Global Software Development in the FreeBSD Project

Diomidis Spinellis
Athens University of Economics and Business
Outline

- FreeBSD
- Methodology
- Findings
  - Global development
  - Productivity and quality
  - Human interactions
FreeBSD

- Complete operating system
- Derived from BSD Unix
- Focus on stability and reliability
  - Yahoo’s s servers
  - Part of Apple’s Max OS X
Development

- ~350 developers

- Kernel
  ~2.8MLoC

- Operating system utilities
  ~9.8MLoC

- Third party application ports
  ~12000

Documentation
Development Facilities

- **CVS** (10 years of historical data; >938k)
- Problem reports (>95k)
- Mailing lists (>100)
- Tinderbox
- Web site
  - Developer’s handbook
  - Release engineering
  - Browsable CVS
  - Mailing lists
  - PRs
Development Process

- Unpaid volunteers
- Teams
  - Core
  - Release engineering
  - Security
  - Ports
  - Donations
- Committer etiquette
- Mentoring
Global Development
Around the Globe
Division of Work

KLoC

Africa  Asia  Australia  Europe  North America  South America
Hold and Advance

Work type per area

KLoC

Main
Branch

Africa | Asia | Australia | Europe | North America | South America

0% | 20% | 40% | 60% | 80% | 100%
... and Fix

Closed GNATS entries per area

- Europe: 41%
- North America: 36%
- South America: 2%
- Asia: 13%
- Australia: 7%
- Africa: 1%
Roads and Buildings

Work type per area

- Africa
- Asia
- Europe
- South America
- North America
- Australia

- ports
- doc
- www
- src
Productivity and Quality

+ Round the clock development
+ Software and hardware diversity
- Lack of face-to-face communication
- Asynchronous communication
? A more rigorous process is required
? Cultural diversity
Productivity

- More developers produce more software (base case)

$r = 0.67$ (95% CI)
Productivity

- ... but developer distance doesn't affect productivity.

\[ r = -0.14 \ (95\% \ CI) \]
Code Style

![Graph showing the relationship between the number of developers and the percentage of non-adhering lines. The Pearson correlation coefficient is r=0.05 (95% CI).]
Defect Density

$r=0.07$ (95% CI)
Human Interactions

- Look for neighbors
  - Cooperating developers
  - Mentors and mentees
Cooperating developers

- Mean distance between
  - any two FreeBSD developers: 6,701km
- 4010 instances of commits
  - Same file
  - Same day
  - Different developers
  - Mean distance:
    6,489km
Mentors and mentees

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1/4</th>
<th>Media</th>
<th>Mean</th>
<th>3/4</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any</td>
<td>0</td>
<td>2.2</td>
<td>7.8</td>
<td>6.7</td>
<td>9.3</td>
<td>19</td>
</tr>
<tr>
<td>M/M</td>
<td>0</td>
<td>0.7</td>
<td>3.8</td>
<td>5.0</td>
<td>8.8</td>
<td>18</td>
</tr>
</tbody>
</table>
GSD Works

- Round the clock development
- No ill effects on
  - Productivity
  - Code quality
  - Bug density
- Distance
  - Doesn’t show up in ad hoc cooperation
  - Appears to matter in mentoring