ENCE 667
Final Report

Project Performance Measurement Techniques in the Development and Implementation of a Data Warehouse

Presented By:
Nizar A. Al Rihani
Michael J. Casey

Overview

- Development of a Data Warehouse for major telecommunications company.
- A professional consultant was hired to develop the Warehouse.
- Two phases were considered.
  - Recurring task: Monthly downloading of data.
  - Development and Implementation

Project Performance Methods Used

- **Goal Programming.** Develop an optimal strategy for downloading data to DW on a monthly basis.
  - Software: LINDO
- **Monte Carlo Simulation.** Assess variability in schedule and cost due to uncertainty.
  - Software: MS Project, @Risk

Phase II – 6 mo. contract

- **5th – 25th of each month:** Receiving and downloading data files into the Warehouse.
- **25th of the current month – 5th of the next month:** Data processing and Warehouse development.

Goal Programming

- **Goal 1:** Downloading data files should be completed in 20 days. Otherwise, there is a penalty of $850/day.
  Goal 2: Data processing and evaluation should be completed in 10 days. Otherwise, there is a penalty of $1000/day.
- The Warehouse receives data files from two main offices, located in New York and Atlanta.
- A linear program is formulated to determine the optimal number of files to be received from each office.

Problem Formulation

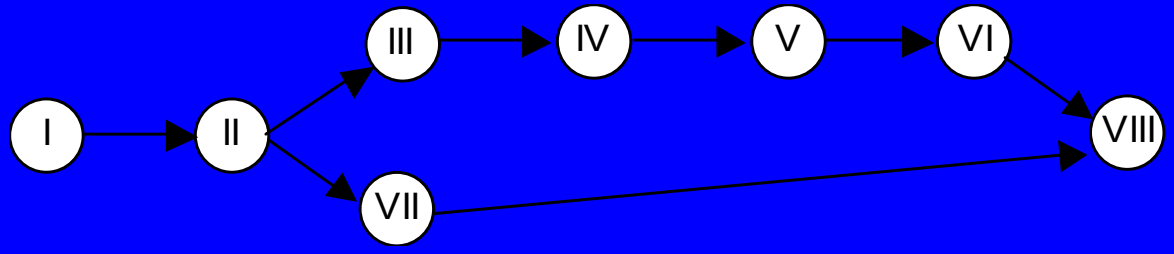
- **Input:**
  MIN 850S1P+1000S2P
  SUBJECT TO
  0.25X1+0.1667X2+S1N-S1P=20
  0.1667X1+0.0833X2+S2N-S2P=10
  455X1+270X2=22680
  END
  GIN X1
  GIN X2
- **Result:** No penalties. 84 files received from Atlanta office, no files received from the New York office.

- ***Input:***
  MIN 850S1P+1000S2P
  SUBJECT TO
  0.25X1+0.1667X2+S1N-S1P=20
  0.1667X1+0.0833X2+S2N-S2P=10
  455X1+270X2=22680
  X1>=1
  X2>=1
  END
- ***Result:*** No penalties. 49 data files received from New York office, one file received from Atlanta office.

Phase I

- ***Phase 1:***
  - ***I. Determine the DW requirements.***
  - ***II. Design and develop the system code.***
  - ***III. Test the system performance.***
  - ***IV. Redesign based on testing.***
  - ***V. Conduct user testing.***
  - ***VI. Integrate the DW with existing systems.***
  - ***VII. Document development & implementation process.***
  - ***VIII. Plan for future work.***

- ***AON Network:***

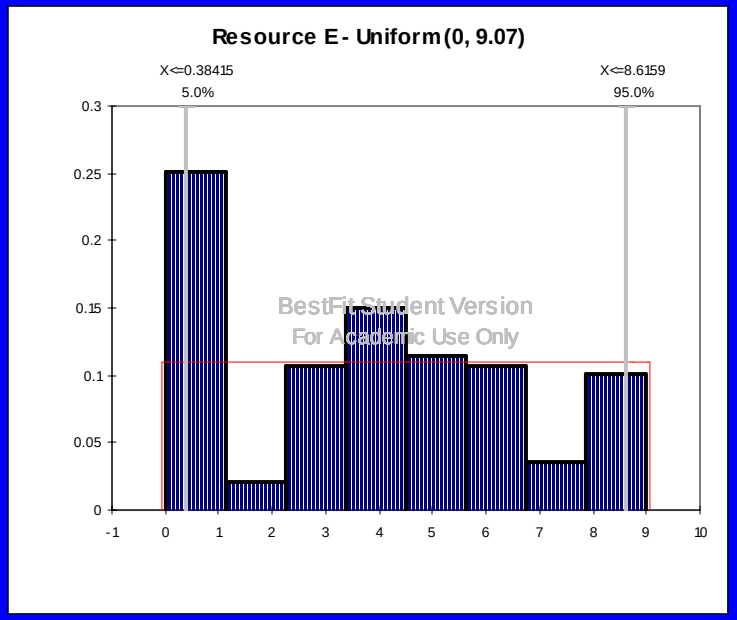


Monte Carlo Simulation

- **Uncertainties:** task durations in coding, DW development and resource utilization.
- Apply Simulation to provide output effect on project duration and cost.
- **Data from real project:** Tasks, monthly time sheet data.

Modeling Uncertainty

- **BestFit** used to fit distributions to employee utilization data.

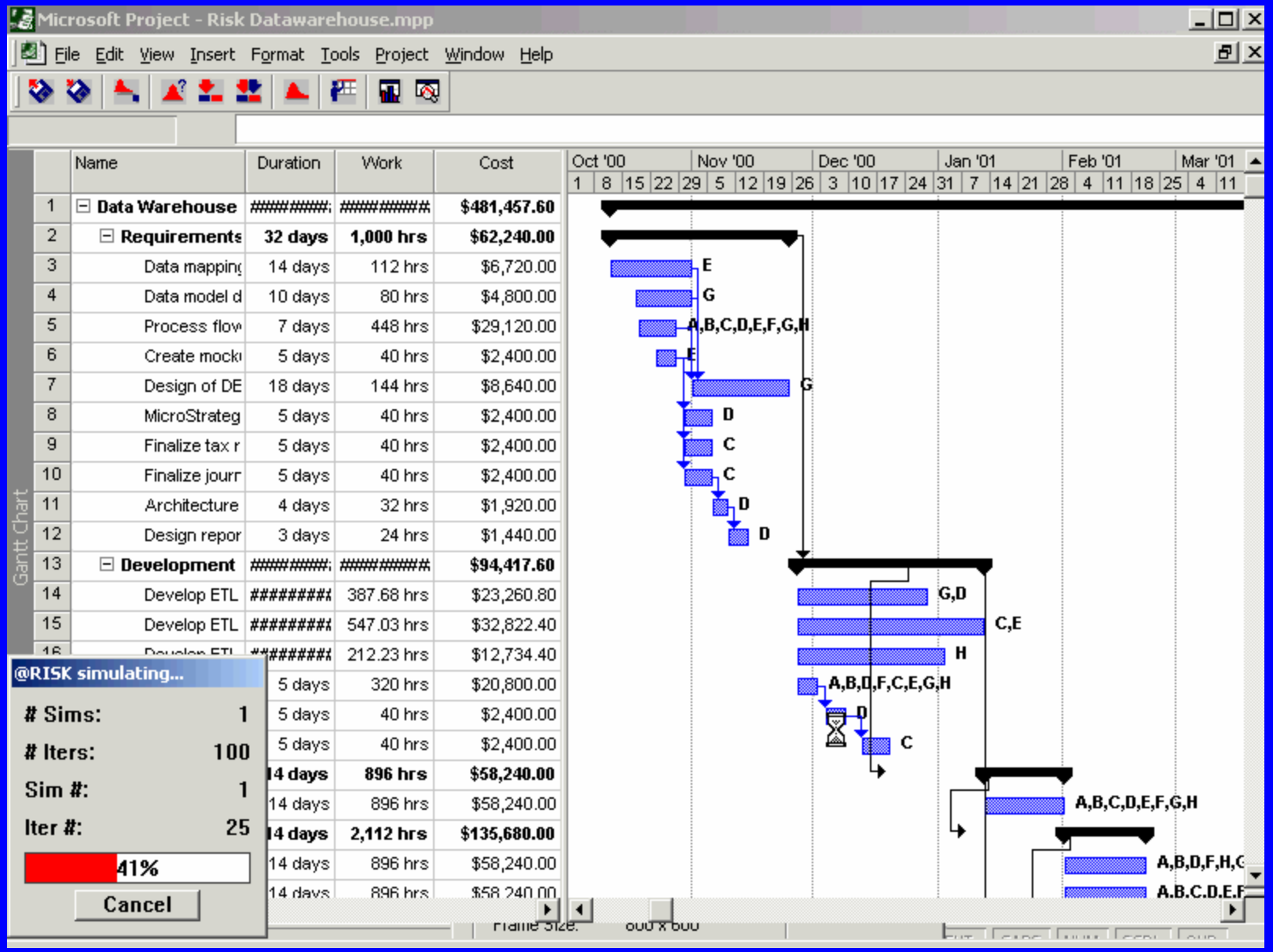


- **@Risk Functions**
  - **Duration (Uniform)**

ID	Name	Duration	Cost	@RISK: Functions
1	**Data Warehouse**	**98.75 days**	**$403,440.00**	**Duration=RiskOUTPUT()**
2	**Requirements Analysis**	**32 days**	**$62,240.00**	
13	**Development DB and Loading**	**18 days**	**$67,360.00**	
14	Develop ETL code and processes for loading the A	18 days	$17,280.00	Duration=RiskUNIFORM(18,36)
15	Develop ETL code and processes for loading the C	18 days	$17,280.00	Duration=RiskUNIFORM(18,36)
16	Develop ETL code and processes for loading the C	15 days	$7,200.00	Duration=RiskUNIFORM(15,30)
17	Consolidate ETL code processes.	5 days	$20,800.00	Duration=RiskUNIFORM(5,10)
18	Add 5 new MicroStrategy reports/change existing	5 days	$2,400.00	Duration=RiskUNIFORM(5,10)
19	Develop new A/R Journal Export module and Peop	5 days	$2,400.00	Duration=RiskUNIFORM(5,10)

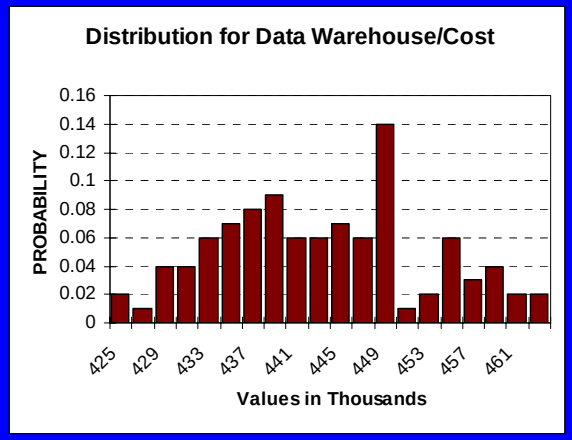
  - **Utilization**

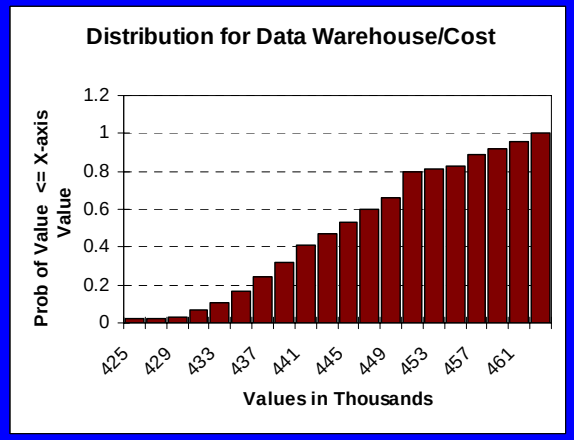
ID	Resource Name	@RISK: Functions
1	A	Max. Units=RiskEXPON(10)
2	B	Max. Units=RiskUNIFORM(0,250)
3	C	Max. Units=RiskTRIANG(0,137.5,215.4)
4	D	Max. Units=RiskTRIANG(0,137.5,196)
5	E	Max. Units=RiskUNIFORM(0,176.3)
6	F	Max. Units=RiskTRIANG(0,0,159)
7	G	Max. Units=RiskUNIFORM(0,176.3)
8	H	Max. Units=RiskTRIANG(0,0,193.8)

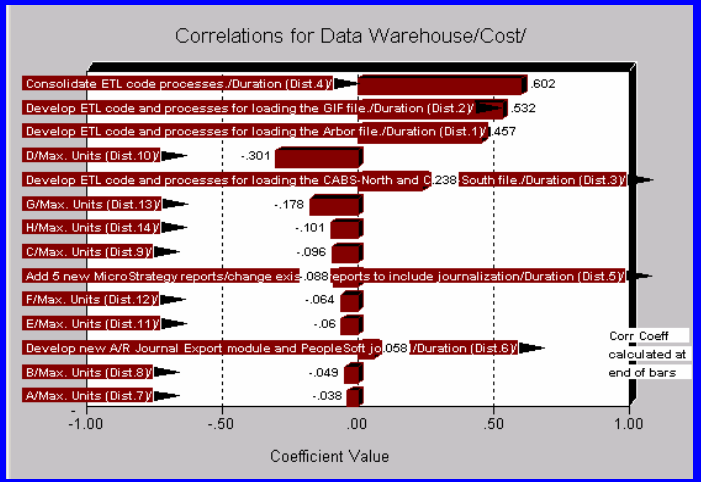

Microsoft Project - Risk Datawarehouse.mpp
File Edit View Insert Format Tools Project Window Help
Name
Duration
Work
Cost
1 Data Warehouse ########## ########## $481,457.60
2 Requirements 32 days 1,000 hrs $62,240.00
3 Data mapping 14 days 112 hrs $6,720.00
4 Data model d 10 days 80 hrs $4,800.00
5 Process flow 7 days 448 hrs $29,120.00
6 Create mock 5 days 40 hrs $2,400.00
7 Design of DE 18 days 144 hrs $8,640.00
8 MicroStrateg 5 days 40 hrs $2,400.00
9 Finalize tax r 5 days 40 hrs $2,400.00
10 Finalize journ 5 days 40 hrs $2,400.00
11 Architecture 4 days 32 hrs $1,920.00
12 Design repor 3 days 24 hrs $1,440.00
13 Development ########## ########## $94,417.60
14 Develop ETL ######### 387.68 hrs $23,260.80
15 Develop ETL ######### 547.03 hrs $32,822.40
212.23 hrs $12,734.40
5 days 320 hrs $20,800.00
5 days 40 hrs $2,400.00
5 days 40 hrs $2,400.00
14 days 896 hrs $58,240.00
14 days 896 hrs $58,240.00
14 days 2,112 hrs $135,680.00
14 days 896 hrs $58,240.00
14 days 896 hrs $58,240.00
Gantt Chart
Oct '00 Nov '00 Dec '00 Jan '01 Feb '01 Mar '01
1 8 15 22 29 5 12 19 26 3 10 17 24 31 7 14 21 28 4 11 18 25 4 11
E
G
A,B,C,D,E,F,G,H
E
G
D
C
C
D
D
G,D
C,E
H
A,B,D,F,C,E,G,H
D
C
A,B,C,D,E,F,G,H
A,B,D,F,H,G
A,B,C,D,E,F
@RISK simulating...
Sims: 1
Iters: 100
Sim #: 1
Iter #: 25
41%
Cancel
Frame Size: 600 x 600

Results

- Cost Distribution data:
- Sensitivity Analysis (tornado diagram)







Conclusions

- Development of a Data Warehouse is a necessary step to optimize profits.
- Monte Carlo simulation is an effective method to measure changes in cost and schedule.
- “LINDO” and “@Risk” are simple, effective and user-friendly software.
- Optimal number and exact locations of data centers should be determined to minimize the cost of data transfer.