
EVALUACIÓN DE DESEMPEÑO

aplicado a

Audio Units de Core Audio

Versión 1.1

Autores:

Eduardo Hopperdietzel

Camilo Muñoz

Docente:

Eliana Scheihing

Asignatura:

Evaluación de Desempeño (INFO273)

Universidad Austral de Chile

2 de julio de 2021

Introducción	3
Metodología	3
Justificación	3
Diseño del Estudio	3
Primeros Resultados	4

Introducción

Core Audio es una interfaz de programación de aplicaciones para el tratamiento de audio del sistema operativo Mac OS X.

Como fue mencionado previamente en el primer informe, en este trabajo nos enfocaremos principalmente en [AUAudioFilePlayer](#), un plugin de tipo generador que permite leer archivos de audio y generar streams en tiempo real y también haremos uso del plugin [MultiChannelMixer](#) para sumar los streams y finalmente reproducirlos en el dispositivo de salida.

Para esta segunda entrega expondremos detalles sobre la metodología de estudio escogida justificando el por qué de su selección así como también se darán los primeros detalles de los resultados obtenidos en primera instancia realizando una prueba temprana.

Metodología

Para el caso propuesto se hará uso de benchmark, el objetivo de esto es ser capaz de medir qué tan bien es capaz de funcionar un computador al trabajar en el procesamiento de pistas de audio, ver cuánta es la cantidad de trabajo que puede soportar de tal modo de determinar la cantidad de pistas que producen sobrecarga u overlay durante el procesamiento de este tipo de datos.

Justificación

Se hace uso de benchmark porque se necesita enviar una carga al computador para saber la capacidad máxima a la cual es capaz de trabajar en el procesamiento de pistas de audio, este tipo de metodología se adapta perfectamente a lo que se desea hacer. El motivo de hacer este estudio de esta manera es que en el mundo real, cuando se trabaja con canciones, por poner un ejemplo, las pistas de audio que se utilizan son muchas y van por cada componente que es parte de la canción por lo que se tiene una pista para la voz, otra pista distinta para la guitarra, otra más para la batería, etc. Esto puede causar una sobrecarga en el sistema u overlay lo cual impide trabajar de forma eficiente con este tipo de datos, por lo que el diseño e implementación de este estudio puede ser bastante útil para ingenieros de sonido, productores musicales, etc.

Diseño del Estudio

Como mencionamos previamente, la idea del estudio es desarrollar un benchmark y utilizarlo para medir el rendimiento de la máquina (número de pistas de audio simultáneas máxima que pueden ser combinadas y reproducidas en tiempo real), para luego analizar los resultados y obtener conclusiones.

La forma en que medimos la cantidad de pistas máximas es la siguiente:

Las interfaces de audio pueden ser configuradas para funcionar a una frecuencia de muestreo fs específica (Ej: 44.1 KHz). El convertidor digital análogo (DAC) de estas, cuenta con un buffer de tamaño N configurable, el cual debe ser rellenado con muestras dentro de un lapso de tiempo limitado. Si no se alcanza a llenar el buffer en ese tiempo, se generan discontinuidades y ruidos en el audio que desafortunadamente son muy perceptibles e impiden un flujo de trabajo adecuado.

Una forma de calcular ese límite de tiempo es dividiendo el tamaño del buffer N por la frecuencia de muestreo fs.

Cuando la tarjeta de sonido solicita llenar su buffer, se gatilla una función callback, la cual contiene como parámetro un puntero al buffer. Dentro de esta función se debe ejecutar todo el procesamiento de audio digital. Por lo tanto, medimos el tiempo antes y después de realizar el procesamiento, y si esta diferencia de tiempo es mayor al límite entonces el sistema se considera saturado. Por lo tanto, nuestro benchmark consiste en iterar hasta encontrar el número máximo de pistas que no sature el sistema.

Primeros Resultados

Actualmente estamos mejorando una segunda versión del benchmark, ya que en la primera obtuvimos resultados poco realistas.

La máquina con la que realizamos las pruebas, permite combinar en promedio 25000 pistas audio mono de 32 bits y 44,1 KHz, cuando en la práctica, utilizando programas de edición de audio, con alrededor de 40 pistas de audio el sistema se satura.

La razón de esta gran diferencia es que no tomamos en cuenta la interfaz gráfica ni los plugins de efectos que se suelen utilizar, como ecualizadores, compresores (de dinámicas), reverbs, delays, etc (como se muestra en la imagen).

Los cuales en algunos casos realizan operaciones intensas como cálculos de transformadas discretas de fourier, convolución, entre otras.



Por lo tanto, decidimos que en esta segunda versión, en vez de únicamente leer pistas de audio y mezclarlas, enlazamos cada una con su propio ecualizador paramétrico de 8 bandas (IIR), un compresor de dinámicas, un efecto chorus, flanger o phaser (aleatorios) y añadiremos 5 buses paralelos que serán compartidos por todas las pistas (3 tipos de reverb distintos y dos tipos de delays), para finalmente aplicar un plugin limitador a la mezcla stereo resultante y enviar el stream al dispositivo de salida.

Probamos esta nueva versión del benchmark en la máquina mencionada en el primer informe, con su tarjeta de sonido integrada y obtuvimos los siguientes resultados:

Max Tracks	Buffer Size	Sample Rate	Bit Depth
58	1024	96000	32
55	512	96000	32
52	256	96000	32
45	128	96000	32
36	64	96000	32
59	1024	96000	24
56	512	96000	24
53	256	96000	24
47	128	96000	24
36	64	96000	24
59	1024	96000	16
56	512	96000	16
54	256	96000	16
48	128	96000	16
36	64	96000	16
61	1024	48000	32
59	512	48000	32
56	256	48000	32
52	128	48000	32
46	64	48000	32
61	1024	48000	24
59	512	48000	24
56	256	48000	24
52	128	48000	24
46	64	48000	24
63	1024	48000	16
59	512	48000	16
56	256	48000	16
54	128	48000	16
48	64	48000	16
61	1024	44100	32
59	512	44100	32
53	256	44100	32

52	128	44100	32
46	64	44100	32
61	1024	44100	24
59	512	44100	24
57	256	44100	24
53	128	44100	24
46	64	44100	24
62	1024	44100	16
59	512	44100	16
56	256	44100	16
53	128	44100	16
45	64	44100	16

Donde:

- **Max Tracks:** Equivale al número máximo de pistas de audio mono mezcladas en tiempo real, con la configuración dada por el resto de las variables.
- **Buffer Size:** Es el número de muestras de audio que solicita la interfaz en cada callback.
- **Sample Rate:** Es la frecuencia de muestreo de los archivos de audio y de la configuración de la interfaz.
- **Bit Depth:** Es la cantidad de bits con la que se representa cada muestra de audio, tanto de los archivos como en la configuración de la interfaz.

Como se puede observar, la cantidad de pistas máximas depende fuertemente del tamaño del buffer, donde se sabe que mientras menor sea su tamaño, menor es la latencia (pero aumenta el uso de CPU). Por suerte la latencia en programas de edición de audio no suele ser tan importante como sería en shows en vivo, por lo tanto se puede optar por tamaños de buffer mayores.

La segunda variable que afecta considerablemente el rendimiento es la frecuencia de muestreo, por ejemplo con una frecuencia de 98 KHz, tamaño de buffer 64 y bit depth 16 se alcanzan a mezclar 36 pistas, en cambio con una fs de 44.1 KHz y los mismos parámetros, se logran mezclar 45.

Esto pareciera no ser muy importante ya que con una frecuencia de muestreo de 44.1 KHz se logra representar perfectamente todo el rango auditivo del humano (de 20 Hz hasta 20 KHz aprox). Sin embargo, existen muchos efectos de audio como la distorsión o saturación, los cuales generan armónicos superiores a este rango, donde al no poder ser representados, reflejan estas frecuencias en el espectro, causando artefactos indeseados, fenómeno que se conoce como aliasing.

Por último, la variable bit depth parece ser la menos influyente ya que genera diferencias de 1 a 2 pistas, lo cual creemos se debe al tipo de procesador, por lo tanto para el siguiente informe, nos gustaría realizar el mismo benchmark en otras máquinas con procesadores diferentes y comparar resultados.