High-Speed Parallel Processing of Protocol-Aware Signatures

Jordi Ros-Giralt, James Ezick, Peter Szilagyi, Richard Lethin

Reservoir Labs
632 Broadway, #803
New York, NY 10012
(212) 780-0527
giralt@reservoir.com

Unclassified, DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited. This material is based upon works supported by the Department of Energy under contract numbers DE-FG02-08ER85046. Any opinions, findings and conclusions expressed in this material are those of Reservoir Labs, and do not necessarily reflect the views of the Department of Energy.

Problem Definition and Previous Work

- **Why: Intelligent signature matching and protocol parsing are key functions of intrusion detection systems that need to keep up with line rates (100 Gbps)**
  - There exists a **trade off** between complexity of the signatures and processing speed.
  - Trend: cyber attacks become smarter, so **need to handle complex signatures.**
  - Trend: network trunks become faster, so **need to handle high speed**.
- **Problem: How to process large number of protocol signatures at high speed?**
  - Need to be able to process large number of signatures in parallel.
  - Need to minimize processing overhead on core processors.
- **How: Use SAT tools, binary decision diagrams (BDD) and deterministic finite automatons (DFA)**
  - Use SALT™ (converts Boolean functions into simplified Conjunctive Normal Form).
  - Compose OR signature, obtain its CNF form and calculate optimal BDD cuts.
  - Offload the resulting BDDs onto hardware accelerated DFA engines.
- **Result: Capability to process large number of signatures at high speed onto hardware DFA engines**
  - Currently building prototype on a heterogeneous 16-core processor/accelerator NPU (Cavium Octeon Plus) and looking at running this on the Octeon II and Tilera TILE64.
  - Experimental results presented in our abstract and poster.

How it Works Through a Low-Scale Example

Ghttpdlog signature expressed in salt:

```
;;;
;;; Ghttpdlog Salt
;;;

$bit1 expr =?
$bit2 expr =?
$bit3 expr =?
$bit4 expr =?
$bit5 expr =?
$conjunction1 and ~$bit1 $bit2 =
$conjunction2 and $bit1 ~$bit3
$bit4 =
$conjunction3 and $bit1 $bit3 $bit5
=
$disjunction or $conjunction1
$conjunction2 $conjunction3 =
$_ eval $disjunction + ; assert
return value true

#done
```

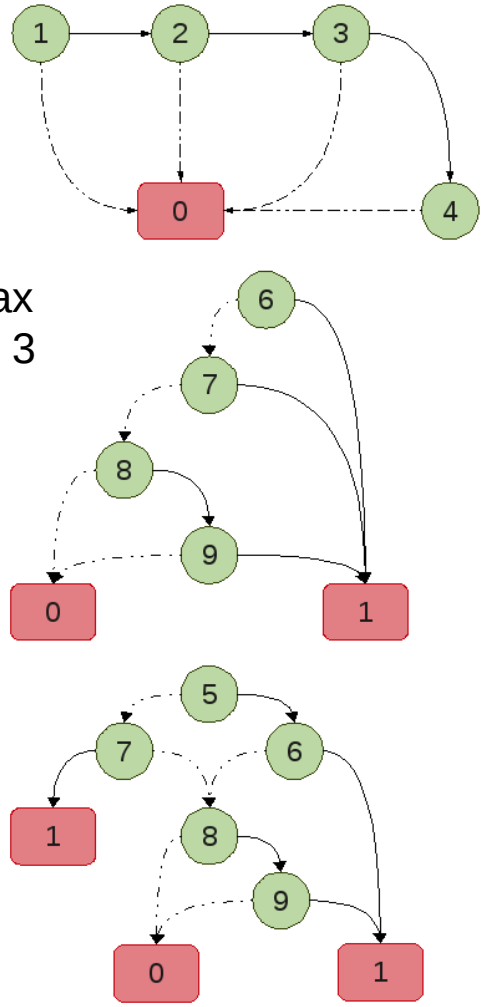
CNF:

```
2 4 5 0
2 1 4 0
2 3 4 0
2 -3 5 0
2 1 -3 0
2 1 5 0
2 1 0
2 1 3 0
-1 4 5 0
-1 3 4 0
-1 -3 5 0
6 0
-7 0
```

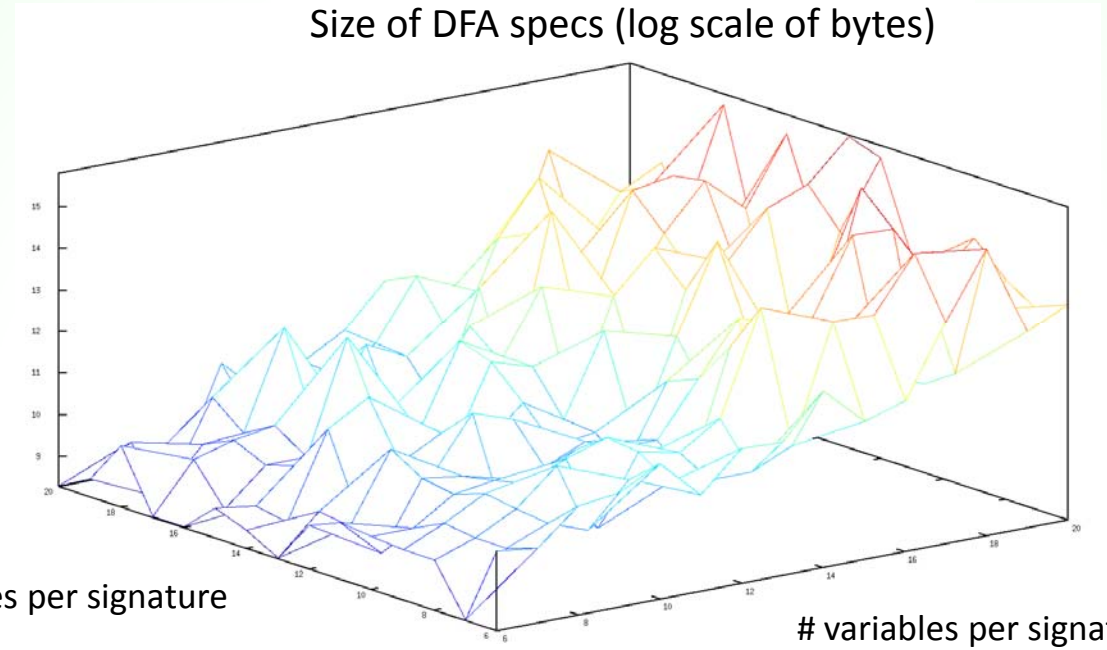
SALT™

$$(b2 \cap b4 \cap b5) \cup (b2 \cap b1 \cap b4) \cup (b2 \cap b3 \cap b4) \cup (b2 \cap \overline{b3} \cap b5) \cup (b2 \cap b1 \cap \overline{b3}) \cup (b2 \cap b1 \cap b5) \cup (b2 \cap b1) \cup (b2 \cap b1 \cap b3) \cup (\overline{b1} \cap b4 \cap b5) \cup (\overline{b1} \cap b3 \cap b4) \cup (\overline{b1} \cap \overline{b3} \cap b5)$$

Optimal min-max cut mapping on 3 DFA engines



Results and Preliminary Conclusions



- **Average case analysis. Using a random signature generator and taking average cases:**
  - Emitting DFA specs for signatures from 6 to 20 variables and from 6 to 20 clauses per signature, the average DFA size obtained is **410340 bytes**.
  - That yields **2.43 signatures per megabyte.**
  - On a DFA with 256MB of memory, we can fit in average about **600 signatures**.
  - Each signature can involve 100s of CPU cycles per connection offloaded from the core processor.
- **Real examples. Examples using real signatures are shown in the poster.**

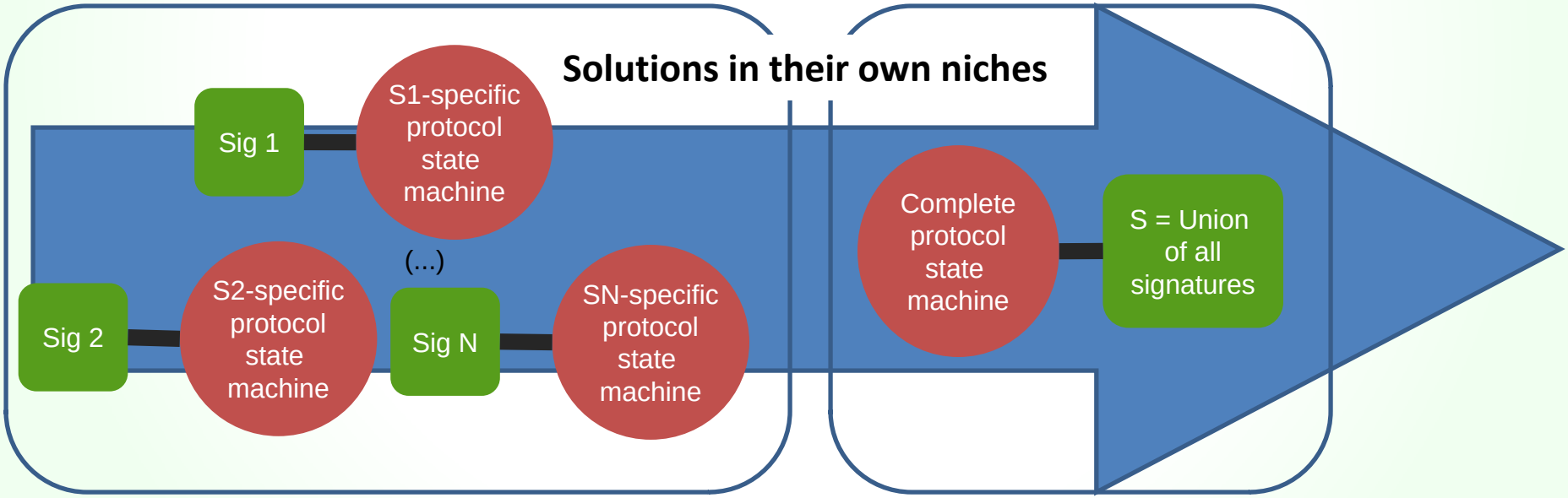
Relationship with Previous Work

Schear's Approach:

- A single signature only uses a small portion of the protocol state machine.
- By customizing the state machine to each signature (removing those elements in the state machine that are irrelevant to the signature), each signature can run much faster .

Our Approach:

- If N is large enough, then the union of Schear's specialized state machines add up to the complete protocol state machine.
- In this case, it pays off to implement one single complete protocol state machine and have all signatures leverage the same machine



Small number of signatures

Large number of signatures