Open Source "Linux Drivers" for ATI Graphics

While the the term "Linux Driver" is actually a misnomer, it is widely used in our field. Linux is an OS kernel (and associated kernel utilities) produced by the "Linux Group" headed by Linus Torvalds. Linux does not encompass the X Window System, or "X" in shorthand. X is the graphics sub-system used by most UNIX systems, including Linux, for handling graphics. X is a specification created twenty-some years ago that has evolved as an industry standard under the auspices of (mostly) the X Consortium through the last formal release known as X11R6.4. Since then, new releases of X seem to be more version changes to cover code updates of the XFree86 or X.org X servers, rather than X specification changes.

Originally, Linux gained its reputation for stability in applications that did not need graphics, hence did not run X. A typical "headless" application was Apache, the Web server. Reports of two or three years continuous up time running Apache was common. Linux with graphics was (is) another story. Reports of instability, poor performance, and lack of support are now quite common if serious graphics operations are involved.

Two major pieces of the X graphics sub-system (technically a UNIX application) are the X server and the "graphics driver code" that is tied directly to the use of graphics chip hardware. This "graphics driver code" is generally considered as also having two parts - the device independent, and device dependent parts. That is, code that is independent of which graphics chip is used (the "dix"), and code that is tied directly to a specific graphics chip (the "ddx"). The ddx part is somewhat analogous to a "driver" written for a MS Windows system. The dix part is incorporated in the X server proper, a large body of code that also handles a large number of tasks associated with communications with the OS kernel as well as application-to-application communications (per ICCCM), resource management, and security.

The X Window System, even though is it can be considered as an OS sub-system much as is the case with the Windows graphics sub-system, is actually an application in UNIX terms, and as such does not - and should not - reside in the kernel, in whole or in part. X was designed for UNIX, and UNIX design dictates that X be an application. The result is that X can be easily moved from one UNIX implementation to another. That is unless portions of X are pulled into the kernel, in which case X would become vulnerable to kernel changes that would otherwise be transparent to X. For Linux, this would be particularly bad, since the Linux kernel changes almost month to month, sometimes substantially.

The graphics "driver" (the previously-defined ddx code) is written to make use of the specific features/capabilities of a particular graphics chip, and to make use of the general features/capabilities of the X server. If the X server cannot support certain features/capabilities needed for an application, the "driver" code for the chip will generally not be able to properly implement those features/capabilities. One example would be the need to drive a monitor that has been physically rotated 90 degrees to portrait mode. Another example is the ability to drive two monitors from the same graphics chip but using different graphical contexts (different resolutions, color depth, etc.) for each xscreen, and have both xscreens fully accelerated. In other words, the X server architecture must be properly designed to support the overall goals of particular X applications. Then, the "driver" must efficiently harness the hardware capabilities in the specific graphics hardware in a manner that is compatible with the X server.

When a graphics manufacturer develops a Linux/UNIX driver for a chip, the driver must be designed to "mate up" with the X server that is to be used. Unfortunately, if the X server is to be an XFree86 or X.org X server, a number of difficulties are encountered. There are many such X servers, and one driver will not work with the majority of them. Once a driver is working with one or more of these X servers, the manufacturer finds that the X servers are constantly changing. Thus the previously working drivers are often "broken" with a new X server release. Then, the Linux Group makes a kernel change, and both the X servers and drivers break. And so on.

An X server that is properly architected and implemented for the X applications targeted, when coupled with ddx's that make maximum efficient use of the graphics hardware being used, should be a package that is very stable and easy to maintain. It should also be easy to move to different UNIX kernels and to different computer platforms and exhibit the same stability and ease of maintenance. If this is the criteria of a well-designed X Window Subsystem, then the freeware-based Linux X servers from XFree86/X.org in combination with the "drivers" from XFree86/X.org and/or the graphics manufacturers fail miserably.

The failure of the "open source community" vehicle to provide quality X graphics software suitable for use in systems that require stability, good performance, and economical maintainability - often over long product life cycles - is not difficult to document. Just download the Linux graphics driver software packages from any of the major graphics chip manufacturers and examine them. Or visit the Linux user forums at those graphics manufacturers Web sites, although some have been discontinued because of the embarrassingly negative comments from users. Also one can delve into the Linux v2.6 kernel history. When even RedHat has moved to isolate itself from the machinations of the Linux kernel group, one suspects things are a bit less than optimal. Clearly the Linux 2.6 kernel has been a "moving target" for quite some time, and the prospects for a stable version seem dim. Perhaps the aversion to specifications evidenced by Linus in this

exchange <http://kerneltrap.org/node/5725> explains some of the fits and starts that have been going on with this kernel. And perhaps it indicates that things will probably not improve.

The constantly changing Linux kernel coupled with the move by the XFree86/X.org "community" over the last few years to move parts of X out of the X server and into the kernel - "Direct Rendering Infrastructure" is just one (but big) example - dooms the Open Source "Linux graphics drivers" activity to one that has no hope of ever providing stable, high-quality, and easily maintained graphics software to the UNIX world. It hasn't happened in the past ten years, and it won't happen in the next ten. Apparently some in the open source X community have begun to grasp this reality, since there have been calls for X to be "dumped," to be replaced with an entirely new approach to graphics on UNIX. After making a hash out of X by ignoring the precepts of UNIX, and the specs of X, now they want to junk X. Cute.

While all of this perversion of UNIX and screwing up X has been going on in the Open Source Community, Xi Graphics has been providing its famed Accelerated-X brand of X servers and graphics drivers (for graphics chips made by ATI, Matrox, 3Dlabs, Intel and others) that delivers high performance, stable operation, and easy maintenance (provided free to Xi Graphics' customers). Additionally, the same X servers and drivers operate on Solaris, HP/UX, and Linux kernels, and on SPARC, PA RISC, PPC, and Intel x86 platforms, both 32- and 64-bit. Moving Accelerated-X to a new kernel or to a new computer platform is basically routine. Moving XFree86/X.org and Nvidia to SPARC must be a bit of a challenge, since it still is not available as this is written. Xi Graphics doesn't do Nvidia, but we do have ATI RADEON 9250 OpenGL v1.5 on SPARC. And ATI FireMV 2400 PCI and PCIe on SPARC. And Matrox cards, too.

Accelerated-X adheres to the fundamental UNIX principle that the X Window System is really an application - albeit an important one from a graphical systems point of view - and therefore it should be as far removed from the kernel as possible. This insures that X cannot mess up UNIX, but it also it insures that moving X from kernel to kernel is a relatively simple task. At least compared to what the XFree86/X.org developers or the graphics chip manufacturers have to contend with if they want to operate and maintain their code on various kernels and platforms.

It is true that Xi Graphics software is not free. But in reality, neither is the freeware open source Linux graphics driver software free. Not if one has to account for the time spent getting it to work, the poor performance, and the downtime because of lack of support available for maintaining it. Many ex-freeware Linux graphics driver users in commercial/industrial environments have commented on the high cost of such "free" software.

Wm E Davis, Nov 2006