SimEDC: A Simulator for the Reliability Analysis of Erasure-Coded Data Centers (Supplementary File)

Mi Zhang, Shujie Han, and Patrick P. C. Lee

The following materials provide supplementary results to our main file.

1 Impact of Cross-Rack Bandwidth

In our main paper, our reliability analysis focuses on the case where the cross-rack bandwidth is 1 Gb/s. In this section, we conduct reliability analysis by varying the cross-rack bandwidth, especially when the cross-rack bandwidth is higher than 1 Gb/s.

1.1 Independent Failures

We vary the cross-rack bandwidth and evaluate the reliability of different erasure code settings. Table 1 shows the results for the cross-rack bandwidth of 400 Mb/s, 2 Gb/s, 5 Gb/s and 10 Gb/s when there are independent failures only. When the cross-rack bandwidth is 400 Mb/s, the erasure code settings under flat placement have PDL equal to (or nearly equal to) one (i.e., the data loss always occurs), while DRC(9,6,3) has PDL equal to 1.26e-2. Thus, it is important to improve reliability by minimizing the cross-rack repair traffic under limited cross-rack bandwidth.

When the cross-rack bandwidth increases to 2 Gb/s, the PDL decreases by two to four orders of magnitude, and some erasure code settings (e.g., RS(14,10)) do not observe any data loss in our simulation process. When the cross-rack bandwidth further increases to 5 Gb/s and 10 Gb/s, SimEDC does not observe any data loss within 20,000 iterations for all erasure code settings when there are independent failures only. The high cross-rack bandwidth implies that the repair time decreases (i.e., the repair performance improves), thereby improving the overall reliability.

1.2 Correlated Failures

We now analyze the reliability by adding permanent correlated failures to our simulation in addition to independent failures (as in our evaluation in the main file), while we set the cross-rack bandwidth as 5 Gb/s and 10 Gb/s. Table 2 presents the results in the presence of both independent and correlated failures. Some erasure code settings do not observe any data loss (e.g., RS(14,10)). However, in some cases, hierarchical placement has higher PDL than flat placement as it sacrifices rack-level fault tolerance (e.g., $r = 3$ versus $r = 9$ in RS(9,6), or $r = 4$ versus $r = 16$ in LRC(16,12)). The relative differences between flat and hierarchical placements are consistent with those that we observe from the case of the cross-rack bandwidth of 1 Gb/s.

1.3 Importance Sampling

We finally analyze the reliability using importance sampling when we set the cross-rack bandwidth as 5 Gb/s and 10 Gb/s; we refer readers to the main paper for our evaluation setup. Table 3 presents the results with and without importance sampling. The PDL when the cross-rack bandwidth is 5 Gb/s and 10 Gb/s is significantly lower than the case where the cross-rack bandwidth is 1 Gb/s (see our main paper). In addition, the relative differences across different erasure codes and placement policies still hold.

• *M. Zhang, S. Han, and P. Lee are with the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong (E-mails: millyz0204@gmail.com, {sjhan,pclee}@cse.cuhk.edu.hk).*

TABLE 1
Reliability under independent failures only for different cases of the cross-rack bandwidth.

Erasure codes	400 Mb/s			2 Gb/s			5 Gb/s			10 Gb/s		
	PDL	**NOMDL**	**BR**	**PDL**	**NOMDL**	**BR**	**PDL**	**NOMDL**	**BR**	**PDL**	**NOMDL**	**BR**
RS(9,6), $r = 9$	9.97e-1	1.28e-6	1.60e-3	2.60e-3	3.31e-9	1.53e-4	0	0	5.80e-5	0	0	2.85e-5
RS(9,6), $r = 3$	2.84e-1	3.71e-7	1.14e-3	5.00e-4	6.36e-10	9.89e-5	0	0	3.83e-5	0	0	1.90e-5
RS(14,10), $r = 14$	1	1.73e-6	3.82e-3	0	0	2.52e-4	0	0	9.20e-5	0	0	4.48e-5
RS(14,10), $r = 7$	1	1.77e-6	3.24e-3	0	0	2.23e-4	0	0	8.23e-5	0	0	4.02e-5
RS(16,12), $r = 16$	1	1.83e-6	4.56e-3	3.00e-4	5.36e-10	2.95e-4	0	0	1.06e-4	0	0	5.14e-5
RS(16,12), $r = 4$	1	1.87e-6	2.85e-3	0	0	2.11e-4	0	0	7.83e-5	0	0	3.83e-5
LRC(16,12), $r = 16$	1	1.32e-6	1.61e-3	1.40e-3	2.00e-9	1.53e-4	0	0	5.80e-5	0	0	2.86e-5
LRC(16,12), $r = 4$	2.45e-1	3.33e-7	1.07e-3	1.00e-3	1.43e-9	9.61e-5	0	0	3.72e-5	0	0	1.85e-5
DRC(9,6,3)	1.26e-2	1.47e-8	7.43e-4	0	0	4.81e-5	0	0	1.90e-5	0	0	9.40e-6

TABLE 2
Reliability under both independent and correlated failures for the cross-rack bandwidth of 5 Gb/s and 10 Gb/s.

Erasure codes	5 Gb/s			10 Gb/s		
	PDL	**NOMDL**	**BR**	**PDL**	**NOMDL**	**BR**
RS(9,6), $r = 9$	0	0	1.78e-3	0	0	1.74e-3
RS(9,6), $r = 3$	8.00e-3	2.35e-7	1.75e-3	2.50e-3	3.18e-9	1.73e-3
RS(14,10), $r = 14$	0	0	1.80e-3	0	0	1.75e-3
RS(14,10), $r = 7$	0	0	1.80e-3	0	0	1.75e-3
RS(16,12), $r = 16$	5.00e-4	8.94e-10	1.83e-3	0	0	1.77e-3
RS(16,12), $r = 4$	0	0	1.79e-3	0	0	1.75e-3
LRC(16,12), $r = 16$	0	0	1.77e-3	0	0	1.74e-3
LRC(16,12), $r = 4$	6.10e-2	6.59e-7	1.75e-3	6.30e-2	7.17e-7	1.73e-3
DRC(9,6,3)	3.50e-3	6.52e-8	1.73e-3	3.50e-3	9.79e-8	1.72e-3

TABLE 3
Reliability measured by importance sampling for the cross-rack bandwidth of 5 Gb/s and 10 Gb/s.

Erasure codes	PDL **with 5 Gb/s**	PDL **with 10 Gb/s**
RS(9,6), $r = 9$	1.69e-6	1.01e-19
RS(9,6), $r = 3$	1.09e-8	1.39e-45
RS(14,10), $r = 14$	2.20e-8	1.43e-18
RS(14,10), $r = 7$	2.58e-9	2.06e-26
RS(16,12), $r = 16$	2.42e-7	6.19e-13
RS(16,12), $r = 4$	1.19e-8	1.37e-21
LRC(16,12), $r = 16$	1.96e-7	1.23e-10
LRC(16,12), $r = 4$	5.01e-18	0
DRC(9,6,3)	3.45e-10	0