MSST 2010

Self-similarity in Parallel I/Os

Qiang Zou, Yifeng Zhu, Dan Feng

Huazhong University of Science and Technology, China
University of Maine

I/O Arrivals in Scientific Applications

- Understanding I/O behavior of parallel scientific applications is important for tuning, managing, or optimizing parallel file systems;
- F. Wang et al (Ref. [8]) examine the I/O burstiness of parallel I/O workloads using a simple methodology. They measure the cumulative distribution functions (CDF) of I/O inter-arrival times and conclude that *I/O activities in the LLNL traces are very bursty in the ior2 benchmark and the f1 application.*

Motivation

➢ Is it appropriate to use a Poisson or Markov model to characterize or predict parallel I/O arrivals with the presence of intensive burstiness in scientific applications?

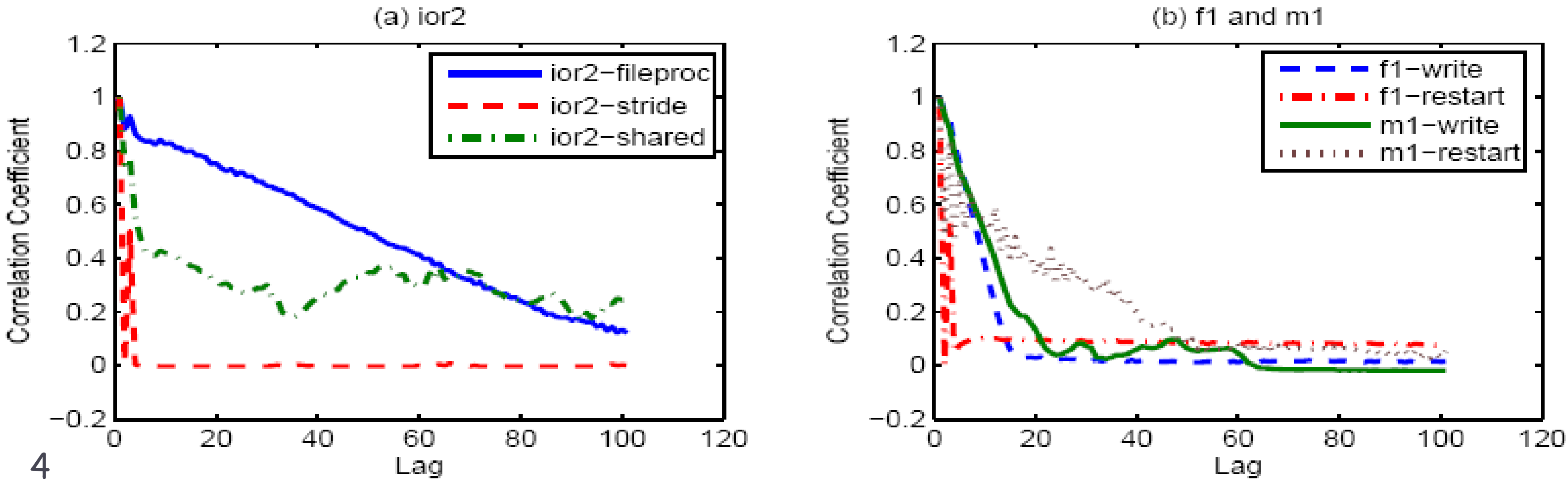
- ✓ In order to model parallel I/O workloads, prior studies usually assumed that *I/O arrival process follows a Poisson distribution*, and I/Os can be generated by using the Markov Model;
- ✓ However, the Markov model does not accurately characterize the ***burstiness*** in parallel I/O workloads;
- ✓ There are very bursty I/O activities as evidenced in the LLNL application workloads, such as the *ior2* benchmark and the *f1* application described in Ref. [8].

Auto-Correlation of Inter-arrival Times

- We use ***auto-correlation functions*** (ACF) to measure if earlier values in a time sequence have some correlation to later values

$$ACF(k) = \frac{c_k}{c_0} \quad \text{where} \quad c_k = \frac{1}{N-k}\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})$$

- It might be appropriate to use Markov model to synthesize *ior2-stride* workload, but not *ior2-fileproc*, *ior2-shared*;
- Examination results above motivate us to further study the self-similarity of parallel I/Os.



Self-similarity Detection

- A covariance stationary stochastic process is self-similar if

$$\lim_{k\to\infty}\frac{ACF(k)}{k^{-\beta}} = c < \infty, \quad \text{for } 0<\beta<1$$

- Second order self-similar $ACF(k) = \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}]$
- Hurst parameter, $H = 1-\beta/2$, measures of the degree of self-similarity.
- Use three methods to estimate H: variance-time plot (VTP), R/S, and Whittle estimator
- Most of the Hurst exponent values are above 0.5 except the I/O events in ior2-stride. This observation indicates the comprehensive *existence of self-similarity* in the LLNL I/O traces studied in this paper.

Traces	Estimation of H				
Streams	1	2	3	4	5
ior2-fileproc	0.67	0.66	0.65	0.69	0.70
ior2-shared	0.72	0.78	0.83	0.84	0.86
ior2-stride	0.52	0.49	0.45	0.40	0.41
f1-write	0.55	0.56	0.54	0.58	0.56
f1-restart	0.51	0.59	0.52	0.50	0.64
m1-write	0.66	0.67	0.67	0.65	0.66
m1-restart	0.67	0.65	0.66	0.64	0.63

Modeling I/O Arrivals

- We find that the parallel I/O traces studied meet the major requirements of a widely used model named *Fractional Brownian Motion.*

- The prediction algorithm can be expressed as

$$\lambda^{2H-1} \sim \frac{-2\alpha \ln(\varepsilon)((1-H)^{1-H}H^{H})^{2}b^{2H-2}}{[b(1-\varepsilon)-\varepsilon]^{2H}}.$$

- Our prediction model can be briefly described as right.

ARRIVAL-RATE-PREDICTION

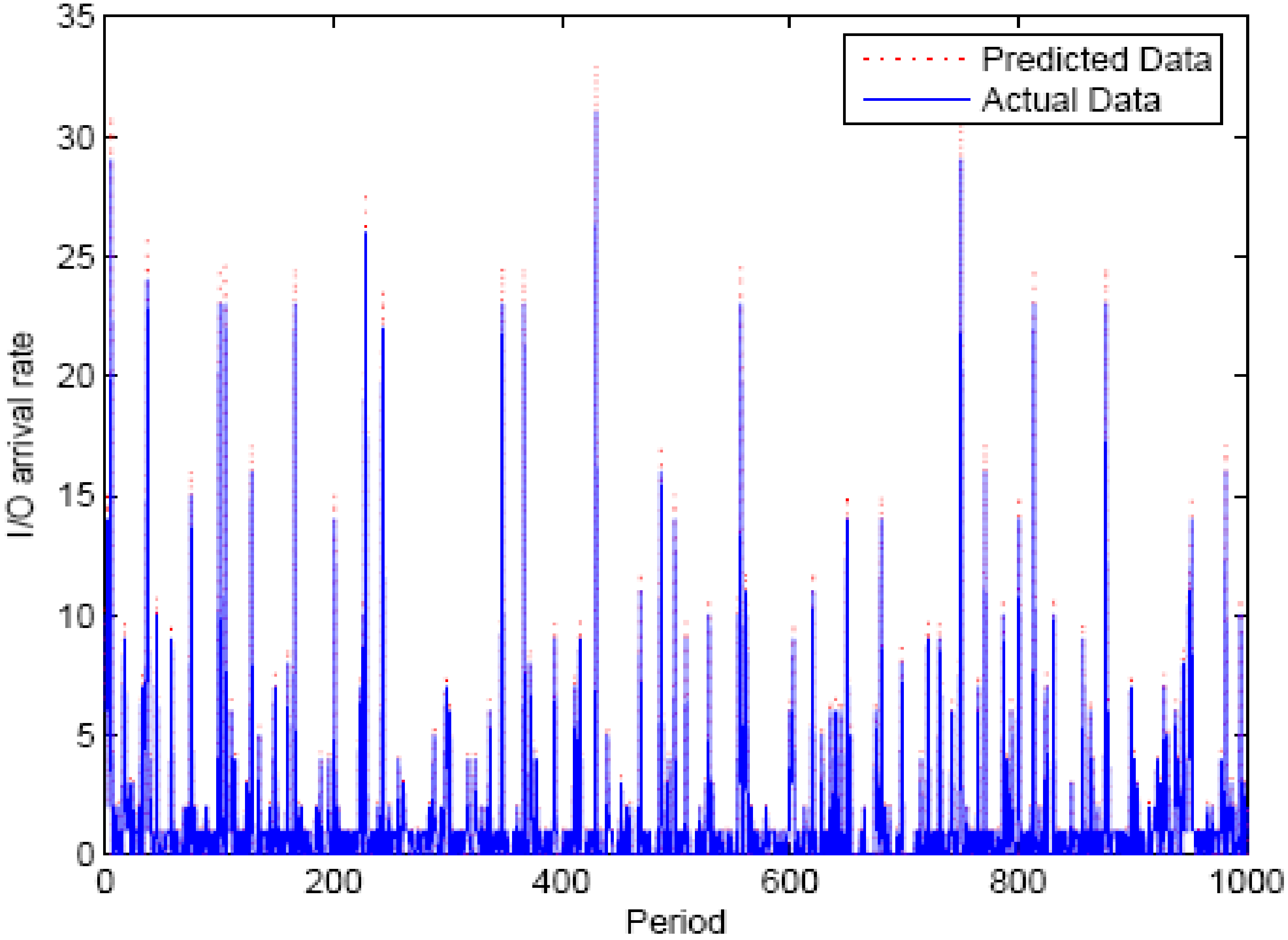
INPUT: The I/O miss probability ϵ, the maximum number of outstanding streams served at per node b, original trace data file f.
OUTPUT: arrival rate $\{\lambda(i); i = 1, 2, \ldots, n\}$.
ALGORITHM:
for each f
  Use maximum-likelihood estimate to estimate the parameter value α for data sets in f;
  Use Pox plot to estimate the Hurst value H
   if $H \notin (0, 1)$ or $H = 1/\alpha$
    then break;
   else
    Set the initial values of ϵ and b, and obtain $\{\lambda(i); i = 1, 2, \ldots, n\}$ using Equation (5)
end for

Evaluation of Our Model

- ➢Our model *vs* realistic LLNL I/O traces

- ➢Results of *one randomly selected stream* in the figure right.

- ➢Both sequences have exactly the same mean arrival rate.



Conclusions

- There are *evident correlations between inter-arrival times* in subtraces collected on most computing nodes.
- Scientific I/O workloads are self-similar for short-term scales. Thus traditional *Poisson or Markovian arrival processes are inappropriate* to model the I/O demands.
- We *develop an accurate analytical model* to model I/O mean arrival rate for I/O workloads with self-similarity.
- Our immediate future work is to collect parallel I/O traces lasting weeks or months and further evaluate self-similarity.