Memory Known Good Die

Lee Nevill
NAND Flash QRA Manager
Micron Technology, Inc.

Memory KGD – Agenda

- Quality
  - Wafer-level testing
  - Costs
- Reliability
  - Wafer-level screening
  - Design for KGD
  - Costs
- Other Challenges
  - The die formerly known as good
  - KGD infrastructure
- Summary

Memory KGD – Quality

- What is quality?
  - Meeting expectations at time 0
- KGD quality is determined by the amount of testing performed at the wafer level
  - Packaged memory will balance wafer-level testing and package-level testing to achieve the lowest possible cost
    - Yield loss vs. assembly cost analysis
  - KGD requires 100% of the testing to be done at wafer level, regardless of cost
- Good enough die (GED) allows for some flexibility
  - Application-specific testing vs. full data sheet

Demanding more quality than the application requires only adds costs.

Memory KGD – Quality

- How do we measure KGD quality?
  - At the vendor?
    - Packaged part monitors
      - Full data sheet
      - Excludes assembly-induced failures
  - At the customer?
    - Test yield after MCP/SIP assembly
      - Application-specific
      - Includes MCP/SiP assembly-induced failures
  - Which is correct?
    - Both, depending on the situation

Quality is relative to how it is measured.

Memory KGD – The Cost of Quality

- Testing memory in wafer form is less efficient than testing it in packaged form
  - Round wafer/square probe card
  - Dead die
  - Equipment reuse
  - Parallelism

All else being equal, it is always more expensive to test memory at the wafer level.

Testing the Limits

FOR PRODUCT SUCCESS

Round wafer – Rectangular probe card

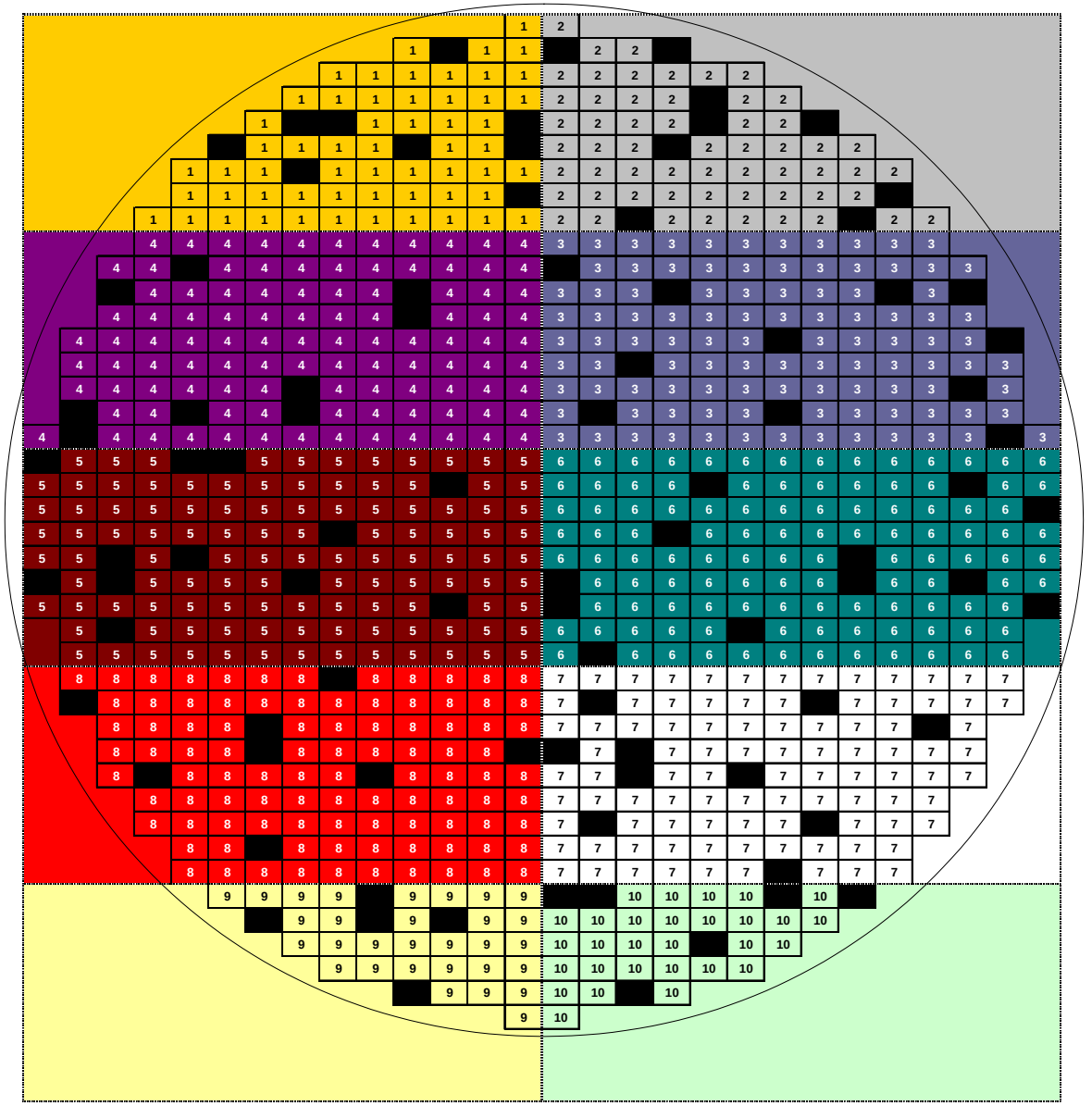
- 300mm DRAM example with 884 dpw
- 126-site probe card
- 10 touchdowns to cover entire wafer
- 884/1,260 = >30% wasted tester sites

Dead die

- Failing die cannot be removed
- Each subsequent probe step becomes less efficient
- Overall yield determines efficiency
  - 10%-15% wasted at maturity
  - >50% early in products life

Limited retest capability

- Individual sites can not be retested
- Retest adds potential for bond pad damage



Memory KGD – The Cost of Quality

- 10 touchdowns @ 126 sites = 1,260 possible
- 884 die per wafer
- Mature yield example
  - 15% dead die from first probe
    - (884 x 0.85)/1,260 = 59.63% efficient
    - Requires **1.67** times the capital
- New product yield example
  - 50% dead die from first probe
    - (884 x 0.5)/1,260 = 35.08% efficient
    - Requires **2.85** times the capital

Wafer-level testing can require up to 3X the capacity compared to packaged part testing. Even at maturity, KGD requires >50% more capacity.

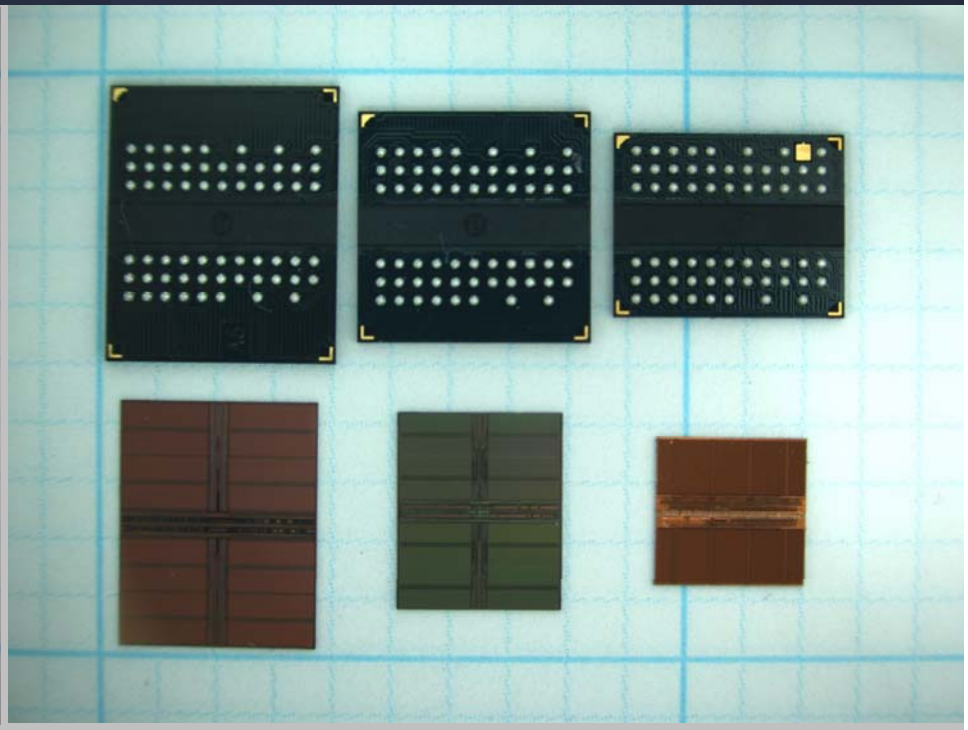
Memory KGD – The Cost of Quality

- New probe cards for each new design
  - High-parallelism, speed-capable probe cards are expensive
    - Any die size/pad layout change makes them obsolete
    - Down sites result in yield loss or entire card out of production
  - High-parallelism, speed-capable interface boards for packaged part handlers are also very expensive, but…
    - They can be reused for many generations and multiple die sizes
    - Down sites are simply disabled

Probe cards are not reusable from one generation to the next, leading to higher costs.

Testing the Limits

FOR PRODUCT SUCCESS



3 Generations of 2Gb NAND Die

- Each die requires new probe cards for KGD
- Same package for all generations
- Same handler kit and interface for all
- Same tester interface can be used for stacked packages

3 Generations of 512Mb DRAM Die

- Each die requires new probe cards for KGD
- Different package outline, but the same ball grid
- Each requires new handler kit (relatively cheap)
- Same tester interface is reused (very expensive)

Memory KGD – The Cost of Quality

- Massive parallel testing of wafers is limited
  - The number of die per wafer sets the maximum number of DUTs
  - Bond pad pitch limits signal density
  - Massively parallel probe cards are very expensive
  - Still have dead die to deal with and down site issues
- Packaged parts are tested with many thousands of units simultaneously
  - >6,000 parts tested simultaneously
  - Dead die are not an issue
  - Equipment is reusable over many generations

The ability to perform massively parallel testing is limited at wafer level and leads to higher costs.

Memory KGD – The Cost of Quality

- Wafer sales hamper the ability to sell multiple bins
  - Speed grading
    - Customer only wants one speed
    - Wafer contains parts of various speeds
  - Parametric grading
  - Density grading
- Wafer sales hamper component recovery
  - Most die on memory wafers are eventually sold
    - Partial die recovered for low-tech applications

Selling a wafer eliminates the ability to generate revenue from down bins and partials.

Memory KGD – Reliability

- What is reliability?
  - Continuing to meet expectations after time 0
- How is KGD reliability measured?
  - At the vendor?
    - Initial reliability qualification studies
      - HTOL/cycling on packaged parts
    - Ongoing monitors on packaged parts
  - At the customer?
    - Final product burn-in (packaged parts)
      - Avoided at all costs
    - Field returns…

Reliability assessment requires accelerated life testing, which is difficult to do prior to assembly.

Memory KGD – Reliability

- KGD reliability is determined by several factors
  - Wafer-level screening
    - Burn-in
    - DC stress modes
  - Design for reliability
    - 2T/2C
    - On-chip ECC
    - Electrical repair
  - Process maturity
    - Lower defects

There are many ways to address the reliability requirements for KGD.

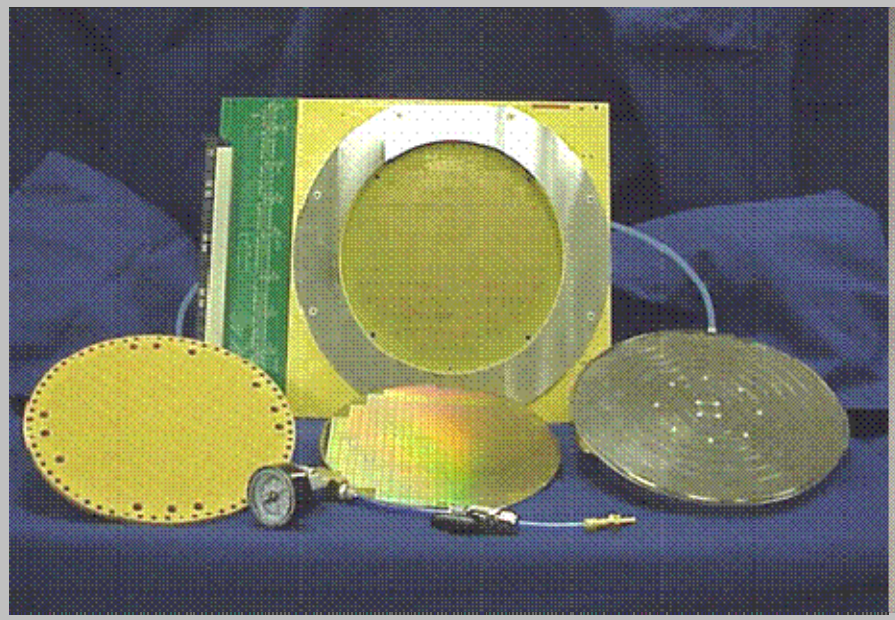
Memory KGD – Reliability

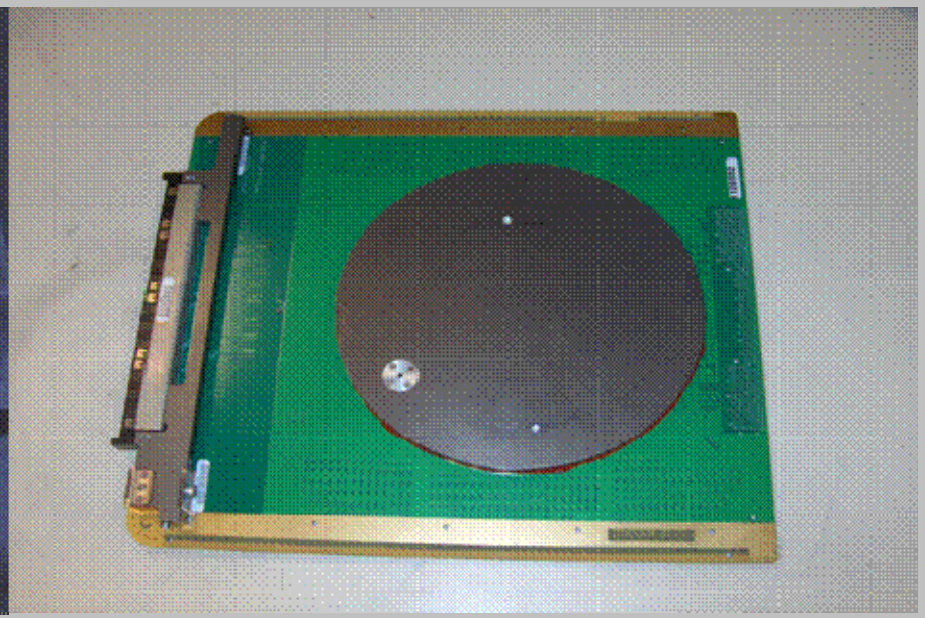
- Wafer-level reliability screening of memory
  - Burn-in
    - Classical burn-in is possible at wafer level, but many challenges remain for high-volume implementation
      - Parallelism limited by die-per-wafer
      - Electrical isolation of dead die
      - New contacts (full wafer probe cards) for every revision
  - Wafer-level DC stress
    - Common among DRAM suppliers
      - DC stress is an effective means of accelerating many years of life
      - Often referred to as wafer-level BI
      - When coupled with other methods, has proven very successful

True wafer-level burn-in has the same limitations as wafer-level testing.

Memory KGD – Reliability

200mm full wafer contact fixture





Memory KGD – Reliability

- Design for reliability
  - 2T/2C proven VERY effective on PSRAMs
    - Differential sensing will sense the “0” correctly even if the “1” has leaked down
    - Array size penalty is large (2X the number of cells)
  - ECC would be even better than 2T/2C
    - All single-bit errors are corrected
    - Array size better than 2T/2C (1.5X for 8/12 Hamming code)
    - Performance penalty can be large
  - Electrical repair
    - Ability to invoke spare elements AFTER assembly
  - Process maturity
    - Yields and reliability are always better on mature processes

Reliability by design is a very effective technique.

Memory KGD – The Cost of Reliability

- Wafer-level BI
  - Same cost issues as wafer-level test
    - DPW limits parallelism
    - Dead die
    - Fixture reuse
- Wafer-level DC stress
  - Cost-effective solution
    - In-situ with wafer test
    - Limited die size impact
    - Proven effective on DRAM-based technologies

True wafer-level burn-in can be done, but wafer-level stress is much more cost-effective and meets the needs of most applications.

Memory KGD – The Cost of Reliability

- Design for reliability
  - Must balance die size impact with benefit
    - 2T/2C works well for small die (pad limited)
    - ECC provides tremendous benefit, but adds die size and hurts performance
    - Electrical repair for DRAM fixes post assembly related issues
- Process maturity
  - Best yields and reliability vs. largest die
- Must have realistic expectations
  - 100 FITs in a server with 1,024 DRAMS → MTBF = **1.1 yrs**
  - 100 FITs in a cell phone MCP with 1 DRAM → MTBF = **856 yrs**

Requesting more reliability than the application requires only adds costs.

Memory KGD – Other Challenges

- It was a known good die when we shipped it; then you…
  - Applied tape to the face of the wafer
  - Flipped it over and ground it to 75µm thick in a chemical slurry
  - Taped the backside to a carrier
  - Ripped off the face tape
  - Cut it out of the wafer
  - Exposed it to UV
  - Ripped it off the carrier tape while pushing it up with needles
  - Glued it down to a material with a different CTE
  - Exposed it to a cloud of high-energy plasma
  - Thermosonically welded dozens of gold wires to it
  - Glued or taped another hunk of silicon on top of it
  - Repeated above process for up to 7 other die
  - Encapsulated it with plastic and silica with differing CTEs at >2,000 PSI
  - Baked it at 185 deg C for 4 hours to make sure it's done
  - Attach solder balls at 240 deg C or coated the leads with solder
  - Cut it out of the matrix or stamped it out of the leadframe
- And you still expect it to be a known good die?
- Instead, it becomes "the die formerly known as good!"

Memory KGD – Other Challenges

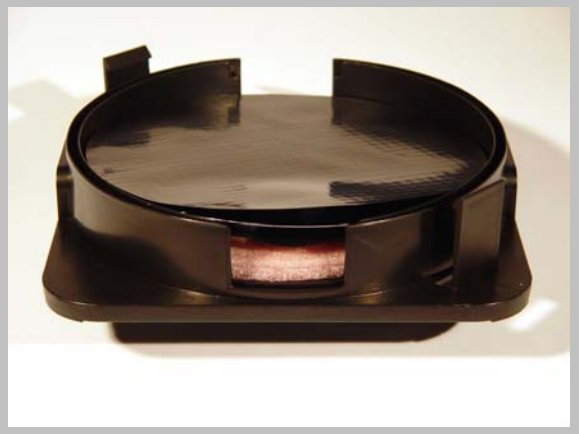
- The assembly process changes things
  - Mechanical stress on the circuits
    - Die thickness reduced from 775µm to <75µm
    - Different materials above/below die
    - CTE mismatch leads to additional mechanical stress
    - Backside surface: polished vs. rough
    - Particle impingement from filler material
  - Exposure to chemicals, heat and pressure
    - Chemical mechanical polish for backside prep
    - UV can alter Flash memory reference cells
    - High-energy plasma can induce electrical charge

No matter how much testing is done at wafer level, the final product needs to be carefully qualified to make sure the assembly process did not change the electrical characteristics of the die.

Memory KGD – Other Challenges

- KGD infrastructure – Shipping materials

Coin Stack



Vendor Box

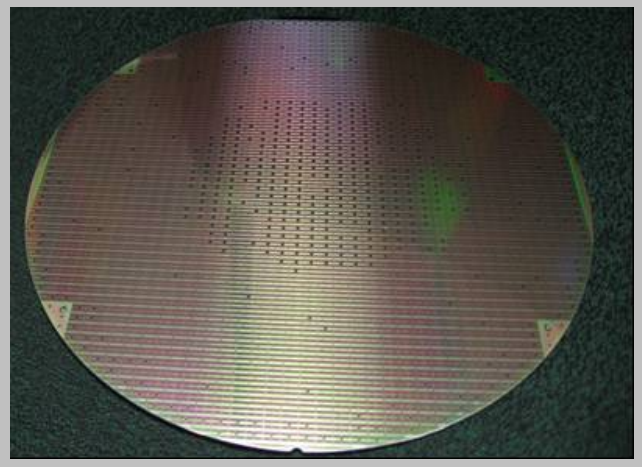
Film Frame Box

- Order entry
- Recycling programs with the ODM

KGD includes seemingly simple things like how the customer wants the product delivered.

Memory KGD – Other Challenges

- KGD infrastructure - Wafer maps
  - Many different forms of electronic maps
    - SEMI standards help, but some customers still require ink!
    - Some customers require their own map format
  - Several different data transfer methods require ongoing support
    - FTP
    - E-mail
    - CD with shipment



Wafer-level testing and reliability screening are only the start of the KGD process.

Memory KGD – Other Challenges

- KGD infrastructure – Assembly methods
  - Wire bonding
    - Copper metallization requires special processing
      - Bond pads are plated or capped
      - In-house assembly techniques need to be fanned out to ODMs
    - Different bond pad construction requires different processing
      - Plating requires different pre-wire bond plasma than aluminum cap
  - Die thinning
    - Thinned wafers are difficult to ship
    - Different tool sets are used for die thinning between ODMs
    - Customers need to understand that thinning can/will affect electrical performance
      - Must evaluate performance AFTER assembly

KGD requires support from functional groups that would normally be considered captive.

Memory KGD – Other Challenges

- KGD infrastructure – Test software support
  - Most KGD customers require test program support
    - Many different test platforms
      - DRAM tester vs. Flash tester vs. logic tester vs. analog tester
    - IP issues with proprietary DFT
    - IP issues with supplied test software
    - Accessibility to die after assembly

KGD customers require test program support.

Memory KGD – Other Challenges

- KGD infrastructure – Failure analysis
  - Material returned in various formats
    - wafer, film frame, half assembled, package, partially deprocessed
  - Electrical failure analysis
    - Must have the ability to correlate back to wafer level
    - Was it bad to start with or did it change during assembly
    - Custom packages require customer fixtures
      - Must work with the ODMs up front to get this capability
      - Can I even “talk” to the die directly
  - Physical failure analysis
    - How do I isolate which die is failing

The ability to perform failure analysis on the customers end product is often limited.

Memory KGD – Summary

- Quality
  - Quality is relative to how it is being measured
  - More expensive to test KGD than packaged parts
  - Strive for quality to meet the application (good enough die)
- Reliability
  - Cost-effective solutions require a combination of techniques
    - Screening + design + realistic expectations
- Other challenges
  - Assembling a known good die can make it a “formerly good die”
    - Need to carefully qualify all new products/processes
  - Being a KGD supplier requires a lot more than testing and screening wafers