NANOCHIP FAB SOLUTIONS
Reprinted from V7/ Issue I /2012

"BIG DATA"— TOO MUCH OF A GOOD THING?

BY
SCOTT WATSON,
JAMINI SAMANTARAY,
JOHN SCOVILLE
AND JAMES MOYNE

Spurred by the need to squeeze additional productivity out of existing assets and successfully ramp new technology nodes, most semiconductor manufacturers are experiencing exponential growth in the volume and variety of data required to support effective operation of their wafer fabs.

APPLIED MATERIALS®

“BIG DATA”— TOO MUCH OF A GOOD THING?

Today, the fault-detection production databases in most state-of-the-art fabs range from 15-30 terabytes, with a limited retention period. Manufacturers cannot collect more sensor data, retain it for longer periods or use it effectively because of issues related to cost, performance and data management. And while the downstream engineering data warehouse is often larger than the production database, it faces the same challenges.

Recent discussions with several leading foundries and integrated device manufacturers (IDMs) indicate they are all struggling with the rate of data-collection growth within their current 300mm wafer fabs. Ironically, they all indicate there is a need for still more data and at higher sampling rates. This data-explosion problem represents one of the more daunting challenges facing the industry—and it’s forecasted to get even worse as we move to sub-20nm and 450mm manufacturing.

Yet, this challenge is not unique to the semiconductor industry. Thanks to the explosion of Internet use for social media, online financial systems, online shopping, etc., other industries have confronted similar challenges and found compelling solutions. Some of their learnings can be applied to help chipmakers effectively manage their own data-growth issues.

IC MANUFACTURING GENERATES SIGNIFICANT DATA VOLUMES

Semiconductor manufacturing has always been data-intensive.

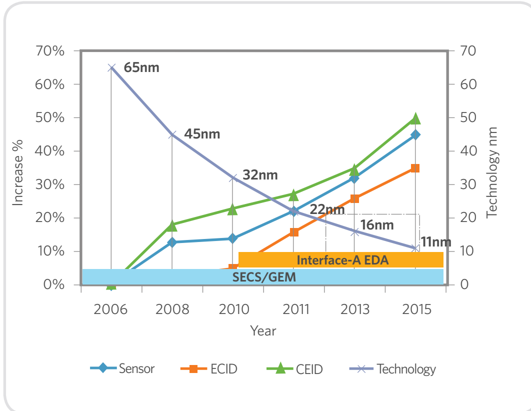

Figure 1: Technology challenges for deep-nano semiconductor. ECID = equipment constant IDs; CEID = collection event IDs (source: Applied Materials FA group).

Initially the focus was around work-in-process (WIP) tracking, metrology and electrical test data. Data collection was mostly manual and primarily used to monitor manufacturing process quality using standard, statistical, process control methodology. Advances in factory automation, tighter process tolerances, improved tool capabilities and the never-ending struggle to improve yields continue to drive these increasing data requirements.

But the ability to collect data does not mean the data is effectively used. It’s been estimated that less than half the data collected is even processed—by computers or humans. Data is simply moved into storage “just in case.” And of the data that is actually processed, more than 90% is never accessed again.

NEXT-GEN MANUFACTURING INCREASES THE CHALLENGE

As we move to sub-20nm technology nodes and 450mm wafer sizes, data volumes will continue along an exponential growth path. Most data collected from tools is gathered at rates of less than 5Hz, with 1Hz being the current norm. But the International Technology Roadmap for Semiconductors (ITRS) predicts that within three years, the requirement will reach rates of 100Hz. While most experts agree 100Hz data collection rates will be the exception, it is expected that 10Hz will become common, and that alone will drive a 10x increase in data volumes.

Additional factors will also come into play. The current roadmap for advanced tool platforms shows a 40%+ growth in the number of sensors required for these advanced technologies (figure 1). Additionally, as technology nodes shorten, the number of ECIDs and CEIDs will increase. So not only will data be captured at a 10x+ faster rate, more sensors will be contributing data. Although single wafer lots are not likely to be standard within the next 10 years, two major factors will drive down lot sizes, and with smaller lot sizes more transactional data are required to manage WIP, which compounds the data problem.

1. Queue times will shorten in sub-20nm technology nodes, requiring a reduction in the exposure time of a wafer to air. It no longer will be possible for a wafer to sit idly waiting for all of the other wafers in the lot to finish processing before moving to the next process.
2. An overall increase in process steps will require dramatic cycle-time reductions that can only be achieved through smaller lot sizes.

WHAT IS THE VALUE?

Data collection is critical to achieving the yields, cycle times and costs the industry demands. The real question is, "How do we get more value out of this data?" For example, Applied Materials regularly engages with customers to help them resolve excursion and yield issues. The answers are almost always in the data. The challenge is having the tools and expertise to turn that data into information that helps solve problems.

Factory operators strive to optimize processes to improve yields of materials and tools. That requires them to effectively use the massive amount of data that will be generated in real time and discover patterns and data trends through offline analysis of data. However, the offline analysis must be performed in a timely manner that allows optimization opportunities to be identified using known rules and heuristics, at speeds which are many times faster than today. Advances in data technologies make this possible.

Even more importantly, there is greater value in using the data to predict and resolve issues before they occur. Predictive technology can be used to analyze data to detect indicators of tool excursions before they happen, determine when tools need preventive and corrective maintenance, predict yield excursions to allow in-line resolution, predict lot arrival times for improved scheduling, and provide many more productivity improvements through other predictive means. Moving from our current reactive state of manufacturing to a proactive, predictive state is where the real value of data capture and analysis will be found.

TURNING DATA INTO INFORMATION

In the semiconductor industry and in many others, turning data into actionable information to support predictive applications is a major challenge that involves both data management and processing. But some solutions may already exist.

For example, take the story of an individual who purchases a headset in Korea with an American Express card. The individual travels extensively and uses the card on a regular basis for hotels, food and transportation but seldom purchases electronics. American Express would immediately catch this anomaly and send an email to confirm the validity of the purchase. It is exactly this type of technology that can be used against the massive volumes of data in a fab to find potential relationships or trends that could be important to rapidly identify, categorize, predict and potentially resolve anomalies with minimal human intervention.

DESCRIPTION OF APPROACH

Three main challenges arise from the explosive growth of data and the processing requirements of advanced analytics, simulation and prediction-based applications:

- How to store and process hundreds of terabytes of data cost-effectively without impacting the online transaction processing (OLTP) requirements of factory automation.
- How to fully analyze the growing quantity of data without increasing engineering staff.
- How to support extreme transaction processing (XTP)-based data processing for predictive analysis, decision tree analysis, and automated and on-demand simulations.

Traditional relational database management system (RDBMS) technologies such as Oracle, SQL Server, DB2, etc. are approaching their limits when processing large sets of data in real time to support complex data analytics in an OLTP environment. Applied Materials is working on a prototype solution to help customers use analytics on very large volumes of data through a combination of distributed databases, map-reduce-based processing, and memory-resident graph databases (figure 2). We envision this solution as running in a private cloud

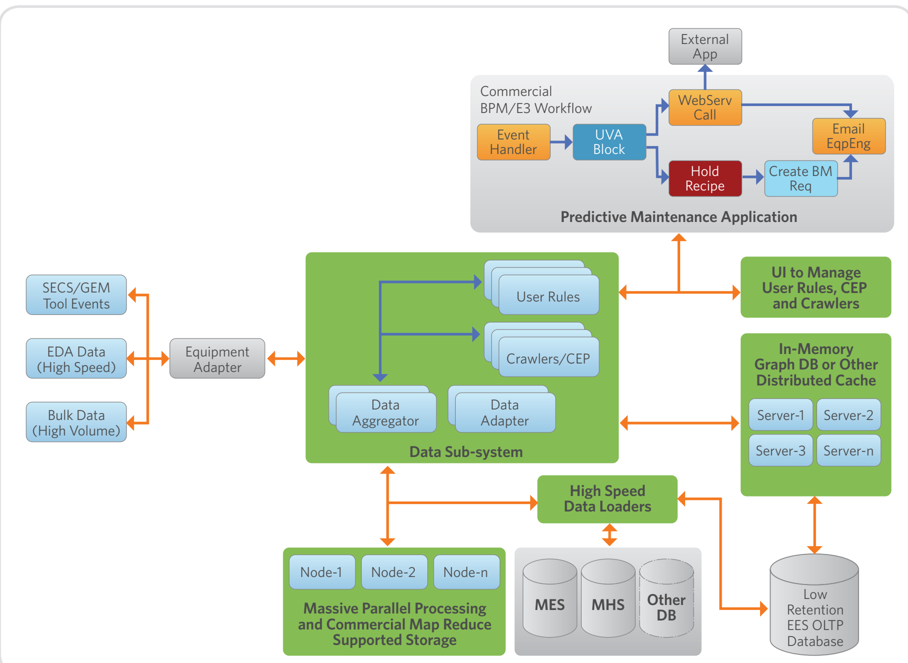

Figure 2: Conceptual architecture of big data analytics.

and supporting dynamic provisioning of system resources. Though the technology is specifically targeted to support large sets of data, older fabs will also benefit by adopting this solution.

Scheduling, dispatching and simulation-related applications may traverse multiple decision trees to determine optimal automation paths in real time. Relational databases are not able to process this type of data fast enough because of overhead in the implementation. Graph databases are designed to help process it efficiently.

Data is captured from multiple sources, including tools and MES and MHS automation software. A "data aggregator" interfaces with multiple systems and writes data in a schema suitable for shared-nothing massive parallel processing and map-reduce processing. "Crawlers" work as batch processes that apply complex analytics to the data and update the distributed memory-resident database with the relevant data. They also interface with a business process management (BPM) server to take the next step with the results. The memory-resident database is a set of preprocessed data created as graph objects based on user-defined analytics and rules (figure 3).

The prediction rules and analytics will interface primarily with this storage for XTP needs. Refer to figure 4 for an example of the automated search of "interesting" data to identify and resolve common failure modes in the factory. The map-reduce-supported data storage and graph-based, memory-resident distributed database will address the three challenges described and give a competitive advantage to next-generation Applied Materials automation products.

SUMMARY

Explosive data growth in the semiconductor industry will continue to be a major challenge, but it is a challenge full of opportunities. The ability to manage large data volumes and move the industry from a reactive to a predictive state has the potential to drive significant value for Applied Materials customers.

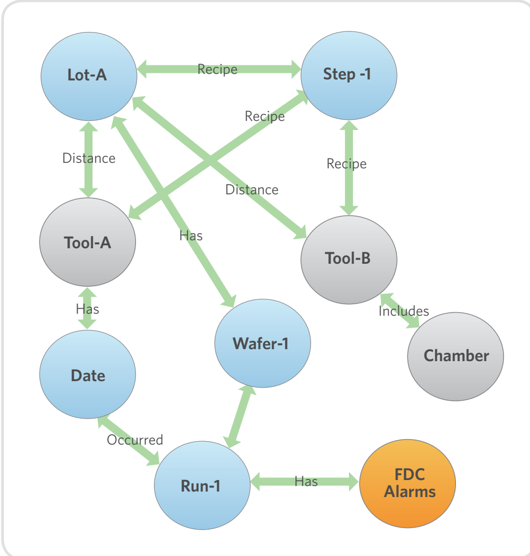

Figure 3: Graph representation of factory data for fast data retrieval and processing.

For additional information, contact scott_watson@amat.com

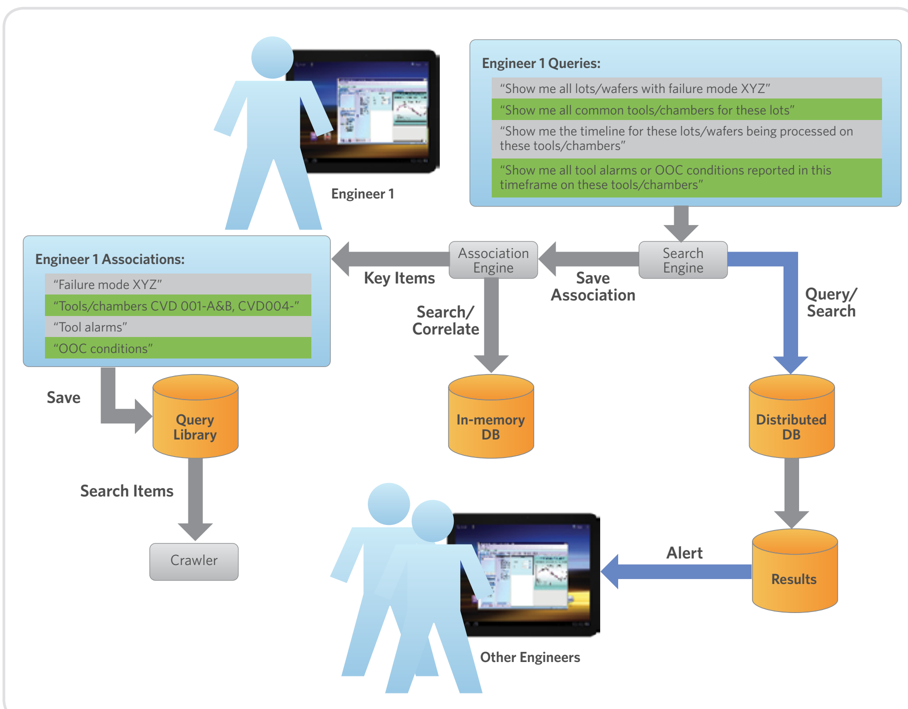

Figure 4: Example of prediction and self-learning by associating common failures.