Modeling Spatially Correlated Survival Data Under Different Distance Functions

Jared Clark

June 18, 2024

Joint work with Dr. Yili Hong, Dr. George Ostrouchov, Dr. Yueyao Wang, and Dr. Jie Min

Table of Contents

1 Introduction

2 Model

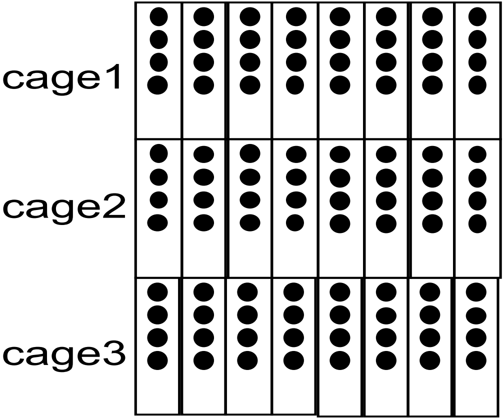
3 Results

Data Introduction

- The data set contains failure information on 19,319 GPUs within the Titan Cray XK7 high perfomance computing (HPC) cluster.
- The GPUs are housed within cabinets.
- Cabinets are spread across 8 rows and 25 columns (200 locations total).
- GPUs experience two failures modes: double bit errors (DBE) and off the bus failures (OTB).
- This work focuses solely on OTB failures, resulting in a censoring rate of 94.2%.

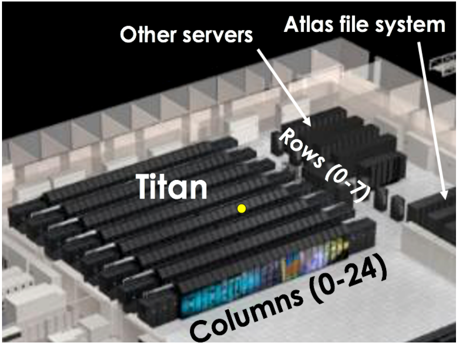
GPU Cabinets

Within a cabinet, 96 GPUs are organized by cage, slot and node.



Cabinet Locations

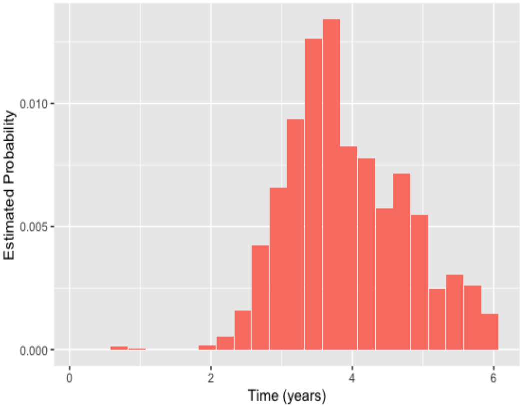
A visual representation of the 200 locations.



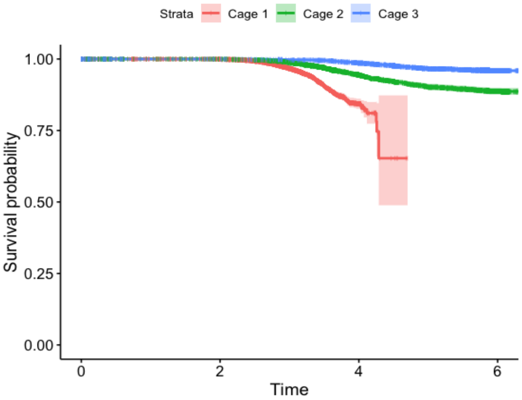
*[1]

[1] G. Ostrouchov, D. Maxwell, R. A. Ashraf, C. Engelmann, M. Shankar and J. H. Rogers, "GPU Lifetimes on Titan Supercomputer: Survival Analysis and Reliability," SC20: International Conference for High Performance Computing, Networking, Storage and Analysis, Atlanta, GA, USA, 2020, pp. 1-14, doi: 10.1109/SC41405.2020.00045.

OTB Failure Time Distribution



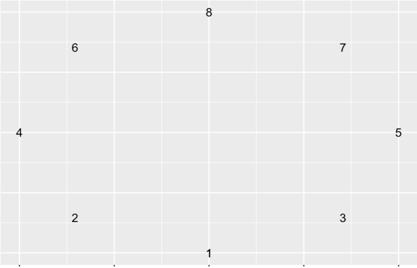
Differences in Cages

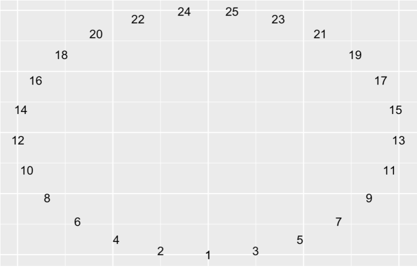


The log-rank test produces a p-value < 0.01.

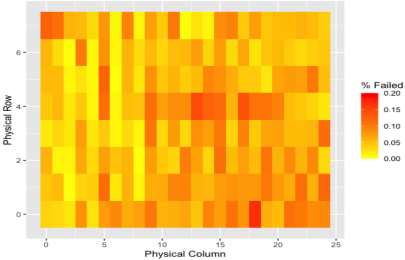
Cabinet Logical Connectivity

The cabinets are connected with cables that form a circuit in both the rows and columns.





Censoring by Location



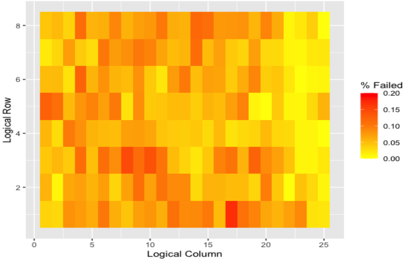


Table of Contents

1. Introduction
2. Model
3. Results

Inferential Model

$$\boldsymbol{s} = X\boldsymbol{\beta} + Z_v\boldsymbol{v} + Z_w\boldsymbol{w} + \boldsymbol{\epsilon}$$

- $\boldsymbol{s}$ is a n-dimensional vector containing the log failure times.
- X is an $n \times (p+1)$ matrix containing the explanatory information associated with cage, slot and node.
- Z_v and Z_w are $n \times m$ matrices linking the observations to their spatial locations.
- The vectors $\boldsymbol{\beta}, \boldsymbol{v}, \boldsymbol{w}$ and $\boldsymbol{\epsilon}$ are unknown with distributions to defined in subsequent slides. These are assumed to be independent.
- Parameters $\sigma, \sigma_v^2, \sigma_w^2, \nu_{v1}, \nu_{v2}, \nu_{w1}, \nu_{w2}, \kappa_v$ and κ_w control the covariance structure and will be furthered defined subsequently.

Prior Specification

The following distributional forms are assumed:

- $\epsilon_i \overset{iid}{\sim} \phi(0, \sigma)$ for $i = 1, 2, ..., n$
- $\boldsymbol{v} \sim N(\mathbf{0}, \sigma_v^2 R_v)$
- $\boldsymbol{w} \sim N(\mathbf{0}, \sigma_w^2 R_w)$
- $\boldsymbol{\beta} \sim N(\mathbf{0}, 500^2 I)$
- $\sigma, \sigma_w^2 \sim Gamma(0.1, 0.1)$
- $\sigma_v^2 \sim Gamma(0.5, 0.5)$
- $\nu_{v1}, \nu_{v2} \sim Gamma(5, 5)$
- $\nu_{w1}, \nu_{w2} \sim Gamma(2, 2)$
- $\kappa_w \sim Beta(0.5, 0.5)$
- $\lambda_v = -\log\left(\frac{2}{\kappa_v} - 1\right) \sim Gamma(1, 2)$

Physical Distance and Correlation

$$d_r^P(r_s, r_t) = |r_s - r_t|, \; d_c^P(c_s, c_t) = |c_s - c_t|$$

$$(R_v)_{st} = \exp\left(-\left(\frac{d_r^P}{\nu_{v1}}\right)^{\kappa_v} - \left(\frac{d_c^P}{\nu_{v2}}\right)^{\kappa_v}\right)$$

$$0 < \kappa_v \leq 2$$

Logical Distance and Correlation

$$d_r^{\mathrm{L}}(r_s^*, r_t^*) = \min\{|r_s^* - r_t^*|, n_r - |r_s^* - r_t^*|\}$$

$$d_c^{\mathrm{L}}(c_s^*, c_t^*) = \min\{|c_s^* - c_t^*|, n_c - |c_s^* - c_t^*|\}$$

$$(R_w)_{st} = \exp\left(-\left(\frac{d_r^L}{\nu_{w1}}\right)^{\kappa_w} - \left(\frac{d_c^L}{\nu_{w2}}\right)^{\kappa_w}\right)$$

$$0 < \kappa_w \leq 1$$

The use of circle distance for logical connections requires a stricter range on κ_w to ensure positive-definiteness.

Logical Distance Visualization

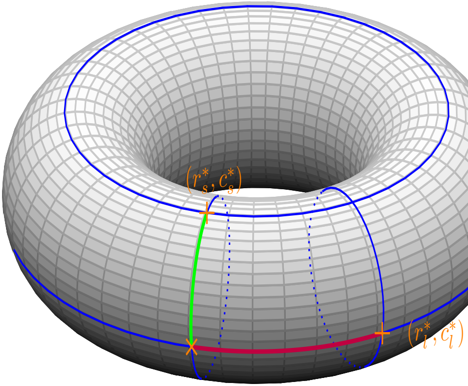


Table of Contents

1 Introduction

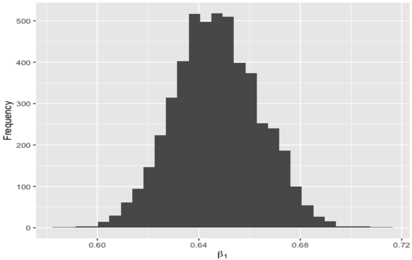
2 Model

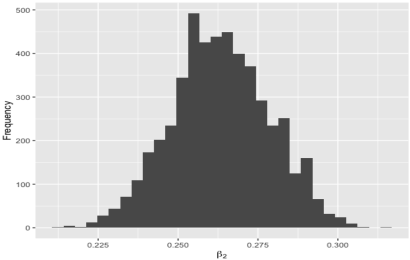
3 Results

Posterior Summary 1

Name	Parameter	Mean	SD	2.5% Quant	97.5% Quant
Intercept	β_0	2.002	0.030	1.947	2.060
Cage 2	β_2	0.263	0.015	0.234	0.293
Cage 3	β_1	0.647	0.017	0.614	0.679
Slot 2	β_3	0.046	0.027	-0.008	0.097
Slot 3	β_4	0.039	0.027	-0.017	0.088
Slot 4	β_5	0.062	0.026	0.011	0.112
Slot 5	β_6	0.052	0.026	0.002	0.101
Slot 6	β_7	0.075	0.026	0.025	0.124
Slot 7	β_8	0.019	0.026	-0.034	0.069
Slot 8	β_9	0.003	0.025	-0.045	0.052
Node 2	β_{10}	-0.272	0.020	-0.311	-0.231
Node 3	β_{11}	-0.302	0.020	-0.311	-0.231
Node 4	β_{12}	-0.063	0.020	-0.102	-0.025

Cage Effect Posteriors

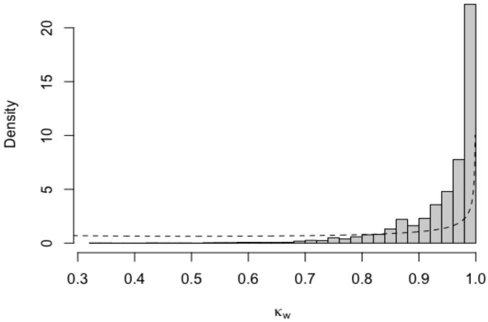




Posterior Summary 2

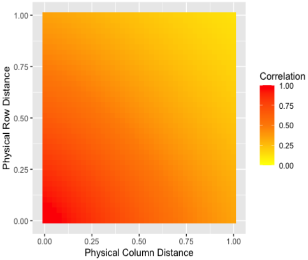
Parameter	Mean	SD	2.5% Quantile	97.5% Quantile
σ^2	0.182	0.009	0.166	0.200
σ_v^2	0.022	0.018	0.004	0.064
ν_{v1}	0.964	0.351	0.424	1.805
ν_{v2}	0.964	0.369	0.426	1.839
κ_v	1.297	0.174	1.016	1.635
σ_w^2	0.012	0.006	0.004	0.028
ν_{w1}	1.876	0.799	0.716	3.834
ν_{w2}	0.598	0.307	0.214	1.346
κ_w	0.945	0.073	0.743	1.000

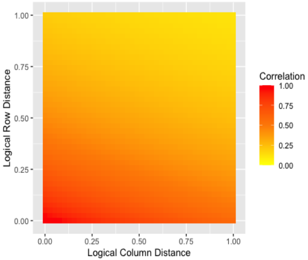
Exponential Correlation



There is a 0.65 posterior probability that $\kappa_w > 0.95$. This suggests that the exponential correlation function might be a good fit for the logical random effects.

Posterior Correlation Functions





Directions for Continued Work

- Compare results under other log location scale likelihoods, such as the smallest extreme value (SEV) distribution.
- Demonstrate model validity through extensive simulation study.
- The multiple spatial random effects can be integrated into existing competing risks models to model both OTB and DBE failures together.

Conclusion

- GPUs located higher in the cabinet, farther away from the cooling unit, fail at a faster rate.
- The relative locations of cabinets in physical space, and the separation along the cable circuit are both structures that create dependency in failure times.
- Any observed slot effects were much smaller than the cage and node effects.