A Holistic Approach to Log Data Analysis in High-Performance Computing Systems: The Case of IBM Blue Gene/Q

Alina Sîrbu and Ozalp Babaoglu

Department of Computer Science and Engineering, University of Bologna,
Mura Anteo Zamboni 7, 40126 Bologna, Italy
alina.sirbu@unibo.it,ozalp.babaoglu@unibo.it

Abstract. The complexity and cost of managing high-performance computing infrastructures are on the rise. Automating management and repair through predictive models to minimize human interventions is an attempt to increase system availability and contain these costs. Building predictive models that are accurate enough to be useful in automatic management cannot be based on restricted log data from subsystems but requires a holistic approach to data analysis from disparate sources. Here we provide a detailed multi-scale characterization study based on four datasets reporting power consumption, temperature, workload, and hardware/software events for an IBM Blue Gene/Q installation. We show that the system runs a rich parallel workload, with low correlation among its components in terms of temperature and power, but higher correlation in terms of events. As expected, power and temperature correlate strongly, while events display negative correlations with load and power. Power and workload show moderate correlations, and only at the scale of components. The aim of the study is a systematic, integrated characterization of the computing infrastructure and discovery of correlation sources and levels to serve as basis for future predictive modeling efforts.

Keywords: Data science, correlation analysis, HPC system monitoring, log data integration, predictive modeling.

1 Introduction

As the size and complexity of high-performance computing (HPC) infrastructures continue to grow driven by exascale speed goals, maintaining reliability and operability levels high, while keeping management costs low, is becoming increasingly challenging. Continued reliance on human operators for management and repair is not only unsustainable, it is actually detrimental to system availability: in very large and complex settings like data centers, accidental human errors have been observed to rank second only to power system failures as the most common causes of system outages [14].

Large computing systems produce large amounts of data in the form of logs tracing resource consumption, errors, events, etc. These data can be put to use for understanding system behavior and for building predictive models to tackle the management challenges. Most studies in this direction have focused on particular subsystems rather than the system as a whole, which is a necessary condition for achieving realistic models with good predictability traits [16]. With recent progress in *Data Science* and *Big Data*, it is becoming increasingly feasible to carry out such a holistic analysis towards improving predictions by considering data from a variety of sources covering different subsystems and measures [8].

In this paper we report the results of a characterization study integrating four datasets from different subsystems in an effort to understand the behavior of a 10-rack IBM Blue Gene/Q [10] installation and quantify the correlations among power, temperature, workload, and hardware/software events as well as among different system components. In certain cases, we report the lack of correlations, which can be just as important as their presence. These results provide a first step towards identifying important features for future predictive studies.

The contributions of this paper are threefold. First, we provide a characterization of a Blue Gene/Q system from thermal, power, workload, and event log perspectives, highlighting significant features for system behavior and the presence and absence of correlations between different components. No correlation in terms of power and thermal behavior was found across components, yet events exhibit significant spatial correlations, indicating possible propagation of errors. Secondly, an integrated analysis of the four datasets searches for correlations among various metrics so as to identify further possible relations for future modeling and prediction studies. This reveals significant positive correlation between power consumption and temperature, and a weaker negative correlation between hardware/software events and power or workload. There are also indications of correlations between workload and power but only at a finer spatial granularity (at rack rather than at system scale). Thirdly, we use the preliminary indications on the importance of different features for explaining system behavior to propose a feature set to be used in future work for event prediction. An important feature of our study is its holistic nature integrating multiple datasets, to an extent not present in the literature, neither in terms of system characterization, nor in terms of correlation and predictive studies.

In the next section we describe the data, while Section 3 contains our analysis for individual and integrated datasets. Section 4 includes related work, and Section 5 discusses future predictive studies and data quality.

2 Dataset description

Our data source is Fermi [7], an IBM Blue Gene/Q system run by CINECA, a consortium operating the largest data center in Italy. Fermi has 163,840 computing cores with a peak performance of 2.1 PFLOPS. Its workload includes large-scale models and simulations for several academic projects, including 3D models of the cell network of the heart, simulation of interaction between lasers and plasmas, neuronal network simulations, models of nano-structures and complex materials. Fermi is organized as 10 *racks*, each with 2 *mid-planes* of 16

Table 1. Four datasets that are analyzed

Dataset	Time span (2014)	Time resolution	Component	Total records
Power	28 Mar – 25 Jul	5 min	Bulk Power Module	9,655,298
Temperature	23 Apr – 25 Jul	15 min	Node-board	2,648,331
Workload	1 May – 27 Jul	NA	System	78,128
RAS	23 Apr – 25 Jul	NA	All	774,555

node-boards with 32 16-core nodes. Each mid-plane is powered by 18 *bulk power modules* (BPM). Logging is based on standard Blue Gene/Q tools [10]. The *Mid-plane Manager Control System* performs environmental monitoring, providing power and temperature logs. The *Machine Controller* handles access to the hardware components and provides so-called *Reliability, Availability and Serviceability* (RAS) logs. Workload is extracted from the *Portable Batch System* scheduler logs, using a custom tool by CINECA. Given that all data used in our analysis originate in logs from standard Blue Gene tools, we consider the information they contain to be correct. Table 1 summarizes the four datasets.

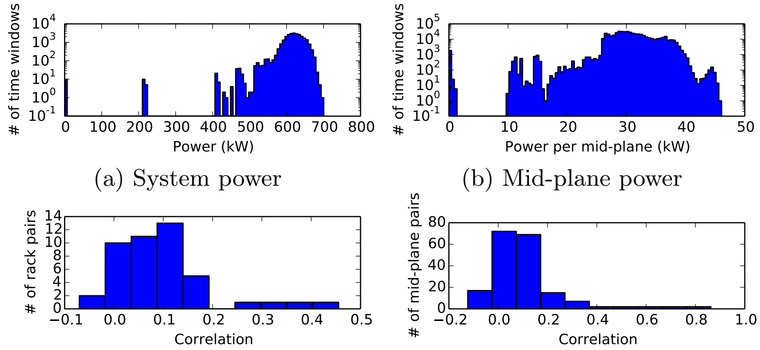
Power logs report input/output voltages and current levels for each BPM, with a 5-min resolution. By summing the input power levels over the different components, we obtained time series of power consumption for individual mid-planes, racks, and for the entire system. Power at the node-board scale cannot be reliably computed since 18 BPM power 16 node-boards (redundant system). *Temperature* logs are reported by the node-board monitor (two sensors/node-board), with a 15-min resolution. From these we computed averaged time series at node-board, mid-plane, rack, and system scales.

Workload data consist of a list of jobs with date of completion, running time, number of cores, queue time, and queue class. Fermi uses six queues, with increasing job length and core count: *serial* (on login nodes only), *debug*, *longdebug*, *smallpar*, *parallel* and *bigpar*. Two other classes — *visual* and *special* — exist, with very few jobs reserved for dedicated users. We computed the CPU time per job and time series of total daily CPU time, number of cores, and queue time. The daily CPU time per queue class was also extracted. Since only the date of job completion (not the exact time) is available in the data, totals are approximate, yet they give a very good indication of the daily load at system scale. No load information at other scales (node-board, mid-plane, rack) was available.

RAS logs consist of hardware and software events from all system components and are labeled FATAL, WARN or INFO, in decreasing order of severity. The dataset contains 163,134 FATAL, 473,982 WARN, and 137,438 INFO events. For each event, the exact time and location are included. From these data, we computed the distribution of inter-event times at system scale and also time series of the number of events in each category at various time and space resolutions.

3 Data analysis

Each dataset alone may provide useful insight into the functioning of Fermi, while an integrated analysis has even greater potential. Hence, in this section we



(a) System power (b) Mid-plane power

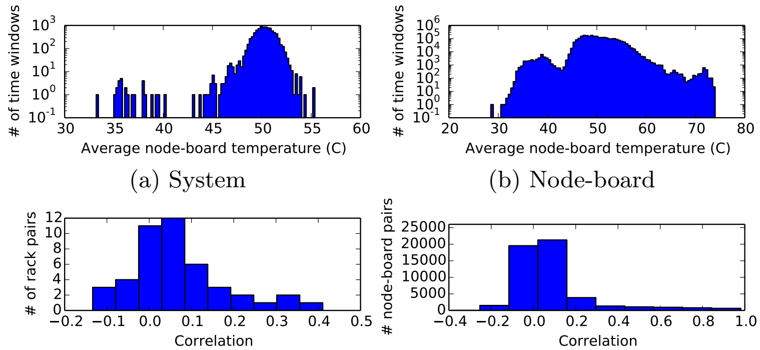
(c) Rack correlations (d) Mid-plane correlations

Fig. 1. Distribution of total power, sampled every 5 min, at system and mid-plane scale and of power correlations between racks and mid-planes.

first study each dataset individually, identifying and comparing their features, then we integrate them to study how metrics from different subsystems correlate. Pearson correlation coefficient is used across the paper to quantify correlations.

3.1 Individual datasets

Power logs. The specifications for a Blue Gene/Q system declare the typical power consumption to be around 65kW, with a maximum of 100kW per rack [13]. However, real consumption varies depending on system load and state of components (e.g., how many nodes are up). Fig. 1a-b displays the distribution of power consumption sampled at 5-min intervals, at system and mid-plane scale. Distributions are centered around the official average values: 650kW at the system scale (10 racks) and 32.5kW at the mid-plane scale, confirming the specifications. Moving from higher to lower scale, the distribution becomes broader. While total consumption is mostly between 50kW and 70kW per rack, with a



(a) System (b) Node-board

(c) Rack (d) Node-board

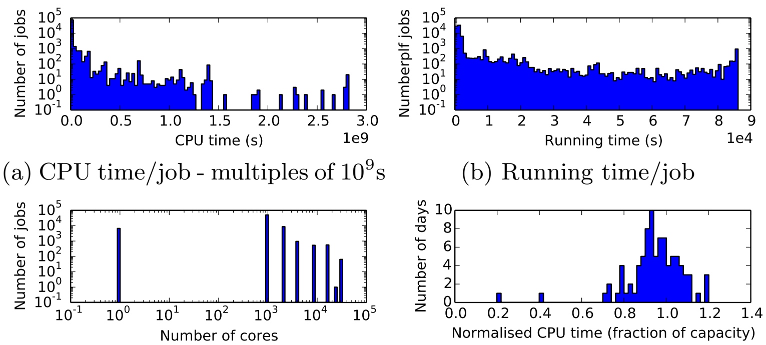
Fig. 2. Distribution of average temperatures, at system and node-board scale, and of temperature correlations among racks and node-boards.

bell-shaped distribution, for individual mid-planes additional peaks emerge with some showing power consumptions up to 46kW, but also frequent values under 20kW. Similar results were obtained at rack scale. This shows that power consumption is very heterogeneous, which needs to be taken into account for modeling. Indications are that while predicting overall system power might be easier due to greater stability in time, finer grained predictions at mid-plane scale might produce more accurate results.

It is interesting to see if power correlates across different components (racks or mid-planes). Traditional load balancing algorithms try to even out the work performed by different processing elements, and power increases with load, so we would expect power to be correlated across different system components under heavy load. Fig. 1c-d shows correlations of power consumption between rack pairs, and between mid-plane pairs. At both scales, correlations are in general very low. Only a few mid-plane pairs have correlation values above 0.5. As we will see later, the observed system load is generally high. The lack of strong correlation for power consumption among components could be interpreted as an effect of energy-aware scheduling [18], yet, this is not the case here since Fermi uses the native IBM LoadLeveler scheduler which is not optimized for power. A different explanation for the weak correlations could be poor design of the applications running on the system: if synchronization requires some program threads to wait, these will keep the nodes occupied but without using them fully. However, given the coarse resolution of the workload dataset, this hypothesis cannot be tested with the current data.

Temperature logs. Fig. 2a-b shows histograms of average temperatures sampled every 15 min. For the overall system, with few exceptions, the distribution is again bell-shaped and narrow with one mode around 50°C. As we zoom in at node-board scale (the lowest available in the data), the distribution becomes again wider with additional peaks appearing at very high and very low temperatures. Individual node-boards can reach up to 75°C, significantly greater than the system average. Similar results were obtained at intermediate scales (rack and mid-plane). This again shows how the system appears to behave differently at different scales, with greater heterogeneity in time at the finer-grained logs. For temperature correlations among different components of the same type (Fig. 2c-d), a pattern similar to power consumption is observed. With very few exceptions, temperature exhibits low correlation across components. Results are consistent across all scales (including mid-plane not shown here). In terms of thermal isolation, this is good news, since having one hot node-board does not imply surrounding node-boards are hot as well. Yet, the fact that power consumption showed a similar pattern, this can be additional evidence that workload is not well balanced or applications need improvement.

Workload logs. An important question in terms of workload regards the types of jobs submitted to the system. Fig. 3a-c displays the distribution of several job attributes: CPU time, running time, and number of cores used. In terms of time requirements, jobs are very heterogeneous as evidenced by a long-tailed CPU time distribution, with a few very heavy jobs and many short jobs



(a) CPU time/job - multiples of 10^9s (b) Running time/job

(c) Core count/job (d) Daily CPU time (normalized)

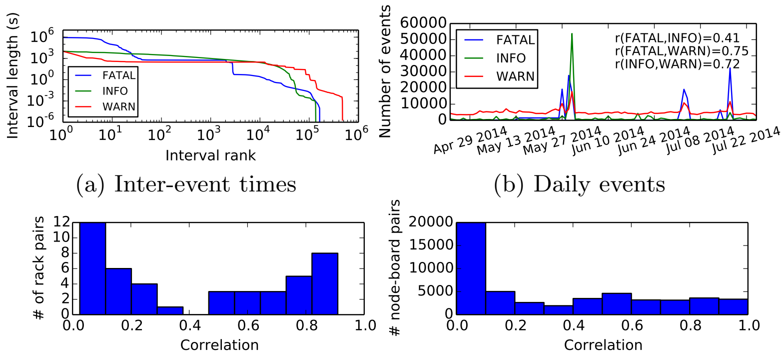
Fig. 3. Workload structure: distribution of CPU time, running time, number of cores per job, and CPU time consumed by jobs completed on the same day (normalized by the overall capacity of Fermi, which is 14,069,376,000 s/day).

present. Effective running times are bimodal, with many short jobs and many long jobs (all running times under 24 hours), and slightly fewer medium-length jobs. The number of cores per job is less heterogeneous, with only eight different values present, most jobs using over 100 cores and up to 32,768. So, in general, jobs are highly parallel. Out of all 78128 jobs submitted, only $\sim$75% were started (running times > 0) and only those will be used in the subsequent sections.

The structure of the workload data enables analysis of patterns in time only at system scale and 24-hour resolution. Fig. 3d shows total CPU time for all jobs *completed* each day, normalized by the overall system capacity. This does not represent the exact system load for that day, but it still is a very good indication. The data contain only the date of job completions not the exact time, making it impossible to compute how many hours each job ran in a given day — jobs completed on one day could have been started the previous day. This is why some days reach capacity exceeding 100%. Again, roughly a bell-shaped distribution is observed, with a mean around 94% usage, indicating very high load levels.

RAS logs. The inter-event times at system scale, for the three event types, do not appear to follow a known distribution (Figure 4a). FATAL events show a few very large and many very small intervals, indicating a pattern with spikes of events in short periods of time with large breaks between them. INFO and WARN events are more evenly spread in time, missing the very large inter-event times, and having a smaller fraction of very short intervals.

Fig. 4b displays the time-series of daily number of events in each category and their relative correlations. WARN and INFO events are more common daily, whereas FATAL events come in spikes and appear in only a few of the monitored days. The 4 larger spikes in FATAL events correspond to issues related to the BPMs which caused shutdown of the entire system several times between 27/05 and 30/05 and shutdown of rack R30 on 04/07 and 17/07. Daily INFO and WARN events are highly correlated, and so are WARN and FATAL events. However INFO and FATAL events seem to appear together less frequently. This could mean that



(a) Inter-event times (b) Daily events

(c) Rack correlation (d) Node-board correlation

Fig. 4. (a) Inter-event times for the three event categories. Intervals between events were ranked in descending order. For each interval, the x axis shows its rank and the y axis its value. (b) Total daily number of RAS events. Correlation of FATAL events for rack pairs (c) and node-board pairs (d).

INFO events could be useful to predict WARN events while WARN events could predict FATAL events at this time resolution. Hence, considering both INFO and WARN events to predict FATAL events could facilitate longer prediction lead time.

A different question is whether events correlate across different components. Fig. 4c-d shows the distribution of correlations between rack and node-board pairs for FATAL events. Similar results were obtained for the other events and at mid-plane scale. Unlike power and temperature, FATAL events have higher correlation across components, with a significant number of pairwise correlations larger than 0.5. This indicates that failures may propagate across components. We studied for various FATAL event types the number of different components (node-boards, power modules, etc.) affected in 5-min windows. We found that most event occurrences do involve a large number of components, sometimes up to a few hundred. So, when trying to predict component failures, one needs to take into account not only their individual behavior, but also that of the others. The way failures propagate can also give indications of the possible causes (e.g., a faulty job running on all components) and enable their automatic identification.

3.2 The big picture

Individual datasets have shed some light into the functioning of the Fermi system, and correlations between components. Here we integrate the four datasets to uncover further correlations between the different components and logs.

A first analysis looks at different measures for the overall system for 24-hour time windows. Figure 5 shows all pairwise correlations between several time series datasets. We note strong correlation between temperature and power, confirming what has been observed in other systems as well [3]. In terms of workload, total daily CPU time, number of cores, and queue time are included. These do not appear to correlate among themselves, while CPU time is the

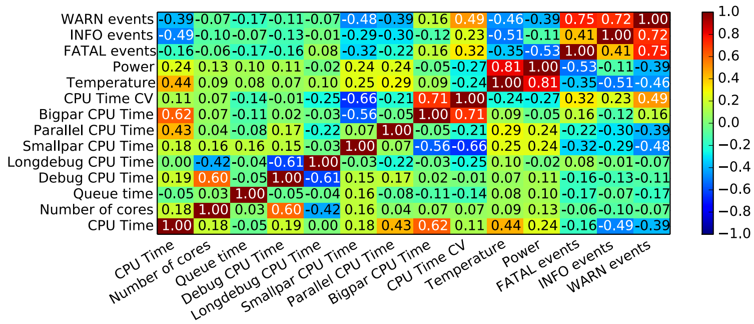


Fig. 5. Correlation between datasets at 24-hour resolution for the overall system.

only one among the three that does correlate with other datasets, although only moderately. Specifically, positive correlation with the temperature is present, so the system does show thermal symptoms of working harder under a high workload. A negative correlation with RAS events also exists, which is somewhat counterintuitive: one would expect more events to appear when the system works harder. However, it is quite possible for large numbers of RAS events to have resulted in system failure, which in turn resulted in fewer completed jobs for that day, explaining the negative correlation. In fact, a closer analysis of the data shows that, in general, a system shutdown (signaled by long periods of missing data in the trace) is preceded by fatal events. In some situations, events may appear also at system restore, which could be due to operator interventions made while the system was down. A negative correlation also appears between power/temperature and RAS events, again rather counterintuitively and due to the same factors as before. So, when trying to predict power consumption or FATAL events, one needs to take into consideration the negative dependence.

The data do not show any correlation between overall workload and power, however correlation could depend on the job class (queue). So we analyzed (same Fig. 5) the daily CPU Time per job class and also the coefficient of variation (CV) of the total CPU Time across the classes. A higher CV means a more unbalanced workload across the queues. The negative correlation between CPU time and RAS events is present for individual queues as well, with strongest effect for *smallpar* jobs. CPU Time CV displays some positive correlation to WARN events, which means that heterogeneity in terms of jobs per queue can be a factor leading to WARN events. However, even at this scale, no link between workload and power

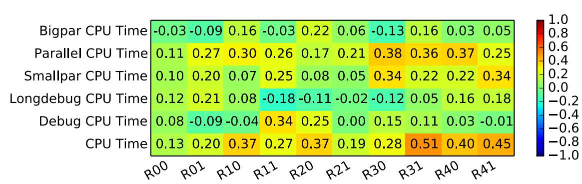
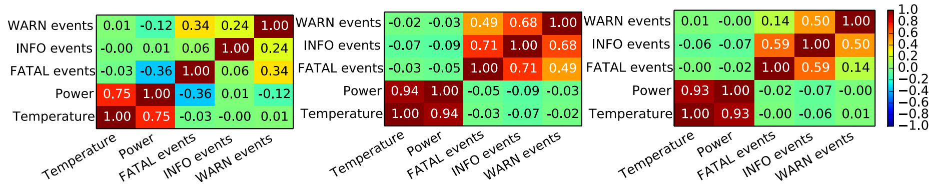


Fig. 6. Correlation between CPU time and power for the 10 racks (R00-R41).



(a) System (b) Rack (c) Mid-plane

Fig. 7. Correlation between datasets at 3 different system scales.

consumption can be found (we also explored other measures, such as job count, core count, queue time per class, with similar results). This suggests that the way workload is distributed on components is important to understand power in this system. Higher correlations might be obtained by zooming in at rack, mid-plane or node-board scales. We can do this for power, but not for workload due to the structure of our dataset. Fig. 6 shows how CPU Time correlates with power consumption per rack. Indeed, higher correlations do appear, indicating structure is important, but still more detailed workload data is required. This suggests the need for changes in the structure of workload logs for Fermi and improving system logging practice, in order to see exactly at which scale correlations appear.

In a second analysis, the resolution of the data was increased to 5 min. Correlations at system, rack, and mid-plane scales are shown in Fig. 7. Due to its coarse time structure, the workload data was excluded. Power and temperature correlation grows with increased time and space resolution. This suggests that for predictions, using only one of the two features might suffice, which is good news since power logs at node-board scale are not available. However, to account for the possibility that temperatures are affected by cooling issues, power should still be monitored, even if not to be used as a modeling feature, but to check that assumed correlation is correct. A sudden decrease in correlation could also flag cooling anomalies. The negative correlation between temperature/power and RAS events is maintained, albeit at a lower value, only at system scale between power and FATAL events.This can be again explained by the existence of periods of system shutdown before or after events. So, correlations will be high over 24 hours, but for 5-min windows, only FATAL events are correlated with power (temperatures take longer to drop, while other event types could have appeared much earlier). Between RAS events, correlations are higher within racks, indicating that propagation of errors might be strongest at the rack scale.

4 Related work

Log analysis for characterization of large computing infrastructures has been the focus of numerous recent studies. The release of two Google workload traces has triggered a flurry of analysis activity. General statistics, descriptive analyses, and characterization studies [12, 15] have revealed higher levels of heterogeneity when compared to grid systems [5]. Some modeling work has also appeared based on these data [19, 2, 17]. While they have provided important insight into Google clusters, focusing only on workload aspects of the system has been limiting. To be effective, it is essential to integrate data from different components and sources. Other traces have also been studied in the past [4], and tools for their analysis developed [9], but again concentrating on a single data type. Here we perform similar analyses for a Blue Gene/Q system but from several viewpoints: workload, RAS, power, and temperature, providing a more complete picture of the system under study.

RAS logs from IBM Blue Gene systems have been included in several earlier studies. In [6] prediction of FATAL events in a Blue Gene/Q machine is attempted while an earlier study of a Blue Gene/L installation is [11]. Both compare several classification tools (SVM, customized KNN, ANN, feature selection, rule-based models). These predictive studies look only at RAS events, while adding further data from other system components could improve prediction accuracy significantly, as noted by the authors themselves. In this paper we provide the first step towards such an analysis, where we perform descriptive analytics mandatory before any prediction can be attempted.

Some integration is performed in a very recent study from Google that models Power Usage Effectiveness using thermal information (temperatures, humidity, etc) and overall system load, using an Artificial Neural Network [8]. Another recent development in this direction is a novel monitoring system [3] designed for a hybrid HPC platform. In this study, several types of data including workload, power, chiller, and machine status are recorded. In principle, these data could be used for future predictive and modeling studies, but they have not been initiated in the reported study. The OVIS project has also developed an integrated monitoring platform called the "Lightweight Distributed Metric Service" [1], recording various system metrics for optimization of application performance. The platform has been tested on several systems, but again steps towards a descriptive and predictive analytics for these data are still missing.

5 Discussion and conclusions

Given the need for a holistic analysis of large computing infrastructures, this paper has presented a characterization study conducted with four datasets describing different subsystems of an IBM Blue Gene/Q installation. Temperature, power consumption, workload, and RAS logs were studied independently to characterize the system and then together to identify correlations between datasets.

The results obtained from correlation analysis will serve as a guideline for a future study aiming to predict in advance FATAL RAS events based on the rest of the data. One possibility would be predicting, for each node-board, the number of FATAL events in the next 24 hours. Alternatively, based on the number of events, we can define discrete failure classes (e.g., NONE, FEW, MANY) to be predicted. We have compiled a set of possible features that may be suitable for this predictive task (Table 2). These cover all datasets with various time resolutions at node-board and system scale.

Table 2. Possible feature set for prediction of FATAL RAS events.

Feature	**Period**	**Scale**
Temperature average and standard deviation	6h	node-board
Correlation between temperature and power	6h	mid-plane
Temperature correlation between node-boards	6h	node-board
CPU time per queue	24h	system
CPU time coefficient of variation across queues	24h	system
Number of WARN, INFO, and FATAL events	24h	node-board
Stdev of number of WARN, INFO, and FATAL events	24h	node-board
Number of WARN, INFO, and FATAL events	24h	system
Stdev of number of WARN, INFO, and FATAL	24h	system
Correlation between temperature and event count	6h	node-board
Correlation between power and event count	6h	system

The first two features are suggested by the fact that power and temperatures are highly correlated. Temperatures can be used a a proxy for power, so that the higher space resolution is employed and the number of features is decreased. This only as long as correlation between temperature and power is high. A decrease in correlation will signal an anomaly, even if the proxy is no longer valid. Large temperature correlations across node-boards could also signify anomalies, since node-board temperatures were uncorrelated in our data. Workload related features are limited to daily CPU time per queue and coefficient of variation across queues, which showed highest correlation with other datasets. Features monitoring all types of RAS events at node-board level account for correlation across RAS event types, while those at system level are justified by correlations across node-boards and propagation of errors. Since prediction is aimed for 24-hour periods, we use event values computed over the same time, but also deviations, to account for varying inter-event patterns. Finally, correlations between power (or temperature at node-board scale) and events should be monitored since large negative correlation could signal component failure. Even if indications are that the features listed will prove important for prediction, final evaluation of the feature set will be performed during the future predictive study itself.

Besides identifying important features, our analysis has also indicated directions for improvement in terms of data collection. Workload data in particular proved to be insufficient for our goals, so we could identify few relations to the other datasets. In the future, at least timestamps for job completion as well as job placement should be included. This additional information will enable analyzing the causes of lack of power correlation across components. Power monitoring was coarse in terms of space resolution, however more data could be extracted from the node-board power rails. Temperatures, on the other hand, could be logged at 5-min intervals rather than 15. We are aiming at prediction with long lead time, so the 15-min interval may be sufficient for applying the model, however finer granularity would allow for more refined training data. In the future we will also use data external to the computing infrastructure, such as the water and air cooling systems, together with data outside the data center, e.g. weather and seismic activity. Cross-correlations will also be investigated, resulting in further features to be added to the proposed set.

6 Acknowledgments

We are grateful to the HPC team at CINECA for sharing with us the log data related to the Fermi system and for helpful discussions.

References

1. Agelastos, A., et al.: The lightweight distributed metric service: a scalable infrastructure for continuous monitoring of large scale computing systems and applications. In: SC'14. pp. 154–165 (2014)
2. Balliu, A., et al.: Bidal: Big data analyzer for cluster traces. In: Informatika (BigSys workshop). vol. 232, pp. 1781–1795. GI-Edition Lecture Notes in Informatics (2014)
3. Bartolini, A., et al.: Unveiling eurora - thermal and power characterization of the most energy-efficient supercomputer in the world. In: Proceedings of the Conference on Design, Automation & Test in Europe. pp. 277:1–277:6. DATE '14 (2014)
4. Chen, Y., Alspaugh, S., Katz, R.H.: Design Insights for MapReduce from Diverse Production Workloads. Technical Report, UC Berkeley UCB/EECS-2 (2012)
5. Di, S., Kondo, D., Cirne, W.: Characterization and Comparison of Google Cloud Load versus Grids. In: IEEE CLUSTER. pp. 230–238 (2012)
6. Dudko, R., Sharma, A., Tedesco, J.: Effective Failure Prediction in Hadoop Clusters. University of Idaho White Paper pp. 1–8 (2012)
7. Falciano, F., Rossi, E.: Fermi: the most powerful computational resource for italian scientists. EMBnet. journal 18(A), 62 (2012)
8. Gao, J.: Machine learning applications for data center optimisation. Google White Paper (2014)
9. Javadi, B., et al.: The Failure Trace Archive: Enabling the comparison of failure measurements and models of distributed systems. Journal of Parallel and Distributed Computing 73(8) (2013)
10. Lakner, G., et al.: IBM System Blue Gene Solution: Blue Gene/Q System Administration. IBM Redbooks (2013)
11. Liang, Y., et al.: Failure Prediction in IBM BlueGene/L Event Logs. IEEE ICDM pp. 583–588 (Oct 2007)
12. Liu, Z., Cho, S.: Characterizing Machines and Workloads on a Google Cluster. In: 8th SRMPDS (2012)
13. Milano, J., et al.: IBM System Blue Gene Solution: Blue Gene/Q Hardware Overview and Installation Planning. IBM Redbooks (2013)
14. Ponemon Institute Research, Emerson Network Power: Cost of Data Center Outages (Dec 2013)
15. Reiss, C., et al.: Heterogeneity and Dynamicity of Clouds at Scale: Google Trace Analysis. In: ACM SoCC (2012)
16. Salfner, F., Lenk, M., Malek, M.: A survey of online failure prediction methods. ACM Computing Surveys (CSUR) 42(3), 1–68 (2010)
17. Sîrbu, A., Babaoglu, O.: Towards data-driven autonomics in data centers. In: International Conference on Cloud and Autonomic Computing (ICCAC) (2015)
18. Valentini, G.L., et al.: An overview of energy efficiency techniques in cluster computing systems. Cluster Computing 16(1), 3–15 (2013)
19. Wang, G., et al.: Towards Synthesizing Realistic Workload Traces for Studying the Hadoop Ecosystem. In: IEEE MASCOTS. pp. 400–408 (2011)