Increased automation is broader than just NFV

The underlying technology theme for Network Function Virtualisation is increased automation. Peel away the layers of detail, and the argument for NFV is easy to grasp: virtualization enables an increased degree of automation in your network and service operations. This in turn enables a corresponding increase in the automation of business processes surrounding the network and therefore reduces overall operating costs – especially the costs associated with changes such as roll out of new services.

This simple economic argument – more automation equals lower costs and better agility - is by no means a new one. It’s been around for decades, embodied both in open approaches from standards organisations and closed ones from “full stack” OSS/BSS vendors.

The impetus for NFV comes from the observation that despite the venerable history of the idea, automation in Communication Service Providers (CSPs) has increased very little because the underlying network supports it very poorly. NFV – and of course, its conjoined sibling, SDN – are radical changes in the underlying technology itself that aim to put more of the network under software control.

There is a danger that this perfectly sensible high-level perspective is producing a distorted architectural orthodoxy which dictates that no further automation is possible until current network technology has been replaced.

The “revolution or evolution” argument about the adoption of virtualisation technology that has played out in the community over the last few years is being resolved by the facts on the ground: CSPs are (some would say unsurprisingly) gradually evolving their networks to increased software control and virtualization.

It stands to reason that opportunities for additional automation of existing processes will therefore also present themselves incrementally at various stages as CSPs undertake this transformational journey – and that these will be broader in scope than just NFV.
Service Assurance: First-class citizen at last?

Despite claiming to accept the importance of designing assurance into services (and building infrastructure that can support it), CSPs have often failed to walk the talk in this respect, and Service Assurance has historically been something of an afterthought.

This reflects its role in the economic lifecycle of a service: it is generally possible to initially bring a service to market without assuring it; it’s only when service quality issues begin to hurt the bottom line through poor customer satisfaction and retention that Service Assurance finds itself in the critical path of revenues.

Pre-empting this pattern is often seen as a mere optimisation because it’s only through the cumulative effect of the pattern over time that the provider suffers long-term, large-scale damage to its competitive standing.

This attitude is changing as markets have matured, forcing providers to play the long game, and although Service Assurance was still not front-and-centre in early NFV reference architectures from the likes of ETSI, it has become much more present in industry dialogue as CSPs start to think seriously about fully orchestrated services in production. This in turn has led to a re-examination of how they deliver assurance in services right now, and the question of how new technologies in play can be harnessed to provide immediate benefits.

Closed-loop Service Assurance: An idea whose time has come

The notion that the relationship between a correctly functioning service and the corresponding network configuration is not static – and so would benefit from being managed in a closed, automatic loop to modify the configuration to maintain the desired behaviour – isn't new. When policy-based network management was in vogue twenty years ago, this idea was much discussed.

However, it was only ever implemented at the device level, never network-wide. This was mainly for two reasons: insufficient trust in automatic management software and fear of unintended emergent behaviours. Both stemmed from a lack of reliable network visibility.

Operators could not reliably know how their network was structured or how it was behaving. This body of knowledge was incomplete, fragmented and many parts of it were manually maintained. How therefore, they quite rightly reasoned, could an automatic control system possibly be trusted – and if couldn’t be trusted, how could they rely on it not to throw the network into chaos?

What has changed since then?

First and most obviously, operators have become more comfortable with the idea of some degree of autonomous control in the network - because it's inherent to promised technical benefits of NFV like elastic scaling and because it's been proven technically to some extent by web-scale enterprise players.

Secondly, economic circumstances for CSPs have become tougher. Margins are under unprecedented pressure and the cost of continuing operation as before, whilst at the same time scaling the network and services to ever greater degrees, is simply unsustainable. In the medium to long term, reducing operating costs through automation is no longer an optimisation to improve profit: it is mandatory for survival.

Both these changes in circumstance clearly apply to automation across the board, but in contrast to full service orchestration, which requires significant advances in service and resource meta-modelling, in activation and provisioning systems and direct support from the underlying network and compute hardware, many of the parts required for service assurance automation are already in place.

As well as the historical machinery of policy-based management, CSPs have been incrementally improving performance measurement, fault management and in many cases, have already begun
to invest in analytics to make sense of the resulting increase in volume and quality of telemetry they can harvest. They now have a good quality, automatically curated view of what their network is doing.

However, closing the loop on service assurance – that is to say, enabling autonomous correction of poor service performance and autonomous pre-emptive change to avoid poor service in the first place – requires more than just measurement.

**The key role of structural network and service visibility**

To close the service assurance loop, it must be possible for the kinds of decisions and actions currently undertaken by network operations staff to instead be undertaken by software. The key inputs used by NOC staff to make decisions certainly include measurement and performance data – and modern analytics goes a long way to providing the kind of interpretation of this data that human experts provide. However, the other key knowledge domain available to operators is the structure of the network and services: the topology.

Astonishingly little progress has been made in automating the assembly of a comprehensive, cross-domain view of the structure of network suitable for machine consumption.

Network inventories, which once promised to achieve such a view, failed on two fronts: they failed to develop reconciliation systems capable of reliably populating the Inventory model and they failed to find a way of making the model easy and cheap enough to change that it could keep up with rapidly evolving network designs and technologies.

As an aside, it’s worth noting that the need to find a technically viable approach to cross-domain topology has more than one driver behind it. After a period of eschewing the need for a global view of structure altogether, NFV reference architectures have evolved to accept that an “active inventory” and an associated up-to-date view of topology is required (AT&T’s oft-cited ECOMP framework calls these “Available and Active Inventory” and “Resource and Service Topology” – see [http://about.att.com/content/dam/snrdocs/ecomp.pdf](http://about.att.com/content/dam/snrdocs/ecomp.pdf)), and consensus is emerging that this should include legacy assets that are involved in the delivery of services on the “next generation” platform.

So, both in the interest of closed-loop service assurance and in the interest of the broader virtualization effort, this is a challenge that must be overcome. The good news for CSPs is that advances in data technology – including scalable event processing, stream computing and graph data representation – now put such a dynamic, comprehensive view of network and service topology within reach. Indeed, some operators have put such systems in production with great success.

**What can we do right now?**

The fact that extensive measurement infrastructure is widely deployed, effective analytics systems are widely available and proven solutions finally exist to the thorny problem of network and service visibility means that a form of “semi-automatic” service assurance, in which the work of operations staff in assuring services is vastly simplified (and thus cost significantly reduced), is possible right now.

Beyond immediate cost reduction, for CSPs, moving in this direction, is also an incremental step towards fully automatic, closed-loop service assurance and ultimately service orchestration. It’s an opportunity to discover and resolve some challenges early and use the knowledge to de-risk their planning for fully orchestrated NFV.

With the delays that have become apparent in achieving the promised benefits of NFV, opportunities to move incrementally (and bring some of these benefits forward) should be of great interest to CSPs.

Service assurance may well present just such an opportunity.