Time-sensitive networking optimizes the performance of advanced automotive and industrial systems

Introduction

Both the automotive and industrial automation sectors are going through massive changes. Both revolutions have been enabled by rapid digitalization and the availability of high-performance silicon to support sophisticated applications that are increasingly driven by artificial intelligence and machine learning. The high computational throughput of the SoCs that are now being distributed around the vehicle and the factory floor makes it possible to augment traditional control with more advanced functions, many of which are likely to be based on machine learning, a technology that requires the exchange of large amounts of real-time data. The key to these architectures is distributed intelligence and control combined with the ability to move large quantities of data not just rapidly but predictably.

An advanced driver-assistance system (ADAS) in a motor vehicle will, for example, call on multiple cameras and sensors to provide data to one or more onboard computers that use a combination of machine learning and digital signal processing to compute trajectories and make decisions. At the same time, within the cabin, occupants now expect to be able to view personal videos or play games using in-car displays or with smart mobile devices connected wirelessly to the vehicle's systems.

In the industrial environment, machine tools are no longer expected to operate as standalone systems but to call upon edge computers to perform detailed analyses of their operation. Those edge computers are expected to incorporate machine-learning algorithms to process video and other sensor data in real time to determine whether the tools are working correctly. That is putting stress onto the networking technologies that are used to interconnect these systems as traditional approaches cannot support the many use-cases that are now appearing.

Automotive and Industrial Evolution

Historically, both the automotive and industrial-automation sectors have relied on comparatively low-speed but deterministic fieldbuses such as CAN or Profibus to exchange data between electronic control units (ECUs) or programmable logic controllers (PLCs). This worked well when the individual controllers were self-contained and only needed to exchange small quantities of sensor and status data.

Because of the high data rates that today's Ethernet implementation are capable of, the network protocol is beginning to replace the older network and fieldbus technologies. Multiple applications can communicate over the same interconnect at high speed. That in the automotive space, for example, is leading to the rise of zonal architecture.

The zonal architecture makes it possible for OEMs to cut overall costs and increase performance by consolidating applications onto a smaller number of high-performance SoCs, with the functions for individual applications distributed across these SoCs. For example, a storage node can capture both instrument data from the vehicle's operations and movements and the content that occupants want to watch or listen to as they travel. Other nodes may relay streaming data to wireless devices used by the passengers so they can play games, make calls, or watch videos on their personal mobile devices. All communications are mediated by zonal gateways that enforce high security between the different applications, so that infotainment applications do not gain access to control functions in the engine, steering, or braking systems.

A similar zonal architecture is appearing in industrial systems, driven by advances in machine vision and machine learning. Rather than each machine tool being an independent unit linked to a core network through a low-bandwidth fieldbus, the machines now need to be able to send video and other rich sensor data to edge computers for processing. AI applications running on these edge computers can quickly determine whether the machine tools are operating correctly or whether changes need to be made to compensate for temperature or other changes to conditions.

The challenge with zonal architectures is ensuring that distributed applications using shared network infrastructure can achieve the same real-time performance as systems based on a more traditional architecture where individual SoCs were able to implement all required functions and just exchange sensor data over dedicated fieldbuses or deterministic networks.

Media players can use high degrees of buffering to overcome transfer delays caused by the best-effort nature of conventional Ethernet. But with multiple independent displays and outputs around the vehicle and in an environment where content is delivered wirelessly on demand, which incurs additional latency, there is a much greater need to ensure that data is delivered in a predictable manner. In the zonal architecture, data may need to pass through multiple nodes and network bridges, which will increase latency and unpredictability. For example, the mass storage ECU may be in a zone that is primarily concerned with ADAS and other real-time controls but, to take advantage of its storage capacity, also acts as a content-delivery system for the onboard infotainment systems in the cabin.
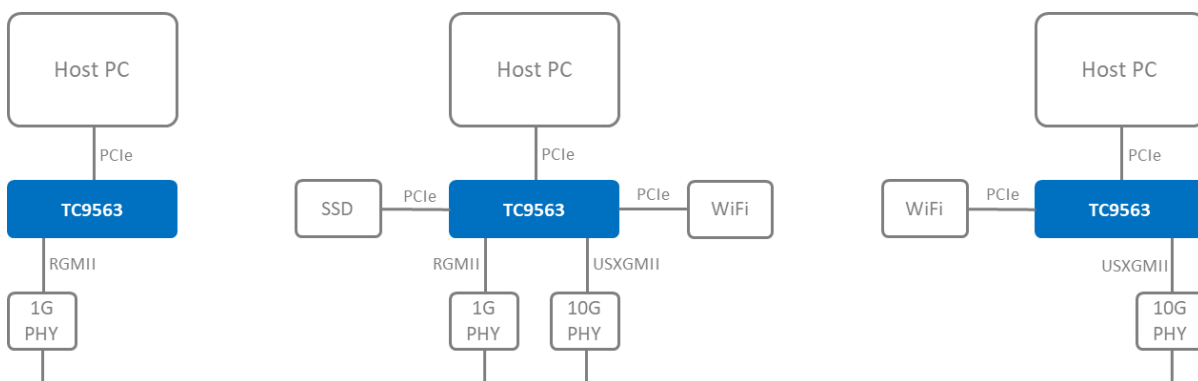
If different high-bandwidth streams interfere with each other to the point that occupants of the car will experience stalls in a video at different times, the overall perception will be of a system that cannot keep up. Therefore, even if the primary application is infotainment, predictable delivery of packets for streaming services is vital.

Though one answer is to increase headroom by moving to higher-speed forms of Ethernet, even this can prove to be insufficient where packets intended for different streams block each other because they happen to enter the network buffer at almost the same time. This behavior has been researched by Toshiba in a reference platform intended to show how applications in automotive and industrial systems are affected by unpredictability in the network.

Network Assessment

In the testbed created by Toshiba engineers, high-speed Ethernet controllers, able to support both 1Gbps and 10Gbps network interfaces are each connected to a host SoC that runs applications code to simulate the behavior of various expected use-cases in the automotive and industrial-automation environments.

In the automotive implementation, one of the controllers is attached to a solid-state drive that is used to emulate the behavior of an in-vehicle mass-storage element. Both also connect to WiFi controllers that are used to implement the basis of an in-vehicle wireless hotspot. Code implemented on the host SoCs was designed to emulate the typical applications in automotive: a zonal gateway; audio-video streaming; wireless-hotspot support; and mass-storage transfers.

Host PC		Host PC			Host PC	
PCIe		PCIe			PCIe	
TC9563		SSD — PCIe — **TC9563** — PCIe — WiFi			WiFi — PCIe — **TC9563**	
RGMII		RGMII	USXGMII		USXGMII	
1G PHY		1G PHY	10G PHY		10G PHY	

A similar reference design also served as a demonstration of bounded latency and low jitter with high-bandwidth Ethernet in industrial-automation environments. In this platform, one of the Ethernet-controller and SoC pairs represents a talker machine that is connected directly to a motor-control board that determines movement of a tool in the y and z coordinates. The other pair of devices is used to represent an industrial controller attached directly to a motor controller for the x-axis. This receives commands, relayed through two Ethernet switches, from the talker that is intended to be used to control movement on the listener's side of the system. The commands transferred between the two SoCs may be mixed in with video traffic from cameras that show the industrial operation being carried out, and which may be

used by a machine-vision application to assess the quality of operations and ensure correct alignment of a machine-tool head.

Control loops are typically designed to expect updates from sensors at a constant rate. They can develop instabilities and generate incorrect positioning commands if the updates and sensor readings the control algorithms are expecting are delayed beyond the processing window. The machine is effectively forced to use data that is likely out of date, and which no longer represents the situation in reality.

The original, baseline Ethernet standard was based on a best-effort principle: any peer with a packet ready to send can access the network and block other activity on that network segment until it is finished. This architectural approach reaches further back into the design of many conventional Ethernet controllers. In these designs, there is a simple queue that will not differentiate between the types of traffic the controllers need to send or receive. A high-priority command encapsulated in a short packet can easily be forced to wait for one or more lengthy video frames to be sent before it is allowed onto the network. These kinds of scenarios are readily demonstrated on platforms such as Toshiba's testbed.

In Toshiba's tests on these reference designs, use of the conventional IEEE 802.1 Ethernet protocol set can lead to a small percentage of the packets that need predictable delivery being heavily delayed. Though the number of delayed packets is small relative to those that are delivered on-time, the effect on a real-time application can be serious. In cases where the low-criticality traffic consumes relatively little bandwidth, less than 1Mbps in a 1Gbps network, the path delay of high-criticality traffic, such as packets used to carry motor-control commands, was measured to be up to 60µs in the experiment, where packets were transferred from talker to listener over three hops through two switches. With larger quantities of low-priority background traffic, consuming up to 840Mbps of bandwidth, the maximum path delay grew more than seven-fold to almost 440µs. The distribution of the path delays became strongly bimodal under these conditions, with many packets experiencing delays of more than 300µs, though a large proportion reached their destination in less than 60µs. In the tests of the effects of these delays in an industrial

scenario, with packets arriving later than a programmed deadline being dropped instead of being used to provide potentially incorrect data to a control loop, the result was a failure to achieve smooth movement and to reach programmed target positions.

Time-sensitive Networking

Enhancements introduced by the IEEE 802.1 committee can ensure Ethernet can operate in time-sensitive networking (TSN) environments and overcome the problems of compatibility that faced earlier proprietary attempts to deliver real-time performance, such as EtherCAT and Profinet IRT/RT.

The TSN protocols for Ethernet provide OEMs and integrators with the ability to use the high-speed network for real-time and streaming-media applications. These standards include IEEE 802.1AS, the Generalized Precision Time Protocol (gPTP), which defines timing and synchronization behavior for Ethernet in TSN environments, and the IEEE 802.1Qbv standard, which adds enhancements to allow traffic to be sent according to a strict schedule.

Timing precision is important on any network where real-time behavior is required. The problem faced by the distributed controllers in a networked environment is that subtle differences in voltage, temperature and other conditions can cause the clocks to drift apart from each other. The gPTP protocol was created to ensure that deviations above a certain level can be corrected automatically and ensure that to a given level of precision, all controllers on the network agree on the current time.

The use of gPTP allows all the nodes on an Ethernet network to agree on a common time, within a certain degree of error. In doing so, the network provides the basis for protocols that ensure nodes can agree time boundaries for delivery of packets with a known end-to-end transmission latency. Though the time protocol on which the standard is based demanded the use of high-precision clocks, the gPTP allows the endpoint devices on an Ethernet network to just agree on a common local time with sub-microsecond precision. To account for path delays across the network, gPTP nodes measure packets' residence time within each bridge and the latency incurred by each hop. These delays are then referenced to the node that is designated as providing the grandmaster-clock function in the bridged network.

Support for gPTP allows the use of time-aware shaping and packet scheduling. The idea behind traffic shaping is to prevent packets with latency constraints from being affected by packets that can be delivered on a best-effort basis. Traffic shaping distributes packets more evenly in time to prevent packets from building up in the queues inside Ethernet interface controllers and therefore being prevented from

being delivered within their latency window.

Some degree of traffic shaping can be delivered with conventional Ethernet as controllers that conform to the IEE 802.1Q can be provided with priorities. Each priority setting is used to add packets to a queue for that designated priority, letting them pre-empt low-priority traffic that might have been waiting in a queue for longer. However, this is not enough to deliver guaranteed on-time delivery. For example, high-priority packets may be held up by long transmissions of low-priority data that have already started when the packet arrives in the queue of a downstream controller.

Many real-time systems use the concept of time-slicing to ensure on-time delivery and processing of data. In these systems, part of the bandwidth is dedicated to regular time slices to which only packets with a defined tag have access. This is the concept that lies behind the IEEE 802.1Qbv time-aware shaper.

The time-aware shaper of IEEE 802.1Qbv requires the use of gPTP to ensure that all Ethernet devices, including those on bridged connections, can synchronize to use the same time windows. The IEEE 802.1Qbv time-aware scheduler separates transmission windows into a fixed-length cycle. Within each cycle, different time slices can be allocated to one or more of the eight IEEE 802.1Q priority levels. In doing so, applications that use these priorities are granted exclusive use of the network for the length of that time window. All other priorities use the time between these time-sliced windows. This service is in addition to an earlier credit-based traffic shaper that was introduced with the IEEE 802.1Qav standard, which can be used to send packets that need better handling of time-sensitive data but where hard timing guarantees are not required. Traffic may use the credit-based system, for example, to deliver video frames where end-to-end latency is not critical and so take lower priority than interactive voice or mission-critical sensor data that employs the IEEE 802.1Qbv priority-based time slices.

Performance Assessment

Though the TSN enhancements to Ethernet provided by standards, such as gPTP and IEEE 802.1Qbv, promise to deliver reliable communications for real-time applications, it is important to be able to demonstrate those features working in environments that mimic real-world use-cases and to understand how the time-aware shaping of TSN performs in scenarios with mixed-traffic patterns.

Toshiba's testbeds for automotive and industrial computing are built around Ethernet controllers, such as the TC9562 and TC9563, that implement gPTP and IEEE 802.1Qbv functionality and so have allowed its engineers to explore the effect of TSN approaches on representative applications. Compared to the situation with time-aware shaping disabled, timing was demonstrated to be far more consistent even

under high traffic conditions. Whereas path delay in the three-hop system ranges from 50µs to 430µs in the non-TSN environment, the distribution for time-sensitive packets reduced to a much narrower normal distribution centered on 25.5µs and with a jitter of just 0.3µs.
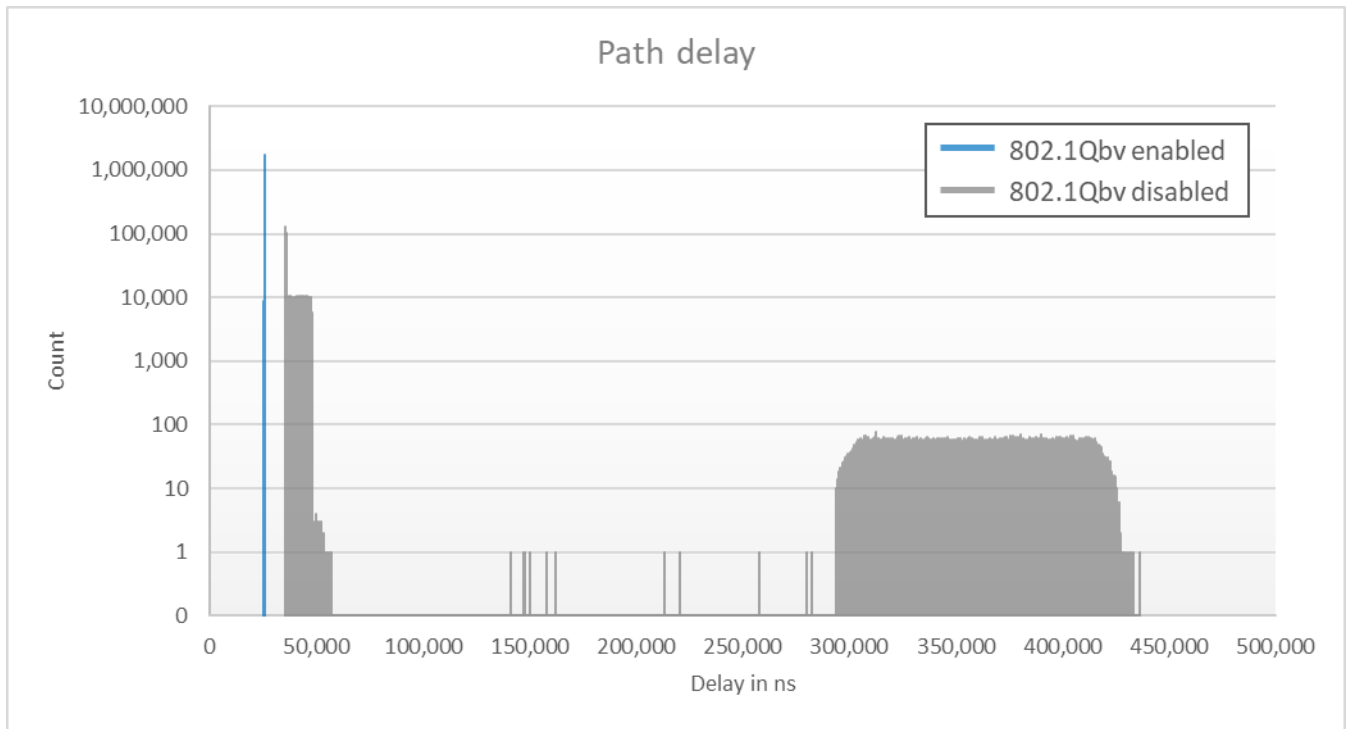
The reduction in overall path delay is assisted by improved behavior in the transmitting node. Whereas the transmission time variance of cyclic traffic can reach almost 8µs when time-aware shaping is disabled, the variance was never longer than 160ns when time-aware shaping was used to schedule high-criticality packets.
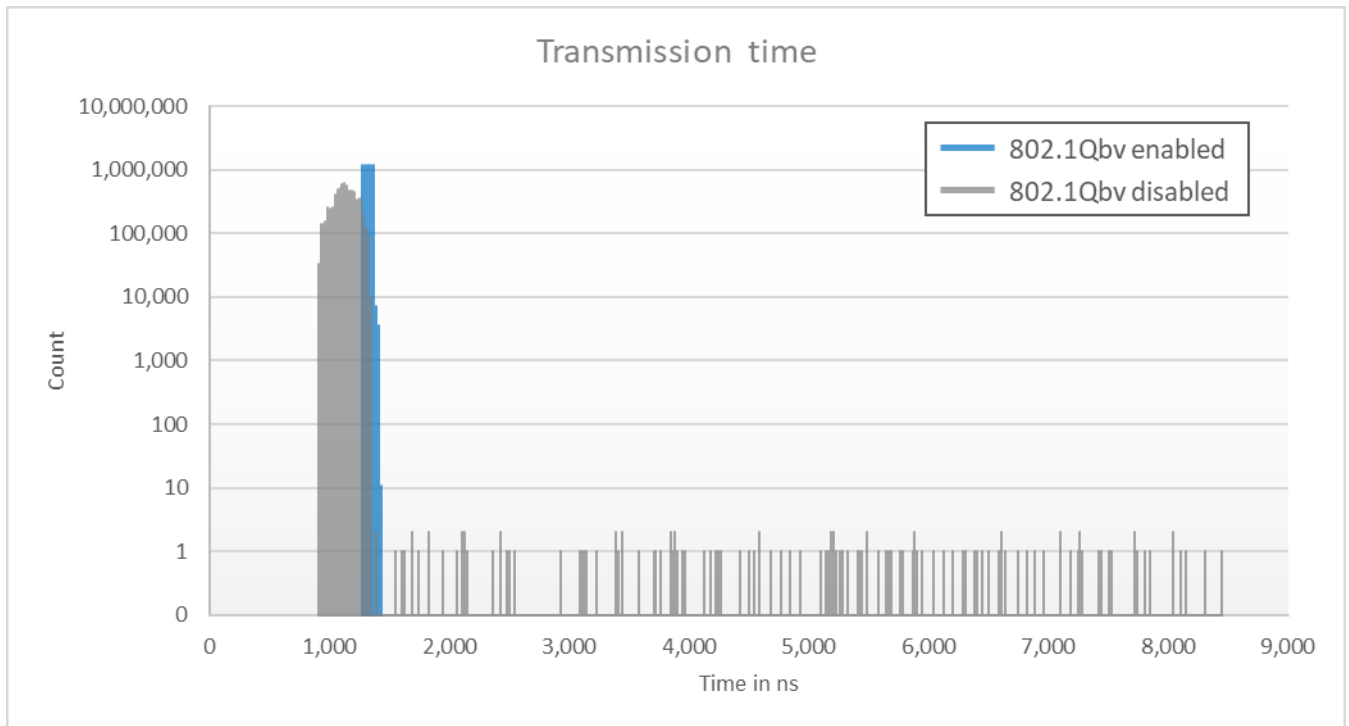
The result of the improvements to latency in the industrial demonstrator meant no packets were late. This allowed smooth control of all machine-tool axes, with no missed positioning targets.

Key Implementation Features

The TC9562 and TC9563 are important elements for building reliable TSN support into automotive and industrial-control systems. Both are highly integrated Ethernet controllers with full support for gPTP, IEEE 802.1Qav, IEEE 802.1Qbv, and other elements needed for reliable real-time communications where high bandwidth is a key requirement. The TC9562 provides support for 1Gbps Ethernet and is highly suited to both automotive and industrial-automation environments. The TC9563 extends the networking capability to two ports both able to support 10Gbps. The device was designed for the increasingly demanding requirements of systems that are based on automotive zonal architectures. The TC9563 is also suitable for industrial systems where there is a requirement for sharing large quantities of video or high-speed sensor data with AI-enabled edge-computing systems.

An on-chip Arm Cortex-M3 provides the ability to run monitoring and control software directly on the Ethernet controller to help improve overall reliability in the system and allow reporting of error conditions that may affect operation. The TC9562 implements a PCIe Gen 2.0 interface and the TC9563 incorporates a PCIe switch with support for one upstream and two downstream Gen 3.0 ports. The PCIe interfaces

provide high-speed communication between peripherals and the host SoC and enable bridging to WiFi controllers.

As well the advanced queuing and packet-delivery support needed for TSN, the TC9562 and TC9563 help optimize the performance of time-critical applications. For example, TC9563 supports single-route I/O virtualization (SR IOV) to support host SoCs that run virtual machines that are expected to host a mixture of real-time and other operating systems. SR IOV reduces the processor load when each of the operating systems makes an I/O call, which increases overall bandwidth and responsiveness, allowing full use of the 10Gbps transmission rate in the TC9563.

Summary

In both the automotive and industrial-automation environments, the combination of real-time performance and high-bandwidth communications is becoming a key requirement. Services such as AI-enhanced control and multiuser multimedia delivery are only possible if applications can rely on packets being delivered within programmed time windows. The TSN enhancements to Ethernet provide the basis for this behavior. Their implementation in devices such as Toshiba's TC9562 and TC9563 ensure that OEMs and systems integrators have access to the TSN support they will need to implement high-bandwidth, real-time networks inside advanced vehicles and factory systems.

References

1. Toshiba Automotive Ethernet Bridge ICs: https://toshiba.semicon-storage.com/eu/semiconductor/product/automotive-devices/automotive-interface-bridge-ics/automotive-ethernet-bridge-ics.html
2. Press release TC9563: https://toshiba.semicon-storage.com/eu/company/news/2022/01/automotive_20220112-1.html
3. At the engineer's bench | Time Sensitive Networking: https://www.youtube.com/watch?v=M7LPbkQL4C4&list=PLvoge_1w_Th8LA-hsgXKnQJtYEfanCu91

Restrictions On Product Use

Toshiba Corporation and its subsidiaries and affiliates are collectively referred to as "TOSHIBA".
Hardware, software and systems described in this document are collectively referred to as "Product".

- TOSHIBA reserves the right to make changes to the information in this document and related Product without notice.

- This document and any information herein may not be reproduced without prior written permission from TOSHIBA. Even with TOSHIBA's written permission, reproduction is permissible only if reproduction is without alteration/omission.

- Though TOSHIBA works continually to improve Product's quality and reliability, Product can malfunction or fail. Customers are responsible for complying with safety standards and for providing adequate designs and safeguards for their hardware, software and systems which minimise risk and avoid situations in which a malfunction or failure of Product could cause loss of human life, bodily injury or damage to property, including data loss or corruption. Before customers use the Product, create designs including the Product, or incorporate the Product into their own applications, customers must also refer to and comply with (a) the latest versions of all relevant TOSHIBA information, including without limitation, this document, the specifications, the data sheets and application notes for Product and the precautions and conditions set forth in the "TOSHIBA Semiconductor Reliability Handbook" and (b) the instructions for the application with which the Product will be used with or for. Customers are solely responsible for all aspects of their own product design or applications, including but not limited to (a) determining the appropriateness of the use of this Product in such design or applications; (b) evaluating and determining the applicability of any information contained in this document, or in charts, diagrams, programs, algorithms, sample application circuits, or any other referenced documents; and (c) validating all operating parameters for such designs and applications. **TOSHIBA ASSUMES NO LIABILITY FOR CUSTOMERS' PRODUCT DESIGN OR APPLICATIONS.**

- **PRODUCT IS NEITHER INTENDED NOR WARRANTED FOR USE IN EQUIPMENTS OR SYSTEMS THAT REQUIRE EXTRAORDINARILY HIGH LEVELS OF QUALITY AND/OR RELIABILITY, AND/OR A MALFUNCTION OR FAILURE OF WHICH MAY CAUSE LOSS OF HUMAN LIFE, BODILY INJURY, SERIOUS PROPERTY DAMAGE AND/OR SERIOUS PUBLIC IMPACT**("UNINTENDED USE"). Except for specific applications as expressly stated in this document, Unintended Use includes, without limitation, equipment used in nuclear facilities, equipment used in the aerospace industry, lifesaving and/or life supporting medical equipment, equipment used for automobiles, trains, ships and other transportation, traffic signalling equipment, equipment used to control combustions or explosions, safety devices, elevators and escalators, and devices related to power plant. **IF YOU USE PRODUCT FOR UNINTENDED USE, TOSHIBA ASSUMES NO LIABILITY FOR PRODUCT.** For details, please contact your TOSHIBA sales representative or contact as via our website.

- Do not disassemble, analyse, reverse-engineer, alter, modify, translate or copy Product, whether in whole or in part.

- Product shall not be used for or incorporated into any products or systems whose manufacture, use, or sale is prohibited under any applicable laws or regulations.

- The information contained herein is presented only as guidance for Product use. No responsibility is assumed by TOSHIBA for any infringement of patents or any other intellectual property rights of third parties that may result

from the use of Product. No license to any intellectual property right is granted by this document, whether express or implied, by estoppel or otherwise.

- **ABSENT A WRITTEN SIGNED AGREEMENT, EXCEPT AS PROVIDED IN THE RELEVANT TERMS AND CONDITIONS OF SALE FOR PRODUCT, AND TO THE MAXIMUM EXTENT ALLOWABLE BY LAW, TOSHIBA (1) ASSUMES NO LIABILITY WHATSOEVER, INCLUDING WITHOUT LIMITATION, INDIRECT, CONSEQUENTIAL, SPECIAL, OR INCIDENTAL DAMAGES OR LOSS, INCLUDING WITHOUT LIMITATION, LOSS OF PROFITS, LOSS OF OPPORTUNITIES, BUSINESS INTERRUPTION AND LOSS OF DATA, AND (2) DISCLAIMS ANY AND ALL EXPRESS OR IMPLIED WARRANTIES AND CONDITIONS RELATED TO SALE, USE OF PRODUCT, OR INFORMATION, INCLUDING WARRANTIES OR CONDITIONS OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, ACCURACY OF INFORMATION, OR NONINFRINGEMENT.**

- Do not use or otherwise make available Product or related software or technology for any military purposes, including without limitation, for the design, development, use, stockpiling or manufacturing of nuclear, chemical, or biological weapons or missile technology products (mass destruction weapons). Product and related software and technology may be controlled under the applicable export laws and regulations including, without limitation, the Japanese Foreign Exchange and Foreign Trade Law and the US Export Administration Regulations. Export and re-export of Product or related software or technology are strictly prohibited except in compliance with all applicable export laws and regulations.

- Please contact your TOSHIBA sales representative for details as to environmental matters such as the RoHS compatibility of Product. Please use Product in compliance with all applicable laws and regulations that regulate the inclusion or use of controlled substances, including without limitation, the EU RoHS Directive. **TOSHIBA ASSUMES NO LIABILITY FOR DAMAGES OR LOSSES OCCURRING AS A RESULT OF NONCOMPLIANCE WITH APPLICABLE LAWS AND REGULATIONS.**