

Contents

Chapter 1	Introduction.....	1
1.1	Problem Statement.....	1
Chapter 2	Related pattern-matching algorithms.....	4
2.1	Aho-Corasick pattern-matching algorithm.....	4
2.2	EGREP matching algorithm.....	6
2.3	Comparison between AC and EGREP.....	7
Chapter 3	Custom automata system for Snort rules.....	10
3.1	Advantages in custom automata system.....	11
3.2	The architecture in custom automata system.....	11
3.3	Supporting Regular Expression.....	14
3.4	Implement header and content custom automata.....	15
3.5	Experiments and comparisons between custom automata system and Snort.....	24
3.6	Implementation in hardware.....	28
Chapter 4	Automata for Regular Expression.....	37
4.1	The architecture in Regular Expression automata.....	37
4.2	Automata in automata.....	39
4.3	Implement Regular Expression automata.....	39
4.4	Compiler and Optimization.....	45
4.5	Experiments and comparisons between Regular Expression automata and EGREP.....	45
Chapter 5	Conclusion.....	49
References	52

Figure and Table Index

Figure Index

Figure 2-1 An example of AC automata	5
Figure 2-2 Function diagrams of EGREP compiling	6
Figure 2-3 Function diagrams of EGREP matching	7
Figure 2-4 Examples of EGREP and AC automata	8
Figure 2-5 Examples of data structure in EGREP and AC automata.....	9
Figure 3-1 Overview of custom automata system for Snort rules when compiling	12
Figure 3-2 Overview of custom automata system for Snort rules when matching.....	13
Figure 3-3 An example of custom header automata.	16
Figure 3-4 An example of custom content automata	17
Figure 3-5 Compiling flowchart of custom content automata	20
Figure 3-6 Matching flowchart of custom content automata.....	22
Figure 3-7 An example of custom content automata while matching	23
Figure 3-8 Matching time of custom automata system when input different data sizes.....	25
Figure 3-9 Hardware architecture for parallel pattern matching.	29
Figure 3-10 Hardware Design of Custom Automata System.....	30
Figure 3-11 An example of group nodes with hash filter ($M=2$)	31
Figure 3-12 An example of dividing every pattern into 4-byte substrings.	32
Figure 3-13 Matching engine in hardware with N 32-bit ALUs ($N=4$)	33
Figure 3-14 Flowchart of custom automata in hardware matching	36
Figure 4-1 Overview of Regular Expression Automata.....	37
Figure 4-2 An example of data structure in Regular Expression Automata	38
Figure 4-3 Concept and example of automata in automata.	39

Figure 4-4 Compiling Flowchart of Regular Expression automata	42
Figure 4-5 Matching Flowchart of Regular Expression automata	44
Figure 4-6 Comparison of matching time for different data sizes	48

Table Index

Table 1 Comparisons between EGREP algorithm and AC algorithm.....	7
Table 2 Matching time of custom automata system when input different data sizes...	24
Table 3 Matching time of Snort system for different packets.....	27
Table 4 Matching time of custom automata system for different packets	28
Table 5 Comparison of matching time (Regular Expression vs. EGREP).	46

