Forensically Important Artifacts Resulting from Usage of Cloud Client Services

Robin Verma, Anuradha Gupta, Ankit Sarkar and

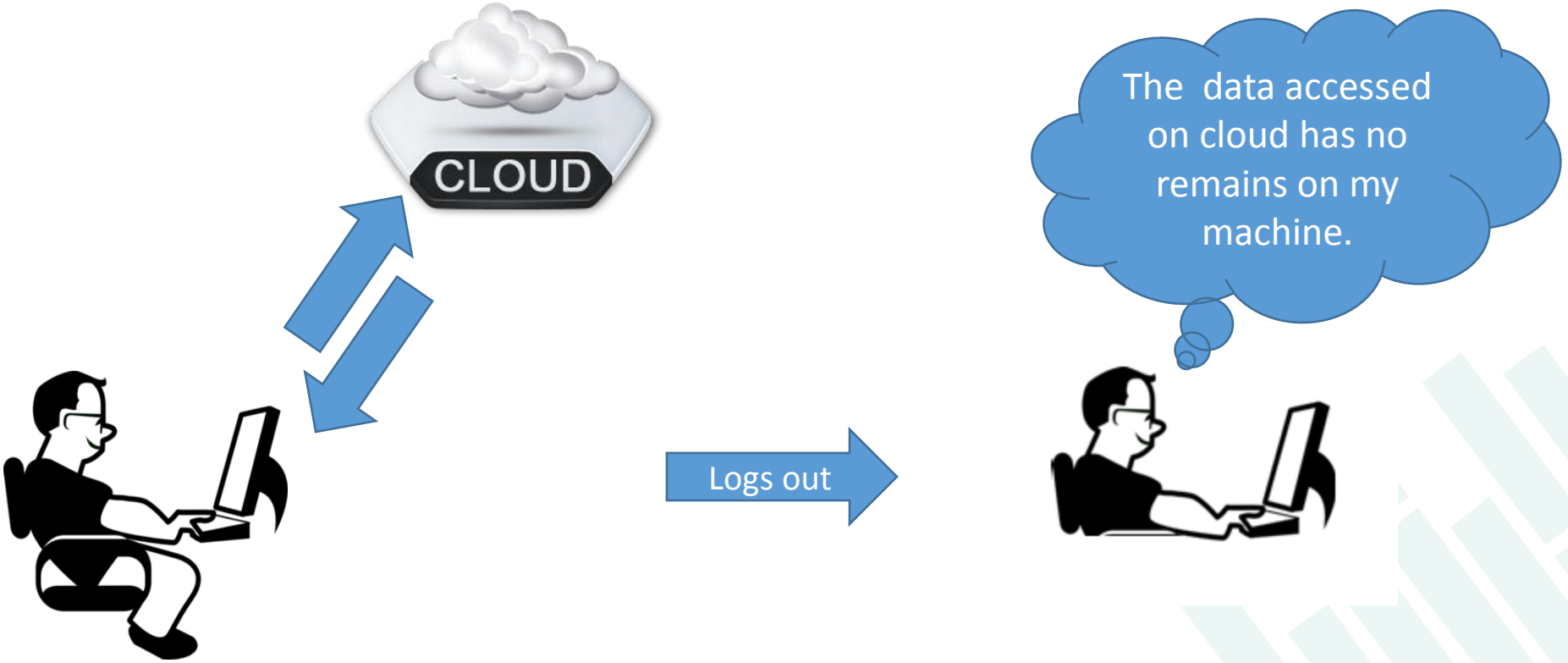
Gaurav Gupta

gauravg@iiitd.ac.in

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Cloud Services Myths

A user after logging off from a cloud based service assumes that all the confidential information is not available to other using the same machine.



Cloud Services Myths

- Is it true?
- Are there any remnant artifacts generated from the usage of cloud bases services available on the client machine?
- Is it possible to correlate information about user, her activities and data available on client machine ?

Approach

Data

- Collect remnant data
- Source: browser cache, RAM dump, registry and Hyberfil.sys

Activity

- Form user-activity

User

- Link it to an account or a suspect

Some of the services

Google Services:

- Gmail
- Google Docs
- Google Groups
- Google Translate
- Google talk
- Orkut

Cache Scenario

- Disk Cache
  - Written on disk, available after power off.
  - Paged out and deleted entries can be retrieved*

- RAM cache
  - System is ON – Live Dump.
  - Cold-Boot attack
  - Page-out on Disk (pagefile.sys & hiberfil.sys)

* Provided they are not overwritten, probability of which decreases with large HDD size.

Gmail

1

User tries to login into her Gmail Account('gxlu'* URL for each login attempt)

Last successful attempt generates

2

User is logged into all of the Google Services(new 'ServiceLoginAuth' URL in cache)

All user activity can be linked until... logout

3

User is logged out of all the services

('Logout' URL updated)

User can Login again

*Not with hosted services like iiitd.ac.in

Google Docs

- Every static document is converted to PNG images, one per page.
- The URL that is generated is like:

https://docs.google.com/viewer?**pid**=explorer&**srcid**=0BkbwV2WQQNVMDQyZDY0YjYtYmZmNy00NGYxLWExYmUtNzJjZDViNWQ0Njkx&**chrome**=true&**docid**=4d97f4be2d72ad50cff9f330935af38e%7Cc7d626e86409c7b11a5ab9f5d555371b&a=bi&**pagenumber**=1&**w**=800

- Important fields in the URL are boldfaced.

URL information from one PNG file

```
1  https://docs.google.com/viewer?pid=explorer&srcid=0B-kbwV2WQQNVYW00LW0
   tNmtTMVE&docid=4cdbc3a1bc17c73835b4985f6b475b1d%7Cbaeed72f4d4753847e94
   6ad9a90226c5&a=bi&pagenumber=2&w=1004
```

```
srcid=0B-kbwV2WQQNVYW00LW0tNmtTMVE
'\xd0\x19\x1b\xc1]\x96A\x03Uam4-m-6kS1Q4'
```

Source ID - unique for the user <BASE64URL>

```
docid=4cdbc3a1bc17c73835b4985f6b475b1d%7Cbaeed72f4d4753847e946ad9a90226c5

4cdbc3a1bc17c73835b4985f6b475b1d

baeed72f4d4753847e946ad9a90226c5
```

Docid Part 1: MD5 hash of the document.

Docid Part 2: MD5 hash, Unique identifier for document + viewer

```
pagenumber=2
```

Page number to collect all PNG's of this document

```
w=1004
```

Page width to collect all PNG's of same size

Inference of Artifacts

ID Name	Data-Type	Example	Information
"srcid"	Base64url (Before Feb'12)	**0B-kbwV2WQQNV**MDQyZDY0YjYtYmZmNy00NGYxLWExYmUtNzJjZDViNWQ0Njkx	First 13 characters are unique for given account. Rest of the 48 characters are ***UUID version 4*[1].**
"srcid"	Base64url (After Feb 25, 12)	**0B-kbwV2WQQN**VbzFCWnhHbktRWmVMNmVhUDUzOFF2UQ	Same as above, but string length is reduced to **43**.
"srcid"	Base64url (After April'12)	**0B-kbwV2WQQNVU**mdYX2djSHpoa2c	String length further reduced to **28**.

[1]**Universally Unique Identifier (UUID)** version 4 contains a totally random string.

Inference of Artifacts

ID Name	Data-Type	Example	Information
"docid" (part-1)	32 characters hexadecimal (***Document-Binding***)	4d97f4be2d72ad50cff9f330935af38e	It's MD5 hash of the uploaded document. Unique identifier of the document.
"docid" (part-2)	32 characters hexadecimal (***Viewer-Binding***)	c7d626e86409c7b11a5ab9f5d555371b	Separated by '|' in the URL, It's dependent on document content as well as the viewer's account information.

String in the url is like:

`docid=4d97f4be2d72ad50cff9f330935af38e%7Cc7d626e86409c7b11a5ab9f5d555371b`

Where '%7C' is html code for '|'.

ID Name	Data-Type	Example	Information
"pagenumber"	Numeric value	**pagenumber**=3	All PNG's associated with a given document can be collected using page-number.
"w"	"	**w**=800	All PNG's of same size can be collected separately.

Google Groups

- The URL obtained in cache is as follows:
  - `http://groups.google.com/groups/profile?hl=en&enc_user=3Ww4IxYAAAD9zlh47LY6GAD_U_Shqhuoo4cocwWvDVg2RHsu8f1bCg`

ID Name	Data-Type	Example	Information
"enc_user"	Base64url	3Ww4IxYAAAD9zlh47LY6GAD_U_Shqhuoo4cocwWvDVg2RHsu8f1bCg	This is unique ID that is linked to a particular Google-account. Every time a new group is subscribed by the account, same ID is generated.

ID Name	Data-Type	Example	Information
"thread"	16 digit Hexadecimal	eda1995409680647	This is unique identifier of the thread/post which is generated on the group.

Orkut

ID Name	Data-Type	Example	Information
"uid"*	Numeric Value	59379498002309011 32 derived from *uid*	19 (sometimes 20) digit number which uniquely identifies given Orkut account.

*it is found in the cache entry starting from "http://talkgadget.google.com/talkgadget/notifierclient...". The user's account name is also included in this cached URL, i.e. (We have replaced the original username with "**xyz**" in the following URL to maintain anonymity of the actual user) –

http://talkgadget.google.com/talkgadget/notifierclient?client=sm&prop=Orkut&nav true&fid=gtn-roster-iframe-id&ts=0&debug=undefined&os=Win32&stime=1336534817898&fb=false&re=false&no=undefined&hc=undefined&ref=undefined&xpc=%7B%22cn%22%3A%22y5a7iw%22%2C%22tp%22%3A1%2C%22ifrid%22%3A%22gtn-roster-iframe-id%22%2C%22pu%22%3A%22http%3A%2F%2Ftalkgadget.google.com%2Ftalkgadget%2F%22%2C%22lpu%22%3A%22http%3A%2F%2Fwww.orkut.co.in%2Frobots.txt%22%2C%22ppu%22%3A%22http%3A%2F%2Ftalkgadget.google.com%2Frobots.txt%22%7D&href=http%3A%2F%2Fwww.orkut.co.in%2FMain%23AlbumList%3Frel%3D1%26uid%3D14692977986891793375&pos=l&css=http%3A%2F%2Fstatic4.orkut.com%2Fcss%2Ftalk%2Fger r002.css&hl=en-US&uj=**xyz**%40gmail.com&hpc=true&hsm=true&hff=true&hrc=true&hotr=true&vp=http%3A%2F%2 ut.co.in%2Fxpc%2Fblank&uqp=false&sl=false&host=1&zx=m8i4873zoioe

Picture sizes in Orkut

PROFILE PIC:

http://img8.orkut.com/images/**medium**/1336561055/24617184/ln.jpg

COMMUNITY PIC:

http://img4.orkut.com/images/**klein**/43/13207843.jpg

FRIEND'S PIC:

http://img4.orkut.com/images/**small**/1241228446/22768592/ep.jpg

'RECENT VISITORS' & 'PEOPLE YOU MIGHT KNOW' PIC:

http://img5.orkut.com/images/**tiny**/1282426547/48665360/gq.jpg

ALBUM PHOTOS:

http://img6.orkut.com/images/**milieu**/1336455606/1336480807388/24617184/ln/Z12x9u79.jpg

Picture uploaded by user

ALBUM Picture:
http://img4.orkut.com/images/milieu/1336470060/1336495261485/24617184/ln/Z10nedbs.jpg?ver=1336495267

1336470060:
It is the ALBUM ID of the album in which the pic exists. It is the Unix Time for upload of the album.

1336495261485:
It is the PICTURE ID of the pic. It is the Unix Time for upload of the picture (last three digits are milliseconds).

24617184:
Unique Shortened USER ID which is common among all the pictures uploaded by the user into any of her albums.

Google Translate

Translation is captured in the cache of the browser like:

http://translate.google.com/translate_a/t?client=t&**text**=voila&**hl**=en&**sl**=fr&**tl**=ar&multires=1&otf=2&ssel=0&tsel=0&sc=1

- Field ‘text’ – input word,
- Field ‘hl’ - home-language in which the Google Translate page is being displayed,
- Field ‘sl’ - code for the source language, and
- Field ‘tl’ - code for target language.

Picasaweb Albums

ID Name	Data-Type	Example	Information
user	Numeric Value (21 digit)	104538393120023502592	This 21 digit number is unique for a particular account on picasaweb.
albumid	Numeric Value (19 digits)	5740856268079783185	This 19 digit number is unique for a particular album made for a particular account on picasaweb.
photoid	Numeric Value (19 digits)	5740856038908974882	This 19 digit number is unique for a particular photo uploaded in the album of a particular account on picasaweb.

Google Talk

Registry Analysis

Google Talk stores its user specific information in the registry key. `KEY_CURRENT_USER\Software\Google\GoogleTalk”` and its sub-keys. We analyzed it to come up with the following findings.

Login Usernames

In the registry, Google Talk maintains a list of user names which have logged in through the Google Talk.

Client

We logged in using the accounts (X and Y) in this order and found the following modifications in `“HKEY_CURRENT_USER\Software\Google\GoogleTalk\Accounts”`

reflects the user details including the last user who successfully logged in and a list of users who logged in using Google Talk. Furthermore, on experimentation, it was found the usernames were arranged in reverse chronological order of last accessed time.

Google Talk

- In addition 2 new subkeys were created under `HKEY_CURRENT_USER\Software\Google\GoogleTalk\Accounts` namely
  - `HKEY_CURRENT_USER\Software\Google\GoogleTalk\Accounts\X`, and
  - `HKEY_CURRENT_USER\Software\Google\GoogleTalk\Accounts\Y`.
- Thus we can conclude that on every successful login, a new subkey of the generic form "`HKEY_CURRENT_USER\Software\Google\Google Talk\Accounts\<username>@gmail.com`" is created.

Account Password

On selecting the option to save the password for a user X, a new value was added to `HKEY_CURRENT_USER\Software\Google\GoogleTalk\Accounts\X`.

- Furthermore, on unchecking the option to save the password, this was deleted.
- Hence, we can conclude that "pw" contains the encrypted password of the concerned username which can be deciphered using windows function named "CryptUnProtectData".

THANK YOU!

gauravg@iiitd.ac.in