MONASH University

MONASH
LIBRARY

2018 ANZREG Conference

EZproxy logs

Proactive security breach management and rich access metrics

GROUP OF EIGHT AUSTRALIA

Linda Farrell

Information Resources Librarian

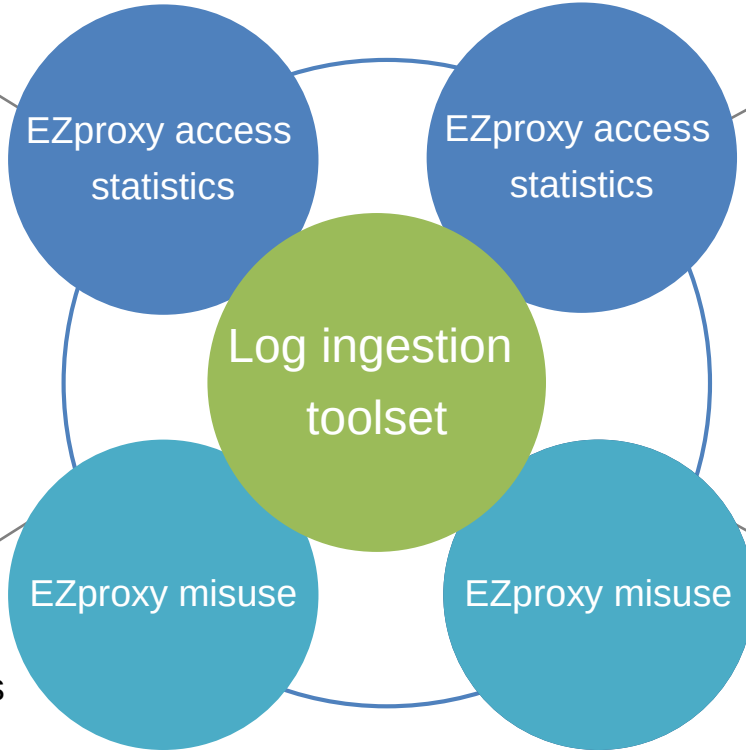
Introduction

Issues

- Maintenance
- Inflated statistics

Solution

- Tool to analyze and visualize log data



Issues

- Excessive downloads
- Compromised accounts

Solution

- Tool for identifying suspicious patterns of behaviour

EZproxy access statistics

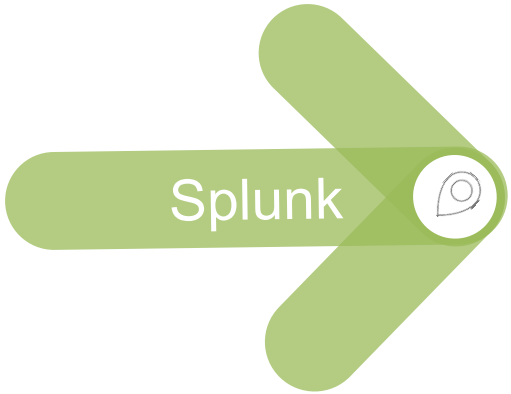
- EZproxy at Monash is locally hosted, administered by the library and IT
- On - and off campus access is through EZproxy where possible
- Monash has always used the EZproxy logs to report on access statistics
- Additional to COUNTER and vendor statistics, for some platforms these are the only statistics available
- In the past, a Python script was used to generate HTML and CSV files

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	**Database title**	**JAN**	**FEB**	**MAR**	**APR**	**MAY**	**JUN**	**JUL**	**AUG**	**SEP**	**OCT**	**NOV**	**DEC**	**Y-T-D**
2	TOTAL SEARCHES	251759	717911	1067543	1357621	1069259	437709	367338	1084747	1666806	940979	721910	223078	9906660
3	academypublisher (ojs.academypublisher.c	2	3	3	1	0	0	0	0	0	0	0	3	12
4	artfilms-digital (www.artfilms-digital.com)	22	0	145	0	0	0	0	0	33	35	37	1	273
5	bmj (bestpractice.bmj.com)	0	0	0	0	11239	0	7009	0	0	8907	0	1	27156
6	cambridge (journals.cambridge.org)	870	1002	3150	5437	3990	1832	1411	0	2763	3074	0	1	23530
7	ceeol (www.ceeol.com)	153	155	1002	1428	1393	498	212	837	1043	988	271	12	7992
8	dbpia (www.dbpia.co.kr)	112	68	355	540	569	207	141	506	376	553	182	80	3689
9	doi (dx.doi.org)	29142	31492	114201	155830	123923	45189	40580	114918	81066	82259	31437	1	850038
10	eblib (www.monash.eblib.com.au)	9149	15620	73786	77624	0	31077	24058	59638	0	40493	15743	1	347189
11	ebscohost (openurl.ebscohost.com)	0	0	86941	0	0	29347	23275	106971	94352	78278	0	6	419170
12	eiu (gfs.eiu.com)	22	0	0	0	0	13	0	45	0	42	0	17	139
13	emeraldinsight (www.emeraldinsight.com)	2887	1500	11471	20742	14911	3228	2464	16608	11993	11388	3806	2	101000
14	factiva (global.factiva.com)	413	696	2473	2356	1784	597	449	2009	1188	752	420	1	13138
15	galegroup (find.galegroup.com)	11434	9781	44256	67768	56249	20249	12519	52404	43506	50620	0	6	368792
16	google (scholar.google.com.au)	6773	9332	0	36435	31406	11796	8423	25832	22757	32609	10082	98	195543
17	ibisworld (clients1.ibisworld.com.au)	463	206	972	2184	2159	579	432	1499	1581	1605	1036	64	12780
18	ieee (ieeexplore.ieee.org)	1306	1539	2240	2729	3399	1472	1581	2885	1885	2764	1367	1	23168
19	igi-global (services.igi-global.com)	290	223	967	1160	1314	346	320	0	780	976	256	1	6633
20	informit (search.informit.com.au)	4687	4951	18373	33843	19899	9911	5763	25091	0	16080	7080	360	146038
21	informs (pubsonline.informs.org)	214	86	215	528	530	116	87	274	405	342	145	2549	5491

EZproxy access statistics

Issues

- Maintenance of Python scripts
- Slow execution of scripts as the logs are getting bigger
- Python libraries no longer supported
- Skewed statistics due to EZproxy misuse



Splunk solution

- Ingestion of logs
- Enrich with faculty data
- Tableau connector
- Improved detection of compromised accounts and excessive downloads

What is Splunk?

- Used for searching, monitoring, and analyzing machine-generated big data, via a web-style interface. Captures, indexes, and correlates real-time data in a searchable repository, from which we can generate graphs, reports, alerts, dashboards, and visualizations.

- Splunk as an analysis tool enables us to create reports and alerts about suspicious patterns of behaviour over a period of time, on which we can act before serious copyright breaches occur.

- Monash University Library implemented Splunk in August 2017

Splunk for EZproxy statistics

column	October 2017	November 2017	December 2017	January 2018	February 2018
sciencedirect		11977	7621	10237	3238
proquest		11450	8908	9263	2232
google		9031	5698	7559	2195
wiley		7474	3947	5278	1635
amh		3723	1516	2895	1345
ovid		5748	2303	4283	1290
tandfonline		5308	3036	4059	1158
ebscohost		8309	8950	6080	1148
scopus		3660	2369	3315	1113
uptodate		3442	834	1368	1084
doi		6033	2399	3256	1001
tg		4581	544	1457	802
webofknowledge		2088	1438	2207	799
bmj		3083	518	975	774
springer		3369	2124	2896	769
westlaw		1309	901	1427	737

EZproxy misuse challenges

Excessive downloads (script or browser plugin)

Compromised accounts (phishing emails)

Procedure at Monash

- Compromised accounts
  - Identify
  - Block account in EZproxy user.txt
  - Inform IT security
  - IT security communicates with user (education)
  - Account password is reset
  - Unblock when completed by IT security
- Excessive downloads
  - Block account in EZproxy user.txt
  - Email user (phone staff member where possible)
  - Unblock when satisfactory reply is received via email

EZproxy misuse challenges

Before splunk>

- Monitoring event logs once or twice a day – very limited insight

Audit Events for 2018-05-13

Date/Time	Event	IP	Location	Username	Session	Other
04:55:04	Login.Success	1.9.247.150	MY 05 Pantai		ZUWM4LZVSKhbJvJ	
06:05:26	Logout				ZUWM4LZVSKhbJvJ	Expired
13:01:20	Login.Success	185.89.248.5	TR 34 Istanbul		uvVrQ1cEi4eEGSt	
13:32:26	Logout				uvVrQ1cEi4eEGSt	Expired
14:38:59	Login.Success	118.139.142.40	AU 05 Monash		KFywL9dHbmmM935	
15:09:26	Logout				KFywL9dHbmmM935	Expired
15:38:26	Login.Success	118.139.142.40	AU 05 Monash		4A6gGA36FRBQ7jH	
16:08:56	Logout				4A6gGA36FRBQ7jH	Expired
19:00:26	Login.Success	185.89.248.5	TR 34 Istanbul		6fxr57jmKL51BKc	
19:34:56	Logout				6fxr57jmKL51BKc	Expired

- Investigated logs after email from vendor about block – time consuming and long delays because of time differences
- Python script to monitor downloads – problematic to maintain especially as parameters have shifted over the years

Splunk for security

EZproxy user access from more than one country or IP (last 7 days)

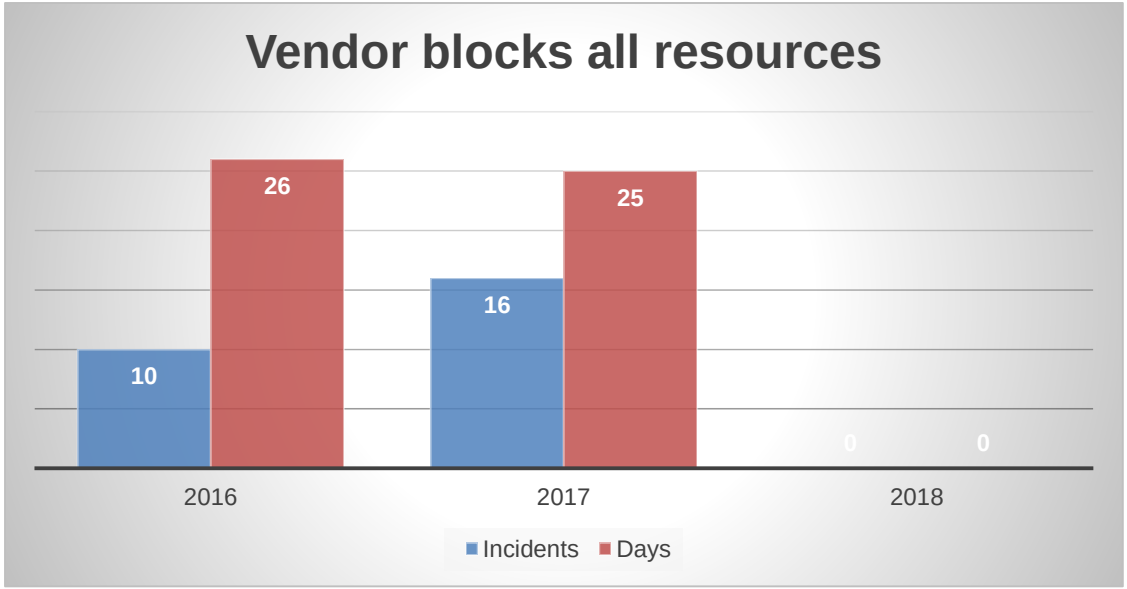
user	CountryCount	Countries	IPCount	IPs
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	3	Australia Malaysia Thailand	13	1.47.109.47 118.139.142.100 180.183.193.62 183.89.82.115 202.28.50.200 61.6.11.88 61.6.121.79 61.6.123.89 61.6.152.198 61.6.183.63 61.6.24.43 61.6.83.64 61.6.9.177
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	3	Australia Netherlands United States	9	103.55.47.195 104.156.210.194 118.138.190.143 130.194.237.41 130.194.237.87 188.226.174.200 198.23.71.92 49.127.42.240 49.127.55.234

Splunk for security

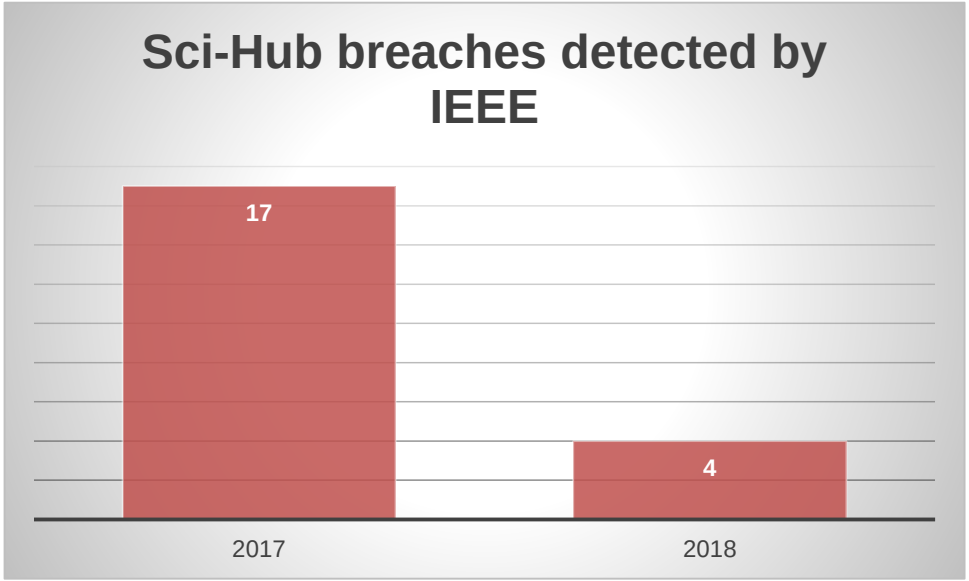
Total downloads by users (last 24 hours)

user	TotalDownload	totalGB
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	616223267	0.5739026395604014
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	604971758	0.5634238552302122
xxxxxxxxxxx	521829527	0.48599161859601736
xxxxxxxxxxxxxxxxxxxxxxxxxxx	350348845	0.3262877883389592
xxxxxxxxxxxxxxxxxxxxxxxxxxx	340297883	0.3169271005317569
xxxxxxxxxxxxxxxxxxxxxxx	240878751	0.22433581855148077
xxxxxxxxxxxxxxxxxxxxxxx	221792927	0.20656075980514288
xxxxxxxxxxxxxxxxxxxxxxx	197421975	0.18386354204267263
xxxxxxxxxxxxxxxxxxxxxxx	190342814	0.17727055959403515
xxxxxxxxxxxxxxxxxxxxxxx	186995592	0.17415321618318558
xxxxxxxxxxxxxxxxxxxxxxxxxxx	181116428	0.16867781803011894
xxxxxxxxxxxxxxxxxxxxxxxxxxx	177938781	0.1657184036448598
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	176031206	0.16394183598458767

Success with Splunk



Success with Splunk



By end of May 2018

- Activity on IEEE detected once in January and May, and twice in February, but avoided downtime because of pre-emptive action (compared to 9 Sci-Hub incidents by May 2017)
- IEEE is seeing an increase in activity worldwide, plus an increase the number of universities targeted. Monash is experiencing a decrease

Success with Splunk

- Increased accuracy of access statistics
- Increased efficiency in identifying compromised accounts
- Improved relationships with providers e.g. IEEE
- Decrease in downtime due to misuse
- Protection of staff and student accounts and identity

The average rate for identifying compromised accounts was 4 per week - 11 accounts identified in the first 3 days of using Splunk

Conclusion

- Thank you
- Contact : linda.farrell@monash.edu
- Questions