LABORATOIRE

INFORMATIQUE, SIGNAUX ET SYSTÈMES
DE SOPHIA ANTIPOLIS
UMR 6070

OPTIMIZING JOBS TIMEOUTS ON CLUSTERS AND PRODUCTION GRIDS

Tristan Glatard, Johan Montagnat, Xavier Pennec

Projet RAINBOW

Rapport de recherche
ISRN I3S/RR–2006-35–FR

Novembre 2006

LABORATOIRE I3S: Les Algorithmes / Euclide B – 2000 route des Lucioles – B.P. 121 –
06903 Sophia-Antipolis Cedex, France – Tél. (33) 492 942 701 – Télécopie : (33) 492 942 898
http://www.i3s.unice.fr/I3S/FR/

RÉSUMÉ :

Ce document présente une méthode pour optimiser la valeur du délai d'expiration (timeout) de tâches de calcul sur une grille. Il est fondé sur un modèle du temps d'exécution des tâches qui considère que la latence du système de gestion de tâches de la grille est une variable aléatoire. Le modèle prend aussi en compte un taux de panne pour modéliser soit des grappes fiables soit des grilles de production qui sont caractérisées par leur pannes. Les systèmes de gestion de tâches sont d'abord étudiés en considérant des distributions classiques de la latence. Plusieurs comportements surviennent en fonction du poids de la queue de la distribution de la latence et de la proportion de pannes. Nous présentons ensuite des résultats expérimentaux fondés sur des mesures de la distribution de la latence et du taux de panne sur la grille de production EGEE. Ces résultats montrent qu'utiliser la valeur optimale du délai d'expiration fournie par notre méthode réduit l'impact des pannes et peut conduire à une accélération de 1.36 sur des systèmes fiables.

MOTS CLÉS :

délai d'expiration, grilles de production, latence, variabilité

ABSTRACT:

This paper presents a method to optimize the timeout value of grid computing jobs. It relies on a model of the job execution time that considers the job management system latency through a random variable. It also takes into account a proportion of outliers to model either reliable clusters or production grids characterized by faults causing jobs loss. Job management systems are first studied considering classical distributions of the latency. Different behaviors are exhibited, depending on the weight of the tail of the distribution and on the amount of outliers. Experimental results are then shown based on the latency distribution and outlier ratios measured on the EGEE grid infrastructure. Those results show that using the optimal timeout value provided by our method reduces the impact of outliers and leads to a 1.36 speed-up for reliable systems without outliers.

KEY WORDS :

timeout, production grids, latency, variability

Optimizing jobs timeouts on clusters and production grids

Tristan Glatard[1,2], Johan Montagnat[1]
[1]I3S, CNRS, UNSA
{glatard,johan}@i3s.unice.fr

Xavier Pennec[2]
[2]INRIA, Asclepios
Xavier.Pennec@sophia.inria.fr

Abstract

This paper presents a method to optimize the timeout value of computing jobs. It relies on a model of the job execution time that considers the job management system latency through a random variable. It also takes into account a proportion of outliers to model either reliable clusters or production grids characterized by faults causing jobs loss. Job management systems are first studied considering classical distributions. Different behaviors are exhibited, depending on the weight of the tail of the distribution and on the amount of outliers. Experimental results are then shown based on the latency distribution and outlier ratios measured on the EGEE grid infrastructure[1]*. Those results show that using the optimal timeout value provided by our method reduces the impact of outliers and leads to a 1.36 speed-up even for reliable systems without outliers.*

1 Introduction

A growing number of distributed applications is relying on large scale workload management systems. These applications usually trigger hundreds, thousands or even more jobs. Although large scale systems provide very high throughput as a direct consequence of the huge amount of resources available, they also introduce *high* and *variable* latencies that drastically degrade the performances of the application competing with other users' computations. The latency corresponds to the duration between the submission date and the one at which the job execution effectively starts. It includes the submission, scheduling and queuing times but also data transfers and delays in the monitoring system. Furthermore, less reliable systems such as production grids are also impacted by *outliers*.

The variability impacts jobs latency in a *normal operation mode*. It mainly comes from the heterogeneity of the infrastructure (endogen hardware and software factors) and from the load imposed to it (exogen factor). In this document, we will model the system latency by a probabilistic distribution.

Outliers correspond to *system faults* that lead to huge latencies prevailing on the ones of the other tasks of the application. Those latency values can be considered as infinite.

Typical faults generating outliers are hardware failures, locally heavy load or scheduling errors leading to a job being queued in a extremely long queue. This outlier mode can be quantified by its proportion of jobs that never return.

[1] http://www.eu-egee.org/

Both variability and outliers penalize applications that rely on the completion of a high number of jobs. Indeed, a single job is then able to slow down the whole application. In case where there are dependencies between the jobs (*e.g.* in case of application workflows), the effects of variability and outliers are even more critical and lead to accumulated performance drops.

From the user point of view, strategies to reduce the impact of variability and outliers include multi-submission [1] (a given job is submitted many times and only the first completion is considered), jobs grouping (many jobs are grouped together to reduce the number of submissions) [2, 3] and timeouting [4, 5]. Multi-submission is an aggressive strategy and job grouping is not always possible. Timeouting and resubmitting abnormally long jobs is a common strategy. Choosing the timeout value is often let to the administrator or the end user. However, a non trivial trade off has to be found as a too long timeout will penalize the jobs completion time too much, while a too short one may be overkilling, causing the unnecessary resubmission of jobs that almost completed. Hence, timeout strategies have been designed in areas as different as TCP throughput optimization [6], HTTP requests [7, 8] or power saving devices [9].

To propose a timeout optimization, we first model jobs execution time in section 2. We then present in section 3 some results of timeout optimization on classical distributions. To show how the optimization behaves on a real infrastructure, we are particularly interested in the asymptotic behavior of the system and on the impact of outliers. We finally present in section 4 some experimental results from a distribution of the latency measured on the EGEE production grid. To facilitate legibility, many of the proofs of the theoretical results are deferred in appendix.

2 Model of the job execution time

We adopted a probabilistic modeling of the large-scale workload manager. This approach has already been successfully reported to tackle related scheduling problems [11, 12]. We will denote random variables with capital letters whereas fixed values will be lowercase. For a random variable V, f_V denotes it probability density function (pdf) and F_V denotes its cumulative density function (cdf).

Let J be the total duration of a job (including all its potential resubmissions) and t_∞ be a user defined timeout value. The system is seen as a black box introducing a positive latency R on the job wall-clock time r in case of normal operation. The outlier ratio is denoted ρ. r is assumed to be a fixed value depending only on the job nature whereas R is a random variable.

We denote with q the probability for a job to timeout. A job timeouts either if it is an outlier or if it faces a latency which is superior to t_∞. Thus:

$$\begin{aligned} q &= \rho + (1-\rho)P(r + R > t_\infty) \\ \text{and then,} \quad q &= 1 - (1-\rho)F_R(t_\infty - r). \end{aligned} \tag{1}$$

If the job timeouts, it is canceled and resubmitted. We neglect the cost of canceling and resubmitting a job as well as the resulting overload on the system, so that consecutive submissions are considered as independent. Let J_i be the duration of the job from its ith submission to its completion. J_i can be recursively defined as:

$$J_i = \begin{cases} r + R & \text{with probability } 1-q \\ t_\infty + J_{i+1} & \text{with probability } q. \end{cases} \tag{2}$$

For the sake of clarity, we will assume that $r = 0$. This hypothesis is not restrictive. In the rest of the equations, it corresponds to the variable change $u = t - r$. In case of real job executions, r would have to

be added to the timeout value. The goal is to express $J = J_1$ (through its cdf F_J) with respect to R and t_∞. J can only be superior to nt_∞ if the job timed-out n times. Thus:

$$P(J > nt_\infty) = q^n \quad \text{so that} \quad P(J < nt_\infty) = 1 - q^n. \tag{3}$$

We have the value of $F_J(x) = P(J < x)$ for every $x = nt_\infty$. We now have to obtain a complete expression of F_J. Interpolating F_J in every $[nt_\infty, (n+1)t_\infty]$ is clearly not suitable. Indeed, those intervals can be quite large with respect to the job execution time and the error coming from the interpolation is likely to produce inconsistent results. We can notice that $F_J(t)$ represents the probability for J to be inferior to t, so that **for all** t **in** $[nt_\infty, (n+1)t_\infty[$:

$$\begin{aligned} F_J(t) &= P(J < t | t \in [nt_\infty, (n+1)t_\infty]) \\ &= P(J < nt_\infty) + P(nt_\infty < J < t \mid t \leq (n+1)t_\infty) \end{aligned}$$

and thus, according to equation 3:

$$F_J(t) = 1 - q^n + P(nt_\infty < J < t \mid t \leq (n+1)t_\infty). \tag{4}$$

Given that $t \leq (n+1)t_\infty$, a job duration J can only be in $[nt_\infty, t]$ if the job timed-out n times (probability q^n) and if it succeeded on the $(n+1)^{\text{th}}$ attempt *i.e.* it was not an outlier (probability $1-\rho$) and $R \leq t - nt_\infty$ (probability $F_R(t - nt_\infty)$). Therefore, $P(nt_\infty < J < t \mid t \leq (n+1)t_\infty) = q^n(1-\rho)F_R(t - nt_\infty)$. We finally get, $\forall t \in [nt_\infty, (n+1)t_\infty[$:

$$F_J(t) = 1 - q^n + q^n(1-\rho)F_R(t - nt_\infty). \tag{5}$$

Given that $R>0$, $F_J(0) = 0$ and $\lim_\infty F_J = 1$. Moreover, F_J is continuous at every nt_∞. Indeed, according to equation 5, the expression of F_J at the lower bound of the segment $[nt_\infty, (n+1)t_\infty]$ is:

$$F_J(nt_\infty) = 1 - q^n + q^n(1-\rho)F_R(0) = 1 - q^n$$

Moreover, at the upper bound of this segment, F_J is:

$$\begin{aligned} F_J((n+1)t_\infty) &= 1 - q^n + q^n(1-\rho)F_R((n+1)t_\infty - nt_\infty) \\ &= 1 - q^n + q^n(1-\rho)F_R(t_\infty) \\ &= 1 - q^n + q^n(1-q) \text{ (given equation 1 with r = 0)} \\ &= 1 - q^{n+1} \end{aligned}$$

However, in general, F_J is not differentiable in $nt_\infty, \forall n$.

Note that if $\rho = 0$, then equation 5 resumes to:

$$F_J(t) = 1 - q^n + q^n F_R(t - nt_\infty) \quad \text{with} \quad q = 1 - F_R(t_\infty).$$

Thus, taking outliers into account corresponds to multiplying F_R by the $(1-\rho)$ factor.

2.1 Illustration for $\rho = 0$

If no outlier are present, the choice of a timeout value can be evaluated by comparing F_R and F_J. Figure 1 displays an example of cdf for R (plain red curve) and J (dashed green curve). Notice the singularities at nt_∞ points. On top, the distribution of R is Gaussian with a mean 300 seconds and standard deviation 100 seconds, truncated above zero to avoid negative latency times. The timeout value is equal to the mean of the original Gaussian (300 seconds). It is of course a very low timeout value leading to many resubmissions. We can graphically notice that for every t, $F_R(t) > F_J(t)$, which means that for every time value t, there is a higher probability that $R < t$ than that $J < t$. In this case, it would thus have been better not to set any timeout value as it highly penalizes the execution.

On the other hand, the bottom of figure 1 displays an example of cdf of R and J in a case where the timeout choice improves the execution. The timeout value is still 300 seconds but the law of R has a longer tail than in the former example. R is actually log-normal, with μ=15 seconds and σ=10 seconds. In this case, it seems that for every t, $F_R(t) < F_J(t)$, which means that the timeout improves the execution time in this case.

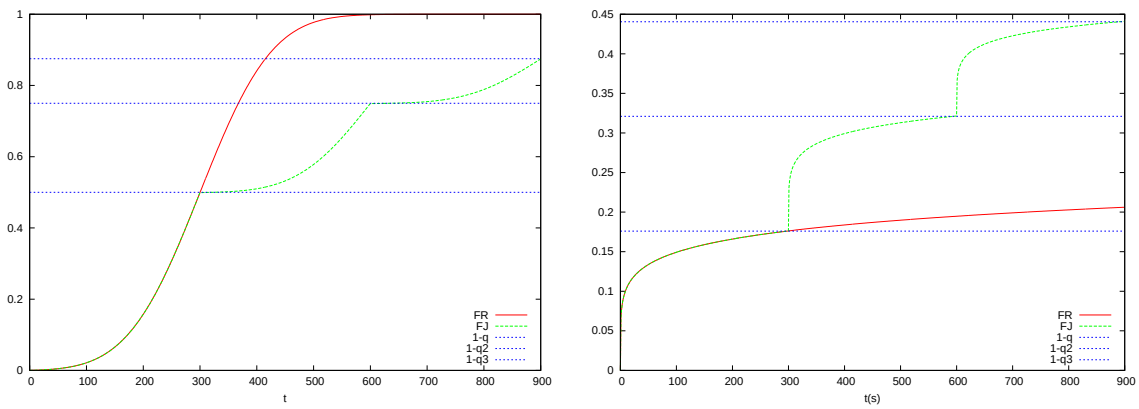

Figure 1. Example of cdf of R and J. Left: bad timeout choice. Right: good timeout choice.

As suggested by those graphical remarks, the impact of some timeout choices on the job execution time may be evaluated only by comparing F_J and F_R. However, apart from those particular cases, it is often not possible to have general results on the comparison between F_J and F_R at every time point t and the configuration displayed on figure 2 is observed. On this figure, the distribution of R is log-normal, with $\mu = 5.5$s, $\sigma = 1$s and a timeout value of 300s. In this case, minimizing the expectation of J with respect to t_∞ is a natural solution to optimize the timeout value.

2.2 Expectation of J

Computing the expectation of a job execution time, general conclusion can be made on its behavior when the timeout value increases, independently from the system latency distribution. As shown in appendix A, the expectation of J is:

$$E_J(t_\infty) = \frac{1}{F_R(t_\infty)} \int_0^{t_\infty} u f_R(u) du + \frac{t_\infty}{(1-\rho)F_R(t_\infty)} - t_\infty. \tag{6}$$

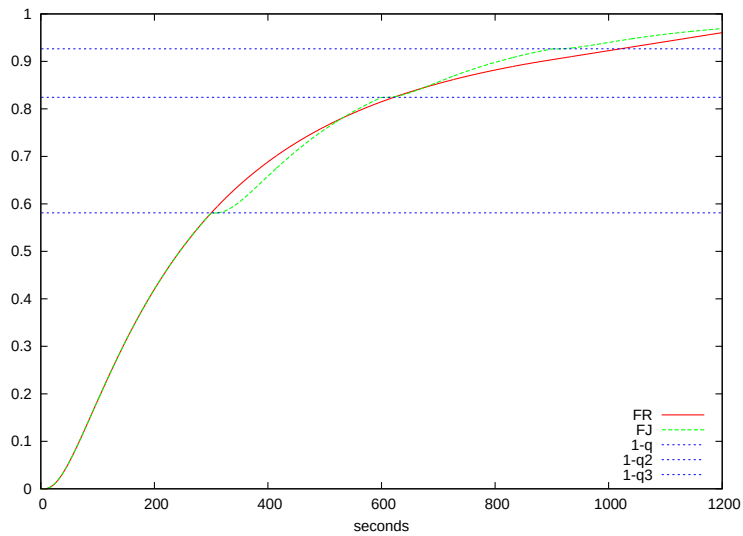

Figure 2. No general result on F_J.

Equation 6 compares to similar expressions derived for modeling completion times probabilistically : equation 6 in [5] and equation 1 in [4]. In both cases, the authors introduced a fixed cost penalty to resubmission that we consider to be zero (there is almost no overhead induced by job resubmission on a large scale system). In [5], the authors also derives higher moments of J and some relevant properties about them (*e.g.* their existence). Our hypotheses are similar to theirs except that they do not take into account outliers that are of major importance on the infrastructures we are targeting. As stated above, this parameter is characteristic of unreliable systems and they are needed to properly model a grid infrastructure. In [4], the authors do take into account the outlier ratio (denoted L). However, the studied hypotheses do not really match ours : the case of a client being unable to hold more than one connection (so-called simple client) is not developed, even if noticeable remarks (such as the fact that the timeout values of all the resubmissions have to be identical) are done.

As shown in in appendix B, E_J has the following limits:

$$\lim_{t_\infty \to \infty} E_J(t_\infty) = +\infty \quad \text{if } \rho \neq 0 \tag{7}$$

$$\text{and} \quad \lim_{t_\infty \to \infty} E_J(t_\infty) = E_R \quad \text{otherwise.} \tag{8}$$

If $\rho \neq 0$, the line $E_R + \frac{\rho}{1-\rho} t_\infty$ is an asymptote of $E_J(t_\infty)$.

The first limit can be explained by noticing that a single outlier may lead to an infinite execution time. When $t_\infty \to +\infty$, the probability for encountering an outlier tends towards 1 and the expected execution time tends towards infinity. The second limit is also intuitive: in absence of outliers, if no timeout value is set, then the system latency would not be disturbed and the expectation of a job duration would resume to the one of the system latency.

Equations 7 and 8 show that ρ as a major impact on the system behavior. The case $\rho = 0$ corresponds to a reliable cluster management system: faults causing jobs loss are very unlikely (highly reliable LAN, robust schedulers). The case $\rho > 0$ is needed to model grid infrastructures. Lower reliability of WANs, scale effects and less mature workload management middlewares lead to a significant number of outliers.

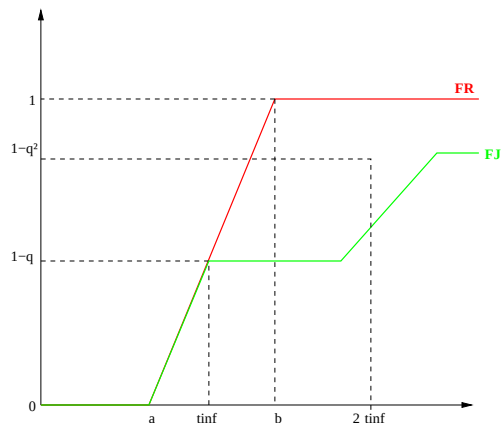

Figure 3. Behavior of F_J for a uniform distribution without outliers.

For instance on the EGEE infrastructure, ρ is in the order of 2%. In case of outliers it is mandatory to set a timeout value.

3 Results on classical distributions

In this section, we study some classical distributions from a theoretical point of view in order to understand how the timeout value impacts the expectation of the job duration both with and without outliers. We explore distributions with light tails (uniform, truncated Gaussian and Weibull with shape parameter >1), heavy tails (Weibull with shape parameter <1 and log-normal) and power tails (Pareto) to show how they exhibit different behavior. The exponential and Weibull distributions will constitute a transition between light and heavy-tailed distributions. The Weibull one is indeed light-tailed for a shape parameter greater than 1 and heavy-tailed otherwise. Light-tailed distributions are the ones that decay faster than the exponential. In this case, there exists a such that: $\lim_{t\to+\infty} e^{at}(1-F(t)) = 0$. On the contrary, heavy-tailed distributions decay more slowly than the exponential : $\lim_{t\to+\infty} e^{at}(1-F(t)) = +\infty$. Power-tailed distributions are a subset of the heavy-tailed ones. In this case, there exists a and b such that $\lim_{t\to+\infty} \frac{1-F(t)}{t^a} = b$.

For each distribution, our goal is to determine the optimal timeout value $\hat{t}_\infty = \arg\min_{t_\infty}\{E_J(t_\infty)\}$. In case of very reliable systems (when no outliers are present), the optimal value of the timeout may be $+\infty$, which means that no timeout should be set. Another singular optimal timeout value is 0. This configuration occurs when the probability for the job to face a null latency is so high that it is interesting to resubmit the job as soon as one knows that it is going to face a non null latency. This later result would only be realistic if it was possible to resubmit an arbitrarily large number of jobs at no additional cost. Obviously, the overhead induced on any real system would finally slow down the process.

3.1 Uniform distribution

In this case, the pdf of the system latency is:

$$f_R(t) = \begin{cases} \frac{1}{b-a} & \text{if } t \in [a,b] \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

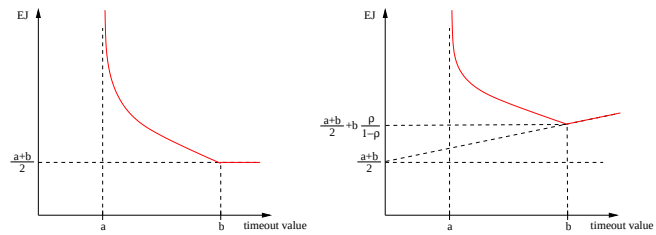

Figure 4. Behavior of $E_J(t_\infty)$ for the uniform distribution without (left) and with (right) outliers.

We can derive from equation 6 the expectation of J:

$$E_J(t_\infty) = \begin{cases} +\infty & \text{if } t_\infty \leq a \\ \frac{t_\infty + a}{2} + t_\infty \frac{b - t_\infty + \rho(t_\infty - a)}{(t_\infty - a)(1-\rho)} & \text{if } t_\infty \in [a, b] \\ \frac{b+a}{2} + t_\infty \frac{\rho}{1-\rho} & \text{otherwise.} \end{cases} \tag{10}$$

The shape of $E_J(t_\infty)$ is depicted on figure 4. The optimal timeout value is b both with and without outliers. Even without outliers, setting the timeout to $+\infty$ is optimal because the expectation of J is constant in [b,+∞[.

If there is no outliers, it can be graphically noticed that setting a timeout always penalizes the execution. Indeed, as figure 3 shows, we have $F_J(t) \leq F_R(t)$, for every timeout value t_∞ and every time point t.

3.2 Truncated Gaussian

Normal distributions are the most commonly used but they do not exclude negative values. In our case, the latency cannot be lower than 0. We are thus considering Gaussian distributions with mean μ and standard-deviation σ truncated above 0. The pdf and cdf of the system latency are:

$$f_R(t) = \begin{cases} \frac{1}{\Phi\left(\frac{\mu}{\sigma}\right)} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} & \text{if} \quad t \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

$$F_R(t) = \frac{\Phi\left(\frac{\mu}{\sigma}\right) - \Phi\left(\frac{\mu - t}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \quad \text{with} \quad \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}u^2} du.$$

As shown in appendix E, the expectation of the job duration is then:

$$\mu + \sigma \frac{\phi\left(\frac{\mu}{\sigma}\right) - \phi\left(\frac{\mu - t_\infty}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right) - \Phi\left(\frac{\mu - t_\infty}{\sigma}\right)} + \frac{1}{1-\rho} t_\infty \left(\frac{\Phi\left(\frac{\mu - t_\infty}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right) - \Phi\left(\frac{\mu - t_\infty}{\sigma}\right)} + \rho \right).$$

with $\phi = \Phi'$.

The curve of E_J is plotted on figure 5. E_J exhibits different behaviors depending on the presence of outliers or not. If there are no outliers ($\rho = 0$), E_J is decreasing towards its limit E_R when $t_\infty \to +\infty$. On the other hand, when $\rho \neq 0$, then E_J exhibits a global minimum reached for $\hat{t}_\infty < +\infty$. The corresponding

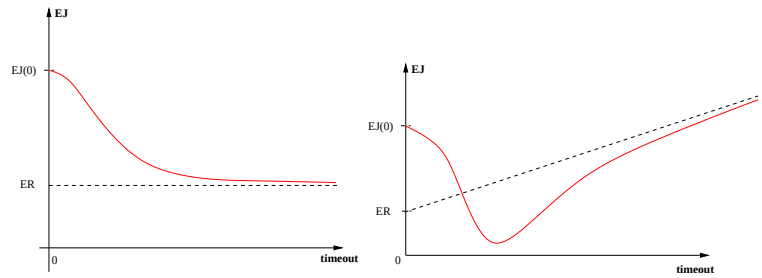

Figure 5. Behavior of the expectation of J for a truncated Gaussian distribution without (left) and with (right) outliers.

proof is based on the fact that the forth derivative of E_J is always positive, so that we can study the existence of a root in the lower order derivatives. It is reported in appendix F.

If the distribution of the system latency is Gaussian and $\rho = 0$, timeouting is not a solution to limit the impact of variability, regardless of the variability order of magnitude. In this case, other solutions such as multi-submissions or job grouping have to be studied.

3.3 Exponential distribution

In this case, the cdf of the system latency is:

$$F_R(t) = 1 - e^{-\alpha t}.$$

And according to equation 6, the expectation of J is:

$$E_J(t_\infty) = \frac{1}{\alpha} + \frac{\rho t_\infty}{(1-\rho)(1-e^{-\alpha t_\infty})}.$$

The curve of $E_J(t_\infty)$ is depicted on figure 6. In case of outliers, E_J is thus increasing and the best timeout value is $\hat{t}_\infty = 0$. If there are no outliers, the expectation of J is independent from t_∞, which is a singular behavior particular to the exponential distribution, as proved in appendix C. This characteristic of the exponential distribution has to be related to the fact that this distribution is the only one to be memory-less. In this case, at a given instant, knowing that a job is still in the system does not give any information about its future behavior and the timeout value thus does not impact the distribution of J. The exponential distribution is a particular case of the Weibull one which is studied in the next section.

3.4 Weibull distribution

The Weibull distribution is typically used to model the failure of technical devices. For this distribution, the cdf of the grid latency is:

$$F_R(t) = 1 - e^{-\left(\frac{t}{\lambda}\right)^k}$$

k is a shape parameter and λ is a scale parameter of the distribution. In the context of failure modeling, k<1 means that the failure rate decreases over time, $k = 1$ means that the failure rate is independent from

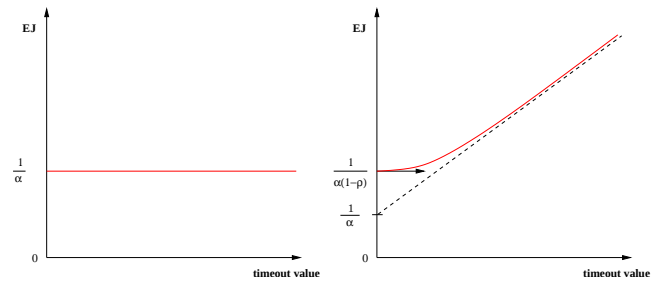

Figure 6. Behavior of E_J for an exponential distribution without (left) and with (right) outliers.

time and k>1 means that the failure rate increases over time. In our situation, the random variable R can be seen as the date at which a job completes, *i.e* R models the success of a job. Note that the exponential distribution of parameter $1/\lambda$ is a Weibull distribution with k=1.

In this case, we have the following results for timeout optimization, *in absence of outliers*:

- If $k > 1$, then setting a timeout value always penalizes the execution, whatever this value is. The optimal timeout value is thus $+\infty$ (no timeout).
- If $k < 1$, then the timeout value has to be as low as possible. The optimal timeout value is 0.
- If $k = 1$, then the distribution of R is exponential and the timeout value does not impact the job execution time.

The corresponding proofs are reported in appendix D. They are based on a comparison between F_J and F_R, following the remarks done in section 2.1.

The obtained results are coherent with the classical interpretation of the shape parameter of the Weibull distribution. Indeed, when k>1, the success rate of a job is increasing over time, which explains that timeouting will penalize the job. On the contrary, when k<1, the success rate of a job is then decreasing over time, and timeouting as soon as possible becomes mandatory.

When outliers are present, we conjecture that the behaviour remains the same when the shape parameter is lower than 1. Indeed, outliers lengthen the tail of the distribution and thus lower the timeout value. Similarly, when the shape parameter is greater that 1, it seems that the expectation of the job execution time with respect to the timeout value reaches a global finite minimum. Its derivative tends towards minus infinity (it seems to be equivalent to $-\frac{\lambda^k(k-1)}{(1-\rho)t_\infty^{k-1}}$) when the timeout value tends towards zero and E_J is thus decreasing on an intervall starting from $t_\infty = 0$. As it is tending towards $+\infty$ when t_∞ tends towards $+\infty$, it has to reach a global finite minimum. However, this behaviour is still not proved.

3.5 Log-normal distribution

The log-normal distribution is a typical example of heavy-tailed distribution. In this section, we assume that R has a log-normal distribution with parameters μ and σ. In this case, the cdf and pdf of the system latency are respectively:

$$F_R(t) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \quad \text{and} \quad f_R(t) = \frac{1}{t\sqrt{2\pi}\sigma} e^{-\frac{(\ln t-\mu)^2}{2\sigma^2}}.$$

The expectation and standard-deviation of R are:

$$E_R = e^{\mu+\frac{\sigma^2}{2}} \quad \text{and} \quad \sigma_R = \left(e^{\sigma^2} - 1\right) e^{2\mu+\sigma^2}. \tag{11}$$

In this case, as reported in appendix G, we can show that E_J is:

$$E_R.\left(\frac{\Phi\left(x_\infty - \sigma\right)}{\Phi(x_\infty)} + e^{\sigma x_\infty - \frac{\sigma^2}{2}}\left(\frac{1}{(1-\rho)\Phi(x_\infty)} - 1\right)\right) \tag{12}$$
$$\text{where} \quad x_\infty = \frac{\ln(t_\infty) - \mu}{\sigma}$$

This expression shows that the minimization of E_J can be performed independently from μ on the transformed variable x_∞. The obtained solution $\hat{x}_\infty(\sigma,\rho)$ only depends on σ and ρ. The optimal timeout value can then be written as:

$$\hat{t}_\infty(\mu,\sigma) = e^{\mu} K(\sigma,\rho) \quad \text{where} \quad K(\sigma,\rho) = e^{\sigma \hat{x}_\infty(\sigma,\rho)} \tag{13}$$

and:

$$\hat{x}_\infty(\sigma,\rho) = \arg\min_{x_\infty} \left(\frac{\Phi\left(x_\infty - \sigma\right)}{\Phi(x_\infty)} + e^{\sigma x_\infty - \frac{\sigma^2}{2}}\left(\frac{1}{(1-\rho)\Phi(x_\infty)} - 1\right)\right).$$

$K(\sigma,\rho)$ is actually the optimal timeout value for $\mu = 0$.

We also have the following limit for $t_\infty = 0$:

$$\lim_{t_\infty \to 0} E_J(t_\infty) \quad = \quad \lim_{x_\infty \to -\infty} E_J(x_\infty) = +\infty.$$

This infinite limit proves that when $\rho \neq 0$, there exists a finite non null optimal timeout value that minimizes E_J. Indeed, in this case, the limit of E_J when t_∞ tends towards infinity is infinite, according to equation 7 and E_J thus has to reach a global minimum.

The existence of a global minimum of $E_J(t_\infty)$ when $\rho \neq 0$ is not straight-forward. Given the infinite limit of E_J when t_∞ tends towards 0 and given that $E_J(+\infty) = E_R$, it resumes to the existence of a t_∞ for which $E_J(t_\infty) < E_R$. If $\sigma > 1$, then $t_\infty = e^{\mu}$ satisfies this relation. Indeed, in this case, $x_\infty = 0$ and according to equation 12, $E_J(x_\infty = 0) = E_R(2\Phi(-\sigma) + e^{-\frac{\sigma^2}{2}})$. A numeric resolution then shows that $E_J < E_R$ if and only if $\sigma \gtrsim 0.9311$. Numeric simulations also suggest that E_J has a global minimum even for lower values of σ. However, an analytic proof still has to be done.

Figure 7 displays a simulation of the optimal timeout value for μ=0, several values of the outlier ratio and σ ranging from 1 to 2. We first can notice that $K(\sigma,\rho)$ seems decreasing with respect to ρ. The timeout value thus has to be reduced when the proportion of outliers is increasing, which is coherent. Moreover, given an outlier ratio, the optimal timeout value for $\mu = 0$ is decreasing as σ is growing. It is also coherent because the standard-deviation of the log-normal distribution is increasing with respect to σ (see equation 11). The optimal timeout value thus has to be reduced as the variability of the infrastructure is growing.

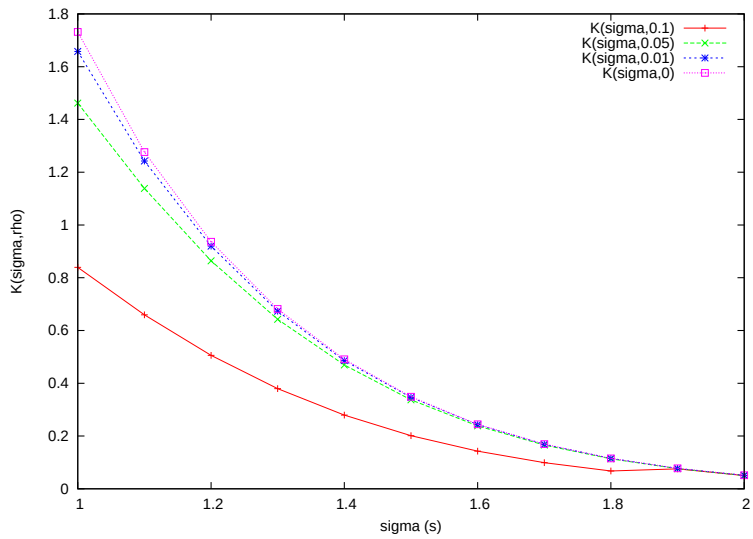

Figure 7. Evolution of the optimal timeout value for μ=0 in the log-normal case.

3.6 Pareto distribution

The Pareto distribution was introduced to represent the distribution of wealth and proved to be very accurate to model a large class of computer systems measurements (jobs durations, size of the files, data transfers length on the Internet...) [13, 14]. It is an example of power tailed distribution. The cdf of the system latency is then:

$$F_R(t) \quad = \quad 1 - \left(\frac{a}{a+t}\right)^{\nu} \quad \text{with} \quad a \quad \text{and} \quad \nu \quad > \quad 0.$$

The expectation is only defined for ν>1. Then:

$$E_R \quad = \quad \frac{a}{\nu - 1}.$$

In this case, the expression of E_J can be directly derived from equation 6 and it is:

$$\frac{a + t_\infty \nu - a\left(\frac{a+t_\infty}{a}\right)^{\nu}}{(1-\nu)\left[\left(\frac{a+t_\infty}{a}\right)^{\nu} - 1\right]} + \frac{t_\infty}{(1-\rho)\left[\left(\frac{a+t_\infty}{a}\right)^{\nu} - 1\right]} + \frac{\rho}{1-\rho} t_\infty.$$

We also have the following limit when the timeout value is null:

$$\lim_{t_\infty \to 0} E_J(t_\infty) \quad = \quad \frac{a}{\nu(1-\rho)}.$$

We then can show (see proof in appendix H) that the expectation of the job duration time is increasing with respect to the timeout value, regardless of the ρ value. The optimal timeout value is thus 0. The behavior of $E_J(t_\infty)$ is depicted on figure 8.

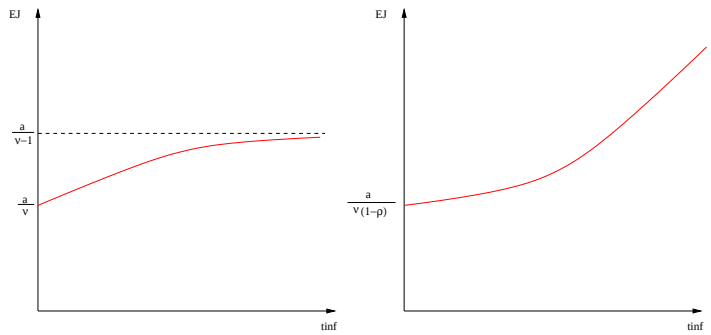

Figure 8. Behavior of E_J for a Pareto distribution of R. Left: no outliers ; Right: $\rho \neq 0$.

Distribution	Without outliers ($\rho = 0$)	With outliers (ρ>0)
Uniform	$+\infty$ (or b)	b
Truncated Gaussian	$+\infty$	$0 < \hat{t}_\infty < +\infty$
Weibull k>1	$+\infty$	$0 < \hat{t}_\infty < +\infty$
Exponential	any	0
Weibull k<1	0	0
Log-normal (μ,σ)	$\hat{t}_\infty = e^\mu K(\sigma) < +\infty$	$0 < \hat{t}_\infty < +\infty$
Pareto (ν>1)	0	0

Table 1. Optimal timeout values for the studied distributions of R.

3.7 Results summary and interpretation

Table 1 displays a summary of the results we obtained for various distributions of the system latency. Those results suggest that the weight of the tail of the distribution of the system latency is a discriminatory parameter for the timeout optimization when outliers are not present. Indeed, only heavy-tailed distributions such as the log-normal, or the Pareto ones lead to finite optimal timeout values. In this case, the optimization speeds the execution up. On the other hand, when the distribution of the system latency decays more rapidly than the exponential, then setting a timeout value always penalizes the execution and the optimal timeout is $+\infty$. The exponential distribution stands in the middle and is not affected by the timeout value.

As noticed in section 2, taking into account the outliers resumes to multiplying F_R by the factor (1-ρ). In this case, the distribution of the system latency becomes heavy-tailed as $\lim_{x\to+\infty} e^{ax}(1-(1-\rho)F_R(x)) = +\infty$ when $a > 0$. Effectively, in this case, the optimal timeout value that we found is always finite, which is coherent with this interpretation.

Results for the Weibull distribution with outliers and for the log-normal one without outliers and with σ<0.94 are only conjectures.

3.8 Performance improvement

In case of reliable systems (without outliers), the expectation of the job duration without timeout equals to the one of the system latency. In this case, the ratio $\frac{E_R}{E_J(\hat{t}_\infty)}$ evaluates the speed-up yielded by the

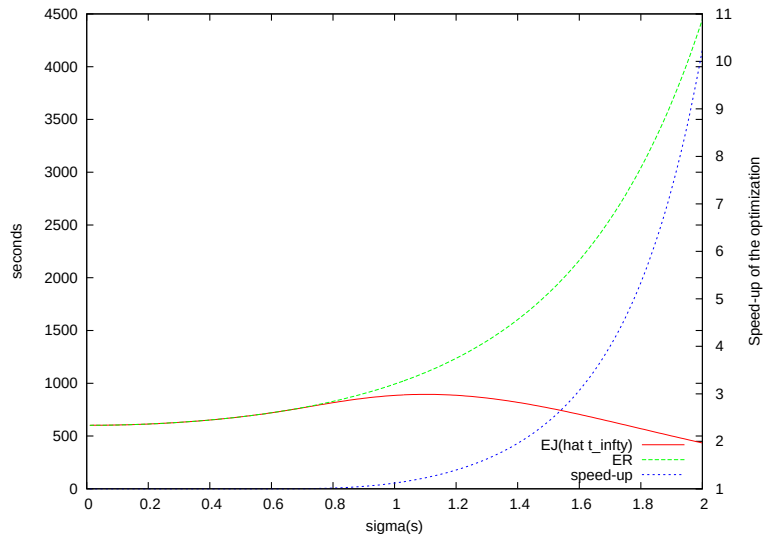

Figure 9. Evolution of the speed-up of the optimization for $\mu = 6.4$s in the log-normal case.

optimization. If the latency of the system is light-tailed, then setting a timeout value always penalizes the execution. The best strategy is thus to set the timeout value to infinity. In this case, the optimization does not provide any speed-up with respect to the expectation of the system latency. Concerning the limit case of an exponential distribution, the expectation of the job duration is independent from the timeout value and the optimization does not lead to any speed-up.

The optimization becomes interesting for heavy-tailed distributions as already suggested. For the log-normal case, figure 9 displays a numerical simulation of the evolution of the speed-up of the optimization with respect to σ for a particular value of μ. We can see that the speed-up is growing with σ. The higher the system latency, the more interesting the timeout optimization. Concerning the Pareto distribution, the optimized expectation of the job duration without outliers is $\frac{a}{\nu}$, whereas the one obtained without setting any timeout is $E_R = \frac{a}{\nu-1}$. The speed-up obtained by the optimization is thus $\frac{\nu}{\nu-1}$. This value is maximal for $\nu = 1$ and decreases towards 1 when ν increases. Under Pareto assumption, the variance of the system latency $\left(\frac{\nu a^2}{(\nu-2)(\nu-1)^2}\right)$ is decreasing with respect to ν. Here again, the more variable the latency of the infrastructure, the more important the speed-up yielded by the optimization.

When outliers are present, the optimization of the timeout prevents the expectation of J to be infinite. The impact of the optimization can then be evaluated by comparing the optimized expectation of the job duration to the one obtained without outliers. In case of a uniform distribution, outliers add the term $b\frac{\rho}{1-\rho}$ to the expectation of the job duration. This term is increasing with respect to the outlier ratio and tends towards infinity when ρ tends towards 1. The exponential distribution and the Pareto one exhibit a similar behavior: the outliers introduce an extra $\frac{1}{1-\rho}$ factor on the expectation of the job duration. Here again, this factor tends towards infinity when the outlier ratio tends towards 1.

4 Experiments

In this section, we present experimental results obtained by measuring the distribution of the latency of the EGEE grid infrastructure on a particular time period. The EGEE grid is a pool of thousands

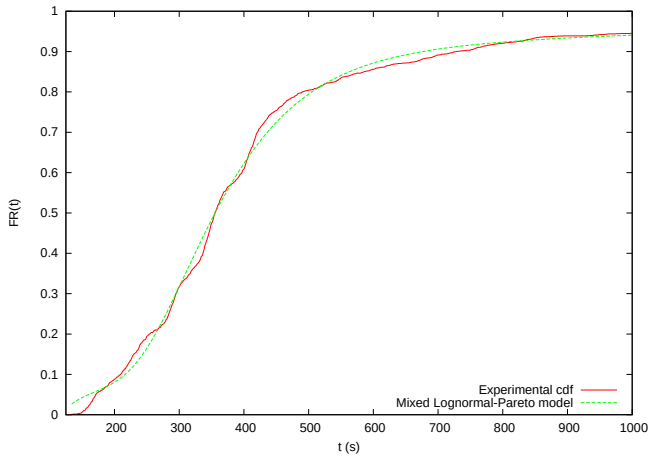

Figure 10. Measured data (green) and best fitting Log-normal-Pareto model (red).

computers (standard PCs) and storage resources accessible through the gLite middleware. The resources are operated in computing centers, each of them running its internal batch scheduler. Jobs are submitted from a user interface to a central Resource Broker (RB) which dispatches them to the computing centers. EGEE is a production infrastructure with more than 25000 CPUs spread in more than 190 computing centers. It is characterized by its high throughput but also by its high latency, high variability and outliers. It is thus an ideal target to test our optimization procedure.

4.1 Measure of the distribution

To measure the distribution of the system latency on the EGEE grid, we submitted probe jobs that only consist in the execution of a `/bin/hostname` and we measure their round-trip time. We maintain a constant number of probes inside the system by submitting a new one as soon as one completed to avoid introducing any extra variability.

To measure the proportion of outliers, we consider a long duration T beyond which jobs are considered as outliers. When the value of T is sufficiently high, the measured outlier ratio, computed as the number of outliers jobs divided by the number of submitted ones is assumed not to depend on T and we use it as an approximation of the true one. However, by definition, outliers are rare phenomena, and the measure of the outlier ratio does not rely on a significant number of events. It is thus likely to be misestimated. Specific statistical methods concerning rare events [15] may be used to have a more precise estimation of the outlier ratio.

Our measure of the distribution of R gathers 2137 probe jobs spread on 3 different days and involving 3 RB. The maximal duration of those jobs was fixed to $t_{\max} = 10000$ seconds. Beyond this value, we consider the job as an outlier. Given those conditions, we obtained an outlier ratio of 2.5 %. In normal operating mode, the measured distribution of R is plotted in green on figure 10. It has an expectation of 393 seconds and a standard deviation of 792 seconds.

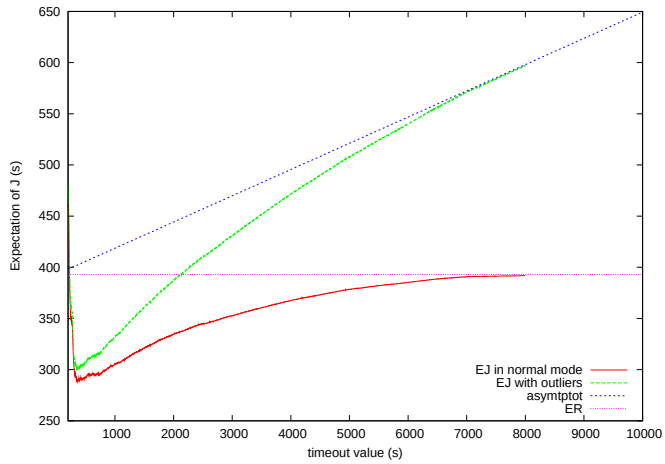

Figure 11. Evolution of the expectation of J in the experimental case.

4.2 Timeout optimization

If we do not take outliers into account, the evolution of E_J with respect to the timeout value is plotted in red on figure 11. E_J then converges towards E_R as predicted by the theoretical analysis. In this case, E_J reaches a minimum for $\hat{t}_\infty = 360$s. At this optimal point, $\hat{E}_J(\hat{t}_\infty) = 289$s. The speed-up w.r.t to an execution without timeout is 1.36.

The evolution of $E_J(t_\infty)$ taking the outliers into account is plotted in green on figure 11. E_J effectively tends to its asymptote. The optimal timeout value $\hat{t}_\infty$ is now 358 seconds and $\hat{E}_J(\hat{t}_\infty)$ has grown to 300 seconds. Setting the optimal timeout value thus limits the impact of the outliers to a 11-seconds loss, whereas it would be highly superior if the timeout value is not properly set, as suggested by figure 11.

4.3 Model of the measured distribution

To relate those experimental results to the ones presented in the previous section, we here model the experimental distribution of the grid latency. The experimental data shown on figure 10 cannot be reasonably fitted with any of the standard distributions described in section 3. However, the distribution appears to be close to a log-normal distribution for low values (up to 500 seconds) and a Pareto distribution beyond. Based on this observation we fitted the experimental data with the following distribution which is an interpolation of the log-normal and Pareto ones, for t in $[t_{\min}, t_{\max}]$:

$$F_R^m(t) = (1-\alpha(t))\,\Phi\left(\frac{\ln(t-t_{\min})-\mu}{\sigma}\right) + \alpha(t)\left(1-\left(\frac{a}{a+t}\right)^{\nu}\right) \quad \text{with} \quad \alpha(t) = \left(\frac{t-t_{\min}}{t_{\max}-t_{\min}}\right)^{k}. \tag{14}$$

$t_{\min}$ denotes the smallest latency measured among the data (the cdf is zero below this value) and $t_{\max}$ the highest one. There are thus five parameters fully describing this model (μ, σ, a, ν, and k). $\alpha(t)$ is a weight function designed so that $\alpha(t_{\min}) = 0$ and $\alpha(t_{\max}) = 1$. The model thus tends to a log-normal distribution in $t_{\min}$ and to a Pareto one in $t_{\max}$.

We have estimated the best fit of the model 14 with the experimental data by minimizing the least square criterion:

$$\arg \min_{(\mu,\sigma,a,\nu,k)} \left\{ \sum_{i=t_{\min}}^{t_{\max}} \left(F_R^{model}(i) - F_R^{exp}(i) \right)^2 \right\}$$

where $F_R^{exp}(i)$ is the value of the measured distribution at time i. The optimal model is displayed on figure 10 (red curve). A Kolmogorov-Smirnov test was made to evaluate the quality of the model. When considering an undersampling of up to 1000 measurements, the Kolmogorov-Smirnov test value is $D_{1000} = 1.35$ (we used $D_n = \sqrt{n} \sup |F_R^{exp} - F_R^m|$), which correspond to a p-value $p = 0.051$. The tests is thus positive. It shows that a simple model (5 parameters) can accurately model the distribution measured over a very complex workload system (EGEE grid infrastructure) even when considering a very large data sample.

5 Conclusion

In this work, we presented a model of the job execution time taking into account the timeout value and resubmissions. We used it to optimize the timeout value of the jobs and we studied the optimization results for various latency distributions. Two cases have been distinguished, given that we consider that the job management system is prone to face outliers (grid) or not (cluster).

We showed that the optimal timeout value is highly dependent on the nature of the distribution of the system latency. In absence of outliers, the heavy-tailed distribution studied (log-normal) leads to a finite optimal timeout value whereas for the light-tailed ones (truncated Gaussian, uniform), setting a timeout value always penalizes the execution.

If we take outliers into account, our model predicts that the expectation of the job execution time is then diverging to $+\infty$ when the timeout value tends towards $+\infty$ for every distribution. The expectation of J has an asymptote whose slope only depends on the outlier ratio. The optimal timeout value is then finite for all the studied distributions, which can be explained by noticing that taking outliers into account lengthens the tail of the distribution.

We finally presented some experimental results from a distribution that was measured on the EGEE production grid. Results showed that even without outliers, a 1.36 speed-up can be achieved by using the optimal timeout value. Considering outliers, using the optimal timeout value is even more critical and the resulting optimal expectation of the job duration is close to the one obtained in absence of outliers.

However, some problems still remain to be able to use this optimization in practice. The most important one is to be able to dynamically measure the distribution of the non-stationary grid latency.

Acknowledgments

We thank Pierre Bernhard and Philippe Nain for useful advises. This work is partially funded by the French research program “ACI-Masse de données” (http://acimd.labri.fr/), AGIR project (http://www.aci-agir.org/). We are grateful to the EGEE European project for providing the grid infrastructure and user assistance.

References

[1] H. Casanova, "On the Harmfulness of Redundant Batch Requests," in *15th IEEE International Symposium on High Performance Distributed Computing (HPDC'06)*, (Paris, France), pp. 255–266, June 2006.

[2] T. Glatard, D. Emsellem, and J. Montagnat, "Generic web service wrapper for efficient embedding of legacy codes in service-based workflows," in *Grid-Enabling Legacy Applications and Supporting End Users Workshop (GELA'06)*, (Paris, France), June 2006.

[3] T. Glatard, J. Montagnat, and X. Pennec, "Probabilistic and dynamic optimization of job partitioning on a grid infrastructure," in *Parallel, Distributed and network-based Processing (PDP06)*, (Montbéliard-Sochaux, France), pp. 231–238, Feb. 2006.

[4] L. Libman and A. Orda, "Optimal Retrial and Timeout Strategies for Accessing Network Resources," *IEEE/ACM Transactions on Networking (TN)*, vol. 10, pp. 551–564, Aug. 2002.

[5] A. van Moorsel and K. Wolter, "Analysis of Restart Mechanisms in Software Systems," *IEEE Trans. on Software Eng. (TSE)*, vol. 32, pp. 547–558, Aug. 2006.

[6] A. Kesselman and Y. Mansour, "Optimizing TCP Retransmission Timeout," in *International Conference on Networking*, vol. 3421 of *LNCS*, (Saint-Denis de la Réunion), Springer, Apr. 2005.

[7] P. Reinecke, A. van Moorsel, and K. Wolter, "A Measurement Study of the Interplay between Application Level Restart and Transport Protocol," in *Intl Service Availability Symposium (ISAS)*, vol. 3335 of *LNCS*, (Munich, Germany), pp. 86–100, May 2004.

[8] W. Xie, H. Sun, Y. Cao, and K. Trivedi, "Optimal Webserver Session Timeout Settings for Web Users," in *Computer Measurement Group Conference (CMGC)*, (Reno, NV, USA), pp. 799–820, Dec. 2002.

[9] P. Rong and M. Pedram, "Determining the optimal timeout values for a power-managed system based on the theory of Markovian processes: offline and online algorithms," in *Design, Automation and Test in Europe*, (Munich, Germany), pp. 1128–1133, Mar. 2006.

[10] T. Glatard, J. Montagnat, and X. Pennec, "Optimizing jobs timeouts on clusters and production grids," Tech. Rep. I3S/RR-2006-35-FR, http://www.i3s.unice.fr/~glatard/RR06.35.pdf, I3S, 2006.

[11] H. Gautama and A. J. C. van Gemund, "Symbolic Performance Estimation Of Speculative Parallel Programs," *Parallel Processing Letters*, vol. 13, no. 4, pp. 513–524, 2003.

[12] J. Schopf and F. Berman, "Stochastic Scheduling," in *Supercomputing (SC'99)*, (Portland, USA), 1999.

[13] M. Harchol-Balter, "Task Assignment with Unknown Duration," *Journal of the ACM (JACM)*, vol. 49, pp. 260–288, Mar. 2002.

[14] N. Ben Azzouna, *Etude des méthodes d'échantillonnage des flux pour la mesure dans les réseaux large bande*. PhD thesis, Université Pierre et Marie Curie - Paris VI, Paris, Sept. 2004.

[15] B. Beauzamy, *Méthodes probabilistes pour l'étude des phénomènes réels*. Société de Calcul Mathématique S. A., Mar. 2004.

[16] T. Glatard, J. Montagnat, and X. Pennec, "Efficient services composition for grid-enabled data-intensive applications," in *IEEE International Symposium on High Performance Distributed Computing (HPDC'06)*, (Paris, France), June 2006.

A Expectation of J in the general case

We have:

$$
\begin{aligned}
E_J(t_\infty) &= \int_0^\infty t f_J(t)dt = \sum_{n=0}^{\infty} \int_0^\infty t f_J^{[n,n+1]}(t)dt \\
&= (1-\rho) \sum_{n=0}^{\infty} q^n \int_{nt_\infty}^{(n+1)t_\infty} t f_R(t - nt_\infty)dt \\
&= (1-\rho) \sum_{n=0}^{\infty} q^n \int_0^{t_\infty} (u + nt_\infty) f_R(u)du \\
&= \frac{1-\rho}{1-q} \int_0^{t_\infty} u f_R(u)du + \frac{(1-\rho)qt_\infty}{(1-q)^2} \int_0^{t_\infty} f_R(u)du \\
&= \frac{1}{F_R(t_\infty)} \int_0^{t_\infty} u f_R(u)du + \frac{(1-\rho)\,(1-(1-\rho)F_R(t_\infty))\,t_\infty}{(1-\rho)^2 F_R(t_\infty)^2} F_R(t_\infty) \\
&= \frac{1}{F_R(t_\infty)} \int_0^{t_\infty} u f_R(u)du + \frac{(1-(1-\rho)F_R(t_\infty))t_\infty}{(1-\rho)F_R(t_\infty)} \\
&= \frac{1}{F_R(t_\infty)} \int_0^{t_\infty} u f_R(u)du + \frac{t_\infty}{(1-\rho)F_R(t_\infty)} - t_\infty
\end{aligned}
$$

B Limits of E_J

$$
\lim_{t_\infty \to +\infty} \frac{1}{F_R(t_\infty)} \int_0^{t_\infty} u f_R(u)du = \int_0^{+\infty} u f_R(u)du = E_R
$$

And, when $\rho \neq 0$:

$$
\lim_{t_\infty \to +\infty} \frac{t_\infty}{(1-\rho)F_R(t_\infty)} - t_\infty = +\infty
$$

Whereas when $\rho = 0$:

$$
\frac{t_\infty}{(1-\rho)F_R(t_\infty)} - t_\infty = \frac{t_\infty(1-F_R(t_\infty)}{F_R(t_\infty)} = \frac{t_\infty \int_{t_\infty}^{\infty} f_R(u)du}{F_R(t_\infty)}
$$

$$
\text{Thus,} \quad \frac{t_\infty}{(1-\rho)F_R(t_\infty)} - t_\infty \leq \frac{\int_{t_\infty}^{\infty} u f_R(u)du}{F_R(t_\infty)}
$$

$$
\text{So that} \lim_{t_\infty \to +\infty} \left(\frac{t_\infty}{(1-\rho)F_R(t_\infty)} - t_\infty \right) = 0
$$

C Distributions for which the timeout value does not impact E_J when $\rho = 0$

Let F be the cdf of a distribution which has this property. Then:

$$\begin{aligned}
&\forall n \in \mathbb{N}, \forall t_\infty \in \mathbb{R}^+,\\
&1-(1-F(t_\infty))^n + (1-F(t_\infty))^n F(t-nt_\infty) = F(t)\\
&\Rightarrow \forall n \in \mathbb{N}, \forall t_\infty \in \mathbb{R}^+,\\
&G(t_\infty)^n G(t-nt_\infty) = G(t) \quad \text{with} \quad G = 1-F\\
&\Rightarrow \forall n \in \mathbb{N}, \forall t_\infty \in \mathbb{R}^+,\\
&nH(t_\infty) + H(t-nt_\infty) = H(t) \quad \text{with} \quad H = ln\,(G)\\
&\Rightarrow \forall n \in \mathbb{N}, \forall t_\infty \in \mathbb{R}^+,\\
&H'(t-nt_\infty) = H'(t)
\end{aligned}$$

H' is thus periodic, with period nt_∞ for every t_∞ and every n. It is thus constant and we have $H'(t) = \alpha$. Thus, $H(t) = \alpha t + \beta$. We thus have $F(t) = 1 - e^\beta e^{\alpha t}$. Moreover, the limit of $F(t)$ has to be 1 when $t \to +\infty$, so that $\alpha < 0$ and $\beta = 0$, which demonstrates that the distribution of R has to be exponential.

D Behavior of E_J in the Weibull case without outliers

We have:

$$\begin{aligned}
& \forall n \in \mathbb{N}, \forall t \in [nt_\infty, (n+1)t_\infty],\\
F_J(t) &= 1 - q^n + q^n F_R(t-nt_\infty)\\
&= 1 - e^{-n\left(\frac{t_\infty}{\lambda}\right)^k} + e^{-n\left(\frac{t_\infty}{\lambda}\right)^k}\left(1 - e^{-\left(\frac{t-nt_\infty}{\lambda}\right)^k}\right)\\
&= 1 - e^{-\left(n\left(\frac{t_\infty}{\lambda}\right)^k + \left(\frac{t-nt_\infty}{\lambda}\right)^k\right)}
\end{aligned}$$

We are going to compare F_J and F_R. To do that, we will actually compare $\ln\left(\frac{1}{1-F_J}\right)$ and $\ln\left(\frac{1}{1-F_R}\right)$. The comparison resumes to study the sign of the following function $f_{t_\infty,n}$, for every n and every t_∞ ($f_{t_\infty,n} = \ln\left(\frac{1}{1-F_J}\right) - \ln\left(\frac{1}{1-F_R}\right)$):

$$\begin{aligned}
&\forall n \in \mathbb{N}, \forall t_\infty > 0,\\
f_{t_\infty,n} : [nt_\infty, (n+1)t_\infty] &\to \mathbb{R}\\
t &\mapsto n\left(\frac{t_\infty}{\lambda}\right)^k + \left(\frac{t-nt_\infty}{\lambda}\right)^k - \left(\frac{t}{\lambda}\right)^k
\end{aligned}$$

If $f_{t_\infty,n}$ is positive, then setting a timeout improves the execution. The derivative of $f_{t_\infty,n}$ with respect to t is: $f'_{t_\infty,n}(t) = \frac{k}{\lambda}\left(\left(\frac{t-nt_\infty}{\lambda}\right)^{k-1} - \left(\frac{t}{\lambda}\right)^{k-1}\right)$.

If $k > 1$: $f_{t_\infty,n}(nt_\infty) = \left(\frac{t_\infty}{\lambda}\right)^k \left(n - n^k\right) < 0$ and $\forall t \in [nt_\infty, (n+1)t_\infty]$, $f'_{t_\infty,n}(t) < 0$. $f_{t_\infty,n}(t)$ is thus negative on this interval and we have:

$$\forall t > 0, \forall t_\infty > 0,$$

$$\begin{aligned}\ln\left(\frac{1}{1-F_J(t)}\right)-\ln\left(\frac{1}{1-F_R(t)}\right) &< 0\\ \text{thus,}\quad F_J(t) &< F_R(t)\end{aligned}$$

It proves that every timeout value penalizes the execution. The timeout thus has to be infinite.

If $k<1$: $f_{t_\infty,n}(nt_\infty)=\left(\frac{t_\infty}{\lambda}\right)^k\left(n-n^k\right)>0$ and $\forall t\in[nt_\infty,(n+1)t_\infty]$, $f'_{t_\infty,n}(t)>0$. $f_{t_\infty,n}(t)$ is thus positive on this interval and we have:

$$\begin{aligned}&\forall t>0,\forall t_\infty>0,\\ \ln\left(\frac{1}{1-F_J(t)}\right)-\ln\left(\frac{1}{1-F_R(t)}\right) &> 0\\ \text{thus,}\quad F_J(t) &> F_R(t)\end{aligned}$$

Moreover, $n\left(\frac{t_\infty}{\lambda}\right)^k+\left(\frac{t-nt_\infty}{\lambda}\right)^k$ is decreasing with respect to t_∞. Thus, $\forall t_\infty>0,\forall t'_\infty>t_\infty$, $F_{J,t'_\infty}(t)<F_{J,t_\infty}(t)$ for every t. The optimal timeout value is thus 0.

E Expression of $E_J(t_\infty)$ in the truncated Gaussian case

According to equation 6,

$$\begin{aligned}E_J(t_\infty)&=\frac{1}{F_R(t_\infty)}\int_0^{t_\infty}u\,f_R(u)du+\frac{t_\infty}{(1-\rho)F_R(t_\infty)}-t_\infty\\ &=\frac{1}{\Phi\left(\frac{\mu}{\sigma}\right)-\Phi\left(\frac{\mu-t_\infty}{\sigma}\right)}\frac{1}{\sqrt{2\pi}\sigma}\int_0^{t_\infty}ue^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2}du+\frac{t_\infty\Phi\left(\frac{\mu}{\sigma}\right)}{(1-\rho)\left(\Phi\left(\frac{\mu}{\sigma}\right)-\Phi\left(\frac{\mu-t_\infty}{\sigma}\right)\right)}-t_\infty\end{aligned}$$

Moreover,

$$\begin{aligned}\int_0^{t_\infty}ue^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2}du&=\sigma^2\int_{-\frac{\mu}{\sigma}}^{\frac{t_\infty-\mu}{\sigma}}ve^{\left(-\frac{1}{2}v^2\right)}dv+\sigma\mu\int_{-\frac{\mu}{\sigma}}^{\frac{t_\infty-\mu}{\sigma}}e^{\left(-\frac{1}{2}v^2\right)}dv\\ &=\sigma^2\sqrt{2\pi}\left(\phi\left(\frac{\mu}{\sigma}\right)-\phi\left(\frac{\mu-t_\infty}{\sigma}\right)\right)+\sigma\mu\sqrt{2\pi}\left(\Phi\left(\frac{\mu}{\sigma}\right)-\Phi\left(\frac{\mu-t_\infty}{\sigma}\right)\right)\end{aligned}$$

Thus,

$$E_J(t_\infty)=\mu+\sigma\frac{\phi\left(\frac{\mu}{\sigma}\right)-\phi\left(\frac{\mu-t_\infty}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)-\Phi\left(\frac{\mu-t_\infty}{\sigma}\right)}+\frac{1}{1-\rho}t_\infty\left(\frac{\Phi\left(\frac{\mu-t_\infty}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)-\Phi\left(\frac{\mu-t_\infty}{\sigma}\right)}+\rho\right)$$

F Behavior of $E_J(t_\infty)$ in the truncated Gaussian case

Let us consider the following transformed variables:

$$v_\infty=\frac{\mu-t_\infty}{\sigma}\quad\text{and}\quad\lambda=\frac{\mu}{\sigma}$$

According to equation 11, we then have:

$$\frac{E_J(v_\infty)}{\sigma}=\lambda+\frac{\phi(\lambda)-\phi(v_\infty)}{\Phi(\lambda)-\Phi(v_\infty)}+(\lambda-v_\infty)\left(\frac{\Phi(v_\infty)}{(1-\rho)\left(\Phi(\lambda)-\Phi(v_\infty)\right)}+\frac{\rho}{1-\rho}\right)$$

To study the behavior of E_J, we are going to consider the function $f^{(1)}(v_\infty) = \frac{(1-\rho)(\Phi(\lambda)-\Phi(v_\infty))^2}{\sigma}\frac{\partial E_J(v_\infty)}{\partial v_\infty}$ which has the same sign as $\frac{\partial E_J(v_\infty)}{\partial v_\infty}$. We are going to show that the third derivative of $f^{(1)}$ is positive. We will then be able to study the sign of $f^{(1)}$. We have:

$$\begin{aligned} f'(v_\infty) &= \phi(v_\infty)\left[k(\lambda) - k(v_\infty) + \rho v_\infty\left(\Phi(v_\infty) - \Phi(\lambda)\right) + \rho\left(\phi(v_\infty) - \phi(\lambda)\right)\right] \\ &+ \Phi(v_\infty)\left(\Phi(v_\infty) - \Phi(\lambda)\right) - \rho\left(\Phi(\lambda) - \Phi(v_\infty)\right)^2 \\ \text{with} \quad & k(v) = v\Phi(v) + \phi(v) \end{aligned}$$

k is actually a primitive of Φ and is thus increasing. Derivating f' with respect to v, we obtain:

$$\begin{aligned} f^{(2)}(v) &= \phi(v)g^{(2)}(v) \\ \text{with } g^{(2)}(v) &= v\left(k(v) - k(\lambda)\right) + \rho v^2\left(\Phi(\lambda) - \Phi(v)\right) \\ & \quad +\rho v\left(\phi(\lambda) - \phi(v)\right) + (1-\rho)\left(\Phi(v) - \Phi(\lambda)\right) \end{aligned}$$

$f^{(2)}$ has the same sign of $g^{(2)}$. By successive derivation of this function, we obtain:

$$\begin{aligned} g^{(3)}(v) &= v\Phi(v)(1-2\rho) + \phi(v)(1-2\rho) + k(v) - k(\lambda) + 2\rho v\Phi(\lambda) + \rho\phi(\lambda) \\ & \quad \text{and} \\ g^{(4)}(v) &= 2(1-\rho)\Phi(v) + 2\rho\Phi(\lambda) \end{aligned}$$

$g^{(4)}(v)$ is thus positive on $]-\infty, \lambda]$, so that $g^{(3)}$ is increasing on this interval. Moreover, $g^{(3)}(-\infty) = -\infty$ and $g^{(3)}(\lambda) = \lambda\Phi(\lambda) + (1-\rho)\phi(\lambda)>0$. $g^{(3)}$ thus has a single root v_0 on $]-\infty, \lambda]$. $g^{(3)}$ is negative for $v<v_0$ and positive otherwise.

$g^{(2)}$ is thus decreasing on $]-\infty, v_0]$ and increasing on $[v_0, \lambda]$. Moreover, $g^{(2)}(-\infty) = +\infty$ and $g^{(2)}(\lambda) = 0$. $g^{(2)}$ thus has a single root v_1 on $]-\infty, \lambda]$ ($v_1<v_0$). $g^{(2)}$ is positive for $v<v_1$ and negative otherwise.

$f^{(1)}$ is thus increasing on $]-\infty, v_1]$ and decreasing on $[v_1, \lambda]$ Moreover, $f^{(1)}(\lambda) = 0$ and $f^{(1)}(-\infty) = -\rho\Phi(\lambda)$. Two cases then have to be studied:

1. $\rho = 0$: in this case, $f^{(1)}(-\infty) = 0$. $f^{(1)}$ is thus positive on $]-\infty, \lambda]$.

2. $\rho \neq 0$: in this case, $f^{(1)}(-\infty)<0$. $f^{(1)}$ thus has a single root v_2 on $]-\infty, \lambda]$ ($v_2<v_1$). $f^{(1)}$ is negative for $v<v_2$ and positive otherwise.

The behavior of $f^{(1)}$, $g^{(2)}$ and $g^{(3)}$ is plotted on figure F.

We can then conclude on the behavior of $E_J(t_\infty)$ by noticing that:

$$\begin{aligned} \frac{\partial E_J(t_\infty)}{t_\infty} &= \frac{\partial E_J(v_\infty)}{v_\infty}\frac{\partial v_\infty}{t_\infty} \\ &= -\frac{1}{\sigma}\frac{\partial E_J(v_\infty)}{v_\infty} \end{aligned}$$

The two cases described above thus resume to:

1. $\rho = 0$: $E_J(t_\infty)$ is decreasing on $[0, +\infty[$

2. $\rho \neq 0$: $E_J(t_\infty)$ is decreasing on $[0, t_2]$ and increasing on $[t2, +\infty[$ (with $t_2 = \mu - \sigma v_2$).

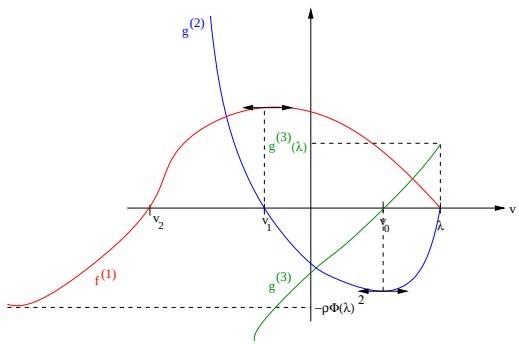

Figure 12. Behavior of $f^{(1)}$, $g^{(2)}$ and $g^{(3)}$ derivatives of $E_J(v_\infty)$ in the truncated Gaussian case.

G Expression of E_J in the log-normal case

If we consider the transformed variable $x_\infty = \frac{\ln(t_\infty)-\mu}{\sigma}$, then equation 6 gives:

$$E_J(x_\infty) = \frac{1}{\Phi(x_\infty)} \int_0^{e^{\sigma x_\infty+\mu}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\ln(u)-\mu}{\sigma}\right)^2} du + e^{\sigma x_\infty+\mu}\left(\frac{1}{(1-\rho)\Phi(x_\infty)} - 1\right)$$

If we then perform the variable change $v = \frac{\ln(u)-\mu}{\sigma} - \sigma$ in the integral term, we obtain:

$$E_J(x_\infty) = \frac{1}{\Phi(x_\infty)} e^{\mu+\frac{\sigma^2}{2}} \Phi(x_\infty - \sigma) + e^{\sigma x_\infty+\mu}\left(\frac{1}{(1-\rho)\Phi(x_\infty)} - 1\right)$$

H Behavior of E_J in the Pareto case

Let us consider the transformed variable $Z = \frac{a+t_\infty}{a}$. We then have, according to equation 14 and after some manipulations:

$$\frac{E_J(Z)}{a} = \frac{\frac{1}{1-\nu}(Z - Z^\nu) + \frac{\rho}{1-\rho}(Z^{\nu+1} - Z^\nu)}{Z^\nu - 1}$$

By derivation, we then have:

$$\frac{(Z^\nu - 1)^2}{a}\frac{\partial E_J(Z)}{\partial Z} = \frac{1}{1-\nu}\left((1-\nu)Z^\nu + \nu Z^{\nu-1} - 1\right) + \frac{\rho}{1-\rho} Z^{\nu-1}\left(Z^{\nu+1} - (\nu+1)Z + \nu Z\right)$$

A(Z)=$(1-\nu)Z^\nu+\nu Z^{\nu-1}-1$ is negative, for every ν>1 and for every Z >1. Indeed, A(1)=0 and A'(Z)=$\nu(1-\nu)Z^{\nu-2}(Z-1)$ is negative. Moreover, B(Z)=$Z^{\nu+1}-(\nu+1)Z+\nu Z$ is positive, for every ν>1 and for every Z >1. Indeed, B(1)=0 and B'(Z)=$(\nu+1)(Z-1)$ is positive. $\frac{\partial E_J(Z)}{\partial Z}$ is thus positive, as well as $\frac{\partial E_J(t_\infty)}{\partial t_\infty}$ and E_J is thus increasing.

I Properties of Φ and link with erf

$$
\begin{aligned}
\phi(t) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \\
\phi'(t) &= -t\phi(t) \\
\Phi(-t) &= 1 - \Phi(t) \\
erf(t) &= 2\Phi(\sqrt{2}t) - 1 \\
\Phi(t) &= \frac{1}{2} + \frac{1}{2} erf\left(\frac{t}{\sqrt{2}}\right)
\end{aligned}
$$