

NGA – звук следующего поколения

Александр Луганский, по материалам Би-би-си и EBU

На фоне достижений в сфере улучшения качественных характеристик изображения часто на второй план отходят усилия по совершенствованию звукового сопровождения, являющегося неотъемлемой компонентой аудиовизуального контента. Хотя и здесь прогресс более чем очевиден, что подтверждается распространением аудиосистем на базе стандартов Dolby Atmos и DTS:X. И тот, и другой стандарт, равно как и созданные в соответствии с ними системы, призваны обеспечить зрителям возможность персонализации и погружения в экранное действие за счет применения объектно-ориентированного аудио.

Все эти технологии объединяются термином «Звук следующего поколения» (Next Generation Audio – NGA), и активное участие в их эволюции принимает британская Би-би-си, а точнее, департамент исследований и разработок этой корпорации, разумеется, в сотрудничестве со всей индустрией. Цель работ – получение новых средств и методов создания и доставки звука, адаптируемого к системе воспроизведения и предпочтениям аудитории, что позволит сделать воспроизведение звука в месте просмотра контента более персонализированным, доступным и с сильным эффектом присутствия (погружения). Именно такой звук принято теперь называть звуком следующего поколения.

NGA, повторюсь, обеспечивает возможность адаптации воспроизведения звука к условиям, в которых оно выполняется, чтобы обеспечить максимально возможное в этих условиях качество аудио. Важно, что звуковой микс создается единожды, а при воспроизведении производится именно адаптация исходного микса к системе звукоусиления и воспроизведения, имеющегося у потребителя, будь то усилитель с акустическими системами или активные (со встроенным усилителем) акустические системы в конфигурации стерео или 5.1, звуковая панель (sound bar) или просто наушники. Надо подчеркнуть – адаптация исходного микса, а не выбор оптимального варианта из набора предварительно созданных миксов.

Ключевым достоинством NGA как раз и является возможность адаптации исходного микса к условиям, в которых находится каждый слушатель. Например, адаптация может заключаться в подстройке громкости диалога и важных звуковых фрагментов, чтобы сделать просмотр телевизионного контента более комфортным для людей с дефектами слуха или для тех, кто смотрит телевизор там, где шумно, в частности, в общественном транспорте. Также NGA упрощает предоставление слушателю альтернативных вариантов звука, в частно-

сти, выбор языка или стиля комментирования спортивного события. Такой персонализацией пользователь может управлять самостоятельно с помощью интерфейса воспроизведения или делая соответствующие настройки в своем профиле, но можно было бы использовать и сенсоры, определяющие место расположения устройства воспроизведения и уровень шума вокруг него, чтобы автоматически корректировать определенные параметры адаптации.

Интересно рассмотреть подходы, которым в своей работе по NGA следуют разработчики Би-би-си. Этих подходов – три. Первый заключается в сотрудничестве с отраслевыми партнерами, включая компании, специализирующиеся на звуковых технологиях, и вещателей. Цель заключается в формировании общих стандартов, рабочих процессов и методов для создания и доставки NGA.

Второй подход предусматривает тесное взаимодействие специалистов внутри самой Би-би-си, направленное на достижение возможности обеспечить аудитории эффективные сервисы NGA в перспективе. И третий подход состоит в подготовке сообщества практиков для производства NGA, чтобы понять, как получить и предоставить аудитории дополнительные опции и как улучшить имеющиеся средства производства звука.

Одним из важных партнеров Би-би-си является Европейский вещательный союз (European Broadcasting Union – EBU). А ключевой частью проводимой работы стал открытый стандарт ADM (Audio Definition Model), определяющий программный материал NGA и служащий платформой для создания, сохранения и обмена NGA.

ADM состоит из набора элементов, каждый из которых используется для описания аспектов звука. Все элементы имеют формат XML, они содержат различные атрибуты и элементы низшего уровня. Элементы связаны друг с другом ссылками (кроме элемента audioBlockFormat).

Как следует из схемы ADM, она делится на три части – контент, формат и файл BW64. Первые две части – контент и формат – формируют метаданные ADM в формате XML, которые обычно содержатся в соответствующем блоке файла BW64. Нижняя часть на схеме, относящаяся к файлу BW64, содержит так называемый блок распределения каналов (блок chpa), представляющий собой справочную таблицу, связывающую метаданные ADM со звуковыми треками в файле.

Часть, обозначенная как контент, описывает техническое наполнение аудио, например, содержится ли в нем диалог, какой используется язык. Также здесь есть метаданные, определяющие громкость.

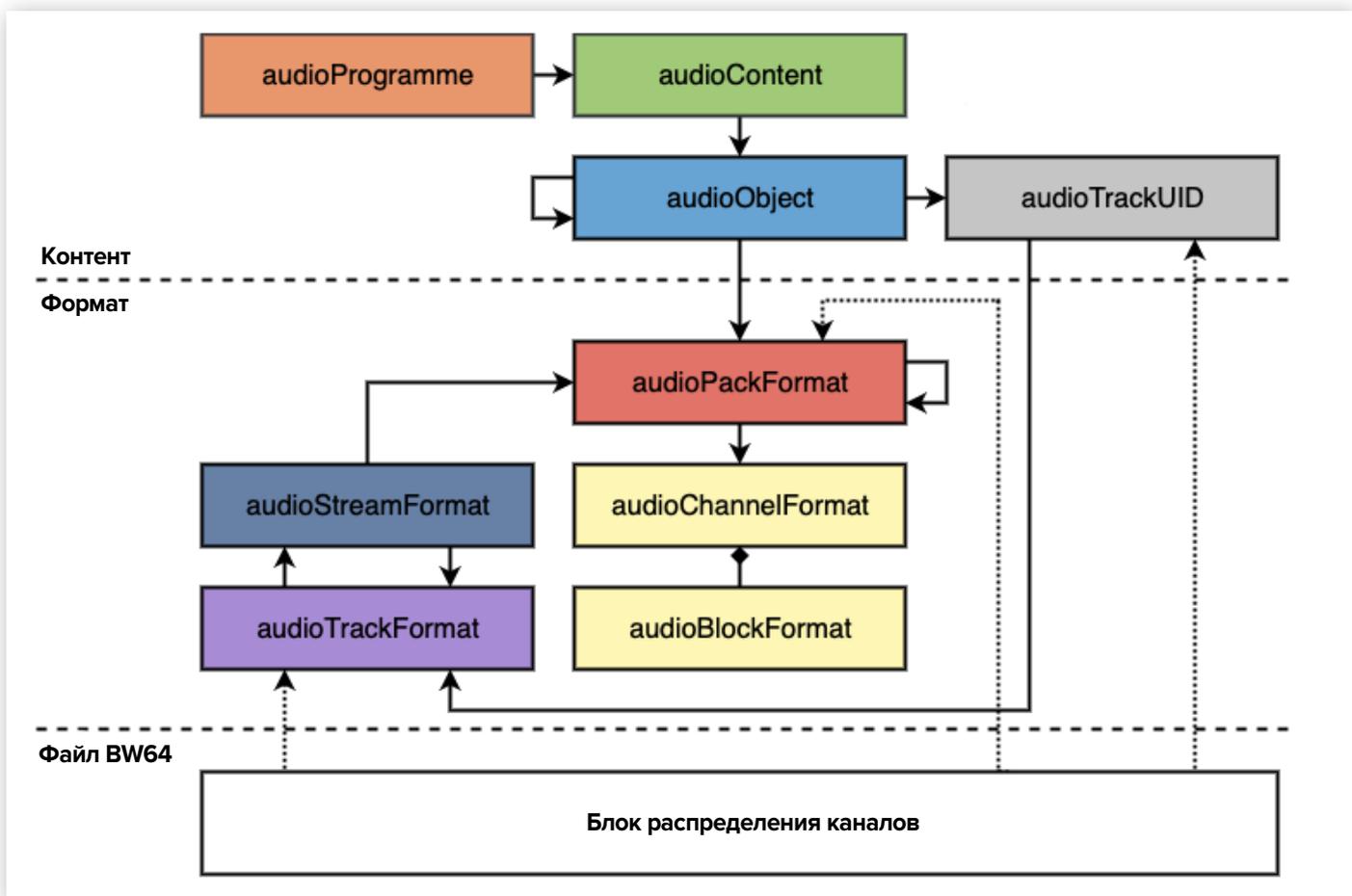


Схема ADM (источник - EBU)

Часть, относящаяся к формату, описывает, каналы какого типа находятся в составе звукового трека и как они сгруппированы друг с другом, например, левый и правый каналы в стереопаре. Элементы в части контента обычно уникальны по отношению к звуку и программе, тогда как элементы в части формата могут использоваться многократно.

Надо отметить, что разработчики Би-би-си не довольствуются только сторонними средствами производства NGA, но и создаются свои собственные. Подробную информацию об этих средствах можно найти на сайте www.bbc.co.uk/rd.

Но создать звук, соответствующий спецификации NGA, это только половина дела. Вторая половина – доставить этот звук до аудитории и обеспечить ей возможность прослушать его так, как задумывалось на стадии создания. Для этого применяются несколько технологий NGA-кодирования, такие как Dolby AC-4, DTS UHD Audio и MPEG-H Audio. С компаниями и организациями, специализирующимися на разработке и развитии этих технологий, равно как и с представителями EBU, Би-би-си находится в постоянном тесном взаимодействии.

Как известно, теория останется только теорией, пока не будет применена и проверена на практике. Аналогично, совершенствование технологий невозможно без практического опыта их использования, для чего необхо-

димы испытания. И они были проведены еще в 2019 году, когда проходил чемпионат Европы по легкой атлетике. Потом, как известно, пришла пандемия коронавируса, что наложило жесткие ограничения и на проведение массовых мероприятий, и на различные эксперименты, с ними связанные.

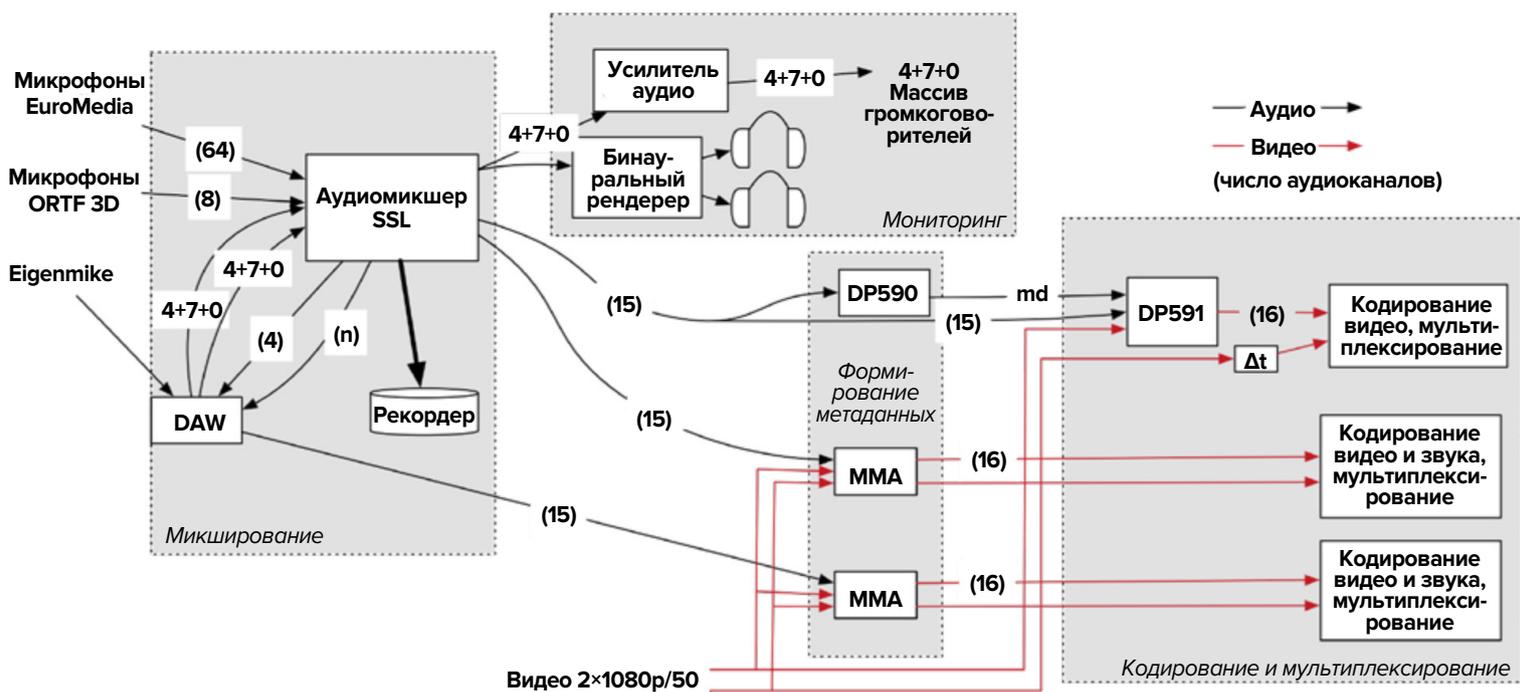
Тем не менее тесты, проведенные в 2019 году Би-би-си в сотрудничестве с партнерами из EBU, дали богатые и полезные результаты. Технология NGA тогда использовалась для формирования 3D-звука в сочетании с комментированием на нескольких языках.

В основе формирования звука NGA лежала пространственная аудиоплатформа, дополняемая четырьмя монофоническими объектами. Платформа обеспечивала формирование аудио в двух форматах: на базе каналов 4+7+0 для последующего кодирования в AC-4 и MPEG-H Audio; на базе сцен (формат Ambisonics высокого порядка) для кодирования в Qualcomm/MPEG-H Audio.

Четыре монофонических объекта – это комментарии и звуковое описание для каждого из двух использовавшихся языков: английского и французского.

Для создания объемного звукового образа применялся массив микрофонов 3D ORTF, расположенный высоко над стадионом – практически под самой его крышей.

А для создания звука на базе сцен прямо на поле, близко к финишной линии, расположили микрофон



Общая схема рабочего процесса NGA, применявшаяся на чемпионате Европы по легкой атлетике 2019 года (источник – отчет Би-би-си WHP 345)

EM32 Eigenmike. Кроме того, было установлено множество направленных микрофонов, распределенных по всей территории стадиона. Они входили в стандартную конфигурацию вещательной системы и тоже использовались для формирования микса. Объединительной транспортной инфраструктурой выступала оптическая сеть MADI.

Аудиомикшер SSL использовался для формирования 15-канального звукового сигнала. В состав этих 15 каналов входили один или два набора из 11 сигналов. Один набор был жестко ограничен форматом 4+7+0 и не содержал канала LFE. Второй был в объемном формате. Еще четыре канала относились к монофоническим объектам. Сюда вошли комментирование на английском с поддержкой интерактивности (регулировка усиления, опционально – панорамы), звуковое описание на английском (интерактивность как опция), комментирование на французском с поддержкой интерактивности (регулировка усиления, опционально – панорамы) и звуковое описание на французском (интерактивность как опция).

Не буду здесь подробно описывать результаты тестирования, приведу лишь общее заключение. Как отмечают те, кто проводил тесты, такая трансляция проводилась впервые с применением данного рабочего процесса и в режиме прямой трансляции. Испытание было успешным, его результаты превзошли ожидания организаторов. Кроме того, были выявлены некоторые важные эксплуатационные аспекты, нуждающиеся в дальнейшем внимании.

Би-би-си не ограничилась испытаниями только на чемпионате Европы по легкой атлетике. Тесты также проводились в 2020 году при содействии съемочной группы сериала «Несчастный случай» и в сотрудниче-

стве с Солфордским Университетом, когда тестировалась персонализация звукового микса. И, наконец, были проведены лабораторные эксперименты, чтобы определить, как пользователи хотят взаимодействовать с такими функциями.

В завершение этого краткого экскурса в область звука следующего поколения имеет смысл сказать, почему разработчики NGA считают свою работу важной. Прежде всего, есть уверенность в том, что NGA позволит повысить качество сервисов, предоставляемых аудитории. Персонализация звукового сопровождения способна существенно улучшить разборчивость и слышимость звуковой программы, а также усилить общее удовольствие, получаемое от прослушивания и просмотра медиаконтента.

При наличии большого числа доступных сегодня способов прослушивания контента автоматическая адаптация микса к системе воспроизведения может способствовать улучшению качества и даже позволит сделать доступным 3D-аудио практически в любых условиях. К тому же технологии кодирования NGA более эффективны, чем те, что у аудиокодеков нынешнего поколения, а потому они могут применяться для повышения качества уже используемого звука форматов стерео и 5.11.

Разработчики NGA уверены, что предыдущие несколько лет, в течение которых разрабатывались соответствующие стандарты и иные нормативные документы, заложили хорошую основу для дальнейшей деятельности в этом направлении, а времена для этого наступают благоприятные, поскольку технология уже созрела настолько, чтобы обеспечить предоставление аудитории соответствующих функций и сервисов.