Capgemini @engineering

intel.

PROJECT BOSE

A smart way to enable sustainable 5G networks



Tribute to Dr. Jagadish Chandra Bose



The project is named after Dr. Jagadish Chandra Bose as a tribute to the 'Father of Radio Science', a great scientist who made several significant discoveries in his life.

Born in 1858, in the Mymensingh district of the Bengal Presidency (present day Bangladesh), Bose was known most significantly for his research on radio development. Bose was among the pioneers of research in radio technology and demonstrated, for the first time ever, wireless communication using radio waves. Bose was against patenting because he believed knowledge should be available to everyone. He always believed in the principle that "it is not the inventor but the invention that matters."

The Institute of Electrical and Electronics Engineers, a New York-based international body, called him the 'Father of Radio Science', since the science behind radio technology was first explained by Bose. His work in radio science was instrumental to another significant discovery he made. Bose was one of the first to employ an interdisciplinary approach — combining botany with physics — to prove that plants had life.

J. C. Bose was an inspiring teacher, a skillful experimenter who designed his own research equipment, a committed institution builder, and a great humanist who believed that the benefits of science should reach society at large. Thus, even today, J. C. Bose stands out as a role model for the younger generation intending to take up careers in academia and research work.

The objective of Project Bose is to realize sustainable 5G networks – and beyond - in a smarter way, by bringing together the best of multiple disciplines like telecommunications, data science, and environmental science to follow in the footsteps of this great scientist.

Table of contents

- **01...** Executive summary
- 02... Introduction
- **03...** Project Bose for a sustainable network
- 04... Capgemini's NWDAF
- **05...** Intel's observability framework
- 06... Intel AI/ML technologies
- **07...** Test results
- **08...** Energy saving results
- **09...** AI/ML performance results
- **10...** Conclusion
- **11...** References
- **12...** Glossary

Executive summary

Capgemini, along with its 5G ecosystem partner Intel, has developed an innovative energy optimization solution for a sustainable 5G network. This solution is the first part of Project Bose, Capgemini's initiative for a green 5G network. This technical project focuses on developing a data-driven intelligent network solution for a 5G RAN and Core network to build a sustainable 5G ecosystem. During current times, where the telco industry is struggling to meet stringent net-zero energy targets, Project Bose, with its holistic approach, will provide a difference on the path to sustainability.

One of the key challenges for telcos today is to tackle the rising carbon footprint in the network. Though there are traditional methodologies available, there is a need for a holistic and innovative solution. Project Bose, as the tagline says, uses a data-driven approach for 'enabling sustainable 5G networks in a smarter way'. The project is built on Capgemini's industry-leading network AI framework NetAnticipate [1] with a newly evolving 5G Network Data Analytics Function (NWDAF) framework [2]. This framework provides various analytic insights for 5G network infrastructure to build a variety of use cases in the 5G ecosystem like service assurance, load balancing, data monetization, and more.

The sustainability solution presented in Project Bose is a result of the collaboration between Capgemini and Intel, combining network and infrastructure expertise from both sides to develop an innovative solution. It focuses on deriving valuable analytic insights from UE (User Equipment) mobility management and network function load conditions by consuming 3rd Generation Partnership Project (3GPP) network data complemented with non-3GPP infrastructure insights. Non 3GPP metrics are exposed through Intel's observability framework.

The five key levers – UE Paging, Smart UPF (User Plane Function) Selection, MICO (Mobile Initiated Connection Only) Mode. NF (Network Function) Selection, and CPU Frequency Tuning cover the entire network deployment in a holistic way and greatly assist in performing closed loop automation for reducing energy consumption, mainly on the 5G core network side. Overall, energy has been reduced by 18% using these levers, and savings validated across various deployment, traffic scenarios and time intervals. **These** energy savings directly translate to around 14% savings in CO2 emissions, also cutting **OPEX by around 12%.** Even while optimizing energy, Project Bose ensured that QoE isn't compromised. There was marginal OoE improvement of 0-5%, observed during our lab experiment, but observations confirm there is scope to improve it further in the future. The results are described in detail in the section "Energy-Saving Results".

Introduction

For many years, the telecom industry's impact on the environment went relatively unnoticed. But its effect is consequential and gaining attention. The industry accounts for 3 to 4% of global CO2 emissions, about twice that of civil aviation. With global data traffic expected to grow by around 60% per year, the industry's share will grow further unless investments in energy efficiency can offset the effect, according to a BCG report [4].

The International Telecommunication Union (ITU) has published guidelines for ICT companies to reduce greenhouse gas emissions at the rate necessary to meet the Paris Agreement's goal of limiting global warming to 1.5°C above pre-industrial levels. [3]

All major telcos across the globe have supported the ITU's vision and have committed to be net zero by 2060 at the latest, as shown in figure 1 below. According to the GSMA [5], about 31% of operators by connections worldwide and 36% of operators by revenue worldwide have set carbon reduction targets to be net zero by 2050. To achieve this timeline, telcos have to halve the emissions by 2030. It is a stringent target and telcos have to act now if they are serious about this net-zero commitment.

According to Fierce Wireless [6], in the U.S., Verizon and AT&T have said they will be carbon neutral across their entire operations by 2035. In 2021, Verizon launched a \$1 billion green bond that will be used for its sustainability efforts. T-Mobile also has a target of reducing greenhouse gas emissions by 95% by 2025 from a 2016 base-year.

Though there are many energy optimization solutions available, operators put most trust in advanced AI-based network optimization for improving their network energy efficiency, compared to other solutions, as per GSMA Intelligence [7]. Unlike other hardware and CAPEX-intensive solutions, such as batteries, diesel generators, and cooling systems, AI-driven solutions offer a scalable option with a favorable return on investment, as shown in figure 2 below.



Figure 1: Telco net-zero target



Source: GSMA Intelligence

Figure 2: Telcos trust advanced AI-driven approach for energy-saving

What can't be measured can't be improved. So to identify and understand the energy usage pattern of the network and to optimize and manage it in an energy efficient way operators need to deploy metering in networks. NGMN has come up with requirement, architecture and protocols for creating a standardized metering solution for sustainable networks [8].

In the last few years, AI and ML driven applications have mostly focused on shutdown solutions in the RAN. RAN optimization is just the first step. In order to maximize the benefits, energy management should be addressed end to end.

Capgemini's Project Bose addresses this gap by providing holistic energy-saving mechanisms for telco operators to achieve their net-zero target. Project Bose provides network AI-enabled energy-saving levers to significantly reduce energy consumption in the 5G network and associated IoT devices, resulting in considerable CO2 emission reductions and cost savings, with no negative impact on end user QoE. It enables operators to have an end-to-end view of the network data and leverage it to provide a holistic energy optimization solution.

This paper presents Project Bose, the sustainable 5G network solution and the underlying NWDAF platform in detail. Capgemini's NWDAF framework [2] is developed in collaboration with our partner Intel who brought in an observability framework for complementary metrics at infrastructure and network level. The test results achieved using various test cases are presented to demonstrate the solution's impact. The paper concludes with future directions which Project Bose can be pursued further.

Project Bose for a sustainable network

The objective of Project Bose is to create a sustainable 5G network - and beyond - using a data-driven approach. In order to achieve this, it has introduced five energy-saving levers that work in tandem to optimize energy consumption in the network and associated IoT devices, reduce CO2 emissions, and deliver cost savings, with no negative impact on end user QoE. It also provides a platform to build new energy-saving levers in the future. Figure 3 below provides an overview of Project Bose.



Figure 3: Project Bose overview

The five energy-saving levers of Project Bose are –

- Directional UE Paging Based on UE mobility analytics information from NWDAF, AMF (Access and Mobility Management Function) optimizes paging decisions and reduces the paging request load through directional UE paging. In directional paging, the paging request is sent to a user or group of users in the direction of their predicted location, rather than broadcasting it in all directions. This leads to power savings in the RAN and connected devices.
- 2. MICO Mode Mobile Initiated Connection Only (MICO) mode is designed for IoT devices that send a small amount of data and do not need to be paged. An example of this could be a smart bin that sends a message to the waste collection company saying it is almost full, so that the bin emptying truck can plan to empty it in the next collection round. Here there is no reason to page the bin as there is no mobile-terminated data that would be required. Based on UE mobility analytics information from NWDAF, AMF decides MICO mode parameters for these types of IoT devices.

- 3. Energy Aware NF Placement Based on NF load prediction information from NWDAF, the orchestrator decides the placement of NFs and other related workloads, so that energy consumption can be optimized in the servers hosting the workloads.
- 4. Smart UPF Selection SMF (Session Management Function), the anchor for selecting appropriate UPFs, performs intelligent load balancing of data traffic in UPFs based on NF load prediction information from NWDAF, thus optimizing energy consumption and improving the user's quality of experience.
- 5. CPU Tuning This feature is related to hardware c-states/p-states management, where frequency tuning saves power by adapting core frequencies to the given load, while c-states allow unused cores to be placed in sleep state when idle.

These five energy levers are implemented on top of Capgemini's NWDAF framework, using insights derived from analytics information exposed by NWDAF.

Capgemini's NWDAF

Capgemini's NWDAF is a 3GPP Release 17 compliant, cloud native 5G network AI framework that uses advanced machine learning (ML) techniques to provide real-time and predictive operational intelligence in the 5G core, enabling analytic-driven service orchestration and closed-loop network automation. The figure illustrates Capgemini's NWDAF high-level architecture:



Figure 4: NWDAF high-level architecture



There are five key components of Capgemini's NWDAF architecture:

- The Data Collection and Coordination Function (DCCF) coordinates the collection and distribution of data requested by NF consumers. It prevents data sources from handling multiple subscriptions for the same data and sending multiple notifications containing the same information due to uncoordinated requests from data consumers.
- The Analytics Logical Function (AnLF) performs inference, derives analytics information – specifically, statistics and predictions based on analytics consumer request – and exposes analytics services to consumers.
- The Model Training Logical Function (MTLF) trains ML models, exposes training services, such as providing trained ML models, and handles lifecycle management of trained ML models. It is based on Capgemini's state of the art ML framework, NetAnticipate [1].
- 4. The Analytics Data Repository Function (ADRF) offers services that enable consumers to store and retrieve data and analytics.
- The Messaging Framework Adaptor Function (MFAF) provides a messaging framework to receive data from the DCCF, and process, format, and send data to consumers or notification endpoints.

The key component that drives NWDAF is the industry-leading AI framework NetAnticipate [1]. It is a highly scalable, cloud-native, selflearning platform for realizing autonomous network operation. NetAnticipate analyzes a substantial number of hidden and hierarchical influencers to predict potential network anomalies, build autonomous decisions, and take preventive actions. An autonomous feedback loop ensures the network self-learns over time to improve the actions it takes. It uses advanced deep-learning algorithms for creating self-learning closed loop automation. NetAnticipate orchestrates various deeplearning algorithms for identifying network anomalies in real-time, predicting anomalies using multi-variate timeseries analytics, and prescribing preventive action to fix an anomaly before it can start affecting the network. Actions taken are improved over time through deep reinforcement learning techniques. Reinforcement learning uses paradigms for sequential decision-making under uncertainty, solely by rewards and penalties, from previous actions it had performed in the network.

Using NetAnticipate capabilities, NWDAF provides analytics information related to network performance, service experience, and UE behavior analytics to the analytics' consumers. For more information about Capgemini's NWDAF framework please refer to [2].

Project Bose's energy-saving levers leverage two types of analytic information from NWDAF – UE mobility prediction and NF load prediction.

- UE Mobility Prediction:
- This analytic information enables NF consumers to predict the UE location in the future and understand the trends in various areas over defined time intervals.
- The UE location information from AMF and OAM is collected with respect to cell ID and geographical location using latitude and longitude.
- This location information is used by AMF to help RAN to assist in directional paging, which leads to energy savings on the 5G RAN side.

- NF Load Prediction:
 - This analytic information enables NF consumers to predict the network function load in the 5G network and understand the trends in various location areas over defined time intervals.
 - The NF network/session load and the resource usage at disk, CPU, and memory level is collected from NRF and OAM.
 - Additional telemetry data is collected at infrastructure level using Intel® Platform Telemetry Insights, explained in detail in the 'Intel's Observability Framework' section.
 - Combining various load metrics from 3GPP and non-3GPP data, NWDAF helps to predict situations for overload, network performance degrade and more.

This analytics information is derived by NWDAF by correlating 3GPP-specified metrics collected from 5G core NFs, AFs, and OAM. In addition, non-3GPP metrics related to power consumption in infrastructure (e.g., servers) are collected using Intel's observability framework [9] and correlated with the 3GPP metrics. This provides deeper insight about energy consumption in the network, infrastructure, and connected devices.



Intel's observability framework

3GPP interfaces such as OAM provide measurements for CPU load, and thermal and cooling information, but do not provide the full set of detailed information available from the infrastructure. 3GPP-defined NF metrics focus on the KPIs for the service provided. such as sessions established per second, with no view from the underlying infrastructure on how factors on the infrastructure may be influencing the behavior of the hosted network functions. To address such gaps in the standards defined. Capgemini's NWDAF has been equipped to provision an alternative creative way to access additional non-3GPP data sources, in particular from the underpinning infrastructure, to enhance the data points suitable for model training. Intel's observability framework has been one such example framework used in Project Bose.

This framework leverages open industry standard interfaces to provide both infrastructure telemetry and cloud native telemetry. It provides additional information to the analytics system, adding infrastructure telemetry and cloud application metrics, logs, and traces to be contextualized with the 3GPP-defined metrics.

- 1. Telegraf For infrastructure telemetry collection [13]
- OpenTelemetry For application metrics, logs, and traces [12]

The logical architecture is shown below for an example application, in this case a 5G UPF.

In the example deployment shown in figure 5, demonstrated at Intel's Kubecon [11], Telegraf provides infrastructure telemetry to the OpenTelemetry agent. This agent associates the infrastructure metrics with the container name spaces and other identifying information. The OpenTelemetry Sidecar agent, running in an unprivileged container, provides UPF application specific metrics, logs, and traces. This framework has the capability of building insight reports which up-level the raw telemetry to reduce monitoring noise.

Intel's Observability Framework Integration with Capgemini's NWDAF

This framework has been integrated with Capgemini's NWDAF with OpenTelemetry Agents running on servers having 5GC-UPF deployments, to collect container level metrics using CAdvisor and Host Metrics, as shown in figure 6 below. These metrics are pushed to Kafka topics and then stored in the database store, which the NWDAF analytic engine can consume.

Intel's observability framework has enriched Capgemini's NWDAF data collection to extend beyond the 3GPP standard. The figure below outlines various energy-centric data collection interfaces from NWDAF: blue lines are the 3GPP-defined interfaces, and dotted red lines are the cloud native observability interfaces, and the infrastructure telemetry interfaces. Additional infrastructure metrics help to provide deeper insights of a network function's load changes and assist in decision making for better energy optimization.

As an example, the enablement of UPF inframetrics collection using an observability framework assists in the smart selection of a UPF use case. Intel's observability framework provides advanced infrastructure level metrics of the UPF, either in a virtual machine or container, to the NWDAF.

The load analytics information of UPF is used by SMF for intelligent load balancing of data traffic in UPF, with UPF being the critical network function that handles the data traffic in a 5G core network. So, the selection of appropriate UPF by SMF for providing the required data throughput, and desired quality of service, plays a key role in the 5G ecosystem. Intelligent load balancing also saves a substantial amount of energy.



Figure 6: Capgemini NWDAF energy-centric data collection



Figure 5: Intel's observability framework integration with 5G UPF

AI/ML technologies

While data collection metrics are enhanced by using Intel's observability framework for deep telemetry information, an equally important part of this solution lies in the predictive analytics using AI/ML models developed in MTLF. The accuracy and latency of the predictions are very important for timely inputs to the 5G core NFs so that decisions can be taken in a near-real time scenario. The models providing the predictions need to be scalable and auto-tunable to adjust to the ever-changing conditions of the network.

Intel's AI/ML technologies help in enhancing these features in the NWDAF solution. Capgemini's NWDAF's MTLF engine, NetAnticipate [1], has integrated some of the following AI-related framework/tools from Intel:

1. Intel Xeon Scalable Processor – built-in AI acceleration

NetAnticipate leverages AI training and inference on third generation Intel® Xeon® Scalable Processors. Intel Deep learning Boost technologies, and Vector Neural Network Instructions (VNNI) with INT8. improve AI performance by maximizing the use of computing resources, improving cache utilization, and reducing potential bandwidth bottlenecks. The industry's first x86 support of Brain Floating Point 16-bit (blfoat16) brings enhanced AI performance. Xeon's built-in AI acceleration provides NetAnticipate with the flexibility to run complex AI workloads on the same hardware as existing workloads such as analytics, network functions, and Edge applications. 2. Intel[®] one API AI Analytics Toolkit (AI Kit): Accelerate E2E AI pipeline

Al Kit provides tools and frameworks to accelerate the end-to-end analytics and Al pipeline on Intel Architecture. Capgemini's deep learning models used in NWDAF load analytics and UE mobility have used Intel® Optimization for PyTorch to improve Al training and inference performance required in the solution. Intel® Distribution of Modin provides performant, parallel, and distributed pandas workflows for Capgemini data pre-processing.

3. BigDL/Chronos: Build scalable time-series AI solution

Chronos provides solutions to build time-series forecast AI models including data-processing and feature engineering, with the support of more than 10 built-in models (TCN, LSTM, Seq2Seq, NBeats, TCMF, MTNet, ARIMA, Prophet) with friendly API, and AutoML for automatic HPO (Hyperparameter Optimization) in a distributed architecture. NetAnticipate leverages Chronos AutoML to train and tune AI models to develop highly accurate load analytics and UE mobility predictions that enable improved energy savings.

Test results

Capgemini and Intel teams have performed various tests to measure the impact of Project Bose. The test results focus mainly on two areas – the comparison of energy consumption measured across the identified levers, and the performance of predictive analytics used to achieve this.

Energy-Saving Results

The integrated test setup of Capgemini's NWDAF and Intel's observability framework was deployed in multiple servers in Capgemini's lab. The five key levers identified were measured across this distributed test setup, both with and without NWDAF at various time intervals and different traffic scenarios over a one month time-period.

Figure 7 below illustrates Capgemini's NWDAF and 5G Next-Generation Core (NGC) deployed on five bare metal servers. There is one central NWDAF, with two NWDAFs deployed at the edge closer to UPF. Capgemini's NGC cluster (with 10NFs) is deployed on three bare metal servers. UE and gNB simulators are used to simulate various test conditions for UE mobility and NF load scenarios. There are 100 UEs simulated with around 4,000 calls per hour.

The observability framework was used to measure infrastructure metrics like per core utilization, temperature, and power consumption to compute overall energy savings at infrastructure level by each of these levers.

The graph in figure 8 below shows a summarized view of the results obtained from the tests conducted in the lab. An average energy saving of 18% has been observed, resulting in around a 14% reduction in CO2 emissions and 12% cost savings.



Figure 7: Project Bose test setup



Figure 8: Energy optimization results

The energy dashboard in figure 9 below shows the detailed result. This dashboard indicates the energy consumption results in two colors - blue indicates without NWDAF, while green indicates with NWDAF. The radar graph on the left side shows a comparison of the energy consumption between the two, due to each of these key levers. The four meters below the radar indicate four critical parameters related to operational efficiency. These include percentage changes in energy consumption, CO2 emission, OPEX, and QoE. While we observed a reduction in energy, CO2, and cost, there has been no negative impact on user QoE - rather, there is a marginal (0-5%) improvement. The charts on the right-hand side show the comparison for the last one hour's data.



Figure 9: Network sustainability dashboard

AI/ML Performance Results

Capgemini, along with Intel, carried out tests on the accuracy of ML predictions and AI performance on various use-cases including NF load and UE mobility. The ML prediction accuracy tests were carried out with and without Chronos AutoTS. AI performance tests were carried out on Chronos versus stock PyTorch. The results show significant improvement on prediction accuracy and training/inferencing performance using Chronos AutoTS.

NF Load Dataset	nf_load	cpu_ usage	memory_ usage	disk_ usage
TCNForecaster	16.72	17.50	17.30	16.43
AutoTS	13.34	14.61	14.76	12.98
Improved Rate	20.22%	16.51%	14.68%	21.00%

Table 1: Comparison of NF load model prediction accuracy (average for each target, in mean square error)

UE Mobility Dataset	Latitude	longitude
TCNForecaster	5.48E-05	5.43E-04
AutoTS	4.72E-06	6.71E-05
Improved Rate	91.39%	87.6%

Table 2: Comparison of UE mobility model prediction accuracy (average for each target, in mean square error)

NF Load Dataset	Training throughput	Inference Latency
Stock PyTorch	580 samples/s	3.10 ms/sample
Chronos	2,079 samples/s	1.11 ms/sample
Improved Rate	258.4%	63.9%

Table 3: Comparison of NF load model training throughput and inference latency

UE Mobility Dataset	Training Throughput	Inference Latency
Stock PyTorch	180 samples/s	2.4 ms/sample
Chronos	3,352 samples/s	0.75 ms/sample
Improved Rate	1762%	68.75%

Table 4: Comparison of UE mobility model training throughput and inference latency Chronos relies on bigdl-nano [14] to speed up the training and inference process leading to the high throughput seen in the tests above. For training, Chronos adopts tcmalloc [15], intel openmp [16], suitable environment variables, and multi-process training for better memory allocation and parallelization performance. For inference latency, Chronos transforms the model to ONNX format and uses onnxruntime [17] to provide lower latency. These improvements are transparent to Chronos users.

Prediction accuracy and latency play a key role while generating analytics from NWDAF. Chronos AutoML on an Intel Xeon processor provides significant gains both in accuracy and throughput of predictions. The predicted load on the NFs is utilized in CPU frequency tuning, smart UPF selection, and NF placement, whereas the predicted UE location serves as an input for the directional paging and MICO mode selection. Also, Chronos AutoML ensures autotuning of model parameters which guarantees that the model is best adapted to the dynamic conditions of the network.

Conclusion

As the telecom industry marches toward becoming net-zero in the next few decades. the complexity of non-standardized energysaving mechanisms is becoming a blocker for achieving the desired energy reduction targets. AI/ML based insights, combined with technology enablers like NWDAF and ORAN RIC in 5G, have unlocked various innovative paths for the telco industry to design and execute such energy efficiency solutions in a standard and economical way. Project Bose is an example of one such innovative solution that was developed in collaboration with Intel. It has shown significant improvement in energy efficiency, as explained in the previous sections. But this is just the beginning and Project Bose will continue to add more innovative use cases in the future, to accelerate operator journeys toward becoming net-zero.

In the next phase of Project Bose, we plan to further strengthen the energy optimization capabilities by using additional metrics, related to slice level quality assurance, packet processing, memory fault monitoring and other analytics information provided by NWDAF. We will also integrate it with benefits that can be derived through ORAN RIC for maximizing the energy saving in 5G networks.



References

- 1. Netanticipate The ultimate Network AI platform, <u>https://bit.ly/3xDqXPL</u>
- 2. Harness the power of the Capgemini engineering NWDAF framework for your 5G ecosystem, https://bit.ly/3Qx01tG
- 3. ICT industry to reduce greenhouse gas emissions by 45 per cent by 2030, ITU, Geneva, 27 February 2020, <u>https://bit. ly/3xF5TZe</u>
- Putting Sustainability at the Top of the Telco Agenda, BCG, June 24, 2021, <u>https://on.bcg.com/39wa3e9</u>
- GSMA's 2021 Mobile Industry Impact Report on Sustainability, GSMA Intelligence, September 2021, <u>https://bit. ly/39xJo0m</u>
- Building sustainable 5G networks for a net-zero world, Fierce Wireless. December 13, 2021, <u>https://bit. ly/3xXQoNr</u>
- The essential role of AI in improving energy efficiency, GSMA, Emanuel Kolta & Tim Hatt, Nov 2021, <u>https://bit. ly/30khLqk</u>

- Green future networks metering for sustainable networks, Version 1.0, NGMN, 25. January 2022, <u>https://bit.ly/3bcLz9W</u>
- 9. Intel Observability framework, Intel, Jan 21, 2021, <u>https://intel.ly/39vbwBv</u>
- The sustainable telco, GSMA, Tim Hatt and Peter Jarich and Emanuel Kolta, , December 2021, <u>https://bit.ly/3zMJv31</u>
- Use Open Telemetry for better troubleshooting, Intel Kubecon Demo, <u>https://intel.ly/39AwoXT</u>
- 12. Open Telemetry, <u>https://opentelemetry.</u> io/
- 13. Telegraf, <u>https://www.influxdata.com/</u> <u>time-series-platform/telegraf/</u>
- 14. <u>https://github.com/intel-analytics/BigDL/</u> tree/main/python/nano
- 15. <u>https://github.com/google/tcmalloc</u>
- 16. <u>https://intel.ly/3zKyy1K</u>
- 17. <u>https://github.com/microsoft/</u> onnxruntime

Glossary

Acronym	Explanation
3GPP	Third Generation Partnership Project
5GC	5G Core Component
AI	Artificial Intelligence
AMF	Access and Mobility Management Function (AMF)
CG	Capgemini
CPU	Central Processing Unit
gNB	Next Generation NodeB
ICT	Information and Communications Technology
ITU	The International Telecommunication Union
MICO	Mobile Initiated Connection Only
ML	Machine Learning
MTLF	Model Training Logical Function
NF	Network Function
NFVi BKC	Network Function Virtualization Infrastructure Best Known Configuration
NGC	5G Next Generation Core
NWDAF	Network Data Analytic Function
OAM	Operations, Administration and Maintenance
ONNX	Open Neural Network Exchange
ORAN	Open Radio Access Network
QoE	Quality of Experience
RAN	Radio Access Network
RIC	Radio Intelligent Controller
SMF	Session Management Function
UE	User Equipment
UPF	User Plane Function

Notices & Disclaimers by Intel

Performance varies by use, configuration and other factors. Learn more at <u>www.Intel.com/</u><u>PerformanceIndex</u>.

See backup for workloads and configurations. Results may vary.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Intel technologies may require enabled hardware, software or service activation.

Your costs and results may vary.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy. © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

¹1 node, 2x Intel(R) Xeon(R) Platinum 8380H CPU @ 2.90GHz with 384 GB (12 slots/ 32GB/ 3200) total DDR4 memory, microcode 0x700001e, HT on, Turbo on, Ubuntu 9.3.0-10ubuntu2 20.04, 5.4.0-42-generic, Temporal Convolution Networks, Python 3, Pytorch, One API, ONNX runtime, AI performance for latency tested using 1 Xeon core, AI performance for training throughput using 56 Xeon cores. AI Performance testing is conducted by Capgemini on 22-Apr-2022

Author



Mayuri Sandeep Tendulkar

Capgemini contributors: Subhankar Pal, Subhash Chopra, Gaurav Jain, Chhaya Vikas Chablani, Atul Agarwal, Pankaj Rawat, Akanksha Dureja

Intel contributors: John Browne, Aibhne Breathnach, Chris Macnamara, Emma Collins, Anand Bodas, Junwei Deng, Herve Mer, Huisuk Hong



About Capgemini Engineering

World leader in engineering and R&D services, Capgemini Engineering combines its broad industry knowledge and cutting-edge technologies in digital and software to support the convergence of the physical and digital worlds. Coupled with the capabilities of the rest of the Group, it helps clients to accelerate their journey towards Intelligent Industry. Capgemini Engineering has more than 55,000 engineer and scientist team members in over 30 countries across sectors including Aeronautics, Space, Defense, Naval, Automotive, Rail, Infrastructure & Transportation, Energy, Utilities & Chemicals, Life Sciences, Communications, Semiconductor & Electronics, Industrial & Consumer, Software & Internet.

Capgemini Engineering is an integral part of the Capgemini Group, a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided every day by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of over 340,000 team members in more than 50 countries. With its strong 55-year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fueled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2021 global revenues of €18 billion.

Get the Future You Want

For more details, contact us:

www.capgemini-engineering.com

Write to us at:

engineering@capgemini.com