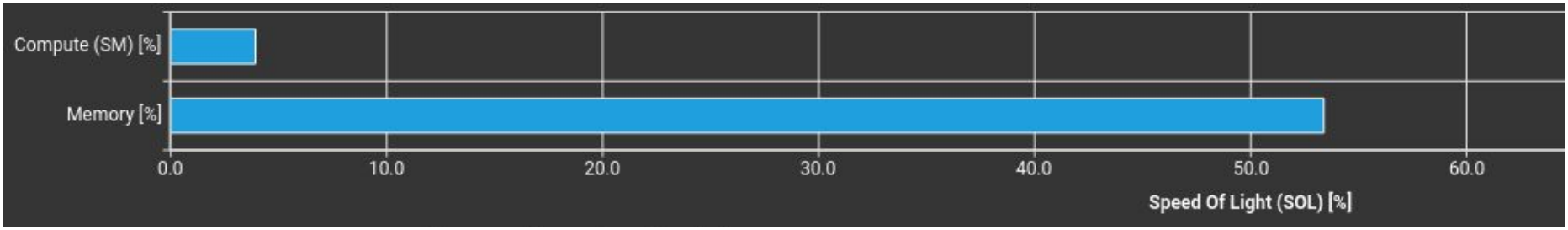
Jordi Manyer (Monash)
Alberto F. Martin (ANU)

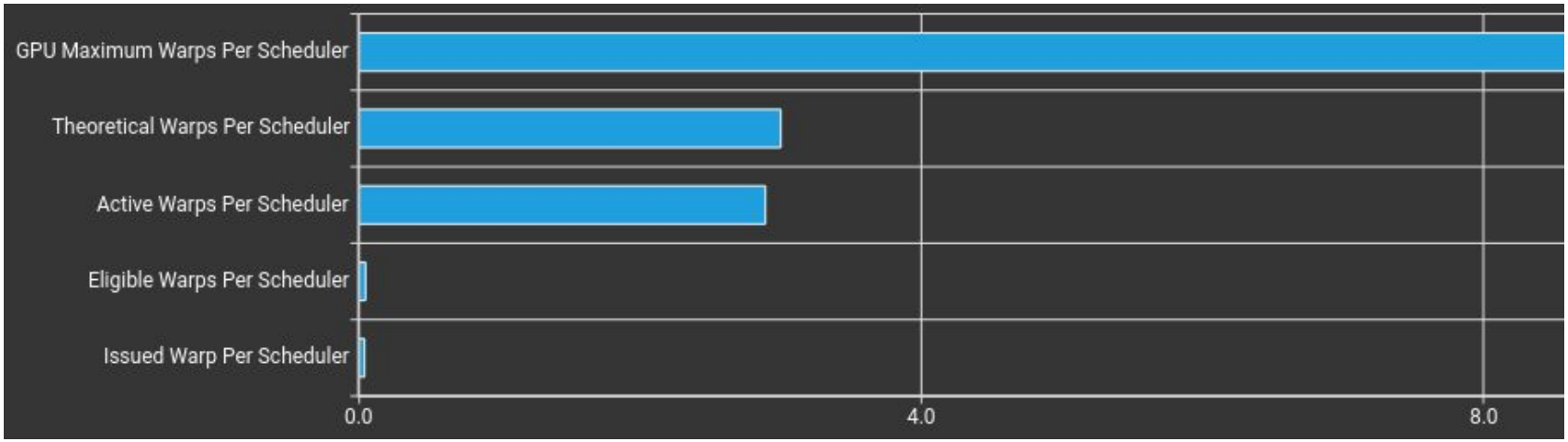
Mentors:
Yue Sun (NCI)
Mark Kittisopikul (JuliaLang)
Mohamed Tarek (Pumas-AI)
James Foster (CSIRO)

- Finite-Element code
- Application: Matrix-Vector product for Finite-Element-based linear systems.
- Method: Sum-Factorization-based matrix-free methods.
- Motif: Small dense tensor index contractions

Evolution and Strategy

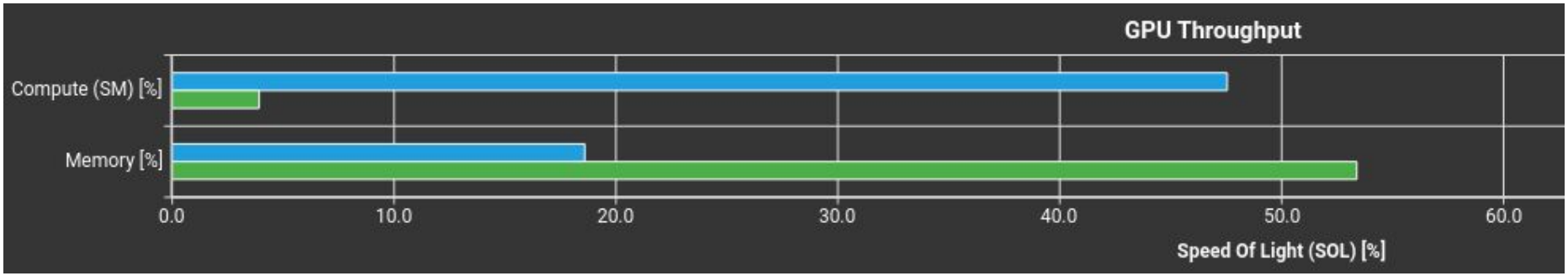
v0 - One thread per mesh cell, parallelizing over cells.





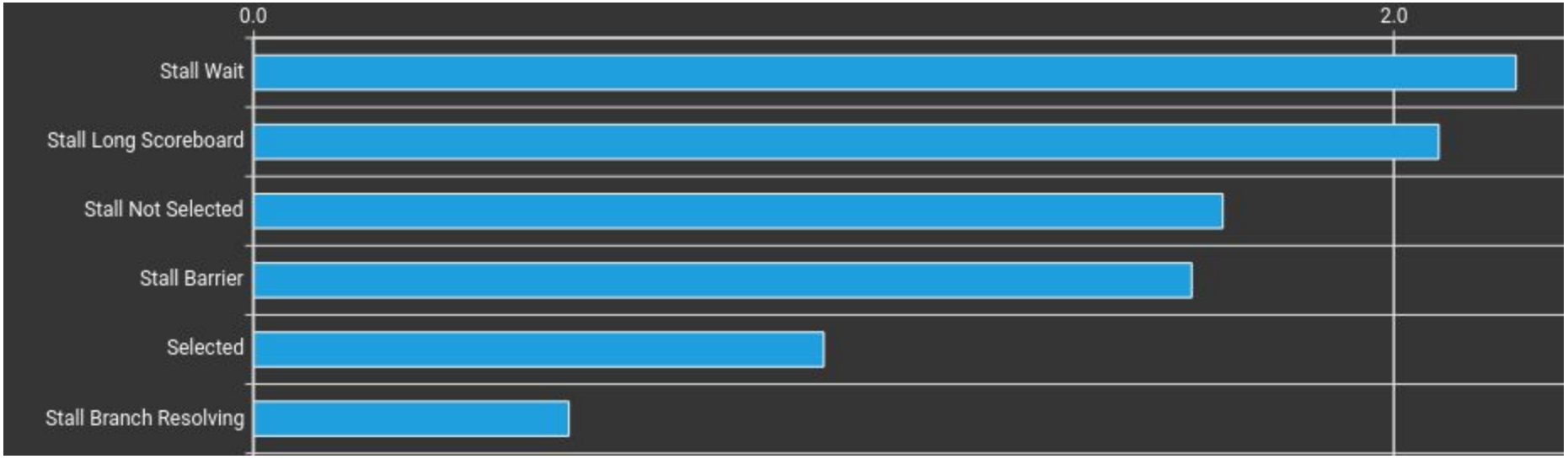
Evolution and Strategy

v1 - Added a second layer of parallelism, using multiple threads per cell to increase throughput while reducing register usage.



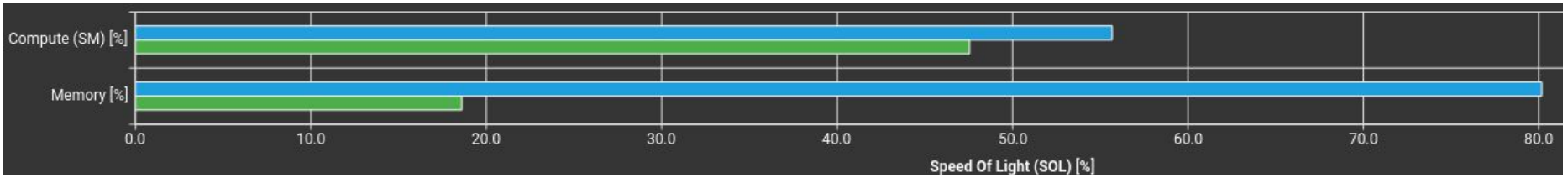
Evolution and Strategy

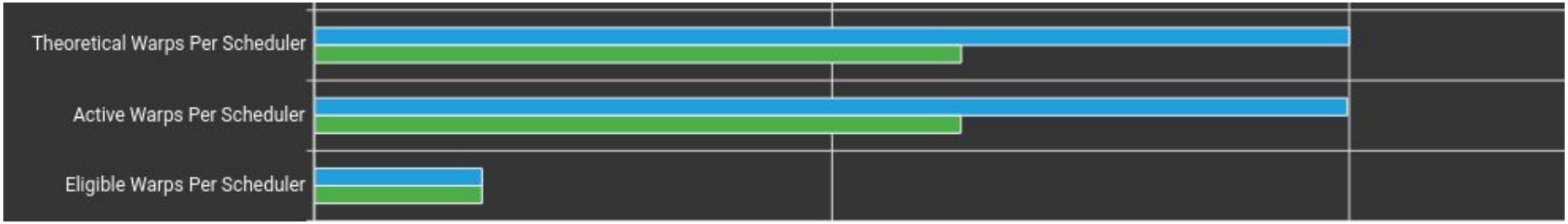
v1 - Added a second layer of parallelism, using multiple threads per cell to increase throughput while reducing register usage.



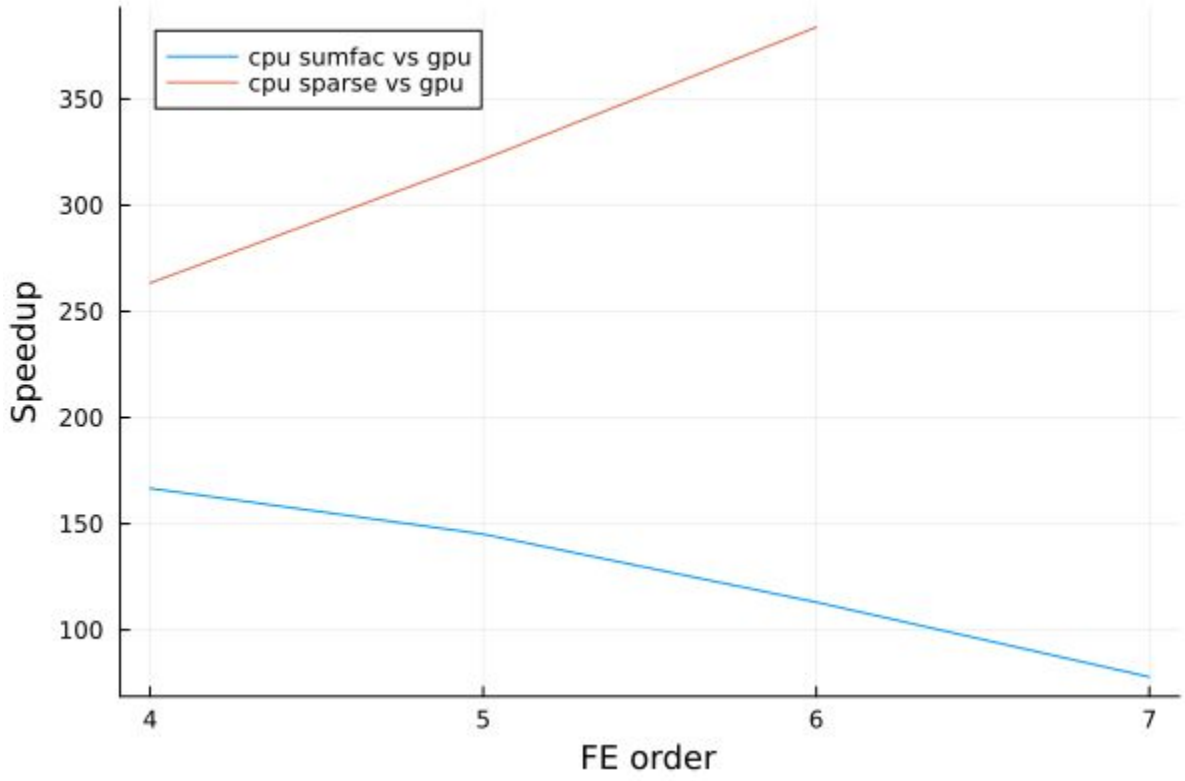
Evolution and Strategy

v2 - Moved part of the memory to shared & constant memory.





Results and Final Profile



Energy Efficiency

INPUTS	
# CPU Cores	1
# GPUs (A100)	1
Application Speedup	100.0x

Node Replacement	6.3x

GPU NODE POWER SAVINGS			
	AMD Dual Rome 7742	**8x A100 80GB SXM4**	**Power Savings**
Compute Power (W)	6,875	6,500	375
Networking Power (W)	290	93	197
Total Power (W)	**7,165**	**6,593**	572

Node Power efficiency	**1.1x**

ANNUAL ENERGY SAVINGS PER GPU NODE			
	AMD Dual Rome 7742	**8x A100 80GB SXM4**	**Power Savings**
Compute Power (kWh/year)	60,225	56,940	3,285
Networking Power (kWh/year)	2,542	814	1,729
Total Power (kWh/year)	**62,767**	**57,754**	**5,014**

$/kWh	$	0.34
Annual Cost Savings	$	1,704.71
3-year Cost Savings	$	5,114.12

Metric Tons of CO2	4
Gasoline Cars Driven for 1 year	1
Seedlings Trees grown for 10 years	59

(source: Link)

What problems have you encountered?

- Issues to get NVIDIA NSight Compute running on Julia.
- Tool bugs: NVIDIA NSight Compute UI crashes my linux local computer at times

Was it worth it?

- Was this worth it? Yes!
- Will you continue development?
  - Generalize/package what we have done here.
  - Continue exploring memory management to enhance performance.
- What sustained resources/support will be critical for your work after the event?

  - Support from the Julia CUDA community
  - Tutorials on kernel optimization.

Please use 100 words to summarize your team's achievements during this Hackathon

We successfully managed to implement our algorithm for GPUs, and to do it in a good enough way to speed up our application.

More importantly, this was our first dive into the world of GPUs. We feel the knowledge we have gained here will serve us incredibly well in the future.

There is still a lot of room for improvement, and we will definitely continue iterate the code towards better performance.

PROMOTING YOUR WORK: AVAILABLE OPPORTUNITIES

- **Papers and Talks:** Please acknowledge the Open Hackathons program and OpenACC Organization in any planned or upcoming papers, presentations, or talks.

  "This work was completed in part at the [Event name], part of the Open Hackathons program. The authors would like to acknowledge OpenACC-Standard.org for their support."
- **Social Media Support:** Please feel free to promote your participation across your social media channels. Tag **@OpenACCorg** and **#OpenHackathons** and we are happy to amplify.
- **Blogs and Technical Write-ups**: Create a blog post or technical article that highlights the work being done and results achieved.
- **Quotes and Testimonials:** Highlight your quote or feedback on our channels (i.e. social, website, etc.).

***Please reach out to Izumi Barker (ibarker@nvidia.com) to discuss marketing options and opportunities.**