

# An Approach for Evaluating and Mitigating Intra-Application I/O Performance Variability Over Parallel File Systems

Eduardo Inacio<sup>1</sup>, Mario Antonio Ribeiro Dantas<sup>1</sup>

<sup>1</sup>Universidade Federal de Santa Catarina

**Abstract.** To meet ever increasing capacity and performance requirements of emerging data-intensive applications, highly distributed and multilayered back-end storage systems have been employed in large-scale high performance computing (HPC) environments. A main component of these storage infrastructures is the parallel file system (PFS), a especially designed file system for absorbing bulk data transfers from applications with thousands of concurrent processes. Load distribution on PFS data servers compose a major source of intra-application input/output (I/O) performance variability. Albeit mitigating variability is desirable, as it is known to harm application-perceived performance, understanding and dealing with I/O performance variability in such complex environments remains a challenging task. In this research, a differentiated approach for evaluating and mitigating intra-application I/O performance variability over PFSs is proposed. More specifically, from the evaluation perspective, a comprehensive approach combining complementary methods is proposed. An analytical model proposal, named DTSMaXLoad, provides estimates for the maximum load in a PFS data server. To complement DTSMaXLoad, modeling conditions and mechanisms hard to represent analytically, the Parallel I/O and Storage System (PIOSS) simulation model was proposed. Finally, for experimental evaluation over real environments, a flexible and distributed I/O performance evaluation tool, coined as IOR-Extended (IORE), was proposed. Furthermore, a high-level file distribution approach for PFSs, called N-N Round-Robin (N2R2), was proposed focusing on mitigating I/O performance variability for distributed applications where each process accesses an individual and independent file. An extensive experimental effort, including measurements on real environments, was conducted in this research work for evaluating each of the proposed approaches. In summary, this evaluation indicated both DTSMaXLoad and PIOSS modeling proposals can represent load distribution behavior on PFSs with significant fidelity. Moreover, results demonstrated N2R2 successfully reduced intra-application I/O performance variability for 270 distinct experimental scenarios, which, ultimately, translated into overall application I/O performance improvements.

**Resumo.** Para atender aos crescentes requisitos de capacidade e desempenho de aplicações que fazem uso intensivo de dados, sistemas de armazenamento altamente distribuídos e com múltiplas camadas têm sido empregados em ambientes de computação de alto desempenho (CAD). Um dos principais componentes dessas infraestruturas é o sistema de arquivos

paralelo (SAP), um sistema projetado para absorver transferências de dados em massa de aplicações com milhares de processos. A distribuição de carga nos servidores de dados dos SAPs compõe uma importante fonte de variabilidade de desempenho de entrada/saída(E/S). Embora reduzir tal variabilidade seja desejável, já que essa é conhecida por prejudicar o desempenho percebido pela aplicação, entender e lidar com a variabilidade nesses ambientes complexos continua sendo um desafio. Nesta pesquisa, uma abordagem diferenciada para avaliar e mitigar a variabilidade do desempenho de E/S intra-aplicação em SAPs é proposta. Uma proposta de modelo analítico, denominado DTSMaXLoad, fornece estimativas para a carga máxima em um servidor de dados. Para complementar o DTSMaXLoad, modelando condições e mecanismos difíceis de representar analiticamente, foi proposto o modelo de simulação Parallel I/O and Storage System(PIOSS). Para avaliação experimental em ambientes reais, foi proposta uma ferramenta de avaliação de desempenho de E/S flexível e distribuída, chamada IOR-Extended (IORE). Finalmente, uma abordagem de distribuição de arquivos de alto nível para SAPs, chamada N-N Round-Robin (N2R2), foi proposta visando a redução da variabilidade do desempenho de E/S para aplicações em que cada processo acessa um arquivo independente. Um extenso esforço experimental foi realizado neste trabalho de pesquisa para avaliar cada uma das abordagens propostas. Em resumo, essa avaliação indicou que as propostas de modelagem DTSMaXLoad e PIOSS podem representar o comportamento da distribuição de carga em SAPs com fidelidade significativa. Adicionalmente, os resultados demonstraram que o N2R2 reduziu com sucesso a variabilidade de desempenho de E/S intra-aplicação para 270 cenários experimentais distintos, que, em última análise, se traduziram em melhorias gerais de desempenho de E/S da aplicação.