



Chetan Arvind Patil  
Contributing Writer | EPDT

# How Silicon Test Data Became a Material Cost Driver

For much of the semiconductor industry’s history, test data has been treated as an inevitable by-product of silicon validation. It is generated to confirm functionality, archived for traceability, then occasionally revisited when yield issues arise. The economic model of testing has reflected this assumption. Costs are dominated by tester depreciation, handler utilisation and test time. Data itself has rarely been viewed as a factor that could materially influence the cost structure of a silicon product.

That assumption no longer holds. As silicon products have grown more complex, the volume, dimensionality and persistence of test data have increased far faster than traditional cost models anticipated. Modern devices generate parametric data across thousands of pins, multiple operating corners and repeated test insertions spanning wafer sort, package test, system-level test, plus ongoing reliability screening. Each additional layer of coverage does not merely add incremental data, it multiplies it.

There are several structural shifts that explain why test data has crossed the threshold from background artifact to cost driver. To start with product architectures have changed. Advanced nodes, heterogeneous integration and domain-specific accelerators demand deeper characterisation and tighter guard-band control. Next, test strategies have evolved. Multi-site testing, adaptive test flows and retest loops increase data density - even when unit volumes remain constant. Furthermore, data retention expectations have expanded. Regulatory requirements, customer audits and long-lifecycle markets now require test data to be stored, searchable and defensible years after production.

Consequently, test data now incurs costs at every stage of its lifecycle. Capturing it requires bandwidth and tester overhead. Preparing it for analysis demands engineering time and custom tooling. Storing it consumes on-premises or cloud infrastructure. Analysing it requires computational resources and specialised

expertise. Even deciding what not to analyse introduces cost, as poor visibility often leads to over-collection by default.

## Breaking down the cost structure of test data analytics

The cost of silicon test data analytics is often underestimated, because it does not originate from a single source. Instead, it accumulates across multiple layers of the test/manufacturing ecosystem. Unlike tester capital expenditure or test time, which are well understood and explicitly budgeted, analytics-related costs are fragmented, making them difficult to quantify and even harder to optimise.

The 1st and often largest contributor is data ingestion and preparation. Raw test data is rarely analytics-ready when it leaves an automated test system. Differences in formats, naming conventions, sampling rates and context across test insertions require significant engineering effort to normalise and cleanse the data. This preparatory work scales with product complexity and test depth. In many organisations, it consumes more resources than the analytical models applied afterward.

The 2nd cost layer resides in storage and infrastructure. High-volume parametric data, waveform captures and historical test records must be retained for traceability, audits and long-term analysis. Whether stored on-premises or in the cloud, this data incurs

Dimension	Traditional Silicon Testing	Modern / Advanced Silicon Testing
<b>Data Volume</b>	Limited parametric data, primarily pass/fail focused.	High-dimensional parametric data across pins, corners, stages.
<b>Data Preparation Effort</b>	Minimal normalisation, largely manual review.	Extensive cleaning, normalisation, and contextual tagging.
<b>Test Insertions</b>	Wafer sort and final test.	Wafer sort, package test, system-level test, reliability screens.
<b>Analytics Compute</b>	Batch analysis, periodic yield review.	Continuous, near-real-time analysis with correlation across stages.
<b>Data Retention Horizon</b>	Short-term storage for yield learning.	Long-term retention for traceability, audits and field analysis.
<b>Decision Latency Impact</b>	Delays primarily affect engineering insight.	Delays directly affect cost-of-test, yield loss and time-to-market.
<b>Organisational Ownership</b>	Primarily test engineering.	Distributed across test, IT, analytics and quality teams.
<b>Economic Treatment</b>	Considered overhead.	Becomes a material cost driver.

Table 1: Silicon test activities and their associated costs

ongoing costs for capacity, redundancy, access control and retrieval. As data volumes grow, storage decisions made for convenience rather than strategy can become recurring financial burdens.

A 3rd component is computational and analytical execution. Advanced yield analysis, correlation across test stages, plus increasingly AI-driven techniques require scalable computational resources. These workloads are not continuous, but are often bursty and unpredictable - complicating capacity planning and cost control. In parallel, specialised talent is required to build, maintain and interpret analytics pipelines, introducing a labour cost dimension that grows with data complexity rather than production volume.

What distinguishes test data analytics from traditional test costs is that inefficiencies compound across these layers. Poor test data preparation increases computational time. Excessive data retention inflates storage needs. Slow analytics delays decisions, indirectly driving retest, over-guard-banding, or yield loss. Together, these effects transform analytics from a supporting function into a structurally embedded cost element within the silicon test flow, one that demands the same level of economic scrutiny historically applied to testers, handlers, test time, etc.

## Advanced silicon multiplies data-related costs

Advanced silicon products, characterised by smaller process geometries (40nm and below), heterogeneous integration and application-specific designs - such as GPUs, AI accelerators and automotive SoCs - fundamentally change the cost dynamics of test data. The depth and persistence of test data amplify analytics costs far beyond linear scaling.

Advanced nodes operate with tighter process windows, narrower voltage margins and stronger sensitivity to variation. To manage this risk, test programs capture more parameters

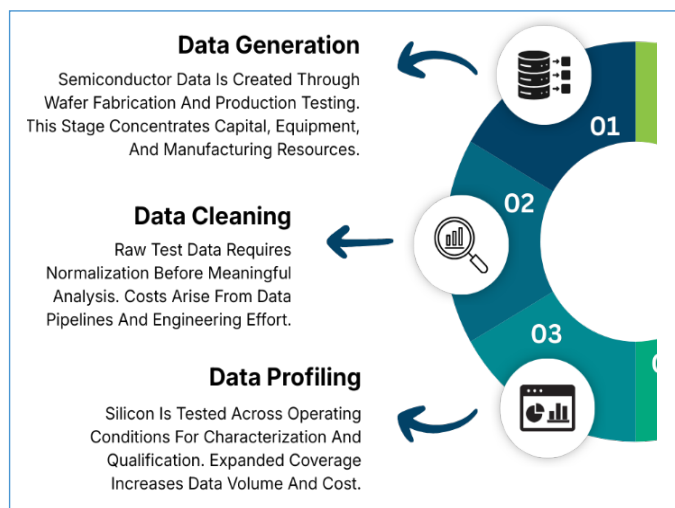


Figure 1: Silicon test data intelligence lifecycle

across more conditions at finer resolution. Each additional data dimension increases raw data volume and the effort required to contextualise, store and analyse it. Heterogeneous integration accelerates this effect. Chiplets, 2.5D/3D packages and multi-die systems (high-bandwidth memory stacks and suchlike), require test data that is correlated across dies, interfaces and assembly stages. Failures must be traced over manufacturing, packaging and test domains, multiplying the retained data's complexity and the analytics burden.

Market requirements further amplify the problem. Automotive-grade chips, data centre processors and AI accelerators face heightened expectations for traceability and reliability. These requirements extend data retention horizons and increase reanalysis frequency, causing analytics costs to persist long after volume production stabilises. Advanced silicon transforms test data from a transient artifact into a long-lived asset with ongoing economic weight. Each layer of complexity, whether architectural, packaging-related, or market-driven, adds multiplicative pressure to data pipelines and analytics systems.

## Need to reframe silicon test data analytics as a cost-engineered system

Silicon test data analytics' economic impact is increasingly defined not by how much data is collected, but by how efficiently insight is extracted and acted upon. Analytics latency introduces costs that are rarely captured in traditional test models, but that directly influence yield stability, binning decisions, retest volumes and time-to-revenue. As data volumes grow, delays in interpretation amplify downstream costs, making them difficult to recover through subsequent optimisation.

Inefficiencies across analytics pipelines further compound this effect. Manual data preparation, fragmented data ownership and inconsistent contextualisation consume engineering capacity that would otherwise be directed toward yield improvement and test optimisation. These inefficiencies scale nonlinearly with product complexity, turning analytics from a supporting capability into a structural driver of test cost inflation, particularly for advanced-node and high-reliability silicon programs. Reframing silicon test data analytics as a cost-engineered system requires applying the same economic discipline historically reserved for testers, handlers and test time. The objective is not analytics sophistication for its own sake, but economically meaningful speed - the ability to convert data into decisions before costs accumulate.

As silicon complexity continues to rise, organisations that engineer analytics systems with cost, latency and efficiency as design constraints will be better positioned to control one of the fastest-growing contributors to overall silicon product economics. The question is no longer whether test data analytics carries cost, but whether that cost is managed with the same rigour applied to every other element of the test flow.