

[КАК СТАТЬ АВТОРОМ](#)

680.84

Рейтинг

Яндекс

Как мы делаем Яндекс

[Подписаться](#)

slavashel

27 ноя 2024 в 07:54

Об инциденте с NTP-серверами

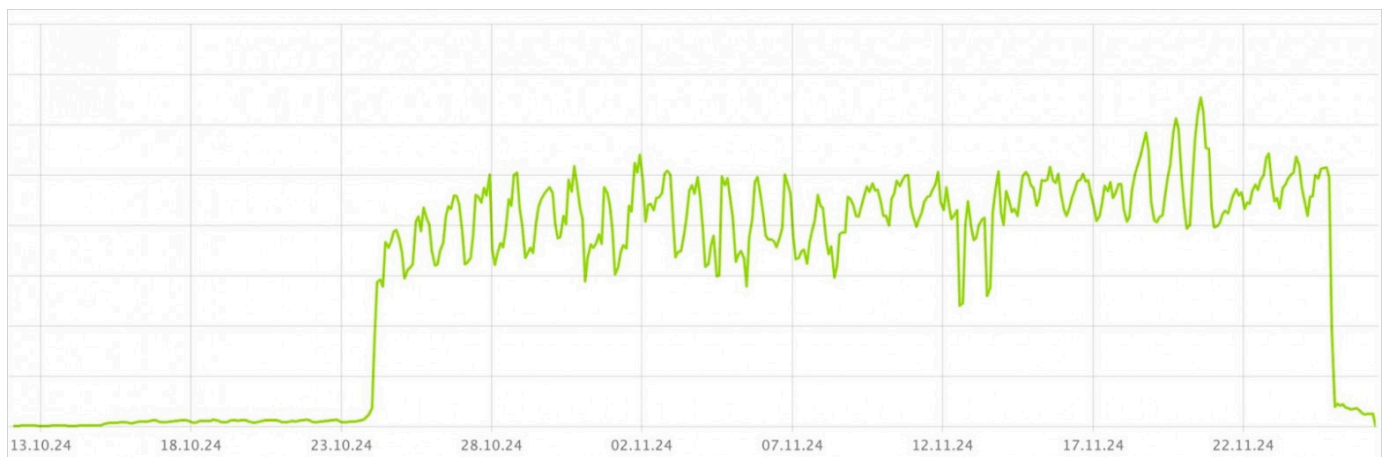
🕒 5 мин 👁 82K

Блог компании Яндекс, Сетевые технологии*, Высокая производительность*, IT-инфраструктура*

[Ретроспектива](#)

Недавно на Хабре вышла [статья](#) об аномальной нагрузке на публичные NTP-серверы в рунете. К сожалению, к этой ситуации привела ошибка в прошивке Яндекс Станций.

Как наши Станции синхронизируют время, что спровоцировало инцидент и как мы планируем поступить дальше — обо всём этом расскажем в статье. Хочется верить, что описание ситуации и принятые меры помогут не только нам, но и коллегам по индустрии избежать подобного в будущем.



Динамика числа отправляемых колонками NTP-запросов

Как Станции обновляют время

Время на наших Яндекс Станциях не должно существенно расходиться с реальным. Не только для того, чтобы часы, таймер и будильник работали правильно, но и для корректной работы других функций, которые требуют высокой точности. Например, алгоритм контрактизации требователен к точности внутренних часов. Благодаря ему на запрос пользователя отвечает только одна, ближайшая колонка, а не из соседней комнаты. Поэтому нам нужно обновлять время на них несколько раз за день.

Для этого несколько лет назад, когда направление умных устройств было во многом ещё экспериментом и стартапом, мы написали простой NTP-клиент для синхронизации часов. Он обращался к публичным серверам проекта NTPPool.org. На тот момент — один раз в минуту. Если же ответ до устройства по каким-то причинам не доходил (или колонка была перезагружена), клиент перезапрашивал информацию с других серверов каждые пять секунд до тех пор, пока не узнавал время.

Вскоре мы, конечно, поняли, что первоначально выбранный период обновлений в одну минуту избыточен, поэтому радикально увеличили его до пяти часов. В остальном никаких серьёзных изменений не было, NTP-клиент работал в течение нескольких лет без каких-либо инцидентов.

Как развивалась проблема

В начале октября мы подготовили обновление прошивки для наших колонок, в котором изменения в смежных с NTP-клиентом утилитах привели к сбою в его поведении. Как мы позднее выяснили, в коде была допущена ошибка, из-за которой колонки начинали перезапрашивать время с NTP-серверов каждые пять секунд вне зависимости от ответа, то есть игнорировали факт корректного получения времени. К сожалению, компонент был не полностью покрыт автотестами, поэтому на этапе тестирования ошибку мы не поймали.

Наш текущий процесс релиза прошивок для колонок выглядит так. После приёмки автоматическими и ручными тестами прошивка отправляется на 1% устройств в A/B-эксперимент, где валидируются основные метрики. После этого прошивка раскатывается на бóльший процент — на 10%, где повторно валидируются основные метрики и принимается решение о раскатке на всех.

15 октября мы раскатали прошивку на 10% устройств. Поэтапная раскатка обновления — это стандартная практика. Она предназначена в том числе для того, чтобы выявлять проблемы новых версий на ранней стадии. И прежде чем увеличивать процент раскатки, мы отматриваем метрики всех ключевых пользовательских сценариев. Но число генерируемых нами NTP-запросов исторически никогда не входило в метрики, требующие валидации, потому что NTP-клиент годами существовал практически без изменений и не приводил к проблемам.

К 24 октября новая прошивка докатилась до 100% колонок. Опять же, критичные для пользователей сценарии покрыты автоматическим мониторингом. Если что-то важное ломается, мы получаем уведомление. Но, как вы уже догадались, для NTP-клиента таких уведомлений настроено не было.

10 ноября мы получили первые жалобы из сообщества на избыточное количество запросов. Подобные обращения мы получали и раньше, но причиной всегда были сетевые ограничения на стороне пользователя (например, работодатель мог ограничивать запросы со стороны колонки в корпоративной сети). Жалоб было мало, действующий регламент поддержки не был рассчитан на подобные ситуации, поэтому обращения рассматривали не в самом высоком приоритете. К 20 ноября мы нашли ошибку, внесли исправление в код и начали готовить новый релиз.

В выходные 23–24 ноября ситуация с NTP-серверами обостряется: доступными остаются лишь четыре сервера. К этому моменту мы уже начали раскатывать релиз с исправлением на 10% устройств. Но фактически обновления происходят по ночам. Мы поняли, что нельзя ждать ближайшей ночи, при этом форсировать этот процесс опасно, поэтому решили пойти на альтернативные, срочные меры.

В воскресенье мы выпустили хотфикс в виде новой рантайм-конфигурации для NTP-клиента, который увеличивал период перезапроса с 5 до 600 секунд, уменьшая нагрузку на серверы в 120 раз. Хотфикс не исправлял проблему полностью, но был единственным быстрым способом снять чрезмерную нагрузку с NTP-серверов. Это был рискованный шаг, потому что увеличение периода перезапроса могло сказаться на работе Станций: если колонка после перезагрузки не смогла с первого раза узнать текущее время, то в течение следующих 600 секунд часы и все другие функции, связанные со временем, будут недоступны.

В тот же день число доступных серверов начинает расти благодаря призыву @kkursor в уже упомянутой выше статье. Сообщество пришло на помощь, хотфикс раскатился, ситуация начала стабилизироваться. Это дало нам время для плановой раскатки обновления с исправлением ошибки в коде.

Что планируем делать дальше

Когда мы начинали, даже частые запросы наших умных колонок к серверам терялись в общем потоке и не создавали избыточной нагрузки на инфраструктуру сообщества. С тех пор устройств стало на порядки больше, а генерируемый ими трафик — существенным для сообщества.

Поэтому мы запланировали выделить ресурсы в общий пул NTP-серверов. Это займёт некоторое время, потому что наши дата-центры удалены от основных точек обмена

трафиком, а для NTP-серверов RTT (Round Trip Time) это — ключевой фактор качества. Мы установим и запустим мощности на основных точках обмена трафиком.

Для наших устройств мы заведём именную зону в соответствии с гайдлайнами проекта NTPPool.org для бóльшей прозрачности. Генерируемый ими трафик будет локализован на наших NTP-серверах, если мы продолжим полагаться на публичную инфраструктуру проекта.

Ещё мы добавим метрики, связанные с NTP, на этап валидации А/Б-экспериментов, а также реалтайм-уведомления о неполадках в этом компоненте. В том числе расширим сценарии мониторинга исходящего и входящего служебного трафика устройства в целом. Также поработаем над каналами коммуникаций и поддержкой, чтобы подобные проблемы быстрее до нас эскалировались.

Мы верим, что запланированные меры помогут сделать этот важный для рунета компонент ещё более надёжным и стабильным.

Особенно хотим сказать большое спасибо участникам сообщества, которые первыми заметили проблему и вовремя привлекли к ней внимание.

Теги: яндекс, яндекс станция, ntp

Хабы: Блог компании Яндекс, Сетевые технологии, Высокая производительность, IT-инфраструктура



Яндекс

Как мы делаем Яндекс

Github



52

0

Карма Рейтинг

Вячеслав Шелковников @slavashel

Head of Alice C++ SDK

Подписаться



Комментарии 442

Публикации

ЛУЧШИЕ ЗА СУТКИ ПОХОЖИЕ

**DmitryOlkhovoi**

20 часов назад

Меня заставили повайбкодить

**Сложный**

18 мин



22K

Кейс

**+89**

96



92

**RationalAnswer**

12 часов назад

Тотальный блэкаут на юге Европы, а также чудеса нейро-лизоблюдия от ChatGPT



11 мин



11K

Дайджест

**+33**

12



62

**lozhnikov**

22 часа назад

Как доказывали теорему о четырех красках. Часть 1

**Простой**

11 мин



2.8K

FAQ

**+29**

23



1

**alizar**

8 часов назад

Ян Лекун, создатель LeNet, формата DjVu и адвокат опенсорса

**Средний**

7 мин



1.3K

Обзор

**+26**

8



6

**MrSotnik**

8 часов назад