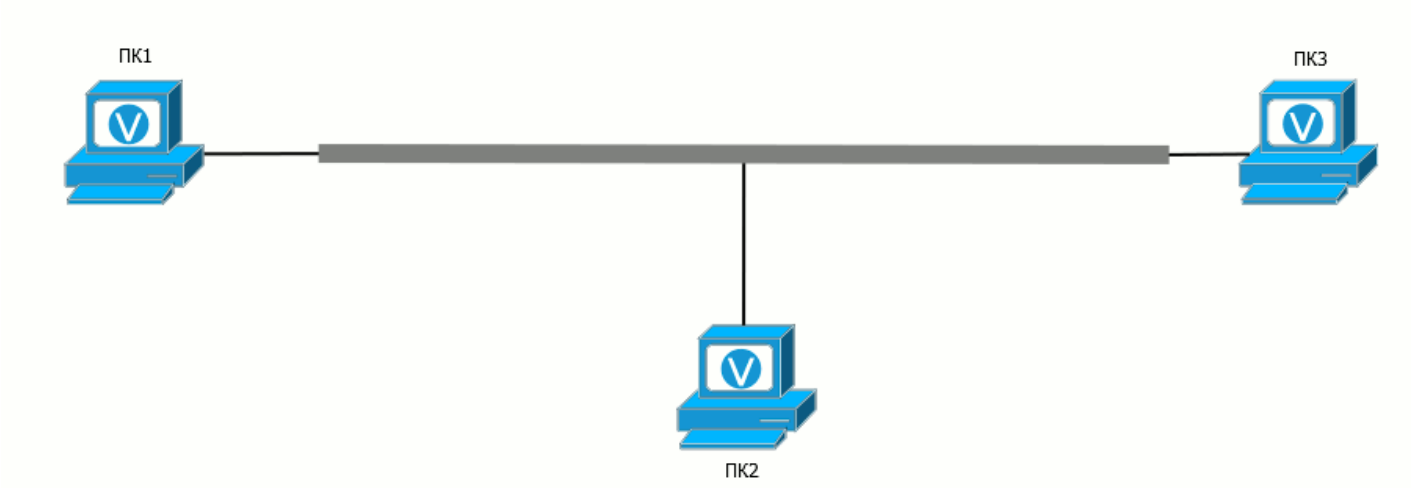
100 метров Ethernet

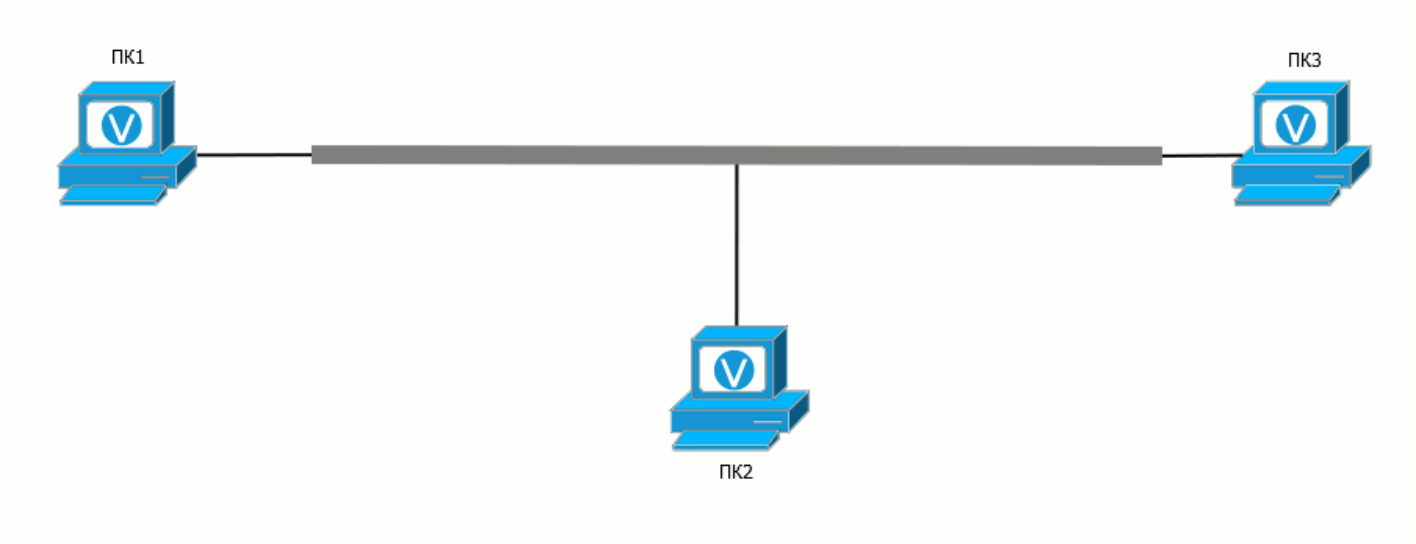
При подготовке к статье с каверзными вопросами я наткнулся на интересный вопрос - откуда взялось ограничение в 100 метров на длину Ethernet-сегмента. Мне пришлось погрузиться глубоко в физику и логику процессов, чтобы приблизиться к пониманию. Часто говорят, что на большой длине кабеля начинаются затухания и данные искажаются. И, в общем-то, это правда. Но есть и другие причины для этого. Попытаемся рассмотреть их в данной статье.

CSMA/CD

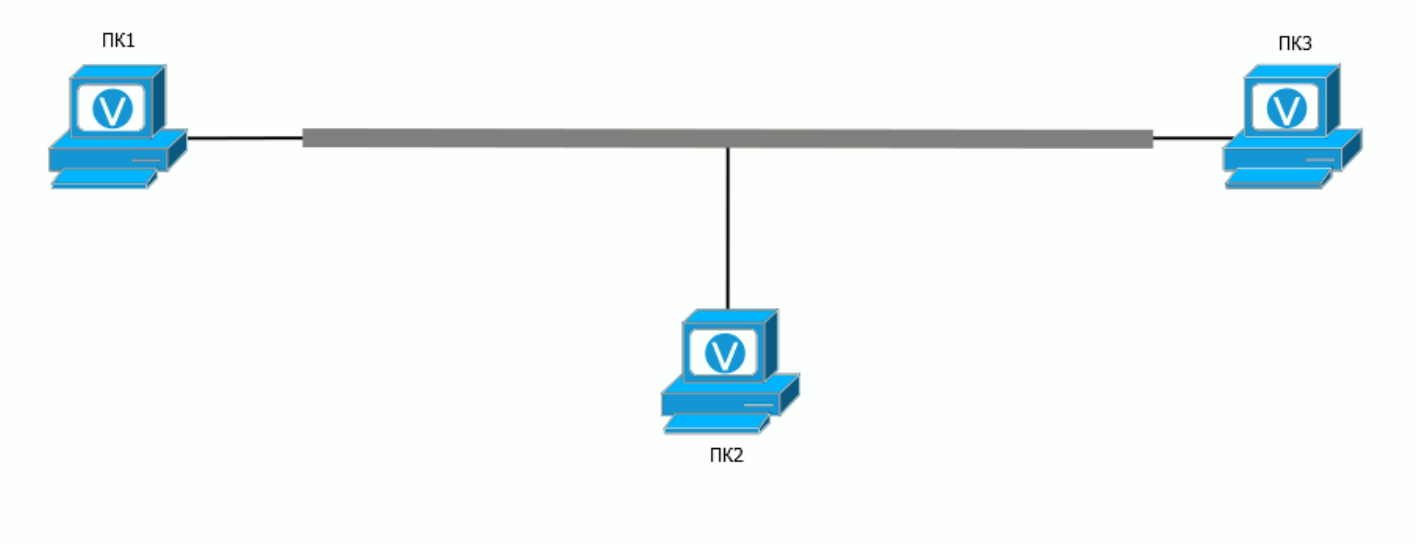
Причина кроется в технологии CSMA/CD - *Carrier Sense Multiple Access with Collision Detection*. Если вдруг кто-то не знает, то это когда у нас одна шина (одна среда передачи данных), к которой подключено несколько станций (*Multiple Access*). Каждая станция следит за состоянием шины - есть ли в ней сигнал от другой станции (*Carrier Sense*). Если вдруг два устройства начали передавать в один момент, то оба они должны это обнаружить (*Collision Detection*). Да, всё это касается полудуплексных сетей. Поэтому если у вас взгляд устремлён исключительно в светлое 10-гигабитное будущее, эта статья не для вас. В первую очередь, я хочу, чтобы все понимали, что скорость передачи сигнала в среде никоим образом не зависит от применяемого стандарта. Хоть в Ethernet (10Мб/с), хоть в 10Gbit Ethernet скорость распространения импульса в медном кабеле - примерно 2/3 скорости света. Как здорово написали в одном холиварном треде: вы можете говорить быстро или медленно, но скорость звука от этого не меняется. Теперь обратимся к сути CSMA/CD. В современных сетях коллизии исключены, потому что у нас уже нет общей шины и практически всегда все устройства работают в полнодуплексном режиме. То есть у нас всего лишь два узла на конце одного кабеля и отдельные пары для приёма и передачи. Поэтому механизма CSMA/CD уже нет в 10Gbit Ethernet. Однако рассмотреть его будет полезно, так же, как например, изучать RIP, который, вроде, никому уже и не нужен, но прекрасно иллюстрирует принцип работы дистанционно-векторных протоколов маршрутизации. Итак, предположим, что к общей шине у нас подключено 3 устройства. ПК 1 начинает передавать данные на ПК3 (запустил импульс в шину). Разумеется, в общей шине сигнал пойдёт не только на ПК3, но всем подряд. ПК2 тоже хотел бы передать, но видит волнения в кабеле и ожидает. Когда сигнал от ПК1 до ПК3 прошёл, может начинать передавать ПК2.



Это пример работы Carrier Sense. ПК2 не передаёт, пока видит сигнал в линии. Теперь другая ситуация. ПК1 начал передавать данные ПК3. А до ПК2 сигнал не успел дойти, он тоже решил начать передавать. Где-то в середине сигналы пересеклись и испортились. ПК1 и ПК2 получили покорёженный сигнал и поняли, что эту порцию данных нужно отправить заново. Каждая станция выбирает случайным образом период ожидания, чтобы снова не начать отправлять одновременно.

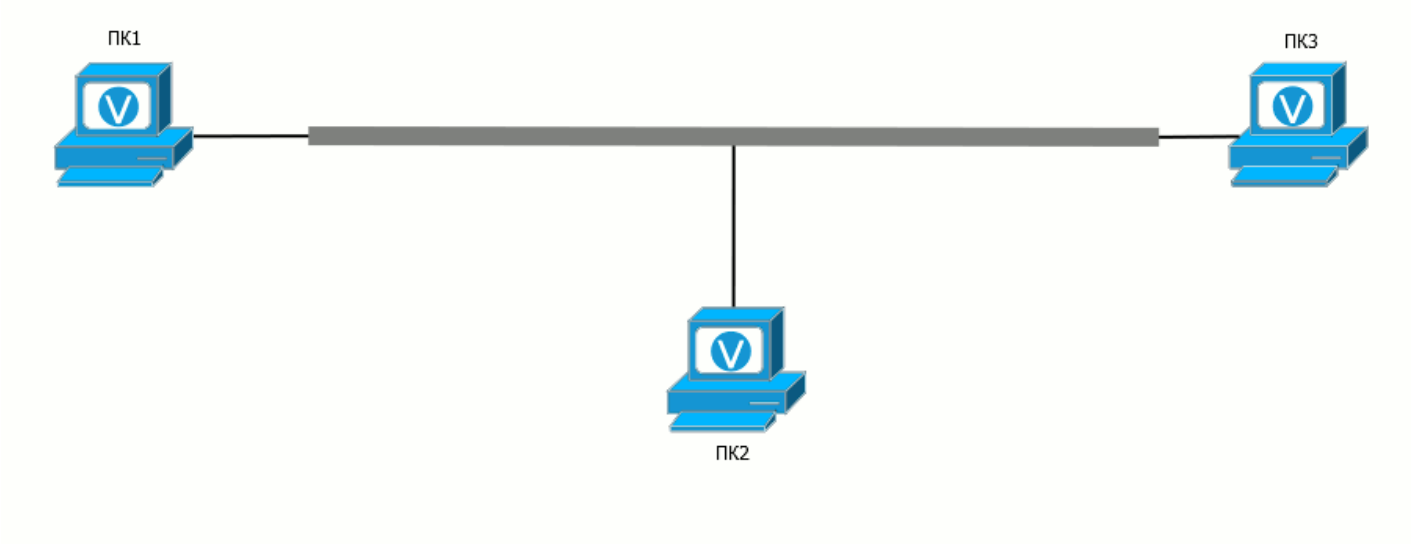


Это пример работы Collision Detection. Чтобы одна станция не оккупировала шину, между кадрами есть промежуток длиной 96 битов (12 байтов), который называется Inter Frame Gap (IFG). То есть, например, ПК1 передал кадр, потом ждёт некоторое время (время, за которое он успел бы передать 96 битов). И отправляет следующий и т.д. Если ПК2 захочет передавать, то он сделает это как раз в таком промежутке. Так же ПК3 и так по очереди. То же самое правило работает и в том случае, когда у вас не общая шина, а один кабель, где к двум концам подключены две станции, и они передают данные в полудуплексном режиме. То есть передавать данные в каждый момент времени может только одна из них. Передаёт ПК2, как только линия освободилась, передаёт ПК1, линия освободилась - передаёт ПК2 и так далее. То есть тут нет какой-то чёткой временной синхронизации, как, например, в TDD, когда для каждого конца выделены определённые промежутки передачи. Таким образом, достигается более гибкое использование полосы: Если ПК1 ничего передавать не хочет, то ПК2 не будет простаивать в ожидании своей очереди.



Проблема

А что если представить себе такую неловкую ситуацию?



То есть ПК1 закончил передачу своей порции данных, но она ещё не успела дойти до ПК2. Последний не видит сигнала в линии и начинает передавать. Бац! Где-то в середине ДТП. Данные покорёжились, сигнал дошёл до ПК 1 и ПК2. Но, обратите внимание на разницу - ПК2 понял, что произошла коллизия и перестал передавать данные, а ПК1 ничего не понял - у него-то передача уже закончилась. Фактически он просто получил битые данные, а свою задачу по передаче кадра как бы выполнил. Но данные потерялись на самом деле - ПК3 также получил искажённый коллизией сигнал. Где-то потом гораздо выше по ступеням OSI отсутствие данных заметит TCP и перезапросит эту информацию. Но представьте, сколько времени на это будет потеряно?

> Кстати, когда на интерфейсах у вас растёт количество ошибок CRC - это верный признак коллизий - приходят битые кадры. Значит, скорее всего, не согласовался режим работы интерфейсов на разных концах.

Вот именно для исключения такой ситуации в Ethernet ввели одно условие: в тот момент, когда первый бит данных будет получен на самой дальней стороне шины, станция ещё не должна передать свой последний бит. То есть кадр должен как бы растянуться на всю длину шины. Это самое распространённое описание, но фактически оно звучит несколько иначе: если коллизия произошла на самом дальнем от отправителя участке шины, то информация об этой коллизии должна достигнуть отправителя ещё до того, как он передал свой последний бит. А это разница в 2 раза, между прочим, по сравнению с первым приведённым условием. Это гарантирует, что даже если случится коллизия, все её участники будут однозначно в курсе. И это очень здорово. Но каким образом этого добиться? И тут мы вплотную приближаемся к вопросу о длине сгемента. Но прежде, чем дать ответ на вопрос про длину, придётся немного окунуться в теорию сетей и для начала введём понятие bit time (термин "битовое время" не прижился). Эта величина означает, сколько нужно времени интерфейсу, чтобы выпульнуть в среду 1 бит. То есть если Fast Ethernet в кабель отправляет 100 000 000 битов в секунду, значит, bit time равен 1b/100 000 000 b/s=10^-8 с или 10 наносекунд. Каждые 10 наносекунд Fast Ethernet порт может отправлять в среду один бит. Для сравнения Gigabit Ethernet отправляет 1 бит каждую наносекунду, старые диал-ап модемы могли отправлять 1 бит каждые 18 микросекунд. Скорострельное оружие Metal Storm MK5 теоретически способно выпускать одну пулю каждые 60 микросекунд. Пулемёт калашникова выпускает 1 пулю каждые 100 миллисекунд.

> Если говорить об IFG, то станция должна делать паузу именно в 96 бит-таймов перед отправкой каждого кадра. Fast Ethernet, например, должен выждать 960 наносекнуд (0,96 микросекунды), а Gbit Ethernet 96 наносекуд

Итак, для выполнения условия вводится понятие кванта или Slot time - минимальный размер блока данных, который можно передавать по сети в Ethernet. И именно этот квант должен растянуться на весь сегмент. Для Ethernet и Fast Ethernet выбран минимальный размер - 64 байта - 512 бит. Для его передачи порту FE понадобится 10 нс*512 = 5120 нс или 5,12 мкс.

> Отсюда и ограничение в 64 байта на минимальный размер Ethernet-кадра.

То есть у блока данных 64 байта будет 5,12 мкс на путешествие по шине и возврат к отправителю в случае коллизии. Попробуем просчитать расстояние в лоб: (5,12 * 10^-6)*(2/3*3*10^8)/2=512 метров. Поясню формулу: время путешествия (5,12 мкс переведённые в секунды) * 2/3 скорости света (скорость распространения сигнала в медной среде в м/с) и делим на 2 - для того, чтобы предусмотреть самый худший случай коллизии, когда сигналу придётся пройти весь путь назад до отправителя. Вроде бы и цифра знакомая - 500 метров, но проблема в том, что ограничение для Fast Ethernet - 100 метров до хаба (200 до самой дальней станции). Здесь вступают в игру задержки на

концентраторах и повторителях. Говорят, что они все просчитаны и учтены в конечной формуле, но следы теряются, сколько я ни пытался найти эту формулу расчёта с результатом в 100 метров, найти не удалось. В итоге известно, чем ограничение обусловлено, но не откуда взялась цифра 100.

Gigabit Ethernet

При разработке Gbit Ethernet встал очень важный вопрос - время передачи одного бита составляло уже 1 нс и на передачу одной порции данных нужно уже всего лишь 0,512 мкс. Даже при расчёте в лоб моей формулой без учёта задержек получается длина 50 метров (и 20 метров с учётом этих величин). Очень мало и потому было решено, вместо уменьшения расстояния (как было в случае с переходом Ethernet->Fast Ethernet), увеличить минимальный размер данных до 512 байтов - 4096 бит. Время передачи такой порции данных осталось примерно таким же - 4 секунды против 5. Тут, конечно, есть ещё момент, что не всегда получается набрать такой размер - 4 кБ данных, поэтому в конце кадра, после поля FCS добавляется недостающий объём данных. Учитывая, что мы давно отказались от общей шины, у нас раздельная среда для приёма и передачи, и коллизий как таковых нет, всё это выглядит костылями. Поэтому в стандарте 10 Gbit Ethernet от механизма CSMA/CD отказались вовсе.

Преодоление ограничений по длине

Итак, всё вышеуказанное касалось устаревших полудуплексных сетей с общей шиной. Какое это имеет отношение к настоящему моменту, спросите вы? Можем тянуть мы километры UTP или не можем? К сожалению, всё-таки стометровое ограничение имеет и другую природу. Даже на 120 метрах с обычным кабелем в большинстве случаев многие коммутаторы не смогут поднять линк. Это обусловлено и мощностью портов коммутаторов и качеством кабеля. Дело и в затухании, и в наводках, и в искажении сигнала при передаче. Обычная витая пара подвержена влиянию электромагнитных помех и не гарантируют защиту передаваемой информации. Но, прежде всего, давайте посмотрим на затухание. Типичная наша витуха UTP имеет минимум по 27 витков на каждый метр и передаёт данные на частоте 100 МГц. Так называемое погонное затухание - это ослабление сигнала на каждом метре среды. Согласно стандартам затухание не должно превышать 24 Дб. В среднем это значение около 22 Дб для обычного UTP-кабеля, что означает затухание изначального сигнала в 158 раз. Получается, что затухание на 1 Дб происходит каждые 4,5 метра. Если же взять длину кабеля в 150 метров, то затухание получается уже примерно 33 Дб и исходный сигнал уменьшится в 1995 раз. Что уже весьма существенно. Плюс к этому добавляется взаимное влияние пар - переходное затухание. Так называется процесс, когда в параллельных проводниках возникают наводки, то есть часть энергии тратится на то, чтобы возбудить ток в соседнем кабеле. Учтём возможные помехи от силовых кабелей, которые могу проходить рядом, и ограничение в 100 метров становится совершенно логичным.

> Почему тогда такого ограничения не было в коаксиальных сетях? Дело в том, что затухание в кабеле зависит от сопротивления/сечения кабеля и частоты. Вспомним теперь, что толстый Ethernet использует кабель с сердечником 2,17 мм. Плюс Ethernet на коаксиальном кабеле работал на частоте 10 Мгц. А чем больше частота, тем выше затухание. Почему вы думаете аналоговый радиосигнал передаётся к антеннам не по такой удобной витухе, а по толстенным фидерам? Кстати, слово Base в стандартах Ethernet означает Baseband и говорит о том, что одновременно может передавать данные через среду только одно устройство, не используется модуляция/мультиплексирование. В противовес ему Broadband накладывает несколько разных сигналов на одну несущую, а с другой стороны каждый отдельный сигнал из несущей извлекается.

На самом деле, учитывая, что затухание обусловлено характеристиками и качеством кабеля, можно достигнуть значительно более радостных результатов, используя более подходящий. Например, с помощью кабеля П-296 или П-270 можно преодолеть даже трёхсотметровый рубеж. Разумеется, это 100 Мб/с в полному дуплексе. Для гигабита уже другие требования. И вообще, чем выше скорость передачи, тем больше параметров приходится учитывать, собственно поэтому в 10Gbit Ethernet поддержка медной среды есть только номинально, а предпочтение отдано оптике.

Итоги и ссылки

В общем, подводя итог всему вышесказанному, цифра в 100 метров - это с хорошим таким запасом, который гарантирует работу даже в полудуплексе на кабеле не лучшего качества. Обусловлена она затуханием и работой механизма CSMA/CD. Данные, использованные в статье:

- Передача сигналов по витой паре
- "Возвращение к основам". Описание различных типов затуханий и проблем витой пары
- Ethernet на витой паре. Преодолеваем барьер в 100 метров.
- Bit time Slot time

- Фундаментальные аспекты технологий локальных сетей

Читайте также обзор GPON vs Ethernet.