© Mark Allen Miller

SEMICONDUCTORS

Reimagining fabs: Advanced analytics in semiconductor manufacturing

Fabs want to streamline the end-to-end process for designing and manufacturing semiconductors. Will innovative analytical tools provide the solution they need?

Ondrej Burkacky, Mark Patel, Nicholas Sergeant, and Christopher Thomas

Across industries, the application of advanced analytics, machine learning, and artificial intelligence is disrupting traditional approaches to manufacturing and operations. While semiconductor companies have been somewhat restrained in applying these technologies, that may soon change—and with good reason. Lead times for bringing integrated circuits to market have been gradually rising with each node. New design and manufacturing techniques account for some of the increase, but more complex inspection, testing, and validation procedures also create delays.

A quick look at the semiconductor value chain shows that fabs need help in multiple areas (exhibit). There has been a 50 percent increase in test and verification time during the design process over the past few years, and new-product introduction and
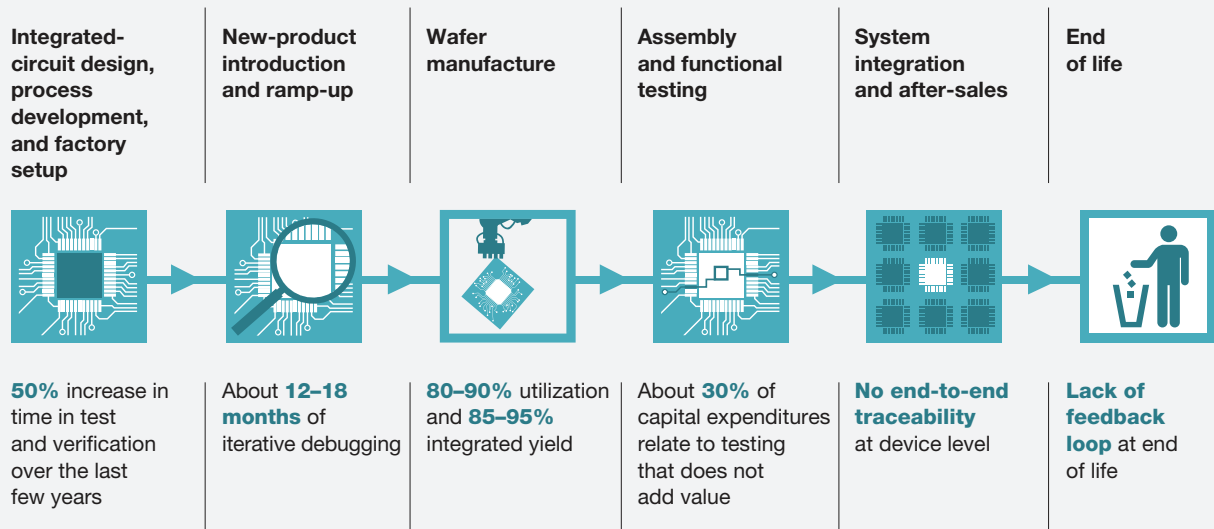
ramp-up now generally involves 12 to 18 months of debugging. Similarly, 30 percent of capital expenditures during assembly and testing relate to tests that do not add value. The problems don't stop after chips enter the market: customers may encounter unexpected performance issues and ask semiconductor companies to help resolve them—a difficult task, since there's no way to trace a chip from design through use. What's more, many fabs don't have efficient processes for recording problems encountered during production, or the steps they took to resolve them.

In many cases, problems arise because important tasks still require frequent manual intervention, despite having some degree of automation. To improve the process, many technology companies are now creating analytical tools that could help

Exhibit

Fabrication plants are struggling to manage increased complexity along the semiconductor value chain.

Current state of semiconductor value chain

Integrated-circuit design, process development, and factory setup	New-product introduction and ramp-up	Wafer manufacture	Assembly and functional testing	System integration and after-sales	End of life

50% increase in time in test and verification over the last few years	About **12–18 months** of iterative debugging	**80–90%** utilization and **85–95%** integrated yield	About **30%** of capital expenditures relate to testing that does not add value	**No end-to-end traceability** at device level	**Lack of feedback loop** at end of life

fabs replace guesswork and human intuition with fact-based knowledge, pattern recognition, and structured learning. In addition to reducing errors, streamlining production, and decreasing costs, these tools might even help fabs discover new business models and capture additional value.

Although analytical tools are just beginning to gain traction at fabs, semiconductor players already have many options from which to choose, since many technology players have recently developed specialized solutions to streamline the chip-manufacture process. We chose three companies from the large pool of innovators to serve as representative examples of nascent disrupters, interviewing their business and technology leaders to gain further insights into their capabilities. Our goal here is not to endorse companies selectively but to provide diverse examples of emerging solutions for semiconductor companies that might be unfamiliar with the new offerings.

Optimizing yield by preventing errors proactively

Advanced data analytics now offer fabs an opportunity to test and flag possible points of failure in virtual or digital-design files. Companies can then correct errors in physical designs and improve yield and reliability without running a single wafer or making a mask. Fabs can also use the same techniques to generate and run virtual and actual test chips, allowing them to identify and eliminate marginalities while simultaneously optimizing processes. Finally, advanced data analytics allow fabs to combine numerous inputs from sensor and tool data with extensive process-level information to create a rich, multivariate data set. They can then rapidly isolate and amplify possible sources of chip or equipment failure, giving them an early warning of potential problems. The tools can learn from prior designs and enhance their ability to detect failures over time. To gain more insight about new tools that may prevent errors, we spoke with Bharath

Rangarajan, CEO of Motivo, an advanced-analytics company that has enhanced the approach to predictive analytics by using proprietary algorithms, machine learning, and artificial intelligence to provide greater insight into diagnosing and preventing complex chip failures.

McKinsey: *Can you talk about some of the problems we're seeing with chip production, particularly error detection?*

Bharath Rangarajan: Each fab has thousands of process steps, which, in turn, have thousands of parameters that can be used in different combinations. With so many factors in play, we see a lot of chip failures or defects. But the frequency of each error tends to be very low, since the parameters are seldom aligned in exactly the same way during design and production. That makes it difficult for even the strongest engineering teams to predict where and when problems will occur.

Since fabs have traditionally had few analytical tools, they've tried to find high-frequency errors by making masks, running test wafers, and performing basic analytics. In other words, they changed a design or process to see if that eliminated a common error. That approach reduces some high-frequency problems in cases where only a few parameters need to be changed, but it doesn't help fabs identify low- and medium-frequency errors, which are much more common. It also doesn't identify the high-frequency errors that can only be resolved by changing numerous parameters—and those are the ones that often decrease yield.

Another problem with the traditional approach to finding errors is that it's hard to learn from past experience. As I mentioned, fabs have been able to eliminate defects by adjusting multiple parameters. That helps them with the current batch, but their tests don't give them insights about what caused the problem. By that, I mean they don't show the exact

change that produced improvement, so it's possible they may repeat the same errors in the future. Fabs have also had some communications problems that lead to errors, since many design teams and process engineers aren't accustomed to describing problems in the same way, or even sharing data, including information about past failures. I can understand why that happens—a lot of times, the design and process people aren't even located at the same site, they speak different languages, and some of them might not even know about a problem.

McKinsey: *How do your tools work?*

Bharath Rangarajan: First, we analyze a customer's physical design—typically a graphic-database system II or Open Artwork System Interchange Standard file—those are the current industry standards for data exchange of integrated-circuit layout. Our tool extracts all features and combinations, from simple geometric patterns to complex structural patterns. Then we determine how these are linked.

After processing this information, we can identify a single point, or node, of failure on a topological network map, as well as the factors that contribute to the failure. For instance, the map will show how a failed node connects to causal nodes, providing a possible point of origin. Our map also helps customers determine what features and nodes to measure and test, which helps optimize yield. That's an improvement from the current practice of randomly selecting points, and it helps increase productivity for metrology and testing. You end up with superior metrology statistics.

There's still a role for some older physics-based models in finding errors, but none of them can predict all possible complications or outcomes for advanced manufacturing processes. And they won't be sufficient as chip complexity increases.

McKinsey: *What sort of results can fabs expect in the field?*

Bharath Rangarajan: With advanced data analytics, we have the potential to alter the current paradigm dramatically. Right now, fabs run multiple batches of wafers and go through multiple costly iteration cycles to eliminate problems. That approach is also time-consuming because of the long cycles needed to process silicon wafers. If companies look more broadly at chip designs, they could reduce the lead time for yield ramps and the number of iterations required to eliminate problems with new products and processes by tenfold. That would have a big impact on timelines and silicon costs. In pilots, two semiconductor companies discovered failures and related failure modes in weeks versus a few quarters.

Enhancing wafer inspection

The inspection tools for semiconductor design and manufacture have become increasingly specialized, with their use limited to one narrow part of the end-to-end process. Fabs may need ten or more large, costly machines to accomplish the hundreds of steps that occur during wafer production, straining capital budgets and floor-space requirements. But what may be most notable are the tools' technological limitations: it can be difficult to transfer data from one device to another, import additional design layers, or program equipment to detect new errors. At many steps, manual inspectors must often review data from the tools—a process that may require the transport of hundreds or thousands of wafers to the inspection and metrology bay, increasing the risk of damage and making it impossible to capture process control and yield data in real time. To learn about new techniques for wafer inspection, we spoke with two officers at Nanotronics, a company that builds automated microscopes that incorporate artificial intelligence: chief revenue officer Justin Stanwix and chief technology officer Julie Orlando.

McKinsey: *Tell us about the use of your technology in chip inspection.*

Julie Orlando: Our microscopes combine nanoscale, micro, and macroscopic imaging with machine learning and artificial intelligence. They can find new defects automatically and share this information across the network. That eliminates the need for image tagging and other tasks that are usually completed manually and are inherently error prone. There's also a convenience factor with our microscopes, since fabs can use them for crystal growth, lithography, etching, and other processes rather than using different tools for these steps, as they've been doing. Another change is that the microscopes can inspect transparent, semitransparent, and opaque chips, as well as microprocessor units, MEMS (microelectromechanical systems) devices, and packaged wafers.

McKinsey: *Can you describe differences from manual inspections in a little more detail?*

Julie Orlando: Our microscope might analyze 100,000 chips within minutes, while a manual inspector could require 30 minutes to look at 50. Fabs can also inspect more layers if they use our microscopes, rather than manual inspections. We worked with one company that inspected 25 layers manually but increased that to 300 with our microscopes. Then there's the improvement in yield and throughput—fabs also see increases when they move from manual inspections to our microscopes.

McKinsey: *How does your software help microscopes share data?*

Justin Stanwix: Our software connects all the microscopes in the fab or fab network, so engineers can develop new algorithms to find correlations between defect-identification data and process-tool parameters. Then, they can incorporate the new

algorithm into the microscope network immediately by updating the software. Our open-software platform and API make this possible, since they allow our microscopes to connect to other tools, including those that a fab might already have.

Connecting the semiconductor and electronics supply chains

Product engineers that want to improve quality often hit an important roadblock: difficulty in obtaining data from other players along the value chain. All too often, they collect only incomplete or piecemeal information about chips within systems or applications, leaving important pieces of the puzzle missing. We discussed better strategies for sharing information with two executives at Optimal Plus, a company that specializes in software for big data analytics: Michael Schuldenfrei, chief technology officer, and Yitzhak Ohayon, vice president of business development.

McKinsey: *Tell us a little about your technology.*

Michael Schuldenfrei: We created a cross-industry platform for connecting OEMs to semiconductor companies along the supply chain. It can track all data for individual products, including where and when they were manufactured, every piece of information from functional and electrical tests, data about the equipment that manufactured them, and usage conditions—things like humidity levels or operating threshold. So, basically, engineers get an end-to-end view of information about the product and its components, making it easier to spot problems. Our platform also allows engineers to pair and match devices coming from specific manufacturing environments. That can really improve reliability in a lot of critical end-user applications.

McKinsey: *Can you tell us about how your platform works?*

Yitzhak Ohayon: For the first step, we clean and normalize data. It has to be complete, accurate, and consistent across all locations and products. Then we enter the data into the platform, where it helps overcome one of the most important data disconnects: the lack of information exchange between the chip manufacturers that conduct wafer-sort testing and the electronic OEMs—either board or system customers—that conduct final testing. After comparing data from these tests through our platform, electronics vendors and semiconductor suppliers that agree to exchange data can determine if the results are highly correlated for a particular chip—a clear signal that it will probably function well—or if discrepancies exist. For example, a chip manufacturer can use data from their electronics customers to determine which test signals predict failure downstream and which signals don't affect the final product. This means that the chip manufacturer can fine-tune their quality screens to optimize yield. In other words, improving the screens reduces the number of devices with borderline quality.

McKinsey: *What sort of results have you seen with your platform?*

Yitzhak Ohayon: In 2016, we analyzed over 50 billion chips. We've seen improvements in time required for testing, operational efficiency, yield, and test escapes.

In one case, we worked with an electronic-equipment OEM that wanted to bring a board-level design to market quickly. The company had encountered a lot of problems with chips and offered to give suppliers any board data they wanted in exchange for limited chip information. After correlating the board test data with the chip data from the original component supplier, we were able to find signatures in the chip-test data that predicted the eventual board failures. These findings reduced

the amount of time it took the customer to analyze the failures. The result was a dramatic improvement in yield and time to market.

Using the same techniques, an electronics OEM reported that it has decreased the time required to achieve acceptable shipping quality by half. It also decreased the number of "faulty" chips with no trouble found on retesting by 50 percent. Those are the chips where customers report problems, often when they're used in combination with other chips, but that work fine when the manufacturer retests them on their own. The amount of time needed to understand product failure dropped from three months to one week. The new techniques also improved testing efficiency, since the number of chips the OEM had to test intensively dropped quite a bit.

McKinsey: *Do you see any barriers to using your technology or similar technologies?*

Michael Schuldenfrei: There are some barriers related to information exchange. For this aspect of our technology to work, semiconductor companies and their customers will have to share information more freely than they do now. It may be difficult to convince suppliers to share information, since they might wonder if customers will use their product data against them during negotiations. Hopefully, our platform can reduce some of those concerns by serving as a kind of third-party intermediary between suppliers and customers. They won't have to exchange information directly, and when we share data, it's all very controlled. We only release information when problems arise—typically quality issues—and that keeps data exchange to a minimum. We've also seen some situations where electronics companies circumvent the problem by reversing the process and providing board test data to their suppliers, so that's another possible solution.

What it will take to move forward
The aspirations are clear: faster translation of product and process into the fab environment,

shorter time to market for new chip designs, lower overall cost resulting from higher and more predictable yields, and traceability through the supply chain and later use for individual chips. Making this a reality will require technical innovation that is now well within our capabilities. But in our view, the semiconductor industry also has to go further by undertaking real work on at least four dimensions.

Talent
We are experiencing a talent drought in semiconductors. Few new graduates with data-analysis capabilities name semiconductors as their chosen field, and fewer still understand the incredible opportunity to innovate and deliver groundbreaking technology improvements. Consider the situation in North America, which has fewer than 10,000 recognized data scientists. Of these, the vast majority work in or in support of a limited number of application areas, most of which involve improving the personalization of advertising or marketing content. This skewed distribution creates problems for semiconductor companies that want to apply machine learning and advanced analytics to their operations. To attract and retain the right talent, semiconductor companies will have to create compelling working environments, where data science is recognized, rewarded, and given the same respect as other technical capabilities.

Organization
Functional and organizational boundaries provide clarity, but they can also hold companies back. For instance, many fabs are struggling to optimize chip design and process technology, but they lack an end-to-end view of manufacturing processes, making it difficult to spot problems and support faster yield ramps. They need to break down boundaries by bringing design and development organizations closer together, under common leadership, to align on goals. Otherwise, even the most compelling approaches to advanced analytics may not deliver the desired results.

Investment

Engineering is at the heart of fabs and chip-design organizations, not data science. That's why fabs are only making a limited investment in advanced analytics, despite the billions of dollars at play. And when semiconductor companies do create data-analytics groups, they tend to incorporate them into the realm of information technology or manufacturing technology, rarely recognizing them as a function in their own right. This needs to change. If semiconductor companies do not significantly invest in analytical capabilities, including the application of machine learning and artificial intelligence, the sector will fall behind.

Collaboration and partnership

Analytics and machine-learning vendors are often hesitant to enter the semiconductor market. In addition to fears that the customer base is consolidating, many believe that semiconductor companies like to develop solutions in-house, alone. This perception may persist because few software or analytics companies now collaborate with semiconductor players, especially in design and operations. In the future, semiconductor players must form active partnerships with technology and research companies to prompt new ideas, applications, and ways of thinking. In the past, the semiconductor sector has enhanced manufacturing and process technologies through collaborative partnerships, such as the one involving International SEMATECH and IMEC, an international research center. It is time that we created an equivalent model for advanced analytics, machine learning, and artificial intelligence.

■ ■ ■

The semiconductor industry presents a unique opportunity to innovate and experiment with advanced analytics, since no other sector creates as much in-process data that provides insights leading to improvement along the entire value chain. Many new companies, including those discussed in this article, have recognized the opportunity and are bringing real data science to semiconductors. The use of such tools, combined with an increased appreciation for data analytics at the leadership level, could turn semiconductor companies into analytics leaders. ■

Ondrej Burkacky is a partner in McKinsey's Munich office, **Mark Patel** is a partner in the San Francisco office, **Nicholas Sergeant** is a consultant in the Silicon Valley office, and **Christopher Thomas** is a partner in the Beijing office.