

***July 1, 2016 – Strategy and Tactics:** The Cook's Tour: The IBM Research Semiconductor Group* ***WildPhotons**: The present is the new future. . . By G Dan Hutcheson*

Focal Points:

The Cook's Tour:

- The IBM Research Semiconductor Group
    - They spun-off its Microelectronics
        - But kept research
    - What IBM is thinking, doing, and the role they expect to play in the future
- Does this high horsepower engine have the drivetrain to move systems down the road?
    - Is this just another toothless fab-lite?
    - Or is IBM inventing a new research model?
        - IBM is often so far ahead that it looks behind
- A tour of the Albany, NY fab
- Why IBM Semiconductor Technology Research represents a new R&D Model
    - The differences between it, the fab-lite research model, and consortia models
- Why systems companies need a deep understanding of devices and interconnect
    - Reverse engineering IBM's vertical research model
    - Apple, Google, and Facebook contrasts
        - How they are evolving their research models
    - IBM's ROI on research

The Cook's Tour: The IBM Research Semiconductor Group is one of only three logic research centers left in the world that does credible broad-spectrum device research. This one's a jewel with a heritage. Older than Moore's Law, IBM's researchers have led the semiconductor industry for decades. So IBM's decision to keep semiconductor research, as it spun-off its Microelectronics division last year, created a question of conflicting track records. It was IBM's versus the dim record of other IDMs going fabless and keeping research. Fab-lite R&D efforts to stay on the leading edge have typically built a process to nowhere. Keeping semiconductor R&D internal while shifting to a foundry for manufacturing may placate shareholders, but it fails when it is an engine without a drive train.

At the same time, IBM is often so far ahead that it looks behind. Sustaining IBM's heritage of leading the industry through its technical insights is critical, not only to their partners, but also to the world. It has always been a great corporate citizen, publishing more insightful research than anyone over the decades. IBM Research was lighting the roadmap's path decades before there was a roadmap. So losing it would be a great loss.

So I was excited to get the invitation to visit. The goal was to find out what IBM is thinking, doing, and the role they expect to play in the future. That meant that this would be no normal Cook's Tour that relied mostly on critiquing what I saw. Instead, this one is more about what I heard.

There is no question that IBM researchers can be innovative. The critical questions are: does this high horsepower engine have the drivetrain to move systems down the road? Is this just another toothless fab-lite … where the R&D capability will quickly fade away? Or … as it has often done in the past … is IBM inventing a new research model?

The first time I witnessed IBM change the research model was back in 1992 when they introduced true partnering with competitors. IBM put together a $1B effort between themselves, Siemens, and Toshiba to develop next generation DRAM technology. It was the first major JDA to successfully bring top-line category, leading-edge products to market: 64Mb and 256Mb DRAMs. Moreover, they successfully overcame major cultural difficulties in getting American, Japanese, and German researchers to work together productively — something few thought possible at the time. The $1B spent collectively saved around $2B in R&D.

This was taken to a higher level, just four days after the 10th anniversary of the IBM, Siemens, and Toshiba partnership — when IBM was again at the center of partnering, with the announcement of the Albany Research Hub. John Kelly, who was director of research at IBM then, realized his vision that government and academia could play a significant role with industry in furthering the scientific research essential to the future of semiconductors while delivering significant economic benefits to the parties involved. In 2009, he would receive the semiconductor industry's highest honor, the Robert N. Noyce Award, for his accomplishments.

But technology never sleeps as it intertwines with economics to make advancements in semiconductors evermore expensive. The value is always there, but it does not come cheap. IBM had beaten the bullet longer than anyone due to its raw innovativeness and the fact that it was not an IDM … It was an Integrated Systems Manufacturer (ISM). The growing problem IBM would face in the 2010 decade was that, while the research was achievable, the scale needed to manufacture semiconductors was not.

Wafer fab ROI is extremely dependent not only on its initial cost, but also on the ability to fill it with large volumes of wafers passing through every month. And that's not just because depreciation eats up margins as most think. Fewer wafers mean fewer cycles of learning, which is as big an issue because it creates a steadily growing yield handicap against competitive fabs running higher volumes of wafers. One of the reasons why fabless companies grew so big was that the emergence of the foundry business model solved the IDM's problem of needing to fill a fab. The high fabless-to-foundry ratio meant that a foundry didn't need to find markets large enough to fill its fabs. Somebody else was finding the markets for the foundries, while the lack of a need to fill a fab meant that fabless companies could be more creative in the markets they addressed – thereby creating steady demand for large volumes of wafers. It was a virtuous relationship. But it also drove the need for even greater scale to be successful.

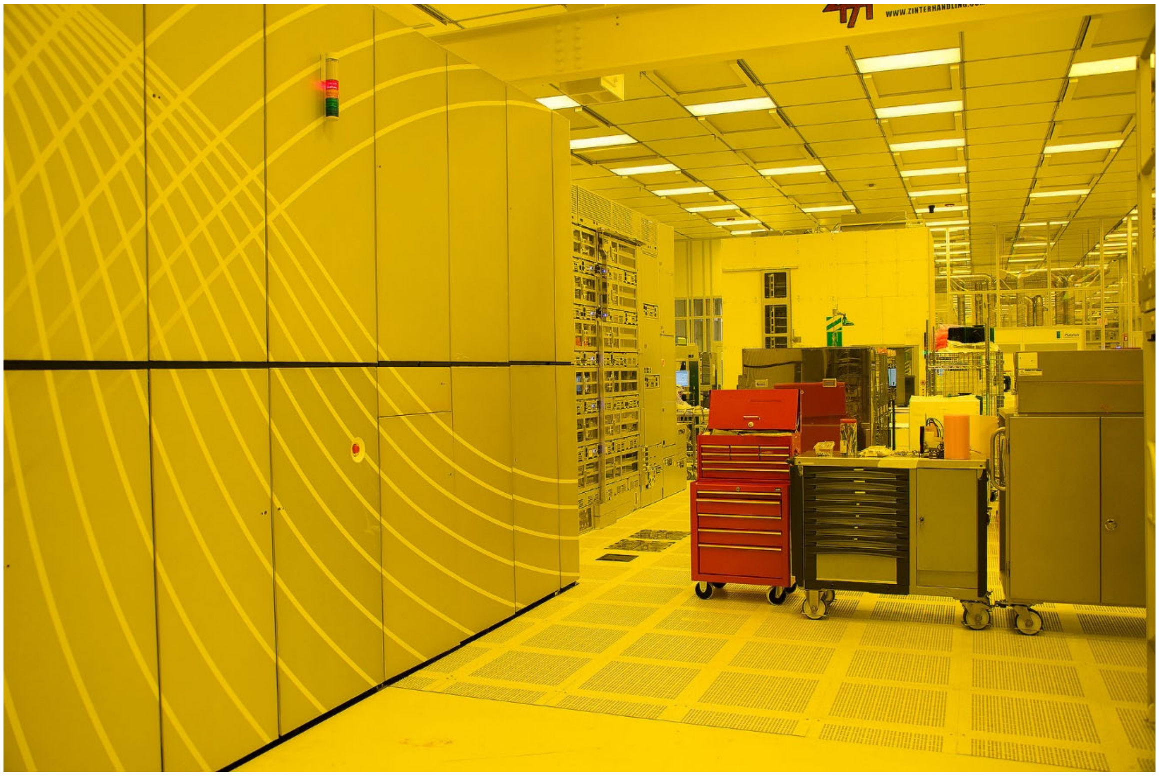
That led to what must have been the hard decision to spin off its microelectronics manufacturing capability to its partner, GLOBALFOUNDRIES. With a company as forward thinking as IBM, its choice to keep a semiconductor research group leaves open the critical question asked above: Is IBM inventing a new research model?

First can they do real research without a real[1] fab? Fab-lite versions tend to be capital constrained, as the R&D is more for financial optics than products. What's different here is that IBM is still a critical research source for its chip making partners and the systems designers within the company.

Walking around the Albany, NY fab, I would say they definitely can do real research here. IBM researchers have access to the most cutting-edge tool sets in the world via their partnership with the State of New York and SUNY. IBM's own EUV lithography tool and track have been operational here since 2014 and they were finishing their latest upgrades, for example.

The only thing missing is the automation systems that only get in the way inside a real research lab[1]. But the tool sets they are using are in fact cutting edge, thanks to the fact that so many chip equipment companies are doing research in the Albany Research Hub as well. The tools themselves are not research playthings. They are the same fully-automated tools that can be found, or will soon be found, in leading-edge manufacturing fabs around the world. So the fab runs like a small manufacturing operation. The evidence is that there are few people around the tools.

Information turns are of critical importance to learning cycles. Here, they are capable of achieving 1 day-per-mask-layer and 8 weeks for a full-flow wafer lot. In other words, this fab runs research like a manufacturing line.



Why IBM Semiconductor Technology Research represents a new R&D Model

The differences between IBM Research's Semiconductor Group and the common fab-lite research model are pretty significant. One is that the group's focus is more on research, leaving development to their partners. Partners like GLOBALFOUNDRIES and Samsung are IBM's drivetrain with the primary advantage being that they have the scale needed to continue investing in building out foundry capability for new nodes. Also important is their foundry's need for IBM's research, which is seldom the case with the common fab-lite research model. This need is both economic and technical – saving hundreds-of-millions in duplicative research, while shortening time-to-market in countless ways.

Another difference is IBM's demonstrated ability to partner well. IDMs converting to fab-lite models come from a background of being extremely secretive, which carries cultural baggage that is counterproductive to partnering. Moreover, IBM's partners already incorporate IBM's research as a critical part of their pipeline. This has not been the case with the fab-lite model, because their foundries had their own separate pipelines. Because IBM's already a part of their foundries' pipeline, they can look farther ahead without duplicating efforts. That's been the whole point of IBM's partnering efforts over the last few decades.

Key Success Factors for Stand-Alone Logic Device Research	IBM Model	Fab-lite R&D Model	Consortia Models

Use a Foundry for manufacturing	Yes	Yes	No
Stop investing in leading edge capacity	Yes	Yes	NA
Maintain leading-edge R&D facility	Mostly R	r&D	R&d
Invest in leading-edge tool sets	Yes	No	Limited
Use OPM[2] for access to tool sets	Yes	No	Yes
Established channel to transfer results	Yes	No	Limited
Ability to partner well	Yes	Maybe	Yes
Full flow process transfer to >1 Foundry	Yes	No	No
Have a system-level need for device knowledge	Yes	No	No
Ability & need to look 50 years out	Yes	No	No

Are the research results transferable? Definitely: there is really little change here other than ownership. GLOBALFOUNDRIES' fab is not far away and GF has researchers and engineers at Albany, NY Research Fab as well. Moreover, IBM has developed partnering into an art form over the last three decades. Fab-lite's results are typically a duplicative waste; since the foundry will do its own research and create its own PDKs. Consortia have an ability to transfer. But they are more project oriented with the effort driven by their customers. Customers typically hand off these projects because either they need independent validation of what they already know or they want to outsource work that will not result in a differentiable advantage. In contrast, IBM Research Semiconductor Group's contribution to a foundry's finished process will be significant, critical, and is differentiating. For example, when their foundry partners needed a finFET process, IBM had been working on the technology since the beginning.

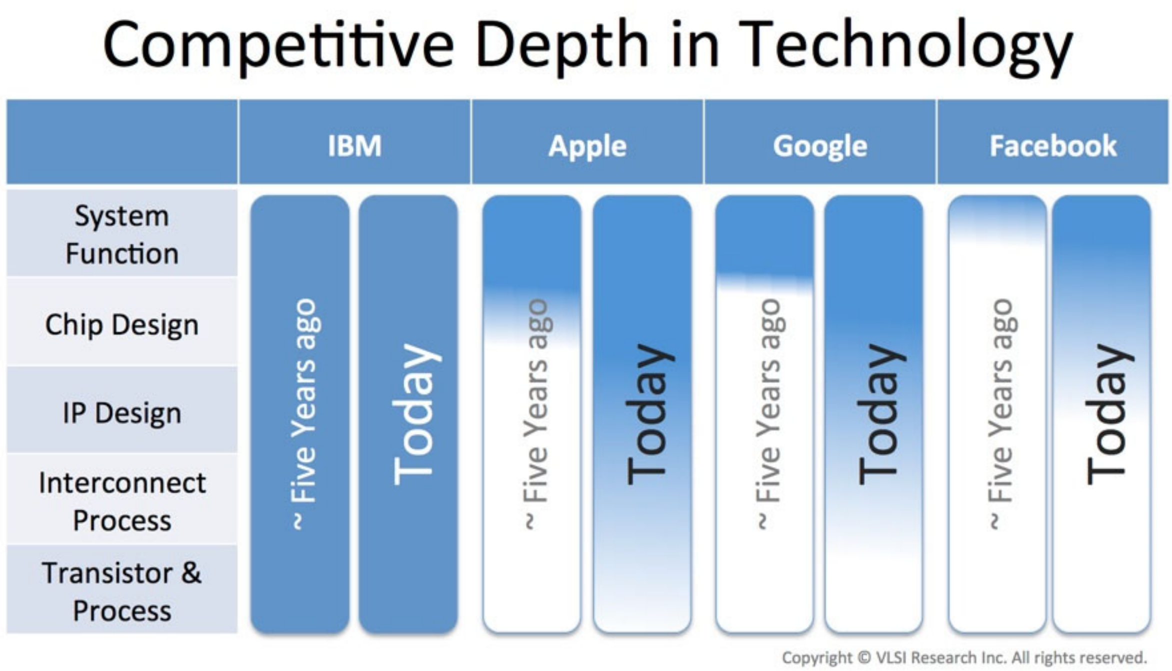
There's also a significant difference between it and the consortia models. This is an important distinction, because the value of consortia dropped significantly when consolidation created fewer chipmakers than consortia. One of the big value propositions of consortia was to do pre-competitive research for lower cost by eliminating the need for duplication. This value falls with the number of companies in the pool. Moreover, consortia are also engines without drivetrains, limiting their value. Plus, their ability to attract leading-edge tool sets varies over time, limiting their quality.

So what's in it for IBM?

So it's clear IBM Research's Semiconductor Group is essential to IBM's future and its partners. It's also clear that IBM can continue to do cutting-edge semiconductor research. That leaves the business case question to be answered. Many in the semiconductor industry would argue that there is no business case today. But the strategic moves of other companies suggest that IBM is ahead in the race.

Think about how several years ago the world's largest fabless producers, such as Nvidia and Qualcomm, began to spend significant resources to reach down into the supply chain to understand how new equipment, materials, and production techniques coming with future nodes would affect designs in development.

Then Apple bought a fabless processor company and soon after started to do the same. More recently Google and Facebook have followed in their steps, first building chip design capability and later hiring process engineers before they came out with their own chips. In short, they have been visibly reverse engineering IBM's vertical research model over the last ten years, which starts at the top of the supply chain and reaches down.



Given these strategic movements, there must be something to having an active semiconductor research capability. The question is why?

Why reverse engineer IBM's vertical research model? The reason is that system architecture advancements are constrained by the process technology advances. It is transistor and interconnect technology as well as variability in manufacturing that ultimately define the performance and functionality of the finished product. If you're fabless, like IBM now is, you also need to have the technical understanding to know which foundry is ahead, as this will determine the competitiveness of the finished product.

But it's more than just performance and functionality. It's also provides assurance that a company will successfully tapeout its first chips on a new node. Early designs on a new node are the most expensive and have the highest probability of a tapeout failure. For a company like

Apple, they risk missing a product announcement cycle for which the revenue and profit loss would far outweigh the upfront cost of early assurance. Top-Tier-Tech (T-cubed[3]) companies attempting to design chips at the leading edge quickly find this out once they start making their own chips.

Don't make the mistake of seeing this as a single design either, because it's never about designing one chip. Success is defined by a consistent ability to successfully design the next chip, the next chip after that, and so on. Otherwise, your products and your company will lose their relevancy. That means you have to look further ahead if you are a systems company. The risks for systems companies today are even greater as radically new computing architectures appear on the horizon.

IBM has over a hundred-year history of innovation. They've had to reinvent themselves multiple times. To do this successfully you have to know where the world is going long before it gets there. You must have vision with a horizon that's far out. For IBM it reaches out thirty to fifty years. On IBM's horizon are exotic technologies like neuromorphic computing, cognitive hardware, and quantum computing. Knowing those are the mountains they will climb allows everyone to align to making it possible. You can't do this from a typical foundry's vision that stretches out two to three nodes, or four to six years.

The practice of looking far ahead has been the case at IBM all along. At IBM research sites around the world, people are figuring out what future system requirements will be. This is tightly intertwined with what the technical requirements will be. They call it DCTO or Design-Technology-Co-Optimization, which is akin to DFM, but starts much earlier and covers far more ground. The task is to take these requirements and translate them into the technical elements they'll need and vice-versa. At each layer, people are looking to the future and then co-optimizing what they find up and down the chain from process to device and interconnect, to chip, to sub-system, to full system and back. Then the Semiconductor Group works on figuring out the value proposition for each of the various elements. Finally, they finish with proof of concepts for manufacturing.

The breakthrough of the 2020s will be getting past atomic dimension limits. There are plenty of possible ways to break through these limits: gate-all-around, the vertical transport FET, nanowires, nanosheets, photonics, and/or 3D multi-chip stacking, just to mention a handful. When it comes to scaling, only a few people commit to 2 to 3 nodes out and most are always in a next-node frame of mind. IBM's horizon on scaling's limit is 2032. So how do they do this?

IBM looks at the future from multiple vectors that end up as a system. They start with the application and then figure out what the system will look like – figuring out the system requirements of Watson for example. From there, they work their way back in time and down into the component details. The reason is that in order to tackle the next big application, you need more performance. The quest for performance starts in the arcane details at the atomic level of transistors and interconnect. Their quest is to take atoms and molecules to technology readiness.

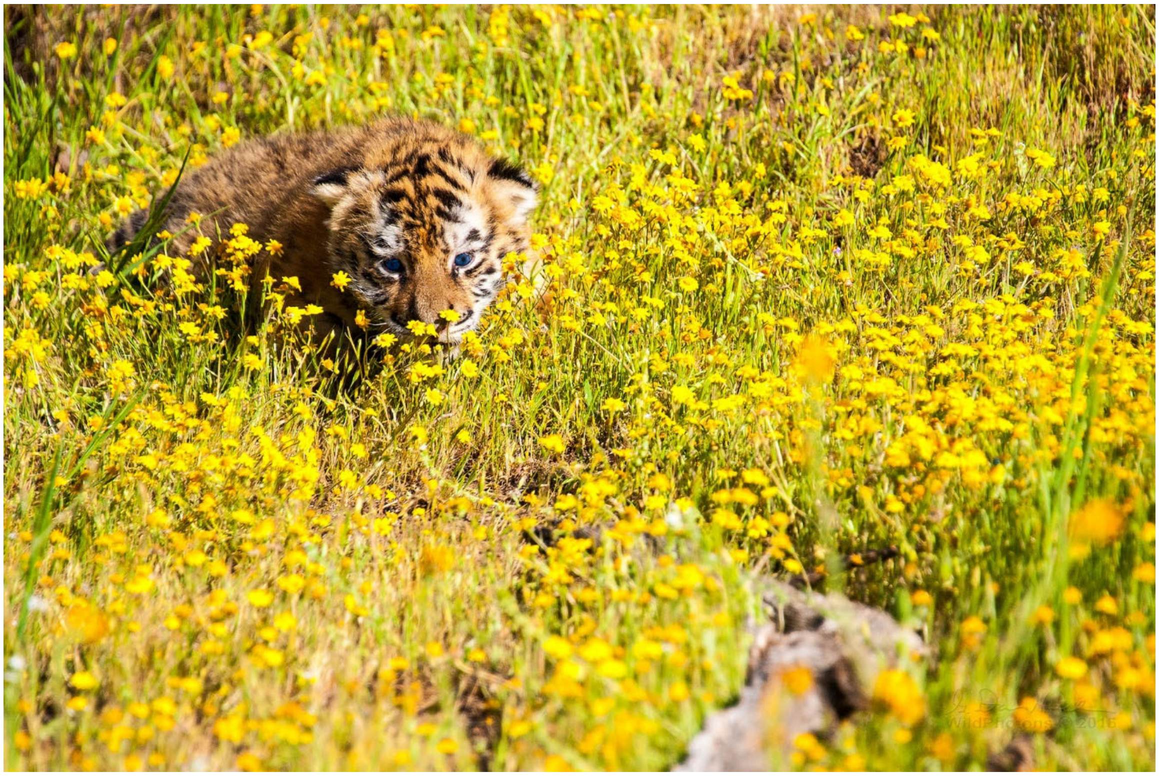
For example, the next big app is big data analytics which, will be needed to process the volumes that will come from IoT edge devices. In order to have the big data analytics, you need a system. In order to have a system, you need an architecture, which for IBM today is POWER. In order to realize that architecture, you need a processor. You also need to understand memory and interface chip requirements which will glue it all together. In order to have a processor, you need a transistor and interconnect. In order to have a transistor and interconnect, you need to understand the materials, tools, and processes that will make them possible. These affect performance parameters like contact resistance. Moreover, it's not just a single transistor in development. Systems engineers want them to figure out how to deal with multiple threshold-voltage gate stacks – each with different work functions. Fail at any one of these and it all falls apart.

So the ROI for IBM is having the technology manufacturing ready for the systems, whatever form they may take, that will be introduced in the 2020s and 2030s. The ROI is ultimately about business continuity … not about the next quarter or the next year. It's closer to be able to double-down on a 100 years of innovation to make it 200 years.

[1] A 'real fab' is one that manufactures semiconductors in volume. In contrast, a 'real lab' is one that can deliver R&D that is either pilot or full-volume production worthy.

[2] OPM: Other Partners Money

[3] While an Apple can be easily labeled as an ODM, it's harder to put this label on a Google, Facebook or Amazon – which are really Internet companies. One thing is clear is that they are on the top tier of technology supply chain, hence my creation of the term T-cubed.



The present is the new future: That where you sit, you create everything that’s going to come

Sarah Jones

A 4 week old tiger cub stops to smell the wildflowers, Panthera tigris, CA. Tigers are largest of the cat species. Males can reach a body length of 11 feet and weigh as much as 670 lb.
Reference number: TDtg_1304_394