

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

[December 25, 2011 at 22:25]

ABSTRACT

Short summary of the contents in English...

*We have seen that computer programming is an art,
because it applies accumulated knowledge to the world,
because it requires skill and ingenuity, and especially
because it produces objects of beauty.*

— Donald E. Knuth [2]

ACKNOWLEDGMENTS

Put your acknowledgments here.

CONTENTS

I INTRODUCTION	1
1 TRAFFIC MEASUREMENT APPROACHES	3
1.1 Capturing Packets	3
1.2 Capturing Flows	3
1.3 Remote Monitoring	3
1.4 Remote Metering	3
2 FLOW EXPORT PROTOCOLS	5
2.1 NetFlow	5
2.2 IPFIX	5
2.3 sFlow	5
3 LANGUAGES AND TOOLS	7
3.1 nfdump	7
3.2 flow-tools	7
3.3 Gigascope	7
4 LEGAL CONSIDERATION	9
II STATE OF THE ART	11
5 FLOWY	13
5.1 Processing Pipeline	13
5.1.1 Splitter	13
5.1.2 Filter	13
5.1.3 Grouper	13
5.1.4 Group-Filter	13
5.1.5 Merger	13
5.1.6 Ungrouper	13
5.2 Python Framework	13
5.2.1 PyTables and PLY	13
5.2.2 Records	13
5.2.3 Filters and Rules	13
5.2.4 Branches and Branch Masks	13
6 FLOWY IMPROVEMENTS USING MAP/REDUCE	15
7 FLOWY 2.0	17
8 FLOWY: APPLICATIONS	19
8.1 IPv6 Transition Failure Identification	19
8.2 Cybermetrics: User Identification	19
8.3 Application Identification using Flow Signatures	19
III IMPLEMENTATION AND EVALUATION	21
9 DESIGN	23
10 IMPLEMENTATION	25
11 PERFORMANCE EVALUATION	27

12 FUTURE WORK	29
13 CONCLUSION	31
IV APPENDIX	33
A APPENDIX	35
BIBLIOGRAPHY	36

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

Part I

INTRODUCTION

You can put some informational part preamble text here

TRAFFIC MEASUREMENT APPROACHES

1.1 CAPTURING PACKETS

1.2 CAPTURING FLOWS

1.3 REMOTE MONITORING

1.4 REMOTE METERING

FLOW EXPORT PROTOCOLS

2.1 NETFLOW

2.2 IPFIX

2.3 SFLOW

LANGUAGES AND TOOLS

3.1 NFDUMP

3.2 FLOW-TOOLS

3.3 GIGASCOPE

LEGAL CONSIDERATION

Part II

STATE OF THE ART

You can put some informational part preamble text here

FLOWY

5.1 PROCESSING PIPELINE

5.1.1 *Splitter*

5.1.2 *Filter*

5.1.3 *Grouper*

5.1.4 *Group-Filter*

5.1.5 *Merger*

5.1.6 *Ungrouper*

5.2 PYTHON FRAMEWORK

5.2.1 *PyTables and PLY*

5.2.2 *Records*

5.2.3 *Filters and Rules*

5.2.4 *Branches and Branch Masks*

FLOWY IMPROVEMENTS USING MAP/REDUCE

FLOWY: APPLICATIONS

8.1 IPV6 TRANSITION FAILURE IDENTIFICATION

8.2 CYBERMETRICS: USER IDENTIFICATION

8.3 APPLICATION IDENTIFICATION USING FLOW SIGNATURES

Part III

IMPLEMENTATION AND EVALUATION

You can put some informational part preamble text here

PERFORMANCE EVALUATION

FUTURE WORK

CONCLUSION

Part IV

APPENDIX



APPENDIX

Put your appendix here.

BIBLIOGRAPHY

- [1] Robert Bringhurst. *The Elements of Typographic Style*. Version 2.5. Hartley & Marks, Publishers, Point Roberts, WA, USA, 2002.
- [2] Donald E. Knuth. Computer Programming as an Art. *Communications of the ACM*, 17(12):667–673, December 1974.

COLOPHON

This thesis was typeset with $\text{\LaTeX} 2_{\epsilon}$ using Hermann Zapf's *Palatino* and *Euler* type faces (Type 1 PostScript fonts *URW Palladio L* and *FPL* were used). The listings are typeset in *Bera Mono*, originally developed by Bitstream, Inc. as "Bitstream Vera". (Type 1 PostScript fonts were made available by Malte Rosenau and Ulrich Dirr.)

The typographic style was inspired by [Bringhurst](#)'s genius as presented in *The Elements of Typographic Style* [1]. It is available for \LaTeX via CTAN as "[thesis](#)".

NOTE: The custom size of the textblock was calculated using the directions given by Mr. Bringhurst (pages 26–29 and 175/176). 10 pt Palatino needs 133.21 pt for the string "abcdefghijklmnopqrstuvwxyz". This yields a good line length between 24–26 pc (288–312 pt). Using a "double square textblock" with a 1:2 ratio this results in a textblock of 312:624 pt (which includes the headline in this design). A good alternative would be the "golden section textblock" with a ratio of 1:1.62, here 312:505.44 pt. For comparison, DIV9 of the `typearea` package results in a line length of 389 pt (32.4 pc), which is by far too long. However, this information will only be of interest for hardcore pseudo-typographers like me.

To make your own calculations, use the following commands and look up the corresponding lengths in the book:

```
\settowidth{\abcd}{abcdefghijklmnopqrstuvwxyz}  
\the\abcd\ % prints the value of the length
```

Please see the file `thesis.sty` for some precalculated values for Palatino and Minion.

145.86469pt

DECLARATION

Put your declaration here.

Bremen, Germany, June 2012

Vaibhav Bajpai