V & V of Flight-Critical Systems

Guillaume Brat, NASA ARC

NASA Aviation Safety Program

NASA Ames Research Center, Code TI

Why?

- By 2025, U.S. air traffic is predicted to increase 2 to 3 times.

NextGen and JPDO

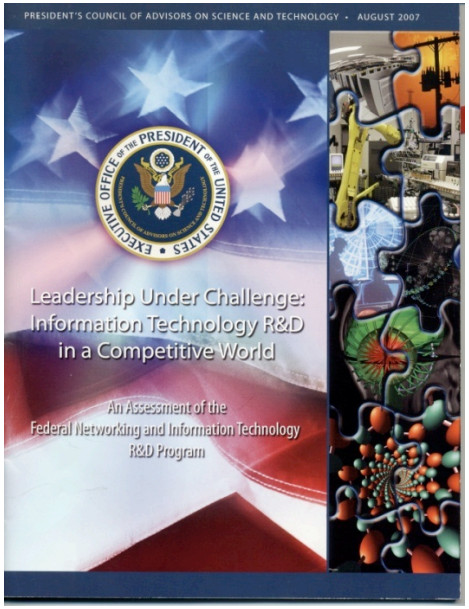
- The Next Generation Air Transportation System (NextGen) is the solution to safely and efficiently manage this growth and allow new aircraft classes and operational concepts.

- The Joint Planning and Development Office (JPDO), coordinating the Departments of Transportation, Defense, Homeland Security, Commerce, FAA, NASA, and the White House Office of Science and Technology Policy, is responsible for managing a public/private partnership to bring NextGen online.

NextGen Complex Technological Developments

From… →	To…
Ground-Based Navigation and Surveillance	Satellite-Based Navigation and Surveillance
Voice Radio Control	Digital Data Exchange
Disconnected Information Systems	Net-Centric Information Access
Human-Centric Air Traffic Control	Automation Assisted Air Traffic Management
Fragmented Weather Forecasting	Probabilistic Weather Decision Tools
Visibility Limited Airfield Parameters	Equivalent Visual Operations
Forensic Safety System	Prognostic Safety System
Inefficient Security Screening	Integrated Security Risk Management
Current Aircraft Environmental Footprint	Reduced Aircraft Environmental Footprint

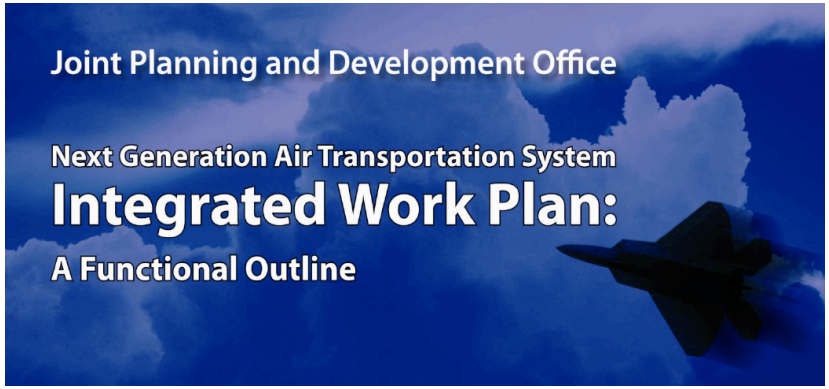
JPDO Identified Critical Gap in V&V Methods



"Developers do not have effective ways to model and visualize software complexity, including the possible range of interactions, especially unexpected and anomalous behaviors that can occur among software and hardware components.

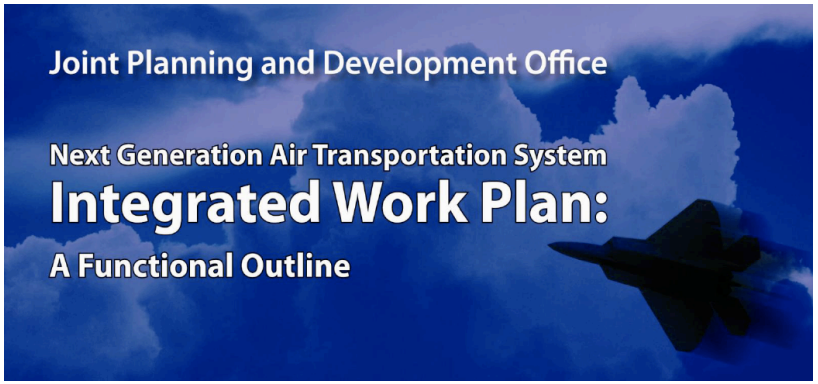
Developers also do not have time- or cost-effective ways to test, validate, and certify that software-based systems will perform reliability, securely, and safely as intended, particularly under attack or in partial failure."

The JPDO Drivers



R-1440 Applied Research on Complex Systems Validation and Verification
Applied research on the methods and algorithms to support the validation and verification of complex systems. Complex systems provide multiple functions that support many different operating models, environments and technologies and therefore require more advanced and integrated validation and verification methods and algorithms beyond those used for less complex systems. This research will support the development of complex systems, their risk assessment and eventual certification decisions.

The JPDO Drivers



EN-3050 Advanced Complex System Validation and Verification Methods

Description: Advanced tools and processes are developed to improve the verification and validation of complex systems and software. Improvements will focus on reducing the time and resources needed to conduct validation and verification as well as improving the quality of the results. ...

What's Changed?

- Long-lifetime of aviation systems, with many subsequent modifications
- An increasing reliance on software, without 'blueprints' and 'load analysis' equivalents during design
  - instead, just test-test-test of the 'finished' design
- 'Strong coupling' between components
  - Requires combined analysis of their interaction
- Significant changes to underlying concepts of operation
  - Can't assume the same human contribution to safety
  - Can't assume the same inherent structural contribution to safety
- Tighter margins and higher safety requirements

Impact: Cost, and Constraints on Innovation

Size Comparisons of Embedded Software

System	Lines of Code
Mars Reconnaissance Orbiter	545K
Orion Primary Flight Sys.	1.2M
F-22 Raptor	1.7M
Seawolf Submarine Combat System AN/BSY-2	3.6M
Boeing 777	4M
Boeing 787	6.5M
F-35 Joint Strike Fighter	5.7M
Typical GM car in 2010	100M

NASA Study

Flight Software Complexity, 4/23/2009

More S/W

Late-on Testing

Boehm, B. 1981 *Software Engineering Economics*, as cited in DAA, 2008

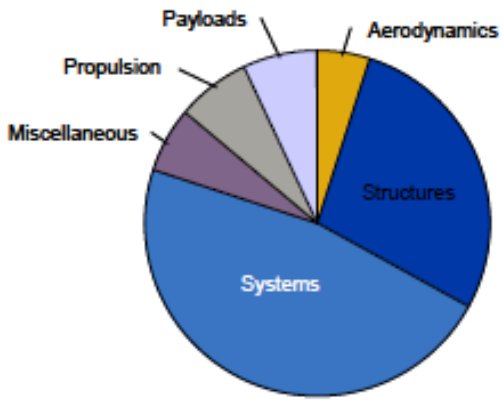


Fig. 1 - Typical Transport Aircraft Development Cost Distribution – Current Generation

Winter, D. (VP, Engineering & IT, Boeing PW)
House Committee on Science and Technology, July 31, 2008

And this is just s/w! Also need to consider human performance, airspace concepts of operation, and new technologies!

V&V cost and Certification

For FAA compliant DO-178B Level A software, the industry usually spends 7 times as much on verification (reviews, analysis, test). So that's about 12% for development and 88% for verification.

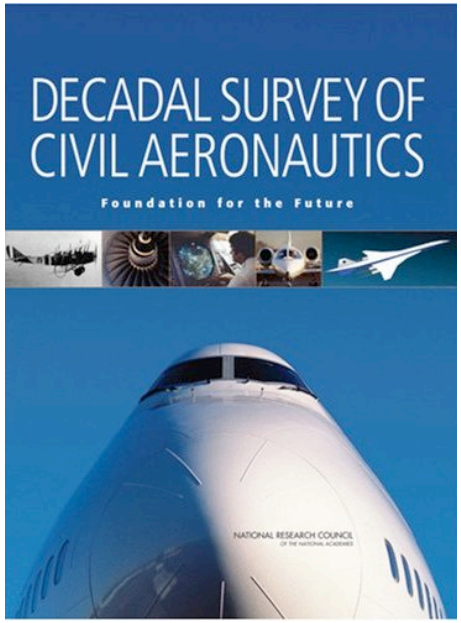
Level B reduces the verification cost by approximately 15%. The mix is then 25% development, 75% verification.

Randall Fulton
FAA Designated Engineering Representative
(private email to L. Markosian, July 2008)

Widely Recognized Concern

Fundamental research is needed to create the foundations for practical certification standards for new technologies

- methods and models are needed for assessing the safety and reliability of complex, large-scale, human-interactive, nondeterministic software intensive systems



NASA's Research Assessment of V&V for NextGen

- NASA Aeronautics' Aviation Safety Program is examining the research required to develop transformative safety V&V methods required to rigorously assure the safety of NextGen developments in a time- and cost-effective manner.
- NASA has completed an assessment of the most critical research activities required to develop these methods. The research activities are organized into four challenge areas.

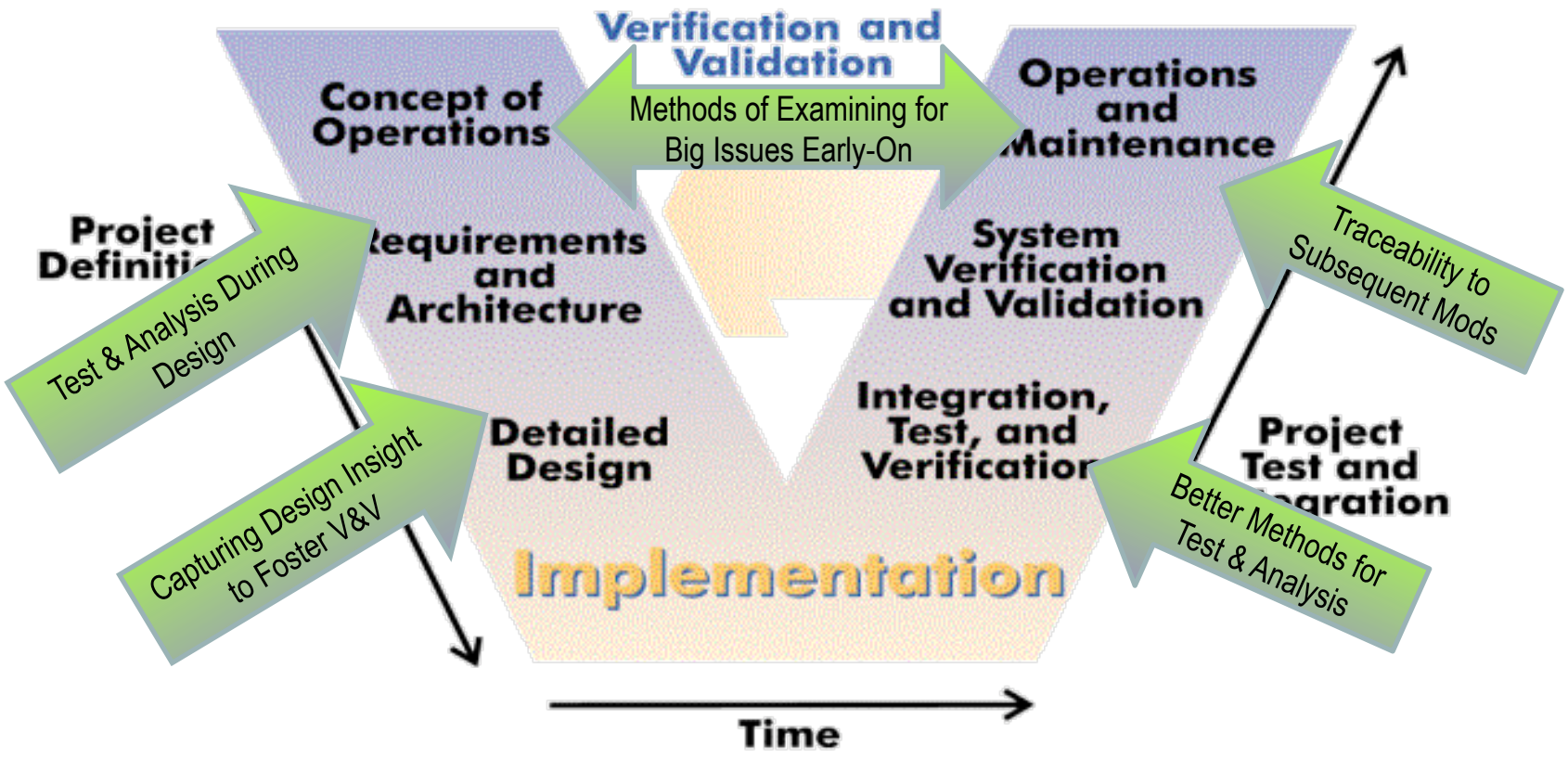
Summary of NASA VVFCS Effort To Date

- Planning effort underway conducted on ARRA funds
  - Document, "Validation and Verification for Flight Critical Systems – Assessment of Critical Research Activities", Nov. 2009:

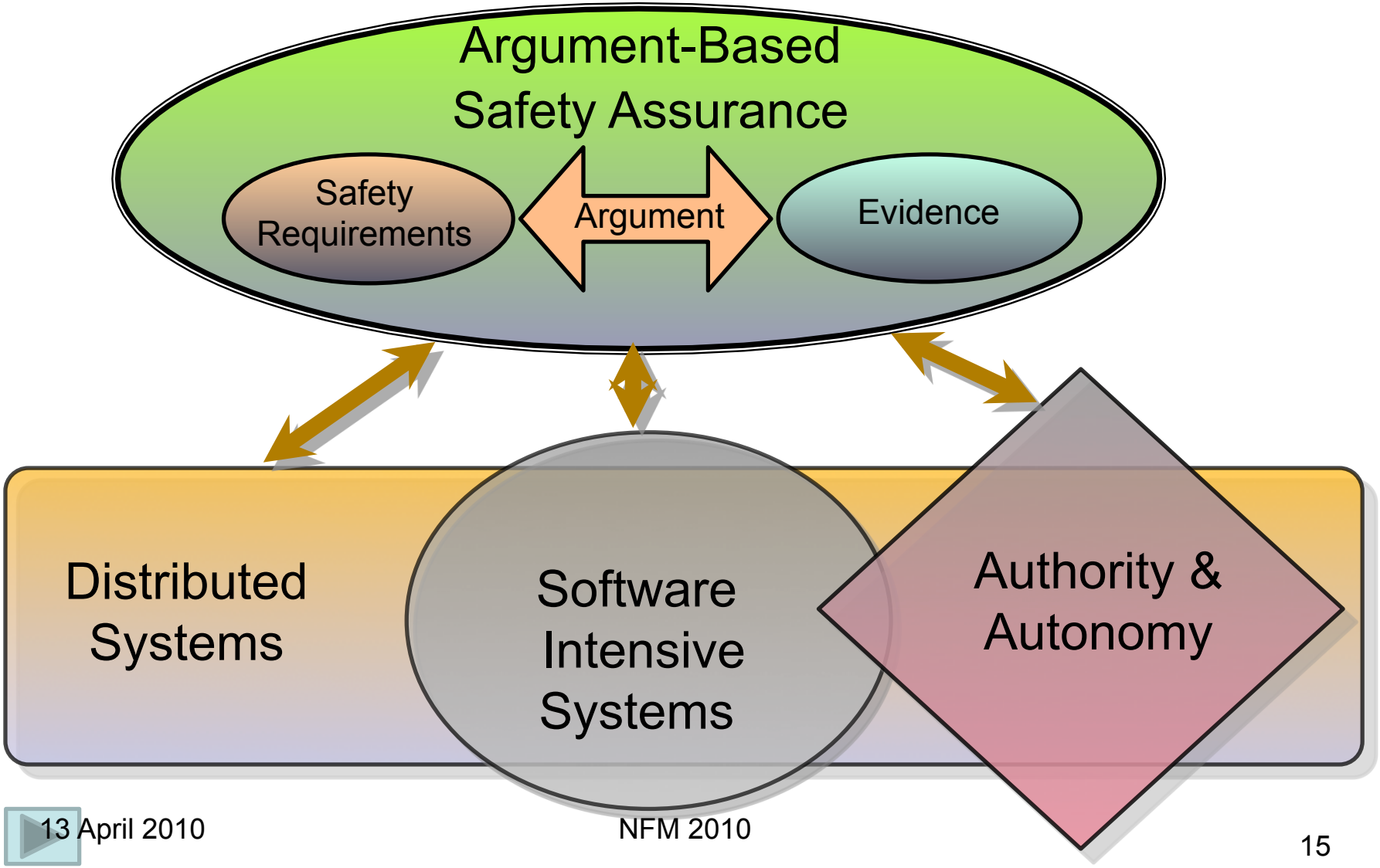
Development of verification and validation tools, methods and techniques that advance safety assurance and certification of complex, networked, distributed flight critical systems operating in the Next Generation Air Transportation System

- Objectives
  - Meet the JPDO's critical interagency needs associated with V&V research in support of NextGen transformation
  - Demonstrate advanced methods to answer relevant questions from aviation community
  - Reduce barriers to innovation associated with safety V&V
  - Develop V&V methods for safety throughout the entire life cycle

What We're Seeking



Four Challenge Areas



Research Area 1: Argument-based Safety Assurance

- NextGen changes the conventional boundaries and layers -- and, consequently, safety assurance
- We envision a framework that explicitly captures:
  - safety goals/claims/objectives, especially for new functions
  - evidence that goals have been met
  - arguments linking evidence to goals
    - assumptions, justifications, and other context
- This framework should support design and integration
  - Used to trace conflicts or gaps in assumptions and evidence of combined functioning during component integration
- This framework should support the entire lifecycle
  - Qualitative early in design
  - Endures beyond first design/implementation, to support modification and integration

Research Area 2: Flight-critical Distributed Systems

- Aviation system is a distributed network of distributed systems
- Multiple levels of distribution exist
  - Multi-core processors (system on a chip)
  - Fault-tolerant mechanisms
  - Airspace concepts of operation: Airborne/Space-based/Ground-based
  - Human/Automation
- We envision methods for ensuring robust system performance at all levels of distribution:
  - Distributed across multiple architecture
  - Distributed across multiple air and ground elements
  - Interactions between components as intended
  - Robust to faults, failures and degradations

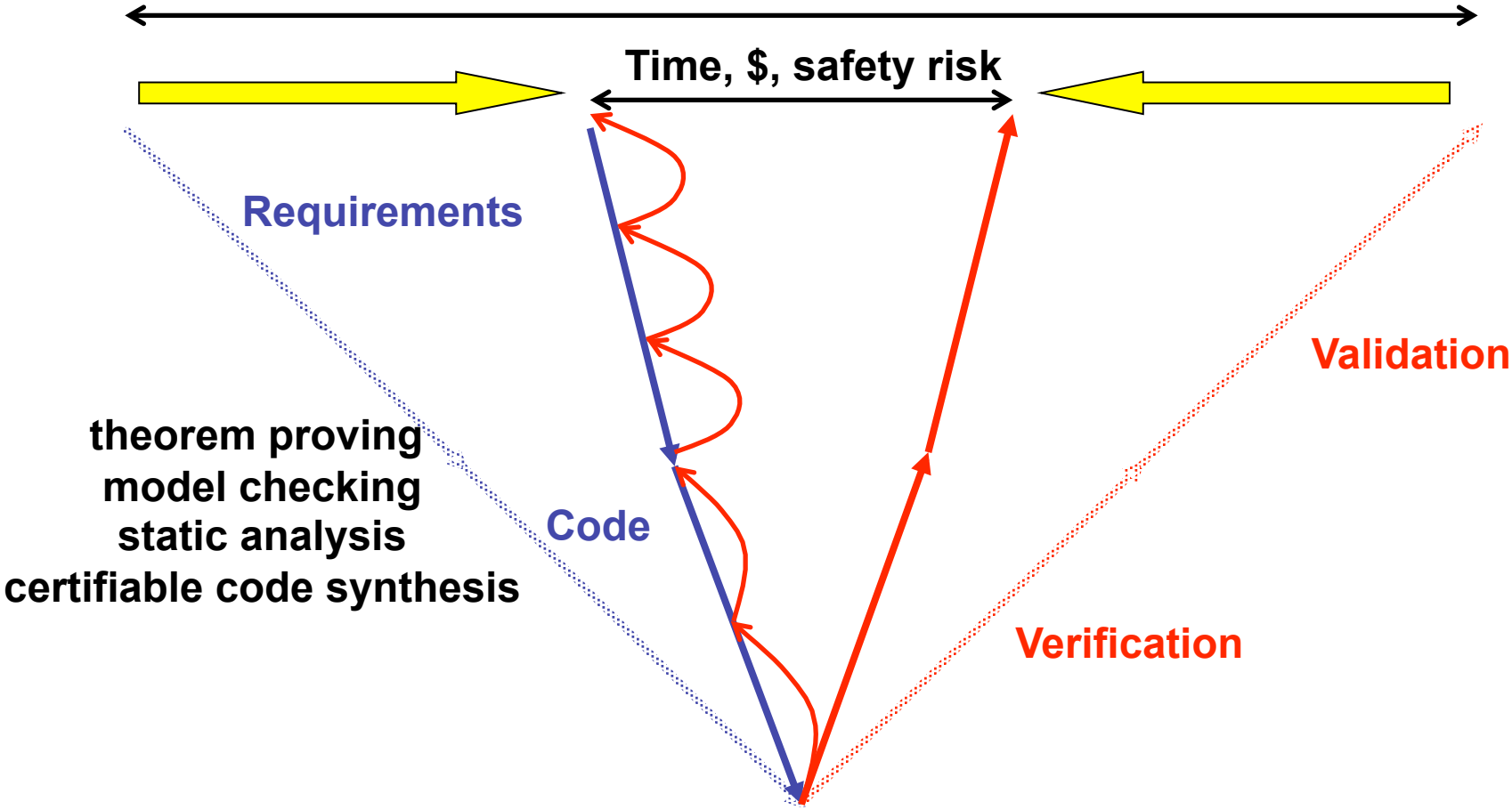
Research Area 3: Authority and Autonomy

- Authority requires both accountability and capability
  - Need authority aligned with autonomous capabilities
  - Need to avoid competing authorities
  - Need to avoid gaps in authority, maintain clearly who/what is in charge
- We envision methods for early-on assessment of 'big issues':
  - Is authority assigned properly?
  - Is authority assigned with correct assumptions regarding capabilities?
  - Are there conflicts or gaps?

Research Area 4: Software-Intensive Systems

- NextGen plans increase reliance on software-intensive systems in both ATC and aircraft systems
  - Software will interact with other software, systems, devices, sensors, and with people
- We envision methods for examining software-intensive systems
  - Appropriate extension of formal methods
  - Increasing capabilities for numerical calculation
  - Generalized capabilities for software testing throughout coding

V&V earlier in life cycle



What's new?

- Widespread use of formal methods
- Compositional verification,
  - especially in the context of heterogeneous systems
- Design for verification and early application of V&V in the life-cycle
- Combination of testing with formal methods and learning techniques
- Development of foundational libraries
- Support for safety cases

Who is involved?

- Experienced research groups in formal methods
  - LaRC: theorem proving + model checking
  - ARC: theorem proving, static analysis, model checking, advanced testing
  - Collaborations with formal method groups in academia and labs
- DFC: practical experience in avionics testing and simulation
- Access to researchers working towards NextGen
  - LaRC and ARC
- Space provided us with great experience V&V-ing unique complex software systems

In Conclusion: Planning Approach

Common Themes

- Make V & V Cost- and Time-Effective
- Support the Entire Lifecycle
- Consider Disturbances & Degradations
- Humans and Software Are Central

Challenge Areas

- Argument-based Safety Assurance
- Distributed Systems
- Autonomy and Authority
- Software-Intensive Systems

Common Test Cases Applied Throughout

- Vehicle System: Integrated Alerting and Notification
- Airspace

Progress

- Completed Research Assessment (Jul-Nov 2009)
- Coordinate planning with other government agencies
  - Held Interagency Coordination Meeting on Sept 7th, 2009
- Present assessment of critical research areas at Aviation Safety Technical Conference (Nov 18, 2009)
  - Near-term research activities (FY09 & FY10)
  - Present Research Assessment for long-term research
- Completed NRA Solicitation NNH09ZEA001N-VVFCS1.
  - Awards decided
  - SOW under negotiation.

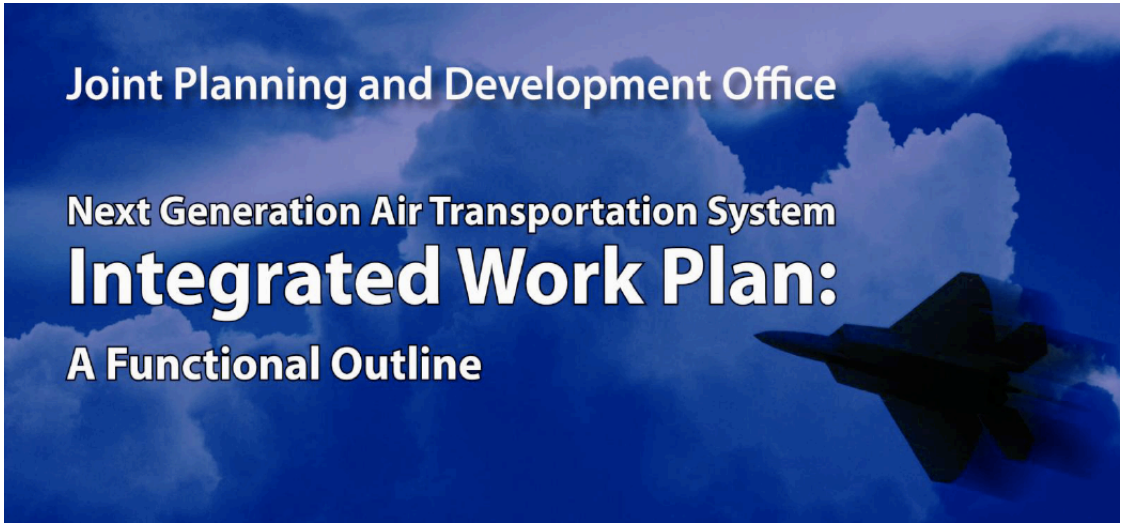
Points of Contact for V&V Assessment of Critical Research Areas

- Douglas Rohn, Acting Director, Aviation Safety Program, douglas.a.rohn@nasa.gov
- John Orme, Technical Integration Manager, Aviation Safety Program, john.s.orme@nasa.gov
- Sharon Graves, Acting Project Manager, sharon.s.graves@nasa.gov
- Guillaume Brat, Acting Project Scientist, guillaume.p.brat@nasa.gov
- Paul Miner, Technical POC for Distributed Systems, p.s.miner@nasa.gov
- Kelly Hayhurst, Technical POC for Safety Assurance, kelly.j.hayhurst@nasa.gov
- Mike Shafto, Technical POC for Authority and Autonomy, mike.shafto@nasa.gov
- Joe Coughlan, Technical POC for SW Intensive Systems, joseph.c.coughlan@nasa.gov
- Jim Disbrow, Technical POC for Testbench, james.d.disbrow@nasa.gov

Backup slides

Myth or Reality

D-2100: Complex System Validation and Verification Tools and Techniques



Is that V&V gap a myth or a reality?

Flight Software Incidents

In August 2005, a Malaysian Airlines Boeing 777 flying from Australia to Malaysia suddenly ascended 3,000 feet, with no input from the flight crew. The pilot disengaged the autopilot and pointed the nose down to avoid a stall, but the plane went into a steep dive. When he throttled back on the engines to reduce the speed, the plane arched into another climb. The flight crew eventually got things under control and returned their 177 passengers safely to Australia.

Wall Street Journal, 08/05

A faulty computer program recently installed on all 777s had provided incorrect information about the plane's speed and acceleration, confusing flight computers.

Complexity of ATM Software

"Software problems are delaying the completion of the world's most advanced air-traffic-control centre". The $570M center is said by National Air Traffic Services (NATS) to be "the largest and most advanced development of its kind in the world". The problems have delayed the opening by 15 months and "stem from the unusually high number of `bugs' which prime-contractor … is having to remove from the 1.82 million lines of software code at the heart of the system."

Peter Ladkin, April 1997

3300 functional requirements
Designed to work on 203 workstations
Defect rate: 15 bugs per 1000 LoC

Clearing 500 bugs per month
"We know where all the bugs are"

Peter Ladkin: "This last statement stands a very, very good chance of being false"

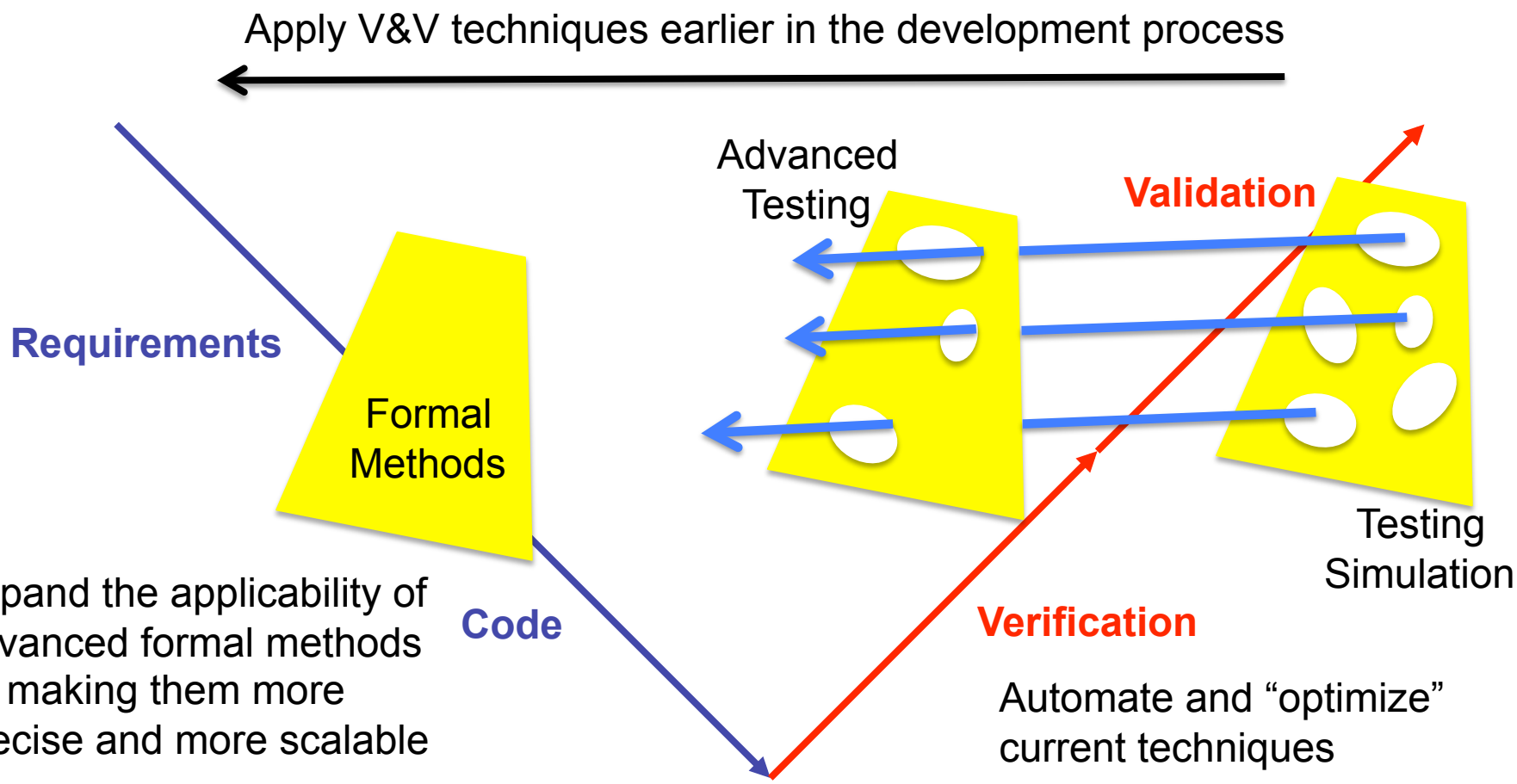
Lessons of AFTI-F16 flight tests

The criticality and number of anomalies discovered in flight and ground tests owing to design oversights are more significant than those anomalies caused by actual hardware failures or software errors

… qualification of such a complex system as this, to some given level of reliability, is difficult ... [because] the number of test conditions becomes so large that conventional testing methods would require a decade for completion …

Dale A. Mackall. Development and flight test experiences with a flight-crucial digital control system. NASA Technical Paper 2857, NASA Ames Research Center, Dryden Flight Research Facility, Edwards, CA, 1988.

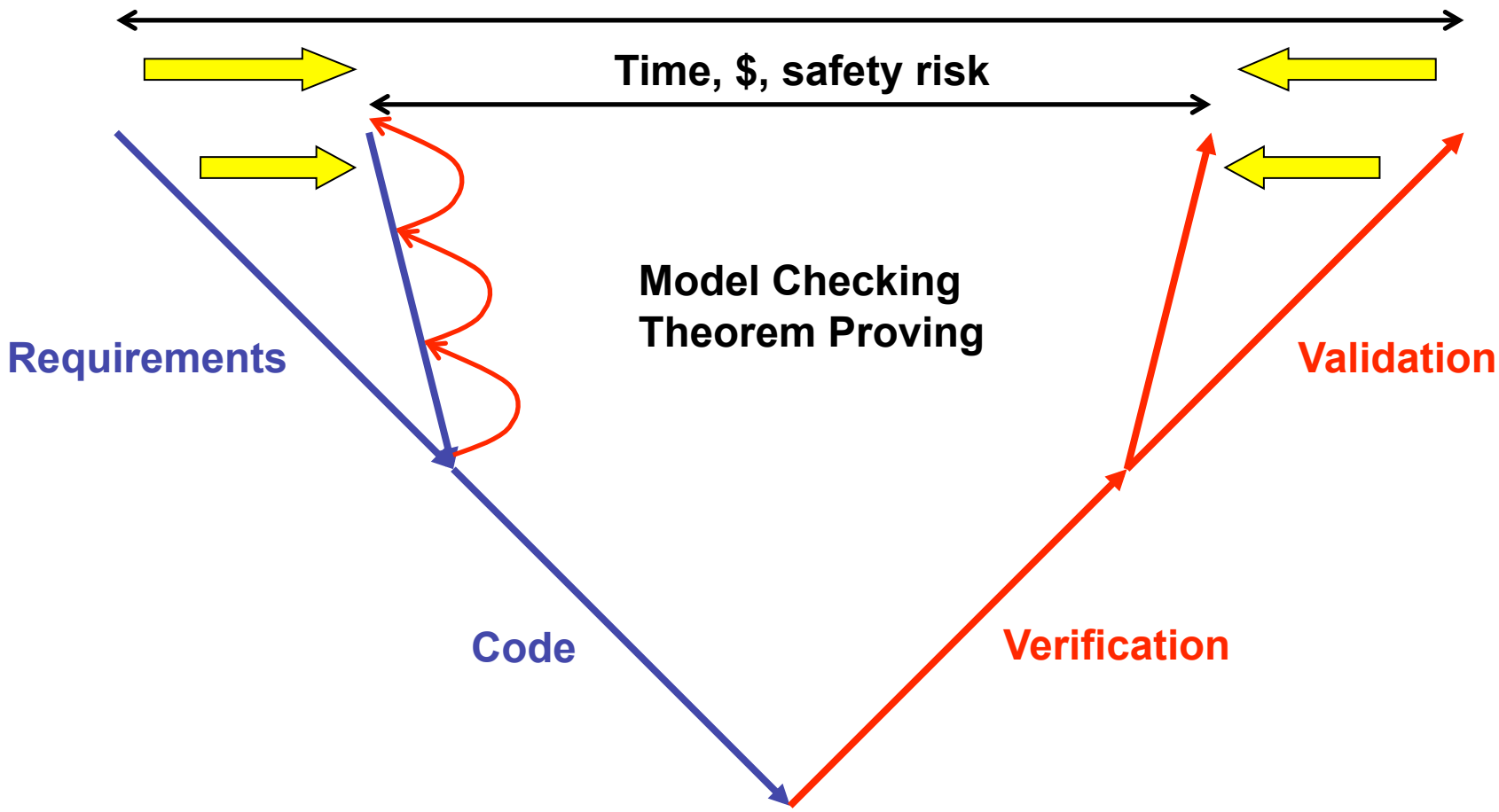
Research Thrust



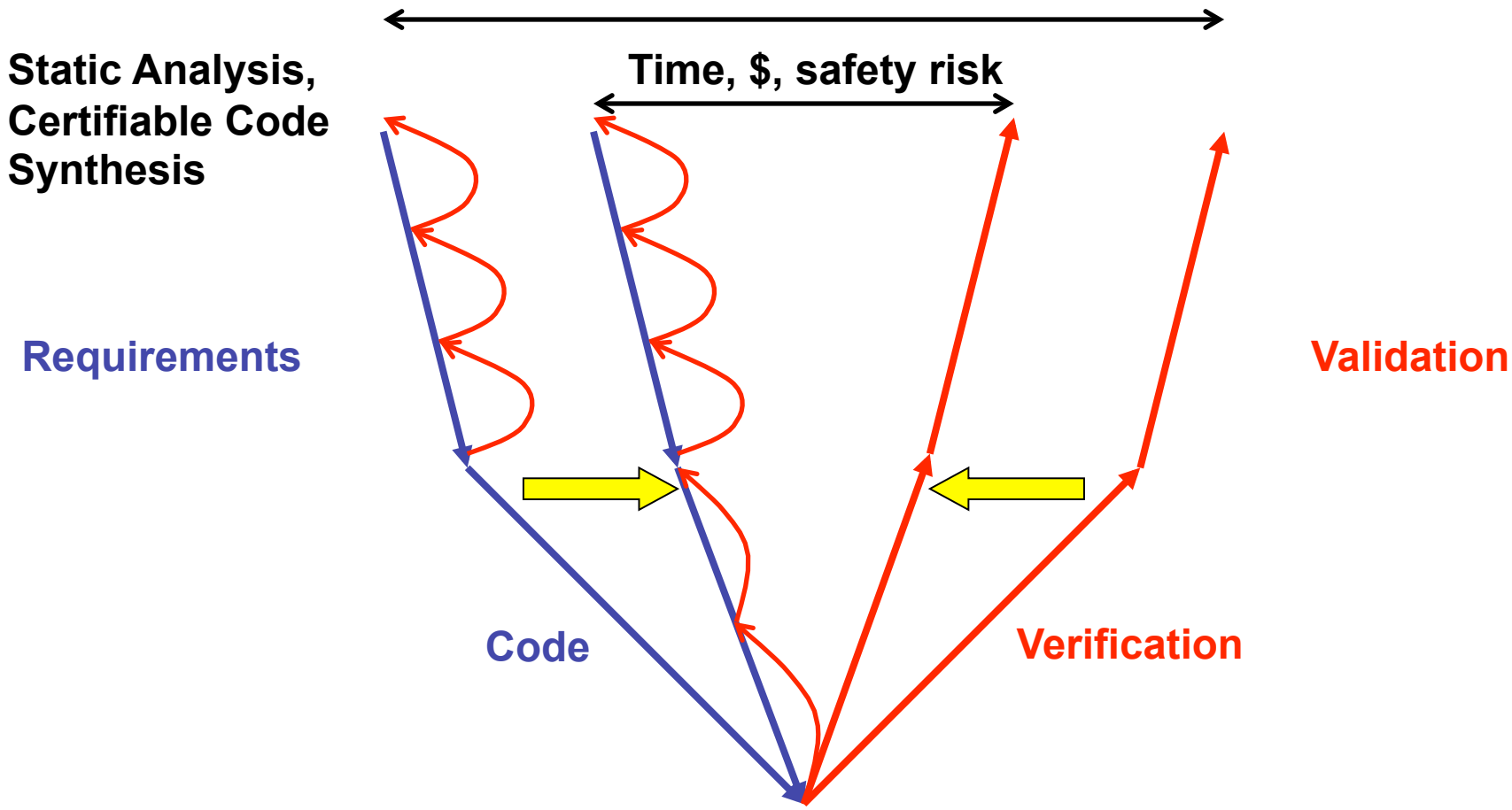
Expand the applicability of advanced formal methods by making them more precise and more scalable

Automate and "optimize" current techniques

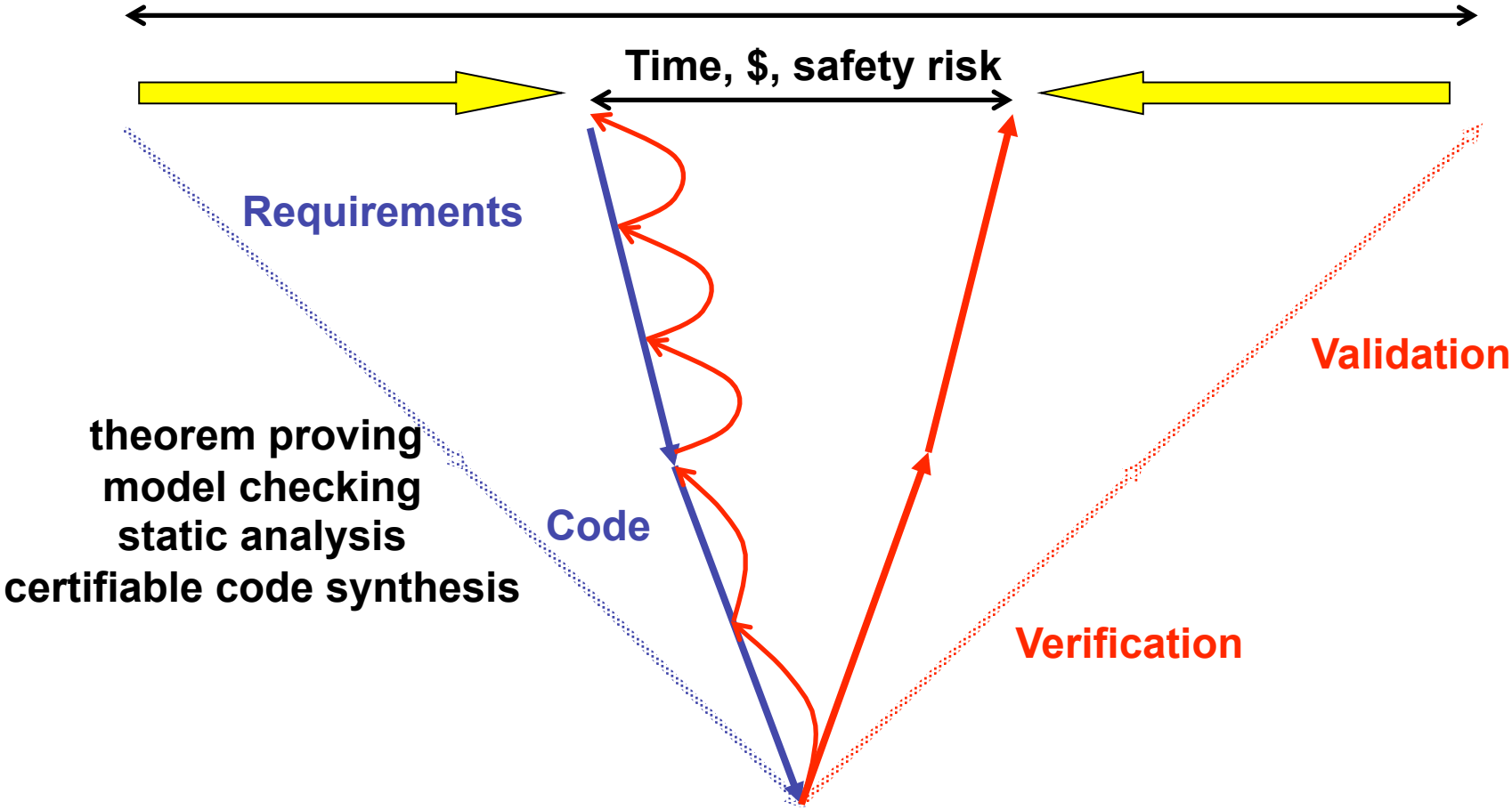
V&V earlier in life cycle



V&V earlier in life cycle



V&V earlier in life cycle



Scalability

- "Today … verification algorithms … suffer from well-known inherent complexity limitations when applied to large systems."
- "First avenue is to develop new abstraction techniques …"
- "Second avenue … involves moving from monolithic verification to compositional techniques."

Joseph Sifakis
2007 Turing award winner
(with E. Clark and A. Emerson)

Compositional Verification

- Use system's natural decomposition into components to break-up the verification task
  - Divide-and-Conquer approach
- Components typically satisfy requirements in specific contexts / environments
  - safety assumptions about contexts
- System safety derives from the ability to compose the components' contexts at the system level

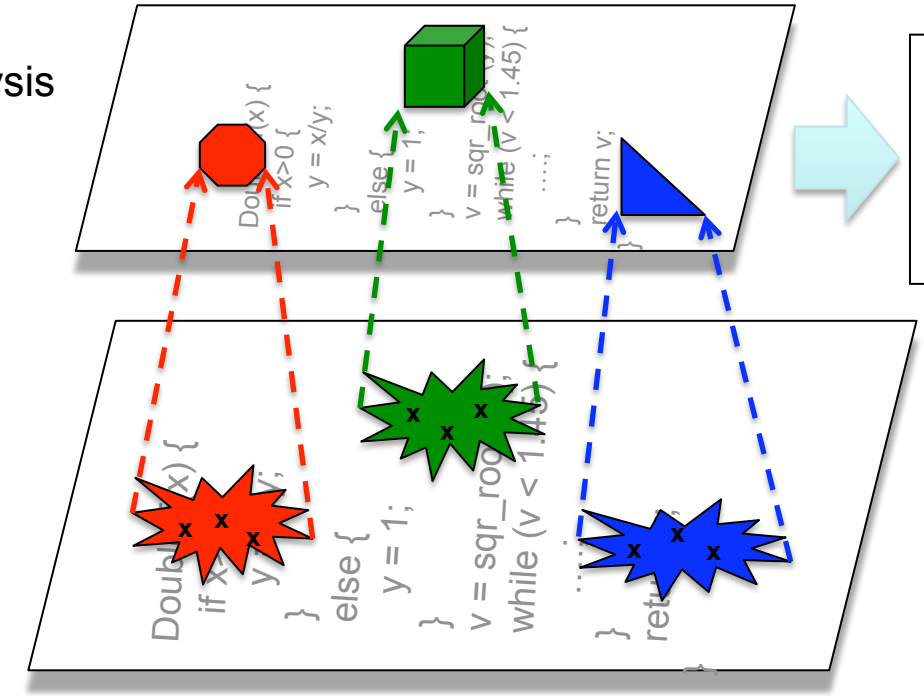
Model Checking Research Themes

- Evaluate and strengthen state of the art model checking capabilities
- Design more abstractions
  - Decision procedures for real number with non-linear operations
- Enable compositional verification for realistic applications
  - Develop and strengthen assumption generation capabilities
- Relate to evidences for safety cases
- Main metrics:
  - Code size
  - Thread interaction level

Static Analysis

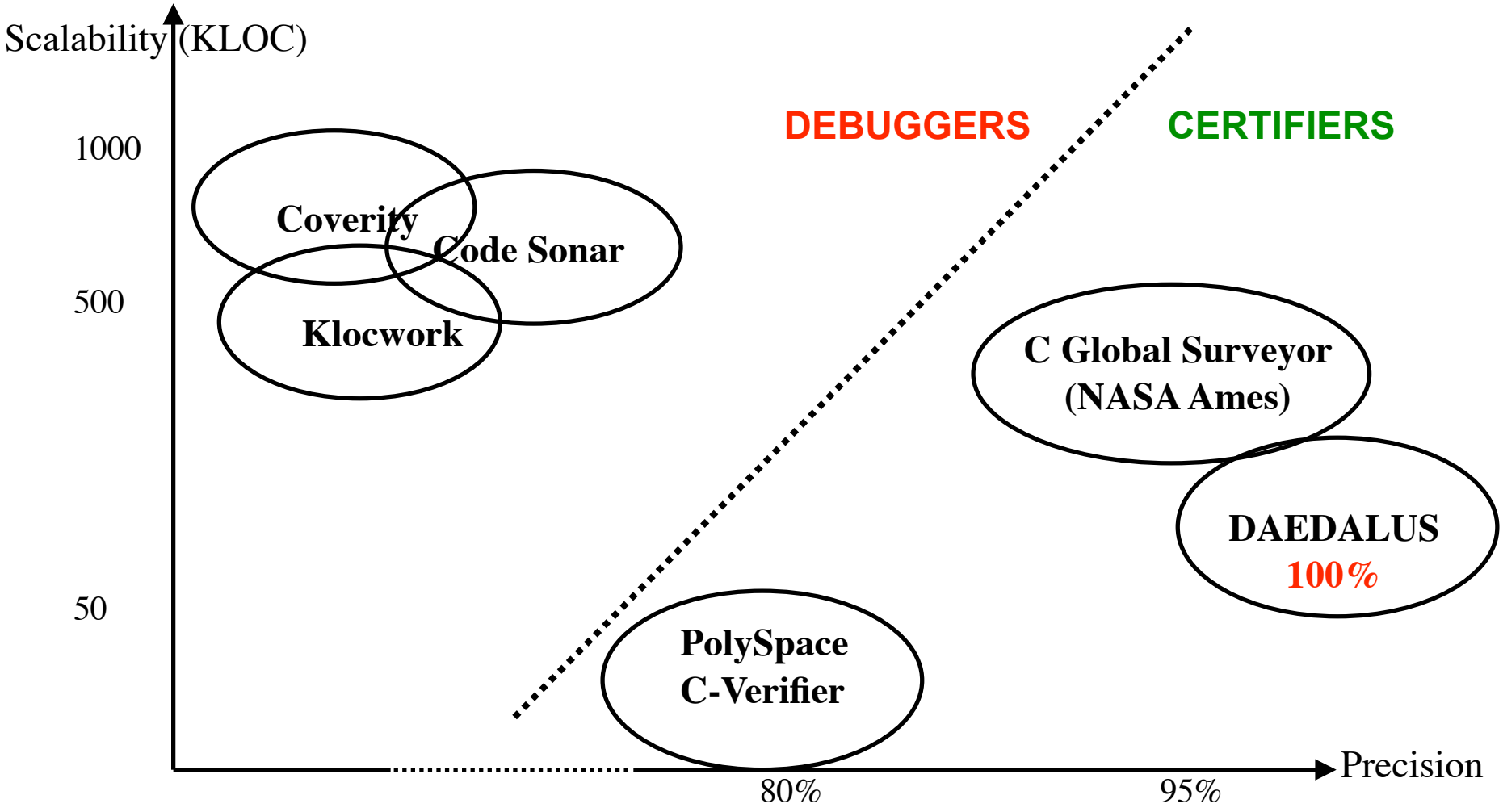
The goal of static analysis is to exercise all data ranges for all paths

Testing exercises some data points and some paths



Operations are:
- safe
- unsafe
- potentially unsafe

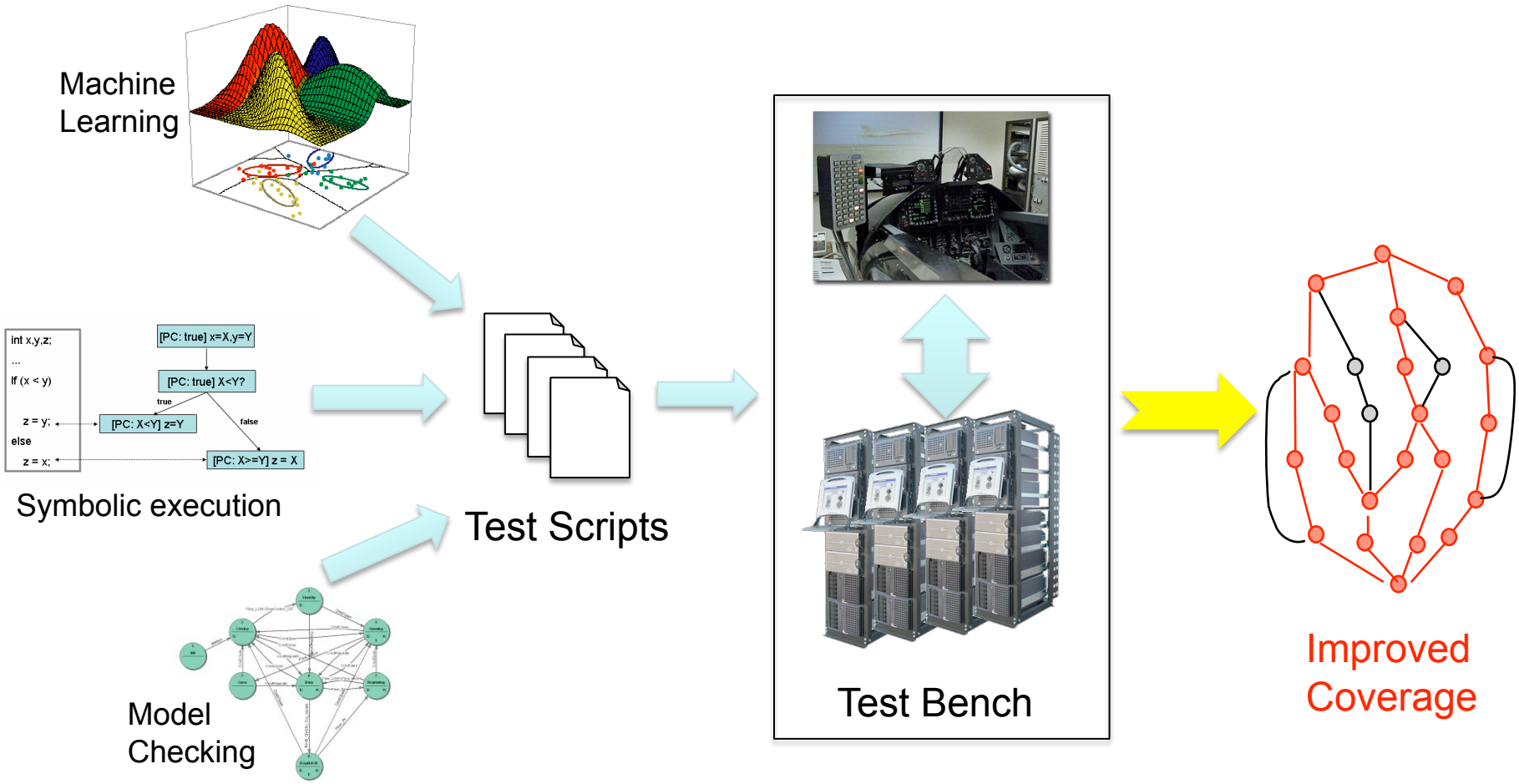
Static Analysis Challenge



Static Analysis Research Themes

- Target more than 90% precision for flight-critical code
  - Less than 10% of operations are deemed potentially unsafe
- Combine with other techniques to increase precision
  - Model checking
- Design more abstract domains
  - Build capabilities for floating-point computation analysis
- Relate to evidences for safety cases
- Main metrics:
  - Code size
  - False positive rate

Advanced Testing Challenges



Advanced Testing Research Themes

- Develop techniques for targeting dynamic areas for validation testing
- Combine with other techniques to improve coverage
  - Integration of testing and symbolic execution
  - Combine formal methods and testing for determining V&V coverage as evidence for building safety cases
  - Use machine learning to drive testing towards boundary cases for adaptive control systems
- Metrics:
  - Effectiveness of error discovery
  - Coverage

Other Research Themes

- Demonstration of formal verification and automated testing for diagnostic and monitoring systems using hybrid abstraction
  - *Milestone inherited from the IVHM project*
- Generation of publicly available verification and validation foundational libraries
- Demonstration of the use of formal methods to build a safety case for an IMA-like flight-critical code
- Verification and validation of automatically generated code, and ultimately, code/model generator
- Concurrent safety verification and validation of the design of hardware and software systems

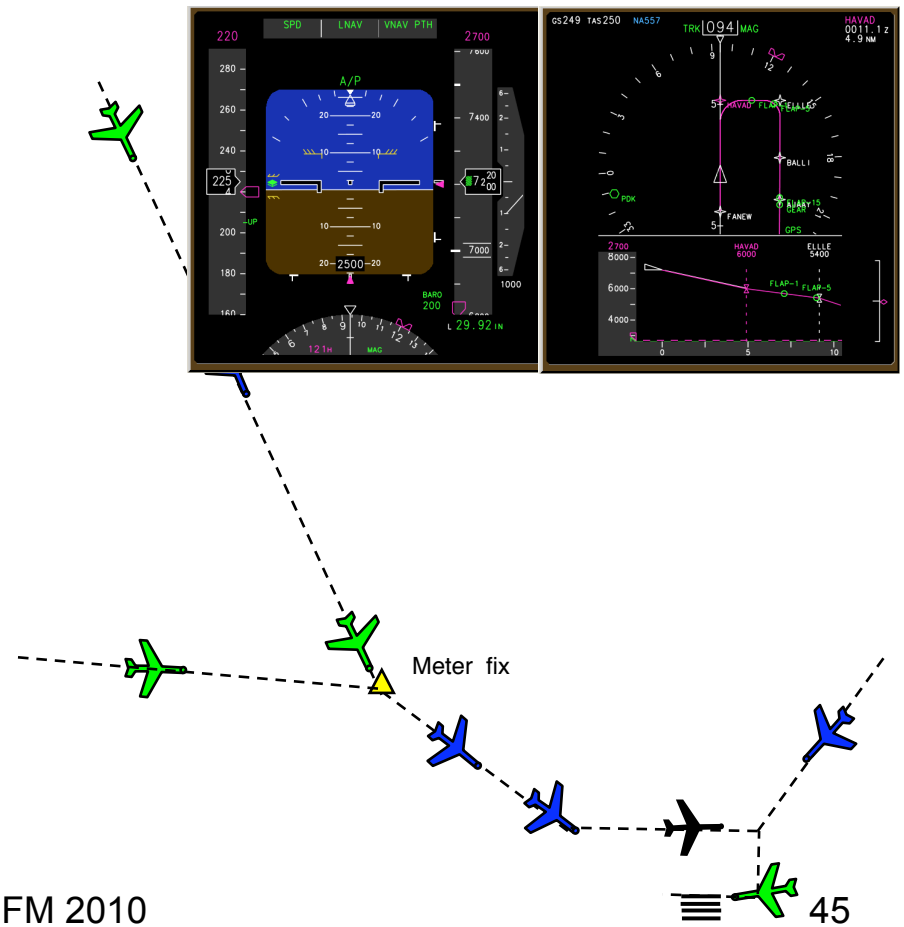
Two potential application domains

- Integrated Alerting and Notification concepts, implemented in Integrated, Modular Avionics (IMA) Architecture
  - Dryden Flight Research Center will provide h/w & s/w in the loop test bench at the highest level of fidelity
- Investigating Congested Airspace Applications
  - Automated conflict detection & resolution
  - Efficient Flows into Congested Airspace (EFICA)

Airspace Case Study

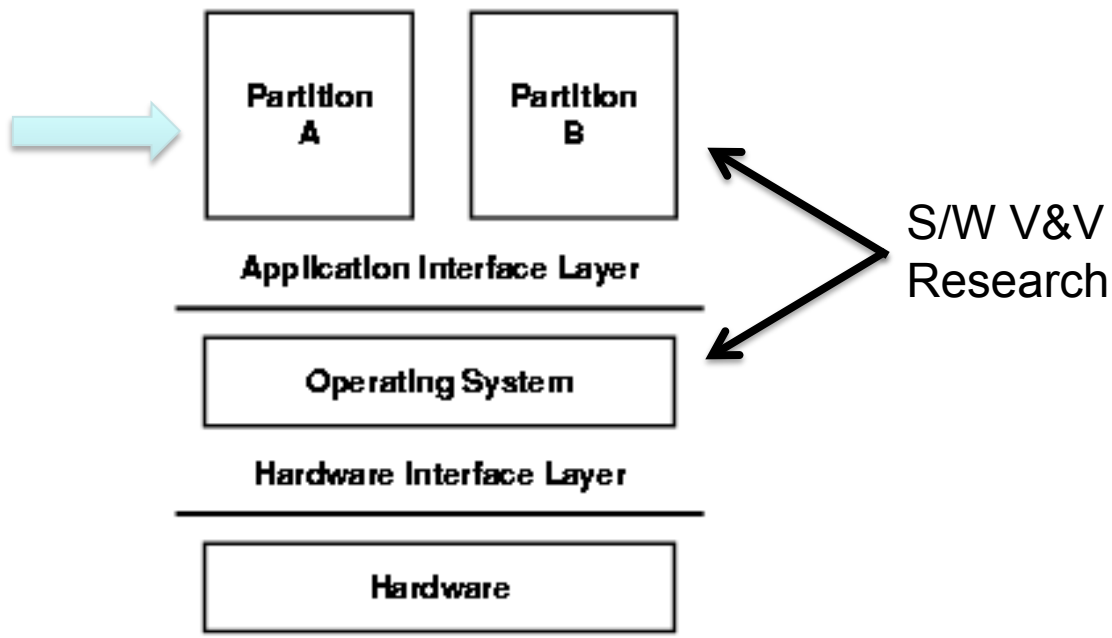
- The airspace-centric case study is a new operational concept for NextGen, which supports high-density merging and spacing operations
  - New procedures and tools for merging and spacing developed by Airspace Super Density Operations project
  - S/W prototypes and algorithms can be used to support S/W V&V research

FMS With Integrated eNAV Guidance

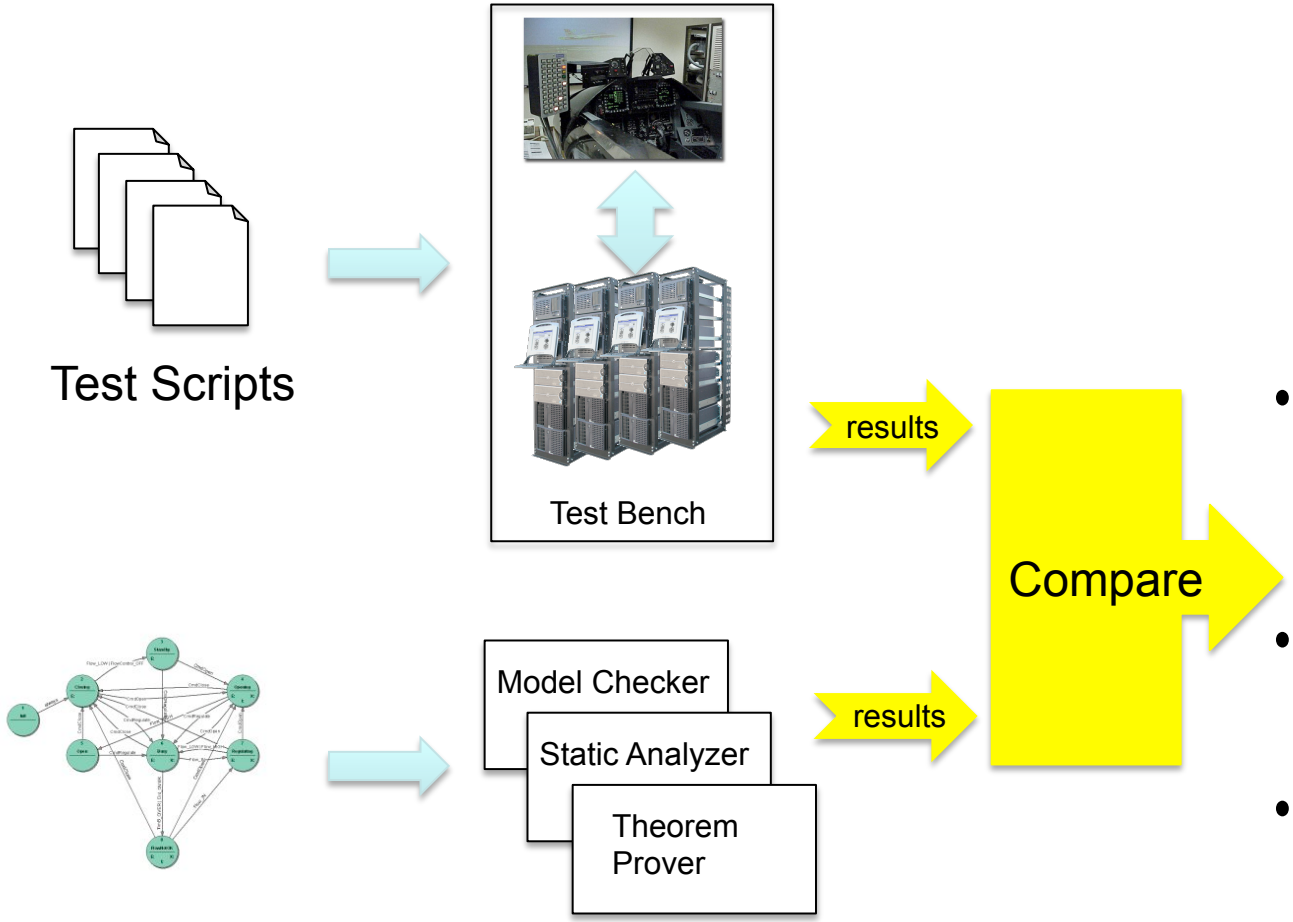


Vehicle Case Study

- Research prototypes developed for IAN will be ported on an IMA platform developed and hosted at Dryden
- It includes models, source code, and executables for the research prototypes developed by IAN



Use of Assessment Environment



- Establish baselines to quantify gains of formal methods
- Establish validity of formal methods
- Ground research in reality