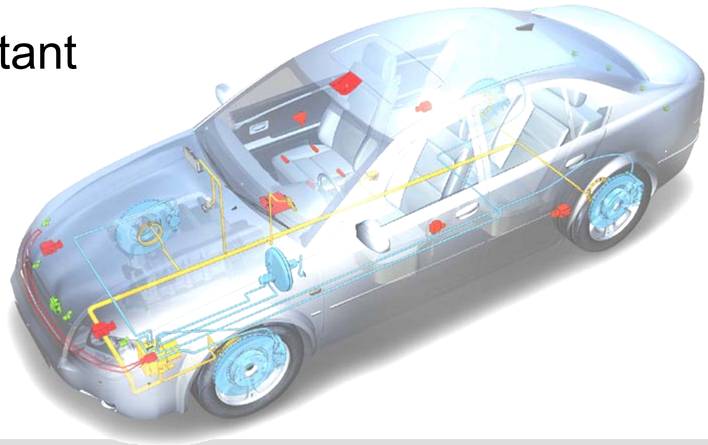
A Bayesian Network for Predicting Defect Correction Costs

Thomas Schulz
Engineering Systems Radar

BOSCH

Background & Motivation

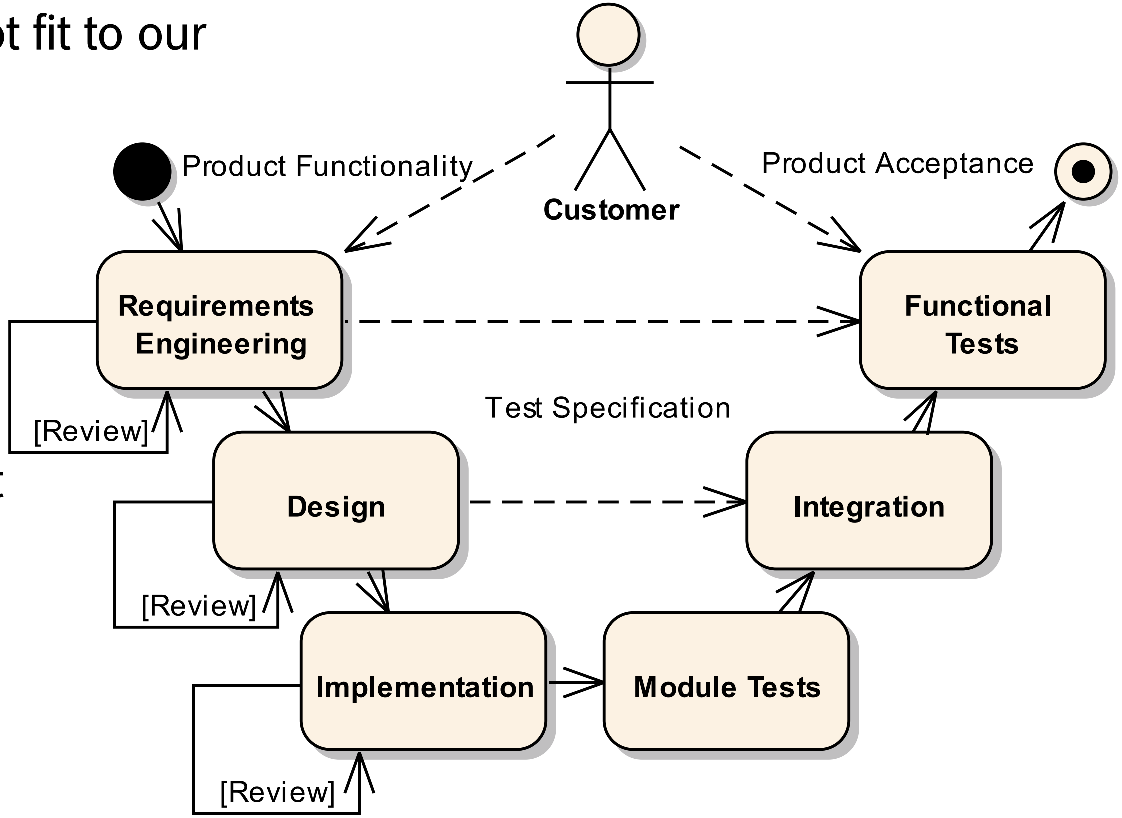
- Today's readmission rate for cars is ~ 70 000 000 cars per year
- Average of 20 to 70 embedded systems per car
- 50% development effort spend on software engineering
- Every 2nd car recall caused by software problems
- Software defect prediction plays an important role in customer satisfaction and overall cost reduction



BOSCH

Software Process Model (SPM) 1/2

- External models did not fit to our processes and data
- Created company specific model (SPM)
- Change model as a Dynamic Bayesian Net
- Calibration based on internal data

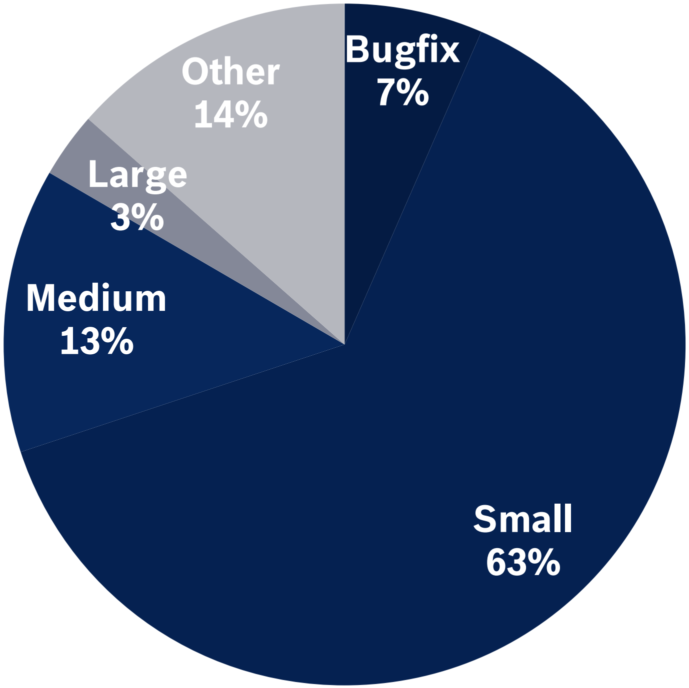


Dataset

- 259 data points, each reflects a single Change of
  - Change & Defect Management system
  - 17 defect correction (Bugfix) entries
  - 242 development entries

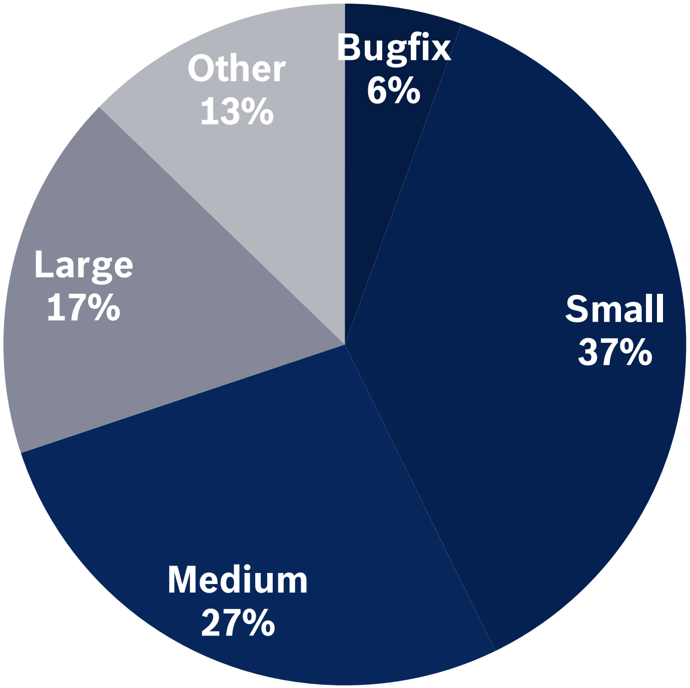
- Focus on the distribution of tasks
  - ~14 Changes per defect

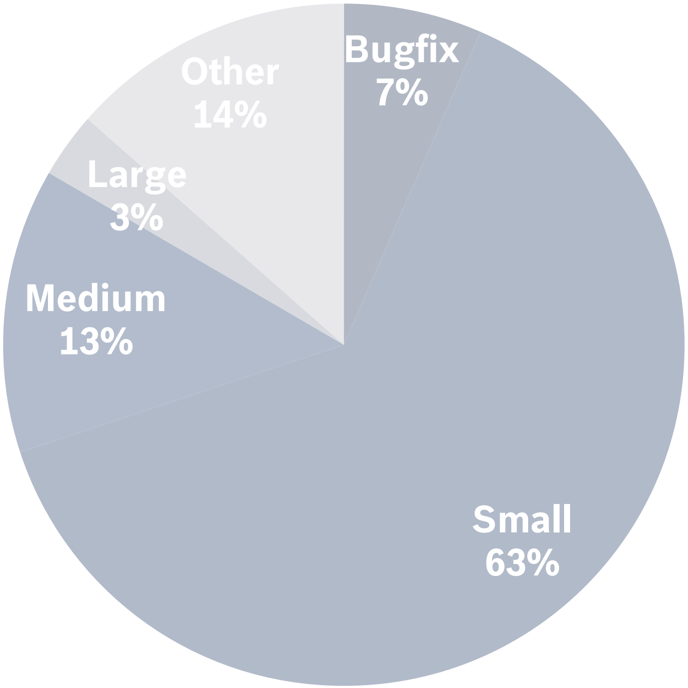
BUT: What does a single defect mean?



Dataset (revised)

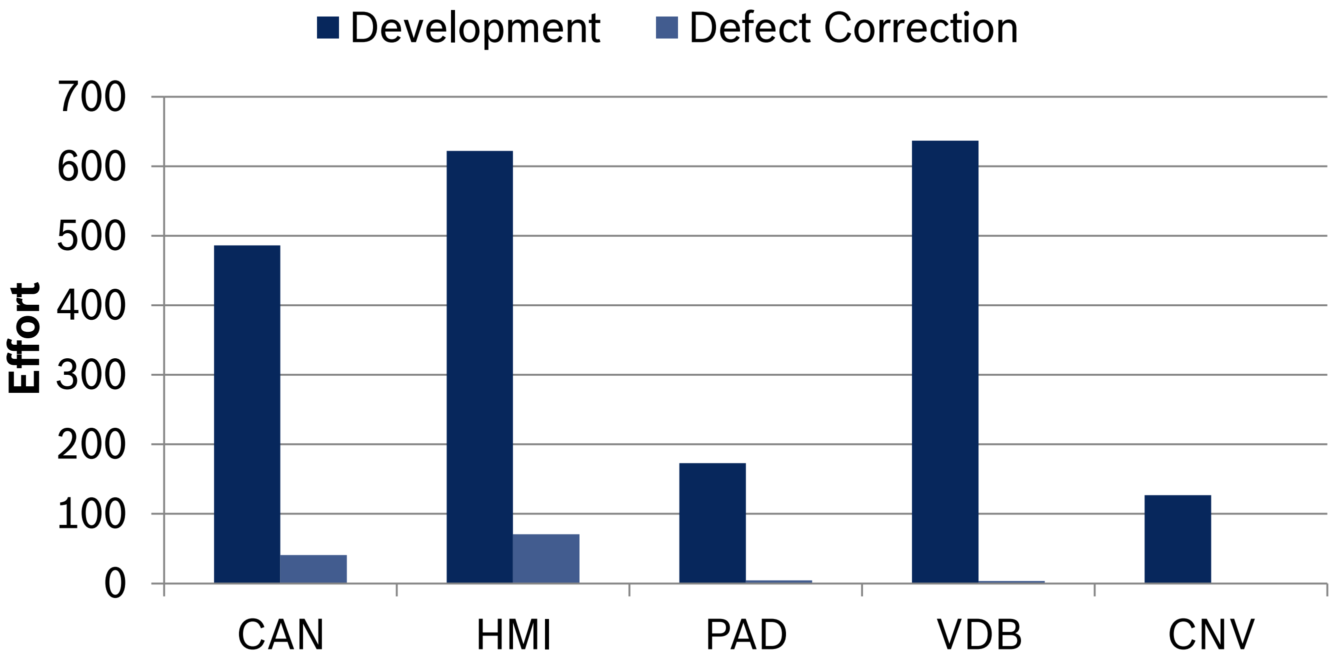
➔ Focus on effort for development and defect correction





Defect Cost Factor (DCF)

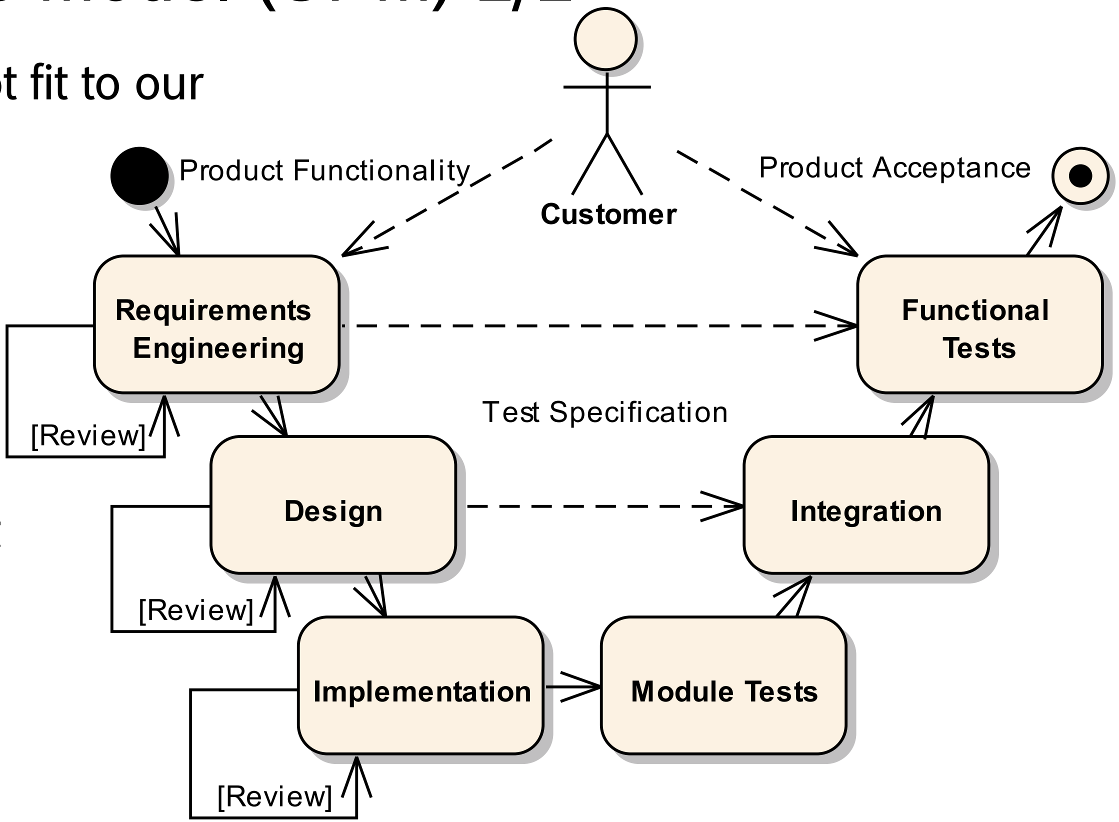
➔ DCF = Defect Correction Effort / Development Effort



DCF	0.0840	0.1137	0.0260	0.0053	0.0063

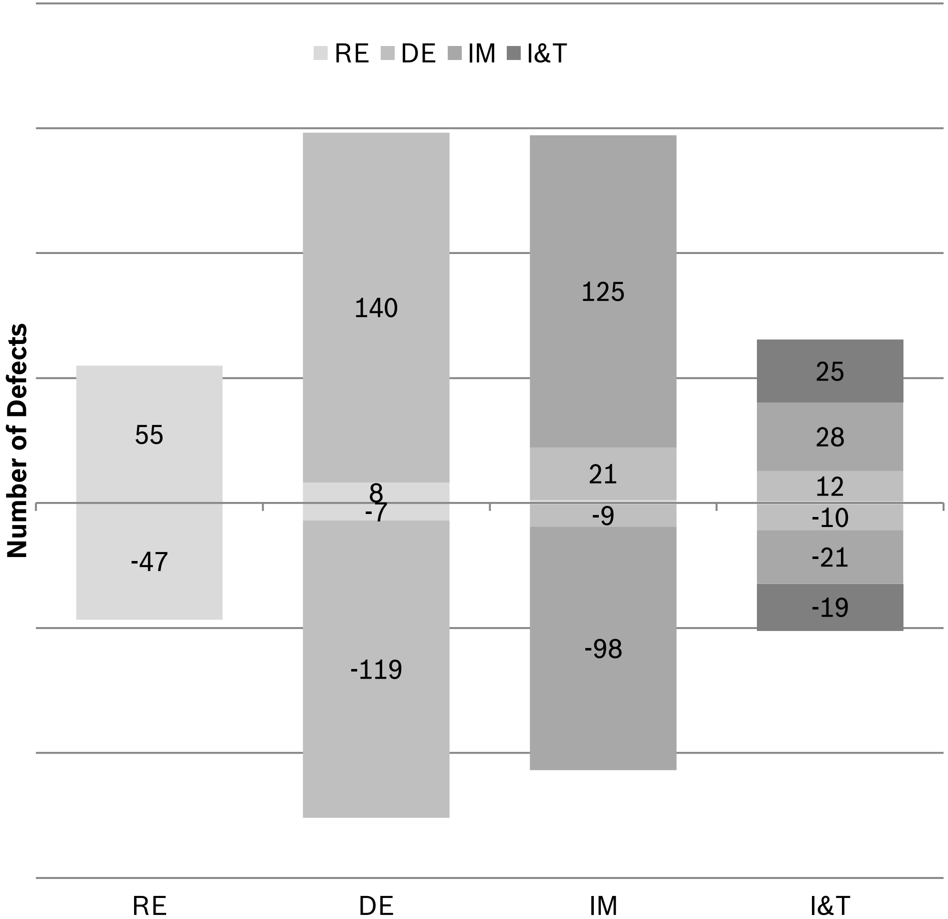
Software Process Model (SPM) 2/2

- External models did not fit to our processes and data
- Created company specific model (SPM)
- Change model as a Dynamic Bayesian Net
- Calibration based on internal data



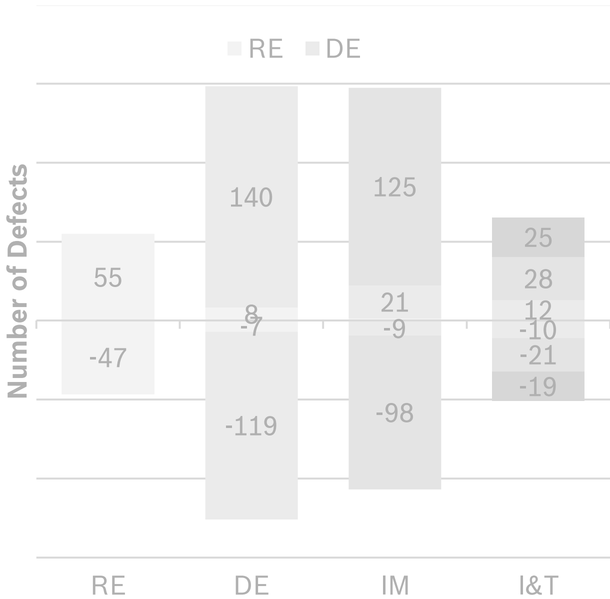
BUT: Optimization potential was too low!

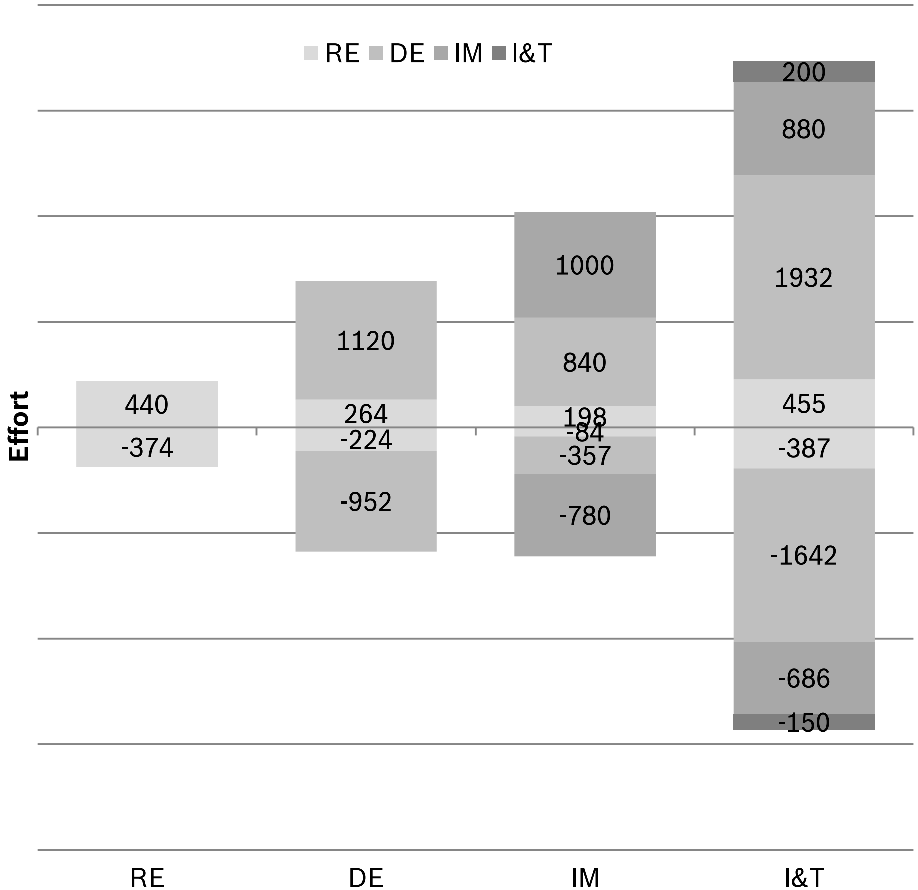
Defect Flow Model



- Defect Flow Model
  - Focus on number of defects
  - Over development phases
- Local optimization
  - Phase specific
- Focus on overall process missing
  - What happens when defects shift over phases?

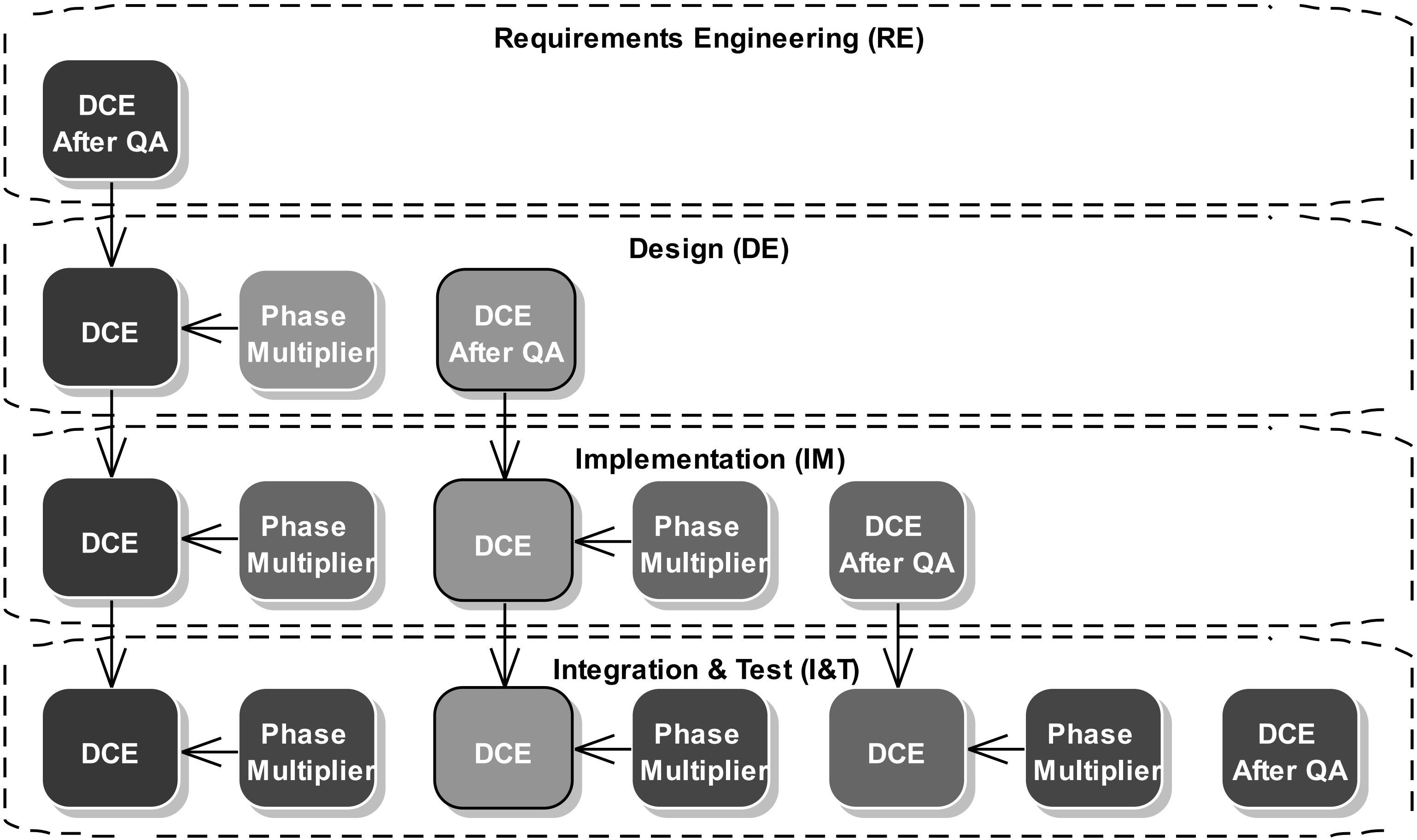
Defect Cost Flow Model 1/2





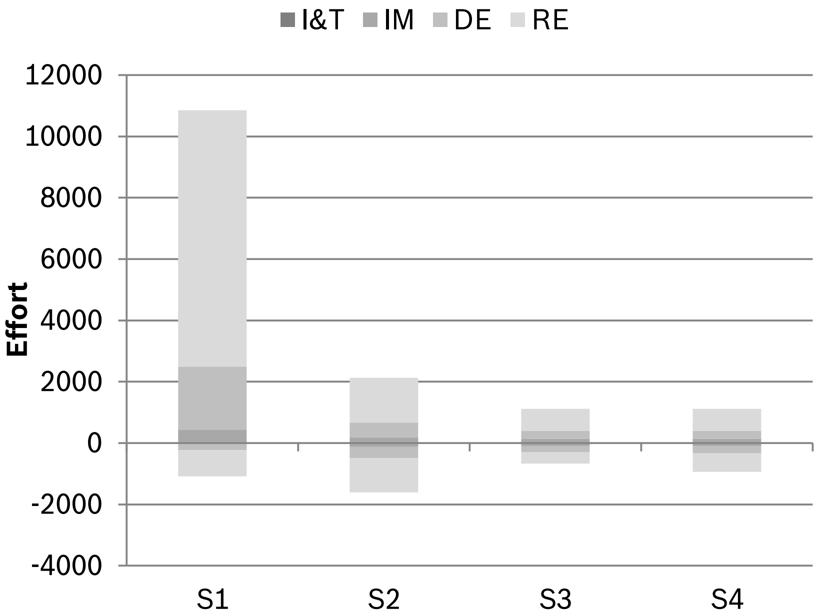
➔ Introducing
- Defect Correction Effort
- Phase Multiplier

Defect Cost Flow Model 2/2



Scenario Simulation

- S1 is at low QA activities, the worst case scenario
- S2 uses a high amount of QA effort typically used
- S3 has very high QA activities focusing on RE and DE
- S4 uses very high QA activities on all development phases



Result	S1	S2	S3	S4
Development effort	1000	1000	1000	1000
QA effort	50	200	320	400
DCE (part of QA)	1224	2343	1328	1608
Residual DCE	9771	528	457	177
Overall	**12045**	**4071**	**3105**	**3185**

Conclusion

- ➔ Defect Cost Flow model enables
  - Process wide optimization
  - Focus on defect correction effort
  - Incorporating phase specific effort multipliers
  - Results in focus on early development phases
- ➔ Increase product quality and customer satisfaction
- ➔ Bayesian Networks can support decision makers
- ➔ Ideal method to monitor process improvement and learn from it

Thank you!

A BN for Predicting Defect Correction Costs