TITEL

DIE NSRL

CHAOSTREFF LUDWIGSBURG

DATUM 08.04.2020

AUTHOR STEFFEN / AMPOFF

info(ampoff)

- Steffen
- Software Engineer
- Ich mag Apfelkuchen, spiel Gitarre in einer Band und übe mich in Krav Maga

Eine Herausforderung

Was liegt da rum?!

Ist das wichtig?

Kann das weg?

Quelle: https://phdlife.warwick.ac.uk/2014/03/14/chaos-creativity-and-data-a-clear-connection/

Beispiele

- Datenträgerrettung
- Digitale Vor-/Nachlässe
- Undisziplinierte Datenablage

NSRL to the rescue!

- NSRL: National Software Reference Library
- Software-Korpus

Quellen: Internet

NSRL?!

- Ausgangsbasis Software-Korpus:
  - Anwendungssoftware, OS, Libraries
  - Metadaten: Hashsummen, Manufacturer, Filesize, OS-Link, …
  - Keine verbotenen Nutzdaten, e.g. Kinderpornographie, Terrorpropaganda
- Informationen, wie z.B. Hashsummen oder Hersteller, werden in *Reference Data Sets* gespeichert
- RDS als zip oder ISO kostenlos und ohne Registrierung als Download verfügbar

NSRL?

- RDS-Typen
  - Modern RDS
    - Software ab 2000; mit (sehr vielen) Dubletten
  - Modern RDS Minimal
    - Software ab 2000; Dubletten entfernt
  - Modern RDS Unique
    - Software ab 2000; nur Einträge, die keine Dubletten besitzen
  - RDS Legacy
    - Software vor 2000
  - Android RDS
  - iOS RDS

NSRL…

```
[[steffen@intern ~/nsrl_minimal_2020-04-04]$ ls
NSRLFile.txt            NSRLOS.txt              RDS_HashCounts.txt      rds_modernm.zip.sha     redis_feed.txt
NSRLMfg.txt             NSRLProd.txt            rds_modernm.zip         README.txt
[steffen@intern ~/nsrl_minimal_2020-04-04]$
```

NSRL…

```
[[steffen@intern ~/nsrl_minimal_2020-04-04]$ ls
NSRLFile.txt            NSRLOS.txt              RDS_HashCounts.txt      rds_modernm.zip.sha     redis_feed.txt
NSRLMfg.txt             NSRLProd.txt            rds_modernm.zip         README.txt
[steffen@intern ~/nsrl_minimal_2020-04-04]$
```

```
[[steffen@intern ~/nsrl_minimal_2020-04-04]$ head RDS_HashCounts.txt
RDS 2.68 March 2020 Hash Counts

Modern:                 104,196,072
Modern (minimal):        26,521,022
Modern (unique):         13,481,033
Legacy:                 107,297,738
Android:                 13,476,831
iOS:                     14,390,472
[steffen@intern ~/nsrl_minimal_2020-04-04]$
```

NSRL!

```
[steffen@intern ~/nsrl_minimal_2020-04-04]$ head NSRLFile.txt
"SHA-1","MD5","CRC32","FileName","FileSize","ProductCode","OpSystemCode","SpecialCode"
"00000079FD7AAC9B2F9C988C50750E1F50B27EB5","8ED4B4ED952526D89899E723F3488DE4","7A5407CA","wow64_microsoft-windows-i..timezones.resour
"000000F694CA9BF73836D67DEB5E2724338B422D","497C460BBA43530494F37DF7DE3A5FF4","46B80AC7","bpa10x.ko",12944,17066,"362",""
"000001BB80E9C6F9CACB6DA82F4D2E3266B9C4C3","3491EE38124BF5382D0828C5209C83B5","6CC040F2","Batman_Seventies.POR",90,196184,"362",""
"0000034F77D9314B1B94DBDA3031BECE1198D067","FE330C56554EF007D38C89764864E365","71C6F991","arm64_49016ecbe73216140477e3b16492e87f_31bf
"000006CD8F63343893C52830FC36118124131E25","41D0DD202B31F022CDB92802567058A5","7AD24105","redbull.erp",8663417,201453,"362",""
"000006E81C829F654163696578D9B1841E8CE167","3F4894B0A067111BC862608E3B6D6205","21AB6D9F","dy3246416.htm",11366,14693,"362",""
"000007B928F4C211CC8ED3C9707196A7C5BA3AF8","68563E2BFC732E10E885BD2DCF49F2EF","34940E24","microsoft-windows-businessscanning-feature-
"00000903319A8CE18A03DFA22C07C6CA43602061","6E2F8616A01725DCB37BED0A2495AEB2","E774FD92","network",7279,182360,"362",""
"0000094B6A7FF7B386E14DE2049478BEA024D206","99001D28B85087B80E5E24B0CA2D15CB","4FF6CD7A","__versions",1984,15116,"362",""
[steffen@intern ~/nsrl_minimal_2020-04-04]$
```

NSRL!

```
[steffen@intern ~/nsrl_minimal_2020-04-04]$ head NSRLFile.txt
"SHA-1","MD5","CRC32","FileName","FileSize","ProductCode","OpSystemCode","SpecialCode"
"00000079FD7AAC9B2F9C988C50750E1F50B27EB5","8ED4B4ED952526D89899E723F3488DE4","7A5407CA","wow64_microsoft-windows-i..timezones.resources_31bf3856ad364e35_10.0.16299.579_de-de_f24979c73226184d.manifest",2520,190718,"362",""
"000000F694CA9BF73836D67DEB5E2724338B422D","497C460BBA43530494F37DF7DE3A5FF4","46B80AC7","bpa10x.ko",12944,17066,"362",""
"000001BB80E9C6F9CACB6DA82F4D2E3266B9C4C3","3491EE38124BF5382D0828C5209C83B5","6CC040F2","Batman_Seventies.POR",90,196184,"362",""
"0000034F77D9314B1B94DBDA3031BECE1198D067","FE330C56554EF007D38C89764864E365","71C6F991","arm64_49016ecbe73216140477e3b16492e87f_31bf3856ad364e35_10.0.17134.81_none_ae8f44b72b46370a.manifest",705,188511,"362",""
"000006CD8F63343893C52830FC36118124131E25","41D0DD202B31F022CDB92802567058A5","7AD24105","redbull.erp",8663417,201453,"362",""
"000006E81C829F654163696578D9B1841E8CE167","3F4894B0A067111BC862608E3B6D6205","21AB6D9F","dy3246416.htm",11366,14693,"362",""
"000007B928F4C211CC8ED3C9707196A7C5BA3AF8","68563E2BFC732E10E885BD2DCF49F2EF","34940E24","microsoft-windows-businessscanning-feature-package~31bf3856ad364e35~amd64~pt-br~6.1.7601.17514.mum",1541,201424,"362",""
"00000903319A8CE18A03DFA22C07C6CA43602061","6E2F8616A01725DCB37BED0A2495AEB2","E774FD92","network",7279,182360,"362",""
"0000094B6A7FF7B386E14DE2049478BEA024D206","99001D28B85087B80E5E24B0CA2D15CB","4FF6CD7A","__versions",1984,15116,"362",""
[steffen@intern ~/nsrl_minimal_2020-04-04]$
```

```
[steffen@intern ~/nsrl_minimal_2020-04-04]$ grep 190718 NSRLProd.txt
190718,"Cumulative Update for Windows Server 2016 (1709) for x64-based Systems (KB4457136)","Sept. 26, 2018","189","608","Multilanguage","Security,Update"
190718,"Cumulative Update for Windows Server 2016 (1709) for x64-based Systems (KB4457136)","Sept. 26, 2018","867","608","Multilanguage","Security,Update"
190718,"Cumulative Update for Windows Server 2016 (1709) for x64-based Systems (KB4457136)","Sept. 26, 2018","868","608","Multilanguage","Security,Update"
[steffen@intern ~/nsrl_minimal_2020-04-04]$
```

Zurück zum Problem…

Welche Datei ist für mich interessant?

Ist sie in einem RDS, dann nicht!

Aber…

Bei 100.000 Dateien 100.000 grep auf ein 3GB File?

Unschön!

Redis? Redis!

NoSQL

Key-Value-Store

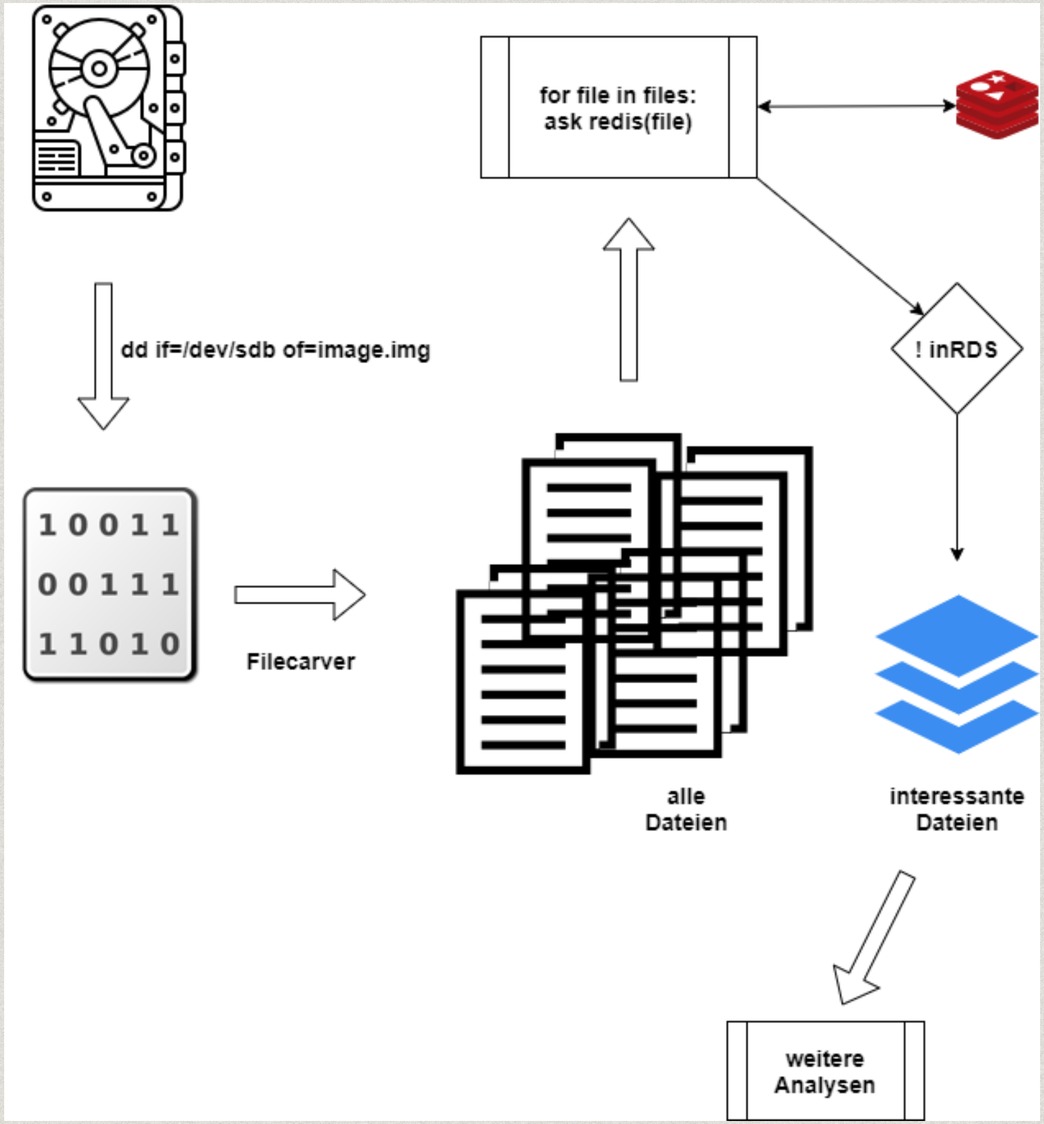
in-memory

BSD-3-Clause-Lizenz

KEY	VALUE
da39a3…afd80709	TRUE
baf550…1936ac	TRUE
109f4f9…066971f	TRUE
…	TRUE

$ redis-cli get 000006E81C829F654163696578D9B1841E8CE167

Beispiel-Workflow



Links

- NSRL: https://www.nist.gov/itl/ssd/software-quality-group/national-software-reference-library-nsrl
- Redis: https://redis.io
- nslredis: https://github.com/steffenfritz/nslredis