INFORME DE SEGURIDAD 2023.12

LibreOffice: Calidad y seguridad del software

LibreOffice nació como una bifurcación de OpenOffice.org, la suite ofimática libre desarrollada por Sun Microsystems. El proyecto OpenOffice.org nació en 2000, un año después de la adquisición de StarDivision (el hogar aleman de StarOffice) por Sun Microsystems, cuando la propia Sun decidió transformar la suite privativa en una de código abierto bajo licencia LGPL.

En 2010, los líderes de la comunidad voluntaria del proyecto OpenOffice.org, preocupados por la gestión de Sun -basada en metodologías de desarrollo obsoletas y un proceso de Control de Calidad excesivamente manual- y por la adquisición de Sun Microsystems por parte de Oracle (Oracle nunca ocultó su idiosincrasia por el software de código abierto) decidieron lanzar un proyecto independiente: LibreOffice.

Sin embargo, la calidad y la seguridad de OpenOffice.org ya eran entonces superiores a las de cualquier software privativo, y en particular a las de Microsoft Office. La base de datos CVE (Common Vulnerabilities and Exposures) informa de un número de problemas un orden de magnitud superior, debido a dos factores: la mayor fragilidad del código fuente privativo, que no se beneficia de los efectos virtuosos del intercambio de conocimientos sobre seguridad, y el mayor número de usuarios, que lo convierte en un objetivo más fácil.

La mayor calidad del software de código abierto fue confirmada por el Informe sobre Código Abierto de Coverity Scan: "En 2013, la calidad de los proyectos de código abierto superó a la de los proyectos privativos en todos los niveles. Para el informe de 2013, analizamos unos 500 millones de líneas de código de unos 500 proyectos privativos escritos en C/C++ y descubrimos que el software de código abierto tiene una menor densidad de defectos que el privativo". Uno de los factores que condujeron a este resultado es el esfuerzo realizado por algunos grandes proyectos -entre ellos LibreOffice- para resolver colectivamente más de 11.000 defectos a lo largo del año"[1].

La calidad del código fuente de LibreOffice

Cuando nació el proyecto LibreOffice, los desarrolladores cambiaron el enfoque respecto a OpenOffice.org, lanzando una actividad de limpieza del código fuente que duró todo 2011, y que desde principios de 2012 ha dado como resultado una suite ofimática de calidad significativamente superior. Como parte de la actividad de limpieza, los desarrolladores también revisaron su enfoque del control de calidad, estableciendo un proceso automatizado basado en tecnologías de vanguardia.

El proyecto LibreOffice utiliza Gerrit como herramienta de revisión de parches dada su integración con Git, el principal sistema distribuido de gestión del desarrollo de software. El código fuente se compila periódicamente mediante una batería de varios Tinderboxes y, si la compilación se realiza correctamente, se somete a una serie de pruebas automatizadas que verifican el comportamiento del software con miles de documentos.

Los archivos de prueba se obtienen de varias instancias públicas de Bugzilla: The Document Foundation, Launchpad (algunos), Freedesktop, Mozilla, GNOME, KDE, Gentoo, Mandriva, Novell, AbiSource y archivos de prueba SVG del W3C. Los documentos más adecuados para las pruebas son los malos, por lo que cargamos y guardamos todo lo que se adjunta a los errores.

Esta actividad automatizada se complementa con el trabajo del equipo de Control de Calidad de LibreOffice, que utiliza herramientas como Bugzilla para gestionar tanto los errores como las regresiones, e informar de ellos a los desarrolladores cuando proceda para que corrijan el código fuente.

INFORME DE SEGURIDAD 2023.12

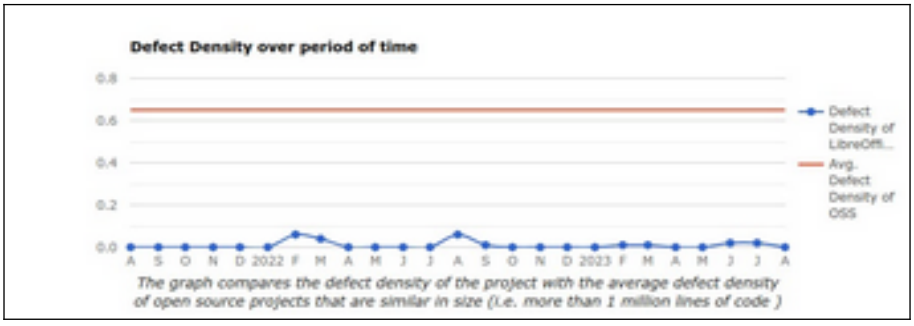
Gestión de defectos en el código fuente de LibreOffice

La calidad del código fuente ha mejorado significativamente desde que los desarrolladores empezaron a utilizar los servicios de Coverity Scan en 2012 [2]. Desde entonces, LibreOffice se ha convertido en uno de los paquetes de software con menos defectos en proporción a las líneas de código fuente. Esta actividad es muy importante en términos de seguridad del software, ya que los defectos en el código fuente se asocian a menudo con los informes CVE (vulnerabilidades y exposiciones comunes).

The first image provided by Coverity Scan represents the current situation of LibreOffice 7.6 Community's source code, with 0 outstanding defects. Over time, LibreOffice developers have fixed 27,929 defects, while 571 defects were dismissed as false positives.

La primera imagen proporcionada por Coverity Scan representa la situación actual del código fuente de LibreOffice 7.6 Community, con 0 defectos pendientes. Con el tiempo, los desarrolladores de LibreOffice han corregido 27.929 defectos, mientras que 571 defectos fueron descartados como falsos positivos.

La segunda imagen proporcionada por Coverity Scan resume la tendencia de los defectos pendientes frente a los fijos durante los últimos 10 años. Desde 2015, el número de defectos pendientes ha sido constantemente 0 o cercano a 0, mientras que el número de defectos fijos ha ido aumentando regularmente (la brecha en 2019 se debe a una revisión completa del software de análisis Coverity Scan).



Esta tercera imagen proporcionada por Coverity Scan ofrece una mejor visualización de la tendencia de la densidad de

defectos en los últimos dos años, de agosto de 2021 a agosto de 2023. Sólo en febrero/marzo y agosto de 2022 la densidad de defectos de LibreOffice fue superior a 0,005 defectos por cada 1.000 líneas de código.

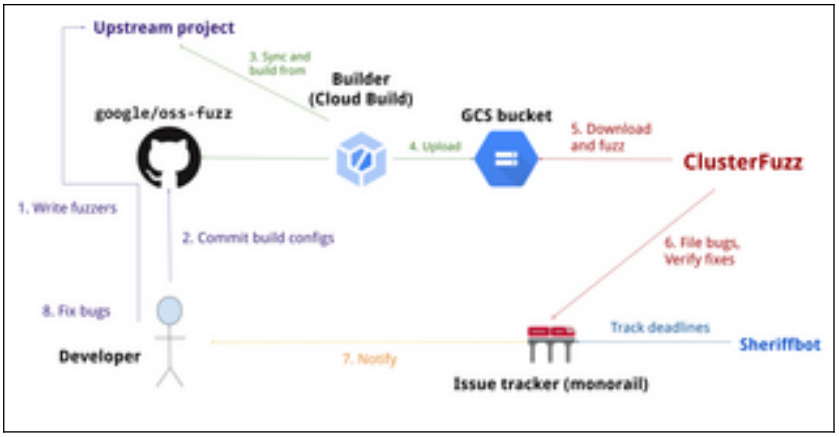
Las cifras proporcionadas por Coverity Scan son un testimonio de la limpieza y refactorización del código fuente de LibreOffice por parte de los desarrolladores desde 2010. También confirman el alcance de la deuda técnica heredada de OpenOffice.org, que se resolvió por completo en cuatro años. Fue un trabajo valioso, comprendido por el mercado sólo a posteriori, cuando por fin quedó claro que la estrategia de desarrollo de LibreOffice es la adecuada.

Coverity Scan permite detectar fallos, vulnerabilidades de seguridad, concurrencias, corrupción de memoria, memoria no inicializada, gestión de errores y fugas de recursos, y contribuye a la seguridad general del software, ya que la reducción del número de defectos en el código fuente proporciona a los desarrolladores una base más sólida y resiliente para su trabajo.

Cabe destacar, sin embargo, que un número bajo de defectos en el código fuente -como el de LibreOffice- no excluye necesariamente la presencia de bugs, regresiones y vulnerabilidades.

Control del código fuente de LibreOffice mediante fuzzing

Fuzzing es una técnica de prueba de software automatizada, ampliamente utilizada por los desarrolladores de LibreOffice, que proporciona datos no válidos, inesperados o aleatorios como entradas a un programa informático. La aplicación es entonces monitorizada por el sanitizador de código, una herramienta de programación que detecta bugs en forma de comportamientos indefinidos o sospechosos, para problemas como cuelgues, fallos en las aserciones de código integradas, o potenciales fugas de memoria.

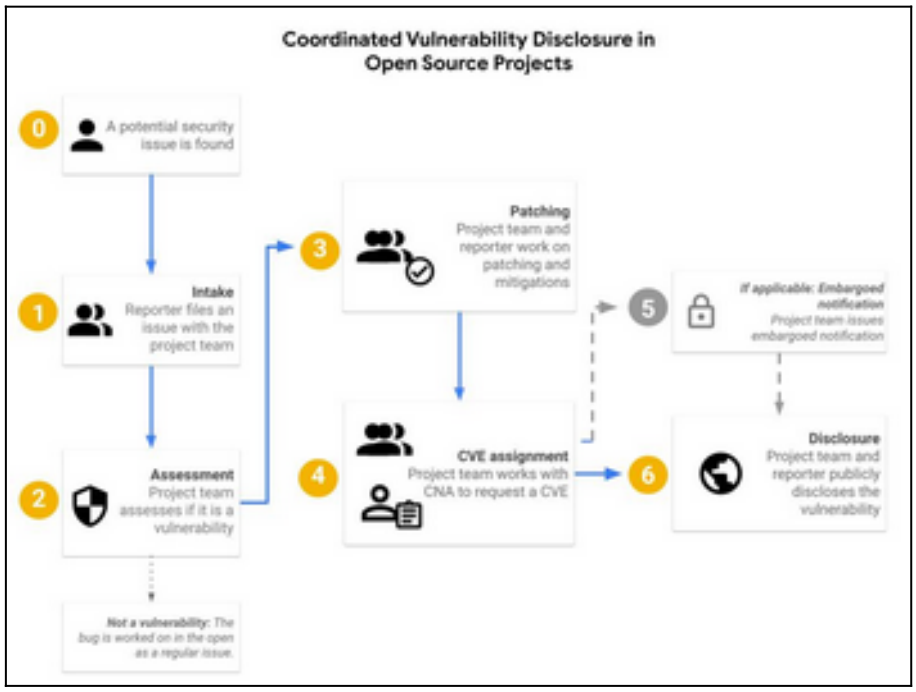


El fuzzing se utiliza como técnica automatizada para exponer las vulnerabilidades de los programas críticos para la seguridad que podrían explotarse con fines maliciosos, para demostrar la presencia

de bugs en lugar de su ausencia.

La principal herramienta de fuzzing adoptada por los desarrolladores de LibreOffice es Google OSS Fuzz, anunciada en 2016. Se trata de una infraestructura de pruebas utilizada para Chrome y otros proyectos de software libre y de código abierto (FLOSS), que combina motores de fuzzing con sanitizadores y proporciona un entorno de ejecución distribuido masivo impulsado por ClusterFuzz. Gracias a OSS-Fuzz, el proyecto está comprobando 50 formatos de archivo y esa comprobación se ejecuta constantemente a medida que se integran nuevos cambios.

Gestión de vulnerabilidades en el código fuente de LibreOffice



El sistema Common Vulnerabilities and Exposures (CVE) proporciona un método de referencia para la información públicamente conocida en torno a las vulnerabilidades y riesgos de seguridad. El FFRDC de Ciberseguridad Nacional de Estados Unidos, operado por la corporación MITRE, mantiene el sistema, con financiación de la División de Ciberseguridad Nacional del Departamento de Seguridad Nacional de Estados Unidos.

Una vulnerabilidad es un punto débil de un sistema informático que permite un acceso injustificado. El identificador CVE es el número único asignado a cada vulnerabilidad por una Autoridad Numeradora CVE (CNA), como The Document Foundation en el caso de LibreOffice. Cuando se investiga una vulnerabilidad o una vulnerabilidad potencial, es útil adquirir un número CVE desde el principio, ya que toda la correspondencia futura puede referirse a él.

Según la base de datos CVE de MITRE Corporation, que puede consultarse en https://www.cve.org/ (antes en https://cve.mitre.org/), LibreOffice se ha visto afectado por 50 CVE en los últimos 10 años. En el mismo periodo, Microsoft Office se vio afectado por 505 CVE, es decir, un orden de magnitud superior al de LibreOffice.

Además, todas las CVE que afectaban a LibreOffice se resolvieron con un parche publicado antes de su divulgación (por convención, la publicación de las CVE en la base de datos se produce entre 30 y 60 días después de que se informe del problema al equipo de seguridad de las aplicaciones afectadas).

El sitio web https://www.cvedetails.com/ ofrece una comparación basada en el Common Vulnerability Scoring System (CVSS), un conjunto abierto de normas utilizadas para evaluar una

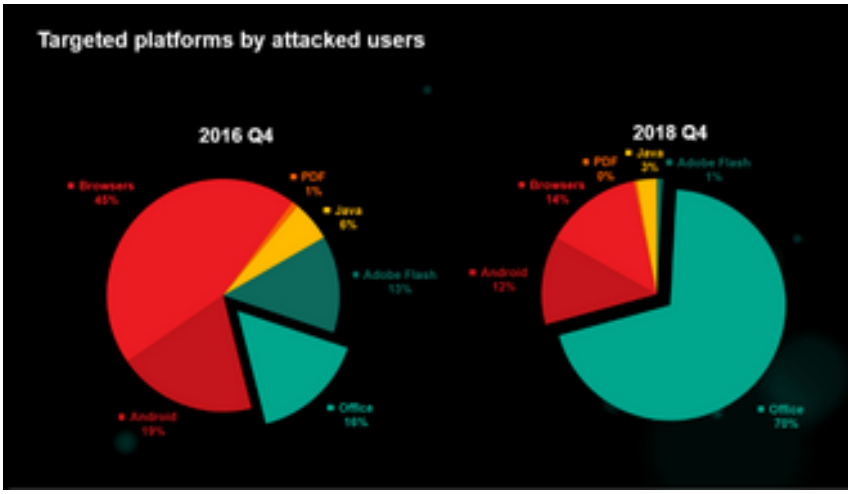
vulnerabilidad y asignarle una gravedad en una escala del 0 al 10 (de ninguna a crítica).

El ESC (Engineering Steering Committee) de LibreOffice cuenta con el apoyo de un equipo de expertos en temas de seguridad más específicos, con especialistas de talla mundial, que a menudo se ofrecen como expertos en seguridad para empresas que desarrollan software en otros sectores (un ejemplo típico es el software para automoción).

Dicho todo esto, hay que dejar claro que el número de vulnerabilidades no puede utilizarse para definir la ventaja competitiva de un programa, y de hecho nunca lo hemos mencionado hasta la fecha, porque no define la calidad de la aplicación en sí, sino que sólo representa el riesgo teórico de que se utilice para acceder a los datos del usuario o comprometer la seguridad del PC. Así pues, los usuarios nunca deberían elegir un software en lugar de otro basándose en el número de vulnerabilidades.

La importancia del formato de archivo ODF nativo de LibreOffice

LibreOffice utiliza el estándar abierto Open Document Format (ODF) como formato nativo, lo que puede ayudar a organizaciones y empresas a reducir su vulnerabilidad ante ataques desde el exterior, en comparación con lo que puede ocurrir con los formatos de archivo privativos.



Los formatos privativos de documentos ofimáticos son una de las vulnerabilidades más explotadas, según una investigación independiente realizada por Symantec en 2011 y Kaspersky Labs en 2018. En la Cumbre de Analistas de Seguridad 2019, Kaspersky dijo que alrededor del 70 por ciento de todos los ataques detectados en el cuarto trimestre de 2018 estaban tratando de abusar de una vulnerabilidad de Microsoft Office, un aumento dramático frente al 16 por ciento detectado en 2016 (la diapositiva es de la presentación original) [3].

La explicación es sencilla. Los formatos de archivo privativos como los DOC, XLS y PPT heredados, y los actuales DOCX, XLSX y PPTX "de transición", pueden contener blobs binarios de datos -que son el vehículo preferido del malware- para permitir la retrocompatibilidad con documentos antiguos, una característica que pretendía proteger a los usuarios de la obsolescencia de los contenidos.

La retrocompatibilidad, y sus correspondientes blobs binarios, no sólo reducen la seguridad de los documentos de Microsoft Office, sino que también impiden que cumplan los estándares. De hecho, aunque las especificaciones "estrictas" de Office Open XML no prevén la integración de blobs binarios al no poder ser representados visualmente por el código XML, sí están permitidos por el actual formato de archivo "de transición", lo que añade más complicación y riesgo.

Por el contrario, la introducción de ODF creó una ruptura en la compatibilidad hacia atrás de

los documentos, que se solucionó con herramientas de software para la conversión de formatos. De este modo, el formato siempre se ha ceñido a la descripción de la norma basada en XML, y nunca ha requerido la integración de blobs binarios.

Por supuesto, el uso del formato estándar ODF no puede garantizar la seguridad del software, aunque puede simplificar la tarea de las herramientas que deben comprobar la existencia de código malicioso. La protección de los usuarios y sus interlocutores se deja en manos de las medidas de seguridad y los programas antivirus adoptados por el individuo o la organización.

En el caso de LibreOffice, el formato estándar Open Document Format es un elemento importante que complementa el trabajo del equipo de expertos en seguridad reduciendo la superficie de ataque. La seguridad de LibreOffice es el resultado de un esfuerzo global de toda la comunidad, desde las empresas del ecosistema hasta los voluntarios que contribuyen al desarrollo, el control de calidad, la documentación y la localización.

Creditos

La seguridad de LibreOffice es el resultado de una enorme cantidad de trabajo realizado por un grupo de personas liderado por Caolán McNamara, que colectivamente se suma a un apoyo significativo y continuo de LibreOffice por parte de varias empresas:

- Red Hat, que lideró las actividades de seguridad durante años
- Collabora, que ha heredado el liderazgo de Red Hat en seguridad
- allotropia, que apoya los esfuerzos relacionados con la seguridad en ámbitos específicos
- Google, que patrocina OSS Fuzz y proporciona muchas CPU caras
- Synopsys, que ofrece gratuitamente el analizador de código estático Coverity Scan
- Adfinis, que financia el hardware para las pruebas de fallos
- The Document Foundation, que está proporcionando la infraestructura para el desarrollo y especialmente el Control de Calidad, incluyendo profesionales para coordinar las actividades.

Además, la seguridad de LibreOffice también está relacionada con el increíble trabajo de los voluntarios en el desarrollo y el Control de calidad, y con muchas contribuciones de empresas centradas en la seguridad tales como Forcepoint.

Referencias

[1] https://www.zdnet.com/article/coverity-finds-open-source-software-quality-better-than-proprietary-code/

[2] https://scan.coverity.com/projects/libreoffice

https://www.zdnet.com/article/kaspersky-70-percent-of-attacks-now-target-office-vulnerabilities/