How to Write an HPC Storage RFP

CUG BOF

May 8, 2007

Peter Rigsbee, Cray

CRAY
THE SUPERCOMPUTER COMPANY

Are HPC Customers Happy with their Storage?

- IDC HPC User Forum surveyed members last summer
  - Reported results in Denver in September

- Q1. Do you have applications constrained by I/O?
  - Yes: 92%
- Q4. How well does your current storage management software meet your requirements?
  - 2.66 (scale of 1-5); only 36% answered 4 or 5
- Q5. Channel bandwidth: how acceptable is your current channel bandwidth level?
  - 2.81 (scale of 1-5); only 31% answered 4 or 5
- Q9a. Please rate the reliability level on your current storage systems.
  - 3.31 (scale of 1-5); only 46% answered 4 or 5

Why Aren't They Happy?

- The typical supercomputer RFP and procurement process doesn't encourage vendors to propose good storage solutions
- Supercomputer RFPs are focused on supercomputers, not on storage
- Message from customer is often:
  - We expect to get at least xx TFLOPS
  - We have a budget of $yy M

- Disclaimer:
  - This is the "typical" supercomputer RFP – there are exceptions to this
  - (and these are probably the customers who responded with 4s and 5s)

What Do You Find in Supercomputer RFPs?

- Reference to “storage” as a single entity
  - Typically referring to this as “scratch space”, or not being specific at all about its purpose
- Two quantitative storage requirements: capacity and aggregate bandwidth
  - Stated in absolute terms, or relative to peak TFLOPS and/or memory size
- Two qualitative storage requirements: reliability and management
  - “The storage system must be fully resilient and redundant”
  - “The storage system must be easily managed”
- Functional requirements for the filesystem
  - POSIX compliance, ACLs
- Maybe some vague requirements or questions like:
  - “explain how disk array performance is impacted by fragmentation”
  - “explain how disk array performance is impacted by a RAID rebuild”
  - “the storage must be upgradeable in terms of capacity and bandwidth”
- At most one I/O benchmark, often something like IOzone, with a small overall weighting, and no restrictions on system setup

What Don't You Find in Supercomputer RFPs?

- Any evaluation credit for exceeding storage requirements
  - Occasionally will see "desirables" with extra credit for more capacity or bandwidth
- Any quantitative requirements on reliability of storage
- Benchmarks reflecting other or mixed usage
  - Small files
  - Small or random I/O requests
  - Metadata performance

What's the Result?

- Vendors try to ensure proposals meet all RFP requirements, and exceed requirements where this gets the biggest payoff
- Typical supercomputer RFPs give more credit to providing:
  - More peak TFLOPS
  - More sustained TFLOPS
  - More memory per FLOP
  - Lower price
- Than they do for any extra or better storage solution

- Vendors are motivated to propose the minimum storage solution, that just meets the RFP requirements, at the lowest cost

But What if That's All We Need!

- It is possible this matches your requirements
  - You just need scratch space – all permanent files are elsewhere
  - All your applications are highly scalable, and only do large block sequential I/O that needs to be striped across lots of disk arrays
  - You routinely delete your scratch files (maybe perhaps reformat your arrays regularly) so there's minimal fragmentation
  - Your users don't use any files in the scratch storage for compiles or anything else that does small I/Os (or you don't care how long it takes them)
- If so, you're all set!
- But we suspect this isn't the case for everyone…

Good RFPs

- There are some RFPs that don't have these faults
- What do these customers do that's different?
- Some high-level differences:
  - Storage sections clearly written by people who know storage
  - Often these are storage-only RFPs for center-wide storage
    - Storage is the focus – not the supercomputer
  - Frequently these are weather/climate sites – who seem to have greater concerns about data management and long-term data availability than other HPC customers
- Some lower-level differences:
  - Consider different types of storage, and specify requirements for each
  - Specify quantitative requirements for more than capacity and bandwidth
  - Have more benchmarks
  - Give vendors who propose something better, credit for doing so

Summary and Discussion

- The Problem:
  - Your users will be more productive and happier with good storage solutions
  - Current supercomputer procurement practices don't lead to you getting good storage solutions
  - Therefore, you're not happy with your storage solutions (as shown by the IDC survey)

- Discussion:
  - Do you agree with these statements and tradeoffs?
  - Can you afford (politically, organizationally) to accept fewer TFLOPS to get better storage?
  - Are there some simple things you can do to improve the situation?
  - Are there things Cray and other vendors can do to help?