Alfredo Giménez | C.V.

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

+1 916 622-7078 • alfredo.gimenez@gmail.com

Technical project leader at Lawrence Livermore National Laboratory, Ph.D. graduate in Computer Science at the University of California, Davis. Expertise in data engineering, visualization/analysis, high-performance computing.

Work Experience

Current

- **Lawrence Livermore National Laboratory** **Livermore, CA**
  Livermore Computing Staff *December 2016–Current*
  Technical project leader for 1) An end-to-end data science platform for analyzing performance of HPC applications and facilities across Livermore Computing, and 2) A multi-lab research effort for exploring hypothetical next-generation memory architectures.

Graduate Student Internships

- **Lawrence Livermore National Laboratory** **Livermore, CA**
  Research Intern *April 2016–December 2016*
  Researched and developed an automated data integration tool for combining multiple streaming heterogeneous data sources for monitoring HPC application and system performance.
- **Lawrence Livermore National Laboratory** **Livermore, CA**
  Research Intern *June 2015–September 2015*
  Developed tools for collection and distributed storage/processing/analysis of performance data from disparate information sources.
- **Lawrence Livermore National Laboratory** **Livermore, CA**
  Research Intern *June 2014–September 2014*
  Researched advanced data analytic methods for identifying memory performance bottlenecks in applications and complex memory architectures.
- **Lawrence Livermore National Laboratory** **Livermore, CA**
  Research Intern *June 2013–September 2013*
  Researched and developed visualization and exploratory analysis methods for on-node memory access information (Initiated open source projects: Mitos and MemAxes)

Other

- **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** **Folsom, CA**
  Software Research and Development *June 2010–October 2012*
  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.

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*

Projects

Led by me

- **Sonar** *Technical Project Lead*
  Data science platform for collection, storage, processing, and analysis of large-scale HPC performance data.
- **Exascale Computing Project - Memory Technologies** *Technical Project Lead*
  Research for predictively designing future on-node memory architectures (https://www.exascaleproject.org/)
- **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/cdwdirect/sos_flow`
  A flexible, scalable, and programmable framework for in-situ observation, introspection, feedback, and control of HPC applications.

Technical Skills

- **Programming Languages:** Python, Scala, Java, C, C++, JavaScript
- **Parallel Programming Models:** Spark, MapReduce, OpenMP, MPI, OpenCL, pthreads
- **Data Engineering:** Kafka, Cassandra, HBase, Hadoop, SQL/NoSQL
- **Data Analysis:** Pandas, Numpy, SciPy, Scikit-Learn, MLlib
- **Visualization/Graphics:** OpenGL, WebGL, DirectX, VTK, D3.js, VisIt, ParaView, Qt, matplotlib, gnuplot
- **Developer Tools:** git, svn, Perforce, cmake, gdb, Visual Studio

Papers

[1] Alfredo Giménez, Todd Gamblin, Abhinav Bhatele, Chad Wood, Kathleen Shoga, Aniruddha 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, Louisana: 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.

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.

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.