

# Alfredo Giménez | C.V.

339 41st Street Apt. A, Oakland, CA 94609

☎ +1 916 622-7078 • ✉ alfredo.gimenez@gmail.com  
🌐 <https://github.com/alfredo-gimenez>

Ph.D. graduate in Computer Science at the University of California, Davis and staff computer scientist at Lawrence Livermore National Laboratory. Expertise in High-Performance Computing (HPC) and large-scale data analysis and visualization. Interested in using data science for understanding and predicting the complex behaviors in data motion, energy consumption, faults, and temperature in HPC environments.

## Education

- **University of California, Davis** **Davis, CA**  
*Doctor of Philosophy, Computer Science* September 2011–December 2017  
Institute of Data Analysis and Visualization. Advisor: Bernd Hamann
- **University of California, Davis** **Davis, CA**  
*Bachelor of Science, Computer Science* September 2006–November 2010

## Experience

### Research

- **Lawrence Livermore National Laboratory** **Livermore, CA**  
*Livermore Computing Staff* December 2016–Current  
Research and development of data collection and analysis methods for analyzing the performance of HPC applications and computing facilities.
- **Lawrence Livermore National Laboratory** **Livermore, CA**  
*Research Intern* April 2016–December 2016  
Research and development for automatic data processing of large disparate data sources for the purpose of creating machine learning models to predict HPC performance behavior.
- **Lawrence Livermore National Laboratory** **Livermore, CA**  
*Research Intern* June 2015–September 2015  
Research and development for large-scale data storage, parallel processing, and analysis of HPC performance data. Developed advanced methods for correlating memory data to application meshes and using memory access samples to predict performance bugs (in collaboration with several other student interns).
- **Lawrence Livermore National Laboratory** **Livermore, CA**  
*Research Intern* June 2014–September 2014  
Research and development for finding patterns and correlations in memory access data within MemAxes. Created Mitos, a data collection and program annotation tool.
- **Lawrence Livermore National Laboratory** **Livermore, CA**  
*Research Intern* June 2013–September 2013  
Research and development of data collection tools for memory access information. Created MemAxes, a visualization/analysis tool for memory accesses on complex hardware topologies.

### Industry

- **Stratovan Corporation** **Davis, CA**  
*Performance Consultation* January 2015–June 2015  
Research/Development for optimizing the performance of Stratovan's segmentation and detection algorithms. Helped them achieve a 4x throughput improvement.

- Intel Corporation**  
 Software Research and Development  
 Research/Development to showcase the capabilities of heterogeneous computing on Intel integrated GPU and CPU hardware using OpenCL. Created a debugging and analysis tool for OpenCL programs that eventually became the OpenCL Code Builder that is currently shipped with Intel's OpenCL SDK. Co-authored a patent for the hardware-agnostic OpenCL line-by-line debugging capability in Code Builder.

**Folsom, CA**  
*June 2010–October 2012*

## Projects

---

### Led by me.....

- Sonar Project Co-Lead**  
 Data science cluster and tools for collection and analysis of large-scale HPC performance data.
  - ScrubJay (Open-source release pending)**  
 Data integration framework for automatic processing of large, heterogeneous data sources.
  - MemAxes** <https://github.com/LLNL/MemAxes>  
 Visualization and analysis of memory accesses on complex hardware topologies.
  - Mitos** <https://github.com/LLNL/Mitos>  
 Memory access sampling and data structure instrumentation interface.

### Led by colleagues.....

- Caliper** <https://github.com/LLNL/Caliper>  
 Generic context annotation tool for HPC performance data collection and integration.
  - spack** <https://github.com/spack/spack>  
 Package manager for complex HPC software dependencies.
  - SOSflow** [https://github.com/cwdirect/sos\\_flow](https://github.com/cwdirect/sos_flow)  
 A flexible, scalable, and programmable framework for in-situ observation, introspection, feedback, and control of HPC applications.

## Technical Skills

---

- Programming Languages:** C, C++, Python, Java, Scala, JavaScript
  - Parallel Programming Models:** Spark, MapReduce, OpenMP, MPI, OpenCL, pthreads
  - Data Analysis:** pandas, numpy, scipy, scikit-learn, MLlib
  - Visualization/Graphics:** OpenGL, WebGL, DirectX, VTK, D3.js, VisIt, ParaView, Qt, matplotlib, gnuplot
  - Database Infrastructure:** Cassandra, HBase, Hadoop, Kafka, SQL
  - Developer Tools:** git, svn, Perforce, cmake, gdb, Visual Studio

## Papers

---

- [1] Alfredo Giménez, Todd Gamblin, Abhinav Bhatele, Chad Wood, Kathleen Shoga, Anirudha Marathe, Peer-Timo Bremer, Bernd Hamann, and Martin Schulz. "ScrubJay: deriving knowledge from the disarray of HPC performance data". In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2017, Denver, CO, USA, November 12 - 17, 2017*. 2017, 35:1–35:12. DOI: 10.1145/3126908.3126935. URL: <http://doi.acm.org/10.1145/3126908.3126935>.
- [2] Alfredo Giménez, Todd Gamblin, Ilir Jusufi, Abhinav Bhatele, Martin Schulz, Peer-Timo Bremer, and Bernd Hamann. "MemAxes: Visualization and Analytics for Characterizing Complex Memory Performance Behaviors". In: *IEEE Transactions on Visualization and Computer Graphics* (May 2017). LLNL-JRNL-. URL: <http://doi.ieeecomputersociety.org/10.1109/TVCG.2017.2718532>.

- [3] H. Xu, S. Wen, A. Gimenez, T. Gamblin, and X. Liu. "DR-BW: Identifying Bandwidth Contention in NUMA Architectures with Supervised Learning". In: *2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. May 2017, pp. 367–376. DOI: 10.1109/IPDPS.2017.97.
- [4] David Böhme, Todd Gamblin, David Beckingsale, Peer-Timo Bremer, Alfredo Giménez, Matthew P. LeGendre, Olga Pearce, and Martin Schulz. "Caliper: performance introspection for HPC software stacks". In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2016, Salt Lake City, UT, USA, November 13-18, 2016*. 2016, pp. 550–560. DOI: 10.1109/SC.2016.46. URL: <https://doi.org/10.1109/SC.2016.46>.
- [5] Chad Wood, Sudhanshu Sane, Daniel A. Ellsworth, Alfredo Giménez, Kevin A. Huck, Todd Gamblin, and Allen D. Malony. "A Scalable Observation System for Introspection and In Situ Analytics". In: *5th Workshop on Extreme-Scale Programming Tools, ESPT@SC 2016, Salt Lake City, UT, USA, November 13, 2016*. 2016, pp. 42–49. DOI: 10.1109/ESPT.2016.010. URL: <https://doi.org/10.1109/ESPT.2016.010>.
- [6] Benafsh Husain, Alfredo Giménez, Joshua A. Levine, Todd Gamblin, and Peer-Timo Bremer. "Relating memory performance data to application domain data using an integration API". In: *Proceedings of the 2nd Workshop on Visual Performance Analysis, VPA 2015, Austin, Texas, USA, November 15, 2015*. 2015, 5:1–5:8. DOI: 10.1145/2835238.2835243. URL: <http://doi.acm.org/10.1145/2835238.2835243>.
- [7] Garrett Aldrich, Alfredo Giménez, Michael Oskin, Richard Strelitz, Jonathan Woodring, Louise H. Kellogg, and Bernd Hamann. "Curvature-Based Crease Surfaces for Wave Visualization". In: *VMV 2014: Vision, Modeling & Visualization, Darmstadt, Germany, 2014. Proceedings*. 2014, pp. 39–46. DOI: 10.2312/vmv.20141274. URL: <http://dx.doi.org/10.2312/vmv.20141274>.
- [8] Alfredo Giménez, Todd Gamblin, Barry Rountree, Abhinav Bhatele, Ilir Jusufi, Peer-Timo Bremer, and Bernd Hamann. "Dissecting On-node Memory Access Performance: A Semantic Approach". In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. SC '14*. New Orleans, Louisiana: IEEE Press, 2014, pp. 166–176. ISBN: 978-1-4799-5500-8. DOI: 10.1109/SC.2014.19. URL: <http://dx.doi.org/10.1109/SC.2014.19>.
- [9] Katherine E. Isaacs, Alfredo Giménez, Ilir Jusufi, Todd Gamblin, Abhinav Bhatele, Martin Schulz, Bernd Hamann, and Peer-Timo Bremer. "State of the Art of Performance Visualization". In: *EuroVis - STARs*. Ed. by R. Borgo, R. Maciejewski, and I. Viola. The Eurographics Association, 2014. ISBN: -. DOI: 10.2312/eurovisstar.20141177.
- [10] René Rosenbaum, Alfredo Giménez, Heidrun Schumann, and Bernd Hamann. "A flexible low-complexity device adaptation approach for data presentation". In: vol. 7868. 2011, 78680F-78680F-12. DOI: 10.1117/12.871975. URL: <http://dx.doi.org/10.1117/12.871975>.
- [11] Alfredo Giménez, René Rosenbaum, Mario Hlawitschka, and Bernd Hamann. "Using R-Trees for Interactive Visualization of Large Multidimensional Datasets". In: *Advances in Visual Computing - 6th International Symposium, ISVC 2010, Las Vegas, NV, USA, November 29 - December 1, 2010, Proceedings, Part II*. 2010, pp. 554–563. DOI: 10.1007/978-3-642-17274-8\_54. URL: [http://dx.doi.org/10.1007/978-3-642-17274-8\\_54](http://dx.doi.org/10.1007/978-3-642-17274-8_54).

## Patents

---

- [1] J. Bottleson and A. Gimenez. *Kernel functionality checker*. US Patent App. 13/995,734. Apr. 2015. URL: <https://www.google.com/patents/US20150121051>.

## Posters

---

- [1] Alfredo Gimenez, Benafsh Husain, David Boehme, Todd Gamblin, and Martin Schulz. "Mitos: A Simple Interface for Complex Hardware Sampling and Attribution". In: *Supercomputing 2015*. Austin, Texas, Nov. 2015. URL: [http://sc15.supercomputing.org/sites/all/themes/SC15images/tech\\_poster/poster\\_files/post191s2-file2.pdf](http://sc15.supercomputing.org/sites/all/themes/SC15images/tech_poster/poster_files/post191s2-file2.pdf).

## Standalone Presentations

---

- [1] Alfredo Gimenez. "Working in the Application Domain". VAPLS 2013 Workshop on Visualization and Analysis of Performance on Large-scale Software. 2013. URL: <http://www.sci.utah.edu/~prosen/vapls2013/slides/5-Gimenez-ApplicationDomain.pdf>.