**Q-Thermal: A Q-Learning-Based Thermal-Aware Routing Algorithm for 3-D Network On-Chips**

* Cooling techniques in 3-D NoCs locate into two categories: the technology level and the algorithmic/architectural level.In the former case, extra equipment is added to the chip to eliminate hotspots. It is an efficient technique but suffers from high area and manufacturing costs [9]. The latter case proposes a runtime thermal management (RTM) technique to keep the system temperature below a certain thermal limit.
* The thermal management techniques [10] in the algorithmic/ architectural level fall into two categories: reactive and proactive.

**The reactive techniques:**  The reactive techniques temporarily either stop or throttle

the activity of the overheated routers. In these techniques, for each router, a thermal sensor is embedded into the chip [11]. As soon as the temperature of a router exceeds a thermal

threshold, the activity of the router is stopped (or limited) until the router’s temperature falls down again

EG: TAAR -Topology aware adaptive routing

**Proactive techniques:**  In the proactive techniques, creating hotspots and reaching to the critical thermal condition are prevented via subtle and smart routing techniques. In fact, the routing mechanism is

done in a way that the traffic load is transferred to the cooler parts and the heat is balanced across the layers of the chip. As a result, none of the routers may reach the thermal threshold and forced to stop their activity.

EG: PTBR : Proactive thermal budget- based beltway routing algorithm

**BACKGROUND ON Q-LEARNING**

Q-learning is a values-based reinforcement learning algorithm that is employed to make the best possible decision in multichoices environments.

* This algorithm learns a policy that helps agents to act optimally in a specific environment. The agent passes several states in its traverse. Each state offers some actions with different costs. The agent tries to improve the quality of its actions in each state using reactions received from the environment.

Here’s your formatted text for pasting into Word:

\*\*Q-learning\*\* is a foundational algorithm in reinforcement learning designed to help agents make optimal decisions in an environment by learning a policy over time. The agent explores multiple states, each offering a set of actions associated with different costs or rewards. The goal of the agent is to improve its actions' quality in each state by receiving feedback from the environment.

In Q-learning, after the agent takes an action \(a\) in state \(s\), it receives a reward \(r\) and moves to the next state \(s'\). The Q-value, denoted as \(Q(s, a)\), represents the estimated utility or cost of performing action \(a\) in state \(s\). Initially, all Q-values are set to predefined values, and the learning process involves updating them based on the following temporal difference rule:

\[

Q[s, a] \leftarrow Q[s, a] + \alpha \left( r + \gamma \cdot \max\_{a'} Q[s', a'] - Q[s, a] \right)

\]

In this equation:

- \(r\) is the reward received for moving from state \(s\) to \(s'\).

- \(\max\_{a'} Q[s', a']\) is the maximum estimated future reward from state \(s'\).

- \(\alpha\) is the learning rate (0 < \(\alpha\) < 1), which controls how quickly the agent adapts to changes. In deterministic environments, setting \(\alpha = 1\) leads to faster convergence. In stochastic environments, a smaller \(\alpha\) provides more accurate solutions but requires more iterations to converge.

- \(\gamma\) is the discount factor, giving weight to future rewards.

* The Q-learning algorithm relies on a Q-table to store Q-values for each state-action pair. Initially, exploration strategies (random actions) are employed to gather data about the environment. As learning progresses, the agent begins to exploit the learned Q-values by selecting actions that lead to higher rewards based on the Q-table.

***Q-Routing***

* Q-routing utilizes a Q-table to manage routing decisions in a 2-D NoC, where actions correspond to possible output channels for packet transmission.
* The algorithm updates Q-values based on latency information exchanged between routers, allowing for dynamic adjustments to routing paths based on current network conditions.
* The proposed thermal-aware routing method for 3-D NoCs builds on Q-routing principles, incorporating thermal information into the decision-making process.

**Q-LEARNING-BASED THERMAL-AWARE ALGORITHM (Q-THERMAL)**

The algorithm is detailed through five parts:

1. how to collect thermal information across the network;
2. how to do thermal-aware lateral routing;
3. how to update thermal information across the network;
4. how to do vertical proactive routing; and finally,
5. how to add a free deadlock mechanism to the thermal-aware routing algorithm.

**Collecting Thermal Information**

* The proposed method collects thermal data during packet transitions, allowing routers to update their Q-values based on the average temperatures experienced along packet paths.
* By embedding temperature information in packet headers, the need for separate learning packets is eliminated, reducing network overhead.
* This approach enables routers to make informed routing decisions based on real-time thermal data, enhancing overall thermal management.

**Thermal-Aware Lateral Routing**

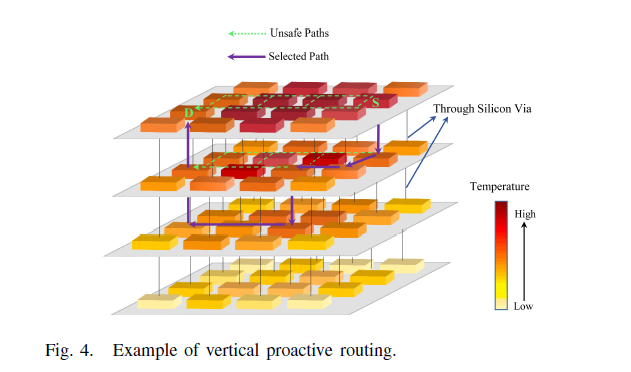
* Routers select output channels based on Q-values, choosing paths with lower temperatures to optimize thermal distribution.
* If Q-values indicate initial temperatures, the routing decision is made randomly, reflecting the learning phase of the algorithm.
* The pseudocode for this routing technique illustrates the decision-making process based on thermal awareness.

**Updating Thermal Information**

* Routers update their Q-values based on the average temperature of packets they handle, ensuring that routing decisions reflect current thermal conditions.
* The algorithm accounts for local temperatures and the number of hops a packet has taken, facilitating accurate thermal information propagation.
* This continuous updating process helps maintain effective routing strategies as thermal conditions change.

**Vertical Proactive Routing**

* The method emphasizes the importance of vertical routing to manage thermal loads effectively, allowing packets to detour to cooler layers when necessary.
* By utilizing TSVs, the routing algorithm can select paths that avoid overheating routers, enhancing thermal safety.
* The routing decisions are influenced by the average temperature, with a higher likelihood of selecting lower layers as temperatures rise.



**Deadlock Freedom**

* The proposed routing method incorporates virtual channels to prevent deadlocks, ensuring smooth packet transmission without network congestion.
* Two distinct virtual channels are used based on the relative positions of source and destination nodes, avoiding problematic routing turns.
* This design enhances network performance by maintaining efficient packet flow and reducing the risk of deadlocks.

**Summary of Q-Thermal**

* The Q-Thermal method stabilizes temperatures across various paths, enabling routers to direct packets toward cooler regions and prevent hotspots.
* The approach leverages recorded temperature data to inform routing decisions, enhancing thermal management in 3-D NoCs.
* The method demonstrates significant improvements in thermal distribution and network performance compared to existing techniques.

**EXPERIMENTAL RESULTS :**

Simulations are carried out through the traffic-thermal co-simulation platform, which uses PAT-Noxim

**Analysis of Thermal Distribution**

* Q-Thermal effectively redistributes traffic away from overheated regions, improving thermal (standard deviation of thermal distribution) distribution by 28% compared to TAAR and 13% compared to PTB3R.
* Q-Thermal improves the standard deviation of thermal distribution by approximately 28% and 13% in comparison with TAAR and PTB3R, respectively.
* The method continuously monitors router temperatures, allowing for proactive adjustments to routing strategies based on real-time data.
* This dynamic approach helps mitigate the formation of hotspots and enhances overall network reliability.

**Analysis of Traffic Load Distribution (statistical traffic load distribution (STLD)**

* Q-Thermal exhibits a slightly denser traffic load in lower layers compared to TAAR, as it proactively steers traffic toward cooler regions.
* The use of additional virtual channels in Q-Thermal improves routing density and adaptability, contributing to better overall performance.
* The analysis highlights the importance of balancing traffic distribution to maintain efficient network operations.

**Analysis of Network Performance (average latency of the network)**

* Q-Thermal consistently achieves lower network latency across various traffic patterns, outperforming TAAR and PTB3R due to its efficient routing strategies.
* The method's ability to update Q-tables in real-time enhances its responsiveness to changing network conditions, reducing delays associated with periodic reconfigurations.
* Overall, Q-Thermal demonstrates a 32% improvement in performance metrics compared to existing routing techniques.

**D. Number of Thermal Hot Spots**

* The implementation of Q-Thermal significantly reduces the number of thermal hotspots, achieving a 38% reduction compared to TAAR and 54% compared to PTB3R.
* This reduction in hotspots contributes to improved network availability and reliability, enhancing the overall performance of the 3-D NoC.
* The findings underscore the effectiveness of the proposed thermal management strategy in maintaining optimal operating conditions.

**Overhead Costs**

* Q-Thermal incurs a slight increase in layout area and power consumption compared to other techniques, primarily due to the Q-table and additional virtual channel.
* Despite these overheads, the benefits of improved thermal distribution and network performance justify the additional costs.
* The analysis indicates that the trade-offs associated with Q-Thermal are acceptable given its significant performance enhancements.
* **Q-Thermal** increases the layout area by 7% and 11% in comparison with **TAAR** and **PTB3R**, respectively.

**CONCLUSIONs**

* The Q-Thermal method effectively records and utilizes temperature data from packet paths to optimize routing decisions in 3-D NoCs.
* The approach leads to better thermal balance, reduced hotspots, and improved network performance compared to previous methods.
* Overall, Q-Thermal presents a promising solution for enhancing thermal management in future 3-D NoC architectures.

Limitations:

**Increased Area and Power Consumption**: One of the notable limitations of the Q-Thermal method is that it slightly increases the area and power consumption compared to previous routing techniques. This increase is primarily due to the implementation of Q-tables and additional virtual channels, which may not be ideal for all applications, especially in resource-constrained environments