DARPATech, DARPA's 25th Systems and Technology Symposium
August 7, 2007
Anaheim, California
Teleprompter Script for Dr. Jagdeep Shah, Program Manager, Microsystems Technology Office

COMMUNICATIONS: THE NEXT FRONTIER

» **JAGDEEP SHAH:**

Smoke signals, telephones ;

sign language,
cell phones ;

fiber optics, internet.

Communications .

Communication is ubiquitous; communication is the central fabric of human existence.

Communication is also the lifeline of the military.

Communication between the soldier in the battlefield and the command center,
between command centers, and around the globe through fiber optic and satellite networks: communication is the *glue* in Network Centric Warfare.

The impressive progress of communication technology in the last few decades leads to a

natural question:
what are the emerging challenges in communications:
10 years from now,
20 years from now?

As we all know, extraordinary advances in nanofabrication technology are leading to ultrahigh densities of functional units on a chip, a *billion* transistors on a single computer chip,
a *quadrillion* active molecular devices in a cubic centimeter…

These ultrahigh-density chips present new opportunities ,
and *new challenges*:
how will these ultradense, nanoscale devices communicate with each other, *and* with the outside world?

This is the next frontier of communications.

It is well known that photonics dominate global communications.

Can photonics enable *intrachip* communications between ultradense functional units on a chip *and* between them and the outside world?

Some say this is impossible.

My name is Jag Shah, and to that I say, “Nonsense!”

DARPA is poised to do just this: bring photonic communications directly onto the chip.

That’s right –
signaling with light directly on your microprocessor chip!

The information age in which we live was ushered in primarily by advances in photonic communication networks.

DARPA has unquestionably played a leading role in this field.

Optical technologies have dominated long-haul communications,
and have been making inroads into shorter-distance communications:
rack to rack,
board-to-board and even chip-to-chip in modern computer systems.

But data communication over short distances remains the domain of electrical signaling.

The microprocessor in your computer,
for example, works by sending electrical signals between transistors on the chip, and off-chip to the peripherals.

So why should we think about intrachip photonic communications for an ultrahigh density chip?

In order to gain some insight, let us consider a high performance microprocessor in 2020
as a <u>*proto-typical futuristic chip*</u> with an ultrahigh density of functional units.

Perhaps the most important challenge facing future generations of microprocessors
is the power dissipation on the chip.

It turns out that the power dissipation per unit area for a typical microprocessor chip exceeds that of a hot plate and, if historical trends continue, would be approaching that of a nuclear reactor in 2020.

This is <u>clearly</u> an untenable trend.

The power dissipation on a microprocessor chip comes from three major sources:
logic circuitry, memory, and communications, in roughly equal proportions.

My colleagues in the electronics and MEMS
are worrying about the first two.

The questions I want to address now are:

How much intrachip communication do we really need?...

And can electrical signaling provide this communication bandwidth within the thermal power budget?

There is a general rule, Amdahl's rule of thumb, that states that a system in which computation and communications are balanced performs the best under most circumstances.

The system balance can be characterized by the ratio of the communications bandwidth
(in Bytes/second)
to the theoretical maximum performance in operations/second or FLOPS.

Trends in microprocessors and supercomputers indicate that this ratio in Bytes/FLOp is decreasing significantly,
indicating a growing system imbalance.

The net result is
a growing gap between the *actual computer performance* and the *theoretical maximum performance* rating in GigaFLOPS that we

hear about.

To appreciate the communication requirements for our proto-typical microprocessor chip circa 2020, let us remember that the TeraFLOPS era is already here,
as evidenced by Intel's recently-announced
80-core Polaris
research chip.

Scaling trends will continue with hundreds or thousands of cores on a chip, and theoretical maximum performance of tens of TeraFLOPS (tens of trillions of operations per second).

Following Amdahl,
a 10 TFLOPS chip will need 10 TB/s or a truly astonishing 100 Tb/s of communications.

This is *twenty times* larger than the 5 Tb/s currently provided by the Global Information Grid, or commercial global networks.

Just think about it:
twenty times the current global communication capacity on a tiny chip measuring no more than a postage stamp!

Truly incredible!

Electronic communications on-chip will no longer be able to keep up with this enormous communications demand within the power dissipation budget.

Can intrachip photonic communication come to the rescue?

Photonics as a communication technology brings many benefits that make it a viable candidate for intrachip communications.

Photonics can carry information on many different wavelengths within the same waveguide with minimal cross-talk.

Photonics can operate at data rates much higher than electronics without distortion or significant loss, and photonics can also exploit space division multiplexing.

In spite of these benefits, the conventional wisdom has been that intrachip photonic communication is not viable because photonic devices are
too big, too power hungry,
too expensive, and incompatible with electronics.

So people say that intrachip photonic communication is
just a dream.

I say that the time is ripe to transform this dream into reality.

The time is ripe because EPIC, DARPA's program for developing electronic and photonic integrated circuits, is successfully developing high-performance photonic devices in Si using processes that are fully compatible with electronics.

Two and a half years into the EPIC program,
we are well on our way to demonstrating a tiny silicon chip with all the photonics and electronics necessary for a 100 Gb/s transceiver,
and a tiny silicon chip that can function as a
high-performance,
wide-frequency channelizer for electronic warfare applications...

Encouraged by the success of the EPIC program, and the huge wave of

activity it has generated in the field of silicon photonics,
DARPA is ready to take the next step.

MTO is about to launch an ambitious new program to develop technology for intrachip photonic communications in a futuristic microprocessor with theoretical maximum performance of 10 TFLOPS, and *beyond*.

The goal of the program, appropriately named UNÍC, is to enable communication bandwidth matched to the peak theoretical performance of the microprocessor,
so that we can
bridge the gap between the *actual* performance and the *theoretical maximum* performance of a computer system.

A number of formidable challenges must be addressed in order to achieve such truly astonishing performance, and we are looking for your ideas to meet these challenges.

The first among these is the issue of power dissipation.

A good way to think about this is the *total energy* dissipated on the chip for transmitting 1 bit of information.

For distances exceeding a few millimeters,
the current value for electronic communications is
~25 pJ/bit and future generations may reduce this by an order of magnitude.

Preliminary studies show that photonics communication links can operate at 100 fJ/bit,
with some credible ideas for bringing it down to
10 fJ/bit.

This would be a *hundred* times lower energy per bit than projected for electronics!

Another way to look at this is that a 10 TFLOPS microprocessor chip requiring 10 TB/s of communication bandwidth to maintain good system balance could dissipate less than *one watt*
on the chip!

Thus, photonics communications can more than keep up as the microprocessor theoretical maximum performance scales to 100 TFLOPS and beyond!

One may naturally ask: what is the fundamental limit on the energy per bit for an optical link?

We recall that Shannon's theorem puts the lower limit at kTlog(2), equivalent to
3 *zeptoJoules* or
3 *millionths* of femtojoule!

The energy of a
1500 nm photon is
130 zeptoJoules.

If 100 photons are required to represent each bit with sufficient accuracy, that is still a *thousand times* smaller than the 10 fJ/bit we were just talking about, indicating that most of the energy is still consumed by factors such as modulator efficiency, detector efficiency and waveguide loss.

We are a long ways from the fundamental limit!

There are other challenges.

Current photonic devices are, by and large,
quite big.

The dimensions of these devices must be reduced so that a sufficiently large number can be packed on a small chip.

One fundamental limit on the size is the
optical wavelength,
roughly 400 nm for telecom wavelengths
in silicon.

This is a good goal to strive for, but even this may not be the ultimate limit!

There is the whole emerging field of
sub-wavelength optics and plasmonics.

These represent some exciting possibilities and exciting research opportunities.

Finally, the need for small device size and low power dissipation must be balanced with the need for operating over a broad spectral range under the temporal and spatial temperature variation expected on a microprocessor chip.

How will the photonic devices maintain their functionality over a wide temperature range?

These are formidable challenges indeed!

But please join me in imagining what we could accomplish if we met these challenges!

By bringing balance between computation and communication,
we would be able to bring microprocessor performance to *unimaginable* heights.

We could have a
chip-scale microprocessor with compute performance exceeding that of a
full rack of the current
IBM Blue Gene,
the basic building block of the highest performance Top 500 supercomputing system.

Imagine a 10 TFLOPS embedded processor that would consume only a few hundred watts, could be flown on micro-UAVs and could process 100 times larger radar images than currently possible –
in real time.

Imagine how such a technology could transform applications such as autonomous operations and
high-bandwidth streaming computing.

Imagine such a processor as a node of a
super supercomputer with ZettaFLOPS of compute power and imagine what it could do for scientific applications,
weather prediction and cryptanalysis.

The possibilities are limitless.

I invite you to bring me your ideas on how to meet the challenges we have discussed.

The payoff is enormous!

Let us *together* begin a journey to meet the challenges and transform this dream into reality!

And now I would like to introduce my colleague, Dr. Amit Lal.