Independent market research and competitive analysis of next-generation business and technology solutions for service providers and vendors

Next-Gen SON: Automation for Service-Centric Mobile Networks

A custom Heavy Reading report produced in association with Cisco

AUTHOR: GABRIEL BROWN, PRINCIPAL ANALYST, HEAVY READING
# TABLE OF CONTENTS

1. **SELF-ORGANIZING NETWORKS FOR MOBILE RAN** ......................................................... 3
   1.1 SON State-of-the-Art and Business Value ................................................................. 3
   1.2 Defining Distributed and Centralized SON ............................................................. 4
   1.3 What Is Next-Gen Centralized SON? ..................................................................... 5

2. **RAN EVOLUTION, AUTOMATION, AND SON** ......................................................... 8
   2.1 RAN Refresh, 5G NR, and SON for Site Automation .............................................. 8
   2.2 Network Consolation and RAN Sharing ............................................................... 9
   2.3 5G RAN Complexity Drives Automation ............................................................. 9
   2.4 SON and Multi-Vendor RAN ............................................................................. 10

3. **USER-CENTRIC, SERVICE-CENTRIC SON** ............................................................. 12
   3.1 Service Optimization Examples ........................................................................ 12
   3.2 New Data Sources for Service-Centric SON ..................................................... 13
   3.3 Highway Optimization Case Study for User-Centric SON .................................. 16

4. **NEXT-GEN SON FOR PROGRAMMABLE RAN** .................................................. 19
   4.1 SON as Development Platform ........................................................................ 19
   4.2 SON Apps and Business Packages .................................................................. 20

5. **MULTI-DOMAIN END-TO-END SERVICE OPTIMIZATION** ................................. 22
   5.1 The End-to-End Vision ..................................................................................... 22
   5.2 RAN and Core Optimization ............................................................................ 23

**ABOUT CISCO CROSSWORK SON SUITE** ................................................................. 25

**TERMS OF USE** ............................................................................................................. 26

---

Use of this PDF file is governed by the terms and conditions stated in the license agreement included in this file. Any violation of the terms of this agreement, including unauthorized distribution of this file to third parties, is considered a breach of copyright. Heavy Reading will pursue such breaches to the full extent of the law. Such acts are punishable in court by fines of up to $100,000 for each infringement.

For questions about subscriptions and account access, please contact support@heavyreading.com.

For questions and comments about report content, please contact Heavy Reading at reports@heavyreading.com.
1. SELF-ORGANIZING NETWORKS FOR MOBILE RAN

Operators have been growing mobile networks for decades by extending coverage, adding capacity, introducing new air interfaces, and serving ever more demanding use cases. The radio access network (RAN) accounts for the majority of accumulated infrastructure and equipment, with hundreds of billions of investments in equipment worldwide since the first Long-Term Evolution (LTE) deployments only 10 years ago. With mobility embedded in myriad business processes, consumer applications, and public services, these networks are critical national infrastructure.

The rate of innovation in mobile networks is also continuing at pace, with enhancements to LTE and the introduction of 5G New Radio (NR) in the near term and ambitious plans for diverse 5G applications over the medium term. The challenge for operators is to deliver advanced capabilities in a way that is cost-effective and scalable. While the picture is positive from the customer and service perspective, operators must leverage technology to deliver functionality at far lower unit costs.

Operators are also challenged to deliver this at speed. Processes that once merited handcrafted configuration by subject matter experts steeped in radio design expertise must now be inherent to the deployment architecture and operating model. Networks need to auto-configure and self-optimize over multiple time periods. New equipment and software must integrate with accumulated layers of RAN technology deployed in the field, as well as with the installed base of customer devices – without bespoke scripting and manual intervention.

Scale is also entering a new dimension. Whereas previously mobile scale was measured in the tens, or hundreds, of millions of devices per operator, to a few billion devices at a worldwide level, the industry is moving into an entirely new era. As Internet of Things (IoT) processes of all kinds – from connected vehicles to critical industrial processes and low power sensors (and more) – are instrumented and connected, the number of endpoints is exploding, placing enormous new demands on the networks that connect them. To reliably connect vast numbers of devices in a single operator network is an entirely new challenge.

The combination of the scale, speed, and cost points needed to deliver new mobile network capability creates the context for operators’ RAN technology strategies. This Heavy Reading report argues that the economic and technical case for advanced services, delivered over complex infrastructure, is underpinned by operators’ ability to automate operating processes in the RAN and across the wider network. Further, next-generation self-organizing network (next-gen SON) technology is fundamental to this ability.

1.1 SON State-of-the-Art and Business Value

SON technology is already established in mobile RANs and has been shown to generate significant cost and performance benefits. Processes such as the remote configuration of new base station equipment use SON to reduce labor and increase accuracy and speed. More demanding functions, such as dynamic adjustment of radio parameters, are also routinely used to improve operating performance and system capacity.

The data in Figure 1 is modeled on midsize European operators. The figure to the left shows an example of average operational savings achievable using SON for RAN automation and how this is distributed across different processes. The figures to the right show how this
translates into overall business value, including CAPEX deferral and increased revenue due to improved network quality. These benefits are compelling enough that most operators now either use SON products or are evaluating deployments and the return on investment case.

**Figure 1: Average OPEX Savings from SON Automation and Overall Business Value**

![Figure 1](image1.png)

*Source: Cisco, Heavy Reading*

### 1.2 Defining Distributed and Centralized SON

Mobile SON can be defined as a software platform operating under a network management layer across the following set of functions: self-configuration, self-optimization, and self-healing. Example SON functions are shown in Figure 2.

**Figure 2: Major SON Functions**

<table>
<thead>
<tr>
<th>Self-Configuration</th>
<th>Self-Optimization</th>
<th>Self-Healing</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Neighbor Cell Lists</td>
<td>• Coverage &amp; Capacity Optimization (CCO)</td>
<td>• Fault Detection &amp; Classification</td>
</tr>
<tr>
<td>• Radio Parameter Planning</td>
<td>• Mass Event Handling</td>
<td>• Cell Outage Compensation (COC)</td>
</tr>
<tr>
<td>• Site Configuration</td>
<td>• Traffic Optimization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Capacity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Load Balancing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Mobility Load Balancing (MLB)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Mobility Robustness Optimization (MRO)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Energy Saving</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Interference Management</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Heavy Reading*

SON is also defined horizontally from distributed SON (D-SON) at the radio base station level to centralized SON (C-SON) that coordinates a greater number of cells and users across a wider coverage area. D-SON operates on short cycle times and is integrated with the radio base station. Thus, it is vendor-specific and not a discrete product. C-SON is an independent solution that can be, and ideally should be, from a different vendor to the RAN equipment.

Heavy Reading expects operators to continue to use a combination of distributed and centralized SON because this produces superior outcomes. Interference management, for example, can be addressed by using D-SON to make quasi real-time changes using inter-cell interference coordination (ICIC) techniques. However, these techniques are limited to
neighboring cells. If interference is caused by a more distant over-shooter cell (perhaps due to inappropriate antenna tilt or power settings), C-SON is needed to coordinate interference management across the wider area. There are many examples of this type. In each case, coordination between D-SON and C-SON is required.

This report focuses on C-SON. These systems are gaining in functionality and offer greater networkwide impact. Moreover, as C-SON solutions gain greater ability to abstract control from access equipment, they are moving from specialist radio products to becoming more horizontal RAN programming platforms with a greater role in end-to-end network automation.

1.3 What Is Next-Gen Centralized SON?

Despite its widespread use and proven value, it is not controversial to say that SON has underdelivered relative to its potential and that it remains an underexploited technology. Generally speaking, SON has been used in silos for important, but arcane, radio optimization tasks that often are not captured in how operators measure customer satisfaction or end-to-end service performance. D-SON, in particular, is a feature of RAN equipment and is not directly linked by service experience monitoring.

Far more can be achieved with a new generation of SON solutions designed with service outcomes in mind. In this report, Heavy Reading identifies and discusses the four key pillars that define next-gen SON:

- Zero-Touch RAN Optimization and Automation
- User-Centric and Service-Centric SON
- Programmable Network Platforms
- Multi-Domain, End-to-End Services

Machine learning (ML) technology reaches across each of these pillars.

Zero-Touch RAN Optimization and Automation

The evolution of the RAN is happening at rapid pace, with the upgrade to 5G setting the agenda. Denser topologies, new architectures, and new operating requirements are forcing change and, in turn, driving complexity. A 5G radio, for example, has around 1,000 configurable parameters versus 300 for LTE. Massive multiple-input and multiple-output (MIMO) radios introduce new optimization requirements, and dual connectivity between LTE and 5G NR demands tight interworking. In addition, more stringent expectations for power efficiency, reliability, and redundancy are forcing a rethink of established operating models. Even more disruptive is the introduction of millimeter wave (mmWave) for mobile use, which is a new order of challenge.

In the near term, narrowband-IoT (NB-IoT) and LTE-M add new classes of device and new performance requirements to the 4G/5G mobile network. SON tools must adapt to the link budget, density, and power consumption implications of IoT in cellular networks and be ready to evolve to serve the new massive IoT specifications currently under study in 5G.
Keeping pace with changes in the RAN is more than enough to keep SON vendors and operators busy – it is an immense multiyear challenge. Yet, just keeping pace is not sufficient. There are many more opportunities to use next-gen SON to make mobile networks more automated, programmable, and optimized for discrete service types.

**User-Centric and Service-Centric SON**

SON is traditionally used to improve network key performance indicators (KPIs) such as call drop rates (CDRs), mobility handover success rates, downlink date rates, and so on. In next-gen SON, the aim is to take this forward to be able to measure, and assure, key quality indicators (KQIs) at the user and service levels. This can be applied to service types (e.g., mission-critical services such as public safety), discrete use cases (e.g., highway or railway performance optimization), specific user groups (e.g., key enterprise accounts), or consumer segments such as Net Promoter Score (NPS) detractors. Example KQIs targeted at these groups include video streaming performance (based on user device type/resolution/buffering), voice quality (drops, silence, repeated calls), and web browsing experience (set-up time, data rate).

To deliver user-centric performance optimization based on KQIs, operators will need to integrate SON with new data sources, such as crowdsourcing, geo-location, CDRs, and call traces. Applying ML to these alternative data sources can reveal simple and cost-effective insight and optimizations that were not previously available. To be truly user- and business-centric, integration with the operator’s customer experience management (CEM) system may also be useful.

**Programmable Network Platforms**

To automate mobile networks, it must be possible to program functional elements (e.g., base stations) via application programming interfaces (APIs). Next-gen SON provides an environment that operators can use to create software modules and algorithms specific to their networks or use cases. In this way, C-SON becomes a development platform that integrates with the network elements downstream and higher layer orchestration and management tools upstream.

In the downstream direction toward the RAN, initiatives such as the Operations Support Systems Interoperability Initiative (OSSii) are important to interoperability with RAN vendor equipment. In the future, initiatives such as xRAN and the O-RAN Alliance will make radio equipment more directly programmable, enabling more granular control by third parties. Toward the transport network, the well-established NETCONF/YANG workflow can be used. Toward the network management layer, integration with orchestration tools such as Open Network Automation Platform (ONAP) or Open Source MANO (OSM) can help enable end-to-end service orchestration. There are many implications of this for the data processing architecture, especially in large networks, which generate very large volumes of data. ML is expected to be influential in addressing this.

**Multi-Domain, End-to-End Services**

SON in the RAN is invaluable, but service quality is determined by end-to-end performance. One of the major opportunities is to extend SON concepts into the mobile core, the transport network, and into the cloud domains where applications are hosted. Extending SON into multi-domain service optimization is important to enabling services such as
network slicing and assuring performance against service-level agreements (SLAs). This is part of making today’s LTE networks 5G-ready and will have a long-term impact on 5G services themselves. Again, this will involve the integration of network telemetry systems with SON and the overarching network management layer.

The starting point is likely to be RAN-core optimization in which the core network adapts to session-level traffic flows, according to factors such as RAN load, congestion, and policy. This is immediately valuable because if the performance issue is very temporary, it is suboptimal to use SON to make hard changes (e.g., antenna tilts) to the RAN, whereas the core network can react in almost real time to issue in the IP layer.

Operator organization and structure are also important for end-to-end optimization. Technology changes invoked by SON need to be supported by business process changes. This is a greater challenge for larger operators with more distributed teams than for smaller, nimbler operators.

**Machine Learning and Next-Gen SON**

ML is the ability for computers to learn without being explicitly programmed. It can identify functional relationships between data that are not easily identifiable by humans or via standard analytics tools to produce algorithms that can be put to practical use. These characteristics make the technology useful to network automation, and ML is emerging as a defining feature of next-gen SON.

Network operators have vast data sets (e.g., derived from probes, telemetry, CDRs, and so on) that are underutilized. Applying ML to this data can reveal insights that can be used to make decisions about how to configure or adapt the network to improve performance. In this sense, it adds a layer of intelligence to the network management and automation toolsets already in use and under development.

Mobile networks exactly fit this profile of having rich data sets from multiple domains. ML can derive decision-making algorithms from these data sets and then apply them to the network. This is valid across RAN, core, transport, and cloud domains. In some cases, ML algorithms will be centrally trained, but distributed to the individual network element or cluster level for the live deployment. For example, RAN optimization algorithms trained on large volumes of aggregated data could be deployed in the C-SON system to make centralized decisions and/or (in the radio base station itself) to make more localized optimizations.

As operators build trust in ML technology, it will be extended to closed-loop decision-making. As more network and customer experience data is generated, so the ML system will evolve, and so the next-gen SON can implement changes autonomically. One of the challenges of using ML for closed-loop automation is to trust the technology even when it proposes unorthodox changes; for this reason, human oversight will be important for some time.
2. RAN EVOLUTION, AUTOMATION, AND SON

The RAN is in continuous evolution, with nearly all operators using a mix of legacy and new equipment and, at the same time, preparing for the next generation. Across a network, deployment and operations are never static. Therefore, from a SON perspective, it is useful to distinguish between what is possible on deployed equipment and what can be done with new equipment soon to be installed.

2.1 RAN Refresh, 5G NR, and SON for Site Automation

The greatest opportunities for a step change in RAN functionality come with the deployment of new, more capable base stations. This occurs continuously in most reasonably sized networks; however, the biggest deployments occur at the once-a-decade transitions to new a generation of RAN. Right now, this means the worldwide upgrade to 5G. Heavy Reading estimates over 4 million 5G-enabled base stations will be deployed over the next 5 years, as shown in Figure 3. Many of these will be deployed in C-Band and mmWave spectrum.

Figure 3: Forecast of Global Number of 5G-Enabled Base Stations (Cumulative)

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macro Cells</td>
<td>7,900</td>
<td>93,200</td>
<td>254,400</td>
<td>712,400</td>
<td>1,655,000</td>
<td>3,101,200</td>
</tr>
<tr>
<td>Outdoor Small Cells</td>
<td>0</td>
<td>66,800</td>
<td>198,800</td>
<td>423,600</td>
<td>718,300</td>
<td>1,146,900</td>
</tr>
<tr>
<td>Total 5G Base Stations</td>
<td>7,900</td>
<td>160,000</td>
<td>453,200</td>
<td>1,136,000</td>
<td>2,373,300</td>
<td>4,248,100</td>
</tr>
</tbody>
</table>

Source: Heavy Reading, December 2018

At the same time as deploying 5G, operators will also seek to refresh their LTE estate or add new LTE bands to the site – or do both at the same time. This is a common strategy for operators in markets deploying advanced LTE solutions ahead of 5G introduction, or where they are installing new 5G equipment. In each case, the intent is to later upgrade new LTE equipment to be 5G capable as 5G device penetration increases.

This strategy is helped by the fact that new LTE base station radio units are software upgradable to 5G NR. For example, operators investing in advanced LTE today, particularly in mid-band Time Division Duplex (TDD) spectrum, such as 2.3 GHz, 2.5 GHz, or 2.6 GHz, are deploying massive MIMO radios that can be upgraded to NR and/or can provide dual-mode NR and LTE service simultaneously via dynamic spectrum sharing. New low band Frequency Division Duplex (FDD) deployments, such as 600 MHz and 700 MHz, can also operate in dual-mode LTE-NR, albeit with more traditional 2x2 MIMO antenna systems.

SON can be extremely helpful in the deployment phase for site automation. Configuring new equipment as it is deployed via SON algorithms linked to the RAN vendor element management systems saves time, removes human errors, and improves configuration consistency across the different coverage areas. Equally important is the opportunity to simplify the whole site reconfiguration that comes with any new deployment or upgrade, which includes the ability to reconfigure legacy equipment and the need to commission and configure the transport connection to the site. This is an example where there is value to SON outside the classic radio parameter optimization.
2.2 Network Consolidation and RAN Sharing

There are already around 7 million physical base station sites worldwide and a considerably higher number of logical sites. On refresh rate alone, there is high demand for site automation. SON can also assist with network consolidation and network sharing in the installed base.

As operators plan major investments in 5G, they are, in some markets, revisiting network sharing. This typically involves either setting up via joint venture companies to run shared infrastructure or managing equipment on the partner’s behalf. In both cases, experience in 4G network shares shows that independent C-SON is needed for visibility and control of site configurations and radio parameters to meet SLAs – for example, by ensuring each operator’s RAN configuration is accurately implemented by the joint venture or host operating company.

Network consolidation is also a critical use case for C-SON. The ability to rapidly and cost-effectively consolidate networks following mergers and acquisitions (M&A) can make or break the economic success of these huge transactions and is one of the ways that C-SON can deliver very large and very fast return on investment. For example, using SON tools to identify sites for consolidation and then reconfigure different RAN vendor equipment to the new parameters can accelerate these programs and improve the quality of the consolidated network through consistent and accurate application of design templates.

2.3 5G RAN Complexity Drives Automation

5G RAN is, by design, more flexible and configurable than LTE because it must serve diverse use cases with diverse parameter optimizations. What is required for high speed rail, for example, is different to a stadium or factory floor environment. A consequence of the need to support multiple use cases and deployment types – a key part of the 5G value proposition – is that it generates the potential for a massive increase in complexity. Addressing complexity in the deployment phase while operating a live 5G network is beyond the capabilities of traditional RAN operating processes. Doing so requires new automation techniques, with decisions driven by predictive analytics and ML.

Some of the major ways 5G RAN differs from classic RAN are as follows:

- **Hyper dense deployments:** To meet the area traffic density requirements of 5G in urban scenarios (e.g., the 3GPP target for dense urban is 125 Gbit/s/km²) will require densification and small cell deployment. One effect of this is the need to mitigate inter-cell interference in dense cell clusters through dynamic control of radios across a coverage zone and perhaps across multi-vendor RAN zones.

- **5G RAN architectures:** The classic macro RAN architecture remains the dominant deployment model but is being enhanced by disaggregation. The base station itself will be split into radio unit, distributed unit, and centralized unit (RU, DU, and CU) modules, creating opportunities for new deployment and operating models, such as cloud radio clusters managed by centralized CUs and optimized by SON. This will also offer new management interfaces that may be leveraged by next-gen SON systems.

- **Dual connectivity LTE-NR:** To accelerate 5G, the first services will be deployed using a 4G core and an LTE control plane. There are many variations to this dual connectivity model, according to factors such as device penetration, service strategy,
and so on, that affect the migration to standalone 5G core. The implication for SON is that the ability to manage tightly integrated LTE and NR RANs is needed immediately and will be important for the long term.

- **Massive MIMO:** Probably the single biggest recent change is cellular radio design is the introduction of multi-user massive MIMO and the related move into TDD spectrum. The beam-forming, beam-steering, and beam-tracking techniques used to allocate capacity to users in LTE and NR introduce many new optimization vectors and represent a new order of challenge for SON.

- **Mobile mmWave:** The first live 5G services in the U.S. have been deployed in mmWave spectrum. This is a radical change from classic cellular system design, and it requires new design templates and optimization toolsets. In the not too distant future, mmWave will also be used for integrated access and backhaul (IAB) devices, bringing requirements for dynamic mesh optimization to SON.

- **Lean design, ultra-reliable low-latency communications (URLLC), and massive IoT:** A feature of the 5G NR interface is that it was developed using lean design principles that allow it to meet immediate requirements and be upgraded with capabilities over time. Examples might be for ultra-reliable services or massive IoT. This means any SON system must also adapt as these features are standardized and come to market.

### 2.4 SON and Multi-Vendor RAN

The three largest RAN vendors control around 75% of the base station market, and each provides its own management and optimization tools. This works reasonably well for baseline optimization features. However, it is can be problematic for operators that need to modify and optimize RAN to their own circumstances and needs, or where they develop and deploy RAN configurations they want to apply networkwide across different vendor equipment.

A critical point is that there is only limited interoperability between RAN vendor base station equipment. This is important where different vendor equipment meets at RAN borders – for example, at geographic coverage boundaries, across national borders, or in network sharing arrangements. The X2 interface is sufficiently standardized to enable handover for mobile devices. However, the only practical way to support inter-vendor RAN management and optimization (e.g., to manage interference) in these scenarios is via a C-SON solution or via manual intervention. Typically, C-SON from an independent third party is preferred.

Independent C-SON is also useful for faster time to market. Operators know their own networks and can design configurations, and even develop custom algorithms, that are better suited to their needs. Working through the timeline of a large OEM vendor to secure such a capability is often a time-consuming process, if the feature request gets developed at all. Moreover, RAN vendor optimization tools are often sold as professional services that generate recurring revenue for the vendor. Multi-vendor C-SON therefore gives operators flexibility and independent control of their network. For some, this is of critical importance.

The major way SON vendors interact with RAN vendor equipment is via OSS interfaces brokered by the OSSii. This is an initiative started by Ericsson, Huawei, and Nokia (the top three RAN vendors) to enable interworking between their OSS and third-party systems to reduce integration costs. Functionality covered includes northbound interfaces for
configuration management (CM), fault management (FM), performance management (PM), event/traces, and management and orchestration (MANO; VNF Manager-NFV Orchestrator [VNFM-NFVO]).

OSSii is useful to independent software vendors (ISVs) because it allows them to create global agreements with the three RAN vendors. Without it, independent C-SON would need custom agreements in each market, likely brokered by operators, leading to inefficiency. There are mixed reports of how enthusiastically the RAN vendors support the initiative (Ericsson is generally said to be the easier of the three to work with; Huawei the most difficult). Generally, however, OSSii is viewed as valuable by operators and essential to independent C-SON vendors.

Another issue is that the OSSii interfaces are at the OSS level into the RAN vendor’s element management system. This restricts how frequently SON can make changes and how fast they are implemented in the RAN. This is one reason why 15-minute cycle times are typical for centralized SON and why third-party D-SON is not practical.

How third-party SON and base station equipment interact may be starting to change. Recently, the xRAN Forum has defined the use of the Internet Engineering Task Force’s (IETF’s) NETCONF/YANG standards to programmatically configure the operational state of radio base stations (specifically DUs and RUs in an Open RAN architecture). The use of YANG models will improve multi-vendor interoperability in the disaggregated RAN and simplify integration into existing network management systems. This work has been moved into the newly formed O-RAN Alliance for ongoing development. It will also be affected by work on RAN architecture in 3GPP. One area of interest is in scenarios where virtual CUs are deployed on edge compute nodes to control more distributed DUs and RUs. In this case, it may make sense to run SON systems at the same edge location, with direct access to the CU (a type of radio controller). Among other things, this would enable ISVs to develop a kind of edge SON version of D-SON.

Cisco Crosswork SON replaces complexity with simplicity. It transforms your network into an elastic, intelligent infrastructure that empowers you to run your RAN automatically, dynamically, and in real time.

Cisco Crosswork SON can help you:

- Reduce drop call rates by up to 30%
- Increase data capacity by up to 20%
- Add new sites in hours without disrupting normal operations
- Reduce operating and capital expenditures by saving on engineering and power resources
- Postpone network expansion by better utilizing spectrum and equipment
- Enable the launch of new services such as narrowband IoT (NB-IoT) and fast deployment of new 5G cells
- Broaden your choice of vendors to help drive down costs

– Sponsored content provided by Cisco
3. USER-CENTRIC, SERVICE-CENTRIC SON

SON is traditionally used to improve network KPIs such as CDRs, mobility handover success rates, connection set-up and failure rates, and so on. This is necessary and provides useful control over RAN performance from the perspective of the operator. However, it does not always accurately reflect the subscriber experience or how well a service is performing. One result is the well-known syndrome that “everything shows as ‘green’ on the console, so we don’t have problem,” when in practice key user groups or services are not experiencing sufficiently good performance. For example, stationary users may experience excellent data throughput, but mobile users in vehicles might experience frequent call drops. In next-gen SON, the aim is to combine network-derived KPIs with other data sources to be able to measure, and assure, KQIs at the user and service levels.

There are several advantages to this approach:

- **Enable better visibility and understanding** of user quality of experience (QoE), which gives a more accurate view of network performance and allows for more useful and actionable benchmarking.
- **Accelerate identification of performance issues** and automatically resolve them, instead of waiting for customer complaints or relying on infrequent and expensive drive test data.
- **Retain customers through improved experience.** Reducing churn and customer acquisition costs and improving net promoter scores has a major impact on mobile operators’ top and bottom lines.

3.1 Service Optimization Examples

Many user and service types can be identified as having special requirements or deserving of attention. **Figure 4** identifies service types that are a focus for operators in today’s LTE networks and have been shown to warrant investment in optimization by producing measurable improvements in user experience. Typically, operators do not address all of these at once, but select one or two lead use cases to address first.

**Figure 4: Examples of Service-Centric Optimization**

<table>
<thead>
<tr>
<th>Service</th>
<th>Importance of Service Performance</th>
</tr>
</thead>
</table>
| Mission-Critical Services| May include public safety or other critical process  
|                          | E.g., traffic lights, tsunami warning, infrastructure monitoring  
|                          | Cost of failure is high; requires prioritization over standard users                               |
| Enterprise User Groups   | High value corporate clients  
|                          | Must meet and report on demanding SLAs                                                             |
| Roadside/Highways        | Need to improve in-car smartphone experience  
|                          | To improve connected car performance longer term                                                  |
| Trains                   | Railways often do not have dedicated cellular infrastructure deployed  
|                          | High speed affects performance; demand can be very bursty                                          |
### Importance of Service Performance

<table>
<thead>
<tr>
<th>Service</th>
<th>Importance of Service Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPS Detractors</td>
<td>Subsets of users with poor service should be identified early</td>
</tr>
<tr>
<td></td>
<td>Seek to retain high average revenue per user (ARPU) users and offset damage from brand detractors</td>
</tr>
<tr>
<td>NB-IoT and LTE-M</td>
<td>Very different coverage and battery life constraints</td>
</tr>
<tr>
<td></td>
<td>Require optimized RAN configuration; can affect smartphone users</td>
</tr>
<tr>
<td>VoLTE</td>
<td>Significantly different quality of service and mobility parameters to data services</td>
</tr>
<tr>
<td></td>
<td>Optimize coverage and VoLTE call duration before fallback to circuit-switched voice</td>
</tr>
</tbody>
</table>

*Source: Heavy Reading*

One common opportunity for service-centric network optimization is voice over LTE (VoLTE). LTE networks were originally deployed and configured for broadband access, so when voice is added, a deep re-optimization is necessary to ensure the service is as reliable as circuit-switched voice. After this redesign, SON is used to maintain service levels with as much automation as possible. Another example is NB-IoT, which is repeating this pattern of a major new service type being deployed on the LTE platform, with distinct performance requirements related to coverage, density, and battery life.

### 3.2 New Data Sources for Service-Centric SON

As noted, to know how a service performs is not the same as knowing how the network performs. To build a more complete picture, next-gen SON should use supplementary data sources in combination with the data derived from the RAN vendor OSS. This is one of the more important aspects of a user-centric and service-centric approach to network optimization. Figure 5 identifies three major categories of data that can be used for this purpose.

*Figure 5: Data Sources for Network and Service Benchmarking*

<table>
<thead>
<tr>
<th>Crowdsourcing</th>
<th>Drive &amp; Walk Test</th>
<th>Network Data &amp; OSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowdsourced app specialists</td>
<td>Specialist equipment</td>
<td>Analysis of call flow data</td>
</tr>
<tr>
<td>Operator custom care apps</td>
<td>Dedicated vehicular systems</td>
<td>Probe data &amp; telemetry from network equipment</td>
</tr>
<tr>
<td>SDKs integrated with third-party apps</td>
<td>Backpack systems for walk test</td>
<td>Advantages: Provides end-to-end view; integrates with OSS for actuation</td>
</tr>
<tr>
<td><strong>Advantages:</strong> Scale and reach; can be low cost (not always)</td>
<td><strong>Advantages:</strong> Reliable, normalized data; traditionally considered “real” benchmarking</td>
<td><strong>Disadvantages:</strong> Doesn’t show radio performance in real-time; expensive to instrument a full network</td>
</tr>
<tr>
<td><strong>Disadvantages:</strong> Affected by variables not related to network performance; hard to normalize data</td>
<td><strong>Disadvantages:</strong> Expensive if used extensively; gives very infrequent updates on network performance</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Heavy Reading*
Crowdsourced User Equipment Data

Data collected directly from user equipment (UE; a.k.a. handsets) – so-called crowdsourced data – gives a useful and compelling view of how customers actually experience network service. There are many variations of this method and many companies involved in collecting this data. This model requires an app to be installed in the user space on the smartphone, with the better-known examples being services such as speedtest.net. Another common approach is for operators to deploy their own benchmarking apps or to integrate data collection into the customer care apps their customers routinely use to check their balance, usage, and so on.

Integrating data collection into other apps is enabled by software development kits (SDKs) available from several suppliers. Such kits have permitted new entrants into this market – developers can now integrate data collection into a shopping app, a game, or a weather app, for example. Sometimes this makes good sense. A large enterprise, for example, may wish to collect (or contract a third party to collect) data from its users so it can monitor its SLA with the service provider. In other cases, there are questions about user consent, especially where “fly-by-night” and “cowboy” developers, with less trusted brands, are involved.

Generally speaking, high quality crowdsourced data is useful. For example, it can identify coverage gaps and geolocate areas where service performance is markedly worse than it should be. The SON system, armed with this information, can then make changes to offset the problem or minimize its impact until new capacity is deployed or hard changes are made to the RAN. However, it is not a panacea. The biggest problem is deriving normalized data because crowdsourcing depends on the performance of the individual model of smartphone (and the antennas, modems, and processors within it). Therefore, different readings are found even among users of the same services at the same location. Another problem is having statistically significant sample sizes. While providers of this data may have large global numbers of data sources, they may not have enough insight into a specific network in a sufficient number of locations. In summary, crowdsourced data is a valuable input into SON when it is of high quality and used appropriately.

Drive and Walk Testing

Drive and walk testing are the classical benchmarking tools and remain very important to operators because they generate normalized data sets on which analysis and action can be applied. They are, however, expensive. Use of the drive test has declined significantly in recent years as a consequence. Operators now tend to drive test tactically where they think they have a problem, where they have deployed new equipment and need to verify performance, or where they have consolidated networks following M&A. Walk tests remain invaluable for a view of indoor and urban pedestrian performance but have many of the same challenges.

From a SON perspective, the infrequency of fresh data from drive and walk tests limits its usefulness to make short cycle time changes; however, it is a welcome source of insight to verify how well major changes are working or to identify major problems. Independent benchmarking firms such as P3 and Root Metrics continue to use drive tests for these reasons.
Network Data, Call Trace, and CDRs

Network-derived data from sources such as probes, CDRs, or call traces can be presented to the SON via network analytics systems and used to augment analysis and decision-making. Probes and call trace data, like drive tests, are traditional and useful and should be part of the next-gen SON solution, but again, come with costs attached. Instrumenting a network with probes can be expensive, for example. Although virtualization and streaming telemetry is addressing this, it remains a work in progress. Similarly, running large numbers of call traces and processing the results can also be expensive – prohibitively so at scale – and this is generally best used in a targeted manner as part of an optimization campaign or to troubleshoot problem areas already known about. Again, these are valuable data sources when used appropriately.

CDRs are also interesting. They do not present real-time data. Typically, CDRs generated by mobile core equipment are batch-processed daily by mediation systems – and are often overlooked as source of value to SON. However, new data analysis techniques, and ML in particular (as discussed in the highway optimization case study in Section 3.3), can make this data a source of useful insight that can enhance SON decisions. Figure 6 shows how CDRs can contribute to and enrich a next-gen SON solution.

Figure 6: SON Enrichment Using CDR Data Sources

Source: Cisco

Machine-Learning Engine: The machine-learning engine enables Cisco Crosswork SON to become a proactive solution that can predict network performance and quality of experience degradation before it happens. It can identify root causes and recommend optimal actions to take before user experience is impacted. Network optimization and user experience continually improve as the machine-learning capabilities acquire more knowledge. Couple this advanced capability with our user-centric SON ability to manage user experience from device to cloud and SONFlex Studio drag-and-drop ease to quickly create new SON RAN automation applications and you have a greater competitive edge.

– Sponsored content provided by Cisco
3.3 Highway Optimization Case Study for User-Centric SON

Passengers and drivers on major highways make up one of the most important user groups for mobile operators. It is a scenario where mobile technology uniquely excels, and performance in vehicles is an influential factor in how consumers select their operator. It often warrants a dedicated category in national RAN benchmark tests published in widely read consumer publications (the German *connect* magazine test is probably the most famous example in terms of the weight it assigns to highway performance). In short, this is a strategic market with opportunities for differentiation. However, highways are challenging to serve effectively, and performance often falls short of customer expectations. But where operators start to focus on this use case, it can quickly become an important determinant of market share.

There are several reasons why highway performance is poor even if the road is ostensibly covered by the RAN. In a typical highway scenario, particularly around urban areas, it is not uncommon for 70% of users in the coverage area to be classical stationary users at home, on the street, and in the workplace and for 30% to be active mobile users in vehicles. Figure 7 shows such an urban highway scenario and how dropped call rates (DCRs) in a cell can vary between these two groups, with a DCR (noted as DR in the figure) of 0.4% for highway users versus an average cell DCR of 0.15%. The challenge is that this difference is not revealed in standard network monitoring. At this 70/30 user ratio, KPIs generally report that the network, in aggregate, is performing acceptably – even if vehicular users do not receive good service.

**Figure 7: Variability in Drop Call Rates in a Cell**

![Figure 7: Variability in Drop Call Rates in a Cell](source: Cisco)
An innovative way to reveal differences in experience between the two user groups is to analyze CDRs collected from sessions within the target coverage area, as shown in Figure 8. This can reveal user experience expressed in the language of call set-up times, data rates, speech quality, handover success rates, and so on.

**Figure 8: CDR Data – Capturing the Connected Journey**

![CDR Data – Capturing the Connected Journey](source: Cisco)

CDR data, as noted, is already available to operators and is non-invasive in that it does not require additional probes to be deployed (although a mediation platform is helpful). The challenge is to analyze the many millions of user events contained within CDR records to derive a view of performance. Big Data techniques, and specifically ML, can be used to interpret such large data sets and turn what is often considered low value data into useful insight.

With insight from the CDR data set revealed by ML algorithms in this highway optimization example, the SON system can take remedial actions such as the following:

- Enhance mobility optimization by adjusting handover thresholds and identifying specific cell groups as preferred handover clusters.
- Identify remote interference from boomer cells not optimally placed to serve the highway and manage their impact via antenna tilt and power controls.
- Improve throughput for highway users via cell load balancing, antenna tilts, etc.
Figure 9 shows the improvement in performance across a series of highways in a European country from a 4-week optimization campaign. This is a real-world example validated with the operator by Heavy Reading. It shows the results of a sizeable optimization campaign. The beauty of SON is that once the campaign is complete, ongoing adjustments can be made, on a largely automated basis, as the network settles down, demand patterns evolve, or new sites are added to the coverage area.

**Figure 9: Highway Optimization Success**

<table>
<thead>
<tr>
<th>Scope of Highway Campaign</th>
<th>Measurable Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Weeks</td>
<td>Average 3G Retainability Highway Commuters: -20%</td>
</tr>
<tr>
<td>7 Highways</td>
<td>Average 3G Voice DCR (KPI): -15%</td>
</tr>
<tr>
<td>2400 Cells UMTS &amp;LTE</td>
<td>Average 4G Retainability Highway Commuters: -38%</td>
</tr>
<tr>
<td>3000 Closed-Loop User-Centric Actions</td>
<td>Average 4G ERAB DCR (KPI): -7%</td>
</tr>
</tbody>
</table>

Source: Cisco
4. NEXT-GEN SON FOR PROGRAMMABLE RAN

To automate mobile networks, operators need programmatic control of the underlying functional elements that combine to deliver the end-to-end service. From a SON perspective, this requires APIs downstream to the RAN equipment and upstream to the management and orchestration systems. These open principles are reasonably well-established in software-defined networking but are still relatively new in radio because RAN vendors retain direct control of their equipment via proprietary interfaces. There is an opportunity to change this and extend the benefits of automation to the mobile access domain.

4.1 SON as Development Platform

Next-gen SON provides an environment that vendors, operators, and ISVs can use to create software modules (e.g., radio optimization algorithms, site configuration scripts, and so on) specific to different networks or use cases. In this sense, next-gen SON can be used to develop and execute RAN scripts, as a component of the wider mobile network automation platform. The concept is shown in Figure 10.

Figure 10: Open SON Automation Platform

In the downstream direction toward the RAN, the SON connects to via the RAN vendor’s OSS, as discussed above. In the future, as Open RAN initiatives take hold, it will be possible for a third-party SON to interact programmatically with radio equipment. The SON will also be able to consume data from other sources – for example, the network analytics and ML systems discussed earlier – to generate algorithmic instructions to network equipment. Toward the network management layer, integration with orchestration tools such as ONAP or OSM is possible. It may also be useful to integrate with customer care dashboards.

The actual functions of the SON are determined by the applications (scripts) it runs, shown as “SON Apps” in Figure 10. These can be designed for 2G, 3G, 4G, and 5G networks, as
required. A library of such apps can be created to create new functionality to meet business needs.

### 4.2 SON Apps and Business Packages

SON apps can be developed by the SON vendor and integrated into the SON platform by the operator using internal resources or by external ISVs. These apps range from fundamental SON functions (e.g., automatic neighbor relation [ANR]) to higher level business packages targeted at specific use cases. Apps could one day be portable across networks and SON platforms.

Figure 11 shows three different categories of SON apps distinguished by size and importance. This example refers to the Cisco SONFlex and SONFlex Studio product; several other SON vendors have the same kind of ideas in their roadmaps.

**Figure 11: Increasing RAN Automation with Programmable RAN**

In this analogy, the jar represents the operator network; the bubbles are the SON apps. To the left are the major SON functions represented by the large bubbles. These are the key things SON should do and cannot be without – for example, ANR and self-healing in the event of cell outages. They are widely applicable apps and are generally integrated into the SON platform. This example generates an automation score of 45%.

The middle jar shows apps tailored to the needs of a specific operator. These are important features that operators want to specify and code themselves so that they can have unique capabilities suited to their network and do not have to wait a long time for vendors to develop their feature requests. This example generates an automation score of 65%. It is a particularly interesting example because mobile operators have in-house radio experts, dedicated to design and optimization, who until now have not been able to code and deploy their own apps into the SON platform. Therefore, they have been restricted to manual scripting and manual deployment of algorithms. Examples might include energy-saving...
scripts matched to the operator’s demand patterns, or low and high frequency load balancing designed according to the operator’s unique spectrum assets.

This ability to write and deploy their own SON apps is important as operators move toward a continuous delivery/continuous integration (CI/CD) approach to RAN. However, it requires operators to be able to not only design their algorithm, but also code it. This mix of radio and coding expertise remains relatively scarce among vendors and operators. Another challenge is that these apps can become quite compute intensive and may consume excessive server capacity; in these cases, it may make sense to integrate complex algorithms back into the base platform over time.

The jar to the right refers to smaller apps that fill in the gaps not addressed by the other methods. In this case, it uses a graphical user interface (GUI)-based tool, with a drag-and-drop style environment, to create custom scripts for non-programmers, such as radio or transport engineers. This example brings the automation score to 80%. An example app might be to create new site configurations that consolidate scripts from different departments involved in commissioning new sites.
5. MULTI-DOMAIN END-TO-END SERVICE OPTIMIZATION

Service quality, as experienced by customers, is determined by the end-to-end performance of the network and by external factors such as device type and application performance. Operators should therefore look at how SON can be incorporated into a multi-domain, service-aware optimization system. This shifts the challenge of performance monitoring and optimization from being focused on individual network domains to focusing on the user’s session experience.

It also extends SON into the network slicing discussion and how to assure performance against SLAs. This is a longer-term role for SON as standards, definitions, and the implementation of network slicing are worked out, initially in 4G and then more formally in 5G. However, if operators are to deliver services that support advanced, business-critical processes, it will be necessary to associate measurable SLAs to the network slice. To commit to SLAs in a mobile network requires control over the RAN and thus a C-SON solution.

The operator’s internal organization is also important. Operators frequently express their desire to have end-to-end visibility and control, but also admit that their organizational structure, with different teams for different network domains and different customer segments, can make this difficult to achieve. This dictates a step-by-step approach for end-to-end optimization and underlines the need for executive level support within the operator, starting with the chief technology officer.

5.1 The End-to-End Vision

The vision itself is compelling. Local orchestrators work, in near-real-time cycles, to optimize the RAN, transport, core, and cloud domains. Meanwhile, a network management layer runs across domains such that it can address a detected problem in one domain by modifying configuration in another to achieve end-to-end optimization. This concept is shown in Figure 12. Because the RAN is often (but not always) the most capacity-constricted part of the network and most subject to variable performance due to load, mobility, interference, etc., next-gen SON is critical to the overall vision of end-to-end service optimization.
5.2 RAN and Core Optimization

A starting point for multi-domain optimization is between RAN and core. Both these domains are defined and standardized by 3GPP and have shared interest and joint participation in managing user sessions. To set up, operate, and tear down a bearer/session in LTE or 5G, for example, requires multiple interactions between the RAN and core at the control plane and user plane levels. This makes RAN-core interaction the logical starting point for multi-domain optimization.
Figure 13: RAN-Core Network Optimization

Figure 13 shows how this works in principle. The SON detects cell-level congestion and the core detects flow level congestion. Both report to the SON, which can then take action in both domains. For example, in the RAN by load balancing or adjusting power controls before radio KPIs deteriorate. Or in the core by taking actions such as rate adapting video streams when it knows RAN performance is constrained.

Another benefit is the core can react in almost real time, whereas the C-SON typically requires 15-minute cycle times. This is particularly useful when the source of degraded performance occurs outside the RAN – for example, on the internet or in the cloud environment that hosts the service. In these cases, which are not uncommon and are often transient, adjusting the RAN via C-SON is overkill, whereas the core operates on much faster cycle times.

As operators gain experience with RAN-core optimization and achieve measurable success, they can start to extend the concept into the transport and cloud/application domains.
ABOUT CISCO CROSSWORK SON SUITE

The information in this appendix was provided by Cisco.

Cisco Crosswork SON is a suite of leading multi-vendor multi-technology (MVMT) self-organizing network (SON) solutions. This is a powerful platform that uses machine learning and a set of applications that automates the Radio Access Network (RAN). Cisco Crosswork SON enables service providers with:

- **5G Readiness** – The Cisco Crosswork SON platform is fully virtualized, microservice based platform, with E2E viability and control over the mobile network.

- **Multi-Vendor, Multi-Technology, Multi-Spectrum, Multi-Architectures** – Cisco Crosswork SON simplifies network densification, single point for any network RAN vendors, technologies, and spectrum mixes, Crosswork SON platform automates it all like a unified single network.

- **CapEx and OpEx Reduction** – In rapid network adjustments, Cisco Crosswork SON reduce time to market of new automation applications, launch of new services deployment, such as 5G or IoT. Cisco Crosswork SON allows operators to deliver the highest level of customer experiences.

- **E2E Automation** – As part of the Cisco Crosswork mobility, it is the most future-proof solution for 5G optimization, supporting integration into multiple network domains such as the transport/backhaul, the core, and the policy domains and enabling end-to-end service optimization solutions.

Cisco Crosswork User-Centric SON optimizes and automates RAN based on new data sources such as Customer Experience systems (KQI), geo-location, crowdsourcing and big data analytics. It enables operators to offer differentiated QoE to their subscribers. It also introduces machine learning capabilities, enabling Crosswork SON to become pro-active, predicting performance degradation and the most effective action to address it.

Cisco Crosswork SONFlex provides simple and open APIs toward Cisco Crosswork SON platform for 3rd party and operator-driven innovation. It enables hassle-free SON development, obscuring the complexities of the underlying vendor equipment and management systems. It also provides the conflict management services, essential to enable multiple SON applications optimizing the network simultaneously.

Cisco Crosswork SONFlex Studio introduces GUI-based SDK environment enabling non-programmers to create light-weight automation applications using Crosswork SONFlex APIs. It reduces OPEX by boosting the effectiveness of the radio engineers who can automate their routine tasks. The resulting apps can be shared across markets and customized to the market-specific challenges and practices.

Cisco Crosswork Multi-Domain Service Optimization (MDSO) allows end-to-end visibility and automation framework for network slices and service optimization. MDSO optimizes user-level service KQIs across all service domains - application servers, packet core, transport and RAN. MDSO assures that network limited capacity is utilized in the best possible manner according to operator’s policies, for optimal service assurance.
TERMS OF USE

LICENSE AGREEMENT
This report and the information therein are the property of or licensed to Heavy Reading, and permission to use the same is granted to purchasers under the terms of this License Agreement ("Agreement"), which may be amended from time to time without notice. The purchaser acknowledges that it is bound by the terms and conditions of this Agreement and any amendments thereto.

OWNERSHIP RIGHTS
All Reports are owned by Heavy Reading and protected by United States Copyright and international copyright/intellectual property laws under applicable treaties and/or conventions. The purchaser agrees not to export this report into a country that does not have copyright/intellectual property laws that will protect Heavy Reading's rights therein.

GRANT OF LICENSE RIGHTS
Heavy Reading hereby grants the purchaser a non-exclusive, non-refundable, non-transferable license to use the report for research purposes only pursuant to the terms and conditions of this Agreement. Heavy Reading retains exclusive and sole ownership of all reports disseminated under this Agreement. The purchaser agrees not to permit any unauthorized use, reproduction, distribution, publication or electronic transmission of this report or the information/forecasts therein without the express written permission of Heavy Reading.

DISCLAIMER OF WARRANTY AND LIABILITY
Heavy Reading has used its best efforts in collecting and preparing this report. Heavy Reading, its employees, affiliates, agents and licensors do not warrant the accuracy, completeness, currentness, noninfringement, merchantability or fitness for a particular purpose of any material covered by this Agreement. Heavy Reading, its employees, affiliates, agents or licensors shall not be liable to the purchaser or any third party for losses or injury caused in whole or part by Heavy Reading's negligence or by contingencies beyond Heavy Reading's control in compiling, preparing or disseminating this report, or for any decision made or action taken by the purchaser or any third party in reliance on such information, or for any consequential, special, indirect or similar damages (including lost profits), even if Heavy Reading was advised of the possibility of the same. The purchaser agrees that the liability of Heavy Reading, its employees, affiliates, agents and licensors, if any, arising out of any kind of legal claim (whether in contract, tort or otherwise) in connection with its goods/services under this Agreement shall not exceed the amount the purchaser paid to Heavy Reading for use of this report.

DISPUTE RESOLUTION
This License will be governed by the laws of the State of New York. In case of a dispute arising under or related to this License, the parties agree to binding arbitration before a single arbitrator in the New York City office of the American Arbitration Association. The prevailing party will be entitled to recover its reasonable attorney fees and costs.

Heavy Reading
P.O. Box 1953
New York, NY 10156
Phone: +1 212-600-3000
www.heavyreading.com