RADiCe: A Risk Analysis Framework for DataCenters

Fabian Mastenbroek[a], Tiziano De Matteis[a], Vincent van Beek[b], Alexandru Iosup[a]

[a] *Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*
[b] *Solvinity B.V., Amsterdam, The Netherlands*

Abstract

Datacenter service providers face engineering and operational challenges involving numerous risk aspects. Bad decisions can result in financial penalties, competitive disadvantage, and unsustainable environmental impact. Risk management is an integral aspect of the design and operation of modern datacenters, but frameworks that allow users to consider various risk trade-offs conveniently are missing. We propose RADiCe, an open-source framework that enables data-driven analysis of IT-related operational risks in sustainable datacenters. RADiCe uses monitoring and environmental data and, via discrete event simulation, assists datacenter experts through systematic evaluation of risk scenarios, visualization, and optimization of risks. Our analyses highlight the increasing risk datacenter operators face due to price surges in electricity and sustainability and demonstrate how RADiCe can evaluate and control such risks by optimizing the topology and operational settings of the datacenter. Eventually, RADiCe can evaluate risk scenarios by a factor 70x–330x faster than others, opening possibilities for interactive risk exploration.

1. Introduction

Datacenters have become essential to support the digitalization of our economy and society [29, 28, 26]. The constantly increasing demand for computing power has let organizations move from in-house to hyper-scale (cloud) datacenters able to serve stakeholders across industry, government, and academia. These stakeholders have come to expect reliable operation and high quality of service, yet demand low cost, high scalability, and corporate responsibility. As a result, datacenter operators and architects are confronted with significant research and operational challenges and are accounted responsible when customers' expectations are not met.

Despite technological advancements and better availability management, failures in datacenters remain a major concern for many operators. The financial consequences of outages can be severe: in 2022, 47% of outages cost between $100,000 and $1 million, and 15% cost over $1 million in terms of violated Service Level Agreements (SLAs) [7]. Furthermore, the recent energy price crisis and efforts to sustainability (e.g., through green bonds emissions) have resulted in operational and societal expenses becoming a primary cost factor for datacenters [12].

Risk analysis is integral to the design and operation of datacenter infrastructure. It enables organizations to set out future objectives, recognize potential risks that they could face, and adequately control them. For example, risk management arises in the design of datacenters, where the *procurement* (i.e., long-term capacity planning) of cloud infrastructure is a critical optimization problem faced by datacenter architects. Simply over-provisioning capacity is expensive. Conversely, under-provisioning could lead to risks of failing to meet SLAs [4].

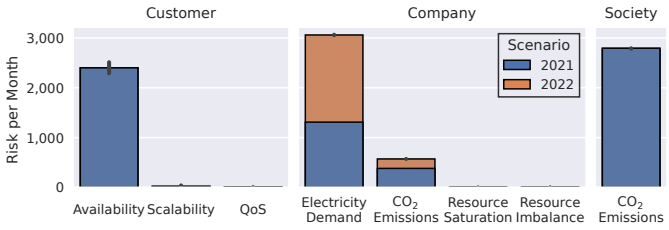

Figure 1: Risk profiles of a private datacenter in 2021 and 2022. Different aspects concern the stakeholders—customers, provider (company), and society.

Risk trade-offs also appear when operating a datacenter. Efforts to improve the energy efficiency of datacenters have reduced the average Power Usage Effectiveness (PUE) [1] (the ratio of the total facility energy to IT energy) significantly since 2007. However, the decline has stagnated over the past few years [20]. Notwithstanding environmental concerns, with a large percentage of the total costs of datacenters going to electricity, consuming too much of it leaves datacenters vulnerable to price spikes: a serious risk, given the sharp energy price increase that occurred in 2022. It remains possible that also organizations with good (low) PUE become highly impacted by energy hikes under specific workloads and resource management policies. Finally, the lifecycle of datacenter infrastructure presents additional risk trade-offs. Through hardware refreshes, datacenter operators can steadily increase the compute capacity while reducing energy consumption, as a result of the increasing energy efficiency of servers. However, replacing equipment too early may cost more in hardware than is saved on energy efficiency [10]. If these trade-offs are fully understood, their management and control could lead to significant reductions in costs, energy usage, and carbon footprint of datacenters.

Despite the importance of risk assessment, relatively few

*Corresponding author
Email address: t.de.matteis@vu.nl (Tiziano De Matteis)

comprehensive approaches and tools exist, leaving many datacenter operators ill-equipped to confidently make informed decisions. For a long time, the community has focused mainly on availability or Quality of Service (QoS) as an indicator of risk [23, 25, 21, 33, 54, 41]. We believe that **risk management in datacenter should consider other business goals, such as sustainability and societal impact**. Figure 1 shows the impact of various risk factors that affect not only the customers but also the datacenter operators and society. In 2021, the primary factors contributing to the risk profile of the datacenter are customer-related factors, such as availability. However, due to the energy price increase in 2022, the electricity expense become the primary risk factor for datacenters. While this may even get worse in the future (e.g., datacenter may be considered accountable for their environmental impact [20]), it is clear that **datacenter providers need an instrument for facilitating risk analysis of datacenters**. Such an instrument must enable automatic analysis to allow continuous risk evaluation and consider the complex interplay of hardware, software, and various risk factors. Finally, we argue that **more quantitative studies, methods, and tools are needed for the society-at-large to gain confidence in datacenter sustainability and impacts**. Otherwise, important decisions may go against stakeholders. In countries like the Netherlands, hyperscale datacenter projects have been stopped based on vague, qualitative statements against the potential impact on the electricity grid or climate[1]

These are fundamental problems for IT-related risks in sustainable cloud infrastructures. To address them, ***we propose* RADiCe*, an open-source framework for quantitative data-driven risk analysis in sustainable datacenters***. Underpinning this system is a trace-based, discrete-event simulator that enables the exploration of different risk scenarios through support for diverse workloads, datacenter topologies, and operational phenomena. The use of discrete-event simulation enables complex and long-term analysis and thus allows prompt answering of critical questions related to risks and sustainability without sacrificing accuracy. Although RADiCe is designed to work across many kinds of datacenters, in this work, we focus on private-cloud, business-critical workloads, and public-cloud operations, representing the majority of workloads in modern datacenters. Our contributions are:

1. We propose RADiCe (Risk Analysis for DataCenters), an open-source risk analysis framework for cloud datacenters (in Section 3). RADiCe introduces a *holistic* approach to risk analysis, where multiple risk factors and scenarios can be considered and evaluated at the same time, obtaining credible and quantitative evidence of risk trade-offs. It can also automatically explore possible topology and operational settings optimizations that reduce risk, and suggest appropriate changes. Compared with pioneering tools that also address risk, such as Capelin [6], RADiCe significantly reduces the effort necessary for the users to identify, evaluate, and optimize applicable scenarios, enables detailed *quantitative* analysis of operational risks for both managerial decisions and detailed policy enforcement, and specifically includes detailed risk components facing cloud customers, cloud operators, and society at large.
2. We evaluate RADiCe over multiple real-world workloads and consider different risk scenarios (in Section 4). Our experiments show many interesting findings, supporting our claim for a need for data-driven risk analysis in datacenters. Importantly, RADiCe enables system-wide, holistic findings, including various sustainability aspects, e.g., CO2 emissions and related costs for companies and society.
3. We release RADiCe as free and open-source software[2], hoping to contribute with a tool, and good practices, for reusable, comprehensive, and inspectable risk management in sustainable datacenters.

2. Background

2.1. Risk Management

Risk management is the systematic process of identifying, analyzing, and controlling risks. Risk, in this context, is defined as the "effect of uncertainty on objectives" in ISO 31000 [2] and can be positive, negative, or even both. Risk originates from various sources, such as legal liabilities, financial uncertainty, threats to IT infrastructure, accidents, and climate change.

Risk management is a multi-step, often iterative, procedure that usually involves the *identification* of risk factors, their *analysis*, and successive *response* [2]. Multiple risk analysis techniques exist, and choosing the appropriate technique often depends on the circumstances and intended use. *Qualitative risk analysis* uses a subjective assessment of the risk probability and impact, relying on the knowledge and interpretation of the assessor. On the other hand, *quantitative risk analysis* [57] is a systematic approach that quantifies both the likelihood and impact of risks numerically, relying on accurate, measurable data to produce insights. *Semi-quantitative risk analysis* uses a combination of both qualitative and quantitative methods to analyze risk. In this work, we use a quantitative approach for risk analysis.

2.1.1. Service Level Agreement and Objective

A *Service Level Agreement* (SLA) is an explicit or implicit contract between the cloud provider and customer that governs the obligations and responsibilities between both parties regarding the provided service. An SLA establishes (i) what kind of service is to be provided and how, (ii) constraints for the level of service (e.g., for availability or performance), (iii) the costs to be paid by the customer to the cloud provider, and (iv) the consequences for not upholding the agreement (usually a financial penalty).

[1] `https://www.datacenterknowledge.com/meta-facebook/scorned-meta-data-center-holland-met-all-environmental-standards`.

[2] A preliminary version is available at `https://github.com/atlarge-research/opendc/releases/tag/project%2Fradice`.

Generally, the SLA employed by the industry includes a set of *Service Level Objectives* (SLOs) that define the expected service between the cloud provider and the customer. Each of these SLO unambiguously expresses the service level guaranteed by the cloud provider for some measurable characteristic, for example, as a "monthly uptime percentage" of at least 99.995% for a Virtual Machine (VM).

2.1.2. Risk Scenarios

The existing approaches for risk analysis are seldom tested with real-world scenarios or operate on simplistic models considering a single workload or topology [4, 60]. We advocate comprehensive experiments using real-world operational traces and diverse risk scenarios to evaluate risk analysis approaches thoroughly. For this reason, we build on the concept of *portfolios of scenarios* introduced in Capelin [5]. A *scenario* represents a point in the datacenter design space to explore. It consists of a combination of workload, topology, and *operational phenomena*. Phenomena can include correlated failures, workload interference, security breaches, etc., allowing the scenarios to more accurately capture real-world operations.

Each portfolio includes a base scenario, a set of candidate scenarios given by the user and/or suggested by the system to enable *multiple* scenario analysis, and a set of targets that prescribe on what grounds and on which time scale the different scenarios should be compared. Targets include the metrics the practitioner is interested in, their desired granularity, and relevant SLOs [47].

Following the taxonomy defined by the performance organization SPEC [33] and illustrated by Figure 2, tools such as Capelin support traditional performance and system-provider metrics. In the Policy Metrics layer, Capelin offers merely proxies for SLO violations [5, § 5.1.7]; in the layer Metrics for Managerial Decisions, Capelin does not support specific metrics, and relies instead on ad-hoc, qualitative warnings (e.g., "high risk" [5, MF2 and MF6]). Addressing an important gap, RADICE proposes an abstraction that also supports the highest two layers in the SPEC taxonomy, adding functionalities for holistic, detailed, and quantitative risk analysis of datacenter risks, focusing on specific types of risk for customers, operators (hosting companies), and societal stakeholders. The higher-level metrics introduced by RADICE facilitate long-term decision-making by abstracting away from specific workloads and infrastructure, enabling important use-cases (e.g., Section 3.1) and complex scenarios (e.g., as Section 4 exemplifies).

2.2. System Model

In this work, we use a generic model for datacenter operation, which is already widely used in academia and industry [48, 56, 6]. Figure 3 shows the model comprising three main components: *workloads*, *resources*, and *management and scheduling*.

Workload

We consider three types of workloads. The first one consists of applications executing in *VMs*, *containers*, or directly on *physical machines*. Then, we consider *app managers*, such as big data frameworks (e.g., Apache Spark or Apache Flink), machine learning frameworks (e.g., TensorFlow or PyTorch), and serverless frameworks (e.g., OpenFaaS), which orchestrate virtualized workflows and dataflow for their users. Finally, we also take into account scientific workloads deployed on virtualized environments. These workloads primarily comprise conveniently (embarrassingly) parallel tasks—e.g., Monte Carlo simulations—forming *batch bags-of-tasks*.

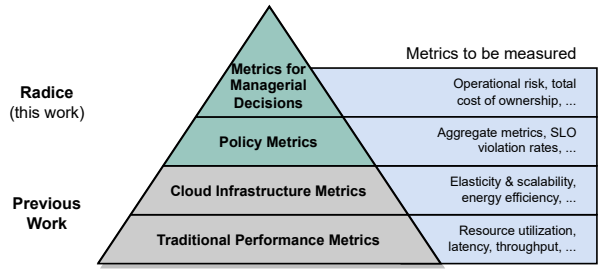

Figure 2: A hierarchical taxonomy of datacenter metrics, [33].

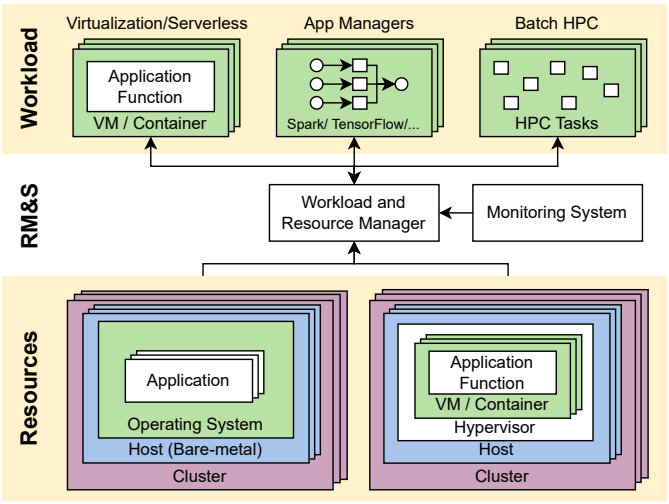

Figure 3: Generic model for datacenter operation.

Resources

Workloads run on physical datacenter infrastructure. We model datacenter infrastructure as a set of physical clusters of possibly *heterogeneous hosts*, each host being a node in a datacenter rack. A host can execute multiple VM- or container-workloads managed by a *hypervisor*. In this work, we model resource consumption of applications (e.g., CPU usage) per discretized time slices. Workloads report at each time slice their resource consumption to the hypervisor, which consolidates the requests and distributes the resources based on some scheduling policy.

Resource Management and Scheduling

A resource manager manages and controls all clusters and hosts and is responsible for the lifecycle of submitted workloads, including their placement onto the available resources [6].

The resource manager is configurable and supports various *policies* to distribute workloads over the available resources.

2.3. Datacenter Simulation

RADICE uses discrete-time simulation to explore and experiment with various scenarios in a timely manner.

The motivation for resorting to simulation is twofold. First, experiments on deployed systems are difficult to adapt or reconfigure. As we want to explore several scenarios, doing this in a real system would require procurement and installation of new resources, which could be expensive. Using simulation instead, we can quickly evaluate alternative scenarios with few costs. Second, experiments on physical infrastructure are time-consuming and expensive, notwithstanding the environmental impact of such experiments. This impact can become unacceptable for even moderately sized infrastructure. RADICE uses the OpenDC platform, a datacenter discrete-event simulator, validated against analytical models and other leading simulators [43]. RADICE leverages its existing feature set that includes: (i) serverless and application manager operation model; (ii) possibility to model a datacenter using a graphical interface; (iii) experiment automation and a comprehensive set of operational telemetry metrics; (iv) the possibility of configuring the resource manager, and experimenting with different scheduling policies. While in this work we build on top of OpenDC, RADICE can also be implemented on other cloud or datacenter simulators, provided that they support the same set of features.

3. The RADICE Approach

RADICE enables a holistic approach to risk *analysis* and *optimization*. In this section, we describe users and use cases of RADICE, detail how risks are modeled, and provide an overview of its architecture and internal components.

3.1. Users and Use Cases

RADICE targets *datacenter operators* that manage the daily operation of the datacenter infrastructure of the cloud provider, and *datacenter architects*, that design the datacenter infrastructure of the service provider. Both roles have different needs and responsibilities. We identify the use cases where RADICE could be useful:

- *risk monitoring*: datacenter operators can utilize the system to monitor the current risk of the datacenter infrastructure in near real-time and make informed decisions on how risk is managed.
- *capacity planning*: RADICE supports datacenter architects during the capacity planning processes for cloud infrastructure, where smart decisions could lead to significant service improvements, cost savings, and environmental sustainability [9].
- *change management*: datacenter operators can utilize the system to evaluate how changes to the datacenter infrastructure (e.g., deploying new workloads) affect the risks they face, such as increased probability of failure events or degraded performance.
- *compliance and auditability*: the system can aid datacenter managers in showing their efforts toward reducing operational risk to (potential) customers and in demonstrating compliance with standards, contracts, and regulations.

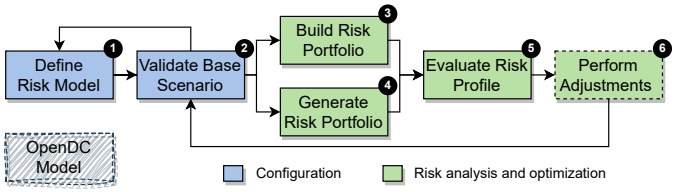

Figure 4: The process for analyzing and optimizing risk in datacenter infrastructure using RADICE.

3.2. Model for Quantifying Risks in Cloud Datacenters

Risks are represented in RADICE as a portfolio of *risk factors*, each modeling an aspect of the operational risk of a datacenter, quantifying it in a specific dimension. These risk factors comprise several SLOs (predicates over metrics), an aggregation time period (e.g., monthly or daily), and an impact function. During every aggregation period, the active SLOs are evaluated, and if in violation, the impact is computed using the impact function. Combined, the impact of all SLO violations sum to a single value, representing the *total risk*.

We distinguish between three categories of risk: customer-facing risks, company-facing risks, and society-facing risks. *Customer-facing risks* are risk factors that directly affect the customer, and include resource availability, QoS, or security. Risk factors that affect the company's sustainability, but do not directly affect customers, are categorized as *company-facing risks*. Operational efficiency is an example of such a risk. Finally, *society-facing risks*, such as environmental sustainability, affect society as a whole but do not directly affect the company.

RADICE considers the financial impact of risks. By expressing risk in terms of monetary value, we ensure a fair comparison between different scenarios since our model can consider the cost of other risk responses, such as risk mitigation and elimination. For example, this might entail the costs of upgrading the datacenter topology or purchasing insurance. Moreover, this approach enables users at different layers of the organization, from technicians to managers, to grasp the impact of both short- and long-term decisions in datacenters.

The strategy for specifying the impact function mostly depends on the type of risk. The impact of customer-facing risks can often be derived from the penalties for SLA violations, as stipulated by the contracts between the company and the customer. Company-facing risks can be quantified based on the operational expenses of the organization. These include electricity usage expenses and increased engineering costs due to problem troubleshooting. An impact function for society-facing risks is often more difficult to define and requires estimations from scientific literature, if available. We further elaborate on this in Section 4.2.

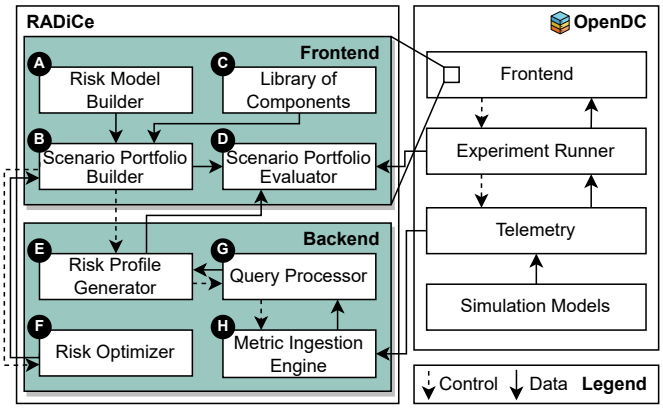

Figure 5: An overview of the high-level architecture of RADiCe.

3.3. Risk Analysis and Optimization with RADiCe

In this section, we present a systematic process for approaching risks in datacenters using RADiCe, for which we show a summary in Figure 4.

RADiCe analysis starts from a model of the datacenter infrastructure in OpenDC, containing the physical and logical topology of the hardware and services in the datacenter [43]. The user defines a risk model for the datacenter (❶). A risk model describes the risk factors that constitute the overall risk in a datacenter, specifies how these factors are quantified and assigns an impact to each factor (see Section 3.2). After defining an initial risk model, it is important to validate it using the results reported by RADiCe (❷), for instance, using historical measurements, and to (potentially) (re-)calibrate the model using the results from the validation step.

Once the risk model is established, RADiCe can be used to explore portfolios of scenarios that show that the current infrastructure is at risk or to reduce overall risk. Users can build such portfolios manually (❸), for example, to explore "what-if" questions about the workload, topology, or operational phenomena. Alternatively, RADiCe can automatically propose portfolios of scenarios that reduce risk in the datacenter (❹), by employing the risk optimization algorithm (Section 3.5).

RADiCe evaluates the portfolios of scenarios constructed by the user or generated automatically and produces corresponding risk profiles, highlighting the risks that threaten the datacenter (❺). These risk profiles assist datacenter operators and architects in deciding how to manage different risk classes. For example, certain risks might be accepted due to their low impact or probability, while other risks may be mitigated by adjusting the datacenter infrastructure. After the risk profiles are evaluated, and adjustments are potentially implemented, the user can re-iterate the analysis (❻).

3.4. RADiCe Architecture

Figure 5 shows the overall architecture of RADiCe, and how it leverages OpenDC capabilities for datacenter modeling and simulation.

Users interact with RADiCe with the *Risk Model Builder* (component Ⓐ in Figure 5). Through this component, the user can construct the risk model applicable to their datacenter design (Step ❶ in Section 3.3), defining risk factors of interest (as SLOs) and the impact of violating them (e.g., a monetary penalty).

Next, through the *Scenario Portfolio Builder* component (Ⓑ), users can visually construct scenarios that are common in the datacenter (e.g., failure events or performance degradation) to explore how they affect the risk of the datacenter (Step ❸ in Section 3.3). These scenarios may consider the workload, topology, scheduler, and operational phenomena (Section 2.1.2). Scenarios can be made out of pre-built components from the *Library of Components* (Ⓒ). This library contains workload, topology, and operational building blocks, facilitating fast and intuitive composition of scenarios. It is pre-populated by the system with industry-standard components (using the Open Compute Project[3] as starting point), but can be augmented by the user with platform-specific components.

Once built, users can explore and evaluate the estimated risk of scenarios in a portfolio via the *Scenario Portfolio Evaluator* (Ⓓ – Step ❺ in Section 3.3). This component provides graphical overviews of the estimated risk, highlighting key selected metrics across scenarios, and gives users quick access to the outcomes of the simulation of built scenarios.

The *Risk Profile Generator* (Ⓔ) converts the risk model specified by the user into a set of queries understood by the *Query Processor*. This component will collect the query results and compute the impact of risk factor violations. In turn, it generates a risk profile of each scenario, estimating the risk of each scenario.

The *Query Processor* (Ⓖ) aggregates metrics received from the simulator based on the active queries. This component instructs the *Metric Ingestion Engine* on what metrics to collect from the simulator. Queries processed by this component are evaluated on the fly during simulation to enable dynamic decision-making based on the near real-time estimated risk and prevent huge volumes of data from being generated. The *Metric Ingestion Engine* (Ⓗ) collects the metrics emitted by the telemetry system in OpenDC during simulations. This component uses a pull-based model that collects metrics from the relevant systems on-demand.

The *Risk Optimizer* (Ⓕ) component generates and optimizes alternative scenarios based on their risk profile.

3.5. Genetic Risk Optimization

Risk optimization is a form of design space exploration that selects candidate solutions based on their risk profile. As we are confronted with a large design space (e.g., in a datacenter topology, we can change the cluster sizes, the type of machines, etc.), exploring all the possible candidate solutions is not feasible even with the use of discrete-time simulation.

To optimize a datacenter for risk using RADiCe, we employ an evolutionary approach inspired by the process of natural selection [46]. Each point in the design space (e.g., a datacenter

[3] https://www.opencompute.org/

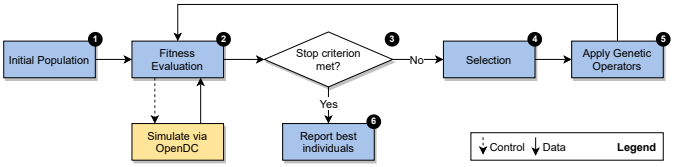

Figure 6: Exploration of the design space through genetic search.

Table 1: Experiment configurations explored in this work.

Experiment	Factors	Optimization	Sect.
Datacenter sustainability	Energy/CO_2 prices, PUE	✓	4.2
Operational phenomena	Availability target	✗	4.3
Datacenter topology	Topology	✓	4.4
Workload	Workload trace	✗	4.5

topology) is encoded as a set of chromosomes that can be altered using genetic operators (such as mutation or crossover). Starting from a population of random scenarios (individuals, ❶ in Figure 6), we evaluate their fitness by estimating their risk, each consisting of multiple simulation runs to take into account variability (❷). The fittest individuals are selected from the current population (❹) and genetic operators are applied to the selection (❺). These operators are used to either converge or diverge the solution by altering the chromosomes of individuals in the population, which helps explore the design space. The procedure is repeated until we reach the maximum number of generations or we converge on a solution (❸). Then RADiCe reports the scenarios with lower risk to the user.

In this paper, we used an evolutionary algorithm for risk optimization, but RADiCe's approach is flexible and can be generalized to incorporate other optimization algorithms, such as metaheuristics (e.g., simulated annealing [38], tabu search [30]) or operation research techniques (e.g., nonlinear programming [11], multi-objective optimization algorithms [42].

4. Evaluation of RADiCe

In this section, we perform an in-depth experimental analysis to demonstrate RADiCe's capabilities. We consider four evaluation scenarios that are synthetically described in Table 1. To improve statistical confidence, all the experiments are repeated 4,096 times, and we report the median figures of the metrics of interest. Then, we discuss the performance benefit of RADiCe over a competitor.

4.1. Evaluation Setup

This section describes the evaluation environment in terms of considered workloads, topology, scheduling policy, operational phenomena, risk models and optimization.

4.1.1. Workload

We define as *baseline* a business-critical workload trace from Solvinity, a private cloud provider. The trace characterizes the operation through a set of VM-level metrics aggregated over 5-minute intervals. The full trace spans three months of datacenter operation. It consists of 1,800 VMs running business-critical applications, collectively consuming 3,063 PFLOPs, with a mean CPU utilization of 5.6% on the original topology. This low utilization is in line with the industry, where utilization levels below 15% are typical [55, 50], as to ensure the datacenter has enough spare capacity in the presence of failures and to reduce the risk of workload interference.

To analyze the impact of workload on risk in datacenters (Section 4.5), we select two other business-critical traces published in the Grid Workloads Archive [35], Bitbrains [51] and Materna [40], which are both similar in scale and structure. Bitbrains is a workload trace spanning one month of operation across 1,250 VMs, while Materna consists of 547 VMs stretching over a period of three months. Finally, we also consider a public cloud trace from Azure [19]. The original trace contains over 2 million VMs. To provide a realistic comparison, we randomly sample approximately 1,800 VMs using the trace sampling tools offered by OpenDC, matching the size of the baseline workload. All the traces used for the evaluations are real-world traces (or subsets of), collected from major datacenter operators. Except for the baseline, traces are also publicly available.

4.1.2. Datacenter Topology

In the evaluation, we consider the topology of the *baseline* workload. This consists of 12 compute clusters comprising approximately 200 physical hosts, spread over three datacenters. These datacenters are connected through a fiber optic ring network, while the servers in the clusters communicate via a low-latency InfiniBand network.

The topology has two sets of clusters: *standard* clusters and *bigmem* clusters. Standard clusters contain 16 servers, with two 8-core CPUs and 128 GB of memory, while bigmem clusters contain 6 servers, each with four 8-core CPUs and 768 GB of memory.

The datacenters in this topology are assumed to have a PUE of 1.55. This value is equivalent to the global average in 2022 [20]. Although hyperscale datacenters report substantially lower PUE values (ranging 1.1–1.4), we focus in this work mainly on *mid-tier providers* of cloud infrastructure where such values are not yet common. The CO_2 emission factor (carbon intensity) for these datacenters, which describes the CO_2 emissions produced per unit of electricity consumed, is assumed to be 556 kg CO_2 per MWh of electricity consumed. This assumption is based on the estimated CO_2 emissions in the entire supply chain for the consumption of gray energy sources in the datacenter country of origin. In Section 4.4, we explore variations of this topology stemming from manual modification and automated optimization.

4.1.3. Metrics and Scheduling Policies

RADICE and OpenDC fully support 30+ common metrics from industry-grade OpenTelemetry SDK, including system (CPU utilization, power, VM counts, etc.), customer (e.g., VM-start latency, availability), and datacenter (CO2 emissions, etc.) characteristics.

OpenDC compute scheduler is designed after the Filter Scheduler[4] from OpenStack, a popular open-source cloud computing platform. It operates in two stages: *filtering* and *weighing*. During the filtering, the scheduler selects available hosts based on a set of user-configured policies (e.g., based on the number of available vCPUs, the remaining RAM, or affinity rules). In the weighing phase, the scheduler uses a selection of policies (e.g., based on the number of VMs allocated to the host, or the available RAM) to assign weights to the hosts that survived the filtering phase. To avoid overloading any single host, the final destination for the VM is chosen randomly from a subset of the highest-ranked hosts. By combining different filters and weighers, the user can re-create in RADICE many of the traditional VM allocation policies described in the literature [53]. For instance, round-robin scheduling can be implemented by ordering hosts based on the number of VMs allocated to them, while a random scheduling policy is achieved by disabling all weighers and setting the number of selected hosts in the filtering phase to infinite. For this evaluation, we emulate the default behavior of OpenStack's FilterScheduler by evenly spreading VMs across all hosts based on available RAM.

4.1.4. Operational Phenomena

We consider two common types of operational phenomena: (1) performance variability due to workload interference and (2) correlated cluster failures.

We assume a common model for workload interference [39, 54], where a set of collocated workloads is assigned a *score* from 0 to 1, with 0 indicating total interference between VMs contending for the same physical CPU, and 1 indicating non-interfering VMs, at a given CPU load level. In our simulation, we randomly generate the score according to the monitoring data (e.g., placement of VMs, CPU utilization, etc.) of the baseline workload.

We consider a model of space-correlated failures for cluster failures [27], in which a failure might trigger more failures within a short time span; these failures constitute a *group*. We limit this phenomenon to hardware failures that crash machines (full-stop failures), with subsequent recovery after some duration. Space-correlated failures in large-scale distributed systems follow a lognormal distribution [27, 36]. For our experiments, we choose parameters inspired by GRID'5000 [27] (public trace also available [36]) and Microsoft Philly [37], scaled to the size of the baseline topology. The failure duration is restricted by a minimum of 15 minutes.

[4]https://docs.openstack.org/nova/latest/admin/scheduling.html#the-filter-scheduler

Table 2: Risk factors considered in this work. The Aggregation Period (A.P.) is one month (M) or one day (D).

Risk Factor	Objective	A.P.	Impact
Availability (Uptime)	≥ 99.5%	M	See Table 3
Scalability (Sched. Latency)	< 1 hour	M	
QoS (VM Interference)	≥ 97.5%	D	See Table 3
Electricity	-	M	€241 per MW/h (Fig. 7)
CO_2 Emissions (Company)	-	M	€81 per tCO_2 (Fig. 7)
Resource Saturation	< 75%	D	€125 per incident
Resource Imbalance	< 0.2	D	€125 per incident
CO_2 Emissions (Society)	-	M	€395 per tCO_2 [49]

Table 3: Refund policy based on Availability (left) and Quality of Service (right) SLAs. We assume customers are charged $0.046 per vCPU-hour, based on the cost of a *m6g.medium* instance in the Frankfurt region of AWS.

Monthly Uptime [%]	Refund
< 95%	100%
< 99%	30%
< 99.5%	10%

Quality of Service [%]	Refund
< 90%	100%
< 95%	30%
< 97.5%	10%

4.1.5. Risk Model

We consider for our experiments eight different risk factors (Section 3.2) present in datacenters and summarize them in Table 2. Our selection covers risks that impact customers, risks related to operational sustainability, and risks facing society.

Customer-facing risks. The key performance indicator customers of datacenter operators are concerned with is availability. We configure RADICE to ensure a monthly uptime target of 99.5% for individual VMs. Customers with a VM failing to reach this uptime target are refunded according to the refund policy in Table 3 (left). This kind of SLA is common in public cloud providers, such as Amazon Web Services[5], Google Cloud[6], and Microsoft Azure[7]. In addition, We want to ensure that customer requests are served promptly (*scalability*). If a VM is not started within an hour of its submission time, the customer is refunded all expenses for that VM while it was pending to be scheduled.

Finally, we also consider QoS as a customer-facing risk. Even if it is rare for cloud providers to offer performance guarantees to customers [47], datacenter operators monitor performance metrics carefully since a low quality of service can lead to reputational damage or loss of customers. To quantify the impact of performance variability due to workload interference, we monitor CPU contention and interference metrics as a proxy for QoS, and refund customers with VM with average CPU interference values above 2.5% per day (equivalent to a QoS value lower than 97.5%), according to the refund policy in Table 3

[5]`https://aws.amazon.com/compute/sla/`

[6]`https://cloud.google.com/compute/sla`

[7]`https://www.microsoft.com/licensing/docs/view/Service-Level-Agreements-SLA-for-Online-Services`

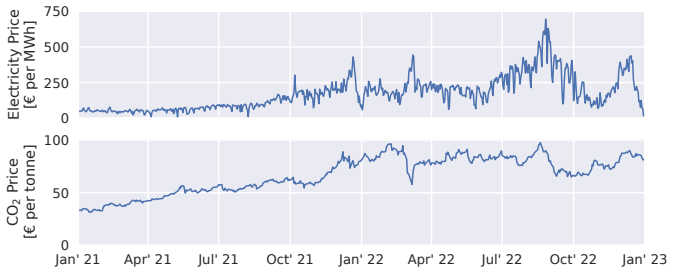

Figure 7: Day-ahead electricity prices in the APX Power Spot Exchange from 2021 to 2023 in € per MW h (Source: APX) and daily close prices for European Union Emissions Trading System (ETS) Futures from 2021 to 2023 in € per tonne CO_2 (Source: ICE).

(right). This policy is derived from existing cloud SLAs and uses commonly accepted thresholds in the industry.

Company-facing risks. To model the operational sustainability of the datacenter operator, we consider resource utilization metrics for hosts. Usually, a high resource utilization or a large imbalance indicates an infrastructure issue and demands an engineer's investigation. We configure RADICE to ensure that the 95th percentile of the host utilization does not exceed 75%, and that the resource imbalance (defined as the standard deviation of the host utilization) does not exceed 0.2. Such values are commonly used in datacenter monitoring systems to alert operators of issues. We assign an impact of €125 to an investigation. This is an optimistic estimation of the costs of an engineer needing to investigate an issue with the infrastructure and taking just two hours to solve the issue.

Furthermore, we consider the expenses for electricity usage and CO_2 emissions incurred by the datacenter operator. We assume the cost of electricity is €241 per MW h, based on the average energy price reported by the Amsterdam Power Exchange (APX) for 2022 (Figure 7, top). The cost of CO_2 emissions is assumed to be €81 per tonne of CO_2, based on the average price reported by International Exchange (ICE) for 2022 (depicted in Figure 7, bottom).

Society-facing risks. Our model also includes risks that society is confronted with as a consequence of datacenter operation. We use the social cost of carbon (SCC) to model the risk of CO_2 emissions due to electricity usage by datacenters. Ricke et al. estimate the global social cost of CO_2 to be $417 per tonne of CO_2 [49] (€395 per tCO_2 as of 2022).

4.1.6. Genetic Algorithm

The evolutionary risk optimization approach described in Section 3.5 is implemented using Jenetics [59], a Java library for building genetic algorithms. Table 4 summarizes the configuration parameters used for the genetic search algorithm.

In our experiments, we limit the number of generations to 100, and the algorithm terminates earlier if the average fitness of the last 10 generations differs by no more than 0.01% from the average fitness of the last 30 generations, indicating convergence. We use tournament selection [45] to choose the best individual from a random sample of five individuals (with replacement) from the population. We employ three types of genetic operators to the selected individuals:

1. *Uniform crossover* combines the genetic information of two parents, swapping two genes with probability 0.2. Empirical studies suggest this improves design space exploration while preserving beneficial information exchange [17].
2. *Uniform mutation* introduces genetic diversity by randomly changing a gene's value with a probability of 0.15.
3. *Gaussian mutation* is applied with a probability of 0.10, adjusting a gene's value based on a Gaussian distribution around its current value, facilitating the exploration of improved solutions near the current individual.

While the optimization results later discussed depend on the specific parameter settings used in this study, the user can tune the genetic algorithm based on their needs. We leave the systematic exploration of these parameters and their impact on the final results for future work.

Table 4: Configuration of the genetic algorithm used for the evaluation.

Parameter	Value
Population Size	30 individuals
Stop Criteria	10 steady generations or at most 100 generations
Selection	Tournament selection [45]
Genetic Operators	Uniform crossover (probability 0.2), uniform mutation (probability 0.15), Gaussian mutation (probability 0.10)

4.2. Analyzing the Sustainability of Datacenters

We look at the sustainability of datacenters from two different perspectives, that of the company having to sustain operation (*operational sustainability*), and that of the environment having to cope with datacenter operation (*environmental sustainability*). Due to the large environmental footprint of datacenters, these perspectives represent a trade-off for infrastructure operators. This is becoming increasingly relevant as governments could start taxing companies based on their societal and environmental footprints [20].

We focus first on the price spikes for electricity and CO_2 bonds that occurred in 2022. In Figure 1 (first page), we depict the risk profile reported by RADICE for 2022, compared to the risk profile of the year before (2021), considering the average energy price registered during the year. We observe that in 2021, the primary factors contributing to the risk profile of the datacenter are availability, covering the costs of potential failures in the datacenter (see Section 4.3), and CO_2 emissions for society. By contrast, in 2022, the prices for electricity and CO_2 emission bonds have increased significantly, leading to electricity expenses becoming the primary risk factor for datacenters in Figure 1. Whereas the availability risk is highly variable and has a much lower median cost (as we show in Section 4.3), the electricity demand and CO_2 emissions of datacenters are relatively stable, leading almost certainly to higher expenses and making this such a serious problem.

These price increases can have even more devastating effects on operators of energy-inefficient datacenters. In Figure 8,

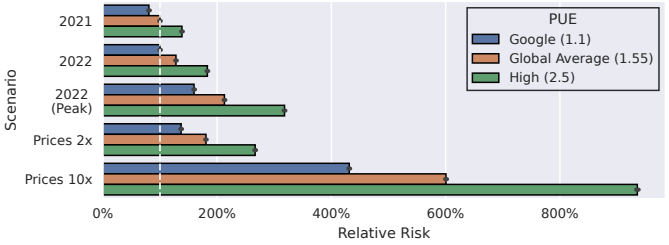

Figure 8: Relative risk for the baseline workload for different price-surge scenarios, compared to 2021. The dashed vertical line represents 100% or the situation of 2021 for a datacenter with an average energy-efficiency.

we illustrate the increased monthly costs that datacenter operators face compared to 2021. We consider multiple price-surge scenarios, as well as the energy efficiency of the datacenter, in terms of PUE. On average, datacenters operate at a PUE of 1.55 [20], while older, less energy-efficient facilities may reach a PUE of 2.5 (equivalent to the average PUE in 2007) or worse. By contrast, hyperscale datacenters, such as those from Google and Facebook, reach considerably lower PUE values of 1.1 or below[8].

Unsurprisingly, we find that datacenters with a high PUE are affected the most by the increasing prices. But, using RADICE, we can actually quantify this risk as being a factor 1.8x–2.2x higher compared to hyperscale facilities. Even if these datacenters still have plenty of opportunities left to improve their energy efficiency, the price increases for electricity and CO_2 bonds pose a real threat. prices increase by another factor 10x compared to 2021, datacenter operator's monthly expenses could rise by 5x–10x.

Finally, we investigate the discrepancy between the price paid by datacenters operators for emitting CO_2, compared to the overall costs incurred by society due to CO_2 emissions, also called the *social cost of carbon*. The European Union Emissions Trading System (ETS) was the first large-scale trading scheme for greenhouse gas emissions. It requires installations to obtain bonds (also called allowances) to cover their emissions, with each bond permitting the emission of 1 tonne of CO_2. In Figure 9, we depict the effect of adjusting the cost of CO_2 on monthly costs incurred by a datacenter operator, where we assume that the risk factors reported by RADICE encompass all costs of the operator, and that the operator runs at an operating margin of 20%. Surprisingly, increasing the CO_2 prices by 5x, to roughly match the social cost of CO_2 (€395 per tCO_2 as estimated by Ricke et al. [49]) consumes the entirety of the datacenter operator's operating margin of 20%. Thus, further price increases of CO_2 bonds could threaten the operational sustainability of datacenters, especially when legislators decide to narrow down the discrepancy between the bond and the societal costs.

[8] `https://www.google.com/about/datacenters/efficiency`

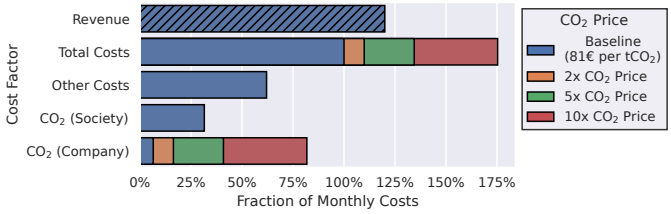

Figure 9: CO_2 costs relative to the other costs incurred by the datacenter operator per month.

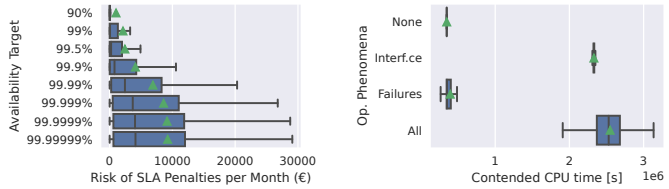

Figure 10: Left: impact of availability target on incurred SLA penalties (99.5% is the baseline availability target). The top-3 risk costs are above €99,000 (outliers have been removed for readability). Right: Total contended CPU time (in s) experienced by VM for different operational phenomena over a timespan of three months.

Key observations:

1. Electricity expenses have become the primary risk in 2022 due to increasing electricity prices.
2. Energy-inefficient facilities are impacted by price increases by a factor 1.8x–2.2x more compared to hyperscale infrastructure operators.
3. The societal impact of the CO_2 emissions is one of the highest risks in 2022, yet only a fraction of that risk is currently carried directly by the company or customer.

4.3. Impact of Operational Phenomena

This experiment investigates the impact of operational phenomena on the risks faced by datacenter operators. Figure 10 (left) depicts the SLA penalties incurred by the datacenter operator per month for different VM availability targets. The left and right limits of the boxes indicate the first quartile (Q1) and the third quartile (Q3), while the whiskers extend to show the rest of the distribution. The middle line of a box plot represents the median, and the green triangle the average. For the baseline scenario, we find that the availability risk is highly variable since the cost can vary significantly depending on how long the downtime is and which machine is affected. To illustrate, the average cost for availability SLA violations is €2,400 per month. In most cases, the company incurs a cost of less than €250 per month, yet in the worst case, the company faces €100,000 in penalties (outliers are not shown in the plot). Furthermore, the risk increases as we increase the availability target, highlighting the difficulty and risk of guaranteeing higher availability of individual VM to customers.

We also investigate the effects of operational phenomena on the quality of service or overall performance. Figure 10 (right) depicts the contended CPU time of all VMs in the baseline

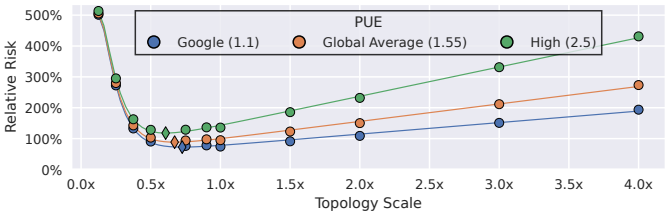

Figure 11: Regression of the relative risk on the topology scale (the number of machines) on datacenter risk in terms of the per-month risk. The diamond markers indicate the minimum risk for each category.

workload in the presence of different operational phenomena. The contended CPU time metric represents the amount of time that a VM was ready to run but was waiting while the hypervisor serviced other workloads and gives an impression of the quality of service experienced by VMs. We observe that workload interference significantly impacts CPU contention metrics, representing 87% of the recorded CPU contention with the presence of failures. Furthermore, we find that failures increase the average CPU contention and variability. This is not unexpected since failures may affect scheduling decisions or mask periods of high interference.

Although workload interference represents a substantial portion of the recorded CPU contention values, it has a negligible impact on risk in our baseline scenario, as shown in Figure 1. This is because resource utilization in the baseline workload is relatively low, and therefore there are fewer opportunities where interference could happen. In Section 4.4, we show that increasing the resource utilization (and thus placing more VM on the same host) eventually leads to a non-negligible QoS risk for the datacenter. We explore the impact of different workloads in Section 4.5.

Key observations:

1. The risk of violating the availability SLA has a high variability. In the considered scenario, the monthly incurred cost is less than €250, whereas in the worst case, the operator could incur €100,000 (400x higher) in penalties due to SLA violations.

2. CPU interference is low risk at low resource utilization (over-provisioned datacenter). Highly-utilized datacenters increase such risk.

4.4. Impact of Datacenter Topology

In this experiment, we examine how scaling a datacenter impacts its risk, and we employ the built-in risk optimization algorithm of RADICE to explore new hardware configurations and cluster sizes, optimizing the datacenter design for risk.

Scaling down datacenter infrastructure, and in turn, operating at a higher resource utilization could enable cost reductions for electricity expenses. However, there is no free lunch because to satisfy customer SLAs, datacenters need to maintain enough capacity to serve customers, even in the presence of failures. Too little capacity and a datacenter will not be able to absorb the full workload as a result of a crash in another datacenter. We highlight this trade-off in Figure 11, where we compare the relative risk of the baseline workload for different topology scales, taking as reference the risk in the baseline topology. Interestingly, the risk of running the baseline workload at half the size of the original datacenter is almost equivalent to our baseline scenario. In this case, risk has become a trade-off between electricity expenses and SLA penalties due to lowered quality of service. The figure also shows that a 0.5x–0.75x scale leads to the lowest per-month risk of all explored scenarios. The figure exhibits a hockey-stick type curve, where risk slowly increases as the scale of the datacenters is increased (due to higher electricity expenses), while the risk rises sharply as the datacenter is downscaled (as a consequence of SLA penalties).

So far, we have only considered the scale of the datacenter topology (the number of machines). In reality, there are other factors that influence the results of our experiments, such as the type of hardware used, the combination of hardware, or the age of the hardware. As discussed in Section 3.5, manually considering each of these factors is infeasible, and instead, an intelligent approach for systematically exploring the design space is necessary. The risk exploration algorithm in RADICE can explore new hardware configurations and cluster sizes. In Figure 12, we depict the risk estimation for the top-4 best-performing topologies discovered by our exploration algorithm. We observe that, on average, the optimized topologies only incur a monthly risk of 0.65x–0.7x compared to the baseline topology, while for the median case only 0.4x–0.6x. The solution set contains two notable characteristics: (1) machines are primarily vertically scaled (e.g., increased CPU count), and (2) hardware models are more recent. Presumably, this is a result of the algorithm optimizing for the energy efficiency of the topology.

Focusing now on the individual topologies, we find that, despite the best-case scenario for C4 incurring the lowest costs of all topologies, its worst-case scenario leads to the highest costs of all optimized topologies. This topology contains less than half the capacity of C1 (in terms of CPU count), which explains the high variability of this solution since such a topology leaves little room for spare capacity in the event of a failure. By contrast, while C2 has a higher median risk, its variability is lower than other solutions. This configuration uses many low-power commodity servers.

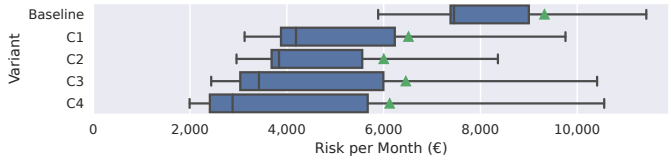

Figure 12: Per-month risk estimation of optimized topology candidates discovered by the risk exploration algorithm in RADICE.

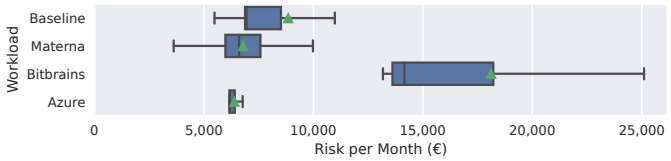

Figure 13: Per-month risk estimation of different workloads running on the *baseline* topology.

Key observations:

1. Amid the soaring electricity prices, datacenter providers must compromise between topology scale and SLA compliance. However, this can lead to risk trade-offs that RADICE can help investigate.
2. RADICE's risk optimization algorithm reduces average risk in datacenters by optimizing the datacenter topology (by a factor of 0.65x–0.7x in the considered workload).

4.5. Impact of Workload

This experiment investigates the impact of introducing new workloads for datacenter operators. Figure 13 shows the risk estimation of running different workloads using our *baseline* topology and highlights significant differences. The Materna workload has a lower overall risk, but a similar variability to the *baseline* workload. The Azure workload has both a lower risk and variability compared to the *baseline* workload. In contrast, the Bitbrains workload has significantly higher risk and variability than the other workloads.

The Materna workload exhibit relative risk contributions of SLA penalties and operating costs (such as electricity and CO_2 emissions) similar to the baseline workload (Figure 14). Interestingly, despite the similar risk profile, the Materna workload is less than one third the size of the baseline workload in terms of VM count. However, VMs in the Materna workload have, on average, double the runtime compared to the baseline workload, meaning the number of active VMs in the system is higher.

Although the baseline and Bitbrains workloads are similarly sized, we observe relatively high host saturation and host imbalance in Figure 14, where we have depicted the risk profile of the Bitbrains workload running on the *baseline* topology. A high host saturation indicates that utilization of the physical hosts is too high on average. In contrast, a high host imbalance indicates a large difference in utilization between physical hosts. This means either the datacenter is underprovisioned for our workload or the scheduler is not doing a good job distributing the load over the physical hosts.

Finally, focusing on the Azure workload, we observe that electricity costs and CO_2 emissions dominate the risk profile for this workload. In contrast, its availability risk is significantly lower than the other workloads. Usually, this indicates that the datacenter running the workload is over-provisioned.

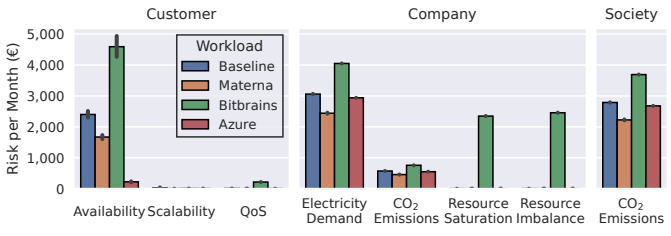

Figure 14: Risk profile for different workloads using the *baseline* topology.

Key observation:
Changing workloads can significantly affect the risk profile of a cloud datacenter. RADICE supports risk analysis for different workloads appearing in cloud datacenters.

4.6. Comparison against CloudSim

We compare the performance of RADICE against CloudSim Plus [24] (v7.1.1), a distribution of the well-known cloud simulation framework CloudSim [14]. Being a datacenter simulator, and not a risk assessment framework, we manually construct in CloudSim Plus a scenario that simulates the *baseline* workload (see Section 4.1.1) running in a datacenter using the *baseline* topology described in Section 4.1.2, and which measures the total energy consumed by the machines in the datacenter. To analyze systems' scalability, we consider multiple scaled-down variants of this scenario for scale factors 50%, 25%, and 5%. We scale both the workload and topology in terms of size (virtual machines and physical machines respectively) according to the scale factor.

We measure both the *runtime* (in ms) and *peak memory usage* (in MB) during simulation, using performance counters exposed by the Java runtime. Our measurements do not include the experiment setup (e.g., reading the workload traces or parsing the topology description) or experiment tear-down. All scenarios are replicated 32 times: this number of replications leads to a sufficiently small error in our measurements, supporting our observations about the overall performance of RADICE with high confidence. The experiments are executed using OpenJDK 17 on Arch Linux (Linux kernel 5.13.19), running on a machine equipped with an AMD Ryzen ThreadRipper 3990x 64-core chip, 128 GB of DDR4 RAM, and 8 TB of fast NVMe storage.

Table 5 lists the measurements of runtime (in s) and peak memory usage (in MB) for both simulators running the selected workloads. RADICE can simulate the selected workloads orders of magnitudes faster than CloudSim Plus, ranging from a factor 70x in runtime for smaller workloads to a factor 330x for the largest workload. In terms of absolute values, RADICE is able

Table 5: Mean simulation runtime and peak memory usage (along with the standard deviation) of both RADICE and CloudSim Plus when simulating various workloads scales.

Scale	**Runtime** [s]		**Peak Memory Usage** [MB]	
	RADICE	CloudSim Plus	RADICE	CloudSim Plus
5%	0.11 ± 0.01	7.53 ± 0.02	59.01 ± 0.05	59.73 ± 0.05
25%	0.66 ± 0.01	159.78 ± 1.08	246.37 ± 0.05	246.32 ± 0.10
50%	1.69 ± 0.01	396.89 ± 0.92	484.69 ± 0.05	483.42 ± 0.10
100%	3.29 ± 0.03	1,099.83 ± 3.01	939.32 ± 0.05	942.05 ± 0.10

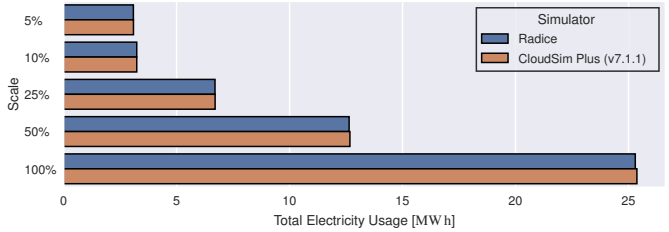

Figure 15: Electricity usage of the datacenter (in MWh) reported by RADiCe compared to CloudSim Plus when simulating equivalent scenarios.

to analyze three months of datacenter operation (of 1,800 virtual machines) in a matter of seconds, whereas CloudSim Plus requires *over 18 minutes* to evaluate the same scenario. Figure 15 shows the electricity usage of the datacenter reported by RADiCe compared to CloudSim Plus when simulating equivalent scenarios. We observe that the electricity usage reported by both simulators is nearly identical. Only for the larger workload scales (50% and 100%), a small difference is visible, resulting in up to the 0.3% of discrepancy. We believe that these small differences are due to the different support for CPU overcommitment in CloudSim Plus and RADiCe. Manual inspection of the results shows that RADiCe can generate an identical VM allocation schedule as produced by the default allocation policy of CloudSim Plus (worst-fit).

The runtime performance of RADiCe opens possibilities for using the tool interactively while discussing operational changes by providing datacenter operators with coarse risk estimates for various "what-if" scenarios. For higher degrees of confidence, RADiCe can be used to automatically consider thousands of risk scenarios that must be otherwise manually built in standard datacenter simulators. For the results presented in this paper, this allowed us to programmatically simulate over 35,000 years of datacenter operations in over 200,000 simulation-runs, in a reasonable running time.

5. Related Work

Risk Analysis for Datacenters. Risk management is a critical process for ensuring high quality of cloud services.

The community has proposed several general methodologies and systems for risk analysis in cloud computing, which we have summarized in Table 6. The AssessGrid project [22] introduces the notion of risk assessment in grids as a decision paradigm, using high-level mathematical models to analyze the impact of SLA violations. Similarly, the OPTIMIS [23, 21] project also employs mathematical models to quantify risks, taking into account historical SLA violations and estimates of the reliability of the environment. Yeo and Buyya describe four objectives for balancing risk in grids and develop two evaluation methods to validate the effectiveness of resource management policies in attaining these objectives [61]. Albakri et al. propose a qualitative method for security in cloud computing environments, which actively involves cloud customers in the evaluation of security risk factors [3]. SEBCRA [25] is a semi-quantitative cloud risk assessment model that focuses on evaluating the impact of cloud-specific risks on organizations' business-level objectives but requires experts' knowledge to establish probabilities and impacts.

Table 6: Frameworks for risk analysis in datacenters.

Framework	T	MT	C	Focus	GUI
AssessGrid [22]	●	M	A	SLA negotiation for Grids	✗
OPTIMIS [23, 21]	●	M	A	Cloud Brokering	✓
Yeo and Buyya [61]	●	M	A	Utility Computing	✗
Albakri et al. [3]	■	-	A	Security	✗
SEBCRA [25]	◗	-	A	Business Objectives	✗
Janus [4]	●	S	I	Network Planning	?
RSS [60]	●	S	I	Network Risk Identification	?
RADiCe (this work)	●	S	A	Infrastructure Optimization, Sustainability	✓

Key: "T": the analysis technique (●: quantitative, ◗: semi-quantitative, ■: qualitative), "MT": the modeling technique ("M": mathematical, "S": simulation – when applicable), "C": reference community ("A": academia, "I": industry), GUI: if they provide or not a Graphical User Interface (✗: No, ✓: Yes, **?**: Unknown).

Risk management is also used for computer networks. For example, Janus [4] is a system for planning network changes used by Google, which uses flow-level Monte-Carlo simulations to evaluate the impact of various risk scenarios based on operator-specified risks and probabilities. The Risk Simulation System (RSS) [60] from Facebook identifies possible issues in the company's backbone and quantifies their potential impact using network simulation and a set of network risk metrics. The design of RADiCe follows an approach similar to these systems but offers a unified model to express, analyze and optimize, various risk factors including sustainability metrics.

Datacenter Simulation. The research community has built many high-quality simulators that provide a rich set of features to build upon [13, 8]. However, many of the existing simulators typically offer a single feature or are very general and thus require repeating the work we propose here for each specific model.

CloudSim [14] is the closest simulator in nature to OpenDC. It focuses on simulating cloud system components, including virtual machines, data centers, and resource provisioning policies. It includes numerous other single-feature simulators, such as iFogSim [32], WorkflowSim [18], and CloudAnalyst [58]. However, the single-feature simulators extend CloudSim each in their direction and can only be combined with extensive engineering. In contrast, OpenDC offers an integrated approach and specific modeling advances for its main features. Similar is the SimGrid framework [16], which serves as the foundation of many simulators, such as SimGrid VM [34], Schlouder [44], and WRENCH [15]. In contrast to OpenDC, it supports not

only clouds but also P2P networks, and runs at a much finer granularity, which in turn enables emulation of specialized applications (e.g., MPI). However, SimGrid and its ecosystem do not support modern workloads (e.g., serverless) and operational phenomena.

6. Conclusion

In this paper, we proposed RADiCe, an open-source framework for risk analysis and optimization in sustainable data-center. RADiCe incorporates discrete-event simulation to model datacenter infrastructure and diverse workloads accurately and timely. It features the ability to define a risk model, explore what-if risk scenarios (such as outages or interference), evaluate risk profiles, and propose changes to the datacenter that reduce risk. RADiCe's unique tuning and exploratory capabilities enable obtaining *new, credible, quantitative* evidence of risk trade-offs.

Future research spans multiple directions. First, we aim to expand our model of IT-related operational risks to include networking, security, and sustainability factors. Specifically, considering sustainability-oriented policies (e.g., using time-shifting to execute workloads during periods of higher green energy availability) could highlight trade-offs between risk (e.g., related to QoS), and the sustainability impact of datacenter operations. Second, we acknowledge that modeling risks such as QoS violations solely based on their financial impact of penalties may overlook other critical risks, such as reputational damage (e.g., the loss of existing customers due to SLA violations). While translating reputational risk into quantifiable costs is often possible, developing effective methods for doing so remains a complex research challenge. Recent studies ([31, 52]) offer only preliminary empirical insights into this issue. We believe further collaboration and discussion among data center operators and stakeholders is necessary to fully address these aspects and integrate them into RADiCe's risk models. Finally, we would like to investigate other risk optimization techniques (e.g., based on operational research or meta-heuristics) and their impact on the obtained trade-offs.

Acknoweldgement

This work was partially supported by EU MSCA CloudStars (g.a. 101086248) and EU Horizon Graph Massivizer (g.a. 101093202). This research is partly supported by a National Growth Fund through the Dutch 6G flagship project "Future Network Services".

References

[1] ISO/IEC 30134-2:2016. 2016. *Information technology – Data centres – Key performance indicators — Part 2: Power usage effectiveness (PUE)*. Standard. International Organization for Standardization, Geneva, CH.

[2] ISO 31000:2018. 2018. *Risk management – Guidelines*. Standard. International Organization for Standardization, Geneva, CH.

[3] Sameer Hasan Albakri, Bharanidharan Shanmugam, Ganthan Narayana Samy, Norbik Bashah Idris, and Azuan Ahmed. 2014. Security risk assessment framework for cloud computing environments. *Security and Communication Networks* 7, 11 (2014), 2114–2124.

[4] Omid Alipourfard, Jiaqi Gao, Jérémie Koenig, Chris Harshaw, Amin Vahdat, and Minlan Yu. 2019. Risk based planning of network changes in evolving data centers. In *Proceedings of the 27th ACM Symposium on Operating Systems Principles, SOSP 2019, Huntsville, ON, Canada, October 27-30, 2019*, Tim Brecht and Carey Williamson (Eds.). ACM, 414–429.

[5] Georgios Andreadis, Fabian Mastenbroek, Vincent Van Beek, and Alexandru Iosup. 2021. Capelin: Data-Driven Compute Capacity Procurement for Cloud Datacenters using Portfolios of Scenarios. *IEEE Transactions on Parallel and Distributed Systems* (2021), 1–1.

[6] Georgios Andreadis, Laurens Versluis, Fabian Mastenbroek, and Alexandru Iosup. 2018. A Reference Architecture for Datacenter Scheduling: Design, Validation, and Experiments. In *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis*. 478–492. https://doi.org/10.1109/SC.2018.00040

[7] Douglas Donellan Andy Lawrence and Lenny Simon. 2022. *Annual outage analysis 2022*. Technical Report. Uptime Institute.

[8] Ilyas Bambrik. 2020. A Survey on Cloud Computing Simulation and Modeling. *SN Computer Science* 1, 5 (2020), 249.

[9] Luiz André Barroso, Urs Hölzle, and Parthasarathy Ranganathan. 2018. *The Datacenter as a Computer: Designing Warehouse-Scale Machines, Third Edition*. Morgan & Claypool Publishers.

[10] Rabih Bashroush and Andy Lawrence. 2020. *Beyond PUE: Tackling IT's wasted terawatts*. Technical Report. Uptime Institute.

[11] D.P. Bertsekas. 1999. *Nonlinear Programming*. Athena Scientific.

[12] Bloomberg. 2022. Big Tech Gets Caught Up in Europe's Energy Politics. [Online; accessed September-2024].

[13] James Byrne, Sergej Svorobej, Konstantinos M. Giannoutakis, Dimitrios Tzovaras, Peter J. Byrne, Per-Olov Östberg, Anna Gourinovitch, and Theo Lynn. 2017. A Review of Cloud Computing Simulation Platforms and Related Environments. In *CLOSER 2017 - Proceedings of the 7th International Conference on Cloud Computing and Services Science, Porto, Portugal, April 24-26, 2017*, Donald Ferguson, Víctor Méndez Muñoz, Jorge S. Cardoso, Markus Helfert, and Claus Pahl (Eds.). SciTePress, 651–663.

[14] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose, and Rajkumar Buyya. 2011. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience* 41, 1 (2011), 23–50.

[15] Henri Casanova, Rafael Ferreira da Silva, Ryan Tanaka, Suraj Pandey, Gautam Jethwani, William Koch, Spencer Albrecht, James Oeth, and Frédéric Suter. 2020. Developing accurate and scalable simulators of production workflow management systems with WRENCH. *Future Generation Computer Systems* 112 (2020), 162–175.

[16] Henri Casanova, Arnaud Giersch, Arnaud Legrand, Martin Quinson, and Frédéric Suter. 2014. Versatile, scalable, and accurate simulation of distributed applications and platforms. *J. Parallel and Distrib. Comput.* 74, 10 (2014), 2899–2917.

[17] Pravir K. Chawdhry, Rajkumar Roy, and Raj K. Pant. 1997. *Soft Computing in Engineering Design and Manufacturing*. Springer-Verlag.

[18] Weiwei Chen and Ewa Deelman. 2012. WorkflowSim: A toolkit for simulating scientific workflows in distributed environments. In *8th IEEE International Conference on E-Science, e-Science 2012, Chicago, IL, USA, October 8-12, 2012*. IEEE Computer Society, 1–8.

[19] Eli Cortez, Anand Bonde, Alexandre Muzio, Mark Russinovich, Marcus Fontoura, and Ricardo Bianchini. 2017. Resource Central: Understanding and Predicting Workloads for Improved Resource Management in Large Cloud Platforms. In *Proceedings of the 26th Symposium on Operating Systems Principles, Shanghai, China, October 28-31, 2017*. ACM, 153–167.

[20] Jacqueline Davis, Daniel Bizo, Andy Lawrence, Owen Rogers, Max Smolaks, Lenny Simon, and Douglas Donnellan. 2022. *2022 Data Center Industry Survey Results*. Technical Report. Uptime Institute.

[21] Karim Djemame, Django Armstrong, Jordi Guitart, and Mario Macías. 2016. A Risk Assessment Framework for Cloud Computing. *IEEE Transactions on Cloud Computing* 4, 3 (2016), 265–278.

[22] Karim Djemame, Iain Gourlay, James Padgett, Georg Birkenheuer, Matthias Hovestadt, Odej Kao, and Kerstin Voß. 2006. Introducing Risk Management into the Grid. In *Second International Conference on e-Science and Grid Technologies (e-Science 2006), 4-6 December 2006,*

Amsterdam, The Netherlands. IEEE Computer Society, 28.

[23] Ana Juan Ferrer, Francisco Hernández-Rodriguez, Johan Tordsson, Erik Elmroth, Ahmed Ali-Eldin, Csilla Zsigri, Raül Sirvent, Jordi Guitart, Rosa M. Badia, Karim Djemame, Wolfgang Ziegler, Theo Dimitrakos, Srijith K. Nair, George Kousiouris, Kleopatra Konstanteli, Theodora A. Varvarigou, Benoit Hudzia, Alexander Kipp, Stefan Wesner, Marcelo Corrales, Nikolaus Forgó, Tabassum Sharif, and Craig Sheridan. 2012. OPTIMIS: A holistic approach to cloud service provisioning. *Future Generation Computer Systems* 28, 1 (2012), 66–77.

[24] Manoel C. Silva Filho, Raysa L. Oliveira, Claudio C. Monteiro, Pedro R. M. Inácio, and Mário M. Freire. 2017. CloudSim Plus: A cloud computing simulation framework pursuing software engineering principles for improved modularity, extensibility and correctness. In *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Lisbon, Portugal, May 8-12, 2017*. IEEE, 400–406.

[25] Josep Oriol Fitó and Jordi Guitart. 2014. Business-driven management of infrastructure-level risks in Cloud providers. *Future Generation Computer Systems* 32 (2014), 41–53.

[26] Flexera. 2023. *State of the Cloud Report*. Technical Report. Flexera. [Online; accessed September-2024].

[27] Matthieu Gallet, Nezih Yigitbasi, Bahman Javadi, Derrick Kondo, Alexandru Iosup, and Dick H. J. Epema. 2010. A Model for Space-Correlated Failures in Large-Scale Distributed Systems. In *Euro-Par 2010 - Parallel Processing, 16th International Euro-Par Conference, Ischia, Italy, August 31 - September 3, 2010, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 6271)*. Springer, 88–100.

[28] Gartner Inc. 2021. Gartner Says Cloud Will Be the Centerpiece of New Digital Experiences. Press Release. [Online; accessed September-2024].

[29] Frank Gens. 2019. Worldwide and Regional Public IT Cloud Services 2019–2023 Forecast. Tech. Rep. by IDC, Doc. #US44202119. [Online; accessed 9-Dec-2021].

[30] Fred Glover. 1989. Tabu Search—Part I. *ORSA Journal on Computing* 1, 3 (1989), 190–206. https://doi.org/10.1287/ijoc.1.3.190

[31] Haryadi S. Gunawi, Mingzhe Hao, Riza O. Suminto, Agung Laksono, Anang D. Satria, Jeffry Adityatama, and Kurnia J. Eliazar. 2016. Why Does the Cloud Stop Computing? Lessons from Hundreds of Service Outages. In *Proceedings of the Seventh ACM Symposium on Cloud Computing* (Santa Clara, CA, USA) *(SoCC '16)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/2987550.2987583

[32] Harshit Gupta, Amir Vahid Dastjerdi, Soumya K. Ghosh, and Rajkumar Buyya. 2017. iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, Edge and Fog computing environments. *Software: Practice and Experience* 47, 9 (2017), 1275–1296.

[33] Nikolas Herbst, André Bauer, Samuel Kounev, Giorgos Oikonomou, Erwin Van Eyk, George Kousiouris, Athanasia Evangelinou, Rouven Krebs, Tim Brecht, Cristina L. Abad, and Alexandru Iosup. 2018. Quantifying Cloud Performance and Dependability: Taxonomy, Metric Design, and Emerging Challenges. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* 3, 4 (2018), 19:1–19:36.

[34] Takahiro Hirofuchi, Adrien Lebre, and Laurent Pouilloux. 2018. SimGrid VM: Virtual Machine Support for a Simulation Framework of Distributed Systems. *IEEE Transactions on Cloud Computing* 6, 1 (2018), 221–234.

[35] Alexandru Iosup, Hui Li, Mathieu Jan, Shanny Anoep, Catalin Dumitrescu, Lex Wolters, and Dick H. J. Epema. 2008. The Grid Workloads Archive. *Future Generation Computer Systems* 24, 7 (2008), 672–686.

[36] Bahman Javadi, Derrick Kondo, Alexandru Iosup, and Dick Epema. 2013. The Failure Trace Archive: Enabling the comparison of failure measurements and models of distributed systems. *JPDC* 73, 8 (2013).

[37] Myeongjae Jeon, Shivaram Venkataraman, Amar Phanishayee, Junjie Qian, Wencong Xiao, and Fan Yang. 2019. Analysis of Large-Scale Multi-Tenant GPU Clusters for DNN Training Workloads. In *ATC*.

[38] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by Simulated Annealing. *Science* 220, 4598 (1983), 671–680. https://doi.org/10.1126/science.220.4598.671 arXiv:https://www.science.org/doi/pdf/10.1126/science.220.4598.671

[39] Younggyun Koh, Rob C. Knauerhase, Paul Brett, Mic Bowman, Zhihua Wen, and Calton Pu. 2007. An Analysis of Performance Interference Effects in Virtual Environments. In *2007 IEEE International Symposium on Performance Analysis of Systems and Software, April 25-27, 2007, San Jose, California, USA, Proceedings*. IEEE Computer Society, 200–209.

[40] Andreas Kohne, Marc Spohr, Lars Nagel, and Olaf Spinczyk. 2014. FederatedCloudSim: a SLA-aware federated cloud simulation framework. In *Proceedings of the 2nd International Workshop on CrossCloud Systems, CCB@Middleware 2014, Bordeaux, France, December 8, 2014*, Yehia Elkhatib and Stefan Walraven (Eds.). ACM, 3:1–3:5.

[41] I. Kunz, A. Schneider, and C. Banse. 2022. A Continuous Risk Assessment Methodology for Cloud Infrastructures. In *2022 22nd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. IEEE Computer Society, Los Alamitos, CA, USA, 1042–1051. https://doi.org/10.1109/CCGrid54584.2022.00127

[42] R.T. Marler and J.S. Arora. 2004. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization* 26 (2004), 369–395.

[43] Fabian Mastenbroek, Georgios Andreadis, Soufiane Jounaid, Wenchen Lai, Jacob Burley, Jaro Bosch, Erwin van Eyk, Laurens Versluis, Vincent van Beek, and Alexandru Iosup. 2021. OpenDC 2.0: Convenient Modeling and Simulation of Emerging Technologies in Cloud Datacenters. In *2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. 455–464. https://doi.org/10.1109/CCGrid51090.2021.00055

[44] Etienne Michon, Julien Gossa, Stéphane Genaud, Léo Unbekandt, and Vincent Kherbache. 2017. Schlouder: A broker for IaaS clouds. *Future Generation Computer Systems* 69 (2017), 11–23.

[45] Brad L. Miller and David E. Goldberg. 1995. Genetic Algorithms, Tournament Selection, and the Effects of Noise. *Complex Syst.* 9, 3 (1995).

[46] Melanie Mitchell. 1998. *An introduction to genetic algorithms*. MIT Press.

[47] Jeffrey C. Mogul and John Wilkes. 2019. Nines are Not Enough: Meaningful Metrics for Clouds. In *Proceedings of the Workshop on Hot Topics in Operating Systems, HotOS 2019, Bertinoro, Italy, May 13-15, 2019*. ACM, 136–141.

[48] Kay Ousterhout, Patrick Wendell, Matei Zaharia, and Ion Stoica. 2013. Sparrow: Distributed, Low Latency Scheduling. In *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles* (Farminton, Pennsylvania) *(SOSP '13)*. Association for Computing Machinery, New York, NY, USA, 69–84. https://doi.org/10.1145/2517349.2522716

[49] Katharine Ricke, Laurent Drouet, Ken Caldeira, and Massimo Tavoni. 2018. Country-level social cost of carbon. *Nature Climate Change* 8, 10 (01 Oct 2018), 895–900.

[50] Arman Shehabi, Sarah J Smith, Eric Masanet, and Jonathan Koomey. 2018. Data center growth in the United States: decoupling the demand for services from electricity use. *Environmental Research Letters* 13, 12 (Dec 2018).

[51] Siqi Shen, Vincent van Beek, and Alexandru Iosup. 2015. Statistical Characterization of Business-Critical Workloads Hosted in Cloud Datacenters. In *15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, CCGrid 2015, Shenzhen, China, May 4-7, 2015*. IEEE Computer Society, 465–474.

[52] Sacheendra Talluri, Leon Overweel, Laurens Versluis, Animesh Trivedi, and Alexandru Iosup. 2021. Empirical Characterization of User Reports about Cloud Failures. In *2021 IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS)*. 158–163. https://doi.org/10.1109/ACSOS52086.2021.00039

[53] Vincent van Beek, Jesse Donkervliet, Tim Hegeman, Stefan Hugtenburg, and Alexandru Iosup. 2015. Self-Expressive Management of Business-Critical Workloads in Virtualized Datacenters. *Computer* 48, 7 (2015), 46–54.

[54] Vincent van Beek, Giorgos Oikonomou, and Alexandru Iosup. 2019. A CPU Contention Predictor for Business-Critical Workloads in Cloud Datacenters. In *IEEE 4th International Workshops on Foundations and Applications of Self* Systems, FAS*W@SASO/ICCAC 2019, Umea, Sweden, June 16-20, 2019*. IEEE, 56–61.

[55] A. Vasan, A. Sivasubramaniam, V. Shimpi, T. Sivabalan, and R. Subbiah. 2010. Worth their watts? - an empirical study of datacenter servers. In *HPCA - 16 2010 The Sixteenth International Symposium on High-Performance Computer Architecture*. 1–10.

[56] Abhishek Verma, Luis Pedrosa, Madhukar Korupolu, David Oppenheimer, Eric Tune, and John Wilkes. 2015. Large-scale cluster management at Google with Borg. In *Proceedings of the Tenth European Con-*

ference on Computer Systems, EuroSys 2015, Bordeaux, France, April 21-24, 2015. ACM, 18:1–18:17.

[57] David Vose. 2008. *Risk analysis: a quantitative guide*. John Wiley & Sons.

[58] Bhathiya Wickremasinghe, Rodrigo N. Calheiros, and Rajkumar Buyya. 2010. CloudAnalyst: A CloudSim-Based Visual Modeller for Analysing Cloud Computing Environments and Applications. In *24th IEEE International Conference on Advanced Information Networking and Applications, AINA 2010, Perth, Australia, 20-13 April 2010*. IEEE Computer Society, 446–452.

[59] Franz Wilhelmstötter. 2021. *Jenetics Library User's Manual*. [Online; accessed September-2024].

[60] Yiting Xia, Ying Zhang, Zhizhen Zhong, Guanqing Yan, Chiunlin Lim, Satyajeet Singh Ahuja, Soshant Bali, Alexander Nikolaidis, Kimia Ghobadi, and Manya Ghobadi. 2021. A Social Network Under Social Distancing: Risk-Driven Backbone Management During COVID-19 and Beyond. In *18th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2021, April 12-14, 2021*, James Mickens and Renata Teixeira (Eds.). USENIX Association, 217–231.

[61] Chee Shin Yeo and Rajkumar Buyya. 2009. Integrated Risk Analysis for a Commercial Computing Service in Utility Computing. *Journal of Grid Computing* 7, 1 (2009), 1–24.