

IORE: Um benchmark para padrões de E/S complexos de aplicações paralelas de larga escala

Eduardo Camilo Inacio, Mario Antonio Ribeiro Dantas

¹Programa de Pós-Graduação em Ciência da Computação (PPGCC)
Universidade Federal de Santa Catarina (UFSC)
Florianópolis – SC – Brasil

eduardo.camilo@posgrad.ufsc.br, mario.dantas@ufsc.br

Resumo. *Esse trabalho introduz o IORE, um benchmark para reprodução de padrões de acesso de E/S paralela complexos. Seu desenvolvimento objetiva facilitar a caracterização e modelagem do desempenho de sistemas de armazenamento paralelo em infraestruturas de computação de alto desempenho. A longo prazo, essa caracterização suportará uma abordagem de configuração automática para sistemas de armazenamento paralelo, o CALLICRATES.*

Introdução

Identificar configurações otimizadas para o subsistema de armazenamento de plataformas modernas de computação de alto desempenho (CAD) tornou-se um grande desafio, dada a complexidade crescente desses ambientes. Motivada por essa questão, uma abordagem de configuração automática de sistemas de armazenamento paralelo em tempo de submissão de trabalhos, chamada CALLICRATES [Inacio and Dantas 2016], foi proposta. A identificação dos fatores com efeito significativo no desempenho de entrada/saída (E/S) é uma informação crucial para o desenvolvimento dessa pesquisa. De fato, a caracterização e modelagem de desempenho é considerada um processo muito importante para o projeto de sistemas mais eficientes [Inacio and Dantas 2014].

Além da adoção de métodos estatísticos de projeto e análise de experimentos, para se obter uma caracterização mais precisa de um sistema, fazem-se necessárias ferramentas voltadas à execução dos experimentos. Principalmente, para geração de carga de trabalho. Tradicionalmente, benchmarks de aplicação, como o BT-IO, MADBENCH2 e FLASH-I/O, são utilizados para avaliar o desempenho de E/S de uma aplicação específica. Benchmarks sintéticos, como o IOR [Shan et al. 2008], por outro lado, oferecem uma maior flexibilidade sobre as cargas de trabalho geradas e, portanto, são mais adequados quando o objetivo é uma caracterização mais ampla e orientada aos fatores de desempenho.

Embora flexível, o IOR foi projetado para reproduzir padrões de acesso regulares, típicos de processos de *checkpointing/restart*, por exemplo. Todos os processos geram requisições de E/S de mesmo tamanho, para blocos contíguos de dados cujo tamanho deve ser múltiplo do tamanho da requisição. Nesse trabalho, se propõe o desenvolvimento de um benchmark para E/S paralela capaz de reproduzir padrões de acesso complexos, chamado IORE (IOR EXTENDED). A capacidade de reproduzir cargas de trabalho irregulares (*ex.*, processos heterogêneos) tem o potencial de auxiliar na compreensão não apenas de aplicações com padrões de acesso irregulares, mas também na análise de interferências entre aplicações com cargas diferentes compartilhando o subsistema de armazenamento.

IORE Benchmark

O projeto do benchmark IORE é baseado no IOR. A Figura 1 apresenta uma visão geral do benchmark. Parâmetros tradicionais oferecidos pelo IOR, como o tamanho da

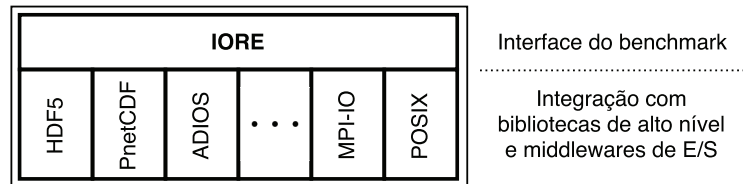


Figura 1. Visão geral do benchmark IORE

requisição e do bloco, serão flexibilizados para permitir requisições e blocos com tamanhos diferentes, incluindo a possibilidade de utilizar distribuições probabilísticas (*ex.*, normal, exponencial, etc.). Além de argumentos na linha de commando, o IORE suportará parametrização por meio de um arquivo com linguagens descritivas, como YAML e JASON, visando facilitar a definição de cargas de trabalho mais complexas (*ex.*, conjuntos de processos com cargas heterogêneas).

A extensão do IORE será facilitada por meio da adoção de interfaces bem definidas. Novas bibliotecas de alto nível e *middlewares* de E/S poderão ser incluídos pela implementação das funções da interface. Tal facilidade possibilitará que as funcionalidades de geração de carga implementadas no IORE sejam aproveitadas automaticamente.

Diferentemente do IOR, o IORE permitirá a execução de kernels híbridos. No IOR, todos os processos são *ranks* do MPI. Considerando a tendência de aplicações modernas combinando diferentes modelos de programação, o IORE suportará kernels combinando *threads* OPENMP e *ranks* MPI.

Considerações Finais

O IORE encontra-se no momento em fase de desenvolvimento. O código-fonte do projeto estará disponível em <https://github.com/lapesd/iore>, em conjunto com scripts de exemplo demonstrando as facilidades implementadas. Além da sua importância para o desenvolvimento do projeto principal dessa tese de doutorado, o CALLICRATES, acredita-se que o IORE possa colaborar em diferentes pesquisas na área de E/S paralela. Tais pesquisas incluem o estudo do comportamento de cargas de trabalho em arquiteturas de armazenamento da próxima geração, análises de interferências entre aplicações paralelas compartilhando o subsistema de armazenamento, novas bibliotecas de alto nível e *middlewares* de E/S, entre outros.

Referências

- Inacio, E. C. and Dantas, M. A. R. (2014). A survey into performance and energy efficiency in HPC, cloud and big data environments. *Intl. J. Net. Virt. Orgs.*, 14(4):299–318.
- Inacio, E. C. and Dantas, M. A. R. (2016). Callicrates: Configuração de sistemas de armazenamento paralelo em tempo de submissão de trabalhos. In *ERAD/RS 2016 XVI Escola Regional de Alto Desempenho*, pages 153–154. SBC.
- Shan, H., Antypas, K., and Shalf, J. (2008). Characterizing and predicting the I/O performance of HPC applications using a parameterized synthetic benchmark. In *SC '08 Proceedings of the 2008 ACM/IEEE conference on Supercomputing*. IEEE.