

Dashing All the Way: Defining the Best Dashboard for Your Program

Tom Merendino

Michael Bandor

Robert Ferguson

Software Solutions Conference 2015

November 16–18, 2015



Software Engineering Institute

Carnegie Mellon University

© 2015 Carnegie Mellon University

Distribution Statement A: Approved for Public Release;
Distribution is Unlimited



Copyright 2015 Carnegie Mellon University

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN “AS-IS” BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material has been approved for public release and unlimited distribution except as restricted below.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

DM-0002741

Agenda



Background

Why a Dashboard?

Dashboard Structure & Categories

Trending & Red/Yellow/Green Thresholds

Dashboard Implementation

**Mapping of Metrics to Dashboard
Example**

Training

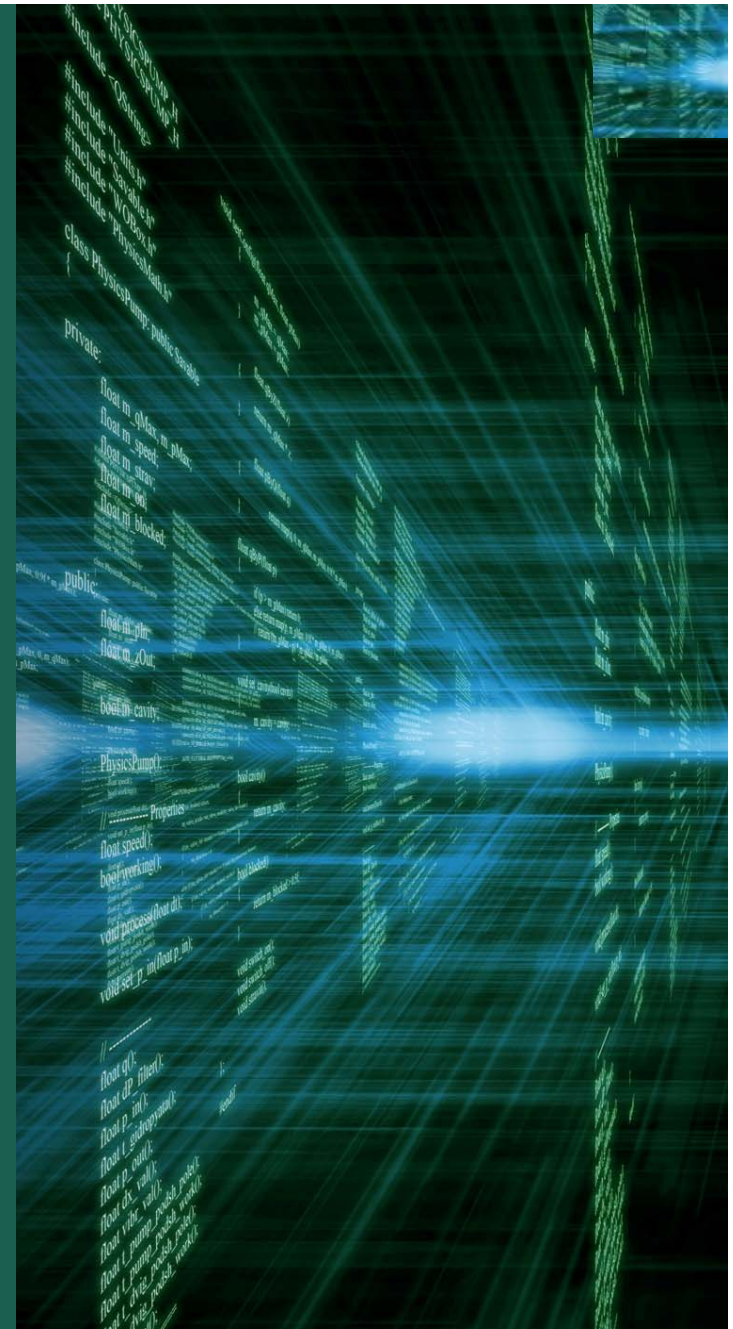
**Metrics Interpretation (Dashboard
Example)**

**Metrics Application (PM & Senior
Leaders)**



Dashing All the Way: Defining the Best Dashboard for Your Program

Background



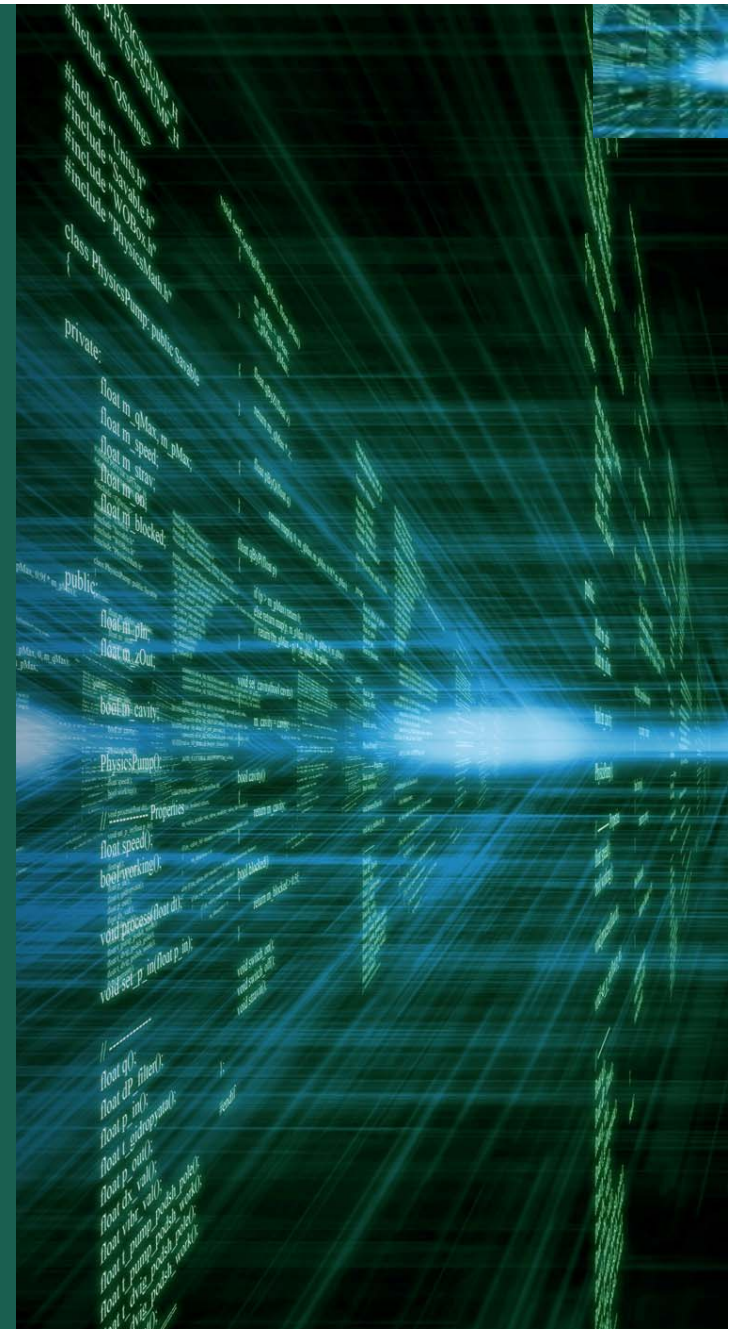


Background

- A DoD program wasn't really tracking many metrics (gov't or contractor) beyond EVMS & cost data
- Program had a Nunn-McCurdy breach
 - DoD directed them to start tracking metrics
- SEI brought in to assist the program (gov't & contractor) with the metrics effort & reporting
- Initial set of metrics identified by the contractor & refined
 - Typically metrics are done in a top-down manner
 - Contractor had some metrics envisioned as an "enterprise" set of metrics
- Gov't metrics plan (dashboard implementation & interpretation) created
- Training provided to the gov't (initial metrics & dashboard training)
- Additional application training (PM & senior leaders) provided
- Additional programs are also looking into implementing this approach

Dashing All the Way: Defining the Best Dashboard for Your Program

Why a Dashboard?





Why a Dashboard?₁

The purpose of a project or program dashboard is to confirm external commitments and to warn leaders when changes to commitments may be required.

Since changing external commitments is undesirable, the dashboard must support forecasting and identify risks to project plans and methods before problems get out of hand. This includes the following critical areas:

- Scheduling
- Resource Allocation
- Scope & Change
- Product Quality
- Effect process performance

Implementation as a single dashboard give the program managers and senior leadership a quick way to assess the overall “health” of their program(s).

It does NOT mean, that additional data beyond the dashboard is unnecessary. The additional data supports the information being presented and provides additional details (drill-down) down to the source metrics where needed



Why a Dashboard?₂

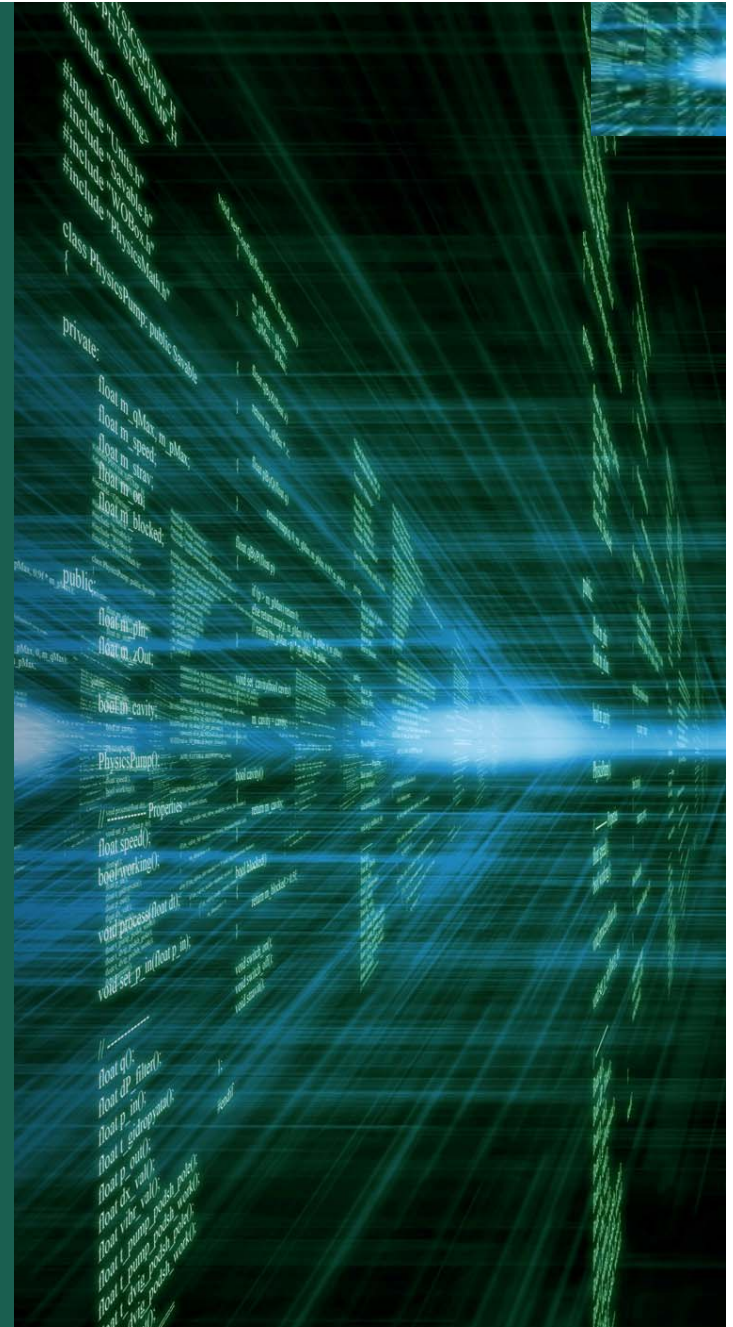
A dashboard should assist with doing the following:

- Forecast milestones and delivery of scope
- Provide clear warnings if a plan is not working or an unplanned event has affected some desired outcome
- Support re-estimation and re-planning by showing the magnitude of the problem



Dashing All the Way: Defining the Best
Dashboard for Your Program

Dashboard Structure & Categories





Dashboard Structure & Categories₁

The purpose of a project or program dashboard is to confirm external commitments and to warn leaders when changes to commitments may be required.

Since changing external commitments is undesirable, the dashboard must support forecasting and identify risks to project plans and methods before problems get out of hand. This includes the following critical areas:

- Scheduling
- Resource Allocation
- Scope & Change
- Product Quality
- Effect process performance

Implementation as a single dashboard give the program managers and senior leadership a quick way to assess the overall “health” of their program(s).

It does NOT mean, that additional data beyond the dashboard is unnecessary. The additional data supports the information being presented and provides additional details (drill-down) down to the source metrics where needed





Dashboard Structure & Categories₂

The first three quadrants (Schedule, Cost & Resources; Scope, Progress & Change; Process & Risk) represents decisions about directing and controlling the work.

- The PM decides on schedule and resource allocation
- The PM must involve both customer and development organizations to make changes in scope
- Requests to change a process may be suggested by the PM, but only the developers and subcontractors can make those changes

The fourth quadrant (Quality) represents the target objectives:

- Do quality checks during the performance of the work reveal potential problem in achieving quality targets for value & performance?
- Do reviews and test demonstrate progress in achieving quality targets for value and performance?





Dashboard Structure & Categories₃

Schedule, Cost & Resources Quadrant

This section focuses on the project or program to make sure it delivers on its forecast date at the forecast cost.

Typical metrics found in this section are cost & schedule (Earned Value where applicable), staffing and other resources.

General questions to think about:

- Has the project/program been predictable in recent milestone performance?
- Are forecasts regularly updated (every program review where applicable)?
- Does the current forecast predict a problem in execution?
- Can the plan be made executable within the project contingency?





Dashboard Structure & Categories₄

Scope, Progress & Change Quadrant

This section focuses on the ability to communicate the scope and progress of a project/program and prepare estimate changes. Changes reflected here typically require more stakeholders to be involved than “predictability” decisions.

Typical metrics found in this section are actual progress measured in size compared to the estimate of size at completion. Other items can include changes in requirements and changes in size estimates.

General questions to think about:

- Is there a change to the estimated size (e.g. more requirements)?
- Does the adjusted size suggest an estimation error (e.g., planning problems)?
- Is the change request too large or too late to accomplish within the budget?
- Will teams (developers, testers, etc..) complete their separate efforts in time for integration?
- Will the PM need to defer portions of the program/project to a later release or eliminate a deliverable?





Dashboard Structure & Categories₅

Process & Risk Quadrant

This section is focused on keeping the costs and quality under control, and to monitor risks to the program/project.

Typical metrics found in this section can include audit results, quality checks (QA), behavioral observations to determine if new technology is properly deployed as well as risk monitoring.

General questions to think about:

- Are the processes being followed (contractor and government)? If not, why not?
- Is performance data available and reviewed when the data suggests?
- What program/project risks might exist because of process program (e.g. QA questions)
- Are technology changes reflected in process changes (e.g. training for new tools, etc..)?





Dashboard Structure & Categories₆

Quality Quadrant

This final section focuses on the ability to deliver a high quality product, as well as to validate claims of progress.

Typical metrics tracked in this section include Key Performance Parameter (KPP) and Technical Performance Measure (TPM) test results, verification and validation results, etc..

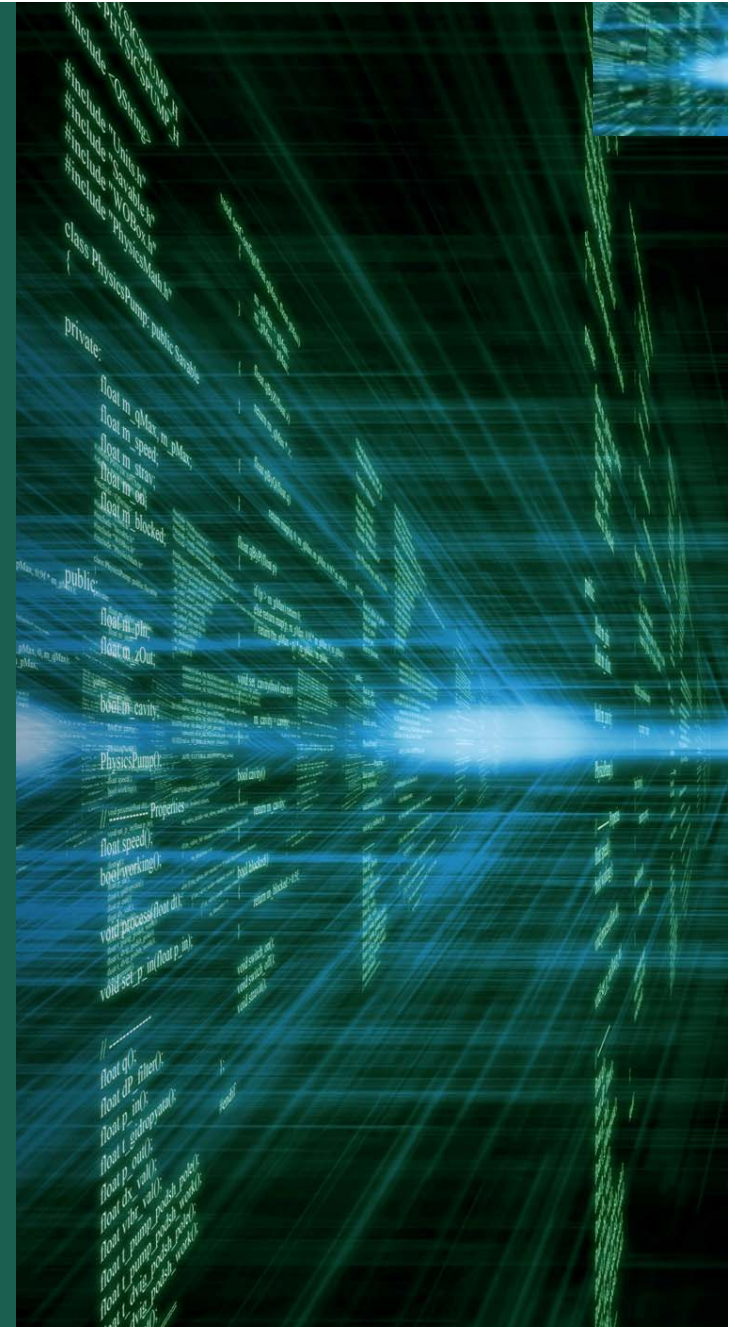
General questions to think about:

- Will the product achieve the desired quality and performance?
- Do validation checks (including Development Test and Operational Test) prove the value in use potential?



Dashing All the Way: Defining the Best Dashboard for Your Program

Trending & Red/Yellow/Green Thresholds





Trending & Red/Yellow/Green Thresholds

The contractor may have already defined the trending criteria (e.g. up, down, no change, etc..) as part of their metrics plan. The same with the Red/Yellow/Green determination.

In cases where that determination has not been made by the contractor, the government will need to set those reporting thresholds. For example:

- Green: at or exceeding the goal
- Yellow: $\leq 5\%$ below the goal or an equivalent change since the last reporting period
- Red: $> 5\%$ below the goal or a an equivalent change since the last reporting period

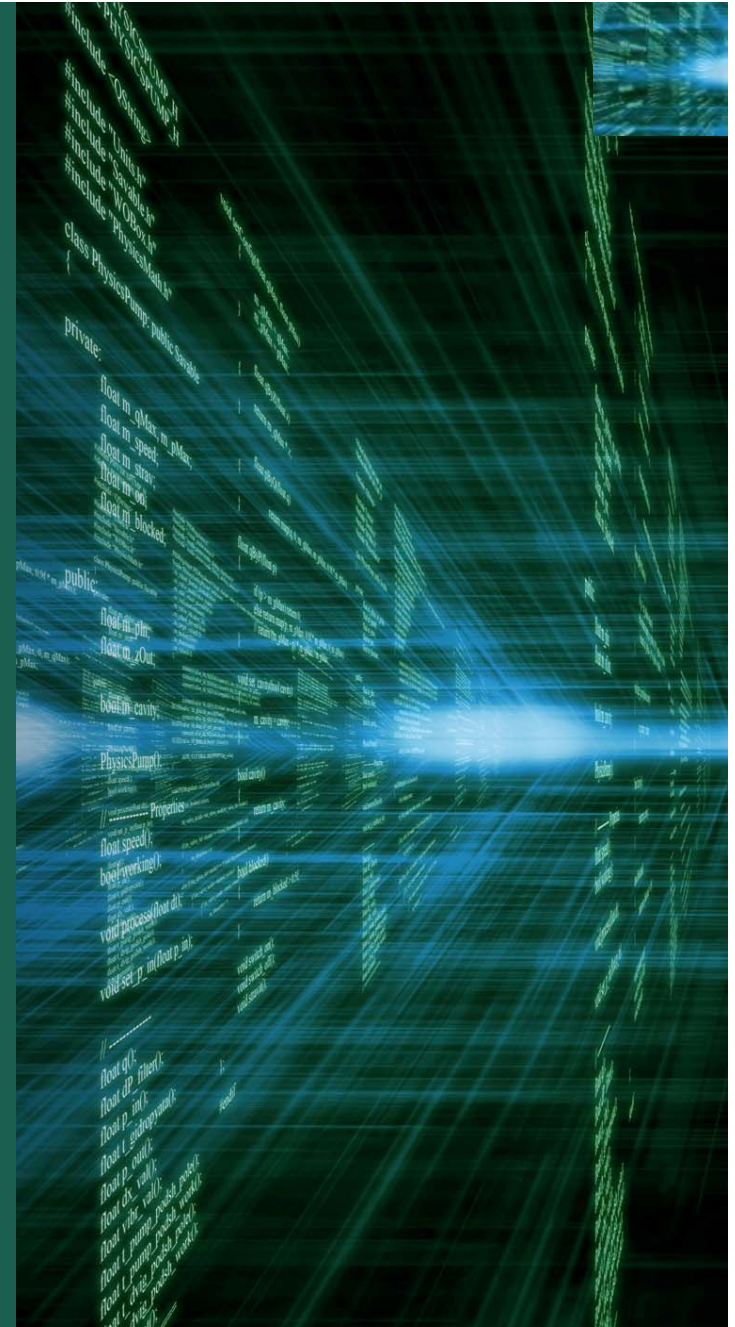
Another possible representation (depending on the data) could be:

- Green: Operating within plan
- Yellow: Using a contingency budget for schedule, cost or product performance
- Red: A change to a plan or sub plan is required to reestablish satisfactory schedule, cost or product performance budgets

For the initial implementation of the dashboard, the first set of criteria will be used as the baseline however as the metrics reporting progresses (e.g. gets more mature), those thresholds should be reevaluated and adjusted as necessary.

Dashing All the Way: Defining the Best Dashboard for Your Program

Dashboard Implementation





Dashboard Implementation

Uses a single, quad chart to present a concise, high-level depiction of the program health and any issues

Major categories:

- Schedule, Cost & Resources
- Scope, Progress & Change
- Process & Risk
- Quality

Data supporting information in section is from the contractor metrics and gov't metrics (as defined)

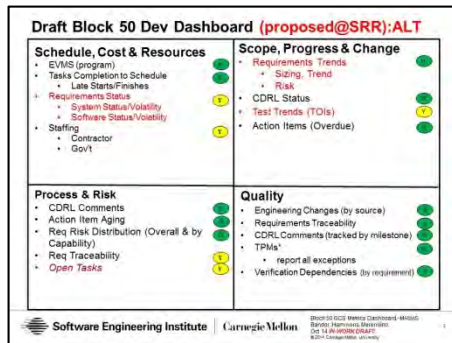
- Used as “intermediate chart and “drill-down” for further detailed information supporting the status shown
- Metrics are not necessarily mapped 1:1 to the dashboard; some are aggregated others are divided across categories

Dashboard contents will evolve as program progresses through major milestones





How It Fits Together

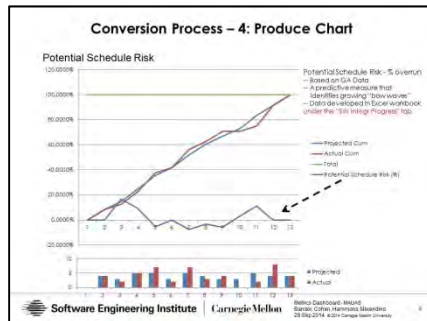


Decomposed into an intermediate chart

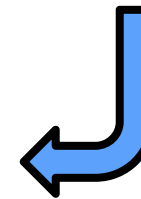
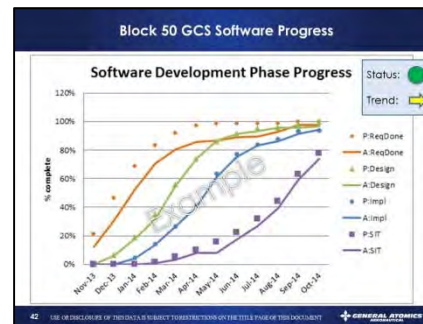
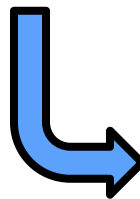


GA BLK 50 Metrics that map to this Label:	Status	Trend	Rationale
SW Development Phase Progress: SW Requirements Progress	🟡	➡	Actual progress in approving SW reqmts is 10% behind projected, which is at the extreme upper end of the yellow threshold. SMEs needed to write reqmts have been overtasked over last several reporting periods. Trend has been flat. May delay future related SW implementation tasks
SW Development Phase Progress: SW Design Progress	🟢	➡	Actual progress is tracking well to projected progress
SW Development Phase Progress: SW Implementation Progress	🟢	➡	Actual progress is tracking well to projected progress
SW Development Phase Progress: SW/HW Integ Test Progress	🔴	⬇	Actual progress is significantly behind projected progress. However, this early in the lifecycle, it is not unexpected that Integration Test progress can have wild swings, period to period. This metric will not presently inform the dashboard value

Supported by analysis

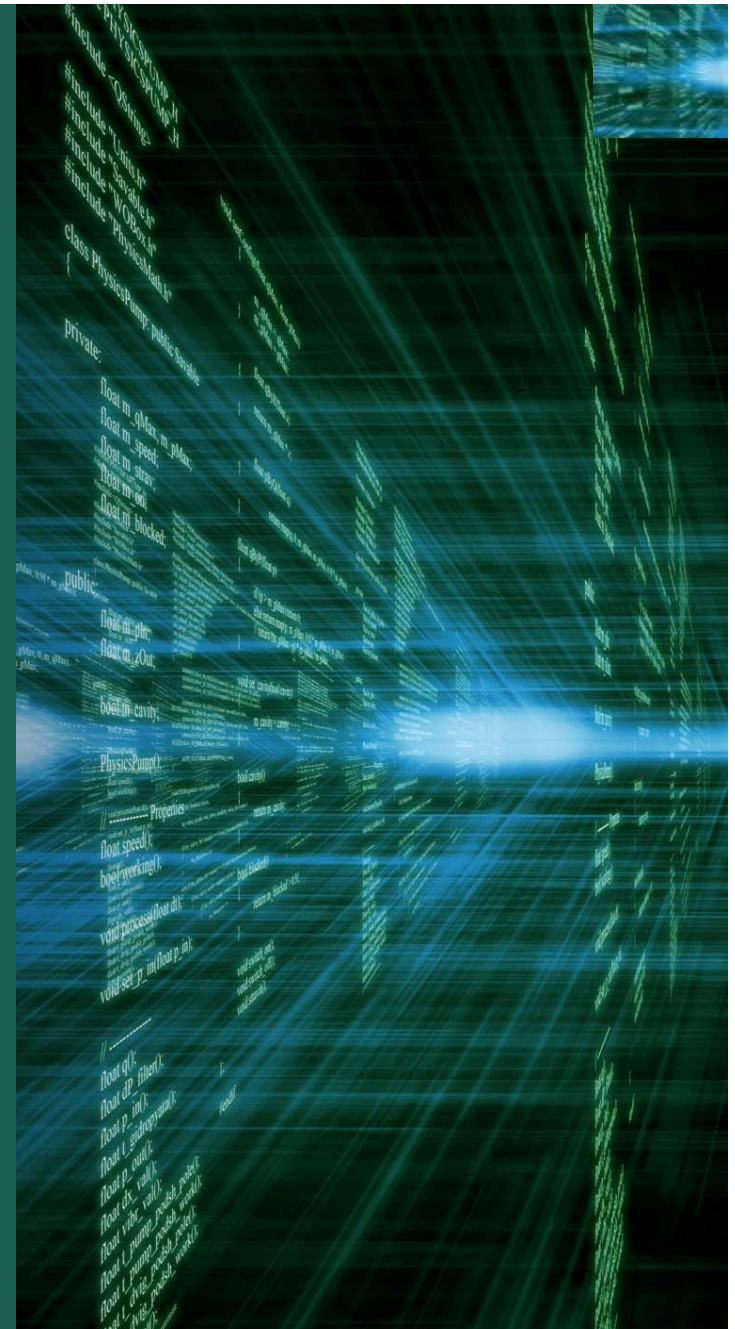


Further supported by the contractor's metrics information



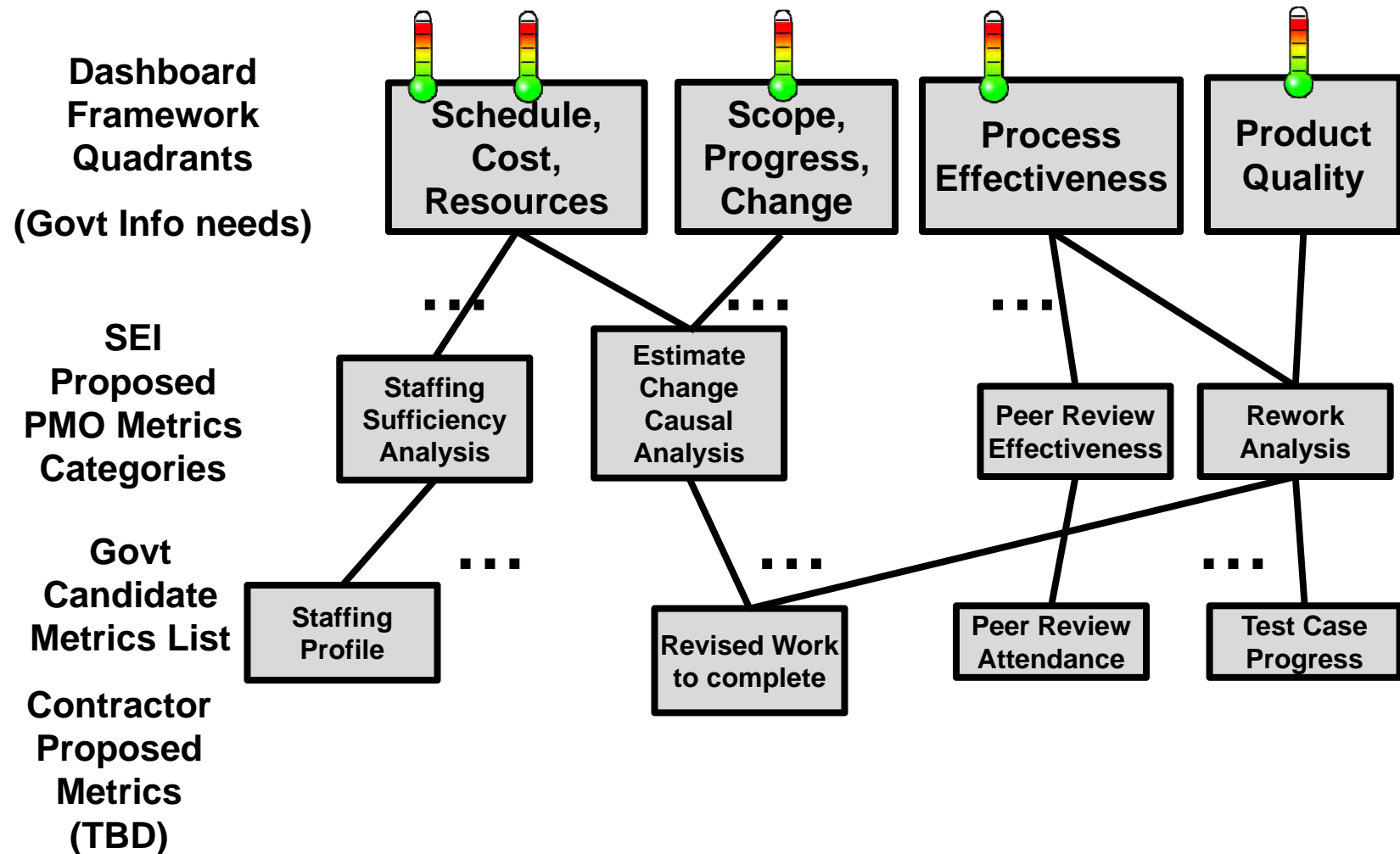
Dashing All the Way: Defining the Best Dashboard for Your Program

Mapping of the Metrics to the Dashboard



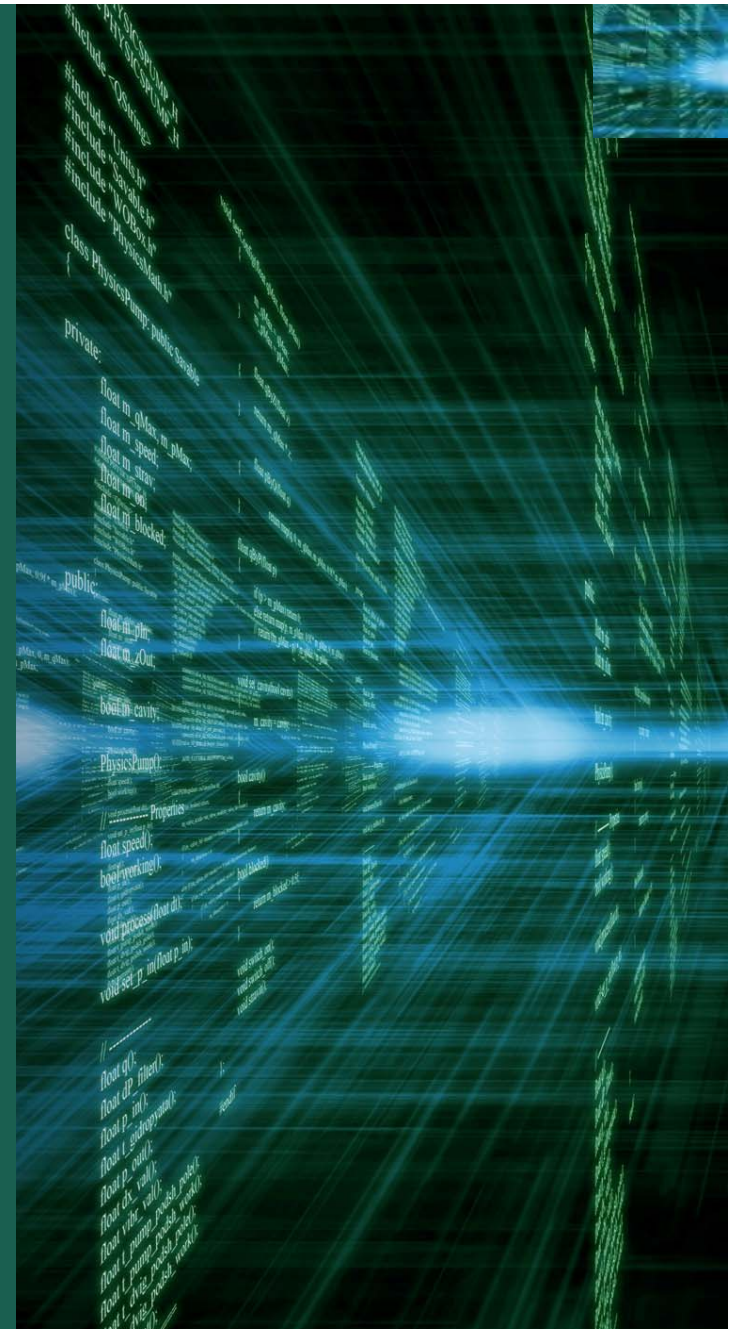


Abstract Depiction of Metrics Mapping Scheme



Dashing All the Way: Defining the Best
Dashboard for Your Program

Training





Training

While many stakeholders exist that have a variety of interests in program metrics, we have identified two primary groups of stakeholders from a training perspective

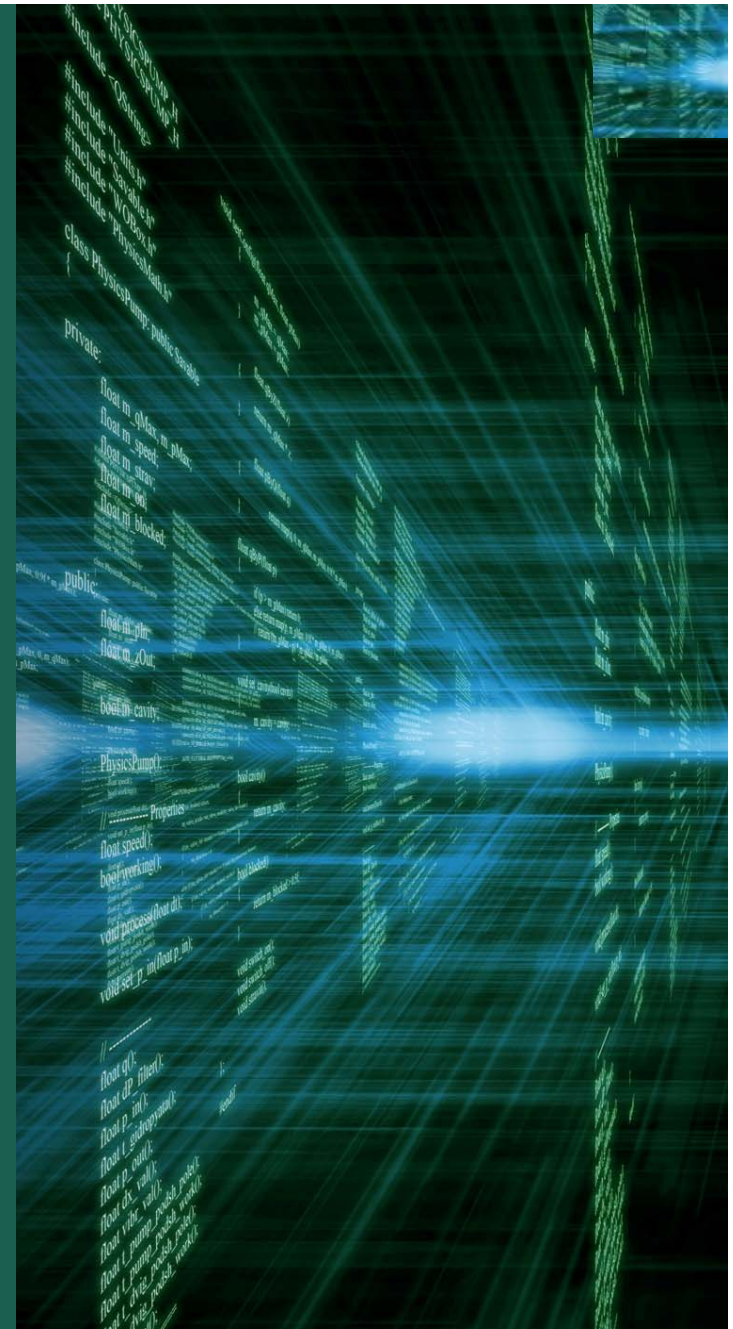
- Metrics implementers / analyzers
- Program leaders / decision makers

We have found that different training approaches are needed to address each of the two primary stakeholder categories:

- Metrics implementers / analyzers need training on the contractor specific metrics, interpretation, and assembly of the underlying “raw” metrics data into dashboard categories
- Program leaders / decision makers need training on interpreting the dashboard level data in the context of program status and progress. They are interested in metrics data to help answer questions such as
 - Are we going to make it to the finish line in time/budget with sufficient quality?
 - Do we understand where the finish line is?
 - What / where are the levers that we can pull? What decisions do we need to make?

Dashing All the Way: Defining the Best Dashboard for Your Program

Metrics Interpretation (Dashboard Example)





Dashboard (proposed)

Schedule, Cost & Resources

- G** EVMS (program)
 - Tasks Completion to Schedule
 - R** [Late Starts/Completions](#)
 - G** Dependencies
 - G** Deferred / Accelerated tasks
 - Requirements Status
 - N** System Status/Volatility
 - N** Software Status/Volatility
 - G** Development Status (HW, SW design, implementation, unit//Integ test)
 - N** Test Status (SIL, GND, FLT)
 - Staffing
 - Y** [Contractor](#)
 - N** Gov't

Scope, Progress & Change

- Requirements Trends
 - R** [Sizing](#), Trend
 - G** Risk
- Development Trends
 - R** Hardware
 - Y** [Software](#)
 - N** SLOC
 - TO Artifacts
- Test Trends
 - G** TOIs
 - R** Trends – SIL, Ground, Flight
- R** CDRL Status
- Y** Problem/Defect Reports
- R** Action Items (Overdue)

Process & Risk

- G** CDRL Comments
- R** [Action Item Aging](#)
- G** Req Traceability
- N** *Req Volatility (future)*
- N** *SLOC (future)*
- N** *TO Comments (future)*
- N** *HW TOIs (future)*
- N** *QA Reports(future)*
 - *Open QA findings*
 - *Defects by phase/severity*
 - *Investigation Aging*

Quality

- G** Engineering Changes (# by source)
- G** Requirements
- G** [Peer Review / CDRL Comments \(by milestone\)](#)
- N** *Problem/Defect Reports: (future)*
 - *SW Defects by Phase & Severity*
 - *SW Defects – by CSCI & Severity*
- N** TPMs*
 - report all exceptions
- G** Verification Dependencies (by req/test point)
- G** Verification Distribution (by CSCI/CSU, HWC1)





Mapping of Contractor Metrics to Dashboard – Schedule, Cost & Resources

Contractor Metrics	Dashboard Label
CPI & SPI	EVMS
Late Starts & Late Completes	Task Completion to Schedule
Staffing Profile (Overall) & Staffing Profile by Discipline (Snapshot)	Staffing (Contractor)
Gov't Staffing (Overall) – Planned vs. Actual	Staffing (Gov't)
Requirements Volatility	Requirements Status (System Status /Volatility)
SW Requirements Volatility	Requirements Status (Software Status / Volatility)
Requirements Progress	Requirements Status (System Status / Volatility); Development Status (HW, SW design, implementation, unit/Integration test)
SW Development Phase Progress	Requirements Status (Software Status / Volatility); Development Status (HW, SW design, implementation, unit/Integration test)
Drawing Release Status	Development Status (HW Design)
SIL, Ground, & Flight Test Progress	Test Status (SIL/Gnd/Flt)





Mapping of Contractor Metrics to Dashboard – Scope, Progress & Change

Contractor Metrics	Dashboard Label
Action Items	Action Items (Overdue)
CDRL Submissions	CDRL Status
Requirements Sizing	Requirements Trends (Sizing)
Requirements Progress	Requirements Trends (Trend)
Requirements Risk Distribution by Capability	Requirements Trends (Risk)
Drawing Release Status	Development Trends (Hardware)
Engineering Changes	Development Trends (Hardware)
SW Development Phase Progress	Development Trends (Software)
SW Defects by Phase & Severity	Problem / Defect Reports
SW Defects by Severity & CSCI	Problem / Defect Reports
SW, HW, & I&I TOIs	Test Trends (TOIs)
SLOC-P	Development Trends (SLOC)
SIL, Ground & Flight Test Progress	Test Trends (Trends – SIL, Ground, Flight)





Mapping of Contractor Metrics to Dashboard – Process & Risk

Contractor Metrics	Dashboard Label
CDRL Comments	CDRL Comments
Requirements Risk Distribution (Overall)	Requirement Risk Distribution (Overall & by Capability)
Requirement Risk Distribution by Capability	Requirement Risk Distribution (Overall & by Capability)
Requirements Traceability	Requirements Traceability
HW TOIs	HW TOIs
SW Requirements Volatility (Overall)	SW Requirements Volatility (Overall & by Type and Source)
SW Requirements Volatility by Type & Source	SW Requirements Volatility (Overall & by Type and Source)
SLOC-P	SLOC
Action Items	Action Item Aging





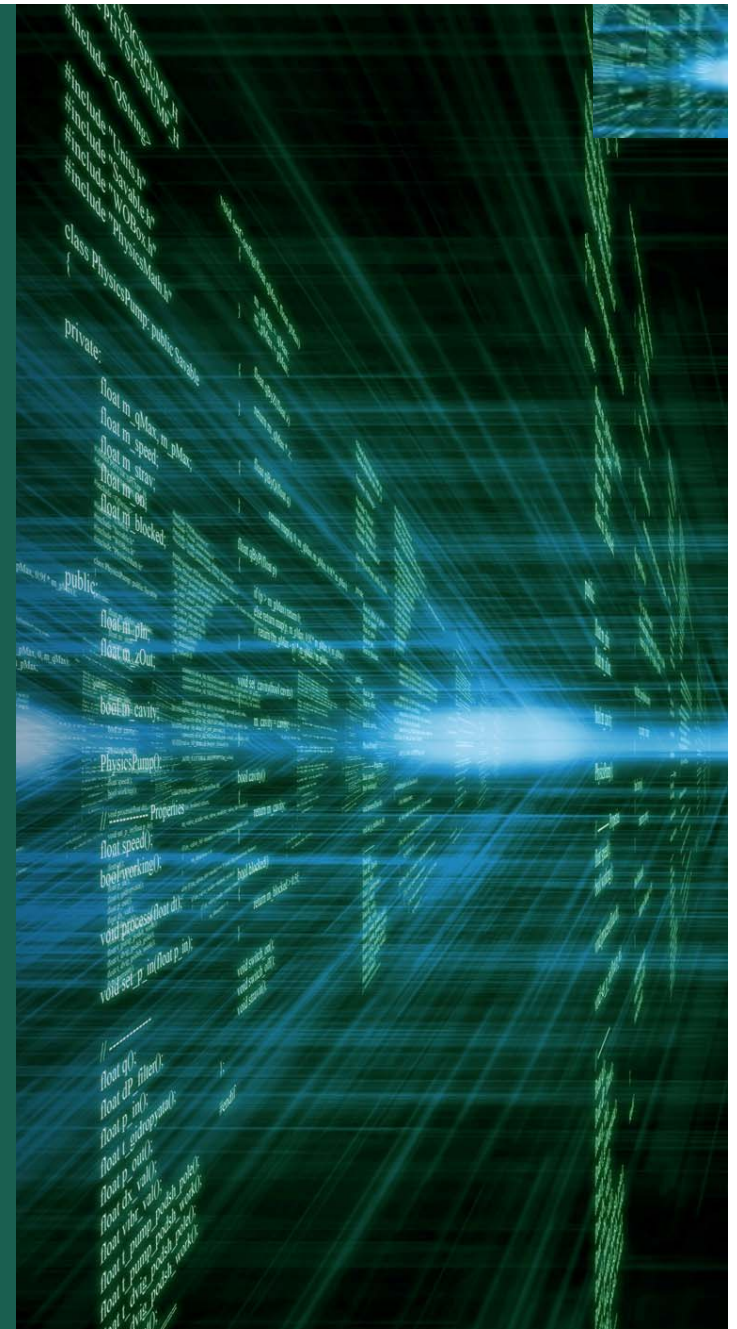
Mapping of Contractor Metrics to Dashboard – Quality

Contractor Metrics	Dashboard Label
Verification Distribution	Verification Distribution (by CSCI/CSU, HWCI)
Verification Dependencies	Ver Dependencies (by req or test point)
Verification Distribution	Ver Distribution (by CSCI, HWCI)
Requirements Traceability	Requirements
Engineering Changes	Engineering Changes (# by source)
HW, SW, and I&T TOIs	Verification Distribution (by CSCI/CSU, HWCI)
SW Requirements Volatility (Overall)	Requirements
SW Defects by Phase & Severity	SW Defects (by Phase & Severity)
SW Defects by Severity & CSCI	SW Defects (by Severity & CSCI)
Tech Orders – New Development	Tech Orders
CDRL Comments	CDRL Comments
System TPMs	TPMS



Dashing All the Way: Defining the Best
Dashboard for Your Program

Metrics Application (PM & Senior Leaders Training)





Metrics Application – PM & Senior Leaders₁

Metrics information developed initially was very detailed (implementation, interpretation, graph construction, etc.)

- Initial training session identified the need for additional training at the PM & Senior Leader level
- Some “nuggets” of info was scatter throughout the gov’t metrics plan, but nothing specific to the PM & Senior Leader level (application)

Additional information (new section) was added to the gov’t metrics plan specific to the PM & Senior leaders

- Allowed them to focus on what was important, why, and any additional related information



Metrics Application – PM & Senior Leaders₂



New training was developed for the PM & Senior Leaders covering (for each metric) the following:

- Purpose of the Metric
- Applicability (where in the enterprise it applied)
- Composite Metrics on the Dashboard
- Interpretation
- Potential Questions
- Related Metrics to Cross-Reference





Metrics Application: Late Starts/Completes₁

Purpose of this Metric: This set of metrics tracks number of contractor tasks that were started late and completed late based on their Integrated Master Schedule (IMS)

Applicability: Hybrid Release & Tech Maturation Event

Composite Metric on Dashboard: No

Interpretation:

- Late Starts and Late Completions are reliable indicators of potential schedule risk and should be carefully tracked and analyzed for causes
- Late Starts may be caused by resources issues (staffing shortfalls, contention for shared resources such as laboratories, test equipment, test aircraft, etc.)
- Late Completions may be caused by prior late starts, tasks encountering unanticipated problems, or tasks that were under-scoped when originally bid/scheduled



Metrics Application: Late Starts/Completes₂



Interpretation (cont.):

- Late Starts and/or Late Completions may also be caused by missed inter-team dependencies
 - These causes can be particularly onerous; should be analyzed further to determine whether
 - such events are single occurrences or
 - there are underlying problems that are likely to drive future recurrences
- Late Starts, Late Completions for tasks on the critical path are red flags of future schedule issues
 - If the WBS is insufficiently detailed to identify a critical path, this is by itself a red flag indicating that there is insufficient understanding of the structure of the work to be done to complete the project.



Metrics Application: Late Starts/Completes₃



Potential Questions:

- Are the occurrences increasing or decreasing?
- If the late start/complete is on a critical path, what is the net effect to the schedule as planned?
- What is the impact on the schedule reserve?
- What are the options to recover schedule reserve?
- What is the impacts on the cost given the schedule slip?
- Does replanning or shifting of tasks/priorities need to occur in order to meet the requested schedule?

Other potential questions?



Metrics Application: Late Starts/Completes₄



Related Metrics to Cross Reference: Correlation of this data should appear in one or more of the following progress measures:

- Cost & Schedule Variance
- Cost & Schedule Performance
- SW Development Phase Progress
- HW Development Phase Progress
- SIL, HILL, Ground & Flight Test Progress

Other potential questions?



Questions?





Contact Information

Mike Bandor

Senior Engineer

Telephone: +1 210-380-5563

Email: mbandor@sei.cmu.edu

Tom Merendino

Sr. Member of the Technical Staff

Telephone: +1 412-268-1154

Email: tjm@sei.cmu.edu

U.S. mail:

Software Engineering Institute

Customer Relations

4500 Fifth Avenue

Pittsburgh, PA 15213-2612 USA

Customer Relations

Email: customer-relations@sei.cmu.edu

Telephone: +1 412-268-5800

SEI Phone: +1 412-268-5800

SEI Fax: +1 412-268-6257

World Wide Web:

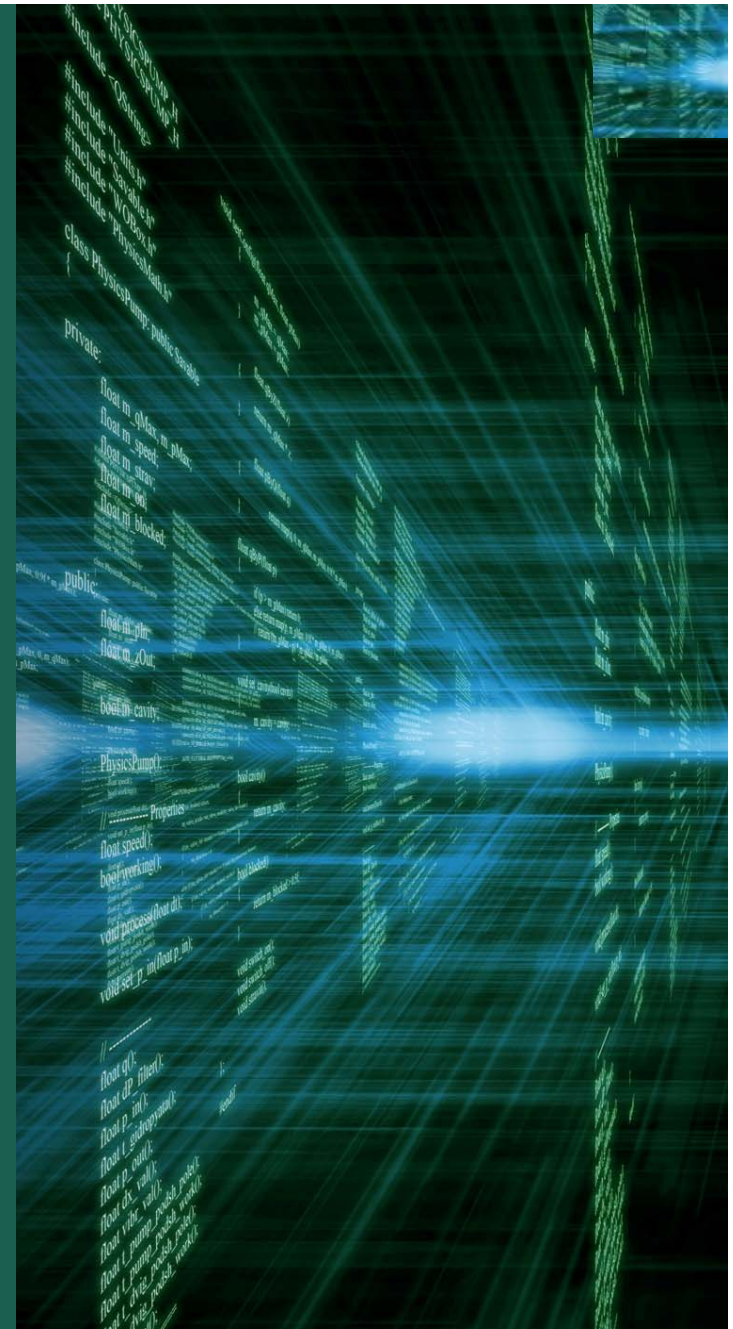
www.sei.cmu.edu

www.sei.cmu.edu/contact.html



Dashing All the Way: Defining the Best Dashboard for Your Program

Backup: Drill-Down Details for the Example Dashboard Quadrants and Metrics Labels





Schedule, Cost & Resources

As noted on a prior chart, metrics that are included in the Schedule, Cost & Resources quadrant focus on the project or program status to make sure it delivers on its forecast date at the forecast cost. Typical metrics found in this section are cost and schedule (Earned Value where applicable), staffing and other resources.

General questions to think about:

- Has the project/program been predictable in recent milestone performance?
- Are forecasts regularly updated (every program review where applicable)?
- Does the current forecast predict a problem in execution?
- Can the plan be made executable within the project contingency?





Interpretation: Task Completion to Schedule – Late Starts/Completions₁

Task Completion to Schedule tracks late starts and late completions

Late Starts and Late Completions are reliable indicators of potential schedule risk and should be carefully tracked and analyzed for causes

Late Starts may be caused by resources issues (staffing shortfalls, contention for shared resources such as laboratories)

Late Completions may be caused by prior late starts, tasks encountering unanticipated problems, or tasks that were under-scoped when originally bid/scheduled

Late Starts and/or Late Completions may also be caused by missed inter-team dependencies. These causes can be particularly onerous and should be analyzed further to determine whether such events are single occurrences or whether there are underlying problems that are likely to drive future recurrences


Late Starts, Late Completions for tasks on the critical path are red flags of future schedule issues. [If the WBS is insufficiently detailed to identify a critical path, this is by itself a red flag indicating that there is insufficient understanding of the structure of the work to be done to complete the project.]




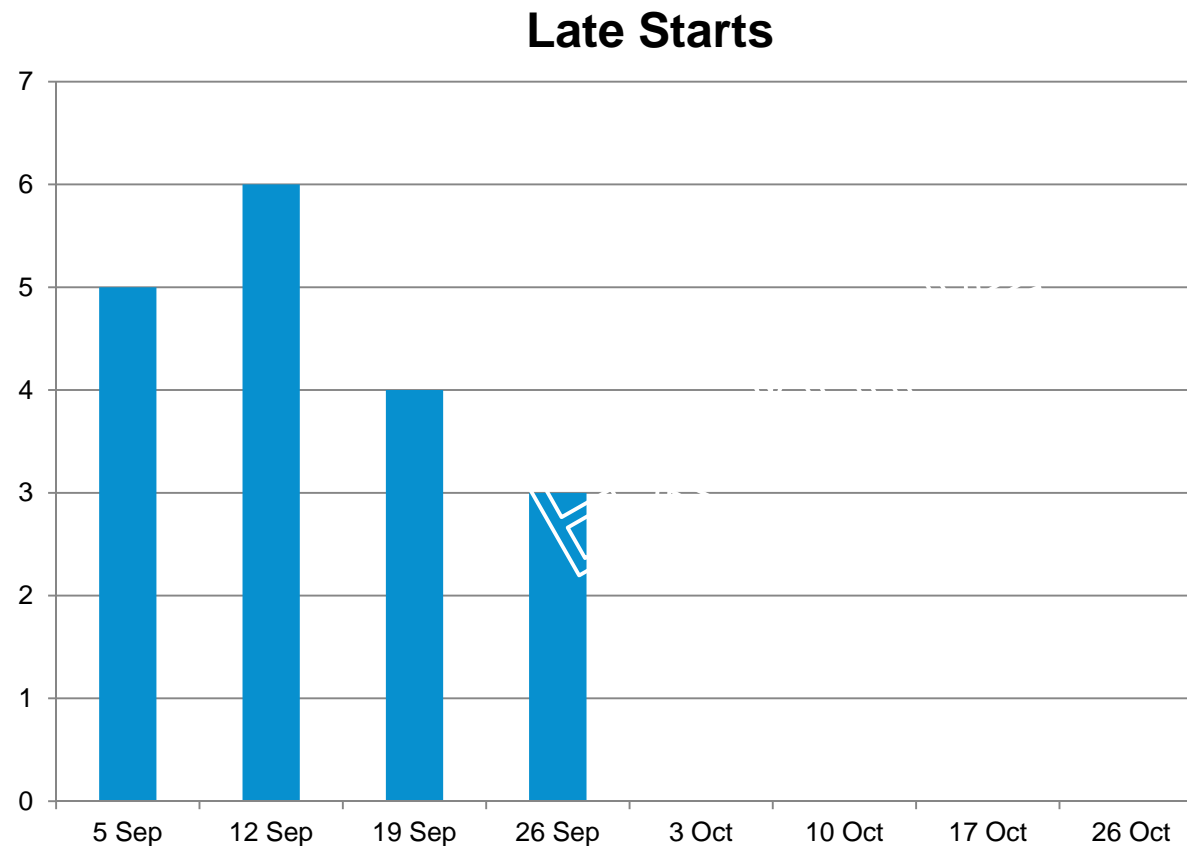


Example: Task Completion to Schedule – Late Starts/Completions₂

: Late Starts

Status: 

Trend: 



■ Late Starts



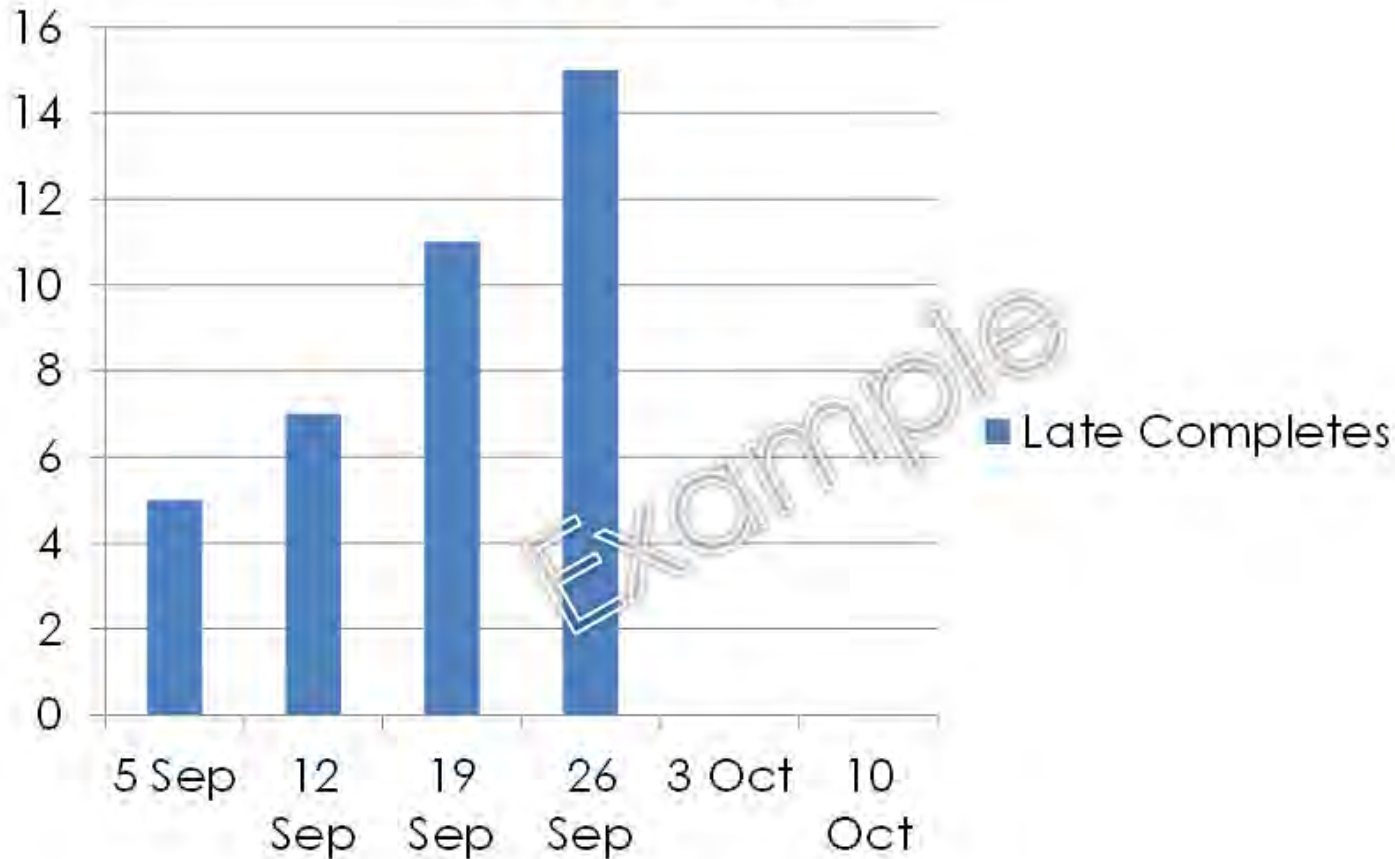


Example: Task Completion to Schedule – Late Starts/Completions.

: Late Completes Late Completes

Status: 

Trend: 





Interpretation: Staffing₁

Applicable to both contractor side and government side staffing

Projected staffing profile and associated skill level requirements will usually vary over the project lifecycle

At any given point in the program, notable variations from projected profile are important to note, and it is always prudent to bore into the details to determine the nature of the skill shortage so that potential consequences may be pinpointed.

Staffing shortfalls are reliable early indicators of future late starts, late finishes, schedule slips

Staffing excesses are early indicators of potential cost overruns

Note that the roll-off of certain skills (e.g. software developers) toward the end of a program may be problematic should problems proliferate to cause reworks (e.g. integration and test)





Interpretation: Staffing₂

Cumulative staffing over time should track pretty closely to cost in most development projects, whose costs are dominated by labor. But note that management and certain 'specialty' labor may be rolled up in overhead and may not be directly reported in staffing

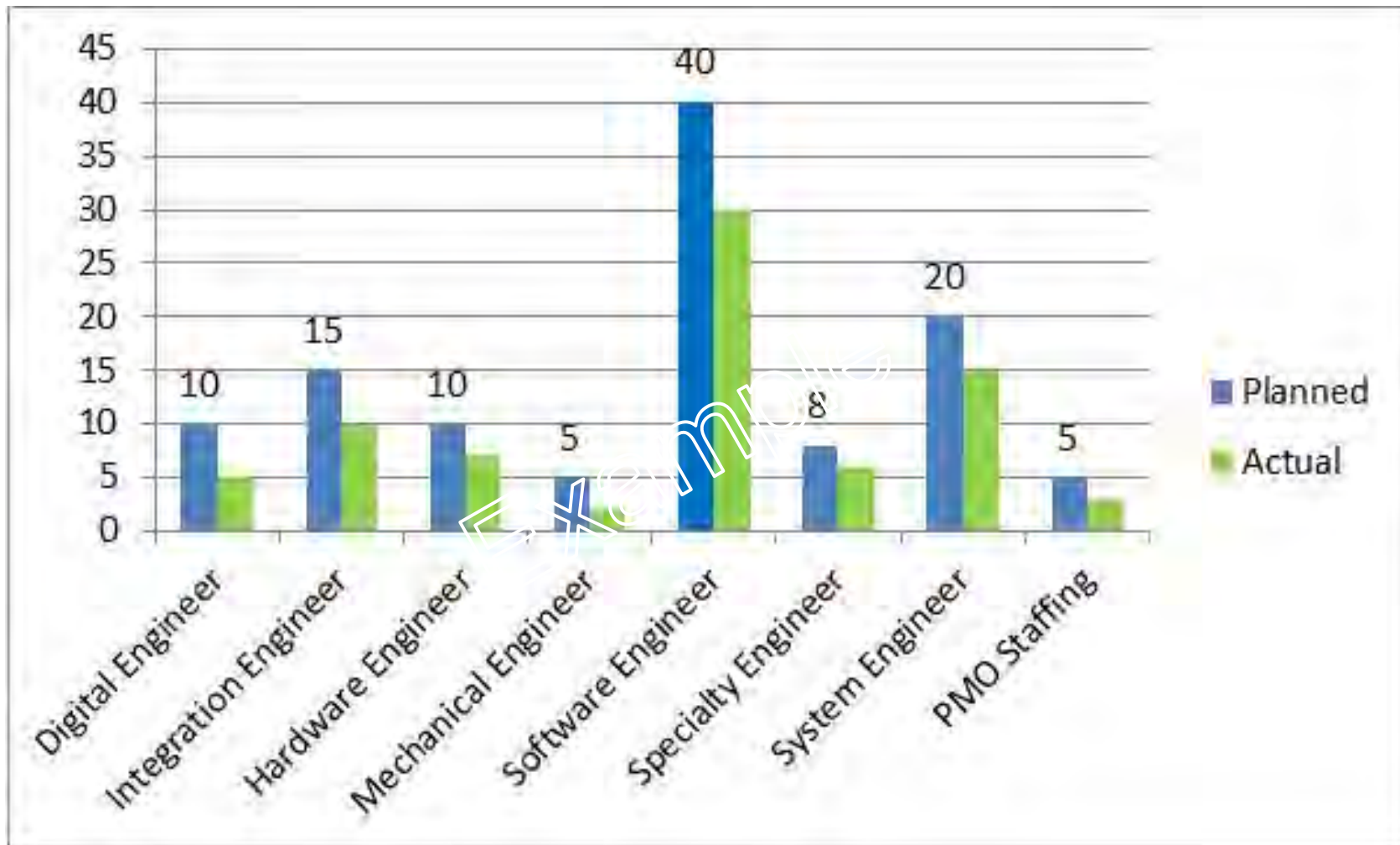
Period to Period Staffing Variations

- Look for consistency from period to period, large variations from projections should draw attention
- Consistent shortfalls are indicative of a growing bow-wave (look at potential schedule risk metric / trend to assess)
- Less margin toward end of schedule to mitigate staff shortfalls, so shortfalls toward end of schedule are potentially more significant than similarly sized shortfalls early in progress



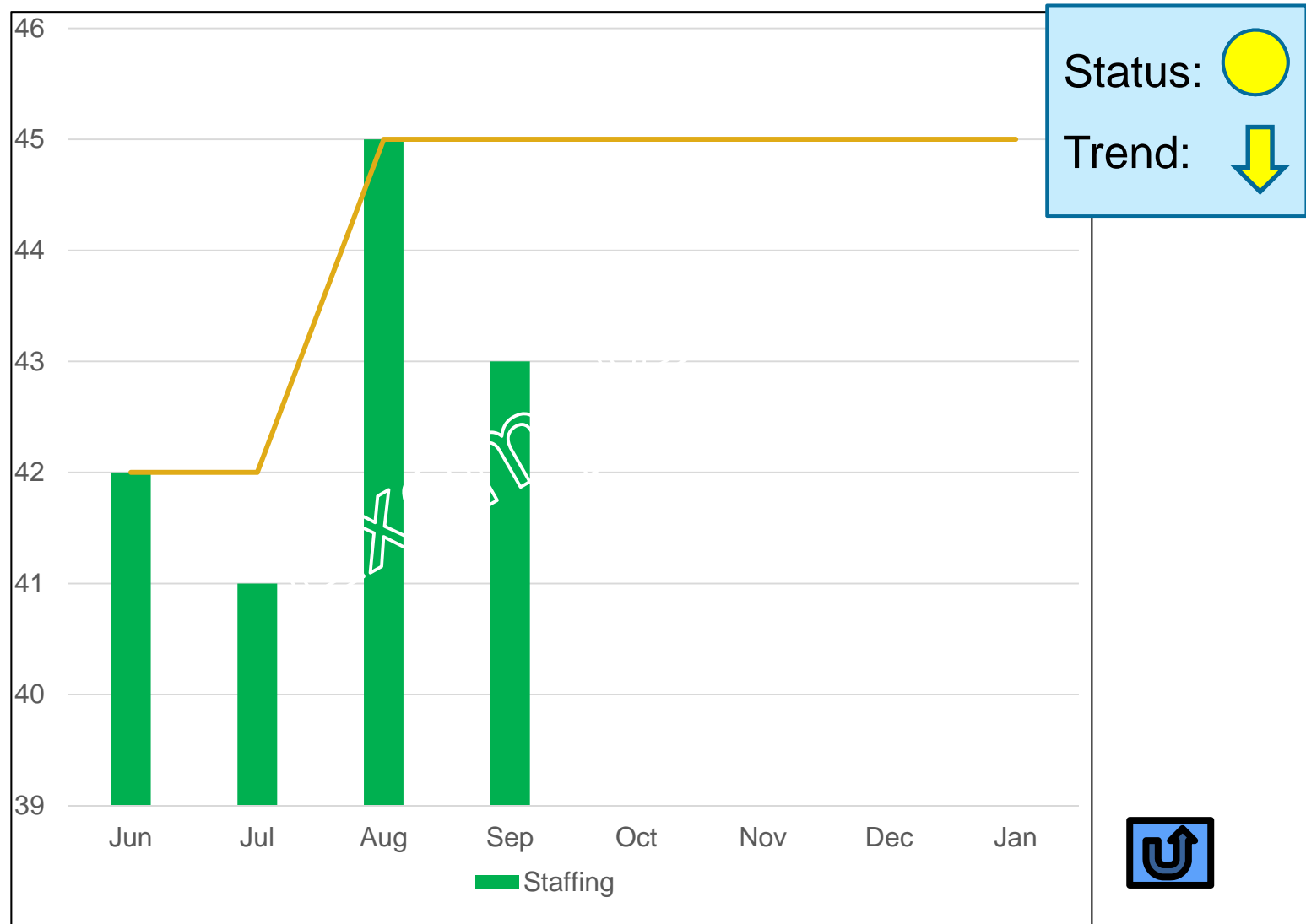


Example: Staffing (Contractor)₁





Example: Staffing (Contractor)₂





Scope, Progress & Change: Trending Analysis – General Information₁

If not prepared to provide estimate/predictions of progress against which Trending Analysis of Scope, Progress and Change can be reported, assume linear progress, and use experience to build up data to support more refined projections in future

Recommend all Trend charts that analyze progress include definition of values measured and actual data, even if reporting and analysis is done in terms of percentages

Cumulative Progress Trends

- Progress may have noise early on due to start up issues, but follow closely anyway
- Look for convergence to projections
- Consistently behind suggests potential scope issues
- Consistently ahead suggests resources may be available to deploy elsewhere if skills are applicable

Period to Period Progress Trends

- Look for consistency from period to period, large variations from projections should draw attention
- Consistent shortfalls are indicative of a growing bow-wave (look at potential schedule risk metric / trend to assess)
- Less margin toward end of schedule to mitigate shortfalls, so shortfalls toward end of schedule are potentially more significant than similarly sized shortfalls early in progress

Applicable to Requirements Trends, Development Trends, Test Trends, Problem/Defect Reports





Scope, Progress & Change: Trending Analysis – General Information₂

Metrics that are mapped to the Scope, Progress & Change dashboard quadrant may be grouped into one of four Trending categories:

1. Metrics reported as a single snapshot value (delta or cumulative) of an assessed item during each reporting period. No associated end-goal value, projected value or projected range of values accompanies each reporting period's snapshot value
2. Metrics reported as a single snapshot value (delta or cumulative) accompanied with an overall end goal target value/range of values
3. Metrics reported as a single snapshot actual cumulative value accompanied with an associated projected cumulative value
4. Metrics reported as a single snapshot actual cumulative value with an associated range of projected cumulative values. The range may be specified as either a:
 - Single threshold value above (or below) which is considered acceptable
 - "Tolerance" band via max-min range.





Scope, Progress & Change: Trending Analysis – General Information₃

For Scope Progress & Change metrics that fall under Trending categories 1 & 2, the program office may need to separately keep track of reported period-to-period value changes in order to analyze trends of assessed items of interest

- Trending analysis results should be sensible for the assessed item at the given point in the lifecycle where the measured value was taken. For example, we might not expect a lot of period-to-period (P2P) change in results of trending analysis for a test metric during the design phase of the lifecycle. However, during the Integration Test lifecycle phase, we will be very focused on understanding the reasons underlying the period-to-period trending behavior for test metrics

For Scope Progress & Change metrics that fall under Trending categories 3 & 4 (i.e., metrics that have actual and projected cumulative values), the program office can perform two types of trending analysis that can assist in projecting the future “health” of the program (also, see next two charts for further description):

1. Potential schedule risk analysis
2. Period-to-period (P2P) Trends analysis with specified Tolerances

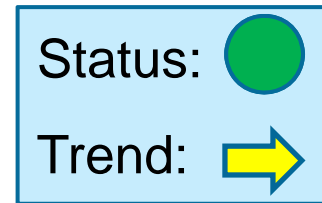
The above types of trending analysis can be done for any Scope, Progress and Change metrics that are reported during the Requirements, Development and Test lifecycle phases





Scope, Progress & Change: General Trending Analysis If Planned vs. Actual Cumulative Values are Available₁

Comprised of two items: “potential schedule risk” and “period-to-period (P2P) schedule risk trend”:




1. Potential Schedule Risk, R [R, as percentage, indicates likelihood of future schedule slips]


- Similar to SPI metric, done in terms of item being assessed (e.g. requirements completed, design docs completed, code unit tests completed, etc.)
- R can be used as a “rough and ready” estimate of likelihood (%) of future schedule slip based on progress to date whenever Projected Cumulative and Actual Cumulative values are available for the assessed item
- Predictive metric, assuming future performance similar to cumulative performance to date:
 - $R \sim \{(Projected\ Cumulative / Actual\ Cumulative) - 1\}$
- Look for convergence to zero
- Identify thresholds A, B ($A < B$) (Note – contractor / govt collaboration needed- these will affect Status!)
- Status green if $R \leq A$
 - i.e., “A” represents a % of schedule risk below which is considered “noise” or acceptable levels of risk
- Status yellow if $A < R \leq B$
- Status red if $R > B$
 - i.e., “B” represents a % of schedule risk above which needs further management analysis and decisions
- A persistent positive value for R suggests likelihood (%) that a “bow wave” is developing, increasing the chance of schedule slips. Smaller positive values for R are “better” in the sense of implying lower risk
- If R consistently negative, suggests potential to complete early and related activities need analysis to determine if this represents an opportunity that may be exploited. Larger negative values for R are “better” in the sense of implying lower risk

Applicable to Requirements, Development, and Test Trends



Scope, Progress & Change: General Trending Analysis If Planned vs. Actual Cumulative Values are Available₂

Status: 

Trend: 

2. Period-to-period (P2P) Schedule Risk Trend with specified Tolerance

- Identify **T** as a Tolerance band for Potential Schedule Risk percentages, **R**
 - **T** is to be used for comparing period-to-period changes in the value of Potential Schedule Risk, **R**
 - (Note – periodic contractor / govt collaboration needed to set/retain value of **T** - will affect Trend report!)
- Value of **R** is as calculated on the previous chart for item being assessed
 - **R_{prior}**, **R_{current}** represent the Potential Schedule Risk percentages as calculated for consecutive reporting periods **P_{prior}**, **P_{current}**, respectively

For consecutive reporting periods **P_{prior}**, **P_{current}** :

- If $|R_{prior} - R_{current}| < T$, then **Trend is flat** and is same color as status
 - i.e., Trend is considered “**Flat**” when difference between two consecutive estimates of Potential Schedule Risk **R** is within Tolerance band
- If $R_{prior} + T < R_{current}$, **Trend is pessimistic** (negative) and is same color as status, explain at status meeting
 - i.e., Trend is considered “**Pessimistic**” when, for two consecutive periods, the current estimate of Potential Schedule Risk is more than a “Tolerance” greater than the prior estimate of Potential Schedule Risk
- If $R_{prior} - T > R_{current}$, Trend is “**Optimistic**” (positive) and is same color as status, explain at status meeting
 - i.e., Trend is considered “**Optimistic**” when, for two consecutive periods, the current estimate of Potential Schedule Risk has reduced below the prior period’s estimate of Potential Schedule Risk less the Tolerance
- Initial recommendation for handling (change via contract / govt collaboration):
 - If three or more periods have **Pessimistic** trend, Trend is Red, detailed examination + written report
 - If three or more periods have **Optimistic** trend, Trend is Green, explain at status meeting

Applicable to Requirements, Development, and Test Trends



Interpretation: Requirements Trends – Sizing₁



Requirements Sizing

The intent of tracking this metric is to use changes to approved system level (B-level) requirements as an early indicator to understand whether the *scope* of the system development effort is changing relative to the scope which was used to plan the IMS/project schedule

Requirements sizing metric intends to track the number of B-Level system requirements being defined and approved during the contractor's System Development Phase. This metric may also include period-to-period changes in wording made to existing approved requirements, even when the total number of requirements has not changed relative to prior periods. Wording changes to approved requirements may imply significant changes in scope which needs to be analyzed and understood for potential management decisions.

For full lifecycle development programs, the contractor specifies a goal for the expected total number of system requirements to be approved through the SRR milestone event. Periodic metric updates to the cumulative number of requirements approved up to a point in time are reported as a single value against the overall goal.

Interpretation: Requirements Trends – Sizing₂



Requirements Sizing

Period-to-period changes in the number of approved requirements after the SRR event may indicate unplanned changes in program scope. Thus, post-SRR, you look to analyze stability in the number of approved period-to-period requirements. You should question the reasons for any changes to ensure the degree of requirements scope creep, if any, is understood and whether management decisions are needed to contain scope.



Interpretation: Requirements Trends – Sizing₃



Requirements Sizing

Significant differences between the goal for the expected number of approved requirements at SRR and the actual total number of approved requirements that exist at the SRR event needs to be analyzed for program scope implications.

- *If it is determined that program scope has changed, increased schedule risk likely exists for future tasks. Engineering investigation and proactive management decisions are likely needed for scope containment.*

Significant changes in the number of approved requirements, or in contents changes made to approved requirements, *after* SRR needs to be analyzed for program scope implications.

- *If it is determined that program scope has changed, increased schedule risk likely exists for future tasks, Engineering investigation and proactive management decisions are likely needed for scope containment.*



Interpretation: Requirements Trends – Sizing₄



Requirements Sizing

For programs restricted to the maintenance lifecycle, the system level “requirements” are essentially specified by the number of system change requests (SCRs) or Defect Reports (DRs) that are authorized for implementation within in a given release.

These SCRs/DRs are identified and authorized before the program to create the release is started. Thus, there is no formal “System Definition” phase in maintenance-only releases during which system requirements are being developed. Rather, you should track changes to program scope via the dashboard metric “Planned Work Deferred to other contracts” to identify a change in scope for a given release. Additionally, there may be other means to determine whether recommendations have been made to defer a number of planned SCRs/DRs. Such deferrals may imply significant changes in release scope, and early management decisions may be needed to approve and accommodate such deferrals.





Interpretation: Requirements Trends – Sizing₅

Requirements Sizing

Note: In the following, a “significant content change” to an approved requirement means any revision to a requirement’s content that impacts the requirement’s original cost or difficulty estimate by at least 5%. *NOTE: this measure may not be presently included within the contractor’s proposed assessed items*

INDICATIONS:

GREEN:

1. Actual total number of approved system requirements, at or after SRR, is within the open range $\pm 2\%$ of the expected total goal **AND**
2. Actual total number of approved system B-level requirements that had “significant content changes” after SRR is less than 2% of the expected goal.

YELLOW:

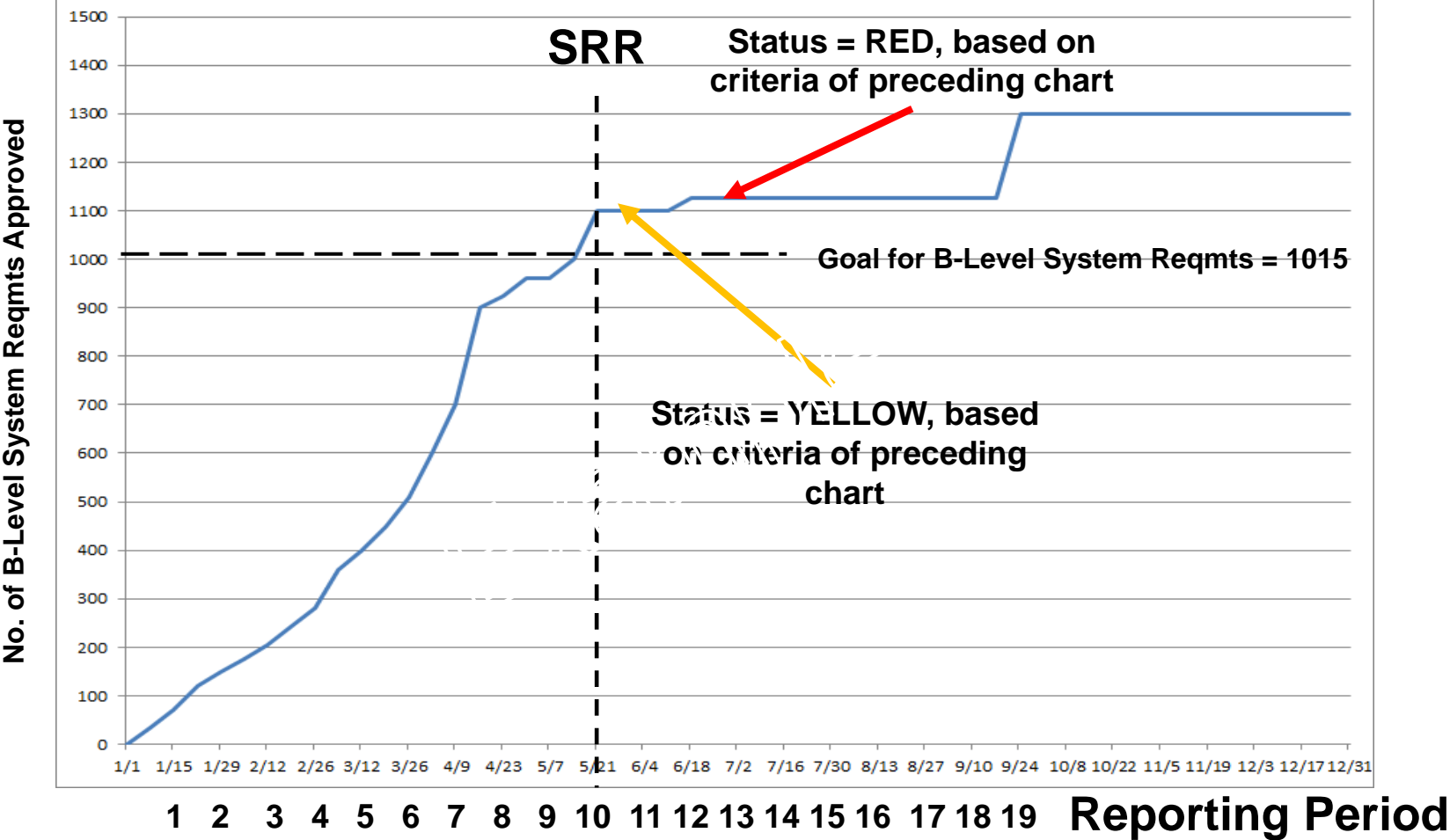
1. Actual total number of approved system B-level requirements, at or after SRR, is in the closed range of $\pm 2\%$ to $\pm 5\%$ of the expected total goal **OR**
2. total number of approved system B-level requirements that had “significant content changes” after SRR is in the closed range of $\pm 2\%$ to $\pm 10\%$ of the expected goal

RED:

1. Actual total number of approved system B-level requirements, at or after SRR, exceeds $\pm 10\%$ of the expected total goal **OR**
2. total number of approved system B-level requirements that had “significant content changes” after SRR exceeds $\pm 10\%$ of the expected goal



Example: Requirements Trends – Sizing





Scope, Progress & Change: Development Trends - Software (General Information -1)

SW Development Phase Progress

The Contractor's proposed metrics for reporting SW Phase Progress should include metrics for 4 distinct development phase components:

1. SW Requirements Phase Progress
2. SW Design Phase Progress
3. SW Implementation Phase Progress
4. SW Integration Test Phase Progress

A program proposes progress metrics for all four SW development phase progress categories noted above.

In addition, contractors employing a form of agile development procedures are likely to include a "Story Point" approach for assigning an amount of effort to develop and test the capabilities implemented during each iteration or sprint. Metrics that show progress against the assigned Story Points in the release can be used as one approach to reporting software development progress.



Scope, Progress & Change: Development Trends - Software (General Information -2)

SW Development Phase Progress - Agile

For contractors employing agile development processes, each iteration or sprint is likely to be a few weeks in duration, with each one being of the same duration.

A total “release’s – worth” of story points may be assigned to a consecutive series of iterations or sprints to represent the development activities that comprise a release.

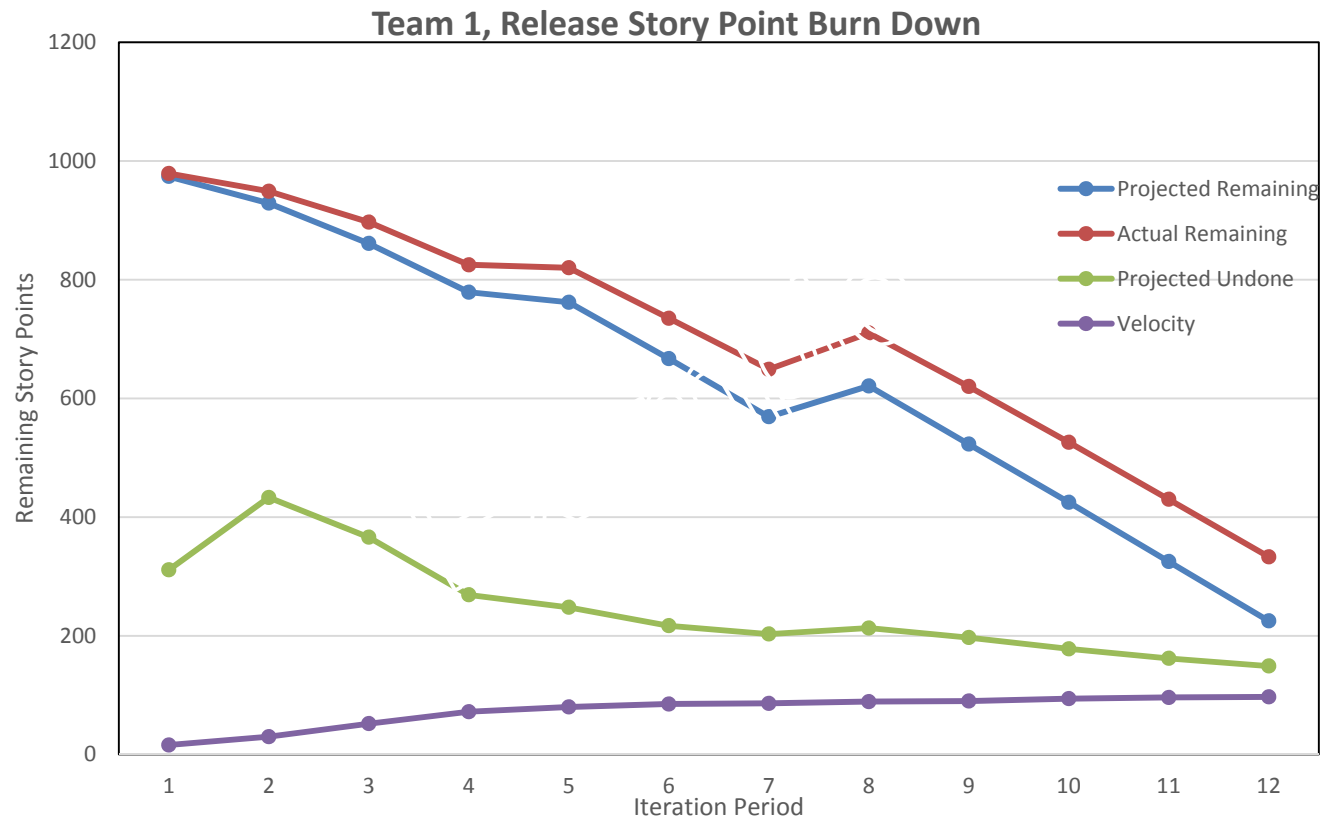
A story point burndown chart showing projected versus actual story points implemented and tested within each iteration / sprint will assist in assessment of software development progress.

If more than one agile team is involved in the contractor’s Release plan, a story point burn down chart for each team may be provided to report progress



Example: Development Trends – Software









SW Development Phase Progress – Story Points Example





Scope, Progress & Change: Software Development Phase Progress₁

Dashboard Label: Development Trends (Software)

Contractor Metrics that map to this Label:	Status	Trend	Rationale
SW Development Phase Progress: SW Requirements Progress			Actual progress in approving SW reqmts is 10% behind projected, which is at the extreme upper end of the yellow threshold. SMEs needed to write reqmts have been overtasked over last several reporting periods. Trend has been flat. May delay future related SW implementation tasks
SW Development Phase Progress: SW Design Progress			Actual progress is tracking well to projected progress.
SW Development Phase Progress: SW Implementation Progress			Actual progress is tracking well to projected progress
SW Development Phase Progress: SW/HW Integ Test Progress			Actual progress is significantly behind projected progress. However, this early in the lifecycle, it is not unexpected that Integration Test progress can have wild swings, period to period. This metric will not presently inform the dashboard value

The Green and Yellow statuses result in reporting Yellow for the Development Trends (SW) Dashboard Label. It is too early in the lifecycle to collect and report SW/HW Integ Test Progress, so Red status for progress is not unexpected at this period and does not impact dashboard status





Interpretation: Development Trends – Software₁

SW Development Phase Progress

For this SW Development Phase Progress metric, the contractor reports four separate “percent-complete” components each period:

- **SW Requirements Progress** – considered “100% done” when software requirements are approved for a specific release
- **SW Design Progress** – considered “100% done” when software design documents for a specific release have been through contractor peer review and are approved
- **SW Implementation Progress** – considered “100% done” when all software for a specific release has successfully passed through unit test
- **SW/HW Integration Test Progress** – considered “100% done” when an entire release has successfully passed through SW/HW Integration Test

Each of the four components has both a “projected cumulative” value and an “actual cumulative” value. This results in eight separate values being reported each period.

The projected and actual cumulative values for percent complete that are reported are estimated by the contractor based on considering funding level expended, elapsed schedule time, and estimated effort to complete

The interpretation approach is based on the Potential Schedule Risk and Period to Period Trends analysis as described earlier in this presentation. This analysis is applied separately to each of the four measured components, examples shown in the following charts

Interpretation: Development Trends – Software₂



SW Development Phase Progress

Potential Schedule Risk analysis:

Identify lower **A** and upper **B** thresholds for Potential Schedule Risk and period-to-period risk tolerance, **T**, for Software Development Progress metric

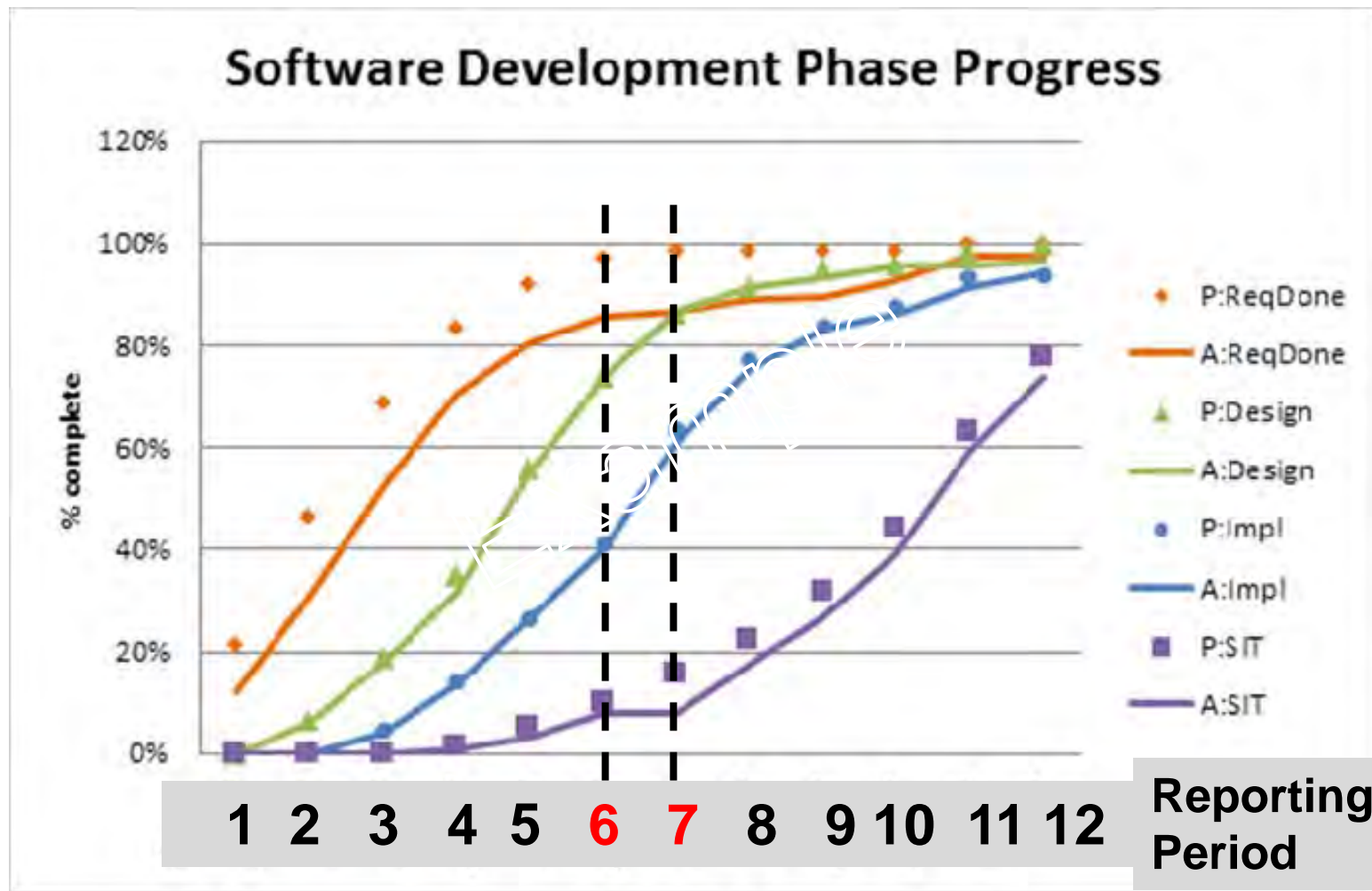
- “**A**” = 5%; “**B**” = 10%; “**T**” = 5%
- Status green if $R \leq A$
 - i.e., “A” represents a % of schedule risk below which is considered “noise” or acceptable levels of risk
- Status yellow if $A < R \leq B$
- Status red if $R > B$
 - i.e., “B” represents a % of schedule risk above which needs further management analysis and decisions

Period-to-Period Risk Trend analysis:

- If $|R_{\text{prior}} - R_{\text{current}}| < T$, then **Trend is flat** and is same color as status
 - i.e., Trend is considered “**Flat**” when difference between two consecutive estimates of Potential Schedule Risk **R** is within Tolerance band
- If $R_{\text{prior}} + T < R_{\text{current}}$, **Trend is pessimistic** (negative) and is same color as status, explain at status meeting
 - i.e., Trend is considered “**Pessimistic**” when, for two consecutive periods, the current estimate of Potential Schedule Risk is more than a “Tolerance” greater than the prior estimate of Potential Schedule Risk
- If $R_{\text{prior}} - T > R_{\text{current}}$, Trend is “**Optimistic**” (positive) and is same color as status, explain at status meeting
 - i.e., Trend is considered “**Optimistic**” when, for two consecutive periods, the current estimate of Potential Schedule Risk has reduced below the prior period’s estimate of Potential Schedule Risk less the Tolerance
- If three or more periods have **Pessimistic** (negative) trend, Trend is Red, detailed examination + written report
- If three or more periods have **Optimistic** (positive) trend, trend is Green, explain at status meeting



Example: Development Trends – Software₃





Interpretation: Development Trends – Software₄

Example SW Development Phase Progress analysis for Periods 6 to 7 on prior chart

Cumulative SW Devel Phase Progress reported at periods 6-7 (from prior chart)	Analysis of SW Dev Phase Progress at reporting period 7
Actual Requirements Progress ₆ = 85% Projected Requirements Progress ₆ = 95% Actual Requirements Progress ₇ = 86% Predicted Requirements Progress ₇ = 95%	<u>SW Requirements Progress₇:</u> $R_6 = [95/85-1] = 12\% > B \rightarrow \text{Status}_6 \text{ is RED}$ $R_7 = [95/86-1] = 10\% \leq B \rightarrow \text{Status}_7 \text{ is YELLOW}$ $\text{Trend}_7 = R_6 - R_7 = 2\% < T \rightarrow \text{Trend}_7 \text{ is FLAT}$
Actual SW Design Progress ₆ = 73% Projected SW Design Progress ₆ = 73% Actual SW Design Progress ₇ = 85% Projected SW Design Progress ₇ = 85%	<u>SW Design Progress₇:</u> $R_6 = [73/73-1] = 0\% < A \rightarrow \text{Status}_6 \text{ is GREEN}$ $R_7 = [85/85-1] = 0\% < A \rightarrow \text{Status}_7 \text{ is GREEN}$ $\text{Trend}_7 = R_6 - R_7 = 0\% < T \rightarrow \text{Trend}_7 \text{ is FLAT}$
Actual SW Implementation Progress ₆ = 40% Projected SW Implementation Progress ₆ = 40% Actual SW Implementation Progress ₇ = 61% Projected Implementation Progress ₇ = 63%	<u>SW Implementation Progress₇:</u> $R_6 = [40/40-1] = 0\% < A \rightarrow \text{Status}_6 \text{ is GREEN}$ $R_7 = [63/61-1] = 3\% < A \rightarrow \text{Status}_7 \text{ is GREEN}$ $\text{Trend}_7 = R_6 - R_7 = 3\% < T \rightarrow \text{Trend}_7 \text{ is FLAT}$
Actual SW/HW Integ Test Progress ₆ = 8% Projected SW/HW Integ Test Progress ₆ = 11% Actual SW/HW Integ Test Progress ₇ = 8% Projected SW/HW Integ Test Progress ₇ = 16%	<u>SW/HW Integ Test Progress₇:</u> $R_6 = [11/8-1] = 37\% > B \rightarrow \text{Status}_6 \text{ is RED}$ $R_7 = [16/8-1] = 100\% > B \rightarrow \text{Status}_7 \text{ is RED}$ $\text{Trend}_7 = R_6 + T = 42\% < R_7 \rightarrow \text{Trend}_7 \text{ is Pessimistic}$





Process & Risk Quadrant

This section is focused on keeping the costs and quality under control, and to monitor risks to the program/project.

Typical metrics found in this section can include audit results, quality checks (QA), behavioral observations to determine if new technology is properly deployed as well as risk monitoring.

General questions to think about:

- Are the processes being followed (contractor and government)? If not, why not?
- Is performance data available and reviewed when the data suggests?
- What program/project risks might exist because of process program (e.g. QA questions)
- Are technology changes reflected in process changes (e.g. training for new tools, etc..)?





Interpretation: Process & Risk -- Action Item Aging

Action item aging provides a means to highlight for further review those actions items that are not being retired in a timely manner

Action item resolution scheduling and delegation should always be done to provide a basis for aging

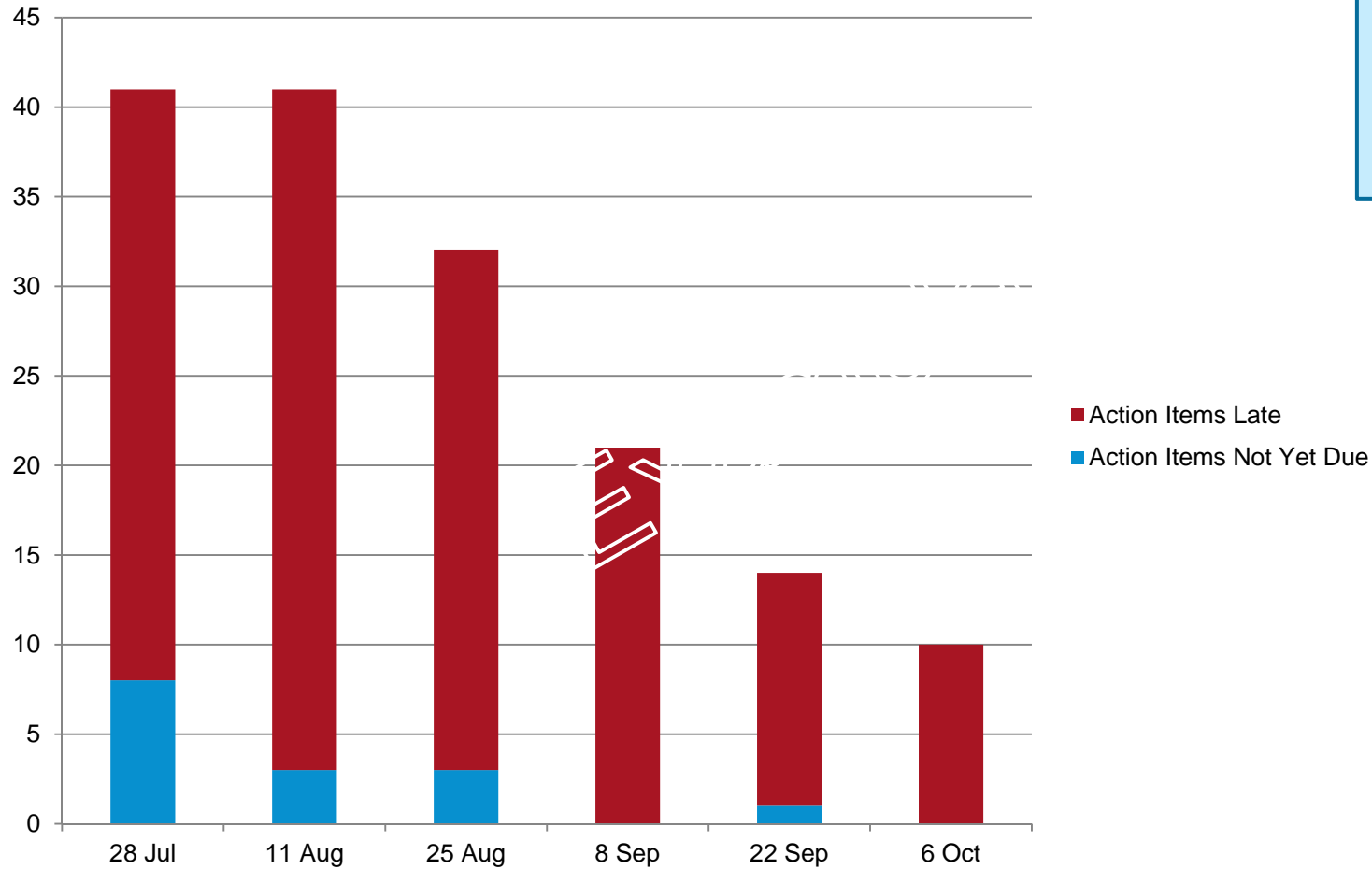
Older unresolved action items should be further investigated to identify highly potentially unrecognized risks, shortage of desired skill sets, or other resource bottlenecks


Recommend keeping track of tasks and actions contingent upon resolution of all actions items to identify potential downstream consequences


Need to identify potential consequence thresholds to identify potentially high consequence (red), potentially medium consequence (yellow), and potentially low consequence (green) actions items

Need to identify aging thresholds to identify lateness of action items. Recommend identification of criteria for declaring old and/or low consequence action items OBE and no longer applicable.

Example: Process & Risk – Action Item Aging



Status: 

Trend: 

Action Items Closed Last 2 weeks = 3





Quality Quadrant

The Quality Quadrant focuses on the ability to deliver a high quality product, as well as to validate claims of progress.

Typical metrics tracked in this section include Key Performance Parameter (KPP) and Technical Performance Measure (TPM) test results, verification and validation results, etc..

General questions to think about:

- Will the product achieve the desired quality and performance?
- Do validation checks (including Development Test and Operational Test) prove the value in use potential?





Interpretation: Quality – Peer Review/CDRL Comments

Peer Review or CDRL comments are direct evidence that CDRLs and/or other important artifacts are being given appropriate levels of scrutiny and adequate attention by SMEs

Number and sources of all comments must be tracked to enable analysis

Timeliness of Peer Review and CDRL comment completion allows more time for comprehensive resolution of comments. Track performance against suspense dates for both contractor-side and government-side comments.

For Artifacts other than CDRLs, Peer Review attendance by mandatory participants and comment closure rates are to be recorded for any artifact that requires a Peer Review sign-off per the associated engineering procedure

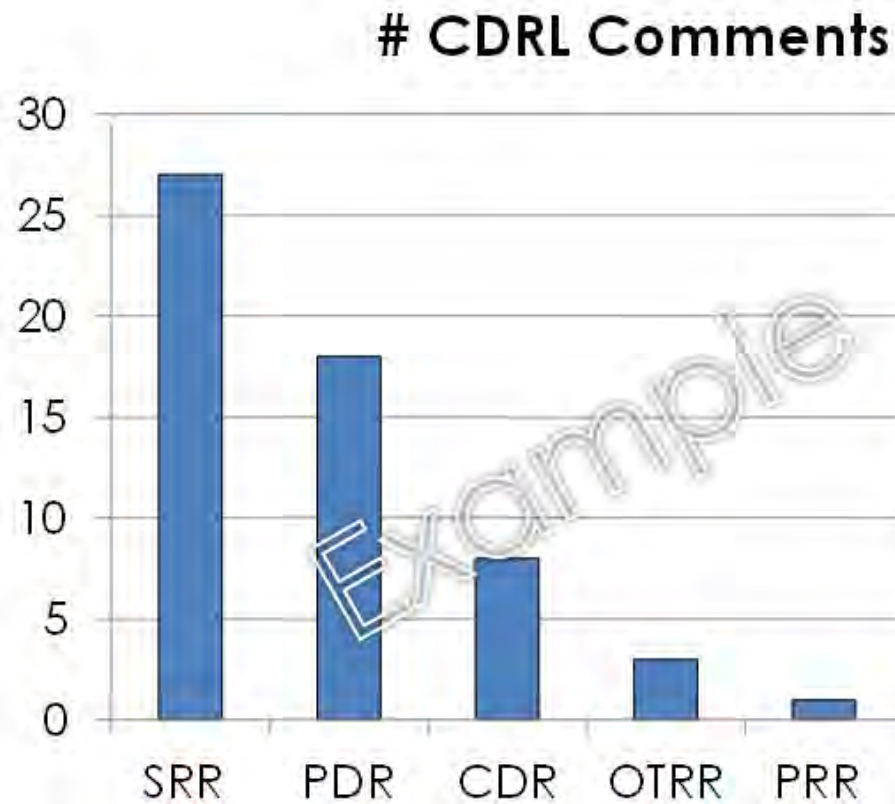
Look for quality comments that include

- Issue identification
- Rationale
- Proposed revisions and/or analysis to derive revision





Example: Quality – Peer Review/CDRL Comments by Milestone



Status: 

Trend: 

CDRL Comments

