Army Software Sustainment Cost Estimating Results

DASA-CE

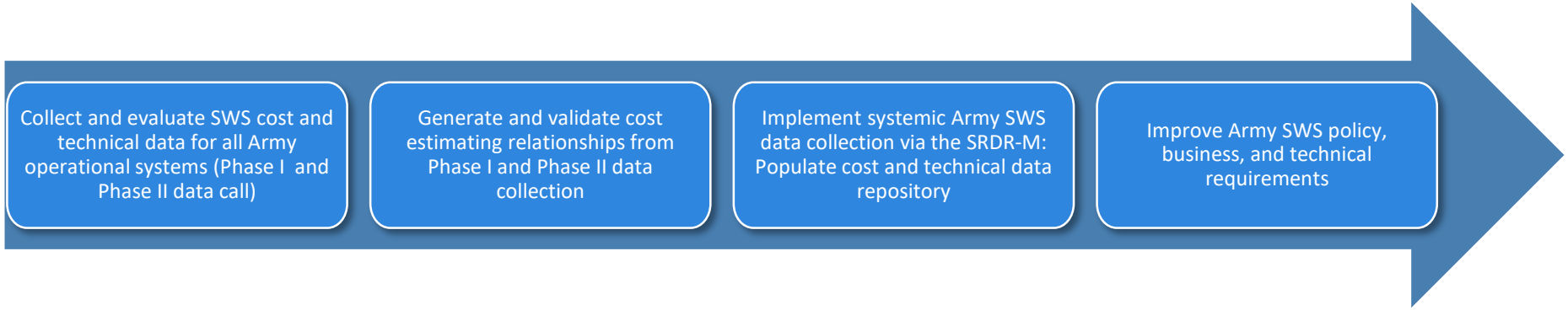
IT CAST

August 2019

SWS Initiative Objective and Strategy

Accurately estimate Army system Software Sustainment (SWS) costs to:

- Effectively project and justify software and system life cycle costs
- Objectively evaluate Army system software sustainment execution costs
- Inform and optimize the allocation of available sustainment resources across the Army



Effective software sustainment cost estimation is the basis for Army system software life cycle cost management

Executive Summary - Accomplishments

- Established Software Sustainment Data Collection Mechanisms
  - Army Software Data Collection Questionnaire
  - SRDR for Maintenance
  - Software Sustainment WBS Used to Collect Sustainment Costs
- Created Comprehensive Software Sustainment Data Repository
  - 192 Systems
  - 700+ Capability Releases
  - 300+ IAVA Releases
  - 3,200+ records on software license data
- Established Robust Foundation for Software Sustainment Fact-Based Decisions
  - Allocations of Costs by WBS Elements
  - Cost & Schedule Estimating Relationships
  - Benchmarks
- Data and Analysis Results provided to DoD and Army Community
  - Benchmarks and CERs Ready for Use

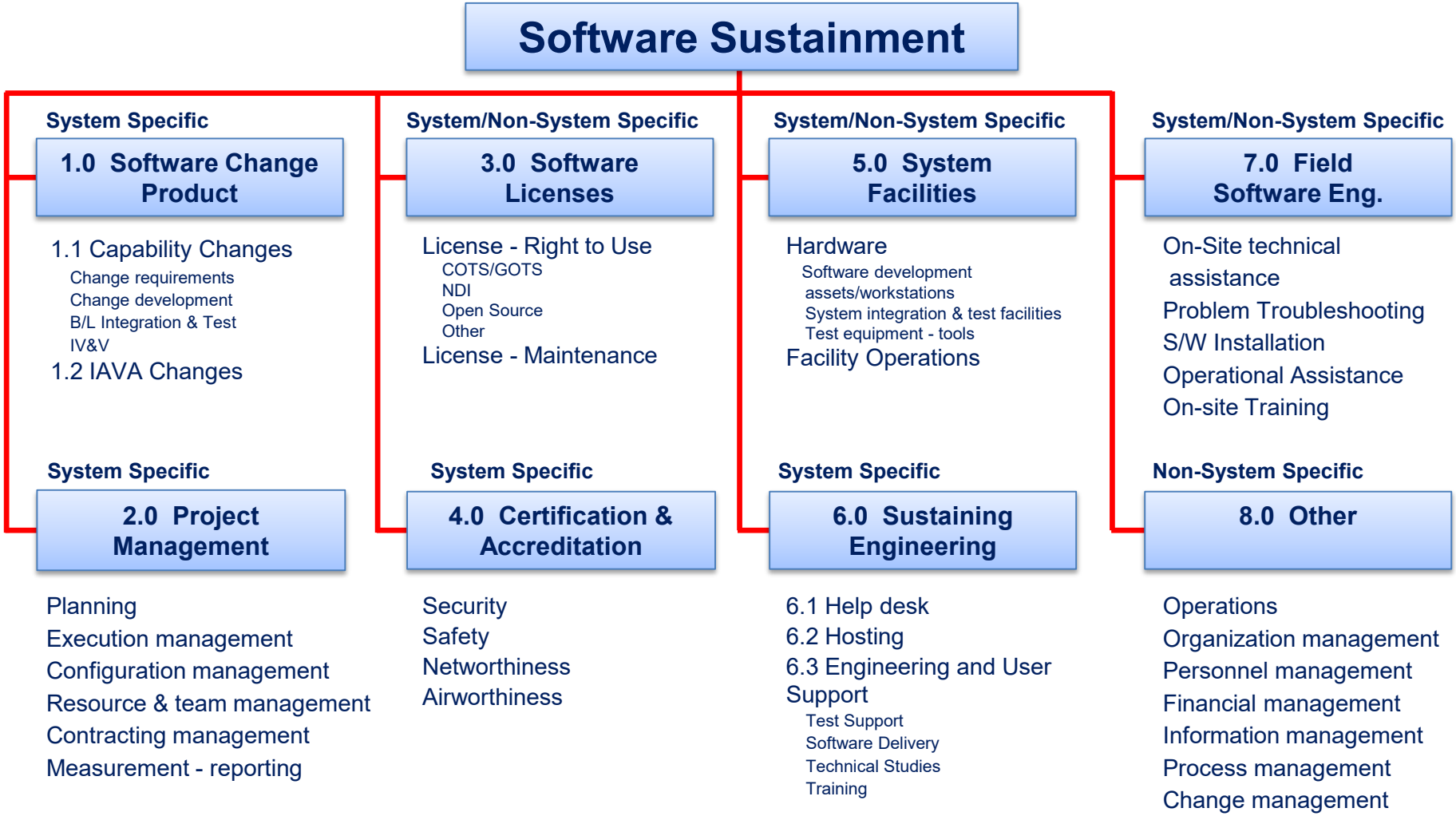
Decision Information

- Decision information must objectively tie investment costs to software product mission capability

- Program-level management must decide
  - Which baseline change requirements to implement
  - Prioritization of capability, maintenance, and security changes
  - Delivery strategy for incremental software releases

- Enterprise-level management must decide
  - Prioritization of resources across the operational system portfolio
  - Tradeoffs between funding and associated mission capability

DASA-CE SWS WBS



Version 5.0

Army Software Sustainment Definition

- Software sustainment (SWS) includes all software change activities and products associated with modifying a software system after a software release has been provided to an external party
- The release is the primary SWS change product - a composite of one or more changes - it can be either a formal release or an engineering release
- SWS includes software enhancements, software maintenance, and cybersecurity updates
- Software maintenance includes defect repair, rehosting, adaptations, updates, and reconfiguration
- SWS may be funded by multiple funding sources
- Costs include both Fixed and Variable costs accrued at both the system and organizational levels
- Costs include both organic (government) and contractor resources

Software Sustainment Data Characterization

Data Demographics

Overview

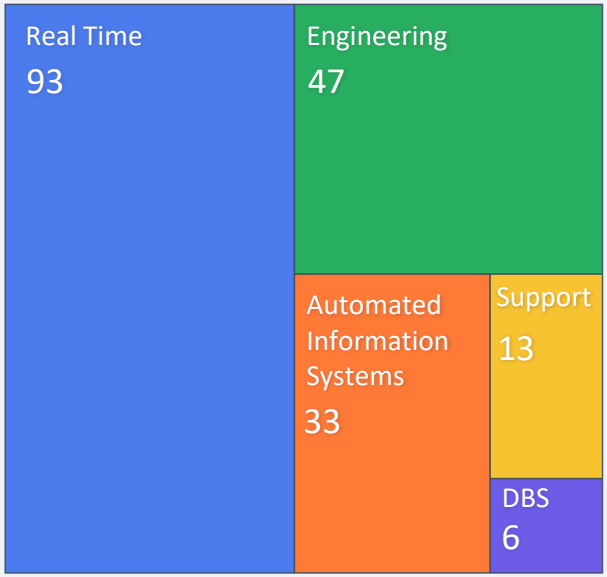
192 Systems

1,040 Releases

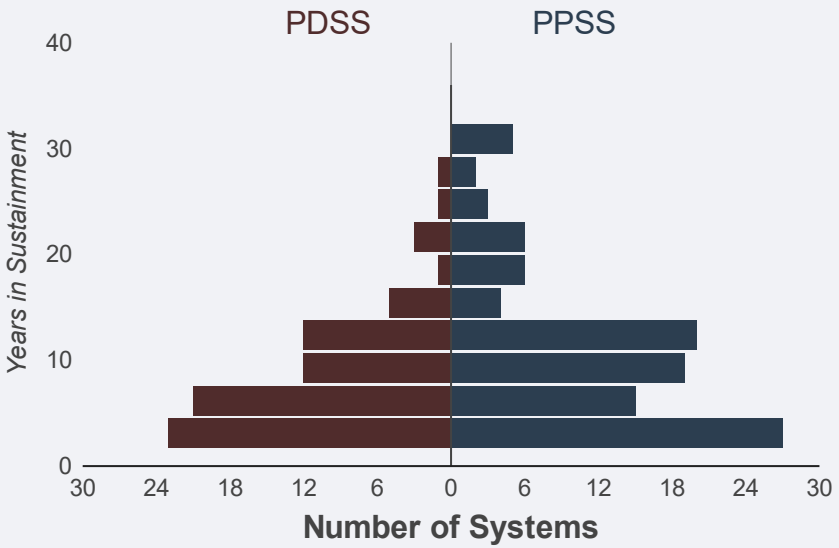
3,434 License Records

411K Data Fields

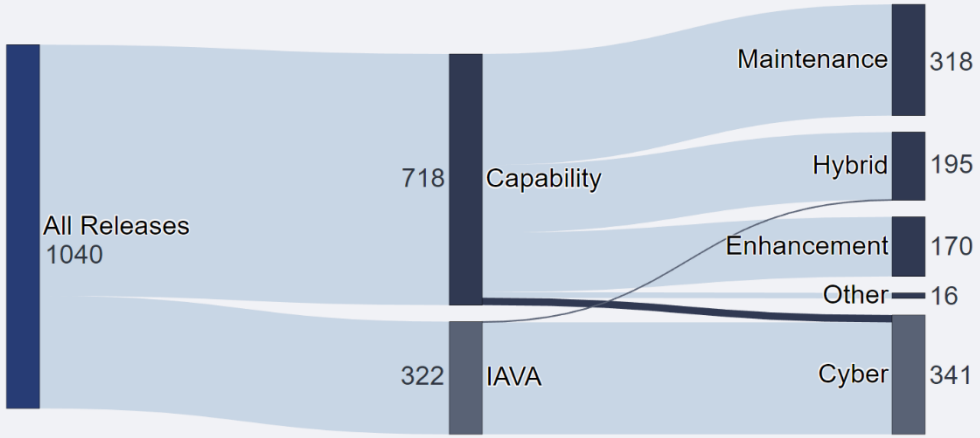
Systems by Super Domain



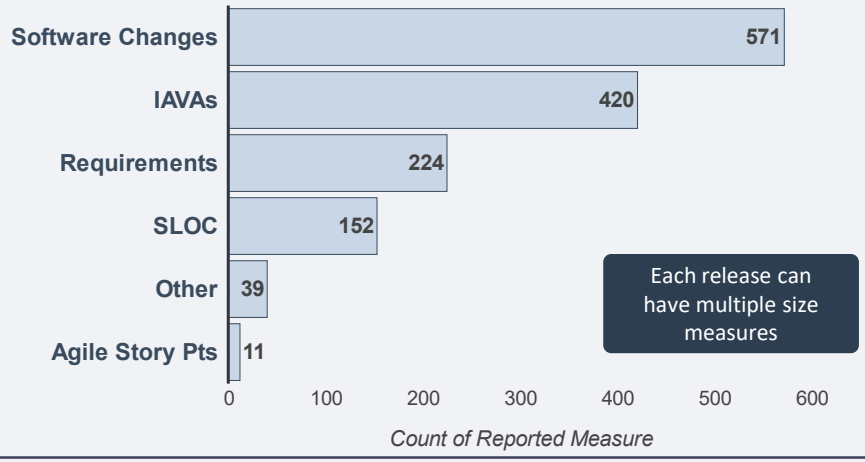
Distribution of System Age



Releases by Change Type

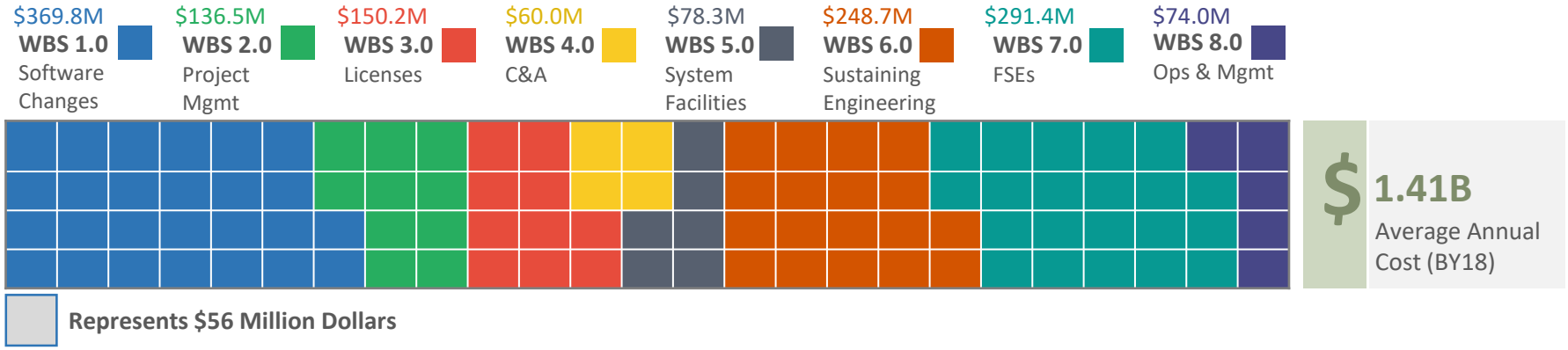


Releases by Size Measure

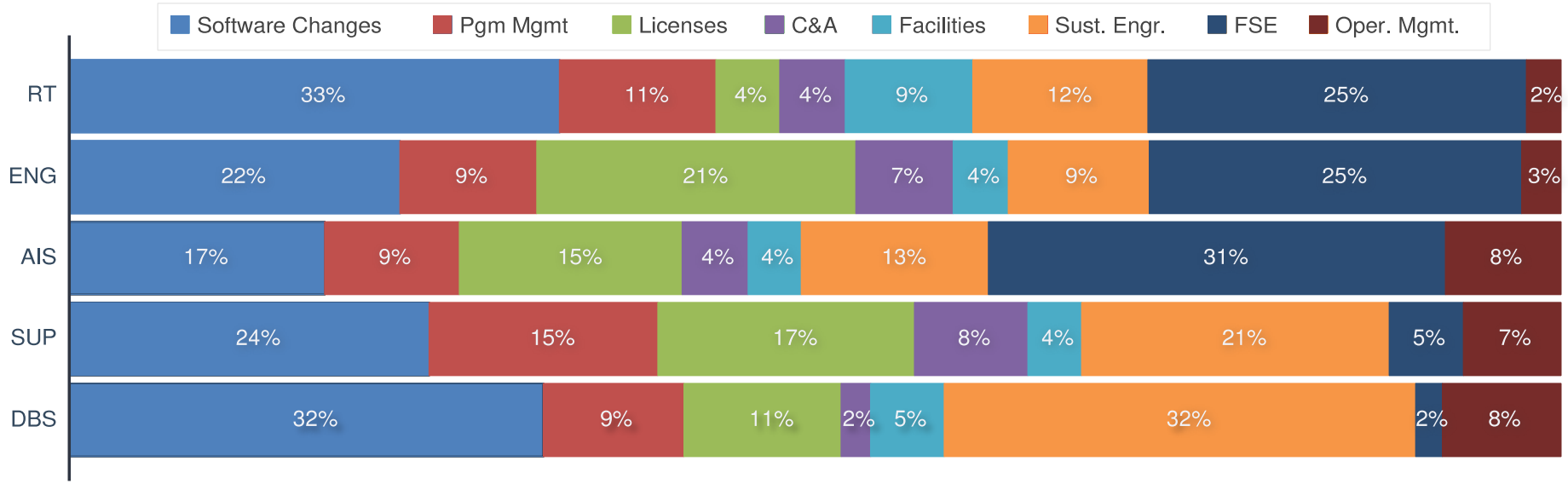


Annual Cost Distribution

Average Annual Cost by WBS & Cost Allocation by WBS



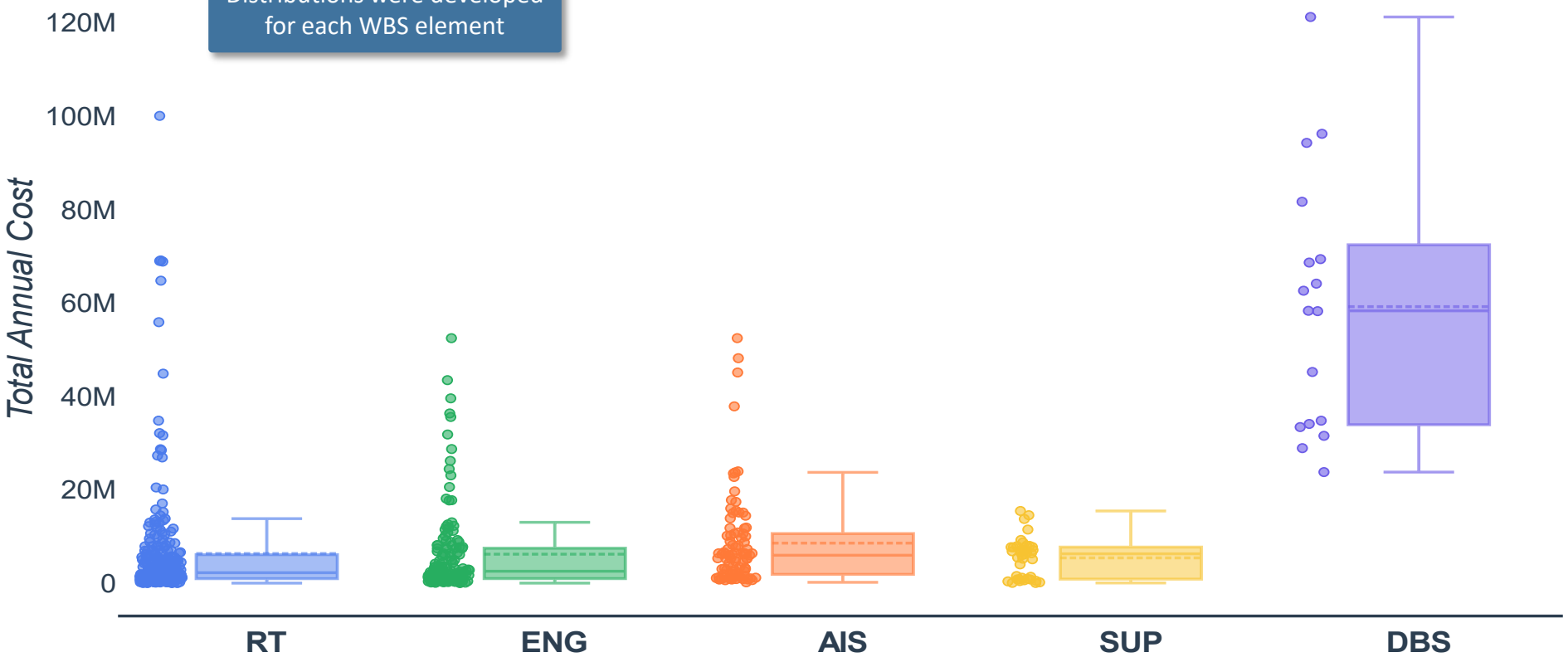
Annual Cost Allocation by WBS



Total Annual Cost Distribution

By Super Domain (BY18$)

Distributions were developed for each WBS element

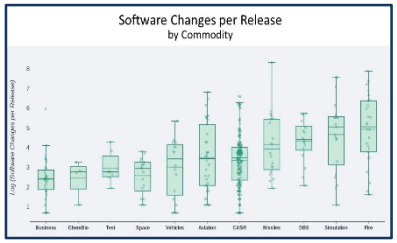


Super Domain	RT	ENG	AIS	SUP	DBS
Count*	240	136	91	40	17
Mean	$6,351,776	$6,202,735	$8,572,860	$5,411,627	$59,221,413
Median	$2,237,870	$2,547,140	$5,972,457	$6,285,354	$58,334,452

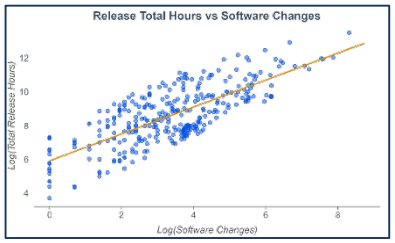
* Up to 3 FYs per system

Specific Analysis Overview
(full detail is available)

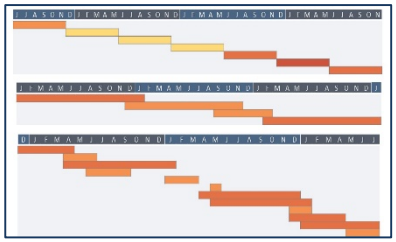
- Benchmarks for capability releases by super domain and commodity
  - # of software changes/release
  - Hours per software change
  - By WBS
  - PDSS vs. PPSS for application super-domains
  - By sustaining organization
  - DSLOC per FTE



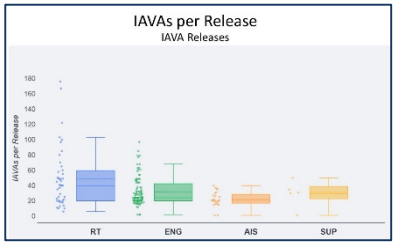
- Cost estimating relationships (CERs) for capability releases
  - Evaluated meta data for impacts on CERs: commodity, change type, # of inter-service partners, and ACAT levels had an impact
  - Data Trimmed and CERs developed
    - CERs for Software Changes (most effective), Requirements, and Lines of Code



- Schedule estimating relationships (SERs) for capability releases
  - Initial SERs had low correlation
  - Data segmented into schedule approaches:
    - Cyclic, Sequential, and Concurrent

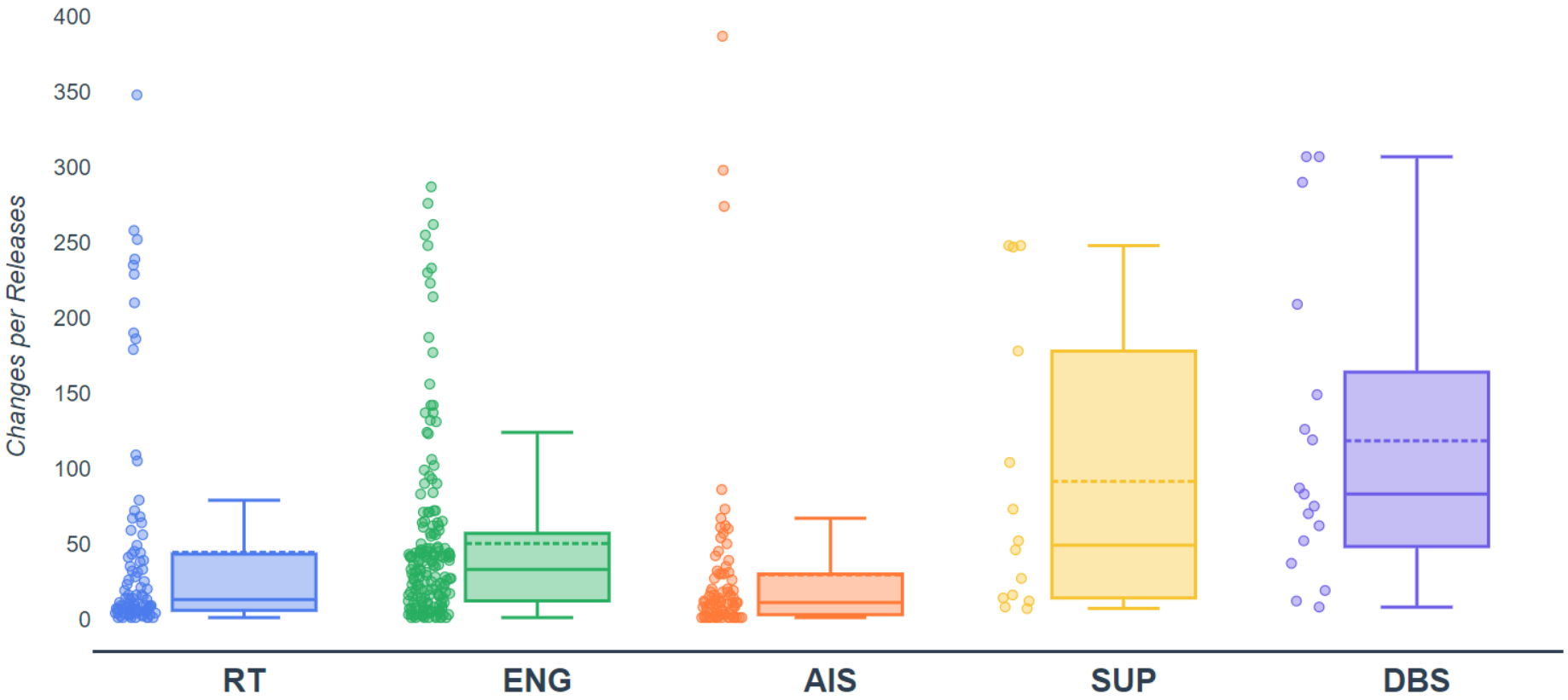


- IAVA Release Analysis
  - Most data is Level-of-Effort
  - Data best described by median by grouping for # of IAVAs and hours per IAVA



Benchmarks for Capability Releases

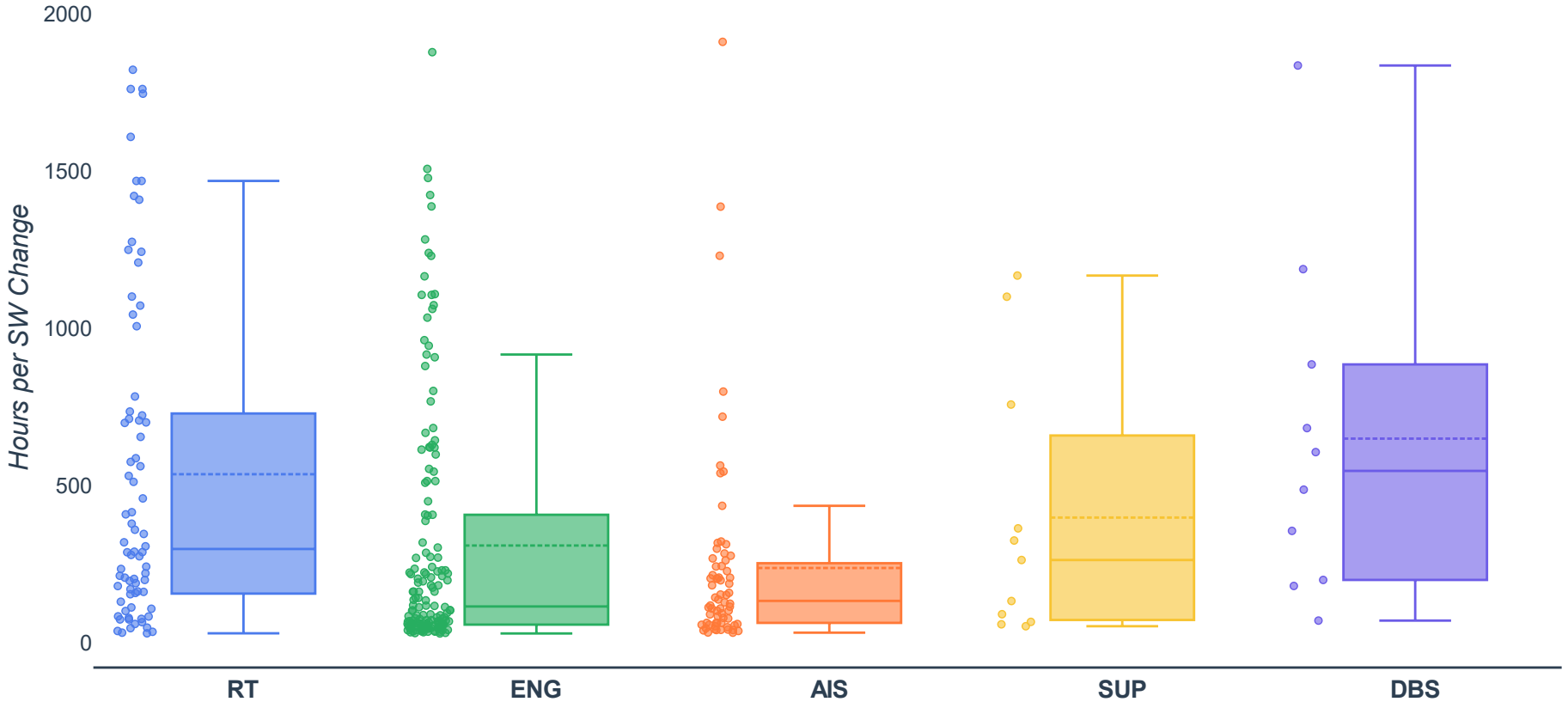
Number of Software Changes per Release
By Super Domain



Super Domain	Count	Mean	Median
RT	89	44	13
ENG	171	50	33
AIS	79	29	11
SUP	14	91	49
DBS	17	118	83

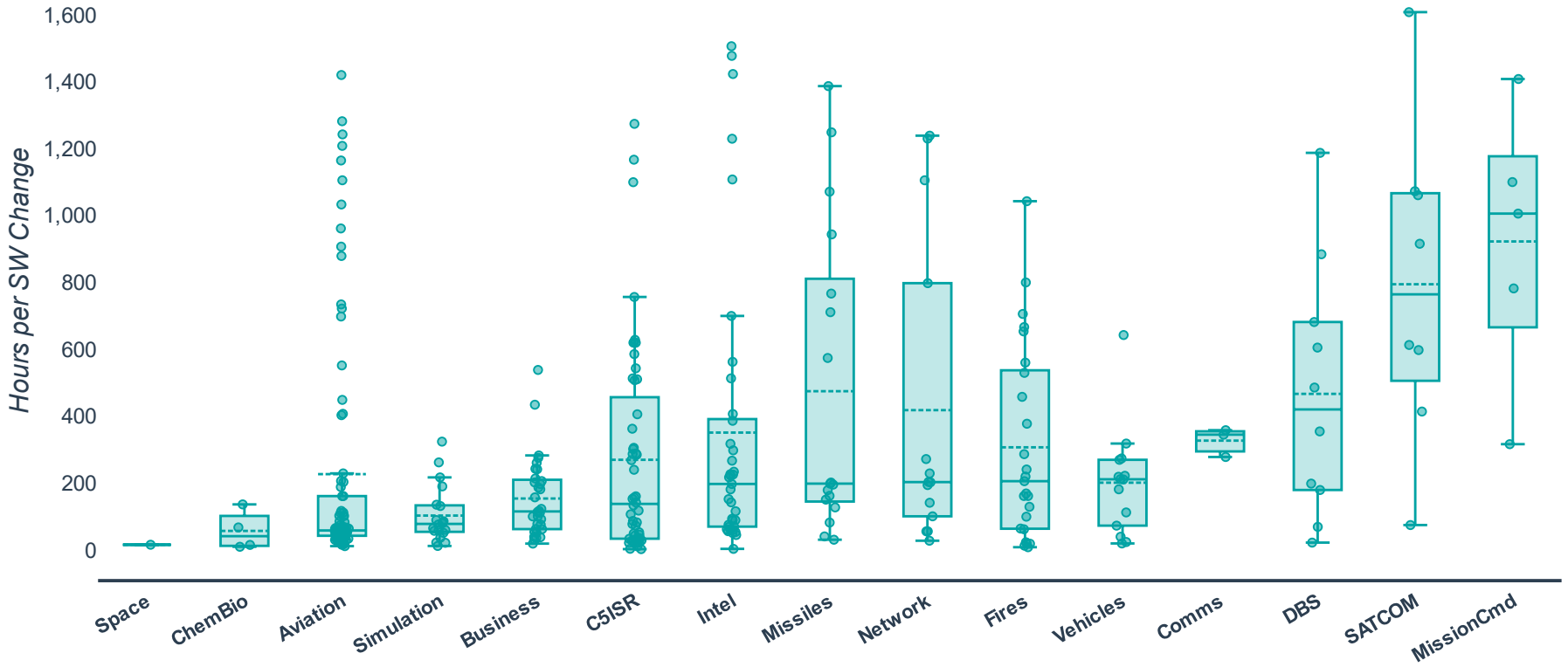
Number of SW Changes/Release can be used to size future releases when program specific data is unknown - The resulting size can be used with the associated cost benchmark or put into a CER

Hours per Software Change
By Super Domain



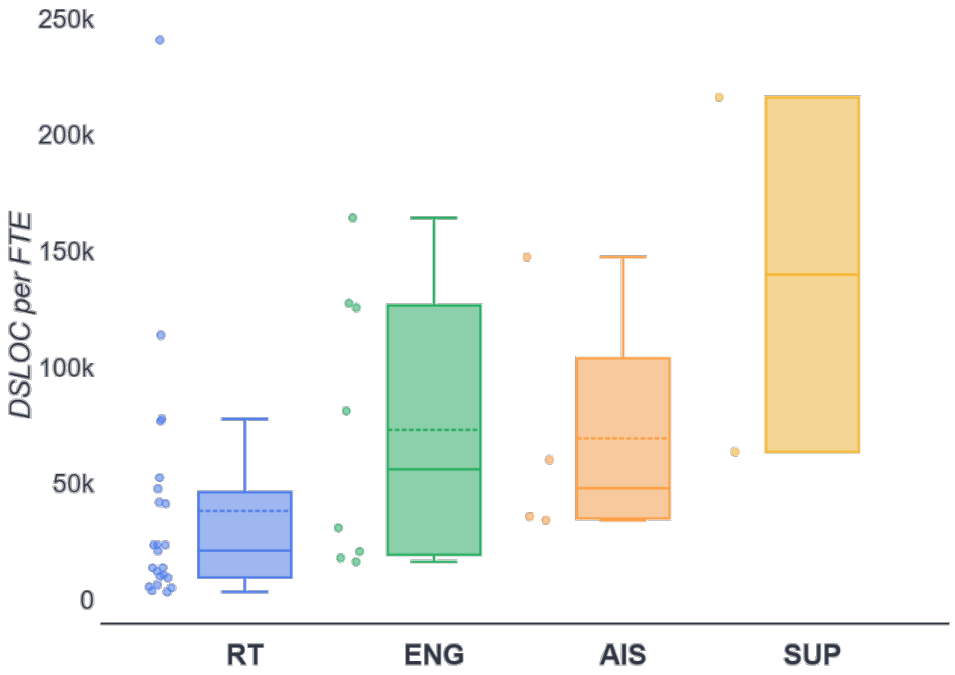
Super Domain	Count	Mean	Median
RT	101	1481.6	407.4
ENG	186	629.8	102.5
AIS	79	244.9	111.5
SUP	14	490.7	197.0

Hours per Software Change
by Commodity



	Space	ChemBio	Aviation	Simulation	Business	C5ISR	Intel	Missiles	Network	Fires	Vehicles	Comms	DBS	SATCOM	MSN CMD
Count	1	4	87	20	32	52	37	17	14	25	14	3	10	8	5
Mean	16.1	57.5	227.2	103.3	154.3	270.0	351.3	475.0	418.4	307.3	201.6	327.3	467.0	794.6	922.5
St Dev	--	58.8	360.3	84.1	118.4	312.7	432.7	456.5	459.0	291.0	160.4	42.9	377.3	471.8	406.5
Median	16.1	41.7	58.8	78.4	115.6	138.2	197.8	198.8	203.3	206.1	211.7	345.1	420.4	764.4	1005.7

DSLOC per FTE
By Super Domain



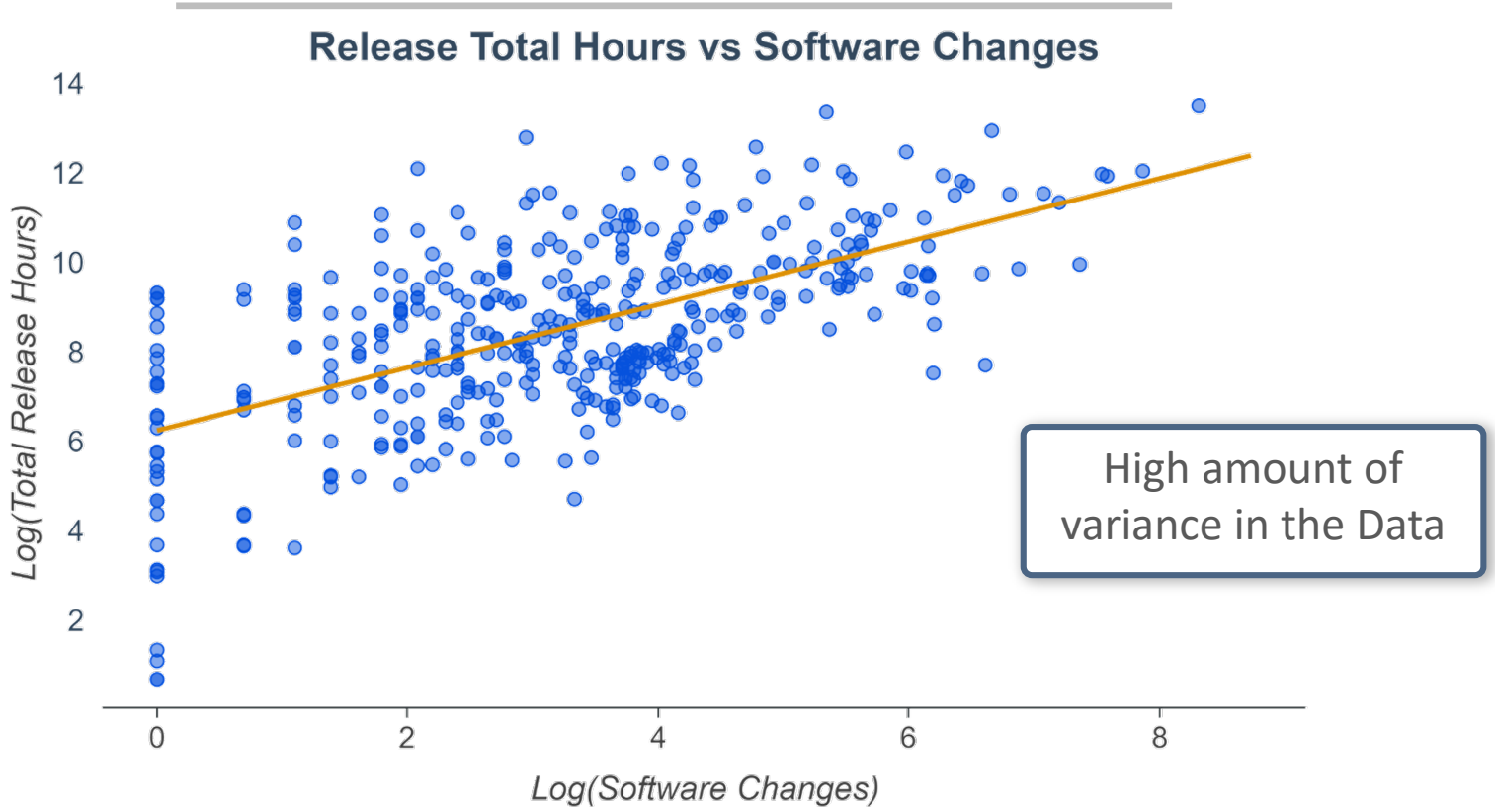
Super Domain	AIS	ENG	RT	SUP
Count	4	8	23	2
Mean	69,492	73,166	38,306	139,953
Min	34,307	16,424	3,496	63,754
1Q	35,486	20,138	9,892	101,853
Median	48,094	56,181	21,221	139,953
3Q	82,099	126,185	44,916	178,052
Max	147,473	164,340	240,813	216,151

For a 100,000 DSLOC baseline, the estimated FTEs are RT=4.7, ENG=1.8, AIS=2.1 & SUP=NA

- DSLOC represents Delivered Source Lines of Code which counts all code equally
- The earliest baseline size reported was used to represent DSLOC
- Full Time Equivalent (FTE) counts were derived by including the following WBS Elements: SW Change Product (1.0), Program Management (2.0), Certification and Accreditation (4.0), and Sustaining Engineering (5.0)
- FTEs were derived by using labor hours per man-year and labor rate reported for each program

Cost and Schedule Estimating Relationships (CER/SER) Capability Releases

All Data CER



Equation

$$\log(Hours) = 0.703 \log SC + 6.2438$$

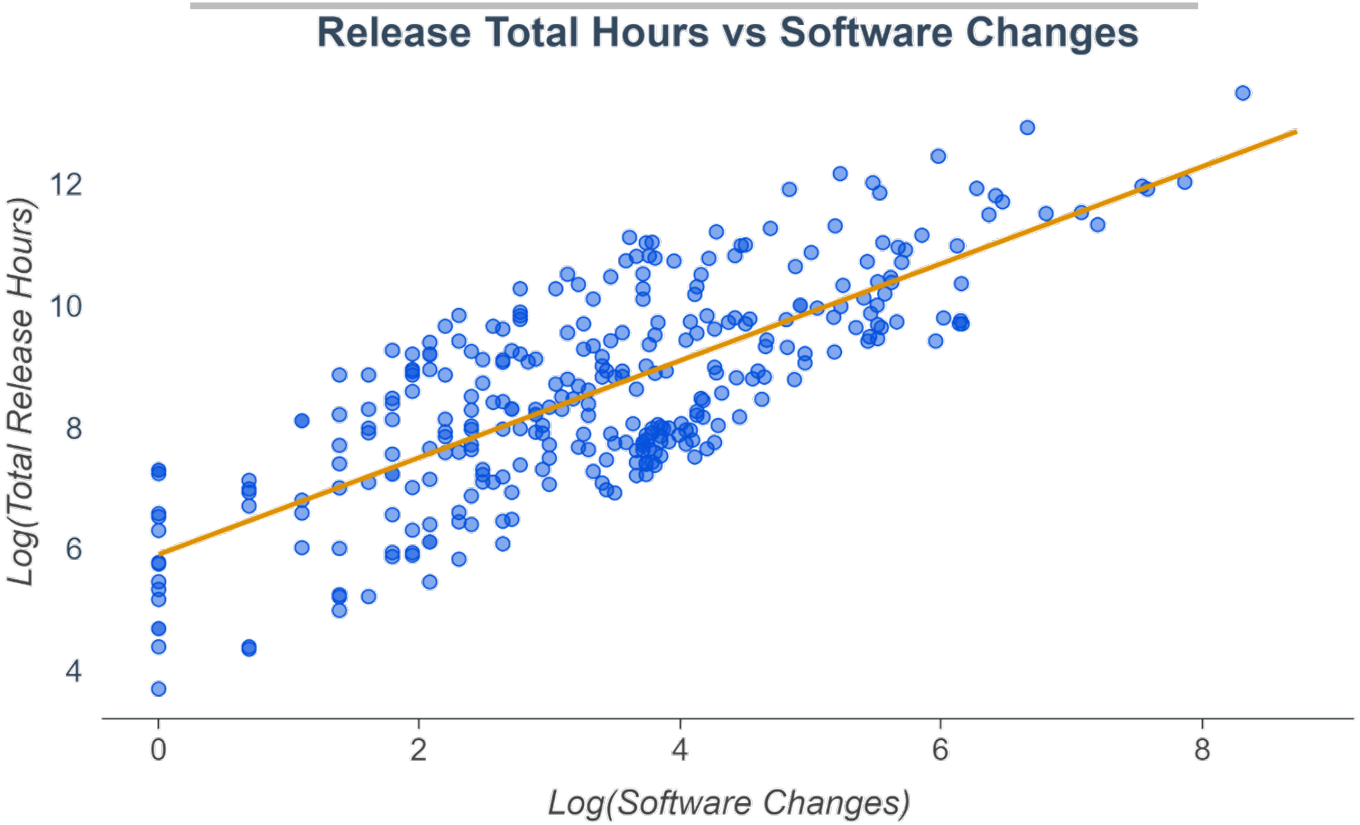
R-Squared	36.0%
Adj. R-Squared	35.8%
Observations	397

Variable	Coef.	Std Err	T-Value	P-Value
Constant	6.2438	0.172	36.231	0.000
Log(SC)	0.0730	0.047	14.900	0.000

Assumptions

- Removed records with:
  - Defense Business Systems (DBS) super domain
  - Hour data outliers or missing data
  - Records with no dependent variable, e.g., SW Change (SC) counts
  - Upper & lower 10% of records based on unit cost
- Both Dependent and Independent variables were transformed using $\log_{10}$
  - Zeros were represented with 0.1
- All categorical variables were represented as dummy variables (0,1)
- Adjusted R^2 was used for model performance comparisons

Trimmed Data CER*



Equation	
$\log(Hours) = 0.7981 \log SC + 5.905$	
R-Squared	57.2%
Adj. R-Squared	57.1%
Observations	317

Variable	Coef.	Std Err	T-Value	P-Value
Constant	5.9052	0.145	40.618	0.000
Log(SC)	0.7981	0.039	20.532	0.000

* Data records trimmed by 10%

Total Hours vs SW Changes -1

Model		Conditions	Obs	Adj R^2	SEE (Hrs)	PRED(30)
$THrs = 463 * (TSC)^{0.69}$		All data	329	0.36	48,385	17.3%
$THrs = 341 * (TSC)^{0.79}$		10% trimmed data	263	0.57	44,842	23.6%
AIS ENG RT SUP	$THr = 242 * (TSC)^{0.73}$ $THr = 386 * (TSC)^{0.73}$ $THr = 736 * (TSC)^{0.73}$ $THr = 698 * (TSC)^{0.73}$	10% trimmed & Super Domains (Categorical)	263	0.62	39,330	20.2%
Aviation Business C5ISR ChemBio Fire Missiles Simulation Space Test Vehicles	$THrs = 1,452 * TSC^{0.66}$ $THrs = 301 * TSC^{0.66}$ $THrs = 364 * TSC^{0.66}$ $THrs = 182 * TSC^{0.66}$ $THrs = 1,531 * TSC^{0.66}$ $THrs = 1,114 * TSC^{0.66}$ $THrs = 577 * TSC^{0.66}$ $THrs = 1,005 * TSC^{0.66}$ $THrs = 1,742 * TSC^{0.66}$ $THrs = 425 * TSC^{0.66}$	10% trimmed & Commodities (Categorical)	263	0.68*	40,886	23.2%
$THrs = 608 * (TSC)^{0.98} / (TReqts)^{0.21}$		10% trimmed	32	0.84	32,228	25.0%
$THrs = 330 * (TSC)^{0.97} / (TReqts\_Imp)^{0.11}$		10% trimmed	65	0.74	63,904	23.1%

* High P-Values for one or more coefficients

Total Hours vs SW Changes -2

Model		Conditions	Obs	Adj R^2	SEE (Hrs)	PRED(30)
$THrs = 296 * (TSC)^{0.94} / (EI\_Mod)^{0.11}$		10% trimmed	41	0.74*	47,326	22.0%
$THrs = 1,219 * (TSC)^{0.75} / (SWBase)^{0.04}$		10% trimmed	69	0.61*	36,567	26.1%
$THrs = 757 * (TSC)^{1.02} / (BL)^{0.36}$		10% trimmed	45	0.74	81,719	15.6%
Cyber Enhance Hybrid Maint Other	$THrs = 332 * TSC^{0.79}$ $THrs = 531 * TSC^{0.79}$ $THrs = 382 * TSC^{0.79}$ $THrs = 281 * TSC^{0.79}$ $THrs = 284 * TSC^{0.79}$	10% trimmed & Change Type (Categorical)	263	0.59*	39,573	21.3%
$THrs = 338 * TSC^{0.77} * Enh\%^{0.10} * Maint\%^{0.02} * Cyber\%^{0.03} * Other\%^{0.01}$		10% trimmed & percentages of Change Types	263	0.60*	26,494	6.8%

IAVA Release Analysis

IAVA Analysis

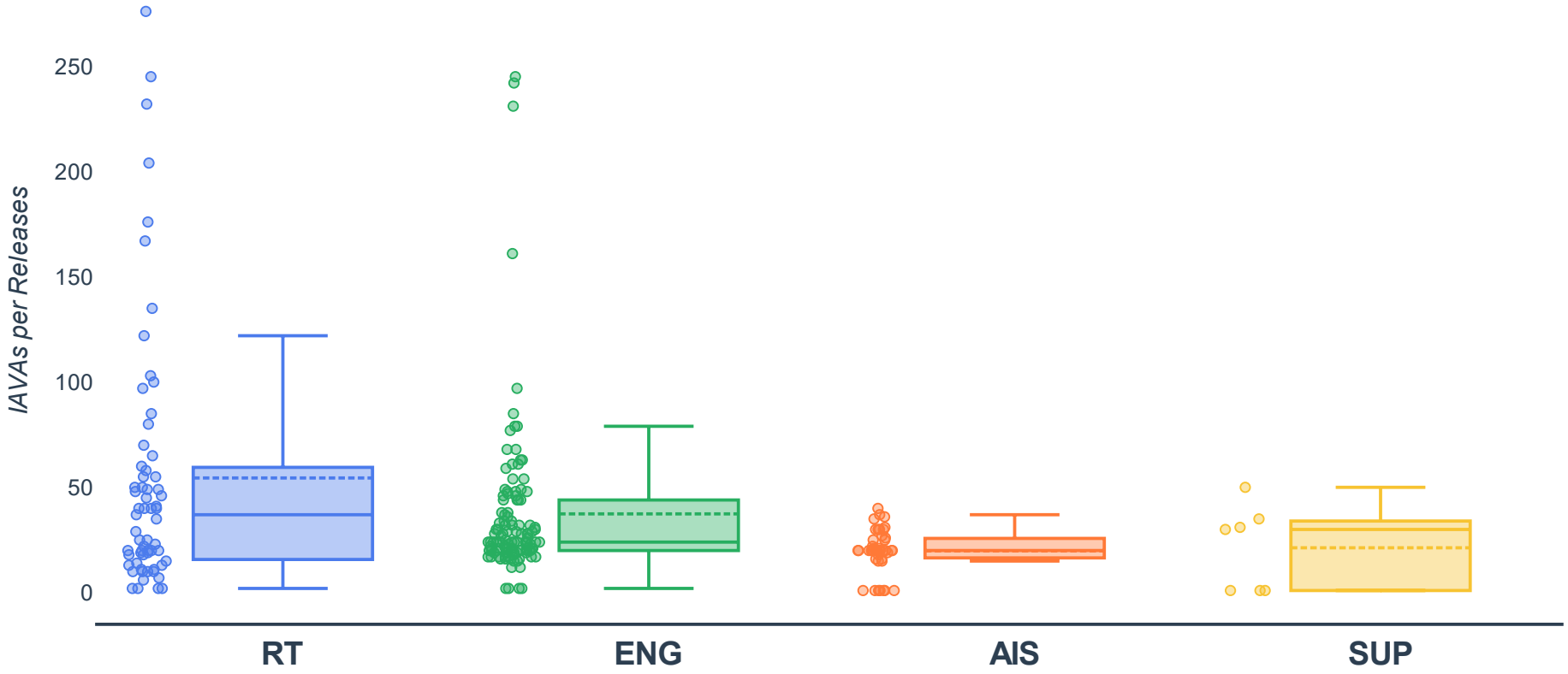
Group	Mean	Median
MaintOrg	243.7	24.2
System	164.0	25.6
Commodity	28.1	30.1
Super Domain	38.0	36.5

There is a central tendency across segmentation groups using the Median

IAVA data is better estimated using descriptive statistics i.e. average cost (hours per IAVA) as compared to regression

IAVAs per Release

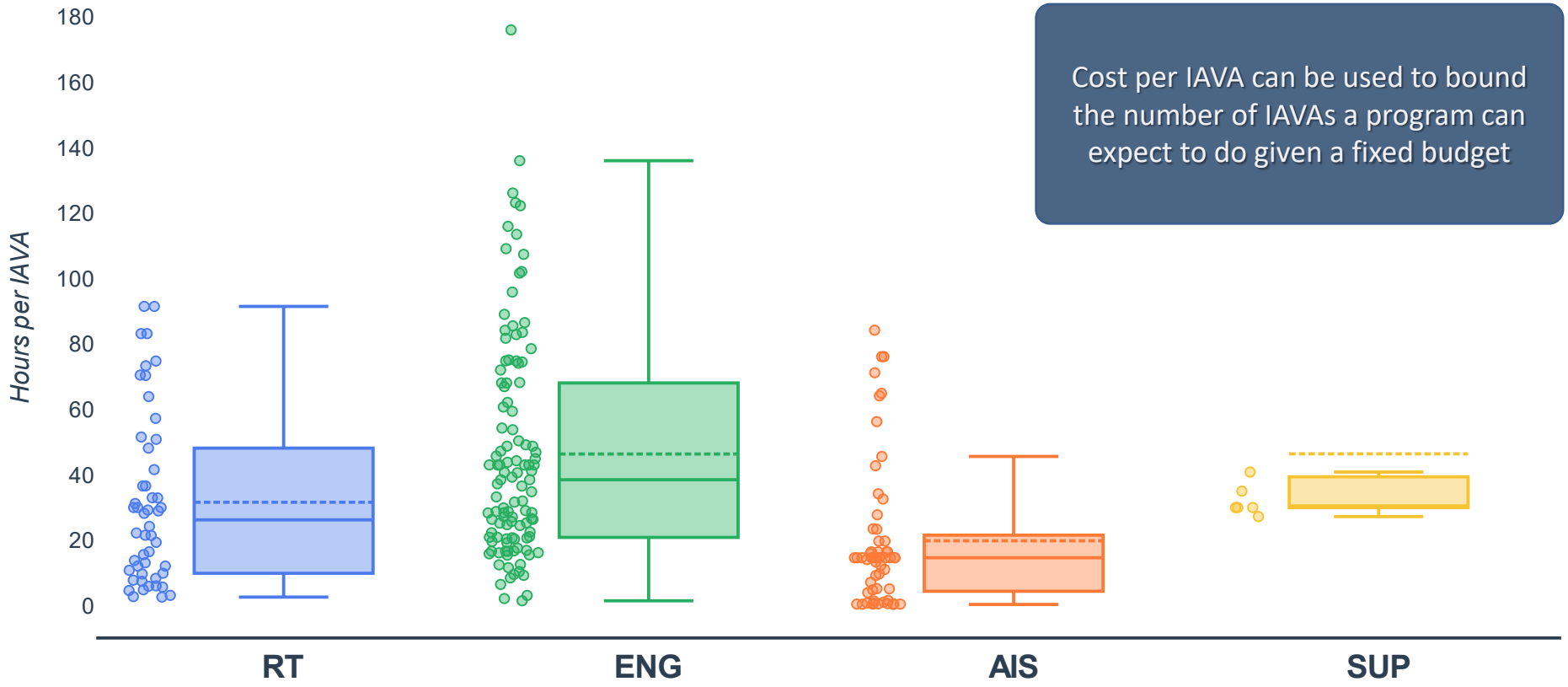
IAVA Releases



Super Domain	Count	Mean	Median
RT	63	55	37
ENG	116	38	24
AIS	43	20	20
SUP	7	21	30

Hours per IAVA

IAVA Releases



Cost per IAVA can be used to bound the number of IAVAs a program can expect to do given a fixed budget

Super Domain	Count	Mean	Median
RT	50	31.6	26.2
ENG	113	46.4	38.5
AIS	56	19.8	14.6
SUP	7	46.4	30.0

Lessons Learned

Hierarchy of Use Cases

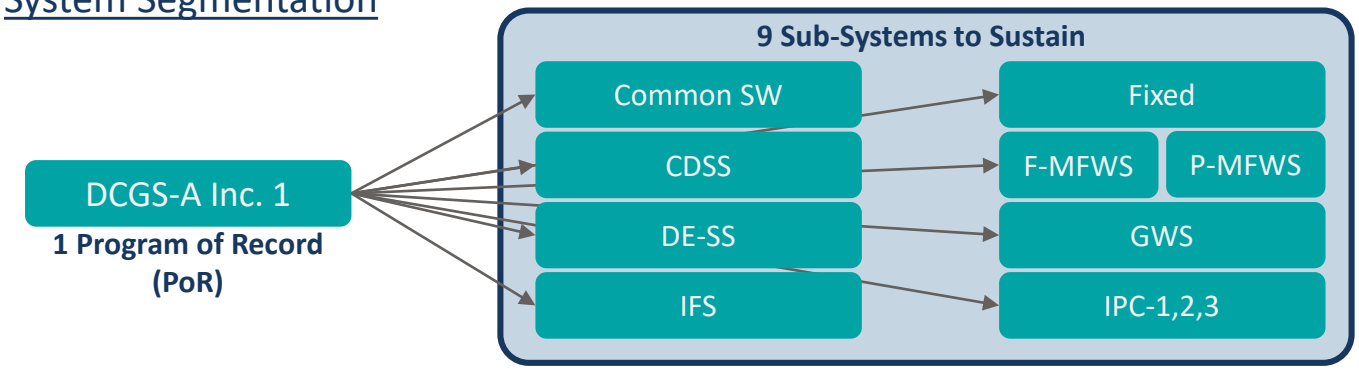
1
- Utilize this analysis structure and findings to support your data collection and analysis
- Build predictive models based on historical actual data collected to predict future efforts

2
- Utilize data SWM PH 2 data repository to filter for analogous systems
- Develop custom regression models based on systems within the analogous set

3
- Use the regression models developed from the SWM PH 2 effort to estimate future effort
- Utilize the benchmarks presented to estimate or compare against your program

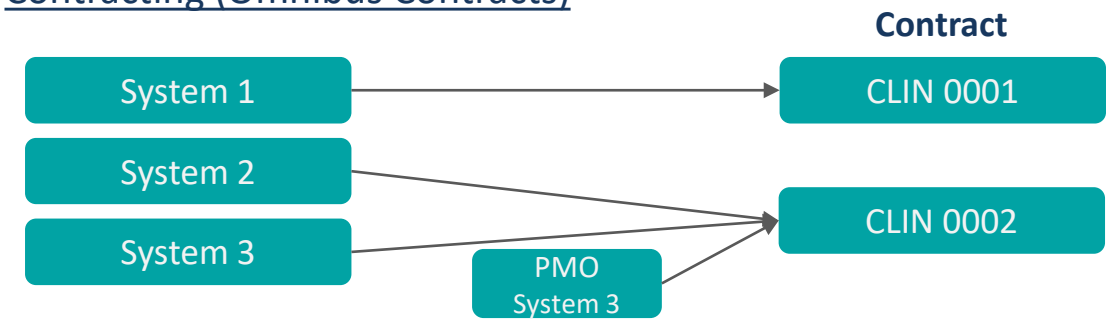
Why Software Sustainment Cost is Difficult to Track

1 System Segmentation



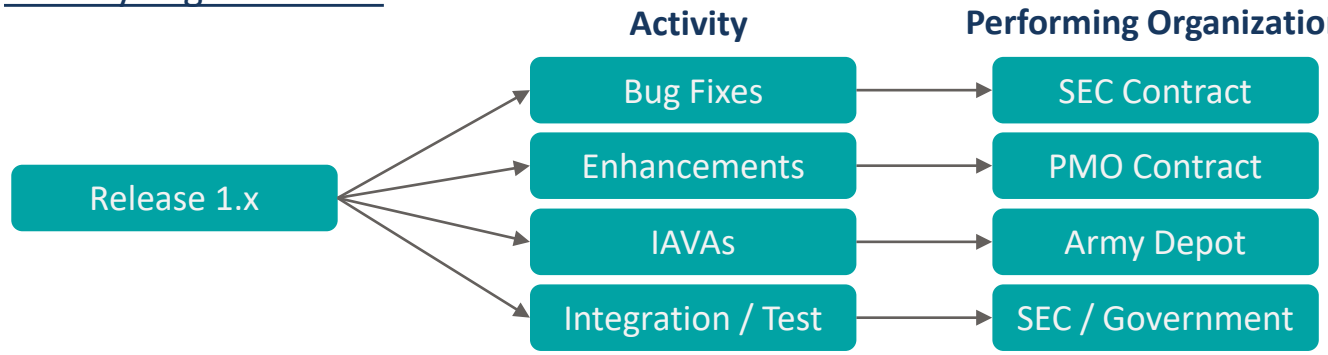
- System segmentation makes it difficult to track software size, licenses, requirements, across the lifecycle
- Each sub-system is separately managed

2 Contracting (Omnibus Contracts)



- Current Omnibus contracts do not provide a means to track cost to an individual system or WBS activity
- Contract CLINs can contain multiple systems and multiple funding streams
- Current contracts do not align to software sustainment activities

3 Activity Segmentation



- Release activities get split across multiple performing organizations
- Annual funding makes it difficult to track release costs since releases may take from one month to several years

Lessons Learned From Analysis

- It often takes multiple iterations with the data provider to clean up the data provided – this may be caused by a misunderstanding of what data is being requested or a lack of complete data
  - It is worth the effort to clean up the submitted data
- Data for some of the WBS elements was reported “unavailable” because the work was funded by different organizations, because costs were applicable to multiple systems, or because data was not tracked at lower WBS levels
- Release data was collected for a full release – yet it is tracked annually
  - Future analysis will evaluate annual release data and aggregate release data that spans multiple fiscal years
- Inner program CERs and SERs show significantly better statistics
  - Project leads at LCMC’s can use same methodology to develop estimates for program funding

Next Steps & Future Research

- Annual data collection
  - The Software Resources Data Reporting for Maintenance (SRDR-M*) closely aligns to the DASA-CE SWS WBS and data requirements
    - Moving forward, the SRDR-M will be utilized to collect SWS data from Army programs and perform analysis
- Additional analysis of data, including:
  - Cost impact of cybersecurity framework (DIACAP vs RMF)
  - Cost of Cybersecurity
  - Analysis of annualized release data
- Expand SER analysis to include all systems in each release duration category (Cyclic, Sequential, Concurrent)
- Additional license analysis
  - Does higher license costs correlate to higher sustainment costs?
  - Does using COTS software save money in sustainment?
- Impact of budget reductions on fixed-cost versus variable-cost funding
- Iterative/Agile versus traditional development is being explored for differences
- New FY18 PPSS data being collected

Concluding Remarks

Importance of Data Collection

- Consistent and accurate technical/cost data allows for more meaningful CERs that are relevant to the changing environment of software sustainment
- Software sustainment data can be used to better inform design decisions and cost analysis
  - DASA-CE and the Army cost community are now able to develop cost products that use analogous program data and technical output to estimate software sustainment
  - This facilitates major milestone estimates, O&S cost targets, Operation Sustainment Reviews, and yearly POM reviews
  - Dataset is hosted on CADE under “Library”

*See http://cade.osd.mil/policy/dids for more information

Contributors

Cheryl Jones

Software Measurement Analyst
US Army AFC
973-724-2644

James Doswell

Senior Operations Research Analyst
ODASA-CE
703-697-1572

Dr. Bradford Clark

Vice President
Software Metrics, Inc.
(703) 754-0115

James Judy

NISEC Division Chief
ODASA-CE
703-697-1612

Dr. Robert Charette

President
ITABHI Corporation
(540) 972-8150

Paul Janusz

Software Measurement Analysis
US Army AFC
973-724-2644

Acronyms -1

ACAT	Acquisition Category
AIS	Automated Information System super domain
BL	Software Change Backlog
BY	Base Year
C&A	Certification and Accreditation
C5ISR	Command, Control, Communications, Computers, Cyber, Intelligence, Surveillance, and Reconnaissance
CADE	Cost Assessment Data Enterprise
CER	Cost Estimating Relationship
COTS	Commercial Off The Shelf
CRED	Uncertainty Estimation Determination
CSCI	Computer Software Configuration Item
Cyber%	Percent of the release that is Cybersecurity updates
DASA-CE	Deputy Assistant to the Secretary of the Army for Cost and Economics
DBS	Defense Business System commodity
DIACAP	DoD Information Assurance Certification and Accreditation Process
DISA	Defense Information Systems Agency
DoD	US Department of Defense
DSLOC	Delivered Source Lines of Code
ECP	Engineering Change Proposal
El_Mod	External Interfaces Modified
ENG	Engineering super domain
Enh%	Percent of the release that is Enhancements to the system
EW	Electronic Warfare

Acronyms -2

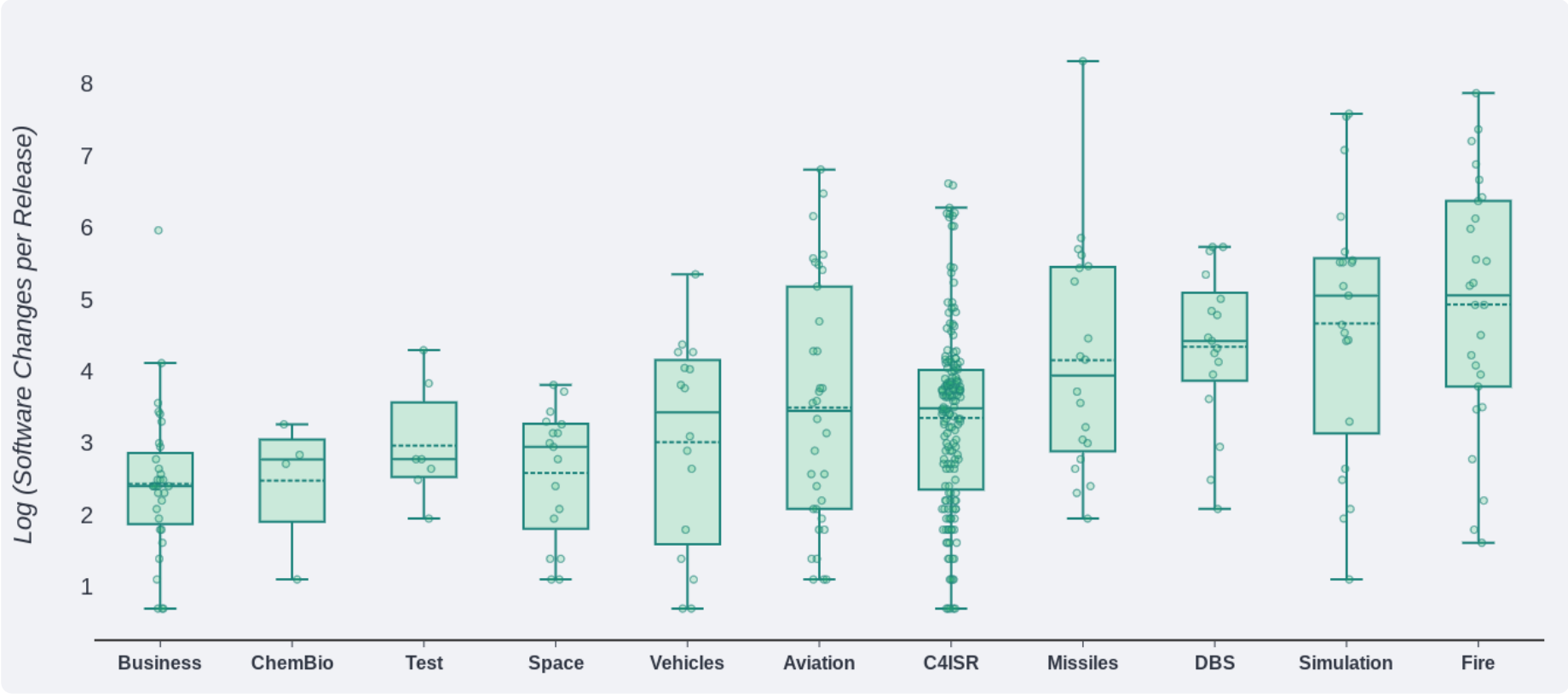
FSE	Field Software Engineering
FTE	Full Time Equivalent
IAVA	Information Assurance Vulnerability Alert
IAVM	Information Assurance Vulnerability Management
ICEAA	International Cost Estimating and Analysis Association
Maint%	Percent of the release that is Maintenance changes
NVD	National Vulnerability Database
O&S	Operations and Sustainment
ODC	Other than Direct Costs
OMA	Operations and Maintenance Army funding
OPA	Other Program Army funding
OSMIS	Operation/Sustainment Management Information System
PDSS	Post-Deployment Software Support
PEO	Program Executive Office
POM	Program Objective Memorandum
PPSS	Post-Production Software Support
PTR	Problem Trouble Report

Acronyms - 3

RDT&E	Research, Development, Testing, and Evaluation
RMF	Risk Management Framework
RT	Real-Time super domain
SC	Software Changes
SEC	Software Engineering Center
SER	Schedule Estimating Relationship
SLOC	Source Lines of Code
SRDR	Software Resources Data Report
SRDR-M	Software Resources Data Report for Maintenance
STIG	Security Technical Implementation Guides
SUP	Mission Support super domain
SW	Software
SWBase	Software Baseline SLOC
SWS	Software Sustainment
TDEV	Time to Develop
THrs	Total release hours
TReqts	Total Requirements in a system
TReqts_Imp	Total Requirements Implemented in a release
TSC	Total Software Changes for a release
WBS	Work Breakdown Structure

Backup

Software Changes per Release
by Commodity



	Business	ChemBio	Test	Space	Vehicles	Aviation	C5ISR	Missiles	DBS	Simulation	Fire
Count	32	4	7	17	16	34	172	20	17	21	26
Mean	2.43	2.47	2.96	2.58	3.01	3.49	3.35	4.15	4.34	4.66	4.93
St Dev	1.05	0.95	0.81	0.91	1.47	1.71	1.28	1.60	1.08	1.83	1.73
Median	2.40	2.77	2.77	2.94	3.43	3.44	3.48	3.94	4.42	5.05	5.05

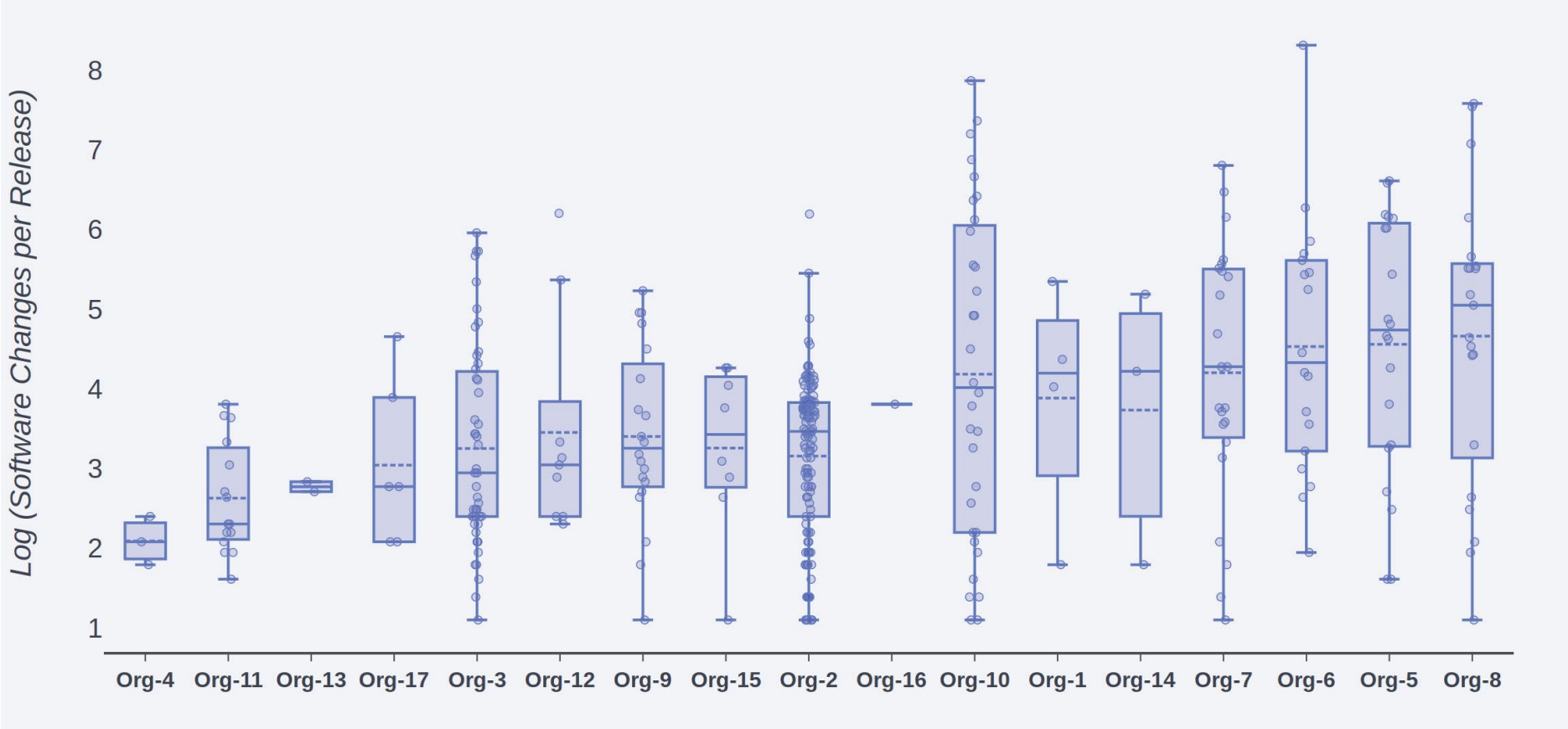
Software Change Definition Variability

- Within WBS 1.0, the effort associated with software releases is captured
- A software release can be sized using the count of the number of software changes
- A software change describes a change where source code/script is altered whether it be added, deleted or modified
- Respondents defined a software change as:
  - Enhancements
    - New capability: ECPs, new requirements
    - Redesign / rewrite: 100% new code, new architecture
  - Maintenance
    - Defect repair: bug fixes, PTR fixes
    - Reconfiguration: threat loads, EW parameters
    - Rehost: migration from Windows to Linux
    - Testing: interoperability testing
    - Update: weapon tables, switch configurations, Operating System
    - Update, Defect repair (see above)
    - Upgrade: upgrade the v “n” to v “n+1”, upgrading applications
  - Cyber
    - Vulnerabilities: enhance security posture not resolved

Since there was significant variability across the programs in the definition of a software change, a more in-depth analysis was conducted to understand the costs of different types of software changes

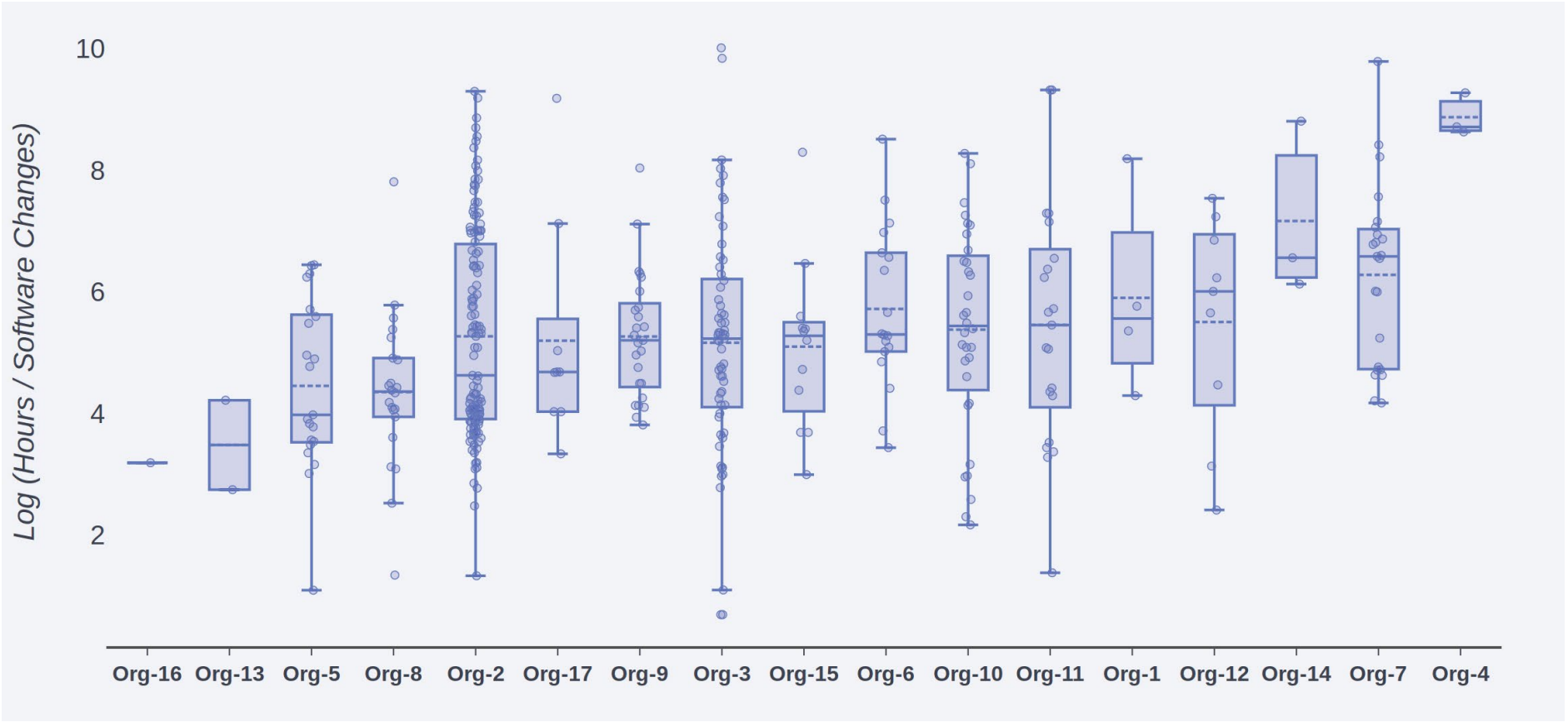
Benchmarks for Capability Releases

Software Changes per Release
By Sustaining Organization



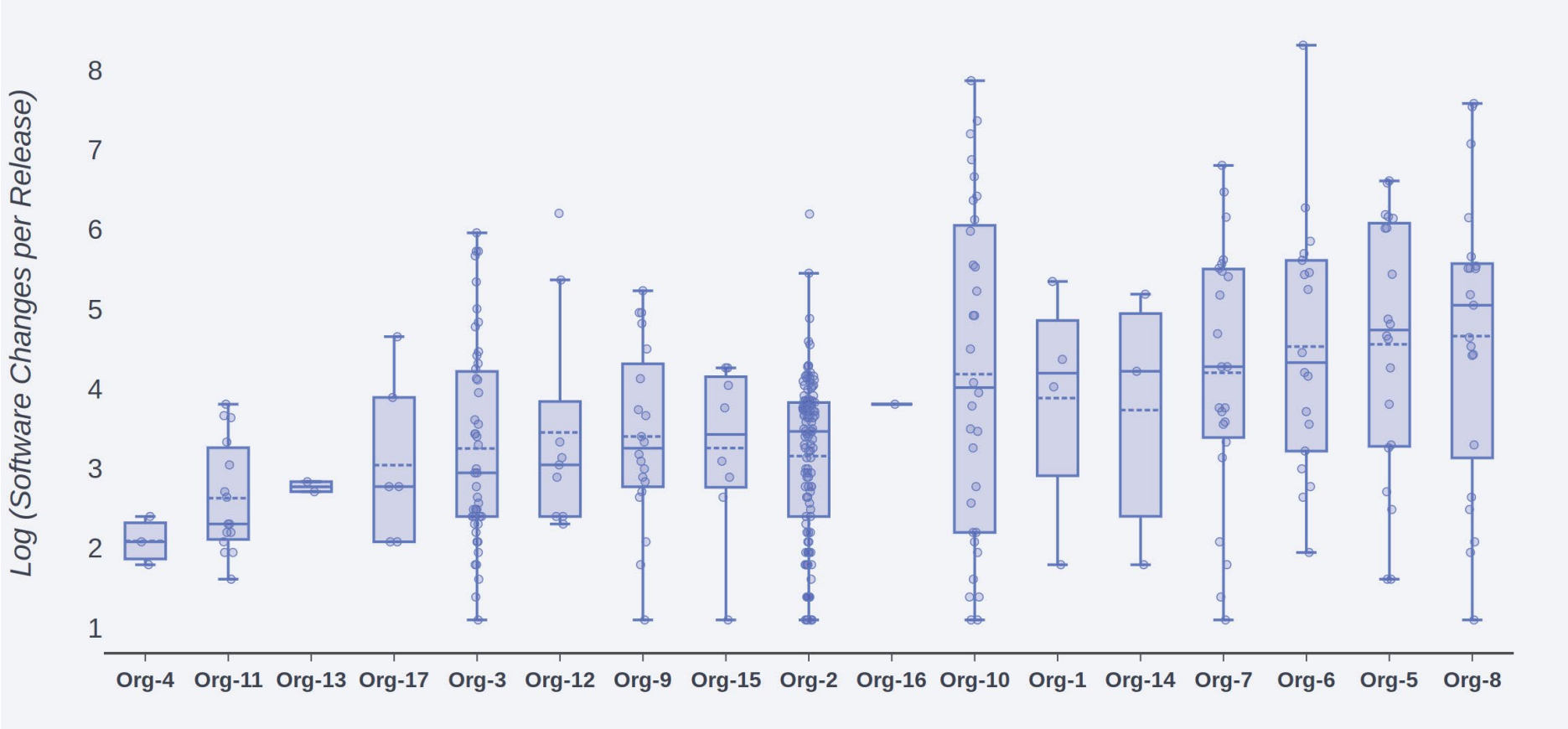
	Org-4	Org-11	Org-13	Org-17	Org-3	Org-12	Org-9	Org-15	Org-2	Org-16	Org-10	Org-1	Org-14	Org-7	Org-6	Org-5	Org-8
Count	1	2	21	22	131	9	25	61	12	18	32	21	4	9	3	23	3
Mean	3.19	3.483	4.454	4.352	5.268	5.197	5.266	5.163	5.099	5.719	5.378	5.456	5.901	5.503	7.164	6.279	8.87
St Dev	N/A	1.039	1.399	1.281	1.712	1.823	1.037	1.88	1.402	1.327	1.675	1.995	1.646	1.802	1.438	1.474	0.349
Median	3.19	3.483	3.978	4.361	4.626	4.682	5.202	5.231	5.278	5.3	5.439	5.454	5.56	6.007	6.559	6.582	8.712

Hours per Software Change
By Sustaining Organization



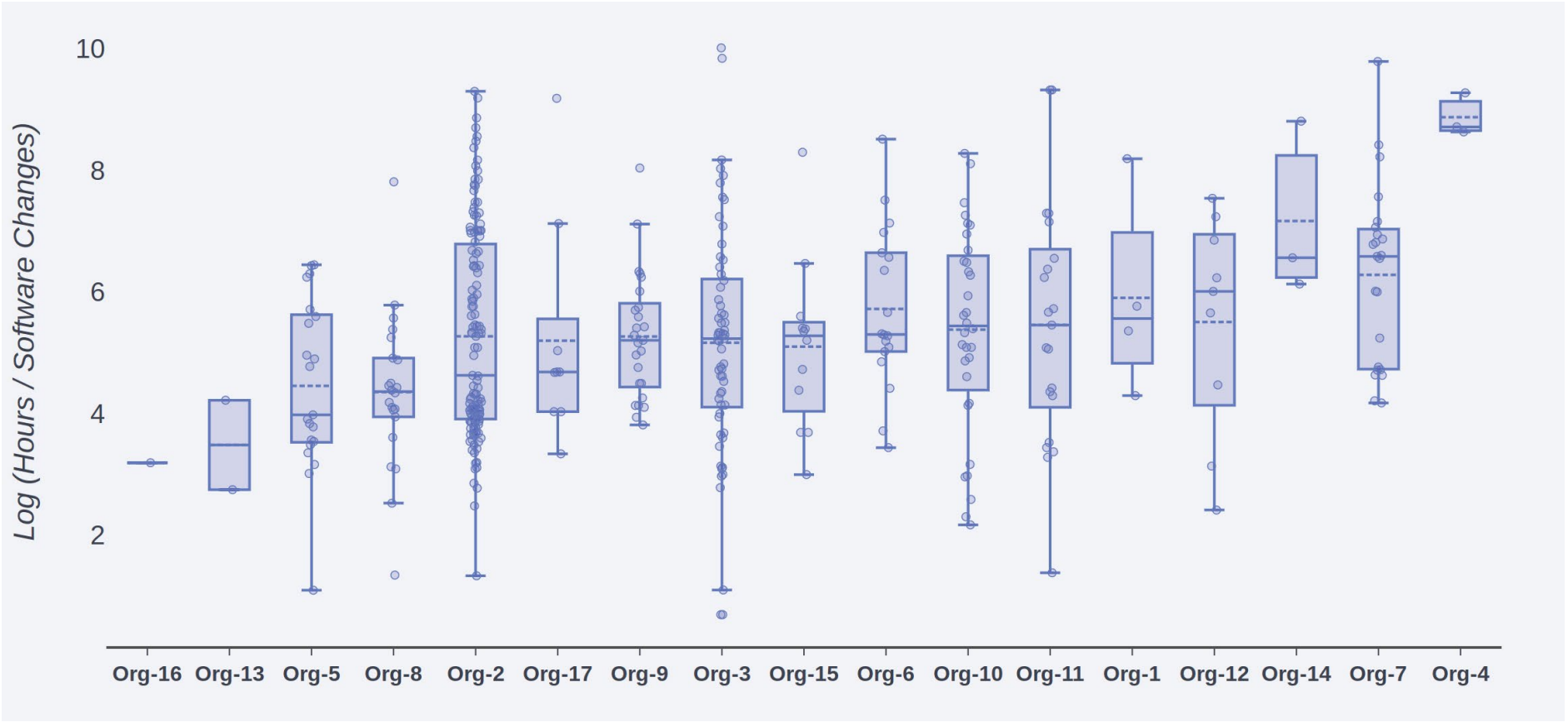
	Org-16	Org-13	Org-5	Org-8	Org-2	Org-17	Org-9	Org-3	Org-15	Org-6	Org-10	Org-11	Org-1	Org-12	Org-14	Org-7	Org-4
Count	1	2	21	22	131	9	25	61	12	18	32	21	4	9	3	23	3
Mean	3.19	3.483	4.454	4.352	5.268	5.197	5.266	5.163	5.099	5.719	5.378	5.456	5.901	5.503	7.164	6.279	8.87
St Dev	N/A	1.039	1.399	1.281	1.712	1.823	1.037	1.88	1.402	1.327	1.675	1.995	1.646	1.802	1.438	1.474	0.349
Median	3.19	3.483	3.978	4.361	4.626	4.682	5.202	5.231	5.278	5.3	5.439	5.454	5.56	6.007	6.559	6.582	8.712

Software Changes per Release
By Sustaining Organization



	Org-4	Org-11	Org-13	Org-17	Org-3	Org-12	Org-9	Org-15	Org-2	Org-16	Org-10	Org-1	Org-14	Org-7	Org-6	Org-5	Org-8
Count	1	2	21	22	131	9	25	61	12	18	32	21	4	9	3	23	3
Mean	3.19	3.483	4.454	4.352	5.268	5.197	5.266	5.163	5.099	5.719	5.378	5.456	5.901	5.503	7.164	6.279	8.87
St Dev	N/A	1.039	1.399	1.281	1.712	1.823	1.037	1.88	1.402	1.327	1.675	1.995	1.646	1.802	1.438	1.474	0.349
Median	3.19	3.483	3.978	4.361	4.626	4.682	5.202	5.231	5.278	5.3	5.439	5.454	5.56	6.007	6.559	6.582	8.712

Hours per Software Change
By Sustaining Organization



	Org-16	Org-13	Org-5	Org-8	Org-2	Org-17	Org-9	Org-3	Org-15	Org-6	Org-10	Org-11	Org-1	Org-12	Org-14	Org-7	Org-4
Count	1	2	21	22	131	9	25	61	12	18	32	21	4	9	3	23	3
Mean	3.19	3.483	4.454	4.352	5.268	5.197	5.266	5.163	5.099	5.719	5.378	5.456	5.901	5.503	7.164	6.279	8.87
St Dev	N/A	1.039	1.399	1.281	1.712	1.823	1.037	1.88	1.402	1.327	1.675	1.995	1.646	1.802	1.438	1.474	0.349
Median	3.19	3.483	3.978	4.361	4.626	4.682	5.202	5.231	5.278	5.3	5.439	5.454	5.56	6.007	6.559	6.582	8.712