

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

[January 10, 2012 at 19:44]

ABSTRACT

Short summary of the contents in English...

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 SQL-based Query Languages	7
3.1.1 NetFlow exports as relational DBMS	7
3.1.2 Data Stream Management System	7
3.1.3 Gigascope	7
3.1.4 Tribeca	7
3.2 Filtering Languages	7
3.2.1 flow-tools	7
3.2.2 nfdump	7
3.3 Procedural Languages	7
3.3.1 FlowScan	7
3.3.2 Clustering NetFlow Exports	7
3.3.3 SiLK Analysis Suite	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
8.4	TCP level Spam Detection	19
	III MOTIVATION	21
	IV WORK PLAN	23
9	DESIGN	25
10	IMPLEMENTATION	27
11	PERFORMANCE EVALUATION	29
12	CONCLUSION	31
	V IMPLEMENTATION AND EVALUATION	33
13	DESIGN	35
14	IMPLEMENTATION	37
15	PERFORMANCE EVALUATION	39
16	FUTURE WORK	41
17	CONCLUSION	43
	VI APPENDIX	45
A	APPENDIX	47
	BIBLIOGRAPHY	48

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 SQL-BASED QUERY LANGUAGES

3.1.1 *NetFlow exports as relational DBMS*

3.1.2 *Data Stream Management System*

3.1.3 *Gigascop*

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*

LEGAL CONSIDERATION

Part II

STATE OF THE ART

You can put some informational part preamble text here

FLOWY

Flowy [1][2] is the first prototype implementation of a stream-based flow record query language [3][4][5]. The query language allows to describe patterns in flow-records in a declarative and orthogonal fashion, making it easy to read and understand.

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
- 8.4 TCP LEVEL SPAM DETECTION

Part III

MOTIVATION

Part IV

WORK PLAN

You can put some informational part preamble text here

PERFORMANCE EVALUATION

CONCLUSION

Part V

IMPLEMENTATION AND EVALUATION

You can put some informational part preamble text here

FUTURE WORK

CONCLUSION

Part VI

APPENDIX



APPENDIX

Put your appendix here.

BIBLIOGRAPHY

- [1] Kaloyan Kanev. Flowy - Network Flow Analysis Application. Master's thesis, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany, 2009.
- [2] 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.
- [3] Vladislav Marinov. Design of an IP Flow Record Query Language. Master's thesis, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany, 2009.
- [4] 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.
- [5] 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.

DECLARATION

Put your declaration here.

Bremen, Germany, June 2012

Vaibhav Bajpai