From Art to Science: Designing Resilient Topologies by Quantifying Network Performance Under Duress

A technical paper prepared for presentation at SCTE TechExpo24

Vaibhav Phatarpekar
Principal Software Development Engineer
Comcast
Vaibhav_phatarpekar@comcast.com

Bob Lutz
Senior Machine Learning Engineer
Comcast
Bob_lutz@comcast.com

Bala Ramachandran,
Senior Director, Network Engineering
Comcast
bala_ramachandran@cable.comcast.com

Cameron Brackmann,
Engineer 2, Software Development & Engineering
Comcast
cameron_brackmann@comcast.com

Table of Contents

List of Figures

List of Tables

1. Introduction

Reliable delivery of services is of utmost importance to Internet Service Providers (ISPs). Seamless delivery of traffic to customers depends on the performance of an ISP's critical infrastructure during adverse network events, such as fiber cuts or equipment failures. Assuming the physical (fiber) topology of a given network is fixed, the performance of the network under duress depends on the design of the logical (routing) topology, but how should an ISP decide among competing logical topology designs? This paper takes a quantitative approach to this question by introducing multiple ways to measure the resilience of logical topology designs.

2. Background

First, we establish some concepts and terminology used throughout the paper. Then we discuss the applications of the work.

2.1. Network Components

We consider two network layers: the *physical layer* and the *logical layer*. In this paper, the physical layer is considered fixed, while the logical layer is subject to design choices we wish to decide between. The physical layer consists of nodes representing sites that contain optical devices, and edges representing segments of fiber. The logical layer consists of nodes representing sites that contain logical devices, and edges representing logical circuits. In general, the logical layer nodes are a subset of the physical layer nodes; the logical layer edges need not be a subset of the physical layer edges, since logical circuits can exist between sites that are not connected by a fiber segment. Figure 1 shows an example of physical fiber topology and a logical topology that maps onto it.

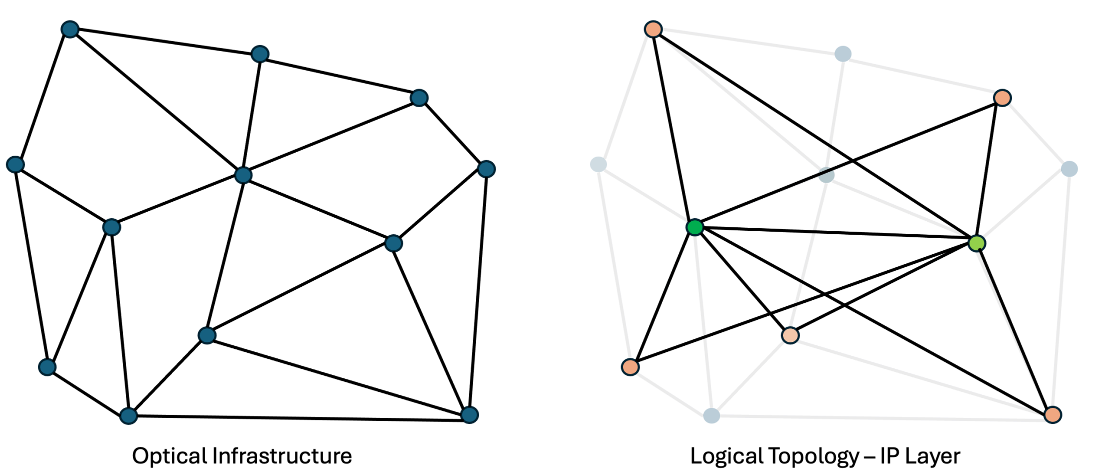

Figure 1 – The static fiber infrastructure (left) and one possible logical topology (right)

2.1.1. Fiber Infrastructure

Often referred to as the physical layer of a network, it represents the underlying infrastructure that facilitates the transmission of data. In a fiber-optic network, information is transmitted as pulses of light that travel through optical fibers made of glass or plastic since light is immune to electromagnetic interference. The transport system manages the transfer of information in this medium.

2.1.2. *Logical Topology*

The logical layer serves as a bridge between the fiber infrastructure and network services. This layer plays a crucial role in routing data, enabling interconnectivity, and ensuring the efficient flow of information within a network. Fiber infrastructure realities must be factored in when designing an efficient logical topology.

2.2. Failure Domain

To ensure uninterrupted delivery of services, protection against failure scenarios is factored into network designs. A *failure set* is a network component or set of components likely to experience concurrent downtime, which may lead to performance impairments. Examples of failure sets include "site down" events due to power outages, router failures, protocol events, etc. The collection of all failure sets is called the *failure domain*. Failure sets are determined based on historical trends and/or business requirements. For example, a new service agreement might require uninterrupted service in case of double fiber failures. In this scenario, all combinations of two fibers going down are considered as the failure domain and all analyses are factored around this failure domain.

Service providers typically protect their networks against isolation from single failures. The likelihood of concurrent multiple failures is low but nonzero, especially if certain failures take time to repair and other failures can occur in the area during the time of repair. Understanding the impact of two or more concurrent failures is helpful for the business to drive design decisions. Traditionally, network engineers have designed networks based on manually interpreting existing network maps, domain knowledge, consideration of geography, risks, etc. When designing for $N > 1$ concurrent failure, manual approaches become cumbersome. In such cases, it is common to leverage graph processing tools to analyze, suggest changes to, or completely restructure an existing topology. Another programmatic approach advantage is the ability to easily derive multiple competing topology designs. This gives rise to the question: how should we compare topology designs in terms of resiliency?

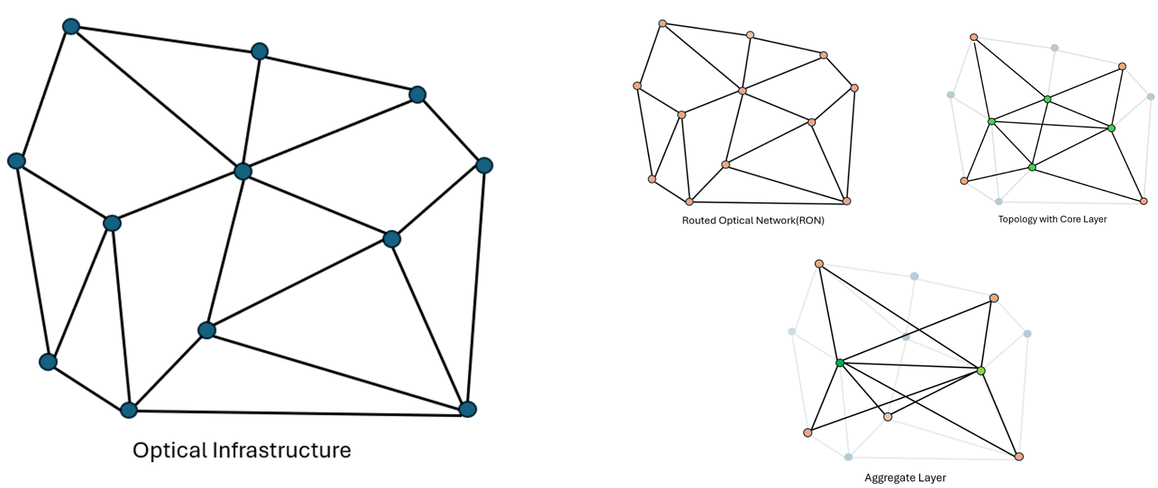

Figure 2 – Candidate logical topologies (right) to be mapped onto existing fiber infrastructure (left)

Figure 2 shows several logical topologies that could be mapped onto a fixed fiber topology. We draw particular attention to the upper-left logical topology, called the *routed optical network* (RON). This design, in which the logical sites and links exactly mimic those in the underlying fiber infrastructure, serves as the “gold standard” for resiliency against which all other designs are compared. However, the RON is often impossible to implement due to technological, operational, and cost constraints.

2.3. Applications

The methods in this paper can be applied to the following problems facing network operators:

- **Comparing competing topologies:** Multiple competing topologies, derived following programmatic approaches, were compared to finalize network designs.
- **Adjustments to existing or candidate topology:** Network design is an iterative process. The general process is to propose an initial design which is then fine-tuned to incorporate feedback from various teams. Quantifying methodologies discussed in this paper are applied successfully to identify optimal changes ensuring minimum impact on gains of initial design. This allows for a more data-driven approach to network design:
    - **Reducing the number of degrees per site at the logical layer:** This was a critical component in reducing the overall cost of network design and ensuring minimal resiliency impact.
    - **Core site replacements:** A core site carries the entire network’s traffic which in turn translates to greater power and space requirements. Feedback can indicate higher costs associated with promoting a candidate site as a core site. Impact volume through exclusion or replacement of such a site assists with determining cost-to-benefit trade-off.
    - **Identify targeted shared risk link group (SRLG) fixes:** The optimal number of SRLG fixes to maintain resiliency was determined where certain design choices were not feasible.
    - **Optimizing the number of logical circuits going over certain fiber links:** Decisions related to designing the optimal number of optimization of circuits to exclude certain fiber links could be made which helped wavelength thresholds.
- **Greenfield deployment:** When designing networks from scratch, competing topology performances can be compared under the assumption that every failure is equally likely, as historical data on failures is not available.
- **Network expansions:** The impact of design choice relative to new fiber construction or a logical circuit can be evaluated for network expansion.

3. Quantifying Resiliency

A key step for any analysis is to quantify gains associated with a choice. A quantifying methodology that condenses all relevant data into a single score is ideal for comparisons, but in most cases requires a certain amount of data entropy. This section discusses possible indicators for determining the volume of impact and a framework that captures resiliency information.

3.1. Indicators for Volume of Impact

To measure the resiliency of a design, we must establish a metric of *volume of impact*, i.e., the severity of a given failure scenario. A straightforward measurement of volume of impact is the number of sites isolated by a failure scenario. However, some sites consume services far more heavily than others, making this metric potentially highly skewed. Ideally, metrics for volume of impact should correlate with service consumption and specify existing and prospective impacts. These metrics can be employed individually or in tandem to provide a complete profile of the resiliency of a given design.

Subscriber count (SC) indicates the actual number of customers serviced in an area. A failure scenario can be measured by the number of subscribers that become isolated. While this might be a good indicator of the *current* impact of a failure scenario, SC does not consider that customer penetration rates may be different in the future. One alternative is to leverage homes passed (HP). HP refers to the number of homes within a service area, regardless of whether those homes contain current subscribers. It is a forward-looking indicator since it also indicates potential impact. Alternatively, the total volume of traffic can also be leveraged as an indicator. Since the volume of traffic is time-dependent, some statistical functions like the 98^{th} percentile can be used to derive conservative estimates.

3.2. Introduction to Failure Profile

The volume of impact is measured for each failure set. However, when comparing topologies, one must compare performance against the entire failure domain, not just a failure set. A topology with a lower probability of high-impact failures is considered better. However, it is important to note that lower-impact failures also engage resources and have associated costs. All failures, including those that do not isolate any subscribers, engage resources for resolution and have associated costs. A quantifying framework should therefore capture how a topology fares in both higher- and lower-impact failures. To address this need, we introduce the *failure profile* (FP) as a framework that comprehensively compares topologies.

Households Isolated	Number of Failures
No Isolations	N1
1-10000	N2
10000-20000	N3
20000-30000	N4
30000-40000	N5
40000-50000	N6
.	.
.	.
.	.
.	.
.	.
200000-300000	N12

Subscriber Count Isolations	Number of Failures
No Isolations	N1
1-10000	N2
10000-20000	N3
20000-30000	N4
30000-40000	N5
40000-50000	N6
.	.
.	.
.	.
.	.
.	.
200000-300000	N12

Total Traffic Volume Isolated	Number of Failures
No Isolations	N1
1-10000	N2
10000-20000	N3
20000-30000	N4
30000-40000	N5
40000-50000	N6
.	.
.	.
.	.
.	.
.	.
200000-300000	N12

Figure 3 – Example failure profile with dummy values

Figure 3 shows an example failure profile. Note that the FP splits failure domain impact volumes into distinct categories. Categories are decided based on how an organization defines severity. This implies that the isolation of 100K customers can be placed in the same category as the isolation of 199K customers if both are considered equally severe. Typically, we will split FP into three categories:

1. Failure sets that do not cause isolations
2. Failure sets that cause isolations but are not considered as high-severity
3. Failure sets that cause isolations associated with high severity.

Some topology designs fare better than others in high-severity failures but worse in low-severity failures. Unbalanced scenarios like this are articulated well in the FP view. The FP can also be enriched by adding more columns to incorporate customized business logic. These columns include the total length of unique fibers, number of aerial fibers, number of leased fibers, etc. In later sections, we discuss how FP is leveraged to determine the probability of being in a severity category and for the derivation of a single resiliency score.

However, FP has certain shortcomings. Even though FP indicates the set of failures, it does not provide any indication of the probability of failures. Since it is not a single score, manual intervention is needed to compare the topologies and hence cannot be leveraged by automation tools or by any brute-force methods of sifting through topologies.

4. Probabilistic Comparison Methods

Just knowing the volume of impact and corresponding failures might not be sufficient to undertake critical tradeoff decisions. Associating probability values with failures helps measure the gain associated with a particular design choice.

Ideally, these probabilities are based on topological and operational data, including failure records for every fiber segment and SRLG. However, data related to fiber topology can be limited. This limitation can be associated with network expansions or acquisitions and the ever-changing nature of the fiber layout, making it difficult to track changes. This section dives into the methodologies to derive scores measuring the resiliency of a topology design. The methodologies listed also indicate how the topologies can be compared even in the absence of empirical records.

4.1. Resiliency Score & Modeling Failure Probabilities

In this section, we will explain how to formally model and measure the average-case behavior of failures in each network design. Suppose that we are given a data representation of all physical links in the network and their locations, such as from a spatial database. An SRLG is a set of links that tend to be disrupted or severed as a single unit.

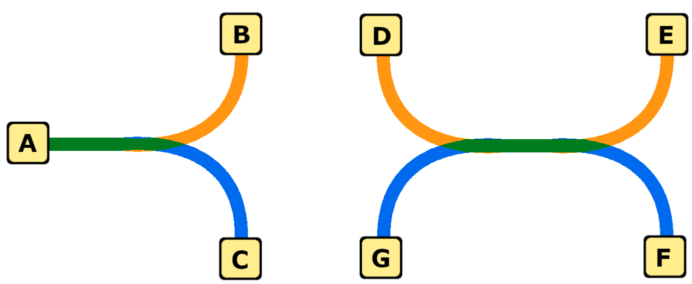

Figure 4 – Two flavors of SRLGs

Figure 4 shows the two main flavors of SRLGs. The first, on the left, occurs when multiple fibers coming out of the same site are close together over some distance. The second, on the right, occurs when multiple fibers meet somewhere along their lengths and remain close together over some distance.

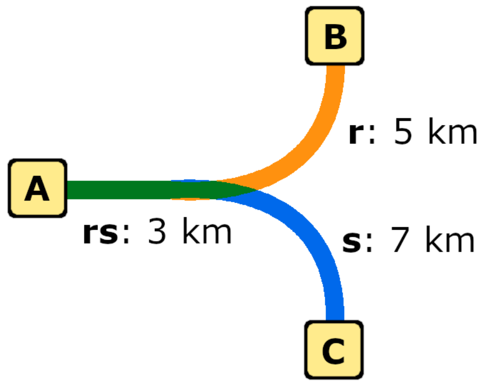

Figure 5 – A small fiber topology

A small fiber topology example is pictured in Figure 5. Here there are three sites, called A, B, and C; two fiber segments, called r and s; and one SRLG, called rs.

4.1.1. Defining the Resiliency Score

We now construct a probability space based on the underlying fiber infrastructure and use it to measure the resiliency of a given logical topology design. We define a *failure possibility* in a fiber topology as a choice of either a single fiber segment or an SRLG. This corresponds with the notion of a *failure set* from Section 2. Thus, the set of failure possibilities for the example topology is

$$\Omega = \{r, s, rs\}.$$

This set corresponds with the notion of the *failure domain* from Section 2. Here we think of the failure possibility r as representing a fiber cut somewhere along the orange segment (excluding the green segment). The failure possibility rs represents a fiber cut somewhere along the green segment.

We will construct a probability space for which the set Ω of all failure possibilities is the sample space. The relevant event space will be the *power set* of failure possibilities, i.e. the set of all subsets of failure possibilities. In the running example, the event space is

$$\mathcal{F} = \{\emptyset, \{r\}, \{s\}, \{rs\}, \{r, s\}, \{r, rs\}, \{s, rs\}, \{r, s, rs\}\},$$

where $\emptyset$ denotes the empty set. We think of an event as representing one or more concurrent fiber cuts. For example, $\{r, s\}$ represents a cut in the orange segment of fiber and another cut in the blue segment.

It remains to assign probabilities P to the events above. Ideally, these will be derived from historical data. For example, operational records of fiber cuts including location and duration could be combined with a data representation of all fiber segments and SRLGs to construct a sampling distribution on the set of events. Even in this approach, however, some modeling is necessary; only events that have occurred in the records will receive nonzero probabilities, leaving some events unaccounted for. We will return to the subject of modeling probabilities in a moment.

We now use the probability space $(\Omega, \mathcal{F}, P)$ to measure the resiliency of a given logical topology design. Let λ denote a measure of *impact* for each event in the event space $\mathcal{F}$. This corresponds to the notion of *volume of impact* from Section 3. For example, $\lambda(f)$ could be the number of subscribers or households isolated when the fiber cuts represented by the event $f \in \mathcal{F}$ occur. Note that the impact λ depends on the particular logical topology design, while the probability space $(\Omega, \mathcal{F}, P)$ depends only on the (fixed) underlying fiber topology.

The *resiliency score* (or *expected impact*) of a logical topology design is

$$R = \sum_{f \in \mathcal{F}} \lambda(f) * P(f).$$

That is, R equals the sum of the impact of each event times the probability of the event, taken over all events in the event space. In other words, R is the expected value $E[\lambda(f)]$ of the impact function λ over the event space.

The resiliency score R measures the average-case impact of a failure event in the network. Thus, a design with a lower score is favored in general over one with a higher score, relative to the chosen impact measure λ. Different impact metrics can prioritize different aspects of a design and therefore result in different design rankings. For example, λ can be chosen to measure

- Number of subscribers or households isolated
- Same as above, but including *homes passed* to account for potential future subscribers
- Amount of traffic isolated
- Indication of a certain type of outcome (e.g. $\lambda(f) = 1$ if f isolates a certain amount of commercial traffic, $\lambda(f) = 0$ otherwise).

4.1.2. Modeling Probabilities via Exposure Length

Historical data on fiber cuts is not necessarily available or usable. Even in the presence of high-quality data, some amount of modeling is likely needed to represent failure scenarios that have not occurred in the timeframe of the data collection. For example, if data collection has persisted for only a year, and a certain fiber segment has not been cut during that year, then the historical data will (inaccurately) suggest that the probability of the fiber being cut is zero. We need a way to estimate the probability of the unrepresented fiber cut. To this end, we will use the *exposure length* of fiber segments and SRLGs to model probabilities on the event space $\mathcal{F}$. This model works on the principle that longer fibers are more likely to be cut because they are more exposed to outside elements.

Recall the sample space of failure possibilities defined in Section 4.1.1. For the running example from Figure 5, the set of failure possibilities is $\Omega = \{r, s, rs\}$. To each failure possibility we assign the length (in km) of fiber where a cut would result in exactly that possibility failing and no others. For example, the exposure length of r is $\ell(r) = 5$ km; the exposure length of s is $\ell(s) = 7$ km; and $\ell(rs) = 3$ km. To convert these into probabilities on the sample space Ω, we can divide each by the sum of the exposure lengths: $p(r) = 5/15 = 0.33$, $p(s) = 7/15 = 0.47$ and $p(rs) = 3/15 = 0.2$.

This gives us a probability distribution on the sample space Ω, but we seek a probability distribution on the event space $\mathcal{F}$ to compute the resiliency score R. While it is possible to define in full generality, we will make several simplifying assumptions. First, we assume that the probability of three or more concurrent failures is vanishingly low, i.e. $P(f) = 0$ for any such event f. Second, we assume that the physical and logical topologies are designed so that a single fiber failing cannot cause any isolations, i.e. $\lambda(f) = 0$ for any event f that results in a single fiber failing and most choices of impact measure λ. Thus, it remains to define probabilities for the events causable by exactly two concurrent fiber cuts. We will call this subset of events $\mathcal{F}_2$. Third, we assume that all fiber cuts are statistically independent.

In the running example, we have

$$\mathcal{F}_2 = \{\{r\}, \{s\}, \{rs\}, \{r, s\}, \{r, rs\}, \{s, rs\}\}.$$

The single-element events $\{r\}$, $\{s\}$ and $\{rs\}$ are included because both fiber cuts can occur on the same segment, e.g. both in the orange segment for the event $\{r\}$. In general, events in $\mathcal{F}_2$ consist of at most two failure possibilities.

We obtain slightly different formulas for $P(f)$ depending on the number of failure possibilities in f:

$$P(f) = \begin{cases} p(\alpha)^2 & \text{if } f = \{\alpha\} \\ 2 * p(\alpha) * p(\beta) & \text{if } f = \{\alpha, \beta\}, \end{cases}$$

where $p(\alpha)$ and $p(\beta)$ are the probabilities defined on the sample space Ω above. Here we are technically computing a conditional probability; the formula for $P(f)$ gives the probability of f occurring *given that exactly two concurrent fiber cuts occur*. This does not materially affect the result, since all such conditional probabilities differ from their unconditional counterparts by a global scalar (the probability of exactly two concurrent fiber cuts occurring).

Returning to the running example, we obtain the following probabilities on the event space $\mathcal{F}_2$:

$$P(\{r\}) = p(r)^2 = \left(\frac{5}{15}\right)^2 = 0.11$$
$$P(\{s\}) = p(s)^2 = \left(\frac{7}{15}\right)^2 = 0.22$$
$$P(\{rs\}) = p(rs)^2 = \left(\frac{3}{15}\right)^2 = 0.04$$
$$P(\{r, s\}) = 2 * p(r) * p(s) = 2 * \frac{5}{15} * \frac{7}{15} = 0.31$$
$$P(\{r, rs\}) = 2 * p(r) * p(rs) = 2 * \frac{5}{15} * \frac{3}{15} = 0.13$$
$$P(\{s, rs\}) = 2 * p(s) * p(rs) = 2 * \frac{7}{15} * \frac{3}{15} = 0.19,$$

where we have rounded to the nearest hundredth. Note that the probabilities indeed sum to 1.

4.1.3. Computing the Resiliency Score

It remains to choose an impact measure λ and compute the resiliency score R for a given design. Suppose that $\lambda(f)$ counts the number of subscribers (in thousands) isolated by event f. Per our earlier assumption, we have $\lambda(f) = 0$ for any event f causing only a single fiber to fail, so $\lambda(\{r\}) = \lambda(\{s\}) = 0$. Suppose that after an analysis of our topology design, we determine that

$$\lambda(\{rs\}) = \lambda(\{r, s\}) = \lambda(\{r, rs\}) = \lambda(\{s, rs\}) = 10.$$

It makes sense that all these values are equal, since all four events result in both r and s being cut. We then have

$$R = \sum_{f \in \mathcal{F}_2} \lambda(f) * P(f) = 0 * 0.11 + 0 * 0.22 + 10 * 0.04 + 10 * 0.31 + 10 * 0.13 + 10 * 0.19 = 6.7.$$

Thus, on average, we expect two concurrent fiber cuts in the example network to isolate 6,700 subscribers. In terms of number of subscribers isolated, designs with $R > 6.7$ perform worse on average than the current design, and designs with $R < 6.7$ perform better.

4.1.4. Comparing Multiple Topology Designs

We will briefly discuss the decision process when comparing candidate topology designs. Recall the routed optical network (RON) defined in Section 2.2. The RON serves as the most resilient logical topology design. Under normal circumstances, and with appropriate choices of impact metric λ, the RON will achieve the lowest possible score. However, it is often impossible or impractical to implement, so it is used primarily as a benchmark against which other designs are measured.

Table 1 – Comparing raw resiliency scores

	Baseline topology	Candidate topology 1	Candidate topology 2	RON
Resiliency score	9.15	8.58	8.50	8.32
Scaled score	0.00	0.78	0.88	1.00

Suppose that our goal is to improve upon an existing topology design, and we are given two candidate topologies whose scores are recorded in Table 1 alongside the current (baseline) design and the RON. We see that the lower-scoring candidate topology is #2, meaning it is the more resilient candidate on average for the chosen impact metric. To emphasize the gap in score between the two candidates, we can use a "scaled" version of the score, also shown in the table:

$$\text{scaled candidate score} = \frac{\text{baseline score} - \text{raw candidate score}}{\text{baseline score} - \text{RON score}}.$$

This scaled score lies between 0 and 1 and can be more easily interpreted as a percentage.

When more granularity is needed, the probability framework can also be used to determine the probability that a random failure event falls into a certain category, e.g. low vs. high severity. This gives a "normalized" version of the failure profile (FP) from Section 3.2, where the likelihood of each failure event is considered, rather than treating all events uniformly.

Table 2 – Probabilities of different failure event categories

	Baseline topology	Candidate topology 1	Candidate topology 2	RON
No isolation	0.1	0.5	0.4	0.6
Low severity	0.4	0.3	0.5	0.3
SEV 1	0.5	0.2	0.1	0.1

Table 2 shows the probabilities of each event category for the four designs from Table 1. In the notation of Section 4.1.1, each entry in the table is simply the sum of the probabilities $P(f)$ over all events f in the listed category. For example, the sum of probabilities of all low-severity events for the baseline topology design is 0.4. In other words, the probability that a random failure event is low severity is 40%.

Note that while candidate topology 2 performs better than candidate 1 in terms of resiliency score based on the chosen impact metric, it does not perform better in all event categories; candidate 1 outperforms candidate 2 in the category of low-severity events. This illustrates the importance of choosing an appropriate impact metric; a metric that emphasizes low severity events might give a resiliency score by which candidate 1 outperforms candidate 2. Choosing an impact metric is a critical and potentially subtle step in the decision process. The best metric is one that most closely reflects the overriding business priorities.

4.2. Monte Carlo Simulation for comparing topologies

For certain scenarios, the formula-based approach described above can be intricate, and a simulation-based approach might enable quicker prototyping and yield results that are sufficiently definitive to be used in applications. This could be the case, for example, when considering more than two concurrent failures at once; when analyzing failure of logical components in addition to physical components; or when considering additional time-based factors like network availability. Monte Carlo simulation (MCS) provides a convenient and flexible interface to estimate scores in these cases.

For a failure domain with pre-defined failure sets, graph processing tools are leveraged to precompute impact for each failure set. Empirical records are leveraged to derive probability density functions for failure sets. In the absence of a probability density function, each failure set in a failure domain is considered equally likely. For analysis, where the same failure domain applies to all the topologies, MCS iterations can be carried out simultaneously for each topology. This indicates performance differences when the same combination of failures is carried out for each topology. If failure domains differ between topologies, for example: if logical layer choices are different, MCS iteration can be run individually on each topology to derive metrics for each topology which can then be compared. MCS provides granularity to run iterations to derive a single score or to identify performance against severity buckets associated with FP. Figure 6 indicates the setup for MCS.

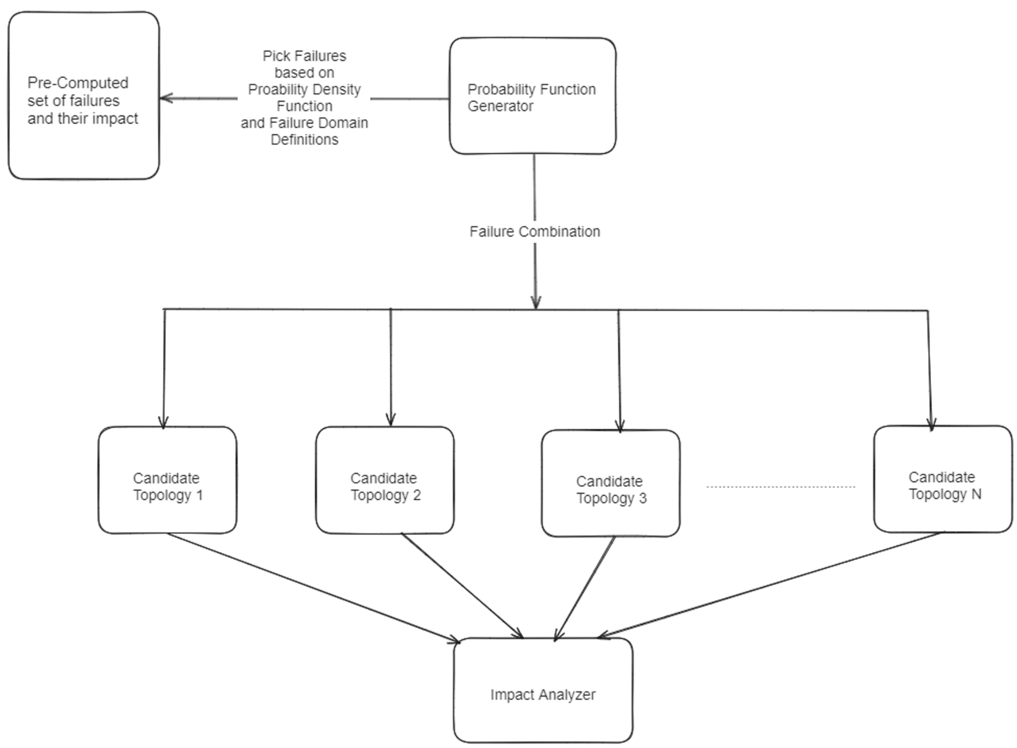

Figure 6 – Monte Carlo simulation flow chart

A *probability function generator* (PFG) picks a failure combination based on defined failure domains and a defined *probability density function* (PDF). In the absence of a probability density function, uniform distribution is leveraged to simulate failures. This failure combination is then applied on the topologies and the results are fed to an impact analyzer which aggregates all the data to generate a report. One failure combination is considered as one MCS iteration. Precomputation enables carrying out a substantial number of iterations in a short duration. Dimensions defined for the report can be based on business requirements. The probability of a core site going down, the probability of a critical site going down, the probability of high severity (SEV1) failures, or a combination of the dimensions can be considered. A table like the one indicated in section 4.1.4 can be built using the below-listed formula:

$$P(\text{Failure falls into certain category}) = \frac{\#\ \text{Failure simulations in the category}}{\text{Total}\ \#\ \text{MC simulations}}$$

5. Future Enhancements

The probability methodologies in Section 4 formally model and measure average-case behavior for a single failure scenario. To enrich this analysis, we can also consider the behavior of events over time. To this end, we can compute *availability*, which considers the amount of time the network is operational ("uptime") in each window. It is given by the following formula

$$\text{Expected availability} = \frac{\text{Expected uptime}}{\text{Total time of the observation window}}$$

If we know the expected number of failures in a year, for example, and the mean time to repair (MTR) for every failure scenario, it is possible to compute expected availability numbers for different topology designs. As with the probabilistic approach, it must be stressed that this computation would indicate *average-case* behavior over the course of a year and is not meant to capture the range of behaviors of the network for use in provisioning or planning.

6. Conclusion

In this paper, we have presented data- and model-driven methodologies to quantify and compare candidate logical topologies. The methodologies provide both formula- and simulation-based approaches to provide a well-rounded perspective of resiliency analysis. This framework also indicates different data points that can be collected to derive probability values for individual failures. All the methodologies discussed in the paper have a modular architecture making it easy to customize based on proprietary scenarios.

Abbreviations

CM	cable modems
FP	failure profile
HP	homes passed
IP	internet protocol
ISP	internet service provider
MCS	monte carl simulation
MTR	mean time to repair
PDF	probability density function
PFG	probability function generator
RON	routed optical network
SC	subscriber count
SRLG	shared risk link group

