

# Service Levels are Bunkum

*Why do Service Level Agreements frequently achieve the opposite effect to that intended? Do the parties really understand what risk they are taking on in signing up to these contracts? This paper looks at some of the wilder myths and fundamental errors perpetrated by this burgeoning business practice.*

Presented by Intercai Consultant  
Steve Hodson

# White Papers

From IML

## Introduction

Service Level Agreements (SLAs) have been with us a long time. James Watt was onto a winner (see panel) back in the 18<sup>th</sup> century and we all know where that led us. More recently they have been used in retail with the likes of the comfortable old favourite '*never knowingly undersold*'. Up to a point, they gave the customer a warm feeling that a service or piece of technology that he or she did not properly understand really would step up to the mark claimed for it by the supplier.

While the service that was to be provided was clear and the SLAs themselves were simple and easily understood, there were some notable successes apart from James Watt. The Government Data Network installed in the late 1980s was a good example. It had its flaws (see the Wall of Death that we discuss below) but by and large both parties to the agreement felt that the targets set were realistic and neither side was pulling the wool over the eyes of the other. However as we make more and more use of specialist suppliers and outsourcers, we have reached for the SLA as a means of keeping the suppliers under control. Riding on the back of this, a whole industry has arisen providing advice, training, software to measure things,

James Watt hit on a good thing when he was repairing one of Newcomen's atmospheric engines back in the 18<sup>th</sup> century. When he went into business with Boulton he was able to offer a guarantee that his engine would use less coal than his competitor's. He even offered to pay for the coal used if his engine didn't live up to this claim. History doesn't record whether it was ever put to the test, but the indications are clear.

handbooks on the management of SLAs and all the rest. SLAs have become as complex as the systems we use them to manage so where has the warm feeling of understanding gone? With this complexity has come the need for armies of people and rooms full of systems just to service the insatiable needs of SLAs that require whole forests full of reports to be provided every month. More dangerously, the SLA thresholds have become enshrined in contractual agreements and are backed up with potentially punitive financial sanctions so the effort has to be put in to monitoring them with the consequent diversion of effort from the original purpose (which, in case you have forgotten, was to provide a reasonable service).

This paper takes a sometimes light-hearted, but always focussed, look at the big wide world of the SLA and the industry that has grown up to support it.

## Invasion of the Marketeers

I suspect that things started to go wrong when the engineers, believing they had done all that was needed with regard to SLAs, turned their attention elsewhere and allowed the marketing men free rein with their invention. We

# White Papers

From IML

have reached a position now where SLAs have lost their original purpose and are used as a marketing tool and, as we all know, at this point any connection with reality or good engineering principles departs through the nearest window.

The Telemangement Forum (TMF) has done some sterling work in its handbook (see

<http://www.tmfcentral.com>) but

although much of its content is admirable and useful, it really addresses much wider issues than just SLAs. There is much muddled thinking about SLAs and that is leading us into expensive and unnecessary activities that, with a little bit of thought, could be avoided. If, in achieving this, we can manage to get back to the original purpose of having an SLA, then that can only be

a good thing, both for service providers and customers.

Consider, in relation to the TMF definition of an SLA (in the panel), a measure that has some variance or variability in its value each time you measure it.

Almost every human activity falls into this category and there is a wide range of mathematics available to characterise it. Each time we measure it we get a different answer, but the answers are distributed in some form around an average value. If now we set

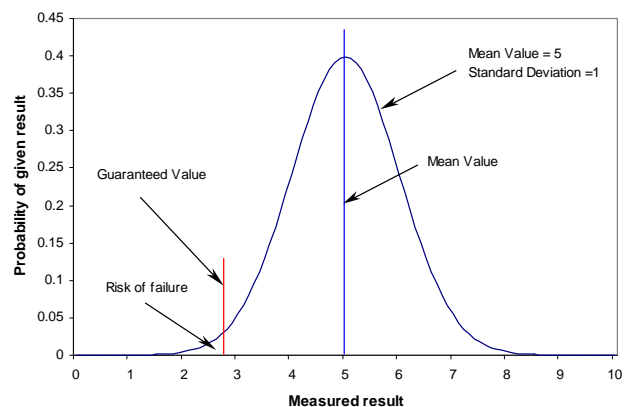
our SLA threshold at the expected level of service we can see, without any maths, that we will fail to hit our target on half of the occasions we measure it. Does this mean that the service is unsatisfactory? The SLA would have you believe 'Yes', whereas a little thought should show that it is 'No'.

Now take this thought a little

further. If we can characterise the distribution of the measures around the average, then it should be possible to set a threshold at a point where a service that is delivering what was expected does not trip the SLA threshold too often. We could even set the risk of this happening. Just think how

The Telemangement forum, in its SLA Handbook, defines an SLA thus:

A SLA is a formal agreement between two parties. It is a contract that exists between the Service Provider and the Customer. It is designed to create a common understanding about service quality, priorities, responsibilities etc. SLAs can cover many aspects of the relationship between the Customer and the SP, such as performance of services, customer care, billing, service provisioning etc. [TMF 701]. However, although an SLA can cover such aspects, agreement on the level of service is the primary purpose of a SLA.



pleased the Financial Director will be if you can tell him how much money he will have to set aside to meet his commitments under the SLA. This ideal world seems a long way from reality and although the rooms may not be filled with smoke any more (although from some of the SLAs that have been signed up to you might think otherwise) the discussions on the content of SLAs have moved into the 'Business' arena rather than retaining the technological input they badly need.

### Let's get the outsourcer to do it

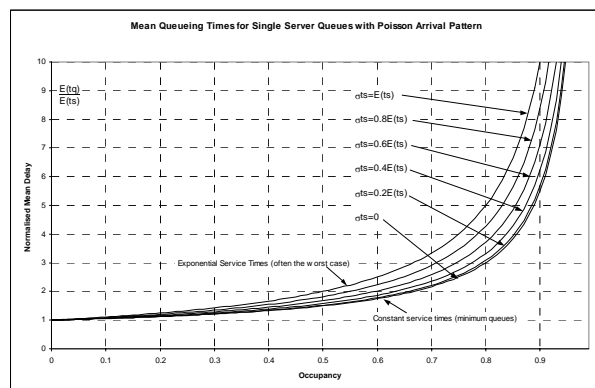
Outsourcing, especially of IT systems, is a field ripe with opportunity for complex, hard to measure and, ultimately often ineffective, SLAs. We frequently see IT functions that have been happily running a system for many years without ever measuring anything much. Then when the outsourcing contract comes along suddenly it is essential to measure over 1,000 different points in the system and expect target levels way above our wildest dreams of achieving. Not only that but the infrastructure is expected, suddenly, to improve after being outsourced and it was supposed to become cheaper. You may think I am joking or exaggerating for the sake of effect. I would it were so. It is real and current and many tens of millions of Euros hang on achieving these unnecessary and unrealistic targets.

In what other field of technology would you find something costing so many millions that had no design effort assigned to it and no budget in the business plan? Perhaps the rooms are full of smoke after all, and not just tobacco smoke.

Our lives are dominated by these targets these days and you don't have to look very far to find more

examples of totally wasted time and money. There is the case of the hugely complex SLA threshold formula that, when modelled properly, applied no incentive or pressure whatsoever on the Service provider. Then again there are the targets for major infrastructures such as health where it is expected that resource occupancy should be as

close to 100% as can be achieved. That is all very well but surely we shouldn't be surprised when the queue lengths and delays shoot to infinity. Has everyone forgotten about queuing theory?



### Lets get a piece of software to do it

Software suppliers will never question the need for a new piece of software – it is what keeps their R&D departments alive. You can buy software today that

manages your SLAs from the bottom up. Some is even so prescient that it tells you in real time when you are about to breach your SLA even before the system itself is aware of it. These are not trivial systems and none is all pervading (which is another way of saying that none of them work together out of the box)

Some systems address inventory management; others look at trouble-ticketing. Some are best at supply chain management and provisioning while others look at performance measurement. All can give rise to the need to measure things so another layer of software (goody goody you can almost see the software developers rubbing their hands in glee) has arisen that takes all the measures from all of these software systems and combines them together into dashboards, reports, risk assessments, trend analyses and a whole load of other things that you never knew you needed. The trouble is that all these systems come from different companies and no-one knows how to make them work together. How many times have you seen major software packages purchased and left to rot for lack of the expertise or time to set them up and integrate them with other systems?

Looking at the mountain of information, the software systems and the army of people needed to run them, you could be forgiven for forgetting that originally all you wanted was to make sure your supplier fixed things efficiently when they went wrong.

Welcome to the club – you have been hi-jacked by the SLA industry.

## **My brain hurts**

The root cause of much of this confusion is some extremely muddled thinking. Customers feel a need to throw the kitchen sink into an SLA when there is no adequate formal description (otherwise known as a specification) for a service. Similarly many SLAs attempt to address shortcomings in procedural definitions by imposing measures on them. Often the SLA is the first document rather than the last to be written.

By taking a highly structured approach to

- (a) Service Specification and
- (b) Process, procedure or operating methods,

it is possible to keep the SLA cleanly focussed on measuring just the minimum needed to keep the supplier honest rather than fill in holes in the above documents. My measure of adequacy of documentation is that a service should be capable of operating to its specification and procedures without the existence of an SLA. If the SLA is needed for a service to be able to operate then it has become a part of the service and loses its independence.

The other big sin in my book that arises from muddled thinking is the hijacking of the Availability parameter by the SLA industry. Availability is a perfectly good term with a specific definition related to numbers of faults and how quickly they are repaired. Performance such as Response time or delay is a

separate parameter using different units and arising from very different causes. In many SLAs these two incompatible parameters are combined into a single 'Availability' parameter. Mixing units has always been a technical no-no, but this approach mixes causes and responsibilities as well and actively prevents adequate analysis of the true causes of a problem.

## **Lets really hammer the Service Provider**

Once the list of SLAs and their targets has been agreed, thoughts in the (smoky) room turn to how to enforce the thresholds in the contract. Again outsourcing is a big culprit. Advantage is taken of the fact that the Service Provider is probably desperate for the contract to introduce swingeing penalties in the event of failure. Customers see this as a major form of initiation ceremony; to impose the harshest possible sanctions on their suppliers. Whilst this might get some initial plaudits in the boardroom it is gambling with the future of the company. Suppose the Service Provider is driven into bankruptcy by the SLAs – and again I am not joking, SLAs frequently go this far – what would happen to the Customer's business when it closes down? I wouldn't want to explain that to the shareholders.

In this case sight is easily lost of the original purpose of the sanction.

Punishment should fit the crime and should start off gently and increase in intensity as the degree of failure increases. There is no point in taking this to an extreme, if the service is so bad that any SLA would be punitive, then obviously the SLA isn't working and some other recourse (such as breach of contract) is needed. SLA sanction schemes, therefore, should always be graduated and capped if they are to be effective.

However, you still have to think about the structure of the scheme. There is no point graduating the pain if the first degree of pain takes you past the capping point (yes, a real case again).

Finally, don't guess. If you don't know what your performance will be don't sign the contract. The risk can be enormous and the penalties potentially very onerous. You wouldn't take such a risk with hardware design, why do it with SLAs. There is a perfectly good set of design mathematics to work out the risk of failure and if the raw data to work this out doesn't exist that is another good reason for refusing a contract.

## **Did I really say that**

We have all seen, and possibly experienced, instances where the existence of targets (SLAs) has the opposite effect to that intended. This result increases exponentially with the complexity of the SLA scheme. There are widely publicised instances in Health (where targets override decisions based on clinical need); Safety (where safety critical actions are delayed in order to meet non-safety specific targets) and our old friend IT outsourcing which is a rich source of stories about the law of unintended consequences.

# White Papers

## From IML

Consider an IT support operation that has a series of measures, all supported by SLAs with fixed penalties. The negotiators no doubt felt they had done a good job by specifying the target levels and the punishment to be meted out in the event of failure. In real life, however, the Service Provider worked hard to keep the most critical service alive up to the point in the month when the downtime exceeded the allowed amount and the penalties were invoked. At that point his only possible rational action was to switch the critical service off and concentrate on the other services that could still damage him in the latter stages of the month. The critical service had lost its power to affect him and the SLA was not having the desired effect (the so-called Wall of Death sanction).

On a similar vein, a call centre was required to measure every type of response that it handled and each category was issued with a pair of targets that specified the 90<sup>th</sup> percentile of response times should be X seconds while the 95<sup>th</sup> percentile should be Y seconds (bigger than X). This sounds a perfectly valid and statistically sound approach. However, since there were only about 5 responses of each type during the month, discriminating between the percentiles was a little difficult.

In cases like these, sight of the original purpose of the SLAs has been lost and it has become an end in itself that often defeats the original purpose.

### **And Finally....**

So, what should we do to avoid falling into these types of traps? It obviously helps to sit back and really think about how to define and document the services and processes at the same time as the SLAs are being defined. There is no need to measure everything that moves and clear thinking at this stage will avoid that. It is also essential to have an understanding of the detail of the service and sanction systems and be able to model them – and in being able to model them oddities such as the inability to work the statistics will be identified.

Finally, it is just as important to 'design' the SLA as it is to design a switching system. You know it makes sense!

Just in case you need help (preferably before you get into a mess) then please do give us a call.