

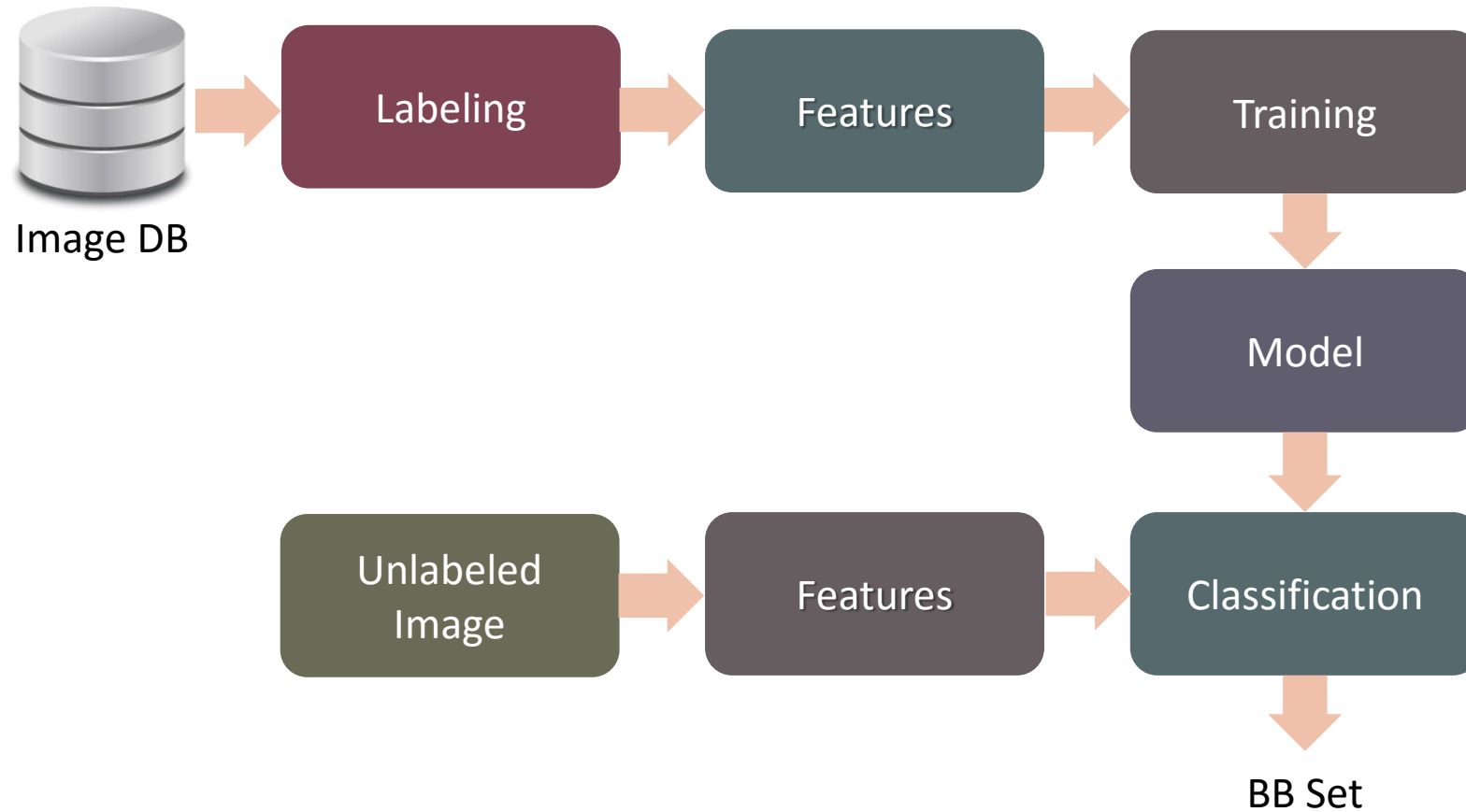
A Scalable and Power-Efficient Sliding-Window Model on Intel Processor Graphics

AVI BLEIWEISS

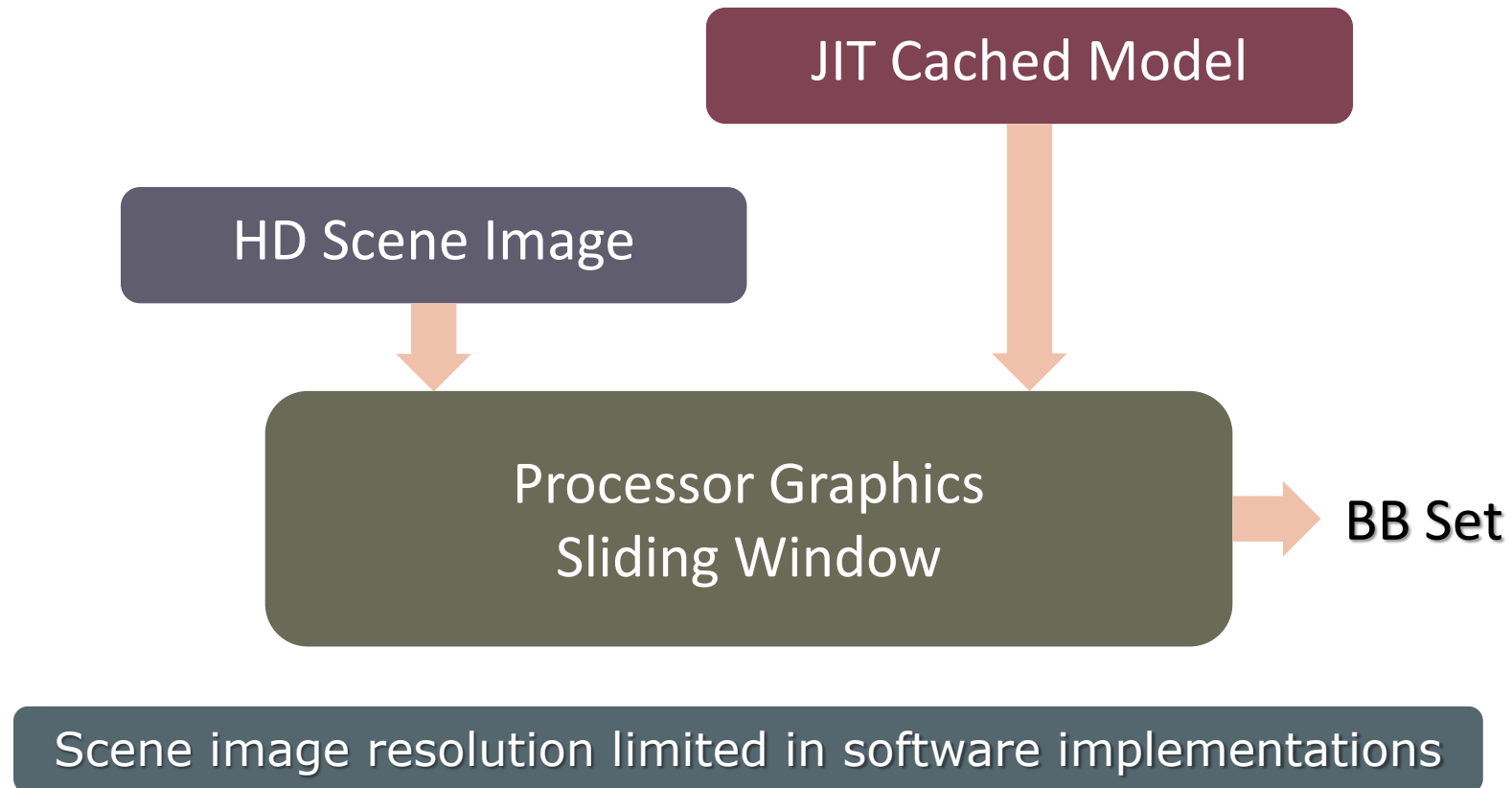
OCTOBER 2012



Software Architecture

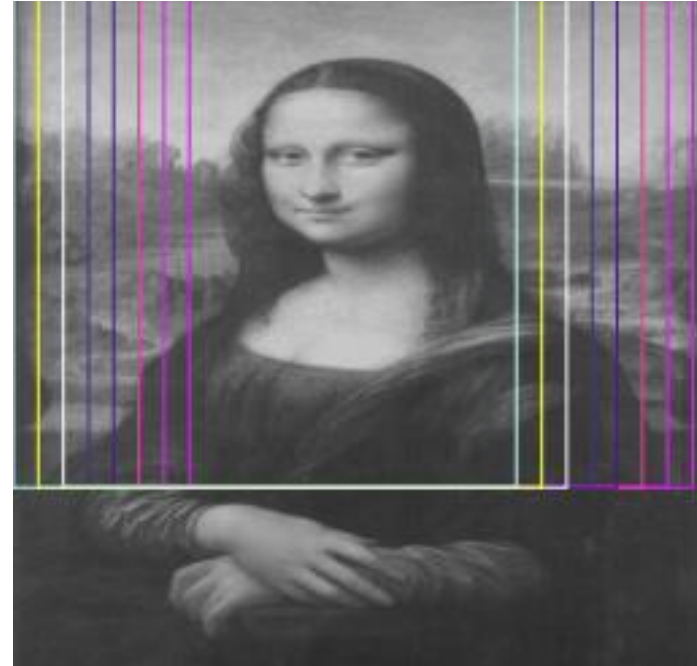


Acceleration



Sliding Window

```
B = 0
for (s in scales)
  for (y in height)
    for (x in width)
      if (classify(x, y, s))
        B = B | bbox(x, y, s)
      endif
    endfor
  endfor
endfor
```



Core to object detection applications

Parallelism

Window threads autonomous

Hundred thousands classifications

Bulk resources shared, read-only

Thread private, split object buffer

Classify Kernel

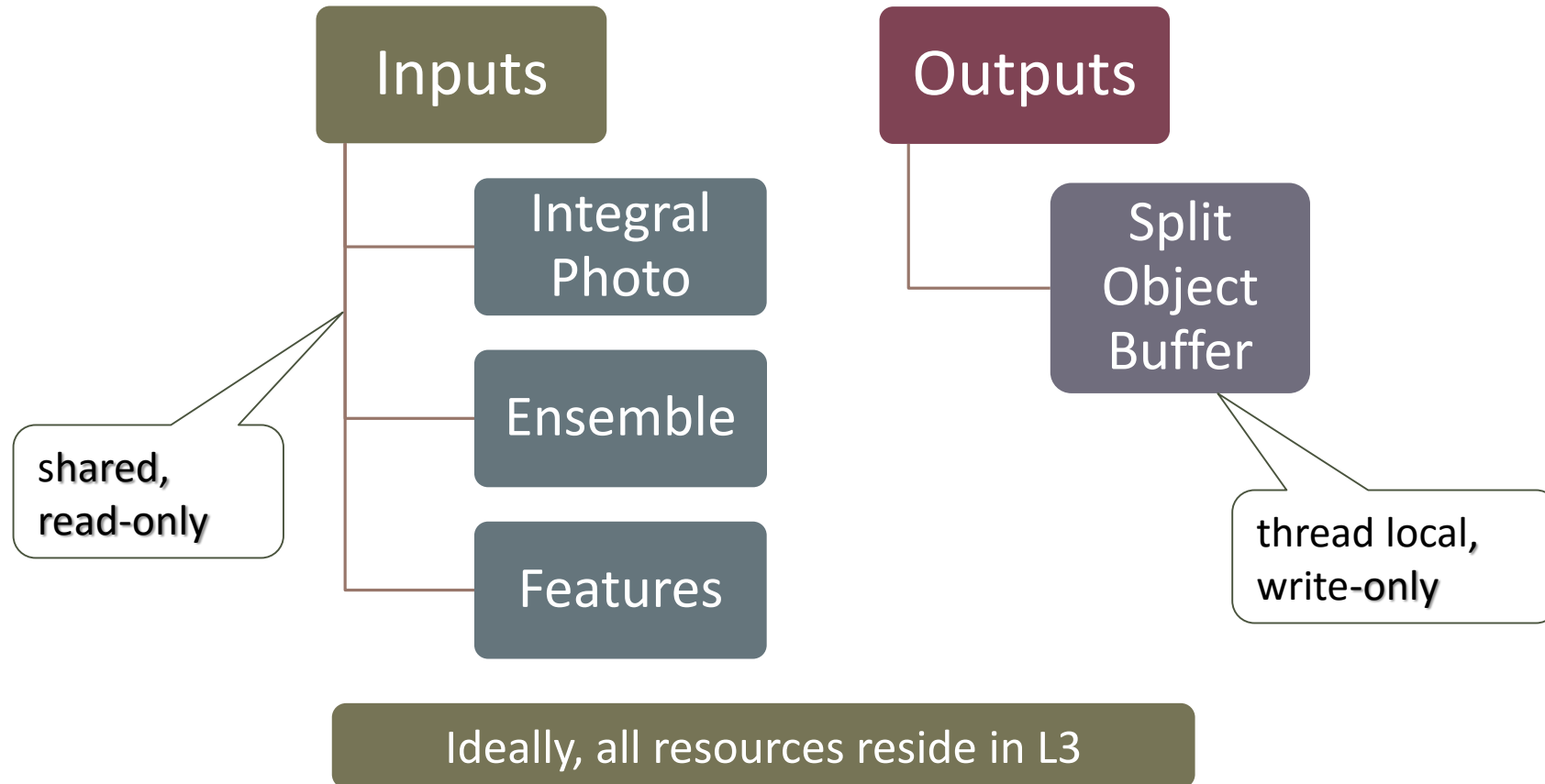
$$H = \sum_{i=0}^N \alpha_i h_i \quad (N = 200)$$

$$h = f\left(\iint \text{Image}(x, y, s)\right) \geq \tau ? +1 : -1$$

Viola-Jones

Series of Integral Image operators, large dot product

Working Set



Feature Compute Sample

```
inline _GENX_  
float r2c2(SurfaceIndex photo, float variance,  
          const Rect r0c0, const Rect r0c1,  
          const Rect r1c0, const Rect r1c1,  
          float x, float y, float scale)  
{  
    float f = .0f;  
    f += Black * integral(photo, r0c0, x, y, scale);  
    f += White * integral(photo, r0c1, x, y, scale);  
    f += White * integral(photo, r1c0, x, y, scale);  
    f += Black * integral(photo, r1c1, x, y, scale);  
    return f / variance;  
}
```

Experimental Setup

Compute platforms

- C-for-Media 2.4 – runtime and inline compiler
- OpenCL 1.0 – offline compiler

Processor Graphics

Former Code Name	Graphics Generation	Frequency (GHz)	Skew	EUs	Max Threads	Machine Instructions
Sandy Bridge (SNB)	PG6	1.3	GT2	12	60	869
Ivy Bridge (IVB)	PG7	1.1	GT2	16	128	859
Haswell (HSW)	PG7.5 [†]	0.8	GT3	40	280	859

[†]pre-silicon, extrapolated results

Scalability

IPS increasingly less relevant

Compute energy

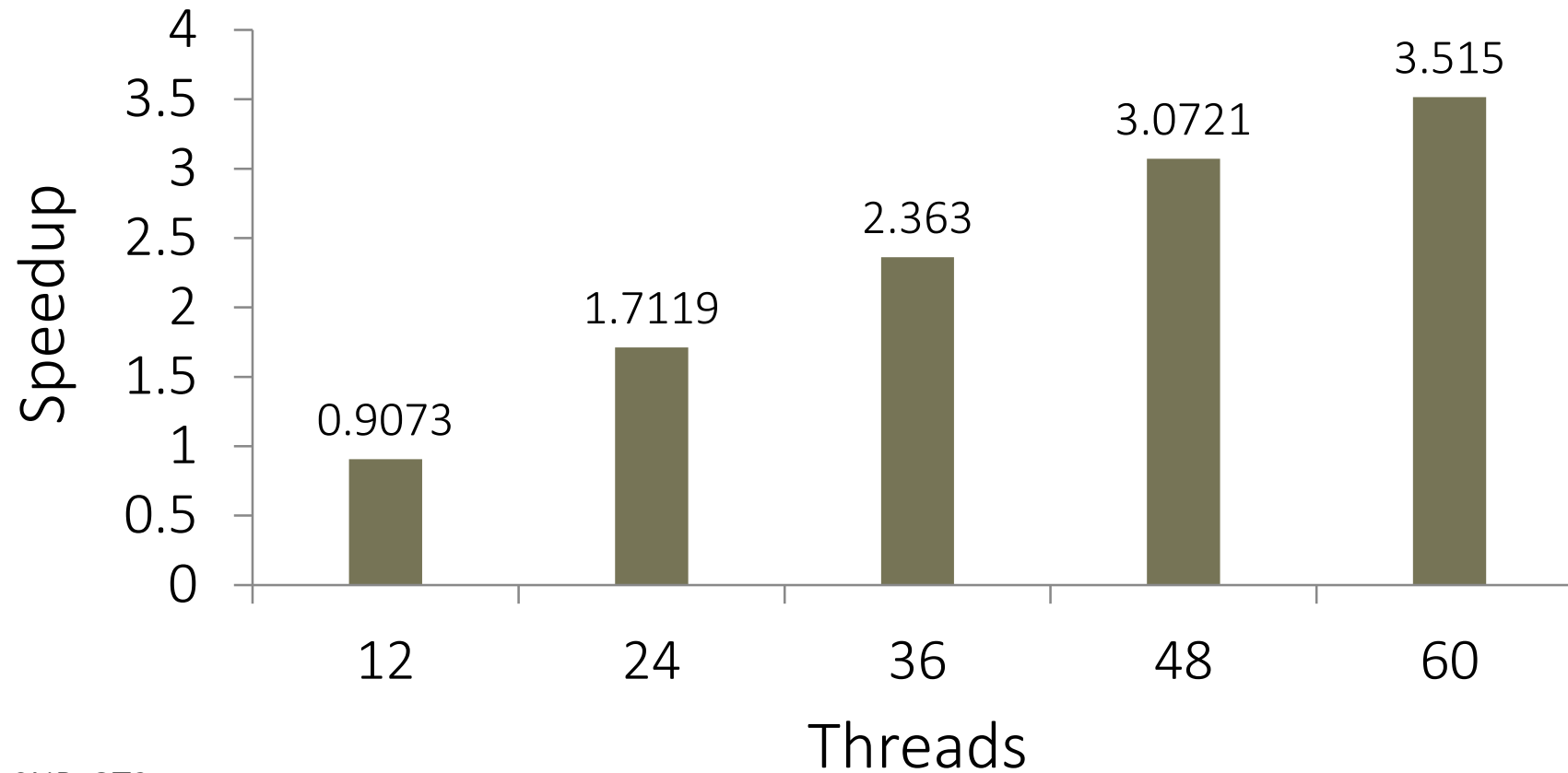
$$\frac{IPS}{Watt} = \frac{Instructions/sec}{Joules/sec} = \frac{Instructions}{Joule} = IPJ$$

Gating, normalized IPJ[†]

SNB	IVB	HSW
1.0	1.12	1.79

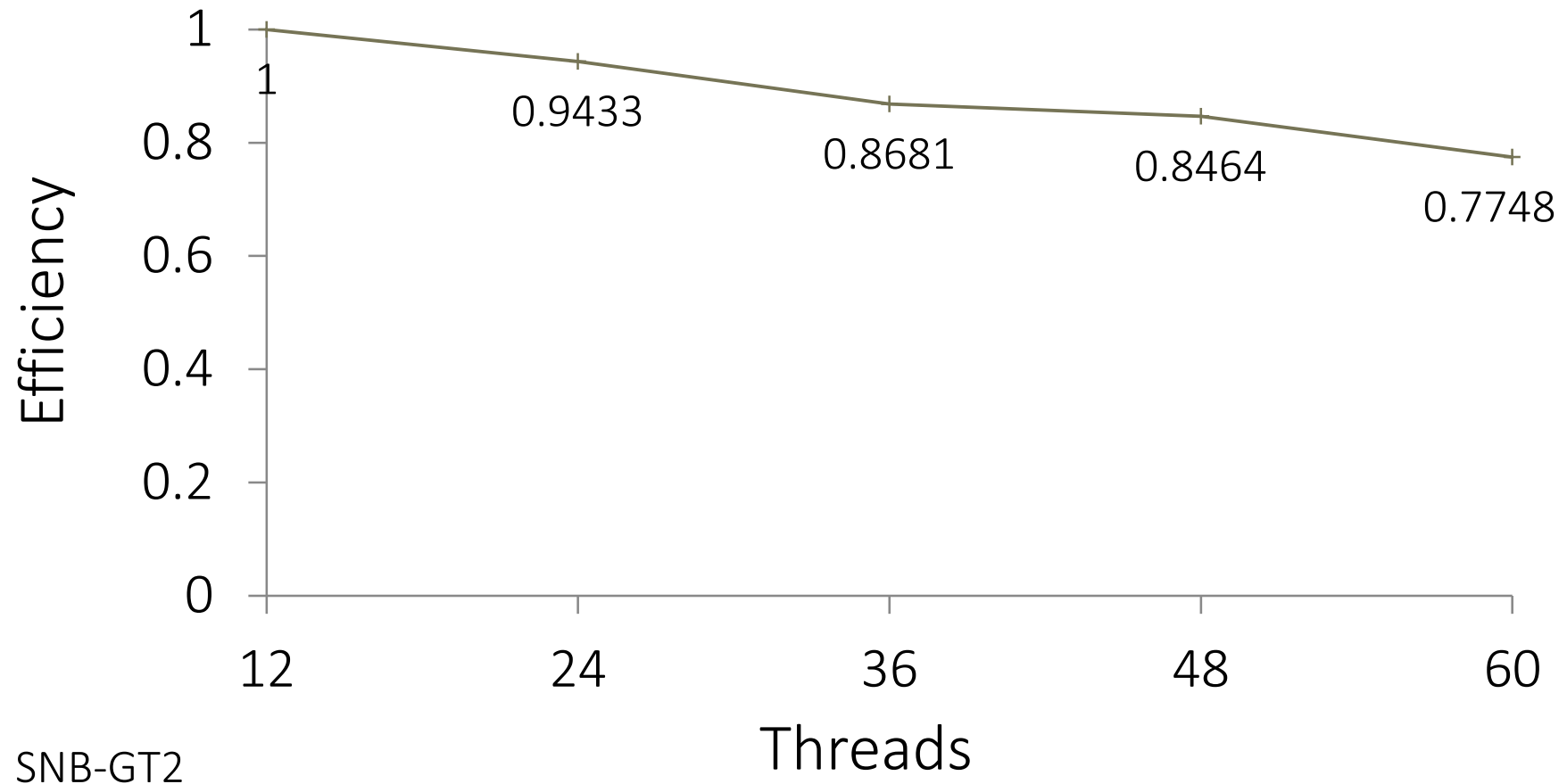
[†]assuming identical power envelope

PG6 vs. 4-Threaded, 4-Core IA

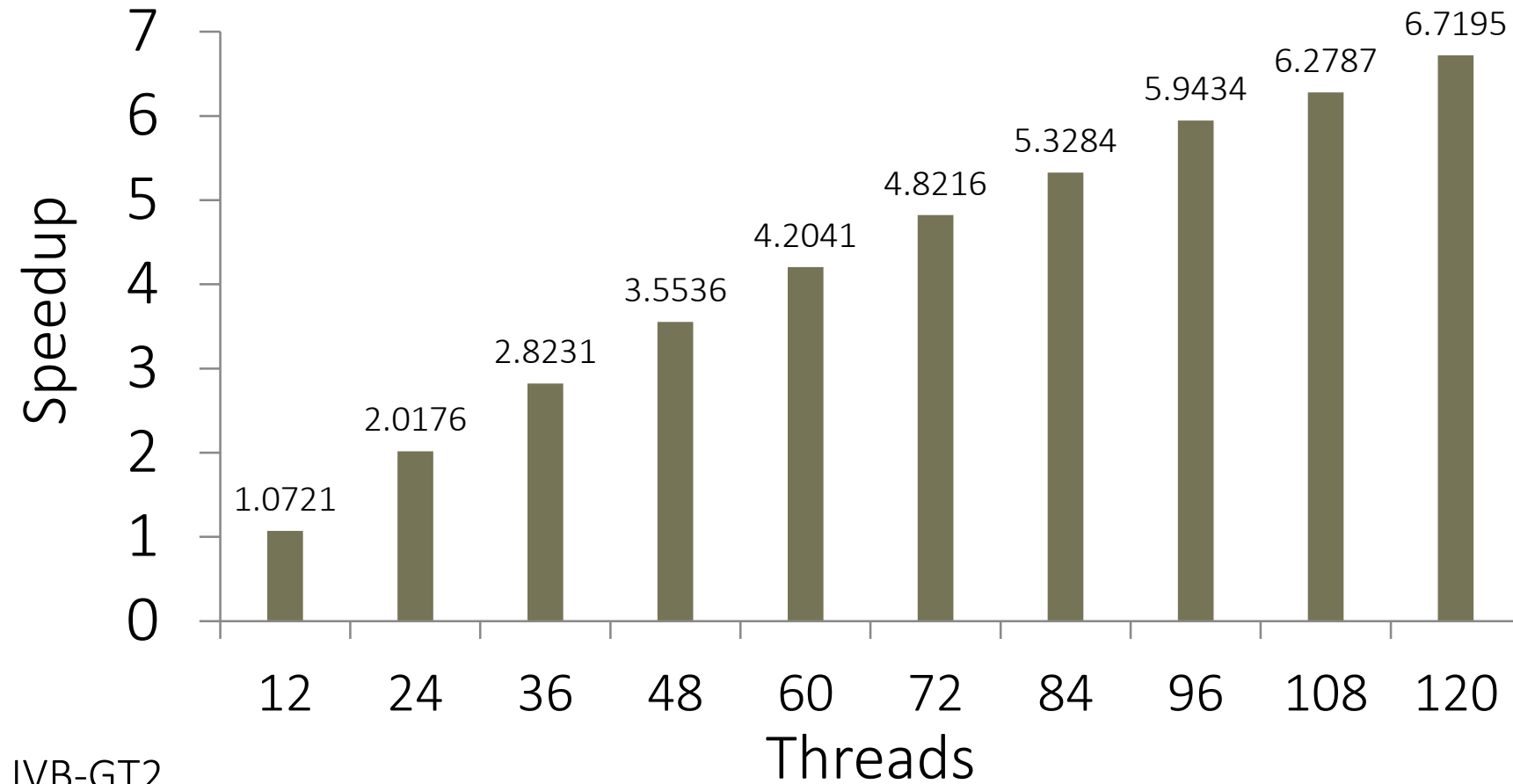


SNB-GT2

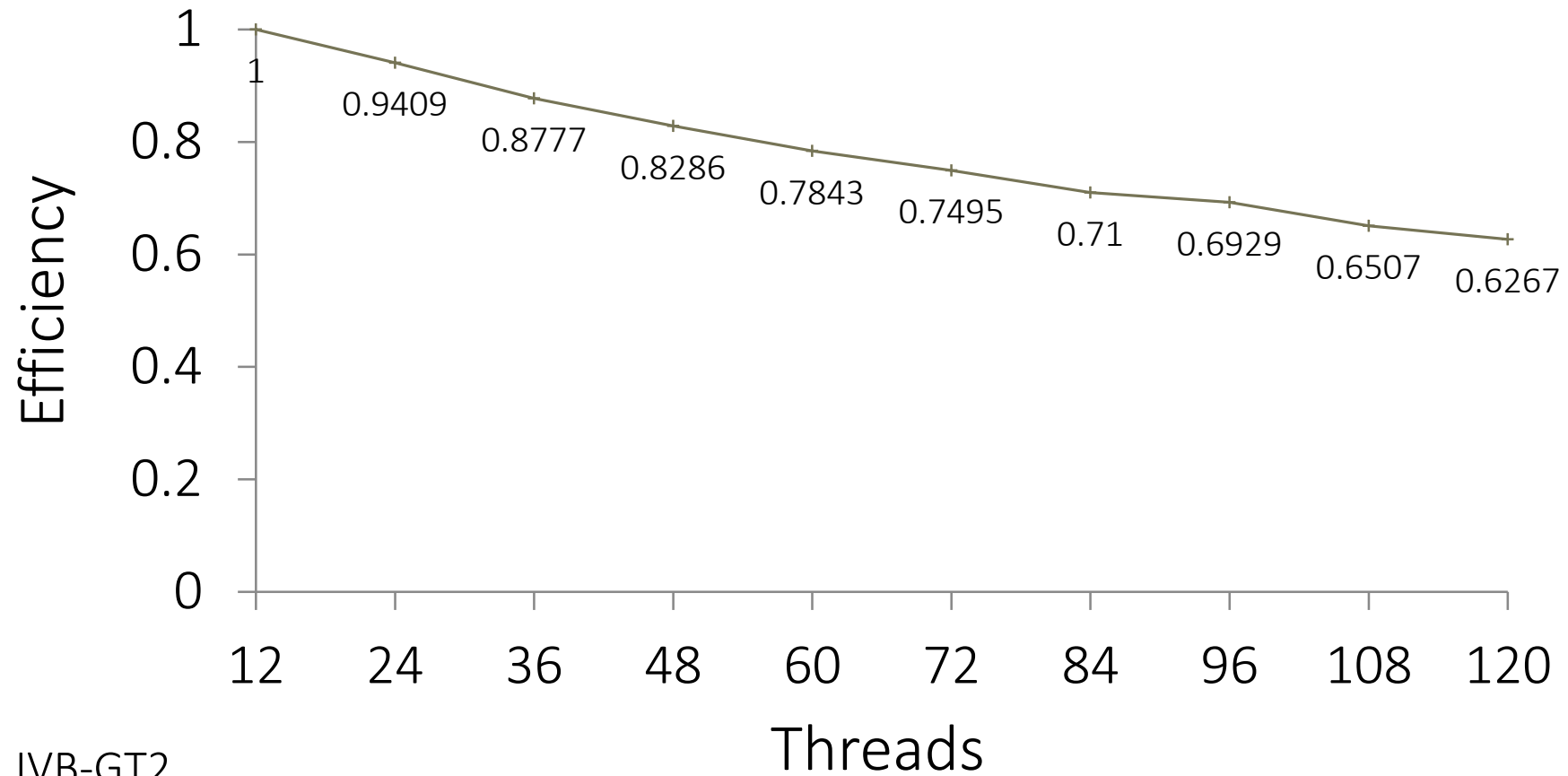
PG6 Parallel Efficiency



PG7 vs. 4-Threaded, 4-Core IA

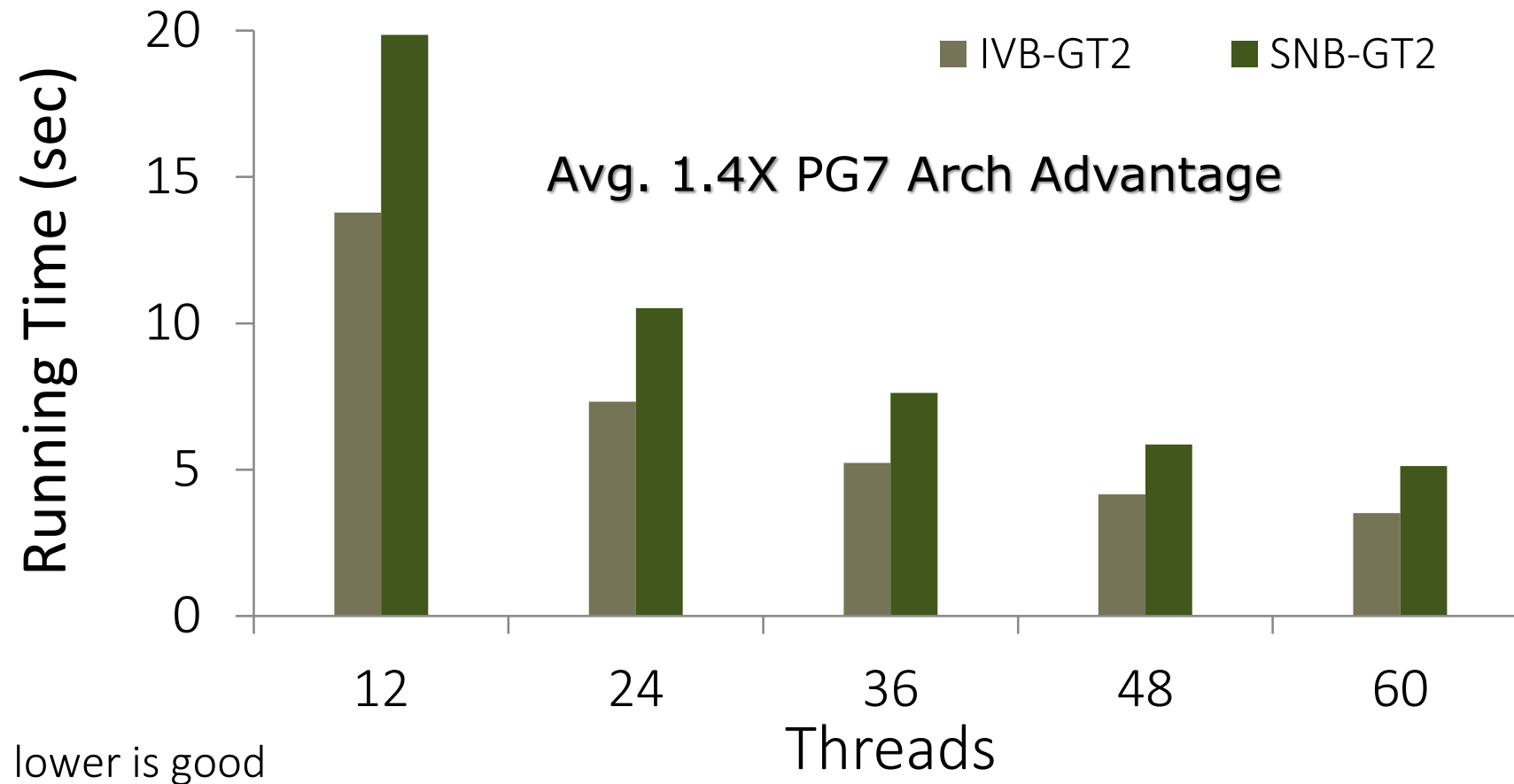


PG7 Parallel Efficiency

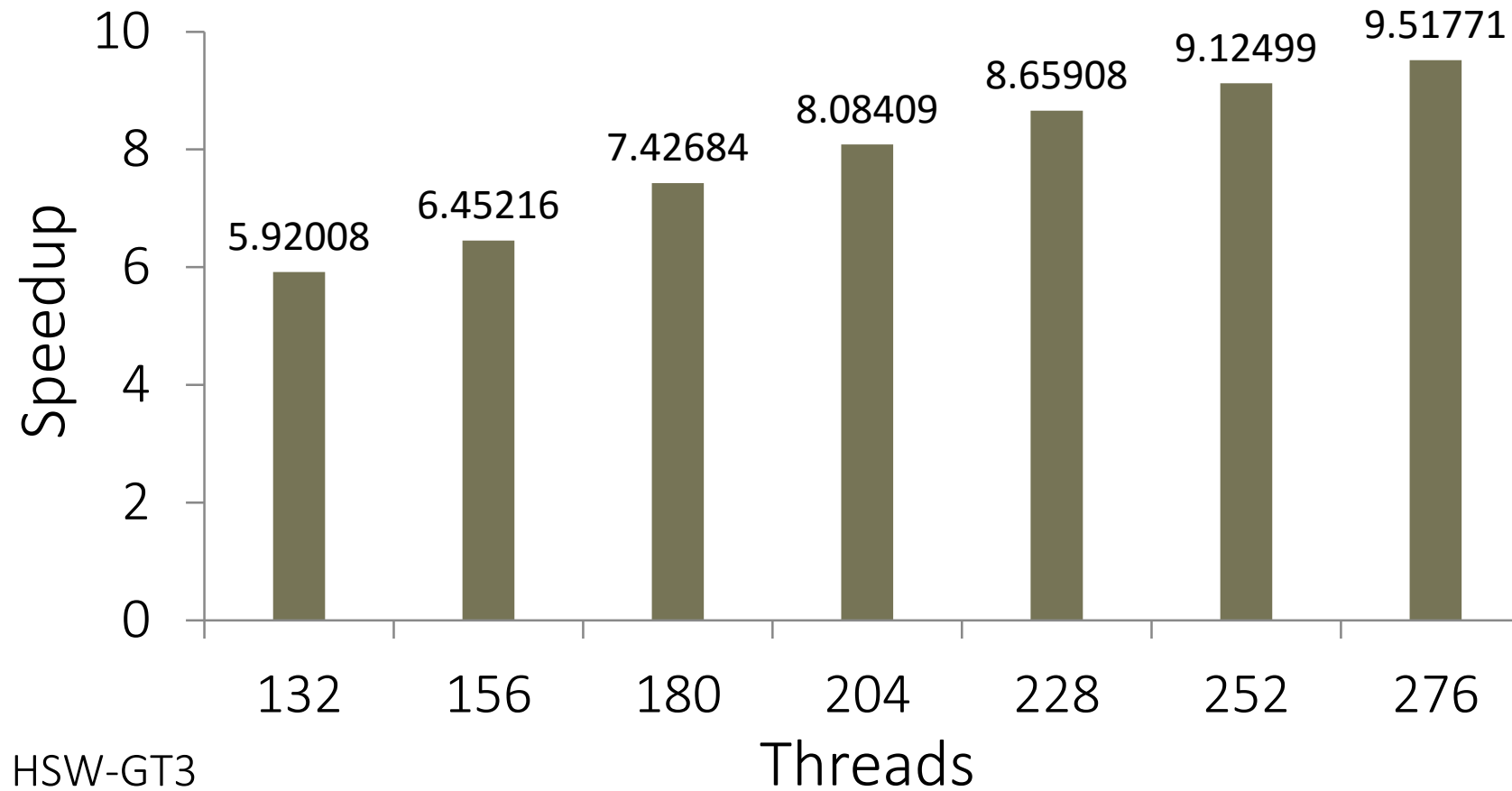


IVB-GT2

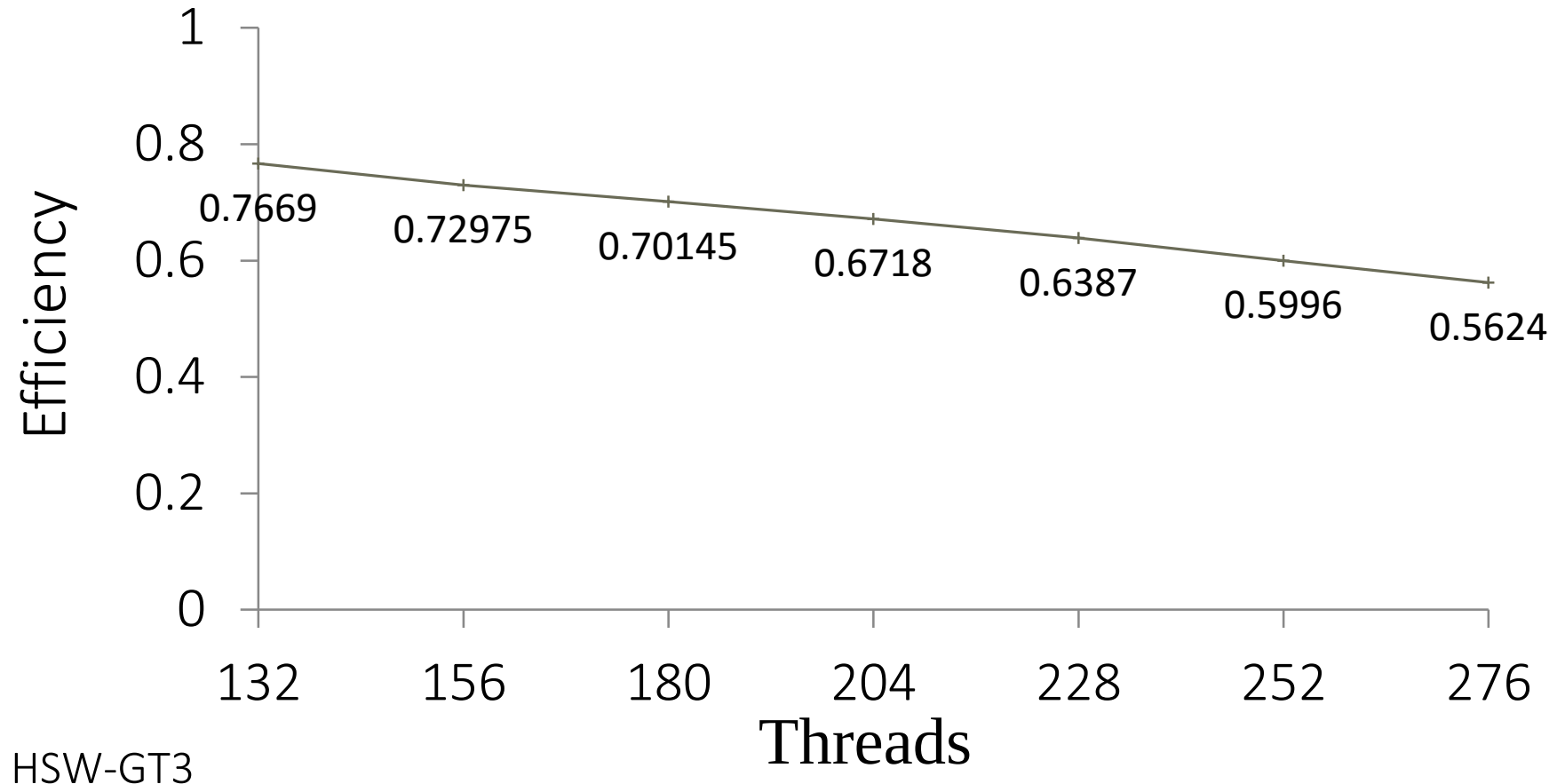
PG7 vs. PG6 Running Time



PG7.5 vs. 4-Threaded, 4-Core IA



PG7.5 Parallel Efficiency



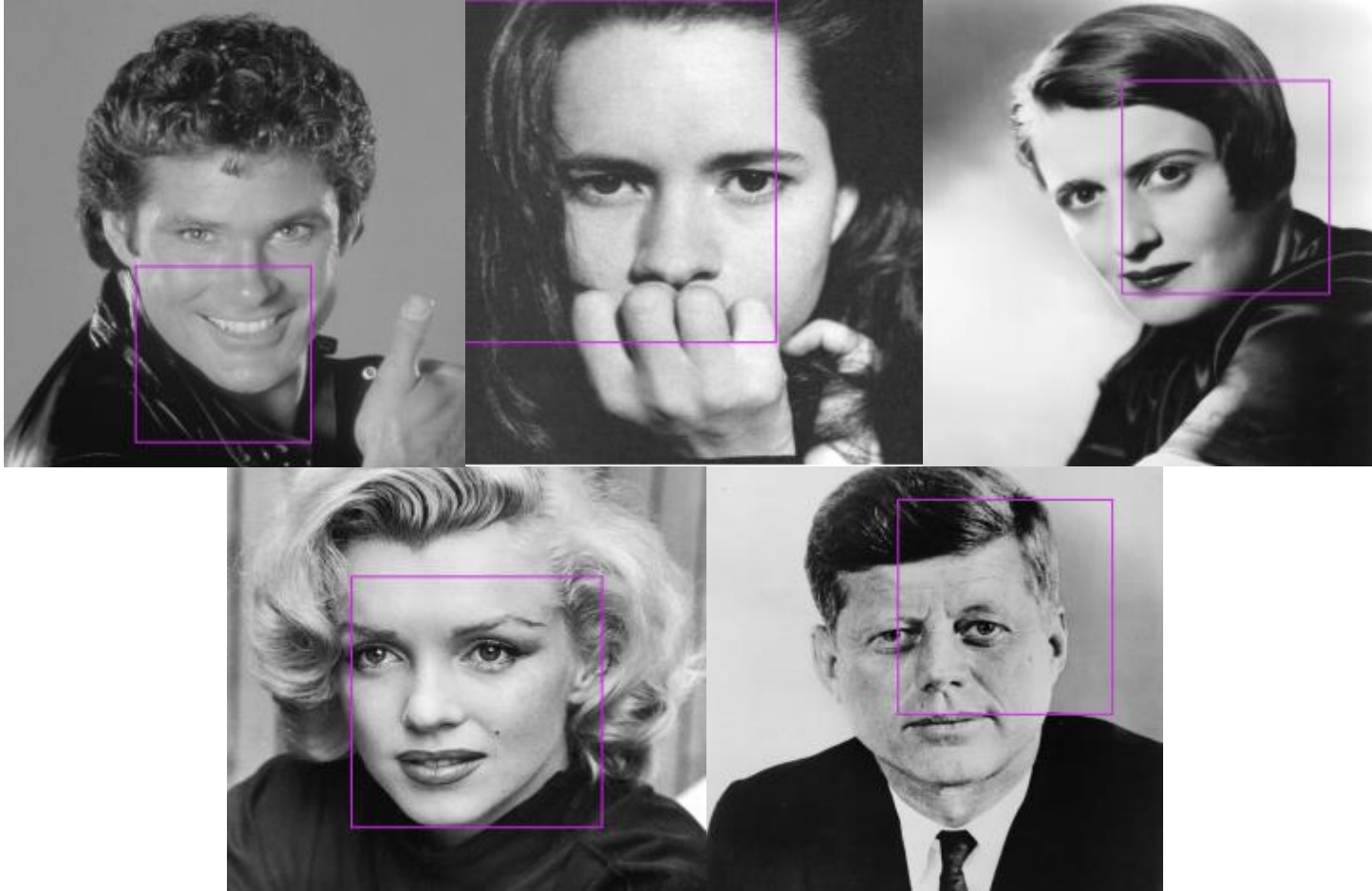
Improvements

Sustainable load balance

Content split object buffer

Kernel optimization

Compact photo representation



Summary

Sliding Window broad value proposition

IVB vs. SNB advantage attributed to L3

HSW vs. IVB 1.4X max parallelism gain

Seek increased scalability on BDW and SKL[†]

[†]BDW – Broadwell, SKL - Skylake