Protected Optical Circuits and SLA Economics

A Study of Decision Criteria for Enterprise Telecommunications

Dan M. Kalin
Sr. Consultant, Telecommunications Practice

Arbitor, Inc.

Summary:

In the telecommunications marketplace, a potential customer will often hear providers talk of protected circuits or circuit protection in order to support an aggressive service level agreement for a higher cost point than unprotected. In contrast to that are providers which offer an aggressive service level agreement with an unprotected offering. This treatment explores the underlying rationale of these issues and promotes the use of a "total cost of failure" economic analysis when making decisions between several different types of provider offering.

The general conclusion is that most companies would realize greater organizational efficiencies, by avoiding outages, than the marginal premium associated with choosing a protected circuit service over an unprotected one. Secondary conclusions are that provider Service Level Agreement (SLA) evaluations should be viewed as indications of the provider's confidence in their ability to provide service, i.e. how much of their service revenue is placed at risk pending compliance with their guarantees?

Background:

Perhaps we should take some initial time to define the term "protection" as it being used in the context of this paper. In concept there are three different levels of protection for dedicated circuits, protected with geographic diversity, protected without geographic diversity, and unprotected.

Protected With Geographic Diversity:

Figure 1 depicts the classic Synchronous Optical Network (SONET) implementation of a protected optical circuit, be it any one of a number of possible increments.

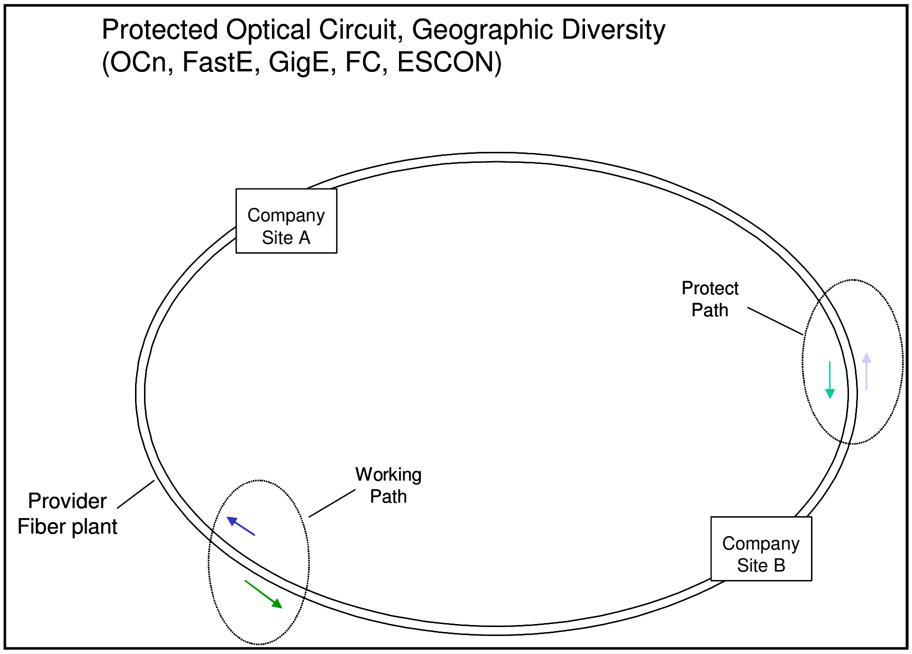

Figure 1: Protected Optical Circuits

Between the two company sites is a physical pair of fiber rings, one for the working path and one for the protect path (Note that the rings are shown as diverse, meaning that the fiber path is not wholly contained in one conduit or spur). That allows for the circuit to operate a primary working path when the system is fully operational. Note that traffic proceeds one direction on each strand, allowing for a mirrored circuit which is in perpetual readiness to substitute should the primary working circuit experience a failure. In this case the working channel direction is shown with dark blue or green arrows traveling in opposite directions. The corresponding protect paths are shown with light blue and green arrows also in opposite directions. When data is presented to the transport network, it is duplicated and sent both directions. The receiving station receives both signals but attends to the working path signal until such time as a failure and within a very short interval switches the protect path into working mode. The industry standard for the switching latency is less than 50 milliseconds, and is often much less. The idea is to provide continuous service (less the switching latency) to where the user traffic between sites is unaffected by a fiber-cut type failure. Figure 2 shows the effect of a fiber cut on data traffic between the two sites.

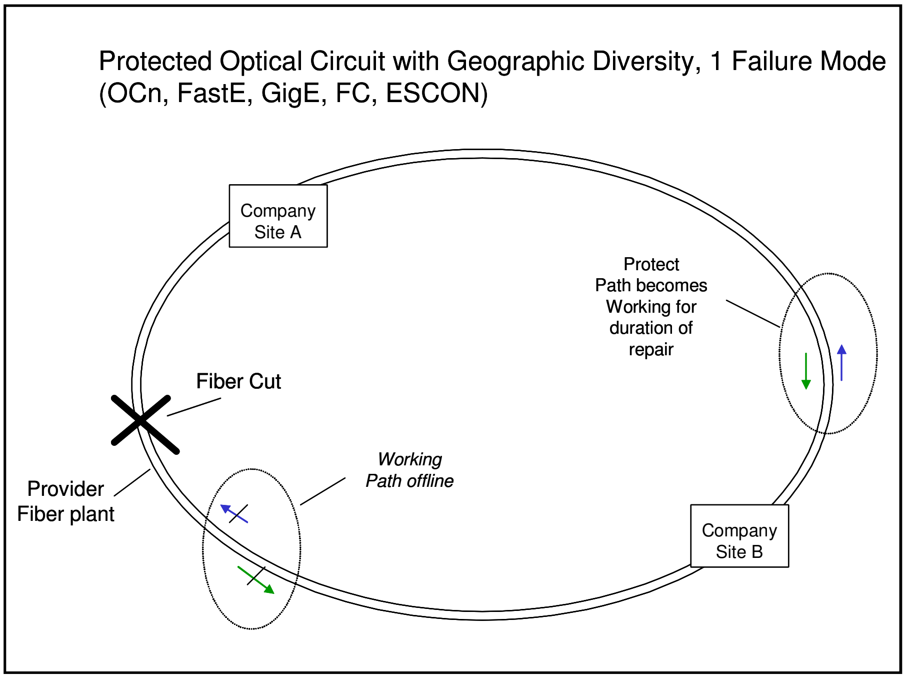

Figure 2: Protected Optical Circuits under Failure

Providers will typically charge approximately a 50-100% premium for a protected service that operates in this fashion. Some providers now offer a “shared” protection plan which allows them to oversubscribe the protect paths and realize lower capital costs. So far, these have not translated into significant price reductions to the customer and will not be covered in this discussion.

In terms of underlying physical reliability, this approach is the most reliable and necessary to truly support a SLA claim of 99.999% availability.

Protected Without Geographic Diversity:

Sometimes a provider will offer a protected service without geographic diversity, sometimes called a “spur” implementation. The reasons for this usually involve the cost of laying fiber along a diverse route, the incremental cost to lay one set of fiber or hundreds along the same path is minimal, as the primary cost is based on linear feet of conduit. So in some instances where a customer site is removed from the fiber plant of the provider, the least capital cost approach for a build is the extruded spur. Figure 3 depicts this approach simplified to show just the spur element of the circuit. Note that the underlying hardware is the same as geographically diverse, except that the protect path rides within the same conduit as the working path.

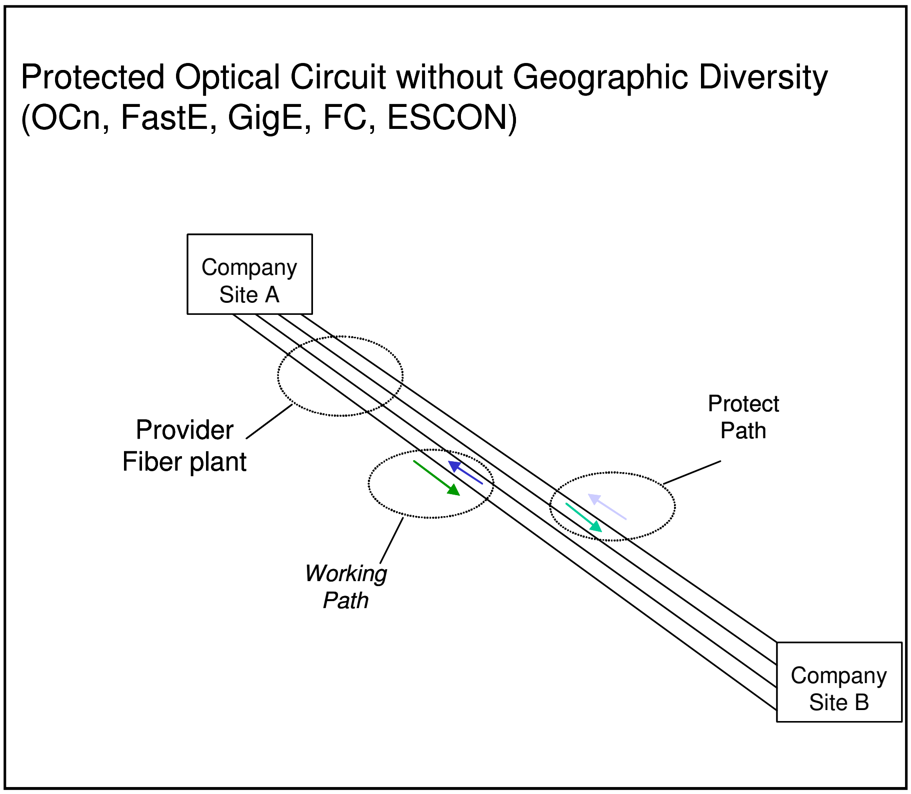

Figure 3: Protected Optical Circuits Without Geographic Diversity

This provides a measure of protection in terms of transceivers and optronics, but does not protect against a fiber cut as shown in Figure 4 below. If the conduit is ruptured by street work, the entire circuit provision is disrupted and non-functional. These types of outages are fairly severe as well, so a company can count on being out of service for many hours with this approach. Fiber cuts are unlikely, but do happen and should be considered when implementing a facility of this type.

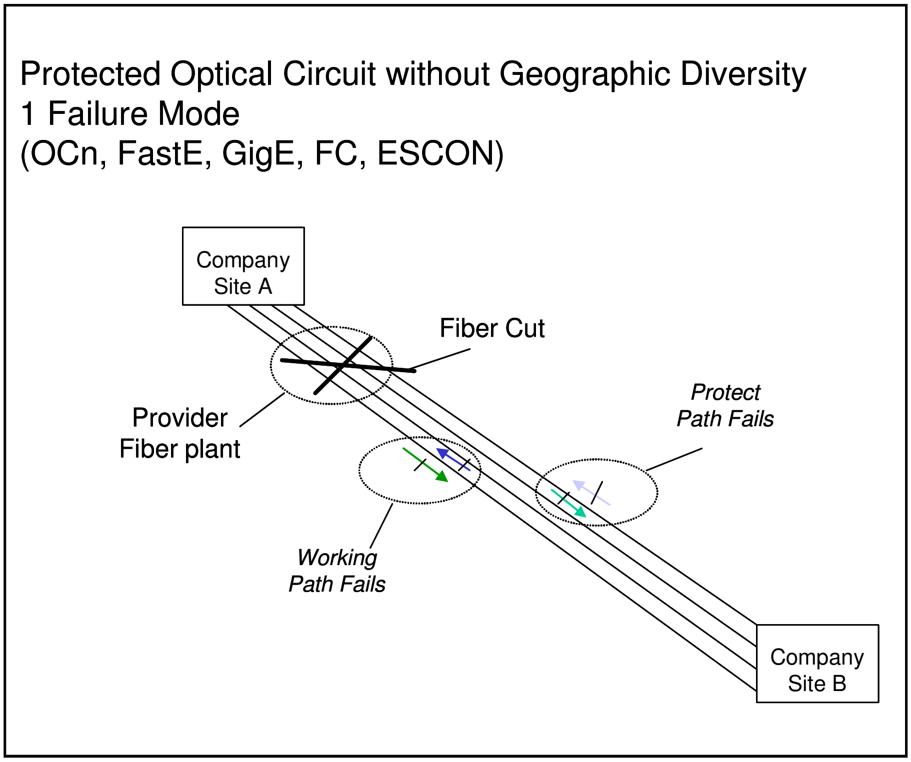

Figure 4: Protected Optical Circuit Failure Without Geographic Diversity

In terms of protection, this kind of approach could reasonably justify a SLA availability metric of 99.99%, as this would enable the provider to do routine maintenance on the circuits without affecting the service. In practice this service is often not differentiated from the fully diverse approach, in terms of pricing, and may not be disclosed to the customer at all without a direct inquiry from the customer.

Unprotected:

Unprotected diversity is usually deployed as a point-to-point implementation and priced to minimize the carrier's capital expense for optronics. Some implementations are actually less fiber-efficient due to stranding fiber plant. Many times that makes financial sense to a provider because fiber is generally a long term capital asset which depreciates over 20 years (often already a sunk expense) while optronics depreciate in less than 5 years and represent new capital spending. By offering unprotected services, the underlying bandwidth capacity of the optronic equipment is twice what it would be offering the same circuits as protected.[1] There are a series of additional issues, like fiber mileage between sites, which could also enter in to this model but they are not significant for the purposes of this discussion and do not affect the conclusions of the analysis.

Figure 5 shows a simplistic diagram of the physical layout of an unprotected circuit. Note that this could be laid in a ring, or series of rings, configuration as well.

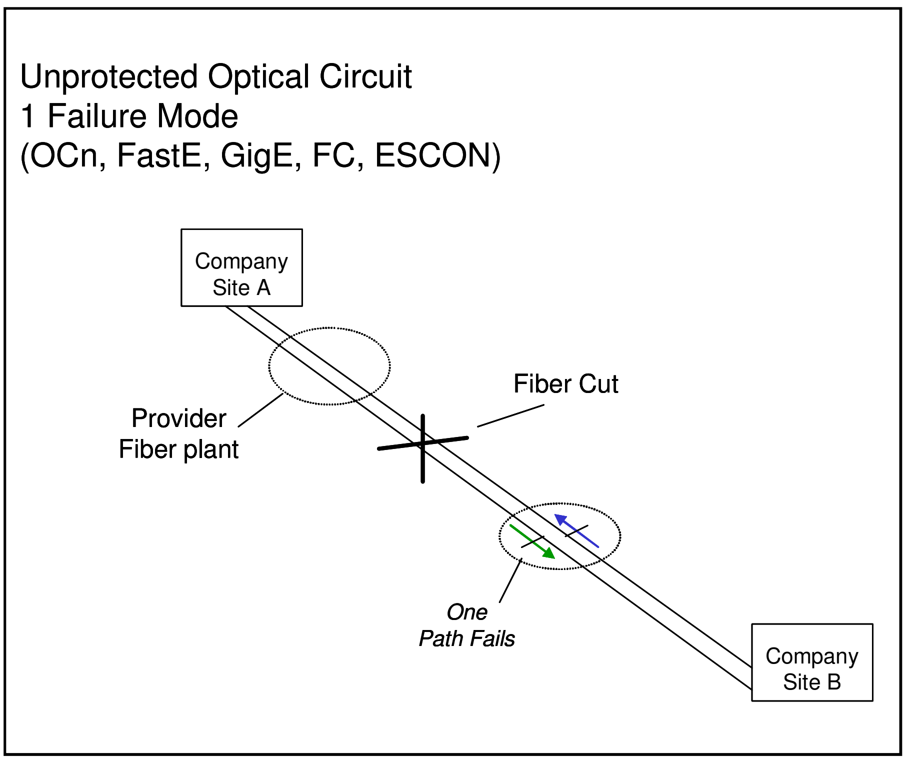

Figure 5: Unprotected Optical Circuit Failure

In the example of a fiber cut, this type of implementation completely fails, and as discussed above, the customer will experience hours of outage. In addition to

[1] While capacity is doubled with an unprotected approach the underlying transport equipment capital cost for protected circuits is only 20-30% more expensive than unprotected on a per-circuit basis. If a provider offers unprotected circuits at 50% the cost of protected, they realize a lower gross profit margin on the unprotected circuits.

the issue of a fiber cut is the issue of optronics failure. When a component in the system fails, the entire circuit fails. Adding to this are the actions of the providers themselves, performing routine maintenance and adding additional circuits to the network. In those circumstances, the customer will often suffer degraded service or outages. Most provider contracts for unprotected services call for "maintenance windows" during which the commitments contained within the applicable SLA do not apply. That is fairly important if a company needs the facility 24/7, as a maintenance window can be an ugly surprise all too often.

Excluding the maintenance windows, this type of implementation can reasonably justify an availability of 99.9%. Pricing for the service ranges from 50% to 70% the price of a protected service.

Service Level Agreements:

At this point the reader pipes up that they have been offered unprotected services with availability commitments of 99.999%, given the above how can that be? In a word, marketing.

It might help to examine what the different levels of availability translate into in terms of minutes per month.[2] Table 1 shows the commonly quoted availability commitments.

Reliability %	< Outage Minutes/Month
98.0	876.6 (14.6 hours)
99.9	43.83
99.99	4.383
99.999	0.438

Table 1: SLA Outage Minutes

Typically, a service provider is making the commitment that the circuit, as defined in the contract, will be down less than the above figures. So a commitment which

[2] We will assume months defined as 30.4375 days per month (including leap year).

exceeds 99.9% doesn't give the service provider time to fix any problems which occur, usually it would take longer than 5 minutes to understand which component in the system had failed. Therefore, either the systems are very reliable, which means aggregate rare failures can be managed when they occur, or the carriers fully expect to fail and structure SLA credits accordingly. Unfortunately, the latter is closer to reality.

The below SLA credits curve in Figure 6 is that of a service provider (Provider A) who provides unprotected OCn and Gigabit Ethernet circuits.

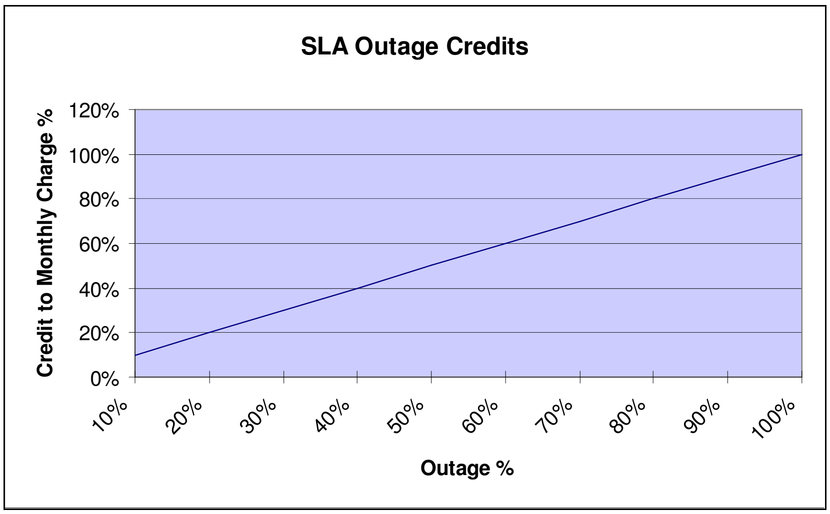

Figure 6: SLA Credits for Provider A

Provider A markets their unprotected service at a total monthly recurring charge of $3,000. Their marketing literature states that the service availability is 99.999%. Their SLA provides the above credit schedule for outages and the customer must request a credit in order to receive one.

Provider B markets their protected (diverse) service at a total monthly recurring charge of $6,500. Their marketing literature also states that the service availability is 99.999%. Their SLA provides the credit schedule in Figure 7 for each outage.[3]

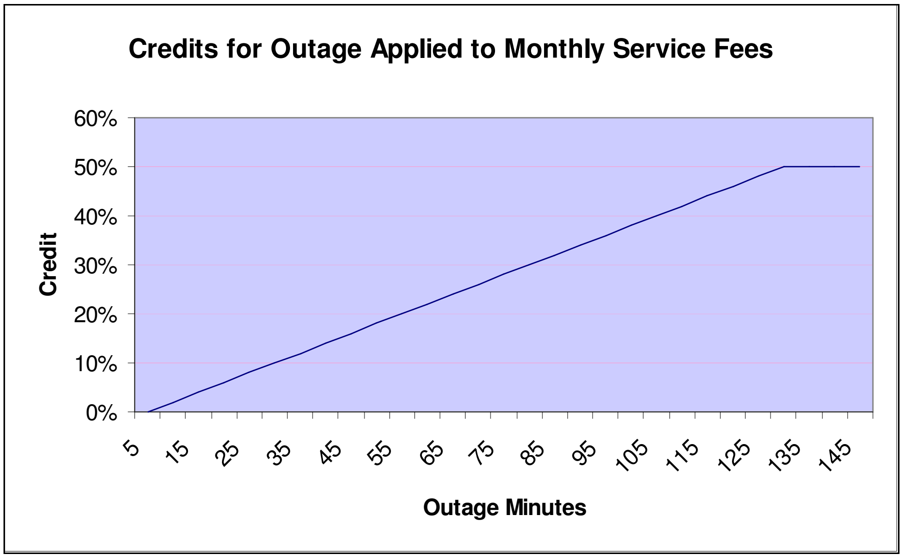

Figure 7: SLA Credits for Provider B

Let's compare the two offerings against the common SLA metrics in Table 2.

Reliability %	< Outage Minutes/Month	Provider A Credit %	Provider B Credit %
98.0	876.6 (14.6 hours)	2.0%	50%[4]
99.9	43.83	0.1%	16%
99.99	4.383	0.01%	0%
99.999	0.438	0.001%	0%

Table 2: Service Provider SLA Credit Comparison

Service provider B reaches the maximum credit per outage (50%) at 130 minutes while Service Provider A reaches the maximum credit only if they provide a full month's outage. Given the pricing levels above and a 45 minute failure, that translates into credits of $3.08 and $1,040.00 for Providers A and B respectively. At the 1 minute of outage (fails 99.999% availability) level it translates into $0.03

[3] Provider B limits the total credit per outage to 50%, but the total monthly credit cannot exceed 100% in the event of more than 2 outages.
[4] Assumes only one outage.

and $0.00 respectively. Clearly the service providers do not expect to provide this level of service, otherwise they could safely place a larger % of their revenue at stake pending performance. Provider A clearly doesn't believe their unprotected service offering will deliver 99.999% availability.

Clearly, the provider SLA commitments available in the marketplace do not reflect the actual performance of their services and should not be relied upon as a key variable to differentiate providers. The whole issue of comparative SLA credits is moot on the surface, as will be shown in the below outage economic analysis, however a comparison is useful with the following approach.

Consider an SLA commitment and the accompanying credits schedule to be an indication of how confident the provider is that they can perform according to their commitments. For the purposes of this discussion we will assume that the service providers intend to remain (or become) economically viable concerns and charge a minimum of a 30% gross profit margin over their operating costs. We will further assume that they will give up the gross profit for periodic non-performance, but not the operating costs. By plotting their SLA credit schedule against the gross margin line we can predict the lowest performance expected by the provider, i.e. they expect to do better in aggregate. Using Provider B as an example, we can determine that they will experience outages less than 83 minutes in duration on average. That metric can easily provide a comparative cost of doing business with Provider B versus others, as well as provide an empirical means to "equalize" providers regardless of protected or unprotected service offerings. We use the term "Failure Metric" to describe this evaluation variable and demonstrate its use in the following economic analysis section.

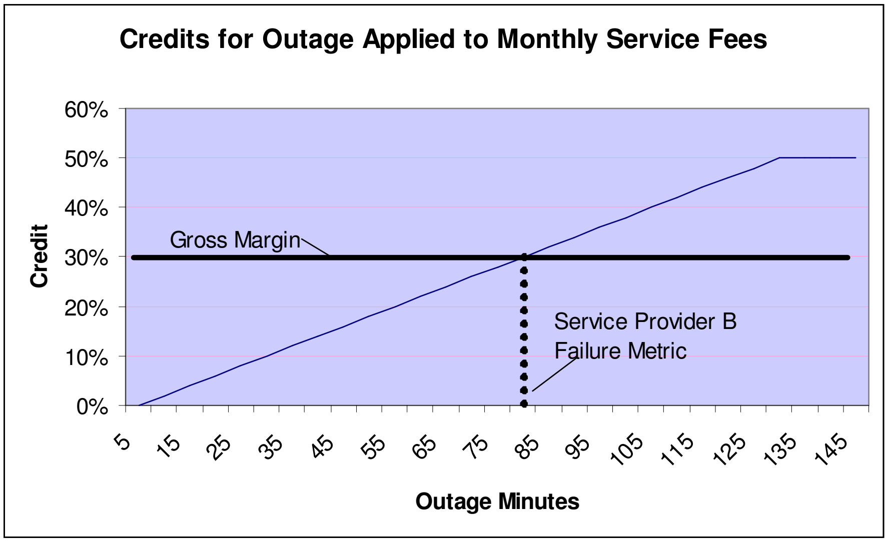

Figure 8: Service Provider B Failure Metric

Economic Analysis:

The cost of network downtime is very significant for most enterprises. An average quoted in telecommunications industry publications is around $1M per hour of downtime.[5] This includes sectors such as Healthcare, Utilities, Transportation, Banking, Pharmaceuticals, Retail, Insurance, Finance, Manufacturing, and Energy. This type of metric assumes a fairly large scale to the enterprise, but the same type of analysis can be done for specific smaller enterprises and can yield discrete values for any proposed telecommunications solution.

Rather than accept an aggregate view, it is possible to construct a top-down cost model or bottoms-up. Which is chosen depends very much on the company structure and how easy it is to gather the appropriate inputs. Note that the most important consideration is an order of magnitude comparison. While it is possible to be very accurate with a lengthy data gathering phase, there is little added value delivered to the final “answer”.

Top-Down Model:

[5] Source: Meta Group cited in Network World 11/2001

A simple top-down model would be to gather the company's financial data in aggregate, in particular the Operating Cost for a month (or other length of time), the percentage of cost attributable to salaries/benefits, the percentage of the firm that would be affected by the outage, and Depreciation of fixed assets. Note that the cost and depreciation values should reflect the same time periods. Figure 9 details how an outage cost per minute calculation can be put together.

Enterprise Operating Costs (1 month)	$ 2,200,000
Enterprise Fixed Asset Depreciation (1 month)	$ 275,000
% of Enterprise affected by outage	50%
% of Operating Costs attributable to Salaries/ben	50%
Total Cost for Salaries/Benefits	$ 1,100,000
Total Cost for Salaries/Benefits per minute	$ 105.41
Adjusted Cost for Salaries/Benefits per minute	$ 68.52
Total Cost for Depreciation per minute	$ 6.27
Total Cost per Minute Outage	$ 74.79
Adjusted for % of Enterprise affected by outage	$ 37.39

Figure 9: Top-Down Model for Outage Cost Metrics

This model contains several adjustments which need to be made simply because of both the behavior of workers facing an outage and the fact that the outage may not affect the entire enterprise. Our studies have shown that workers faced with an outage will spend somewhere between 30 minutes and one hour trying to recover from the outage. When the outage exceeds that amount of time, workers redeploy their time into other tasks which can be accomplished without the failed link. That significantly reduces the effect of an outage that a simple calculation would indicate, although over time the inefficiencies accumulate. The total also needs to be adjusted to cover just the portion of the firm affected.

Since the company now has a usable metric for cost, $37.39/minute of outage, they can evaluate whether the added cost of diverse protection is better for their needs.

For simplicity's sake, lets assume a perfect provider world, without marketing adjustment of availability statistics. In this case, an unprotected circuit will experience less than 43.8 minutes of outage versus 0.438 minutes of outage for the diverse protected offering. The price for unprotected is $3000 per month, the price for diverse protected is $6500 per month. The actual cost of the service will be the sum of the monthly charge plus the monthly expected outage costs.

In this case the following results would accrue.

Calculation (Provider A)	
Subtract 99.9% minutes from 100% (minutes)	43.8
Multiply by Company outage cost/min	$1638
Add to monthly charge for adjusted monthly cost	$4638
Decision Result:	Unprotected

But wait, this isn't a perfect world, the provider is expecting worse performance (per the above analysis) so the cost needs to be increased accordingly. Provider A is expecting outages to be less than 13,149 minutes (Provider A Failure Metric), so the actual expected cost will be greater than a $1621 differential, closer to $490K per outage, which would be a ringing endorsement of the protected service. At the end of the day, companies should realize most outages exceed 4 hours (>$8974 based on the company's cost estimation), and that statistical tools provide an aggregate look. How confident/bullish the customer is regarding provider predicted performance will definitely affect whether to use the nominal SLA commitment or actual credit schedule as a predictor of provider performance.

Bottoms-up Model:

A bottoms-up approach would require the total affected employee salaries/benefits, affected fixed asset depreciation, and an adjustment factor for the redeployment of resources. One advantage of a bottoms-up approach is that actual discrete data is obtained and is therefore more accurate, the only modeled parameter is the redeployment adjustment factor. Figure 10 details how an outage cost per minute calculation can be put together.

Total Cost for Salaries/Benefits per month	$ 1,350,000
Total Cost for Salaries/Benefits per minute	$ 129.37
Total Cost for Depreciation per minute	$ 5.53
Total Cost per Minute Outage	$ 134.90
Adjusted Cost for Redeployment per minute	$ 87.68

Figure 10: Bottoms-up Model for Outage Cost Metrics

The below uses the same method as above to derive a decision.

Calculation (Provider A)	
Subtract 99.9% minutes from 100% (minutes)	43.8
Multiply by Company outage cost/min	$3,840
Add to monthly charge for adjusted monthly cost	$6,840
Decision Result:	Protected

Since the "perfect world" result is very close to the $6500 charged for Provider B's protected product, the same calculation should be done using Provider B's information for a complete look. The SLA credits approach yield an even more lopsided judgment in favor of the protected solution.

Calculation (Provider B)	
Subtract 99.999% minutes from 100% (minutes)	43.8
Multiply by Company outage cost/min	$38
Add to monthly charge for adjusted monthly cost	$6,538
Decision Result:	Protected

Provider Failure Metrics:

In the Provider A and Provider B service offering cases that we have presented, Provider B is making an attempt to justify their SLA claim, while Provider A is not. Provider B is experiencing something less than 83 minutes of average outage per month, based on the Provider Failure Metric Model. Provider A is more cynical, if we can make that judgment, as their SLA has little in the way of credits and seems to merely be an attempt to say they have credits rather than a means of supporting their performance claims. Provider A's failure metric is less than or equal to 13,149 minutes. Given that data, it is hard to justify their SLA as a reason to award them business.

We consider that it would be in the interest of the user community to promote the use of Failure Metrics as an analytic decision making tool for telecommunications service contracting. The more that it is utilized by enterprise customers, provider SLAs will become more indicative of a provider's actual performance commitment.

Service Providers can also realize cost differentiation benefits from adopting its' use in crafting legitimate product SLA commitments, ending the use of SLAs as an industry shell game.

The SLA data used in this case are very similar to Providers that exist within the actual marketplace.