
Winning a public performance benchmark and cost efficiency can go together

Dublin, Ireland 2020

Author: Jose Rodriguez, Technical Solutions Consultant, Aspire Technology

Table of Contents

1. Abstract	p3
2. State of play	p5
3. Operator challenges	p6
3.1 End to end optimization	p6
3.2 The optimization process	p7
4. Solution	p8
4.1 Aspire vision	p8
4.2 The new end to end optimization process	p8
4.3 End to end troubleshooting example	p11
5. Success Stories	p13
5.1 Time savings	p13
5.2 Realizing improvements	p14
6. Conclusion	p15
7. References	p16

1. Abstract

Public performance benchmarks are a priority for many Communication Services Providers (CSPs) around the world, and an integral part of their business strategy every year. Preparing for them can be expensive, labor intensive and inefficient due to the need for extensive active testing (e.g. drive and walk tests) and the optimization process that will require improvements.

The metrics benchmarked reflect customer experience and as such are completely service centric – for example they include YouTube video start time, voice quality, call setup time and Webpage download time. In order to get improvements in such metrics, active testing information is not enough so very often engineers need to gather additional network information (cell trace, probes, PM) to get the complete picture to come to a solution. This process is slow, manual and reactive; and it is not scalable: engineers find it challenging to understand the prevalence of a given behavior and move to more pro-active holistic network wide analyses.

The above is made worse if we consider the very few failures in today's networks (e.g. unsuccessful voice calls for drive tests in the cities in the 2019 UK national public benchmark were 0.05% and 0.11% for EE and Vodafone) and how close some operators are with each other (e.g. YouTube video start times were 1.4 and 1.3 seconds respectively for EE and Vodafone). Every issue counts – as the slightest improvement can make an operator win the benchmark.

At Aspire we understand this process and have developed a solution to take it to another level. Our 'Drive Test & Probe Analytics' platform automates the collection, parsing, analysis and presentation of drive test and probe logs. We use Machine Learning assisted models to identify issues or abnormal behaviors and to fingerprint the entire network, allowing operators to scale and reduce the need for manual repetitive tasks and for continuous active testing.

The solution has proven to be very efficient. The time required to analyze a 6-hour drive test drive with 320 calls and 10 issues for deep-dive went from 3 working days to 1 for one of our European customers. In Telekom Srbija, our solution was part of the strategy for winning the national public performance benchmark for the first time in 2019.

At Aspire we understand this process and have developed a solution to take it to another level. Our 'Drive Test & Probe Analytics' platform automates the collection, parsing, analysis and presentation of drive test and probe logs.

We use Machine Learning assisted models to identify issues or abnormal behaviors and to fingerprint the entire network, allowing operators to scale and reduce the need for manual repetitive tasks and for continuous active testing.

2. State of Play

Network performance benchmarking campaigns have become a priority for many mobile Communications Services Providers (CSPs) around the world. Winning the national benchmark or simply improving the results from the previous year leads to marketing campaigns, positive press exposure and brand awareness [1] [2] and has therefore become an integral part of their company business strategy.

The growth of companies such as umlaut, formerly known as P3, is a good measure of how common these campaigns have become. Their performance scoring methodology is used in over 200 mobile networks in more than 120 countries [3]. Another good example is the public benchmark in Serbia organized by Ratel – Serbian Regulatory Agency for Electronic Communications and Postal Services. Its public portal [4], developed by Systemics PAB, is available online and provides a good measure on the methodology, and high level of details on performance drivers and scoring in these initiatives.

The way operators optimize their networks to prepare and maximize their benchmarking score is two-fold. They ensure access to a service and the ability to retain it, e.g. the percentage of unsuccessful voice calls for drive tests in the cities in the 2019 UK umlaut public campaign were 0.05% and 0.11% for EE and Vodafone respectively [5]. This is a very low number so each one of these failures matter. Operators make sure they resolve each one.

Once the service is guaranteed from an accessibility and retainability point of view, the focus is on quality. For voice it is call setup time and voice quality, for data it is time to content, video resolution and speed. Again, operators look at each one of these metrics individually as even the slightest improvement can make them gain a position in the final score. Keeping the same example as before, YouTube video start times for drive tests in the cities were 1.4 and 1.3 seconds respectively for EE and Vodafone – that's very close.

Certainly, network and spectrum assets together have a major influence on how high an operator can score. These benchmarks are usually done nationwide so if an operator doesn't have coverage in a certain area and the competitor does there is not much to be done. Or if the capacity is very imbalanced it will be impossible to match the delivered speeds.

Nevertheless, coverage is often very similar across competitors, especially in urban areas so the differentiation will come from how well the network behaves. Capacity may be more different depending on the case, but it barely affects important metrics such as call setup time, voice quality or time to content.

Load is another factor to consider that will affect performance. Its impact has a direct correlation not only with the operator's network and spectrum assets but also with how efficiently the traffic is managed and with the mechanisms the network has to react to load increases and potential capacity shortages.

3. Operator Challenges

This section presents some of the main technical and operational challenges for CSPs when preparing for a public performance benchmark.

3.1 End to end optimization

The first challenge to overcome is related to the ways of working. In order to troubleshoot and optimize technologies such as Voice Over LTE (VoLTE) and different data applications (Web Browsing, YouTube, Facebook) it is essential to have a clear view and understanding of the end to end service, covering from the device behavior to the application server. Traditionally the focus has been in the radio domain, but this is no longer enough. Two examples to illustrate this:

1. In most networks, services such as YouTube or WhatsApp voice calls are not differentiated in any way in the radio access network, they go through the same 'pipe', so to speak.
2. For many services speed is not a determining factor in driving performance, e.g. voice quality, small file transfers and Web Browsing will be impacted much more by end to end latency and packet loss.

This may sound straight forward but has proven a big challenge. As many other big businesses, mobile operators are not used to work end to end, i.e. across multiple organization silos. The teams in the performance departments are domain-centric plus it is extremely difficult to find 'end to end experts', given the number of nodes and the increasing complexity of networks and services. Something very similar happens with available tools: they were designed to serve a set of ways of working that are not valid anymore – very few tools in the market today provide an end to end view, they focus on single data sources from each specific domain and do not tackle the required correlation and data analytics needed to understand and improve nowadays' services.

Service providers and vendors do realize the need for such an end to end approach to troubleshooting and optimization but many of them are failing to take it to practice, at least in an efficient way.

In order to dig into how they tackle this challenge, it shall be mentioned that, to prepare for the public benchmark, operators do extensive internal testing beforehand, replicating the same methodology, services and routes than the ones to be used in the public campaign. To analyze and drive improvements based on these private benchmarks, there is usually a 'Central Performance Team' that acts as first line of analysis, performs issue triage and involves other groups or experts as needed. The Central Performance Team would investigate as many as possible individual failures and voice/data calls with poor quality of service, as per one of the measured metrics. This is a tedious and very manual process; certainly, the number of calls analyzed and the depth of the analysis will depend on the size and efficiency of the team.

Here is where the difficulty comes: to perform an adequate analysis engineers need to look at the end to end picture, as explained above, and to do so the data collected from the device in the drive test and the data available in the network from the radio and core domains need to be parsed, correlated, presented and analyzed. This correlation and end to end analysis is completely manual in most cases leading to long times to action and huge inefficiencies. Great (and expensive) experts spend hours collecting and parsing such information. Additionally, given the increasing size and complexity of the data and services, natural human limitations come into play which limit the undesired behaviors that can be identified – some things will undoubtedly be missed.

3.2 The optimization process

Apart from the very technical aspect of performing efficient end to end troubleshooting and optimization, there is an even bigger problem: the process from performing active testing to realizing improvements is long, manual and inefficient. Presented below such process, highlighting where the bottlenecks and challenges are.



Figure 1: Optimization process today

1. Drive test data collection

- Usually the raw files are uploaded to a server, manual intervention is required for collection and parsing.

2. Manual analysis & issue identification

- Nowadays there is usually a level of automation here in the form of a summary of the test results. Typical failures or undesired behaviors are identified automatically.
- However, this is done using the testing device or User Equipment (UE) as unique data source and there is little to no flexibility as to which scenarios, patterns or behaviors are marked or highlighted as an issue. As services and network complexity expand there are more patterns to be highlighted and new issues to be found.

3. Troubleshooting and data sources correlation

- Once the issues have been identified, engineers will perform a root cause analysis on as many as possible to decide on any required actions.
- At this point, correlation with other data sources is triggered. Very often the point of view of the UE is not enough so it will be needed to fetch further radio or core network data.
- Only fetching this extra information is already time consuming and manual, which multiplies when the correlation and analysis that comes after is added.

4. Categorization & prevalence assessment

- Once the root cause of the issue is identified, there may already be a resolution action, or a domain expert may need to be involved.
- At the same time, engineers would try to assess how common the identified issue is to understand if needs to be prioritized or if the corrective action needs to be applied somewhere else in the network.
- However, this is very challenging due to the very nature of this process: being a manual analysis and data correlation, most of the time it is simply not possible to move to a holistic network view where all tests or network nodes are checked for a particular set of symptoms.

5. Action & improvement verification

- Finally, certain actions are taken to prevent the issue from occurring again.
- Here there is a similar challenge to the one described in the previous step: very often it is not possible to 'query' the entire network or set of tests to see if the issue is still happening so engineers cannot be certain if the implemented solution has worked, unless new drive tests are performed, which takes engineers back to point 1.

4. Solution

4.1 Aspire vision

At Aspire we have first-hand experience on the above challenges and sub-optimal processes. We have suffered the long processing times in drive test tools and the nightmare of correlating a UE log with core network probe traces to troubleshoot a VoLTE drop call, for example.

We were driven to enable our engineers to work smarter by removing all manual and repetitive tasks from their day to day job. But it is more than that – we are convinced that there is much more to do with the available data than simply reacting to a set of common issues implementing localized solutions. We had a vision:

- 1** Foundation for **advanced analytics & ML models** for gap analysis and closed loop issues resolution
- 2** Network wide services & experience **troubleshooting & RCA**, moving away from call by call analysis
- 3** Actionable **intelligence & engineering logic**, focused on solutions rather than highlighting problems
- 4** Continuous support of **new technologies & services**, e.g. VoLTE, IoT, 5G, OpenRAN
- 5** Flexible addition of new **data sources & correlation logic**, e.g. PM, FM, complaints, crowdsourcing, BSS

Figure 2: Aspire performance optimization automation vision

4.2 The new end to end optimization process

Aspire Drive Test & Probe Analytics solution allows operators and service providers to radically improve how they tackle end to end optimization processes, and particularly how they prepare for public performance benchmarking campaigns.

The new process is presented below:



Figure 3: The optimization process with Aspire Drive Test & Probe Analytics solution

Winning a public performance benchmark and cost efficiency can go together

1. Drive test and probe data collection

- Collection and parsing are done automatically, no manual intervention is required.
- Other data sources such as probes are collected at this point, automatically, so they are available for analysis from the beginning.
- Correlation between data sources is performed using Aspire proprietary models.

2. Auto correlation, analysis & fingerprinting

- Our logic would automatically look for any predefined patterns, behaviors or issues using the collected data.
- The amount and type of patterns detected evolve and improve over time as more data goes into the system and through user inputs, they are in continuous enrichment.
- Each detected issue gets categorized with a domain and reason, and recommendations are delivered where possible.
- Advanced analytics and ML models come into play here: as the collected data grows, the system will get trained and start identifying new abnormal patterns on its own, learning from corrective actions and suggesting recommendations.

3. AI assisted end-to-end optimization and analysis

- It's at this point where expert engineers come into play. They will define the required actions to fix the highlighted issues or deep dive into the specific patterns or behaviors identified by the solution.
- Thanks to the automatic collection and correlation of different data sources, there will be no need for the experts to go to other systems to collect extra information to complement their investigations.
- They will be able to check the end to end signaling flow in a single view and visualize the propagation of messages from the air interface towards the network and applications. This facilitates at the same time the correlation of radio metrics with network and application performance, one of the main challenges during these investigations.
- Our solution allows as well to visualize both sides of voice calls at the same time, allowing for faster troubleshooting. This is a very common scenario where for example two cars call each other during drive tests.
- As mentioned in the step 2, the correlation engine, logic and pattern recognition are in continuous improvements - as more data goes into the system and together with punctual user feedback the models keep learning and more and better AI assisted recommendations are possible.

4. Continuous analysis of probe data & fingerprinting

- An important limitation with the traditional process had to do with issue prevalence assessment and improvement verification.
- Thanks to the continuous analysis of network wide probe logs, signatures or patterns can now be marked and verified across the entire network.
- This will heavily reduce the need for additional drive tests to verify a solution, reproduce issues or check other regions
- Additionally, thanks to having everything in a single system, all drive test data can be aggregated and analyzed as a whole: trending, benchmarking, GEO-based optimization are now possible.

Winning a public performance benchmark and cost efficiency can go together

Log ID	City	Cluster	Logs	Result	Domain	Reason	Events	Service	Service Type	Call Type	Matching Log ID	Star
350-1		POC	TEMS IRIS	BLOCKED	Network	SIP Cancel Request (DL)	⊗ ⚠	Voice	VoLTE	MT	377-1	2020
409-1		POC	TEMS IRIS	BLOCKED	Network	SIP not established at Release	⊗	Voice	VoLTE	MO	-1-1	2020
377-1		POC	TEMS IRIS	BLOCKED	Radio	SIP not established at Release	⊗ ⚠	Voice	VoLTE	MO	350-1	2020
402-1		POC	TEMS IRIS	DROPPED	Radio	Failed LTE Reestablishment	⊗ ⚠	Voice	VoLTE	MT	403-1	2020
398-1		POC	TEMS IRIS	DROPPED	Network	Early Call End	⊗	Voice	VoLTE	MO	393-1	2020
393-1		POC	TEMS IRIS	DROPPED	Radio	Reestablishment Reject	⊗ ⚠ ⚠	Voice	VoLTE	MT	398-1	2020

Call result (drop, block or success) and domain with issue

Set of events identified per call

Figure 4: DT & Probe Analytics solution: Call summary view

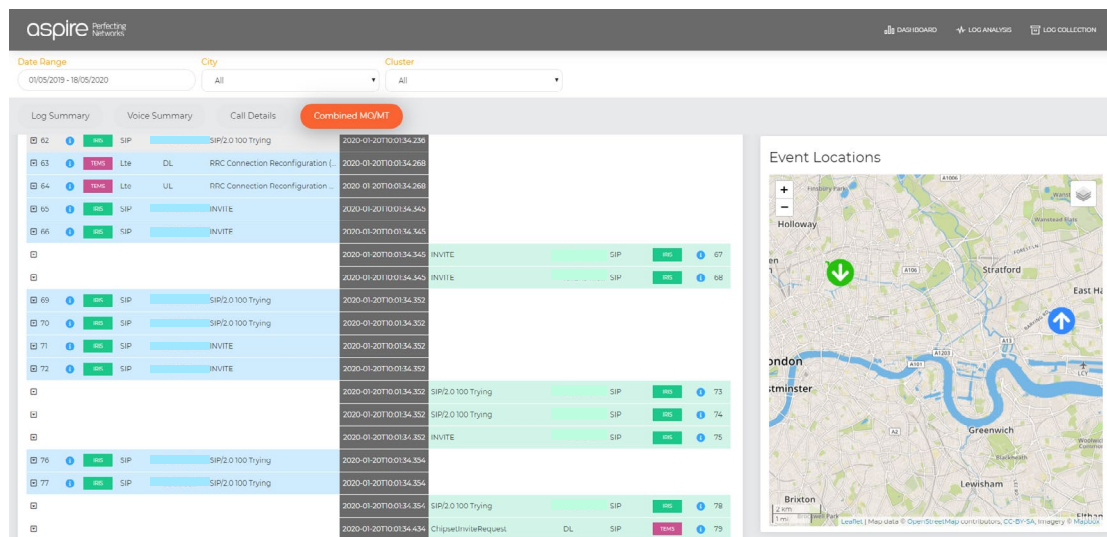


Figure 5: DT & Probe Analytics solution: MO & MT signalling flow view

Winning a public performance benchmark and cost efficiency can go together

4.3 End to end troubleshooting example

In order to illustrate the benefits to be achieved with Aspire’s methodology, presented below an example during which, during the preparation of a customer’s network for the benchmark, the automated analysis of drive test logfile pointed to three segments of calls in terms of call setup time:

- Well performing calls (72%), with call setup time between 1s and 2s.
- Moderately affected calls (20%), with call setup time between 4s and 5s.
- Heavily affected calls (8%), with call setup time above 7s.

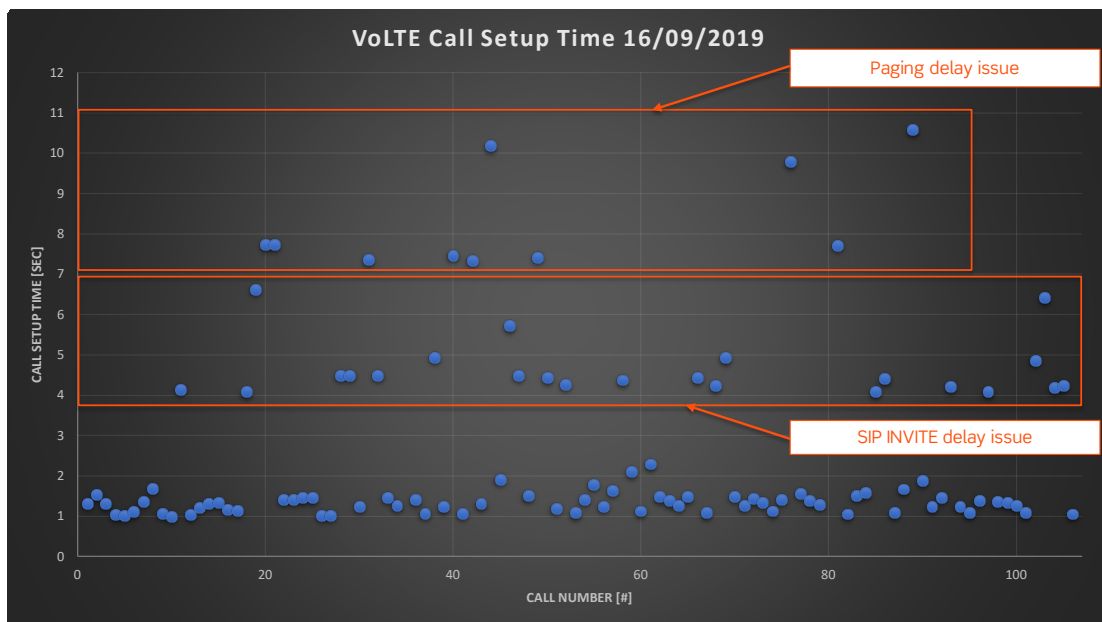


Figure 6: VoLTE calls segmentation – call setup time criteria

The deep dive to moderately affected calls, pointed to issues with SIP Invite procedure, where calls with a SIP INVITE interchange between IMS nodes with a constant propagation delay of around 3 seconds:

10	2019-09-12 16:18:55.129000	ANRITSU	SIP		SIP/2.0 100 Trying	confidential
11	2019-09-12 16:18:55.132000	ANRITSU	SIP		INVITE	confidential
12	2019-09-12 16:18:55.138000	ANRITSU	SIP		SIP/2.0 100 Trying	confidential
13	2019-09-12 16:18:55.220000	TEMIS	SIP	DL	ChipsetInviteTrying	
14	2019-09-12 16:18:55.263000	TEMIS	Lte	UL	Measurement Report (UL-DCCH)	No measurement configured for measid
15	2019-09-12 16:18:56.102000	TEMIS	Lte	UL	Measurement Report (UL-DCCH)	No measurement configured for measid
16	2019-09-12 16:18:58.236000	ANRITSU	SIP		INVITE	confidential

Propagation delay ~3.1 sec

Figure 7: SIP INVITE delay issue, solution snapshot

Winning a public performance benchmark and cost efficiency can go together

Analyses further indicated to a Diameter issue which was confirmed and fixed afterwards as a problem in a charging system node.

The deep dive to heavily affected calls on the other hand, pointed to paging delay issues, where calls where the first paging was unsuccessful, and the second paging was coming very late, adding between 5 and 8 seconds extra:

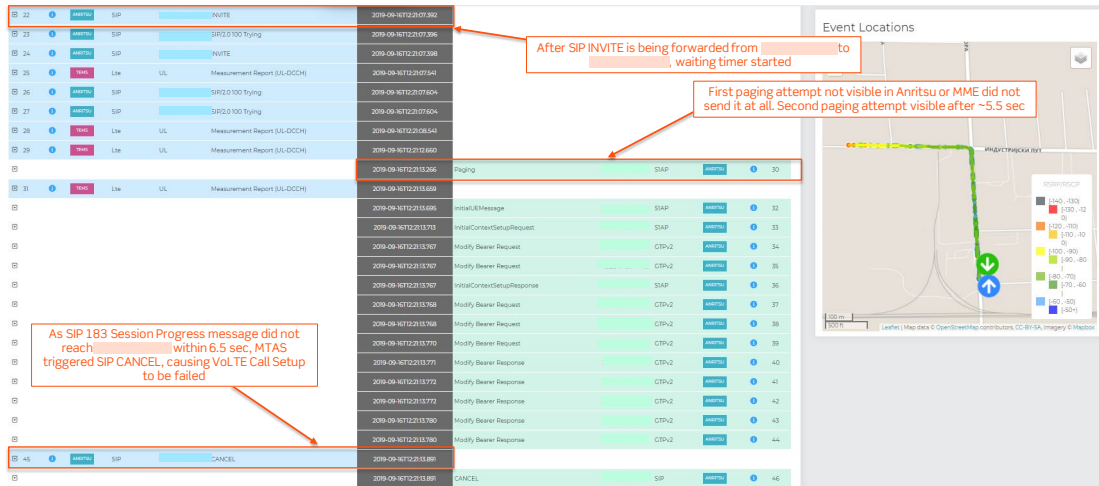


Figure 8: Paging delay issue, solution snapshot

On the paging delay problem, a temporary solution was put in place during the investigation, as the extra delay was creating call setup failures when the MTAS waiting timer to receive the next SIP message was expiring - the timer was extended from 6.5 to 8 seconds to eliminate the failures. Afterwards, after a thorough investigation with the domain experts, a software bug in one of the PCRF nodes was found and corrected.

5. Success Stories

5.1 Time saving

Using a typical scenario is one of our major customers, we were able to measure the achievable gains from a time perspective. For a 6-hour drive test with approximately 360 voice calls and a total of 10 issues that require end to end deep-dive investigation:

Traditional process without Aspire:

- 3 working days, purely manual process
- Missed issues and limited abnormal behaviors
- Same slow repetitive and process applies for the new drive-tests needed for improvement verification and issue reproduction

With Aspire Drive Test & Probe Analytics solution:

- 1 working day, automatic process
- All issues and additional patterns identified, including recommendations
- Network wide analysis for issue prevalence and verification, thanks to having all drive tests accessible in one platform plus continuous analysis of network wide probe logs
- ML/AI model continuous improvement to reduce the need for active testing and to keep improving issue identification and recommended/performed actions

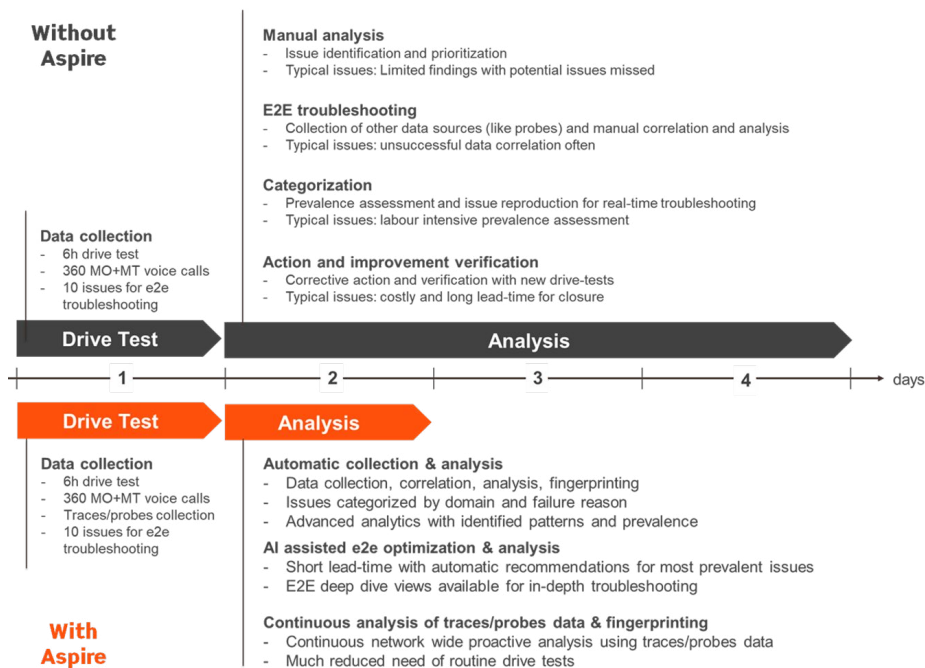


Figure 9: Typical drive test and analysis scenario (with and without Aspire)

Winning a public performance benchmark and cost efficiency can go together

5.2 Realizing Improvements

Another example where Aspire's solution proved successful took place during the 2019 Public Performance Benchmark in Serbia. Our customer, Telekom Srbija (MTS), introduced the Drive Test & Probe Analytics with the objective to serve as the differential factor for MTS to win the benchmark for the first time.

The initiative was a success as it not only enabled engineers to work more efficiently by automating labor intensive and repetitive tasks, but it uncovered issues affecting important metrics for winning the benchmark [6].

Apart from benefits in terms of time savings and an improved day to day optimization process, the advanced analytics and issue identification models proved useful very fast as the solution was able to detect abnormal behaviors affecting service performance. Furthermore, outcome of analysis were recommendations for signaling procedures optimization, where implementation of changes contributed to winning the first place in category of call setup time, with significant 480ms margin over closest competitor.

Conclusion

Service centric optimization requires data analytics and automated ways of working, it's unfeasible from a cost and effort perspective to do in other way. In nowadays ultra-competitive environment, it is essential for operators to make this transition successfully.

Talk to our experts to understand how you can prepare for the next public benchmark efficiently, and how to benefit from our proven solutions and ways of working.

Aspire has been supporting CSPs achieving outstanding results based on the expertise gained through hundreds of multi-technology projects with different vendors. This has given us a solid background of use-cases and lessons learned that can be applied to new scenarios in a tailored way depending on business objectives and challenges.

Engage with Aspire Technology now

The next step is to call Aspire Technology on +353 (1) 9022376, email soltuions@aspiretechnology.com or visit us at aspiretechnology.com

References

- [1] EE and Vodafone Come Top of New 2020 UK Mobile Benchmark, <https://www.ispreview.co.uk/index.php/2020/01/ee-and-vodafone-come-top-of-new-2020-uk-mobile-benchmark.html>
- [2] EE retain crown in mobile benchmarking survey for 2019, <https://www.choose.co.uk/news/2020/ee-win-mobile-performance-survey/>
- [3] Umlaut Benchmarking, <https://www.umlaut.com/en/benchmarking>
- [4] Ratel | Mobile networks in Serbia <http://benchmark.ratel.rs/en>
- [5] Umlaut Benchmarking, United Kingdom, <https://www.umlaut.com/en/benchmarking/united-kingdom>
- [6] MTS – the best network in Serbia, <https://mts.rs/About-Telekom/Media-center/a88689-mts-the-best-network-in-Serbia.html>

Winning a public performance benchmark and cost efficiency can go together

Contact Us

111 Q-House,
76 Furze Road,
Sandyford Industrial Estate,
Dublin,
Ireland
Tel: +353 (1) 9022376
info@aspiretechnology.com
www.aspiretechnology.com