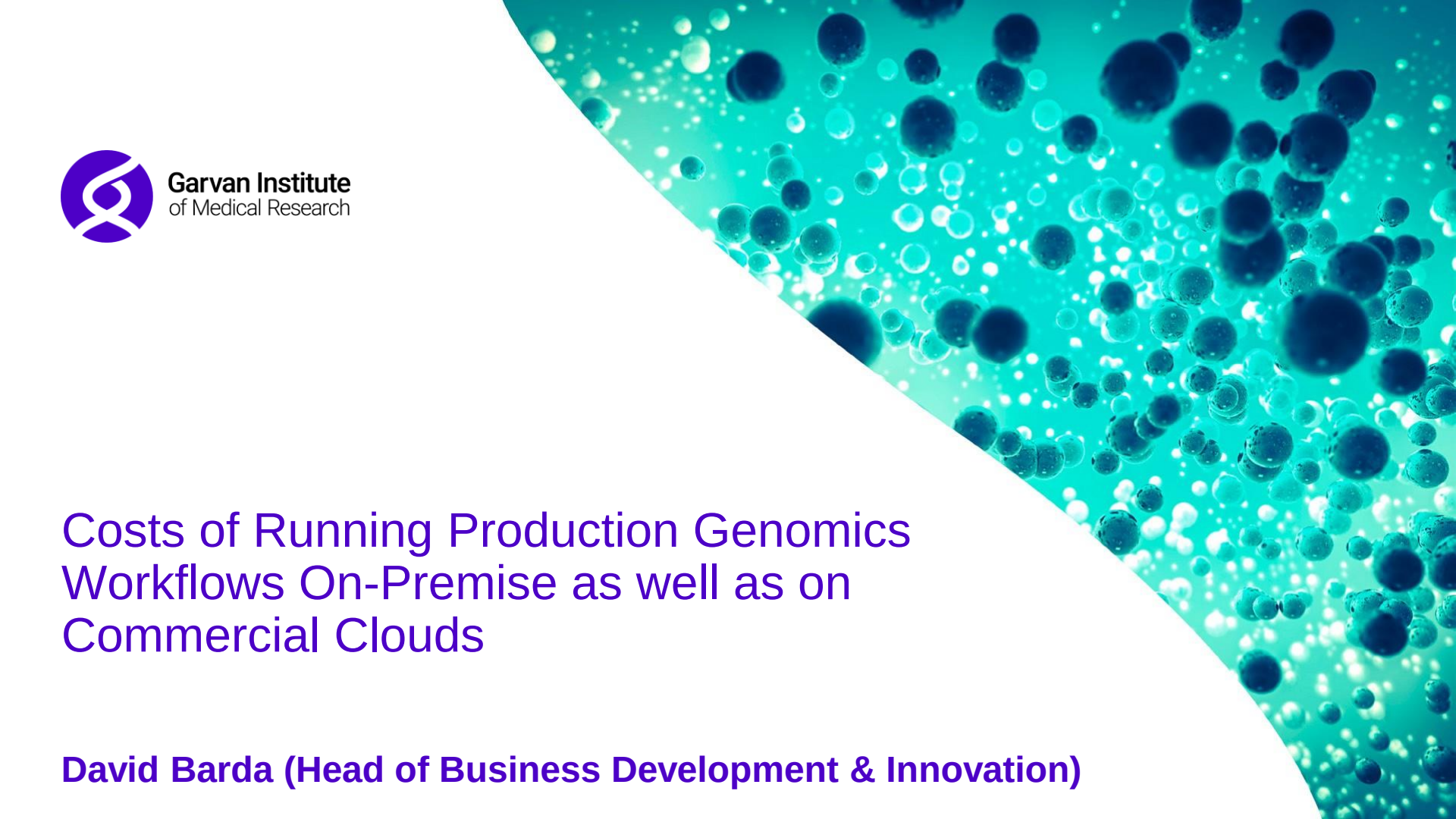


Garvan Institute
of Medical Research

Costs of Running Production Genomics Workflows On-Premise as well as on Commercial Clouds

David Barda (Head of Business Development & Innovation)

Questions we set out to answer

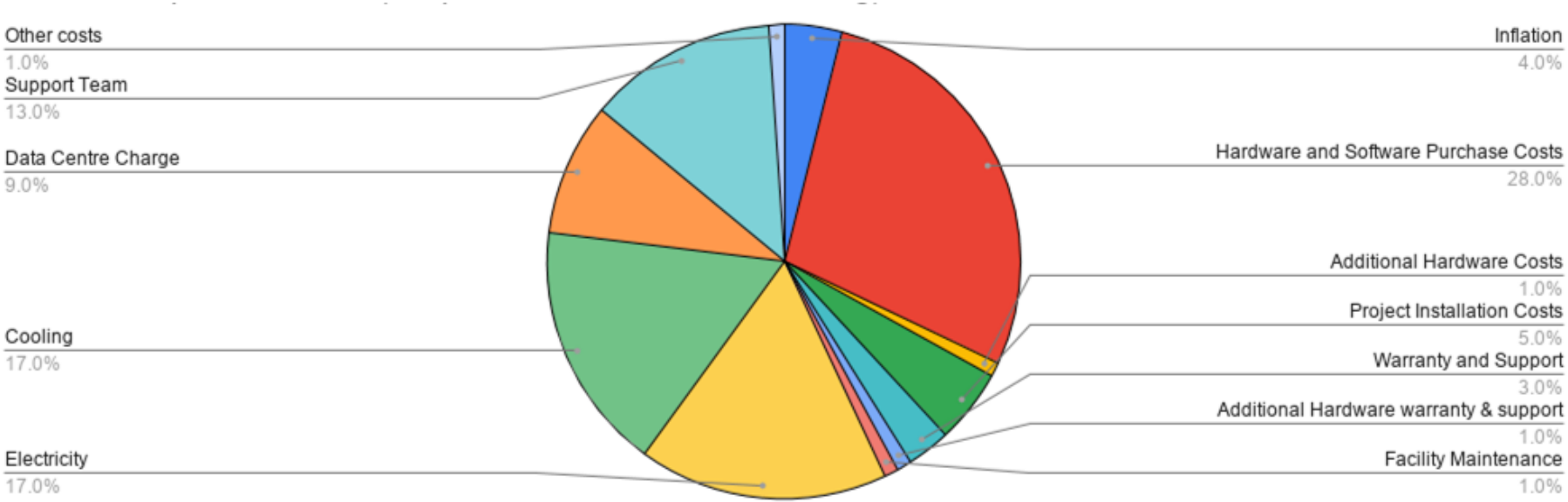
1. Since we built our 5000 core on-premise computational infrastructure using grants - what does it cost us to run these workflows at Garvan, and how do these costs compare when running the identical workflows on AWS, Azure and the NCI?
2. What are the underlying costs of running these analyses that include, power, maintenance, staffing etc?

Things to think about

- Capital versus Operational expenditure
- Data egress
- Central grant/funding agencies
- Ease of funding - At Garvan budgets are allocated on an annual basis. If research computing costs (HPC and others) is to move into Opex, how do Garvan researchers secure long-term funding commitments to insulate themselves from the vagaries of the annual budgeting cycle?
- How can researchers more accurately predict their cost exposure?
- Cloud is a new provisioning and usage model for the researchers. There is now an increased (overwhelming) level of choice. What tools should be developed to help users decide on what services to use?

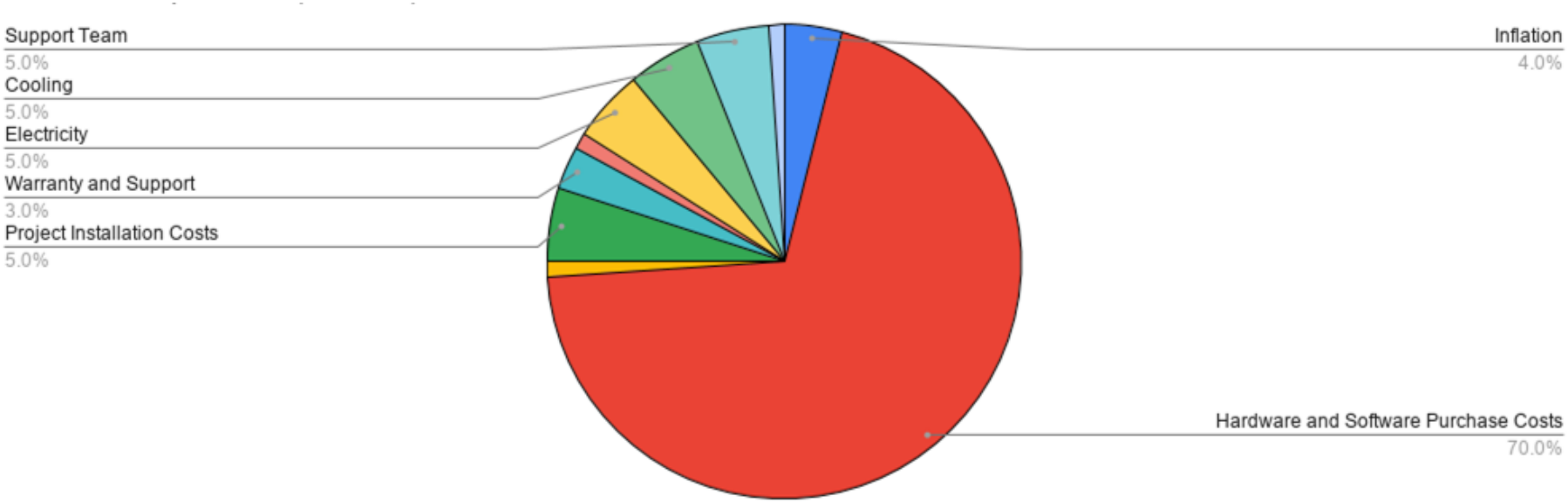
Lifetime on-premise costs (Red Oak Consulting)

Garvan Institute of Medical Research



Red Oak Consulting
Expert advice, exceptional delivery

Lifetime on-premise costs (Garvan)



Whole Genome Workflow Costs

Garvan Institute of Medical Research

Whole Genome Workflow	AWS Cost	Azure	NCI	Garvan
Wall time (hours)	21	22	12	11
CPU cost ($AUD)	$10	$14	$17	$3 (estimated)
Egress	$30	$28	zero	zero
Total	$40-$45	$43-$46	$17	$3 (estimated)

Final Thoughts

Garvan is cheapest - but that's probably because of:

- the number of unknowns
- we have zero redundancy of staff - making us very vulnerable
- we only have hardware support and pay no software support
- we may end up moving to cloud as we struggle to compete with giant tech companies to recruit engineers.