Test Data Management in a Complex Environment

June 8, 2005

Glenn Claudi-Magnussen
Southwest Test Workshop 2005
The Objective

• All wafer test data -- and needed associated data -- is available to everyone who needs it, in the tools or format desired, at the time the user wants it.
A Complex Environment

• Internal and external manufacturing sites (sources of data)
  – differing conventions, practices, etc.

• Internal and external users (consumers of data)
  – examples of external consumption: probe feedback to foundries, probe data to external assembly for inkless assembly

• Various tester types, running various operating systems

• Numerous test data formats

• Production and engineering data (often co-mingled)

• Thousands of files per day
A Complex Environment
Elements of Test Data Management

• Data Acquisition
• Data Transport
• Data Storage and Archiving
• Data Sorting
• Data Preprocessing
Data Acquisition

• Data formats
  – Different equipment often have different standard data formats
    • STDF, ATDF, CSV, SPD, etc.
  – Data format from custom test equipment is often at the mercy of the Test Engineer
  – A number of suppliers support STDF, but there is variation from implementation to implementation

Consistency and completeness of data is determined when it is captured.
Data Acquisition

• Data content
  – Equipment configuration can affect data consistency
    • regional date formats, data capture settings, etc.
  – Automated capture significantly helps with data consistency
    • OCR, barcodes, electronic interfaces
  – Some formats don’t contain all the content desired (especially internally-developed formats)
Data Acquisition

• Test procedures can help address a number of issues
  – Identify data as production vs. engineering vs. QA samples
  – Label data for first pass, retest, equipment setup, etc.
    • P1 – first pass test
    • P2 – repeat first pass test (data replaces P1)
    • F1 – retest of failures (if applicable)
    • CORR – correlation data
  – Drive consistency in manually entered attributes

• Business decisions can affect usefulness of data
  – Lot number uniqueness
  – Part numbering
Data Acquisition

• Extracts from other systems
  – WIP tracking system
    • lot number / part number association
    • equipment information
    • lot genealogy
    • yield
  – Supplier data
    • lot number / part number association
    • yield
Data Transport – Alternatives

- **Direct transfer (file copy or move)**
  - Easy, but not robust
  - Suitable for moving data off the tester

- **Web interface**
  - User friendly, but not ideal for automation
  - Good solution for user access to raw data in archive

- **FTP**
  - Widely supported and easy to automate
  - Doesn’t handle interruptions well, but enhanced FTP tools are available

- **EDI / RosettaNet**
  - Addresses concerns, but not everyone supports it
Data Transport – Skyworks’ Solution

• Tester to Server and Within Systems – automated file copy with cleanup

• Site to Site and Between Systems – 3rd party enhanced FTP software
  – Scheduled FTP transfers
  – Successful transfer is verified before file is available
  – Clean up after successful transfer
  – Pull and push files, so we have control

• Mapped network drive for access to raw files

• Web-based and client-server tools for using data

Optimal solution is very dependent on the situation.
Data Storage and Archiving

• Storage options: file server, database, removable media (tapes, CDs, DVDs)

• Central and/or distributed storage
  – Network performance and site business rules are factors

• Considerations
  – Business rules for data retentions
  – Accessing the data (indexing, directory structure, security, etc.)
  – Store original and/or processed data
  – Storage space required
  – Don’t forget about backups!
Data Storage and Archiving – Skyworks’ Solution

• Central storage of original data, plus storage at data collection sites

• Fault-tolerant file server

• Hierarchical directory structure
  - part number
  - mode (production, engineering, QA sampling)
    - operation (probe, PCM)
    - data type
    - month code (e.g., 2005_06)
  - test location

• Automated processes to compress files, clean up, etc.

• Daily backups
Data Sorting

• Automatic data processing step to determine what needs to be done with the data, and direct the data accordingly

• Sorting is based on where the data came from and the content of the data and (potentially) from outside data
  – Example: Internal production probe data sent to foundry
    • data source identifies it as internally collected
    • data content identifies it as production and probe
    • outside data (reference database) identifies foundry

Data sorting outside the data transport process provides greater flexibility.
Data Sorting

File

Parse File

Look Up Lot Information

Lot Found?

Yes

Retrieve Lot Information

No

Process as Engineering Data

Use File Attributes to Determine Destinations

Copy/Move to Target Locations

Reference Database

Directory Attributes
- Data Type
- Test Mode (production, engineering, QA)
- Operation (probe, PCM)
- Test Location

File Attributes
- Part Number
- Data Type
- Test Mode (production, engineering, QA)
- Operation (probe, PCM)
- Test Location
- Tester Type
- Data Validity
Data Sorting

• Handling duplicate files
  – Files can be deleted here if they are identical to files in archive
  – If files in archive are compressed, need to uncompress to compare

• Handling different files with the same name
  – Add timestamp to filename to avoid overwriting and losing data

• Logging
  – Extremely useful for troubleshooting

• Monitoring
  – If all data passes through a sorting step, this becomes a good point to monitor incoming data streams
Data Preprocessing

• Preparing data for loading into test data analysis tools
• Isolates data loaders (often vendor supplied) from data cleanup
• Allows incorporation of additional data
  – External test limits
  – Lot and part information
• Simplifies cleanup and standardization of content
  – Case standardization
  – Lot number / part number cleanup

Preprocessing maximizes the usefulness of data analysis tools.
Data Preprocessing

1. File
   - Parse File
   - Look Up Lot Information
     - Lot Found?
       - Yes: Retrieve Lot Information
       - No: Process as Engineering Data
         - Generate Output File(s)
           - Output File in Standard Format

2. Directory Attributes
   - Data Type
   - Test Mode (production, engineering, QA)
   - Operation (probe, PCM)
   - Test Location

3. Test Data Structure
   - Lot Attributes
   - Part Attributes
   - Testing Attributes (tester, operator, etc.)
   - Test Attributes (test names, limits, etc.)
   - Parametric Data
   - Bin Data
Preprocessing Considerations

• Handling and identifying retest data

• Handling and identifying data from QA sampling

• Data collected during test program development
  – Test sequence, conditions, limits, etc., change during test development, usually with no test program revision change

• Incomplete data
  – Could be legitimate (partially probed wafer) or garbage (empty file)

• Invalid data
  – Lot number 11111, 12345, glenn123, unknow.n, etc.
  – Dates in the future or distant past
  – Invalid characters
Preprocessing Considerations

• Enforce rules based on your data analysis tools
  – Mandatory information, formatting, etc.

• Production vs. engineering vs. QA sampling data
  – Unintentional mixing of data will skew analysis results
  – If in doubt, don’t identify it as production

• Sample loading of data
  – Can improve system performance and reduce storage requirements for data analysis tools
  – Needs to be applied uniformly

• Error handling

• Logging / notifications
Closing Thoughts

• Analyze your own situation and requirements
  – What are your data sources? Who are the consumers? What do they need?

• Know the data and the processes

• Plan for change
  – New equipment, new suppliers, new data analysis tools, etc.

• Monitoring is essential
  – “An unwatched process always degrades” – Bill Mann
  – Applications, data loading, disk space, etc.