ГЛАВА
2 КОМПЬЮТЕР КАК СТУДИЯ

[Уайатт, Эмиес]

Цифровая звуковая рабочая станция (англ. *digital audio workstation*, DAW; по-русски чаще говорят просто «станция») — это программа, которая заменяет сразу магнитофон, монтажный стол, микшерный пульт и стойку с приборами обработки. Reaper, Pro Tools, Nuendo, Logic и другие отличаются интерфейсом и набором функций, но устроены одинаково. Они читают звуковые файлы с диска, складывают их по правилам, записанным в проекте, пропускают через плагины — подключаемые программы обработки — и отдают результат на выход. В этой главе примеры даются в основном на Reaper и Pro Tools, но всё сказанное относится к любой станции.

Зачем знать, как станция устроена внутри, если можно просто нажимать кнопки? Затем, что у работы в компьютере есть свои ловушки, и на слух их причину не найти. Актёр на записи жалуется на эхо в наушниках. Голос после подключения параллельной обработки вдруг звучит как из трубы. Сцена, сведённая в одной программе, в другой звучит иначе, хотя ничего не меняли. Аккуратно нарисованные изменения громкости исчезают сами собой.

Как станция проигрывает звук

Файлы на диске не трогаются. Когда монтажёр режет, двигает или укорачивает фрагмент (в Reaper он называется *item*, в Pro Tools — *clip*, в других программах — *region*), меняется только запись в проекте: «взять файл такой-то с такой-то секунды, длиной столько-то, поставить на такое-то место». Это называется *неразрушающим монтажом* (*non-destructive editing*). Оригинал всегда цел, и любое решение можно отменить. Обработка, которая создаёт новый файл (например, шумоподавление, выполненное заранее), тоже обычно сохраняет исходник, а новый файл просто подставляет на его место.

Во время воспроизведения станция работает порциями. Она берёт из файлов кусок звука длиной, скажем, 256 отсчётов на каждую дорожку, пропускает его через все плагины, складывает и отдаёт звуковой карте. Пока карта проигрывает эту порцию, программа уже считает следующую. Такая порция называется *буфером* (*buffer*). Если компьютер не успел досчитать следующую порцию к моменту, когда закончилась текущая, звук прерывается: появляется щелчок или короткий провал. В Reaper такие сбои видны в окне производительности (*Performance Meter*), в других программах встречаются названия *xrun* (общее название для переполнения и опустошения буфера, англ. *overrun* и *underrun*) или *dropout* — «выпадение».

Отсюда простое правило: чем больше буфер, тем больше у компьютера времени в запасе и тем меньше риск сбоев. Но у большого буфера есть цена — задержка.

Задержка

Звук, поданный на вход, проходит через входной буфер, потом через выходной, а кроме того через сами преобразователи. Поэтому общая задержка от микрофона до наушников примерно равна

$$t \approx \frac{2B}{f_s} + t_{\text{АЦП/ЦАП}}, \qquad (2.1)$$

где B — размер буфера в отсчётах, f_s — частота дискретизации, а последнее слагаемое — задержка аналого-цифрового и цифро-аналогового преобразователей, обычно около 1–2 мс. При 48 кГц буфер 64 отсчёта даёт около 4 мс, 256 отсчётов — около 12 мс, 1024 — около 44 мс (рис. 2.1). Это лишь нижняя оценка: драйвер и звуковая карта часто добавляют собственные буферы, поэтому реальная величина бывает заметно больше. Точное значение программа обычно показывает в настройках звукового устройства, а надёжнее всего измерить его, записав сигнал с выхода обратно на вход.

Задержка важна только тогда, когда человек слышит себя через компьютер: при записи голоса, шумов или закадрового текста. При монтаже и сведении никто себя не слышит, и задержка значения не имеет: программа просто начинает

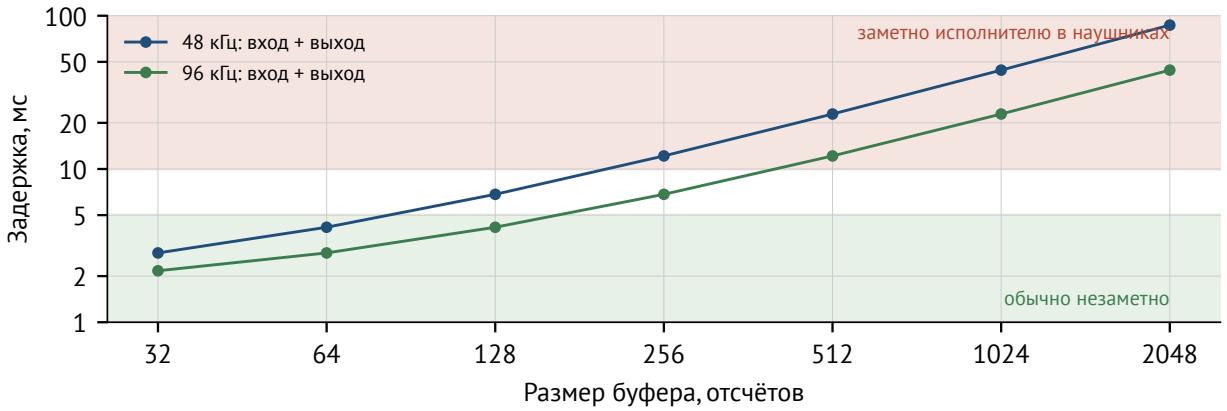

Рис. 2.1. Примерная задержка от входа до выхода по формуле (2.1) при разном размере буфера; задержка преобразователей принята 1,5 мс. Границы заметности условны: для речи в наушниках задержка в несколько миллисекунд обычно незаметна, а свыше примерно 10 мс многие исполнители начинают её слышать.

воспроизведение на долю секунды позже нажатия кнопки.

Как это выглядит на практике, хорошо видно на записи. Идёт переозвучание, проект уже тяжёлый: на соседних дорожках стоят шумоподавители и реверберации. Буфер маленький, и в записи начинают проскакивать щелчки — компьютер не успевает. Буфер увеличивают до 1024 отсчётов, щелчки пропадают, но теперь актёр жалуется, что слышит себя «как в эхо», сбивается и начинает говорить медленнее. Задержка выросла до 40 с лишним миллисекунд, а свой голос человек слышит ещё и изнутри, через кости черепа, без всякой задержки. Две копии голоса с таким сдвигом мозг воспринимает как эхо.

Противоречие разрешается тем, что запись и сведение требуют разного:

- при записи — маленький буфер, 64–128 отсчётов, и минимум тяжёлых плагинов на дорожке, куда идёт запись;
- при монтаже и сведении — большой буфер, 512–2048 отсчётов, и тогда можно включать много обработки.

Размер буфера задаётся в настройках звукового устройства: в Reaper — в разделе *Audio → Device*, в Pro Tools — в окне *Playback Engine* (параметр *H/W Buffer Size*). Помогают и сами станции. Reaper заранее, с опережением, просчитывает дорожки, на которые сейчас не идёт запись (*anticipative FX processing*), поэтому тяжёлые плагины на них почти не мешают записи. Многие звуковые карты умеют направлять входной сигнал прямо в наушники, минуя компьютер (*direct monitoring*, прямой мониторинг). Тогда исполнитель слышит себя без задержки, а компьютер только записывает.

Запись в станции

Голос, шумы или реплики под изображение станция записывает одинаково, и у записи есть несколько приёмов, без которых студийная работа почти невозможна.

Дубли. Одну и ту же фразу редко записывают с первого раза, и станция хранит все попытки, а не только последнюю. В Pro Tools каждая попытка может лечь в отдельный *плейлист* (*playlist*) дорожки, в Reaper — отдельным *дублем* (*take*) внутри фрагмента или на отдельной полосе дорожки. Потом из них собирают лучший вариант по кускам: начало фразы из третьего дубля, конец из пятого. Такую сборку называют *компингом* (*comping*, от *compiling* — «составление»). Поэтому неудачный дубль не удаляют сразу: через неделю именно в нём может найтись нужная интонация.

Врезка. Если в хорошем дубле не получилось одно слово, всю фразу не переписывают. Включают *врезку* (*punch-in*): станция начинает воспроизведение чуть раньше нужного места, актёр слышит подводку и входит в ритм, а запись включается только на заданном отрезке и выключается после него. Запас перед местом записи называют *предпрокруткой* (*pre-roll*). Без неё актёру пришлось бы начинать с полуслова, и врезка была бы слышна.

Запись по кольцу. Короткий отрезок изображения проигрывается по кругу, и каждый проход станция сохраняет как новый дубль, не останавливаясь. Актёр пробует фразу снова и снова, а

звукорежиссёр потом выбирает лучшие попытки. Это прямое продолжение старой практики колец, о которой шла речь в главе 4. В Pro Tools для этого есть режим записи по петле (*Loop Record*), в Reaper — запись с выделенным отрезком и повтором, при которой каждый проход становится отдельным дублем.

Сигналы для актёра. Чтобы попасть в изображение, актёру нужно знать, когда начинать. Обычно перед репликой звучат три коротких сигнала-отсчёта с интервалом в секунду, а сама реплика начинается там, где прозвучал бы четвёртый. Иногда вместо них или вместе с ними на экране идёт полоса, которая доходит до края кадра в момент начала реплики. Когда новый голос заменяет существующий, актёр слушает в наушниках исходную реплику как образец ритма и интонации. При этом исходный звук должен идти только в наушники, а не в записываемый микрофон.

Задержка плагинов и её компенсация

Некоторым плагинам нужно «заглянуть вперёд». Лимитер с упреждением (*lookahead*) должен увидеть пик до того, как тот придёт, чтобы плавно убавить уровень. Эквалайзер с линейной фазой работает с длинными фрагментами сигнала. Шумоподавители анализируют звук кусками. Все они выдают результат с опозданием на несколько миллисекунд, а иногда и на сотню.

Если такой плагин стоит на одной дорожке, а на соседних нет, обработанная дорожка отстанет от остальных. Чтобы этого не произошло, станция выполняет *компенсацию задержки плагинов* (*plugin delay compensation*, PDC): она узнаёт у каждого плагина его задержку и задерживает все остальные дорожки на ту же величину. В итоге всё снова совпадает. В Reaper компенсация работает всегда, в Pro Tools её включают в меню *Options → Delay Compensation*.

Компенсация работает почти всегда, но есть случаи, когда она не справляется: плагин неправильно сообщает свою задержку, задержка меняется при переключении режима во время воспроизведения или сигнал идёт по необычному пути, например через внешний прибор. Тогда две копии одного сигнала, которые должны совпадать, расходятся во времени, и при сложении возникает *гребёнчатый фильтр* (*comb filter*): провалы на определённых частотах, которые на графике выглядят как зубья гребёнки (рис. 2.2). При сдвиге 1,5 мс первый провал приходится на 333 Гц, и голос звучит пусто, как из трубы. Если сдвиг ещё и меняется во времени, провалы «гуляют» по частотам, и звук приобретает характерный свистящий, «реактивный» призвук, который в музыке используют намеренно под названием *фланжер* (*flanger*). Подробнее об этом явлении — в главе 12 «Монтажа диалогов».

Чаще всего это случается при *параллельной обработке*. Монтажёр отправляет диалог на отдельную шину с компрессором и подмешивает её к исходному сигналу, чтобы голос стал плотнее. Без шины всё звучит нормально, с шиной голос становится глухим и «трубным». Брак ли это? Да: на шине стоит плагин с задержкой, которую станция по какой-то причине не скомпенсировала, и две копии голоса складываются со сдвигом. Проверить это можно простым приёмом. Плагин на шине переводят в нейтральное состояние, при котором он не меняет звук, но продолжает вносить задержку (например, порог компрессора ставят выше сигнала), а у шины переворачивают полярность. Если компенсация

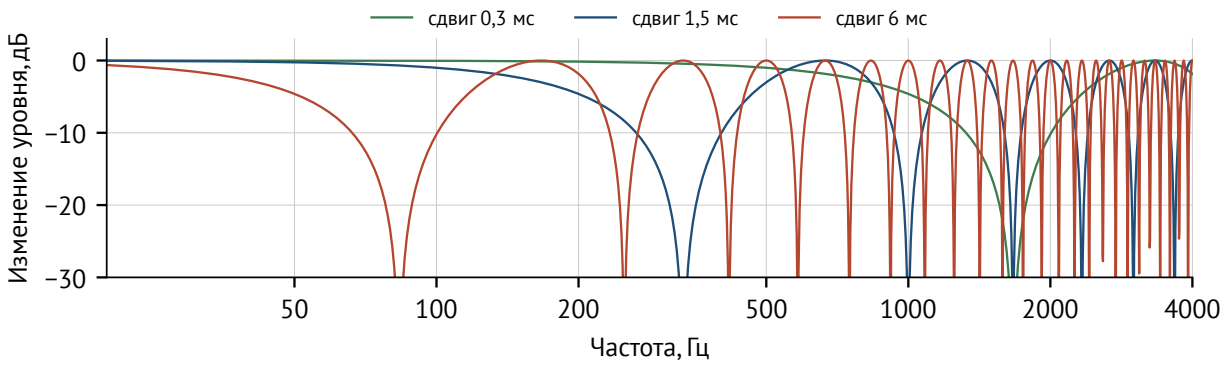

Рис. 2.2. Что происходит, если одна из двух одинаковых копий сигнала отстала на несколько долей миллисекунды. Провалы появляются на частотах $f = (2k-1)/(2\tau)$, где τ — сдвиг. Расчётные кривые для двух копий одинакового уровня.

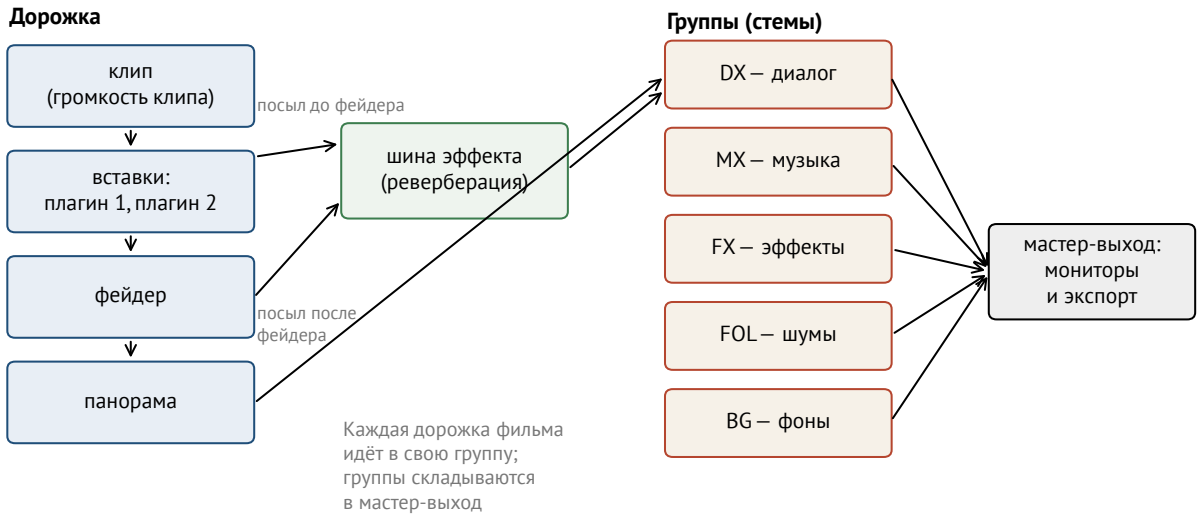

Рис. 2.3. Упрощённая схема пути сигнала в станции. Звук клипа проходит через вставки, фейдер и панораму дорожки, затем попадает в группу своего типа, а группы складываются в мастер-выход. Посылы отводят копию сигнала на шины эффектов.

точна, исходный сигнал и шина погасят друг друга, и звук почти исчезнет. Если остаётся громкий окрашенный остаток, что-то сдвинуто. Тогда проверяют, включена ли компенсация, заменяют подозрительный плагин или вручную задерживают исходную дорожку на нужную величину. Простое дублирование дорожки вместе с плагином для такой проверки не годится: обе копии задержаны одинаково и погасят друг друга при любой компенсации.

> **Для любопытных**
>
> Если две одинаковые копии сигнала сдвинуты на время τ, их сумма имеет амплитудную характеристику $|1 + e^{-j2\pi f\tau}| = 2\,|\cos(\pi f\tau)|$. Она обращается в ноль на частотах $f_k = (2k-1)/(2\tau)$: при $\tau = 1$ мс это 500, 1500, 2500 Гц и так далее. На частотах $f = k/\tau$ сумма удваивается (+6 дБ). Задержку в отсчётах переводят в миллисекунды делением на частоту: 64 отсчёта при 48 кГц — это 1,33 мс.

Путь сигнала

Чтобы сведение было предсказуемым, нужно ясно представлять, через что проходит звук. В большинстве станций путь одинаков (рис. 2.3).

Первое звено — сам фрагмент. У него есть своя громкость, свои плавные нарастания и спады (*fades*). Громкость фрагмента меняют, когда нужно выровнять отдельную реплику или эффект, не трогая всю дорожку.

Второе звено — *вставки* (*inserts*), то есть плагины, через которые проходит весь сигнал дорожки по очереди. Порядок имеет значение: эквалайзер до компрессора и после него работает по-разному, потому что компрессор реагирует на то, что ему подали.

Третье — фейдер и панорама. Фейдер задаёт уровень дорожки в сведении, панорама — положение звука между громкоговорителями, в стерео или в многоканальном поле.

Затем сигнал попадает на *шину* (*bus*). Шина — это просто место, где складываются несколько сигналов. В Reaper шиной служит обычная дорожка, на которую направлены другие, и самый удобный способ её создать — *папка* (*folder*): дорожки, вложенные в папку, автоматически идут через неё. Кроме основного пути есть *посылы* (*sends*): дорожка отправляет копию сигнала на другую шину, например на общую реверберацию. Посыл может быть *до фейдера* (*pre-fader*), тогда его уровень не зависит от фейдера дорожки, или *после фейдера* (*post-fader*), тогда реверберация убавляется вместе с самим звуком. Для реверберации обычно нужен второй вариант. Первый нужен, например, для отдельного сигнала в наушники исполнителя, который не должен меняться, когда звукорежиссёр двигает фейдеры.

В фильме дорожки собирают в группы по типу звука. Каждая группа сводится в свою шину, а затем выводится отдельным файлом — *стемом* (*stem*). Обычный набор с принятыми англий-

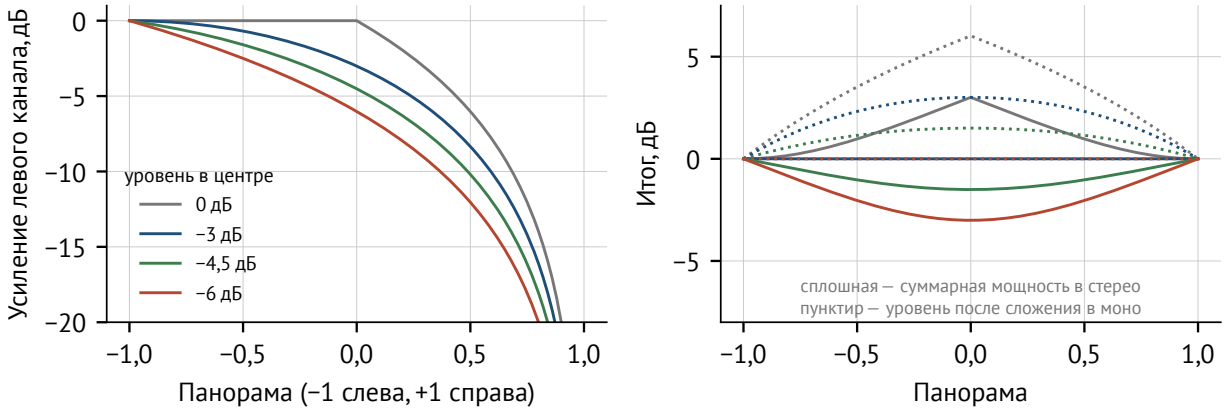

Рис. 2.4. Законы панорамы. Слева: усиление левого канала при перемещении звука; подпись — насколько убавлен каждый канал в центре. Справа: суммарная громкость в стерео (сплошные линии) и уровень после сложения каналов в моно (пунктир). Ни один закон не даёт постоянства сразу в обоих случаях.

скими сокращениями: диалог (DX, от *dialogue*), музыка (MX, *music*), эффекты (FX, *effects*), шумовое озвучание (Foley) и фоны (BG, *background*). Отдельные стемы нужны, чтобы потом сделать международную версию без диалога, телевизионную версию с другим балансом или исправить одну группу, не пересводя весь фильм. Об этом подробнее — в главе 14.

Закон панорамы

Есть одна настройка, которую легко не заметить, хотя она влияет на весь баланс. Когда звук перемещают панорамой из одного канала в другой, программа должна решить, насколько тише он становится в центре. Это называется *законом панорамы* (*pan law*). Если ничего не убавлять, звук в центре станет на 3 дБ громче, чем при крайнем положении, потому что звучит из двух громкоговорителей сразу. Если убавить каждый канал на 3 дБ, сохраняется суммарная мощность двух каналов. В простой модели стереовоспроизведения это соответствует постоянной громкости, но реальное впечатление зависит от громкоговорителей, помещения и места слушателя. Если убавить на 6 дБ, звук не станет громче при сложении каналов в моно (рис. 2.4).

Значения −3 и −4,5 дБ — частые компромиссы, но программы по умолчанию используют разные законы, а встречаются и 0, и −2,5, и −6 дБ. Отсюда типичная неприятность. Эффекты для сцены собрали в одной станции, а сводят в другой; звуки, стоящие в центре, вдруг стали громче или тише на пару децибел, и баланс, выверенный по ушам, разошёлся. Никто ничего не трогал — просто у двух программ разные законы панорамы. То же случается внутри одного проекта, если у части дорожек закон задан отдельно. Поэтому закон выбирают один раз для проекта, проверяют его в настройках (в Reaper — в свойствах проекта и дорожки, в Pro Tools — в настройках сессии, параметр *Pan Depth*) и согласовывают с теми, кому передаётся материал. Для многоканального звука к этому добавляются свои правила, о которых пойдёт речь в главе 13.

Автоматизация

Почти любой параметр — громкость, панораму, посыл, параметр плагина — можно менять во времени. Программа записывает его изменения как линию (*envelope*, *automation lane*) и потом воспроизводит. Записывать её можно мышью, рисуя точки, или движением регулятора во время воспроизведения. Для второго способа у станций есть режимы, названия которых почти совпадают у всех программ:

- *Read* (чтение) — автоматизация только воспроизводится;
- *Touch* (касание) — запись идёт, пока вы держите регулятор, а после отпускания он возвращается к записанной линии;
- *Latch* (защёлка) — запись начинается при касании и продолжается до остановки, сохраняя последнее положение;
- *Write* (запись) — записывается всё подряд, даже если регулятор не трогали.

Большинство неприятных сюрпризов связано с режимами *Write* и *Latch*. Звукорежиссёр аккуратно нарисовал изменения громкости реплик, а после очередного прослушивания они исчезли: дорожка осталась в режиме *Write*, и программа молча перезаписала линию текущим положени-

ем фейдера. Спасают история отмены или сохранённая версия проекта, а на будущее — привычка возвращать дорожки в режим *Read* сразу после записи. Для правки уже готовой автоматизации удобен режим, при котором движение регулятора не заменяет линию, а сдвигает её целиком (в Reaper — *Trim/Read*, в Pro Tools — режим *Trim*).

Вторая ловушка связана с перемонтажом. Изображение перемонтировали, звук сдвинули вслед за ним, а изменения громкости остались на старых местах и теперь приглушают не ту реплику. Причина в том, что линия автоматизации принадлежит дорожке, а не фрагменту, и при сдвиге фрагментов остаётся на месте. В большинстве станций можно включить перемещение автоматизации вместе с фрагментами (в Reaper — *Move envelope points with media items*, в Pro Tools — *Automation Follows Edit*). Но надёжнее задавать громкость отдельных реплик и эффектов на самих фрагментах: она переедет вместе с ними при любом перемонтаже. Автоматизацию дорожки оставляют для того, что относится к сцене в целом.

Плагины

Плагины подключаются к станции через стандартные форматы:

- *VST3* (*Virtual Studio Technology*, разработка фирмы Steinberg) — самый распространённый, работает на Windows и macOS;
- *AU* (*Audio Units*) — формат Apple, только на macOS;
- *AAX* (*Avid Audio eXtension*) — формат программ Avid, прежде всего Pro Tools;
- *CLAP* (*CLever Audio Plug-in*) — более новый открытый формат, его поддерживают Reaper и ряд других программ.

Один и тот же плагин обычно выпускается в нескольких форматах, и звучит он в них одинаково. Формат важен для совместимости. Проект с плагинами AU, открытый на компьютере с Windows, откроется без них, потому что AU там не существует, а сессии Pro Tools нужны версии AAX. Поэтому, передавая проект другому человеку, заранее выясняют, на чём он работает, а то, что должно звучать именно так, сохраняют в виде обработанного файла.

Плагин может работать в реальном времени (обработка пересчитывается при каждом воспроизведении, и её можно менять) или применяться к фрагменту заранее, создавая новый файл. Второе разумно для тяжёлой обработки, например шумоподавления, или когда результат уже утверждён. Есть и промежуточный вариант — *заморозка* (*freeze*): станция один раз пересчитывает дорожку со всеми плагинами в файл и проигрывает его, но может в любой момент «разморозить» её и вернуть плагины.

Отдельный вопрос — нагрузка на процессор. Плагины, стоящие на дорожке последовательно, ждут результата друг друга, и такую цепочку нельзя разделить между ядрами процессора, а независимые дорожки считаются параллельно. Поэтому нагрузка, распределённая по многим дорожкам, даётся легче, чем та же нагрузка в одной длинной цепочке, например на мастершине.

Компьютер и хранение

Для звука компьютеру нужно меньше, чем кажется: сто дорожек 48 кГц, 24 бита — это около 14 МБ в секунду, а любой современный твердотельный диск выдаёт в сотни раз больше. Узким местом оказывается процессор, нагруженный плагинами, и видео, которое занимает гораздо больше ресурсов, чем звук (поэтому для работы удобнее его облегчённая копия, глава 4).

Ценнее всего в компьютере — не программы, а проект и материал. Надёжное правило резервного копирования называют «3–2–1»: три копии данных, на двух разных носителях, одна из них — в другом месте (на внешнем диске дома или в облаке). Кроме копий материала полезно сохранять последовательные версии проекта: Reaper и другие станции умеют автоматически сохранять копии проекта с отметкой времени.

> **Историческая справка**
> В первые годы цифрового монтажа станции часто были готовыми комплексами, где аппаратура и программа шли от одного производителя, а плагины считались не процессором компьютера, а дополнительными платами цифровой обработки сигнала (*digital signal processing*, DSP), поэтому их число на канал было ограничено. Работа не обходилась без внешнего микшерного пульта и коммутационной панели. Мощность обычных компьютеров постепенно сделала эти приборы необязательными, и сегодня небольшой проект целиком помещается в ноутбук.