7.3 Zachary Tatlock

Zachary Tatlock has not received prior NSF funding.
Project personnel

1. Zachary Tatlock; University of Washington; PI
2. Thomas E. Anderson; University of Washington; co-PI
3. Michael D. Ernst; University of Washington; co-PI
4. Suzanne Millstein; University of Washington; Other professional (staff programmer)
5. Calvin Loncaric; University of Washington; Graduate student (research assistant)
6. Chandrakana Nandi; University of Washington; Graduate student (research assistant)
7. Pavel Panchekha; University of Washington; Graduate student (research assistant)
8. Stuart Pernsteiner; University of Washington; Graduate student (research assistant)
9. Konstantin Weitz; University of Washington; Graduate student (research assistant)
10. James R. Wilcox; University of Washington; Graduate student (research assistant)
11. Doug Woos; University of Washington; Graduate student (research assistant)
12. Steve Anton; University of Washington; Undergraduate student
13. Alex Sanchez-Stern; University of Washington; Undergraduate student
14. Daryl Zuniga; University of Washington; Undergraduate student
15. Juliet Oh; International High School; High school student
Collaborators

1. Collaborators for Zachary Tatlock; University of Washington; PI
   1. Sonya Alexandrova; University of Washington
   2. Thomas Anderson; University of Washington
   3. Maya Cakmak; University of Washington
   4. Adam Chlipala; Massachusetts Institute of Technology
   5. Michael D. Ernst; University of Washington
   6. Dan Grossman; University of Washington
   7. Jon Jacky; University of Washington
   8. Dongseok Jang; University of California, San Diego
   9. David Lazar; Massachusetts Institute of Technology
  10. Sorin Lerner; University of California, San Diego
  11. Calvin Loncaric; University of Washington
  12. Eric Mullen; University of Washington
  13. Pavel Panchekha; University of Washington
  14. Stuart Pernsteiner; University of Washington
  15. Daniel Ricketts; University of California, San Diego
  16. Valentin Robert; University of California, San Diego
  17. Alex Sanchez-Stern; University of Washington
  18. Michael Stepp; University of California, San Diego
  19. Ross Tate; Cornell University
  20. Emina Torlak; University of Washington
  21. Xi Wang; University of Washington
  22. James R. Wilcox; University of Washington
  23. Doug Woos; University of Washington
  24. Nickolai Zeldovich; Massachusetts Institute of Technology

2. Collaborators for Thomas E. Anderson; University of Washington; co-PI
   1. Ivan Beschastnikh; University of British Columbia
   2. Ken Birman; Cornell
   3. Matthew Caesar; University of Illinois Urbana-Champaign
   4. Justin Cappos; NYU-Poly
   5. Dave Choffnes; Northeastern
   6. Doug Comer; Purdue
   7. Chase Cotton; Delaware
   8. Mike Dahlin; Google
   9. Colin Dixon; Brocade Networks
  10. Lucas Dixon; Google
  11. Mike Ernst; University of Washington
  12. Nick Feamster; Georgia Tech
  13. Mike Freedman; Princeton
  14. Andreas Haeberlen; University of Pennsylvania
  15. Daniel Halperin; University of Washington
  16. Zack Ives; University of Pennsylvania
17. Ethan Katz-Bassett; University of Southern California
18. Arvind Krishnamurthy; University of Washington
19. Taesoo Kim; Georgia Tech
20. William Lehr; MIT
21. Boon Thau Loo; University of Pennsylvania
22. Harsha Madhyastha; University of Michigan
23. David Mazieres; Stanford
24. Antonio Nicolosi; Stevens Institute of Technology
25. Simon Peter; University of Texas-Austin
26. Dan Ports; University of Washington
27. Sylvia Ratnasamy; University of California at Berkeley
28. Timothy Roscoe; ETH Zurich
29. Jonathan Smith; University of Pennsylvania
30. Ion Stoica; University of California Berkeley
31. Zach Tatlock; University of Washington
32. Robbert van Renesse; Cornell
33. Michael Walfish; NYU
34. Xi Wang; University of Washington
35. Hakim Weatherspoon; Cornell
36. David Wetherall; Google
37. Christopher Yoo; University of Pennsylvania

3. Collaborators for Michael D. Ernst; University of Washington; co-PI

1. Jenny Abrahamson; Microsoft
2. Ruth Anderson; University of Washington
3. Thomas E. Anderson; University of Washington
4. Shay Artzi; Zappix
5. Vinay Augustine; ABB
6. Richard Bailey; University of Washington
7. Magdalena Balazinska; University of Washington
8. Paulo Barros; Federal University of Pernambuco
9. Ivan Beschastnikh; University of British Columbia
10. Ravi Bhoraskar; Facebook
11. Yuriy Brun; University of Massachusetts
12. Yingyi Bu; Couchbase
13. Brian Burg; Apple
14. Juan Caballero; IMDEA Software Institute
15. Aaron Cammarata; VoidALPHA
16. John Cheng; Veracient LLC
17. Raymond Cheng; University of Washington
18. Seth Cooper; Northeastern University
19. Forrest Coward; Microsoft
20. Marcelo d’Amorim; Federal University of Pernambuco
21. Drew Dean; SRI International
22. Werner Dietl; University of Waterloo
23. Stephanie Dietzel; Tableau Software
24. Kellen Donohue; Google
25. Leonard Eusebi; Charles River Analytics
26. Gordon Fraser; University of Sheffield
27. Vijay Ganesh; University of Waterloo
28. Colin S. Gordon; Drexel University
29. Dan Grossman; University of Washington
30. Sean Guarino; Charles River Analytics
31. Philip J. Guo; University of Rochester
32. Seungyeop Han; University of Washington
33. Irfan Ul Haq; IMDEA Software Institute
34. Pingyang He; University of Washington
35. Reid Holmes; University of British Columbia
36. Pieter Hooimeijer; University of Virginia
37. Bill Howe; University of Washington
38. Wei Huang; Google
39. Laura Inozemtseva; University of Waterloo
40. Jon Jacky; University of Washington
41. Darioush Jalali; University of Washington
42. Ralph E. Johnson; University of Illinois
43. René Just; University of Massachusetts
44. Andrew Keplinger; Left Brain Games
45. Adam Kiezun; Broad Institute
46. Gene Kim; University of Washington
47. Andrew J. Ko; University of Washington
48. Karl Koscher; University of California at San Diego
49. Arvind Krishnamurthy; University of Washington
50. Wing Lam; University of Illinois
51. Dominic Langenegger; ETH Zurich
52. Jingyue Li; DNV Research & Innovation
53. Nuo Li; ABB
54. Calvin Loncaric; University of Washington
55. Alberto Lovato; Università di Verona
56. Hao Lü; Google
57. Damiano Macedonio; Julia Srl
58. Thomas Maddern; Veracient LLC
59. Kelly McLaughlin; XPD Analytics
60. Massimo Merro; Università degli Studi di Verona
61. Ana Milanova; Rensselaer Polytechnic Institute
62. Suzanne Millstein; University of Washington
63. Nathaniel Mote; University of Washington
64. John Murray; SRI International
65. Kıvanç Muşlu; Microsoft
66. David Notkin; deceased
67. Robert Ordóñez; Southern Adventist University
68. Johan Östlund; Uppsala University
69. Pavel Panchekha; University of Washington
70. Timothy Pavlik; University of Washington
71. Jeff H. Perkins; MIT
72. Stuart Pernsteiner; University of Washington
73. Paul Pham; The Evergreen State College
74. Amarin Phaosawasdi; University of Illinois
75. Zoran Popović; University of Washington
76. Alex Potanin; Victoria University of Wellington
77. Brian Robinson; ABB
78. Franziska Roesner; University of Washington
79. Todd W. Schiller; Bridgewater Associates
80. Sigurd Schneider; Saarland University
81. Will Scott; University of Washington
82. Michael Sloan; University of Washington
83. Ciprian Spiridon; Julia Srl
84. Eric Spishak; Google
85. Fausto Spoto; Università di Verona
86. Siwakorn Srisakaokul; University of Illinois
87. Luke Swart; University of Washington
88. Zachary Tatlock; University of Washington
89. Javier Thaine; University of Washington
90. Emina Torlak; University of Washington
91. Ben Treibelhorn; Seattle University
92. Mohsen Vakilian; Google
93. Paul Vines; University of Washington
94. Brian Walker; University of Washington
95. Xi Wang; University of Washington
96. Ronald Watro; Raytheon BBN
97. Konstantin Weitz; University of Washington
98. James R. Wilcox; University of Washington
99. Steven A. Wolfman; University of British Columbia
100. Doug Woos; University of Washington
101. Edward X. Wu; Extrahop
102. Jochen Wuttke; Google
103. Cheng Zhang; Shanghai Jiao Tong University
104. Sai Zhang; Google
105. Yoav Zibin; Google
Collaboration plan

The PIs are Zachary Tatlock, Michael Ernst, and Tom Anderson. We have an established, productive history of collaboration and co-authored the PLDI 2015 paper [WWP+15a] that inspired this grant proposal. We all work in the same department, with offices within a few feet of each other. We co-advice students and work together on multiple joint projects across a number of distinct areas of computer science, including verification, software engineering, and distributed systems. We participate in multiple joint weekly meetings, both on the Verdi project and on other current research topics. In particular, Tatlock and Ernst co-advice students Pavel Panchekha and Konstantin Weitz. Anderson and Ernst co-advice student Doug Woos, with Tatlock as an informal advisor. Tatlock is an informal advisor to Ernst’s student Calvin Loncaric, and Ernst is an informal advisor to Tatlock’s students Stuart Pernsteiner and James Wilcox. All of these students are interested in formal verification of distributed systems and will work on the proposed project. Ernst and Anderson also co-advised Ivan Beschastnikh (who graduated in 2013 and is currently a professor at UBC), whose work spans distributed systems and software engineering.

The PIs will jointly define and participate in the graduate seminars on systems, programming languages, and software engineering, which are held each quarter. In Spring 2015, Zach Tatlock, along with our colleague Xi Wang, will develop and teach a graduate seminar, CSE 599, on systems verification focusing on operating systems, distributed systems, and networking. This seminar will be project-based, and students will be encouraged to use Verdi to verify their own systems or work on further developing a part of the Verdi framework.

The Software Engineering and Programming Languages Research Lab is a place where students “hang out,” allowing not only in-depth interaction among the team members but also informal interactions and idea-sharing with other students in these areas. The PIs bring different styles, knowledge, and approach to the research table. This allows us to help students develop along multiple dimensions. We anticipate many technical challenges in this ambitious project — but luckily collaboration won’t be an obstacle.

7.4 Division of Work

As described in the proposal, we will investigate 10 research questions. The primary faculty leads for each question are listed below, but we stress that everyone on the team is a generalist. Although we each have different areas of expertise, we all intend to pitch in on whatever sub-project is needed to push the goals of the work forward. We strive to foster the same attitude among our graduate students.

RQ1 (Anderson, Tatlock)  RQ6 (Ernst, Tatlock)
RQ2 (Anderson, Tatlock)  RQ7 (Anderson, Tatlock)
RQ3 (Anderson, Ernst)  RQ8 (Ernst, Tatlock)
RQ4 (Ernst, Tatlock)  RQ9 (Ernst, Tatlock)
RQ5 (Anderson, Ernst)  RQ10 (Anderson, Ernst)

Each of these questions has a multi-year research and development agenda. Instead of trying to complete one area and then the next, our plan is to work on them in parallel as the results from many of these questions can help developing solutions to the rest. We will split the research agenda in each area into a series of six-month challenges that we can integrate into the mainline Verdi framework if successful, and learn from our failures if not. The idea is to repeatedly course-correct, to go from working system to working system.
7.5 **Collaboration Mechanisms**

We plan to continue the model we have used in the recent past:

- Graduate students will be co-advised in almost all cases. Not only do students get the benefit of multiple perspectives, but this provides an almost daily opportunity for faculty to discuss issues with respect to the project.

- Graduate researchers will typically work on sub-projects in small teams of 2–3; in our view, student collaboration across project topics is as advantageous as faculty collaboration.

- Undergraduate researchers will be paired with a graduate student mentor, overseen by one other faculty member. We have found that undergraduates benefit from more frequent interaction than is often possible with over-committed faculty. Furthermore, graduate students learn a lot from the experience of being an advisor.

- We will have weekly project meetings, including faculty and students, to discuss research progress and plan future work. We believe students should see the process of how decisions are made, rather than making those decisions behind closed doors.

- We will hold weekly PI meetings to track progress against progress goals, diagnose and fix student issues, and strategize industrial adoption.

- We will organize and hold an annual off-site meeting to coordinate with practitioners from our industrial partners.
Data management plan

The goal of this project is to enable the creation of formally-proven distributed system implementations. To evaluate the space of design alternatives, we plan to build proofs and extract runnable source code from them; we will also run experiments on the running systems. All source code, along with configuration files and instructions for reproducing published measurements of the system, will be made public, hosted via GitHub and accessible from the project website at the University of Washington.

The principal investigators have a long history of making research software and data publicly available, so that other scientists can reproduce and build on the work. Anderson has led the development of several widely used open source software artifacts over the past few years, most notably BitTyrant [PIA+07] and OneSwarm [IPKA10]. Both have been downloaded more than a hundred thousand times. Most recently they have led an effort to produce an easy to configure public anti-censorship tool, uProxy, to enable client browsers to work together to evade censorship [Upr]. uProxy is open source software, and it has more than ten thousand enrolled beta users for an initial release in the next few months. They have also led the iPlane Internet topology mapping project [MIP+06]. iPlane has collected and published Internet topology maps continuously since 2006. The iPlane system measures millions of paths per day and the results have been used by more than thirty research groups. Ernst recently led the development of the Checker Framework (http://checkerframework.org/), which is downloaded about 300 times per day, and his Daikon (http://plse.cs.washington.edu/daikon/) and Randoop (http://mernst.github.io/randoop/) projects are even more popular. Dozens more software downloads are listed at http://homes.cs.washington.edu/~mernst/software/ and http://homes.cs.washington.edu/~mernst/pubs/.

As NSDI steering committee chair, Anderson was instrumental in encouraging NSDI to create a special award for those papers whose source code and data sets are placed in the public domain. Tatlock is the Artifact Evaluation Committee Co-Chair for PLDI 2016. In this role, he will encourage the authors of all accepted papers to submit their implementations to the committee which will attempt to independently replicate the paper’s results before the conference is held.

Policies for access and sharing. We will make the source code for our tools and all test configurations available for public download using the BSD or Apache license. To ensure the broadest impact of our work, we will make the source code available for re-use without restriction.

Reproducibility of experimental results. We will publish our results and algorithms using peer-reviewed venues such as PLDI, POPL, ICSE, FSE, OSDI, and NSDI, as well as through technical reports freely available through our department web sites. All published experimental results will be accompanied by public source code, configuration files, and data sets, allowing others to directly reproduce our results.

Plans for archiving data. Data archival is essential for the repeatability of experiments and longitudinal studies. We will ensure archival of the measured data reported in our published work. We expect these data sets to be small enough to be served out of the same GitHub and department repositories for the project source code.

Standards, data format and content. There is no accepted standard for storing and maintaining proofs (other than the source code of the proof in Coq) and experimental data from distributed systems. We will ensure that the data set and source code distribution includes sufficient documentation to allow for the published data to be reconstructed by a third party.
References


[DHJ+]
Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall, and Werner Vogels. Dynamo: Amazon’s highly available key-value store. SOSP ’07.

[DLE03]

[ECGN01]

[EPG+07]

[FGL12]

[GAM+07]

[GEG13]

[GL00]

[GMT09]

[Gol14]

[GRF13]

[Hay98]

[Hig11]


