A Fast Transient Simulation Strategy by Positively Utilizing On-Chip Inductance

Takayuki WATANABE†
† School of Management and Information of Innovation, University of Shizuoka
52-1 Yada, Suruga-ku, Shizuoka, Japan
Phone/Fax: +81-54-264-5433
Email: watanat@u-shizuoka-ken.ac.jp

1. Introduction

With the progress of semiconductor manufacturing technology, the number of transistors integrated in one chip reaches more than 1 billion. The speeding-up of the operating frequency and the scaling of the fabrication process expose many problems, such as the interconnect delay due to the parasitic capacitance, inductance and resistance, and the manufacturing variation. Therefore, physical design at the transistor level and its analysis technique become increasingly important.

Recently, several fast linear transient simulators, such as KSIM [1] and INDUCTWISE [2], have been developed for the analysis of on-chip power distribution grids, which are modeled as a large number of RL(K)C elements. In INDUCTWISE, the Modified Nodal Analysis (MNA) equation is transformed into the Nodal Analysis (NA) equation. Using the NA approach, the size of equations to be solved simultaneously is smaller than one by the MNA approach. Therefore, INDUCTWISE is much faster than SPICE which uses the MNA approach. However, these approach still have to solve the simultaneous equations with the large-scale sparse matrix. Also, the linear simulators cannot be used to analyze signal nets including nonlinear transistors.

Next, the Two-Stage Newton-Raphson (TSN) algorithm [3] is successfully applied to VLSI circuits with strong-coupling effects including linear and nonlinear elements. In this method, the linear and nonlinear portions are decoupled by the linear-nonlinear interfaces which are modeled as Norton equivalent circuits inside the Newton-Raphson iteration. The algebraic multigrid (AMG) method is used to solve the linear portion. However, the Newton-Raphson iteration for the whole circuit is still needed.

On the other hand, the Latency Insertion Method (LIM) is one of the fast transient simulation method based on the "leapfrog" algorithm [4]. In the LIM simulation, the node voltage vector and the branch current vector are computed alternately. It can analyze large scale RLC circuits very efficiently with much lower calculation cost of solving simultaneous equations in contrast to the implicit numerical integration used in SPICE-like simulator. However, this scheme has a limitation of the circuit structure to be analyzed. That is to say, every branch must have an inductor and every node must be connected with the grounded capacitor [5]. Therefore, if the circuit does not meet this condition, dummy latency is inserted into the original circuit.

This paper discusses a fast transient simulation strategy of the VLSI interconnect circuits based on the LIM. The PCB/Package interconnects are suitable for the LIM simulation because they usually have large parasitic capacitance and inductance [5]. However, on-chip interconnects are often modeled as RC circuits when inductive effects are negligible [6]. In such cases, a small inductance is inserted into the branch to generate the dummy latency according to [4]. There are no rules to decide the value of dummy inductors even though this value affects the accuracy of the transient response.

In this paper, a rule to decide the value and the insertion position of the dummy inductor will be given. After insertion of the inductor, the whole circuit which is partitioned by the dummy inductor can be simulated by using a method similar to the Block-LIM algorithm [7][8].

2. Latency Insertion Method

2.1 Standard LIM algorithm

In this section, we briefly introduce the LIM [4] and the Block-LIM[7][8]. The circuit to be analyzed by the LIM requires that each branch has an inductor and each node has a grounded capacitor. For the branch which consists of a series of an inductor, a resistor and a voltage source as shown in Fig. 1(a), branch current is updated, as in

\[
i_{ab}^{n+1} = i_{ab}^n + \frac{\Delta t}{L} (v_a^{n+1/2} - v_b^{n+1/2} - R i_{ab}^n + e_{ab}^{n+1/2}), \quad (1)
\]

(a)                        (b)

Fig. 1. Branch and node in LIM. (a): branch. (b): node.
where the superscript \( n \) indicates the time index.

Next, each node has a parallel combination of a capacitor, a conductance and a current source to the ground as shown in Fig. 1 (b). Then, the node voltage is updated as

\[
y_{a}^{n+1/2} = \left( \frac{C}{\Delta t} y_{a}^{n-1/2} + h_{a}^{n} - \sum_{k=1}^{M} i_{ak} \right) + \left( \frac{C}{\Delta t} + G \right), \quad (2)
\]

where \( M_{a} \) is the number of branches connected to the node \( a \).

Finally, the transient simulation is done by the alternate "leapfrog" updates of branch-currents and node-voltages according to (2) and (4) at alternate half-time-steps. To simulate stably, the size of time-step \( \Delta t \) is determined based on the minimum values of inductance and capacitance in the circuit as follows,

\[
\Delta t \leq \frac{1}{\sqrt{L_{C}}}. \quad (3)
\]

### 2.2 Block-LIM algorithm

The application of LIM is limited from the circuit structural point of view. If there are branches or nodes without inductance or grounded capacitance, minute valued inductance or capacitance must be inserted to each branch or node. Then, the time-step size in the simulation would be too small due to the minute valued inductance and capacitance in the circuit as follows,

\[
\Delta t \leq \frac{1}{\sqrt{L_{C}}}.
\]

In the Block-LIM [7][8], a large linear circuit is described in Fig. 1 (b). Then, the node voltage is updated as

\[
E_{i} = \begin{bmatrix} G & E \end{bmatrix} + \begin{bmatrix} C & 0 \end{bmatrix} \frac{d v}{dt} = \begin{bmatrix} h \end{bmatrix}, \quad (4)
\]

where \( \mathbf{E} \) represents the incidence matrix. From Fig.1, \( R \), \( L \), \( C \), and \( G \) are stamped as the diagonal elements in the sub-matrices \( \mathbf{R} \), \( \mathbf{L} \), \( \mathbf{C} \), and \( \mathbf{G} \) respectively. If the branch has no inductors, the resistance is moved and stamped in the sub-matrix \( \mathbf{G} \). Also, in the case that the node has no grounded capacitors, the conductance is stamped in the sub-matrix \( \mathbf{R} \). Then, the updating formulas are obtained by

\[
i_{n+1}^{a} = L_{-1}^{-1} \begin{bmatrix} 1 & \frac{1}{L} - R \end{bmatrix} i_{n}^{a} + E_{a}^{T} v_{n+1/2}^{a} + e_{n+1}^{a}, \quad (5)
\]

\[
y_{n+1/2}^{a} = \left( \begin{bmatrix} 1 \end{bmatrix} C + G \right)^{-1} \left( \begin{bmatrix} 1 \end{bmatrix} C v_{n-1/2}^{a} - E_{a}^{T} v_{n}^{a} + h_{n}^{a} \right). \quad (6)
\]

When all of the branches and the nodes in the given circuit have inductance and capacitance, the inverse matrix can be obtained easily. If there are some branches and nodes without any inductance and capacitance, their parts have to be treated as the circuit blocks.

### 3. Positive Utilization of On-Chip Inductance

#### 3.1 The value and the position of the dummy inductor

Generally, on-chip interconnects are classified into local, intermediate and global wiring. They are often modeled as RC lumped circuits. Since global interconnects have a lower resistance because of a larger cross-sectional area, inductive effects become pronounced in the high-frequency. In [6], the rule to require RLC modeling instead of RC modeling was defined as follows,

\[
\frac{t_{r}}{2\sqrt{L'C'}} < \frac{2}{R}\sqrt{\frac{L}{C}}, \quad (7)
\]

where \( t_{r} \) and \( l \) are an input rise time and a wire length, respectively. Also, \( R', L' \) and \( C' \) are per-unit-length line parameters. If \( l \) is outside the condition of Eq.(7), RC modeling is available.

Conversely, in the case of the RC model, the insertion of an inductor whose value is outside the condition of Eq.(7) has no affect on the transient response. Therefore, the largest inductance value is selected within possible to use the RC modeling.

From the second half of the conditions (7), the rule to require RLC modeling was rewritten as follows,

\[
L > \frac{R^{2}C}{4}, \quad (8)
\]

where \( R = lR', L = ll' \) and \( C = lC' \). Here, \( R \) and \( C \) are total amount of resistance and capacitance of the wire. Therefore, the largest inductance value within possible to use the RC modeling is \( R^{2}C/4 \). Here it will be shown that a larger \( R \) and \( C \) result in a larger \( L \). Since the time-step size in the LIM simulation is determined from (3), the insertion position of the inductor is chosen at the branch which connected with the largest capacitance. Finally, if the wire is modeled as \( n \)-segment RC ladder circuit, the following \( L \) is selected as the value of the dummy inductor.

\[
L = \frac{R^{2}C}{4n}, \quad (9)
\]

#### 3.2 Circuit partitioning and transient simulation algorithm

After insertion of the inductor, two sub-circuits are connected by the branch inserting the inductor as shown in Fig. 2(a). In our method, the whole circuit is partitioned by this RL branch as shown in Fig. 2(b). After the circuit partitioning, the two sub-circuits and the single RL branch can be simulated by using a method similar to the Block-LIM algorithm.

In our method, the unknown variable vector of the two sub-circuits and the RL branch current are computed alternately. For example, at the step \( n+1/2 \), the two sub-circuits...
connected with the constant current sources whose values were obtained as the RL branch current at the step n are solved by the SPICE-like algorithm. Next, at the step n+1, the RL branch current is updated according to (1).

Fig. 2. Partitioning the whole circuit into two sub-circuits and a single RL branch. (a): an original circuit after insertion of the dummy inductor. (b): sub-circuits after partitioning.

4. Numerical Results
In this section, the validity of our proposed method is evaluated based on actual layout design data of a two-stage inverter chain shown in Fig.3. In a 0.18μm CMOS process, RC models of the interconnect between two inverters are extracted by using Star-RCXT [9]. The RC extractions are performed for 3 different lengths of the signal wire, such as 100μm, 1mm and 5mm. Fig.4 shows the extracted RC model for the 5mm case. The circuit structure of "SIGNAL" portion in Fig. 4 varies with the lengths of the signal wire. In the 1mm case, the "SIGNAL" portion becomes the circuit structure as shown in Fig. 5. In the 100μm case, the equivalent circuit of the "SIGNAL" portion is not generated.

In the 5mm case, a dummy inductor whose value is 2.47nH is inserted into the branch at R29. In the 1mm case, a dummy inductor whose value is 0.108nH is inserted into the branch at R3. These values are calculated according to (9).

First, only the linear network shown in Fig. 4 is analyzed. Transient simulations are performed using the HSPICE by varying the value of the dummy inductor. The excitation is a pulse with rise time of 0.05ns. The resulting transient response waveforms are shown in Fig. 6 and Fig. 7. They indicate that the values of the dummy inductors are quite appropriate.

Next, similar simulations with the CMOS transistors are performed. The resulting transient response waveforms are shown in Fig. 8 and Fig. 9. They indicate that the dummy inductors have no effect on the transient responses.

Finally, only the linear network shown in Fig. 4 is analyzed by our proposed algorithm described in Section 3.2. The resulting transient response waveforms are shown in Fig. 10 and Fig. 11. The waveform calculated by our proposed method is in good agreement with the results of HSPICE.
5. Conclusion

In this paper, a fast transient simulation strategy of the VLSI interconnect circuits based on the LIM was discussed. Although on-chip interconnects are often modeled as RC circuits when inductive effects are negligible, the proposed method inserts the dummy inductor whose value is selected within possible to use the RC modeling.

Acknowledgment

This work was supported by Grant-in-Aid for Young Scientists (B), No.22700058 from MEXT Japan. This work is partly supported by VLSI Design and Education Center (VDEC), The University of Tokyo. The author would like to thank Prof. Hideki Asai for his pioneering research.

References