

Hogyan működik: Linux-hangrendszer leírása

Posztolva 2010. április 8-án, 10:30-kor

Forrás: <http://tuxradar.com/content/how-it-works-linux-audio-explained>

Gond van a Linux hangrendszerével és nem az, hogy nem mindig működik. A helyzet az, hogy túlbonyolított. Azonnal tiszta lesz, ha leülsz egy darab papírral és megpróbálsz felvázolni a zenei fájljaidat a hangszóróig elszállító technológiák kapcsolatait: a vázlat hamar összegubancolódott spagettihez hasonlóvá válik. Ez baj, mert nincs semmi olyan lényeges dolog, amitől a hangszolgáltatás sokkal bonyolultabb lenne, mint más technológia. Egy ponton belép a linuxos dobozodba és egy másik ponton elhagyja.

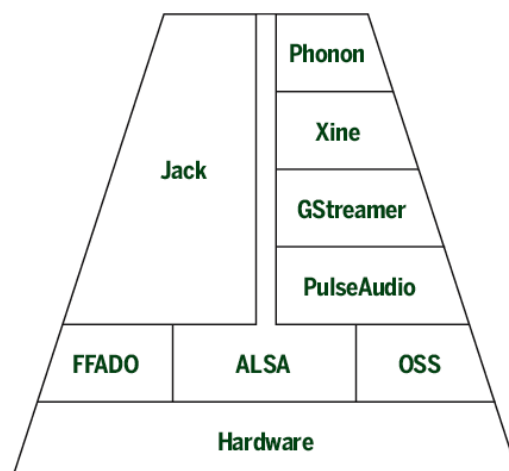
Ha már elegend van a zűrzavarból és szeretnéd megérteni, hogyan áll össze a kép, itt vagyunk, segítünk – olvass tovább és megtudod, hogyan működik a Linux audio.

Ha megrajzolnánk a számítógépedet a hálózat többi gépével összekötő hálózati keretrendszert leíró OSI-modelt, egyértelmű elrendezést találnánk, minden eleme saját folyamatokkal és funkciókkal rendelkezvén. A rétegek között nagyon kicsi az átfedés és biztosan nem találsz olyan végfelhasználói folyamatot a hetedik rétegen, ami az első réteg soros adatfolyamának impulzusaiba kavarna bele.

Éppen ez az, ami viszont a Linux hang keretrendszerénél előfordulhat. Még világosan meghatározott alsó szint sincs, számos hang-technológia kavar önállóan a kernellel és a hardverrel. A Linux hangrendszere nem a hálózati modellhez, hanem inkább a földkéreg rétegeihez hasonlít, ahol az alsóbb rétegek esetenként a felszínre törnek zavart, bajt okozva és a felső rétegeket az eredetileg elrejtett alap technológiák helyére tolva.

Például az Open Sound Protocol (Nyílt Hangprotokoll) régebben a kernel szintjén ült, közvetlenül kommunikálva a hardverrel, de mostanra már csak egy kompatibilitási réteg van, ami az ALSA-n ül. Az ALSA rendelkezik kernel szintű veremmel és magasabb szintű API-val a programozók számára, meghajtók és hardverek jellemzőit összekeverve térhatású hang lejátszásának lehetőségével, vagy MP3 codec-kel. Mindeközben a legtöbb disztribúció a legfelső szinten a PulseAudio-hoz és Gstreamer-hez ragaszkodik, és ezzel a Szent András hasadék romboló hatásához mérhetően potenciális instabilitást kockáztatnak.

Íme a Linux által tipikusan használt hangrétegek szokásos elrendezése egyszerűsítve. A réteg minél lejjebb található, annál közelebb van a hardverhez.



ALSA

BEMENETEK: PulseAudio, Jack, GStreamer, Xine, SDL, ESD

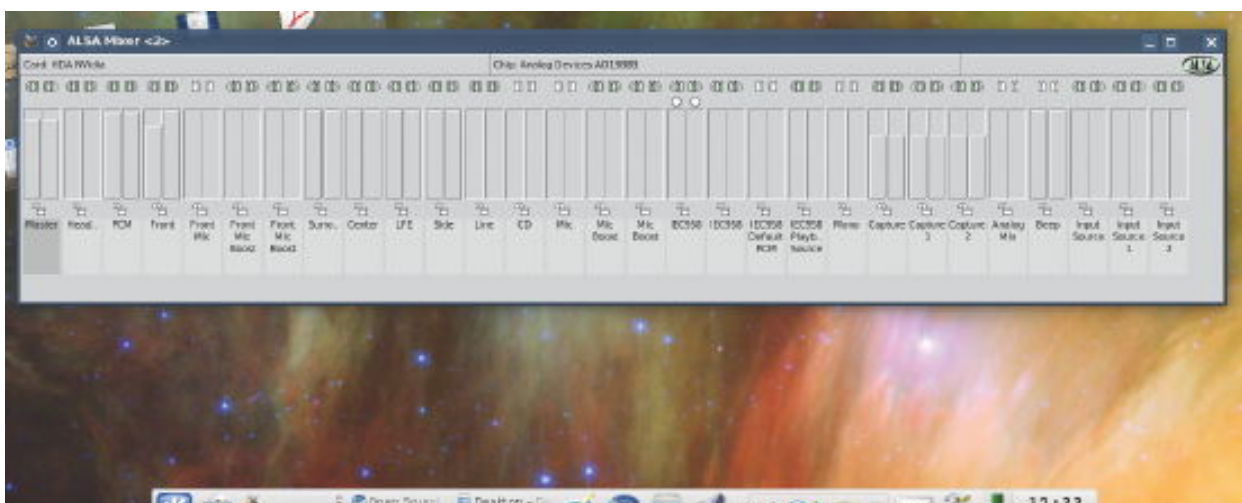
KIMENETEK: harver, OSS

Ahogy Maria von Trapp¹ mondta: „Kezdjük a legelején.” Ami a Linux audio-t illeti, a kezdet az Advanced Linux Sound Architecture, azaz ALSA. Ez kapcsolódik a kernelhez és a rendszer többi része számára biztosítja a hangképességet. Ugyanakkor sokkal többre tör, mint a szokásos kernel-meghajtók; képes keverni, biztosítja a kompatibilitást a többi réteg felé, a programozóknak API-t biztosít és olyan alacsony szinten és stabilan működik, hogy vetekszik a Windows-on és az OS X-en található ASIO- és a CoreAudio-félékkel.

Az ALSA-t az OSS leváltására tervezték. Ám az OSS valójában nem halt meg, köszönhetően az ALSA kompatibilitási rétegének, amit a régebbi, OSS-es alkalmazások futtatásának biztosítására terveztek. Legegyszerűbb úgy tekinteni az ALSA-ra, mint a Linux hangrendszer meghajtó rétegére. A hang hardverének szüksége van egy megfelelő, snd_ előtaggal jelölt kernelmodulra, amit be kell tölteni, hogy bármi történjen. Hogy a rendszered halljad, szükség van egy ALSA kernel driver-re, ezért voltak a laptopok némák, amíg valaki úgy nem gondolta, hogy csinál meghajtót hozzá. Szerencsére a legtöbb disztró automatikusan beállítja az eszközöket és a modulokat.

Az ALSA felelős a hangeszközök képességeinek szoftver API-vá való lefordításáért, amit azután a rendszer többi része használ a hangok kezelésére. Úgy tervezték, hogy kiküszöbölje az OSS (és a többi akkor létezett hangmeghajtó) hiányosságait, amik közül a legjelentősebb az volt, hogy egyszerre csak egy alkalmazás érhet el a hardvert. Ezért van szükség az ALSA-ba beépített szoftver elemekre a hanggal kapcsolatos kérések kezeléséhez és a hardver képességeinek azonosításához.

Ha például miközben AmaroK-kal zenét hallgatva szeretnél játszani, ALSA kell, hogy a két hangfolyamot kezelj és szoftverben összekeverj, vagy a hangkártyádon hardveres keverőt kell használni ugyanezen hatás eléréséhez. Az ALSA képes egyszerre nyolc hangeszközt kezelni és a hardver MIDI-funkcióit alkalmanként elérni, noha (ez a konkrét hardver hangmeghajtójának képességeitől is függ és) egyre kevésbé fontos a számítógépek teljesítményének növekedésével.



Az ALSA-mixer képernyőképe mutatja mindazt, ami rossz a Linux hangrendszerében – nem kellene ilyen bonyolultnak lennie.

Amitől az ALSA más mint a tipikus kernel modulok, vagy eszközmeghajtók az, hogy a felhasználó részben beállíthatja. Na és itt kezd a Linux-hang összetetté válni, mivel az ALSA beállításainál majdnem 1. A Muzsika hangja írója

Hogyan működik: Linux-hangrendszer leírása

minden megváltoztatható saját konfigurációs fájl készítésével – onnantól, hogyan keverje össze a hangfolyamokat és melyik kimeneten lépjenek ki a rendszerből, egészen a mintavételi arányig, a bit-mélységig és a valósídejű effektusokig.

Az ALSA viszonylagos átláthatósága, hatékonysága és rugalmassága segítette elő, hogy a Linux-hang szabványává és azzá a réteggé váljon, amin majdnem az összes többi hang keretrendszernek át kell haladnia, hogy a hangelőállító hardverrel kommunikáljon.

PulseAudio

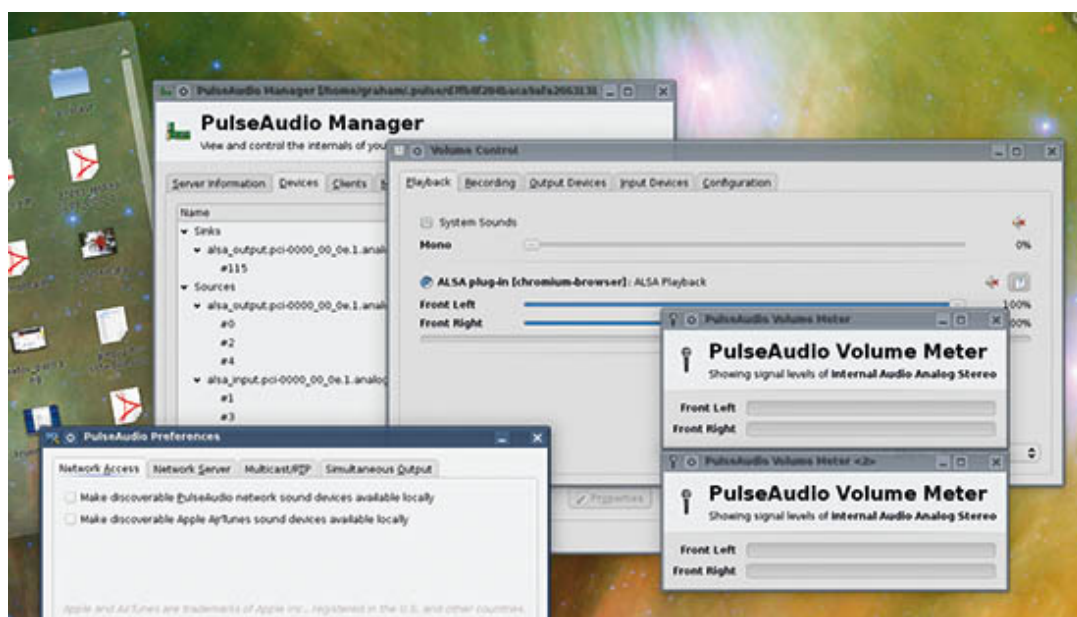
BEMENTEK: GStreamer, Xine, ALSA

KIMENETEK: ALSA, Jack, ESD, OSS

Ha azt gondolnád, hogy stabilan az ALSA-ra támaszkodva a dolgok egyszerűbbé válnak, sajnos tévedsz. Az ALSA majdnem mindent kezel ahhoz, hogy a hang be- és kijusson a gépedből, de egy újabb összetett rétegen kell áthámozni magad. Ez a PulseAudio birodalma – egy kísérlet a hardver és a szoftver képességei közötti, a helyi és a távoli gépek közötti, valamint a hangfolyamok tartalma közötti szakadék áthidalására. A hálózati hangokkal azt teszi, amit az ALSA a többes hangkártyával és rugalmassága miatt sok Linux-disztró számára kezd valamiféle szabvánnyá válni.

Ahogy az ALSA-nál láttuk, ettől a rugalmasságtól lesz ez is komplex, de a probléma sokkal fokozottabban jelentkezik, mivel sokkal inkább felhasználói irányultságú. Ezért az átlag felhasználó nagyobb eséllyel gabalyodik bele a hálójába. A legtöbb disztró a beállításokat kéznél tartja; pl. a legfrissebb Ubuntu-ban talán még észre sem veszed, hogy a PulseAudio telepítve van. Ha a hangerő beállítandó, a keverő applet-re kattintasz, amit látsz az az ALSA-panelnek tűnik, de valójában azt látod, ahogy az ALSA a PulseAudio-hoz fordul, majd vissza az ALSA-hoz – egy virtuális eszközt látsz.

Elsőre nem látszik, hogy bármi újat adna a Linux-hanghoz, ezért viszolyognak tőle annyian. Nem egyszerűsít, vagy tesz hatékonyabbá, de számos fontos tulajdonságot ad a hangrendszerhez. Emellett, a hangalkalmazásoknak szintén mindenhol jelen lévő réteg, függetlenül azok képességeitől, vagy a hardveredtől.



A PulseAudio nagy teljesítményű, de gyakran kárhozott, amiért a Linux-hangot sokkal komplikáltabbá teszi.

Hogyan működik: Linux-hangrendszer leírása

Ha minden alkalmazás PulseAudio-t használna, a dolgok egyszerűbbek lennének. A fejlesztőknek nem kellene törődniük a többi rendszer komplexitásával, mivel a PulseAudio platformokon átnyúló kompatibilitást hoz. De ez az egyik ok, amiért annyi más hangmegoldás is van. Az ALSA-val ellentétben a PulseAudio képes több platformon is futni, közöttük más POSIX platformokon és a Microsoft Windows-on. Emiatt, ha egy PulseAudio-t és nem ALSA-t használó alkalmazást készítesz, az alkalmazás más platformokra portolása sokkal könnyebb.

Ugyanakkor, a Linux-ok esetén a PulseAudio és az ALSA szorosan kapcsolódik egymáshoz, mivel az előbbinek szüksége van az utóbbira. A PulseAudio, mint egy hardver, virtuális eszközként állítja be magát az ALSA-hoz kapcsolódva. Ettől a PulseAudio olyan mint a Jack, mert az ALSA és a asztal között ülve adatokat csatornáz oda-vissza transzparens módon. Rendelkezik saját terminológiával is. Pl. a „Sink” (lefolyó) a végső cél, ami lehet másik gép a hálózaton, vagy hangkimenet a hangkártyádon, a virtuális ALSA-nak hála. A sink-eket tápláló PulseAudio-elemeket forrásoknak (sources) hívják – ez lehet a rendszer szokások hangelőállító alkalmazása, a hangkártya bemenete, vagy hálózati hangfolyam, amit egy másik PulseAudio-gép küld.

A Jack-kel ellentétben az alkalmazások közvetlenül nem adnak hozzá, vagy távolítanak el forrásokat, és sokkal finomabban lehet kezelni a hangfolyamokat. Pl. PulseAudio-mixerrel minden, a PulseAudio-n átmenő alkalmazás viszonylagos szintjét be lehet állítani, függetlenül attól, hogy rendelkezik-e, vagy sem saját tolokával. Kiváló eszköz zajos honlapok kordában tartására.

GStreamer

BEMENETEK: Phonon

KIMENETEK: ALSA, PulseAudio, Jack, ESD

A GStreamer-től a Linux-hang még zavarosabbnak néz ki. Azért, mert a GStreamer, akár a PulseAudio, mintha nem adna semmi pluszt a katyvaszhoz. Egy újabb hang-keretrendszer, ami a PulseAudio előtti években sok fejlesztőt tett a követőjévé, különösen a Gnome asztali környezetben. Egyike a kevés eljárásnak, amivel jogvédett codec-eket lehet könnyen Linux-asztal alatt használni. A GTK-fejlesztői is ezt választották a hangrendszer környezetként és még a Palm Pre-en a hangot kezelő verzióját is találhatsz.

A GStreamer a PulseAudio fölötti szintre illeszkedik (amit sok disztribúció hangkimeneteként használ), de az alkalmazások alatti szinten. A GStreamer különleges, mert nem csupán a hanghoz tervezték – kiegészítők segítségével, számos adatfolyam-médiумok számos formátumát támogatja, köztük a videót is.

Az MP3-lejátszás például szokásosan egy kiegészítő codec segítségével történik, ami GStreamer-plugin-ként adódik hozzá a rendszerhez. A Fluendo, hivatalosan Linux alatt bejegyzett kereskedelmi MP3-codec, akár a többi jogvédett codec, mint pl. az MPEG-2, H.264 és MPEG, GStreamer-kiegészítőként érkezik.

Jack

BEMENETEK: PulseAudio, GStreamer, ALSA,

KIMENETEK: OSS, FFADO, ALSA

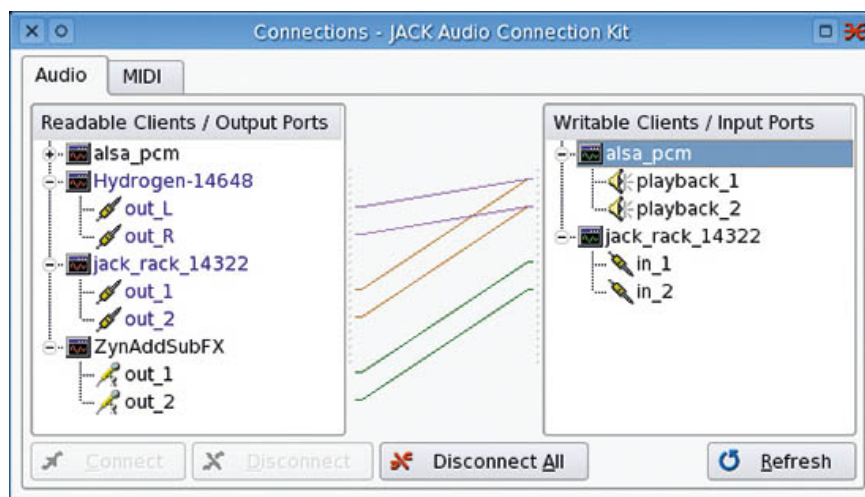
A PulseAudio-hoz hasonló nyílt eszközök előnyei ellenére, mindegyik abból a feltételezésből kiindulva továbbítja a hangot, hogy az azonnal a kimenetre kerül. A Jack a közbenső réteg – a programozásban ismert

Hogyan működik: Linux-hangrendszer leírása

távoli eljárás hívás audio-megfelelője, lehetővé téve, hogy a hangalkalmazásokat különféle elemekből rakják össze.

A legjobb példa a virtuális hangstúdió, ahol egy alkalmazás felel a hangok kinyeréséért és egy másik a feldolgozásáért effektusokkal, mielőtt a végeredményt jelentő hangfolyam a kibocsátáshoz véglegesítő feldolgozásra kerülne. Egy igazi hangstúdióban ezen kapcsolatok a létrehozására jack-eknek is hívott kábelhálózatokat használnának. A Jack ugyanezt végzi szoftveresen.

A Jack a „Jack Audio Connection Kit” rövidítése. Kis késleltetésűnek készítették, ami annyit tesz, hogy



A Jack-kel alkalmazások hangkimenetét más bemenetekre kötheted kézzel – akár egy igazi hangstúdióban.

nem hajt végre olyan változtatást a hangon, ami gátolná a továbbhaladását. De ahhoz, hogy a Jack hasznos lehessen, kell egy olyan alkalmazás, ami a Jack kapcsolatait lekezeli. Ezért nem tekinthető az ALSA-hoz, vagy a PulseAudio-hoz hasonló egyszerű helyettesítőjének, és egy olyan rendszer tetején kell futnia, ami előállítja a hangot és a fizikai bemeneteket szolgáltatja.

A legtöbb Jack-kompatibilis alkalmazással a hangot és a bemeneteket szabadon irányíthatod bármerre, kedved szerint. Például a VLC kimenetén a hangfolyamot lejátszás közben átcsövezheted közvetlenül az Audacity-hez felvételre.

Vagy elküldheted a JackRack-re, amellyel alkalmazással valós idejű effektusok tárházát alkalmazhatod a késleltetésről, a visszhangosításon át, a sejtelmes hangkódolásig.

Az audio-munkaállomások számára ez a sokoldalúság szuper. Pl. az Ardour Jack-et használ a belső és külső kapcsolatokhoz és a Jamin véglegesítő feldolgozó csak Jack folyamatok részeként használható. Olyan mintha teljes ellenőrzésed lenne a stúdiód kábelezése fölött. Linux alatti használata annyira sikeres volt, hogy készítették Jack-et az OS X alá is hasonló célból.

FFADO

BEMENETEK: Jack

KIMENETEK: hang-hardver

A fél- és teljesen professzionális hang világában számos hangeszközt kötnék a befogadó berendezéshez FireWire porton keresztül. Ez a módszer számos előnnyel bír. A FireWire gyors és az eszközök központi

Hogyan működik: Linux-hangrendszer leírása

táplálásúak lehetnek. A szabványos FireWire port stabil és elég kiforrott, sok laptop és asztali gép rendelkezik ilyennel. Külső felvételekhez vihetsz magaddal FireWire-eszközt lappal, amiket azután a stúdióba visszatérve bedughatsz az asztali gépbe.

Azonban, az USB-vel ellentétben, ahol szabvány a hangok külön meghajtó nélküli kezelése, a FireWire-eszközök saját meghajtót igényelnek. A FireWire-protokoll összetettsége miatt elég nehéz ALSA-interfészt készíteni, ezért saját rétegre van szükségük, Eredetileg ez a FreeBOB projekt feladata volt. Azt használta ki, hogy a legtöbb FireWire-hangeszköz alapja ugyanaz a hardver volt.

A FreeBOB jogutódja az FFADO, ami a meghajtói környezetet megnyitotta számos más FireWire-eszköz felé. A Version 2-t 2009. végén adták ki és számos eszközt támogat, mint az Alesis, az Apogee, az ART, a CME, az Echo, az Ediol, a Focusrite, az M-Audio, a Mackie, a Phonic és a Terratec. Hogy melyik eszköz működik és melyik nem, megjósolhatatlan, ezért vásárlás előtt le kell ellenőrizni. Ugyanakkor sok eszközgyártó segítette a meghajtók fejlesztését, a fejlesztőknek eszközöket adva használatra és tesztelésre.

Az FFADO másik jó tulajdonsága, hogy a hardver néhány DSP-keverési tulajdonságát integrálták a meghajtóba, grafikus keverővel kiegészítve a be- és kimenetek szabályozásához. Ez más, mint az ALSA keverője, mert ezzel a hangfolyamok késleltetés nélkül vezérelhetők a hardveren, és pontosan erre van szükség élő előadások felvétele során.

Más hangrétegektől eltérően az FFADO csak a Jack és a hangeszköz között keveri a hangot. Nincs hátsó kijárata PulseAudio-, vagy GStreamer-felé, hacsak nem csatlakoztatod azokat a Jack-hez. Így az FFADO nem használható általános hangréteggént zene, vagy film lejátszásához, hacsak nem vagy kész telepítéssel és Jack-kel foglalkozni. Azt is jelenti, hogy a meghajtó nincs túlterhelve különféle protokollok támogatásával, már csak azért sem, mert a komolyabb hangeszközök alából rendelkeznek Jack-támogatással. Ettől lesz a stúdiók számára az egyik legjobb választás.

Xine

BEMENTEK: Phonon

KIMENETEK: PulseAudio, ALSA, ESD

Kezdünk belemenni a Linux-hang geológiájának magjába. A Xine egy őskövétek. Maradvány, miután sok hangréteget elmosott az idő. A legtöbben a nevet egy nagy tudású DVD film- és médialejátszóból ismeri, amit kora ellenére a legtöbb disztribúcióban továbbra is csomagol, ez a Xine fennmaradásának oka.

A Xine-t, készítésekor a fejlesztők felbontották a média lejátszásához szükséges végfelhasználói könyvtárra és a felhasználói műveletekhez egy megjelenített alkalmazásra. A könyvtár az, ami megmaradt, köszönhetően annak, hogy a csomag sok olyat képes lejátszani, mint az AVI, a Matroska és az Ogg, illetve az általuk tartalmazott tucatnyi formátumot, mint az ACC, Flac, MP3, Vorbis és WMA. Ezt úgy éri el, hogy sok más könyvtár képességeit gyűjti össze. Ennek eredményeként a Xine úgy működik, mint egy mindent egyben keret a fejlesztők számára, akik a legszélesebb kompatibilitásra törekednek anélkül, hogy a szerzői jogok és licencek jogállása miatt kellene aggódniuk.

A Xine képes ALSA-val és PulseAudio-val kommunikálni, továbbá számos alkalmazás képes közvetlenül kapcsolódni hozzá. A legnépszerűbbek a Gxine front-end és a Totem, és a Xine a KDE-s Phonon végén is ott csücsül, tehát az Amaroktól kezdve a Kaffeine-ig mindenben benne van.

BEMENETEK: Qt and KDE-alkalmazások

KIMENETEK: GStreamer, Xine

A Phonon-t a fejlesztők és a felhasználók életének megkönnyítésére tervezték, a rendszer összetettségének csökkentésével. A KDE-4-es alkalmazások számára egy újabb hangkezelő réteggént kezdte életét, de a Qt fejlesztőinek annyira megtetszett az ötlet, hogy magukévá tették, közvetlenül behívva a KDE alapjául szolgáló Qt-keretrendszerbe.

A keresztplatformos alkalmazásoknál ez a fejlesztők számára nagyon előnyös. Lehetővé tette Qt-val zenejátszó készítését Linux alatt, és annak egyszerű lefordítását OS X, illetve Windows alá anélkül, hogy a lejátszás minősége, a hangeszköz képességeinek kihasználása, vagy a cél operációs rendszer hangkezelése miatt aggódni kellett volna. Ezt mind automatikusan elrendezte a Qt és a Phonon, átadva a hangot például OS X alatt a Core Audio API-nak, vagy Windows esetén a DirectSound-nak. A Linux platformon (a Phonon eredeti KDE-verziójától eltérően) a Qt Phonon-ja, alapvetően a transzparens codec-támogatása miatt, a hangot a Gstreamer-nek adja át.

A Phonon-támogatás csendben kimaradt a Qt keretrendszerből. Sokat kritizálták a rendszert, általában azért, mert túlzottan egyszerűsítő jellegű és nem ad semmi újat, ám a KDE valószínűleg ragaszkodni fog hozzá a KDE4 élettartamáig.

A maradék

Sok további hangtechnológia van még, közte az ESD, az SDI és a PortAudio. Az ESD (Enlightment Sound Daemon) sokáig a Gnome alap hangkiszolgálója volt. Időközben a Gnome-ot átalakították, hogy a libcanberra-t használja (ami képes ALSA-val, Gstreamer-rel, OSS-szel és PulseAudio-val kommunikálni) és 2009 áprilisában az ESD-t ejtették a szükségletek közül. Akkor van még az Arts, a KDE ESD-megfelelője, ami nem kapott olyan széleskörű támogatást és több gondot okozott, mint amennyit megoldott. A többség már KDE 4-re váltott, így nincs napirenden többé.

Ellenben az SDL továbbra is ott virít az SDL-könyvtárban, mint hang kimeneti komponens, amit többszáz keresztplatformos játék készítésre használtak. Sok funkciót támogat, emellett kiforrott és stabil.

A PortAudio egy további keresztplatformos hangkönyvtár, ami elérhetővé teszi az SGI-t, a Unix-ot és a Beos-t. A legismertebb, PortAudio-t használó alkalmazás az Audacity hangszerkesztő, így érthető, hogy esetenként miért kiszámíthatatlan a Jack-támogatásának hangkimenete és minősége.

Akkor van még az OSS, az Open Sound System. Az sem ingatta meg, hogy a 2.4-es kernel óta nem alap hangtechnológia a Linux-ban, ami részben annak köszönhető, hogy sok régebbi alkalmazás függ tőle és, ellentétben az ALSA-val, nem csak Linux-rendszereken működik. Van még FreeBSD-s verziója is. 1992-ben jó rendszer volt, de majdnem mindig az ALSA-t ajánlják a leváltására.

Az OSS meghatározta, hogyan legyen Linux alatt a hang és különösen az eszközök elérése az ioctl-fán keresztül, pl. /dev/dsp-ként. Az ALSA-nak van OSS-kompatibilitási rétege a régebbi, OSS-hez ragaszkodó alkalmazások számára, miközben megmarad az aktuális ALSA-szabványon belül.

Az OSS-projekt kísérletezett nyílt és jogvédett fejlesztéssel is és a 4 Front Technologies még mindig aktívan fejleszti üzleti vállalkozásként. Az OSS 4.2-es Build 2002-jét 2009 novemberében adták ki.