A Cost Analysis of Windows Vista Content Protection

Executive Summary

Windows Vista includes an extensive reworking of core OS elements in order to provide content protection for so-called "premium content", typically HD data from Blu-Ray and HD-DVD sources. Providing this protection incurs considerable costs in terms of system performance, system stability, technical support overhead, and hardware and software cost. These issues affect not only users of Vista but the entire PC industry, since the effects of the protection measures extend to cover all hardware and software that will ever come into contact with Vista, even if it's not used directly with Vista (for example hardware in a Macintosh computer or on a Linux server). This document analyses the cost involved in Vista's content protection, and the collateral damage that this incurs throughout the computer industry.

Executive Executive Summary

The Vista Content Protection specification could very well constitute the longest suicide note in history [Note A].

Introduction

This document looks purely at the cost of the technical portions of Vista's content protection [Note B]. The political issues (under the heading of DRM) have been examined in exhaustive detail elsewhere and won't be commented on further, unless it's relevant to the cost analysis. However, one important point that must be kept in mind when reading this document is that in order to work, Vista's content protection must be able to violate the laws of physics, something that's unlikely to happen no matter how much the content industry wishes that it were possible [Note C]. This conundrum is displayed over and over again in the Windows content-protection requirements, with manufacturers being given no hard-and-fast guidelines but instead being instructed that they need to display as much dedication as possible to the party line. The documentation is peppered with sentences like:

> "It is recommended that a graphics manufacturer go beyond the strict letter of the specification and provide additional content-protection features, because this demonstrates their strong intent to protect premium content".

This is an exceedingly strange way to write technical specifications, but is dictated by the fact that what the spec is trying to achieve is fundamentally impossible. Readers should keep this requirement to display appropriate levels of dedication in mind when reading the following analysis [Note D].

Disabling of Functionality

Vista's content protection mechanism only allows protected content to be sent over interfaces that also have content-protection facilities built in. Currently the most common high-end audio output interface is S/PDIF (Sony/Philips Digital Interface Format). Most newer audio cards, for example, feature TOSlink digital optical output for high-quality sound reproduction, and even the latest crop of motherboards with integrated audio provide at least coax (and often optical) digital output. Since S/PDIF doesn't provide any content protection, Vista requires that it be disabled when playing protected content [Note E]. In other words if you've sunk a pile of money into a high-end audio setup fed from an S/PDIF digital output, you won't be able to use it with protected content.

Say you've just bought Pink Floyd's "The Dark Side of the Moon", released as a Super Audio CD (SACD) in its 30th anniversary edition in 2003, and you want to play it under Vista. Since the S/PDIF link to your amplifier/speakers is regarded as insecure for playing the SA content, Vista disables it, and you end up hearing a performance by Marcel Marceau instead of Pink Floyd.

Similarly, component (YPbPr) video will be disabled by Vista's content protection, so the same applies to a high-end video setup fed from component video. But what if you're lucky enough to have bought a video card that supports HDMI digital video with HDCP content-protection? There's a good chance that you'll have to go out and buy another video card that really *does* support HDCP, because until quite recently no video card on the market actually supported it even if the vendor's advertising claimed that it did. As the site that first broke the story in their article The Great HDCP Fiasco puts it:

> "None of the AGP or PCI-E graphics cards that you can buy today support HDCP [...] If you've just spent $1000 on a pair of Radeon X1900 XT graphics cards expecting to be able to playback HD-DVD or Blu-Ray movies at 1920x1080 resolution in the future, you've just wasted your money [...] If you just spent $1500 on a pair of 7800GTX 512MB GPUs expecting to be able to play 1920x1080 HD-DVD or Blu-Ray movies in the future, you've just wasted your money".

(The two devices mentioned above are the premium supposedly-HDCP-enabled cards made by the two major graphics chipset manufacturers ATI and nVidia). ATI was later subject to a class-action lawsuit by its customers over this deception. As late as August of 2006, when Sony announced its Blu-Ray drive for PCs, it had to face the embarrassing fact that its Blu-Ray drive couldn't actually play Blu-Ray disks in HD format:

> "Since there are currently no PCs for sale offering graphics chips that support HDCP, this isn't yet possible".

The same goes for high-resolution LCD monitors. One of the big news items at CES 2007 was Samsung's 1920x1200 HD-capable 27" LCD monitor, the Syncmaster 275T, released at a time when everyone else was still shipping 24" or 25" monitors as their high-end product. The only problem with this amazing HD monitor is that Vista won't display HD content on it because it doesn't consider any of its many input connectors (DVI-D, 15-pin D-Sub, S-Video, and component video) secure enough. So you can do almost anything with this HD monitor except view HD content on it.

If you have even more money to burn, you can go for the largest (conventional) computer monitor made, the Samsung's stupidly large (for a computer monitor) 46" SyncMaster 460PN. Again though, Vista won't display HD content on it, turning your $4,000 purchase into a still-image picture frame (oddly enough, this monitor has been advertised as "HDTV ready" by retailers even though you can't display HD images on it, although in practice the term "HD-ready" has been diluted close to meaninglessness).

In order to appropriately protect content, Vista will probably have to disable any special device features that it can't directly control. For example many sound cards built on C-Media chipsets (which in practice is the vast majority of them) support Steinberg's ASIO (Audio Stream I/O), a digital audio interface that completely bypasses the Windows audio mixer and other audio- related driver software to provide more flexibility and much lower latency than the Windows ones. ASIO support is standard for newer C-Media hardware like the CMI 8788. Since ASIO bypasses Windows' audio handling, it would probably have to be disabled, which is problematic because audiophiles and professional musicians require ASIO support specifically because of its much higher quality than the standard Windows channels.

Indirect Disabling of Functionality

As well as overt disabling of functionality, there's also covert disabling of functionality. For example PC voice communications rely on automatic echo cancellation (AEC) in order to work. AEC requires feeding back a sample of the audio mix into the echo cancellation subsystem, but with Vista's content protection this isn't permitted any more because this might allow access to premium content. What is permitted is a highly-degraded form of feedback that might possibly still sort-of be enough for some sort of minimal echo cancellation purposes.

The requirement to disable audio and video output plays havoc with standard system operations, because the security policy used is a so-called "system high" policy: The overall sensitivity level is that of the most sensitive data present in the system. So the instant that any audio derived from premium content appears on your system, signal degradation and disabling of outputs will occur. What makes this particularly entertaining is the fact that the downgrading/disabling is dynamic, so if the premium-content signal is intermittent or varies (for example music that fades out), various outputs and output quality will fade in and out, or turn on and off, in sync. Normally this behaviour would be a trigger for reinstalling device drivers or even a warranty return of the affected hardware, but in this case it's just a signal that everything is functioning as intended.

Decreased Playback Quality

Alongside the all-or-nothing approach of disabling output, Vista requires that any interface that provides high-quality output degrade the signal quality that passes through it if premium content is present. This is done through a "constrictor" that downgrades the signal to a much lower-quality one, then up- scales it again back to the original spec, but with a significant loss in quality. So if you're using an expensive new LCD display fed from a high- quality DVI signal on your video card and there's protected content present, the picture you're going to see will be, as the spec puts it, "slightly fuzzy", a bit like a 10-year-old CRT monitor that you picked up for $2 at a yard sale [Note F]. In fact the specification specifically still allows for old VGA analog outputs, but even that's only because disallowing them would upset too many existing owners of analog monitors. In the future even analog VGA output will probably have to be disabled. The only thing that seems to be explicitly allowed is the extremely low-quality TV-out, provided that Macrovision is applied to it.

The same deliberate degrading of playback quality applies to audio, with the audio being downgraded to sound (from the spec) "fuzzy with less detail" [Note G].

Amusingly, the Vista content protection docs say that it'll be left to graphics chip manufacturers to differentiate their product based on (deliberately degraded) video quality. This seems a bit like breaking the legs of Olympic athletes and then rating them based on how fast they can hobble on crutches.

The Microsoft specs say that only display devices with more than 520K pixels will have their images degraded (there's even a special status code for this, STATUS_GRAPHICS_OPM_RESOLUTION_TOO_HIGH), but conveniently omit to mention that this resolution, roughly 800x600, covers pretty much every output device that will ever be used with Vista. The abolute minimum requirement for Vista Basic are listed as 800x600 resolution (and an 800MHz Pentium III CPU with 512MB of RAM, which seems, well, "wildly optimistic" is one term that springs to mind). However that won't get you the Vista Aero interface, which makes a move to Vista from XP more or less pointless. The minimum requirements for running Aero on a Vista Premium PC are "a DX9 GPU, 128 MB of VRAM, Pixel Shader 2.0, and minimum resolution 1024x768x32", and for Aero Glass it's even higher than that. In addition the minimum resolution supported by a standard LCD panel is 1024x768 for a 15" LCD, and to get 800x600 you'd have to go back to a 10-year- old 14" CRT monitor or something similar. So in practice the 520K pixel requirement means that everything will fall into the degraded-image category.

(A lot of this OPM stuff seems to come straight from the twilight zone. It's normal to have error codes indicating that there was a disk error or that a network packet got garbled, but I'm sure Windows Vista must be the first OS in history to have error codes for things like "display quality too high").

Beyond the obvious playback-quality implications of deliberately degraded output, this measure can have serious repercussions in applications where high-quality reproduction of content is vital. For example the field of medical imaging either bans outright or strongly frowns on any form of lossy compression because artefacts introduced by the compression process can cause mis-diagnoses and in extreme cases even become life-threatening. Consider a medical IT worker who's using a medical imaging PC while listening to audio/video played back by the computer. This scenario is already very common, the CDROM drives installed in workplace PCs inevitably spend most of their working lives playing music or MP3 CDs to drown out workplace noise.

Now obviously CDs aren't (yet) regarded as premium content and so won't trigger Vista's content-protection measures, that's merely an example to illustrate how common it is for users to play back audio/video content while working. Let's say that instead of listening to music while they work, the user may have a humorous video that a workmate sent them or that they grabbed from YouTube, playing in the background that, and that unbeknownst to them this video is protected premium content. As a result, the video image will be subtly altered by Vista's content protection, potentially creating exactly the life-threatening situation that the medical industry has worked so hard to avoid. The scary thing is that there's no easy way around this - Vista will silently modify displayed content under certain (almost impossible-to-predict in advance) situations discernable only to Vista's built-in content-protection subsystem [Note H].

An interesting potential security threat, suggested by Karl Siegemund, occurs when Vista is being used to run a security monitoring system such as a video surveillance system. If it's possible to convince Vista that what it's communicating is premium content, the video (and/or audio) surveillance content will become unavailable, since it's unlikely that a surveillance center will be using DRM-enabled recording devices or monitors. I can just see this as a plot element in Ocean's Fifteen or Mission Impossible Six, "It's OK, their surveillance system is running Vista, we can shut it down with spoofed premium content".

Elimination of Open-source Hardware Support

In order to prevent the creation of hardware emulators of protected output devices, Vista requires a Hardware Functionality Scan (HFS) that can be used to uniquely fingerprint a hardware device to ensure that it's (probably) genuine. In order to do this, the driver on the host PC performs an operation in the hardware (for example rendering 3D content in a graphics card) that produces a result that's unique to that device type.

In order for this to work, the spec requires that the operational details of the device be kept confidential. Obviously anyone who knows enough about the workings of a device to operate it and to write a third-party driver for it (for example one for an open-source OS, or in general just any non-Windows OS) will also know enough to fake the HFS process. The only way to protect the HFS process therefore is to not release any technical details on the device beyond a minimum required for web site reviews and comparison with other products.

This potential "closing" of the PC's historically open platform is an extremely worrying trend. A quarter of a century ago, IBM made the momentous decision to make their PC an open platform by publishing complete hardware details and allowing anyone to compete on the open market. Many small companies, the traditional garage startup, got their start through this. This openness is what created the PC industry, and the reason why most homes (rather than just a few offices, as had been the case until then) have one or more PCs sitting in a corner somewhere. This seems to be a return to the bad old days of 25 years ago when only privileged insiders were able to participate.

Elimination of Unified Drivers

The HFS process has another cost involved with it. Most hardware vendors have (thankfully) moved to unified driver models instead of the plethora of individual drivers that abounded some years ago. Since HFS requires unique identification and handling of not just each device type (for example each graphics chip) but each variant of each device type (for example each stepping of each graphics chip) to handle the situation where a problem is found with one variation of a device, it's no longer possible to create one-size-fits-all drivers for an entire range of devices like the current Catalyst/Detonator/ForceWare drivers. Every little variation of every device type out there must now be individually accommodated in custom code in order for the HFS process to be fully effective, resulting in a re-balkanisation of drivers that have only just become available in a clean, unified form in the last few years.

If a graphics chip is integrated directly into the motherboard and there's no easy access to the device bus then the need for bus encryption (see "Unnecessary CPU Resource Consumption" below) is removed. Because the encryption requirement is so onerous, it's quite possible that this means of providing graphics capabilities will suddenly become more popular after the release of Vista. However, this leads to a problem: It's no longer possible to tell if a graphics chip is situated on a plug-in card or attached to the motherboard, since as far as the system is concerned they're both just devices sitting on the AGP/PCIe bus. The solution to this problem is to make the two deliberately incompatible, so that HFS can detect a chip on a plug-in card vs. one on the motherboard. Again, this does nothing more than increase costs and driver complexity.

Further problems occur with audio drivers. To the system, HDMI audio looks like S/PDIF, a deliberate design decision to make handling of drivers easier. In order to provide the ability to disable output, it's necessary to make HDMI codecs deliberately incompatible with S/PDIF codecs, despite the fact that they were specifically designed to appear identical in order to ease driver support and reduce development costs.

Denial-of-Service via Driver/Device Revocation

Once a weakness is found in a particular driver or device, that driver will have its signature revoked by Microsoft, which means that it will cease to function (in informal terms, your device gets bricked, i.e. turned into a brick). Details on exactly what happens are a bit vague here, the specs contain sentences like "the related driver would have to be revoked and a new driver would have to be deployed", however presumably some minimum functionality like generic 640x480 VGA support will still be available in order for the system to boot.

What this means is that a report of a compromise of a particular driver or device will cause all support for that device worldwide to be turned off until a fix can be found [Note I]. Again, details are sketchy, but if it's a device problem then presumably the device turns into a paperweight once it's revoked. If it's an older device for which the vendor isn't interested in rewriting their drivers (and in the fast-moving hardware market most devices enter "legacy" status within a year or two of their replacement models becoming available), all devices of that type worldwide become permanently unusable.

An example of this might be nVidia TNT2 video cards, which are still very widely deployed in business environments where they're all that you need to run Word or Outlook or Excel (or, for that matter, pretty much any non-gaming application). The drivers for these cards haven't been updated for quite some time for exactly that reason: You don't need the latest drivers for them because they're not useful with current games any more (if you go to the nVidia site and try and install any recent drivers, the installer will tell you to go back and download much older drivers instead). If a TNT2 device were found to be leaking content, it seems unlikely that nVidia would be interested in reviving discontinued drivers that it hasn't touched for several years, creating instant orphanware of the installed user base.

The threat of driver revocation is the ultimate nuclear option, the crack of the commissars' pistols reminding the faithful of their duty. The exact details of the hammer that vendors will be hit with is buried in confidential licensing agreements, but I've heard mention of multi-million dollar fines and embargoes on further shipment of devices alongside the driver revocation mentioned above.

This revocation can have unforeseen carry-on costs. Windows' anti-piracy component, WGA, is tied to system hardware components. Windows allows you to make a small number of system hardware changes after which you need to renew your Windows license (the exact details of what you can and can't get away with changing has been the subject of much debate). If a particular piece of hardware is deactivated (even just temporarily while waiting for an updated driver to work around a content leak) and you swap in a different video card or sound card to avoid the problem, you risk triggering Windows' anti-piracy measures, landing you in even more hot water. If you're forced to swap out a major system component like a motherboard, you've instantly failed WGA validation. Revocation of any kind of motherboard-integrated device (practically every motherboard has some form of onboard audio, and all of the cheaper ones have integrated video) would appear to have a serious negative interaction with Windows' anti-piracy measures.

The details of what will happen if a motherboard contains unused onboard audio capabilities and an additional sound card alongside it, and the motherboard drivers are revoked, is unknown. Windows can't tell that there's nothing connected to the cheap onboard audio because the user prefers to use their M-Audio Revolution 7.1 Surround Sound card instead, so it'll probably have to revoke the motherboard drivers even though they're not used for anything. Since virtually all motherboards contain onboard audio, this could prove quite problematic.

An entirely different DoS problem that applies more to HDMI-enabled devices in general has already surfaced in the form of, uhh, "DVI amplifiers", which take as input an HDMI signal and output a DVI signal, amplifying it in the process. Oh, and as a side-effect they forget to re-apply the HDCP protection to the output. These devices are relatively simple to design and build using off- the-shelf HDMI chips. Beyond the commercially-available models, individual hardware hackers have built their own protection-strippers using chip samples obtained from chip vendors. If you have the right credentials you can even get hardware evaluation boards designed for testing and development that do this sort of thing. Even more accessible than that are HD players with non-HDMI digital outputs, for example ones that contain an HD-SDI (SMPTE 292M) interface. HD-SDI is an unencrypted digital link typically used in TV studios but also available from various non-US sources as after-market sidegrades for standard HD players, providing better-than-HDMI image quality without the hassle of HDCP.

Now assume that the "DVI amplifier" manufacturer buys a truckload of HDMI chips (they'll want to get as many as they can in one go because they probably won't be able to go back and buy more when the chip vendor discovers what they're being used for). Since this is a rogue device, it can be revoked... along with hundreds of thousands or even millions of other consumer devices that use the same chip. If they're feeling particularly nasty, they can recycle the HDMI chips from junked TVs to ensure that the maximum possible damage to the consumer base occurs. Engadget have a good overview of this doomsday scenario.

Exactly what will happen when a key is leaked depends on how the attackers handle it. The way HD-DVD/Blu-Ray keying works is that a per-device key is used to decrypt the title key on the disk, and the title key is then in turn used to decrypt the content. So the chain of custody is Device key -> Title key -> Content. This level of indirection allows an individual device to be disabled by revoking the device key without making the disk unplayable on all devices, since other device keys can still decrypt the title key and thus the content (I've simplified this a bit to cut down the length of the explanation, see the AACS specification for more details).

The device key is tied to a particular device/player/vendor, but the title key is only tied to the content on disk. You can probably see where this is going... by publishing the device key, the attacker can cause general mayhem by forcing device revocation. On the other hand by publishing the title key the attacker can release the content in an untraceable manner, since it's not known which device key was used to leak the title key. In addition since there's no way to un-publish the title key (encrypted content + title key = unencrypted content), at that point it's game over for the content.

Decreased System Reliability

> "Drivers must be extra-robust. Requires additional driver development to isolate and protect sensitive code paths" — ATI.

Vista's content protection requires that devices (hardware and software drivers) set so-called "tilt bits" if they detect anything unusual. For example if there are unusual voltage fluctuations, maybe some jitter on bus signals, a slightly funny return code from a function call, a device register that doesn't contain quite the value that was expected, or anything similar, a tilt bit gets set. Such occurrences aren't too uncommon in a typical computer. For example starting up or plugging in a bus-powered device may cause a small glitch in power supply voltages, or drivers may not quite manage device state as precisely as they think. Previously this was no problem - the system was designed with a bit of resilience, and things will function as normal. In other words small variances in performance are a normal part of system functioning. Furthermore, the degree of variance can differ widely across systems, with some handling large changes in system parameters and others only small ones. One very obvious way to observe this is what happens when a bunch of PCs get hit by a momentary power outage. Effects will vary from powering down, to various types of crash, to nothing at all, all triggered by exactly the same external event.

With the introduction of tilt bits, all of this designed-in resilience is gone. Every little (normally unnoticeable) glitch is suddenly surfaced because it could be a sign of a hack attack, with the required reaction being that (from the spec) "Windows Vista will initiate a full reset of the graphics subsystem, so everything will restart". The effect that these tilt bits will have on system reliability should require no further explanation.

Content-protection "features" like tilt bits also have worrying denial-of- service (DoS) implications. It's probably a good thing that modern malware is created by programmers with the commercial interests of the phishing and spam industries in mind rather than just creating as much havoc as possible. With the number of easily-accessible grenade pins that Vista's content protection provides, any piece of malware that decides to pull a few of them will cause considerable damage. The homeland security implications of this seem quite serious, since a tiny, easily-hidden piece of malware would be enough to render a machine unusably unstable, while the very nature of Vista's content protection would make it almost impossible to determine why the denial-of- service is occurring. Furthermore, the malware authors, who are taking advantage of "content-protection" features, could claim protection under the DMCA against any attempts to reverse-engineer or disable the content- protection "features" that they're abusing.

Even without deliberate abuse by malware, the homeland security implications of an external agent being empowered to turn off your IT infrastructure in response to a content leak discovered in some chipset that you coincidentally happen to be using is a serious concern for potential Vista users. Non-US governments are already nervous enough about using a US- supplied operating system without having this remote DoS capability built into the operating system. And like the medical-image-degradation issue, you won't find out about this until it's too late, turning Vista PCs into ticking time bombs if the revocation functionality is ever employed.

Like the medical-imaging degradation example given earlier, it's possible to imagine all sorts of scenarios in which the tilt bits end up biting users. Consider a warship operating in a combat zone and equipped with Vista PCs for management of the vessel's critical functions which does nothing more wrong that to suffer a severe jolt from a near miss, scrambling the bus just enough to activate the tilt bits (without causing any other real damage). In one infamous incident in September 1997, Windows NT managed to disable the Aegis missile cruiser USS Yorktown ("NT Leaves Navy "Smart Ship" dead in the water", Government Computer News, 13 July 1998). Now Windows Vista can do the same thing via a by-design feature of the OS [Note J]. This issue, unless it can be clearly resolved, would make the use of Vista PCs unacceptable for any applications that have any hint of unusual environmental conditions such as high altitude, environmental variations, shock, and so on.

Some contributors have commented that they can't see the revocation system ever being used because the consumer backlash would be too enormous, but then the legal backlash from not going ahead could be equally extreme. The only real indication that we have for how committed Microsoft really are to this is the amazing speed with which Microsoft released a patch for the WMDRM (Windows Media DRM) vulnerability, which they rushed out at a speed that even the most virulent worm never produced. This would seem to indicate that they're pretty serious about this, since they prioritised it above any conventional non-DRM-related security problem.

Increased Hardware Costs

> "Cannot go to market until it works to specification... potentially more respins of hardware" — ATI.
>
> "This increases motherboard design costs, increases lead times, and reduces OEM configuration flexibility. This cost is passed on to purchasers of multimedia PCs and may delay availability of high-performance platforms" — ATI.

Vista includes various requirements for "robustness" in which the content industry, through "hardware robustness rules", dictates design requirements to hardware manufacturers. The level of control the content producers have over technical design details is nothing short of amazing. As security researcher Ed Felten quoted from Microsoft documents on his freedom-to-tinker web site about a year ago:

> "The evidence [of security] must be presented to Hollywood and other content owners, and they must agree that it provides the required level of security. Written proof from at least three of the major Hollywood studios is required".

So if you design a new security system, you can't get it supported in Windows Vista until well-known computer security experts like MGM, 20th Century-Fox, and Disney give you the go-ahead (this gives a whole new meaning to the term "Mickey-Mouse security"). It's absolutely astonishing to find paragraphs like this in what are supposed to be Windows technical documents, since it gives Hollywood studios veto rights over Windows security mechanisms.

As an example of these "robustness rules", only certain layouts of a board are allowed in order to make it harder for outsiders to access parts of the board. Possibly for the first time ever, computer design is being dictated not by electronic design rules, physical layout requirements, and thermal issues, but by the wishes of the content industry. Apart from the massive headache that this poses to device manufacturers, it also imposes additional increased costs beyond the ones incurred simply by having to lay out board designs in a suboptimal manner. Video card manufacturers typically produce a one-size- fits-all design (often a minimally-altered copy of the chipset vendor's reference design, as illustrated by one product review which shows five virtually identical cards from different vendors with the only noticeable difference being the logo on the heatsink), and then populate different classes and price levels of cards in different ways. For example a low-end card will have low-cost, minimal or absent TV-out encoders, DVI circuitry, RAMDACs, and various other add-ons used to differentiate budget from premium video cards. You can see this on the cheaper cards by observing the unpopulated bond pads on circuit boards, and gamers and the like will be familiar with cut-a-trace/resolder-a-resistor sidegrades of video cards.

An example of omitting components from a high-end card to create a mid-range card clearly shows the large red rectangular area to the far left of the card, which is where the manufacturer has omitted a component to produce a lower- cost model. The same thing is visible in another card. Conversely, an (at the time it was released) top-of-the-line card with optional components fitted shows an additional chip to the left of the large square heatsink+fan that handles video encoding and can be added or removed (along with other optional components) to create different levels of cards at different price points. The automotive industry does the same thing, you have one basic model of each car type and 10,000 extras and options to suit everyone's needs and pockets.

In some cases the addition of extra circuitry isn't merely a convenient price- differentiation mechanism but is required for the device to function. Most never video cards have dual video outputs, and the higher-end ones tend to have dual-DVI out. However, many devices only provide have a single TMDS (Transition Minimized Differential Signaling, a high-speed serial data format) output for DVI signalling. The second output is provided by a DVO (Digital Video Out, not to be confused with Intel's similarly-named SVDO) port in combination with an external TMDS transmitter. In addition some high- resolution displays require multiple DVI/TMDS links because single-channel DVI doesn't have enough bandwidth to support very high resolutions, requiring external TMDS transmitters. You can see this in this image of the dual-link DVI output used to drive Apple's 30" Cinema Display (this actually requires two dual-link TMDS transmitters to support a second display, but I'll spare you the technical details of that one). The important point in all of this is the phrase "external TMDS transmitter", none of which meet the robustness requirements since they have direct access to the high- quality digital signal. Perversely enough, it's mostly the high-resolution displays advertised as suitable for HD content that require the external TMDS circuitry that makes them unable to meet the robustness requirements.

This problem is a nasty catch-22 from which there's no escape. In theory it would be possible to add a DVI-to-HDMI (with HDCP) encoder to bypass this (a typical example would be the Silicon Image Sil139x or Sil193x devices, which were specifically designed for this application. Silicon Image TMDS transmitters are widely used on graphics cards), but HDMI doesn't have the bandwidth to carry the high-definition images that this monitor displays. Even without explicit image degradation via constriction, the requirement to use the lower-quality HDMI link to carry what should be a DVI signal means that image quality is lost, and to make it even more painful the resulting graphics cards will be more expensive because it costs extra to add the quality- downgrading HDMI transmitter. In other words consumers will be paying extra in order to get a lower-quality image.

Even with lower-resolution monitors, the fact that the data signal is present in unprotected form when it enters the external encoder means that it probably won't meet the robustness requirements. (Exactly how this is meant to work is unspecified in any documentation that I've been able to get my hands on. It appears to be impossible to output a content-provider approved protected signal from a PC while also meeting the robustness requirements).

Vista's content-protection requirements eliminate the ability to accomodate different feature sets in a one-size-fits-all design, banning the use of separate TV-out encoders, DVI circuitry, RAMDACs, and other discretionary add- ons because feeding unprotected video to these optional external components would make it too easy to lift the signal off the bus leading to the external component. So everything has to be custom- designed and laid out so that there are no unnecessary accessible signal links on the board. This means that a low-cost card isn't just a high-cost card with components omitted, and conversely a high-cost card isn't just a low-cost card with additional discretionary components added, each one has to be a completely custom design created to ensure that no signal on the board is accessible.

This extends beyond simple board design all the way down to chip design. Instead of adding an external DVI/TMDS chip, it now has to be integrated into the graphics chip, along with any other functionality normally supplied by an external device. So instead of varying video card cost based on optional components, the chipset vendor now has to integrate everything into a one- size-fits-all premium-featured graphics chip, even if all the user wants is a budget card for their kid's PC (although given the popularity of graphics- intensive computer games, it's more likely that they'd be getting the budget card for their own PC).

A further example of external meddling in hardware vendors' product development and distribution can be found in the document that specifies what happens when a product is compromised in some way even though it's previously been found to be fully compliant with the robustness requirements:

> "Company shall promptly redesign the affected product [...] if such redesign is not possible or practical, cease manufacturing and selling such product"

This indicates that no matter how much dedication you show to the party line, it still won't help you when the chips are down. Some years ago a friend of mine was working for a company that was building a custom IT solution for a government department. When the day came time to sign off on it, everyone in the entire department who had signing authority called in sick rather than end up being the one who put their name to it. I can just imagine the corporate sick day at ATI, nVidia, Intel, VIA, SiS, when it came time to put someone's name to this gem, which gives Hollywood veto rights over your production lines and sales and distribution channels.

Increased Cost due to Requirement to License Unnecessary Third-party IP

> "We've taken on more legal costs in copyright protection in the last six to eight months than we have in any previous engagement. Each legal contract sets a new precedent, and each new one builds on the previous one" — ATI.

Protecting all of this precious premium content requires a lot of additional technology. Unfortunately much of this is owned by third parties and requires additional licensing. For example HDCP for HDMI is owned by Intel, so in order to send a signal over HDMI you have to pay royalties to Intel, even though you could do exactly the same thing for free over DVI (actually you could do it better, since DVI is provides a higher-quality link than HDMI). Similarly, since even AES-128 on a modern CPU isn't fast enough to encrypt high-bandwidth content, companies are required to license the Intel- owned Cascaded Cipher, an AES-128-based transform that's designed to offer a generally similar level of security but with less processing overhead.

The need to obtain unnecessary technology licenses extends beyond basic hardware IP. In order to demonstrate their commitment to the cause, Microsoft have recommended as part of their "robustness rules" that vendors license third-party code obfuscation tools to provide virus-like stealth capabilities for their device drivers in order to make it difficult to interfere with their operations or reverse-engineer them (for example the spec requires "use of techniques of obfuscation to disguise and hamper attempts to discover the approaches used"). Vendors like Cloakware and Arxan have actually added "robustness solutions" web pages to their sites in anticipation of this lucrative market. This must be a nightmare for device vendors, for whom it's already enough of a task getting fully functional drivers deployed without having to deal with adding stealth-virus-like technology on top of the basic driver functionality.

The robustness rules further complicate driver support by disallowing features such as driver debugging facilities in shipping drivers. Most Windows XP users will at one time or another have encountered a Windows crash message indicating that some application that they were using has terminated unexpectedly, and would they like to send debugging information to Microsoft to help fix the problem. Some device vendors even implement their own custom versions of this debugging support in their drivers, an example being ATI's VPU Recover, which captures graphics diagnostic and debugging information to send to ATI when a graphics device problem occurs. Since this debugging functionality could leak content or content-related security information, it can no longer be used with audio or video components, considerably complicating vendors' driver support and software enhancement processes (the ATI product manager referenced in the "Sources" section lists these additional testing and support costs as "potentially the highest cost of all").

Unnecessary CPU Resource Consumption

> "Since [encryption] uses CPU cycles, an OEM may have to bump the speed grade on the CPU to maintain equivalent multimedia performance. This cost is passed on to purchasers of multimedia PCs" — ATI.

In order to prevent tampering with in-system communications, all communication flows have to be encrypted and/or authenticated. For example content sent to video devices has to be encrypted with AES-128. This requirement for cryptography extends beyond basic content encryption to encompass not just data flowing over various buses but also command and control data flowing between software components. For example communications between user-mode and kernel-mode components are authenticated with OMAC message authentication-code tags, at considerable cost to both ends of the connection. The initial crypto handshake is:

```
  driver -> application: cert + nonce
application -> driver: RSA-OAEP-SHA512( nonce || key || seqNo1 || seqNo2 )
```

In this step the driver supplies its certificate to the calling application via DxgkDdiOPMGetCertificate() and a 128-bit nonce via DxgkDdiOPMGetRandomNumber(). This is either a COPP or an OPM certificate, with COPP being the older Windows XP content protection and OPM being the newer Windows Vista one. There's also a third type of certificate which the driver uses if it has a UAB (User-Accessible Bus). The certificates contain a 2048-bit RSA key which is used to encrypt a 40-byte payload containing the nonce provided by the driver, a 128-bit session key, and two 32-bit initial sequence numbers (they start at random values), the first number is for status messages via DxgkDdiOPMGetInformation() and the second for command messages via DxgkDdiOPMConfigureProtectedOutput().

Once the keys are set up, each function call is:

```
  in = OMAC( nonce || seqNo || data )
out = OMAC( nonce || seqNo || data )
```

(I've used conventional bits-on-the-wire notation for this, the values are actually fields in a structure so for example the sequence number is provided in the ulSequenceNumber member). This is very similar to the protocol used in SSL or SSH (in practice some steps like cipher suite negotiation are omitted, since there's a hardcoded set of ciphers used). Finding SSL being run inside a PC from one software module to another is just *weird*.

Needless to say, this extremely CPU-intensive mechanism is a very painful way to provide protection for content, and this fact has been known for many years. Twenty years ago, in their work on the ABYSS security module, IBM researchers concluded that the use of encrypted buses as a protection mechanism was impractical.

In order to prevent active attacks, device drivers are required to poll the underlying hardware every 30ms for digital outputs and every 150 ms for analog ones to ensure that everything appears kosher. This means that even with nothing else happening in the system, a mass of assorted drivers has to wake up thirty times a second just to ensure that... nothing continues to happen (Steve Gibson in his Security Now podcast with Leo Laporte calls Vista "an operating system that is insanely paranoid"). In addition to this polling, further device-specific polling is also done, for example Vista polls video devices on each video frame displayed in order to check that all of the grenade pins (tilt bits) are still as they should be. We already have multiple reports from Vista reviewers of playback problems with video and audio content, with video frames dropped and audio stuttering even on high-end systems [Note K]. Time will tell whether this problem is due to immature drivers or has been caused by the overhead imposed by Vista's content protection mechanisms interfering with playback.

An indication of the level of complexity added to the software can be seen by looking at a block diagram of Vista's Media Interoperability Gateway (MIG). Of the eleven components that make up the MIG, only two (the audio and video decoders) are actually used to render content. The remaining nine are used to apply content-protection measures.

On-board graphics create an additional problem in that blocks of precious content will end up stored in system memory, from where they could be paged to disk. In order to avoid this, Vista tags such pages with a special protection bit indicating that they need to be encrypted before being paged out and decrypted again after being paged in. Vista doesn't provide any other pagefile encryption, and will quite happily page banking PINs, credit card details, private, personal data, and other sensitive information, in plaintext. The content-protection requirements make it fairly clear that in Microsoft's eyes a frame of premium content is worth more than (say) a user's medical records or their banking PIN [Note L].

In addition to the CPU costs, the desire to render data inaccessible at any level means that video decompression can't be done in the CPU any more, since there isn't sufficient CPU power available to both decompress the video and encrypt the resulting uncompressed data stream to the video card. As a result, much of the decompression has to be integrated into the graphics chip. At a minimum this includes IDCT, MPEG motion compensation, and the Windows Media VC-1 codec (which is also DCT-based, so support via an IDCT core is fairly easy). As a corollary to the "Increased Hardware Costs" problem above, this means that you can't ship a low-end graphics chip without video codec support any more.

The inability to perform decoding in software also means that any premium- content compression scheme not supported by the graphics hardware can't be implemented. If things like the Ogg video codec ever eventuate and get used for premium content, they had better be done using something like Windows Media VC-1 or they'll be a non-starter under Vista or Vista-approved hardware. This is particularly troubling for the high-quality digital cinema (D-Cinema) specification, which uses Motion JPEG2000 (MJ2K) because standard MPEG and equivalents don't provide sufficient image quality. Since JPEG2000 uses wavelet-based compression rather than MPEG's DCT-based compression, and wavelet-based compression isn't on the hardware codec list, it's not possible to play back D-Cinema premium content (the moribund Ogg Tarkin codec also used wavelet-based compression). Because *all* D-Cinema content will (presumably) be premium content, the result is no playback at all until the hardware support appears in PCs at some indeterminate point in the future. Compare this to the situation with MPEG video, where early software codecs like the XingMPEG en/decoder practically created the market for PC video. Today, thanks to Vista's content protection, the opening up of new markets in this manner would be impossible.

The high-end graphics and audio market are dominated entirely by gamers, who will do anything to gain the tiniest bit of extra performance, like buying Bigfoot Networks' $250 "Killer NIC" ethernet card in the hope that it'll help reduce their network latency by a few milliseconds. These are people buying $500-$1000 graphics and sound cards for which one single sale brings the device vendors more than the few cents they get from the video/audio portion of an entire roomful of integrated-graphics-and-sound PCs. I wonder how this market segment will react to knowing that their top-of-the-line hardware is being hamstrung by all of the content-protection "features" that Vista hogties it with?

Unnecessary Device Resource Consumption

> "Compliance rules require [content] to be encrypted. This requires additional encryption/decryption logic thus adding to VPU costs. This cost is passed on to all consumers" — ATI.

As part of the bus-protection scheme, devices are required to implement AES-128 encryption in order to receive content from Vista. This has to be done via a hardware decryption engine on the graphics chip, which would typically be implemented by throwing away a GPU rendering pipeline or two to make room for the AES engine.

Establishing the AES key with the device hardware requires further cryptographic overhead, in this case a 2048-bit Diffie-Hellman key exchange whose 2K-bit output is converted to a 128-bit AES key via a Davies-Meyer hash with AES as its block transformation component. In programmable devices this can be done (with considerable effort) in the device (for example in programmable shader hardware), or more simply by throwing out a few more rendering pipelines and implementing a public-key-cryptography engine in the freed-up space.

Needless to say, the need to develop, test, and integrate encryption engines into audio/video devices will only add to their cost, as covered in "Increased Hardware Costs" above, and the fact that they're losing precious performance in order to accommodate Vista's content protection will make gamers less than happy.

Final Thoughts

> "No amount of coordination will be successful unless it's designed with the needs of the customer in mind. Microsoft believes that a good user experience is a requirement for adoption" — Microsoft.

> "The PC industry is committed to providing content protection on the PC, but nothing comes for free. These costs are passed on to the consumer" — ATI.

At the end of all this, the question remains: Why is Microsoft going to this much trouble? Ask most people what they picture when you use the term "premium-content media player" and they'll respond with "A PVR" or "A DVD player" and not "A Windows PC". So why go to this much effort to try and turn the PC into something that it's not?

In July 2006, Cory Doctorow published an analysis of the anti-competitive nature of Apple's iTunes copy-restriction system which looked at the benefits of restrictive DRM for the company that controls the DRM. The only reason I can imagine why Microsoft would put its programmers, device vendors, third-party developers, and ultimately its customers, through this much pain is because once this copy protection is entrenched, Microsoft will completely own the distribution channel. In the same way that Apple has managed to acquire a monopolistic lock-in on their music distribution channel (an example being the Motorola ROKR fiasco, which was so crippled by restrictions that a Fortune magazine senior editor reviewed it as the STNKER), so Microsoft will totally control the premium-content distribution channel. Not only will they be able to lock out any competitors, but because they will then represent the only available distribution channel they'll be able to dictate terms back to the content providers whose needs they are nominally serving in the same way that Apple has already dictated terms back to the music industry: Play by Apple's rules, or we won't carry your content. The result will be a technologically enforced monopoly that makes their current de-facto Windows monopoly seem like a velvet glove in comparison [Note M].

The onerous nature of Vista's content protection also provides a perverse incentive to remove the protection measures from the content, since for many consumers that'll be the only way that they can enjoy their legally-acquired content without Vista's DRM getting in the way. This is already illustrated in the "Quotes" and "Footnotes" sections, where the people bypassing HD-DVD protection measures aren't hardcore video pirates but ordinary consumers who can't even play their own legitimately-acquired content. The sheer obnoxiousness of Vista's content protection may end up being the biggest incentive to piracy yet created. Even without overt "piracy" (meaning bypassing restrictions in order to play legally-purchased media), it makes very sound business sense for companies to produce hardware that bypasses the problem, just as they have already with region-free play-anything DVD players. Perhaps Hollywood should heed the advice given in one of their most famous productions: "The more you tighten your grip, the more systems will slip through your fingers".

Overall, Vista's content-protection functionality seems like an astonishingly short-sighted piece of engineering, concentrating entirely on content protection with no consideration given to the enormous repercussions of the measures employed. It's something like the PC equivalent of the (hastily dropped) proposal mooted in Europe to put RFID tags into high-value banknotes as an anti-counterfeiting measure, completely ignoring the fact that the major users of this technology would end up being criminals who would use it to remotely identify the most lucrative robbery targets.

To add insult to injury, consider what this enormous but ultimately wasted effort could have been put towards. Microsoft is saying that Vista will be the most secure version of Windows yet, but they've been saying that for every new Windows release since OS security became a selling point. I don't think anyone's under any illusions that Vista PCs won't be crawling with malware shortly after the bad guys get their hands on them (there were already Vista exploits up for sale before the OS even hit the shelves). But what if the Vista content-protection technology had instead been applied towards malware protection? Instead of a separate protection domain for video playback, we might have a separate protection domain for banking and credit card details. Instead of specialised anti-debugging techniques to stop users getting at even one frame of protected content, we could have those same techniques combatting malware hooking itself into the OS. The list goes on and on, with all of the effort being misapplied to DRM when it could have been used to combat malware instead. What a waste. What a waste.

The worst thing about all of this is that there's no escape. Hardware manufacturers will have to drink the kool-aid (and the reference to mass suicide here is deliberate [Note N]) in order to work with Vista: "There is no requirement to sign the [content-protection] license; but without a certificate, no premium content will be passed to the driver". Of course as a device manufacturer you can choose to opt out, if you don't mind your device only ever being able to display low-quality, fuzzy, blurry video and audio when premium content is present, while your competitors don't have this (artificially-created) problem.

As a user, there is simply no escape. Whether you use Windows Vista, Windows XP, Windows 95, Linux, FreeBSD, OS X, Solaris (on x86), or almost any other OS, Windows content protection will make your hardware more expensive, less reliable, more difficult to program for, more difficult to support, more vulnerable to hostile code, and with more compatibility problems. Because Windows dominates the market and device vendors are unlikely to design and manufacture two different versions of their products, non-Windows users will be paying for Windows Vista content-protection measures in products even if they never run Windows on them.

Here's an offer to Microsoft: If we, the consumers, promise to never, ever, ever buy a single HD-DVD or Blu-Ray disc containing any precious premium content [Note O], will you in exchange withhold this poison from the computer industry? Please?

Acknowledgements

This document was put together with input from various sources, including a number that requested that I keep their contributions anonymous (in some cases I've simplified or rewritten some details to ensure that the original, potentially traceable wording of non-public documents isn't used). Because it wasn't always possible to go back to the sources and verify exact details, it's possible that there may be some inaccuracies present, which I'm sure I'll hear about. No doubt Microsoft (who won't want a view of Vista as being broken by design to take root) will also provide their spin on the details.

In addition to the material present here, I'd be interested in getting further input both from people at Microsoft involved in implementing the content protection measures and from device vendors who are required to implement the hardware and driver software measures. I know from the Microsoft sources that contributed that many of them care deeply about providing the best possible audio/video user experience for Vista users and are quite distressed about having to spend time implementing large amounts of anti-functionality when it's already hard enough to get things running smoothly without the intentional crippling. I'm always open to further input, and will keep all contributions confidential unless you give me permission to repeat something. If you're concerned about traceability, grab a disposable account at Yahoo, Gmail, or some similar provider and contact me through that. If you're worried about being identified via the machine you connect to the email provider with, use an Internet cafe to send the message - just use standard common-sense precautions. If you want to encrypt things, my PGP key is linked from my home page.

(In case the above hints aren't obvious enough, if you work for nVidia, ATI, VIA, SiS, Intel, ..., I'd *really* like to get your comments on how all of this is affecting you).

Sources

Because this writeup started out as a private discussion in email, a number of the sources used were non-public.

The best public sources that I know of are:

- Output Content Protection and Windows Vista from WHDC.
- Windows Longhorn Output Content Protection from WinHEC.
- How to Implement Windows Vista Content Output Protection from WinHEC.
- Protected Media Path and Driver Interoperability Requirements from WinHEC.

  (Note that the cryptography requirements have changed since some of the information above was published. SHA-1 has been deprecated in favour of SHA-256 and SHA-512, and public keys seem to be uniformly set at 2048 bits in place of the mixture of 1024 bits and 2048 bits mentioned in the presentations).

An excellent analysis from one of the hardware vendors involved in this comes from ATI, in the form of Digital Media Content Protection from WinHEC. This points out (in the form of PowerPoint bullet-points) the manifold problems associated with Vista's content-protection measures, with repeated mention of increased development costs, degraded performance and the phrase "increased costs passed on to consumers" pervading the entire presentation like a mantra.

In addition there have been quite a few writeups on this (although not going into quite as much detail as this document) in magazines both online and in print, one example being PC World's feature article Will your PC run Windows Vista?" which covers this in the appropriately-titled section "Multimedia in chains". Audience reactions to these proposals at WinHEC are covered in Longhorn: tough trail to PC digital media published in EE Times, unfortunately you need to be a subscriber to read this but you may be able to find accessible cached copies using your favourite search engine. The EFF has an overview of the effects of Vista's revocation mechanisms in "Protected Media Path, Component Revocation, Windows Driver Lockdown".

Use, Modification, and Redistribution

This document is licensed under the Creative Commons Attribution 2.5 License, http://creativecommons.org/licenses/by/2.5/. This means that you can copy, distribute, display, and perform the work, and make derivative works, provided that you credit the original author and provide a link back to the original work (at the URL given in the title). To quote the Creative Commons site, "This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered, in terms of what others can do with your works".

Appendices and Footnotes

The more formal section of the document ends here. The following sections contain various informal comments, thoughts, and other odds and ends. For people doing translations of this document, it's probably not worth trying to translate these sections.

Mini-FAQ

This document seems to produce various reactions that come up repeatedly. To respond to the more frequently-expressed views, I've added this mini-FAQ.

This is just Microsoft-bashing.

It's bad-technology bashing. If this had been done by Linus Torvalds, Steve Jobs, Alan Cox, or Theo de Raadt, I'd have said the same thing about it. As far as I'm concerned computers are tools to get a job done and not a platform for religious wars, and if something's bad I'll say so regardless of who's doing it. Just for the record I run various versions of Windows on ... [counting] ... seven of my machines (the rest are a mixture of Linux, FreeBSD, and occasionally Solaris), so I'd be a rather unlikely Microsoft detractor if I have their software all over my machines.

This is a biased writeup.

Perhaps, but then I challenge anyone to read the specifications given in the "Sources" section above and write a positive analysis of Vista's content protection. Someone has to point out these problems, and it happened to be me in this case, but I think anyone with technical skills who reads the relevant documents would come to a similar conclusion.

This is all a pile of FUD.

The process that leads to comments like this tends to be (1) Quickly skim through this document, (2) Decide that it sounds a bit implausible (possibly even before performing step 1), (3) Post a rant saying it's FUD. To pick one particular example, a Digg reader's reaction to the section of text that states there isn't sufficient CPU power available for both decompression and encryption was:

> I'm sorry, where does this come from? You do realize that this is completely uncited, and very likely wrong? Entire paragraphs that follow are based on this magical detail pulled out of thin air. [...] I'm no fan of this asinine DRM bullshit, but the scenarios and postulates put forth in this article are complete rubbish.

Referring to the very first source listed in the "Sources" section shows that this is picked not from thin air but from Microsoft's own documentation:

> The problem with regular AES is that it takes about 20 CPU clocks to encrypt each byte. This is OK for compressed or semi-compressed video, but for the multiple HD uncompressed case, it is too much even for a 2006 processor [referring to the fastest CPU available at the time the document was written].

and then again:

> In the case of premium content, whether video can play back smoothly when using regular AES with uncompressed video will be a function of the resolution of the uncompressed video and the power of the processor. It is unlikely to work well in 2006 for uncompressed HD premium content

If you don't believe what you've read here, go back to Microsoft's own documentation and read that (in fact read the Microsoft documents no matter what you believe, because they're quite scary). If you still think it's FUD then you can at least post informed comments about it.

Microsoft is only doing this because Hollywood/the music industry is forcing them to.

"We were only following orders" has historically worked rather poorly as an excuse, and it doesn't work too well here either. While it's convenient to paint an industry that sues 12-year-old kids and 80- year-old grandmothers as the scapegoat, no-one's holding a gun to Microsoft's head to force them do this. The content industry is desperate to get its content onto PCs, and it would have been quite easy for Microsoft to say "Here's what we'll do with Vista, take it or leave it. We won't seriously cripple our own and our business partners' products just to suit your whims". In other words they could make it clear to Hollywood who's the tail and who's the dog.

Here's an illustrative story about what can happen when the content- industry tail tries to wag the dog. About 10-15 years ago, music companies told a bunch of NZ TV stations that they had to pay fees in order to screen music videos. The TV stations disagreed, saying that they were providing free advertising for the music companies, and if they didn't like that then they'd simply stop playing music videos. So they stopped playing all music videos.

After a few weeks, cracks stated to appear as the music companies realised just how badly they needed the TV channels. One of the music companies bought an entire prime-time advertising block (at phenomenal cost, this wasn't a single 30-second slot but every slot in an entire prime-time ad break) just to play one single new music video.

Shortly afterwards, music videos reappeared on TV. The details of the settlement were never made public, but I imagine it consisted of a bunch of music company execs on their knees begging the TV stations to start playing music videos again and let's please never bring this matter up again.

It's the same with Microsoft, the content industry needs them as badly (or more badly) than Microsoft needs the content industry. Claiming that they're only following orders from Hollywood is a red herring - if Microsoft declined to implement this stuff, Hollywood would have to give in because they can't afford to lock themselves out of 95% of the market, in the same way that the music companies couldn't afford to cut out their primary advertising channel.

You're just upset because you can no longer steal content under Vista.

Yes, someone really did send me email with this claim in it. It's silly enough that I just had to include it for the amusement value :-).

Open Questions

There are a number of open questions about Vista's content protection that probably won't be able to be answered until some months after its wide deployment when users can report on real-life experiences, because no-one seems to know how certain things will work.

Question 1.

How easy is it to get HD content around the outside of Vista's content- protection? Looking at the block diagrams in the specification documents is:

User-space application

Vista content-protection interface

Vista content playback subsystem

Vista device drivers

Device hardware

Reading the specs, user-space applications are expected to call down into the Vista content-protection interface to play back content (one document actually uses the metaphor of the user-space application simply acting as a remote control for the Vista content-protection and playback subsystem). The question is, can a user-space application that chooses to opt out perform an end-run around the higher-level Vista interface and go directly to the low- level interface to get its content out without Vista's content-protection getting in the way? User feedback on Microsoft's own forums indicates that even using third- party playback software like the nVidia or Cyberlink decoders instead of the Vista one will result in playback being disabled when (in this case) the Vista Media Centre trial license expired.

Question 2.

How will all of this affect users who want to prepare HD content, protected or not? Given that the intent of Vista's content-protection is to ensure that no HD content ever leaves the system in usable form, how do you prepare the HD content? More importantly, since Vista happens to be a multitasking OS, how do you guarantee that as your HD content is being prepared, the presence of some other protected content somewhere in the system doesn't cause it to be silently degraded for "protection" purposes? Just how deep does the protection extend? If it's on a per-task or even per-thread level then any cross-task or cross-thread mechanism (e.g. process thread injection) can be used to compromise the content protection. On the other hand if it's "all your content are belong to us" whenever protected content is present then innocent content will be degraded along with protected content.

Question 3.

If you build it, they will come. Once the DRM mechanisms are in place, there's every reason to believe that any kind of content subject to any kind of copyright will try and take advantage of it. After all, why not? The tools are there, there's no reason not to use them. We already have so-called Enterprise DRM (E-DRM) that's intended to control access to documents like Microsoft Word, PDF documents, CAD files, and so on (about 20 years ago during the heyday of the DoD Orange Book this stuff was known as ORCON, originator- controlled access control). Now that DRM is integrated into Vista as a core technology there's no knowing how far this can be taken in the future. What will computing be like in a few years time?

Question 4.

I've both read on the web and received via email endless reports of people unable to play HD-DVD and Blu-Ray content on Windows PCs, both Vista (beta) and XP. Has anyone actually been able to *play* HD-DVD or Blu-Ray content (i.e. the material that Vista classes as premium content) under Windows? If so, what HD drive, player software, graphics card, and monitor did you use?

(So far I've recevied *zero* reports of anyone actually being able to play HD content. Anyone?).

Glossary

This document was originally written for a technical audience and so used a number of technical terms that would have been familiar to its target audience but not to the general public. This glossary provides a few basic definitions, for more details see your favourite online source, for example Wikipedia.

DRM
: "Digital Rights Management" or "Digital Restrictions Management" or "Defective Recorded Media" (when applied to CDs, since it deliberately introduces defects into the media). Combine all three and you have a general idea of what DRM is.

HD
: High definition, technically meaning video content of 1920 x 1080 (1080p) resolution, but more generally anything with better than generic TV-quality resolution. In their specs, Microsoft regard anything with more than 520K pixels or 800x600 resolution as premium content that needs to be downgraded before displaying it to the user.

HD-DVD
: One of the proposed successors to DVDs, capable of storing HD content.

(More definitions to come).

AEC (automatic echo cancellation)
AES-128
AGP
CRT
DVI
DMCA
HDCP
HDMI
HFS (hardware functionality scan)
IDCT
IP (intellectual property)
JPEG2000
MPEG
OMAC
PCIe
RAMDAC
S/PDIF
TOSlink
VGA
VPU
WGA

Quotes

A few fun quotes, included for amusement value.

"I propose that each copy of the OS should ship with an orange jumpsuit and sensory deprivation goggles, since all Vista users have been unilaterally declared 'enemy combatants' by the content apparatchiki" — Daniel Nevin.

"Windows Vista? And what a vista! All you see as you look around your garden is a 60foot high brick wall" — Crosbie Fitch.

"welcome to the new world of DRM where expensive pieces of hardware across the world could potentially be remotely rendered useless by over-zealous copyright holders. Way to go, Hollywood!" — Chip Mulligan.

"I can not only say that the idea [of tilt bits] is basically insane, but I can also see hardware manufacturers refusing to implement tilt bits, or more likely, faking their functionality" — Dave Walker.

"I purchased a new DVD/SACD player (w/HDMI out), surround-sound receiver/amp (non-HDMI i/o - they're still too expensive for me), & LCD TV with HDMI input. My DVD/SACD player was connected to the SSamp via a nice single simple optical cable (& HDMI cable to the TV). I figured that would be all I need, keeping a digital path all the way to the SSamp (& TV). Wrong! It worked beautifully until I played my one & only SACD. No sound came forth! Huh? I read the DVD/SACD player manual: in brief small print, "When playing SACDs, audio is output only from the 5.1ch RCA analogue outputs" — Anthony May.

"I can't playback HD because I need to upgrade my 2 (SLI'd) Nvidia Quadro 4500's (~$2000) to a $200 FX7600GT because it supports HDCP. I can't wait till someone cracks this DRM/HDCP/AACS crap" — "Sy".

"Thanks alot, Muslix64 [the author of the purported HD-DVD crack] you're not the only one with a monitor/vid card that doesn't support hdcp, your work is greatly appreciated" — "yodoso".

"The funny thing is that I cant see how HDCP will actually even prevent piracy. In fact the only thing I can see it doing is encouraging piracy because everyone whose bought a new computer/monitor/HDTV in the last few years which don't have HDCP are now screwed out of the several thousand dollar purchases. So instead of buying new products they will turn to pirated/cracked Blu-ray/HD-DVDs which will work without the HDCP" — "Gizza".

"The HDCP scheme will serve to make the illegal product the most full featured and least restrictive, and thus the most attractive to the consumer. Add in the expense of buying new equipment to view the legal content (when existing equipment is perfectly capable) and the performance drain imposed by in-line encryption/decryption and they've put out the biggest incentive to piracy yet" — "Greg".

"The HDCP (high-definition content protection) overlords are coming to get us. They are basically saying you can't watch video unless you have a digital monitor and a special video card that supports the end-to-end content protection they have built; So that you, the un-trusted-consumer-who-bought- their-expensive-product, can't possibly make backup copies or anything else with that fancy new HD-DVD or Blu-ray disc you have" — "verifex".

"Digital rights management technology will still fail to prevent widespread infringement. In a related development, pigs will still fail to fly. I predict that every year, and it turns out to be true every year" — Ed Felten.

"Microsoft wasted no time; it issued a patch three days after learning about the hack. There's no month-long wait for copyright holders who rely on Microsoft's DRM. This clearly demonstrates that economics is a much more powerful motivator than security" — Bruce Schneier on Microsoft's DRM re-enabling patch for FairUse4WM.

"As a not-so-long-ago electronics design engineer, I can imagine the rage & pain felt by engineers & their employers [...] This is total insanity from anyone's perspective except the content providers, and they don't care because it's everyone else who's picking up the tab for it!" — Anthony May.

> "Good job, industry! Spend an incredible amount of time and effort developing the next generation of video quality only to step on it BEFORE THERE'S EVEN A DECIDED UPON STANDARD in the name of Copy Protection which will just be outflanked by a couple of 14 year old hackers and distributed over BitTorrent anyway" — "SweetMercury".
>
> "Sony, MS, movie studios... here's the deal. You've screwed up so bad that i'm not buying either HD drive option until they're so cheap that I end up getting one included with my computer because it was the minimum optical drive" — "zweben".
>
> "There has to be a whole new division at Microsoft. The "Office of Consumer Apology" or something. Responsible for "I'm sorry your content didn't PlayForSure. That isn't meant to be literal you know" and "yes, I know you're supposed to be able to play HD at full resolution, but you see, your cable has a kink in it, which changed the electrical characteristics slightly and, well, I guess I'm just sorry" — Blake Ramsdell.
>
> "your latest girly moan bitch rant is making the rounds on every news site just about isn't it? are you on cnn yet? are women throwing their panties at you?" — A friend (who requested anonymity).

Footnotes

Note A: This comment was inspired by Sir Gerald Kaufman's similar comment about the British Labour Party's 1983 election manifesto, which resulted in Labour turning in its worst election results since its founding (it was so bad that Labour's opponents in the election reprinted and distributed it themselves. Maybe Apple could take a hint from this and use Microsoft's content-protection details in their advertising for OS X). At 44 pages, Microsoft's "Output Content Protection and Windows Vista" document squeezes out Labour's 37-page manifesto to take the crown.

Note B: This document uses "cost" in the sense of "penalty", "damage", "harm", "injury" and "loss" rather than the more financial "expense", "outlay", and "price". A full financial analysis would require a top-to-bottom internal audit of the design, development, production, distribution, support, and legal costs for each vendor involved, something for which even the vendors themselves would have difficulty producing a precise figure.

Note C: In order for content to be displayed to users, it has to be copied numerous times. For example if you're reading this document on the web then it's been copied from the web server's disk drive to server memory, copied to the server's network buffers, copied across the Internet, copied to your PC's network buffers, copied into main memory, copied to your browser's disk cache, copied to the browser's rendering engine, copied to the render/screen cache, and finally copied to your screen. If you've printed it out to read, several further rounds of copying have occurred. Windows Vista's content protection (and DRM in general) assume that all of this copying can occur without any copying actually occurring, since the whole intent of DRM is to prevent copying. If you're not versed in DRM doublethink this concept gets quite tricky to explain, but in terms of quantum mechanics the content enters a superposition of simultaneously copied and uncopied states until a user collapses its wave function by observing the content (in physics this is called quantum indeterminacy or the observer's paradox). Depending on whether you follow the Copenhagen or many-worlds interpretation of quantum mechanics, things then either get weird or very weird. So in order for Windows Vista's content protection to work, it has to be able to violate the laws of physics and create numerous copies that are simultaneously not copies.

(Someone has pointed out that Microsoft is trying to implement a quantum encryption channel in software that attempts to make premium content non- observable, detecting problem states and discontinuing transmission if any are observed).

Note D: I'll make a prediction at this point that, given that it's trying to do the impossible, the Vista content protection will take less than a day to bypass if the bypass mechanism is something like a driver bug or a simple security hole that applies only to one piece of code (and can therefore be quickly patched), and less than a week to comprehensively bypass in a driver/hardware-independent manner. This doesn't mean that it'll be broken the day or week that it appears, but simply that once a sufficiently skilled attacker is motivated to bypass the protection, it'll take them less than a day or a week to do so.

(In a recent development, a sort of re-run of the DeCSS/Xing player story from a few years ago has occurred in which someone appears to have figured out how to extract HD-DVD and Blu-Ray keys from the PowerDVD player software, allowing all(?) HD disk content to be decrypted and played back on any HD display, without content-protection measures getting in the way (although the manufacturers of PowerDVD claim that they've done nothing wrong and won't be updating the player. The fact that the legally-purchased content wouldn't play on a legally-purchased player because the content protection got in the way appears to have been the motivating factor for the crack. The time taken was about a week, that information wasn't revealed until after I made my prediction above).

Note E: There is SCMS, but that has all the effectiveness of a "Keep out" sign.

Note F: As an example of an experience that's likely to become commonplace once more "premium content" is rolled out, Roger Strong reports from Canada that "I've just had my first experience with HD content being blocked. I purchased an HP Media Center PC with a built-in HD DVD player, together with a 24" 'high definition' 1920 x 1200 HP flat panel display (HP LP2465). They even included an HD movie, 'The Bourne Supremacy'. Sure enough, the movie won't play because while the video card supports HDCP content protection, the monitor doesn't. (It plays if I connect an old 14" VGA CRT using a DVI-to-VGA connector)". "muslix64" tells a similar tale: "when I disable my HD monitor, I can watch the movie, on my old VGA screen, but, what is the point of having a HD monitor and not being able to watch a HD movie on it". muslix64 was so upset at not being able to play his legitimately-purchased movies on his legitimately-purchased monitor attached to his legitimately-purchased player that he broke the AACS protection just to be able to see his own movies, see Note D above.

Note G: The question of how content producers other than the major studios who can afford expensive custom equipment are supposed to create and manipulate high-definition content has been raised by a number of readers. For example one contributor who works with people in the content industry comments that "I have seen [smaller content producers] going from just recording weddings and the like, to ones that have gone all the way to make a full featured movie. They have gone through problems like where to edit HD material, which cameras to use, which format, etc. Their decisions have been based on availability of equipment to make their projects, not really costs". It has been suggested that the large content producers are quite happy with this situation, since it prevents any competition from more innovative, creative, and agile newcomers.

Note H: Philip Dorrell has a neat cartoon that illustrates this problem.

Note I: There is some confusion over exactly how much functionality gets disabled when a revocation occurs. The HDCP requirements are quite clear that once this happens (in technical terms once a revoked device's key selection vector (KSV, effectively it's unique ID) appears on a revocation list) the device is effectively dead since it won't be supplied with content any more (in informal terms, your device gets "bricked", i.e. turned into a brick). However the behaviour of devices subject to revocation in a mixed-content environment is made very unclear in the specs. Some documents imply that it's an HDCP-style kill switch ("Vista will [...] revoke any driver that is found to be leaking premium content [...] if the same driver is used for all the manufacturer's chip designs, then a revocation would cause all that company's products to need a new driver"), while others indicate that the device will still work, but be unable to render premium content. Exactly how well this hope can be realised in practice (if it can be realised at all) remains to be seen.

Note J: I see some impressive class-action suits to follow if this revocation mechanism ("bricking") is ever applied. Perhaps Microsoft or the content providers will buy everyone who owns a device that inadvertently leaks content and is then disabled by the revocation process replacement hardware for their system, although that will in turn trigger the WGA time-bomb.

For anyone who's read "Guns of August", the situation seems a bit like pre- WWI Europe with people sitting on step 1 of enormously complex battle plans that can't be backed out of once they're triggered, no matter how obvious it is that going ahead with them is a bad idea. Driver revocation is a lose/lose situation for Microsoft, they're in for some serious pain whether they do or they don't. Their lawyers must have been asleep when they let themselves get painted into this particular corner - the first time a revocation takes out a hospital, foreign government department, air traffic control system, or whatever, they've guaranteed themselves a front-row seat in court proceedings for the rest of their natural lives.

(Several people have suggested that this was deliberate in order for the lawyers to guarantee themselves lifetime employment, but this seems highly unlikely. Firstly, lawyers have an obligation to protect their clients, so deliberately getting a client into trouble in order to generate more work would be a severely career-limiting move. Secondly, they're corporate in- house counsel rather than independent counsel, so they'll get paid anyway. Making more work for themselves would not be a big priority for them).

Note K: Some insider comments indicate that it'll be mid-2007 at least before Vista's non-Microsoft graphics and sound drivers are finished enough to be stable and reliable. Vendors are frantically rushing to get drivers ready in time for Vista's release (they didn't even make it onto the RTM media and will have to be downloaded after the install), but even those have been described as 'beta-quality at best'. No doubt we'll hear more of this at Vista's public release.

Note L: The Enterprise and Ultimate editions of Vista do feature this type of encryption, but the features of these high-end versions will never get into the hands of typical users. In addition it's an all-or-nothing encryption where (to quote Microsoft) "all user and system files are encrypted" when what really counts is swap-file encryption, since that's what contains copies of sensitive in-memory data. The OpenBSD approach of generating a random swap- file encryption key at boot time and encrypting any memory data that gets paged to disk is the correct way to handle this.

Note M: Video and audio playback aren't the only areas in which Vista's inner control freak comes to the fore. A Gamasutra article Vista Casts A Pall On PC Gaming looks at Vista's new Game Explorer "feature", which subjects all games to parental controls. Any games vendor who can't afford to obtain an extremely expensive ESRB rating has their software treated as "Not Rated", the equivalent of the MPAA's X- rating for films which was originally intended to mean "Not Rated" but has since become synonymous with hardcore porn. Obviously any parent would block Not-Rated content, which means that anyone who can't afford to pay the ESRB (in other words any small, independent game producers, including the ones most likely to produce free and low-cost family-rated games) can't work with Vista's Game Explorer. This seems like yet another area of Vista in which the words "anticompetitive" and "class-action" will feature prominently in the future.

Note N: The "kool-aid" reference may be slightly unfamiliar to non-US readers, it's a reference to the 1978 Jonestown mass- suicide in which Jim Jones' followers drank Flavor Aid laced with poison in order to demonstrate their dedication to the cause. In popular usage the term "kool-aid" is substituted for Flavor Aid because it has more brand recognition. There's also an earlier, less well-known link to fruit juice laced with LSD, I'll avoid the obvious comment linking that and some of the thinking behind Vista's content protection.

Note O: If I do ever want to play back premium content, I'll wait a few years and then buy a $50 Chinese-made set-top player to do it, not a $1000 Windows PC. It's somewhat bizarre that I have to go to communist China in order to find vendors who actually understand the consumer's needs.

A reductio ad absurdum solution to the "premium-content problem", proposed by a Slashdot reader, is to add support to Windows Vista for a black-box hardware component that accepts as input encrypted compressed premium content and produces as output encrypted (or otherwise protected) decoded premium content. In other words, move the entire mass of hardware, driver, and software protection into a dedicated black box that's only used in media PCs where it's (arguably) required.

Now compare this add-on black box to the canonical Chinese-made $50 media player. Why would anyone buy the black box (which will almost certainly cost more than $50) merely as an add-on to their already-expensive PC when they can buy a complete dedicated media player that does the same thing and more?

Author: Peter Gutman
Copied from: http://www.cs.auckland.ac.nz/~pgut001/pubs/vista_cost.html