

# Senao Networks Delivers SASE Appliance for Network Edge

**Senao SA9820, based on Intel Atom® x7000RE processor series, is optimized for AI-based network security workloads, cost, and power consumption; tests show performance for phishing and portable execution use cases**



Network security has evolved and become a critical concern for organizations, who are now urged to consider new network security use cases to provide better defense against fast-changing cybersecurity threats.

Technologies, including artificial intelligence, need to evolve very quickly to stay ahead of always changing cybersecurity threats such as ransomware, zero-day attacks, phishing attacks and distributed denial of service (DDoS) attacks. The negative impact of these cybersecurity threats can be minimized using artificial intelligence / machine learning / deep learning (AI/ML/DL) combined with a dynamic rules database rather than a static rules database. New AI-driven detection capabilities can recognize and react to cyber threats faster.



Senao Networks, an Intel® Industry Solution Builders' Network Builders Community member, has developed the Senao SA9820, a network security appliance for enabling AI-powered network security use cases in branch offices or network edge locations to fortify these network locations with a system that can proactively monitor and adapt to evolving cyber threats in real-time.

## Senao SA9820 Delivers SASE with AI-Based Network Security

The Senao SA9820 is a SASE-enabled network desktop appliance for edge deployment. The system is based on the Intel Atom x7000RE processor series for the Edge. The SA9820 provides ample network I/O ports including two ports of 1GbE Base-T, two ports of 2.5GbE Base-T, and two ports of 1GbE SFP connection for network connectivity. The SA9820 uses an external power adapter.



**Figure 1.** Senao SA9820 front view.

The SA9820 has fast memory supporting up to 16GB LPDDR5 (8GB x 2), which is essential for supporting network security workloads. For storage, the device can support one 32GB eMMC flash and additional features via PCIe M.2 slots.

### Performance from Intel Atom x7000RE Processor Series for the Edge

For the SA9820, Senao chose the Intel Atom x7000RE processor series for the Edge because it is a powerful and energy-efficient choice for networks. With up to eight efficient cores, these processors excel in per core performance, offering robust processing power in compact and low-power designs with a TDP of between six watts and 12 watts. By combining low power consumption, advanced processing features, and integrated graphics, the Intel Atom x7000RE processor series is well-suited for the dynamic and demanding nature of edge SASE applications.

### Network Security AI Workload Accelerated by Integrated GPU for AI Processing

The Intel Atom x7000RE processor series for the Edge offers enhanced AI processing power via an integrated Gen 12 UHD GPU (iGPU). This iGPU functionality uses the Intel® Xe Architecture to boost the performance of AI applications by delivering faster inferencing for network security workloads.

The iGPU delivers up to 32EU and 1.2 GHz of turbo frequencies. This capacity allows a high degree of parallelization for AI workloads, combined with built-in AI acceleration with Intel® Deep Learning Boost (Intel® DL Boost) and INT8 support. The use of the iGPU offers significant AI performance to meet SASE workload requirements. It brings the overall total cost of ownership (TCO) down compared to board designs with discrete GPUs.

The Intel Atom x7000RE processor series for the Edge also features 2.5GbE ports and up to nine lanes of PCIe 3.0. It also comes with predictable performance with turbo disabled. It can achieve real-time capability with Intel® Time Coordinated Computing (Intel® TCC) and Time-Sensitive Networking (TSN).

By leveraging cutting-edge network and security acceleration technologies within Intel Atom x7000RE processor series for the Edge, the Senao SA9820 appliances meet the demand for multiple network services and applications, including Internet access, SaaS access, public cloud connection, and private cloud connection as shown in Figure 2.

The scalable, modular design supports easy network upgrades and expansions for businesses, ensuring efficient traffic management, robust security, and cloud-based management that can ideally promote the seamless integration of enterprises to achieve continuous performance.

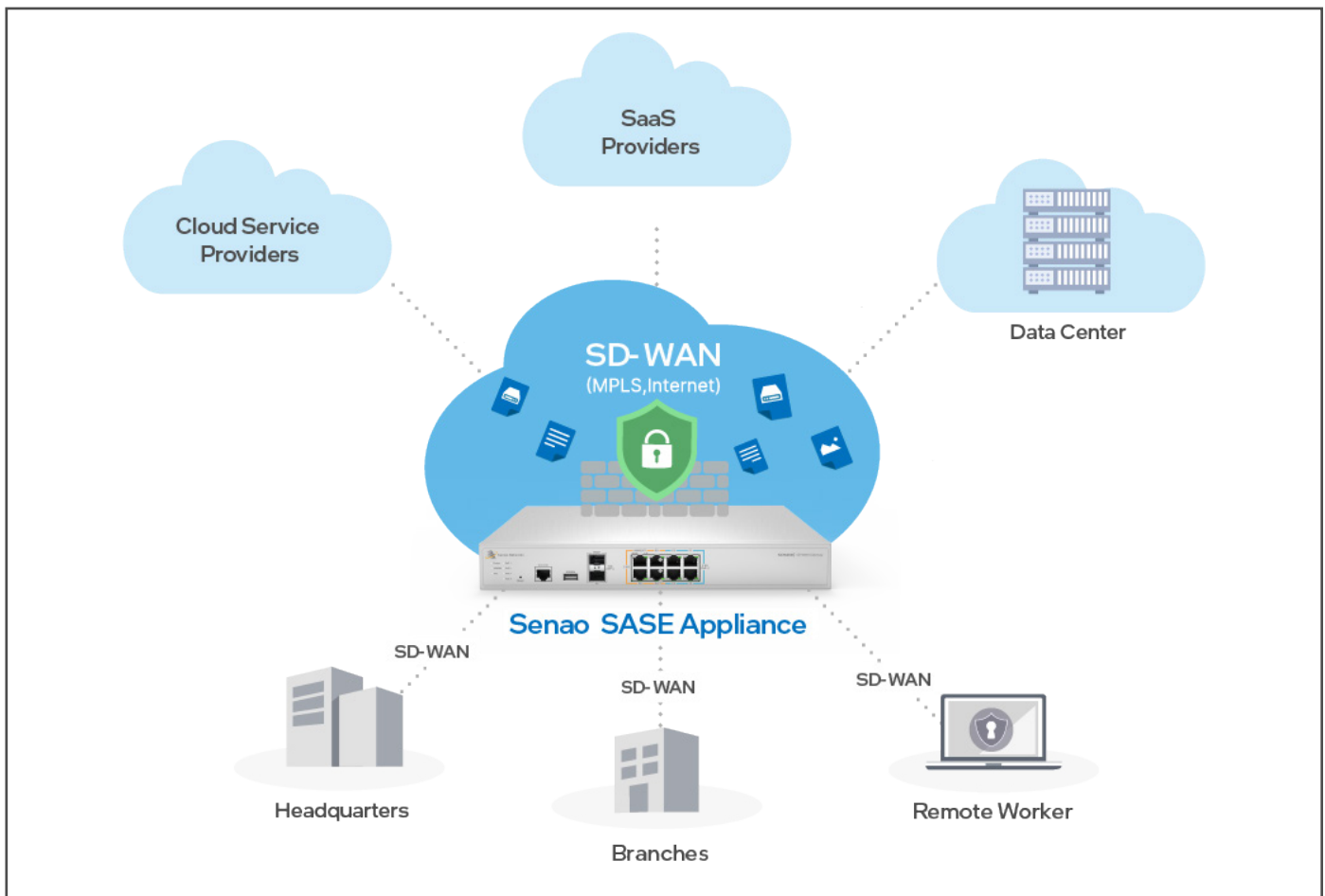


Figure 2. SA9820 in typical branch office network security AI applications.

## Senao SA9820 Performance Tests

Senao Networks measured the AI inference performance<sup>1</sup> of the Senao SA9820 in two typical network security use cases: portable executable (PE) detection and email phishing. The tests used the same hardware and software, with the exception of the load (BERT vs. MalConv) and AI frameworks (see complete configurations in the footnotes).

Three different AI frameworks were used to measure the latency of each network security use case: TensorFlow, Intel® Extension for TensorFlow (ITEX), and OpenVINO.

The Intel Atom x7000RE processor series for the Edge inside the Senao SA9820 was configured in two ways to examine appliance performance using the CPU and the iGPU. The iGPU offered significantly lower latency in both FP32 and INT8 tests when compared to the CPU.

The PE tests were done with the open-source MalConv model (<https://github.com/elastic/ember/tree/master/malconv>). The open-source MalConv model was converted to the FP32 model and then used Intel® Neural Compressor to quantize it to the INT8 model. The CPU testing with [codes and methods can be found in this downloadable PDF](#). It is recommended to use a virtual environment (venv) to run the test during the testing.

All package dependencies for CPU testing are listed in the requirements.txt as shown below:

```
# cat requirements.txt
tensorflow==2.15.0
openvino-dev==2024.2.0
nncf==2.9.0
neural-compressor==2.5
```

Meanwhile, ITEX is used to accelerate the MalConv model on the iGPU to further improve AI inference performance. Here are the reference steps to install ITEX:

Step 1: Install the iGPU driver for the Intel Atom x7000RE processor series for the Edge; the steps can be found at <https://dgpu-docs.intel.com/driver/client/overview.html>

```
# Install Compute Runtime, Media, and Mesa packages.
$ sudo apt install -y \
    intel-ocl-icd intel-level-zero-gpu level-zero \
    intel-level-zero-gpu-raytracing \
    intel-media-va-driver-non-free libmfx1 \
    libmfxgen1 libvpl2 \
    libegl-mesa0 libegl1-mesa libegl1-mesa-dev \
    libgbm1 libgl1-mesa-dev libgl1-mesa-dri \
    libglapi-mesa libgles2-mesa-dev libglx-mesa0 \
    libigdgmm12 libxatracker2 mesa-va-drivers \
    mesa-vdpau-drivers mesa-vulkan-drivers va- \
    driver-all vainfo hwinfo clinfo
```

Step 2: Install oneAPI Base Toolkit 2024.0; the reference steps and methods can be found at <https://networkbuilders.intel.com/solutionslibrary/lanner-delivers-ai-security-appliance-for-network-edge>

```
# Install oneAPI Base Toolkit Packages
$ wget https://registrationcenter-download.intel.com/akdlm//IRC_NAS/20f4e6a1-6b0b-4752-b8c1-e5eacba10e01/1_BaseKit_p_2024.0.0.49564.sh
$ sudo sh 1_BaseKit_p_2024.0.0.49564.sh
```

Step 3: Install package dependencies for iGPU testing as listed in the requirements-igpu.txt as shown below:

```
# cat requirements-igpu.txt
tensorflow==2.14.1
openvino-dev==2024.2.0
nncf==2.9.0
neural-compressor==2.5
intel-extension-for-tensorflow[xpu]==2.14.0.2
```

After the drivers and software are installed, the iGPU can be used to perform AI inference with zero code changes. Table 1 shows these test results.

Framework	Platform		Latency (ms)		Performance improvement INT8 vs FP32 w/oneDNN
			FP32 w/oneDNN	INT8	
TensorFlow (2.15.0)	Senao Networks SA9820	1 CPU	110.26	53.32	2.07x
ITEX (2.14.0.2)		iGPU	17.62	9.64	1.54x

**Table 1.** Portable execution latency results from CPU and iGPU using ITEX and TensorFlow frameworks.

The results show that the inference time can be reduced to 53.32 ms by performing model quantization to INT8 using the Intel Neural Compressor, a 2.07x performance improvement in latency after the quantization. The AI inference latency can be further improved to 9.64 ms in iGPU. That is an 11.44x boost from using 1 CPU (FP32) to using iGPU (INT8) with zero code changes.

### Phishing Performance

A similar test that used the Bidirectional Encoder Representations from Transformers (BERT) model was performed. One of the common use cases BERT is trained for is to detect phishing emails from text within the email. This test uses OpenVINO as a deep learning framework instead of TensorFlow. The detailed codes and methods can be found in this edge-developer-kit-reference-scripts [GitHub page](#).

Step 1: Git clone the edge-developer-kit-reference scripts

```
$ git clone https://github.com/intel/edge-developer-kit-reference-scripts.git
```

Step 2: Change directory to usecases/netsec/opencvino-bert/.

```
$ cd edge-developer-kit-reference-scripts/usecases/netsec/opencvino-bert/
```

Step 3: Intel has prepared a script to automate the installation. Running the script installs the dependencies.

```
$ bash setup_ov.sh
```

Note: OpenVINO 2024.2.0 was the latest version of OpenVINO when this document was written. For the latest releases and features of OpenVINO and more information, visit <https://www.intel.com/content/www/us/en/developer/tools/opencvino-toolkit/overview.html>. It is recommended to run the testing in a virtual environment. For this case, the script will create a "bert\_ov\_venv" in the same directory when setup\_ov.sh is completed.

The BERT base model was benchmarked with 512 max tokens inference time test on the Senao SA9820. The model was tested in 1 CPU core environment, with the command

```
$ numactl -C 0 benchmark_app -m models/quantized_bert_based_cased.xml -d CPU -hint latency -shape "[1, 512]"
```

The capabilities of Intel Atom x7000RE processor series for the Edge include the iGPU acceleration to test the model and it can be done using commands with no code change needed inside the benchmark\_app code with this command:

```
$ numactl -C 0 benchmark_app -m models/quantized_bert_based_cased.xml -d GPU -hint latency -shape "[1, 512]"
```

The benchmark test is shown in Table 2 and is performed with OpenVINO 2024.2.0 for the BERT base case model with 512 input tokens:

The AI inference latency can be reduced to 120.91ms in the iGPU testing environment. That is a 19.37x boost from using 1 CPU (FP32) to using iGPU (INT8) without any code changes.

Framework	Platform		Latency (ms)		Performance improvement INT8 vs FP32 w/oneDNN
			FP32	INT8	
OpenVINO (2024.2.0)	Senao Networks SA9820	1 CPU	2342.46	1268.99	1.84x
		iGPU	186.50	120.91	1.54x

Table 2. BERT test results using OpenVINO model.



## Conclusion

The benchmark test results shown in this paper demonstrate that the Senao SA9820 desktop type SASE is a good choice for network security use cases for different AI workloads. The Senao SA9820 can benefit from Intel's latest built-in AI acceleration with Intel Deep Learning Boost and iGPU featured in Intel Atom x7000RE processor series for the Edge. All NetSec AI use cases discussed are available in a single package NetSec SW package offered by Intel. Contact your representative account managers for more information.

## Learn More

[Senao Homepage](#)

[Senao SD-WAN Products \(includes SA9820\)](#)

[Intel Atom® x7000RE processor series](#)

[Intel Industry Solution Builders](#)



<sup>1</sup>SUT Config 1: Senao Networks SA9820 with 1x Intel Atom® x7433RE with eight cores. Total DDR memory was 16 GB (1 slot/16GB/4800 MHz); microcode v2.2; Intel® Hyper-Threading Technology - enabled. BIOS version: INSYDE Corp. Cradle\_v0.01. Software: OS was Ubuntu 20.04.4; kernel was Linux 6.5.0-41-generic. Benchmark/workload software: BERT base model (cased), with 512 input tokens; Compiler was GCC 11.4.0; Libraries were DPDK 22.11.0. Other software: OpenVINO 2024.2.0, Torch 2.3.1. Test conducted by Senao Networks Inc. on July 20, 2024.

SUT Config 2: Senao Networks SA9820 with 1x Intel Atom® x7433RE with eight cores. Total DDR memory was 16 GB (1 slot/16GB/4800 MHz); microcode v2.2; Intel® Hyper-Threading Technology - enabled. BIOS version: INSYDE Corp. Cradle\_v0.01. Software: OS was Ubuntu 20.04.4; kernel was Linux 6.5.0-41-generic. Benchmark/workload software: MalConv from Netsec software release 24.04 tokens. Compiler was GCC 11.4.0; Libraries were DPDK 22.11.0. Other software: OpenVINO 2024.2.0, TensorFlow is 2.15.0 for CPU testing. Intel® Extension for TensorFlow (ITEX) is 2.14.0.2 for iGPU testing. Test conducted by Senao Networks Inc. on July 20, 2024.

## Notices & Disclaimers

Performance varies by use, configuration and other factors.

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 optimizations, for Intel compilers or other products, may not optimize to the same degree for non-Intel products.

Your costs and results may vary.

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

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

See our complete legal [Notices and Disclaimers](#).

Intel is committed to respecting human rights and avoiding causing or contributing to adverse impacts on human rights. See Intel's [Global Human Rights Principles](#). Intel's products and software are intended only to be used in applications that do not cause or contribute to adverse impacts on human rights.

© Intel Corporation. Intel, the Intel logo, Xeon, the Xeon 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.