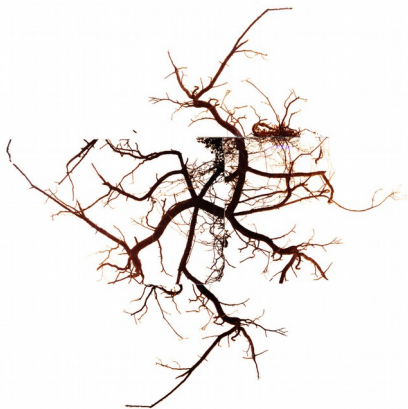
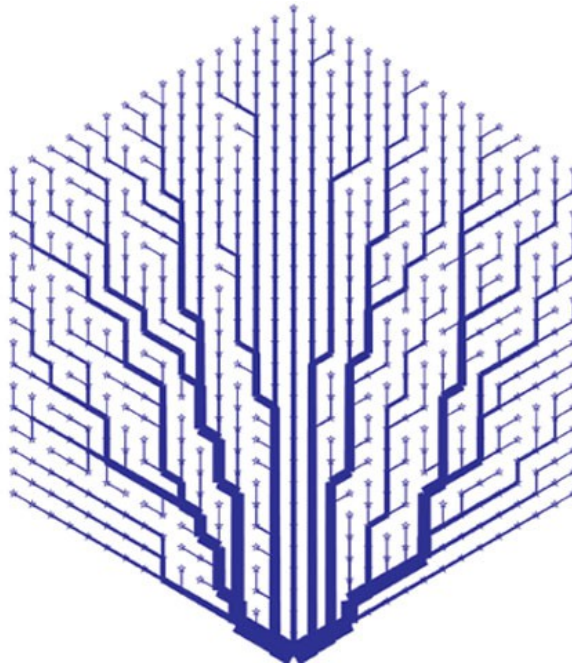
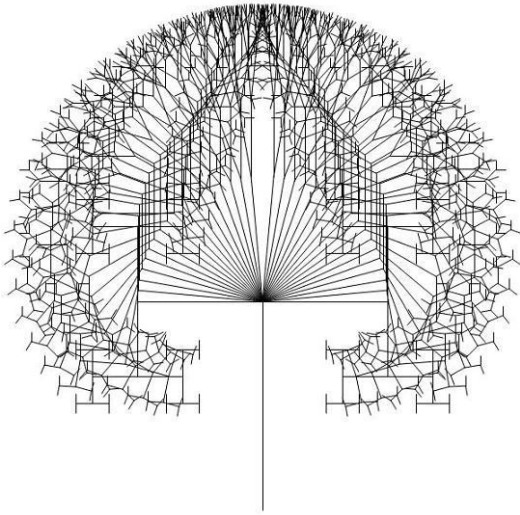
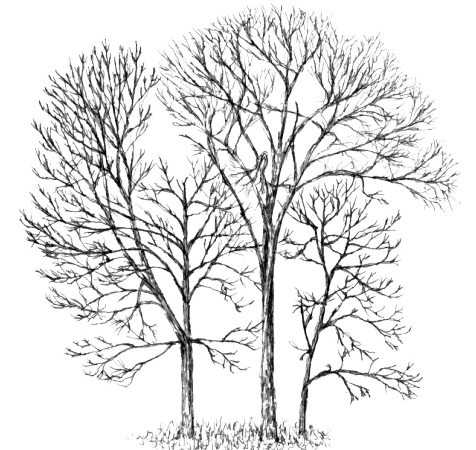


# Effect of Branching Strategy on Software Quality

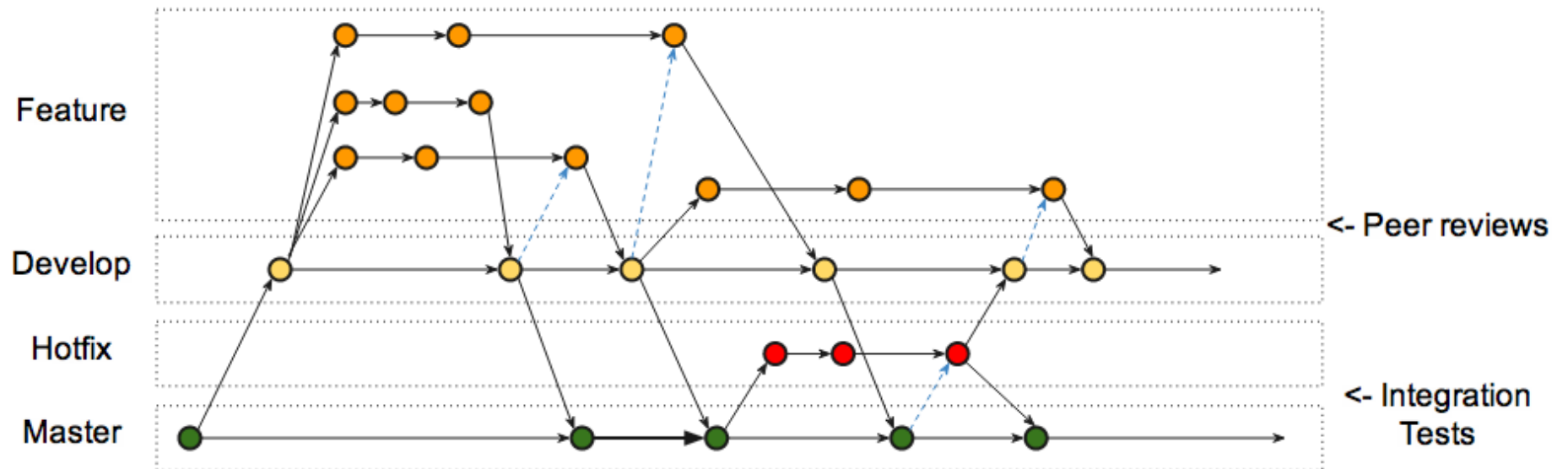


## Authors

Emad Shihab  
Christian Bird  
Thomas Zimmermann



# Branches in Git



Allows development of features with Isolation

May lead to integration failures due to merge conflicts and unseen dependencies

# Branching Strategies

- Branching according to Software Architecture
- Branching according to Organizational Structure

# Contributions

- **Define Metrics** to capture the effects on Software Quality
- **Quantify Effects of Branching** on Software Quality
- **Examine Effect of Mismatch** between branching structure and organizational structure.
- **Provide Recommendations** of proper usage of branch.

# Data Collection



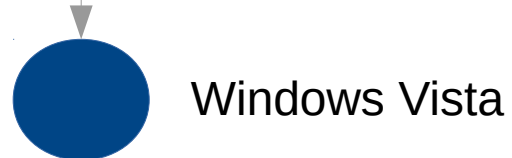
Windows Vista™



# Data Collection

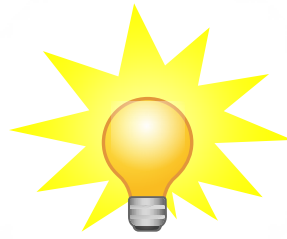


- Historical development data for each binary.
- Large releases **Vista** and **Win7**.

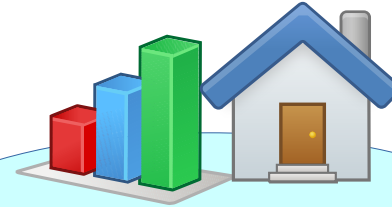


# Research Questions

- How much and in what ways does branching affect software quality?
  - Evaluate quality at the component level and the branch level through the use of a number of measures of branch use.
  - Represents the testable hypotheses.
  - Represents rationale that underlies each of these hypotheses.



Effects on Component Quality



Architectural and Organizational  
Congruence

Branch  
Activity

Branch  
Scatter

Branch  
Depth

According to  
System  
Architecture

According to  
Organizational  
Structure





# Hypotheses

## Effects on Component Quality

- H1: Branch Activity
  - Software components with high branching activity have more failures

# branching changes / # development changes



# Hypotheses

## Effects on Component Quality

- H2: Branch Scatter
  - a. Software components **spread across many branch families** have more failure
  - b. Software components **developed equally in multiple branch families** have more failure

$\# \text{BranchFamilies}(\text{component}) / \# \text{development changes}$

The more scattered in branches, the more is integration failure



# Hypotheses

## Effects on Component Quality

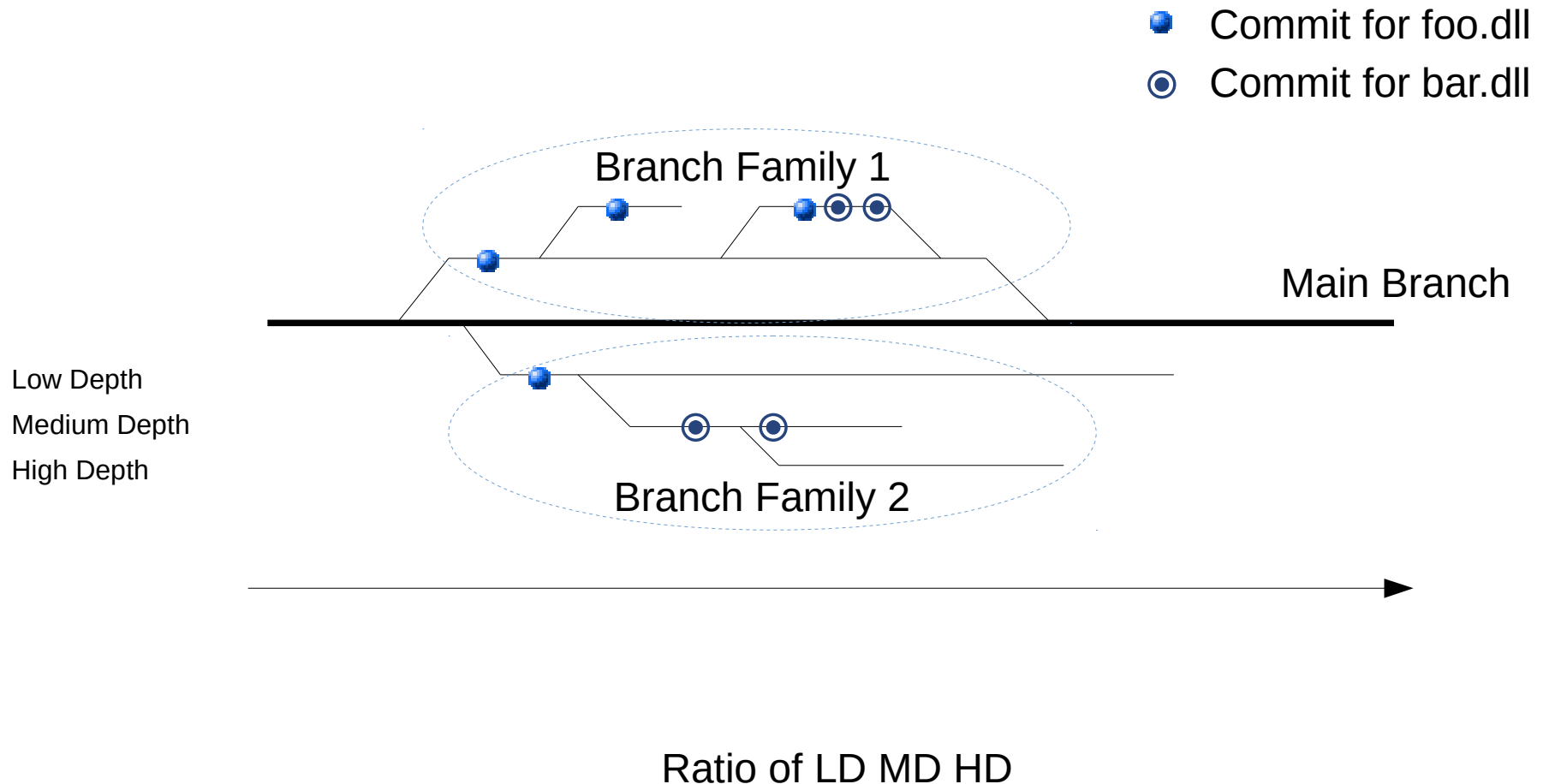
- H3: Branch Depth
  - a. Software components developed **primarily in deeper branches** have more failures
  - Software components developed **evenly in LD, MD and HD** have more failures



# Hypotheses

## Effects on Component Quality

- H3: Branch Depth





# Hypotheses

## Architectural and Organizational Congruence

- H4: Branching According to Architectural Structure
  - a. Branches with higher architectural mismatch have more failures
  - b. Branches with higher organizational mismatch have more failures



# Hypotheses

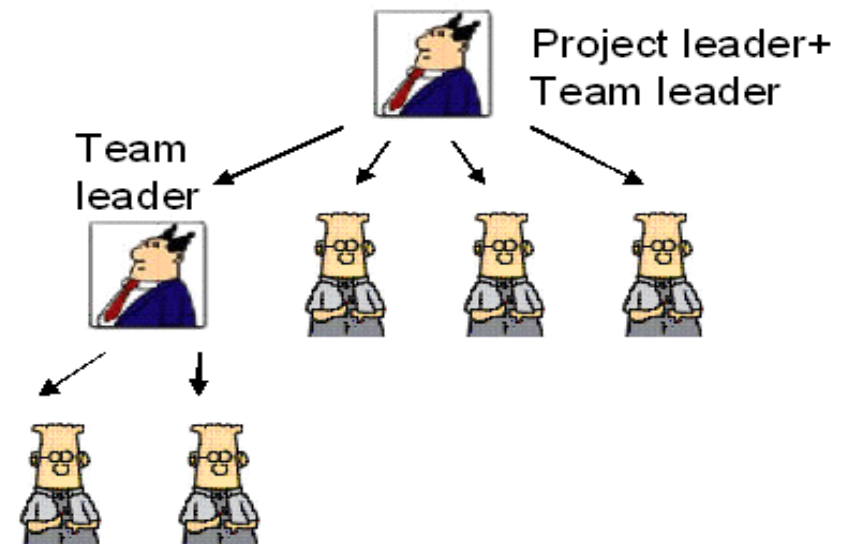
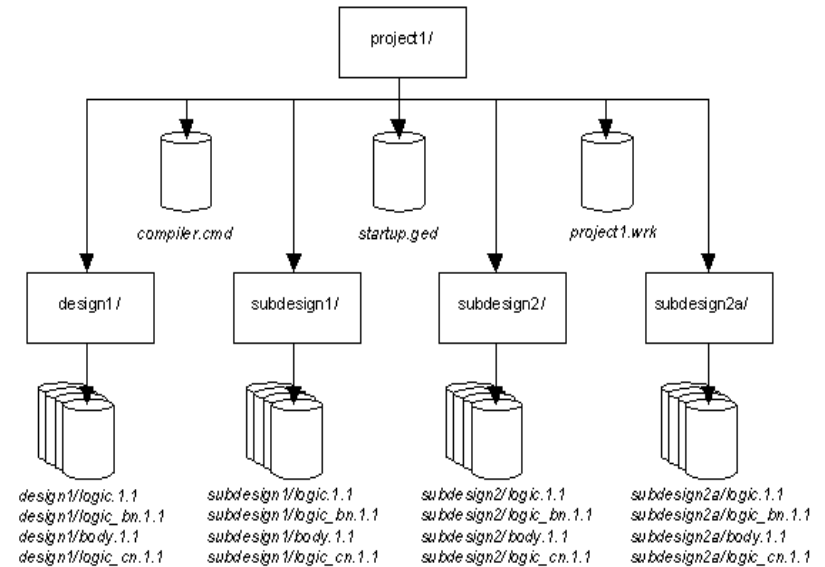
## Architectural and Organizational Congruence

- Architectural Mismatch:

- # Individual system
- # Subsystem
- # Area
- # Component
- # Sub-Components

- Organizational Mismatch:

- # Managers
- # Development Lead
- # Engineers



# Analysis Techniques

- Multiple Linear Regression
  - Model outcome of response variable
  - **Model the relationship**
- Measures are in component level
- Variance Inflation Factor (VIF) analysis

# Case Study Result

Post-release failure model for **Vista**

	Base Model	Model 2	Model 3	Model 4	Model 5
log(Size)	+	+	+	+	+
log(Churn)	+	+	+	+	+
log(Complexity)					
log(Dev. Changes)	+	+	+	+	
log(No. Files)	+	+	+	+	+
Branch Activity		+	+	+	+
log(Branch Scatter)			+		
BS Entropy			+	+	+
Low Branch Depth				+	+
log(High Branch Depth)					
Branch Depth Entropy				-	+
Branch Groups					-
R2	72%	75%	77%	77%	79%



# Case Study Result

## Post-release failure model for **Windows7**

	Base Model	Model 2	Model 3	Model 4	Model 5
log(Size)	+	+	+	+	+
log(Churn)	+	+	+	+	+
log(Complexity)	+	+	+	+	+
log(Dev. Changes)	+	+	+	+	
log(No. Files)	+	+	+	+	+
Branch Activity		+	+	+	+
log(Branch Scatter)					
BS Entropy			-	-	-
Low Branch Depth					
log(High Branch Depth)					
Branch Depth Entropy					
Branch Groups					+
R2	17%	18%	19%	19%	36%

# Effect of Branching on Software Quality

Model fit of Architectural and Organizational mismatch

	Vista	Windows7
Arch mismatch	0.426**	0.308**
Org mismatch	0.543**	0.321**
Org+Arch	0.594**	0.385**
(p<0.01**, p<0.05*)		

# Practical Implementation

- Reduce branch activity
- Reduce scattering development
- Branch structure alignment closer to organizational structure

# Threats

- Pre-release failures might in some cases be a measure of quality.
- Both projects are developed and by Microsoft