Delivering Systems That Work:
Institutionalizing Performance Measures across the DoD Acquisition Life Cycle

Mr. Sean Brady
Office of the Deputy Assistant Secretary of Defense for Systems Engineering

83rd MORS Symposium
June 22-25, 2015

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Overview

- **PMO and OSD Decision Analysis Support: why do we measure?**
  - Data-driven decisions at every level (IPT Lead, Chief SE, PM, PEO, DAE, Congress)
  - Statue, Policy and Guidance

- **DASD(SE)'s Development of a Metrics Assessment Framework**
  - toward consistent performance measurement across the DoD Acquisition System
  - toward improved decision making

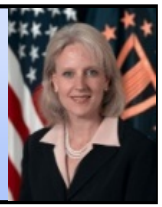
- **Impacting Decisions: DASD(SE) Metrics Efforts**
  - Schedule Realism and System Maturity
  - Knowledge Point: Performance Measures for OT Readiness
  - Agile Metrics
  - Enterprise Benchmarking
  - Software Parametric Statistical Analysis

- **Conclusions & Future Goals**

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

DASD, Systems Engineering

DASD, Systems Engineering
Stephen Welby
Principal Deputy Kristen Baldwin



Major Program Support
James Thompson

Supporting USD(AT&L) Decisions with Independent Engineering Expertise

- **Engineering Assessment / Mentoring of Major Defense Programs**
- **Program Support Assessments**
- **Overarching Integrated Product Team and Defense Acquisition Board Support**
- **Systems Engineering Plans**
- **Systemic Root Cause Analysis**
- **Development Planning/Early SE**
- **Program Protection**

Engineering Enterprise
Robert Gold

Leading Systems Engineering Practice in DoD and Industry

- **Systems Engineering Policy and Guidance**
- **Technical Workforce Development**
- **Specialty Engineering (System Safety, Reliability and Maintainability, Quality, Manufacturing, Producibility, Human Systems Integration)**
- **Security, Anti-Tamper, Counterfeit Prevention**
- **Standardization**
- **Engineering Tools and Environments**

Providing technical support and systems engineering leadership and oversight to USD(AT&L) in support of planned and ongoing acquisition programs

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Performance Measurement is Crucial to DoD

- **The DoD engineering challenge:**
  acquire, within budget and schedule, the beyond-cutting-edge systems that will provide U.S. warfighters overwhelming superiority in the full spectrum of conflicts

- **DoD develops & delivers incredibly effective but increasingly complex weapon systems to the Warfighter**

- **Increasing complexity of Warfighting capability demands more effective life cycle SE and quantitative insight**

- **DoD's senior leadership depends on data to drive multi-billion dollar decisions which impact warfighting capability**

- **DASD(SE) is committed to using a quantitative SE approach to**
  - mentor major PMOs and system developers; shape program plans; monitor execution
  - inform DoD leadership of technical risks, opportunities, and impacts to schedule & performance at major decisions
  - reduce time-to-Warfighter and cost for System and Software acquisition

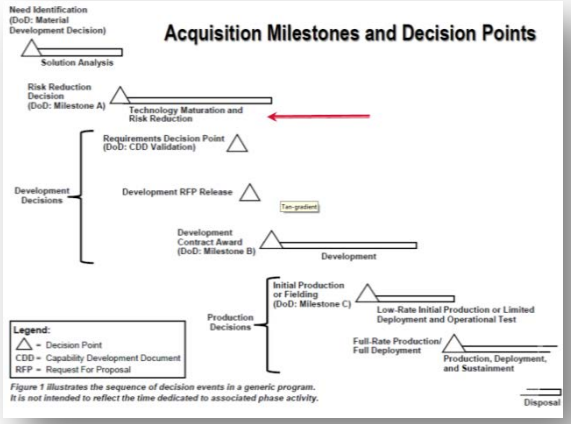
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

D, MPS Mission

Inform Decision Makers to Understand and Mitigate Risk (Better Buying Power 3.0)

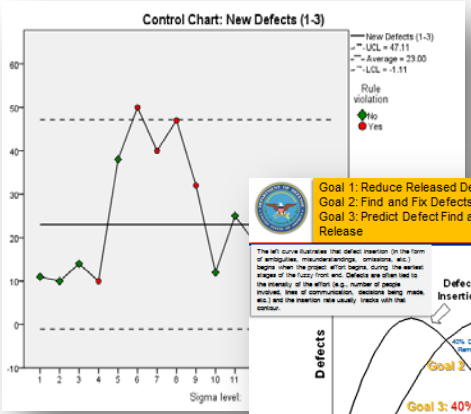
Knowledge Points

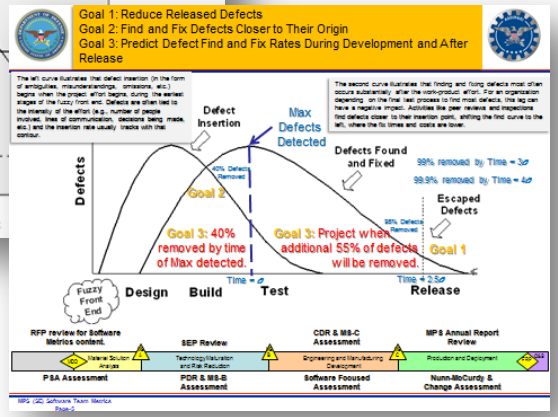
- What information is relevant at decisions maker engagement points?
- DoDI 5000.02 provides some key knowledge points including acq milestones, decisions points, and SE Technical Reviews



Inflection Points

- What changes in metrics alert decision makers to emerging problems?





- **Both are relevant in the context of:**
  - Knowledge sought
  - Decisions to be made
  - Data / documentation available

Represented as Goals, associated Questions, and performance Measures (GQM)

MPS supports data-driven decisions to reduce risk.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Why do we measure?

- **Weapon Systems Acquisition Reform Act (2009)**
  - S.454-10; d.(1): The development and tracking of detailed measurable performance criteria as part of the systems engineering master plans…

- **Systems Engineering Plan (SEP) Outline, 20 April 2011**
  - Directs programs to present their strategy for identifying, prioritizing, and selecting metrics for monitoring and tracking program SE activities and performance
  - **Sect 3.6. - "Technical Performance Measures and Metrics"**
    - Provides an overview of measurement planning and metrics selection process
    - Include approach to monitor execution-to-plan and identification of roles, responsibilities, and authorities
    - Minimum set of TPMs and intermediate goals and plan to achieve them with dates
    - Examples include TPMs in areas of software, reliability, manufacturing, integration, and test

Performance measures are foundational to DASD(SE)'s mission.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Performance Measurement Shortfalls & DASD(SE) Initiatives

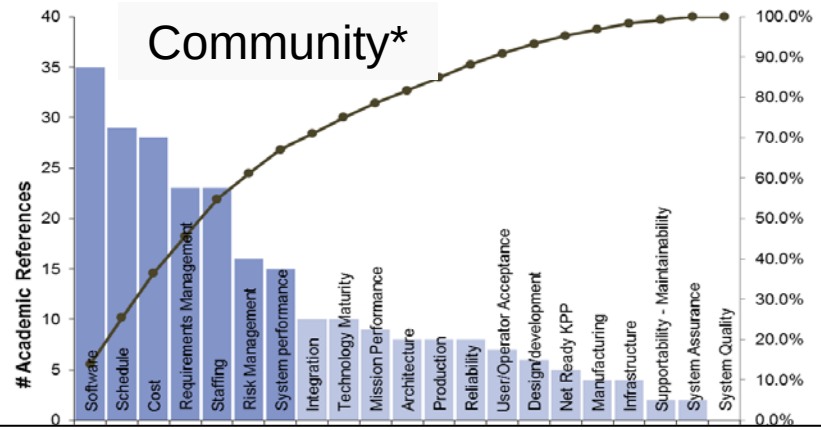
- **Background**
  - By statute, DASD(SE) is responsible for "the *development and tracking of detailed measureable performance criteria* as part of the [Systems Engineering Plans]"
    - Performance measurement over a system's lifecycle is at the core of our mission

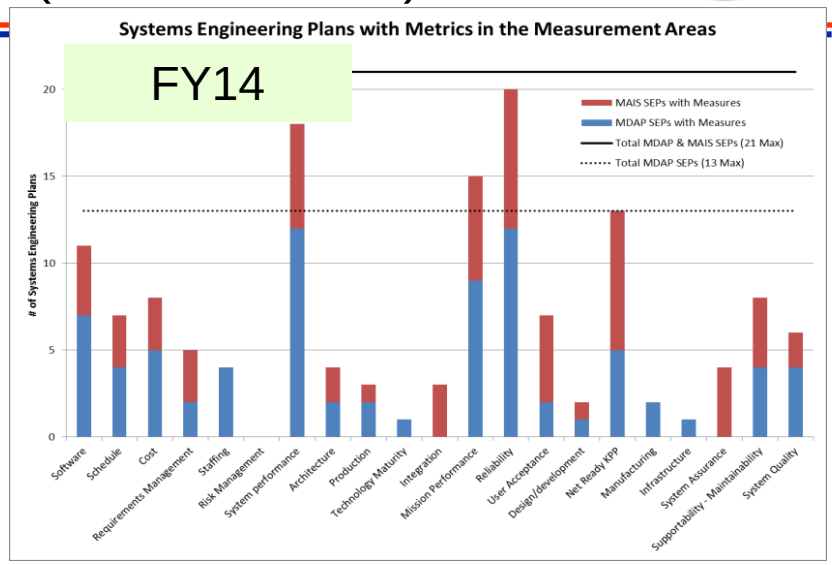
- **Performance Measurement Shortfalls**
  - DASD(SE) identified systemic issues in Annual Report to Congress and various forums (e.g., NDIA, MORS)
  - Lack of sufficient **predictive metrics** and **quantitative management**
    - DASD(SE) mentoring:
      - Introduced programs to defect prediction and measurement techniques as a means to plan, measure and control software quality, and assess maturity
  - Lack of **end-to-end performance measurement, developer/tester disconnect** and **insufficient integration testing**
    - DASD(SE) mentoring:
      - MPS developed a framework that aids programs' identification of effective measures to track;
      - Helped programs develop the right performance measures

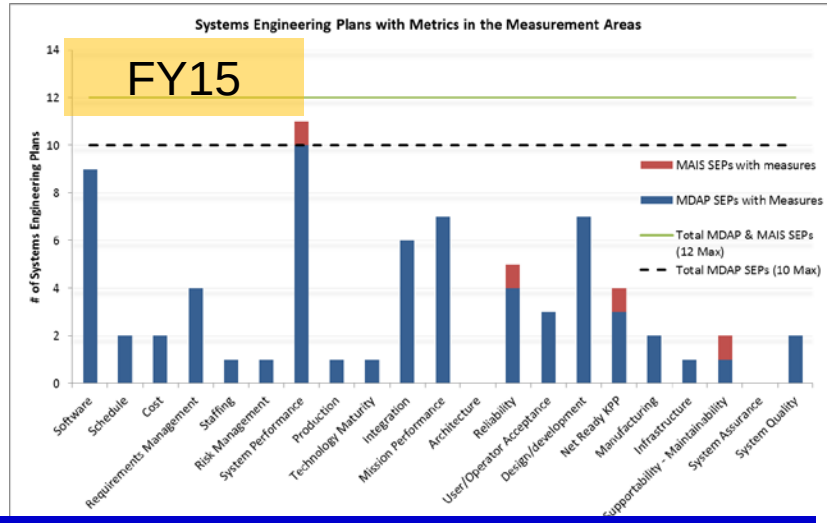
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Enterprise-level Insights & Trends:

-- Measurement Categories in Practice --
Community vs. SEP Samples (FY14 vs. FY15)







Top 10 Measure Categories

Community Sources:	FY14 SEPs:	FY15 SEPs:
1. Software	1. Reliability	1. System Performance
2. Schedule	2. Sys Performance	2. Software
3. Cost	3. Mission Performance	3. Mission Performance
4. Requirements	4. Net Ready KPP	4. Design/Development
5. Staffing	5. Software	5. Integration
6. Risk	6. Supportability – Maintainability	6. Reliability
7. Sys Performance	7. Cost	7. Requirements Management
8. Integration	8. Schedule	8. Net Ready KPP
9. Tech Maturity	9. User Acceptance	9. User Acceptance
10. Msn Performance	10. System Quality	10. Schedule

Differences between Community recommended measures and "real world" measures in SEPs may be driven by law, DoD policy & guidance (e.g., reliability, NR KPP).

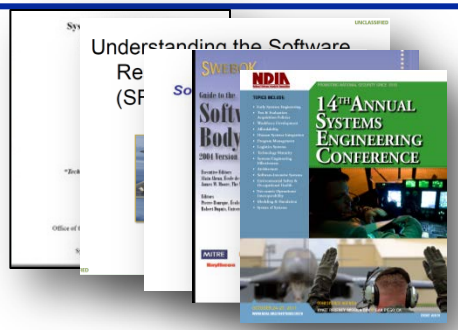
*Community: Govt (P&G, reports, training) / Industry (reports, standards) /Academia (papers)

Developing the right measures and tracking performance is a challenge in DoD programs. Identifying high-level trends across portfolio and enabling improved guidance.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

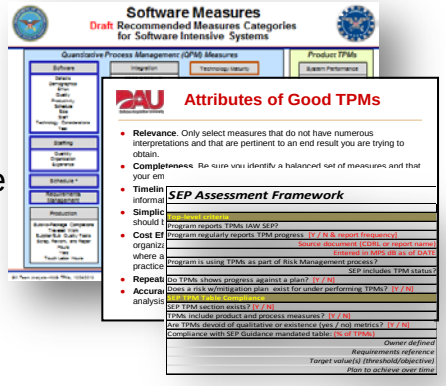
Performance Measures Study
Approach to Develop & Maintain a Practical Assessment Framework

Lean Six Sigma project is focusing on improving metrics in SEPs.



43 total sources reviewed

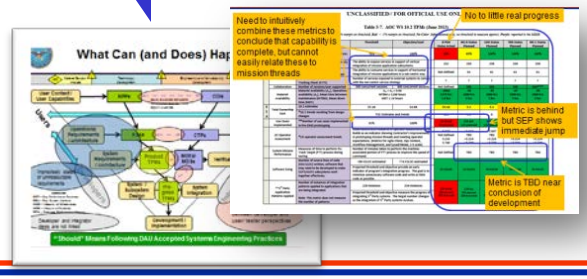
1. What metrics categories exist in DoD, Academia & Industry literature?

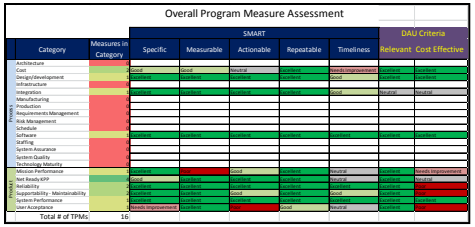


2. Created an Analysis Framework to evaluate SEPs' performance measures

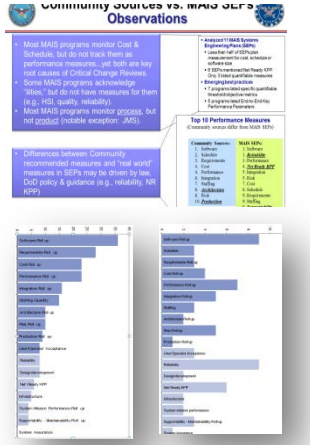
Framework = categories + criteria

3. Applied framework to select programs & study





4. Identified high-level trends across portfolio and recommended guidance

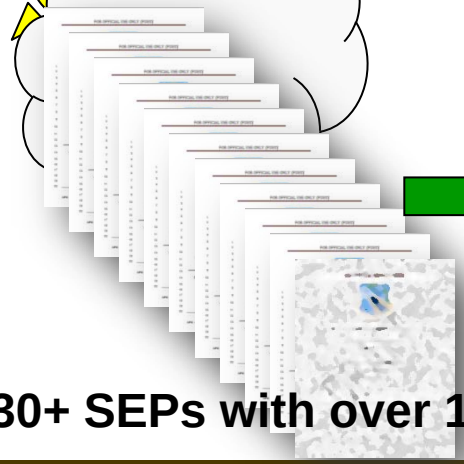


Compare/contrast Literature TPMs to Program SEPs' TPMs

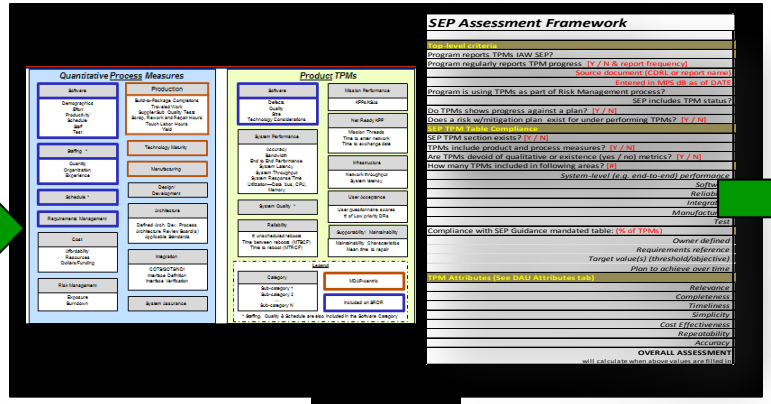
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

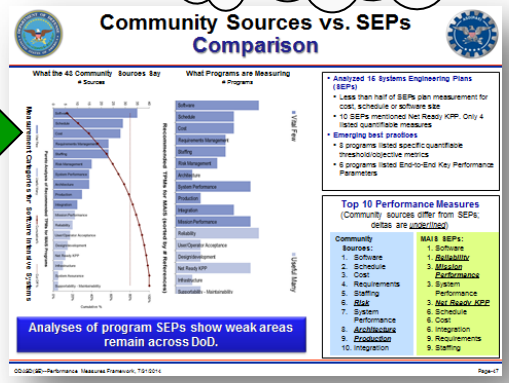
Study Outcomes & Framework Outputs
"Inside the Black Box"

High-level analysis and trends



30+ SEPs with over 1,000 metrics

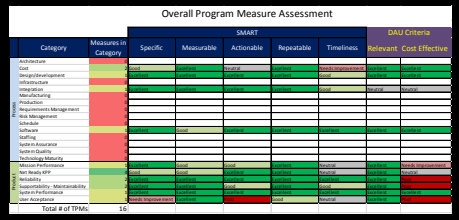




Detailed-level

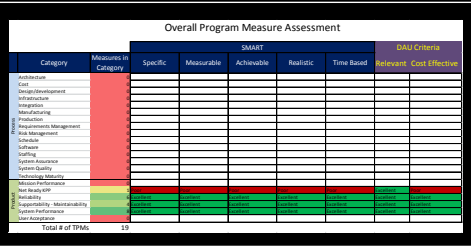
Practical application

Prgm A SEP



Not enough TPMs; No threshold / objective values; Measuring too late; Limited ability to influence program; Expensive to collect

Prgm B SEP

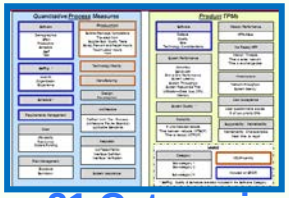


No mission performance metrics; exclusively focused on "Product" measures; NR KPP unmeasurable

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

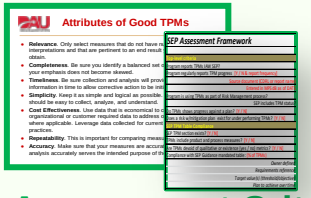
SEP Performance Measures Assessment Framework Purpose & Effect

Measures Assessment Framework



21 Categories

+

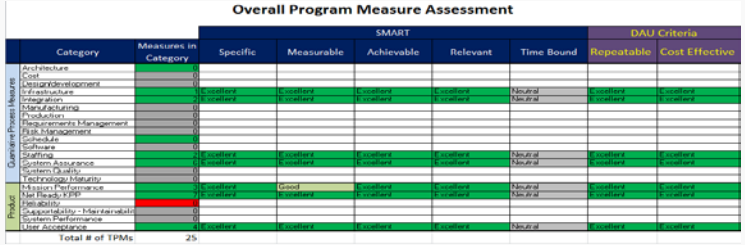


Assessment Criteria

MPS tool assesses 2 basic questions:

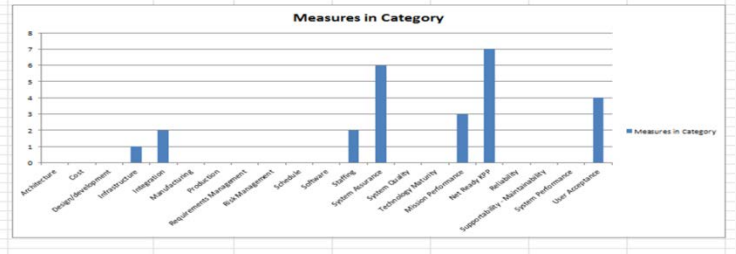
1. Does the program have measures in the appropriate categories?
2. How good are the measures the program selected?

Before

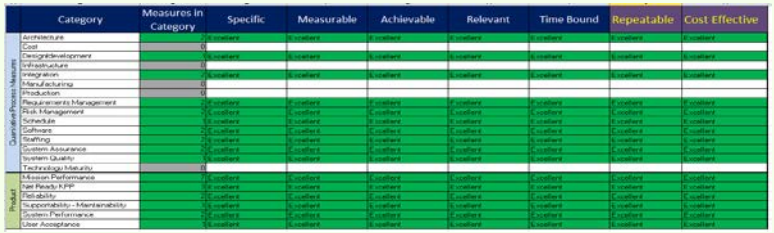


Overall Program Measure Assessment

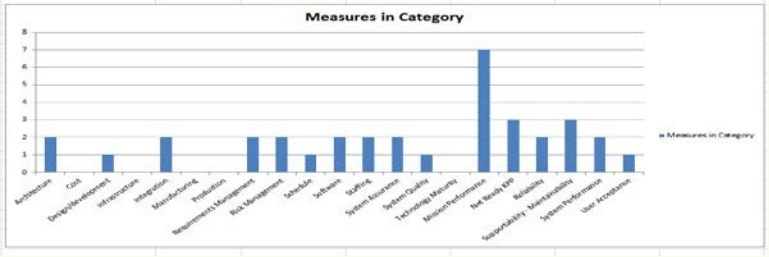
	Category	Measures in Category	SMART: Specific	Measurable	Achievable	Relevant	Time Bound	DAU Criteria: Repeatable	Cost Effective
Quantitative Process Measures	Architecture	0							
	Cost	0							
	Design/development	0							
	Infrastructure	1	Excellent	Excellent	Excellent	Excellent	Neutral	Excellent	Excellent
	Integration	2	Excellent	Excellent	Excellent	Excellent	Neutral	Excellent	Excellent
	Manufacturing	0							
	Production	0							
	Requirements Management	0							
	Risk Management	0							
	Schedule	0							
	Software	0							
	Staffing	2	Excellent	Excellent	Excellent	Excellent	Neutral	Excellent	Excellent
	System Assurance	0	Excellent	Excellent	Excellent	Excellent	Neutral	Excellent	Excellent
	System Quality	0							
	Technology Maturity	0							
Product	Mission Performance	3	Excellent	Good	Excellent	Excellent	Neutral	Excellent	Excellent
	Net Ready KPP	3	Excellent	Excellent	Excellent	Excellent	Neutral	Excellent	Excellent
	Reliability	1							
	Supportability - Maintainability	0							
	System Performance	0							
	User Acceptance	4	Excellent	Excellent	Excellent	Excellent	Neutral	Excellent	Excellent
	Total # of TPMs	25							



After (for SEP Signature)



	Category	Measures in Category	Specific	Measurable	Achievable	Relevant	Time Bound	Repeatable	Cost Effective
Quantitative Process Measures	Architecture	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Cost	0							
	Design/development	1	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Infrastructure	0							
	Integration	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Manufacturing	0							
	Production	0							
	Requirements Management	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Risk Management	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Schedule	1	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Software	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Staffing	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	System Assurance	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	System Quality	1	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Technology Maturity	0							
Product	Mission Performance	7	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Net Ready KPP	3	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Reliability	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Supportability - Maintainability	3	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	System Performance	2	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	User Acceptance	1	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent



In FY14, MPS Performance Measure assessment tool assisted 20 programs in improving their individual measurement plans.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Performance Measures

Recommended Measurement Categories for DoD Acquisition Programs

Quantitative Process Measures

Software
- Demographics
- Effort
- Productivity, Agile Velocity
- Schedule
- Staff
- Test

Staffing *
- Quantity
- Effort Hours
- Experience
- Turnover Rate

Schedule *

Requirements Management

Cost
- Affordability
- Resources
- Dollars/Funding
- CPI

Risk Management
- Exposure
- Burndown

Production
- Build-to-Package Completions
- Traveled Work
- Supplier/Sub Quality Tests
- Scrap, Rework and Repair Hours
- Touch Labor Hours
- Yield

Technology Maturity

Manufacturing

Design/ Development

Architecture
- % DoDAF drawings complete
- Quality Attributes
- Flexibility, Stability

Integration
- COTS/GOTS/NDI Components
- Interface Definition
- Interface Verification
- Interface Stability

System Assurance

Infrastructure

Product TPMs

Software
- Defects
- Quality
- Size

System Performance
- Accuracy | Lethality
- Bandwidth
- System Latency
- System Throughput
- System Response Time
- Utilization—Data bus, CPU, Memory
- SWAP-C | Range

System Quality *

Reliability
- # unscheduled reboots
- Time between reboots (MTBCF)
- Time to reboot (MTRCF)
- MTBF, MTTF

Mission Performance
- Mission Thread & End-to-End Performance
- e.g. Probability of Detection

Net Ready KPP
- Network Management
- Time to enter network
- Time to exchange data

User Acceptance
- User questionnaire scores
- User acceptance scores

Supportability/ Maintainability
- Maintainability Characteristics
- Mean time to repair

Legend

Category
- Sub-category 1
- Sub-category 2
- …
- Sub-category N

MDAP-centric

Included on SRDR

* Staffing, Quality & Schedule are also included in the Software Category

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

"Brady Matrix"
A Key Output of Measure Assessment Framework

Metrics Assessment Analysis

September version of SEP Rev 1.1, April 2014

Overall Program Measure Assessment

Updated Annex B: Risk, counted as a TPM for Integration and Risk Management.

SIL Integration / Infrastructure: to be addressed in SEP update post contract award

Software Requirements Stability TPMs will be addressed in SEP update post contract award

Contractor IMS and Risks to be addressed in SEP update post contract award

Contractor Software TPMs will be included

	Category	M/O	Measures in Category	DAU Criteria: Repeatable	DAU Criteria: Cost Effective	SMART: Specific	SMART: Measurable	SMART: Achievable	SMART: Relevant	SMART: Time Bound
Quantitative Process Measures	Architecture	opt	0							
	Cost	mand	1	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Poor
	Design/development		0							
	Infrastructure	opt	1	Excellent	Excellent	Poor	Neutral	Poor	Excellent	Poor
	Integration	opt	0							
	Manufacturing	mand	1	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Poor
	Production	opt	0							
	Requirements Management	opt	1	Excellent	Excellent	Poor	Neutral	Poor	[illegible]	Poor
	Risk Management	mand	1	Excellent	Neutral	Excellent	Good	[illegible]	Excellent	Good
	Schedule	mand	3	Excellent	Excellent	Needs Improvement	Good	Poor	Excellent	Poor
	Software	opt	1	Excellent	Excellent	Poor	Neutral	Poor	Excellent	Poor
	Staffing	mand	1	Excellent	Excellent	Neutral	Neutral	[illegible]	[illegible]	Poor
	System Assurance	opt	0							
	System Quality	opt	1	Excellent	Excellent	Neutral	[illegible]	[illegible]	[illegible]	[illegible]
	Technology Maturity	opt	2	Excellent	Excellent	Excellent	[illegible]	[illegible]	[illegible]	[illegible]
Product	Mission Performance	mand	0							
	Net Ready KPP	mand	0							
	Reliability	mand	3	Excellent	Excellent	Excellent	[illegible]	[illegible]	[illegible]	[illegible]
	Supportability - Maintainability	opt	0							
	System Performance	opt	19	Excellent	Good	Good	[illegible]	[illegible]	[illegible]	[illegible]
	User Acceptance	opt	0							
	Total # of TPMs		31							

End to End Mission Performance measures are comprised of Mobility, Lethality, Availability, and Survivability requirements

Net Ready KPP is JROC approved with Yes/No Compliance Criteria

Page-3

SEP Measures / TPM Categories / Assessment Criteria

1. Cut & Paste Performance Measure List Here in Column A

2. Select a Category from the dropdown

Performance Measure	Category	SMART Attributes: Specific	Measurable	Achievable	Realistic	Time Based	Additional DAU Recommendations: Relevant	Cost Effective
[illegible]	Cost	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Disagree	Strongly Agree	Strongly Agree
[illegible]	Cost	Disagree	Disagree	Disagree	Strongly Agree	Disagree	Strongly Agree	Strongly Agree
[illegible]	Design/development	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Strongly Agree	Strongly Agree
[illegible]	Integration	Strongly Agree	Disagree	Disagree	Strongly Agree	Agree	Neutral	Neutral
[illegible]	Mission Performance	Strongly Agree	Disagree	Disagree	Disagree	Disagree	Strongly Agree	disagree
[illegible]	Net Ready KPP	Disagree	Disagree	Disagree	Strongly Agree	Neutral	Strongly Agree	Neutral
[illegible]	Net Ready KPP	Agree	Disagree	Disagree	Strongly Agree	Neutral	Strongly Agree	Strongly Agree
[illegible]	Net Ready KPP	Agree	Disagree	Disagree	Strongly Agree	Neutral	Strongly Agree	Disagree
[illegible]	Net Ready KPP	Agree	Disagree	Disagree	Strongly Agree	Neutral	Strongly Agree	Disagree
[illegible]	Reliability	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Neutral	Strongly Agree	Strongly Disagree
[illegible]	Reliability	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Neutral	Strongly Agree	Strongly Disagree
[illegible]	Software	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Neutral	Strongly Agree	Strongly Agree
[illegible]	Supportability - Maintainability	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Neutral	Strongly Agree	Strongly Disagree
[illegible]	Supportability - Maintainability	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Neutral	Strongly Agree	Strongly Disagree
[illegible]	System Performance	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree
[illegible]	User Acceptance	Disagree	Disagree	Disagree	Agree	Disagree	Strongly Agree	Strongly Disagree

Each TPM assessed against criteria

SMART Criteria
(Source: Peter Drucker)
- **S**pecific
- **M**easurable
- **A**chievable
- **R**ealistic
- **T**ime Based

DAU Criteria
- Relevance
- Cost Effectiveness

Practical framework highlights measurement risk areas & tester-developer disconnect

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

SEP Coverage of Metrics Categories FY14 & 15

		Category																						
		Quantitative Process Measures																	Technical Performance Measures					
FY	Program SEPs	Software	cost	System Quality	requirements management	staffing	schedule	System Assurance	Integration	Manufacturing	architecture	Design/development	risk management	Technology Maturity	Infrastructure	Supportability - Maintainability	Production	Testing	RAM	system performance	Net ready kpp	Mission Performance	User Acceptance	% QPM Categories Covered by SEP	% TPM Categories Covered by SEP
14	Program SEP 1	■		■	■	■	■	■	■		■	■	■						■	■	■	■	■	59%	100%
14	Program SEP 2	■	■						■			■							■	■	■	■	■	24%	100%
14	Program SEP 3		■					■			■								■	■	■	■	■	18%	100%
14	Program SEP 4	■	■	■	■	■				■	■								■	■		■	■	41%	80%
14	Program SEP 5	■	■	■	■	■	■	■											■	■	■	■		41%	80%
14	Program SEP 6		■	■		■	■			■		■							■	■	■	■		35%	80%
14	Program SEP 7			■				■	■										■		■	■	■	18%	80%
14	Program SEP 8	■	■		■														■	■	■	■		18%	80%
14	Program SEP 9		■	■						■									■	■		■	■	18%	80%
14	Program SEP 10	■			■	■	■				■		■						■		■	■		35%	60%
14	Program SEP 11			■				■	■										■	■	■			18%	60%
14	Program SEP 12																		■	■	■			0%	60%
14	Program SEP 13	■	■																■	■				12%	40%
14	Program SEP 14			■															■	■				6%	40%
14	Program SEP 15																			■				0%	20%
14	Program SEP 16	■	■	■	■	■	■			■			■	■	■				■	■				59%	40%
14	Program SEP 17																		■		■			0%	40%
14	Program SEP 18	■	■		■				■	■		■		■			■		■	■	■	■	■	47%	100%
14 & 15	Program SEP 19	■							■			■				■			■	■	■	■	■	24%	100%
15	Program SEP 20																		■	■		■	■	0%	80%
15	Program SEP 21	■		■	■				■			■						■	■	■		■		35%	60%
15	Program SEP 22	■	■				■												■	■		■		18%	60%
15	Program SEP 23	■							■			■							■	■		■		18%	60%
15	Program SEP 24	■										■							■	■		■		12%	60%
15	Program SEP 25															■			■	■	■			6%	60%
15	Program SEP 26	■							■	■		■							■	■				24%	40%
15	Program SEP 27	■			■				■			■							■	■				24%	40%
15	Program SEP 28	■																	■	■				6%	40%
	% SEPs Covered by category (FY14)	50%	50%	45%	35%	30%	25%	25%	30%	25%	20%	25%	15%	10%	5%	5%	5%	0%	95%	85%	65%	65%	45%		
	% SEPs Covered by category (FY15)	80%	10%	10%	20%	0%	10%	0%	50%	10%	0%	60%	0%	0%	0%	20%	0%	10%	100%	100%	20%	60%	20%		

***Both %'s include the 2 Program SEPs in 14 &

Opportunity to increase visibility into development progress

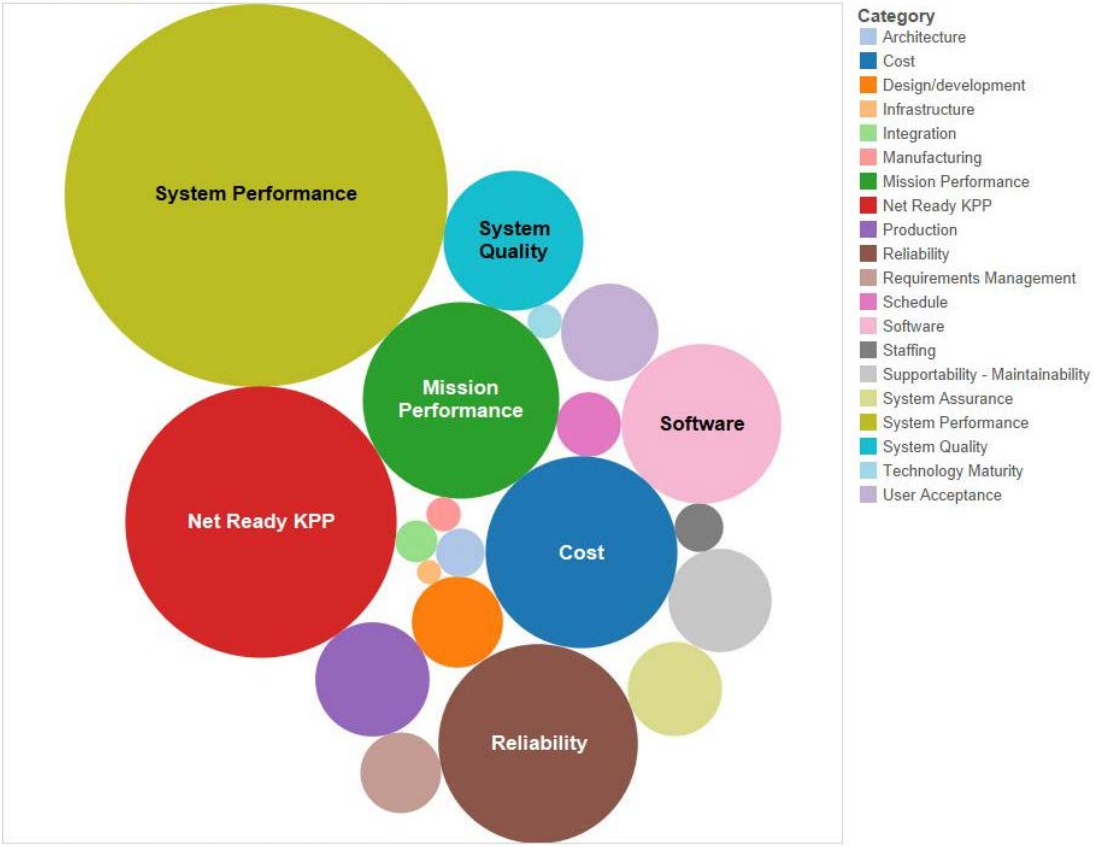
Quantitative Process Management
How far have you progressed in developing the product? *(e.g. design, cost, schedule)*

Technical Performance Measures
How well does your product do what it is supposed to do? *(e.g. throughput, CPU/memory use)*

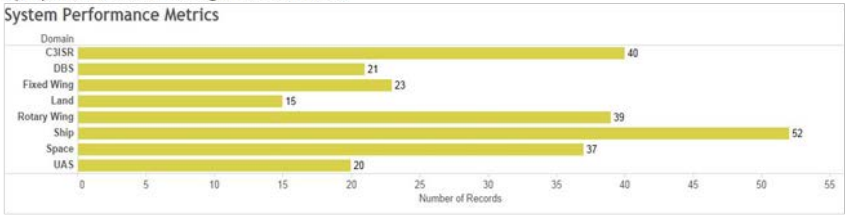
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

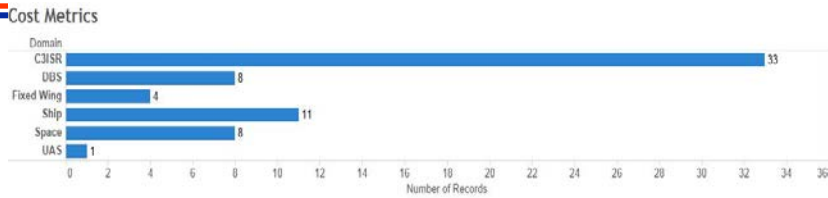
Metrics Distribution in FY14 SEPs
FY14 Software Focus Area Summary

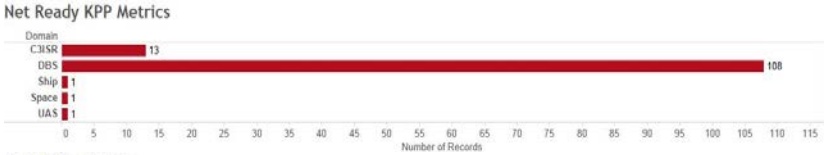
Distribution of Gathered Metrics

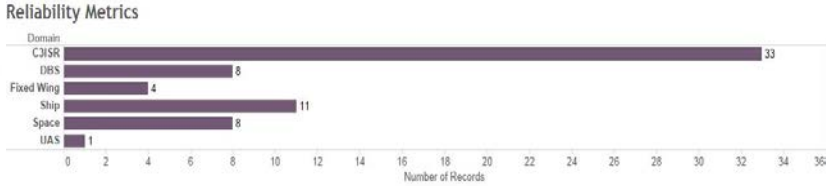


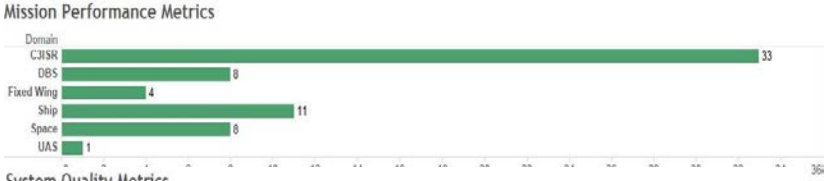
Color shows category and proportional size shows gathered metrics.

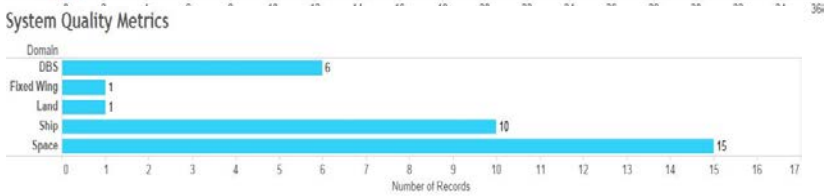


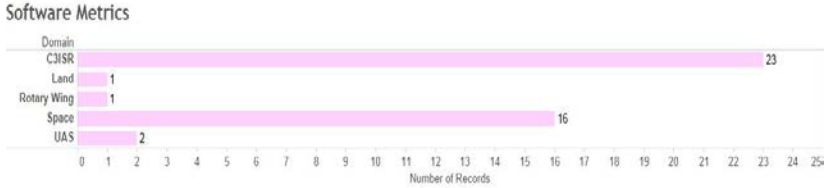






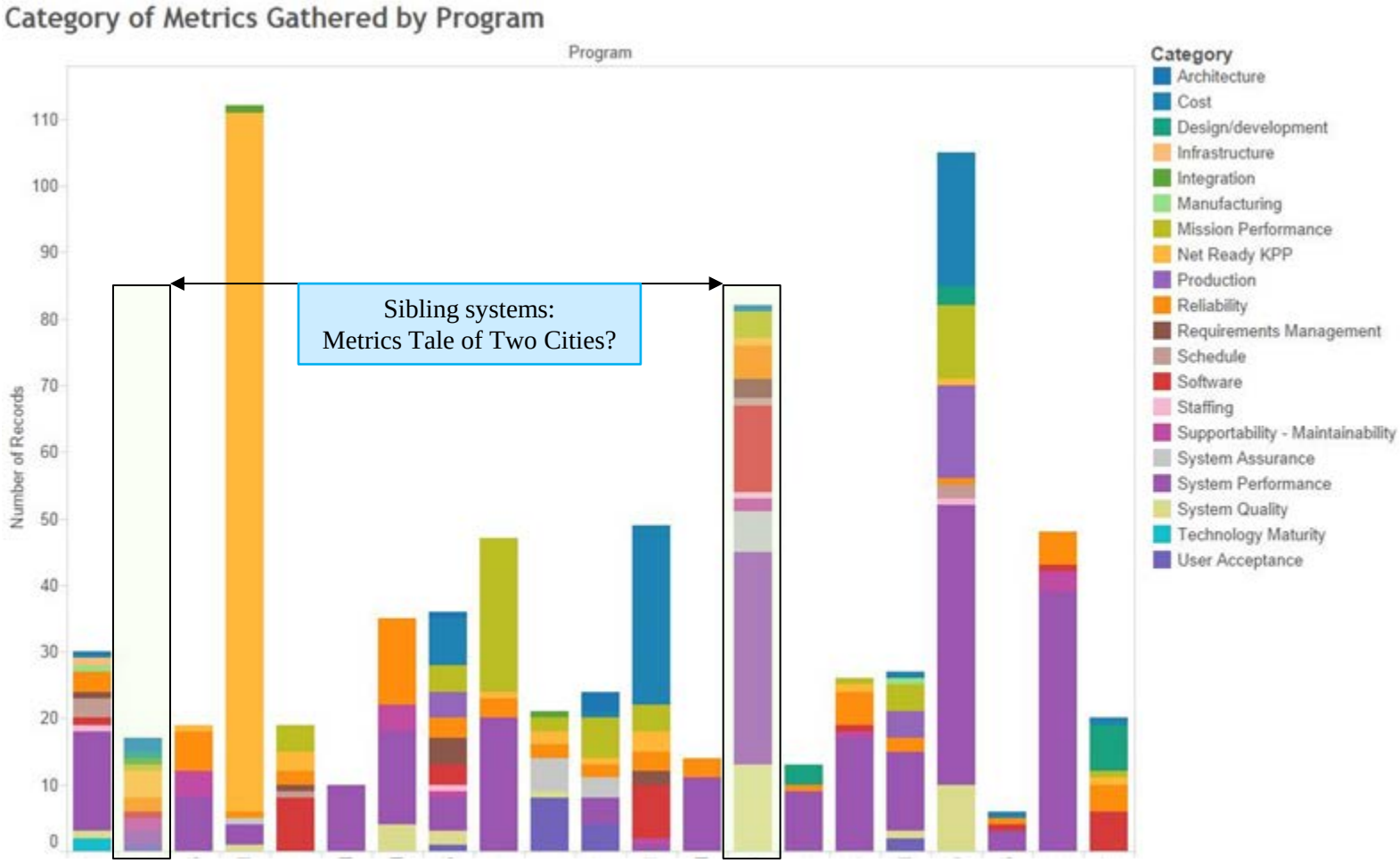






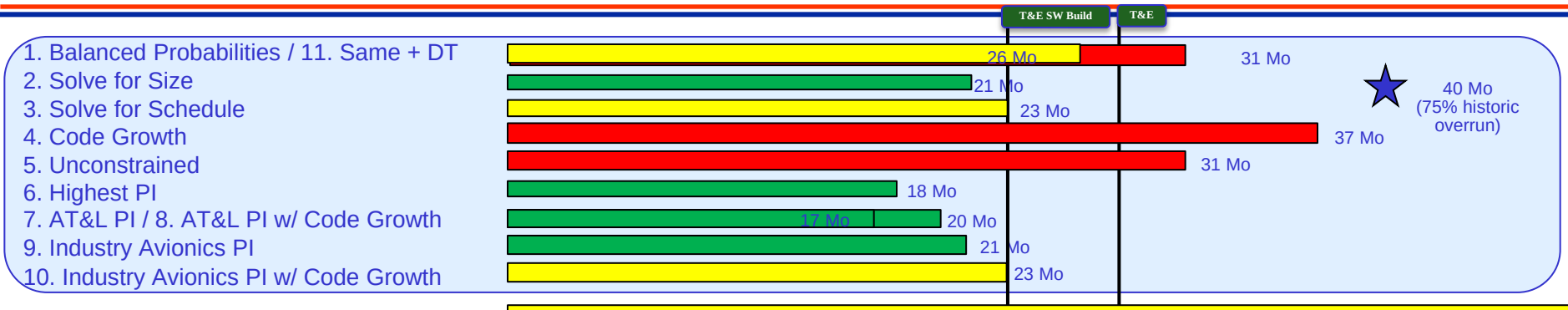
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Metrics Breakdown by Program
FY14 Software Focus Area Summary



Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

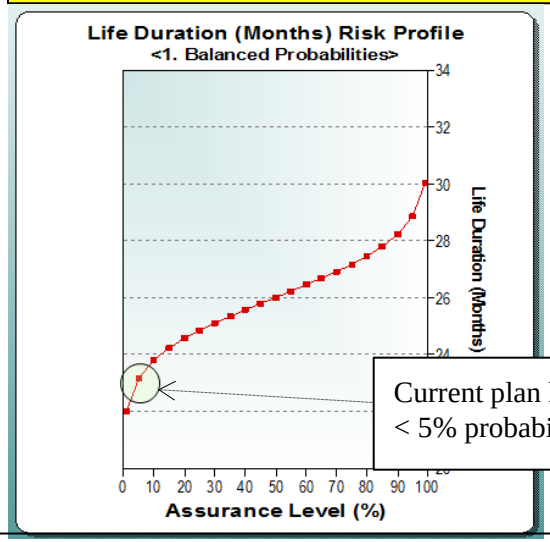
What if? MPS aids decisions, helps establish realistic schedules

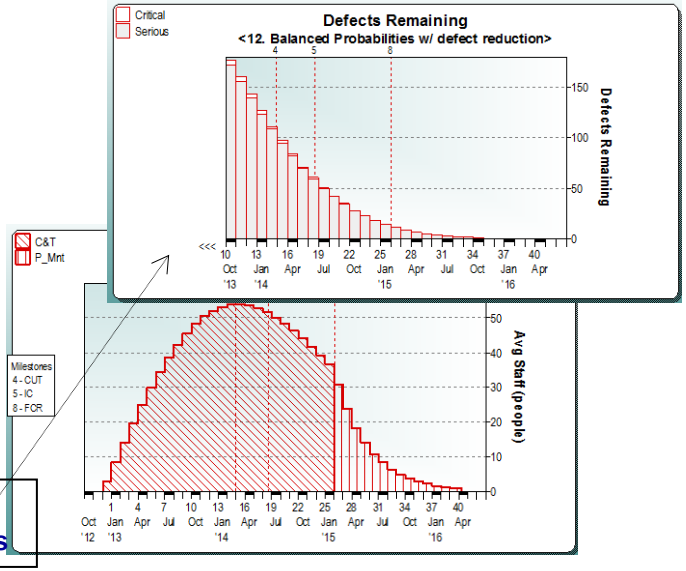


Scenario	Prob%*
Prime's Plan[µ]	N/A
1. Balanced Probabilities	5%
2. Solve for Size	90%
3. Solve for Schedule	50%
4. Code Growth[#]	<1%
5. Unconstrained	<1%
6. Highest Productivity	>99%
7. AT&L (Rotary/UAS) PI	>99%
8. AT&L w/ Code Growth	99%
9. Industry (Avionics) PI	85%
10. Industry PI Code Growth	50%

Estimates (yellow, red, green) of T&E Build overlaid on program schedule

1. Balanced: 80% probability not exceeding schedule/50% probability peak staff; very tight schedule
2. Solve for Size: reduces size to meet plan; reflects 70% of planned size; risks delivery of all planned features
3. Solve for Schedule (23-mon): high staffing and quality risks; likely unrealistic staffing (~2X planned staffing)
4. Code Growth: historical prime code growth average over 75%, misses schedule





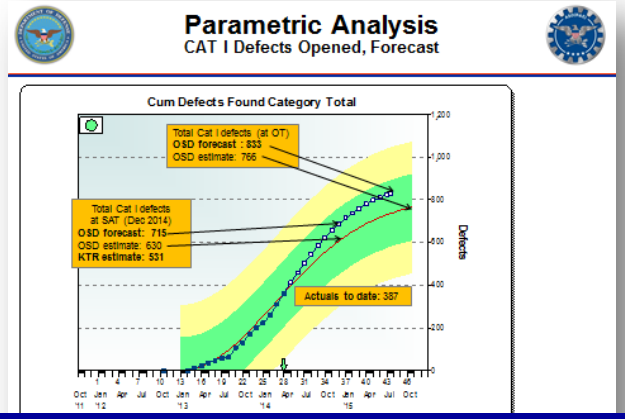
- **Extending project to remove defects:** ~37 months predicted to ensure SW is mature for T&E

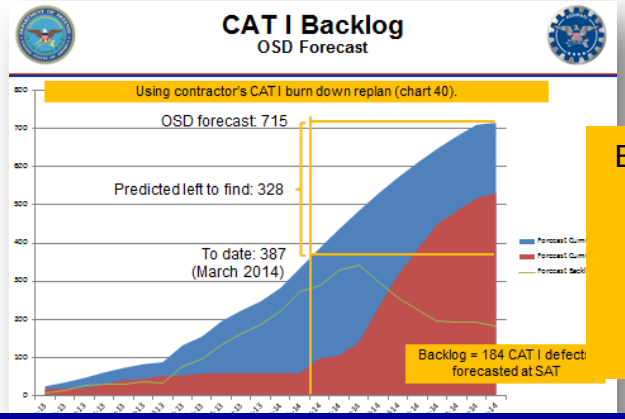
SW quality must be used as early-warning system for release management & Test Readiness

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

System Maturity Analysis

- **Defects are not closing at a rate that will ensure no CAT Is by entry into System Acceptance Test (SAT) per contractor's plan**
  - Contractor defect metric is currently yellow/red; re-planning to "get well"
- **OSD defect projection exceeds contractor's capacity to fix by SAT and indicates risk of delivering immature software to OT&E**
- **Recommendations**
  - PMO open a risk for CAT I defect closure by SAT and DT/OT events
  - **Status:** PMO and contractor agreed to increase projection by 100 defects



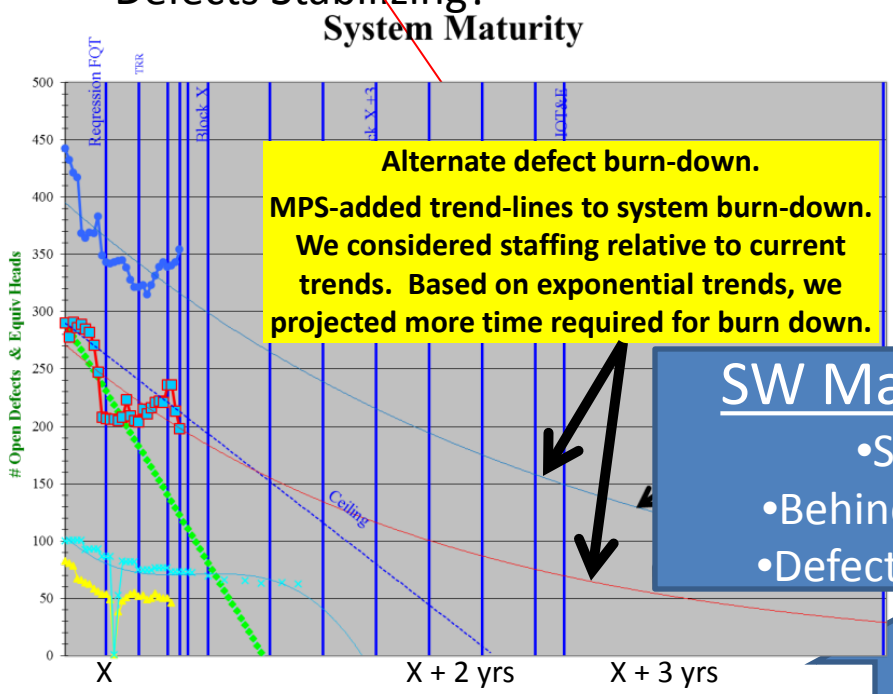


Best month defect burn down = 91 defects/month → Mar 2015 (~3 months late)
Average defect burn down = 48 defects/month → Apr 2015 (~4 months late)

OSD analysis predicts defect burn down risk to test events; up to 300 additional defects and 3-4 months to fix per OSD's CAT I defects predictions.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

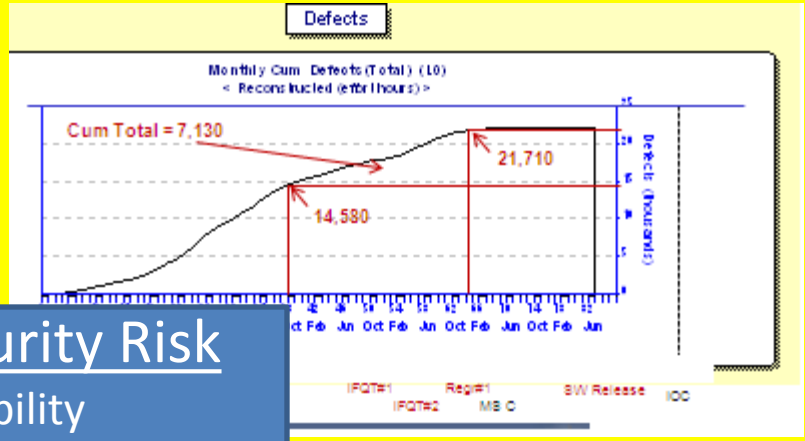
Defects Stabilizing?



Alternate defect burn-down.

MPS-added trend-lines to system burn-down. We considered staffing relative to current trends. Based on exponential trends, we projected more time required for burn down.

Defect Projection



Does the model reflect observed maturity risk?

SW Maturity Risk

- Stability
- Behind on Testing
- Defect Burn-down

Un-modeled factors contributing to maturity risk

Complicating Factors

- Late Formal CM (and CR process)
- Lab (late stand up)
- QA
- SW Supplier Quality
- M&S (Lab Fidelity)
- Staffing
  Loss of Knowledge, Skill, Ability and its effect on test
- Loosening of Allowable Defects for IOT&E

Execsum: Enough capacity? Sufficient schedule? Is X months to fix bugs/test insufficient mitigation?
Risk: If SW maturity problem is discovered, it is too late to add staff (Brooke's Law)
First Order Impact – Unsuccessful IOT&E
Recommend (1) Maintain/add staff (2) Monitor metrics submitted to OSD -- until adequately demo maturity

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

OSD Recommended Agile Metrics - Frequent Learning & Decision Support

- Team and aggregate velocity
- Team task & Aggregate story point burn-down
- Capabilities/Features completed
- Capabilities/Features accepted
- Product Backlog Stability
- Rate of closure vs discovery of defects

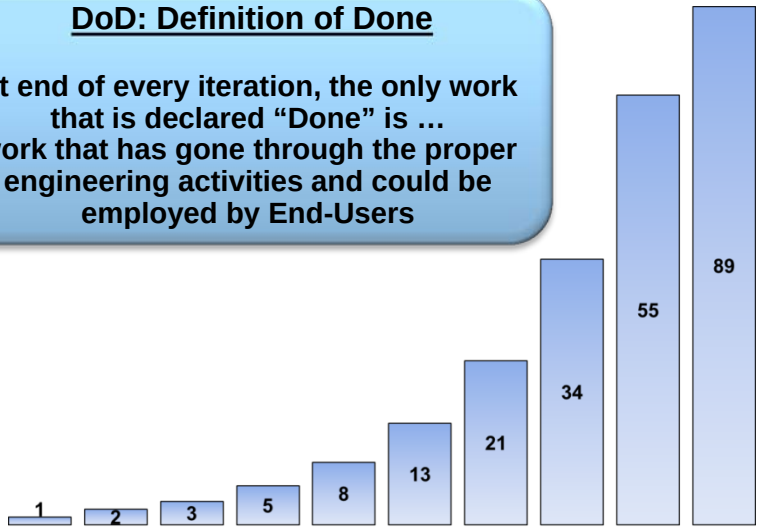
Core

- Rate of change of risk exposure
- Story point efficiency
- Percentage of key stakeholder groups represented at sprint demos
- Peer review effectiveness
- Hours per defect to resolve
- Staffing planned vs actual (by month/sprint)
- Staffing stability
- Agile competency

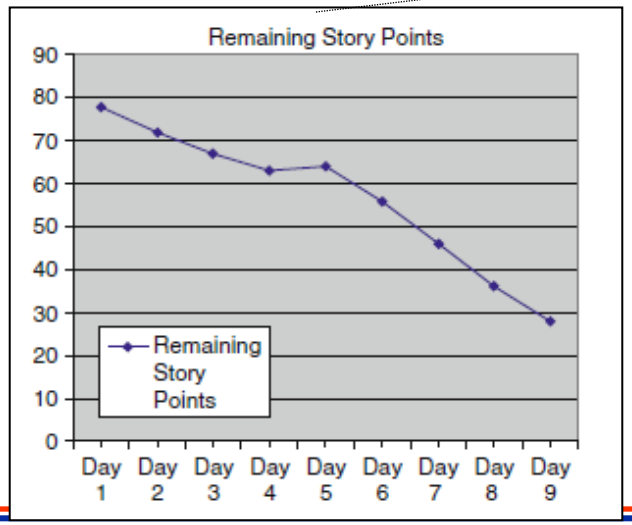
Agile tools and team & exec-level Agile metrics support rapid, near real-time decisions at IPT, PMO, CEO/CIO and MDA levels

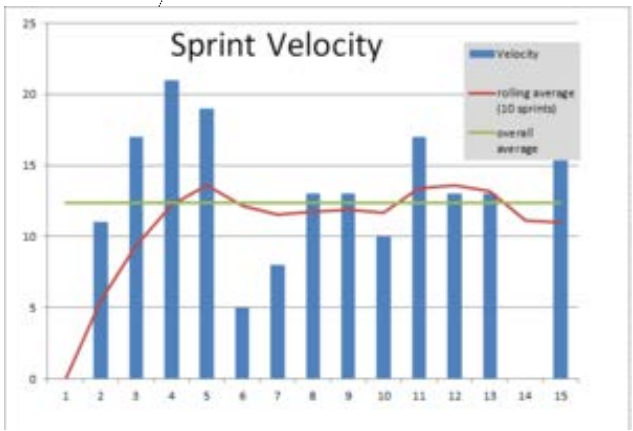
DoD: Definition of Done

At end of every iteration, the only work that is declared "Done" is … work that has gone through the proper engineering activities and could be employed by End-Users



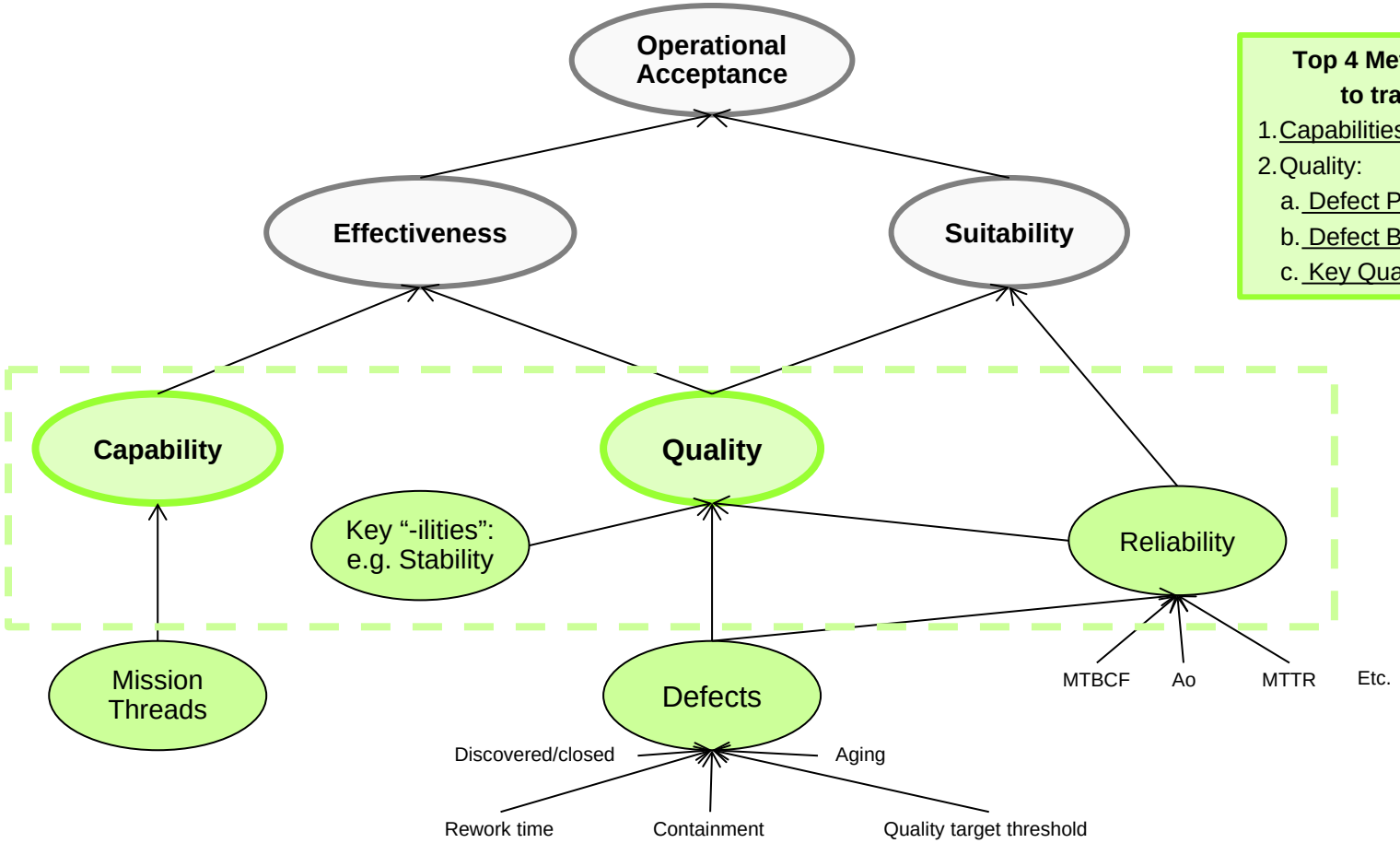
Fibonacci series often used for **story points**





Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Influence Diagram
Performance Measures for OT Readiness



Top 4 Metrics to Assess Readiness to transition from DT to OT

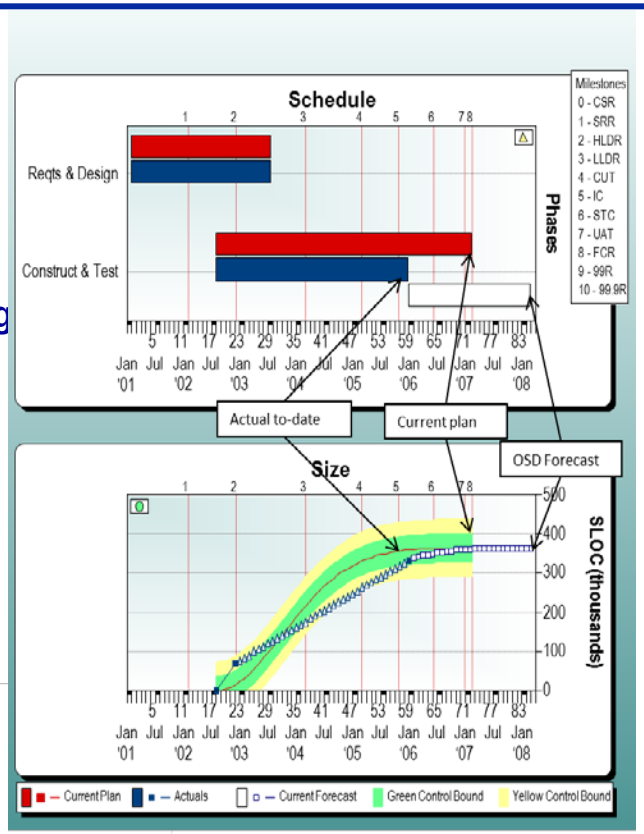
1. Capabilities met
2. Quality:
   a. Defect Profile Risk Assessment
   b. Defect Backlog (No Pri 1 or 2 defects)
   c. Key Quality Attributes "-ilities"

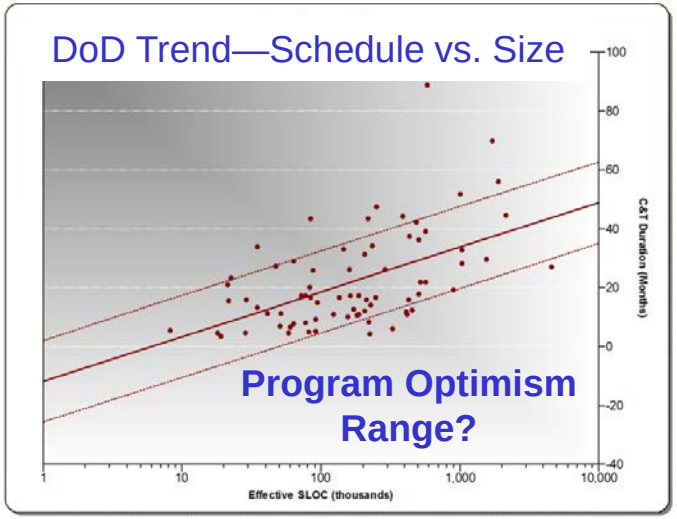
"To meet Better Buying Power goals to inform decision makers of program risk, we need to ***characterize a system's technical and SW maturity*** *(e.g. quality and stability) and* ***quantitatively assess its readiness at key knowledge points*** *in the development life cycle. OT readiness is one of several critical points."* (Mr. James Thompson, Director, MPS)

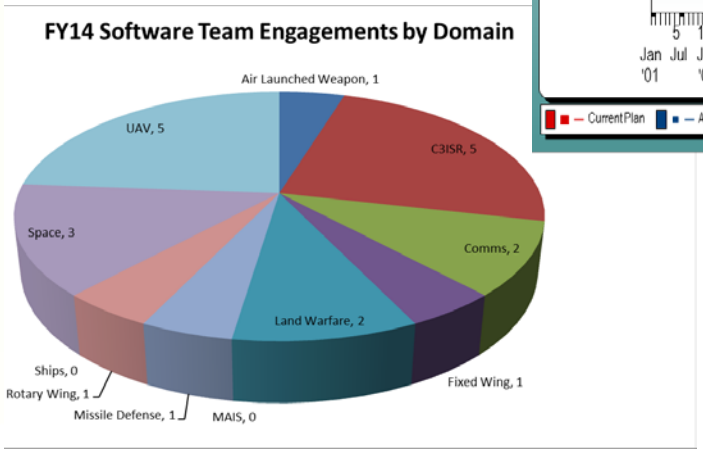
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Enterprise Benchmarking and Software Parametric Statistical Analysis

- **SW benchmarks: a foundation supporting DASD(SE) decisions**
  - underpin OSD independent parametric models
  - identifying statistically anomalous behavior
- **OSD SW Team gathers data and extracts core SW metrics**
  - for SW assessment and parametric analyses
  - e.g. size, effort, staffing, schedule, defects, demographics, etc.
- **Validated metrics form critical basis of comparison; benchmarking DoD trend-lines**
- **Parametric analyses of supported top findings in 21 program engagements**
  - Analyzed 52 SW releases
- **Analytical support to Congressional reporting on key Warfighting capability**
  - Unprecedented support: 413 excursions on complex SW helped deliver RTC predictions

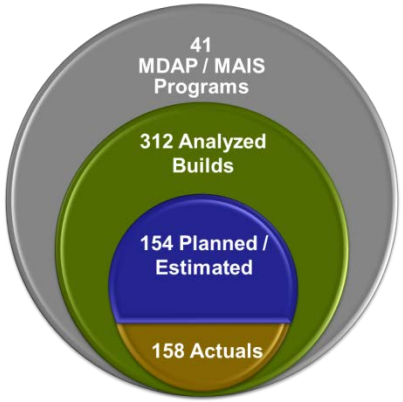




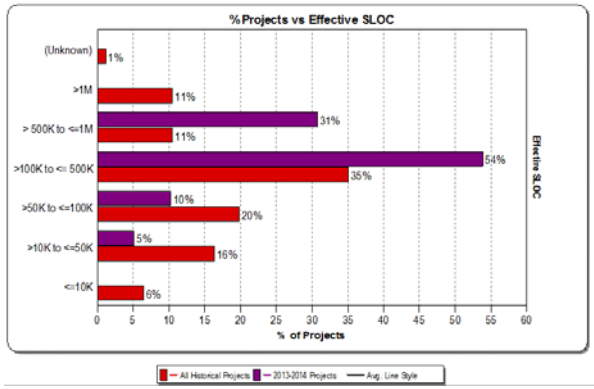


Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

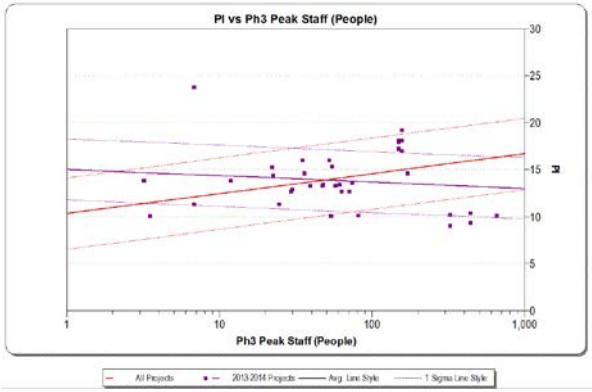
Sample of Software Benchmarks



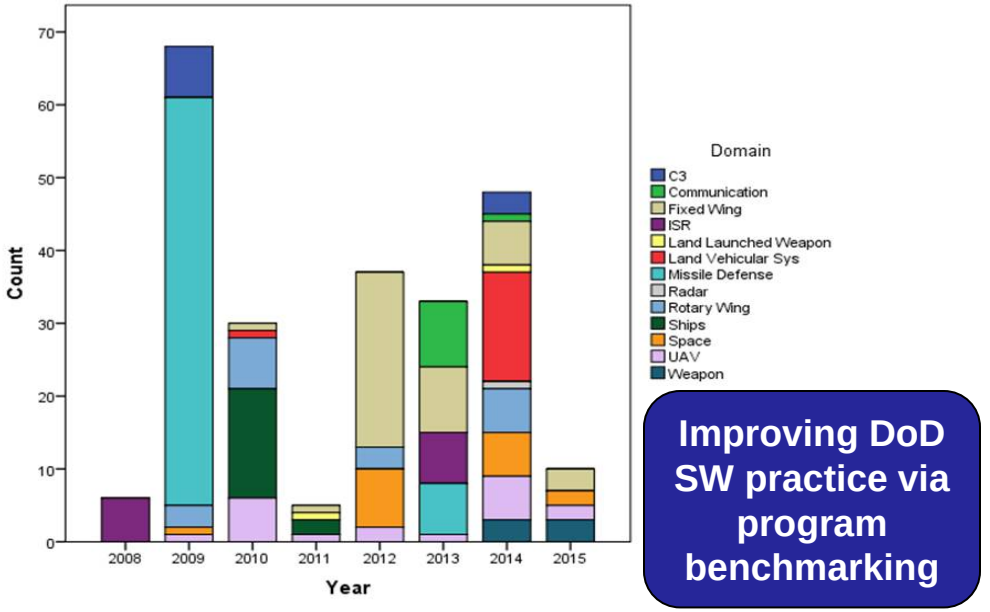
Increasing statistical power year over year —adding dozens of closed and planned projects



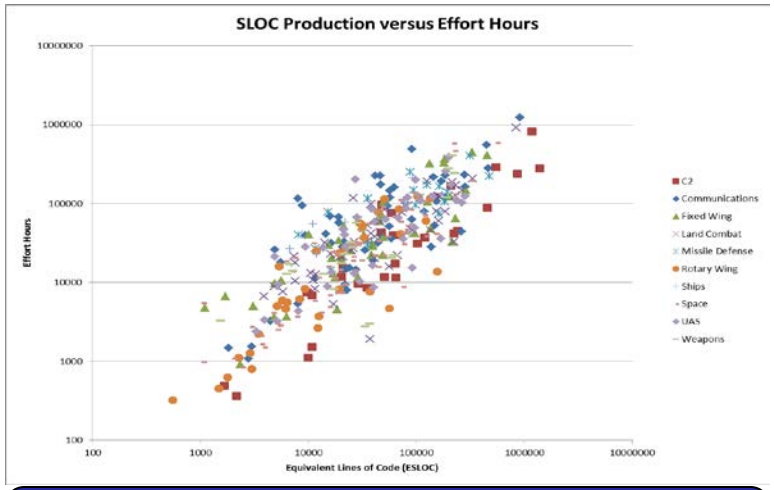
FY14 projects averaged 13% larger than historic trends.



DoD defied industry trends with lower productivity for larger software staffs.



Improving DoD SW practice via program benchmarking



Based on SRDR data for 374 subsystems DoD SW programs roughly follow the industry productivity heuristic of 1-2 SLOC per hour.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Conclusions & Future Goals

1. **We defined best practice performance measures**
   - 21 categories of expected performance measures
2. **We developed the SEP measurement assessment framework**
   - In use internally to support ODASD(SE)/MPS assessment of SEPs' metrics
3. **We quantitatively analyzed programs to identify trends and gaps**
   - Most MAIS programs measuring **Process** but not **Product** performance; MDAPs are the reverse
   - General lack of quantifiable threshold/objective values & end-to-end performance
4. **Goal: Strengthen *consistent* use of SE & SW metrics DOD-wide**
   - Characterize SE & SW development status at key knowledge points & define inflection points
   - Address tracking of metrics & enable path to growth (flexibility, tech refresh, modularity, quality…)
   - Training and fielding of tools: Metrics Best Practices, Assessment Framework, OSD Metrics Guide
5. **Goal: Quantitative, enterprise wide-assessments about state of DoD software practice**
   - Employ new analysis tools and OSD benchmarks
6. **Goal: Promote & advance quantitative SE approach by enhancing**
   - performance forecasting throughout the life cycle and support rapid decision making
   - structuring & leveraging of program data; trend analyses; correlation of results across programs

Performance measurement remains critical to DoD

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Additional Reference Material

- **Abstract**
- **Who are we?**
- **Performance Measures**
- **Zipper Chart: Tester Developer Disconnect**
- **DASD(SE) Portfolio**

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Abstract

This presentation details how the Office of the Deputy Assistant Secretary of Defense for System Engineering uses metrics and quantitative analysis to shape program plans, monitor execution, and assess technical risk in support of the acquisition decision making process. We discuss our analysis of defense, industry, and academia recommendations for process and product performance measures.

Based on these analyses, we developed a set of expected performance measure categories and an assessment framework to analyze acquisition programs' performance. Over the last year, we assessed performance measures on a variety of programs to identify best practices and improve System Engineering Plans (SEPs). Given our framework, we provided guidance and mentored a wide variety of major programs on performance measures, including Agile metrics and software maturity and test readiness metrics. The framework also serves as the basis for enterprise benchmarking, to include software parametric statistical analyses.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Deputy Assistant Secretary of Defense for Systems Engineering (DASD(SE)) Mission

Systems Engineering focuses on engineering excellence – the creative application of scientific principles:

- **To design, develop, construct and operate complex systems**
- **To forecast their behavior under specific operating conditions**
- **To deliver their intended function while addressing economic efficiency, environmental stewardship and safety of life and property**

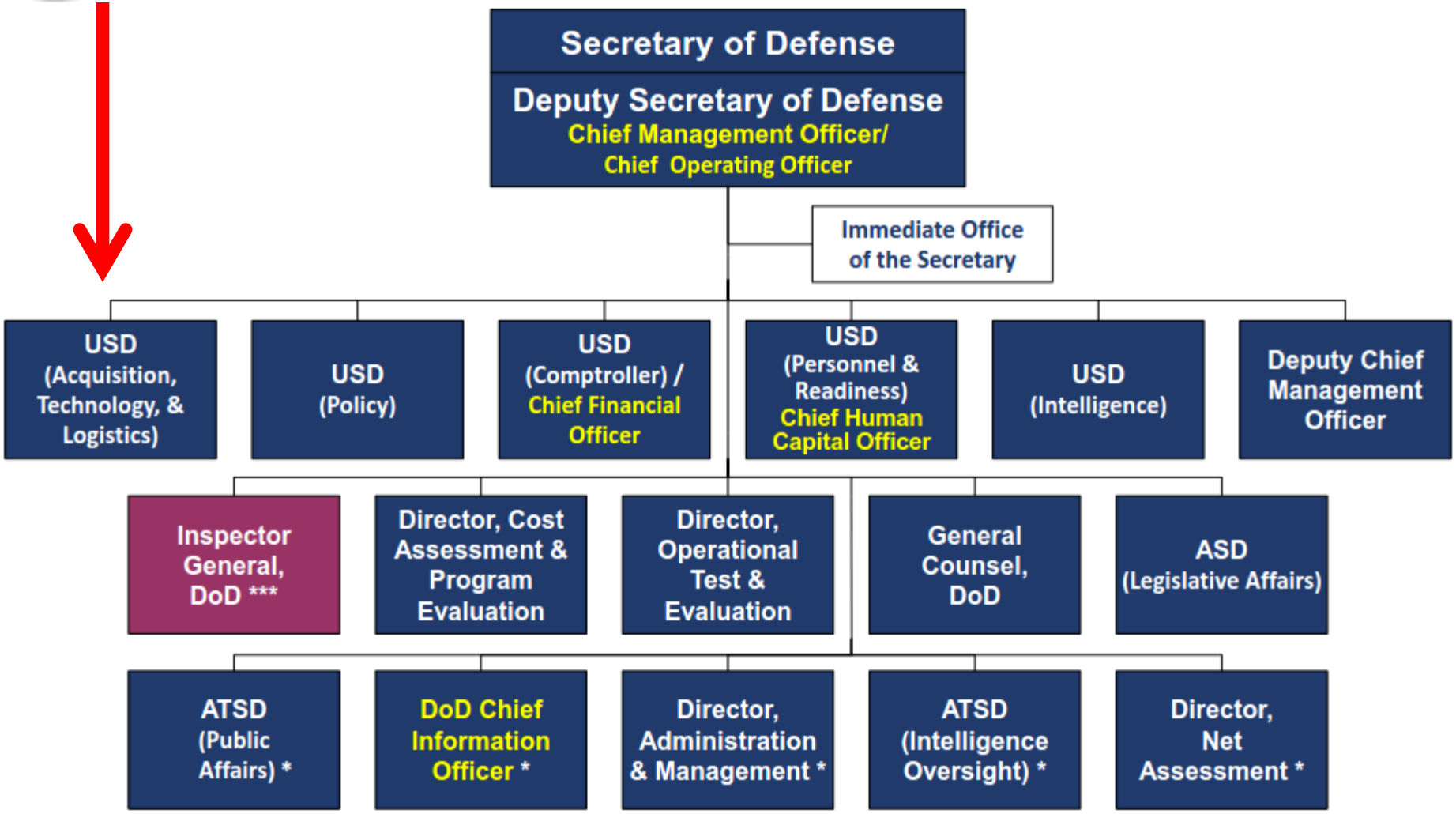
DASD(SE) Mission: Develop and grow the Systems Engineering capability of the Department of Defense – through engineering policy, continuous engagement with component Systems Engineering organizations and through substantive technical engagement throughout the acquisition life cycle with major and selected acquisition programs.

A Robust Systems Engineering Capability Across the Department Requires Attention to Policy, People and Practice

- ***US Department of Defense is the World's Largest Engineering Organization***
- ***Over 108,000 Uniformed and Civilian Engineers***
- ***Over 39,000 in the Engineering (ENG) Acquisition Workforce***

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

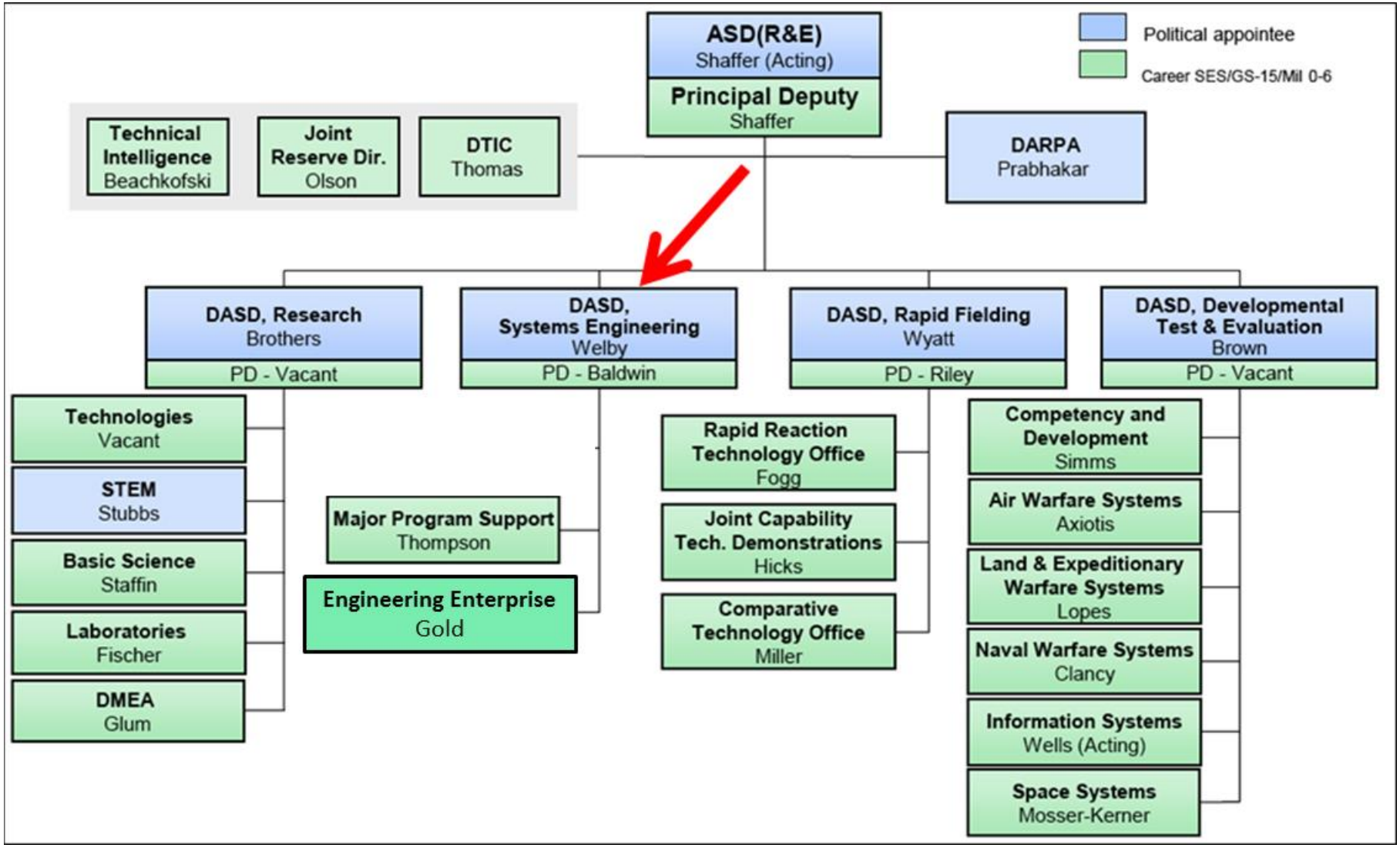
Where DASD(SE) sits in OSD



** All positions shown are Presidentially Appointed, Senate-confirmed (PAS) except those with * which are SES positions*
*** As of February 2013*

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

ASD(R&E) Hierarchy

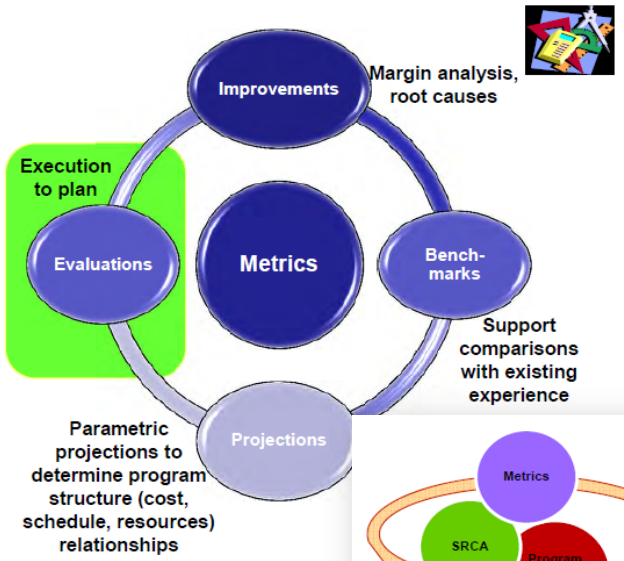


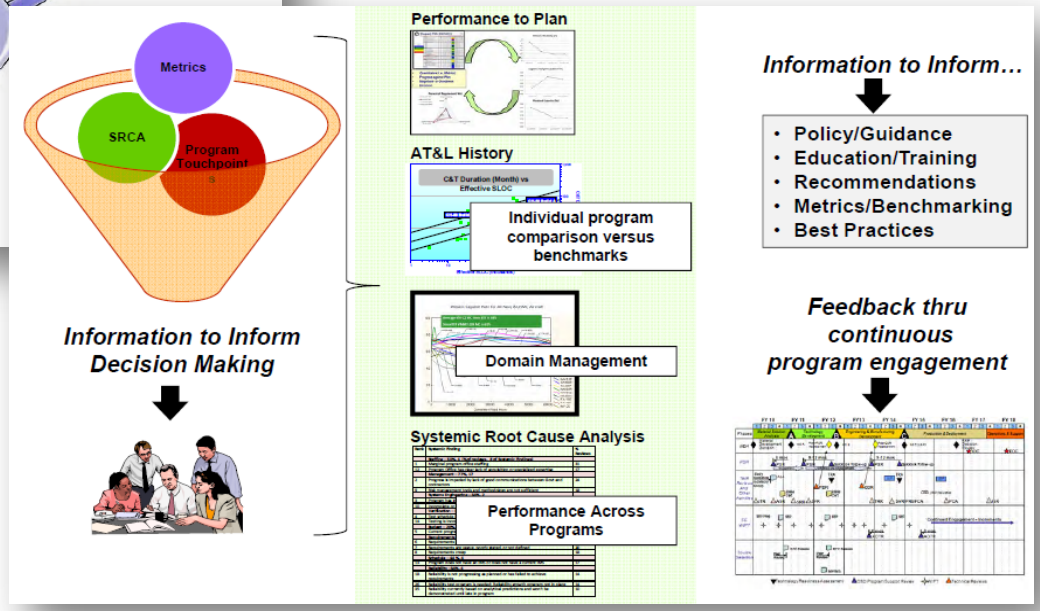
Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

D,MPS Mission

"Inform Decision Makers to Understand and Mitigate Risk"
- Better Buying Power 3.0

- Emphasize quantitative understanding consistent with Industry practice of systems engineering
- Make visible relationships between system/equipment design objectives and performance
- Harness and use existing information for timely and better decisions at the appropriate levels
- Enable data-driven decisions





"To meet Better Buying Power goals to inform decision makers of program risk, we need to ***characterize a system's technical maturity*** *and* ***quantitatively assess its readiness at key knowledge points*** *in the development life cycle.*

-Mr. James Thompson, Director, MPS

Reduce Risk: Data-driven Decisions at Every Level.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Recommended Metrics for SW Development

- **Requirements**
  - Periodic (e.g. monthly) & Cumulative volatility
  - Total # SW requirements
  - Requirements deferred to later builds
  - SW requirements growth
- **Technical Performance**
  - Metrics related to KPPs, KSAs, MOEs, MOPs
  - Latency by CSCI and aggregate
  - CPU utilization
  - Bandwidth usage as percent of planned hardware capacity
  - Quality attributes (-ilities)
  - End-to-end mission performance (e.g., time to perform mission critical function, # simultaneous users)
- **Cost**
  - EVM
  - Tradeoff analysis (as related to performance and schedule)
- **Risk**
  - Risk Burn Down
  - Risk Cube
  - Risk Curve
- **Progress/Schedule**
  - Size (Planned vs. Actual)
  - SW build completion date (plan v. actual)
  - Capabilities/features (plan v. actual)
  - SW requirements verified (plan v. actual)
    - Safety-critical SW requirements
  - Requirements documents status (e.g., # ICDs defined)
  - Design artifacts (e.g., # use cases complete—actual v. plan)
  - Tests (completed v. planned)
  - Productivity (planned v. actual)
- **Resources**
  - Staffing levels (planned vs actual)
- **Product Quality**
  - Defects by severity/priority
  - Defect backlog
  - Defect/maturity target
  - Defects open and closed
  - Defects aging
  - Average resolution time in hours
  - Technical debt
  - Defects inherited/deferred by build
  - Life-cycle phase containment

Top 4 Metrics to Assess Readiness to transition from DT to OT

1. Capabilities met
2. Quality:
   a. Defect Profile Risk Assessment
   b. Defect Backlog (No Pri 1 or 2 defects)
   c. Key Quality Attributes "-ilities"

Common set of metrics for all programs.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Recommendations
Performance Measures for OT Readiness

1. Capabilities met

- Status of end-to-end performance / SW-related mission thread TPMs (from SEP)
  - Go/No-go criteria: all measures have met threshold; Present any “non-green” measures
  - Examples: Time to decision, time to perform mission critical function, latency by CSCI and aggregate, bandwidth usage as percent of planned hardware capacity, # simultaneous users
- Status of MOEs/MOSs/CTPs (from TEMP)
  - Go/No-go criteria: all measures have met threshold; Present any “non-green” measures
    - Related measures already linked to TEMP
    - Status of KPPs/KSAs/TPMs
    - Requirements verification status [Go/No-go : X% of all operational & system requirements met]

Tester-developer disconnect & the “SE trace”

TPMs should trace to the relevant requirements (operational and system), KPPs, KSAs, and capability measures (e.g. MOEs). Program managers will ensure consistency between system TPMs and operational Critical Test Parameters.

2. Quality

a. Risk Assessment based on Defect Profile

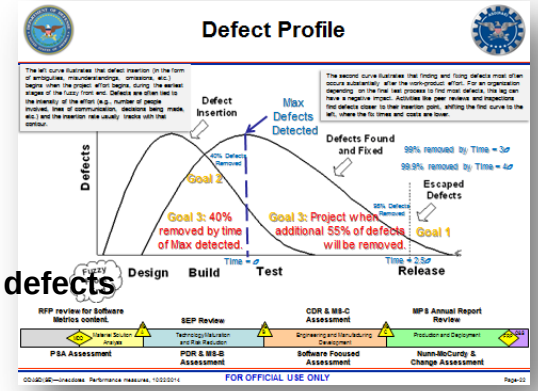
- Defect profile – discovered vs. closed over time (includes defect backlog)
- Other defect information to understand maturity risk: e.g. defect aging
- Go/No-go criteria
  - » Developmental test readiness level: 95% of expected defects have been found & fixed
  - » Operational test readiness level: 99% of expected defects have been found & fixed

b. Defect Backlog

- No open Priority 1 or 2 (showstopper) defects
- No excessive (user-acceptable) number of open Priority 3 & 4 (workaround/nuisance) defects

c. Key Quality attributes on track

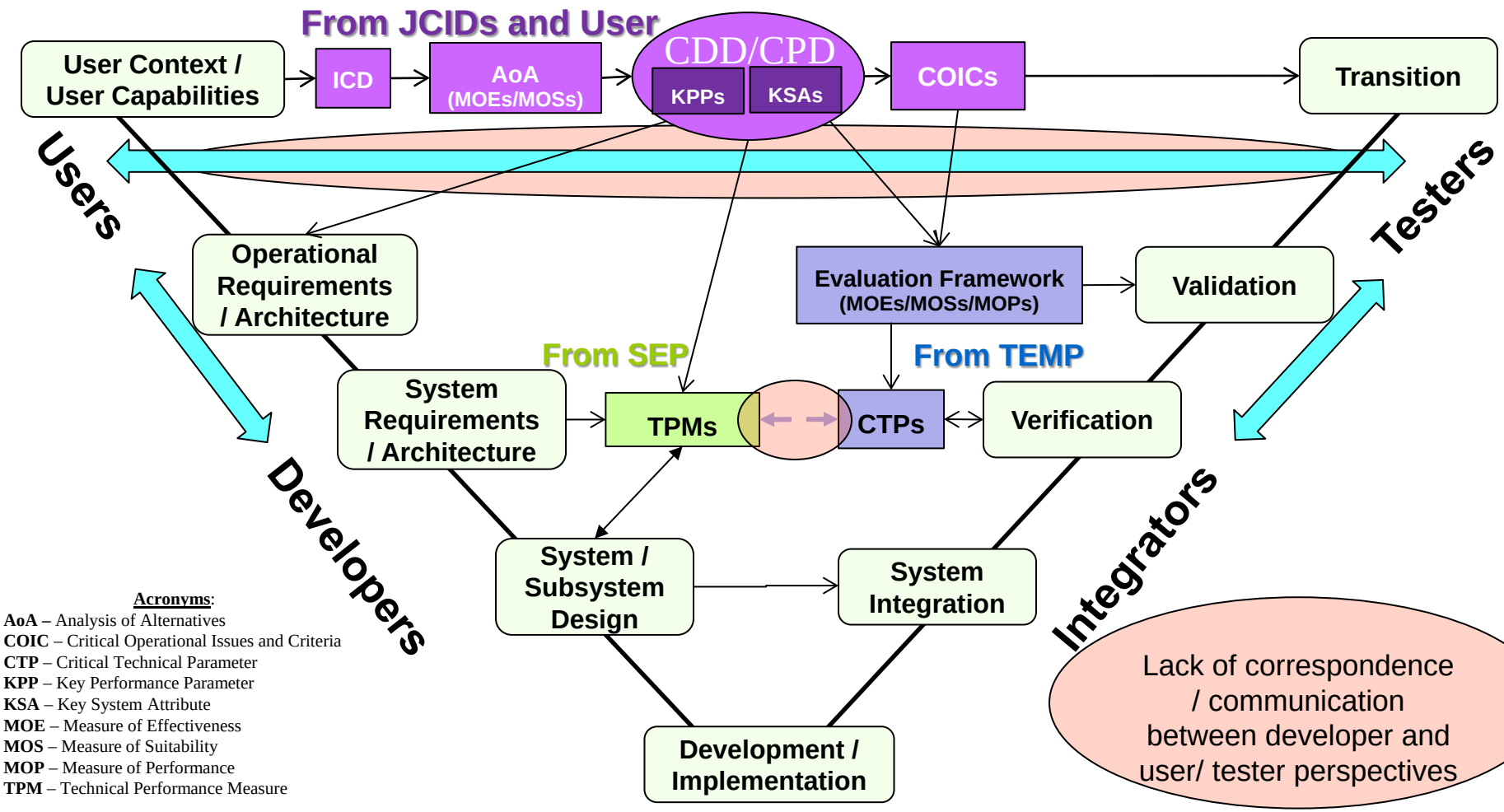
- e.g. System stability (# anomalies per test hour)



For transitioning from DT to OT, the 2-4 metrics are related to results of DT testing.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Performance Measurement Disconnects What can (and does) happen?

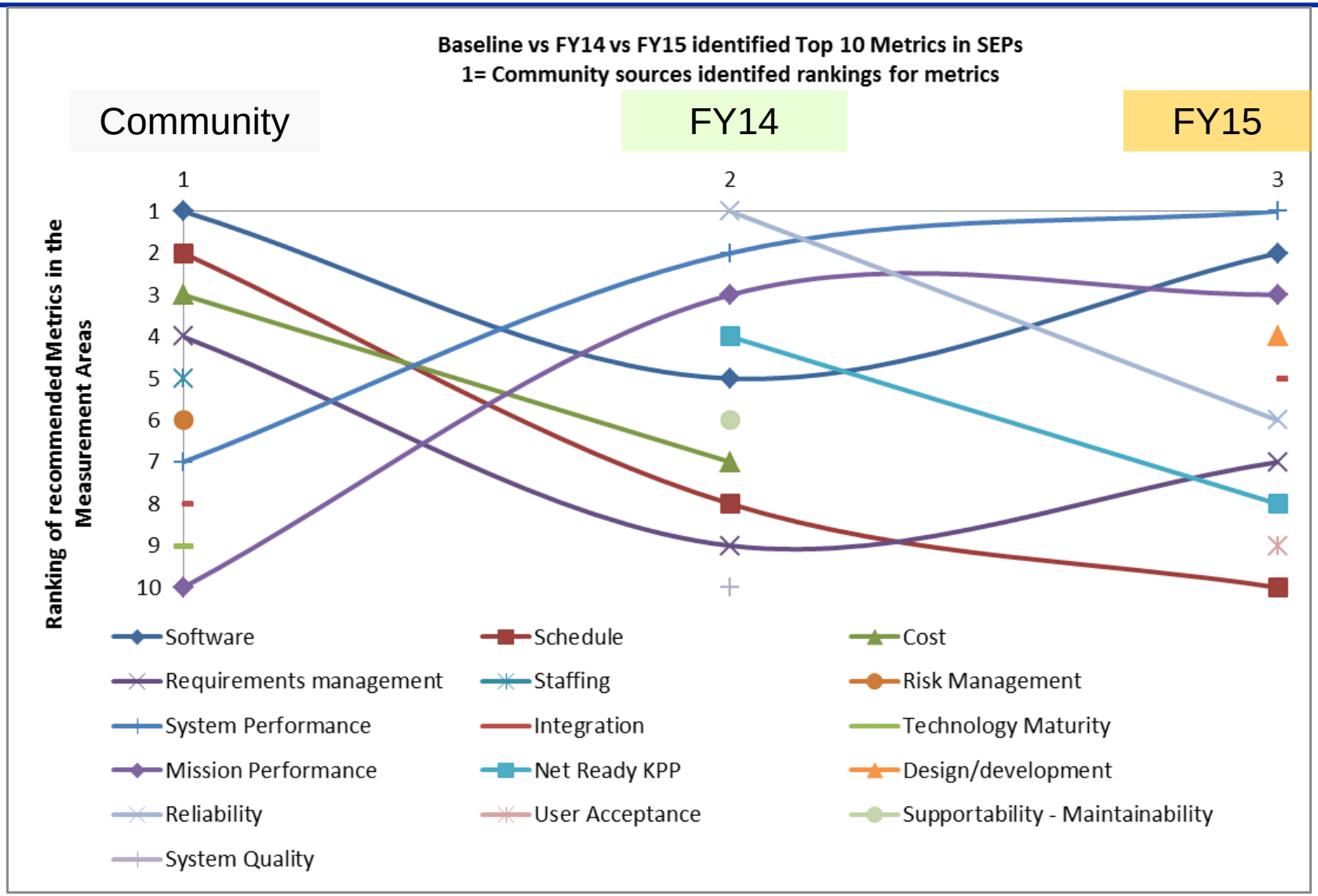


Acronyms:

AoA – Analysis of Alternatives
COIC – Critical Operational Issues and Criteria
CTP – Critical Technical Parameter
KPP – Key Performance Parameter
KSA – Key System Attribute
MOE – Measure of Effectiveness
MOS – Measure of Suitability
MOP – Measure of Performance
TPM – Technical Performance Measure

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Top 10 Metric Categories Over Time
Community vs. FY14 vs. FY15



Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

US Department of Defense (DoD)

- **World's largest engineering organization**
  - Over 99,000 Uniformed and Civilian Engineers
  - Over 39,000 in the Engineering (ENG) Acquisition Workforce

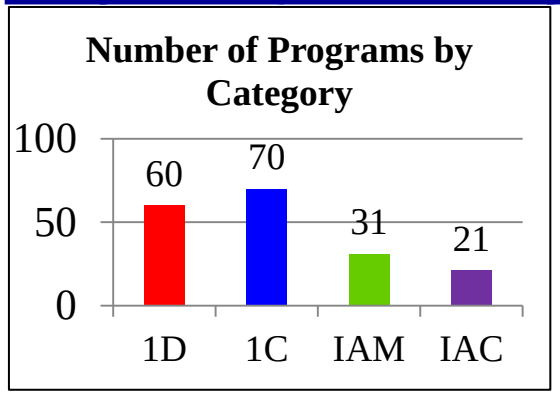
- **DoD Systems Engineering focuses on engineering excellence**
  - Design, develop, construct and operate complex systems
  - Forecast their behavior under specific operating conditions
  - Deliver their intended function while addressing economic efficiency, environmental stewardship and safety of life and property

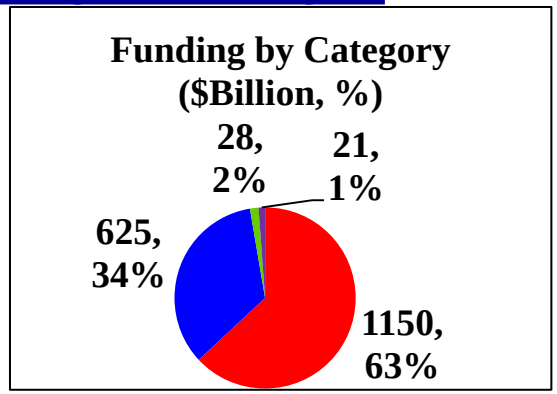
A Robust Systems Engineering Capability Across the Department Requires Attention to Policy, People and Practice

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Preponderance of Acquisition Funding is for ACAT 1D and 1C Programs

- **97% of acquisition funding is in MDAP 1D and 1C programs**
- **Software applications are components or sub-components of large, complex systems**
- **Software must:**
  - Support system / component / sub-component **requirements**
  - Support overall program / component / sub-component **schedules**
  - Support **integration** with other system software and hardware
- **Software acquisition for 1D and 1C programs poses some of our <u>toughest systems engineering challenges</u>**



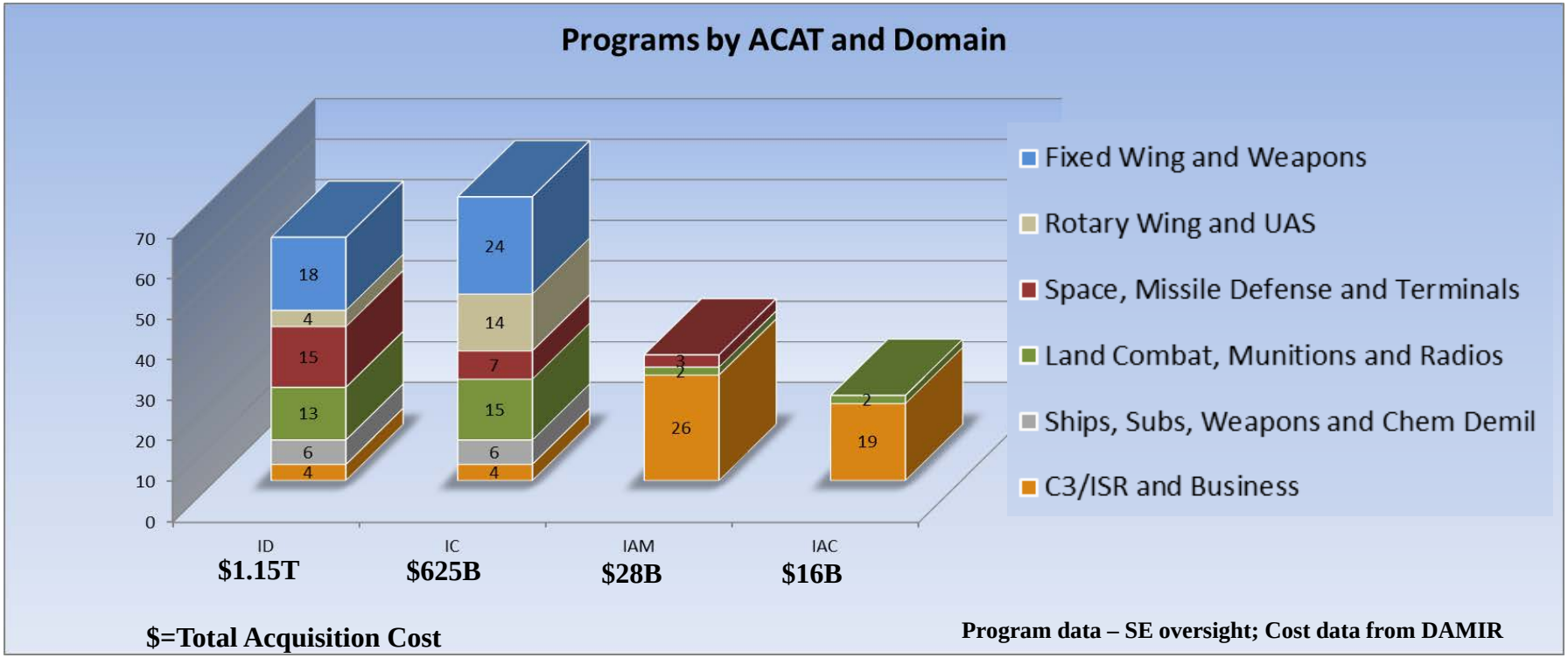


Program data – SE oversight; Cost data from DAMIR

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

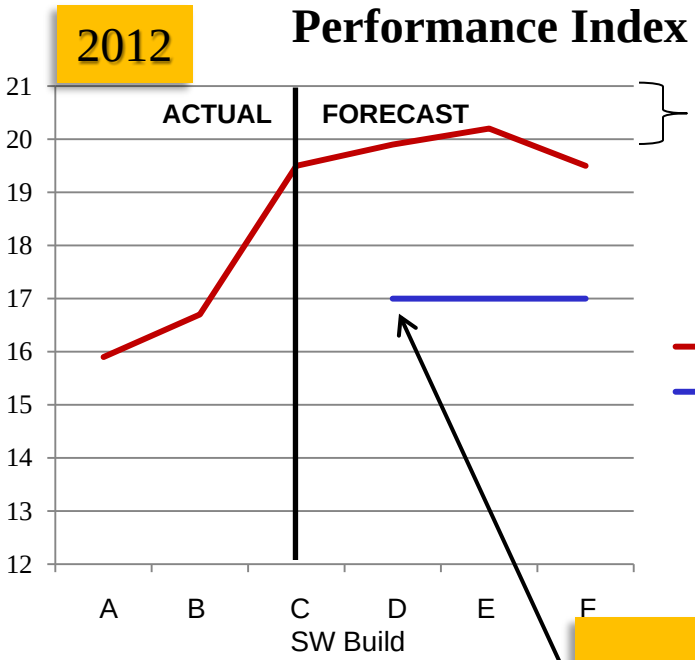
DASD(SE) Portfolio

- System Engineering oversight of 182 Programs with acquisition costs of $1.8T
- Approve System Engineering Plans (SEPs)
- Assess preliminary and critical design reviews (PDR, CDR)



Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.

Productivity Analysis
PMO Plan and OSD Estimates

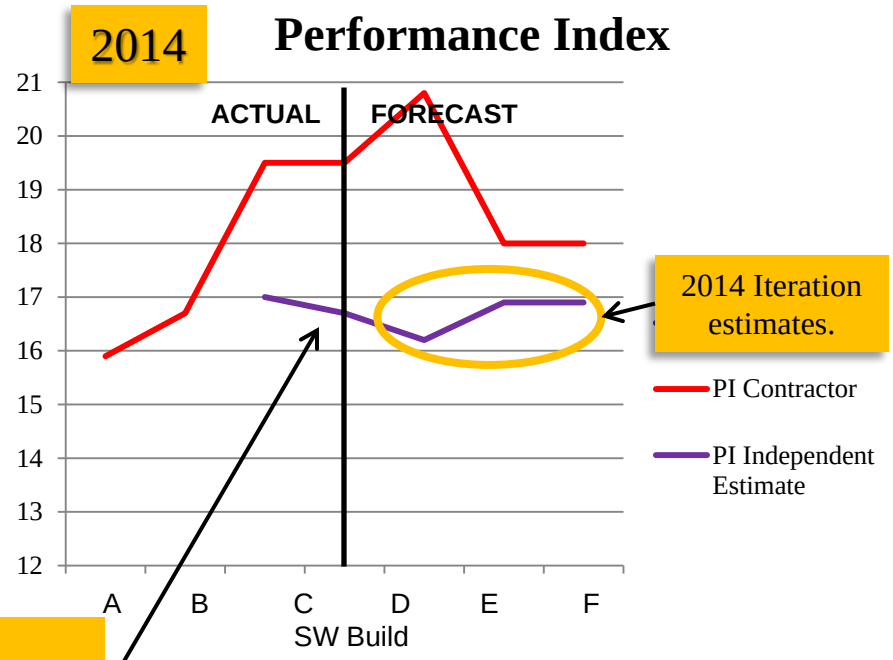


One Index Point which represents a gain of 1.27 times in process productivity.

3 index point difference (1.27^3) doubles (or halves) process productivity.

Contractor

Independent Estimate



SW Build PI
2012 independent estimate = 17
Actual as of Jun 2014 = 16.7

"So-what": OSD's quantitative estimates of contractor performance in 2012 was very close to 2014 actual performance.

Distribution Statement A – Approved for public release by DOPSR on 6/4/2015, SR Case # 15-S-1638 applies. Distribution is unlimited.