Extended Large (3-D) Integration TEchnology

Seventh Framework Programme
FP7-ICT-2007-1
Project Number: FP7-ICT-215030

D2.5: 3-D IC’s Modelling and Simulation: Conclusion Report

Version 3.0
Extended Large (3-D) Integration TEnology

<table>
<thead>
<tr>
<th><strong>Project Name</strong></th>
<th>Extended Large (3-D) Integration Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project Number</strong></td>
<td>ELITE-215030</td>
</tr>
<tr>
<td><strong>Document Title</strong></td>
<td>3D IC’s Modelling and Simulation: Conclusion Report</td>
</tr>
<tr>
<td><strong>Work Package</strong></td>
<td>WP2</td>
</tr>
<tr>
<td><strong>Dissemination Level</strong></td>
<td>CONFIDENTIAL</td>
</tr>
<tr>
<td><strong>Lead Beneficiary</strong></td>
<td>Lancaster University (ULANC)</td>
</tr>
<tr>
<td><strong>Document Editors</strong></td>
<td>Roshan Weerasekera, Matt Grange, Dinesh Pamunuwa (ULANC), Christine Fuchs (LETI), Luca Bortesi and Loris Vendrame (MICRON)</td>
</tr>
<tr>
<td><strong>Version</strong></td>
<td>3.0</td>
</tr>
</tbody>
</table>
Abstract

In this report we discuss the modelling and analysis of 3-D ICs within the scope of ELITE project and their usage in design space exploration of general systems.

We present a set of TSV parasitic extraction models for various TSV structures laid out in different substrates, taking into account the physical proximity of neighbouring TSVs. The proposed models can be used in system-level performance design space explorations. The RF and low-frequency behaviour of TSVs has also been modelled and measured from test structures. We then discuss 3-D signalling conventions by exhaustively quantifying the trade-offs between standard CMOS drivers and receivers as well as shielding techniques over the TSVs. Then, the methodology to model the thermal behaviour of a die stack is discussed and a stand-alone tool for thermal performance estimation is presented. We follow with the network-level communication analysis where we show how the average distance as a design metric can aid in the partitioning and optimization of large multi-core systems. We then discuss the scalability of 2-D and 3-D networks-on-chip and vertical clocking schemes. Finally we use all of the physical, circuit, and communication-level analyses to create system-level models for performance of computational units.

Keywords

3-D Integration, Through-Silicon Via (TSV), Parasitic parameter extraction, Signalling, Thermal Compact Models, Computational efficiency, Flash memory, Integrated Circuit (IC) design.
# Table of Contents

1. INTRODUCTION .............................................................................................................................. 7

2. ELECTRICAL MODELLING OF TSVS ............................................................................................... 8
   2.1. INTRODUCTION ....................................................................................................................... 8
   2.2. RELATED WORK ...................................................................................................................... 8
   2.3. 3-D INTEGRATION TECHNOLOGY ....................................................................................... 9
   2.4. PHYSICAL MODELLING OF TSVS ....................................................................................... 11
   2.5. TSV MODELLING IN THE RADIO FREQUENCY RANGE ....................................................... 27
   2.6. LOW-FREQUENCY CAPACITANCE MEASUREMENTS .......................................................... 44
   2.7. PHYSICAL MODELLING CONCLUSIONS ............................................................................. 52

3. SIGNALLING CONVENTIONS FOR 3-D ICS .................................................................................. 54
   3.1. INTRODUCTION ..................................................................................................................... 54
   3.2. RELATED WORK .................................................................................................................... 54
   3.3. TSV PARASITICS AND ELECTRICAL MODEL .................................................................... 55
   3.4. PERFORMANCE COMPARISON: ENERGY, AREA, SI AND BANDWIDTH ............................ 70
   3.5. DISCUSSIONS AND CONCLUSIONS ................................................................................... 72

4. COMPACT THERMAL MODELLING OF 3-D ICS ........................................................................ 73
   4.1. INTRODUCTION ..................................................................................................................... 73
   4.2. RELATED WORK .................................................................................................................... 74
   4.3. THERMAL ANALYSIS OF INTEGRATED CIRCUITS .......................................................... 74
   4.4. DISCUSSION AND CONCLUSIONS ..................................................................................... 82

5. COMMUNICATION PLATFORM ARCHITECTURES .................................................................... 83
   5.1. OPTIMAL NETWORK ARCHITECTURES .............................................................................. 84
   5.2. SCALABILITY OF 3-D NOC ARCHITECTURES ................................................................... 96
   5.3. PERFORMANCE ANALYSIS OF MULTI-CLOCK 3-D NOCS ............................................. 106
   5.4. CONCLUSIONS ....................................................................................................................... 112
6. SYSTEM-LEVEL MODELLING OF SILICON SYSTEMS ................................................................. 113
   6.1. INTRODUCTION ........................................................................................................... 113
   6.2. RELATED WORK .......................................................................................................... 113
   6.3. MODELLING PERFORMANCE ...................................................................................... 114
   6.4. EFFECTIVE COMPUTATIONAL EFFICIENCY ............................................................... 116
   6.5. SCALING CIRCUITS, DEVICES AND INTERCONNECTS ............................................. 120
   6.6. DESIGN SPACE EXPLORATION ................................................................................... 120
   6.7. COST .............................................................................................................................. 124
   6.8. APPLICATIONS OF THE MODEL ............................................................................... 125
   6.9. CONCLUSIONS ............................................................................................................ 131

D2.5 CONCLUSIONS .................................................................................................................. 133

REFERENCES ................................................................................................................................ 134
1. Introduction

WP2 in the ELITE project addresses all the steps necessary for pre-fabrication estimation of performance, signal integrity, power integrity, thermal behaviour and cost of 3-D ICs. A hierarchical approach has been adopted in this work package, with models at different levels of abstraction proceeding from physical to circuit to network to system level to enable efficient and detailed analyses at various points in the spectrum defined by computational efficiency vs. accuracy. The modelling activity is supported by simulation methodologies for investigating overall performance metrics and are used to carry out detailed design space explorations to provide understanding and insight into the design of 3-D integrated circuits in various application domains.

Figure 1 shows the hierarchy of tasks carried out in WP2 and their usage in the design of 3-D ICs.

Figure 1 ELITE WP2 Modelling and Simulation Package Task Dependencies

The models as well as their usage have previously been reported in various forms to the general public and through internal reports and deliverables to ELITE consortium. The deliverables submitted within WP2 are:

- **D2.1**: Extraction and Simulation Methodology
- **D2.2**: 3D IC Analysis
- **D2.3**: 3D IC Simulation
- **D2.4**: 3D Signalling Conventions

In order to make this report a complete stand-alone source for system-level performance estimation models and their usage, we reproduce relevant portions of the previously reported works together with updates as per the discussions and comments received from ELITE partners.

This report is organized as follows: in Chapter 2, we discuss TSV electrical models for various configurations and geometries in different substrate types. We follow with RF modelling and measurements and low-frequency capacitance measurements from the fabricated test structures. In Chapter 3 we give an appropriate equivalent electrical model for TSVs for signal integrity analysis and compare different signalling conventions for 3-D ICs. Chapter 4 is dedicated to describing the thermal simulation methodology for the ELITE prototype, and also the development of a full 3-D thermal estimation tool based on the circuit approximation of heat propagation and an underlying Spice engine. Next, in Chapter 5, the optimal organization of multi-processor systems in a mesh structure and its communication platform is discussed in detail, with special emphasis on 3-D. Additionally, a multi clocking scheme for 3-D ICs has also been investigated. Based on the parasitic and thermal models proposed, a set of hierarchical models are proposed for physical and system level performance estimation of 2-D and 3-D ICs in Chapter 6. Further guidelines for 2-D and 3-D IC system-level cost comparisons are also given.
2. Electrical Modelling of TSVs

2.1. Introduction

Key to the performance benefits of 3-D ICs is the electrical behaviour of TSVs, which do not exhibit the slow diffusive propagation seen across global planar wires. Hence the global architecture of high-performance 3-D ICs has to reflect the improvements in latency and bandwidth afforded by the availability of TSVs, latency and bandwidth being functions of the TSV density, physical dimensions and material constants in the first instance. Parasitic parameter extraction of TSVs is therefore a critical step in the successful physical design as well as design-space exploration of 3-D ICs. Accurate prefabrication estimates of interconnect parasitics can be obtained by field-solvers but running a full-chip simulation is computationally infeasible (elaborated Figure 2). Further, the electric potential on a TSV creates a depletion layer in the surrounding substrate and the related analysis is more the province of semiconductor process and device modelling tools. The combination of small and large physical dimensions (barrier thicknesses in the range of a few nanometres and TSV length in the range of micrometres), and the complexity of a coupled network of TSVs mean that even field solvers dedicated to interconnect extraction can run into severe convergence problems. Simulations to extract the electrical parasitics of a TSV bundle with such a tool can easily run for several days on a high-performance desktop.

The proposed usage of the models reported in this study is design space exploration and early chip planning prior to full place and route, when accurate pre-characterised structures are not available for capacitance and resistance estimation, and perhaps even technology details have not been finalized. With this goal in mind we propose a set of analytic models for parasitic parameter extraction of TSVs that cover a wide range of technological and design parameters. This work builds on models described initially in [4], and presents a set of closed-form equations for resistive \((R)\), capacitive \((C)\) and inductive \((L)\) parameters of TSVs including coupling terms for TSV structures laid out in varying topologies in different substrate types. We show how the functional form of our models is much more compact but still more accurate and far more general in its ability to handle different structures such as rows and bundles as well as different substrate types than previously reported models. Specifically we consider isolated TSVs, TSVs laid out in a row and TSVs laid out in a bundle with all associated coupling terms for three substrate resistivities representative of wafers typically used in analogue, digital and mixed-signal designs. The generality, accuracy and closed-form nature of the parasitic models mean they can readily be used for estimation of performance and signal integrity of TSV-based 3-D ICs before precise layout information is known.

The rest of this chapter is organised as follows. Firstly, we discuss prior related work. Then we describe the physical structure of TSVs in different 3-D integration schemes and discuss the nature of capacitive and inductive coupling in a row of parallel TSVs as well as a TSV bundle. Next, semi-empirical closed-form equations for resistive, capacitive and inductive terms are proposed and their accuracy in the simulated range which is deemed to be representative of widely available technologies is discussed. We then discuss the RF modelling and measurements of TSV test structures and follow with the low-frequency capacitance measurements. We end with our conclusions.

2.2. Related Work

Several prior works address the microwave characterization of specific TSV structures and analysis of their high frequency behaviour [5], [6]. Characterisation necessarily depends on fabrication of a test structure and test vehicles for characterization are generally not available for design space exploration. Others [7], [8] have employed full-wave numerical simulations which are computationally expensive and memory intensive for any but the simplest structures. Design space exploration requires parasitics for TSVs of varying geometries, including different inter-via spacings within a bundle.
In [5] and [6] the high frequency dependence of resistance and inductance of TSVs arranged in a row are modelled, but only for a specific geometry, and further capacitance is not discussed. The equivalent circuit proposed in [6] is rather complex and can be reduced to a simple $R$, $L$, and $C$ model as discussed in this work and in [10], [11]. Khalil et al. in [10] proposed an empirical delay model, but no explicit formulae are given for parasitic parameter estimation. Although the propagation delay of a TSV is a function of its physical dimensions, describing a TSV using an equivalent circuit provides circuit intuition that allows not only propagation delay calculation, but also power and energy calculations and signal integrity analyses. Other recent works [11], [12] discuss trends in TSV parasitics for specific technologies. Some validated models for isolated TSVs have been proposed in [13], [14] without considering coupling between adjacent TSVs. Typically TSVs will be laid out in close proximity within a bundle to maximize area efficiency and therefore inter-via coupling has to be considered. In [15] the authors describe general parasitic models for TSVs arranged in a row but direct closed-form formulae are not presented for the various capacitive and inductive components; instead a set of equations need to be solved. The authors of [16] were among the first to propose empirical closed-form models but considered only isolated and two coupled TSV structures for a single substrate resistivity.

As described in [1], well-established practice in physical design automation is to capture complex interconnect structures as a collection of simpler representative shapes, where each such “primitive” element has a well-defined parasitic signature. The overall effective parasitics of the system are obtained by extrapolating the parameterized primitives using a combination of table look-up and analytic equations [2], [3]. Previously we have proposed a set of self-consistent closed-form expressions that focused on parasitic extraction of TSVs in highly resistive substrates [4], [9]. In this work we model three different substrate types representative of typical digital, analogue and mixed-signal processes, and also consider the effect of the depletion region around the TSV. We also consider extended dimensional ranges and include the isolation layer thickness and signal frequency as parametric variables, making the models more general and complete. Finally in [13] the functional form presented in this work for inductance has been experimentally validated with measurement results for square shaped isolated TSV structures.

![Figure 2: Two Layer Face-to-Back (F2B) CMOS Die Stack](image)

### 2.3.3-D Integration Technology

The past decade has seen an explosion in research within academia and industry into different aspects of 3-D integration based on die and wafer stacking; institutes such as IBM [17], IMEC [18], LETI [19], Samsung [20], Elpida [21], NEC [21] and Tezzaron [22] have vigorously explored and developed 3-D integration technologies using diverse processing and stacking technologies. All wafer-level techniques share four major steps: formation of TSVs, thinning of wafers, precision alignment of wafers or dies, and bonding [23]. These approaches can be categorized based on assembly orientation, TSV process sequence, and the handling of the substrate, including any sacrificial handles.

Wafers and dies can be bonded using a multitude of techniques, such as with the use of organic glues, oxide bonding, or metal bonding. The bonding can either be face-to-back (F2B), shown in Figure 2, or face-to-face (F2F). The TSVs in a F2F orientation would be very short and resemble standard via structures in terms of their parasitics,
but any subsequent layer would always require long TSVs that traverse the entire substrate. The total thickness of the substrates used for integrated circuits is typically between 300 μm and 800 μm, but they are thinned before being stacked in order to control the aspect ratio of the TSVs as well as to make the height of the total stack small enough to be compatible with existing packaging standards.

Epitaxial (EPI) and Silicon-on-Insulator (SOI) are widely used substrate types in IC manufacturing and both can be stacked, but for handling ease and providing a convenient etch stop, SOI wafers are generally preferred. EPI is the standard substrate used for digital applications, with a relatively heavily doped substrate (typically resistivity ρ=0.02 Ωcm) in order to reduce the voltage differences across the substrate and thereby reduce the chances of latch-up, and an epitaxial layer with a higher resistivity (typically ρ= 10 Ωcm). Wafers with lightly doped substrates (typically ρ= 7 Ωcm) on the other hand are used for radio frequency circuits, as the higher resistivity provides isolation between components. Mixed-signal circuits can use either, depending on whether the latch-up or isolation problem is deemed to be more critical [24]. Hence the substrate conductivity can vary over a range spanning more than four decades, for example from 1 S/m to 10s S/m (resistivities from 100 Ωcm to 10000 Ωcm). Typically though, the effect of the substrate conductivity on capacitance can be approximated to three regimes captured by high, medium and low resistivities. We adopt this approach and while it is an approximation, capturing three regimes significantly increases the allowable exploration space, as previous works essentially consider only the highly resistive regime.

Figure 3: TSV structures connecting adjacent wafers have different electrical properties depending on proximity to other TSVs and layout: (a) an isolated TSV, estimation of $R_{tsv}$ from (1), $L_{tsv}$ from (2), $C_{tsv}$ from (4), (5) or (6) depending on the substrate type. (b) Coupled TSVs in a row, estimation of $R_{tsv}$ from (1), $L_t$ from (2) and $L_m$ from (3). $C_c$ can be estimated from (8), and total capacitance $C_t$ of each TSV from (7), where component terms are chosen depending on the substrate type and coefficients are selected based on whether or not the TSV is at the edge from Table IV. (c) A TSV bundle, estimation of $R_{tsv}$ from (1), $L_t$ from (2) and $L_m$ from (3). Coupling capacitances $C_{c,d}$, $C_{c,p}$, and $C_{c,t}$ can be estimated from (8) with appropriate coefficients being substituted from Table V; total capacitance $C_t$ of each TSV can be estimated from (7), where component terms are chosen depending on the substrate type and coefficients are selected based on whether the TSV is at the corner (NW, NE, SW, SE) or middle (N, W, E, S) of an outer row, or at the centre (M) from Table V.

TSV formation is essentially the same as a contact hole process step, but a major difference is that much deeper holes are etched in silicon. These holes can be fully filled or annular with either copper or tungsten, and surrounded by an adhesive layer and a dielectric barrier [25]. The insulating dielectric (SiO2 or Si3N4) acts as a barrier to the Cu TSV, preventing the migration of Cu ions into the Si substrate which can degrade device performance by inducing leakage currents. It also electrically isolates the Cu cylinder from the substrate, providing improved isolation to power and ground planes. Further, a thin annular TiN layer is usually deposited between the Cu and SiO2 layers, which acts as an adhesion layer and also concentrates the current in the Cu bar due to its high resistivity [26]. The fully filled narrow vias are used to achieve high density vertical interconnections whereas lower density implementations may involve larger U-filled or annular via structures with a polymer or air filling. In general the TSV cross-section can be either square [13] or circular [5], and they can also have a uniform [5], [11] or non-uniform tapered [27] cross-section along their length. Diameters and pitches of TSVs range from 10-100 μm for medium to low density processes with silicon thicknesses around 50-300 μm, to less than 1-10 μm with corresponding silicon thicknesses less than 50 μm in very high density processes. IBM have recently demonstrated
TSVs with a diameter of 0.14 μm and interconnection densities around 108 I/Os per cm² [28] for SOI wafers. In the case of SOI wafers, the substrate may be thinned down to a thickness of 1-2 μm to allow very narrow TSVs.

### 2.4. Physical Modelling of TSVs

Depending on the TSV density requirement and other placement constraints a TSV structure in a 3-D chip stack may be relatively isolated with no significant capacitive coupling to other TSVs, or there may be significant coupling to adjacent TSVs (refer Figure 3). Also they may be laid out in a row, or within a regular matrix depending on the application and technology constraints (see Figure 3 (c) for a 3 × 3 bundle). We model all of these likely structures, and present closed-form equations to extract resistive, capacitive and inductive components including coupling terms as a function of the TSV geometry for three common substrate resistivities. The baseline has been established by simulating all structures in a 3-D/2-D quasi-static electromagnetic-field solver specifically used for parasitic extraction of electronic components [29]. These extracted values are functions of the structure’s geometry and the material constants. The underlying variables have been combined into dimensionless quantities according to the principles of dimensional analysis [30] in order to reduce the number of independent variables where possible, and joined in functional forms suggested by insight into field theory and empirical validation.

#### Table 1 Electrical Parameter Estimation Models for TSVs

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resistance</strong></td>
<td></td>
</tr>
<tr>
<td>Low-frequency resistance</td>
<td>( R_{tsv} = \frac{\rho l_v}{\pi r_v^2} ) (1)</td>
</tr>
<tr>
<td><strong>Inductance</strong></td>
<td></td>
</tr>
<tr>
<td>Low-frequency self Inductance of a TSV</td>
<td>( L_s = \frac{\mu l_v}{2\pi} \ln \left( 1 + \frac{2.84 l_v}{r_v} \right) ) (2)</td>
</tr>
<tr>
<td>Low-frequency mutual Inductance of any two TSVs</td>
<td>( L_m = 0.199 \mu l_v \ln \left( 1 + 0.438 \frac{d_v}{l_v} \right) ) (3)</td>
</tr>
<tr>
<td><strong>Capacitance</strong></td>
<td></td>
</tr>
<tr>
<td>Capacitance of an Isolated TSV in Low Resistive Substrates</td>
<td>( C^{LRS}<em>{tsv} = \frac{k_1^L \varepsilon</em>{SiO2} \varepsilon_0 l_v}{\ln \left( 1 + k_2^L \frac{d_v}{r_v} \right)} ) (4)</td>
</tr>
<tr>
<td>Capacitance of an Isolated TSV in Medium Resistive Substrates</td>
<td>( C^{MRS}<em>{tsv} = \frac{k_1^M \varepsilon</em>{SiO2} \varepsilon_0 l_v \left( \frac{l_v}{r_v} \right)^{k_2^M}}{\ln \left( 1 + k_3^M \frac{l_v}{r_v} \right)} ) (5)</td>
</tr>
<tr>
<td>Capacitance of an Isolated TSV in Medium Resistive Substrates incl. ( d_b )</td>
<td>( C^{MRS}<em>{tsv} d_b = \frac{k_1^M \varepsilon</em>{SiO2} \varepsilon_0 l_v \left( \frac{l_v}{r_v} \right)^{k_2^M}}{d_b \frac{k_4^M}{r_v} \ln \left( 1 + k_5^M \frac{l_v}{r_v} \right)} ) (6)</td>
</tr>
<tr>
<td>Capacitance of an Isolated TSV in High Resistive Substrates</td>
<td>( C^{HRS}_{tsv} = \frac{63.34 \varepsilon_0 l_v}{\ln \left( 1 + 5.26 \frac{l_v}{r_v} \right)} ) (7)</td>
</tr>
<tr>
<td>Total Capacitance of a TSV in a Coupled Configuration</td>
<td>( C_t = C_{tsv} + k_1 C_{tsv} \left( \frac{p_v}{r_v} + k_3 \frac{p_v}{l_v} \right) \left[ k_4 \frac{l_v}{r_v} + k_6 \frac{p_v}{r_v} + k_8 \right] ) (8)</td>
</tr>
</tbody>
</table>
Coupling Capacitance between any two TSVs in a Coupled Configuration

\[
C_c = \frac{k_1 \varepsilon_0 l_v}{\ln \left( \frac{k_2 p_v}{r_v} \right)} \left[ 1 + k_3 \left( \frac{p_v}{r_v} \right)^{k_s} + k_5 \left( \frac{l_v}{r_v} \right)^{k_6} \right] + k_7 \left( \frac{p_v}{r_v} \right)^{k_8} \left( \frac{d_b}{r_v} \right)^{k_{10}} \right]
\] (9)

Frequency Dependence of MRS Capacitance Models

\[
C_{MRS}(f) = C_{HRS} k_f^f + (C_{MRS} - C_{HRS} k_f^f) e^{\left( \frac{1}{k_2} + \frac{f}{k_3 f_{1v}} \right)} \] (10)

In the following section we present models for estimating parasitics of various TSV structures. For ease of reference all of the equations are gathered in one place in Table 1 and can be related directly to Figure 3. The coefficients for the capacitance equations depend on the configuration (i.e. whether in a row or bundle), barrier thickness, and substrate type, which are defined in Tables 3-7.

**TSV Specification**

In most cases TSVs can have a non-uniform tapered cross-section along their length; but for the purpose of modelling, we consider a general TSV process with a uniform circular cross-section with an annular dielectric barrier (SiO2) surrounding the Cu cylinder, neglecting the surface roughness. A moderate taper of up to about 10% does not cause a significant deviation. The TiN barrier layer has been neglected for the sake of simplicity, since its inclusion has a negligible effect on the parasitic parameters within the geometries considered in this work. Finally it is assumed that the substrate immediately surrounding the TSV structures is not explicitly grounded similar to [15]. For capacitance extraction of an isolated TSV, we consider the voltage reference in the field solver to be 2 TSVs situated a distance that is at least an order of magnitude greater than its largest dimension. For row and bundle configurations the adjacent TSVs serve as the reference.

![Figure 4 Cross-sectional view of a TSV including the dielectric barrier, depletion layer and the Si bulk region. The equivalent circuit for the structure is also shown.](image)

**RLC Extraction of an Isolated TSV**

In compliance with the general transmission line model for a lossy wire, the electrical model of an isolated TSV is composed of series resistive (\(R_{tsv}\)) and inductive (\(L_{tsv}\)) components, and shunt capacitive (\(C_{tsv}\)) and conductive (\(G_{tsv}\)) components. These parameters are a function of its geometry, *i.e.* radius, length, and barrier thickness, as well as the material properties of the conductor and surrounding dielectric layers. The two shunt components represent the properties of the SiO2 dielectric barrier and Si substrate. Since a TSV is laid out within the substrate which is doped to be either p-type or n-type and surrounded by a dielectric barrier, this eventually acts as a cylindrical MOS capacitor. Therefore when an electric potential is applied to a TSV, a depletion layer is formed within the substrate; when the surrounding substrate is fully depleted it can be modelled as a second dielectric layer in the substrate. Now the effect of the SiO2 barrier and depletion region can be represented by the capacitive terms (\(C_{d pai}\) and \(C_{d pir}\))
respectively while the substrate requires a capacitive (\(C_{si}\)) and a conductive (\(G_{si}\)) term dependent on the dopant concentration (see Figure 4).

It is well known that as the frequency increases the current crowds towards the periphery, a phenomenon known as the skin-effect. Typically, with increasing frequency, resistance increases while inductance decreases. Since the TSVs are relatively very short the resistance increase due to the skin-effect is insignificant in digital applications for frequencies up to 50 GHz, as the m\(\Omega\) resistance is completely swamped by the driver impedance.

**Resistance:**
The resistance of a TSV is a function of the conductivity of the material and the cross-sectional area. It is not dependant on the properties of the surrounding materials. This relationship can be described using the traditional formula (1) given in Table 1.

The variation of \(\frac{R_{\text{par}}}{\rho l}\) with cross-sectional area is shown in Figure 5. The figure also shows the closeness with which the model in (1) tracks the simulated results; it is accurate to within 97% for the simulated ranges \(5 \mu m \leq r_{c} \leq 50 \mu m\) and \(50 \mu m \leq l_{c} \leq 250 \mu m\).

**Inductance:**
The inductance (\(L_{\text{tsv}}\)) of an isolated TSV is a function of the geometrical parameters of radius and length, and effective permeability of the surrounding medium. Due to the nature of the electromagnetic field distribution, the dependence of inductance on \(d_{0}\) is negligible. For the given range, it can be estimated from (2) in Table 1.

The empirical constants in (2) have been estimated for the ranges, \(5 \mu m \leq r_{c} \leq 40 \mu m\) and \(50 \mu m \leq l_{c} \leq 400 \mu m\). In Figure 6 we show the variation of \(L_{\text{tsv}}\) with \(\frac{l_{c}}{r_{c}}\) and the model predicted values against the field-solver extracted values. This functional form is validated with measurement results in [13] for TSVs with a square cross-section. Our semi-empirical model predicts the self-inductance with a minimum accuracy of 97% for aspect ratios \(\frac{L}{r_c}\) in the range \(0.5 \leq \frac{l_{c}}{r_{c}} \leq 300\) while the self inductance model proposed in [16] has a minimum accuracy of 89%.

[13]
To accurately extract the capacitance of the TSV the MOS capacitance effect as explained previously must be considered in conjunction with the electro-static behaviour. Most widely used electromagnetic field-solvers are incapable of handling the physical domain related to this effect. In order to carry out the capacitance extraction we estimated the depletion layer analytically by solving Poisson’s equation. This value is then used as a parameter in the parasitic extraction tool assuming the Cu cylindrical bar is surrounded by two dielectric layers, one of SiO\(_2\) (\(\varepsilon_r = 3.9\)) and the other of Si (\(\varepsilon_r = 11.9\)).

Using the same methodology as in [14], [15], [30] we derive the expression for dielectric barrier layer thickness. Poisson’s equation in cylindrical co-ordinates is:

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2} + \frac{\partial^2 \phi}{\partial z^2} = -\frac{qN_a}{\varepsilon_{Si}}
\] (11)

where \(\phi\) is the potential, \(N_a\) the p-type bulk doping concentration, \(q\) the elementary charge of an electron (1.60218×10\(^{-19}\) C) and \(\varepsilon_{Si}\) the permittivity of Si. Assuming that the electric potential does not vary with the angle around the TSV and the axial distance, (9) reduces to:

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \phi}{\partial r} \right) \approx -\frac{qN_a}{\varepsilon_{Si}}
\] (12)

Figure 5 The variation with cross-sectional area \(\pi r_c^2\) of (a) \(R_{TSV}/\rho l_v\), and (b) percentage error between simulated and predicted values using (1)
Solving (12), the electrical potential ($\phi_s$) at the Si-SiO$_2$ barrier can be expressed as:

$$\phi_s = \frac{qN_a}{2\varepsilon_{si}} \left[ r_d \ln \left( \frac{r_d}{r_b} \right) - \frac{r_d^2}{2} + \frac{r_b^2}{2} \right]$$  \hspace{1cm} (13)

where $r_b = r_v + d_b$, $r_d = r_v + d_b + t_{dpl}$ and $t_{dpl}$ is the depletion layer thickness. When the depletion layer thickness reaches its maximum, $\phi_s = 2\phi_F$ which is equal to $2\frac{k_B T}{q} \ln \left( \frac{N_d}{n_i} \right)$, where $k_B$ is Boltzmann’s constant ($1.3807 \times 10^{-23}$ m$^2$kg$s$^{-2}K$^{-1}$), $T$ the absolute temperature and $n_i$ the intrinsic carrier concentration of Si ($1.45 \times 10^{10}$ cm$^{-3}$ at 300 K). Now the maximum depletion layer thickness ($t_{dpl}$) is the solution of:

$$2 \frac{k_B T}{q} \ln \left( \frac{N_d}{n_i} \right) = \frac{qN_a}{2\varepsilon_{si}} \left[ \frac{t_{dpl}}{2} \left( 2r_v + 2d_b + t_{dpl} \right) + \left( r_v + d_b + t_{dpl} \right)^2 \ln \left( 1 + \frac{t_{dpl}}{r_v + d_b} \right) \right]$$  \hspace{1cm} (14)

Si substrates are typically identified by resistivity ($\rho_{si}$) rather than doping concentration ($N_d$). Therefore, we use an empirical relationship given in [31] to convert substrate resistivity to the corresponding doping concentration. For boron-doped silicon, the doping concentration can be estimated from:

$$\text{Figure 6: The variation with } \frac{L_v}{r_v} \text{ of (a) } \frac{L_{exp}}{\mu l_v}, \text{ and (b) the percentage error between simulated and predicted values using (2).}$$
\[ N = \frac{1.330 \times 10^{16}}{\rho} + \frac{1.082 \times 10^{17}}{\rho[1 + (54.56 \rho)^{1.105}]} \]  

For \(5 \mu m \leq r_0 \leq 35 \mu m\) and \(1 \mu m \leq d_b \leq 5 \mu m\), the depletion layer thickness is estimated by solving (14) numerically for various substrate resistivities. The variation with resistivity for various combinations of TSV radius and dielectric thickness is shown in Figure 7. It is evident that for the considered range of dimensions, the depletion layer thickness does not vary significantly for very small substrate resistivities. However, for higher resistivities above 1 cm, the variation of depletion layer thickness is significant. In order to model this effect, we propose an analytical equation to avoid numerical calculations; the relationship between \(\rho_{si}\) and \(t_{dpl}\) can be expressed in the form of:

\[ t_{dpl} = \gamma_1 + \left(\frac{\rho_{si}}{\gamma_2}\right)^{\gamma_3} \]  

Where \(\gamma_1\), \(\gamma_2\), and \(\gamma_3\) are empirically estimated coefficients defined in Table II. The depletion layer thickness is used as an input parameter in the field-solver simulations. Without this critical step an error of up to 60% can occur in the capacitance prediction through neglecting the MOS capacitive effect [35].

![Figure 7 Variation of depletion layer thickness with substrate resistivity for different radii and dielectric barrier thicknesses](image)

Table 2 Empirical Coefficients for Table 2

<table>
<thead>
<tr>
<th>Range</th>
<th>(\gamma_1)</th>
<th>(\gamma_2)</th>
<th>(\gamma_3)</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(10^{-4} \leq \rho_{si} \leq 10^{-2})</td>
<td>0.0008</td>
<td>12.4848</td>
<td>0.6225</td>
<td>3.0</td>
</tr>
<tr>
<td>(10^{-2} \leq \rho_{si} \leq 10^{-1})</td>
<td>0.0015</td>
<td>3.9043</td>
<td>0.7552</td>
<td>1.2</td>
</tr>
<tr>
<td>(10^{-1} \leq \rho_{si} \leq 10^{1})</td>
<td>-0.0508</td>
<td>16.7230</td>
<td>0.4237</td>
<td>1.0</td>
</tr>
<tr>
<td>(10^{1} \leq \rho_{si} \leq 10^{4})</td>
<td>-0.1322</td>
<td>13.7684</td>
<td>0.4075</td>
<td>5.9</td>
</tr>
<tr>
<td>(10^{4} \leq \rho_{si} \leq 10^{6})</td>
<td>-4.4725</td>
<td>0.1356</td>
<td>0.2579</td>
<td>8.7</td>
</tr>
</tbody>
</table>

The capacitance of an isolated TSV \(C_{tsv}\) is a function of its geometry, i.e. radius, length, and thickness of barrier, as well as the material properties of the surrounding dielectrics and the substrate (Figure 4). Therefore, \(C_{tsv}\) is the series...
combination of the capacitances formed by the dielectric barrier \( (C_{ax}) \), the depletion layer \( (C_{dpl}) \) and the substrate \( (C_{sl}) \). It can be mathematically expressed as:

\[
C_{tsv} = \left( \frac{1}{C_{ax}} + \frac{1}{C_{dpl}} + \frac{1}{C_{sl}} \right)^{-1}
\]  

(17)

Among these, \( C_{ax} \) and \( C_{dpl} \) are comparatively larger than \( C_{sl} \). In this work, we provide empirical models for the effective capacitance \( (C_{tsv}) \) rather than the individual components, as electrical level simulations only require self and coupling terms. We discuss three different substrate types, namely low resistive substrates, medium resistive substrates and highly resistive substrates identified in the equations with superscript terms LRS, MRS and HRS respectively. It is worth noting here that the capacitance for an isolated TSV in a low resistive substrate is the worst-case capacitance that can be expected; when the substrate resistivity increases the isolated TSV capacitance reduces, as is to be intuitively expected.

(a) **Low Resistive Substrates**: When the substrate resistivity is very low \( (\rho_{st} < 0.1 \text{ cm}) \), \( C_{si} \) is very high and \( C_{sl} \) is essentially short-circuited (the substrate acts as a conductor), and no depletion layer is formed. Therefore, \( C_{tsv} \) is equal to \( C_{ax} \), which can be estimated empirically from (4) in Table 1. It takes the exact form of the co-axial cable capacitance equation with the fitting coefficients, \( k_1^L \) and \( k_2^L \). When the fitting parameters are the same as those for the coaxial cable case, i.e. \( k_1^L = 2\pi \) and \( k_2^L = 1 \), the error is about 5%. For the empirically estimated coefficients we propose \( (k_1^L = 2.87\pi \) and \( k_2^L = 1.39 \) the predicted capacitance value is approximately 97% accurate for the dimensional ranges 5 \( \mu \text{m} \leq r_v \leq 35 \mu \text{m}, 0.2 \mu \text{m} \leq d_b \leq 1 \mu \text{m} \) and 20 \( \mu \text{m} \leq l_v \leq 100 \mu \text{m} \). The coaxial cable equation has been used in [32] and verified with simulations for TSVs in low resistive substrates but the error margin is not given.

A plot of the variation of \( \frac{C_{tsv}}{\varepsilon_0 l_v} \) with \( \frac{d_b}{r_v} \) for isolated TSVs in a low resistive substrate is shown in Figure 8.
(b) **Medium Resistive Substrates:** In this case, all the surrounding material regions contribute to the overall capacitance and the equation for the capacitance of an isolated TSV in a medium resistive substrate takes the form (5) as described in Table 1. Given that the thickness of the isolation oxide layer $d_b$ is usually a constant for a given technology, we also present a simplified equation (6) for a default $d_b$ value of 0.2 µm. We adopt the practice throughout of presenting two sets of equations for $C_{MRS}$ for all configurations (isolated, row and bundle), one for a fixed $d_b$ value of 0.2 µm and one that explicitly contains $d_b$ as a parametric variable. The reason is the advantage of the reduced error for the coupled configurations where design space exploration to identify trends is the objective.

Table 3 Coefficients for isolated TSV capacitance for (5) for MRS ($\rho_{si} = 10 \Omega cm$)

<table>
<thead>
<tr>
<th>Eq</th>
<th>$k_1^M$</th>
<th>$k_2^M$</th>
<th>$k_3^M$</th>
<th>$k_4^M$</th>
<th>% Std. Dev.</th>
<th>% Avg. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5)</td>
<td>4.6047</td>
<td>0.0721</td>
<td>0.1927</td>
<td>-</td>
<td>6.71</td>
<td>9.71</td>
</tr>
<tr>
<td>(6)</td>
<td>4.6900</td>
<td>-0.1746</td>
<td>-0.8265</td>
<td>0.0255</td>
<td>6.01</td>
<td>8.98</td>
</tr>
</tbody>
</table>

The constants $k_1^M$, $k_2^M$ and $k_3^M$ for the most typical silicon resistivity of 10 cm are defined in Table 3 for the ranges 5 µm ≤ $r_v$ ≤ 35 µm, 0.2 µm ≤ $d_b$ ≤ 1.0 µm and 20 µm ≤ $l_v$ ≤ 140 µm. The variation of $\frac{C_{TSV}}{\epsilon_0 l_v}$ with $\frac{r_v}{d_b}$ is shown in Figure 8.

Figure 8 The variation with $\frac{d_b}{r_v}$ of (a) $\frac{C_{TSV}}{\epsilon_0 l_v}$, and (b) the percentage error between simulated and predicted values using (4) for an isolated TSVs in low resistive substrates.
for simulated and model-predicted values. The model proposed for this substrate type shows a 24% maximum error but that is only for a small number of pathological cases; the average error is around 3.6%.

![Graph](image)

**Figure 9** The variation with $\frac{l_v}{r_v}$ of (a) $\frac{C_{tsv}}{\varepsilon_0 l_v}$, and (b) the percentage error between simulated and predicted values using (4) for an isolated TSVs in medium resistive substrates

(c) **High Resistive Substrates**: As the substrate resistivity is very high, i.e. $\rho_{sil} > 1000 \ \Omega \text{cm}$, $G_{sil}$ is nearly zero, and the depletion layer thickness is significant. Hence the Si substrate and TSV form a capacitor with the SiO2 barrier and the substrate as the combined dielectric. In this case, the TSV capacitance $C_{tsv}$ can be expressed using the empirical equation (7) with a maximum error contained to 6% for the simulated range, $5 \ \mu m \leq r_v \leq 40 \ \mu m$ and $20 \ \mu m \leq l_v \leq 140 \ \mu m$. Even though (7) does not contain a term that represents dielectric barrier thickness $d_b$, the proposed empirical self-capacitance formula has a maximum error contained to within 8% for the range of $0.2 \ \mu m \leq d_b \leq 1 \ \mu m$ and for the range the effect of $d_b$ is not measurable in the simulator.
The self ($L_s$) and mutual ($L_m$) inductance terms for a system of $n$ TSVs are defined by:

$$L_{\text{bundle}} = \begin{bmatrix}
L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\
L_{2,1} & L_{2,2} & \cdots & L_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
L_{n,1} & L_{n,2} & \cdots & L_{n,n}
\end{bmatrix}$$

(18)

where diagonal elements represent the self inductance terms, and off diagonal elements the mutual inductance terms. The self inductance of a TSV is not affected due to the presence of neighbouring TSVs, and it can be estimated from the equation derived for an isolated TSV in (2).

Figure 11 shows the inductive coupling between the middle (M) TSV and all other surrounding TSVs in a high resistive substrate (from field solver simulations). It shows that inductive coupling is long range and therefore the inductance matrix is well populated, with all elements being non-negligible.

The mutual inductance ($L_m$) between any two TSVs, for the physical geometry ranges $5 \mu m \leq r_v \leq 40 \mu m$, $20 \mu m \leq l_v \leq 140 \mu m$ and $5 \mu m \leq s_v \leq 200 \mu m$, is defined by (3) in Table 1, where $d_v = s_v + 2(r_v + d_v)$ is the centre-to-centre distance between the lines. This model maintains an accuracy of 92% over the simulated range of physical dimensions similar to the model proposed in [16] (see Figure 12).
Figure 11 Inductive coupling between M TSV and surrounding TSVs in a $7 \times 7$ bundle. Values in the graph is normalized to middle TSV self inductance and physical dimensions are $r_v = 20 \, \mu m$, $l_v = 80 \, \mu m$ and $s_v = 20 \, \mu m$. 
Capacitance Extraction of Coupled TSVs

This section describes the capacitance models for \( n \) coupled TSVs including self and coupling terms. All self and coupling capacitance terms for a system of coupled TSVs are given by:

\[
C_{\text{bundle}} = \begin{bmatrix}
C_{1,1} & -C_{1,2} & \ldots & -C_{1,n} \\
-C_{2,1} & C_{2,2} & \ldots & -C_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
-C_{n,1} & -C_{n,2} & \ldots & C_{n,n}
\end{bmatrix}
\]  

(19)

In (19), the diagonal element \( C_{ii} \) represents the sum of the self \((C_{i0})\) and inter-via coupling capacitances \((C_{i,j})\) as given in (20):

\[
C_{ii} = C_{i0} + \sum_{j=1}^{n} C_{i,j}
\]  

(20)

The capacitance matrix is sparse; the main diagonal and adjacent diagonals representing coupling terms to nearest neighbours are populated while the other entries are vanishingly small in comparison.

Figure 13 shows the capacitive coupling between the M TSV of a 7×7 bundle and its surrounding TSVs. This figure clearly demonstrates that coupling terms to nearest neighbours are much more significant than those to non-adjacent TSVs. Therefore the representative structure for capacitance extraction in the general case can be restricted to a 3×3
bundle. However we also consider the case of three coupled TSVs, as technological constraints may dictate that the TSVs are laid out in a row, when the field distribution is different.

We present two self-consistent equations for total and coupling capacitances in (8) and (9) from which the capacitance matrix for n coupled TSVs can be obtained. The total capacitance \( C_t = C_{t,i} \) formula is given in (8) in Table 1, where \( C_{t,i} \) is the capacitance of an isolated TSV (3) depending on the substrate resistivity:

\[
C_{t,i} = \begin{cases} 
C_{t,i}^{MRS} & \text{for medium resistive substrates} \\
C_{t,i}^{HRS} & \text{for high resistive substrates}
\end{cases}
\]

(21)

In (8) the constants \( k_2 \) and \( k_3 \) are negative, and therefore as \( \rho \) approaches infinity, \( C_t \) approaches \( C_{t,i} \), providing a self consistent model.

The formula for the coupling capacitance \( C_c = C_{c,i,j} \) terms of (19) for \( i \neq j \) is of the form of (9). In very low resistive substrates, the substrate acts as a conducting medium and hence there exists no coupling between TSVs. Therefore, \( C_c \) is equal to \( C_{t,i}^{LRS} \) and \( C_c \) is zero.

When TSVs are arranged in a row, the general unit of representation is three parallel TSVs. We define \( C_{t,3} \), \( C_{c,3} \) and \( C_c \) as the total capacitance of the middle conductor, the self capacitance of the conductor at the edge, and the coupling capacitance between two conductors in a row, respectively. The constants corresponding to self and coupling capacitances of three coupled TSVs are defined in Table 5 for medium and high resistive substrates. The absolute maximum error for each case shown in Table 5 is less than 15%, but the average error is less than 2% for medium resistive substrate case and less than 1% for high resistive substrates.

In the case of a 3×3 bundle, the symmetry can be exploited to reduce the number of terms to be investigated. With reference to the naming convention given in Figure 3(c) the distances from M TSV to N, E, S and W TSVs are the same, while the distances to NE, NW, SE and SW TSVs are also equal. Therefore, the capacitance formulae for the total capacitance of M, N, and NE TSVs \( (C_{t,i,j}) \), and the coupling terms to their nearest neighbours \( (C_{c,i,j}) \) as defined in Figure 3(c) are a representative unit for a TSV bundle of any size, as the coupling terms to non-adjacent TSVs are negligible. The corresponding constants \( k_1, k_2, ..., k_8 \) for total and coupling terms in a TSV bundle as defined in Figure 3(c) are mentioned in Table 6. Similar to a row, the average absolute error is under 18% for all terms.

Figure 13 Capacitive coupling between M TSV and surrounding TSVs in a 7×7 bundle. Values in the graphs are normalized to middle TSV total capacitance and physical dimensions are \( r_v=20 \mu m \), \( l_v=80 \mu m \) and \( s_v=20 \mu m \).
Also similar to a row, we present a model for a default isolation layer thickness of \( d_b = 0.2 \, \mu m \) with the same motivation in Table 7. Again this results in an improvement in the error to under 10%. Table 7 also presents the coefficients for high resistive substrates with an average absolute error under 8% for all terms. The usage of the capacitance estimation models (8) and (9) for TSVs arranged in a row and a bundle is summarized in Table 8.

**Table 4 Constants for self and coupling capacitance terms of three-parallel TSVs for medium resistive substrates (5 \( \mu m \leq r_v \leq 25 \, \mu m \), 50 \( \mu m \leq l_v \leq 100 \, \mu m \) and 20 \( \mu m \leq s_v \leq 100 \, \mu m \)) and high resistive substrates (5 \( \mu m \leq r_v \leq 25 \, \mu m \), 20 \( \mu m \leq l_v \leq 140 \, \mu m \) and 20 \( \mu m \leq s_v \leq 140 \, \mu m \)) for varying \( d_b \) (0.1 \( \mu m \leq s_v \leq 100 \, \mu m \))**

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( k_1 )</th>
<th>( k_2 )</th>
<th>( k_3 )</th>
<th>( k_4 )</th>
<th>( k_5 )</th>
<th>( k_6 )</th>
<th>( k_7 )</th>
<th>( k_8 )</th>
<th>( k_9 )</th>
<th>( k_{10} )</th>
<th>Avg. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRS 10 ( \Omega \cdot cm )</td>
<td>( C_{L_{1k}} )</td>
<td>-3.5195</td>
<td>-0.0499</td>
<td>-0.2092</td>
<td>-2.2521</td>
<td>0.3988</td>
<td>2.871</td>
<td>0.4703</td>
<td>3.5586</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( C_{E_{1b}} )</td>
<td>-1.8599</td>
<td>-0.0367</td>
<td>-0.1997</td>
<td>-2.4033</td>
<td>0.3556</td>
<td>3.6245</td>
<td>0.46</td>
<td>2.9971</td>
<td>-</td>
<td>-</td>
<td>15.31</td>
</tr>
<tr>
<td>( C_{E_{c}} )</td>
<td>-3.936</td>
<td>-0.07005</td>
<td>0.45033</td>
<td>1.0842</td>
<td>0.5091</td>
<td>1.0241</td>
<td>0.6224</td>
<td>2.7724</td>
<td>0.7587</td>
<td>0.4488</td>
<td>20.80</td>
</tr>
</tbody>
</table>

**Table 5 Constants for self and coupling capacitance terms of three-parallel TSVs for medium resistive substrates (5 \( \mu m \leq r_v \leq 25 \, \mu m \), 50 \( \mu m \leq l_v \leq 100 \, \mu m \) and 20 \( \mu m \leq s_v \leq 100 \, \mu m \)) and high resistive substrates (5 \( \mu m \leq r_v \leq 25 \, \mu m \), 20 \( \mu m \leq l_v \leq 140 \, \mu m \) and 20 \( \mu m \leq s_v \leq 140 \, \mu m \))**

<table>
<thead>
<tr>
<th>Substrate resistivity</th>
<th>( k_1 )</th>
<th>( k_2 )</th>
<th>( k_3 )</th>
<th>( k_4 )</th>
<th>( k_5 )</th>
<th>( k_6 )</th>
<th>( k_7 )</th>
<th>( k_8 )</th>
<th>( k_9 )</th>
<th>( k_{10} )</th>
<th>% Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium (10 ( \Omega \cdot cm ))</td>
<td>( C_{L_{1k}} )</td>
<td>12.6656</td>
<td>-0.0837</td>
<td>-0.4129</td>
<td>-0.1907</td>
<td>-0.7697</td>
<td>0.0067</td>
<td>2.6992</td>
<td>6.0228</td>
<td>6.51</td>
<td>9.42</td>
</tr>
<tr>
<td>( C_{L_{1b}} )</td>
<td>20.2233</td>
<td>-0.0067</td>
<td>-0.1885</td>
<td>1.3038</td>
<td>0.2626</td>
<td>20.4562</td>
<td>-3.6585</td>
<td>-0.6956</td>
<td>4.76</td>
<td>6.29</td>
<td></td>
</tr>
<tr>
<td>( C_{E_{c}} )</td>
<td>13.1579</td>
<td>0.6567</td>
<td>3.3072</td>
<td>-1.346</td>
<td>-7.5764</td>
<td>-2.2163</td>
<td>10.2532</td>
<td>0.1475</td>
<td>6.84</td>
<td>7.94</td>
<td></td>
</tr>
<tr>
<td>High (10 ( \Omega \cdot cm ))</td>
<td>( C_{L_{1b}} )</td>
<td>3.3353</td>
<td>-0.0483</td>
<td>-0.2046</td>
<td>0.3421</td>
<td>0.1804</td>
<td>8.6351</td>
<td>-2.4159</td>
<td>-0.1084</td>
<td>2.65</td>
<td>3.41</td>
</tr>
<tr>
<td>( C_{E_{1b}} )</td>
<td>7.7772</td>
<td>-0.0427</td>
<td>-0.1801</td>
<td>1.2417</td>
<td>0.0170</td>
<td>2.0284</td>
<td>-2.5074</td>
<td>-1.1822</td>
<td>2.08</td>
<td>2.81</td>
<td></td>
</tr>
<tr>
<td>( C_{E_{c}} )</td>
<td>19.3154</td>
<td>0.8010</td>
<td>2.8188</td>
<td>-2.6038</td>
<td>1.2397</td>
<td>-0.2059</td>
<td>0.7218</td>
<td>0.2656</td>
<td>1.93</td>
<td>2.18</td>
<td></td>
</tr>
</tbody>
</table>

**Table 6 Constants for self and coupling capacitance terms of 3 x 3 TSV bundle (see Figure 3(c) for capacitance definitions). Simulated ranges are: 5 \( \mu m \leq r_v \leq 25 \, \mu m \), 50 \( \mu m \leq l_v \leq 100 \, \mu m \) and 20 \( \mu m \leq s_v \leq 100 \, \mu m \) for medium resistive substrates and 5 \( \mu m \leq r_v \leq 25 \, \mu m \), 20 \( \mu m \leq l_v \leq 140 \, \mu m \) and 20 \( \mu m \leq s_v \leq 140 \, \mu m \) for high resistive substrates. For varying \( d_b \) (0.1 \( \mu m \leq s_v \leq 100 \, \mu m \))**

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( k_1 )</th>
<th>( k_2 )</th>
<th>( k_3 )</th>
<th>( k_4 )</th>
<th>( k_5 )</th>
<th>( k_6 )</th>
<th>( k_7 )</th>
<th>( k_8 )</th>
<th>( k_9 )</th>
<th>( k_{10} )</th>
<th>% Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRS 10 ( \Omega \cdot cm )</td>
<td>( C_{L_{1k}} )</td>
<td>-2.416</td>
<td>0.0160</td>
<td>-0.133</td>
<td>16.247</td>
<td>0.595</td>
<td>39.852</td>
<td>2.592</td>
<td>0.464</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( C_{L_{1N}} )</td>
<td>-0.787</td>
<td>0.009</td>
<td>-0.100</td>
<td>41.920</td>
<td>-0.438</td>
<td>-15.608</td>
<td>-0.671</td>
<td>1.037</td>
<td>-</td>
<td>-</td>
<td>16.18</td>
</tr>
<tr>
<td>( C_{E_{1NE}} )</td>
<td>-1.3771</td>
<td>-0.0002</td>
<td>-0.072</td>
<td>14.949</td>
<td>-0.436</td>
<td>2.777</td>
<td>0.341</td>
<td>-4.4761</td>
<td>-</td>
<td>-</td>
<td>13.92</td>
</tr>
<tr>
<td>( C_{E_{1d}} )</td>
<td>4.546</td>
<td>0.811</td>
<td>0.912</td>
<td>0.0425</td>
<td>0.916</td>
<td>6.974</td>
<td>0.145</td>
<td>1.282</td>
<td>0.330</td>
<td>13.03</td>
<td></td>
</tr>
<tr>
<td>( C_{E_{1l}} )</td>
<td>29.789</td>
<td>0.754</td>
<td>6.190</td>
<td>2.251</td>
<td>-0.188</td>
<td>0.449</td>
<td>3.680</td>
<td>0.0138</td>
<td>5.608</td>
<td>0.302</td>
<td>11.30</td>
</tr>
<tr>
<td>( C_{E_{1p}} )</td>
<td>-61.907</td>
<td>0.018</td>
<td>83.063</td>
<td>2.270</td>
<td>27.351</td>
<td>0.013</td>
<td>32.893</td>
<td>0.007</td>
<td>4.90</td>
<td>-0.0060</td>
<td>18.04</td>
</tr>
</tbody>
</table>
Table 7 Constants for self and coupling capacitance terms of 3 x 3 TSV bundle (see Figure 3(c) for capacitance definitions). Simulated ranges are: 5 μm ≤ r_v ≤ 25 μm, 50 μm ≤ l_v ≤ 100 μm and 20 μm ≤ s_v ≤ 100 μm for medium resistive substrates and 5 μm ≤ r_v ≤ 25 μm, 20 μm ≤ l_v ≤ 140 μm and 20 μm ≤ s_v ≤ 140 μm for high resistive substrates.

<table>
<thead>
<tr>
<th>Substrate resistivity</th>
<th>k_1</th>
<th>k_2</th>
<th>k_3</th>
<th>k_4</th>
<th>k_5</th>
<th>k_6</th>
<th>k_7</th>
<th>k_8</th>
<th>% Std dev.</th>
<th>% Avg. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium (10 Ωcm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) C_{t,M}</td>
<td>7.8700</td>
<td>-0.1555</td>
<td>-0.4629</td>
<td>14.8789</td>
<td>0.0654</td>
<td>0.0215</td>
<td>3.0253</td>
<td>-0.1754</td>
<td>4.90</td>
<td>5.97</td>
</tr>
<tr>
<td>(b) C_{t,N}</td>
<td>12.677</td>
<td>-0.1523</td>
<td>-0.4946</td>
<td>-2.3195</td>
<td>-1.8998</td>
<td>0.0115</td>
<td>3.0436</td>
<td>9.4361</td>
<td>4.35</td>
<td>5.47</td>
</tr>
<tr>
<td>(c) C_{t,N}</td>
<td>1.5765</td>
<td>-0.1193</td>
<td>-0.5828</td>
<td>57.7142</td>
<td>-0.0184</td>
<td>0.1815</td>
<td>2.5785</td>
<td>-2.0938</td>
<td>3.62</td>
<td>5.01</td>
</tr>
<tr>
<td>(d) C_{c,d}</td>
<td>4.1468</td>
<td>0.8935</td>
<td>3.6220</td>
<td>0.4482</td>
<td>-1.5836</td>
<td>-0.5117</td>
<td>-1.9408</td>
<td>0.4982</td>
<td>18.13</td>
<td>18.05</td>
</tr>
<tr>
<td>(e) C_{c,l}</td>
<td>14.527</td>
<td>0.8807</td>
<td>11.841</td>
<td>-0.3443</td>
<td>22.6152</td>
<td>-7.7184</td>
<td>-0.7118</td>
<td>-1.4476</td>
<td>12.20</td>
<td>9.40</td>
</tr>
<tr>
<td>(f) C_{c,p}</td>
<td>19.027</td>
<td>1.4098</td>
<td>21.473</td>
<td>-0.4478</td>
<td>6.1884</td>
<td>-5.9570</td>
<td>-0.8268</td>
<td>-1.5024</td>
<td>15.35</td>
<td>8.15</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(g) C_{c,M}</td>
<td>0.1505</td>
<td>-0.0071</td>
<td>-0.0291</td>
<td>0.1849</td>
<td>-1.9371</td>
<td>6.9577</td>
<td>-0.0131</td>
<td>-0.0354</td>
<td>3.10</td>
<td>3.89</td>
</tr>
<tr>
<td>(h) C_{c,N}</td>
<td>0.6876</td>
<td>-0.0390</td>
<td>-0.0583</td>
<td>1.8076</td>
<td>-0.2229</td>
<td>11.356</td>
<td>0.0402</td>
<td>-13.181</td>
<td>3.16</td>
<td>3.42</td>
</tr>
<tr>
<td>(i) C_{c,N}</td>
<td>0.3406</td>
<td>-0.0345</td>
<td>-0.0686</td>
<td>5.0708</td>
<td>-0.1530</td>
<td>-5.6346</td>
<td>-0.3859</td>
<td>-0.7643</td>
<td>2.94</td>
<td>3.67</td>
</tr>
<tr>
<td>(j) C_{c,d}</td>
<td>18.117</td>
<td>28.457</td>
<td>-1.734</td>
<td>-2.178</td>
<td>0.6000</td>
<td>-0.518</td>
<td>-0.470</td>
<td>0.188</td>
<td>1.45</td>
<td>7.60</td>
</tr>
<tr>
<td>(k) C_{c,l}</td>
<td>10.191</td>
<td>0.5490</td>
<td>-0.014</td>
<td>0.796</td>
<td>0.054</td>
<td>-1.157</td>
<td>-0.018</td>
<td>-0.600</td>
<td>1.92</td>
<td>1.77</td>
</tr>
<tr>
<td>(l) C_{c,p}</td>
<td>3.180</td>
<td>0.5440</td>
<td>-0.199</td>
<td>0.586</td>
<td>0.122</td>
<td>0.540</td>
<td>2.176</td>
<td>0.110</td>
<td>2.23</td>
<td>2.18</td>
</tr>
</tbody>
</table>

The usage of the capacitance estimation models (8) and (9) for TSVs arranged in a row and a bundle is summarised in Table 8.

Table 8 Usage of Capacitance estimation models for Coupled TSV Structures

<table>
<thead>
<tr>
<th>Capacitance Terms</th>
<th>MRS</th>
<th>HRS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Three TSVs</td>
<td></td>
</tr>
<tr>
<td>TSVs</td>
<td>(8) with C_{tsv} from (5) and k_1, k_2, ..., k_8 from Table 5 row (a)</td>
<td>(8) with C_{tsv} from (7) and k_1, k_2, ..., k_8 from Table 5 row (d)</td>
</tr>
<tr>
<td></td>
<td>(8) with C_{tsv} from (5) and k_1, k_2, ..., k_8 from Table 5 row (b)</td>
<td>(8) with C_{tsv} from (7) and k_1, k_2, ..., k_8 from Table 5 row (e)</td>
</tr>
<tr>
<td></td>
<td>(9) with k_1, k_2, ..., k_8 from Table 5 row (c)</td>
<td>(9) with k_1, k_2, ..., k_8 from Table 5 row (f)</td>
</tr>
<tr>
<td></td>
<td>(8) with C_{tsv} from (5) and k_1, k_2, ..., k_8 from Table 7 row (a)</td>
<td>(8) with C_{tsv} from (7) and k_1, k_2, ..., k_8 from Table 7 row (g)</td>
</tr>
<tr>
<td></td>
<td>(8) with C_{tsv} from (5) and k_1, k_2, ..., k_8 from Table 7 row (b)</td>
<td>(8) with C_{tsv} from (7) and k_1, k_2, ..., k_8 from Table 7 row (h)</td>
</tr>
<tr>
<td></td>
<td>(8) with C_{tsv} from (5) and k_1, k_2, ..., k_8 from Table 7 row (c)</td>
<td>(8) with C_{tsv} from (7) and k_1, k_2, ..., k_8 from Table 7 row (i)</td>
</tr>
<tr>
<td></td>
<td>(9) with k_1, k_2, ..., k_8 from Table 7 row (d)</td>
<td>(9) with k_1, k_2, ..., k_8 from Table 7 row (j)</td>
</tr>
<tr>
<td></td>
<td>(9) with k_1, k_2, ..., k_8 from Table 7 row (e)</td>
<td>(9) with k_1, k_2, ..., k_8 from Table 7 row (k)</td>
</tr>
<tr>
<td></td>
<td>(9) with k_1, k_2, ..., k_8 from Table 7 row (f)</td>
<td>(9) with k_1, k_2, ..., k_8 from Table 7 row (l)</td>
</tr>
</tbody>
</table>

Frequency Dependence of Capacitance Models in MRS

As mentioned at the start of this section all parameters (resistance, inductance and capacitance) generally vary with frequency, with varying levels of significance. The frequency dependence of resistance and inductance has been relatively well documented and our interest in this article is in capturing the frequency dependence of capacitive parasitics, which is generally not widely reported. Essentially as the frequency increases, for an isolated TSV, C_{tsv}^{MRS} decreases and asymptotically converges to C_{tsv}^{HRS}. This dependency as frequency increases from the default (low) frequency is captured in Eq. (6) of Table 1 which maintains integrity in the limits of very low (DC) and very high frequencies as discussed above. The coefficients for this model as well as its accuracy for the range 1 GHz to 50
GHz for all total and coupled components in all three configurations of isolated, row and bundle are given in Table 9.

Table 9 Coefficients for frequency-dependent TSV capacitances in Medium Resistive Substrates for all TSV topologies as defined in equation (10). Equations cover the geometrical ranges $5\mu m \leq r_v \leq 35\mu m$, $20\mu m \leq l_v \leq 100\mu m$, $20\mu m \leq s_v \leq 100\mu m$ and $0.1\mu m \leq d_b \leq 1\mu m$.

<table>
<thead>
<tr>
<th>Topology</th>
<th>TSV Capacitance</th>
<th>$k_1$</th>
<th>$k_2$</th>
<th>$k_3$</th>
<th>% Avg.</th>
<th>%Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolated</td>
<td>$C_{T,\text{self}}$</td>
<td>1.07</td>
<td>1.59</td>
<td>0.05</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Row</td>
<td>$C_{T,M}$</td>
<td>1.55</td>
<td>0.97</td>
<td>0.03</td>
<td>10.86</td>
<td></td>
</tr>
<tr>
<td>Row</td>
<td>$C_{T,Rd}$</td>
<td>1.47</td>
<td>0.8</td>
<td>0.03</td>
<td>13.13</td>
<td></td>
</tr>
<tr>
<td>Row</td>
<td>$C_c$</td>
<td>1.43</td>
<td>2.02</td>
<td>0.07</td>
<td>10.25</td>
<td></td>
</tr>
<tr>
<td>3x3 Bundle</td>
<td>$C_{T,M}$</td>
<td>1.35</td>
<td>1.06</td>
<td>0.06</td>
<td>8.84</td>
<td></td>
</tr>
<tr>
<td>3x3 Bundle</td>
<td>$C_{T,N}$</td>
<td>1.33</td>
<td>0.99</td>
<td>0.06</td>
<td>8.72</td>
<td></td>
</tr>
<tr>
<td>3x3 Bundle</td>
<td>$C_{T,NE}$</td>
<td>1.3</td>
<td>0.99</td>
<td>0.05</td>
<td>7.12</td>
<td></td>
</tr>
<tr>
<td>3x3 Bundle</td>
<td>$C_{c,d}$</td>
<td>0.83</td>
<td>1.02</td>
<td>0.03</td>
<td>5.57</td>
<td></td>
</tr>
<tr>
<td>3x3 Bundle</td>
<td>$C_{c,l}$</td>
<td>1.38</td>
<td>0.98</td>
<td>0.11</td>
<td>8.7</td>
<td></td>
</tr>
<tr>
<td>3x3 Bundle</td>
<td>$C_{c,p}$</td>
<td>1.2</td>
<td>0.29</td>
<td>0.12</td>
<td>9.38</td>
<td></td>
</tr>
</tbody>
</table>

TSV Compact Modelling Summary
Compact closed-form equations for calculating resistive, inductive and capacitive parasitic parameters of coupled TSVs in 3-D ICs were reported in this work. These models have been compared against existing models and shown to have a similar accuracy or better than others in the literature and for simpler and more intuitive functional forms suggested by the underlying physical behaviour. The general functional form has been validated experimentally, showing the usefulness of the models and the ability to characterise them for a given technology. They can be used for equivalent circuit formulation when calculating performance, SI and PI related metrics in design space and architectural explorations of 3-D ICs with good accuracy. They can also be directly utilised in physical design automation tools targeting a wide range of technological parameters comprising material properties as well as geometrical dimensions.
2.5. TSV modelling in the Radio Frequency range

RF modelling methodology

For RF TSV modelling, a specific test structure is required to enable RF testing; this structure, called a Dual Via Chain (DVC) is composed of two transmission line segments and two TSVs separated by a BRDL (Back Redistribution Line). Figure 14 gives the schematic view of a DVC device: RF pads are in coplanar configuration (GSG for Ground / Signal / Ground), with a 100 µm pitch, while access lines, also in GSG configuration are generally designed to be matched to 50 Ω. RF pads and access lines will both be de-embedded with a specific calibration operation. Metallic TSV pads are situated from each TSV part (at bottom and top) and a variable length line of BRD links the two TSV (with $L_{BL}$ length): this line has to be long enough to prevent coupling between the SV and short enough to introduce weak loss.

![Figure 14: Schematic view of the Dual Via Chain structure](image)

This Dual Via Chain, when being characterised or simulated, allows obtaining the scattering matrix ($S_{ij}$) of one TSV, after a specific calibration and a de-embedding operation. This measurement is compared to 3D simulation realized by ad hoc full wave 3D software; then by fitting measurements, the material parameters can be adjusted. A Spice-like model can be then extracted with frequency dependence for R, L, C and G parameters (see figure below). This model is a basement for building the parametric model.

![Figure 15: Methodology for RF TSV modelling](image)

The next paragraphs describe each part of the methodology steps (RF characterization, full 3D simulations, and parameters extraction) and give results for a 60 µm deep and 15 µm wide TSV, integrated in the ELITE project process flow.
**RF characterisation**

RF characterisation is realized with a Vector Network Analyser (VNA), which provides \([S_{ij}]\) scattering matrix of the Device Under Test (DUT) versus frequency. RF measurements are performed with RF probes in GSG configuration on the RF pads. Two specific operations are performed to remove all parasitic contribution from cables, probes and access lines. First a calibration is used to remove effects from cables and probes, in order to adjust reference planes close to the contact pads. Then a de-embedding step is done to eliminate effects from access lines and pads, and reference planes are adjusted close to the TSV pads (see figures below). Finally, in a third step, effects of the intervias line are precisely removed by using the characterization of DVC with different BRDL lengths: we extract BRDL propagation parameters and compute TSV \([ABCD]\) matrix by converting scattering matrix in ABCD matrix, and using the following relation:

\[
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix}_{Tot} = \begin{bmatrix}
A & B \\
C & D
\end{bmatrix}_{TSV} \begin{bmatrix}
A & B \\
C & D
\end{bmatrix}_{InterviaLine} \begin{bmatrix}
A & B \\
C & D
\end{bmatrix}_{TSV}
\]

where \([ABCD]\)Tot is the matrix of the DVC without pads and access lines.

![Diagram](image)

**Figure 16: Definition of the DUT Scattering matrix in RF mode / Calibration and de-embedding operations in RF characterisation**

**Layout of RF structures**

To calibrate and perform the de-embedding operations described above, in addition to the Dual Via Chain structure, various types of specific devices must be incorporate to the layout: design kit devices such as RF pads, thru and coplanar lines, and de-embedding devices such as open and short structures. As an example, the CPW (Coplanar Wave guide) transmission lines enable the extraction of the material characteristics (such as oxide permittivity, metal conductivity and material thickness). Including those specific RF propagating devices directly on the layout, intrinsic performance of the TSV will be easily extracted. The ELITE Test Vehicle layout of RF devices is shown below, including all the devices previously described.
Full 3D EM simulations

Simulation of the dual via chain structure

The scattering parameters (Sij) can be obtained by simulations using a full wave 3D software, here we used HFSS from Ansoft/Ansys. The Dual Via Chain structure is simulated between from few MHz to 5 GHz and all process characteristics can be depicted within the simulation: material parameters and geometry dimensions.

Figure 17: a) Layout of the RF devices for ELITE Test Vehicle structure  
  b) Details of the Dual Via Chain structure
Scattering matrix is given below for a 60 µm deep and 15 µm diameter TSV, which are the nominal dimensions of the TSV. $S_{11}$ is below -30 dB on the whole frequency range, and $S_{21}$ remains below -0.018 dB, leading to consider that one TSV induces weak RF losses.

**Influence of the intervia line**

Intervia line length influence is very important as shown in figure below: increasing this line length (from 10 to 30 µm) increases the RF loss, but using 50 µm long line decreases $S_{21}$ magnitude due to smaller coupling between TSV through silicon substrate. For 50 µm, device matching is also improved. A trade-off is then determined by optimization to find the optimal length (50 µm), short enough to minimize losses, but long enough to prevent substrate coupling. In a same way, the effect of the silicon thickness is shown below. As forecasted, a smaller thickness (50 and 60 µm) decreases the RF loss since the RF signal has shorter path to go through.

**Figure 18:** a) Dual Via Chain design for 3D simulations  b) Zoom on CPW access lines  c) Cross section and d) top view of the TSV
Building a \( \pi \)-model and parameters extraction

A \( \pi \)-shape model was chosen to accurately describe TSV behaviour (see below) and elements of the model (R, L, C, G) can be extracted after specific de-embedding operations to remove pads and access lines, as previously described. R and L represent the TSV serial resistance and inductance, while C and G, in parallel symbolise the parasitic capacitance and conductance in the substrate. To explain the frequency dependent evolution of C and G it is necessary to develop a physical model, where oxide and silicon contributions on capacitance are split into 2 parts: \( C_{\text{Ox}} \) and \( C_{\text{Si}} \). Consequently, \( G_{\text{Si}} \) represents the parasitic conductance due to silicon losses. \( C_{\text{Ox}} \) is the oxide capacitance due to the SiO\( _2 \) layer between copper and Silicon in the TSV side walls. Oxide capacitance can be expressed as following, where \( t_{\text{Si}} \) is the silicon thickness and \( r_{\text{TSV}} \) the TSV radius:

\[
C_{\text{Ox}} = \frac{2\pi \varepsilon_0 \varepsilon_{\text{Si}} t_{\text{Si}}}{\ln \left( \frac{r_{\text{TSV}}}{r_{\text{TSV}} - r_{\text{Ox}}} \right)}
\]

![Figure 19: Scattering parameter for the nominal TSV](image1)

![Figure 20: Scattering parameter for the nominal TSV when: a) intervia line increase  b) Silicon thickness increases](image2)
And the following relations can easily be demonstrated:

\[ \lim_{\omega \to 0} C = C_{Ox} \]
\[ \lim_{\omega \to \infty} C = \frac{C_{Ox}C_{Si}}{C_{Ox} + C_{Si}} \]

Leading to:

\[ C = \frac{C_{Ox} \left[ G_{Si}^2 + C_{Si} \omega^2 (C_{Si} + C_{Ox}) \right]}{G_{Si}^2 + \omega^2 (C_{Si} + C_{Ox})^2} \quad \text{and} \quad G = \frac{G_{Si} C_{Ox} \omega^2}{G_{Si}^2 + \omega^2 (C_{Si} + C_{Ox})^2} \]

For the ELITE TSV, with a diameter of 15 µm and a silicon thickness of 60 µm, we briefly gives afterward the main process characteristics: a SiO₂/SiN isolating oxide layer is deposited on the TSV side walls, with 330 nm of thickness. Metal layer material is copper with a barrier layer of 10 nm thick TaN/Ta and a copper seed layer (120nm). Concerning Metal1 layer, on frontside, and BRDL on backside, copper is used to realize pads, access lines and intervia lines. The R, L, C and G parameters variations of the π-model were extracted (see the following figure). For one TSV, resistance increases from DC value, (220 mΩ) up to 0.86 Ω at 20 GHz owing to the skin effect in the TSV copper side walls. TSV static resistance is a combination of resistance of the TSV and resistance of the top and bottom TaN/Ta/Cu pads. \( \delta \) is the frequency where the skin effect begins, is close to 95 MHz, given by

\[ \delta = \frac{1}{\pi \mu_0 \sigma_{Cu} \delta^2} \]

DC value of the inductance is close to 140 pH; L steeply decreases until \( \delta \) due to slow wave effect (the self-inductance falls down). And from 95 MHz, as eddy currents become effective, L moderately decreases and reaches 136 pH at 10 GHz. Capacitance hugely varies on the whole frequency range: at low frequency, as previously described, C is roughly equal to the oxide capacitance, \( C_{Ox} = 370 \text{ fF} \), which expression was previously given. \( C_{Si} \) can be deduced from the limit of C at 20 GHz and is very low, close to 8.5fF. This capacitance physically represents the silicon contribution capacitance. Using the models proposed in section 2.4 for the same TSV dimensions, the capacitance at 1 GHz is calculated to be 306 fF and at 20 GHz 8.9 fF. This exhibits a close match to these RF measurements. Conductance exhibits quite weak value (750 µS) due to the silicon substrate conductivity (standard value for the ELITE wafers). We will demonstrate in the next section how substrate conductivity influences in a strong way the TSV performance. 306 DC 8.9 HRS
Simulations based on a Design of Experiment (DOE) were carried out to identify the most significant parameters, and it was demonstrated that the Silicon conductivity has a great influence on TSV performance, in particular for the silicon capacitance, see below where it is shown than the 2 mains contributors are $\sigma_{Si}$ and $t_{Si}$.

**Substrate conductivity influence**

![Figure 22 Parasitic parameters of the π-model](image)

![Figure 23: Quantification and distribution of technological factors effects on C_{Si}](image)
In 3D integration, according to the final product applications (see also Chapter 2, related work section), TSVs can be processed in substrates with different conductivities: for instance, epi-substrates are habitually used in CMOS and BiCMOS processes to play the role of "ground planes", standard silicon is generally dedicated to bipolar and CMOS technologies and High Resistivity (HR) substrates are chosen for microwave devices and RF MEMS (Micro-ElectroMechanical Systems). Since the TSV can be used in systems for all these applications, the influence of silicon conductivity on TSV behaviour is a real issue. We give below the simulation result of Sij parameters up to 20 GHz for a dual via chain structure in 3 substrate configurations: HR, standard and P+ doped Silicon (conductive silicon with very low resistivity):

For reflexion parameter, the gap between doped and HR silicon is tremendous due to great mismatch of the structure when TSVs are processed in high conductive silicon. In one hand, for HR substrate and below 10 GHz, DVC reflexion remains lower than 20 dB; hence its global impedance is supposed to be quite close to 50 Ω. On the other hand, we observe a great mismatch for epi-silicon, with $S_{11}$ values greater than 10 dB from DC to 20 GHz, leading to decrease DVC impedance down to 8 Ω. Transmission remains lower than 0.26 dB for HR silicon on the whole frequency range, but, as expected, reaches high values for doped silicon because of the conductive and dielectric losses in the substrate: more than 10 dB at 20 GHz. For standard Si, $S_{21}$ is quasi-constant with less than 2.5 dB on the whole frequency range.

![Graph of Reflexion and Transmission](image-url)
RF characterisation
RF devices were characterised in the 0.1 to 20 GHz frequency range. An IF band was chosen very low to obtain accurate measurements and 178 sites of measured were enable on the wafer (including centre and edges).

Coplanar lines
In a first step the coplanar lines were characterised and the figure below shows the result for Front Side (FS) lines: it compares 3 lines with different lengths: 60, 140 and 220 µm; please note that the access lines are included, leading to a total length of 260, 340 and 420 µm respectively. $S_{11}$ is lower than 20 dB for $F<5$ GHz, and reach a -18 dB at 20 GHz, showing that impedance is quite close to 50 Ω. The RF loss is 1.6 dB at 20 GHz for the 220 µm long line. Values for $S_{21}$ are respectively 1.21 dB, 1.41 dB and 1.62 dB at 20 GHz.

Figure 25 Substrate conductivity impact on TSV parameters
“Embedded” lines which are so called due to their position in the backside (BS) of the wafer (with front side access only) are also characterised and the comparing with the FS line is given below (with equivalent lengths). The process is different since a Ta/TaN layer and a copper layer of 0.3µm thick are deposited; we observed quasi no difference in the matching impedance (0.7 dB for S_{11}) even if the FS line has better matching; but an additional loss of 0.5 dB at 10 GHz is observed for the “embedded” line. This additional loss can be explained by the presence of 2 TSV (placed at the beginning and end of the line since measurements are only possible from front side), and a thinner thickness of copper layer in backside (0.3 µm versus 2 µm in front side).

TSV chains

TSV performance through dual via chain measurements: 3 chains of 2, 4 and 6 TSV respectively were characterised and results are given below. The intervia line is constant with 20µm between pads. At 10 GHz, RF loss is 1.26 dB, 1.53 dB and 1.8 dB respectively. We can deduce that an isolated TSV almost induces 0.135 dB of loss at 10 GHz. Dual via chain matching is always lower than 16dB on the whole frequency range.
When comparing the TSV chains with frontline with comparable lengths, we remark that the $S_{21}$ module differences increases when the number of TSV increases: 0.18, 0.33 and 0.46 dB respectively for 2, 4 and 6 TSV (figure below). This increasing in the difference is due to the RF loss in the intervia backside lines between TSV which has higher loss than FS line.

**Figure 28: RF performance of the TSV chains (2, 4 and 6 TSV)**

TSV π-shape model parameters extraction

From the model presented above, parasitic elements of the model are extracted after de-embedding of access and pads. R, L, C and G for one TSV are presented below for 3 measurement sites. The resistance remains lower than 90 mΩ on the whole frequency range even if a smooth skin effect is observed. The DC value is 15 mΩ due to metallic loss in the copper cylinder and Ta/TaN side walls, and the skin effect frequency $F_\delta$ is equal to 93.8 MHz. The self-inductance varies between 32 and 35 pH for $F > F_\delta$ which confirms that TSV is a good candidate to compete standard wire bonding solution.

**Figure 29: Comparison between TSV chains and equivalent lengths FS lines**
The oxide capacitance (extracted at very low frequency) is close to 300fF and the silicon capacitance ($C_{Si}$) is very low, around 4 fF. The conductance increases up to 0.38mS which is a very small value thanks to the good performance of dielectric properties of the silicon, since standard silicon is used to process TSV.

**Figure 30:** Parasitic a) resistance and b) inductance of one TSV (for 3 measurement sites)

Process dispersion

We also study the process dispersion through 180 measurement sites on the whole wafer, and we give below extracted TSV resistance versus frequency. $R$ varies between 100 and 300 mΩ for min/max value and the dispersion at 10GHz is comprised between 140 and 220mΩ, including wafer edges where dispersion is higher than wafer centre.

**Figure 31:** Parasitic capacitance and conductance of one TSV (for 3 measurement sites)
To find a good agreement between characterization and simulation, retro-simulations with process values adjustment are necessary. Initial values are provided by morphological analysis, electric material properties measurements as given in the picture below.

A good fit is obtained regarding the $S_{ij}$ parameters as showed below for the reflection and the transmission coefficients as well as the Smith representation. Results given here are for the entire dual via chain device, with node-embedding operation except pads de-embedding.
By fine tuning of several process parameters and materials such as silicon resistivity, silicon thickness, seed layer thickness and conductivity...etc., the best fit is obtained, as seen on the figure below, for $\sigma_{Si}=25$ S/m and $t_{Si}=55$ µm. Due to the very weak value of R, it is difficult to get perfect fit between simulation and measurements. As a conclusion for retro-simulation, we can note that:

- a perfect fit between simulation and measurements is difficult to reach (especially for R), due to very low values on the whole frequency range
- we demonstrated that results have a high sensitivity to metallic properties ($\sigma_{Cu}$) and to $t_{Si}$
- from the theoretical values, we have to adjust our values as the following: $\sigma_{Si}$ 25 S/m instead of 12 S/m and $t_{Si}$: 55 µm instead of 60 µm

Figure 34: Retro-simulations with initial process values of the dual via chain device
RF TSV Modelling

The following graphs compare measurements and model for each parameter of the π-shape model. Resistance can be modelled with the following equation (for $F > F_0$): 

$$R_\delta = R_{\text{Barrier}} \frac{t_{\text{sl}}}{\sigma_{\text{Cu}} \pi (2r_{\text{TSV}} + \delta - \delta^2)}$$

where $R_{\text{Barrier}}$ describes the resistance of the Ta/TaN layer in the TSV side walls. Skin effect begins at very low frequency ($F_0 = 93.8$ MHz) since the TSV are fully filled with copper.

The inductance computed by model is 25.7 pH, as shown on the graph below with the validated following equation:

$$L = \frac{\mu_0}{2\pi} t_{\text{sl}} \ln \left( \frac{t_{\text{sl}} + \sqrt{t_{\text{sl}}^2 + r_{\text{TSV}}^2}}{r_{\text{TSV}}} \right) + \tau_{TSV} - \sqrt{\frac{t_{\text{sl}}^2 + r_{\text{TSV}}^2 + \frac{r_{\text{TSV}}}{4}}{}}.$$
The oxide capacitance, computed from:

\[ C_{\text{ox}} = \frac{2\pi \varepsilon_0 \varepsilon_{\text{ox}} t_{\text{Si}}}{\ln \left( \frac{r_{\text{TSV}}}{r_{\text{TSV}} - t_{\text{ox}}} \right)} \]

gives 370 fF with a thickness of 330 nm and a permittivity of 5 for the oxide layer. The silicon capacitance is given by:

\[ C_{\text{Si}} = \frac{1.5 \cdot \sigma_{\text{Si}}^{0.29} \cdot \varepsilon_0 \cdot t_{\text{Si}}}{\ln \left( 1 + \frac{8 \cdot t_{\text{Si}}}{2 \cdot r_{\text{TSV}}} \right)} \]

which gives 5.79 fF at very high frequencies. And by re-building the capacitance behaviour using:

\[ C = \frac{C_{\text{ox}} \left( G_{\text{Si}}^2 + C_{\text{Si}} \omega^2 (C_{\text{Si}} + C_{\text{ox}}) \right)}{G_{\text{Si}}^2 + \omega^2 (C_{\text{Si}} + C_{\text{ox}})^2} \]

we obtained the model given in the graph below in blue dashed line. In the first graph (a), C is computed with the process values described in Figure 33, while the second graph is obtained with the optimised process values given in the conclusion of the retro-simulation part.

**Figure 36:** TSV Resistance and inductance: measurements and model (in blue dashed line)

**Figure 37:** TSV capacitance a) with initial process value b) with optimised process parameters values.
Finally the conductance is rebuilt with $G_{Si}$, calculated as following:

$$G_{Si} = \frac{C_{Si} \sigma_{Si}}{\varepsilon_0 \varepsilon_{Si}}$$

and using:

$$G = \frac{\sigma_{Si} \varepsilon_{Si}^2 \omega^2}{\sigma_{Si}^2 + \omega^2 (C_{Si} + C_{ox})^2}$$

As shown in the graph below is very difficult to fit the G value on the whole frequency range since G is extremely weak, with less than 0.4 mS even for F=20 GHz. A better fit could be obtained by tuning the silicon conductivity, but the others parameters will be changed in a way that the gap between simulations and measurements would increase.

![Graph showing TSV conductance](image)

**Figure 38: TSV conductance**

**TSV RF modelling Conclusions**

We performed measurements of dual via chain to extract parasitic parameters of one TSV with aspect ratio of 4 from 0.1 up to 20 GHz, and we showed that performance of such TSV are very good with extremely low resistance value (<100 mΩ) and an inductance better than 35 pH on the whole frequency range. The oxide capacitance is equal to 300 fF with a thin oxide layer of 330 nm and this value could be improved by using thicker oxide. Full ElectroMagnetic 3D simulations are in agreement with the measurements, showing that our tool is able to depict process in an extreme precise way. By modelling the TSV behaviour in RF, we can depict and predict its performance with parametric equations depending on the process parameters and the TSV dimensions. Retro-simulations and modelling step showed that silicon thickness is likely 55 µm instead of 60 µm, and that the silicon conductivity should approach 25 S/m to match even better with characterization.
2.6. Low-frequency Capacitance Measurements

This section discusses measurements and test structures created to characterise self and coupling capacitance of TSVs at low frequencies.

**TSV capacitance experimental electrical characterisation**

Within the ELITE project mask set, some suitable test structures have been designed for the evaluation of TSV self and mutual capacitance. To simplify the necessary process steps and mask requirements, a simple but accurate test structure can be created by interconnecting a chain of TSVs with a metal line to a testing pad. In this way, only the front end mask set is needed and the measurements can be performed before grinding; in this way the capacitance at the bottom of the TSVs have to be taken into account when compared to the model.

Several test structures are devised to measure coupling and self capacitance of the TSVs and the de-embed the effects of the metal lines, pads and connections:

- Three interdigitated TSV combs (illustrated in Figure 39(a)) with 30 and 40 TSVs per arm (see Figure placed at different spacing (shown in Figure 39(a)) for measuring coupling capacitance.
- An additional single comb of 60 TSVs was also created for self capacitance estimation.
- A stand-alone pad of size = 86 µm×86 µm is used to de-embed the non-TSV test structure.
- Some test structures are also available in the CAREFEA0 module (see Figure 40 and Figure 41) which contains one or more TSVs differently placed and connected with different metal routing that can be used to crosscheck other measurements.

![Interdigitated Comb Structure](image)

Figure 39 (a) an illustration of the interdigitated comb structure and (b) the actual interdigitated combs of 30 and 40 TSVs connected on the left-pad and right-pad respectively. The considered metal length (Lmetal) connecting the TSV and used for de-embedding purpose is the one on oxide field, i.e. the TSV diameter has not been considered. Different pitches are shown.

The test structures have been designed under numerous constraints with the following features in mind:

- Structures are designed to minimise the additional parasitic capacitance induced by the metal lines (the metal line width is 0.4 µm).
- The amount of TSVs and the layout has been chosen to obtain a suitable signal to provide accurate measurements with the LCR meter.
• Exploration of the coupling effects (for the interdigitated combs) are enabled by spacing TSVs from the minimum design rules to a relatively long distance. The TSVs are spaced at 40 µm, 60 µm and 100 µm.

The interdigitated test structures can also be used for self capacitance evaluation by simultaneously biasing the two combs with the same signal and by accounting for the parasitic metal line capacitance. The total comb length varies from about 1500 µm to 12000 µm. As the mask set was designed to minimize the front end process steps, it did not include a p+ ring implant for the substrate contact, so the contact reference point has to be taken from the substrate itself.

We anticipate that no significant coupling capacitance has been detected even with the TSVs at the minimum pitch since the substrate under test (p substrate, CK 10 – 13 Ω-cm -> ~10^15 atoms/cm^3) is sufficiently conductive to screen the coupling effect. Each comb structure can be considered isolated and therefore in the following sections we focus on self-capacitance measurements. The isolated pad is used to take into account the unwanted parasitic capacitance contribution of the measurement point.

The additional test structures available in the CAREFERA0 module are used to cross-check results and are not designed to minimize the interconnect metal capacitances (instead they are created to reduce the resistance).

![Figure 40 The CAREFERA0 CAD layout](image)

![Figure 41 The CAREFERA0 module consist of different TSV numbers (1, 7, 15, 26) placed in various ways in the module and connected with a large metal (10 µm) and a large metal landing (25 µm diameter) above the TSV. Moreover a large (50 µm×107 µm) metal extends from the pad as detailed in the layout.](image)
The schematic cross sections along different portions of the test structure are shown below to illustrate conductive and dielectric regions of the measured structures.

Figure 42 (a) 3-D drawing and (b) cross-section of stack, pad, metal and TSV.

Figure 43 (a) cross-sections of the (a) pad and metal layers and (b) TSV and metal only.

Figure 44 cross-sections of test structures for the (a) metal over TSV in a comb structure and (b) pad and metal for de-embedding.
The CV measurements have been performed with an LCR meter on the test structures (shown in Figure 39, Figure 40 and Figure 41). The measurements were conducted under different DC bias, AC small signal value, frequency and light conditions in order to investigate the possible presence of oxide charge and measurement quality/reproducibility and die-to-die variation. CV measurements are needed due to the metal/oxide/semiconductor structure intrinsically present in the TSV.

In the following a summary, the most significant results are reported for the 60 TSV comb unless otherwise specified and all measurements have been plotted considering the DC bias applied to the pad (independently from how they have been measured, i.e. DC bias applied to the pad or to the substrate to improve the quality of the measured signal). The measurement conditions are indicated in the figure captions. Due to the absence of a p+ implant in the design, the quality of substrate contact from the chuck is sometimes poor, however this has a negligible impact on the behaviour of the trends of the measured characteristics.

Die to Die variation @ 100kHz - 50mV

![Graph showing die to die variation at 100kHz with 50mV applied.]

Figure 45 Die to die variation. (Note: the measurement reproducibility is very good: repeating the measure on the same die the characteristic is perfectly superimposed).
Figure 46 AC signal frequency impact: some influence detected. At the plateau the major differences are related to the fact that at very low or very high frequencies (around the instrumentation limit) the phase departs from the ideal -90 degrees. The kink shift is possibly attributed positive charge trapped into oxide.

Figure 47 AC signal level shows no impact.
Figure 48 Light / frequency dependence: the microscope light on induced generation and recombination forces a reduction of the capacitance kink suggesting that charge traps and / or interface states are present in this measured wafer.

Figure 49 Different TSV structures provide the same results.
We can summarise the experimental CV measurement results as follows:

- the overall measurement reproducibility provides a certainty of the results;
- the die-to-die variation is not negligible especially considering the test wafers are only for trials;
- in general the frequency and AC level dependencies are related to the well-known limitation of the LCR experimental setup (calibration, quality of the measured signal vs. noise, etc.);
- the TSV CV minimum peak at about -25 to -20V can be attributed (considering also the light and frequency dependence) to a positive oxide charge present in the oxide layers surrounding the TSV since this peak it is not present in the pad-only measurement. This interpretation is supported by the literature; in fact charge effects in the TSV have already been reported in \[\text{124}\] and \[\text{14}\].

In order to allow model calibration / crosscheck we concentrated the extraction methodology on the single TSV capacitance at 0V DC bias, where the bias dependence is negligible.

By using all the different test structures and the following model: \( C_{\text{meas}} = (N_{\text{TSV}} \times C_{\text{TSV}}) + (L_{\text{metal}} \times C_{\text{line}}) \) it has been possible to extract the single TSV self capacitance by a linear regression fit while de-embedding the parasitic capacitance of the pad and the metal lines of the test structures.
From the linear fit in the two cases we obtained:

<table>
<thead>
<tr>
<th>N_{TSV}</th>
<th>C_{TSV} (pF)</th>
<th>C_{line} (fF/µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.3943</td>
<td>0.1059</td>
</tr>
<tr>
<td>40</td>
<td>0.3936</td>
<td>0.1098</td>
</tr>
</tbody>
</table>

Thus from our measurements we can consider $C_{TSV} = 394$ fF (for a single TSV). As a cross check, applying the previous coefficients to the single comb of 60 TSV we obtain a total capacitance (TSV+lines) of 24.92 pF that compares very well with the measured value of 24.76 pF.
As a further cross check by comparing different structures on the CAREFERA0 module with varying number of TSVs (1, 7, 15, 26), a self-capacitance value for a single TSV of 375 fF is extracted. The difference with the previous results (394 fF) is attributed to the large metal contribution in these test structures which prevents a clear distinction between the two capacitance contributions.

The measured value (394 fF) compares well with a calculated one of about 380 fF using the cylindrical plus plate (for bottom) capacitance formula and technological data. Considering the uncertainty in dielectric constants and non-uniformity of the dielectric layer thicknesses, this result is of interest. The bottom capacitance is not negligible and it is in the 20 fF range.

Low-frequency Measurement Summary
We have presented and discussed the test structures designed to measure the low-frequency capacitance of TSVs used in the ELITE project. Measurements for self- and coupling capacitances were discussed for TSVs in comb and chain structures were discussed and distinguished from the metal lines and pads using isolated de-embedding structures. Trends in the via capacitance for chains of 30 and 40 TSVs were also plotted and fitted using a linear equation. The study can be summarised as follows.

- Test structures to measure self- and coupling capacitances were designed to minimise the parasitic nature of the metal and pad structures.
- The good reproducibility of the CV measurements allows sensible assessment of the die-to-die variation. Experimentally, the CV die-to-die variation is not negligible especially considering that the test wafers were only designed for trials.
- Results across different TSV structures are consistent.
- The coupling capacitance between TSVs in comb structures was shown to be negligible for the substrate used in the test wafers (10-13 Ohm-cm).
- The self-capacitance has been found to track closely with the cylindrical plus plate capacitance formula.

2.7. Physical Modelling Conclusions
Closed-form compact models have been created for TSVs in low, medium and high resistivity substrates for TSVs in isolation, in a row and a bundle. The models cover a large geometrical range with radii, length, barrier thickness, pitch and frequency between 1-50 GHz as parameters. The models are on average more than 90% accurate to field solver simulations and at worst 80% accurate for the extended barrier thickness range in medium substrates, motivating their use for design space exploration with orders of magnitude greater computational efficiency. The models are packaged as a stand-alone tool and available on the web to the public.

S-parameter simulation and modelling of TSVs at high frequencies has also been undertaken with comparison to measured results. The equivalent RLGC circuit was developed from pre- and post-measurement simulations to enable the highest accuracy. Additionally, the low-frequency capacitance was measured with test structures including coupling and self capacitance for numerous configurations with an LCR meter. Measurements showed good reproducibility and coupling was negligible on the test wafers. The self-capacitance of the vias closely tracks the cylindrical plus parallel plate capacitance model.

Modelling and Measurement Correlation
The measurement methodology was conducted with the verification of the physical models in mind. Although the measurement structures consist of complex test-specific geometries not considered in the physical models (access pads, lines and inter-via connections), there is a close correlation to the compact models presented in section 2.4 when their effects are de-embedded. For the ELITE TSV dimensions (15 µm diameter, 60 µm length and 330 nm dielectric barrier), we find that the capacitance estimation provided by the compact models is close to the low-frequency isolated capacitance measurement and when the frequency models are used, it is accurate to the RF measurements at 20 GHz.
In section 2.5, the parasitics are extracted from s-parameter simulations and converted to an equivalent $RLGC \ \pi$-model. The extracted circuit parameters have a close match to the physical model parameters extracted from Ansoft Q3D using a completely different methodology. The convergence in the simulation results using different techniques provides a firm foundation for the compact and RF model development.

The measured self-inductance for a single TSV varies between 32 pH and 35 pH across frequencies from 0-20 GHz (see Figure 36), where the model for low-frequency self-inductance presented in equation (2) estimates 25.3 pH for the same geometries. Considering the difficulty of de-embedding the access lines and surrounding test structures, this is a very close match. The measured DC resistance value is 15 mΩ (see Figure 30) for a single TSV and the calculation from equation (1) yields 6.1 mΩ. This discrepancy is attributed in the first instance to non-idealities associated with de-embedding.

Furthermore, the DC capacitance is measured from the RF test structures and found to be roughly 300 fF, where the models estimate the capacitance at 1 GHz to be 306.6 fF. The high-frequency capacitance converges to roughly 4 fF in the measurements and when equation (10) is applied to the isolated capacitance model in (4) the estimated high-frequency capacitance is 8.9 fF. The low-frequency capacitance measurements conducted on a chain of vias for a medium resistivity substrate in section 2.6 additionally found that the coaxial cable capacitance equation has a close match to the measured and de-embedded capacitance of a single TSV. This has a close correlation to the isolated capacitance equation we have presented for a low-resistive substrate (which may be attributed to the location of the LCR meter reference on the bottom of the substrate).

The close collaboration between the measured $RLC$ parasitics discussed in sections 2.5 to 2.6 and the parameters calculated from the closed-form compact models presented in section 2.4 motivate their use for accurate design space exploration and early-performance estimations of TSV-integrated 3-D die stacks.
3. Signalling Conventions for 3-D ICs

We show using predictive technology models for bulk CMOS transistors that the 3-D interconnect provides over an order of magnitude less energy-per bit when compared to off chip I/O drivers. Optimal techniques for signalling over Through Silicon Vias in 3-D die stacks are then investigated for Low-voltage differential (LVDS) and low-voltage single-ended (LVSE), voltage mode (VM) and current mode (CM) drivers and receivers implemented in a 65 nm CMOS technology and SPICE simulations using electrical models of coupled TSVs.

3.1. Introduction

The key performance benefit in 3-D ICs is provided by the reduced interconnect length, and the electrical properties of the TSVs, which are much faster than on-chip wires. In order to realize this potential, it is essential to maximise the data-rate over TSVs, especially as they are a scarce resource and may have to support a higher data-rate than longer planar wires which are slower but more numerous. Therefore, it is imperative that signalling conventions for TSVs are tailored to their electrical characteristics, taking into account the low resistance and high coupling to other TSVs, which increase cross-talk-on-delay effects and adversely impact signal integrity (SI) as we show in a previous report [33] and in ELITE D2.2 [34].

The analysis of signalling conventions for TSVs, which to the authors’ best knowledge has not been addressed at all in the literature to date, is the topic of the ELITE deliverable D2.4 in WP2. In this task, we examine voltage-mode (VM), current-mode (CM) and low-voltage single-ended and differential signalling techniques for TSVs in order to maximise data rate and preserve signal integrity, and present results for a 65 nm technology based on Spice simulations. The TSV parasitics are extracted from a field solver, while device models are based on predictive technology parameters [35]. We show how current-mode signalling supports a higher data-rate and reduces jitter, at the cost of increased energy consumption. Low-voltage differential signalling (LVDS) increases the signal-to-noise ratio and demonstrates limited delay variation, at the cost of high static power consumption. Low-voltage single-ended signalling (LVSE) reduces the power and decreases the effective coupled noise amplitude, but at the cost of decreased signalling speed. We quantify the trade-off between energy and latency for these modes of signalling over TSV interconnects. The main contribution of this study is in providing analysis results and design guidelines for signalling over TSVs, necessary for chip designers to better understand signal characteristics in 3-D ICs in the design planning phase.

This chapter is organised as follows. In section 3.2 we discuss existing work on TSV modelling, as well as relevant work in the extensive body of literature on signalling over on-chip interconnect. In section 3.3 we propose an electrical model for delay and crosstalk estimation. Next the performance, limitations and opportunities in voltage-mode signalling including the judicious use of shielding is investigated. Then we discuss CM signalling, LVDS and LVSE signalling respectively in a TSV bus structure to show their potential for higher performance and improved SI over voltage-mode signalling; design guidelines are also proposed. We then investigate the effect of TSV parasitics on the I/O driver of a flash device using IBIS models. Finally we present our conclusions.

3.2. Related Work

The majority of work published so far on TSVs in 3-D integration has focused on technological aspects and electrical characterization of isolated structures. An initial work that looks at SI effects in 3-D ICs is [36], which discusses crosstalk between interconnects on different layers within a 3-D IC for different substrate types. The authors of [37] propose a TSV model based on measurement results and discuss signal transfer characteristics. Pak et al. in [38] discuss electrical characteristics of two TSVs for various structural and material parameters based on full-wave simulations. They investigate the effect of TSV parasitics on data rate with increasing number of dies in the stack as well as material and structural parameter variations. They also discuss the variation of crosstalk amplitude with different shielding arrangements within a densely packed arrangement of TSVs. In [39] analytical models for TSV resistance and capacitance are proposed with an assumption of a co-axial transmission line model for a TSV; further inter-layer I/O circuits to cope with different parasitic loads in the TSV itself and micro-bumps and achieve optimum power and signal drive are discussed. The authors of [40] empirically derive a delay model for a TSV in a bundle as a function of its geometrical and material parameters, but do not propose an electrical model that can be used in a circuit simulation to explore SI or delay effects with drivers. Savidis et al. in [41] discuss trends
in TSV parasitic parameters extensively and show that in a TSV the L/R time constant is higher than its RC time constant. We have previously derived in a complete set of analytic equations to predict the capacitive and inductive parasitic parameters in a $3 \times 3$ matrix of TSVs that can be used for SI and delay related analysis early in the design flow. These models are detailed in a previous deliverable [34] and a recent article [4]. There we have shown that the output impedance of a CMOS inverter completely masks the rather small combined resistive and inductive components of the TSV impedance ($R_{TSV} + j\omega L_{TSV}$), and the capacitive parasitics dominate, with capacitive crosstalk coupling a noise pulse with an amplitude over $V_{dd}/2$ on a quiet victim and having a significant effect on delay on a switching victim when all the neighbouring aggressors in a $3 \times 3$ matrix switch simultaneously.

Performance (SI and delay) related modelling and analysis techniques for on-chip interconnects as well as signalling conventions have been extensively researched by the IC design community, with a significant body of work having been published in the literature proposing solutions for mitigating crosstalk noise and improving data rate over parallel wire structures. Shielding is a simple and effective way of reducing capacitive coupling between adjacent conductors thereby reducing the effect of cross-talk on delay and jitter [40]. Repeater insertion is a ubiquitous solution in reducing delay over highly resistive interconnects [41]; wire-sizing is also used to control wire resistance. A methodology as well as analytic guidelines for simultaneous co-optimisation of repeater insertion including number and size, shielding and wire-sizing for maximising throughput over an area-limited interconnect resource was proposed in [42]. Repeater insertion is power-intensive, and a driver that dynamically adjusts its drive strength to compensate for switching-pattern dependent crosstalk to reduce jitter and power consumption while maintaining a similar worst-case latency as with inverters was proposed in [43]. This work also provides a comprehensive overview of other alternatives to inverters and repeater insertion, such as boosters [44], skewed inverters [45] [46] and transition-aware receivers [47] that have been proposed in the literature. CM signalling [48] has also been proposed as an alternative to conventional VM signalling for on-chip wires, which can be up to three times faster but with increased static power consumption [49]. Apart from the CM signalling, LVDS has drawn the attention of designers as an efficient signalling alternative compared to conventional rail-to-rail signalling [50]. In this study we implement a simple CM receiver proposed in [49] to show how CM signalling together with the LVDS transceiver proposed in [50] is suitable for signal propagation over TSVs and reduce latency and jitter, as well as increase noise resilience when compared to VM signalling. Low-voltage signalling circuits are widely used for on- and off-chip wires to increase signalling speed and noise immunity. A number of low-voltage signalling circuits are proposed and studied in [51] and the trade-offs for on-chip interconnect discussed. The LVDS and LVSE circuits detailed in [50] and [49] respectively are implemented in this study and simulated with TSV parasitics to determine their performance in 3-D ICs.

### 3.3. TSV Parasitics and Electrical Model

The dimensions vary depending on the process and application; we consider pitches of 15 µm and 25 µm in our study, well within the current capabilities of many processes [19] [52] and relevant for the ELITE project. The cross-section is assumed to be uniformly circular, with a radius of 10 µm and length 50 µm. The substrate conductivity ($\sigma$) varies depending on the application, with values being typically around 10 S/m for RF processes, so that the relatively high resistance provides better isolation between components. However digital applications use a much more heavily-doped substrate, to ensure equi-potentiality throughout the bulk substrate as much as possible, reducing the possibilities of latch-up [24]. In this work we have considered three different substrate conductivities 100 kS/m, 10 S/m and zero, representing typical values used in digital and RF processes and a hypothetical highly resistive substrate respectively. The topology considered is three parallel coupled TSVs, a representative unit in a row of TSVs. A field solver has been used to extract the electrical parameters (resistance $R$, capacitance $C$ and inductance $L$) in the context of the equivalent circuit shown in Figure 53.

---

1 The CZ substrate to be used by NMX in the ELITE prototype for example has a substrate conductivity of 10 S/m.
The extracted parasitics for the structure of Figure 53 are given in Table 10 for the considered pitch values of 15 μm and 25 μm. It can be seen that when the substrate conductivity is very high (i.e. a heavily-doped substrate) the total capacitance is the highest, with the majority of the capacitance being to ground, as can be expected since the substrate voltage is very nearly at 0 V. For a highly resistive substrate, the capacitance is the lowest, but the coupling component dominates, being 3 to 6 times greater than the self component. For a lightly-doped substrate, the total capacitance is in between these values, but the ratio of coupling to self capacitance is even higher, from 60 to 90.

The TSV resistance and inductance values including the mutual component are not affected at all by substrate conductivity and can be estimated from the compact models proposed by us in [33] [4], while the capacitive models proposed there can be used for highly resistive substrates. Further, through field-solver simulations it was established that TSV resistance does not change significantly due to the skin effect up to a signal frequency of 50 GHz.

### Table 10 TSV capacitances for different inter-via spacings and substrate conductivities (Rtsv=2.78 mΩ)

<table>
<thead>
<tr>
<th>Substrate Conductivity, σ (S/m)</th>
<th>TSV Pitch/(μm)</th>
<th>15</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>C=3.01 fF,Cc=17.79 fF</td>
<td>C=2.85 fF,Cc=8.90 fF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L=17.13 pH,Lm=23.27 pH</td>
<td>L=17.13 pH,Lm=11.90 pH</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>C=1.89 fF,Cc=114.21 fF</td>
<td>C=3.37 fF,Cc=88.57 fF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L=17.13 pH,Lm=23.27 pH</td>
<td>L=17.13 pH,Lm=11.90 pH</td>
</tr>
<tr>
<td>100k</td>
<td></td>
<td>C=564.81 fF,Cc=0.04 fF</td>
<td>C=564.88 fF,Cc=0.01 fF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L=17.13 pH,Lm=23.27 pH</td>
<td>L=17.13 pH,Lm=11.90 pH</td>
</tr>
</tbody>
</table>

### Isolated TSV

In order to determine the appropriate electrical model for a TSV, we carried out simulations to determine the relative contribution of the resistive (R), capacitive (C) and inductive (L) parasitic components of the TSV to the shape of the output waveform in typical delay and SI calculations as described in detail in [33] and [34]. The driver and receiver are CMOS inverters implemented in a representative 65 nm technology through the use of predictive parameters [35]. In the simulations the TSV model was driven by an inverter sized to be 50× a minimum-sized inverter, and loaded by a minimum-sized inverter. In a minimum-sized inverter the NMOS and PMOS transistors have channel lengths of 65 nm, and widths of 1.5× and 2.3×1.5× the channel length respectively, with 2.3 being the n-type to p-type mobility ratio. The supply voltage for this technology is 1.1 V.

The output waveform was simulated while sweeping the entire range of $R$, $L$ and $C$ values as determined by the field solver including any potential increments in resistance due to the skin effect. For rise times down to 1 ps, no discernible change could be observed when varying the inductance. Varying the capacitance revealed a significant spread in the 50% latency of the output waveform. The resistance is swamped by the output impedance of the driver, while the self inductance is significant only when the TSV resistance is comparable to the driver resistance, as is usually the case for on-chip wires. As driver resistances are typically in the kΩ range and the TSV resistance in the

Figure 53 Electrical Model of Coupled TSVs
mΩ range, \( R_{\text{driver}} \gg R_{\text{tsv}} \) and the inductive effect is negligible. Further with technology scaling driver resistance scales at a higher pace than the TSV resistance. These results appear to show that the electrical model for an isolated cylindrical TSV can be reduced to a purely capacitive model. The resistive and inductive parasitics are small enough to be neglected in any delay simulation, which reduces the complexity of the electrical model significantly.

**Coupled TSVs**

As capacitive and inductive coupling can have detrimental effects on bandwidth and signal integrity, the crosstalk between adjacent structures must be examined to determine the most efficient use of area and TSV sizing to maximize signal throughput and reliability. The three-parallel TSV structure of Figure 53 was examined to determine parasitic coupling effects. As with on-chip interconnects, the capacitive coupling terms to nearest neighbours dominate over the coupling terms to nonadjacent lines, which are mostly insignificant. In the case of inductance, the coupling terms are relatively significant within the entire structure in so far as no one term dominates another, because magnetic field lines tend to permeate the length and breadth of the global structure, again analogous to the on-chip case. Simulations reveal that over the entire considered range of TSV dimensions and substrate conductivities, the effect of inductive coupling on the victim net does not appear to be large enough to justify the modelling of parasitic mutual inductance in a three parallel TSV system for delay and SI calculations. Inductance may of course be significant when considering the cumulative effect in a large network of TSVs as inductive coupling is long range, but as long as there is judicious interspersing of signal and power/ground vias, the analysis we have carried out here can be extended to much larger networks. Capacitive coupling however needs to be considered at the outset. Therefore the simulation model for coupled TSVs reduces to a network of coupled capacitors representing self- and mutual-capacitances.

**Voltage-Mode (VM) Signalling**

A simulation setup as shown in Figure 54 was used to analyse the SI effects in VM signalling over TSVs. All TSVs (victim and aggressors) are driven by a 50× inverter \((V_1)\) and loaded by a minimum-sized inverter \((V_2)\). As the TSV resembles a lumped capacitive load, a cascaded driver with optimal stages needs to be used. To analyse the effect of driver size on the energy and delay characteristics of voltage mode signalling for TSV parasitics, a parametric analysis was conducted and the results plotted in Figure 55. For driver sizes of under 20× the minimum-sized inverter, the total capacitance of the coupled TSV presents too large a load for a single small inverter to drive effectively. This results in large rise and fall times during the transitions to charge the TSV capacitance. In Figure 55(a), this effect can be seen by an initial sharp knee in the energy curve, which falls outside the typical trend of energy growth with driver size for an inverter. This behaviour is a result of the increased short-circuits current caused by the long rise and fall times of the inverter in the transition period. As the driver size increases, the short-circuit current also increases due to a subsequent drop in resistance through the PMOS and NMOS transistors, which is exacerbated by the long transition period. Until the size of the driver is adequate to drive the TSV load (above \( H_D = 20\times \)), the short-circuit current dominates the energy consumption. Above this point, the dynamic switching behaviour contributes the most to the energy performance of the voltage mode circuit. Table 11 shows the coupled noise voltage amplitude on a quiet victim at the far end of the TSV when both aggressors are switching and the worst-case delay when the two adjacent TSVs switch simultaneously in the opposite direction to the victim. The table is populated with the noise amplitude and delay values corresponding to different \( C_s \) and \( C_c \) values to demonstrate the effect of worst-case coupled voltage and noise-on-delay effect for different substrate characteristics and TSV geometries.
Table 11 Worst-case coupled noise amplitude and delay for various $C_S$ and $C_C$ combinations corresponding to different pitches and substrate types

<table>
<thead>
<tr>
<th>$C_S$(fF)</th>
<th>$C_C$(fF)</th>
<th>Peak noise/$V_{dd}$</th>
<th>Worst-case delay/(ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>0.594</td>
<td>82</td>
</tr>
<tr>
<td>3</td>
<td>89</td>
<td>0.557</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>0.271</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>0.177</td>
<td>20</td>
</tr>
<tr>
<td>50</td>
<td>20</td>
<td>0.105</td>
<td>12</td>
</tr>
</tbody>
</table>

The victim and aggressor drivers were driven by pseudo-random bit sequences with rise/fall time and period of 10 ps and 200 ps respectively, and the resultant eye-diagram plotted in Figure 56(a). In this particular case the delay varies between 8 ps and 55 ps. As the TSV resistance is several orders of magnitude less than that of the driver, the TSV capacitance dominated by the coupling component dictates the delay. Shielding is widely used in bus structures to alleviate crosstalk and consequent coupled noise and delay variation in interconnects. Shown in Figure 56(b) is the eye diagram at the output when both neighbouring TSVs act as shields; the best-case delay and worst-case delay are 28 ps and 29 ps respectively. The cost of shielding is reduced utilisation of TSVs, a scarce resource in 3-D ICs.
The voltage mode unshielded scheme demonstrates the greatest coupled noise amplitude on a silent victim out of all the signalling circuits investigated. This crosstalk can be significant in noise sensitive systems as the voltage transferred between TSVs can be up to 55% of $V_{dd}$. Shielding eliminates the capacitive coupling entirely, but at the cost of signal TSVs. However, the impact of the loss of TSVs could be mitigated if these TSVs were used as redundant signal lines in case of failure or as thermal vias.

**Current-Mode (CM) Signalling**

The theoretical analysis in [49] has shown that CM signalling can be up to three times faster than VM signalling, but with an increase in static power consumption. The fundamental difference in CM and VM signalling is the termination resistance at the receiver end; low for CM signalling and several orders of magnitude larger than the driver resistance in VM signalling [48]. A circuit schematic of a basic CM signalling link is shown in Figure 57 [51]. The driver is a typical inverter, but the receiver is composed of NMOS (MN2) and PMOS (MP2) transistors with drain and gate shorted so that they act as active load resistors connecting net “TSV_out” to ground and $V_{dd}$ respectively. The inverter formed by MN3 and MP3, INV3, restores rail-to-rail output.

![Figure 56 Eye diagram at far-end of TSV for a pseudo-random bit sequence when $C_S = 3f$ and $C_C = 89fF$.](image-url)
The driver resistance and termination resistors act as a voltage divider and the output voltage at the far end of the TSV charges/discharges between two levels as the input switches between logic 1 and 0; the steady-state output voltage levels are $\frac{\alpha}{1+2\alpha}$ and $1+\frac{\alpha}{1+2\alpha}$, where $\alpha$ is the ratio between driver resistance and termination resistance. As the termination resistance approaches infinity, i.e. the receiver becomes a simple inverter, the swing becomes rail-to-rail and the system reverts to VM signalling. The speed improvement of CM signalling is due to the reduced swing, which results in less time being taken to charge and discharge the interconnect capacitance.

The design criteria to achieve functionality of this CM receiver is that sizing of the termination transistors and INV3 should be carried out so that the reduced voltage swing is sufficient to switch INV3, providing rail-to-rail signal swing on net “out_CM”. The driver, INV1, is an $H_d \times$ inverter while MN2 and MP2 are sized to be $H_r \times$ and 2.3 $H_r \times$ a minimum sized device respectively. Figure 58 shows the input and output voltage waveforms of the CM link depicted in Figure 57 with $H_d = 50$. Shown in the same plot is the output of a comparable VM link. In order to be able to compare between different signalling systems, the VM signal link comprises the same circuitry as the CM link except for MN2 and MP2; the output after INV3, is called “out_VM” in this instance and plotted in the figure. The simulation results reveal that the CM link is 50% faster than the VM link.
In order to see the effect of coupling in CM signalling, a CM link flanked on either side by identical CM signalling links has been simulated. Drivers are sized to be 50× while self and coupling capacitances of TSVs have been set to 3 fF and 89 fF respectively corresponding to a substrate resistivity of 10 S/m (matching the wafer specifications supplied by Numonyx [54]) and pitch of 25 μm (see Table 10). The termination resistors form a voltage divider with the driver output resistance, and effectively reduce the RC time constant of the system as compared to VM, reducing the effect of crosstalk on delay. Also in CM signalling the coupled noise voltage on a silent victim discharges from both the receiver and driver side whereas in VM it discharges only from the driver side. Therefore the termination resistance affects the coupled noise amplitude and pulse width. Increasing $H_r$ from 1 to 50 reduces the noise amplitude from 0.55 V to 0.47 V. As the coupled noise in the VM link is 0.61 V, a 20% reduction in crosstalk noise can be observed with CM signalling. Figure 59 shows the eye diagrams of the output after INV3 for CM and VM signalling links. As can be seen, CM signalling reduces delay variation or jitter by approximately 50%.

![Figure 58 Output voltage waveforms for the CM link shown in Fig. 4 and comparable VM link with $C_{tsv} = 100\text{fF}$ and $H_d = H_r = 50$](image-url)
The CM delay can be further reduced through reduction of the termination resistances (by increasing the size of MN2 and MP2), which reduces the swing of the signal at the input of INV3. However there is a limit defined by the point at which the low voltage swing is inadequate to drive INV3 to restore rail-to-rail swing; for $H_d = 50$, this limit is $H_r = 52$. Moreover, increasing $H_r$ causes the static power dissipation to increase, as it creates a low resistance current carrying path from the supply rail to ground. This trade-off has been quantified in Fig. 7, where the worst-case delay and average energy consumption per cycle for CM signalling has been normalised to the worst-case delay and energy consumption in VM signalling.

Figure 59 $H_d = H_r = 50$ Eye diagrams for (a) CM and (b) VM after INV3, with $H_d = 50$, $H_r = 50$, $C_s = 3f$ and $C_c = 89fF$. 
Low-voltage Differential Signalling (LVDS)

Low-voltage Differential Signalling has been used extensively as an interface between on-chip signals and off-chip circuits. By reducing the swing of the output voltage, the emitted noise can be significantly lowered. This can reduce noise-on-delay effects and coupled noise amplitude, allowing for higher signalling speeds. Additionally, differential signalling has a much higher signal-to-noise ratio, and can effectively filter out common mode noise from outside sources. However, LVDS requires two signalling lines to send a single data bit. This drawback is already a significant issue in on-chip interconnects due to the increased area overhead of routing additional signals; in a 3-D environment, the TSVs are an even scarcer interconnect resource. In reality the overhead is not exactly 2:1, as with on-chip wiring, return paths are required for a group of signal lines. Thus the overhead of the differential signalling is roughly 1.8 for planar wires [55]. Another drawback is the increased complexity of the driver and receiver circuits which increases area, transistor count, power consumption and additional design time. Low-voltage circuits frequently require additional low-voltage supply rails which in turn introduces routing and layout issues. Routing multiple voltage sources through several chip layers in a 3-D circuit can become complex. For these reasons, we simulate a low-voltage differential driver and receiver capable of producing and restoring low-voltage signals without additional supplies to maintain a fair comparison with other signalling modes. The simulation arrangement is shown in Figure 61.
The LVDS driver and receiver circuit [50] is shown in Figure 62. The circuit has been optimised for a row of coupled TSVs in a 65 nm CMOS process. Two TSVs are required for one signal, where each TSV has the same self and coupling capacitances as shown in Figure 61. The driver circuit produces a 300 mV differential voltage output centred on $V_{dd}/2$ from a single rail-to-rail input using controlled gains and inverters on the driver side. The receiver senses a voltage transition from the differential signal over the TSVs and subsequently restores full rail-to-rail voltage from the differential voltage signal using a minimum-sized inverter at the output. The operation of the circuit is further explained in [50]. In our analysis, noise related SI metrics are measured at the near- and far-end of the TSV, between the driver and receiver, while the final input to output 50% rise time is measured after inverter V2, shown in Figure 61, at the output.

The LVDS simulation setup for a row comprising two TSVs per signal for victim driver and receiver (V) and the surrounding aggressors (A, B). $C_s=3fF$, $C_c=89fF$.

The transistors are optimised to compare with the VM scheme’s delay and the inverters are sized to match the voltage and current mode signalling circuits. The PMOS to NMOS ratio and a minimum sized transistor are kept consistent with previous schemes. The reduced voltage swing and differential signalling lowers the effect of capacitive crosstalk and improves common mode noise rejection by providing the reference in the signal itself. However, this LVDS circuit requires 18 transistors in total to produce the low-voltage transitions, which is
significantly more than the other signalling schemes considered, which contributes to an increase in static power consumption.

Pseudo-random bit patterns were used to produce the eye diagrams shown in Figure 63. Input rise/fall times and pulse widths are consistent with CM and VM schemes. The results reveal that LVDS shows a much more open and predictable eye, which is indicative of low delay variation and reduced coupled noise amplitude. However, in terms of worst-case coupled rise time, the LVDS driver is much slower than VM and CM, taking 89 ps to reach 50% compared to 55 and 32 ps for the VM and CM schemes. This is largely due to the increased overhead from the circuitry requiring the translation of $V_{dd}$ to a low-voltage signal and sensing delay of the level-restoring circuit. The coupled noise amplitude on a silent victim is just 68 mV, which is significantly less than all other schemes.

Low-Voltage Single-Ended (LVSE) Signalling

As the additional wiring overhead of LVDS can be undesirable even for the planar on-chip interconnect, there are a number of alternative single-ended low-voltage signalling techniques. These circuits generally provide a reduction in power consumption and EMI but at the cost of increased delay. As with the LVDS case, to attain a low-voltage signal, driver designs frequently incorporate additional low-voltage supplies to simplify the circuitry. Unfortunately, for multi-layer 3-D devices this extra overhead in the form of power rail distribution is even less desirable with the scarcity of TSVs. Thus the signalling circuitry with the least routing and design complexity and the smallest area will be more attractive for future 3-D devices.
A low-voltage single-ended driver is described in detail in [51]. In this design, the low-voltage output signal is produced by the driver without any additional low-voltage supply rails, which provides a good comparison to the other TSV signalling schemes investigated in this study. The LVSE driver reduces the voltage swing from a rail-to-rail input by interchanging the PMOS and NMOS transistors in a typical inverter. The PMOS pulls low and the NMOS pulls high in Tx shown in Figure 64. This limits the swing and offsets the output voltage from V_{dd} and ground by a value, V_{tn}. In our representative 65 nm 1.1V CMOS technology, this equates to a voltage swing of just 340mV centred around V_{dd}/2. The receiver restores the full output voltage using a minimum sized inverter; where transistors P3 and N3 isolate the low-voltage signal from the rail voltage and feedback transistors P1 and N1 help bring the output of the inverter formed by P2 and N2 back to full rail-to-rail voltage.

Output waveforms of the circuit in operation are shown in Figure 65. The LVSE system was simulated with the same transistor parameters as the other schemes. The inverters are sized to H_D=10 and H_D=36 for INV1 and INV2 respectively, with the low-voltage transmission inverter, Tx, sized to H_TX=70. The pass transistors N3 and P3 were sized to 13\times, with N1 2\times and P1 2\times as minimum sized transistors. This setup produced the minimum delay for the circuit (investigated by a series of parametric sweeps in Cadence). The delay is measured from the final output, “OUT”, whilst the coupled noise amplitude is measured at the far-end of the TSV itself.

The simulation results for this LVSE demonstrated a much larger variation in delay and a severely reduced eye opening for rise/fall times of 10 ps and period of 200 ps. The delay from best- to worst-case coupled switching patterns varied from 62 ps to 132 ps. This large variation was unacceptable for the same pseudo-random input bit patterns as the other signalling schemes, thus the pulse width was extended to 150 ps from 100 ps. The resultant eye diagrams are shown in Figure 66. Even for a reduced input data rate, it is clear that this signalling scheme has the highest variation in delay, but does not suffer from any large overshoot effects. This demonstrates that this signalling scheme is more appropriate for lower data-rate applications for reduced energy consumption from the interconnect.

Figure 64 - LVSE driver and receiver circuit diagrams

Figure 65 LVSE output waveforms showing the input, voltage at the TSV and voltage at the output
The coupled noise amplitude resulting from the worst-case switching pattern on a silent victim was the second best at 162 mV. Using (22), the maximum data rate per TSV in LVSE signalling is significantly less than the other schemes, at just 7.2 Gbps per TSV. However, the power is significantly reduced, at roughly half the usage (0.62%) of the conventional voltage mode driver. Although the coupled noise amplitude and energy per bit of LVSE is small in comparison to other schemes, the delay is far greater. However, with crosstalk, and more importantly, thermal issues in 3-D integration, a low-voltage, low-power signalling schemes such as this may prove enticing in future systems where reliability, power and reduced heat are of greater concern than layer-to-layer signalling speed.

**Flash Device Input/Output Buffers (IBIS)**

In order to provide a comprehensive analysis of the signalling characteristics of the ELITE 3-D system demonstrator, realistic models of input and output drivers of the intended Flash device are simulated with ELITE TSV dimensions and parasitic properties. In this section, we explain the ELITE specific TSV simulated pitches, radii and lengths used in the field solver. In addition to project-specific TSV properties, a model of Input/Output drivers (IBIS) from Numonyx’s “One NAND” 70 nm product line are simulated in spice simulations. By introducing the I/O capabilities of the intended Flash device into the TSV simulations, the signalling characteristics of the ELITE 3-D demonstration prototype can by analysed.

The ELITE TSV structure is simulated in a field solver for a range of pitches, for two sizes to represent the expected process technology and dimensions to be used in the prototype. The extracted parasitics are summarised in Table 12.
The IBIS file provided by Numonyx is a black-box model of the I/O drivers and receivers of any external pin on the Flash memory device. The IBIS model uses an I/V lookup table, to simulate the drive strength and slew rate of the I/O capabilities of the memory device. In this case, the R, L, C parasitics of the pin and wire bond to the die are not included in the model, however the pad parasitics at the die are considered in the model. This de-embeds the parasitics of the package, allowing a direct analysis of a bare die-to-die interconnect using TSVs. The equivalent circuit of an IBIS model is shown in Figure 67.

Given the driver and receiver characteristics of the memory device, and the electrical models of the ELITE TSVs, simulation of signalling performance within the ELITE demonstrator is possible. The first simulation comprises a data transfer from the bottom flash die (in a stack of two) to the top die using the asynchronous data I/O bus model in the IBIS file. A row of three data bits is considered, where the central bit is the victim and the two adjacent signalling lines are aggressors. The simulation setup is shown in Figure 68 below. The top die receives a simulated input pulse at the asynchronous data input of an IBIS buffer in output mode. This IBIS buffer is then connect directly to the TSV, where the TSV transmits the signal from the top layer to the bottom die and to an identical IBIS buffer in receive mode. This effectively demonstrates the capability of one flash die to send information to another flash die in the stack.

Table 12 – ELITE specific TSV parasitics extracted from the field solver used in IBIS simulations

<table>
<thead>
<tr>
<th>TSV parasitics Case 1:</th>
<th>C_s_centre (fF)</th>
<th>C_c (fF)</th>
<th>C_s_side (fF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_v = 7.5 μm, L_v = 50 μm, P_v = 50 μm, d_b = 0.3 μm, σ = 10 Ω·cm</td>
<td>14.942</td>
<td>45.434</td>
<td>36.649</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TSV parasitics Case 2:</th>
<th>C_s_centre (fF)</th>
<th>C_c (fF)</th>
<th>C_s_side (fF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_v = 6 μm, L_v = 50 μm, P_v = 50 μm, d_b = 0.3 μm, σ = 10 Ω·cm</td>
<td>15.511</td>
<td>38.6</td>
<td>36.764</td>
</tr>
</tbody>
</table>

Figure 67 Equivalent circuit of an IBIS driver. Package and pin parasitics as well as drive strength and slew rate of the I/O drivers of a device are contained within this industry-standard IBIS specification.
The first simulation draws comparisons between the signalling schemes detailed earlier in this report and the realistic I/O implementations for a flash device. The switching pattern input in the row of TSVs generates worst, best and nominal switching patterns. An output waveform of the input-to-output delay of the IBIS driver is shown in Figure 69. The extracted 50% rise times for different switching patterns are detailed in Table 13 below. Initial results clearly point to a marginal effect of delay on the flash I/O data drivers. This is a straightforward observation as the maximum delay of the TSV is in the order of tens of picoseconds, compared to the nanosecond delay of the I/O driver.

![Diagram of the simulation setup for IBIS simulations of a Flash memory die stack.](image)

**Figure 68** Diagram of the simulation setup for IBIS simulations of a Flash memory die stack. The top layer receives a stimulus signal at the pad which drives the TSV using an IBIS asynchronous data buffer. The test signal is sent to the second die to an IBIS data buffer in receive mode to provide realistic loading.

![IBIS driver with simulation input of 100 ps rise. 50% delay is measured from the input of the IBIS model to the output of the TSV before the load representing the next flash die.](image)

**Figure 69** IBIS driver with simulation input of 100 ps rise. 50% delay is measured from the input of the IBIS model to the output of the TSV before the load representing the next flash die.

<table>
<thead>
<tr>
<th>Switching Pattern</th>
<th>Nominal Switch</th>
<th>Worst-Case Switch</th>
<th>Best-Case Switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay (50%)</td>
<td>1.777</td>
<td>1.784</td>
<td>1.768</td>
</tr>
<tr>
<td>Typical[ns]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst-Case</td>
<td>3.799</td>
<td>3.808</td>
<td>3.799</td>
</tr>
</tbody>
</table>

Table 13 - Switch-pattern based delay using IBIS drivers of a flash device
The maximum delay from the worst-case switching pattern to the best-case switching pattern is only 16 ps for IBIS drivers in a typical corner configuration. When the buffers are put into a minimal corner configuration, meaning the worst-case performance given temperature or voltage inputs, the TSV has a slightly larger impact on delay. This however is entirely masked out, when the total variation in delay from minimal to maximal corners of the parasitics and drivers are considered (0.8598 ns to 3.799 ns). The TSV can be entirely neglected in any delay or SI considerations if the conventional I/O drivers are used to drive the vertical interconnect. A typical I/O driver is intended to drive output capacitances in the order of Pico farads due to pin parasitics and output traces. The TSVs relatively small capacitance is almost entirely negligible due to the large drive strength of the IBIS buffer.

As the ELITE specification uses the I/O pads of the memory device to connect the TSVs to each die, the IBIS driver represents the drive strength and characteristics expected of the ELITE demonstrator. These characteristics have been shown to completely mask out the effects of the TSVs, regardless of switching pattern. For all delay or signal integrity places, the TSVs can be neglected from the electrical modelling of the vertical interconnect in the ELITE prototype. However if TSVs are to be used throughout the die for intermediate 3-D signalling, using less powerful drivers will require the use of the signalling design guidelines provided in this report.

3.4. Performance Comparison: Energy, Area, SI and Bandwidth

Conventional voltage mode signalling, voltage mode signalling with shielding, current mode signalling, low-voltage differential signalling and low-voltage single-ended signalling schemes were investigated for TSVs. The TSVs were arranged in a row and inter-via coupling capacitances and self capacitances of 89 fF and 3 fF respectively were used from field solver results as representative values for the structures envisaged within ELITE. The circuits were simulated in Cadence Spectre with input rise/fall times of 10 ps and period of 200 ps. The delay, energy, bandwidth, and coupled noise amplitude were extracted from the simulations for each scheme.

**Bandwidth and delay variation**

Output waveforms for VM, VM shielded, LVDS and CM signalling schemes were used to calculate the worst-case coupled delay of each scheme. The worst-case switching pattern assumes both neighbouring TSVs switch in the opposing direction to the victim TSV. The best-case pattern occurs when the aggressors switch simultaneously in the same direction as the victim. The output waveforms for all signalling schemes are plotted in Figure 70. The greatest delay is seen for the unshielded VM and LVSE drivers, while the least delay is for the CM and shielded VM cases, the latter at the loss of a reduced number of signal TSVs for the shields. The LVDS circuit does not see much change from the best-case to the worst-case switching pattern due to its relatively robust noise rejection.

![Figure 70](image_url)

**Figure 70** - Best and worst-case coupled rise times for CM, VM, VM shielded, LVDS and LVSE signalling schemes. $C_s=3$ fF and $C_c=89$ fF.
Eye diagrams at the output and TSV for each configuration reveal that shielded VM has the most reliable signalling, with LVDS showing only a minor amount of variation. The shielded VM driver demonstrates no variation in delay whilst the unshielded VM shows significant change between 8 ps and 55 ps for best and worst-case patterns. The CM scheme also shows a significant change in delay, from 4 ps to 22 ps. The LVDS circuit produces the least amount of variation, from 38 ps to 52 ps, due to the nature of its low-swing signalling and noise rejection circuitry. The maximum data rate for each case is shown in Table 14, where the data rate through a single TSV, in Gbps, is calculated from the worst-case delay and the ratio of signalling TSVs to shield TSVs. For the LVDS and shielded VM cases, two TSVs are required for a single data signal. The LVSE showed the greatest amount of variation, necessitating a wider pulse width to plot the eye diagrams, which indicates that this scheme may not be appropriate for high-speed applications.

**Signal Integrity**

The maximum coupled noise amplitude on a silent victim was measured before the final load inverter and is listed in Table 14 for each signalling mode. The shielded VM circuit shows insignificant variation in noise amplitude with switching pattern, whilst the CM configuration shows noise of up to 46% of $V_{dd}$ for a small receiver size. Increasing the size of the receiver slightly reduces the voltage overshoot; however due to the nature of the current mode driver nominally resting at roughly one third of $V_{dd}$, the effect of coupled noise is much more significant. The LVDS coupled noise amplitude is the least at 68 mV and the LVSE also has a relatively small coupled noise of 162 mV. The low-voltage schemes may prove desirable for noise immunity however the LVDS circuit is the only one that can reject common mode noise from external sources. This may prove a significant advantage when coupling is considered from the on-chip planar interconnect to the TSVs.

**Energy and area**

The energy consumption per bit for each signalling configuration was calculated from the spice simulation data to provide quantitative trade-offs between the different signalling modes. The energy per bit is normalized to the single-ended non-shielded voltage mode circuit. The information is shown in Table 14, where the CM receiver, MP2 and MN2, is sized to demonstrate the effect on delay and power for two configurations. The LVDS circuit exhibits the poorest performance in terms of energy and delay due to the complexity and the transistor count of the circuit. Increasing the LVDS driver size for optimal delay incurs a large power penalty, while reducing the driver size increases the latency beyond comparable levels to the other modes.

A fair metric for comparison between the different signalling strategies we have examined is the data rate per unit area, defined by:

$$\text{Maximum Data rate (per TSV)} = \frac{1}{t_d} \cdot \frac{N_s}{N_T}$$  \hspace{1cm} (22)

where $t_d$ is the worst-case delay, $N_s$ the number of signal TSVs, and $N_T$ the total number of TSVs (i.e. signal and shield lines). In the data rate comparison we have considered parallel TSVs as a representative unit to compare the following signalling strategies: VM without shielding, VM with every other TSV being a shield (i.e. a pattern of signal-ground-signal-ground or SGSG), LVDS (a pattern of SXSX where ‘X’ is the reference signal in a signal pair), LVSE and CM with different receiver terminations. The results are given in Table 14, which summarises the maximum achievable data rates as well as energy consumption per cycle.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Configuration</th>
<th>Data rate (Gbps)</th>
<th>Normalised Energy per Cycle</th>
<th>Coupled Noise Amplitude (mV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM</td>
<td>SSSS</td>
<td>18</td>
<td>1</td>
<td>612</td>
</tr>
<tr>
<td></td>
<td>SGSG</td>
<td>17</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CM</td>
<td>5</td>
<td>20</td>
<td>1.8</td>
<td>471</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>22</td>
<td>2.52</td>
<td>427</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>27</td>
<td>3.78</td>
<td>357</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>31</td>
<td>4.80</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>36</td>
<td>5.63</td>
<td>259</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>40</td>
<td>6.27</td>
<td>220</td>
</tr>
<tr>
<td>LVDS</td>
<td>SXSX</td>
<td>11.2</td>
<td>7.8</td>
<td>68</td>
</tr>
<tr>
<td>LVSE</td>
<td>SSSS</td>
<td>7.3</td>
<td>0.62</td>
<td>162</td>
</tr>
</tbody>
</table>

[71]
3.5. Discussions and Conclusions

Due to the active area loss and complexity of incorporating relatively large TSVs within multi-layer 3-D chip stacks, each TSV must be used as efficiently as possible to meet increasing demands on device performance and cost. The issue is further complicated by the ratio of coupling to self capacitance within a row of TSVs that is much higher than for on-chip wires, exacerbating noise and noise-on-delay effects. Innovative techniques that take advantage of the physical characteristics of TSVs must be developed in order to select an optimal signalling scheme. We have considered a host of signalling techniques including combinations of shielding, optimal VM signalling, low-voltage differential and single-ended signalling, and CM signalling, and made a comparison using data-rate, energy, delay variation and coupled noise as metrics, in order to understand optimal signalling schemes for TSVs.

We have shown, using realistic values for TSV parasitics and device models within a representative 65 nm technology that CM signalling can provide a higher data-rate than all cases, although at a cost of higher power consumption. LVDS signalling maximizes noise rejection and has the least delay variation but also at the cost of greater power and reduced data rate from requiring two TSVs per signal. LVSE signalling can improve power consumption significantly, but at a significant loss of speed. The signalling speed and noise over TSVs can be significantly improved with VM shielding but at a loss of area for grounded TSVs. However as can be seen from Table 14, this result only in a relatively small drop in data-rate, implying that shielding may be a simple and effective strategy for improving noise resilience while maintaining a high bandwidth if the applications is able utilise the higher data rates. An ideal situation is if the shield wires comprise existing power/ground lines interspersed with signal lines. As with on-chip signalling over planar wires, conventional rail-to-rail voltage mode signalling with inverters represents a simple and effective scheme with a reasonably high performance being coupled with a low-energy and area footprint, with all other schemes trading off one or more of the metrics of speed, noise, power, energy and area to achieve an improvement in the rest. We have also investigated the performance of IBIS models of the I/O drivers of an existing flash device to characterise signalling over TSVs in the ELITE prototype. We have shown that the drive strength of the I/O driver is designed for much larger capacitive loads than the TSV, such that the parasitics of the TSV have a negligible effect on the performance of the flash device. The quantification of these trade-offs in this analysis provides guidelines for 3-D Integrated Circuit designers.
4. Compact Thermal Modelling of 3-D ICs

In this section we examine the thermal behaviour of stacked ICs and construct general models and a stand-alone tool to estimate thermal effects in 3-D packages. We use the dimensions, materials and power maps of Flash memory dies and an associated memory controller to understand the thermal behaviour of the ELITE demonstrator in simulations using a Computational Fluid Dynamic (CFD) solver. We show that in low-power applications such as a NAND Flash memory stack, the thermal effects are minimal, where existing packaging technologies for memory devices are sufficient to handle the additional die layers. Using the compact thermal models verified by the CFD simulations, we show how they can be used to provide a general outlook of the limitations and performance of 3-D systems under thermal constraints.

4.1. Introduction

As the minimum feature size decreases with each technology generation, the overall die area remains relatively constant to increase the chip functionality. Typically, the total chip power scales downwards as a function of the feature size and supply voltage ($CV^2$). However, the demand for higher throughput and more functionality on a single chip has led to a trend of increasing power consumption despite the lower energy per operation. Higher power density not only spells trouble for IC package designers, but it also has a further effect on the leakage power of the transistors. At feature sizes below 65 nm, the gate leakage for bulk-CMOS transistors in sub-threshold operation can exceed 50% of the total power on the chip [57]. Leakage increases as transistors shrink (see Figure 71) and because of the strong thermal-dependence of gate leakage, it only worsens at higher temperatures.

So far we have discussed the current thermal design challenges for all high-performance CMOS ICs. Migrating high-performance designs to 3-dimensional packages will become even more difficult. 3-D integration inherently means a smaller form-factor and as a result, increased power density due to die thinning and tightly compacted active devices. Furthermore, active dies sandwiched between the top and bottom dies will have indirect thermal resistances to the ambient which may exacerbate thermal crowding and hotspots. For the above reasons, the design challenges in 3-D ICs will increase due the unfavourable thermal properties of stacked devices.

Consequently, adequate and accurate models of thermal behaviour will be essential from the early design stages through to the packaging stages. Thermal behaviour of integrated circuits is an essential consideration in the design of consumer electronics. Excessive temperature in a CMOS ICs can result in decreased reliability and device life,
increased delays in wires due to higher resistivity and longer device switching times due to degraded carrier mobility among others. Therefore, the overall performance, power and reliability of an integrated circuit will rely directly the thermal behaviour as we show in Figure 72.

![Thermal Design Dependence](image)

**Figure 72 - Thermal design dependence**

By establishing the thermal limitations of a particular device, the power, performance and reliability can be correctly determined and designed for. The work in this report concentrates on the thermal behaviour of stacked ICs in order to establish the performance and power limitations of future CMOS designs. We examine the thermal profile of the ELITE demonstrator as the prototype was designed to be constructed (a Ball-Grid Array (BGA) stack with NAND dies and a NAND controller). We also examine the behaviour of general and higher power CMOS ICs. The rest of the chapter is organised as follows. We first discuss related work and then explain our methodology for thermal simulation and modelling. We follow by discussing the setup and simulations to extract the thermal behaviour in the ELITE demonstrator and general stacked ICs. Finally we discuss compact thermal models and our simulation tool and end with our conclusions.

### 4.2. Related Work

The complete methodology to develop electrical and thermal models for various microsystems packages can be found in [58]. Early studies of thermal behaviour in 3-D ICs [59] [60] use 1-dimensional compact thermal models based on physical materials in the stack to predict temperature behaviour in a multi-tier 3-D system. 1-D models provide limited accuracy and cannot model the distribution of power across a die, nor can they capture the effect of hotspots in the package. In [61] the authors compare the thermal performance of a four-die NAND flash memory stack in a CFD simulator for different stacking orientations and power. Isolated CFD simulations are useful for analysing the behaviour of a particular design but do not provide an outlook of thermal performance in general packages. In [62] the authors describe the use of vias as thermal conduits to pass heat through a stack of dies which can reduce local hotspots but in turn cause increased routing congestion and large area overhead. 3-D ICE [63] and HotSpot [64] both offer thermal estimation of 3-D chip stacks but are not intended to model the underlying hardware features which generate the power profile.

### 4.3. Thermal Analysis of Integrated Circuits

Generally there are two methodologies to characterise the thermal properties of an integrated circuit: highly-accurate and intensive computational fluid dynamic (CFD) solver simulations based on structural layout including materials, dimensions and power maps or comparatively simple but reasonably accurate compact thermal models (CTM). We first describe our IC simulations using a CFD solver in FloTHERM and then detail our development of a CTM tool to analyse more general thermal effects and trends in 2-D and 3-D devices.
Simulation Tools and Environment
IC design tools incorporate solvers which directly import layout information and process technology to provide accurate thermal maps of a 2-D chip design. This is particularly useful for characterising the thermal integrity of a device before the packaging stage. These tools are ideal if layout and process information is available, but do not provide the thermal properties of a general system, or give a complete picture of thermal trends for other devices or applications. Tools that can import a full 3-D IC stack and provide a combined thermal response are still under development.

Simulation of the 3-D Flash Memory Demonstrator
To examine the thermal effects of the proposed ELITE 3-D Flash device, we have used the CFD solver from FloTHERM with materials, dimensions and power maps expected of the two Numonyx/Micron Flash memory dies and the Hyperstone controller. Our simulations obtain the steady-state behaviour of the ELITE prototype under worst-case conditions. Transient information is not required in this case as the package design usually depends on the maximum thermal design power (TDP) given by the worst-case die power/temperature.

![Image](image_url)

Figure 73 - The JEDEC standard test box used for thermal simulations in FloTHERM. The IC package is mounted on 4-layer thermal test PCB which is secured by plexiglass fixtures in a still air environment at 25 C.

Due to the large dimensions of the test enclosure (10s of cm) compared to the relatively small dimensions of the TSVs (10s of μm), we do not model the intra-die interconnect as individual structures. Their inclusion results in lengthy simulations times, memory overruns and numerous convergence issues due to the mesh resolution. Instead, the bonding interface with TSVs is modelled as a solid layer with a certain ratio of copper (TSVs) to die attach (epoxy-type) which alters the average thermal conductivity. The effect of the TSV on the thermal conductivity is small due to the large dimension of the die surface as compared to the copper vias (although modelling individual vias in hotspots may be important). Otherwise, all aspects of the package are modelled, including a multi-layer PCB with thermal vias, a solder ball grid array, substrate and package overmold.

The CFD simulations follow the JEDEC standard No. 51-2A (JESD2512A) for the thermal evaluation of IC packages. The JEDEC standard calls for a still air test enclosure consisting of Plexiglas, illustrated Figure 73. The package-under-test is mounted on a 4-layer thermal PCB consisting of copper ground and power planes sandwiched between FR-4 dielectric material and connectivity trace layers on the top and bottom. A thermocouple is mounted below the PCB to monitor the ambient reference temperature and junction temperatures are monitored at various places within the package and the PCB.
The ELITE demonstrator specifications were taken from internal reports and typical dimensions of a 16 GB Flash memory die and a memory controller. Figure 74 is a diagram of the general setup of the ELITE demonstrator in a ball-grid array (BGA) package. The 144 I/O BGA package is mounted directly on the PCB with an array of thermal vias directly underneath to aid in the dissipation of heat to both sides of the PCB and replicate densely packed IO and PCB interconnect. The BGA substrate is a mix of silicon and copper traces, so a typical thermal conductivity for BGA substrates is used. According to the ELITE specification there is an additional silicon substrate layer which serves as an interposer and intermediate layer between the active dies and the BGA package. The next two layers are the two 16 GB NAND Flash dies and the top layer is the smaller controller. In the final specification, the controller is wire-bonded to the silicon intermediate layer which the two Flash dies are also connected to with TSVs. The silicon layer is wire-bonded to the BGA package with all the required connections.

We use a die model that approximates the thermal conductivities of the interconnect, dielectric and the die surface to reduce the computational complexity of the simulations. We add an epoxy-like adhesive layer (die attach) to the die model and the bulk silicon back end. The model for the dies is shown below in Figure 75.

A typical Flash memory die will consume around 100 mW on average. To accurately produce the thermal profile of the 3-D stack a power map must provide correspondence to the switching activity of the underlying hardware. Typically, the majority of the power may concentrate in certain areas on the chip where higher power components are located, such as clock buffers. These areas are known as hotspots which can cause significant problems in 3-D stacks, especially if they overlap (very possible if identical dies are stacked). In Figure 76 we show the power map used for the Flash dies in the stack. On the left side of the figure a micrograph of a 16 Gb NAND flash die is dissected into its principle power consuming components. The right side of Figure 76 shows the breakdown of the power consumption as a percentage of the total die power (in this case 100 mW). Typically the logic and drivers (charge pumps, decoders and I/O) consume more power than the memory arrays themselves. We have distributed the power based on the functional components in a NAND Flash device in this manner.

Figure 74 - Diagram of the ELITE demonstrator thermal model

Figure 75 – The breakdown of the die material model properties is shown (above) and the overall dimensions shown for the bottom Flash die (below).
The Flash controller power information is extracted from data sheets on Hyperstone’s product information page. According to the data sheet of the F-series Flash memory controllers, the maximum power consumption is roughly 50 mW, motivating our power map in a similar manner as the Flash dies.

**Thermal profile of the demonstrator**

We conduct our simulations with an ambient still air temperature of 25°C and the die power maps and materials as mentioned in the previous section. We take readings at all junctions (where dies or materials meet) with reference to an ambient thermocouple located at the bottom of the enclosure as specified by the JEDEC standard. The package-under-test is shown in Figure 77. We mainly focus the CFD simulations on flash memory and controller stacks as it falls under the objective of the ELITE 3-D specification.

Figure 76 - The micrograph of a typical die is shown on the left where the architectural components of the system are noted. A power map for the die was constructed based on the estimated power contribution of each component of the system based on circuit-level descriptions. Die micrograph [85].

The Flash controller power information is extracted from data sheets on Hyperstone’s product information page. According to the data sheet of the F-series Flash memory controllers, the maximum power consumption is roughly 50 mW, motivating our power map in a similar manner as the Flash dies.

**Thermal profile of the demonstrator**

We conduct our simulations with an ambient still air temperature of 25°C and the die power maps and materials as mentioned in the previous section. We take readings at all junctions (where dies or materials meet) with reference to an ambient thermocouple located at the bottom of the enclosure as specified by the JEDEC standard. The package-under-test is shown in Figure 77. We mainly focus the CFD simulations on flash memory and controller stacks as it falls under the objective of the ELITE 3-D specification.

Figure 77 – (a) The 3-Dimensional model of the ELITE prototype with (b) temperature contour distribution on a standard JEDEC 4-layer thermal test PCB

Figure 78 plots the temperature rise in a stack of two flash memory dies and a controller with a total power of 1 Watt. The simulation shows the rise in temperature within the stack where die closest to the heatsink (at the top) has a lower temperature than the die beneath it. This may be an important design consideration as within the die stack, the electrical performance is very high between any of the dies (i.e. signalling from die 1 to die 4 in a stack is similar to die 1 to die 2 from an electrical perspective). However, some dies may have more aggressive feature sizes.
or power profiles than others resulting in higher consumption and leakage currents. Thus, the die stack should be partitioned from a thermal perspective such that the high performance dies are closest to the heat sinks. In this instance, the low power of the flash memory and controller chips do not necessitate any special considerations and it is expected that large memory stacks are thermally feasible (where the temperature has a relatively linear relationship with the power and number of dies.)

![Thermal profile of the flash stack](image1)

**Figure 78 - Thermal profile of the flash stack**

Although the power of the stack is mostly consumed in the charge pumps and I/O located on one end of the die stack, the temperature distribution is mainly even across the package as Figure 79 shows. At low power densities seen in flash memory systems, the overlap of hotspots does not cause excessive crowding and the standard package is sufficient to spread the heat. This may change for higher-power stacks.

![Top view of stacked package](image2)

**Figure 79  Top view of stacked package. Although most of the power is concentrated on the left side of the package, the distribution is mainly even because the heat spreading is sufficient at flash memory power levels.**

We additionally examine the limitations of more general 3-D stacks. Figure 81 compares the thermal performance of 2-D and 3-D stacks under forced convection using a high velocity 85 CFM heatsink. A high density I/O ball grid array package is mounted to the PCB with three different die configurations: a 400 mm$^2$ 2-D, a 2-layer 3-D where
each die is 200 mm² and a 4-layer 3-D where each die is 100 mm². The power per die is then increased (where the 3-D configurations split the total chip power evenly over each die layer). Since the maximum temperature of DRAM is roughly 100 °C, we impose an upper limit on the temperature and increase the power until that value is reached at any point in the stack. The plots in Figure 81 establish the maximum thermal capability of a partitioned system, where 3-D designs have a much lower limit than a 2-D system. Despite this, we show later in the system-performance studies how 3-D systems can still achieve higher performance efficiency than an equivalent 2-D design. The simulation geometry and cross-sectional contour for the simulation is shown in Figure 80.

Figure 80 Cross-sectional view of the thermal package under test. Tests were conducted to find the maximum power in 2-D and 3-D stacks under forced convection at high fan speeds.

The CFD simulations conducted in this section aim to model the thermal behaviour of a stacked flash memory-controller device and establish the upper limits on power for general 3-D stacks. The analysis is also used to verify the accuracy of the compact thermal simulator discussed in the next section.

Figure 81 The growth in temperature as a function of power for three systems: a 400 mm² 2D, a 200 mm² 2-layer 3D2 and a 4-layer 100 mm² 3D4 package. A theoretical upper limit of 100 °C is imposed as the maximum operating temperature of commercially available DRAM.

**Compact Thermal Models**

Thermal compact models of 2-D and 3-D packages have been developed for many applications in IC design, as we summarise in the related works section of this chapter. The thermal behaviour of any system has a close resemblance to the electrical behaviour of a circuit. Figure 82 shows the electrical equivalent components of a thermal system for any physical system. The thermal resistance, $R_{th}$, can be determined from the thermal conductivity of the material, $k$, the length, $L$, and cross-sectional area, $A$, of that material:
A power source on a die can be represented by an ideal current source, $q$, and the ambient temperature of the environment, $T_a$, is a constant voltage source. The current injected into a node, $T_j$, induces a voltage across the equivalent resistance, which can be calculated by the thermal “Ohm’s law”:

$$T_j = qR_{th} + T_a$$  \hspace{1cm} (24)
electrical equivalent circuit (see Figure 82) of the entire stack which is written to a SPICE equivalent circuit file. The MATLAB routine calls NG-SPICE with the equivalent circuit and performs a DC operating point analysis to find the steady state voltages and writes the 3-D I/V matrix to a raw file. Finally, the MATLAB program reads the raw data and plots the results as temperature contours for each die layer. The code flow for the simulator is described in Figure 83.

![Flow-chart for 3-D Compact Thermal Model-based package simulator.](image)

The tool (shown in Figure 84) allows any order of die stacks with a host of user-definable input parameters and IC package solution. As a contrast to computationally intensive CFD simulations, this application is orders of magnitude faster and requires no expert knowledge of simulators or CAD drawing to execute. Unlike other available tools based on compact thermal modelling, this thermal tool runs on multiple operating systems, requires no installation and is also made available with a simple web interface. This work enables fast exploration of the 3-D thermal design space for the average user. We show in the system-level modelling section how these models can be used to determine maximum performance limitations for realistic processor applications under various physical constraints.
4.4. Discussion and Conclusions

We have conducted CFD simulations of the ELITE demonstrator and shown the thermal profile of the chip stack with realistic power, material and ambient conditions for a standard thermal test structure. The demonstrator shows a small increase in temperature with each layer, but the design issues are minor and should be easily mitigated by current packaging technology for memory devices. We have also shown thermal simulations and trends for more general 3-D stacks, where the temperature increases linearly with the number of chips in the stack and die power. This is largely due to the higher thermal resistance of the package rather than the die-to-die resistance. Finally, we have discussed compact thermal models and a stand-alone tool we have developed in the ELITE project which shows thermal trends in 3-D ICs given a large number of user input parameters.

Figure 84 – User Interface and output for the stand-alone thermal solver. Figure 85 Temperature contour from the top of the die. Although the power distribution is greater on the left side (where logic, I/O, sense...
5. Communication Platform Architectures

The shift from single-core to multi-core processors has long since occurred and new designs are migrating towards many-core systems. The performance bottlenecks imposed by unfavourable on-chip interconnect delays and congestion in shared buses has finally enabled the packet-switching Network-on-Chip framework to become a viable architecture for high throughput computing. By standardising the communication hardware, multifunctional resources can be seamlessly integrated into a system with the promise of scalable performance for the future.

In this chapter we move one level up in the modelling hierarchy to explore the communication behaviour of parallel systems to optimise the configuration of on-chip mesh networks. First, we derive the correct model for average distance in mesh networks and show how it can be used to optimise the placement of hotspots and configure asymmetric networks under different clocking schemes. We then develop a cycle-accurate simulation model to investigate the performance and scalability of various 2-D and topologies under physical constraints, different vertical communication routing schemes and traffic patterns.
5.1. Optimal Network Architectures

A general expression for the average distance for meshes of any dimension and radix, including unequal radices in different dimensions, valid for any traffic pattern under zero-load condition is formulated rigorously to allow its calculation without network-level simulations. The average distance expression is solved analytically for uniform random traffic and for a set of local random traffic patterns. Hot spot traffic patterns are also considered and the formula is empirically validated by cycle true simulations for uniform random, local, and hot spot traffic. Moreover, a methodology to attain closed-form solutions for other traffic patterns is detailed. Furthermore, the model is applied to guide design decisions. Specifically, we show that the model can predict the optimal 3-D topology for uniform and local traffic patterns. It can also predict the optimal placement of hot spots in the network. The fidelity of the approach in suggesting the correct design choices even for loaded and congested networks is surprising. For those cases we studied empirically it is 100%.

Introduction

Two important metrics of performance for NoCs are latency and throughput, generally functions of network characteristics such as topology; interconnect characteristics, routing scheme and switch architecture, as well as application characteristics primarily defined by traffic pattern. The average distance is a NoC performance metric that depends on the topology and the traffic pattern only, under the assumption that the network operates well below its saturation point. We express it as a closed formula comprising a sum over all source-destination node pairs for arbitrary n-dimensional radix-k mesh networks and for arbitrary spatial traffic patterns. A traffic pattern is defined as the packet exchange probability for each source-destination pair. Average distance is an upper bound on the performance for all possible routing, switching, and flow control algorithms. For instance, a routing algorithm is optimal for a given traffic pattern if packets on average do not travel more than the average distance through the network.

We define the distance of communication as the minimum number of switch-points or nodes that a packet has to traverse from a source node to a destination node. It is measured in hops. The average distance \( \bar{H} \) is the average distance of all packets in the network under a given traffic load, and a function of the dimension and radix in a mesh topology. It is a useful basic metric providing insight into the performance of the overall network.

The starting point in our analysis is the rigorous formulation of an expression for the average distance in k-ary n-dimensional meshes, including unequal radices in the different dimensions, that is valid for any traffic pattern. We proceed to evaluate this expression for Uniform Random Traffic (URT), Local Random Traffic (LRT), and Hot Spot Traffic (HST) and verify through network simulations that the resulting formula accurately predicts the average latency for unloaded networks. The upshot of this is that the upper bound on performance given by the average latency of a k-ary n-dimensional unsaturated network for any traffic pattern can be estimated from the general model we propose, without running network simulations, which saves both model development time and computation time.

In the case of URT, the general expression can be solved to yield a closed-form expression for the delay, which is proven to be more exact and more general than other expressions available in the literature. We also propose empirical closed-form solutions for three specific LRT patterns based on the Response Surface Method (RSM) and illustrate the general methodology to obtain closed-form solutions for an arbitrary traffic pattern.

Due to the simplicity of the model, it is particularly useful in finding optimal network architectures to minimize average communication delay under any traffic pattern, including empirically validated traffic models for irregularly sized networks with multiple constraints. For relatively simple traffic patterns such as URT, LRT or empirically-validated models where a closed-form expression can be obtained for the delay, the optimization problem takes a few seconds of computation time. In our results we show solutions for optimal architectures under different traffic patterns and boundary conditions that can be obtained from a few iterations at most with minimal computation time using any general purpose programming language such as Matlab or C, rather than computationally expensive cycle-accurate packet-level simulations. However even for those cases where the traffic model does not allow such a formulation, the general model for average distance we propose has a computational complexity \( O(N^2) \) in the worst-case where \( N \) is the number of nodes in the network, meaning that a brute-force search to find the optimal
network configuration for network sizes up to say a thousand nodes is feasible on a desktop. Notably our model exhibits 100% fidelity for all simulations carried out for three types of traffic (namely URT, LRT, and HST), where the optimum architecture to minimize latency for unloaded networks remains optimal even under congestion up to the point of saturation.

The main contribution of this study is in providing a generally valid model for the average distance and accompanying analysis that provides insight into network behaviour under common traffic loads and constraints.

Related Work
The analytic formulæ for average distance provided in the literature either cover only a special case, or the assumptions made are not fully explained leading to misunderstanding. Agarwal [65] gives

$$\bar{H} = \frac{n}{3} \left(k - \frac{1}{k}\right)$$ (26)

as the average distance in an n-dimensional mesh with radix k.

The formula assumes that k is the same in all dimensions and no derivation or further motivation is provided.

Liu et al. [66] provide a formula for a $k_1 \times k_2$ 2-D mesh:

$$\bar{H} = \frac{1}{3} \left(k_1 - \frac{1}{k_1}\right) + \frac{1}{3} \left(k_2 - \frac{1}{k_2}\right)$$ (27)

Here the case with different radices is covered but only for two dimensions. The derivation of the formula is somewhat unclear because it contains an approximation step, that replaces $k_1 k_2 - 1$ by $k_1 k_2$ without motivation or explanation. It seems that the authors intended this formula only to be an approximation. They exclude the self-traffic case (when a node is allowed to send packets to itself) which is in contradiction to Agarwal’s assumptions.

Dally and Towles [67] offer

$$\bar{H} = \begin{cases} \frac{nk}{3} & k \text{ even} \\ n \left(\frac{k}{3} - \frac{1}{3k}\right) & k \text{ odd} \end{cases}$$ (28)

for $k + 1$ nodes in each dimension. This expression is in contradiction to both Liu’s and Agarwal’s formulæ. However, closer inspection shows that, under the corrected assumption of radix k (rather than $k + 1$), the odd case formula is identical to Agarwal’s formula and the even case is an upper bound of $\bar{H}$, approaching the true value asymptotically for large k.

Holsmark [68] devotes appendix I of his Licentiate thesis to the clarification of the average distance. He concludes with

$$\bar{H} = \frac{k_1 + k_2}{3}$$ (29)

for a $k_1 \times k_2$ mesh assuming no self-traffic. Agarwal and Dally et al. include the self-traffic case. Taking this into account, it turns out that Holsmark’s formula is consistent with Agarwal’s expressions, but he covers only 1-D and 2-D meshes.

We provide a derivation for the average distance in n-dimensional meshes with the general case of unequal radices along the different dimensions. This is important as many practical on-chip networks serving a few tens or hundreds of cores in multiprocessor systems are often irregular, and the number of nodes along the x, y, and z dimensions in a 3-D system for example, are seldom equal. We show how this formula is exactly correct, and is a generalization of all the models mentioned above.

Koohi et al. [69] present abstract performance models in a spirit similar to ours. They propose power and throughput models for uniform, local, hotspot, and first matrix transpose traffic models. Their approach is more empirical since they use simulation results as starting point and analyse the effect of combinations of different traffic models to derive comprehensive throughput and power models for mixed traffic patterns. In contrast, our approach focuses on distance and latency, deriving an analytical formula which is further validated through simulation. Primarily we focus on the unloaded case but we show that design decisions based on minimum latency provided by our model also hold true in loaded networks. While Koohi et al. consider only 2-D networks, we have special interest in 3-D topologies.
In a seminal paper from 1990 Dally [70] studied the communication performance of k-ary n-dimensional tori, a similar class of topologies that we cover (meshes rather than tori). The paper analyses average latency in networks with different dimensions (mostly between 2-20) under different cost constraints for routers and wires, ranging from “wires are free and infinitely fast” to “limited wires per router” and a linear delay wire model. The paper identifies the optimal dimension constrained by the cost and wire model assumed. It almost exclusively deals with uniform random traffic, except a short section where hot spot traffic is briefly discussed. Our work makes more specific assumptions on wire cost and wire delays based on realistic implementations for on-chip planar interconnects and in a 3-D stack. But we cover general traffic patterns and show how the model facilitates topological exploration and hot spot placement.

A significant body of research comprises the development of network performance models for packet latency in congested networks. Even in the narrower scope of Networks-on-Chip the published literature on this topic is substantial. Most of the work (e.g. [69-77]) make very specific assumptions about routing (mostly deterministic, dimension order routing) and switching (mostly wormhole switching). The majority focuses on average delay [69-74] while some work targets worst-case delay [75], [76]. Sometimes even more specific assumptions are made such as single flit buffers [78] or one dedicated virtual channel for each flow [79].

In contrast with all this and similar work, we do not offer a delay model under congestion or for a specific switch architecture, but we start with an analytical formula for average distance which is valid for all k-ary n-dimensional meshes and for any traffic pattern, but abstracts entirely from routing and switching techniques. We then derive average distance formulas for specific traffic patterns and we illustrate the usefulness of this abstract, ideal formulae. In particular, we use the expression for average distance to investigate optimal architectures for minimizing delay when the link delays are not necessarily equal.

Calculating Average Distance
First, we derive a general average distance formula for k-ary n-dimensional networks and for arbitrary traffic patterns. Then, we derive a closed form solution for Uniform Random Traffic. This is straight forward and the result is consistent with earlier published formulas. We follow by deriving a closed form formula for a specific type of Local Random Traffic, which is a more complicated derivation. Then we introduce HotSpot Traffic but do not provide a closed formula due to the intractable dependency on the precise location of the hotspots. But later on we show how an optimal placement of hotspot nodes can be found by minimizing the average distance.

The average distance of a network is the ratio between the total distance that the packets emitted by all switches travel and the total number of packets emitted. The total distance that the packets emitted by all switches, in a network with nodes N, travel is:

$$D_t = \sum_{A \in N} \sum_{B \in N} p_{A,B} \times d_{A,B}$$ (30)

where A and B are any given nodes, $d_{A,B}$ is the Manhattan distance between the node A and node B, and $p_{A,B}$ is the probability that a packet is sent from node A to node B. $p_{A,B}$ defines the traffic pattern and can be an arbitrary function e.g. for local random traffic $p_{A,B}$ is a function of the distance $d_{A,B}$. This equation is a general formulation for any k-ary n-dimensional network.

The total number of paths $N_p$ traversed by the emitted packets considered here is:

$$N_p = \sum_{A \in N} \sum_{B \in N} p_{A,B}$$ (31)

Hence the Average Distance $D_{avg}$ is given by:

$$D_{avg} = \frac{D_t}{N_p} = \frac{\sum_{A \in N} \sum_{B \in N} p_{A,B} \times d_{A,B}}{\sum_{A \in N} \sum_{B \in N} p_{A,B}}$$ (32)
In order to demonstrate the formula (34), we calculate the average distance over a single dimension by considering a 1-D array shown in Figure 86. Each node connects to every other node, and each such path has an associated distance. Shown in the figure are the distances $h_{i,j}$ from the first node ($S_{1,1}$) to the other nodes in the array. The average distance for a k-ary 1-D array is:

$$H_{1\times k} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{k} P_{i,j} \times |i-j|}{\sum_{i=1}^{k} \sum_{j=1}^{k} P_{i,j}}$$

(33)

1) Uniform Random Traffic

For Uniform Random Traffic, each node generates the same number of packets uniformly distributed over time, and all destination nodes are equally likely. Hence the probabilities are the same for all source-destination pairs, leading to $p_{i,j} = p_{URT}$ where $p_{URT}$ is a constant. Therefore, (33) can be simplified thus:

$$H_{URT 1\times k} = \frac{p_{URT} \sum_{i=1}^{k} \sum_{j=1}^{k} |i-j|}{p_{URT} \sum_{i=1}^{k} \sum_{j=1}^{k} 1} = \frac{\frac{k^2}{3} - \frac{k}{3}}{k^2}$$

(34)

leading to

$$H_{URT 1\times k} = \frac{k}{3} - \frac{1}{3k}$$

(35)

In the case of URT, since the probability is not a function of distance that the packets travel, the average dimension for an n-D network can be calculated by adding the average distance for each dimension. Therefore for an n-D array, the average distance for URT can be expressed using the following closed-form equation:

$$H_{URT} = \sum_{i=1}^{n} H_{URT 1\times k} = \frac{k_1}{3} - \frac{1}{3k_1} + \frac{k_2}{3} - \frac{1}{3k_2} + \ldots + \frac{k_n}{3} - \frac{1}{3k_n}$$

(36)

When the radix along each dimension is identical, it is true that $k_1 = k_2 = k_3 = \ldots = k_n = k$, and (36) reverts to (26) and the odd case of (28).

For $n = 2$ equation (36) simplifies to (27). Although Liu et al. [66] intended an approximation, they ended up providing an exact formula for the case that includes a node sending traffic to itself. Even though in their setup they excluded the self-traffic case, the approximation they included as a corrective measure resulted in exactly the right formula including self-traffic.
Finally it can be seen that taking the different assumptions into account (i.e. with and without self-traffic) equation (36) reverts to equation (27) for \( n = 2 \) as follows. Since the average distance is computed by dividing the sum of all distances by the number of paths, we correct for these different assumptions by multiplying with the number of paths in the self-traffic case \( ((k_1 k_2)^2) \) and dividing by the number of paths in the case with no self-traffic \( ((k_1 k_2)^2 - k_1 k_2) \). Since the distance for self-traffic is 0, the sum of all distances does not differ in the two cases. Setting \( n = 2 \) and applying this correction to (31) results in:

\[
\frac{1}{3} \left( k_1 - \frac{1}{k_1} + k_2 - \frac{1}{k_2} \right) \left( k_1^2 k_2^2 - k_1 k_2 \right) = \frac{1}{3} (k_1 + k_2)
\]

which turns out to be identical to Holsmark’s equation (27).

### Local Random Traffic

For local traffic, a variety of probabilistic models can be considered, with the probability that a given node being the destination is some inverse function of the distance from the source: \( p_{A,B} = \frac{1}{\lambda_A (d_{A,B})} \) where \( \lambda_A \) is the normalization constant defined by \( \sum_{B \in N} p_{A,B} = 1 \), and \( d_{A,B} \) is the distance between nodes A and B.

For \( f(d_{A,B}) = d_{A,B}^2 \), Table 15 compares the average distance estimated by (32) and RTL network simulations, for different traffic models and network sizes. The RTL simulator employs buffer-less 5-port (2-D) and 7-port (3-D) switches, where one port serves the independent packet generating resources. The switches are buffer-less and the routing decisions are based on an address minimizing “hot potato” deflection algorithm. The simulation invokes a configurable mesh to instantiate any size network in 2 or 3 dimensions. In order to comply with the RTL cycle-accurate simulations, we excluded self-traffic from formula (32) by considering \( A \neq B \). To validate the model, simulations are performed at low packet injection rates to ensure an unloaded network. The results show a good agreement between the formula and the simulation results, where the discrepancy between the model and simulation is a result of stochastic deviation in the generation of packet destinations and rounding errors. Further, for URT the closed-form solution as given in (36) for 3-D and the general expression (32) show exactly the same result.

### Table 15 Comparison of Average Distance Calculated by the Formula and Simulations

<table>
<thead>
<tr>
<th>Network Size</th>
<th>Probability Model</th>
<th>Average Distance</th>
<th>%Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Formula</td>
<td>Simulation</td>
</tr>
<tr>
<td>5x5x5</td>
<td>URT</td>
<td>4.8300</td>
<td>4.813</td>
</tr>
<tr>
<td>6x6x6</td>
<td></td>
<td>5.8600</td>
<td>5.888</td>
</tr>
<tr>
<td>7x7x7</td>
<td></td>
<td>6.8772</td>
<td>6.971</td>
</tr>
<tr>
<td>8x8x8</td>
<td></td>
<td>7.8900</td>
<td>7.931</td>
</tr>
<tr>
<td>9x9x9</td>
<td></td>
<td>8.9000</td>
<td>8.976</td>
</tr>
<tr>
<td>10x10x10</td>
<td></td>
<td>9.9090</td>
<td>9.894</td>
</tr>
<tr>
<td>4x8x16</td>
<td></td>
<td>9.9090</td>
<td>10.0008</td>
</tr>
<tr>
<td>5x5x5</td>
<td>LRT: ( \alpha = 1.0 )</td>
<td>3.7900</td>
<td>3.81</td>
</tr>
<tr>
<td>6x6x6</td>
<td></td>
<td>4.5900</td>
<td>4.555</td>
</tr>
<tr>
<td>7x7x7</td>
<td></td>
<td>5.3900</td>
<td>5.418</td>
</tr>
<tr>
<td>8x8x8</td>
<td></td>
<td>6.1900</td>
<td>6.146</td>
</tr>
<tr>
<td>9x9x9</td>
<td></td>
<td>7.0000</td>
<td>6.969</td>
</tr>
<tr>
<td>10x10x10</td>
<td></td>
<td>7.8060</td>
<td>7.855</td>
</tr>
<tr>
<td>5x5x5</td>
<td>LRT: ( \alpha = 1.5 )</td>
<td>3.1800</td>
<td>3.163</td>
</tr>
<tr>
<td>7x7x7</td>
<td></td>
<td>4.4781</td>
<td>4.498</td>
</tr>
<tr>
<td>4x8x16</td>
<td></td>
<td>5.3757</td>
<td>5.301</td>
</tr>
</tbody>
</table>
Closed-Form Formula for LRT: To provide a closed form solution for non-URT traffic, we propose an empirical equation for average distance of a 3-D NoC, using the Response Surface Method (RSM) [80], which is a collection of mathematical and statistical techniques useful for the modelling and analysis of problems in which a response of interest is influenced by several variables. This same methodology we demonstrate for our local traffic pattern can be applied to any custom traffic pattern. The second-order approximation function with three variables using RSM is defined as:

\[ y = b_0 + \sum_{i=1}^{m} b_i x_i^2 + \sum_{i=1}^{m} \sum_{j=1}^{m} b_{ij} x_i x_j + \varepsilon \tag{37} \]

where \( x_i \) and \( x_j \) are the design variables, \( b_0, b_i, b_{ij} \) are called regression coefficients, \( \varepsilon \) is the error, and \( m \) is the number of variables. For various values of \( x_i, i=1,...,m \) the dependent variable \( y \) is found from experiment. The relationship between a set of independent variables and the response \( y \) defined by the regression coefficients is determined using the method of least squares. In general, (37) can be written in matrix form:

\[ Y = bX + E \tag{38} \]

where \( Y \) is defined to be a matrix of measured values (of size \( p \times 1 \)), \( X \) to be a matrix of independent variables, \( b \) to be the regression coefficient matrix (of size \( p \times 1 \)), and \( E \) to be the error matrix ( of size \( p \times 1 \)). Then, the solution of (13) is:

\[ b = (X^TX)^{-1}X^TY, \tag{39} \]

which are the regression coefficients for (37).

Using the RSM model, for a 3D NoC with size \( k_x \times k_y \times k_z \), the average distance can be expressed as follows:

\[ H_{f_{lt}} = b_0 + b_x k_x + b_y k_y + b_z k_z + b_{xx} k_x^2 + b_{yy} k_y^2 + b_{zz} k_z^2 + b_{xy} k_x k_y + b_{xz} k_x k_z + b_{yz} k_y k_z \tag{40} \]

For example, regression coefficients for local traffic model LRT: \( \alpha=0.5, 1.0, 1.5 \) are shown in Table 16. In some cases when either of \( k_x, k_y \) or \( k_z \) is 2 the maximum error in the functional form can be as high 13%, however the error for all cases on average is shown to be lower than 1%.

**Hotspot Traffic**

To represent less than ideal networks, we also consider that some nodes in the network will attract more traffic than others. In this case the probability of sending packets originating from a particular node to hotspots is higher than that for non-hotspot nodes. If the fraction of packets generated at any given node that have a hotspot destination is \( p_{hst} \), the fraction of packets addressed to non-hotspots will be \( (1 - p_{hst}) \). Packets for hotspots and non-hotspot can also be assigned using URT and LRT. In this analysis we assumed that 80% of packets from one node is uniformly distributed among two hotspots while the rest is uniformly distributed among all the non-hotspots. We consider different strategies for hotspot placements within the network and use our model to demonstrate optimum placements, which can be a significant factor in reducing congestion and communication bottlenecks and improving overall system performance.
Table 16 Coefficients for the closed-form equation for Average Distance for a network of size \(k_x \times k_y \times k_z\), where \(k_x, k_y \in 2, \ldots, 10\) and \(k_z \in 2, \ldots, 30\).

<table>
<thead>
<tr>
<th>b</th>
<th>LRT: (\alpha=0.5)</th>
<th>LRT: (\alpha=1.0)</th>
<th>LRT: (\alpha=1.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\gamma=1)</td>
<td>(\gamma=0.5)</td>
<td>(\gamma=0.25)</td>
</tr>
<tr>
<td>(b_0)</td>
<td>-0.4915</td>
<td>-0.4455</td>
<td>-0.4224</td>
</tr>
<tr>
<td>(b_x)</td>
<td>0.3556</td>
<td>0.3472</td>
<td>0.3430</td>
</tr>
<tr>
<td>(b_y)</td>
<td>0.3556</td>
<td>0.3472</td>
<td>0.3430</td>
</tr>
<tr>
<td>(b_z)</td>
<td>0.2804</td>
<td>0.1423</td>
<td>0.0732</td>
</tr>
<tr>
<td>(b_{xx})</td>
<td>-0.0061</td>
<td>-0.0054</td>
<td>-0.0051</td>
</tr>
<tr>
<td>(b_{yy})</td>
<td>-0.0061</td>
<td>-0.0054</td>
<td>-0.0051</td>
</tr>
<tr>
<td>(b_{zz})</td>
<td>-0.0013</td>
<td>-0.0008</td>
<td>-0.0005</td>
</tr>
<tr>
<td>(b_{xy})</td>
<td>0.0022</td>
<td>0.0027</td>
<td>0.0029</td>
</tr>
<tr>
<td>(b_{xz})</td>
<td>0.0022</td>
<td>0.0019</td>
<td>0.0014</td>
</tr>
<tr>
<td>(b_{yz})</td>
<td>0.0030</td>
<td>0.0019</td>
<td>0.0014</td>
</tr>
<tr>
<td>%Avg. Error</td>
<td>0.49</td>
<td>0.44</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Application of the Model

In design space exploration it is often of interest to find the topology that minimizes delay under various constraints imposed by technological, physical and system-level requirements. For the unsaturated case, delay is a straightforward function of the average distance, and such constrained optimization problems can be solved accurately with the proposed analytical model by treating it as the objective cost function. In this section we demonstrate how our model for average distance can be used to provide performance comparisons between networks and optimise the topological configuration of nodes in 2-D and 3-D meshes for any traffic pattern.

A 3-D mesh is an increasingly common topology with the advent of 3-D Integrated Circuits (IC). Equal radices in each dimension, translating to a cube with the same number of nodes or switches however, is an unrealistic arrangement; rather, the number of nodes in each dimension are likely to be different, especially in the vertical direction. For example, recent work [81] has shown that the lower physical delay associated with the vertical interconnects in a 3-D stacked IC can enable higher data rates in the vertical dimension by clocking die-to-die links at greater frequencies than the horizontal dimensions. This is largely due to the relatively lower parasitics of through silicon vias (TSV) used to connect vertical die layers as compared to long planar wires used on a 2-D IC.

1) Optimisation of Network Topology

The most efficient network topology for minimizing latency under different vertical and horizontal clocking constraints can be found by solving the constrained optimization problem of minimizing

\[
D_{3D} = \frac{\sum_{A \in N} \sum_{B \in N} p_{A,B} \times (|x_A - x_B| + |y_A - y_B| + r|x_A - z_B|)}{\sum_{A \in N} \sum_{B \in N} p_{A,B}}
\]

subject to \(k_x \times k_y \times k_z = N\) where \(N\) is the total number of nodes. Here \(d_{AB}\) has been obtained by multiplying the distance in each dimension by the corresponding clock period, under the assumption that the periods in the x and y dimensions are equal and normalized to 1, and that the period in the z dimension is shorter, where \(0 < r \leq 1\).

The optimum network size under the given constraint is found using an extensive brute-force search algorithm, and the solutions for different \(N\) are shown in Table 17. As described earlier, closed-form equations for different traffic models were obtained considering three different \(r\) values. These equations are used for to find the optimal network sizes which considerably reduces the computational load when compared to the general expression in equation (32).
Table 17 Optimum network sizes for different traffic models under planar clock speeds, $\gamma$, of 0.5 and 0.25 times the 3-D link clock. Uniform vertical and horizontal clock speeds through a network intuitively perform at their optimum latency in symmetrical configurations of $N \times N \times N$. $\delta$ is the ratio between the average distance for optimum network size and the average distance for cubic solution.

<table>
<thead>
<tr>
<th>N</th>
<th>URT</th>
<th>LRT: $\alpha=0.5$</th>
<th>LRT: $\alpha=1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma=0.5$</td>
<td>$\gamma=0.25$</td>
<td>$\gamma=0.5$</td>
</tr>
<tr>
<td></td>
<td>$N_{opt}$</td>
<td>$\delta$</td>
<td>$N_{opt}$</td>
</tr>
<tr>
<td>27</td>
<td>2x2x7</td>
<td>0.96</td>
<td>2x2x7</td>
</tr>
<tr>
<td>64</td>
<td>2x4x8</td>
<td>0.98</td>
<td>2x3x11</td>
</tr>
<tr>
<td>125</td>
<td>4x4x8</td>
<td>0.95</td>
<td>3x3x14</td>
</tr>
<tr>
<td>216</td>
<td>4x5x11</td>
<td>0.96</td>
<td>3x4x18</td>
</tr>
<tr>
<td>343</td>
<td>5x5x14</td>
<td>0.97</td>
<td>4x4x22</td>
</tr>
<tr>
<td>512</td>
<td>5x7x15</td>
<td>0.97</td>
<td>5x5x21</td>
</tr>
<tr>
<td>729</td>
<td>7x7x15</td>
<td>0.95</td>
<td>5x6x25</td>
</tr>
<tr>
<td>1000</td>
<td>7x8x18</td>
<td>0.95</td>
<td>6x6x28</td>
</tr>
</tbody>
</table>

In order to assess the validity of the topological solution given by our model for the uncongested and congested cases, we use RTL cycle-accurate simulations to measure average distance for networks under varying loads. Figure 87(a) plots the average distance in hops as the packet injection rate per cycle increases. Intuitively, in a network with uniformly clocked horizontal and vertical links, the optimum configuration for minimum communication distance under uniform traffic is always a symmetrical network. The unloaded average distance for three 64-node network topologies from our model is shown as a dashed base-line. This is the absolute minimum unloaded average latency achievable in that particular architecture. In this case a 64-node 3-D mesh will be best organized as a $4 \times 4 \times 4$ network for minimum average latency. Our model matches the baseline simulation result to within a 1% error when the injection rate is sufficiently low as to negate any contention issues.
As we increase the injection rate, contention becomes more prevalent and the average latency increases as a result of non-ideal routing conditions and link bandwidth limitations. Despite loading the networks to the point of saturation (where the network bandwidth is exceeded by the packet injection rate), Figure 87(a) demonstrates that the topology predicted as optimum by our traffic pattern-based average distance model is still valid for loaded networks as a network topology deemed as optimal in the unloaded case will be sub-optimal under congestion only if the lines intersect or cross. Figure 87(b) plots the same result for a local traffic pattern in a larger 256-node network.

1) **Hotspot Node Placement Optimization**

In a 3-D IC with TSVs providing the vertical switch-to-switch links, it is likely that the off-chip I/O will only be able to serve the outermost dies. Ball-Grid Array (BGA) or other area-array packages allow the I/O pads to be dispersed across the entire die surface, meaning that any outer node in the network could interface with off-chip devices. In a high-throughput multi-tier 3-D NoC, it is further likely that mesh nodes on the bottom layer, which have direct access to the BGA I/O will face higher traffic due to the on/off-chip communication. To model this effect we place nodes which attract high-traffic from the network (called hotspot nodes) on the bottom layer of a 3-D NoC in several configurations. The placement of the hotspot network nodes is crucial to the overall performance of the network.
the system (for example, placing hotspot nodes on the edge of a network will limit the surrounding link bandwidth due to the unused switch links).

The model we developed in section 0 can be used to estimate the average unloaded latency and the optimum placement (giving minimum average distance) of hotspots in a network. We place two hotspots on the bottom layer in three configurations as depicted in Figure 88(b): (HS1) hotspots are placed on the edge of the network in opposing corners, Figure 88(c) (HS2) hotspots are placed in opposing corners one node removed from the edge, and Figure 88(d) (HS3) the hotspots are diagonally adjacent at the centre of the network. The probability that any node in the network will send a packet to either of the hotspots is 80%, where the remaining 20% of the generated packets have a uniform probability to the other non-hotspot network nodes.

Figure 88 - Example of hotspot node placements HS1, HS2 and HS3 for the bottom layer of a 6×6×6 network

Similar to Figure 87, we have used our analytical model to predict the optimum placement of hotspot nodes, given by the minimum average latency for several network configurations and found that in the RTL simulations even under heavily congested networks with hotspot traffic, the optimum node placement to achieve the lowest average packet latency is consistent with the model prediction at zero load. Figure 89 plots the growth of latency with injection rate for the three hotspot placement schemes in a 7×7×7 network. Despite the minimal difference in unloaded latency provided by our model between the HS2 and HS3 placement schemes, the simulation results indicate no crossover points even at heavy loads. These results demonstrate the model’s usefulness in predicting network optimal topologies and node placement without computationally expensive packet-level simulations. Further to this, the model can quickly predict the optimum placement and configuration for hotspot traffic under
different vertical and horizontal clocking schemes. This efficient design space exploration is enabled by the availability of a closed form comparison metric.

To assess the fidelity of our model, we have conducted RTL simulation sweeps for unloaded and loaded networks for twelve different mesh configurations of sizes ranging from 8 to 1000 nodes each with the three hotspot placement schemes shown by Figure 88 and six packet injection rates of between 0.0001 and 0.8 packets-per-node-per-cycle (see Table 18 for a range of results). We consider the model to correctly predict the optimal network topology under loading if the latency curve with increasing injection rate corresponding to the zero-load optimal configuration is always below the latency curve corresponding to any other topology. Formally, fidelity can be defined as meeting the following condition:

\[
\overline{H}_{nwt_{1}}(IR) < \overline{H}_{nwt_{1}}(IR) \quad \text{for any injection rate (IR) where } \overline{H}_{nwt_{1}}(0) < \overline{H}_{nwt_{1}}(0) \text{being the average distance for the network configuration } nwt_{1} \text{ at zero load. The latter condition essentially defines } \overline{H}_{nwt_{1}} \text{ as the optimal network under zero load. If the data points cross-over at any point with a stable network or } \overline{H}_{nwt_{1}}(IR) > \overline{H}_{nwt_{1}}(IR), \text{ this is a failing of the model prediction.}
\]

For the simulations we have conducted under local, uniform and hotspot traffic, we found that the optimum placement predicted by the model is valid throughout all the network configurations when the injection rate is below the threshold of saturation; for the network.

These exhaustive findings report that although the model cannot predict the exact latency of a topology under congestion (which depends on low-level architectural features such as switch design, buffering and routing algorithms among others) it can accurately and repeatedly determine the optimum network topology and placement of hotspots under any network load. This assumption holds true for the simulations we have conducted with our traffic models, however we have yet to test it under other traffic scenarios such as burst mode or wormhole routing which may reduce the fidelity of the model.
Table 18 Simulation sweeps to assess the fidelity of the model. The optimum hotspot placement calculated by our model is shown to be valid for any traffic congestion in stable networks for over 100 data points.

<table>
<thead>
<tr>
<th>HSP</th>
<th>4x4x4</th>
<th>6x6x6</th>
<th>7x7x7</th>
<th>8x8x8</th>
<th>10x10x10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HS1</td>
<td>HS2</td>
<td>HS1</td>
<td>HS2</td>
<td>HS1</td>
</tr>
<tr>
<td>Model</td>
<td>4.96</td>
<td>6.9</td>
<td>6.12</td>
<td>6.06</td>
<td>4.06</td>
</tr>
<tr>
<td>0.0001</td>
<td>4.46</td>
<td>6.6</td>
<td>6.23</td>
<td>5.66</td>
<td>3.48</td>
</tr>
<tr>
<td>0.0005</td>
<td>4.46</td>
<td>6.6</td>
<td>6.23</td>
<td>5.66</td>
<td>3.48</td>
</tr>
<tr>
<td>0.0007</td>
<td>4.72</td>
<td>6.9</td>
<td>6.88</td>
<td>5.88</td>
<td>3.93</td>
</tr>
<tr>
<td>0.0009</td>
<td>4.82</td>
<td>7.5</td>
<td>8.08</td>
<td>6.38</td>
<td>4.25</td>
</tr>
</tbody>
</table>

[95]
5.2. Scalability of 3-D NoC Architectures

This section examines the performance and scalability of different communication topologies for 3-D Network-on-Chips using Through-Silicon-Vias for inter-die connectivity. VHDL simulations are conducted for two communication schemes based on a 7-port switch and a centrally arbitrated vertical bus using different traffic patterns. The scalability of the 3-D NoC is examined under both communication architectures and compared to 2D NoC structures in terms of throughput and latency in order to quantify the variation of network performance with number of nodes and derive key design guidelines.

Introduction

The performance bottleneck imposed by on-chip interconnects in the deep submicron regime of process technology has been widely documented [82], as have the benefits of standardization of on-chip communication [83]. Implementing a standardized communication architecture such as a packet-switched NoC for massively integrated multiprocessor systems provides an abstraction of the global interconnection link and can greatly reduce design effort, potentially at the cost of some area and possibly power and performance penalties. The suitability of the NoC as a communication architecture depends on the overall system; i.e. the number of autonomous functioning blocks, the degree of parallelism, and area and performance requirements dictate its usefulness. The precise quantification of the performance and overhead of the network is of paramount importance in making this decision. The potential of the 2-D NoC for interconnecting multi-processor systems has been widely researched, and most recently an 80 tile, 100M transistor system with 1.28 TFLOP peak performance has been demonstrated in a 65nm, 1V technology [84]. However a major new paradigm for continued Moore’s law integration is 3-D chip stacks based on a variety of vertical interconnection techniques [85], [86]. 3-D integration provides opportunities for cost reduction and yield improvement in integration of different technologies such as CMOS, DRAM and MEMS circuits through the ability to implement them over multiple die layers on the same chip. It can also reduce form factor in applications where size is critical, while effective heat dissipation and temperature control can be a challenge. To get the most benefit out of 3-D chip stacks in multiprocessor systems, the communication architecture has to support efficient and high throughput vertical communication. In this section we examine the scalability of the NoC for such systems.

We build on previous work published in the literature and investigate the scalability of the three NoC architectures of 2-D mesh, 3-D mesh (with switch connectivity between layers) and 3-D bus (with bus connectivity between layers) with respect to performance. This is accomplished by quantifying latency and throughput of the different architectures for various network sizes, under two representative traffic patterns, uniform random and local. The Nostrum network communication protocol used in this study is based on a buffer-less hot-potato routing algorithm [87], while the bus protocol utilizes a centrally arbitrated least-served-first priority scheme.

Cycle accurate RTL simulations are carried out to capture latency and throughput in terms of hops between nodes and hops per node per cycle respectively, for symmetric networks up to 1000 nodes. Various injection rates are used to study saturation points for the two traffic patterns. The main contribution of this work is in providing a novel and extensive analysis into the scalability of 3-D NoC architectures as the number of nodes and layers grows. This study quantifies performance to aid the design of the communication architecture of future massively integrated 3-D systems.

Link and Network Layer Protocols

The NoC protocol employed in this study is a hot-potato implementation with switch architecture as described in [87]. The routing strategy is based on non-minimal and load dependent deflection type packet switching, with adaptive per hop routing. A relative addressing scheme is implemented which simplifies the duplication of identical switches when network structures of varying sizes are designed.
The switches employed in all of the network configurations in this study are buffer-less, and directly connected with their associated resource with a 1:1 resource to switch ratio. A packet cannot be stored in a switch, and thus in each cycle the packets must be moved from switch to switch or deflected back to a resource. To reduce the complexity of the switches, deterministic routing is favoured over adaptive routing in buffered networks [88]. In bufferless networks, deflection routing is advocated [89] because it is possible to design fast and small switches with simplified control circuitry. The implemented switch uses 128-bit input and output channels for each switch-to-switch Physical Channel (PC) and resource to switch connection. For 2-D structures a 5-port switch is used as in Figure 89(a), while for 3-D structures a 7-port switch, as shown in Figure 89(b), is used where the inter layer connectivity is implemented with two simplex channels having the same switch-to-switch bit-width as the horizontal channels.

A TDMA vertical bus to provide connectivity between network layers is also designed as seen in Figure 89(c). It is a centrally arbitrated bus employing a least-served-first priority scheme for packet arbitration. The matrix arbiter circuit in [67] has been extended to deal with up to 10 layers. The bus has a 10 packet deep FIFO buffer at each input to prevent packet loss for the maximum network size. The bus runs on a clock 16 times faster than the network clock to provide serialization and de-serialization time, potentially circumventing area limitations for vertical TSV connectivity between layers. Hence the bus can accommodate a reduction in the bit width of the vertical channel up to a ratio of 1/16 and still receive and deliver a packet to a destination, on any layer, within one cycle of the network clock. It is connected to 6-port switches through a 128-bit Input/Output PC. In this study, every switch on a given layer is connected to a separate bus, and a single bus connects all switches that are vertically in-line. Thus, a 5x5x5 network will have 25 buses in total.

The packets are generated by each resource and injected into the network with an injection rate of up to 0.9 packets per node per cycle. For this study, a packet is considered to be one flit long. Each resource has a FIFO buffer to temporarily store packets if they cannot enter the network due to congestion. These packets are queued and re-injected into the network with a higher priority than newly generated packets. The implemented network does not drop packets. The packet headers generated by the resource contain final destination addresses and the switches make routing decisions on the fly based on this information.
Any NoC structure is comprised of switches located in general at the corners, edges, surface and centre of the physical structure. For example a 2-D switch has 4 links to other switches in 4 directions, namely, North, South, East and West. A 3-D switch has two additional, Up and Down links giving a total of 6 bi-directional links, whereas a bus architecture adds only one extra link to the switch, to give a total of 5 bi-directional links. There is also an additional bidirectional port from switch to resource in each case. For a given NoC structure, not all of the links can be used for routing packets. For example, only half of the switch-to-switch links in the corner of the network are connected, halving the available switch bandwidth due to the unconnected edge ports. The total number of connected links, in any NoC structure depends on the topology, dimension and size of the network. Figure 90 depicts the number of links as a function of network size in a mesh topology to highlight the differences between a 2-D 5-pot switch, a 3-D 7-port switch, and a 3-D bus architecture as defined in equations (42), (43) and (44) respectively.

For a 2-D \( n \times n \) network:
\[
L_{2D} = 4n(n - 1)
\]
(42)

For a 3-D \( n \times n \times n \) network:
\[
L_{3DN} = 6n^2(n - 1)
\]
(43)

For a 3-D Bus \( n \times n \times n \) network:
\[
L_{3DB} = 4n^2(n - 5) + 32
\]
(44)

All three curves for the different architectures show a sharp rise as network size increases and then begin to level off as the number of corner and edge nodes (i.e. nodes with some ports unconnected) become outweighed by the fully connected nodes. This trend of growth of links per node allows for higher throughput as the network size increases due to the increased bandwidth available to route a packet in the network. Additionally, it is clear from this figure that the 7-port switch has an advantage over the other topologies due to the increased number of switch-to-switch links per node available to route the traffic in the network for a given network size. The 3-D bus architecture has the least number of links due to the fact that all switches in a given vertical line share a common uni-directional link. The 7-port design benefits from separate PCs between every layer to handle multiple packets simultaneously within one cycle on any given vertical line from the bottom to the top layer. The bus in contrast can only deal with a single incoming or outgoing packet within a cycle of the network. This trend is evident in the simulation results, as the bus’ available bandwidth is necessarily limited by its design.

Traffic Patterns and Simulation Setup
The simulations were performed for a 2-D network with size \( n \times n \), for \( n = 3, 5, 8, 12, 15, 19, 23, 27 \) and \( 32 \) nodes, with packet injection rates, \( r \), varying from 0.1 to 0.9 packets per node per cycle in step increments of 0.1. For the 3-D 7-port switch and 3-D bus architectures, the network size was \( n \times n \times n \) for \( n = 2, 3, 4, 5, 6, 7, 8, 9 \) and 10 with the
same injection rate as in the 2-D case. The number of simulated nodes in the 2-D mesh follows the 3-D setup to provide a fair comparison. The data sample used is extracted following the warm-up phase of the network and preceding the cool-down phase to ensure reliable results. For example, Figure 91 shows the average latency growth for a 5×5×5 3-D mesh network with an injection rate of 0.5. The warm-up phase occupies the first 200 clock cycles. After 500 clock cycles the network is fully stable until the simulation ends. In our experiments samples were obtained after the network reached stability but for certain combinations of network size and injection rate, the network never becomes stable. These cases are identified in the discussion.

![Figure 91 – Variation of average normalised latency for a 5×5×5 mesh with 0.5 injection rate under the local traffic model, illustrating the warming up phase](image)

The metrics used to quantify the performance of the network in each case are raw latency, normalized latency, and throughput, defined in (45), (46) and (48) respectively.

\[ T_{\text{Raw}} = \frac{C_{\text{final}} - C_{\text{init}}}{HC} \] (45)

The raw latency, \( T_{\text{Raw}} \), is the distance traveled by a packet from the source to the destination address in terms of hop counts, denoted by \( HC \). \( C_{\text{final}} \) and \( C_{\text{init}} \) represent the final and initial clock cycles respectively. When the network is at zero-load, the raw latency is equivalent to the minimum distance. The normalized latency, \( T_{\text{norm}} \), is the ratio of the raw latency (i.e. the actual distance travelled by a packet) to the minimum distance with zero-load, \( T_{\text{zero,load}} \), defined as:

\[ T_{\text{norm}} = \frac{T_{\text{Raw}}}{T_{\text{zero,load}}} \] (46)

The average normalized latency is the mean of the normalized latencies for the collected samples defined as:

\[ T_{\text{norm,avg}} = \text{mean}(T_{\text{norm}}) \] (47)

In each cycle, packets arrive at all nodes at a rate based on the injection rate and the congestion level of the network. The throughput per node per cycle, \( \lambda \), is defined in (48), where \( P_{\text{total}} \) is the total number of packets received over the simulated range, \( N \) is the number of nodes in the network and \( C \) is the number of cycles in the sampling region.

\[ \lambda = \frac{P_{\text{total}}}{NC} \] (48)

The generated traffic patterns attempt to load the network in a realistic manner as would be encountered on a multi-processor SoC. A hybrid of two traffic patterns are employed for all cases, uniform random traffic (URT) and local traffic models. The URT model stipulates that each node in the network is equally likely to become the destination address of any packet emitted from a resource. From a design perspective, it is customary to place frequently communicating resources close to each other to maximize efficiency. This increases the performance of the
communication in terms of power, timing and resource management. The local traffic pattern generated for each resource replicates this localized communication behaviour. The destination address is assigned randomly and then a localized probability formula is applied to the address, as shown in Equation (49) and (50). This formula increases the probability that the final destination of the packet will be local to the sending resource rather than farther away. This follows the principle that the resources will be arranged within the network such that their nearest neighbours are devices with which they communicate with most frequently.

Figure 92 – Load per channel for 2-D, 3-D mesh and 3-D Bus architectures
The localized probability for local traffic patterns is defined in (49), where D is the maximum distance in the network and (9) is a normalizing factor guaranteeing that the sum of all probabilities is 1:

\[ P(d) = \frac{1}{A(D)2^d} \] (49)

\[ A(D) = \sum \frac{1}{2^d} \] (50)

In order to compare traffic patterns, a fully URT pattern without localization was also implemented. This pattern means that any address within a given network is equally likely to be assigned to each packet being generated. To maintain stability in the network, the injection rate is lowered as the number of nodes is increased. The injection rate as a function of size is determined by Equation (56).

The load, \( \gamma \), that each node puts on the network is shown in (51), where \( r \) is injection rate from 0.1 to 0.9 and HC is the average hop count in the network:

\[ \gamma_{\text{ntwrk}} = r \times HC \] (51)

We assume that a packet loads the network with each hop by 1, because it occupies 1 link per hop. Equation (51) shows that the load depends on the injection rate of each node and on the average distance of each packet. HC grows with the size of the network. In \( k \)-dimensional meshes when \( n \) is even, we have

\[ HC = \frac{k \times n}{3} \] (52)

In 3-D meshes HC is coincidentally \( HC = 3n \times 3 = n \), whereas for a 2-D mesh \( HC = 2n \times 3 \). Hence the total load in the network is

\[ \gamma_{\text{ntwrk},2D} = 2r \times \frac{n^3}{3} \] (53)

\[ \gamma_{\text{ntwrk},3D} = r \times n^4 \] (54)

The network capacity can be expressed as the number of links for 2-D and 3-D meshes as given in Equation (42) and (43) respectively. As the network grows, the network capacity grows and the network load grows. However, the load under uniform random traffic grows faster than the network capacity. Consequently, the injection rate has to decrease as the network grows. The load per channel, \( \gamma_{\text{Chnl}} \), is defined as:

\[ \gamma_{\text{Chnl},3D} = \frac{\gamma_{\text{ntwrk},3D}}{L_{3d}} \] (55)

The load per channel for 2-D-mesh and 3-D mesh and 3-D bus architectures for varying injection rates is shown in Figure 92(a), (b) and (c) respectively. Substituting equations (2), (13), into (14) leads to

\[ r = \frac{6 \times \gamma_{\text{Chnl},3D} \times (n - 1)}{n^2} \] (56)

Now in order to maintain stability, the load per channel, \( \gamma_{\text{Chnl},3D} \), is set constant. For example if \( \gamma_{\text{Chnl}} \) is set to 0.5, i.e. a constant load of 0.5 packets per channel, the injection rates, \( r \), corresponding to various network sizes are given in Table 19 Injection rate for channel load of 0.5 in 3-D nxn.xn mesh.

**Results**

The simulation results produced in this study highlight the effect of increasing the number of nodes and injection rates for different NoC topologies on network performance. The results quantify the normalised and raw latency, and throughput versus injection rate for a 2-D mesh, 3-D 7-port mesh, and a 3-D 6-port mesh with vertical bus connectivity for localized and uniform traffic. Figure 93(a), (b) and (c) are plots of normalized latency, throughput and raw latency, respectively, versus network size for injection rates ranging from 0.1 to 0.9 in a 2-D mesh. Figure 93 (d) is a plot of throughput versus injection rate for different network sizes. These localized 2-D simulations show that as network size increases, the normalized latency quickly reaches a saturation point whereas the throughput matches the injection rate for stable input conditions. For injection rates of 0.1 to 0.4, once the size has increased beyond 225 nodes, scaling the network size up to 1024 nodes incurs no significant performance penalty in terms of latency or throughput due to the localized traffic pattern. Figure 93(d) shows that the 2-D network investigated in
this study fails to match injection rate with throughput above 0.5 packets per node per cycle; i.e. the network becomes congested and unstable. The instability of injection rates above 0.5 was confirmed by an examination of the transmitter output buffers from each resource, where the network becomes unstable when the buffer fills to capacity. Figure 93 (d) clearly shows the effect of an overly congested network on the throughput as the number of nodes grows. These 2-D simulations show that under a localized traffic model, a bufferless 2-D mesh network can maintain relative stability, and provide acceptable performance for up to 1000 nodes with injection rates below 0.5.

Table 19 Injection rate for channel load of 0.5 in 3-D n x n x n mesh

<table>
<thead>
<tr>
<th>n</th>
<th>N</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>0.75</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>0.67</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>0.56</td>
</tr>
<tr>
<td>5</td>
<td>125</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>216</td>
<td>0.42</td>
</tr>
<tr>
<td>7</td>
<td>343</td>
<td>0.37</td>
</tr>
<tr>
<td>8</td>
<td>512</td>
<td>0.33</td>
</tr>
<tr>
<td>9</td>
<td>729</td>
<td>0.30</td>
</tr>
<tr>
<td>10</td>
<td>1024</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Figure 93 – 2-D mesh performance
Figure 94(a), (b) and (c) show the normalized latency, throughput and raw latency against network size for varying injection rates in a 7-port 3-D NoC with local traffic. The 3-D 7-port mesh copes with a higher level of traffic better than the 2-D mesh. Where the 2-D network becomes congested at an injection rate of 0.5 as the number of nodes grows, the 3-D 7-port maintains throughput and a constant latency for all network sizes. The increased number of links per node in the 7-port 3-D design allows the network to remain stable for a wider range of network sizes and injection rates. This topology manages to match throughput to injection rate up to 0.6 packets per node per cycle. This result is interesting as it demonstrates the capability of a 3-D 7-port mesh to handle higher injection rates than a 2-D mesh for a given number of nodes. The plot in 96(a) however shows that the average normalized latency increases sharply with injection rates beyond 0.6 as the packets spend more time in output buffers at their origin resource due to congestion. Figure 94 (d) is interesting because it correlates the injection rate to the throughput per node per cycle directly, so the saturation point for each network size can be determined. This understanding is paramount to designers as it determines the boundary conditions for maximum performance for different sized networks.

Figure 94 – 3-D mesh performance

Figure 95(a), (b) and (c) plot the average latency, throughput, and raw latency against network size, respectively for the 3-D mesh network with the vertical bus. The 3-D bus has a markedly worse performance in comparison with the 7-port 3-D topology and the 2-D mesh. Figure 95 (a) shows that the average normalized latency values are lower than the corresponding values for both the 2-D and 3-D cases when the network has a lower injection rate. This is largely due to the bus only requiring 1 hop for a packet to reach its destination on any vertical layer. The bus performs well below 200 nodes for low injection rates in terms of latency; however, the throughput plotted in Figure 95(b) begins to drop for injection rates of 0.2 and above for any network size greater than 3x3x3. This is due to...
packets being deflected horizontally around the network, as contention for the vertical bus link increases with the number of layers. This effect increases as the network size grows as the number of users requesting access to the bus grows in proportion to the number of layers. Finally Figure 95(d) clearly shows that the bus struggles to match throughput to injection rates of over 0.3. The local traffic model is not friendly to the bus topology either, as the advantage the bus provides in sending a packet in one vertical hop to any layer is not fully utilized. Although the bus appears to be the worst of the three architectures under localized conditions, it may have advantages in other traffic conditions, such as when a stack of memory lies above a processor, where the bus will provide equal access time to any layer of memory. The results in Figure 96(a), (b) and (c) relate to the uniform random traffic model showing plots of the average normalized latency, raw latency and normalized throughput versus increasing network size respectively. Figure 96 (a) shows that the bus has the lowest latency and that a conventional 2-D network has the highest, with the 3-D 7-port lying in-between. Figure 8(c) shows that the 3-D 7-port switch behaves reasonably similar to the localized pattern in that it reaches a saturation point quickly for throughput and further increases in network size add little performance penalty. However the 2-D mesh and 3-D bus architectures are the worst for this traffic pattern. They both have decreasing throughput with increasing network size. In the 2-D case, the packet must travel over a long distance, and the network easily becomes congested. The bus again suffers from the reduced link bandwidth per node when it has to deal with traffic patterns with injection rates greater than 0.1. However, different from the local traffic simulations, the bus has outperformed the 2-D mesh in terms of throughput and latency. The uniform random traffic allows the bus to utilize its ability to transport a packet any distance in just one hop, and this shows through in these results. These figures appear to highlight that under uniform random patterns, the 7-port switch is the best option for large sized networks in terms of packet throughput, but trade-offs such as area overhead and power may render the bus a viable alternative, especially when the vertical links are limited, and a scarce resource.
Figure 96 – 3-D URT performance
5.3. Performance Analysis of multi-clock 3-D NoCs

The physical performance of a 3-Dimensional Network-on-Chip (NoC) mesh architecture employing Through Silicon Vias (TSV) for vertical connectivity is investigated with a cycle-accurate RTL simulator. The physical latency and area impact of TSVs, switches, and the on-chip interconnect is evaluated to extract the maximum signalling speeds through the horizontal and vertical network links. The relatively low parasitics of TSVs compared to the on-chip 2-D interconnect allow for higher signalling speeds between chip layers. The system-level impact on overall network performance as a result of clocking vertical packets at a higher rate through the TSV interconnect is simulated and reported.

Introduction

This section examines the performance of a 7-port 3-D NoC mesh employing higher clock speeds for vertical (layer-to-layer) communication links for switch to switch connections. Physical properties and models of the TSVs, on-chip wires, and synthesized switches are studied to understand the physical limitations in clocking and provide appropriate boundary conditions for the simulations. Extensive simulations are then carried out to explore the performance benefits of clocking the vertical link communication at a higher rate than the horizontal links. Cycle accurate RTL-level simulations are conducted with Uniform Random Traffic (URT) on a 7-port switch using the Nostrum network communication protocol based on a bufferless hot-potato routing algorithm [88] for symmetrical \( n \times n \) configurations. The performance of the NoC in terms of throughput, latency and area is extracted from the simulations in order to quantify the variation of network performance with different vertical and horizontal clock speeds for a varying number of network nodes to derive key design guidelines.

The main contribution of this work is in providing a novel and extensive analysis into the performance of 3-D NoC architectures with a realistic application of the physical limitations of 3-D integrated circuits employing TSVs. This study quantifies performance to aid the design of the communication architecture of future massively integrated 3-D systems.

Physical Layer and Design Trade-offs

The technology that allows the highest density of vertical interconnects and is the most promising for tackling interconnect related communication woes in very large designs is based on through-silicon or through-hole vias that are fabricated either as part of the front-end wafer processing (front-end-of-line: FEOL) [90] or a back-end process (back-end-of-line: BEOL) at the die or wafer level. A pitch on the order of 0.14 \( \mu m \) is currently state-of-the-art in TSV technology [28]. Physical models for the equivalent circuit representation of TSVs are reported in [9], which recommends a lumped capacitive model with inter-via coupling for representative TSV lengths. The resistance of the TSV is of the order of m\( \Omega \), which results in a very low signal transmission latency compared to on-chip wires which run for a considerably longer length. The authors of [91] come to a similar conclusion. Because of this relatively low resistance, the effect of capacitive crosstalk within a TSV bundle can be overcome by a sufficiently high driver strength, which leads to the result that the highest bandwidth can actually be obtained by having the highest possible TSV density [92], unlike for the on-chip case, where the highly resistive wires dictate an optimal point that does not correspond to the highest density [93]. Even though it is desirable from a performance point of view to have the highest density, there are associated cost penalties. TSVs with different material composites and widely varying densities have been reported in the literature, from structures with pitches over 100 \( \mu m \) with a ‘U’ filled cross-section to 0.14 \( \mu m \) pitch copper filled structures [28]. The higher densities are usually only enabled by costly BEOL processing. It is important to recognize that the via density is governed by cost constraints (dictated by poor yield for the more advanced processes) as well as performance requirements. Another important factor is that TSVs will take away logic area, and reduce the area utilization efficiency. As an example, the die, resource and switch size for an aggressive design in a nanometre scale CMOS technology is approximately 20 mm, 2 mm and 1 mm on a side respectively [95]. Based on some reference designs we have synthesized, we estimate the sizes of a simpler 5-port 2-D and 7-port 3-D switch to be 0.6 mm and 0.7 mm on a side respectively in 180nm UMC. The area utilization is quantified in Figure 97, with the exclusively 2-D NoC serving as a baseline. Under the assumption that every switch has vertical connections, the extra area consumed by the TSV bundle may present an unacceptably high overhead for certain bit-width and pitch combinations. Hence the main design choices related to implementation in 3-D integration revolve around trade-offs associated with TSV density (i.e. parallelism), area
overhead and cost. Potential solutions to minimize cost and area overhead without imposing bandwidth penalties are to reduce the number of switches with vertical connectivity, and also to reduce the bundle size, with parallel-to-serial and serial-to-parallel conversion when a packet crosses a layer. This is made feasible by the fact that the TSVs have considerably better electrical performance than the horizontal inter-switch link, and support a higher signalling speed.

Figure 97 – Area utilization efficiency on a single layer with TSV bundle size for different pitches

1) **Switch Properties**

The aim of this section is to obtain representative performance limitations of the physical interconnect and hardware of a 3-D NoC which will provide a performance indicator and reality check for the system-level simulations conducted in this study. A maximum die size of 2 cm×2 cm using a 32nm CMOS process technology is considered for all physical aspects of this investigation. To determine the switch, interconnect and resource areas, and subsequent switch-switch wire lengths, a maximum network size of 32×32 was selected. This represents the maximum number of nodes simulated for 2-D and 3-D networks within this exploration. This upper limit on the number of nodes allows for a resource, switch, and interconnect to occupy equally divided tiles of 624 μm×624 μm on a single die. This area allocated for each tile in the network is kept constant, but the number of nodes reduces along with the footprint of the die. The area allocated for the 2-D and 3-D nodes is identical to provide a fair comparison. As this study intends on mapping the physical capabilities of the communication hardware of the network, the resources are not considered in depth.
The switch analysis establishes the physical length of the 2-D network links and the capability of a bufferless switch to be pipelined sufficiently enough to route packets through the vertical TSVs at higher frequencies. The switches used in this study were designed for functionality thus were not optimised for speed or layout. The area and speed of both the 2-D and 3-D switches can be considerably improved with customized layout tools and pipelining stages, which has been widely documented in academia and industry, but falls outside the scope of this investigation. Using the methodology applied in [93] we have estimated the area and interconnect density for a 2-D 5-port and 3-D 7-port switch in 32 nm technology. Analysing the switch speeds and sizes establishes where the bottleneck in the network performance lies. The single-stage 5-port 2-D switch is reported to occupy an area of 600 μm × 600 μm and consist of 40,029 gates in 0.18μm technology from a synthesized VHDL model. By applying equation (57) [95], the size of the switch in 32 nm technology was estimated to occupy an area of 106 μm × 106 μm, where $A$ denotes the area, the feature size and $\alpha_s$ a constant to account for integration losses or gains in scaling. The layout of a tile is show in Figure 98. The 2-D switch occupies roughly 3% of the area allocated for each tile.

$$A_{\text{new}} = \frac{A_{\text{old}}}{\left(\frac{Y_{\text{old}}}{Y_{\text{new}}}ight)^2 \alpha_s}$$

The gate depth from the synthesis report was used to find the critical path delay of the switch. Using the fan out of 4 (FO4) delay allows the maximum operating frequency of the design to be roughly estimated. In this case, the FO4 delay for 32 nm technology is approximately 16 ps, the critical path for our synthesized 2-D switch is approximately 3.52 ns, which results in a maximum operating frequency of 284 MHz at 32 nm. As the implementation is currently a single stage design, it is assumed that the switch can be pipelined to improve upon the critical path delay and maximize the clock frequency to bring the switch designs into the GHz realm. The single-stage 7-port 3-D switch design was also synthesized for 0.18μm technology and reported to occupy an area of 700 μm × 700 μm, consisting of 50K gates. Applying the same methodology as in the 2-D case we have estimated the switch area to be 124 μm × 124 μm and the operating frequency as 250 MHz at 32 nm.

**Network Links**

The resistance and inductance of a wide range of TSV dimensions is negligible [96] compared to the parasitics of an appropriately sized driver and on-chip global wires, the delay through the TSV is much smaller than signals on each 2-D network layer. It is therefore feasible to attain faster signalling speeds through the vertical interconnect than the 2-D horizontal communication grid due to these physical properties. The physical parameters of future 32nm technology are obtained from the International Technology Roadmap for Semiconductors (ITRS) 2009 and rail-to-rail CMOS signalling is considered. The insertion of repeaters is beneficial when the line delay is at least 7 times greater than the buffer delay [9]. Given the allocated size of each functioning block in the network and the results from the estimated switch areas, a maximum switch to switch wire length of 518 μm is obtained. In order to
determine appropriate driver sizes and the number of repeaters required, Bakoglu’s equation (58) derived in [40] was used.

\[ t_{2D} = k \left( \frac{0.7 R_d}{h} \left( C_s + hC_d + 4.4 \frac{C_c}{k} \right) + \frac{R_w}{k} \left( 0.4 \frac{C_s}{k} + 1.51 \frac{C_c}{k} + 0.7 hC_d \right) \right) \] (58)

Using projected parasitic values for 32nm technology on a 518μm length of semi-global wire, as detailed in Table 20, the ideal driver size and repeater count was calculated as H=115 and k=3 respectively for the worst case coupled switching scenario. This results in an analytically determined 50% rise time of 58.6ps. This calculated figure was confirmed to be accurate with Spectre simulations for a coupled wire with the same parasitic properties.

\[ t_{3D} = 0.69 R_d \left( C_s + K_1 C_{lat} + K_2 C_{diag} \right) \] (59)

A 16x16 bundle of TSVs serves as the vertical connectivity for each switch between chip layers. The parasitic elements of a TSV, shown in Table 20, were calculated from the compact models developed in [4] to determine R, L, C values for a 10μm diameter TSV with a pitch of 30 μm between adjacent TSVs. Circuit simulations in Spectre of a 3x3 bundle under the worst case switching scenario were compared with the delay model presented in [96] to evaluate the rise time information of TSVs in a bundle. The delay in an n×n bundle of TSVs can be quickly determined with reasonable accuracy in this fashion. A 50% rise time of 4.08 ps was predicted by equation (59) with K1 and K2 as 9.0 and 10.6 respectively. Running a spice simulation with the same values produced a 5.77 ps rise time, which confirms the accuracy of the model. When compared to the horizontal link delay, the vertical packets arrive some 50ps prior to the horizontal packets. This allows enough room for a vertical clock of up to 8 times faster than the horizontal links.

### Table 20 Parasitic Values for 2-D on-chip wires, drivers and TSVs

<table>
<thead>
<tr>
<th>On-chip parameters</th>
<th>TSV Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Width = 70 nm</td>
<td>Diameter = 10 μm</td>
</tr>
<tr>
<td>Length = 0.518 mm</td>
<td>TSV Length = 20 μm</td>
</tr>
<tr>
<td>C_s=20 fF/mm</td>
<td>C_s=7.37 fF</td>
</tr>
<tr>
<td>C_d=0.10 fF</td>
<td>C_{lat}=1.31 fF</td>
</tr>
<tr>
<td>C_c=61 fF/mm</td>
<td>C_{diag}=0.45 fF</td>
</tr>
<tr>
<td>R_w=1870 Ω/mm</td>
<td>R_{tsv}=4.5 Ω</td>
</tr>
<tr>
<td>R_d=30.48 kΩ</td>
<td>L_{tsv}=6.18pH</td>
</tr>
</tbody>
</table>

### Link and Network Layer Protocols

The RTL cycle-accurate simulator has been designed to enable vertical packets entering or exiting the vertical ports (Up/Down) of the 7-port switch to be clocked integer multiples faster or slower than the horizontal ports (N,E,S,W,R). For this study, the vertical packets are clocked up to four times for every one cycle of the horizontal clock to observe the effect of faster vertical packet routing. The switch runs on the faster vertical clock and the routing algorithm is not modified in any way to take advantage of the speed of the vertical links. Packets may only be injected into the network from the resources on the horizontal clock edges. The packets are generated by each resource and injected into the network with an injection rate of up to 0.9 packets per node per cycle. For this study, a packet is considered to be one flit long. The NoC protocol, routing algorithm and basic switch design are further elaborated in section 5.2.

### Traffic Pattern and Simulation Setup

The data samples are extracted from the output files following the warm-up phase of the network and preceding the cool-down phase to ensure reliable results. For example, Figure 91 shows the average latency growth for a 5x5x5 3-D mesh network with an injection rate of 0.5. The warm-up phase occupies the first 200 clock cycles. After 500 clock cycles the network is fully stable until the simulation ends. In these experiments samples were obtained after the network reached stability but for certain combinations of network size and injection rate, the network never becomes stable. Simulations were performed for an n×n×n 3-D network for n = 2, 4, 6, 8 and 10 for a symmetrical vertical and horizontal clock, and a vertical clock with twice the frequency. Injection rates for these cases were simulated at 0.1 and 0.2 packets per node per horizontal cycle. Additionally, simulations were performed for a 10x10x10 network with a vertical clock of 2, 3 and 4 times the speed of the horizontal network links. The injection rate for these cases was fixed at 0.3 packets per node per horizontal cycle.

[109]
The metrics used to quantify the performance of the network in each case are raw latency, normalized latency, and throughput, defined in (45), (46) and (48) respectively.

The generated traffic pattern attempts to load the network to provide a clear indicator of the impact of altering the clock frequencies of the vertical network links. A Uniform Random Traffic pattern is employed in all of the simulations conducted in this study. The URT model stipulates that each node in the network is equally likely to become the destination address of any packet emitted from a resource. Random addresses are generated by each resource according to the current network size and configuration and then assigned to outgoing packets given a definable injection rate. Although capable of simulating localized traffic patterns, the simulations were designed to characterise the performance variation of the vertical clocking schemes without complicating the design space, thus the localization protocol was disabled.

**Results**

The results detail the performance contrast in clocking the vertical TSV links at a higher rate than the on-chip horizontal links. Figure 99(a) and (b) plot the average raw latency as the number of nodes and layers is increased for injection rates of 0.1 and 0.2 packets per node per horizontal cycle respectively. In these cases the vertical links are fixed at twice the speed of the horizontal links and compared against a uniformly clocked 3-D and 2-D network of comparable sizes. The 2-D results were omitted from Figure 99(b) for clarity, as its performance was degraded. For smaller networks it is clear that the increased vertical links do not have a large effect on the delivery time of the packets in the 3-D case. As the data packets are equally likely to head in any given direction and the vertical links only represent two out of the possible seven output ports of the switch, the likelihood of a packet going vertically rather than horizontally is less in a small network due to the physical structure of the switch and limited network layers. In these cases, the effect of clocking twice as fast vertically is less prevalent but still advantageous in terms of latency.

As the network becomes larger, consisting of more chip layers, the raw latency curves begin to separate, with the faster clocked network proving less latent than the uniformly clocked network by 9.5 and 10% for injection rates of 0.1 and 0.2 respectively. It is clear that the 2-D network performs comparably worse to the 3-D variants and the latency grows exponentially as the number of nodes in the network grows. The throughput per node per cycle for both 3-D cases remains constant, matching the injection rate. However, in the 2-D case, above 225 nodes, the throughput declines. It is clear that the faster clocked network performs best in all cases in these simulation results.
Figure 99 – Raw latency growth as the number of network nodes increases for injection rates of 0.1(a) and 0.2(b) for vertical links twice as fast as horizontal links.

Figure 100 – Raw Latency against increasing vertical clock speeds for a 10×10×10 network with an injection rate of 0.3.
In the next simulation set, the network size, horizontal clock speed and injection rate is fixed, whilst the speed of the vertical links is increased by integer multiples. This simulation shows the effect of higher clock rates on a large network. Figure 100 is a plot of the average raw latency against increasing clock speeds for the vertical communication links. It can be seen from this figure that there is a distinct improvement in latency when the clock rates are increased for a large network under URT. We see an improvement of roughly 70% in latency by increasing the clock rate four times. This demonstrates a significant savings in packet delivery time for a large chip stack. Eventually there will be a limit to the amount of savings provided by clocking the vertical links faster as the speed of the horizontal communication grid will eventually prove a bottleneck to the network. This can be seen by the levelling off of latency with higher clock speeds in the curve.

Discussion
This investigation quantifies the effect of utilising faster vertical links on the overall performance of a bufferless 3-D NoC mesh. The physical analysis of the switch, interconnect and TSVs provides realistic boundary conditions for the system-level simulations, demonstrating that signalling speeds through the vertical network links can be achieved at up to eight times the speed of the horizontal communication wires. The VHDL model of a 7-port 3-D switch presented in Part II was modified to enable clocking of vertical packets at integer multiples of the horizontal clock to exploit the faster TSVs. A URT traffic pattern was employed in a large scale 3-D NoC and the results show that packets arrive significantly faster than a NoC with a uniform clock. The savings in average raw latency were quantified in the results as network size scales.

Further performance improvements are expected when the switch architecture and routing algorithm is modified to favour vertical routing in times of deflection or heavy traffic at the desired destination port. The ability to route packets vertically with less latency penalty can significantly improve the communication performance for multi-layer chip stacks either in a bus configuration or a NoC. Additionally, the faster clocking of TSVs can be used to reduce the area footprint of a TSV bundle by providing fast serialization and de-serialization hardware to lessen the impact of the relatively large TSV structures.

5.4. Conclusions
This chapter models the behaviour and performance of future massively parallel computation systems using the network-on-chip infrastructure from theoretical models for average distance, to cycle-accurate simulations, to the physical performance of 3-D designs. In section 5.1, the correct model for average distance in any dimension or radix network was derived and used to optimise the design of 2-D networks based on core partitioning, hot spot location and traffic patterns. Our study shows that the model for average distance exhibits 100% fidelity when used as a metric to optimise the configuration of mesh-based networks.

In section 5.2, the scalability of the Network-on-Chip architecture was addressed by developing cycle-accurate simulation models of 2-D, 3-D and hybrid-bus switches under local and uniform random traffic. We demonstrate that the 7-port 3-D switch exhibits the best performance in terms of throughput and latency under heavy load, while the 5-port 2-D switch reaches saturation much earlier. The 6-port bus switch can reduce the latency of packets significantly, but cannot provide sufficient throughput under load due to the shared vertical links.

Finally, in section 5.3, the network architecture and switch hardware discussed in Part II was mapped to the physical domain using 2-D and 3-D wire models and scaled synthesis results in a 32 nm CMOS process. Using these results a multi-clock switch was designed to exploit the faster vertical communication over TSVs (up to 8× faster than the 2-D links in our study) and its relative performance demonstrated. By exploiting the speed of the vertical links, the average latency in a mesh can be significantly reduced and higher throughputs are possible, at the cost of additional hardware overhead.
6. System-Level Modelling of Silicon Systems

Hierarchical models from physical to system-level are proposed for architectural exploration of high-performance silicon systems to quantify the performance and cost trade-offs for 2-D and 3-D implementations. We show that 3-D systems can reduce interconnect delay and energy by up to an order of magnitude over 2-D, with an increase of 20-30% in performance per-watt for every doubling of stack height. Contrary to previous analysis, the improved energy efficiency is achievable at a favourable cost. The models are packaged as a standalone tool and can provide fast estimation of coarse grain performance and cost limitations for a variety of processing systems to be used at the early chip-planning phase of the design cycle.

6.1. Introduction

As design complexity increases with each generation of processors, more emphasis must be spent on the planning stages of the product design. For any new technology or architecture to become viable, it must conclusively demonstrate significant value at an acceptable risk and cost. To quantify the trade-offs between the many design choices available in the early chip planning phase of a silicon-based processor, we have created hierarchical models to extract the performance limitations and the computational efficiency of 2-D and 3-D ICs under realistic cost and physical constraints. We base our models for computation on the underlying physical tenets of the three fundamental operations of any digital processing system: computation (logic), storage (memory), and communication (interconnect). Using these models and the accompanying tool the following features in single-chip processor design can be examined:

- **Scaling:** We model planar wires, devices, logic and memory from 180 nm down to 17 nm to extract the performance per Watt for 2-D and 3-D systems up to 16 layers as technologies scale. We also consider the effect of TSVs scaling from 20 µm down to under 1 µm. As an example we show that scaling the 65 nm Intel 80 core processor to 17 nm will increase the computational efficiency by 80%.

- **Architecture:** System-level design choices are captured by parameters in our model for computation including the layer partitioning, on-chip memory distribution, interconnect sharing ratio, memory technology (DRAM, Flash, SRAM), memory locality, data width and the ratio of on to off-chip transactions. We show that by partitioning the 65 nm Intel 80 architecture over four 3-D layers, an equal performance-per-Watt can be achieved as the same system developed in 45 nm.

- **System Constraints:** Specific systems can be modelled under the physical constraints of thermal, power, area, and frequency to aid in the optimization of a given topology for maximum performance under set of realistic application constraints. Using our models, we quantify the optimal number of layers for a processing system under thermal and power budgets dictated by common computing applications.

- **Yield and Cost:** We model and compare the variable and fixed costs for 2-D and 3-D processor architectures as a function of feature size and provide the necessary tools to assess the system costs. We conclude that for large silicon systems a 3-D implementation can recover total cost at lower volumes than a 2-D implementation despite the added expense of stacking.

The rest of this chapter is organized as follows: we first discuss related work in section 6.2. In section 6.3 we introduce and explain the derivation of our performance model including its physical basis. We then in section 6.6 explore the design space by varying architectural parameters to demonstrate the flexibility of our model to determine performance limitations of both 2-D and 3-D processor topologies. In section 6.7 we discuss the cost of different implementations. Section 6.8 discusses the applications of our models for a concrete processing system and finally we end with our conclusions in section 6.9.

6.2. Related Work

There are a number of available software suites that can provide early chip planning estimates, where many firms have in-house early-estimation tools. Cadence InCyte [103] for example allows a global view of a system including
area, cost, performance and power using IP from manufacturers but is limited to 2-D designs and is restricted to specific applications. CACTI [104] provides an integrated environment for timing, power and area modelling, but is restricted to memory modelling. The authors of [35] offer predictive models of devices and 2-D wires, but no system-level analysis tool. 3-D ICE [63] and HotSpot [64] both offer thermal estimation of 3-D chip stacks but are not intended to model the underlying hardware features which generate the power profile. Our contribution is in providing a set of hierarchical models to analyse general 2-D and 3-D systems under physical, performance, and cost constraints before detailed knowledge of the design is known.

The potential of stacked 3-D processors has been investigated extensively by academia and industry. The authors of [94] provide an excellent review of 3-D stacking technology and its benefits to high-end processors. In [106] the authors conclude that stacking DRAM on a single processor is a viable solution to overcome the increasing performance gap between memory and logic. The authors of [107] demonstrate that a 12-core stacked IC with no L2 cache outperforms an 8-core 2-D system with a large on-chip L2 cache by about 14% while consuming 55% less power. A similar performance result for stacked DRAM and logic was arrived at in [108]. However, none of these works present a general model to encapsulate the global physical architecture to enable design space exploration. In [110] the authors stipulated that the performance advantage of 3-D over 2-D ICs would shrink as feature size decreased below 65 nm, however recent work [111] analysing the performance of a 16-core stacked IC and our own study has shown that in fact the advantage of 3-D ICs grows in lower technology nodes. An outlook of processor power and cooling strategies for 2-D and 3-D ICs as feature size scales is given in [112]. A number of thermal management strategies such as thermal via floorplanning, interposers and microchannel liquid cooling between die layers have been proposed to mitigate increased power density in 3-D ICs. In [113] cooling strategies are explored for multi-tier 3-D CPUs and test structures fabricated to compare parallel plate, microchannel, and pin fin structures in multiple configurations. Costing of a multi-tier 3-D processor is discussed in [114] where the authors show that different partitioning strategies can result in lower cost for 3-D ICs. As a proof of concept, a 64-core stacked IC with 47,940 TSVs was fabricated to provide 63 GB/s of memory bandwidth at 277 MHz under a 47°C thermal window [115].

6.3. Modelling Performance

To introduce our contribution we look back to a metric first introduced by T. Claasen in [116] which describes the absolute maximum computational capability of a silicon system with a term called the intrinsic computational efficiency (ICE). The ICE of a system is calculated by filling the entire silicon area with the most fundamental computational circuitry, in this case 32-bit adders. Figure 101 shows Claasen’s original ICE projection as feature size decreased and our own projection extending his work down to 17 nm. Two multi-core high-throughput processors are shown to exemplify the difference between a realistic system and the ideal upper bound of the ICE metric, which does not account for the I/O, interconnect, memory or general purpose control logic.
Our model attempts to build upon the ICE metric by factoring into account the expense of interconnect, memory, and I/O by introducing several design dependent parameters reflective of architectural features in a realistic processing system. To include the effects of memory and interconnect, we first define the principle parameters of a computational system which enable us to model a variety of purpose-built processors.

- **Ratio of computation to memory \( \mu \):** We distinguish between the temporal ratio \( \mu_T \) and the spatial ratio \( \mu_S \). The relative number of memory accesses for each operation is \( \mu_T \), while \( \mu_S \) is the number of memory words in the system for each operator. With memory we mean SRAM, caches and the like but also off-chip DRAM. If \( \mu_T = 1 \), for each operation there is 1 memory access. Typical values will be between 1 and 3. On the other hand, the amount of memory is usually much higher than the operators. Hence, typical values for \( \mu_S \) are between 1000 and 10000 as we discuss later.

- **Ratio of on-chip versus off-chip memory \( \omega \):** If \( \omega = 1 \), all memory is on-chip; if \( \omega = 0 \), all memory is off-chip. In a 3-D topology, with on-chip we mean all dies in the 3-D stack.

- **Memory distribution factor \( \Delta \):**
  - \( \Delta = 0 \): completely distributed memory. The distance between a computation unit and the memory is always 0;
  - \( \Delta = 1 \): completely central memory where the distance between a computation unit and the memory is always the diameter of the system (or off-chip)
  - \( \text{e.g. } \Delta = 0.05 \): models a cache system where 95% of all memory accesses are local and 5% are far away.

This parameter models the communication required to write to and read from memory. If the memory is completely distributed, we assume all memory reads and writes are local and no long-range communication is required. Obviously, this is a simplification but any specific architecture-application pair can be characterised by a \( \Delta \) value between 0 and 1, denoting the amount of global on-chip communication occurring.

We explore different 2-D and 3-D topologies but we typically compare systems with the same total silicon area. E.g. if the total area is 400 \( \text{mm}^2 \), the configurations considered are:
• 2D: one plain silicon die of size $20 \times 20 \text{mm}^2$;
• 3D2: 2 dies stacked upon each other, each 200 mm$^2$;
• 3D4: 4 dies stacked upon each other, each 100 mm$^2$;
• 3D8: 8 dies stacked upon each other, each 50 mm$^2$;
• 3D16: 16 dies stacked upon each other, each 25 mm$^2$.

6.4. Effective Computational Efficiency

We define the Effective Energy (EE) for a 32-bit addition as:

$$EE_{arch}^{tn} = E_{32}^{tn} + \mu T (\omega (e_1 + \Delta \times E_{int_{arch}}^{tn}) + (1-\omega)(e_1 + E_{int_{arch}}^{tn} + E_{offchip}))$$

for a given technology node, $tn$, and a given architecture, $arch$, \{2D, 3D2, 3D4, 3D8, 3D16\}. The three main terms correspond to the energy consumption of an addition, of on-chip memory access and of off-chip memory access, respectively.

- $e_1$ is the amount of energy it takes to read or write one 32-bit word in on-chip SRAM.
- $E_{int_{arch}}^{tn}$ is the energy it takes to transport one 32-bit word from a nonadjacent on-chip memory to the local cache either over a 1 mm horizontal bus or from one vertical level to the next via a set of TSVs. For example, if the total silicon area is $400 \text{mm}^2$, we have

$$E_{int_{arch}}^{tn} = \begin{cases} (10+10)e_1(tn) & \text{if } arch = 2D \\ (7.07+7.07)e_1(tn) + e_2(tn) & \text{if } arch = 3D2 \\ (5+5)e_2(tn) + 2e_1(tn) & \text{if } arch = 3D4 \\ (3.5+3.5)e_2(tn) + 4e_1(tn) & \text{if } arch = 3D8 \\ (2.5+2.5)e_2(tn) + 8e_1(tn) & \text{if } arch = 3D16 \end{cases}$$

- $e_2$ is the energy it takes for a 32-bit word to be transported 1 mm horizontally in a given technology.
- $e_3$ is the energy it takes to move a 32-bit word from one vertical level to the next via a set of TSVs.
- $E_{offchip}$ is the energy to read/write the off-chip memory. It includes the energy consumed in the I/O drivers, the inter-chip communication as well as the memory controller.

The purpose of $E_{int}$ is to capture the communication energy in different architectures to get from an arbitrary point in the system to a particular point at the system boundary. For a 2-D 20×20 mm$^2$ die, the distance is on average 10 mm in each dimension, hence it is 20 mm. For a 3-D structure we have to traverse half of the vertical levels on average. E.g. for a 3D4 stack we have to traverse 2 vertical levels.
Thus, the effective energy, \( EE \) gives the required energy for a 32-bit addition if memory access and communication is taken into account. The factors \( \mu_T \), \( \omega \) and \( \Delta \) are abstractions of architectural choices and features. Based on \( EE \) in (1) we define the Effective Computational Efficiency (ECE) as

\[
ECE_{\text{arch}}^{\text{in}} = \frac{1}{EE_{\text{arch}}^{\text{in}}}
\]

the energy envelope of 1 Joule; or the amount of computations per second that can be carried out within the power envelope of 1 Watt. Figure 102 shows our model for computational efficiency as technology node scales. We have compared several recently implemented processors to demonstrate our model’s correlation to real-world systems. Clearly, general-purpose processors exhibit lower ECE than domain-specific processors such as DSPs due to the large overhead of control to implement the required features in a desktop processor. The abstract system-level parameters in our model allow for any processor architecture to be represented.

![Figure 102 The Effective Computational Efficiency of recent multi-core processors](image)

Figure 102 The Effective Computational Efficiency of recent multi-core processors [84; 97; 99; 100; 101; 119] (where markers indicate actual performance data) compared to our model. The limitation of 2-D systems shown by our model for 2-D computational efficiency is for a specific set of parameters (memory locality, bus width etc.). The model parameters can be altered to represent virtually any system, where the difference between our ideal and the actual implementation is dictated by the efficiency of the control circuitry.

**Effective Computational Density**

Similarly, the area cannot be filled with computational units only. We need to take memory and interconnect into account as well. We define the Effective Area (EA) as follows.

\[
EA_{\text{arch}}^{\text{in}} = A_{32}^{\text{in}} + \mu_S \omega a_1 + \sigma A_{\text{int}}^{\text{in}}_{\text{arch}}
\]

EA is defined similarly to EE but the off-chip component is omitted since we do not include the area for off-chip memory. Again, with “off-chip” we really mean “out-of-package”. Different dies in a 3-D stack are considered “on-chip”.

- \( a_1 \) is the area for a 32-bit memory word. Depending on the geometry we assume either SRAM or DRAM memory. For a 2-D system we use the area of embedded SRAM, while for a 3-D system we use DRAM.
Concretely we use 60F² area for one SRAM cell [117] and between 8F² and 4F² for DRAM cells [118], where F is the minimum feature size.

- $A_{\text{int}}^m$ is the interconnect area required for transporting a 32-bit word to memory (a 1 mm long 32-bit bus for 2-D systems or the area of 32 TSVs for a 3-D IC).
- $\sigma$ is the interconnect sharing factor. If $\sigma=1$, no sharing takes place and every operator has its own, private interconnect across the system. If $\sigma=0$, the interconnect is optimally shared and the interconnect area per operator is 0. In analogy to $\mu_S$, it gives the ratio of area occupied by operators versus interconnect. Typical values are between 0.01 and 0.1. For instance the Tilera TILE64 [119] with 64 cores has 8 32-bit operators per core. For 512 operators with an 8×8 mesh interconnect, the sharing factor is $\sigma = 16/512 = 0.031$. Note that only the global interconnect for the dataflow is counted, while the local interconnect and global control lines are ignored.
Table 21 Notation and metrics of comparison

<table>
<thead>
<tr>
<th>Abstraction of architectural design parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$l_n$</td>
<td>minimum feature size of a technology node (nm)</td>
</tr>
<tr>
<td>arch</td>
<td>architecture of system (2D, 3D2, 3D4, 3D8, 3D16)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>ratio of on- to off-chip memory ($\omega=1$; all memory is on-chip, $\omega=0$; all off-chip)</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>memory distribution factor ($\Lambda=1$; all centralized memory, $\Lambda=0$; all local)</td>
</tr>
<tr>
<td>$\mu_T$</td>
<td>number of memory accesses per h/w operation ($\mu_T=1$; one memory access per operation)</td>
</tr>
<tr>
<td>$\mu_S$</td>
<td>amount of memory per h/w operator (typically $\mu_S = 1000 - 10000$)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>interconnect sharing factor ($\sigma=1$; no sharing, $\sigma=0$; completely shared)</td>
</tr>
<tr>
<td>$n$</td>
<td>number of die layers for 3-D architectures</td>
</tr>
<tr>
<td>area</td>
<td>area of die</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technology and architecture dependent parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{12}$</td>
<td>energy for a 32-bit add operation</td>
</tr>
<tr>
<td>$e_1$</td>
<td>energy for a 32-bit read/write to local SRAM</td>
</tr>
<tr>
<td>$e_2$</td>
<td>energy to transport a 32-bit word over 1 mm on a planar on-chip bus</td>
</tr>
<tr>
<td>$e_3$</td>
<td>energy to transport a 32-bit word over 1 vertical layer across TSVs</td>
</tr>
<tr>
<td>$a_1$</td>
<td>area for a 32-bit memory word in SRAM or DRAM</td>
</tr>
<tr>
<td>$a_2$</td>
<td>area for a 1 mm long 32-bit planar on-chip bus</td>
</tr>
<tr>
<td>$a_3$</td>
<td>area for 32 TSVs</td>
</tr>
<tr>
<td>$E_{offchip}$</td>
<td>energy to read/write to off-chip memory. Includes I/O drivers, inter-chip communication and memory chip energy consumption.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary comparison metrics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$ICE$</td>
<td>number of 32-bit add operations per Joule</td>
</tr>
<tr>
<td>$ICD$</td>
<td>number of 32-bit adders per mm²</td>
</tr>
<tr>
<td>$A_{in\text{arch}}$</td>
<td>interconnect area required to transport a 32-bit word from a non-adjacent on-chip memory to the local cache: $\sqrt{\frac{\text{area}}{n}} a_2 + \frac{n}{2} a_3$</td>
</tr>
<tr>
<td>$E_{in\text{arch}}$</td>
<td>interconnect energy required to transport a 32-bit word from a non-adjacent on-chip memory to the local cache: $\sqrt{\frac{\text{area}}{n}} e_2 + \frac{n}{2} e_3$</td>
</tr>
<tr>
<td>$EE_{arch}$</td>
<td>Effective Energy for a 32-bit addition: $EE_{arch} = E_{12} + \mu_T (\omega e_1 + \Delta E_{in\text{arch}} + (1-\omega)(e_1 + E_{in\text{arch}} + E_{offchip}))$</td>
</tr>
<tr>
<td>$ECE_{arch}$</td>
<td>amount of computation achieved with 1 Joule: $\frac{1}{EE_{arch}}$</td>
</tr>
<tr>
<td>$EA_{arch}$</td>
<td>Effective Area for a 32-bit addition without off-chip memory: $EA_{arch} = A_{in\text{arch}} + \mu_S \omega a_1 + \sigma A_{in\text{arch}}$</td>
</tr>
</tbody>
</table>
6.5. Scaling circuits, devices and interconnects

To provide a physical foundation for our system-level model, we based our work on realistic circuit-level parameters extracted from published data, SPICE simulations, parasitic extraction tools, and consistent scaling methodologies [55]. The technology parameters cover global planar 2-D wires, TSVs, logical operations, memory transactions, thermal properties, transistors and leakage power.

Interconnect

We model planar on-chip wires and off-chip transactions for technologies from 180 nm down to 17 nm (including repeater insertion strategies, drivers and transistor characteristics) in a similar manner to [121]. The models and tabulated parasitics for global and off-chip transactions are described in detail in Chapter 1 and 2 respectively. We have also conducted field solver simulations of cylindrical, copper-filled TSVs to extract the relevant RLC parasitics. The dimensions vary depending on the technology node; the cross-section is assumed to be uniformly circular, with radii of 10, 8, 6, 4, 2 and 1 µm and a constant length of 50 µm. The pitch of the TSVs is twice the radius to match planar global wire spacing trends. We use the extracted parasitics with the same methodology as the planar wires, where the bus width and switch factor match the 2-D parameters. Driver and receiver energy is also considered for the TSVs for any number of layers.

Logical Operation and DRAM Scaling

We model various logic operations such as a 32-bit addition or SRAM read, by using published data [122], for a particular technology node and scaling the energy and area for future or past generations. Dally in [120] publishes the energy per add operation of a 32-bit adder in 130 nm 1.2 V technology as 5 pJ. A reasonable approximation, including leakage, for the energy and area in other technology nodes can be obtained according to well-practiced scaling methodologies [55], based on transistor feature size, supply voltage and thermal-leakage dependencies of transistors. The off-chip DRAM transaction energy is not a simple function of the feature size, and depends on the DRAM architecture, its peripheral circuitry and also characteristics of off-chip drivers, terminations, and chip, package and board trace parasitics. We have used the Micron System Power Calculator [118] to estimate the average off-chip read/write power for different generations of DRAM, from SDRAM to DDR3. We have matched the DRAM generation to the technology node, such that 180 nm corresponds to SDRAM and 17 nm to DDR3. The flexibility in our memory model can allow for any type of memory to be integrated in the stack such as Flash or SRAM with a simple modification of the transaction energy.

Power, Leakage and Thermal Models

As the power density in 3-D ICs increases linearly with the number of active die layers, the thermal performance is an essential design consideration and subsequently forms a staple component of our physical models. We have created compact thermal models based on material dimensions, conductivities and cooling strategies to provide a power-related temperature analysis in our 2-D and 3-D architectures. The FloTHERM® Computational Fluid Dynamic (CFD) solver was used to verify the accuracy of the models under a JEDEC still-air thermal test environment. We extracted leakage power trends across technology nodes using SPICE simulations with Predictive Technology Models [35] for bulk CMOS transistors for a temperature range of 0°C to 200°C to develop temperature-dependent current sources for operational die power in our analysis. Leakage is of especial concern with 3-D systems (particularly for lower technology nodes) due to the non-uniform temperature distribution across dies in the package, leading to a large variation of leakage power and hence reducing the overall maximum thermal ceiling and performance. We mainly use forced convection air cooled heatsinks at the top of the package with a typical ball-grid array package, but we also consider the added heat removal benefit of microchannel liquid cooling between die layers.

6.6. Design space exploration

Varying the parameters (further described in section 6.3) in our model such as the amount of on-chip cache versus off-chip memory transactions and memory distribution factor (effectively how far away is the cache resource to the computational unit), can allow for virtually any processor architecture to be represented. Figure 103 shows the ECE for various topologies when A, representing the proportion of centralized memory, varies between 0 and 1. For all topologies a centralized memory drops down ECE significantly from over 60 GOPS/W to about 5-10 GOPS/W. Hence, there is a benefit from distributing memory, but only a distribution of A < 0.2 has a significant effect. This
benefit from distribution is more pronounced for a 2-D topology. Going from $\Delta=1$ to $\Delta=0.1$ improves $ECE$ for 3D16 by a factor 5, while the improvement is 8 for 2D. Intuitively the reason for this is that the energy of transporting data across the chip to a central memory is much lower for a 3D topology. Hence, if it is difficult to decentralize most memory accesses, the penalty will be lower for 3-D. However, the impact of centralized memory on performance becomes steeper for more advanced technologies. The effect is apparent for a 3D16 topology (see Figure 104). While the difference in performance between $\Delta=0$ and $\Delta=1$ is a factor 5.4 for 180 nm technology, it grows to a factor of 34 for a 17 nm technology. Hence, even if a 3-D topology can mitigate the cost of centralized memory, it is still growing exceedingly as technology advances due to the inverse effect on the performance of logic versus interconnect as a result of scaling.

![Figure 103](image)

Figure 103  The effect of on-chip memory locality on the $ECE$, where decreasing from 1 to 0 reduces the interconnect distance between operator and memory

For the purpose of a comparative study between 2-D and 3-D topologies, we have distributed a single processor over 3-D layers of 2, 4, 8 and 16, so each 3-D die in a 16-layer processor will have 1/16th the power of the total 2-D processor. To compare architectures, we consider concrete system configurations under power, frequency, area, and thermal constraints. Our scenarios mainly model cache systems where 95% of the memory transactions are on-chip ($\omega=0.95$) and local ($\Delta=0.05$). Furthermore, to represent a realistic system we require a certain amount of on-chip memory, in this case we set our $\muS$ parameter to 2000 words per operator, similar to the Tilera TILE64 [119], processor. Adjusting the amount of cache in either direction will directly affect the number of operators that can be squeezed into the die. We assume an interconnect sharing ratio, $\sigma$, similar to the interconnect area of a large mesh-based NoC or many-core processor such as the TILE64.
Performance

Our model exposes the potential of 3-D stacked systems, which mainly stems from (1) the possibility to integrate dense DRAM tightly into the multicore architecture, and (2) from the more power efficient interconnection in the third dimension, which essentially is due to shorter geometric distances. Theoretically a 3D16 topology offers 2.4 times higher performance per Watt than a 2-D topology and for every doubling of the stack height, we see a 20 to 30% increase of the performance per Watt figure. This relationship is encapsulated in Figure 105(a), where the maximum throughput is shown for increasing power constraints in a 400 mm² 17 nm processor. However, when cooling limitations are considered, we find that 3-D ICs above eight layers will struggle to dissipate more than 15 W with conventional air-cooled heatsinks, where stacks of up to four layers may consume up to 50 W of total power. In most realistic cases, the operational power of a device will be dictated by the technology, packaging, environment and the application rather than their absolute maximum limitations.

To understand realistic performance limitations of 2-D and 3-D architectures within a given domain, we have constrained each topology to operate below an absolute maximum temperature of 100°C at any point in the structure for an upper power limit of 100 W. 3-D topologies are constrained by their thermal performance more so than 2-D systems operating under the same power budget. Figure 105(b) plots the maximum throughput versus the number of die layers given the maximum thermal ceiling for each topology. Further inter-die cooling strategies such as fins, interposers and microchannel cooling have been shown to reduce temperatures of air cooled systems by up to 30% and it is likely, as the authors of [113] have also concluded, that additional cooling, such as liquid microchannels between die layers will be required for high performance logic-on-logic die stacks.

Therefore, in Figure 105(b) we have shown the effect of different cooling strategies as a percentage improvement over air-cooled heatsinks to depict what may be achievable in the future with 3-D systems. The optimal topology for throughput at this particular maximum thermal design power (TDP) mainly lies at a 2-layer 3-D system and when additional cooling is considered the apex intuitively shifts to larger 3-D stacks as the maximum power-per-die improves.
It is clear that large 3-D systems operating at their thermal ceiling are sharply limited by the lower power constraints necessary to maintain the thermal integrity of the package and logic. Leakage contribution to the total power consumption increases as feature size reduces and is more prevalent in 3-D topologies due to higher temperatures, which in turn degrades the theoretical maximum performance in larger stacks. However, the thermal junction-to-ambient package resistance in a two-layer 3-D system is still low enough that a two-layer 3-D system can attain higher throughput than an equivalent 2-D system. Furthermore, we find that a four-layer 3-D computational system at 35 nm has a performance advantage of 16% over the same system instantiated in a 2-D package two technology generations lower at 17 nm. This means that a processor design in 3-D with smaller numbers of layers can achieve equal or higher performance without significant investment in further technology node shifts.

Figure 105 (a) The maximum throughput (GOPs) for our topologies constrained by power and (b) The maximum throughput with a thermal ceiling of 100°C and 100 W. Dashed lines show additional heat removal beyond conventional air cooling.

![Graph showing maximum throughput](image)

Table 22 The maximum operations per second (GOPs) for a 50 nm 400 mm² processor invoked in 2-D and 3-D topologies for different applications constrained by the maximum power under a thermal envelope. The maximum performance is shown in bold-face font.

<table>
<thead>
<tr>
<th>App.</th>
<th>P(W)</th>
<th>T(°C)</th>
<th>2D</th>
<th>3D2</th>
<th>3D4</th>
<th>3D8</th>
<th>3D16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile</td>
<td>5</td>
<td>65</td>
<td>164</td>
<td>201</td>
<td>240</td>
<td>357</td>
<td>257</td>
</tr>
<tr>
<td>Laptop</td>
<td>25</td>
<td>75</td>
<td>1054</td>
<td>1295</td>
<td>1471</td>
<td>715</td>
<td>289</td>
</tr>
<tr>
<td>Desktop</td>
<td>65</td>
<td>85</td>
<td><strong>2445</strong></td>
<td>2383</td>
<td>1668</td>
<td>822</td>
<td>321</td>
</tr>
<tr>
<td>Server</td>
<td>150</td>
<td>120</td>
<td><strong>2867</strong></td>
<td>2797</td>
<td>1977</td>
<td>965</td>
<td>402</td>
</tr>
</tbody>
</table>
6.7. Cost

3-D integration incurs significant costs related to stacking that are over and above the cost of a pure 2-D implementation, including the cost of the TSVs, test, pick and place of Known Good Dies (KGD) and bonding, and any additional cooling. Offset against this is the fact that the individual dies occupying the different layers are a fraction of the area that would be occupied by a single 2-D die, resulting in more dies per wafer as well as increased yield due to the smaller die area, with a significant drop in cost (per die in the 3-D stack), generally acknowledged to drop off as the 4th power of die area [55].

The other main issue is that a significant (fixed) investment is required in shifting technology nodes, related to infrastructure as well as design. This investment is usually amortized over many production runs for large-volume processors and the cost-per-unit asymptotically approaches the variable cost (costs that are proportional to the volume of a given product) with increasing volume. As we have shown though, the performance gain achievable by reducing feature size with its accompanying costs can be matched or bettered by a 3-D implementation without a tech shift. Our focus in this section therefore is to answer primarily two questions: first, as the complexity of the system increases and the total silicon area grows, is there a point at which a system implemented in 3-D becomes more cost effective than a 2-D implementation, and if so, what is that area, and the corresponding system architecture? Second, is there a volume of units sold at which the total unit cost of a 2-D system including the design related Non-Recurring Engineering (NRE) cost of moving to that node (costs that cannot be billed directly to a single product) becomes equal to the total unit of an equivalent 3-D system implemented in the older technology, which does not have the NRE cost associated with feature size reduction. We call this volume the cost-equilibrium volume.

In carrying out a comparative cost analysis we divide the variable cost into a die cost that includes material, labour, and process costs and a test cost. The die cost is a function of the wafer cost, number of dies per wafer and die yield. The yield has material-defected related ($Y_m$), systematic ($Y_s$) and random ($Y_r$) components which are complex functions of die area, and process related parameters including defect density and other statistically estimated quantities. We use typical values for MPU product families as reported in the ITRS [105] for these yields.

![Figure 106 Die-to-wafer variable cost](image_url)
The stacking cost for 3-D systems is estimated by factoring in the required extra mask layers and associated processing costs, the cost of pick-and-place and bonding as a fraction of the wafer cost. The test associated with selecting KGDs to bond onto the base wafer, and the yield drop due to stacking is also considered. Some of the model parameters can be sourced from the open literature \[105\] while the main imponderable is the 3-D stacking cost as a fraction of the wafer cost. Based on information gathered from our involvement in 3-D integration projects, we have carried out investigations for a technology that can be approximated by a stacking cost that is 20% of the wafer cost.

The first question we posed is answered in Figure 106, which shows the costs of 3-D systems implemented using Die-to-Wafer (D2W) stacking normalized to the 2-D system cost for that particular silicon area. This normalized view clearly shows the cost effectiveness of 3-D vs. 2-D; for smaller areas, 3D integration is more expensive than a 2-D implementation and the greater the number of layers in the stack, the higher the cost. However, as the total silicon area increases, having more 3-D layers lowers the unit cost. That is because the cost increases approximately as the fourth power of die area, and for large areas a very low yield in a pure 2-D implementation can be contrasted with the much higher yield of the smaller individual dies in the stack, which more than compensates for the extra 3-D bonding cost and reduced yield in the stacking. For yield parameters provided in \[105\], the cost-equilibrium point for D2W cost is approximately 170 mm$^2$.

This cost-equilibrium point changes with the 3-D stacking cost as well as defect density; the higher the defect density, the smaller the cost-equilibrium silicon area. It should be noted that the defect density of 0.13 per cm$^2$ used in this study is on the low side for cutting-edge technologies, and can be quite a bit higher, in which case 3-D would be even more attractive from a cost point of view.

To answer the second question, we estimated the design related portion of the NRE cost as being 75% higher when moving from 65 nm to 45 nm \[123\], which results in Figure 107. The inset shows the cost-equilibrium point for 2D systems as being approximately a million units. For the yield parameters used, the variable cost for implementing a 20 mm×20 mm system in 45 nm technology is about 24% of the cost in 65 nm technology, whereas it is 80% for implementing the system in a 2-layer 3-D configuration under the same 65 nm technology. The main graph shows what the cost-equilibrium point is for various 3-D implementations for a range of 3-D stacking costs and shows for example that even for a 3-D stacking cost of up to 40% of the wafer cost, approximately 900k units must be sold before the 45nm 2-D implementation becomes more cost-effective than a 65nm 4-layer stack.

6.8. Applications of the model

Modelling implemented processors
Our hierarchical models for performance reflect global system-level design choices and are not intended to model specific processing systems, which would require detailed information of the control, logic, layout and application.
The models we have created are intended to provide an early outlook of performance and cost for a wide-range of system design choices and technologies, which can aid in the partitioning and architectural organisation of a processing system before the physical design process. Similar works and in-house tools can provide fine-grained system-specific performance, area and cost information but are limited to that particular design space and do not provide a general model for computational efficiency for 2-D and 3-D devices. Quantifying the trade-offs between the investment in feature size reduction or 3-D stacking must be done for general systems and not necessarily for one particular design.

As an example, we model the Intel 80 core “teraflops” research chip to demonstrate the ability of our tool to model similar implemented processors. The 80 core device was designed to demonstrate the potential of a tiled multi-core processor design under a desktop power envelope. The 12.64 mm × 21.72 mm 65 nm die was laid out as an 8x10 packet-switching network-on-chip mesh to interconnect the 80 tiled-resources consisting of two high-performance floating point units and 2KB of data SRAM. The processor achieved its peak performance per Watt (19.4) at 394 GFLOPs. Given this information we can populate a list of equivalent input parameters for our model of the processing unit as shown in Table 23. These parameters are then used to examine the performance limitations of that system under various constraints controlled by the user such as the number of layers the system can be partitioned into, the locality of the memory to the computational units, the technology node and TDP limitations. This provides a fast first-estimate of the capabilities of a particular system under different technological and architectural design choices which would normally be considered in the planning phase of a new product. This output, coupled with the cost comparison between 2-D and 3-D designs detailed in section 6.7, can allow fast and early optimization of a processing system for a given application, streamlining the planning stages and increasing emphasis for design and implementation of the device.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Type</th>
<th>Intel 80 Core</th>
<th>Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td># of Layers</td>
<td>1 (2-D)</td>
<td>1-16 (3-D)</td>
</tr>
<tr>
<td>A</td>
<td>Die Area</td>
<td>12.64×21.72 mm</td>
<td>275 mm2/N</td>
</tr>
<tr>
<td>tn</td>
<td>Tech. node</td>
<td>65 nm</td>
<td>180-17 nm</td>
</tr>
<tr>
<td>b</td>
<td>Data width</td>
<td>32-bit</td>
<td>32</td>
</tr>
<tr>
<td>μs</td>
<td>Memory/Operator</td>
<td>2K SRAM/2 FPU</td>
<td>1KB/Op</td>
</tr>
<tr>
<td>μt</td>
<td>Memory/Operation</td>
<td>App. Specific</td>
<td>01-Mar</td>
</tr>
<tr>
<td>σ</td>
<td>Bus Sharing Ratio</td>
<td>8x10 mesh/160 FPU</td>
<td>18/160=0.11</td>
</tr>
<tr>
<td>Δ</td>
<td>Memory Distribution</td>
<td>NoC Mesh</td>
<td>0.01-0.1</td>
</tr>
<tr>
<td>ω</td>
<td>On/off-chip mem</td>
<td>All on-chip</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>Power (W)</td>
<td>20-230</td>
<td>App. Specific</td>
</tr>
</tbody>
</table>

Shown in Figure 108 is our model for the Intel 80 core processor as compared to the same system partitioned over 2 (3D2) and 4 (3D4) layers. The reduced interconnect energy caused by partitioning the system in the vertical dimension demonstrates significant improvements in the overall energy efficiency. We show that older technology nodes can achieve a similar performance to a 2-D device in a more aggressive technology node if implemented in 3-D and that the performance increasingly favours 3-D topologies as feature size reduces. This plot provides a technological comparison both in feature size reduction and 3-D integration without adjusting any of the architectural parameters in the model. Further examination of the memory architecture or interconnect organisation can provide essential insight into the performance limitations of a wide range of applications. The underlying physical parameters on which our global system-level model is founded upon are well-defined and on their own can provide fast and accurate design space explorations. For instance, models for TSVs, 2-D planar wires, and logical operations are available and accurate for a wide range of geometrical input parameters. Our global methodology for assessing processor performance is an amalgamation of all of the underlying physical models in order to provide a complete application perspective and quantify performance between different processor designs.
Trade-offs in multi-chip packages

So far we have used our models to define the performance limitations of a standalone system with all data transactions on-chip. In this section we show how our model can be used to quantify trade-offs between multi-die 2-D and 3-D packaged systems. Using the system-level computational models and the thermal simulation tool and the signalling parameters presented in Chapter 3, we compare the performance of a multi-core logic memory processing system in 2-D and 3-D packages. We again consider a 400 mm² single-die processor using parameters similar to the 80-core Intel teraflops experimental prototype with off-chip DDR3 DRAM in OLGA flip-chip packages.

Figure 108 The computational efficiency in GOPs/Watt of the Intel 80 core floating point mesh-based processor as a function of feature size. We model the processor given the parameters in Table 23 in 2-D and 3-D implementations. The plot quantifies the efficiency of partitioning a similar system to the Intel processor in 3-D layers of 2 and 4 dies. The same performance per Watt can be achieved in lower technology nodes in 3-D and the efficiency gap only increases as feature size reduces.
Instead of the 32-bit adder, we model a Floating Point Multiply-Add Accumulator (FPMAC) \cite{84}, which is the unit used in the Intel 80 core. For each FPMAC operation, the processor must read and write to and from the local SRAM cache, which is transferred over a given distance across the die on the global interconnect. When a memory miss occurs, or off-chip storage is necessitated the result must be transferred over the global interconnect wires through I/O drivers and off-chip traces to the destination DRAM die off-chip. For the 3-D implementations we partition the 2-D processor into two layers of 200 mm\(^2\) each but keep the DRAM off-chip as in the 2D case (3D2). We also consider two more 3-D topologies with the DRAM integrated in the stack: a 2-layer 3-D processor of 200 mm\(^2\) (3D2 w/DRAM) and a 4-layer 3-D processor of 100 mm\(^2\) (3D4 w/DRAM). The ratio of on- to off-chip memory transactions, \(\omega\), is set to 0.75.

Figure 109 The energy per 32-bit operation broken down into the logical operation and the on- and off-chip interconnect memory transaction. The values will change depending on the application (how many read/writes per operation) but the trends between the 2-D system (a) and the 2-layer 3-D system with integrated DRAM in the stack (b) will remain similar.
Figure 109 plots the energy for a 32-bit FPMAC operation including the expense in the interconnect and memory as technology node scales downwards. Figure 109 (a) shows the energy for a 2-D operation with off-chip DRAM broken into its constituent components. The off-chip energy is a function of the package-type and off-chip trace characteristics and does not scale with feature size. As the logical FPMAC operation decreases with each feature size reduction, the contribution of the off-chip transaction to the total power becomes greater. Although the on-chip energy decreases in the 2-D case, the global interconnect capacitance does not scale as sharply as the CMOS logic, which also results in an upper saturation limit which cannot be lowered. The two layer 3-D implementation with integrated DRAM is shown in Figure 109 (b) for the same model parameters. The off-chip transaction to the DDR3 in this case only has to traverse the vertical layers, thus the lower energy and higher performance as detailed in section 6.6 can be taken advantage of. The 3-D interconnect energy also scales favourably with the smaller transistors, becoming almost negligible below 35 nm. The on-chip interconnect energy is also favourable compared to the 2-D case because of the shorter global wires resulting from the partitioning of the 400 mm\(^2\) die over two layers. Further partitioning can achieve a reduction of energy by roughly 20-30% for every doubling of the stack, but the thermal design troubles become more prevalent.

Figure 110 The computational efficiency of a multi-core processor with off-chip DRAM is scaled for different topologies and technology nodes.
By taking the inverse of the effective energy per 32-bit FPMAC operation (EE\(_{32}\)), we find the computational efficiency of that system. Figure 110 plots this metric as feature size scales for the 2D and 3D2 systems with off-chip DDR3, and the 3D2 and 3D4 stacks with on-chip DRAM. The 2D package forms the baseline for comparison as it represents the lowest efficiency due to the longer on-chip wires and off-chip drive characteristics. Partitioning the 2D system over two layers but maintaining the off-chip DRAM only results in a marginal increase in efficiency (roughly 20%). The expense of the off-chip transactions dominates in these topologies. However, when the DDR3 is brought onto the die stack and the off-chip communication is replaced by the TSV interconnect, the increase in computational efficiency is over an order of magnitude greater. This value will depend largely on the application, where the amount of off-chip transactions will dictate the maximum efficiency of the 2D implementations.

To understand the maximum absolute performance of each topology, we again use our thermal simulation tool with a ceramic ball grid array (CBGA) package and heatsink with forced convection of 2.5 m\(^3\)/s to determine the maximum temperature affordable in each topology. When the input power results in a temperature in excess of 100°C (with an ambient temperature of 20°C) at any point in the package, we consider that the absolute maximum thermal dissipation power (TDP). We use this power to determine the number of operations per second afforded by each topology (i.e. Max. performance (OP/sec) = \(EE_{32} \times Max\ TDP\)).

Figure 111 plots both the maximum performance and the computational efficiency for each topology in a 32 nm CMOS process. The 2D system can obtain the highest absolute performance (in excess of 650 Giga ops) where the maximum absolute performance for the 3-D topologies decreases as the stack height increases. This is a result of the cooling limitations inherent in multi-die packages (in particular this CBGA we have chosen, where other package types may increase the maximum GOPs but the ratio of performance between topologies will remain largely the same). With more active layers, the power density increases but the thermal resistance of the package does not change. A 2-D package allows the die direct access to the heatsink and the substrate to board contact (also acting as a heatsink) which can afford much higher input power. The higher power enables the 2D system to compensate for the increased energy efficiency in the 3-D processor, resulting in favourable absolute performance.

However Figure 111 also plots the efficiency of each implementation on the right axis, where it is clear that the ECE of 3-D topologies far exceeds the 2D device. The product will rarely touch the maximum TDP of the package, where most systems will be designed to meet a power budget dictated by the battery-life or application. This means that processor memory systems designed for low power hand-held devices are more likely to operate within the thermal budget of a 3-D stack (sub-40 W for a 4-layer stack) and achieve a greater absolute performance than the equivalent 2-D device.
6.9. Conclusions

We have developed the concepts of effective computational efficiency (ECE) and effective computational density (ECD) to study the limits of performance of 2-D and 3-D topologies with technology down to 17 nm. Our model provides an abstraction of real systems in order to provide an upper bound on the performance. As such, we have not considered the control structure including logic, local interconnect and registers (which is less significant in comparison with global communication). The lower the overhead of the control structure, the closer the performance of a real system to our predicted upper bound, which is encapsulated by the DSP [101] in Figure 102.

Another limitation is our focus on throughput as the main performance characteristic, while ignoring latency. Latency is much harder to capture at an abstract level since it is influenced strongly by many details of the architecture, arbitration policies and resource management strategies. In real systems the theoretical limits of throughput are often not achieved because raw capacity is over-provided and a lot of control logic is used to keep critical latency figures low. It can be noted however, that a main benefit of 3-D topologies is the lower latency of memory transactions since high capacity memory can be located much closer to the computation units. This may mean that 3-D systems come closer to their intrinsic performance limits than 2-D topologies.

In summary, although our model constitutes an idealization of systems, it still expresses correct trends and bounds of real systems and we draw the following main conclusions from our study:

- 3-D systems can attain 2 to 3 times higher ECE due to lower interconnect power;
- 3-D systems have one order of magnitude higher memory density due to DRAM integration which means they can accommodate more computation units in a given area with the same amount of memory;
- This allows for much higher performance but causes also very high power density. Die stacks of over four layers will mainly be suitable in low-power mobile applications or high-density memory stacks such as Flash memory and a controller.
• The same performance with the same power can be realized in 3-D topologies with much smaller area and at lower frequency.

• A four-layer 3-D system can provide higher throughput than a 2-D system up to two technology nodes lower.

• The added expense and yield loss associated with 3-D stacking can be compensated by higher individual die yields and reduced NRE investments allowing 3-D systems above 170 mm² to reach their cost-equilibrium point earlier than a 2-D system.

The models which we have developed can be used to provide an early estimate of the performance limitations and capabilities of various processing systems before fine-grained layout and technological details are known. The quantification of both performance and cost for 2-D and 3-D systems as well as accurate parasitic models for a wide range of through silicon vias and 2-D wire geometries can provide designers with the framework to make realistic comparisons between the overwhelming number of CMOS design choices available in early-chip planning phase.

To enable the general community access to the tools, we have packaged all of our models as an openly available web-enabled and stand-alone program (see Figure 112) with a range of user input parameters. Constraints can be introduced such as the number of die layers, the maximum die power, and frequency or area to extract performance for concrete systems. The parameters can be adjusted to match existing processors and view their performance across technology nodes or in different 3-D configurations.
D2.5 CONCLUSIONS

The main contribution stems from a complete modelling and analysis methodology, which not only covers the physical behaviour of the interconnect structures and devices, but also the global circuit and application-based performance for realistic future massively integrated systems. The work contributes to state-of-the-art IC design in five main areas: (1) compact closed-form models for TSVs under numerous physical and topological constraints, (2) the exhaustive analysis of the signalling behaviour over TSVs and multi-die packages, (3) the development of a thermal simulation tool for 3-D packages, (4) extensive network-level studies to aid in the optimisation of partitioning and communication design, and (5) global system-level performance models to quantify the performance limitations of 2-D and 3-D high performance computational units.

The models for TSVs offer more topologies (isolated, rows and bundles) and process-related parameters (barrier thickness, radius, length, spacing and frequency) at either better or comparative accuracy to other models available in the literature. The models also show close correlation to the measurements conducted on ELITE test structures. The circuit-level performance of TSVs is then exhaustively defined for a number of standard on-chip signalling modes and compared to other package technologies, which significantly extends the available knowledge in the area of signal integrity and performance design. The model for average distance is shown to be an excellent metric for optimising network performance, exhibiting 100% fidelity in our studies. Finally, the scalability study and system-level efficiency models amalgamate all of the underlying physical models and studies to provide a global perspective on performance for future processing systems, where the limitations of 2-D and 3-D systems are concretely defined.

To disseminate the results, the models have been packaged in stand-alone tools and made openly-available on the web for the wider electronics and design community. The tools enable design space exploration of parasitic trends in TSVs, thermal performance for 3-D IC stacks and the limitations of computational efficiency for 2-D and 3-D processors.
References


[33] [135]

[34] Roshan Weerasekera and Dinesh Pamunuwa, "3-D IC Analysis," Internal, ELITE WP2 Modelling and Simulation 2009.


A Rahman and R Reif, "Thermal analysis of three-dimensional (3-D) integrated circuits (ICs)," In Proc Int. Interconnect Technology Conference, 2001, pp. 157--159.


[83] W J Dally and B Towles, "Route packets, not wires: on-chip interconnection networks," In


