<u>Review Report</u>

on

<u>Milestone-Proposal: The Engineering Data Analysis System sNOVA, 1989-1997</u>

Date: 01/31/2025

SUMMARY:

This proposal (hereafter, "the Proposal") intends to describe the achievements and importance of the sNOVA system, an Engineering Data Analysis (EDA) system from Macronix International Co., Ltd (or MXIC, a Taiwan-based, non-volatile semiconductor memory company). Since 1989, MXIC has devoted to promoting computerization and automation in its semiconductor manufacturing. The development of sNOVA was based on IBM NOVA System, an IT software solution used primarily in semiconductor manufacturing during the 1990s, to assist in engineering data analysis, process control, and yield management. Key functions of IBM NOVA System include *Statistical Process Control (SPC)*, *Defect & Yield Analysis*, *Equipment & Process Data Integration*, *Process Optimization via DOE (Design of Experiments)*, and *Basic Predictive Modeling*. Compared to faster, smarter, and fully automated modern semiconductor EDA systems, IBM NOVA System in 1990s heavily relied on manual processing, basic statistical tools, batch-processed data analysis, and expensive on-premise systems. After more than 20 years of enhancement and development, MXIC claims that the sNOVA system is an entirely in-house developed system by MXIC.

GENERAL COMMENTS (GC):

[GC1] The contents of the Proposal are somehow far off the mark. The Proposal is entitled "Milestone-Proposal: The Engineering Data Analysis System sNOVA, **1989-1997**". However, the Proposal describes very few or nothing about the topics of Engineering Data Analysis (EDA) **during Years 1989 to 1997**. <u>(Q1) What are the difficulties, challenges, and/or opportunities in the adoption and development of an early EDA system in semiconductor manufacturing in 1990s?</u> In those years between 1989 and 1997, many technological and computational constraints have limited the development of engineering data analysis systems in semiconductor manufacturing. Examples of challenges to adopt or develop an EDA system in that era are listed as below:

<u>1. Limited Computing Power</u>

- Limited Processing Power
  - mainframes or personal computers (PCs) like 386, 486, and Pentium I
  - slow floating-point calculations
  - limited numerical simulations
- Limited Memory & Storage
  - limited-sized random access memory (RAM), often <64MBs (megabytes)
  - hard drives in the hundreds of MBs to a few GBs (gigabytes)
  - big data storage & processing infeasible
- No Hardware for Machine Learning (ML) and Artificial Intelligence (AI)
  - deep learning (DL) impractical due to the lack of parallel processing

<u>2. Software & Algorithm Constraints</u>

- Early-Staged Development in ML and AI
  - basic statistical methods (e.g., regression, DOE, SPC) for engineering analysis
  - lacking practical implementation in ML and AI due to hardware limitations
- Without High Speed Networking, Cloud Computing, and Data Mining Solutions

  - only on-premise processing with limited networked computing capabilities
  - data mining and distributed computing in their infancy
- Software Limitations
  - early statistical tools (SAS, MATLAB) with limited visualization & automation
  - no Python, TensorFlow, or modern AI libraries; only FORTRAN/C/MATLAB

3. Data Volume, Complexity, and Integration Constraints
- No Internet of Things (IoT) & Limited Sensor Networks
  - limited or primitive tool sensors in semiconductor manufacturing
  - proprietary and expensive for automated data collection implementation
- No Common Environment for Integrating Applications and Sharing Information
  - existing many and various data silos/legacies in semiconductor manufacturing
  - limited data warehouse/data mart/big data solutions
- Limited Industry Standards for Fab Integration & Automation
  - in 1989, only SECS (SEMI Equipment Communication Standard, an RS-232 based protocol) available for equipment-to-host data communications
  - in 1992: SEMI Generic Equipment Model (GEM) defined, using TCP/IP High-Speed SECS Message Services (HSMS) protocols
  - in 1998: SEMATECH Computer-Integration Manufacturing (CIM) Framework defined, specifying the interfaces and behavior of functional software components for an open, multi-supplier CIM system environment
- Limited Multivariate, Time-series Big Data Analysis Techniques
  - process/tool sensors data, context parameters, engineers' experience and knowledge, …
  - difficult to manage and analyze large volumes of complex data from multiple sources or constraints

4. Industrial Challenges in Semiconductor Manufacturing
- Limited Automation in Data Processing
  - manual inspection common for defect analysis
  - SPC charts only, no AI-based anomaly detection
- Ineffective Analysis and Simulations
  - only simplified analytical models for semiconductor process modeling
  - slow and expensive computing for Monte Carlo simulations
- No Advanced Predictive Maintenance
  - only simple heuristics for equipment failure prediction, without AI-driven models
- No Virtual Metrology & Online Monitoring Methodology
  - largely reactive rather than predictive in process control
  - relying on historical data analysis rather than real-time adjustments

5. Needs for New Paradigm Shifts for Problem-Solving and Decision-Making
- From Experience-based to Data-driven
  - people not optimizing choices with data
  - engineering data and analysis not an integral part in engineers' practices
- Lack of Cross-Disciplinary Expertise
  - lack of multi-disciplined profession overlapping various areas of engineering
  - Hard to integrate and collaborate with a team of multidisciplinary expertise

[GC2] Following [GC1], <u>(Q2) what are the critical/challenging issues and problems that the MXIC sNOVA team faced during the development of sNOVA in Years 1989 to 1997?</u> Please consider to add some words/sentences/paragraphs about the answers to this question in the "What obstacles (technical, political, geographic) needed to be overcome" part.

[GC3] Following [GC2], (Q3) how were these issues/problems tackled or resolved? Please consider to add some words/sentences/paragraphs about the answers to this question in the "What obstacles (technical, political, geographic) needed to be overcome", "Historical Significance of the Work", and "Why was the achievement successful and impactful" parts.

[GC4] During the development of sNOVA in Years 1989 to 1997, (Q4) were any lessons-learned, new designs or findings ever shared to the semiconductor industry, such as in SEMI standards or in the SEMATECH CIM framework, or in SEMICON conferences ? Please consider to add some words/sentences/paragraphs about the answers to this question in the "Historical Significance of the Work" and "Why was the achievement successful and impactful" parts.

[GC5] In addition to the key functions provided by IBM NOVA System, such as *Statistical Process Control (SPC)*, *Defect & Yield Analysis*, *Equipment & Process Data Integration*, *Process Optimization via DOE (Design of Experiments)*, and *Basic Predictive Modeling*, and so on, (Q5) any innovative ideas or novel implementations were made by the MXIC team during Years 1989 to 1997? Please consider to add some words/sentences/paragraphs about the answers to this question both in the "abstract" part that describing the significant of the technical achievement and in the "What features set this work apart fro similar achievements?" part.

[GC6] In the "Historical Significance of the Work" part of the Proposal, the claims of "*sNOVA leads global semiconductor manufacturing into the era of AI and Big Data*", "*This groundbreaking innovation set the foundation for Taiwan's prominent role in global chip manufacturing*" and "*sNOVA is Taiwan's pioneer of AI and Industry 4.0 in the semiconductor industry*" were made. Other than the quote from Dr. C.C. Wei's "very innovative [1]", (Q6) are there any rigorous supports or convincing evidences for such claims? Please consider to add some words/sentences/paragraphs about the answers to this question in the "Historical Significance of the Work" part.

[GC7] Following [GC6], (Q7) why and how did the development of sNOVA during Years 1989 to 1997 lead global semiconductor manufacturing into the era of AI and Big Data? Please consider to add some words/sentences/paragraphs about the answers to this question in the "Historical Significance of the Work" part.

[GC8] In the "Historical Significance of the Work" part of the Proposal, the descriptions in the paragraph of "*The market share of Taiwan in global semiconductor foundry capacity increase year by year*" are neither direct supports nor significantly related to (from Fig. 1 & Fig. 2??) the historical significance of the development of sNOVA. Please consider to add some words/sentences/paragraphs to explain in the "Historical Significance of the Work" part. (Q8) Why does the development of sNOVA in 1990s lead to the increasing Taiwan's market share in global semiconductor foundry capacity? Why "foundry"? Why only "foundry"? Please consider to add some words/sentences/paragraphs about the answers to this question in the "Historical Significance of the Work" part.

[GC9] The three paragraphs of "Creating Value" of the "What features set this work apart from similar achievements?" part are not easy to understand. (Q9) Which to create values? And what values were created? Did the development of sNOVA in Years 1989 to 1997 make sNOVA itself an AI system, a value created during the development of sNOVA? Or "The system must be adjusted in a timely manner … or AI will not be able to continue to create value.", in the last paragraph in the Creating Value" part. What is "the relationship between the sNOVA system and product quality"? Why was it "identified as a critical issue in the production process from the outset"? Please consider to add some words/sentences/paragraphs about the answers to this question in the "Creating Value" of the "What features set this work apart from similar achievements?" part.

[GC10] In the first paragraph of "sNOVA leads global semiconductor manufacturing into the era of AI and Big Data", the Proposal describes the genesis of sNOVA in 1989 that "a Statistical Process Control (SPC) engineer and head of the in-house SPC system project, led the team to develop the prototype of the engineering data analysis system sNOVA". While in its following paragraph, the Proposal describes "As a pioneering system in Taiwan, sNOVA not only ushered in the era of AI and Big Data in global semiconductor manufacturing but also established Macronix as the world's first paperless fab from its inception." <u>(Q10) How were the evolutions in the development of sNOVA from SPC to AI and Big Data? Why and how are SPC vs. AI vs. Big Data related, especially in the development of the EDA system, sNOVA?</u> Please explain.

[GC11] Some of the descriptions/wordings/claims of the Proposal should be more rigorous, self-evident, and with strong supports/validation/proofs. For example, in the "The system has effectively improved the product defect rate from PPM to PPB" part, the Proposal said "Macronix has also become <u>the world's first memory company to measure product defect rates in parts per billion (PPB)</u> rather than parts per million (PPM) (Fig. 1)." The problem is that it is NOT obvious to justify the claim of "the world's first memory company …" Also, as device geometries shrank into sub-micron level and the need for higher purity materials became critical, the semiconductor industry, including companies like Toshiba, began emphasizing contamination control measured in parts per billion (ppb) in the 1980s and 1990s.

[GC12] Following [GC11], the other examples are the use of the terms "<u>paperless</u>", "<u>establishing the first paperless fab in the world</u>", "<u>fully computerized</u>", "… transformed <u>the automation of the entire semiconductor industry chain</u>" in the Proposal, which is NOT good as a semiconductor-professional proposal in IEEE. In a contemporary standard 300mm fab, full automation of wafer processing, data collection, and material handling are a must for efficient and effective fab operations. However, in Years of 1989 to 1997, a paperless fab (200mm or less) is a dream but impossible, due to the challenges listed in [GC1] and more issues from the unreadiness of processing and metrology tools, lacking of collaboration software, and so on.

[GC13] Following [GC11], the other improper statements are such as "The sNOVA concept rooted in efforts to rival Japanese companies which had a dominant share in the global semiconductor industry. They were accustomed to rigorous work ethics and high-quality standards, but <u>relatively less focused on using scientific methods</u> to analyze engineering problems encountered." appearing in the first paragraph of the "What features set this work apart from similar achievements" part. Such statements are neither rigorous nor good at all. For example, the famous robust design methods, Taguchi methods, are welcomed by worldwide statisticians and economists. Taguchi methods are widely used in fab EDA applications with the optimized design of experiments (DoE) to eliminate variation of the final product quality.

TYPO/GRAMMATICAL ISSUES (TG):

Line 2, Paragraph 2, the "Allowing statistics, IT and engineering talent to cooperate effective
Macronix gradually found common ground ~~between~~ the parties
=> Macronix gradually found common ground <u>among</u> the parties

Line 4, Paragraph 2, the “Allowing statistics, IT and engineering talent to cooperate effective
…, and production data from the first three to five years ~~was~~ effective …
=> …, and production data from the first three to five years <u>were</u> effective …

—— End of the Report ——