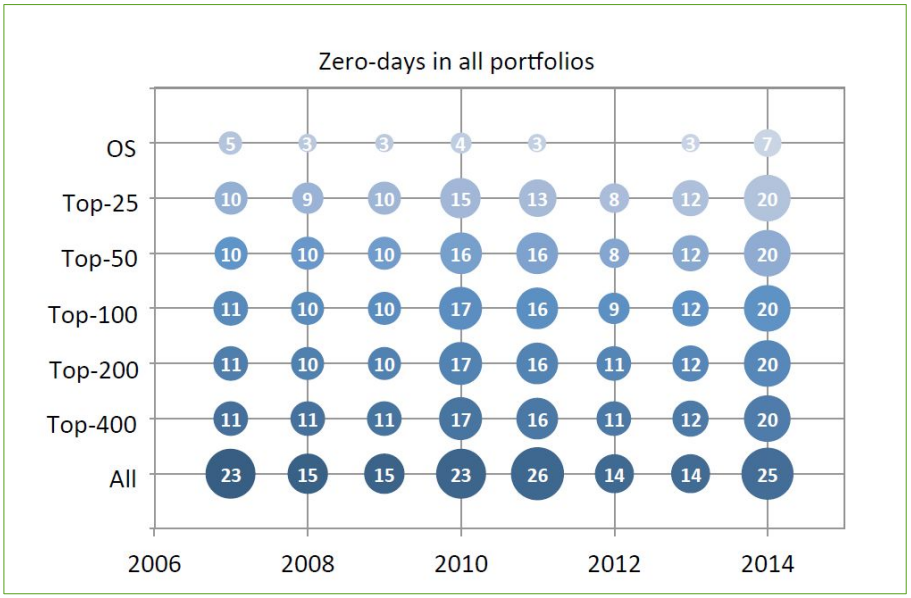
WHITE PAPER

LibreOffice: Qualità e Sicurezza del Software

LibreOffice è un fork di OpenOffice.org, la suite libera di produttività per ufficio sviluppata da Sun Microsystems. Il progetto è nato nel 2000, a un anno di distanza dall'acquisizione di StarDivision - l'azienda tedesca che sviluppava StarOffice - da parte di Sun Microsystems, quando la stessa Sun ha deciso di trasformare la suite proprietaria in suite open source, con licenza LGPL.

Nel 2010, i leader della comunità di volontari del progetto OpenOffice.org, preoccupati sia dalla gestione Sun - basata su metodologie di sviluppo ormai superate e su una eccessiva manualità del processo di Quality Assurance - sia dall'acquisizione di Sun Microsystems da parte di Oracle - un'azienda che non ha mai nascosto la propria idiosincrasia per il software open source - hanno deciso di lanciare un progetto indipendente, annunciando LibreOffice.

Intendiamoci, la qualità e la sicurezza di OpenOffice.org erano già superiori a quelle di qualsiasi software proprietario, e in particolare di Microsoft Office. Il database dei CVE (Common Vulnerabilities and Exposures) riporta un numero di problemi di un ordine di grandezza superiore, dovuto a due fattori: la maggiore fragilità del codice sorgente proprietario, che in quanto "offuscato" è frutto dell'attività di una singola azienda, e non beneficia degli effetti virtuosi della condivisione della conoscenza in materia di sicurezza, e la maggiore diffusione del programma, che lo rende un obiettivo più facile.



La maggiore qualità del codice sorgente del software open source, che viene confermata anche dai risultati comunicati da Coverity Scan (su cui torneremo più avanti), spiega anche la distribuzione degli attacchi Zero Day, concentrati a livello dei primi 25 portafogli software (dove è concentrata la maggior parte del software proprietario). Il grafico, pubblicato all'interno della Vulnerability Review 2015 di Secunia, mostra come la concentrazione degli attacchi Zero Day in questa fascia sia addirittura aumentata negli ultimi anni.

Secondo l'Open Source Report 2013 di Coverity Scan: "Se il software sta cambiando il mondo, come sostiene Marc Andreessen, allora il software open source è all'avanguardia di questo cambiamento. Secondo lo studio su Future of Open Source Software condotto da Black Duck Software and North Bridge Venture Partner, l'adozione e il supporto del software open source nelle aziende non sono mai state così elevate. Su dieci partecipanti alla ricerca, otto hanno scelto il software open source per la sua qualità".

Dopo una lunga ed esauriente serie di analisi, il report conclude: "Nel 2013, la qualità dei progetti open source ha sorpassato quella dei progetti proprietari, a tutti i livelli. Per il report 2013, abbiamo analizzato circa 500 milioni di linee di codice di circa 500 progetti proprietari scritti in C/C++, e abbiamo rilevato che il software open source ha una densità dei difetti inferiore a quella del software proprietario. Uno dei fattori che hanno portato a questo risultato è lo sforzo che hanno fatto alcuni grandi progetti - NetBSD, FreeBSD, LibreOffice e Linux -

per risolvere, collettivamente, più di 11.000 difetti nel corso dell'anno".

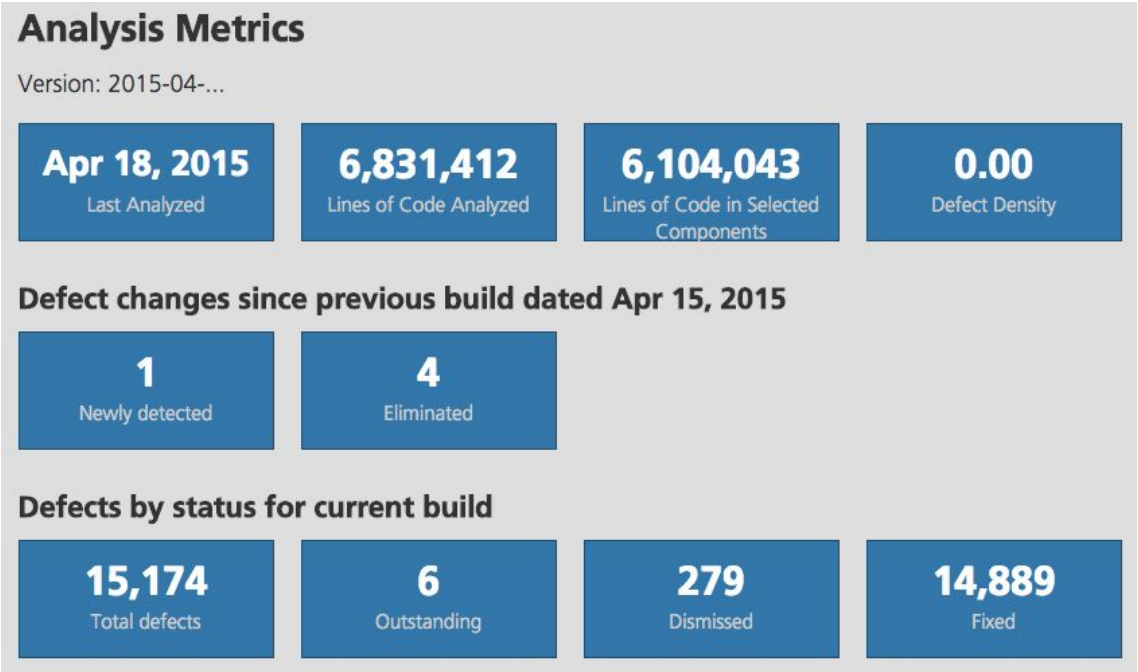
La Qualità del Codice Sorgente di LibreOffice

Quando è nato il progetto LibreOffice, gli sviluppatori hanno modificato le strategie di sviluppo rispetto a OpenOffice.org, lanciando un'attività di pulizia del codice sorgente che è durata per tutto il 2011, e dall'inizio del 2012 ha consentito di avere una suite per ufficio di qualità significativamente migliore. Nell'ambito dell'attività di pulizia, infatti, gli sviluppatori hanno rivisto anche l'approccio alla quality assurance, impostando un processo automatizzato basato su tecnologie allo stato dell'arte.

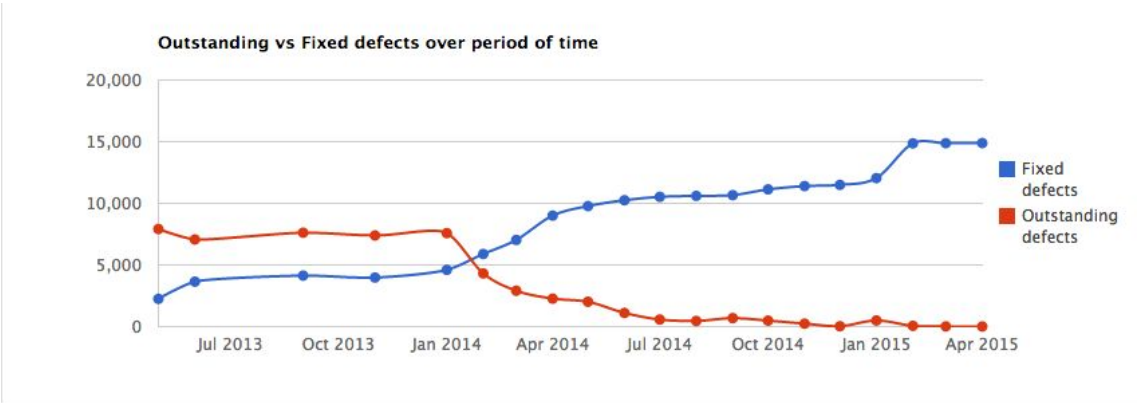
Il progetto LibreOffice utilizza Gerrit come strumento di revisione delle patch per la sua integrazione con Git, il principale sistema distribuito per la gestione dello sviluppo del software. Periodicamente, il codice sorgente viene compilato da una batteria di Tinderbox, e se la compilazione ha successo viene sottoposto a una serie di test automatizzati che verificano il comportamento del software con migliaia di documenti. A tutto questo si aggiunge l'attività del team di quality assurance, che utilizza strumenti come Bugzilla per gestire sia i bug che le regressioni, e per segnalarle agli sviluppatori nel modo più opportuno.

I Difetti del Codice Sorgente di LibreOffice

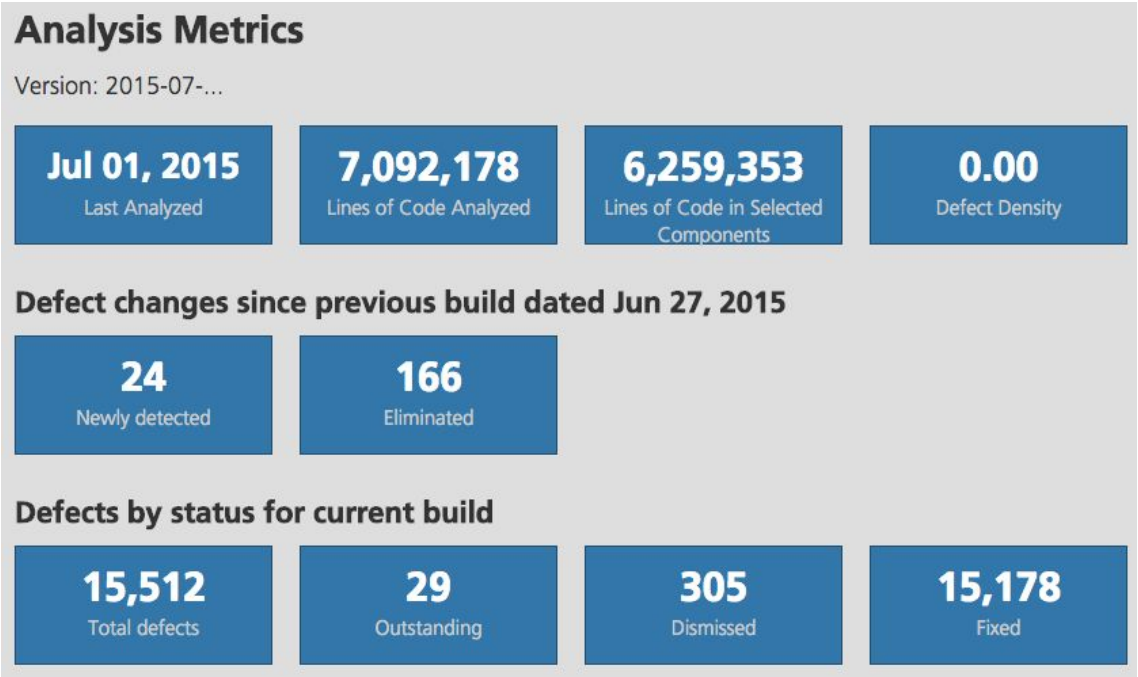
La qualità del codice sorgente è migliorata in modo significativo da quando gli sviluppatori hanno iniziato a utilizzare i servizi di Coverity Scan, nel 2012. In tre anni, LibreOffice è arrivato a essere uno tra i software con il minor numero di difetti in proporzione alle linee di codice sorgente. Questa attività è importante soprattutto in funzione della sicurezza, in quanto i difetti del codice sorgente sono spesso associati alle vulnerabilità e alle esposizioni.



Questa schermata rappresenta la situazione del codice sorgente di LibreOffice 5.0, che verrà annunciato all'inizio di agosto 2015, nel momento in cui il codice sorgente stesso è stato "congelato".



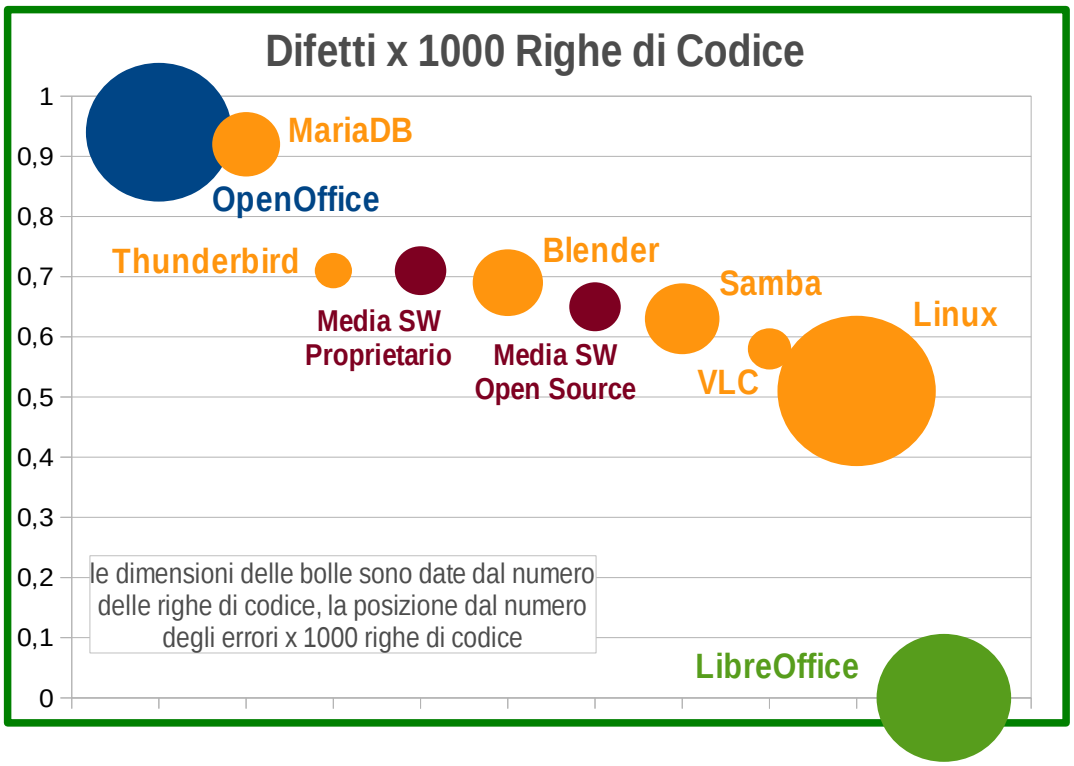
Questa schermata, invece, sintetizza l'andamento dei difetti risolti verso quello dei difetti ancora "aperti", che sono solo 6 su quasi 7 milioni di righe di codice sorgente. Naturalmente, i difetti ancora aperti sono noti (si tratta di software open source), e non rappresentano un problema per la sicurezza.



Questa schermata, infine, rappresenta la situazione al 1° luglio 2015 del codice sorgente di LibreOffice 5.1, che verrà annunciato alla fine di gennaio 2016. Non ci sono sostanziali variazioni rispetto a LibreOffice 5.0, se non che le linee di codice sorgente - per l'aggiunta di alcune funzionalità - adesso superano di poco i 7 milioni, e i difetti ancora aperti sono 29 (ma ci sono ancora più di sei mesi di lavoro prima dell'annuncio di questa versione del programma).

In entrambi i casi, ovvero sia per LibreOffice 5.0 sia per LibreOffice 5.1, la densità dei difetti è scesa sotto lo 0, pari a 1 difetto ogni 100.000 righe di codice, con un miglioramento di due ordini di grandezza rispetto alla situazione di partenza di OpenOffice.org (0,94). Inoltre, il dato è significativamente migliore rispetto alla media dei difetti per 1.000 linee di codice sorgente del software proprietario, che è di 0,71 (fonte: Coverity Scan).

Coverity Scan esegue una scansione settimanale del codice sorgente di LibreOffice, in base alla quale aggiorna la pagina web del progetto e invia una newsletter a una lista di destinatari (tra cui gli sviluppatori di LibreOffice). Ogni scansione settimanale analizza una nuova versione del codice sorgente, ancora in fase di sviluppo.



Il grafico riassume la situazione alla data del 1° luglio 2015, e mette a confronto diversi progetti di software open source con le medie del software open source e del software proprietario.

E' importante ricordare che Coverity Scan offre un contributo importante alla sicurezza del software, perché permette di identificare una percentuale elevata dei difetti che possono dare origine a vulnerabilità ed esposizioni.

La Sicurezza del Codice Sorgente di LibreOffice

Secondo il database delle Common Vulnerabilities and Exposures (CVE), che si trova all'indirizzo http://www.cvedetails.com/ o https://nvd.nist.gov/home.cfm, negli ultimi tre anni LibreOffice è stato colpito da 9 CVE. Nello stesso periodo, Microsoft Office è stato colpito da 101 CVE, ovvero un numero per trimestre superiore a quello che ha colpito LibreOffice in tre anni.

Inoltre, tutti i CVE che hanno colpito LibreOffice sono stati risolti con una patch del codice sorgente, che è stata rilasciata prima della pubblicazione del CVE (per convenzione tra gli addetti ai lavori, la pubblicazione dei CVE avviene

60 giorni dopo la comunicazione della vulnerabilità o dell'esposizione ai team di sviluppo delle applicazioni coinvolte dal problema di sicurezza).

In nessun caso, le vulnerabilità o le esposizioni che hanno colpito il codice sorgente di LibreOffice - e in precedenza quello di OpenOffice.org - non hanno dato origine a un attacco Zero Day. In un solo caso, all'epoca di OpenOffice.org, la patch è stata applicata dopo la pubblicazione del CVE, ma si trattava di un problema minore.

Questi risultati sono il frutto complessivo di tutte le attività descritte fino a questo punto, e del fatto che la qualità del codice sorgente open source non è più una sorpresa (indipendentemente dagli incidenti, che possono sempre succedere, tanto che la qualità è sempre una preoccupazione e non è mai uno strumento di vantaggio competitivo).

Punteggio CVE	**LibreOffice**	**Microsoft Office**
10	1	1
9,3		74
9		1
7,5	4	
6,9		3
6,8	2	5
5,5		1
5		3
4,3	2	13
TOTALE	**9**	**101**

La tabella elenca i CVE che hanno colpito LibreOffice e Microsoft Office in base alla loro gravità, espressa su una scala da 10 per i più gravi a 1 per i meno gravi. Come si può verificare dalla distribuzione dei CVE, LibreOffice è stato colpito da un solo problema grave, e da alcuni problemi di media o bassa gravità.

Il progetto, in ogni caso, dispone degli strumenti e delle risorse per affrontare un eventuale attacco Zero Day. Lo sviluppo viene coordinato da un comitato di esperti nelle varie discipline che afferiscono allo sviluppo stesso - Engineering

Steering Committee, o ESC - che si riunisce settimanalmente per discutere le problematiche correnti della versione in corso di sviluppo, e delle versioni rilasciate sul mercato.

L'ESC è affiancato da un team di esperti nelle problematiche più specifiche della sicurezza, che viene coordinato da RedHat. Questo team si avvale della collaborazione di specialisti di livello mondiale, che in molti casi collaborano come volontari nell'ambito della loro attività come esperti di sicurezza per aziende che sviluppano software in altri settori (un esempio tipico è quello del software per il settore automobilistico).

La sicurezza dei dati del formato nativo di LibreOffice

LibreOffice adotta come formato nativo dei documenti il formato standard aperto Open Document Format, che può contribuire a ridurre la vulnerabilità delle organizzazioni rispetto agli attacchi provenienti dall'esterno, rispetto a quanto può succedere con i formati proprietari dei file. In particolare, riduce il numero dei personal computer che rischiano di essere infettati da virus, spyware e adware.

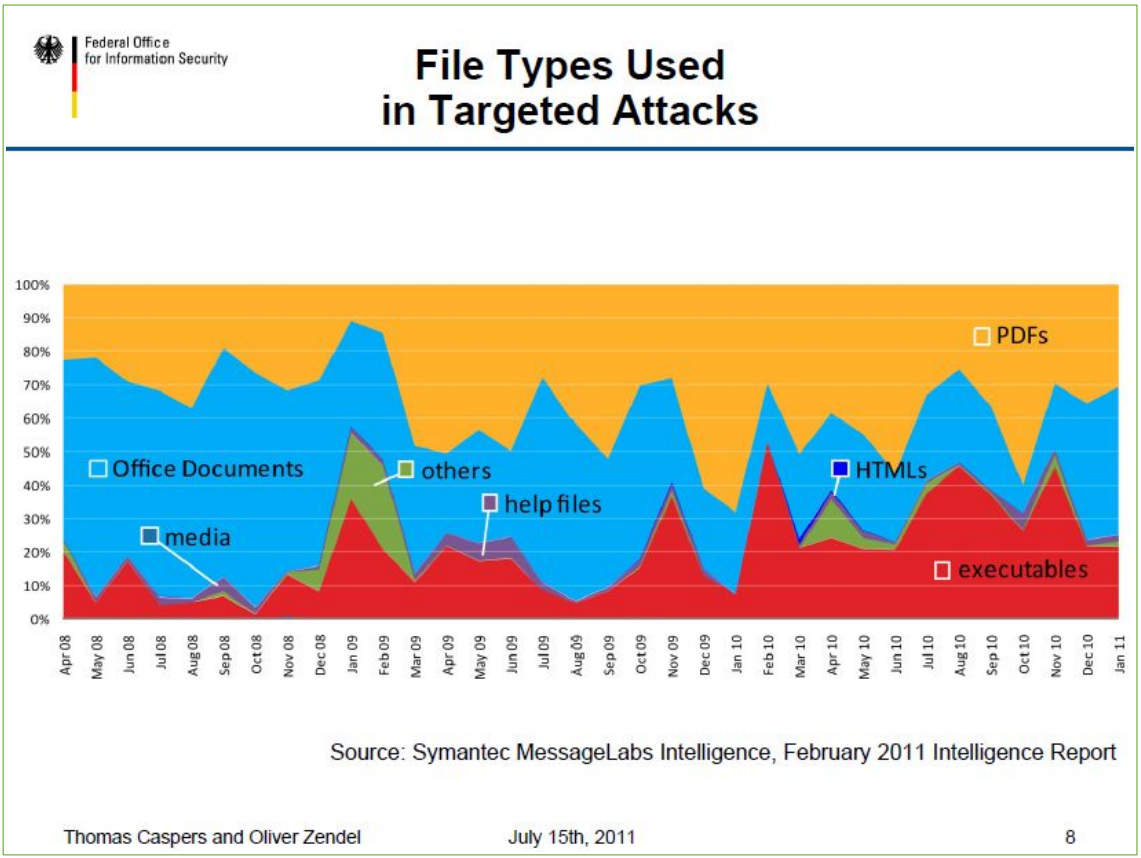
Infatti, i formati proprietari dei documenti da ufficio sono una delle tre vulnerabilità più sfruttate dagli attacchi provenienti dall'esterno. Una ricerca tedesca del 2011 ha rilevato che l'efficacia degli antivirus nei confronti degli attacchi perpetrati attraverso l'uso di file in formato proprietario è limitata. Tre antivirus su quattro hanno fatto registrare una percentuale di riconoscimento pari o inferiore al 20%.

La spiegazione è semplice. I formati proprietari dei file (come DOC e XLS, ma anche DOCX e XLSX, che contengono comunque componenti binari) sono stati concepiti quando la difesa contro gli attacchi provenienti dalla rete non era una priorità, e quindi non faceva parte delle specifiche, perché la maggior parte dei computer era offline. Inoltre, la natura proprietaria di questi formati si traduce nell'impossibilità di analizzarli utilizzando routine pubbliche di validazione in grado di verificare che il documento non contenga nulla di sospetto.

I file binari, e quelli che contengono porzioni di codice binario, giustificando questa presenza con le esigenze di retrocompatibilità, semplificano il compito a chi vuole nascondere al loro interno del codice in grado di sferrare un

attacco, in quanto questo codice è quasi sempre composto da stringhe di 0 e 1.

La stessa ricerca tedesca del 2011 - realizzata da Symantec e MessageLabs - ha rilevato che i formati proprietari dei documenti sono quelli più utilizzati, con il formato standard PDF, come mezzo di trasporto per gli attacchi alla sicurezza. La slide che segue è tratta dalla presentazione della ricerca, effettuata nel corso di una conferenza del Comitato Tecnico OASIS per la gestione del formato ODF, nel corso della quale sono stati discussi proprio i problemi di sicurezza.



Naturalmente, l'uso del formato ODF non riesce - da solo - a trasformare un software poco sicuro in un software sicuro, ma semplifica il compito a chi deve controllare che il documento non nasconda nessun tipo di codice malevolo. La protezione degli utenti e dei loro interlocutori sta al complesso delle misure di sicurezza e ai programmi antivirus adottati dall'organizzazione.