Detecting data loss from cloud sync applications

Jake Williams
CSRgroup Consultants
@MalwareJake
jwilliams@csr-group.com
BLUF

• Just to level set expectations....
• After researching this problem, I don’t have a failsafe exfil detection method for cloud sync apps
  – One may not exist using current tools
• I will share some strategies I’ve used with varying degrees of success
$whoami

- Chief Scientist at CSRgroup
  - Incident Response/Forensics
  - Penetration Testing
  - Exploit Development
- SANS Instructor and author – Malware, Cloud Forensics, Offensive Forensics
- Two time winner of the DC3 Forensics Challenge
- PhD Candidate (Computer Science)
Outline

• The Problem
• Proof of Concept
• Is DLP the Answer?
• Partial Solutions
• Conclusion
The Problem

- Cloud synchronization applications
  - Dropbox
  - SkyDrive
  - Others
- Data placed in special folders is automatically sent to the cloud
  - And replicated to other devices associated with a user’s account
Problem of Feature?

• For some home users, automatic synchronization of files is a *feature*
• For business users it is often a *liability*
  – Your CEO thinks it’s a feature
  – And the CISO knows it’s a liability
Why do you care?

- I don’t manage a corporate network
  - But I do respond to incidents on them
- For obvious reasons, victims want to know if data was exfiltrated
  - And if so, what data
- That’s where we come in as Incident Responders and Forensicators
Install Privileges

• Some cloud synch applications install in the user’s profile directory
  – Removes the need for users to have Power User/Admin privileges

• Most apps have moved to a standard “Program Files” installation over the last year
  – Good news for defending against rogue deployments
Logging

• Cloud sync providers offer varying levels of logging
• Logs are commonly found in SQLite databases and flat files
• Most write to user’s Roaming profile section
• Good news for enterprise wide searches when Roaming profiles are used
Authorized or Not?

• Business leaders need to carefully evaluate the risks associated with these applications
  – Data leakage channel
  – Increased attack surface
  – Malware command and control (C2)
  – Authorized by the firewall
Regulated Data

• Businesses that store or process regulated data should be especially wary of cloud sync
  – HIPAA
  – GLB
  – PII
  – PCI

• Apps may copy regulated data outside of the enterprise without users even being aware
eDiscovery Risks

• Allowing synchronization of business computer to personally owned devices may open these to eDiscovery

• Best Practice: Inform users in writing of eDiscovery risk on personal devices before allowing installation/use of cloud sync
DropSmack

- Proof of concept penetration testing tool
  - Source released at Blackhat EU 2013
- Allows attackers to use Dropbox for C2
  - Supports programmatic exfiltration of data over Dropbox
- The C2/exfil channel is **not** Dropbox specific
  - Only requires the automatic synch of files
DropSmack Use Case

- Highly secure corporate network
- Exploit road warrior outside the defenses of the corporate network
- Infect files in the user’s sync folder
- Wait for user to open files inside corporate network
- Use DropSmack to provide C2, data exfil
DropSmack Use Case

DropSmack Tasking Diagram

Local Dropbox Share

Dropbox Server

Remote Dropbox Share

User Laptop (@Home)

User Workstation (@Work)
New DropSmack (v2)

- Releasing update at Blackhat USA in July
- Operates on more than just just Dropbox
  - Box
  - SkyDrive
  - Google Drive
  - Amazon Cloud Drive
  - Spider Oak
  - SugarSync
  - JustCloud
Offensive Techniques

• Attackers could install cloud sync application on the victim machine
  – msiexec /q FTW!

• NSM shows data exfil to cloud sync provider
  – Attacker’s IP kept out of high volume connection logs

• PoC only, haven’t seen used operationally
Offensive Techniques (2)

- Amazon S3 supplies PowerShell tools
- Write-S3Object CmdLet allows you to upload individual files-folders to S3 Buckets
  - Supports transfers of up to 5GB/file
- I know penetration testers who are using this currently
  - I suspect APT actors are too
  - Malicious insider potential as well
Demo

• Since a seeing is believing...
The Threat

• To review:
  – Malicious insider – purposefully exfiltrates data using cloud sync
  – Clueless insider – unintentionally shares regulated/sensitive data via cloud sync
  – Malicious actor -
DLP Solutions

• Do DLP solutions cover Dropbox?
  – Sort of (read not really)

• Solutions focus on blocking:
  – Dropbox application from running
  – Dropbox communicating on the network

• Problems
  – Sledgehammer approach
  – Application specific
DLP - Symantec

• One forum post suggests file inspection using “Copy to Hard Drive” option
  – No other references to this option found in documentation

• Another post suggests blocking web communications
  – Sledgehammer approach
DLP - McAfee

• McAfee support engineer highlights the problem
  “there's no rule available to block local file transfers”

• Suggests network blocking rule
DLP - Others

• Core problem: if a user can save a file to their profile, they can save it to the cloud
  – DLP solutions must be cloud sync app aware
• Covering default sync directories is a start
  – But not the end
• Searches for DLP solutions that actually address this issue came up empty
  – But it’s a hard problem
NGFW

• Next generation firewalls may offer some capabilities against data loss through cloud providers
• Proxy SSL traffic and decrypt it
• Must integrate with DLP solutions
  – Distinguish legitimate file sync from data exfil
Amazon S3

• Amazon S3 offers cloud storage
  – Dropbox uses S3 as storage backend
• S3 also used by many CDNs
  – NGFW useful to differentiate S3 CDN traffic from file sync
Cheaper Alternative

• Squid proxy can be used for SSL inspection
• Log excerpts below show Dropbox access via a web browser
• Similar information available from standalone application
NSM Solution

• Security Onion supports Bro and ELSA out of the box
  – Point and click installation

• Since becoming a convert, I use bro in all of my IR jobs
  – Using bro to detect cloud sync data exfil is a sledgehammer approach 😞
  – Still useful
Netflow for Exfil Detection

• Workstations normally receive more data than they send
  – Machines used by attackers to pivot into the network and exfil traffic violate this pattern
• Machines using cloud sync often violate this pattern due to high bandwidth backup traffic
  – Look more granularly at the traffic
  – Look at cloud sync traffic in relative isolation
Bro Logs - DNS

• Bro DNS logs are very useful for locating machines running cloud sync software

• Fields
  – id.orig_h -> source IP
  – query -> hostname
  – answers -> shows the answers IPs resolved
Bro Logs - Connections

- Bro DNS logs are very useful for locating machines running cloud sync software
- Fields
  - id.orig_h -> source IP
  - id.resp_h -> dest IP
  - orig_ip_bytes -> bytes sent
  - resp_ip_bytes -> bytes received
Bro Logs - Mashup

- Use DNS logs to find IP addresses used by cloud sync providers
- Use connection logs to look for hosts with unusually high outbound traffic to those IP addresses
  - You define “unusually high”
  - No traffic inspection, could be legitimate sync traffic or exfil
LAN Sync

- Some sync providers offer LAN Sync
- Allows multiple computers on same LAN to sync with one another directly
  - Only one copy comes in from cloud
  - Saves bandwidth
- This is an attack surface I’m actively investigating
LAN Sync (2)

- Clients broadcast constantly to find new friends on the network
  - Reminiscent of NetBIOS 😊
- Easy to find in PCAP
  - Helps to locate illicit cloud sync app installations
LAN Sync - Dropbox

- Dropbox uses TCP and UDP 17500
  - Two broadcasts every 30 seconds
- Unique ID (user) plaintext in broadcast
  - Probably not useful for detecting exfil
LAN Sync - SpiderOak

• SpiderOak uses TCP and UDP 21327 and 21328
  – Two broadcasts every 30 seconds
• Some invariants in payload, may be ID
User Profiles

• Examine user profiles for existence of default sync folders
  – Remember that users can configure additional sync folders / change default

• Scan files currently in sync folders for sensitive information
  – I use document parsers + keyword list for this
User Profiles (2)

- Connect to C$ on user machines
- If investigating a domain and roaming profiles are enabled, go to profile store
- Only finds files currently in the sync folder
Application Logs

• Some application logs may include information about files synchronized
  – Useful for files that no longer exist on host machine or in cloud

• Many apps write logs into user’s Roaming profile directory
  – Potential for one-stop shopping

• File name/size only
  – No content, but some apps store file hashes
Log Examples

• JustCloud

29-06-2013 13:45:24:1374 - Include/Exclude for current backup:
{"CheckRestrictions":[{},{}],"changesMade":false,"includeFolders":{"C:\\Users\\Jake\\Desktop":true,"C:\\Users\\Jake\\Pictures":true,"C:\\Users\\Jake\\Documents":true},"includeFiles":{},"excludeFolders":{"C:\\USERS\\JAKE\\SYNC\"}:true},"excludeFiles":{}}
29-06-2013 13:45:24:1374 - Excluded file types for current backup:
29-06-2013 13:45:25:4010 - queued 0 files (0 Bytes) for drive A:\
29-06-2013 13:45:25:4478 - queued 4 files (73.33 KB) for drive C:\
29-06-2013 13:45:25:4478 - finished queueing files
Log Examples

- SugarSync

```
[7/1/2013 9:13:54 AM:866600] [error] [client.cloudtab] [3820] [qsccloudtab.cpp:417 DCL-1977: Enter]
[7/1/2013 9:13:54 AM:866600] [error] [client.cloudtab] [3820] [qsccloudtab.cpp:424 DCL_1977: pData != NULL]
[7/1/2013 9:13:54 AM:866600] [error] [client.dragdrop] [3820] [qscdraganddrop.cpp:613 DCL_1977:
QScFolderDragAssistant: Enter]
[7/1/2013 9:13:54 AM:866600] [error] [client.dragdrop] [3820] [qscdraganddrop.cpp:616 DCL_1977:
QScFolderDragAssistant: iteration = 0]
[7/1/2013 9:13:54 AM:866600] [error] [client.config] [3820] [qscsyncconfigurationmanager.cpp:176 DCL_1977:
QScFolderDragAssistant: isPathSynced: Enter with path = C:/Users/Jake/Desktop/atombomb]
[7/1/2013 9:13:54 AM:866600] [info] [client.config] [3820] [Root folders: 2]

[7/1/2013 9:13:56 AM:395400] [info] [fs.sync.root] [3892] [starting fs enumeration of root
/sc/6938908/26731539_16578 at "C:\Users\Jake\Desktop\atombomb" to merge contents]
[7/1/2013 9:13:56 AM:395400] [info] [fs.sync] [3892] [[fs] add "C:\Users\Jake\Desktop\atombomb\hydrogen.zip"
reserved dsid /sc/6938908/26731539_16581]
[7/1/2013 9:13:56 AM:395400] [error] [files.statusmanager] [2752] [qscstatusmanager.cpp:1178
QScFileNotificationListener::onNotification: received invalid key for pNotification 0x99af5cc
```
Application Databases

• Many cloud sync apps store information in SQLite database files
  – More often than plaintext logs, these tend to store file hashes

• Use SQLite and Python to dump databases to text and search for keywords
  – Use more surgical searches only if keyword hits are found
Database Examples

- JustCloud
  - mpcb_file_cache.db

<table>
<thead>
<tr>
<th>path</th>
<th>fileName</th>
<th>hash</th>
<th>size</th>
<th>mtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>{syncfolder}</td>
<td>{syncfolder}</td>
<td></td>
<td>0</td>
<td>-8588290789708405808</td>
</tr>
<tr>
<td>{syncfolder}\</td>
<td>JustCloud Quick Start Guide.pdf</td>
<td>9c92b30d88e47418651552e04056176</td>
<td>991271</td>
<td>-8588290790150197808</td>
</tr>
<tr>
<td>{source:11210377}\</td>
<td>C:</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>{source:11210377}\C:| Users</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>{source:11210377}\C:\Users\</td>
<td>Jake</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>{source:11210377}\C:\Users\Jake\</td>
<td>Desktop</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>{source:11210377}\C:\Users\Jake\Desktop\</td>
<td>desktop.ini</td>
<td>9e36cc3537ee9ee1e3b10fa4e761045b</td>
<td>282</td>
<td>-8588357149342454653</td>
</tr>
<tr>
<td>{source:11210377}\C:\Users\Jake\Desktop\</td>
<td>regshot.exe</td>
<td>aaa8fbcace9c4999a77d63e0fa80f85</td>
<td>73728</td>
<td>-8590074634774775808</td>
</tr>
</tbody>
</table>
Web Applications

• Many cloud sync providers also support access via web applications
  – Web apps may keep records of files that were deleted on the host
• In some cases, the deleted file itself may be available for a period of time
  – Like a ‘Recycle Bin’ in the cloud
Web Applications (2)

• If writing a warrant/eDiscovery subpoena, request access to suspect’s online storage accounts
  – Common procedure for webmail

• Logs on the web application have the best historical record of sync activity
  – Most online logs are not editable by the user
  – Many log source IP addresses
Dropbox RSS

- Dropbox is unique in offering RSS
- If configured, you could at least monitor file names being synchronized
Dropbox Deleted Files

- Dropbox allows the display/restoration of deleted files
- All accounts get storage of 30 days of deleted files
Dropbox Deleted Files (2)

- Dropbox Pro and Dropbox Business may use an add-on called PackRat
  - PackRat offers unlimited recovery of deleted files/folders (not retroactive)
- Also offers version history for sync’d files
  - Especially interesting in eDiscovery cases
  - Depending on your goals, either mandate these settings on or off
Dropbox Events

- Perpetual log of Dropbox activity

Events gives you a timeline of everything that’s happened in your Dropbox since the beginning of time.

- a guy invited you to the shared folder 'a' 6/29/2013 8:29 PM
- Alissa Torres joined the shared folder 'Memory-Images' 6/28/2013 9:09 PM
- You joined the shared folder 'SRL - Malware' 2/14/2013 12:10 AM
- You joined Dropbox 2/14/2013 12:10 AM
SugarSync Deleted Files

- SugarSync allows remote restoration of deleted files or download directly from them.
SugarSync Logs

• The web interface lists files added
  – Not as granular as Dropbox Events
  – Doesn’t show files shared or deleted
Conclusions

- There’s currently a technology gap for detecting data exfiltration via cloud sync apps
- NSM helps
  - But only when clients have it in place
- Local log files on some apps fill some gaps
  - Limited information though
- Web app logs are the best resource
  - When you can get *legal* access to them
Questions?

Thanks for your time!

Jake Williams
@MalwareJake
jwilliams@csr-group.com