F

A complete system integration of stream-based IP flow-record querier

VAIBHAV BAJPAI

Masters Thesis

School of Engineering and Science
Jacobs University Bremen
Bremen, Germany

June 2012

ABSTRACT

Short summary of the contents in English…

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

IPFIX Internet Protocol Flow Information Export

HDF Hierarchical Data Format

LALR Look-Ahead LR Parser

PLY Python Lex-Yacc

HDFS Hadoop Distributed File System

API Application Programming Interface

Part I

INTRODUCTION

You can put some informational part preamble text here

1 TRAFFIC MEASUREMENT APPROACHES

1.1 CAPTURING PACKETS

1.2 CAPTURING FLOWS

1.3 REMOTE MONITORING

1.4 REMOTE METERING

2 FLOW EXPORT PROTOCOLS

2.1 NETFLOW

2.2 IPFIX

2.3 SFLOW

3 LANGUAGES AND TOOLS

3.1 SQL-BASED QUERY LANGUAGES

3.1.1 *NetFlow exports as relational DBMS*

3.1.2 *Data Stream Management System*

3.1.3 *Gigascope*

3.1.4 *Tribeca*

3.2 FILTERING LANGUAGES

3.2.1 *flow-tools*

3.2.2 *nfdump*

3.3 PROCEDURAL LANGUAGES

3.3.1 *FlowScan*

3.3.2 *Clustering NetFlow Exports*

3.3.3 *SiLK Analysis Suite*

4 LEGAL CONSIDERATION

Part II

STATE OF THE ART

You can put some informational part preamble text here

5 FLOWY

Flowy [3][4] is the first prototype implementation of a stream-based flow record query language [5][1][6]. The query language allows to describe patterns in flow-records in a declarative and orthogonal fashion, making it easy to read and flexible enough to describe complex relationships among a given set of flows.

5.1 PYTHON FRAMEWORK

Flowy is written in Python. The framework is subdivided into two main modules: the validator module and the execution module. The validator module is used for syntax checking and interconnecting of all the stages of the processing pipeline and the execution module is used to perform actions at each stage of the runtime operation.

5.1.1 *PyTables and PLY*

Flowy uses PyTables [7] to store the flow-records. PyTables is built on top of the Hierarchical Data Format (HDF) library and can exploit the hierarchical nature of the flow-records to efficiently handle large amounts of flow data. The `pytables` module provides methods to read/write to PyTables files. The `FlowRecordsTable` class instance within the module exposes an iterator interface over the records stored in the HDF file. The `GroupsExpander` class instance within the same module on the other hand exposes an iterator interface over the group records and facilitates ungrouping to flow records.

In addition, Flowy uses Python Lex-Yacc (PLY) for generating a Look-Ahead LR Parser (LALR) parser and providing extensive input validation, error reporting and validation on the execution modules.

5.1.2 *Records*

Flow-records are the principal unit of data exchange throughout Flowy's processing pipeline. The prototype implementation allows the `Record` class (defined in the `record` module) to be dynamically generated using `get_record_class(...)` allowing future implementations to easily plug in support for Internet Protocol Flow Information Export (IPFIX) or even newer versions of NetFlow [8] exports. The `FlowToolsReader` class instance (defined in `ftreader` module) provides an iterator over the records defined in `flow-tools` format.

This can be plugged into the `RecordReader` class instance (defined in `record` module) to instantly get `Record` class instances.

5.1.3 *Parsers and Statements*

The `parser` module holds definitions for the lexer and parser. The statements when parsed are implicitly converted into instances of classes defined in the `statement` module. The instances contain meta-information about the parsed statement such as the values, line numbers and sub-statements (if any).

5.2 PROCESSING PIPELINE

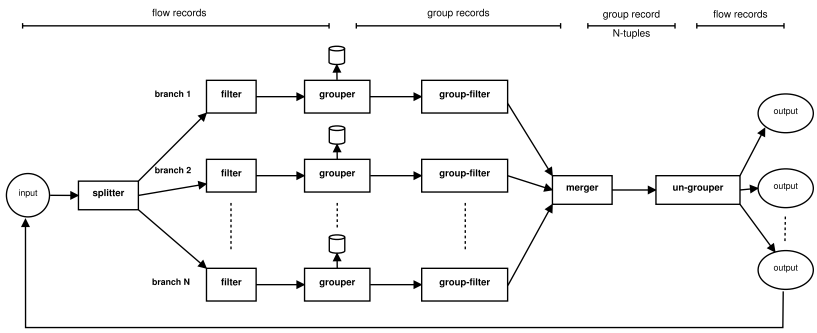

Figure 1: Flowy: Processing Pipeline [1]

The pipeline consists of a number of independent processing elements that are connected to one another using UNIX-based pipes. Each element receives the content from the previous pipe, performs an operation and pushes it to the next element in the pipeline. Figure 1 shows an overview of the processing pipeline. The flow record attributes used in this pipeline exactly correlate with the attributes defines in the IPFIX Information Model specified in RFC 5102 [9]. A complete description on the semantics of each element in the pipeline can be found in [5]

5.2.1 *Splitter*

The splitter takes the flow-records data as input in the `flow-tools` compatible format. It is responsible to duplicate the input data out to several branches without any processing whatsoever. This allows each of the branches to have an identical copy of the flow data to process it independently.

5.2.1.1 *Splitter Implementation*

The splitter module handles the duplication of the Record instances to separate branches. Instead of duplicating each flow-record to every branch (as specified in the specification), the implementation follows a pragmatic approach by filtering the records beforehand against all the defined filter rules to determine which branches a flow-record might end up in and saves this information in a record-mask tuple of boolean flags. The go(...) method in the Splitter class then iterates over all the (record, record-mask) pairs to dispatch the records to corresponding branches marked by their masks using the split(...) method. The class uses branch names to branch objects mapping to achieve the dispatch.

5.2.1.2 *Splitter Validator*

The splitter_validator module handles the splitter processing stage. The SplitterValidator class within the module uses the Parser and FilterValidator instances passed to it to create a Splitter instance and its child Branch instances.

5.2.2 *Filter*

The filter performs *absolute* filtering on the input flow-records data. The flow-records that pass the filtering criterion are forwarded to the grouper, the rest of the flow-records are dropped. The filter compares separate fields of a flow-record against either a constant value or a value on a different field of the *same* flow-record. The filter cannot *relatively* compare two different incoming flow-records

5.2.2.1 *Filter Implementation*

The filter module handles the filtering stage of the pipeline. Since in the implementation the filtering stage occurs before the splitting stage, a single Filter class instance suffices for all the branches. Within the filter module, each filtering statement is converted into a Rule class instance, against which the flow-records are matched. The Rule instances are constructed using the (branch mask, logical operator, arguments) tuple. After matching the records against the rules, the record's branch mask is set and is then used by the splitter to dispatch the records to the filtered branches.

5.2.2.2 *Filter Validator*

The filter_validator module handles the filter processing stage. The FilterValidator class within the module uses the Parser instance passed to it to create a Filter instance once the check on semantical constraints have passed. The constraints involve checking whether records fields referenced in the filter definition exist, whether filters references in composite filter definitions exist and whether duplicate filter definitions are defined.

5.2.3 *Grouper*

The grouper performs aggregation of the input flow-records data. It consists of a number of rule modules that correspond to a specific subgroup. A flow-record in order to be a part of the group should be a part of at-least one subgroup. A flow-record can be a part of multiple subgroups within a group. In addition a flow-record cannot be part of multiple groups. The grouping rules can be either absolute or relative. The newly formed groups which are passed on to the group filter can also contain meta-information about the flow-records contained within the group using the aggregate clause defined as part of the grouper query.

5.2.3.1 *Grouper Implementation*

The `grouper` module handles the grouping of flow-records data. The `Group` class instance contains group-record's field information required for absolute filtering. It also contains the first and last records of the group required for relative filtering of the group-records. The `AggrOp` class instance handles the aggregation of group-records. The allowed aggregation operations are defined in `aggr_operators` module. Custom-defined aggregation operations are also supported using `-aggr-import` command line argument.

5.2.3.2 *Grouper Validator*

The `grouper_validator` module handles the grouper processing stage. The `GrouperValidator` class within the module uses the `Parser` and `SplitterValidator` instances passed to it to create a `Grouper` instance once the check on semantical constraints such as the presence of referenced names and non-duplicate names have passed. Three aggregation operations: `union(rec_id)`, `min(stime)`, `max(etime)` are added by default to each `Grouper` instance.

5.2.4 *Group-Filter*

The group-filter performs *absolute* filtering on the input group-records data. The group-records that pass the filtering criterion are forwarded to the merger, the rest of the group-records are dropped. The group-filter compares separate fields (or aggregated fields) of a flow-record against either a constant value or a value on a different field of the *same* flow-record. The group-filter cannot *relatively* compare two different incoming group-records

5.2.4.1 *Group-Filter Implementation*

The `groupfilter` module handles the filtering of group-records. The `GroupFilter` class within the module iterates over the flow-records within the group and applies filtering rules across them. The filtering rules reuse the `Rule` class from the `filter` module. The flow-records are then added to the time index and stored in a pytables file for further processing. For groups that do *not* have a group-filter defined for them, run through a `AcceptGroupFilter` class instance.

The `timeindex` module handles the mapping of the time intervals to the flow-records. The time index is used by the merger stage to learn about the records that satisfy the Allen relations. The `add(...)` method in the `TimeIndex` class is used to add new records to the time index. The `get_interval_records(...)` method on the other hand is used to retrieve records within a particular time interval.

5.2.4.2 *Group-Filter Validator*

The `groupfilter_validator` module handles the group-filter processing stage. The `GroupFilterValidator` class within the module uses the `Parser` and `Grouper` instances passed to it to create a `GroupFilter` instance. The check for the referenced fields is performed against the aggregate clause defined in grouper statements. The class instance uses the `AcceptGroupFilter` instance in case a branch does *not* have a group filter defined for it.

5.2.5 *Merger*

The merger performs relative filtering on the N-tuples of groups formed from the N stream of groups passed on from the group-filter as input. The merger rule module consists of a number of a submodules, where the output of the merger is the set difference of the output of the first submodule with the union of the output of the rest of the submodules. The relative filtering on the groups are applied to express timing and concurrency constraints using Allen interval algebra [10]

5.2.5.1 *Merger Implementation*

The `merger` module handles the merging of stream of groups passed as input. It is implemented as a nested branch loop organized in an alphabetical order where every branch is a separate `for-loop` over its records. During iteration, each branch loop executes the rules that matches the arguments defined in the group record tuple and subsequently passes them to the lower level for further processing. The `Merger` class represents the highest level branch loop and as such it must iterate over all of its records since it does not have any rules to impose restrictions on the possible records. The `MergerBranch` on the other hand represents an ordinary branch loop with rules.

5.2.5.2 *Merger Validator*

The `merger_validator` module handles the merger processing stage. The `MergerValidator` class within the module uses the `Parser` and `GroupFilterValidator` instances passed to it to create a `Merger` instance once the check on referenced fields and branch names has passed. In addition, the validator also ensures semantic checks on Allen algebra such as whether the Allen relation arguments are correctly ordered, whether the Allen rules with the same set of arguments are connected by an OR and whether each branch loop is reachable by an Allen relation (or a chain of Allen relations) from the top level branch.

5.2.6 *Ungrouper*

The ungrouper unwraps the tuples of group-records into individual flow-records, ordered by their timestamps. The duplicate flow-records appearing from several group-records are eliminated and are sent as output only once.

5.2.6.1 *Ungrouper Implementation*

The `ungrouper` module handles the unwrapping of the group-records. The generation of flow-records can also be suppressed using the `-no-records-ungroup` command line option. The `Ungrouper` class instance is initialized using a merger file and an explicit export order.

5.2.6.2 *Ungrouper Validator*

The `ungrouper_validator` module handles the ungrouper processing stage. The `UngrouperValidator` class within the module uses the `Parser` and `MergerValidator` instances passed to it to create a `Ungrouper` instance. This processing stage does *not* require any validation.

6 FLOWY IMPROVEMENTS USING MAP/REDUCE

Flowy, although clearly setting itself apart with its additional functionality to query intricate patterns in the flows demonstrates relatively high execution times when compared to contemporary flow-processing tools. A recent study [2] revealed that a sample query run on small record set (around 250MB) took 19 minutes on Flowy as compared to 45 seconds on `flow-tools`. It, therefore is imperative that the application will benefit from distributed and parallel processing. To this end, recent efforts were made to investigate possibility of making Flowy Map/Reduce aware [2]

6.1 MAP/REDUCE FRAMEWORKS

Map/Reduce is a programming model for processing large data sets by automatically parallelizing the computation across large-scale clusters of machines [11]. It defines an abstraction scheme where the users specify the computation in terms of a `map` and `reduce` function and the underlying systems hides away the intricate details of parallelization, fault tolerance, data distribution and load balancing behind an Application Programming Interface (API).

6.1.1 *Apache Hadoop*

Apache Hadoop is a Map/Reduce Framework written in Java that exposes a simple programming API to distribute large scale processing across clusters of computers [12]. However in order to make Flowy play well with the framework, the implementation either has to use a Python wrapper around the Java API or translate the complete implementation to Java through Jython. Even more since Flowy uses HDF files for it's I/O processing, staging the HDF files properly in the Hadoop Distributed File System (HDFS) [13] and then later streaming them using Hadoop Streaming utility would still be an issue as suggested in [2]

6.1.2 *The Disco Project*

Disco is a distributed computing platform using the Map/Reduce framework for large-scale data intensive applications [14]. The core of the platform is written in Erlang and the standard library to interface with the core is written in Python. Since the `map` and `reduce` jobs can be easily written as Python functions and dispatched to the worker threads in a pre-packaged format, it is less difficult to setup Disco to utilize Flowy as a `map` function. In addition, the usage of HDF files for I/O processing pose no additional modifications whatsoever since the input data files can be anywhere and supplied to the worker threads in absolute paths.

6.2 PARALLELIZING FLOWY

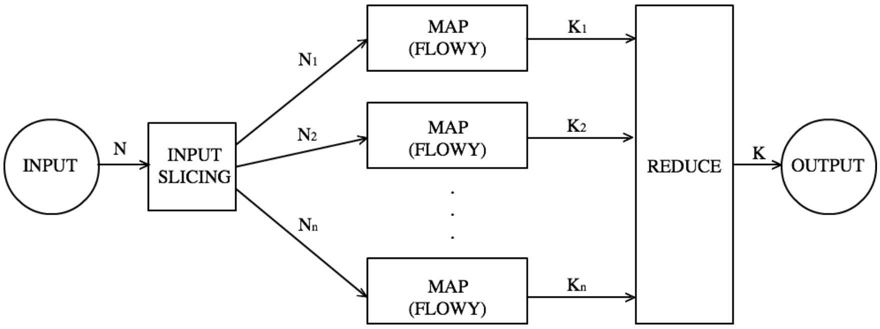

Figure 2: Parallelizing Flowy using Map/Reduce [2]

In an attempt to parallelize Flowy, it was run as a map function on a successful single node Disco installation as shown in **??**. Although the setup on a multiple node cluster would be theoretically almost the same, Flowy has not yet been tested in such a scenario.

6.2.1 *Slicing Inputs*

6.2.1.1 *Using* only *Filters*

6.2.1.2 *Using Groupers*

6.2.1.3 *Using Mergers*

6.2.2 *Flowy as a Map Function*

7 FLOWY 2.0

8 FLOWY: APPLICATIONS

8.1 ipv6 transition failure identification

8.2 cybermetrics: user identification

8.3 application identification using flow signatures

8.4 tcp level spam detection

Part III

MOTIVATION

Part IV

WORK PLAN

You can put some informational part preamble text here

9 DESIGN

10 IMPLEMENTATION

11 PERFORMANCE EVALUATION

12 CONCLUSION

Part V

IMPLEMENTATION AND EVALUATION

You can put some informational part preamble text here

13 DESIGN

14 IMPLEMENTATION

15 PERFORMANCE EVALUATION

16 FUTURE WORK

17 CONCLUSION

Part VI

APPENDIX

A APPENDIX

Put your appendix here.

BIBLIOGRAPHY

[1] Vladislav Marinov and Jürgen Schönwälder. Design of a Stream-Based IP Flow Record Query Language. In *Proceedings of the 20th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management: Integrated Management of Systems, Services, Processes and People in IT*, DSOM '09, pages 15–28, Berlin, Heidelberg, 2009. Springer-Verlag.

[2] Peter Nemeth. Flowy Improvements using Map/Reduce. Bachelor's thesis, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany, May 2010.

[3] Kaloyan Kanev. Flowy - Network Flow Analysis Application. Master's thesis, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany, August 2009.

[4] Kaloyan Kanev, Nikolay Melnikov, and Jürgen Schönwälder. Implementation of a stream-based IP flow record query language. In *Proceedings of the Mechanisms for autonomous management of networks and services, and 4th international conference on Autonomous infrastructure, management and security*, AIMS'10, pages 147–158, Berlin, Heidelberg, 2010. Springer-Verlag.

[5] Vladislav Marinov. Design of an IP Flow Record Query Language. Master's thesis, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany, August 2009.

[6] Vladislav Marinov and Jürgen Schönwälder. Design of an IP Flow Record Query Language. In *Proceedings of the 2nd international conference on Autonomous Infrastructure, Management and Security: Resilient Networks and Services*, AIMS '08, pages 205–210, Berlin, Heidelberg, 2008. Springer-Verlag.

[7] Francesc Alted and Mercedes FernÃ¡ndez-Alonso. PyTables: Processing And Analyzing Extremely Large Amounts Of Data In Python. 2003.

[8] B. Claise. Cisco Systems NetFlow Services Export Version 9. RFC 3954 (Informational), October 2004.

[9] J. Quittek, S. Bryant, B. Claise, P. Aitken, and J. Meyer. Information Model for IP Flow Information Export. RFC 5102 (Proposed Standard), January 2008. Updated by RFC 6313.

[10] James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26:832–843, November 1983.

[11] Jeffrey Dean and Sanjay Ghemawat. Mapreduce: simplified data processing on large clusters. In *Proceedings of the 6th conference on Symposium on Opearting Systems Design & Implementation - Volume 6*, pages 10–10, Berkeley, CA, USA, 2004. USENIX Association.

[12] T. White. *Hadoop: The Definitive Guide*. Definitive Guide Series. O'Reilly, 2010.

[13] K. Shvachko, Hairong Kuang, S. Radia, and R. Chansler. The Hadoop Distributed File System. In *Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on*, pages 1 –10, May 2010.

[14] Prashanth Mundkur, Ville Tuulos, and Jared Flatow. Disco: A Computing Platform for Large-Scale Data Analytics. In *Proceedings of the 10th ACM SIGPLAN workshop on Erlang*, Erlang '11, pages 84–89, New York, NY, USA, 2011. ACM.

DECLARATION

Put your declaration here.

Bremen, Germany, June 2012

Vaibhav Bajpai