Unified Socket Interconnect in the Test Ecosystem of High Pin Count and Large Body Size Packages

MP Divakar, Ashok Kabadi & Prathiksha Dhanpal

Ampere Computing

Mesa, Arizona • March 2–5, 2025

Contents

- Background
- Test Cost Reduction Coverage at TestConX
- Post-Silicon Test Flow for Processor Packages
- Problem Statement
- Socket Standardization Coverage at TestConX
- CPU Sockets -Applications and Requirements
- Challenges for Unified Socket Interconnect
- Ampere Approach to Standardizing Socket Interconnects
- Ampere's Solution to Unified Socket Interconnect
- Implementation Example
- Conclusions

Background-1

- Our focus here is on the test ecosystem of larger body size LGA/BGA processor packages with high pin count
- Ampere has been offering high compute core count processors with I/O count of 4926, 5964, 7228, etc.
- Package body size has increased from 67.5 mm in the first generation to 87 mm in the second; future generations may exceed 100 mm package dimension
- Disaggregation of processor functions has led to multichip topologies that are harder to control package warpage which impacts connectivity with the socket contacts

Background-2

- Disaggregated packages with exposed Silicon pose challenges in designing socket loading mechanisms which vary in each area of test (Validation, ATE*, SLT**, HTOL*** etc.)
- High pin count results in many 100's kgf (kilogram force) which is required to be managed ensuring the integrity of the exposed die
- Ampere's yearly cadence of releasing new products requires shorter test development schedules and minimal-to-no respin of test system designs
- Our motivation to standardize the socket type is a multi-objective optimization effort –meeting schedules, cost, repeatability, test coverage, serviceability, durability, reuse across multiple generations, etc.

*ATE: Automated Test Engineering, **SLT: System Level Test, ***HTOL: High Temperature Operating Life

Test Cost Reduction Coverage at TestConX

- TestConX (formerly BiTS) is perhaps the only industry forum where the test costs are actively debated
- In previous conferences, the coverage on cost reduction has omitted addressing the fragmented socket type in the test ecosystem
- There is opportunity for cost reduction by standardizing one type of socket

From BiTS 2016 Keynote
"Chip Overtest: Are ICs Tested too Much?

BiTS 2016 — Keynote – Dale Ohmart — Chip Overtest Are ICs Tested too Much?

Approaches to Test Cost Reduction

BiTS 2016 — Keynote – Dale Ohmart — Chip Overtest Are ICs Tested too Much?

Evaluating the "Magic Bullets"

Unified Socket Interconnect in the Test Ecosystem of High Pin Count and Large Body Size Packages

5 2025

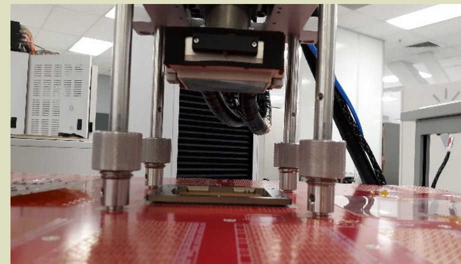
Post-Silicon Test Flow for CPU Packages

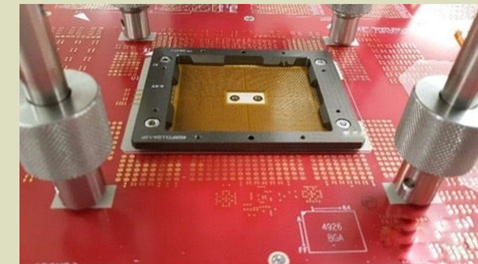
Bring up / Validation Phase

- Low volume
- ATE/FT1
- Validation and Characterization
- Final Test (FT)
- Bench top System-Level-Test (SLT)
- Burn-in

Production Phase

- High volume
- Final Test (FTx)
- System-Level-Test (SLT)





Problem Statement-1

- Disparate types of sockets
  - In the first generation of our products, for each of the foregoing test areas the socket interconnects were spring pin-based for ATE, FT and Burn-in whereas validation systems used elastomer sockets
- SI characteristics of each socket type is different
  - Electrical length of the interconnect were different in each type
  - Elastomer had a shorter height
  - Difficult to correlate data
  - Anomalies in one platform could not be validated / replicated in another

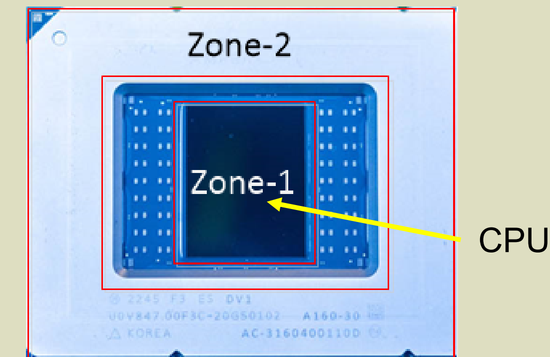
Opportunity:

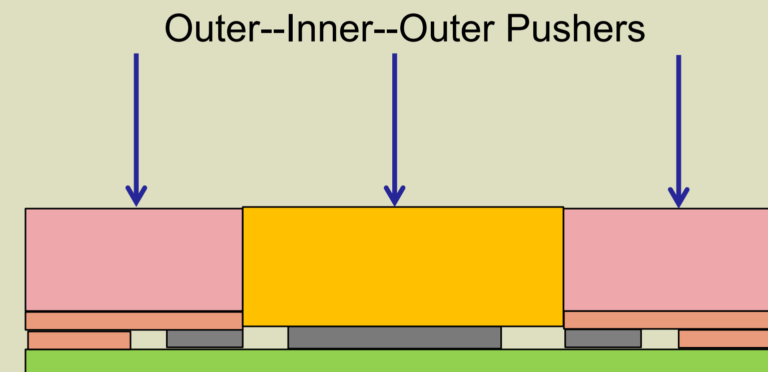
- Reduce the number of socket types – standardize to one socket type for Validation, ATE, FT, SLT and Burn-in
- Reduce post-silicon validation/test cost significantly
- Correlate SI data
- Minimize/eliminate test systems’ design change

TestConX 2025

Problem Statement-2

- Mechanical challenges
  - Matching package warpage for reliable connectivity and lowest contact resistance
  - Zonal loading for exposed die packages – independent control of each zone to assure die integrity
  - Adaptability for in-line socket cleaning
  - Thermal management
- Cost challenges
  - For package w/ high pin-count (>10,000), the cost/socket are too high (>$12,000/unit)
  - For lower utility test functions like validation and burn-in, certain socket types are becoming cost-prohibitive





Socket Standardization Coverage at TestConX

- There has been discussion on socket standardization at previous TestConX conferences
- “One size fits all” is debatable, is it “one type fits all?”

BiTS 2018 — Session 5 Presentation 2 — Semiconductor Device Manufacturer

Need Standards

- What are the major influences on socket life?
  - Temperature cycling, power cycling, current carrying capability, how long the current is applied for, force distribution, electrical performance specially high speed data rates, package size, pitch, number of pins, routing of the signal, power and ground pins.
- How should life cycle be defined across the industry?
  - What is critical?
    - Socket Electrical, Mechanical, Thermal Performance, Lifecycle and Cost are critical vectors
    - Socket specs has to be defined for that socket rather than the technology itself. It's all the other interaction that actually determines the performance of the socket in the specific application
  - What is controversial?
    - Disconnect between the socketing technology and the whole socket performance itself
  - What is unnecessary?
    - One size fits all for the technology will not work

Burn-in & Test Strategies Workshop www.bitsworkshop.org March 4-7, 2018

From BiTS 2018
“Life Cycles of Sockets: Specification vs. Reality and Setting Standards”

TestConX 2019 — Session 5 Presentation 4 — Panel Discussion

Mission

Alignment of expectation between User and suppliers

Socket manufacturer:
Focused on mechanical life of contact
- Interface Wear
- Force
- Cres variation
- Cleaning intervals
- Room temp
- Temp ramp
- CCC
- RF/High speed data

Socket user:
Focused on production yield
- First pass yield
- METS test
- Temp cycling
- Power cycling
- Disconnect between contact and whole socket performance
- Contamination
- Contact Performance Maintenance System
- Use of actual devices vs. surrogates

TestConX Workshop www.testconx.org March 3-6, 2019

From TestConX 2019 Panel Discussion Update
“Life Cycles of Sockets”

CPU Sockets – Applications and Requirements

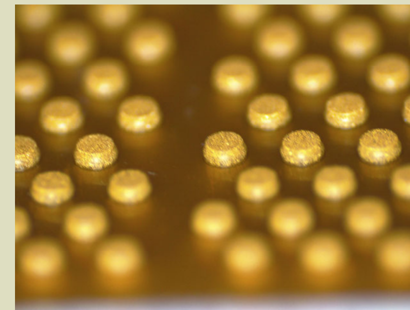
Socket Type	Application	Main Requirements
Validation/Characterization	• Used on Validation/debug board for initial bring-up and validation	• Easy to remove and install • Cycles – several hundreds • Meet SI requirements (low profile) • Low cost -- <$600/unit
Final Test (FT)	• Run test patterns in one or more stages prior to SLT	• Use on ATE handler • Cycles – Tens of thousands • Low cost -- <$600/unit
System-Level-Test (SLT)	• Run tests to assess performance and reliability of the CPU emulating its end application environment	• Use on ATE handler • Cycles – Tens of thousands • Low cost -- <$600/unit
Burn-in	• Screen for infant mortality and early life failure rates (ELFR)	• Pass high temp test – 125 C
Production	• End application in the servers	• Extremely low cost -- $100/unit

Challenges for a Unified Socket Interconnect (Validation, FT, SLT and Burn-in)

- Meet high cycle life (>20,000 insertions/removal) and still meet cost target of $x00/socket
- Minimum maintenance
- Warpage in large body size and high pin count packages
  - Packages with exposed die
- Accommodate ATE/SLT handler requirement
- Meet SI requirements – Low height (<0.9 mm) w/ reliable interconnection
- Meet high temperature requirements (125 C)
- Meet current carrying capability of more than 4 amps/contact
- Meet environmental testing requirement

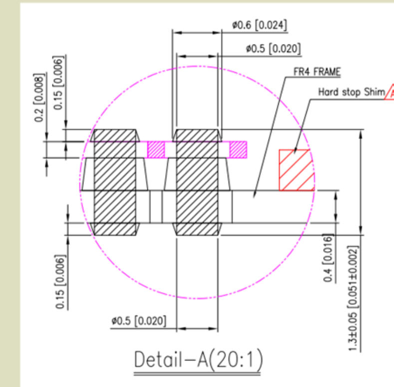
Ampere Approach to Standardizing Socket Interconnects

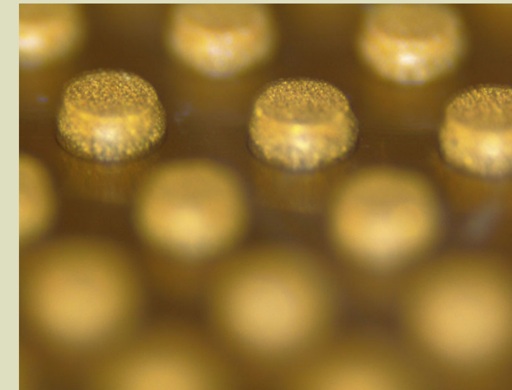
- We evaluated both spring pin and elastomer types of sockets for our first-generation CPU products (Ampere Altra / Altra Max)
- We investigated contact resistance and signal integrity characteristics as well as the repeatability of the interconnection
- Based on our investigation favors elastomeric sockets in the context of large body size CPU processors
- We developed internal requirements for elastomer sockets



Ampere’s Solution to Unified Socket Interconnect

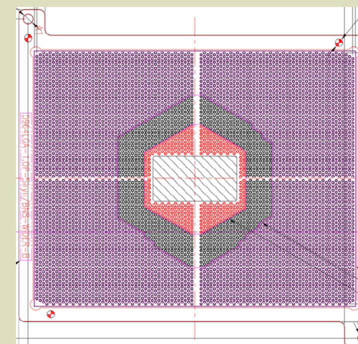
1. Use of low-profile elastomer socket contacts –option to use higher profile to accommodate larger bottom side passives on the CPU package
2. Higher working range of 0.4 mm helps package warpage issues
3. Innovative dual-loading mechanism for packages with open dies (collaborative effort with the socket vendor)

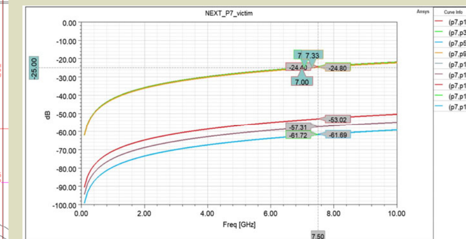




Ampere’s Solution to Unified Socket Interconnect

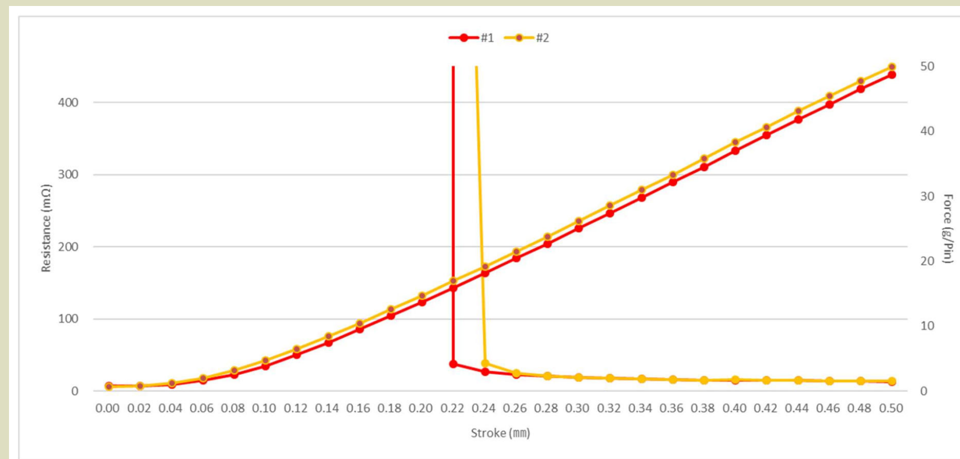
4. Profiling of socket contact array to accommodate package warpage
5. Extensive environmental, high temperature and cycle testing (initial contact resistance for each pin documented)
6. Signal Integrity simulations and actual measurements

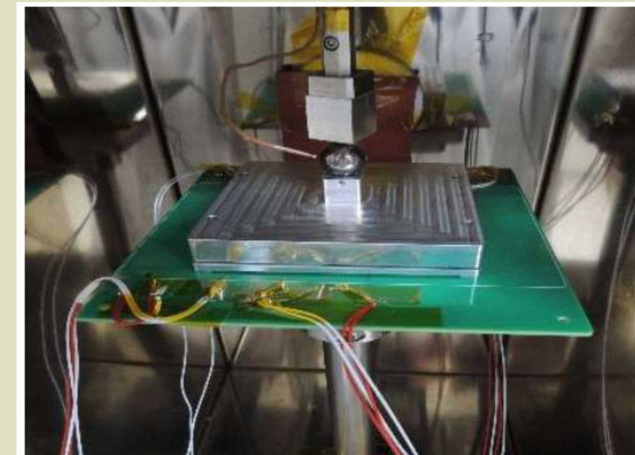




1 & 2. Low Profile Elastomer Contacts with Higher Range

- Depending on the package body size and pad pitch, we have worked with 0.3 mm to 0.4 mm total stroke in the elastomer contacts
- Force-stroke-resistance characterized at room and elevated temperature for 50K+ cycles

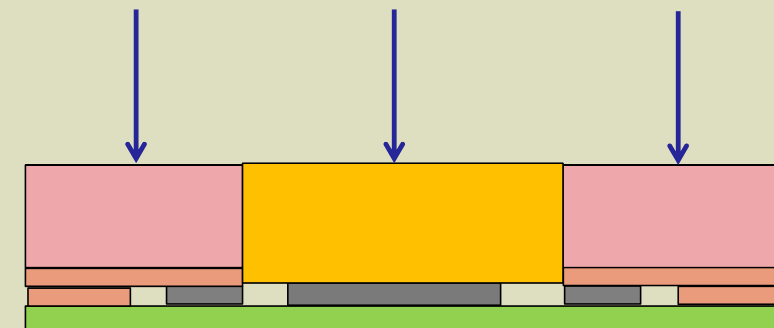




3. Innovative Dual Loading Mechanism

- For exposed die packages, zoned loading mechanism that preserves the integrity of the Silicon while ensuring adequate contact with the socket contacts is critical
- At the same time, meeting thermal requirements using fan-based cooling using a heatsink lid is challenging
- Ergonomics of load application for ~7300 contacts each with ~35gm requires efficient lid design with mechanical advantage
- Stress simulations using FEA tools validated the design approach of the loading mechanism

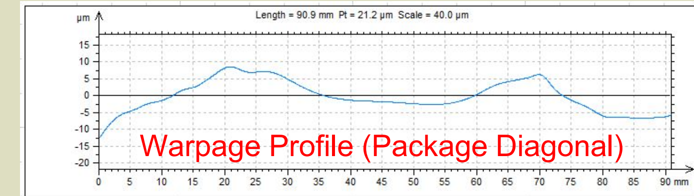
Outer--Inner--Outer Pushers

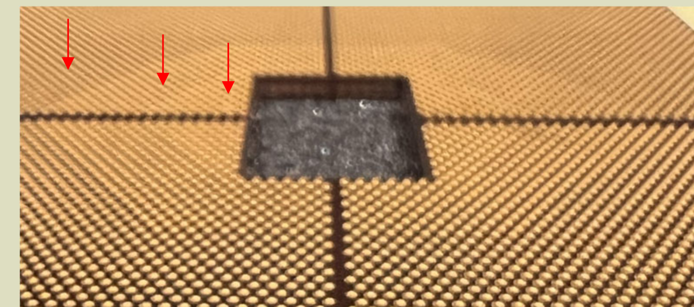


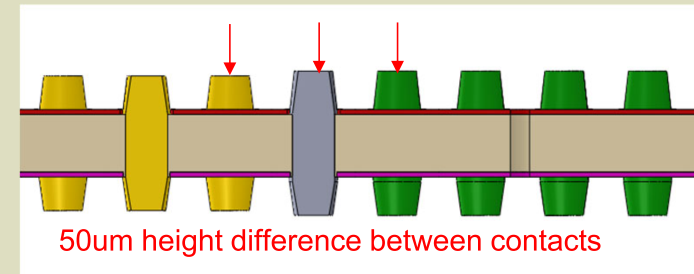
Partially Lidded CPU

4. Profiling Socket Contact Arrays

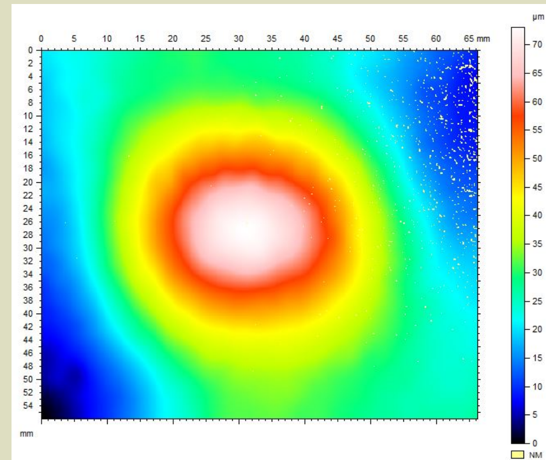
- Contact heights varied to match package warpage
- Contact resistance compared with baseline (non-profiled) elastomer in validation platforms using functional CPUs
- Single molding tool used to fabricate the whole socket with different contact height zones





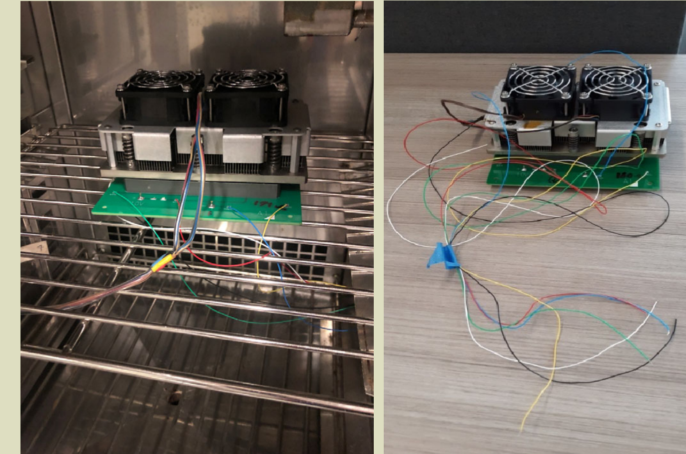


Warpage Profile (Top View)



5. Environmental, High Temperature and Cycle Tests

- Our elastomer socket qualification test plan covers:
  - Force-Stroke-Resistance measurement
  - Mechanical cycles (100x) and contact resistance
  - Repeated lid actuation (1000x)
  - Thermal gap pad durability (from above)
  - Room temperature resistance stability
  - Maximum current carrying capability/contact (CCC)
  - High temperature resistance stability (125 C, 100 hr)
  - Thermal Cycling (thermal shock & thermal cycling)
  - Temperature humidity testing (85C/85%, 120 hr)
- Contact resistances are continuously monitored



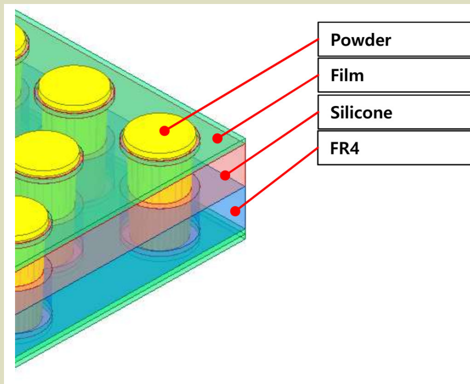
Zone F-A	Zone A-B	Zone B-C	Zone C-D	Zone D-E	Zone E-F
208.47	81.20	84.88	74.21	75.86	75.16
197.07	78.56	81.71	71.86	73.48	72.98
194.00	77.72	80.77	71.13	72.75	72.06
192.26	77.51	80.48	71.00	72.62	71.78
190.39	76.80	79.70	70.35	71.97	71.02
188.46	75.79	78.63	69.39	71.01	70.00

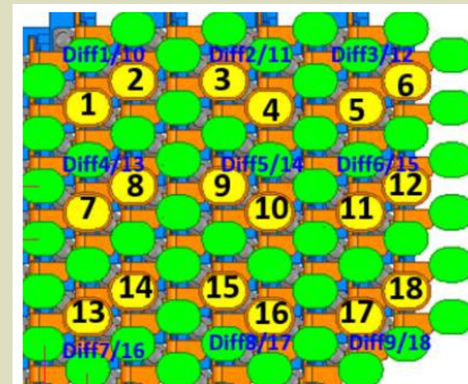
5. Sustained High Temperature Tests for Burn-in (HTOL)

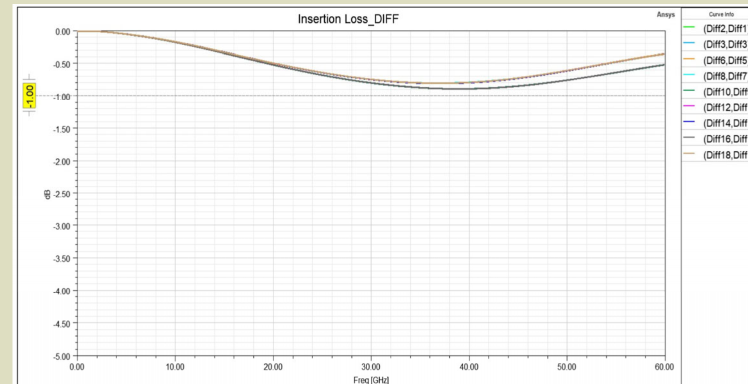
- Test requirements for HTOL:
  - The test set up mimics Ampere's end applications
  - Chamber temperature at 135 C for 1200 hr
  - Up to 4 Amps / contact
  - Each elastomer had a minimum of two zones with max current applied and many zones without applied current
- Contact resistances
  - Both zones were measured for contact resistance every 12 hours

6. Signal Integrity Simulations and Measurements

- Typical elastomer contacts have 25 mohm initial contact resistance that do not degrade significantly over temperature and use cycles
- Max current per pin 5 A
- Insertion loss of -3 dB 69 GHz provides excellent margin; -10 dB return loss @ 44 GHz
- Ampere team has developed internal characterization elastomer fixture that enables testing critical nets of its processors

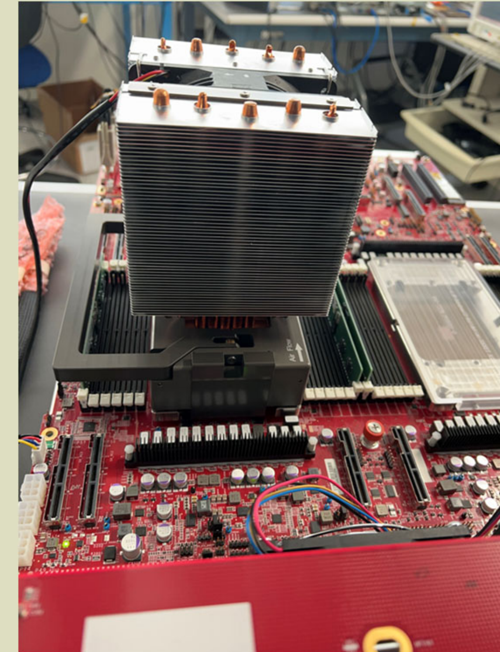






Implementation Examples

- The image shows the socket nest on the right and the heatsink lid assembled on the socket nest
- The lid incorporates zoned loading mechanism in which a lower magnitude of the load is applied on the exposed die ensuring its integrity
- The lid uses cam mechanism to efficiently apply more than 200 kgf with mechanical advantage meeting the ergonomics of the test operator
- The heatsink incorporated in the lid is capable of meeting >450 W of thermal dissipation



Conclusions

- The types of sockets used in the semiconductor test ecosystem remain disparate and fragmented leading to increase in cost and development times
- There are difficulties in correlating measurements when different types of sockets are used, particularly in the characterization phase
- We investigated standardizing the socket interconnect type across all phases of semiconductor test for processor packages
- Our investigation addressed mechanical, electrical, serviceability, durability and cost considerations in all area of test –validation, ATE, FT, SLT and HTOL
- For Ampere CPUs, we found elastomer socket interconnects to be reliable and cost-effective solutions to address all areas of semiconductor test
- For other applications, spring pin-based contacts may be a better choice

Presentation / Copyright Notice

The presentations in this publication comprise the pre-workshop Proceedings of the 2025 TestConX workshop. They reflect the authors' opinions and are reproduced here as they are planned to be presented at the 2025 TestConX workshop. Updates from this version of the papers may occur in the version that is actually presented at the TestConX workshop. The inclusion of the papers in this publication does not constitute an endorsement by TestConX or the sponsors.

There is NO copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies: as such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author/s or their companies.

The TestConX logo and 'TestConX' are trademarks of TestConX.