BugTraq

Back to list | Post reply

XXE (Xml eXternal Entity) attack Oct 29 2002 11:23PM
Gregory Steuck (greg-xxe nest cx) (1 replies)

-----BEGIN PGP SIGNED MESSAGE-----
Hash: SHA1

Gregory Steuck security advisory #1, 2002

Overview:
XXE (Xml eXternal Entity) attack is an attack on an application that parses
XML input from untrusted sources using incorrectly configured XML parser.
The application may be coerced to open arbitrary files and/or TCP connections.

Legal Notice:

Disclaimer:
The information in this advisory is believed to be true though
it may be false.
The opinions expressed in this advisory and program are my own and
not of any company. The usual standard disclaimer applies,
especially the fact that Gregory Steuck is not liable for any damages
caused by direct or indirect use of the information or functionality
provided by this advisory or program. Gregory Steuck bears no
responsibility for content or misuse of this advisory or program or
any derivatives thereof.
Anything in this document may change without notice.

Details:
External entity references allow embedding data outside the main file into
an XML document. In the DTD, one declares the external reference with the
following syntax:
<!ENTITY name SYSTEM "URI">

XML processor behavior as specified is
[http://www.w3.org/TR/REC-xml#include-if-valid]:

"When an XML processor recognizes a reference to a parsed entity, in
order to validate the document, the processor must include its
replacement text. If the entity is external, and the processor is not
attempting to validate the XML document, the processor may, but need
not, include the entity's replacement text..."

Now assume that the XML processor parses data originating from a source under
attacker control. Most of the time the processor will not be validating,
but it MAY include the replacement text thus initiating an unexpected
file open operation, or HTTP transfer, or whatever system ids the XML
processor knows how to access.

Suspect systems:
The buzz on the street is "web services". They accept XML encoded
data over the network, sometimes from untrusted clients. So, the
prime targets are SOAP and XMLRPC implementations. Yet, there are
many more XML based protocols and vulnerability does not necessary
lie with the servers. Pick any "XML based network protocol" and
try to apply the attack methodology.

Suggested fix:
Most XML parsers allow their user to explicitly specify external
entity handler. In case of untrusted XML input it is best to prohibit
all external general entities.

Successful exploitation may yield:
* DoS on the parsing system by making it open, e.g.
file:///dev/random | file:///dev/urandom | file://c:/con/con
* TCP scans using HTTP external entities (including behind firewalls
since application servers often have world view different
from that of the attacker)
* Unauthorized access to data stored as XML files on the parsing
system file system (of course the attacker still needs a way to
get these data back)
* DoS on other systems (if parsing system is allowed to establish
TCP connections to other systems)
* NTLM authentication material theft by initiating UNC file access to
systems under attacker control (far fetched?)
* Doomsday scenario: A widely deployed and highly connected application
vulnerable to this attack may be used for DDoS.

Products review:
Several SOAP and XMLRPC implementation were found vulnerable. I will
be contacting their respective authors directly. It will be up to
those authors to publish the patches and/or advisories.

The following implementations were found NOT vulnerable and the reasons

contributing to their resistance were researched.

Java:
Apache XML-RPC server is NOT vulnerable in the default configuration
due to its use of MinML parser which doesn't support external entities.
Yet should be vulnerable if used with a full blown parser like Xerces
or Crimson. To make it invulnerable in all configurations it needs to
explicitly setup an EntityResolver that aborts having found external
entities.

Marquée XML-RPC also uses MinML and thus is NOT vulnerable.

XMLRPC-J uses freeDOM that only supports Minimal XML which
lacks entity references (http://www.docuverse.com/smldev/minxml.jsp)

WebLogic 6.1sp3 SOAP implementation was NOT found vulnerable. It
appears to be using a parser that ignores entities altogether. Ignorance
is bliss...

Python:
Python 2.2 SimpleXMLRPCServer does NOT seem to be vulnerable. It can use
multiple different parsers:
* xmllib.XMLParser is the default one shipped with Python. It
doesn't implement processing of doctype definition and thus doesn't
understand external entities defined in there
* ExpatParser is used when expat python-expat is installed,
it understands the references but seems to replace them with
empty strings unconditionally. This negates the attack.
* SGMLOP parser, judging by comments in its source doesn't recognize
external entities
* FastParser was not available for inspection

Acknowledgments:
Even though the issue was discovered and researched independently I
cannot claim to be the first one to realize the risks associated with
XML external entities. E.g. RFC 2518 discusses the issue in section
17.7 Implications of XML External Entities.

-----BEGIN PGP SIGNATURE-----
Version: GnuPG v1.0.6 (OpenBSD)
Comment: Processed by Mailcrypt 3.5.6 and Gnu Privacy Guard <http://www.gnupg.org/>

iEYEARECAAYFAj2/FZkACgkQCxVCvY31obB6vQCbBlV+v0jDRQQ7GcNxYRtajtAf
FxUAnRCDfjLy2692iGF3Ewmxzo/VXYmz
=t4QF
-----END PGP SIGNATURE-----

[reply]

Re: XXE (Xml eXternal Entity) attack Oct 30 2002 09:15AM
Miles Sabin (miles milessabin com)