MOS VLSI Circuit Simulation by Hardware Accelerator Using Semi-Natural Models

Victor V. Denisenko,
Research Laboratory of Design Automation
Shadenko st., 89 Taganrog, 347924 RUSSIA
e-mail:den@rlda.rostov-na-donu.su

ABSTRACT - The accelerator is destined to circuit-level simulation of digital and analog/digital MOS VLSI'c containing of up to 100 thousand transistors (with 16 Mb RAM host-machine). The acceleration factor obtains of 3-4 order with respect to the Spice-2 program on VAX-11/780. The basic idea of the accelerator is to use real transistors instead of their mathematical models. In addition the accelerator concurrently uses 16-processors and programmable communications between processors and distributed memory, the waveform relaxation method and Spice-like input language.

In most practical cases VLSI digital circuits time-domain verification is carried out using switch level timing simulation. Switch-level verification and simulation both have low reliability, because they are based on significant simplification of the problem. Many of the essential features are not taken into account, for example: switch delay's dependence upon the input signals combination; effects of long transmission lines, non linearity of load capacitors and waveforms peculiarity, parasitic feedbacks, the influence of temperature and radiation. The attempts of taking these effects into account without solving the system of nonlinear differential equations lead to loss of reliability. On the other hand, the standard ways of solution for such a system of equations requires too many CPU time. These problems become more serious for the multivariant circuit simulation, optimization, statistical analysis, and for the percent of parametrical defect determination.

A radical solution to these problems would be to use a hardware accelerator. There are several commercial implementation of specific hardware accelerators for VLSI circuit simulation: the SPICE ACCELERATOR system by Weitek Corporation (has the acceleration factor 10-40 in comparison with Spice-2 program on VAX-11/780); the autonomic system series for circuit simulation SX-250, SX-1000 and SX-2500 by Shiva Multisystem Corporation (a factor of 20 to 100); the accelerator project constructed by Brunel University is assumed to increase the simulation rate to 100 times as much. The accelerator project Awsim-2 [1] has the accelerator factor to 50000, but it uses the greatly simplified mathematical models. It is assumed, that VLSI circuit does not contain the float capacitors, such as drain-gate capacitors, all of MOS-transistor capacitors are linear, source and drain resistances are neglected. These assumptions are of principle importance and they are the essential signs to differ the classical circuit simulation from its simplified variants. That is why these accelerators may not be considered as a circuits simulators without any conditionality.

This paper describes the project of hardware accelerator, which is an equivalent to SPICE-simulation and is about 1000...10000 faster than SPICE-2 with VAX11/780. The accelerator will take up to 16 Mbyte RAM as the number of transistors becomes about 100,000. Besides, the use of the real transistors instead of their

mathematical models allows to achieve the 99,9% - fidelity of volt-ampere characteristics for MOS-transistors inside the wide interval of temperature, radiation, light, humidity, pressure and vibration influence. The real devices as the models are subjected themselves by the external influence without terminating of simulation process.

1. BASIC PRINCIPLES OF ACCELERATOR DESIGN

The high operation characteristics of accelerator arise from the original simulation methods based on the use of real circuit elements instead of their own models. In several points it is similar to Realmodel, Realchip, Realfast systems (Daisy Systems corp., Valid Logic Systems Inc.), VEE-Test (Hewlett-Packard Co.), PDM (Plessy Semiconductors Ltd.), but in our case we have circuit (not logic) level simulation.

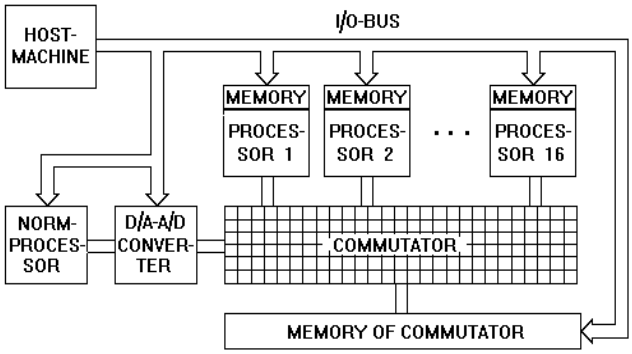

Fig. 1. The architecture of hardware accelerator with host-machine

The basic principles of the accelerator include exploiting the real transistors instead of their mathematical models and modification of Waveform Relaxation method (WR) [2]. Besides, the widely known methods of modeling have been used such as hardware realization of VLSI circuit element models in each processor; concurrence of a 16-processor system with the programmable communications between processors and distributed memory (fig. 1). Each processor consists of the processor for transistor simulation and processor [3] for two-pole elements simulation (fig. 2). The interprocessor commutator was organized in the form of the closed surface (fig. 2) and it allows to realize the overwhelming majority of connection between processors in accordance with simulated subcircuit.

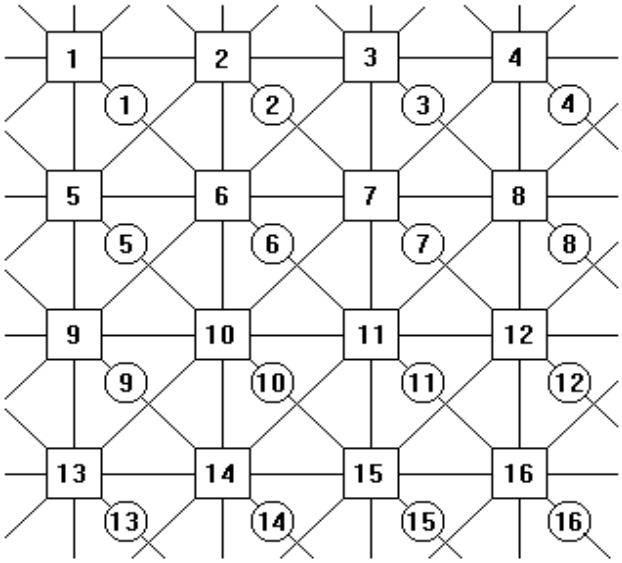

Figure 2. The commutator structure; square and round is processor for transistor and two-pole simulation

On the hardware/software level we use independent choice of timesteps for different subcircuits; piecewise-linear approximation of waveforms that reduces the expenditures of memory and increases the channel capacity (bandwidth) of interconnection between host-machine and accelerator; subcircuit simulation with different speed-up (and precision). We use also event control of the simulation process; partition of VLSI into subcircuits across the weak feedback paths; SPICE-input using the

subcircuits language construction for the circuit description language.

2. MOS-TRANSISTOR MODEL

The accelerator employs the semi-natural models of MOS transistor (fig. 3) which are based on the real (reference) transistors T_n and T_w , manufactured by the same technology as VLSI circuit is to be designed. In this case the most of transistor parameters can be considered as a build-in parameters of the models and we are not need in identification stage. We must only control the transistor width and length as well as drain and source resistances and capacitances.

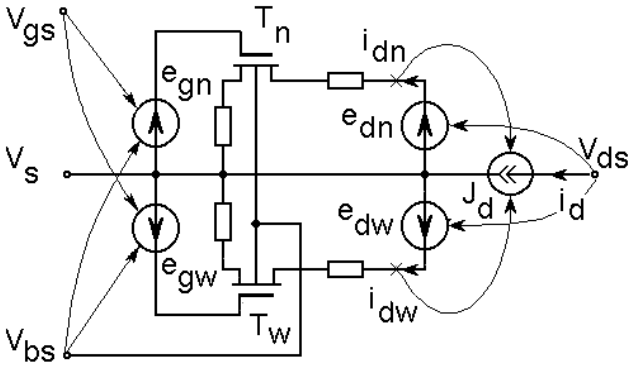

Fig.3. The semi-natural MOS transistor model

The controlled sources on fig. 3 reproduce the following dependence:

$$e_{gn} = V_{gs} - \delta \frac{\pi \varepsilon_{si}}{4 C_{ox}} \left(\frac{1}{W_w} - \frac{1}{\Delta W} \right) (2\varphi_F - V_{bs});$$

$$e_{gw} = V_{gs} - \delta \frac{\pi \varepsilon_{si}}{4 C_{ox}} \left(\frac{1}{W_n} - \frac{1}{\Delta W} \right) (2\varphi_F - V_{bs});$$

$$J_d = M_w i_w + M_n i_n; \qquad e_{dn} = e_{dw} = V_{ds};$$

$$M_n = \frac{W - W_w}{W_n - W_w}; \qquad M_w = 1 - M_n;$$

$$\frac{1}{\Delta W} = \frac{1}{\Delta W_n} + \frac{1}{\Delta W_w} - \frac{1}{W},$$

where δ , ε_{si}, C_{ox} , φ_F, are the widely known parameter of Spice program models; W_w , W_n - the channel width of transistors T_n and T_w; W is the channel width of transistor to be simulated. The values of other parameters and variables follow from fig. 3 and the above equations.

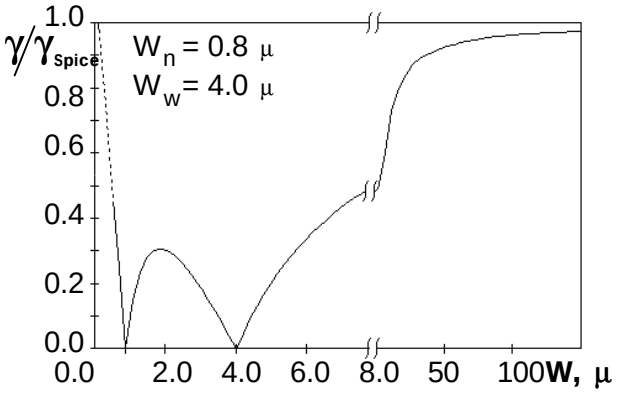

Fig.4. The Dependence of methodical error γ from the channel width **W**

The resistors on fig. 3 have the negative resistance and they are used to compensate the transistor own resistances. The dynamic simulation is implemented by the similarity theory methods, by means of dynamical processes scaling in 10^6 - 10^7 times.

The use of "real" transistors allows to take easily into account the number of short-channel effects, for instance: avalanche carrier generation into the drain p-n-junction and multiplying in lateral bipolar transistor, an oxide charging by hot electrons, the closing of source and drain space charge regions, and the electron tunneling through the oxide. These effects are not considered in MOS models of SPICE and other widely used simulation programs. The particularity of the model is its static characteristics 99,9% fidelity when $W = W_n$ or $W = W_w$ (fig.4). Here is: γ - the methodical error of semi-natural model; $\gamma_{\mathbf{Spice}}$- the error of

the Spice MOS-transistor model (LEVEL=2). The curve on fig.4 have built in assumption, that the methodical error γ_{Spice} is proportional to $|W - W_n|$ and $|W - W_w|$.

3. ACCELERATOR SOFTWARE

To simulate large scale circuit it is divided into subcircuits, that contain no more than 16 transistors and 32 two-pole elements. Each subcircuit is simulated by hardware tools and sewing the subcircuitstogether is realized by WR which is modified for the accelerator. These modifications include VLSI's decomposition into subcircuits across weak feedback paths; in necessary cases, the dynamic damping is used that realized on hardware tools.

The host-machine (fig. 1) ought to interact between the designers, data base and other design levels; translate the input information; decompose the VLSI circuit model; adjust of the interprocessor commutator; identify the model parameters; operate with the model library; rank the subcircuits; make the eventual control of simulation process; iterate the subcircuit models; communicate data with accelerator; control the global simulation process. The input language of the system is identical to one of Spice program. The workstation translator perform the syntax and lexical check of the input description, produces the lists of element model parameters in the format suitable for accelerator, and makes the LSI interconnects description ready for subsequent processing by the algorithm of interprocessor commutator setting and circuit partition.

The characteristic property of circuit partition algorithm is that circuit is decomposed by means of splitting into a number of functional units. To maximize the utilization of processor each subcircuit must contain 16 transistors and 32 two-terminal elements. To make this, the special procedure is used that integrates small subcircuits into the larger ones. The interprocessor commutator of the accelerator cannot realize all possible combinations of connections therefore the algorithm of commutator setting is used that select the optimal distribution for subcircuit transistors among the simulating processes.

All mentioned algorithms are executed ones in the translation stage, so they do not affect on the speed of simulating system. In each WR iteration the sampling is used for the analog signals taken from subcircuit terminals and subsequent recovery of these signals by means of piecewise-linear interpolation. The residual norm estimation unit, that is also hardware realized, checks the convergence of WR iterations.

4. EXPERIMENTAL RESULTS

The commercial using of above named ideas demands the semi-custom analog-digital VLSI technology which is nontraditional for computer systems. Such a technology has appeared 2-3 years ago inUSA and Japan and now the most rate of development is observed in this technology among all the others. Therefore we have fabricated the processor for simulation of the two-terminal elements and one MOS-transistor model on the basis of CMOS standard integrated circuits (registers, memories, DAC's, OPAmps and other).

It has the following experimental performance: the band of resistor parameters 0.5 Ohm...8 MOhm; the bandwidth of the capacitor and resistor parameters is 12 bit plus the 3-bit scale factor. The timing scale factor is about 10^6. It allows to increase the transistor model capacitors up to 1-10 mcF and avoid the influence of parasitic capacitances of the bread board.

During the experiment the WR-method was used to realize the iterations which connect the solutions for independent subcircuits of the circuit to be analyzed.

The experiment has confirmed the following results, earlier obtained theoretically:

1. The time-spent for simulation of each subcircuit does not depend on its size and complexity. It takes about 1 ms and 10 ms when relative errors are 10 and 1 percent respectively. This corresponds to speedup in 1...10 thousand times with respect to Spice program realized on VAX-11/780.

2. The noise and small accelerator's width of a bus do not alter the convergence rate of WR method.

3. The use of 30...50 sampling points and 1M RAM for each 10 thousand transistors is enough to represent the most typical waveforms with acceptable error (about 5-10 percent).

4. The electromagnetic emanation and the noise level from the host-machine and interface can be easily decreased to the neglected threshold.

REFERENCES

1. D.M. Lewis. IEEE Trans. on CAD No. 11 (1988) 1134.
2. E. Lelarasmee, A.E. Ruehli, A.L. Sangiovanni-Vincentelli. IEEE Trans. on CAD No. 3 (1988) 131.
3. V.V. Denisenko. Proc. 3-rd Int. Des. Autom. Workshop. July 19-20, Moscow (1993) 187.