CHAMPLAIN COLLEGE

LCDI

The Senator Patrick Leahy
Center for Digital Investigation

Zeitgeist Forensics

Written by
DJ Palombo
Researched by
DJ Palombo

175 Lakeside Ave, Room 300A
Phone: 802/865-5744
Fax: 802/865-6446
http://www.lcdi.champlin.edu

February 2014

Disclaimer:

This document contains information based on research that has been gathered by employee(s) of The Senator Patrick Leahy Center for Digital Investigation (LCDI). The data contained in this project is submitted voluntarily and is unaudited. Every effort has been made by LCDI to assure the accuracy and reliability of the data contained in this report. However, LCDI nor any of our employees make no representation, warranty or guarantee in connection with this report and hereby expressly disclaims any liability or responsibility for loss or damage resulting from use of this data. Information in this report can be downloaded and redistributed by any person or persons. Any redistribution must maintain the LCDI logo and any references from this report must be properly annotated.

Contents

Introduction

As Linux is beginning to gain a larger market share of the computing world, there is a larger need for forensic processes that will allow an examiner to determine what a user was doing on a computer at a specific time and date, as well as which user was opening what programs. Zeitgeist is a tool built into the Gnome Desktop Environment that will create a log of user activity in a SQLite database, which is used to help predict the user activity. The goal of this project was to determine if there was significant forensic value to this SQLite database.

Background:

There has been no prior publicized research done on this topic.

Purpose and Scope:

The overall goal of this project was to determine what can be used from a Linux environment for forensic analysis. As Linux is growing in popularity, it is possible that more forensic researchers will be encountering computers that are running Linux. As this can become a larger issue, it is important that the examiner be able to create a timeline on a Linux system, much the same as the examiner can do on a Windows or Mac based system. Zeitgeist is integrated into the Gnome Desktop Environment (GDE), and many users are unaware that it is operating in the background. Because of this, many users do not disable it or attempt to remove information from it. The information available may be able to direct a forensic investigator to the actions of a particular user on the system, and assist in creation of a timeline and pattern analysis.

Research Questions:

1. What user activity is stored in the Zeitgeist SQLite database?
2. How accurate is the information stored in the database?
3. When are the entries into the database created?
4. Is there a way to differentiate between actions of different users?
5. Is there a way to differentiate between actions through the terminal versus the graphical user interface?

Terminology:

SQLite Database: A relational, standalone database that is integrated into many programs, operating systems, and embedded systems.

Zeitgeist: A service installed in the Gnome Desktop Environment that logs user activities and events on the system.

Linux: A UNIX-like operating system that is openly distributed as open source. There are many different distributions, each being a variation off the same Linux kernel.

Terminal: A command line interface that allows a user to interact with a system in text-only mode. Similar in function to a Windows Command Prompt.

Forensic Tool Kit (FTK): Computer forensic software created by AccessData.

Forensic Tool Kit (FTK) Imager: Standalone version of FTK used for disk imaging purposes.

Methodology and Methods

To achieve the goals set out in the Research Sections portion of this report, a virtual machine needed to be set up in order to have a standalone Linux environment that could be created without contamination. This was achieved through a VMWare Workstation. The operating system that was selected to be used as a basis was Linux Mint 12 (GNOME), as it uses the Gnome Desktop Environment, and therefore has Zeitgeist installed. User activity was created in this virtual machine, with each action recorded in an Excel spreadsheet, logging the action taken and the time it was taken. The virtual hard drive was then loaded into FTK Imager in order to extract the SQLite database, which was then loaded into FTK for parsing and further data extraction. This information was then compared to the Excel spreadsheet that was created from the user activity to verify the information being collected from the database.

Equipment Used

Equipment Used	Purpose
Desktop computer with 16 GB of RAM	Used to allow for a virtual machine to be running, as well as to allow for use of forensic tools
VMWare Workstation	Used to create a virtual machine for testing purposes
FTK Imager	Used to mount a read only copy of the virtual hard drive (.vmdk) to extract a copy of the SQLite database used by Zeitgeist
FTK	Used to load the SQLite database into to view the tables stored within the database

Data Collection:

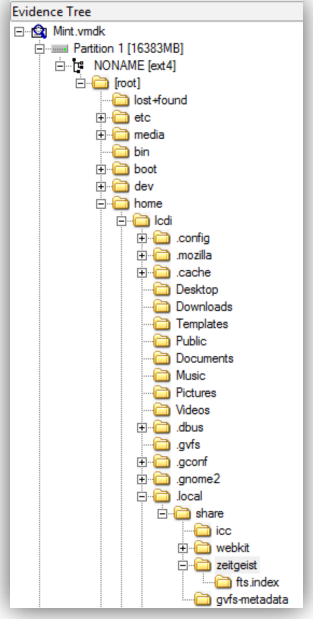
The table below is taken from the Excel document that was created during the evidence creation process.

Table 1: Excel document containing programs and date and time launched

terminal	11/5/13 4:10 PM
VMWare tools mounted	11/5/13 4:12 PM
archive manager	11/5/13 4:12 PM
Installing vm tools	11/5/13 4:13 PM
terminal	11/5/13 4:15 PM
software manager	11/5/13 4:17 PM
terminal	11/5/13 4:22 PM
solitaire	11/5/13 4:31 PM
solitaire	11/5/13 4:34 PM
solitaire	11/5/13 4:38 PM
solitaire	11/5/13 4:40 PM
solitaire	11/5/13 4:43 PM
solitaire	11/5/13 4:46 PM
solitaire	11/5/13 4:48 PM
terminal	11/5/13 4:53 PM
log in/desktop	11/6/13 1:16 PM
terminal	11/6/13 1:17 PM

gedit (from terminal)	11/6/13 1:18 PM
Saved testfile from gedit to Desktop	11/6/13 1:19 PM
Firefox	11/6/13 1:20 PM
solitaire	11/6/13 1:59 PM
solitaire	11/6/13 2:07 PM
Firefox	11/6/13 2:21 PM
Opened & modified testfile	11/6/13 2:26 PM
terminal	11/6/13 2:27 PM
Firefox From Terminal	11/6/13 2:27 PM
powered on	11/12/13 12:50 PM
system settings	11/12/13 12:52 PM
created new user	11/12/13 12:52 PM
log out lcdi	11/12/13 12:54 PM

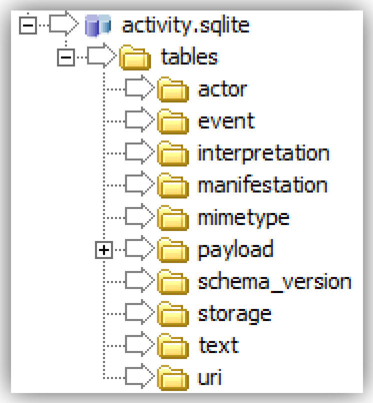
To collect the data from the virtual machine, the virtual hard disk (VMDK) was mounted into FTK Imager as an image file.



```
Evidence Tree
⊟ Mint.vmdk
   ⊟ Partition 1 [16383MB]
      ⊟ NONAME [ext4]
         ⊟ [root]
              lost+found
            ⊞ etc
            ⊞ media
              bin
            ⊞ boot
            ⊞ dev
            ⊟ home
               ⊟ lcdi
                  ⊞ .config
                  ⊞ .mozilla
                  ⊞ .cache
                    Desktop
                    Downloads
                    Templates
                    Public
                    Documents
                    Music
                    Pictures
                    Videos
                  ⊞ .dbus
                    .gvfs
                  ⊞ .gconf
                  ⊞ .gnome2
                  ⊟ .local
                     ⊟ share
                          icc
                        ⊞ webkit
                        ⊟ zeitgeist
                             fts.index
                          gvfs-metadata
```

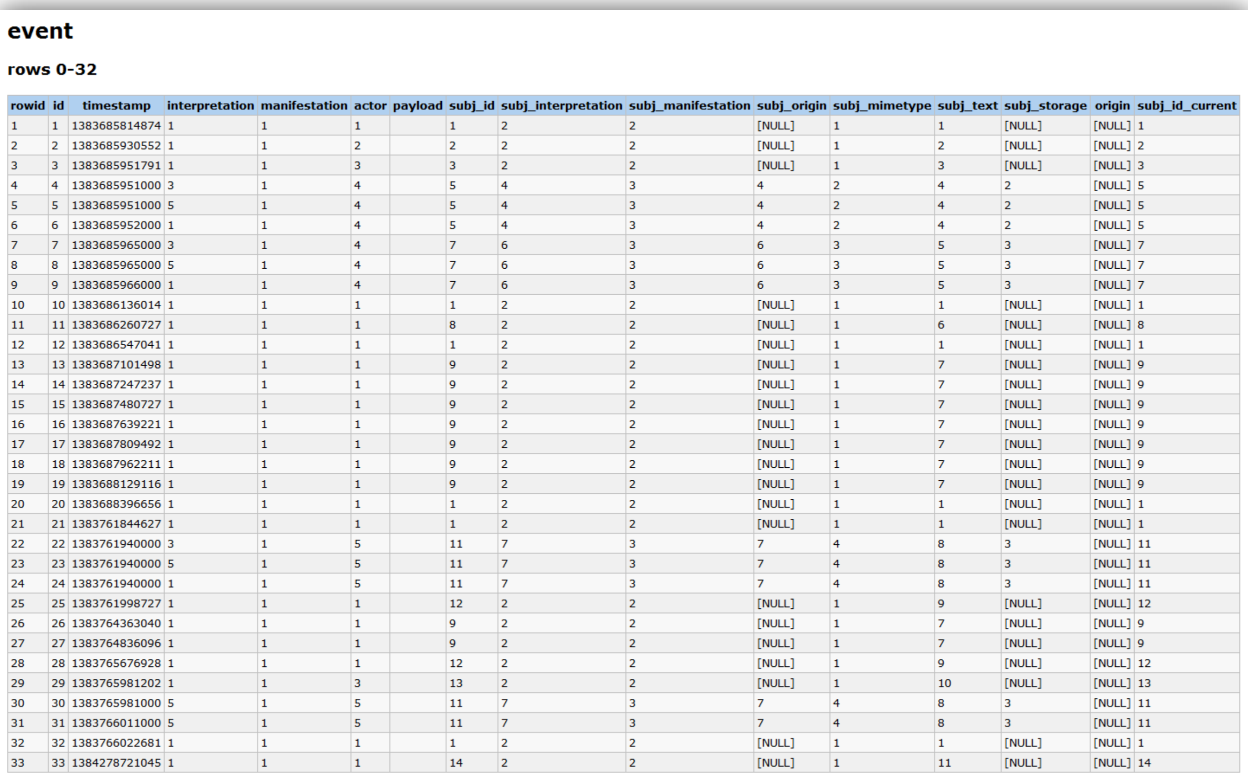
Analysis

The SQLite database, when put into FTK, displayed multiple folders that made up the database.



activity.sqlite
- tables
  - actor
  - event
  - interpretation
  - manifestation
  - mimetype
  - payload
  - schema_version
  - storage
  - text
  - uri

Each of these folders contains a table that houses data on the user activity from Zeitgeist. Each of these individual tables does not mean much as a standalone record. The records are all interacting with each other, and the tables reference each other in order to provide a full picture of the user activities. The main table in the database is the file in the "event" folder of activity. SQLite. The file is titled rows_0000000_xxxxxxx.html, where the value of the x's is one less than the number of entries in that file. For example, in one of the tests, the file is named rows_0000000_0000032.html, meaning that it had 33 separate entries of user activity in the table. This example is displayed below for reference.

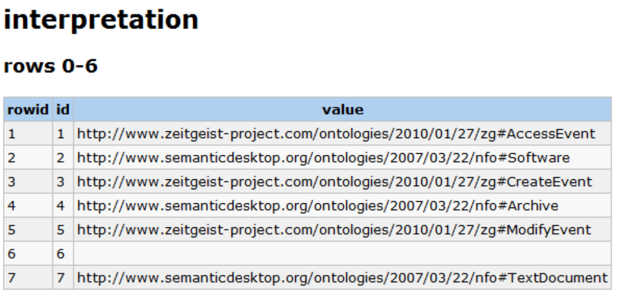


event

rows 0-32

rowid	id	timestamp	interpretation	manifestation	actor	payload	subj_id	subj_interpretation	subj_manifestation	subj_origin	subj_mimetype	subj_text	subj_storage	origin	subj_id_current
1	1	1383685814874	1	1	1		1	2	2	[NULL]	1	1	[NULL]	[NULL]	1
2	2	1383685930552	1	1	2		2	2	2	[NULL]	1	2	[NULL]	[NULL]	2
3	3	1383685951791	1	1	3		3	2	2	[NULL]	1	3	[NULL]	[NULL]	3
4	4	1383685951000	3	1	4		5	4	3	4	2	4	2	[NULL]	5
5	5	1383685951000	5	1	4		5	4	3	4	2	4	2	[NULL]	5
6	6	1383685952000	1	1	4		5	4	3	4	2	4	2	[NULL]	5
7	7	1383685965000	3	1	4		7	6	3	6	3	5	3	[NULL]	7
8	8	1383685965000	5	1	4		7	6	3	6	3	5	3	[NULL]	7
9	9	1383685966000	1	1	4		7	6	3	6	3	5	3	[NULL]	7
10	10	1383686136014	1	1	1		1	2	2	[NULL]	1	1	[NULL]	[NULL]	1
11	11	1383686260727	1	1	1		8	2	2	[NULL]	1	6	[NULL]	[NULL]	8
12	12	1383686547041	1	1	1		1	2	2	[NULL]	1	1	[NULL]	[NULL]	1
13	13	1383687101498	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
14	14	1383687247237	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
15	15	1383687480727	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
16	16	1383687639221	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
17	17	1383687809492	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
18	18	1383687962211	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
19	19	1383688129116	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
20	20	1383688396656	1	1	1		1	2	2	[NULL]	1	1	[NULL]	[NULL]	1
21	21	1383761844627	1	1	1		1	2	2	[NULL]	1	1	[NULL]	[NULL]	1
22	22	1383761940000	3	1	5		11	7	3	7	4	8	3	[NULL]	11
23	23	1383761940000	5	1	5		11	7	3	7	4	8	3	[NULL]	11
24	24	1383761940000	1	1	5		11	7	3	7	4	8	3	[NULL]	11
25	25	1383761998727	1	1	1		12	2	2	[NULL]	1	9	[NULL]	[NULL]	12
26	26	1383764363040	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
27	27	1383764836096	1	1	1		9	2	2	[NULL]	1	7	[NULL]	[NULL]	9
28	28	1383765676928	1	1	1		12	2	2	[NULL]	1	9	[NULL]	[NULL]	12
29	29	1383765981202	1	1	3		13	2	2	[NULL]	1	10	[NULL]	[NULL]	13
30	30	1383765981000	5	1	5		11	7	3	7	4	8	3	[NULL]	11
31	31	1383766011000	5	1	5		11	7	3	7	4	8	3	[NULL]	11
32	32	1383766022681	1	1	1		1	2	2	[NULL]	1	1	[NULL]	[NULL]	1
33	33	1384278721045	1	1	1		14	2	2	[NULL]	1	11	[NULL]	[NULL]	14

The rowID and ID both have consistently held the same value through all of the test scenarios that were run. The timestamp column displays the specific date and time that the activity occurred. This timestamp is in epoch time (also known as Unix time or POSIX time), which is the amount of milliseconds since midnight (UTC) on January 1, 1970. This time can be converted using a formula in Excel, and the conversion can take into account the changes in time zone. The interpretation column tells what the event actually is. In the example event sheet listed above, the only activity it displays in the interpretation is 1, 3 and 5, which is Access Event, Create Event, and Modify Event.

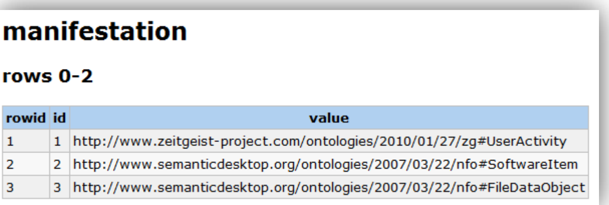


interpretation

rows 0-6

rowid	id	value
1	1	http://www.zeitgeist-project.com/ontologies/2010/01/27/zg#AccessEvent
2	2	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#Software
3	3	http://www.zeitgeist-project.com/ontologies/2010/01/27/zg#CreateEvent
4	4	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#Archive
5	5	http://www.zeitgeist-project.com/ontologies/2010/01/27/zg#ModifyEvent
6	6	
7	7	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#TextDocument

The next column in the "event" table is manifestation. Manifestation relates what created the action, whether it was the user, software, or a file.

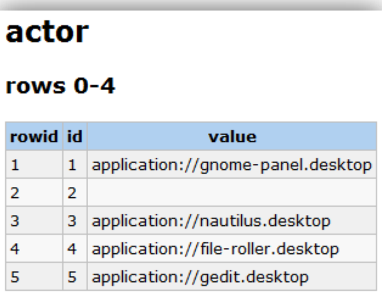


manifestation

rows 0-2

rowid	id	value
1	1	http://www.zeitgeist-project.com/ontologies/2010/01/27/zg#UserActivity
2	2	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#SoftwareItem
3	3	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#FileDataObject

In most instances, the user is creating the activity. This also helps refute any defense that a piece of software or malware was responsible for any actions taken, as it would show up in the manifestation column. The next column that is presented, titled "actor," pertains to how the change to the system was made.

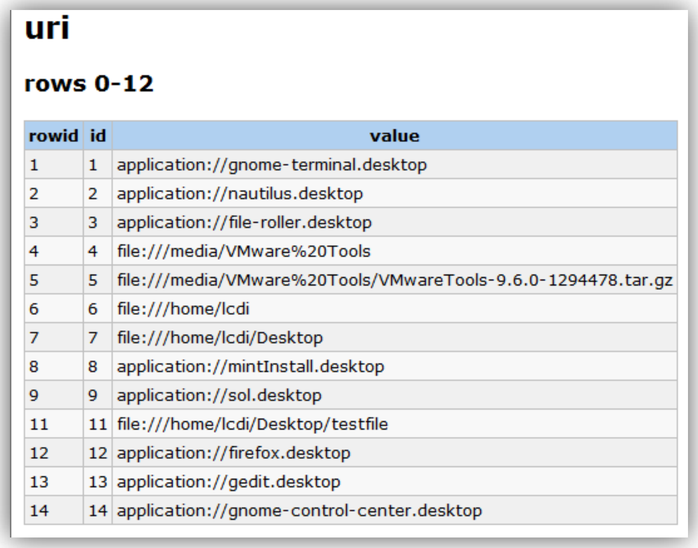


actor

rows 0-4

rowid	id	value
1	1	application://gnome-panel.desktop
2	2	
3	3	application://nautilus.desktop
4	4	application://file-roller.desktop
5	5	application://gedit.desktop

The majority of actions were made from the gnome panel, which is the equivalent to the start menu on Windows systems. Nautilus desktop is a group of files that are being run off the desktop. File-roller is a compressed file manager and files can be launched from that system. Gedit is a text editor program and files can be opened and edited in gedit. The next column listed is subj_id. This column does not link directly by name to the database; the information is stored in the "uri" folder.

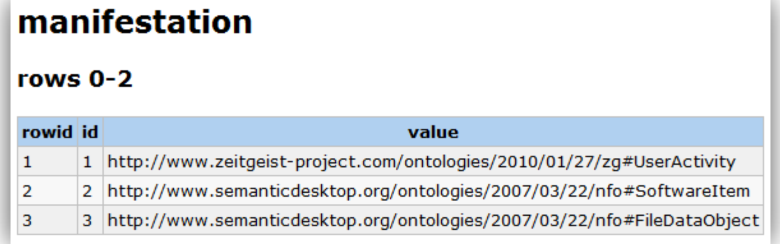


uri

rows 0-12

rowid	id	value
1	1	application://gnome-terminal.desktop
2	2	application://nautilus.desktop
3	3	application://file-roller.desktop
4	4	file:///media/VMware%20Tools
5	5	file:///media/VMware%20Tools/VMwareTools-9.6.0-1294478.tar.gz
6	6	file:///home/lcdi
7	7	file:///home/lcdi/Desktop
8	8	application://mintInstall.desktop
9	9	application://sol.desktop
11	11	file:///home/lcdi/Desktop/testfile
12	12	application://firefox.desktop
13	13	application://gedit.desktop
14	14	application://gnome-control-center.desktop

This table shows a list of all the programs and files that have been accessed. The items are each inserted into the table the first time they are used, and then the same ID number is used each subsequent time. There is a clear distinction made between files and applications in the value column, where applications begin with application:// and files and folders begin with file:///, where the added "/" is the leading slash for denoting a folder in a Linux folder structure. Again, referring back to the main chart, subj_interpretation is the next column of interest. This column seems to be calling back to the interpretation folder, although there is another column referring to that chart. The difference between the two columns is that the first seems to be the action taken (what changes are being made). In the column subj_interpretation, it seems that it is referencing what is causing the action to be taken, be it software, a specific program, or a blank line that has not been populated in any of the test environments that have been made. The next table referenced is the subj_manifestation table.



manifestation

rows 0-2

rowid	id	value
1	1	http://www.zeitgeist-project.com/ontologies/2010/01/27/zg#UserActivity
2	2	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#SoftwareItem
3	3	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#FileDataObject

This table is seemingly referring to what is being logged, if it is the action of a piece of software, an action of the user, or an action of the file system. The majority of actions will be from software items, as software allows for the user to interface with the system in the most common and simple way. None of the remaining data has any clear linkage to other tables in the database.

Results

Through the research that was conducted into Zeitgeist, we were able to yield some results of forensic value. The explanation of the database table displayed above will allow the reader to view a Zeitgeist database and determine user activity through that database. The entire database must be exported to allow for the examination to occur. The testing has been run multiple times in order to verify results. The methodology was to create a virtual machine, and log every event that was created through use of the system. After using this process for a period of time, the database was exported using the process detailed earlier, and the analysis was performed using the methodology explained through this document. The database was compared with the notes that were taken during the use of the system. The notes consistently matched the output from the activity. SQLite file.

Conclusion

The Zeitgeist database stores most of the user activity, with the exception of log in and log off information. Other user activity will all be logged in a separate user database under each user's home folder, which helps separate the activity to ensure you are looking at a specific user's information. The information located in these databases is accurate down to the minute. As there was no way to accurately record seconds, the extent of the research and experimentation was down to the minute level. The database entries were created when the action occurred, which is as expected. Different user activities can be differentiated as they all have varying activity. SQLite databases for each specific user. There does not seem to be a way to see if the user was using terminal or the GUI to launch programs.