

# Training by the Research Computing Group



**Digital Research  
Alliance** of Canada

Join our weekly mailing list



<https://training.westdri.ca/contact>

# Training team



MARIE-HÉLÈNE BURLE  
[marie.burle@westdri.ca](mailto:marie.burle@westdri.ca)

- Background in evolutionary & behavioural ecology (SFU)
- 18 years training and support experience (university courses, Research / Student Learning Commons, WestGrid, RCG)
- Passionate open-source advocate, certified Software Carpentry instructor



ALEX RAZOUMOV  
[alex.razoumov@westdri.ca](mailto:alex.razoumov@westdri.ca)

- Background in computational astrophysics (UBC)
- Postdocs in Urbana-Champaign, San Diego, Oak Ridge, & Halifax
- 30+ years training and support experience (university courses, SHARCNET, WestGrid, RCG)
- National Visualization Team lead since 2014, certified Software Carpentry instructor

Both can be reached at [training@westdri.ca](mailto:training@westdri.ca)

# Research computing training at SFU and beyond

- RCG runs (one of) the largest academic supercomputers in Canada
- Our funding:
  - ✓ supports training researchers in their actual computational workflows, to be run on our systems (max science, max hardware utilization)
  - ✗ does not support undergraduate academic programs (teaching university courses)
- In practice, we often see graduate students lacking basic computational skills (and trying to use our clusters), highlighting the need for earlier training and closer collaboration with departments
- SFU is one of the Alliance's national host sites  $\Rightarrow$  our training and support are available to all academic researchers across Canada
  - anyone with a valid Canadian university email address can register for our courses
  - webinars and online materials open to anyone
  - we answer national support tickets and contribute to the national docs <https://docs.alliancecan.ca>




# Topics: full-day courses

Details/abstracts at <https://training.westdri.ca/courses>

- Remote computing basics
  - Bash command line
  - Introduction to HPC
  - + Specific remote computing tools
- Programming basics
  - Python • R • Julia
  - Scientific Python
  - Webscraping or GIS in Python or R
  - Selected Python libraries: Polars, PyTables, ...
- Parallel coding (next slide)
  - Julia • Chapel • Python • R
  - GPU computing with Chapel, Python, Julia
- Deep learning
  - PyTorch • JAX+Flax
  - Neural nets for specific research workflows
- Virtualization
  - Intro to Apptainer containers
  - Intro to cloud VMs (OpenStack)
- Scientific visualization (15+ workshops)
  - ParaView • VisIt • 3D vis for the humanities
  - Python/R plotting libraries
  - In-situ vis, remote and large-scale parallel rendering, scripting, animation, interactive web visualization, other advanced sci-vis topics
- RDM topics
  - Version control with Git
  - Distributed datasets: git annex, or DataLad, or DVC
  - Managing files and formats: scientific file formats, large datasets, parallel I/O, container overlays, etc.
- Publishing
  - R Markdown and Quarto • Interactive dashboards

# Parallel coding: our bread and butter

- 10-15 years ago we used to teach MPI (message-passing interface) and OpenMP in C/C++/Fortran (most common HPC languages)
- Today's "Intro to HPC" covers MPI and OpenMP when discussing parallel frameworks
  - but we don't teach actual coding in this course
  - focus on basic cluster tools, software environments, running codes (serial, shared-memory, distributed-memory, GPU, and their various combinations), efficient resource utilization, and HPC best practices
-  **CHAPEL**
  - fast, modern programming language for both shared- and distributed-memory systems
  - high-level, easy-to-use abstractions for data and task parallelism
  - one syntax across 3 levels of hardware parallelism: **multi-core CPUs**, **multi-node clusters**, and **GPUs**
  - often hides the distinction between multithreading and multiprocessing from the programmer
  - PGAS (Partitioned Global Address space) programming model + **computation follows data**
    - lower-level network communication hidden from the programmer (unlike in MPI)
  - focus on data parallelism through two numerical examples: one embarrassingly parallel (2D Julia set fractals) and one tightly coupled (heat diffusion) ⇒ work on bottlenecks, load balancing, scaling

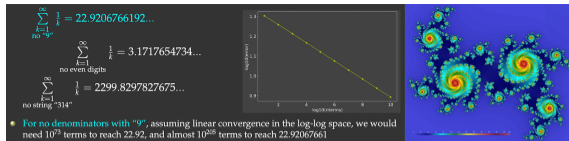
# Parallel coding (continued)



- fast, high-level programming language for scientific computing and data science
- just-in-time compilation + interactive shell
- we take you on a journey to parallelize two problems: slow series and 2D Julia set
- multi-threading with Base.Threads and ThreadsX.jl
- multi-processing with Distributed.jl + arrays with DistributedArrays.jl and SharedArrays.jl
- work on thread safety, race conditions, benchmarking, load balancing



- aim: make it as fast + scalable as Julia or Chapel for numerical problems
- explore ~dozen different approaches: most fail for the slow series, and we analyze why
- eventually settle on a combination of **Numba for compilation** and **Ray for distributed computing**
- cover parallelizing with both Ray core and Ray data
- our latest GPU modules: Arrays with CuPy, Numba JIT compilation to CUDA kernels, DataFrames with Polars and RAPIDS



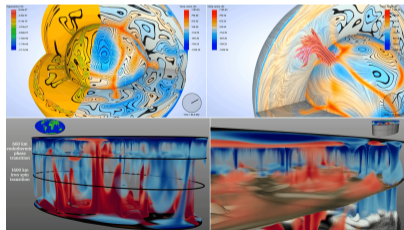
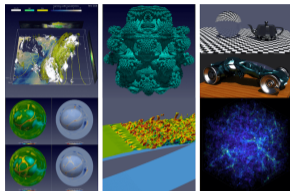
# 70 – 90 events each year across all formats

Current events at <https://training.westdri.ca> | Events | select the season

1. Biweekly webinars since March 2016 (during academic year)
  - advanced and introductory topics **not already covered in our courses**
  - focus on new and/or impactful open-source tools and techniques that will benefit researchers
  - most webinars delivered by the two of us; last year partnered with UBC
  - 150+ past webinars archived at <https://training.westdri.ca> (chronologically and by topic)
  - watch <https://training.westdri.ca/blog> for fall schedule, or subscribe to our newsletter
2. Weekly online courses since February 2023 (during academic year)
3. Annual week-long, in-person summer schools since 2017
  - typically every June @SFU: 5 full days, at the Big Data Hub, ~ 70 registrants
  - this year June 1-5 @SFU, details and register at <https://sfu26.netlify.app>
  - would love to run a couple per year and distribute them geographically; previously organized and taught these at UBC, UVic, U of Manitoba, U of Saskatchewan, U of Calgary, UNBC
4. National and international training events: HSS/Visualization Winter Series, national courses, DHSI (since 2016), multi-year PIMS/M2PI collaboration, IHPCSS
5. Local free workshops and bootcamps available upon request for departments and research groups

# Researcher support and outreach

- Official support [support@tech.alliancecan.ca](mailto:support@tech.alliancecan.ca) from the Alliance, mostly for specific problems on our clusters  $\Rightarrow$  tickets answered by HPC / ARC analysts across the country
- Followup from our training, support emails and tickets
- Visualization support
  - I lead the National Visualization Team <https://ccvis.netlify.app>
  - answer most visualization tickets in the national queue
  - HPC background  $\Rightarrow$  can visualize very large (GBs to TBs and larger) datasets
  - recent sci-vis projects with researchers at UBC (in-situ visualization), UQAM (climate model visualization), UVic (climate CFD, optimization), UofA (engineering CFD)
  - happy to help with any visualization projects
- Near-annual *Visualize This!* contests
  - nationally since 2016 <https://ccvis.netlify.app/contests>
  - in 2021 hosted the international IEEE Vis Contest <https://scivis2021.netlify.app>
  - currently on pause: we have data and technical expertise; seeking collaborative outreach partners
- Subscribe to our newsletter <https://training.westdri.ca/contact>



# Opportunities for you

- We offer training on any research computing topic in any format mentioned earlier
  - we constantly develop new materials, can look at new topics and specific computational problems
  - couple of restrictions
    1. in general, we do not cover commercial (not open-source) tools
    2. we probably do not have specialized knowledge in your field  $\Rightarrow$  with few exceptions, cannot teach domain-specific topics & tools, but we could develop a joint course with you (domain knowledge from you and computational expertise from us)
- We can teach guest lectures or collaborate on graduate-level university courses
- Email [training@westdri.ca](mailto:training@westdri.ca) to request a workshop or suggest a topic
- All our training is free for Canadian academic researchers

# Other

- For other RCG offerings, check <https://docs.alliancecan.ca> | Services
  - HPC
  - cloud
  - database servers
  - Globus file transfer, Nextcloud storage
  - research software
- Special interest in numerical methods  $\Rightarrow$  collaborate on courses and materials?
  - computational fluid dynamics, conservation methods, Riemann solvers
  - numerical radiative transfer (all regimes from free streaming to diffusion, 6D)
  - serial / parallel ODE and PDE solvers, stiff systems and implicit methods
  - computational linear algebra
  - adaptive techniques
  - ...