

# Reliability model of fault-tolerant data processing system with primary and backup nodes

P A Rahman<sup>1</sup>, E Yu Bobkova<sup>2</sup>

<sup>1</sup>Department of Automated Technological and Informational Systems, Ufa State Petroleum Technological University, Sterlitamak, 453118, Russia

<sup>2</sup>Department of Design and Humanities, Samara Cossack Institute of Food Industry and Business, Samara, 443031, Russia

E-mail: pavelar@yandex.ru

**Abstract.** This paper deals with the fault-tolerant data processing systems, which are widely used in modern world of information technologies and have acceptable overhead expenses in hardware implementation. A simplified reliability model for duplex systems and the offered by authors advanced model for data processing systems with primary and backup nodes based on a three-state model of recoverable elements, which takes into consideration different failure rates of passive and active nodes and finite time of node activation, are also given. A calculation formula for the availability factor of the dual-node data processing system with primary and backup nodes and calculation examples are also provided.

## 1. Introduction

In modern world a rapid development of information technologies and their implementation in different spheres of human activity is observed. Almost every day a person has to deal with information. He creates, stores, processes and transmits it using computers and mobile devices. Medium and large-scale enterprises use specialized data storage and processing systems. On their basis a set of information systems operate and assist the business processes of an enterprise.

Data processing systems are widely used in modern enterprises, especially high-availability clusters for database systems, which provide fault-tolerant data processing and storage. A cluster is a set of computers linked by a high-speed communication network and logically united by special software for distributed data processing. In practice the dual-node high-availability clusters with shared storage are used because of the acceptable reliability / cost ratio. To avoid the database access conflicts at any time only one node is active and it processes user requests, the other one is passive in a standby mode. For such systems, it is important to know their reliability for estimation of risks to the business processes. In this situation, the development of reliability models and the analysis of reliability indexes for fault-tolerant data processing systems is quite an important task.

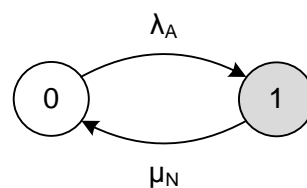
What concerns the reliability models, on the one hand, there are a number of academic books on the reliability theory [1, 2], in which the generalized reliability models of technical systems are discussed, but there are no specific examples related to modern data processing systems, including the high-availability clusters. On the other hand, a number of specialized books [3, 4], dedicated to reliability of computing systems and networks, discuss data processing systems, but the given reliability models for duplex systems are too simplified and provide overestimated values for reliability indexes.



Within the scope of the scientific research work of authors regarding reliability of data processing and transmission systems [5] a scientific task of development of a specialized reliability model for the dual-node data processing system with primary and backup nodes has been raised. This model is also considered for future application to obtained results in designing of data processing systems for industrial enterprises.

## 2. A simplified reliability model of the duplex system with independent nodes

In the well-known simplified reliability model of the duplex system [1-4] we consider data processing nodes as the simplest repairable elements with two states: up-state and down-state (Figure 1). From up-state a healthy node can pass to down-state with failure rate  $\lambda_A$ . From down-state a faulty node can pass to up-state with repair rate  $\mu_N$ .



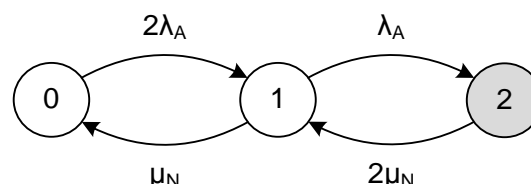
**Figure 1.** Reliability model of repairable element.

In the duplex system with two independent nodes both of them may be in one of two states independently and the system is considered to be ready for user requests, when at least one node is in up-state. One user request also can be processed on both of healthy nodes simultaneously.

Now let us introduce the following state-space for the duplex system and conditions for transition from one state to another according to a simplified reliability model:

- State 0 (up / up) – both nodes are up, and they are processing user requests. From this state system can pass to state 1 with rate  $2\lambda_A$  (failure of any healthy node).
- State 1 (up / down) – one node is up, and it is processing user requests, the other one is down. From this state the system can pass either to state 2 with rate  $\lambda_A$  (failure of the remaining healthy node) or to state 0 with rate  $\mu_N$  (recovery of the faulty node).
- State 2 (down / down) – both nodes are down, and user requests are not processed. From this state the system can pass to state 1 with rate  $2\mu_N$  (recovery of any faulty node).

Figure 2 shows Markov chain, which represents the state-space and transitions conditions according to a simplified reliability model:



**Figure 2.** A simplified reliability model of the duplex system with independent nodes.

Accordingly, the differential equations system of Kolmogorov-Chapman for this Markov chain is as follows:

$$\left\{ \begin{array}{l} P_0(0) = 1; \quad P_1(0) = 0; \quad P_2(0) = 0; \\ P_0(t) + P_1(t) + P_2(t) = 1; \\ \frac{dP_0(t)}{dt} = -2\lambda_A P_0(t) + \mu_N P_1(t); \\ \frac{dP_1(t)}{dt} = 2\lambda_A P_0(t) - (\lambda_A + \mu_N) P_1(t) + 2\mu_N P_2(t); \\ \frac{dP_2(t)}{dt} = \lambda_A P_1(t) - 2\mu_N P_2(t). \end{array} \right.$$

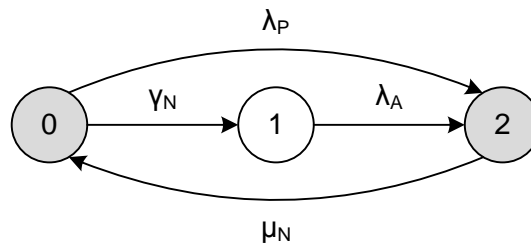
Considering states 0 and 1 as system healthy states and  $t \rightarrow \infty$ , we obtain the following simple formula for the stationary availability factor of the duplex system with independent nodes:

$$K_{DS} = P_0(\infty) + P_1(\infty) = \frac{\mu_N(\mu_N + 2\lambda_A)}{(\mu_N + \lambda_A)^2}. \quad (1)$$

### 3. Advanced reliability model of system with primary and backup nodes

In the advanced model of the dual-node system, we consider the finite time of node activation (node initialization and reconfiguration before it will be ready for user requests), and different failure rates for active and passive nodes. We regard nodes as special elements with three states: active, passive and down (Figure 3). Only in the active state the node can process user requests.

From the passive state the node can pass either to the down state with passive failure rate  $\lambda_P$  or to the active state with activation rate  $\gamma_N$ . From the active state the node can pass to the down state with active failure rate  $\lambda_A$ . From the down state the node can pass to the passive state with recovery rate  $\mu_N$ .



**Figure 3.** Reliability model of three-state element.

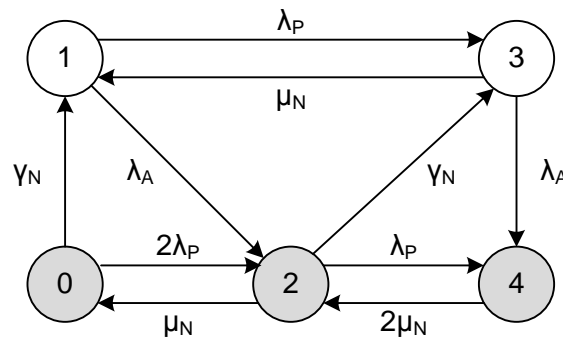
In the advanced model we also suppose that only one of the nodes can be active and the other node must be passive. This is a well-known approach in modern high-availability systems with shared storage, where only one data processing node has access to the system database for processing user requests. So, if both of the nodes are in the passive state and ready to become active, then only one of them (assigned by a system engineer as primary) will pass to the active state, the other node (assigned as backup) will stay in the passive state. As for failure and recovery we assume that both nodes are independent. The advanced model is a further development of reliability models obtained by authors earlier [5].

Now let us introduce the following state-space for the system with primary and backup nodes, and conditions for transition from one state to another according to the advanced reliability model:

- State 0 (passive / passive) – both nodes are passive, user requests are not processed. From this state the system can pass either to state 1 with rate  $\gamma_N$  (activation of the passive node assigned as primary), or to state 2 with rate  $2\lambda_P$  (failure of any passive node).

- State 1 (active / passive) – one node is active and the other one is passive, user requests are processed. From this state the system can pass either to state 2 with rate  $\lambda_A$  (failure of the active node), or to state 3 with rate  $\lambda_P$  (failure of the passive node).
- State 2 (passive / down) – one node is passive and the other one is down, user requests are not processed. From this state the system can pass either to state 3 with rate  $\gamma_N$  (activation of the passive node), or to state 4 with rate  $\lambda_P$  (failure of the passive node), or to state 0 with rate  $\mu_N$  (recovery of the faulty node).
- State 3 (active / down) – one node is active and other one is down, user requests are processed. From this state the system can pass either to state 4 with rate  $\lambda_A$  (failure of the active node), or to state 1 with rate  $\mu_N$  (recovery of the faulty node).
- State 4 (down / down) – both nodes are down, user requests are not processed. From this state the system can pass to state 2 with rate  $2\mu_N$  (recovery of any faulty node).

In Figure 4 Markov chain, which represents the state-space and transitions conditions according to the advanced reliability model, is shown:



**Figure 4.** Advanced reliability model of system with primary and backup nodes.

Accordingly, the differential equations system of Kolmogorov-Chapman for this Markov chain is as follows:

$$\left\{ \begin{array}{l} P_0(0) = 1; \quad P_1(0) = 0; \quad P_2(0) = 0; \quad P_3(0) = 0; \quad P_4(0) = 0; \\ P_0(t) + P_1(t) + P_2(t) + P_3(t) + P_4(t) = 1; \\ \frac{dP_0(t)}{dt} = -(2\lambda_P + \gamma_N)P_0(t) + \mu_N P_2(t); \\ \frac{dP_1(t)}{dt} = \gamma_N P_0(t) - (\lambda_A + \lambda_P)P_1(t) + \mu_N P_3(t); \\ \frac{dP_2(t)}{dt} = 2\lambda_P P_0(t) + \lambda_A P_1(t) - (\mu_N + \lambda_P + \gamma_N)P_2(t) + 2\mu_N P_4(t); \\ \frac{dP_3(t)}{dt} = \lambda_P P_1(t) + \gamma_N P_2(t) - (\mu_N + \lambda_A)P_3(t); \\ \frac{dP_4(t)}{dt} = \lambda_P P_2(t) + \lambda_A P_3(t) - 2\mu_N P_4(t). \end{array} \right.$$

Where,

$\lambda_A$  is the failure rate of the active node.

$\lambda_P$  is the failure rate of the passive node.

$\gamma_N$  is the activation rate of the passive node.

$\mu_N$  is the recovery rate of the faulty node.

Considering states 1 and 3 as system healthy states and  $t \rightarrow \infty$ , we obtain the following advanced formula for the stationary availability factor of the system with primary and backup nodes:

$$K_{AP} = \frac{\gamma_N}{(\gamma_N + \lambda_A) \left( 1 + \frac{\lambda_A(\gamma_N + \lambda_P)(2\mu_N\lambda_P + (\gamma_N + 2\lambda_P)(\lambda_A + \lambda_P))}{2\mu_N(\gamma_N + \lambda_A)(\mu_N + \gamma_N + 2\lambda_P)(\mu_N + \lambda_A + \lambda_P)} \right)}. \quad (2)$$

Note. In case of infinite activation rate  $\gamma_N \rightarrow \infty$  (immediate activation of the passive node), a formula for the availability factor is simplified:

$$K_{AP} \rightarrow \frac{2\mu_N(\mu_N + \lambda_A + \lambda_P)}{2\mu_N(\mu_N + \lambda_A + \lambda_P) + \lambda_A(\lambda_A + \lambda_P)}.$$

Moreover, in this case if failure rates for passive and active nodes are also the same  $\lambda_A = \lambda_P$ , then the formula is simplified to the expression for duplex systems:  $K_{AP} \rightarrow \frac{\mu_N(\mu_N + 2\lambda_A)}{(\mu_N + \lambda_A)^2}$ .

#### 4. Availability factor calculation example

A data processing system with primary and backup nodes is given. The failure rate of the active node is  $\lambda_A = 1/8760 \text{ h}^{-1}$  (on average one failure per year). The recovery rate of the node is:  $\mu_N = 1/24 \text{ h}^{-1}$  (on average one recovery in 24 h).

For passive node failure and activation rates let us consider three cases:

- 1) Hot-spare nodes with fast activation: passive node failure rate  $\lambda_P = 1/8760 \text{ h}^{-1}$  and activation rate  $\gamma_N = 1200 \text{ h}^{-1}$  (on average one activation per 3 s).
- 2) Warm-spare nodes with intermediate activation: passive node failure rate  $\lambda_P = 1/17520 \text{ h}^{-1}$  and activation rate  $\gamma_N = 20 \text{ h}^{-1}$  (on average one activation per 3 minutes).
- 3) Cold-spare nodes with slow activation: passive node failure rate  $\lambda_P = 0 \text{ hour}^{-1}$  and activation rate  $\gamma_N = 1/3 \text{ h}^{-1}$  (on average one activation per 3 h).

According to the simplified model of the duplex system in all three cases the following value for the availability factor is obtained by using formula 1:

$$K_{DS} \approx 0.9999925349$$

According to the advanced model of the system with primary and backup nodes the following values for the availability factor are obtained by using formula 2:

In case of hot-spare nodes with fast activation:

$$K_{AP} \approx 0.9999924397$$

In case of warm-spare nodes with intermediate activation:

$$K_{AP} \approx 0.9999886897$$

In case of cold-spare nodes with slow activation:

$$K_{AP} \approx 0.9996543268$$

It can be easily noticed that the advanced model (obtained by the authors), which takes into consideration additional reliability parameters of the system with primary and backup nodes, gives more low and realistic estimation of the availability factor rather than the well-known simplified model.

## 5. Summary

Within the scope of this article a well-known reliability model of duplex system and an advanced model for the data processing system with primary and backup nodes obtained by the authors are presented. The formula for availability factor estimation and calculation examples are also given.

Scientific results obtained by the authors were used for designing of data processing systems of Moscow Power Engineering Institute, Nuclear Power Plant 'Balakovo', Joint Stock Company 'Krasnyi Proletariy' and several other enterprises.

## 6. Acknowledgements

Authors thank professor A I Kayashev, Sterlitamak branch of Ufa State Petroleum Technological University and professor I I Ladygin, Moscow Power Engineering Institute, for scientific support and theoretical base, also S N Khorkov, a chief system administrator of MPEI corporate network, and R M Gafarov, a chief of information technologies department of Joint Stock Company 'Krasnyi Proletariy', for technical support and experimental base.

## 7. References

- [1] Cherkesov G N 2005 *Reliability of Hardware and Software Systems* (Saint-Petersburg: Piter)
- [2] Polovko A M and Gurov S V 2006 *Basis of Reliability Theory* (Saint-Petersburg: BHV-Petersburg)
- [3] Shooman M L 2002 *Reliability of computer systems and networks* (New York: John Wiley & Sons, Inc.)
- [4] Koren I and Krishna C M 2007 *Fault-Tolerant Systems* (San Francisco: Morgan Kaufmann Publishers)
- [5] Rahman P A, Muraveva E A and Sharipov M I 2016 *Key Engineering Materials* **685** 805–810