Join the revolution

A domain focused operating system and a domain specific reprogrammable chip could revolutionise networking technology. By **Graham Pitcher**.

Above: The Tofino chip is being made by TSMC on its 16FF+ process

The networking world has largely been defined by the development of fixed function chips. While these devices have been one of the means by which data rates have increased, this has been achieved at the expense of configurability.

Now, parts of the industry are looking to regain control over how networks are defined and one company hoping to shake things up is Barefoot Networks, whose Tofino chips bring reprogrammability into networking.

Ed Doe, vp of product marketing and strategy, said: "Networking today has been dictated by making things faster. As a result, it's been about functionality built into chips. And whatever is in that chip is the network you get.

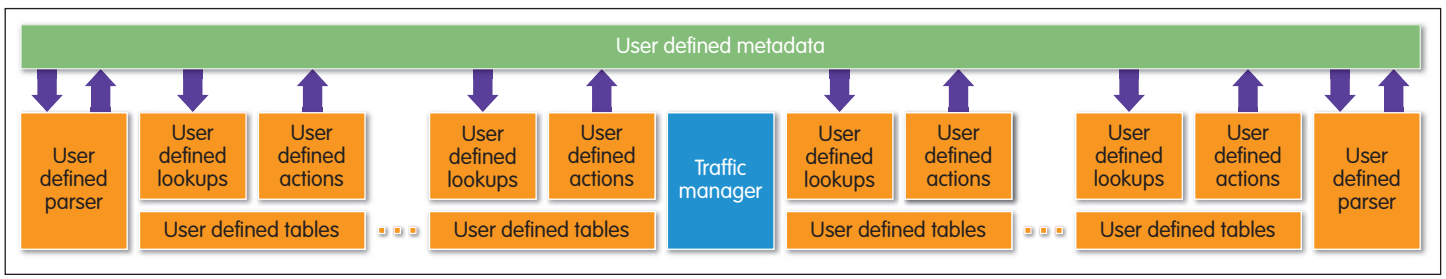
"Because networks have been driven by chips," he continued, "even large operators – such as BT, for example – don't get to make decisions about functionality."

One of the issues, said Doe, is that different users have different needs. "Everyone thinks the likes of Google and Amazon might need the same network, but they don't. While Google is focused on search, Facebook is handling videos and images. They have different needs, but lack the opportunity to influence the network. We are trying to return power to the end user."

Barefoot is taking advantage of a recently developed open source programming language called P4 – short for Programming Protocol Independent Packet Processors; the title of a conference paper presented in 2014. P4 is said by Doe to have significant advantages compared with Verilog, but to be focused specifically on network data forwarding, with elements that allow the user to customise a network according to their particular needs. But the P4 program needs to target something and Barefoot has developed the Tofino networking chip for this. "It comes with an architecture that supports the use of P4," Doe said.

The combination of a targeted operating system and a programmable device is set to allow communications technology to catch up with developments in other technologies. "The key is bringing the ability to customise all the way down to the network level," Doe said.

What Barefoot, amongst others, is looking to do is to move the focus away from fixed functionality. "People have been building chips for speed and everything in the pipeline is fixed. So, if you build a chip for VxLAN, it won't support the next protocol. We



Below: Barefoot says its technology can be used to support a user defined forwarding plane

asked whether we could come up with an architecture that achieves that without power, performance and price penalties."

The answer to that question was 'yes', and Barefoot has developed the Tofino range, based on what it calls a programmable ASIC fabric. "Network designers can use the same approach in creating a chip," said Doe, "but each stage in the design process is programmable and definable." Essentially, the chip uses the 'match and action' approach in which headers are 'matched' and 'actions' determined.

The process starts with programming using P4. "The program is then compiled," Doe continued, "and the data plane configured."

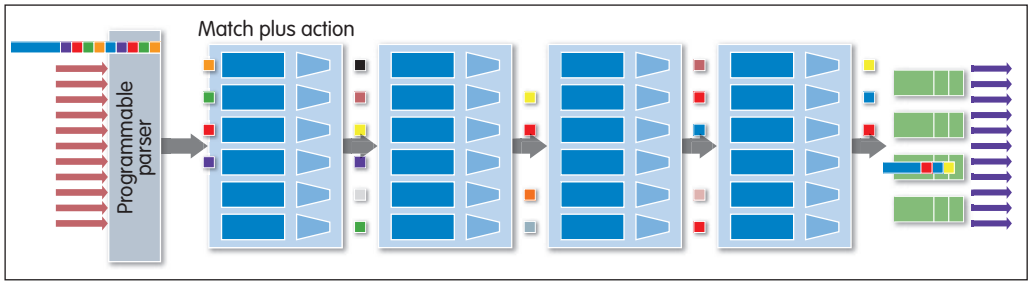
The parser, for example, allows the Tofino chip to be programmed with new protocols. "IPv6 was written as a standard in 1996," Doe said. "And yet, 20 years later, we are still talking about the transition. Could anyone then have understood what would happen in 20 years?

"So, if there's an IPv7, for example, and the need to increase the number of bits in the header, the parser would be able to extract them."

Tofino is reprogrammable, but not in the sense of one time use only. "Users can develop a new program and upgrade the chip," he said. "It's useful, because most networking companies don't know what their infrastructure needs will be in the future. By investing in a programmable architecture, networking providers will be able to add new features over time."

Even though each stage in the design process is programmable, the end result is still a pipelined architecture. "And we can do that without adding any overhead in term of chip size or power consumption," Doe asserted.

Intuitively, you might expect Tofino to be realised using FPGA technology, but that's not the case. While FPGAs are flexible, Doe says they aren't optimised for use in networking applications; they come with a large overhead and are not power efficient. "FPGAs tend to be a solution of last resort," Doe claimed, "so we have designed our own high end chip. While FPGAs operate in a similar mode, with a program compiled to the fabric, we have developed a domain specific processor."



PISA, the Protocol Independent Switching Architecture, uses memory and ALU blocks to implement a 'match and action' scheme

Barefoot's approach not only allows for reprogrammability, but also brings a significant increase in data throughput. "FPGAs may only operate at a couple of hundred Mbit/s," Doe suggested. "Our Tofino parts can run at Tbit/s – an order of magnitude improvement."

According to Doe, FPGA providers are seeing the benefits of a programming language such as P4 and are becoming involved with the P4 Language Consortium (p4.org).

Barefoot's architecture is called the protocol independent switch architecture, or PISA, and can be considered as hard blocks in a programmable fabric. "It's like RISC vs CISC when compared to fixed function devices," Doe offered, "but you can also think of it in terms of GPUs. When first introduced, these used to features fixed functions and support a particular resolution – VGA, for example.

"GPUs have been through a transition and now feature programmable architectures that are just as efficient as their predecessors. Similarly, languages such as OpenGL were developed. We're trying to bring the same transition to the networking domain. Just as GPUs have opened new markets, we think the same thing could happen in networking, with the language, compiler and programmable architecture."

"The key is bringing the ability to customise all the way down to the network level"
Ed Doe

Tofino is said by Doe to be the fastest family of Ethernet chips currently available. At the top of the range is a 6.5Tbit/s device which can support 256 25Gbit Ethernet channels. This is complemented by devices running at 3.3, 2.5 and 1.9Tbit/s. "The 6.5Tbit/s part has four pipelines," Doe pointed out, "while the 3.3Tbit/s device has two. The other parts are created using lower clock frequencies and through speed binning. It proves that chips don't have to be fixed function to be cost and power efficient and fast."

Speeds and feeds – port density – are of interest to big network operators, Doe contended. "Our programmable approach allows the life of hardware to be extended because it does away with 'rip and replace' if an operator needs to go to 100G, which could be in a couple of years."

The approach has also been enabled by process technology. "If you roll back 20 years," Doe offered, "packet forwarding logic used to occupy most of the chip. Now, it's a very small amount, so we can add programmability for no cost." Tofino is being produced on TSMC's 16FF+ process and devices have been sampling since last year.

"OEMs can take advantage of the Tofino approach," Doe concluded. "Today, they have to do things the hard way – fixed chips for specific functions. With Tofino, they can build a platform; it's an opportunity to change the delivery model so they don't have to obsolete equipment every couple of years."