IOWN GLOBAL FORUM

INNOVATIVE OPTICAL AND WIRELESS NETWORK GLOBAL FORUM VISION 2030 AND TECHNICAL DIRECTIONS

Overview

Empowered by the evolution of photonic networking technologies and the emergence of Silicon Photonics technologies, the Innovative Optical Wireless Network Global Forum (IOWN GF) aims to develop technologies that will lead to quantum leap improvements in communication capacity and latency, data efficiency, computing scalability, and energy efficiency. A new communication and computing full stack architecture will be developed. Leveraging IOWN GF technologies, smart world applications will be created and augmented with beyond-human cognitive-communication capacity, beyond-human response speed, linear computing scalability, and superior energy efficiency. The accomplishment of this vision requires global collaboration and endeavor cross communication, computing, software platform, and applications.

1. The World of 2020 and 2030

1.1. Paradigm Shifts in Connected World

The world today has experienced faster than ever growth, thanks to advancements in communication and computing technologies. Moving forward, another quantum leap in computing and communication capabilities is expected to empower the world toward a new era of growth. The mission of IOWN GF is to develop fundamental technologies on communication, computing, data, and energy efficiency that would bring in quantum leap performance improvement and enable a much smarter world with advanced applications, including those with digital twin computing.

IOWN GF's technology quantum leaps will enable several paradigm shifts in smart and connected world applications. One such paradigm shift is in data collection from digital to natural.

While digital technologies have enabled innovations in many fields, today, digital technologies are optimized for the human cognitive system. For example, the frame rate for motion video, e.g. 30 frames per second (FPS), is chosen considering human's visual acuity for motion. In addition, sampled audio and video data are compressed with mechanisms that take advantage of the masking effect of human's cognitive system. For human cognition, such encoded audio and video can be considered as fine quality. However, for use cases that require beyond human cognition, the quality is far from enough. Technologies should evolve to capture, cognize, and control the physical world without being constrained by human's capabilities and achieving this will lead to innovations in many fields. For example, manufacturing robot monitoring systems may detect anomaly from sound beyond audible frequency. Moreover, we expect that, driven by the automobile industry's investment in Automobile Driver Assistance System (ADAS), sensors for beyond-human sensing will rapidly evolve.

Besides cognitive capability, response speed is another factor that would require beyond human capabilities. It takes about 100 msec for an ordinary person to react upon seeing an event, therefore 100 msec turn-around-time (TAT) are sufficiently responsive for applications designed for human usages. However, for applications beyond human usages, e.g., emergency stop systems, further shortening on TAT is very much needed.

1.2. Smart and Connected World Applications

To visualize how IOWN GF technology can shape the future smart and connected world, we use some examples of smart world applications to illustrate as showed in Table 1. For the purpose of discussion, we classify smart world applications into four types. Detailed definitions of the four types can be found in Appendix 1. Each row of the table represents an area in terms of use case area. Please note that this table is far from exhaustive and is not indented to define applications that IOWN GF targets or prioritizes. IOWN GF will conduct use case exploration activities, adding new use cases based on inputs from its members.

Additional Note: Writing this whitepaper amid the COVID-19 pandemics, we see the urgent needs for innovative technologies that bring big data, communication and computing together, to combat this unprecedented pandemic as well as future crisis. Many technologies today have displayed remarkable power in saving human lives, from infection tracking, to emergence response, to automated shipping, to virtual meetings. The technology platforms developed in IOWN GF will directly multiply the power and effectiveness of our ability in combating future pandemics and other nature disasters.

Table 1: Examples of Smart and Connected World Use Cases

TYPE / AREA	COGNITIVE SYSTEMS	AUTONOMOUS SYSTEMS	XR (AUGMENTED / MIXED REALITY)	SYSTEMS FOR PREDICTION, SIMULATION, OR PROACTIVE CONTROL
Smart Energy	Detection of energy waste situation	Energy consumption control using cognitive systems		Energy demand and supply prediction (virtual powerplant)
Smart Area Management (Smart City)	People/traffic counting Structural health monitoring Damage survey Border monitoring Crime/accident detection	People flow/traffic control using cognitive systems Autonomous public safety, border control, and rescue systems	Security surveillance with XR	People flow/traffic optimization Earthquake/flood simulation predicting damaged areas
Smart Mobility	Street accident warning	Street accident prevention Collaborative autonomous vehicles, picking robots, drones	Extended vision for remote driving	Mobility demand and supply prediction
Smart Industries (Manufacturing/ Agriculture)	Manufacturing robot monitoring Agricultural field monitoring	Emergency stop systems for plants Precise fertilizer/pesticide application Collaborative industrial/agricultural robots	Navigation for novice workers	Process time estimation, process optimization Growth/disease prediction for harvest
Smart Health Care and Eldercare	Remote surgical operation Patient monitoring	Collaborative autonomous care service systems	Guidance for elderly people with XR	Disease prediction

TYPE / AREA	COGNITIVE SYSTEMS	AUTONOMOUS SYSTEMS	XR (AUGMENTED / MIXED REALITY)	SYSTEMS FOR PREDICTION, SIMULATION, OR PROACTIVE CONTROL
Smart Entertainment	Natural UI for video game Live event viewing with immersive experience, e.g. music and sports		Video game with XR Event viewing with XR Virtual trip experience	Gaming, e.g. role play, in digital twin world

1.3. Key Requirement Dimensions and Performance Levels

Most of the above use cases have already been implemented at performance levels achievable with today's technologies. However, we expect a boost in applications capabilities and usage scenarios with the quantum leap in computing and communication infrastructures enabled by IOWN GF technologies. To be more specific, we expect IOWN GF technologies could bring in new levels of performance in the following four dimensions.

A. Cognitive and Communication Capacity

Cognitive Capacity is defined as how finely, precisely, and multisensory the system can capture objects in the physical world and process the captured data. In cognitive systems, the process typically includes cognizing the status of monitored objects.

As mentioned earlier, today's digital systems are designed for the human cognitive systems, which are very limited in comparison with today's sensing technologies. Some application systems can be greatly upgraded by capturing data beyond the human's cognitive capacity. For instance, in manufacturing robot monitoring, motion captured at 120 FPS may help the system detect anomaly. Cognitive capacity can be defined in three performance levels as showed in Table 2.

Table 2: Cognitive Capacity Performance Levels

LEVEL	DEFINITION
Level 1	Low Quality Examples: 10 FPS motion, blocky or blurred images
Level 2	Human-Level (High-Quality Digital) Examples: 30 FPS motion, 20-20,000 Hz sound, encoded audio/image at fine-for-humans quality
Level 3	Beyond Human (Natural) Examples: 120 FPS motion, 20-150,000 Hz sound, extra senses: Brix, heat, precise shape, precise location, precise spatial position/geolocation

It should be noted that the sensor's capacity is not the only factor that determines the cognitive capacity. Rather, it is determined by the capacity of an end-to-end system that covers sensors, networks, and computing infrastructures. A more detailed discussed in this regard can be found in Section 2 of the paper.

B. Response Speed

Response Speed is defined as the responsiveness of a system to an event, such as status change of a controlled object. It can be represented with turn-around-time (TAT). Response Speed can be defined in three performance levels as showed in Table 3. The 0.1 second human response time is considered as Level 3. As will be discussed in Section 2, even achieving this level of response speed in a networked application system is technically challenging.

Table 3: Response Speed Performance Levels

LEVEL	DEFINITION
Level 1	Act after several minutes
Level 2	Act in a second
Level 3	Human Level: Act in 0.1 second
Level 4	Act in 10 milliseconds

C. Scalability in Computing

Scalability in Computing is defined as the system's capability of accommodating varying and uncertain workload while achieving high resource utilization. In many cases, we adopt the scale-out approach where the computing capacity is scaled up or down by increasing or decreasing the number of computing resources. To address dynamic workloads and enable efficient usage of the computing resource, a system that achieves dynamic linearly computing scaling is required. Scalability in computing can be defined in three performance levels as showed in Table 4.

Table 4: Scalability in Computing Performance Levels

LEVEL	DEFINITION
Level 1	Semi-static computing scaling
Level 2	Dynamic computing scaling
Level 3	Dynamic computing scaling with linear computing efficiency

D. Energy Efficiency

Energy Efficiency is defined as how energy efficient the system can run. Three performance levels can be defined as showed in Table 5. To explain the three levels of energy efficiency, we use an example in camera-based cognitive systems. Suppose we use a high-end server with two GPGPU cards for AI inference in the cognitive system. One node should be able to handle 600 image frames per second or accept 60 cameras, and the average energy consumption of one node would be around 600-900W (assuming PUE is 2).

A traditional on-premise approach would deploy at least one server regardless of the number of connected cameras. Despite intelligent energy saving mechanisms by modern computing system, the energy consumption could still be high when under-utilized. This scenario corresponds to Level 1.

If we deploy the cognitive system in a cloud infrastructure and dynamically adjust the resource allocation, we may be able to make resource consumption linearly related to the number of connected cameras. This would be a great improvement from the on-premise approach. This scenario corresponds to Level 2. However, even with this approach, we would still have to use 10-15 watts per camera, i.e., "one light bulb per camera".

Energy consumption will be greatly improved with an event-driven approach. If we deploy the cognitive system on a serverless computing platform and add a filtering mechanism that forwards an image to the serverless platform only when it contains objects that have to be cognized, resource consumption will be a function of the event occurrence rate. This scenario corresponds to Level 3. With this approach, energy consumption would be 1-1.5 watt-second per event. To further reduce energy consumption in Level 3, technologies such as silicon photonics can be used. Energy-efficient computing with silicon photonics should reduce energy consumption to "LED light level".

Table 5: Energy Efficiency Performance Levels

LEVEL	DEFINITION
Level 1	Dominated by very large fixed cost
Level 2	Unit-driven consumption, e.g. 10-15 W per camera = "light bulb level"
Level 3	Event-driven consumption, e.g. 300mWsec per object detection event = "LED light level"

1.4. Evolution from 2020 to 2030

Today's systems are at around performance Level 1 or Level 2 across the four dimensions. IOWN GF's quantum leaps in communication capacity/latency, data efficiency, computing scalability, and energy efficiency will lead to a level up in system capabilities. Table 6 explains how systems will be upgraded with IOWN GF technologies. These upgrades can only be achieved with full-stack architectural design which is the goal of IOWN GF.

Table 6: System upgrades enabled by IOWN GF technologies

	2020 (NOW)	2030 (IOWN GF'S VISION)
Cognitive-Communication Capacity	Level 1 - Level 2 - Data is captured and digitized based on human cognitive capacity - High compression ratios are often applied due to the network capacity constraints. - The lack of a precise calibration mechanism for time and geolocation causes data from different data sources disoriented.	Level 3 - High-density and multi-sensory data are captured by beyond human sensors - High communication capacity allows sending massive sensed data without high compression. - Reduced latency and jitter enables precision time and geolocation stamping, which eventually allows us to construct a digital replica of the physical world, i.e. digital twins
Response Speed	Level 1-Level 2 - Data is sent over TCP/IP. Sending several hundreds of kilobytes requires several round trips of communication.	Level 3-Level 4 - Accelerated L3/L4 protocols and data transfer functions enables instant data transfer - The instant data transfer would eventually reduce buffering in data source nodes.
Scalability in Computing	Level 1 – Level 2 Semi-static computing scaling, mostly within a single data center and registered to a limited kind of workloads	Level 3 Dynamic computing scaling across device, edge, center cloud with linear computing efficiency.

	2020 (NOW)	2030 (IOWN GF'S VISION)
Energy Efficiency	Level 1-Level 2 - Server resources are scaled up/down based on the lifecycle of sensor nodes. - AI inference for an HD full-motion camera consumes tens of watts, which is equivalent with light bulb's energy consumption	Level 3 - Computing scaling architecture leveraging Silicon Photonics interconnect will enable energy consumption on a per-event basis. - AI inference for an HD full-motion camera consumes single digit watts, which is equivalent with LED light's energy consumption

2. IOWN GF Technology Directions

The exponential growth of data and the exponential increase in system complexity as projected for the 2030s bring in new challenges to the communication and computing system. For instance, today a mobile user typically consumes approximately 8 GB data/month on 4G mobile network. Once upgraded to 5G, the data usage is tripled to 25 GB/month. By 2030, majority of mobile users will switch to 5G. It is predicted that a mobile user's data usage will be 160 GB/month or more in 2030, a 20-fold increase, based on screen time/day and usage of high-resolution video applications. Per user data usage growth will propagate capacity requirements throughout the communication network, from backhaul all the way to data centers. Projected higher user data consumption will drive higher energy usage. Mobile user's ever-increasing time-sensitive applications require ultra-low network latency (<5 ms). It is also worth noting that mobile user traffic consists of only part of overall wireless network traffic. Traffic generated by autonomous vehicles and connected devices will likely be several magnitudes larger than human generated mobile traffic.

To meet the challenges from the exponential growth of data, the system needs to be able to handle 10x or even 100x increase in traffic. The communication and computing latency need to be 10x reduced to meet extreme requirements of advanced use cases. The power consumption needs to be 10x to 100x reduced for economic and environmental sustainability. The computing system needs to be able to handle extreme, complex, and diverse workloads and data models that could require fundamental changes in computing platforms across edge and center clouds. In this section, we describe how IOWN GF will accomplish these technical challenges.

2.1. Shift Toward Data-centric Communication and Computing Infrastructure

Driven by the rapid evolution of sensing and AI technologies, data generation rate (e.g., by sensors, applications, content providers, etc.) will always outpace communication network capacity. It is therefore important to design mechanisms on data generation, data processing, data storage, data sharing, data distribution, and data transmission to optimize data efficiency, including data usage efficiency, data processing efficiency, and data movement efficiency. More specifically, network and computing infrastructure can be enhanced with data pipeline infrastructure that reduces workload on underlying network and computing infrastructure with functions such as pub/sub data brokerage and statistical aggregation. In addition, data pipeline infrastructure shall also enrich data, e.g., precise network clock and geolocation stamp can be provided to sensed data across massive number of sensors for precise modeling of the physical world in a digital twin system.

The above will lead to a fundamental shift to data-centric from today's application and Internet-centric, i.e. application specific data and compute silos running over the Internet creating decoupling of data, computing and communication. The network and computing infrastructure can be designed around data, i.e., data plane can be the hub to connect communication planes and computing planes. There may be multiple communication and computing planes, each of which is for a specific type of data sources or computing functions. However, data plane, as the hub, should allow data exchange among different communication and computing planes.

The infrastructure shall be equipped with integrity, security, and privacy mechanisms to ensure trusted data collection, secure data sharing and protected data usage.

A data-centric communication and computing architecture can be established to achieve the features described above. Please refer to section 3.1.1 for specific work items on this direction.

2.2. Full-stack Communication Acceleration

Despite the great boost in data link capacity driven by advancements in optical communication and radio access technologies, computing and application systems cannot fully exploit this capacity boost due to several factors that are inherent in today's common practice in networking.

One factor is the use of TCP/IP, whose speed is severely slowed as communication round trip time (RTT) increases. One snapshot by a high-density sensor will generate a block of hundreds of kB data and sending it through TCP/IP requires several round trips of communication, resulting in several tens of msec delay.

Another factor is the dependency on the evolution of packet processing technologies. Network virtualization is inevitable for flexible and speedy infrastructure operation and network service insertion, i.e. inserting gateways such as firewalls, proxies, load balancers, and tunneling gateways, is also inevitable for security and performance management. As network virtualization and network service insertion are often implemented today with software, processing the projected 10x-100x increase in communication throughput would require new technology innovations.

To be more specific, as communication throughput and latency bottleneck shifting from L2 to L3 and L4, we need to define a full-stack functional and implementation architecture that achieves high-speed communication across layers.

We expect that a high-speed communication stack envisioned in IOWN GF will empower many applications with instant massive data exchange, which would eventually create more demand on optical transport and wireless access networks. Anticipating such demands, IOWN GF will put effort into developing optical and wireless technologies that achieve much higher capacity and speed.

Please refer to section 3.1.2 for more specific work items.

2.3. Computing Scaling Across Device, Network, Edge and Center Cloud

With the diminishing gains in computing performance by increasing operation frequency, computing scaling is largely relay on distribute computing workload across different computing units. The complexity of computing scaling is ever increasing with the emergence of new devices, e.g. accelerators and persistent memories, new use cases, new data patterns, new traffic types, new algorithms, and new performance requirements.

In particular, the complexity of computing scaling imposes technical challenge in data-intensive computing applications, such as deep learning. Data-intensive computing applications cannot be practically implemented without accelerators, such as GPGPU, TPU, and VPU, and the performance benefit from an accelerator depends heavily on how fast the application can transfer data to the accelerator. On the other hand, if computing workload is distributed across different hosts, the performance can be easily sacrificed by inefficient packet processing at layer 3 and layer 4 (see Section 2.2).

There are existing solutions in computing industry, such as RDMA, InfiniBand, and Converged Ethernet. However, these technologies are mainly for data transfer within one data center.

That said, we should define a computing scaling platform that goes beyond inter-data centers and expands across device, network, edge, and center cloud. The goal of computing scaling is to achieve dynamic linear increase on computing capability with respect to the number of computing units and in the meantime minimize power consumption. This requires joint design effort on computing platform, inter-connect communication links and networks and software platform, and should leverage the outcome of Full-Stack Communication Acceleration (see 2.2).

Please refer to section 3.1.3 for more specific work items.

2.4. Sustainable Growth with Energy Efficiency

For an economical and technological sustainable growth and considering society and environmental responsibilities, energy efficiency should be considered in both component level and system level.

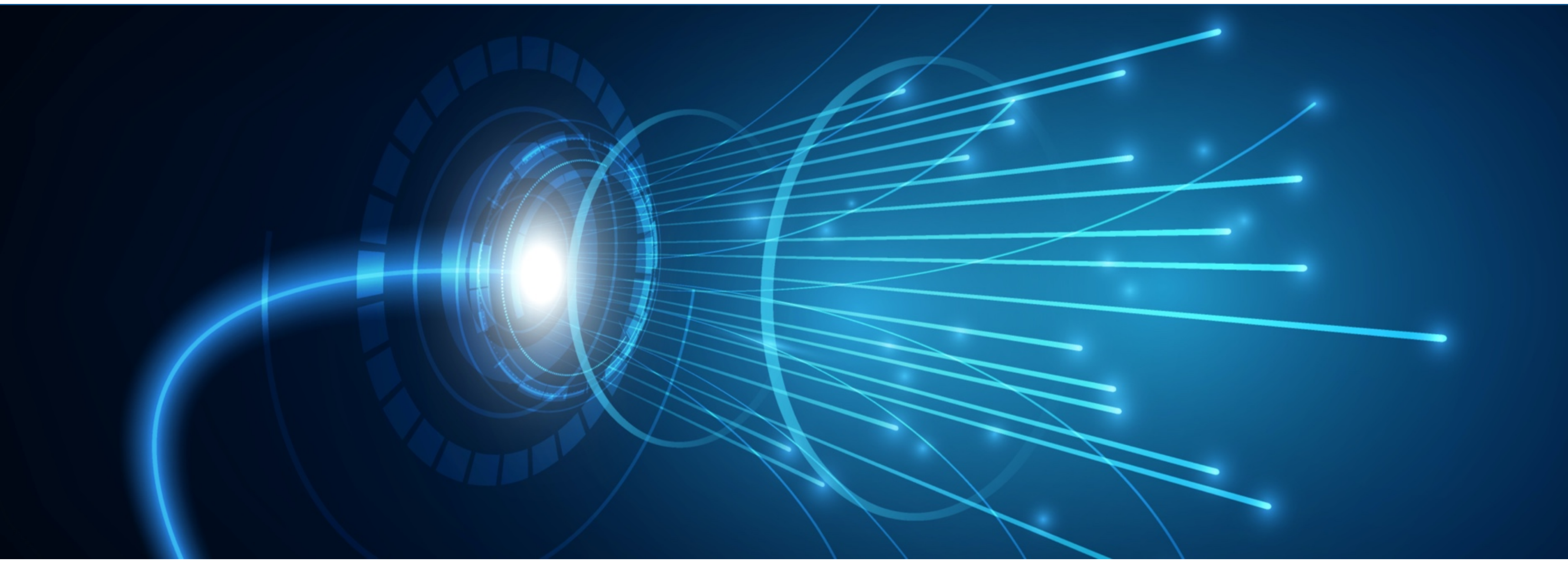
In the component level, factors that could contribute to energy efficiency include: advanced DSP schemes, continuing enhancement on processor design, further shrinking on process nodes, silicon photonics with advanced packaging, energy conscious air interface design, energy conscious access network and transport network design, etc.

However, component level enhancement is not enough. A fundamental architectural work on computing and communication system infrastructure leveraging component-level advancements is needed to achieve a quantum leap in energy efficiency.

One aspect of the architectural work is to enable a paradigm shift from constant rate to event driven, which is emerging in the sensing industry. Many of today's components are designed based on a constant rate paradigm. For example, IP cameras, which are used as data source nodes in many application systems, capture and generate data at constant frame rates, image sizes, and image compression rates. Users are struggled with a choice of these parameters, because those parameters greatly impact the network bandwidth and storage consumption. As a solution for this, the sensing industry has come up with a new paradigm, event driven. In this paradigm, there is no notion of frame rate or image size. An image, which is just a partial region of the camera frame, is captured upon the occurrence of an event. This new paradigm will allow us to reduce the network and storage resource consumption while achieving higher cognitive capacity and response speed. However, this cannot be achieved with static network bandwidth allocation and computing resource allocation. We should define network and computing infrastructure with new service models that well support this event-driven paradigm.

The computing and communication system infrastructure should support proper scheduling and decisions on when and how to generate the data, where and how to move the data, where to do the computing, where to store data, when to allocate/deallocate/upscale/downscale computing resource to an application, etc.. Overall, energy efficiency should be considered across communication, data and computing in both component level and system level.

Please refer to section 3.1.4 for more specific work items.



3. IOWN GF Technology Work Plan

3.1. Work Items (Tentative)

The work items listed in this document are just tentative. Any member can propose any work item that is aligned with IOWN GF's vision and technology directions presented in this whitepaper.

3.1.1. Data-centric Communication and Computing

- A functional architecture for Data-Centric Communication and Computing infrastructure that is designed around data, optimized for data efficiency, enriched for data and equipped with security, privacy and data integrity features. The infrastructure should provide the following functions:
    - event/data brokerage, which should accommodate very large data objects, very strict latency requirements, and a very large number of subscribers and publishers
    - controlling publisher's data generation and injection based on overall system condition and the importance of the data
    - processing, analyzing, and reducing data at network edge instead of sending all the data to the center cloud
    - storing and sharing data among communication and computing nodes in the network
    - precision time and/or geolocation stamping
    - assuring data security, privacy, and integrity
    - data life cycle management
    - data brokerage between IP and non-IP nodes
    - data brokerage accessible through multiple networks and/or networks service providers
- Protocols, interfaces, and data formats for the defined architecture

3.1.2. Full-Stack Communication Acceleration

- A functional and implementation architecture for fast and efficient communication that can fully exploit the speed of layers 2 and 1. This should include the following subitems:
  - a data transfer protocol optimized for the instant transfer of large data objects, which should significantly reduce the negative effect of distance and leverage ultra-broadband data link
  - a communication stack for high-speed and low-latency communication between strongly interdependent two nodes, e.g. an IoT controller and an IoT device, a main CPU node and a remote GPU node. (In such cases, layer 3 could be null or replaced with a label switching layer.)
  - efficient L3 and L4 processing that can address network virtualization and service insertion demands
  - security, privacy, and integrity protection for the above
  - fast end point authentication
  - implementation architecture leveraging Silicon Photonics technologies
- A functional and implementation architecture for optical transport and radio access that can accommodate demands for much higher capacity and speed
- Interfaces and protocols for the above

3.1.3. Computing Scaling Across Device, Edge, and Center Cloud

- A functional and implementation architecture for computing scaling platform that goes beyond inter data centers and expands into across device, network, edge and center cloud. Targeted at dynamic linear computing scaling for various types of data and computing workloads. This should include the following subitems:
  - architecture for computing scaling platform for data-intensive computing requiring accelerator
  - application design patterns of applications that run on the above platform
  - method of creating a pool of data-intensive computing resources such as accelerators and persistent memories and sharing it across data centers
  - integration with the outcome of Full-Stack Communication Acceleration
- Interfaces and protocols for the above

3.1.4. Sustainable Growth with Energy Efficiency

- A novel communication and computing architecture for event-driven computing and networking, where energy consumption is roughly a linear function of the occurrence rate of an event, e.g. a human appearing in the monitored area, such as:
  - end-to-end network service model and architecture for event-driven communication among data source nodes and data consumers
  - computing service model and architecture for event-driven data-intensive computing
  - dynamic orchestration and resource scheduling mechanism supporting the above
  - interface specifications for the above event-driven communication and computing
- A novel computing architecture for paradigm shift from traditional server-box-oriented computing to box-less disaggregated computing on silicon photonics-based data plane. This should include the following items:
  - data plane architecture supporting a paradigm shift from NIC-to-NIC to module-to-module, i.e. direct data transfer between modules without transiting NICs
  - orchestration and operation systems for box-less disaggregated computing, where modules are dynamically reconfigured and assembled to provide computing functions

  - security and performance management mechanisms for the above
  - interface specifications for the defined architecture

3.2. Operation Principles

- IOWN GF should respect other communities' activities. The main focus of IOWN GF should be to define an end-to-end and full-stack system architecture that leads to quantum leaps in our key dimensions. Specifying technical details for a specific layer or component should be carried out by subject matter expert groups outside of this forum. However, if there is no such a group, we will work on technical details.
- IOWN GF should also run use case exploration to fully understand user's requirements and to raise people's awareness about IOWN GF's potential impact on future societies and businesses.

4. IOWN GF Roadmap

4.1. Long-Term Roadmap

As shown in Table 7, IOWN GF will run activities in phases, each of which is 18 months long. Activities after the completion of phase 3 will be discussed by December 2024.

Table 7: IOWN GF Long-Term Roadmap

	MONTH OF COMPLETION	DELIVERABLES
Phase 1	December 2021	Prospective use cases for early adoption and their KPIs, Release 1 Functional and deployment architectures, Release 1
Phase 2	June 2023	A series of technical specifications for Release 1 Prospective use cases for early adoption and their KPIs, Release 2 Functional and deployment architectures, Release 2
Phase 3	December 2024	Activities necessary for the successful roll out of IOWN GF Release 1 should be identified and conducted A series of technical specifications for Release 2

4.2. Short-Term Roadmap for Phase 1

IOWN GF will publish Call for Proposals with this whitepaper. We expect members to propose use cases and technologies that will help to achieve IOWN GF vision. Hopefully, with collective efforts of members, IOWN GF will start use case and technology development work in June 2020. In the initial phase, the two working groups, Use Case Working Group and Technology Working Group, will run together to make initial progress.

Table 8: Short-Term Roadmap for Phase 1

MONTH	MILESTONES
April 2020	Release of IOWN GF Whitepaper (this document) and Call for Proposals The submission deadline for proposals will be the end of June. Kick-off meeting including Whitepaper briefing session (will be an online meeting)
July 2020	Start use case and technology development work based on received proposals
December 2020	Initial work on prospective use cases for early adoption (This should facilitate architectural work by Technology WG.)
September 2021	Completion of working group draft of the Phase 1 deliverables
December 2021	Release of Phase 1 deliverables

5. Summary

IOWN GF targets to develop a next generation communication and computing infrastructure, which excels in:

- communication capacity/latency
- data efficiency
- computing scalability
- energy efficiency

With IOWN GF, smart world applications will be created and augmented with:

- beyond-human cognitive-communication capacity
- beyond-human response speed
- dynamic linear computing scalability
- much superior energy efficiency

The accomplishment of the IOWN GF vision requires global collaboration to develop and promote a new full-stack communication and computing architecture. IOWN GF will start its global collaboration from the following four areas:

- Shift toward Data-centric Communication and Computing Infrastructure
- Full-Stack Communication Acceleration
- Computing Scaling Across Device, Edge, and Center Cloud
- Sustainable Growth with Energy Efficiency

Members are encouraged to propose any ideas about use cases and technologies that contribute to the fulfillment of the forum’s vision.

Appendix 1: Types of Smart and Connected Applications

Cognitive Systems

A cognitive system is defined as an application system that consists of data source nodes, cognitive functions, and post-cognition functions as defined below:

Data source node: A connected node that has some sensing functions to generate data about monitored something (things, humans, and areas) and send generated data to cognitive functions

Cognitive function: A function of analyzing data received from data source nodes to produce intelligence about the monitored.

Post-cognition function: A function of processing intelligence data produced by cognitive function to achieve application objectives

Autonomous Systems

Autonomous system is defined as an application system that consists of data source nodes, cognitive functions, control functions, and controlled objects as defined below:

Data source node: same as that in Cognitive Systems

Controlled node: a connected node that receives commands from Control functions and executes the received commands. There may be a node that functions as both data source node and controlled node.

Cognitive function: same as that in Cognitive Systems

Post-cognition function: A function of processing intelligence data produced by cognitive function, generating commands, and sending the generated commands to controlled nodes to achieve application objectives

XR System

XR System is defined as application system that consists of input device, data generation function, and output device as defined below:

Input device: A device that captures motion images of the reality and sends them to data generation functions. Images may be accompanied by some additional data such as thermograph, geolocation and time.

Data generation function: A function of receiving data from input devices, generating response data, and sending the response data to output devices

Output device: A device that receives response data from data generation functions, constructs images to be displayed, and display the constructed images

Systems for Prediction, Simulation, and Pro-active Control

System for Prediction, Simulation, and Pro-active Control are defined as systems that consists of the following functions.

Data collection function: a function of collecting data from data source nodes or cognitive systems as defined above

Data organization function: a function of organizing collected data for prediction, simulation, or pro-active control

Prediction, simulation, and pro-active control function: a function of achieving application objectives through analytical computing using organized data. Application objectives may include prediction, simulation, and pro-active control.

We expect that the emerging computing paradigm with digital twins will advance systems of this type.