

# **JCL in a Nutshell**

## **Exploring the Math of Joint Cost & Schedule Risk Analysis Through Illustrative Examples**

2010 SCEA Conference  
San Diego, CA  
Eric Druker

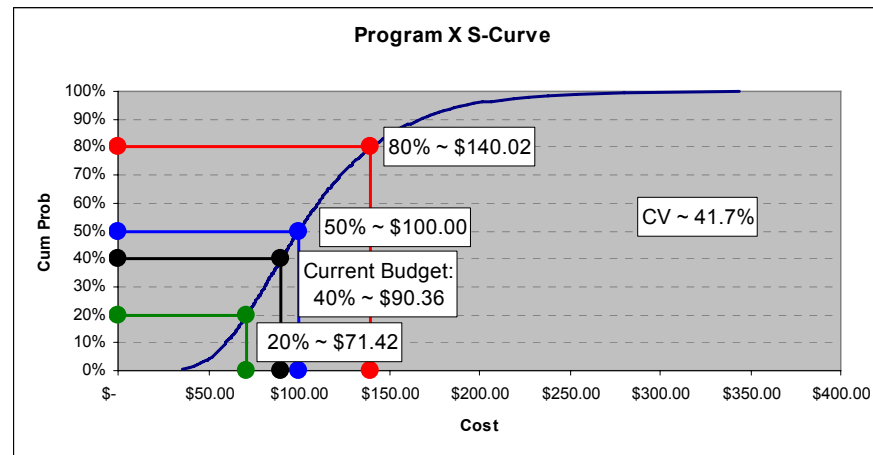
*This document is confidential and is intended solely for the use and  
information of the client to whom it is addressed.*

## Outline

- ▶ Introduction to Joint Cost & Schedule (JCL) Risk Analysis
- ▶ JCL Paradoxes
  - Typical Cost Risk Analysis Example
  - The Merge Bias
  - The Correlation Effect
- ▶ Schedule Risk Analysis Correlation Guideline Recommendations
- ▶ Conclusion

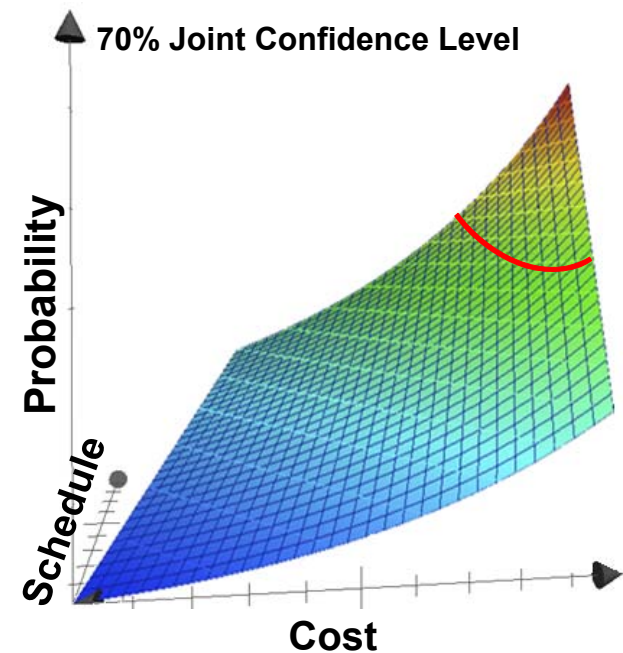
## What is Joint Cost & Schedule Risk Analysis?<sup>1</sup>

- ▶ **Definition:** A Joint Cost and Schedule Risk Assessment, sometimes known as Integrated Cost & Schedule Risk Assessment or Joint Confidence Level (JCL) Analysis, generates a joint probability distribution relating cost and schedule in a way that allows the analyst to determine the confidence level for meeting both target budgets and schedules simultaneously.
  - But what does this mean?
- ▶ Traditional cost or schedule risk analysis generates a distribution of potential final costs and durations from which confidence levels for budgets and schedules can be derived
  - These confidence levels are generated and reported separately



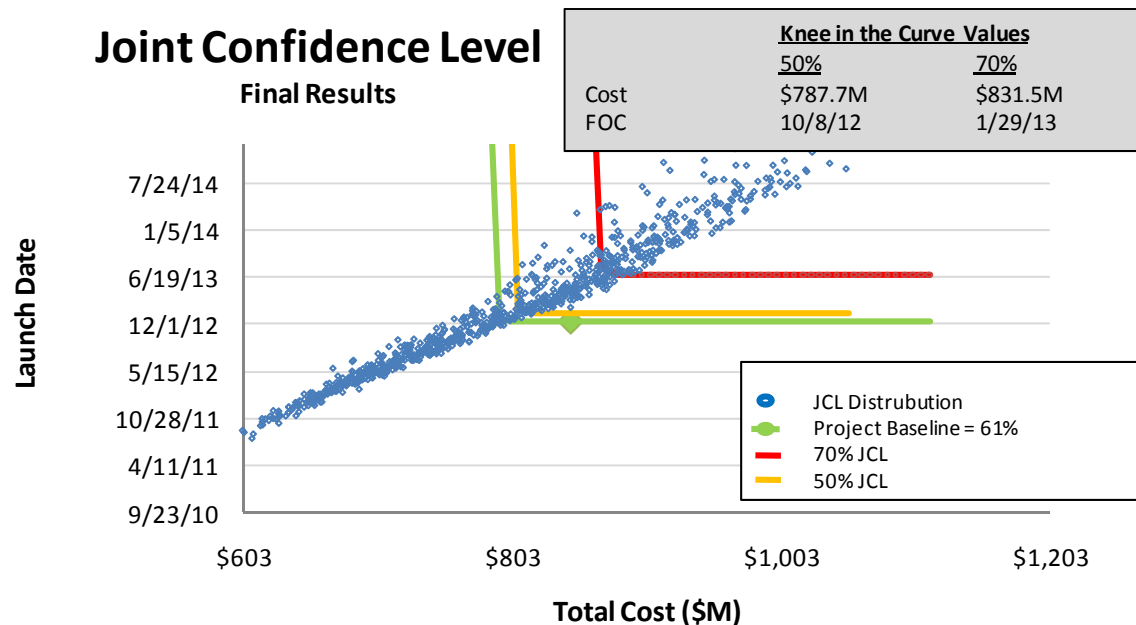
## What is Joint Cost & Schedule Risk Analysis?

- ▶ Joint Cost & Schedule Risk Analysis creates a bivariate distribution of cost and schedule
  - Thus, the confidence level of any cost and schedule pair represents the probability of the program finishing ***both at-or-under cost and on-or-ahead of schedule***
  - These are known as Joint Confidence Levels
- ▶ There are several methods for performing joint cost & schedule risk analysis
- ▶ At early program phases, parametric cost and schedule estimates/risk analysis can be combined to produce joint confidence levels
  - Using Monte Carlo or Copula methods
- ▶ When the program matures, and artifacts such as an integrated master schedule are developed, the build-up method can be used
  - This presentation will focus on paradoxes related to the build-up method



## What is Joint Cost & Schedule Risk Analysis?

- ▶ Another way to display Joint Cost & Schedule Risk Analysis results is through a scatter plot
  - Each point on the scatter plot represents 1 iteration of a Monte Carlo simulation performed on the JCL model
  - From this any % JCL can be uncovered



## What Goes In To a Build-Up Joint Cost & Schedule Risk Analysis?

### ► The Integrated Master Schedule (IMS)

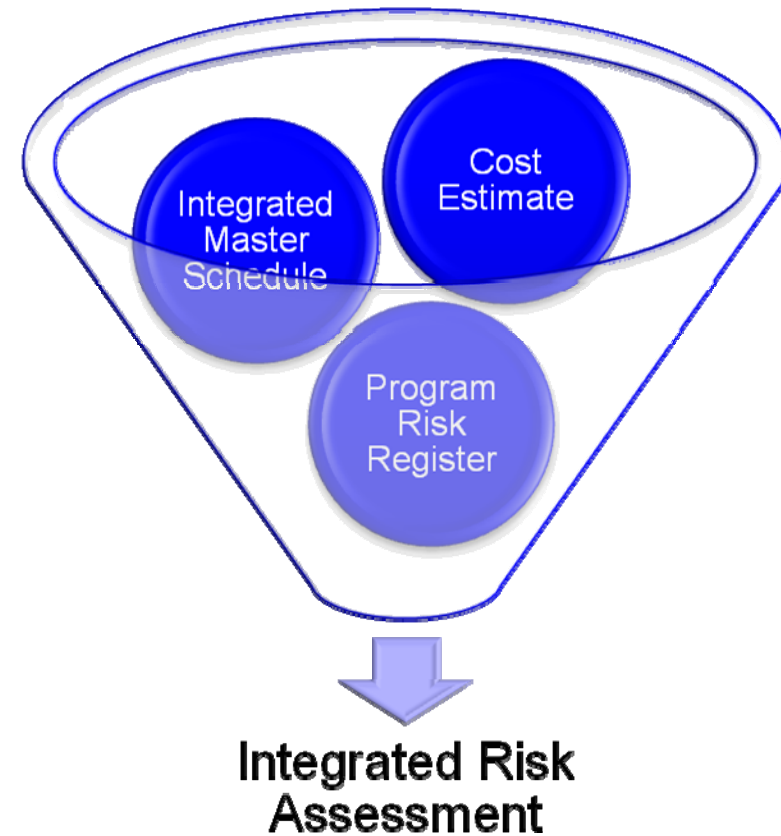
- A schedule health check is performed on the IMS to ensure logic structure
- Uncertainty around schedule tasks (at a level where there is sufficient insight) is quantified

### ► The Cost Estimate

- The cost estimate is loaded into the IMS at a summary level
- Uncertainty around the estimate is quantified; broken into:
  - Time-Dependent Costs: Increase as Schedule Grows
  - Time-Independent Costs: Independent of Schedule

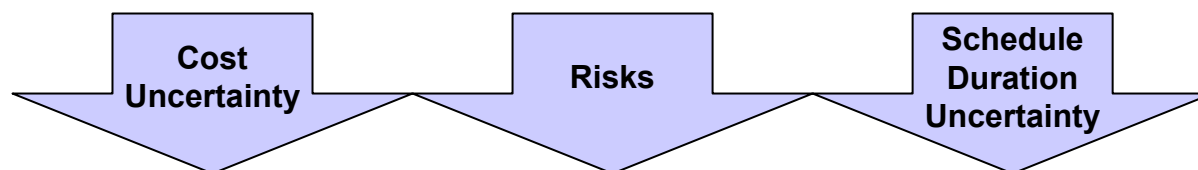
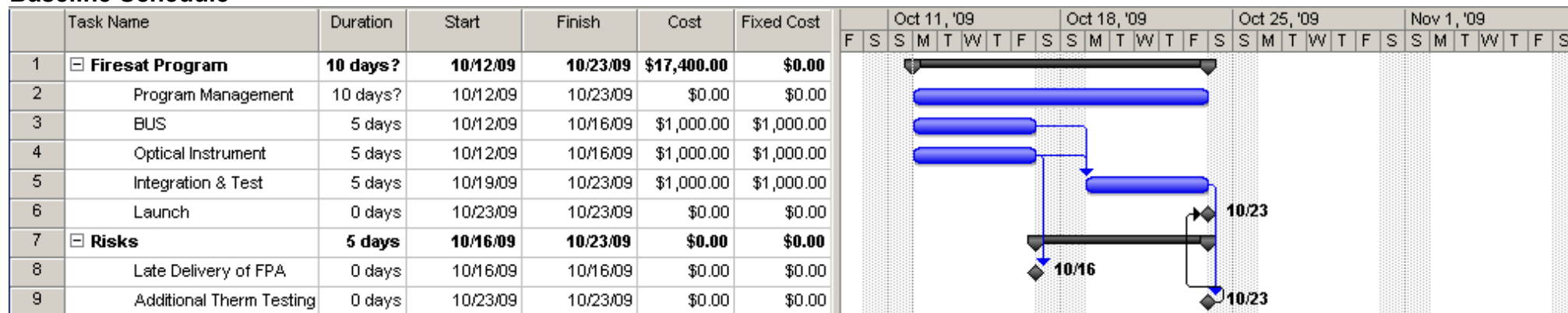
### ► Program Risk Register

- Risks managed as a part of the program's risk management plan are quantified in terms of cost and schedule impacts and mapped to tasks in the IMS as probabilistic events

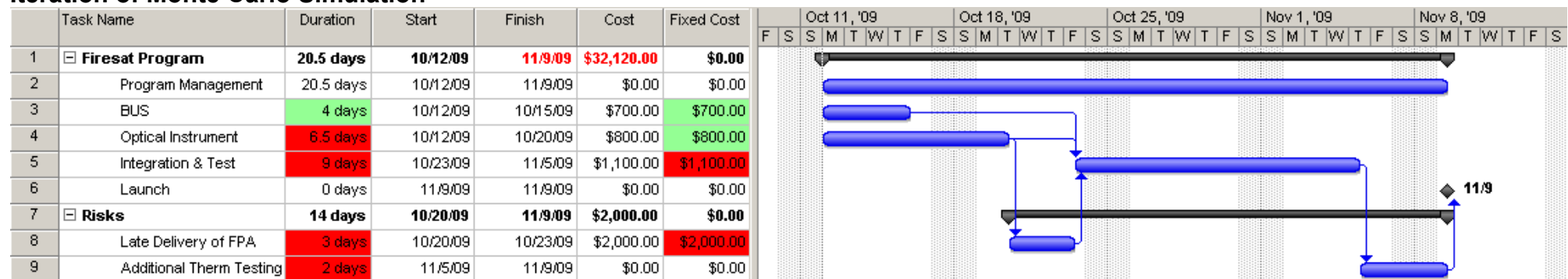


# Example JCL Model

## Baseline Schedule

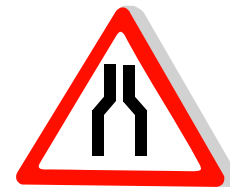


## Iteration of Monte Carlo Simulation



## The JCL Paradoxes

- ▶ The buildup method seems simple enough, why dedicate a paper to it?
  - There are two paradoxes of the build-up methodology that risk analysts (both cost and schedule) need to be aware of
- 1. Under most circumstances, the inclusion of parallel tasks in the schedule will cause the deterministic schedule to be at a low confidence level<sup>2</sup>
  - Known as the merge bias<sup>3</sup>
  - True even when symmetric uncertainty is applied
- 2. Correlation between schedule duration distributions has a direct effect on both the mean and variance of the risk adjusted program completion date and must be accounted for
  - In cost risk analysis, correlation only affects the spread of the cost distribution
- ▶ It is important that cost and schedule risk analysts understand both of these paradoxes as they are important to performing JCL analysis
  - Guidelines for correlation between schedule distributions will be included



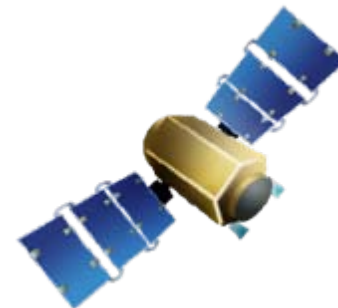
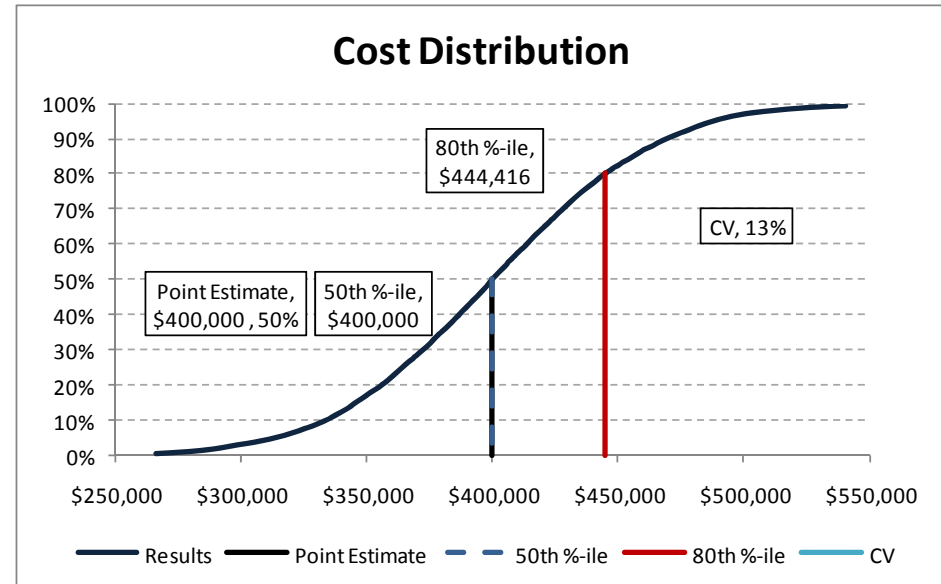


## A Typical Cost Risk Analysis

- ▶ Assume a satellite consists of 5 components, with the cost estimate for each distributed as a normal distribution with a mean of \$80K and a CV of 30%
- ▶ With no correlation, the total cost of the satellite has a mean of \$400K and a CV of 13%
  - Decrease of CV to 13.4% caused by “Square Root of n Effect”

$$13\%_{TotalCV} = \frac{30\%_{InputCVs}}{\sqrt{5}_{InputDistributions}}$$

- ▶ Cost risk analysis is fairly simple as it only deals with the summation of random variables
  - As would schedule risk analysis be if it only dealt with serial tasks
- ▶ Let's look at how schedule parallelism turns the above logic on its head



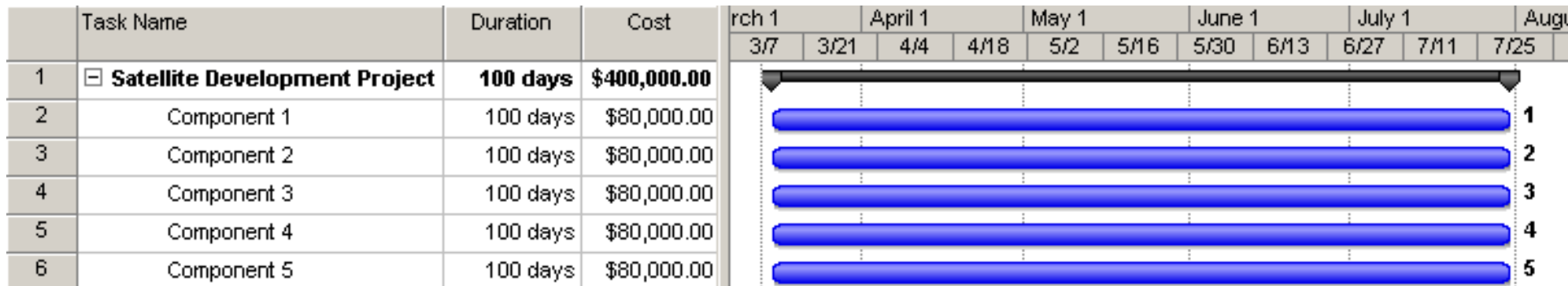
## The Merge Bias: A Simple Example

- ▶ Suppose you and your friend are driving from separate locations to meet up for dinner and you need to decide at what time to make reservations
- ▶ After talking, you both decide you will leave work at 5:00 pm
  - Each of you estimates that it will take 30 minutes to get to the restaurant (at the median)
  - Reservations are made for 5:30 pm, what is the probability you both will make dinner on time?
- ▶ If each driver has a 50% probability of arriving at dinner on time then, assuming independence, the probability the party as a whole arrives on time is 25%
  - Same as the probability of getting two heads in flips of a fair coin
- ▶ This example is intuitively simple, let's return to our satellite program

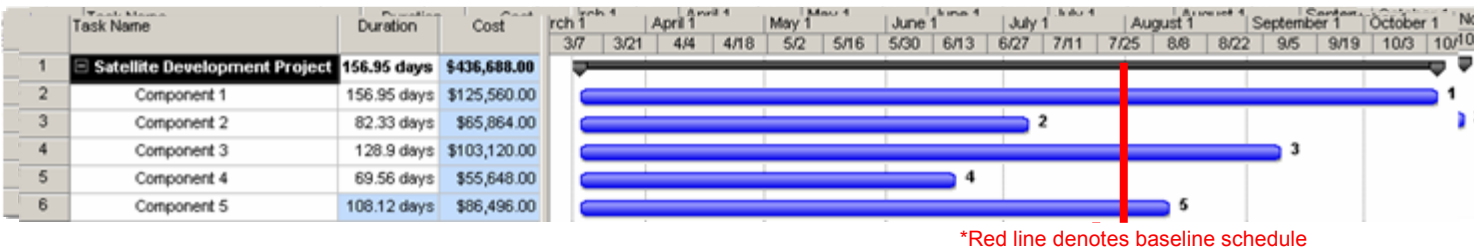


## The Merge Bias: A Satellite Development Program Example

- ▶ Satellite programs generally have several components being constructed in parallel
  - To complete the satellite, all components must be complete
- ▶ In this satellite example, 5 components are being built in parallel
  - What is the distribution of the completion date of the system? Of it's final cost?
- ▶ For simplicity, let us assume the duration to build each element is normally distributed with a mean of 100 days and a standard deviation of 30 days
  - In terms of cost, each schedule element spends at the rate of \$100 per hour
  - We'll use a Monte Carlo simulation to perform this JCL analysis



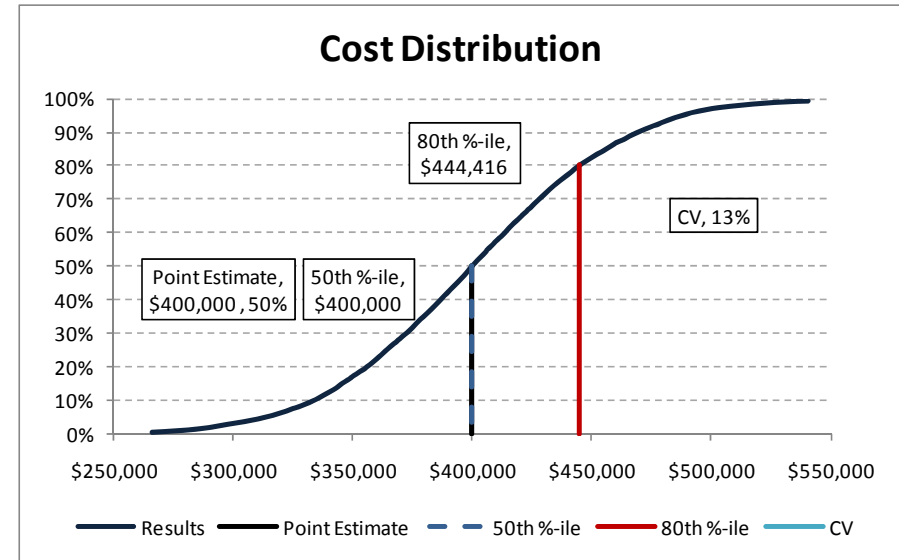
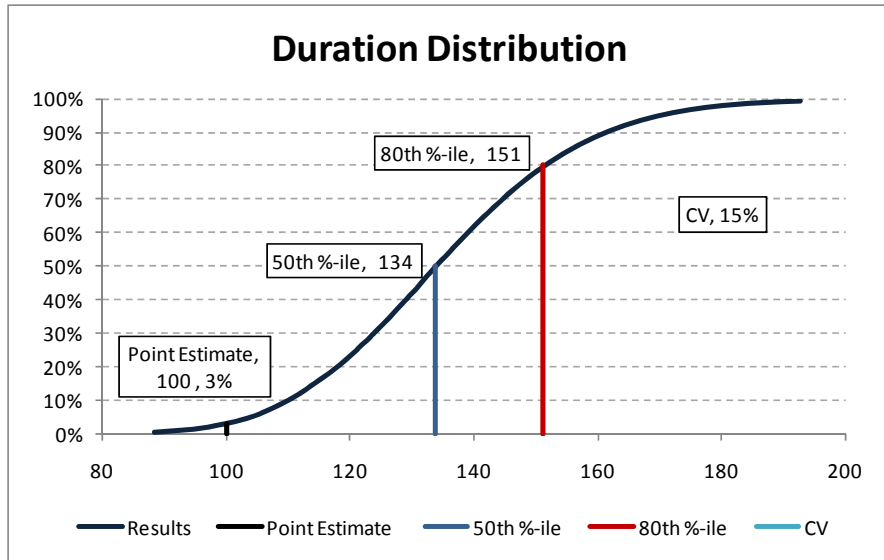
## Monte Carlo Iterations



- ▶ The satellite's finish date is the *maximum* of the finish dates of the 5 components
  - For most iterations (97%) there is one component that overruns
  - Probability of finishing on time is  $50\%^5 = 3\%$
- ▶ At the same time, cost is still driven by the individual, *symmetrically distributed* component durations
  - Thus we are actually finding the distribution of the *sum* of random variables
  - This is analogous to traditional cost estimating
- ▶ The next slide will show the results of 10,000 runs of this simulation

Iteration	Duration	Cost
1	131 days	\$441,240
2	149 days	\$342,832
3	116 days	\$355,368
4	112 days	\$300,560
5	151 days	\$413,392
6	127 days	\$439,088
7	159 days	\$468,152
8	168 days	\$395,536
9	127 days	\$395,712
10	157 days	\$436,688
<b>Average</b>	<b>141 days</b>	<b>\$402,597</b>

## Results



- ▶ The deterministic schedule is at 3% confidence with 34% schedule risk to the median completion date
  - 3% confidence =  $(50\% \text{ probability each component finishes on time})^{(5 \text{ components})}$
  - 34% represents 5<sup>th</sup> order statistic of 5 iid  $N(100, 30)$  random variables
  - Is this really realistic? Without correlation we may be overstating risk (more on this later)
- ▶ Point cost estimate is at 50% confidence with 0% cost growth to the median

## The Merge Bias and Order Statistics

- ▶ In cost risk analysis, results can be cross-checked using the central limit theorem
  - Method of Moments can (and should) be used to cross-check mean and standard deviation
- ▶ In schedule risk analysis there is no simple way to mathematically check the expected finish date of a complex schedule network
- ▶ Serial tasks are simple; since they can be summed the math is the same as traditional cost risk analysis
- ▶ Parallelism introduces a significant layer of complexity as the finish date of the parallel schedule network is constrained by the maximum of finish dates
  - Calculating the maximum of schedule distribution requires the use of order statistics
- ▶ Closed form equations for order statistics are very limited
  - For example, the distribution of the  $k^{\text{th}}$  statistic from *iid*, continuous random variables is:

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1} (1-F(x))^{n-k} f(x).$$

## Complex Schedule Networks

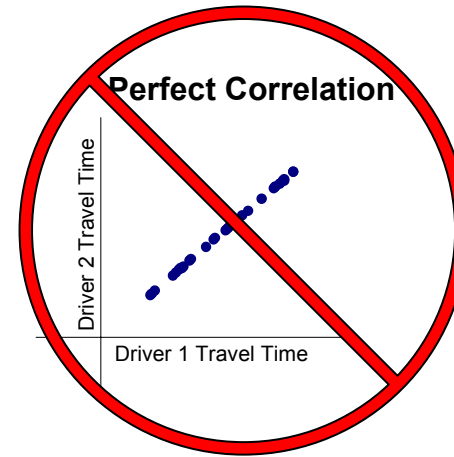
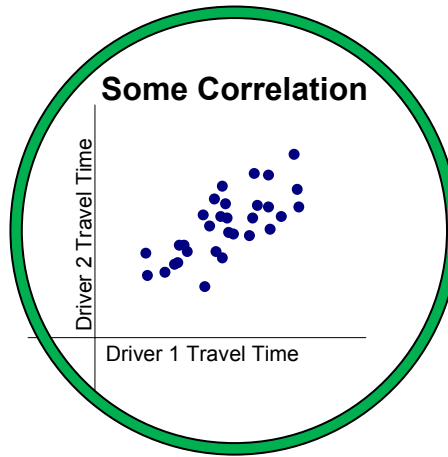
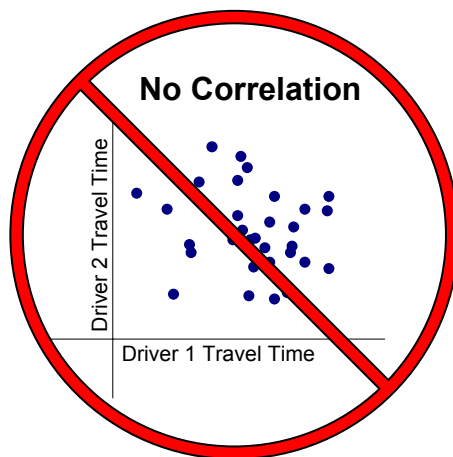
- ▶ Unfortunately, real-world schedules have far more than just parallel tasks to deal with:

Factors Influencing Schedule Risk	
Mix of Parallel/Serial Tasks	Types of Task Constraints (Start no Earlier Than, Must Start On, etc.)
Resource Availabilities	Types of Task Relationships (Start to Start, Start to Finish, Finish to Finish, etc)
Cross-links/dependencies	Schedule Margin, Schedule Slack, Schedule Reserves

- ▶ Due to the complexity of large schedules, Monte Carlo analysis is the only reasonable way to evaluate schedule risk using the build-up method
- ▶ It is important that risk analysts effectively communicate to PMs how the topology of their schedule will affect the results
  - Schedule parallelism, cross-links and high uncertainty factors will all cause the baseline plan to be at a low confidence level unless significant reserves and slack are in the schedule
  - Margin should be determined based on results from risk analysis
- ▶ Without correlation however, this effect (the “Merge Bias”), is likely to be overstated
  - Similarly the CV of the cost distribution will be understated
- ▶ To examine how correlation affects schedules, let’s return to our dinner example

## Correlation (cite CEBok)

- ▶ Although the two parties are driving separately, there are factors likely to affect both:
  - Traffic, difficulties in finding parking spaces and weather will affect both drivers similarly
- ▶ Similarly, each driver has factors that are independent
  - The risk of breaking down or getting in an accident will be independent between drivers
- ▶ Since there are similarities and differences between the drives, it is impossible to argue that travel times to dinner are uncorrelated or perfectly correlated
  - Rather, they must be somewhere in between





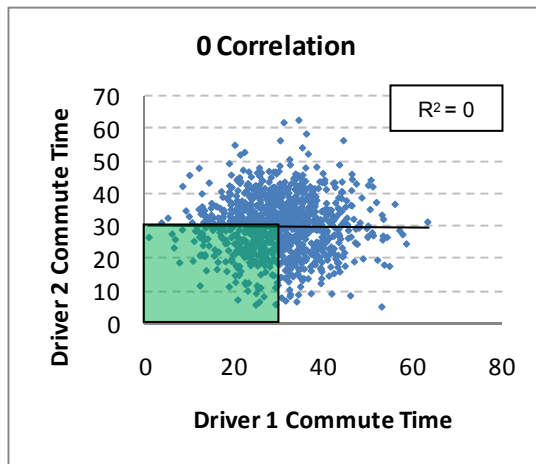
## Correlation: A Simple Example

- ▶ In the dinner example, with zero correlation, the probability of both parties arriving on time was 25%
- ▶ As correlation rises, the probability of both parties arriving on time increases
  - This is because extreme disparities (one driver arriving very early, one very late) are reduced
  - Similarly, the CV of the arrival time increases

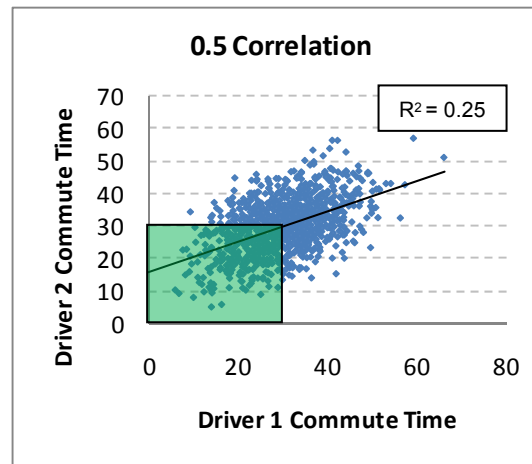
### Assumptions

5000 Monte Carlo  
Runs. Commute  
times distributed as  
 $N(30 \text{ minutes}, 9 \text{ minutes})$

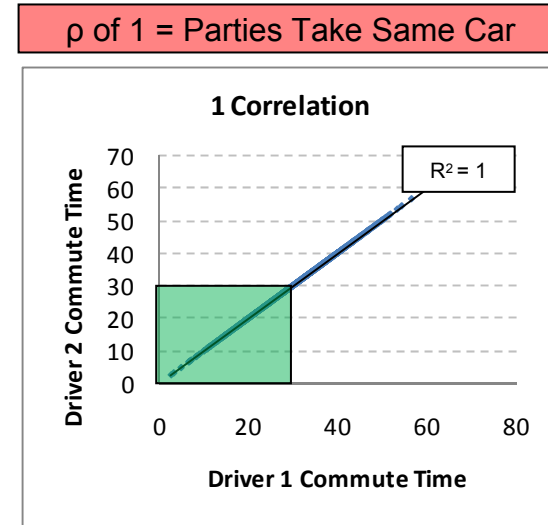
Green box denotes  
both parties arriving  
on time for dinner



**0 Correlation**



**0.5 Correlation**



**1 Correlation**

$\rho$  of 1 = Parties Take Same Car

**Probability of On Time Arrival**

**25%**

**33%**

**50%**

**Average Arrival Time**

**35 Minutes**

**34 Minutes**

**30 Minutes**

**CV of Arrival Time**

**21%**

**24%**

**30%**

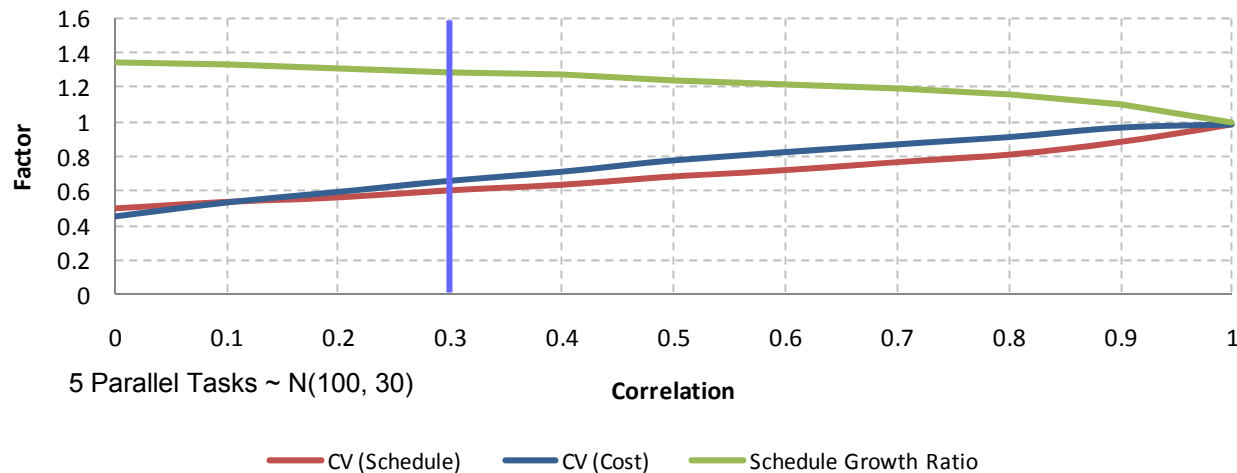
## Today's Schedule Risk Analysis Standards

- ▶ Most schedule risk analysts ignore correlation between schedule tasks using the rationale that “it’s covered in the schedule logic”
  - The correlation inferred by that statement only addresses schedule slips
  - I.e. If development takes longer, integration will be *delayed*
- ▶ Injecting correlation between schedule task duration distributions accounts for this omission
  - I.e. If development takes longer, integration will be delayed *and is likely to take longer*
  - This is because difficulties (or, conversely, successes) in overcoming the problem or schedule underestimations are likely to be systemic to a program
- ▶ Another way to induce schedule correlation is the “Risk Factor Method”<sup>5</sup>
  - The risk factor method correlates schedule durations by applying the same “Risk Factor” distributions across schedule tasks
  - May not always be appropriate
    - Certainly durations for subsystem production and integration & test are correlated, but do these activities have the same distribution?

**No Cost or Schedule Risk Analysis is Valid If It Has Not Addressed Correlation**

## Effects of Correlation on Satellite Example

CV & Schedule Growth Ratios vs. Correlation



### Graph Explanation

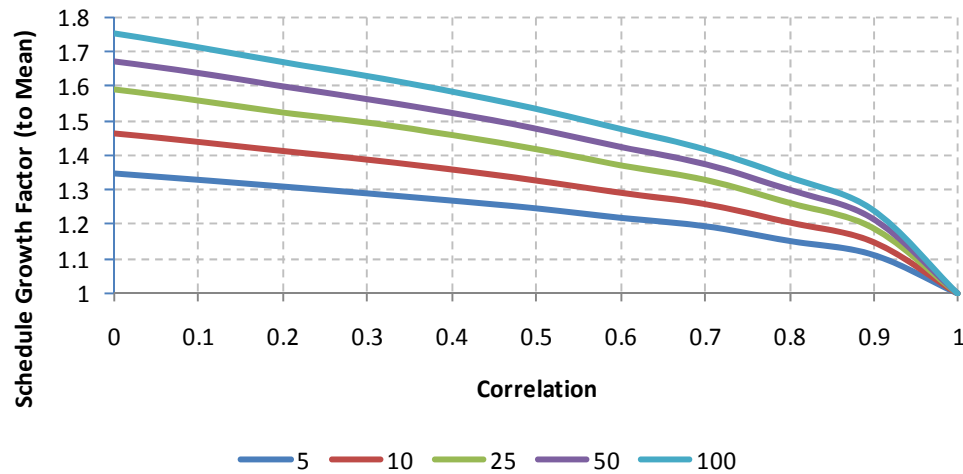
With a correlation of .3 (blue line) between schedule tasks:

- CV of the schedule duration distribution is 18% (or **60%** of the input task duration CVs of 30%)
- CV of the cost distribution is 19% (or **65%** of the input task duration CVs of 30%)
- The average schedule duration is 129 days (or 129% of the baseline schedule duration of 100 days)

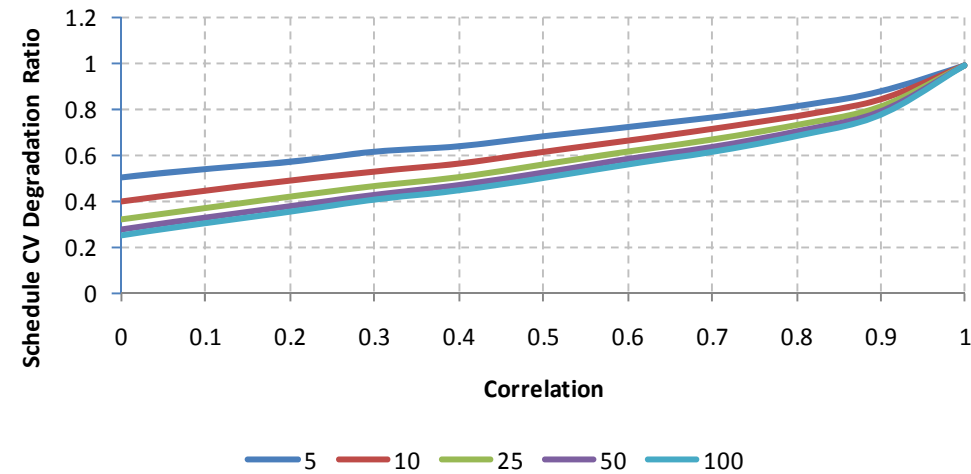
- ▶ Above graph represents effects of injecting correlation between component durations from the spacecraft example
  - As correlation increases, CV of both cost and schedule distributions increases and expected finish date of the schedule improves
- ▶ If there are similar factors affecting schedule tasks it is practically inconceivable that they could be uncorrelated
  - Next slide will examine schedule risk as the number of parallel tasks varies

## Effects of Correlation (Generalized)

Average Schedule Growth vs # of Parallel Tasks & Correlation






Schedule CV Degradation vs # of Parallel Tasks & Correlation



- ▶ The above graphs show how correlation affects the risk adjusted schedule as the number of parallel tasks varies (still assuming 30% CV for each task)
  - Less correlation and more parallel tasks equal more schedule risk
- ▶ Next slide will present guidelines for correlation between schedule tasks

## Correlation Recommendations

- ▶ When possible, data driven approaches should always be used to determine correlation between schedule task durations
  - For example: historical schedule growth between satellite subsystems
- ▶ If a data driven approach is not feasible and the schedule is of a reasonable size, the following guidelines should be used

Correlation (including example basis for selection)*	$\rho$	Pic
<b>Weak</b> (different personnel working different component)	<b>0.25</b>	
<b>Medium</b> (same personnel working different component or different personnel working same component)	<b>0.50</b>	
<b>Strong</b> (same personnel working on the same component)	<b>0.75</b>	

- ▶ When data is not available, or it is infeasible to directly assess correlation, it is recommended that a correlation of 0.3 be injected between schedule distributions
  1. This correlation is industry standard for cost risk analysis<sup>4</sup> to prevent to prevent sqrt(n) effect
    - Mitigates ~30% of CV degradation
  2. Acts as the knee in the curve for schedule risk: Mitigates the same % of schedule CV degradation for all serial networks (~30%), slightly less for all parallel networks (~15%)
    - Simulation must be run to determine exact effect, likely to be 15% < x < 30%

\*It is extraordinarily rare for tasks on the same project to be completely uncorrelated ( $\rho = 0$ ). Similarly, if two tasks are perfectly correlated ( $\rho = 1$ ) they should be functionally linked

## Conclusion & Recommendations

- ▶ As a community, risk analysts need to understand, and be able to communicate, the two JCL Paradoxes
  1. Schedule parallelism, a high number of cross-links between activities and large uncertainties on task durations will lead to a high risk adjusted finish date
    - ...and low confidence in the deterministic schedule
  2. Correlation is a significant driver of schedule risk and must be accounted for in all schedule risk analysis
    - Ignoring correlation leads to an overstatement of schedule risk and an underestimation of both the cost and schedule CVs
- ▶ It is hoped that the correlation guidelines provided in this presentation can begin the conversation regarding how correlation should be handled in both JCL and schedule risk analysis

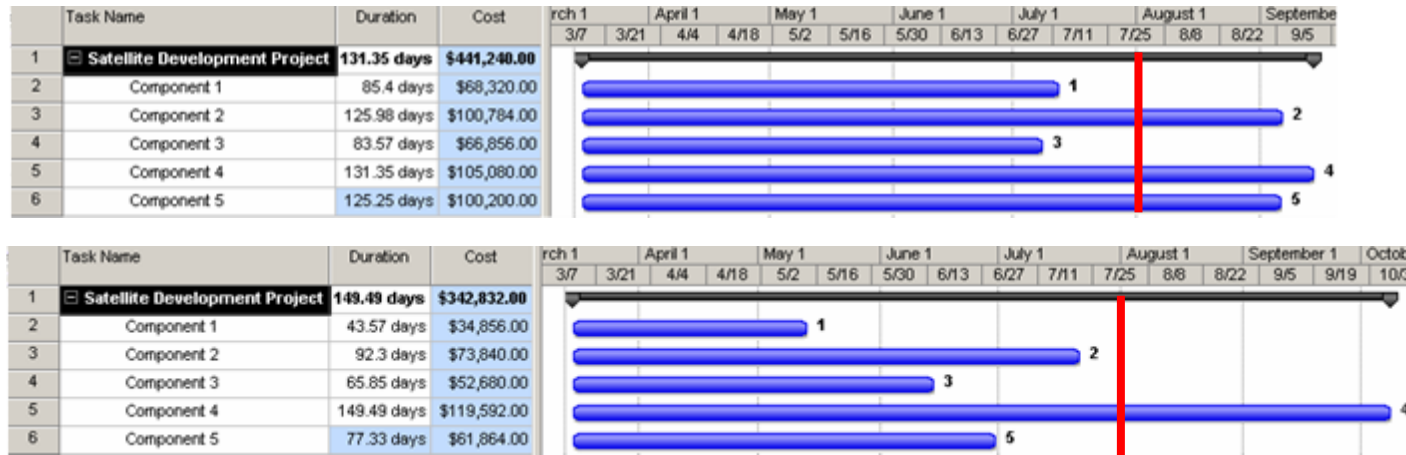
## Works Cited

- <sup>1</sup> Druker, Eric. “Emerging Practice: Joint Cost & Schedule Risk Analysis.” 2009 Institute for Operations Research and the Management Sciences Annual Meeting. San Diego, CA. October 2009
- <sup>2</sup> Coleman, Richard. Summerville, Jessica. “The Relationship Between Cost Growth and Schedule Growth.” *Acquisition Review Quarterly*. Spring 2003
- <sup>3</sup> Hulett, David. *Practical Schedule Risk Analysis*. Burlington: Gower Publishing Company, 2009.
- <sup>4</sup> Book, Stephen. “Why Correlation Matters in Cost Estimating.” 32<sup>nd</sup> Annual Department of Defense Cost Analysis Symposium. Williamsburg, VA. February 1999.

# Backup



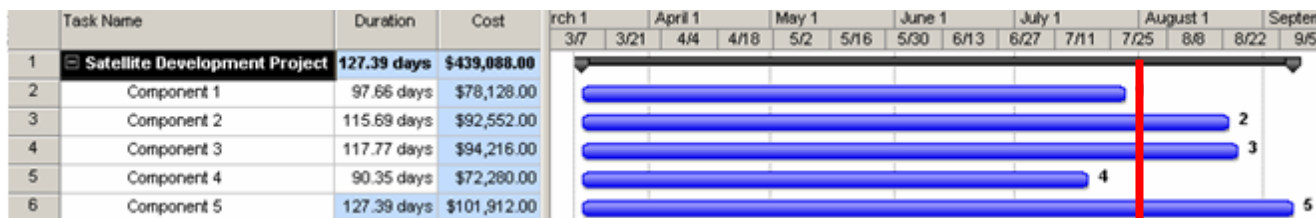
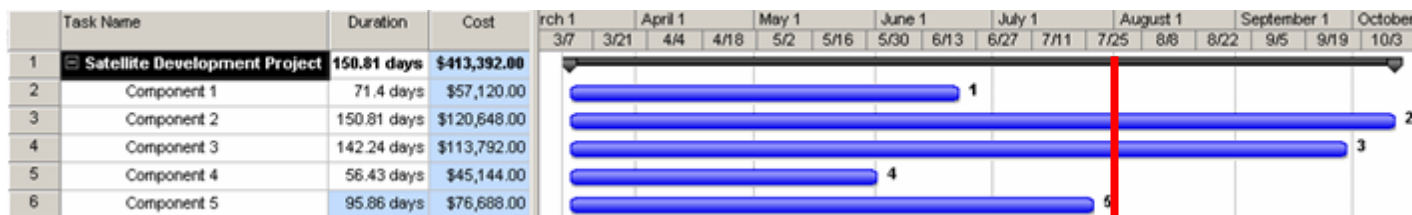
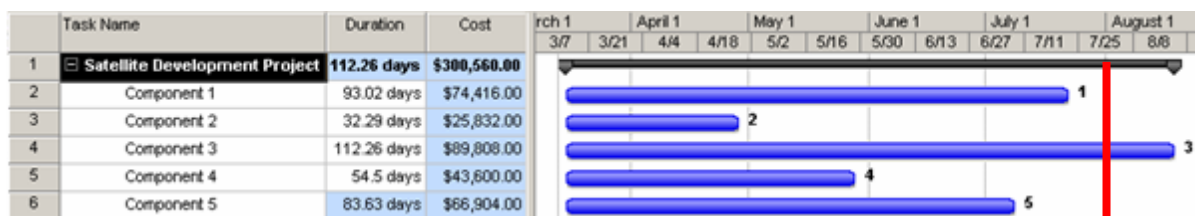
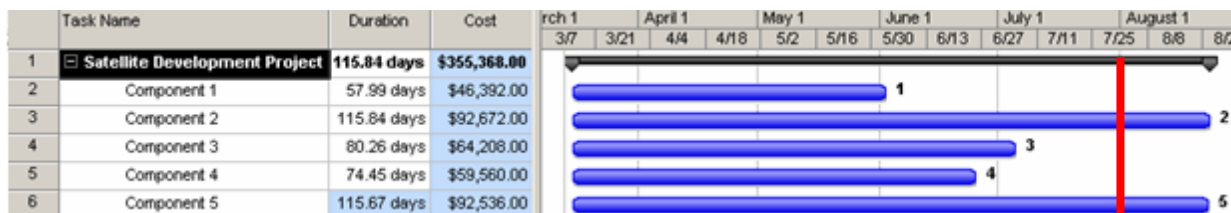
## Monte Carlo Iterations (separated)



- ▶ For most iterations (97%) there is one component that overruns
  - Thus we are actually finding the distribution of the *maximum* of random variables
  - Even with symmetric uncertainty, this will always be greater than the critical path length
- ▶ At the same time, cost is still driven by the individual, *symmetrically distributed* component durations
  - Thus we are actually finding the distribution of the *sum* of random variables

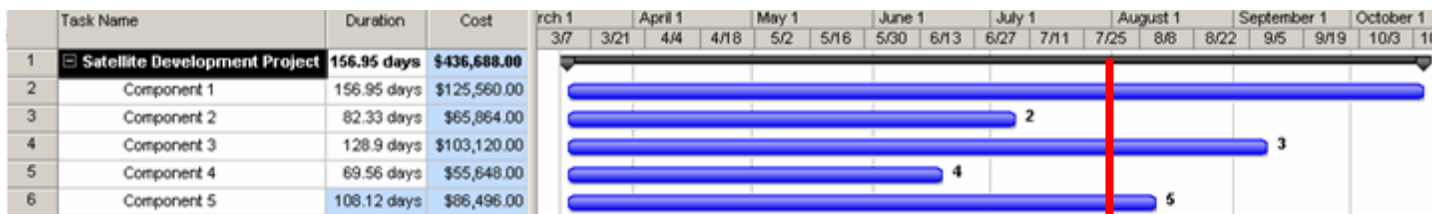
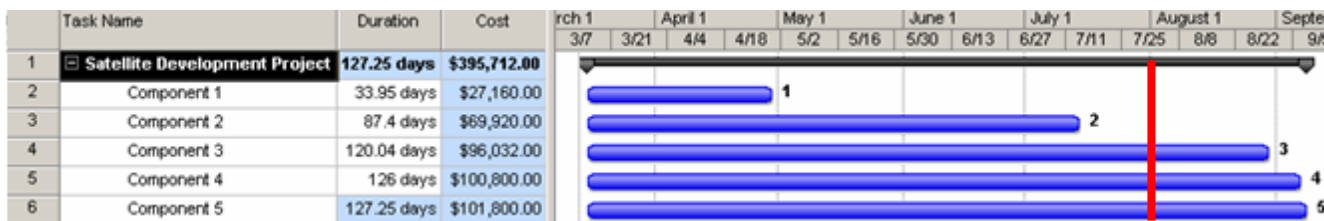
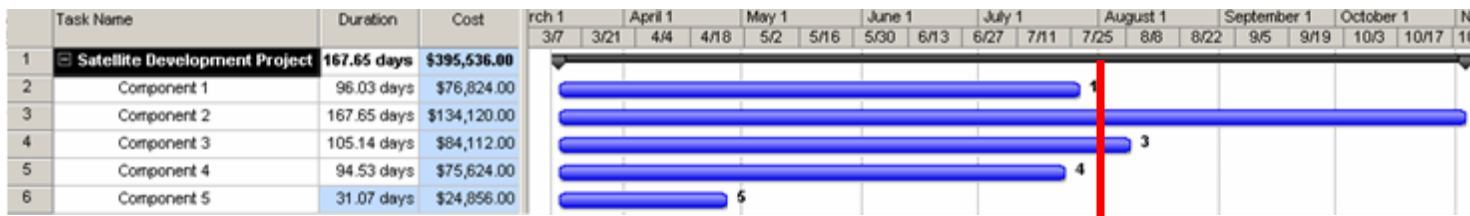
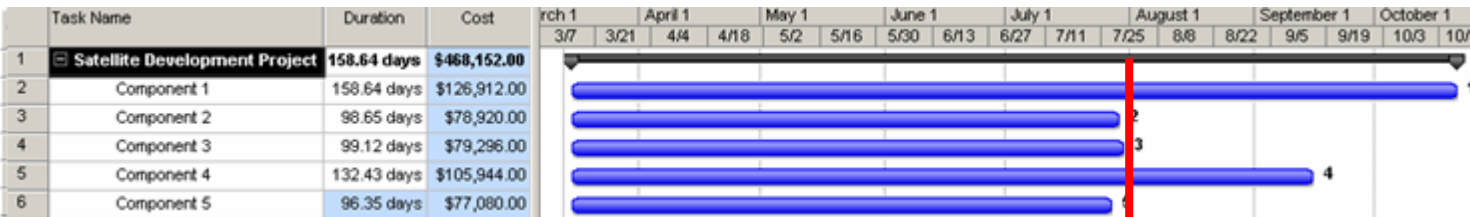
Iteration	Duration	Cost
1	131 days	\$441,240
2	149 days	\$342,832
3	116 days	\$355,368
4	112 days	\$300,560
5	151 days	\$413,392
6	127 days	\$439,088
7	159 days	\$468,152
8	168 days	\$395,536
9	127 days	\$395,712
10	157 days	\$436,688
<b>Average</b>	<b>141 days</b>	<b>\$402,597</b>

## Monte Carlo Iterations (separated)



Iteration	Duration	Cost
1	131 days	\$441,240
2	149 days	\$342,832
3	116 days	\$355,368
4	112 days	\$300,560
5	151 days	\$413,392
6	127 days	\$439,088
7	159 days	\$468,152
8	168 days	\$395,536
9	127 days	\$395,712
10	157 days	\$436,688
Average	141 days	\$402,597

## Monte Carlo Iterations (separated)



Iteration	Duration	Cost
1	131 days	\$441,240
2	149 days	\$342,832
3	116 days	\$355,368
4	112 days	\$300,560
5	151 days	\$413,392
6	127 days	\$439,088
7	159 days	\$468,152
8	168 days	\$395,536
9	127 days	\$395,712
10	157 days	\$436,688
Average	141 days	\$402,597