Can we program machines we can build?

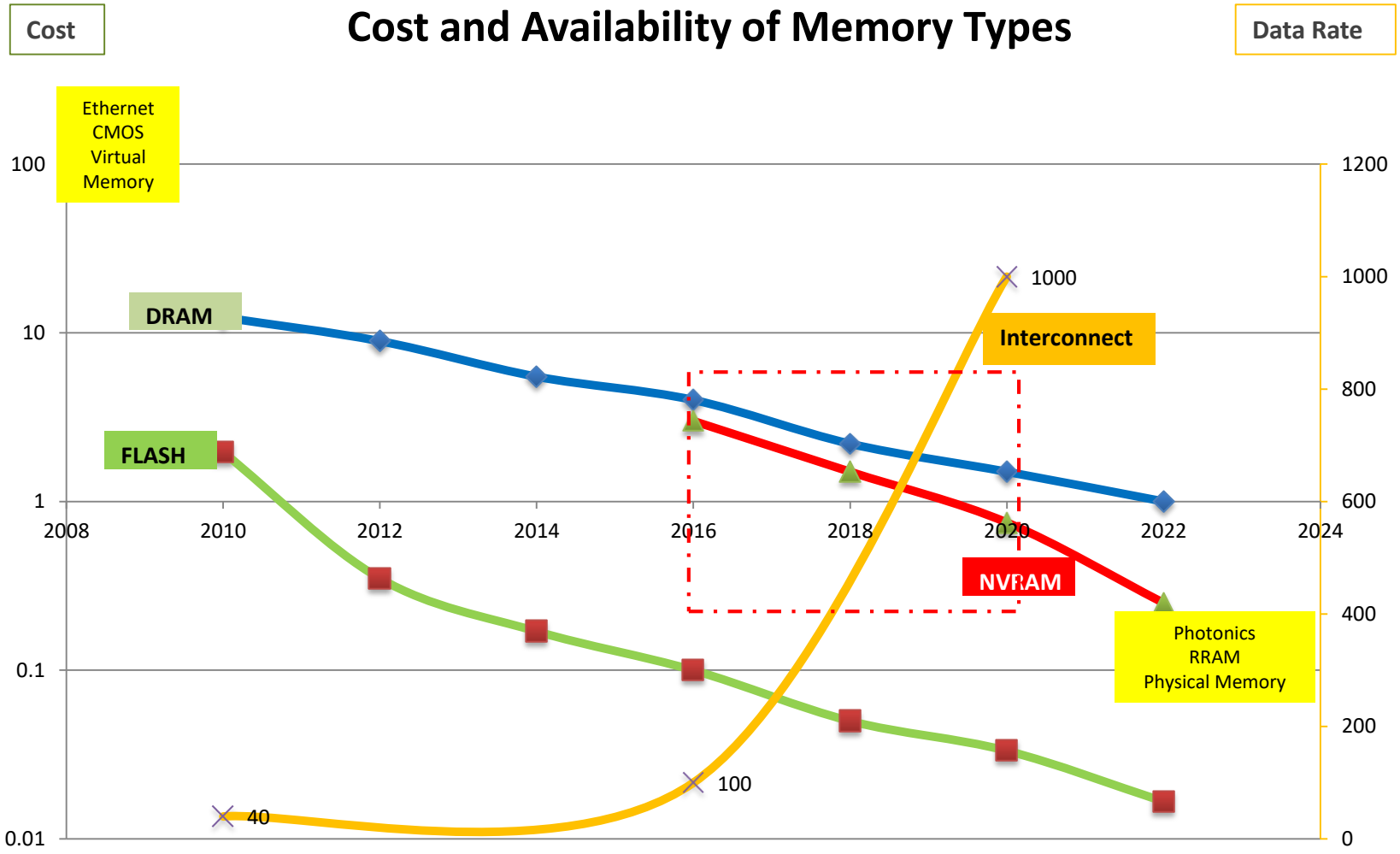
Renu Raman

HPTS

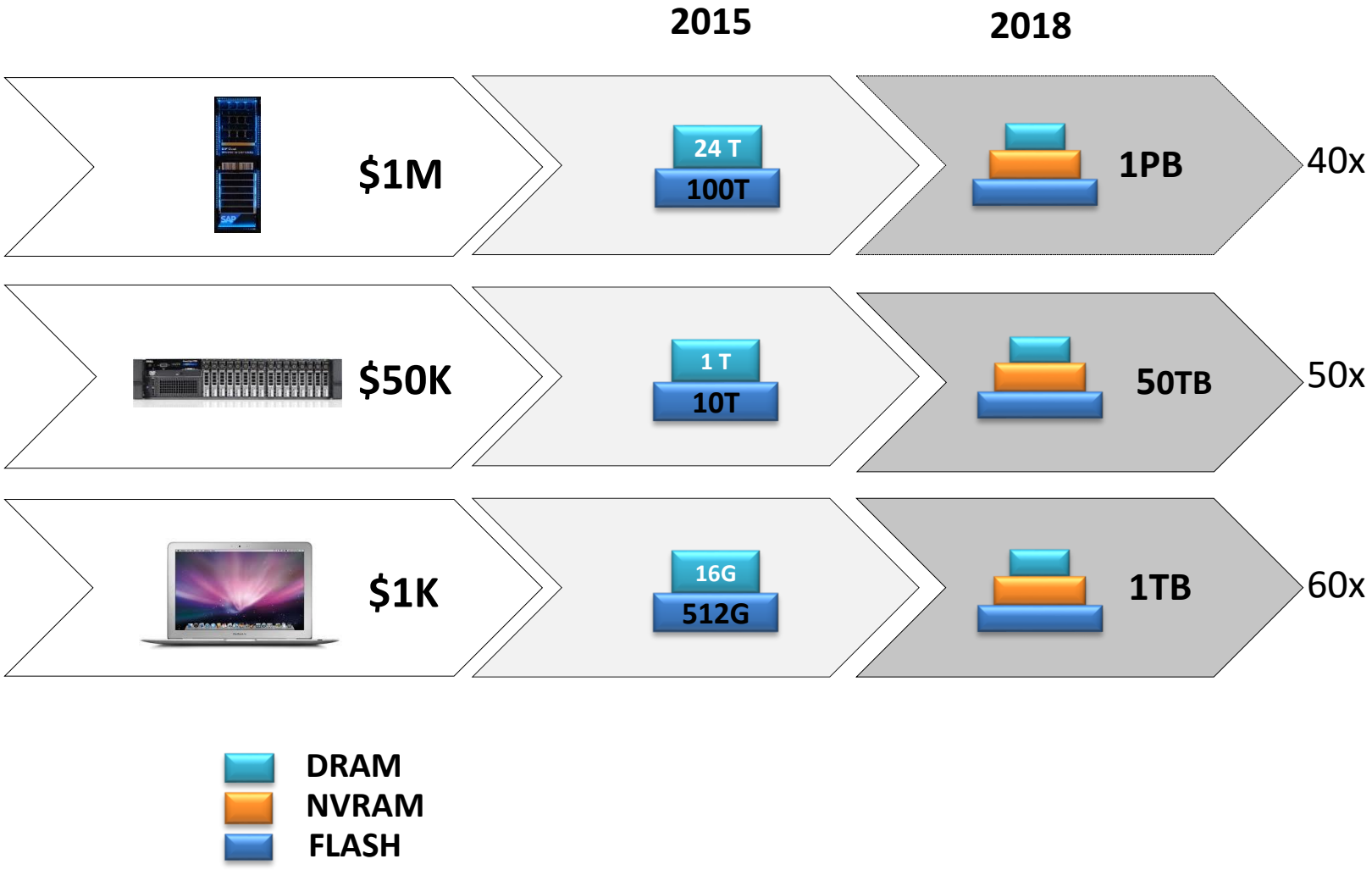
IT'S MEMORY STUPID

BUT THERE'S MORE TO MOORE…

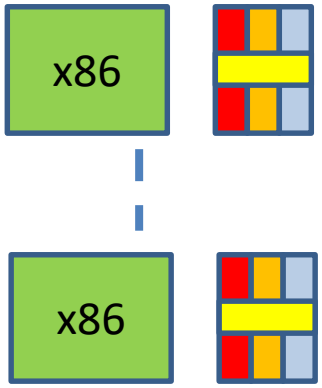
Technology Trends

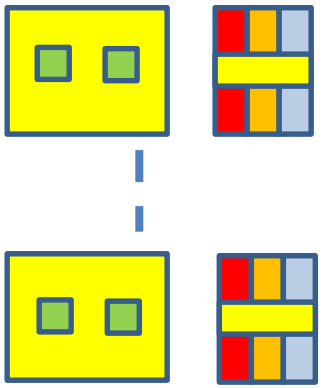


Implies….
Memory only systems



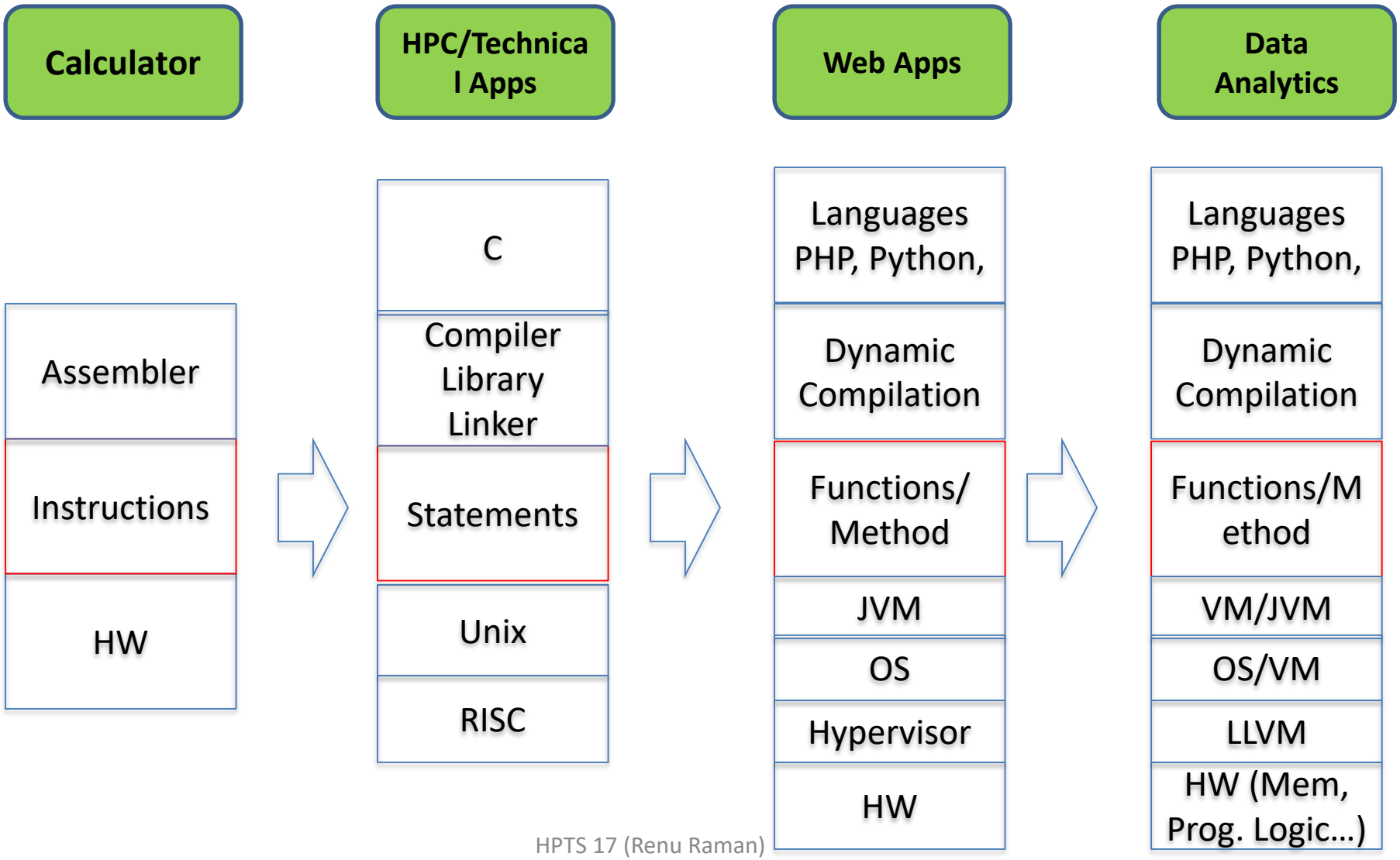
By 2020: $1M can get you…



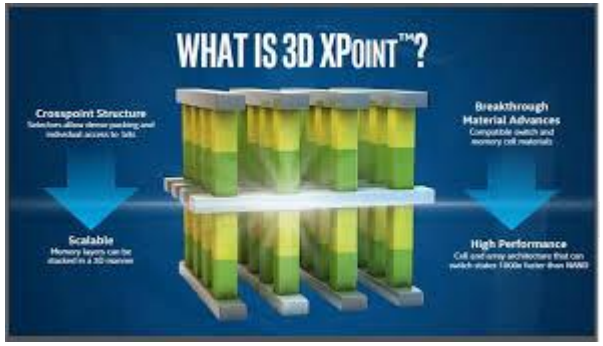


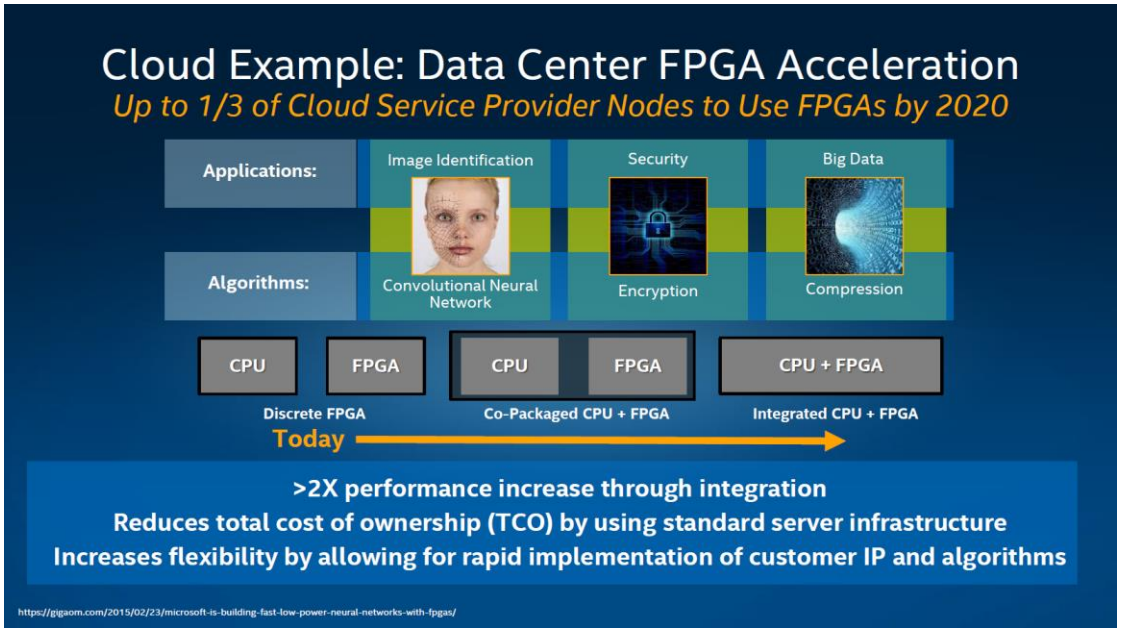
	48U	48U	48U
Master CPU cores	8000	8000	8192
Slave CPU cores		10240	65536
DSP Slices (M)		3.74784	28.016
LUTs (Millions)		382.1568	3516
L1 Memory	80	80	128
L2 Memory	640	640	1024
L3 Memory		1920	2048
Storage (TB)	2048		

Apps – That drove Platform adoption



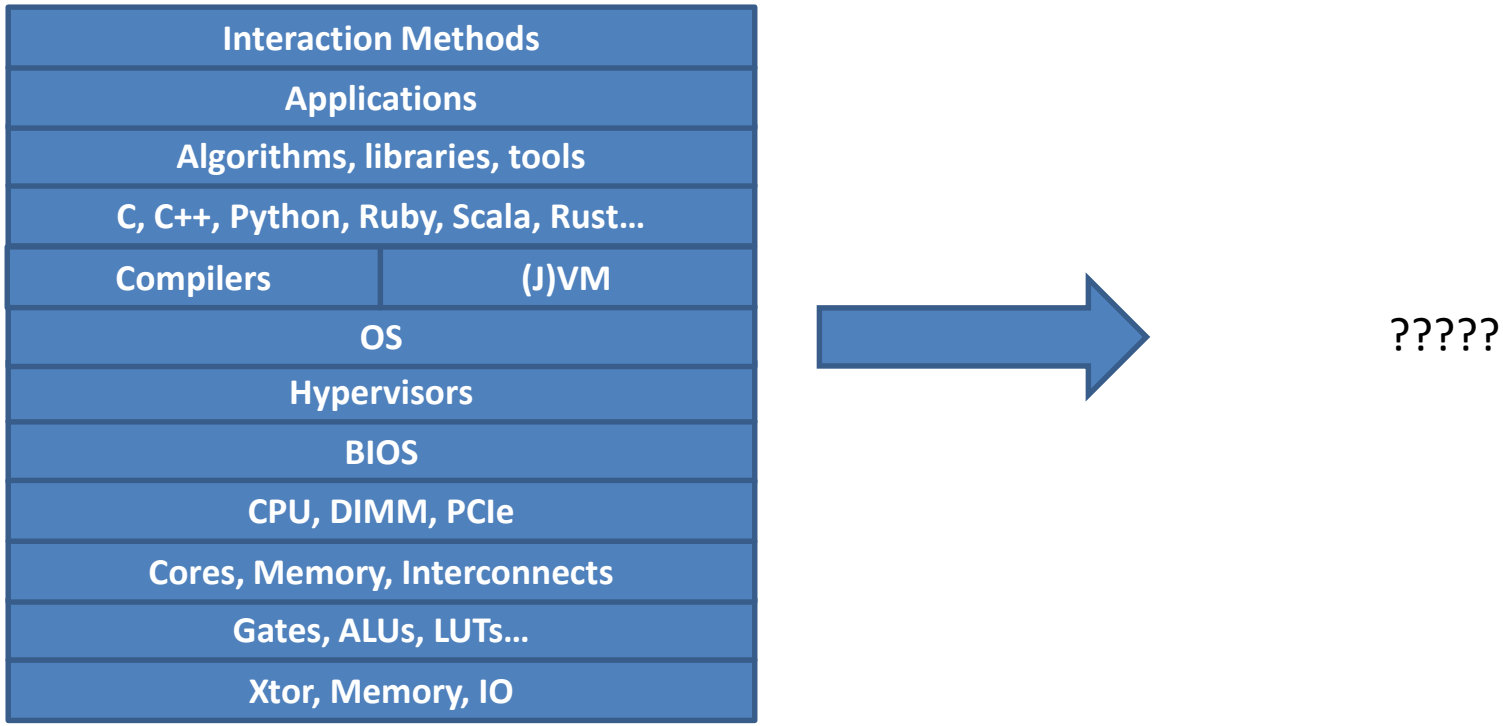
Memories - Changing Compute Landscape?





- 3D xpoint and other similar memories could enable a programmable structure in the metal stack. That could change the FPGA architecture and design (2020+)

Simpler is better



Simpler is better

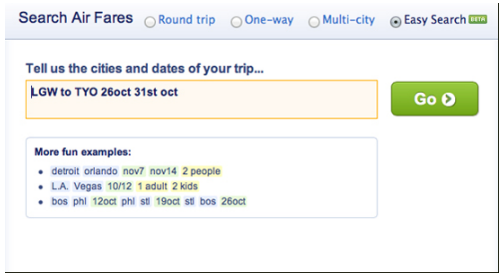
Programs vs Interactions

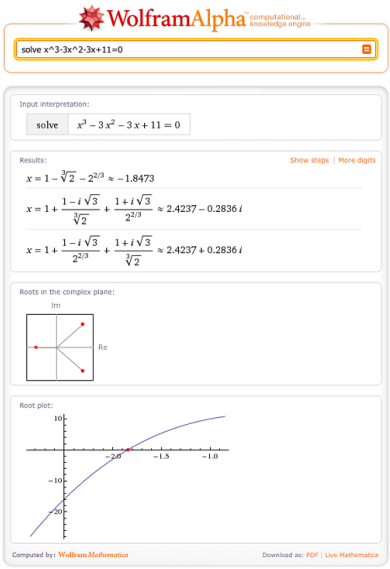
```
#include<stdio.h>
#include<conio.h>
main()

{
 clrscr();
 printf("I am a new programmer");
 getch();

}
```

Example of a Java program

```
class SomeNumbers
{
    static int square (int x)
    {
        return x*x;
    }

    public static void main (String[] args)
    {
        int n=20;
        if (args.length > 0) // change default
            n = Integer.parseInt(args[0]);
        for (int i=0; i <= n; i++)
        {
            System.out.print("The square of " + i + " is ");
            System.out.println(square(i));
        }
    }
}
```

```
% Transmission loss
clear all; close all; clc

rho=.6; c= 486; % c [m/s] and rho have been computed
                %for a temperature of 600 Fahrenheit

% cross-sectional area of pipe
a1=.025;S1=pi*a1^2; % m
a2=.015;S2=pi*a2^2; % m
a3=.025;S3=pi*a3^2; % m
L1=.5; L2=0.15; L3=0.1; % m

freq=linspace(1,11000,1000);

for ii=1:1000
    w=freq(ii); k=w/c;
    f(ii)=w/(2*pi);

    % With muffler
    Tout = [ 1 ; 1 ]; % dummy values which don't influence
                      % the final result of the transmission loss
    T1=[ cos(k*L1)   j*rho*c/S1*sin(k*L1) ;  j*S1/rho/c*sin(k*L1)  cos(k*L1) ];
    T2=[ cos(k*L2)   j*rho*c/S2*sin(k*L2) ;  j*S2/rho/c*sin(k*L2)  cos(k*L2) ];
    T3=[ cos(k*L3)   j*rho*c/S3*sin(k*L3) ;  j*S3/rho/c*sin(k*L3)  cos(k*L3) ];

    T_tot_1=T1*T2*T3;

    Tin = T_tot_1*Tout;

    Powin(ii) = .5 * Tin (1,1) * Tin (2,1);
    Powout(ii) = .5 * Tout (1,1) * Tout (2,1);

end
loss = -10*log10(Powout./ Powin);

plot(f,loss)
xlabel('Frequency [ Hz ]')
ylabel('dB')
title('Transmission Loss')
```





Thought – Compute Gap

Interactive Cloud	Cancer Genomics	Machine Learning	Graph Processing	Multimedia Analysis	Computer Vision	Software Radio

Efficiency Layer

Concept Languages
Productivity Languages (Python, Scala)
~~Efficiency Languages (C++/C/assembler, CUDA/OpenCL, Java)~~ New Programming Frameworks (Tensorflow, DAWN…)

- Productivity programmers want to concisely express problem in application domain
- Efficiency programmers focus on packaging efficient frameworks and libraries for use in productivity layer
- Expressive Programmers Just Think

Some Reminders…

- We always overestimate the change that will occur in the next two years and underestimate the change that will occur in the next ten. Don't let yourself be lulled into inaction. Bill Gates
- "If you throw enough hardware at the problem, you can get interactivity" – BigDAWG Stonebraker
- "Analyses is dead, Synthesis is King" – Anastasia Ailamaki
- Machine learning needs parallel hardware and parallel hardware needs machine learning - Roland Memsevic (Univ. of Montreal)
- People who are really serious about software should make their own hardware. Alan Kay
  - People who are really serious about hardware should make their own software (Yours Truly☺)