ITPro Today™

COMPUTE ENGINES

What the Windows Source Code Leak Means to You

Paul Thurrott | Feb 16, 2004

Last Thursday morning, I received an excited Instant Messaging (IM) alert from a friend at Microsoft: "Have you seen this?" he asked. He then sent me a file named "winver.c," reportedly part of the Windows 2000 source code. The source code for Win2K, as well as for Windows NT, he said, had leaked to the Internet. The file I was looking at was a source code listing for a short program written in the C language; it was described as the "Windows version program" and was written in March 1989 by someone identified as toddla. Several other C source code listings were leaked, including one purportedly written by NT architect David Cutler.

The notion that Microsoft's crown jewels might leak publicly wasn't surprising to me; after all, the company had opened its source code to an increasingly large portion of the public in recent years through its Shared Source program, a response to the open-source threat of Linux. Since first announcing the Shared Source program, Microsoft has regularly extended the program's reach, and now many governments, corporations, educational institutions, hardware and software development partners, and even individuals have signed nondisclosure agreements (NDAs), giving them limited-rights access to the source code for various Windows versions. The software giant has even publicly acknowledged that it was considering opening the source code to Microsoft Office also.

Microsoft disseminates its valuable source code to other institutions and individuals for various reasons. Historically, the company's hardware and software partners have received

access to the source code to ensure that the products they develop work seamlessly with Microsoft systems. Under the Shared Source program, the reasons are a bit more varied. But one reason Microsoft has opened up its source code is to fend off competition from Linux and other open-source solutions, which provide users with modifiable source code. Microsoft doesn't let its Shared Source partners change the Windows source code and potentially make their own modified versions of Windows. Instead, the source code access provides suspicious governments with the evidence they need to prove that Microsoft isn't inserting back doors, especially US governmental back doors, into its software. And Microsoft has shown itself to be more, ahem, open to the notion of providing governments with specially tailored Windows versions when needed, as the company did recently with Thailand, although those will be developed inhouse, as the need arises.

But here's what we know so far about the leak. Contrary to early reports, only a small portion of the source code for Win2K Service Pack 1 (SP1) and NT 4.0 leaked. Experts differ about how much code leaked--I've seen estimates in the 1 to 15 percent range--but using the code to build a working version of Windows would be impossible. I did obtain the leaked Win2K source code so that I could analyze it and confirm it was real, but I've never seen the NT source code. I'll be destroying my copy of the source code after completing my analysis and have no intention of publishing major portions of it, of course.

At this time, a software company called Mainsoft is the most likely source of the leak, which means the leak had no ties to the Shared Source program. Mainsoft has had Windows source code access for years; longtime Windows & .NET Magazine UPDATE readers might recall my August 2000 revelation that Microsoft had hired the company to explore Linux ports of Office and Microsoft Internet Explorer (IE), for example--but uses the information for integration software development purposes.

The leaked source code I've seen includes code for the Windows Explorer shell, among other things, and an interesting wealth of documentation that shows Microsoft's developers how to move pre-IE 4.0 Windows shell code to the then-new IE integrated shell. The code occupies about 147MB of space and includes about 12,900 files, mostly C, C++, and assembly source files, as well as C and C++ header files. And for you conspiracy theorists, sorry, the code doesn't appear to include any proof that Microsoft stole source code from UNIX, Linux, or

other sources in a bid to make its systems better. Open-source enthusiasts probably spent the weekend poring over the code just to find such evidence.

On a technical note, the source code is clean and well coded but is often devoid of useful comments. It's also quite frank in some places, with occasional swearing and name calling, usually aimed at Microsoft's own products. But what really stands out, is how often Microsoft must insert a minor coding change to accommodate the idiosyncrasies of one application. These hacks, as they're called in the code, are often aimed at third-party applications, letting the applications work after a bug or previous feature they've relied on has been eliminated. This is a good example of Microsoft going out of its way to ensure that its partners products work with Windows, a task the company has never received a lot of credit for.

When the source code leak was first reported, security experts opined that it would have damaging effects on Microsoft's credibility and could lead to a new generation of software exploits that take advantage of hackers' newfound knowledge of the Windows source code. However, little networking or security code is included in the leaked source I've examined, and because the code comprises such a small portion of the entire source-code base, it will be impossible to figure out the complex interworkings of code that make up the complete OS and find some systemic flaw. So from a technical standpoint, I think that, for now at least, the Windows source code leak shouldn't affect any rollout decisions, though arguably you'd be better off going with Windows Server 2003 and Windows XP over Win2K right now for various unrelated reasons anyway.

Indeed, with Microsoft's recent emphasis on upgrading to Windows 2003 and XP for security reasons, there's been some question about Microsoft's plans to adequately support Win2K going forward. For example, although both of these newer systems will get the improved Windows Firewall in service pack updates later this year, and Windows 2003 will get the roles-based Security Configuration Wizard, Microsoft hasn't said much about offering such improvements to Win2K users. For whatever it's worth, I do know that the company intends to soon reveal various Win2K security improvements that it will roll out this year, but I'm a little worried about its public silence thus far. I'm further concerned what the Win2K source code leak will do to put these plans on the back burner. It would be sad to see Microsoft take

advantage of this episode to formalize its desire to deemphasize Win2K, years before the company should do so.