру Под прямым доступом понимается прямой доступ по SSH по публичному IP-адресу (если необходимо, с ограничением IP-адресов Исполнителя) или доступ через организованный site to site VPN.	Нет		Да
1.5.10.	Прохождение аудита у службы ИБ Включает участие в конференц-коллах с представителями ИБ Заказчика, ответы на вопросы по устройству системы и применяемым практикам, обработку результатов аудита, проведенного ИБ Заказчика.	Нет	Да, до 3 часов	Да, до 5 часов
1.5.11.	Возможность срочного (приоритетного) исправления ошибок в upstream-компонентах Подразумевает исправления в стандартных Open Source-компонентах кластера Kubernetes, в том числе control-plane, cloud-provider'ы, Prometheus и прочих.	Нет		Да, оплачивается отдельно (см. п. 5.2 и 5.4)
1.5.12.	Возможность глубокой кастомизации и доработки функциональных возможностей	Нет		Да, оплачивается отдельно (см. п. 5.2.)

¹ Подробнее термин “Время реакции” в частности и процесс коммуникации в целом описан в Регламенте, см. п. 1 Договора.

1.6. Поддерживаемые варианты размещения

№	Тип инфраструктуры	Автомасштабирование кластера	CSI (дисковая подсистема)	LoadBalancer (сервисы с типом LoadBalancer)	Минимально необходимый тариф
Варианты размещения для Cloud-кластеров					
1.6.1.	Amazon AWS	Да	Да	Да	Standard
1.6.2.	Google Cloud Engine	Да	Да	Да	Standard
1.6.3.	Яндекс.Облако	Да	Да	Да	Standard
1.6.4.	OVH Cloud	Да	Да	Не гарантируется	Standard
1.6.5.	Mail.ru Cloud	Да	Да	Да	Standard
1.6.6.	Selectel Cloud	Да	Да	Да	Standard
1.6.7.	OpenStack	Да	Не гарантируется	Не гарантируется	Silver
1.6.8.	VMware vSphere ¹	Да	Не гарантируется	Не гарантируется	Platinum
Варианты размещения для NonCloud-кластеров					
1.6.9.	Bare-metal	Нет	Не гарантируется	Не гарантируется	Standard
1.6.10.	Виртуальные машины	Нет	Не гарантируется	Не гарантируется	Standard
Варианты размещения для Hybrid-кластеров					
1.6.11.	Размещение в OVH	Да	Да, только на Cloud узлах	Не гарантируется	Standard
1.6.12.	Размещение в Selectel	Да	Да, только на Cloud узлах	Да	Standard

¹ Необходимо наличие доступа к VMWare vSphere API.

2. Обновление кластеров

Кластер Kubernetes и установленные в нем компоненты постоянно развиваются, за время эксплуатации в них обнаруживаются и исправляются функциональные проблемы и проблемы безопасности, появляются новые возможности. Managed Kubernetes от Фланта — это продукт, который тестируется и постоянно обновляется.

2.1. Типы обновлений

В процессе оказания Услуги различаются следующие **типы обновлений**:

- **Значительное обновление** — обновление минорных версий Kubernetes. Например, с 1.15 на 1.16 (но не патч-версий, т.е. не с 1.15.2 на 1.15.3), обновление версии Ingress-контроллера, обновления, требующие вывода узлов из работы (режим maintenance), любые другие обновления, вызывающие простой приложений, запущенных в кластере.
- **Автоматическое обновление** — массовое, автоматическое обновление компонентов кластера, выполняемое под наблюдением инженера.

По своему **составу** обновления подразделяются на:

- **Релиз (release)** — плановое обновление компонентов кластера, связанное с расширением или изменением функционала, а также набором патчей для исправления ошибок.
- **Хотфикс (hotfix)** — внеплановое обновление компонентов кластера с исправлением критических ошибок или иных функциональных проблем, которые уже привели или с высокой вероятностью приведут к нестабильной работе кластера.

Единичные Релизы или Хотфиксы всегда выпускаются в виде автоматических обновлений. Во время значительных обновлений помимо смены версии Kubernetes может быть установлено одновременно несколько релизов и/или hotfix'ов.

2.1.1. Каналы обновлений

Исполнитель проводит всестороннее тестирование новых Релизов и Хотфиксов, включающее в себя юнит- и интеграционное тестирование отдельных частей системы, end-to-end-тестирование кластеров в приближенных к реалистичным условиям, ручную проверку значительных изменений. Однако, следование лучшим практикам SRE требует применения подходов, позволяющих снизить потенциальный негативный эффект проблемы, которая не была выявлена на стадии тестирования.

К кластерам как элементам инфраструктуры предъявляются различные требования. Продуктивный кластер, в отличие от кластера разработки, имеет более высокие требования надежности, на нем желательно снизить частоту обновлений или внесения любых изменений без острой необходимости. Компоненты должны быть максимально протестированы. В то же время низкая скорость изменений компонентов кластера приводит к низкой скорости появления нового функционала.

Чтобы иметь возможность выбрать оптимальный баланс скорости изменений и надежности, предлагается пять **каналов обновлений** компонентов кластера. Обновления, сгруппированные в виде релиз-версий или хотфикс-версий, последовательно применяются к кластерам каждого канала обновлений, от наименее стабильного — Alpha, к наиболее стабильному — RockSolid. Таким

образом, на канале обновлений Alpha — самая высокая скорость появления нового функционала при наибольшей частоте обновлений и наименьшей стабильности. На канале обновлений RockSolid — самая низкая скорость появления нового функционала, наименьшая частота обновлений и наивысшая стабильность. Условно равномерное распределение всех обслуживаемых кластеров Kubernetes по каналам обновлений позволяет в итоге считать, что к моменту применения обновлений к кластерам канала обновлений RockSolid эти обновления уже были последовательно применены не менее чем к 80% других кластеров.

№	Канал обновлений	Описание	Стандартное время обновлений	Требование к версиям канала обновлений	Уведомление о времени и составе изменений
2.1.1.1.	Alpha	Наименее стабильные версии	Без ограничений	Без ограничений	В момент обновления
2.1.1.2.	Beta	Стабильные версии, уже работавшие на части кластеров, обслуживаемых Исполнителем.	Без ограничений	1 сутки успешной работы на канале обновлений Alfa для релизов. 2 часа успешной работы на канале обновлений Alfa для hotfix.	В момент обновления
2.1.1.3.	EarlyAccess	Стабильные версии, работавшие не менее, чем на 10% кластеров, обслуживаемых Исполнителем.	Чт, 12:00-13:00 (GMT+3) для релизов. Без ограничений для hotfix.	1 сутки успешной работы на канале обновлений Beta для релизов. 2 часа успешной работы на канале обновлений Beta для hotfix.	Не позднее, чем за 2 часа до обновления
2.1.1.4.	Stable	Стабильные версии, работавшие не менее, чем на 40% кластеров, обслуживаемых Исполнителем.	Ср, 12:00-13:00 (GMT+3) для релизов. Рабочий день, 14:00-15:00 (GMT+3) для hotfix.	5 суток успешной работы на канале обновлений EarlyAccess для релизов. 1 сутки успешной работы на канале обновлений EarlyAccess для hotfix.	Не позднее, чем за 24 часа до обновления

2.1.1.5.	RockSolid	Стабильные версии, работавшие более, чем на 80% кластеров, обслуживаемых Исполнителем.	Вт, 12:00-13:00 (GMT+3) для релизов . Рабочий день, 14:00-15:00 (GMT+3) для hotfix .	12 суток успешной работы на канале обновлений Stable для релизов . 5 суток успешной работы на канале обновлений Stable для hotfix .	Не позднее, чем за 7 дней до даты обновления
----------	------------------	--	---	--	--

2.2. Параметры обновлений кластера

№	Название	Standard	Silver	Gold	Platinum
Автоматические обновления (п. 2.1)					
2.2.1.	Проведение автоматических обновлений в рабочее время (п.1.4) без возможности выбора времени для окна обновлений			Да	
2.2.2.	Проведение автоматических обновлений в рабочее время (п.1.4) с возможностью выбора времени для окна обновлений	Нет		Да	
2.2.3.	Проведение автоматических обновлений в нерабочее время (п.1.4) с возможностью выбора времени для окна обновлений	Нет		Да, только для канала обновлений RockSolid¹	
Значительные обновления (п. 2.1)					
2.2.4.	Проведение значительных обновлений в рабочее время (п.1.4) с возможностью выбора времени для окна обновлений			Да	
2.2.5.	Проведение значительных обновлений в нерабочее время (п.1.4) с возможностью выбора времени для окна обновлений	Нет		Да, за дополнительную стоимость (см. п. 5.1)	
2.2.6.	Проведение других работ в нерабочее время (п.1.4) с возможностью выбора времени для окна работ	Нет		Да, за дополнительную стоимость (см. п. 5.1)	
Каналы обновлений (п. 2.1.1)					

2.2.7.	Возможность использования канала обновлений EarlyAccess	Да			
2.2.8.	Возможность использования канала обновлений Stable	Нет	Да		
2.2.9.	Возможность использования канала обновлений RockSolid	Нет	Бесплатно для 1/3 количества кластеров ² , если общее количество кластеров более 3-х; в остальных случаях — за дополнительную стоимость (см. п.5.5)	Бесплатно для 1/3 количества кластеров ³ , если общее количество кластеров более 3-х; в остальных случаях — за дополнительную стоимость (см. п.5.5)	Бесплатно, но не более чем на 1/2 количества всех кластеров ⁴

¹ Любые изменения системы потенциально могут приводить к ошибкам, что справедливо и для изменений, проводимых в ночное время. Несмотря на то, что все обновления выполняются под наблюдением инженера, возникновение нештатных ситуаций может потребовать привлечения ведущих инженеров и архитекторов. Учитывая это, обновление в ночное время доступно только для наиболее стабильного канала обновлений – RockSolid, содержащего изменения, гарантированно работающие на более чем 80% кластеров обслуживаемых Исполнителем.

² Учитываются кластеры на тарифе Silver

³ Учитываются кластеры на тарифе Gold

⁴ Учитываются кластеры на тарифе Platinum

3. Соглашение об уровне сервиса (SLA)

Соглашение об уровне сервиса закрепляет условия и способ определения уровня сервиса в момент времени, гарантируемый уровень сервиса (см. п. 3.2) и штрафы за нарушение гарантируемого уровня сервиса (см. п. 3.3).

Уровень сервиса определяется по данным мониторинга не по кластеру в целом, а по отдельным его компонентам, которые сгруппированы в типы доступности (см п. 3.1). Для обеспечения гарантированного уровня сервиса некоторые типы доступности могут требовать включения специального режима работы компонентов (модулей) кластера Kubernetes — режима высокой доступности (далее — режим **High Availability**, режим **HA**).

Особенности и условия работы компонентов в режиме **High Availability**:

- Компонент запускается в двух (или более) экземплярах, чтобы обеспечить возможность выхода из строя одного экземпляра без потери доступности.

- Включение режима требует дополнительных вычислительных ресурсов.
- По умолчанию, если в кластере более одного мастер-узла, режим High Availability включается для всех компонентов.
- При необходимости режим High Availability можно отключать для отдельных компонентов (например, аутентификации, мониторинга и т.п.), снижая затраты на вычислительные ресурсы, однако уровень доступности (Service Level, SL) для таких подсистем не будет гарантирован.
- Ядро кластера (Control Plane) считается работающим в режиме High Availability, если оно запущено не менее чем в трех экземплярах.

3.1. Типы доступности

Для того, чтобы точно определять виды деградаций и степень их влияния на работоспособность сервисов, выделяется несколько типов недоступности, для которых впоследствии возможно определить параметры соглашения об уровне сервиса (SLA).

3.1.1. Synthetic Availability

Определяет доступность (пред)запущенных приложений в кластере, компонентов кластера, не связанных с доступностью Control Plane кластера:

- ранее запущенных приложений (контейнеры продолжают выполняться);
- наличие Kubernetes services (правил в iptables);
- DNS (возможность резолвинга).

SLA для Synthetic Availability действует, только если соблюдены базовые условия:

- узлы кластера продолжают быть доступны;
- есть сетевая связанность между узлами;
- приложение продолжает функционировать корректно.

При этом сохраняют работоспособность приложения, которые соответствуют следующим условиям:

- приложение само кэширует DNS;
- приложение способно работать в условиях отсутствия шедулинга (не работает scheduler) и автоскейлинга (не работает autoscaler), а также в отсутствии возможности добавления/изменения endpoint'ов в сервисах;
- для работоспособности приложения не требуется доступность kube-apiserver.

Не требует включения режима **High Availability** для компонентов.

3.1.2. Nginx Availability

Определяет уровень доступности Ingress Nginx и его составляющих, а это означает, что входящие HTTP-запросы доставляются конечным приложениям.

SLA на Nginx Availability не распространяется на сетевую доступность, которая зависит от инфраструктуры.

Требует включения режима **High Availability** для компонентов.

3.1.3. Node Group Availability

В Managed Kubernetes узлы группируются по функциональному назначению и управляются как связанная группа. Это означает, что все узлы группы имеют одинаковые метаданные, определяемые в параметрах ресурса NodeGroup из полей labels, annotation или taints. Такая группа узлов называется Node Group.

Node Group Availability определяет, что в каждой Node Group одновременно доступно не менее чем N-1 корректно работающих узлов, где N — это минимально доступное количество узлов в NodeGroup, определенное в конфигурации.

SLA на Node Group Availability не распространяется на случаи выхода узла из строя по причине некорректной работы приложения, установленного клиентом (например, из-за утечек памяти).

Не требует включения режима **High Availability** компонентов.

3.1.4. Control Plane Availability

Означает, что доступен деплой в кластер и функционируют процессы самовосстановления. Определяет уровень доступности ключевых управляющих компонентов Kubernetes:

- etcd;
- kube-apiserver;
- kube-controller-manager;
- kube-scheduler;
- kube-dns;
- других компонентов, от которых зависит работа API (admission controller из vertical-pod-autoscaler, cainjector из cert-manager и др.).

Требует включения режима **High Availability** для компонентов.

3.1.5. M&A Availability

Означает, что корректно функционирует мониторинг и автоскейлинг. Определяет уровень доступности Prometheus и функционала автоскейлинга, а именно компонентов и источников данных:

- Prometheus (функционирование TSDB и alerting);
- Trickster;
- prometheus-metrics-adapter;
- vertical-pod-autoscaler;
- horizontal-pod-autoscaler;
- ключевых источников данных, таких как:
 - kube-state-metrics;
 - ingress-nginx;
 - node-exporter'ы и kubelet (с доступных узлов).

Требует включения режима **High Availability** для компонентов.

3.1.6. Extensions Availability

Означает, что дополнительные компоненты Kubernetes доступны. Определяет уровень доступности компонентов, напрямую не влияющих на возможность доставки приложений в среду выполнения и их работу в кластере:

- Dex (возможность SSO-аутентификации в кластере);
- Grafana;
- Dashboard;
- longterm-prometheus;
- cluster-autoscaler;
- cloud-controller-manager;
- machine-controller-manager (возможность заказа узлов);
- openvpn.

Требует включения режима **High Availability** для компонентов.

3.1.7. Общие условия SLA для всех типов доступности

- 3.1.7.1. Любой тип доступности не распространяется на ситуации полной недоступности сети между узлами и полной недоступности площадки эксплуатации кластера.

3.1.7.2. Доступность каждого компонента кластера в отдельности и типа доступности в целом определяется по наилучшему значению от нескольких агентов проверки за каждые 30 секунд.

3.2. Гарантируемый уровень сервиса

В зависимости от типа доступности, конфигурации кластера и выбранного канала обновлений в рамках SLA может быть гарантирован различный уровень доступности.

№	Тип доступности	Standard	Silver	Gold	Platinum
3.2.1.	Synthetic Availability	99,90%	<ul style="list-style-type: none"> 99,95%, при использовании канала обновлений Stable или RockSolid 99,90%, при использовании канала обновлений EarlyAccess 	<ul style="list-style-type: none"> 99,99%, при использовании канала обновлений RockSolid 99,95%, при использовании канала обновлений Stable 99,90%, при использовании канала обновлений EarlyAccess 	
3.2.2.	Nginx Availability	<ul style="list-style-type: none"> 99,90%, при включении режима HA и размещении на выделенных узлах SL не обеспечивается в других случаях (менее 2 выделенных узлов, не включен режим HA) 	<ul style="list-style-type: none"> 99,95%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений Stable или RockSolid 99,90%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений EarlyAccess SL не обеспечивается в других случаях (менее 2 выделенных узлов, не включен режим HA) 	<ul style="list-style-type: none"> 99,99%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений RockSolid 99,95%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений Stable 99,90%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений EarlyAccess SL не обеспечивается в других случаях (менее 2 выделенных узлов, не включен режим HA) 	
3.2.3.	Node Group Availability	99,50%	<ul style="list-style-type: none"> 99,90%, при использовании канала обновлений Stable или RockSolid 99,50%, при использовании канала обновлений EarlyAccess 	<ul style="list-style-type: none"> 99,95%, при использовании канала обновлений RockSolid 99,90%, при использовании канала обновлений Stable 99,50%, при использовании канала обновлений EarlyAccess 	
3.2.4.	Control Plane Availability	<ul style="list-style-type: none"> 99,50%, при использовании 3 мастеров на выделенных узлах SL не обеспечивается в других случаях (менее 3 мастеров) 	<ul style="list-style-type: none"> 99,90%, при использовании 3 мастеров на выделенных узлах и канала обновлений Stable или RockSolid 99,50%, при использовании 3 мастеров на выделенных узлах и канала обновлений EarlyAccess SL не обеспечивается в других случаях (менее 3 мастеров) 	<ul style="list-style-type: none"> 99,95%, при использовании 3 мастеров на выделенных узлах и канала обновлений RockSolid 99,90%, при использовании 3 мастеров на выделенных узлах и канала обновлений Stable 99,50%, при использовании 3 мастеров на выделенных узлах и канала обновлений EarlyAccess SL не обеспечивается в других случаях (менее 3 мастеров) 	
3.2.5.	M&A Availability	<ul style="list-style-type: none"> 99,50%, при использовании выделенных узлов и включении режима HA 	<ul style="list-style-type: none"> 99,50%, при использовании выделенных узлов и включении режима HA SL не обеспечивается в других 	<ul style="list-style-type: none"> 99,95%, при использовании выделенных узлов, включении режима HA и использовании канала обновлений RockSolid 	

		<ul style="list-style-type: none"> • SL не обеспечивается в других случаях 	случаях	<ul style="list-style-type: none"> • 99,50%, при использовании выделенных узлов и включении режима HA • SL не обеспечивается в других случаях
3.2.6.	Extensions Availability	<ul style="list-style-type: none"> • 99,00%, при использовании выделенных узлов и включении режима HA • SL не обеспечивается в других случаях 	<ul style="list-style-type: none"> • 99,50%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений Stable или RockSolid • 99,00%, при включении режима HA, размещении на выделенных узлах и использовании канала обновлений EarlyAccess • SL не обеспечивается в других случаях 	<ul style="list-style-type: none"> • 99,50%, при использовании выделенных узлов, включении режима HA и использовании канала обновлений Stable или RockSolid • 99,00%, при использовании выделенных узлов, включении режима HA и использовании канала обновлений EarlyAccess • SL не обеспечивается в других случаях

3.3. Штрафы за нарушение уровня сервиса

В случае нарушения SLA по нескольким типам доступности одновременно в рамках одного кластера учитывается только штраф с наибольшей суммой по типу доступности.

№	Название	Standard	Silver	Gold	Platinum
3.3.1.	Максимальная сумма штрафа за месяц обслуживания за нарушение SLA ¹	Не более 25% от ежемесячной стоимости обслуживания	Не более 50% от ежемесячной стоимости обслуживания	Не более 75% от ежемесячной стоимости обслуживания	
3.3.2.	Размер штрафа за 1 час простоя	1% от общей ежемесячной стоимости обслуживания данного кластера	3% от общей ежемесячной стоимости обслуживания данного кластера	5% от общей ежемесячной стоимости обслуживания данного кластера	

¹ Исключая простой по объективным и независящим от Исполнителя обстоятельствам, например, простой площадки, проблемы с оплатой площадки, форс-мажорные обстоятельства и т.п.

4. Стоимость и порядок оплаты

4.1. Стоимость обслуживания кластера

№	Название	Standard	Silver	Gold	Platinum
4.1.1.	Базовая цена на Cloud cluster ¹ , ₽/мес.	25 000			
4.1.2.	Базовая цена на NonCloud cluster ¹ , ₽/мес.	35 000			
4.1.3.	Базовая цена на Hybrid cluster ¹ , ₽/мес.	35 000			
4.1.4.	Контактные лица со стороны Заказчика ² , ₽/мес.	15 000			

4.1.5.	Дополнительные контактные лица со стороны Заказчика, третье и последующее контактное лицо, ₽/мес.	10 000			
4.1.6.	Доплата за тариф для Cloud cluster, ₽/мес.	0	10 000	40 000	80 000
4.1.7.	Доплата за тариф для NonCloud cluster, ₽/мес.	0	20 000	50 000	90 000
4.1.8.	Доплата за тариф для Hybrid cluster, ₽/мес.	0	20 000	50 000	90 000
4.1.9.	Доплата за тариф для Backup cluster, ₽/мес., в инфраструктуре, идентичной основному кластеру.	0	0	0	0
4.1.10.	Доплата за тариф для Backup cluster, ₽/мес., в инфраструктуре, не идентичной основному кластеру.	0	Согласно типу инфраструктуры и требуемому тарифу по ценам из пп. 4.1.6, 4.1.7, 4.1.8		
4.1.11.	Доплата за тариф для Migrate cluster, ₽/мес., в инфраструктуре, идентичной основному кластеру	0	0	0	0
4.1.12.	Доплата за тариф для Migrate cluster, ₽/мес., в инфраструктуре, не идентичной основному кластеру	0	Согласно типу инфраструктуры и требуемому тарифу по ценам из пп. 4.1.6, 4.1.7, 4.1.8		
4.1.13.	Доплата за тариф, безлимитный тариф (любое количество кластеров, любой тип кластера) ³ , ₽/мес.	-	150 000	250 000	400 000

¹ В базовую цену обслуживания кластера включено обслуживание рабочих узлов на общую сумму не более 10 000 рублей и не более 3-х выделенных мастер-узлов.

² Обязательная опция для первого кластера Заказчика, включает до 2 контактных лиц. Контактные лица не привязываются к конкретному кластеру. Для второго и последующих кластеров опция не является обязательной.

³ При условии размещения кластеров не более чем на двух разных видах инфраструктуры, включая и разные типы кластеров, и разные типы облаков для Cloud cluster.

4.2. Стоимость обслуживания узлов в кластере

№	Название	Standard	Silver	Gold	Platinum
4.2.1.	Cloud-узел, ₽/мес.	500			
4.2.2.	VM-узел, ₽/мес.	1 000			
4.2.3.	Hard-узел, ₽/мес.	2 000			
4.2.4.	Special-узел, ₽/мес.	5 000			

4.3. Расчет стоимости и порядок оплаты

- А. Отчетный период для данной Услуги составляет один календарный месяц. В случае оказания Услуги в течение неполного календарного месяца, отчетный период сокращается пропорционально фактическому количеству дней оказания Услуги.
- В. Расчет стоимости Услуг в отчетном периоде производится согласно правилам и условиям, определенным в настоящих Условиях оказания услуг, а также в Заказе, пропорционально количеству дней, входящих в отчетный период.
- С. Оплата Услуги производится по окончании очередного отчетного периода одним платежом в размере 100% стоимости Услуги не позднее 5 (пятого) банковского дня с момента окончания отчетного периода на основании счета выставленного и направленного Исполнителем по email и/или ЭДО.
- Д. Порядок сдачи-приемки Услуги определен Договором.
- Е. Все цены указаны без НДС. Заказчик доплачивает НДС, согласно действующим ставкам в соответствии с НК РФ, дополнительно к стоимости Услуг.

5. Дополнительные услуги

№	Название	Standard	Silver	Gold	Platinum
5.1.	Проведение значительных обновлений в нерабочее время	Не применимо	10 000 рублей за час работ, минимум 3 часа		
5.2.	Проведение дополнительных работ в рабочее время	8 000 рублей в час, минимум 2 часа			
5.3.	Проведение дополнительных работ в нерабочее время	16 000 рублей в час, минимум 3 часа			
5.4.	Работы по разработке, адаптации, доработке и исправлению upstream-версий Kubernetes и другого ПО экосистемы Kubernetes	6 000 рублей в час, минимум 10 часов			
5.5.	Стоимость использования канала обновлений RockSolid	Не применимо	20 000 ₺ за каждый кластер сверх 1/3 от общего количества кластеров	10 000 ₺ за каждый кластер сверх 1/3 от общего количества кластеров	Не применимо