Sustainable Cloud Adoption
A Financial Perspective

Cloud Cost Optimisation Maturity Model

A Strategic Blue White Paper

Strategic Blue

Introduction

There are many assessment frameworks, migration methodologies and maturity models to support cloud adoption. When considering maturity models our experience at Strategic Blue is that these *can* provide a useful tool in providing guidance and initial structure but they must be used with flexibility and with a clear focus on the environment in which they are being used. They should be considered as a structured illustration of *potential* building blocks from which a roadmap is built and then continually reevaluated and refined to ensure it remains relevant.

We believe this approach is vital for two key reasons. Firstly, as a general point maturity models often have their roots in promoting a service need, guiding through a sales journey or evaluating risk. This inherently introduces a bias, an awareness of which must be maintained when either developing or using such models. Secondly, the power of cloud is its ability to support evolution and innovation and that dynamic nature should not be crippled by adherence to static models that solely create discrete snapshots in time. The model itself must allow flexibility and continuous evaluation to ensure it remains a useful tool rather than a bureaucratic barrier to agility.

Strategic Blue is a finance organisation that has specialised exclusively in supporting hyperscale cloud adoption for over 10 years. In that time we have grown with our customers as the capability of the Cloud Vendors has exploded. This growth has come with an increasing ability to support a wide range of cloud use cases and an increasing trust in its ability to support complex, business-critical systems. With the growth in cloud adoption, we have seen increasing governance and assurance demands on and expectations from the business, not just IT.

The pace of development has been phenomenal and this creates opportunities and challenges in equal measure. The impact of and momentum behind Digital Transformation is such that overcoming these challenges is now a requirement because those that can are creating such powerful disruption that those who cannot keep pace cannot survive.

This Cloud Cost Optimisation Maturity Model was created to help organisations adopt cloud in a sustainable and cost effective way. We use it to guide development of our suite of services. Our customers and partners use it to guide their internal discussions, tactical and strategic planning. Together we collaborate to understand what we can each do to bring required value at a particular time to ensure that the paradigm shift cloud allows can deliver capability without sacrificing financial governance and assurance.

Five Pillars of Cloud Cost Optimisation

Five pillars of cost management are defined to first understand where cost is incurred, eliminate waste of unnecessary consumption and optimise procurement of the necessary. For continuous improvement, technical *and* financial factors should be aligned in the design of consumed services and a framework created to evaluate the efficiency of the end to end cost management and optimisation process.

Cost Visibility & Attribution

Historical usage clarity for, budgeting and forecasting, tracked at the required granularity.

Usage & Utilisation

Identify and minimise waste associated with underutilised resources, services and licences to consume only what is needed when it is needed.

Commitment Savings

Minimise spend at the premium, pay-as-you-go rate without undue risk or compromise.

Design for Cost

Embedding financial considerations into the technical architecture design process to ensure capability and value delivered are evaluated in equal measure.

Monitoring & Improvements

Benchmarking of the efficiency and value of cost optimisation and management processes to drive continuous improvement.

We have seen its elements combined to good effect in different ways at different times and in different circumstances. The model's structure presented here is therefore only an attempt to combine its elements in an understandable manner to demonstrate the potential.

Organisational development opportunities within each pillar are provided via progression against four key characteristics. This is the extent to which they are: proactive vs reactive; standardised vs ad-hoc (at local levels within or globally across the organisation); automated vs manual; and quantitative vs qualitative in the evaluation of their overall effectiveness.
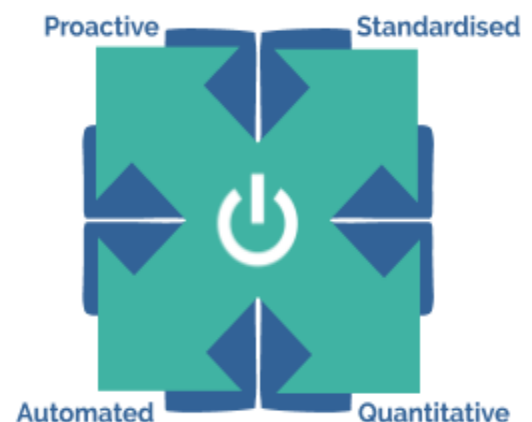
At the very start of cloud adoption technical capability typically takes precedence over cost optimisation. This is not a long term sustainable position. Beyond this initial point the model frames levels of "maturity" to explore how these four development characteristics manifest in tangible actions as summarised below.

	Maturity Level Reactive, Ad-hoc, Manual, Qualitative → Proactive, Standardised, Automated, Quantitative				
	1 ▶	**2 ⚙**	**3**	**4 🔍**	**5**
Cost Visibility & Attribution *Cost transparency*	High Level Tracking	Budgeting and Forecasting	Anomaly Alerting	Low Level Tracking	Overspend Risk Management
Usage & Utilisation *Cost Cutting & Avoidance*	Remove totally unused resources	Rightsize necessary resource	Automated identification of wasted resource	Redeployment and repricing	Quantified waste reduction
Commitment Savings *Cost Optimisation*	Localised and ad-hoc	Defined guidance baseline	Embedded dynamic process	Quantifiable impact	Global, organisation wide process
Design for Cost *Cost Containment*	Fully technical in focus	Static consideration of high level factors	Defined standard for high level factors	Cloud native approaches adopted	Quantifiable impact of design policies
Monitoring & Improvements *Value Creation*	No cost management process exists to monitor	Cost metrics defined as process KPIs	Automated collection and review of KPIs	Outcome metrics defined to enhance KPIs	Cost Management process delivered at lower cost

Each pillar will be explored with respect to these characteristics on the understanding that relevance of each element to each organisation will vary. This will be defined by the individual characteristics and objectives of the organisation and as it progresses its cloud adoption journey. Evaluation of relevance should consider factors such as the financial benefit, the organisational and technical risk, and the investment of personnel, time or money to deliver.

Organisations can progress at different rates across each of the pillars but the greater the alignment of maturity in each pillar, the greater the effectiveness of the overall process.

Cost Visibility & Attribution

Cost Transparency of historical trends, current costs and cost owner is the foundation to any cost management or optimisation process. Without this, it is impossible to determine where and how the cost is generated, to evaluate the value delivered, to forecast, budget or proactively mitigate the risk of overspend.

Tracking: Cost visibility is provided at an overall level, aggregated by service. Cloud usage is categorised using high-level Cloud Vendor organisational structures (e.g. AWS Account, Azure Subscription or Google Cloud Project). This is used to provide preliminary cost attribution. Cost management is predominantly reactive as cloud spend is unpredictable.

Planning: Each cloud service can be associated with an expected cost that is based on cost calculators, completion of pilots or profiling existing use. This affords more measurement and control for high-level forecasting and budget planning.

Alerting: The results of planning are used to integrate forecasts and budgets into cloud spend monitoring to trigger alerts for high and/or anomalous expenditure as potential overspend situations.

Attribution: An organisation-wide cloud spend categorisation process is followed to enable more specific attribution of costs, typically enabled via tags. The high-level tracking aggregated by cloud service and high-level Cloud Vendor organisational structures is augmented by more granular attribution to internal organisational units. This increases accountability for spend, understanding of value returned and enables focused remediation work when overspend is identified.

Overspend: This includes prevention, detection and remediation controls and procedures to mitigate the risk of inadvertent overspend. Identity and Access Management provide guardrails as to who can incur spend, how and to what level. Billing alerts provide notification to spend owners, with appropriate escalation. Thresholds set, and associated procedures defined, to investigate and control escalating costs. Insurance policies implemented to cover significant, erroneous overspend.

Usage & Utilisation

Eliminating the waste of purchasing resources and services at too high a level or that are not required at all is the most obvious **Cost Cutting and Cost Avoidance** method. To be effective this needs monitoring of usage and utilisation rates and the technical ability to adjust service and resource adoption levels required to meet point in time demand.

Quick Wins: Idle resources, orphaned servers, unassigned IP address, unattached storage systems and old snapshots, for example, are identified and removed in ad-hoc cleanup exercises to eliminate obvious waste. Storage auto-tiering is enabled for blob based storage systems to ensure storage performance and durability is aligned to specific data access requirements ensuring premium rates are only paid where necessary.

Rightsize: Utilisation rates of server instances are verified to confirm whether they are appropriately sized, be that too large or too small. Autoscaling is used dynamically to adjust the configuration to meet the demand required at any point in time.

Initial Optimisation: Monitoring for and alerting against idle and unallocated resources used to proactively trigger remedial actions prior to any scheduled reviews. This includes *server instances* to use the latest instance types and services to improve performance vs price; *storage* to verify auto-tiering is working effectively and using optimum resources; and operating system and software *licences*. For services which run continuously licences are generally better purchased outside of the Cloud Vendor. For services used infrequently, licences can be bought via the Cloud Vendor for simplicity or flexibility.

Advanced Optimisation: At a *server instance* level this includes migration to the latest instance types to reduce cost, scheduling for non-persistent services to align with vendor pricing models (e.g. AWS spot instances) or to commitments for persistent services. At an *application level,* this includes technical optimisation of code (i.e. improved code efficiency and/or containerisation) to reduce infrastructure usage and associated costs. Engagement with stakeholders should ensure that applications are delivering the required value and without unused capability that may have an associated and therefore unnecessary cost to deliver.

Mature Optimisation: The actions taken in initial and advanced optimisation continue but with the results being quantified and tracked to keep the time when any resource operates at a sub-optimal level to a minimum.

Commitment Savings

Cloud vendors provide incentives, in the form of discounts, for commitments to persistent usage or spend at either a resource level or globally across a range of services. As a **Cost Optimisation** strategy, these discounts can be significant and further improved with different payment profiles. With 1 or 3 year terms typical, care is needed to understand the scale and hence the *risk* of making such commitments.

Ad-hoc: Commitment based savings are made at a local level within the organisation with decisions made on a system-by-system or project basis rather than through a standardised process or policy.

Baseline: A baseline has been created of which services (compute and database at a minimum) should have a commitment attached and which services should be Pay As You Go. This baseline is used to prompt suitability for commitment discussions for relevant services.

Dynamic: A proactive commitment process forms part of new deployments and change/release management to regularly update the reservation baseline to reduce system costs. There is a regular review of the reserved services. This should include any service for which the Cloud Vendor offers commitment options.

Quantified: The impact of all savings activities can be quantified in impact on savings made, impact on cash flow and potential future flexibility implications. The commitment process is supported with active engagement with Cloud Vendors to explore opportunities to gain local and global offers for further discounts and savings.

Organisational: Commitment decisions are made on the aggregated consumption across all departments, applications, business units and projects rather than at an individual level. Combining usage in this way maximises commitments whilst minimising the risks of over-committing or significantly impacting flexibility as organisational units provide a degree of risk mitigation for each other. Commitments can be made at the resource level or now globally across all types of consumption for flat discounts against public pricing.

Design for Cost

A useful **Cost Containment** approach is to design cloud resource and service adoption not just to meet the functional and normal non-functional requirements, but also to consider how best to meet those core requirements in the most cost-effective way from the outset.

1 Level

Functionality: At this early maturity level cost is not a factor in making design decisions, functionality and other technical capabilities such as security are the principal concerns.

Initial Cost: Factors such as location, minimising network traffic flows between locations and instance generations are used to determine resources to deploy prior to deployment.

2 Level

3 Level

Advanced Cost: The localised, ad-hoc initial cost considerations are standardised globally across the organisation with clearly defined policies against which cost-effectiveness of design decisions can be assessed. Policies are routinely updated based on Cloud Vendor pricing changes, product releases and service enhancements.

Cloud Native: Infrastructure deployment is automated through code (IAC) with automation and orchestration to provide Continuous Integration and Deployment (CI/CD) pipelines. The potential offered by higher-level services, such as FaaS, PaaS & SaaS, which can include serverless technology as well as 3rd party SaaS services, is evaluated and incorporated where deemed appropriate.

4 Level

Quantifiable: The impact of design decisions and the effectiveness of cost-conscious design policies can be reliably quantified. This expands beyond options within and comparison between Cloud Vendors to include wider factors such as scope and scale of change (complexity and risk), speed of adoption and deployment (time to value), future portability between cloud vendors (lock-in) and in-house staffing costs (time, training, recruitment or redundancy).

5 Level

Monitoring and Improvements

The process of cost management and optimisation itself needs to be monitored and maintained to ensure it plays its part and does not act to limit opportunities for *Value Creation*. For some organisations, this may form part of a regulatory, investor or other compliance obligation providing an audit trail for actions and initiative but cost alone cannot be the only measure of value in the cloud.

Infancy: Cost management and optimisation processes themselves are so limited at this level of maturity there is nothing meaningful to measure or analyse.

Cost Metrics: A set of key performance indicators are defined to gauge the impact of the actions taken in each of the other management and optimisation pillars. These are periodically collated and reviewed with impact assessed in terms of cost savings,

Automation: The collection and presentation of cost metrics is automated to simplify the process and reduce the opportunity for human error in their collation.

Outcome Metrics: As the management and optimisation process matures in the other pillars, cost metrics can be replaced by outcome-based metrics, such as cost per order placed on a retail site. In this way, leadership teams are provided more useful insight on returns on investments.

Streamlining: The process by which initiatives are measured and presented is made more efficient to reduce effort, time and thereby costs to provide insights. Available near real-time in an interactive manner such insights can support scenario modelling to validate new organisational initiatives or demonstrate financial viability to investors or auditors for example.

Conclusion

To maximise effectiveness, cost management initiatives should include elements of cost transparency, cutting, avoidance, containment and optimisation but without losing focus on value creation.

This paper has introduced the components of such a management approach for organisations to incorporate into their tactical and strategic adoption plans for cloud. It provides a reference for the cloud adoption journey when the requirement for increased sophistication at each layer of cost management becomes more important. In this way this approach will support operational success within, and organisational success from using, the cloud.

Combining mechanisms for financial governance and assurance with technical capability will make cloud a sustainable platform on which organisations can deliver, with the agility to adapt, freedom to innovate and ability to disrupt.

The Author

Jon Bryant

Financial Solution Architect

Jon joined Strategic Blue in December 2018 and leads the technical guidance for customers to optimise their cloud platform costs. He has worked for 25 years across the IT industry, with the last 10 years as an Architect. Jon has a wide and diverse set of knowledge and experience across public and private sector organisations covering many technical services and processes.

Cost Management & Optimisation

delivering the power of cloud without sacrificing financial control

Strategic Blue

Founded in 2010 we are a financial services organisation specialising exclusively in cloud procurement, cost management and optimisation. Our professional consultancy and managed support services provide Cloud Financial Operations, "FinOps", services to cloud customers both direct and through partners.

Our FinOps services allow cloud customers to buy in the way they want, understand what they have bought and continually optimise cloud costs. By using economies of scale across our customer base our customers benefit from additional discounts not possible direct from cloud vendors. Many of our services are provided without management fees.

As financial experts in cloud we help organisations develop the sophistication of their cloud procurement, governance and assurance processes. We understand the necessary integration between technical and financial teams. Our Financial Solutions Architects have technical engineering backgrounds to bridge that gap either supporting internal teams or working with members of our Teaming Partner Network.

Find out more:
strategic-blue.com

Contact us:
+44 203 872 2583
contact@strategic-blue.com