Raytheon
Customer Success Is Our Mission

Statistically-Based
Test Optimization Strategies

23rd Annual NDIA Test & Evaluation Conference
March 13, 2007

Neal Mackertich, Ph.D
Raytheon Integrated Defense Systems

Best Practices in Statistically-Based Test Optimization

Why, What , When & How: There are lots of meaningful testing questions and opportunities to explore…

The focus of this presentation is to share with you three industry cited best practices around the use of statistically - based test optimization strategies:

- Usage-based Statistical Testing
- Combinatorial Design Methods
- Critical Parameter Management

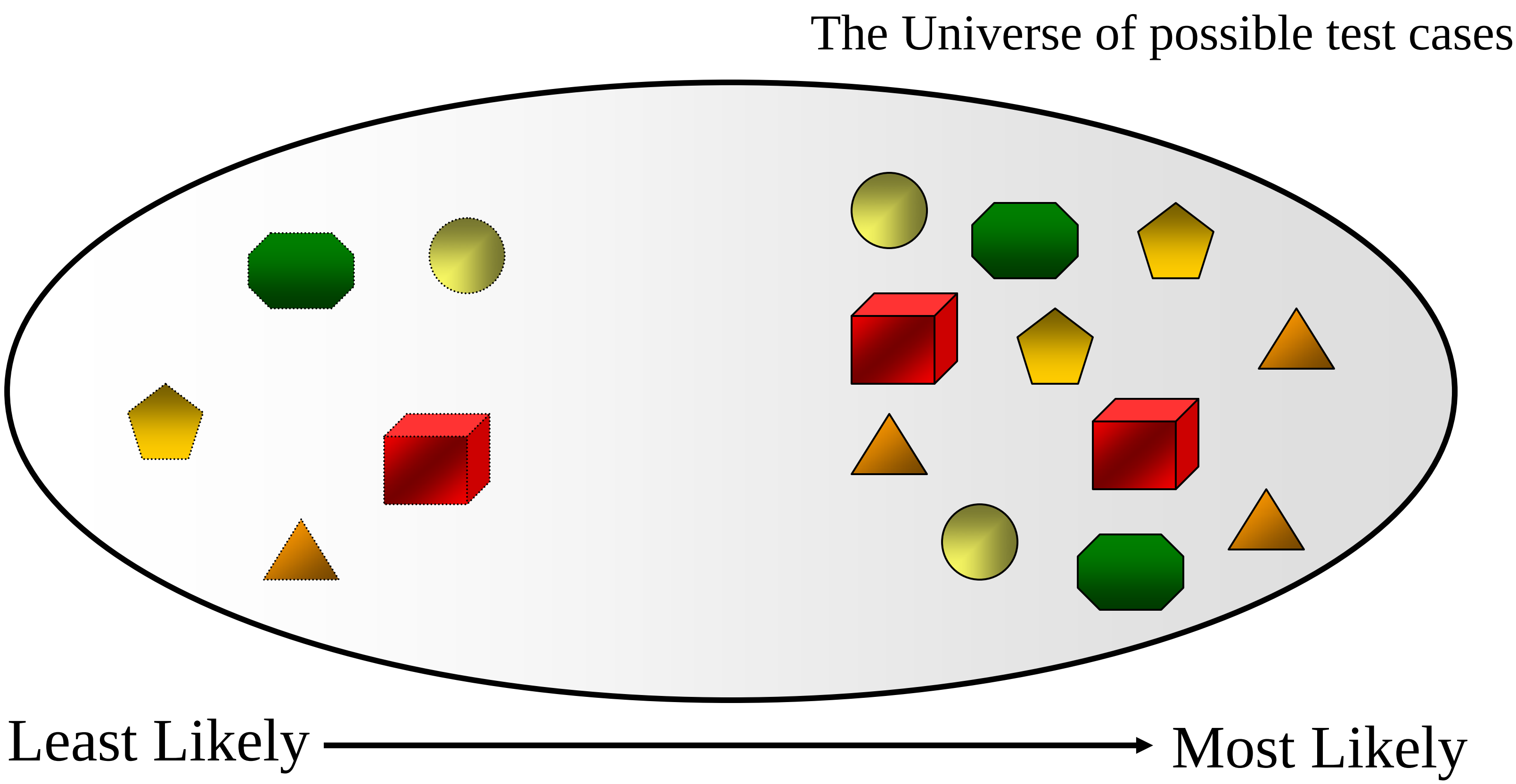
What to Test?

Okay, let's say we need to test a system or subsystem…

How do we typically determine what test cases to run?

Usage-Based Statistical Testing

- Use-Based Statistical testing emphasizes the operational scenarios most likely to occur



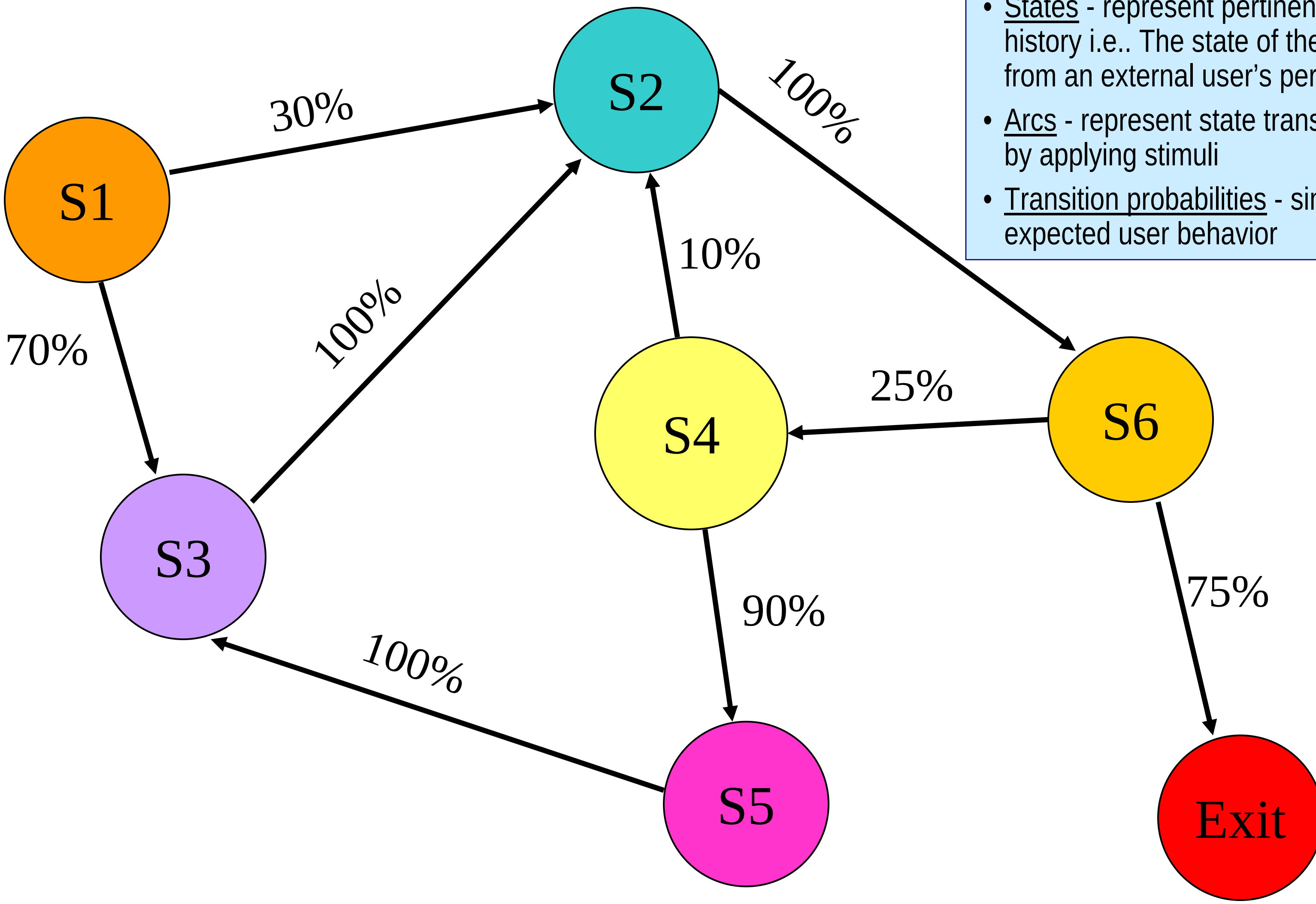
Raytheon

Usage-Based Statistical Testing

- Testing based on the way the system / product is to be operational used.
- A “Use Case” state diagram (in the form of a Markov Model) is used to generate a representative, random sample usage so statistical methods can be applied to model System behavior.
- Specific test case generation is accomplished using input stimuli selected via a random walk through the Markov chain.

Raytheon

Sample Usage Model



- States - represent pertinent usage history i.e.. The state of the software from an external user's perspective.
- Arcs - represent state transitions caused by applying stimuli
- Transition probabilities - simulate expected user behavior

Usage-Based Statistical Testing- Case Study Results

- Escaping defects were less than previous methods (1.16 defects per KSLOCs Vs. 6-defects/ KSLOCs) from previous testing method.

- Development costs met budgets.

- Major functions were integrated quickly.

- Poll was taken with the leads about whether to use this statistical testing on a follow-on program; results were unanimous to use statistical testing for next project which has already started!!

What are Combinatorial Design Methods?

- A testing methodology in which a subset of all possible combinations is chosen such that all N-way combinations are tested.
- Covering all 2 way combinations would require that for any two factors A and B, where Ai and Bi are valid levels for A and B, there is a test for all Ai and Bi combinations.

Why not DOE?

The use of proven statistical / combinatorial methods can be very helpful here…

But why not just use Design of Experiments (DOE) techniques?

CDM Advantages

- The use of N-way combinations provide reasonable, balanced coverage across the test space.
- More realistic than full/fractional designs
  - Compatible with constraints
  - Compatible with factors at different levels
  - Can account for previous test
- Drastically reduces the total number of test cases when compared to all combinations.
- Since generating test cases is very quick and simple, there are no major barriers to using CDM as part of the testing process.
- Can be used in almost all phases of testing.

Raytheon

Black-Box Testing Scenario

Definition	Black-box type testing geared to functional requirements of an application.
Example Application	Graphics manipulation function that converts from one format to another.
Inputs	Source Format (GIF, JPG, TIFF, PNG) Dest. Format (GIF, JPG, TIFF, PNG) Size (Small, Med, Large) # colors (4 bit, 8 bit, 16 bit, 24 bit) Destination (Local drive, network drive) Windows Version (95, 98, NT, 2000, Me)
Outputs	Correct conversion (True or False)
Constraints	If Destination Format is GIF, then # colors cannot be 16 bit or 24 bit.
All Combinations	1920 Test Cases
2 Way Combinations	22 Test Cases

Regression Testing Scenario

Definition	Re-testing after fixes or modifications of the software or the environment in which the software operates.
Example Application	Graphics functionality on previous page must be expanded to support a new file format (WMF) and a new OS (Windows XP).
Inputs	Source Format (GIF, JPG, TIFF, PNG, WMF) Dest. Format (GIF, JPG, TIFF, PNG, WMF) Size (Small, Med, Large) # colors (4 bit, 8 bit, 16 bit, 24 bit) Destination (Local drive, network drive) Windows Version (95, 98, NT, 2000, Me, XP)
Outputs	Correct conversion (True or False)
Constraints	If Destination Format is GIF, then # colors cannot be 16 bit or 24 bit
All Combinations	1620 Test Cases
2 Way Combinations	32 Test Cases

CDM- Case Study Application

- Testing program requirements
  - Factory Acceptance Test (FAT) Dry Run
  - FAT
  - Site Acceptance Test (SAT) Dry Run
  - SAT

- Not realistic to do exhaustive testing of all 144 possible System test scenarios

- Quasi-Exhaustive strategy invented
  - 100% of tests for FAT Dry Run
  - 10% of tests, selected at random, for FAT
  - 50% of tests, selected at random, for SAT Dry Run
  - 10% of tests, selected at random, for SAT

Raytheon

CDM Case Study- A Comparison of Strategies

Quasi-Exhaustive Strategy	FAT Dry Run	FAT	SAT Dry Run	SAT
Number of test cases	144	15	72	15
Time spent writing test procedures	49.00 hours	0.00 hours	8.00 hours	0.00 hours
Time spent performing data paths test procedures	73.25 hours	8.75 hours	55.00 hours	12.25 hours
Number of persons involved in testing	1	3	3	4
Total Labor Hours	370.5 hours (46.3 days)			

CDM Strategy	FAT Dry Run	FAT	SAT Dry Run	SAT
Number of test cases	12	12	12	12
Time spent writing test procedures	16.00 hours	0.00 hours	3.00 hours	0.00 hours
Time spent performing data paths test procedures	7.25 hours	7.25 hours	10.00 hours	10.00 hours
Number of persons involved in testing	1	3	3	4
Total Labor Hours	118.0 hours (14.8 days)			

Bottom Line Comparison (Time-to-Market Savings)

- CDM Strategy was superior to Quasi-Exhaustive Strategy
  - Schedule savings = 68%
  - Cost savings (labor) = 67%

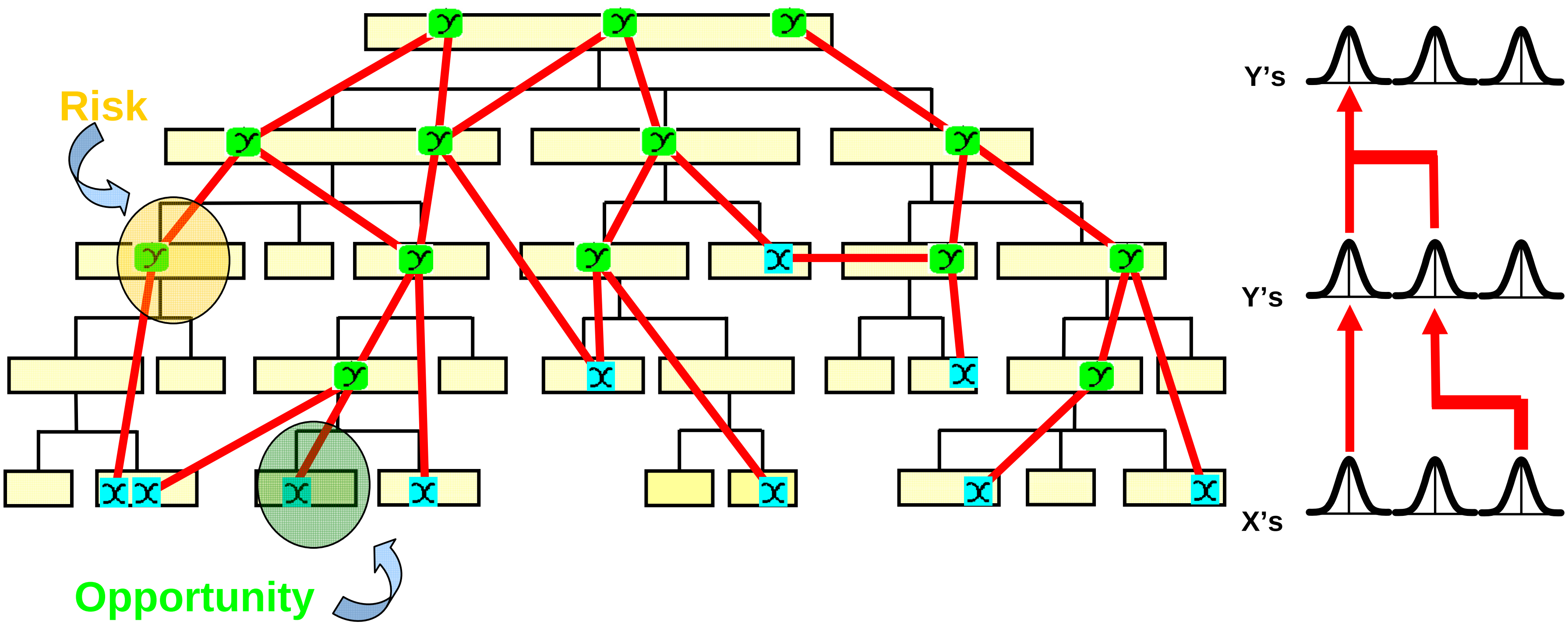
What is Critical Parameter Management?

- A disciplined methodology for managing, analyzing, and reporting technical product performance.
- A process for mathematically linking system parameters for sensitivity analysis and optimization of critical performance threads.
- A strategic tool for improving product development by unifying and integrating systems, design, manufacturing and test activities.

CPM = TPMs + Other parameters critical to functionality, cost, schedule or customer

Performance Analysis

"The System Can…."

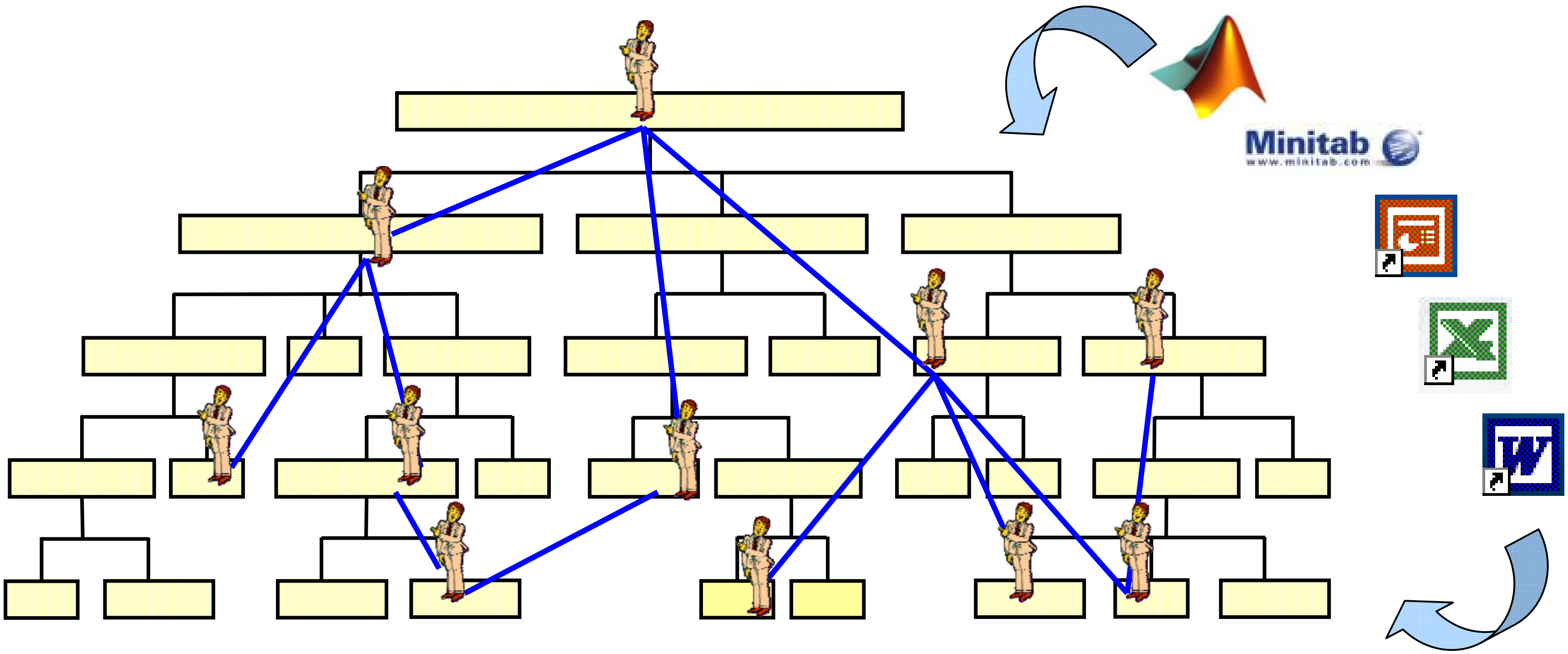


CPM Statistically Flow-Up Design, Supplier and Manufacturing Capabilities Exposing Performance Risks and Opportunities

Raytheon

Whitepaper Management and Task Delegation

Engineering Documents and Whitepaper Analyses



Attach Engineering Documents, Models, and Whitepaper Analysis.

Connect people to analyses, requirements, and performance measures.

CPM Program Benefits

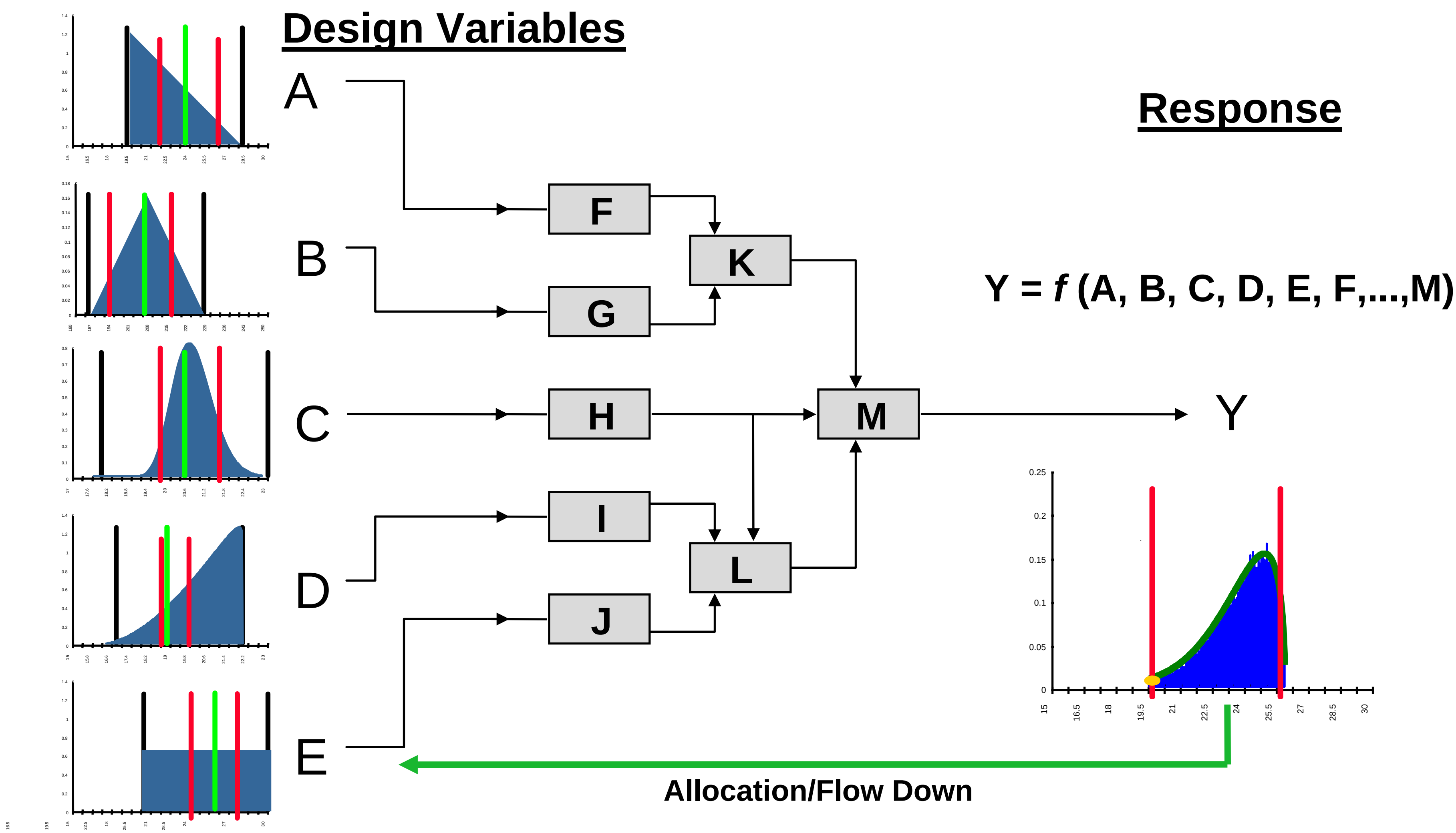
- Facilitate Analysis
  - Statistical modeling & optimization of the performance – cost trade space
  - Real-time System-level sensitivity analysis
  - Connects analyses between system, subsystem and component levels

- Improve Collaboration
  - Shares technical analysis and knowledge
  - Links ownership to parameters
  - Mathematically connects Program teams and parameters to understand requirement flow-down
  - Captures and leverages invested intellectual capital for future business reuse

- Enable TPM Management and Reporting
  - TPM design margins are statistically tracked over product lifecycle
  - Automated, real-time TPM data gathering / report generation
  - Reconciliation of requirement allocation and engineering design capability

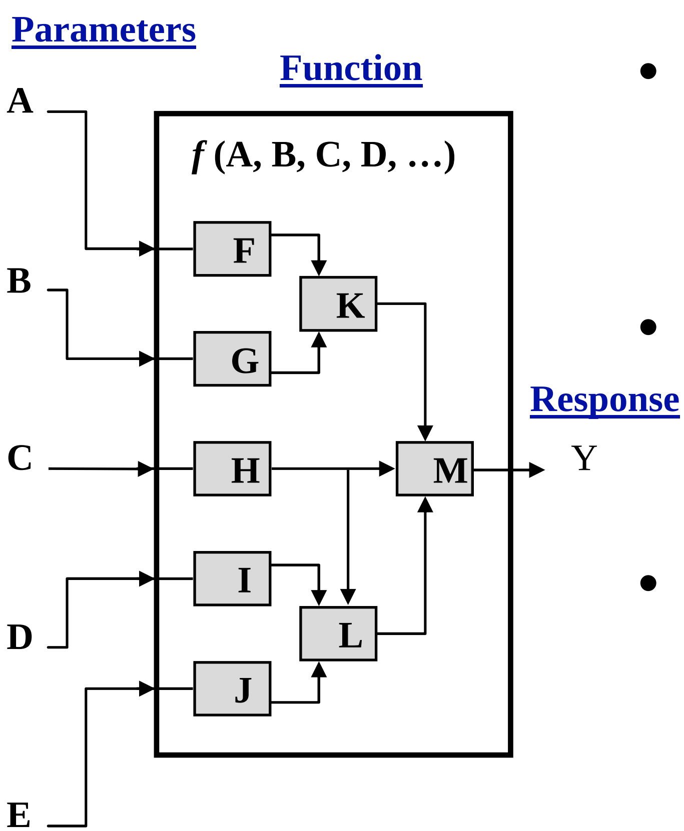
Raytheon

DFSS Statistical Performance Analysis

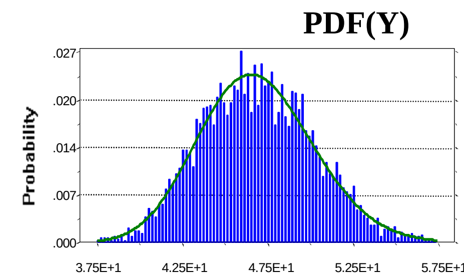


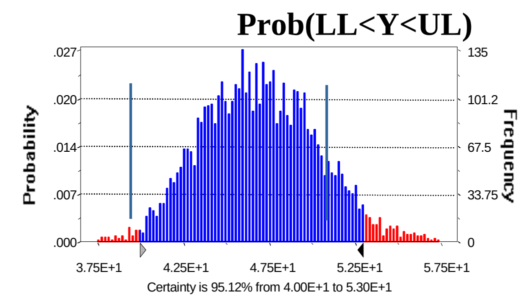
Raytheon

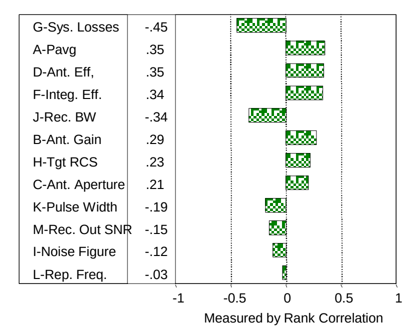
Statistical Performance Analysis Results In…



- **A prediction of the Response Statistical Properties**
- **A prediction of the Probability of Non-Compliance**
- **An assessment of the Contribution of Parameter Variation to Response Performance and Cost**







Results from Crystal Ball® Monte Carlo SW

Traditional TPM Stoplight Reporting

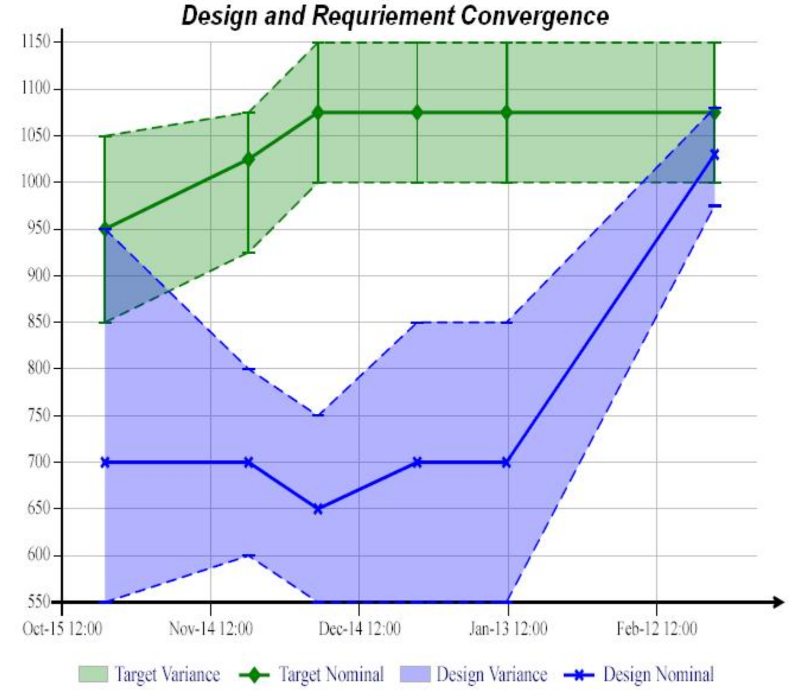
TPM Number	Description	Aug '05	Sep '05	Oct '05	Nov '05	Dec '05	CDR Jan '06
TPM-001	Single Pulse Sensitivity	G	G	G	G	G	G
TPM-002	Search Sensitivity	G	G	G	G	G	G
TPM-003	Range Accuracy	G	G	G	G	G	G
TPM-004	Angle Accuracy	G	G	G	G	G	G
TPM-005	RCS Accuracy	G	G	G	G	G	G
TPM-006	Phase Stability	Y	Y	Y	Y	Y	G
TPM-007	Polarization Isolation	G	G	G	G	G	G
TPM-008	Ellipticity	Y	Y	Y	Y	Y	G
TPM-009	Range Sidelobe Level	G	G	G	G	G	G
TPM-010	Range Resolution	G	G	G	G	G	G
TPM-011	2-Way Notch Depth (combined)	Y	G	G	G	G	G
TPM-012	Receive Pattern Sidelobe Level	Y	G	G	G	G	G
TPM-013	Weight	G	G	G	G	G	G

G	Meets Requirement with Margin
Y	Meets Requirement with No Margin
R	Does Not Meet Requirement

Previous TPM Reporting / Tracking method is vague and ambiguous with respect to the design margin for each metric. TPM report was only tracked and managed on a monthly basis for the tracking book.

Statistically-Based TPM Reporting / Management

- Statistically track design capability and requirement by establishing upper and lower limits
- Monitor design capability and requirement convergence over product lifecycle:
  - IPDR, PDR, CDR etc.
  - Design, Manufacturing, Integration, Test

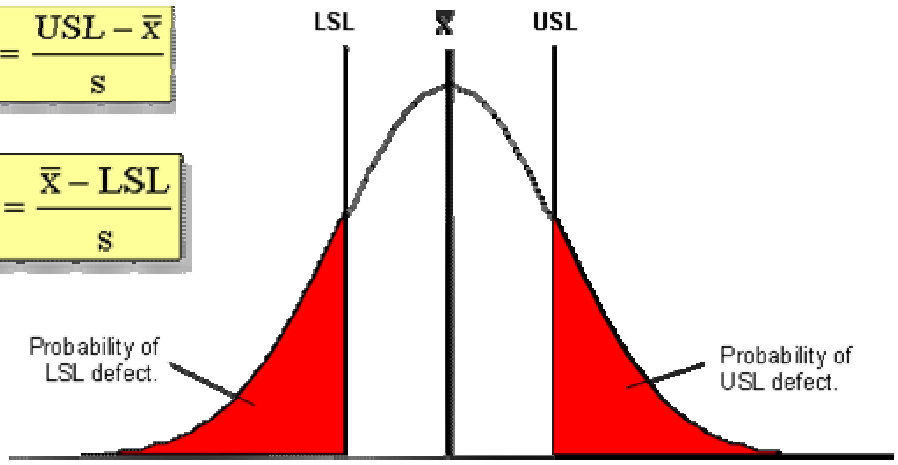


Probability of Non-Compliance

$$PNC = PNC(Z_{USL}) + PNC(Z_{LSL})$$

$$Z_{USL} = \frac{USL - \bar{x}}{s}$$

$$Z_{LSL} = \frac{\bar{x} - LSL}{s}$$



Cp and Cpk

$$Cpk = \min\left(\frac{USL - \bar{x}}{3 * s}, \frac{LSL - \bar{x}}{3 * s}\right)$$

$$Cp = \frac{USL - LSL}{6s}$$

New Automated TPM Reporting Format

TPM design margins are statistically tracked real-time

Block Name: TPMs | **BOM ID#:** | 2/9/2006

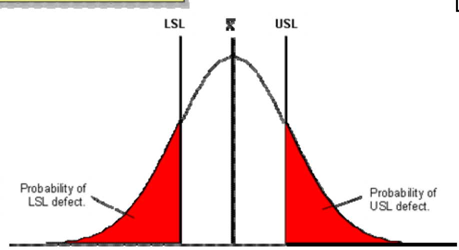
	CFR (Y)	Spec Code	Owner		Units	Target Value Lower Limit	Target Value Nominal	Target Value Upper Limit	Design Model Prediction Nominal	Design Model Prediction Mean	Design Model Prediction Std. Dev.	Design Model Prediction Cp	Design Model Prediction Cpk	Validation Test Data Mean	Validation Test Data Std. Dev.	Validation Test Data Cp	Validation Test Data Cpk
	TPM 1	R				0.9400	0.9600	0.9990	0.9672	1.9040	0.6238	0.0158	-0.9836	0.0000			
	TPM 2	R				6.0000	8.0000	11.0000	7.0000	9.5000	0.5833	1.4286	0.3571	0.0000			
	TPM 3	Y				43.0000	45.0000	49.0000	43.0000	42.7500	0.3750	2.6667	-0.7222	0.0000			
	TPM 4	G				14.0000	17.0000	19.0000	16.0000	16.5000	0.4167	2.0000	1.5000	0.0000			
	TPM 5	G				0.1000	0.1200	0.1600	0.1250	0.1250	0.0008	12.0000	9.5000	0.0000			
	TPM 6	R				4.0000	6.8000	8.0000	7.0000	7.0000	0.0833	8.0000	3.5000	0.0000			
	TPM 7	Y				0.4000	1.0000	1.2000	0.6000	0.7000	0.0333	4.0000	2.8333	0.0000			
	TPM 8	G				0.4000	0.4500	0.5000	0.4500	0.4550	0.0042	4.0000	3.4333	0.0000			

Meaningful stoplights based on statistical sensitivity

Compares requirement allocation vs. current design capability

Includes sensitivity and statistical information

$$PNC = PNC(Z_{USL}) + PNC(Z_{LSL})$$



Captures complete product lifecycle: requirements, design, manufacturing, testing, validation, etc. (other columns not shown)