Proceedings of the 8th WSEAS International Conference on APPLIED MATHEMATICS, Tenerife, Spain, December 16-18, 2005 (pp55-58)

Poisson counting process with variable servers' number.

A. SURKOV
Sankt-Petersburg University
7/9, University emb., Sankt - Petersburg
RUSSIA

Abstract: - In this article there is an attempt to investigate a counting process with variable number of Poisson working nodes producing output. Servers' quantity here is a stochastic variable. Some formulae for the counting process are received and compared with well known results of corresponding evaluations using degenerate probability as server distribution. In calculation virtual time concept was used as time mapping (shrinking) from system with variable quantity nodes to single server one.

Key-Words: Queue system, Poisson process, variable servers' number, virtual time

1 Introduction

Modern distributed systems middleware offers resource hiring from others cites which may belong to other organizations. For example, this is a good practice in GRID environments [1]. One can build virtual organization using borrowed resources. Well known example in business life is outsourcing. Computational resources in such environments can be offered or revoked at arbitrary time what makes these systems mutable. At another point of view, system designer often is forced to apply dynamical hiring policy of some kind as a result of nonzero cost in general case of borrowed resources. These and other reasons are supporting appearance of systems with variable servers' number. So in near future systems with variable servers' quantity seems to be the common case.

One can often see mass production using number of similar machines. These machines in real life can sometimes stop its work due to any reasons, but production is still going. Or master may decide start extra machines in a rush times.

These and others similar situations derive models where a system produces items (goods, events e. t. c.) and this production is based on a group of identical producing nodes (servers). It can de modeled with a counting process. To represent effects of adding or removing servers author suggests to consider counting process made of a group of identically distributed servers varying in its quantity. We assume that these servers have Poisson distribution due to its simplicity.

1.1 Related works

The most close to the theme seems investigations of systems with variable parameters: compound (for example [2]) process, Cox processes (for example [3]), mixed processes, phase processes e. t. c. We'll look at them here from time shrinking point of view.

Analyze of systems with variable servers' number are performed in many scientific fields. The Chord system (for example [4]) allows nodes join and leave system at random. Fault tolerant systems are widely discussed. But all these papers investigate its specific questions (content preservation in Chord, availability with fault tolerance e. t. c.) without looking at nodes number dependent queuing dynamic.

Idea of state-dependent time in queuing is used, for example, in stability analyze [5]. But it differs from time–shrinking approach used in the paper.

2 Problem Formulation

In the paper the counting process is named as $A(t)$, and it is made of independent identical distributed Poisson service processes $X_i(t)$, with mean μ. Servers' quantity is provided with random variable U with mean τ, variance σ, probability distribution function F, and probability generation function Y.

The problem is to obtain some characteristic for the counting process (probability distributions, queue mean and variance) and show that they are correspond to known ones in suitable conditions. We'll compare achieved results with known properties of Poisson processes using degenerate probability distributions of their quantity U and put some generalizations.

Because of Poisson nature of the service process, we don't pay any attention to entering and/or leaving service node's behavior within entering/leaving intervals and at its boundaries.

3 Problem Solution

3.1. Virtual time within a process

Probability and time are close bound:

$$u_i = T_i / T, \quad T = \sum_{i \in E} u_i T = \sum_{i \in E} T_i \qquad (1)$$

Here $\{T_i\}$ are the time slices while the system stays in one of disjoint system states indexed with i, and u_i is the probability of the state. E is a union of all possible states.

Let U is time independent. Let these slices are indexed with servers' number within it (i-th epoch has i servers). So the process within time slices $\{T_i\}$ can be seen as usual i-server Poisson process. We may combine several epochs with the same servers' number into single one and interchange such epochs due to Poisson's probability distribution memory less nature of nodes' output.

<u>Proposition 1.</u> The throughput of independent identical distributed Poisson i-servers process can be modeled with a single server Poisson process with a special timing, which is i times longer then real (hereafter referred to as 'virtual time'):

$$T_i^v = i * T_i. \qquad (2)$$

For example, formally one can get this effect with simple variable substitution like following in some distribution function H:

$$H_i(t) = 1 - e^{-i\mu t},\ it \to t^v \Rightarrow$$
$$H_i(t) = H_1(t^v) = 1 - e^{-\mu t^v},$$

When $i = 0$, then no work is done and no increment of the counting process' output is expected due to absence of anything capable to perform it. Hence such epoch need not to be modeled. So received zero modeling (virtual) tine seems to be valid.

For virtual time like (1) and using (2) one can get:

$$T^v = \sum_{k=1,+\infty} T_k^v = \sum_{k=1,+\infty} kT_k = \sum_{k=1,+\infty} k(u_k T) = \qquad (3)$$
$$= T \sum_{k=0,+\infty} k u_k = TE(U) = \tau T.$$

So we found that virtual modeling single server system time is E(U) times larger than real one in variable servers' number system.

Now let $U = U(t)$ is time dependent. We have to fragment T^v and T into related intervals $[T^v, T^v + dT^v)$, and $[T, T + dT)$ due to unlimited Poisson's divisibility. Then one can use (3) for every fragment with stable $U(t)$ within it to calculate virtual time: $dT^v(t) = \tau(T)dT$. The result is obtained aggregating all intervals using integration:

$$T^v = \int_o^T \tau(t)dt.$$

If one try to formally insert this virtual time as a process time into single server's Poisson probability generation function, then (this method doesn't consider any other probability characteristic of U except mean, because this way this information is lost):

$$g(z,t) = \exp[-\mu(1-z)\int_o^t \tau(\xi)d\xi],$$
$$E(A(t)) = \frac{\partial}{\partial z} g(z,t)\Big|_{z=1} = \mu \int_o^t \tau(\xi)d\xi$$
$$V(A(t)) = \left[\frac{\partial^2}{\partial z^2} + \frac{\partial}{\partial z} - \left(\frac{\partial}{\partial z}\right)^2\right] g(1,t^v) =$$
$$= \mu \int_o^t \tau(\xi)d\xi.$$

As we'll see later, the mean is correct but the variance doesn't account for servers' variance (see (8)).

3.2 Probability generation function for stationary servers' number distribution

Let F is servers' number probability distribution function: $F(x) = P\{U \le x\}$. As probability distribution function, F is monotonic, and $F:[0,\infty) \mapsto [0,1]$.

Let's go from 0 to infinity along x with distribution function $F(x)$. In integer case we'll get stepping function with steps corresponding to u_i in (1) at integer points i corresponding to i-servers state. For example, single step function at $x = s$ makes deterministic distribution corresponding to s-servers system. Looking at (3) we came to proposition 2.

<u>Proposition 2</u>: Function $F:[0,\infty) \mapsto [0,1]$ maps virtual single server system time to unity interval of real system time.

Indeed:

$$[0,1] = \sum_{i=0}^{\infty} u_i = \sum_{i=0}^{\infty} dF(i)$$

There first equality follows from probability norm, and second one follows from definition F(x) in integer case.

In equilibrium state for a single server Poisson process the probability to get i counts within all virtual time scale (which corresponds unity real time interval) we may find as:

$$A_i = \int_o^{\infty} \frac{(\mu\xi)^i}{i!} e^{-\mu\xi} dF(\xi).$$

Then probability generation function will be:

$$G(z) = \eta(\mu(1-z)). \qquad (4a)$$

Here $\eta(s)$ is Laplas-Stiltjes transform of F(t).
For real time interval t we use ξt instead of ξ above and get:

$$G(z,t) = \eta(\mu(1-z)t). \qquad (4b)$$

So the mean and variance of output production are:

$$E(A) = \frac{\partial}{\partial z} G(z)_{z=1} = \frac{\partial}{\partial z} \int_o^{\infty} e^{-\mu\xi(1-z)t} dF(\xi) \Bigg|_{z=1} \qquad (5)$$

$$= \mu t \int_o^{\infty} \xi dF(\xi) = \mu\tau t$$

$$V(A) = (\mu t)^2 \int_o^{\infty} \xi^2 dF(\xi) + \mu\tau t - (\mu\tau t)^2 =$$

$$= \mu^2 t^2 \sigma^2 + \mu\tau t$$

Here σ^2 is the servers' distribution variance.

Let's build generation function another way. Because of theorem of total probability:

$$P\{A(t) = i\} = \sum_{k=0}^{\infty} P\{A(t) = i \mid U = k\} P\{U = k\}.$$

If there are only k servers in the system then:

$$P\{A(t) = i \mid U = k\} = P\{\sum_{l=1,..,k} X_j(t) = i\}.$$

Corresponding to k-servers system $P\{A(t) = i \mid U = k\}$ probability generation function $G_k(z,t)$ and its coefficients $g_{k,j}(t)$ are calculated as:

$$P\{A(t) = j \mid U = k\} = g_{k,j}(t).$$

$$G_k(z,t) = \sum_{j=0}^{\infty} g_{k,j}(t) z^j = [h(z,t)]^k.$$

Later equality follows from the fact that generation function of k-convolution is k-power of convoluted function $h(z,t)$ — generation function for $X_i(t)$, $h(z,t) = e^{-\mu(1-z)t}$.
Finally these formulae make a compound distribution:

$$G(z,t) = \sum_{i=0}^{\infty} P\{A(t) = i\} z^i = \qquad (6)$$

$$= \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} P\{A(t) = i \mid U = k\} P\{U = k\} z^i =$$

$$= \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} g_{k,j}(t) u_k z^i = \sum_{k=0}^{\infty} u_k \sum_{i=0}^{\infty} g_{k,j}(t) z^i =$$

$$= \sum_{k=0}^{\infty} u_k [h(z,t)]^k = Y(h(z,t))$$

Here Y(z) is a generating function corresponding to U. $G(z,t) = Y(e^{-\mu(1-z)t})$. It corresponds (4b) using relation between Poisson probability generation function and Laplas-Stiltjes transform for stochastic intervals.

3.3 Varying in time servers' distribution

Let's now $U = u(t)$ is a function of time.

Performance of a Poisson counting process is proportional to its runtime. Therefore we have to calculate its relative length:

$$U_i(T) = \frac{1}{T} \int_o^T u_i(t) dt.$$

It may be considered as constant distribution (for given T) and one can insert it into (6):

$$G(z,T)=\sum_{k=0}^{\infty}\left(\frac{1}{T}\int_{o}^{T}u_i(t)dt\right)[h(z,T)]^k = \qquad (7)$$
$$=\frac{1}{T}\int_{o}^{T}Y(h(z,T),t)dt \, ,$$

Here (for Poisson) $Y(z,t)=\sum_{j=0}^{\infty}u_j(t)z^j$, so $G(z,T)=\frac{1}{T}\int_{o}^{T}Y(e^{-\mu(1-z)T},t)dt$

Using the fact that

$$\left.\frac{\partial}{\partial z}Y(e^{-\mu(1-z)T},t)\right|_{z=1} = \mu T e^{-\mu(1-z)T}\Big|_{z=1}\left.\frac{\partial Y(x,t)}{\partial x}\right|_{x=1}$$

(here $x=e^{-\mu(1-z)T}$), one can get:

$$E(A(T))=\frac{\partial}{\partial z}G(z,T)_{z=1} = \qquad (8)$$
$$=\left.\int_{o}^{T}\frac{\partial}{T\partial z}Y(e^{-\mu(1-z)T},t)dt\right|_{z=1} = \mu\int_{o}^{T}\tau(t)dt \, ,$$
$$V(T)=\frac{1}{T}\left[\frac{\partial^2}{\partial z^2}+\frac{\partial}{\partial z}-\left(\frac{\partial}{\partial z}\right)^2\right]\int_{o}^{T}Y(e^{-(1-z)T},t)dt \,\Big|_{z=1}$$
$$=\mu^2 T\int_{o}^{T}\sigma(t)dt+\mu\int_{o}^{T}\tau(t)dt \, ,$$

One can see that the mean and variance are like (5):

4 Conclusion

4.1 Comparison with existing results

Fixed server counting process be achieved from variable servers' model using degenerate probability p^s:

$$p_j^s = \begin{cases}1, & j=s>0\\ 0, & j\neq s\end{cases}. \qquad (9)$$

For distribution in (9) formula (2) gives:

$$t^v = st \, .$$

The mean and variance coincide with known values for Poisson process are:

$$E(A(t)) = \mu t^v = \mu st \, ,$$
$$V(A(t)) = \mu t^v = \mu st \, .$$

Generation function for a counting process is the same as for s-server Poisson servicing process.

$$G(z,t) = Y^s(h(z,t),t) =$$
$$=\left(\exp(-\mu t(1-z)\right)^s = \exp[-\mu st(1-z)] \, .$$

These results correspond to well known for s-server Poisson processes.

4.2 Discussion

In this paper server number was controlled with a stochastic variable/process. One can use functional control instead and investigate system's behavior in control, decision or another theory's framework.
If working nodes are not Poisson then boundary behavior effects appear: we cannot group different epochs with the same servers' quantity together and swap epochs to order them in increasing servers' number.

References:

[1] I. Foster, C. Kesselman, S. Tuecke «The Anatomy of the Grid: Enabling Scalable Virtual Organizations», http://www.globus.org/alliance/publications/papers/anatomy.pdf.

[2] E. Gudzovska–Novak, S. Ritter, G. Taucher–Scholz, G. Kraft, Compound Poisson processes and clustered damage of radiation induced DNA double strand breaks, *ACTA physica polonica*, Vol. 31, No 5, 2000, pp. 1109–1123

[3] J. Moller, R. Waagepetersen, Statistical inference for Cox processes, www.math.auc.dk/~rw/papers/cox2.RR.ps, 2001, 31p.

[4] Liben-Nowell D., Balakrishnan H., Karger D. «Analyze of the evolution of Peer-to-Peer systems»; Proc. of the 21-th symposium on distributed computing, 2002.— pp. 233–242

[5] Foss S., Konstantopulos T., An overview of some stochastic stability methods