Western Digital®

presents

INTERNATIONAL STORAGE TECHNOLOGY CONFERENCE

3rd Year

ISTC 2017

Western Digital.

Implementation of machine learning to increase yield, and to improve on-time delivery and reliability of the products

Presenter's Name: Manish Kumar Keshri

Title : Engineer, Systems Design Engineering

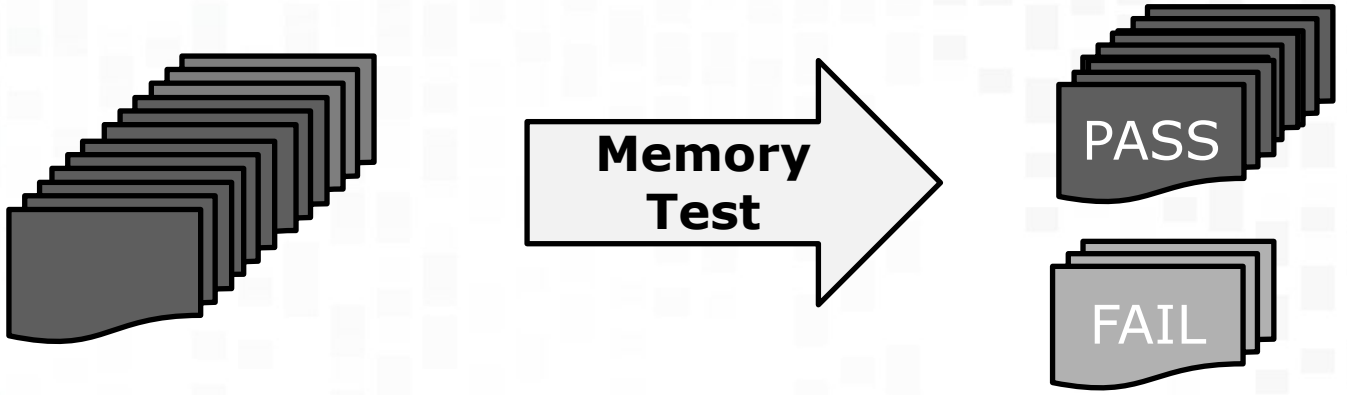
June 14, 2017

Co-Authors: Gururaj S B, Anwar Jabbar

Agenda

- Background
  - Selection process of memory units for product qualification tests, DAT Qualification, Firmware development, NAND Memory Characterization
- Suggested Approach
  - Clustering of units into Perfectly passing, Barely passing, Barely failing or Perfectly failing, based on memory parameters' values in Memory Test/Die Sort
  - Inclusion of memory units which are more prone to failures for tests, development and qualification
- Method
  - Use of Machine Learning techniques to cluster the units

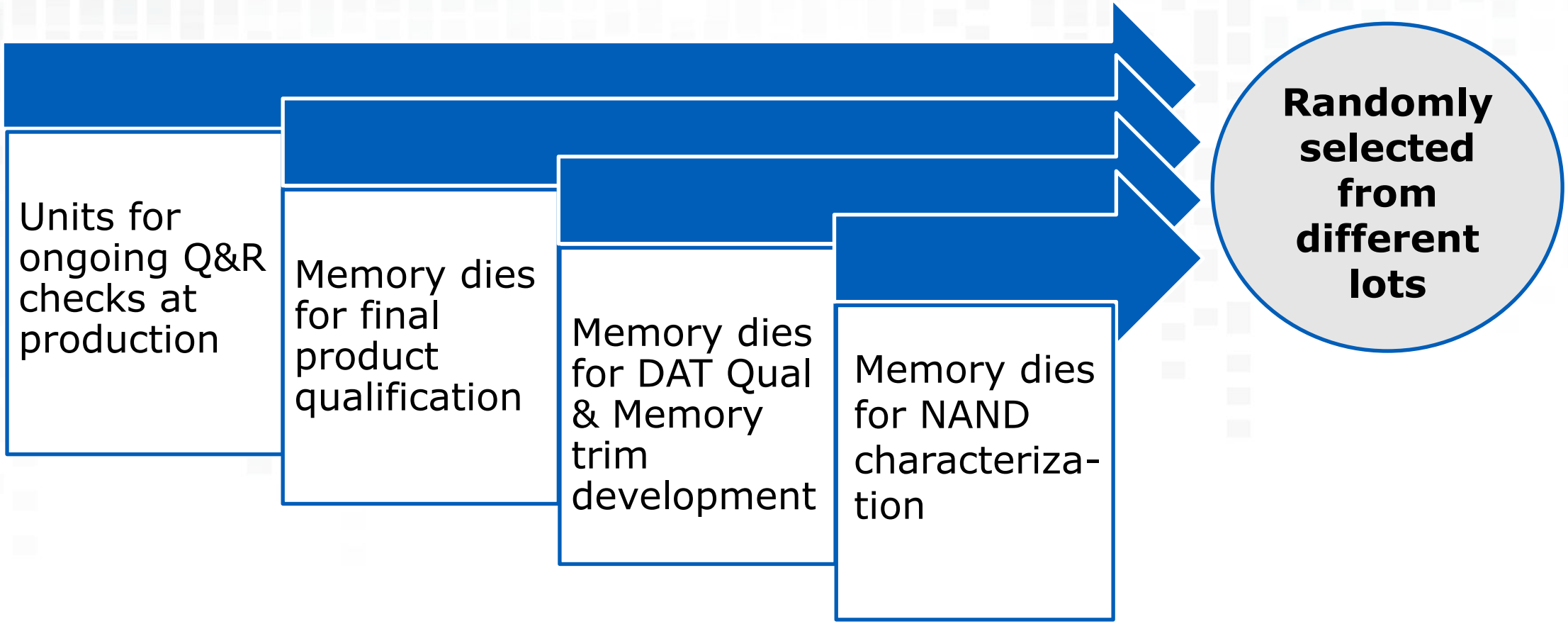
Background



Binary result(Pass/Fail) in Memory Test/Die sort

Determining quality of memory

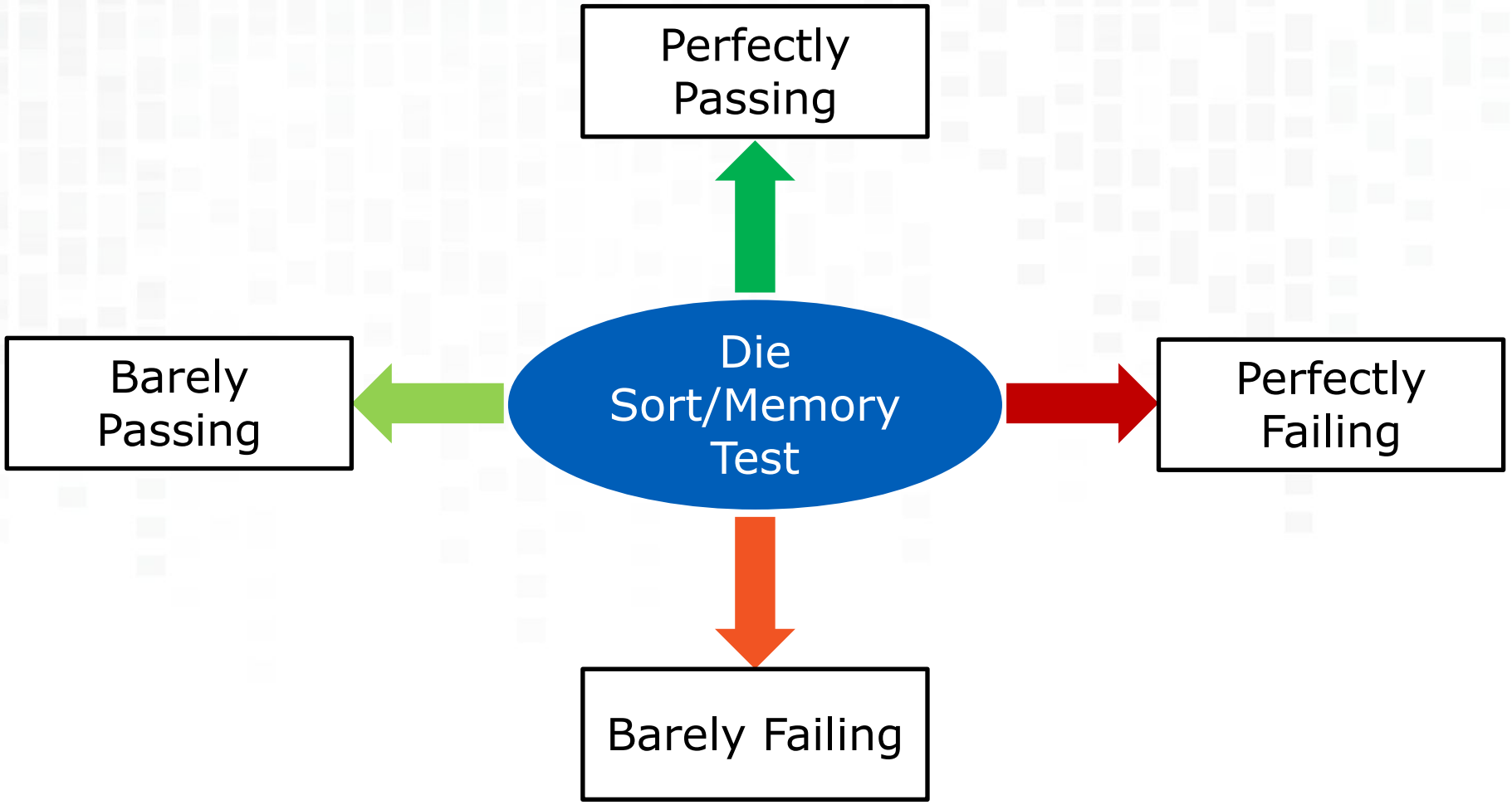
Background



Limitations of current process

- We might get all best units
  - Quality Tests on such units compromises the reliability
  - Trim developed on such units will not be robust
  - Most of the memory defects cannot be caught beforehand
- Yield loss on tests
  - Some percentage of units not meeting the limitations of test parameters

Suggested Approach



Advantages

- Identification & inclusion of Barely passing units in Product qualification, Quality Tests and Memory Trim development
  - More reliable products
  - Robust trim and FW development
  - Early shipping
  - Accurate DPPM

- Extend method to classify memory quality based on target controller capability without additional insertion steps in memory test & possible yield improvement

Memory Test data collected on 6 Memory dies passing Memory Test with current scheme

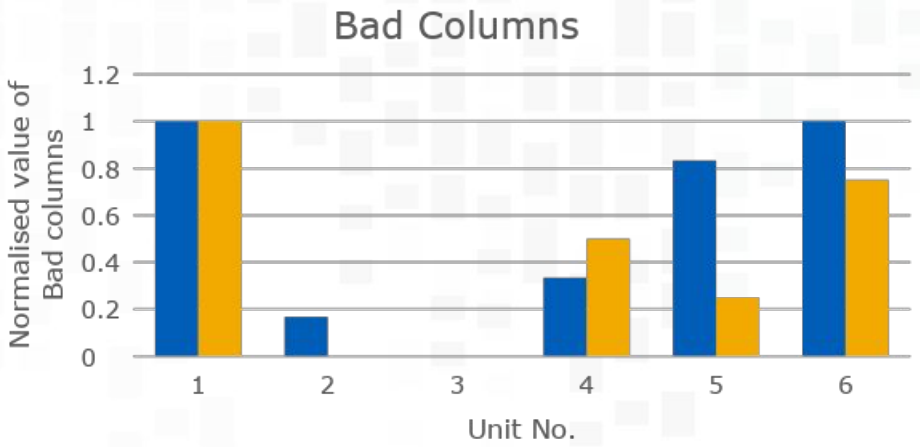
Unit No.	Bad Blocks		Bad Columns		Blocks barely passing MLC bit ignore criteria		Blocks barely passing SLC bit ignore criteria		Blocks barely passing VSGD criteria	
	Plane 0	Plane 1	Plane 0	Plane 1	Plane 0	Plane 1	Plane 0	Plane 1	Plane 0	Plane 1
1	16	0	6	5	1	0	63	0	0	0
2	0	4	1	1	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0	0
4	10	27	2	3	2	1	0	3	0	0
5	3	3	5	2	0	0	0	0	0	0
6	4	10	6	4	0	0	0	0	0	4

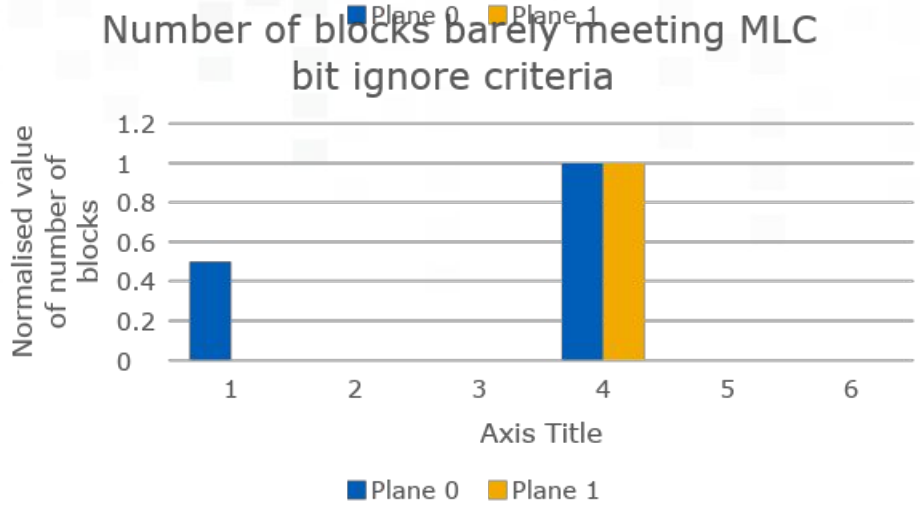
**Data source: Memory Test on BinZZ BiCS3 DAT6 Memory Dies*

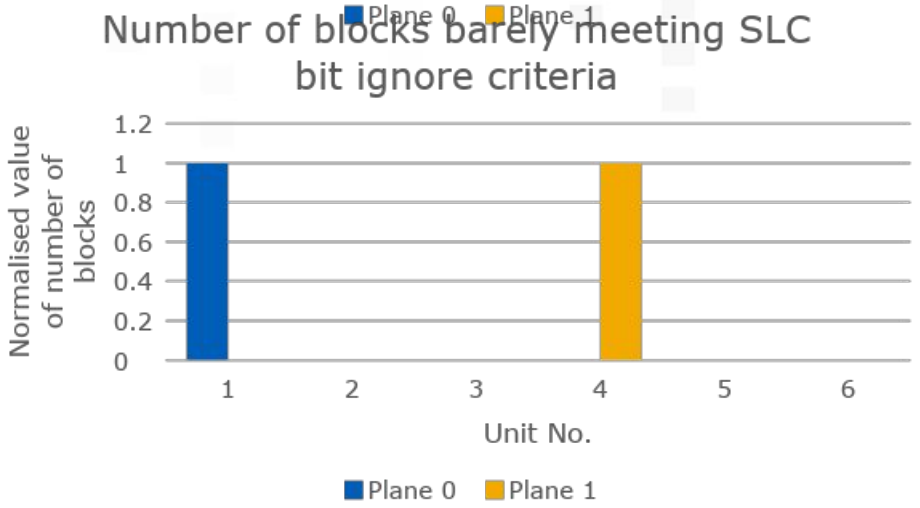
Categorization based on calculated index

Unit No.	Bad Block(1)		Bad Columns(0.9)		blocks barely passing MLC bit ignore criteria(0.4)		blocks barely passing SLC bit ignore criteria(0.5)		blocks barely passing VSGD criteria(0.7)		Index	Category
	Plane 0	Plane 1	Plane 0	Plane 1	Plane 0	Plane 1	Plane 0	Plane 1	Plane 0	Plane 1		
1	1	0	1	1	0.5	0	1	0	0	0	3.5	Barely Passing
2	0	0.1482	0.1667	0	0	0	0	0	0	0	0.2982	Perfectly Passing
3	0	0	0	0	0	0	0	0	0	0	0	Perfectly Passing
4	0.625	1	0.3333	0.5	1	1	0	1	0	0	3.675	Barely Passing
5	0.1875	0.1111	0.8333	0.25	0	0	0	0	0	0	1.2736	Perfectly Passing
6	0.25	0.3704	1	0.75	0	0	0	0	0	1	2.8954	Perfectly Passing
	Index = Σ(normalized value*weight)					IF index>3, Category= "Barely Passing"; else Category= "Best unit"						

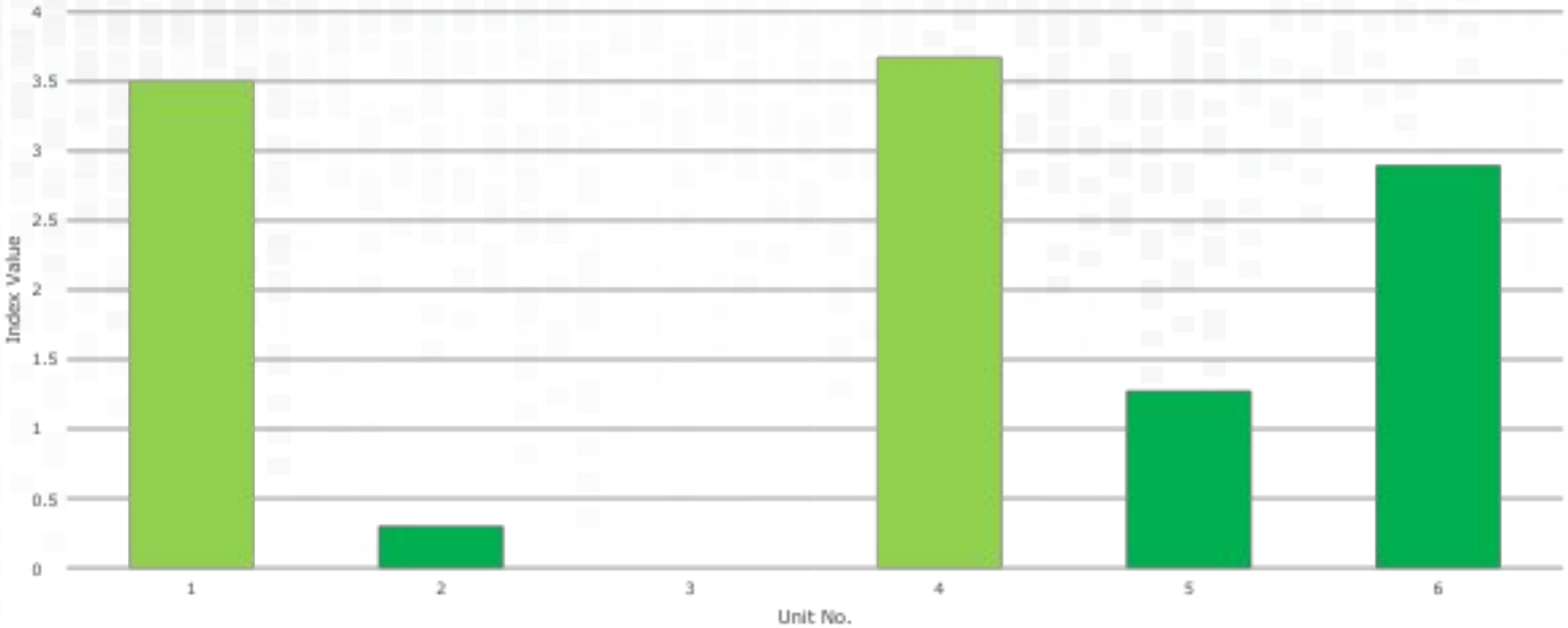
Graph: Normalized value of blocks for each parameter







Categorization based on calculated index



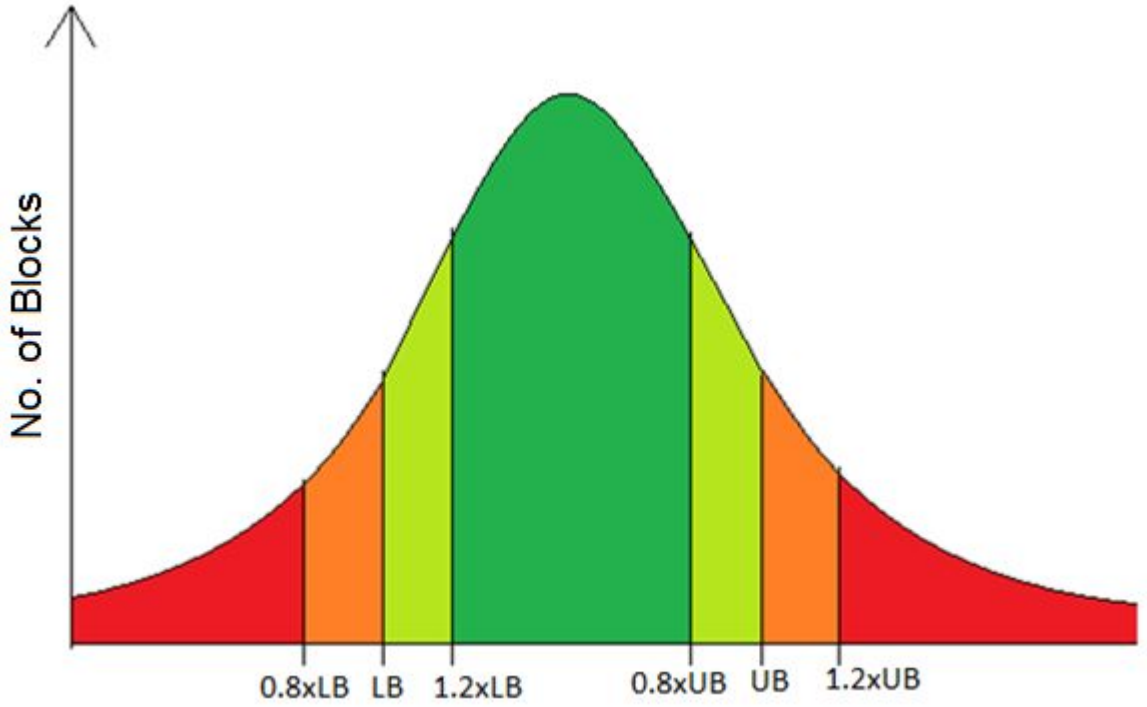
Barely Passing Unit

Perfectly Passing Unit

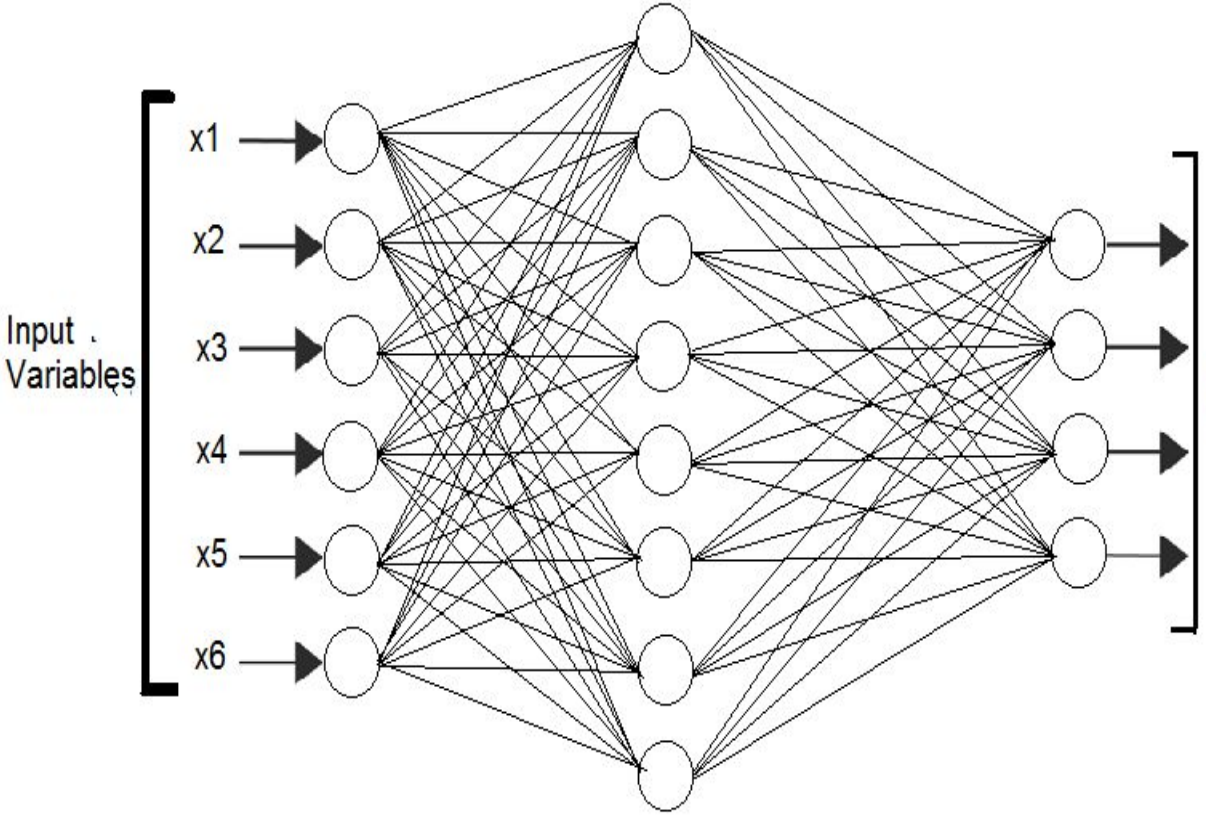
Classification of units

- For each parameter, find the number of blocks barely meeting or missing the window range criteria
- Assign weights to each parameter

  in ANN Matrix
- Train the classifier
- Produce results



Building the classifier:



$h_\theta(x) \in \mathbb{R}^4$

$h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, when a unit is perfectly meeting all parameter values

$h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, when a unit is barely meeting some/many parameters' values

$h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, when a unit is barely failing due to some parameter values

$h_\Theta(x) \approx \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$, when a drive fails due to significant deviations in parameter(s) values

A Feed Forward neural network

What parameters to select for clustering?

- More than 200 parameters in MT/Die sort: High Complexity
  - Suggestion to use parameters associated with common Quality Tests Failures:

Some common memory defects countered in quality tests	Associated parameter(s) in memory test with the defect	Allowed window range for the parameter value
WL-WL short	F_WLLD_IDT	250nA to 2000nA
	F_WLLD_ICM	1000nA to 2500nA
	F_VPGM_STRPCG	8V to 20.4V
Broken WL	F_VPGM_WL2SUB	8V to 20.4V
Slow to program	F_DR8_FREAD	2.4us to 8us
Slow to erase	F_LPCTRL_DVCGERV	0.05V to 0.4V
	F_LPCTRL_DVCGERV_SLC	0.05V to 0.4V

Challenges

- Assigning the weights to each parameter
- Placement of algorithm in product flow hierarchy

Western Digital.

Thank you!

What is the problem in random selection?

- One cycle of Read through CNE on 5 units to get FBC count per 2K bits read:
- Unit 1: A failure unit
- Unit 2 and 4: Perfect Units. They have very low FBC Count
- Unit 3 and 4: Barely passing units. Very susceptible to failures in Quality Tests

Unit	Read with FBC>100	Read with FBC>122
1	582984	582984
2	0	0
3	118	0
4	0	0
5	22	0

On random selection, we might get all perfectly passing units and then,

☐Quality Tests on such units will not be that reliable

☐Trim developed on such units will not be robust

Limitations of current process

Meaning of best units and barely passing units:

- Best units:
  - Meeting all/most of the parameters in tightened limits(tightened to 20% of the margins)
- Barely passing units:
  - Passing Memory Test/Die Sort
  - Barely meeting some/many parameters within the limit

Advantages

- Increased Reliability
- Improved on-time delivery of products
- Increased yield
- Accurate DPPM

Advantages

- Use of Barely passing units for OQRM, Quality Tests and Trim development
  - More reliable products
  - Robust trim and FW development
  - Early shipping
  - Accurate DPPM

- Changes in Trim, FW to include barely failing units
  - More yield

- MT scheme customization
  - Full array MT can be done for Barely passing LOT while sample MT can be performed for perfectly passing units, leading to accurate and effective memory screen