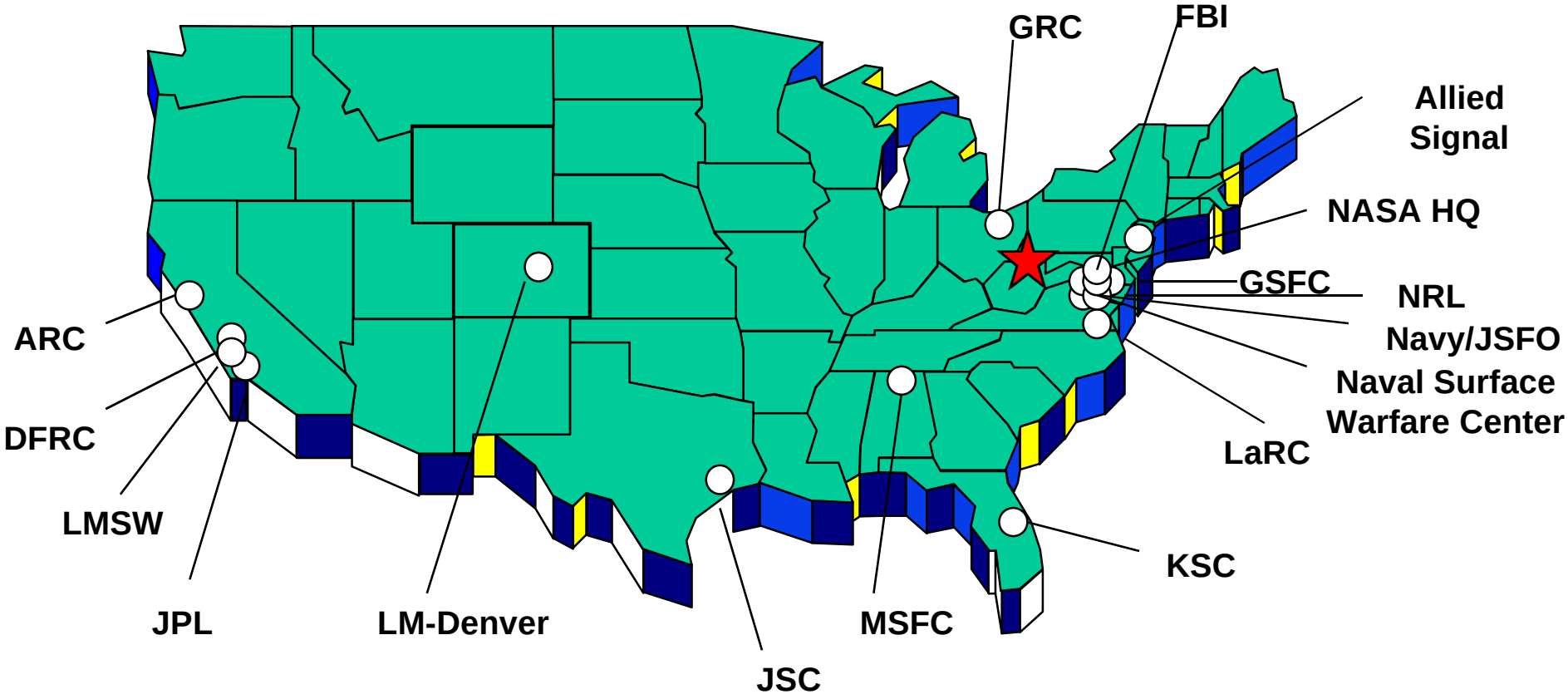
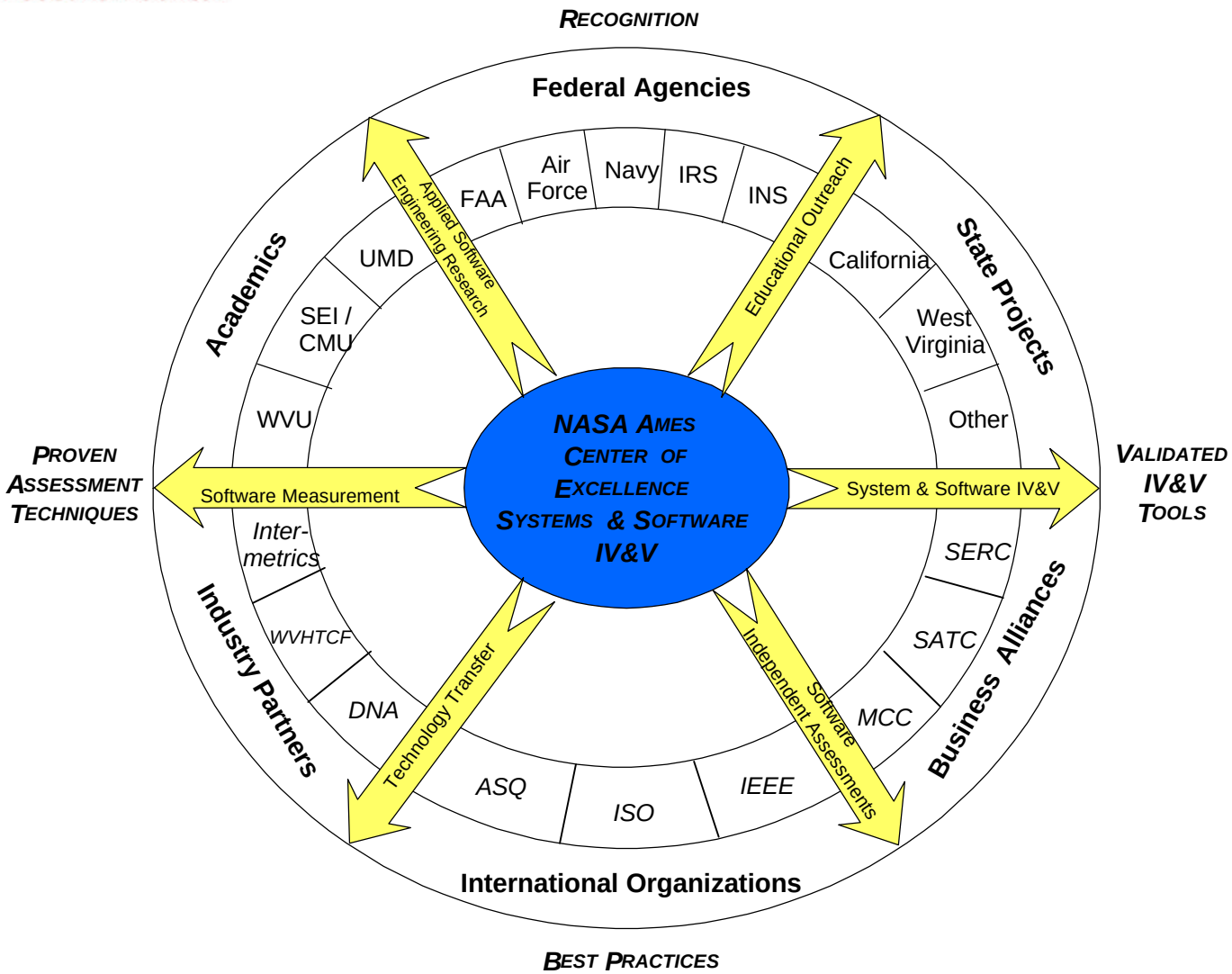
Software Measurement Frameworks to Assess the Value of Independent Verification and Validation

Dr. Nancy S. Eickelmann
WVU/Software Research Laboratory
NASA Independent Verification and Validation Facility
Phone: (304) 367-8444
E-mail: Nancy.Eickelmann@ivv.nasa.gov
Website: http://research.ivv.nasa.gov/~ike
December 1999

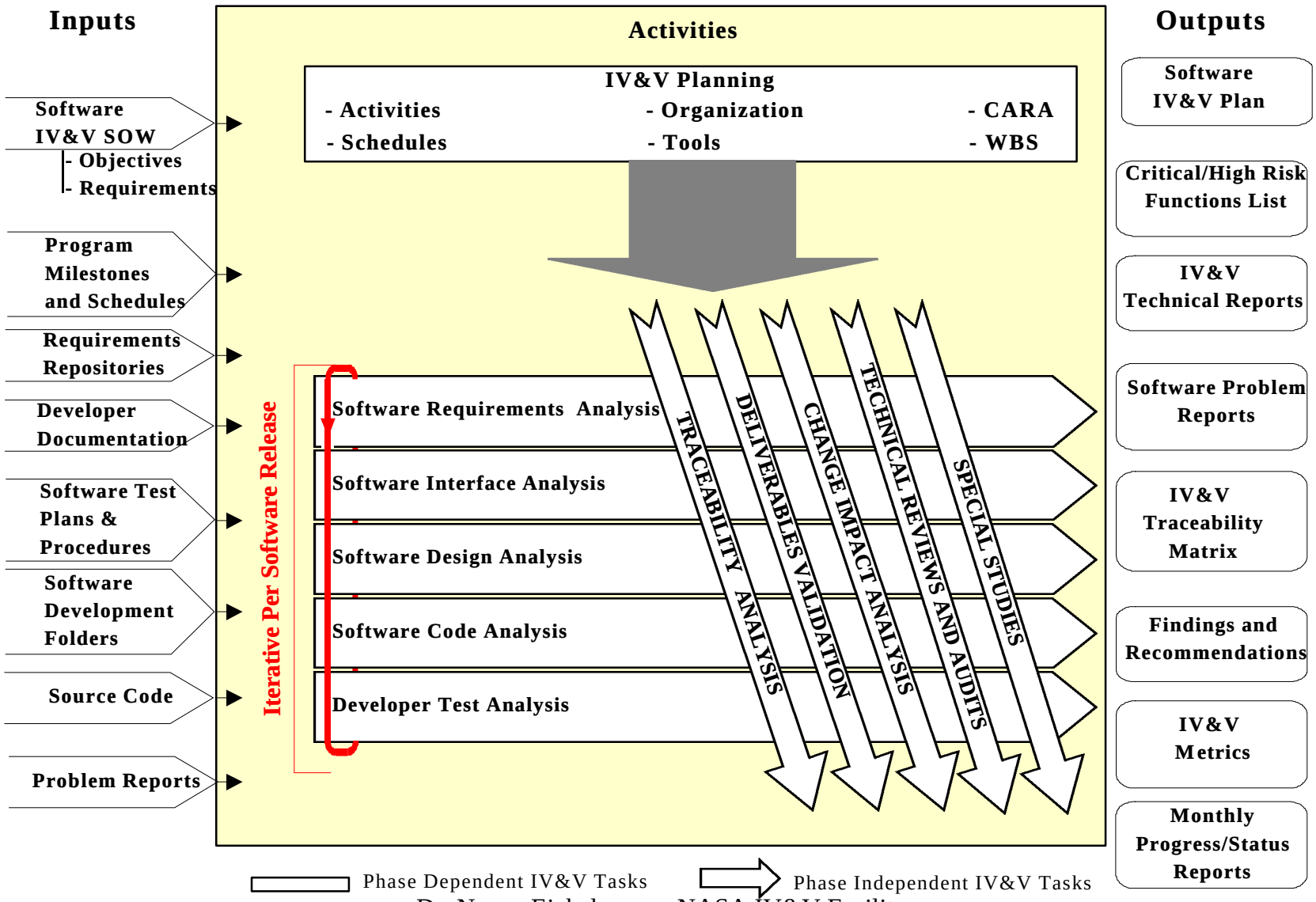
IV&V Infrastructure Supports Programs Across Country



NASA IV&V Core Customers, Processes and Practices



Iterative IV&V Methodology



IV&V Goals

- IV&V processes provide an objective assessment of software products and processes throughout the software life cycle.
- This assessment demonstrates whether the system and software requirements are ***correct, complete, accurate, consistent***, and ***testable***.
- Provides confidence through objective evidence that reliability, safety and quality goals will be met.

- Other objectives of performing V&V are to:
  1) Facilitate *early* detection and correction of software errors;
  2) Enhance management *insight into process and product risk*; and
  3) Support the software life cycle processes to ensure compliance with program *performance, schedule, and budget* requirements.

Measuring Economic Value

Financial and accounting measures report short-term and historical results. They fail to communicate whether an organization is creating economic value through…

- process improvements
- technological product innovations
- investments in tools, methods, and models

The Balanced Scorecard provides meaningful measures of the dynamics among organizational entities, their relationships and behaviors in achieving NASA IV&V mission through the use of...

- cause and effect graphing
- identifying performance and outcome measures
- linking these measures to financial performance

Balanced Score Card

NASA IV&V
Balanced Score Card

Strategic Goals: "To become an international leading force in the field of software engineering for improving: safety, quality, reliability, cost performance of software systems."

Finance

"How can we reduce costs and not compromise our mission?"

Customer

"To achieve our goals how do we want our customers to perceive us?"

Core Business Processes

"To satisfy our customers what business processes do we need to excel at to differentiate us and create a COE?"

Learning & Growth

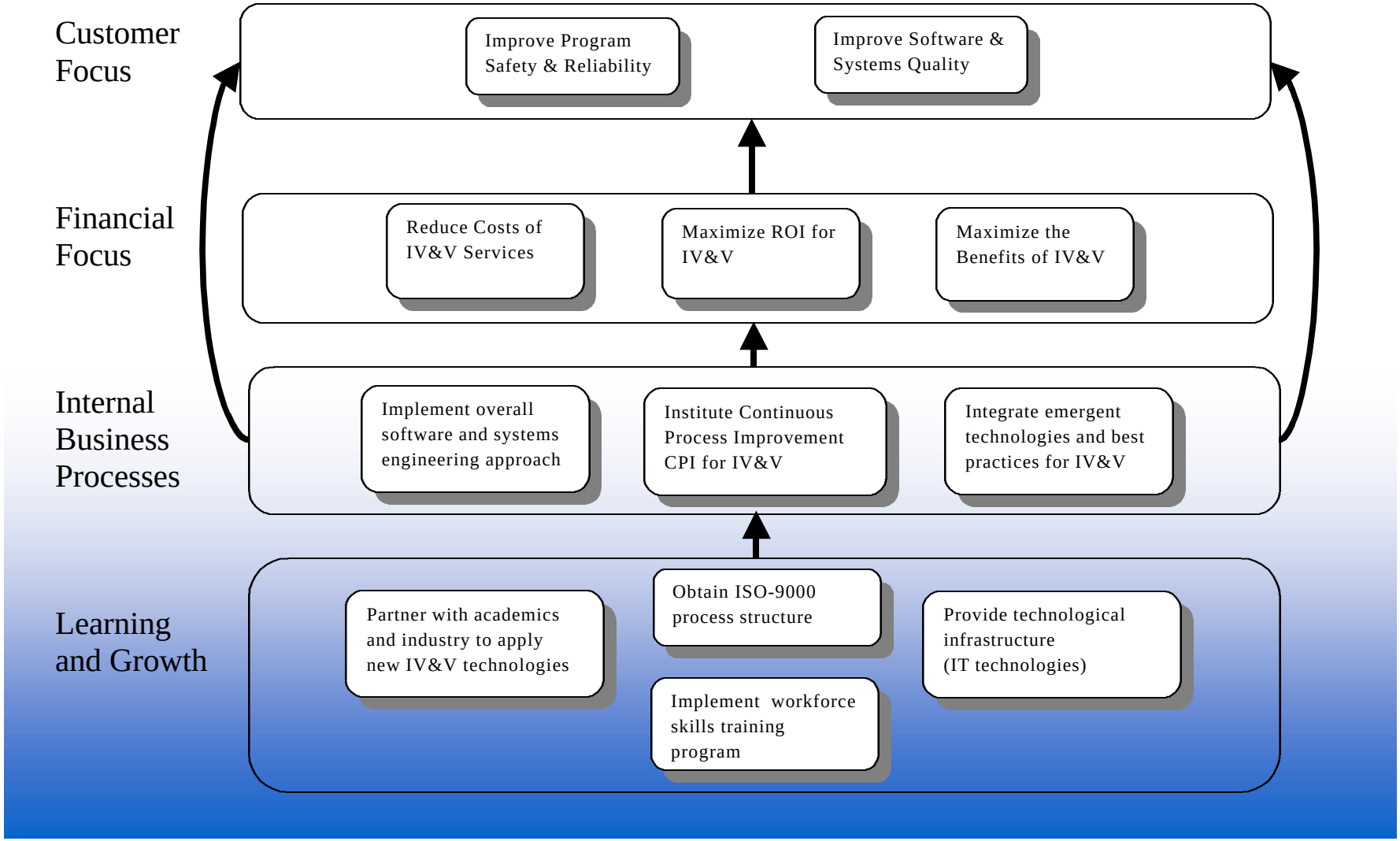
"What infrastructure do we need to sustain our ability to change and improve?"

ROI Research will provide objective measures for evaluating:

- **safety**
- **quality**
- **reliability**
- **cost**

NOTE: A Balanced ScoreCard (BSC) is a strategic measurement framework that makes visible the critical linkages between strategic goals and operational realities. BSCwas developed by Kaplan and Norton (Harvard Business School). Dr. Eickelmann has customized the framework for application in the NASA IV&V organization.

NASA IV&V
BSC Objectives

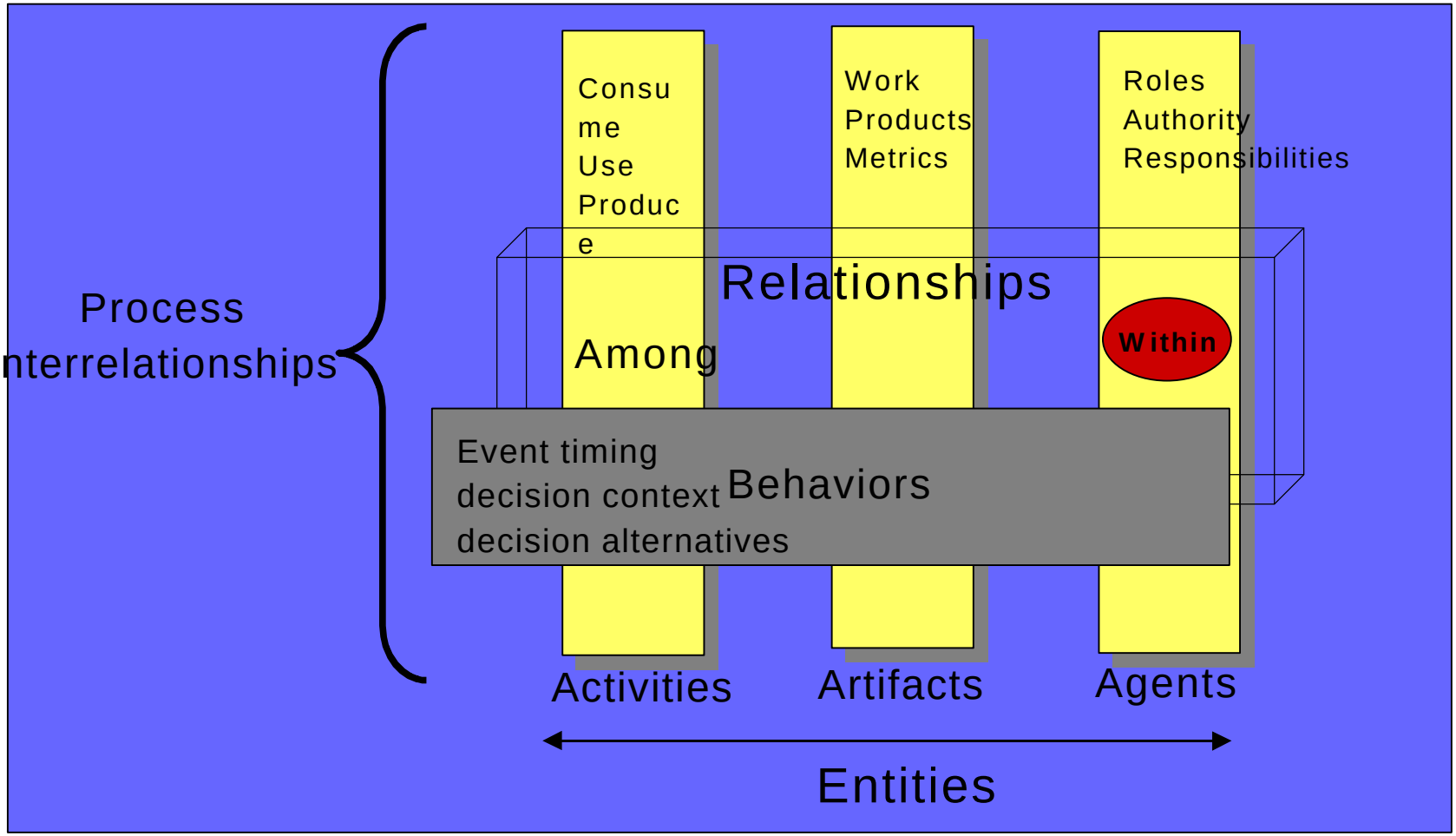


Commonly Applied IV&V Developed and COTS Tools

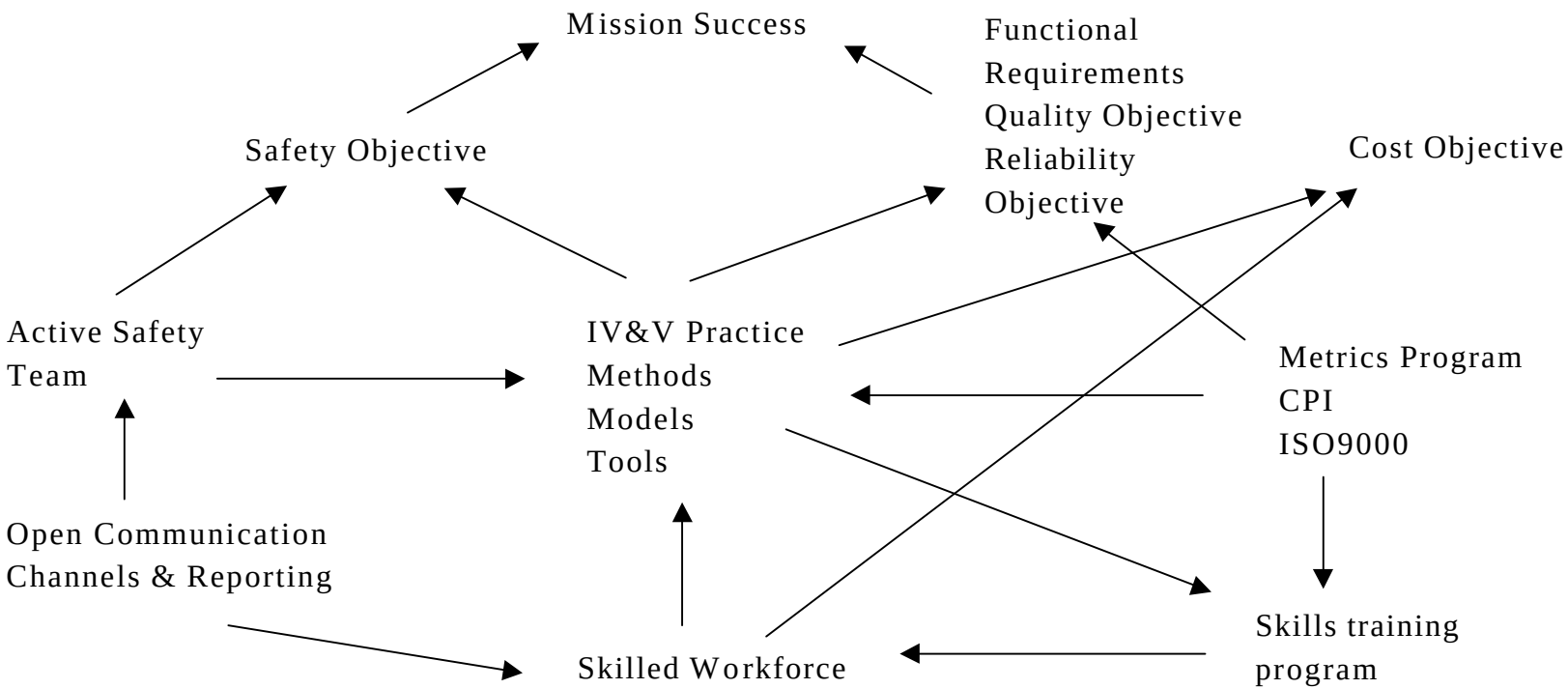
IV&V Developed Tools	Issue Tracking (PITS)	Risk Management (RMS)	Req. Analysis (ARDB)	Test Management (TMDB)	IV&V Effort (IVVEE)	Req. Measurement (ARM)	Code and I/F Analysis (SIAT)	IV&V Shuttle Toolset	MATRIXx Utilities	Criticality/Risk Analysis (CARA)	IV&V Config. Testbed	Flight Simulators	EWatch Dist. System Analyzer
AATT	X	X		X									
CLCS	X	X			X	X				X			
EOS	X	X	X	X						X	X		
ISS	X	X	X		X		X		X	X	X	X	
NOAA				X									
Shuttle	X				X		X	X		X		X	
X33	X	X		X	X				X	X	X	X	
Other	X	X		X	X	X	X			X		X	

- Information Management
- Analytical

Process Definition Document
Metric Identification



Influence Diagram of Cause-Effect Graphing



NASA IV&V
BSC Measures

BSC Financial Measures	
Leading Indicators	IT - NPV & IRR
Lagging Indicators	ROI, ROM, ROA Cost reductions Value added

BSC Customer Measures	
Leading Indicators	Quality of service IV&V Responsiveness Surveys Benefit expectations
Lagging Indicators	# of Internal and External IV&V contracts (cross service)

NASA IV&V Strategic Goal Measures	
Safety	In-Flight Anomalies Post release defects
Quality	Severity 1 Defects
Reliability	MTTF,MTTR, MTBF
Performance	IT - Lag time
Cost	ROI, ROM, ROA Cost reductions Value added-NPV

BSC Infrastructure Measures	
Leading Indicators	Information dissemination lag time Communications effectiveness R&D project penetration % R&D technology transfer ratio
Lagging Indicators	Survey of customer satisfaction Appropriate staffing levels Strategic job coverage Strategic information availability Staff climate survey

BSC Core Process Measures	
Leading Indicators	Activity based costing Issue resolution cost Technology utility IT cost-benefit profiles
Lagging Indicators	TTEF–NPV+ IRR

Dr. Nancy Eickelmann - NASA IV&V Facility

Mapping IV&V Activities to BSC

Internal Business Processes	Objectives	Measures	Targets	Initiatives
	Minimize test effort	# test cases test pass/fail rates	Minimum set	Test Effectiveness
	Track issues to disposition	# Total Issues Open/Closed	Desired rate of closure	Web-based tracking system

IV&V Developed Tools	Issue Tracking (PITS)	Risk Management (RMS)	Req. Analysis (ARDB)	Test Management (TMDB)	IV&V Effort (IVVEE)	Req. Measurement (ARM)	Code and I/F Analysis (SIAT)	IV&V Shuttle Toolset	MATRIXx Utilities	Criticality/Risk Analysis (CARA)	IV&V Config. Testbed	Flight Simulators	EWatch Dist. System Analyzer
AATT	X	X		X									
CLCS	X	X			X	X				X			
EOS	X	X	X	X						X	X		
ISS	X	X	X		X		X		X	X	X	X	
NOAA				X									
Shuttle	X				X		X	X		X		X	
X33	X	X		X	X				X	X	X	X	
Other	X	X		X	X	X	X			X		X	

- Information Management
- Analytical

Learning and Growth	Objectives	Measures	Targets	Initiatives
	Effective Communications	Reporting and response times	X+2(Y+!)	Web-based issue tracking
	World Class Research Staff	PhDs Publications	Top graduates Best in Class	Recruiting and Retention
	State of the Art Skills	Ongoing Training and certificates	100%	NASA HQ training program

Balanced Scorecard Perspectives

	Objectives	Measures	Targets	Initiatives
Customers (Internal External)	No Losses	# Severity 1 &2	Remove < FRR	Formal Methods
	Reduce Risk	# IFA's	No Severity 1	Risk Management
	Manage Risk	Fault tolerance	Performance	Risk Mitigation

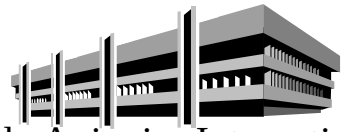
	Objectives	Measures	Targets	Initiatives
Financial	High ROI	Loss of Shuttle	> 100:1	ROI-CSIP
	Reduce Costs	Program costs	10% per FY	

	Objectives	Measures	Targets	Initiatives
Internal Business Processes	Minimize test effort	# test cases	Maximum Coverage	Test Effectiveness
	Track issues to disposition	# Total Issues Open/Closed	Rate of closure	Web-based tracking system

	Objectives	Measures	Targets	Initiatives
Learning and Growth	Skilled workforce	Degreed professionals	100%	Retention Strategies
	World Class Research Staff	PhDs Publications	Top graduates Best in Class	Recruiting and Retention
	State of the Art Skills	Ongoing Training and certificates	100%	NASA HQ training program

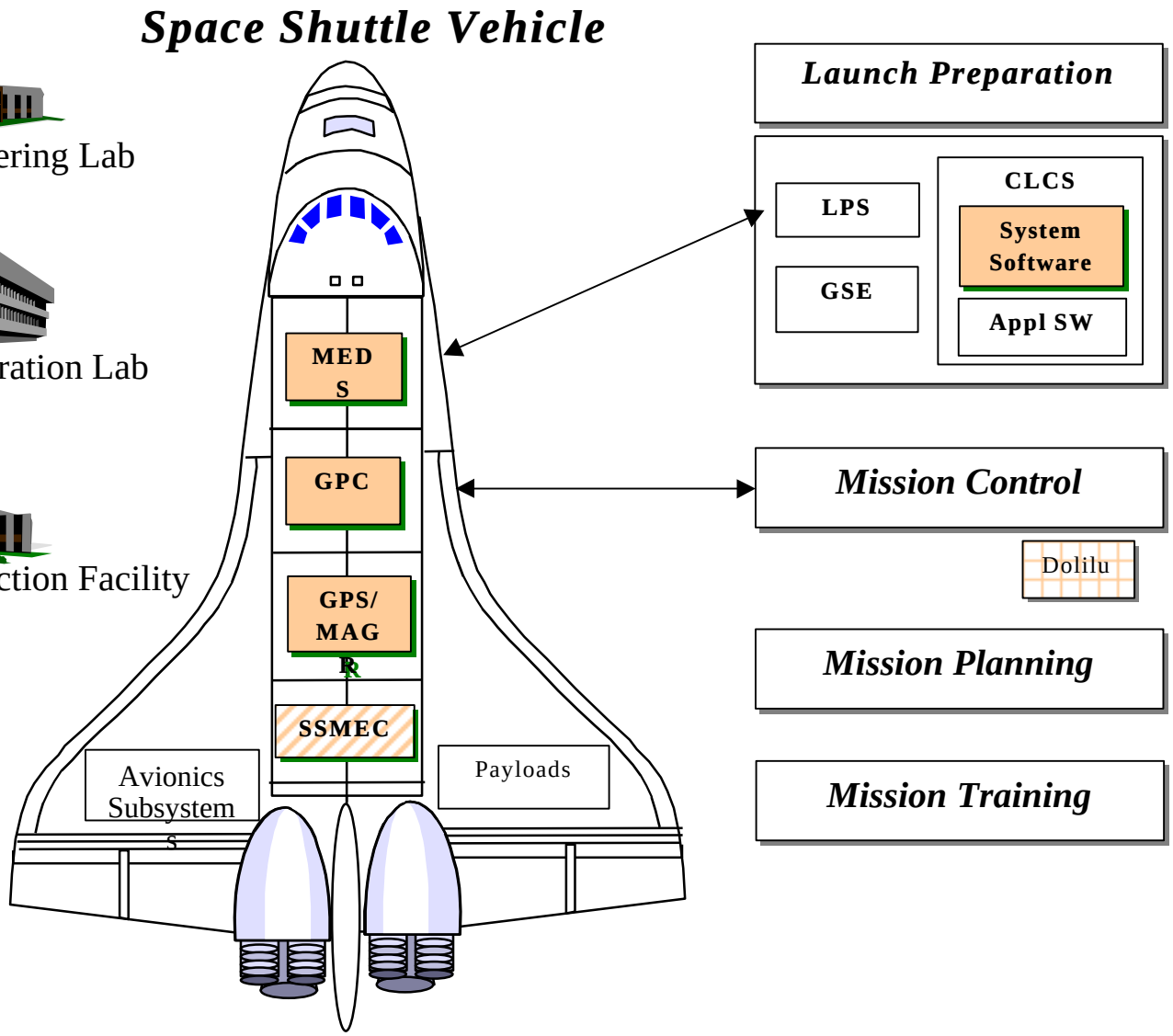
Space Shuttle Program Software Elements

JSC Avionics Engineering Lab



Shuttle Avionics Integration Lab

SW Development/ Production Facility



IV&V Activities

- Past IV&V
- No IV&V
- IV&V
- Partial IV&V

Summary

- BSC is a ***strategic measurement framework*** that NASA IV&V is applying to track strategic goals relative to operational activities for the Space Shuttle program.
- BSC ***measures economic value*** of operational activities relative to strategic goals
- BSC ***minimizes metrics required*** to evaluate product quality, safety, reliability, and cost.

Software Measurement Frameworks to Assess the Value of Independent Verification and Validation

Dr. Nancy S. Eickelmann
WVU/Software Research Laboratory
NASA Independent Verification and Validation Facility
Phone: (304) 367-8444
E-mail: Nancy.Eickelmann@ivv.nasa.gov
Website: http://research.ivv.nasa.gov/~ike
December 1999

Motivation
Research Roadblocks and Directions

ROI Directions

Clinger-Cohen Act of 1996 and the Government Performance and Results Act (GPRA) mandates the use of …

> "a framework of modern technology management based on practices followed by leading private-sector and public-sector organizations that have successfully used technology to dramatically improve performance and meet strategic goals." (Source GAO/NSIAD-98-181).

ROI Roadblocks

Four significant barriers to measuring financial performance related to information technologies included:

- Difficulty of measuring economic benefits
- Inability to determine returns
- Lack of good metrics
- Incomplete records/accounting of investments

ROI Research Initiative applies the strategic measurement framework developed at the Harvard Business School known as the Balanced Score Card BSC.

NRC Based Financial Measures
1999 Space Shuttle ROI

Severity 1 & 2 Issue Tracking Report flight software.

Shuttle Cost Risks (Severity 1)

- Loss of life = Priceless
- Loss of Shuttle = *$2Billion**
- Loss of hardware = $?$
- Causes a Stand-down = $?$

Shuttle Cost Risks (Severity 2)

- Mission objectives not met
- Manifest redeployment based on Standard Flight Cost $?$

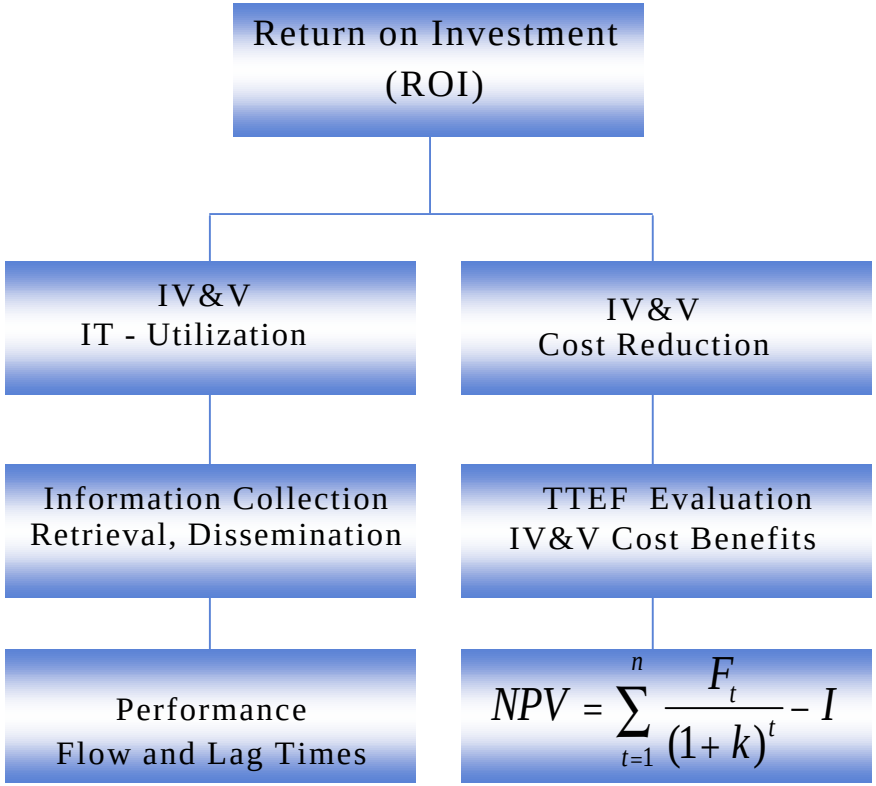
FY98 Shuttle IV&V Cost = $3.037 Million

ROI = 863:1

1) FY98 Shuttle IV&V costs: $3.037M
2) Average Shuttle mission cost: $409M for FY98 (was $590M in FY93); taken from FY2000 budget package chart

*Amount shown is in 1993 dollars and was reported by the National Research Council

NASA IV&V Linking from Financial - Process - Infrastructure Measures



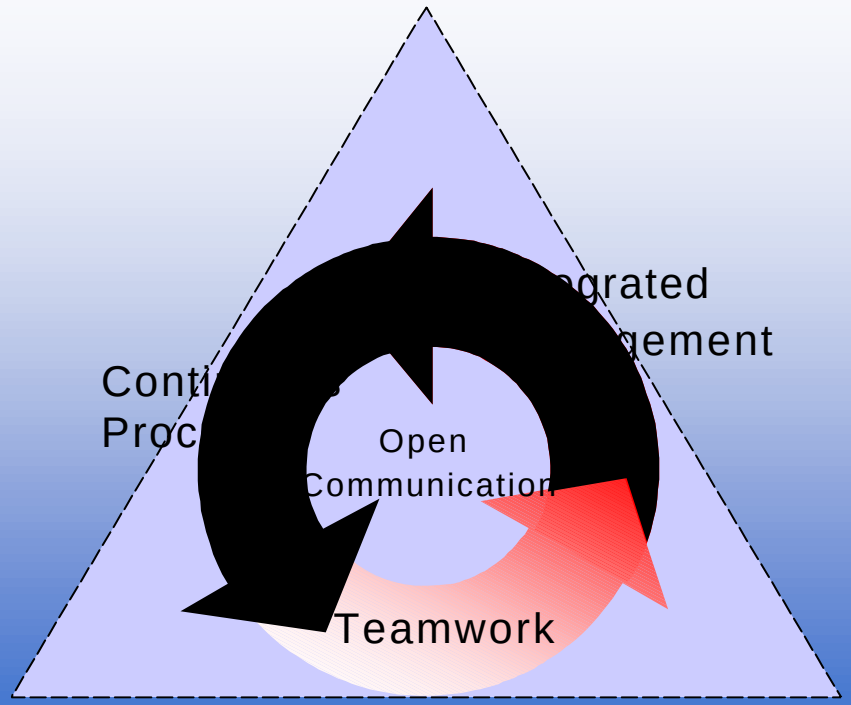
The BSC strategic measurement framework links cost benefit analysis to evaluating the effectiveness of technologies in achieving strategic goals.

Internal Process Measures for IV&V as Risk Management

IV&V brings a unique contribution with respect to Risk Management.

The core principle of risk management is ***open communication*** that enables:

- free-flowing information between all project and organizational levels
- facilitating formal, informal, impromptu communication
- applying consensus and valuing individual contributions

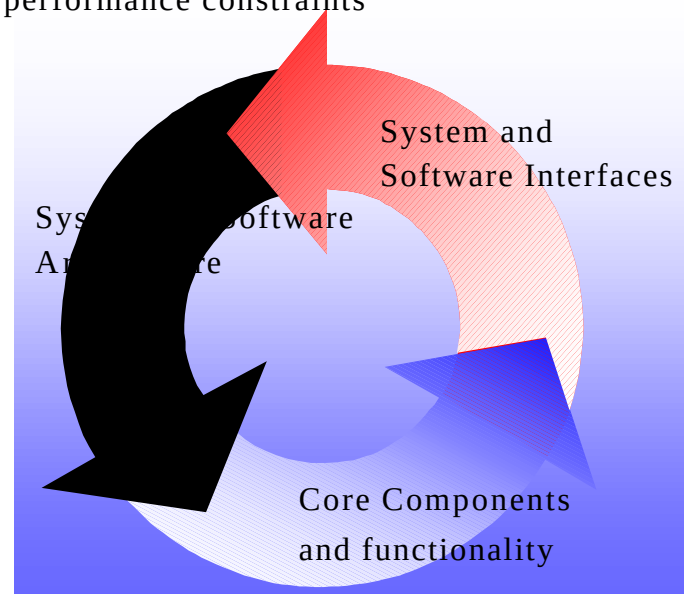


Source: Continuous Risk Management Guidebook

Space Shuttle Product Lines Process Reduces Risk

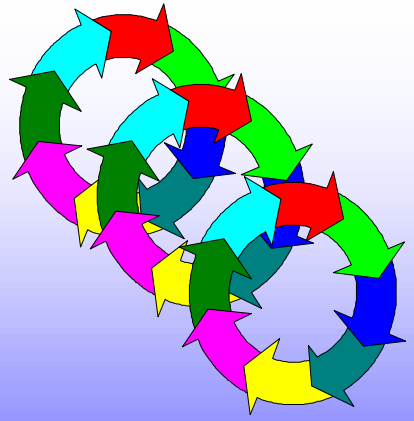
Domain Engineering for the Space Shuttle Product Line

•Requirements Engineering for every OI
•Architectural Integrity through negotiated modifications
•Maintained extensibility and modifiability of design through restriction of violations of architectural structural, functional, and performance constraints



Application Engineering

•Reuse of core assets for every OI
•Architectural Integrity through modularity and encapsulation
•Maintained extensibility and modifiability of design



Multiple Operational Increments are developed using a core set of software CSCIs.

The core functionality is modified to accommodate specific mission requirements relative to launch and load.

A primary product line benefit is a result of test environment reuse. Test suites are regressed and reused and test plans are readily extensible to V&V system and software modifications to core assets.

Dr. Nancy Eickelmann - NASA IV&V Facility

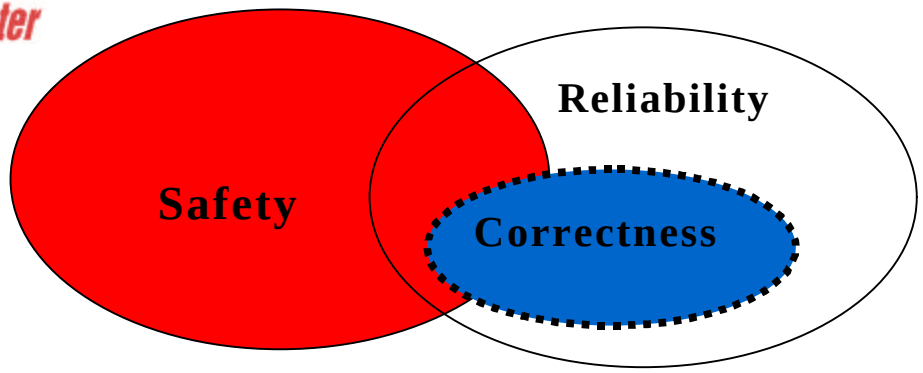
IV&V Challenges

Challenges

- Software Modularity
- COTS
- Program Mngt
- Software Engineering
- SW Reuse
- Maintainability
- Fail Safe
- Risk Control

- Technology Transfer
- Configuration Management
- Process IV&V across Series
- Validation of Flight Software
- Configuration Management for all developed elements
- Establish Program-wide Standards

Customer Themes for Mission Success
Safety-Reliability-Quality



- The relationships among these customer themes are significant as they are not independent of one another and therefore must be analyzed based on their degree of covariance and interaction.
- Safety requires that unsafe states cannot be entered from any point of function of the system.
- It is possible for the systems to function reliably that is without failure and still enter unsafe states of operation.
- A system can be completely correct and defect free and still enter unsafe states.

Example Customer Measures ROI

Objectives	Sample Measures
Customer partnership and involvement	*percent projects using integrated project teams* *percent joint IT customer/supplier service lever agreements*
Customer satisfaction	*percent customers satisfied with IT product delivery* *percent customers satisfied with IT problem resolution* *percent customers satisfied with IT maintenance and support* *percent customers satisfied with IT training* *percent products launched on time* *percent service-level agreements met*
Business process support	*percent IT solutions supporting process improvement projects* *percent users covered by training to use new IT solutions* *percent new users able to use applications unaided after initial training*

*GAO /AIMD-97-163

Example Financial Measures and Strategic IT Goals

Objectives	Sample Measures
Enterprise mission goals	*percent mission improvements (cost, time, quality, quantity) attributable to IT solutions and services* *percent planned IT benefits projected v. realized*
Portfolio analysis and management	*percent IT portfolio reviewed and disposed* *percent old applications retired* *percent applications retirement plan achieved* *percent reusability of core application modules* *percent new IT investment v. total IT spending*
Financial and investment performance	*percent and cost of services provided in-house v. industry standard* *IT budget as percent of operational budger and compared to industry average* *pet present value, internal rate of return, return on investment, return on net assets*
IT resource usage	*percent consolidated/shared resources across units* *percent cross-unit shared databases and applications* *percent hardware/software with interoperability capabilities*

*GAO /AIMD-97-163

Example Business Process Measures - ROI

Objectives	Sample Measures
Applications development and maintenance	*number of function points delivered per labor hour* *number of defects per 100 function points at user acceptance* *number of critical defects per 100 function points in production* *percent decrease in application software failures, problems* *mean time to resolve critical defects* *cycle time for development*
Project performance	*percent projects on time, on budget* *percent projects meeting functionality requirements* *percent projects using standard methodology for systems analysis and design*
Infrastructure availability	*percent computer availability* *percent communications availability* *percent applications availability* *on-line system availability*
Enterprise architecture standards compliance	*number of variations from standards detected by review and audit per year* *percent increase in systems using architecture* *percent staff trained in relevant standards*

*GAO /AIMD-97-163

Example Infrastructure Measures ROI

Objectives	Sample Measures
Workforce competency and development	*percent staff trained in use of new technologies and techniques* *percent staff professionally certified* *percent IT management staff trained in management skills* *percent IT budget devoted to training and staff development*
Advanced technology use	*percent employees skilled in advanced technology applications* *number of dollars available to support advanced technology skill development*
Methodology currency	*currency of application development methods in used* *percent employees skilled in advanced application development methods* *percent projects developed using recognized methods and tools*
Employee satisfaction and retention	*percent employee satisfaction with the capability of the existing technical and operating environment to support mission* *percent employee turnover by function*

*GAO /AIMD-97-163

IT - BSC Measures

1. Identify achievable IT goals by NASA Center
2. Classify goal structure
   - Short-term
   - Medium-term
   - Long-term
3. Categorize the impact of proposed IT strategies
   - Very high impact
   - High impact
   - Low impact
   - Very low impact
4. Estimate probability of IT success (0 to 1)
   - 10% up to 100%
5. Identify appropriate IT measures for each NASA Center based on ***their*** strategic goals and mission

TTEF

test effectiveness based on a relative test detection rate and test failure rate given the optimistic or pessimistic inaccuracy of a specific test strategy

test productivity based on the rate of execution and verification of successful test inputs with a test output that results in an observed failure or detection

test schedule compression based on full life cycle analysis

It then relates these factors in a properly scoped cost benefit analysis methodology applying ROI, NPV and IRR.

TTEF - NPV - IRR

The TTEF evaluates the degree of achievement of the primary goals including test effectiveness, test productivity, and test schedule compression using a comparative analysis...

$$NPV = \sum_{t=1}^{n} \frac{F_t}{(1+k)^t} - I$$

...and is designed to capture the principle benefits associated with a technology as well as the complete cost structure that has been extended to use NPV and IRR.

TTEF Example

- A specific example is chosen that uses commercially available test tools and object-oriented modeling and specification documentation tools.

> "The availability of these already integrated tools enabled the team to measure the costs and benefits of reusing object models for testing without having to deal with the sometimes confusing cost and productivity factors..." Communications of the ACM, September 1994/Vol..37, No. 9.

- A comparison is conducted between current practice for evaluating technologies and our Test Technology Evaluation Framework (TTEF).

Naïve
Cost Benefit Analysis

$$ROI = \frac{152 hours}{0.5 hours}$$

- Industry average test case productivity (Capers Jones) 20 to 300 test cases per month
- COCOMO Man Month is equal to 152 hours
- Cost of model development 24 minutes
- Test benefit with automated OMT 151.5 hours saved

Cost Shifting Versus Cost Savings

StP/OMT Claims

- Cost savings (benefit) of one person month.
- Cost of model production 24 minutes.
- ROI of 304:1

Reality

- Cost shifting to model development and evaluation of model fidelity.
- Cost shifting to tools, training, updates, reuse of models.
- Test effectiveness was not evaluated...

Initial Costs

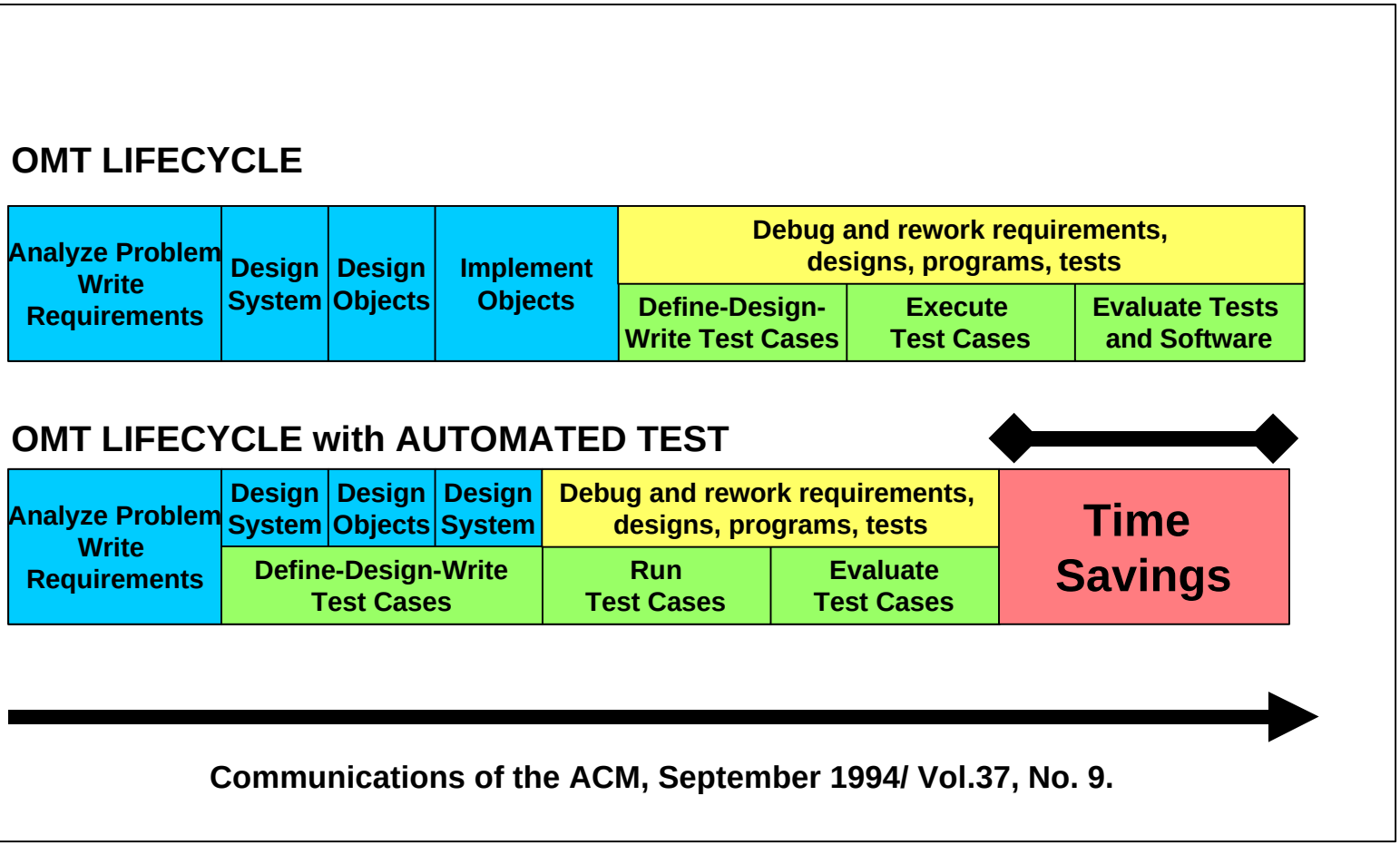
Poston's Case study used OMT technologies and evaluated automated specification-based testing technologies:

- StP/OMT (Integrated Development Environments, IDE) - Development tool based on James Rambaugh's object modeling technology was used to develop the specification Single license cost: $6,500. Plus one year maintenance cost $1,300.

- StP/T (Integrated Development Environments, IDE) - Test case generation tool based on OMT specifications was used to generate the test cases from the specification Single license cost: $10,530 plus $1,580 maintenance.

- XRUNNER - Software test tool test execution, capture replay was used to run the test cases. Single license costs: $9,503 plus $2,100 annual maintenance fee.

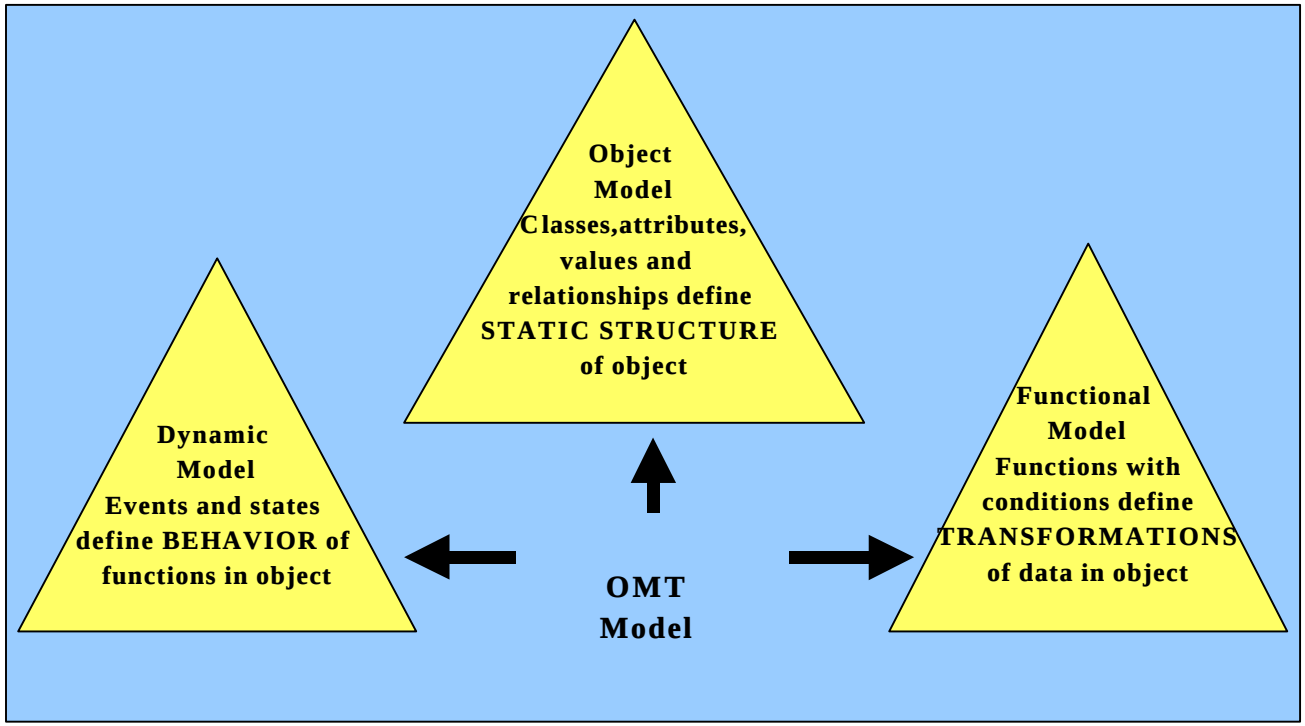
OMT Lifecycle Savings



Communications of the ACM, September 1994/ Vol.37, No. 9.

Automated OMT and Test Tools

Rumbaugh's OMT uses three separate views to graphically describe the system. OMT provides support to create models to represent system objects, dynamics and functional aspects. Evaluate benefits w.r.t. the technology window...



Scoping Test Cost Benefits Using TTEF

The software test process and test process automation include six canonical functions that are used to evaluate test technology impacts:

- test execution
- test development
- test failure analysis
- test measurement
- test management
- test planning

The software test technology impacts are evaluated for impacts in the following categories in the context of test functionality and lifecycle context:

- test effectiveness
- test productivity
- test schedule compression

Test Automation Cost

• The costs associated with adopting Poston's technology (or similar automated testing tools) are associated with the paradigm shift to OO technology and anticipated reuse benefits w.r.t. the technology window:

- Equipment purchase StP/OMT, StP/T, Xrunner*
- Reuse libraries*
- Training of personnel*
- OMT Specification development and analysis
- Verification of OMT model fidelity
- Maintenance and archiving of reusable assets
- Software updates

** Initial costs in adopting test automation technologies.*

Test Automation Benefits

• The benefits associated with adopting Poston's technology (or similar automated testing tools) are associated with the paradigm shift to OO technology and anticipated reuse benefits w.r.t. the technology window:

- Test personnel time savings
- Automated test case generation and execution
- Test schedule savings
- Design model reuse*
- Test specification reuse*

** Benefits may accrue to future projects based on reuse savings.*

NPV Calculations

$$NPV = \sum_{t=1}^{1} \left(\frac{152.5\,\text{hours} \times \text{labor rate}_1}{(1+k)^1} \right) - 30\,\text{minutes} \times \text{labor rate}$$

Naïve NPV = $2,763.00

- Assumes one time period
- Applies only labor savings costs
- Labor rate - $20.00/hour
- Discount rate 10%

Full Costs NPV = $(20,854.00)

Assumes one time period

Applies capital ($20,036) and labor

Labor rate - $20.00/hour

Discount rate 10%

NPV Calculations

TTEF NPV = $1819.00 (in year one with reuse over 10 projects)

- Automated specification based testing initial cost estimate - $20,036.00 per seat
- Annual maintenance costs - $ 3,680.00
- Additional savings through reuse strategies - $2772.00 per additional program (example uses 10 programs)
- Discount rate 10%
- *Improved test effectiveness savings - difference between old rate and new, not calculated**

Cost-Benefit Analysis

- A case study was performed to evaluate the measurement framework that provides guidelines for how the data will be interpreted, this work has been accepted for publication the results demonstrate:
- The value of the technology was significantly overstated when the scope of the cost-benefit analysis was underspecified, ROI 304:1.
- The value of the technology was understated when the scope was applied correctly but the underlying reuse process was not taken into account, ROI 0.1517:1.
- Finally when both the scope of the cost-benefit analysis and the underlying process and reuse paradigm are incorporated into the analysis meaningful quantification of the value of the technology results.