WHITE PAPER

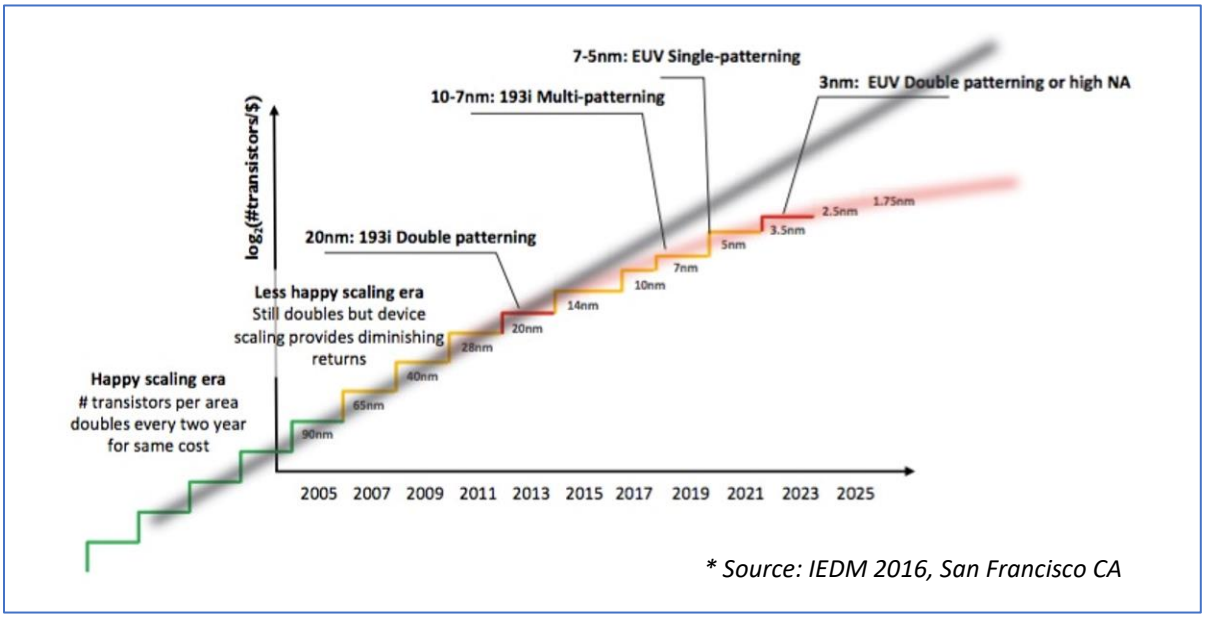
Chronos Link- A Solution to Congested SoC Interconnect

By S. Giaconi & G. Rinaldi, Chronos Tech

Introduction:

Continuing advances in semiconductor device fabrication technology have yielded a steady decline in the size of process nodes. For example, 20 nanometer (nm) process nodes were first introduced in 2012 but were quickly succeeded by 14 nm Fin Field-Effect Transistor (FinFET) technology in 2014 while 5 nm process nodes are projected for 2020.

The decrease in process node size allows a growing number of intellectual property (IP) cores or IP blocks to be placed on a single System on Chip (SoC). At the same time, modern technology requires large data throughput, forcing increased connectivity and large data transfers between various IP blocks. It is also common in modern SoC chips to include multiple clock domains, in order to leverage multi-core implementations; thus, one or more clock signals may need to be distributed across the chip in a manner that minimizes clock skew.

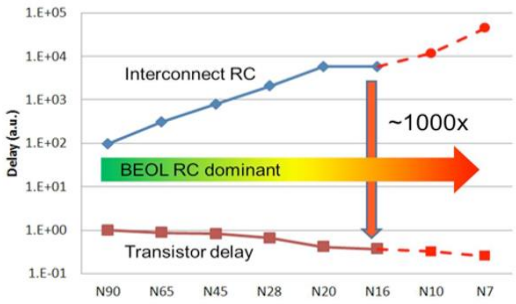


** Source: IEDM 2016, San Francisco CA*

Conventionally, a balanced clock signal distribution is achieved by inserting buffers. For example, pursuant to a traditional ASIC design flow, after floor planning and placing various IP blocks, a clock tree may be synthesized, and buffers may be added along the signal path from a clock source to various IP blocks according to the clock tree strategy. In fact, timing closure for a clock signal that is distributed over a large and complex SoC design typically requires the strategic placement of numerous buffers. Moreover, the distribution of a clock signal is also highly susceptible to both systematic and random variations. In particular, proper timing closure must account for the effects of on-chip variations that arise as a result of different Process, Voltage, and Temperature (PVTs) conditions and operation modes, which would otherwise introduce additional clock skews. As such, the most laborious and time-consuming aspect of conventional ASIC design tends to be clock alignment. Clock tree synthesis and timing closure generally require significant manual intervention. In addition, the mechanisms used to balance the clock across an ASIC chip generally consume a majority of the power in any conventional SoC design.

As transistor speed increases with technology scaling so do process variation and delay uncertainties; the consequence is an increase in number of conditions (corners) in which the design needs to be able to work to guarantee a minimum required yield for the whole chip.

On the other hand, smaller technology comes with higher timing constant for the Back-End of Line (BEOL) resulting in complex SoCs that are becoming interconnect-limited.



"As performance mismatch between transistors and interconnects continue to increase, designs have become interconnect-limited"

** IEDM 2013 by Geoffrey Yeap, IEDM 2013 Key Notes*

Design houses have also to address extremely demanding customer requirements that range from the traditional Low Power constraints to newest addition such as Data Security and 3D interconnect (TSV, etc.), bringing the already hard challenge to a complete new level.

All the above factors contribute to a constant struggle for semiconductor companies in achieving a reliable and lucrative SoC ramp-up which becomes limited by a stretched Time to Market (TTM), and a challenging minimum Yield requirement in the latest very expensive, and not yet fully stable, FinFET technology.

Chronos solution:

Chronos technology was specifically designed with the goal of enabling the next generation of complex SoC in latest FinFET nodes. It aims at enabling robust and secure on-chip and off-chip communication, while drastically reducing interconnect overheads. Such unique characteristics allow substantial area reduction, effortless integration of IPs, resilience to PVT and enhanced security; addressing the very source of the limitations in current SoC design.

Our solution is founded on the synergy of four main technologies:

- Delay-Insensitive channels
- Clockless temporal compression
- SoC compatibility
- *AccuGauge* probing technology

Delay-Insensitive (DI) channels move the validity of the data from the clock edge to the data itself using a choice of data encoding and "handshake" protocol. To practically leverage the benefits of DI channels in silicon we rely on the usage of Quasi-Delay-Insensitive (QDI) circuits. In a QDI circuit, there is therefore no need to provide a clock signal to determine a starting time for a computation. Instead, the arrival of data to the input of a sub-circuit triggers the computation to start. Consequently, the next computation can

be initiated immediately when the result of the first computation is completed. The benefits of delay insensitivity have been long known to address resilience to PVT, and with that, simplification of timing closure as well as significant reduction in TTM. In the past, this set of benefits was usually offset by the significant increase in routing requirements and area, not to mention the resistance in using a non-standard design methodology (as well as set of tools) decreasing the engineering confidence level on a successful tape-out. Chronos Technology comes with a solution to all the above concerns, mitigating the risk factors while leveraging the benefits.

Chronos clock-less temporal compression mitigates the routing and area limitation previously discussed, and, actually, moves the needle in the opposite direction. It enables the serialization of the data at higher speeds according to a specific compression ratio. This ratio can be adjusted at the time of synthesis allowing to trade routability vs link speed. Below we show the compression table for one of our technology libraries called Chronos Robust.

Chronos Compression Table

RATIO	Orig. Bus Width		Chronos Bus Width	% Diff
1	1,000	→	2,500	150%
2	1,000	→	1,250	25%
3	1,000	→	835	-17%
4	1,000	→	625	-38%
5	1,000	→	500	-50%
6	1,000	→	420	-58%
7	1,000	→	360	-64%
8	1,000	→	315	-69%
16	1,000	→	160	-84%
32	1,000	→	85	-92%

Fine tuning of compression ratio and library type can be performed on a channel by channel basis. The maximum compression ratio per channel is limited by the Maximum Equivalent Speed (MSPD) of the signaling traveling within a Chronos channel. MSPD is a parameter depending on the process technology in which the SoC is implemented.

SoC1 - High Frequency Dominant

Original bus Freq [MHz]			Original Ports	Compression Ratio	Chronos ports	% Diff	Repeater Dist. (um)
15%	@	960	7,500	3	6,263	-17%	100
25%	@	500	12,500	6	5,250	-58%	100
30%	@	240	15,000	8	4,725	-69%	145
30%	@	80	15,000	32	1,275	-92%	109
Total			50,000		17,513	-65%	

SoC2 - Low Frequency Dominant

Original bus Freq. [MHz]			Original Ports	Compression Ratio	Chronos ports	% Diff	Repeater Dist. (um)
10%	@	500	5,000	6	2,100	-50%	110
20%	@	240	10,000	8	3,150	-69%	145
30%	@	100	15,000	16	2,400	-84%	174
40%	@	10	20,000	32	1,700	-92%	870
Total			40,000		9,350	-81%	

** Assuming a MSPD of 3.0GHz*

Above are shown two different examples of routing reduction applying Chronos technology. The first one (SoC1) represents a high frequency dominant SoC such as a mobile phone or a tablet, where busses can reach very high frequencies, in this example the original 50,000 routes at top level are reduced to 17,513, bringing around 65% reduction in congestion. Compression can be pushed even further when the SoC is

Low Frequency dominant (second example SoC2) such as in IoT or microcontroller application. In this case the overall routing congestion is reduced by over 80%.

The benefits of a compressed DI channel would be greatly overshadowed if the technology was not able to interface seamlessly with standard SoC protocols commonly used in modern designs. Chronos integrates directly with the most common modern interfaces (i.e. AMBA AXI, ACE, OCP, etc.) as well as with any custom protocol used within dedicated IP and/or NOC with flow control. Chronos is also able to interface with not timing critical interfaces such as control registers, fuses, Interrupts, etc., to greatly reduce routing congestion at top level.

AccuGauge probe is the last piece of the puzzle which enables Chronos technology to assess the maximum performance of each channel on silicon on a specific PVT condition, enabling accurate tradeoff in terms of power, speed and reliability. AccuGauge enables the testing and margining of each individual channel simplifying EV, qualification, and DPM analysis. This unique probe can also be used to throttle on the fly voltage and speed while still guarantying functionality for each link.