System level Reliability modeling of Liquid cooled Data Centers

1st Sidharth Rajeev, 2nd Srikanth Rangarajan, 3rd Achyud, 4th Tiwei Wei, 5th Bahgat Sammakia
dept. Mechanical Eng. Binghamton University, Binghamton, USA
dept. Mechanical Eng. PURDUE University (of Aff. 3), Lafayette
srangar@binghamton.edu

***Abstract*—The focus of this research lies in developing a methodology for designing a cooling system for data centers, with a primary emphasis on reliability and efficiency. In this study, we conduct a thorough examination of various stochastic models that are employed for the purpose of predicting reliability. Following this, we explore the relationship between system availability and the reliability of an innovative component that has been integrated into the system. Within the framework of cooling systems utilized in data centers, this innovative component is commonly referred to as the cold plate. In the case of a liquid-cooled cold plate, we articulate the specific failure modes that the system may potentially exhibit. To facilitate this analysis, we utilize the conventionally employed Failure Modes and Effects Analysis (FMEA). In particular, three failure modes—Erosion, Corrosion, and Clogging—are meticulously analyzed within this discussion. We present several of the most effective models currently available for quantifying these failure modes.**

***Index Terms*—reliability, cold-plate, failure mode, data center cooling, modeling**

I. Introduction

As the trend in AI and cloud based computing rises, it is very evident that the global usage of computational performance is going to grow exponential in the coming decade. According to Sherwood [1] big tech companies are spending more on AI infrastructure, nearly 60% more than the same quarter of year 2023. More computational demand also means a rise in thermal dissipation power(TDP) for the newer generations of GPU's and CPU's. To manage and enable high performance computing (HPC), a more advance cooling technique will be required. This calls for a shift from a traditional air cooling to liquid cooling technology for data centers. A transition to liquid cooling is not an easy task as it requires a lot of complex subsystems which are interdependent on each other. As it sounds, having more complexity introduces reliability issues which undermines the benefits of a liquid cooling system. Fig.1 is a hypothetical plot on where each cooling techniques in data center stands. An ideal transition from air cooling to liquid cooling would be er thermal performance with no loss in reliability (state 1 to state 3). This transition can be made possible with careful planning, system modeling, predictive reliability modeling, and component level performance analysis.

Organizations are establishing data centers in Arctic locales to leverage ecological benefits. This strategic initiative is driven by the rising demand for cooling solutions within data centers. As the requirement for increasingly sophisticated data centers escalates, the concurrent management of energy consumption and operational reliability emerges as a considerable challenge. The impending era calls for enhanced focus on efficiency, the dependability of cooling systems, and the accessibility of extensive statistical failure data. There exists a relative deficiency of academic literature addressing the reliability of cooling systems in data centers compared to studies centered on energy efficiency. A thorough analysis of energy efficiency and reliability within data centers is articulated in the review paper [2]. Energy consumption models utilized for cooling systems have diverse applications, including energy consumption management, optimization, heat recovery, and thermal regulation. These models are pivotal for augmenting the efficiency of cooling systems in data centers. The authors specify five distinct reliability metrics, namely mean time between failure (MTBF), mean time to repair (MTTR), availability, severity, and risk (measured in terms of financial losses resulting from failures), for assessing data center reliability. A practical case study presented in [3] examines various cooling architectures in data centers, emphasizing the variances in availability, cost, and operational exergy consumption. The investigation highlights the necessity of integrating both energy efficiency and reliability in the design and functionality of data centers to attain optimal performance. The manuscript elaborates on the implementation of a hybrid modeling approach that merges Stochastic Petri Nets (SPN) and Reliability Block Diagrams (RBD) to assess the dependability of cooling infrastructures in data centers. This methodology aids in quantifying the effects of different design selections on both energy efficiency and reliability, offering a comprehensive framework for informed decision-making. While improving reliability often necessitates the incorporation of redundancy, this may result in heightened resource utilization and energy consumption, potentially adversely affecting sustainability. Consequently, data center designs must judiciously navigate the trade-offs between energy efficiency and reliability to optimize both operational efficacy and environmental ramifications. Control algorithms significantly impact not only the efficiency of a system but also its reliability. A comparative analysis of system-level and component-level control strategies, as presented in [4], highlights this influence. The study employs the fluid network model(FNM) tool to investigate the impact of various control combinations on data center cooling energy

ARPA-E Cooler chip program

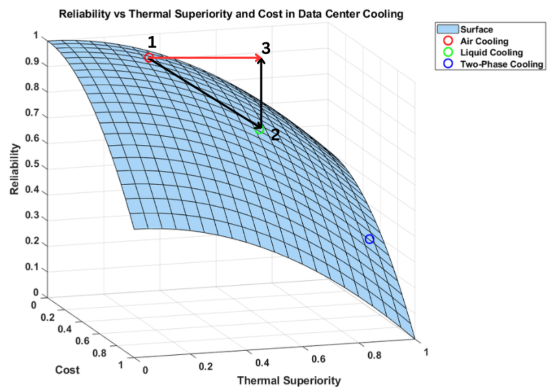

Fig. 1: Ideal transition from air cooling to liquid cooling

consumption and IT equipment reliability. When employing system-level controls exclusively, the desired operational conditions are maintained for approximately 50% of the simulated duration. However, integrating system-level with ON/OFF controls yields a reliability rate of 100%, while also achieving the lowest Power Usage Effectiveness (PUE) for the cooling system, measured at 1.13. This synergistic approach effectively captures system dynamics, minimizes the risk of over-cooling, and ensures that IT equipment consistently operates within the specified conditions throughout the simulation period. This also suggests that a system level approach and modeling can have a huge benefit on improving the cooling performance and reliability.

The cooling systems implemented in data centers function continuously and constitute approximately 30 to 50% of total energy consumption [3], with their primary objective being the maintenance of temperatures that facilitate optimal server performance. AS per [5] an unanticipated disruption in operations incurs costs for data centers that are projected to be around \$9000 per minute. This financial burden is expected to increase with the proliferation of computational devices, thereby necessitating a comprehensive evaluation of the reliability and availability of cooling systems. A term maximum Allowable Downtime (MADT) has been introduced as a metric within the domain of data center reliability research to establish standards for cooling systems. Wang *et al.* [5] utilized reliability block diagrams as a methodological approach to evaluate the reliability of cooling systems. This approach enabled improvements in reliability through the integration of redundancies where considered necessary. Cho *et al.* [6] in their research discuss the thermal performance of data centers under fault conditions. This study concentrated predominantly on computer room air handling (CRAH) units, with the principal conclusion underscoring the importance of backup systems and the necessity for prompt response from these systems during outages.

This manuscript concentrates on the methodology for evaluating and modeling the dependability of the cooling apparatus and the cold plate. The ARPA-E COOLERCHIP initiative has delineated stringent objectives [7] that must be attained to enhance the efficacy of data centers. Table I encapsulates the targets from the SOPO. Furthermore, the cold plate under investigation exemplifies a direct-to-chip liquid cooling technology.

TABLE I: ARPA-E Coolerchip Specifications

Specification	Requirement
Cooling System Energy Efficiency	$\leq$ 5%
Rack Power Density	$\geq$ 80 kW/m^3
Target Ambient Condition	**Dry Bulb Temp:** 40 °C **Humidity:** 60%
Thermal Resistance (Chip-to-Coolant)	$\leq$ 0.01 K/W
System Availability	> 99.982% (Tier 3 Uptime)

II. Preliminaries

To achieve these stringent targets (Table I), it is essential to design systems that are inherently reliable and robust. Although this is a complex task, scientific tools can help to understand the critical links within a system. One such approach involves using block representations to model systems and their components. To start with, an idea of the architecture of the system should be fixed. A block diagram of the architecture of the cooling system is depicted in Fig. 2. The system consists of a primary cooling loop which is the facility water side and the secondary loop recirculates through the cold plate and exchanges the heat picked up to the primary loop. Subsequently, the thermal energy is discharged to the ambient environment through the utilization of a dry cooler.

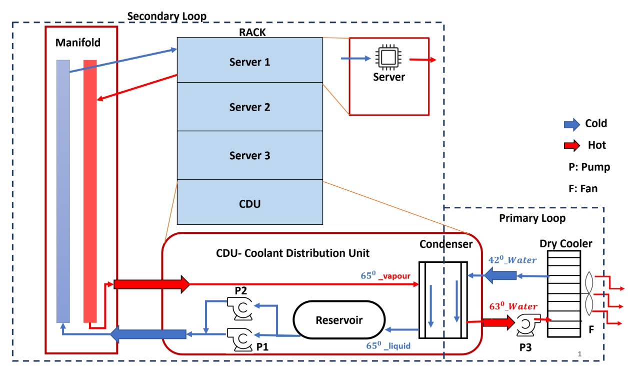

Fig. 2: Cooling System Architecture

For instance, a parallel arrangement, depicted in Fig. 3, employs multiple components with identical functionalities, where some are designated as standby. In this configuration, the system continues to function as long as at least one component remains operational, thereby improving the reliability of the system. The reliability of a parallel system is quantified by Equation (1).

$$R_S = 1 - \prod(1-R_i) = 1-(1-R_1)(1-R_2)\cdots(1-R_n) \quad (1)$$

In the series arrangement illustrated in Figure 4, components are connected in sequence, such that failure of any single component results in failure of the entire system. This configuration is highly susceptible to a single point of failure, significantly reducing the overall system reliability. The system's reliability is determined by multiplying the reliabilities of its individual components, as quantified by Equation (2).

$$R_S = \prod(R_i) = (R_1)(R_2)\cdots(R_n) \quad (2)$$

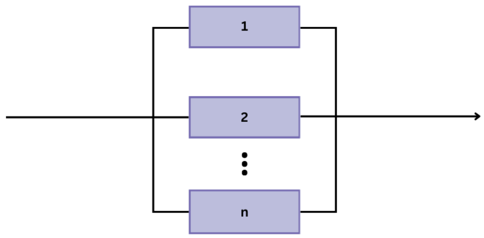

Fig. 3: Parallel Configuration

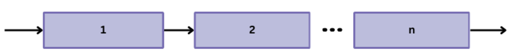

Fig. 4: Series Configuration

Utilizing the principles of series and parallel arrangements to delineate components, we construct a block representation of the cooling architecture under consideration, as illustrated in 2. The methodology employing a block-diagram framework is depicted in fig.5, wherein HX denotes the heat exchanger responsible for transferring thermal energy from the secondary to the primary system. The components are depicted as discrete blocks, interconnected in accordance with their interdependencies and the flow of information. This representation gives a thorough understanding of the system and aids in the simplification across diverse configurations.

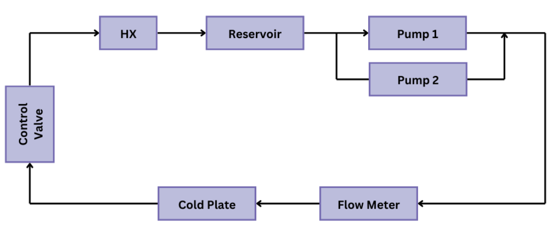

Fig. 5: Block-diagram representation of the cooling loop

An alternative methodology to a block-diagram base model is a state-based model, wherein the emphasis is placed on the comprehensive state of the system, and transitions between states are dictated by predetermined probabilities. One such methodology is the Markov Chain model. A Markov chain model has been formulated and is depicted in fig. 6. In this framework, state 0 signifies the fully operational state of the system. States 1 and 2 pertain to conditions in which the system's functionality is compromised, such as a defective cold plate or a failure of the pump. The transition from each state is linked to a failure rate denoted as λ and/or a repair rate indicated as μ. For instance, λ_1 denotes the failure rate (transition probability) from state 0 to state 1. For the sake of simplicity, it is assumed that the transition probability is independent of temporal factors.

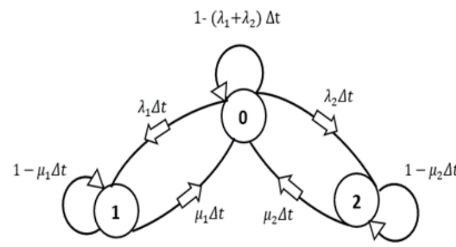

Fig. 6: Markov chain showing 3 states

It is also important that we discuss a little about failure mode modeling. Failure mode modeling constitutes a rigorous methodology used to identify, scrutinize, and improve prospective failure modes across diverse systems and processes. By using empirical evidence and statistical techniques, failure modeling helps predict the occurrences and timing of failures, thereby enabling organizations to proactively mitigate vulnerabilities prior to their escalation into critical issues. Through the amalgamation of historical failure data, and comprehensive analytical frameworks, failure mode modeling not only reinforces the reliability of systems but also enhances safety, performance, and operational life. Adding failure modeling on top of a system reliability model gives the added benefit of improving the overall reliability of the system. In Section III C, we discuss some modeling techniques for failures pertaining to a jet impinging cold plate.

III. Methodology

A. *Cooling System Architecture*

The principal aim of a cooling system is to effectively transfer the thermal energy generated at the cold plate to the external environment. Furthermore, this procedure must be conducted with a significant level of efficiency. A variety of methodologies exist for accomplishing this goal, and we are utilizing a specific configuration in which the thermal energy from the secondary loop is exchanged with the primary loop through a heat exchanger. The design is intended to emulate the research conducted by [8]. The ambient temperature is set from ARPA-E specifications and we back calculate the total energy utilization. Figure 5 presents the block diagram that illustrates our cooling system framework.

From the definitions of power density and total cooling efficiency from ARPA-E [7], we need to select the optimum number of servers and minimum TDP per servers.

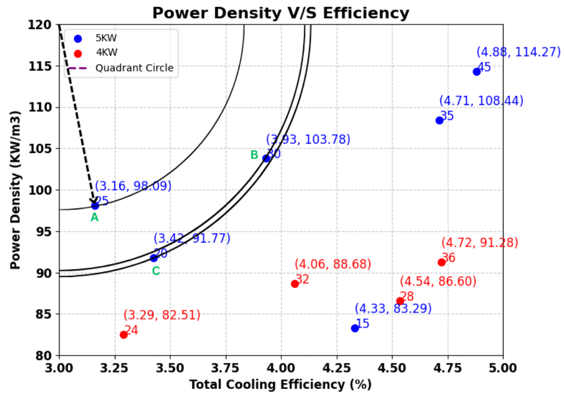

Fig. 7: Plot showing the points obtained using Euclidean distance, optimal points are marked as A,B and C

The Euclidean distance serves as a quantitative assessment of the linear separation between two distinct points within a Euclidean framework. This metric is extensively used in optimization scenarios where the primary objective involves the minimization or maximization of spatial separations. The analytical formulation of the minimization problem is represented as 3. In the illustration provided in fig.7, we establish a reference point characterized by an efficiency of 3% and a power density of 120 (Kw/m3). Our objective is to enhance the problem by reducing the distance to the reference point in pursuit of the most optimal solution. This methodology facilitates a thorough examination of prospective solutions, thereby guaranteeing that the chosen alternative is in closest proximity to the stipulated efficiency and power density criteria. Through this approach, we can proficiently obtain the most feasible options that correspond with our efficiency and power density objectives, ultimately providing enhanced performance in real-world applications. The best three options are tabulated in Table II.

Minimize the function:

$$f(x, y) = (x - x_0)^2 + (y - y_0)^2 \tag{3}$$

TABLE II: Power Density and Efficiency Data

Point	Power Density (KW/m^3)	Efficiency (%)	Server Power (KW)	Rack Partial Power (KW)
A	98.09	3.16	5	25
B	103.78	3.93	5	30
C	91.77	3.42	5	20

B. System Reliability Models

A reliability model for a liquid cooling system within data centers is designed to evaluate and forecast the operational efficacy and reliability of the cooling infrastructure throughout its service life. The model encompasses a range of failure modalities and elements that influence system performance, such as pump malfunctions, leaks in piping or fittings, material degradation due to erosion or corrosion, and blockages resulting from contaminants. By conceptualizing these components as part of a series-parallel network, one can ascertain the aggregate reliability of the system. This model is essential for pinpointing critical failure junctures and refining maintenance schedules. Furthermore, it facilitates the assessment of redundancy measures, including the incorporation of auxiliary pumps or alternative pathways for fluid flow, to guarantee uninterrupted cooling. Reliability models also incorporate environmental parameters such as temperature, pressure, and coolant flow rates, ensuring that the system can satisfy the elevated performance requirements of contemporary data centers while concurrently minimizing both downtime and energy consumption. Notwithstanding its advantages, the validation of these models presents significant challenges due to a deficiency in experimental data.

1) 3-State Lumped Model: A straightforward approach to reliability modeling combines block diagram methods with state-based models. This approach is advantageous due to its simplicity and quick implementation. As illustrated in Fig.6, the system is represented by three distinct states. The methodology involves consolidating multiple components into a single state for simplicity. One state represents a fully operational system, while another state accounts for a novel component with limited data availability. The analytical solution provided by [9] is employed to evaluate the system availability.

2) Multi-state Model: In the 3 state model we lump the components in to a single state. Even though it is time efficient process, it might affect the accuracy of the results. In the multi-state approach, we keep all the components as it is for the analysis. The number of components in the system can be higher which makes this model complicated. Analytically solving for a solution for this model is practically impossible and we use an iterative process to solve the problem.

3) Monti-Carlo Markov Chain (MCMC): The most advanced approach to reliability modeling incorporates randomness into the problem by employing Monte Carlo simulations alongside a system's Markov chain model. In this approach, a predefined number of simulations are run, with each simulation's outcome determined by the prior probability distribution of event occurrences. These probability distributions model the likelihood of transitions between states in the Markov chain. A simple discrete sampling algorithm is employed to determine state transitions in a Markov Chain based transition matrix. The results from the simulations are aggregated to evaluate the system's reliability, providing a robust estimation by accounting for the inherent randomness and variability in system behavior [10].

- Monte Carlo simulations involve sampling from probability distributions to model stochastic behavior.
- The Markov chain defines state transitions and their probabilities.
- Aggregated simulation results provide a probabilistic measure of reliability.

C. Cold plate Failure Modes Modeling

A cooling system consists of a heat collection system (Coldplate), heat carrying system (coolant distribution system) and heat rejection system. Of the three, we focus on the heat collection system, that is, the cold-plate. The cold plate is a novel component, it uses jet impingement combined with vapor venting through micro-channels to effectively collect heat.

Failure model effect analysis (FMEA) is a method by which we can identify potential failures that can affect the performance and reliability of the component. Table III shows the different failures the component can have and the respective Risk Priority number (RPN) is obtained from the assigned criticality, occurrence, and detectability. The values are assigned based on discussions with the respective team members, and a certain level of intuition is involved. The uppermost value attributed to each matrix is designated as six. In terms of criticality, this signifies that the occurrence of this failure can significantly hinder operation; regarding occurrence, a value of six denotes a high probability of occurrence, and with respect to detection, a score of six indicates considerable difficulty in identifying the failure.

TABLE III: FMEA

Sl.no	Failure Mode	Criticality	Occurrence	Detection	RPN
1	Delamination	6	3	3	54
2	Cracks in silicon	6	2	5	60
3	Loose powered fouling	6	5	2	60
4	Erosion	6	3	3	54
5	Corrosion	5	2	6	60
6	Leakage	5	2	2	20
7	Non uniform injection	1	4	4	16
8	Crack in Manifold	1	2	2	4

Crack in silicon and delamination are more likely to occur due to manufacturing process and or during assembly. The principal mechanisms of failure that necessitate our attention encompass erosion, corrosion, and clogging (fouling). . We will examine various models of failure modes that have been formulated by researchers in other studies to assess the implications of these failure mechanisms. This comprehensive analysis will provide insights into the applicability and reliability of these models in predicting failure outcomes, ultimately guiding improvements in system design and maintenance strategies.Through this exploration, we aim to establish a framework that not only highlights the critical factors contributing to system failures but also offers practical recommendations for mitigating these issues in real-world applications.

1) *Erosion:* Erosion modeling involves predicting material loss or surface degradation caused by mechanical interactions such as particle impact, fluid flow, or a combination of both. Using the models, our aim is to quantify the erosion rate based on parameters such as impact velocity, particle size, material hardness, and impact angle.

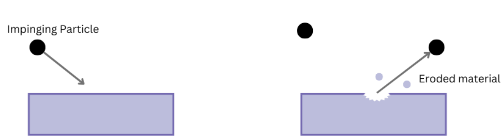

Fig. 8: Schematic of erosion of surface due to particle impingement

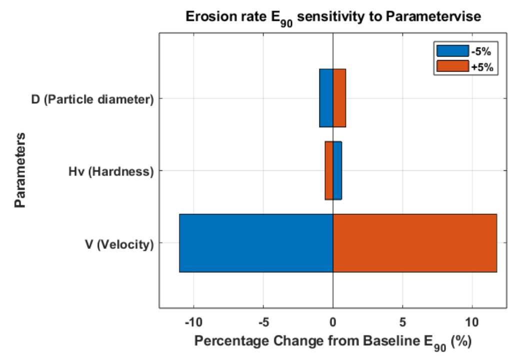

Fig. 9: Sensitivity analysis on erosion model

Oka's model [11] and [12]is a widely used empirical approach for predicting erosion rates due to solid particle impact. It incorporates key factors such as impact velocity (v), particle size (D), impact angle (α), and material properties like hardness (Hv). The model employs material-specific constants to account for different erosion mechanisms, such as cutting and deformation, which vary with the impact angle. One of its strengths is its ability to predict erosion behavior across a range of conditions, making it a versatile tool for engineering applications. Oka's model provides valuable insights into optimizing material selection and design in erosive environments. The erosion rate as a function of α is given by:

$$E(\alpha) = g(\alpha)E_{90} \tag{4}$$

where:

$$g(\alpha) = (\sin\alpha)^{n_1}\left(1 + Hv(1 - \sin\alpha)\right)^{n_2} \tag{5}$$

and:

$$E_{90} = K(Hv)^{k_1}(v)^{k_2}(D)^{k_3} \tag{6}$$

where k1 , k2 and k3 are exponent factors, which are affected by other parameters, respectively. K denotes the particle property factor that is influenced by particle shape (angularity) and hardness, which has no correlation among different types of particles and other factors. Figure 9 is a tornado plot to signify the sensitivity of the erosion rate to the models parameters. The erosion rate is more sensitive to velocity of impingement compared to the particle size and hardness of the surface. This implies that a reduction of jet velocity can significantly reduce the erosion there by improving the life of the cold-plate.

2) *Corrosion:* Corrosion, an electrochemical process, results in the degradation of materials due to environmental interactions. In liquid-cooled cold plates, corrosion presents a significant challenge to maintaining long-term reliability and performance. The interaction between the coolant, metal surfaces, and contaminants can cause material degradation, leading to reduced thermal efficiency and potential system failures, such as leaks. Utilizing Arrhenius-type models allows for the quantification of corrosion rates based on temperature and chemical conditions. These models provide valuable insights by linking operating parameters, such as temperature and coolant composition, to the cold plate's lifespan, facilitating predictive maintenance and system optimization. Equation 7

is the mathematical representation of the model where, E_a is the activation energy and T is the surrounding temperature in kelvin.'A' is the model constant, which is specific to the material and or operational parameters. A sensitivity analysis done on the model 7 is shown in fig. 10. The corrosion rate is very sensitive to temperature and activation energy. A low activation energy implies a faster reaction and hence a higher corrosion rate.

$$R = A \cdot e^{-\frac{E_a}{RT}} \tag{7}$$

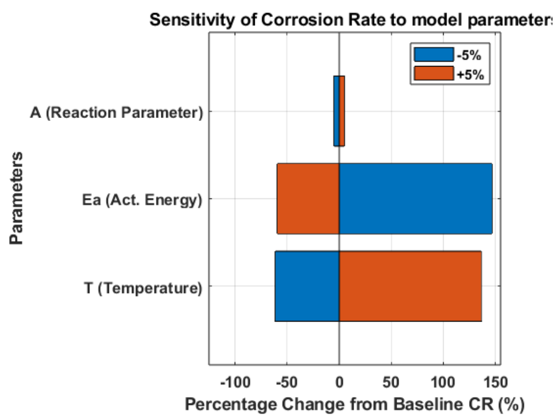

Fig. 10: Sensitivity analysis for corrosion model

3) Clogging: Due to the utilization of the laser sintering technique in the fabrication of the cold plate, a minor proportion of copper particulates may be incorporated within the refrigerant circuit. The mechanism of clogging primarily hinges on the geometric properties of both the particles and the channel. The research conducted by [13] establishes the likelihood of particle obstruction as a function of temporal progression. For sieve clogging, particle size larger than the channel dimension, the mean time to clog (t^*) depends on the flow rate (Q) and the particle concentration (c) and its relation is given by 8. It is to be noted that the concentration of large particle is of consideration. For agglomerate clog, a certain number of particle should flow through the channel for it to clog and this average number of particle is termed as critical particle number (CPN). The CPN for a channel with height H, width W and a particle size D is obtained using the equation 9. The figure 12 presents a sensitivity analysis pertaining to the CPN, in which the physical parameters of the channel—including width, and height- and the particle diameter are analyzed as variables. The analysis reveals that the CPN exhibits a greater sensitivity to the particle diameter in contrast to the geometric characteristics of the channel.

$$t^* \propto \frac{1}{Qc} \tag{8}$$

and

$$N^* = \frac{W^3 H}{6a^2 D^4} \tag{9}$$

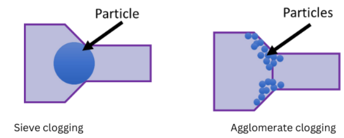

Fig. 11: Schematic of clog formation in a channel

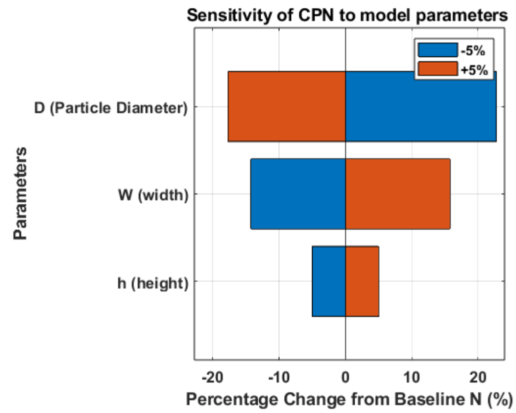

Fig. 12: Sensitivity analysis for clogging model

IV. Results

A. Reliability Model

Three different models have been developed to evaluate the overall system reliability. Table IV shows the cooling system reliability evaluated by each model when the failure rate of the novel component is set at thousand failed components in a million. There is no measure of which model is the best to use. The computational time varies with each model. The MATLAB R2024a codes for the models were run using a computer with Intel i7 8700K, 32GB memory. The computational time for each model are also shown in table IV. MCMC takes the highest computational time because of the large sample size required. Also because of the random variable, the results obtained will be in a range.

TABLE IV: System Availability obtained from different models with fixed failure rate for novel component

Model	**3 State Model**	**Multi State**	**MCMC**
Availability (%)	99.983	99.9824	99.9816 - 99.9837
Computational time (ms)	12	150	65800

B. Failure Modes Modeling

1) Erosion Modeling: Fig.14 and fig.13 are parametric plots of the erosion rates for normal impingement and at a parallel impingement.The erosion rate exponentially increasing with the velocity. This is also seen in the sensitivity analysis 9 done on the model parameters. The model has a lot of model constants which are dependent on the particles, and on the surface. Since we cannot exactly quantify the results without experimental data, hypothetically if the erosion rate is high, a solution to this problem lies in altering the variables under our control. This would be the flow velocity and then the particle size. Controlling the particle size requires a finer filter which can affect the pumping power. Apart from that, comparing the erosion rate for same flow rate and particle size, normal impingement has a higher erosion rate as compared to the later.

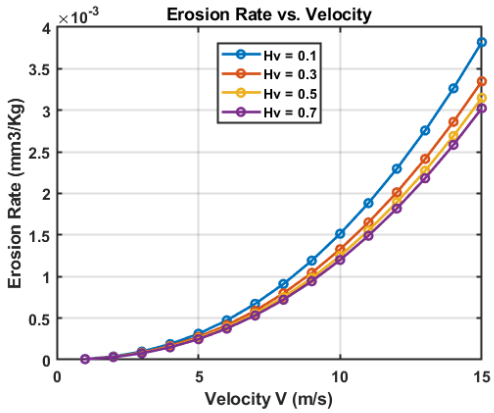

Fig. 13: Parametric plots on erosion rate 90 degree impingement

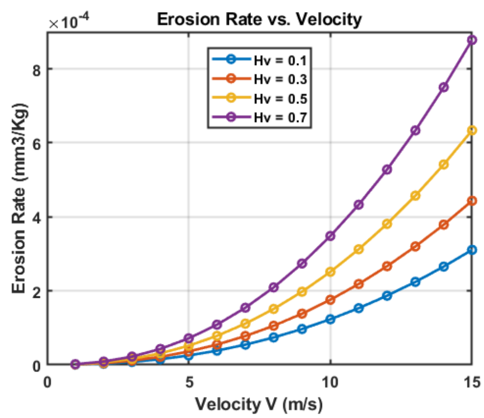

Fig. 14: Parametric plots on erosion rate at Zero impingement

2) Corrosion Modeling: The rate at which corrosion occurs is significantly influenced by the activation energy as well as the prevailing temperature conditions within the environment. The level of activation energy required for a specific electrochemical reaction to transpire is, in fact, contingent upon the degree of corrosiveness presented by the surrounding environment. In scenarios where the environment is characterized by inert properties, one can deduce that the activation energy will manifest as relatively high, which consequently indicates that the rates at which corrosion occurs will be markedly low. This phenomenon can be interpreted as meaning that the electrochemical reaction necessitates a greater amount of energy in order to advance through its various stages and ultimately reach completion. Thus, it can be argued that environmental conditions play a crucial role in determining the efficiency and speed of corrosion processes, which are fundamental to understanding material degradation. Overall, the interplay between activation energy and environmental factors is essential for predicting corrosion behavior in various industrial applications and contexts.

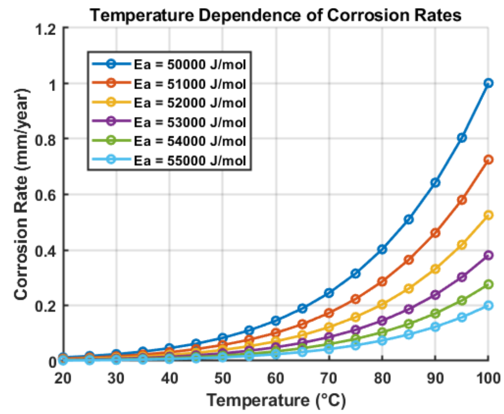

Fig. 15: Parametric plots on corrosion rate

3) Clogging Modeling: The clogging model differes in the application for getting a clog. Fig.8 is a plot on mean clogging time and the concentration of particles in the fluid. Revering to fig.9, depicts the variation of CPN with particle diameter. The clogging is significantly affected by the concentration of particles and is affected by the size of the particle. These two data can provide the filtration level we need to separate out the particle sizes so that the channels have higher mean time to clog and CPN. This essential meas higher operational life of the cold-plate.

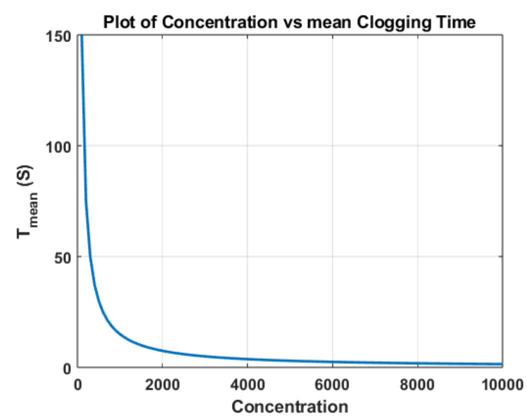

Fig. 16: Sieve clogging time with increase in concentration

Conclusion

The availability of cooling systems is an essential determinant for data centers, and the formulation of a dependable system necessitates meticulous planning and execution. A reliability model has been constructed for this objective, and the evaluation of system reliability is conducted. Through the application of Failure Mode and Effects Analysis (FMEA), it is possible to pinpoint critical failures associated with the cold plates. We undertake a thorough examination of the

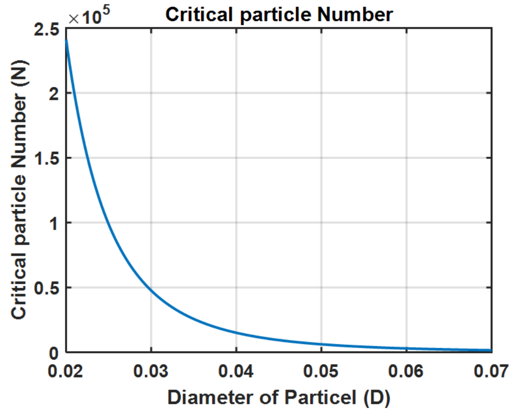

Fig. 17: Particle size on Aggregate clogging

failure modes and engage in a discussion of various empirical models designed to quantify these failure modes. A significant limitation of the models lies in the validation phase. There exists a paucity of data, if any, to substantiate the findings and to calibrate the models for our particular scenario.

Acknowledgment

Huge gratitude to Professors Scott Schiffers, Justin Weibel, and also Ryan Enright for their invaluable expertise and dedication, driving our project's success forward. Sincere thanks for your unwavering support and collaboration. Furthermore, we extend heartfelt thanks to all members of our research team for their tireless efforts and significant contributions. Your collaboration and dedication have played a vital role in advancing our project.

References

[1] S. News, *Meta, amazon, microsoft: Massive ai capex spending in quarterly earnings*, https://sherwood.news/tech/meta-amazon-microsoft-massive-ai-capex-spending-quarterly-earnings/, Accessed: 2024-12-12, 2024.

[2] K. M. U. Ahmed, M. H. J. Bollen, and M. Alvarez, "A review of data centers energy consumption and reliability modeling," *IEEE Access*, vol. 9, pp. 152 536–152 563, 2021, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3125092. [Online]. Available: https://ieeexplore.ieee.org/document/9599719/?arnumber=9599719 (visited on 10/26/2024).

[3] G. Callou, P. Maciel, D. Tutsch, and J. Araújo, "Models for dependability and sustainability analysis of data center cooling architectures," in *IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN 2012)*, ISSN: 2325-6664, Jun. 2012, pp. 1–6. DOI: 10.1109/DSNW.2012.6264697. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6264697 (visited on 12/12/2024).

[4] R. Khalid and A. P. Wemhoff, "Thermal control strategies for reliable and energy-efficient data centers," *Journal of Electronic Packaging*, vol. 141, no. 4, pp. 041 004–, 2019. DOI: 10.1115/1.4044129.

[5] J. Wang, Q. Zhang, S. Yoon, and Y. Yu, "Reliability and availability analysis of a hybrid cooling system with water-side economizer in data center," *Building and Environment*, vol. 148, pp. 405–416, 2019, ISSN: 0360-1323. DOI: https://doi.org/10.1016/j.buildenv.2018.11.021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360132318307169.

[6] J. Cho, B. Park, and Y. Jeong, "Thermal performance evaluation of a data center cooling system under fault conditions," *Energies*, vol. 12, no. 15, p. 2996, 2019. DOI: 10.3390/EN12152996.

[7] ARPA-E. "Funding opportunity announcement: Cooling system innovation." Accessed: 2024-12-14. (2024), [Online]. Available: https://arpa-e-foa.energy.gov/FileContent.aspx?FileID=59084ec2-841e-4e3d-a40e-1f95e42b1f33.

[8] S. Khalili, S. Rangarajan, V. Gektin, H. Alissa, and B. Sammakia, "An experimental investigation on the fluid distribution in a two-phase cooled rack under steady and transient information technology loads," *Journal of Electronic Packaging*, vol. 142, no. 41002, Sep. 4, 2020, ISSN: 1043-7398. DOI: 10.1115/1.4048180. [Online]. Available: https://doi.org/10.1115/1.4048180 (visited on 12/11/2024).

[9] "Three-state markov chain based reliability analysis of complex traction power supply systems | IEEE conference publication | IEEE xplore." (), [Online]. Available: https://ieeexplore.ieee.org/document/9660623 (visited on 12/09/2024).

[10] D. A. Spade, "Chapter 1 - markov chain monte carlo methods: Theory and practice," in *Principles and Methods for Data Science*, ser. Handbook of Statistics, A. S. R. S. Rao and C. R. Rao, Eds., vol. 43, ISSN: 0169-7161, Elsevier, 2020, pp. 1–66. DOI: https://doi.org/10.1016/bs.host.2019.06.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169716119300379.

[11] Y. I. Oka, K. Okamura, and T. Yoshida, "Practical estimation of erosion damage caused by solid particle impact: Part 1: Effects of impact parameters on a predictive equation," *Wear*, 15th International Conference on Wear of Materials, vol. 259, no. 1, pp. 95–101, Jul. 1, 2005, ISSN: 0043-1648. DOI: 10.1016/j.wear.2005.01.039. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0043164805000979 (visited on 08/29/2024).

[12] Y. I. Oka and T. Yoshida, "Practical estimation of erosion damage caused by solid particle impact: Part 2: Mechanical properties of materials directly associated with erosion damage," *Wear*, 15th International Conference on Wear of Materials, vol. 259, no. 1, pp. 102–109, Jul. 1, 2005, ISSN: 0043-1648. DOI: 10.1016/j.wear.2005.01.040. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S004316480500102X (visited on 08/29/2024).

[13] H. M. Wyss, D. L. Blair, J. F. Morris, H. A. Stone, and D. A. Weitz, "Mechanism for clogging of microchannels," *Physical Review E*, vol. 74, no. 6, p. 061 402,

Dec. 11, 2006, Publisher: American Physical Society. DOI: 10.1103/PhysRevE.74.061402. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.74.061402 (visited on 11/14/2024).