

# A hybrid approach to computing the measure of success of computing tasks in a grid<sup>1</sup>

**Christer Carlsson**

IAMSR, Åbo Akademi University,  
Joukahaisenkatu 3-5 A, FIN-20520 Turku  
e-mail: christer.carlsson@abo.fi

**Robert Fullér**

IAMSR, Åbo Akademi University,  
Joukahaisenkatu 3-5 A, FIN-20520 Turku  
e-mail: rfuller@abo.fi

**József Mezei**

Turku Centre for Computer Science,  
Joukahaisenkatu 3-5 B, FIN-20520 Turku  
e-mail: jmezei@abo.fi

*Abstract:* In this paper we will show a hybrid (probabilistic and possibilistic) approach for assessing the success of a computing task in a grid environment.

## 1 Introduction

Grid computing is a form of distributed computing where clusters of networked, loosely-coupled computers, are acting in concert to perform very large and/or a large number of tasks [4]. However, running applications on the Grid environment poses significant challenges due to the diverse failures encountered during execution. If one computing node fails during the job execution, the whole job fails and has to be restarted. To handle resource failures and avoid restarting the job from the initial state, the fault-tolerance mechanisms checkpointing and migration have been developed. The Resource Management System makes a snapshot of the jobs execution state,

---

<sup>1</sup>Appeared in: *Proceedings of the 9-th International Symposium of Hungarian Researchers on Computational Intelligence (CINTI 2008)*, November 6-8, 2008, Budapest, Hungary

transfers the snapshot to another computing node, and resumes the job execution after all nodes failed have been replaced with nodes that are working [5]. Users negotiate for resource usage through a Grid resource broker which queries resource providers on their behalf to find suitable resources. They require a job execution with a desired level of priority and quality [3]. Failure intensity usually increases with age for mechanical equipment. Power law and loglinear Poisson processes are often used to model failure intensity. The distinguishing feature of Poisson processes is that the previous history of failure times  $t_1, \dots, t_n$  does not affect failure intensity [1].

## 2 Probability of success of a computing task

We consider an approximation to the probability that a particular computing task in a grid is successful. This happens if there will always be at least a single idle node available in the system in the case of a node failure. Let  $S = 1$  denote the event that the task is successful, and  $S = 0$  the opposite event. We formulate the probability of the success as the sum of the probabilities  $P(\text{"none of the nodes allocated to the task fail"}) + \sum_{m=1}^{m_{\max}} P(\text{"}m\text{ of the nodes allocated to the task fail \& at least }m\text{ idle nodes are available as reserves"})$ . Here  $m_{\max}$  is an upper limit for the number of failures considered. The value can be chosen by judging the size of the contribution of each event, determined by the corresponding probability. Thus, the sum can be simplified by considering only those events that do not have vanishingly small probabilities. A conservative bound for the success probability can be derived by assuming that the  $m$  failures take place simultaneously, which leads to

$$\begin{aligned}
 P(S = 1) &= 1 - P(S = 0) \\
 &= 1 - \sum_{m=1}^{m_{\max}} P(m \text{ failures occur \& less than } m \text{ free nodes available}) \\
 &= 1 - \sum_{m=1}^{m_{\max}} P(m \text{ failures occur})P(\text{less than } m \text{ free nodes available}) \\
 &\geq 1 - \sum_{m=1}^{m_{\max}} P(m \text{ failures occur})P(\text{less-than-}m\text{-anytime})
 \end{aligned}$$

Here *less-than- $m$ -anytime* stands for the event *less than  $m$  free nodes available at any time point*.

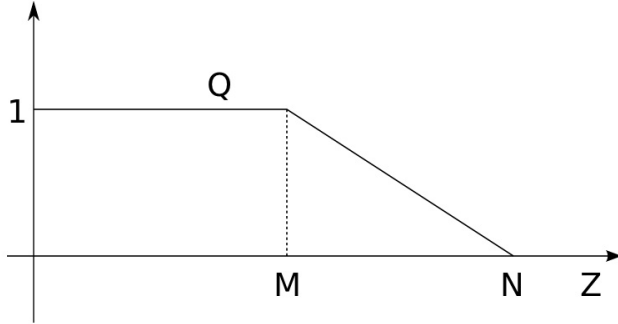


Figure 1: A linear possibility distribution for maximal number of failures.

### 3 A hybrid approach to success of computing tasks

Let us suppose that the number of all nodes in a grid is equal to  $Z$ . Let  $N$  be an upper limit for the number of failures for all classes and at any time during the simulations (this number is estimated by the broker). We will suppose that  $N = m_{\max}$ . Let  $M$  be the most possible value for the number of failures. Let  $S^* = 1$  denote the event that the task is successful. Let  $Q$  be the possibility distribution for maximal number of failures. That is,  $Q(m)$  is interpreted as the degree of possibility of the statement that the maximal number of failures is equal to  $m$ .

We use the notation

$$\Pi(Q = m) = Q(m).$$

That is,  $Q(m)$  denotes the degree of possibility to which  $m$  is considered to be the maximal number of failures.

Let  $\Omega$  be a finite set. We recall that a function  $\Pi: \Omega \rightarrow [0, 1]$  is said to be a possibility measure on  $\Omega$  if [2]

- $\Pi(\emptyset) = 0$
- $\Pi(\Omega) = 1$
- $\Pi(U \cup V) = \max\{\Pi(U), \Pi(V)\}$  for any  $U, V \subset \Omega$ .

It follows that the possibility measure on finite set is determined by its behavior on singletons:

$$\Pi(U) = \max_{\omega \in U} \Pi(\{\omega\}).$$

Now we compute the probability of success taking into consideration the possibility distribution for maximal number of failures derived from the broker's observations,

$$\begin{aligned}
P(S^* = 1) &= \\
&= 1 - \sum_{m=1}^N \Pi(Q = m) P(m \text{ failures occur \& less than } m \text{ free nodes available}) \\
&= 1 - \sum_{m=1}^N \Pi(Q = m) P(m \text{ failures occur}) P(\text{less-than-}m\text{-anytime}) \\
&\geq 1 - \sum_{m=1}^N \Pi(Q = m) P(m \text{ failures occur}) P(\text{less-than-}m\text{-anytime}) \\
&= 1 - \sum_{m=1}^M P(m \text{ failures occur}) P(\text{less-than-}m\text{-anytime}) \\
&\quad - \sum_{m=M+1}^N \Pi(Q = m) P(m \text{ failures occur}) P(\text{less-than-}m\text{-anytime})
\end{aligned}$$

Here, for simplicity we used the same notation  $P$  for the hybrid measure for success.

For example, if  $Q$  is linear

$$Q_{\text{linear}}(m) = \begin{cases} 1 & \text{if } m \leq M \\ 1 - \frac{m - M}{N - M} & \text{if } M \leq m \leq N \\ 0 & \text{if } N \leq m \leq Z \end{cases}$$

then using the equality,

$$Q_{\text{linear}}(m) = 1 - \frac{m - M}{N - M} = \frac{N - m}{N - M}$$

we get

$$\begin{aligned}
P(S^* = 1) &= 1 - \sum_{m=1}^M P(m \text{ failures occur}) P(\text{less-than-}m\text{-anytime}) \\
&\quad - \sum_{m=M+1}^N \frac{N - m}{N - M} \times P(m \text{ failures occur}) P(\text{less-than-}m\text{-anytime})
\end{aligned}$$

Since

$$P(S = 1) = 1 - \sum_{m=1}^N P(m \text{ failures occur})P(\text{less-than-}m\text{-anytime})$$

we get  $P(S^* = 1) \geq P(S = 1)$ .

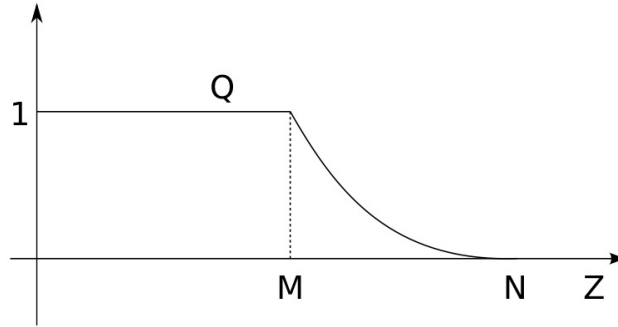


Figure 2: A quadratic possibility distribution for maximal number of failures.

Depending on the observations, the broker could also use a quadratic function for the possibility distribution of maximal number of failures.

$$Q_{\text{quadratic}}(m) = \begin{cases} 1 & \text{if } m \leq M \\ \left(\frac{N-m}{N-M}\right)^2 & \text{if } M \leq m \leq N \\ 0 & \text{if } N \leq m \leq Z \end{cases}$$

In this case, then we get

$$P(S^* = 1) = 1 - \sum_{m=1}^M P(m \text{ failures occur})P(\text{less-than-}m\text{-anytime}) - \sum_{m=M+1}^N \left(\frac{N-m}{N-M}\right)^2 \times P(m \text{ failures occur})P(\text{less-than-}m\text{-anytime})$$

The probability of success with a quadratic possibility distribution is bigger than in the linear case. It follows from the relationship

$$Q_{\text{quadratic}}(m) < Q_{\text{linear}}(m)$$

if  $m > M$ . That is, the quadratic possibility,  $Q_{\text{quadratic}}(m)$ , that  $m$  is the maximal number of failures diminishing more quickly than the linear possibility,  $Q_{\text{linear}}(m)$ , that that  $m$  is the maximal number of failures for any  $m > M$ .

## 4 Summary

We presented a hybrid probabilistic and possibilistic technique for assessing the success of a computing task in a grid environment. The probability of success in a hybrid environment is bigger than in the pure probabilistic environment since the hybrid approach takes into consideration the possibility distribution for maximal number of failures derived from the broker's observations.

## 5 Acknowledgment

This work has been partially supported by the EU within the 6th Framework Programme under contract IST-031772 *Advanced Risk Assessment and Management for Trustable Grids*, (AssessGrid).

## References

- [1] R.D. Baker: Calculating the expected failure rate of complex equipment subject to hazardous repair, *International Journal of Production Economics* 67 (2000) 53-61.
- [2] C. Baudrit, I. Couso, D. Dubois: Joint propagation of probability and possibility in risk analysis: Towards a formal framework, *International Journal of Approximate Reasoning*, 45 (2007) 82-105.
- [3] K. Djemame, I. Gourlay, J. Padgett, G. Birkenheuer, M.Hovestadt, O. Kao, K. Voss: Introducing Risk Management into the Grid, in: *Proceedings of the Second IEEE International Conference on e-Science and Grid Computing*, 2006 28-35.
- [4] I. Foster: What is the Grid? A Three Point Checklist, *GRIDToday*, 2002.

- [5] Kerstin Voss: Recursive Evaluation of Fault Tolerance Mechanisms for SLA Management, in: Proceedings of the Fourth International Conference on Networking and Services, 2008 223-229.