<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Cyber-physical management for heterogeneously integrated 3D thousand-core on-chip microprocessor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>P. D., Sai Manoj; Yu, Hao</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2013</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/18210">http://hdl.handle.net/10220/18210</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2013 IEEE. This is the author created version of a work that has been peer reviewed and accepted for publication by IEEE International Symposium on Circuits and Systems (ISCAS) 2013, IEEE. It incorporates referee’s comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: <a href="http://dx.doi.org/10.1109/ISCAS.2013.6571898">http://dx.doi.org/10.1109/ISCAS.2013.6571898</a>.</td>
</tr>
</tbody>
</table>
Cyber-Physical Management for Heterogeneously Integrated 3D Thousand-core On-chip Microprocessor

Sai Manoj P.D and Hao Yu
School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798
haoyu@ntu.edu.sg

Abstract—Though 3D TSV/TSI technology provides the promising platform for heterogeneous system integration with design drivers ranged from thousand-core microprocessor to millimeter-cubic sensor, the fundamental challenge is lack of light to deal with significantly increased design complexity. From device level, new state of variables from different physical domains such as MEMS, microfluidic and NVM devices have to be identified and described together with conventional states from CMOS VLSI; and from system level, cyber management of states of voltage-level and temperature has to be maintained under a real-time demand response fashion. Moreover, a cyber-physical link is required to compress and virtualize device level state details during system level state control. This paper shows device-level 3D integration by example of MEMS and CMOS VLSI. In addition, a cyber-physical thermal management for 3D integrated many-core microprocessors is discussed.

I. INTRODUCTION

With the increasing demand of cloud computing for big-data, design of high-throughput data servers has obtained recent interest significantly. The big-data processing at exascale is obviously beyond traditional single-core or multi-core microprocessors. Many-core microprocessors with thousand-core become the emerging need with many recent explorations [1], [2], [3]. The primary challenges come from the low bandwidth and high power density in 2D integration. Moreover, such a complicated computing system requires new means of states identification, reduction and management.

3D integration applies vertical stacking of layers one above other by through-silicon-via (TSV) or through-silicon-interposer (TSI). As such, the communication bandwidth can be improved with small interconnection latency. Moreover, as the loss of I/O is reduced with more data transferred, the communication power can be reduced as well. The other advantage of 3D comes from heterogeneous integration, i.e., devices made from different technologies such as nano-scale non-volatile memory (NVM), MEMS, and even microfluid [4], [5], [6], [7]. Thereby, one can build a smart cubic microsystem with multiple functionalities. For example, Fig.1 shows one possible 3D integrated thousand-core on-chip microprocessor with TSV based I/O to connect structured memory and core blocks such as data-bus or clock. The NVM device is considered here to replace the main memory by DRAM. In addition, microfluid works as active cooling channel to dissipate heat [7].

The primary limitations for 3D integrations, especially with applications in thousand-core on-chip, are as follows. Firstly, one needs to identify new physical-domain states. The new nano-scale NVM device such as spin-based STT-RAM may show dynamics not determined by traditional electrical voltages or currents, but by magnetization angles or doping density [6]. Moreover, a reliable utilization of TSV/TSI needs a multiple physical-domain model to characterize cross-coupled electrical-thermal-mechanical delay. Secondly, one needs to reduce the number of states as too many timing violation, power integrity and thermal reliability to check layer by layer. The essential state extraction by macromodeling is required to virtualize the system complexity [4]. Lastly, one needs to perform smart state management for power and thermal. For example, the problem is different now when providing the power supply from many power converters with many voltage levels to thousand cores. Moreover, the long heat dissipation path may require integrated active cooling scheme [7] or new power gating scheme [5].

In this paper, we discuss potential challenges and solutions to build 3D thousand-core system for big-data cloud computing. We show a heterogeneously integrated 3D thousand-core on-chip microprocessor design from perspective of cyber-physical management. Section 2 explains the overall architecture of 3D thousand-core microprocessor and the need for cyber-physical management. Physical modeling in terms of state identification for NVM devices and TSV/TSI is explained in Section 3. Section 4 illustrates how macromodels is formulated for complexity reduction. Section 5 explains cyber system management for 3D thousand-core on-chip microprocessor with adaptive flow-rate cooling and power gating. Conclusions are drawn in Section 6.

II. HETEROGENEOUS 3D THOUSAND-CORE SYSTEM

One heterogeneously integrated 3D thousand-core system is shown in Fig. 1 for big-data cloud computing. Firstly, many-core microprocessors are organized in a network-on-chip mesh, where core and core communicate by routers. The main memory is designed using NVM devices. Each core visit its local block memory by TSV or TSI with I/O links. Digital power and temperature sensors are realized on-chip to monitor real-time power and thermal profiles, which provide feedback to system to control the power and temperature. For example, one can adjust the flow-rate of microfluid according to the temperature gradient profile.

In this paper, we show a design methodology from cyber-physical perspective to realize such a heterogeneously integrated 3D thousand-core system. Firstly, building a physical model to consider new physical-domain states introduced from non-traditional devices such as nano-scale NVM devices is shown followed by building a physical-model by considering multiple physical-domain states for TSV or TSI delay under coupling from thermal temperature and mechanical stress. Next, with the use of structured and parameterized macromodeling, we show how to extract the essential states to reduce complicated physical model of 3D thousand-core system. Lastly, by the use of macromodels in a close-feedback-loop with prediction,
managing the system states such as power and thermal in a cyber-physical fashion is shown.

III. PHYSICAL DEVICE MODELING

The heterogeneous integration of different technologies from different physical domains results in the challenges in physical models, to identify new physical-domain states or multiple-physical states. In this section, we first discuss physical modeling for the nano-scale NVM devices such as STT-RAM [6] by identifying the new physical-domain state, and also show physical modeling of TSV with cross-coupled delay model from electrical, thermal and mechanical domains.

A. New Physical-domain State

Traditional electric devices are mainly described by modified nodal analysis (MNA) with nodal voltages and branch currents ($V_n$, $j_i$). For nano-scale NVM devices, there are new states to be determined. For example, we need to know magnetization angle to fully describe the dynamics of STT-RAM. Moreover, doping ratio is needed for memristor and crystallization rate for PCM [6]. At the same time, one needs to develop a new MNA state description for both CMOS and NVM devices.

As shown by Fig. 2, we add new branch currents associated with new NVM devices, which are described by introducing new state variables $s_{n}$, determine the conductance of all NVM devices. Note that incident matrix for capacitor, resistor, inductor and current source are denoted by $E_c$, $E_g$, $E_l$, $E_i$, and additional state variables, $s_n$ for NVM device are linked by incident matrix $E_m$.

Considering a STT-RAM device, which has two sandwiched ferromagnetic layers and oxide layer in between [6], it needs a new state variable $\theta$, angle of magnetization between two magnetic layers to describe giant-magneto-resistance (GMR). As such, GMR becomes

$$R(\theta) = R_L + \frac{R_H - R_L}{2} (1 - \cos(\theta))$$

(1)

The new state vector becomes $X = [v_n, j_i, j_m, \theta_m]^T$ instead of $X = [v_n, j_i, j_m]^T$ [6]. One new MNA [6] can be derived correspondingly to fully describe the dynamics of such a hybrid NVM and CMOS system.

B. Multiple Physical-domain State

TSV or TSI is the essential component to integrate NVM memory and microprocessor core in the proposed 3D thousand-core system. As global data-bus or clock with relevant I/Os between memory and core blocks, TSV or TSI can be fabricated in a structured fashion reliably. At the same time, unlike 2D, TSV becomes the path for heat dissipation and also stress passing. All multi-physical domain effects can have significant impact on the electrical delay of TSVs. The electrical model of TSV is thereby not accurate if no thermal and mechanical behavior are considered.

1) TSV Delay with Temperature: TSVs are generally surrounded by a liner material (SiO₂ or Si₃N₄), of very small radius to avoid diffusion of metal atoms into silicon substrate and to provide isolation. The TSV structure with isolation, for example, in a 3D clock-tree [8], is shown in Fig. 3.

$$\frac{1}{C_t} = \frac{1}{C_{ox}} + \frac{1}{C_{dep}}; R_t = \frac{\rho h}{\Pi r_{metal}}.$$

(2)

Here, $C_{ox}$ and $C_{dep}$ are liner capacitance and depletion capacitance of TSV respectively, with $C_{ox} = \frac{2\varepsilon_{si}}{\ln(\frac{r_{metal}}{r_{ox}})}$ and $C_{dep} = \frac{2\varepsilon_{ox}}{\ln(\frac{r_{metal}}{r_{dep}})}$. Note that $\rho$ is the resistivity of the TSV metal and $h$ is height of TSV; $\varepsilon_{si}$ and $\varepsilon_{ox}$ are dielectric constants of silicon and silicon oxide; and $r_{metal}$, $r_{ox}$ and $r_{dep}$ are the outer radius of TSV metal, silicon and depletion regions. As $r_{dep}$ depends on temperature surrounding TSV, it results in non-linear TSV capacitance with temperature and hence has non-negligible impact on delay.

A typical C-V curve for TSV with liner is shown in Fig. 4, which can be divided into three regions, based on variation of capacitance, separated by flat band ($V_{FB}$) and threshold voltage ($V_T$). Based on this electrical-thermal coupled model, one can derive the Elmore delay model when using TSV as link between memory and core [8].

Fig. 4: Typical C-V curve of TSV MOSCAP with non-linear temperature dependence

2) TSV Delay with Stress: TSVs can exert stress on the silicon substrate due to differences in coefficient of thermal expansion (CTE) between TSV and silicon material. The exerted mechanical stress from TSVs will alter the mobility of the devices present on the
essential states, which can be achieved by macromodeling \[4\]. Of physical states, one need to reduce unnecessary states and extract to manage the states, in terms of timing, power and temperature, for and also multiple physical-domain states. As such, it becomes difficult tremendously increased complexity from new physical-domain states

\[\sigma\]
substrate by changing the lattice structure. The impact of stress (\(\sigma\)) with respect to parameter in a structured fashion by expanding the original state vector in frequency (s) domain

\[x(p, s) = \sum_{i=1}^{\infty} \sum_{i_1}^{\infty} (x_1^{i_1} \cdots p_s(s)(dp_1)^{i_1} \cdots (dp_s)^{i_s})\]

\(x_{ap}(0) = [x_0^{(0)}, x_1^{(1)}, \ldots, x_p^{(1)}, \ldots, x_{ap}^{(2)}, \ldots]_{K,P} \ldots]\)

Reorganize (4) by considering sensitivities

\[sx_{ap}(s) = A_{ap}x(s) + b_{ap}u(s)\]

\[y_{ap}(s) = C_{ap}x_{ap}\]

As such, one can result in a compact state representation with both sensitivity \((x_1^{(1)}, \ldots, x_p^{(1)})\) and nominal response \((x_0^{(0)})\) from the original state equation.

Such a structured and parameterized macromodeling is deployed for a 2-layer 3D design is performed in [4]. With compact macromodeling, one can perform simultaneous TSV density optimization to reduce thermal and power hot-spots as shown in Fig. 6. The macromodeling based design shows 127X faster compared to the approach without reduction of states.

V. CYBER SYSTEM MANAGEMENT

The management of system states such as power and temperature requires the the use of macromodel. Based on the data from macromodel and sensor, one can perform prediction and correction to generate the real-time response to track the power and temperature profiles. As shown in Fig. 7, one can predict temperature/power demand by macromodel. When corrected by sensor measured data,
The physical device model has been explored to consider new state threshold, resulting in large data retention overhead. Though power-control with static and run-time power gating is depicted in Fig. 9. In terms of power and temperature, with prediction and correction with calibration from sensor measured data. A number of design examples have been deployed to support the aforementioned design methodology in 3D thousand-core system, including STT-RAM model, thermal-stress delay model of TSV in clock-tree at physical level; but also microfluid cooling and NEMS based power gating at system level.

**ACKNOWLEDGEMENT:** This project is supported by Singapore MOE TIER-2 (MOE2010-T2-2-037) fund. The authors thank for the fruitful collaboration with Prof. Chuan Seng Tan, Prof. Wei Zhang and Prof. Chip-hong Chang at Nanyang Technological University.

**REFERENCES**


