Oral History of Les Kohn

Interviewed by:
Douglas Fairbairn
Uday Kapoor

Recorded via Zoom January 26, 2021

CHM Reference number: X9400.2021

Kapoor: Welcome, Les Kohn, to this oral history interview. My name is Uday Kapoor. I am a member of the Computer History Museum's Oral Histories Program. With me is my colleague, Doug Fairbairn, and we will both be interviewing you, Les. So welcome.

Kohn: Thank you.

Kapoor: So first, a short introduction. Les Kohn is very well-known in the Valley as an architect's architect and an entrepreneur. He was born in Oakland in 1956 and grew up in San Mateo. And that's really all that I have in your early part of your life, but we can get started with that. It seems you were a prodigy. You were not quite four years old when your dad taught you to how to use a slide rule. So—

Kohn: Well, yes, kind of.

Kapoor: So, please tell us a little bit about your early childhood.

Kohn: Well, my dad was an engineer. Actually, he worked at a company designing high-performance mechanical calculators at the time. So, he was always showing me various gadgets, including the slide rule. I think, at that point, I didn't really understand what multiplication was, but he just showed me how to line up the various scales and what the results looked like.

Kapoor: So, for you, it was like a toy.

Kohn: Yes.

Kapoor: That's wonderful. And you, in fact, within a few years after that, you started with your personal computer. And we can talk a little bit about that. A three-bit Digicomp.

Kohn: That's right. That was another one of these toys of that era, which-- it had like three sliding mechanical flip flops, and they had these little pegs that you would stick onto different positions, and another sliding thing that was like a clock, and you could cause it to cycle through various states.

Kapoor: So, tell us about your family then. You mentioned your dad. How about your siblings? Were they all like you? What kind of schooling did you get? So talk a little bit about your early life.

Kohn: Yes, sure. Well, I have two brothers, both younger than me. One of them is also an engineer in Silicon Valley, working at Juniper, and the other one is programming, but took a little bit different path to get there. When growing up, we went to public schools and, luckily, they did have a nice gifted program which I took advantage of to get some early exposure to computer stuff.

Kapoor: Okay. So it looks like you had real desire and a fascination with computers from early life.

Kohn: Yes.

Kapoor: You said, there was a poster of a CDC 6600 on your wall.

Kohn: Right. Which I think I got from a neighbor that was throwing it away. Yes, it was something, I guess, I was always attracted to. I can also recall reading articles in National Geographic about the ILLIAC IV, this massive parallel computing system, and I was always intrigued by these things with big consoles of flashing lights and that kind of stuff.

Fairbairn: Right. Les, were you-- did you get into any other mechanical stuff? Erector sets? Anything else?

Kohn: Sure. I mean, I had the Lego blocks and, yes, I had this-- I forget the-- yes, the erector set, that kind of stuff.

Fairbairn: And was your father sort of the major influence in that? Is that something you just sort of picked up?

Kohn: I mean, certainly he was an influence, and I think it was-- it was just always something that I can recall being interested in, in science and math and that kind of stuff.

Kapoor: So you mentioned that you played with a six relay and a motorized rotary switch computer in school.

Kohn: Right.

Kapoor: Can you tell us a little bit about that?

Kohn: Yes, that was in sixth grade, in this gifted program, and they-- they bought this thing. I don't exactly remember how we found it, but it was what you could get at that time for learning how computers work. It had these six relays which you could hook up with like diodes and resistors to generate different kinds of logic operations, and then there was this-- the output devices were a set of lights plus this mechanical switch that had a motor on it, and you could program it such that it would turn and then stop at a particular position, and that was your output. And the whole thing was connected together with patch cords.

Kapoor: I see. So what was the gifted program like? Who were your colleagues? Did you stay in touch with them until later?

Kohn: Yes. It was a program-- basically the whole school district-- San Mateo School District-- they would-- I guess they'd do this testing and then select. I think we had about thirty kids for that class. And then they bussed them all to a central school where we could all be together in class, and then they'd offer special enhanced stuff. So like there was some math teacher that they brought in from the high school to teach us some advanced math stuff, and like they gave us these extra toys to play with and so on. And I do-- I made one really good friend there who I actually went to school with at Caltech. And we--

we've kept in touch. He's in New York, now, so I don't talk to him regularly. Definitely there were some interesting people there.

Kapoor: I can imagine. You mentioned something about that they brought in a math teacher. Were there other exposures to programming, for example?

Kohn: So beyond this computer, in the seventh grade, they also introduced me to someone who was--had contacts with Stanford and was able to get me into their 360/67 humongous data center. I don't exactly recall how I picked it up, but I wrote a Fortran program to compute prime numbers and punched it. They took me there. I had basically one day to go in there, key punch it, and then run it. It was all batch, right? So you run it. Something doesn't work. You get back this printout and you--

Kapoor: Right.

Kohn: --iterate on that until you get something.

Kapoor: You mentioned also a Heathkit analog computer kit that you played with.

Kohn: Right. That was in eighth grade. The school district-- I was kind of into electronic kits, at that time. I also had a-- my dad bought me an oscilloscope kit and a multimeter. And then there was this analog computer that the school district bought, which also was all vacuum tube based. So, it's basically these vacuum tube op amps and it had a meter on it that was the output device, and you could program it to solve differential equations and that kind of stuff.

Kapoor: Right. So, in terms of overall concepts, in terms of theoretical concepts, you were getting that training in school, and then you were curious, and you were applying it to all these experiments. Is that how it was working?

Kohn: Yes. I mean, it wasn't so much that there was a lot of training, because, at that time, the teachers were not really up on computer stuff. But I got exposed to-- I got to play with these computers, and I kind of picked up stuff on my own. You know, read books and that kind of stuff.

Kapoor: Okay. You must have done a lot of reading. Yes, of course.

Kohn: Yes.

Kapoor: You developed computer games and simulations on IBM 1620 in the junior college.

Kohn: Right.

Kapoor: PDP 8/11, IBM 360.

Kohn: Right. Yes, that was another lucky thing that-- our junior college had a couple of computers at that time, and they-- one of them was a [IBM] 1620, which was a scientific computer and also programmed with punch cards. But, in that case, you could go in and play-- sit at the console of the computer and run programs. And it had kind of a Selectric typewriter output device that you could interact and type stuff with. And the other great thing about that computer was it had-- it was a decimal computer, and it had all the multiply and add tables stored in the core storage. And you could watch-- you could give it this humongous multiply and you could see on the lights how it was cycling through the memory when was doing multiplies and stuff.

Fairbairn: But you were doing all this work when you were in high school, right?

Kohn: Actually, I think this started in eighth grade. Yes. Right before high school.

Fairbairn: You had-- you were fortunate to get access to a lot of--

Kohn: Yes, I was really lucky.

Fairbairn: --computing stuff and other things. That's great. Were there any other major developments in high school? Teachers or whatever that kind of steered you in certain directions?

Kohn: Well, in high school, they had a very good math/science department. And I also got early access to the junior college math courses, so I took calculus in tenth grade and second-year stuff in eleventh grade, before graduating. So that plus access to-- DEC had an office in-- I don't recall whether it was in Mountainview or something, but, anyway, I managed to get access to some of the early PDP-8, PDP-11 computers. And then our school district bought a really low-end 360 system that they used to run timesharing BASIC across all the schools in the school district.

Kapoor: That's wonderful.

Fairbairn: So, it sounds like from math and other studies you were two or three years ahead. You were already doing second-year college in high school in eleventh grade.

Kohn: Yes, probably. And physics was another really-- passion that I got involved in. And then, when I went to Caltech, I--

Kapoor: Yes, so that's an interesting segue, because I noticed that, and I was going to ask you how did you-- what inspired you to get into physics?

Kohn: Fundamentally, I like to look at-- understand things at a deep level. At some point, I realized that there's kind of two worlds that humans are exposed to. There's the external world, which is basically described by physics at the most fundamental level, and then there's the internal world, what's going on inside your brain. That's actually more closely related to what computers are going. How they compute things. So I guess that's how I ended up on these two branches, physics and computers.

Kapoor: Great. How about mathematics?

Kohn: Well, mathematics was something that I would say-- I wasn't so interested in pure mathematics, but as a tool for understanding physics and computing.

Fairbairn: How did you end up going to Caltech? Were there other choices?

Kohn: Yes. Yes. Caltech, of course, was well-known for physics, being a top physics school. It's also a small school, which allows you, as an undergraduate, to have a lot more access to faculty and research stuff. So that was attractive. I applied to MIT. Got into MIT, but I just decided to select Caltech because of the size and the physics orientation.

Kapoor: So you were not so much curious about physics in terms of modern physics or theoretical physics, but applied physics.

Kohn: No. No, actually I was quite interested in theoretical physics.

Kapoor: I see.

Kohn: Because my dad gave me a book. I don't know whether it was junior high school or not, but it was a book on quantum mechanics.

Kapoor: Okay.

Kohn: And the whole-- it was quite a shock to me because I think in like fifth or sixth grade, they showed these atoms with these little balls of electrons spinning around. And when I found out that that whole model just doesn't work, that was quite a shock and really got me interested into fundamental theoretical physics.

Kapoor: Right. But it seems you didn't pursue that later, right?

Kohn: Right. Not after graduating. Because I didn't realize-- when I was going to school at Caltech, to really do physics professionally requires-- first of all, you definitely have to have a PhD, and I was getting a little bit tired of school at that point. But one of the things that kind of turned me off was they would teach you these really advanced techniques that you need to do leading-edge physics, but they would-- in school, they would have to apply it to a trivial problem. So they say, "Use tensor formulation to solve force between two electrons." Which is a completely wrong way of solving that problem, but that's what you had to do on the homework. And so that kind of turned me off, and I decided to get a real job, do computing.

Kapoor: Okay.

Fairbairn: It looked like a lot of your actual hands-on stuff; you actually did a lot of computer--

Kohn: Sure. And, in fact, at Caltech, I took a number of computer science and EE digital electronics classes. And I did a research project with a couple fellow students there that really kind of set me on this path in my career.

Kapoor: Right. So, at Caltech, you did a summer job-- Citibank EFT terminal software, writing 8080 assembly language.

Kohn: Yes, that was actually at a start-up company. The founder of that company, he actually had an earlier company that did one of the early stock terminals-- stock trading terminals. And he knew-- he had a lot of contacts at Citibank. And they gave him a contract to develop this EFT terminal, and so we did it based on 8080 and some proprietary hardware encryption engine that they had.

Kapoor: I see. And, as you were saying, you took computer science classes. It seems like you got into programming languages.

Kohn: Right. Yes, so Caltech had-- they had a professor-- they didn’t have the world’s best computer science department, because that wasn’t really their focus. But they did have a professor who was involved in some of the early structured programming, which was just kind of coming onto the scene at that time. And so, he taught Pascal as a programming language and we did a compiler development in that class, and it really kind of got me interested in programming languages and how to improve programming productivity. But I think, at that time, what was really unsolved was how to apply these programming languages to system programming, because people were mainly doing application programming in Fortran and stuff, but they were using assembly language for operating system.

Kapoor: Right.

Kohn: And so that was something that our research project that I mentioned-- that was kind of the goal of that project was to develop a programming language for system programming, and it also tried to be structured programming and included this concept of generic types, so you could define things like a linked list type that could work with any type of list element. That kind of thing.

Fairbairn: Who was that professor?

Kohn: Per Brinch Hansen, I think.

Fairbairn: Okay. Did you work with Chuck Seitz or Carver Mead?

Kohn: Well, so Carver Mead, I took this digital design from him and lab where we built kind of a 2-D Pong game. And just on my senior year is when Ivan Sutherland-- he just joined Caltech at that time, so I did get introduced to him, but I didn’t really take classes from him.

Fairbairn: Did you another student, Jim Rowson, there?

Kohn: I don't recall.

Fairbairn: Probably a couple years-- probably a graduate student at the time but working with Carver and Seitz.

Kohn: When I went to Intel, I kind of reconnected with some of the stuff that Carver was doing with Intel and some of the other guys at that time.

Kapoor: So, transitioning from school to first job.

Kohn: Right.

Kapoor: Can you tell us about how that happened?

Kohn: So, after I decided not to go to grad school, I moved back to my parents' house in San Mateo and just started looking around for various job openings.

Kapoor: You said you were sort of tired of school. Is that what--

Kohn: Yes. I got a little bit burned out at the end.

Kapoor: Okay.

Kohn: I interviewed around at a few different companies, and one of the companies that had an opening was National Semiconductor, which had a job to-- they were looking for someone to help them develop a processor based on these bit-slice-- like these 2901-type bit-slice designs. They had been looking at this-- at that time, there was this UCSD Pascal implementation, which was basically a stack-based compiler. And the idea was to implement this stack-based instruction set directly into the 2901 implementation. So I joined them to work on that project.

Kapoor: You call it configurable data flow.

Kohn: So after I got there, I got, I don't know, interested in moving beyond just this kind of stack-based architecture, which it was pretty clear was not a scalable performance solution. In fact, even relative to register-based machines, it's not a particularly efficient design. So, it's very easy to generate code for but it doesn't have good performance.

So, I started to look at alternative kinds of architectures, processing architectures. And I guess the big challenge in general purpose processor architecture was how to scale performance with parallel execution. So I just looked at all different kinds of implementations. And one of the ideas that I was playing around with was a way to connect multiple data processing operations into a data flow graph and basically execute all of them in parallel. But, at the time, it was really beyond the scope of what you could really look at implementing.

Kapoor: Yes, it’s interesting. You mentioned that, at Ambarella--

Kohn: Yes. So many years later, I actually ended up coming back and implementing it.

Kapoor: So, at that time, National, of course, was not in the microprocessor business, right?

Kohn: Well, they were. They were in the business, but they weren’t particularly successful.

Kapoor: Okay.

Kohn: They were selling an early 8-bit microcontroller and they had ambition to try to take on Intel. They hired one of the key designers from Intel to come over, and I ended up working with him, along with one of the students that worked on this programming language project with me at Caltech, Dan O’Dowd. So, I managed to entice Dan to join National, and the two of us took on the architecture for this microprocessor project that was supposed to-- at the time, I guess, the big thing that National wanted to go after was the 8086, but we kind of realized that 16-bit architecture was already running out of gas and it made sense to go to a 32-bit type of architecture. So, at that time, the most famous 32-bit processor was the VAX-11/780. And we ended up doing a machine that was, I think, similar in style to a 780, memory-to-memory plus register architecture. But, of course, it had to be optimized to fit into a microprocessor implementation.

Kapoor: Right. So you needed good floating point.

Kohn: Floating point was part of it. Yes. We pretty much wanted to be able to run any kind of software that you would run on a VAX, to be able to run it on this microprocessor. So that meant we needed both floating point and virtual memory architecture. We ended up with a three-chip implementation. A CPU which did all the integer operations, a memory management unit that did the address translation, and then the floating-point unit.

Kapoor: Who was the person from Intel that joined, that you just mentioned?

Kohn: Zvi Soha.

Kapoor: Oh, okay.

Kohn: Yes.

Kapoor: I think I have met him. Yes.

Fairbairn: So how did you-- in retrospect, what was your evaluation of that major project? Was it the best you could do at the time? Was it overdone for what was cost-effective? How did you feel, independent of the marketing success?

Kohn: I think it was probably the best you could do for that type of architecture at that time. The CPU budget-- we only had sixty thousand transistors to play with, right? That was already considered a huge chip by National standards. They were selling a lot of flip flops and stuff. I think, in that respect, technically, it was a good product, but on the business side, it didn't really fit into the culture of the company and the kind of salesforce they had and the market position they had. So, it was difficult to make it commercially successful.

Kapoor: What kind of tools did you have to build the models and check the performance and run programs?

Kohn: Yes, we had a VAX-11/780. And for the physical design, it was just at the point where they were-- they had this advanced computer-aided design that involved basically manually drafting and then digitizing all the--

Kapoor: Yes, of course.

Kohn: [inaudible] and stuff.

Kapoor: So then of course, Israel comes into being.

Kohn: Right.

Kapoor: How did that happen? Why did you select Israel for building?

Kohn: Of course, Zvi Soha was from Israel.

Kapoor: Right.

Kohn: At that time, Intel had actually set up, I think, an early design center there and was expanding there. I think the idea was get access to a lot of good engineers that you might not be able to hire in Silicon Valley.

Kapoor: I see.

Kohn: That's why Zvi--

Kapoor: It's interesting that I visited the design center in Israel in Herzliya, the National design center. I was with Synertek, and we were thinking of second-sourcing. So, this is a few years later, but I met a lot of the people that were there. The people that you had hired and built the team.

Kohn: Yes. It was a really good team.

Fairbairn: What was your role in setting up that design center and what were the goals? It sounds like Zvi, who's the sort of driver in terms of the location, but sort of what was the driving force in terms of-- from a business point of view or a design point of view?

Kohn: Well, the driving force was to build a team that could develop leading-edge microprocessors that could compete with Intel. What I was involved with there was basically-- since this thing was built from scratch, I basically went over there when they first got the building and helped them from a technical point of view, helped them set the direction for how to-- like the microarchitecture design of the chip is something that Dan and I worked out the basic ideas. It was a micro-coded machine, so we did some of the early microcode development for it. Helped them build the simulation models for it and just made sure, as architects do, that everyone on the team understands what the chip is supposed to be doing and answer questions and help out.

Fairbairn: So, you had done the architecture with Dan O'Dowd?

Kohn: Yes.

Fairbairn: But then all the implementation and detail work was done at the Israeli design center. Is that correct?

Kohn: For the chip design, yes.

Fairbairn: Yes. Right.

Kohn: We also-- sorry.

Kapoor: It's interesting that there was a fear-- this was, of course, a few years before that, about the location because of all the wars.

Kohn: That's right.

Kapoor: And I think a lot of the hard-- from Dov Frohmann-Bentchkowsky on the Intel side to sell that idea to Intel management to establish that-- fab and the design there. And so I think you were probably-- had reaping the rewards of that.

Kohn: Yes, I mean, actually, even at the time that I first went over there, it was still-- I guess what you read about in the papers about how bad it is, when you actually go there, you get a lot different perspective because what I quickly realized is you were much more likely to get killed in a car crash the way people were driving than by a bomb. It's something you have to see firsthand.

Fairbairn: Did you actually live over there for a couple years? What was your--

Kohn: I was there for a total of about six months. It wasn't continuous, because I would kind of go there for a month or two, then come back, and so on.

Fairbairn: What would you say was the major learning that you had from doing this chip? What are the lessons that you took from that and said, "Well, I'm not going to do that," or, "I will certainly do that the next time," or whatever. So what was the--

Kohn: Sure. Well the first thing is I learned how to design a microprocessor. This was the first time I'd ever been involved in something like that, in understanding the full scope of all the steps that you have to go through to build a complex chip. But the thing that I really learned, that I decided we had to do differently in the future, was when I started working on the next generation of that chip, to scale the performance up. That this micro-coded-style implementation of complex instructions, you end up hitting this bottleneck. So, the thing that you have to do to scale beyond the traditional microcode machine is you have to go to a pipeline design so that you can execute one instruction per clock. Once you start to do that, this whole micro-code concept breaks down, because now you have a whole bunch of instructions in flight, and there's no single stage that you can just execute a microcode program to work on an instruction.

So I realized that this pipelining limitation would be much more challenging for the CISC instruction sets. And then I happened to get some papers from-- I think it was from Dave Patterson on their RISC project at Berkeley, the first RISC project that they had. And I looked at that and realized, yes, this makes much more sense for a pipeline implementation and that should be the direction for the next projects.

Kapoor: So that was the level of complexity at that time, not VLIW and--

Kohn: That's right. You're just trying to get a basic pipeline machine going.

Fairbairn: You also worked on a network workstation there? Is that correct?

Kohn: Right. Yes. During the development of this microprocessor, it was clear that this kind of processor, performance-wise, was actually a lot better than the mainframes and VAX-11/780s that we were using to design it. So, it was pretty clear that this was going to change the direction of computing in the future, and so we should rethink what's the right kind of computing systems that make sense for microprocessor-based processing. So, in looking at the cost of the computing system, at that time the hard disk was this super expensive thing. I can remember even, I think, in 1986, which was many years after this project, I bought a 10-megabyte hard disk for a thousand dollars. So that was kind of a killer cost. So, what we decided to do was to develop an architecture where each person has his own workstation, but you share the file storage across them. So in order to do that, you need to network. So, we had developed this concept, and National, after some political battles, they actually let us kick off this program, and we hired a software team to work on it. We did hire someone from Xerox PARC who was familiar with the kind of GUI interface that they were developing, so we kind of got motivated to work on that kind of stuff.

Fairbairn: Who was that that you hired from PARC?

Kohn: His name was Peter something. I don't remember his last name, unfortunately. Been too many years. But anyway, we had a bitmap display. We had a joystick instead of a mouse, unfortunately, but we had this UI system. We had the compiler that was going to use these kind of language concepts that we had been working on at Caltech, and we had our own OS. So it was a super ambitious program but very fun to work on and learned a lot of things not to do, next time.

Fairbairn: So did National think they were going to go into the workstation business or was this--

Kohn: Yes. At that time, National, they had a mainframe business selling 370-compatible mainframes. They were a great company to try lots of different things, not all of them successful.

Kapoor: So, in terms of the National Design Center or the whole establishment there, unlike Intel where, I think, people were fairly loyal, there was a migration out. And I think when I got to know many of the people there that ended up in the EDA world or other start-ups-- they went to Motorola and started the Motorola Group in Israel. So, I think they are very talented people, but they migrated out, would you say?

Kohn: Yes. I mean, there was a VP at the time, Pierre Lamond. I don't know if--

Kapoor: Oh, Pierre, yes.

Kohn: But he was a kind of a champion for these high-tech projects, and then, at some point, he left and I think a lot of people couldn't make it there anymore.

Fairbairn: Tell me about Pierre. What was your relationship with him and sort of was he--

Kohn: Yes, it was interesting because I was just a kid out of school. He was probably like four levels above me, but, because of this project-- this microprocessor family was one of his pet projects-- I did get to talk to him and kind of--

Kapoor: And he oversee--

Kohn: I managed to convince him to start that software project that I talked about. Yes.

Kapoor: I'm sure he also recognized a creative person when he saw one.

Kohn: Yes.

Kapoor: And he was on the board of Cypress Semiconductor. I remember he--

Kohn: And, also an adivsorr at, C-Cube [Microsystems], where I worked.

Fairbairn: Yes, I mean, you sort of worked together in various forms.

Kohn: Yes, I don't think we really worked together but, yes, we did cross paths, I guess.

Fairbairn: What happened with that project?

Kohn: The 16000?

Fairbairn: The workstation project.

Kohn: So, after Pierre left and I kind of realized that it would be pretty difficult to finish that project given the direction the company was going, so I started to look outside. One of the things I wanted to do was… I knew National was a great place to explore things that you would never be able to do at a company like Intel coming out of school. But I really wanted to learn what was involved in a really successful microprocessor company. So that's why I decided to go to Intel. After that, I think the group kind of disbanded. People just went different directions. Dan ended up starting this Green Hills Software company, which is still developing leading-edge software tools today for embedded applications.

Kapoor: Yes, he's been there for a long time.

Kohn: Yes. I don't know if you guys have interviewed him, but he's a pretty impressive guy.

Fairbairn: No. I saw him-- you referenced him, and then-- it sounds like he was a couple years older than you at Caltech?

Kohn: One year.

Fairbairn: One year? I figured he was the one that got you to National, but it was the other way around. Is that right?

Kohn: Right.

Fairbairn: You went to National, and he had gone somewhere else.

Kohn: Yes, he actually-- he went to grad school for, I think, about a year, and then also got tired of school. And then he went to a start-up that a Caltech guy had founded that actually was developing the software for these early Mattel handheld games. I don't know if you remember these things. Anyway, yes, he was at that company when I got him to come to National.

Kapoor: So, when you went to Intel, Why did they hire you? You were looking at alternative architectures?

Kohn: Actually, at that time, Intel-- they had this kind of 32-bit microprocessor crisis.

Kapoor: Right.

Kohn: The interesting thing about Intel is the senior management of the company. They all came from the semiconductor process and device physics side of things. And so, they didn't really understand computer architecture. They were the leader in microprocessors, but it wasn't from the work of the senior management. What was happening is, as-- basically, for the 32-bit generation, they decided to do this 432 project. I don't know if you're familiar with that. But it was more or less completely the wrong direction for them. But they didn't see it for a long time. They didn't realize that that was true. And so, they started this 8086 project as kind of a panic thing to get something out quick that could hold their position in the market until this 432 would come in and take over. Of course, it didn't work out that way, so they realized that they needed to look at other 32-bit alternatives, and they hired a guy named Glen Myers. I don't know if you're familiar with him.

Kapoor: Yes.

Kohn: From IBM. And he was supposed to develop the next generation mainstream 32-bit architecture for Intel, which was actually the project that I got hired into, as an architect on that project. But, as it happens, at the same time, they had another project which was started by the 432 architects where they had made a deal with Siemens to do a joint development on this next generation 432 architecture. So that was basically going on at the same time as this project, and, in addition, there was a 32-bit DSP project going on.

So, there was like several 32-bit architecture projects all going on at the same time. And, at some point, Intel management realized that was too many. But, when I came in, I wanted to do this RISC project because I had also realized from what I was doing at National that that should be the right direction for the 32-bit. But because there were so many 32-bit projects going on at the same time, I could never get that off the ground. Whenever I tried to propose it, they would say, "Well, sorry. We made the decision to go with this one, and that's what we're going to do." At some point, they decided to merge the project that Glen Myers was developing, which was kind of a-- I would say another VAX-style architecture with this-- the next generation 432 project. And, at that--

Kapoor: I remember-- I was there at the same time, at Intel, and I was in the HIMO operation, which is the high-integration microprocessor group, and I think you were part of the HPMO-- high-performance.

Kohn: Right.

Kapoor: And I remember you walking around with Jean-Claude Cornet.

Kohn: Yes.

Kapoor: He's another character.

Kohn: Definitely.

Kapoor: That was like...

Fairbairn: So what was this project that merged?

Kohn: The internal code name was the P-4 for the project that was the VAX-style architecture, and then there was this P-5 project, I think, was the-- or was it P-7? Anyway, it was the Oregon next-generation 432 project that was being done with Siemens combined with this Santa Clara VAX-style architecture project. And when they combined the two things together, they moved the whole thing up into Oregon, and I was offered to move up there. I was a little bit reluctant to do that, so I kind of was commuted there, lived back-and-forth between Santa Clara and Oregon. And I met a lot of the team members. They had a really, really great team in Oregon, from chip design point of view, but, unfortunately, they were working on the wrong architecture. So, I ultimately decided not to move up there.

Kapoor: So I became in charge of the P-6 project, which is the one with the graphics.

Kohn: Uh-huh. Yes. I remember that one. Yes.

Fairbairn: In all of these discussions of the microprocessors, where did the issue of, "Hey, let's make it compatible with all the previous generation stuff" go on?

Kohn: Right. So yes. In addition to those 32-bit projects, there was another 32-bit project which was the compatible one, which was the 386. And, again, at that time, this was not viewed as a long-term architecture. It was more like a stopgap that was just to hold the line until this beautiful next generation thing would take over. I did get involved in that. First, on the MMU side, the paging part, I kind of helped define what that was going to look like based on the stuff that we had done at National. And then I got involved in the 387, which was the floating-point part of it. The original scheme that I was trying to do there was to make the 387 be a micro-coded floating-point accelerator that was actually a RISC machine that had floating-point on it.

Fairbairn: You were going to do RISC, one way or another here.

Kohn: Right. The idea was you had-- people could actually program it, offload stuff from the 386 and run this code on this floating-point accelerator for high-performance processing. That project did get started, and it was being worked on in Israel. The Intel Israel team. But, at some point, they decided that that schedule was too late, and the cost was too high, so we ended up going back to just a hard-- well, it wasn't really hard-wired, but it was a dedicated floating-point coprocessor for the 387.

Kapoor: Itanium was started much later, right?

Kohn: Yes. Itanium, the early stage of that project was right at the time I was leaving Intel. It always struck me as a super ironic thing, that the team that developed the 386/486, that was so successful in being a compatible evolution of the architecture, would go down this path of-- incompatible thing. Meanwhile, the Oregon team, which had been developing the incompatible architecture, took the x86 architecture and took it to the next level.

Fairbairn: So where did the 860 come in?

Kohn: After the 386 generation-- 386, 387-- was done, after this NEC lawsuit that I was involved in--

Kapoor: I was also involved in that. I didn't realize you were, too.

Kohn: That was another super educational thing to do one time.

Fairbairn: All right. Let's start with that. What did you learn from that and what did you find interesting or boring about that?

Kohn: Well, a few things. First is, by the time you go through a lawsuit, you could spend so many years doing that that you never really get any real result out of it, any real help from it, so don't bother. But I also learned that these lawyers are actually-- there are some pretty smart lawyers out there, and the way they can twist your words if you're not super careful, it's amazing.

Kapoor: I was going to mention that Tom Dunlap was in charge of that from Intel.

Kohn: Right.

Kapoor: And he didn't know anything about microcode.

Kohn: Right.

Kapoor: So I was helping him. We took a picture of the chip and all that and trying to explain to him what microcode was.

Kohn: Yes. So there was this outside law firm that was handling the case, and I was kind of the technical coordinator for that. One of the key things that they were trying to prove in that case was that microcode was copyrightable, because, at that time, I think software-- it was not really clear how you were going to do IP protection for software. I remember that they brought in this guy that-- he was like an ex-attorney general, and I had to explain to him what is microcode and why is it like software? That was really fun. And also, the other really fun thing was to go over to Japan and take depositions of these engineers, which was one of my early Japan experiences. But, like I said, in the end, after you do all this work, and we had Dave Patterson as an expert witness. We had to do all kinds of prep work and stuff for that. Just so many people working for so long and then, at the very end of the case, it turns out that the judge had owned a few shares of Intel stock in some fund that he didn't even realize, and so the whole thing was disqualified. They had to retry it. It's just such a waste of time and money.

Fairbairn: All right. So then you moved on to the 860?

Kohn: Yes. So I think partially as a reward for helping them on this lawsuit case, they finally let me do a RISC project which-- I had been looking at like high-performance floating-point-- what's the next step for that. So, with the 860 project, the goal was basically to do pipeline floating point, which was, I guess, just a little bit novel at that time, but also to expand the parallelism to the next level, which was basically superscalar execution -- pipeline RISC processor plus pipeline floating point, and also doing some SIMD operations. It combined multiply and accumulate operations and also a special set of SIMD graphics instructions for doing low-level pixel interpolation in buffers.

Fairbairn: What was the goal of that product and how did it fit with the other things going on at Intel? What was the application part of it?

Kohn: High-performance scientific computing and graphics was the kind of goal. And we actually-- when the chip came out, it kind of-- it was like the first million transistor microprocessor and it caught people's attention. It was like my fifteen minutes of fame. We got a lot of interest from, first, all these mini-computer companies that were doing scientific computing at that time. There was a number of companies that were doing like parallel scientific computing. VLIW scientific computing. And they all kind of jumped on this kind of microprocessor chip for those kinds of applications. We also had designs at SGI that were using it for the 3-D-- the floating-point part of the pipeline.

Fairbairn: So, were you familiar with the world of graphics and graphics transformations and so forth?

Kohn: Yes.

Fairbairn: Was that something you learned along the way?

Kohn: Yes. It was one of the targets that we had for it. I mean, obviously I ended up getting a lot more into that later on, but I understood the--

Fairbairn: Yes, I was wondering whether--

Kohn: --that part of it.

Fairbairn: So how did that-- I read somewhere that the sort of real performance-- you had a hard time meeting the performance goals or the expectations versus the theoretical in real-world applications. Is that-- what's your evaluation of that?

Kohn: Well, you have to understand that that project, it was not the mainstream project for Intel, and so it was extremely, let's say, understaffed in terms of-- both on the chip design point of view, but especially on the software side and the compilers and stuff. So we ended up having to outsource all the compiler development, including-- we had Green Hills work on the compiler. But, at that time, there was very limited understanding in the compiler world of how to take advantage of these kinds of highly-pipeline architectures. And then, on the 860 side, I have to say, the instruction set wasn't particularly compiler-friendly either. Those two things together made it difficult to have a compiled solution that worked efficiently. Now if you hand-coded stuff, which I did for a few things-- like I did a jpeg decompression thing hand-coded-- you could actually get pretty good performance, but it wasn't easy enough for people to use.

Fairbairn: So, what you're learning from that experience, what did you take away from that?

Kohn: That you definitely need to work early on on how you're going to handle the tool side of it, how to develop good compilers for it. And, also, the importance of doing a lot of verification work during the development process, which we just couldn't afford to do with the size of the team we had.

Fairbairn: High-level performance verification, that kind of thing? That sort of assisted verification?

Kohn: Both. Both performance and just logic verification.

Kapoor: And you did some due diligence for Intel in an acquisition of video compression technology from Sarnoff?

Kohn: Right. So that was a project that I think RCA had developed for doing these videodiscs, early videodisc stuff. And it was a kind of early vector quant kind of compression algorithms. And they had implemented a chip for it and then they, I guess, decided to get out of that business and ended up selling that team to Intel. And that's how I started looking at video compression and how that might fit into future microprocessor stuff.

Kapoor: So, then the next step was a start-up. How did that happen?

Kohn: Right. So, after the 860 project was out there, introduced into the market, because of some of the team dynamics, we decided to try our luck in the start-up world. A number of the early-- like the other 860 project leader-- Sai-Wai Fu and I-- spun out and my brother and some other people-- some Caltech guys and some other people that we knew-- we formed this company to do an 860-based workstation. Because that was another kind of frustration, at that time, was that it's difficult to do that kind of a program inside of Intel, get the right people to work on it.

Kapoor: They were also going through a financially rough time. I think there was a layoff and everything, I remember.

Kohn: Well, there was in the earlier part of the eighties.

Kapoor: Right.

Kohn: It was kind of just at the turning point of when they started-- the 486 really started to take off. But, anyway, so we did spin out. The original plan was that we would get Intel to go in with us and bring in some other companies that had been working on 860 workstations to fund this thing. But after we got out, it became apparent, first of all, that Intel just decided to go in another direction. I think partially because they didn't feel comfortable continuing 860 without the original people there anymore. And, also, the 486 was starting to take off. So that kind of killed the 860 workstation program. So we then refocused on add-in boards and developed kind of early graphics add-in card solutions that we got funding from this company, Hercules, one of the early add-in card companies for PC.

Fairbairn: What was the name of this company?

Kohn: Hercules.

Fairbairn: No, the name of the company you formed.

Kohn: Oh, it was called Aquest.

Fairbairn: Okay.

Kohn: All the start-ups I've founded, they all begin with "a." This Hercules deal that we made was not very favorable for us, and they ended up being able to sell cards against-- to our customers, right? So we do all the business development, get the customer up and running, and then they'd come in and cut the price. Undercut us. So it was not really a viable business.

Fairbairn: Did you get any venture financing?

Kohn: No, we didn't. We didn't. We did attempt to, but I think it was-- it was too difficult given what Intel's direction was with 860. So, if you're going to leave a chip company, don't do a company based on the chips that you just worked on.

Fairbairn: You just took all the people that made it viable.

Kohn: Exactly.

Fairbairn: So basically, you suffered business setbacks. You couldn't make a deal. Did the company fold? Did you just leave? What was the--

Kohn: Well, we decided to shut it down, basically. We actually didn't go bankrupt, but it was just so difficult to scale the business given the deal that we had made and what was happening with the 860, so we just decided to move on.

Kapoor: So Sai-Wai also found some other-- I think you went to Sun after that.

Kohn: Mm-hm. Sai-Wai went to Weitek.

Kapoor: Okay. I think his wife was also at Intel.

Kohn: That's right, yes. Annd.... She was there for much longer than us.

Kapoor: So then tell us about Sun.

Kohn: Right. So, I ended up looking for what to do after Aquest, and then Sun was also, I guess, having some challenges in microprocessor development. They had developed this 32-bit SPARC, several generations, but the high-performance chips that were supposed to compete with Intel were having problems. So, I came in to lead the architecture for the first 64-bit microprocessor, which was the UltraSPARC I. It was also an interesting set of dynamics, because there was a lot of different microprocessor projects going on, including several things in Sun Labs. And, let's say, again, I think the senior management of the company didn't understand chip design.

Kapoor: Right.

Kohn: So, it was difficult for them to figure out what's the right next direction to go in. So, the good thing about this project was there was a lot of resource and staffing available for it, and I was able to apply all the things that I had learned working at Intel and National-- all the mistakes that I had made-- and really focus on making this project a success.

Fairbairn: You were able to apply-- you were finally able to do something RISC-like.

Kohn: RISC-like with the proper amount of staffing and the proper methodology, etc. And working with a really good team, of course.

Fairbairn: That always helps.

Kohn: Sorry?

Fairbairn: That always helps.

Kohn: Yes. And, of course, it was very fun to work with Andy Bechtolsheim. The thing that was really nice about the Sun project was you were also working with the system team in house, right? And all the trade-offs could be made knowing exactly what kind of system you wanted to build out of it.

Kapoor: Were you also interacting with Bill Joy?

Kohn: And, yes, Bill Joy on the architecture side. Yes.

Kapoor: In terms of architecture at that time, what was the thinking? Performance-wise-- RISC-like, superscalar. Was there any attention to better compilers or better architectures? Were there any considerations? You have your constraint to follow.

Kohn: Right. When I joined the V-9, which was the 64-bit instruction set, that already had been pretty much defined, along with the memory management unit system architecture stuff. But the part of the instruction set that I did work on was this Viz instruction set, which was another SIMD. Like next generation SIMD instruction set. It was addressing both video processing, to some extent, but also graphics, that kind of stuff. And that was something-- it was still pretty controversial within Sun that this thing should even be done. I remember there were several graphics groups that were all fighting with each other. I aligned with one of them that believed in this kind of instruction set.

Kapoor: Dave Ditzel was also there at that time.

Kohn: Yes. He was on the Lab side, working on several other microprocessor projects.

Kapoor: But you didn't run into Roger Ross, in terms of any controversy about the 64-bit architecture, right? Because was involved from an international standpoint. SPARC International.

Kohn: No. It was pretty much-- that part was pretty much worked out.

Kapoor: Okay.

Fairbairn: What would you say were the most innovative or unique aspects of this UltraSPARC I chip that you did?

Kohn: Well, it had a fully-pipelined L2 cache which allowed you to run fairly large dataset scientific applications pretty efficiently. And this is something that, again, had a good compiler team, and they managed to do a pretty good job of taking advantage of this, along with the fully pipelined floating point, four-way superscalar processing. I think the overall memory architecture, memory hierarchy, was well-designed for large applications.

At that time, a lot of people were focused on benchmark results and, still today, a lot of people focus on small benchmarks, and it's difficult to really understand how the system is going to perform with large applications. And so, one of the things that we spent a lot of time on was developing a simulation environment that could scale up to running like operating system plus application code and measure performance in real applications. So that gave us a much better understanding of the architecture tradeoffs. Microarchitecture tradeoffs, to optimize.

Kapoor: So a lot of emphasis on web and networking kind of issues, because Sun was very--

Kohn: Well, web was still pretty early at the time. In fact, I remember that Eric Schmidt almost banned web access. It definitely included full OS-- multiprocessing. That was another key thing that Sun was just starting to develop at that time. Large-scale multiprocessing servers. So, there was a lot of work on both the cache coherency protocols and, again, the operating system/performance verification for that.

Fairbairn: So, it sounds like this chip development went largely as you envisioned coming in. You came in to do a project and you got to do that project and it was along the lines that you had come in to do. Is that correct? Is that a proper description?

Kohn: I would say so. I mean, it was a pretty challenging environment, because, like I said, the microprocessor side of things wasn't doing that well when we came in. There were all these competing projects and a lot of arguments about the direction. But, in the end, yes, it did work out pretty well.

Fairbairn: I mean, it became sort of the mainstream engine for Sun moving forward, right?

Kohn: Yes.

Fairbairn: What was the learning from this? Other than if you get a good team and you get control of all the pieces it really can work out?

Kohn: Yes, I mean, it was a validation of the methodology that was used, I think, for the project. The other thing that, personally, I took away from it was that conventional microprocessor design was kind of reaching its limits. It was clear that you could do things like out-of-order execution, which was another huge step in complexity, but, beyond that, there wasn't a whole lot more you could do without just going to multiple processors, right? It became kind of less interesting to me as a future direction, and that's why I decided to move into more application-specific processing. The other thing that was kind of difficult about it was you had to work with a really large team. UltraSPARC wasn't particularly large by Intel standards, but you're talking about hundreds of people kind of projects, and the dynamics of that is not quite as fun to me as smaller projects.

Kapoor: So then what was the step that caused you to move on to C-Cube?

Kohn: Well, partly because of this application-specific focus, and, of course, it's working in video compression, which was an area that I was interested in and looked quite promising. That needs a lot of performance. And a smaller company and all that kind of stuff.

Fairbairn: So, before you get onto that, I just wanted to pick up on the last comment you made. How would you describe the chip development environment in Sun versus Intel. You said it was a big team but still smaller than what Intel threw at it. Was that sort of structurally-- how would you compare them? Was Sun more efficient? Was it just different? What was the--

Kohn: Yes, I think they were more efficient. I think since then, like the things that we've done at C-Cube and Ambarella to develop pretty complicated chips with much smaller teams show that, yes, you can be a lot more efficient than what Intel is doing. And to some extent, these things-- it's like all these bureaucracies, right? They always grow over time and people tend to make things overly complicated.

Kapoor: And I think it has some historic precedents. Meaning Intel and other, like National, probably had large teams because of the methodologies and tools.

Kohn: And, of course, they are very successful in that business. They could afford to have large teams. And, in fact, when I went back to Sun the second time, the teams had also gotten much larger and I think were less efficient.

Fairbairn: All right. So, anyway, you were moving on to C-Cube.

Kohn: Yes.

Fairbairn: Were you recruited there? Did you seek them out?

Kohn: Yes. Well, actually, Sai-Wai Fu had joined C-Cube before me, so, yes, he recruited me in there.

Kapoor: And Bill O'Meara was there, right?

Kohn: Yes.

Kapoor: Okay.

Fairbairn: Did you know him before?

Kohn: Briefly. I mean, I think he was leaving as I joined. We didn't overlap too much.

Fairbairn: What were you brought in to do?

Kohn: Originally, I actually-- sorry?

Fairbairn: What was the state of C-Cube at the time?

Kohn: Yes. So C-Cube, they were kind of an early pioneer in the MPEG-based video compression space. My boss there was a guy named Didier LeGall, who was the chairman of the Video MPEG Committee that developed the standard for MPEG-1 and MPEG-2. So, he was an expert on video compression. He had a team there that had been involved in some of the standards development and had kind of the state-of-the-art video compression algorithms understanding. So their business, basically, was two main things. One was a set of video decompression chips for consumer applications like video cd or DVD or set-top box. And then they had this compression family of chips which, at that time, was mainly being used for broadcast applications.

So they had actually developed the first generation of MPEG-2 compression for DirectTV, that was used for all the original DirectTV deployment and other kinds of early digital broadcasting systems. This thing was developed as a multichip. It was like a humongous board full of chips and FPGAs to do a single channel of standard definition video compression. So, the project I was hired in was to basically come in and develop a single-chip replacement for that board full of stuff and have better quality at the same time. So, I originally started working there as a consultant, but, after a few months, I decided it was pretty interesting, so I decided to join them.

Kapoor: Yes. And you were there for about six years.

Kohn: Yes. That's right. So while I was there, I met Fermi Wang, who was the guy that-- he was basically running the software side of all the projects that I was working on there, and he became a general manager and very interested in the business side of things. So, the two of us decided to go out and do a start-up together around 2000.

Kapoor: That was Afara?

Kohn: Yes.

Kapoor: Okay.

Kohn: Along with-- I mean, Afara was actually originally started by Kunle Olukuton out of Stanford.

Kapoor: Okay. Right.

Kohn: And we joined him.

Fairbairn: What were the major takeaways from your C-Cube experience?

Kohn: Well, first, that you could do these large projects with a much smaller team, especially if you understand-- the key thing is to understand the problem that you're trying-- the algorithms that you're using to solve a particular problem really well, so that you can optimize the architecture for that. And then you cannot waste time developing a lot of general-purpose things that aren't particularly efficient are very, very complicated to implement. So that was one key thing. I think we got to try a few things. Like we had an early 3-D technology project that, unfortunately, didn't make it out. I just learned a lot about how to come up with really efficient-- both from internal processing but whole-memory system hierarchy. Efficient processing algorithms for graphics and video processing.

Kapoor: So, your desire to do multiprocessing came to fruition in Afara, right? I mean, that was the basic-

Kohn: Yes. What was intriguing to me about Afara was that it represented another way to scale performance without having to implement a super complex pipeline. The idea was to use multi-threaded, fairly simple pipelines, but use multi-threading to scale the performance up. And that multi-threading could cover the memory latency that you have with cache misses more effectively than what you can do with an out-of-order pipeline. And I think this has actually been proven out when you look at like the way GPUs work. Modern GPU architectures. They use a super large number of threads to basically do that exact same thing. But we were using it for server applications, particularly for web processing. That's, again, a case where you have aggregation of a large number of clients that are coming in, and so optimizing the throughput of those clients works. You don't need to have super low latency for each client.

Fairbairn: Going into this, it seems like you-- I mean, Sun was the dominant supplier in the web systems/web server business, at the time, right?

Kohn: Yes, they were.

Fairbairn: So that was your primary target you were aiming against?

Kohn: Yes, although I think, at that time, PC architecture was starting to become very competitive in that space. Because you had the Pentium Pro and the follow-on chips that were actually quite competitive with any processors that Intel-- I mean, Sun-- had developed.

Fairbairn: So you were going to do a new processor and build the system around that--

Kohn: Yes.

Fairbairn: --to really focus on this.

Kohn: Mm-hm. Of course, it was during the peak of the dot-com boom, right?

Fairbairn: Yes.

Kohn: The thing everyone had to do.

Fairbairn: But it's still a major undertaking--

Kohn: Yes.

Fairbairn: --to develop a new microprocessor, compilers, operating system and insert it into an application, right?

Kohn: Yes. I mean, we were leveraging a lot of open-source development that was, I guess, really becoming popular, at that point. So we were using Linux. We were using open-source web servers and so on. And we used SPARC architecture, so we had-- the compilers and stuff for that were all there.

Fairbairn: I see.

Kapoor: Yes, I was going to ask about the compiler, because they get more complex as you go.

Kohn: Well, the beauty of this multi-threading stuff is you don't need a good compiler for that. The way the application is written naturally has all the threading in it, because that's what they have to use anyway for scaling.

Kapoor: Okay. So, anything else? I think you then got absorbed into Sun. I was there. I was one of the people that did the due diligence.

Fairbairn: So before you do that, what happened? It looks like you were--

Kohn: We had raised one round of funding at the peak of the dot-com thing. And then 2001, September 11th happened. Fermi and I were actually on the way to New York City, driving into New York City, when the planes hit. We saw the smoke coming out of the Twin Towers. But that basically just killed the fundraising for us, and we ended up having to sell the company to Sun.

Fairbairn: I had a similar experience. I also had a VC presentation scheduled for that day, and it didn't go well. Well, number one, it was delayed by a day, and, number two, it didn't pan out after that.

Kohn: Tough time.

Kapoor: So, of course, on the Sun side, they were running into the clock speed and power issue, and that was really the motivation.

Kohn: Right.

Fairbairn: Okay. So, Sun swooped in to scoop you up. What specifically were they looking for and what solution did you provide?

Kohn: They were interested in this multi-threaded microprocessor that we were developing, although I have to say, again, there was a lot of competing projects going on. And so, I did have to kind of fight to keep that project going intact, or as intact as possible. And, fortunately, we were able to do that. It was, let's say-- it was scaled back a bit from what we were originally trying to build, but it did come out and it, I think, was pretty competitive at the time that it came out with other processors or for these kinds of applications.

Kapoor: And then their roadmap actually changed to that. They killed some of the other projects and continued next generation on that T-1 [ph?] and so on.

Kohn: Right.

Fairbairn: So did you find Sun a familiar environment, coming back after--

Kohn: Actually, it was both a familiar and quite different from-- since I was there the first time. I mean, a lot of the people were there that I knew, but the company had grown quite bureaucratic, unfortunately. I don't want to get into all the dirt, but let's say, coming from a start-up company where you're used to moving fast and making quick decisions, it was quite a shock. And after a year, I decided not to stay there anymore.

Fairbairn: But you stayed long enough to establish that as sort of the main path for Sun going forward.

Kohn: Right. Mm-hm. Yes.

Fairbairn: So you did that twice. You sort of got them onto new paths with your two times at Sun. That's quite an accomplishment.

Kapoor: Yes.

Fairbairn: So that also convinced you that you didn't want to do a big company. You enjoyed the smaller environment?

Kohn: Right. So Fermi and I decided, yes, it's time for another start-up.

Fairbairn: And how did that evolve? How did you come to a conclusion as to what the next step was?

Kohn: So after we both left Sun, we became EIRs [Entrepreneurs in Residence] at Benchmark because the former C-Cube CEO, Alex Balkanski, was a partner there. And so, we looked around at various opportunities and we ended up coming to an opportunity that looked intriguing, which was… At that time, there was a new video compression standard that was just being finalized called H.264. And this standard, we realized, could enable you to record high-quality video onto a regular flashcard memory, so used very low-cost storage. You could record an hour or more of video onto a 4GB card. Enabled you to essentially combine the functions of video recording-- digital camcorder kinds of functions-- with digital still camera technology. Digital still camera was really taking off around that time, and people were replacing old film cameras with digital still cameras. But the video cameras were still tape-based or big clunky things that you had to have a separate camera to carry around. So we said why not combine those two things together and just have a much more convenient camera device? So that's the premise that we started the company [Ambarella] with.

Fairbairn: So you were going to develop a chip to--

Kohn: That would basically provide high-quality image processing for the camera side and you could either store those images as JPEG files for still images or as H.264 video files.

Fairbairn: And your previous experience, you had a pretty clear view of how to do that? Was there any--

Kohn: Well, we certainly understood, let's say, the general processing required for video compression. We had, actually, no experience in any image sensor processing when we started the company, so that's something we learned a lot in the job. And the other thing is, because H.264 was a new standard, people really didn't believe you could implement it in a portable power envelope. So it was, again, one of these things that, at the time, when we started the company, the implementation of a single HD channel was like a refrigerator-size box with tons of DSP chips and FPGAs and stuff. And we managed to get that down to something that could run on a battery-operated device.

Fairbairn: And that was your goal? I mean, that was your goal going in and you were able to accomplish that? How long did that take?

Kohn: Yes. So we had a really, really good chip team this time. And we got the first chip back within about two years from when we started the company.

Fairbairn: So this was a small high-performing team.

Kohn: Right.

Fairbairn: For a small high-performing chip. Or a large-- how big was this chip?

Kohn: Oh, let me see here. I don't recall exactly. The first chip that we did was on .13 micron, so pretty loose by today's standards. But I believe it was around maybe 70, 80 millimeters in size.

Fairbairn: Oh, very--

Kohn: Sorry?

Fairbairn: Very cost-effective. Not physically too large.

Kohn: Yes, I mean, of course, we ended up, when we got to high-volume production, we got it down to like 20 millimeters or so. The novel thing that we did, that was a mistake in hindsight, was we used this embedded DRAM technology. In order to make the DRAM usage efficient, we wanted to have a large on-chip memory for that. So, we decided to use this embedded 1T-DRAM technology which never really worked properly. Luckily, we managed to survive it.

Fairbairn: Did you eventually put it off-chip or use static technology?

Kohn: We replaced it with SRAM, yes.

Fairbairn: Yes. And so was that chip-- you launched that chip as the opening salvo? Was that a big success?

Kohn: Yes. So, it turned out that where the big success was for the first chip was in broadcast encoding. So, as I mentioned, the existing systems were these humongous refrigerator-size boxes that were very expensive and very inefficient. Because we knew people that had been working on the broadcast encoders, and when they evaluated this chip against their large boxes, they found the video quality was actually better in the single chip. So that thing just took off. It was an instant success in that space. Even though it's not super high-volume, it's very profitable because it's broadcast professional equipment. And so, it funded, really, the company to grow for the first few years.

Fairbairn: Are you still making that chip?

Kohn: No, not that one. We still sell a few broadcast chips.

Fairbairn: So, what was the evolution? That was the one that launched. What was the next one?

Kohn: After the first chip, we did this kind of cost-down of that chip that was actually used in a consumer camcorder device. As it turns out, our original idea of combining the still camera and the video camera, from the business point of view, was challenging because the still camera space was dominated by these large Japanese camera companies. You know, Nikon, Canon, that kind of company. And they were very, let's say, conservative and slow-moving. One of the big challenges we had in the early days was that they were all using these CCD image sensors. And CCDs are nice for a still camera but they're not good for video. We wanted them to switch to CMOS sensors.

It eventually happened, but it took a long time. So we ended up really focusing more on the digital video camera, for the first few years. Like these pocket camcorders. And we were pretty successful in getting that market going until basically the cellphone started to do video, good video. After that happened, we ended up kind of migrating into other segments that cellphones couldn't address. I think the first really big success was the sports camera. GoPro pioneered that segment, and we worked with them very closely in the early days to develop these very high-quality but very portable and robust cameras.

Fairbairn: So, is that because they, in fact, did want to combine still and video and use CMOS sensors?

Kohn: Well, they were more video-centric than still. Of course, you can take still. But they really needed something that was low power, because you had a very small physical size, and it's in a sealed case, right? So the cooling is extremely difficult. So that was a perfect fit for our technology. And then, of course, we did still provide very high-quality video from that.

Kapoor: So the company is located here?

Kohn: So the company-- actually, even at the very beginning of the company, we decided to try to use a kind of hybrid Silicon Valley/Asian model. Fermi is from Taiwan, and there are other people who have a lot of contacts there, so we set up an office in Taiwan from the beginning. And we actually have the bulk of our team there, right now. We also have teams in China and Italy. Engineering teams.

Kapoor: And you're using the fab and TSMC?

Kohn: Well, we started with TSMC for the first few generations, and then we ended up moving to Samsung.

Fairbairn: So how did you wind up in Italy?

Kohn: We acquired a team in 2015. We had decided that we wanted to really focus on opportunities in automotive as one of our major markets because there's a large number of cameras in the car. And then, of course, the other major development is that computer vision and AI processing is being used for all the camera stuff. So, we wanted to understand all the algorithms that are involved in doing a full autonomous driving stack. Because, as I mentioned, we'd like to really optimize architecture based on a good understanding of the core algorithms. And there was a team in Parma, Italy called VisLab which was started by a professor there who actually started working on autonomous driving in the mid-nineties and was one of the early pioneers in the space. His team was involved in the DARPA Grand Challenges and did things like drive autonomously from Parma to Shanghai.

Fairbairn: Wow.

Kohn: This multi-month road trip. They had a very good team that understood the full stack. Tthey, at that time, in 2015, were running into the problem that they had a trunk full of PCs to do the processing for all the cameras. They needed to figure out how to shrink that down to something that was reasonable, so we were a natural fit. And we've been working together to come up with the ultimate chip for that application.

Kapoor: How about drones? Have you also been--

Kohn: Yes, after the sports camera market, we ended up also becoming involved in the drone market for camera processing, and we're still involved in it now today. It's not as big of a market, obviously, as sports camera or the other markets that we're focusing on now.

Kapoor: Yes, I was just watching a cricket match between India and Australia, and a lot of the drones are used for sports interviewing and stuff. It's really fascinating.

Kohn: Yes. I mean, they definitely have their niche. Very, very interesting applications. I guess because of the regulations and-- it hasn't become as mainstream as--

Kapoor: Right.

Fairbairn: So it was through this group in Italy that you got into sort of the AI processing? How did that happen?

Kohn: Well, we were working on AI processing from more generic computer vision and AI processing before that, but we really got into understanding the full stack with this team. Of course, they also brought a lot of computer vision expertise, as well, particularly in stereo processing. It's really across the company at this point, all the AI development.

Fairbairn: Is that an increasing element in your chip, moving forward, then?

Kohn: It's the dominant element, yes. The other stuff is becoming a small, small piece of it.

Kapoor: So that's a good segue into the future of architectures. What can you tell us a little bit from your perspective? What's evolving and so on?

Kohn: Well, I think it's pretty clear now that all the really high-performance scaling that's going on is happening outside of the CPU. GPU was the first really mainstream example of that. But I think, as we go into AI, it's time to move beyond GPU into architectures that are optimized for AI processing. It's one of these things that you can just soak up an arbitrary amount of compute power. And so, you'd better come up with the most efficient solution for that. And I think people look at AI-- like the convolutional neural net-- and I hear a lot of people say, "Oh, it's just a matrix multiply, and anybody can do a matrix multiply." But, actually, throwing a bunch of multipliers and adders down doesn't solve the problem.

You have to be able to handle a wide range of different configurations of networks, and there's all different kinds of bottlenecks that show up when you do that. And keeping the data paths busy is really where the key challenge is. So, if you look at the efficiency of these networks running on a GPU, you'll see that, when you really look at how many multiply accumulates, they're actually doing versus the raw hardware capability, you'll see it's not very good efficiency. So, we're all about trying to optimize efficiency and still provide a high degree of flexibility so that you can run different kinds of networks. And that's a big challenge.

Fairbairn: But that involves the memory architecture and--

Kohn: And definitely memory is a big part of it. Big part of it. Yes.

Fairbairn: So, you've been at this for fourteen years. This is--

Kapoor: Yes.

Kohn: Yes. My longest job yet.

Fairbairn: By a long way. So is this-- do you see yourself continuing to be challenged--

Kohn: Yes.

Fairbairn: --in architecture alternatives and so forth going forward?

Kohn: I mean, yes, we're working on some monster-sized chips right now, which are quite interesting and fun.

Kapoor: So, do you have any words of wisdom for the generation that is starting their careers? What kind of things they should work on? Of course, you have a very nice perspective.

Kohn: Yes, it's interesting. We often talk about whether we would do another chip company. It's getting hard, because the industry is becoming more and more consolidated, and the level of investments required is becoming larger and larger. So, I don't know whether there will be a renaissance at some point, but right now I would say it's probably difficult to do that as a start-up, and software is a lot more attractive opportunity.

But I do think that we have to be careful that you don't just become…. In software, it's very easy to stay at a super high level and use libraries of code that you don't really understand what's going on inside those libraries. Everything is fine as long as you can find the right function in the library to do it for you, but if you have to do something new-- and, at some point, that always happens, right? We need people who still understand how the stuff really works underneath. And that's like at compiler level, operating system level, and what's going on inside these chips.

Kapoor: And various applications, whether it's the medical field or data analytics or different domains. The kind of solutions that have to be found, right?

Kohn: Yes. If you really want to come up with an efficient solution for something, you need to have a good understanding of how things work at a deep level to really understand the bottlenecks.

Fairbairn: I think the launch of neural networks and so forth, stuff that finally sort of came to the fore after people realized what could be done, has actually launched a large number of "AI chip companies."

Kohn: Yes, that's true.

Fairbairn: Few will be successful, but at least sort of this new architectural domain, getting out of the CPU world and even the GPU world, has at least opened up opportunities for people to explore, but, as usual, there will probably only be a few that are successful.

Kohn: That's right. Yes. It's like the same thing happened in the video compression space. There was-- I don't remember how many MPEG compression chip companies. Maybe twenty or thirty. And probably only a couple actually survived<inaudible 01:52:42>.

Fairbairn: Yes.

Kapoor: So any other last-minute words that you want to talk about?

Kohn: Well, I don't--

Fairbairn: Has Ambarella gotten too big and become bureaucratic? Or are you--

Kohn: Well, you know, We've had our challenges over the years as markets shift around. I mentioned the cellphone was one big challenge. And consumer things tend to-- another big challenge for us was consumer grows fast but it saturates out and then transitions to something else. So we haven't grown so fast that we've become super bureaucratic. We've had to stay fairly nimble.

Kapoor: So how many people are there?

Kohn: Total of around seven hundred people.

Kapoor: Okay.

Fairbairn: And most of those are in Asia?

Kohn: Yes. But I do think, if you're lucky enough to get to work on interesting projects, that you can keep going for a long time. A lot of people think you can only be an engineer for maybe-- after your thirty-five or something you're going to have to become a manager or something. Not necessarily the case.

Fairbairn: You're proving them wrong. A lifetime of engineering innovation.

Kapoor: So that was wonderful. Thank you, Les.

Kohn: Sure, my pleasure.

Fairbairn: Yes, great for all the insights and hearing your story. It sounds like you've made a tremendous contribution to various companies in the industry, and we appreciate your time and sharing it with us.

Kohn: Let's hope it's useful for some people.

Fairbairn: I think it will be. Thanks very much.

Kohn: See you.

Kapoor: Thank you very much.

END OF THE INTERVIEW

