

### 29.3 A 3nm FinFET 2.2Gsearch/s 0.305fJ/b TCAM with Dynamically Gated Search Lines for Data-Center ASICs

Sushil Kumar<sup>\*1</sup>, Gajanan Jedhe<sup>\*1</sup>, Chetan Deshpande<sup>\*1</sup>, Agastya Gogoi<sup>1</sup>,  
Phoebe Su<sup>2</sup>, Kim Soon Jway<sup>3</sup>, TzeYing Seoh<sup>3</sup>

<sup>1</sup>MediaTek, San Jose, CA

<sup>2</sup>MediaTek, Hsinchu, Taiwan

<sup>3</sup>MediaTek, Singapore, Singapore

\*Equally Credited Authors (ECAs)

Packet classification and forwarding are fundamental tasks for data-center network (DCN) components, such as switches and routers, which are used to efficiently manage and direct network traffic. Packet classification involves examining packet headers to identify and apply specific policies, such as access control or quality of service; while packet forwarding determines the next hop for each packet based on routing tables. Ternary content-addressable memory (TCAM) facilitates these two tasks by performing a parallel in-memory search to compare an incoming packet's header against the rules stored in TCAM. TCAM provides fast parallel-lookup functionality making it an indispensable foundation IP for DCN ASICs. TCAM do have one major cost, the parallel all-entry search is power intensive; this impacts the operational efficiency, reliability, and environmental footprint of network devices. To reduce DCN-TCAM power, a 3nm Fin-FET 2.2Gsearch/s 0.305fJ/b TCAM is designed with (1) a dynamically gated search-line (DGSL) architecture that enables a 37.4% power saving and a 46.6% peak-current reduction; (2) an asymmetric-split architecture (ASA) for additional bit-width dependent power savings.

Packet classification relies on a set of rules defined in access control lists (ACL), which are stored in a TCAM in order of priority; so that similar rules are likely to be programmed close to each other as shown in Fig. 29.3.1. A TCAM search finds the highest-priority match; if a search key matches a particular entry, then it is likely that additional matches will be in nearby entries, demonstrating matching-rule spatial locality. A TCAM macro is divided into banks, where each bank represents a priority level. A packet header matching a rule in the lowest-priority bank, *B7*, is unlikely to also match a rule stored in bank *B0*, the highest-priority bank. We leverage this statistically significant application scenario to save power.

The 512-entry TCAM is split into eight 64-entry banks. The 220b search data input is divided into two stages S1 and S2 of 110b each. Each bank is also split into matching-width stages: consisting of a  $64 \times 110$  TCAM array and a local IO (LIO), which is controlled by the local-control (LC) block. The memory-array compare results are provided to the hit generation (HG) block to produce a HIT output. The LIO and LC receive global signals from the global IO (GIO) and global control (GC) blocks, respectively. When clock (CK) and search enable (SR) inputs are active, the search-data input (SDI) is compared against all 512 stored entries to generate HIT[511:0] outputs, where each HIT is high if the SDI *matches* and low if *mismatches* (aka. *miss*) an entry. Read and write operations are the same as SRAM; hence, not shown for brevity.

Figure 29.3.2 illustrates a baseline-TCAM bank design showing the search line (SL) and match line (ML) architecture. The GIO comprises of global search-line (GSL) logic to generate complimentary GSLT and GSLC signals. GIO *G1* and *G2* drive stages S1 and S2 for all banks. GC clock generation is configured to operate both S1 and S2 in parallel. Within each bank, complimentary SL signals, SLT and SLC, distribute the search key to each TCAM cell. Each TCAM cell consists of two 6T-SRAM bit cells X and Y to store a ternary value, as shown in the truth table in Fig. 29.3.2. The search-line to stored ternary-value comparison is implemented with a 4T NMOS stack that drives ML: it is pulled low for a *mismatch* or remains high for *match*. The ML is shared across multiple TCAM cells and is pre-charged high before a comparison begins. The ML connects to the match-line sense amplifier (MSA) to generate the HIT output for each stage. HIT outputs from S1 and S2 are combined using an AND at the input to a flip flop (FF): generating a HIT output with a 2-cycle search latency.

The most-common and worst-case power-consumption scenario is when most entries result in a *miss*. For each compare operation, all SLs toggle and all MLs are charged and subsequently discharged by a *miss*. For the  $512 \times 220$  TCAM, the power breakdown is shown in Fig. 29.3.2: SL (37%) and ML (35%) power are dominant. While many techniques to reduce ML power are well known, an efficient technique to reduce SL power has been difficult to achieve. In this paper, we introduce a dynamically gated SL (DGSL) architecture that enables significant SL power and peak-current reduction.

Figure 29.3.3 shows the DGSL 2-stage pipeline TCAM bank, utilizing 2-stage ML power saving technique. S1 and S2 operate as pipelines, where the S1 comparison result, for each entry, enables or disables subsequent S2 comparisons. Statistically, a majority of TCAM entries result in a *miss*,  $S1HIT = 0$ ; thus, disabling S2 comparisons for those entries to save ML power. The match result on ML is captured by the MSA, which is controlled by the MSA-enable (MSAE) signal. The  $S1HIT$  output from MSA is pipelined to generate EN\_S2ML of

an entry which is used as gating signal for precharge (PRCHG). The gated-precharge (PC) controls the S2-ML, preventing ML toggle if  $S1HIT = 0$ . DGSL TCAM has a 3-cycle search latency due to pipelining.

To enable SL power saving through DGSL architecture, we introduce a  $S1HIT$  wired-or logic that drives a F/F to generate EN\_S2SL gating signal for the search-clock (SCLK) clock-gate (CGSC). All 64  $S1HIT$  outputs from stage 1 are fed to wired-or logic to generate S1MISS and S1NOMISS signals.  $S1HIT$  signals are distributed across an entire bank for each entry; hence, S1MISS is a dynamic node that is pre-charged high, by MSAE low at the start of a search. A weak keeper is used to hold S1MISS high during the evaluation phase, in the absence of a pull-down path. If one, or more,  $S1HIT$  across entries is logic-1, then the S1MISS signal is discharged through the corresponding pull-down NMOS transistors ( $N_{S1\_}$ ); hence, S1NOMISS will be logic-1. Otherwise, if all entries in the bank are a *miss*, then all  $S1HIT$  are logic-0 allowing S1MISS to remain high, which results in S1NOMISS=0. S1NOMISS is fed to  $P1NM$  F/F to generate EN\_S2SL, which acts as a clock gating signal for S2 SCLK for next cycle. If EN\_S2SL=1, then S2SCLK receives SCLK, which enables SLT/SLC to receive GSLT/GSLC values and S2 stage performs a search operation; like S1 it generates S2HIT which is captured by P2 F/F as HIT. If EN\_S2SL=0, then S2SCLK is disabled, setting  $S2\_SLT = S2\_SLC = 0$ , independent of S2\_GSLT/S2\_GSLC. All pipeline FFs ( $P1^*$  and  $P2^*$ ) are clocked with the HIT clock (HCLK). For banks that *miss* for all S1 entries, the S2 SL toggling is prevented with DGSL, thus saving SL power. To achieve above 2GHz operation the number of TCAM entries per bank needs to be small. This allows a greater probability of all misses in a bank thus increasing SL power savings.

Figure 29.3.4 show the timing diagram and simulation waveform captures, where the critical control signals of DGSL TCAM prevent ML and SL toggling of S2, in comparison to the baseline case. DGSL TCAM power varies linearly with the number of banks that miss for all entries, the *all-miss* case. The DGSL scheme achieves an average power savings of up to 37.4%, with all 8 banks having an *all-miss* result. If a bank's S1 has an *all-miss*, then S2SCLK and S2\_SL\* toggling is prevented. If no bank's *all-miss*, then the 2-stage pipelined ML power saving enables 16.7% power reduction.

Figure 29.3.5 illustrates a DGSL TCAM with smaller peak-current compared to the baseline during a search operation, resulting in a lower IR drop. The pie-charts show the break-down of the peak-current across ML, SL, and other components. For the baseline, the ML pre-charge peak-current dominates. The DGSL's peak-current reduction also depends on the number of banks having *all-miss*, as shown by the bar chart in Fig. 29.3.5. For the DGSL TCAM, 2 pie-charts are given: 1) if '*all-miss* in no bank': in this case peak-current is dominated by the SL, due to a reduction in ML peak-current by preventing S2-ML toggle and achieving 19.7% lower peak-current than the baseline; 2) if '*all-miss* in all banks': DGSL prevents SL-toggling, which enables further peak-current reduction to 46.6%; shifting the dominant peak event back to ML, from SL.

Further, on a DCN ASIC, we observe that the total TCAM power is distributed across instances of varying bit widths, up to 220b. To enhance chip-level TCAM power saving, we introduce an asymmetric-split architecture (ASA) for the 2-stage TCAM, which depends on the SDI bit width to achieve additional power savings from the DGSL scheme in Fig. 29.3.5. The  $512 \times 220$  design features an equal 110b width, between S1 and S2 stages. The 110b width dictates performance; it is optimized for a target frequency of 2.2GHz. To maximize for smaller bit-width instances, we allocate a maximum allowable bit width to S2 and the remaining to S1. A representative ASA for a  $512 \times 164$  TCAM instance is shown in Fig. 29.3.5. In a symmetric configuration, two stages would be split into 82b, but with ASA we allocate 110b to S2 and the remaining 54b to S1. For a  $512 \times 164$  2-Stage TCAM, ASA with DGSL achieves a 51% power saving compared to the baseline, which is higher than 37.4% for  $512 \times 220$  while maintaining 2.2Gsearch/s performance and a similar area density.

The TCAM is fabricated in a 3nm Fin-FET technology. Figure 29.3.6 shows a Shmoo plot for a  $512$ -entry  $\times 220$ b DGSL TCAM. We demonstrate 2.2GHz performance from 0.75V at room temperature. Compared to previously published TCAMs [1-5], the DGSL TCAM has the highest reported memory density,  $4.97\text{Mb/mm}^2$ , and search speed, 2.2Gsearch/s. The DGSL TCAM search energy is 0.305fJ/b achieving a 37.4% reduction from the baseline TCAM with similar size and power conditions. Figure 29.3.7 presents the die micrograph and chip summary.

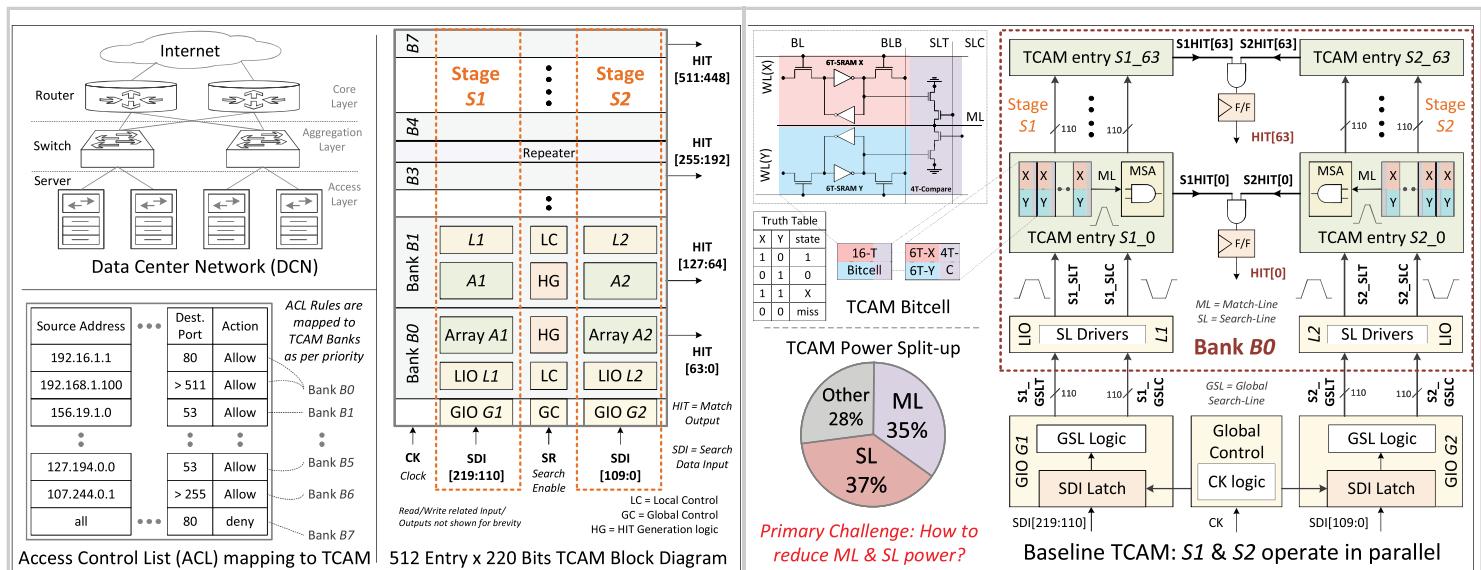


Figure 29.3.1: (Left) DCN with sample ACL rule storage in TCAM and (Right) block diagram of a 512 x 220 TCAM.

Figure 29.3.2: (Top-left) TCAM bit cell with its corresponding truth table baseline. (Bottom-left) TCAM power breakdown. (Right) baseline TCAM architecture.

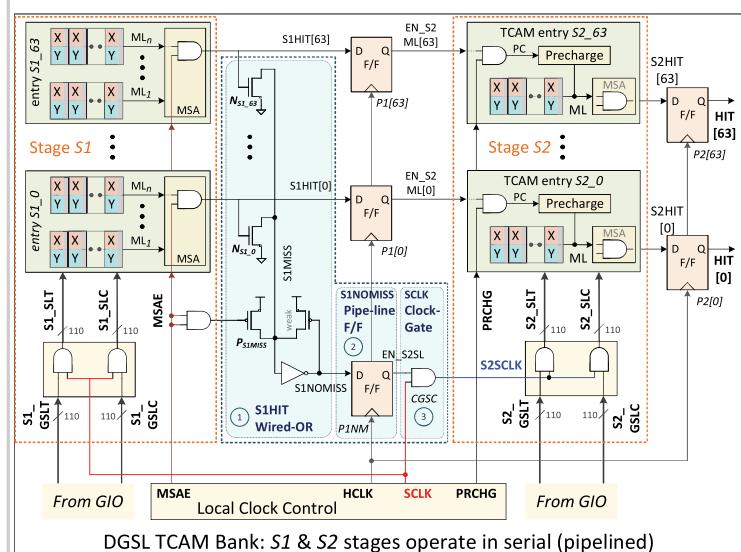


Figure 29.3.3: DGSL TCAM bank with SL power-reduction components: wired-OR logic, a pipeline FF, and clock-gating logic.

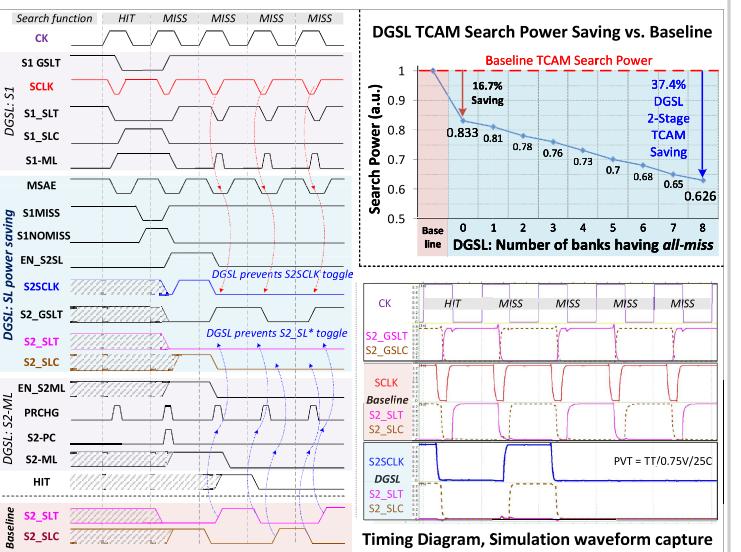


Figure 29.3.4: DGSL TCAM timing diagram, and its search power reduction compared to the baseline case. At-speed captured waveforms.

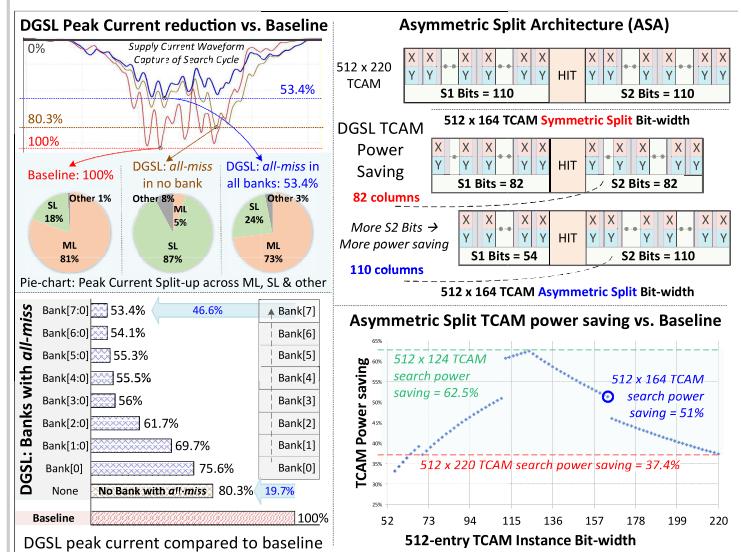


Figure 29.3.5: (Left) DGSL peak current reduction vs. baseline case; (Right) ASA and its power saving vs. bit width.

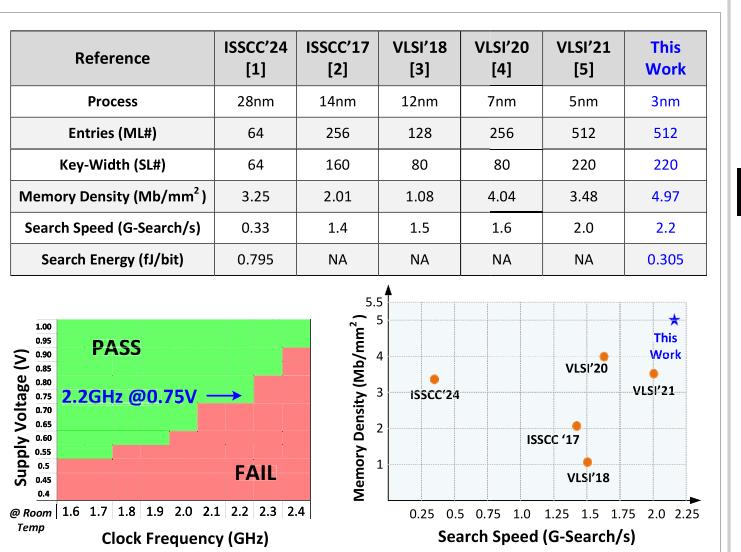


Figure 29.3.6: Comparison to prior work, voltage vs. frequency Shmoo, and plot of memory density vs. search speed.

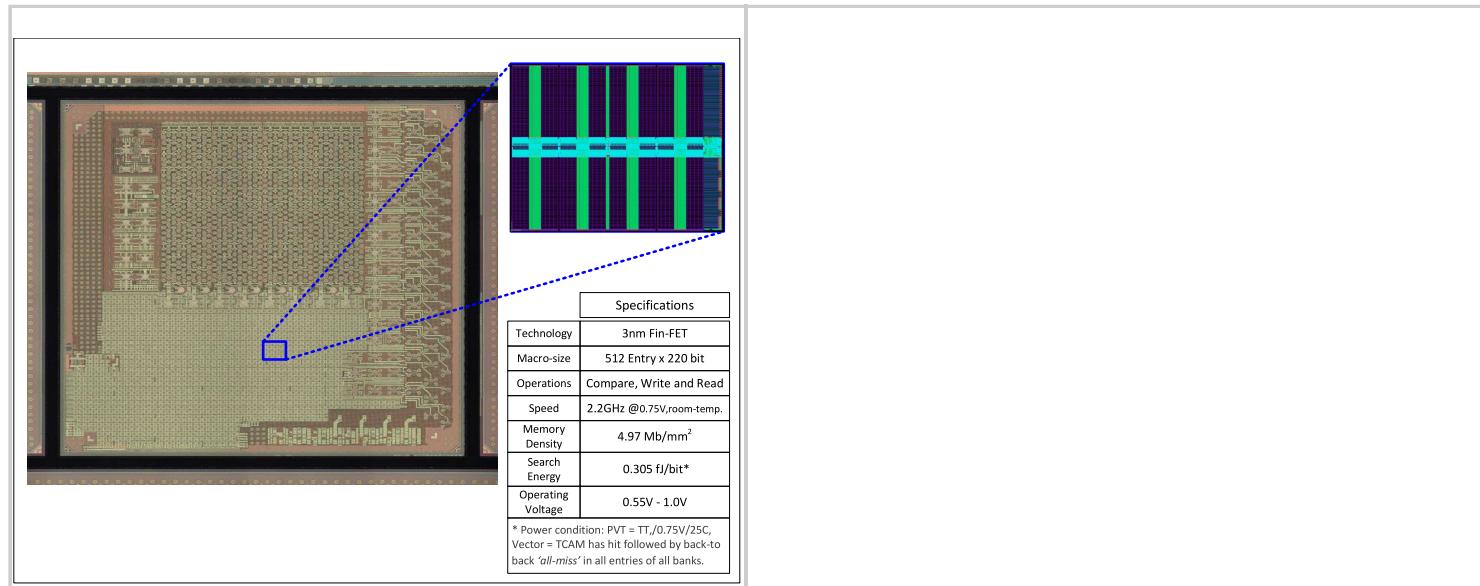


Figure 29.3.7: 3nm chip micrograph, layout view, and design summary table.

#### References:

- [1] Z. Yue et al., "A 0.795fJ/bit Physically-Unclonable Function-Protected TCAM for a Software-Defined Networking Switch," *ISSCC*, pp. 276-278, 2024.
- [2] I. Arsovski et al., "1.4Gsearch/s 2Mb/mm<sup>2</sup> TCAM using two-phase-precharge ML sensing and power-grid preconditioning to reduce Ldi/dt power-supply noise by 50%," *ISSCC*, pp. 212-213, 2017.
- [3] M. Yabuuchi et al., "12-NM Fin-FET 3.0G-Search/s 80-Bit × 128-Entry Dual-Port Ternary CAM," *IEEE VLSI*, pp. 19-20, 2018.
- [4] M. Yabuuchi et al., "A 7nm Fin-FET 4.04-Mb/mm<sup>2</sup> TCAM with Improved Electromigration Reliability Using Far-Side Driving Scheme and Self-Adjust Reference Match-Line Amplifier," *IEEE VLSI*, pp. 1-2, 2020.
- [5] C. Deshpande et al., "A 5nm Fin-FET 2G-search/s 512-entry x 220-bit TCAM with Single Cycle Entry Update Capability for Data Center ASICs," *IEEE VLSI*, pp. 1-2, 2021.