WNIG Use Case study Ubiquitous computing

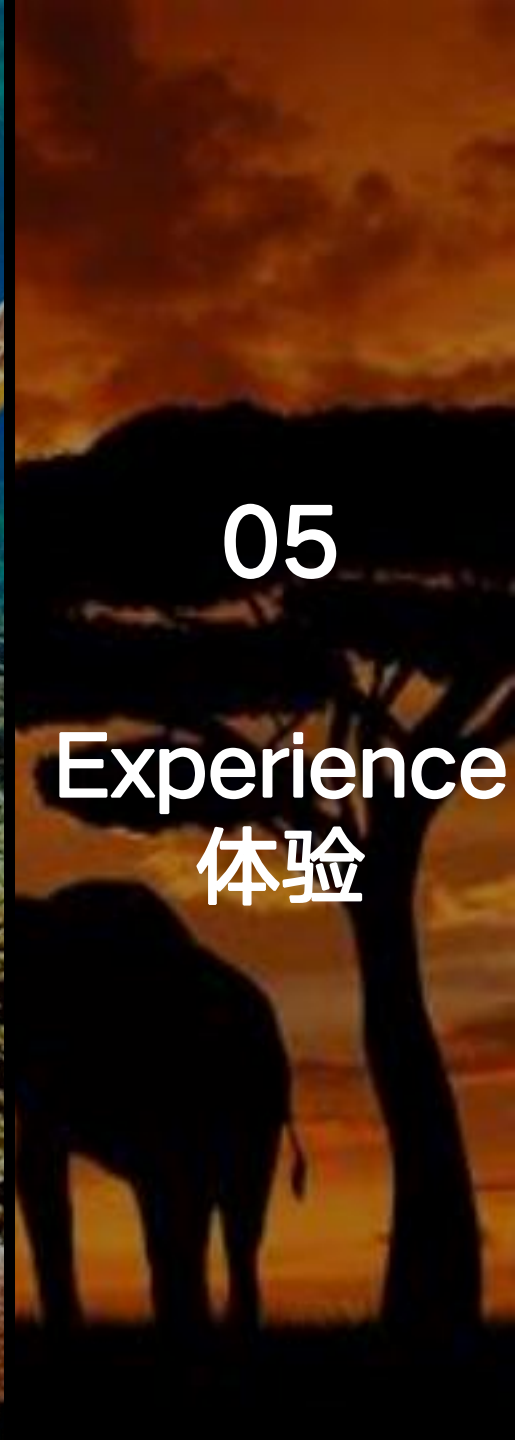
TPAC 2023 WNIG


01
Scenarios
场景


02
Improve
ments
改进


03
Audio
visual
音视频


04
Immersive
沉浸式Web


05
Experience
体验

USE CASE 1: SUL

- WebRTC is not cure-all for any cases. Uplink streaming: Developers use the traditional RTMP method.
- HLS as the Downlink streaming, Can WebTransport solve challenges?
- Unable to transparently transmit H.265 video, leading to use WebSocket and to implement varied proprietary methods
- WEAK NETWORK: The adaptive mechanism does not have developer-friendly flexibility; Not all scenarios have the priority for lower latency, some need high-quality video

USE CASE 2: COMPL.TASK -STREAMING

- Complex tasks such as ChatGPT and AIGC, the user waiting time can reach 50–150 seconds
- The waiting time is too long. There is a streaming mechanism (40% of the generation is rendered first, and the remaining 60% can be expected)
- Example: flexible application of streaming mechanism brings high experience
- Acquisition. No intention to introduce into W3C and community. The mechanism could be helpful for web developers. (existing web spec. Pipeline?)

USE CASE 3: COMPL.TASK DISTRIBUTION – Intro.

WebGL/WebGPU distributed rendering is a distributed 3D rendering technology, similar to grid computing, which allows multiple computers or devices to work together to complete complex 3D rendering tasks.

- each computer or device can become a node.
- Data between nodes can be carried out through a central node
- Ex.: Unity rendering in cloud: splits the rendering task and schedules it to multiple GPU devices, pushes the rendered image to the user through stream.

Distributed WebGL/WebGPU: Applied to scenarios requiring high-performance graphics rendering: virtual reality, game development, scientific visualization, architectural design.

- By distributing rendering tasks to multiple computers or devices
- making full use of GPU resources, accelerating the speed of graphics rendering
- improving the quality of rendering, to reduce the load on a single computer or device,
- improving the reliability and stability of the system

USE CASE 3: COMPL.TASK DISTRIBUTION - Facts

WebGL/WebGPU/WebRTC/WebTransport

- ☐ Limited terminal hardware resources, challenges to run large 3D applications on Web
- ☐ Runtime performance is poor
- ☐ Common practice: 3D applications on remote servers/cloud through virtual machines or application containers
- ☐ Transmit video streams, present them to the front end through WebRTC/WebTransport and other technologies

Cost of 3D application cloudization is high

- ☐ Relies on heavy components such as WebRTC, which makes network deployment more complex suitable for ultra-large applications.
- ☐ For small and medium-sized applications, the ROI is relatively low.

USE CASE 3: COMPL.TASK DISTRIBUTION – GAP ANALYSIS

node discovery, registration and task scheduling

- Each node registers the current machine GPU hardware information, load status, etc. to the central node.
- Each node can initiate tasks and then be scheduled by the central node.
- Note: GPU hardware information acquisition involves privacy information, which is not supported by the current browser.

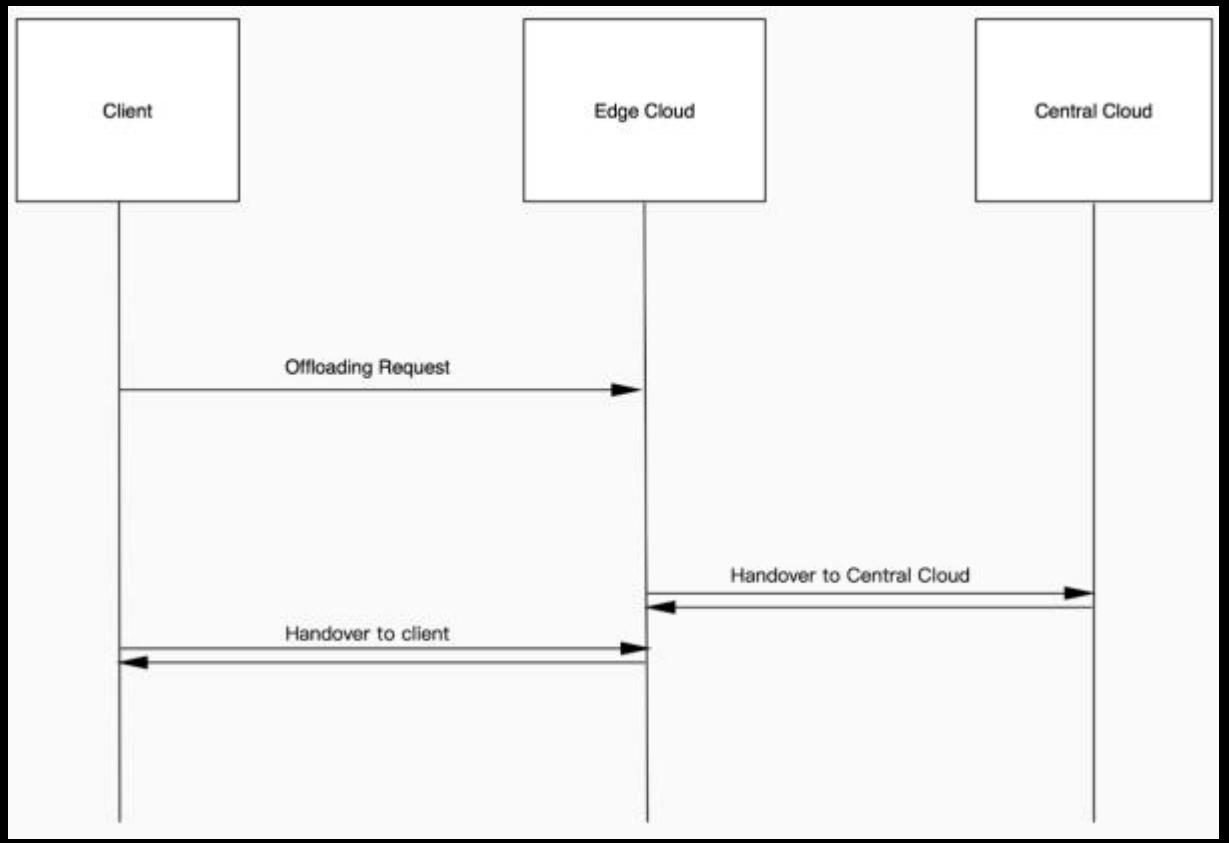
To split WebGL/WebGPU rendering tasks

- he GPU pipelines supported by WebGL/WebGPU include Render Pipeline and Compute Pipeline.
- Only native GPU can be applied. Consider supporting Distributed Render Pipeline and Distributed Compute Pipeline, scheduling expensive rendering and computation to other machines for collaborative rendering and computation,
- merging the rendering results and rendering them to the web browser.

USE CASE 4: Ubiquitous Computing

Clould–edge–mobile coordination explainer
Ubiquitous computing

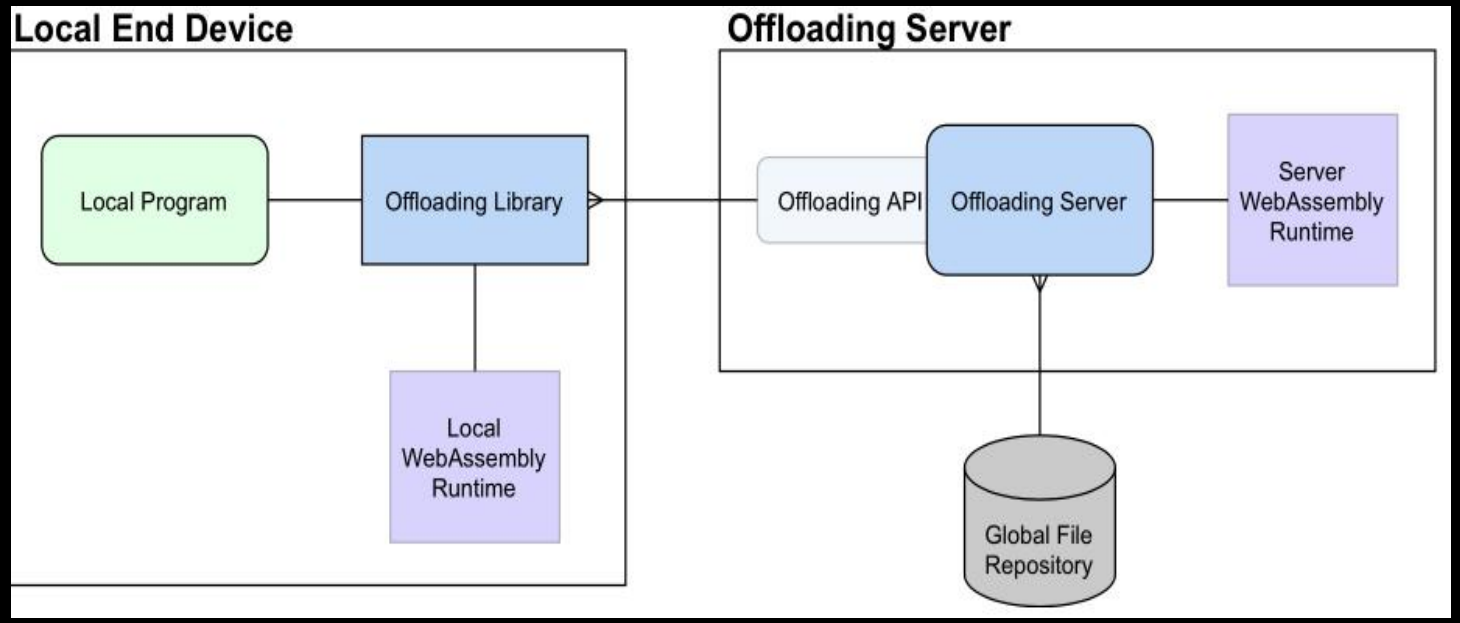
- Large internet enterprises are designing proprietary solution for in–house micro–service platform; WASM/JS/Deno computing models have lower awareness.
- Proprietary product will be used for the runtime of its own business deployment; benchmark: CF and other products
  - Client should be able to offload computing intensive work to the edge cloud
  - The edge cloud should be able to handover computing intensive work back to the client



USE CASE 4: Ubiquitous Computing

Offloading

- ❑ Resource limited for intensive computing
- ❑ Battery consumption and application performance
- ❑ To solve the heat dissipation issues
- ❑ 5G-Advanced and edge computing



Challenges

- ☐ discover and locate the Edge Node to get its IP addr.
- ☐ hardwares with different architectures
- ☐ Node cannot store all information
- ☐ Data and State synchronization

Solutions

WebAssembly: WasmEdge from CNCF and wamser can provide a high-performance WASM runtime

USE CASE 5: Net Info

❑ **Example: L4S**

❑ 3GPP, IETF, W3C
Developers need a practical and end-to-end solution, rather than technologies implemented in different standard bodies

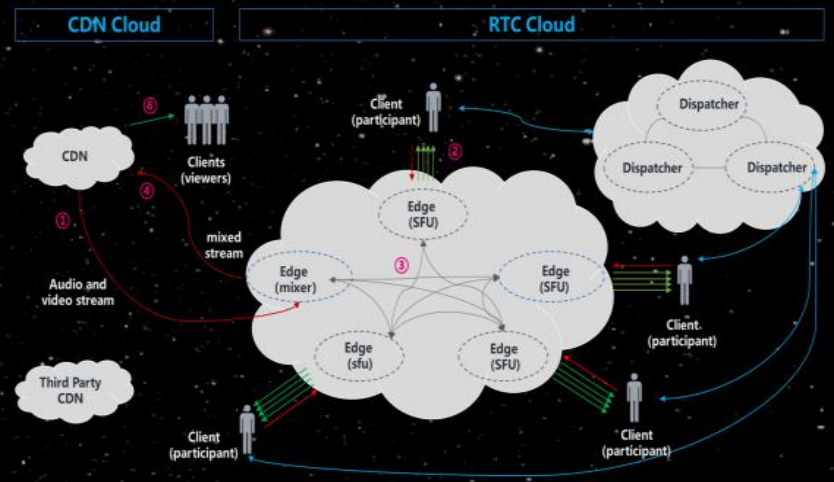
CORSS-LAYER research and development

❑ W3C is the place to expose developer-friendly interface to dig out the NETWORK functions (UL/DL, signal strength in data field, congestion statics data field)

USE CASE 6: Immersive Web

- ❑ the livestreaming multimedia ecosystem is experiencing robust growth, needs fulfill a broad range of user demands
- ❑ Cloud Box: Customized on-line private media room
  - ➢ The live and on-demand media streaming (sports events, movies, TV, etc.) are pushed from CDN via RTMP to the RTC edge node (mixer) which distributes them to the RTC cloud
  - ➢ The RTC edge nodes (mixer) mix the live and on-demand media streaming with real-time media streaming pushed by the participants and forwards to the CDN via RTMP
  - ➢ the cloud box pull the mixed media streaming via RTMP, FLV, WebRTC, etc., from CDN
  - ➢ Join Breakout session for other use cases: Metaverse Convention Center(Migu), Livestreaming e-commerce, fundamental livestreaming cloud services(Alibaba), Educational livestreaming, livestreaming e-commerce, next-gen Conferencing(Bytedance)

Real-time interactive Cloud Box for sports events, concert, movies, TV, etc.



RTC and CDN cloud for real-time interactive Cloud Box

New use cases and tasks => Rechartering

Charter proposal draft

Changes:

- Updated mission statement to include exploration of solutions to improve network resource allocation and efficiency, and leveraging of new technologies like edge computing
- Added Motivation and Background section
- Under Scope section, included Client-Edge-Cloud Coordination
- Added two study tasks around trust model and network resource utiliztation in Scope section


Ubiquitous computing,
Serving Web with advanced network
Call for action: Proposal for Web and Network IG https://www.w3.org/Telco