

Large scale PCAP Analysis using Apache Hadoop

Wolfgang Nagele
Global Information Infrastructure Manager



We do big data ...

- K-root (15.000 qps)
 - 1.5TB of compressed PCAP data every month
 - And that is only queries
- F-reverse (6.000 qps)
- AS112 (2.000 qps)
- Auth DNS (26.000 qps)
- RIS (BGP updates from back in 2000 onwards)
- You get the idea ...

Why not libtrace, PacketQ, <you name it>

- Vertical scaling does not work for terabytes of data
- Running those tools in parallel is hard
 - This is what Hadoop is good at

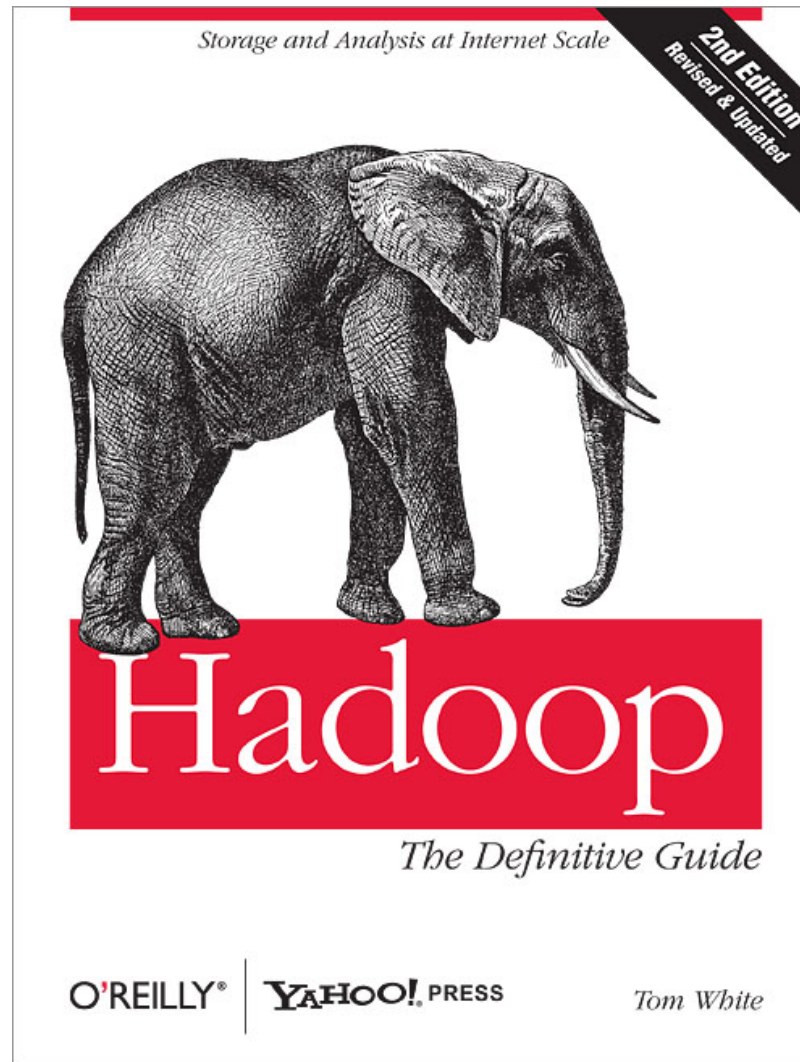
What is HDFS?

- Open-source implementation of Google Filesystem (GFS) as detailed in a whitepaper
 - <http://labs.google.com/papers/gfs-sosp2003.pdf>
- Hadoop Distributed Filesystem (HDFS)
 - Namenode holding filesystem registry
 - Datanodes holding filesystem blocks

What is MapReduce?

- Another whitepaper from Google:
 - <http://labs.google.com/papers/mapreduce-osdi04.pdf>
- Essentially: A programming pattern
 - Allows distribution of large computational tasks

Start with a good read ...



Native PCAP reading in Java

- Open source under the LGPL

- Available at:

<http://github.com/RIPE-NCC/hadoop-pcap>

Live Demo: The data

```
wnagele@bastion1:~$ hadoop fs -du /datasets/k-root-pcap-attack-201106/
Found 19 items
87371727800 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/ams-ix
6731756482 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/apnic
3046072188 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/bix
3571662900 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/cern
773 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/copy.sh
23731100686 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/delhi
103503797822 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/denic
935385937 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/emix
2348458368 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/ficix
3758556675 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/grnet
461275326 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/isnic
152632258729 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/linx
16702579240 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/mix
99042920194 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/nap
897337586 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/nskix
616068024 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/poznan
991508061 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/qtel
181967770 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/tix
127280923995 hdfs://namenode1.hadoop.ripe.net/datasets/k-root-pcap-attack-201106/tokyo
wnagele@bastion1 ~]$
```

590GB total

Live Demo: Create table

```
wnagele@bastion1:~  
[wnagele@bastion1 ~]$ hive  
Hive history file=/tmp/wnagele/hive_job_log_wnagele_201110152104_869372884.txt  
hive> CREATE EXTERNAL TABLE pcaps (ts bigint,  
> protocol string,  
> src string,  
> src_port int,  
> dst string,  
> dst_port int,  
> len int,  
> ttl int,  
> dns_queryid int,  
> dns_flags string,  
> dns_opcode string,  
> dns_rcode string,  
> dns_question string,  
> dns_answer array<string>,  
> dns_authority array<string>,  
> dns_additional array<string>)  
> PARTITIONED BY (node string)  
> ROW FORMAT SERDE 'net.ripe.hadoop.pcap.serde.PcapDeserializer'  
> STORED AS INPUTFORMAT 'net.ripe.hadoop.pcap.io.DnsPcapInputFormat'  
> OUTPUTFORMAT 'org.apache.hadoop.hive ql.io.HiveIgnoreKeyTextOutputFormat';  
OK  
Time taken: 4.889 seconds  
hive> █
```

Live Demo: Add partitions

```
wnagele@bastion1:~  
Time taken: 0.27 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='bix') LOCATION '/datasets/k-root-pcap-attack-201106/bix/';  
OK  
Time taken: 0.362 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='cern') LOCATION '/datasets/k-root-pcap-attack-201106/cern/';  
OK  
Time taken: 0.211 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='delhi') LOCATION '/datasets/k-root-pcap-attack-201106/delhi/';  
OK  
Time taken: 0.284 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='denic') LOCATION '/datasets/k-root-pcap-attack-201106/denic/';  
OK  
Time taken: 0.856 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='emix') LOCATION '/datasets/k-root-pcap-attack-201106/emix/';  
OK  
Time taken: 0.307 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='ficix') LOCATION '/datasets/k-root-pcap-attack-201106/ficix/';  
OK  
Time taken: 0.164 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='grnet') LOCATION '/datasets/k-root-pcap-attack-201106/grnet/';  
OK  
Time taken: 0.215 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='isnic') LOCATION '/datasets/k-root-pcap-attack-201106/isnic/';  
OK  
Time taken: 0.28 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='linx') LOCATION '/datasets/k-root-pcap-attack-201106/linx/';  
OK  
Time taken: 0.269 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='mix') LOCATION '/datasets/k-root-pcap-attack-201106/mix/';  
OK  
Time taken: 0.287 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='nap') LOCATION '/datasets/k-root-pcap-attack-201106/nap/';  
OK  
Time taken: 0.173 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='nskix') LOCATION '/datasets/k-root-pcap-attack-201106/nskix/';  
OK  
Time taken: 0.309 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='poznan') LOCATION '/datasets/k-root-pcap-attack-201106/poznan/';  
OK  
Time taken: 0.157 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='qtel') LOCATION '/datasets/k-root-pcap-attack-201106/qtel/';  
OK  
Time taken: 0.263 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='tix') LOCATION '/datasets/k-root-pcap-attack-201106/tix/';  
OK  
Time taken: 0.224 seconds  
hive> ALTER TABLE pcaps ADD PARTITION (node='tokyo') LOCATION '/datasets/k-root-pcap-attack-201106/tokyo/';  
OK  
Time taken: 0.152 seconds  
hive> █
```

Live Demo: Run query

```
[screen 0: bash] wnagele@bastion1:~  
[wnagele@bastion1 ~]$ hive  
Hive history file=/tmp/wnagele/hive_job_log_wnagele_201110152123_1839526960.txt  
hive> SELECT COUNT(1) AS queries FROM pcaps WHERE dns_question LIKE '%[REDACTED]%' AND node='isnic';  
Total MapReduce jobs = 1  
Launching Job 1 out of 1  
Number of reduce tasks determined at compile time: 1  
In order to change the average load for a reducer (in bytes):  
  set hive.exec.reducers.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducers.max=<number>  
In order to set a constant number of reducers:  
  set mapred.reduce.tasks=<number>  
Starting Job = job_201110081356_0029, Tracking URL = http://namenode1.hadoop.ripe.net:50030/jobdetails.jsp?jobid=job_201110081356_0029  
Kill Command = /usr/lib/hadoop/bin/hadoop job -Dmapred.job.tracker=namenode1.hadoop.ripe.net:9001 -kill job_201110081356_0029  
2011-10-15 21:24:39,813 Stage-1 map = 0%, reduce = 0%  
2011-10-15 21:24:44,854 Stage-1 map = 1%, reduce = 0%  
2011-10-15 21:24:45,864 Stage-1 map = 2%, reduce = 0%  
2011-10-15 21:24:48,886 Stage-1 map = 3%, reduce = 0%  
2011-10-15 21:24:49,896 Stage-1 map = 4%, reduce = 0%  
2011-10-15 21:24:52,998 Stage-1 map = 5%, reduce = 0%  
2011-10-15 21:24:54,008 Stage-1 map = 6%, reduce = 0%  
2011-10-15 21:24:55,025 Stage-1 map = 7%, reduce = 0%  
2011-10-15 21:24:59,059 Stage-1 map = 9%, reduce = 0%  
2011-10-15 21:25:03,132 Stage-1 map = 10%, reduce = 0%  
2011-10-15 21:25:04,142 Stage-1 map = 11%, reduce = 0%  
2011-10-15 21:25:07,171 Stage-1 map = 12%, reduce = 0%  
2011-10-15 21:25:08,182 Stage-1 map = 13%, reduce = 0%  
2011-10-15 21:25:09,195 Stage-1 map = 13%, reduce = 4%  
2011-10-15 21:25:11,216 Stage-1 map = 14%, reduce = 4%  
2011-10-15 21:25:12,226 Stage-1 map = 15%, reduce = 4%  
2011-10-15 21:25:13,237 Stage-1 map = 16%, reduce = 4%  
2011-10-15 21:25:16,320 Stage-1 map = 17%, reduce = 4%  
2011-10-15 21:25:17,331 Stage-1 map = 18%, reduce = 4%  
2011-10-15 21:25:20,364 Stage-1 map = 19%, reduce = 6%  
2011-10-15 21:25:21,376 Stage-1 map = 20%, reduce = 6%  
2011-10-15 21:25:24,411 Stage-1 map = 21%, reduce = 6%  
2011-10-15 21:25:26,460 Stage-1 map = 22%, reduce = 6%  
2011-10-15 21:25:28,483 Stage-1 map = 23%, reduce = 6%  
2011-10-15 21:25:29,497 Stage-1 map = 23%, reduce = 7%  
2011-10-15 21:25:30,511 Stage-1 map = 24%, reduce = 7%  
2011-10-15 21:25:32,537 Stage-1 map = 25%, reduce = 7%  
2011-10-15 21:25:34,563 Stage-1 map = 26%, reduce = 7%
```

See if we received target traffic at Reykjavík instance

Live Demo: Query in progress

Hadoop job_201110081356_0029 on namenode1

Hadoop job_201110081356_00...

namenode1.hadoop.ripe.net:50030/jobdetails.jsp?jobid=job_201110081356_0029&refresh=30

TeuxDeux Mail Reader Bookmark on De... Bookmarks

Job Counters	Launched reduce tasks	0	0	1
	Launched map tasks	0	0	721
	Data-local map tasks	0	0	721
org.apache.hadoop.hive.ql.exec.FilterOperator\$Counter	PASSED	0	0	0
	FILTERED	3,648,360	0	3,648,360
FileSystemCounters	HDFS_BYTES_READ	122,484,142	0	122,484,142
	FILE_BYTES_WRITTEN	50,019,512	0	50,019,512
org.apache.hadoop.hive.ql.exec.MapOperator\$Counter	DESERIALIZE_ERRORS	0	0	0
Map-Reduce Framework	Combine output records	0	0	0
	Map input records	3,648,360	0	3,648,360
	Spilled Records	631	0	631
	Map output bytes	6,941	0	6,941
	Map input bytes	121,954,032	0	121,954,032
	SPLIT_RAW_BYTES	174,156	0	174,156
	Map output records	631	0	631
	Combine input records	0	0	0

Map Completion Graph - close

Reduce Completion Graph - close

Live Demo: Result

```
[screen 0: bash] wnagele@bastion1:~  
2011-10-15 21:26:56,965 Stage-1 map = 68%, reduce = 21%  
2011-10-15 21:26:59,017 Stage-1 map = 69%, reduce = 21%  
2011-10-15 21:27:00,043 Stage-1 map = 69%, reduce = 22%  
2011-10-15 21:27:01,075 Stage-1 map = 70%, reduce = 22%  
2011-10-15 21:27:03,157 Stage-1 map = 71%, reduce = 22%  
2011-10-15 21:27:05,209 Stage-1 map = 72%, reduce = 22%  
2011-10-15 21:27:06,234 Stage-1 map = 73%, reduce = 23%  
2011-10-15 21:27:08,291 Stage-1 map = 74%, reduce = 23%  
2011-10-15 21:27:10,375 Stage-1 map = 75%, reduce = 23%  
2011-10-15 21:27:12,430 Stage-1 map = 76%, reduce = 24%  
2011-10-15 21:27:14,482 Stage-1 map = 77%, reduce = 24%  
2011-10-15 21:27:15,511 Stage-1 map = 78%, reduce = 25%  
2011-10-15 21:27:17,581 Stage-1 map = 79%, reduce = 26%  
2011-10-15 21:27:19,706 Stage-1 map = 80%, reduce = 26%  
2011-10-15 21:27:21,764 Stage-1 map = 81%, reduce = 26%  
2011-10-15 21:27:23,826 Stage-1 map = 82%, reduce = 26%  
2011-10-15 21:27:24,856 Stage-1 map = 83%, reduce = 26%  
2011-10-15 21:27:26,918 Stage-1 map = 84%, reduce = 27%  
2011-10-15 21:27:29,009 Stage-1 map = 85%, reduce = 27%  
2011-10-15 21:27:31,069 Stage-1 map = 86%, reduce = 27%  
2011-10-15 21:27:33,132 Stage-1 map = 87%, reduce = 28%  
2011-10-15 21:27:34,165 Stage-1 map = 88%, reduce = 28%  
2011-10-15 21:27:36,223 Stage-1 map = 89%, reduce = 29%  
2011-10-15 21:27:38,284 Stage-1 map = 90%, reduce = 29%  
2011-10-15 21:27:40,372 Stage-1 map = 91%, reduce = 29%  
2011-10-15 21:27:41,405 Stage-1 map = 92%, reduce = 29%  
2011-10-15 21:27:42,442 Stage-1 map = 92%, reduce = 30%  
2011-10-15 21:27:43,473 Stage-1 map = 93%, reduce = 30%  
2011-10-15 21:27:45,542 Stage-1 map = 94%, reduce = 31%  
2011-10-15 21:27:47,602 Stage-1 map = 95%, reduce = 31%  
2011-10-15 21:27:48,636 Stage-1 map = 96%, reduce = 31%  
2011-10-15 21:27:51,429 Stage-1 map = 97%, reduce = 31%  
2011-10-15 21:27:52,463 Stage-1 map = 98%, reduce = 31%  
2011-10-15 21:27:54,555 Stage-1 map = 99%, reduce = 32%  
2011-10-15 21:27:55,589 Stage-1 map = 100%, reduce = 32%  
2011-10-15 21:28:00,781 Stage-1 map = 100%, reduce = 33%  
2011-10-15 21:28:01,818 Stage-1 map = 100%, reduce = 100%  
Ended Job = job_201110081356_0029  
OK  
0  
Time taken: 221.317 seconds  
hive> |
```

No target traffic at
ISNIC instance

Live Demo: Conclusions

- Works well at scale
 - 100+ CPU cores
- High processing overhead
 - Example took 200 seconds total
 - Only 50% of it spent on actual computation
 - Small input files (only 500MB total)
 - See Screencast – 80GB in 3 minutes

Step by Step Screencast

- Zero setup using Amazon EC2



<http://goo.gl/8uvlX>

Questions?

wnagele@ripe.net

