NASDAQ:AEHR Q3 2026 Earnings Call Transcript

Generated on 11/25/2026

Operator | Conference Operator:

Greetings. Welcome to the Air Test System's Fiscal 2026 Third Quarter Financial Results Conference Call. At this time, all participants are in a listen-only mode. A question and answer session will follow the formal presentation. If anyone should require operator assistance during the conference, please press star zero on your telephone keypad. Please note, this conference is being recorded. I will now turn the conference over to your host, Jim Byers of Pondell Wilkinson Investor Relations. You may begin.

Jim Byers | Pondell Wilkinson Investor Relations:

Thank you, Operator. Good afternoon and welcome to Airtest Systems' third quarter fiscal 2026 financial results conference call. With me on today's call are Airtest Systems President and Chief Executive Officer, Gane Erickson, and Chief Financial Officer, Chris Tiu. Before I turn the call over to Gane and Chris, I'd like to cover a few quick items. This afternoon, right after market closed, Airtest issued a press release announcing its third quarter fiscal 2026 results. That release is available on the company's website at air.com. This call is being broadcast live over the Internet for all interested parties, and the webcast will be archived on the investor relations page of the company's website. And I'd like to remind everyone that on today's call, management will be making forward-looking statements that are based on current information and estimates and are subject to a number of risks and uncertainties that could cause actual results to differ materially from those in the forward-looking statements. These factors are discussed in the company's most recent periodic and current reports filed with the SEC. These forward-looking statements, including guidance provided during today's call, are only valid as of this date, and Airtest Systems undertakes no obligation to update the forward-looking statements. And now with that, I'd like to turn the conference call over to Gane Erickson, President and CEO.

Gane Erickson | President and Chief Executive Officer:

Thanks, Jim. Good afternoon, everyone, and welcome to our third quarter fiscal 26 earnings conference call. I'll start with an update on the key markets driving our business and strong demand we're seeing, particularly from AI and data center infrastructure. Chris will then review our financial results and we'll open up the call for questions. We're very pleased with the strong momentum in our business across multiple market segments, highlighted by more than $37 million in quarterly bookings and a book-to-bill ratio exceeding 3.5x. Our effective backlog, which includes the backlog of $38.7 million at the end of the fiscal third quarter, plus additional bookings received since the end of the quarter, is now over $50 million, a new company record. After generating approximately $20 million in bookings in our fiscal first half, we're already two and a half times that in second half bookings, and now expect to come in on the high side of the $60 to $80 million in second half bookings I mentioned last quarter. Demand continues to accelerate across both package level and wafer level burn-in, driven by increasing semiconductor complexity, power requirements, and deployment in mission-critical AI, networking, automotive, and industrial applications. As devices become more advanced, the need for comprehensive test and burn-in is becoming essential to ensure reliability and performance. This is driving growing adoption of our solutions across multiple markets. So let me start with wafer-level burn-in. During the quarter, we continue to make progress in growing our installed base and expanding to new customers with our wafer-level burn-in solutions. AI wafer-level burn-in is really hot right now, I guess pun intended. We received a $14 million follow-on production order from our lead wafer-level AI accelerator processor customer for multiple new fully automated Fox XP wafer-level burn-in systems to be used in data

center training and inference applications. The order included multiple additional Fox XP wafer-level test and burn-in systems, each configured to test nine 300-millimeter wafers in parallel, along with a set of AERS proprietary Fox wafer pack full wafer contactors and a fully integrated Fox wafer pack auto-aligner with each system to enable hands-free operation and high-volume production. In addition, the order included multiple additional Fox wafer pack auto-aligners to upgrade the customer's existing installed base of Fox XP systems to full automation. AIR is the first company to successfully demonstrate and ship a wafer level burn-in solution for AI processors. Our Fox XP systems configured for very high power, high current AI processors began shipping last year and provides the highest power per wafer capability available in the market, delivering up to thousands of amperes of current per wafer. This order further expands their installed base of Fox XP systems and adds full automation across their production lines, highlighting the growing importance of wafer-level burn-in to ensure the long-term reliability of today's very high-power, high-current AI processors. We're also actively engaged with multiple additional AI processor companies on benchmark evaluations and expect to make meaningful progress with those opportunities. Our benchmark evaluation program with a top tier AI processor supplier continues to make good progress, but it's taken longer than we originally expected. This was due to a technical misunderstanding on the clock configurations, which created some challenges with the initial wafer pack designs. While we wish we had been able to catch this earlier, we're taking device data now on their wafers with the current wafer pack design and redesigning the wafer packs to meet the new requirements. We expect to continue to provide them with additional data on this wafer pack design as well as the improved one over the next several months. We have several other companies ranging from suppliers of data center-focused AI accelerator processors to edge AI processors and CPUs that are providing us with information on their devices and roadmaps and asking about our wafer-level burn-in capabilities and recommendations for burn-in of their next generation devices. There is significant interest in doing wafer-level burn-in for devices that are expected to put in advanced packages such as TSMC's COAS-based packages that include other dyes such as HBM DRAM stacks, other compute AI processors, and photonic or electrical-based transceiver chipsets. Weeding out bad devices before they're packed together with these other devices is significantly cheaper than the yield loss if these are burned in at package level and the entire multi-chip package is thrown away. For burn-in of silicon photonics devices, we recently announced a major new customer win, a major new silicon photonics customer with an initial order for multiple high-power FoxXB wafer-low burn-in systems for devices aimed at the hyperscale data center optical interconnect market. This customer is developing advanced silicon photonics-based transceivers for data center networking and optical I.O. applications to address the rapidly accelerating demand for high-speed fiber optic communication links in hyperscale AI and cloud data centers. These multiple systems are for both engineering qualification and high-volume production and include a Fox XP wafer-low burn-in system configured to test nine wafers in parallel, a fully integrated wafer pack auto-aligner, multiple Fox NP wafer-low burn-in systems, and multiple full sets of Fox wafer pack full wafer contactors for production, engineering, and new product introduction. These systems are all scheduled to ship in this fiscal fourth quarter ending May 29-26. They've also provided a forecast for multiple additional XP production systems over the next year as they ramp capacity to support next-generation hyperscale data center deployments. We believe this win positions AI to participate in what could be a significant multi-year expansion of silicon photonics production driven by the growth of fiber optic interconnects in hyperscale AI data centers. Additionally, we received a follow-on order from our lead silicon photons customer for both a new high-power Fox XP wafer-level system and an upgrade of an existing system to our latest high-power fully automated configuration. We now have fully integrated our systems and aligners with their autonomous guided robots that carry around the 300-millimeter FOOPS so the customer can operate in a fully lights-out, hands-free operation. They, too, have given us a forecast for additional production systems as they ramp into next calendar year. As data center architectures scale to support AI, cloud computing, and high performance networking, fiber optic interconnects offer significant advantages over copper wiring, including higher data rates, lower power consumption, longer reach, improved thermal performance, and reduced electromagnetic interference. These advantages are driving rapid adoption of silicon photonics transceivers across hyperscale and enterprise data centers worldwide and increasing demand for cost-effective production-proven burn-in solutions that can ensure device quality and long-term reliability at volume. AIR is the market leader in wafer-level burn-in for silicon photonics transceivers, with a large installed base at leading global semiconductor and photonics companies. The company's RR Fox XP

platform enables high parallelism, high temperature, and high-power wafer-level burn-in, allowing customers to stabilize their devices a critical manufacturing process step in the laser diode emitters for these devices, as well as to identify early life failures before packaging to significantly reduce the cost of test. In gallium nitride and silicon carbide power semiconductors, we've been working with our lead GAN production customer on a significant number of new devices aimed at multiple markets that include automotive, automated bus conversion, data center, and electrical infrastructure. This continues to be a great partnership, and we continue to work on and believe we have solved the key challenges with full wafer burn-in of GaN devices on silicon. Wafer-level burn-in of their GaN devices for both qualification and production burn-in is an extremely valuable capability that is critical to their roadmap and plan, and we're both very excited to see them meet their growth projections. We continue to see GaN and silicon carbide power semiconductors as critical to the electrification of the world's infrastructure in addition to key market opportunities such as data center power delivery, electric vehicles, and charging infrastructure. We want a new customer in silicon carbide this quarter with a company in Taiwan focused on the Asian and particularly greater EV market, greater China EV market, sorry. They placed an order for a small configured Fox XP system For qualification and production, key elements of their decision included our ability to demonstrate all the capabilities they needed with our systems in Fremont, California, as well as the feedback they received from customers who had data and confidence in AIR's wafer-level burning systems used for testing and burning silicon carbide wafers across a large number of silicon carbide suppliers. We see an uptick in activity and forecasts from the silicon carbide players. This makes sense as we see major OEM EV suppliers in Japan and Germany roll out a number of new EVs later this year. These EV suppliers understand the value and need for wafer-level burn-in of these six devices before they're put into modules containing many devices in parallel for the EV engine drive inverters. This is well understood in the industry, and AIR is seen as the market leader and proven solution for wafer-level burn-in silicon carbide devices used in EV inverters by a significant number of EV suppliers. We're still conservative about forecasts from customers. And while we have plenty of capacity and believe we have the world's most cost effective and highest performance wafer level burden solution on the market, we're not yet counting on significant revenue from this segment to return yet. However, it could still be a very good performing segment for us next year. We'll see. Now let me talk about wafer-level burn-in for memory. Our engagement with a key memory supplier continues to progress with additional wafer testing just this last week. We've been able to achieve the correlation they're asking for are now in discussions about test system specifications needed for their next-generation flash memories, and in particular their high-end with flash devices. We hope to close on this in the next few months which would lead to a development agreement to supply systems and wafer packs to them after a 12 to 18 month development of our new memory optimized blades for our Fox XP and NP multi-wafer test and burn-in platform. But we're also now in discussions with other key memory suppliers that also produce high bandwidth memory. The new DRAM standard being used in AI GPUs in addition to standard DRAM and flash memories. The HBM memories, as they're referred to, are embedded into multi-chip packages with advanced substrates, such as the CoAS packaging from TSMC. NVIDIA's roadmap is aggressively pushing toward higher capacity, faster HBM standards to address the memory wall in AI training and inference. The upcoming roadmap transitions from HBM3e to 4 in 2026. and then from HBM4E and HBM5 in the following years, with capacity per GPU expected to increase from 80 gigabytes in the A100 class to over a terabyte in the Rubin Ultra by 2027 for semi-analysis. We are seeing the added potential for HBM insertions with our FOX multi-wafer test and burn-in system roadmap that extends to flash, high bandwidth flash, DRAM, and HBM memories. This is a key focus for AIR this year, to drive to an agreement to work with these customers in the development of the enhancements needed to extend our FOX systems to these markets. This is a market that we believe could drive orders in fiscal 27 with ramps in fiscal 28. Now turning to package level burn-in. Let me start by highlighting that we're trying to change our own vocabulary from package part burn-in to package level burn-in. This may seem subtle, but to give a little background, traditionally there was one semiconductor integrated circuit per single package. The package was used to protect the die from elements and wire out to a standard pattern of pins or pads that allowed easy handling and assembly onto a printed circuit board. This pattern or pitch between pins is much, much larger than the pitch on the individual die. So contacting the devices is very different for us between our package level and wafer level solutions. Historically, about 20 years ago, there was a package concept called multi-chip packages where multiple individual die were wire bonded into a single package. This was driven at

the time for size and performance. Typically, this was much more expensive and generally this faded out in time to other smaller package sizes. Recently, in the last handful of years, there have been three major drivers of the need for new multi-chip packages. but this has been called advanced packaging or modules rather than MCPs. One driver, which is the biggest one, is that the multi-decade long trend that was referred to as Moore's Law has come to an end. This law was the number of transistors was doubling every one and a half to two years while the die size was staying the same and therefore costs were staying flat or decreasing. This allowed higher and higher performance, smaller die, and therefore lower cost die to be made via process improvements or die shrinks. This drove the industry for 40 years or so until around 2010 plus or minus when shrinks started to slow materially. Then as several applications such as AI processors, extremely high density memory such as flash and DRAM, power semiconductors were being driven by massive markets such as data center, AI and electric vehicles. The extremely high value and need for multiple devices in the same package came to fruition. This time it was functionality and feasibility that drove this. We now refer to these devices in two camps, really three camps, wafer level, die level, and package level. Where package level includes both single die per package and also multi-chip modules or advanced package multi-die packages such as those found in AI GPUs with HBM DRAM stacks. multi-stack flash SSDs, and also multi-diacyl and carbide modules for EV inverters and charging infrastructure. At least I hope this helps as we talk through this and make it more clear what the difference is between wafer level and package level. You may catch me still saying package part at times as old habits are hard to break, but we'll try to refer these as package level from now on. Okay. During the quarter, we announced a key production win with our lead package level hyperscale customer. This customer is a premier large-scale data center provider and selected air for production burn-in of their next generation significantly higher power AI processor with an initial production order of our high power Sonoma systems. This next generation AI ASIC is expected to move to production later this year and is believed to be even higher volumes than the first device that this customer is ramping our Sonoma systems on right now. We also expect a significant near-term follow-on order from this customer for package-level burn-in systems to support their high-volume manufacturing of their custom AI processors today, the current one used in data center training and inference. They are forecasting a substantial expansion of Sonoma Systems purchases beginning the second half of calendar 2026 and continuing into 2027. We believe it's likely that there is overlapping ramps between the current and next-generation devices, which should significantly expand both our install base and long-term consumable opportunity with this customer. We're also engaged with multiple potential customers for package-level qualification tests of AI accelerators, ASICs, network processors, and edge AI processors for automotive and robotics. These engagements also represent opportunities to move to production burden over time And interestingly, about half of these have also expressed interest in wafer-level burn-in in addition to our package-level burn-in solutions. Yesterday afternoon, in fact, we received an order from a brand new customer for Sonoma to be used for reliability qualification of their new AI processor. But they may also do production burn-in with this device, which they can do with the exact same platform using Sonoma. This momentum reinforces our leadership in high-power burn-in for AI processors. The broader demand environment remains very strong. Industry forecasts indicate that hyperscale data center capacity expected to nearly triple by 2030, driven by both new builds and upgrades to existing infrastructure. This is driving substantial growth in high-performance semiconductors and, in turn, demand for advanced burn-in solutions. As we've noted before, As our install base of systems continues to grow, our consumables, which includes our wafer pack, full wafer contactors for wafer level, and our burn-in board and modules for package level burn-in, can continue to grow beyond our systems. While this year has been lighter in terms of consumable sales, particularly wafer packs, we believe it's an outlier. Some customers had bought systems ahead of the need and have grown into capacity, and this seems to be running its course. We believe over time our consumables business will consistently be at 30% or more of our total revenue, and our margins will increase as sales of these value-add consumables grow. To support growing demand, we're continuing to scale manufacturing capacity. In addition to our Fremont expansion, this quarter we'll begin shipping Sonoma systems from one of our current contract manufacturers, adding capacity of more than 20 additional Sonoma systems per month. This meaningfully increases our ability to support future growth. With expanding AI infrastructure deployments and our recent manufacturing capacity enhancements, we believe we're well positioned to support significant growth both in our wafer level and package level burden systems as customers ramp production. With strong second half booking so far

and a strong funnel of additional orders expected this quarter, We believe we're well-positioned to exit the fiscal year ending May 29th with a strong backlog and deliver significant revenue growth in fiscal 27. We currently expect full-year fiscal 26 revenue to be on the high side of the $45 to $50 million range provided last quarter. We also expect our bookings for the second half of the fiscal year to be on the high side of the $60 to $80 million range provided last quarter. More broadly, we believe we have a clear path to sustain long-term growth as our installed base expands across AI, silicon photonics, power semiconductors, memories, and other high-performance applications. As semiconductor performance and reliability requirements continue to rise, burn-in is becoming increasingly critical across a growing set of applications. We believe AIR is uniquely positioned as the only provider offering both wafer-level and package-level burn-in solutions at scale. That'll turn it over to Chris.

Chris Tiu | Chief Financial Officer:

Thank you again, and good afternoon, everyone. I'll begin with bookings and backlog, then walk through our third quarter financial performance, cash position outlook, and investor activity. The company recognized bookings of $37.2 million in the third quarter of fiscal 2026, significantly higher than the $6.2 million in the second quarter as we have received multiple purchase orders for Fox Systems, wafer packs, and Sierra auto aligners from different customers for AI and silicon photonics and silicon carbide applications. At the end of the quarter, our backlog was 38.7 million. During the first five weeks of the fourth quarter, we received an additional 12.2 million in bookings. This increase was driven primarily by a major new silicon photonics customer for wafer-level burn-in with an initial order for multiple FOX systems for both engineering qualification and high-volume production, which we recently announced. With this recent bookings, our effective backlog, which includes our quarter end backlog plus additional bookings received since the end of the third quarter, has now grown to a record of 50.9 million, providing strong visibility for the remainder of fiscal 2026 and positioning us for significant growth for fiscal 2027. Our strong bookings include increased demand for both our wafer level and package part, package level burn-in solutions. We believe this reflects the proven value of these differentiated solutions, which are increasingly integral to the production and reliability strategies of our customers in the AI, data center, and other key markets we serve. Turning to our Q3 performance, while we did not provide quarterly guidance, our third quarter revenue of $10.3 million was in line with internal expectations due to delayed orders. Q3 revenue was slightly below consensus and down 44% from $18.3 million prior year period. The decline was primarily driven by lower shipments of FOX systems and wafer packs for wafer-level burning business, partially offset by stronger demand for our Sonoma systems and BIMs from our hyperscale customer. Contacted revenues, which include wafer packs for wafer-level burning business and BIMs and BIPs for package-level burning business, total $3 million, representing 29% of total revenue in the third quarter. This compares to $5.9 million, or 32% of revenue, in Q3 last year. Non-GAAP growth margin for the third quarter was 36.5%, compared to 42.7% a year ago. The year-over-year decline reflects lower overall sales volume and a less favorable product mix, as last year's quarter included a higher proportion of high-margin wafer pack revenue. Non-GAAP operating expenses in the third quarter was $6.3 million, flat from $6.3 million in Q3 last year. We continue to invest significant resources in our AI benchmark and memory projects. During the quarter, we recorded an income tax benefit of $0.8 million, resulting in an effective tax rate of 19.9%. Non-GAAP net loss for the third quarter, which excludes the impact of stock-based compensation and acquisition-related adjustments, was $1.5 million, or a loss of $0.05 per diluted share, compared to net income of $2 million, or $0.07 per diluted share, in the third quarter of fiscal 2025. Non-GAAP net loss for the third quarter exceeded consensus by $0.02. Turning to cash flow. We used 3.7 million in operating cash during the third quarter. We ended the quarter with 37.1 million in cash, cash equivalents and restricted cash, up from 31 million at the end of Q2. The increase was primarily due to proceeds from our at-the-market or ATM equity program. During the third quarter fiscal 2026, we raised 10.5 million in gross proceeds through the sale of about 269,000 shares. Since the end of Q3, we've raised another 19.5 million gross proceeds through the sale of about 477,000 shares. And with the 9.9 million we raised in Q2, we have now fully utilized the 40 million available under the ATM and have sold over 1.13 million shares at an average price of $35.38. We also announced this afternoon that we'll be changing our

fiscal year from the last Friday of May to the last Friday of June effective after our fiscal year ends on May 29th, 2026. Our new fiscal year, 2027, will begin on June 27th, 2026 and end on June 25th, 2027. Continue with the 4-4-5 calendar. As a result, we'll have one month of financial results from May 30th to June 26, 2026, which will be reported as a transition period when we file our quarterly form on 10Q in the first quarter ending September 25, 2026. We believe our new fiscal year will align more closely with the reporting periods of our customers and our peers in the semiconductor test equipment industry. Moving to our outlook, for the full year fiscal 2026 ending on May 29, 2026, We currently expect total revenue to be on the high side of the $45 million to $50 million range provided last quarter, and non-GAAP net loss prediluted share to be between negative 13 cents and negative 9 cents for the full fiscal year. We expect our gross margin to improve as our manufacturing activity increases to support higher sales volume and better absorb our fixed costs. We also expect to return to profitability on a non-GAAP basis in the fourth quarter of fiscal 2026. Lastly, looking at the investor relations calendar, Airtest will be participating in two investor conferences over the next couple of months. We'll be meeting with investors at the Craig Hallam Institutional Investor Conference taking place in Minneapolis on May 28th, and we'll be presenting a meeting with investors on June 2nd at the William Blair 46th Annual Growth Conference taking place in Chicago. We hope to see some of you at these conferences. That concludes our prepared remarks. We're now happy to take your questions. Operator, please go ahead.

Operator | Conference Operator:

Thank you. At this time, we will be conducting a question and answer session. If you would like to ask a question, please press star 1 on your telephone keypad. A confirmation tone will indicate your line is in the question queue. You may press star 2 if you would like to remove your question from the queue. For participants using speaker equipment, it may be necessary to pick up your handset before pressing the star keys. One moment, please, while we poll for questions. Once again, please press star 1 if you have a question or a comment.

Operator | Conference Operator:

Our first question comes from Mark Shooter with William Blair.

Operator | Conference Operator:

Please proceed.

Mark Shooter | Analyst, William Blair:

Hey, Gain. Hey, Chris. You have Mark Shooter on here for Jed Door Timer.

Jim Byers | Pondell Wilkinson Investor Relations:

Hey, Mark. Hey, Mark.

Mark Shooter | Analyst, William Blair:

Congrats on all the progress, especially with the hyperscaler. I'm curious how you guys are looking at this internally, and what percentage of GPUs or ASICs or XPUs do you think are burnt in today? And how do you guys size the vector space?

Gane Erickson | President and Chief Executive Officer:

That's a really good question, and I think we're still getting our arms around it a little bit here. I would say that we've been a little bit surprised at how – many devices are not yet doing production burn-in. You know, one of the things that we mentioned strategically when we purchased NCAL, what, 18 months ago or so, Intel does a type of burn-in, and they were known for it, called qualification reliability burn-in, which all processors go through, in fact, all semiconductors. It's what determines their lifetime reliability specs and that they will last long enough, et cetera. It's sort of a one-time deal you do with a large number of devices to do the statistics on it. Then certain devices go through a screening in production to weed out infant fatalities because the failure rate is higher than the market will bear. So InCal was doing this with a large number of AI customers, but actually prior to that wasn't doing any production burn-in. When we acquired them, we've now, because of the capacity we have in terms of people and infrastructure, we've been able to capture this large hyperscaler and are engaged with multiple others. But one of the things that I've been surprised at is that how many of the, I guess, particularly the ASIC suppliers don't do production burn-in yet or are talking about doing it. And that goes for a lot of different devices that are out there from edge, robotic, ASIC network processors, and even, you know, I've got to always be careful with GPU because everybody just associates GPU only with NVIDIA. But, you know, not all devices are burnt in still today. And so there are certain ones that are, there are certain ones that aren't. And even within a company, they may have some of their products are burnt in and others aren't. However, the common theme is they're all moving to burn in. The data is out now that there's solutions out there like Sonoma or the wafer-level burn-in of our Fox system that can cost-effectively do it. And so now there's a very viable alternative to doing it at the system level or the rack level. We've said in the past that many of these guys would actually build it all the way to the rack. And then at the system integrator, they would burn it in for a week or two and weed out the infant mortality to ship it. Or in some cases with the ASIC suppliers, they just shipped it into their data centers and and dealt with the fallout. So it's growing. I'm trying to think if I'd try and put a percent, I think on ASICs, it might be, you know, by unit, like skew. I mean, I don't know if it's 20%. Maybe it's 5% of the number. So most ASICs are not printed. I would say on The AI accelerators that are out there across the wide variety, you know, maybe half. But then what's happening is the processors are getting higher power from generation to generation and breaking all the tools that are out there. So even the tools that were out there, and I'm not giving any inside information whatsoever, but just what's classically understood. like NVIDIA's processors of three couple generations ago compared to their current ones, their power is substantially more which would require new tools. And the ones that they're working on and others in a year and out, and again, just what's publicly available, break the current tools. And so there's a continuous roadmap. And so even within our Sonoma platform, We're continuing to add capabilities. One of the key features we have is the ability to adapt it and add higher and higher current and power as you go forward. So there's, you know, how many times you hear a CEO say you're at the early innings, but this is still at the kind of the beginning phases of this. And over time, people will be buying a lot more burn-in systems as a percentage, meaning to cover the percentage of total, and then just ensure quantity.

Mark Shooter | Analyst, William Blair:

I appreciate all the color gain. That's very helpful. To zero in a bit on your hyperscaler customer, can you bring us a little into the room a bit here? And what was the decision process to go with package level, right? Not package part anymore. It's package level versus wafer level. And do you see a transition potentially with this customer to move to wafer level? And if you get a new customer... Do you think that they'll make the same decision, or is there a track towards wafer-level? Try to help us out with that.

Gane Erickson | President and Chief Executive Officer:

Okay, so to be fair, two, three years ago, if you would have asked me, I've said this before, can you do wafer-level burn of AI processors? I think we would have said absolutely not. We didn't have the power in the system, and the belief was that there weren't the test modes that we now understand there are to be able to do it. And now as we've gone from customer to customer across a wide variety, there's commonalities about it that allow us to be able to confidently tell them we can do wafer-level burn-in. So prior to that, it was whether you did package-level burn-in or not, or did it at, say, the rack level. So people first step is, do I do burn-in? Then they're going to default to thinking, I'm going to do it at the package level. But then what we're seeing, and I mentioned this before, we have customers, I don't want to get too carried away here, but the last two customers that we're in in the last two weeks, Alberto is our package level Vernon VP, and Vernon really runs kind of the wafer level side of things. The customer will come in and say, I want to talk about package level. And about halfway through the tour, they're like, what is that? We talk about wafer level. They're like, whoa, whoa, whoa, whoa. How do I do that? And so we kind of joke about it around here. It's like, ah. But the reality is we don't care which side you go to. We have both. Specifically on the hyperscaler, and I've said this out loud before, the first device they ran with us, it's not their first device, but it's the first one they went to production on, is on Sonoma. Their second device, they just awarded us with production for that one and are planning the ramp of that with us right now. They're already on the roadmap talking about the third device, and they've asked us about the DFT to specifically put into the third device because they would like to consider that for wafer level on our Fox systems. So I think that's sort of a progression that we will see, and I would actually imagine large customers that have multiple different product lines Some they would do wafer level on and some they might do package level on. It becomes particularly valuable when you have a package that has multiple processors in it and all the HBM memory. So in those particular ones, I mean, the co-op substrate is more expensive than the silicon itself of the processor, which sounds crazy. So they would be very interested in doing the wafer level to screen out the dye before they have to throw away everything else. So I think there's a progression over time where people will move towards wafer level on the things they can, default to package level where they can't.

Mark Shooter | Analyst, William Blair:

That's really helpful. Thank you, Ken. Thanks, Mark.

Operator | Conference Operator:

The next question comes from Christian Schwab with Craig Hallam. Please proceed.

Christian Schwab | Analyst, Craig-Hallam:

Hey, good afternoon. Thanks for a tremendous amount of detail regarding the different target markets and your success in each one of them. The most common question I receive is, is there a way to gauge over a multi-year time frame? Obviously, you gave guidance for this year in support of substantial growth the following year with bookings in hand and others to come. But have you had enough time to give some thought to the range of potential outcomes over a multi-year timeframe? that you guys could do in combination of your target markets and potential entry into the market, memory market down the road?

Gane Erickson | President and Chief Executive Officer:

So the short answer is we have. The long answer is we're just really cautious about trying to get too carried away with our projections. But the numbers are very significant. If you just, because, you know, particularly now that there's some HS that hung down with memory kind of angle on this thing too, if you look at the dollar

spend that people are going to do on, whether you call it compute or AI or, you know, if you look at the compute capability, right, that are going into training and inference in data centers, inference in edge, automotive, robotics, you know, the number of different applications and the way people are using it and deploying it. The amount of silicon wafers is staggering. And, you know, why people talk about these enormous dollars. Those devices, a processor has always been burnt in. It feels like I'm contradicting what I said earlier. You know, it's widely known that Intel and AMD, the primary processor suppliers of the world, burnt in every one of their processors and always have. When the first GPUs were coming out, those were using graphics, they were not burnt in. And the initial people that are all related to AI are our foundries, and they're out looking for burning capability. There were no burning systems in the foundry OSAP models. And so people weren't spending on it. They spent enormous amounts of money on tests, and it's growing, and they're going to be spending a significant portion of their test budget on burn-in going forward. I hear things, I mean, I hear it constantly from the customers rotating through. So, you know, the TAMs are, you know, multi-hundreds of millions of dollars for, you know, package-level burn-in, wafer-level burn-ins, If you say it displaces, package level is even higher. The average actual price per unit time of wafer level is actually more expensive than package level. But the yield pays for all of it. And so it's cheaper to the customer to spend more money. And so the TAMs are larger there. If you look at the memory side of things, if you look at the memory spend of the number of fabs that are coming out in the next you know, five years, what percentage of budget is for their test budget, these are big numbers. And so, you know, the spend is in Vernon is probably total spend measured in multiple billions of dollars per year in the next couple of years, you know, on an annual basis. And, you know, the question is, well, then wait a minute, how come you guys aren't, you know, $500 million? And the answer is, We think that we have a very good opportunity to significantly grow our package-level and wafer-level business across the biggest segments that are driving burden. And one of the reasons we're leading with putting infrastructure and capacity in place to be able to have the conversations we're having with these customers, they're throwing out some really big numbers. And somebody in legal... going to warn me, you're getting carried away here, but it's an awesome place to be. And it's not only silicon carbide for EVs. Lots of people are wondering if the EVs are ever going to make it. As you guys know the history, it's like people got ahead of themselves. And I was even saying it. It's like, come on, you guys. We're not all going to be driving EVs. But the TAMs in these segments are significantly larger than anything we ever talked about on the power semiconductor side.

Christian Schwab | Analyst, Craig-Hallam:

Great. That gives me enough to work with. No other questions. Thanks, Kane.

Gane Erickson | President and Chief Executive Officer:

Thanks, Christian. Thank you.

Operator | Conference Operator:

The next question comes from Max Michaelis with Lake Street Capital Markets. Please proceed, Max.

Max Michaelis | Analyst, Lake Street Capital Markets:

Hey, guys. Thanks for taking my questions. First, I want to start out here. We look at the demand environment from the package level and wafer level. So demand seems strong on both sides of the business here. But, I mean, to me it looks like wafer level is outpacing on the demand side and maybe the order side. Can you let me know if I'm wrong there? But anything else you can add as well?

Gane Erickson | President and Chief Executive Officer:

The challenge with our business and for all of our shareholders is we know how to be lumpy. And, you know, by having more markets, and more customers, it can make it less lumpy. But the ASP of a production order, you know, a set, you know, in wafer level burn-in, you know, can be $10 to $20 million in an order, let's say, okay? Package level can be that big or bigger, too, okay? So when they come in, it looks like, oh, right now, we see demand on both significant. Now, the engagement level And the work to get a wafer level burn-in is definitely harder than package level. And the obvious reason is, in many cases, we're already testing the part for the call on our tool. So now they have to just say, oh, I need to buy a whole bunch of them and add automation and go to production. Does that make sense? On wafer level, what we found is that there's a learning process by both sides a little bit, but to understand how they can use our tool to be able to test their part. And in some cases, they're like, okay, I know if I just did this, it would make it a lot easier, but it's too late. I already taped out this part. That would be an example of this benchmark I'm in right now. It's like they're having to use some little fancier wafer pack to do it. And if they just did some specific DFT, they could use a very simple wafer pack, the same wafer pack we're using for like silicon photonics or silicon carbide and some of these others. Their vocabulary with us is, oh, I'll be able to do that for the next gen. But can you just work around it with the current one? Well, it's kind of harder. The other one, as I mentioned, I want to get a little too carried away. I mean, I get pretty techie on these things. But, you know, we had a miscommunication on the clocks, which is something really simple, candidly. But if you do them wrong, it doesn't work. And so we've had to jury rig some stuff to actually get it to work, and we're going to spin it to make it work. Nobody's freaking out about it because this isn't rocket science, but it would be something we would never mess up again with that customer because we now both have the same vocabulary. Second one is always easier. And so there's a little bit more startup thing with the wafer-level burn-in, but if you're technically astute and engaged and you look at it, you're not going, oh, this isn't going to work. You just go, okay, gosh, that's too bad. Okay, now let's keep going. And so there's a learning process. We're getting faster at it. And I think over time, wafer-level burn-in, like the silicon carbide or the silicon photonics customer we won this last quarter, it was just, yes. I mean, there was no on-wafer benchmark. It went from can you do it to how fast can you deliver. I think that is a natural progression. You'll see it in our package level, and you'll see it in our wafer level over time where customers will engage. They'll know we can do it, and they'll just think, let's go.

Max Michaelis | Analyst, Lake Street Capital Markets:

Perfect. Thank you. That's it for me. Thank you. Thanks, Max.

Operator | Conference Operator:

Once again, if you have a question or a comment, please press star 1. The next question comes from Larry Shlebina with Shlebina Capital. Please proceed.

Larry Shlebina | Analyst, Shlebina Capital:

Hi, Dane. Your contract manufacturer that you're starting up, when does that start and when will it be fully capable of doing your 20 Sonomas a month?

Gane Erickson | President and Chief Executive Officer:

They're in the process of building the first batch, I would say, is the best way of looking at it. It's a little more complicated than the way I described, but there's actually two contract manufacturers together, and then one feeds into the other one. The one that feeds into the other one did their prototypes. They sent to us. We were

going through kind of a an acceptance process to validate it to work out any kinks. Then those go to the other contract manufacturer for final system integration and shipping. The other one is when we were out there, we visited them last September, I think. We did kind of an audit of facility power infrastructure and cleanliness, and they did a kind of a remodel similar to ours if people have seen it. You know, it's all white and fancy, clean floors. more clean room space so that we can actually build these things in a clean room area. They had facilities. They were doing some stuff for solar, as it turns out. And so we were able to leverage from that. And that is in place now. And we think first products would be ready to ship to customers this quarter, you know, through May. And what we want to make sure is they're ready to go by late summer, when we see this Sonoma ramp hitting.

Larry Shlebina | Analyst, Shlebina Capital:

Right, that was the root of my question. Are you keeping any capacity? Are you planning on producing those systems in Fremont as well?

Gane Erickson | President and Chief Executive Officer:

Oh, yes, for sure. This is in addition to, we've kind of talked about like about a 20 system per month capacity here from an infrastructure and footprint perspective. We'd actually still need to hire some more people, maybe take on a shift. But we'll use – we use that facility for, like, large volume orders of the same SKU, if you will, make it simple. And then we'll use – we'll continue to make Sonoma systems here, and all of the XPs will be built out of here, all the Fox products.

Larry Shlebina | Analyst, Shlebina Capital:

And then did I hear you say that your first expected XP sales to an HBM customer will be – This calendar year or this or next fiscal year, 27?

Gane Erickson | President and Chief Executive Officer:

Yeah, I didn't quite say anything. I was a little more elusive than that on purpose. What I will tell you is that we have identified some interesting opportunities with HBM, probably the new 4E, that it has some interesting challenges that people would really like to do this wafer level burn-in on. And between our Fox system as it stands and the roadmap that we've been working on, as people know, with a team of people here for a memory extension to the Fox system to add what we would call channel modules into the Fox that make it memory focused, we think that there's some real overlap there. That just, as you know, Larry, you follow this a lot, that's an uptick. I thought HBM had a past flash, and it is in parallel with flash now.

Larry Shlebina | Analyst, Shlebina Capital:

I would say that would be an uptick. Yep, I agree. A little bit of an uptick.

Gane Erickson | President and Chief Executive Officer:

In order to be a good uptick, I'll agree with you. But right now, I'm excited about the discussions.

Larry Shlebina | Analyst, Shlebina Capital:

Yep. So the flash engagement, is that? Do you think that will bear fruit on the enterprise side here shortly before HPF gets underway, the effort that you're going to have on that front?

Gane Erickson | President and Chief Executive Officer:

That's a good question. You know, I think it really is up to the customer. Kind of the timing of what we would build would be something that would be a superset that could do both. So, yes, if HPF were delayed a little bit, maybe we would intercept their standard products. They've asked us to build it. The definition discussion has been to do both. In some ways, HBF is easier. Because if you start saying it's all flash, a lot of times what happens is people say, well, I want to be able to test everything I've ever had before. And then as the interfaces evolve, they tend to converge in voltages and speed or whatever. And if you say, well, I want legacy, it's like, well... okay, I've got to support this old voltage or something on a device you don't make anymore. So part of the challenge for us is to try and kind of converge on what do you really need going forward? You know, where are you going to spend the money? They'll probably never buy a system for legacy products from us in general. So I think that's one of the challenges we get to work through.

Larry Shlebina | Analyst, Shlebina Capital:

Well, that's all I have. Boy, you got a lot of irons on fire.

Gane Erickson | President and Chief Executive Officer:

It is. It is so much fun, you guys. I'm telling you. Yeah, Vernon and Alberto and the R&D teams and the, you know, poor Nick and our wafer pack team is very busy right now. And we're doing some things to offload that, adding additional resources. We're hiring anybody looking for a great job with a company that's growing, you know, let us know. We've got a lot of recs out there and we're looking for great people.

Larry Shlebina | Analyst, Shlebina Capital:

It sounds like it is a lot of fun, and congratulations. I know you've been working at it for a good while to get to this point.

Gane Erickson | President and Chief Executive Officer:

Thanks, Larry.

Larry Shlebina | Analyst, Shlebina Capital:

Take care.

Gane Erickson | President and Chief Executive Officer:

Thank you.

Operator | Conference Operator:

Once again, if you have a question or a comment, please indicate so by pressing star 1 on your phone.

Gane Erickson | President and Chief Executive Officer:

All right, Operator, if there's no other questions, we'll end on that really happy note. And as always, if you guys have any questions, please feel free to reach out to us. If you happen to be in the Bay Area and want to try and stop by, we're always happy to, you know, give a short tour to key investors and things like that. And we look forward to a great quarter talking to you next quarter. I guess with our new fiscal year, Now, our quarterly earnings will be the same time, the next time, and then there'll be, I guess, a one-month push out or something like that. But it should work out. This will be a good thing for our customers, which, honestly, that's the key to all of this. All right. Thank you very much, folks. Bye-bye.

Operator | Conference Operator:

Thank you. This concludes today's conference, and you may disconnect your lines at this time. Thank you for your participation.