Packing more into less
NANOCHIP Fab Solutions is now delivered in an environmentally friendly online version. Printed copies are available upon request. All traditions so designated or otherwise indicated as product names or services are trademarks of Applied Materials, Inc. in the U.S. and other countries. All other product and service marks contained herein are trademarks of their respective owners.

Human Factors in Automation Systems

A Letter from Ali Salehpour

Every so often it happens. Our industry reaches a major inflection point and suddenly everything changes. We experienced it when the Internet became ubiquitous and PCs drove chip demand. We experienced it again when smartphones and other mobile devices shifted chip demand. And as mobile devices become thinner and thinner, we’ve reached a new inflection point. The line between where wafer processing ends and packaging begins is now blurred, and the result is nothing less than the transformation of the outsourced semiconductor assembly and test (OSAT) industry.

Not so long ago, product packaging was a pretty straightforward proposition. The first iPhone used just 2 wafer-level packages. The latest iPhone 6 uses 26! Today’s consumers clearly want it all: thin, mobile devices that provide complex functions in a convenient form factor. It’s a challenge our industry has to meet, but how?

That’s the focus of this issue of Nanochip Fab Solutions. In these pages, you’ll learn why chip manufacturing can no longer be neatly divided between front-end and back-end processes. We discuss how OSATs are implementing fab-like advanced control processes and wafer fab equipment to deliver on the demand for more complex packaging. And you’ll learn why fan-out wafer scale packaging, through-silicon vias (TSVs), and fab automation technologies are key enablers for this industry transition.

Automation technology has wider implications for both front-end and OSAT manufacturers, and it’s another theme we explore in multiple articles. When it comes to fab productivity, increasing the yield and output side of the equation has a far greater impact than simply reducing costs. You’ll read how using automation software to optimize every production step—from planning to dispatching—can overcome fab inefficiencies and enable manufacturers to meet the demand of fast-changing markets.

An outstanding example is RF supplier Qorvo in Richardson, Texas, which we profile in this issue. By investing in fab automation software and hardware, Qorvo has empowered its workforce and increased production of bulk acoustic wave (BAW) filters 10x over the last 5 years using refurbished 150mm and 200mm tools. That growth will undoubtedly continue as demand for 4G smartphones proliferates worldwide—and Applied’s commitment to supporting legacy tools will enable Qorvo to meet it.

In addition, new Applied tools are being developed to help customers maximize the value of big data. One example, discussed in this issue, is our Next-Generation Fault Detection and Classification (NG-FDC) solution. It includes enhancements that significantly reduce setup times, improve fault detection with fewer false alarms, and leverage big data capabilities to decrease response times and increase depth of analysis.

No matter how this current inflection point reshapes our industry, however, Applied’s essential goal remains the same: to solve problems for customers, enabling their success. Whether you’re an OSAT who depends on highly reliable legacy tools, or a fab manager who needs state-of-the-art solutions, our product and service portfolio is designed for you. As mobile and wearable devices continue to slim down, we will continue to scale up—with a pipeline of highly differentiated offerings that allow you to incorporate the latest and most sophisticated technologies into your packaging and production.

Change can be daunting. But it’s inevitable. As the following articles will show, change is not something you have to face alone. Applied Materials will be there with you, through this inflection point and the next. Looking ahead to see what’s around the corner is what we do.
Often this means consolidating user experience into an interface or into an integrated data set that allows meaningful decisions. But the results of these patchwork efforts may be ineffective or inefficient, particularly when it comes to homegrown factory automation systems.

A fundamental flaw in the patchwork approach is lack of adherence to human factors in engineering. While general solutions are built on the premise that more information provides better insight and better problem resolution, the all-too-real phenomenon of cognitive saturation is given little regard in system design.

This discussion is centered on the design of an integrated workflow management system to help with the disposition of inline statistical process control (SPC) violations. But a similar line of thought can be applied to fault detection and classification (FDC) and contamination-free manufacturing (CFM). These factory processes are also examples of complex multivariate scenarios that are difficult to isolate.

**BOTTOM-LINE IMPACTS**

There are two fundamental ways in which any system can negatively impact a company’s bottom line, either directly and indirectly. The first is obvious: overt occurrences such as wafer breakage and particle contamination.

The second is more subtle, but more impactful: interpretation of signals in the fab. How well this is managed in automation system design has a direct

When you consider the design of factory automation systems today, you have to ask yourself: Is the system an extension of its users or are the people who use it an extension of the system? The distinction is important, because we often adopt the principles and lessons of previous system deployments, and allow them to shape user behavior.
impact on profit margins, because margin-eroding events, driven by scrap rates and excursions, are primarily products of human error. Efficiently identifying the underlying causes of variability in wafer processing and being able to measure cause and effect—not just action and effect—enable us to reduce the frequency of some of these problems.

Automation systems can help reduce human error by:
1. Using proper detection screens to capture events.
2. Effectively bridging communication gaps that exist within modules and across teams to facilitate the rapid and accurate communication of information to those who need to act.
3. Being sensitive to signals that something is amiss. Once such measures are in place, identifying opportunities to automate and streamline workflows more effectively becomes possible.

HOW DID WE GET HERE?

Most automation systems used in semiconductor manufacturing are homegrown efforts or a composite of internal and specialized products from various suppliers. Some components may be state of the art while others are far less sophisticated. Although the semiconductor fab has existed for over four decades, most automation systems are still designed as a patchwork of capabilities. These capabilities typically include a material execution system (MES), a recipe management scheme, SPC, equipment interface, dispatching and material-handling systems, computer-controlled maintenance management system, routines for spec and quality management, and an enterprise resource planning (ERP) system, to name just a few.

All these systems, however, have a singular supporting role: they are all used to manufacture wafers that are sold to customers. They are built on finite business rules. They tend to be predictive in some cases (with varying degrees of success) but most are often little more than interrupt-driven systems.

Human beings, on the other hand, come with an entirely different set of characteristics. Two people can interpret the same stimuli differently, yielding far different outcomes. People frequently try to fix production problems by modifying automation rules based on their individual observations and beliefs. Those individual differences may clash with the intentions of the manufacturing team.

Efforts have been made to integrate the highly methodical rule-driven semiconductor production environment with the cognitively varied understandings of the individuals who work within it. For example, the industry has adopted so-called Out-of-Control Action Plans (OCAPS) as the standard method to deal with these circumstances. But it’s not adequate. Is there a better way to bridge this gap between man and machine? For the inline SPC world, bridging the gap would require the ability to quickly make good decisions according to the wafer measurements taken over the course of manufacturing a semiconductor product.

Several foundational elements are required for such a system. To begin with, it must adhere to consistent and proper SPC practices. Without a consistent approach to SPC, additional variability would be introduced, and if a cause isn’t attributed to some source, then a falsification could become part of the dataset used to construct the control limits of an SPC chart.

For example, say a wafer is measured at a metrology tool in a fully automated facility. The metrology tool indicates a problem that points to the SPC samples that shouldn’t be considered in the calculation. In order for this to be possible, a concise, easily defined, and auditable action trail must exist that all users can see.

**Figure 2.**

**POTENTIAL CAUSES OF VARIABILITY AMONG WAFERS**

<table>
<thead>
<tr>
<th>Process</th>
<th>Incoming</th>
<th>Metro Tool</th>
<th>Process Tool</th>
<th>Recipe</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSL</td>
<td>-3</td>
<td>1</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>USL</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

This raises the following question: how well adapted are people to interpreting which sources cause the variability? Consider the inline SPC world.

Wafer measurement faults can be the product of a number of root causes; possibly more than one at a time. The user, typically a process technician or process engineer, is required to identify the possible sources of variability and adjust them accordingly.

The user has to overcome a number of hurdles at this point. There is the possibility that a violation occurred because of an incoming problem from an upstream process. This can stem from mismatched SPC, from tool matching at the process and metrology levels, from mismatched specifications, or from a simple process adjustment such as a recipe modification.

First, the user must consider the health of the current tool that processed the wafer by asking questions such as: Is the tool matched? Was there a fault on the tool? Was an experiment run on the wafer previously?

An even better approach would be for users to have the ability to modify the previous points on the SPC chart once additional data arrives confirming the actions of the first violation. In effect, the user would be able to understand whether the results of the actions taken have met the success criteria, and thus confirm that the proper decision was made. This would allow for a more accurate flagging of the SPC samples that shouldn’t be considered in the calculation. In order for this to be possible, a concise, easily defined, and auditable action trail must exist that all users can see.

**SOURCES OF VARIABILITY**

If you ask four different process experts what the violation rate (i.e., the variability) of a mature process should be, you will get four different answers. If you introduce a statistician into the mix, you will add an additional answer that will likely not match any of them. The reason for this is simple: sources of variability are not well understood nor are they well managed.

**Figure 2.**
Inefficiency also breeds a culture of “Verify that it’s not my tool and process, and then move on with life because it’s not my problem.” The result is a hesitancy to make decisions if there is a perceived risk in doing so.

HUMAN INTERPRETATION OF SOURCES OF VARIABILITY

Recently, Applied Materials looked closely at ways to handle human variability. Let’s look at the ways a human being interprets information, and also consider the overall effects of making a decision for better or worse compared to not making a decision at all.

In semiconductor manufacturing there are many faults whose effects on the overall quality of the product can’t be quantified until much farther down the line. Often, there is little evidence to relate the data to performance. Out of 160 parameters, for example, maybe only 15 have a known correlation to an electrical test parameter that affects device power consumption or reliability.

Until someone raises a flag, you as a user will learn little from your persistent SPC violations. Most aren’t correlated to electrical test parameters and there is little opportunity to sensitize a culture to them. Therefore, if you overload users with signals that have little discernable effect, they will begin to dismiss the meaning of the violation, because in this instance doing nothing is less risky than taking action.

The interpretation of users is the most important variable there is. If we were simply a matter of evaluating data thresholds, we would be able to automate everything. But again, the biggest source of variability comes from the inability of people to make proper observations in a standardized manner. Unfortunately there are no timely and effective ways to resolve this through training.

A well-designed automation system would provide the ability to audit the interpretations of its users, who tend to gravitate to one source or another depending on the understanding of the violation.

Figure 3 is an example of the variability in SPC signatures by user for a given process. Process engineer 1 (PE1) and process engineer 2 (PE2) appear to respond to the same process differently. The same can be seen for process technician 1 (PT1) and so on.

The solution is to measure this variability as a standard KPI and institute a mentoring program within the module to connect process experts with users who demonstrate a bias. The interpretations of users will become more homogeneous in a very short period of time. Continuous improvement programs (CIPs) will then allow shift behavior to change in lockstep for a given process across the entire module staff, providing measurable and more meaningful results.

Applied Materials has invested in a number of solutions that target human-error and variability reduction across inline SPC, including tool qualification, fault detection, defect and contamination management, and electrical test. Field results have demonstrated a 4% reduction in human variability across all modules in less than 3 months.

This approach is closely coupled with user training that focuses on problem-solving with a new dimension. It also integrates the failure and response-detection methods, in this case inline SPC, with the equipment and process failure modes and effects analyses (FMEAs). By reducing human error and variability, fabs can rapidly roll out methods of mitigating future scrap events. Their automation systems become extensions of the people who use them, bringing new levels of efficiency to the fab, while preserving the human factor.

For additional information, contact selim_nahas@amat.com.
Driven by fast growth in smartphones, RF supplier Qorvo Inc. has been expanding production of BAW filters at a rapid rate. Howard Witham, vice president of Texas operations, sat down with Nanochip Fab Solutions at the company’s Richardson, Texas, fab to talk about how fab automation tools have helped an experienced workforce deal with rapid capacity expansions.

When RFMD and TriQuint Semiconductor merged at the beginning of 2015, it brought together two synergistic companies—each with roughly a billion dollars in revenues—and created a merged entity, Qorvo, Inc., which may approach $3 billion in sales this year. Qorvo’s rapid growth is reflected in expansions at all of its manufacturing operations. The company’s Richardson, Texas, fab has increased production of bulk acoustic wave (BAW) filters by 10x in the last 5 years, and demand is expected to continue to increase as smartphones using 4G wireless networks proliferate around the world.

Brad Shaffer, senior analyst for mobile electronics at IHS in Tempe, Arizona, said BAW filters, already one of the very high growth areas in smartphone components, will gain further momentum as LTE wireless networks in China come on line.

The LTE networks operate at frequencies above 1.5–2 GHz where BAW filters have performance advantages over surface acoustic wave (SAW) filters. “As LTE gets closer to where Wi-Fi fits into the spectrum, there are possible interference issues,” Shaffer noted.

Avago Technologies and Qorvo share about 95% of the market for BAW filters (see figure 1), and are the only BAW suppliers capable of meeting the huge volumes that smartphone vendors Apple and Samsung require. “Avago and Qorvo are both investing a lot of money to meet demand,” Shaffer said, noting that Avago has stated it may be late 2016 before it fully catches up to demand for its filters.

“BAW is growing significantly and at a very fast pace, and that is where you are seeing some of these utilization issues,” Shaffer said.

Howard Witham, the Richardson site manager, said Qorvo has invested about $220 million over the last 2 years in the Richardson fab. About 400 employees were added to the manufacturing staff, which raised employment there to 1,200. The Richardson team has brought in about 40 tools for a pilot 200mm line, and expanded the 150mm line by 3x in the last 2 years.

Qorvo’s Richardson fab has two quite different product lines: BAW filters and power amplifiers made on gallium arsenide (GaAs) and gallium nitride on silicon carbide (GaN on SiC) substrates. The compound semiconductor line is largely used to make high-power RF amplifiers for infrastructure (base station) and defense customers.

“We are expecting a big pull in 2016. As the fab gets filled out, how do you scale? We can look to a 200mm line for the BAW filters, and are running some 200mm now. Over the last 5 years, it has been mandatory that all of our 150mm equipment be convertible to 200mm,” Witham said.

TriQuint bought the Texas Instruments defense electronics business in 1998, and soon after moved its Texas operations into the Richardson fab, which had been making DRAMs for a Hitachi-Texas Instruments (TI) joint venture called TwinStar. Many of Qorvo’s current workers had previously worked at the TI gallium arsenide-based defense electronics line.

“The people from the TI defense group were outstanding. They had a healthy chip on their shoulders, and were very eager to show that they had the capability to move into a high-volume manufacturing environment on a commercial product,” Witham said.

When he moved from managing fabs for STMicroelectronics to TriQuint in 2010, the company was in the early stages of BAW production. The Richardson fab’s GaAs and GaN-on-SiC defense and infrastructure line used only about one-quarter of the 50,000-square-foot fab, processing about 200 wafers per week.
The company quickly began ramping BAW filter production as smartphone sales zoomed up. In those early days of 2009 and 2010, the challenge was to keep up with demand. “We had yield and reliability challenges, moving a young technology into a robust manufacturing state,” Witham said.

One answer was to invest in fab automation, including advanced process control (APC) and a new manufacturing execution system (MES).

**PAINFUL EXCURSIONS**

At last year’s Advanced Process Control (APC 2014) meeting in Ann Arbor, Michigan, Witham detailed how he brought modern process control techniques to Richardson.

“My feeling at that time [2010] was that we were about 15 years behind, and that we had about 2 years to show some progress,” Witham said. He recalled coming into “an environment where every day our MES was crashing."

One challenge was to convince management to release the capital expenditures needed to modernize the fab with new factory automation hardware and software. “We were begging for APC and process control, for improvements in our MES. Our management said ‘OK, if you want a new MES, you have got to commit to a return on that investment. You have to put a dollar value on it,’” Witham said.

The TriQuint senior managers were receptive because they had gone through several “extremely painful” excursions that resulted in million-dollar financial losses. Problems—ranging from a helium leak in a cryogenic pump to fluctuations in CMP carrier speeds to an irregular argon supply—were slowing down the yield ramps at Richardson.

“The engineering team started documenting things that they were guessing at before. The team scrubbed its quality database and found that about 8% of the scrapped wafers were preventable. “We figured out how to do automated optical inline inspection in a fab that was not looking at anything inline. Nothing we did was novel, but I can tell you that we got huge benefits very quickly,” Witham said, adding that “matching equipment becomes possible when you have these tools.”

**RIGHT TOOLS FOR THE JOB**

Initially concerned that the Richardson engineers’ experience with defense ICs would prove a handicap, Witham found the opposite to be the case. “I thought I would have a fight on my hands, but they were so ready. They tell me now that if I ever wanted to reverse directions on these automation tools, that I had better watch it,” he said.

Process control is a big reason beyond the sheer boost in capacity with the larger wafers. Witham said as Qorvo moves from “vintage” 100mm equipment to 150mm tools, it can support improved recipe downloading, easier operation for the direct labor force, and data capture.

With the newer 150mm refurbished tools, “the engineers can see what is happening with the equipment. Operators don’t have to make decisions with poor data; we can set up algorithms to make the decisions for them,” Witham said. The 150mm compound semiconductor line paves the way for what could be a boom in GaN-based power amplifiers, as customers beyond the defense sector turn to GaN for improved power efficiency. GaN is increasingly used in LEDs, led by Cree, Sumitomo Electric, and others. And power transistors are taking advantage of the increased power densities and heat dissipation with GaN. A U.S. power electronics research consortium, based near Albany, New York, and led by General Electric, is now getting underway.

“We see more GaN in the future,” Witham said, noting that there is a “lot of pull from our customer base. We have a very competitive GaN amplifier product line now and over the next two or three years we see that potentially really taking off in the commercial space.”

**BAW FILTER BOOM**

While Qorvo’s compound semiconductor production line continues to thrive, much of its growth has come from making BAW filters. They are made on a silicon wafer that acts as a simple carrier, with nothing active in the wafer. Above the silicon, Qorvo deposits aluminum nitride and other films, which form a piezoelectric layer that acts as an acoustic resonator, filtering out unwanted frequencies. The goal is to pass on 100% of the in-band frequency. “This is not like GaAs or silicon where the active component is in the wafer. What we do is all up above the silicon. With these acoustic resonators, the piezoelectric layers are not easy to do,” Witham said. At the high frequencies in which BAW filters are effective, the piezo layer must be only hundreds of nanometers thick, and the acoustic Bragg reflector is created by stacking thin layers of alternating stiffness and density.

Physical vapor deposition (PVD) is a key process step, and finding used tools has been a challenge, forcing Witham to buy some new PVD equipment at higher prices. (See “Demand for 200mm Tools Outstrips Supply,” Nanochip Fab Solutions, Vol. 10, Issue 1, 2015.)

Howard Witham, vice president, Texas operations, at Qorvo.
Witham said that during the recent 9x production expansion at the Richardson fab, the company’s suppliers did not have a shortage of cores, although some of the tools were in “touch shape” before refurbishment. Buying largely brand-new equipment is not an option given the company’s wafer cost targets.

**FROM LEMONS TO LITHOGRAPHY**

“We did not feel the shortage [for used cores], and maybe that’s because this RF market is now big enough that we get more attention from the suppliers. I felt the biggest pinch on the steppers. We found some lithography tools on the used market, but we got some real lemons. Fortunately, we have at least two former litho vendor field service reps working here at Qorvo, and if you can give them a frame, they can build you a stepper. That’s the dice you are rolling. That’s what can be the difference [between success and failure].”

The equipment refurbishment team has worked with equipment OEMs, including Applied Materials. Qorvo also relies on local refurbishers who work hand-in-hand with the in-house equipment technicians. By mixing large OEM and local refurbishment suppliers, and converting 150mm tools to 200mm, Witham believes he will be able to add enough capacity while staying within his equipment cost targets.

**FAST AND SLOW YIELD RAMPS**

One challenge Qorvo faces is how to meet the demands of the fast-changing smartphone market, where the RF modules are upgraded regularly. A BAW filter might be in production for only 12-14 months before production switches to a new design, and profitability depends on winning key sockets and then meeting demand with a fast yield ramp (see figure 2). Witham said smartphone makers factor in manufacturing capacity as one criterion in their component selection process. “If I have manufacturing clout compared to competitors, that’s a big plus for the smartphone manufacturers. They have to believe you have put in the necessary capacity. At Qorvo, we have to bet early. We can still be under capacity, but then we are scrambling to put in new capacity and scrambling to get every bit out of our existing tools as we work with our supply chain to put in new capacity.”

The key to success is not being overloaded, which can put the fab into overdrive mode. That can result in quality issues, while keeping volumes high and wafer costs low. “Volumes are everything when it comes to the best financial performance. Semiconductors are a very heavily capitalized industry, and we want to keep this fab full so we can contribute to a good P&amp;L.”

At CS MANTECH, a recent compound semiconductor manufacturing technology conference, Witham detailed the profit difference between slow and fast yield ramps, with a 4.45% difference in margin impact over the product’s relatively short run. In his keynote speech at the conference, Witham said that “the fab must have excellent speed for NPD (new product development) prototypes during development and also for fast engineering learning cycles when working to improve yields.”

Witham explained the difficulty. “You have to win [the BAW filter sockets] every year. The competition is fierce to win sockets, with new specs, and more stringent limits on battery power. Every year you have to win at all the big smartphone makers.”

**A SIGNAL IS NEEDED**

Witham added that “if you think the company will win something big, you want to learn about your win within your capacity lead time. You have to get some signal in order to make a $200 million investment, and our customers understand that. The equipment manufacturers are not going to sit there with their own supply, waiting for us, and it is the same for all of us in the [production] chain.”

Suppliers must have a certain size, and a willingness to play in a smartphone game where decisions are necessarily made late, with precious little spare time to add capacity. “You tend to go late, and when you do, you have got to be firing on all cylinders. You cannot afford any significant hiccups. You can’t hire people too soon, and you can’t buy equipment too soon,” Witham said.

One plus has been the availability of a skilled workforce in the Dallas area. Several IDMs have either closed their fabs there or reduced their footprint. Along the way, they have learned to use new productivity tools to get the most out of a fleet of mostly refurbished tools, coping with steep ramps of a truly unique product.

For additional information about Qorvo, Inc., visit www.qorvo.com.
If you’ve ever thought how convenient it would be to have a mobile device with a large display that you could fold and put in your pocket, you’re not alone. A recent Applied Materials survey\(^1\) showed that as many as four out of five people feel exactly the same way.

That attitude is reflected in the strong growth anticipated in the markets for flexible and curved displays over the next several years. Market analysts at Touch Display Research, for example, forecast that flexible and curved displays will achieve 16% of global display revenues by 2023, compared with 1% in 2013.\(^2\) IHS iSuppli, meanwhile, projects that global shipments of flexible displays will reach 792 million units in 2020, up from 3.2 million in 2013.\(^3\)

Some curved displays, such as Samsung’s Galaxy Edge smartphone with its beautiful curved edges, are already available. But what technology will be required to create flexible, bendable displays in new form factors, and with adequate resolution, display quality, and ruggedness? And most importantly, what technology will do that at costs that enable display and end-device manufacturers to enjoy attractive margins?

**MATERIALS ARE A KEY ENABLER**

Materials have a key role to play in the advance toward flexible displays for mobile devices, televisions, advertising signage and other uses. A major materials challenge is how to move away from rigid glass encapsulation without compromising the operation or reliability of the display’s thin-film transistor (TFT) backplane when the device is flexed, folded, or bent.

For example, it’s very difficult to maintain the required cell gap between the color filter and the TFT backplane when the display is flexed. In this regard active-matrix organic light-emitting diode (AMOLED) display technology is promising, because the rigid glass encapsulation otherwise required to protect the organic material from moisture and air can be replaced by layers of thin films.

But it’s important to note that this isn’t the only challenge. For any OLED display to get to mass production, especially in large sizes, all these challenges must be addressed whether it is a flexible display or not: electroluminescence (EL) evaporation and its impact on the lifetime of organic materials; low deposition efficiency; low yield because of defects; and the scalability of evaporation technology, which directly impacts the cost of volume production.

Early adopters of flexible OLED displays have tended to use a delamination process to fabricate the transistor backplane on a flexible thin film. First, the thin film is weakly bonded to a glass substrate and run through typical OLED processing steps to build the transistors. Then, once processing is complete, the thin film that contains the transistor backplane is delaminated, or removed, from the glass substrate and used to build the flexible display.

However, this delamination process is complex and costly. Therefore, its use in high-volume manufacturing without further technological breakthroughs is questionable.
THIN-FILM ENCAPSULATION (TFE)

The primary goal of many display manufacturing technologies is to achieve uniformity and performance stability for both the TFT backplane and the OLED emission layers. Effective encapsulation with layers of thin films is critical to prevent the degradation of AMOLED displays caused by moisture and particles (see figure 1). The encapsulation directly affects the life span and lighting performance of the AMOLED device.

Applied Materials offers thin film encapsulation (TFE) processes that support the manufacturing requirements for flexible OLED displays (see figure 2) while providing key barrier protection.

Applied Materials’ AKT TFE product line includes a multilayer solution (see figure 3) that extends the lifetime of flexible OLEDs. The multilayer concept reduces water permeation by decoupling defect sites in the barrier films and increasing the permeation channel length.

It also combines diffusion barrier films made from SiN that have very low water and oxygen penetration (see figure 4) with HMDSO buffer layers (for OLED mobile applications) that release stresses in the film stack and isolate any particle contamination from upstream processes (see figure 5). Particles lead to dark spots and delamination issues.

These high-performance films, deposited at <100°C, address the tendency of OLED materials to degrade when exposed to the outside environment. In addition, the AKT TFE systems’ unique vision-alignment technology ensures accurate and precise mask positioning and deposition. This allows display manufacturers to eliminate photolithography andetch process steps and reduce production costs.

Samples of HMDSO buffer layers typically demonstrate high optical transmittance (>95% at 300nm and above) and low stress. The samples passed the device lifetime test (SiN/HMDSO 3 layers) and passed 100,000 test cycles with 1” diameter bending. Further, without any high-stress points, this buffer layer provides excellent particle coverage and leaves no voids or diffusion channels.

Applied Materials’ TFE tool architecture will be based on a cluster tool design to facilitate high-throughput multilayer deposition without breaking vacuum (see figure 6).

COLLABORATION IS A MUST

The development of flexible, bendable displays brings many technical challenges that require advances in manufacturing processes and materials. To implement these new technologies as rapidly as possible and at a cost that can drive mass consumer adoption, tight collaboration between equipment and materials suppliers, and panel and device makers, is a must.

By leveraging its expertise in precision materials engineering, Applied Materials is helping to solve these technology hurdles so that customers can make flexible displays a reality.

For additional information, contact kerry_cunningham@amat.com.

Manufacturing complexity in the outsourced assembly and test (OSAT) industry is exploding, driven both by demands for higher functionality, thinner form factors, and longer battery life in handheld devices, and by competitive dynamics (see figure 1).

On the technology front, OSAT factories are moving into more complex packaging technologies that blur the line between where wafer processing ends and packaging begins. In order to meet the challenges of 2.5D and 3D wafer-level architectures, they are becoming more like modern wafer fabs. This means traditional packaging methods are no longer necessarily relevant, and that fab-like characteristics, life cycles, and activities will have to be learned and deployed in the OSAT environment. Fab automation technologies are a key enabler for this transition.

Return on investment can be fast, too. One user realized a payback in just a few months from improved throughput and MES transaction automation.

BY SHEKAR KRISHNASWAMY
In fact, it’s no longer possible to neatly divide the manufacturing of handheld devices such as smartphones and tablets into the traditional categories of front-end-of-the-line (FEOL), where transistors are fabricated, and back-end-of-the-line (BEOL) where interconnections, packaging, and assembly take place. A so-called mid-end-of-the-line (MEOL) approach (figure 2) is evolving that has both FEOL and BEOL (figure 2) is evolving that has both FEOL and BEOL characteristics. It came about because in vertical architectures the interconnections between layers, such as through-silicon vias (TSVs) and bumps, must be built using FEOL tools and processes. Companies in the OSAT industry, however, have traditionally provided relatively low-margin, commoditized test and packaging services in support of BEOL requirements. To remain competitive in the wafer-level packaging (WLP) era, they must now provide their customers with higher-level engineering resources and fab-like manufacturing capabilities.

In addition, advanced packaging applications are growing faster than the semiconductor industry and carry relatively high margins. So new, formidable competitors are now encroaching on OSAT turf.7 Several leading wafer foundries and integrated device manufacturers (IDMs)—seeing MEOL as a profitable extension of their existing capabilities—either already have, or are building, wafer fab-like packaging facilities to address it. One way OSAT companies can overcome these technological and competitive challenges is through greater use of automation to reduce errors and waste, provide increased flexibility and responsiveness, and drive higher levels of output.

Monitoring and controlling such complex production operations is orders of magnitude beyond what OSAT companies have had to do in the past. Before, they could operate their factories with spreadsheet-based applications. Now, WLP requires much higher levels of statistical data analysis and precise, automated control of equipment, processes and factory operations, similar to wafer fabs.

Realistically, meeting these much more complex requirements can only be achieved by implementing modern automation and control strategies (see figure 3). At the equipment level, the essential capabilities needed for WLP competitiveness are automated recipe management (ARM) to reduce human errors and achieve better yields, and integrated fault detection and classification (FDC) to improve equipment availability and reduce scrap.

At the process level, statistical process control (SPC) is needed for faster data analysis and better quality. Advanced process control (APC) solutions can lead to higher yields and reduced scrap. At the factory level, automated real-time product dispatching is a key enabler for high-volume production (see figure 4). The need for real-time dispatching occurs when a tool is available and lots are in a queue waiting to be processed.

Software such as Applied’s APF Real-Time Dispatcher (RTD) can determine which lot to process first to achieve the overall highest throughput. It takes into account factors such as equipment capability, lot due dates and priorities, desired cycle time, equipment setup and maintenance requirements, and ancillary resources such as masks or reticles.

Figure 2: TSV-MEOL/BEOL process in overall TSV process flow. (Source: ICEP-IAAC 2012 Proceedings; STATS ChipPAC-SW Yoon et al)
AUTOMATION HELPS OVERCOME PRODUCTIVITY CHALLENGES IN WAFER-LEVEL PACKAGING

Figure 4. Uses and benefits of real-time dispatching.

Uses of Real-Time Dispatching
- Process repeatability
- Automation of best practices
- Optimization of manufacturing batch size, minimizing of setups, lot sequencing
- Load balancing
- Variability reduction
- Reduction of WIP bubbles
- Management of plan changes

Results
- Increased fulfillment of orders
- Reduced overall lot cycle time
- Reduced cycle time variability
- Manpower redaction
- Improved linearity, WIP balance and tool utilization
- Increased product throughput
- Increased operator conformance

High-Volume Manufacturing

Short-interval scheduling automation strategies can help OSAT companies further increase overall productivity by eliminating process and equipment inefficiencies known as “white space.” White space is the term for small gaps in processing that, in aggregate, sap factory productivity. To eliminate white space, a comprehensive and realistic look at anticipated factory production is required. Short-interval scheduling can accomplish this. The scheduling methodology is based on the processing of large amounts of good data, highly realistic mathematical modeling of factory operations, and the assumption that tasks will be executed according to prescribed schedules.

A FOUNDATION FOR THE FUTURE

Automated solutions provide OSAT companies not only with a framework for flexible, high-quality, high-output, and profitable production today, they also provide a foundation that can accommodate future changes in production strategies.

An example is the introduction of mobile technology into WLP factories to enhance manufacturing productivity and quality. One Applied Materials customer, a global leader in flash memory storage solutions for a wide range of applications and devices, implemented mobility-based solutions to increase productivity, quality, and reliability in one of its Asian assembly and test facilities.

A large proportion of that company’s products were manufactured using 2.5D and 3D packaging technologies, applying factory automation techniques such as manufacturing execution systems (MES), statistical and advanced process controls, equipment automation, and advanced scheduling solutions. However, the effective use of these techniques was hindered by reliance on manual methods and procedures. In such a complex environment, this can lead to procedural errors in the execution of the manufacturing process, resulting in product scrap and quality issues.

In this particular factory, the customer’s manufacturing operation involved processing hundreds of part numbers at multiple machines under different conditions. The processing parameters for each operation and each part-type were encapsulated into a process recipe, and hundreds of different recipes could be qualified on any particular machine. Using an incorrect recipe invariably caused product scrap, leading to higher manufacturing costs and potential customer satisfaction issues.

A key opportunity for scrap reduction and quality improvement came from the deployment of an automated recipe management system (RMS) that extended to mobile devices. This eliminated human processing errors caused by using incorrect process recipes. The application was installed on both tablets and smartphones.

A user interface was provided to capture the sequence of automated and manual events. Initially this was meant only for configuration and diagnostics, but it provided enough value that it eventually became the universal operator interface, enabling rapid user training.

Alarm management was a key component of the application. Until the deployment of the RMS, operators were performing many unnecessary steps. The RMS alarm management function captured all these events along with their frequencies, and the information was used to educate operators about the need to eliminate unnecessary steps.

Some of the key results seen from the first phase of the project were:
- Product cycle time reduction and fab throughput improvement
- Streamlined equipment monitoring and alert management
- Dramatic reduction in product quality incidents related to equipment unit issues
- Increased operator efficiency thanks to mobile access to key transaction information right next to their machines

In addition, throughput improvement and MES transaction automation alone provided a payback for the first phase of the project in just a few months per machine.

CONCLUSION

The need for WLP is growing strongly, and is introducing great complexity into the operations of OSATs. Successfully addressing this complexity requires the use of innovative automation strategies to increase manufacturing flexibility, efficiency, and quality to meet customer demands and remain competitive.

For additional information, contact shekar_krishnaswamy@amat.com.

MOBILITY DRIVES THIN PACKAGING

Shaving a fraction of a millimeter from the height of chip packages is a big deal in smartphones. The need for thin devices is dictating the evolution of chip-scale packaging.

For the packaging industry, the smartphone changed everything, and experts agree that more changes—potentially big ones—are underway.

Bill Chen, a senior fellow at ASE Group, said when personal computers ruled chip demand, the main components had fairly standard package formats. The pace of change has picked up in the smartphone era. Now, “the people who design smartphones compete by the functions they can provide, and probably more importantly, by how they look. You want them to be thin, and you want all the functions.”

That need for thin packages drove the industry to wafer-level chip-scale packages (WL-CSP), a hybrid between front-end and back-end processes in which solder balls are created beneath the chips before the wafer is diced. ASE estimates that about 30% of the die in smartphones now are in WL-CSP packaging, nearly all in a fan-in configuration in which the solder balls are spaced out directly underneath the die.

Jan Vardaman, president of the packaging consultancy at TechSearch International (Austin, Texas), said Apple’s first iPhone used 2 wafer-level packages (WLPs) in 2007, but today’s much-thinner iPhone 6 has about 26 WLPs. “Several years ago, Samsung did not use WLPs in its mobile phones,” Vardaman said. “Today, around a dozen WLPs are found in the latest Samsung smartphone. Even the latest portable communication gadget, the Apple Watch, contains as many as a dozen WLPs.”

Traditionally, WLPs have been used for devices with low pin counts and small die sizes. The number of chips packaged in fan-in WLPs is still

According to industry research firms Yole Development and System Plus Consulting, approximately 30% of the chips in 3 of the market-leading smartphones use wafer-level chip-scale packages.

“I would say the back end is no longer at the back end. We are at the front of the whole effort of getting semiconductors into systems.”

— Bill Chen, senior fellow, ASE Group
MOBILITY DRIVES THIN PACKAGING

Future Evolutions in Discrete Device Packaging
WLP: Different Levels and Technologies

MOBILITY DRIVES THIN PACKAGING

Growing: TechSearch predicts a nearly 9% CAGR for the 5-year period from 2014 to 2019. Underway is another significant evolution that could shake up the chip industry. Fan-out wafer-scale packaging, some experts believe, could push down the share held by flip chip packages and push out the need for the more expensive 3D through-silicon via (TSVs) in consumer applications, including smartphones. Outsourced semiconductor assembly and test companies (OSATs) and foundries alike are scrambling to provide this next significant wave of packaging innovation. (See sidebar on page 30.)

Fan-out packaging already includes a high-density (HD-FO) category, with sub-10mm lines and spaces. The advanced fan-out processes and TSV connections, are pushing Applied Materials and other tool makers to develop new techniques that can be used for 300mm, 200mm, and smaller wafer sizes. “We want to push our tools and see how far they will go,” said Mike Rosa, director of strategy and technical marketing for emerging technologies at Applied. “One challenge with the increasing device density on thinner wafers is dicing. The dicing lanes are becoming smaller, and if there is any chipping it can impact the active area on these small dies,” said Rosa.

Rosa added that Applied is “pushing the limits of conventional TSV processes to accommodate the trend toward smaller CDs and higher aspect ratio TSVs.”

OSATs ADOPT PROCESS CONTROL

The rise of wafer-scale packaging, where front-end-like steps are used, has been a big part of the transformation of the OSAT sector. “More packaging-style processes are being completed in the front end. The line between front-end and back-end is being blurred with the growth in technologies like wafer bonding,” Rosa said.

Shekar Krishnaswamy, a senior manager of Advanced Productivity Solutions at Applied Global Services who works with OSATs on factory automation issues, said the OSAT factory is becoming more fab-like, with more packaging-style processes being completed in the traditional front end. “The OSATs are using wafer fab equipment and technology—lithography, etching, plating, CMP, and others. Certainly the automation solutions they use are essentially the same [as front-end fabs].”

Gone are the days when packaging plant workers would tend to low rows of humming, clacking wire bond machines. And OSATs are putting in the same manufacturing execution, material handling, and process control systems long used by front-end wafer fabs in order to maximize return on their assets.

“The OSATs have had to quickly change,” said Krishnaswamy. “End customers now are asking their OSAT suppliers to provide the same kinds of advanced process control (APC) and yield tracking tools as the wafer fabs. As customers turn to more complex processes, such as fan-out wafer scale packaging and TSVs, those process control demands are set to increase,” he added.

Indeed, while integrated device manufacturers (IDMs) once did almost all their own wafer probe testing, for example, now that important function often is done by the OSATs. To leverage their investments in expensive test computers, OSATs use one test head for wafer probing and another for final test after assembly is completed.

“The same expensive test computer can be used one day as a tester, and another day with a different handler for final test. The OSATs have become very good at extracting value from their physical assets,” Krishnaswamy said.

Speaking at the 2015 Advanced Process Control (APC) conference, Thomas Sonderman, vice president of the software business unit at Rudolph Technologies, Inc. (Tewksbury, Massachusetts), said OSATs are scrambling to modernize their automation processes.

In an overview of the last 25 years of APC, he noted that the big system companies are seeking to link their entire supply chain together, using dashboard-style displays to manage the flow of components.

“A lot of information gets generated, including wafer-fab and assembly-and-test information. The goal is to link all of this together. We need data democracy, and we need to think about how to prevent barriers in the first place,” Sonderman, who previously worked in APC development at GLOBALFOUNDRIES, said.

“With China, the OSATs have the mindset that if our customers want APC and fault detection, we want it. They have the motivation, but they don’t necessarily want to give you a lot of money for it.” Sonderman added that the Chinese OSATs prefer off-the-shelf process control and automation solutions. “They don’t want to build an army [of internal software developers],” he said.

AS MOORE’S LAW SLOWS...

Early in his career, ASE’s Chen helped develop IBM’s pioneering C4 flip chip packaging technology (C4, or controlled collapse chip connection, was an early flip chip technology that involved depositing solder bumps on the top side of the wafer during final wafer processing). He said the back-end processes are picking up the pace as Moore’s Law scaling becomes more difficult. “Speaking from the packaging side, I would say the back end is no longer at the back end. We are at the front of the whole effort of getting semiconductor into systems. Now, what we do is much more application specific and customer specific,”

Rozalia Beica, chief technology officer at Yole Development ( Lyon, France), concurred, saying the advances in packaging technology have provided the semiconductor and systems companies with more flexible ways to advance their products. “With Moore’s Law slowing down at the advanced nodes, things are getting more expensive and yields are becoming more difficult to achieve. So many companies are using, say, 65nm technology and not 45nm technology. They are using advanced packaging to put the right solutions together without trying to do everything on 300mm wafers at the most advanced nodes.”

SMALL FORM FACTOR SOLUTIONS

One of those companies is Silicon Labs (Austin, Texas), which combines microcontrollers, low-power radio frequency (RF) circuitry, and on-chip flash memory into products aimed at Internet of Things applications. (See “Silicon Labs Tackles IoT Solutions,” NanoChip FebVol, Vol. 10, Issue 1, 2015.)

Daniel Cooley, Silicon Labs’ vice president of marketing for IoT products, said many of the chips used in IoT or wearable applications are extremely small. For example, a recent 8-bit IoT controller from Silicon Labs is in a chip-scale package (CSP) with a 1.8 by 1.66mm footprint, which is about one-fourth the footprint of a conventional quad-flat no leads (QFN) package.

“With the IoT, you have to keep costs down. But you also have to get all of the IOs out in a small form factor product, and that is a hard process. There is a lot of current going through very small pipes: 300 or 400 milliamperes while transmitting. Doing that in a small form factor is hard,” Cooley said.

Over the past two years, the major OSATs have improved their ability to handle the very small CSPs, with small solder ball pitches. “Two years ago, handling 3mm pitch balls
A TSV-last process flow for MEMS devices is suited for outsourcing to an OSAT, and has relatively low stress between the via and silicon. (Source: ASE Group.)

MOBILITY DRIVES THIN PACKAGING

Watches and Wearables

The entire wearable category is prompting closer collaboration, with packaging decisions made early in the design cycle with packaging partners. Cooley said Slicen Labs is learning that chips used in wearables are “a little more complicated” in terms of flexibility and thermal constraints. Also, “wearables all connect to your phone, so there is a need to optimize for the antenna radiation pattern. And the package has to be able to adapt to a wide array of PCB materials.”

Dick James, a senior fellow at reverse-engineering firm Chipworks (Ottawa, Canada), said his company did a teardown of the Apple Watch recently. While most of the chips are in fairly standard packages of one kind or another, the entire PCB board is encased in an over-molded cover to prevent moisture and radio frequency waves from compromising the watch.

“That is the first time we have seen that total encapsulation, and we think they did that because it is a wearable and likely to see moisture ingress challenges. RF shielding is also needed. In the unlikely event of sticking it in a microwave, the shielding could prevent the watch from exploding,” James said.

Thinner MEMS Sensors

Rosa, who earned his doctorate in the MEMS field, said the move to thinner MEMS sensors in smartphones is having a major impact.

“MEMS are a special challenge, because the package often includes the MEMS device, a layer of control logic, and a capping layer. The whole structure might be 300 microns, so all those layers need to be thinned.”

In part, that need motivated Applied Materials to offer customers a series of modifications to its Endura platform, enabling the handling of wafers down to 100 microns thick.

“It should be noted that in most cases, customers are not passing around super-thin wafers. The customers usually bond the wafer to be thinned to another wafer, either glass or silicon, before thinning,” Rosa added.

“People are participating in research and development efforts in Singapore and Europe to reduce the size of the TSVs. TSVs are used to connect the MEMS and ASIC wafers in back-side-illumination (BSI) CMOS image sensors, some power ICs, and in several other applications. To reduce stress, polysilicon vias are often used, and the number of TSVs in these applications is small, sometimes only 6 to 10, to connect the MEMS structure with the control logic wafers, or to connect the MEMS to the PCB,” Rosa said.

One motivation for TSV use in these devices is to get away from wire bonding, where the pads take up valuable real estate. “If a customer goes from wire bonds to TSVs, they can save 25 to 30% of the die area. That is huge. They can increase the numbers of die per wafer by 25 to 30% with TSVs,” he said.

With an overall goal of putting more TSVs into a certain area, researchers are investigating how to make the critical dimensions of the TSVs smaller, with a smaller diameter and a higher aspect ratio (width to depth).

“As chips get smaller, the CD of the TSVs decreases, and that increases the challenge of how to perform a reliable line-and-fill without voids. If the diameter goes from 10 to 5 microns, we really have to improve the barrier and seed layers,” Rosa said.

The thickness of the barrier and seed layers correlates strongly with the quality of gap fill, which is related to the choice of chemistry, hardware, and process control capabilities.

For a detailed discussion of TSV processes, see “Optimizing TSV processes and integration for volume manufacturing” by David Erickson, Isaac Che, and Sesh Ramaswami (all of Applied Materials), Chip Scale Review, January–February 2015.

TSVs Face Competition, Need Time

Gaurav Sharma, a senior member of the technical staff at GLOBALFOUNDRIES, joined that company this year after spending more than a decade in various advanced packaging operations. Sharma said TSVs—both in interposers and in vertical 3D ICs—are gaining more traction from a production point of view. “TSVs are in volume production, but for niche applications,” he said, noting that high-bandwidth memory graphics, and very high performance ASICs are turning to TSVs.

An area where TSVs provide advantages is for the “huge bandwidth requirements coming up with Internet traffic,” Sharma said. While these applications do not have volumes nearly as high as devices used in mobile systems, he believes there is serious interest in 3D TSVs by some pretty significant players who want to bring this technology into production.

“The challenge facing TSV adoption arises from the relatively high cost of the complex process, especially compared to the up-and-coming HD-FO chip scale packages. “When we look at TSVs for mobility, there are alternative technologies that can do much the same thing at a much better price,” he said.

Sharma noted that designers are also hesitant about 3D TSVs, which require that the active silicon circuits be designed around a keep-out zone occupied by the TSVs. “What I have learned from various customer engagements is that designers are resistant to this idea,” he said.

Rozalia Beica, the Yole CTO, believes 3D TSVs will gain a stronger foothold, but it will take time. “I personally am a strong believer in TSVs and have worked on them since 2006,” she said.

“If you look at other technologies, flip chip was developed by IBM in 1960, but it didn’t go into volume production until 1980. Similarly, the time for MEMS to go from development to mass adoption took 20 years. TSV technology is not slower; it is just that there was such big hype around it that now everyone is waiting for it to be adopted.”

Indeed, companies ranging from AMD to Xilinx in logic, and SK Hynix and Toshiba in memories, have introduced products that include some form of TSVs. “Personally, I don’t think TSVs are necessarily slower [to proliferate] than other technologies,” Beica said.

For additional information, contact mike_rosa@amat.com.


Mike Rosa, PhD, director of strategy and technical marketing for 200mm products at Applied Materials.
**FAN-OUT IS A GAME CHANGER**

In one sense, fan-out chip-scale packaging is just an evolutionary step, an advance on the fan-in CSP used in so many relatively low pin count chips. But look a bit deeper, and fan-out is perhaps the biggest thing to hit the semiconductor industry since immersion lithography and high-k dielectrics.

Jan Vardaman, president of packaging consultancy at TechSearch International (Austin, Texas) said fan-out packaging can be used for single or multiple die. “Fan-out can be a disruptive technology for several reasons. It can be used for multiple die, there is no substrate, and it can be done in a foundry.”

TSMC has been talking about its version of fan-out packaging, called integrated fan-out (iFanO), for several years, and is one of a handful of companies already in production with fan-out packaging. In a financial results conference call with analysts in mid-October, TSMC co-CEO Mark Liu said TSMC’s iFanO technology “will enter high-volume production with our 16-nanometer technology next year. We are currently working on the second-generation iFanO technology for several projects of systems integration on 10 nanometer and 7 nanometer.”

Liu and co-CEO C.C. Wei fielded multiple questions on TSMC’s iFanO technology, and said it can bring greater than 20% reduction in overall package thickness, a 20% speed gain, and 10% better power dissipation. TSMC has completed construction of the new facility in Longtan, near Hsinchu, Taiwan, for iFanO volume production, with manufacturing equipment move in “on schedule” and volume ramp-up scheduled for the second quarter of next year.

The TSMC executives did not mention Apple or its A10 application processor by name, but they said a large proportion of TSMC’s 16nm wafers will go through the fan-out packaging steps, with one customer accounting for much of it. According to Wei, “probably it will be adopted by the mobile processors first, and then applied to all other applications.”

First developed independently by Freescale and Infineon some 15 years ago, and subsequently licensed and developed by multiple OSATs, fan-out packaging is starting to see volume production. Vardaman said Intel, by virtue of its acquisition of Infineon’s wireless operation, is an early adopter. Intel’s Wireless Division uses fan-out packaging for an LTE multichip part that measures just 5.32 by 5.04 mm.

Some form of fan-out packaging has been used by Marvell and Maxim as well, Vardaman said. Amkor is bringing up a FO-WLP line in Korea, and ASE Group is building a fan-out line in Kaohsiung, Taiwan, with other major OSATs already in production or planning their versions of fan-out packaging. Cadence is among the EDA vendors who have released design tools that bridge the front-end and back-end design processes for packaging.

ASE senior fellow Bill Chen said ASE’s original motivation to license fan-out, and then later to work on its own implementation, was to increase the number of solder balls that could be placed under the increasingly tiny die. “The original fan-out motivation was to accommodate the die shrinks. Then, being creative engineers, we realized we can do a lot of other things with it, and create very thin packages,” he said.

Chen said the process involves dicing the chips on a silicon wafer, and then very precisely positioning them on a thin “reconstituted” or carrier wafer, which is then molded. The redistribution layer is created, and then solder balls are formed on top, just as in a wafer-level chip scale package. “The idea is very simple, to produce a reconstituted wafer and put it through the line,” he said, adding that “there are no laws of physics we are fighting, we are just fighting engineering, which means the challenges will be solved gradually. Fan-out is something that can be processed through wafer fab equipment—we need to pick the wafer up, place it, mold it, and bake it so the molding compound is cured. There is not any major wafer thinning needed like with TSVs. Clearly, fan-out technology is becoming more and more significant as we move further into the era of SIP and heterogeneous integration”

Rozalia Beica, CTO at Yole Development, said TSMC has been telling its suppliers to be ready for volume fan-out processing beginning next year, perhaps in the 100,000 wafers-per-month range.

Beica estimated that an application processor can be combined with memories and other devices, with a 40% reduction in form factor compared with conventional flip chip packaging. And because fan-out does not require a substrate, the cost savings is substantial.

“There is a big difference in costs. Fan-out is lower cost than flip chip, which requires an organic substrate. And compared with silicon interposer, there is an even bigger cost difference, though maybe at this point there might be more performance with a silicon interposer,” she said.

How far fan-out packaging will go remains to be seen. But if multiple chips can be accommodated, it may cause some design teams to reconsider the need to create complex SoCs on advanced silicon, particularly for mobile and IoT systems. The analysts are predicting major shifts in how many chips will require more expensive types of packaging, such as flip chip, BGA, interposers, and 3D integration.

“Everyone in the industry is scrambling to figure out the consequences” of fan-out packaging, Vardaman said, noting that “packaging is much more interesting than it used to be.”

— David Lammers
Fault detection (FD) is pervasive in the industry and is now a key capability in the ongoing effort to improve quality and reduce cost. The next generation of FD will significantly reduce setup times, improve detection with fewer false alarms, and take advantage of big data capabilities to decrease response times and increase depth of analysis.

It’s hard to believe, but fault detection (FD) has been a part of our industry for over 20 years and an integral component of microelectronics manufacturing as a whole for at least the past decade. Manufacturers rely on FD to minimize scrap, improve product quality, detect quality degradation, and determine when equipment may need to be shut down for maintenance, among other benefits. Today’s typical fab employs some form of FD on almost all processes.

But while FD provides significant benefits, there are also cost and performance issues associated with current FD deployment and operation that present challenges to both users and suppliers. For example, Paul Ewing, an FD deployment expert and part of Applied Material’s FD advanced services deployment team, said, “It often takes up to two weeks to correctly configure univariate FD for a process tool, including collecting data, refining limits, and correlating limits violations to actual events of importance.” Also, there are often too many false alarms or missed alarms associated with a particular FD model.

The problem is illustrated in figure 1. With just this single sensor trace, the FD engineer must investigate several features and develop multiple models, each with limits. For the entire fab, there are often thousands, if not millions, of models and limits to manage. Clearly, an opportunity exists for improvement in FD setup, execution and maintenance capabilities.
In determining how to address the problem it is important to consider the human factor (see related article “Human Factors in Automation Systems” on page 2 of this issue of Nanochip Fab Solutions). Much of the cost of FD deployment results from the time and occasional error associated with humans executing repetitive tasks that could benefit from automation; however, it is important to continue to harness human expertise in process, equipment, and sensor knowledge. FD improvements should provide the optimal cost-benefit balance in level of automation.

Applied Materials has been researching this problem for several years. We have spoken to customers and FD deployment experts, collected statistics on the benefits and costs of FD deployment, researched new techniques in fault detection and classification in microelectronics and other industries, and innovated as needed to address specific issues.

As a result, several improvements to our FD capability are being implemented that will allow us to address key issues and provide a higher-quality FD solution. Some of the more noteworthy improvements are summarized in table 1, and are collectively part of the Applied Materials Next-Generation Fault Detection and Classification (NG-FDC) solution. A few of these features are explored in this article, with references provided for further reading.

### NEXT-GENERATION FAULT-DETECTION AND CLASSIFICATION CAPABILITIES: A DEEPER LOOK

Automated, expert-driven trace transition detection and feature selection, with ranking

If we analyze the trace of figure 1 we see that the FD modeling engineer must complete a number of tasks in order to provide a high-quality FD solution for this single trace. The engineer must first define the regions that need to be monitored and their boundaries, denoted as “steps” in figure 1. He must then decide if traces should be aligned according to specific region boundaries before analysis, or if the lack of alignment is actually a sought-after anomaly. He must also determine what FD model or models are best suited to assess a fault in a particular region.

For example “max” and “range” are selected in step 4, while “mean” and “sigma” are the methods of choice in step 12. Warning, alarm, and control limits must be applied to these models. The choice of boundaries, alignment, models, and limits are usually derived from analysis of multiple traces of the same sensor, combined with process and equipment knowledge as to what boundaries, steps, and features are important. NG-FDC features will partially automate this process while ensuring that process and equipment knowledge is incorporated into the final FD model set.

As illustrated in figure 2, NG-FDC will utilize techniques such as moving windows and wavelets to determine region boundaries. 

Once the regions are identified, several techniques can be used to determine which features should be extracted and modeled. One is a “Monte Carlo” approach, where existing model types such as mean, standard deviation, and slope are applied to the region to determine the level of variability of the feature they capture. The model-types are then ranked. More complex techniques such as binning and structural feature extraction can also be employed.

The output is a list of features with a ranking that indicates the level of signal-to-noise that would be captured by feature-monitoring. Leveraging this automation, the expert would then select region boundaries, plus the features to model within these boundaries, incorporating process and equipment knowledge. In this way model quality is guaranteed from both an analytical and process and equipment perspective, while model setup times are significantly reduced.

Trace ranking and move to “supervised” models

In traditional FD all sensors and regions are candidates for monitoring, which results in an unwieldy number of models, many of which are relatively useless. It is up to the expert to pare down this list, as effort that can be time-consuming and error-prone. This modeling process is called “unsupervised” because models are developed without directly correlating trace data to quality data (such as metrology or yield).

With NG-FDC there is often an opportunity to reduce the set of features that need to be extracted, by determining which sensors and trace regions are associated with a particular issue that needs to be monitored, such as a metrology or yield excursion. The variability in the trace information is analyzed with respect to quality variability. Sensors, sensor trace regions, and features can then be ranked according to their impact on quality variability.

Techniques such as guard-band statistics and hidden Markov models (HMM) are useful to determine these critical sensors and regions. This process of incorporating quality or other “output” data into the determination of sensors, critical features, models, and model limits is part of the NG-FDC move from “unsupervised” to “supervised” modeling techniques.

Model management

A key issue in FD performance over time is the ability of models to continue to accurately reflect the operation of the tool, minimizing false positives (i.e., false alarms) and false negatives (i.e., missed excursion

---

**Table 1. Improved FD features included in the Applied Materials NG-FDC solution.**

<table>
<thead>
<tr>
<th>FEATURE</th>
<th>SOLUTION</th>
<th>PRIMARY BENEFITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trace transition detection and alignment</td>
<td>Moving window, wavelets, etc. †</td>
<td>Reduce setup times</td>
</tr>
<tr>
<td>Feature selection ranking</td>
<td>Unsupervised and supervised; Monte Carlo, binning, structural feature extraction, etc. †</td>
<td>Reduce setup times, reduce false positives (FPs) and false negatives (FNs)</td>
</tr>
<tr>
<td>Automated limits generation and management</td>
<td>Auto charting, defining training sets, golden tool limits, netalert triggering based on context information and setpoint changes; innovative techniques such as hidden Markov models and change-point analysis may have some applicability. †</td>
<td>Reduce setup and maintenance costs; Reduce FPs and FNs</td>
</tr>
<tr>
<td>Model management and maintenance</td>
<td>Leverage innovative techniques used for management and maintenance of virtual metrology models. †</td>
<td>Reduce FPs and FNs</td>
</tr>
<tr>
<td>Fault classification (FC)</td>
<td>Leverage supervised techniques used for virtual metrology and predictive maintenance, such as partial least squares (PLS) and support vector machines (SVM); Leverage clustering techniques (both unsupervised and supervised) for optimizing classification schemes.</td>
<td>Improve capabilities by providing the “C” in FDC, thereby reducing mean-time-to-repair (MTTR).</td>
</tr>
<tr>
<td>Innovative features to expand FDC capabilities</td>
<td>Wafer topology prediction and sensitivity analysis, method for finding sensors related to a trace feature.</td>
<td>Improve capabilities; improve overall FDC quality and usability</td>
</tr>
<tr>
<td>Support for big data frameworks: Volume</td>
<td>Support FD-data collection and analysis on big data infrastructures such as Hadoop.</td>
<td>Improve capabilities; improve overall FDC quality and usability</td>
</tr>
<tr>
<td>Velocity</td>
<td>Support FDC-data collection and analysis on big data infrastructures such as Hadoop.</td>
<td>Reduce setup times; improve overall FDC quality and effectiveness; enable a platform to support future FDC and prediction technology capabilities.</td>
</tr>
<tr>
<td>Variety</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volatility</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† Semi-automated solution that allows incorporation of expert knowledge

---

**Figure 2. Illustration of moving window approach to identify trace boundaries:** (a) A moving window is used with a difference function to capture transitions; (b) The size of the window is determined by signal stability, noise, and other properties so as to generate the best difference function; (c) Normalization of the signal is needed because multiple traces might have different values and value change profiles; and (d) the transition points are mapped back onto the original trace (here utilizing color for demarcations) to identify regions and boundaries for analysis.

---
NEXT-GENERATION FAULT DETECTION

IMPROVES QUALITY AND REDUCES COST

detections across PMs and other events that alter the state of the tool. Fortunately, techniques are being developed for advanced capabilities like virtual metrology that can be leveraged back into FD model maintenance.22-24 These “supervised” techniques allow feedback of information such as false positives and false negatives to be used for model optimization. They can provide decision points on when to adjust models or limits in response to a change in offset in the tool or process, and when to rebuild models from the ground up. Additionally, they enable troubleshooting of faults to determine the critical sensors associated with a particular excursion, as illustrated in figure 3.

Incorporating revolutionary capabilities

While a large part of NG-FDC is focused on the automation and improvement of traditional FD features, a parallel focus is to incorporate new and innovative features to make NG-FDC more effective and easy to use, and to endow it with additional capabilities. For example, a wafer (or panel) topography prediction capability is being developed that utilizes FD information collected for a process to predict wafer topography (e.g., film thickness). As illustrated in figure 4, process sensor value or recipe set point adjustments can be simulated to determine the sensitivity of particular parameters to this topography. Utilizing this capability, product quality and yield degradation from topographical issues such as non-uniformity can be reduced.

Another example of FD innovation that can be incorporated into NG-FDC is a technique developed to determine sensor traces related to a target sensor being analyzed. The correlation is determined based on (1) the location where sensor values change and (2) the change in trace signature with respect to the target sensor. This technique is useful to identify sensors or sensor groups that may be better suited to monitor a particular fault, thereby providing stronger signals and insight into fault classification and root cause.25-26

SUPPORT FOR BIG DATA FRAMEWORKS AND CAPABILITIES

The big data revolution provides us with opportunities to leverage improvements in the “five V’s” of big data: volume, velocity, variety (emerging of data sources), veracity (data quality) and value (algorithm).27 While a large part of NG-FDC is focused on the next generation of fault detection and prediction capabilities, there is also a parallel focus to incorporate new and innovative features to make NG-FDC more effective and easy to use, and to endow it with additional capabilities.

NG-FDC systems can leverage big data ecosystems such as Hadoop to provide FDC advancements from each of the five “V” perspectives.27-29 “Data volume” improvements support improved models that mine larger quantities of data from both depth (archive length) and breadth (number of sensors). Improvements in the “velocity” of data collection and analysis allow for finer granularity and increased complexity of analysis without increased development time. The improvements in data merging (“variety”) allow direct access to quality data (e.g., yield and metrology) alongside trace and FD output data. This will facilitate the move from “unsupervised” to “supervised” modeling in NG-FDC. Finally, data quality improvements (“veracity”) will support more sophisticated modeling techniques (“value”). These will reduce the occurrence of false positives and false negatives in NG-FDC systems, and pave the way for more complex predictive solutions such as yield prediction feedback for fab-wide control to hit yield targets, or even improve yield.

THE FUTURE IS NEAR

Each of the features included in NG-FDC will improve the performance and lower the cost of an FD solution by (a) decreasing setup time and reducing false positives, (b) simplifying management of models and limits, and (c) expanding capabilities. Collectively they provide a foundation not only for NG-FDC, but also for key emerging technologies that rely on FDC, such as virtual metrology and predictive maintenance.

Acknowledgments: The authors would like to thank Brad Schulze, Deepak Sharma, Kommisetti Subrahmanyam and Jianping Zou for their support in the development of this article.

(22) This fact was underscored at the 2015 Integrated Measurement Association (IMA) APC Council meeting, attended by key users. Three main points of consensus from this meeting: (1) FD limits management is a key concern of fab APC managers, (2) there is a need for some level of automation in the FD model-building processes, while retaining process and equipment expertise, and (3) no comprehensive solution is currently available.


For additional information, contact michael_d_armacost@amat.com.
SURFACE TEXTURES OF CHAMBER PARTS

CAN BOOST TOOL OUTPUT AND AVAILABILITY

Given the many issues involved in semiconductor manufacturing, it’s easy to overlook the small things that can make a big contribution to operational excellence—things like the surface textures of your chamber process kit parts. But advanced surface technology, available as an option within Applied Materials service agreements, can significantly increase the availability of Applied Materials Endura tools while reducing overall cost of ownership (COO).

As semiconductor manufacturing technology evolves, particle limits and sizes are constantly reduced to meet production yield targets. Usually this requires more frequent maintenance, including process kit changes, to maintain acceptable particle levels, resulting in more chamber downtime, reduced tool availability, and increased COO.

Applied Materials LavaCoat2 process kit parts can improve chamber particle performance and device yield, increasing tool output and availability. LavaCoat2 is a proprietary surface-preparation technology that texturizes parts (see figure 1) so that deposited films adhere to their surfaces more readily than to surfaces texturized by other methods, such as grit-blasting and twin-wire arc spraying (TWAS).

While traditional process kit textures such as TWAS provide standard roughness and adhesion of deposited film, LavaCoat2 is an engineered texture that provides greater tensile strength, resulting in reduced particle levels and longer process kit life.

LavaCoat2 replacement parts can lengthen the mean time between cleans (MTBC) by up to 2x and help lower defect densities in the chamber by up to 50%. This can significantly boost output and yield.

At one site, a LavaCoat2 deposition ring delivered significantly lower particle performance for a customer’s Applied Materials EpiCote II TaSiO chamber. Particles >0.06 µm in size were reduced by 90% when compared with a conventional TWAS deposition ring.

Figure 2 shows how LavaCoat2 replacement parts eliminated particle spiking and reduced the overall baseline of defects in an Applied Materials TxZ (CVD TiN) chamber that had been using standard textured parts.

MORE TOOL OUTPUT

Applied Materials Total Kit Management (TKM) programs provide customers worldwide with a turnkey offering that encompasses spares, cleaning and coating, and logistics services for efficient management of their inventories of process kits and components. The program offers a kit-management solution that includes genuine spare parts, cleaning and coating specifications, and advanced textures. When combined with Applied Materials services, the TKM program can help customers further optimize tool performance and overall factory operations.

LavaCoat2 replacement parts are the latest service enhancement from Applied’s Surface Technology Group. This engineering team is dedicated to driving continuous improvements in parts to support Applied’s tool-specific performance commitments to its customers with service contracts.

For additional information, contact john_mangini@amat.com.

Given the many issues involved in semiconductor manufacturing, it’s easy to overlook the small things that can make a big contribution to operational excellence—things like the surface textures of your chamber process kit parts. But advanced surface technology, available as an option within Applied Materials service agreements, can significantly increase the availability of Applied Materials Endura tools while reducing overall cost of ownership (COO).

As semiconductor manufacturing technology evolves, particle limits and sizes are constantly reduced to meet production yield targets. Usually this requires more frequent maintenance, including process kit changes, to maintain acceptable particle levels, resulting in more chamber downtime, reduced tool availability, and increased COO.

Applied Materials LavaCoat2 process kit parts can improve chamber particle performance and device yield, increasing tool output and availability. LavaCoat2 is a proprietary surface-preparation technology that texturizes parts (see figure 1) so that deposited films adhere to their surfaces more readily than to surfaces texturized by other methods, such as grit-blasting and twin-wire arc spraying (TWAS).

While traditional process kit textures such as TWAS provide standard roughness and adhesion of deposited film, LavaCoat2 is an engineered texture that provides greater tensile strength, resulting in reduced particle levels and longer process kit life.

LavaCoat2 replacement parts can lengthen the mean time between cleans (MTBC) by up to 2x and help lower defect densities in the chamber by up to 50%. This can significantly boost output and yield.

At one site, a LavaCoat2 deposition ring delivered significantly lower particle performance for a customer’s Applied Materials EpiCote II TaSiO chamber. Particles >0.06 µm in size were reduced by 90% when compared with a conventional TWAS deposition ring.

Figure 2 shows how LavaCoat2 replacement parts eliminated particle spiking and reduced the overall baseline of defects in an Applied Materials TxZ (CVD TiN) chamber that had been using standard textured parts.

MORE TOOL OUTPUT

Applied Materials Total Kit Management (TKM) programs provide customers worldwide with a turnkey offering that encompasses spares, cleaning and coating, and logistics services for efficient management of their inventories of process kits and components. The program offers a kit-management solution that includes genuine spare parts, cleaning and coating specifications, and advanced textures. When combined with Applied Materials services, the TKM program can help customers further optimize tool performance and overall factory operations.

LavaCoat2 replacement parts are the latest service enhancement from Applied’s Surface Technology Group. This engineering team is dedicated to driving continuous improvements in parts to support Applied’s tool-specific performance commitments to its customers with service contracts.

For additional information, contact john_mangini@amat.com.
The complexity of semiconductor manufacturing makes it difficult to optimize production for increased efficiency and output while maintaining high product quality and reducing overall costs. Yet it is imperative that semiconductor manufacturers and outsourced assembly and test (OSAT) suppliers find ways to do so, because production inefficiencies can keep them from fully capitalizing on the opportunities presented by today’s fast-changing, highly competitive markets.

One strategy is to approach the tasks of production planning, scheduling, dispatching, and reporting in an integrated, flexible, and open fashion, with the goal of optimizing production every step of the way. This means from initial enterprise-level master planning down to the factory floor dispatching of specific lots of work-in-progress (WIP) in real time.

Enabled by an integrated suite of automation software that incorporates a state-of-the-art optimization engine, such an approach would allow fabs to anticipate and better respond to fluid conditions, thereby controlling the speed of WIP processing and reducing waste. It could help manufacturers overcome the impediments to higher productivity they face every day, including:

- Disruption, delays and cost associated with having to analyze and validate changes to dispatching rules and scheduling policies directly on the factory floor, in response to changing conditions.
- Inability to simulate dispatching rules with sufficient accuracy to build realistic “what-if” models that reduce bottlenecks, increase overall throughput, and ensure that cycle time commitments are met.
- Lack of cohesion resulting from the difficulty of coordinating the efforts of engineers who work separately in planning, dispatching and shop floor “silos,” making it difficult to achieve faster and better operational decision-making and reporting.

Manufacturers often see the need to optimize production, but cannot do so efficiently for various reasons. For example, their systems may not provide good, real-time data from the fab (as opposed to data used in spreadsheet planning). Or there may be a lack of integration between optimization and execution mechanisms, combined with an inability to automatically initiate the right actions based on certain critical events and timing.

Other factors may include the lack of an integrated way to visualize inputs and outputs and limited diagnostic information to help trace why and how decisions are being made.

To begin to address these issues, it is useful to think of production planning, scheduling, dispatching, and reporting as a hierarchy of functions, organized by specific tasks and requirements and with different time frames (see figure 1).
The master planning function deals with enterprise-level planning issues. Master planning is generally the responsibility of a central planning organization whose main goal is to interface with customers and address their overall needs. Its scope includes overall sourcing strategies and capacity analyses to decide which fabs will be used to produce parts for a customer. The master planning group typically looks 1-18 months ahead.

Factory planning, meanwhile, performs capacity analyses for tool groups with the goal of delivering product by the due date while balancing factory constraints. Tasks in the purview of this group include decisions to set up tools in certain ways, whether to bring on more personnel for given work shifts, and so forth. These planners typically look 1-4 months ahead.

The scheduling group makes sure that key tools and tool groups are available and WIP is balanced so that they can meet the factory objectives. Another key goal is to know how to load the products across the tool group to achieve factory planning objectives (e.g., to meet lot due dates, cycle time targets, setup time goals, etc.). The time frame here is generally 12-24 hours ahead, essentially a work shift or two.

Finally, the dispatching group is real time-oriented, with the goal of executing tool schedules to meet fabrication process requirements and maximize production throughput.

**AVOIDING SUBOPTIMAL SOLUTIONS**

Most manufacturers use a disparate mix of software applications for planning across the hierarchy. These tools may include open source solutions, spreadsheets, commercial products, and homegrown applications. However, these heuristics-based tools trade quality for speed. That is, while they may produce results quickly in response to changes on the factory floor, those rules may not represent the best of all possible solutions.

For example, a new dispatching rule that is adequate for a given process step may be so narrowly written that it precludes opportunities to balance production down the line, as can happen when manufacturers inadvertently create downstream bottlenecks by trying to push more WIP through a given set of tools.

Although optimization algorithms may be used to improve the performance of these heuristics-based programs, they don't offer a comprehensive infrastructure for developing and maintaining solutions. Also, these systems require additional databases and applications to be supported. For example, additional custom code is required to prepare input data, set up and run optimization programs, and validate and post-process optimization results.

By contrast, an integrated automation framework can offer users the biggest potential productivity gains (see figure 2). Applied Productivity Family (APF) software can be used to build a framework to optimize planning across the hierarchy. It encompasses production planning, dispatching, automation, and scheduling software for equipment and process-control systems.

A fast, robust mathematical optimization program — referred to here as a “solver” — works with the APF software to holistically analyze and implement production requirements. With it, manufacturers can simulate, select, schedule, and dispatch production resources and WIP with optimal productivity and efficiency in real time.

The APF software currently supports the CPLEX and Gurobi commercial solvers as well as the COIN open source solver. This gives users a unique, fully integrated software framework (see figure 3) that is scalable and flexible, and enables them to protect their IP.

**MASTER PLANNING EXAMPLE**

The productivity advantages this framework offers can be illustrated by a hypothetical look at how the master planning function might be conducted at an OSAT company.

Assume that the OSAT receives orders from a semiconductor manufacturer every week for packaging, testing and assembly, along with specific detailed requests for the production of multiple lots needed over each of the next 10 weeks.

The OSAT must respond to the customer by specifying what it can actually produce week-by-week within that time frame and committing to do it. With manual systems it takes a few days to formulate this response and commitment, but an automation framework is expected to reduce that response time to a few hours.

The commitment would have to consider not only the customer’s specific requests (generally for multiple lots of product per week), but also additional work for that customer and other customers that may be in process or anticipated in the same time frame.
OPTIMIZING PRODUCTIVITY ANALYSIS FOR BETTER FACTORY PERFORMANCE

It is also possible that the customer may change the production request. If that happens, the OSAT would have to quickly determine whether it can meet the new requirements.

Taking into account all of the variables involved in making production commitments, modeling and verifying optimum production scenarios, and then actually carrying them out in a dynamic environment is a formidable task.

Conceptually (see figure 4), a master planner would follow these steps:

1. Initiate the APF system.
2. Specify the data to be used.
3. Configure master planning scenarios to be run.
4. Initiate plan execution.

The planner would consider higher-level defined objectives such as customer end-product demand, process and device mix, wafer/die schedule, suggested lot release dates, and weekly commitment plans.

Planning inputs would include constraints and other information such as customer orders in the queue, upstream WIP equipment availability, potential capacity shortages, long cycle times, wafer costs, materials requirements, arrivals and inventories, line imbalances, and high levels of variability. All assembly and test operations would be considered, including load boards and handling equipment in the testing area.

The optimization model operates according to predefined logic and “what-if” scenarios executed by the optimization solver, based on automation workflows defined in Applied’s APF framework. The solver minimizes lack of responsiveness to the customer request, and gives a reason for any unsatisfied demands, such as wafer lots that didn’t arrive on time or tools that were over capacity.

As a result, the OSAT may realize specific improvements. These can potentially include better dispatch rules and scheduling, improved bottleneck management, greater or different use of automation systems, and overall higher efficiency and productivity.

CONCLUSION

In the manufacturing, assembly, testing, and packaging of semiconductors, meeting customer needs for quality, delivery, and cost has never been easy. But today’s complex technology and demanding economics make it more difficult than ever before.

Now more than ever, factories and tools alike must be planned, scheduled, and used as efficiently as possible. New solver-based optimization capabilities integrated with already powerful planning, scheduling, dispatching, and reporting automation software will help make this possible from the enterprise level down to the factory floor.

For additional information, contact shekar_krishnaswamy@amat.com.

Figure 4. By incorporating a solver, APF software enhances master planning accuracy.

GLOBAL CUSTOMER CONTACT CENTER

NOW AVAILABLE IN MORE REGIONS

BY NANOCHIP STAFF

Customers in North America, Taiwan, and Europe have 24/7 support capability with a single point of contact for service and parts requests.

The GCC center is staffed by trained service and spares experts who understand how to address customers’ needs. When customers contact the center, requests are efficiently and effectively relayed through the network of Applied Materials support services experts. This communication chain ensures that the right customer engineers (CEs) for the job respond quickly. CEs specifically authorized to work on a customer’s tool can access a comprehensive tool service history for the customer’s fab. They can also search Applied’s extensive global database of service and repair best-known methods (BKMs) and the order status of current parts.

The GCC center also provides the convenience of an always-available backup for local service teams. If a local Applied on-site CE or other service professional is not immediately available, GCC center staff will quickly assess the situation and route the issue within Applied’s service network for a prompt response. The GCC center provides standardized, simplified service order creation, issue tracking, management, and, if necessary, rapid escalations to help resolve customer issues.

“The goal of the GCC center is to simplify and consolidate our contact center services. We want to make it easier to do business with Applied Materials and to deliver greater customer satisfaction,” said Applied Materials GCC Center Director Elizabeth Brogdon.

“With more than 1,000 unique tool-types in Applied’s installed base, and over 30,000 tools installed in customer fabs around the globe, we’ve designed the advanced capabilities of the GCC center to help us address customers’ parts and services requirements faster, more efficiently, and as cost-effectively as possible.”

Customers in the following regions can contact the GCC center at these email addresses:

<table>
<thead>
<tr>
<th>Region</th>
<th>Email Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td><a href="mailto:GCC_AMNA@amat.com">GCC_AMNA@amat.com</a></td>
</tr>
<tr>
<td>Taiwan</td>
<td><a href="mailto:GCC_AMT@amat.com">GCC_AMT@amat.com</a></td>
</tr>
<tr>
<td>Europe</td>
<td><a href="mailto:GCC_AME@amat.com">GCC_AME@amat.com</a></td>
</tr>
</tbody>
</table>

For additional information about the GCC center, contact your local Applied Materials sales or service manager.

Applied Materials has expanded the operations of its new Global Customer Contact (GCC) center to include customers in Taiwan and Europe. The GCC center, first launched in North America in September, offers a streamlined, single point of contact for all Applied equipment service- and parts-related needs. In the next phase of expansion, regionalized GCC access for Japan, Korea, China, and Southeast Asia will begin in early 2016.

Open around the clock, seven days a week, the GCC center allows Applied Materials customers to use a single, assigned phone number or email address to schedule service, order parts, check order status, request a quote, and escalate or expedite a support request. Service is offered in local languages by people who provide expert help. This centralized contact point for service and support means that all requests can be accurately logged, tracked, and dispatched quickly.

SUBFAB EXHAUST MANAGEMENT EVOLVES TO MEET NEW CHALLENGES

Subfab utilities and systems aren’t glamorous, and they usually aren’t considered as critical to a fab’s success in the same way production tools are. Yet significant economic and reputational penalties may apply if a fab’s emissions violate increasingly strict environmental standards and expectations—so failing to upgrade the subfab is a risk not worth taking.

BY ANDREAS NEUBER, JOHN DICKINSON, DUSTIN HO, AND ANDREW HERBERT

Abatement of fluorinated greenhouse gases (F-GHGs) generated during semiconductor manufacturing has come a long way in recent years, and it continues to evolve. In 2009, Applied Materials introduced its first zero footprint pre-pump plasma technology for abatement of F-GHGs from dry etching chambers. The Applied Aeris-G product addresses the need for an efficient abatement capability for this application.

The idea is to convert perfluorocarbon compounds (PFCs) and other oxides/poly etch waste gases into water-soluble species, which then can be removed by the facility scrubber. Applied Aeris-G units focus on greenhouse gas reduction and treat PFCs/F-GHGs upstream of the pump, where there is no pump purge gas present to dilute the effluent.

The smaller volume of gas and the higher concentration of target compounds mean that destruction efficiency is higher. Incineration by-products, such as volatile organic compounds (VOCs) and nitrogen oxides (NOx), is also reduced compared to post-pump burning and plasma systems.

In addition, treating the process effluent prior to the addition of pump purge gases results in significantly lower operating costs for abatement compared to traditional solutions. A typical Aeris-G unit uses just 1-3 kW of electrical energy per chamber to accomplish the abatement process, as compared to post-pump solutions that typically use 8–12 kW of equivalent energy, usually because they rely on natural gas to destroy the unwanted compounds via combustion.

Since the introduction of pre-pump plasma technology, more than 1,000 first-generation Aeris-G abatement units have been installed at fabs around the world, with a demonstrated mean-time-between-failure (MTBF) reliability of more than 100,000 hours.

The Applied Aeris product family has been continually adapted to meet the requirements of a growing number of etch processes and, more recently, deposition technologies. This often involves optimization of the chemical reactions that occur during abatement. In many instances CO2 or H2 must be added to achieve the optimum conversion into less harmful chemicals and to avoid recombination. The Aeris-G uses water vapor to provide the needed O and H atoms when converting, for example, CxFy into HF and CO2; NF3 into N2 and HF; or SF6 into SO2 and HF (see figure 1).

The Applied Aeris product family has evolved in other ways as well (see figure 2). One development is a higher-power replacement plasma source to address increasing PFC gas flow rates and wider process windows.

Kits are now available to synchronize Applied Aeris abatement operation with the actual needs of the process to further reduce operating costs. One example is the synchronization of abatement operation to process cooling water (PCW) needs at any given point in time (see figure 3).

Figure 1. The figure shows a chemical reaction that typically occurs during pre-pump plasma abatement of F-GHGs in the gaseous effluent from dry etch chambers. At left are the F-GHG molecules to be treated. In the center, a plasma is applied to dissociate them. At right they combine to form compounds that have less environmental impact.

Figure 2. The figure shows drawings of the Applied Materials Aeris abatement product family. At left is the Aeris-G greenhouse gas product, with an installed base of >1,000 units. Center, a higher-power replacement plasma source for use with wider process windows. At right is a design under evaluation for abatement of hazardous effluent from deposition chamber-types.

Figure 3. Photo of a PCW kit used to synchronize abatement operation to process cooling water needs.
First, a little background. A number of the processes used to build high-aspect-ratio structures, which are based on low thermal rates, such as ALD or CVD with low material conversion efficiency, often require the use of energetic materials. Low thermal budget processes, such as metalorganic precursors, allow a reaction to take place with SiO₂ deposits in the foreline and the pump. This reaction converted the waste steam into SiF₄, which passed through the pump and was later removed in the local scrubber. Thus, the pump was able to run longer before maintenance was required. Moreover, there was no need to add any reactive gas to the system to facilitate these transformations, although processes other than HARP may call for it.

SUCCESSFUL RESULTS WITH HARP PROCESS

A first solution for the HARP process has been studied and tested successfully. Users of this process are no doubt familiar with the foreline and gate valve clogging it can bring, as well as the need for more frequent pump maintenance. The Applied Materials plasma-cleaning process enabled highly reactive F radicals to be recombined into F₂ molecules, with an associated loss in reactivity. Then, a second use of the plasma dissociated the F₂ molecules, allowing a reaction to take place with SiO₂ deposits in the foreline and the pump. This reaction converted the waste steam into SiF₄, which passed through the pump and was later removed in the local scrubber. Thus, the pump was able to run longer before maintenance was required. Moreover, there was no need to add any reactive gas to the system to facilitate these transformations, although processes other than HARP may call for it.

LOOKING FORWARD

Pre-pump plasma abatement technology continues to evolve in lockstep with increasingly stringent requirements for greener and more sustainable manufacturing. The next opportunity for this efficient, cost-effective, and space-saving solution will be for so-called “dirty” deposition chemistries.

For additional information, contact john.dickinson@amat.com or andreas_neuber@amat.com.

The semiconductor industry is in a dilemma, well described by Ben Eynon, senior director of engineering development at Samsung Austin Semiconductor, at the Advanced Process Control (APC) conference held in Austin in mid-October. Scaling to 10nm and beyond, Eynon said, means that chipmakers and their equipment and materials vendors “must attack everything simultaneously. The problems we have are large and must be solved perfectly and quickly in order to stay competitive.”

Eynon told the APC audience “we have to keep our equipment clean and our processes right on target to keep these defects down...whoever can keep the defects down, wins. We face process limitations and integration complexity. And it takes a lot of cross-communication to make things work well.”

All of the like tools in a fleet must be matched, and stay matched. The time required for feedback and feed-forward must be minimized. And standard operating procedures in the 10nm era and beyond need to change to custom operating procedures where each wafer is treated as an individual with different process requirements.

Chipmakers need “more sensors on everything so we can monitor more than we do today. We need feedback to adjust the next process, based on the just-prior process. I need to know if I run this stuff this certain way, I can be sure of what I will get for electrical test results and yield,” Eynon said.

He then shifted to the crux of the problem. With so much data being pulled off the tools that already is “spilling onto the floor,” Eynon acknowledged that the industry faces a general reluctance to share that data and lack the necessary protocols to do so securely.

“The tool vendors have their log files. We have our process. Our customers have their designs. Everyone has their own IP that they don’t want to share. We have to figure out how to handle IP better than we do today [to make sharing happen],” he said.

INTEL’S BREAD AND BUTTER

Steve Chadwick, senior engineer at Intel’s manufacturing IT operation, described a similar conundrum. After outlining several ways that Intel engineers use proprietary software to analyze some five billion data points each day, Chadwick was asked a question about sharing data in the cloud. He replied that Intel “has its own compute farms, both distributed and centralized,” and there is no problem taking expense reports, for example, out to the cloud. But Chadwick also said Intel would be very cautious about taking its manufacturing IP to the cloud. “You are talking about a company’s bread and butter, our R&D, and there is going to be a lot more study before we do that.”

Eynon added that Samsung also uses proprietary software to improve security, including its own email system. “That’s our mentality in semiconductor. I believe we are going to be using our own hardware to do our number crunching for a long while. Every time we hear about another case of identity theft, we realize this kind of activity is not going away. We are not going to swing to an attitude of ‘here is our IP, go crunch it.’”

James Moyne, a professor at the University of Michigan and semiconductor industry consultant on advanced factory automation, noted that all industries face security issues, but the chip industry is particularly concerned that IP not leak. “Compared to other industries, we are really skittish,” he said. While chipmakers are increasingly sharing data with their tool vendors, only a fraction of that data is being sent outside for remote analysis.

Nick Ward, director of marketing for the services group at Applied Materials, said that “Applied Materials completely errs on the side of caution” when dealing with the security of customer data by keeping tight controls and safeguards in place.

“We are dealing with proprietary protocols and proprietary approaches to data. As an industry, we need to develop a set of standards for protecting the shared data. Being able to integrate with certain systems and tools will speed time to problem resolution and create greater cost savings for manufacturers,” Ward said.

However, Moyne cautioned that companies need to keep working on the problem, or lose out. “The tighter the control specs and the more sophisticated the algorithm, the more suppliers and customers need to algorithm, the more sophisticated the algorithm, the more suppliers and customers need to be concerned about another case of identity theft, we realize this kind of activity is not going away. We are not going to swing to an attitude of ‘here is our IP, go crunch it.’”

James Moyne, a professor at the University of Michigan and semiconductor industry consultant on advanced factory automation, noted that all industries face security issues, but the chip industry is particularly concerned that IP not leak. “Compared to other industries, we are really skittish,” he said. While chipmakers are increasingly sharing data with their tool vendors, only a fraction of that data is being sent outside for remote analysis.

Nick Ward, director of marketing for the services group at Applied Materials, said that “Applied Materials completely errs on the side of caution” when dealing with the security of customer data by keeping tight controls and safeguards in place.

“We are dealing with proprietary protocols and proprietary approaches to data. As an industry, we need to develop a set of standards for protecting the shared data. Being able to integrate with certain systems and tools will speed time to problem resolution and create greater cost savings for manufacturers,” Ward said.

However, Moyne cautioned that companies need to keep working on the problem, or lose out. “The tighter the control specs and the more sophisticated the algorithm, the more suppliers and customers need to be concerned about another case of identity theft, we realize this kind of activity is not going away. We are not going to swing to an attitude of ‘here is our IP, go crunch it.’”

James Moyne, a professor at the University of Michigan and semiconductor industry consultant on advanced factory automation, noted that all industries face security issues, but the chip industry is particularly concerned that IP not leak. “Compared to other industries, we are really skittish,” he said. While chipmakers are increasingly sharing data with their tool vendors, only a fraction of that data is being sent outside for remote analysis.

Nick Ward, director of marketing for the services group at Applied Materials, said that “Applied Materials completely errs on the side of caution” when dealing with the security of customer data by keeping tight controls and safeguards in place.

“We are dealing with proprietary protocols and proprietary approaches to data. As an industry, we need to develop a set of standards for protecting the shared data. Being able to integrate with certain systems and tools will speed time to problem resolution and create greater cost savings for manufacturers,” Ward said.