

### **Forward Looking Statements**

This presentation contains forward-looking statements that involve risks and uncertainties relating to projections regarding industry growth and customer demand for the Company's products. Actual results may vary from projected results. These risks and uncertainties include without limitation, acceptance by new customers for wafer level and packaged part test and burn-in systems and contactors, world economic conditions, the Company's ability to maintain sufficient cash to support operations, and the potential emergence of alternative technologies, which could adversely affect demand for the Company's products. See the Company's recent 10-K and 10-Q reports filed with the SEC for a more detailed description of the risks facing the Company's business. The Company disclaims any obligation to update information contained in any forward-looking statement to reflect events or circumstances occurring after the date of this presentation.



## **Presenting Today**



Gayn Erickson CEO, Aehr Test Systems



### **Aehr Test Systems Company Overview**

#### Semiconductor Test & Burn-in for over 45 Years!

 World-wide leader in wafer-level test and burn-in systems

 Unique full-wafer test and burn-in systems and contactors

 Technology leader in massively parallel and high-power test and burn-in systems



High Power Multi-Wafer Test & Burn-In System



TAH<del>\_</del>E

High Power Test & Burn-In System





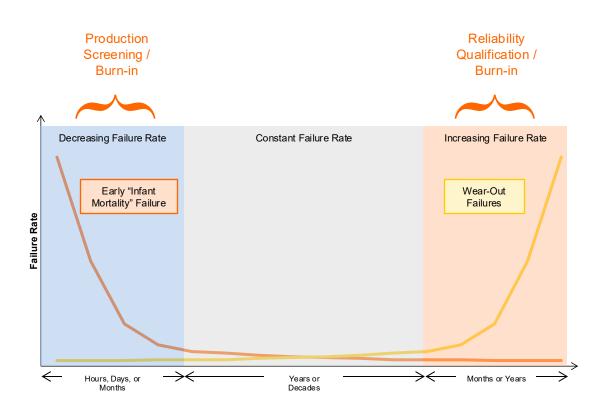
Single Wafer Stepping Test & Burn-In System



Medium Power Test & Burn-In System



### **Burn-in Testing – The Bathtub Curve**



- Production Burn-in: semiconductor components are subjected to elevated voltages and temperatures for a duration of time (2 – 48 hours) to screen for reliability and early failure in production before shipping to customers
- Reliability Qualification Burn-in: semiconductor components subjected to elevated voltage and temperatures for 1000 hours to validate and meet industry standards for long term reliability via High Temperature Operating Life (HTOL) tests



### **Worldwide Customer Base**



Aehr has been a leader in burn-in test solutions for over 45 years with thousands of systems shipped worldwide

(Partial Customer List)



### **Proprietary Wafer Level Enabling Technology**

- Aehr's FOX-XP is capable of both functional burn-in and test solutions leverages proprietary aligner and contactor technology
- Multi-wafer technology enables customers to achieve an overall decrease in test equipment cost and fab footprint while increasing die yield and throughput



- May be configured with up to 18 Blades, enabling 18 wafers to be tested in parallel – driving cost efficiency and throughput
- High performance thermal chucks allow uniform temperature control of the wafers
- Footprint similar to single wafer automated test equipment – reducing lab test space



- Houses the wafer and distributes signals and power to the wafer itself
- The WaferPak contactor is capable of over 50,000 contacts in a single touchdown on up to 300mm wafers
- Consumable input into the test system driving recurring revenue from the installed base



- Integral piece of test cell as it loads the wafer in the WaferPak at immense levels of precision
- By perfectly setting the wafer in a cartridge it ensures perfect contact
- Performs wafer alignment "offline" which eliminates the need for one wafer prober per wafer during long burn-in and test times



#### Aehr Wafer Level Test & Burn-in Patents





- WaferPak temperature control methods
  - Vacuum & pressure-based WaferPaks
    - Maintaining probe contact over temperature
      - Electrical components in WaferPak
        - Individual DUT power supplies
          - Per Die Current Protection
            - Redundant power supplies
              - Portable WaferPaks
                - And more . . .



### Sonoma High Power Reliability Test of Al xPUs

- High-power test & burn-in for production, reliability qualification, and life-time tests
- Up to 88 processors with independent test resources and high-power liquid cooling per device
- Qualification and Production configurations
- Shipping in volume with multiple systems installed at most test houses today







### **Aehr Test Systems Market Drivers**

Need for cost-efficient wafer level, singulated die, and packaged part burn-in, stabilization, & testing is creating significant revenue opportunities for Aehr Test in the following key markets

- Artificial Intelligence Processors and Processing Infrastructure driving explosive spend in data center processing, edge processors, communication infrastructure, and power conversion infrastructure drives AI processors, memory, data storage, Silicon Photonics I/O, and power conversion semiconductors like Silicon Carbide & Gallium Nitride
- <u>Electric Vehicle & Electrification of Transportation Infrastructure</u> driving motor control, charging infrastructure, and power conversion using Silicon Carbide & Gallium Nitride semiconductors
- <u>Electrification of the World's Power Infrastructure and Shift to Clean Energy</u> driving efficient and economical electrical power conversion using Silicon Carbide & Gallium Nitride semiconductors
- <u>Data Center Infrastructure and unstoppable growth in Data Storage</u> driving Silicon Photonics,
   Flash Memory Solid State Data Storage, and new Photonics Assisted Hard Disk Drive Storage
- Worldwide 5G Infrastructure build out using Silicon Photonics fiber optic transceivers and new Optical Network Switches
- Heterogeneous Integration of semiconductors and 3D fabrication and stacking driving technology and cost roadmaps pushing known good die with test and burn-in of device in wafer form prior to packaging



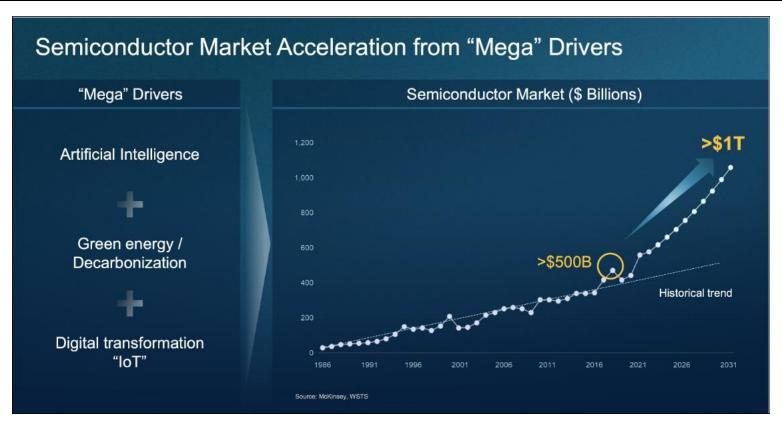








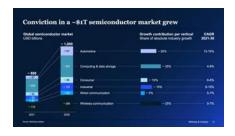
### **Aehr Market Growth Drivers**





### **Aehr Test Systems Market Growth Drivers**

- Reliability test needs to grow with or ahead of the market as
  - Semiconductors are getting less reliable
    - √ Larger die size and therefore more likely to have defects
    - √ Smaller geometries
    - ✓ More compound semis (SiC and GaN for power, and optical for communication)
  - Semiconductors going into more applications where quality, long term reliability, safety, and security are critical
    - ✓ Al Processing Custers (reliability critical due to impact on training, inference, cost, and even safety)
    - Autos / EVs (reliability more financially impactful, out of warranty recalls, processing, communication, ADAS, autonomous driving, power train, BMS, charging)
    - ✓ Electrification of WW infrastructure (charging, power storage, power transmission and conversion, efficiency in data centers, industrial)
    - ✓ Semi content increasing in value in non consumer applications (AI, data center, Autos/EVs)
- Semiconductors packaged as modules, Heterogeneous Integration, packages optics due to more than Moore, Huang's Law
  - Data and power density, drives greater spend in package test and burn-in, reliability qualification and wafer and die level test and burn in







### Al Processor Cluster Reliability is Critical

- **Impact on Training:** Al workloads, especially for training large language models, often utilize massive processor clusters. Failures in these clusters can have a cascading effect, impacting the entire training process, leading to training interruptions, data loss, and significant delays in model development.
- Impact on Inference: For deployed AI models, particularly in real-time applications, unreliable
  processor clusters can result in incorrect predictions, system downtime, and potentially serious
  consequences.
- **Increased Costs:** Unreliable hardware can lead to increased operational costs, including downtime, maintenance, and energy consumption due to inefficient resource utilization. Lack of reliability drives need for unnecessary redundancy.
- **Safety Concerns:** In critical applications like autonomous vehicles or medical diagnosis, unreliable Al systems can have severe safety implications.
- Complexity of Interconnects: As AI clusters scale, reliable high-speed interconnects are crucial to ensure seamless communication and data transfer between processors, minimizing bottlenecks and maintaining system stability.



### **Key Trends Driving Wafer Level Burn-in**

# Decreasing Semiconductor Reliability

- Smaller Geometries
- Larger Die Sizes
- Compound Semiconductors
- Optical Semiconductors

# Increasing Reliability Needs

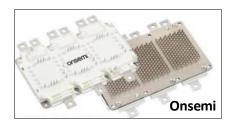
- Artificial Intelligence Processing / Processors
- Electric Vehicles
- Automotive Electronics
- Electrification Infrastructure
- Data Communication & Storage Infrastructure
- Mobile Electronics Devices

## Increasing Known Good Die

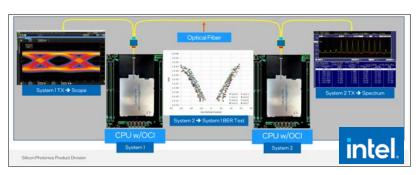
- Heterogeneous Integration (Extends Moore's Law)
- Multi-Die Modules (Power Density / Efficiency)
- Photonics Integration (Extends Data Rates)
- Stacked Die Packaging (Density and Cost)



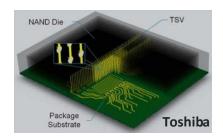
### Multi-Die Packages Drive Wafer Level Test & Burn-in



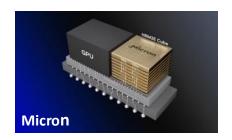
EV Traction Inverters (24-die module for Power)



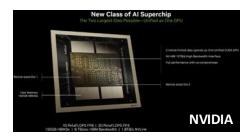
Intel Co-Packaged Optics Demo at OFC 2024



16-die NAND flash memory for Enterprise Data Storage



HBM3E 8 Die DRAM Stack for Al and Supercomputing



Blackwell AI MCM announced at GTC 2024



### **Key Trends Driving Package Part Burn-in**

# Decreasing Semiconductor Reliability

- Smaller Geometries
- Larger Die Sizes
- Compound Semiconductors
- Optical Semiconductors

## Increasing Reliability Needs

- Artificial Intelligence Processing / Processors
- Electric Vehicles
- Automotive Electronics
- Electrification Infrastructure
- Data Communication & Storage Infrastructure
- Mobile Electronics Devices

## Increasing Power and Complex Packages

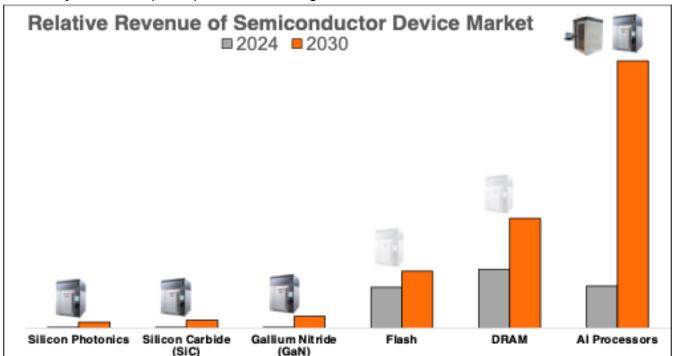
- Multi-Die COWOS Packaging
- Integration of multiple Compute Processors and Memory Stacks create higher power and multiple thermal zones (1600W Parts in 2026)
- Challenges with device handling and BIB movement creating
   Automation need



### Relative Size of Targeted Semiconductor Device Markets

These are relative annual revenues of these semiconductor devices.

Historically, annual capital spend on test ranges from 2% to 5% of the device annual revenue.



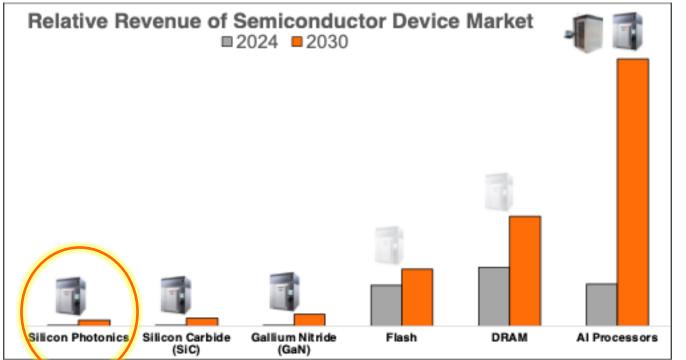




### Relative Size of Targeted Semiconductor Device Markets

These are relative annual revenues of these semiconductor devices.

Historically, annual capital spend on test ranges from 2% to 5% of the device annual revenue.





Estimated market sizes based on third party sources and Aehr estimates.



### Optical I/O is Coming...

- "TSMC silicon photonics tech first co-package optics (CPO) samples ready for NVIDIA, Broadcom in 2025" – Tweaktown, December 2024
- Ayar Labs CEO: Optical Chiplets Coming to SOCs Soon HPCwire, Oct 2024
- "Intel's 4 TB/s Integrated Optical I/O Chiplet Called 'Important Milestone'" inside HPC, June 2024
- "Optical communication is a key area that we think is very, very important for us to reach anything like zettascale type of computing capability." "AMD is working with DARPA on packaging optics solutions into chips" – Dr. Lisa Su, CEO, AMD March 2023 (HPCwire)
- "Performance metrics in applications like machine learning could ultimately pave the way for high-density integration and optical I/O to connect GPUs." – Vivek Raghuraman, Director of R&D, Broadcom, March 2023
- "In-package optical I/O technology to accelerate data movement and enable future AI" – Yole Group, April 2023





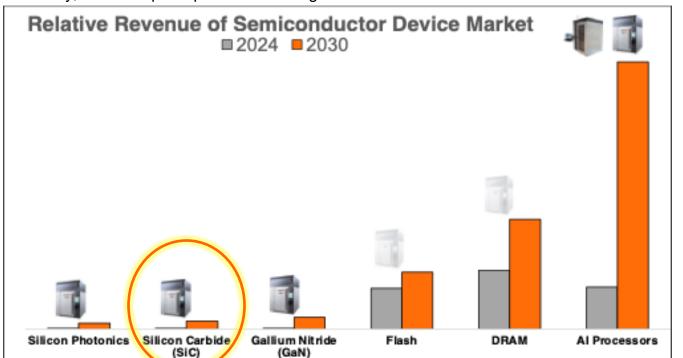




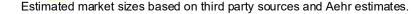
### Relative Size of Targeted Semiconductor Device Markets

These are relative annual revenues of these semiconductor devices.

Historically, annual capital spend on test ranges from 2% to 5% of the device annual revenue.

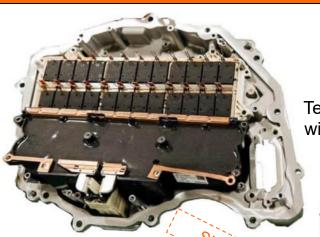








### SiC and Multi-Die Packages / Modules

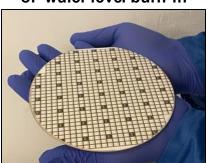


Tesla Traction Inverter with 24 2-die modules



Lucid Traction Inverter with 3 10-die modules

Package/Module yield loss cost much greater than cost of wafer level burn-in



Onsemi Traction Inverter with 1 24-die module



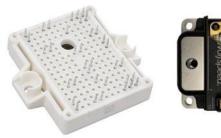


Onsemi Traction Inverter with 3 8-die modules



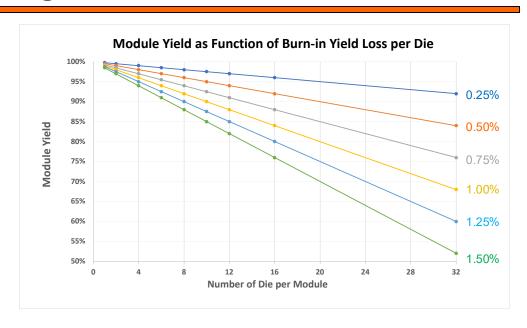
### SiC and Multi-Die Packages / Modules











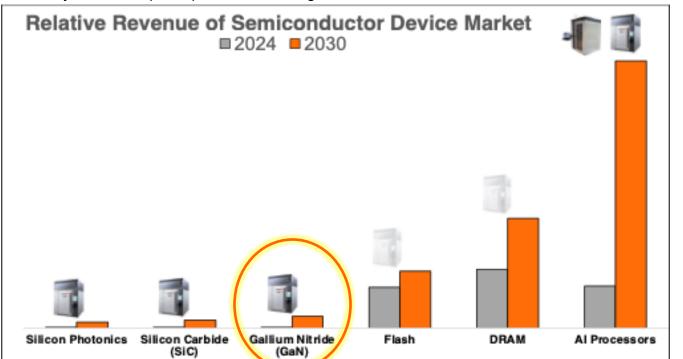
- Infant Mortality Yield loss of Modules linear with yield loss per die times number of die per module
- Cost of yield loss much greater than cost of burn-in test
- This is why the industry is driving to Wafer Level Burn-in



### Relative Size of Targeted Semiconductor Device Markets

These are relative annual revenues of these semiconductor devices.

Historically, annual capital spend on test ranges from 2% to 5% of the device annual revenue.









### **GaN Power Semiconductor WLTBI Win**

- Closed first gallium nitride (GaN) semiconductor order for production wafer level test and burn in system.
- Customer is a leading automotive semiconductor supplier and a key player in the gallium nitride power semiconductor market.
- Expands Aehr production wafer-level burn-in market for power semiconductors beyond silicon carbide applications to GaN, which offers a much broader application range than silicon carbide
- While the largest market segment for silicon carbide (about 70%) is for the electric vehicle and EV charging infrastructure markets, GaN is very diversified and is not dominated by EVs or autos with many more customers and broader market for GaN semiconductors
- Aehr FOX-XP solution for GaN semiconductors allows fully automated wafer handling and test of 6" to 12" GaN wafers using Aehr's proprietary WaferPak full wafer Contactors and integrated FOX WaferPak Aligner





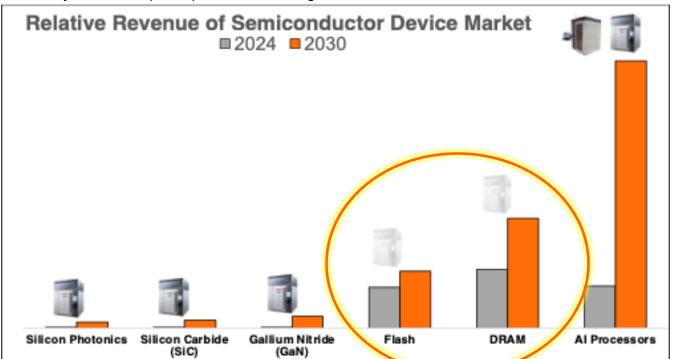
High Power / High Voltage Multi-Wafer Test & Burn-In System (Shown with Integrated WaferPak Aligner)



### Relative Size of Targeted Semiconductor Device Markets

These are relative annual revenues of these semiconductor devices.

Historically, annual capital spend on test ranges from 2% to 5% of the device annual revenue.









### Flash Memory WLTBI Opportunity

- Aehr has been engaged with multiple Flash memory companies related to our FOX wafer level test and burn in systems for their high-volume production
- Aehr has secured a commitment from one of the major Flash memory suppliers to evaluate the FOX-XP system with our proprietary WaferPak full wafer contactors for their flash devices
- This application is for 100% test and burn-in of devices to be used in mission critical applications such as enterprise storage
- We see this as a multi-year program but expect to have preliminary results and feedback during our FY25 which ends May 30, 2025
- We see the NAND Flash market as a key new market opportunity for our systems and WaferPaks with long term potential to also move into DRAM wafer level test and burn-in

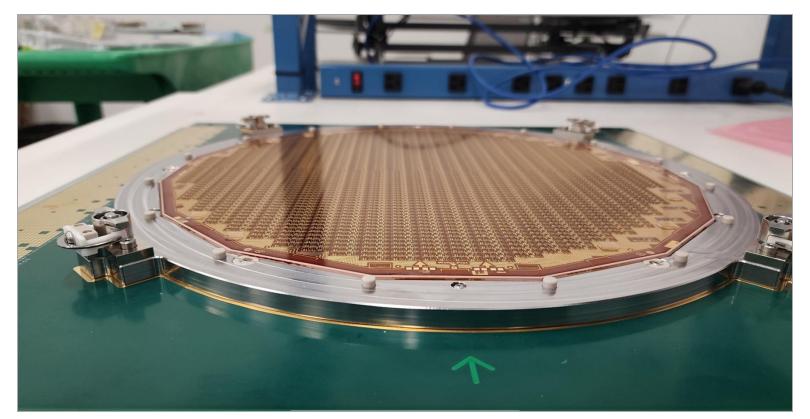




Multi-Wafer
Test & Burn-In System
(Shown with Integrated
WaferPak Aligner)

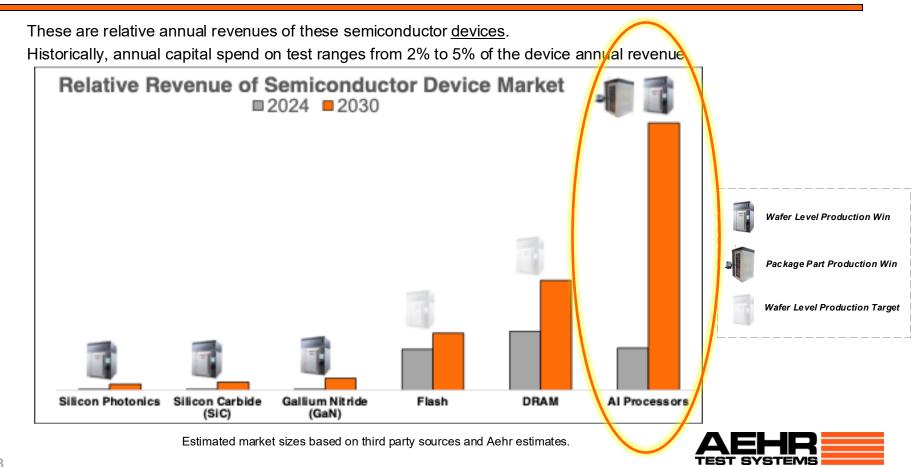


## Aehr Proprietary MEMs Fine Pitch Full Wafer Contactor



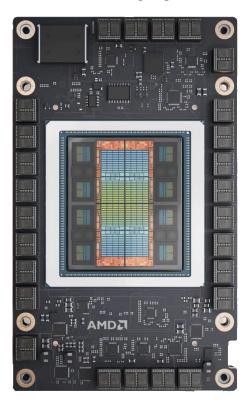


### Relative Size of Targeted Semiconductor Device Markets



### Al Accelerator Example of Multi-Chip Packaging

#### AMD Instinct™ MI325X Accelerator



- Multiple Compute Chiplets
- 8 12-Die HBM3 DRAM Stacks
- Single OAM Substrate
- Roadmap for new Chiplet Architecture Accelerators each year from 2023 to 2026
- 8 MI325X Cluster:



### **Higher Reliability Needed in AI Processing?**

Meta article from July 2024: The Llama 3 Herd Models:

Component	Category	Interruption Count	% of Interruptions
Faulty GPU	GPU	148	30.1%
GPU HBM3 Memory	$\operatorname{GPU}$	72	17.2%
Software Bug	Dependency	54	12.9%
Network Switch/Cable	Network	35	8.4%
Host Maintenance	Unplanned Maintenance	32	7.6%
GPU SRAM Memory	$\operatorname{GPU}$	19	4.5%
GPU System Processor	$\operatorname{GPU}$	17	4.1%
NIC	$\operatorname{Host}$	7	1.7%
NCCL Watchdog Timeouts	Unknown	7	1.7%
Silent Data Corruption	$\operatorname{GPU}$	6	1.4%
GPU Thermal Interface $+$ Sensor	$\operatorname{GPU}$	6	1.4%
SSD	$\operatorname{Host}$	3	0.7%
Power Supply	$\operatorname{Host}$	3	0.7%
Server Chassis	$\operatorname{Host}$	2	0.5%
IO Expansion Board	$\operatorname{Host}$	2	0.5%
Dependency	Dependency	2	0.5%
CPU	Host	2	0.5%
System Memory	Host	2	0.5%

Table 5 Root-cause categorization of unexpected interruptions during a 54-day period of Llama 3 405B pre-training. About 78% of unexpected interruptions were attributed to confirmed or suspected hardware issues.



### Al Accelerator/Processor WLTBI Win

- Aehr worked with and successfully captured the first production orders (> \$10 million in initial order) from an AI Accelerator company to move their AI processor system level test and burn-in to wafer level on the Aehr FOX-XP system
- Our proprietary WaferPaks and new high-power FOX-XP system allow delivery of precise voltages at extremely high currents up to thousands of amperes
- Our proprietary WaferPaks and new high-power FOX-XP system thermally control thousands of watts of power per wafer while also delivering signals required to determine functionality of good and bad devices.
- There are significant benefits of production test and burn-in of their Al accelerators while still in wafer form before they are integrated into the end application product, which would prove to be more cost effective and significantly more scalable than doing this screening later in their manufacturing process.





High Power Multi-Wafer Test & Burn-In System (Shown with Integrated WaferPak Aligner)



### **Testing without Compromise**

#### Reliability, Stress, and DFT Testing without compromise

- Solutions for package parts, modules, panels, or wafers allow testing at optimal process point
- Confirm which devices received desired test with per device measurements, monitoring, & feedback
- 100% traceability with die location (wafer) or device ID read back (module) and electronic tracking ensures knowledge of "good" devices
- Thermal range, uniformity, and capacity permit reduced test times & confidence in target test conditions
- Vast system resources allow for minimal sharing (higher sample size, higher yields, fewer hostage failures)
- Economical solutions and customizations allow required testing to be performed at the lowest cost



