OCP Global Summit
November 8, 2021 | San Jose, CA
Reliability of cloud scale hardware through performance and health monitoring using on-chip telemetry

Alex Burlak
VP Test & Analytics, proteanTecs
proteanTecs Leading a New Category

Deep data health & performance monitoring for advanced electronics

Founded in 2017 by industry leaders and co-founders of Mellanox

Addressing industry-wide challenges of scale

Silicon proven in 28/16/7/5nm

Customers in multiple key segments including Datacenter, Automotive, and Communications

New category with a multi-disciplinary approach

Global Footprint

- California
- New Jersey
- Germany
- Israel
- India
- Taiwan

HQs, Sales & support, R&D
Challenges in High Performance Resilience

- HW failures and service disruption
- No in-mission monitoring solution
- Expensive redundancies
- Constant replacements

Inconclusive Investigations

- Undetected defects

Advanced Electronics

Facebook:
“Silent data corruption due to silicon latent defects and aging. (1,000 DPPM)”

Google:
“Rare, short-term computational errors on systems that passed all manufacturing tests successfully”

Microsoft:
“Frequent unexplained HW failures with “No Issue Found” at high rates”

CAR A I D

“Current FuSa guidelines do not cover HW/system reliability during the operational lifetime”

3. “Circuit Reliability Mitigation Techniques & EDA Requirements”; Georgios Konstadinidis, Google, EDPS 2019
4. “Improving Cloud Scale Hardware Fault Diagnostics”; Neeraj Ladhani, Rama Bhimanadhuni, Microsoft, Mar. 2020 OCP Virtual Summit
5. “Automotive semiconductor reliability and its changed importance for complex system engineering, Andreas Aal, Volkswagen, Dr. Oliver Aubel, Globalfoundries, IRPS 2021”
Visibility at Every Stage

**Chip Design**
- Variability Analysis
- Production Visibility

**Chip Production**
- Higher Confidence Volume Ramp
- Outlier Detection
- Performance Yield

**System Production**
- Performance Optimization
- Workload Analysis
- Environmental Monitoring

**In-Field**
- Performance Monitoring
- Alert on Faults before Failures
- Predictive Maintenance

**Faster Time to Market**
**10X Lower DPPM**
**Reduced Costs**
**Failure Prevention**
Multi-Pillar Solution

Deep Data
Universal Chip Telemetry™ (UCT) with on-chip Agents

Machine Learning
Agent fusion and inference with ML algorithms

Cloud & Edge Analytics
Advanced analytics for actionable insights

Automated insertion tools
Universal Chip Telemetry™ (UCT)

On-chip Agents

- Parametric measurements
- High coverage & high resolution
- Minimal PPA penalty
- Operate in mission-mode
- Sense the surrounding electronics
- Application optimization to HW

Operational Monitoring
Interconnect Performance Monitoring
Performance and Degradation Monitoring
Classification and Profiling
Proteus
Targeted Analytics

Chip Production

- Fast Ramp-Up with Higher Confidence
- DPPM Reduction
- Shift-Left Decisions

System Production

- Correlation Between Value Chain Stages
- Power Optimization
- System & Application Optimization

In-Field

- Predictive Maintenance
- Continuous Performance Monitoring
- RMA Reduction with Fast Time-to-Resolution
Estimator Based Outlier Detection

- Personalized chip assessment
- Multi dimensional outlier detection
- Yield reclamation
- Fast RMA with pinpoint Root Cause Analysis

**BKM: Population based**

- Measured IDDQ only
- PAT Limit Spec Limit
- One dimension only identifies outliers outside of PAT spec limits
  (circled in red: outliers unidentified by BKM)

**proteanTecs: Dynamic limit per chip**

- Measured IDDQ vs. Estimated IDDQ
- Detected Outliers within PAT spec
- Returned as RMA
- Undetected Outliers within PAT spec
- Dynamic limits per chip
- Additional dimension to identify outliers within PAT spec limits
  (circled in red: outliers identified by proteanTecs)
System Performance Monitoring

Bring up & production EOL software optimization and performance tuning

- Agent data recorded throughout multiple phases of system operation
  - Configure
  - Operation Start 1
  - Operation End 1
  - Idle 1
  - Idle 2
  - Operation Start 2
  - Operation End 2
  - Idle 3
  - Idle 4

- Correlation and merging of different Agent measurements
  - Margins of millions of paths
  - IR drop
  - Workload/stress: V*T*Toggle rate
  - Cycle to cycle clock jitter
Health Monitoring with Alerts on Faults Before Failure

• UCT-based continuous performance and health monitoring
  • In mission-mode
  • High coverage critical path monitoring based on millions of internal paths and low margin alert
  • Correlated to system environment and application induced stress
• Workload management for achieving longer product lifetime
• High reliability through personalized Predictive Maintenance

*Based on Physics-of-Failure and Time-to-Failure
High Demand in Multiple Markets

Serving customers in key industry segments

- Datacenters
- Communications
- Automotive
- AI
- Storage

In use and silicon proven

- 28nm
- 16nm
- 7nm
- 5nm

Trusted by industry leaders

- Hyperscalers
- Fortune 100 corporations
- OEMs
- Disruptive startups
- ASIC and IP vendors