Long Range Dependent Job Arrival Process and Its Implications in Grid Environments

Hui Li[*]

Leiden Institute of Advanced
Computer Science, Leiden University
Niels Bohrweg 1, 2333 CA, Leiden, The Netherlands
hui.li@computer.org

ABSTRACT

Job arrivals can be described as point processes and it is shown that correlations and fractal behavior can be reliably revealed using the count/rate representation. Using real workload data from production Grids, we show that the second order properties such as the autocorrelation function (ACF) and the scaling behavior can be well reconstructed by a Multifractal Wavelet Model (MWM). A so-called controlled-variability integrate-and-fire (CV-InF) algorithm is applied to transform rates into interarrivals so that a full description of the arrival process can be obtained. The additive nature of rates makes it possible to model different patterns separately and aggregate them back to form a unified process. We further quantify the performance impacts of autocorrelated job arrivals in Grid scheduling using model-driven simulation. It is shown that autocorrelations in the arrival processes can cause performance degradation both at the local and the Grid level.

Keywords

Point Processes, Long Range Dependence, Performance Evaluation, Computational Grids

1. INTRODUCTION

Performance evaluation of scheduling studies requires representative workloads to produce dependable results. Real traces on production Grids are being collected and have not yet become widely available[1]. Models for generating synthetic Grid workloads, on the other hand, are still emerging and subject to further research. Consequently a large number of literature on Grid scheduling studies either use traces available from parallel supercomputers[2] [11], or make simplified assumptions on workload characteristics (Poisson arrivals [13], fixed-interval arrivals [4], or constant background loads [16]). However, it is shown that workload characteristics on data-intensive Grids are significantly different than those on conventional parallel supercomputers. As has been studied in [6], pseudo-periodicity, long range dependence, and the "bag-of-tasks" behavior with strong temporal locality are the main properties that characterize data-intensive workloads. Therefore, it is important that representative models be developed to capture the salient properties of Grid workloads. In [7] a matching pursuit approach is proposed to model pseudo-periodic job arrival patterns. In this paper we focus on long range dependence and fractal behavior in job arrival processes, and how different patterns can be aggregated and transformed. By using the developed model we conduct model-driven simulation of Grid scheduling strategies and quantify the performance differences of various correlation structures in job arrivals.

The contribution of this work is three-fold. Firstly the Multifractal Wavelet Model (MWM) [12], introduced in the context of network traffic, is applied successfully to model the long range dependent and fractal job arrivals in Grids. Secondly, we adopt a controlled-variability integrate-and-fire algorithm to convert a rate process to an interarrival process, through which a full description of the arrival process can be obtained. The additive nature of rates can be utilized to facilitate model composition and aggregation, which makes our approach a coherent solution for modeling Grid job arrivals at different levels. Thirdly, initial results on the performance implications of autocorrelated job arrivals in Grid scheduling are presented.

The rest of the paper is organized as follows. Section 2 introduces the point process and its several representations, namely, the interarrival process and the count/rate process. Section 3 defines several important notions such as scaling, long range dependence (LRD), and (multi)fractals. Section 4 discusses the Multifractal Wavelet Model and explains its recursive structure. A deep understanding is obtained by laying out the relationship among cascading, wavelets, and scaling. Section 5 presents the controlled-variability integrate-and-fire algorithm for transforming a rate function to an interarrival process. Section 6 discusses the additive nature of rates and the advantages gained thereby from a modeling perspective. Section 7 introduces the methodol-

[1]Grid Workload Archive. http://gwa.ewi.tudelft.nl/.

[2]Parallel Workload Archive. http://www.cs.huji.ac.il/labs/parallel/workload/.

ogy and experimental design for simulation of Grid scheduling algorithms using synthetic traces. Section 8 presents the model fitting results using real workload data from the LCG production Grid and several representative data-intensive clusters. Simulation results for the scheduling strategies are presented and analyzed as well. Conclusions and future work is discussed in Section 9.

2. POINT PROCESSES

Job arrivals can be described as a (stochastic) *point process*, which is defined as a mathematical construct that represents individual events as random points at times $\{t_n\}$. There are different representations of a point process. An *interarrival time process* $\{I_n\}$ is a real-valued random sequence with $I_n = t_n - t_{n-1}$ and it completely describes the point process. The sequence of counts or the *count process*, on the other hand, is formed by dividing the time axis into equally spaced contiguous intervals of T to produce a sequence of counts $\{C_k(T)\}$, where $C_k(T) = N((k+1)T) - N(kT)$ denotes the number of events in the kth interval. The normalized version of the sequence of counts is called the *rate process* $R_k(T)$, where $R_k(T) = C_k(T)/T$.

In general, forming the sequence of counts loses information because the inter-event times within interval T are not preserved. Nevertheless, this representation is of primary importance because it preserves the correspondence between the discrete time axis of the count process $\{C_k(T)\}$ and the absolute "real" time axis of the underlying point process. We can readily associate correlations in the process $\{C_k(T)\}$ with correlations in the point process. The interarrival time process, on the other hand, eliminates the direct correspondence between absolute time and the index number therefore it only allows rough comparisons of correlations [9]. As is shown in [6], measures based on interarrivals are not able to reliably reveal the correlation structures of the underlying job arrival processes and count based measures should be trusted instead. Figure 2 illustrates such a case that the rate process is LRD while the interarrival process is short range dependent with quickly vanishing autocorrelation lags.

3. SCALING, LRD, AND FRACTALS

Physical processes can be observed at a large range of *scales*. For instance, one can represent the network traffic as the number of bytes at the level of milliseconds, seconds, up to minutes. On clusters and Grids the number of job arrivals can be aggregated and averaged every second, every minute or even every hour. *Scaling*, or scale invariance, means the lack of any special characteristic scale (all scales have equal importance). The power law relationship is intrinsic in understanding the scaling behavior and it occurs in many of the following presentations, such as the first-order statistics (marginal distribution), the second-order statistics (slow decaying variance, ACF), and nonlinear transformations (spectrum, wavelet coefficients). The scaling component can be estimated via some of these transformations, showing as a slope in the log-log plot [1].

Exact *self-similar* and *long range dependent* (LRD) processes are the two most important classes of the general scaling processes. A process $X(t)$ is said to be long range dependent (LRD) if either its autocorrelation function (ACF) or power spectrum satisfies the following conditions:

$$R(k) \sim c_r k^{\alpha-1}, k \to \infty, \text{ or } S(f) \sim c_f f^{-\alpha}, f \to 0. \quad (1)$$

The autocorrelation function $R(k)$ decays so slowly that $\sum_{k=-\infty}^{\infty} R(k) = \infty$ and $S(0) = \infty$. Frequency-domain characterization of LRD leads to a class of so-called $1/f$-*like processes* or $1/f$ *noise* [18].

Scaling with one single component is referred as *monofractal*. If multiple scaling components are found within one process, or the scaling is time-depedent, the process is called *multifractal*. As an example the count process can exhibit *biscaling* [1]. The scaling concentrated at lower scales indicates the fractal nature of the sample path. The alignment at higher scales, on the other hand, resembles that of a stationary short range dependent (SRD) process. Such a process is illustrated in Section 8 and models that can capture its behavior are discussed in the following section.

4. WAVELET ANALYSIS AND SYNTHESIS

Multiplicative cascades [5] form a very important paradigm for generating multifractal processes. Our discussion in this section starts with the simplest multiplicative process called *binomial cascade*. Binomial cascades based models have been proposed to capture the LRD and multifractal nature of a variety of processes, including disk and network traffic [12, 17]. We choose the multifractal wavelet model (MWM) introduced by Riedi et al. [12] because it provides a coherent wavelet framework for analysis and synthesis of the scaling behavior. Moreover, with the wavelet energy decay estimated from the original process, MWM can potentially model the scaling behavior with multiple exponents (e.g. biscaling). We discuss the discrete wavelet transform (DWT) in the framework of filter banks (introduced by S. Mallat [10]), based on which the structure of the multifractal wavelet model (MWM) is well explained.

Due to its inherent multiresolution properties, wavelets provide a natural framework for analyzing the scaling behavior. Like the Fourier transform that decomposes signals with sinusoidal functions, the wavelet transform projects the signal onto the so-called *wavelets* [10, 14]. A *wavelet function* $\psi(t)$ is a function that can be scaled and shifted: $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$. There also exists a *scaling function* $\phi(t)$, which can be scaled and shifted as well. As a result the signal is decomposed into a sum of weighted scaling functions and wavelets:

$$X(t) = \sum_k c_{j_0,k}\phi_{j_0,k} + \sum_{j \leq j_0} \sum_k d_{j,k}\psi_{j,k}(t), \quad (2)$$

where $c(j_0, k)$ and $d(j, k)$ are referred as *scaling coefficients* (or approximations) and *wavelet coefficients* (or details), respectively.

A discrete wavelet transform (DWT) of a signal can be calculated by passing the signal recursively through a set of lowpass and bandpass filters [14]. The scaling function $\phi(t)$ mentioned above corresponds to a lowpass filter while the wavelet function involves a highpass (or bandpass) filter. In the Haar case the dilation equation for the scaling function can be defined as $\phi(t) = \phi(2t) + \phi(2t-1)$. This occurs when $\phi(t)$ is the box function, namely $\phi(t) = 1$ for $0 \leq t < 1$ and $\phi(t) = 0$ otherwise. The wavelet equation, on the other hand, can be written as $\psi(t) = \phi(2t) - \phi(2t-1)$. Explicitly, there is $\psi(t) = 1$ for $0 \leq t < \frac{1}{2}$ and $\psi(t) = -1$ for $\frac{1}{2} \leq t < 1$, which is called the *Haar wavelet*. Given the dilation and the wavelet equation the scaling and wavelet coefficients in

Equation 2 have the following recursive structure:

$$c_{j,k} = \frac{1}{\sqrt{2}}(c_{j+1,2k} + c_{j+1,2k+1}), \qquad (3)$$

$$d_{j,k} = \frac{1}{\sqrt{2}}(c_{j+1,2k} - c_{j+1,2k+1}). \qquad (4)$$

An attractive feature of wavelet-based analysis lies in the fact that the long range dependent, non-stationary process turns into stationary, nearly uncorrelated or short range dependent wavelet coefficients $d(j,k)$ [2, 1]. In the case of scaling the energy of these coefficients is power law dependent on the scale j, denoted by

$$\frac{1}{n_j} \sum_{k=1}^{n_j} |d_{j,k}|^2 \propto 2^{j\alpha}. \qquad (5)$$

This property leads to a wavelet-based scaling exponent estimation tool called the *Logscale Diagram* [1], which is used extensively in our experimental studies.

The multifractal wavelet model (MWM) is capable of generating *stationary*, *positive*, and *multifractal* processes with non-homogeneous scaling. The MWM synthesis procedure resembles the recursive structure of computing the scaling and wavelet coefficients. By arranging Equation 3 and 4 to

$$c_{j+1,2k} = \frac{1}{\sqrt{2}}(c_{j,k} + d_{j,k}), \qquad (6)$$

$$c_{j+1,2k+1} = \frac{1}{\sqrt{2}}(c_{j,k} - d_{j,k}), \qquad (7)$$

one constraint can guarantee the positivity of the process:

$$|d_{j,k}| \leq c_{j,k}. \qquad (8)$$

A multiplicative model can be built that automatically satisfies constraint (8) by defining $d_{j,k} = A_{j,k} \times c_{j,k}$ with $A_{j,k} \in [-1, 1]$. The recursive structure can be applied to generate data samples: the finest-scale scaling coefficients form the output MWM process.

A key characteristics of MWM is that the correlations and fractal behavior of the output process can be controlled by the wavelet energy decay of the data. This property is essential to reproduce the scaling behavior of the original process. A simple way to control energy decay is to fix the energy at the coarsest scale ($j = 0$) and set the ratios of energy for other scales with $R_j = \frac{var(d_{j-1,k})}{var(d_{j,k})}$. For a stationary LRD 1/f process, we can see from Equation 5 that $R_j = 2^\alpha$ is a constant. For the real world data it is shown in [12] that

$$R_j \propto \frac{var(A_{(j-1)})}{var(A_{(j)})(1 + var(A_{(j-1)}))}, \qquad (9)$$

and this recurrence can be solved recursively. We can see that the wavelet energy decay can be controlled by the multipliers $A_{(j)}$, to be exact, the probability densities for $A_{(j)}$. For details about the choices of $A_{(j)}$, the data fitting procedure, and the multifractal analysis of MWM we refer to [12]. In our experimental studies we use a symmetric beta (β) distribution for $A_{(j)}$. The model is therefore called β *multifractal wavelet model* (βMWM)[3].

---

[3]The Matlab codes for MWM can be obtained from Digital Signal Processing group at Rice University, which is available at `http://www-dsp.rice.edu/software/mwm.shtml`.

5. CONVERSION FROM RATES TO INTER-ARRIVALS

Although correlations and the scaling behavior can only be reliably revealed using the count/rate process, it is necessary to generate a point process in the form of interarrival times so that a full description can be obtained for modeling purposes. A simple method of transforming a rate function into interarrivals is the *integrate-and-fire* (InF) algorithm. The InF algorithm generates an event each time the integral of the rate $\mu(t)$ reaches a value of unity. It then resets the integrated value to zero whereupon the process begins anew, so the $(k + 1)$st event can be obtained from

$$\int_{t_k}^{t_{k+1}} \mu(t)dt = 1. \qquad (10)$$

This is a direct conversion from a rate process to a point process therefore the stochastic and fractal nature is completely determined by the rate process.

A more sophisticated method derived from above is the so-called *controlled-variability integrate-and-fire* (CV-InF) algorithm [15]. After generating the event t_{k+1} according to Equation 10, the $(k + 1)$st interarrival time ($t_{k+1} - t_k$) is multiplied by a Gaussian random variable with zero mean and variance σ^2. Therefore t_{k+1} is now replaced by

$$t_{k+1} + \sigma(t_{k+1} - t_k)\mathcal{N}(0, 1). \qquad (11)$$

CV-InF introduces a second source of randomness that can be specified and controlled via σ, which is independent from the rate process. Within the limit $\sigma \to 0$ CV-InF turns into the standard integrate-and-fire (InF) algorithm. Within the limit $\sigma \to \infty$, on the other hand, it leads to a homogeneous Poisson process and none of the stochastic nature of the rate process will be preserved. As σ increases from zero, the fractal characteristics of the rate process is progressively lost. A small σ value (compared to the average interarrival time) is desirable if one want to preserve the fractal behavior of the rate process at the same time introduce certain randomness in the interarrival process.

6. THE ADDITIVE NATURE OF RATES

Given the same count interval T, the rate processes can be added together to create an aggregated rate process. This additive nature of rates is very attractive from a modeling perspective. It suggests that the whole arrival process can be divided into rate processes by Virtual Organizations (VOs), users, or patterns, being modeled individually, and aggregated back to form a whole unified process. The VO or user names can be included in the synthetic traces, which is valuable for scheduling studies that take VO/user policies into account [4]. Distinctive patterns, such as pseudo-periodicity and long range dependence, have been identified for VOs both at the cluster and Grid level [6]. This indicates that VO is an appropriate level for modeling different patterns. The rate representation not only preserves the correlation structures of the underlying arrival process but also enables aggregation, which is not possible with interarrivals. Section 8 uses an example of VO aggregation to illustrate the additive nature of rates.

7. MODEL-DRIVEN SIMULATION

We simulate the Grid resources using GridSim [3], which is a discrete event simulator based on *simjava*. The first setup

Trace	Location	Architecture	Scheduler	#CPUs	Period	VOs	#Jobs
LCG1	Grid wide	clusters	Grid Broker	~30k	Nov 20-30, 2005	cms, atlas	188,041
LCG2	Grid wide	clusters	Grid Broker	~30k	Dec 19-30, 2005	cms, atlas	239,034
NIK05	NIKHEF, NL	PC cluster	PBS/Maui	288	Sep - Dec, 2005	atlas	63,449
LPC05	LPC, FR	PC cluster	PBS/Maui	140	Feb - Apr, 2005	biomed	71,271
RAL05	RAL, UK	PC cluster	PBS/Maui	1000	Oct - Nov, 2005	hep1, atlas	332,662

Table 1: Summary of workload traces used in the experimental study.

is a single space-shared cluster with one queue and multiple computer nodes. The scheduling algorithm is first-come first-served (FCFS). The second setup consists of multiple FCFS-based clusters and a single Grid resource broker. The algorithm used by the broker is called Minimum Completion Time (MCT). MCT assigns each incoming job to the cluster with the minimum expected completion time for that job. The clusters are assumed not empty and have their own local background workloads.

The Grid-level as well as background workloads are generated by models. We consider that a job is an independent computationally-intensive task requiring one single CPU. Since our goal is to investigate the performance implications of correlations in the arrival process, we model the job run time as an exponentially distributed random variable. Job arrivals, on the other hand, require three different models to produce the desired correlation structures. The first model is a Poisson process, which has no dependency (NoD) in the time series. The second model is a 2-state Markov moderated Poisson process (MMPP2) and it is considered short range dependent (SRD). The third model is the one presented in this paper for modeling long range dependence (LRD). Firstly the rate process is generated using the multifractal wavelet model, then it is being transformed into interarrivals via the CV-InF algorithm. The performance metrics used are *slowdown* and *utilization*. *Slowdown* is defined as the average job response time (run time plus queue wait time) divided by the average job run time. *Utilization* means the mean system utilization and it is calculated as the proportion of system's resources which are busy. Analysis of the simulation results are presented in the following section.

8. EMPIRICAL EVALUATION

In this section we present experimental studies using real workload data which exhibit long range dependent and multifractal arrival patterns. Table 1 presents a summary of workload traces under study. *LCG1* and *LCG2* are two traces from the LHC Computing Grid. The LCG production Grid currently has approximately 180 active sites with around 30,000 CPUs and 3 petabytes storage, which is primarily used for high energy physics (HEP) data processing. There are also biomedical applications running on the Grid. Almost all the jobs are computationally-intensive tasks, requiring one CPU to process a certain amount of data. The workloads are obtained via the LCG Real Time Monitor[4] for two periods: *LCG1* consists of jobs of eleven consecutive days from November 20th to 30th in 2005, while *LCG2* is

from December 19th to 30th in the same year. *cms* and *atlas*, two HEP experiments running on LCG, are the two main VOs whose arrival processes exhibit long range dependence. At the cluster level we use traces from three data-intensive clusters, namely, *NIK05* and *LPC05*, and *RAL05*. They are located at the HEP institutes in the Netherlands, France, and UK, respectively. All of these sites participate in LCG, however, it should be noted that these clusters are involved in multiple different collaborations and have their own local user activities. Arrival processes of *atlas* on *NIK05*, *RAL05* and *biomed* on *LPC05* are long range dependent, while *hep1* on *RAL05* shows multifractal (biscaling) behavior. Since no single statistic is able to completely characterize a point process, we assess the goodness of fit based on a range of statistical measures reflecting the first order and the second order properties.

8.1 Model Fitting

As an representative example for LRD, the fitting results for *cms* on *LCG1* are shown. Plotting the data series will give a good visual indication of the fitting. As is shown in Figure 1, for the rate process we can see that the fitted MWM model visually resembles the trace data. The interarrival process obtained by the CV-InF algorithm ($\sigma = 0.1$), on the other hand, is visually similar to the real data as well in terms of amplitude burstiness. By examining the complementary cumulative distribution function (CCDF) in Figure 2, we can see that the model is able to approximate the distributions of real data for both rates and interarrivals. Figure 2 also shows the autocorrelation function (ACF) and the Logscale Diagram. The autocorrelations in the rate process are excellently reconstructed by the MWM model. It is clearly observed that the interarrival process is weakly-correlated, which empirically proves that the correlation structures can only be reliably revealed by the count/rate process. The synthetic interarrival process is indeed short range dependent, matching the correlations of the original process very well. As to the second-order scaling exponents calculated by the Logscale Diagram, we can see that the scaling behavior (long range dependence) is almost perfectly reproduced by the MWM model with a scaling exponent $\alpha \approx 1.1$. More fitting results for long range dependent processes such as *atlas* on *LCG1* and *biomed* on *LPC05* are not shown here due to space limits. Generally speaking the autocorrelations and the scaling behavior can be well reconstructed by MWM modeling and the marginal distribution of interarrivals can be approximated by the synthetic process converted by the CV-InF algorithm.

Figure 3 shows the fitting results for *hep1* on *RAL05* with a multifractal (biscaling) arrival behavior. At $scale$[5]$= 6$ the rate process is short range dependent with quickly vanishing

---

[4] The Real Time Monitor is developed by Imperial College London and it monitors jobs from most of the major Resource Brokers in LCG, therefore the data it collects is representative at the Grid level. `http://gridportal.hep.ph.ic.ac.uk/rtm/`.

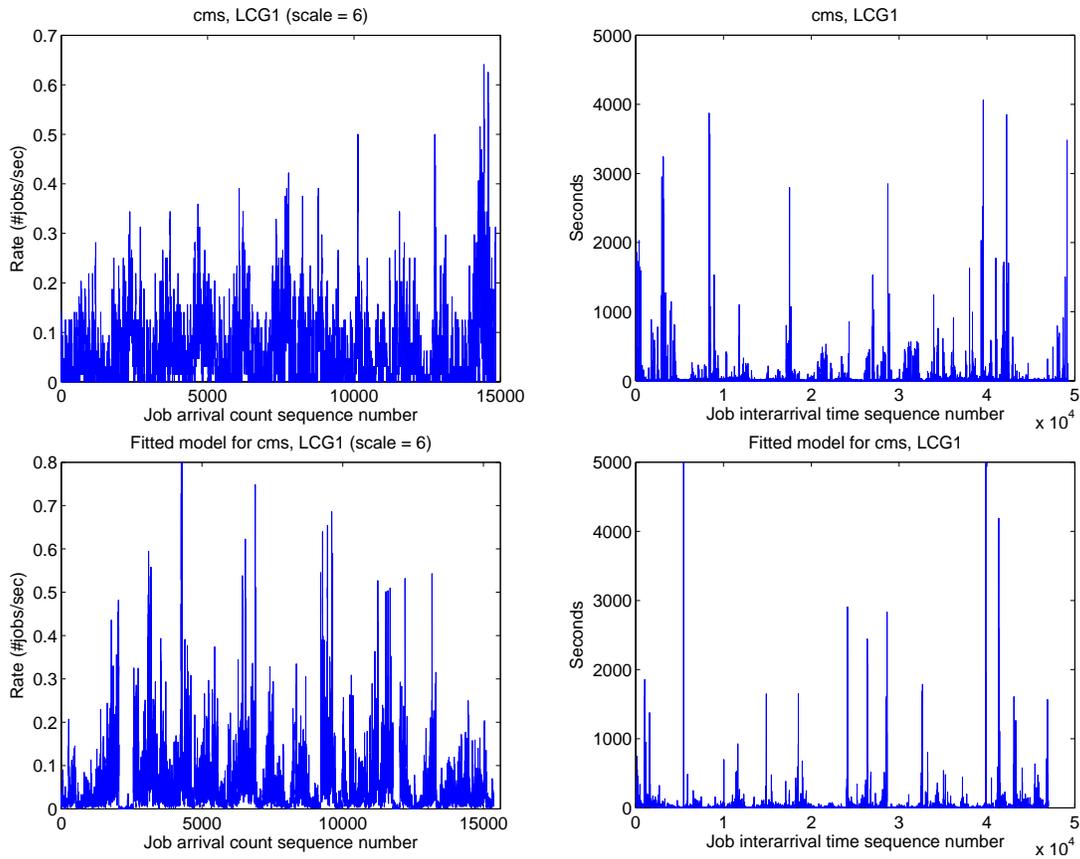[5] A dyadic scale is used so scale j means $T = 2^j$ seconds in the count/rate based measures.

Figure 1: Plotting the rate and the interarrival processes of the original data and the synthetic traces, respectively ($\sigma = 0.1$ in the CV-InF algorithm).
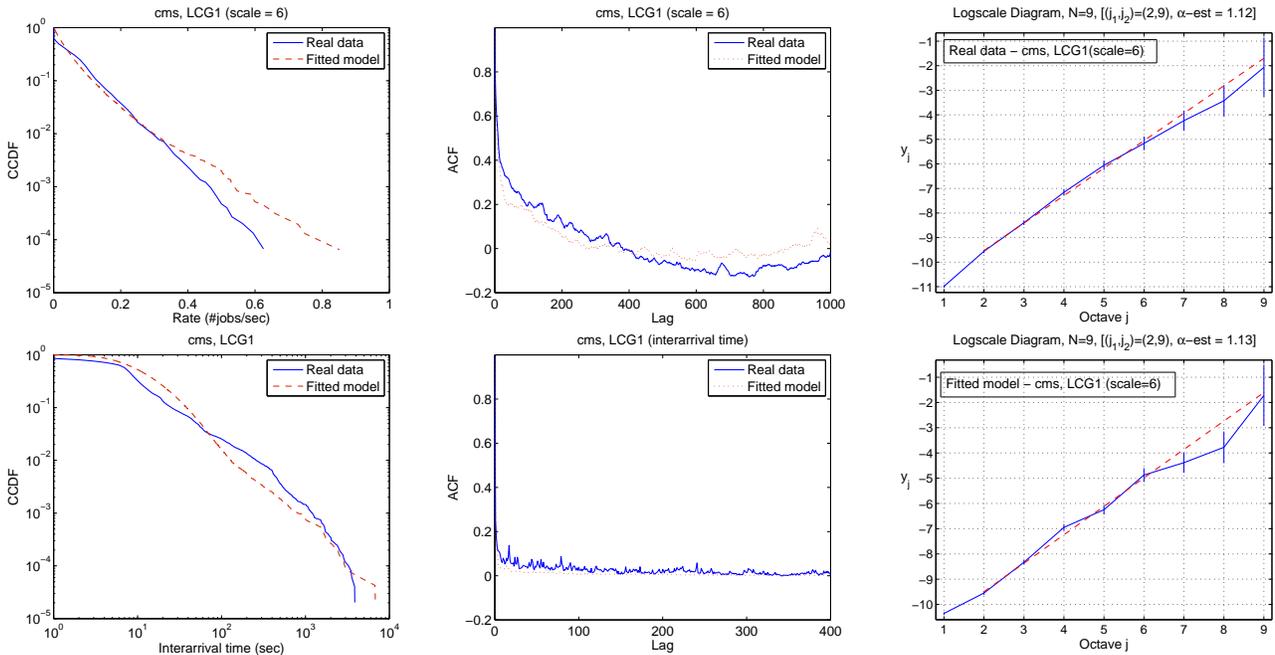

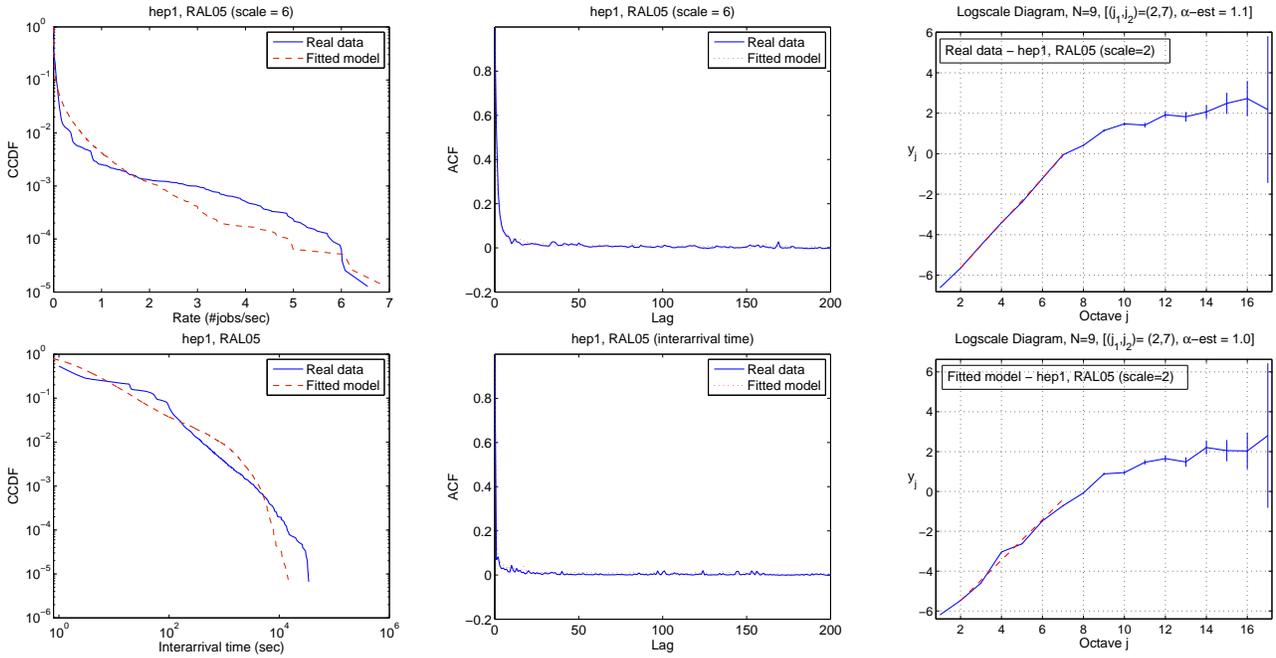Figure 2: βMWM fitting results for *cms* on *LCG1* ($\sigma = 0.1$ in the CV-InF algorithm).

Figure 3: β**MWM fitting results for** *hep1* **on** *RAL05* ($\sigma = 0.1$ **in the CV-InF algorithm).**

autocorrelation lags. It becomes interesting when we look at the scaling behavior in the Logscale Diagram. There are two alignment regions: in the smaller scale region ($scale < 7$) it indicates the fractal nature of the sampling path. In the larger scale region ($scale > 7$), on the other hand, the close-to-horizon alignment suggests weak autocorrelations [15]. It is clearly observed that the biscaling behavior can be well reconstructed by the MWM model. The success of MWM lies in its inherent structure of controlling the wavelet energy decay. If we recall the Equation 9 in Section 4, the energy decay ratios of wavelet coefficients are approximated from the data scale by scale. By this way it is possible for the MWM to reproduce a range of scaling behavior, including exact self-similarity, long range dependence, and multifractals.

8.2 VO Aggregation of Rates

Figure 4 shows an example of VO aggregation to illustrate the additive nature of rates. The job arrival process of *lhcb* on *LCG1* is pseudo-periodic, which shows a sequence of equally-spaced peaks in the rate ACF plot. The job arrival process of *cms* on *LCG1* is long range dependent with a slowly decaying ACF lags. The addition of these two rate processes is shown as the solid line in the right part of Figure 4. This aggregated process embodies properties of both contributing processes: a slow-decaying rate ACF lags with equally-spaced peaks. We model the pseudo-periodic process using a matching pursuit approach [7] and the long range dependent process using the MWM model described above. The aggregation of synthetic data from the two models, showing as a dotted line in the right part of Figure 4, matches well with the aggregation of the original data. The additive nature of rates is very attractive in modeling the arrival processes. Different patterns can be modeled separately and added back to a whole unified process, in which desired attributes such as VOs and users can be included.

For a full model the CV-InF algorithm can be applied to convert a rate process to an interarrival process.

8.3 Performance Analysis of Correlations

Model	Parameters
Poisson	$\mu = 10$
MMPP2	$\sigma_1 = 0.04$, $\sigma_2 = 0.01$, $\lambda_1 = 8.0$, $\lambda_2 = 1.0$
MWM	$p = [3.3, 5.3, 6.6, 7.5, 6.7, 7.1, 4.8, 3.0, 2.2, 1.4]$, $\mu_c = 0.28$, $\sigma_c = 0.33$

Table 2: Model parameters used in experiments. MWM parameters are fitted using *biomed, LPC05.*

Following the experimental setup described in Section 7, we drive the simulation of scheduling using the synthetic traces generated by models. The model parameters are provided in Table 2. From the ACF plots of the rate processes shown in Figure 5 (left) we can see that the selected models generate the desirable correlation structures in job arrivals, namely, NoD (Poisson), SRD (MMPP2), and LRD (MWM). All the interarrival processes are scaled to maintain the same mean and CV for the comparison studies. The run time process is drawn from an exponential distribution with an average run time equal to 5,000 seconds. In the single cluster case we set the number of processors to 100. Performance is measured by the average job slowdown as a function of system utilization, which is shown in Figure 5 (middle). We can see that the impacts of autocorrelations is very large: the bigger the ACF, the worse the performance. Similar results have been reported in a clustered web server environment [19]. The cause of such performance degradation can be the high degree of temporal burstiness in a LRD process. Bursty arrivals, which is the opposite of smoothness (e.g. Poisson), result in a long queue of waiting jobs. Consequently it leads to much longer queueing delays (big-
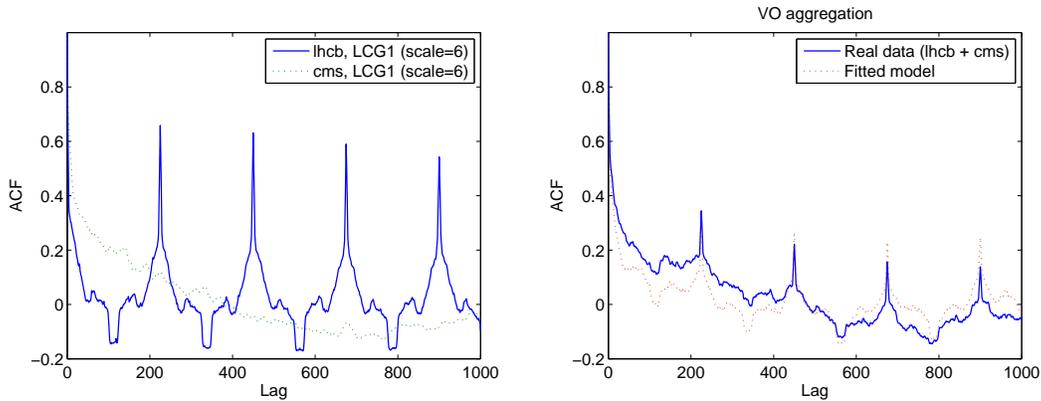
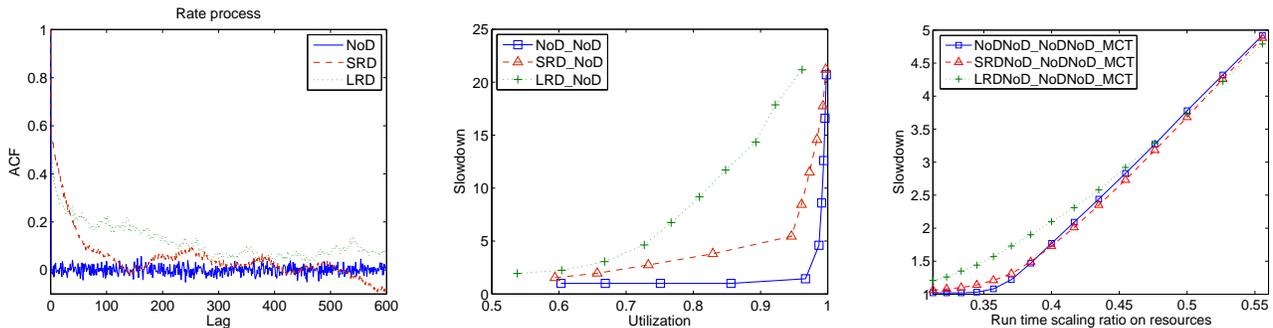Figure 4: VO aggregation of long range dependent and pseudo-periodic rate processes.

Figure 5: Performance analysis of correlations in job arrivals and its impacts in scheduling. In the single cluster case (middle) "NoD_NoD" represents "Arrival_Runtime", the correlation structures in the synthetic workloads. In the single broker case (right) "NoDNoD_NoDNoD_MCT" represents the correlation structures ("ArrivalRuntime") of the Grid workload, the local background workloads, and the broker algorithm, separated by "_". Run time scaling ratio is defined as the job MIPS rating versus resource MIPS rating.

ger slowdown for jobs) and overall lower system utilization. In the single broker case the performance is measured by the average job slowdown as a function of the run time scaling ratio on resources. The run time scaling ratio is the job MIPS rating versus resource MIPS rating and a higher ratio indicates a larger average run time. By varying the run time scaling ratio we get the curves as shown in Figure 5 (right). Although not as big as in the single cluster case, performance degradation is observed for large autocorrelations in the lower range of the scaling ratios. As the average run time increases the average slowdowns of all workloads converge. We can conclude that autocorrelations in the arrival processes results in worse system performance, both at the local and the Grid level.

9. CONCLUSIONS AND FUTURE WORK

In this paper we apply the Multifractal Wavelet Model (MWM) in modeling long range dependent and multifractal job arrival processes. These patterns are identified in data-intensive clusters and Grids by using the count/rate representation of the original processes. The second order properties such as the autocorrelation function and the scaling behavior can be well reconstructed by MWM modeling. A controlled-variability integrate-and-fire (CV-InF) algorithm is used to convert a rate process into an interarrival process so that the arrival process can be fully generated. Together

with the doubly stochastic models for middle to short range dependence [8] and the matching pursuit model for pseudo-periodicity [7], we are able to characterize and model the main patterns that are identified for the Grid job arrival processes. The additive nature of rates makes it flexible in model composition and aggregation. We further conduct simulation studies to quantify the performance impacts of correlated job arrivals in Grid scheduling and it is shown that autocorrelations can cause large performance degradations.

The effects of autocorrelations in system performance have important implications in the performance evaluation studies of Grid scheduling algorithms. Most of the previous work either use supercomputer traces, or make simplified assumptions on workloads (Poisson arrivals, fixed-interval arrivals, or constant background loads). We show that real workloads on production Grids exhibit long range dependence and have a high degree of temporal burstiness. Under such workload conditions the performance evaluation results can be highly different and it has been quantified in this paper. We argue that more realistic workloads should be taken into account in future Grid scheduling studies, and the modeling approach proposed in this paper provides one way of generating synthetic traces.

Future work includes a comprehensive study of performance impacts of workload correlations in Grid scheduling,

including both job arrivals and run times. How to improve scheduling under LRD and temporal burstiness is also a very interesting research direction. For a broader influence and adoption, especially in the Grid scheduling community, we are preparing to release tools and software developed for Grid workload modeling[6].

Acknowledgments

10. REFERENCES

[1] P. Abry, M. S. Taqqu, P. Flandrin, and D. Veitch. *Self-Similar Network Traffic and Performance Evaluation, K. Park and W. Willinger, editors*, chapter Wavelets for the analysis, estimation, and synthesis of scaling data. Wiley, 2000.

[2] P. Abry, D. Veitch, and P. Flandrin. Long-range dependence: revisiting aggregation with wavelets. *Journal of Time Series Analysis*, 19(3):253–266, May 1998.

[3] R. Buyya and M. Murshed. Gridsim: A toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. *Concurrency and Computation: Practice and Experience (CCPE)*, 14, 2002.

[4] C. Dumitrescu, I. Raicu, and I. Foster. Di-gruber: A distributed approach to grid resource brokering. In *proceedings of Supercomputing (SC)*, 2005.

[5] C. J. G. Evertsz and B. B. Mandelbrot. *Chaos and Fractals: New Frontiers in Science, Heinz-Otto Peitgen, Hartmut Jrgens and Dietmar Saupe, editors*, chapter Multifractal Measures, pages 849–881. Springer, New York, 1992.

[6] H. Li. Workload dynamics on clusters and grids. Technical Report 2006-04, Leiden Institute of Advanced Computer Science, Leiden University, short version appeared in IEEE IPDPS, 2006.

[7] H. Li, R. Heusdens, M. Muskulus, and L. Wolters. Analysis and synthesis of pseudo-periodic job arrivals in grids: A matching pursuit approach. In *proceedings of IEEE CCGrid*, 2007.

[8] H. Li and M. Muskulus. Analysis and modeling of job arrivals in a production grid. *ACM SIGMETRICS Performance Evaluation Review*, 34:59–70, 2007.

[9] S. B. Lowen and M. C. Teich. *Fractal-Based Point Processes*. John Wiley and Sons, Inc., 2005.

[10] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[11] R. Ranjan, A. Harwood, and R. Buyya. Sla-based cooperative superscheduling algorithms for computational grids. *ACM Transactions on Autonomous and Adaptive Systems*, 2007, to appear.

[12] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk. A multifractal wavelet model with application to network traffic. *IEEE Transactions on Information Theory*, 45(3):992–1019, April 1999.

[13] S. Song, K. Hwang, and Y.-K. Kwok. Trusted grid computing with security binding and trust integration. *Journal of Grid Computing*, 3:53–73, 2005.

[14] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.

[15] S. Thurner, S. B. Lowen, M. Feurstein, C. Heneghan, H. G. Feichtinger, and M. C. Teich. Analysis, synthesis, and estimation of fractal-rate stochastic point processes. *Fractals*, 5:565–595, 1997.

[16] S. Venugopal and R. Buyya. A set coverage-based mapping heuristic for scheduling distributed data-intensive applications on global grids. In *proceedings of the 7th IEEE/ACM International Conference on Grid Computing(Grid06)*, 2006.

[17] M. Wang, T. Madhyastha, N. Chan, S. Papadimitriou, and C. Faloutsos. Data mining meets performance evaluation: Fast algorithms for modeling bursty traffic. In *proceedings of 18th International Conference on Data Engineering*, 2002.

[18] G. W. Wornell. Wavelet-based representations of the 1/f family of fractal processes. *Proc. of IEEE*, 81(10), 1993.

[19] Q. Zhang, N. Mi, A. Riska, and E. Smirni. Load unbalancing to improve performance under autocorrelated traffic. In *proceedings of IEEE ICDCS*, 2006.

---

[6]Tools for workload analysis and modeling in production grids. http://www.liacs.nl/home/hli/gwm.