Sophos Anti-Virus and the millennium

Paul Ducklin

Part # tr00047n / 970908

Introduction

As the manufacturers of the anti-virus software product SWEEP, we regularly receive requests asking whether we are "Year 2000 Aware" or "Millennium Compliant". This document aims to answer three questions:

- What does Sophos understand by "Millennium Compliance", and what are the potential problems faced by software in the year 2000?
- What significance do dates have for SWEEP, and how does the product address potential date problems in the year 2000?
- What can users of Sophos' products do to verify for themselves that SWEEP is indeed "Millennium Compliant"?

What is 'millennium compliance'?

We are currently in the 1997th year of the Christian Era, which is used and understood as the basis for dates in most countries of the world. Because 1997 is a lengthy number to say aloud, and because 100 years is a long time by human standards, we have become accustomed to referring to the last two digits of the year. We rely on context to disambiguate statements such as "I was born in '63", or "my passport expires in '06".

This tendency to use only two digits for dates has extended to computer programming, especially in legacy applications whose implementors never imagined that their code might still be in use in the next century.

In the example above, a human observer would have no trouble computing the number of years between the two dates as the value 2006 minus 1963, or 43 years. A simplistic computer program, however, might subtract 63 from 6, producing -57. Clearly, a program of this type does not recognise 1 January 2000 as the day following 31 December 1999, and will produce incorrect results in date calculations which cross this boundary.

There are other date concerns relating to the year 2000, even for programs which correctly place 1 January 2000 after 31 December 1999. These revolve around the fact that all years evenly divisible by four are leap years **except centuries**. Centuries are leap years only if they are evenly divisible by 400. 1700, 1800 and 1900 were not leap years; 2000 will be.

Some software is said to report 2000 as a non-leap year, and any program which does this will skip over 29 February 2000. This implies that all date calculations performed across this "missing day" will be wrong by one day.

For an application to be "Millennium Compliant", we believe that its date-based functionality should be able to perform consistently for dates prior to, during and after the year 2000, and that it should be able to recognise 2000 as a leap-year.

Note that the correct behaviour of a program successfully designed by its manufacturer to be "Millennium Compliant" may depend on other software in the system operating correctly too.

What significance do dates have for SWEEP?

Fortunately, SWEEP's central function is the detection and disinfection of computer viruses, so dates and date calculations are relatively unimportant to the product. Even if SWEEP is used with system software that is not Millennium Compliant, and which presents incorrect date information to SWEEP, the ability of SWEEP to detect and disinfect viruses **will not be compromised**.

SWEEP uses date-related calculations in only two places. The first is performed when SWEEP starts up, and involves a comparison between the current date (as reported by the system software) and SWEEP's compilation date (as recorded in the program file at build time). SWEEP's internally-stored compilation date is stored in the form YYYYMMDD, using four characters for the year.

Obviously, the system date should always be equal to or later SWEEP's own compilation date. If it is not, SWEEP reports this with the message "System clock date is incorrect". SWEEP will also report "Useful life of SWEEP has been exceeded" if the system date is more than four months after its compilation date. This helps warn users that the program should be replaced with a more recent version that detects more viruses.

The second date calculation performed by SWEEP is carried out when external virus identity files (known as IDE files) are loaded. IDE files are used in emergencies to add detection and disinfection of specific new viruses to SWEEP. The files consist entirely of printable ASCII hexadecimal characters, and can thus easily be downloaded from the Sophos BBS, FTP server or Web site, and can even be received by fax and typed in locally.

Virus identities provided in IDE files are typically built in to the next release of SWEEP. This means that any IDE file older than three months is almost certainly redundant, and SWEEP reports this with a message such as "File C:\SWEEP\NEWVIRUS.IDE is older than 90 days".

This calculation is performed by comparing the date stamp on each IDE file in use against the current system date. Because SWEEP's internal date calculations are Millennium Compliant, SWEEP will produce correct results if it is presented with correct input by the operating system.

How can SWEEP's millennium compliance be verified?

Testing SWEEP's handling of dates is easy. Install the current version of SWEEP onto a test machine (with SWEEP for DOS, the program SWEEP.EXE can be run directly from the distribution diskette) and run SWEEP. Confirm that no date-related messages appear.

Set the system date back by one year, and run SWEEP again. The software should report "System clock date is incorrect". Set the system date to one year into the future and run SWEEP again. It should now report "Useful life of SWEEP has been exceeded". The same thing should happen if the date is set to 31 December 1999, and should continue to happen if the date is then advanced by one day to 1 January 2000.

Similar tests can be done with an IDE file (IDE files can be downloaded from the Web — see http://www.sophos.com). Place the IDE file into the same directory as SWEEP.EXE, and then repeatedly run SWEEP while advancing the system date up to and past the year 2000. SWEEP should continue to report "File ... is older than 90 days" as the date crosses the "Millennium Boundary" into the next century.