Defect Measurement and Analysis of JPL Software: A Case Study

John D. Powell & John N. Spagnuolo Jr.

Southern California ISPA Workshop
03/05/2003

Agenda

• Introduction

• Models and Data

• Defect Metrics Case Study on a JPL Project

• Conclusion
Introduction

• Purpose
  – Predictable software quality via the development of Defect Prediction Models
  • Predict defects early in project’s life cycle
  • Applicable across projects throughout JPL
    – Flight
    – Ground
    – Instrument

High Level Goals

• Make use of JPL defect data to:
  – Support Analysis to:
    • Determine Trends
    • Determine Exceptions to Trends
      – Find Explanations (i.e. Critical Discriminators) of Exceptions
  – Support Decision Making by:
    • Enhancing predictive capability
      – Prediction tools for Managers
      – Guidance by SQI Personnel
    • Determining appropriate Corrective Actions to resolve delta’s
      – delta = predictions - actuals
    • Process Improvement
**Defect Prediction Models**

![Diagram showing Defect Prediction Models]

- **Defects by Size**
  - \# defects = \( f(\text{KSLOC}, \text{Critical Discriminators}) \)

- **Defects by Criticality by Size**
  - \# defects = \( f(\text{KSLOC}, \text{Critical Discriminators}, \text{Criticality}) \)

- **Defects by Test Type by Size**
  - \# defects = \( f(\text{KSLOC}, \text{Critical Discriminators}, \text{Test Type}) \)

---

**Terminology - 1**

- **Software Development Set (SDS)**
  - Logical portions of the software system used for segregation of SQI M&B metrics data

- **Defect Collection and Categorization**
  - Individual Defects must be analyzed during collection
    - Defects must be categorized by SDS
    - Defect must be categorized by characteristics of SQI M&B
    - Defect Prediction models
      - Size
      - Criticality
      - Test type
  - Defect Data Repository must be analyzed for trends
Terminology - 2

- Definition of “Repair Hours” per PFR
  - Fix
    - Find and Write up Problem
    - Analyze Problem
    - Determine Solution
    - Code Solution
    - Unit Test
  - Large Scale Testing
    - Integration
    - Regression
  - Documentation
    - Additions
    - Corrections

Classification Hierarchy
Project 12
High Level Statistics

- Project 12 illustrates the advantages of consistent & conscientious defect metrics collection.
- Number of PFR’s - Test 453, Operations 127
- Defect density is 1.2 / LKSLOC
- Average Hours to repair - Test 12.3 hrs, Ops 18.6 hrs

11/13/2003
John D. Powell & John N. Spagnuolo Jr.

Project 12
High Level Observations

- On Average: OPS PFR’s cost more to fix than TEST PFR’s
- Number of PFR’s per SDS not proportional to Number of Work Hours per SDS
- Trend Breakers
  - Personnel Turnover and Code Breakage
  - Highly Interconnected SDS’s
- Trends don’t necessarily carry over from TEST to OPS

11/13/2003
John D. Powell & John N. Spagnuolo Jr.
On Average: OPS PFR's cost more to fix than TEST PFR's

### TEST / OPS Work Hours's

<table>
<thead>
<tr>
<th>Hours to fix</th>
<th>FrequencyOPS</th>
<th>FrequencyTEST</th>
<th>Cumulative%OPS</th>
<th>Cumulative%TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>20</td>
<td>30</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>20</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>15</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>12</td>
<td>40%</td>
<td>50%</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>6</td>
<td>50%</td>
<td>60%</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>2</td>
<td>60%</td>
<td>70%</td>
</tr>
<tr>
<td>240 More</td>
<td></td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

11/13/2003

John D. Powell & John N. Spagnuolo Jr.

---

Personnel Turnover and Code Breakage
Increase Defect Density - TEST

### LKLOC vs. # of pfr's

<table>
<thead>
<tr>
<th>LKLOC</th>
<th># of pfr's</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>9,006</td>
</tr>
<tr>
<td>TOG</td>
<td>4,382</td>
</tr>
<tr>
<td>Terminal</td>
<td>107,354</td>
</tr>
<tr>
<td>FTP Server</td>
<td>7,656</td>
</tr>
<tr>
<td>CCM</td>
<td>20,553</td>
</tr>
<tr>
<td>Common Software</td>
<td>19,798</td>
</tr>
<tr>
<td>Data Terminal System</td>
<td>1,737</td>
</tr>
<tr>
<td>UL Class</td>
<td>37,752</td>
</tr>
<tr>
<td>UI Server</td>
<td>21,322</td>
</tr>
<tr>
<td>Telnet</td>
<td>172,354</td>
</tr>
</tbody>
</table>

11/13/2003

John D. Powell & John N. Spagnuolo Jr.
Determination of CDs for Work

Hours per PFR

- Expected Reason for outliers
  - High Interconnectivity
- Confirmed Expectation with Project 12 personnel
- Empirical justification of outliers
  - Interconnectivity ratings from Project 12 personnel
  - Used clustering algorithm for empirical confirmation

Empirical Data supporting Outlier Status

\[ V_i = \{ \log(LKSLOC), \log(\text{Average PFR fix time for SDSi}), \log(\text{Coupling measure SDSi}) \} \]
OPS and TEST: Differences and Similarities

Number of PFR’s vs LKSLOC

<table>
<thead>
<tr>
<th>Feature</th>
<th>OPS</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TDG</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>TVS</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>FTP Server</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>COM</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Common Software</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Earth Terminal</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>UI Client</td>
<td>37</td>
<td>8</td>
</tr>
<tr>
<td>UI Server</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Telemetry</td>
<td>12</td>
<td>22</td>
</tr>
</tbody>
</table>

11/13/2003

John D. Powell & John N. Spagnuolo Jr.

Average Number of Work Hours vs LKSLOC

<table>
<thead>
<tr>
<th>Feature</th>
<th>OPS</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TDG</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>TVS</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>FTP Server</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>COM</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Common Software</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Earth Terminal</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>UI Client</td>
<td>37</td>
<td>8</td>
</tr>
<tr>
<td>UI Server</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Telemetry</td>
<td>12</td>
<td>22</td>
</tr>
</tbody>
</table>

11/13/2003

John D. Powell & John N. Spagnuolo Jr.
Prediction of Work Hours Spent on Selected SW Operations

- Use:
  - Estimated Work Hours per SDS
  - Complexity Estimates of SW Operations in SDS
    - Control Operations
    - Computational Operations
    - Device Dependent Operations
    - Data Management Operations
    - UI Management Operations
    - Interconnectivity Operations
  - LKSLOC of SDS’s

Chart for Work Hours Spent on Selected SW Operations - TEST

| LKSLOC  | 1.8 | 4  | 8  | 9  | 20 | 21 | 27.8 | 37 | 103 | 173 | 1.8 | 4  | 8  | 9  | 20 | 21 | 28 | 37 | 103 | 173 |
|---------|-----|----|----|----|----|----|------|----|-----|-----|-----|----|----|----|----|----|----|----|----|-----|-----|
| Control Operations - WH | 0.3 | 0  | 5  | 4  | 18 | 14.5 | 53  | 24 | 13  | 7  | 22 | 22 | 22  | 30 | 211 |
| Computational Operations - WH | 0.3 | 0  | 5  | 4  | 18 | 14.5 | 53  | 24 | 13  | 7  | 22 | 22 | 22  | 30 | 211 |
| Device Dependent Operations - WH | 0.3 | 0  | 5  | 4  | 18 | 14.5 | 53  | 24 | 13  | 7  | 22 | 22 | 22  | 30 | 211 |
| Data Management Operations - WH | 0.3 | 0  | 5  | 18 | 22 | 21.8 | 30  | 181 | 0.3 | 5  | 5  | 53  | 18  | 66  | 18  | 66  |
| UI Management Operations - WH | 0.3 | 0  | 5  | 4  | 18 | 21.8 | 30  | 181 | 0.3 | 5  | 5  | 53  | 18  | 66  | 18  | 66  |
| Interconnectivity Operations - WH | 0.3 | 0  | 5  | 4  | 18 | 21.8 | 30  | 181 | 0.3 | 5  | 5  | 53  | 18  | 66  | 18  | 66  |

Conclusions

- Desired defect data recording techniques
  - Requires little additional effort
    - Work Hours to Fix a defect
    - Consistency in filling PFR fields
  - Significantly increases power of analysis
    - Enhancing predictive capability
    - Determining appropriate Corrective Actions to resolve delta’s (predicted – actual)
    - Process Improvement for subsequent projects
Desired Defect Data

- Relationships
  - Density (by Size)
  - Complexity Correlation
  - Interconnectivity Correlation

- Individual Defect Characteristics
  - Criticality
  - Test Phase of Discovery
  - Development Phase Introduction/Cause
  - Effort to Repair

- Groupings
  - SDS
  - Version

Progress Metrics

- **2886 Defects** Analyzed and Categorized
- **12 Projects** for which Defect have been Analyzed and Categorized
- **28 SDS's** for which Defect have been Analyzed and Categorized
- Total Effort Working Defect = **11.9 work months**
- 1.13 M&B FTE = < **0.7** Defect FTE