ATLASSIAN

Обзор результатов реагирования на инцидент

Сбой в обслуживании в апреле 2022 г.

Данный перевод предоставлен исключительно для удобства. В случае любых неопределенностей или противоречий между переводами английская версия имеет преимущественную силу.

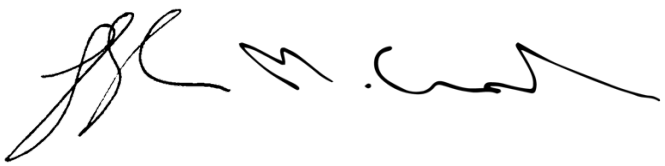
Сообщение от наших соучредителей и соуправляющих директоров

Мы хотим засвидетельствовать факт сбоя в обслуживании клиентов, произошедшего в начале этого месяца. Мы понимаем, что наши продукты критически важны для вашего бизнеса, и со всей серьезностью подходим к исполнению своих обязательств. Вся ответственность лежит на нас. Ни на ком другом. Мы прилагаем усилия для того, чтобы вернуть доверие клиентов, которые испытали на себе последствия инцидента.

Одна из основных ценностей Atlassian гласит: «Открытая компания, никакой ерунды». Эта ценность проявляется, в частности, в том, что мы открыто обсуждаем инциденты и относимся к ним как к возможности для обучения. Мы публикуем этот обзор результатов реагирования на инцидент для наших клиентов, сообщества Atlassian и технического сообщества в целом. Компания Atlassian гордится своим процессом управления инцидентами, в основе которого лежит идея о том, что невозможно наладить надежное обслуживание в крупном масштабе без культуры, в которой не ищут виноватых, и нацеленности на совершенствование технических систем и процессов. Хотя мы прилагаем все усилия, чтобы избежать инцидентов, мы также верим, что инциденты — мощный стимул для улучшения.

Хотим заверить, что облачная платформа Atlassian позволяет нам удовлетворять разнообразные потребности более 200 000 клиентов независимо от их размера и отрасли. До этого инцидента наше облако работало без отказов стабильно на протяжении 99,9 % времени и с превышением времени безотказной работы, гарантированного в SLA. Мы сделали долгосрочные вложения в развитие нашей централизованной платформы и отдельных ее возможностей, реализовав масштабируемую инфраструктуру и наладив регулярный выпуск улучшений системы безопасности.

Мы благодарим наших клиентов и партнеров за то, что вы продолжаете доверять нам и сотрудничать с нами. Мы надеемся, что, прочитав этот документ и ознакомившись с описанными в нем мерами, вы убедитесь в том, что компания Atlassian продолжит предоставлять облачную платформу мирового класса и многоцелевой портфель продуктов, которые удовлетворят потребности любой команды.



Скотт и Майк

Краткое описание

Во вторник, 5 апреля 2022 года, начиная с 7:38 UTC 775 клиентов Atlassian потеряли доступ к своим продуктам Atlassian. Часть клиентов продолжали испытывать на себе последствия сбоя до 14 дней. Для первой группы клиентов восстановление было выполнено 8 апреля, а к 18 апреля были постепенно восстановлены все сайты клиентов.

Произошедшее не было следствием кибератаки или несанкционированного доступа к данным клиентов. В компании Atlassian действует комплексная программа управления данными; заключенные соглашения SLA были опубликованы, а минимальные показатели, гарантированные в них, обеспечивались даже с запасом.

Хотя это серьезный инцидент, ни один клиент не потерял данные более чем за пять минут работы. Кроме того, более 99,6 % клиентов и пользователей продолжали использовать наши облачные продукты во время работ по восстановлению без каких-либо проблем.

> В рамках этого документа для описания клиентов, чьи сайты были удалены из-за инцидента, мы будем использовать формулировки «клиенты, которые испытали на себе последствия инцидента» и «затронутые (инцидентом) клиенты». В этом обзоре результатов реагирования на инцидент приводятся точные сведения об инциденте, описаны шаги, которые мы предприняли для восстановления, и наш подход к предотвращению подобных ситуаций в будущем. В этом разделе дается общее описание инцидента, а более подробная информация содержится в других частях документа.

Что произошло?

В 2021 году мы завершили приобретение и интеграцию отдельного приложения Atlassian для Jira Service Management и Jira Software под названием Insight — Asset Management. Функциональность этого отдельного приложения была затем встроена в Jira Service Management и стала недоступна для Jira Software. В связи с этим нам требовалось удалить отдельное устаревшее приложение на тех сайтах клиентов, где оно было установлено. Наши технические команды использовали существующий скрипт и процедуру для удаления экземпляров этого отдельного приложения, но возникли две проблемы.

- **Отсутствие взаимопонимания.** Между командой, которая запрашивала удаление, и командой, которая его проводила, возникло недопонимание. Вместо того чтобы предоставить идентификаторы *удаляемого приложения*, команда предоставила идентификаторы *всего облачного сайта*, на котором необходимо было удалить приложения.

- **Отсутствие системных предупреждений.** Интерфейс API, с помощью которого выполнялось удаление, принял идентификаторы сайта и приложения и не обнаружил ошибки. То есть при получении идентификатора сайта удалялся сайт; при получении идентификатора приложения удалялось приложение. Система не генерировала предупреждение, чтобы подтвердить тип запрашиваемого удаления (сайта или приложения).

Выполненный скрипт прошел стандартный процесс проверки коллегами, которые уделяли основное внимание тому, какая конечная точка вызывается и как. Перекрестная сверка предоставленных идентификаторов облачного сайта не проводилась, поэтому не было проверено, к чему они относятся — к Insight App или ко всему сайту. Между тем скрипт содержал идентификатор всего сайта клиента. В результате 883 сайта (представляющих 775 клиентов) были удалены в период между 07:38 UTC и 08:01 UTC во вторник, 5 апреля 2022 года. *См. раздел «Что произошло».*

Как мы отреагировали?

Как только факт инцидента был подтвержден 5 апреля в 08:17 UTC, мы инициировали процедуру управления серьезными инцидентами и сформировали многофункциональную команду по управлению инцидентами. Команда реагирования на инциденты, включающая сотрудников со всего мира, работала круглосуточно и без выходных, пока инцидент не был разрешен, а все сайты — восстановлены, проверены и возвращены клиентам. Кроме того, руководители команды по управлению инцидентами каждые три часа устраивали собрания для координации направлений работы.

С самого начала мы поняли, что восстановление сайтов сотен клиентов, использующих несколько продуктов одновременно, сопряжено с рядом сложностей.

В начале инцидента мы точно знали, какие сайты затронул инцидент, и нашей главной задачей было наладить связь с подтвержденным владельцем каждого затронутого сайта, чтобы сообщить о сбое.

Однако контактные сведения некоторых клиентов оказались удалены, из-за чего клиенты не могли, как обычно, подавать заявки в службу поддержки. Также из-за этого мы не могли оперативно связаться с ключевыми контактными лицами клиентов. *Подробнее см. в разделе «Общий обзор направлений работы по восстановлению».*

Что мы делаем, чтобы предотвратить подобные ситуации в будущем?

Чтобы избежать такой ситуации в будущем, мы приняли ряд срочных мер и внесли или планируем внести существенные изменения в четырех перечисленных ниже областях.

1. **Реализация единообразного процесса обратимого удаления во всех системах**. В целом описанная выше операция удаления должна быть запрещена, либо для нее должна быть внедрена многоуровневая защита от ошибок, включающая поэтапное выполнение удаления и испытанный план отката для обратимого удаления. Мы внедрим глобальный запрет на удаление данных и метаданных клиентов, для которых не реализована процедура обратимого удаления.

2. **Улучшение нашей программы аварийного восстановления с целью автоматизации восстановления в ситуациях, когда выполняются операции удаления, затрагивающие несколько продуктов и (или) несколько сайтов, используемых большим числом клиентов.** Мы будем использовать автоматизацию и сделанные из этого инцидента выводы, чтобы быстрее привести программу аварийного восстановления в соответствие с показателем целевого срока восстановления (RTO), определенным в нашей политике касательно инцидентов такого масштаба. Мы будем регулярно отрабатывать действия по аварийному восстановлению, включающие восстановление всех продуктов для большой группы сайтов.

3. **Пересмотр процедуры управления инцидентами для крупномасштабных инцидентов**. Мы усовершенствуем нашу стандартную операционную процедуру для крупномасштабных инцидентов и будем отрабатывать ее, моделируя инциденты такого масштаба. Мы доработаем наши инструменты и планы подготовки специалистов, чтобы создать условия для параллельной работы большого количества команд.

4. **Создание сборника сценариев информирования о крупномасштабных инцидентах**. Мы будем признавать факт инцидента на раннем этапе в нескольких каналах. Мы будем публиковать сообщения для информирования общественности об инцидентах в течение нескольких часов. Нам нужно лучше доводить информацию до затронутых клиентов, поэтому мы улучшим процесс резервирования данных ключевых контактных лиц и модернизируем инструментарий поддержки, чтобы клиенты могли напрямую связаться с нашей службой технической поддержки без действительного URL-адреса или идентификатора Atlassian.

Подробный список наших задач приведен в полном обзоре результатов реагирования на инцидент ниже. *См. «Как мы будем совершенствоваться».*

Содержание

Обзор облачной архитектуры Atlassian

Чтобы понять, какие факторы способствовали возникновению инцидента, о котором говорится в этом документе, целесообразно сначала получить представление об архитектуре развертывания продуктов, сервисов и инфраструктуры Atlassian.

Архитектура облачного хостинга Atlassian

Поставщик облачных услуг Atlassian — компания Amazon Web Services (AWS). Atlassian использует высокодоступные центры обработки данных AWS в нескольких регионах мира. Каждый регион AWS представляет собой отдельное географическое местоположение с несколькими изолированными и физически разделенными группами центров обработки данных — так называемыми зонами доступности (AZ).

Для разработки своих продуктов и компонентов платформы мы используем вычислительные сервисы, сервисы хранения данных, сетевые сервисы и сервисы аналитики AWS. Благодаря этому мы можем задействовать возможности обеспечения избыточности, предлагаемые AWS, такие как зоны доступности и регионы.

Архитектура распределенных сервисов

В этой архитектуре AWS размещен ряд наших сервисов, которые обеспечивают работу платформы и продуктов, составляющих наши решения. Сюда относятся возможности платформы, которые применяются во множестве продуктов Atlassian, такие как поддержка мультимедиа, управление удостоверениями и функции электронной коммерции, пользовательские интерфейсы, такие как редактор, а также возможности для конкретных продуктов, такие как сервис Jira Issue и Confluence Analytics.

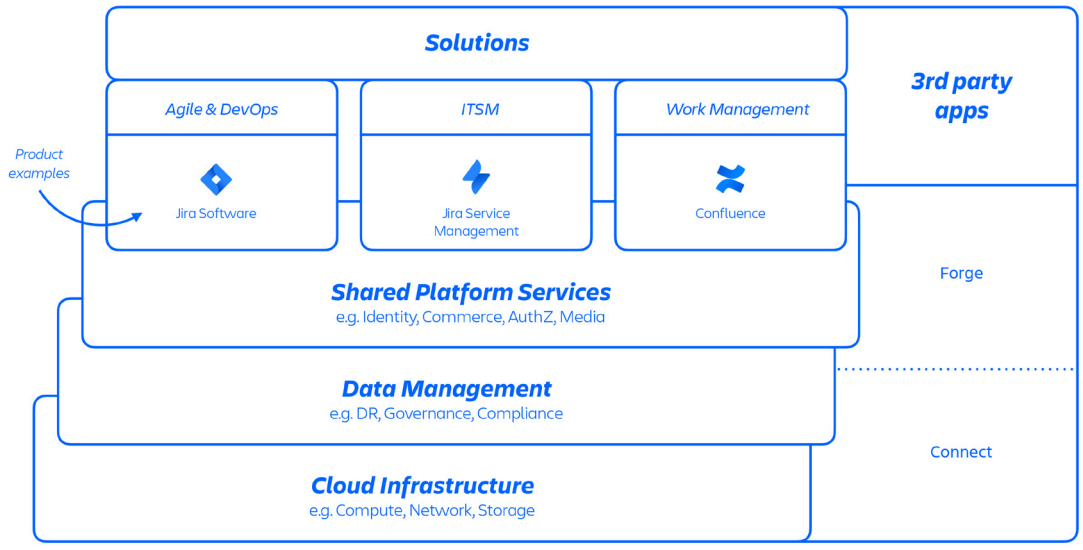


Рисунок 1. Архитектура платформы Atlassian.

Разработчики Atlassian предоставляют эти сервисы с помощью разработанной компанией платформы как услуги (PaaS) под названием Micros, которая автоматически координирует развертывание общих сервисов, инфраструктуры, хранилищ данных и возможностей управления ими, включая требования к контролю за обеспечением безопасности и соответствия нормативным требованиям (см. *рис. 1* выше). Как правило, продукт Atlassian состоит из нескольких сервисов, помещенных в контейнер, которые развертываются на AWS с помощью Micros. Продукты Atlassian используют базовые возможности платформы (см. *рис. 2* ниже), такие как маршрутизация запросов к хранилищам двоичных объектов, аутентификация/авторизация, транзакционный пользовательский контент (UGC), а также хранилища на базе модели «сущность — связь», озера данных, общее ведение журналов, трассировка запросов, возможности наблюдения и аналитические сервисы. Эти микросервисы разработаны с использованием утвержденных комплектов технологий, стандартизированных на уровне платформы:

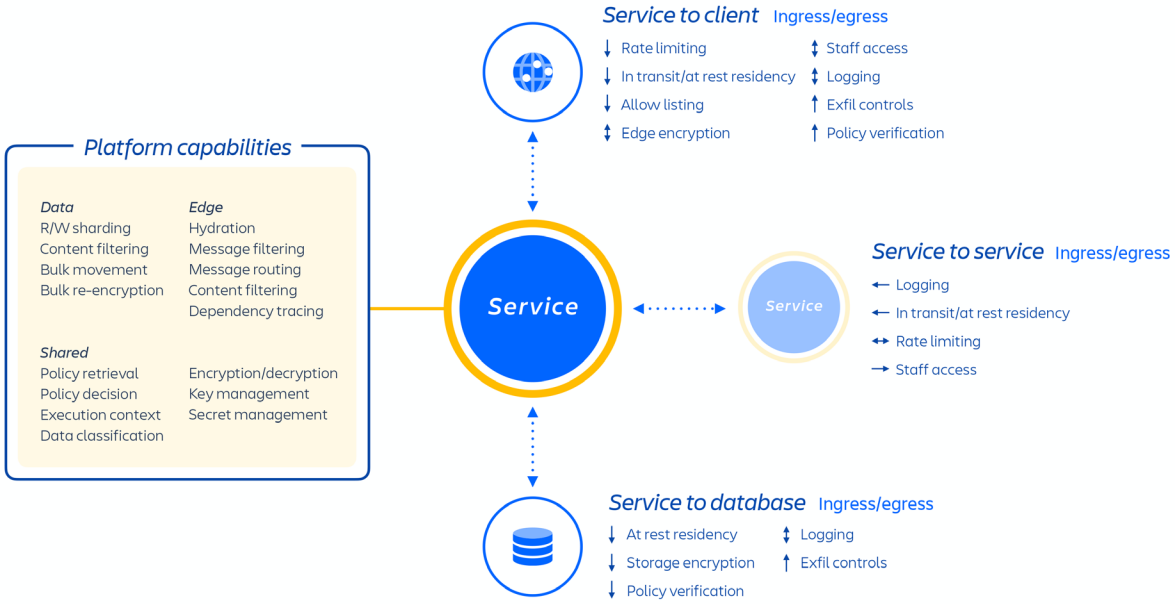


Рисунок 2. Обзор микросервисов Atlassian.

Многоклиентская архитектура

На базе облачной инфраструктуры мы создали архитектуру микросервисов с поддержкой нескольких держателей, а также общую платформу, поддерживающую наши продукты.
В архитектуре с несколькими держателями один сервис обслуживает множество клиентов, в том числе базы данных и экземпляры вычислительных ресурсов, необходимые для запуска наших облачных продуктов. Каждый сегмент (по сути, контейнер — см. *рис. 3* ниже) содержит данные нескольких держателей, но данные каждого держателя изолированы и недоступны другим держателям.

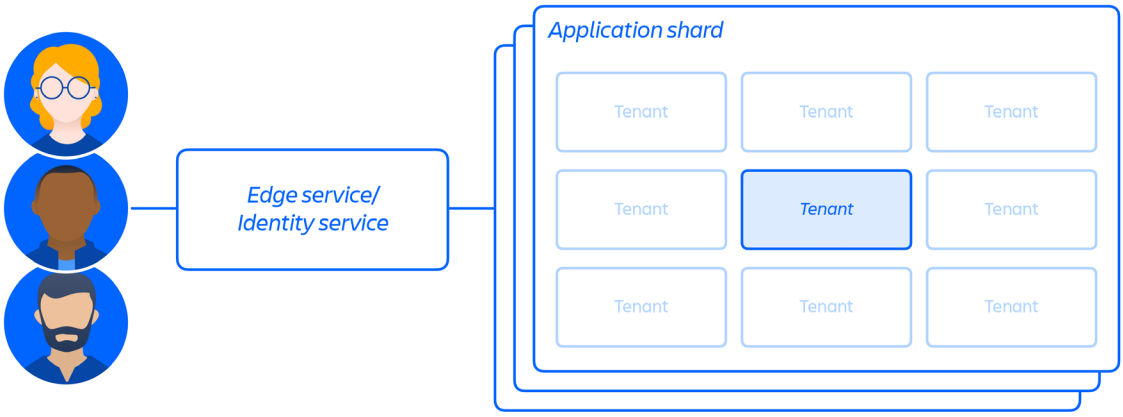


Рисунок 3. Как мы храним данные в архитектуре с несколькими держателями.

Создание и жизненный цикл держателей

При создании нового клиента ряд событий запускает оркестрацию распределенных сервисов и выделение ресурсов хранилищ данных. Эти события обычно можно сопоставить с одним из семи этапов жизненного цикла.

1. В коммерческие системы мгновенно поступают актуальные метаданные и информация о правах доступа, относящаяся к этому клиенту, после чего система оркестрации выделения согласовывает «состояние выделенных ресурсов» с состоянием лицензии посредством серии событий держателя и продукта.

События держателя

Эти события влияют на держателя в целом и могут быть следующими.

- Создание: держатель создается и используется для новых сайтов
- Уничтожение: держатель полностью удаляется

События продукта

- Активация: после активации лицензированных продуктов или сторонних приложений
- Деактивация: после деактивации определенных продуктов или приложений
- Приостановка: после приостановки работы определенного существующего продукта, в результате чего становится недоступен определенный сайт, владельцем которого является держатель
- Отмена приостановки: после отмены приостановки работы определенного существующего продукта, в результате чего становится доступен сайт, владельцем которого является держатель

Обновление лицензии: содержит информацию о количестве рабочих мест лицензии для определенного продукта, а также о ее статусе (активна/неактивна)

2. Создается сайт клиента и активируется соответствующий набор продуктов для клиента. Модель сайта представляет контейнер, содержащий несколько продуктов, предоставленных по лицензии конкретному клиенту (например, Confluence и Jira Software для <имя-сайта>.atlassian.net). Это (см. *рис. 4* ниже) важно понимать в контексте данного сообщения, поскольку в результате этого инцидента был удален именно контейнер сайта и в этом документе проблема рассматривается исходя из понятия сайта.

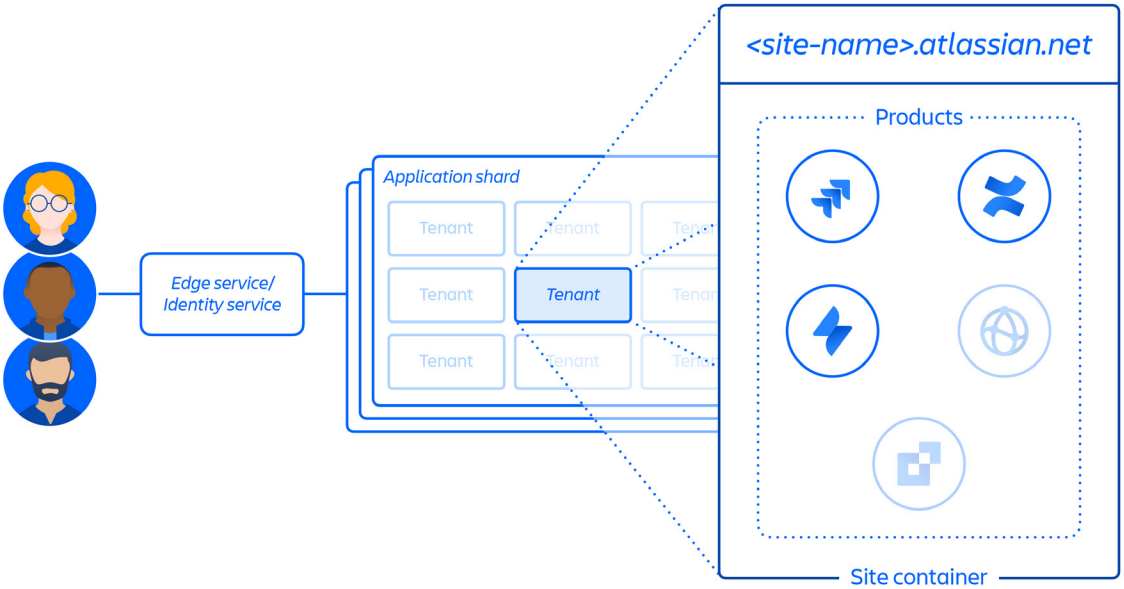


Рисунок 4. Обзор контейнера сайта.

3. Продукты предоставляются на сайт клиента в указанном регионе.

   При предоставлении продукта большая часть его содержимого размещается в регионе поблизости от того места, где пользователи получают к нему доступ. Чтобы оптимизировать производительность продукта, мы не ограничиваем перемещение данных, если он размещен глобально, и можем перемещать данные между регионами по мере необходимости.

   Для некоторых наших продуктов мы даем возможность выбрать вариант размещения данных. Клиенты могут решить, будут ли данные о продуктах распределены по всему миру или же они будут храниться в одном из определенных нами географических мест.

4. Создаются и сохраняются основные метаданные и конфигурация сайта и продуктов клиента.

5 Создаются и сохраняются идентификационные данные сайта и продуктов, такие как пользователи, группы, права и т. д.

6 Выделяются ресурсы баз данных продуктов на сайт, например семейство продуктов Jira, Confluence, Compass, Atlas.

7 Предоставляются лицензированные приложения продуктов.

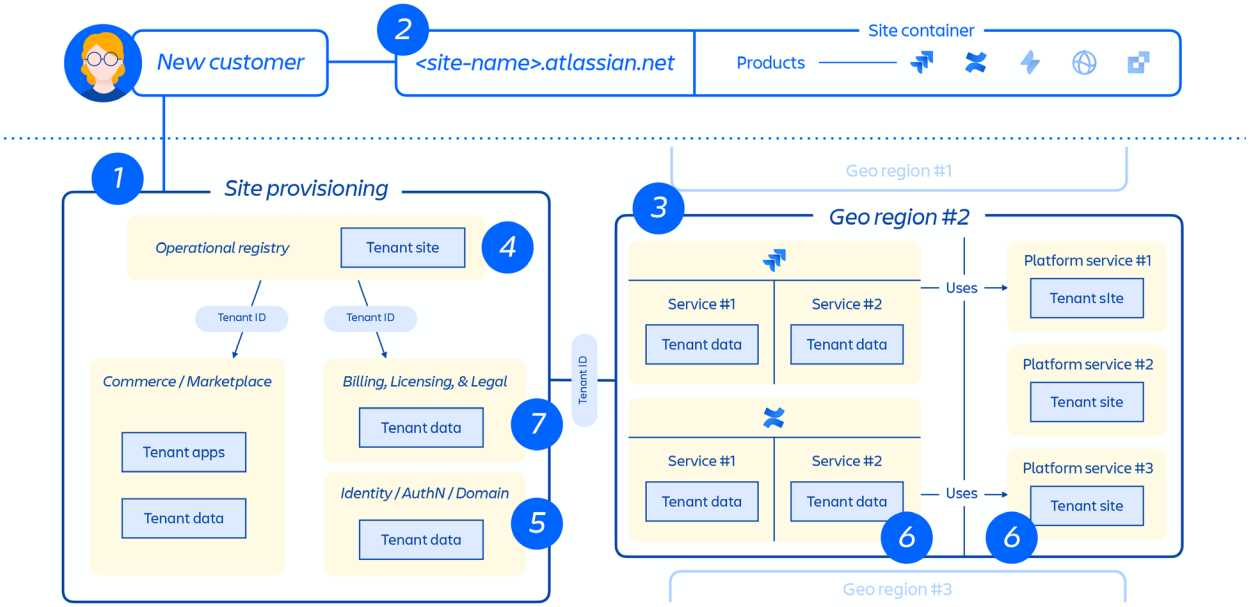


Рисунок 5. Обзор выделения сайта клиента в распределенной архитектуре.

На рис. 5 выше показано, как развертывается сайт клиента в нашей распределенной архитектуре, а не просто в рамках отдельной базы данных или хранилища. На нем изображено множество физических и логических местоположений, в которых хранятся метаданные, данные конфигурации, данные продуктов, данные платформы и другая связанная с сайтом информация.

Программа аварийного восстановления

В нашей программе аварийного восстановления обобщены все действующие меры. призванные обеспечить устойчивость к отказам инфраструктуры и восстанавливаемость хранилища сервисов из резервных копий. Обозначим два понятия, важных для понимания программ аварийного восстановления.

- **Целевой срок восстановления (RTO):** как быстро можно восстановить и вернуть данные клиенту во время аварийной ситуации?

- **Целевая точка восстановления (RPO):** насколько актуальны данные после восстановления из резервной копии? Сколько данных будет потеряно с момента последнего резервного копирования?

Во время этого инцидента мы не уложились в целевой срок восстановления, но выполнили обязательства по целевой точке восстановления.

Отказоустойчивость

Мы готовимся к сбоям на уровне инфраструктуры, например к потере всей базы данных, сервиса или зон доступности AWS. В рамках этой подготовки мы создаем реплики данных и сервисов в нескольких зонах доступности и регулярно выполняем тестирование аварийного переключения.

Восстанавливаемость хранилища сервисов

Мы также готовимся к восстановлению после повреждения данных в хранилище сервисов в связи с такими рисками, как программы-шантажисты, злоумышленники, дефекты программного обеспечения и операционные ошибки. В рамках этой подготовки мы создаем неизменяемые резервные копии и проводим тестовое восстановление хранилищ сервисов из резервных копий. Мы можем взять любое отдельное хранилище данных и вернуть его к состоянию, зафиксированному в предыдущий момент времени.

Возможность автоматического восстановления множества продуктов для множества сайтов

На момент инцидента у нас не было возможности выбрать множество сайтов клиентов и вернуть все их взаимосвязанные продукты из резервных копий к состоянию, зафиксированному в предыдущий момент времени.

Наши возможности были ориентированы на инфраструктуру, повреждение данных, события, связанные с отдельным сервисом, или удаление отдельных сайтов. В прошлом мы сталкивались с подобными сбоями, поэтому готовились к ним. У нас не было перечней процедур на случай удаления на уровне сайтов, которые можно было бы быстро автоматизировать в масштабе этого события. Для этого требовалось согласовать использование инструментов и автоматизацию для всех продуктов и сервисов.

В следующих разделах мы подробнее расскажем об этой сложности и о том, что компания Atlassian предпринимает для развития и оптимизации своих возможностей для поддержки этой архитектуры в любом масштабе.

Что произошло, хронология и восстановление

Что случилось

В 2021 году мы завершили интеграцию отдельного приложения Atlassian для Jira Service Management и Jira Software под названием Insight — Asset Management. Функциональность этого отдельного приложения была затем встроена в Jira Service Management и стала недоступна для Jira Software. В связи с этим нам требовалось удалить отдельное устаревшее приложение на тех сайтах клиентов, где оно было установлено. Наши технические команды использовали существующий скрипт и процедуру для удаления экземпляров этого отдельного приложения.

Однако возникли две критические проблемы:

- **Отсутствие взаимопонимания.** Между командой, которая запрашивала удаление, и командой, которая его проводила, возникло недопонимание. Вместо того чтобы предоставить идентификаторы удаляемого приложения, команда предоставила идентификаторы всего облачного сайта, на котором необходимо было удалить приложения.
- **Отсутствие системных предупреждений.** Интерфейс API, с помощью которого выполнялось удаление, принял идентификаторы сайта и приложения и не обнаружил ошибки. То есть при получении идентификатора сайта удалялся сайт; при получении идентификатора приложения удалялось приложение. Система не генерировала предупреждение, чтобы подтвердить тип запрашиваемого удаления (сайта или приложения).

Выполненный скрипт прошел стандартный процесс проверки коллегами, которые уделяли основное внимание тому, какая конечная точка вызывается и как. Перекрестная сверка предоставленных идентификаторов облачного сайта не проводилась, поэтому не было проверено, к чему они относятся — к приложению или ко всему сайту. Скрипт был протестирован в промежуточной среде согласно нашим стандартным процедурам управления изменениями, однако в результате тестирования не было выявлено, что предоставленные идентификаторы неверны, так как они отсутствовали в промежуточной среде.

При запуске в рабочей среде скрипт изначально был выполнен для 30 сайтов. Первый запуск в рабочей среде прошел успешно, и приложение Insight App было удалено на этих 30 сайтах без каких-либо других побочных эффектов. Дело в том, что при передаче идентификаторов для этих 30 сайтов недопонимания не возникло, и среди них были верные идентификаторы Insight App.

В скрипте, который выполнялся следующим в рабочей среде, содержались идентификаторы сайтов, а не идентификаторы приложения Insight. Он был применен к группе из 883 сайтов. Скрипт был запущен 5 апреля в 07:38 UTC и завершился в 08:01 UTC. Скрипт удалял сайты последовательно на основе входного списка, поэтому сайт первого клиента был удален вскоре после запуска скрипта в 07:38 UTC. В результате этого 883 сайта были моментально удалены, и нашим техническим группам не поступил сигнал тревоги.

Для затронутых инцидентом клиентов стали недоступны следующие продукты Atlassian: семейство продуктов Jira, Confluence, Atlassian Access, Opsgenie и Statuspage.

Как только мы узнали об инциденте, наши команды занялись восстановлением сайтов всех клиентов, испытавших на себе последствия инцидента. На тот момент, по нашей оценке, было затронуто примерно 700 сайтов (всего инцидент затронул 883 сайта, но мы не стали учитывать сайты, которыми владела компания Atlassian). Значительная часть из этих 700 сайтов приходилась на неактивные, бесплатные или небольшие учетные записи с небольшим количеством активных пользователей. Исходя из этого, мы изначально оценили количество затронутых клиентов приблизительно как 400.

Сейчас, имея более четкое понимание ситуации и отталкиваясь для полной прозрачности от официального определения клиентов Atlassian, мы утверждаем, что последствия сбоя испытали на себе 775 клиентов, однако большинство пользователей относилось к тем 400 клиентам, которые входили в изначальную оценку. Для части этих клиентов последствия сбоя сохранялись в течение 14 дней. Для первой группы клиентов восстановление было выполнено 8 апреля, а для всех остальных — 18 апреля.

Как мы координировали действия

Первая заявка в службу поддержки была создана затронутым клиентом 5 апреля в 07:46 UTC. Наша внутренняя служба мониторинга не выявила проблему, потому что для удаления сайтов использовался стандартный рабочий процесс. В 08:17 UTC мы инициировали процесс управления серьезными инцидентами, сформировав многофункциональную команду по управлению инцидентами, и через семь минут, в 08:24 UTC, был объявлен критический уровень опасности. В 08:53 UTC наша команда подтвердила, что заявка в службу поддержки и запуск скрипта были связаны между собой. Как только мы поняли сложность восстановления, мы присвоили инциденту наивысший уровень серьезности. Это произошло в 12:38 UTC.

В команду по управлению инцидентами входили сотрудники из нескольких команд Atlassian, в том числе технические специалисты, специалисты из службы поддержки клиентов, команды по управлению программами, отдела связей и многих других команд. Основная команда проводила собрания каждые три часа, пока инцидент не был разрешен, а все сайты — восстановлены, проверены и возвращены клиентам.

Для управления ходом восстановления мы создали новый проект Jira (SITE) и рабочий процесс для отслеживания восстановления каждого сайта несколькими командами (техническими специалистами, специалистами по управлению программами, специалистами из службы поддержки и т. д.). Такой подход позволил всем командам легко выявлять и отслеживать проблемы, связанные с восстановлением любого отдельного сайта.

8 апреля в 03:30 UTC мы также ввели запрет на изменение кода во всех проектах технического отдела на время инцидента. Это позволило нам сосредоточиться на восстановлении сайтов клиентов, устранить риск возникновения несогласованности данных клиентов вследствие изменений, свести к минимуму риск других сбоев и снизить вероятность того, что изменения, не связанные с инцидентом, отвлекут команду от восстановления.

Хронология инцидента

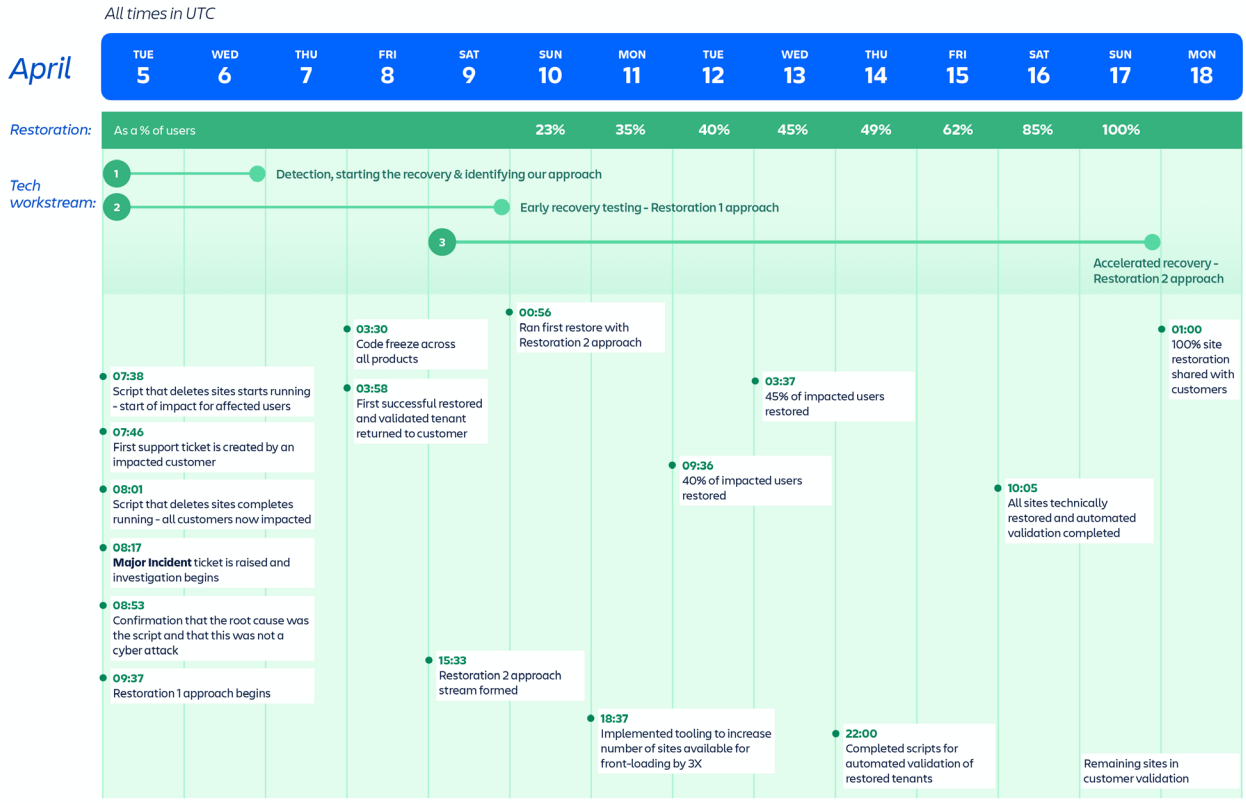


Рисунок 6. Хронология инцидента и ключевые этапы восстановления.

Общий обзор направлений работы по восстановлению

Процесс восстановления был разделен на три основных направления работы: обнаружение, предварительное восстановление и ускорение. Ниже каждое направление работы описано по отдельности, однако во время восстановления работа велась параллельно по всем направлениям.

Направление работы № 1: обнаружение, начало восстановления и формирование нашего подхода

Временная метка: дни 1–2 (5–6 апреля)
В 08:53 UTC 5 апреля мы обнаружили, что скрипт Insight App вызвал удаление сайтов. Мы подтвердили, что это произошло не в результате кибератаки или злоумышленного деяния внутри компании. Соответствующие команды по инфраструктуре продуктов и платформ были оповещены и привлечены к реагированию на инцидент.

В начале инцидента мы поняли следующее.

- Восстановление сотен удаленных сайтов — это сложный многоэтапный процесс (подробно описанный выше в разделе об архитектуре), для выполнения которого требуются усилия нескольких команд и несколько дней.
- У нас была возможность восстановить отдельный сайт, но мы не предусмотрели возможности и процессы для восстановления большой группы сайтов.

Поэтому нам потребовалось разбить процесс восстановления на несколько параллельных направлений работы и автоматизировать его, чтобы как можно скорее вернуть затронутым клиентам доступ к продуктам Atlassian.

В направлении работы № 1 было занято большое количество команд разработчиков. Они занимались следующими задачами.

- Определение этапов восстановления и поэтапное восстановление групп сайтов в конвейере.
- Разработка и совершенствование автоматизации, чтобы команды могли выполнять поэтапное восстановление большего количества сайтов в группе.

Направление работы № 2: предварительное восстановление и первый подход к восстановлению

Временная метка: дни 1–4 (5–9 апреля)
Мы выяснили, что привело к удалению сайтов 5 апреля в 08:53 UTC, в первый же час после завершения работы скрипта. Мы также определили процедуру восстановления, которая ранее

использовалась для восстановления небольшого количества сайтов в рабочей среде. Однако процедура восстановления удаленных сайтов в таком масштабе не была четко определена.

Чтобы быстро приступить к восстановлению, первые этапы реагирования на инцидент были разделены между двумя рабочими группами.

- Группа, на которую была возложена ручная работа, утвердила необходимые этапы и вручную выполнила восстановление для небольшого количества сайтов.
- Рабочая группа, которой поручили автоматизацию, взяла существующую процедуру восстановления за основу и разработала автоматизацию для беспроблемного поэтапного восстановления более крупных групп сайтов.

Обзор первого подхода к восстановлению (см. *рис. 7* ниже):

- Для каждого удаленного сайта было необходимо создать новый сайт, после чего восстановить данные каждого подключенного продукта, сервиса и хранилища.
- Новому сайту присваивались новые идентификаторы, такие как cloudId. Все эти идентификаторы считаются неизменяемыми, то есть во многих системах внедряются в записи данных. Из-за изменения этих идентификаторов нам пришлось обновлять большие объемы данных, что представляло особую сложность в сторонних приложениях экосистемы.
- Чтобы изменить новый сайт для воспроизведения состояния удаленного сайта, нужно было учитывать сложные и зачастую непредсказуемые зависимости между этапами.

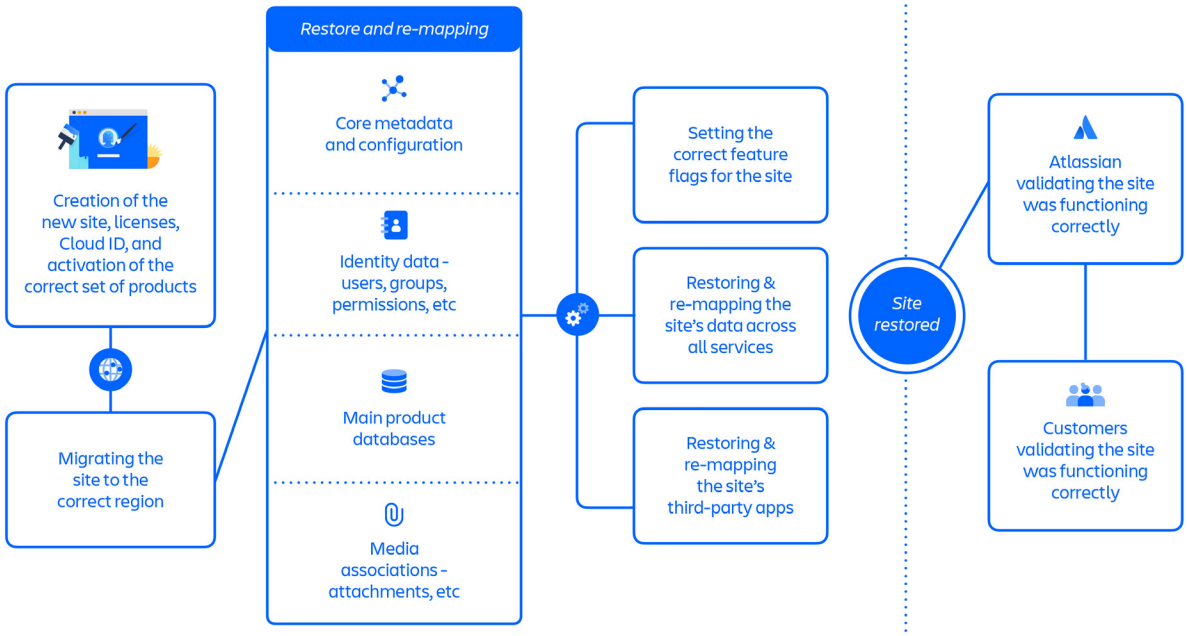


Рисунок 7. Ключевые этапы в рамках первого подхода к восстановлению.

Первый подход к восстановлению состоял примерно из 70 отдельных этапов, которые в совокупности выстраивались в следующий, по большей части последовательный процесс:

- создание нового сайта, лицензий и идентификатора Cloud ID, активация правильного набора продуктов;
- перенос сайта в правильный регион;
- восстановление и повторное сопоставление основных метаданных и конфигурации сайта;
- восстановление и повторное сопоставление идентификационных данных сайта: пользователей, групп, прав и т. д.;
- восстановление основных баз данных продуктов сайта;
- восстановление и повторное сопоставление связей для мультимедийных файлов сайта: вложений и т. д.;
- включение надлежащих возможностей для сайта с помощью флажков возможностей;
- восстановление и повторное сопоставление данных сайта для всех сервисов;
- восстановление и повторное сопоставление сторонних приложений сайта;
- проверка правильности работы сайта со стороны компании Atlassian;
- проверка правильности работы сайта со стороны клиента.

После оптимизации восстановление группы сайтов в рамках первого подхода занимало около 48 часов. Этот подход использовался для восстановления 112 сайтов 53 % затронутых пользователей в период с 5 по 14 апреля.

Направление работы № 3: ускоренное восстановление и второй подход к восстановлению

Временная метка: дни 4–13 (9–17 апреля)
Если бы мы использовали первый подход к восстановлению, нам потребовалось бы три недели, чтобы восстановить сайты для всех клиентов. Поэтому 9 апреля мы предложили новый подход для ускорения восстановления всех сайтов — второй подход к восстановлению (см. *рис. 8* ниже).

Второй подход к восстановлению отличался повышенным параллелизмом между этапами, который был достигнут благодаря снижению сложности и уменьшению количества зависимостей, характерных для первого подхода.

При втором подходе восстановление подразумевало повторное создание (или отмену удаления) записей, связанных с сайтом, во всех соответствующих системах, начиная с записи в каталоге служб. Основная идея нового подхода заключалась в *повторном использовании всех старых идентификаторов сайтов*. Благодаря этому количество этапов сократилось более чем наполовину по сравнению с предыдущим процессом. Мы отказались от этапов, в рамках которых старые идентификаторы сопоставлялись с новыми идентификаторами и требовалось согласовывать действия с каждым поставщиком сторонних приложений для каждого сайта.

Однако в случае отказа от первого подхода в пользу второго значительно возрастали издержки реагирования на инцидент.

- Многие скрипты и процессы автоматизации, разработанные в рамках первого подхода к восстановлению, было необходимо изменить для реализации второго подхода.
- Командам, выполняющим восстановление (в том числе координаторам инцидента), нужно было вести параллельные процессы восстановления в рамках обоих подходов, пока мы тестировали и утверждали процесс, предусмотренный вторым подходом.
- Чтобы использовать новый подход, нам нужно было протестировать и утвердить предусмотренную им процедуру, прежде чем ее масштабировать. Это потребовало бы заново проделать ту же работу по утверждению, что ранее была выполнена для первого подхода.

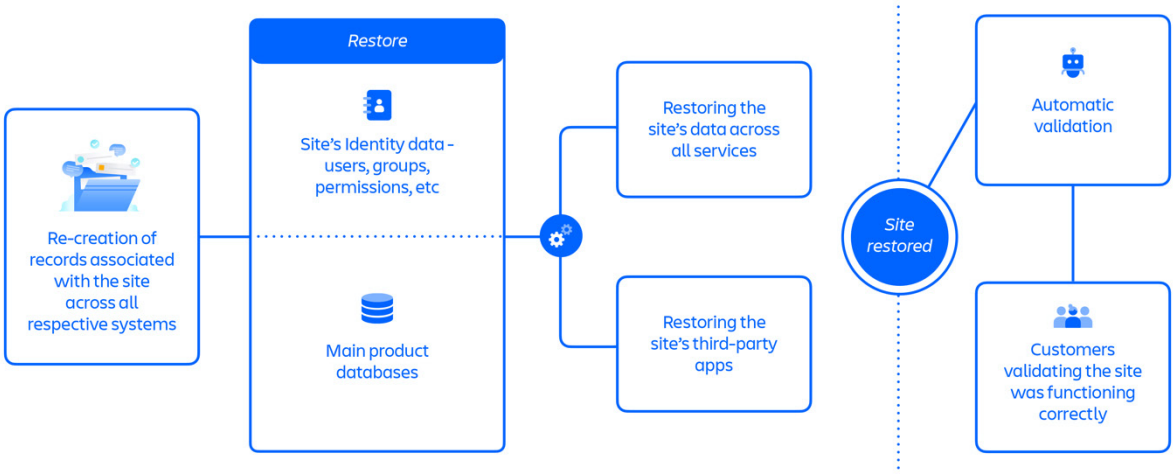


Рисунок 8. Ключевые этапы в рамках второго подхода к восстановлению.

Схема выше служит иллюстрацией второго подхода к восстановлению. Он включал более 30 этапов, которые выполнялись по большей части параллельно в рамках следующего процесса:

- повторное создание записей, связанных с сайтом, во всех соответствующих системах;
- восстановление идентификационных данных сайта: пользователей, групп, прав и т. д.;
- восстановление основных баз данных продуктов сайта;
- восстановление данных сайта для всех сервисов;
- восстановление сторонних приложений сайта;
- автоматическая проверка;
- проверка правильности работы сайта со стороны клиента.

Чтобы ускорить процессы, мы также приняли меры для предварительной подготовки и автоматизации восстановления, поскольку ручное восстановление не было бы эффективным в масштабе больших групп. Процесс восстановления, предполагающий последовательное выполнение этапов, мог занять больше времени при восстановлении больших баз данных и базы пользователей с правами. Среди принятых мер по оптимизации можно упомянуть следующие.

- Мы разработали инструменты и средства контроля, необходимые для *предварительной подготовки* и автоматизации длительных этапов, таких как восстановление баз данных и синхронизация идентификационных данных, чтобы завершить их прежде других этапов восстановления.
- Технические команды разработали автоматизацию для этапов, которые им были поручены, благодаря чему удалось беспроблемно восстанавливать сайты большими группами.
- Также была автоматизирована проверка правильности работы сайтов после завершения всех этапов восстановления.

При применении второго, ускоренного подхода на восстановление сайта уходило около 12 часов. Таким образом восстановление было выполнено примерно для 47 % затронутых пользователей с 771 сайтом в период с 14 по 17 апреля.

Минимальная потеря данных после восстановления удаленных сайтов

Для резервирования баз данных мы применяем полное резервное копирование в сочетании с инкрементным резервным копированием, что позволяет возвращать хранилища данных к состоянию на любой конкретный момент времени в рамках периода хранения резервных копий (30 дней). Мы определили основные хранилища данных продуктов для большинства клиентов, затронутых инцидентом, и решили использовать в качестве безопасной точки синхронизации точку восстановления, созданную за пять минут до удаления сайтов. Неосновные хранилища данных были восстановлены до той же точки или путем воспроизведения записанных событий. Благодаря тому, что для основных хранилищ использовалась фиксированная точка восстановления, нам удалось добиться непротиворечивости данных во всех хранилищах данных.

Так как у нас отсутствовали согласованные политики, а резервные снимки состояния баз данных извлекались вручную, у 57 клиентов, для которых восстановление было выполнено на ранних стадиях реагирования на инцидент, некоторые базы данных Confluence и Insight были восстановлены до точки, отстоящей *более* чем на пять минут от момента удаления сайта. Это несоответствие было обнаружено в процессе проверки после восстановления. После этого мы восстановили оставшиеся данные, связались с клиентами, которых это касалось, и теперь помогаем им вносить изменения для дальнейшего восстановления данных.

Резюмируем:

- Во время этого инцидента мы выполнили обязательства в отношении целевой точки восстановления (RPO), равной одному часу.
- Потери данных в результате инцидента ограничиваются пятью минутами, предшествующими удалению сайта.
- У небольшого количества клиентов базы данных Confluence или Insight были восстановлены до состояния, зафиксированного более чем за пять минут до удаления сайта, однако мы можем восстановить данные и в настоящее время работаем над этим с клиентами.

Сообщения об инцидентах

Важная составляющая нашего процесса информирования об инцидентах — контакты с клиентами, партнерами, СМИ, отраслевыми аналитиками, инвесторами и технологическим сообществом в целом.

Что случилось

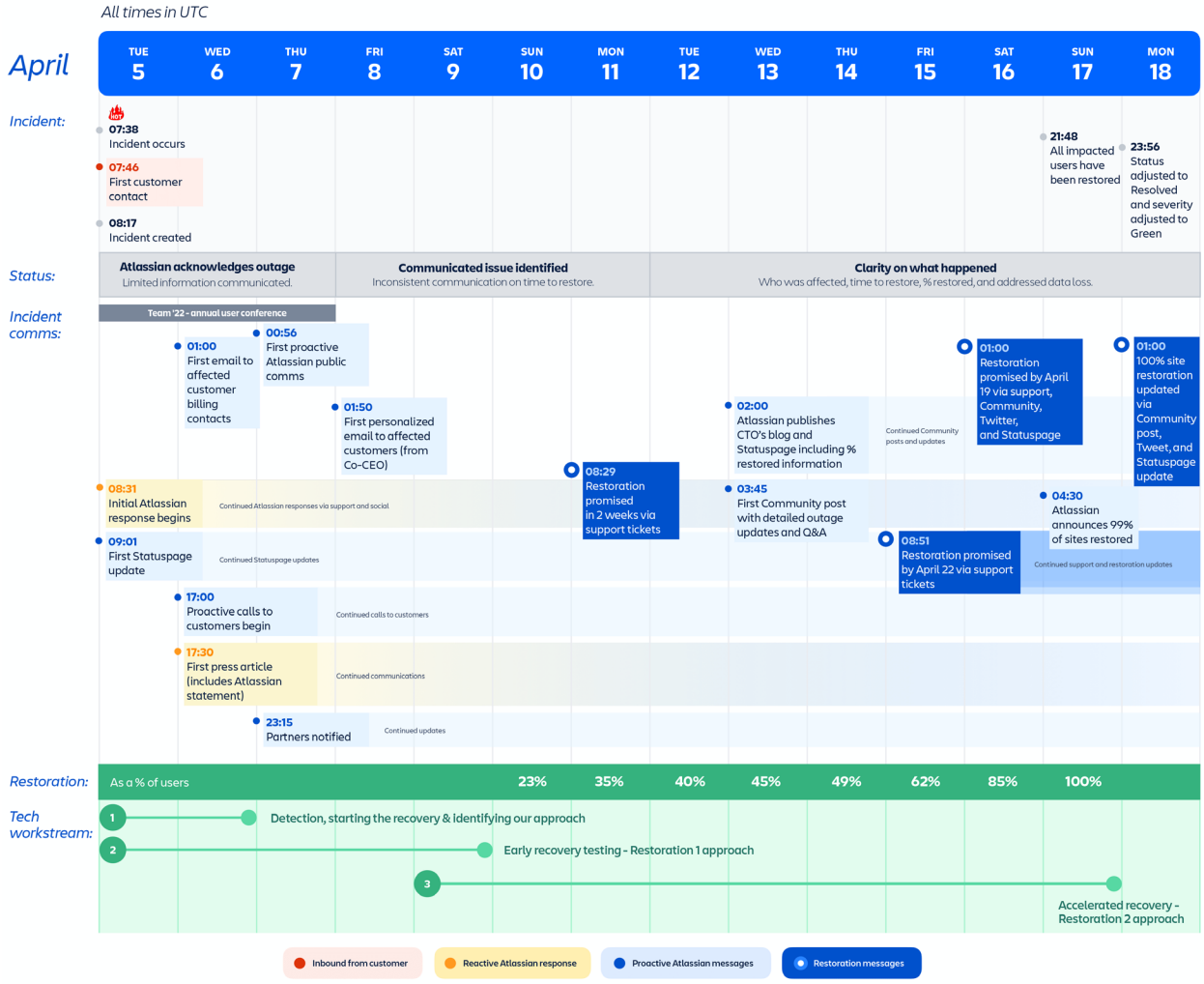


Рисунок 9. Хронология ключевых этапов информирования об инциденте.

Временная метка: дни 1–3 (5–7 апреля)

Первые этапы реагирования

Первая заявка в службу поддержки была создана 5 апреля в 7:46 UTC, и служба поддержки Atlassian в ответ подтвердила факт инцидента к 8:31 UTC. В 9:03 UTC было опубликовано первое сообщение Statuspage, из которого клиенты узнали, что мы расследуем инцидент. А в 11:13 UTC мы подтвердили через Statuspage, что выявили первопричину и работаем над исправлением. К 1:00 UTC 6 апреля в ответ на изначальные заявки клиентов были отправлены сообщения, в которых утверждалось, что сбой произошел из-за скрипта, связанного с обслуживанием системы, и мы ожидаем минимальных потерь данных. 6 апреля в 17:30 UTC компания Atlassian сделала заявление в ответ на запросы СМИ. 7 апреля в 00:56 UTC компания Atlassian опубликовала в Twitter свое первое сообщение для широкой публики, в котором признавала факт инцидента.

Временная метка: дни 4–7 (8–11 апреля)

Увеличиваем охват информирования и высылаем адресные сообщения

8 апреля в 1:50 UTC компания Atlassian отправила затронутым клиентам электронные письма с извинениями от соучредителя и соуправляющего директора Скотта Фаркуара.
В последующие дни мы работали над восстановлением удаленной контактной информации и созданием заявок в службу поддержки для всех затронутых сайтов, на которые еще не было подано ни одной заявки. В это время наша служба поддержки продолжала регулярно информировать клиентов о мерах по восстановлению через заявки в службу поддержки, связанные с каждым затронутым сайтом.

Временная метка: дни 8–14 (12–18 апреля)

Повышается прозрачность; завершается восстановление

12 апреля компания Atlassian опубликовала сообщение от технического директора Шри Вишваната, в котором он делился более подробными техническими сведениями о том, что произошло, кого затронул инцидент, потеряны ли данные и как продвигается работа по восстановлению, а также информировал о том, что полное восстановление всех сайтов может занять до двух недель. Статья в блоге сопровождалась еще одним заявлением для прессы за авторством Шри Вишваната. Мы также привели ссылку на его личный блог в сообществе Atlassian в публикации от технического руководителя Стивена Дизи, где впервые поднималась тема инцидента. Там же впоследствии стали появляться дополнительные сообщения и ответы на вопросы от широкой публики. В обновлении этого сообщения от 18 апреля было объявлено о полном восстановлении всех сайтов затронутых клиентов.

Почему мы не отреагировали публично раньше?

1. Нашим приоритетом было непосредственное общение с затронутыми клиентами посредством Statuspage, электронных сообщений, заявок в службу поддержки и в индивидуальном порядке. Однако нам не удалось связаться со многими клиентами, потому что при удалении их сайтов была утрачена контактная информация. Нам следовало расширить охват информирования гораздо раньше, чтобы затронутым клиентам и конечным пользователям стало известно о реакции на инцидент и сроках его разрешения.

2. Хотя мы сразу знали, что стало причиной инцидента, архитектурная сложность и уникальные обстоятельства этого инцидента не позволили нам быстро определить масштаб и точно оценить сроки разрешения. Вместо того чтобы ждать, когда у нас сложится полная картина ситуации, мы должны были откровенно заявить, что знаем, а что — нет. Если бы мы дали общий (хотя бы приблизительный) прогноз касательно восстановления и однозначно сказали, к какому времени рассчитываем составить более полное представление, наши клиенты смогли бы лучше планировать свои действия в условиях инцидента. Это в первую очередь касается системных администраторов и контактных лиц по техническим вопросам, которые непосредственно взаимодействуют с заинтересованными сторонами и пользователями в своих организациях.

Опыт поддержки и взаимодействие с клиентами

Как упоминалось ранее, скрипт удалил из наших рабочих сред не только сайты, но также идентификаторы и ключевую контактную информацию клиентов (например, URL-адреса сайтов Cloud и контактную информацию системных администраторов сайтов). Это была существенная потеря, так как URL-адрес Cloud и контактные данные системных администраторов сайтов используются в качестве основных идентификаторов во всех наших основных системах (например, в системах поддержки, лицензирования, оплаты) для целей обеспечения безопасности, переадресации запросов и расстановки приоритетов. Утратив эти идентификаторы, мы поначалу утратили способность систематической идентификации клиентов и взаимодействия с ними.

Как это повлияло на поддержку наших клиентов?

Во-первых, большинство затронутых клиентов не смогло связаться с нашей службой поддержки посредством обычной онлайн-формы для связи. Чтобы воспользоваться этой формой, пользователь должен войти в систему с помощью своего идентификатора Atlassian и указать действительный URL-адрес Cloud. Без этого пользователь не сможет отправить заявку в службу технической поддержки. В обычных условиях цель проверки — обеспечить безопасность сайта и определить очередность рассмотрения заявок. Однако это привело к непредвиденным последствиям для клиентов, затронутых сбоем: они лишились возможности отправлять приоритетные заявки в службу поддержки сайта.

Во-вторых, из-за того, что вследствие инцидента данные системных администраторов сайтов оказались удалены, мы не могли со своей стороны установить связь с пострадавшими клиентами. В первые дни инцидента мы инициировали связь с затронутыми клиентами через контактных лиц по счетам и техническим вопросам, зарегистрированных в Atlassian. Однако быстро выяснилось, что данные многих контактных лиц по счетам и техническим вопросам неактуальны. Поскольку мы потеряли контактные данные системных администраторов сайтов, у нас не было полного списка активных и утвержденных контактных лиц, с которыми можно было бы взаимодействовать.

Как мы отреагировали?

В первые дни инцидента перед нашими командами поддержки стояли три одинаково важные задачи, выполнение которых позволило бы ускорить восстановление сайтов и устранить пробелы в наших каналах связи.

Во-первых, нужно было получить достоверный список утвержденных контактных лиц клиентов. Поскольку наши технические команды были заняты восстановлением сайтов клиентов, команды, которые взаимодействуют с клиентами, сосредоточились на восстановлении проверенной контактной информации. Мы использовали все имеющиеся в нашем распоряжении механизмы (системы оплаты, предыдущие заявки в службу поддержки, другие защищенные резервные копии пользователей, связь с клиентами напрямую и т. д.), чтобы восстановить список контактов. Необходимо было сформировать заявку в службу поддержки в отношении каждого затронутого сайта, чтобы оптимизировать адресное информирование и скорость реагирования.

Во-вторых, нужно было восстановить рабочие процессы, очереди и соглашения SLA, имеющие отношение к этому инциденту. Удаление идентификатора Cloud и невозможность правильной аутентификации пользователей также повлияли на нашу способность обрабатывать связанные с инцидентом заявки в службу поддержки через обычные системы. Заявки неправильно отображались на соответствующих дашбоардах, в очередях приоритетов и запросов. Мы быстро сформировали многофункциональную команду (из специалистов службы поддержки, команды по продукту и ИТ-отдела) для разработки и внедрения дополнительной логики, SLA, состояний рабочих процессов и дашбоардов.

Поскольку это нужно было сделать в нашей рабочей системе, на полную разработку, тестирование и развертывание ушло несколько дней.

В-третьих, нужно было многократно увеличить масштаб ручных проверок для ускоренного восстановления сайтов. Пока технические команды выполняли первоначальное восстановление, стало ясно, что для быстрого восстановления сайтов понадобятся возможности наших глобальных команд поддержки, которые будут проводить ручное тестирование и проверки. Процедура проверки должна была сыграть важнейшую роль в возвращении восстановленных сайтов нашим клиентам, после того как техническая команда ускорит восстановление данных. Нам пришлось создать независимый поток стандартных операционных процедур, рабочих процессов, процессов передачи работы и кадровых ведомостей, чтобы мобилизовать более 450 специалистов поддержки для проведения проверок и организовать смены, чтобы работа шла круглосуточно и клиентам быстрее вернулись восстановленные сайты.

Несмотря на то что эти ключевые приоритеты были четко определены к концу первой недели, нам было сложно предоставлять *содержательную* информацию, поскольку из-за сложности процессов восстановления было неясно, к какому сроку инцидент будет разрешен. Нам следовало признать, что мы не можем своевременно назвать точную дату восстановления сайтов, и быстрее предоставить возможность обсуждения ситуации в индивидуальном порядке, чтобы наши клиенты могли строить планы.

Как мы будем совершенствоваться?

Мы немедленно внедрили запрет на массовое удаление сайтов до тех пор, пока не будут внесены соответствующие изменения.

Пересматривая внутренние процессы по результатам инцидента, мы хотим подчеркнуть, что инциденты происходят не из-за людей. Скорее, системы позволяют людям совершать ошибки. В этом разделе кратко излагаются факторы, поспособствовавшие возникновению этого инцидента. В нем также рассказывается, что мы планируем сделать, чтобы быстрее устранить эти недостатки и проблемы.

Вывод 1. Во всех системах нужно реализовать универсальный процесс обратимого удаления

Необходимо в целом запретить операции удаления, подобные той, что привела к данному инциденту, или же для них необходимо внедрить многоуровневую защиту от ошибок. Наше основное улучшение заключается в глобальном запрете на удаление данных и метаданных клиентов, для которых не реализована процедура обратимого удаления.

a) Удаление данных должно быть обратимым и только таким.
Удаление всего сайта должно быть запрещено, а для обратимого удаления должна действовать многоуровневая система защиты, которая предотвращала бы ошибки. Мы внедрим политику обратимого удаления, запрещающую внешним скриптам или системам удалять данные клиентов в рабочей среде. Согласно нашей политике обратимого удаления данные будут храниться достаточно долго, чтобы восстановить их можно было быстро и беспроблемно. Данные будут удаляться из рабочей среды только по истечении срока хранения.

Действия:

- **Реализовать обратимое удаление в рабочих процессах выделения ресурсов и во всех соответствующих хранилищах данных.** Команда, отвечающая за платформу держателей, будет следить за тем, чтобы удаление данных выполнялось только после деактивации. С той же целью в этой сфере будут развернуты и другие контрольные средства. В долгосрочной перспективе команда по платформе держателей будет играть ведущую роль в дальнейшей разработке правильного процесса управления состоянием данных держателей.

b) Для обратимого удаления должен быть налажен стандартизированный и утвержденный процесс проверки.
Операции обратимого удаления сопряжены с высоким риском. В связи с этим для таких операций должны быть налажены стандартизированные или автоматизированные процессы проверки, в которых будут предусмотрены откаты и процедуры тестирования.

Действия:

- **Сделать обязательным поэтапное выполнение всех действий по обратимому удалению.** Все новые операции, требующие удаления, сначала будут протестированы на наших собственных сайтах, чтобы подтвердить правильность подхода и проверить работу автоматизации. Как только мы завершим проверку, мы постепенно включим клиентов в этот процесс и продолжим тестирование на наличие ошибок и только потом будем применять автоматизацию ко всей выбранной пользовательской базе.

- **Для операций обратимого удаления должен существовать проверенный план отката.** Перед запуском любой операции по обратимому удалению данных в производственной среде необходимо тестировать восстановление удаленных данных, и для любой такой операции должен существовать проверенный план отката.

Вывод 2. В рамках программы аварийного восстановления следует автоматизировать восстановление в ситуациях, когда выполняются операции удаления, затрагивающие множество продуктов и (или) сайтов, используемых большим количеством клиентов

На странице Управление данными в Atlassian подробно описаны наши процедуры управления данными. Для обеспечения высокой доступности мы предоставляем и поддерживаем синхронную резервную репликацию в нескольких зонах доступности AWS. Аварийное переключение между зонами доступности выполняется автоматически и обычно занимает 60–120 секунд. Мы регулярно устраняем перебои в работе центров обработки данных и другие распространенные нарушения без последствий для клиентов.

Кроме того, мы поддерживаем неизменяемые резервные копии, которые обеспечивают устойчивость к событиям, вызывающим повреждение данных. Это позволяет восстанавливать данные на предыдущий момент времени. Резервные копии хранятся в течение 30 дней, и Atlassian постоянно тестирует и проверяет возможность восстановления резервных копий из хранилища. При необходимости мы можем восстановить данные всех клиентов в новой среде.

С помощью этих резервных копий мы регулярно выполняем откат к предыдущей версии для отдельных клиентов или небольших групп клиентов, случайно удаливших собственные данные. Однако у нас не было перечней процедур на случай удаления на уровне сайтов, которые можно было бы быстро автоматизировать в масштабе этого события. Для этого требовалось согласовать использование инструментов и автоматизацию для всех продуктов и сервисов.

Процесс, который мы (пока) не автоматизировали, — это восстановление данных большого количества клиентов в существующей (и используемой в настоящее время) среде без влияния на других клиентов.

В нашей облачной среде в каждом хранилище данных находятся данные множества клиентов. Поскольку данные, удаленные во время этого инцидента, составляют лишь часть данных в хранилищах, которыми продолжают пользоваться другие клиенты, нам приходится вручную извлекать и восстанавливать отдельные фрагменты из резервных копий. Восстановление сайта каждого клиента — это длительный и сложный процесс, требующий внутренней проверки, а затем окончательной проверки со стороны клиента.

Действия:
Ускорить восстановление множества сайтов и продуктов для большего числа клиентов. Программа аварийного восстановления отвечает требованиям наших текущих стандартов: целевая точка восстановления (RPO) не превышает одного часа. Мы будем использовать автоматизацию и сделанные из этого инцидента выводы, чтобы быстрее привести программу аварийного восстановления в соответствие с показателем целевого срока восстановления (RTO), определенным в нашей политике касательно инцидентов такого масштаба.

Включить этот случай в тестирование аварийного восстановления и автоматизировать его проверку. Мы будем регулярно отрабатывать действия по аварийному восстановлению, включающие восстановление всех продуктов для большой группы сайтов. С помощью этого тестирования аварийного восстановления мы будем проверять актуальность перечней процедур по мере развития нашей архитектуры и появления новых нестандартных случаев. Мы будем постоянно совершенствовать наш подход к восстановлению, продолжать автоматизировать этапы процесса восстановления и сокращать время восстановления.

Вывод 3. Следует усовершенствовать процесс управления инцидентами для крупномасштабных событий

Наша программа по управлению инцидентами хорошо подходит для управления серьезными и незначительными инцидентами, с которыми нам довелось столкнуться за прошедшие годы. Мы часто моделируем реагирование на инциденты меньшего масштаба и меньшей продолжительности, в которых обычно задействовано меньше людей и команд.

Однако в самый разгар этого инцидента над восстановлением сайтов клиентов одновременно работали сотни технических специалистов и сотрудников службы поддержки клиентов. Наша программа и команды по управлению инцидентами не были готовы к серьезности, масштабу и продолжительности инцидента такого типа (см. *рис. 10* ниже).

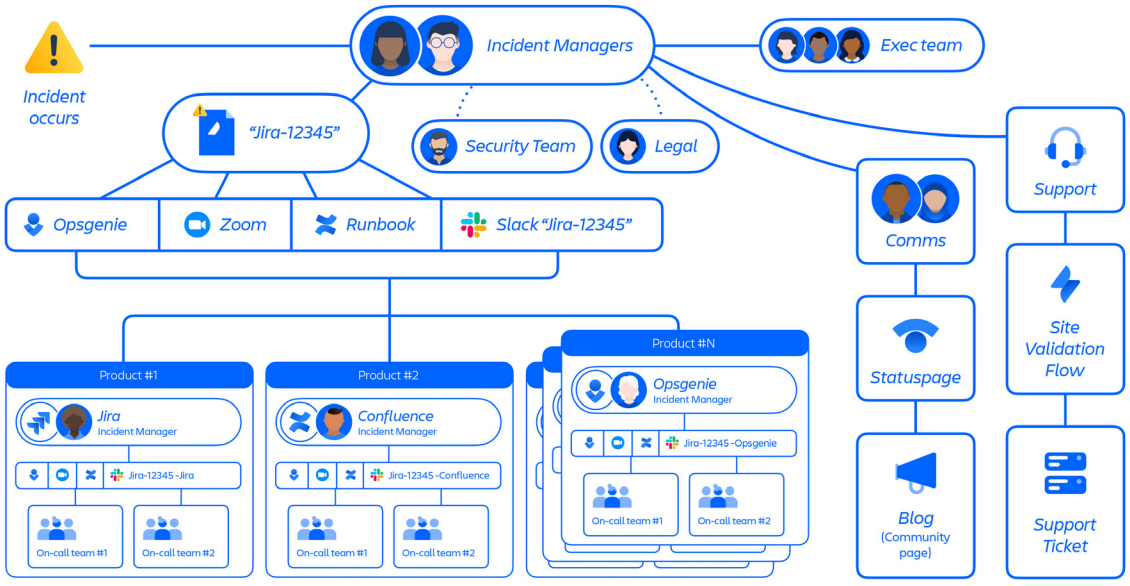


Рисунок 10. Обзор процесса управления крупномасштабными инцидентами.

Наша процедура управления крупномасштабными инцидентами будет лучше структурирован и будет часто отрабатываться

У нас разработаны сборники сценариев на случай инцидентов на уровне продуктов, но не на случай событий такого масштаба, когда в работу одновременно вовлечены сотни людей из разных подразделений компании. В наших инструментах по управлению инцидентами реализована автоматизация для создания каналов связи, таких как чаты Slack, конференции Zoom и документы Confluence, однако нужно, чтобы автоматически создавались и вспомогательные каналы для разграничения направлений работы по восстановлению во время крупномасштабных инцидентов.

Действия:
Разработать сборник сценариев, определить инструменты на случай крупномасштабных инцидентов и отрабатывать их применение с помощью моделирования. Необходимо определить и задокументировать типы инцидентов, которые могут считаться крупномасштабными и требуют реагирования на таком уровне. Необходимо наметить ключевые шаги по координации и создать инструменты, которые помогут менеджерам инцидентов и другим специалистам оптимизировать реагирование и начать восстановление. Менеджеры инцидентов вместе с командами будут регулярно моделировать инциденты, проводить учения и дорабатывать инструменты и документы, чтобы продолжать совершенствоваться.

Вывод 4. Следует усовершенствовать наши процессы обмена информацией

а) Мы лишились важных идентификаторов клиентов, что повлияло на связь и взаимодействие с затронутыми клиентами.
Скрипт удалил из наших рабочих сред не только сайты клиентов, но и ключевые идентификаторы клиентов (например, URL-адреса сайтов и контактную информацию системных администраторов). По этой причине (1) клиенты лишились возможности подавать заявки в службу технической поддержки через обычный канал предоставления поддержки; (2) нам потребовалось несколько дней, чтобы составить достоверный список ключевых контактных лиц клиентов (таких, как системные администраторы сайтов), испытавших на себе последствия сбоя, и начать работать с ними; (3) рабочие процессы поддержки, SLA, дашбоарды и процессы эскалации изначально не функционировали должным образом из-за беспрецедентного характера инцидента.

Во время сбоя эскалация заявок клиентов происходила по нескольким каналам (электронная почта, телефонные звонки, заявки от управляющих директоров, LinkedIn и другие социальные сети, а также заявки в службу поддержки). Из-за разрозненности инструментов и процессов в командах, взаимодействующих с клиентами, мы реагировали недостаточно быстро, испытывали сложности с комплексным отслеживанием этих эскалаций и составлением отчетов по ним.

b) У нас не было достаточно подробного сборника сценариев для информирования об инцидентах, чтобы справиться с таким уровнем сложности.
У нас не было сборника сценариев для информирования об инцидентах, в котором излагались бы принципы, роли и обязанности для достаточно быстрой мобилизации единой многофункциональной команды по информированию об инцидентах. Мы не смогли быстро и системно подтвердить наличие инцидента по нескольким каналам, особенно в социальных сетях. Следовало более широко распространять сообщения о сбое и повторять главное: потери данных удалось избежать и сбой произошел не по причине кибератаки.

Действия:

- **Усовершенствовать резервное копирование ключевой контактной информации.** Следует хранить резервные копии авторизованной контактной информации клиентов за пределами экземпляра продукта.
- **Модернизировать средства поддержки.** Мы должны создать механизмы прямой связи со службой технической поддержки для клиентов, у которых нет действительного URL-адреса сайта или идентификатора Atlassian.
- **Система и процессы эскалации заявок клиентов:** инвестировать в единую систему эскалации на основе учетных записей и рабочие процессы, которые позволяют хранить несколько рабочих объектов (заявок, задач и т. д.) в отдельном объекте клиентской учетной записи. Благодаря этому можно будет эффективнее согласовывать работу всех наших команд, взаимодействующих с клиентами, и мы получим более четкое представление о ней.
- **Ускорить перевод управления эскалациями на круглосуточный режим работы.** Мы должны предусмотреть возможность управления эскалациями в режиме 24/7 в планах по расширению географического охвата. Для этого в каждом крупном географическом регионе должен быть выделен специальный персонал, а также должны быть разработаны вспомогательные роли для оказания поддержки соответствующих продуктов, эксперты по продажам и руководители.
- **Дополнить наш сборник сценариев по информированию об инцидентах на основании нового опыта и регулярно пересматривать его.** Мы должны пересмотреть сборник сценариев, чтобы четко определить роли и линии коммуникации внутри компании. Необходимо применить принципы DACI к инцидентам и назначить дублеров для каждой роли на случай болезни, отпуска или других непредвиденных событий так, чтобы соответствующие специалисты были доступны в любой момент времени в любой день недели. Необходимо проводить ежеквартальную проверку, чтобы обеспечить постоянную готовность.

Действия (продолжение)
Составлять все сообщения об инцидентах на основании принятого шаблона. Следует рассказать, что произошло, указать, кого затронул инцидент, обозначить график восстановления, сообщить долю восстановленных сайтов и предполагаемый объем потерянных данных (а также степень нашей уверенности в этом предположении) и привести четкие инструкции по обращению в службу поддержки.

Заключительные замечания

Сбой устранен, все сайты клиентов восстановлены, но наша работа не окончена. На данном этапе мы внедряем описанные выше изменения, чтобы усовершенствовать наши процессы, повысить отказоустойчивость и предотвратить повторение подобной ситуации.

Atlassian — это компания, которая не прекращает учиться, и наши команды, безусловно, извлекли множество тяжелых уроков из этого случая. На основе полученного опыта мы внесем долгосрочные изменения в наш бизнес. В конечном итоге благодаря этому опыту мы станем сильнее и повысим качество обслуживания.

Мы надеемся, что сделанные из этого инцидента выводы будут полезны другим командам, которые прилагают множество усилий, чтобы бесперебойно обслуживать своих клиентов.

Я хочу поблагодарить тех, кто читает эту публикацию и учится вместе с нами, а также тех, кто является частью нашего обширного сообщества и команды Atlassian.

Шри Вишванат, технический директор

ATLASSIAN