File and Storage Tech. research for HPC

Greg Ganger

PDL Director

Storage systems: fun quotes

"I/O certainly has been lagging in the last decade"

- Seymour Cray, 1976

"Also, I/O needs a lot of work"

- David Kuck, 1988

"Scalable I/O is perhaps the most overlooked area of HPC R&D"

- Suggested R&D topics doc for 2005-2009

Carnegie Mellon
Parallel Data Laboratory

So, kudos and thank you!

- We all know storage/FS is a crucial area
  - information is at the heart of it all, and storage is what allows the information to persist…
- Yet, it's been a difficult area to focus on
  - redheaded stepchild in gov't funding
    - and, thus, in research and education
  - this will be first "storage program" I'll see
    - in contrast with networking, architecture, etc.
- Consequence: BIG problems on horizon
  - but way behind on developing solutions…

Current landscape

- Most government funding comes from non-storage-focused programs
  - networking, architecture, distributed systems
- So, too much goes to
  - topical fads: e.g., FSs for P2P WAN overlays
  - small stuff: e.g., local file systems
- Too little goes to
  - almost everything that this community cares about…

Three domains of storage challenges

- High-performance computing
  - scratch: scalability, performance, tune-ability
  - observations/results: reliability, archiving, sharing
- Business IT organizations
  - robustness, automation/mgmt, TCO, disaster tolerance, regulatory compliance, ILM, security
- Personal and consumer electronics
  - new applications, ease of use, data indexing, cost, battery lifetimes, physical robustness, …

Some HPC IO/FS research areas

- Suggested R&D topics doc was excellent
  - I'm not going to repeat it
  - instead, I will highlight and comment on a few points
  - I'm going to focus on longer-range challenges, expecting that others have hit things like standardization stuff already
- #1 = Benchmarks
  - amazing how much a field can be driven by clear goals
  - also, companies would compete based on results
- #1a = Tools and testbeds for research

Some HPC IO/FS research areas

- Emergent issues from scalability
  - Fault tolerance
    - have to expect and work thru more problems
    - awkward for data but harder (and more important) for metadata
  - Performance tuning
    - need evolved and standardized instrumentation plus tools for interpretation
    - both as system size scales and as workload variation grows
  - Diagnosing and correcting problems

Some HPC IO/FS research areas

- Shared storage infrastructures add challenges
  - Performance insulation and predictability
    - e.g., mixing sequential streams can yield random-like perf.
    - such uncertainty can make I/O libraries hard to use properly
  - QoS and performance “guarantees”
  - Security
    - data protection, availability protection
  - Diagnosis of problems
    - If/when other users/apps are concurrent
  - Administration
    - must schedule against many users instead of just one at a time
  - likely must adopt heterogeneity from evolution

A couple thoughts on leverage

- The industry folks care about those things too
  - an opportunity for leverage, for all involved
  - offers the tech. transfer path for the big challenges
    - remember, complexity management cannot be after the fact… must be defined in from the beginning
  - also how HPC gets to tap into COTS products more
    - heard consistently from industry folks that this is important
- Don't think "industry funding that, so we shouldn't"
  - think "industry is interested in that space, we should get involved and expand it to include our needs as well"

Some current CMU/PDL activities

- PASIS: scalable protocols for survivable storage
- Ursa Minor: versatile shared cluster storage
  - directions: performance instrumentation and predictions, performance insulation and QoS, scalable and robust metadata services, pNFS
- Computational databases: science via DBMSs
  - combines with automated storage manager tuning
- Self-* storage: umbrella targeting storage admin
  - much of above, plus automating diagnosis, repair, etc.
- Self-securing devices: intrusion-tolerant systems
- Attribute-based, context-aware file indexing