American Journal of Applied Sciences 9 (10): 1618-1624, 2012
ISSN 1546-9239
© 2012 Science Publication

Reconfigurable Hardware Architecture for Network Intrusion Detection System

Kaleel Rahuman A. and G. Athisha
Department of ECE
PSNA College of Engineering and Technology, Dindigul, TamilNadu, India

Abtract: Intrusion rule processing in reconfigurable hardware enables intrusion detection and prevention. The use of reconfigurable hardware for network security applications has great strides as Field Programmable Gate Array (FPGA) devices have provided larger and faster resources. This proposes architecture called "BV-TCAM" is presented, which is implemented for an FPGA-based Network Intrusion Detection Systems (NIDS). The BV-TCAM architecture combines the Ternary Content Addressable Memory (TCAM) and Bit Vector (BV) algorithm to effectively compress the data representation and throughput. A tree bitmap implementation of the BV algorithm is used for source and destination port lookup while a TCAM performs lookup for other header fields, which can be represented as a prefix or exact value. With the aid of small embedded TCAM, packet classification can be implemented in relatively small part of the available logic of an FPGA. The BV-TCAM architecture has been modelled by VHDL. Simulations were performed by MODELSIM. This architecture have to be synthesized and implement our design using Xilinx FPGA device.

Key words: Network Intrusion Detection Systems (NIDS), Bit Vector (BV), Ternary Content Addressable Memory (TCAM), Field Programmable Gate Array (FPGA), reconfigurable hardware

INTRODUCTION

The world is now network together; People, business and governments share information and communicate nearly instantaneously. Individuals use the networks for today's everyday tasks, such as banking, shopping, investing or transferring pictures to friends. With sensitive information now available on-line, measures must be taken to ensure security and privacy the electronics database of customer's credit card numbers, address and phone number must be secured against identity theft (Michael, 2005). Similarly, medical institutions must secure patient information to protect medical information and maintain privacy.

In order to protect networked systems, intrusion detection and prevention is necessary. Intrusion detection determines when harmful activities are being attempted. Intrusions are defined as attempts to compromise the confidentially, integrity, availability, or bypass the security mechanisms of a computer or network. They are caused by attackers accessing the system from the internet, authorized users of the system who attempt to gain additional privileges for which they are not authorized and authorized users who misuse the privileges given them. Intrusion detection is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of intrusions.

Types of intrusions: Intrusion can take several forms. They can occur as abnormal, unauthorized, or unwanted system usage. Examples related to networking.

Unauthorized access: Unauthorized access occurs when an individual's gains access to a system they have no right to use. For example, a user may view web pages containing proprietary information that they have not been authorized to view.

Authorized access: An intrusion can occur even if the credentials of the individual accessing the system are correct. For example, an intruder can fraudulently obtain account information such as login names and passwords. This is known as masquerading. The system believes the intruder is authorized. This is the most difficult type of intrusion to detect, since the detector must consider what is being accessed and what operations are being performed.

Corresponding Author: Kaleel Rahumanm A., Department of ECE, PSNA College of Engineering and Technology, Dindigul, TamilNadu, India

Spam: Spam is an unwanted electronic message from individuals or companies who send the message to people that may not desire to receive the messages. These messages generally try to sell items, such as medication, loan application or pornography. Phasing is a heinous form of spam where a message supposedly from an authoritative institution, such as bank, e-commerce site, or government agency, direct the recipient to replay to the messages or go to a web page and enter sensitive information. The messages can be quite persuasive, claiming account will be deactivated unless information is verified.

Virus: A virus is a piece of malware hidden in files or emails. Once activated by the host, the virus replicates itself and spreads to additional hosts. Viruses generally spread via email, requesting that the recipient view an attachment. Clever virus writers write code to search an infected host's address to find additional recipients. The virus assumes the host identity when sending new email messages, increasing the likelihood that the target becomes spoiled.

Methods of the detection:

Signature based: Signature based intrusion can be recognized by comparison with known patterns. Somewhat confusingly, this is also often referred to as misuse detection. A database of known cases of misuse (signatures) is maintained and incoming events are compared with these signatures to determine if they are representative of a misuse. The detection system would maintain a large database of signatures, which could be updated as new misuses are identified. Such systems are inherently limited by the size of the database and as such are prone to issuing false negatives- i.e., indicating normal behaviour for unknown misuses.

Anomaly detection: Where a baseline of normal network activity is compiled and deviations from this baseline then imply that an anomalous event has misuse. Such system can detect new misuses provided that they are sufficiently different from normal behaviour. Given the range of normal activity and the quantity of data occurred and this is taken to be a involved, this type of detection mechanism is prone to issuing false positives-i.e., indicating abnormal behaviour for innocent events that deviate from known patterns of activity (Michael, 2005).

Performance metrics for classification algorithms:

- Search speed: Faster links require faster classification.
- Low storage requirements: Small storage requirements enable the use of fast memory technologies like Static Random Access Memory (SRAM)
- Ability to handle large real-life classifiers
- Fast updates: As the classifier changes, the data structure need to be updated
- Scalability in the number of header fields used for classification
- Flexibility in specification: A classification algorithm should support general rules, including prefixes, operators and wildcards

BV-TCAM architeture: Our design combines and optimizes the Ternary Content Addressable Memories (TCAMs) and Bit Vector algorithms for packet header classification in NIDS (Spitznagel *et al*., 2003).

Network intrusion detection systems require header classification to report all matches, not just one. In usual applications, TCAM is associated with priority encoder than only reports the ID of the matched entry with the highest priority. In this application, an un-encoded TCAM are used. That is, the number of output bits equals the number of TCAM entries and each bit indicates the matching status of the corresponding TCAM entry (Clark and Schimmel, 2003). Just like the BV output, the un-encoded TCAM output forms another bit vector and each bit indicates the match to the corresponding rule field(s) or not. So the idea is that the header fields are partitioned in a way that some of them are classified using TCAM while others are classified using Bit Vector algorithms (Lakshminarayanan *et al*., 2005). This design exclude source and destination port fields from TCAM while keeping IP address and protocol fields in TCAM. We order the rules in the same sequence; hence we can intersect all the output bit vectors to get the set of matches. This method optimizes the size of the TCAM, as it does not expand the number of TCAM entries (Lie, 2002).

Our design combines and optimizes the TCAM and Bit Vector algorithms for packet header classification in NIDS. Network intrusion detection systems require header classification to report all matches, not just one (Fig. 1). In usual applications, TCAM is associated with priority encoder than only reports the ID of the matched entry with the highest priority. In this application, an un-encoded TCAM are used. That is, the number of output bits equals the number of TCAM entries and each bit indicates the matching status of the corresponding TCAM entry. Just like the BV output, the un-encoded TCAM output forms another bit vector and each bit indicates the match to the corresponding rule field(s) or not. So the idea is that the header fields are partitioned in a way that some of them are classified using TCAM while others are classified using Bit Vector algorithms. This design exclude source and destination port fields from TCAM while keeping IP address and protocol fields in TCAM. We order the rules in the same sequence; hence we can intersect all the output bit vectors to get the set of matches. This method optimizes the size of the TCAM, as it does not expand the number of TCAM entries (Baker and Prasanna, 2004).

TCAM overview: Figure 2a shows a 1-T dynamic RAM cell and Fig. 2b shows 6-T dynamic TCAM cell. In a DRAM cell, the Bit Line (BL) is connected to the capacitor when the Word Line (WL) is high thus enabling read and write functionality. In a dynamic TCAM, there are two DRAM storage cell (M3 and M6) and the read write functionality remains the same. The additional four transistors (M1, M2, M4 and M5) make up the comparison logic used for the match operation.

The Match Line (ML) is pre-charged to Vdd. The storage nodes are loaded with complementary data. The Search Lines (SL) are charged to the value which is being searched. The comparison logic essentially performs an XNOR operation. In the case of SL data matching the stored data, the ML does not discharge. If the bits mismatch, then the ML is discharged through two of the four search transistors. When l bits are placed in parallel with a common ML, if any one bit mismatches, the ML will discharge. Only if all stored (l) bits and complements match l-pair search lines will the ML remain charged (Yu *et al*., 2004).

The ternary nature of the TCAM cell is evident when”0” is stored on both capacitors. This turns off both the lower transistors in the comparison logic. Regardless of the values on the SLs, the ML is unable to discharge. This effectively represents a “don’t care” condition being stored, hence the name “ternary”:”1”,”0” and “don’t care”. Table 1 shows the different states that can be stored in the TCAM cell. The SLs can both be set to “0” as well. This is equivalent to searching for a “don’t care” condition, which will always match since the ML cannot discharge through either discharge path in the comparison logic (Clark and Schimmel, 2003).

Figure 3 shows a simplified schematic of TCAM search architecture. In the event of multiple matches, a Priority Encoder (PE) allows the ML with the highest priority to be encoded as a matching address. The PE is composed of a Multiple-Match Resolve (MMR) and an Address Encoder (MAE).The MMR allows only the ML with the highest priority to pass the MAE and outputs a multiple match detection signal if there is more than one match represent. The MAE is Read-Only Memory (ROM) that encodes the address of the output from the MMR. These additional components are not found in RAMs; hence special attention must be paid to testing both these components and the comparison logic.

TCAM architecture: A typical TCAM Integrated Circuit (IC) consists of three major types of circuits:

- TCAM arrays to store ternary data
- Peripheral components to facilitate read, write and search functionality
- Design of Test (DFT) components that enable testing of the various other components

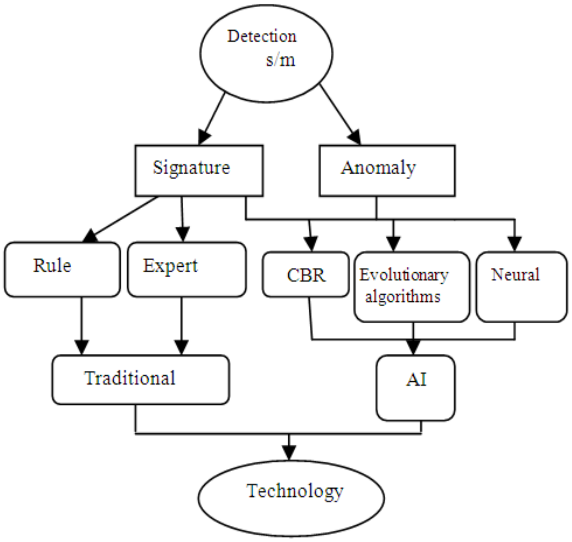
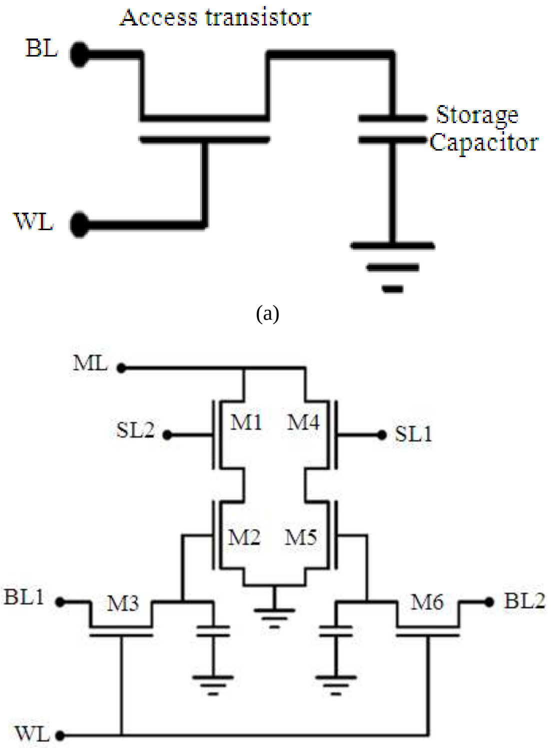


Fig. 1: Overview of detection mechanisms



(a)

(b)

Fig. 2: (a) DRAM cell and (b) TCAM cell

Table 1: TCAM Storage and Search States

Value	BL1/SL1	BL2/SL2
Zero (0)	0	1
One (1	1	0
Don't Care (X)	0	0
Not Used	1	1

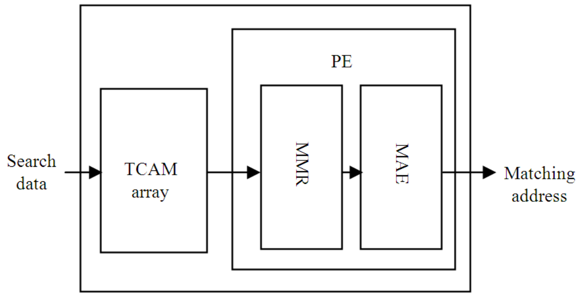


Fig. 3: Simplified TCAM Search Architecture

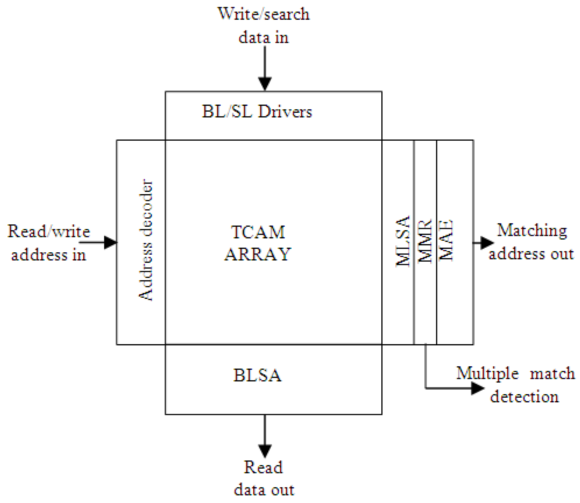


Fig. 4: TCAM architecture overview

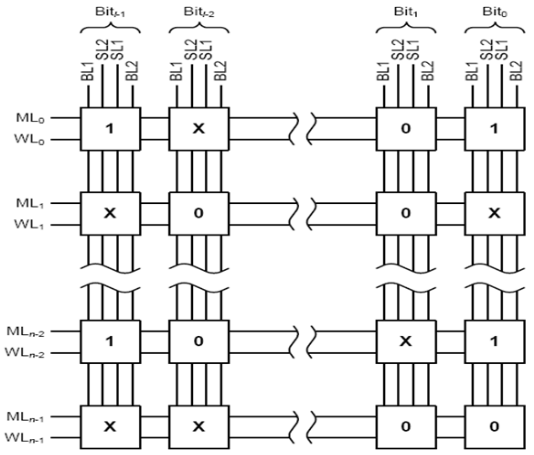


Fig. 5: TCAM Array organization

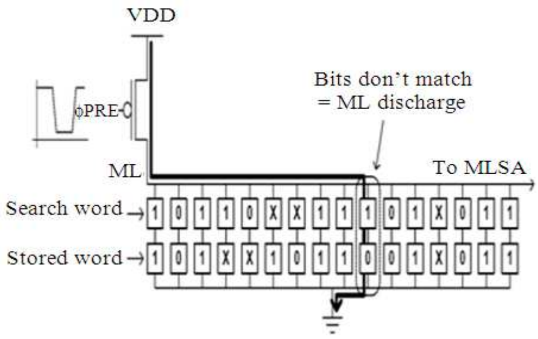


Fig. 6: TCAM Word during Search Operation

Figure 4 shows a simplified block diagram of a TCAM illustrating the first two types of circuits. The TCAM array consists of TCAM cells that make up TCAM words. Peripheral to the array is its associated address decoders, line drivers, buffers, sense amplifiers. (Spitznagel *et al*., 2003). Also, the peripheral circuitry includes the MMR and MAE. The DFT components include row/column redundancy, on-chip testing circuits, scan chains and any other circuits specifically included for testing purpose.

In memory semiconductors, it is common practice to segment the entire memory into blocks and banks. This allows power to be saved by disabling portions of the memory that are not in use. Also, signal drivers can be made smaller since they only have to drive local busses and not global busses.

TCAM array: Figure 5 shows a more detailed diagram of the TCAM array. Each memory block in a TCAM chip is implemented as an array of TCAM cell, where each TCAM cell is capable of storing a ternary state.

Typically, a horizontal row forms a word of a TCAM-based table. Within a word, a bit is located by its column number. All the TCAM cell in a row share a common WL and a common ML. Similarly, the TCAM cell in the same column share a common set of BLs and a common set of SLs.

TCAM word: A TCAM word consists of l TCAM cell all connected to a common WL and ML. During a write operation, an l- bit word is stored in the TCAM cells at the address specified by the user. The read operation returns an l - bit word stored at the given address. The search operation compares an l -bit search word against the bb-bits stored in the TCAM word and if the two match, the address of the TCAM word is returned.

During a search operation, the ML is precharged and the l- bit search word is applied to the TCAM cells on the SLs. Any mismatching cells cause the ML to discharge, indicating a mismatch. This is illustrated in Fig. 6. The " x"("don't care") condition shut off the search path for that cells, thus preventing an ML discharge regardless of the applied data. If the ML remains charged, the stored word and search word match.

MATERIALS AND METHODS

WL and BL drivers and BLSAs: The WL drivers are same as in any RAM. The WLs are driven by large buffers activated by an address decoder. There are existing test algorithms that are quite suitable for testing address decoders with a high fault coverage. The BL drivers are tristate buffers that can be activated during write operation and set to high-Z state during read operations. The BL Sense Amplifiers (BLSAs) can be voltage or current mode and are single-ended in dynamic memories or double ended (differential) in static memories.

The clock is 1, the reset is 0 and the input is given 1 and the output received 72 bit and the 2 eight bits. i.e., the header file and datas. The Maximum Frequency is 104.548MHz and the minimum period is 9.565ns.

All these components are highly established and researched and are identically implemented in TCAMs as they are in RAMs.

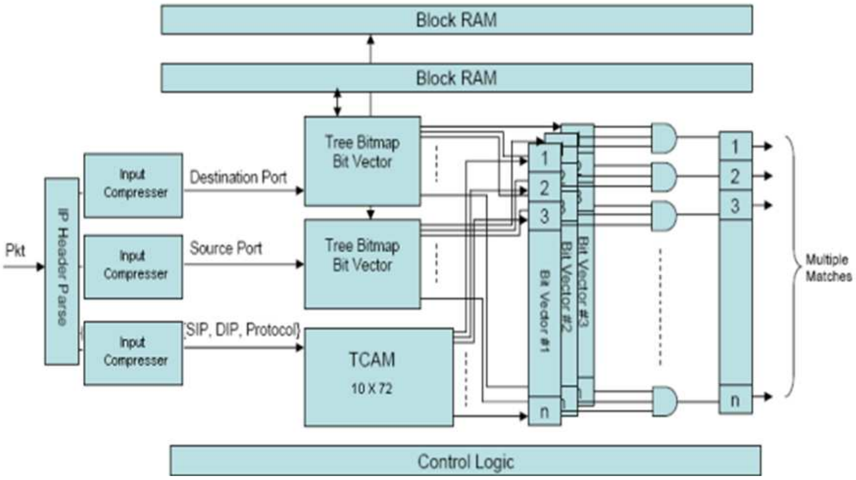
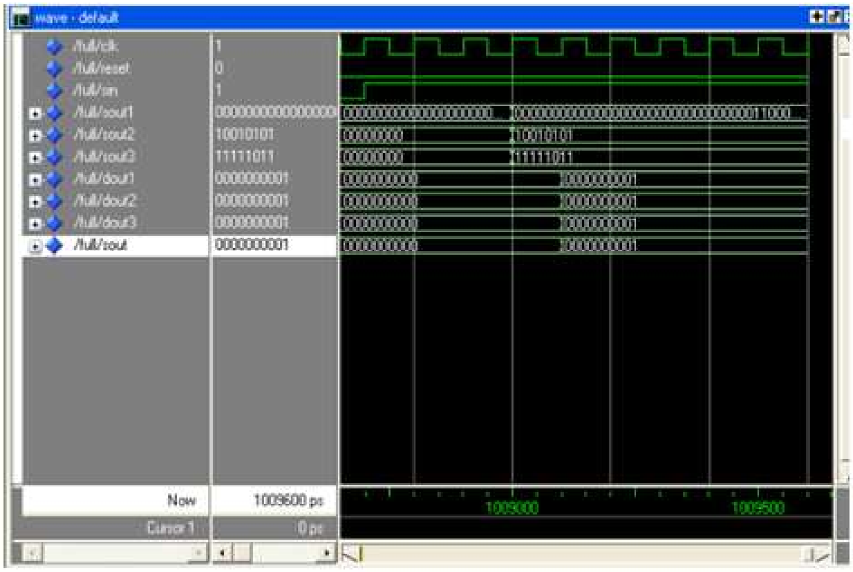


Fig. 7: BV-TCAM Circuit block diagram



```
wave - default
/full/clk     1
/full/reset   0
/full/sin     1
/full/sout1   00000000000000000
/full/sout2   10010101
/full/sout3   11111011
/full/dout1   0000000001
/full/dout2   0000000001
/full/dout3   0000000001
/full/sout    0000000001
Now           1009600 ps
Cursor 1      0 ps
```

Fig. 8: Simulation results of a BV-TCAM (MODELSIM)

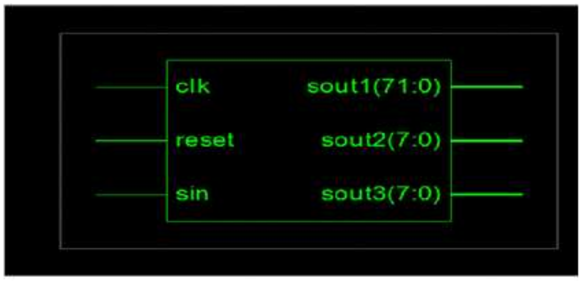


Fig. 9: RTL schematic of BV-TCAM

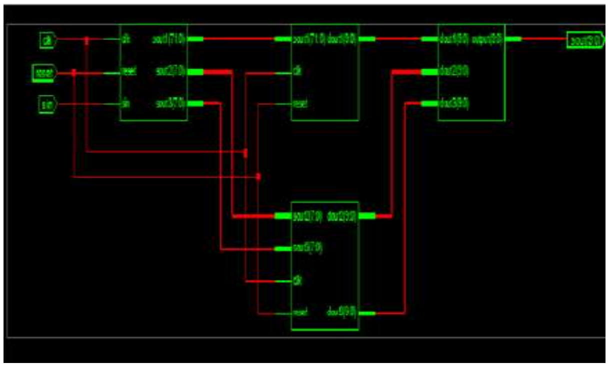


Fig. 10: RTL schematic of BV-TCAM

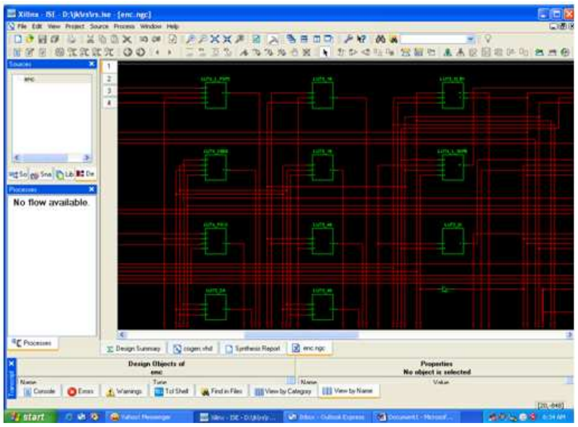


Fig. 11: Technology schematic of BV-TCAM

SL drivers and MLSA: The SL drivers are buffers very similar to the BL drivers, except that they do not need to be tristate, since the SLs are never read from, only written to.

This block indicates the inputs and outputs of the architecture. The above Fig. 7-11 indicates the total number of flip flops and blocks allotted in the chip. The above schemetic indicates Total area occupied in the chip.

The ML Sense Amplifiers (MLSAs) are responsible for detecting whether or not its associated ML has charged or discharged during a search operation. The MLs are charged during every search operation. Limiting the minimum charge necessary to make an accurate detection of the ML's state is critical to reducing the energy required per search operation. Consequently, there are numerous MLSA designs, all aimed at lowering the charge used on the ML during search operation.

For the purpose of TCAM array testing, the MLSA is assumed to provide correct results since a malfunctioning MLSA would result in grossly incorrect test results (always matching, or always mismatching), which is easily detected.

FPGA: Field Programmable Gate Array (FPGA) technology enables a full-custom, high speed and flexible hardware implementation of data processing circuits. FPGA use Configurable Logic Blocks (CLBs) that contain look-up tables and flip-flops. Boolean logic functions are implemented using the CLBs. The speed of a design is determined by the location of the CLBs used and the routing delays incurred by connecting them. Fast, on-chip memories are also spread throughout current FPGAs .By allowing soft modification of the hardware, the non recurring expense of FPGA design is lower than for ASIC designs (Haoyu and Lockwood, 2005). For Low volume applications or prototyping, FPGAs are most cost-effective solution. The design cycle for a FPGA is considerably shorter than a fully-custom hardware solution.

RESULTS

The BV-TCAM architecture for network intrusion detection has been modeled by VHDL. The simulation waveforms are shown above. Simulations have been done in MODELSIM. The code has been synthesized using XILINX tool. Synthesis report is also given here for BV-TCAM architecture.

Timing summary:
Speed grade: -6: Minimum period: 9.560ns (Maximum Frequency: 104.588MHz):

- Minimum input arrival time before clock: 15.290ns
- Maximum output required time after clock: 7.958ns
- Maximum combinational path delay: No path found

DISCUSSION

The main contribution of present research concerned our design combines and optimizes the Ternary Content Addressable Memories (TCAMs) and Bit Vector algorithms for packet header classification in NIDS. The BV-TCAM architecture for network intrusion detection has been modeled by VHDL and FPGA implementation with fairly low system complexity and fast processing. The Maximum Frequency is 104.548MHz and the minimum period is 9.565ns.

CONCLUSION

The main contribution of present research concerned Ternary Content Addressable Memories (TCAMs) architecture for network intrusion detection has been modeled by VHDL and very efficiently classifies header rules for NIDS in an FPGA have two aspects. First, while using TCAM as a component, we avoid the need to expand the size of rule set by only using TCAM to classify the fields that is represented as prefix or exact value. We further compress the number of entries needed in TCAM due to the fact that the number of distinct combined values of these fields is very much less than the total number of rules. Second, after the port ranges are transformed to prefixes, we use a Tree Bitmap approach to implement the multi-bit tree Bit Vector algorithm. During the parallel operation and data structure size compression, the architecture is optimized for both storage effectiveness and throughput. It is fit for straightforward FPGA implementation with fairly low system complexity.

REFERENCES

Baker, Z. and V. Prasanna, 2004. Time and Area Efficient pattern matching on FPGAs. Proceedings of the ACM/SIGDA 12th International Symposium on Field Programmable Gate Arrays (FPGA’ 04), ACM New York, NY, USA, pp: 223-232. DOI: 10.1145/968280.968312

Clark, C. and D. Schimmel, 2003. Efficient reconfigurable logic circuits for matching complex network intrusion detection patterns. Proceedings of the International Conference on Field-Programmable Logic and Applications, (FPL’ 03), Lisbon, Portugal, pp: 956-959.

Haoyu, S. and J.W. Lockwood, 2005. Efficient Packet Classification for Network Intrusion Detection Using FPGA. Proceedings of the 2005 ACM/SIGDA 13th international symposium on Field-programmable gate array, Feb-20-22, ACM New York, NY, USA pp: 238-245. DOI: 10.1145/1046192.1046223

Lakshminarayanan, K., A. Rangarajan and S. Venkatachery, 2005. Algorithms for advanced packet classification with ternary CAMs. Proceedings of the Conference on Applications, Technologies, Architectures and Protocols for Computer Communications, (SIGCOMM '05), ACM New York, NY, USA., pp: 193-204. DOI: 10.1145/1080091.1080115

Lie, H., 2002. Efficient Mapping of Range Classifier into Ternary-CAM. Proceedings of the 10th Symposium on Symposium on High Performance Interconnects Dec, 10-02, IEEE Xplore Press, pp: 95-100. DOI: 10.1109/CONECT.2002.1039263

Michael, E.A., 2005. Architectures for Rule Processing Intrusion Detection and Prevention Systems. A thesis presented to the sever Institute of Washington University in partial fulfilment of the requirements for the degree of Master of science. Saint Louis, Missouri.

Spitznagel, E., D. Taylor and J. Turner, 2003. Packet Classification using extended TCAMs. Proceedings of the 11th IEEE International Conference on Network Protocols, Nov, 4-7, IEEE Xplore Press, pp: 120-131. DOI: 10.1109/ICNP.2003.1249762

Yu, F., R. Katz and T.V. Lakshman, 2004. Efficient multimatch packet classification and lookup with TCAM. IEEE Micro, 25: 50-59. DOI: 10.1109/MM.2005.8

Zachary, K., Baker, Z.K. and V.K. Prasanna 2004. Automatic synthesis of efficient intrusion detection systems on FPGAs. Field Programmable Logic Appli., 3203: 311-321. DOI: 10.1007/978-3-540-30117-2_33