A Methodology for the Characterization of Process Variation in NoC Links

Carles Hernández, Federico Silla, and José Duato
Universidad Politécnica de Valencia, Grupo de Architecturas Paralelas, Departamento de Informática de Sistemas
Camino de Vera s/n 46022 Valencia, España
Email: carherlu@gap.upv.es fsilla@disca.upv.es jduato@disca.upv.es

***Abstract*—Associated with the ever growing integration scales is the increase in process variability. In the context of network-on-chip, this variability affects the maximum frequency that could be sustained by each link that interconnects two cores in a chip multiprocessor. In this paper we present a methodology to model delay variations in NoC links. We also show its application to several technologies, namely 45nm, 32nm, 22nm, and 16nm. Simulation results show that conclusions about variability greatly depend on the implementation context.**

I. Introduction

Chip Multiprocessors are the common choice for current processor manufacturers to increase the performance of their designs while keeping their energy consumption inside the power budget. Instead of increasing the frequency of a single processor, the number of cores in a die is increased. An example of this trend is a prototype chip consisting of 80 cores by Intel [6]. In this case, because it is not possible to interconnect such amount of cores by using a bus or a crossbar due to scalability concerns, a network-on-chip (NoC) is used. In general, both academia and industry agree that NoCs are the best option for interconnecting a high number of cores [10], which will be the case as VLSI technologies continue leveraging larger integration scales.

Unfortunately, reducing transistor size does not only allow placing many computing cores in a single die, but also introduces a side effect concern: process variation. Process variation causes that the behavior of devices in a manufactured chip does not match the initial design. Device variability arises as one of the most important challenges to tackle in the development of new on-chip system architectures as technology scales down from 65nm to 16nm [11].

Variability sources can be divided into front-end and back-end ones. The front-end phase of the IC fabrication process is related with the steps involved in the creation of devices whereas the back-end stage comprises steps involved in the wiring definition. The main sources of front-end variation are L_{gate} (transistor channel length) and threshold voltage variations. These variations are mainly a consequence of deviations introduced in the photolitographic process and due to random dopant fluctuations, respectively. On the other hand, back-end variation sources are capacitance and resistance variations, that arise as a consequence of variations in wire dimensions, defects introduced by the Chemical Metal Planarization (CMP) process, etc.

Many recent works analyze the impact of process variation in NoCs. Most of them focus on the impact of within-die variation in devices without considering variations in the links [5][9]. However, links in the NoC are also affected by variability, although they use semi-global metal layers, where metallizations are much wider than in lower layers, and therefore they should be much less affected by process variation than active devices located on the silicon surface. Unfortunately, the effect of variability on NoC links is not negligible. On one hand, although there are examples of repeaterless NoC links [7], they are usually repeated interconnects as the one shown in Figure 1, and thus they will also suffer from L_{gate} variations and dopant fluctuations in the repeater transistors. On the other hand, NoC links will also suffer from the variability introduced by the CMP process that causes surface imperfections because of dishing and erosion. Actually, this process may be a very important source of timing variability [1].

Process variation in NoC links causes that links in the network present different delays, despite that they were initially designed to be identical. Thus, some links will not be able to switch at the intended frequency, reducing performance. This will be the case when NoC operating frequency is set by link delay. In this regard, recent implementation works on NoCs have proved that the critical path of the network is rapidly moving from switches to inter-switch links as technology scales below 65nm [16]. Additionally, it is expected that clock cycle in future switches would be reduced as a consequence of pipelining them because of a more complex switch architecture, due to, among other reasons, the use of virtual channels in order to support coherency protocols.

Characterizing process variability in NoC links is therefore becoming mandatory in order to later explore architectural and technological approaches to minimize its impact on network performance. Some studies have already partially analyzed variability in NoC links. For example, in [1] the authors identify the main sources of variability in NoC links and provide an analytical expression of timing variability from the variation of parameters involved in the interconnect delay. However, this study lacks the influence of spatial features of variability in NoC links. Other studies [9][5] present similar deficiencies. In general, none of them present a global analysis of process variation in NoC links.

In this paper we present a methodology for characterizing within-die process variation in NoC links. This methodology starts from a given NoC layout and additional technology information and provides the delay for every link in that NoC. Although this methodology is based on many previous analysis of process variability and some other mathematical artifacts, its novelty is based on the fact that, up to our knowledge, none of the previous works has proposed such a comprehensive analysis of process variation in NoC links and none of them has glued together so many different sources of process variation as in the proposed methodology. This characterization is intended to later help computer architects

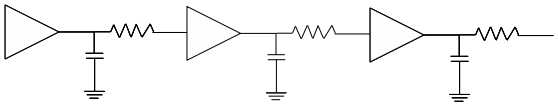


Fig. 1. Diagram of a repeated link composed of three sections.

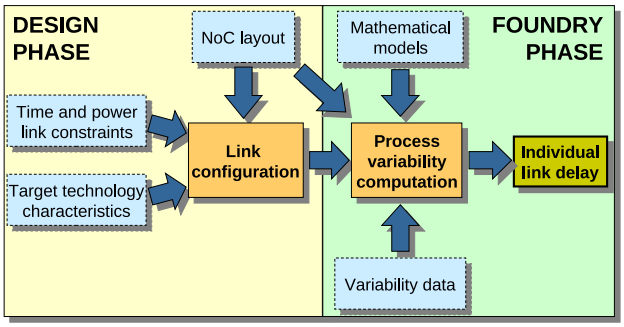


Fig. 2. Flowchart of the proposed methodology.

and designers to better design and tune on-chip networks.

The paper is organized as follows. Section II presents the proposed methodology for characterizing process variation in NoC links. Section III presents an introduction to link design and the example NoC that will later be used as a case study for our methodology. In Sections IV, V, and VI, the pieces that compose our methodology are individually presented. Section IV describes how back-end variability has been modeled while Sections V and VI present the models for random and systematic components of front-end variability, respectively. Once the different components of the methodology are introduced, Section VII describes how all of them work together in order to provide delay data for each link in the network. Finally, Section VIII draws some conclusions.

II. A Methodology for Characterizing Link Delay Variability in NoCs

In this paper we propose a methodology intended for the analysis of process variability in NoC links. This methodology, implemented as a tool set, is displayed in Figure 2. As can be seen, this methodology has several inputs, some of them related with the NoC design phase and some of them related with the chip manufacturing stage. The first set of inputs is composed of the NoC layout synthesized at design time, the characteristics of the target technology used to implement the chip, and time and power constraints for the links in the network. The second set of inputs includes parameter variation data and several mathematical models. The output of the methodology is the delay for each of the links in the network. More precisely, this methodology provides accurate delay data for each of the wires in every network link.

The methodology is divided into two steps. In the first one, the best link configuration (number of repeaters and their size) is computed according to the design time set of inputs. In the second step, link delay variability data is obtained for all the links in the network according to the link characteristics from the previous step by performing Monte-Carlo Spice simulations.

The proposed methodology takes into account several sources of variability, namely back-end variability, caused by the CMP process, and front-end variability caused by L_{gate}, and threshold voltage variations. Actually, these variability

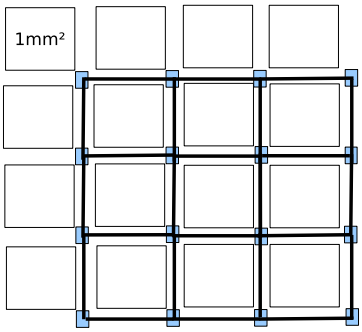


Fig. 3. Scheme of the network layout and link distribution for a 4x4 multicore chip using a 2D mesh NoC. Small grey squares denote switches.

sources can also be classified into systematic and random components since their consequences manifest in a different way. More concretely, variations in L_{gate} and variations due to the CMP process belong to the systematic category. On the opposite, a fraction of the total threshold voltage variation has random nature (another fraction is related with variations in L_{gate}). Systematic variation is characterized by a spatial correlation, meaning that differences in neighboring areas are expected to be low. On the contrary, random variation may cause different operation characteristics to adjacent areas.

It is commonly accepted that random and systematic variations are uncorrelated ($\sigma^2 = \sigma^2_{rand} + \sigma^2_{sys}$) [9]. According to this, the delay in NoC links can be represented by Equation 1. In this equation T_{nom} represents the nominal component of the delay whereas ΔT_{rnd} and ΔT_{sys} represent the timing deviations caused by the random and systematic components of variation, respectively.

$$T_{link} = T_{nom} + \Delta T_{rnd} + \Delta T_{sys} \quad (1)$$

The following sections will present each of the steps of the methodology as well as how each source of variability has been modeled.

III. NoC Link Design Background

When designing a link several concerns must be taken under consideration for the sake of efficiency. More specifically, power and area must be optimized for a target link delay.

Repeater insertion is an efficient method to reduce interconnect delay and signal transition times. Actually, this mechanism allows to minimize link delay by the optimal insertion and sizing of repeaters. However, minimizing delay involves high-sized repeaters and consequently higher power consumption. Consequently, links are designed to reach a given frequency[1] with the minimum possible power dissipation. This is achieved by inserting the proper number of minimum sized repeaters [8].

Repeater insertion requires knowing both link length and target link delay. Regarding link length, in this paper we have used the 65nm real implementation NoC layout [3] as a testbench where to apply our methodology. According to this implementation, all cores are identical, and their size is 1mm^2. Additionally, the gap between cores is 0.2mm. Therefore, links connecting NoC switches are 1.2mm long. Figure 3 shows the location of cores and links in an example die. As we will consider larger integration scales than 65nm, all layout

[1]Note that in this paper NoC operating frequency is set by link delay.

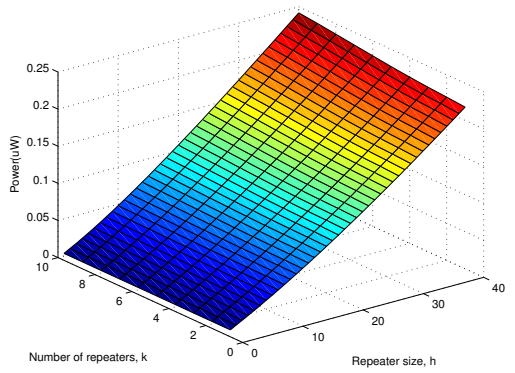


Fig. 4. Total power dissipation in a repeated interconnect as a function of h and k. Operating frequency is 2 GHz.

dimensions will be scaled by the factor "1/s", where s is the ratio between the smaller technology node and the node size in [3]. Table I shows the size of cores and links for each of the technologies considered in this paper.

Regarding link delay, it is possible to trade it for power. This is shown in Figure 4, which represents how the power consumption of a link wire varies with the number of sections (k) and repeater size (h). Data plotted in this figure has been obtained by simulating a 0.83mm long wire for a 45nm technology. Simulations have been carried out using SPICE and the PTM model for 45nm [14]. The simulated wire is assumed to be placed in a semiglobal interconnect metallization layer, whose characteristics are also shown in Table I. Data shown in this table for 45nm technology has been obtained from [13]. Data for the rest of technologies has been scaled from [13]. Wire capacitance that correspond to these features has been calculated according to the expression given in [12]. Wire segments have been modeled by using a 5-pi wire model. As shown in Figure 4, on one hand power consumption rapidly increases with repeater size. On the other hand, when the number of sections is increased, power consumption shows a slight increment when the size of repeaters is above 20. Note that high-sized repeaters are considerably faster.

In this work we will apply our methodology to NoC links synthesized using 45nm, 32nm, 22nm, and 16nm technologies. Therefore, it is required to know the link configuration (optimum number of repeaters and their size) for each of these technologies. To do so, we have simulated these links by using SPICE in the same way as described above. Figure 5 shows the link design space for the four technology nodes considered. In the plots, bars represent the delay value for a given link configuration, determined by the number of sections (k) and the size of its repeaters (h). As expected, high-sized repeaters provide the minimum delay. However, setting a maximum delay constraint value equal to 0.5ns[2], the optimal configuration is given by the minimum repeater size satisfying that constraint. When several possibilities with minimum repeater size exist, we chose the option with fewer repeaters. Table II shows the link configurations we will use in the rest of the paper for the technologies considered. Note that the number of repeaters does not increase with technology scaling because link length is also scaled.

[2]Link delay remains constant to fairly compare among technologies the effects of process variation with technology scaling.

Technode(nm)	45	32	22	16
Link length(mm)	0.83	0.59	0.41	0.3
Core area (mm^2)	0.48	0.24	0.11	0.06
Vdd(V)	1	0.9	0.8	0.7
Width(um)	140	99.6	68.4	49.78
Spacing(nm)	140	99.6	68.4	49.78
Thickness(nm)	280	199.1	136.9	99.6
Heigth(nm)	290	206.2	141.78	103.11
Dielectric	2.5	2.3	2.1	2.0

TABLE I
DATA FOR EACH OF THE TECHNOLOGIES CONSIDERED

Technode(nm)	45	32	22	16
h	5	5	5	5
k	5	4	4	5
Delay(ns)	0.46	0.44	0.44	0.43

TABLE II
LINK CONFIGURATIONS FOR THE SCALED LINKS

IV. MODELLING BACK-END VARIABILITY

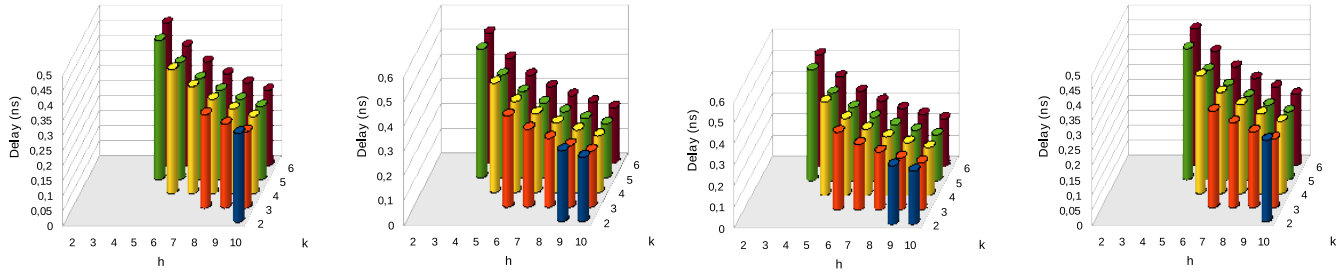
As previously stated, the Chemical Metal Planarization (CMP) process is one of the sources of timing variability. The CMP process causes surface imperfections in the wires as a consequence of dishing and erosion. Wide wires, as the ones located in the semi-global layers, are strongly affected by dishing [17] causing considerable changes in the interconnect resistance. In concordance with [1] and [4], we consider a 3σ[3] resistance variation of 15% for all technologies considered in this study. Moreover, as the degree of dishing and erosion strongly depends on the pattern density of the metallization [19] and NoC links are built in a regular layout, it is possible to assume that all wires in the link and all links in the NoC will be affected by the CMP process in a similar way.

The resistance variation produced by the CMP process has been introduced in the 5-pi link model previously mentioned. Simulation results show that resistance variation effects on delay are negligible. Delay variation in all technologies remain below 0.1% of the mean delay. These results seem to be contradictory with the results in [1] where the authors measure a delay variation, as a consequence of dishing, of around 9% of the mean delay. Delay variations (ΔT) are computed as σ/μ, where μ represents the mean and σ the delay standard deviation. However, it is necessary to clarify that links in [1] are designed to have minimum delay. More concretely, they consider, for a 65nm technology, a 5mm link with 3 repeaters of size 100. However, interconnects designed to have minimum delay are not the best option for NoC links due to their high power consumption. On the contrary, when links are designed for minimum power operating at a given frequency, delay is dominated by repeater delay and consequently variations in wire resistance have a negligible contribution to the resulting link delay.

V. MODELLING FRONT-END RANDOM VARIATION

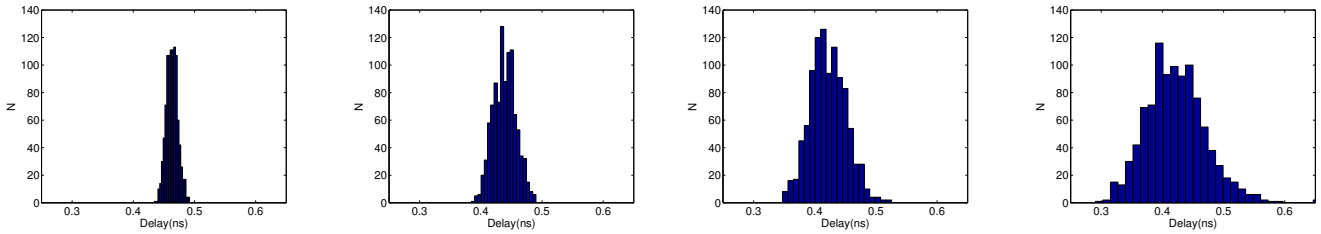
The main source of random variation in NoC links is threshold voltage variation due to Gaussian random dopant fluctuations (RDF). RDF will increasingly affect deep submicron technologies scaling from 45nm down to 16nm. Table III shows the values of $\sigma_{V_{th}}$ and $\sigma_{L_{gate}}$ for the technologies considered as provided by the ITRS report [4]. The $\sigma_{V_{th}}$ values shown in Table III represent the total threshold

[3]3σ is the usual way to express parameter variation where σ stands for the standard deviation.



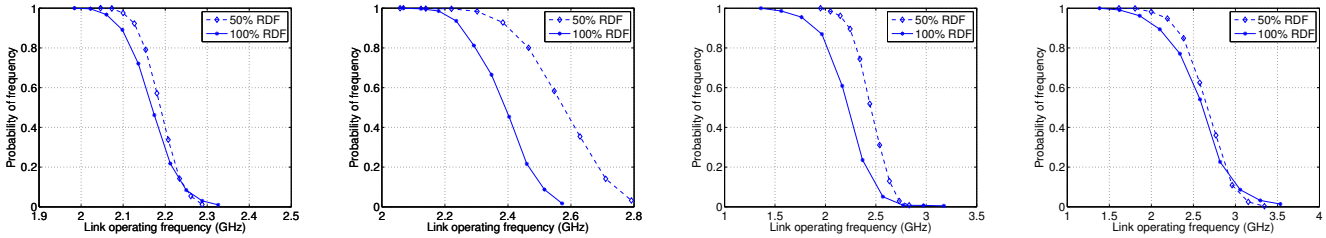
(a) 45nm (b) 32nm (c) 22nm (d) 16nm

Fig. 5. Link design space for the technologies considered. K is the number of repeaters and h is repeater size.



(a) 45nm (b) 32nm (c) 22nm (d) 16nm

Fig. 6. Link delay variation due to front-end random variability.



(a) 45nm (b) 32nm (c) 22nm (d) 16nm

Fig. 7. Maximum operation frequency of a link.

Technode(nm)	45	32	22	16
$3\sigma V_{th}$	40%	58%	81%	112%
$3\sigma L_{gate}$	12%	12%	12%	12%

TABLE III
VTH AND LGATE VARIATION ACCORDING TO ITRS

Technode(nm)	45	32	22	16
ΔT_{rnd} 50% RDF (%)	2	4.23	6.61	11.26
ΔT_{rnd} 100% RDF (%)	2.76	5.32	10.64	16.44

TABLE IV
RANGE OF DELAY VARIATION AS A CONSEQUENCE OF RDF

voltage variation. In [20] it is shown that at least 50% of that variation is due to RDF for a 45nm technology node. Note that these values represent the $\sigma_{V_{th}}$ of minimum size devices. Interconnect repeaters usually have higher gain in order to be able to drive high load values as a consequence of link length. To compute $\sigma_{V_{th}}$ for repeaters, according to [18] we can use Equation 2, and then relate the σ_{Vth_0} value of the minimum size device with the $\sigma_{V_{th}}$ of a repeater of size h, as shown in Equation 3, which clearly shows that $\sigma_{V_{th}}$ can be minimized by increasing the width of repeaters, represented by h. However, as previously stated, because of power consumption reasons it is necessary to keep the gain of repeaters as low as possible.

$$\sigma_{V_{th}} \propto \frac{1}{\sqrt{W_{eff}L_{eff}}} \quad (2) \qquad \sigma_{Vth} = \frac{\sigma_{v_{th_0}}}{\sqrt{h}} \quad (3)$$

Figure 6 shows how delay variation associated with the RDF drastically increases with technology scaling. In this figure we consider that 50% of the overall threshold voltage variation is caused by RDF according to [20]. Links in this figure are designed as explained in Section III. As shown in Figure 6, delay uncertainty increases up to a factor of 6 when fabrication processes move from 45nm to 16nm. Note that it is expected that the RDF fraction of the total threshold variation increases for future fabrication process. For this reason, Table IV shows how delay variation increases when the RDF fraction varies between 50% and 100%.

Finally, Figure 7 shows the fact that, as a consequence of random variations, not all wires in a given link will be able to operate at the same frequency. For example, for a 45nm technology all link wires will work at 2GHz whereas only 50% of them will operate at 2.15GHz. In this case there is no significant differences among wires for that link. On the contrary, when using a 16nm technology, all wires in the link will work at only 1.5GHz, while 50% of the wires will be able to operate at a frequency higher than 2.5GHz.

VI. Modelling front-end systematic variability

The systematic component of front-end variation is strongly related with the photolitographic process. Lens aberrations may lead to an important systematic spatial non-uniformity of L_{gate} over the reticle field [17]. According to [4], 3σ L_{gate} variation can be as high as 12% for fabrication processes from 45nm to 16nm. However, it is not enough knowing the maximum percentage of variation in L_{gate}. It is required to know how variations in L_{gate} influence variations in link delay and also how variations in L_{gate} are spatially distributed in the exposure field. Thus, our front-end systematic variability characterization is based on i) modeling the spatial distribution of the variation of L_{gate}, ii) assigning a value of L_{gate} to each repeater according to both the link configuration and the spatial distribution computed in i), and iii) analyzing how the values of L_{gate} assigned to repeaters affect link delay.

A. Modelling the spatial distribution of L_{gate}

In order to model the spatial non-uniformity of L_{gate} as well as its correlation we have used Gaussian Random Fields (GRF) [2]. When using GRF, with stationary and isotropic [2][9] fields, the variance (σ_i^2) of the random field $L(x, y)$, representing transistor gate length (L_{gate}) in the (x,y) die position, depends only on the euclidean distance between two given locations. Then, the gate length distribution (L) only depends on a correlation function. The correlation model we have used is the spherical model proposed by [9], which is derived from the measurements of [15]. Equation 4 shows the correlation function for this model, where $r = \|l - l'\|$ is the distance between two given locations, l and l', and X_L is a characteristic correlation length depending on the photolitographic process. Basically, $\rho(r)$ states how much the gate length of a transistor located at l is similar to the gate length of a transistor located at l'. If the distance between locations l and l' is larger than X_L then the gate length values of both transistors are totally independent. On the contrary, if both transistors are located closer than X_L then their gate length values are correlated.

$$\rho(r) = \begin{cases} 1 - (3r/2X_L) + (r/X_L)^3/2 & \text{if } (r \leq X_L) \\ 0 & \text{if } (r > X_L) \end{cases} \quad (4)$$

Understanding X_L requires further explanation. Chips are printed using a reticle field of a given size. Spatial patterns of L_{gate} variation are consequently spread over the entire exposure field. This behavior is shown in Figure 8. In this figure, the exposure field is enclosed by the thick square, while chip border is denoted by the thin square. Additionally, in this figure L_{gate} values are correlated for an area that spans half of the reticle field (grey circles). Thus, when a small chip is printed per field (left), spatial features of variability will cover half of the chip. In this case the values of X_L in equation 4 would be 0.5. However, it is possible to print four small chips in a single shot by increasing the exposure field (middle). In this case, X_L must be set to 1 as spatial features of variability are spanned across the entire chip. On the other hand, if we print a chip four times bigger than the initial one (right) by using an enlarged exposure field, then X_L would remain half of the chip size ($X_L = 0.5$).

Once the characteristics of the correlation function are defined, L_{gate} values are computed. Note that as there are infinite points on the chip surface, discretizing the surface is required. The resolution used should be the coarsest one (minimum computational complexity), that allows distinguishing repeaters in the network. Thus, to decide on the resolution to use, the 65nm real NoC layout previously mentioned [3] has been considered. According to the sizes in [3], the 4x4 chip in Figure 3 would require a 4.7mmx4.7mm die. Discretizing this die by using 0.1mmx0.1mm square points will allow us to later differentiate repeaters because with this resolution 1.2mm links will have twelve points, which is enough for the number of repeaters in Table II[4]. Thus, chip surface will be divided into a grid of 47x47 points. When larger integration scales than 65nm are considered, all layout dimensions will be properly scaled. However, the resolution will be kept constant. Obviously, for larger chips the grid would also be larger. Once chip surface has been discretized, L_{gate} values are computed for each of the points in the grid. Left side of Figure 9 shows an example of such computation. Darker areas correspond to smaller L_{gate} values and thus faster devices.

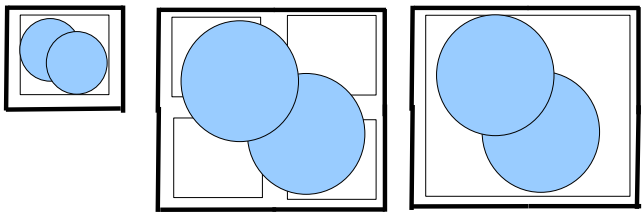


Fig. 8. Exposure field.

B. Asigning an L_{gate} value to each repeater

Once we know how L_{gate} varies across the chip, the next step in our methodology is mapping computed values of L_{gate} to each of the repeaters of every link in the NoC. To do so, we need to know the location of those repeaters. This means both knowing the exact location of each link and also knowing how many repeaters each link has. As the location and length of each link is known according to the initial layout in [3], and the number of repeaters per link was calculated in Section III using SPICE, the location of every repeater is also known. Then, the L_{gate} values computed in the previous stage are mapped to the repeaters in the network. Note that only L_{gate} values for grid points that match the location of a repeater will be used.

C. Computing delays for each link

Once the L_{gate} value for every repeater is known, delay data for each link can be collected. In order to compute the exact delay value for each link, computed L_{gate} values for each of the repeaters of a given link are introduced in the 5-pi wire model and simulated using SPICE with the PTM models previously mentioned. Note that as we are using SPICE, variations in L_{gate} automatically produce variations in the systematic component of threshold voltage. Moreover, it is worth to point out that the delay computation is independently performed for each of the links in the network.

Table V shows, for the four technologies considered in this paper, the influence on link delay of front-end systematic

[4]Although this resolution is enough, the effect of using higher resolutions needs to be further researched.

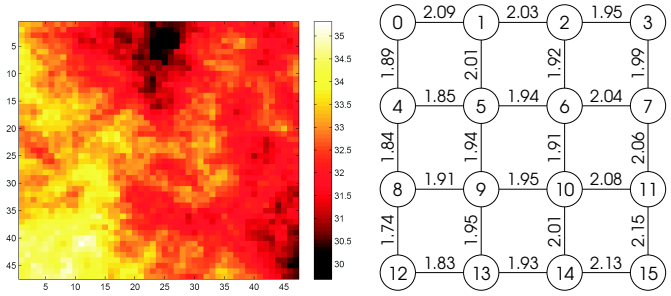


Fig. 9. Map of L_{gate} variation for a 4x4 chip XL = 1 (left) and a 4x4 NoC showing link frequency in GHz (right).

Technode(nm)	45	32	22	16
ΔT_{sys} (%)	4.31	4.34	6.27	9.31

TABLE V
LINK DELAY DIFFERENCES DUE TO SYSTEMATIC VARIATION

variation. As can be seen, this component of process variation may contribute to link delay variability as much as 9% for the links considered. Data in Table V are the result of averaging the link delay standard deviation for 100 instances of 8x8 2D mesh NoCs.

VII. COMBINING ALL SOURCES OF VARIATION

As the final stage of our methodology we have to combine all variability contributions to obtain the link frequency characterization. To do so, all the variability components previously analyzed are combined in the link model presented in Section III and simultaneously simulated in order to collect variability data for each of the links in the network. Right side of Figure 9 shows an example of the output provided by our methodology for the 4x4 2D mesh that originated the L_{gate} variation map on the left side of the same figure. Numbers close to each link in the figure are the maximum frequency achievable by that particular link. Note that frequencies in Figure 9 are the result of applying all sources of variability previously detailed to links that were initially intended to operate at 2GHz.

On the other hand, the second result provided by our methodology is a distribution of the operating frequency of link wires as shown in Figure 7. Note that this frequency distribution is individually provided for every link in the network. Additionally, as it was explained in Section II, link frequency will be influenced by both components of variability, systematic and random. Therefore, the distribution of the operating frequency of link wires will be centered at a different frequency for each link (systematic component) and will have a different shape (random component). Actually, random variability data will be useful for tuning mechanisms based on using or not single link wires [21].

VIII. CONCLUSION

In this paper a methodology to analyze the impact of process variation in link delay for NoCs has been presented. Up to our knowledge, no previous work has performed such a comprehensive analysis of process variation in NoC links. Although the proposed methodology should be applied to additional NoC layouts with very different characteristics (link length and/or spatial link distribution), the results of this work confirm that link variation cannot be neglected anymore. In this sense variability models for NoC architectures should also consider link variation in addition to variation in cores and switches. Actually, our methodology could be easily extended in order to consider variations in cores and switches just by taking into account their netlist.

Other important results of this work are the ones related with the impact of the different variability sources on delay variation. First, it is specially important the fact that resistance variation has a negligible impact in NoC links when links are designed with power and timing constraints. Second, we have measured an increasing importance of random variation in delay uncertainty with technology scaling. This is the confirmation that random variation will play a very important role in future technologies. Consequently, new approaches able to face the presence of random variation will be required to better exploit link wire characteristics.

ACKNOWLEDGMENT

This work was supported by the Spanish MICINN under Grants "CONSOLIDER-INGENIO 2010 CSD2006-00046" and TIN2009-14475-C04-01, and the European Comission in the context of the NaNoC project, Grant Agreement 248972.

REFERENCES

[1] M. Mondal et al. "Provisioning On-Chip Networks under Buffered RC Interconnect Delay Variations", ISQED '07.
[2] B. Hargreaves, H. Hult and S. Reda, "Within-die process variations: How accurately can they be statistically modeled?", ASPDAC 2008.
[3] A. Pullini et al, "NoC Design and Implementation in 65nm Technology", *NOCS '07*.
[4] "International Technology Roadmap for Semiconductors", 2007 Edition, available online at http://www.itrs.net/Links/2007ITRS/Home2007.htm
[5] B. Stefano et al. "Process variation tolerant pipeline design through a placement-aware multiple voltage island design style", *DATE '08*.
[6] Y. Hoskote et al. "A 5-GHz Mesh Interconnect for a Teraflops Processor," in *IEEE Micro Magazine*, Sept-Oct. 2007, pp. 51-61.
[7] A.J. Jose "Near speed-of-light on-chip interconnects using pulsed current-mode signalling", Symposium on VLSI circuits 2005.
[8] G. Chen and E. Friedman, "Low-power repeaters driving RC and RLC interconnects with delay and bandwidth constraints", *IEEE Trans. Very Large Scale Integr. Syst.*, 2006
[9] Smruti R. Sarangi et. al, "VARIUS: A Model of Parameter Variation and Resulting Timing Errors for Microarchitects", *IEEE Transactions on Semiconductor Manufacturing*, 2008.
[10] R. Kumar et al, "Interconnections in multi-core architectures: Understanding mechanisms, overheads, and scaling", ISCA, June 2005.
[11] J.D. Owens et al., "Research challenges for on-chip interconnection networks", IEEE Micro, 2007.
[12] Shyh-Chyi Wong et al. "An empirical three-dimensional crossover capacitance model for multilevel interconnect VLSI circuits", *IEEE Transactions on Semiconductor Manufacturing*, 2006.
[13] "The Nangate Open Cell Library", 45nm FreePDK, available online at https://www.si2.org/openeda.si2.org/projects/nangatelib/.
[14] "Predictive Technology Model", available online at http://www.eas.asu.edu/ ptm.
[15] P. Friedberg et al., "Modeling Within-Die Spatial Correlation Effects for Process-Design Co-Optimization", in *ISQED*, 2005.
[16] D. Ludovici et al., "Assessing Fat-Tree Topologies for Regular Network-on-Chip Design under Nanoscale Technology Constraints", in *DATE* 2009.
[17] Michael Orshansky et al., "Design for Manufacturability and Statistical Design: A Comprehensive Approach", Springer 2006.
[18] A. Asenov et al., "Intrinsic threshold voltage fluctuations in decanano MOSFETs due to local oxide thickness variations", IEEE Transactions on Electron Devices, January 2002
[19] S. Lakshminarayanan et al., "Electrical characterization of the copper CMP process and derivation of metal layout rules", IEEE Transactions on Semiconductor Manufacturing, November 2003.
[20] Kenyon, C. et. al, "Managing Process Variation in Intel's 45nm CMOS Technology." Intel Technology Journal, January 2008.
[21] C. Grecu et al., "NoC Interconnect Yield Improvement Using Crosspoint Redundancy", in DFT 2006.