Relative Timing

Kenneth S. Stevens, *Senior Member, IEEE*, Ran Ginosar, *Member, IEEE*, and Shai Rotem

***Abstract*—Relative timing (RT) is introduced as a method for asynchronous design. Timing requirements of a circuit are made explicit using relative timing. Timing can be directly added, removed, and optimized using this style. RT synthesis and verification are demonstrated on three example circuits, facilitating transformations from speed-independent circuits to burst-mode and pulse-mode circuits. Relative timing enables improved performance, area, power, and functional testability of up to a factor of 3× in all three cases. This method is the foundation of optimized timed circuit designs used in an industrial test chip, and may be formalized and automated.**

***Index Terms*—Asynchronous design, dynamic logic circuit, high performance, low-power design, performance tradeoffs.**

I. Introduction

THE design of RAPPID, the asynchronous instruction length decoder, took more than two years to complete [1]. The primary goal was to investigate whether asynchronous design could improve performance in high-end microprocessors. This naturally led to the effort, reported in this paper, to study and develop circuits, computer-aided design (CAD), and methodology most suitable for aggressive timed asynchronous circuit design.

Initial designs and methods were based on the CAD available at that time. The circuits were specified and synthesized using speed-independent (SI) or burst-mode (BM/XBM) methodologies [2]–[4], as well as metric timed circuit design [5]. We quickly discovered that many of the circuits that achieved our performance goals contained some form of timing assumptions—either the fundamental mode assumption of burst-mode or gate-level metric timing. The performance was improved by studying the natural delays of the circuits to employ timing that simplified the designs by reducing series transistors and logic levels.

Unfortunately, all the asynchronous methodologies at that time had what we considered an impediment to conceptualizing, optimizing, validating, and interfacing timed circuits. The timing assumptions were all *implicit*. We felt that in many cases, the key performance was achieved through careful management and design of the *timing* of the circuits as much as the behavior. We therefore studied ways to make the timing of circuits explicit. This effort resulted in the *relative timing* (RT) style reported here.

Manuscript received March 5, 2001; revised June 29, 2001 and January 7, 2002.

K. S. Stevens and S. Rotem are with Strategic CAD Labs, Intel Corporation, Hillsboro, OR 97124 USA.

R. Ginosar is with the VLSI Systems Research Center, The Technion, Haifa, Israel.

Digital Object Identifier 10.1109/TVLSI.2002.801606

Relative timing proved to be a very effective method of substituting aggressive pulse-mode self-resetting circuits for the original full-handshake speed-independent designs in RAPPID. This novel method also allowed us to design and verify speculative asynchronous state machines. However, this effort required a new way of thinking about asynchronous designs and required a new set of tools.

In the absence of RT CAD tools, the manual flow is quite inefficient for the design of large systems. Now we face the question of how our manual method can be formalized into an effective CAD methodology and tools. We propose that new formal methodologies and tools be developed to support this method. This paper presents our methodology and lessons in order to motivate further CAD development. We start with simple, contrived examples that demonstrate basic principles, and move to a key RAPPID circuit that has been improved substantially with relative timing.

II. Motivation and Description

The design of timing in digital circuits is an extremely difficult challenge. The conventional clocked digital design methodology solves this problem by decomposing the circuit into cycle-free combinational logic (CL) stages and interstage clocked latches; the clock cycle is simply tuned to accommodate the worst case propagation delay in the CL stages. The behavior of the combinational logic can be specified and synthesized without considering timing. Delay-insensitive (DI) asynchronous circuits are analogous to clocked CL design in the sense that both types are independent of time—the behavior will be correct for arbitrary gate and wire delay.

High-performance circuits, both clocked and asynchronous, benefit from more aggressive timing methodologies. Clocked circuits can be considerably enhanced using local self-timing [6]–[8]. Timed asynchronous circuits can also have significantly enhanced performance.

Asynchronous design consists of handshake protocols that ensure the validity of data [9], [10]. Asynchronous design methodologies, apart from DI, make timing assumptions in the protocols, function logic, or data transmission [11]. If the assumptions are invalid in the physical implementation, the circuits can glitch and fail to operate correctly. SI circuits assume indistinguishable skew on wire forks, burst mode assumes fundamental mode (the circuit will stabilize internally before new inputs arrive), and bundled data assumes that all data is stable before the handshake signal arrives. Ensuring that the timing assumptions hold in timed design, such as burst mode, can be challenging [12].

The design style we investigated explicitly specifies the*effect* of delays in a circuit in terms of assertions on relative ordering

of events (e.g., a goes high before b goes low). Our application of relative timing is based on the unbounded delay model commonly used by many asynchronous synthesis and verification tools. The circuits are then designed to meet the relative orderings and validated that the constraints are part of the natural delays in the system.

A number of benefits emerged from making RT constraints explicit in our designs. Timing relationships are no longer hidden by a design style or tool. RT can unify the asynchronous methodologies as well as provide support for ad hoc manual designs. The bundled and burst-mode assumptions, for example, can usually be made explicit with a small number of RT constraints, as shown in Section IV-B4. The explicit nature of the constraints can simplify interfacing, synthesis, and performance verification. RT is not restricted to any particular specification style and supports arbitrary designs. Since timing can directly effect the quality and robustness of the circuits, each assumption can be individually evaluated, and its application can be aggressive or conservative.

Many timing CAD tools and methodologies exist; asynchronous design itself is a timing methodology. Ordering signals temporally is not novel. This ordering can be achieved through graph transformations that reduce concurrency similar to the theory developed by Vanbekbergen [13]. Timed Petri nets, timed finite-state machines, and other bounded-delay formalisms have been used to reason about timed circuits in [14]–[20]. Component databooks include waveforms showing relative signal orderings, and orderings have been applied to micropipeline latches and controllers [21]–[23]. These methodologies can achieve extremely efficient circuits; indeed, the tag unit in RAPPID, used as the primary example in this paper, was first specified, synthesized, and validated using the metric tool ATACS [24].

However, we do feel that the RT methodology used in RAPPID applies timing top-down in a novel way that is intuitive and flexible, creating compact, testable, high-performance low-power circuits in a style that can be automated by CAD. Further, this methodology supports both automatic and user-specified timing transformations. Initial RT solutions based on this work applied to synthesis [20] and verification [25] show remarkable results and potential for an automated RT design flow.

III. Rappid Relative Timing Design

Relative timing had a significant impact on the RAPPID results. The timed asynchronous circuits, when compared to similar clocked logic in a commercial synchronous implementation, showed a 3× improvement in throughput, a 2× improvement in latency, and half the energy per operation, at a 20% area penalty [1]. Although harder to quantify, we feel that relative timing was also key in achieving the 95% stuck-at testability in RAPPID with our functional built-in self-test method through removing redundancies that naturally result through fixed signal orderings induced by timing.

Most of the RT circuits in RAPPID were designed by hand. The RT transformations modified many behavioral aspects of the specifications, concurrency in particular. However, the essential functionality of the controllers—synchronization and ordering—remained. This effort, while time consuming, helped us better understand timing, timed technology mapping, and what types of transformations appeared most beneficial. Various forms of handshaking were investigated, including protocols without direct acknowledgment. These pulse-based protocols can at times significantly improve the simplicity and latency of asynchronous circuits.

Most of our implementations were mapped onto standard static and domino library cells. Domino circuits are a restricted class of generalized C-elements [26], where only a single term exists in the reset function. The combination of state-holding, low transition latency, and low activity factor of the domino gates made them the best circuit alternative we investigated.

Section IV describes the method we developed for designing and optimizing relative-timed circuits. What began as a number of circuit experiments evolved into a manual flow. Automated tool support for these flows was painfully lacking, so we began mentoring development of RT CAD. Early engagement with the Petrify team led to automation of synthesis using relative timing [20]. Verification using RT constraints was added to the verification tool Analyze [27] in-house. This tool was used to optimize the constraints in a slow, error-prone manual loop. Theory automating the verification and RT constraint optimization is under development [25]. We encourage researchers to further formalize and develop new CAD for automating RT design.

IV. Examples

A. Notation and Terminology

Table I shows some notations used in this paper. The process logic CCS [28] is used in this paper, where "." is the sequential operator, "+" is the nondeterministic choice operator, "|" is parallel composition, and "$\backslash\{\underline{a}\}$" is the restriction operator applied to signal $\underline{a}$, which disallows independent $\underline{a}$ and a transitions. Restricting signal a only permits the internal τ synchronization of the "handshake" between $\underline{a}$ and a.

All simulations have been made using synchronous standard library cells in a 0.18-μ process. The output of each circuit drives a 0.18×25 μ gate load. The circuits are simulated using SPICE and the values normalized against one of the circuits in terms of area and energy. A more complete modeling of some of these circuits and parameters can be found in [29].

The circuit examples in this paper contain static and domino gates normally employing a single pMOS device. Asynchronous tools such as ATACS [5], 3D [4], [30], and Petrify [2] can typically synthesize set–reset flops and the appropriate functions [Fig. 1(a)]. We can often apply technology mapping into single-variable reset (equivalently set) functions and implement them using standard footed domino gates as in Fig. 1(b). When the reset variable is not used in the set function, an unfooted domino gate is used instead [Fig. 1(c)].

B. C-Element

We use a simple C-element example to demonstrate the concepts and methods of applying relative timing to synthesis and verification. A simple two-input generalized C-element and its CMOS implementation are shown in Fig. 2(a). The formal def-

TABLE I
NOTATION CONVENTIONS

Signal	Description	Example
input signal	underline	input
output signal		output
inverted (asserted low)	over-bar	$\overline{z}$
rising transition	up arrow	a↑
falling transition	down arrow	b↓
timing arc	dashed arc	- - →
behavioral arc	solid arc	→

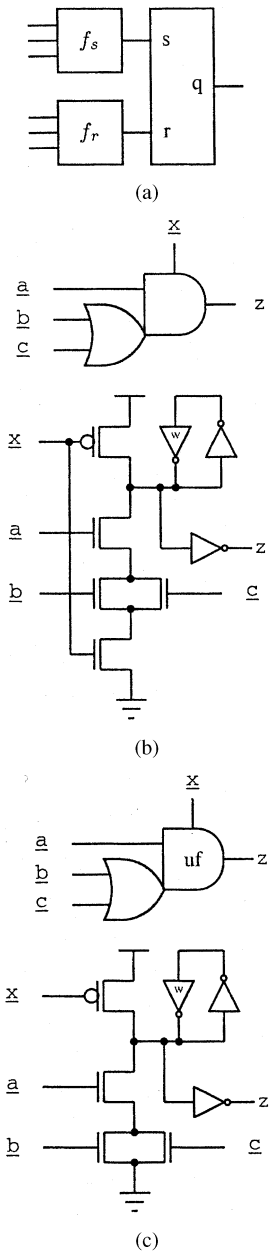

Fig. 1. (a) Set–reset flop and functions. (b) Footed domino gate (symbol and circuit) implementing a set–reset flop with $f_r = \overline{\mathbf{x}}$, $f_s = \mathbf{x} \times \mathbf{a} \times (\mathbf{b} + \mathbf{c})$. (c) Unfooted domino gate implementing $f_r = \overline{\mathbf{x}}$, $f_s = \mathbf{a} \times (\mathbf{b} + \mathbf{c})$.

inition in CCS is $C = (\underline{\mathrm{a}} \mid \underline{\mathrm{b}}).\mathrm{z}.C$, which reads "$C$ is defined as single transitions showing on inputs a, b in parallel (at any order), followed by a transition on the output z, then followed recursively by C again" [28]. An equivalent signal transition graph (STG) [31], [32] representation of the specification is shown in Fig. 3(a) .

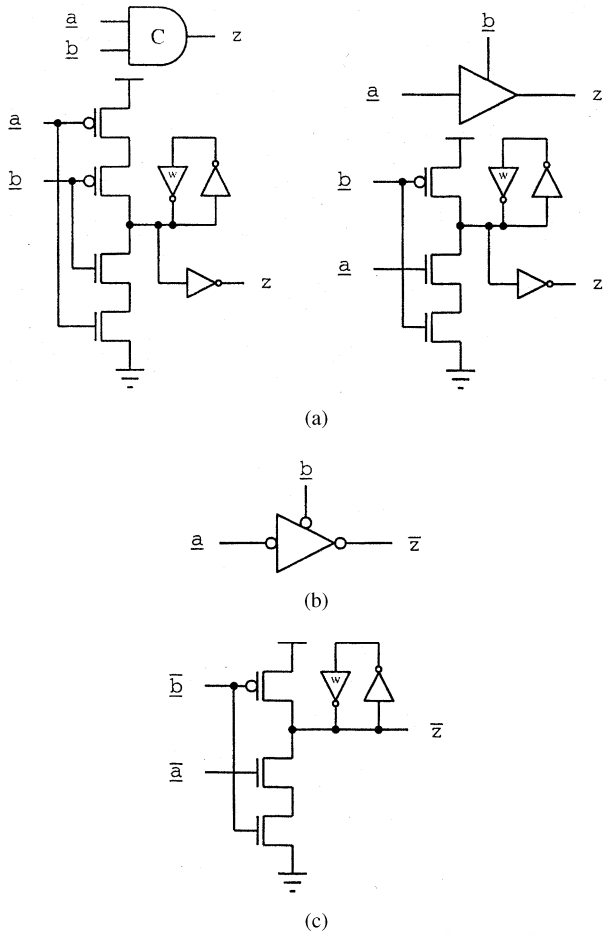

Fig. 2. Generalized C-elements: (a) GC, (b) GC-RT for $\mathbf{a}\downarrow \prec \mathbf{b}\downarrow$, and (c) for $\mathbf{a}\uparrow \prec \mathbf{b}\uparrow$.

1) Relative Timing Synthesis: RT synthesis optimizes a circuit by adding timing arcs to a behavioral specification. Both timing and causality affect the behavior of an RT circuit. Behavioral arcs must be synthesized into gates, and timing relations enforce a specific ordering between concurrent events, resulting in concurrency reduction in the specification.

Relative timing assumptions come in two forms: local and global. Local timing constraints can automatically be generated by moving behavioral arcs based on various assumptions such as *lazy transition systems* [20]. Global assumptions are dictated by the response of the environment. These assumptions can be applied manually, as in Section V-C, or automatically, as in the burst-mode assumption that a circuit will stabilize before a new input burst arrives [4], [30], [33].

RT synthesis supports the creation and strengthening of timing assumptions by moving the relative positions of the heads and tails of arcs in a specification. If timing arcs are restricted to relative translations of behavioral arcs, aggressive timing optimizations can be performed on a circuit while ensuring a consistent, compatible result. The new specification can now be synthesized, and timing assumptions and requirements can be back-annotated. In this section, we show some simple, intuitive transformations on C-elements. In Section V, we show aggressive application of relative timing in a large circuit.

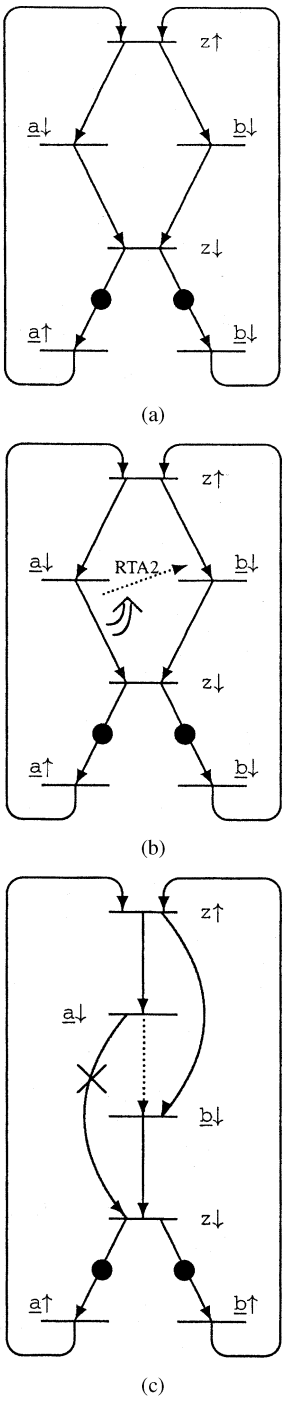

Fig. 3. Relative timing transformations on the Petri-net of a C-element: (a) initial spec, (b) relative timing arc RTA2 $\underline{a}\downarrow \prec \underline{b}\uparrow$ effectively "translates" arc $(\underline{a}\downarrow, z\downarrow)$ to $(\underline{a}\downarrow, \underline{b}\uparrow)$, and (c) new spec, $\times$ arc is redundant.

2) Synthesis Examples: Assume that the environment always produces transitions on $\underline{a}$ before transitions on $\underline{b}$..This relative timing assumption is expressed as follows:

$$\text{RTA1: } \underline{a} \prec \underline{b}$$

the C-element can be reduced to a buffer $C = \underline{b}.z.C$ using this assumption. Fig. 3(b) shows the STG when the assumption is limited to the falling edges

$$\text{RTA2: } \underline{a}\downarrow \prec \underline{b}.\downarrow$$

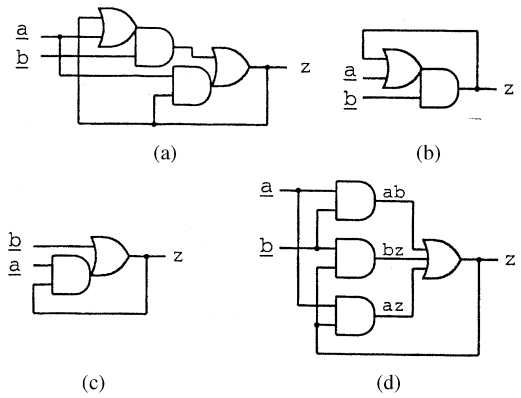

Fig. 4. Static C-elements: (a) speed-independent, (b) with RT assumption $\underline{a}\downarrow \prec \underline{b}\downarrow$, (c) with RT assumption $\underline{a}\uparrow \prec \underline{b}\uparrow$, and (d) burst-mode C-element with hazards.

the dashed arc represents the timing assumption RTA2. Note that the timing arc supersedes the behavioral arc from $\underline{a}\downarrow$ to $z\downarrow$. Relative timing effectively moves the tail of this arc from one event $(z\downarrow)$ to a predecessor of the event $(\underline{b}\downarrow)$, as indicated by the double arrow in Fig. 3(b). The new timing arc makes the behavioral arc redundant, as shown in Fig. 3(c). In the corresponding circuit, the reset function contains only $\underline{b}\downarrow$, and the C-element can be implemented as the Fig. 2(b) footed domino gate GC-RT

$$C = (\underline{a}\uparrow | \underline{b}\uparrow).z\uparrow.\underline{a}\downarrow.\underline{b}\downarrow.z\downarrow.C$$

given a similar assumption on the positive edges

$$\text{RTA3: } \underline{a}\uparrow \prec \underline{b}\uparrow$$

the circuit can be mapped to the domino gate in Fig. 2(c) by inverting the inputs and employing the nonbuffered $\overline{z}$ output.

Static C-element implementations can be synthesized with Petrify. The STG of Fig. 3(a) produces the speed-independent circuit (SIC) shown in Fig. 4(a). Timing assumptions RTA2 and RTA3 lead to the simpler static circuits of Fig. 4(b) and (c), respectively. Note that these two circuits are actually subcircuits of the speed-independent circuit. 3D synthesizes the circuit of Fig. 4(d).

In general, applying relative timing for synthesis means that new (timing) arcs are inserted, rendering other arcs redundant. This could also be considered as moving the head, tail, or both ends of behavioral arcs to predecessors. This effectively reduces concurrency in the specification, allowing a simpler implementation by removing transistors and gates.

3) Relative Timing Verification: This section introduces the method developed to verify a large, relative-timed asynchronous circuit called RAPPID [1]. An implementation I *conforms* to a specification S $(I \succeq_c S)$ when an implementation is an acceptable construction of the specification [16], [27], [34]. In this section, implementations can be assumed to be parallel compositions of the untimed behavioral specifications of the gates. Relative timing predicates can be added to implementations and specifications to reduce their concurrency by pruning states in a state graph (SG) that are unreachable due to timing. Thus, a specification S conforms to an implementation I with RT predicate R when $I \wedge R \succeq_c S$.

Early in this effort, the Analyze verifier was enhanced to support RT predicates on both implementations and specifications. Circuits can then be verified using SI and DI unbounded delays with RT constraints.

RT verification has two aspects. First, RT constraints reduce concurrency in the implementation by disallowing transitions to failure states. Second, the set of RT constraints are optimized and merged through a set of transformations.

The following algorithm was applied to generate RT constraints and verify RAPPID and the circuit examples in this paper. Step 1) generates RT constraints that remove a single failure state, as will be shown in the following example. This capability was added to our verifier. Step 2) optimizes the constraint by reducing additional concurrency beyond the single failure state. Step 3) adds the new optimized RT constraint to the set, removing any constraints covered by the new constraint. Steps 2) and 3) were done manually.

1) Verify conformance using current RT predicates.
   - If failure free, report RT constraints.
   - If failure cannot be fixed through timing, quit.
   - If failure exists, create RT constraint(s) that remove this failure.
2) Optimize new constraint(s).
   - Remove concurrency by increasing coverage of the SG by the RT constraint.
   - Iterate optimization, terminating when:
     i) further concurrency reduction would remove states required by the specification;
     ii) slack in constraint is no longer positive;
     iii) an arc edge touches a primary input or output.
3) Add optimized constraint to RT constraint set, remove covered constraints, and iterate.

Section IV-B-IV illustrates the procedure used in RAPPID to determine how and when to increase coverage of an RT constraint.

4) Verification Example: Consider the static C-element (SC) in Fig. 4(d). This circuit is implicitly hazard-free under burst-mode or fundamental-mode assumptions. However, it is not hazard-free in a speed-independent environment. If the environment responds quickly, $\underline{\mathtt{b}}\downarrow$ may immediately follow $\mathtt{z}\uparrow$ before node $\underline{\mathtt{az}}$ rises, resulting in a hazard.

Fig. 5 shows a state graph of the SC C-element circuit. The "bottom" symbols in the left and right corner of the diamonds label error states. Transitions $\mathtt{ab}\downarrow_1$ and $\mathtt{bz}\downarrow_1$ lead to the error state on the right.

Using the method described in Section IV-B3, we first try to eliminate the right error. Verification will identify any arcs that lead to error states. The following two constraints eliminate the right error state:

$$\text{RTC4: } \mathtt{az}\uparrow_4 \prec \mathtt{ab}\downarrow_1$$
$$\text{RTC5: } \mathtt{az}\uparrow_3 \prec \mathtt{bz}\downarrow_1 .$$

If one signal must precede another and both exit from a single state, then the later arc will never be taken (e.g., $\mathtt{ab}\downarrow_1$ in RTC4). RTC4 and RTC5 therefore disallow entrance to the right error state.

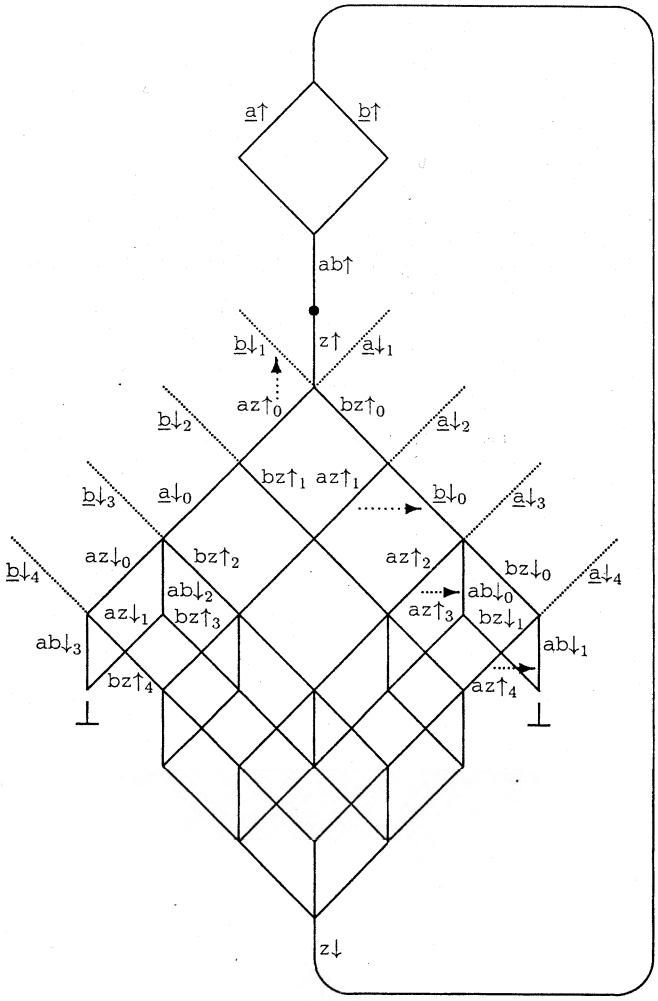

Fig. 5. Relative timed burst-mode SC state graph.

One representation of the timed precedence of RTC4 is the dashed arc between $\mathtt{az}\uparrow_4$ and $\mathtt{ab}\downarrow_1$ in the SG of Fig. 5. We now try to strengthen this constraint to cover more of the graph. While there may be many methods to optimize the instance-based constraints, our hand methodology used two main iterations.

First, instance information is removed from the constraints when possible. The generalized constraint $\mathtt{az}\uparrow \prec \mathtt{ab}\downarrow_1$ is equivalent to RTC4, as it adds no new timing arcs to the SG. Generalizing the right side as well results in the constraint $\mathtt{az}\uparrow \prec \mathtt{ab}\downarrow$, effectively adding a second timing arc $\mathtt{az}\uparrow_2 \prec \mathtt{ab}\downarrow_0$ to the SG. This constraint now removes two states from the graph: the error state and the state of RTC5. Hence RTC5 is covered by the optimized RTC4 constraint.

Second, the generalized RT constraint can be strengthened based on slack calculations.[1] The constraint is strengthened by moving the left and/or right transition to earlier transitions in the SG. Hence the right side of constraint RTC4 can be strengthened to transition $\underline{\mathtt{b}}\downarrow$ and $\mathtt{bz}\uparrow$. A simple unit delay model can be used to calculate slack, where local gates are assigned a single delay and input transitions are assigned a value k, where $k \geq 1$. The

[1]Slack is the difference between the latest arrival of the first signal and the earliest arrival of the second.

TABLE II
COMPARISON OF C-ELEMENT IMPLEMENTATIONS. FALL, RISE, AND ENERGY COLUMNS USE WORST SIC PERFORMANCE AS BASE, WITH ALL OTHER NUMBERS A MULTIPLE OF THAT DELAY OR ENERGY. ENERGY IS AVERAGE FOR A COMPLETE CYCLE (RISE AND FALL). TEST COLUMNS SHOW COSMOS STUCK-AT FAULT COVERAGE ON ALL FANOUTS, WITH REDUCED PATTERNS IN RTA2 COLUMN DUE TO ENVIRONMENT RESTRICTIONS

Circuit	HF in SI Env.	Fall Delay	Rise Delay	Switching Energy	Area # Transistors	Exhaustive Testability	RTA2 Env. Testability
SIC	Yes	1.00	0.98	1.00	16	100%	90%
SIC-RT	No	0.59	0.52	0.74	8	n/a	100%
SC	No	0.56	0.51	2.03	18	100%	92%
GC	Yes	0.77	0.55	0.86	10	100%	100%
GC-RT	No	0.52	0.52	0.72	9	n/a	100%

following example illustrates the strengthening of az $\uparrow \prec$ ab $\downarrow$ starting from the common signal z $\uparrow$:

$$\begin{aligned} &\mathtt{z}\uparrow\mathtt{az}\uparrow \prec \mathtt{z}\uparrow\{\mathtt{bz}\uparrow\ \underline{\mathtt{b}}\downarrow\}\mathtt{ab}\downarrow \quad k \\ &\mathtt{z}\uparrow\mathtt{az}\uparrow \prec \mathtt{z}\uparrow\underline{\mathtt{b}}\downarrow \quad k-1 \\ &\mathtt{z}\uparrow\mathtt{az}\uparrow \prec \mathtt{z}\uparrow\mathtt{bz}\uparrow \quad 0 \end{aligned}$$

this indicates that if $k > 1$, the best strengthening is RTC6; otherwise, the weaker az $\uparrow \prec$ ab $\downarrow$ should be used. Applying the same method to the left error state generates RTC7

$$\begin{aligned} &\text{RTC6: } \mathtt{az}\uparrow \prec \underline{\mathtt{b}}\downarrow \\ &\text{RTC7: } \mathtt{bz}\uparrow \prec \underline{\mathtt{a}}\downarrow \end{aligned}$$

the RT implementation now conforms to the specification. Precisely, SC $\wedge$ RTC6 $\wedge$ RTC7 $\succeq_c C = (\underline{\mathtt{a}} \mid \underline{\mathtt{b}}).\mathtt{z}.C$. All signals in these constraints are either primary inputs or directly enabled by the primary output, simplifying hierarchical validation.

In general, RT verification allows one to manipulate the initial constraints to arrive at a minimal set of constraints that are easiest to verify in a hierarchical system. Constraints that have overaggressively reduced the slack can be weakened back to the original failure state. If any initial constraint contains unachievable timing, then the circuit is an invalid implementation of the specification.

RTC6 and RTC7 implement a "weak" form of the fundamental-mode requirement of burst mode. Because Analyze uses bisimulation semantics, hazardous behavior inside a circuit that does not propagate to the outputs is permissible due to the *observational equivalence* property [27], [28]. (This is not the case when using verifiers based on weaker formalisms.) RTC6 and RTC7 prune arcs $\underline{\mathtt{a}}\downarrow_{1,3,4}$ and $\underline{\mathtt{b}}\downarrow_{1,3,4}$ but transitions $\underline{\mathtt{a}}\downarrow_2$ and $\underline{\mathtt{b}}\downarrow_2$ of Fig. 5 remain. Given RTC6 and RTC7, if $\underline{\mathtt{a}}\downarrow_2$ occurs, az $\uparrow$ will either glitch or not fire. This does not create an observable failure because signals bz and ab are asserted holding z high, and the output will not lower until $\underline{\mathtt{b}}\downarrow$ and bz $\downarrow$ occur. Hence the additional "strong" burst-mode RT constraints az $\uparrow \prec \underline{\mathtt{a}}\downarrow$ and bz $\uparrow \prec \underline{\mathtt{b}}\downarrow$ are unnecessary.

5) C-Element Summary: Table II summarizes the five alternative designs from Figs. 2 and 4. The circuits are all sized near the optimal power/performance point. All designs were simulated to drive the same load. If a circuit is hazard-free in an SI environment, then no timing is required for correct operation. The SIC is slower than all others. Applying the RTA2 assumption to this design leads to a circuit (SIC-RT) that is half the size and enhances performance by almost a factor of two. The static SC requires the largest circuit and is fast, but doubles the power. The reduced domino C-element (GC-RT) improves fall times over the GC circuit by 50% (due to simplification of the pullup stack) and rise times by 5%.

Static circuits tend to expend more energy than domino circuits. This is largely due to the extra switching activity in the static designs, as can be observed by the SC circuit, which expends twice the power of the SIC circuit because all four gates toggle for every output transition. When activity factors are similar, the domino circuits have a slight edge. The GC-RT circuit uses only 3% less energy than the static SIC-RT circuit because the reduced device sizes in the domino gates are offset by the short circuit current when the gates switch. Testability was measured in COSMOS using a functional test methodology, where only valid timed signal orderings allowed by the environment can be supplied to the circuit. The table shows that the static and SI circuits are fully testable for complete patterns, but not when timing constraints reduce signal interleavings (in column RTA2). The RT optimized versions of these circuits are fully testable.

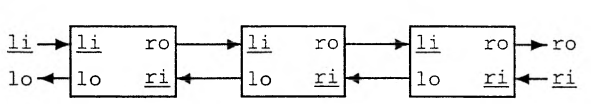

Fig. 6. FIFO block diagram containing three cells.

V. TIMING EVOLUTION IN A RING

In this section, we trace the development of a simple first-in–first-out (FIFO) controller, similar to a micropipeline [35]. These controllers can be connected in series as shown in Fig. 6. This circuit is a simplified abstraction of a part of the RAPPID design [1] and closely follows the actual steps used to derive the final circuit. We begin with a speed-independent design and review a succession of progressively simpler circuits, enabled through careful application of relative timing assumptions.

A. Speed-Independent FIFO Cell

A simple FIFO cell can be specified in CCS as follows:

$$\begin{aligned} \text{LEFT} &= \underline{\mathtt{li}}\uparrow .\underline{\mathtt{c}}.\mathtt{lo}\uparrow .\underline{\mathtt{li}}\downarrow .\mathtt{lo}\downarrow .\text{LEFT} \\ \text{RIGHT} &= \quad \mathtt{c.ro}\uparrow .\underline{\mathtt{ri}}\uparrow .\mathtt{ro}\downarrow .\underline{\mathtt{ri}}\downarrow .\text{RIGHT} \\ \text{FIFO} &= (\text{LEFT} \mid \text{RIGHT})\backslash\{\underline{\mathtt{c}}\}. \end{aligned} \tag{1}$$

The specification in (1) consists of two handshake processes, LEFT and RIGHT. The $\underline{\mathtt{c}}$ event synchronizes the two processes so that $\underline{\mathtt{ri}}$ must go low and $\underline{\mathtt{li}}$ must rise before both processes may proceed. This process-based specification is equivalent to the Petri-net of Fig. 7.

The SI circuit in Fig. 8 was synthesized using Petrify [2] and is a hazard-free implementation of (1).

B. Burst-Mode FIFO Cell

The SI FIFO pays a considerable delay penalty to achieve speed independence. The trace $\underline{\mathtt{li}}\uparrow$, y1 $\uparrow$, lo $\uparrow$, y2 $\uparrow$, ro $\uparrow$ shows that $\underline{\mathtt{li}}\uparrow$ produces $\underline{\mathtt{lo}}\uparrow$ after two complex gate and inverter delays and $\underline{\mathtt{ro}}\uparrow$ after four. Perhaps the performance can be improved if the circuit can ensure that concurrent outputs are

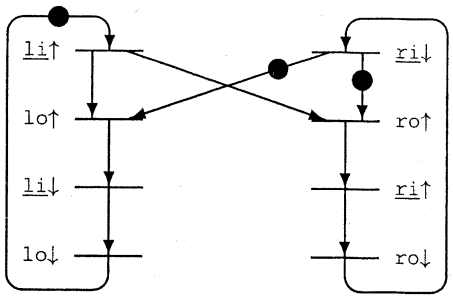

Fig. 7. FIFO specification Petri-net.

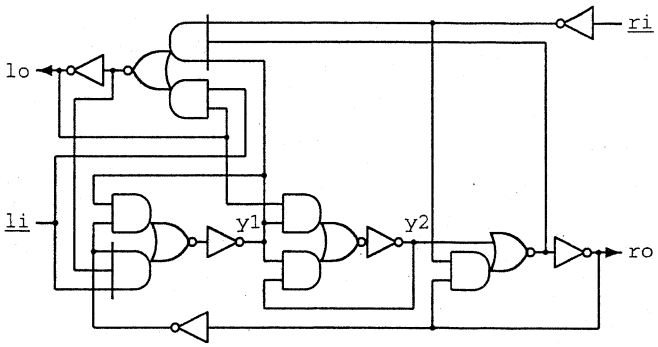

Fig. 8. Speed-independent FIFO cell (SI).

generated faster than they can be acknowledged by the environment. This assumption can be formulated as follows:

$$\text{RTA8: } \texttt{lo} \uparrow \prec \underline{\texttt{ri}} \uparrow$$
$$\text{RTA9: } \texttt{ro} \uparrow \prec \underline{\texttt{li}} \downarrow .$$

A new specification is generated by adding these two relative timing assumptions to (1)

$$\text{FIFO} \wedge \text{RTA8} \wedge \text{RTA9} \qquad (2)$$

where FIFO is the specification from (1). This is equivalent to the Petri-net of Fig. 9, where the dashed arrows represent relative timing constraints.

Note that the two constraints RTA8 and RTA9 are in a form where outputs precede inputs and these outputs are concurrently enabled from the same pair of inputs. This is a burst-mode constraint where the input burst is $\{\underline{\texttt{li}} \uparrow \underline{\texttt{ri}} \downarrow\}$ and the output burst is $\{\texttt{lo} \uparrow \texttt{ro} \uparrow\}$. This burst-mode timing assumes that the variance in the generation of the concurrent outputs is always less than the response time of the environment.[2]

The RT-BM circuit of Fig. 10 is derived in [20] using the new RT synthesis capabilities of Petrify and implements (2). (The C-element here is synthesized as an OR gate in [20].) RT verification by Analyze extracts the timing in the physical circuit and creates additional orderings that must hold for the circuit to operate correctly

$$\text{RTC10: } \texttt{x} \downarrow \prec \texttt{lo} \downarrow$$
$$\text{RTC11: } \texttt{x} \downarrow \prec \texttt{ro} \downarrow .$$

[2]Also applying burst-mode constraints on input set $\{\underline{\texttt{li}} \downarrow \underline{\texttt{ri}} \uparrow\}$ results in a C-element—the micropipelines implementation.

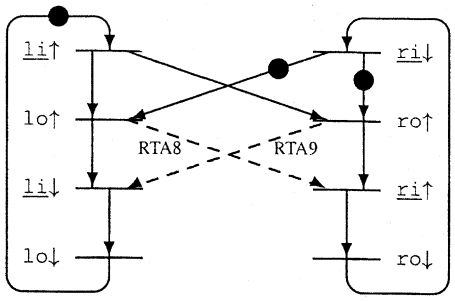

Fig. 9. FIFO specification Petri-net, with RT assumptions RTA8 and RTA9 represented as dashed arcs.

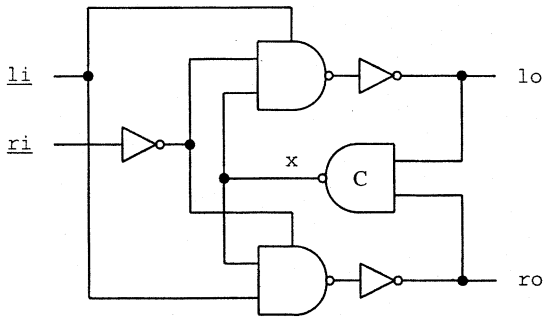

Fig. 10. Relative timed burst-mode FIFO (RT-BM).

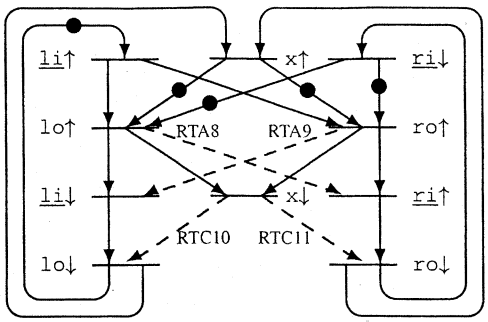

Fig. 11. Petri-net for circuit of Fig. 10 and constraints RTC10–RTC11.

These constraints, as well as the state variable x, are shown graphically in Fig. 11. The burst-mode implementation achieves a 2.6× average speedup over the SI circuit. Constraints RTC10–RTC11 apply only to the physical implementation and must be validated given actual circuit delays.

C. Right Before Left

Assume that we connect the circuit of (2) into a ring with a single token. The token will always arrive at an idle cell due to circuit delays if the ring is sufficiently large. Hence the handshake in process RIGHT will always complete before a new handshake in process LEFT. The SI or RT-BM circuits can safely be used in a large ring. However, the global timing of RTA12 can improve the circuit in terms of power, performance, area, and testability

$$\text{RTA12: } \underline{\texttt{ri}} \downarrow \prec \underline{\texttt{li}} \uparrow$$

this assumption can be graphically represented as shown in Fig. 12. The arcs from $\underline{\texttt{ri}} \downarrow$ to $\texttt{ro} \uparrow$ and $\texttt{lo} \uparrow$ are now redundant and have been removed from the figure.

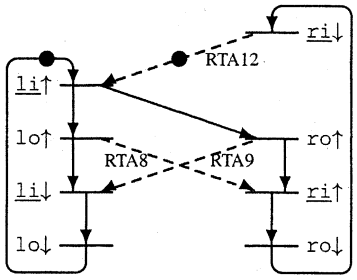

Fig. 12. Net representing addition of RT assumptions $\underline{\mathtt{ri}}\downarrow \prec \underline{\mathtt{li}}\uparrow$.

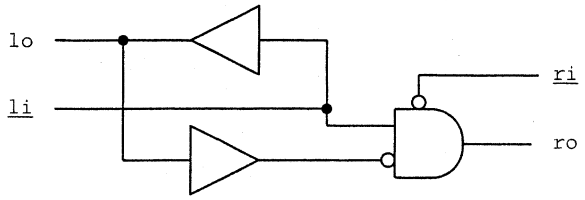

Fig. 13. Aggressive relative timed FIFO (RT-Agr).

The dashed arcs are not *causal* arcs; $\underline{\mathtt{ri}}$ must go low before $\underline{\mathtt{li}}$ can rise and $\underline{\mathtt{ri}}$ cannot delay $\underline{\mathtt{li}}$. This represents a major change in the operation of the circuit; the LEFT process is no longer synchronized directly with the RIGHT process except through system timing. The design must guarantee that the token appears on the dashed arc before $\underline{\mathtt{li}}\uparrow$.

The circuit in Fig. 13 can be synthesized with 3D and Petrify from (2) adding assumption RTA12. The rising edge of signal $\underline{\mathtt{li}}$ must be delayed sufficiently through lo and the buffer to ensure that the domino AND gate is not disabled before it is fully set. This results in a number of RT constraints on races in the circuit that can be derived as was done for RTC4–RTC7 in the SC circuit. This circuit shows 15% and 3× improvement in average case performance over the RT-BM and SI circuits, respectively, and energy is also improved by factors of 26% and 1.9×.

D. Pulse-Mode FIFO Cell

RTA12 now constrains the specification sufficiently to derive a pulse-mode circuit. Through transitivity, ro ↓ must precede $\underline{\mathtt{li}}\uparrow$. We can use this weaker constraint to discard $\underline{\mathtt{ri}}$, the backward handshake signal, altogether. We show how this can be accomplished through transformations on the circuit of Fig. 13.

Three elements of the ring are shown in Fig. 14. Observe that the lo signal is nothing more than a delayed version of the $\underline{\mathtt{li}}$ signal. Shuffling the lo devices and bubbles results in the circuit of Fig. 15, which has only forward-moving signals without any intercellular feedback. The shuffling that removes acknowledgment is directly based on RTA12 that dissociates the LEFT process from the RIGHT. This shuffling turns output lo and input $\underline{\mathtt{ri}}$ into local signals.

Note that signal $\overline{\underline{\mathtt{li}}}$ in Fig. 15 is just $\underline{\mathtt{li}}$ inverted. A transition $\underline{\mathtt{li}}\uparrow$ creates a short period when both $\underline{\mathtt{li}}$ and $\overline{\underline{\mathtt{li}}}$ are high, which will set the output of the domino AND gate. The duration of both inputs to the domino AND gate's being high depends on the delay in the $\overline{\underline{\mathtt{li}}}$ path. This signal pair can be combined into a single wire $\underline{\mathtt{li}}$ if the signal on this wire operates as a pulse. The final circuit derivation can be seen in Fig. 16.

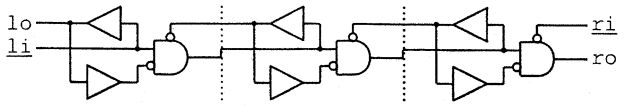

Fig. 14. Aggressive relative timed FIFOs.

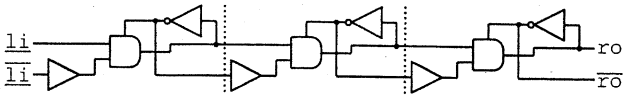

Fig. 15. Shuffled aggressive relative timed FIFO cells.

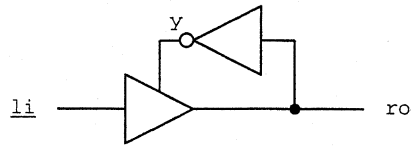

Fig. 16. Relative timed pulse-mode FIFO (pulse).

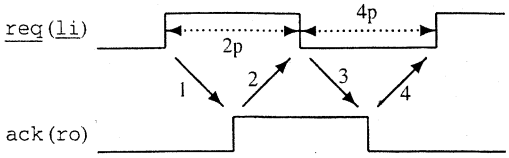

Fig. 17. Four cycle and pulse handshake protocol constraints.

The following specification removes the direct handshake signals lo and $\underline{\mathtt{ri}}$ of (1) and adds RTA12

$$\begin{aligned} \text{LEFTP} &= \underline{\mathtt{li}}\uparrow .\mathtt{c}.\underline{\mathtt{li}}\downarrow .\text{LEFTP} \\ \text{RIGHTP} &= \mathtt{c.ro}\uparrow .\mathtt{ro}\downarrow .\text{RIGHTP} \\ \text{PULSE} &= (\text{LEFTP} \mid \text{RIGHTP})\backslash\{\mathtt{c}\} \\ &\quad \wedge\, \mathtt{ro}\downarrow \prec \underline{\mathtt{li}}\uparrow . \end{aligned} \tag{3}$$

Designing reliable pulse-mode circuits is very difficult [36]. We can observe some of the constraints of pulse circuits by understanding how we have derived the pulse-mode circuit in this example. Fig. 17 shows a four-phase request-acknowledge handshake. Constraints 1–4 are causal with speed-independent signaling. By removing the ack signal (lo and $\underline{\mathtt{ri}}$ in Fig. 14), we are left with only the request signal that requires constraints 2p and 4p. These constraints contain both minimum and maximum metric bounds. However, the actual requirements for the size of these bounds can be represented with relative timing arcs between the inputs and outputs of a pulse-mode circuit ($\underline{\mathtt{li}}$ and ro in Fig. 16). Interestingly, these arcs correspond to a protocol very similar to the standard request acknowledge handshaking.

The pulse on $\underline{\mathtt{li}}$ of Fig. 16 causes the output pulse ro, as required by (3). If we map $\underline{\mathtt{req}}$ to $\underline{\mathtt{li}}$ and ack to ro in Fig. 17, we see that arc 1 is causal. However, this circuit can fail if the pulse is so short that the ro(ack) pulse does not occur. We can therefore impose an RT constraint that requires ro ↑ (ack ↑) before li ↓ (req ↓). This makes arc 2 in Fig. 17 an RT constraint, and slightly restricts the specification. (It may be possible to not restrict the specification if an internal signal toggles, which ensures that the domino gate has changed state.) The circuit will

TABLE III
COMPARISON OF FIFO IMPLEMENTATIONS. ALL DELAYS ARE IN TERMS OF SI CIRCUIT WORST CASE DELAY, ENERGY IN TERMS OF SI CIRCUIT. ENERGY ACCOUNTS FOR A COMPLETE FOUR-PHASE CYCLE. SYNCHRONOUS TESTING IN COSMOS REQUIRED EXTRA TEST GATE FOR PULSE CIRCUIT

Circuit	HF in SI Env.	Worst Delay	Average Delay	Switching Energy	Area # Trans.	SI Env. Testability	RTA14 Env. Testability	Pulse-Mode Testability
SI	Yes	1.00	0.67	1.00	42	88%	79%	n/a
RT-BM	No	0.34	0.26	0.81	32	80%	77%	n/a
RT-Agr	No	0.34	0.22	0.53	18	n/a	100%	n/a
Pulse	No	0.22	0.22	0.32	15	n/a	n/a	100%

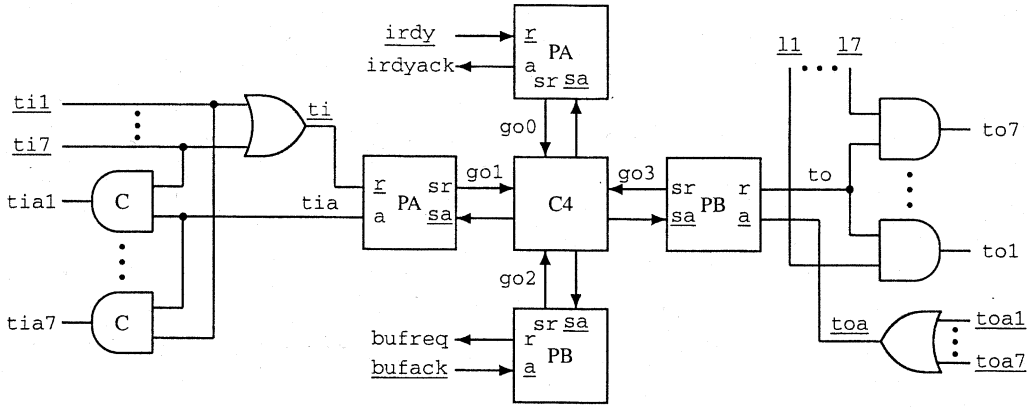

Fig. 18. SI tag unit. Assumes TAGIN (**ti**) handshakes are mutex.

also fail if the li(req) pulse is too long. If ro ↓ (ack ↓) and y ↑ have occurred before li ↓ (req ↓), then an additional pulse on ro might be generated. Therefore, arc 3 in Fig. 17 is a necessary RT constraint for the circuit to work. Finally, arc 4 is assumed to hold from RTA12, which drove this example. We therefore have a system of causal and relative timing relations that must hold in the pulse-mode circuit that directly mimic a four-phase handshake.

E. Ring Summary

Some consequences of evolving a simple FIFO-like controller from a speed-independent to a pulse-mode circuit are summarized in Table III. The different circuits are characterized in terms of performance, power, area, and testability. The worst case latency of the SI circuit is from three to five times longer than the circuits that use timing. The SI circuit is not fully testable, and the testability degrades as the circuit is placed in an environment where concurrency is restricted. The more aggressive timing assumptions tend to increase the performance of the circuits, reduce the area and power, and increase functional testability. Note that the bulk of the improvement in performance has been achieved with the simple burst-mode transformation; simple timing assumptions can often have significant impact on the quality of the circuit. The additional savings awarded by going to pulse mode are much less pronounced, except that the variation is eliminated. Indeed, the "aggressive" RT controller may already be considered a pulse-mode circuit. Power is improved for each transformation, as the pulse circuit shows a 40% reduction over RT-Agr. We feel that functional testability is increased using relative timing because many of the redundant coverings are removed when the circuits are optimized for time.

VI. TAG UNIT EXAMPLE

The FIFO ring is a simplified example used for illustration. Typically, such an application would have synchronizations coming from multiple paths. The tag unit example from RAPPID [1] shows how relative timing can be employed to generate extremely high-performance pulse-mode implementations.

Decoding of variable-length instructions is inherently a serial process, since the length of any instruction directly depends on the lengths of all previous instructions since the last branch. The performance of decoding variable length instructions directly depends on how fast this serial process operates [1]. A critical component in RAPPID is the tag unit. The tagging control signals interconnect the tag units to form a 4×16 torus, synchronizing the serial ordering of instructions by passing a tag along the toroidal rings.

Fig. 18 shows a single speed-independent tag unit. An input tag arrives on at most one of the inputs ti1–ti7. The tag is synchronized with irdy and steered to one of to1–to7 depending on instruction length l1–l7. A bufreq, irdyack, and the corresponding tia are also issued concurrently. The four-input C4 allows four processes to complete their four-cycle handshake concurrently and begin a new transfer when all interfaces are synchronized. The three behaviors in the boxes are specified as follows:

$$\text{PA} = \underline{\texttt{r}} \uparrow .\texttt{sr} \uparrow .\underline{\texttt{sa}} \uparrow .(\texttt{sr} \downarrow .\underline{\texttt{sa}} \downarrow \mid \texttt{a} \uparrow .\underline{\texttt{r}} \downarrow).\texttt{a} \downarrow .\text{PA}$$

$$\text{PB} = \texttt{sr} \uparrow .\underline{\texttt{sa}} \uparrow .(\texttt{sr} \downarrow .\underline{\texttt{sa}} \downarrow \mid \texttt{r} \uparrow .\underline{\texttt{a}} \uparrow).\texttt{r} \downarrow .\underline{\texttt{a}} \downarrow .\text{PB}$$

$$\text{C4} = (\underline{\texttt{go0}} \mid \underline{\texttt{go1}} \mid \underline{\texttt{go2}} \mid \underline{\texttt{go3}}).\texttt{sa}.\text{C4}.$$

The two passive PA processes synchronize the four-phase handshake after r requests are received, while the two PB processes are active and synchronize before handshaking. There-

fore, when the ti and irdy requests arrive and the bufreq and to cycles have completed, the ti and irdy signals will be acknowledged and the to and bufreq cycles will start. This is accomplished in the specification by renaming the signals and composing the processes as follows:

$$\begin{aligned}
\text{IRDY} &= \text{PA}[\underline{\text{irdy}}/\underline{\text{r}},\ \underline{\text{irdyack}}/\underline{\text{a}},\ \underline{\text{go0}}/\underline{\text{sr}}] \\
\text{TAGIN} &= \text{PA}[\underline{\text{ti}}/\underline{\text{r}},\ \underline{\text{tia}}/\underline{\text{a}},\ \underline{\text{go1}}/\underline{\text{sr}}] \\
\text{TAGOUT} &= \text{PB}[\underline{\text{to}}/\underline{\text{r}},\ \underline{\text{toa}}/\underline{\text{a}},\ \underline{\text{go3}}/\underline{\text{sr}}] \\
\text{BUFREQ} &= \text{PB}[\underline{\text{bufreq}}/\underline{\text{r}},\ \underline{\text{bufack}}/\underline{\text{a}},\ \underline{\text{go2}}/\underline{\text{sr}}] \\
\text{TAGUNIT} &= (\text{IRDY} \mid \text{TAGIN} \mid \text{TAGOUT} \mid \text{BUFREQ} \\
&\quad \mid \text{C4}) \backslash \{\underline{\text{go0}},\ \underline{\text{go1}},\ \underline{\text{go2}},\ \underline{\text{go3}},\ \underline{\text{sa}}\}. \qquad (4)
\end{aligned}$$

The SI implementation of these processes using ATACS is shown in Fig. 19. Processes PA and PB result in very efficient implementations. However, the large OR gates, C-elements, and necessity of passing through three state machines from the input to the output of the tag path create significant latency in this implementation.

The circuit used in RAPPID is shown in Fig. 20. This efficient circuit is very similar to the pulse FIFO (Fig. 16) derived in Section V. The extra gates are used to steer the tag paths (ti to to) based on the instruction length and to synchronize with the instruction issue buffers. The backward handshake signals in the tag path have been removed, and the forward-going signals are pulses. The request and acknowledge protocols on the irdy and bufreq paths are combinations of four-phase and pulse-mode signaling, with irdyack and bufreq being pulses

$$\begin{aligned}
&\text{2p} \quad \text{RTA13:}\{\texttt{bufreq}\uparrow, \texttt{irdyack}\uparrow\} \prec \texttt{to}\downarrow \\
&\text{2p} \quad \text{RTA14:}\{\texttt{to}\uparrow, \texttt{irdyack}\uparrow\} \prec \texttt{bufreq}\downarrow \\
&\text{2p} \quad \text{RTA15:}\{\texttt{to}\uparrow, \texttt{bufreq}\uparrow\} \prec \texttt{irdyack}\downarrow \\
&\text{3} \quad \text{RTA16:}\underline{\texttt{ti}}\downarrow \prec \texttt{to}\downarrow \\
&\text{3} \quad \text{RTA17:}\underline{\texttt{ti}}\downarrow \prec \texttt{bufreq}\downarrow \\
&\text{3} \quad \text{RTA18:}\underline{\texttt{ti}}\downarrow \prec \texttt{irdyack}\downarrow \\
&\text{4} \quad \text{RTA19:}\{\texttt{bufreq}, \underline{\texttt{bufack}}\uparrow, \texttt{irdyack}, \underline{\texttt{irdy}}\downarrow\} \\
&\qquad \prec \underline{\texttt{ti}}\uparrow \\
&\text{4} \quad \text{RTA20:}\{\texttt{to}, \texttt{bufreq}, \underline{\texttt{bufack}}\uparrow, \overline{\texttt{ba}}\downarrow\} \prec \underline{\texttt{irdy}}\uparrow \\
&\qquad \text{RTA21:}\texttt{irdyack}\downarrow \prec \underline{\texttt{irdy}}\downarrow \\
&\qquad \text{RTA22:}\texttt{bufreq}\downarrow \prec \underline{\texttt{bufack}}\downarrow \\
&\qquad \text{TAGS} = \underline{\texttt{b1}}.\underline{\texttt{ti}}\uparrow.\texttt{c1}.(\underline{\texttt{ti}}\downarrow \mid\mid \underline{\texttt{c2}}.\texttt{to}\uparrow.\texttt{to}\downarrow).\text{TAGS} \\
&\qquad \text{BUF} = \underline{\texttt{c1}}.\underline{\texttt{c2}}.\texttt{bufreq} \\
&\qquad\qquad .(\texttt{bufreq} \mid \underline{\texttt{bufack}}.\underline{\texttt{bufack}}).\text{BUF} \\
&\qquad \text{IRDY} = \underline{\texttt{irdy}}.(\underline{\texttt{b2}}.\texttt{c2}.\texttt{irdyack} \\
&\qquad\qquad .(\texttt{irdyack}. \mid \underline{\texttt{irdy}}).\text{IRDY} \\
&\qquad\qquad + \underline{\texttt{nott}}.\underline{\texttt{irdy}}.\underline{\texttt{nott}}.\text{IRDY}) \\
&\qquad \text{MUTEX} = (\texttt{b1}.\texttt{b2} + \texttt{nott}.\texttt{nott}).\text{MUTEX} \\
&\qquad \text{TAG} = (\text{TAGS} \mid \text{BUF} \mid \text{IRDY} \mid \text{MUTEX}) \\
&\qquad\qquad \backslash \{\underline{\texttt{c1}}, \underline{\texttt{c2}}, \underline{\texttt{b1}}, \underline{\texttt{b2}}, \underline{\texttt{nott}}\} \\
&\qquad\qquad \wedge \text{RTA13} - \text{RTA22}. \qquad (5)
\end{aligned}$$

The specification for the RAPPID tag circuitry is shown in (5). The processes are behavioral pulse-based specifications without

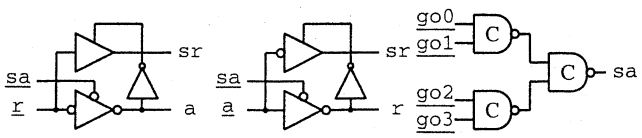

Fig. 19. Speed-independent tag unit circuits: (a) PA, (b) PB, (c) C4.

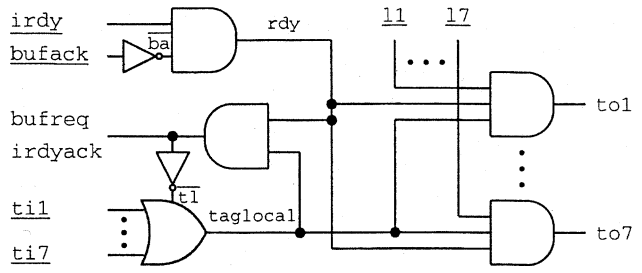

Fig. 20. Simplified RAPPID tag unit.

timing. For example, the lowering edge of the pulse signal ti ↓ and the output pulse in process TAGS are concurrent. The timing assumptions necessary to create a failure-free circuit can be classified by type according to Fig. 17. Type 4 assumptions on the ti and to signals are encoded into the specification since the TAGIN and TAGOUT processes have been combined. The synchronizations c1 and c2 encode causal transitions of type 1. RTA13–RTA15 encode type 2p transitions—minimum pulse-width constraints on to, bufreq, and irdyack. Assumptions RTA16–RTA18 are type 3, ensuring that the input pulse lowers before the output pulse. RTA19 and RTA20 are type 4 assumptions, which require the logic to stabilize before the next TAGIN arrives. Assumptions RTA21 and RTA22 simply constrain the ordering of the pulsed handshake signals. (Such constraints easily could have been placed in the specification, but have been included as RT assumptions because they are guaranteed by timing rather than by a causal relation.)

Equation (6) shows the complete set of RT constraints placed on the circuit and system for the simplified RAPPID implementation to be valid. These constraints were generated and verified through Analyze [27]. RTC23 and RTC24 are the type 2 constraints, RTC25–RTC27 are type 3 (the same as RTA16–RTA18 in the specification), RTC28–RTC31 the type 4 constraints, and type 4p RTC32–RTC33 constraints. Note that a single delay path constraint may include several RT constraints as we have used them here

$$\begin{aligned}
&\text{2} \quad \text{RTC23:}\texttt{to}\uparrow \prec \texttt{taglocal}\downarrow \\
&\text{2} \quad \text{RTC24:}\{\texttt{irdyack}\uparrow, \texttt{to}\uparrow, \overline{\texttt{tl}}\downarrow\} \prec \texttt{rdy}\downarrow \\
&\text{3} \quad \text{RTC25:}\underline{\texttt{ti}}\downarrow \prec \texttt{to}\downarrow \\
&\text{3} \quad \text{RTC26:}\underline{\texttt{ti}}\downarrow \prec \texttt{br}\downarrow \\
&\text{3} \quad \text{RTC27:}\underline{\texttt{ti}}\downarrow \prec \texttt{irdyack}\downarrow \\
&\text{4} \quad \text{RTC28:}\texttt{rdy}\downarrow \prec \texttt{taglocal}\uparrow \\
&\text{4} \quad \text{RTC29:}\texttt{rdy}\downarrow \prec \overline{\texttt{ba}}\uparrow \\
&\text{4} \quad \text{RTC30:}\{\texttt{taglocal}\downarrow, \overline{\texttt{tl}}\uparrow\} \prec \underline{\texttt{ti}}\uparrow \\
&\text{4} \quad \text{RTC31:}\texttt{taglocal}\downarrow \prec \texttt{rdy}\uparrow \\
&\text{4p} \quad \text{RTC32:}\{\underline{\texttt{ba}}\uparrow, \overline{\texttt{ba}}\downarrow\} \prec \underline{\texttt{irdy}}\uparrow \\
&\text{4p} \quad \text{RTC33:}\texttt{taglocal}\downarrow \prec \overline{\texttt{tl}}\uparrow. \qquad (6)
\end{aligned}$$

TABLE IV
COMPARISON OF RAPPID TAG UNIT WITH THE SI VERSION. CYCLE TIME OF SI CIRCUIT IS BASE CASE FOR DELAYS. AREA IS THE NUMBER OF TRANSISTORS; TESTABILITY REFERS TO THE COMPLETE RAPPID TAG UNIT AND STEERING LOGIC

Circuit	Tag Latency	Cycle Time	Cycle Energy	Area # Trans.	RAPPID Testability
SI	0.53	1.00	1.00	297	n/a
RAPPID	0.21	0.39	0.54	97	98.6%

While the circuit of Fig. 20 may be easier to verify using the metric timing of ATACS, we feel that explicitly attaching many, if not all, of the timing constraints as RT predicates makes the specification and circuit timing requirements more perspicuous. Each interface has a simple behavioral definition, which is refined by timing assumptions as predicates. Incorporating the assumptions into the specification removes much of the clarity of the resulting synchronizations and orderings. Representing the complete behavior constraints or timing constraints as a Petri-net, as was shown in Section V, can be elucidating for understanding small examples, but can be confusing and impractical for larger, real-world examples such as the tag unit in RAPPID. This is particularly the case for pulse-based implementations where the set of timing constraints can be quite large.

Table IV compares the two implementations. The RT circuit has a 3.1× area, 1.9× power, and 2.5× latency and throughput improvement over the speed-independent circuit. Since this circuitry is in the critical path of the RAPPID length decoder, the improvements in this example directly resulted in improvements to RAPPID [1]. The area impact on RAPPID from the RT circuit is arguably much higher than the transistor count comparison since this circuit is wire-limited and can be scaled. If slow parts are used, higher scaling factors must be employed to meet the target performance. If the slower SI tag unit had been used in RAPPID, the area would have ballooned significantly to meet the performance goals. The area savings in terms of the 50% reduction in wire count from removing the backward handshake is also significant. Since RAPPID tagging uses point-to-point signaling connected in a torus, removing the backward acknowledgment path resulted in a savings of 14 wires per tag unit. This reduced the network bisection of the tag logic by a total of 224 tag wires.

VII. CONCLUSION

The development of circuits requires correct operation in two domains—behavioral and temporal. Our experiments indicate that the design, synthesis, and verification of circuits can be significantly enhanced if both temporal and behavioral domains can be explicitly represented and merged. Relative timing is a means of combining behavioral and temporal information. The state space of the untimed circuit is reduced by removing unreachable relative signal orderings that are induced through time constraints.

Relative timing is a useful way of reasoning about designs. The waveforms in databooks are presented in such a way as to highlight the relation between signals and transitions. One can use relative timing to architect systems, as well as synthesize controllers, and verify the correctness of systems. Synthesis and verification algorithms can be designed to directly support this concept, where time is represented as a relationship similar to a behavioral or causal relation.

RT can be applied as aggressively or conservatively as desired. Races due to the environment in burst-mode and in speed-independent implementations due to inverter delays can be discovered and explicitly listed with the circuit. Indeed, relative timing is a superset of asynchronous methodologies such as DI, SI, and burst mode.

Relative timing does not preclude metric or absolute timing. Metric timing must eventually be applied in the implementation against the RT constraints to prove that they hold. Further, many of the RT constraints require a certain amount of slack, or setup and hold times, in the precedence relations. The robustness and reliability of the circuits can depend directly on the amount of slack on the RT constraints.

The quality of the RAPPID results in terms of throughput, power, area, testability, and latency was largely due to the timing employed in the circuits [1]. This benefit is shown through applying relative timing to the examples in this paper, and in the early tools that have formalized some of these translations.

ACKNOWLEDGMENT

The authors are grateful for the helpful and constructive comments from the referees. H. Hulgaard and S. Burns participated in timing verifications. J. Cortadella and M. Kishinevsky were the first to introduce automatic RT into the CAD tool Petrify. P. Beerel and H. Kim have been key contributors to RT verification and optimization.

REFERENCES

[1] K. Stevens, S. Rotem, R. Ginosar, P. Beerel, C. Myers, K. Yun, R. Kol, C. Dike, and M. Roncken, "An asynchronous instruction length decoder," *IEEE J. Solid-State Circuits*, vol. 36, pp. 217–228, Feb. 2001.

[2] J. Cortadella, M. Kishinevsky, A. Kondratyev, L. Lavagno, and A. Yakovlev, "Petrify: A tool for manipulating concurrent specifications and synthesis of asynchronous controllers," *IEICE Trans. Inform. Syst.*, vol. E80-D, no. 3, pp. 315–325, 1997.

[3] S. M. Nowick, "Automatic synthesis of burst-mode asynchronous controllers," Ph.D. dissertation, Dept. of Computer Science, Stanford Univ., 1993.

[4] K. Y. Yun, "Synthesis of asynchronous controllers for heterogeneous systems," Ph.D. dissertation, Stanford Univ., 1994.

[5] C. J. Myers, "Computer-aided synthesis and verification of gate-level timed circuits," Ph.D. dissertation, Dept. of Electrical Engineering, Stanford Univ., 1995.

[6] K. J. Nowka and T. Galambos, "Circuit design techniques for a gigahertz integer microprocessor," in *1998 IEEE Int. Conf. Computer Design: VLSI in Computers & Processors (ICCD98)*, Oct. 1998, pp. 11–16.

[7] D. Sager, G. Hinton, M. Upton, T. Chappell, T. D. Fletcher, S. Samaan, and R. Murray, "A 0.18 μm CMOS IA32 microprocessor with a 4 GHz integer execution unit," in *Int. Solid State Circuits Conf.*, Feb. 2001, pp. 324–325.

[8] S. Schuster, W. Reohr, P. Cook, D. Heidel, M. Immediato, and K. Jenkins, "Asynchronous interlocked pipelined CMOS circuits operating at 3.3–4.5 GHz," in *Int. Solid State Circuits Conf.*, 2000, pp. 292–293.

[9] C. L. Seitz, "System timing," in *Introduction to VLSI Systems*, C. A. Mead and L. A. Conway, Eds. Reading, MA: Addison-Wesley, 1980, ch. 7.

[10] D. E. Muller and W. S. Bartky, "A theory of asynchronous circuits," in *Proc. Int. Symp. Theory of.Switching*, Apr. 1959, pp. 204–243.

[11] S. Hauck, "Asynchronous design methodologies: An overview," *Proc. IEEE*, vol. 83, no. 1, pp. 69–93, Jan. 1995.

[12] S. Chakraborty, "Polynomial-time techniques for approximate timing analysis of asynchronous systems," Ph.D. dissertation, Stanford Univ., 1998.

[13] P. Vanbekbergen, G. Goossens, F. Catthoor, and H. J. De Man, "Optimized synthesis of asynchronous control circuits from graph-theoretic specifications," *IEEE Trans. Computer-Aided Design*, vol. 11, pp. 1426–1438, Nov. 1992.
[14] C. J. Myers, T. G. Rokicki, and T. H.-Y. Meng, "POSET timing and its application to the synthesis and verification of gate-level timed circuits," *IEEE Trans. Computer-Aided Design*, vol. 18, pp. 769–786, June 1999.
[15] W. Belluomini and C. J. Myers, "Timed circuit verification using TEL structures," *IEEE Trans. Computer-Aided Design*, vol. 20, Jan. 2001.
[16] R. Alur and D. L. Dill, "A theory of timed automata," *Theoret. Comput. Sci.*, vol. 126, no. 2, pp. 183–235, 1994.
[17] H. Hulgaard, "Timing analysis and verification of timed asynchronous circuits," Ph.D. dissertation, Dept. of Computer Science, Univ. of Washington, 1995.
[18] R. Negulescu and A. Peeters, "Verification of speed-dependences in single-rail handshake circuits," in *Proc. Int. Symp. Advanced Research in Asynchronous Circuits and Systems*, 1998, pp. 159–170.
[19] S. Chakraborty, K. Y. Yun, and D. L. Dill, "Practical timing analysis of asynchronous systems using time separation of events," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1998.
[20] J. Cortadella, M. Kishinevsky, S. M. Burns, A. Kondratyev, L. Lavagno, K. S. Stevens, A. Taubin, and A. Yakovlev, "Lazy transition systems and asynchronous circuit synthesis with relative timing assumptions," *IEEE Trans. Computer-Aided Design*, vol. 21, pp. 109–130, Feb. 2002.
[21] P. Day and J. V. Woods, "Investigation into micropipeline latch design styles," *IEEE Trans. VLSI Syst.*, vol. 3, pp. 264–272, June 1995.
[22] S. B. Furber and P. Day, "Four-phase micropipeline latch control circuits," *IEEE Trans. VLSI Syst.*, vol. 4, pp. 247–253, June 1996.
[23] S. S. Appleton, S. V. Morton, and M. J. Liebelt, "Two-phase asynchronous pipeline control," in *Proc. Int. Symp. Advanced Research in Asynchronous Circuits and Systems*, Apr. 1997, pp. 12–21.
[24] C. Myers, "Timed circuits: A new paradigm for high-speed design," in *Proc. Asia and South Pacific Design Automation Conf.*, Feb. 2001.
[25] H. Kim, P. A. Beerel, and K. S. Stevens, "Relative timing based verification of timed circuits and systems," in *Proc. Int. Symp. Advanced Research in Asynchronous Circuits and Systems*, Apr. 2002.
[26] A. J. Martin, "Programming in VLSI: From communicating processes to delay-insensitive circuits," in *Developments in Concurrency and Communication*. ser. UT Year of Programming Series, C. A. R. Hoare, Ed. Reading, MA: Addison-Wesley, 1990, pp. 1–64.
[27] K. S. Stevens, "Practical verification and synthesis of low latency asynchronous systems," Ph.D. dissertation, Univ. of Calgary, Calgary, Alta., Canada, 1994.
[28] R. Milner, "Communication and concurrency," in *Computer Science*. London, U.K.: Prentice-Hall, 1989.
[29] M. Shams, J. C. Ebergen, and M. I. Elmasry, "Modeling and comparing CMOS implementations of the C-element," *IEEE Trans. VLSI Syst.*, vol. 6, pp. 563–567, Dec. 1998.
[30] S. M. Nowick and D. L. Dill, "Exact two-level minimization of hazard-free logic with multiple-input changes," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 986–997, Aug. 1995.
[31] T.-A. Chu, "Synthesis of self-timed VLSI circuits from graph-theoretic specifications," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1987.
[32] T. Murata, "Petri nets: Properties, analysis and applications," *Proc. IEEE*, vol. 77, pp. 541–580, Apr. 1989.
[33] W. S. Coates, A. L. Davis, and K. S. Stevens, "Automatic synthesis of fast compact self-timed control circuits," in *IFIP Working Conf. Design Methodologies*, Apr. 1993, pp. 193–208.
[34] D. L. Dill, "ACM distinguished dissertations," in *Theory for Automatic Hierarchical Verification of Speed-Independent Circuits*. Cambridge, MA: MIT Press, 1989.
[35] I. E. Sutherland, "Micropipelines," *Commun. ACM*, vol. 32, no. 6, pp. 720–738, June 1989.
[36] V. Narayanan, B. A. Chappell, and B. M. Fleischer, "Static timing analysis for self resetting circuits," in *Int. Conf. Computer-Aided Design (ICCAD-96)*, Nov. 1996, pp. 119–126.

Kenneth S. Stevens (S'83–M'84–SM'99) received the B.A. degree in biology in 1982 and the B.S. and M.S. degrees in computer science from the University of Utah, Salt Lake City, in 1982 and 1984, respectively. He received the Ph.D. degree in computer science from the University of Calgary, AB, Canada, in 1994.

From 1984 to 1991, he has held research positions at the Fairchild/Schlumberger Laboratory for AI Research, the Schlumberger Palo Alto Research Laboratory, and Hewlett-Packard Laboratories, Palo Alto, CA. He became an Assistant Professor at the Air Force Institute of Technology, Dayton, OH, in 1994. Since 1996, he has been an Adjunct Professor. Since 1996, he has been with Intel Corporation's Strategic CAD Labs, Hillsboro, OR. His primary expertise includes asynchronous circuits, VLSI, architecture, hardware synthesis and verification, and timing analysis. He has received seven patents and has been the principal author for three papers that received the best paper award and has served on technical program committees for conferences and workshops.

Ran Ginosar (S'79–M'82) received the B.Sc. degree in electrical engineering and computer engineering *(summa cum laude)* from The Technion—Israel Institute of Technology, Haifa, in 1978 and the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1982.

After working with AT&T Bell Laboratories for one year, he joined the Faculty of The Technion in 1983. He was a Visiting Associate Professor with the University of Utah in 1989–1990 and a Visiting Faculty Member with the Strategic CAD Lab at Intel in 1997–1999. He is the Head of the VLSI Systems Research Center at The Technion. His research interests include asynchronous systems and electronic imaging.

Shai Rotem was born in Haifa, Israel, in 1954. He received the B.Sc. degree from The Technion—Israel Institute of Technology, Haifa, in 1980.

He has been with Intel Corporation since 1980, in positions of VLSI design and architecture of data communication controllers and microprocessors, as well as CAD design and research in formal verification and asynchronous design. He is currently a Principal Engineer in the Mobile Processor Group's architecture team, responsible for future mobile microprocessor definition.