Security Level:

Tableless Routing

www.huawei.com

Author/ Email: Author's name/Author's email
Version: V1.0(20YYMMDD)

HUAWEI TECHNOLOGIES CO., LTD.

Tableless Routing--Background

- IP Routing VS. Tableless Routing

Table-based Forwarding

1. Delay: On-chip Table Lookup costs ~100cycle, Off-chip costs ~n*100cycle

2. Area cost: The routing table is too large to store on-chip memory. A small part of table is stored on-chip, while a large part of table is stored off-chip, which leads to a large access delay.

3. Bandwidth: Table lookup performance becomes the bottleneck of forwarding (Memory wall problem). The future bandwidth improvement becomes much harder as the routing table scales up inside a small chip.

Lower delay

Super-large scale

Much less storage

Super-high bandwidth

Break memory wall

Computation-based Forwarding

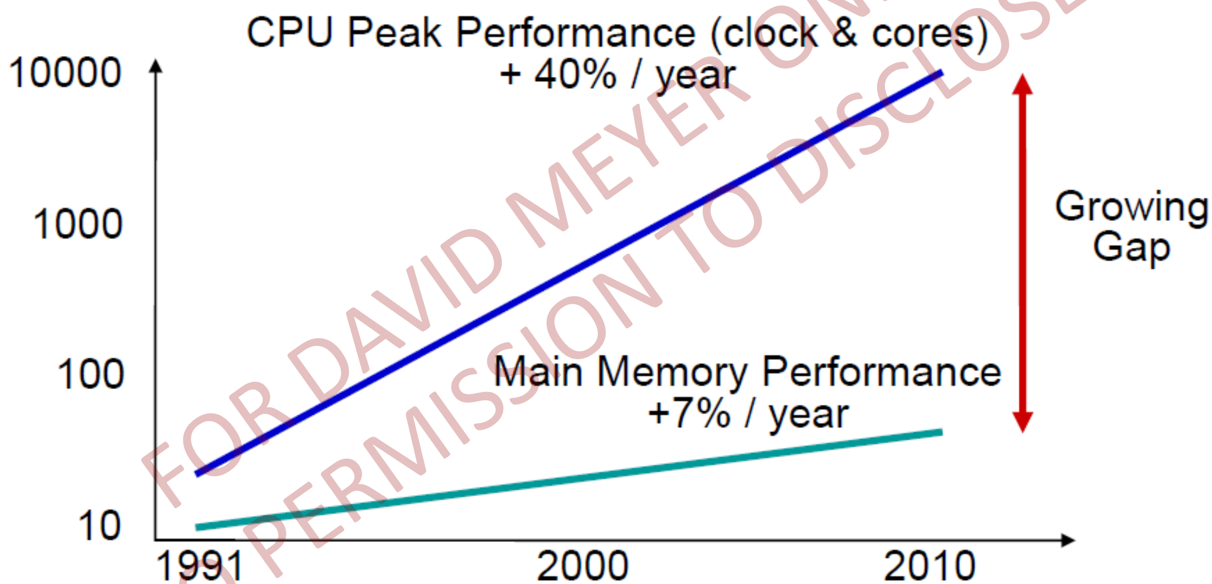
1. Delay: Fixed-logic computation, based on registers (~1-10cycles)

2. Area cost: No table storage. No need to worry about the routing entry scales. The computation complexity and parameter storage is not limited by the routing entry scale. Support supper-large scale routing with small storage.

3. Bandwidth: It becomes a popular trend to design specialized hardware for computation (e.g., TPU/NPU/GPU); Tableless routing will save most storage area of tables, and could use the saved storage area to make more powerful parallel computation pipelines to further improve the forwarding bandwidth.

Tableless Routing--Background

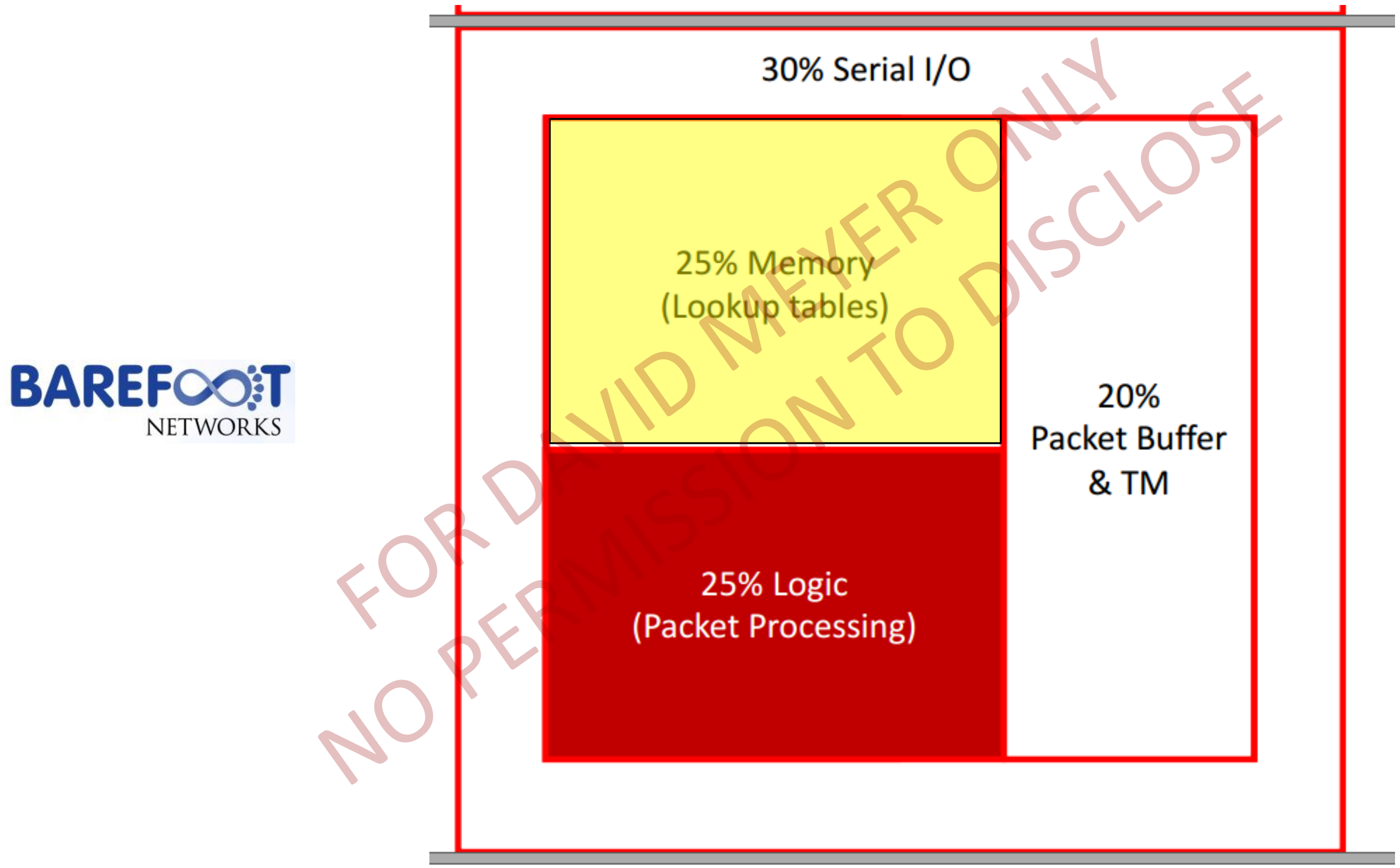
The Memory Wall



Access latency to main memory today up to 300 cycles

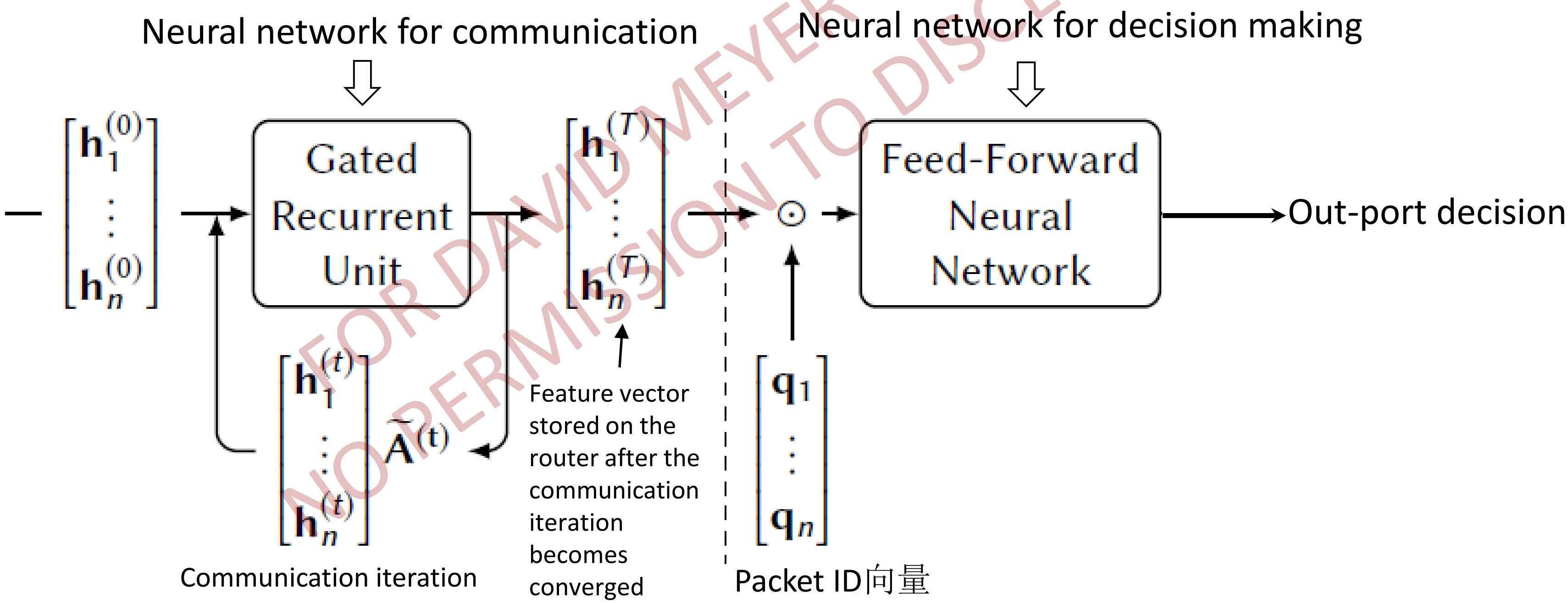
Assume 2 Flops/clock ticks ➔ 600 Flops wasted while waiting for one main memory access!

Tableless Routing--Background



Tableless Routing—GNN framework

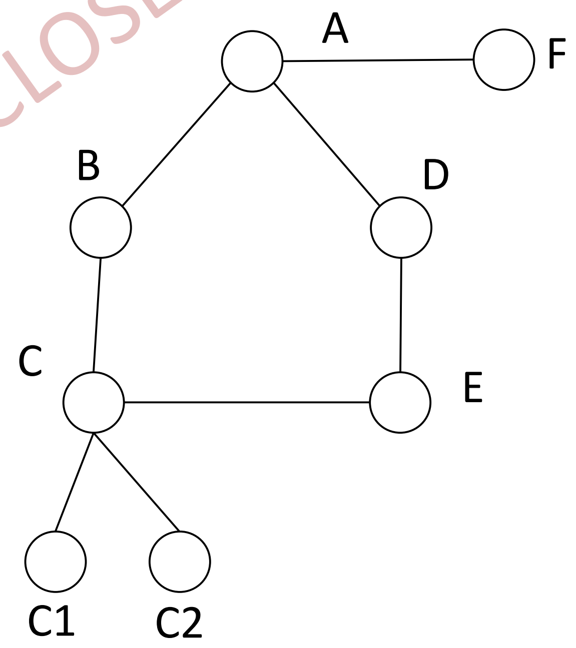
- GNN (Graph Neural Network) framework



Tableless Routing—GNN

- GNN Features:
  - 1. Each node (i.e., router) stores a feature vector
  - 2. Each node updates its feature vector based on the feature vectors of its neighbor nodes:

    $X_i = \sum f_i\,(X_j \mid j \in Neighbour(X_i))$, f_i function is implemented by a neural network (i.e., communication neural network)
  - 3. Fixed-Point theory of GNN: with finite iteration steps, the feature vector can converge to a stable value (i.e., fixed point)
  - 4. The neural networks can be first trained in the controller offline and then deployed on the routers



Tableless Routing—GNN

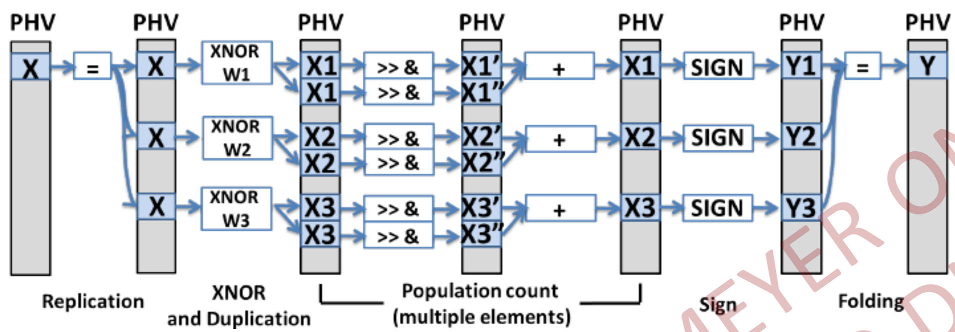
- Advantage of GNN
  - 1. Flexible optimization objectives
    - Shortest path
    - QoS optimization (delay), Load balancing
  - 2. Highly-robust training: leading to a faster convergence
    - Use the dropout training method to simulate random link failures in the topology. Enable a faster convergence speed when link fails.
    - Use the incremental training method to deal with the case of adding new nodes into the network. Enable a fast convergence speed when network scales up.

Tableless Routing

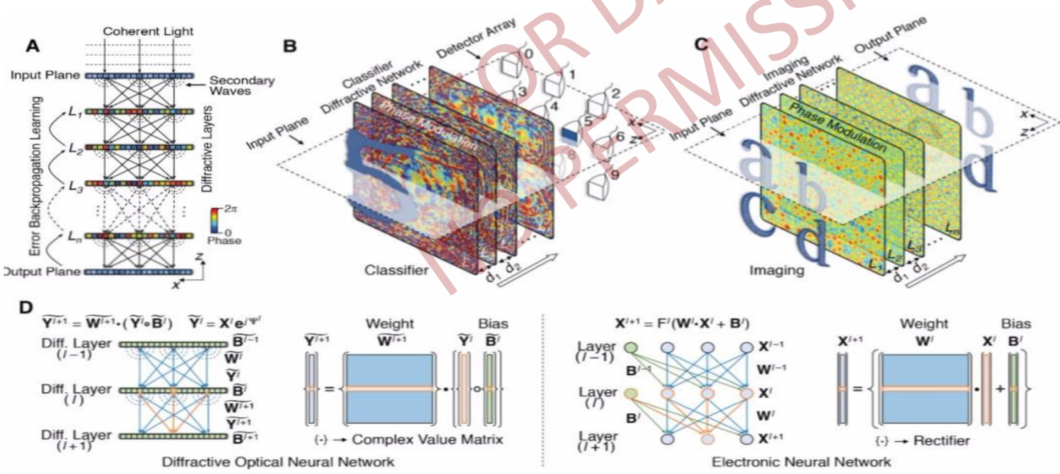
- Key Concepts:
- Packet ID
  - The packet with the same packet ID will follow the same route
  - Packet ID can be set flexibly as destination IP, flow ID, flowlet ID…
- Communication protocol: GNN
- Neural network forwarding
  - Line-speed neural network computation
    - **Recent plan**: two-layer BNN (only XOR&ADD required), supported by existing hardware (P4), with some precision loss
    - **Medium term plan**: three-layer DNN, supported by current specialized AI accelerator (GPU/NPU/TPU/…)
    - **Long term plan**: large-scale DNN, supported by new next-generation computation hardware, such as the Optical Neuromorphic Computing

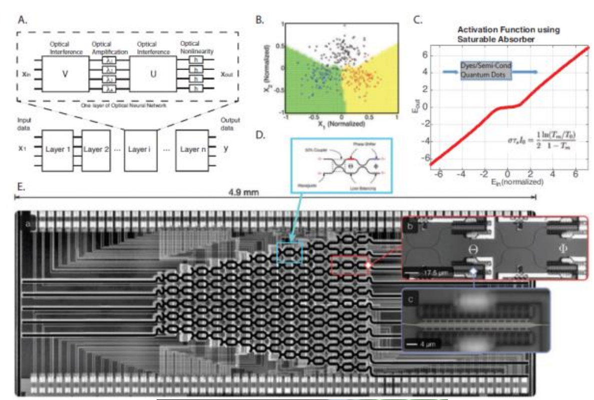
Tableless Routing

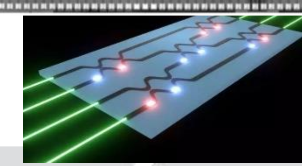
- Recent (Line-speed verified in P4, In-network neural network)



- Medium term: add the AI accelerator, e.g., the Cambrian chip
- Long term: All-optical machine learning using diffractive deep neural networks (Nature 2018)







Tableless Routing– Reliability Challenge

- Simulation results show that, **when trained the GNN with supervised training, its error rate less than 5%.**

- It means that more than 95% pairs of nodes learn the shortest path. However, **the left 5% does not have any guarantee on the reachability of the learned path** (because each node selects the forwarding port independently). **The nature of Neural Network makes such errors inevitable.**

- Our solution: borrowing some techniques of reachability assurance in IoT, **we propose a Tableless reliable forwarding method with 100% reachability guarantee**.

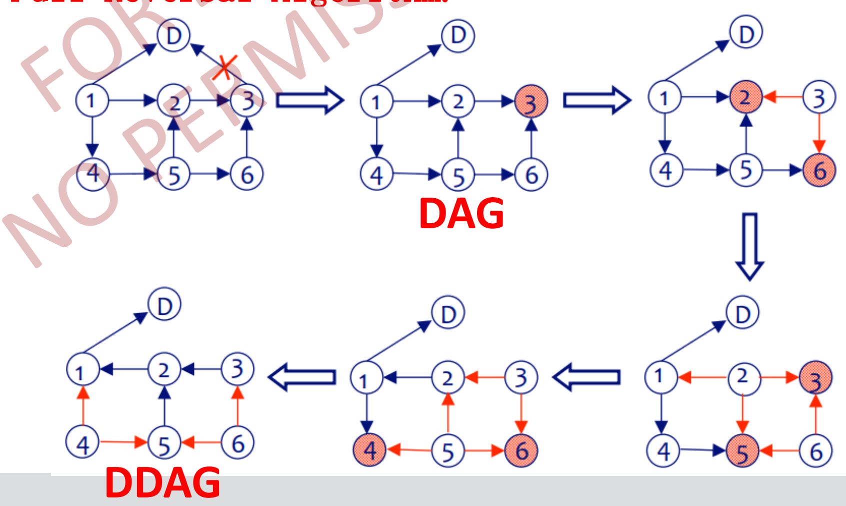
Key Technology- Link Reversal

- DAG: Directed Acyclic Graph
- DDAG: Destination-oriented DAG, i.e., for a destination d, the nodes other than d can follow the link directions to reach d.

- Theorem 1: there exists an algorithm to transform a random DAG to a DDAG by changing the link directions in finite steps.

- Full Reversal Algorithm: let a sink node denote the node with no out-edges, and the destination as d. Select the sink nodes beyond d in the DAG in any order to reverse all its in-edges.

- Theorem 2: The DAG can be always transformed to a DDAG within finite steps following the Full Reversal Algorithm.

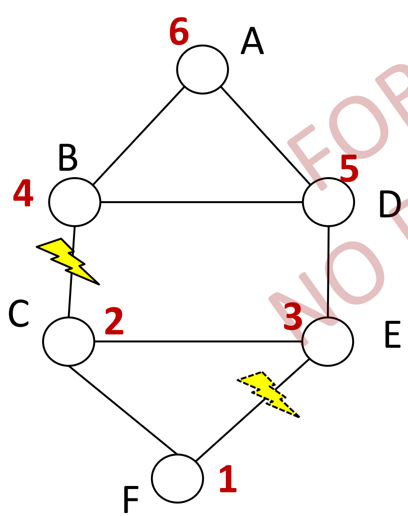


Tableless Routing

- Tableless LRR: implement Link Reversal Routing in a Tableless way
  - Calculate the Node Value by inputting the Packet ID & feature vector into the decision-making neural network
  - Select the next hop based on the Node Value as follows:
    - Step 1: select the neighbor node with a smaller value than the current node value to forward; if there are many, select the one with min value
    - Step 2: if no neighbor node's value is smaller than the current node value, "increase the current node value" as Max(Neighbor Node Value)+1, go to Step 1.

⇩

Push a key-value pair into the packet head,
Key is the node ID, value is the newly-increased value



Example: Suppose a packet ID from A->F,

The figure shows a possible value assignment

1. Default shortest path is A->B->C->F

2. If Node B fails, the routing path is A->D->E->F

3. When a packet arrives at Node B, if link B-C fails,
the routing path becomes A->B (B increases to value 7)->D->E->F

4. When a packet arrives at Node B, if both link B-C and E-F fail,
the routing path becomes A->B (increases to value 7)->D->E->C->F

Tableless Routing-- Summary

- Key features:
- 1. No Table & Computation-based forwarding
- 2. Flexible optimization objectives:
  - 1. shortest-path routing
  - 2. Load balancing, QoS routing…
- 3. 100% reachability guarantee
  - for any number of link/node failures or neural network errors (if the network is still connected after link/node failures)

- Key technologies:
  - GNN
  - Tableless LRR

Thank you

www.huawei.com