CLAIMS

1. A computer-implemented method for implementing online video games or applications over a network including at least one public network component comprising:

generating control signals from a client in response to an input device of the client being actuated by a user as the user is playing a low latency video game or application;

receiving the control signals related to the low latency video game or application transmitted from the client operated by the user over the network which includes at least one public network component;

routing the control signals to a server and responsively executing program code of the low latency video game or application to generate a plurality of video frames resulting from execution of the video game or application;

encoding the plurality of video frames or portions thereof according to a first encoding format to generate a plurality of encoded video frames or portions, the first encoding format being optimized for low latency transmission to a client device over a current network communication channel including the at least one public network component;

transmitting the plurality of encoded video frames or portions to the client device over the network communication channel,

the client device decoding the video frames and rendering the video on a display,

wherein the operations of generating control signals, receiving the control signals, routing the control signals to the server, responsively executing the program code of the low latency video game or application to generate the plurality of video frames, encoding the plurality of video frames according to the first encoding format optimized for low latency transmission to generate the plurality of encoded video frames, transmitting the plurality of encoded video frames, decoding the video frames on the client and rendering the video on the display are performed with a latency such that the user has the perception that the video game or application is responding instantly to the user input;

concurrently encoding the first plurality of video frames according to a second encoding format, the second encoding format having a relatively higher-quality compressed video and/or a lower compression ratio than the first encoding format;

storing the first plurality of video frames encoded in the second encoding format on a storage device;

making the first plurality of video frames encoded in the second encoding format available to the client device and other client devices for playback of the video stream; and

responsively transmitting the first plurality of video frames encoded in the second encoding format to the client and/or other clients upon receiving a request from a user.

2. The method as in claim 1 wherein the first encoding format comprises frames which are dependent on other frames and the second encoding format comprises some frames which are not dependent on other frames.

3. The method as in claim 2 wherein the second encoding format includes I-Frames and wherein the first encoding format does not include I-Frames.

4. The method as in claim 3 wherein the first encoding format comprises only P-Frames.

5. The method as in claim 1 wherein the first plurality of video frames are frames of a video game, generated as a user of the client device plays an online video game and wherein the first encoding format is selected to reduce latency as the frames compressed according to the first encoding format are transmitted to the client device over the current communication channel.

6. The method as in claim 5 wherein the second encoding format is selected to archive the video frames of the online video game, the video frames to be made available for subsequent playback from the archive.

7. A computer system having a memory for storing program code and a processor for processing the program code to perform the operations of:

generating control signals from a client in response to an input device of the client being actuated by a user as the user is playing a low latency video game or application;

receiving the control signals related to the low latency video game or application transmitted from the client operated by the user over the network which includes at least one public network component;

routing the control signals to a server and responsively executing program code of the low latency video game or application to generate a plurality of video frames resulting from execution of the video game or application;

encoding the plurality of video frames or portions thereof according to a first encoding format to generate a plurality of encoded video frames or portions, the first encoding format being optimized for low latency transmission to a client device over a current network communication channel including the at least one public network component;

transmitting the plurality of encoded video frames or portions to the client device over the network communication channel,

the client device decoding the video frames and rendering the video on a display,

wherein the operations of generating control signals, receiving the control signals, routing the control signals to the server, responsively executing the program code of the low latency video game or application to generate the plurality of video frames, encoding the plurality of video frames according to the first encoding format optimized for low latency transmission to generate the plurality of encoded video frames, transmitting the plurality of encoded video frames, decoding the video frames on the client and rendering the video on the display are performed with a latency such that the user has the perception that the video game or application is responding instantly to the user input;

concurrently encoding the first plurality of video frames according to a second encoding format, the second encoding format having a relatively higher-quality compressed video and/or a lower compression ratio than the first encoding format;

storing the first plurality of video frames encoded in the second encoding format on a storage device;

making the first plurality of video frames encoded in the second encoding format available to the client device and other client devices for playback of the video stream; and

responsively transmitting the first plurality of video frames encoded in the second encoding format to the client and/or other clients upon receiving a request from a user.

8. The computer system as in claim 7 wherein the first encoding format comprises frames which are dependent on other frames and the second encoding format comprises some frames which are not dependent on other frames.

9. The computer system as in claim 8 wherein the second encoding format includes I-Frames and wherein the first encoding format does not include I-Frames.

10. The computer system as in claim 9 wherein the first encoding format comprises only P-Frames.

11. The computer system as in claim 7 wherein the first plurality of video frames are frames of a video game, generated as a user of the client device plays an online video game and wherein the first encoding format is selected to reduce latency as the frames compressed according to the first encoding format are transmitted to the client device over the current communication channel.

12. The computer system as in claim 11 wherein the second encoding format is selected to archive the video frames of the online video game, the video frames to be made available for subsequent playback from the archive.

13. A non-transitory machine-readable medium having program code stored thereon which, when executed by a machine, causes the machine to perform the operations of:

generating control signals from a client in response to an input device of the client being actuated by a user as the user is playing a low latency video game or application;

receiving the control signals related to the low latency video game or application transmitted from the client operated by the user over the network which includes at least one public network component;

routing the control signals to a server and responsively executing program code of the low latency video game or application to generate a plurality of video frames resulting from execution of the video game or application;

encoding the plurality of video frames or portions thereof according to a first encoding format to generate a plurality of encoded video frames or portions, the first encoding format being optimized for low latency transmission to a client device over a current network communication channel including the at least one public network component;

transmitting the plurality of encoded video frames or portions to the client device over the network communication channel,

the client device decoding the video frames and rendering the video on a display,

wherein the operations of generating control signals, receiving the control signals, routing the control signals to the server, responsively executing the program code of the low latency video game or application to generate the plurality of video frames, encoding the plurality of video frames according to the first encoding format optimized for low latency transmission to generate the plurality of encoded video frames, transmitting the plurality of encoded video frames, decoding the video frames on the client and rendering the video on the display are performed with a latency such that the user has the perception that the video game or application is responding instantly to the user input;

concurrently encoding the first plurality of video frames according to a second encoding format, the second encoding format having a relatively higher-quality compressed video and/or a lower compression ratio than the first encoding format;

storing the first plurality of video frames encoded in the second encoding format on a storage device;

making the first plurality of video frames encoded in the second encoding format available to the client device and other client devices for playback of the video stream; and

responsively transmitting the first plurality of video frames encoded in the second encoding format to the client and/or other clients upon receiving a request from a user.

14. The machine-readable medium as in claim 13 wherein the first encoding format comprises frames which are dependent on other frames and the second encoding format comprises some frames which are not dependent on other frames.

15. The machine-readable medium as in claim 14 wherein the second encoding format includes I-Frames and wherein the first encoding format does not include I-Frames.

16. The machine-readable medium as in claim 15 wherein the first encoding format comprises only P-Frames.

17. The machine-readable medium as in claim 13 wherein the first plurality of video frames are frames of a video game, generated as a user of the client device plays an online video game and wherein the first encoding format is selected to reduce latency as the frames compressed according to the first encoding format are transmitted to the client device over the current communication channel.

18. The machine-readable medium as in claim 17 wherein the second encoding format is selected to archive the video frames of the online video game, the video frames to be made available for subsequent playback from the archive.

19. The method as in claim 1 wherein the video frames are encoded and transmitted in response to the control signals with a finite latency of 80 ms or less.

20. The method as in claim 1 wherein the video frames are encoded and transmitted in response to the control signals with a finite latency of 150 ms or less.

21. The system as in claim 7 wherein the video frames are encoded and transmitted in response to the control signals with a finite latency of 80 ms or less.

22. The system as in claim 7 wherein the video frames are encoded and transmitted in response to the control signals with a finite latency of 150 ms or less.

23. The machine-readable medium as in claim 13 wherein the video frames are encoded and transmitted in response to the control signals with a finite latency of 80 ms or less.

24. The machine-readable medium as in claim 13 wherein the video frames are encoded and transmitted in response to the control signals with a finite latency of 150 ms or less.