Big Data Predictive Analytics for Proactive Semiconductor Equipment Maintenance

Sankalp Agrawal

Overview

- Manufacturing Industry generates about a third of all data today and the modern semiconductor manufacturing is one of the most contribution to this data volume.
- terabytes of data produced in just ~500 steps in semiconductor chip manufacturing.
- Analyzing different predictive analytical algorithms and data techniques to achieve near zero downtime and improve OEE.

Understanding the problem

Large amount of data

- Data volumes from manufacturing growing at exponential rate.
- Terabytes of data accumulated on a single day.
- Compilation of data from various sensors and removing unnecessary variables from the data

No PdM model applied yet

- Still not applied to complexity of data volumes.
- Semiconductor processing are usually still controlled by Statistical Process Control (SPC) and Fault Detection and Classification (FDC),
- These methods fail to predict tool failure.

equipment productivity

- Average equipment productivity stands at 44%.
- Overall equipment effectiveness ranges from 40-60%.
- Unscheduled and scheduled shutdown accounts for around 15%.

Big Data

(Collection of big data, consolidation of data, data pre-processing)

- Semiconductor manufacturing has always been data intensive.
- Due to the sheer size of manufacturing data being generated it is becoming increasingly difficult to analyze it using relational databases.
- With big data, predictive preventative maintenance can optimize maintenance planning while minimizing consequential costs due to faulty equipment.
- Traditional relational database management system (RDBMS) technologies are approaching their limits when processing massive sets of data in complex data analytics.

GROWTH OF DATA GENERATED

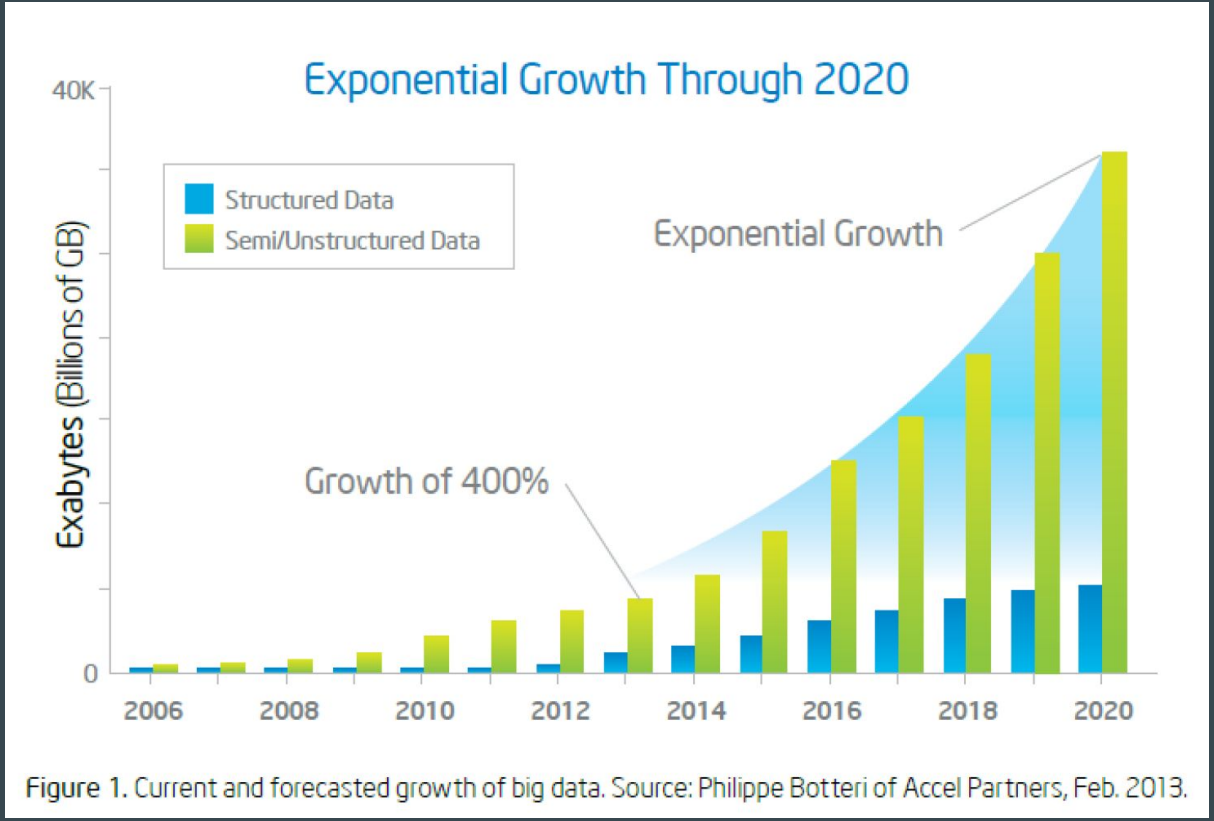


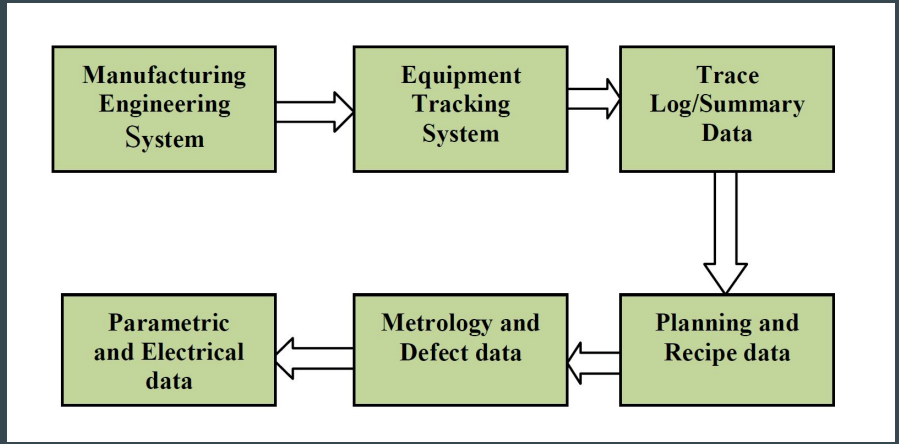
Figure 1. Current and forecasted growth of big data. Source: Philippe Botteri of Accel Partners, Feb. 2013.

Data Sources

the majority of maintenance operations in the semiconductor industry are still based on either historical reliability of fabrication equipment, or on diagnostic information from equipment performance signatures extracted from in-situ sensors.

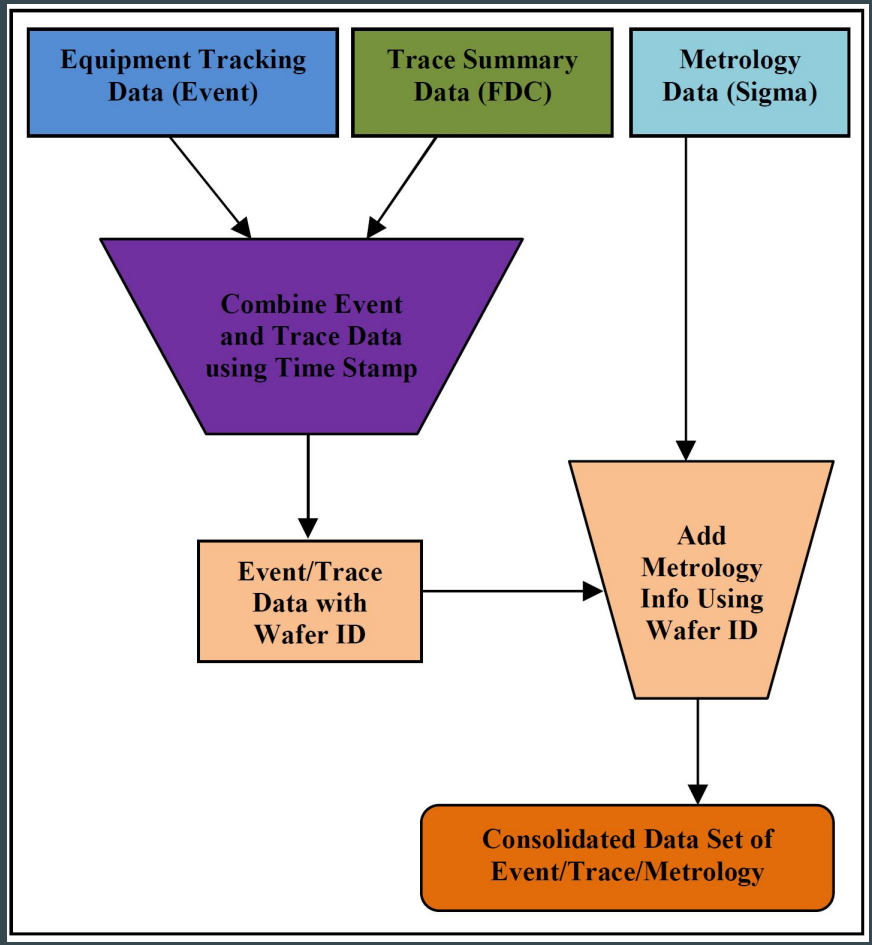
For this study, we started collecting the data from the following data sources:

- Fault Detection and Classification (FDC) data
- Trace Log data
- Equipment Tracking (ET) Data
- Metrology/Probe/Param Data



Data consolidation

Data consolidation is the key step in merging all these 4 data sets.



Big Data Architecture

Following softwares were used for processing of data:

- Apache Hadoop software
  - Using Hadoop architecture, we can gather and process data from a variety of sources, including equipment sensors, maintenance logs, trace, FD and equipment tracking.

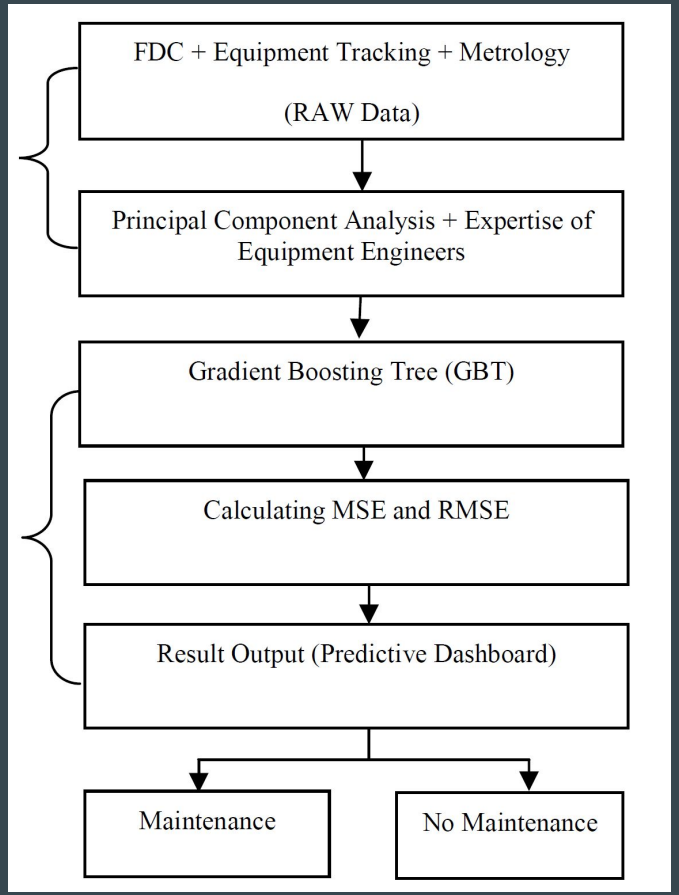
- Not only SQL (NoSQL) databases
  - NoSQL databases relax the constraints of a traditional RDBMS to deliver higher performance and scalability.

- Memory-Resident Graph Databases
  - Graph databases (GDB) are now a viable alternative to Relational Database Systems, especially in the field of predictive analytics.

Research Process:

Data consolidation and data processing

Predictive modelling and predictive dashboard



Predictive Models

The following predictive Models were used for analyzing and predicting future maintenance issue:

- ❏ Dynamic Data Clustering
- ❏ Random Forests CART (Classification and Regression Tree)
- ❏ Multivariate Analysis (MVA)
- ❏ Bayesian Networks
- ❏ Artificial Neural Networks (ANN)
- ❏ Self-Organizing Maps (SOM)
- ❏ Support Vector Machine
- ❏ Health Score Model (Binary logit Model)
- ❏ Lifespan Analysis (Cox Regression Model)
- ❏ Time Series Models