

Wind Tunnels for ASI

Arun Darlie Koshy (Sr Engineer, HPE)

Jean Marie-Verdun (Sr Distinguished Technologist, HPE)



October 13-16, 2025
San Jose, California



Hello World

- Key themes
 - We are far from AGI (& arguably ASI)
 - Tests are key for production AI/ML models
 - Private/Personal models long-haul may win
- Jumpstart
 - Clone <https://github.com/opencomputeproject/ocp-diag-windtunnel>
 - Setup on preferred gear

Cracks in Foundation models

“Wilbur Wright could pack a punch of enormous rhetorical power in.. a few words.. with these five.. “Unfortunately, the wind usually blows.” .. Wright had just gutted the experimental strategy Samuel Langley had been following for more than ten years.” [1]

- hallucinations, jail breaks & indirect prompt injection – have no ‘fix’ [2]
- power/compute consumption indicates trouble (human brain ~ 20 watts) [3]
- scaling will continue to pay-off yet not yield AGI [4]
- math : curve-fitting vs causality vs...

Test for intelligence and bubbles

“In bubbles, it is often a good trade to make the “wrong” directional bet. Level one thinking would have you buying puts.. the better bet is to sell puts. The pros are statistically passive.” [5]

(And Validation i.e. OCP T&V)

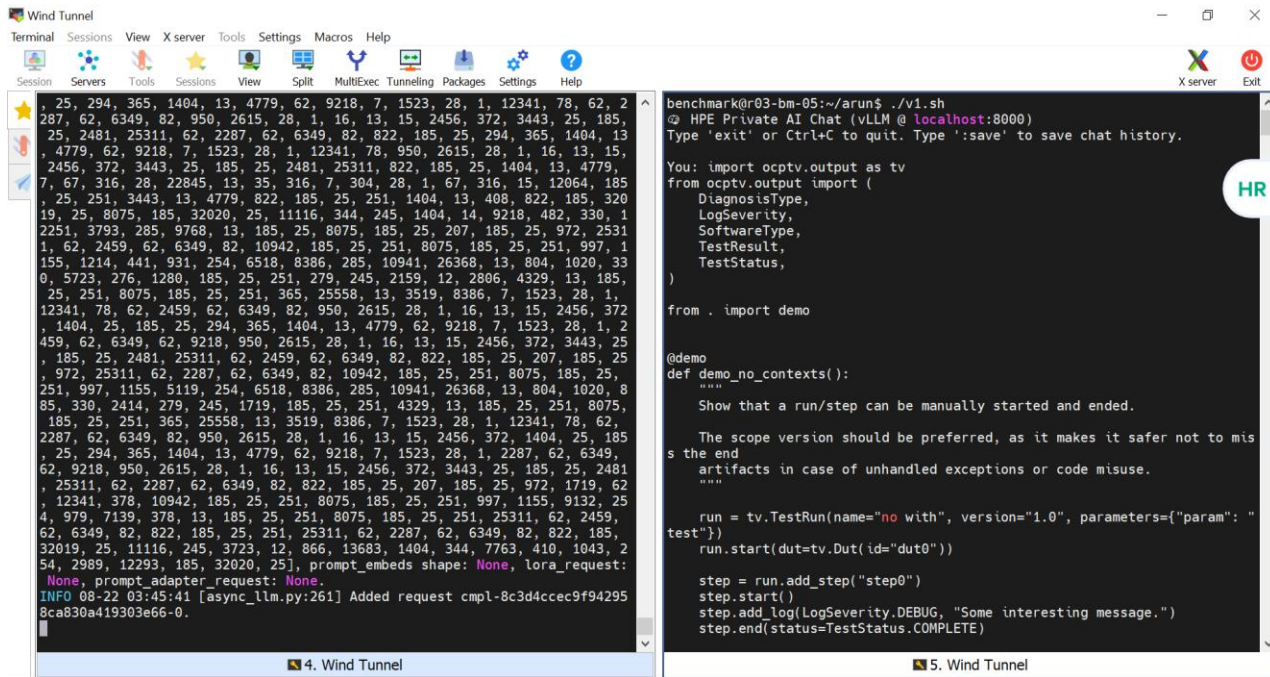
- Private agents tailored for contributions and tests for
 1. <https://github.com/opencomputeproject/ocp-diag-core>
 2. <https://github.com/opencomputeproject/OSF-OSFCI>

(Big-picture)

- LLMs broke the Turing test [6]
- <https://arcprize.org>
- <https://www.claymath.org/millennium-problems/> machine solves from ground-up with thesis defense to peer human experts

Demo

"For a successful technology, reality must take precedence over public relations, for nature cannot be fooled." [7]



4. Wind Tunnel

```
25, 294, 365, 1404, 13, 4779, 62, 9218, 7, 1523, 28, 1, 12341, 78, 62, 2
287, 62, 6349, 82, 950, 2615, 28, 1, 16, 13, 15, 2456, 372, 3443, 25, 185,
25, 2481, 25311, 62, 2287, 62, 6349, 82, 822, 185, 25, 294, 365, 1404, 13
4779, 62, 9218, 7, 1523, 28, 1, 12341, 78, 950, 2615, 28, 1, 16, 13, 15,
2456, 372, 3443, 25, 185, 25, 2481, 25311, 822, 185, 25, 1404, 13, 4779,
7, 67, 316, 28, 22845, 13, 35, 316, 7, 304, 28, 1, 67, 316, 15, 12064, 185
25, 251, 3443, 13, 4779, 822, 185, 25, 251, 1404, 13, 408, 822, 185, 320
19, 25, 8075, 185, 32020, 25, 11116, 344, 245, 1404, 14, 9218, 482, 330, 1
2251, 3793, 285, 9768, 13, 185, 25, 8075, 185, 25, 207, 185, 25, 972, 2531
1, 62, 2459, 62, 6349, 82, 10942, 185, 25, 251, 8075, 185, 25, 251, 997, 1
155, 1214, 441, 931, 254, 6518, 8386, 285, 10941, 26368, 13, 804, 1020, 33
0, 5723, 276, 1280, 185, 25, 251, 279, 245, 2159, 12, 2806, 4329, 13, 185,
25, 251, 8075, 185, 25, 251, 365, 25558, 13, 3519, 8386, 7, 1523, 28, 1,
12341, 78, 62, 2459, 62, 6349, 82, 950, 2615, 28, 1, 16, 13, 15, 2456, 372
1404, 25, 185, 25, 294, 365, 1404, 13, 4779, 62, 9218, 7, 1523, 28, 1, 2
459, 62, 6349, 62, 9218, 950, 2615, 28, 1, 16, 13, 15, 2456, 372, 3443, 25
185, 25, 2481, 25311, 62, 2459, 62, 6349, 82, 822, 185, 25, 207, 185, 25
972, 25311, 62, 2287, 62, 6349, 82, 10942, 185, 25, 251, 8075, 185, 25,
251, 997, 1155, 5119, 254, 6518, 8386, 285, 10941, 26368, 13, 804, 1020, 8
85, 330, 2414, 279, 245, 1719, 185, 25, 251, 4329, 13, 185, 25, 251, 8075,
185, 25, 251, 365, 25558, 13, 3519, 8386, 7, 1523, 28, 1, 12341, 78, 62,
2287, 62, 6349, 82, 950, 2615, 28, 1, 16, 13, 15, 2456, 372, 1404, 25, 185
25, 294, 365, 1404, 13, 4779, 62, 9218, 7, 1523, 28, 1, 2287, 62, 6349,
62, 9218, 950, 2615, 28, 1, 16, 13, 15, 2456, 372, 3443, 25, 185, 25, 2481
25311, 62, 2287, 62, 6349, 82, 822, 185, 25, 207, 185, 25, 972, 1719, 62
12341, 378, 10942, 185, 25, 251, 8075, 185, 25, 251, 997, 1155, 9132, 25
4, 979, 7139, 378, 13, 185, 25, 251, 8075, 185, 25, 251, 25311, 62, 2459,
62, 6349, 82, 822, 185, 25, 251, 25311, 62, 2287, 62, 6349, 82, 822, 185,
32019, 25, 11116, 245, 3723, 12, 866, 13683, 1404, 344, 7763, 410, 1043, 2
54, 2989, 12293, 185, 32020, 25], prompt_embeds shape: None, lora_request:
None, prompt_adapter_request: None.
INFO 08-22 03:45:41 [async_llm.py:261] Added request cml-8c3d4ccec9f94295
8ca830a419303e66-0.
```

5. Wind Tunnel

```
benchmark@r03-bm-05:~/arun$ ./v1.sh
@ HPE Private AI Chat (vLLM @ localhost:8000)
Type 'exit' or Ctrl+C to quit. Type ':save' to save chat history.

You: import ocptv.output as tv
from ocptv.output import (
    DiagnosisType,
    LogSeverity,
    SoftwareType,
    TestResult,
    TestStatus,
)

from . import demo

@demo
def demo_no_contexts():
    """
    Show that a run/step can be manually started and ended.

    The scope version should be preferred, as it makes it safer not to mis
s the end
    artifacts in case of unhandled exceptions or code misuse.
    """
    run = tv.TestRun(name="no with", version="1.0", parameters={"param": "
test"})
    run.start(dut=tv.Dut(id="dut0"))

    step = run.add_step("step0")
    step.start()
    step.add_log(LogSeverity.DEBUG, "Some interesting message.")
    step.end(status=TestStatus.COMPLETE)
```

Call to action

- Contribute frameworks, tests, agents for OCP Wind Tunnel
- Broaden beyond T&V context / scope
- Pool resources together for the community
- Develop highly specialized set of tests for ML/AI agent contributions
- <https://github.com/opencomputeproject/ocp-diag-windtunnel>

References

- [1] – James Tobin. To Conquer the Air: The Wright Brothers and the Great Race for Flight.
- [2] - Mark Russinovich, Ahmed Salem, Santiago Zanella-Béguelin, and Yonatan Zunger <https://cacm.acm.org/practice/the-price-of-intelligence/>
- [3] - Chris Eliasmith. How to Build a Brain (Oxford Series on Cognitive Models and Architectures)
- [4] - Yann LeCun, François Chollet (and many others) in varied public settings
- [5] - Andrew Mack, Euan Sinclair, Martin P. O'Connell (context: trading derivatives)
- [6] - Celeste Biever, <https://www.nature.com/articles/d41586-023-02361-7>
- [7] - R. P. Feynman, <https://www.nasa.gov/history/rogersrep/v2appf.htm>

Thank You!