



Alex Eleftheriadis, Ph.D.
Chief Scientist & Co-founder

Recent Advances in Multipoint Video Server Architectures: Scalability, Simulcasting, and SFUs



Alex Eleftheriadis, Ph.D.
Chief Scientist & Co-founder

history of multipoint video



Alex Eleftheriadis, Ph.D.
Chief Scientist & Co-founder



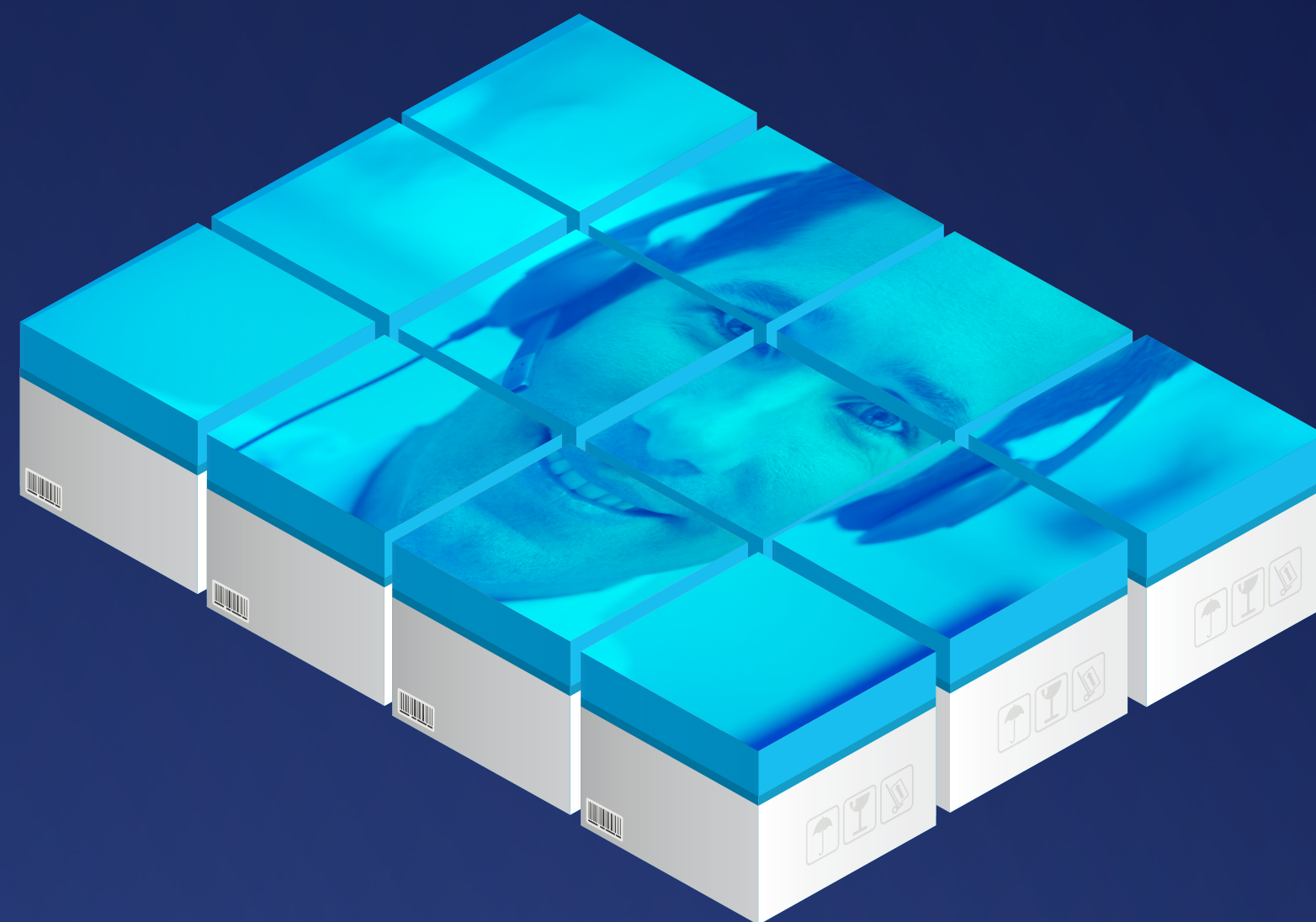
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

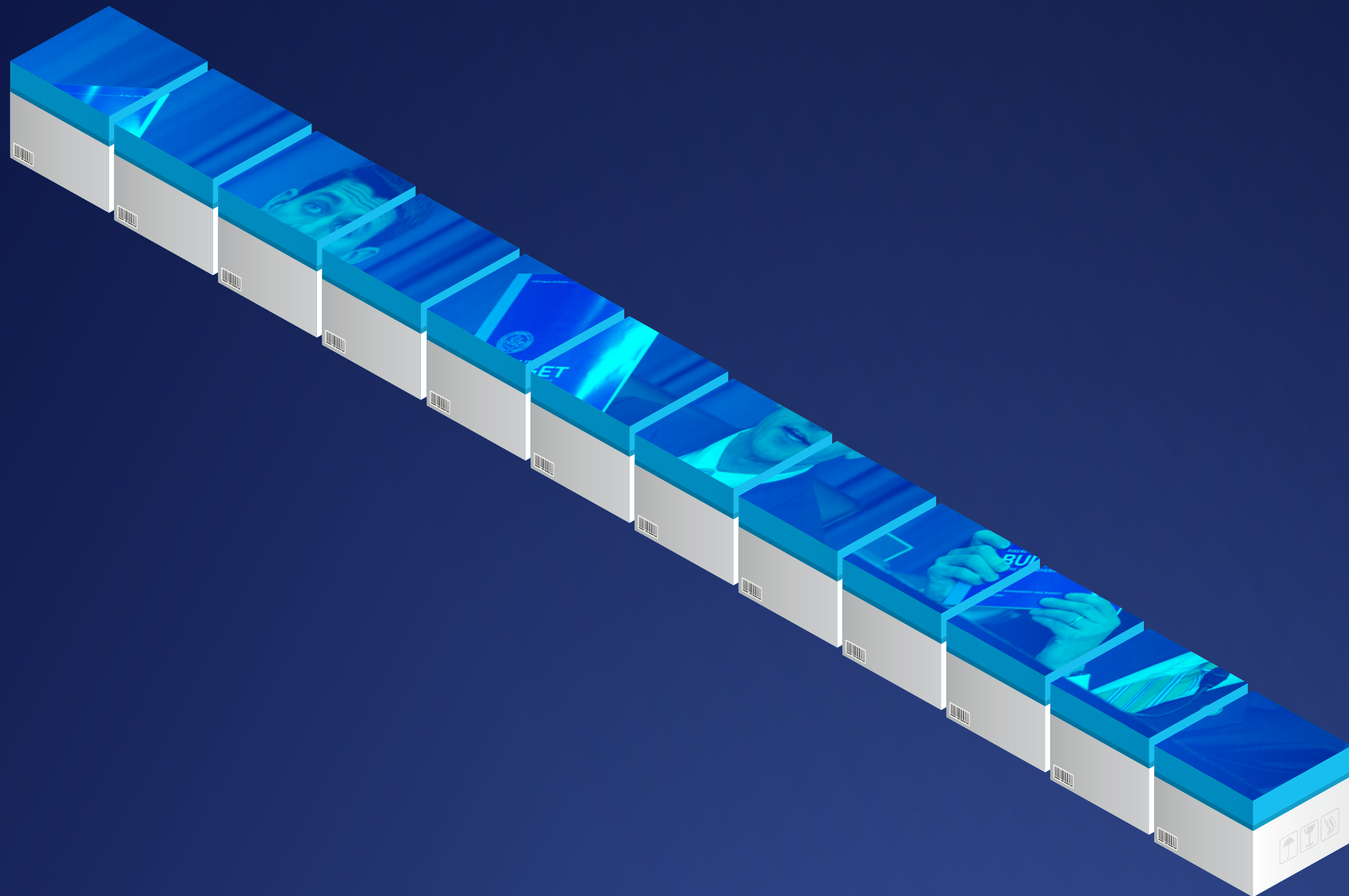


Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder









Vidyo®

Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



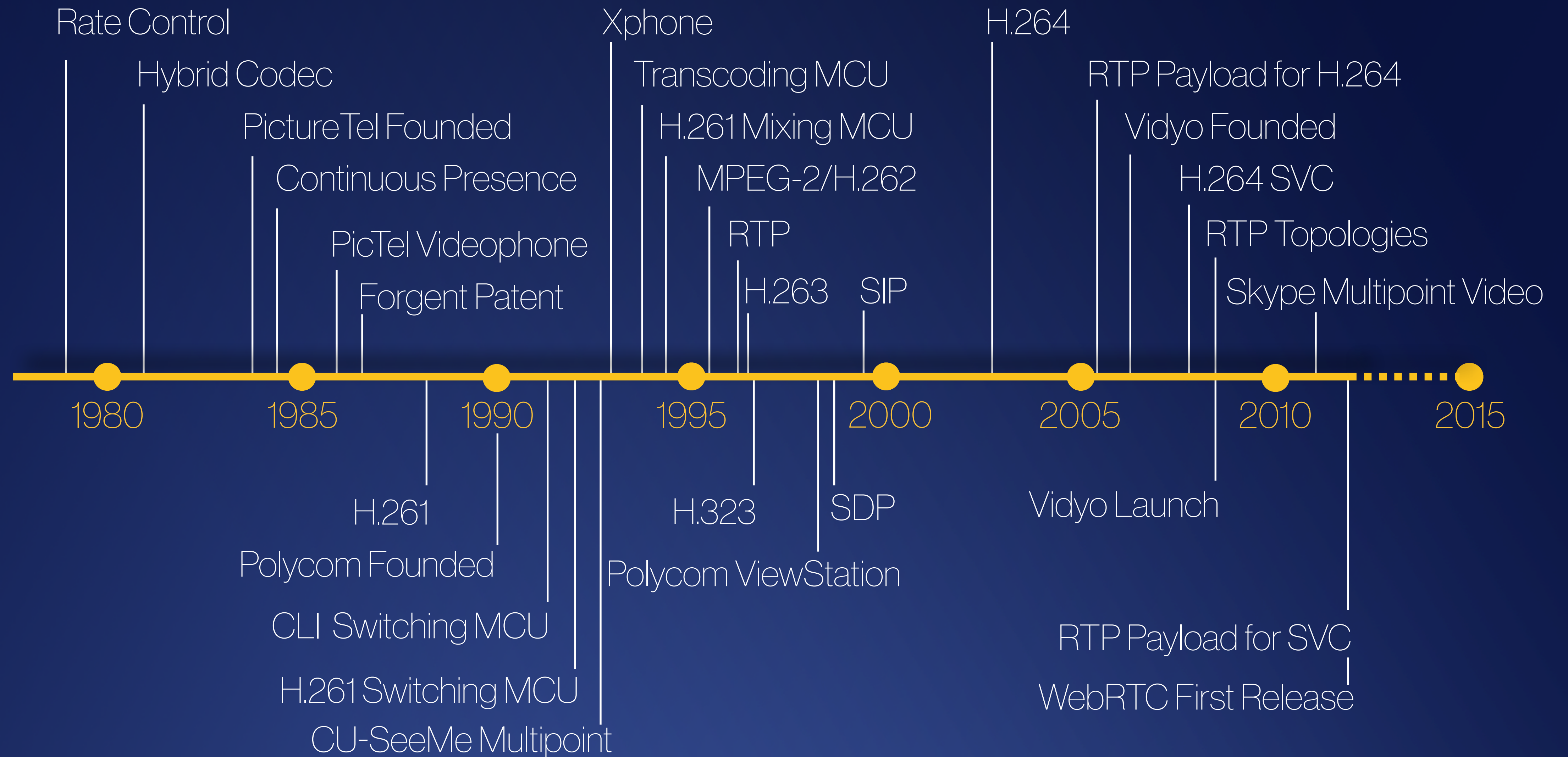
timeline



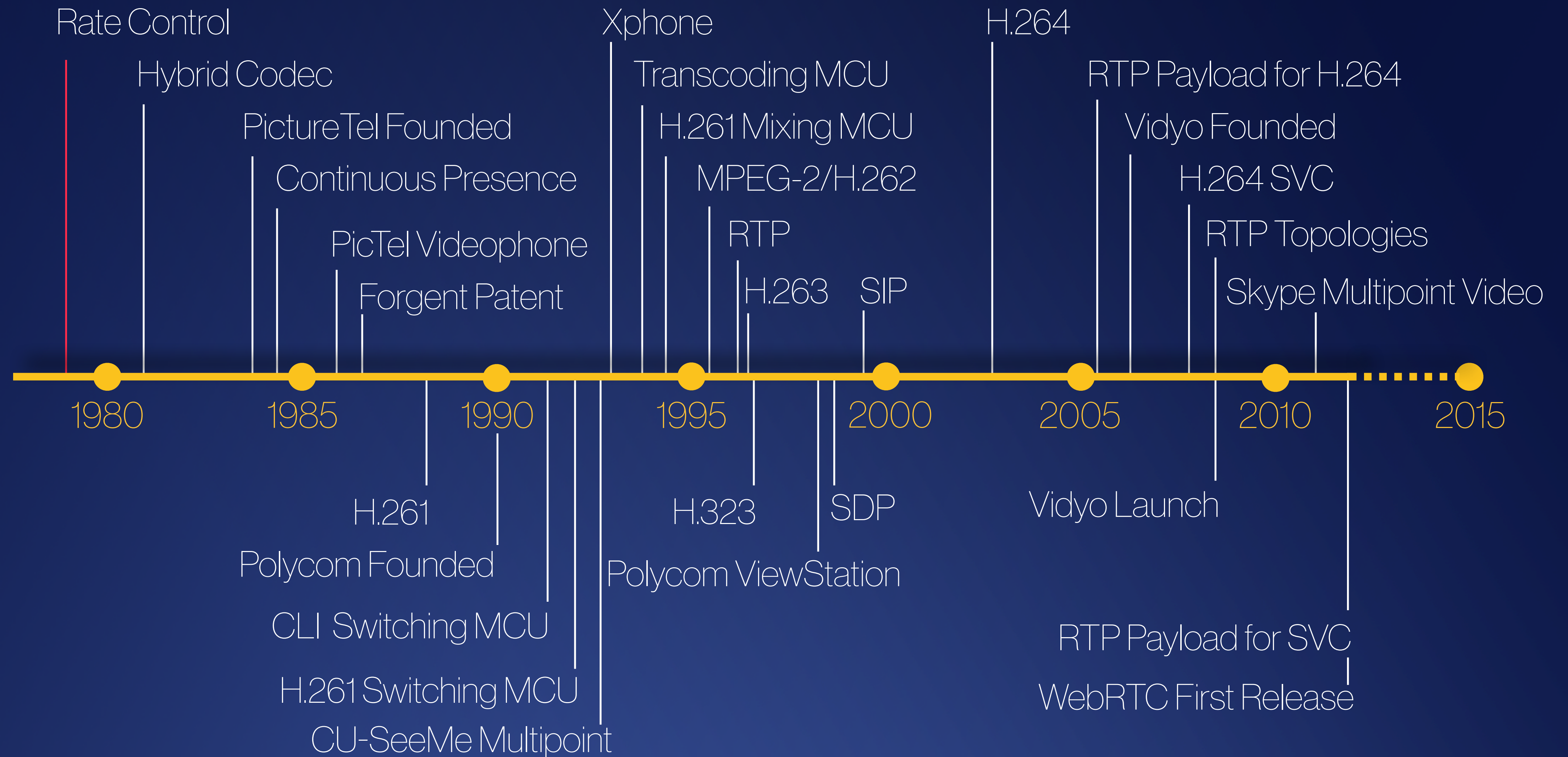
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Rate Control

December 1978
US Pat. Nr. 4,302,775
Compression Labs, Inc.

First CLI video patent,
rate control through
adaptive quantization

United States Patent [19] [11] 4,302,775
Widergren et al. [45] Nov. 24, 1981

[54] DIGITAL VIDEO COMPRESSION SYSTEM AND METHODS UTILIZING SCENE ADAPTIVE CODING WITH RATE BUFFER FEEDBACK

[75] Inventors: Robert D. Widergren, Saratoga; Wen-Hsiung Chen, Sunnyvale; Stanley C. Fralick, Saratoga; Andrew G. Tescher, Claremont, all of Calif.

[73] Assignee: Compression Labs, Inc., San Jose, Calif.

[21] Appl. No.: 969,991

[22] Filed: Dec. 15, 1978

[51] Int. Cl.³ H04N 7/12; H04N 9/32; G06F 15/20; G08C 9/00

[52] U.S. Cl. 358/136; 358/133; 340/347 DD; 364/514; 364/515; 364/582

[58] Field of Search 364/514, 515, 576, 582; 358/12, 13, 133, 138, 260, 261; 340/347 DD

References Cited

3,795,763 3/1974 Golding et al. 358/135 X
3,984,626 10/1976 Mounts et al. 358/135
4,005,411 1/1977 Morris 340/347 DD
4,047,221 9/1977 Yasuda et al. 358/136
4,051,530 9/1977 Kuroda et al. 358/136
4,054,909 10/1977 Kojima et al. 358/133
4,060,797 11/1977 Maxwell et al. 325/419 X
4,125,861 11/1978 Mounts et al. 358/133
4,168,513 9/1979 Hains et al. 358/261
4,179,710 12/1979 Ishiguro et al. 358/13 X

OTHER PUBLICATIONS

Image Data Compression by Predictive Coding II: Encoding Algorithms Bahi & Kobayashi: IBM J. Res. Develop., Mar. 1974, pp. 172-179.
Frame-to-Frame Coding of Television Pictures Using Two-Dimensional Fourier Transforms: Haskell: IEEE Transactions on Info. Theory: vol. IT-20, No. 1, pp. 119-120; Jan. 74.
Spatial Transform Coding of Color Images: Pratt: IEEE Transactions on Comm. Technology, vol. Com-19, No. 6, Dec. 71, pp. 980-992.
Goertzel et al., Two-Dimensional Data Compression & Decompression System; Aug. 7, 1979.
Application of Fourier-Hadamard Transformation to

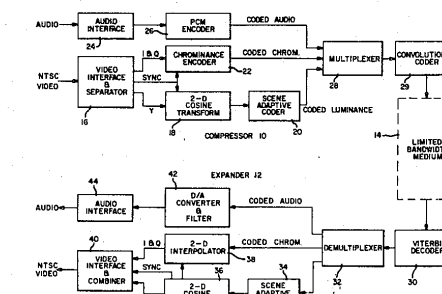
Bandwidth Compression-Pratt & Andrews Proc. Polytechnic Institute of Brooklyn, 1969, pp. 56-68.
Hadamard Transform Image Coding, Pratt, Kane, Andrews, Proc. IEEE, vol. 57, No. 1, Jan. 69, pp. 58-68.
Television Bandwidth Reduction by Encoding Spatial Frequencies, Andrews & Pratt, Journal SMPTE, vol. 77, No. 12, Dec. 1968, pp. 1279-1281.
Television Bandwidth Reduction by Fourier Image Coding, Andrews & Pratt, Paper Delivered to 103rd Technical Conference SMPTE, May 5-10, 1968.
Transform Image Coding, Andrews & Pratt: Proc. Symposium on Computer Processing in Communications, Polytechnic Institute of Brooklyn, Apr. 8-10, 1969, pp. 63-84.

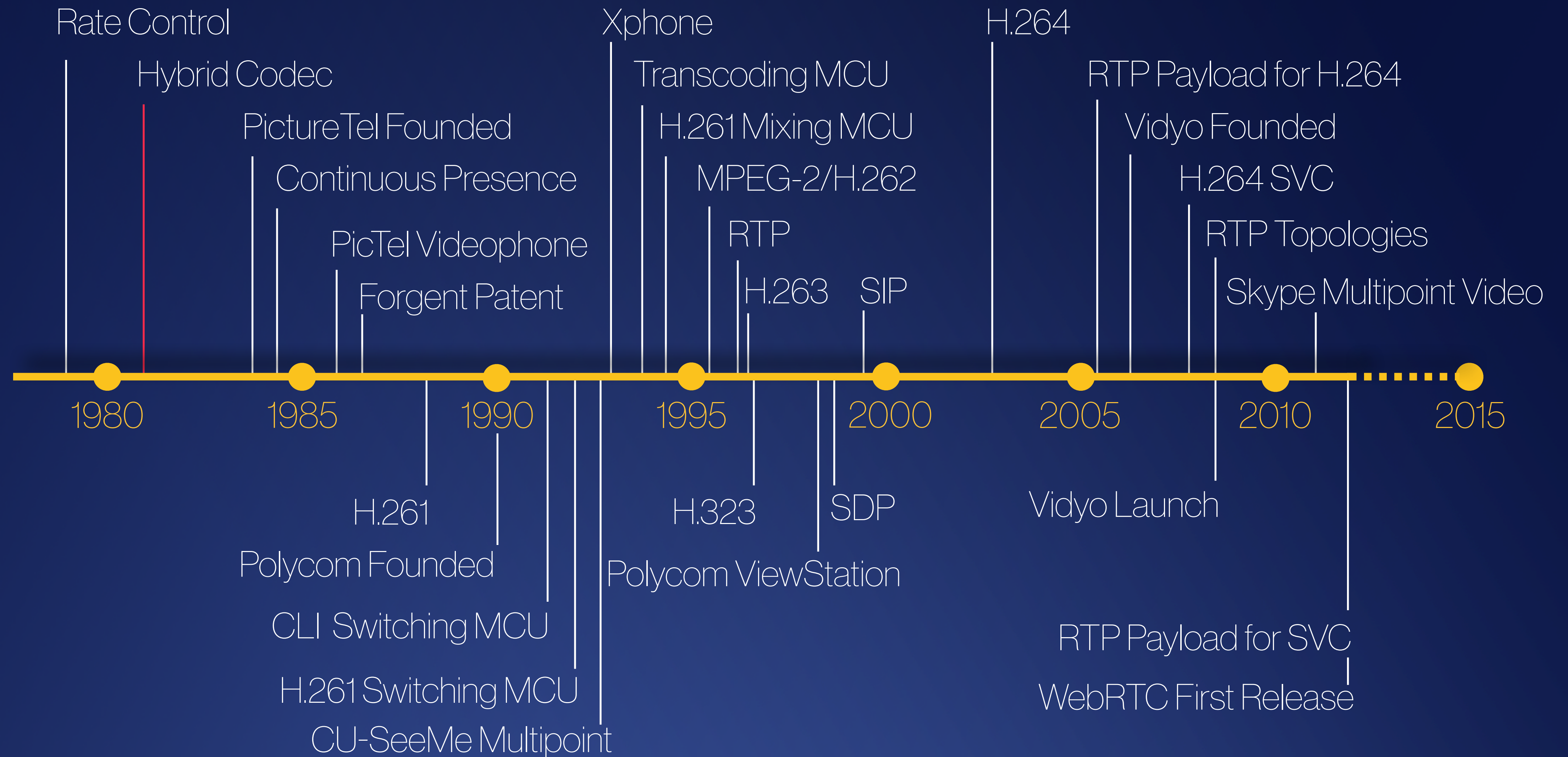
Primary Examiner—James W. Moffitt
Assistant Examiner—Aristotelis M. Psitos
Attorney, Agent, or Firm—David B. Harrison

ABSTRACT

[57] A digital video compression system and its methods for compressing digitalized video signals in real time at rates up to NTSC color broadcast rates are disclosed. The system compressor receives digitalized video frames divided into subframes, performs in a single pass a spatial domain to transform domain transformation in two dimensions of the picture elements of each subframe, normalizes the resultant coefficients by a normalization factor having a predetermined compression ratio component and an adaptive rate buffer capacity control feedback component, to provide compression, encodes the coefficients and stores them in a first rate buffer memory asynchronously at a high data transfer rate from which they are put out at a slower, synchronous rate. The compressor adaptively determines the rate buffer capacity control feedback component in relation to instantaneous data content of the rate buffer memory in relation to its capacity, and it controls the absolute quantity of data resulting from the normalization step so that the buffer memory is never completely emptied and never completely filled. In expansion, the system essentially mirrors the steps performed during compression. An efficient, high speed decoder forms an important aspect of the present invention. The compression system forms an important element of a disclosed color broadcast compression system.

7 Claims, 30 Drawing Figures





Hybrid Codec

December 1981

Jain & Jain

IEEE Trans. on Comm.

Block-based motion
estimation with
transform coding



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Displacement Measurement and Its Application in Interframe Image Coding

JASWANT R. JAIN, MEMBER, IEEE, AND ANIL K. JAIN

Abstract—A new technique for estimating interframe displacement of small blocks with minimum mean square error is presented. An efficient algorithm for searching the direction of displacement has been described. The results of applying the technique to two sets of images are presented which show 8–10 dB improvement in interframe variance reduction due to motion compensation. The motion compensation is applied for analysis and design of a hybrid coding scheme and the results show a factor of two gain at low bit rates.

I. INTRODUCTION

A LARGE number of image transmission and storage applications, e.g., teleconferencing, videotelephone, television and satellite image transmission, medical imaging for computer aided tomography and angiocardiology, etc., contain images of moving objects. The motion captured in such a multiframe sequence of images includes translation and rotation of objects with respect to the camera.

For interframe image coding, large levels of compression could be achieved if only one knew the trajectories traversed by the various objects. Then one could simply code the initial frame together with the trajectory information of each pixel. In practice, a significant component of the motion in a scene can be approximated by piecewise translation of several areas of a frame with respect to a reference frame. Utilization of the knowledge of motion or displacement of pixels in successive frames for image coding is called motion compensation.

Displacement measurement and motion compensation have been applied for interframe image data compression with improved results [1]–[8]. Limb and Murphy [1] and Rocca, Brofferio *et al.* [2]–[4] have considered techniques for estimating translation of a block of pixels. Netravali and Robbins [5] take the approach of predicting the displacement of each pixel recursively from its neighboring pixels which have already been coded. In this paper we present a method of displacement measurement which estimates displacement on a block by block basis. Application of this method in interframe hybrid coding (with and without frame skipping and interpolation) is shown. The results presented here are based on [6] and [7] and the algorithm developed is quite different from other displacement measurement techniques including the one reported in [8]. A detailed bibliography and dis-

ussion of other interframe coding techniques are given in [9] and [10].

A new technique for displacement measurement is described in Section II. This technique is based on an efficient 2-dimensional search procedure. The results of applying this technique for measurement of the displacement on two sets of images are reported in Section III. We find 8–10 dB reduction in variance of the interframe difference signal as a result of motion compensation. Motion compensation is applied for analysis and design of interframe hybrid coding methods in Section IV. Summary and conclusions are presented in Section V. In the Appendix we give a proof of the convergence of the search algorithm.

II. A DISPLACEMENT MEASUREMENT ALGORITHM

In this section, we describe a method of measuring interframe motion for digitized images. First, we approximate the interframe motion by piecewise translation of one or more areas of a frame relative to a reference frame. The segmentation of an image into areas, each of which is undergoing approximately the same translation, and the measurement of the magnitude and the direction of the translation of each area is a difficult task. Cafforio and Rocca [3] describe a method for segmentation and measurement of the displacement of a single moving object in a stationary background. Then, extension of the method to more than one moving object has also been shown. The method becomes increasingly complex as the number of moving areas increases and the size of the image grows larger, since the information concerning segmentation as well as translation is to be coded. Coding of segments with arbitrary boundaries increases the complexity as well as the length of the code to be utilized.

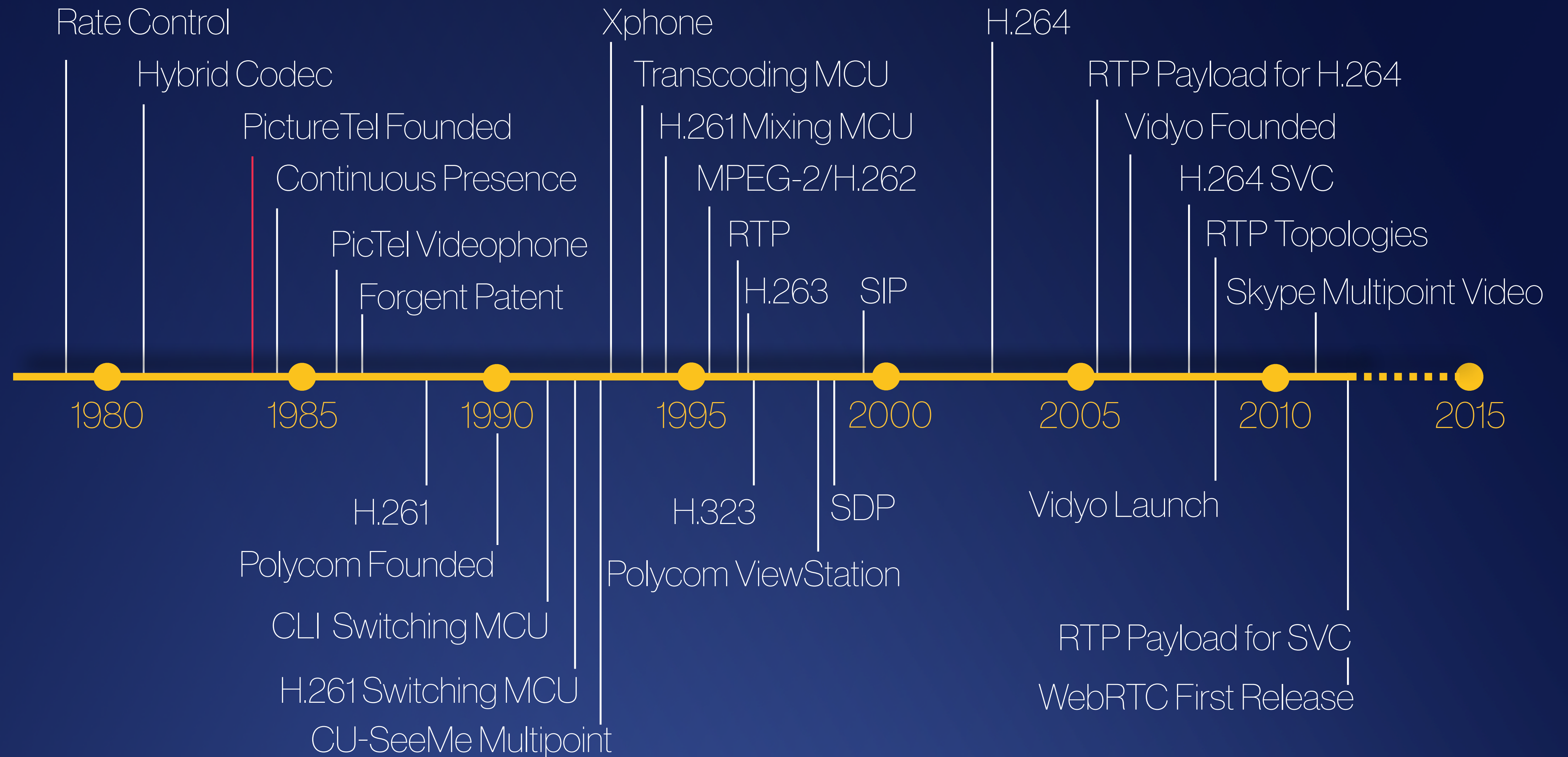
A simpler method is to segment an image into fixed size, small rectangular blocks and to assume that each of these areas is undergoing independent translation. If these areas are small enough, rotation, zooming, etc., of larger objects can be closely approximated by piecewise translation of these smaller areas. It avoids the problem of coding the segmentation information and only the displacement vector of each block needs to be coded.

A method which has been used for the measurement of displacement between two given images, particularly for aerial guidance, is area correlation [11], [12]. This consists of calculating the cross-correlation function of the two images. The location of the peak of the correlation function gives the displacement vector. The cross-correlation function is usually calculated via the fast Fourier transform (FFT). To improve the accuracy of this method some high-pass filtering (which

Manuscript received January 5, 1981; revised August 28, 1981. This paper was presented at the Picture Coding Symposium, Ipswich, England, July 1979.

J. R. Jain is with the Defense Division, Systems Control, Inc., Palo Alto, CA 94304.

A. K. Jain is with the Signal and Image Processing Laboratory, Department of Electrical and Computer Engineering, University of California, Davis, CA 95616.



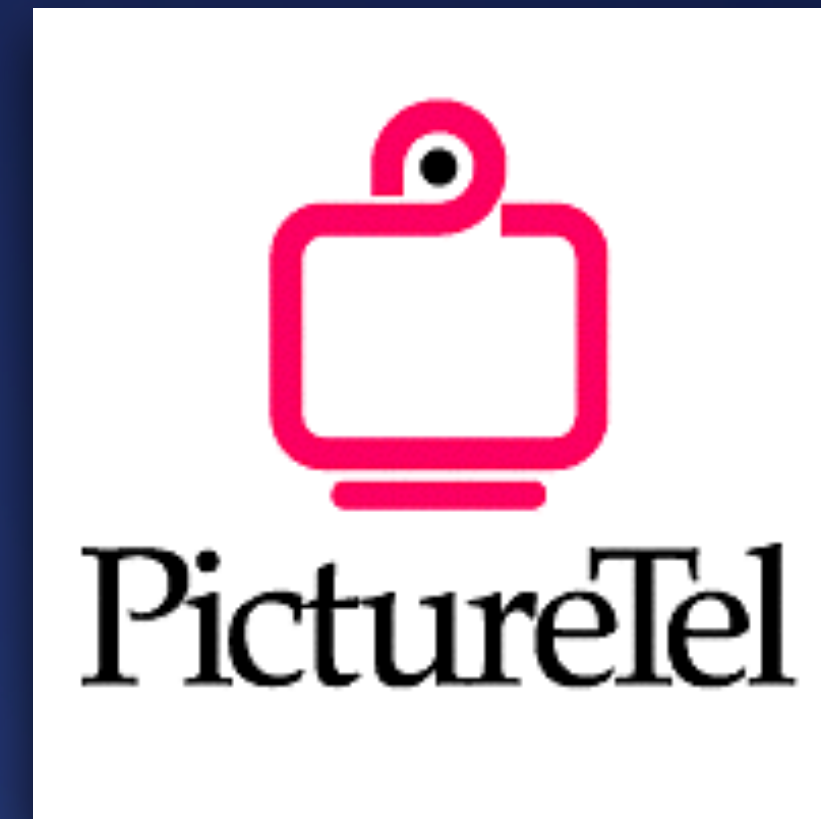
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

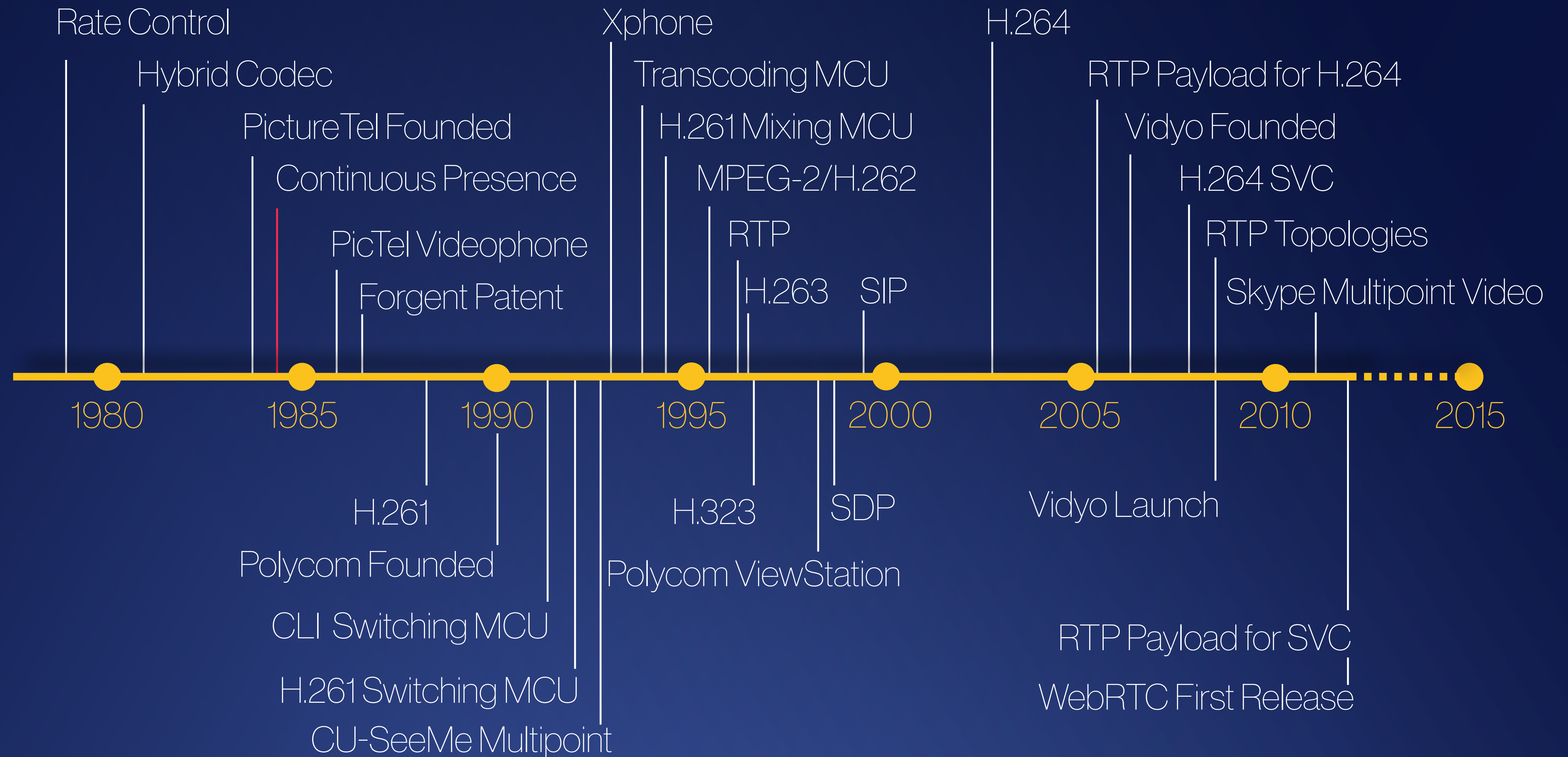
PictureTel Founded

August 1984
(as "PicTel")

Hinman, Bernstein, and Stalin

with MIT'ers Dertouzos,
Papadopoulos, and Soley





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Continuous Presence

April 1985
Sabri & Prasada
Proc. of the IEEE

P2P “continuous presence”
using multiplexing of
multiple cameras
(like telepresence)
or switching



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Video Conferencing Systems

SHAKER SABRI, MEMBER, IEEE, AND BIRENDRA PRASADA, SENIOR MEMBER, IEEE

Invited Paper

In this paper several video conferencing systems are discussed. These include single person-camera, voice-switched, split-screen, continuous presence, and virtual space systems. Special emphasis is placed on the more recent video conferencing systems, i.e., the continuous presence and the virtual space systems. The role of digital signal processing in the video conferencing environment is discussed. The interaction between service definition, the video conferencing system, and the digital signal processing requirements is highlighted.

I. INTRODUCTION

Teleconferencing systems can generally be classified into three main categories. The first category is audio conferencing where the system provides a shared aural space between participants using the telecommunication network. The second category is the visually augmented audio conferencing system. In this type of system, a shared visual space, in addition to the aural space, is provided. The visual space provides the means for displaying, modifying, and interacting with images. These images can be text, graphics, or photographic still images and can, in some cases, involve some form of animation. The third category is video conferencing. This type of system provides means for communicating live (moving) pictures of conference participants thus expanding the shared visual space. It also subsumes the first and second categories.

Each of the main categories; namely, audio, visual, and video conferencing, is generic. In each main category a number of different services can be supported.

This paper deals with video conferencing systems, i.e., the systems that support an enhanced visual space which includes live images of one or several participants. Audio and visually augmented audio systems are addressed in [1].

A. Historical

One of the earliest experimental video conferencing systems was installed between two locations of Bell Laboratories at Murray Hill, NJ and Holmdel, NJ [2]. The Bell Labora-

Manuscript received October 3, 1984.
S. Sabri is with Bell-Northern Research, Nuns' Island, Verdun, Que., Canada H3E 1H6.
B. Prasada is with Bell-Northern Research and INRS-Telecommunications, University of Quebec, Nuns' Island, Verdun, Que., Canada H3E 1H6.

ories system used one full-duplex analog TV channel for transmission. Each of the conference rooms had three cameras to cover the participants. A voice-switching strategy was used to transmit the output of the camera covering the speaker to the other site. When nobody was speaking at a site the output of a fourth camera with a wide-angle lens covering all the participants was transmitted.

The Bell Laboratories system was followed by several experimental systems which included the CONFRAVISION system of British Telecom Research [3], Australian Post Offices split-screen system [4], Bell Canada's system which used voice-switched as well as split-screen systems in different locations, systems at NTT [5], and other telephone administrations. All these systems utilized analog TV transmission technology and standard TV cameras and monitors.

These were followed by video conferencing systems based on digital transmission. Bell Canada utilized digital transmission at 45 Mbits/s in their Montreal-Toronto link. In France, commercial video conferencing service at 2 Mbits/s began in 1976 [6]. The current European system is based on COST 211 codecs jointly designed and standardized by 10 European countries [7], [8]. The 2-Mbit/s codec and satellite-based networking characterize the European systems. NTT in Japan has implemented video conferencing with 6.3-Mbit/s (and also 1.5-Mbit/s) codecs and digital transmission facilities [9]. In North America, besides the public service of AT&T [10], Bell Canada, and Telecom Canada [11], a number of private networks have emerged [12], [13]. The North American systems are operating at transmission rate of 1.5 Mbits/s.

B. Approach in the Paper

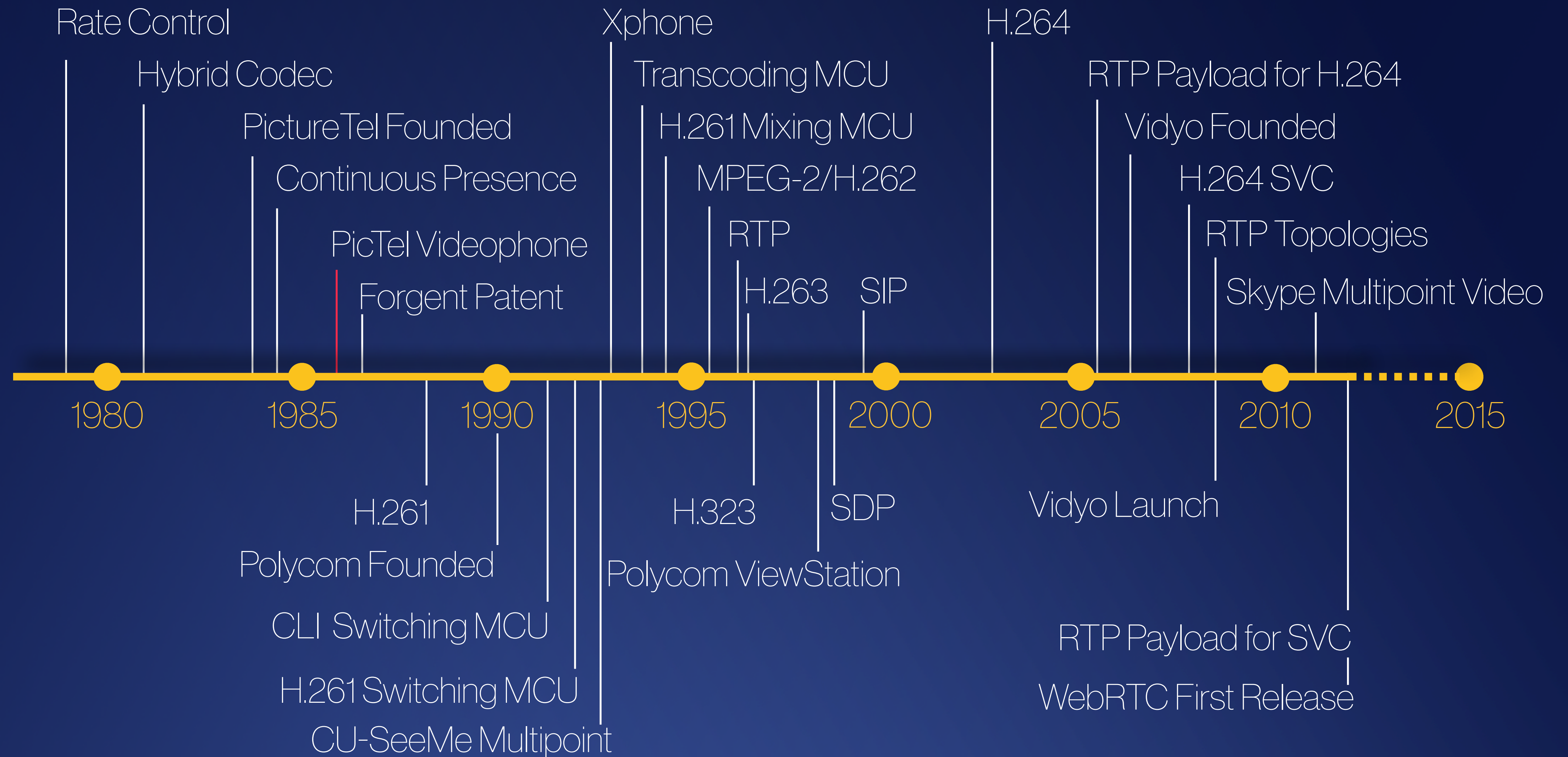
The approach followed in this paper is to focus on the generic video conferencing systems and services that are the basis of all the different systems in this fast-emerging field.

The generic video conferencing systems fall into one of the following classes:

i) *Single Person-Camera System (SPC)*: A single person-camera (PC) is used to capture a view of conference participants and the resulting video signal is transmitted and displayed at the other end.

ii) *Voice-Switched System (VSW)*: In this system several person-cameras are used to capture pictures of the par-

0018-9219/85/0400-0671\$01.00 ©1985 IEEE



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

PicTel Videophone

March 1986

2x52 Kbps
7.5 fps, 11 boards
\$150K for a pair

MARCH 31, 1986 COMPUTERWORLD 73

NEW PRODUCTS

TI adds Travelmate series terminals to Silent 700 line

Texas Instruments, Inc. of Dallas has announced the addition of the TI Silent 700 Travelmate series of portable display terminals to its Silent 700 family.

The Travelmate series includes the Travelmate, the Travelmate 1200 and the Travelmate DT. All three offer user interface modules that, according to the vendor, can be programmed to provide customized solutions for users' needs. The interface modules are plug-in application cartridges that provide 32K bytes of read-only memory and 24K bytes of random-access memory per cartridge.

All three standard Travelmate terminals are said to feature a retractable, 16-line by 80-column LCD screen, built-in 45 char./sec. thermal printer, full-size keyboard, data communications interfaces and specific application and text editing functions.

The Silent 700 Travelmate and Travelmate 1200 were designed for portable data communications applications. They come with internal 300 and 300/1,200 bit/sec. AT&T compatible modems, respectively. The Travelmate DT was designed for desktop applications in which the terminal has direct connection to the host computer system via the RS-232C interface said to support up to 9.6K bit/sec. communications.

The Travelmate terminal costs \$1,095; the Travelmate 1200 lists for \$1,395; and the Travelmate DT costs \$995, the vendor stated.



TI's portables can be programmed to provide customized solutions for users.

Videophone system bows from Pictel

Pictel Corp. of Peabody, Mass., recently introduced a videophone system incorporating its proprietary video image compression technology.

The system is said to include Minix videophones and a network controller manufactured by Datapoint Corp. According to the vendor, the system can support seven terminals in a star configuration network.

Pictel executives claimed that their proprietary video coder/decoder computer will transmit high-quality images and voice over low-cost 56K bit/sec. transmission lines, such as AT&T's Accunet Switched 56 Service.

A spokesman stated that Pictel's Motion Compensated Transform compression algorithm measures and compensates for changes in motion, reducing the system's need to regenerate images of moving objects.

Voice is transmitted over a separate audio line.

Software able to be upgraded

According to a spokesman, Pictel offers an advantage over earlier videoconferencing products by being software upgradeable. With the Pictel system, users will be able to plug in a software card to gain future enhancements, he explained.

The commercial product will operate at 52K bit/sec. over full-duplex digital lines and will operate at faster speeds with improved quality over dedicated and T1 lines.

The product will be available by the end of the year at a cost of \$150,000 for a system including five Datapoint terminals, a local-area network connecting terminals within a building and the video coder/decoder.

The Datapoint terminals accommodate IBM Personal Computers and compatibles, but the system is not currently capable of transmitting data, the spokesman said. He added that graphics can be compiled on a micro, and pictures of the graphics can be transmitted.

DEC introduces LG family

Two line matrix printers available for text, graphics

Digital Equipment Corp. of Maynard, Mass., has introduced the LG family of printers, including the LG01 for text and the LG02 for text and graphics.

The LG01 and the LG02 are said to be 600 line/min line matrix printers. According to the vendor, the LG01 offers data processing and correspondence-mode printing as well as multiple char./in. print selections. It is said to be compatible with both existing U.S. and European character sets.

The LG02 text/graphics printer reportedly offers the capability to create bar codes, custom forms and logos; superscript and subscript modes; and prints in landscape mode. In addition, the LG02 printer provides all of the text capabilities of the LG01 printer, the vendor stated.

According to a spokesman, the number of moving, mechanical parts of both printers has been reduced, providing electrical reliability. The vendor also claimed that the printers require no scheduled maintenance calls.

Both printers are said to be compatible with DEC VAX computers from the VAX 8600 to Microvax II systems as well as with the company's Micro/PDP-11/73, Micro/PDP-11/83 and Micro/PDP-11/84 computer systems.

The LG01 costs \$11,950. The LG02 costs \$14,000.

HP library implements GKS

Hewlett-Packard Co. of Palo Alto, Calif., has introduced a two-dimensional graphics library that is said to be an implementation of Level 2B of the industry-standard Graphical Kernel System (GKS).

The HP-GKS library was designed for use with HP 9000 technical computers and graphics hardware under the HP-UX operating system.

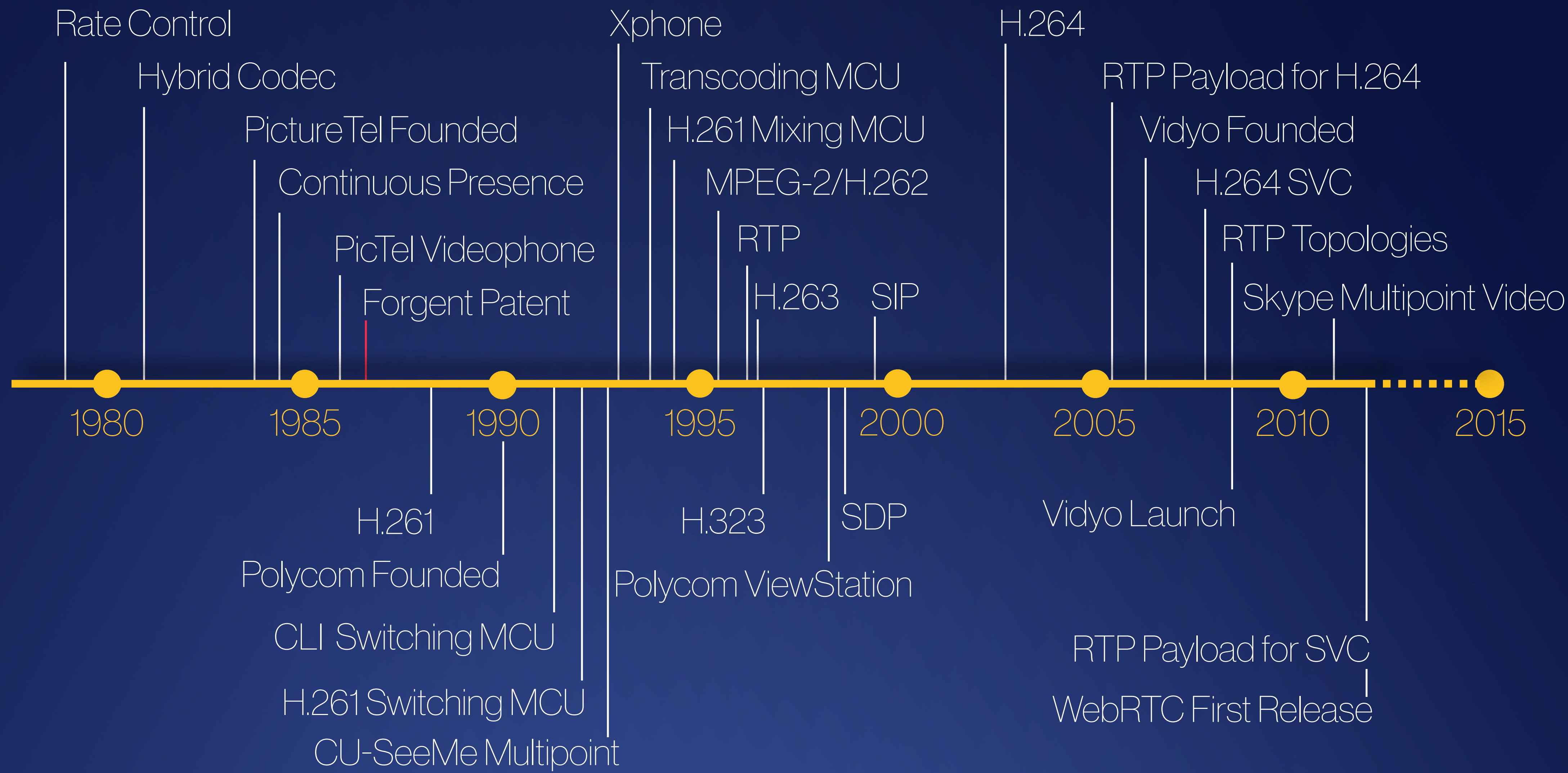
It is a tool that applications developers can use in drafting, process control, mapping, simulation analysis and presentation graphics. The applications developed may achieve performance up to 50,000 2-D vector/sec. on high-performance graphics display hardware.

Features of HP-GKS include out-

segment operations that can be used to create, copy, associate and delete segments; metafile input and output defined using the GKS metafile format; attribute binding that allows the attributes to be specified individually for each primitive or bundled together; tables that can be defined for general use; and input operations that are available for six classes of input in request mode.

HP-GKS is said to support a dynamic workstation configuration and provide device-independent access to graphics workstations. In addition, applications that have already been written to the GKS standard are portable to the HP 9000 systems.

ME-GKS for the HP 9000 Series

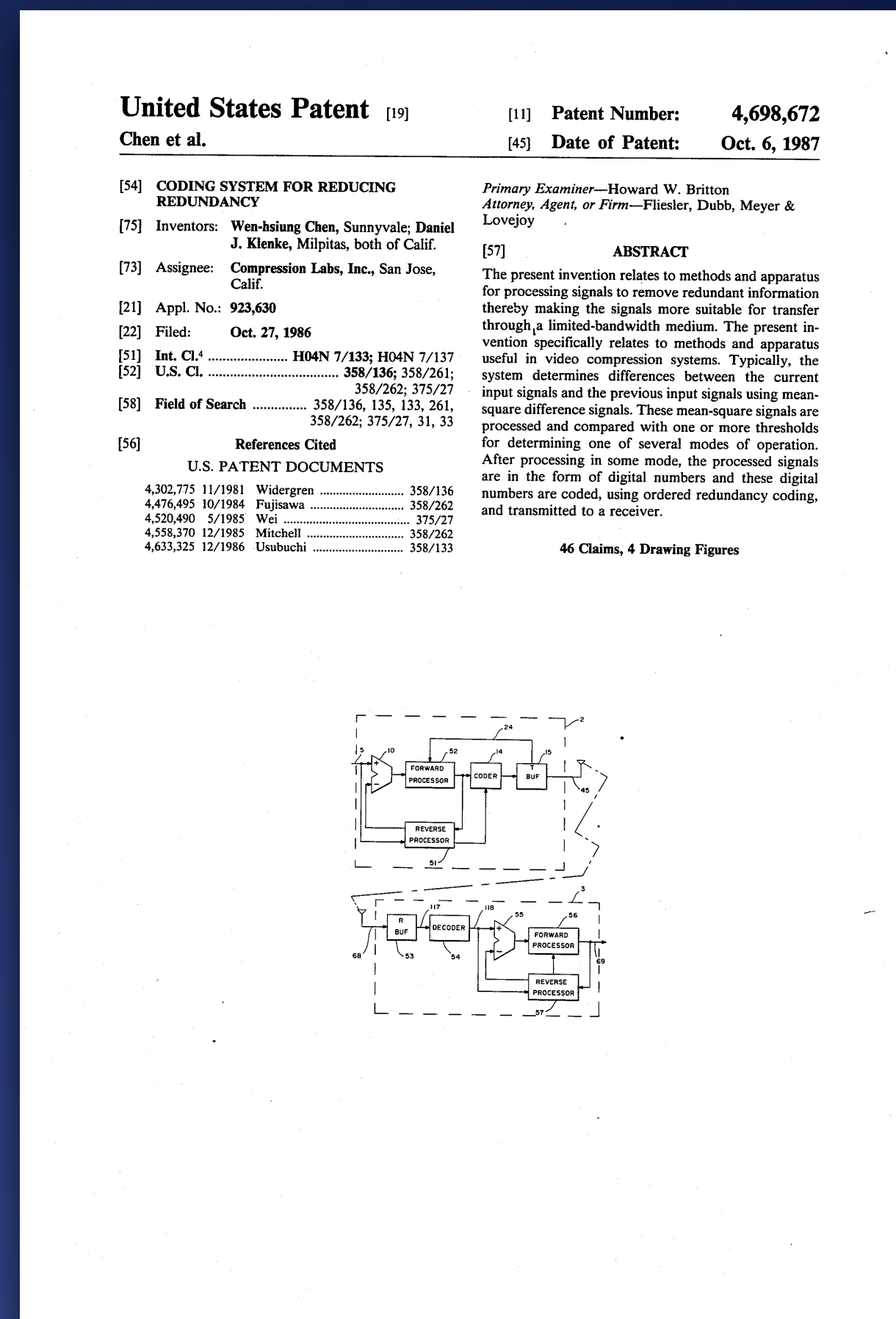


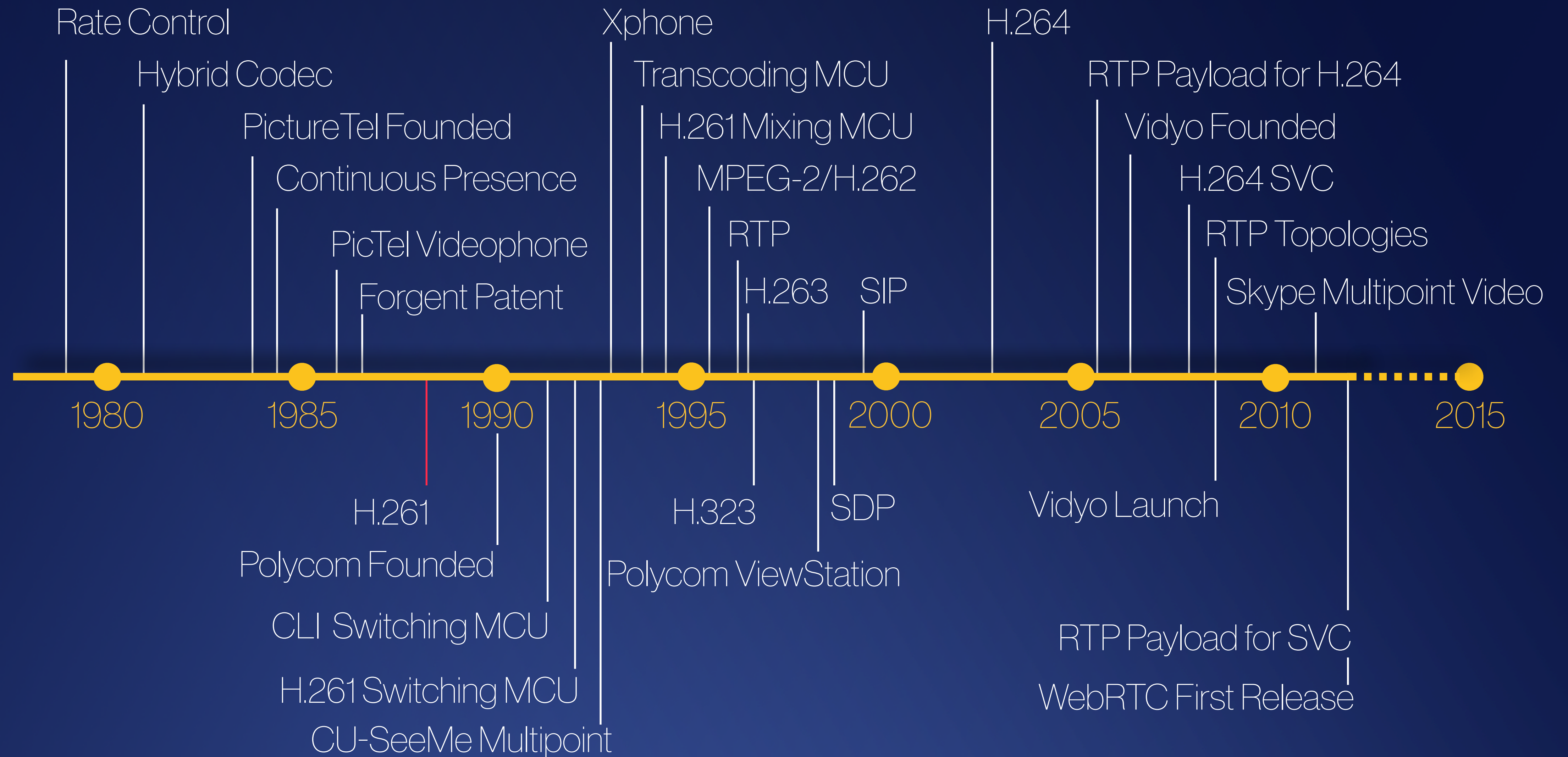
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Forgent Patent

October 1986
US Pat. Nr. 4,698,672
Compression Labs, Inc.

“JPEG patent”



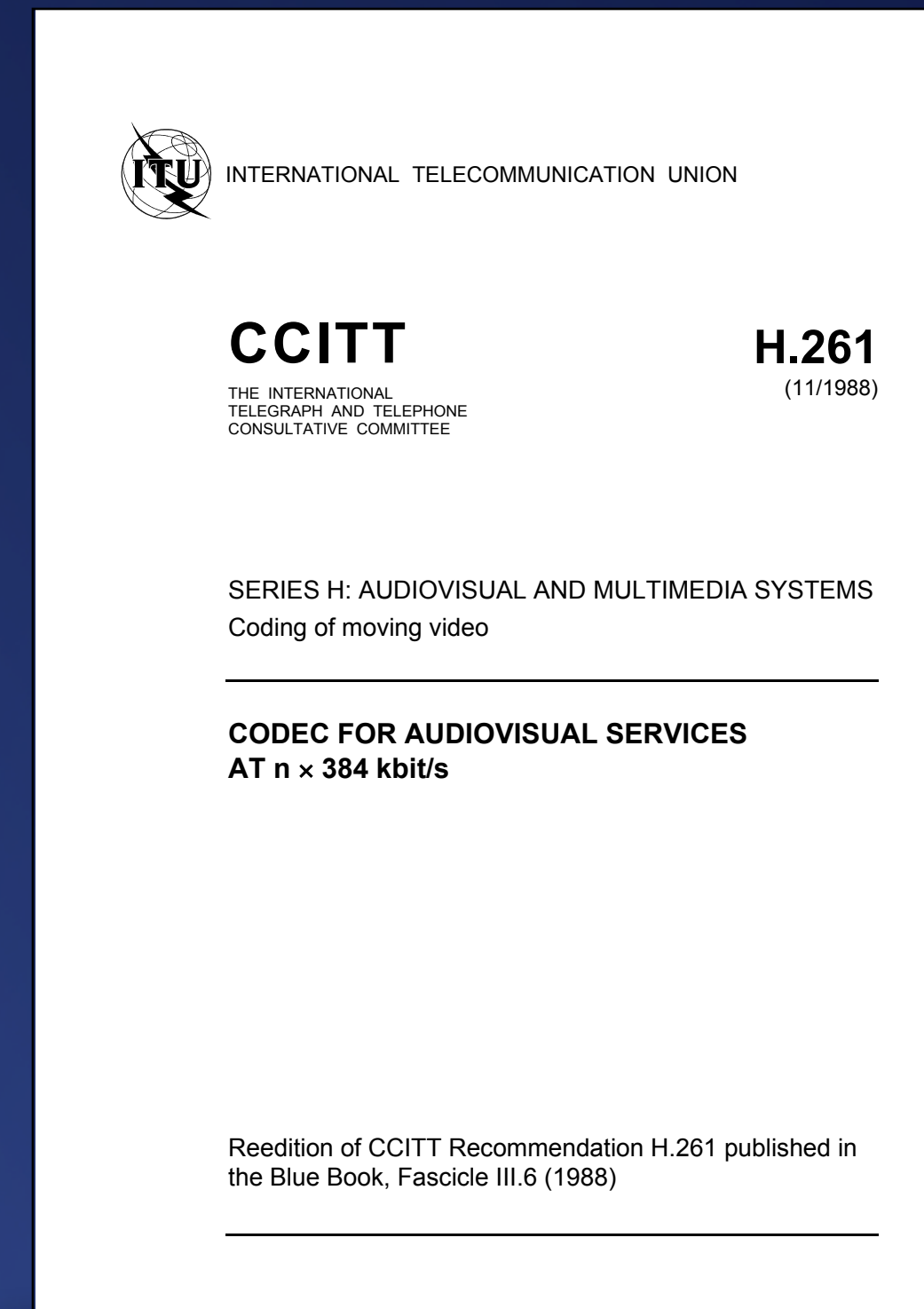


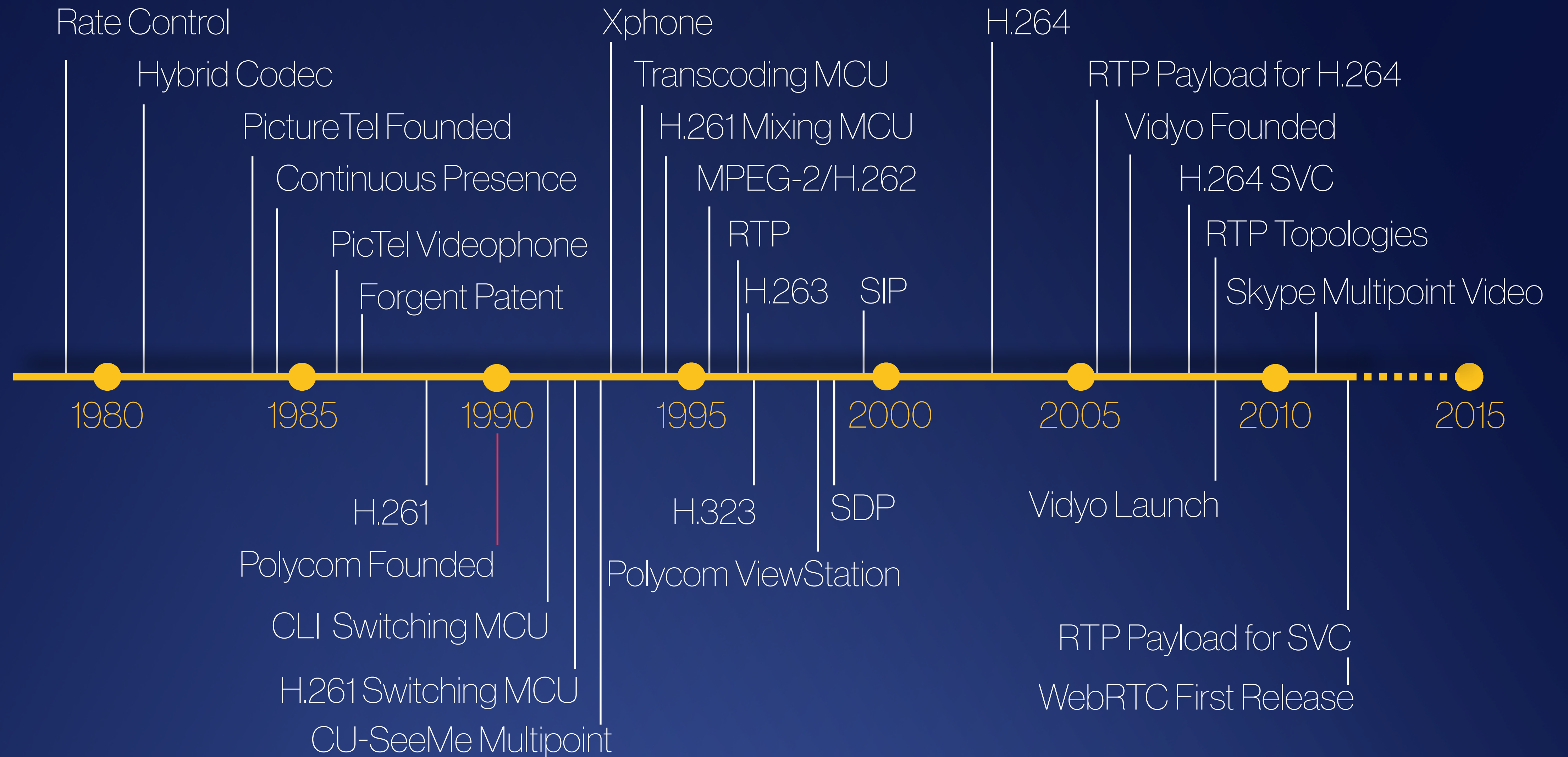
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

H.261

November 1988

First digital video
compression standard
for communications



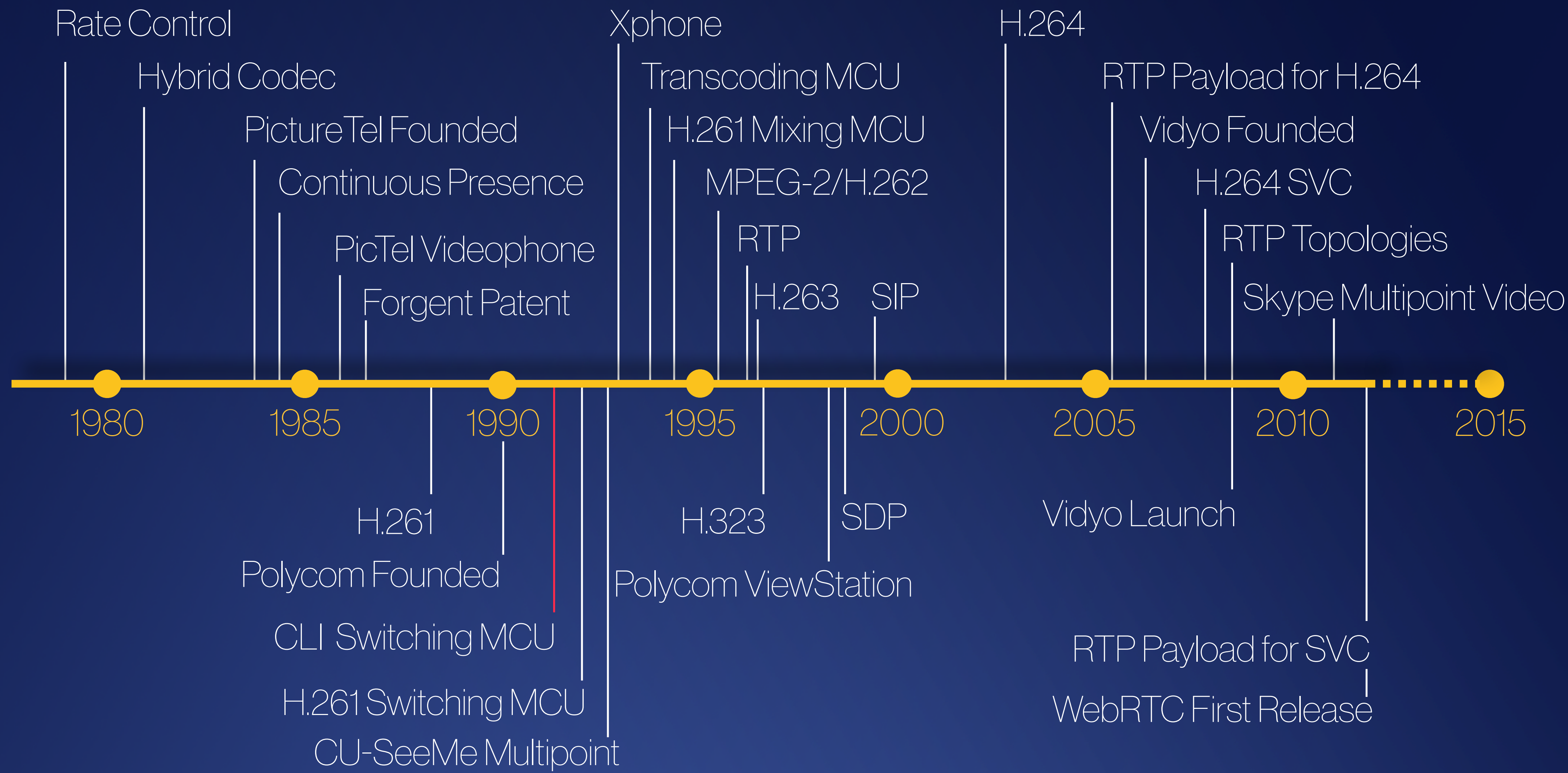


Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Polycom Founded

January 1990
Hinman and Rodman





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

CLI Switching MCU

February 1991

8-way
or 14-way with
cascading

MANAGEMENT STRATEGIES

MANAGING PEOPLE AND TECHNOLOGY: USER GROUPS AND ASSOCIATIONS

Worth Noting

“For almost all companies, communications networks are as much a competitive tool as having a modern factory building.”

Bob Brown
Telecommunications/
networking associate
Milrose, Mass.

Association Watch

Timeplex, Inc. will hold its first Users Network Conference from Feb. 27 to March 1 in Orlando, Fla. The conference is designed to give Timeplex's North American customers the opportunity to shape the company's product direction and strategies.

Timothy Zerbe, vice-president of advanced technology at Timeplex, will deliver the keynote address on "Technology Trends in the '90s." The conference will also feature discussions on network and bandwidth management, as well as local area network internetworking and disaster recovery.

Sessions are planned for Timeplex's LINX + multiplexers, Timepac packet switches, Timepath cell-relay products and Time/LAN internetworking Fiber Distributed Data Interface bridges, routers and concentrators. For information, contact Jo Ann Turli at (201) 573-6475.

Service Systems International, Ltd. recently announced the formation of the S2000 Service Management System Users Group, which held its first meeting last fall in Overland Park, Kan.

At the meeting, Service Systems unveiled Release 3.0 of the wide-area version of its S2000 for the IBM Application System/400. The software is a service management package that automates most of the daily functions of a service operation.

Service Systems can be reached at (915) 661-0190. □

Survey finds some workers unclear on company ethics

Respondents split on acceptance of some gifts.

By Bob Wallace
Senior Editor

HOUSTON — An informal survey of attendees at the recent International Communications Association (ICA) Winter Seminar here revealed that although users and vendors must abide by comprehensive codes of ethics, some of those policies are not fully understood.

A total of 71 users and vendors answered all or part of the 11-question survey, which was conducted by Mark Smith, deputy general counsel for Northwest Mutual Life Insurance Co. of Milwaukee. Smith stressed that the survey was not an official ICA undertaking.

"We wanted to give attendees an idea of where [users and vendors] stood as far as the ethics process and corporate requirements are concerned," Smith said. Respondents were not required to sign the questionnaires. Attendees were asked if their companies have a code of ethics, set of guidelines or policy statement relating to business conduct and conflicts of interest. Sixty-three respondents said their firms follow such guidelines, seven indicated they did not know, and one said no such formal program existed at his firm.

Fifty-seven attendees said their company's code or statement addresses employee acceptance of gifts, travel, entertainment, fees or compensation from firms doing business or seeking to do business with their company. Three said their guidelines address these areas, and five said they did not know.

Of those responses, 18 said acceptance of items is absolutely prohibited, while 49 said they are allowed provided they fall within certain guidelines.

Most respondents said restrictions or limitations are placed on what may be accepted in terms of type, value or quantity. Many users said they are permitted to accept items less than \$10, \$15 or \$25 in value.

(continued on page 18)

Videoconferencing propels Bendix/King

Aerospace firm relies on video sessions to keep departments, suppliers and customers in sync.

By Moussem Mohy
Staff Writer

OLATHE, Kan. — Bendix/King, a manufacturer of aircraft radios and flight instruments, is using videoconferencing to gain a competitive edge in its aerospace program by streamlining manufacturing operations and coordinating the flow of goods from suppliers.

The use of videoconferencing allows design engineers at sites here, in Florida and the Far East to confer on an as-needed basis, giving the company a strategic edge over competitors by significantly trimming product development time.

"Videoconferencing is a breakthrough technology that's changed the way our company does business," said Larry Ehlers, General Aviation Avionics Division manager of plant engineering. "Quicker and better communications with our teaming partners, customers and vendors was the key motivation for installing videoconferencing; reduced travel was only a minor consideration."

In one instance, the network enabled Bendix/King to speed up development of a Traffic Alert and Collision Avoidance System (TCAS) for which several airlines had asked different manufacturers to submit prototypes. The product's design was completed a year before major trials' products, thereby allowing Bendix to capture nearly 60% of the \$1 billion TCAS market.

Ehlers said Bendix/King uses videoconferencing throughout the development phase and has begun using it to coordinate de-

sign and manufacturing operations as well.

Videoconferences are held three times a month with management teams in Singapore to coordinate manufacturing operations and to design the next generation of airborne communications and navigation equipment. Monthly videoconferences are also held between Bendix/King and its two largest customers, The Boeing Co. and McDonnell Douglas Corp.

Larry Ehlers

Videoconferencing is now used for software demonstrations for the company's data processing and maintenance departments. Finance, project and inventory reviews are also completed via videoconferences.

Ehlers said this frees up time and money. While the division president previously traveled to Kansas and Florida for two separate reviews, both reviews can now be completed from a single location.

Bendix/King, a division of Allied-Signal Aerospace Co., was the first organization to install

(continued on page 18)



Bendix/King videoconferencing room in Olathe, Kan.

Establishing trust helps projects succeed

BY WAYNE ECKERSON

More than ever, network managers are being required to cost-justify every project they want to undertake. Not surprisingly, many strategic projects never get off the ground because the benefits they provide cannot easily be quantified in dollars and cents.

Projects that generate soft benefits — such as increases in productivity, enhanced customer service or improved quality — usually provide significant paybacks. Companies that fail to fund these strategic projects save money in the short term but sacrifice their long-term competitiveness.

Firms that apply technology in innovative ways often have senior executives who are visionary and willing to take risks. They know that technology, if applied properly, can position their company to be effective competitors in the long run.

Too few executives, however, have that degree of confidence in information technology. Many have been burned by technology projects in the past that have promised much and delivered little.

Not surprisingly, some executives are extremely reluctant to risk their reputations and precious corporate resources on another strategic technology scheme.

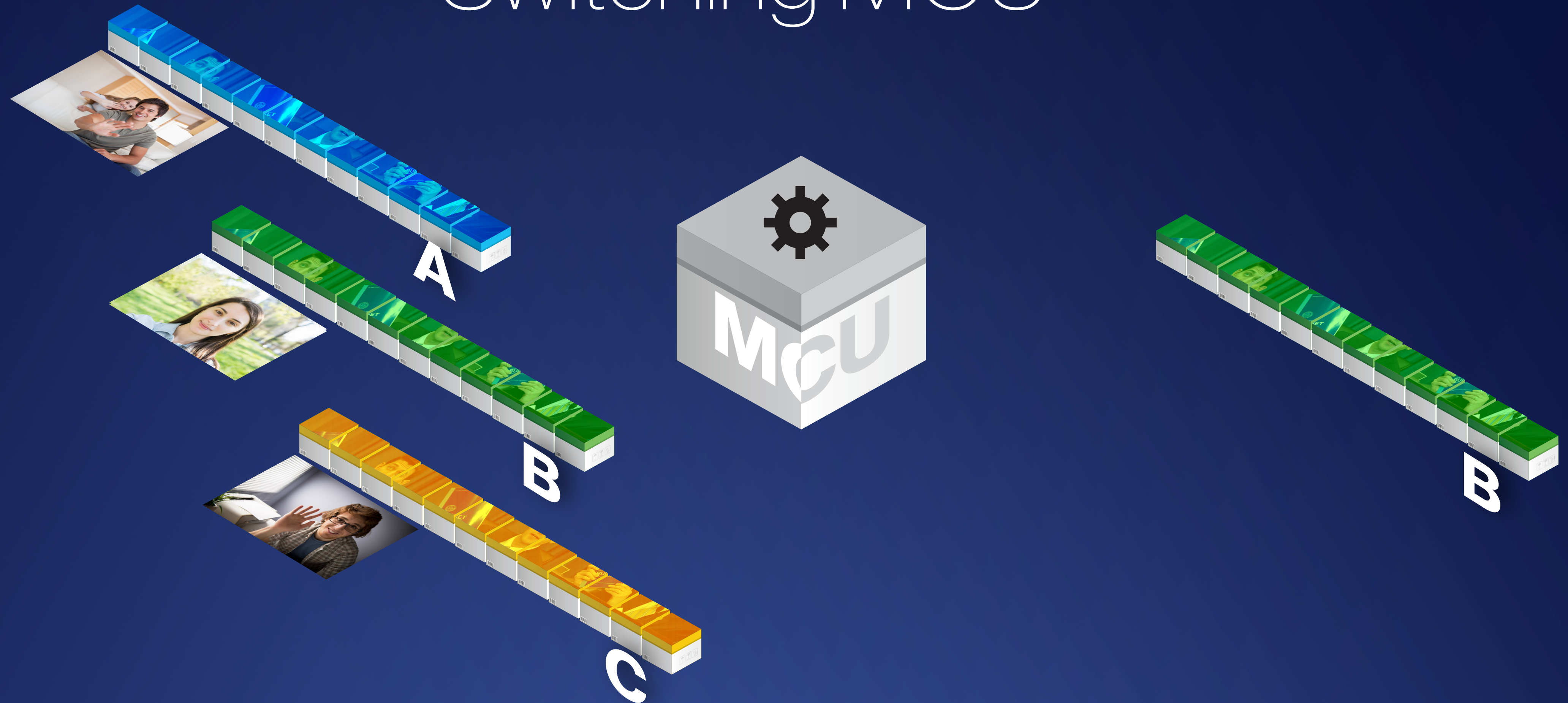
According to Bud Mathiasel, executive director of Ernst & Young's Center for Information Technology and Strategy in Boston, it all boils down to the credibility of those backing the project.

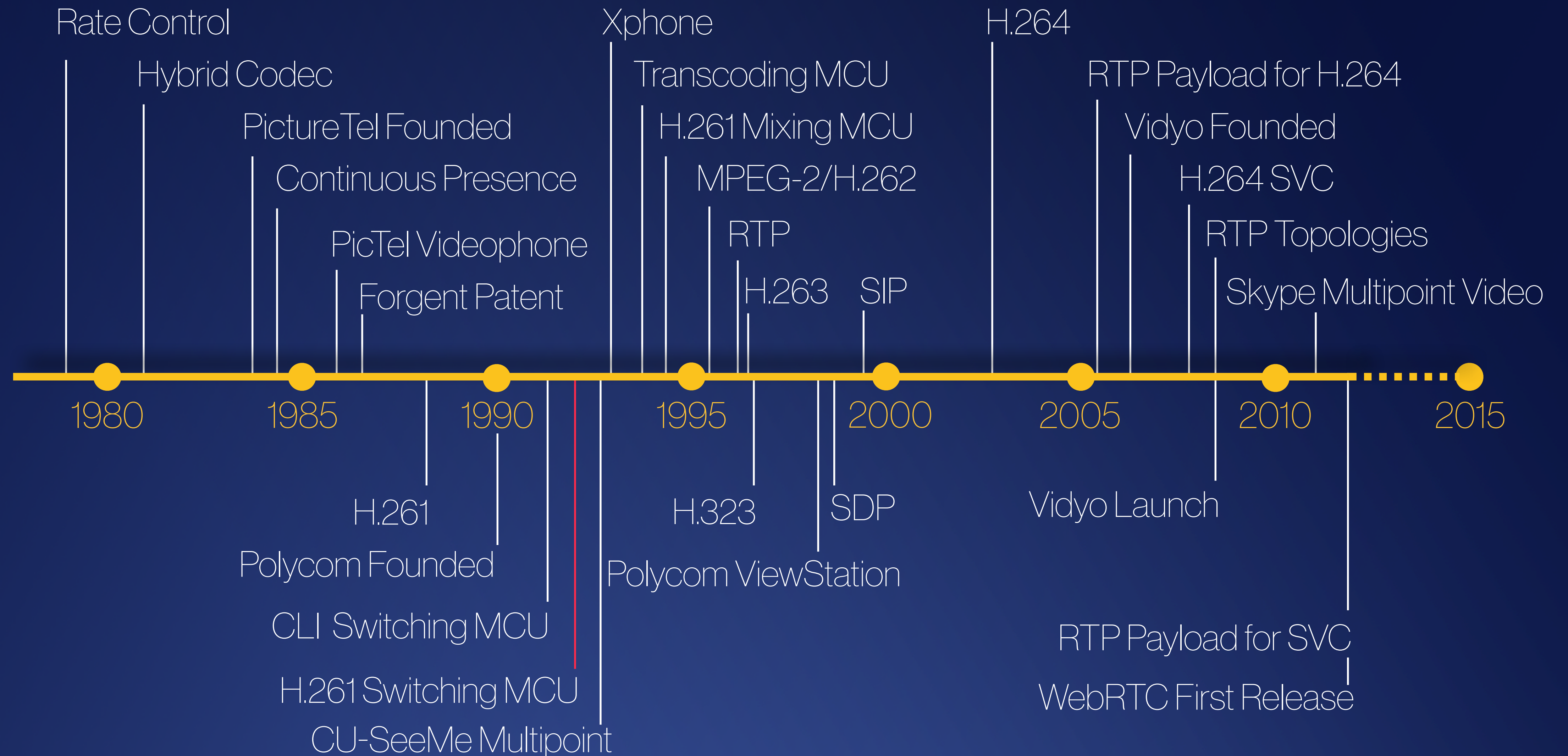
"Executives will approve projects if they trust the people behind them and those people have a good track record," said Mathiasel, who was formerly chief information officer at The Walt Disney Co.

(continued on page 18)

NETWORK WORLD • FEBRUARY 4, 1991 • 17

Switching MCU





H.261 Switching MCU

May 1992
Clark

IEEE Comm. Magazine

Switching using H.261
(MIAS ESPRIT Project)



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Multimedia Conferencing

Multipoint Multimedia Conferencing

The expansion of ISDN and the introduction of standardized audio and video coding algorithms are accelerating the development of multipoint multimedia communications.

William J. Clark

In general terms, telecommunications systems have traditionally been designed to operate on a point-to-point basis, either between two end users or between a user and a centralized facility such as a database. Where communication is required between a number of locations (for instance, facsimile transmission between a headquarters office and a number of branch offices), this has been achieved by either a store-and-forward facility or sequential transmission from one site to each of the others. While such techniques may be quite adequate for non-real-time applications, in cases where communication between people is in progress by voice or picture, simultaneous reception of information by a number of sites is necessary. Multipoint services provide for this real-time transmission between three or more locations. The generic term "teleconferencing" is also often used, although this also covers the case where a number of people are involved on a purely point-to-point link. One of the features of teleconferencing services is that users often have a need to employ many different types of media in order to more accurately simulate a face-to-face meeting. Besides hearing and seeing other participants, they must be able to transmit documents, files, and still images, and also replicate some of the human interactions of meetings such as chairmanship.

With the increase in businesses operating on a multi-site basis, the need for standardized multipoint multimedia services is apparent. This article describes the development of multipoint multimedia services for conferencing with emphasis on the use of Integrated Services Digital Network (ISDN) as part of the work carried out in the European collaborative projects Multipoint Interactive Audiovisual Communication (MIAC) and Multipoint Interactive Audiovisual System (MIAS).

History

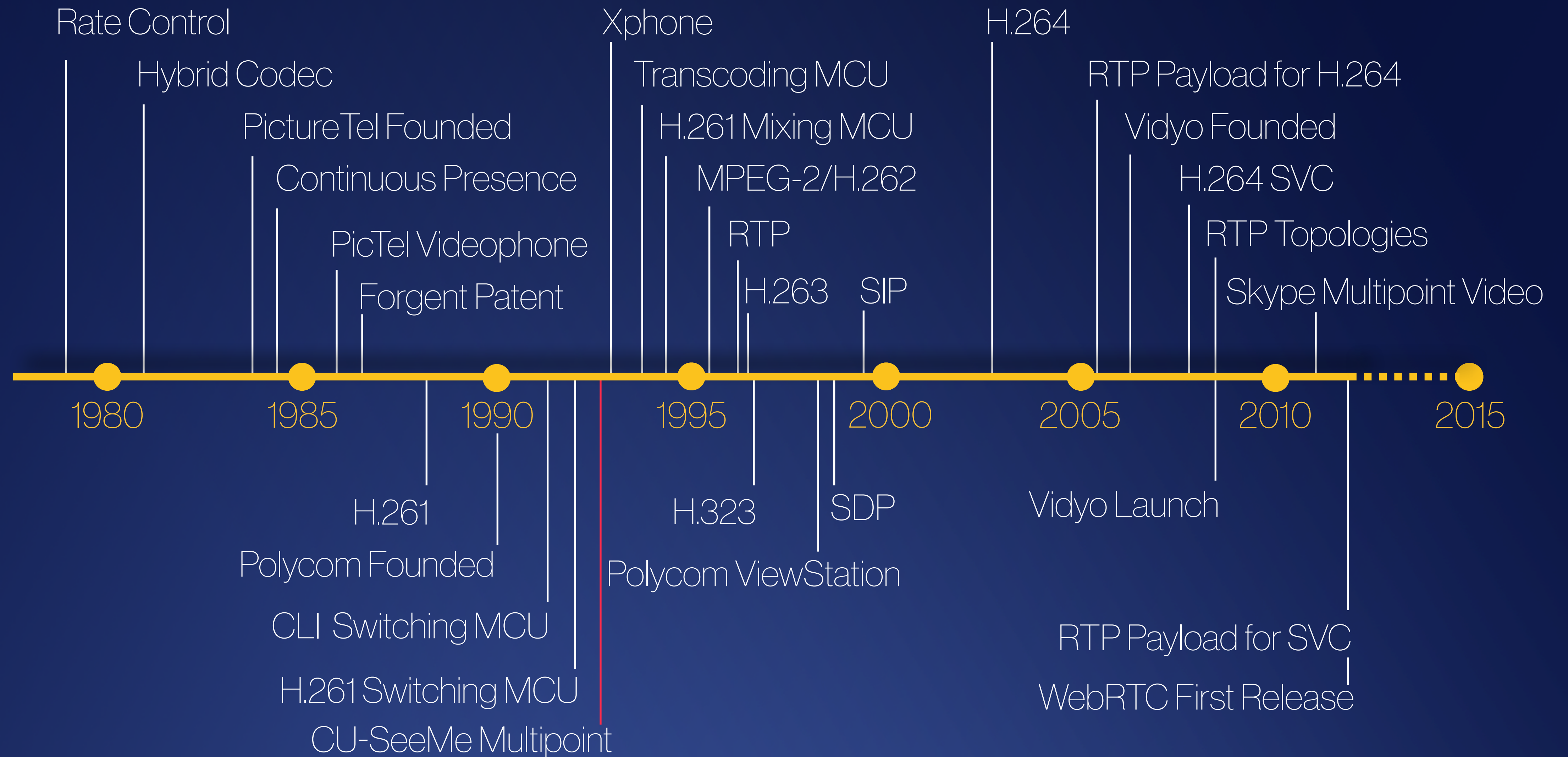
From the earliest days of the telephone, service operators have taken the opportunity to provide additional services. The first use of what might be called multipoint telephony took place dur-

ing the last century when the sound of a live performance from the opera stage was sent to a number of subscribers. In this system, the sound was transmitted in one direction only, and due to the lack of amplification only a limited number of receivers could be used. The first multipoint audioconferences took place in the 1930s, linking housebound and hospitalized students in a school district of Iowa [1].

In such simple audioconferences, only audio is transmitted. Once the idea of linking a number of sites for audio had been established, it was natural to attempt to add some form of pictorial or graphical information. Basically, three types of additional information can be considered: still images, such as facsimile and slow-scan TV (now known as still-picture television); computer data, such as files; and moving video images, generally derived from a camera. The combination of these information sources with audio are known as audiographic teleconference and videoconference, respectively. For historical reasons these have tended to be considered as separate services, although, as described later, one can envisage systems providing a single service across a whole range of facilities.

Audiographic teleconferencing started in the late 1960s, using the telephone network to provide audio together with some supplementary facility. An early example of this was the use of audio and facsimile by the National Aeronautics and Space Agency (NASA) to coordinate the Apollo program. In the late 1970s, various experiments were conducted in the use of still-picture TV in conjunction with audio. In 1980, British Telecom Laboratories (BTL) implemented trials of applications for still-picture TV, including telemedicine and teleconferencing. Resulting from this experience, a successful system was developed, known as IMTRAN [2], which provided for the transmission of body scanner images from a hospital to a remote consultant.

In 1981, BTL with the Open University conducted a human-factor trial of the CYCLOPS system [3]. This equipment allowed users to draw interactively on the face of a television screen by means of a light pen and exchange simple images, drawings, or alphanumeric characters, together with an

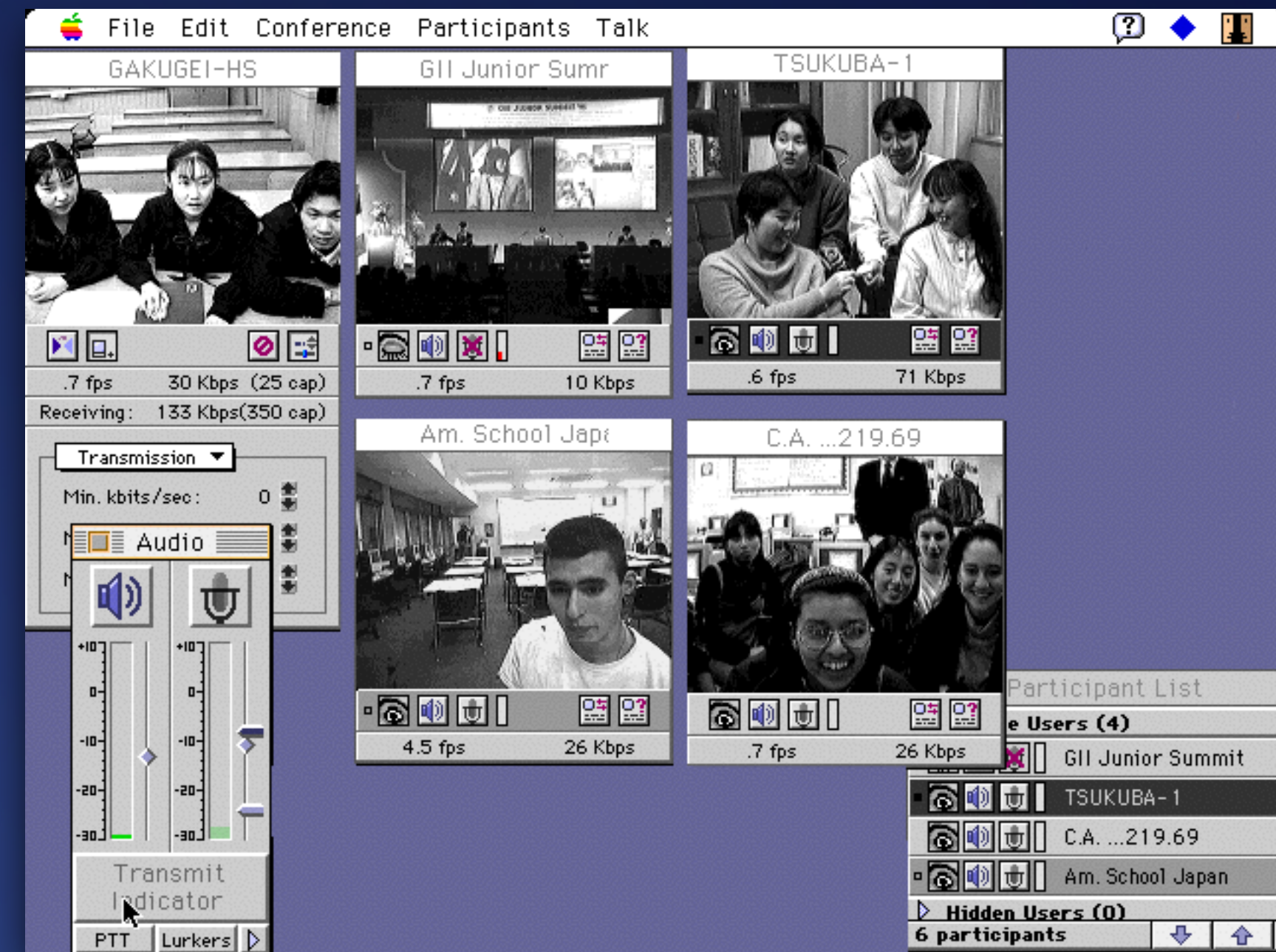


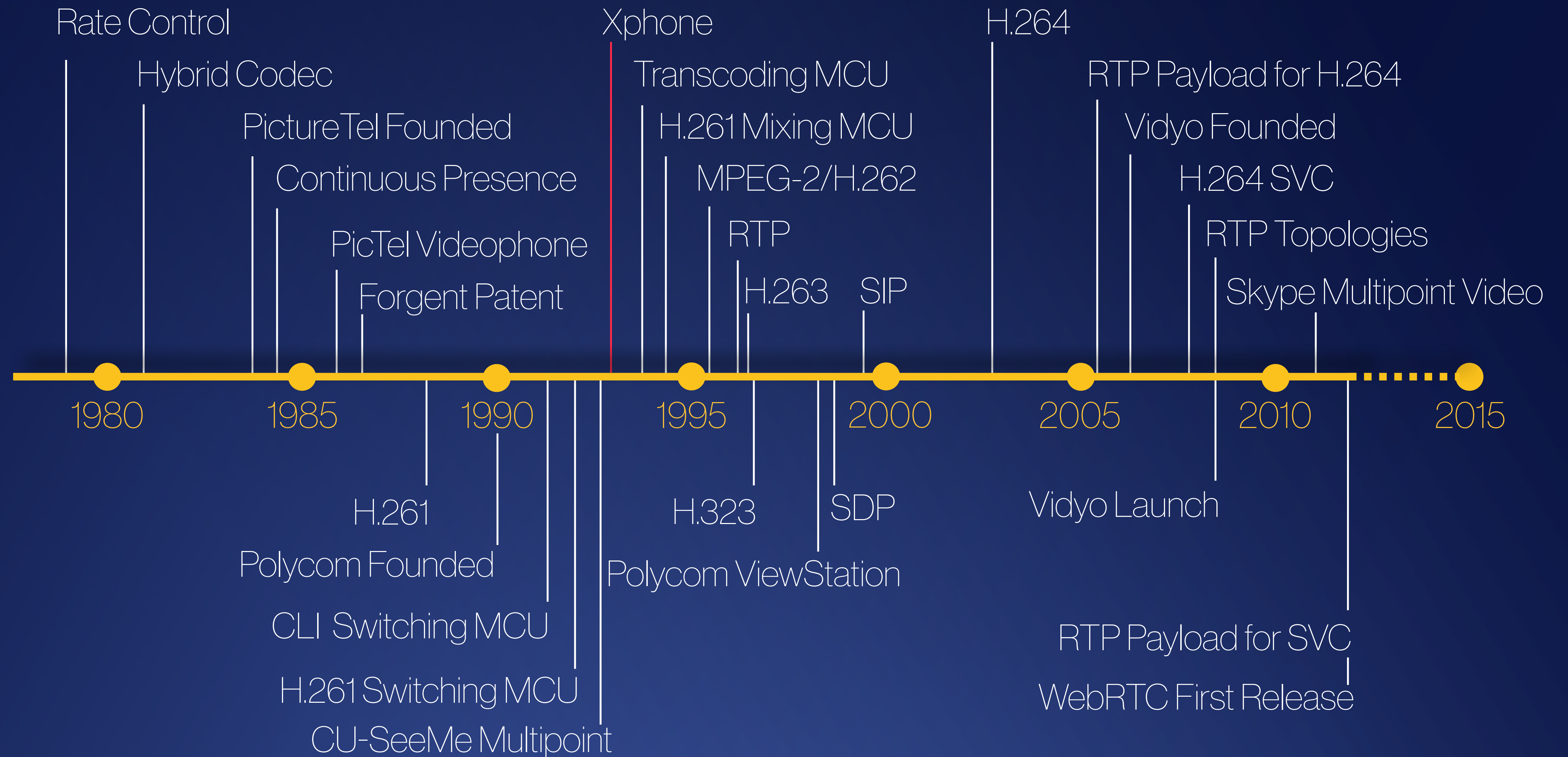
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

CU-SeeMe Multipoint

April 1993
CU-SeeMe v0.19 for Mac

Multipoint using “reflector”.
First multi-stream endpoint.





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Xphone

August 1993

Eleftheriadis, Pejhan, and
Anastassiou
1st ACM Multimedia Conf.

P2P audio+video using
motion-JPEG hardware on
Sun workstations



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Algorithms and Performance Evaluation of the Xphone Multimedia Communication System

Alexandros Eleftheriadis, Sassan Pejhan and Dimitris Anastassiou
Electrical Engineering Department, Columbia University

Abstract—We describe and evaluate the performance of the algorithms used in Columbia University's "Xphone" multimedia communication system. The system assumes a "best-effort" operating system and network, and provides synchronized video/audio acquisition/playback (locally or across a network) with minimized and bounded end-to-end delay. Synchronization is achieved using an algorithm based on time-stamps and device state information. The effects of jitter (delay variation) are mitigated using silence detection; the end-to-end delay is kept bounded using a restart mechanism. Finally, for live video sources, we describe a source bit-rate adaptation algorithm that maximizes the video image quality to the available network bandwidth and video display window size.

Keywords—Multimedia communication systems, media synchronization, source rate control, application development systems.

I. INTRODUCTION

One of the enabling technologies for multimedia systems is video compression algorithms. Recent advances in compression technology for images and video (JPEG, MPEG-1, MPEG-2) have resulted in bandwidth reductions of two orders of magnitude, down to 1–2 Mbit/sec. In addition, the work of international standardization organizations and the increased interest in video applications for computers and consumer electronics products have resulted in VLSI implementations of these algorithms which can be used for the development of real systems [1; 2; 3; 6; 13; 14; 19].

Video coding, however, is just one of the components of a multimedia system. The support of continuous, high-volume and real-time data (like video or audio) in both computers and networks represents a tremendous shift in design methodology, resulting in a re-evaluation of basic principles. Time dependency of information as a concept existed only in dedicated systems (e.g. the telephone network, or embedded systems); with multimedia, it becomes an issue for practically any application. The focal point of multimedia research is to provide bit-pipe characteristics (guaranteed bandwidth, low and constant delay, accurate synchronization) to packet-based systems, using algorithms and architectures that can be widely deployed.

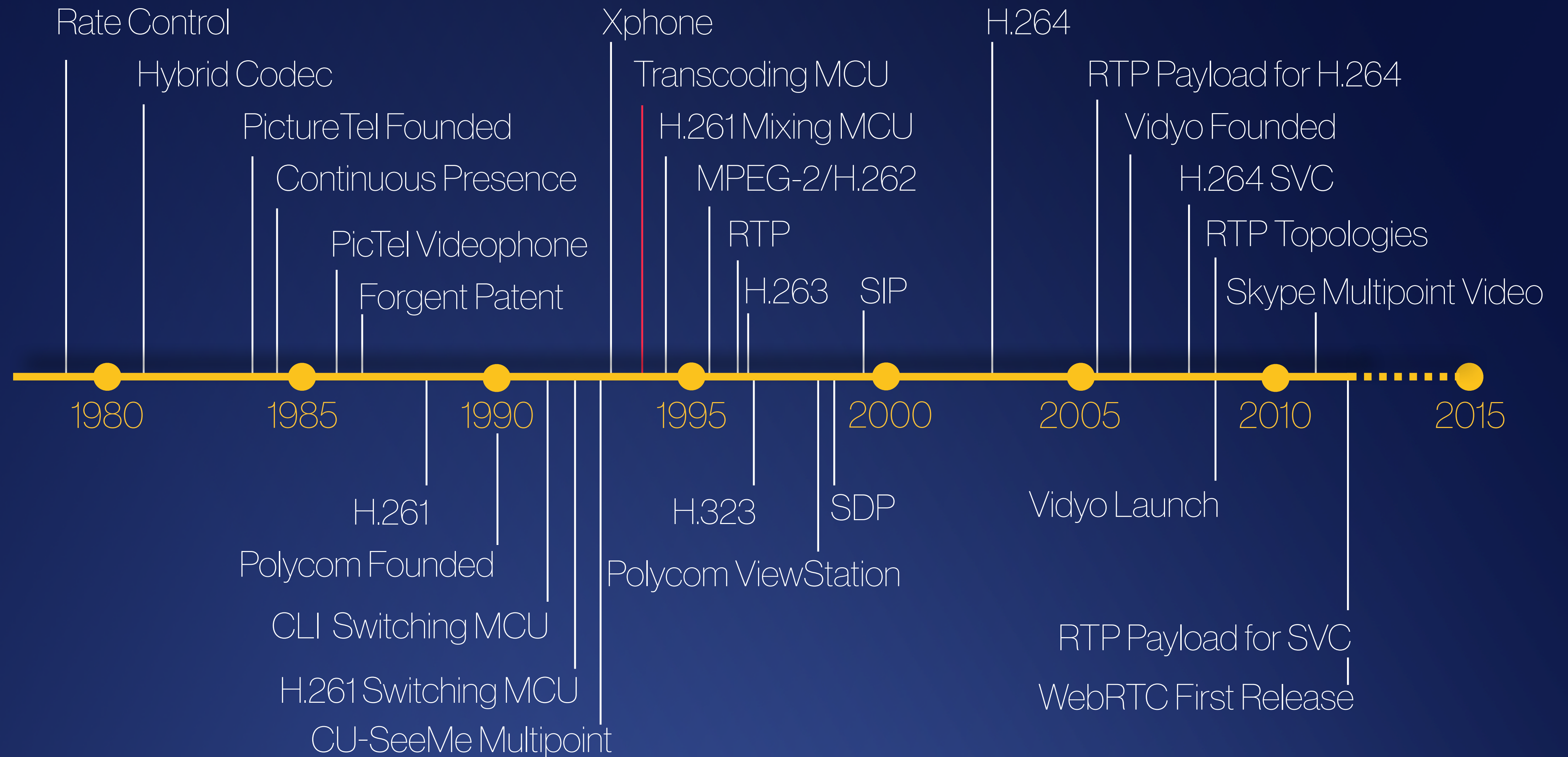
The availability of some kind of real-time support from the underlying operating system and network is important

Work supported in part by the New York State Center for Advanced Technology in Computers and Information Systems, and in part by the Center for Telecommunications Research at Columbia University. The authors are with the Electrical Engineering Department, Columbia University, New York, NY 10027 {elef, sassan, anastas}@ctr.columbia.edu.

for high-quality, wide-area multimedia communications, and is currently a very active area of research (see e.g. [9; 10; 20] and references therein). It is nevertheless possible to provide multimedia communication even in environments where delay uncertainty prevails (best-effort systems), albeit with some quality degradation. In addition, algorithms employed in non-real-time and real-time systems can be the same; although the latter will definitely perform better, the techniques used to achieve this performance can be similar (especially if the real-time support is not "hard"). Throughout this paper we assume the use of a best-effort operating system and network; in other words, no time-related guarantees are provided.

A number of systems and techniques have appeared in the literature, addressing various aspects of multimedia systems. Early efforts provided audio communication only [4]. Some systems use analog video and audio communication [5], with the corresponding self-evident limitations in terms of media integration in user applications. A significant volume of work has been reported at the system architecture level [8; 16; 18], describing the interface between applications and multimedia services and the latter's structure. In the area of media synchronization, a number of techniques have been proposed. These include incorporation of time constraints and scheduling of multimedia documents [7; 12; 15], media synchronization for database access applications (where a high end-to-end delay is acceptable) [8; 17], and synchronization for interactive multimedia communications [11]. In the first and second areas, the proposed techniques are basically used to derive time-stamps (or their equivalent) with no further analysis of how these time-stamps will be enforced; in addition, strong assumptions are usually made in terms of the performance of the underlying network and host equipment [17]. In the third area, which is more directly related to our work, the techniques described in [11] are only applicable in token ring networks (where the network access time is known prior to transmission) and require very tight coupling of software/hardware layers (the authors use their own operating system).

In this paper we describe the architecture and associated algorithms of the Xphone multimedia communications system, which has been developed to support the use of multimedia



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Transcoding MCU

January 1994

Willebeek-LeMair, Kandlur,
and Shae

*19th Conf. on Local Computer
Networks*

First reference to a
transcoding gateway

On Multipoint Control Units for Videoconferencing

M. H. Willebeek-LeMair, D. D. Kandlur, and Z.-Y. Shae

IBM T. J. Watson Research Center
Yorktown Heights, NY 10598, USA.
mwlml, kandlur, zshae@watson.ibm.com

Abstract

This paper examines the issues involved in the design of conference servers that support multiparty, multimedia conferences. These servers, called Multipoint Control Units (MCUs) in the telephony world, coordinate the distribution of audio, video, and data streams amongst the multiple participants in a videoconference. The MCU is responsible for the processing of video and audio so that a conference participant can hear and see one or more of the other participants in the conference. It is also responsible for handling and forwarding the data streams from the participants. This paper presents different approaches to the design of an MCU to implement these functions. It also describes the design of a related device - a transcoding gateway that enables conferencing between participants using different video/audio equipment.

1 Introduction

In recent years, with the emergence of improved communication technologies with wider coverage and accessibility, videoconferencing has become one of the major new growth applications. Videoconferencing standards are being developed and more and more videoconferencing products are appearing in the market.

Videoconferencing solutions are currently evolving from several directions. On the one side, there are the circuit-switched (e.g., Narrowband ISDN or the Switched-56Kbps phone lines) types of solutions, which being motivated by the telephony industry, can be likened to it. On the other side are the packet-based network (e.g., Ethernet and Token Ring legacy LANs)

solutions, which are designed to carry real-time traffic over existing computer communications networks. The advent of ATM [1] (Asynchronous Transfer Mode) might eventually allow these two approaches to converge.

Due to the stringent bandwidth, delay, and jitter requirements of real-time audio and video data, the solutions for the circuit switched and packet-based networks differ considerably. These differences include the encoder/decoder (CODEC) technology for video compression and decompression, the methods used to guarantee network performance, and in the provisions within the end-stations to handle the real-time traffic.

Videoconferences may be point-to-point or multipoint.

Point-to-Point. In a point-to-point videoconferencing a user is able to connect to only one other participant and communicate via video, audio, and shared data applications.

Multi-Point. A multi-point conference involves more than two participants and multimedia data is multicast from each participant to all others.

Each of the above scenarios involves the integrated communication of video, audio, graphics, and text data.

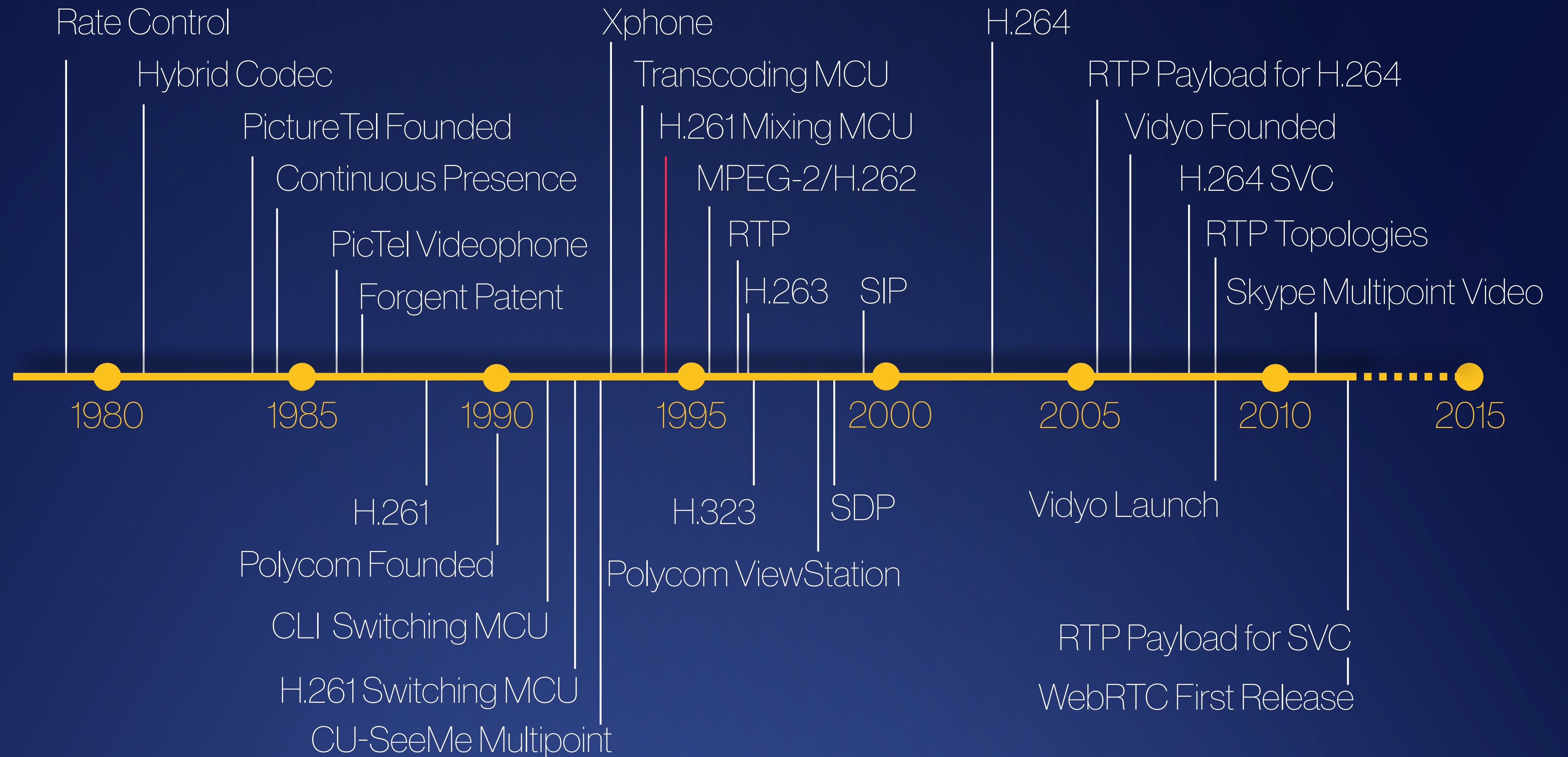
Approaches used to support multipoint conferences may be categorized as either distributed or centralized. In a distributed approach each end-station receives the video and audio streams from all, or some, of the participating end-station sources in the conference. Each end-station then composes these multiple incoming streams as desired. This approach is advantageous since it allows more flexibility and control at each end-station and minimizes the distance that streams need to travel between source and destination. It requires additional processing capability



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Transcoding MCU





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

H.261 Mixing MCU

August 1994

Lei, Chen, and Sun
IEEE Trans. on CSVT

H.261 “mixing” bridge
(Bellcore)

Video Bridging Based on H.261 Standard

Shaw-Min Lei, Member, IEEE, Ting-Chung Chen, Senior Member, IEEE, and Ming-Ting Sun, Senior Member, IEEE

Abstract— Multi-point ISDN videoconferencing with video bridging in network-based servers represents a viable new network service. This paper presents a detailed technical analysis of a continuous presence video bridge using the H.261 video coding standard. We first compare the pros and cons of coded domain versus pel domain video bridges. The architecture and the required operations of a coded domain bridge using H.261 are then investigated. We derive the bounds of the bridge delay and the required buffer size for the implementation of the bridge. The delay and the buffer occupancy of the video bridge depend on the order, complexity, and the bit-distribution of the input video sources. To investigate a typical case, we simulate the delay and the buffer occupancy of a video bridge. We also provide a heuristic method to estimate the delay in a typical case. Several techniques are discussed to minimize the bridge delay and the buffer size. Finally, we simulate intra slice coding and show that the delay and the buffer size can be reduced significantly using this technique.

I. INTRODUCTION

MULTI-point videoconferencing is a natural evolution of two-point videoconferencing. To establish multi-point connections, coded video streams from the participants are sent to a Multi-point Control Unit (MCU). In a “switched presence” MCU, either a signal selected by the conference chairman or a signal selected based on audio channel activity is broadcast to all participants [1]. In a “continuous presence” MCU, multiple video streams from the participants are combined so that each participant can see the selected multiple participants all the time. The “switched presence” MCU has been standardized recently [2], [3] while the “continuous presence” MCU is still under active research. In this paper, we focus on video combining in the “continuous presence” MCU to support network-based multi-point multimedia videoconferencing.

In the MCU, multiple coded video sources can be decoded, combined in the pixel domain, and then encoded for distribution. In some special cases, however, the video sources can be combined directly in the coded domain without transcoding. Video combining in the coded domain offers shorter end-to-end delay, better picture quality, and lower MCU cost. It happens that the current videophone/videoconferencing standard H.261 [4] is able to provide such video combining without any transcoding. The MCU can combine four QCIF (Quarter Common Intermediate Format) videos into one CIF video in the coded domain for continuous presence multipoint videoconferencing. Each user terminal is operated in an asymmetric mode which transmits QCIF pictures but receives CIF pictures at four times of transmission bit rate. A QCIF video combiner

will provide users a continuous presence view of up to four conferees at one time. Although the concept of the QCIF combiner is simple, many implementation issues need to be studied in more detail. The issues that need to be answered include the combiner architecture, required size of buffers, end-to-end delay, picture quality, and system complexity.

Two examples of continuous presence applications are shown in Fig. 1. In Fig. 1(a), conferees are involved in a multi-way videoconference. QCIF videos from the conferees are transmitted to the MCU with a transmission rate R . For each conferee, the MCU combines the selected four QCIF videos into a CIF video and transmits it back to the conferee at a transmission rate of $4R$. Fig. 1(b) is a remote classroom application where the MCU combines 4 QCIF remote-site videos into a CIF video so that the teacher can interact with 4 remote sites simultaneously. The video from the teacher is broadcast to all the remote students.

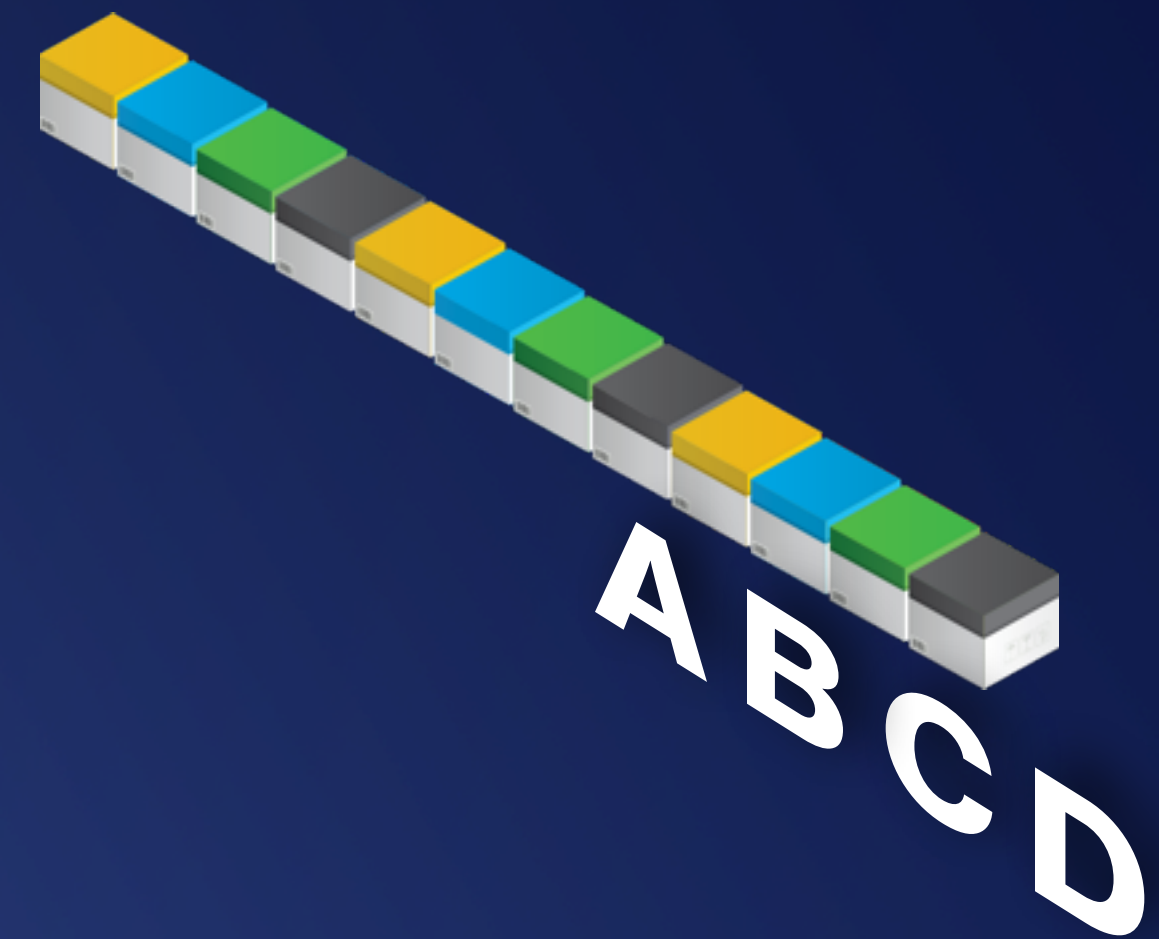
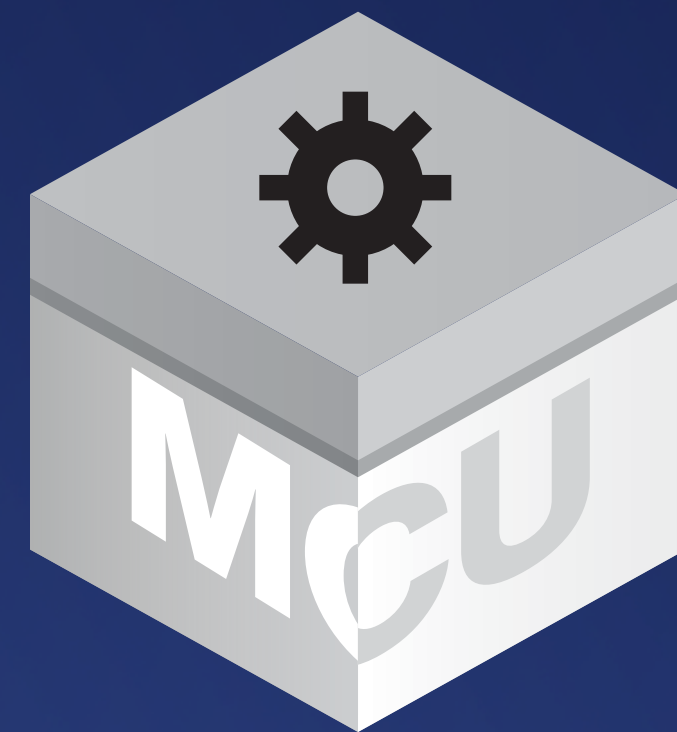
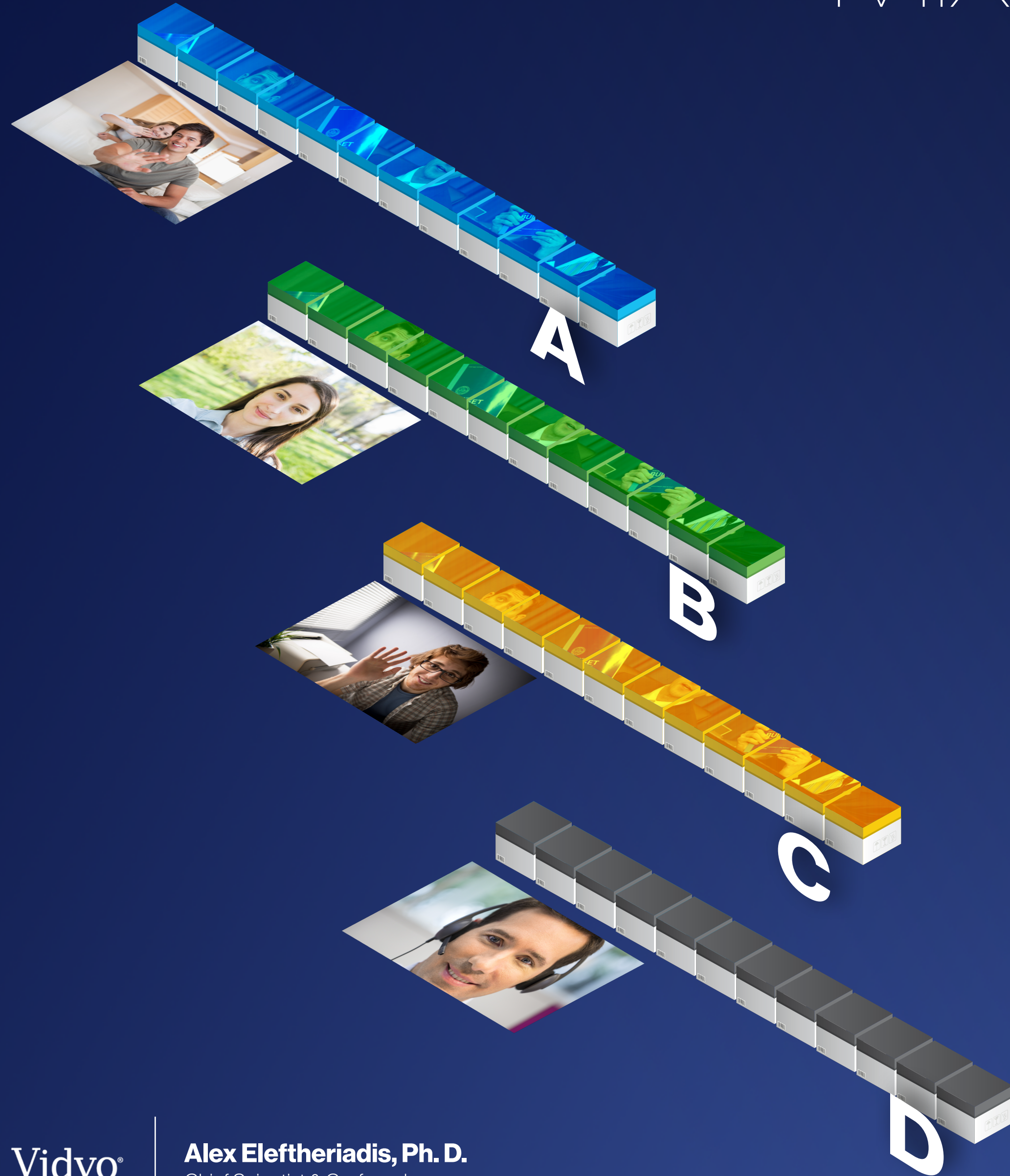
The organization of this paper is as follows. Section II describes and compares pixel-domain and coded-domain combining. The remaining sections are focused on the coded-domain QCIF combiner. Section III describes the architectures of the QCIF combiner. Section IV presents a theoretical analysis for the delay and buffer size. Section V shows some simulation results. Section VI discusses techniques to improve the end-to-end delay. Finally, conclusions are provided in Section VII.

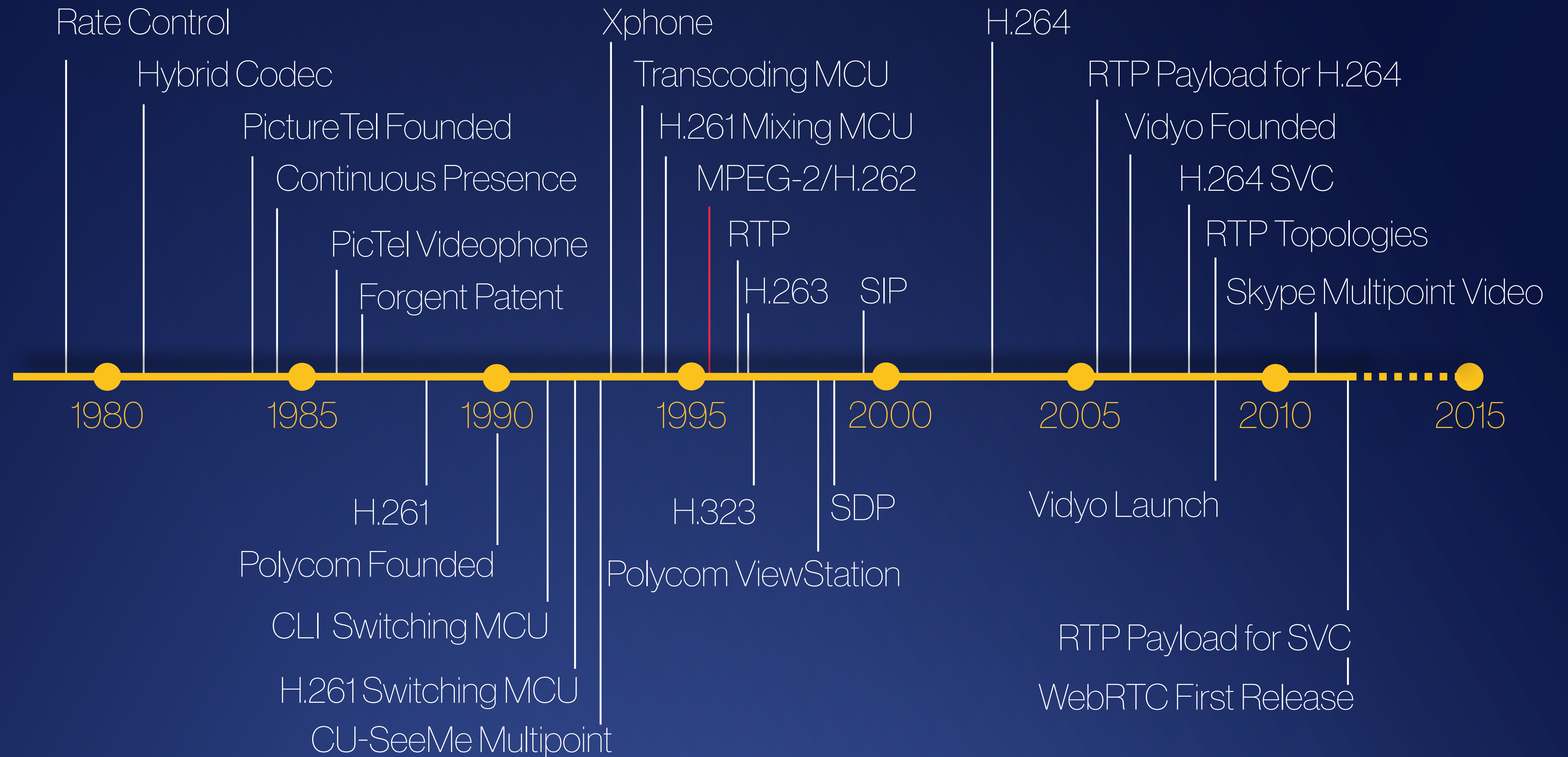
II. CODED-DOMAIN VS. PEL-DOMAIN VIDEO COMBINING

Since the raw video data rate is large and the network bandwidth is limited, video data usually have to be compressed for network transmission. Therefore, video bridging in the network would normally require video decoding, combining in the pel domain, and encoding for retransmission. For some coded bit streams, however, video combining can be done in the coded domain. We define coded-domain video combining as a process that does not decode the compressed data down to the pel domain. The compressed data are either not decoded at all or only partially decoded for combining. In the undecoded case, e.g., the proposed QCIF combiner, the video bridge only has to process data headers and concatenate the remaining data stream without modification. In this case, we are mainly dealing with data coded as the variable-length codes (VLCs). We will refer to this case by *VLC-domain approach*. In the partially decoded case, video bridging is usually done after decoding variable-length codes, e.g., DCT-domain video compositing [5]. For pel domain video combining, the compressed data have to be fully decoded and then, after combining in the pel domain, encoded again. Table I gives a qualitative summary of the pros and cons of these

Manuscript received June 22, 1993; revised November 5, 1993 and January 26, 1994. This paper was recommended by King N. Ngan.
The authors are with Bellcore, Red Bank, NJ 07701-5699 USA.
IEEE Log Number 9404692.

Mixing MCU



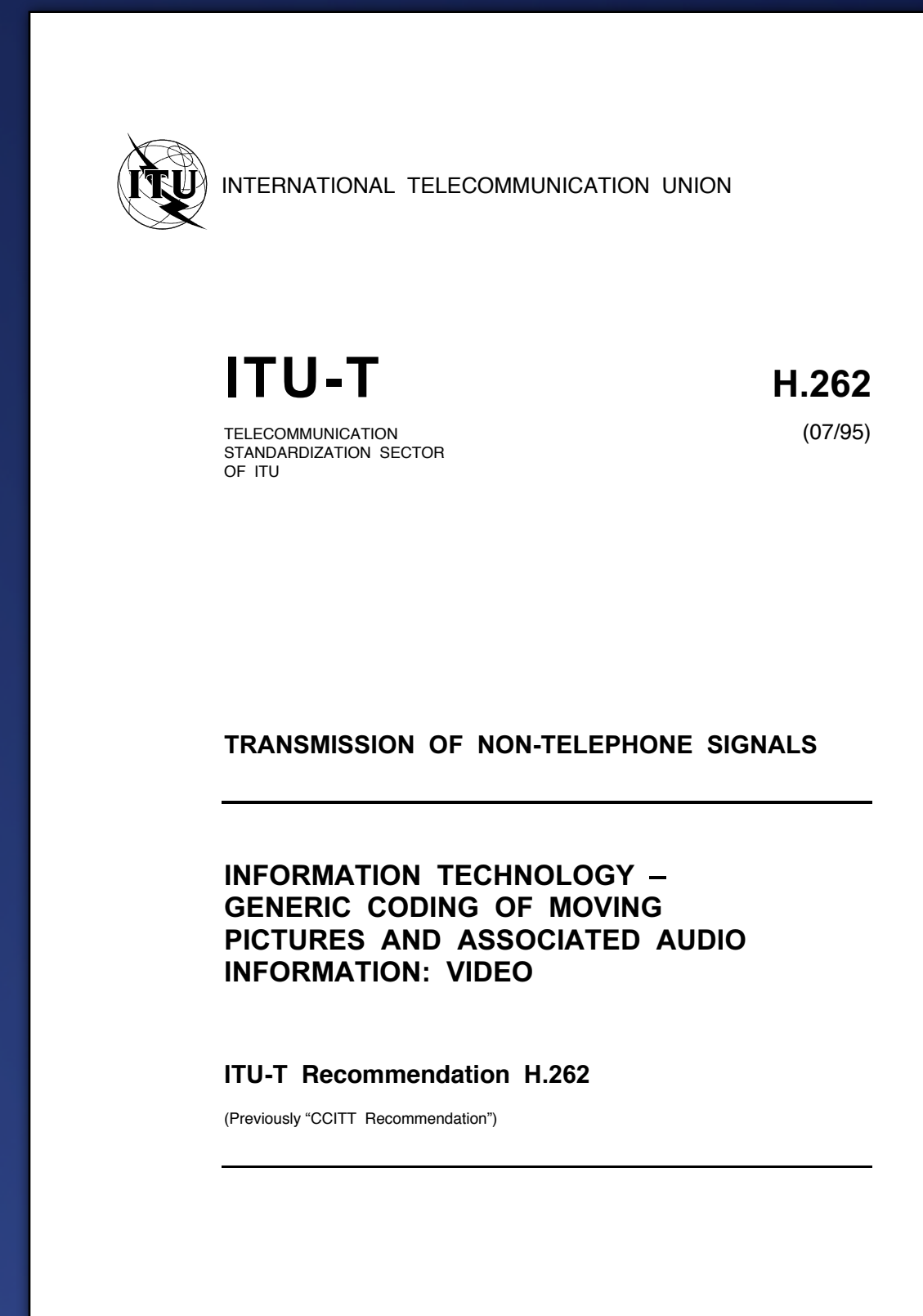


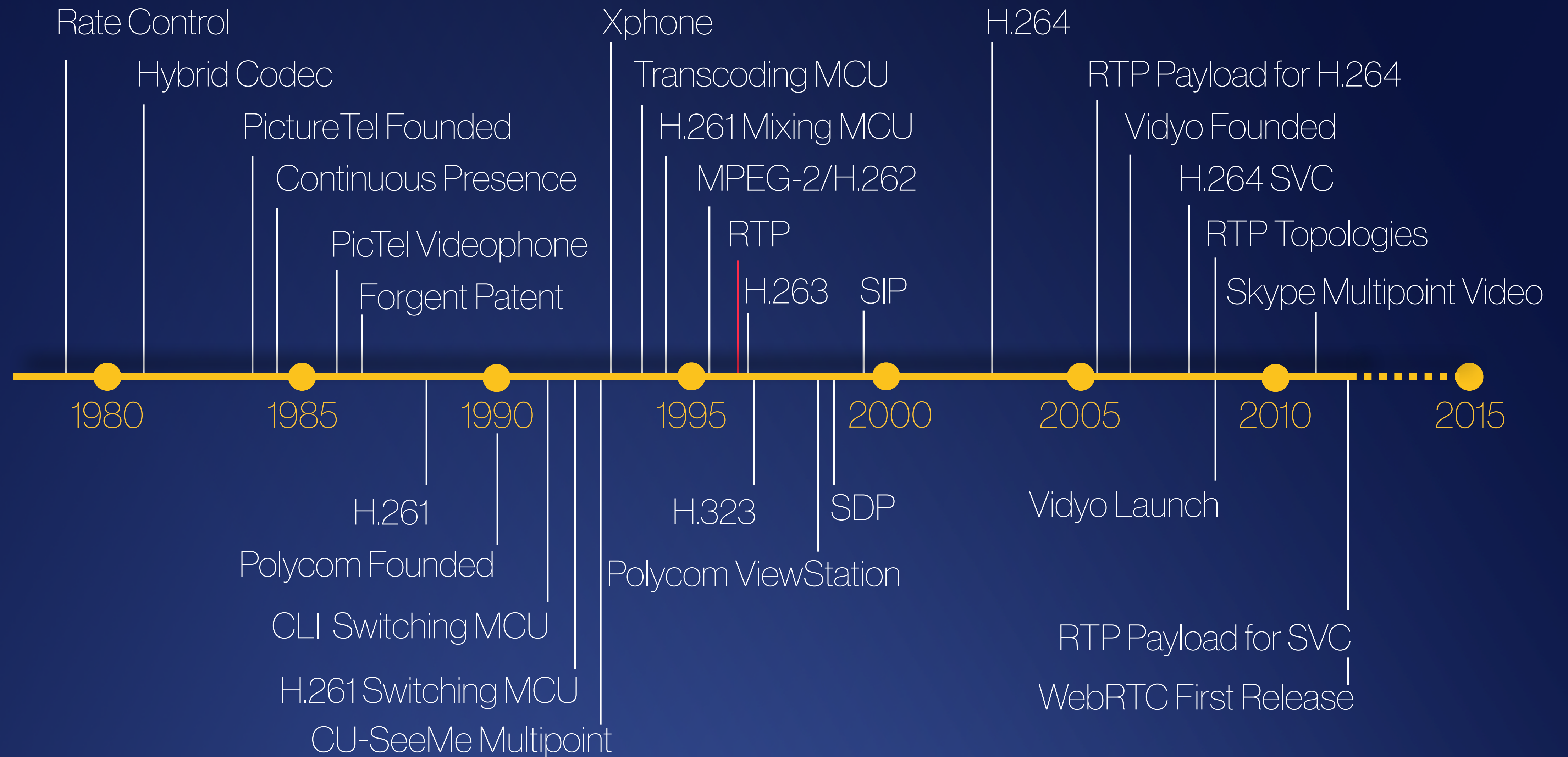
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

MPEG-2/H.262

July 1995

First digital video compression standard for broadcast, cable, and satellite TV, and DVD



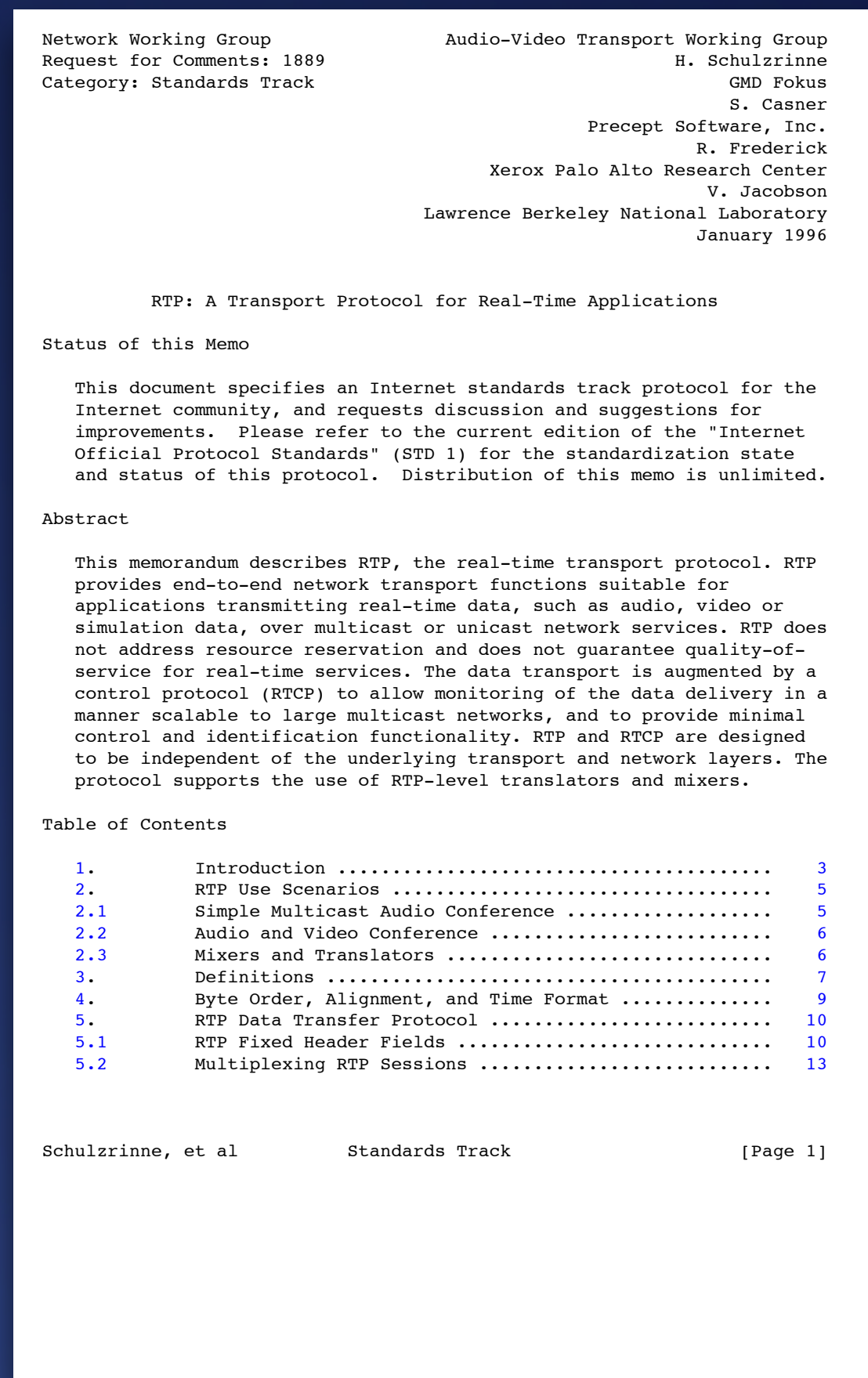


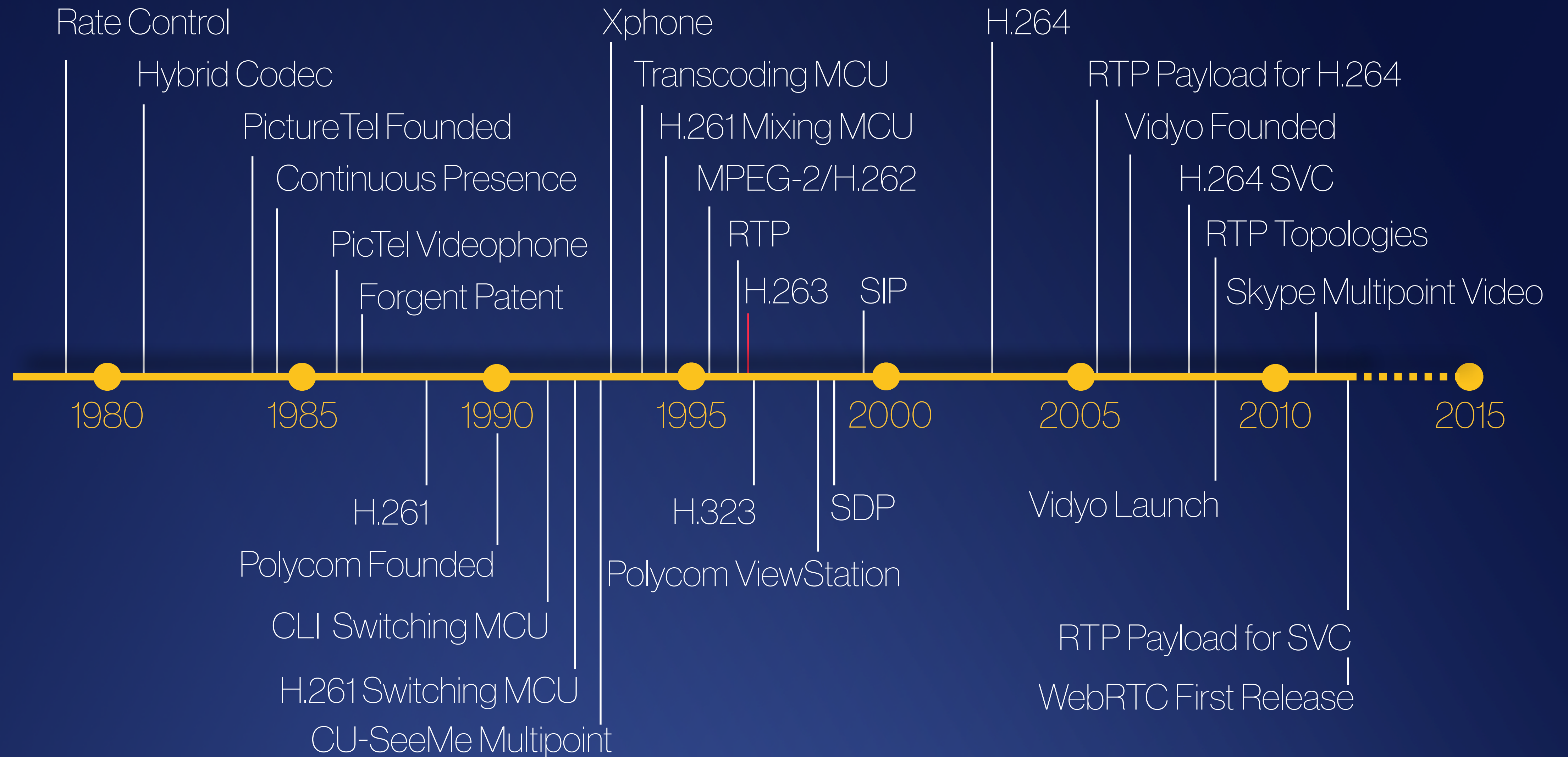
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

RTP

January 1996
Schulzrinne, Casner, Frederick,
and Jacobson
RFC 1889

Original RTP version



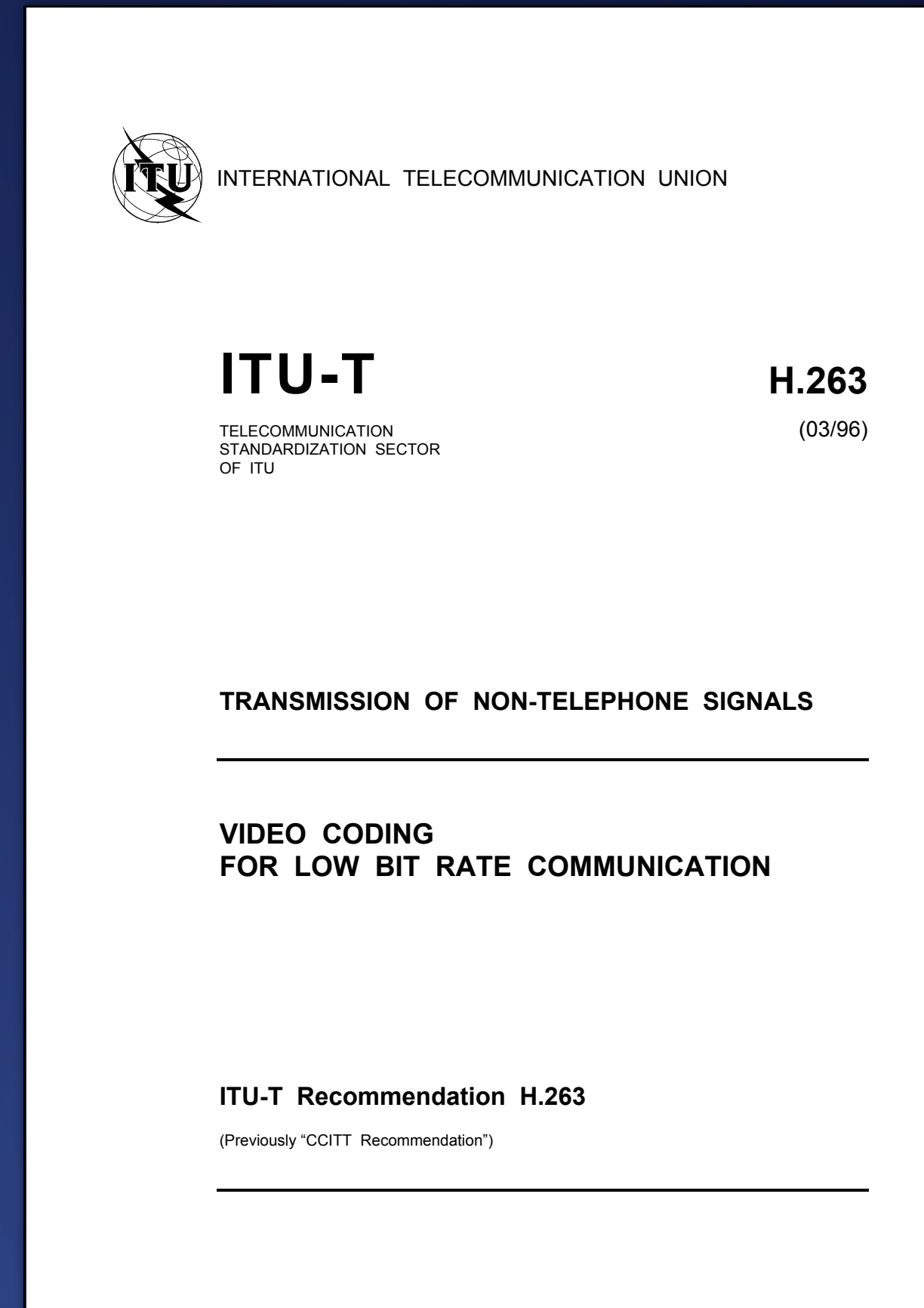


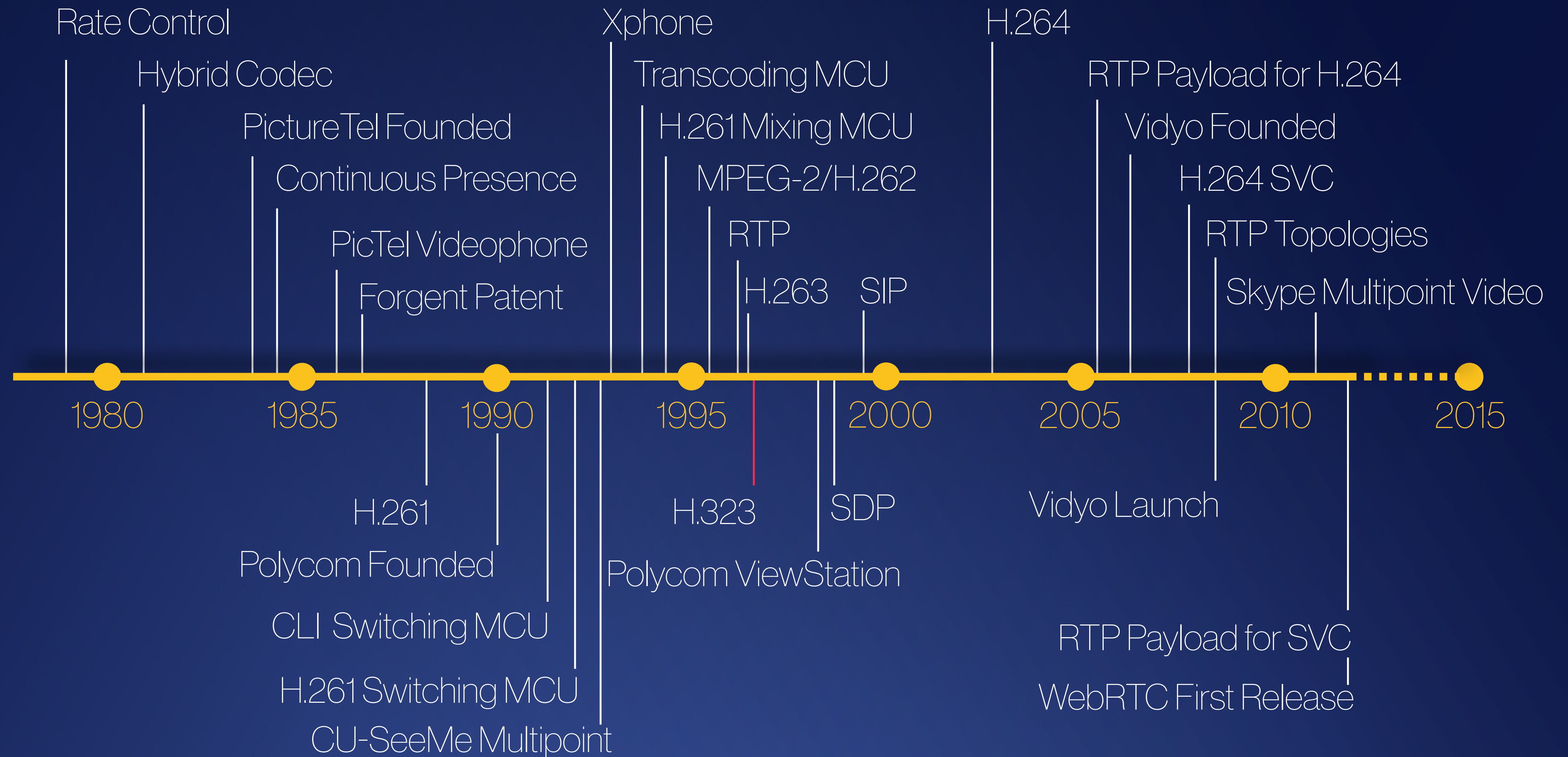
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

H.263

March 1996

2nd generation video compression
standard for communications






Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

H.323

November 1996

Packet-based multimedia
communication systems

Superseded by a more recent version

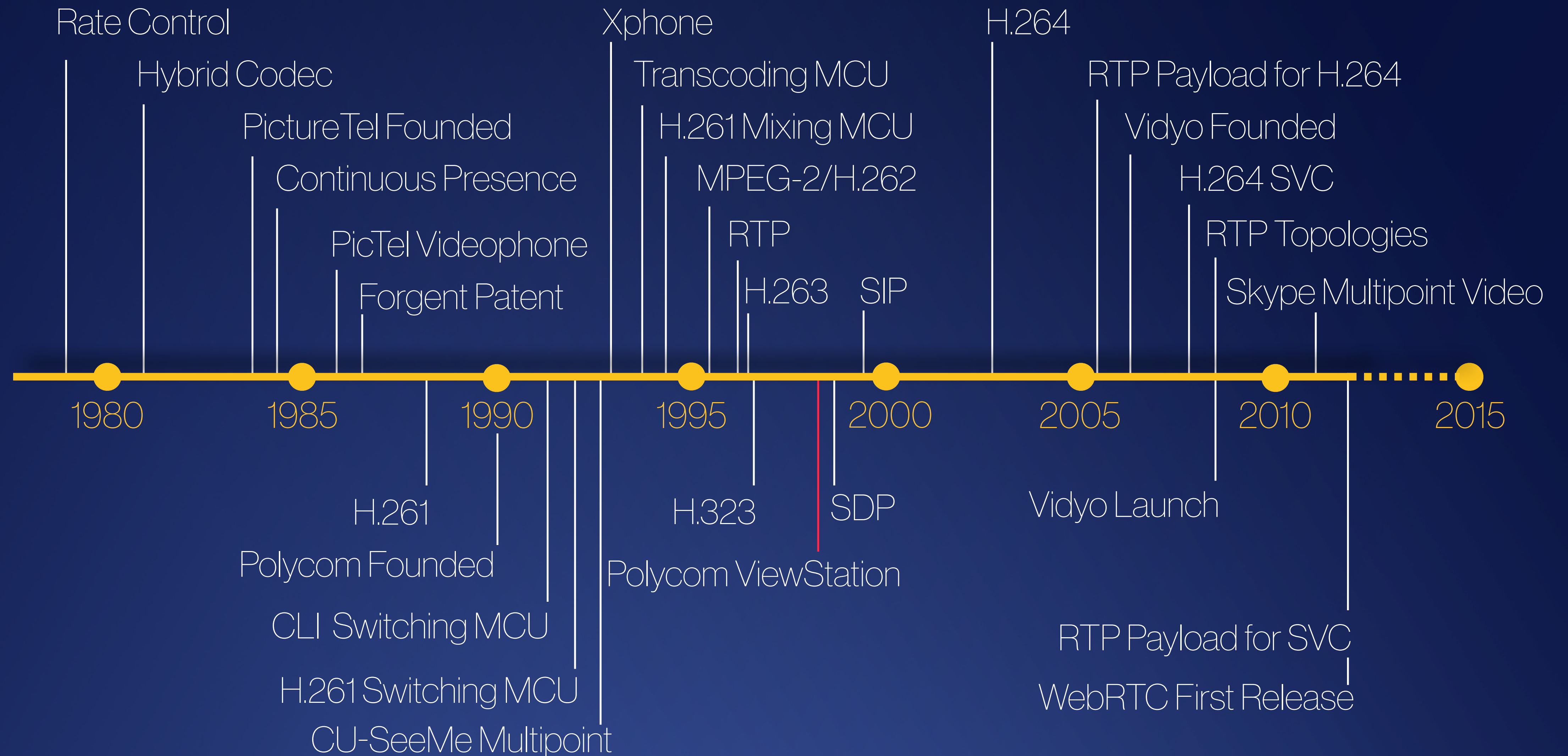
 INTERNATIONAL TELECOMMUNICATION UNION

ITU-T **H.323**
TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU (11/96)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS
Infrastructure of audiovisual services – Systems and terminal equipment for audiovisual services

Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service

ITU-T Recommendation H.323
Superseded by a more recent version
(Previously CCITT Recommendation)



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Polycom ViewStation

January 1998

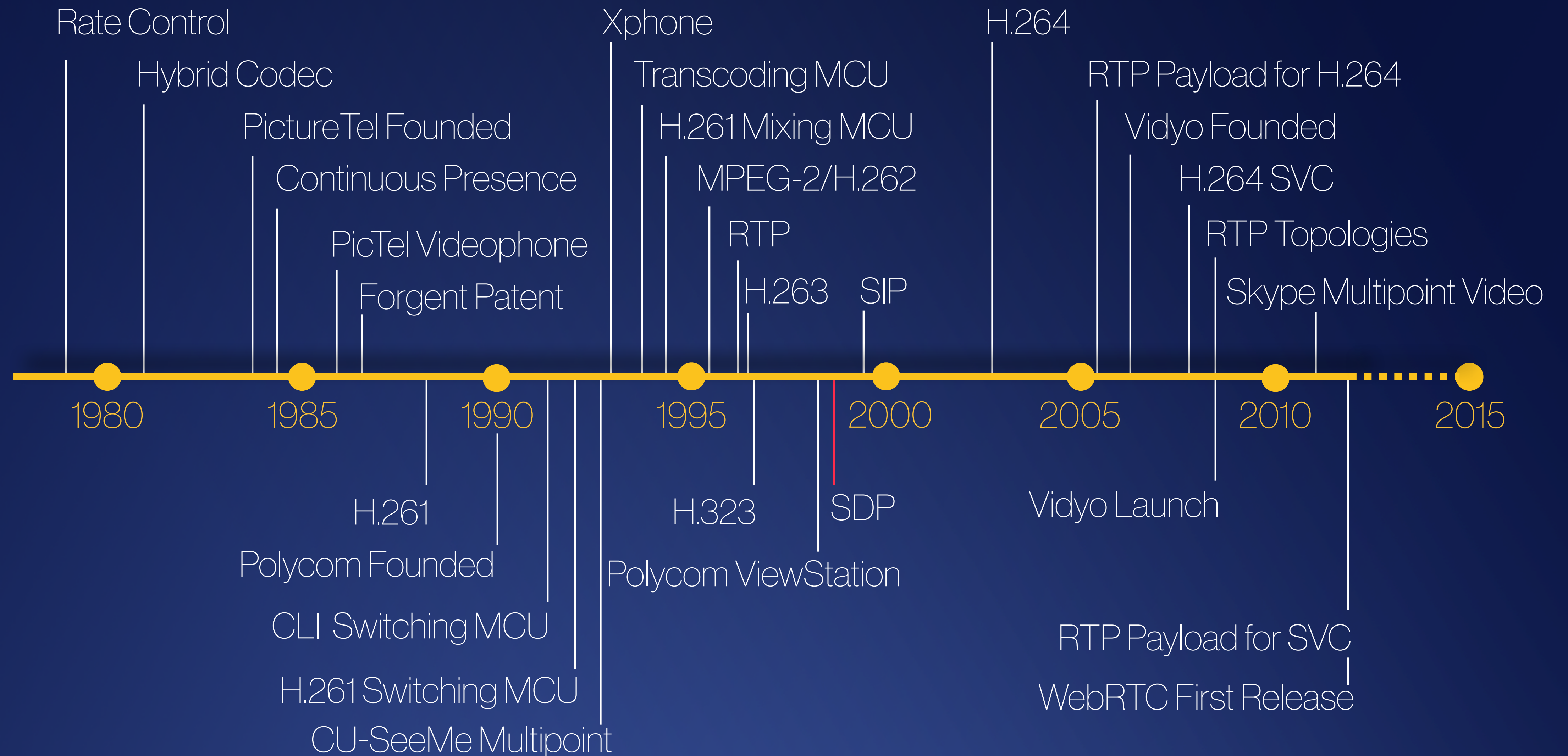
First Polycom video product

H.323 over IP

H.261 and H.263 video

4-way embedded multipoint
(auto, discussion/continuous
presence, presentation, active
speaker switching)





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

SDP

April 1998
Handley and Jacobson
RFC 2327
Session Description Protocol

Network Working Group
Request for Comments: 2327
Category: Standards Track

M. Handley
V. Jacobson
ISI/LBNL
April 1998

SDP: Session Description Protocol

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1998). All Rights Reserved.

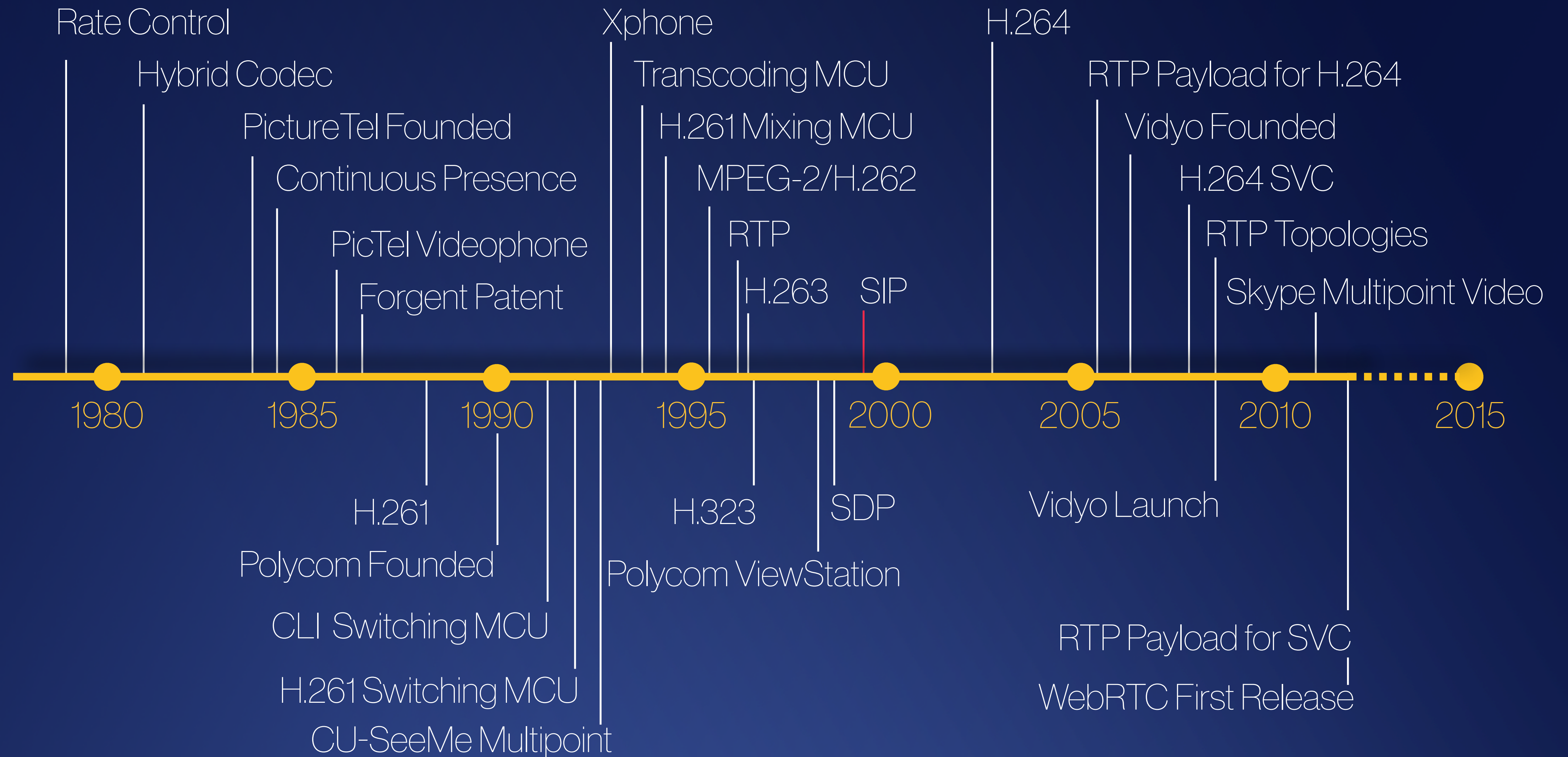
Abstract

This document defines the Session Description Protocol, SDP. SDP is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation.

This document is a product of the Multiparty Multimedia Session Control (MMUSIC) working group of the Internet Engineering Task Force. Comments are solicited and should be addressed to the working group's mailing list at confctrl@isi.edu and/or the authors.

1. Introduction

On the Internet multicast backbone (Mbone), a session directory tool is used to advertise multimedia conferences and communicate the conference addresses and conference tool-specific information necessary for participation. This document defines a session description protocol for this purpose, and for general real-time multimedia session description purposes. This memo does not describe multicast address allocation or the distribution of SDP messages in detail. These are described in accompanying memos. SDP is not intended for negotiation of media encodings.



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

SIP

March 1999
Schulzrinne, Handley, Schooler,
and Rosenberg
RFC 2543

Session Initiation Protocol

Network Working Group
Request for Comments: 2543
Category: Standards Track

M. Handley
ACIRI
H. Schulzrinne
Columbia U.
E. Schooler
Cal Tech
J. Rosenberg
Bell Labs
March 1999

SIP: Session Initiation Protocol

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1999). All Rights Reserved.

IESG Note

The IESG intends to charter, in the near future, one or more working groups to produce standards for "name lookup", where such names would include electronic mail addresses and telephone numbers, and the result of such a lookup would be a list of attributes and characteristics of the user or terminal associated with the name. Groups which are in need of a "name lookup" protocol should follow the development of these new working groups rather than using SIP for this function. In addition it is anticipated that SIP will migrate towards using such protocols, and SIP implementors are advised to monitor these efforts.

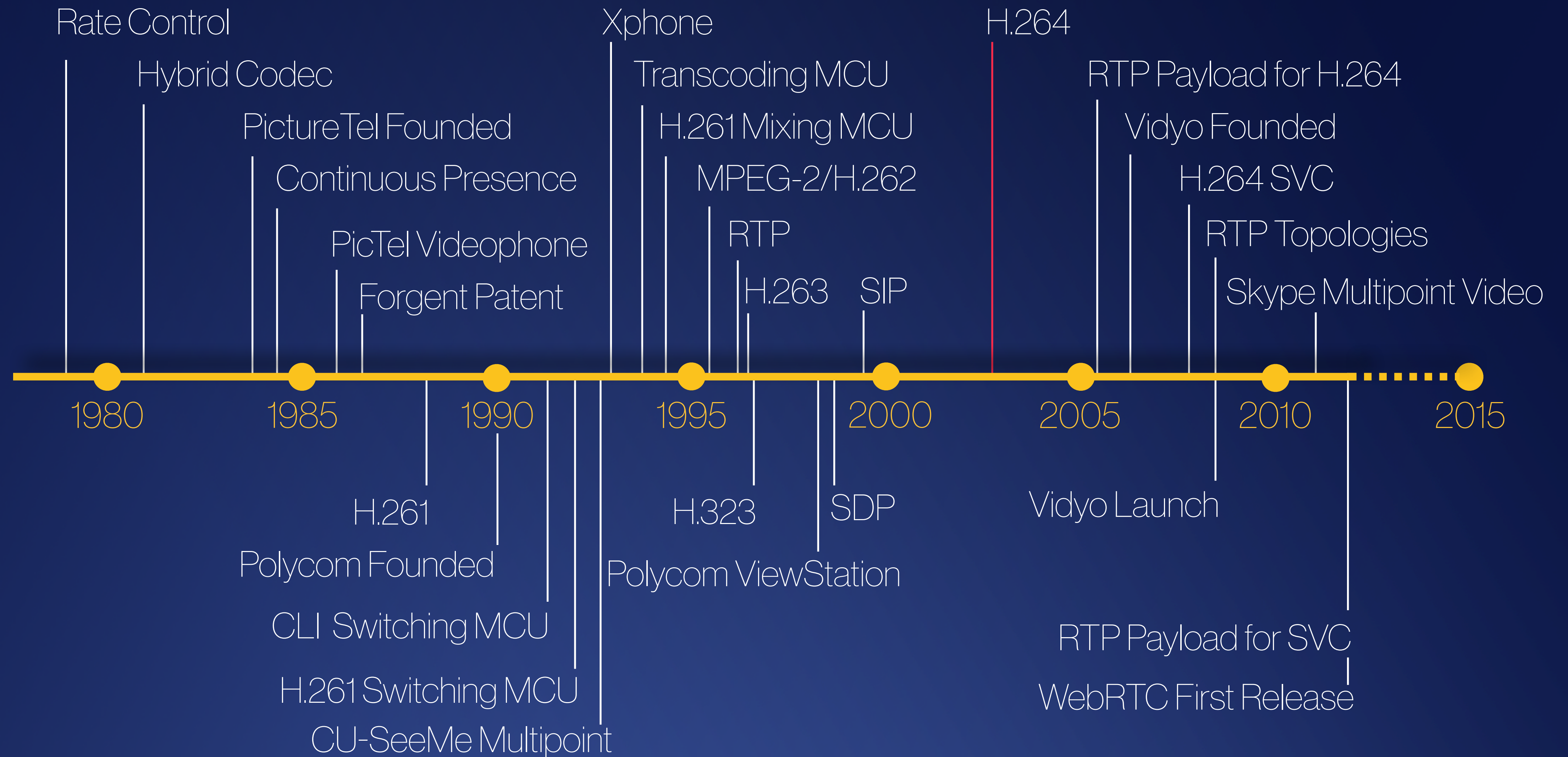
Abstract

The Session Initiation Protocol (SIP) is an application-layer control (signaling) protocol for creating, modifying and terminating sessions with one or more participants. These sessions include Internet multimedia conferences, Internet telephone calls and multimedia distribution. Members in a session can communicate via multicast or via a mesh of unicast relations, or a combination of these.

Handley, et al.

Standards Track

[Page 1]



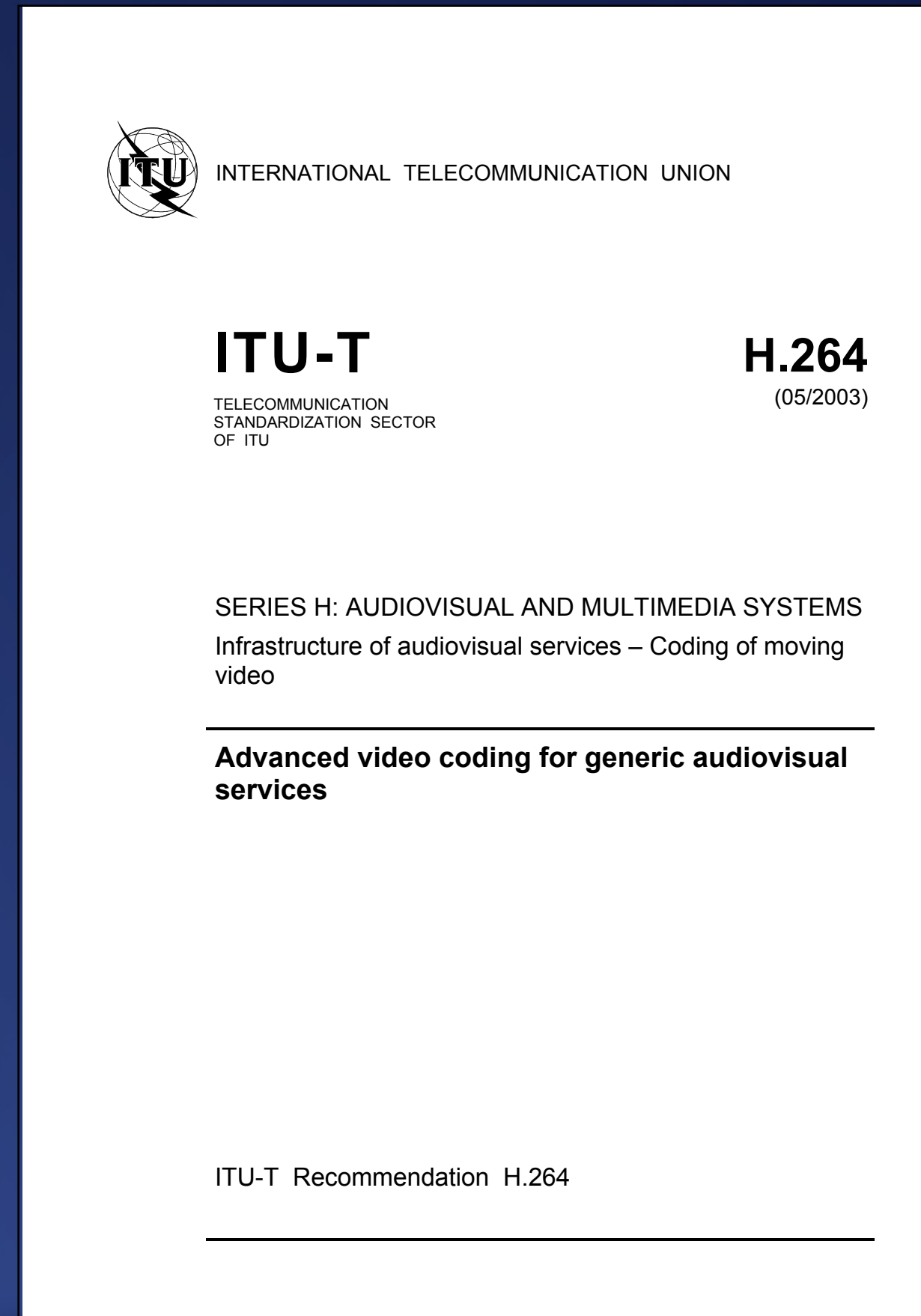
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

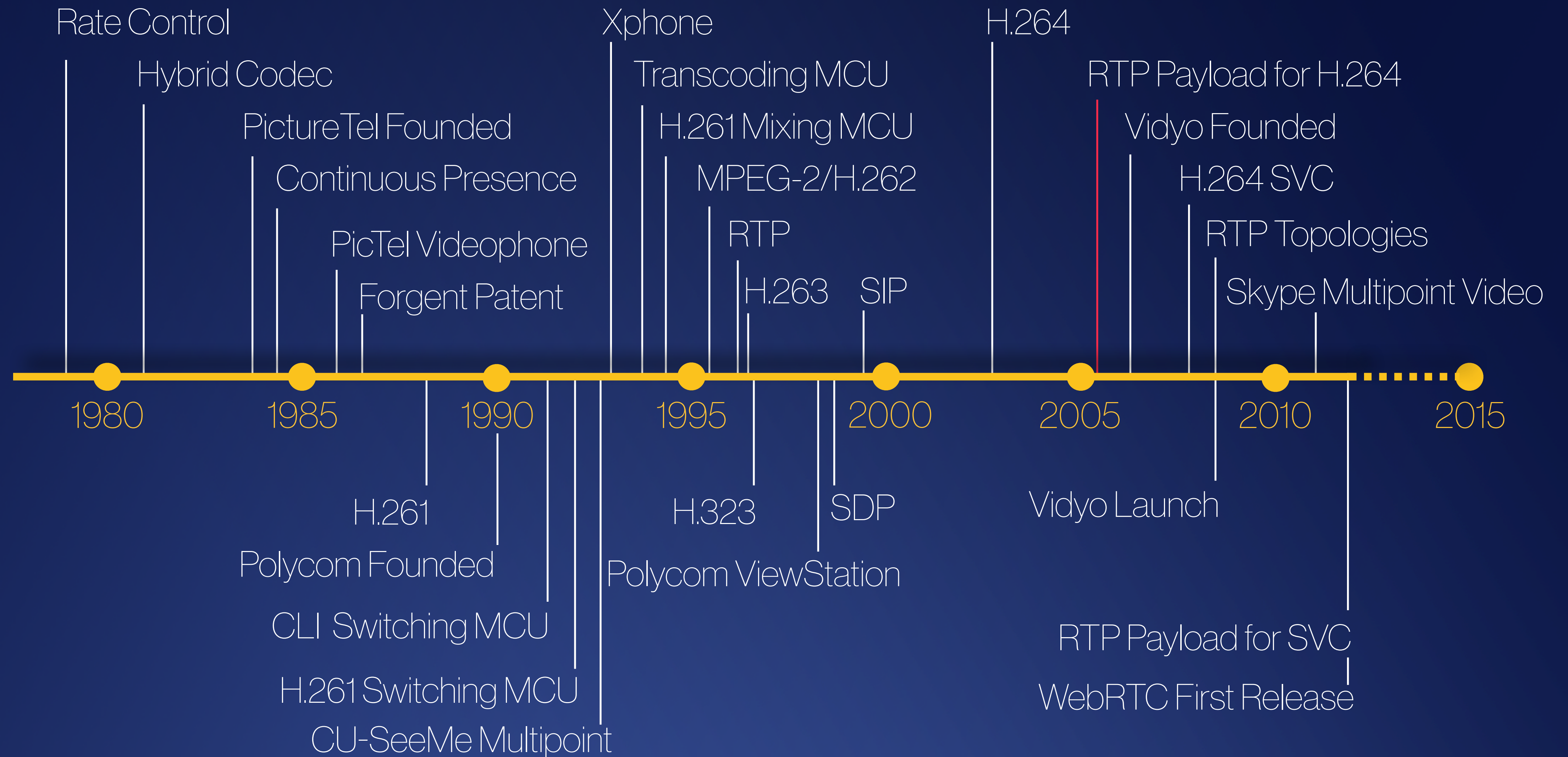
H.264

May 2003

H.264/MPEG-4 Part 10 AVC

The ubiquitous video codec





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

RTP Payload for H.264

February 2005
Wenger, Stockhammer,
Westerlund, and Singer
RFC 3984

How to packetize H.264 video for
RTP transport,
plus signaling parameters

Network Working Group
Request for Comments: 3984
Category: Standards Track

S. Wenger
M.M. Hannuksela
T. Stockhammer
M. Westerlund
D. Singer
February 2005

RTP Payload Format for H.264 Video

Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2005).

Abstract

This memo describes an RTP Payload format for the ITU-T Recommendation H.264 video codec and the technically identical ISO/IEC International Standard 14496-10 video codec. The RTP payload format allows for packetization of one or more Network Abstraction Layer Units (NALUs), produced by an H.264 video encoder, in each RTP payload. The payload format has wide applicability, as it supports applications from simple low bit-rate conversational usage, to Internet video streaming with interleaved transmission, to high bit-rate video-on-demand.

Table of Contents

1. Introduction.....	3
1.1. The H.264 Codec.....	3
1.2. Parameter Set Concept.....	4
1.3. Network Abstraction Layer Unit Types.....	5
2. Conventions.....	6
3. Scope.....	6
4. Definitions and Abbreviations.....	6
4.1. Definitions.....	6
5. RTP Payload Format.....	8
5.1. RTP Header Usage.....	8
5.2. Common Structure of the RTP Payload Format.....	11
5.3. NAL Unit Octet Usage.....	12

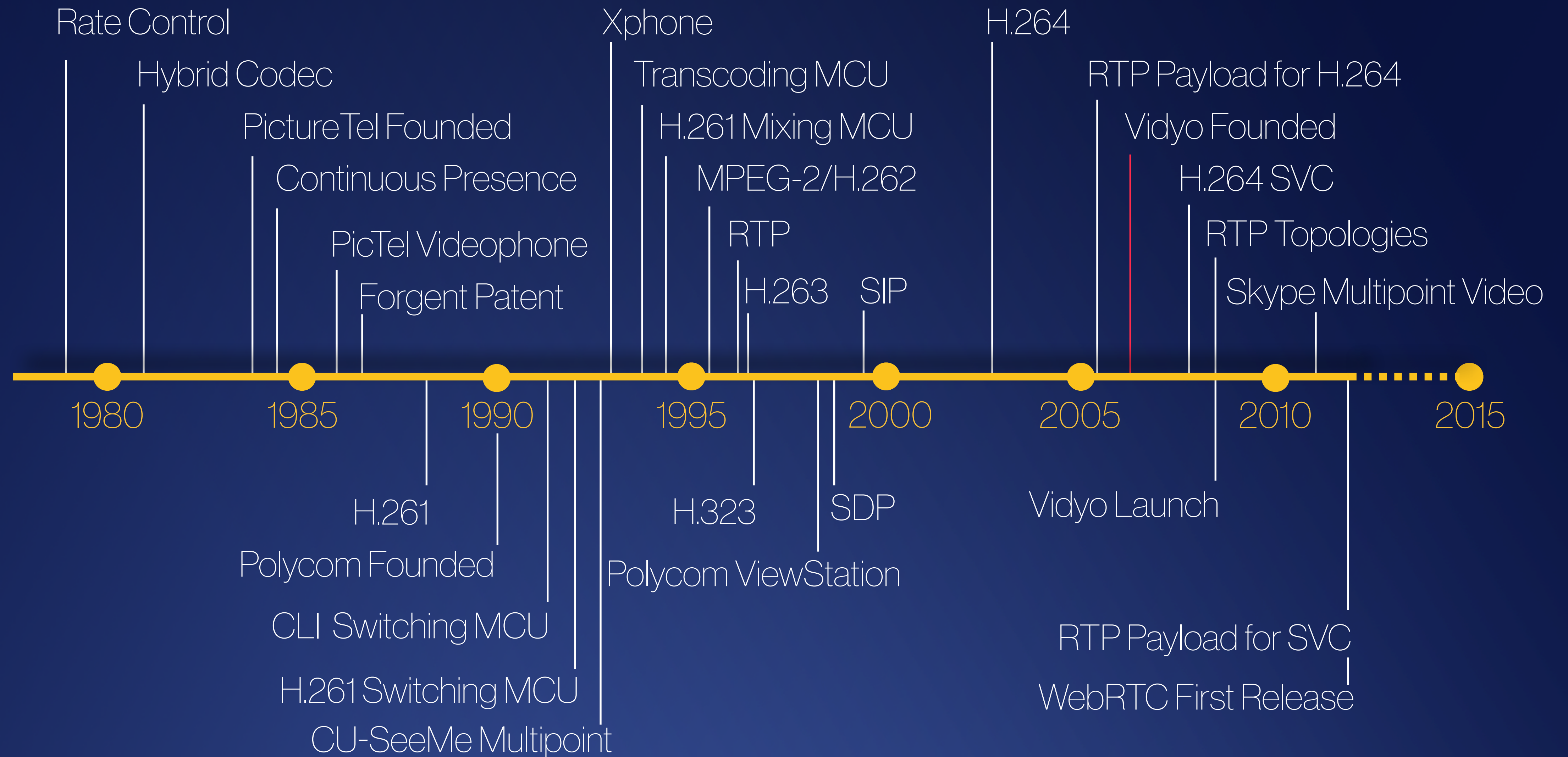
Wenger, et al.

Standards Track

[Page 1]



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

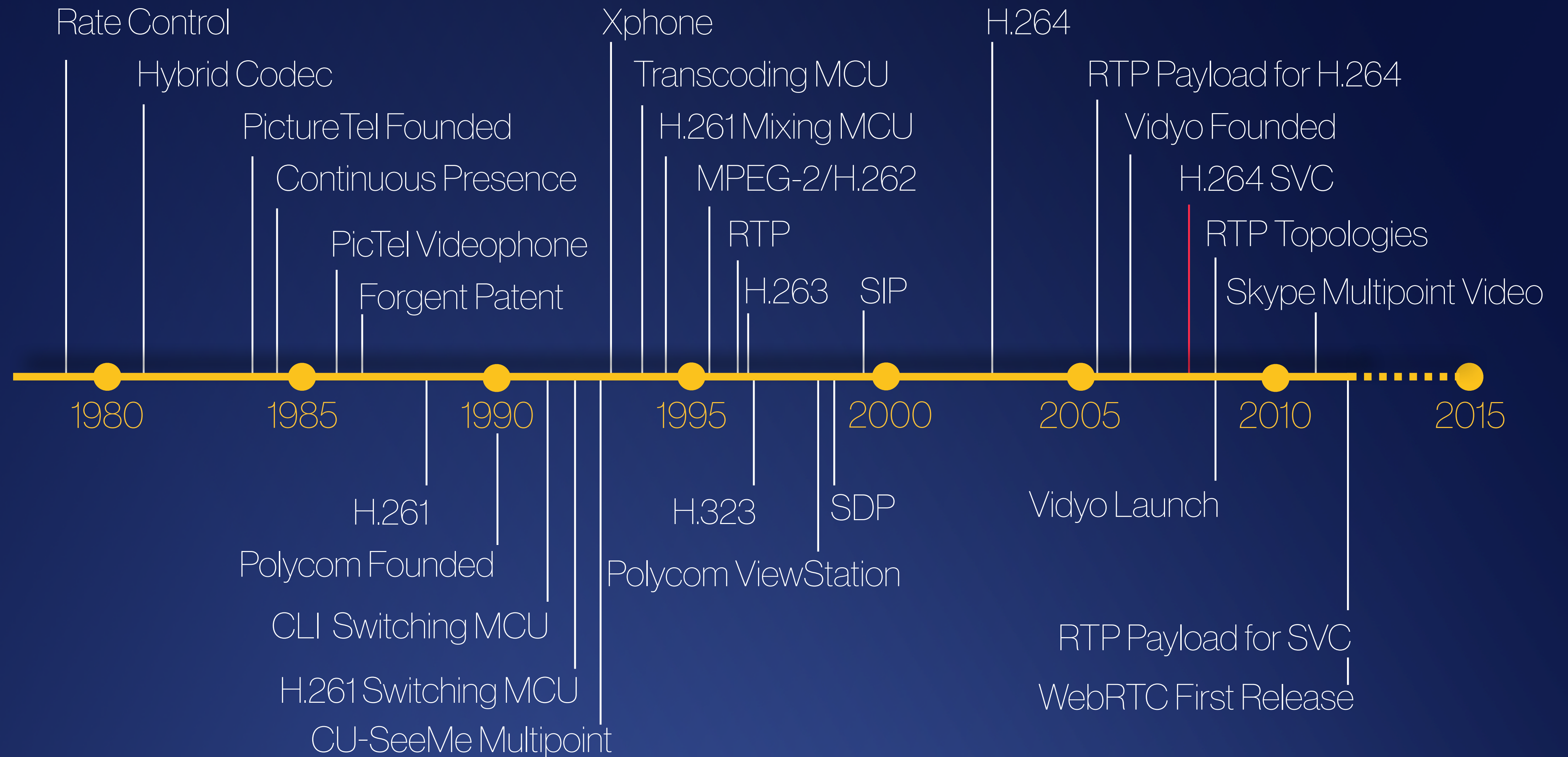
Vidyo Founded

April 2005
(as “Layered Media”)



Vidyo®

Shapiro, More, and Eleftheriadis



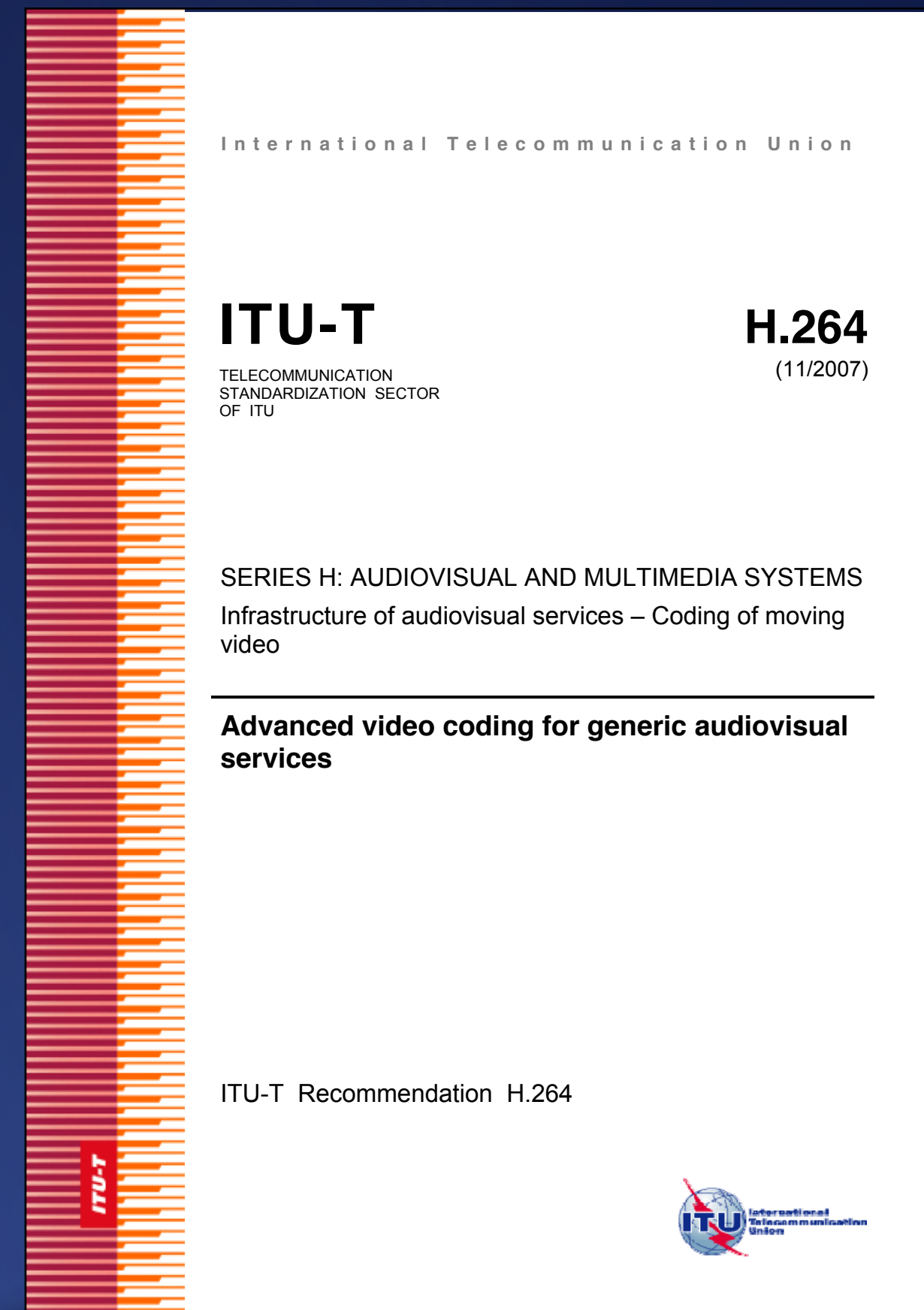
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

H.264 SVC

November 2007

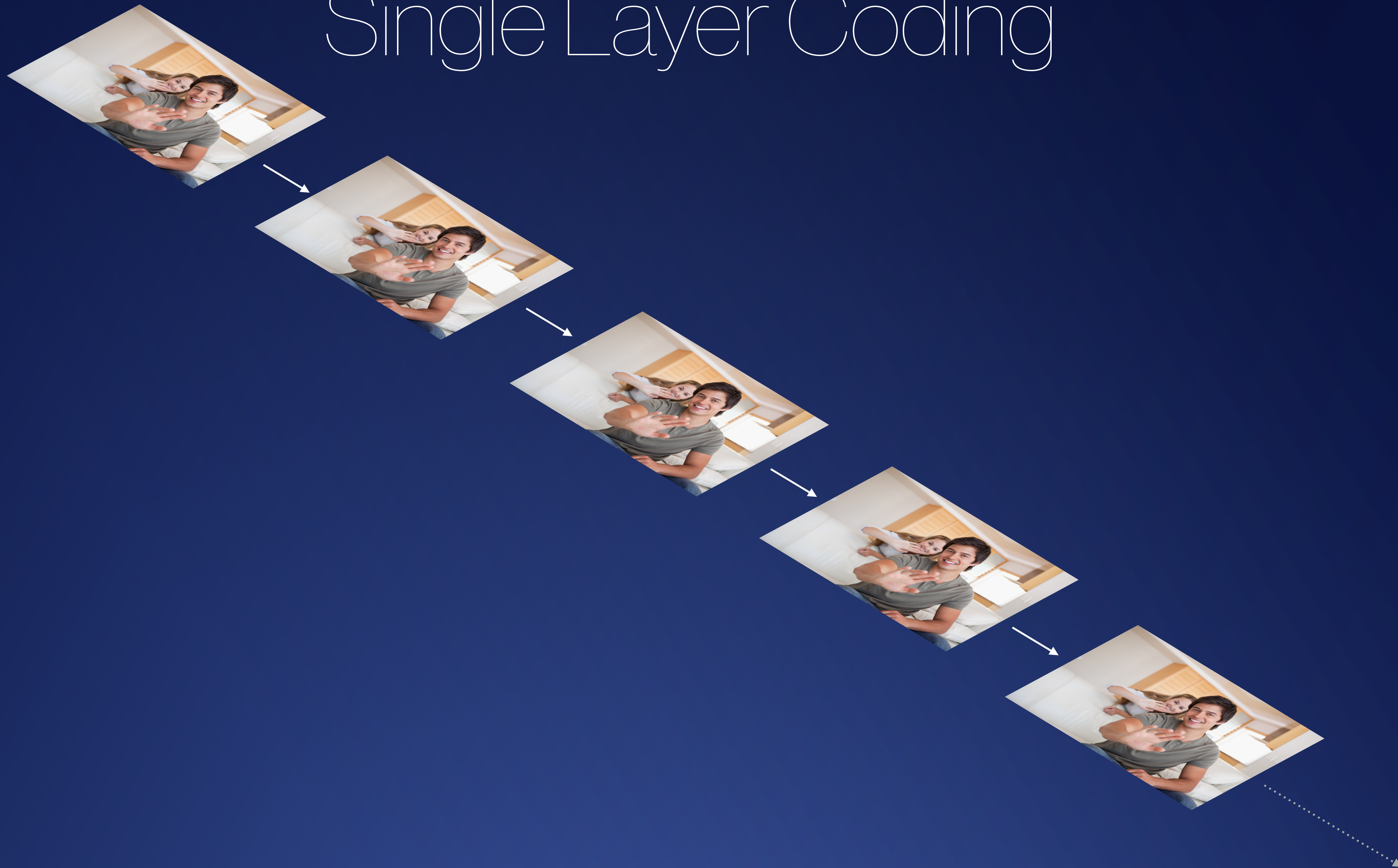
H.264/MPEG-4 Part 10
AVC Annex G

Scalable Video Coding



SVC

Single Layer Coding

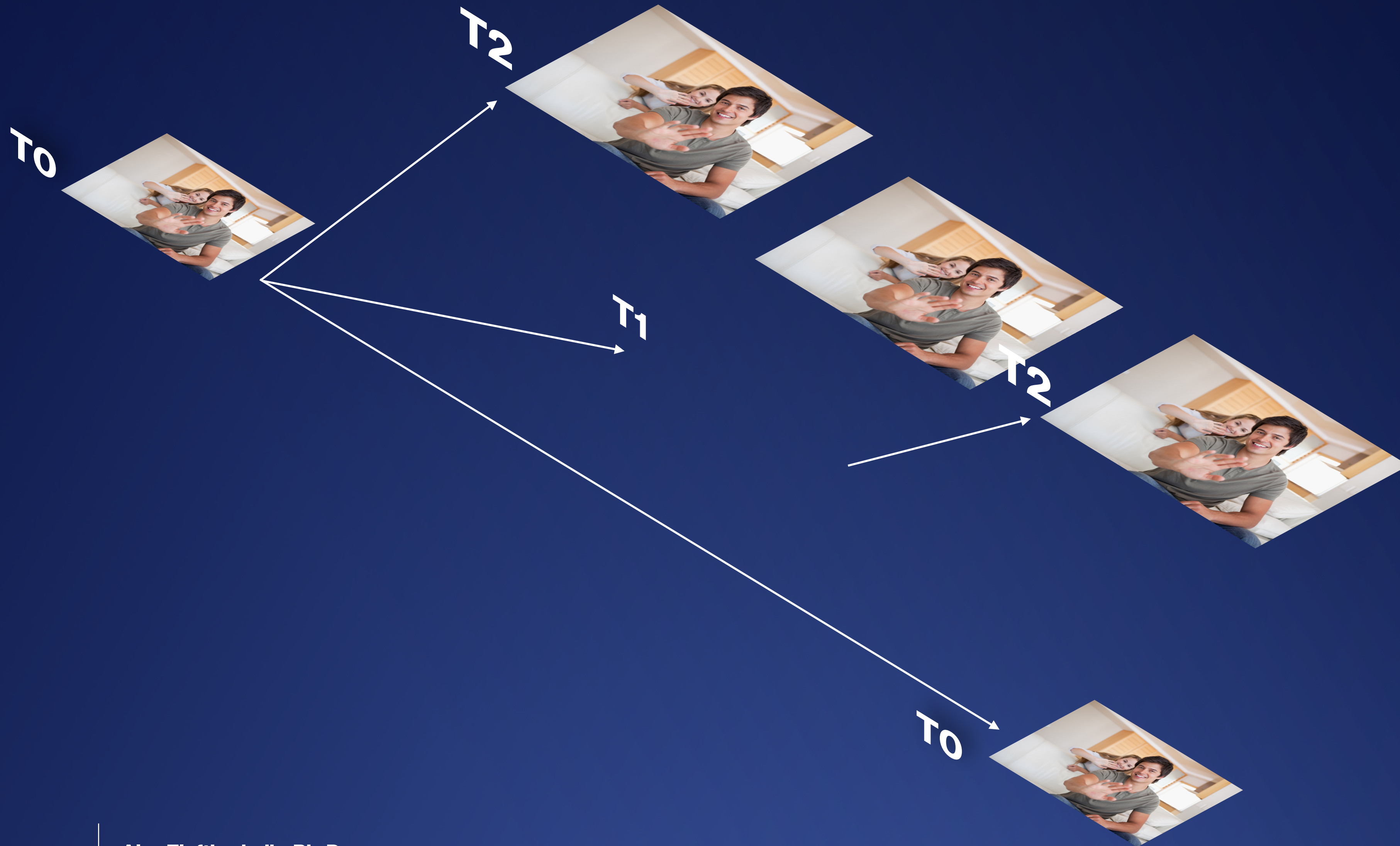




Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Proprietary, Confidential & Patent Pending Information

Temporal Scalability



Spatial Scalability



Spatial Scalability



Pros & Cons



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Pros & Cons

~ 20% more bits than single layer

superb error resilience - can take
more than 20% packet loss rate

adaptability

Scalability is already in:

H.264

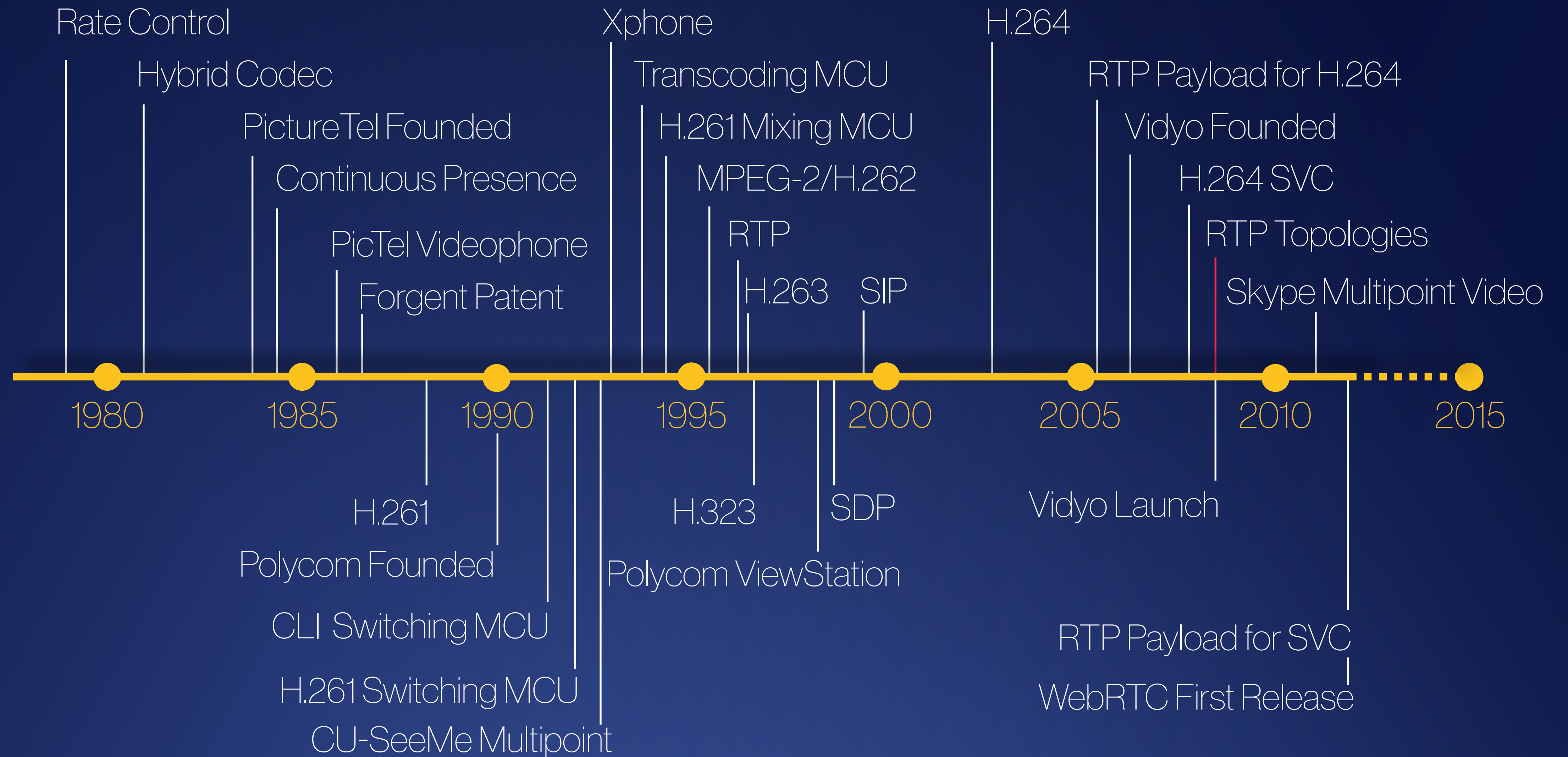
VP8

H.265/HEVC

H.265 v.2 - Scalable HEVC (October 2014)

VP9 (in progress - Vidyo/Google)





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

RTP Topologies

January 2008
Westerlund and Wenger
RFC 5117

Multipoint topologies used in
RTP-based environments,
particularly centralized topologies
for videoconferencing

Network Working Group
Request for Comments: 5117
Category: Informational

M. Westerlund
Ericsson
S. Wenger
Nokia
January 2008

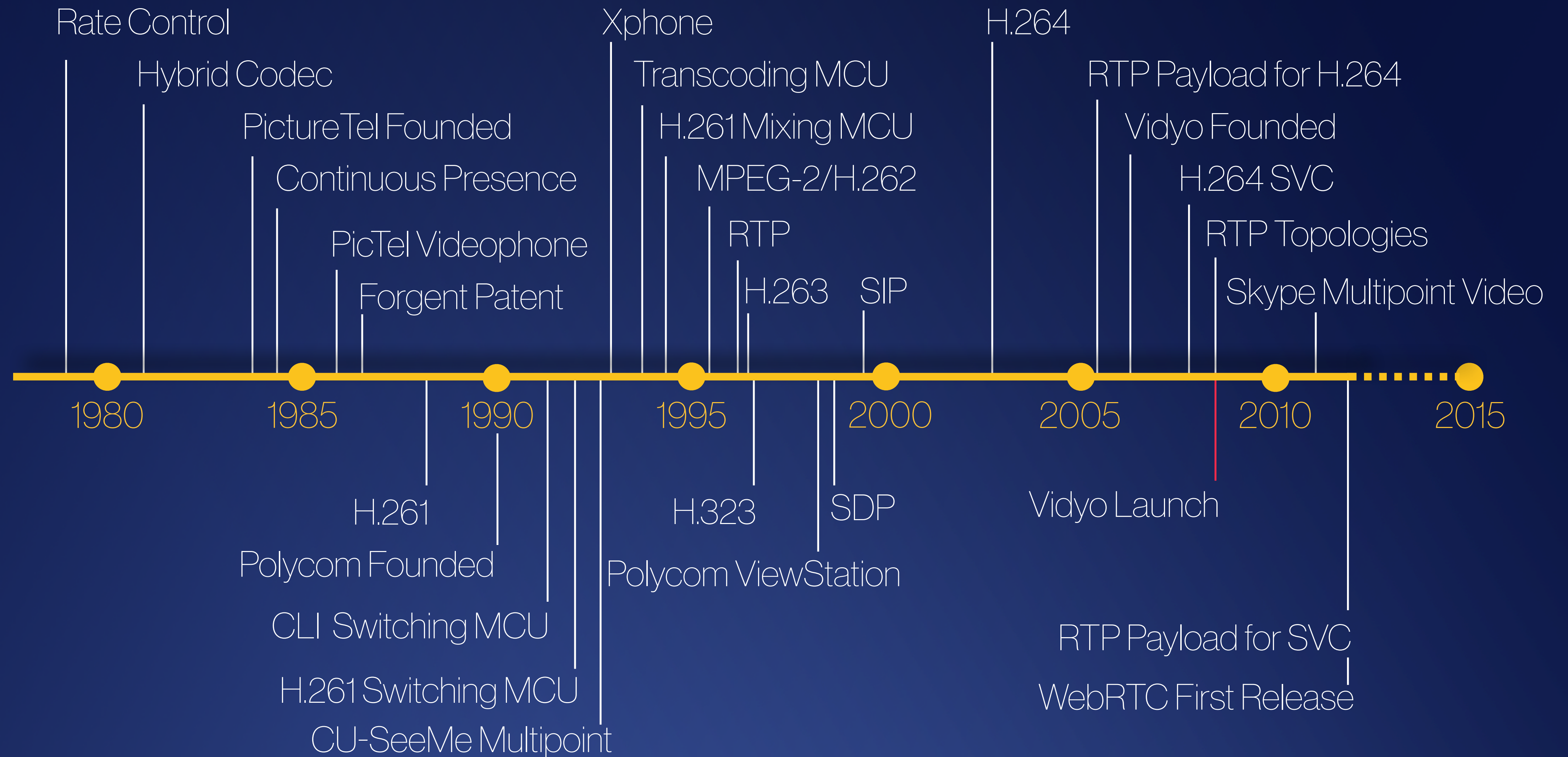
RTP Topologies

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Abstract

This document discusses multi-endpoint topologies used in Real-time Transport Protocol (RTP)-based environments. In particular, centralized topologies commonly employed in the video conferencing industry are mapped to the RTP terminology.



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Vidyo Launch

January 2008

Vidyo launches SVC and the VidyoRouter[®] architecture



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

A new type of server - VidyoRouter



A new type of server - VidyoRouter



Endpoint design



multi-stream
+
composition in endpoint
≈
browser



Perfect match for WebRTC



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

What did we gain?



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

What did we gain?

Error Resilience

Rate Matching

Personalized Layout

Low Delay

Error Localization

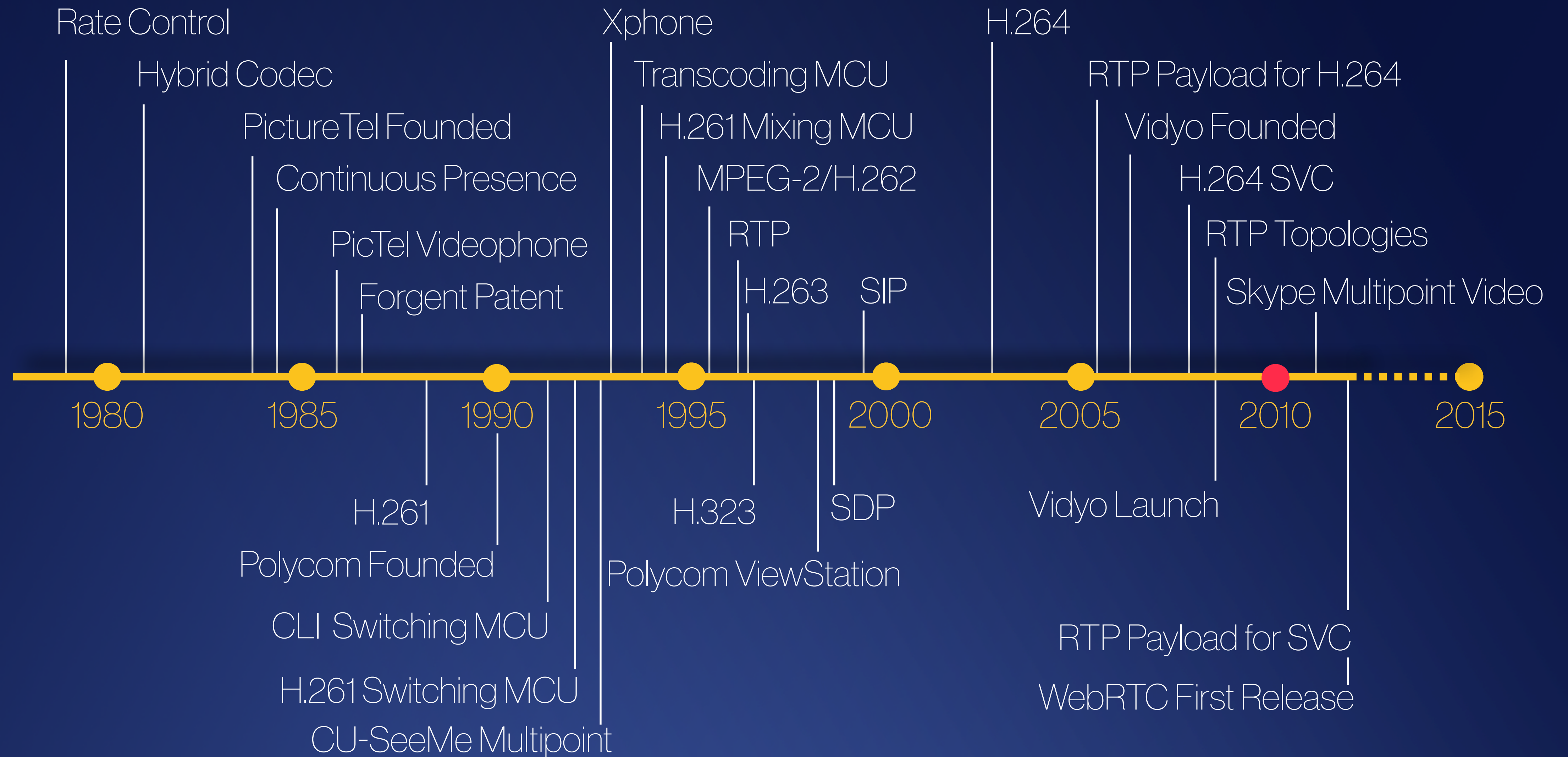
Low Complexity

Cascading

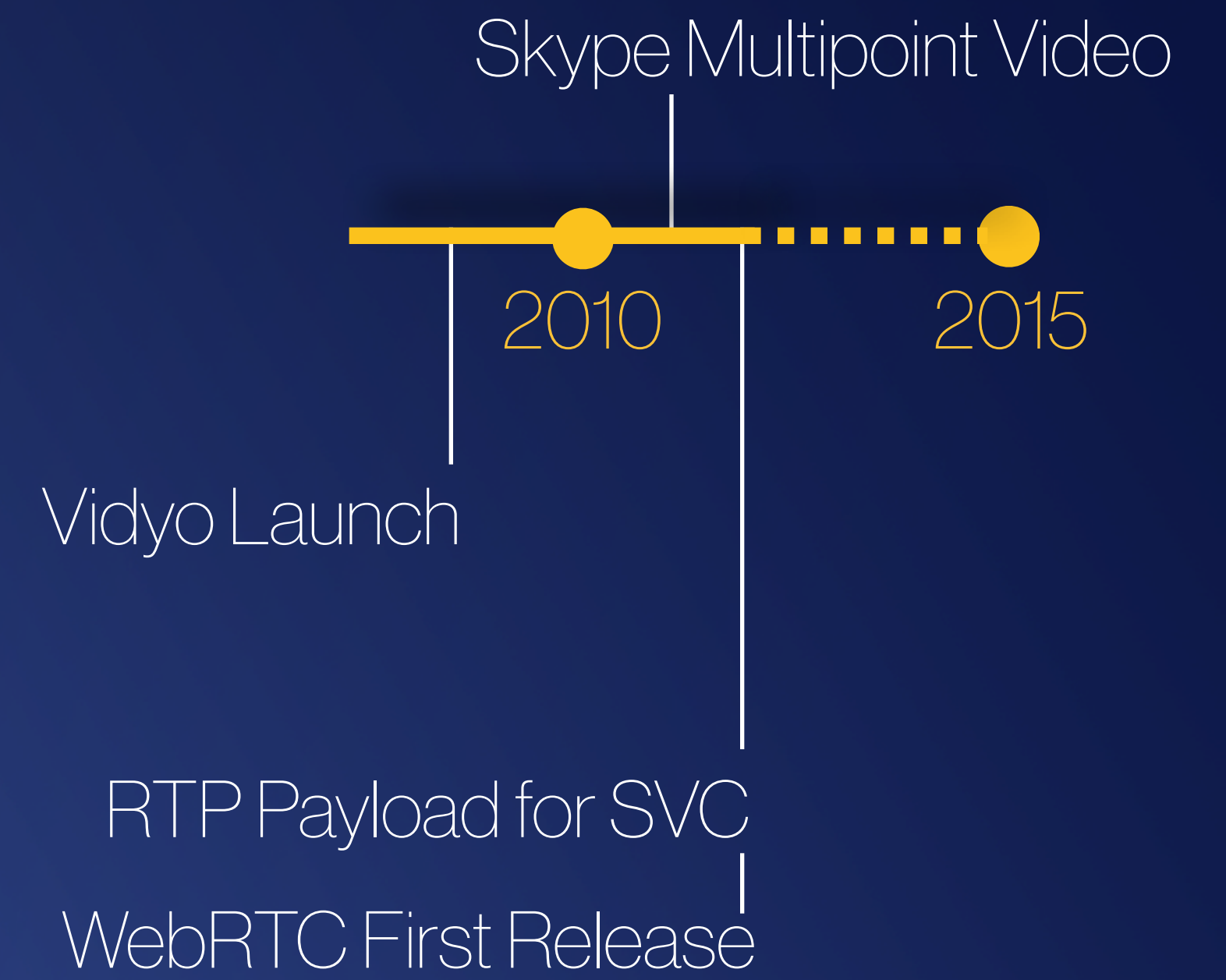
All with **no signal processing** at the server



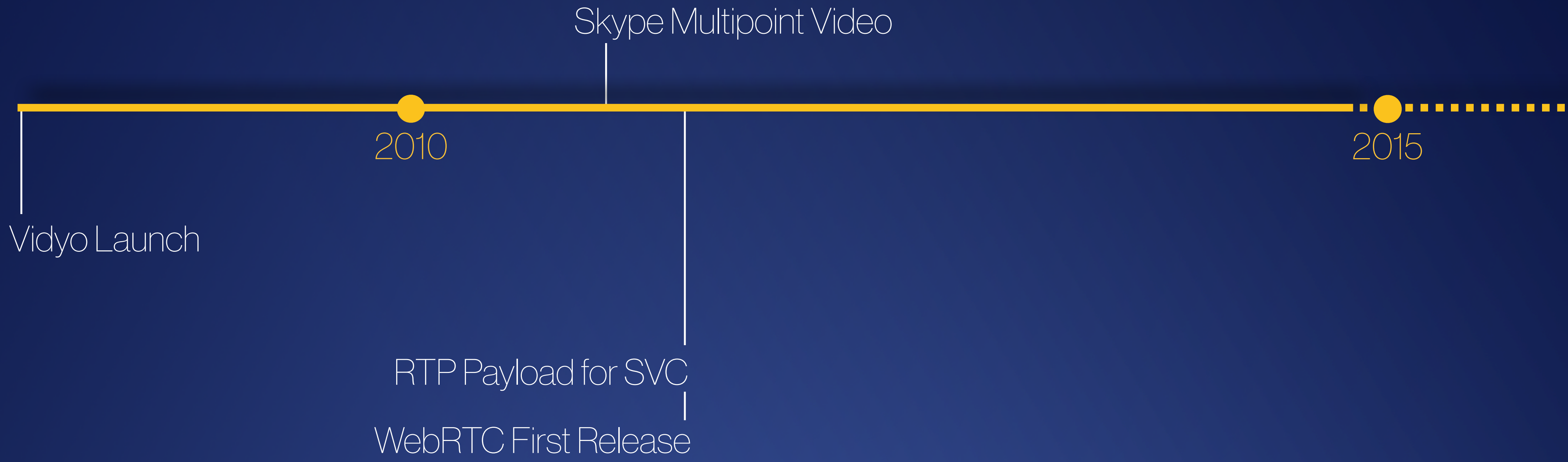
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



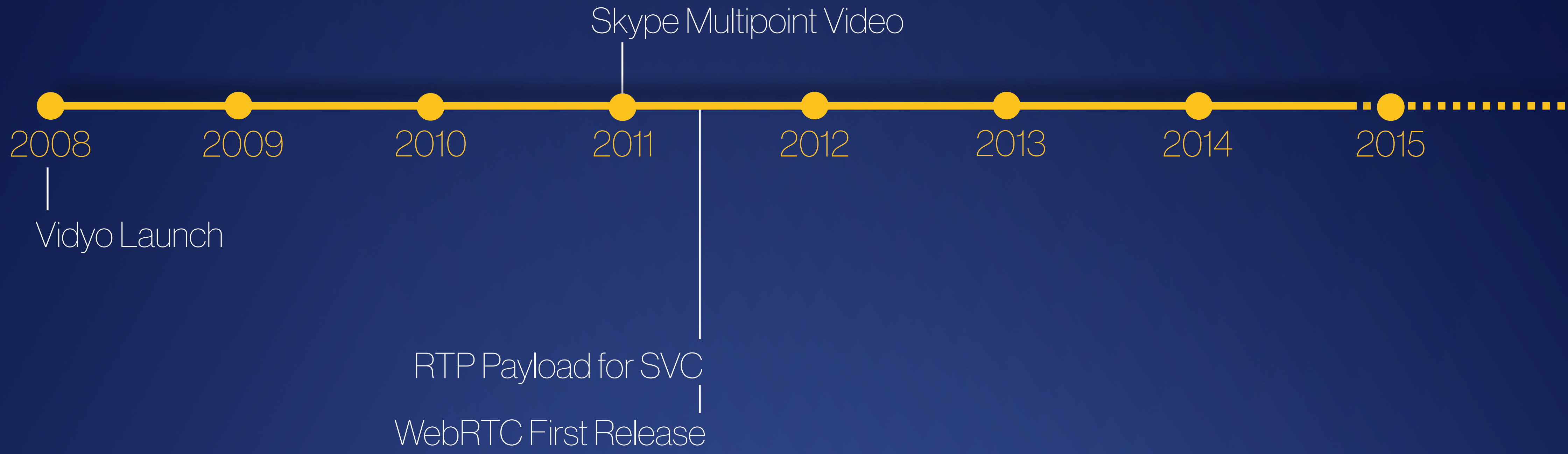
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



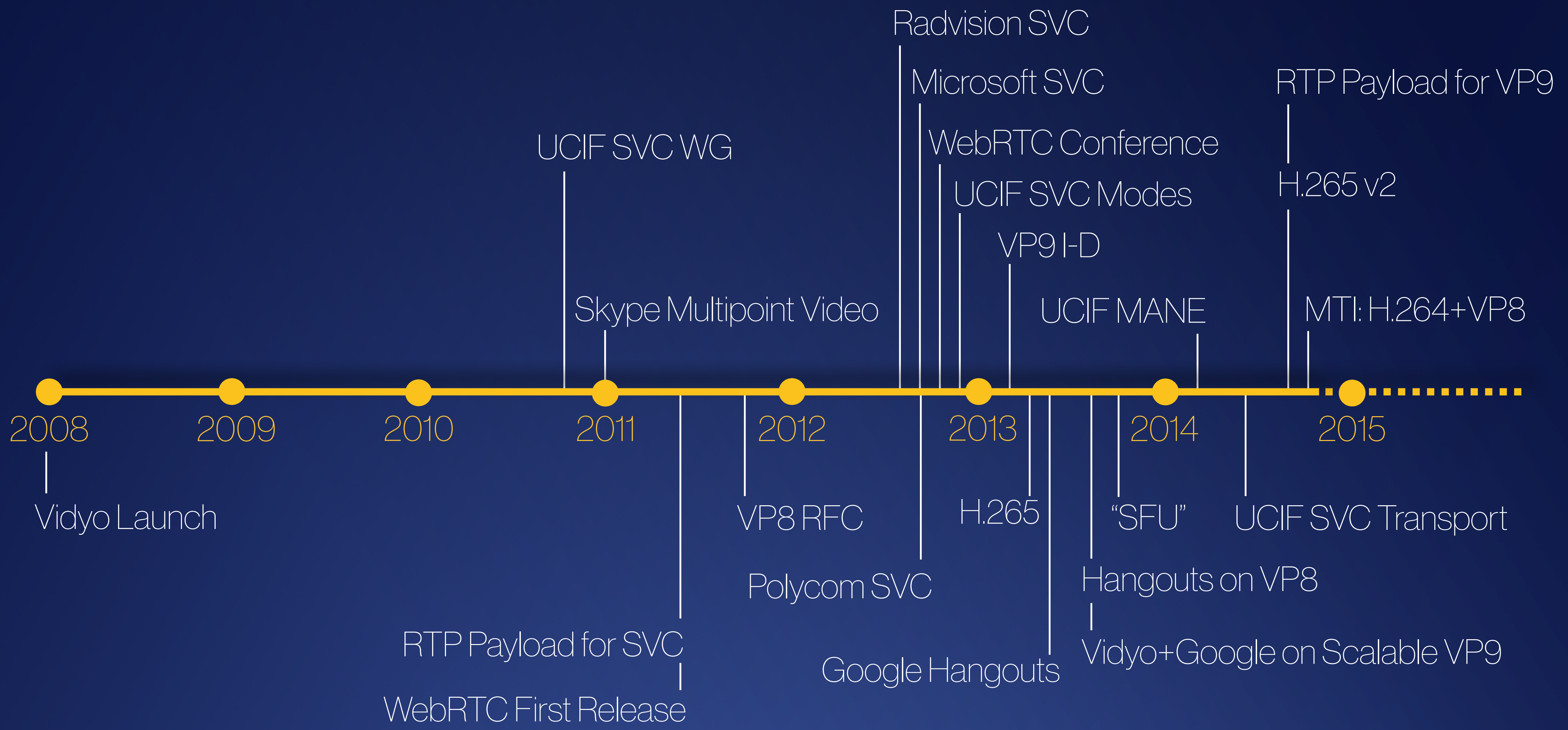
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



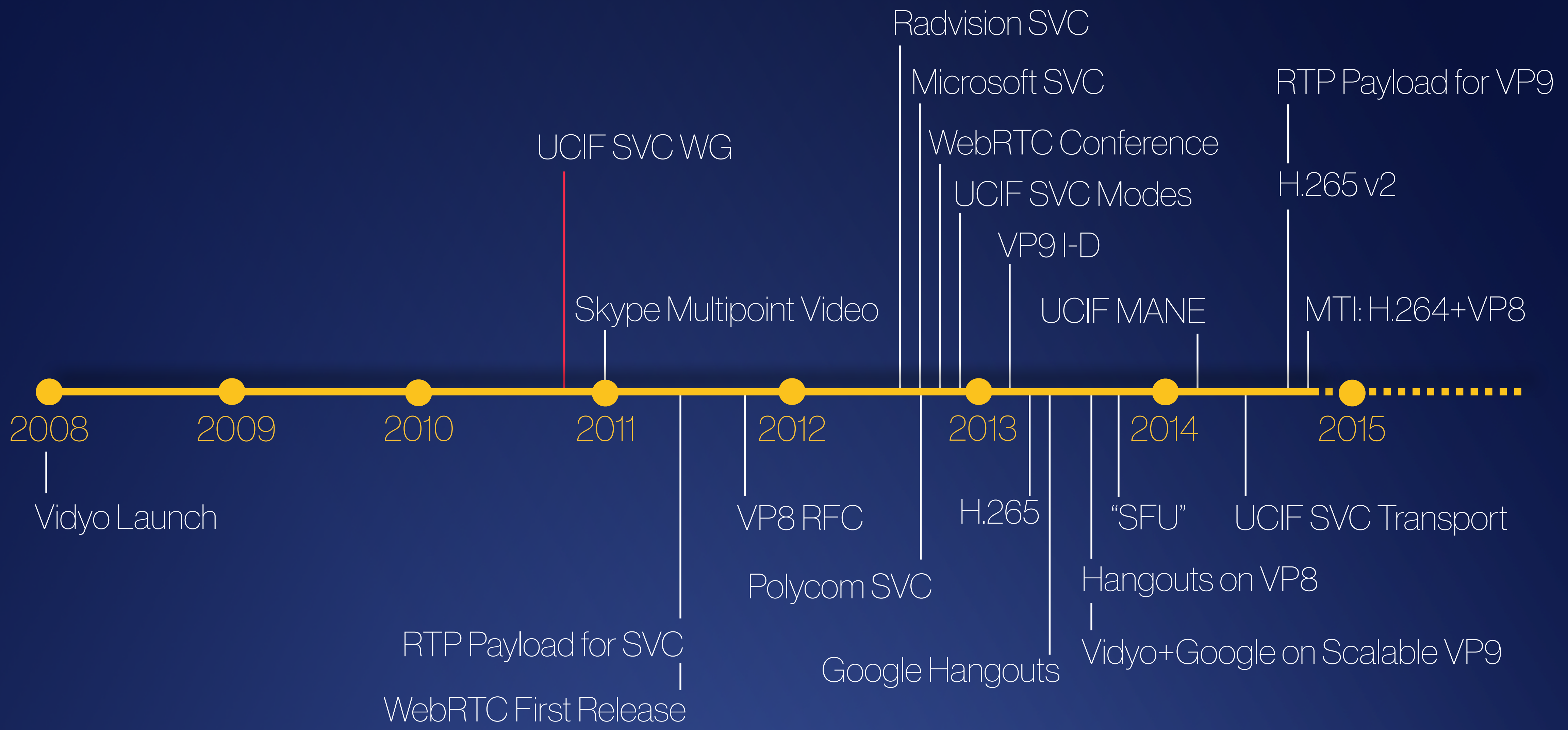
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Radvision SVC
 Microsoft SVC
 WebRTC Conference
 UCIF SVC Modes
 VP9 I-D
 UCIF MANE
 RTP Payload for VP9
 H.265 v2
 MTI: H.264+VP8



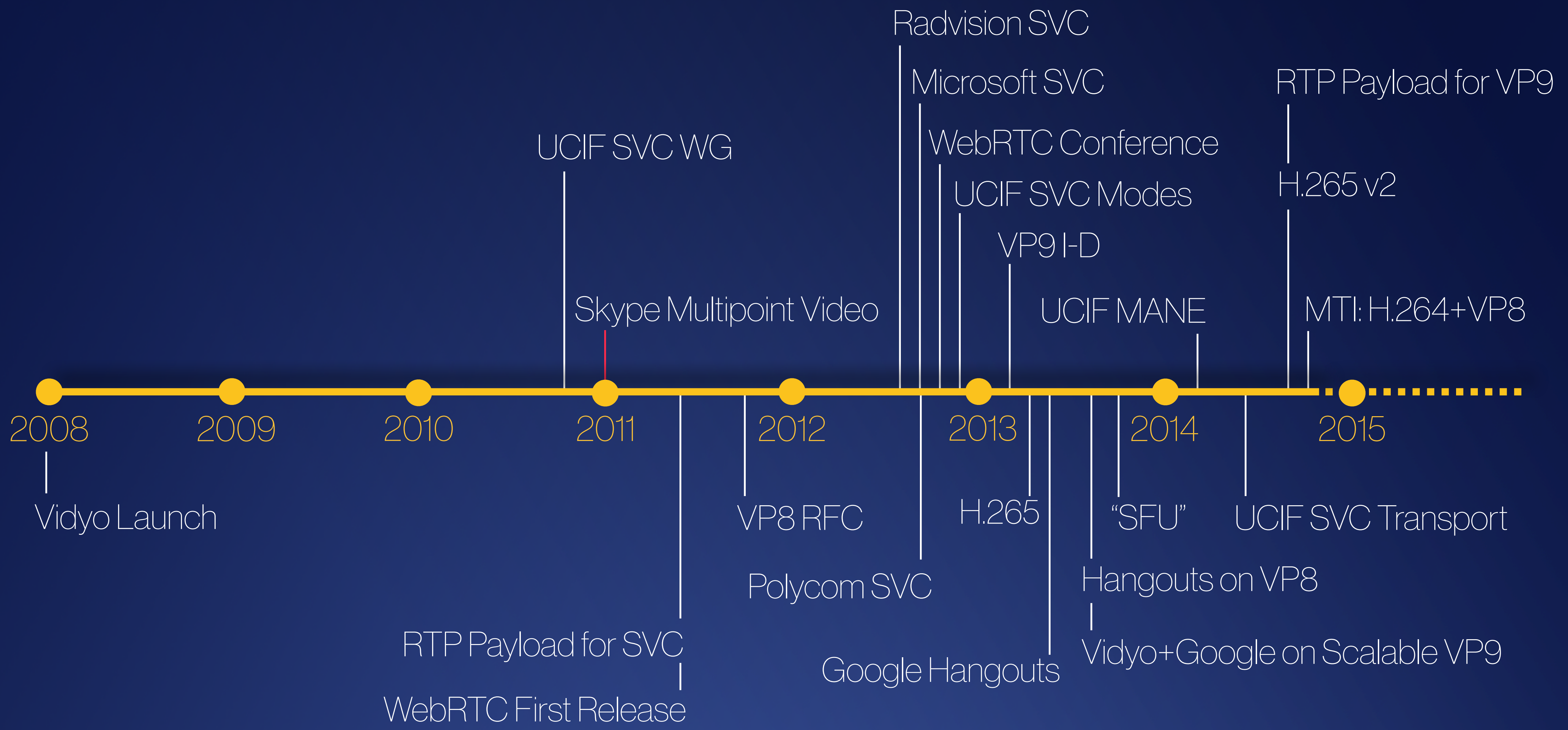
Alex Eleftheriadis, Ph. D.
 Chief Scientist & Co-founder

UCIF SVC WG

November 2010
UCI Forum SVC
Technical Working Group
formed



Certification and interop
for SVC systems



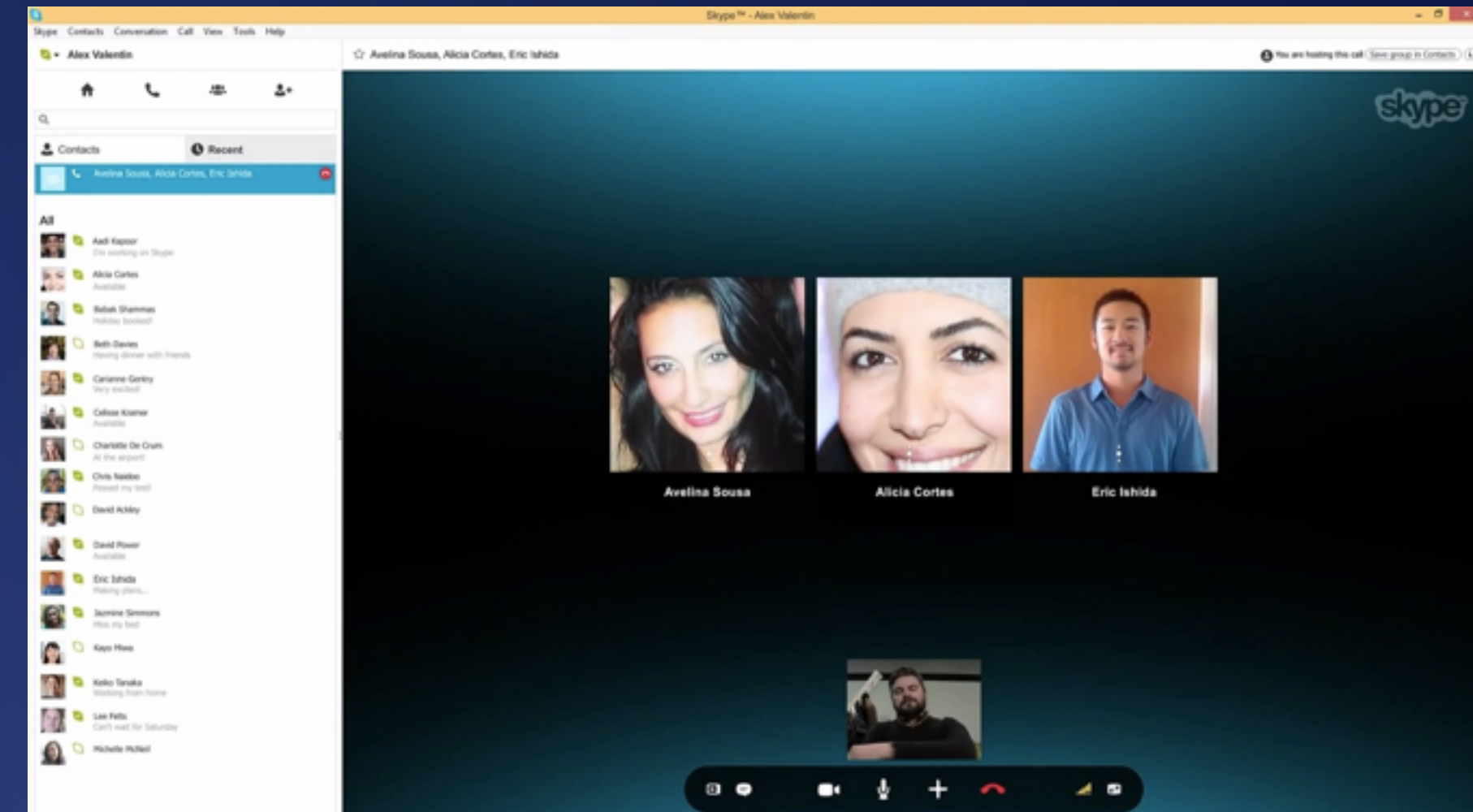
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

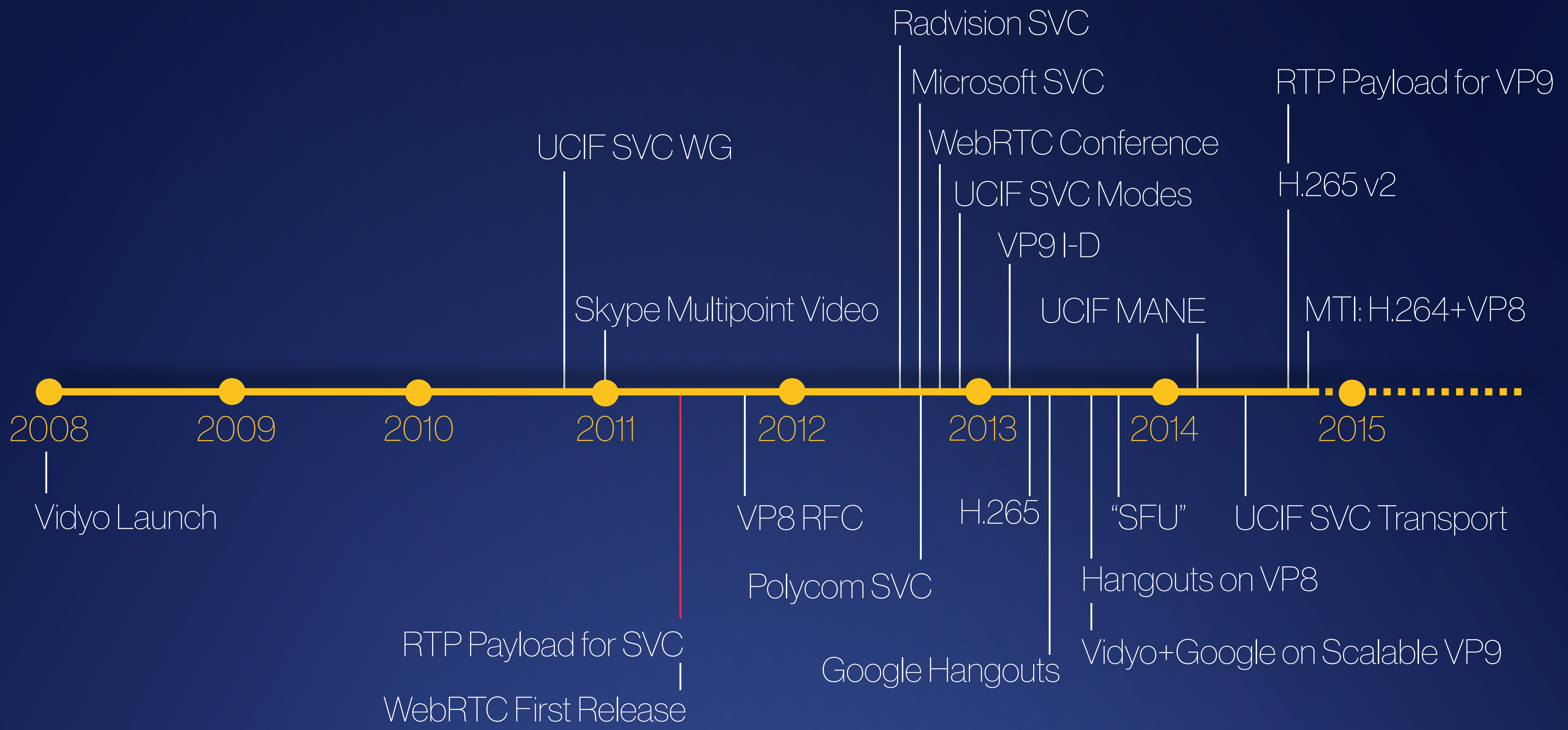
Skype Multipoint Video

January 2011

“Group Video Calling” with
up to 5 participants as
Premium service

Up to 10, and free as of April 2014





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

RTP Payload for SVC

May 2011
Wenger, Wang, Shierl,
and Eleftheriadis
RFC 6190

How to packetize H.264 SVC video
for RTP transport,
plus signaling parameters

Internet Engineering Task Force (IETF)
Request for Comments: 6190
Category: Standards Track
ISSN: 2070-1721

S. Wenger
Independent
Y.-K. Wang
Huawei Technologies
T. Schierl
Fraunhofer HHI
A. Eleftheriadis
Vidyo
May 2011

RTP Payload Format for Scalable Video Coding

Abstract

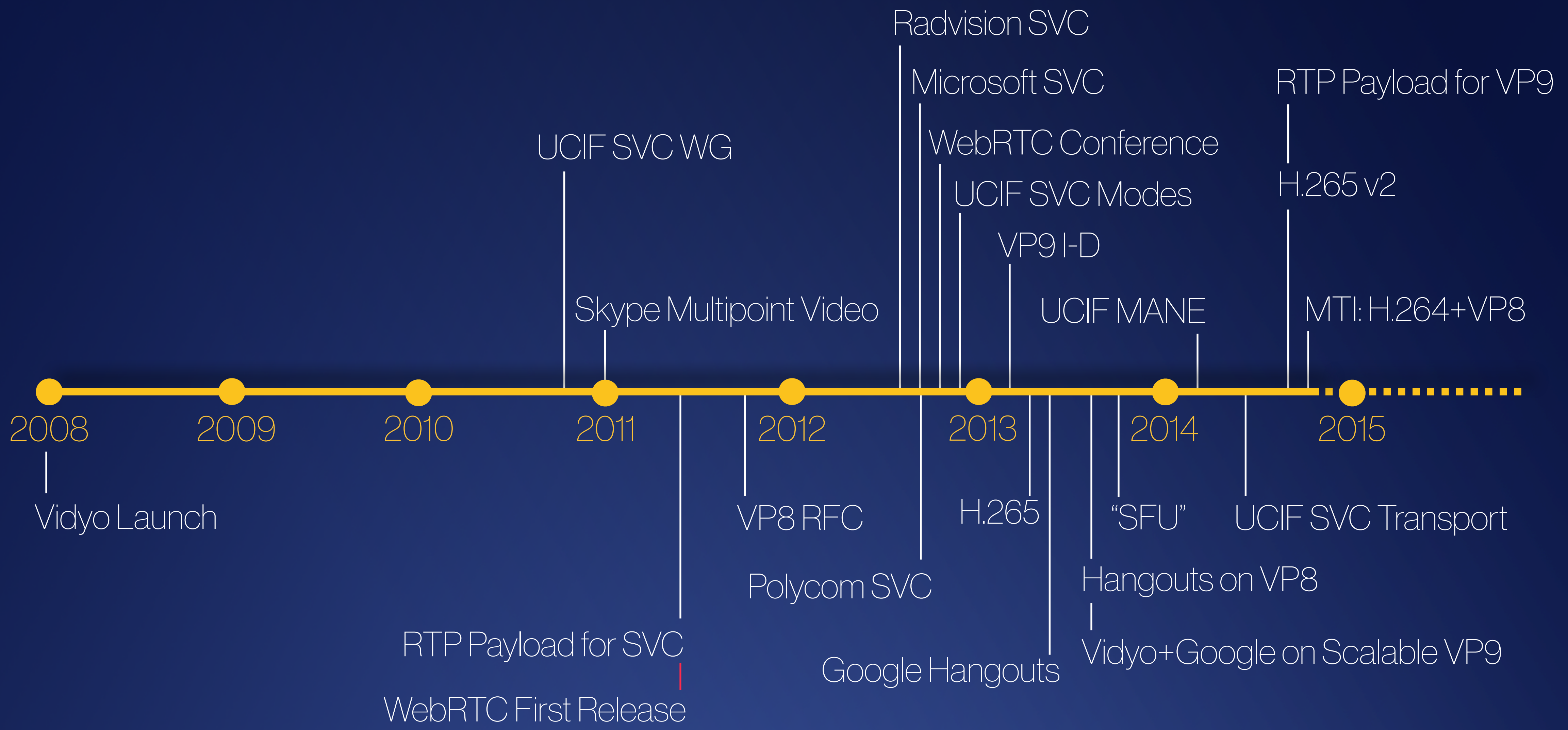
This memo describes an RTP payload format for Scalable Video Coding (SVC) as defined in Annex G of ITU-T Recommendation H.264, which is technically identical to Amendment 3 of ISO/IEC International Standard 14496-10. The RTP payload format allows for packetization of one or more Network Abstraction Layer (NAL) units in each RTP packet payload, as well as fragmentation of a NAL unit in multiple RTP packets. Furthermore, it supports transmission of an SVC stream over a single as well as multiple RTP sessions. The payload format defines a new media subtype name "H264-SVC", but is still backward compatible to RFC 6184 since the base layer, when encapsulated in its own RTP stream, must use the H.264 media subtype name ("H264") and the packetization method specified in RFC 6184. The payload format has wide applicability in videoconferencing, Internet video streaming, and high-bitrate entertainment-quality video, among others.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in [Section 2 of RFC 5741](#).

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <http://www.rfc-editor.org/info/rfc6190>.



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

WebRTC First Release

May 2011

Google releases WebRTC
source code

Google release of WebRTC source code

From: Harald Alvestrand <hta@google.com>

Date: Wed, 1 Jun 2011 01:04:28 +0200

Message-ID: <BANLkTinXEJypAc2tVvAySzUGOtG9kLZNXA@mail.gmail.com>

To: public-webrtc@w3.org

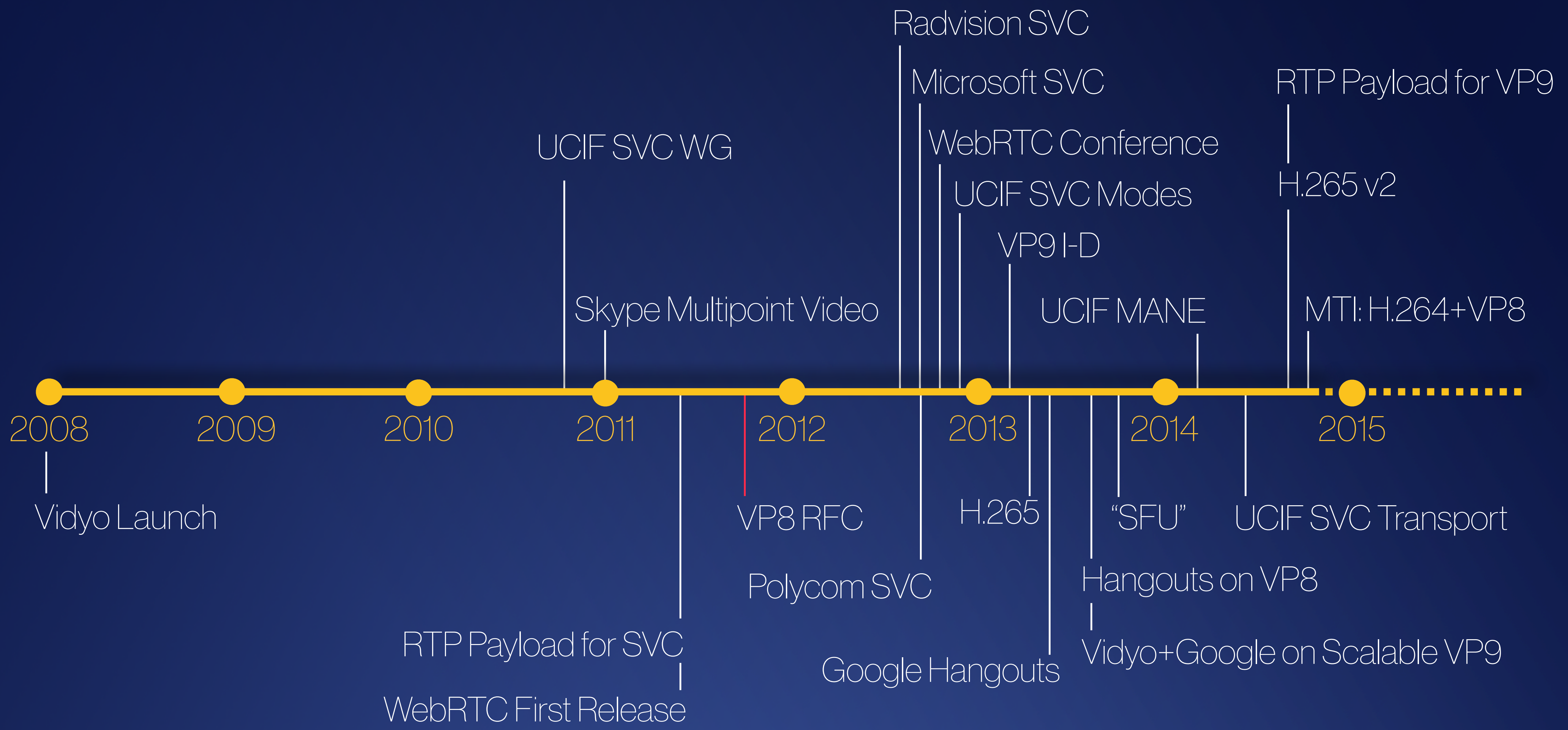
Today, Google made available WebRTC, an open source software package for real-time voice and video on the web that we will be integrating in Chrome. This is our initial contribution to achieve the mission of making audio and video available in all browsers, through a uniform standard set of APIs. This initial release will provide the functionality we envision for WebRTC/RTCWEB, as detailed at <https://sites.google.com/site/webrtc/>. Working with the browser community and working groups like this, our goal is to expand the available APIs over the next few months for developers to create voice and video applications on the web.

The underlying components we're releasing are stable and the interfaces for this initial release are consistent with the discussions in this working group. We will continue to provide working implementations for consideration and feedback to collectively ensure stable standards are finalized. Google is committed to fully supporting these standards and we look forward to your input in the coming months.

Harald, speaking for Google.

Received on Tuesday, 31 May 2011 23:05:13 GMT

This archive was generated by [hypermail 2.2.0+W3C-0.50](#) : Tuesday, 31 May 2011 23:05:14 GMT



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

VP8 RFC

November 2011
Bankoski, Koleszar, Quillio,
Salonen, Wilkins, and Xu
RFC 6386

VP8 data format and
decoding guide
(independent submission)

Independent Submission
Request for Comments: 6386
Category: Informational
ISSN: 2070-1721

J. Bankoski
J. Koleszar
L. Quillio
J. Salonen
P. Wilkins
Y. Xu
Google Inc.
November 2011

VP8 Data Format and Decoding Guide

Abstract

This document describes the VP8 compressed video data format, together with a discussion of the decoding procedure for the format.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

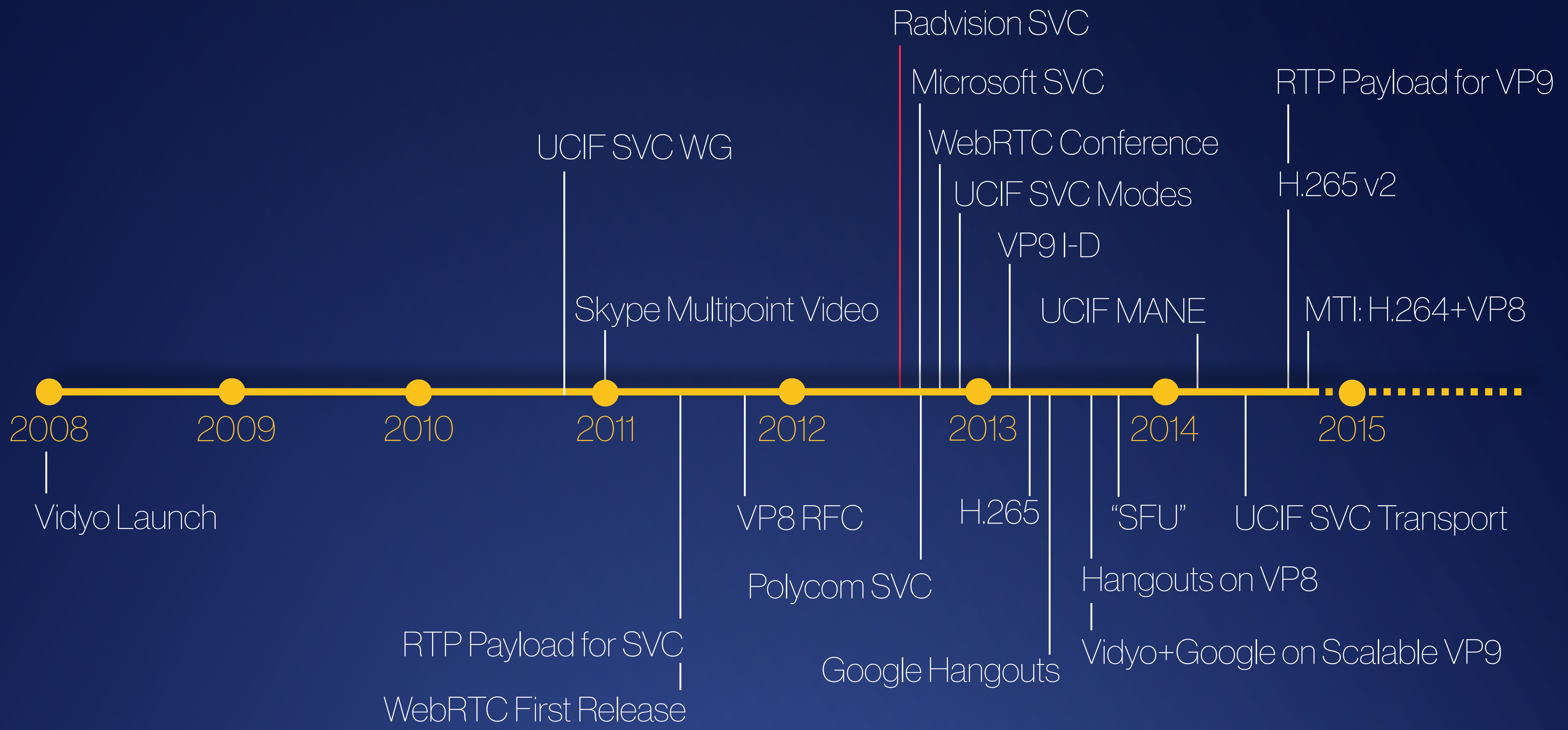
This is a contribution to the RFC Series, independently of any other RFC stream. The RFC Editor has chosen to publish this document at its discretion and makes no statement about its value for implementation or deployment. Documents approved for publication by the RFC Editor are not a candidate for any level of Internet Standard; see [Section 2 of RFC 5741](#).

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <http://www.rfc-editor.org/info/rfc6386>.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Radvision SVC

September 2012

Elite 5000 Series MCU Ver. 7.7
SVC for error resilience



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

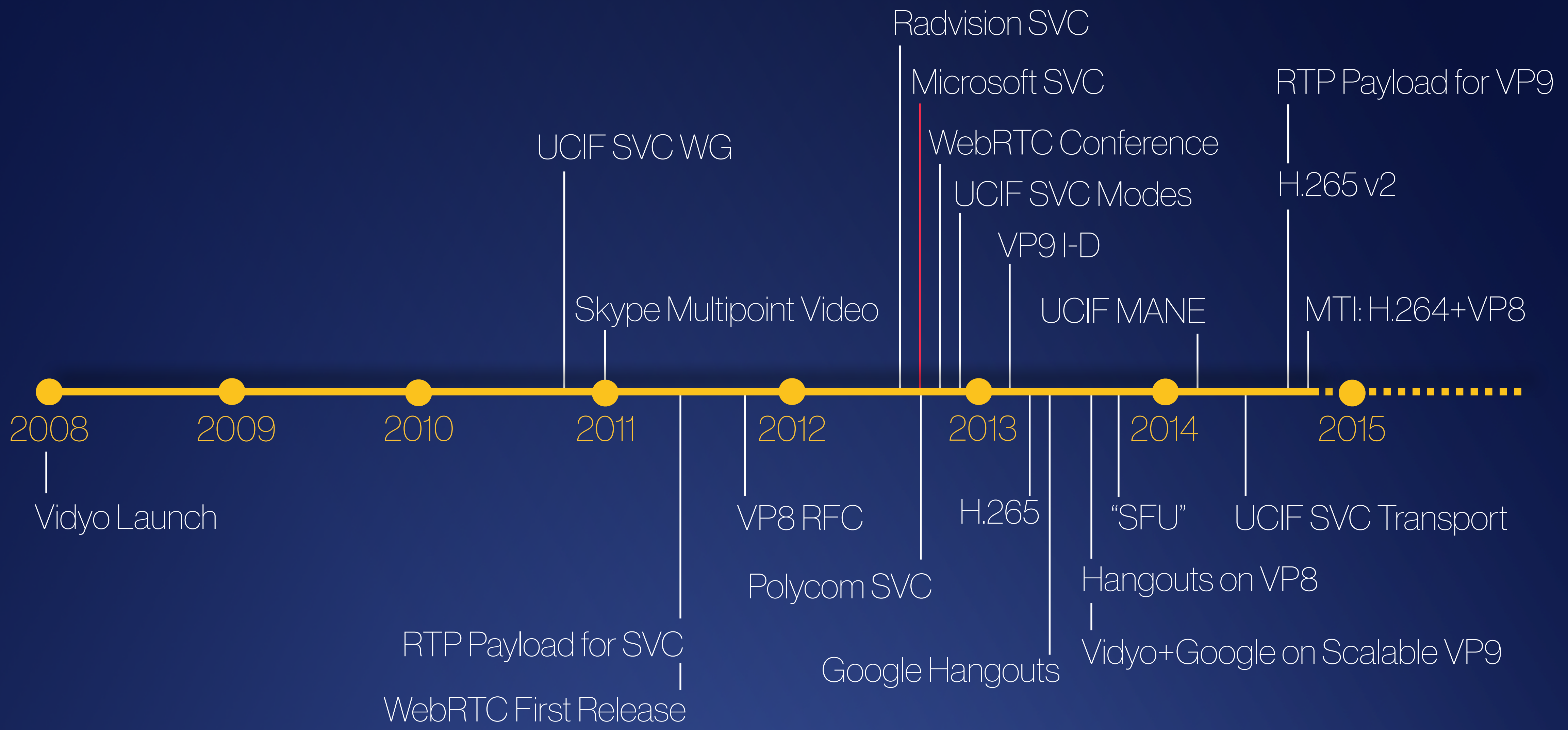


SCOPIA Elite MCU 5000 Series



SCOPIA Elite Provides the Highest Quality Experience and Most Advanced Conferencing Capabilities Available

- Supreme High Definition**
With 1080p High Definition processing, multistream telepresence connectivity, unlimited conferences, personal video layout per participant and AAC audio, SCOPIA Elite delivers supreme High Definition conferencing. Utilization of the very latest state of the art DSP technology offers uncompromised HD.
- Flexible Resources**
Mixed endpoint applications are supported with ease and efficiency. Enhanced Definition at 4X capacity affords excellent value. On-demand, dynamic and automatic resource allocation supports any combination of endpoint capabilities without complex configuration changes.
- Optimized Quality**
Encoder per participant ensures an optimal quality experience for any endpoint on any network. H.264 Scalable Video Coding (SVC) delivers superior performance over compromised networks.
- Maximum Usability**
SCOPIA Elite was designed by users and administrators for usability and simplicity. On-screen information overlays, easy conference creation and entry through the enhanced video auto attendant and IVR make attending conferences simple. The industry leading management interface makes seemingly complicated tasks simple; users can view important statistics at a glance saving time, money and headache.
- Designed to Scale**
Unequalled scalability is achieved through patented, distributed multipoint conferencing. SCOPIA Elite utilizes SVC, enabling enhanced connections between MCUs in poor network conditions. The Advanced Telecommunications Computing Architecture (ATCA) delivers investment protection through a future proof platform.



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Microsoft SVC

October 2012

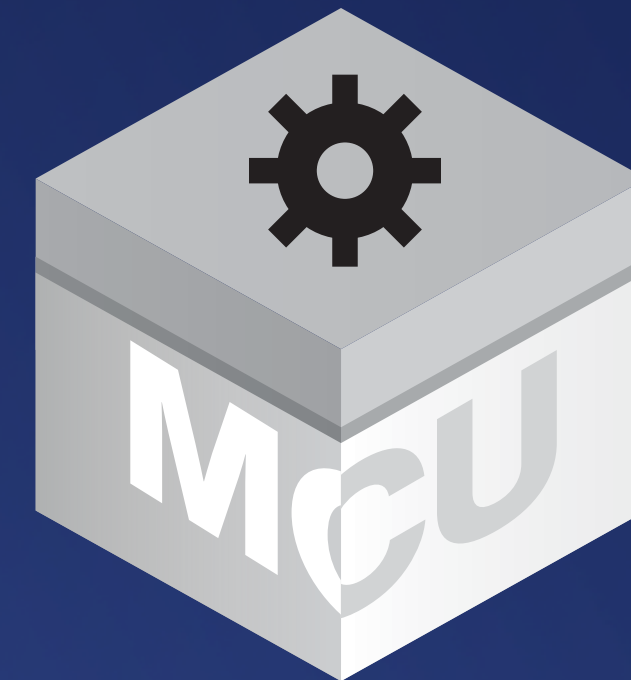
Simulcast SVC (temporal only)
on Lync 2013



Simulcasting vs. Scalable Coding

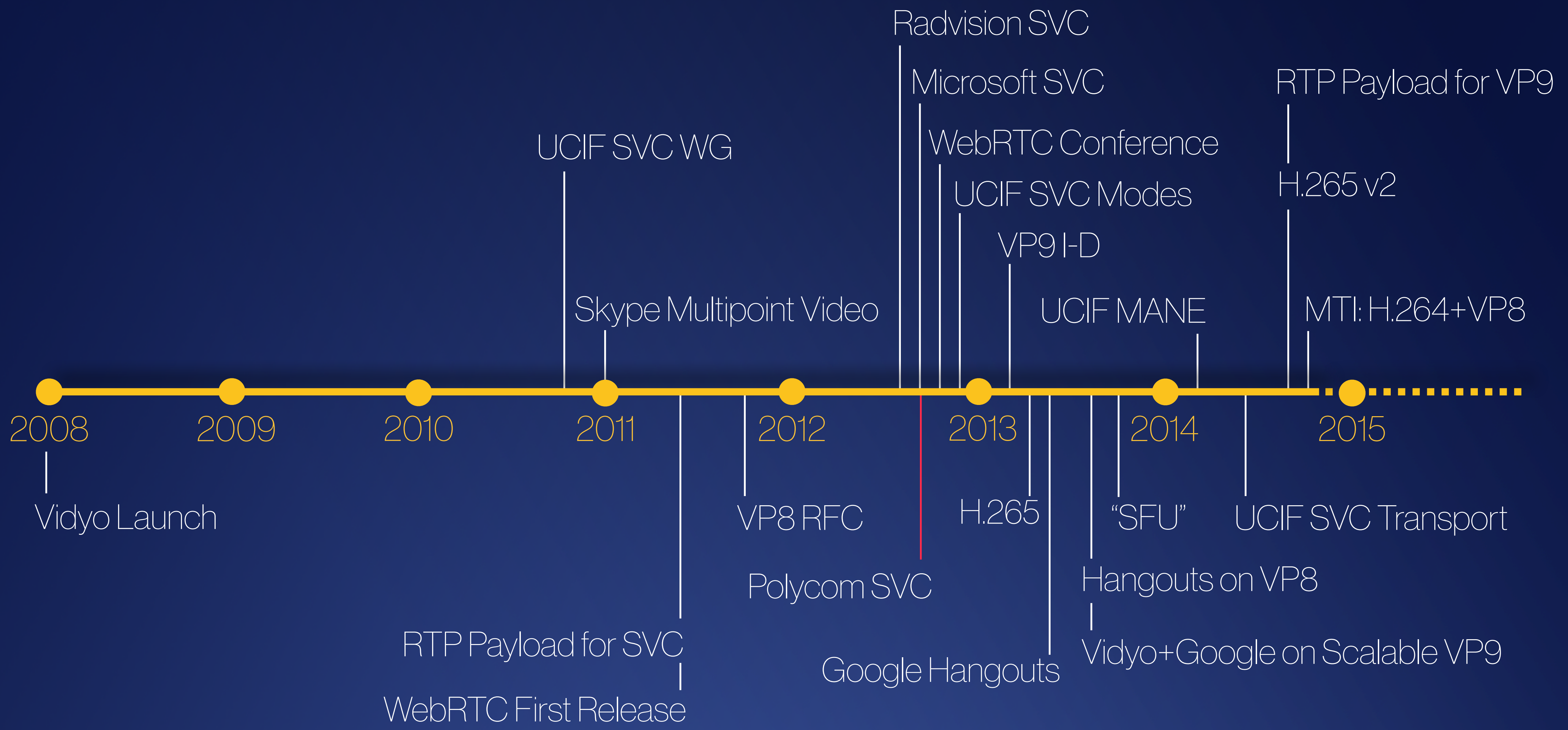


Simulcast Server



“Selective Forwarding Unit”

SFU

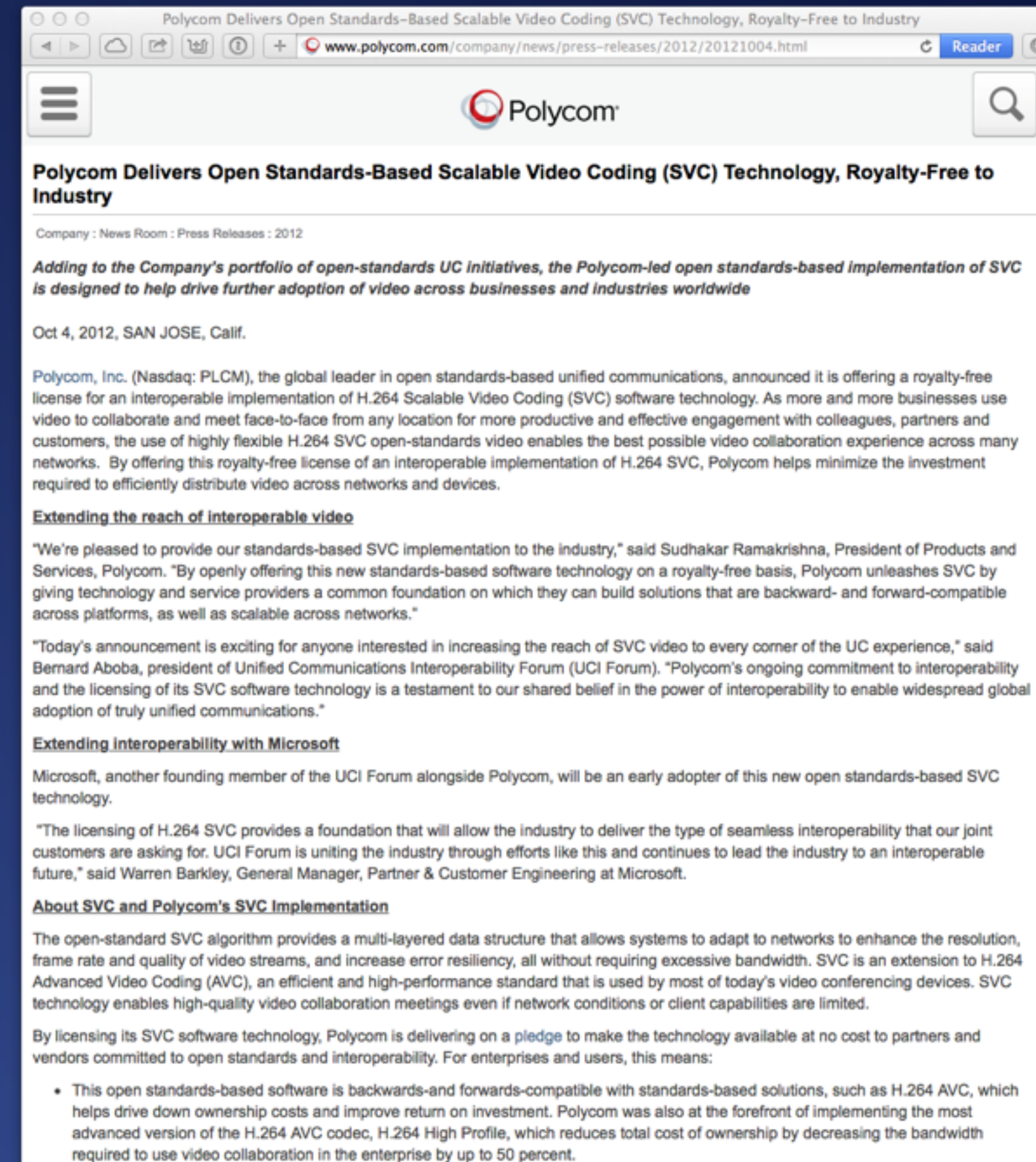


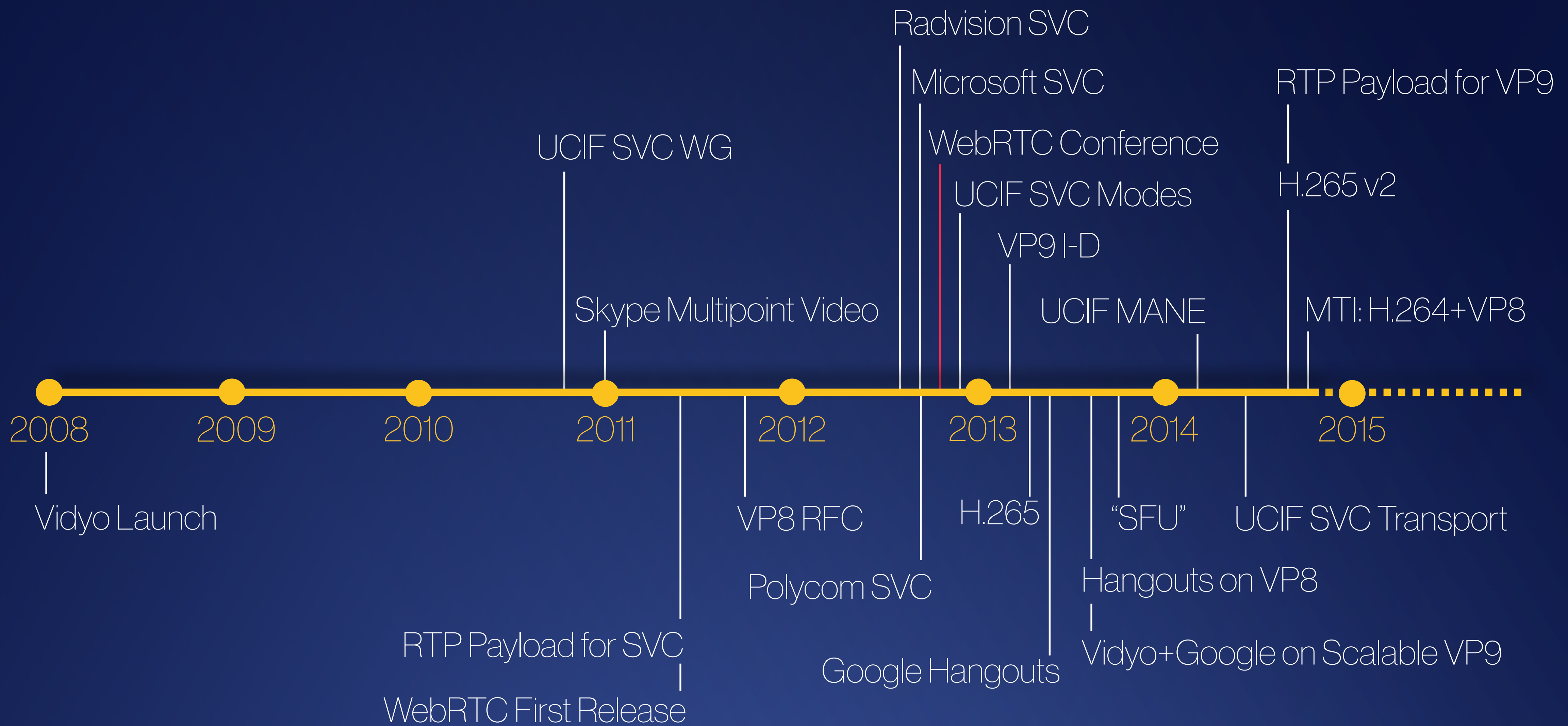
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Polycom SVC

October 2012

Announces availability of royalty-free SVC implementation, with Microsoft as early adopter





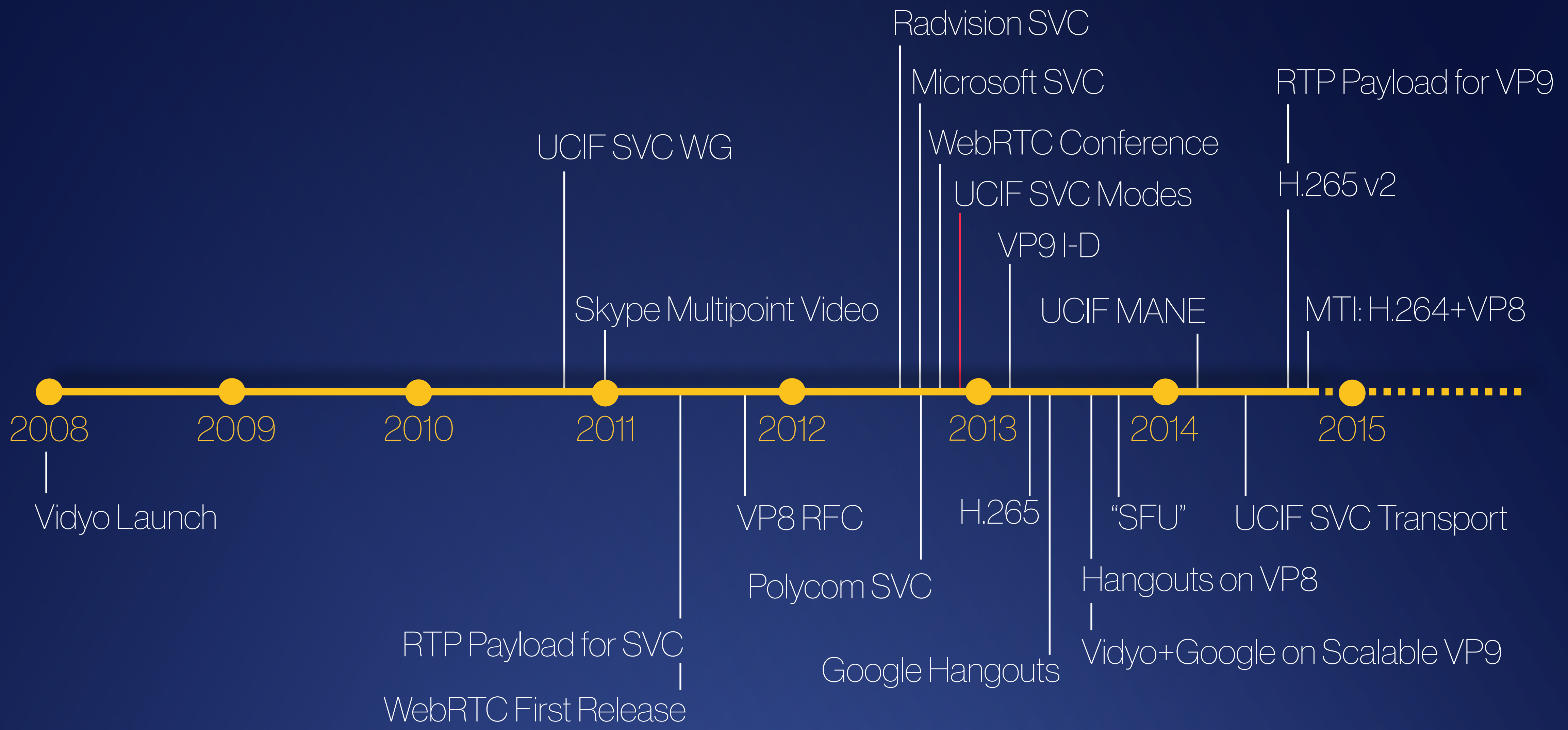
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

WebRTC Conference

November 27-29, 2012

First WebRTC Conference and Expo
South San Francisco





Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

UCIF SVC Modes

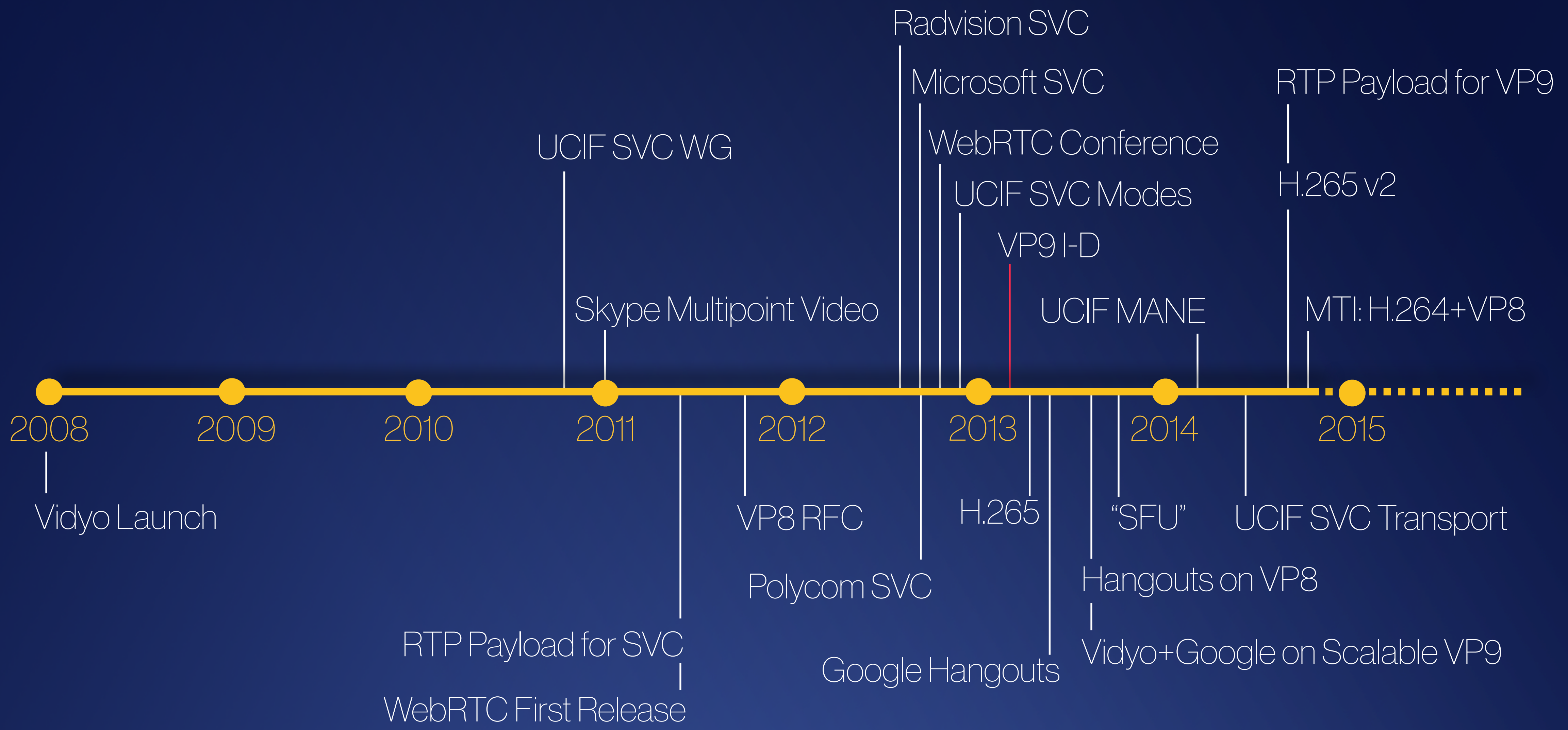
December 2012

AVC and SVC Video
Bitstream Modes

(Lync does Mode 1, Vidyo Mode 2s)



Unified Communication
Specification for
H.264/MPEG-4 Part 10
AVC and SVC Modes
Version 1.0



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

VP9 I-D

February 2013
Grange and Alvestrand
draft-grange-vp9-bitstream
Overview of VP9 Bitstream

Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 22, 2013

A. Grange
H. Alvestrand
Google
February 18, 2013

A VP9 Bitstream Overview
draft-grange-vp9-bitstream-00

Abstract

This document describes VP9, a video codec being developed specifically to meet the demand for the consumption of video over the Internet, including professionally and amateur produced video-on-demand and conversational video content. VP9 is an evolution of the VP8 video codec that is described in [bankoski-rfc6386] and includes a number of enhancements and new coding tools that have been added to improve the coding efficiency. The new tools that have been added so far include: larger prediction block sizes up to 64x64, various forms of compound INTER prediction, more modes for INTRA prediction, pel motion vectors, 8-tap switchable sub-pixel interpolation filters, improved motion reference generation, improved motion vector coding, improved entropy coding including frame-level entropy adaptation for various symbols, improved loop filtering, the incorporation of the Asymmetric Discrete Sine Transform (ADST), larger 16x16 and 32x32 DCTs, and improved frame level segmentation. VP9 is under active development and this document provides only a snapshot of the current state of the coding tools as they exist today. The finalized version of the VP9 bitstream may differ considerably from the description contained herein and may encompass the exclusion or modification of existing coding tools or the addition of new coding tools.

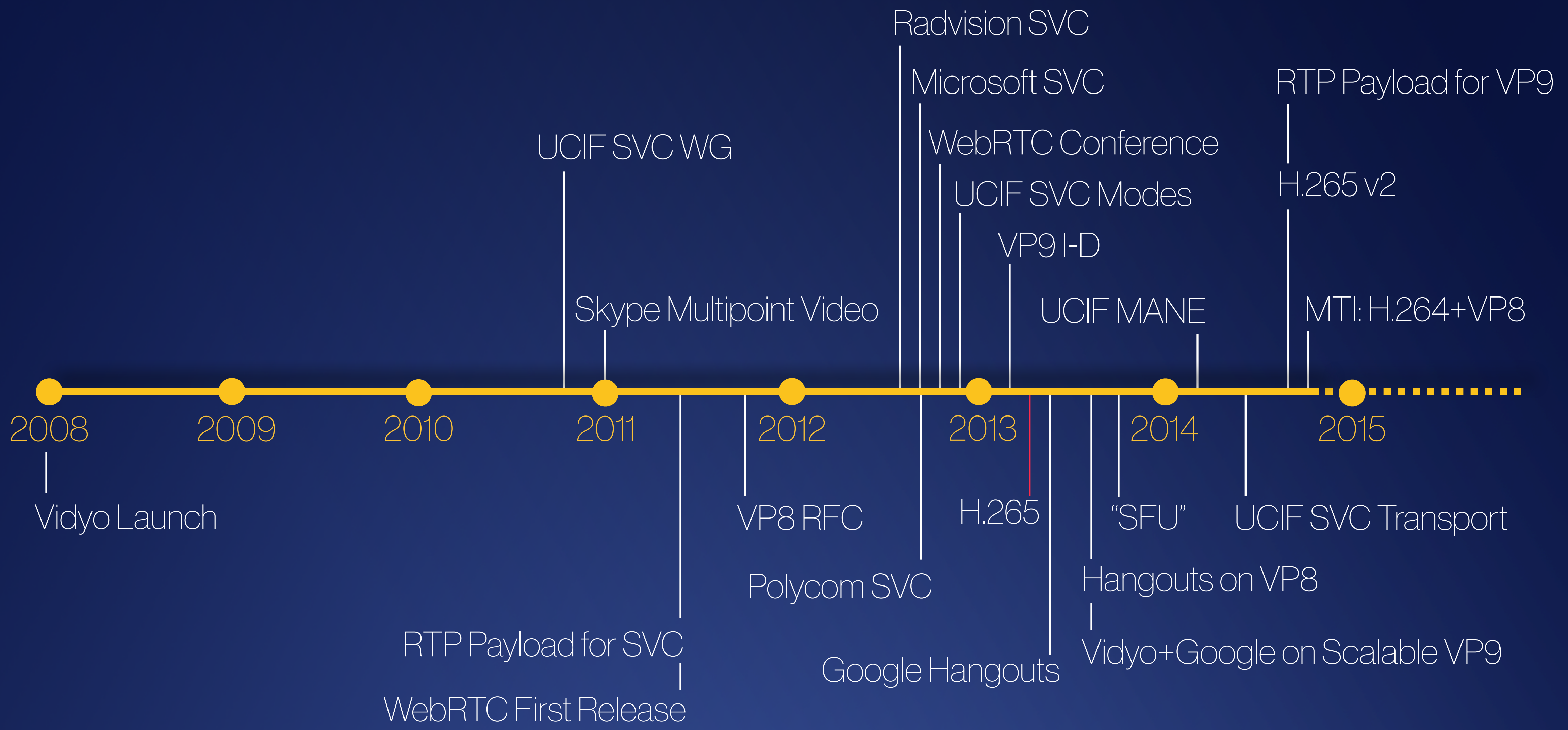
Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

Grange & Alvestrand Expires August 22, 2013 [Page 1]



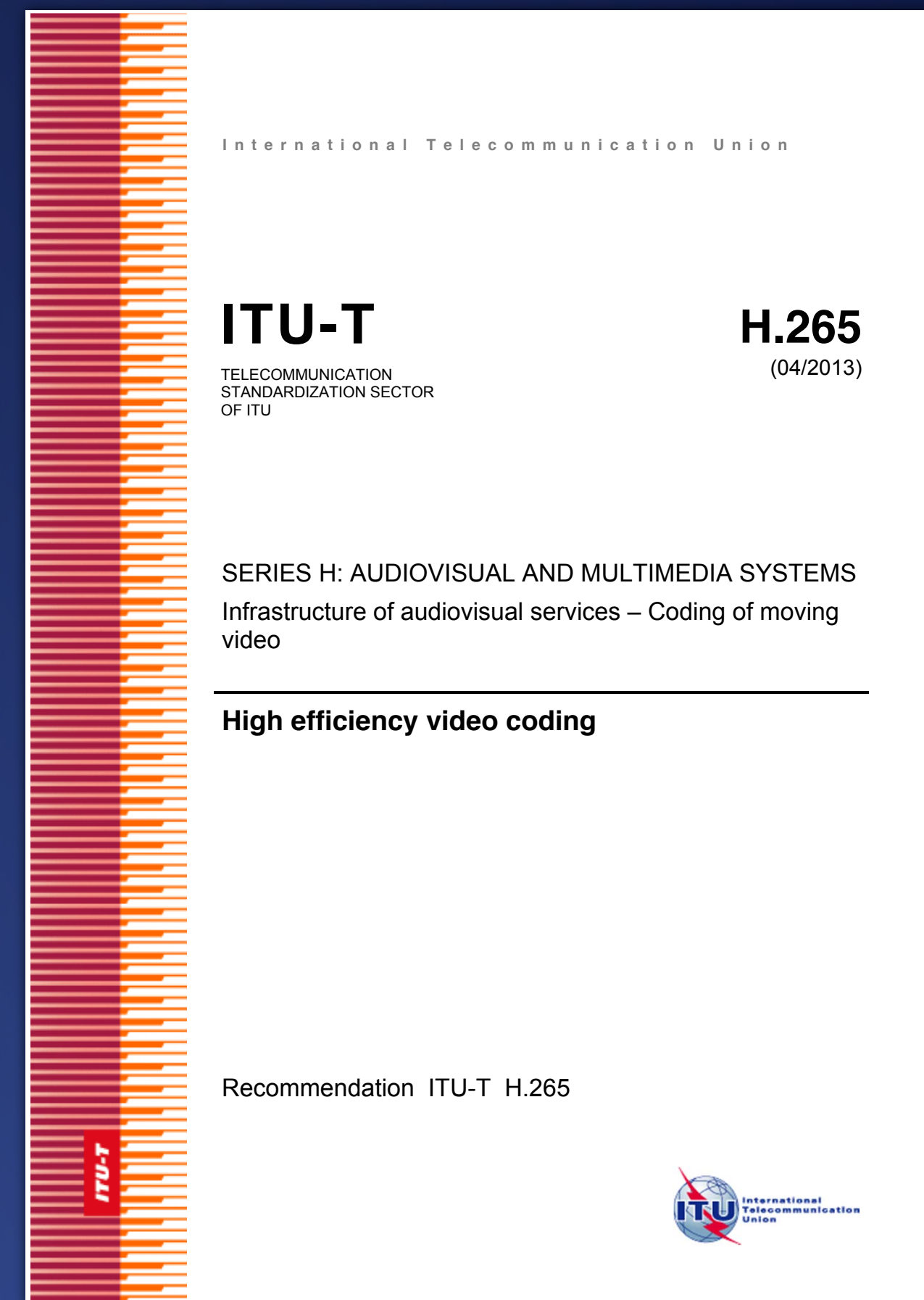
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

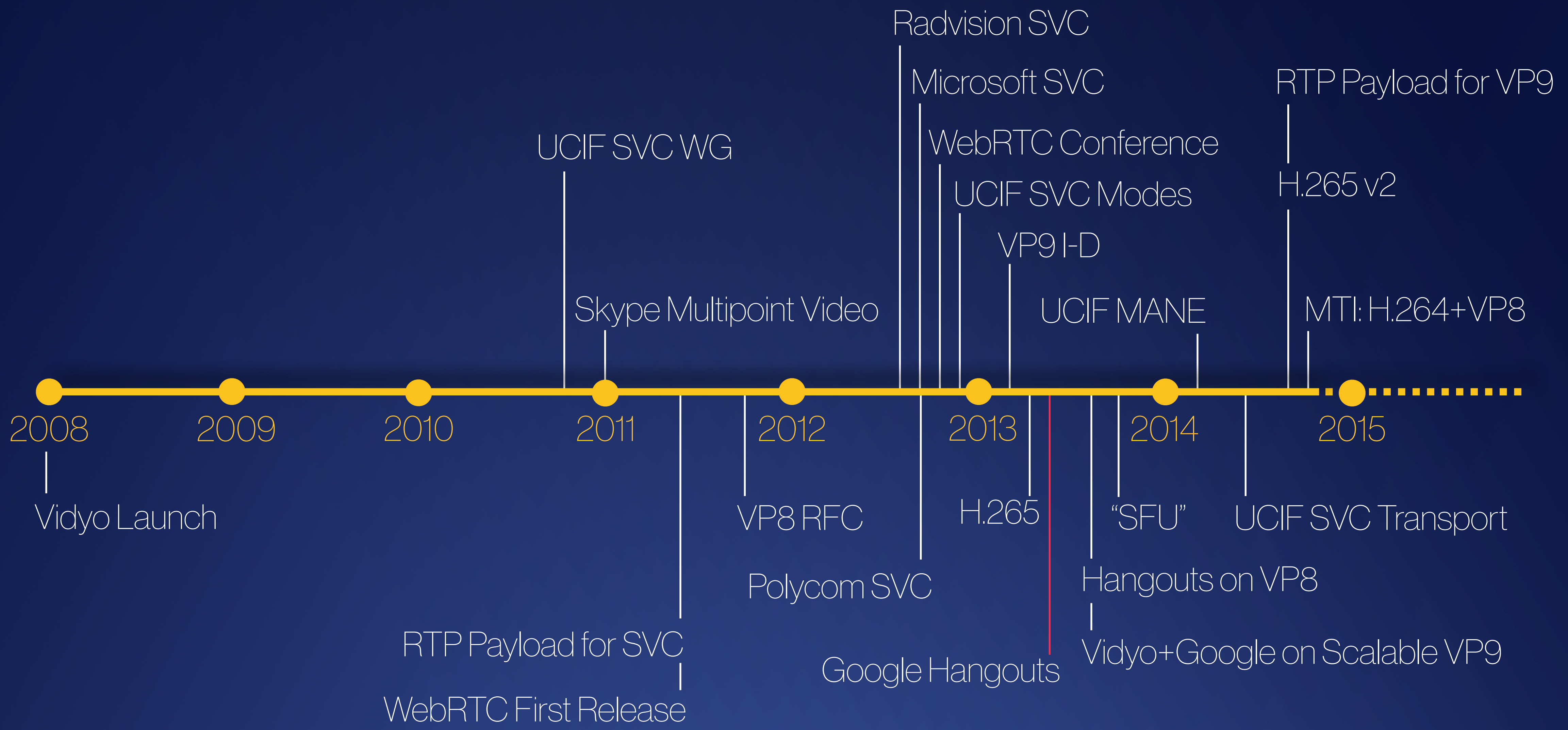
H.265

April 2013

HEVC - High Efficiency Video Coding

50% bitrate reduction from H.264
ideal for 1080p and 4K
built-in temporal scalability





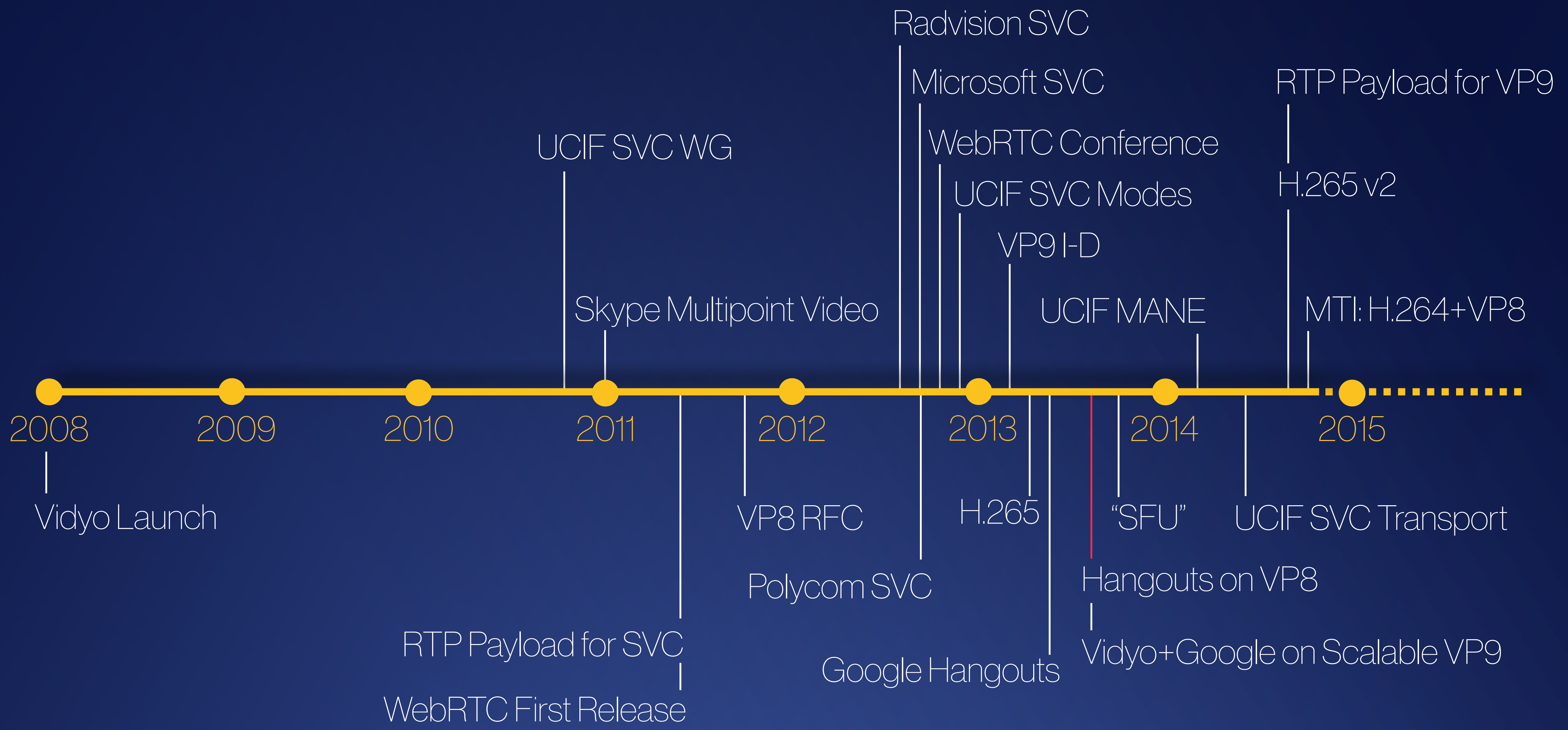
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Google Hangouts

May 2013

Using H.264 SVC



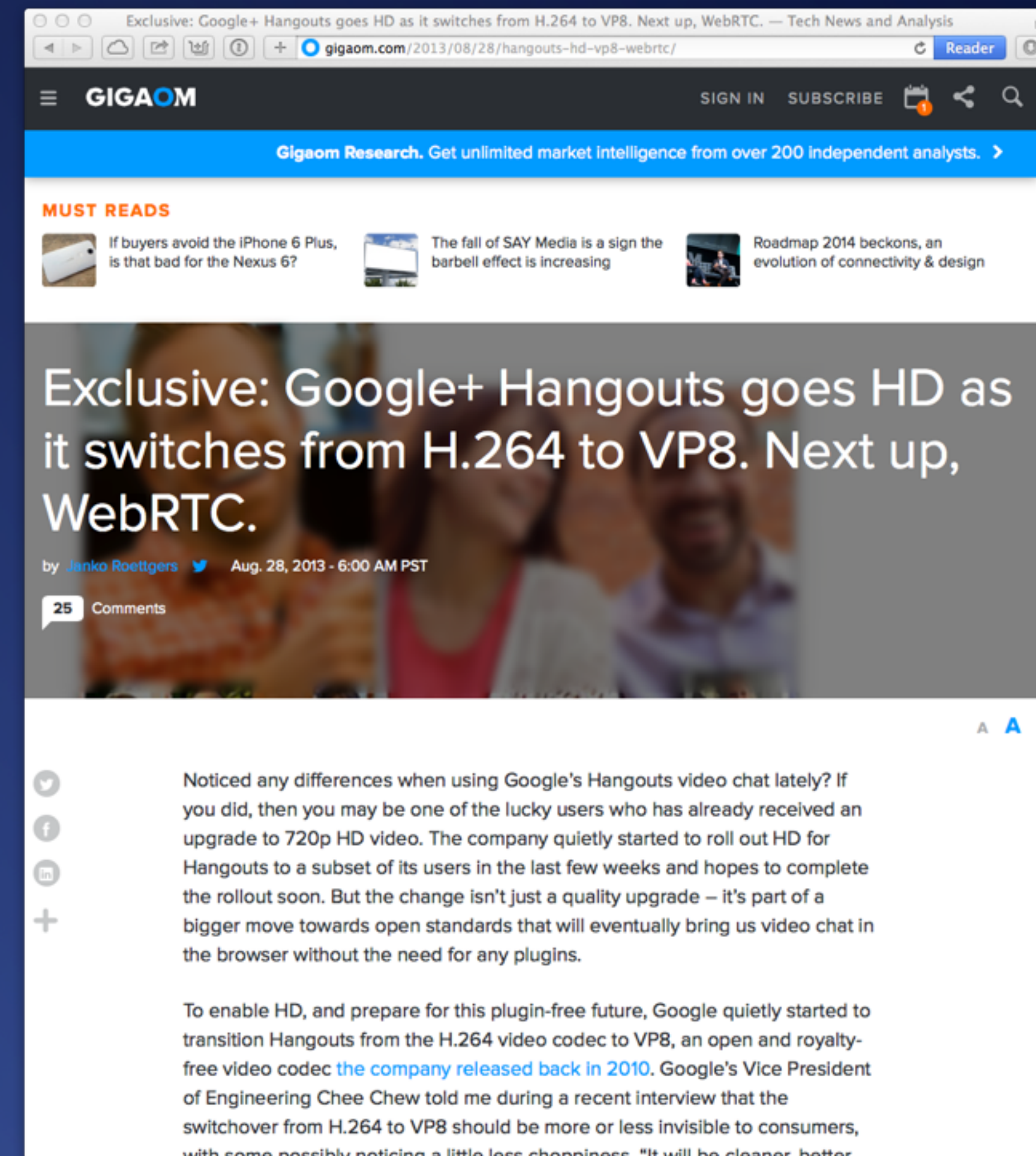


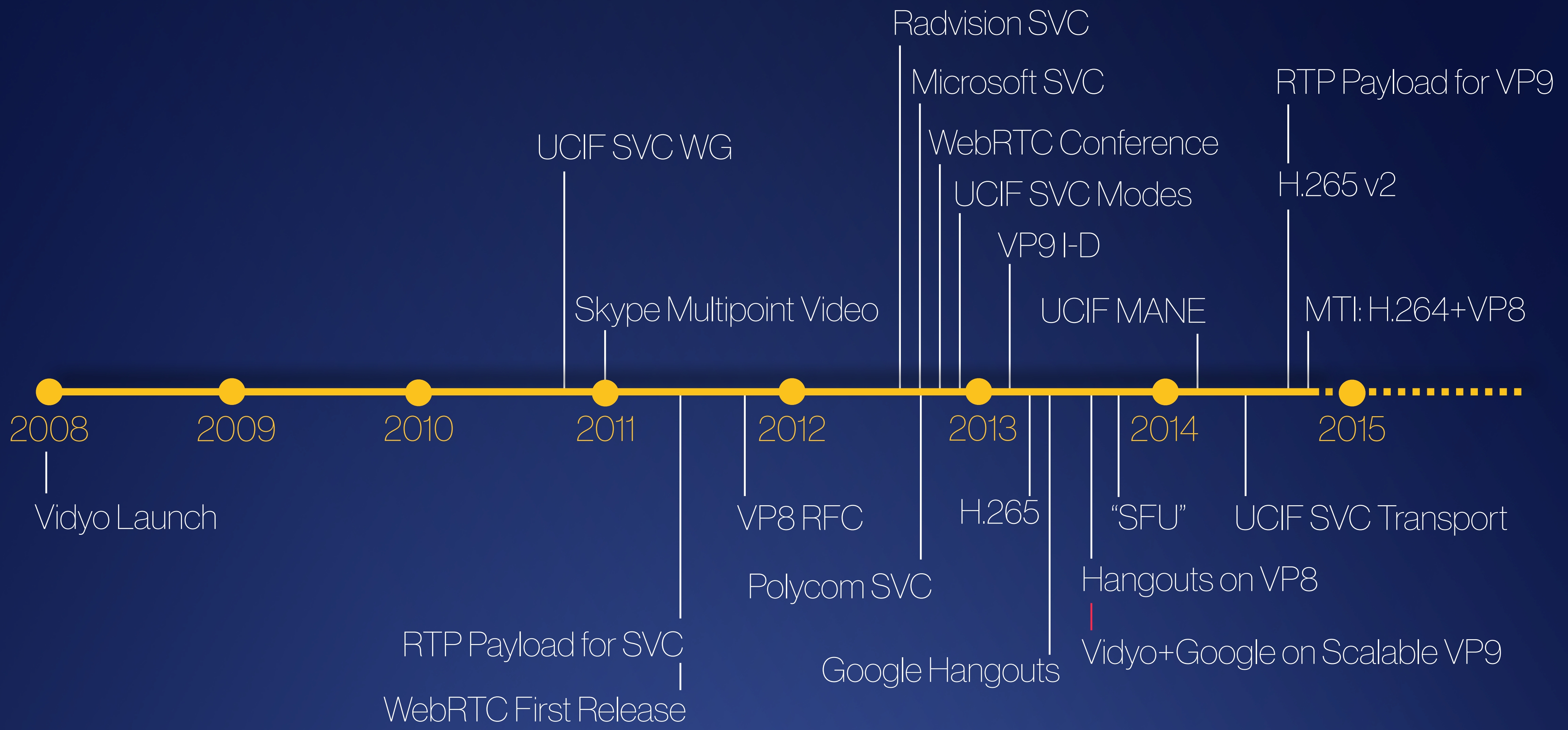
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Hangouts on VP8

August 2013

Switches to VP8



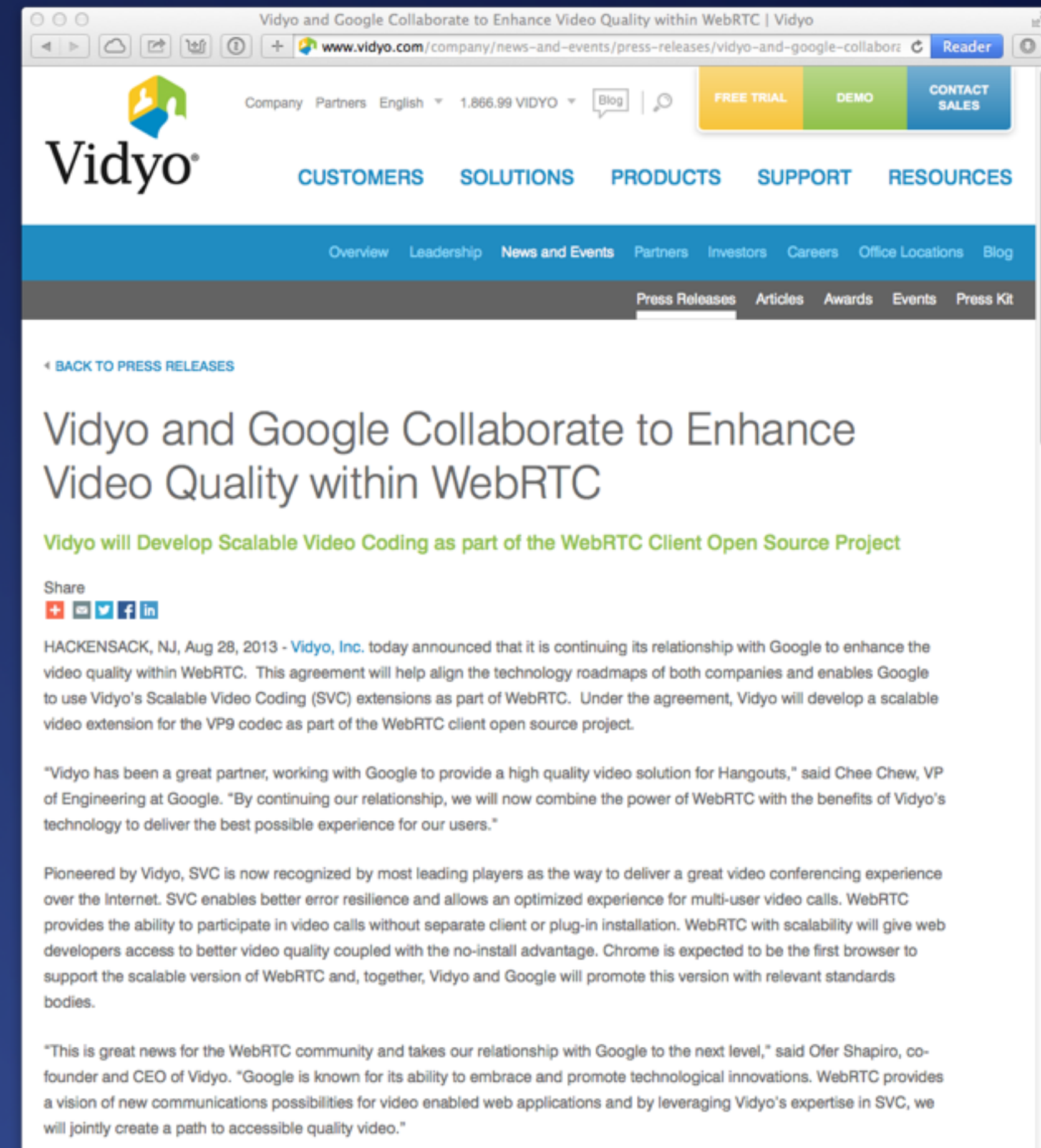


Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

Vidyo+Google on Scalable VP9

August 2013

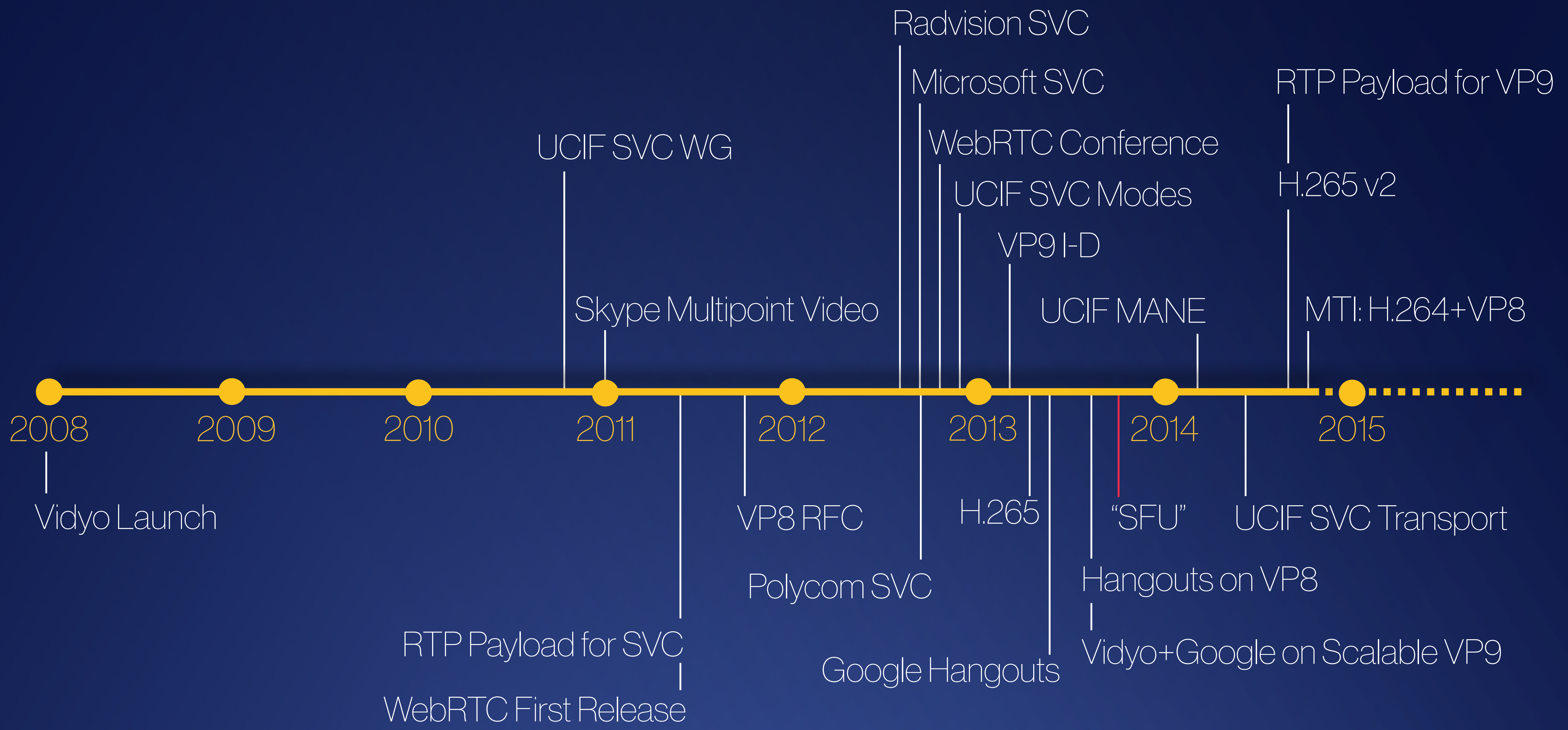
Scalable Video Coding for VP9
in WebRTC



The screenshot shows a web browser displaying a press release from Vidyo. The page title is "Vidyo and Google Collaborate to Enhance Video Quality within WebRTC | Vidyo". The URL is "www.vidyo.com/company/news-and-events/press-releases/vidyo-and-google-collaborate". The page features a navigation menu with links for "CUSTOMERS", "SOLUTIONS", "PRODUCTS", "SUPPORT", and "RESOURCES". Below the navigation, there are buttons for "FREE TRIAL", "DEMO", and "CONTACT SALES". The main content area has a sub-header "Vidyo and Google Collaborate to Enhance Video Quality within WebRTC" and a sub-headline "Vidyo will Develop Scalable Video Coding as part of the WebRTC Client Open Source Project". The article text discusses the collaboration between Vidyo and Google to enhance video quality within WebRTC, mentioning the use of Vidyo's Scalable Video Coding (SVC) extensions and the development of a scalable video extension for the VP9 codec. It also includes a quote from Chee Chew, VP of Engineering at Google, and a quote from Ofer Shapiro, co-founder and CEO of Vidyo.



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

“SFU”

October 2013
Westerlund and Wenger
*draft-ietf-avtcare-rtp-topologies-
update-01*

Introduction of the term
“Selective Forwarding Unit” - SFU

Network Working Group
Internet-Draft
Obsoletes: 5117 (if approved)
Intended status: Informational
Expires: April 25, 2014

M. Westerlund
Ericsson
S. Wenger
Vidyo
October 22, 2013

RTP Topologies
draft-ietf-avtcare-rtp-topologies-update-01

Abstract

This document discusses point to point and multi-endpoint topologies used in Real-time Transport Protocol (RTP)-based environments. In particular, centralized topologies commonly employed in the video conferencing industry are mapped to the RTP terminology.

This document is updated with additional topologies and is intended to replace [RFC 5117](#).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 25, 2014.

Copyright Notice

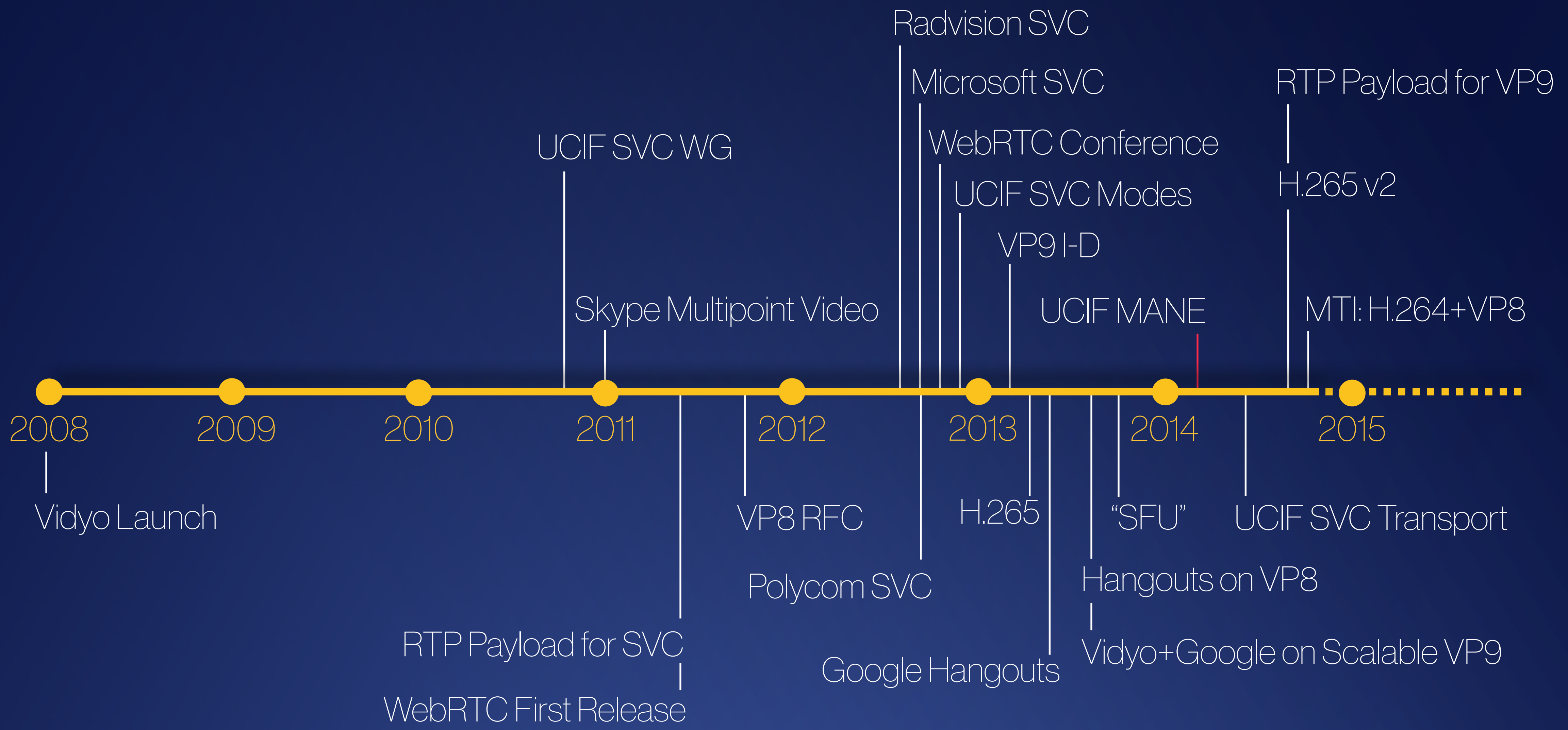
Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

Westerlund & Wenger Expires April 25, 2014 [Page 1]



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



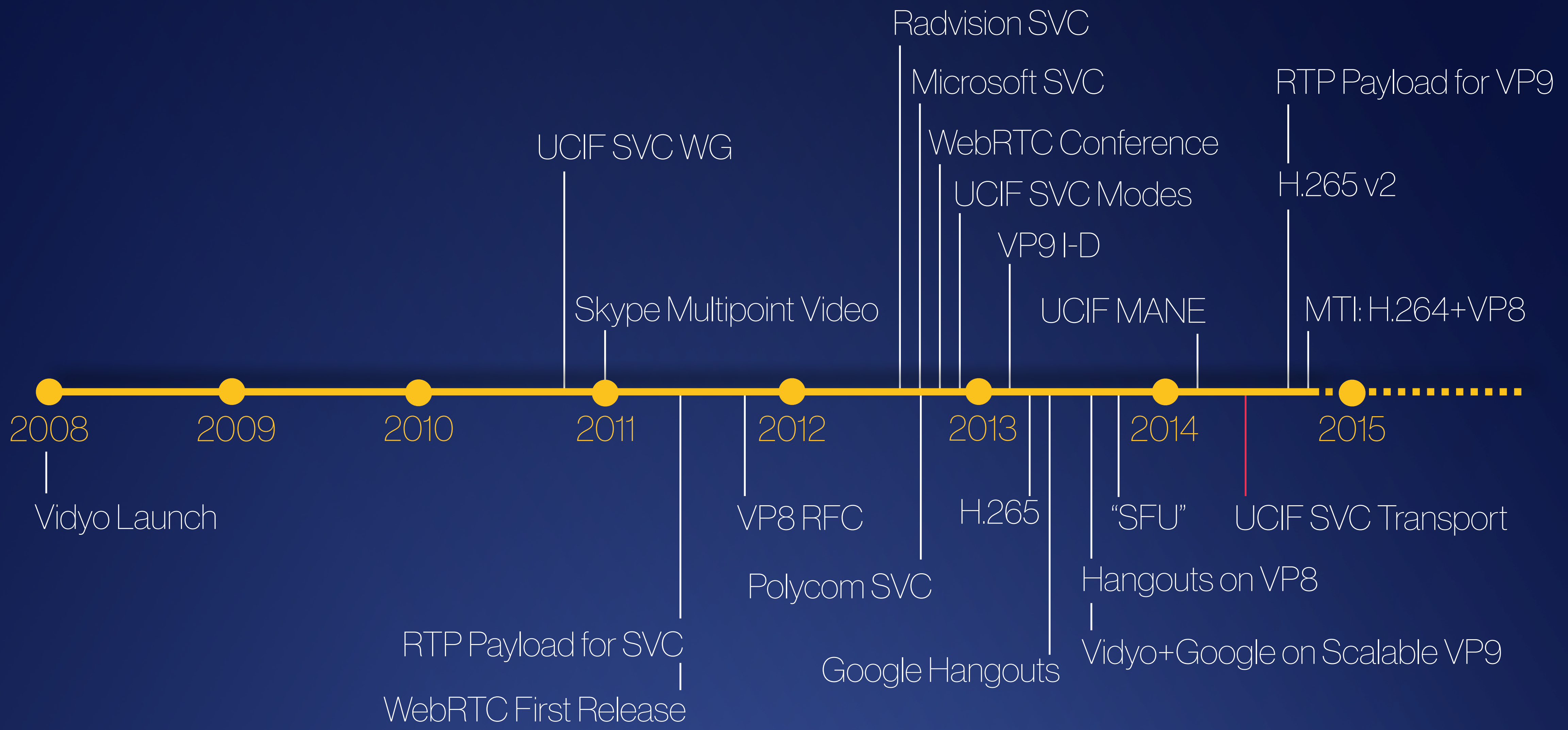
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

UCIF MANE

February 2014

Investigate SFU implementations,
as well as existing standards and
draft proposals

The screenshot shows a web browser window with the URL www.imtc.org/uc/mane-task-group/. The page features the IMTC logo and the text "The International Multimedia Telecommunications Consortium". A navigation menu includes "About", "Carrier Technologies", "Unified Communication", "Test and Certification", "Requirements and Education", "Documents", "Membership", "Events", and "Press". The main content area is titled "Media Aware Network Element (MANE) Activity Group" and includes a "Member Login" button, a "Subscribe to our interest list" form with fields for "First name", "Last name", and "Email", and a "Go" button. A "Chairs" section lists Bernard Aboba as the chair, with his title "Principal Architect, Skype/Lync, Microsoft", phone number "+425-818-4011", and email "bernard.aboba@microsoft.com". The footer contains a "Privacy Policy" link and the address "Bishop Ranch 6, 2400 Camino Ramon, Suite 375, San Ramon, CA 94583 | Phone: +1.925.275.6600".



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

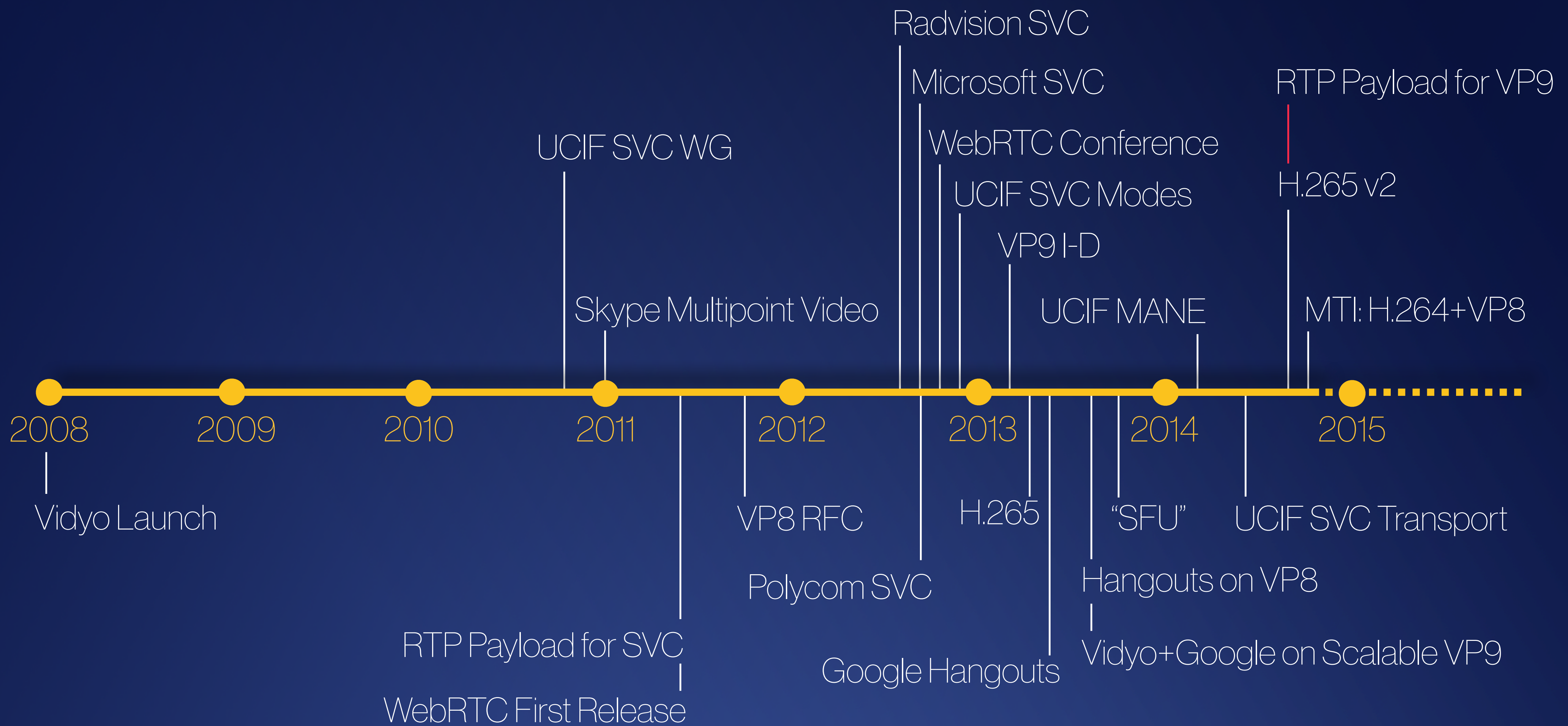
UCIF SVC Transport

June 2014

SVC RTP Transport



Unified Communication
Specification for
H.264/MPEG-4 Part 10
Scalable Video Coding RTP Transport
Version 1.0



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

RTP Payload for VP9

October 2014
Uberti, Holmer, Flodman,
and Lennox
draft-uberti-payload-vp9

How to packetize VP9 video for
RTP transport,
plus signaling parameters

Payload Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 30, 2015

J. Uberti
S. Holmer
M. Flodman
Google
J. Lennox
Vidyo
October 27, 2014

RTP Payload Format for VP9 Video
draft-uberti-payload-vp9-00

Abstract

This memo describes an RTP payload format for the VP9 video codec. The payload format has wide applicability, as it supports applications from low bit-rate peer-to-peer usage, to high bit-rate video conferences. It includes provisions for temporal and spatial scalability.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 30, 2015.

Copyright Notice

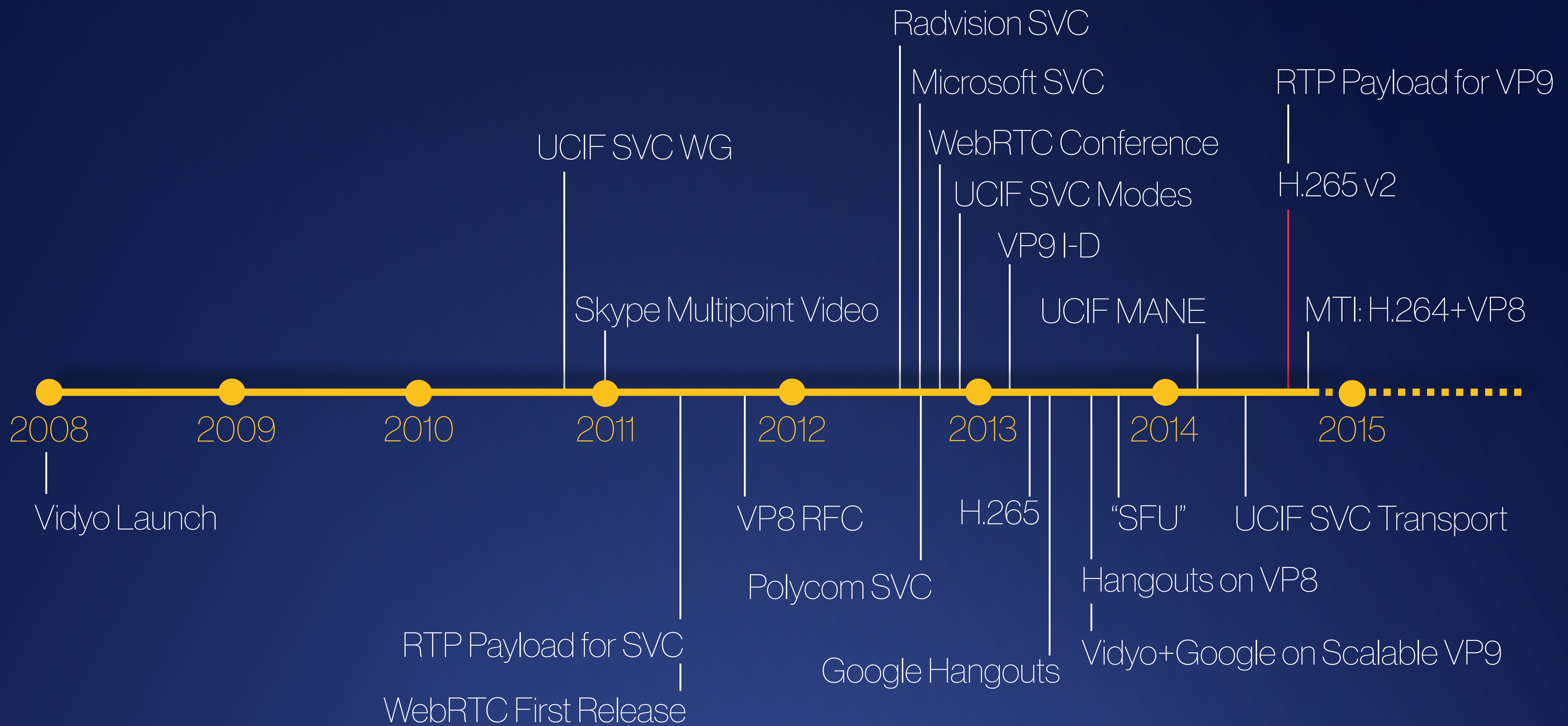
Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

Uberti, et al. Expires April 30, 2015 [Page 1]



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



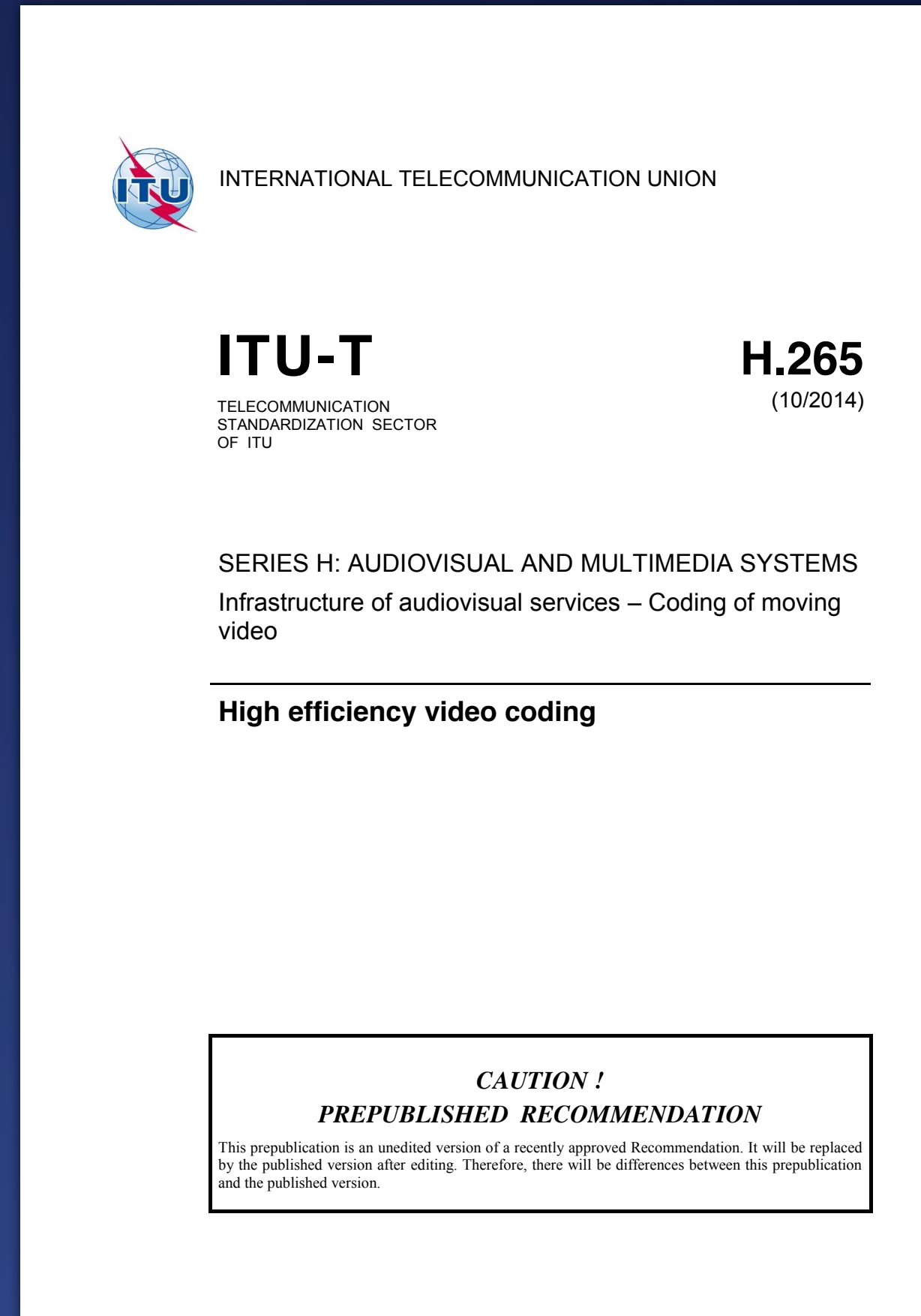
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

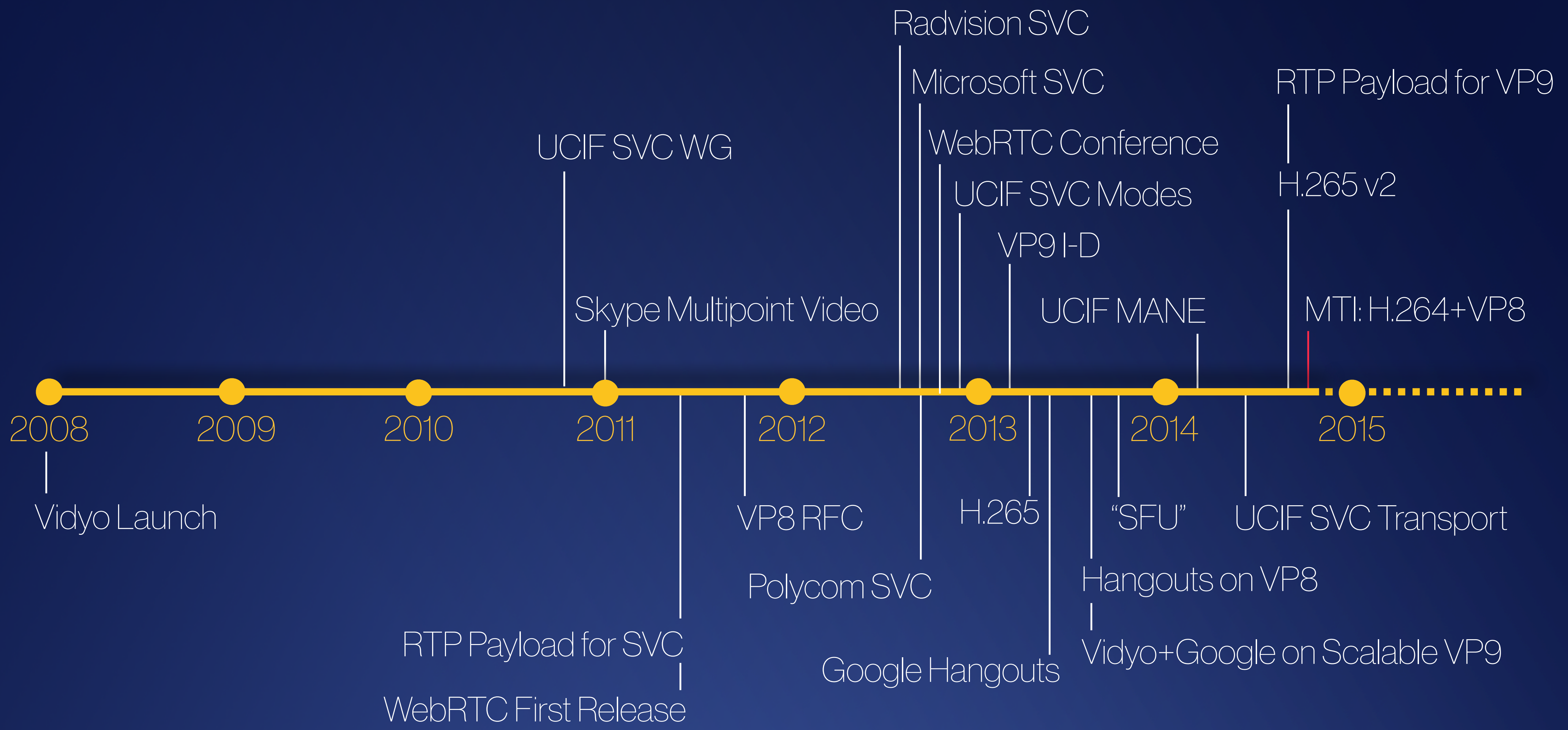
H.265 v2

October 2014

HEVC version 2

Spatial and quality scalability
extensions
(SHVC)



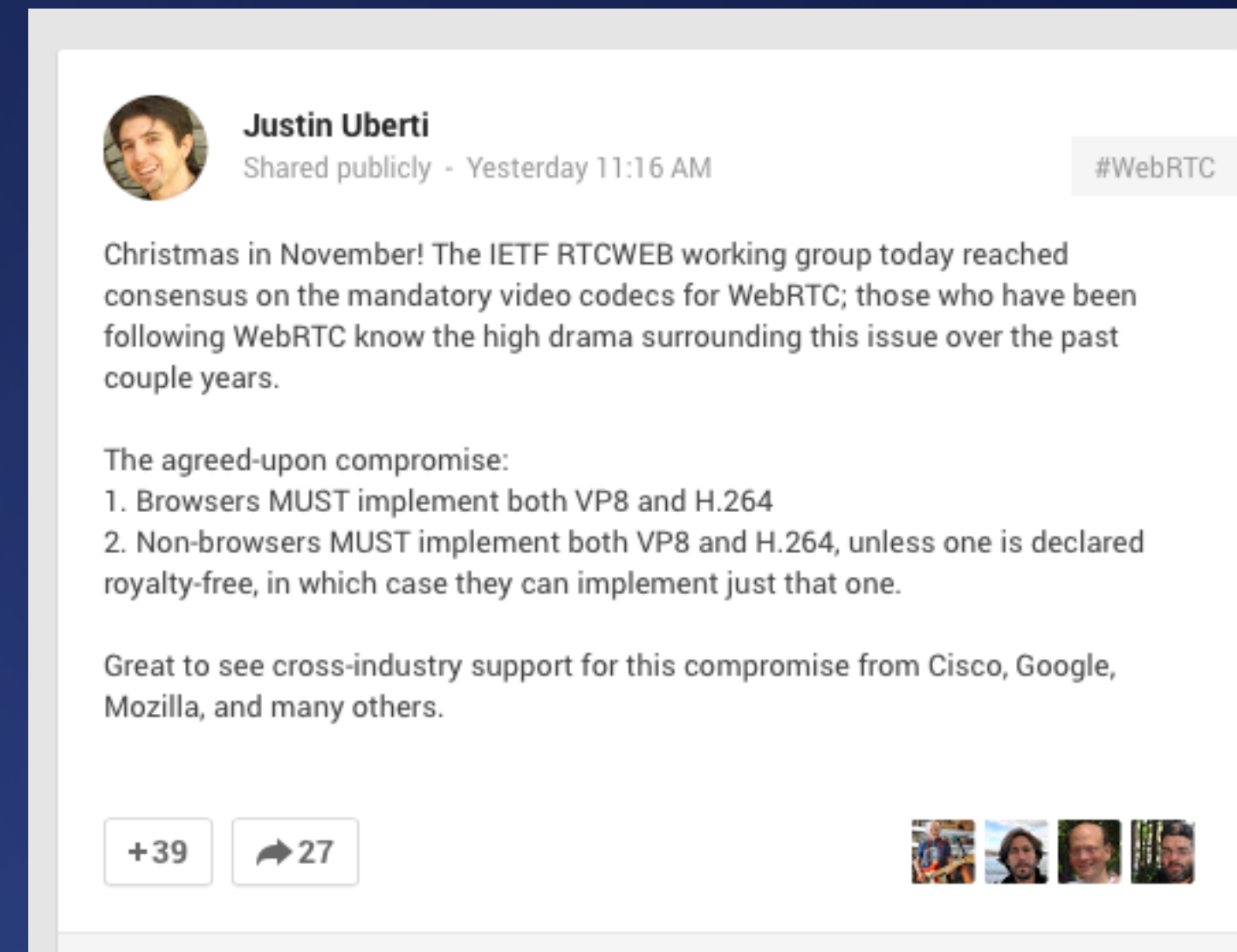


Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

MTI: H.264+VP8

November 13, 2014

Mandatory-to-Implement Codec:
both H.264 and VP8



Justin Uberti
Shared publicly · Yesterday 11:16 AM #WebRTC


Christmas in November! The IETF RTCWEB working group today reached consensus on the mandatory video codecs for WebRTC; those who have been following WebRTC know the high drama surrounding this issue over the past couple years.

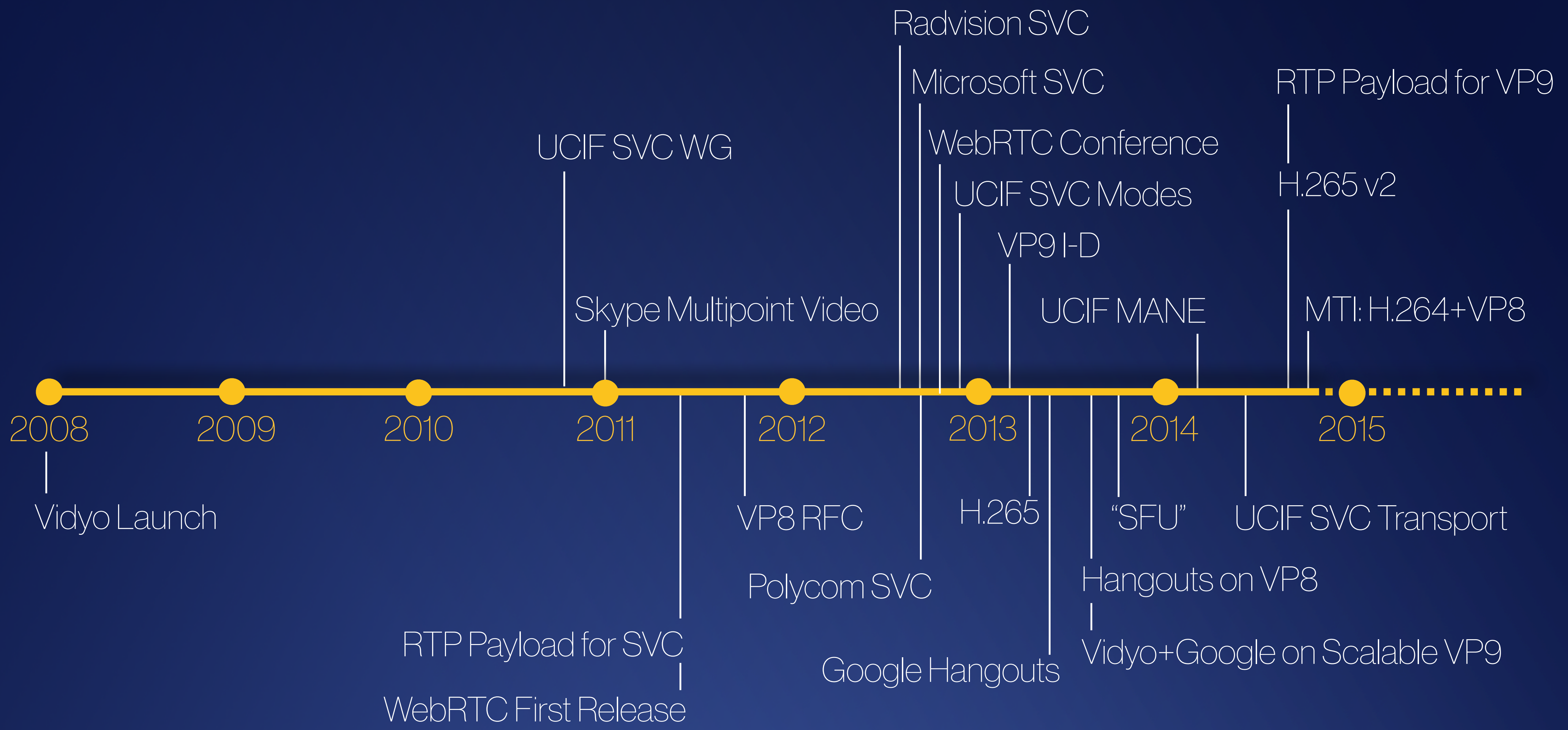
The agreed-upon compromise:

1. Browsers **MUST** implement both VP8 and H.264
2. Non-browsers **MUST** implement both VP8 and H.264, unless one is declared royalty-free, in which case they can implement just that one.

Great to see cross-industry support for this compromise from Cisco, Google, Mozilla, and many others.

+39 ↪ 27

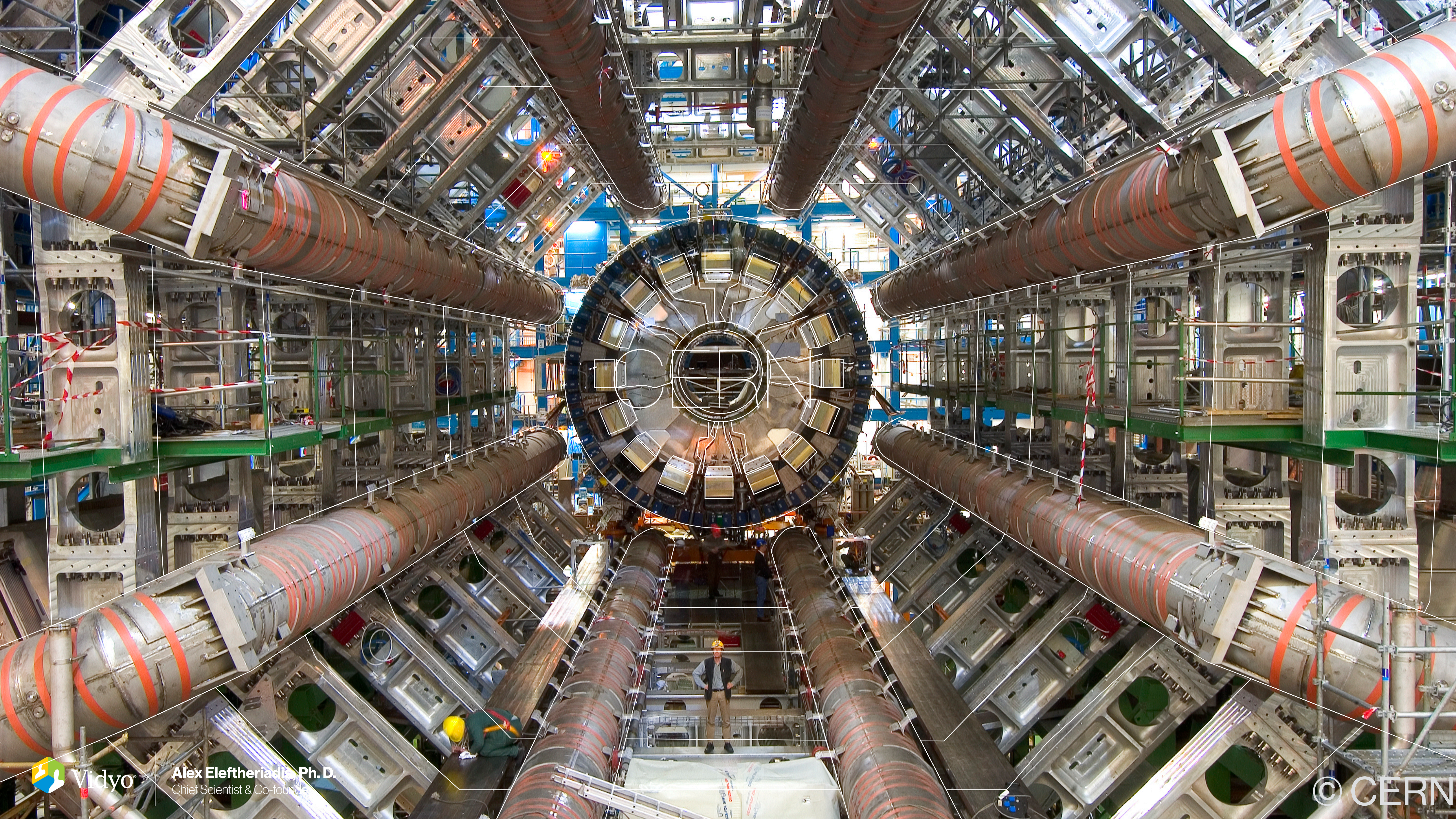




Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder



Vidyo at CERN

~ 18,000 users

800

simultaneous connections

up to **168** simultaneous **H.323/SIP**

12 phone access points worldwide



2 Vidyo Panoramas
(CERN site)

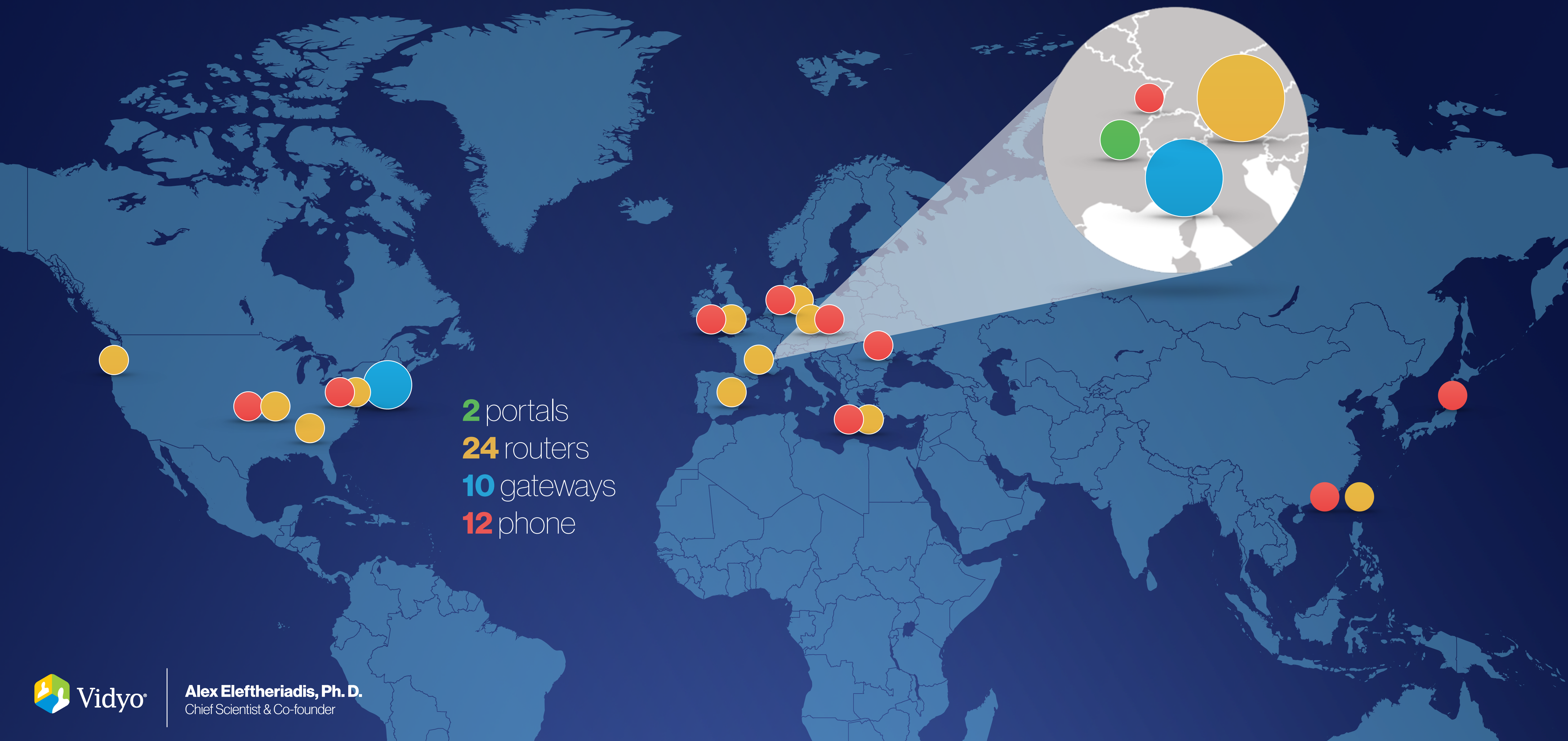
12 simultaneous recordings

CERN Vidyo Worldwide Service Topology

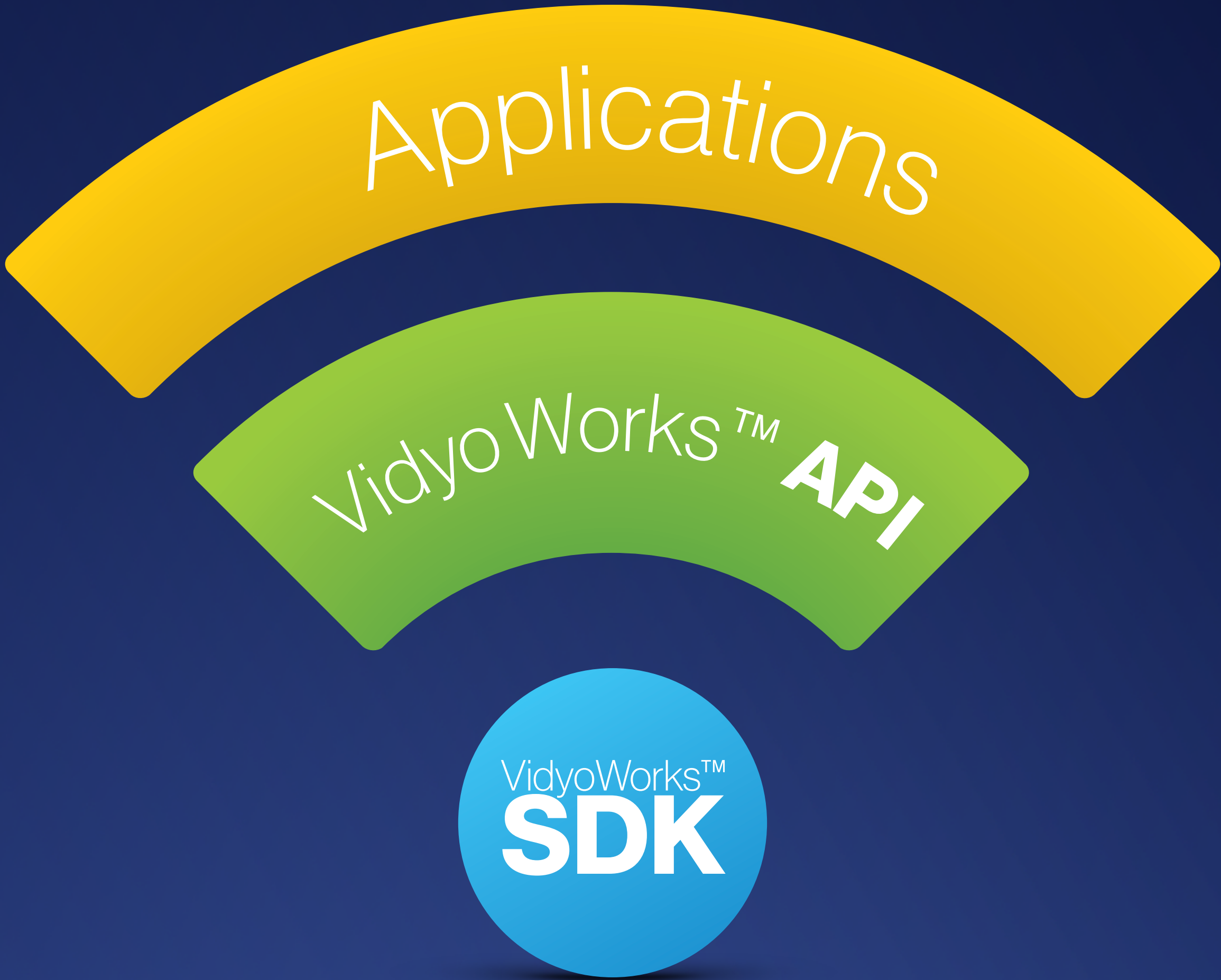


Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

CERN Vidyo Worldwide Service Topology



2 portals
24 routers
10 gateways
12 phone



WebRTC & VidyoWorks



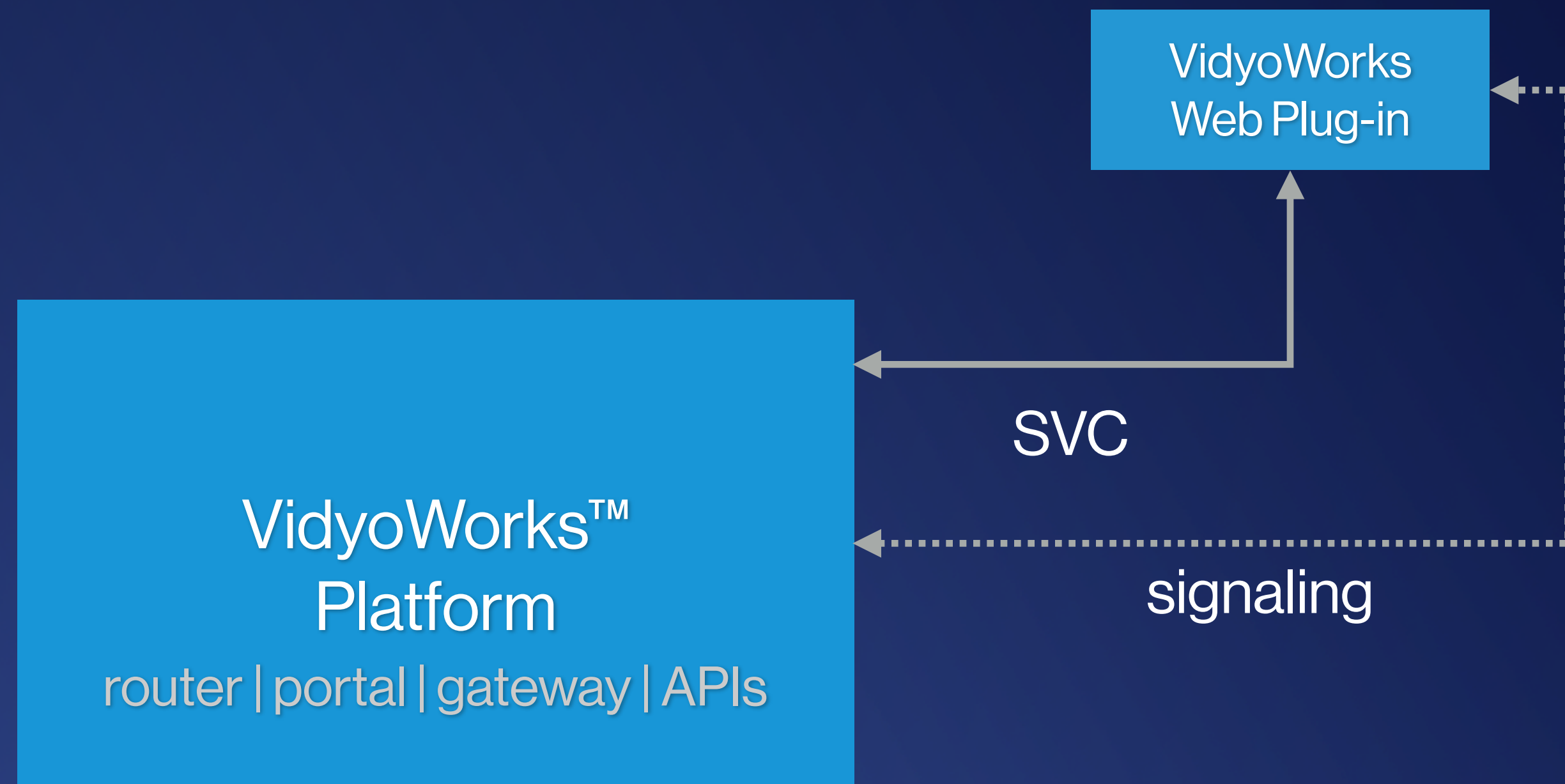
Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder

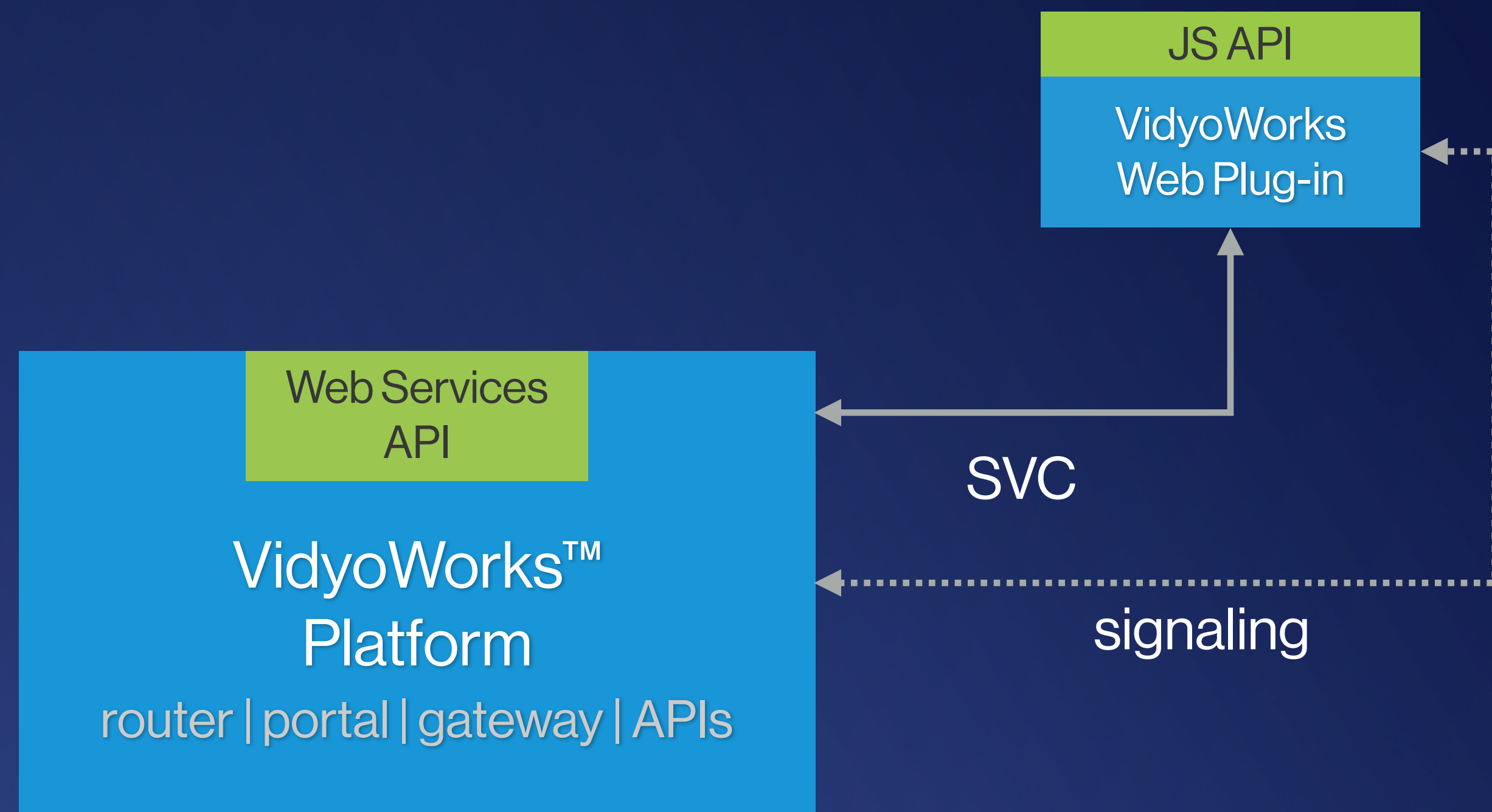
VidyoWorks™ Platform

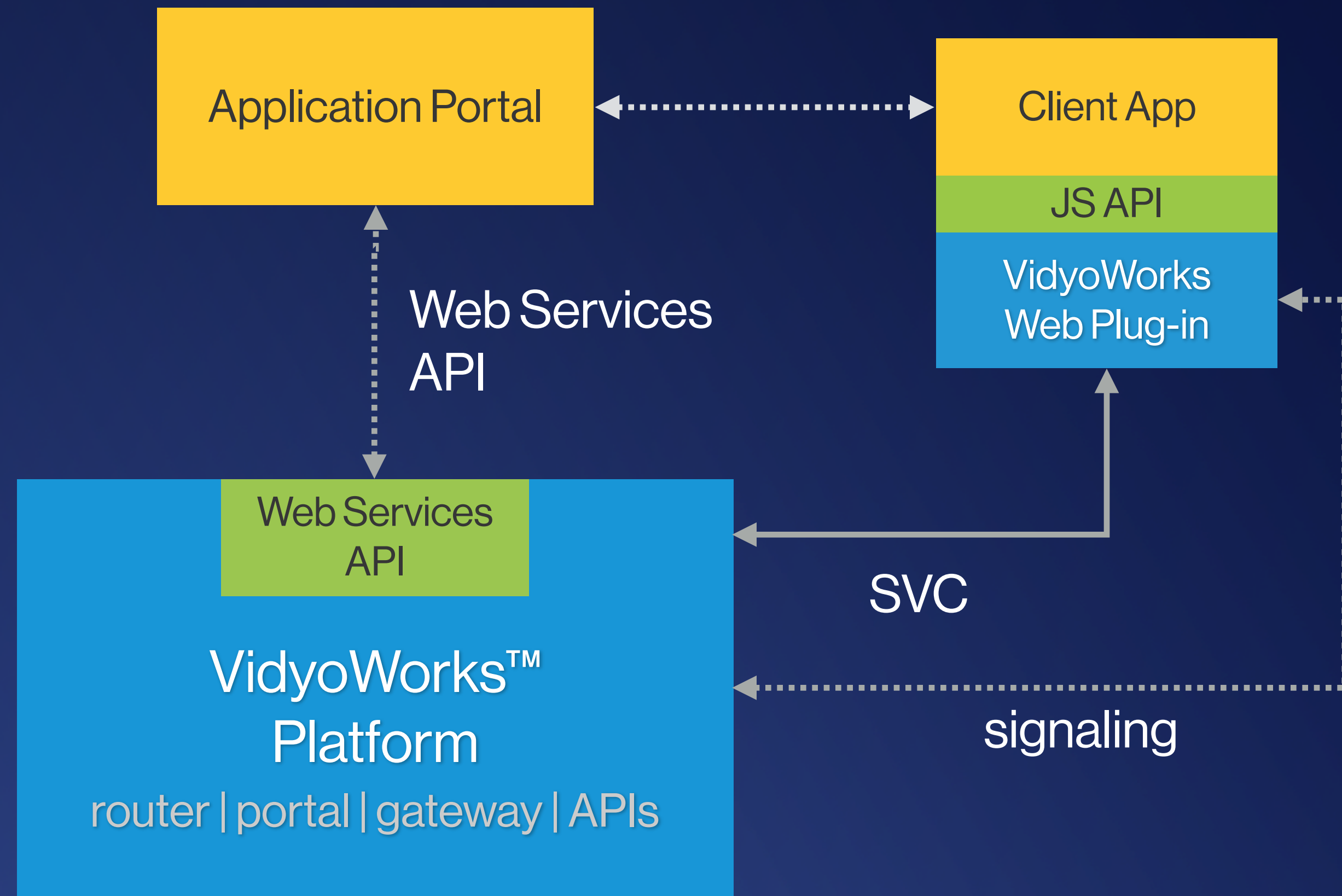
router | portal | gateway | APIs



Alex Eleftheriadis, Ph. D.
Chief Scientist & Co-founder







Application Portal

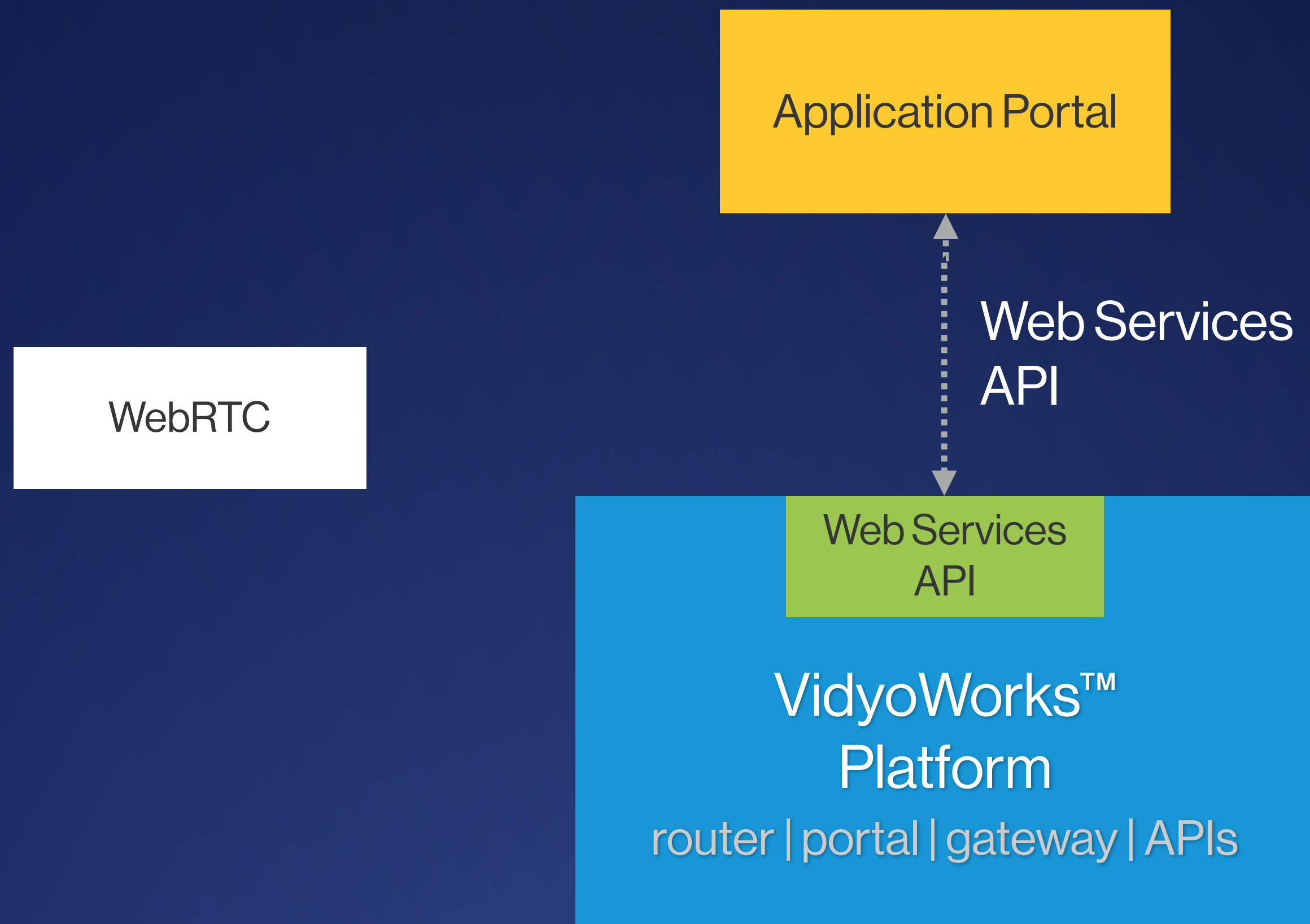
Web Services
API

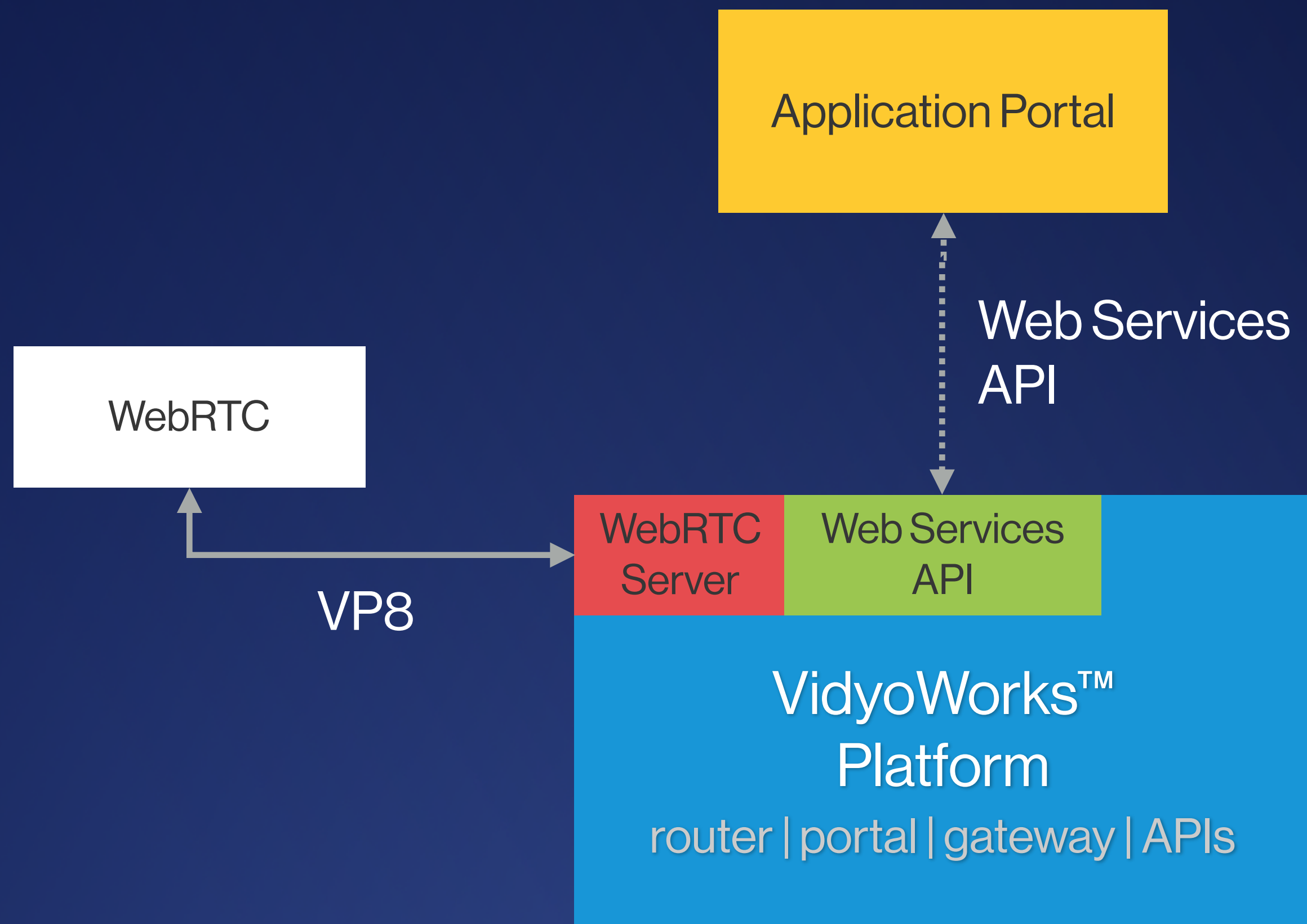
Web Services
API

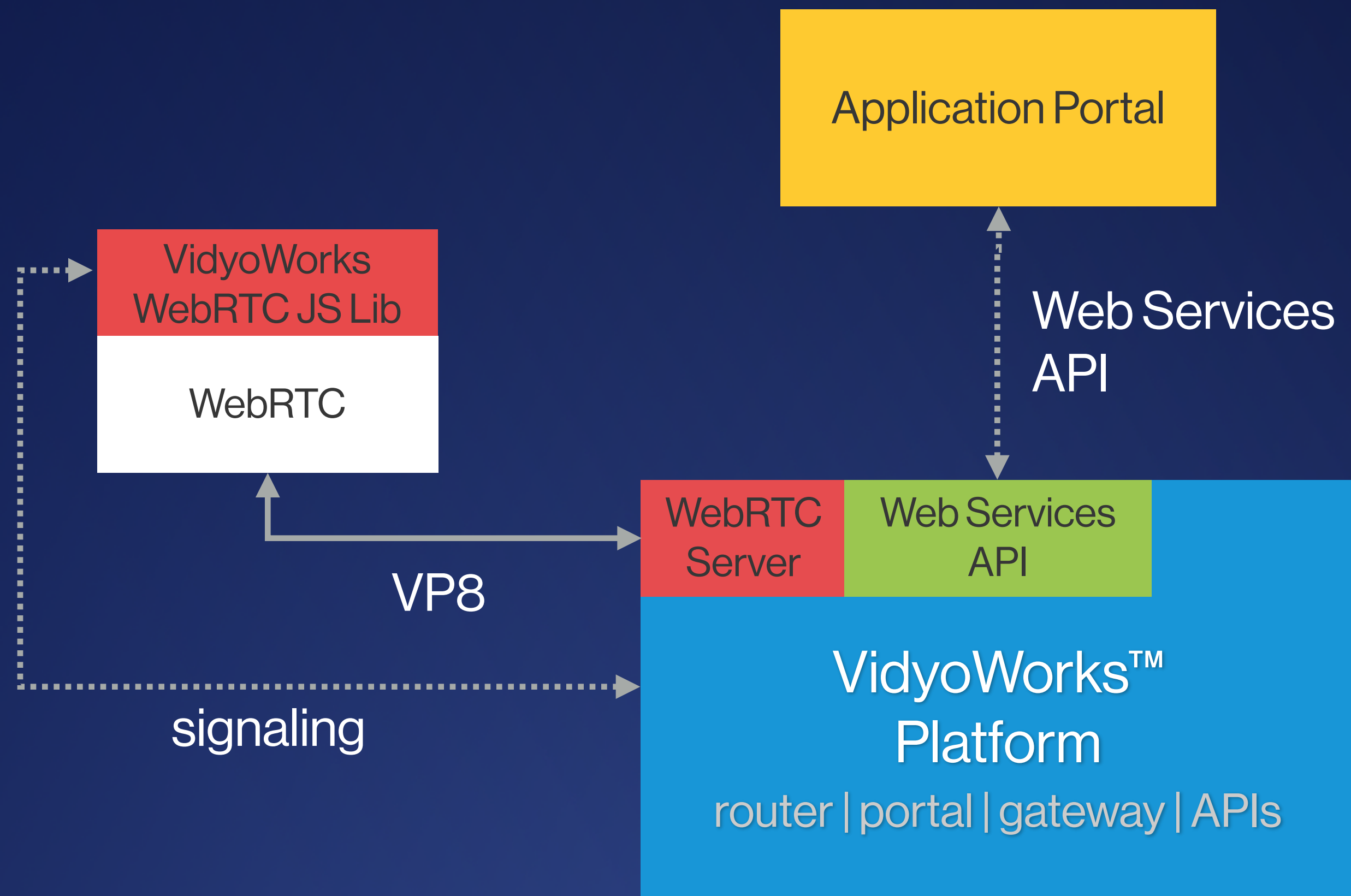
VidyoWorks™
Platform

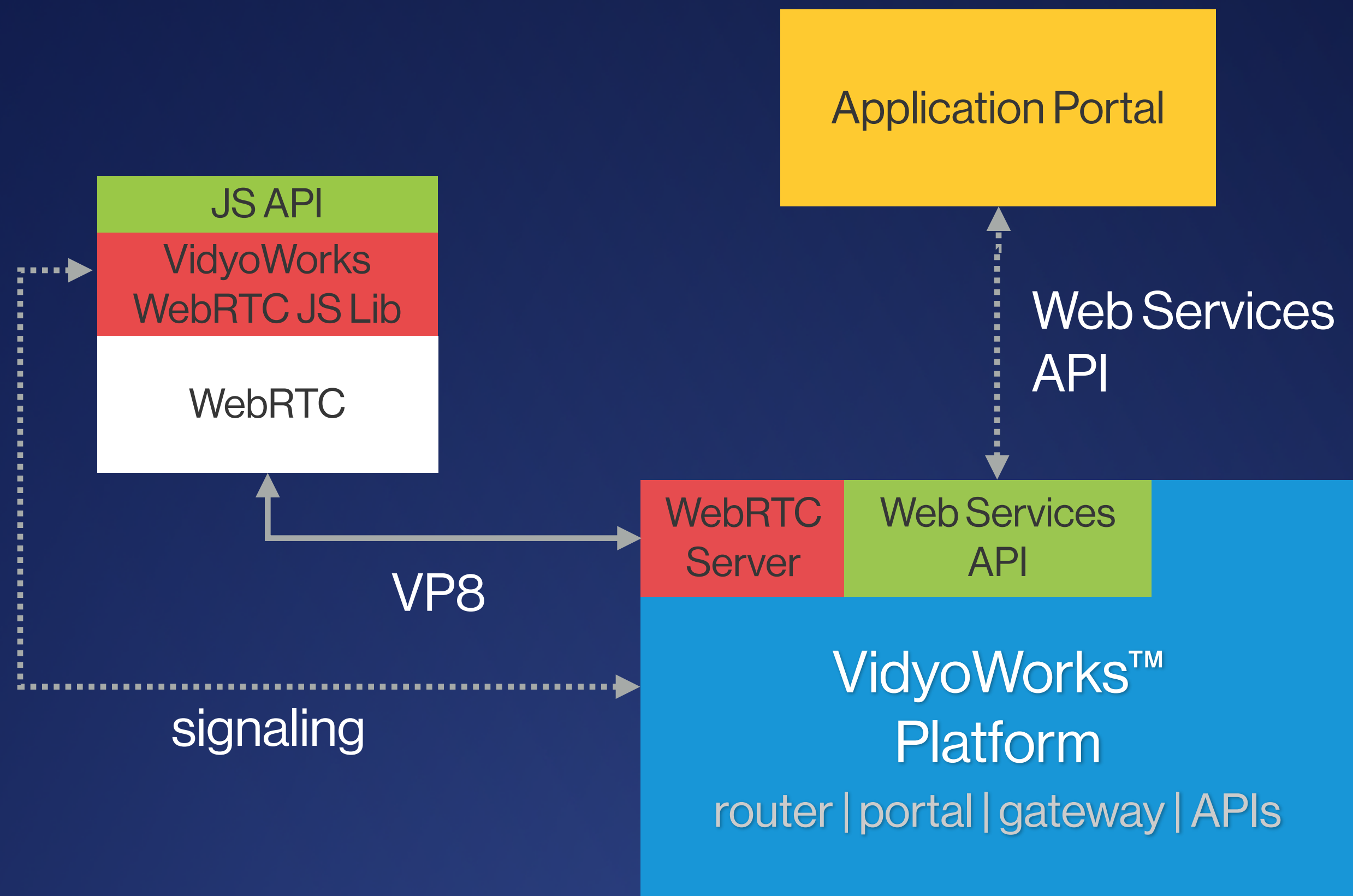
router | portal | gateway | APIs

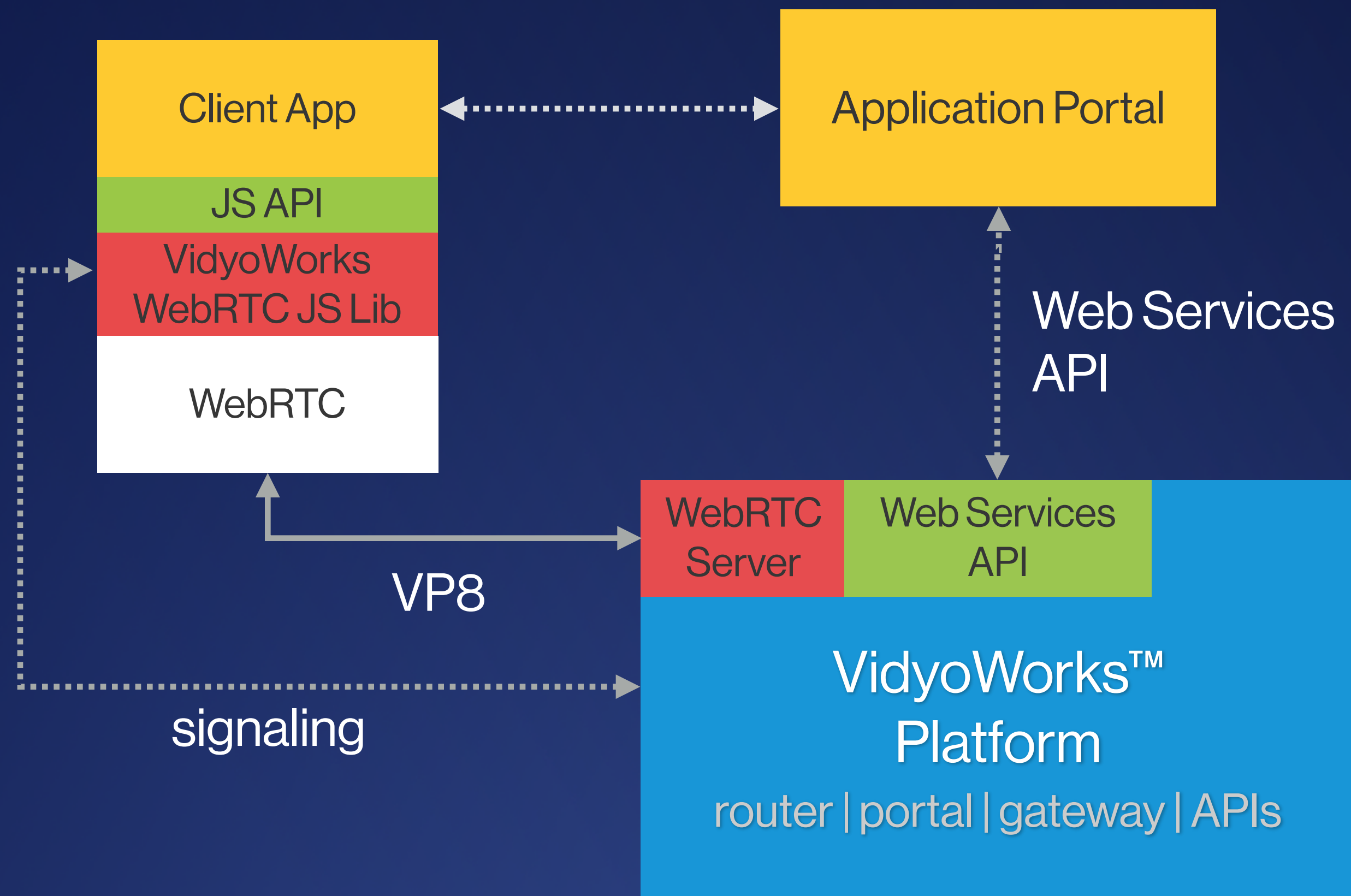












Version 3.0 with sample API released Nov. 2014

Version 3.2 with Vidyo-branded client
to be released May 2015



Alex Eleftheriadis, Ph.D.

alex@vidyo.com



Vidyo®