MPAM: Reliable, Low-Latency, Near-Threshold-Voltage Multi-Voltage/Frequency-Domain Network-on-Chip with Metastability Risk Prediction and Mitigation

Chuxiong Lin[1], Weifeng He[1], Yannan Sun[1], Lin Shao[1], Bo Zhang[2], Jun Yang[3], Mingoo Seok[2]

[1]Shanghai Jiao Tong University, Shanghai, China, [2]Columbia University, New York, USA, [3]Southeast University, Nanjing, China

Emerging applications like a drone and an autonomous vehicle require system-on-a-chips (SoCs) with high reliability, e.g., the mean-time-between-failure (MTBF) needs to be over tens of thousands of hours [1]. Meanwhile, as these applications require increasingly higher performance and energy efficiency, a multi-core architecture is often desirable. Here, each core operates in an independent voltage/frequency (V/F) domain, ideally from the near-threshold voltage (NTV) to super-threshold, while communicating with one another via a network-on-chip (NoC) [2]. However, this makes it challenging to ensure robustness in clock domain crossing against metastability. Metastability becomes even more critical to NTV circuits since metastability resolution time constant τ grows super-linearly with voltage scaling [3]. Conventionally, an NoC uses multi-stage (4 stages in [4]) synchronizers to improve MTBF, but they increase latency and cannot completely eliminate metastability. Recently, [5] proposed a novel NTV flip-flop, which has a lower probability of having metastability. Another recent work [6] proposed to detect the *necessary condition* of metastability and mitigate it by modulating the RX clock and also requesting retransmission to guarantee data correctness. However, as it detects a necessary condition, not actual metastability, it tends to overly request retransmission, hurting latency, throughput, and energy efficiency.

In this work, we propose a new technique titled ***m***etastability risk ***p***rediction ***a***nd ***m***itigation (MPAM). MPAM uses three conterminous clocks with different phases to ***predict*** if a synchronizer will have a metastability condition soon and adjusts the clock phase ***on the fly before*** the next clock edge, which prevents the synchronizer from having metastability condition, let alone metastability. Moreover, it allows us to use a single flip-flop *in lieu of* a multi-stage synchronizer, improving the latency, throughput, and energy efficiency of a router. We prototyped a 2-by-2 NoC test chip in a 40-nm LP. With MPAM enabled, the NoC has the bit error rate (BER) of $<10^{-13}$, reducing the metastability condition rate by at least ten orders (theoretically increasing MTBF by the same orders). This complete prevention of metastability conditions allows us to use a single flip-flop as a synchronizer, improving the latency by >34.5% and throughput by >7.4% over the MPAM-disabled NoC baseline. The MPAM's area overhead is only 3.4%.

Fig. 1 shows the proposed 2-by-2 NoC, containing four processing elements (PE0 to PE3) and four routers (R0 to R3). Each pair of a PE and a router (and their local clock generator) shares the same V/F domain. A source PE sends a packet, which travels several routers to a destination PE. We can formulate the typical packet latency per router as $t_{ROUTER} = t_{FIFO} + t_{OUT}$, where t_{FIFO} is the latency of a dual-clock FIFO and t_{OUT} is that of the router datapath (Fig. 1 top right). t_{OUT} is two clock cycles in our design. On the other hand, with N-stage synchronizers, t_{FIFO} can be formulated as $t_{FIFO} = 1 + N$, where the "1" is for the write pointer logic latency. This makes $t_{ROUTER} = 3 + N$. Note that the synchronizer's latency (N) takes a large portion of the total router latency (40% if N=2, 57% if N=4).

To increase the MTBF against metastability, the FIFO integrates two MPAMs, MPAM1 for reading and MPAM2 for writing (Fig. 1 bottom). Let us use MPAM1 for the following description (Fig. 2). To predict if it will have a metastability condition soon, MPAM1 monitors the edge of the write clock *wclk* to the read clock *rclk*. Specifically, it firstly takes *rclk* and generates three conterminous clocks, i.e., $rclk_{lead}$ with the smallest, $rclk_{int}$ with the intermediate, and $rclk_{lag}$ with the largest phase addition. Then, it compares the edge of the divided version of *wclk* (*wclk_div*) with the edges of the three delayed clocks using a 2-b phase detector. The detector is based on in-situ error detectors with M stages (ED, [6]) (Fig. 2 top left), producing 00 if *wclk_div* is earlier than $rclk_{lead}$ or later than $rclk_{lag}$, 01 if *wclk_div* is between $rclk_{lead}$ and $rclk_{int}$, or 10 if it is between $rclk_{int}$ and $rclk_{lag}$. If the detector outputs 01 or 10, we can predict that the FIFO will have a metastability condition (i.e., the edges of *wclk* and *rclk* get close) (Fig. 3 top). Once a metastability condition is predicted to happen, MPAM1 selects one of $rclk_{lead}$, $rclk_{int}$, or $rclk_{lag}$ as the synchronizing clock *srclk* such that the new *srclk* has the sufficient phase difference to *wclk,* and thereby evading to have the metastability condition. Note that this *pre*-metastability-condition mitigation differs from the prior work [6]. The prior work lets the metastability condition to occur while this work predicts the metastability condition and avoid it before it actually happens.

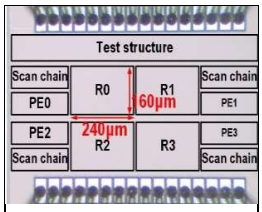

Fig. 7. Die micrograph.

The above process of the prediction and avoiding needs to be completed timely in case *wclk* approaches the next edge of *srclk*. However, the phase detector with the M stages incurs the prediction latency of M+1 cycles. (Note that we use the M stages to avoid the metastability of the *err* signal.) Therefore, we need to add a sufficient amount of phase difference ΔP between $rclk_{lead}$ and $rclk_{int}$ and between $rclk_{int}$ and $rclk_{lag}$. Specifically, we have found that ΔP should be larger than $(M+1)\cdot D_e/D$, where D_e is the accumulated phase shift between *srclk* and *wclk* over D cycles [7].

Fig. 3 middle gives an exemplary case, where *wclk* and *rclk* are plesichronous, having 100-ppm frequency mismatch, i.e., the f_{wclk} to f_{rclk} ratio *k* is 0.9999. This *k* leads to $D_e/D=2\pi E\text{-}4$ and as we use M=3, we set ΔP to be larger than 8πE-4. The phase difference between *wclk_div* and delayed *rclk*s, would gradually decrease to 0 and return to 2π and repeat it periodically. At $t=t_1$, the edge of $rclk_{lead}$ gets too close to that of *wclk_div* by less than a detection window W_I (❶). The phase detector predicts a metastability condition would happen if the current *rclk* keeps being used, so we select $rclk_{lag}$ for *rclk* (❷) since it has the largest phase difference with *wclk_div*. Later on, at $t=t_2$ (❸), the phase detector finds $rclk_{lag}$ is going to cause the metastability condition and thus chooses $rclk_{lead}$ for *srclk* (❹). At $t=t_3$, we take $rclk_{int}$ as *srclk* (❺, ❻). This process ensures that *srclk* maintains a good phase margin with respect to *wclk*, preventing the metastability condition. Fig. 3 bottom shows the case for *k*=2.4999. While the frequency difference is larger, MPAM can still deal with it.

We prototyped the NoC test chip in a 40nm LP CMOS (Fig. 7). Each router takes 0.038 mm^2, and the area overhead of MPAMs is 3.4%. We first measured the rate of metastability *condition* in the NoC when disabling MPAM and enabling MPAM, respectively. To do so, we set PE0 to receive packets from other PEs. PE0 operates at 1MHz (200MHz) and 0.4V (1.0V) while other PEs runs at 13 different clock frequencies from $¼\cdot f_{RX}$ to $4\cdot f_{RX}$ (f_{RX} is the frequency of PE0) and different VDDs. The measurement result shows that when enabling MPAM, the NoC experiences neither metastability condition nor error for 10^{13}-bit data transmission (Fig. 4 top). This indicates MPAM can reduce the metastability condition rate by $>10^{10}$ times as compared to the MPAM-disabled case. We further measured the performance of MPAM across W_I delay variations, f_{TX} variations, and voltage scaling (Fig. 4 bottom), and confirmed its robustness.

Finally, by leveraging this excellent metastability prevention capacity, we aim to replace the multi-stage synchronizer with only one flip-flop. We have found that a 3 to 8-stage synchronizer is needed for the MTBF of 10^5 hours (Fig. 5 top left). The MPAM-enabled NoC configured with the 1-stage synchronizers (i.e., flip-flops) can reduce the packet latency by 34.5% (58.0%) over 3(6)-stage synchronizer at 10% data injection rate (Fig. 5 top right). At a higher injection rate, packets are congested, and the latency is not dominated by synchronizers and the throughput becomes more important. At 100% injection rate, MPAM improves the throughput by 7.4% (13.4%) over 3(6)-stage synchronizer (Fig. 5 bottom left) since MPAM can quickly grant the FIFO access with fewer synchronizer cycles. It also benefits energy efficiency across VDDs (Fig. 5 bottom right). Fig. 6 compares prior works, showing the MPAM largely improves the metastability condition rate and latency over the state of the arts.

References:

[1] E. Petritoli *et al.*, Sensors, 2018. [2] J. P. Cerqueira *et al.*, JSSC, 2020. [3] L. Clemenz *et al.*, JSSC, 1995. [4] C. Liu *et al.*, A-SSCC, 2017. [5] G. Shin, *et al.*, CICC, 2020. [6] C. Lin *et al.*, ISSCC, 2020. [7] W. J. Dally, *et al.*, ASYNC, 2010.

2022 IEEE Custom Integrated Circuits Conference (CICC) | 978-1-6654-0756-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/CICC53496.2022.9772849

978-1-6654-0756-4/22/$31.00 ©2022 IEEE

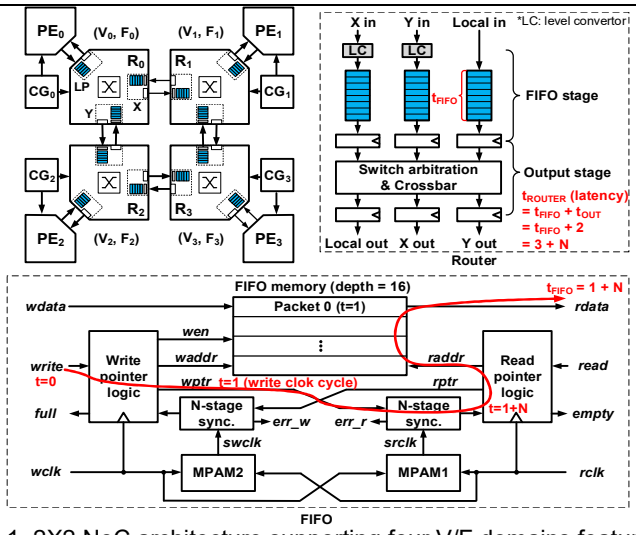

Fig. 1. 2X2 NoC architecture supporting four V/F domains featuring the proposed MPAM-based FIFOs. The router latency t_{ROUTER} is formulated to t_{ROUTER} = 3 + N, where N is the stage count of a synchronizer.

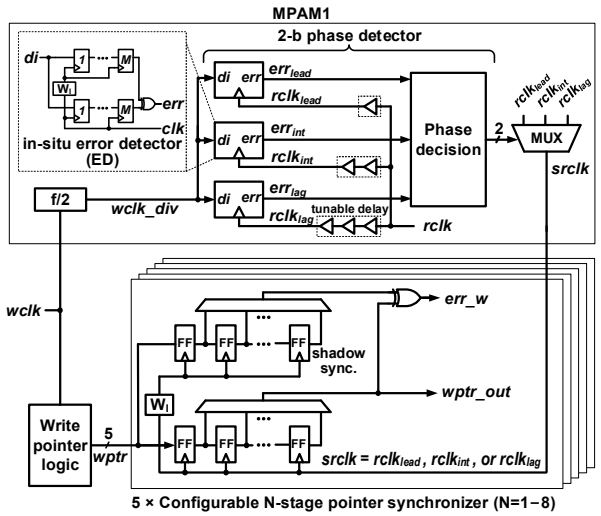

Fig. 2. The MPAM1 circuits with the configurable N-stage pointer synchronizers. The pointer synchronizer contains an extra shadow synchronizer for detecting the metastability condition during pointer synchronization.

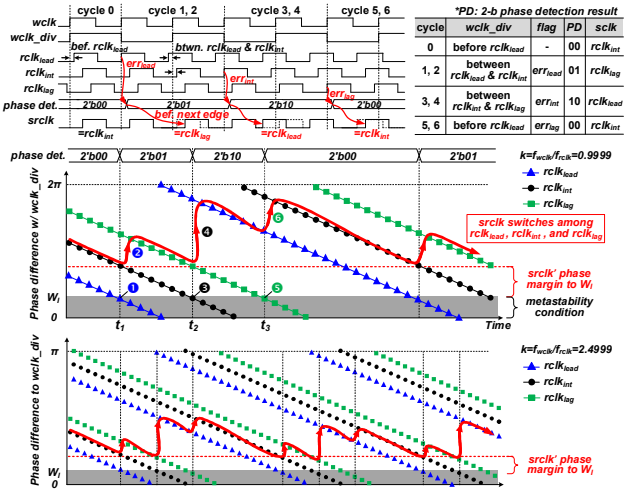

Fig. 3. The MPAM uses three clocks with different clock phases for predicting an imminent metastability risk. Once it detects such a risk, it performs a *pre*-metastability-condition mitigation process, where it chooses one of the three clocks as the synchronizing clock *srclk* so that *srclk* has the largest phase margin.

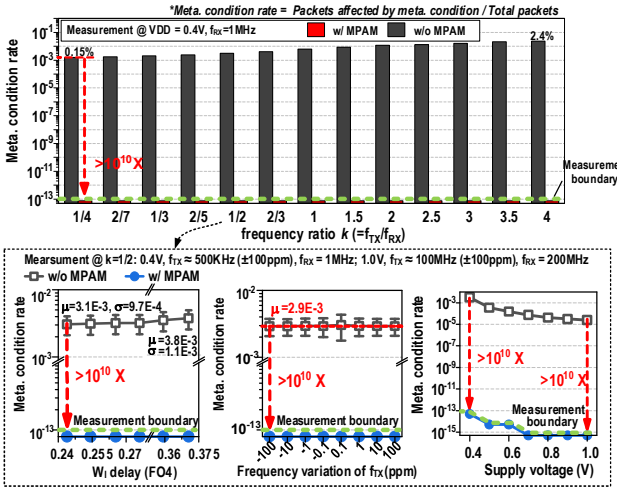

Fig. 4. Measured metastability condition rates across different frequency ratios ($k=f_{TX}/f_{RX}$). MPAM can mitigate metastability condition rate by >10^{10}X. It maintains the mitigation performance across detection window (W_i) variations, 100 part-per-million (ppm) frequency variations, and supply voltage variations.

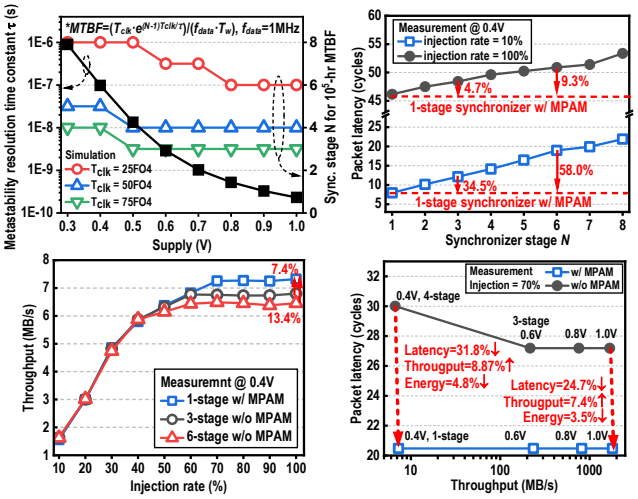

Fig. 5. The metastability resolution time constant τ and the required synchronizer stages to achieve the MTBF of 10^5 hours. We simulated it across supply voltages and clock periods (T_{clk}). Measurement results show the MPAM reduces packet latency and improves network throughput for VDD=0.4V-1V.

	S. Paul, VLSI'13	K. J. Lee, JSSC'16	C. Lin, ISSCC'20	This work
Technology	22nm Tri-gate	65nm	65nm LP	**40nm LP**
NoC	2×2	4×4	2×2	**2×2**
V/F domains	1	4	4	**4**
Operating range	0.34V – 0.85V	0.65V – 1.2V	0.5V – 1V	**0.4V – 1V**
Max clock frequency	151MHz – 1GHz	50MHz – 250MHz	7.3MHz – 175MHz	**2.5MHz – 416.7MHz**
Router power (mW)	1.3 – 28.5	-	0.07@0.5V	**0.046 – 10.58**
Design goal	Timing error	Packet congestion	Metastability	**Metastability**
Proposed scheme	Error detection	Communication pattern prediction	Metastability condition detection	**Metastability risk prediction**
	Flit replay	Intelligent task scheduler	Post-metastability clock modulation	**Pre-metastability clock modulation**
Metastability condition rate reduction	-	-	500 – 1600X @ 0.5V 6.5 – 40X @ 1V	**>10^{10} X**
Latency reduction	Flit replay increases latency	-	Retransmission increases latency	**34.5% (over 3-stage)**
				58.0% (over 6-stage)
Throughput improvement	-	31%	19.5%*	**7.4% (over 3-stage)**
				13.4% (over 6-stage)
Energy efficiency improvement	14.6%	-	21.1%*	**3.5% (over 3-stage)**
				8.6% (over 6-stage)
Area overhead	2.8%	-	4.4%	**3.4%**

* Compared to a baseline with the conservative data retransmission scheme

Fig. 6. Comparison table. MPAM largely advances the reliability of NoC against metastability while improving latency, throughput, and energy efficiency.