Workload Characteristics of a Multi-cluster Supercomputer

Hui Li[1], David Groep[2], and Lex Wolters[1]

[1] Leiden Institute of Advanced Computer Science (LIACS),
Leiden University, The Netherlands
hli@liacs.nl, llexx@liacs.nl
[2] National Institute for Nuclear and High Energy Physics (NIKHEF),
The Netherlands
davidg@nikhef.nl

Abstract. This paper presents a comprehensive characterization of a multi-cluster supercomputer[3] workload using twelve-month scientific research traces. Metrics that we characterize include system utilization, job arrival rate and interarrival time, job cancellation rate, job size (degree of parallelism), job runtime, memory usage, and user/group behavior. Correlations between metrics (job runtime and memory usage, requested and actual runtime, etc) are identified and extensively studied. Differences with previously reported workloads are recognized and statistical distributions are fitted for generating synthetic workloads with the same characteristics. This study provides a realistic basis for experiments in resource management and evaluations of different scheduling strategies in a multi-cluster research environment.

1 Introduction

Workload characterization of parallel supercomputers is important to understand the system performance and develop workload models for evaluating different system designs and scheduling strategies [1,2]. During the past several years, lots of workload data has been collected [3], analyzed [4,5,6], and modeled [7,8,9]. Benchmarks and standards are also proposed for job scheduling on parallel computers [10].

In previously studied workloads [4,5,6,7], some characteristics are similar. For example, most of the workloads are collected from large custom-made production facilities (IBM SP2, SGI Origin, etc) in supercomputing centers. Jobs typically request "power-of-two" number of processors and have different arrival patterns in different periods (e.g. peak and none-peak hours in a daily cycle). Some characteristics, such as job attribute distributions and correlations, vary across different workloads [4,5,11]. Other characteristics are studied and reported separately, such as job cancellation rate [9] and conditional distributions (e.g. actual runtime distributions conditioned on requested runtime [4]). In this paper we compare our workload with previous reported ones on a per characteristics basis.

[3] Distributed ASCI Supercomputer-2 (DAS-2). ASCI stands for Advanced School for Computing and Imaging in the Netherlands.

D. Feitelson, L. Rudolph, and U. Schwiegelshohn (Eds.): JSSPP 2004, LNCS 3277, pp. 176–193, 2005.

This paper presents a comprehensive workload characterization of the DAS-2 [12] supercomputer. The DAS-2 system is interesting in that it is built using the popular COTS (Commodity Off The Shelf) components (e.g. Intel Pentium processors and Ethernet networks) and consists of multiple distributed clusters serving the participating universities. Not like other production machines, DAS-2 is dedicated to parallel and distributed computing research thus has much lower system utilization. We analyze twelve-month workloads on DAS-2 clusters in year 2003. Characteristics include system utilization, job arrival rate and interarrival time, job cancellation rate, job size (degree of parallelism), job runtime, memory usage, and user/group behavior. Correlations between metrics are also identified and studied.

The contributions of this paper reside in the following. Firstly, our study is based on cluster workloads. Cluster computing is a popular alternative in the HPC community and to our knowledge, not much work has been done in characterizing cluster workloads. Secondly, the system we study is a research facility. This provides an interesting comparison point to the well studied production workloads. Thirdly, we present a comprehensive characterization of the DAS-2 workloads. We not only analyze most of the metrics appeared in previous work, but also extensively study the correlations between different characteristics. Moreover, we fit the observed data with statistical distributions to facilitate synthetic workload generation. This research serves as a realistic basis in modeling cluster workloads, which contributes as input for evaluations of different scheduling strategies in a multi-cluster research environment [13].

The rest of the paper is organized as follows. Section 2 provides an overview of the DAS-2 system and workload traces used in our study. Section 3 analyzes the overall system utilization. Section 4 describes the job arrival characteristics, including job arrival rate, job interarrival time and job cancellation rate. Distributions are fitted for job interarrival times and job cancellation lags. Section 5 describes job execution characteristics. This includes job size, job actual runtime, memory usage, and correlations between them. Distributions and/or conditional distributions are also provided. Section 6 describes user/group behavior and its implications in modeling and predictions. In section 7 conclusions are presented and future work is discussed.

2 The DAS-2 Supercomputer and Workload Traces

The DAS-2 supercomputer consists of five clusters located at five Dutch universities and is primarily used for computing and scientific research. The largest cluster (Vrije Uni-

Cluster	Location	#CPUs	Period	#Job entries
fs0	Vrije Univ. (VU)	144	01-12/2003	219618
fs1	Leiden Univ.	64	01-12/2003	39356
fs2	Univ. of A'dam (UvA)	64	01-12/2003	65382
fs3	Delft Univ. of Tech.	64	01-12/2003	66112
fs4	Utrecht Univ.	64	02-12/2003	32953

Table 1. DAS-2 clusters and workload traces.