A cost-benefit framework for making architectural decisions in a business context

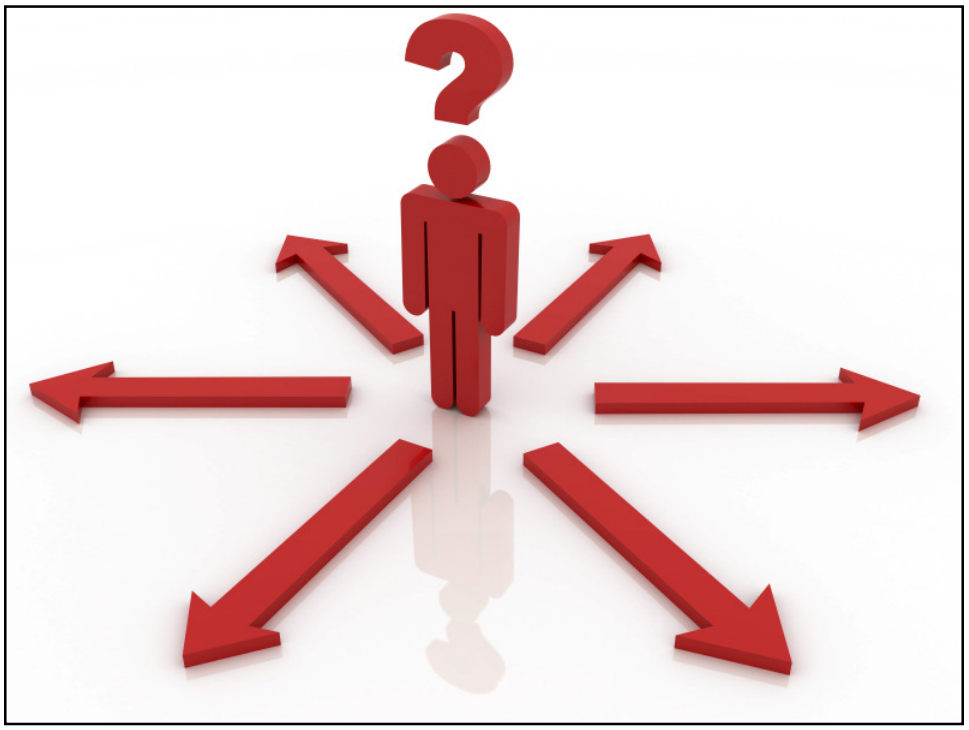
Jeromy Carriere
Chief Architect, VistaPrint
Rick Kazman and Ipek Ozkaya
Software Engineering Institute

Background

- VistaPrint
- A belief in the promise of early-lifecycle cost-benefit analysis
  - Software is seldom (never?) developed from scratch
  - Ad-hoc or code-level cost-benefit is state-of-the-practice

Motivation

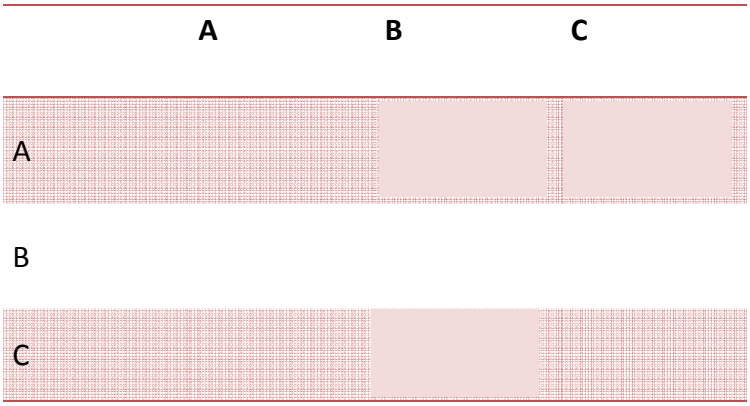
- Easy to place a value on revenue-generating projects
- Not easy to do the same for architectural transformations or refactoring
  - Risk reduction
  - Developer productivity
- We want architecture-based cost/benefit judgment



Approach

- Evaluate a completed project using
  - Tickets that document the work
  - Before and after codebase
  - Dependency structure matrix → coupling
- Train a model that can predict benefit based on a stream of classified changes

The Dependency Structure Matrix

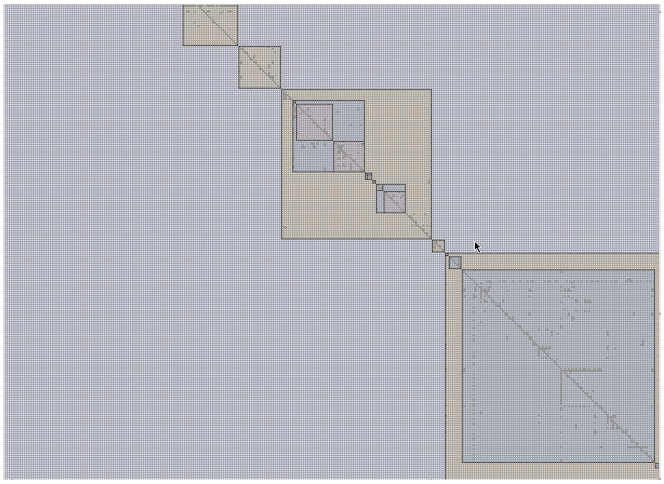


The Dependency Structure Matrix

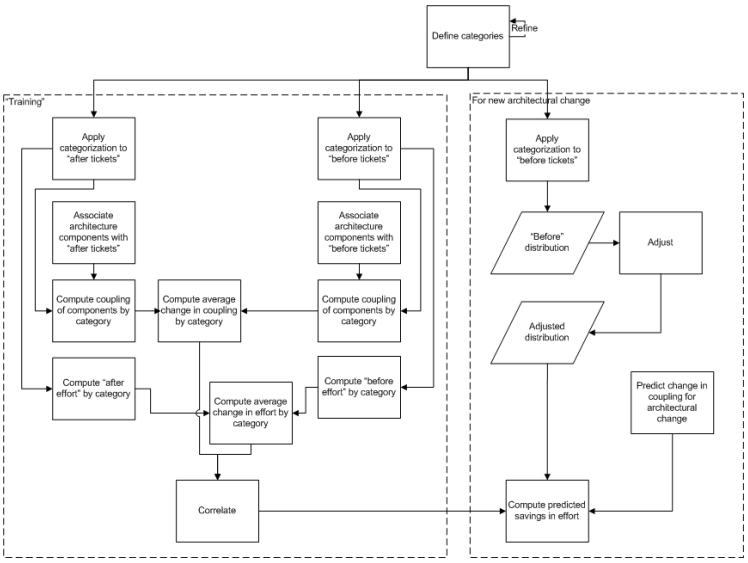
	A	B	C
A		Strength of B's dependency on A	
B	Strength of A's dependency on B		Strength of C's dependency on B
C		You get the idea…	

Hidden slide – the previous builds to this

A (small) piece of VistaPrint's DSM



Idealized process



Classification

- Price data
- Price test
- Tool
- Price logic
- Website UI

Result … the bad news

Classification	Time	LOC	Coupling
Price Test	-3%	11%	-19%
Price Data	61%	166%	-100%
Website UI	-54%	-73%	99%
Tool	-79%	-76%	271%
Price Logic	31%	5%	248%

In other words

Classification	Time	LOC	Coupling
Price Test			
Price Data			
Website UI			
Tool			
Price Logic			

Why?

- The classification was wrong
- Ticket scope wasn't normalized

Result … the good news

	Number of tickets	Total LOC	LOC per ticket	Average time per ticket
Before	129	31035	240	8336
After	156	25162	161	5781
Improvement			67%	69%

- Okay, then

Anecdotal evidence

- Easier integration of new users of pricing (decreased coupling to concrete implementation)
  - Price a basket
  - Third party pricing
- Easier integration of new pricing logic (consolidation)
  - Shipping pricing
- Correctness
  - Actually pricing accurately
- New capabilities
  - Managing discounts
  - Predictive modeling for cross-sell discount
  - Dynamic discounts
- Reduced error rate
- Improved testability
  - Unit testing
  - QA automation

So?

- Effort by LOC by category isn't informative or architecturally significant
- Anecdotal evidence and coarse effort analysis is positive
- Consider *propagation cost* to look for correlation

Propagation cost

- Start with the DSM, call it M
- Compute M_p – visibility based on pricing only
- Then,

$$V = M^0 + M + \sum_{i=2}^{4} M^{i-1} M^P$$

- Transform V to 1's and 0's
- Compute propagation cost as density of V

Propagation cost example

M^0

	A	B	P1	C	P2
A	1				
B		1			
P1			1		
C				1	
P2					1

M

	A	B	P1	C	P2
A	1		1		
B		1			1
P1			1	1	1
C		1		1	1
P2			1		1

M_P

	A	B	P1	C	P2
A			1		
B					1
P1			1	1	1
C					1
P2			1		1

M^2

	A	B	P1	C	P2
A	1		2	1	1
B		1	1		2
P1		1	2	2	3
C		2	1	1	3
P2			2	1	2

M^3

	A	B	P1	C	P2
A	1	1	4	3	4
B		1	3	1	4
P1		3	5	4	8
C		3	4	2	7
P2		1	4	3	5

M^4

	A	B	P1	C	P2
A	1	4	9	7	12
B		2	7	4	9
P1		7	13	9	20
C		5	11	6	16
P2		4	9	7	13

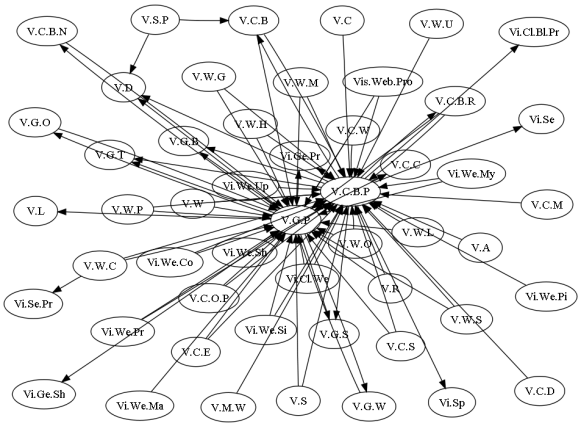
Propagation cost example

V

	A	B	P1	C	P2
A	1	1	1	1	1
B		1	1	1	1
P1		1	1	1	1
C		1	1	1	1
P2		1	1	1	1

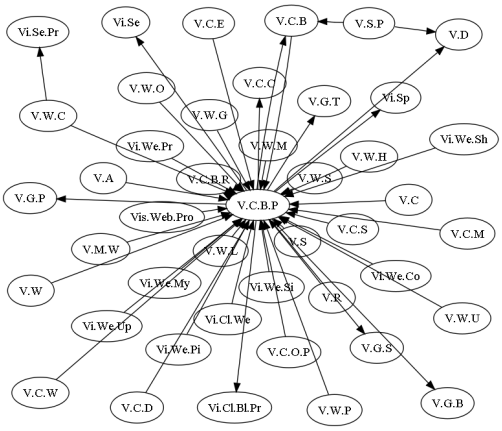
Propagation cost = 21/25 = 84%

Before (/during)



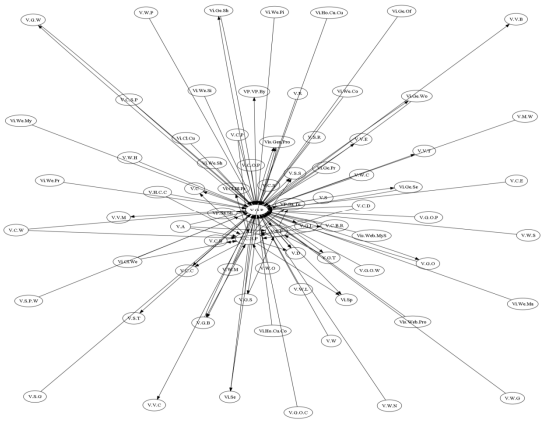
- Propagation cost: 20%

Before



- Propagation cost: 15%

After

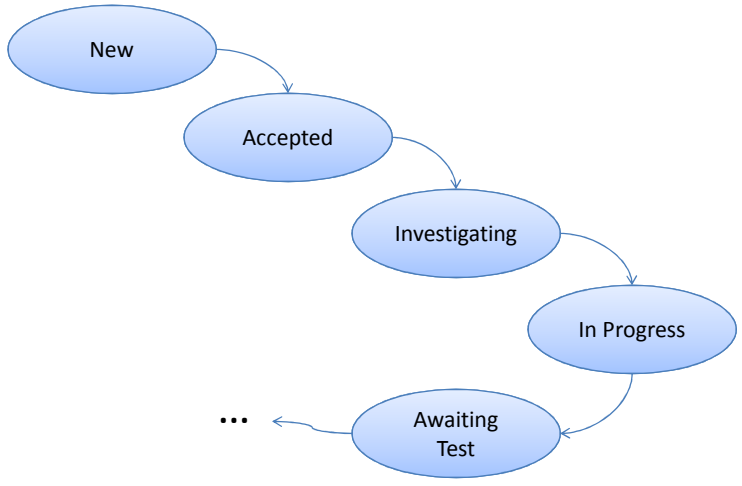


- Propagation cost: 6%

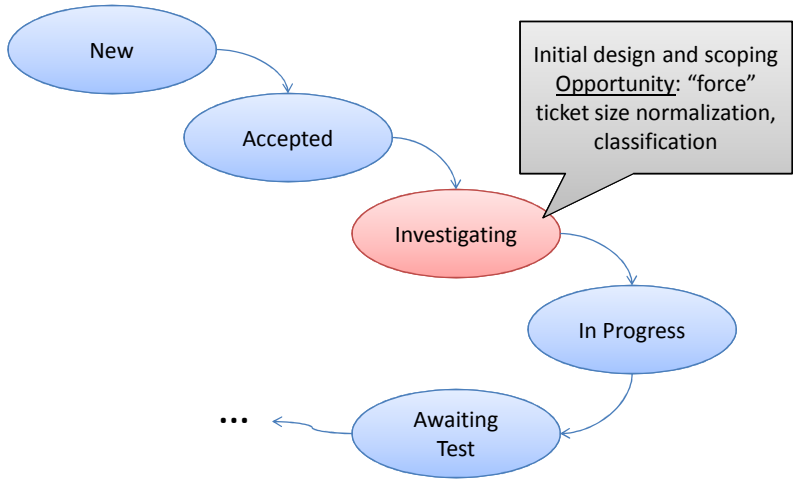
Actions

- Collect different data: actionable, timely
- New ticket classification, in-the-moment
  - Project-/area-specific
  - Universal: arrival frequency, projected effort
  - ISO/IEC 14764 maintenance classification probably not a fit
- Actual effort

Workflow



Workflow



Future work

- Immediate
  - Add propagation cost to metrics tracking
  - Finish defining new measures
  - Collect new measures during the process
  - Apply the model to upcoming projects, refine
- Medium term
  - Option model
- Speculative
  - Concept clustering on ticket descriptions