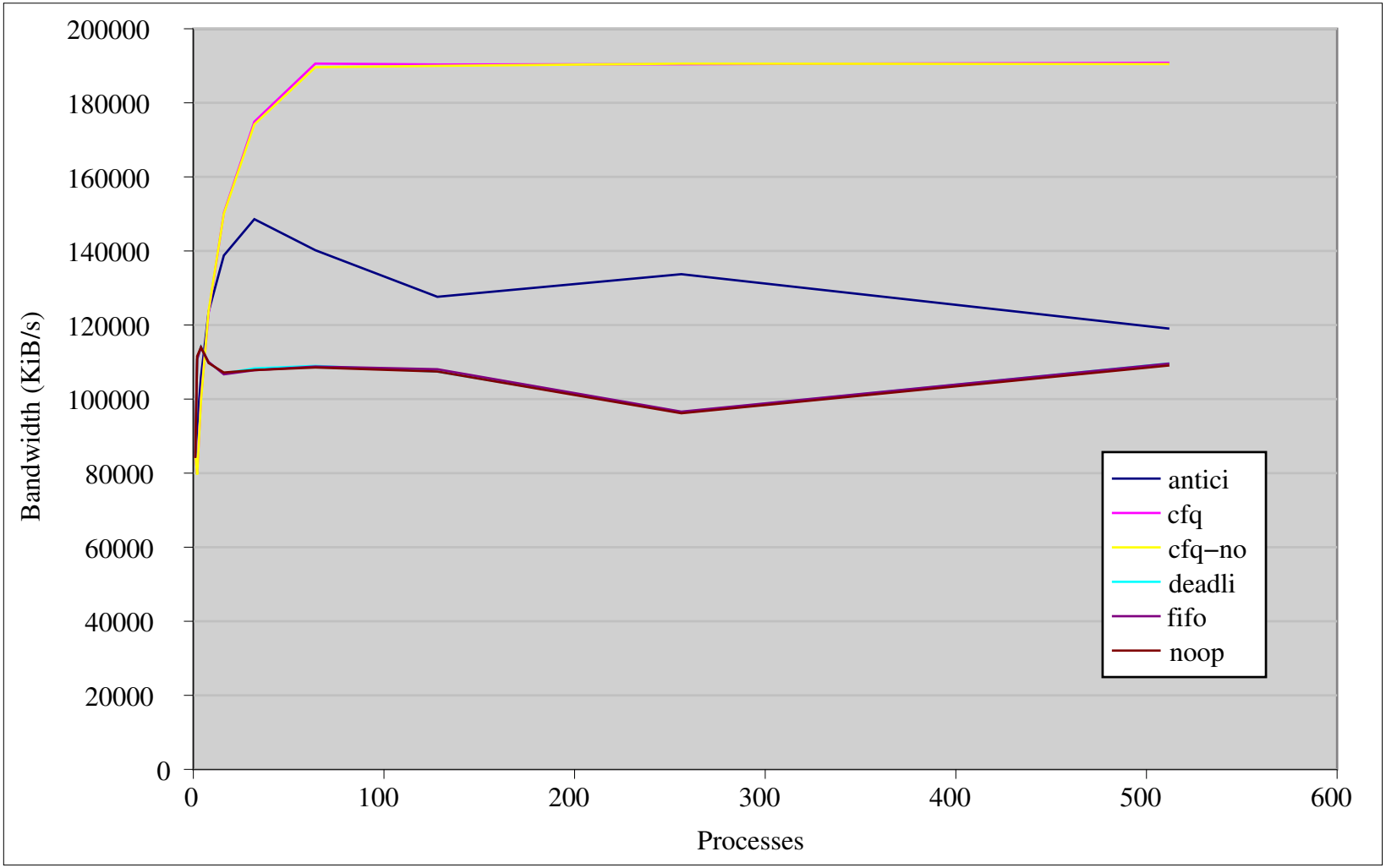
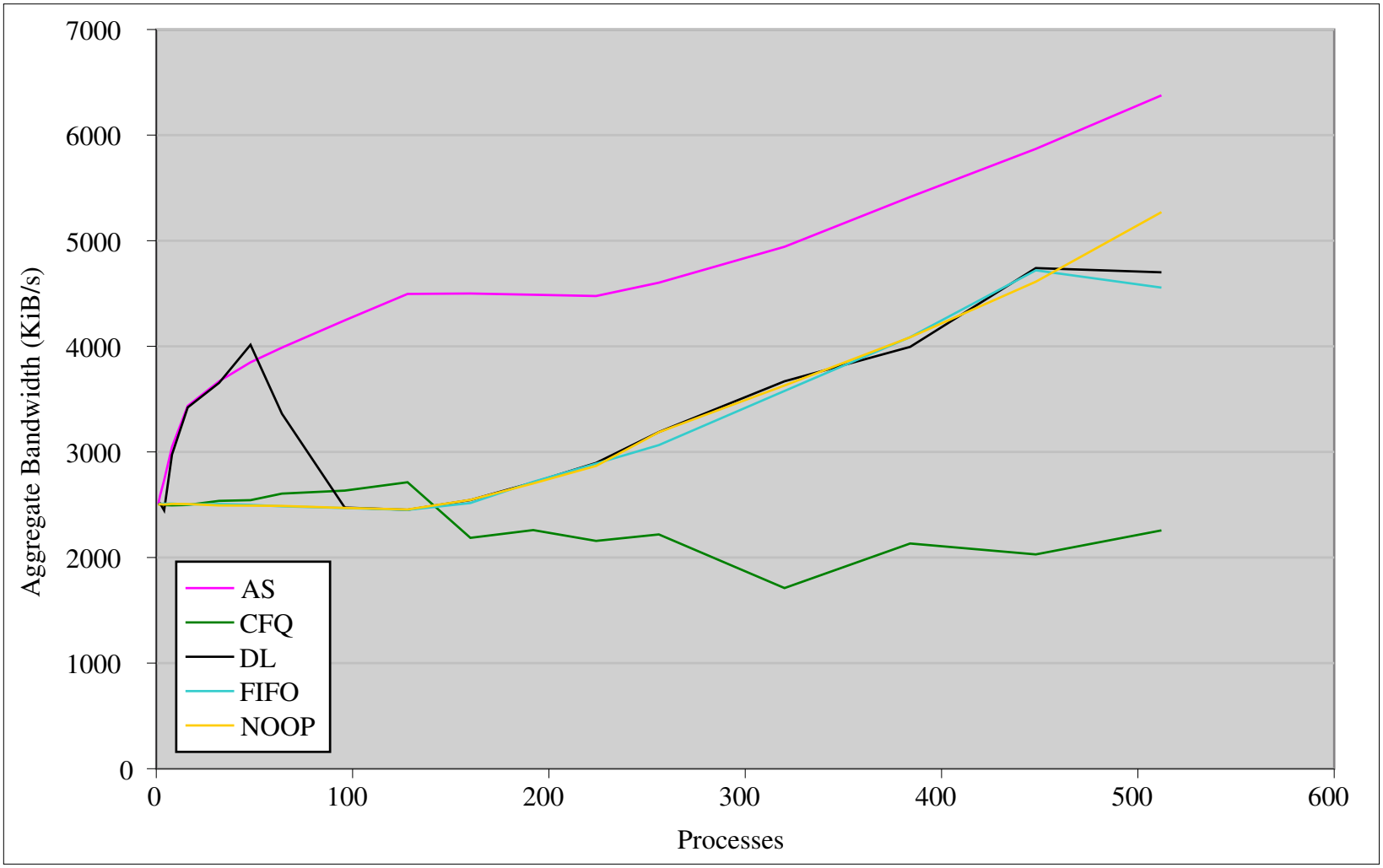
Aaron Carroll, Joshua Root
{aaronc,jmr}@gelato.unsw.edu.au

Supported by UNSW, NICTA, HP and Google through the Gelato Federation

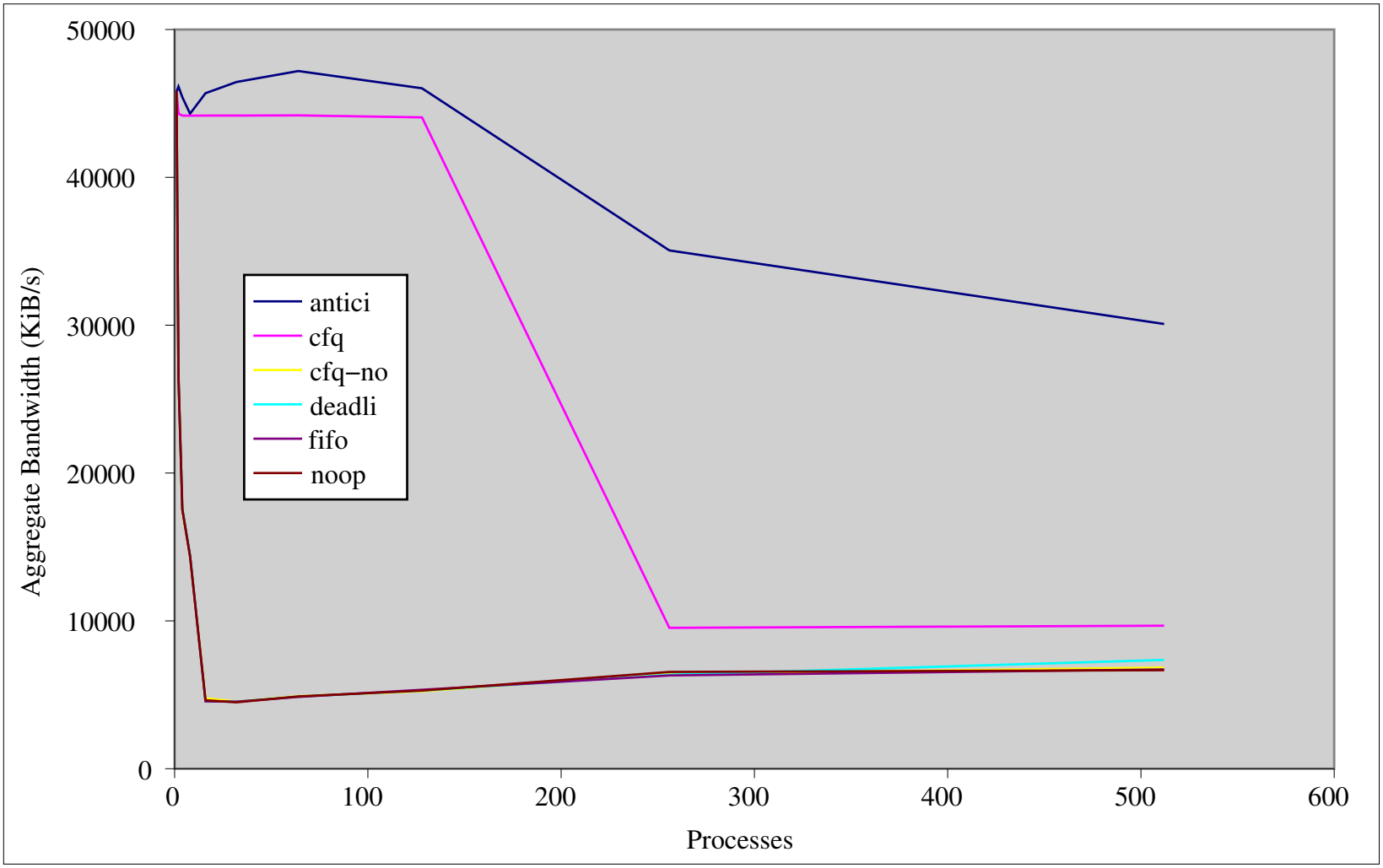
RAID 0, sequential asynchronous



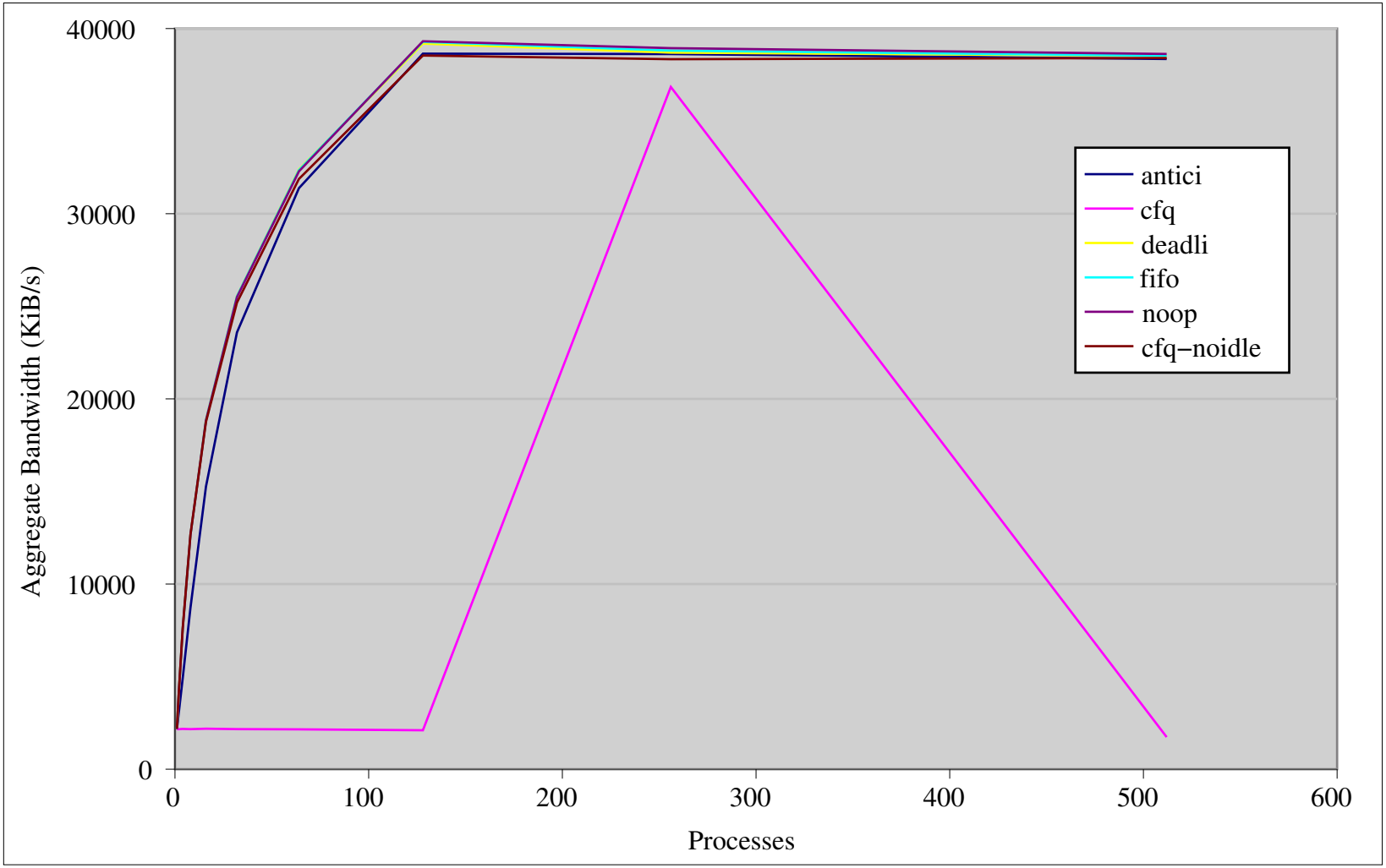
Single disk, random



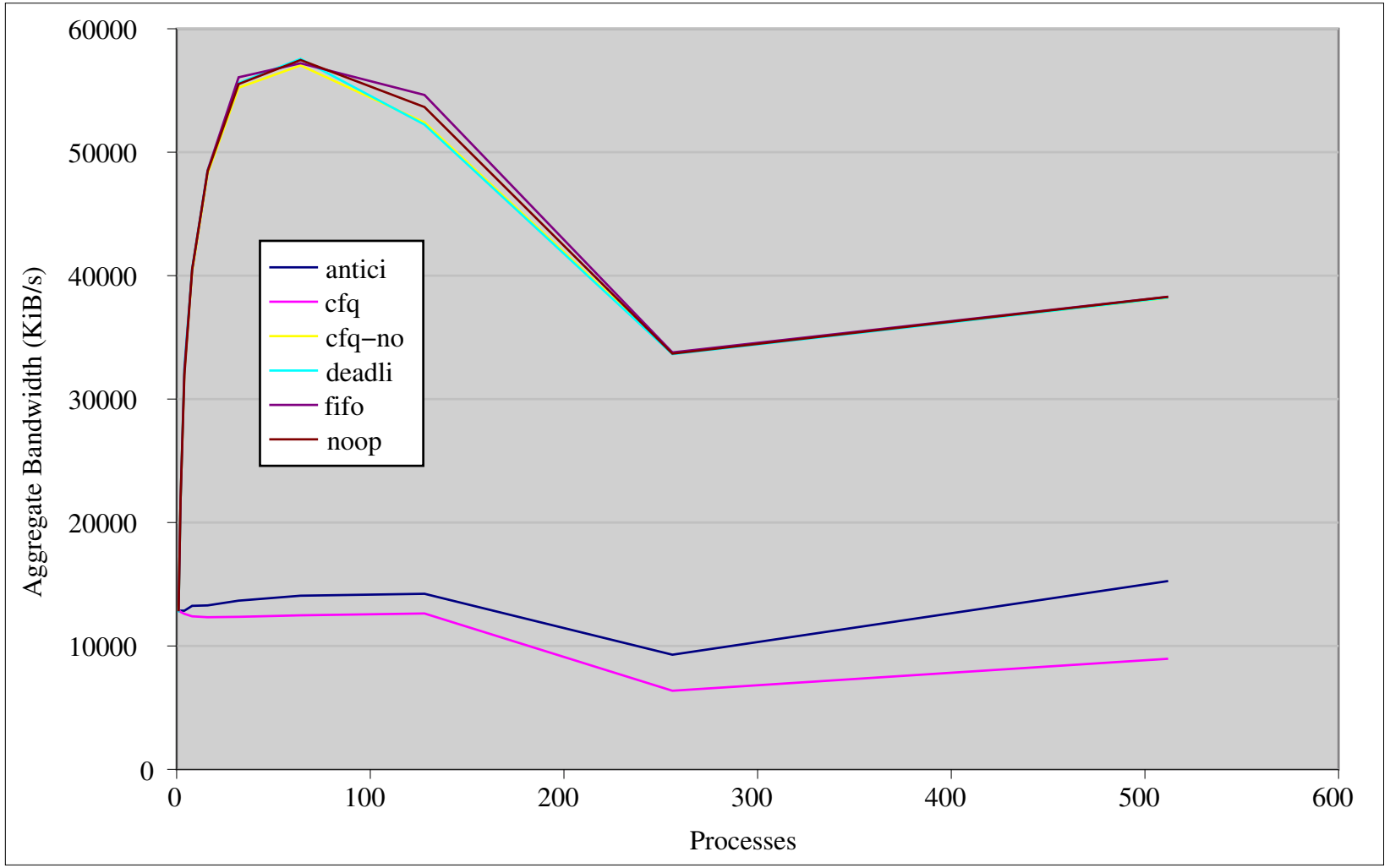
Single disk, 1MiB sequential



10 disk RAID 0, random



10 disk RAID 0, 1MiB sequential



What Parameters are Useful?

- Queue depth
- Underlying storage device
- RAID topology
- ...?

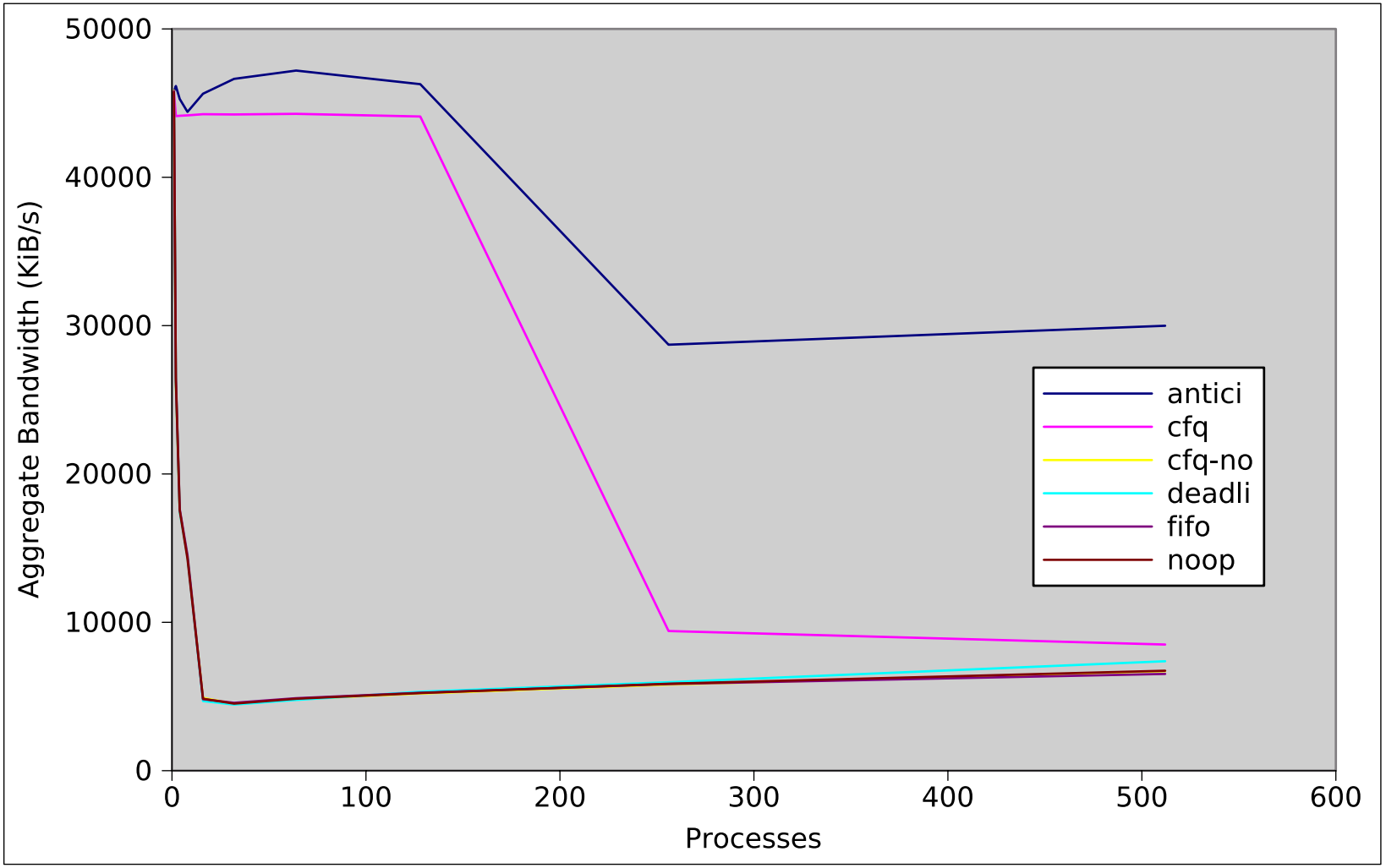
How To Discover the Parameters?

- ➜ User input
- ➜ Measurements
- ➜ Ask other layers

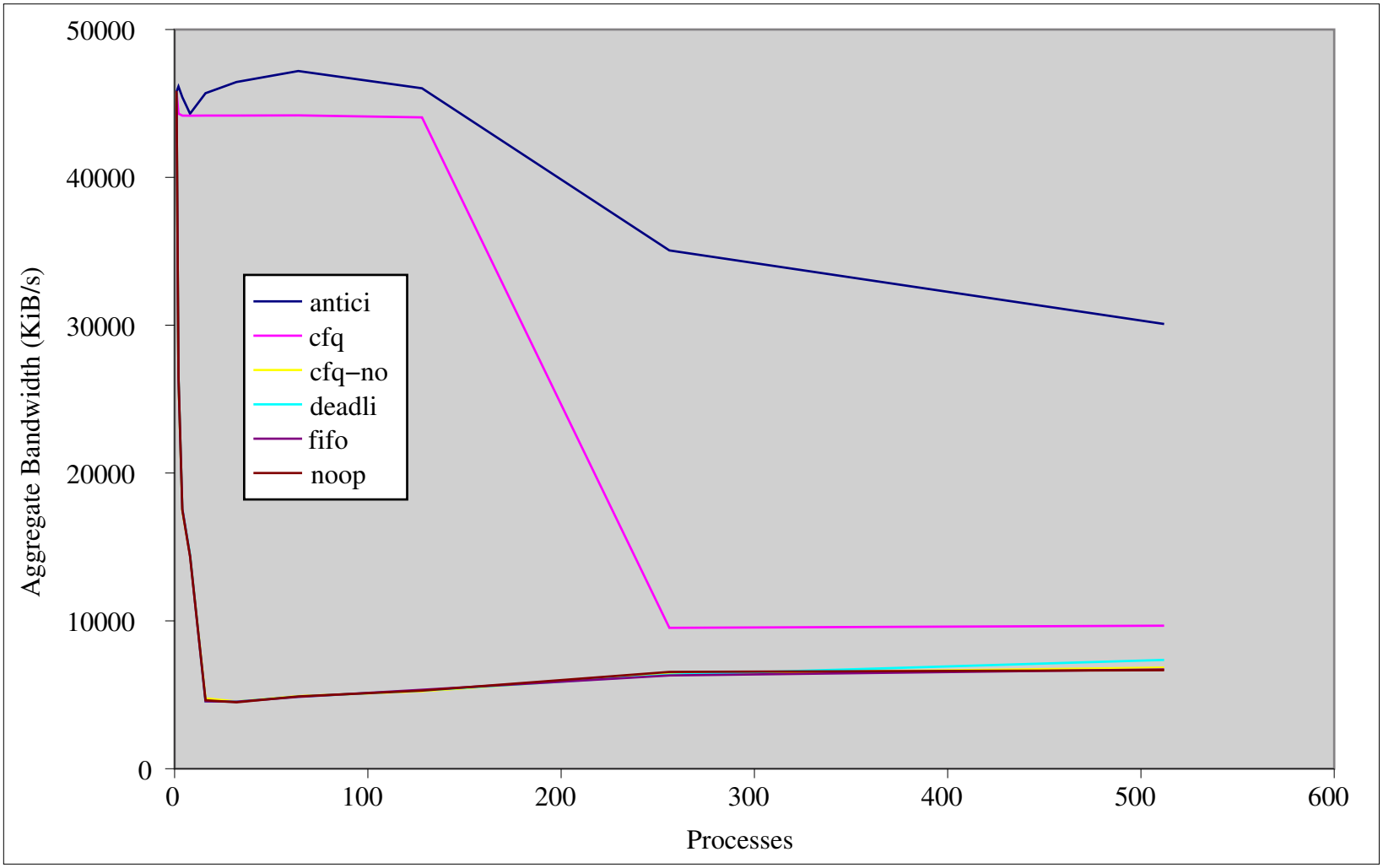
QUEUE DEPTH

- ➜ Measurements or API
- ➜ Measurement a bit suspicious
- ➜ ... but why is queue depth useful?
- ➜ For anticipation, it’s not
- ➜ What we really want is device type

Single disk, 1MiB sequential, TCQ depth 64



Single disk, 1MiB sequential, TCQ depth 1



DEVICE TYPE

- ➜ Ask the driver
- ➜ What is a suitable level of abstraction?
  - ① Random access
  - ② True parallelism
  - ③ Just device type (H/W RAID, SSD, etc)

RAID Topology

- Stripe boundaries and width
- Request → disk mapping
- Take measurements
- Per-spindle scheduling
- Better resource accounting

AUTOMATIC TUNING

- ➜ Seek profile → anticipation expiry
- ➜ Read/write ratio + deadlines