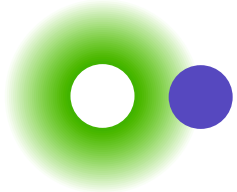
Dynamic Voltage and Frequency Scaling
The Laws of Diminishing Returns



Etienne Le Sueur and Gernot Heiser

HotPower'10, Vancouver, Canada, 2010
etienne.lesueur@nicta.com.au

What is DVFS?

Dynamic Voltage and Frequency Scaling

What is DVFS?

$$P = CfV^2$$

What is DVFS?

dynamic power consumption

$$P = CfV^2$$

What is DVFS?

dynamic power consumption

$$P = CfV^2$$

constant

What is DVFS?

dynamic power consumption

$$P = CfV^2$$

constant

frequency

What is DVFS?

NICTA

dynamic power consumption

voltage ***squared***

$$P = C f V^2$$

constant

frequency

In the past...

"Under some conditions, we observe energy savings of 30% for a 4% performance loss."

Snowdon et al. [2009]

In the past...

"... in most of the traces the potential for energy savings is good. The savings range from about 5% to about 75%, with most data points falling between 25% to 65% savings."

Weiser et al. [1994]

In the past...

"Energy savings of 22% ... to complete the same task are possible without a substantial reduction in application performance..."

Weissel et al. [2002]

In the past...

"... Hence, on this system, by lowering the frequency to a point where the workloads can be adequately served without sacrificing latency, energy is saved."

Miyoshi et al. [2002]

The whole story

$$P = CfV^2 + P_{static}$$

The whole story

total power consumption

$$P = CfV^2 + P_{static}$$

The whole story

total power consumption

$$P = CfV^2 + P_{static}$$

dynamic power consumption

The whole story

total power consumption

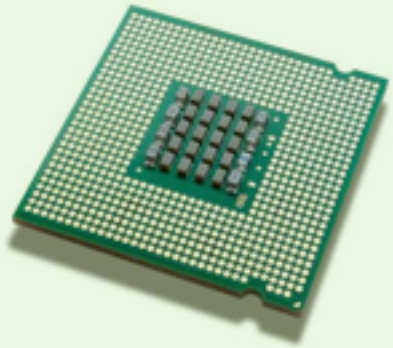
static power consumption

$$P = CfV^2 + P_{static}$$

dynamic power consumption

Static power consumption

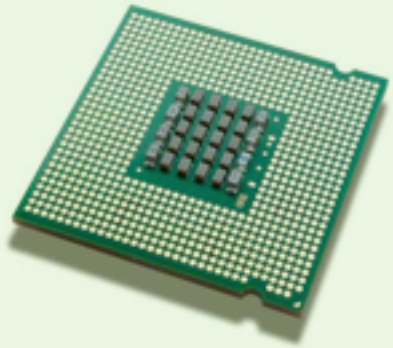
Static power consumption



CPU leakage

Static power consumption

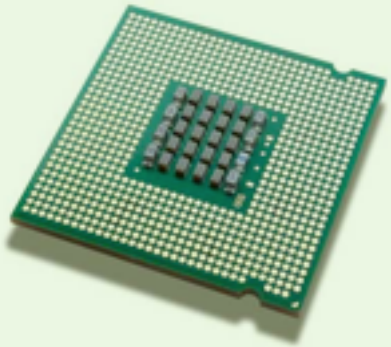
Hard drives



CPU leakage

Static power consumption

Hard drives

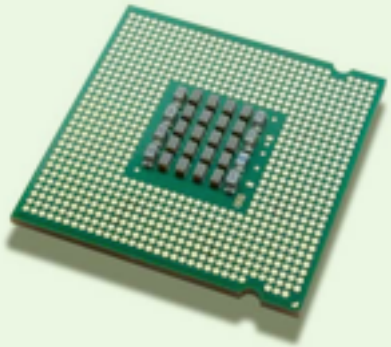


CPU leakage

Memory

Static power consumption

Hard drives



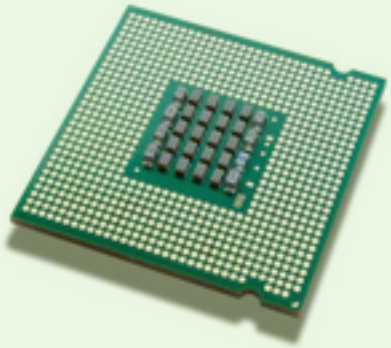
CPU leakage

Memory

Power supply losses

Static power consumption

Hard drives



CPU leakage

Memory

Power supply losses

etc.

How can DVFS save energy?

```
400c30:       48 8b 11              mov    (%rcx),%rdx
400c33:       48 8b 42 48           mov    0x48(%rdx),%rax
400c37:       48 89 41 10           mov    %rax,0x10(%rcx)
400c3b:       48 89 4a 48           mov    %rcx,0x48(%rdx)
400c3f:       48 8b 51 08           mov    0x8(%rcx),%rdx
400c43:       48 8b 42 50           mov    0x50(%rdx),%rax
400c47:       48 89 41 18           mov    %rax,0x18(%rcx)
400c4b:       48 89 4a 50           mov    %rcx,0x50(%rdx)
400c4f:       48 83 c1 40           add    $0x40,%rcx
```

[SPEC CPU2000, 181.mcf, gcc 4.2, high optimisation]

How can DVFS save energy?

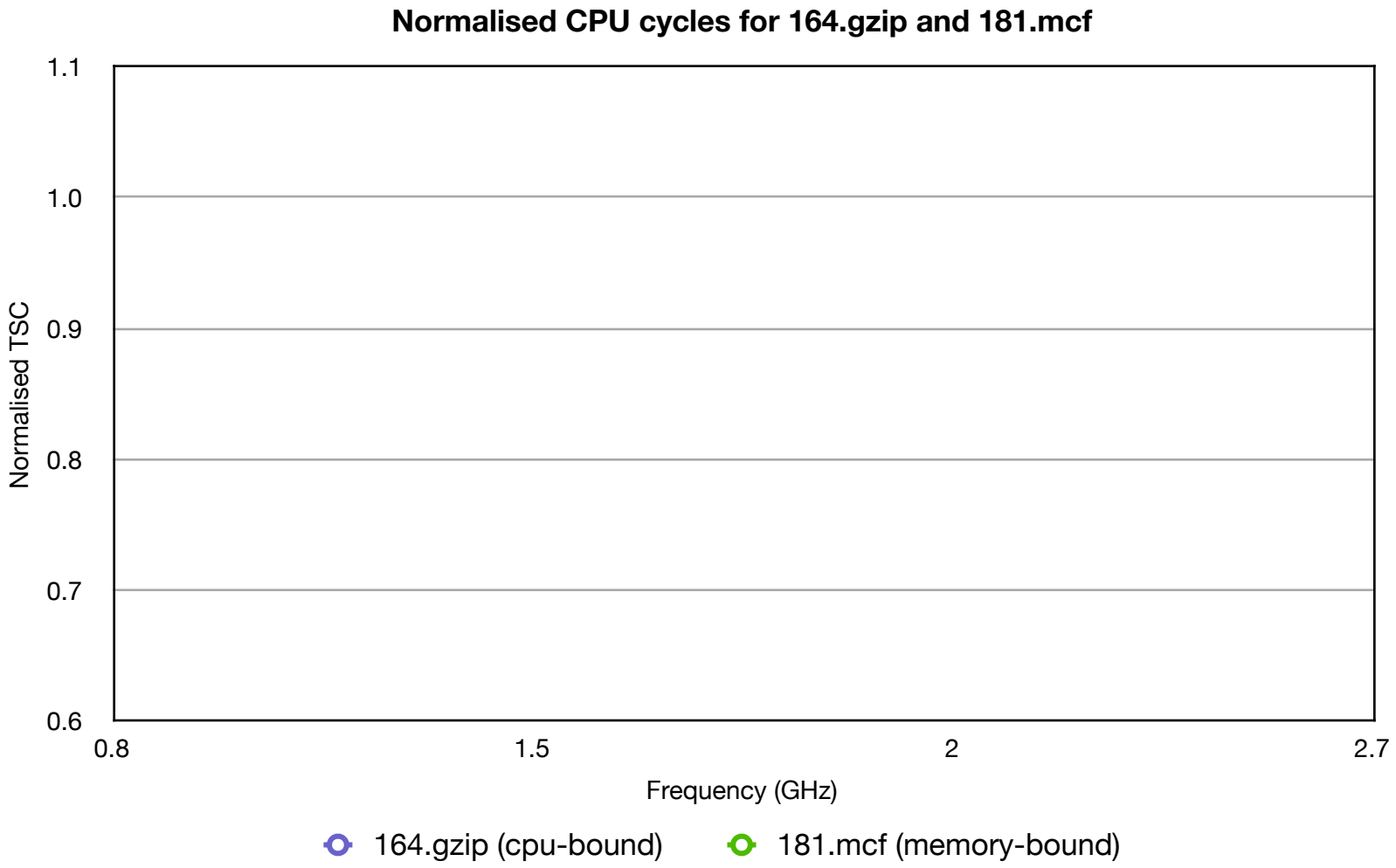
```
400c30:       48 8b 11              mov    (%rcx),%rdx
400c33:       48 8b 42 48           mov    0x48(%rdx),%rax
400c37:       48 89 41 10           mov    %rax,0x10(%rcx)
400c3b:       48 89 4a 48           mov    %rcx,0x48(%rdx)
400c3f:       48 8b 51 08           mov    0x8(%rcx),%rdx
400c43:       48 8b 42 50           mov    0x50(%rdx),%rax
400c47:       48 89 41 18           mov    %rax,0x18(%rcx)
400c4b:       48 89 4a 50           mov    %rcx,0x50(%rdx)
400c4f:       48 83 c1 40           add    $0x40,%rcx
```

[SPEC CPU2000, 181.mcf, gcc 4.2, high optimisation]

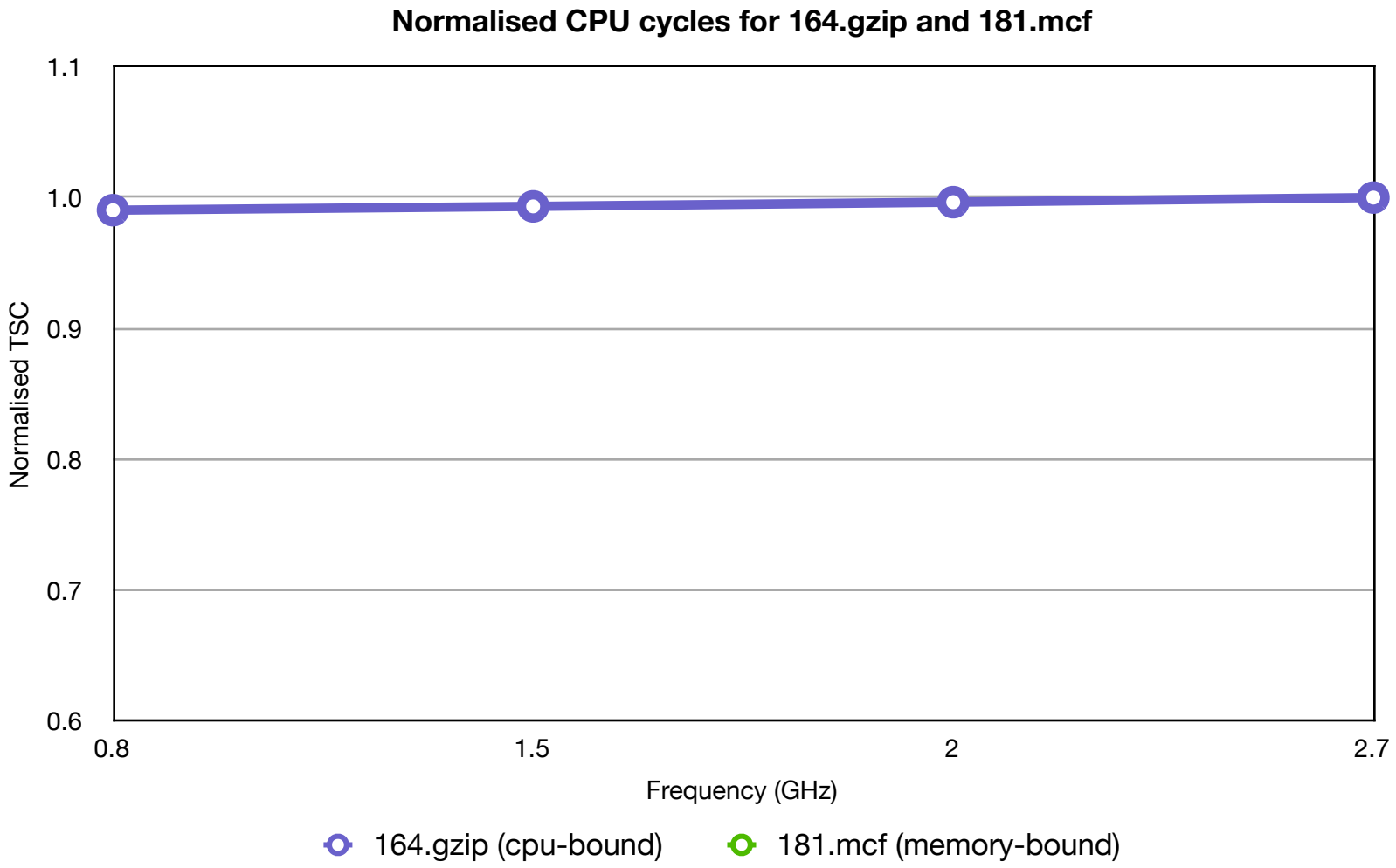
... such workloads can be ***memory-bound***

How can DVFS save energy?

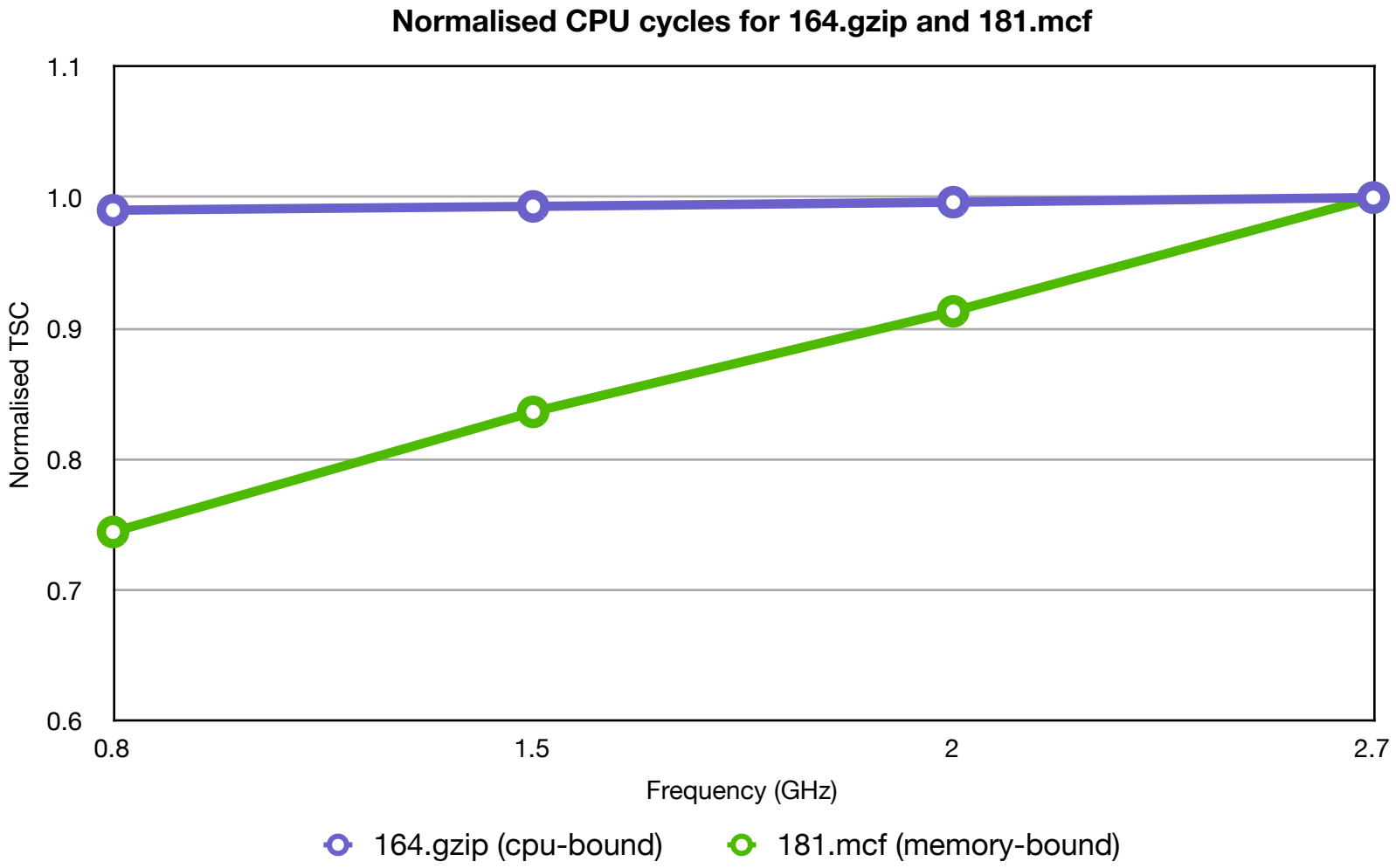
How can DVFS save energy?



How can DVFS save energy?

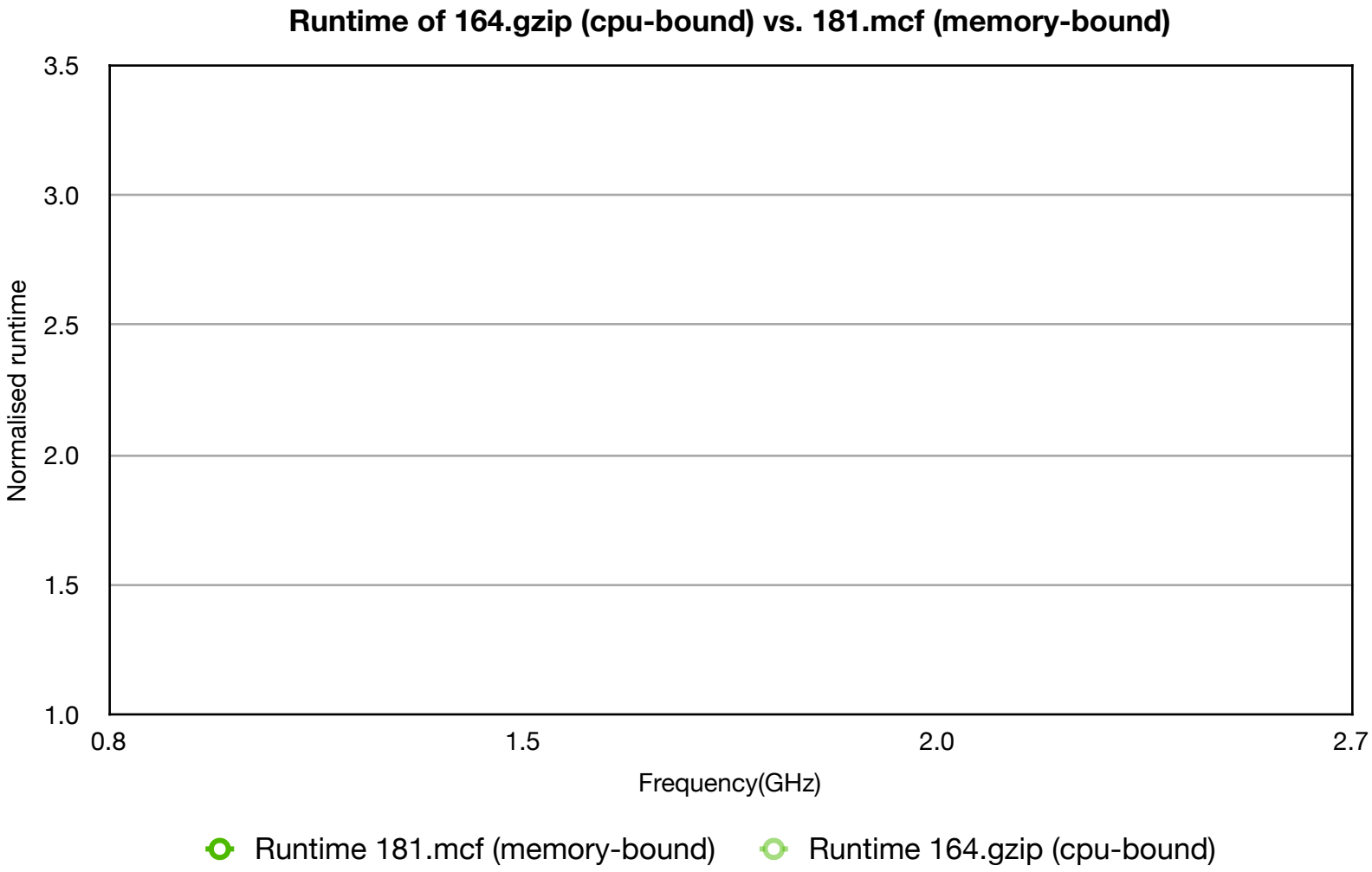


How can DVFS save energy?

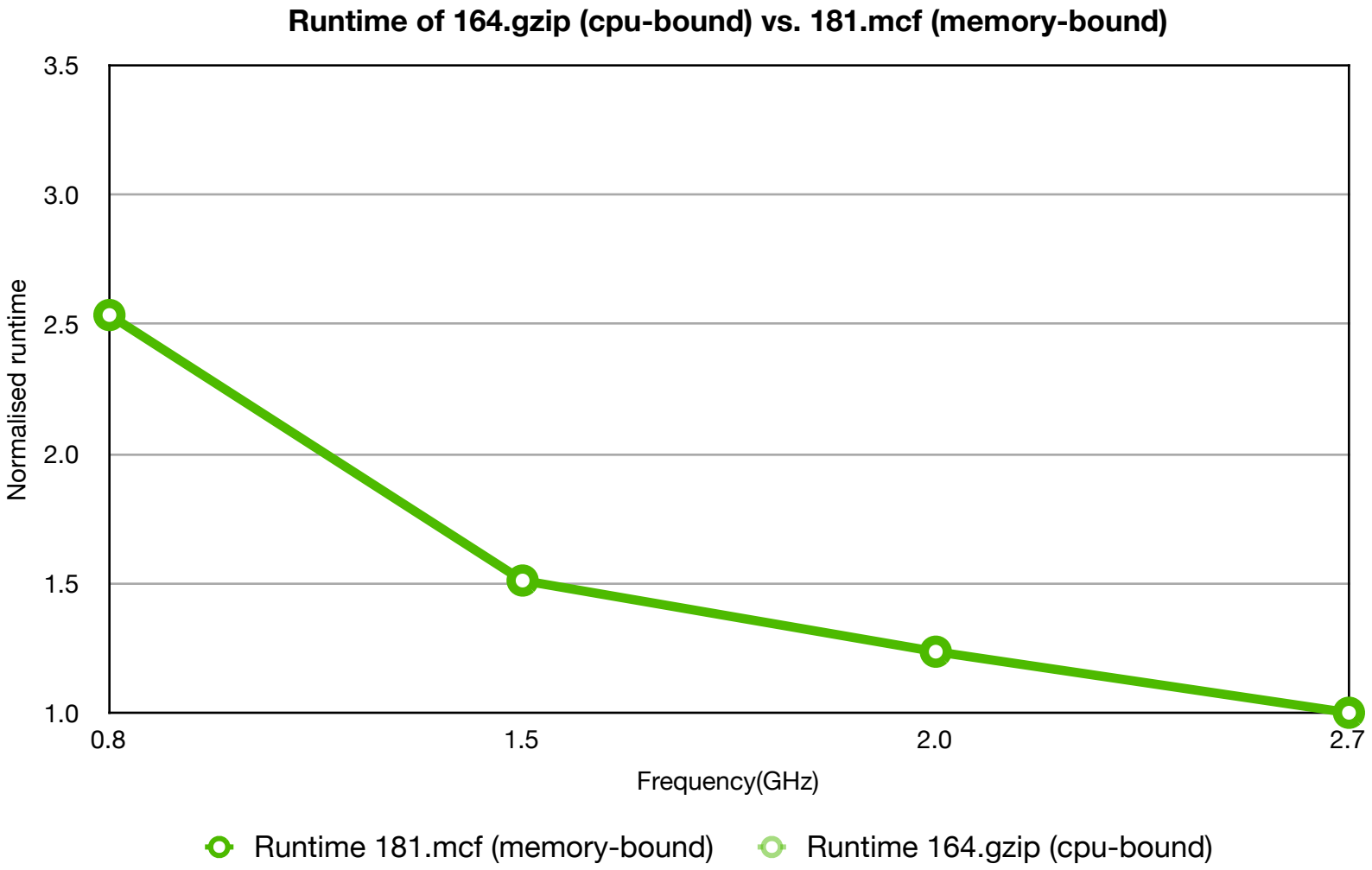


How can DVFS save energy?

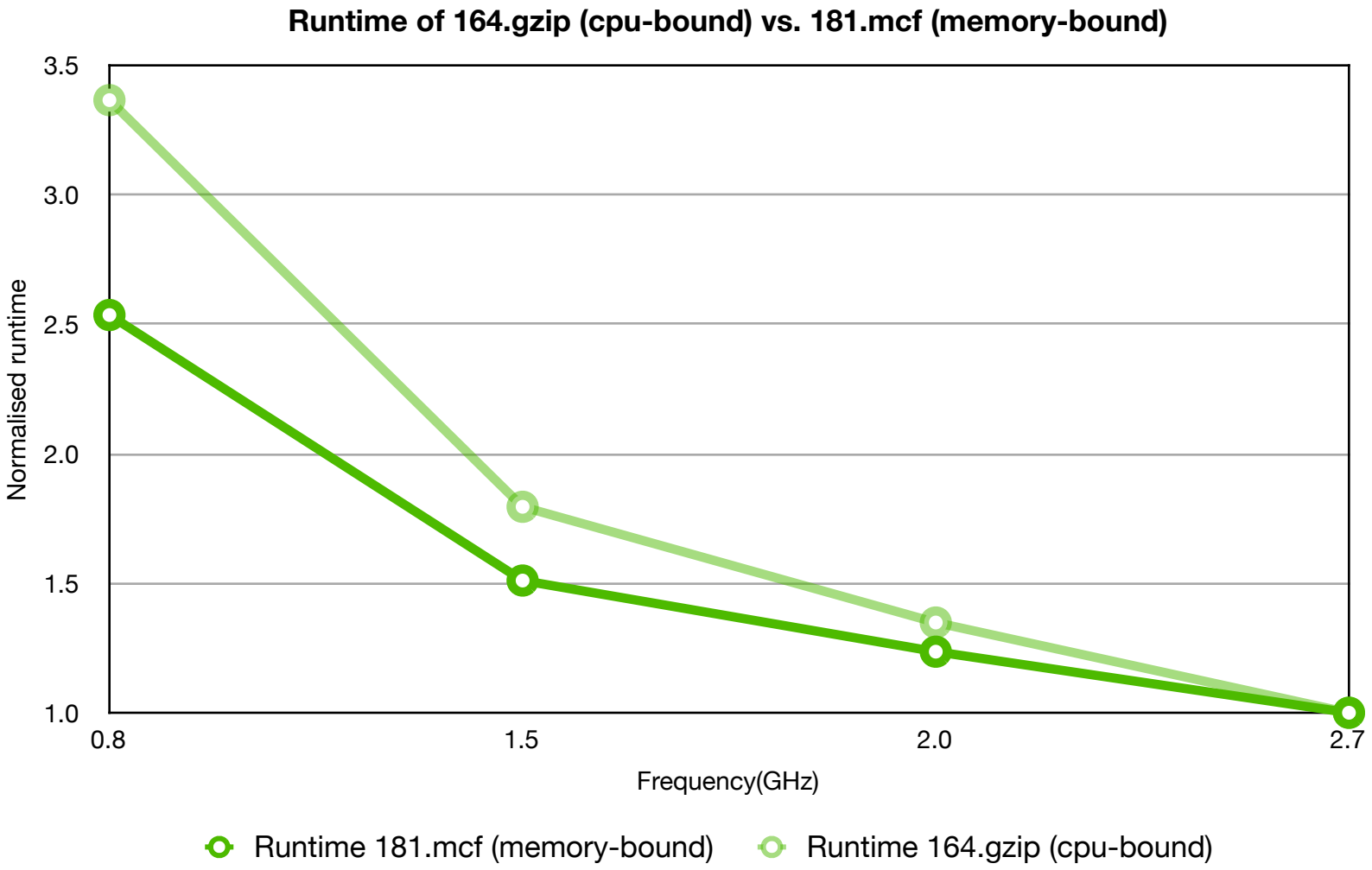
How can DVFS save energy?



How can DVFS save energy?



How can DVFS save energy?



Our analysis

NICTA

3 generations of ***AMD Opteron*** CPUs
in server-class systems

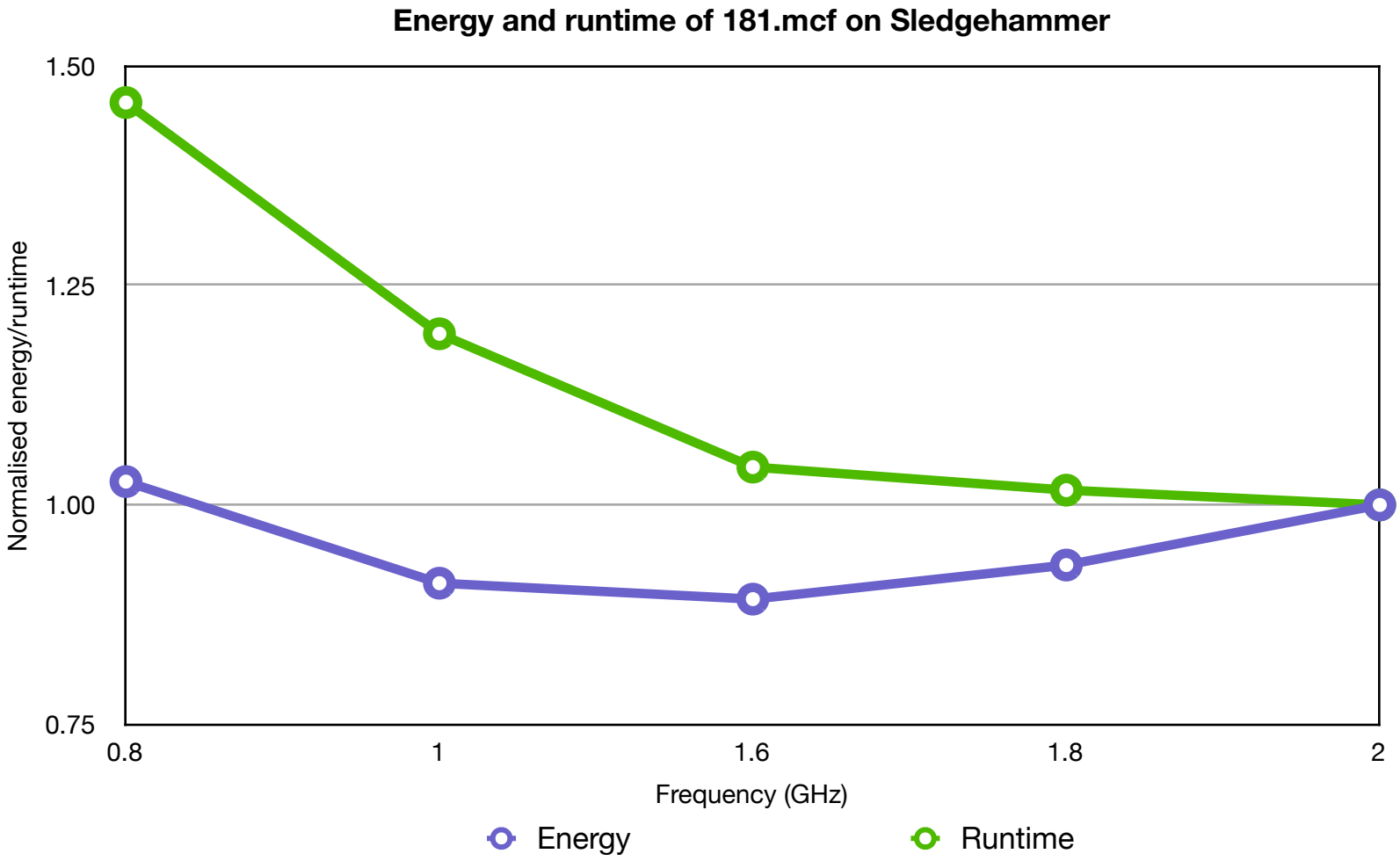
Our analysis

Die codename	Sledgehammer		
Year	2003		
Core count	1		
Frequency range	0.8 - 2.0GHz		
Voltage range	0.9 - 1.5		
Process	130nm		
TDP	89W		
Die area	193mm^2		
Transistor count	106M		

Sledgehammer

NICTA

Sledgehammer

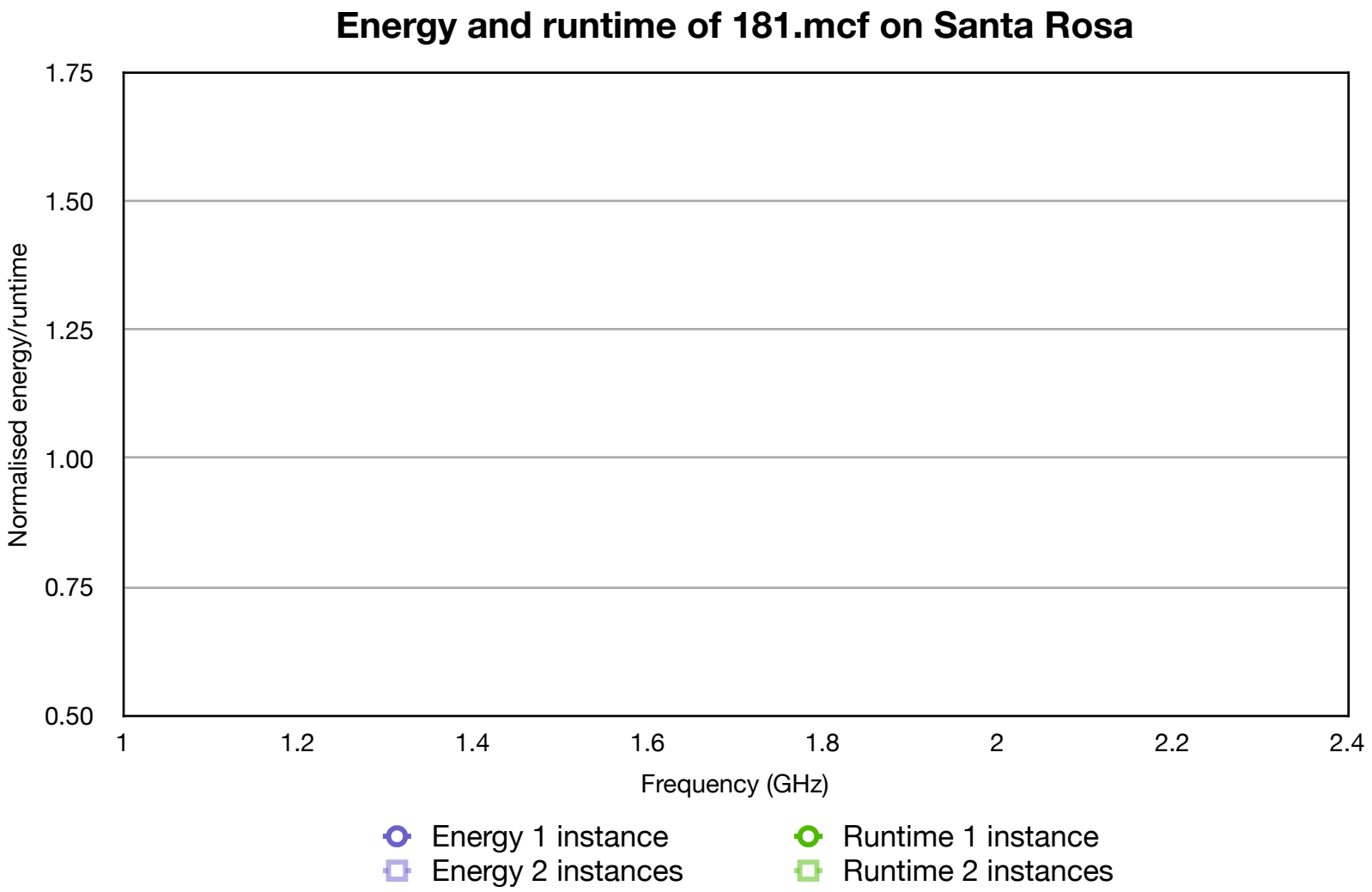


Our analysis

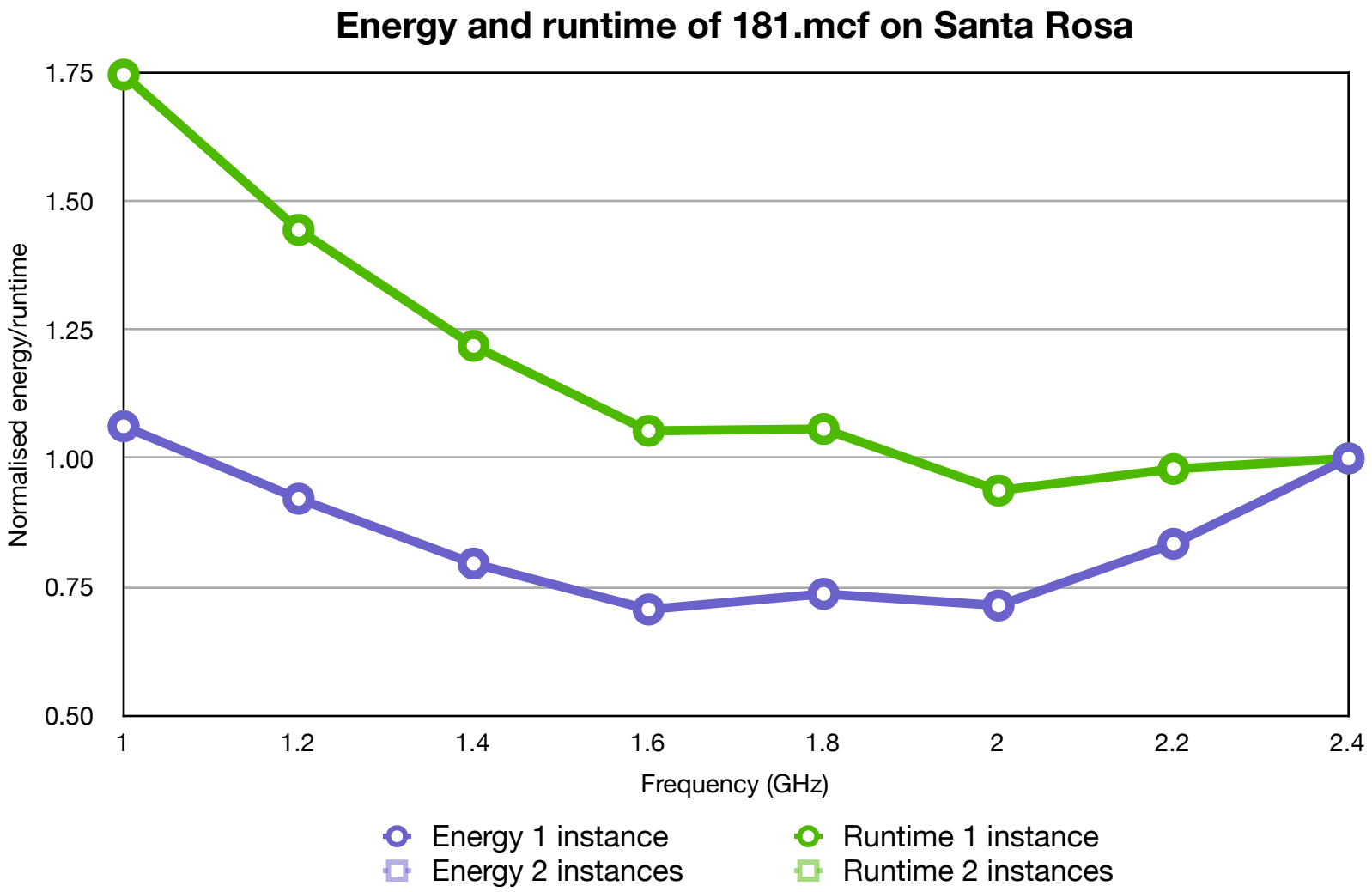
Die codename	Sledgehammer	**Santa Rosa**
Year	2003	**2006**
Core count	1	**2**
Frequency range	0.8 - 2.0GHz	**1.0 - 2.4GHz**
Voltage range	0.9 - 1.5	**0.9 - 1.35V**
Process	130nm	**90nm**
TDP	89W	**95W**
Die area	193mm^2	**230mm^2**
Transistor count	106M	**243M**

Santa Rosa

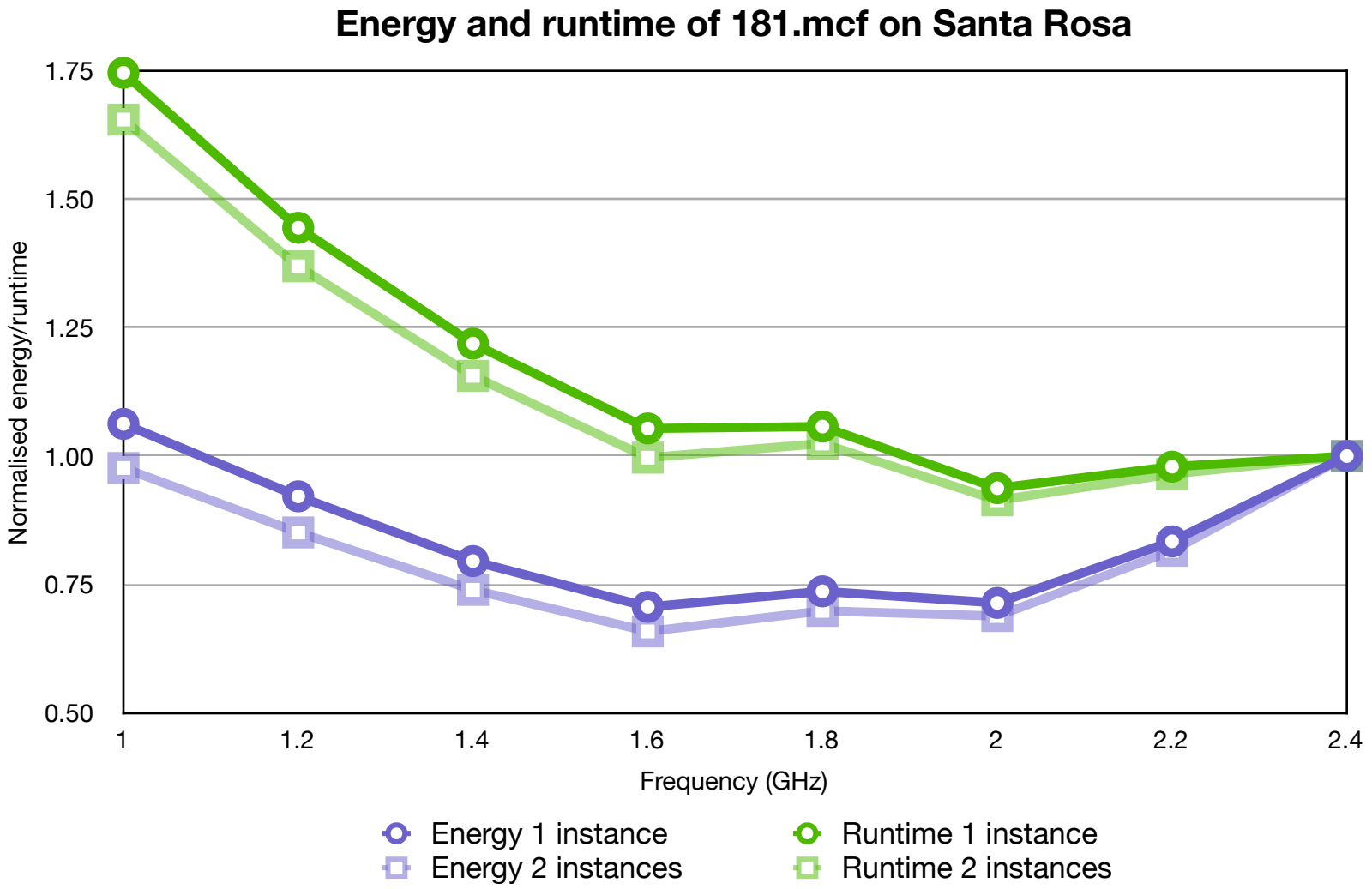
Santa Rosa



Santa Rosa



Santa Rosa



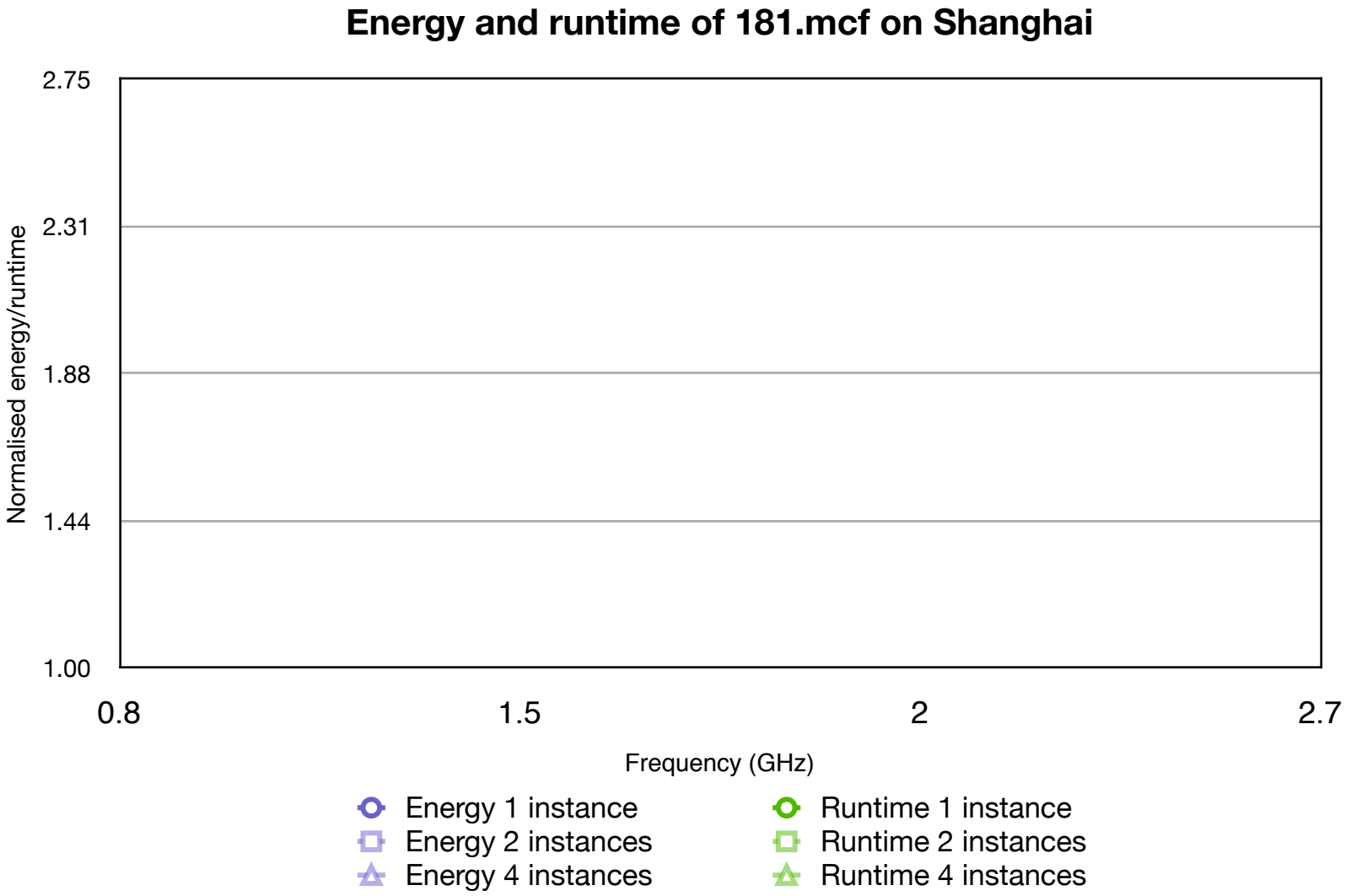
Our analysis

Die codename	Sledgehammer	Santa Rosa	**Shanghai**
Year	2003	2006	**2009**
Core count	1	2	**4**
Frequency range	0.8 - 2.0GHz	1.0 - 2.4GHz	**0.8 - 2.7GHz**
Voltage range	0.9 - 1.5	0.9 - 1.35V	**1.0 - 1.35V**
Process	130nm	90nm	**45nm**
TDP	89W	95W	**75W**
Die area	193mm^2	230mm^2	**285mm^2**
Transistor count	106M	243M	**463M**

Shanghai

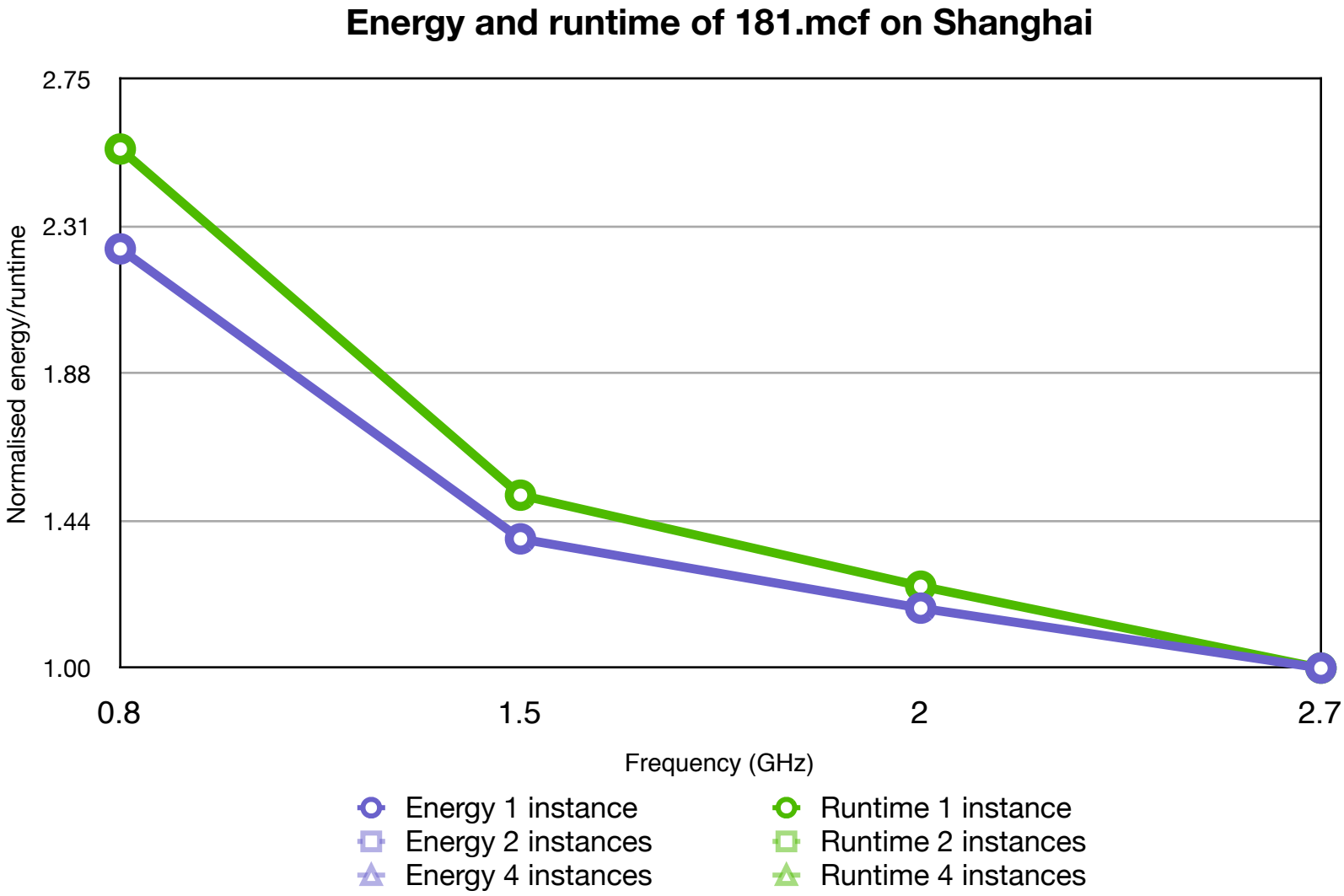
Shanghai

NICTA



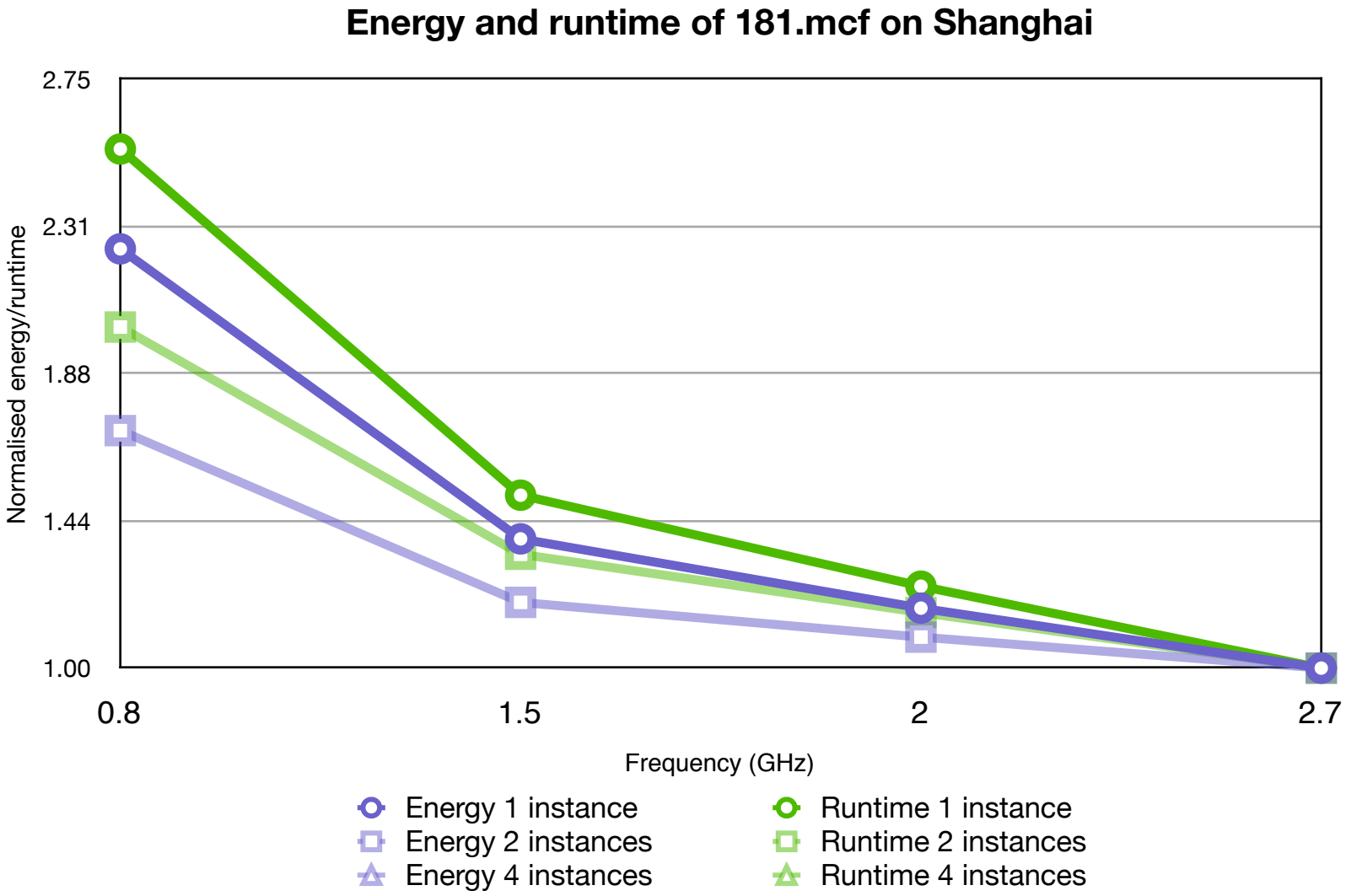
Shanghai

NICTA



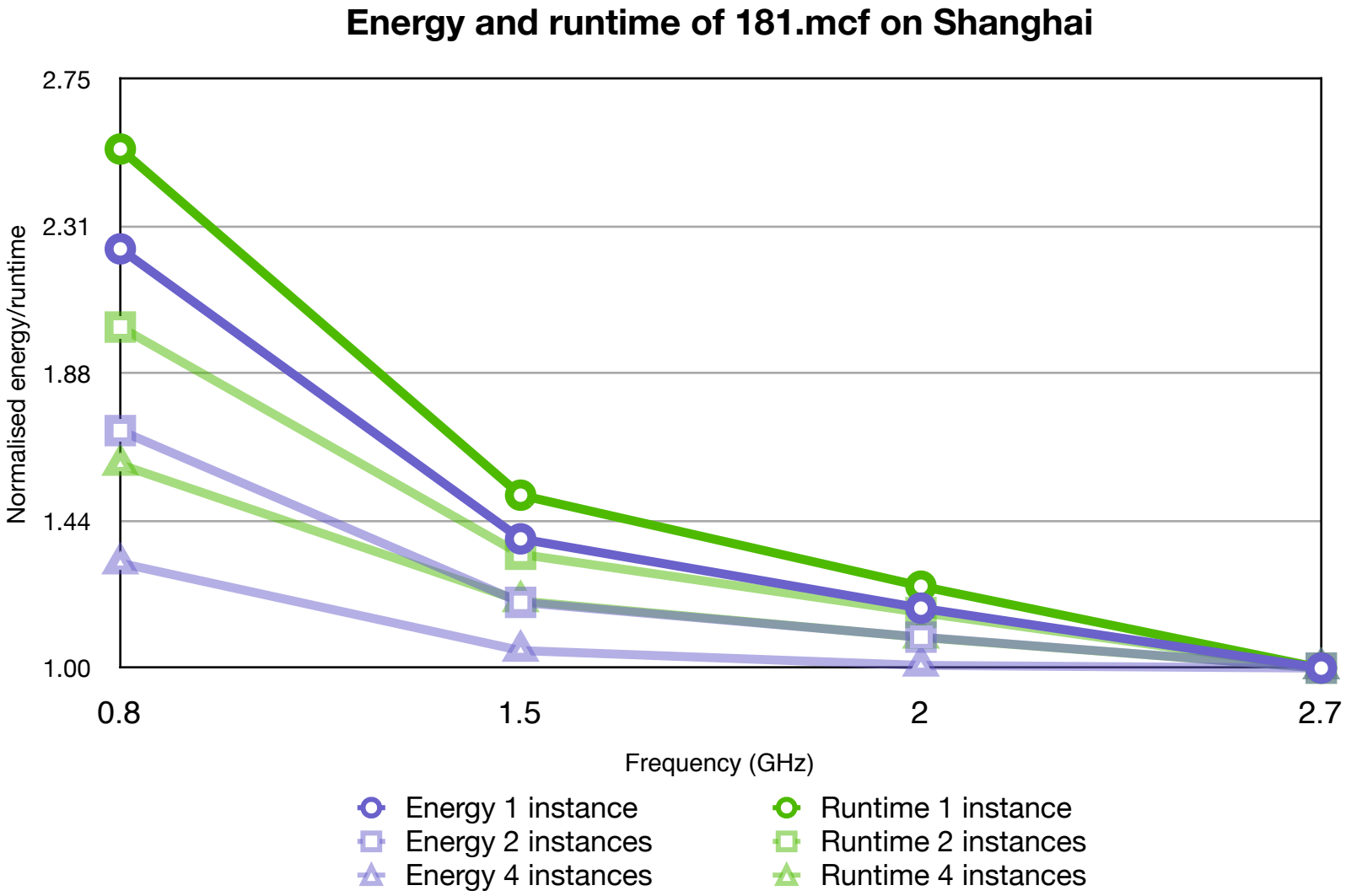
Shanghai

NICTA



Shanghai

NICTA



Results

NICTA

DVFS on Shanghai is ***ineffective for saving energy!***

Results

NICTA

DVFS on Shanghai is ***ineffective for saving energy!***

... but ***why?***

Static power

DVFS can only change ***dynamic*** power consumption and static power is increasing!

What are the trends?

What are the trends?

- scaling of transistor technology;

What are the trends?

- scaling of transistor technology;
- increasing memory performance;

What are the trends?

- scaling of transistor technology;
- increasing memory performance;
- improved idle/sleep modes; and

What are the trends?

- scaling of transistor technology;
- increasing memory performance;
- improved idle/sleep modes; and
- multi-core processors.

Scaling of transistor technology

NICTA

In ~7 years:

Scaling of transistor technology

NICTA

In ~7 years:

- 130 nm to 45 nm

Scaling of transistor technology

NICTA

In ~7 years:

- 130 nm to 45 nm
- 106 M to 463 M transistors

Scaling of transistor technology

NICTA

In ~7 years:

- 130 nm to 45 nm
- 106 M to 463 M transistors
- 193 mm^2 to 285 mm^2

Scaling of transistor technology

NICTA

In ~7 years:

- 130 nm to 45 nm
- 106 M to 463 M transistors
- 193 mm^2 to 285 mm^2
- 1 MiB to 8 MiB total SRAM cache

Scaling of transistor technology

NICTA

In ~7 years:

- 130 nm to 45 nm
- 106 M to 463 M transistors
- 193 mm^2 to 285 mm^2
- 1 MiB to 8 MiB total SRAM cache
- Single to quad-core (more later...)

Scaling of transistor technology

NICTA

In ~7 years:

- 130 nm to 45 nm
- 106 M to 463 M transistors
- 193 mm^2 to 285 mm^2
- 1 MiB to 8 MiB total SRAM cache
- Single to quad-core (more later...)

Static power is increasing significantly, dynamic power is decreasing

Increasing memory performance

NICTA

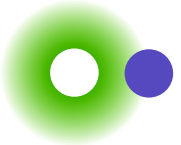
In ~7 years:

Increasing memory performance

NICTA

In ~7 years:

- Much larger CPU SRAM caches (fewer cache-misses)



NICTA

Increasing memory performance

In ~7 years:

- Much larger CPU SRAM caches (fewer cache-misses)
- DRAM throughput: 3.2GB/s to 5.33GB/s

Increasing memory performance

In ~7 years:

- Much larger CPU SRAM caches (fewer cache-misses)
- DRAM throughput: 3.2GB/s to 5.33GB/s
- Larger DRAM prefetch distance: 2 to 3 cache-lines

Increasing memory performance

NICTA

In ~7 years:

- Much larger CPU SRAM caches (fewer cache-misses)
- DRAM throughput: 3.2GB/s to 5.33GB/s
- Larger DRAM prefetch distance: 2 to 3 cache-lines
- Dual to triple channel DDR memory-controllers

Increasing memory performance

NICTA

In ~7 years:

- Much larger CPU SRAM caches (fewer cache-misses)
- DRAM throughput: 3.2GB/s to 5.33GB/s
- Larger DRAM prefetch distance: 2 to 3 cache-lines
- Dual to triple channel DDR memory-controllers

Fewer pipeline stalls for memory references means code is less memory-bound

Improved idle/sleep modes

NICTA

In ~7 years:

Improved idle/sleep modes

NICTA

In ~7 years:

- From a single ‘halt’ mode (ACPI C1)

Improved idle/sleep modes

NICTA

In ~7 years:

- From a single ‘halt’ mode (ACPI C1)
- To multiple stepped low-power sleep modes (C1 - 4)

Improved idle/sleep modes

NICTA

In ~7 years:

- From a single ‘halt’ mode (ACPI C1)
- To multiple stepped low-power sleep modes (C1 - 4)

Push towards race-and-halt

Multi-core processors

Today:

Multi-core processors

NICTA

Today:

- Single off-chip voltage regulator module (VRM)

Multi-core processors

NICTA

Today:

- Single off-chip voltage regulator module (VRM)
- One or more on-chip PLL clock generator modules

Multi-core processors

NICTA

Today:

- Single off-chip voltage regulator module (VRM)
- One or more on-chip PLL clock generator modules

Per-core DVFS adds complexity to hardware and DVFS algorithms

How much energy can we really save?

NICTA

Up to now, we've ignored the performance loss...

We can:

How much energy can we really save?

NICTA

Up to now, we've ignored the performance loss...

We can:

- use a different benchmarking methodology;

How much energy can we really save?

NICTA

Up to now, we've ignored the performance loss...

We can:

- use a different benchmarking methodology;
- use a different metric;

How much energy can we really save?

NICTA

Up to now, we've ignored the performance loss...

We can:

- use a different benchmarking methodology;
- use a different metric;
- or both.

Padding methodology

Padding methodology

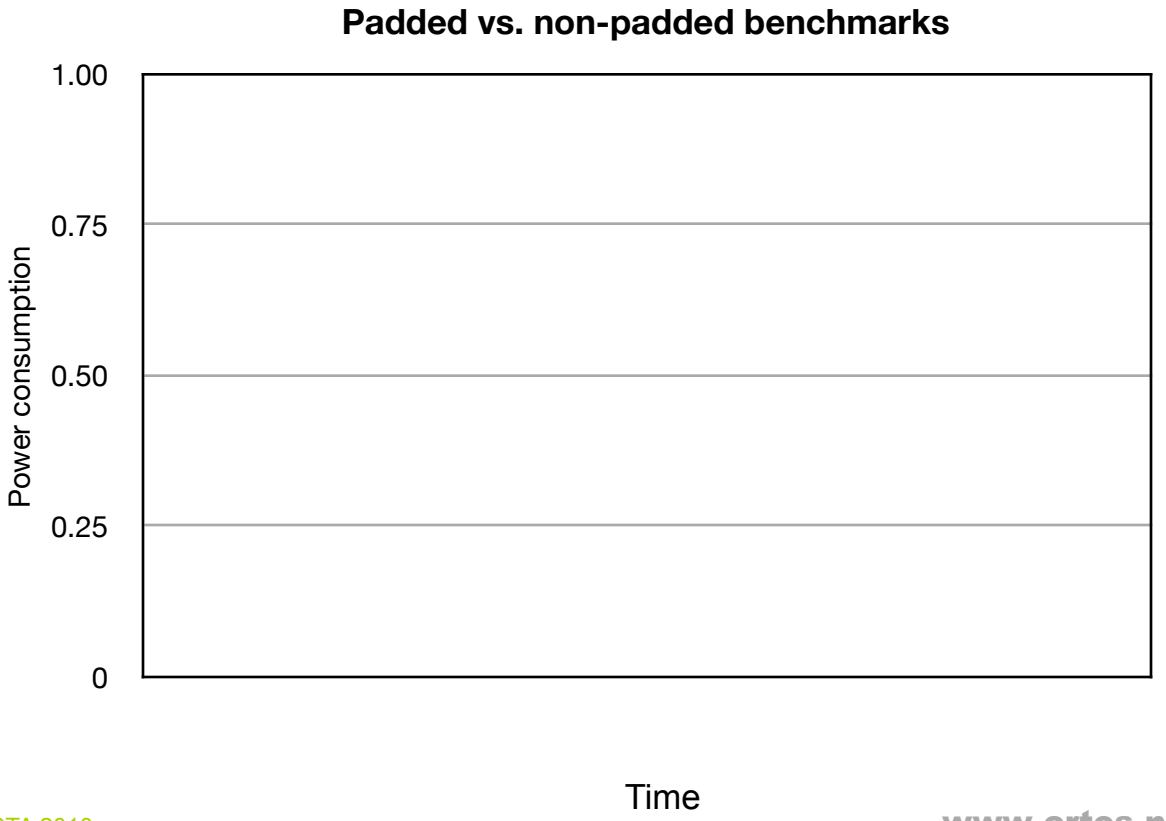
- Extend shorter executions;

Padding methodology

- Extend shorter executions;
- Add idle energy

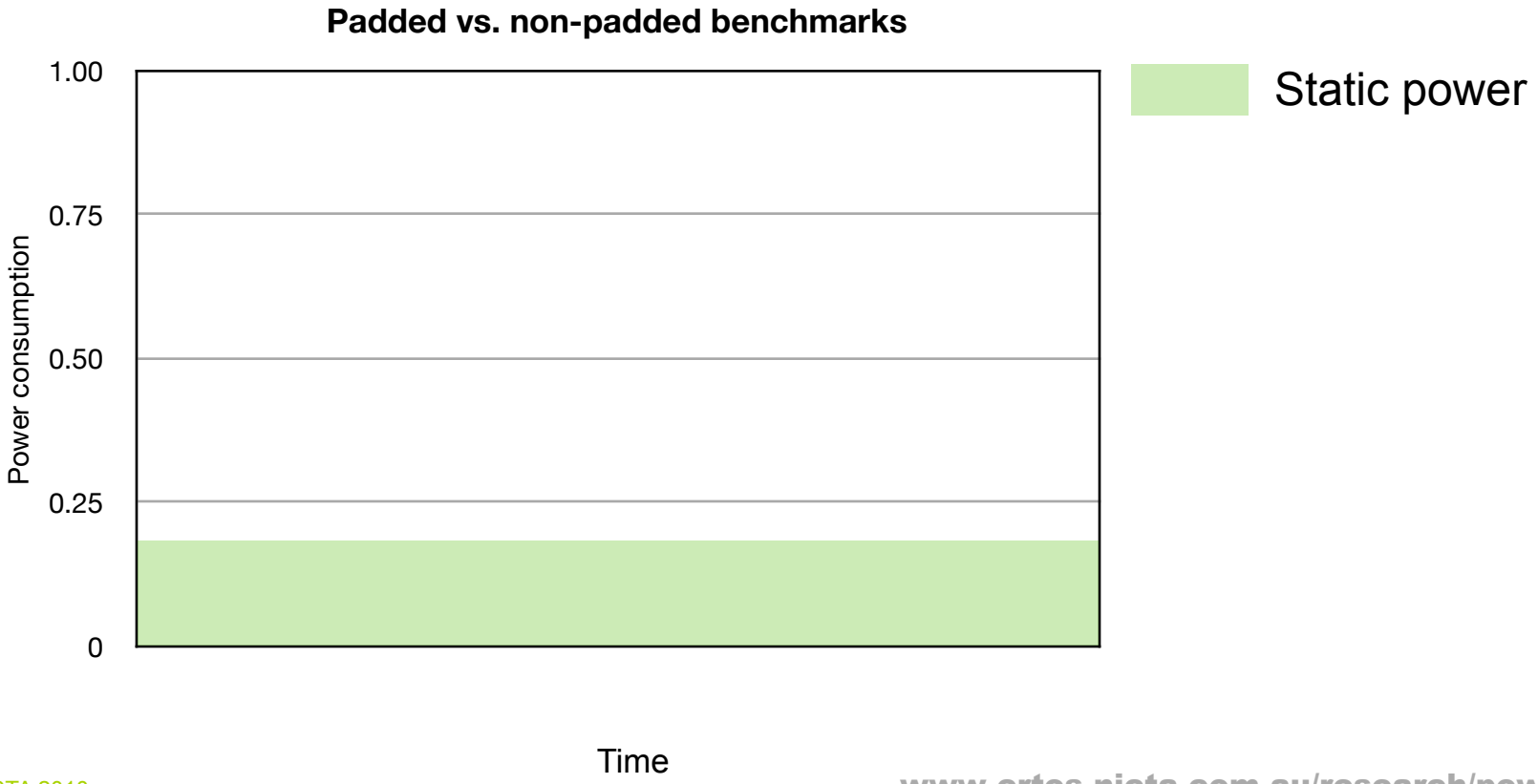
Padding methodology

- Extend shorter executions;
- Add idle energy



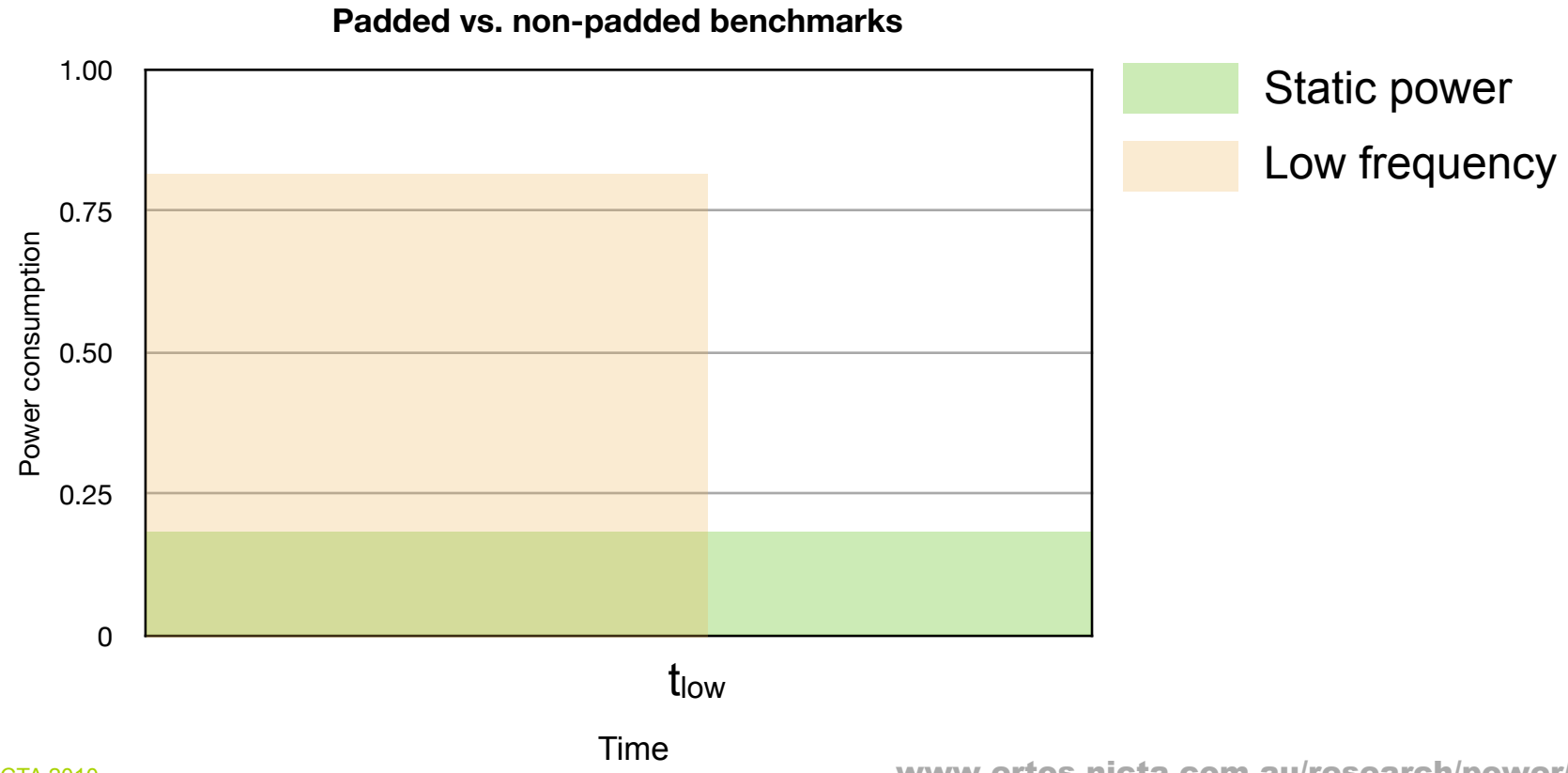
Padding methodology

- Extend shorter executions;
- Add idle energy



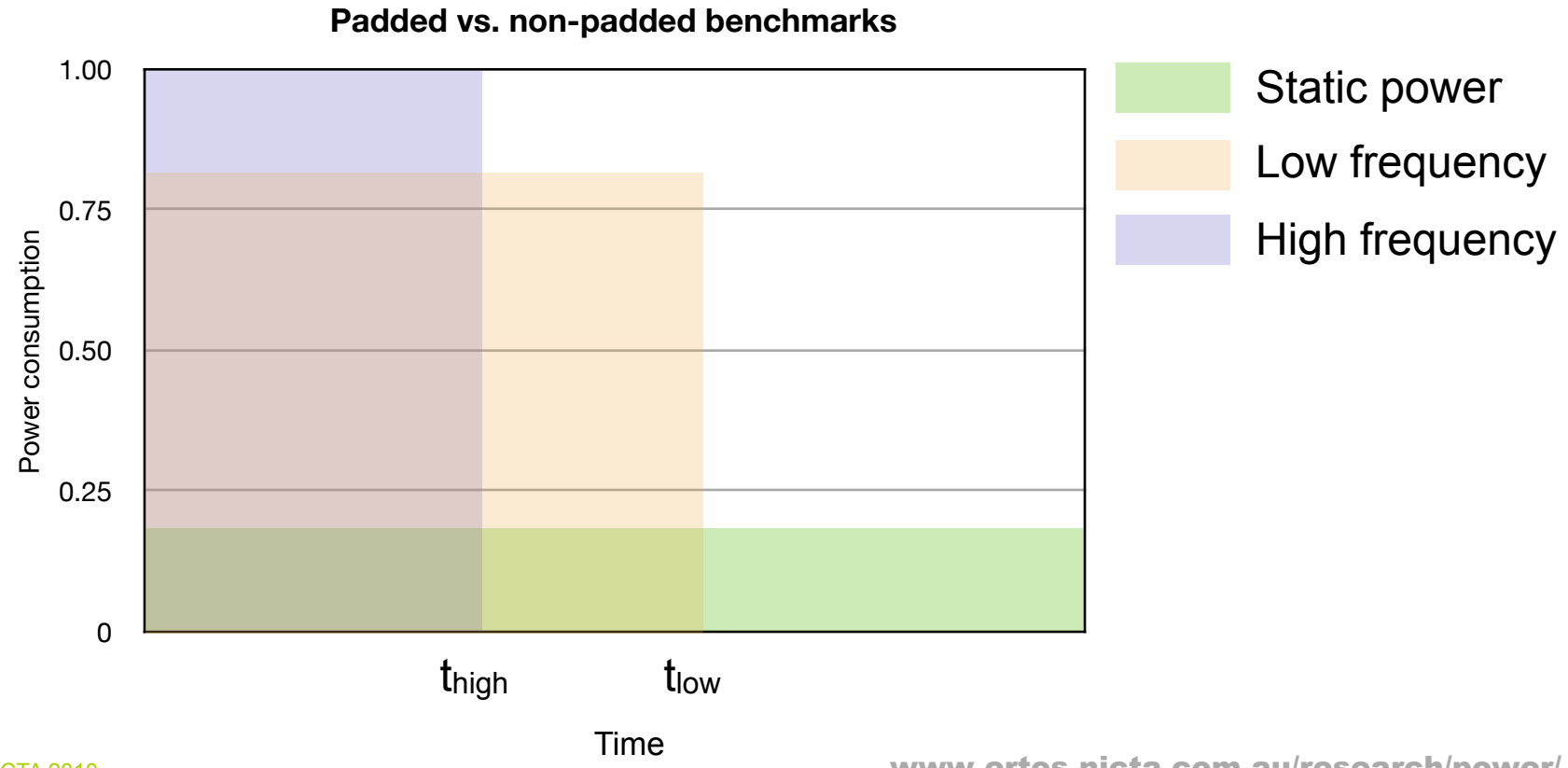
Padding methodology

- Extend shorter executions;
- Add idle energy



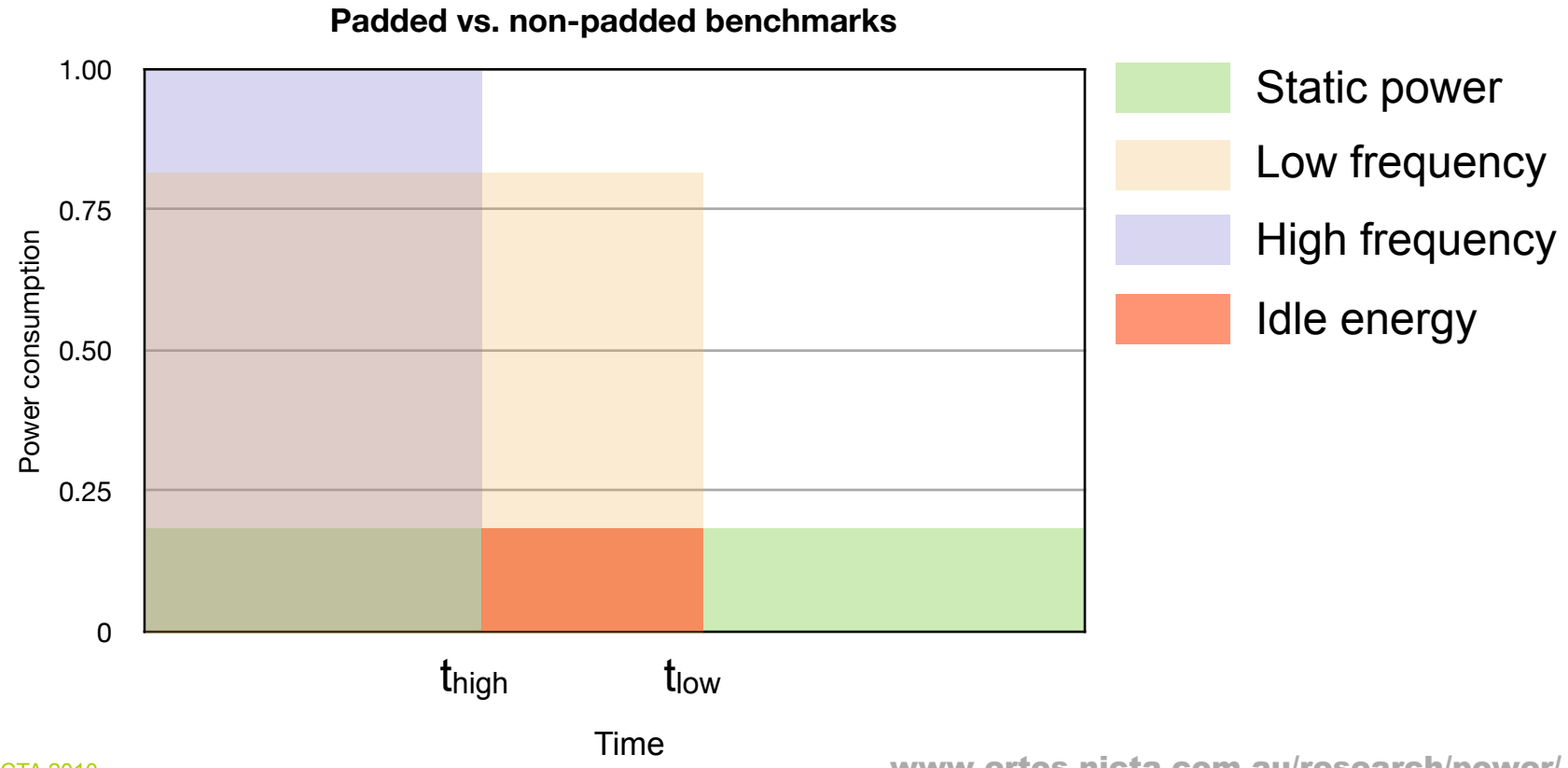
Padding methodology

- Extend shorter executions;
- Add idle energy



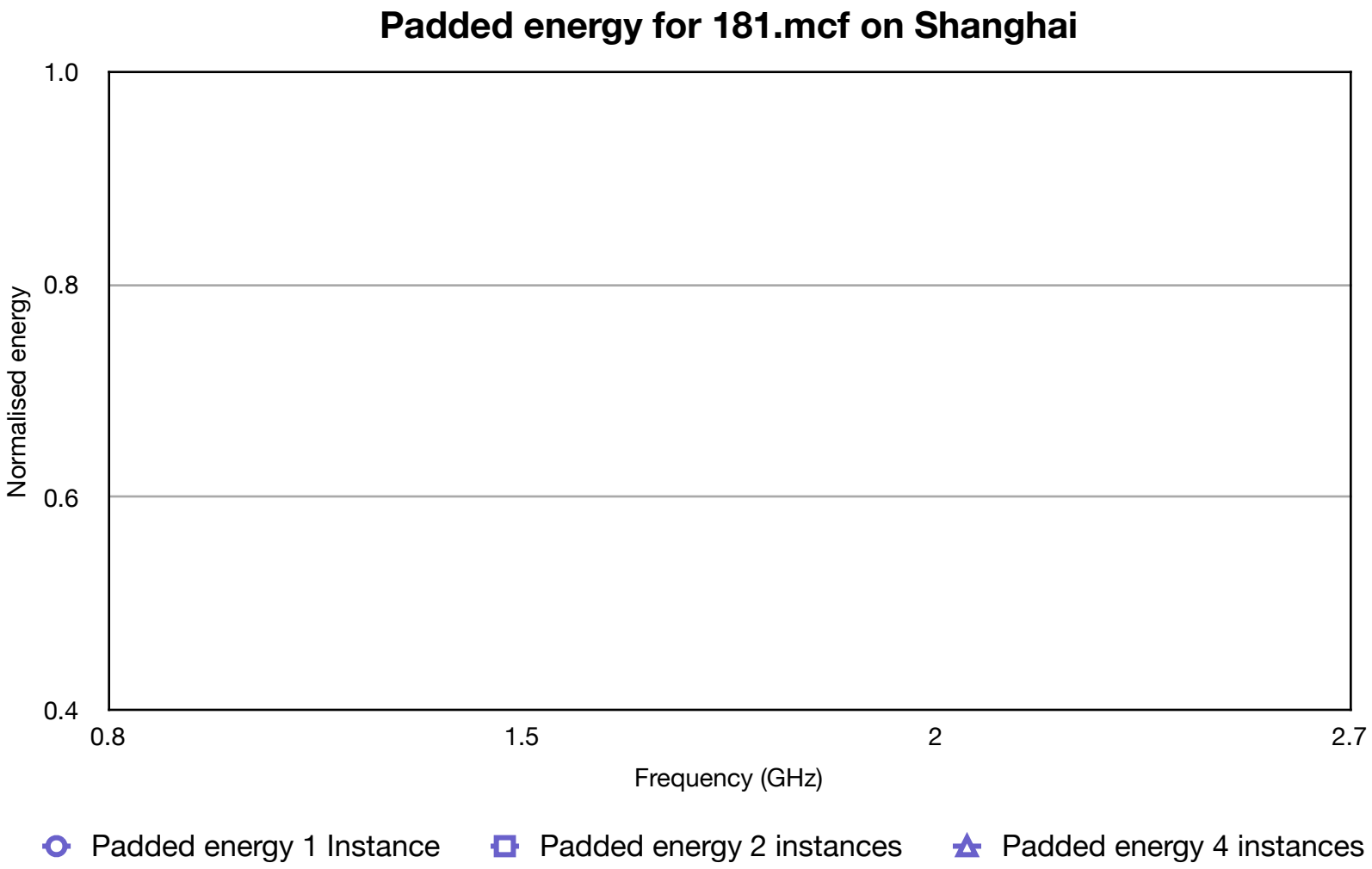
Padding methodology

- Extend shorter executions;
- Add idle energy

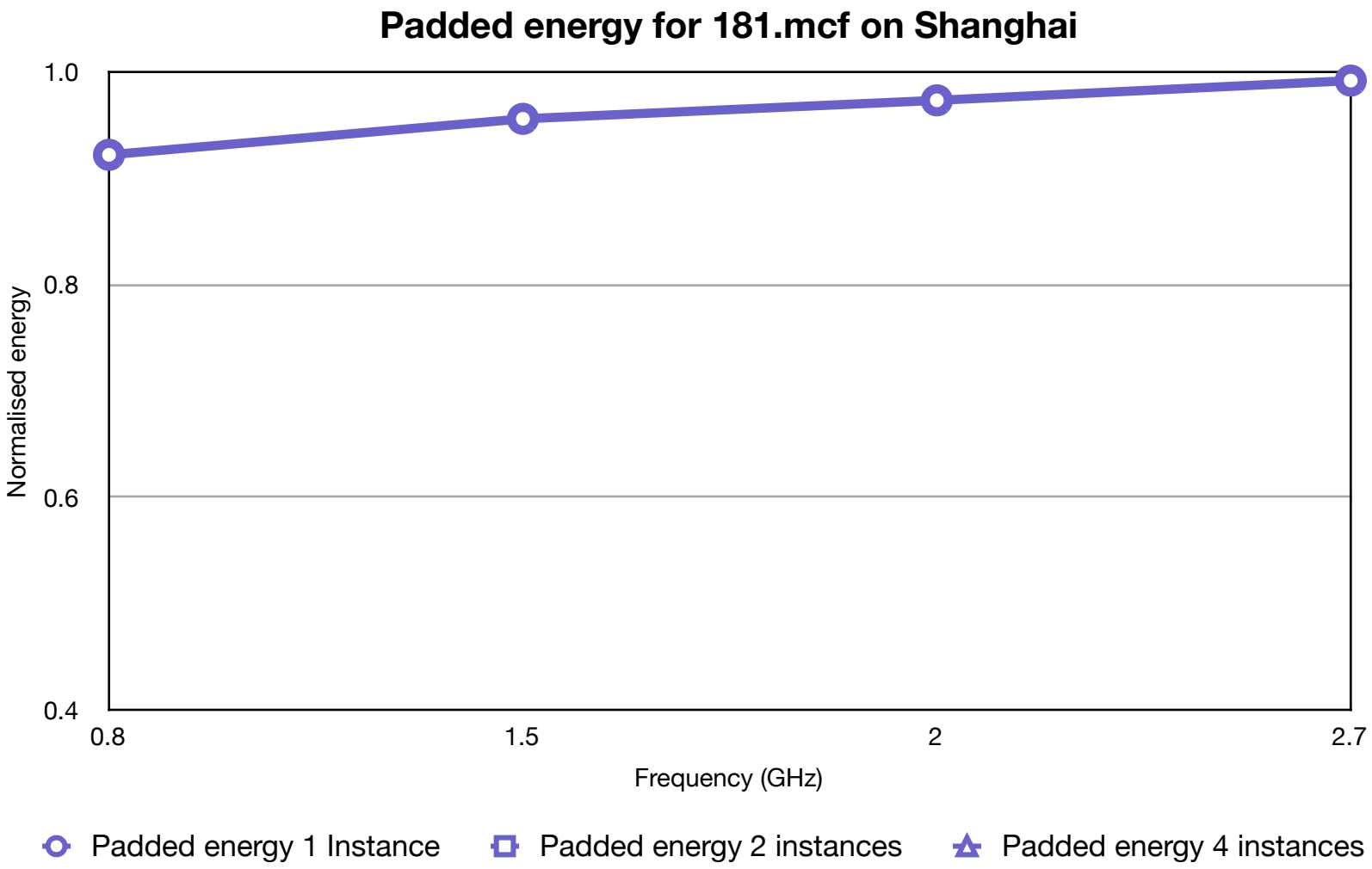


Padding methodology

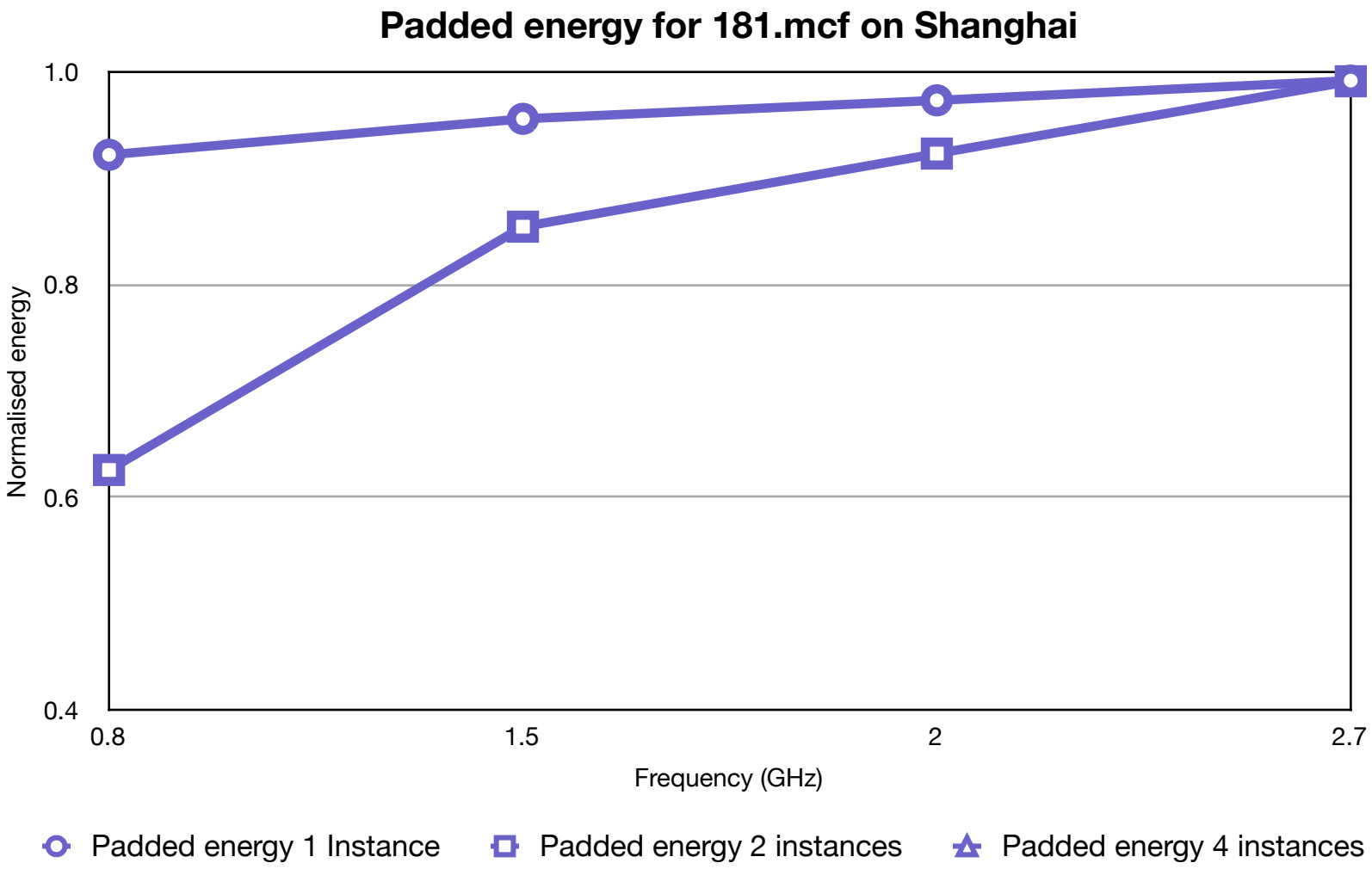
Padding methodology



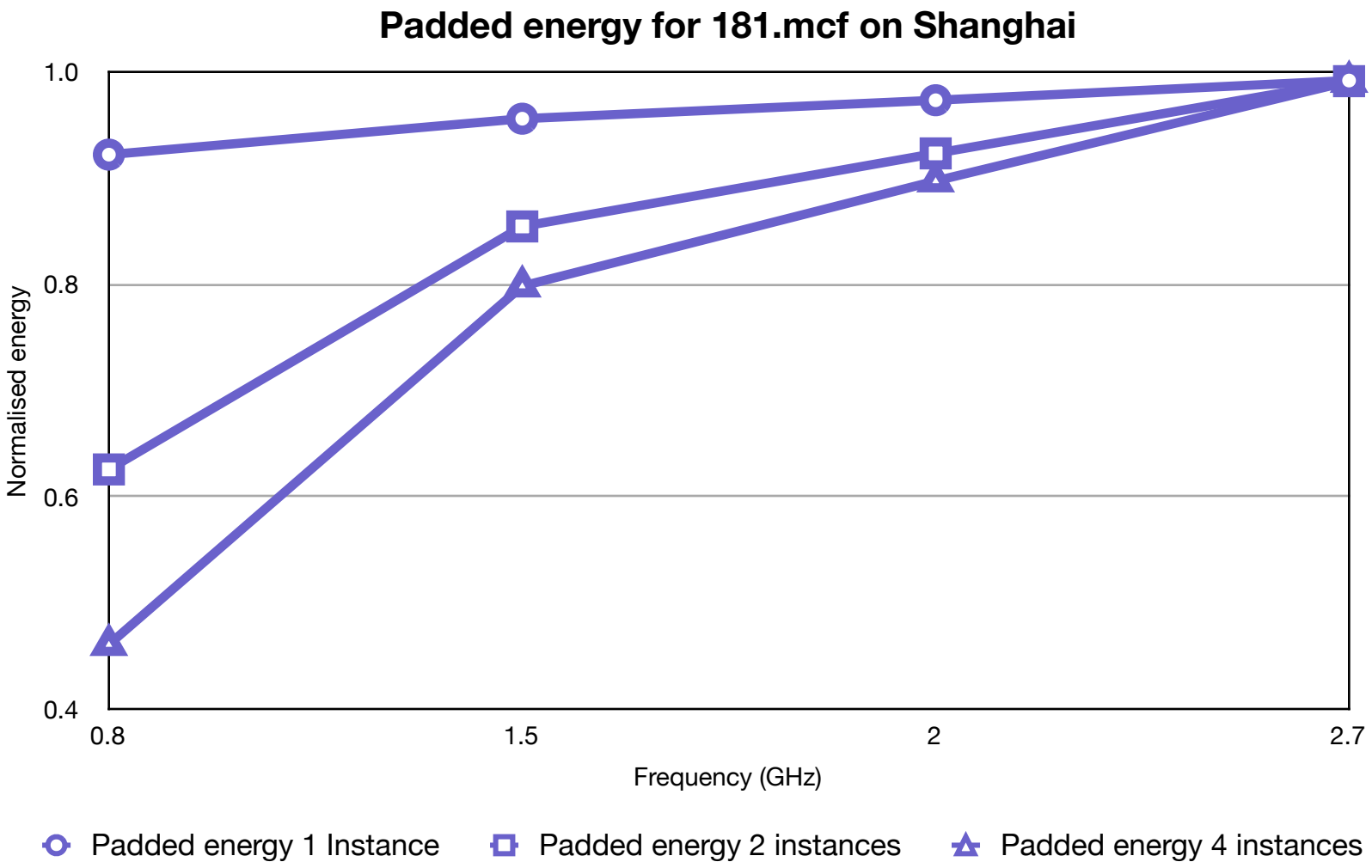
Padding methodology



Padding methodology



Padding methodology



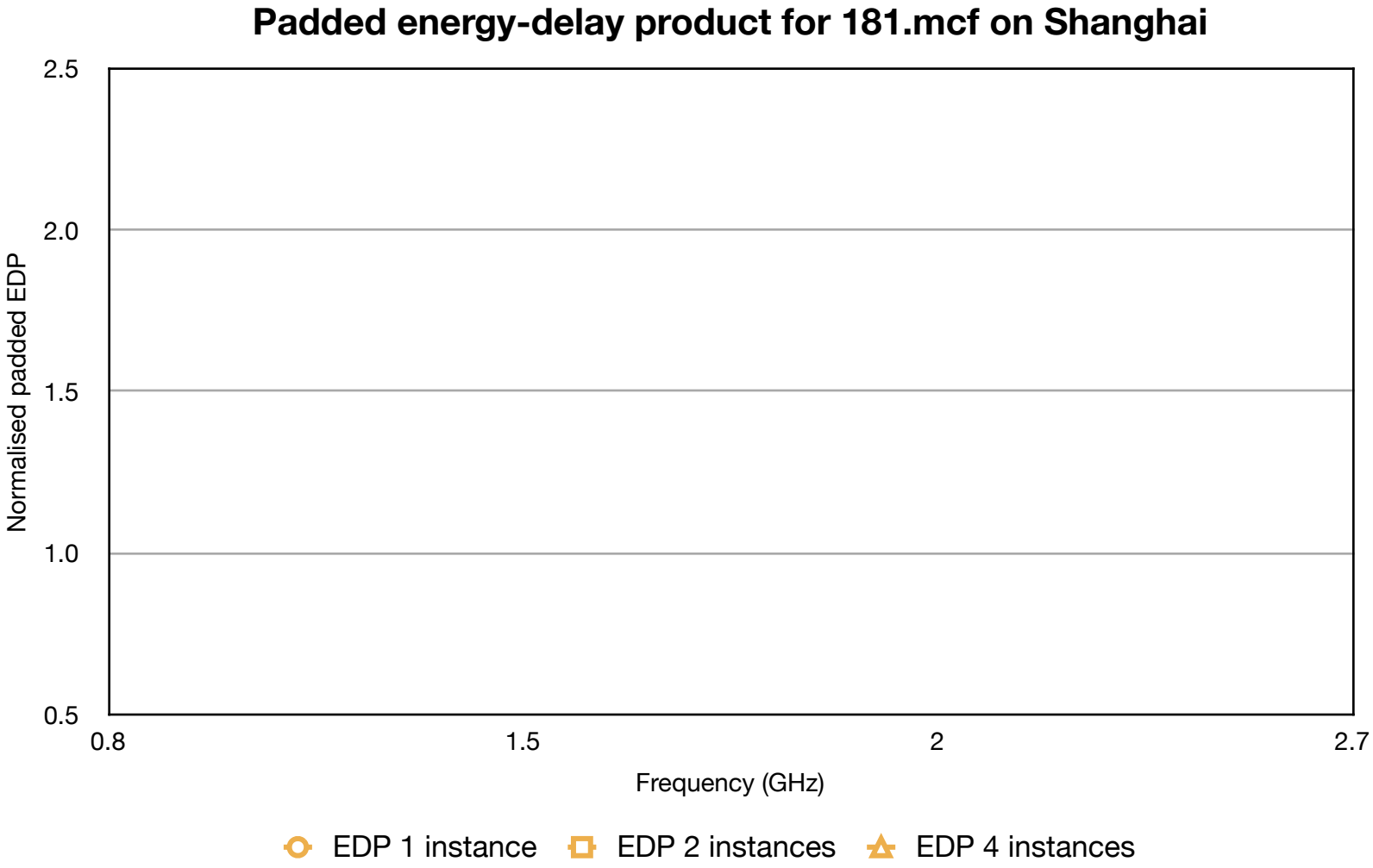
Padding methodology

NICTA

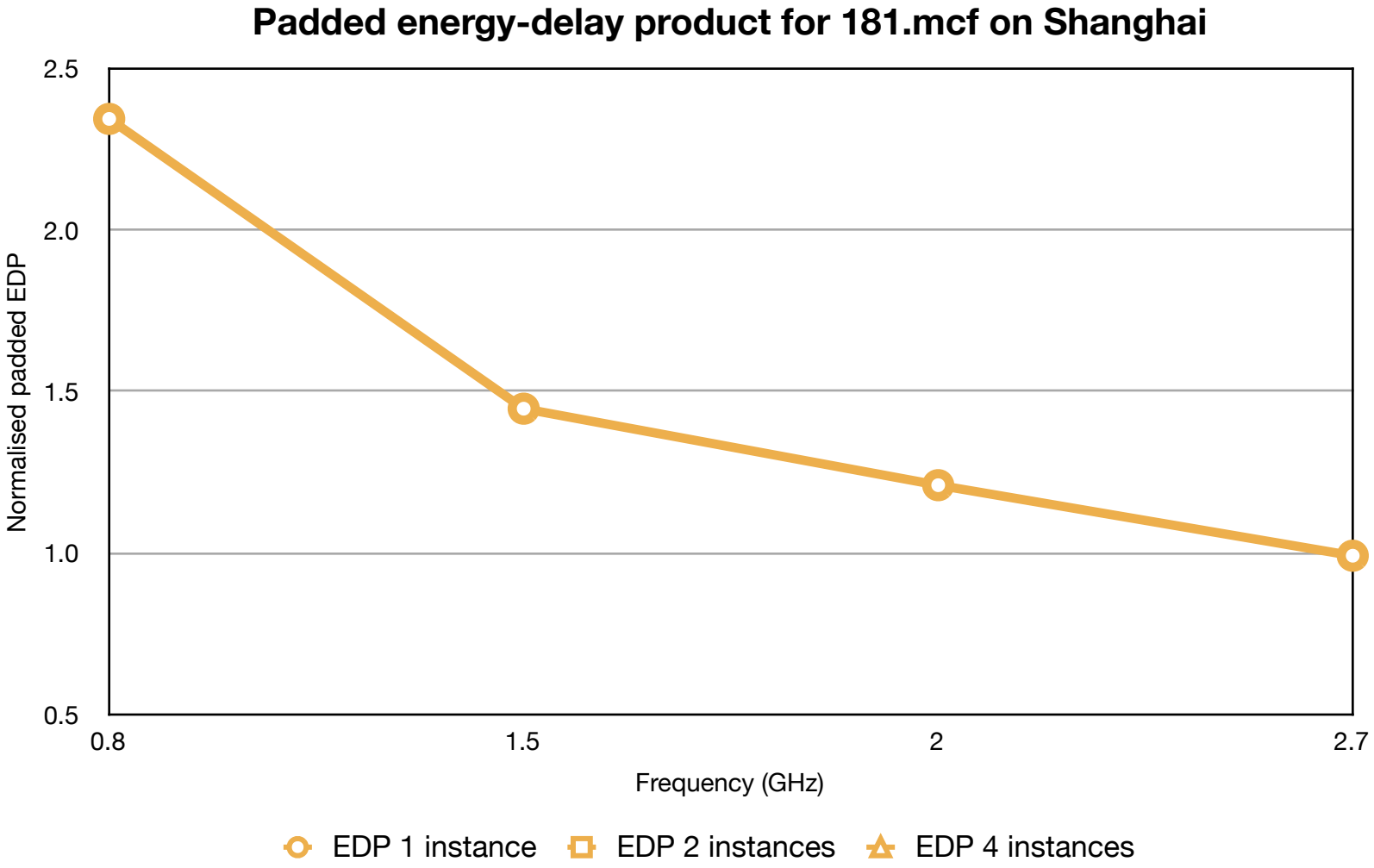
What about performance?

Energy-delay product

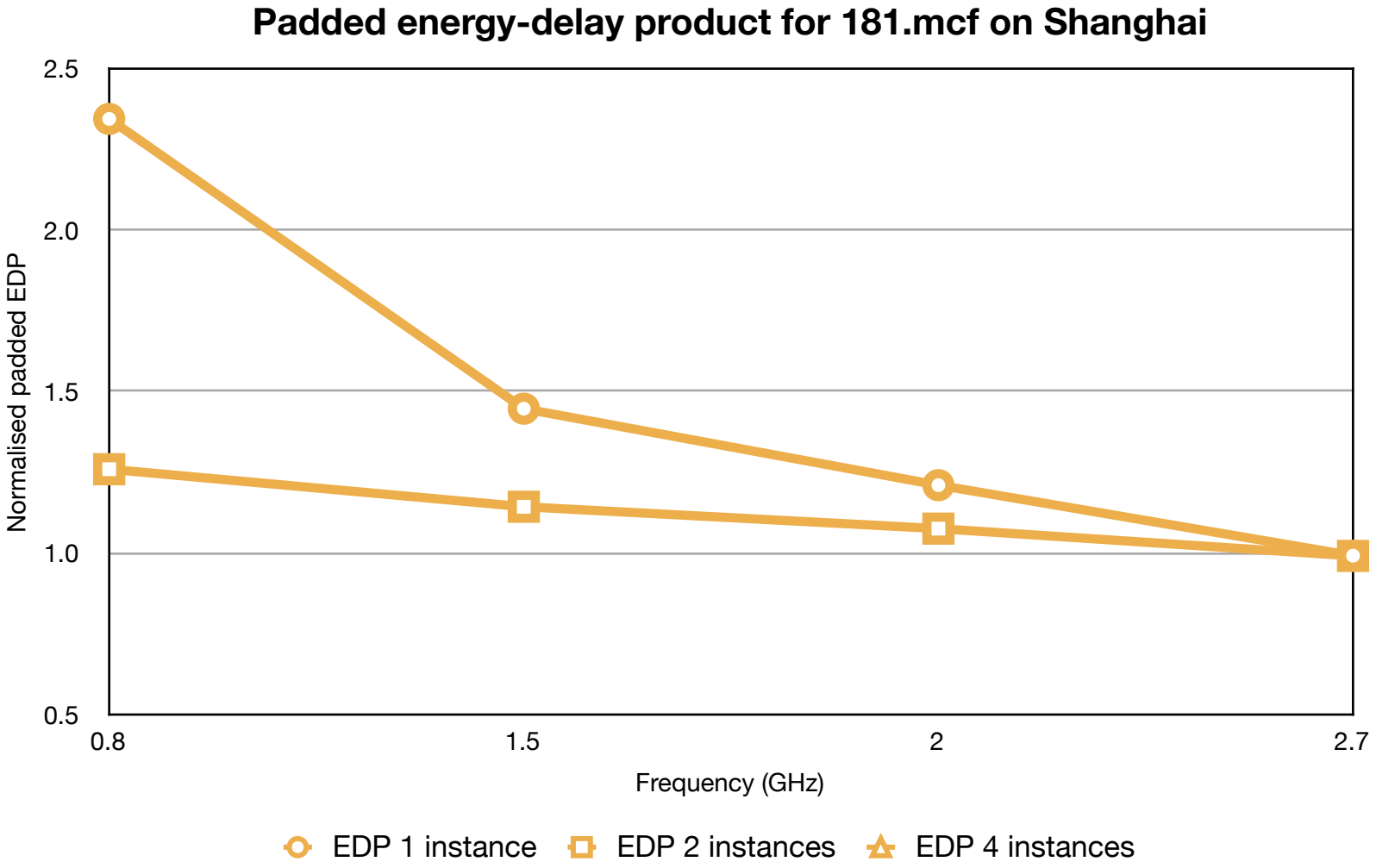
Energy-delay product



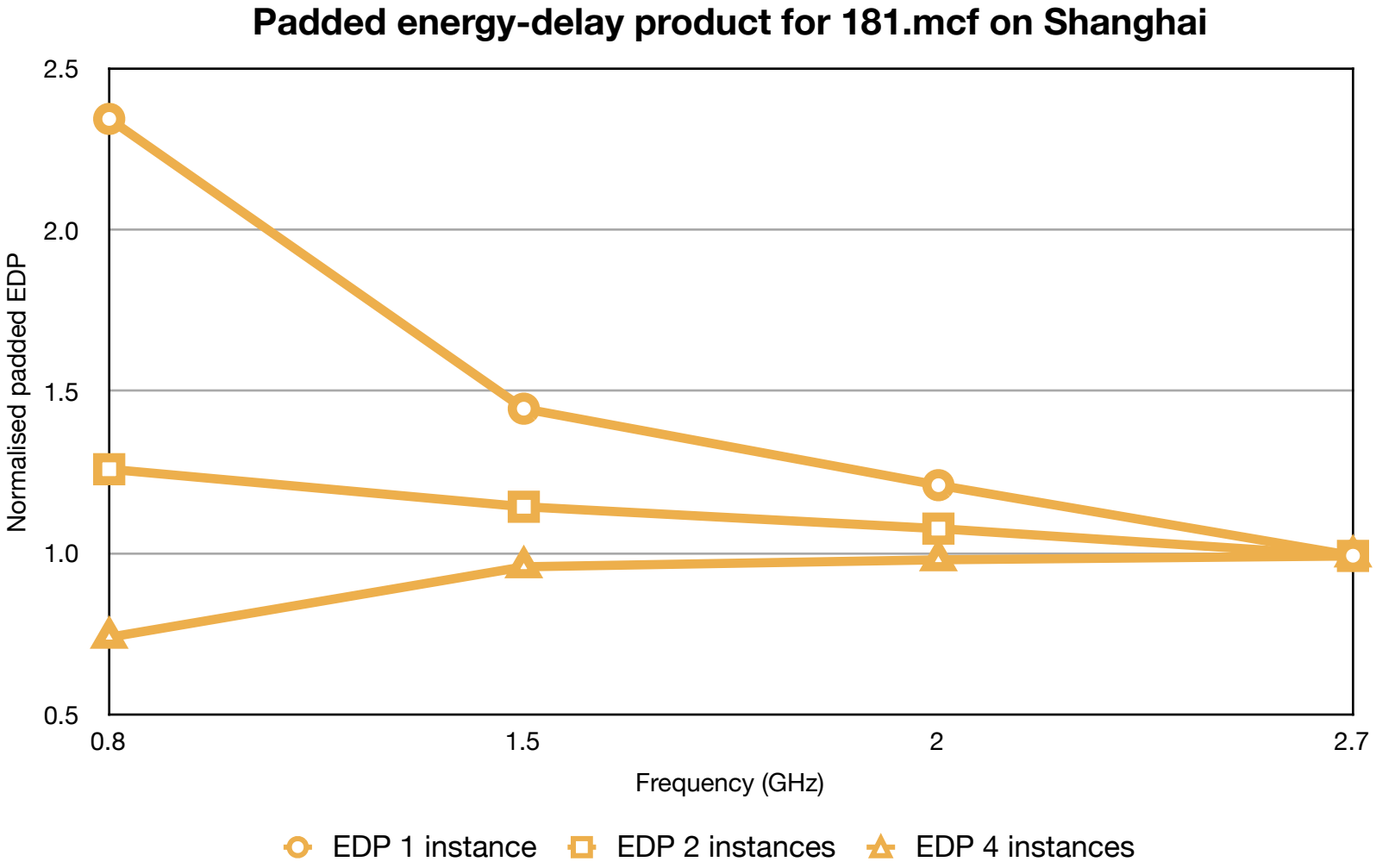
Energy-delay product



Energy-delay product



Energy-delay product



The future...

The future...

- 32 nm process already in production;

The future...

- 32 nm process already in production;
- 22 nm and smaller are on the horizon;

The future...

- 32 nm process already in production;
- 22 nm and smaller are on the horizon;
- caches will get bigger;

The future...

- 32 nm process already in production;
- 22 nm and smaller are on the horizon;
- caches will get bigger;
- leakage power will rise;

The future...

- 32 nm process already in production;
- 22 nm and smaller are on the horizon;
- caches will get bigger;
- leakage power will rise;
- memory performance will continue to improve; and

The future...

- 32 nm process already in production;
- 22 nm and smaller are on the horizon;
- caches will get bigger;
- leakage power will rise;
- memory performance will continue to improve; and
- entry/exit costs to/from sleep modes will improve.

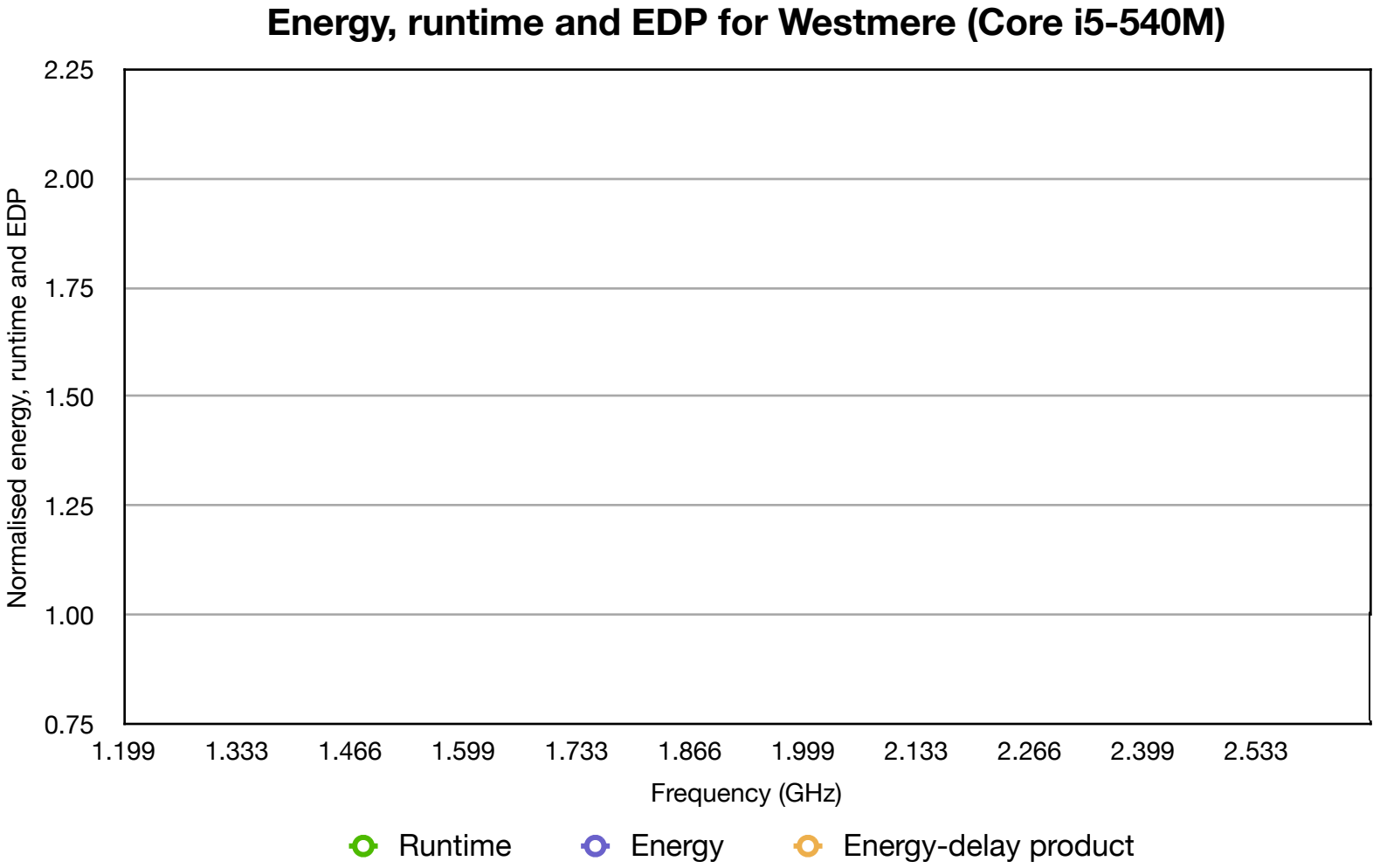
The future...

- 32 nm process already in production;
- 22 nm and smaller are on the horizon;
- caches will get bigger;
- leakage power will rise;
- memory performance will continue to improve; and
- entry/exit costs to/from sleep modes will improve.

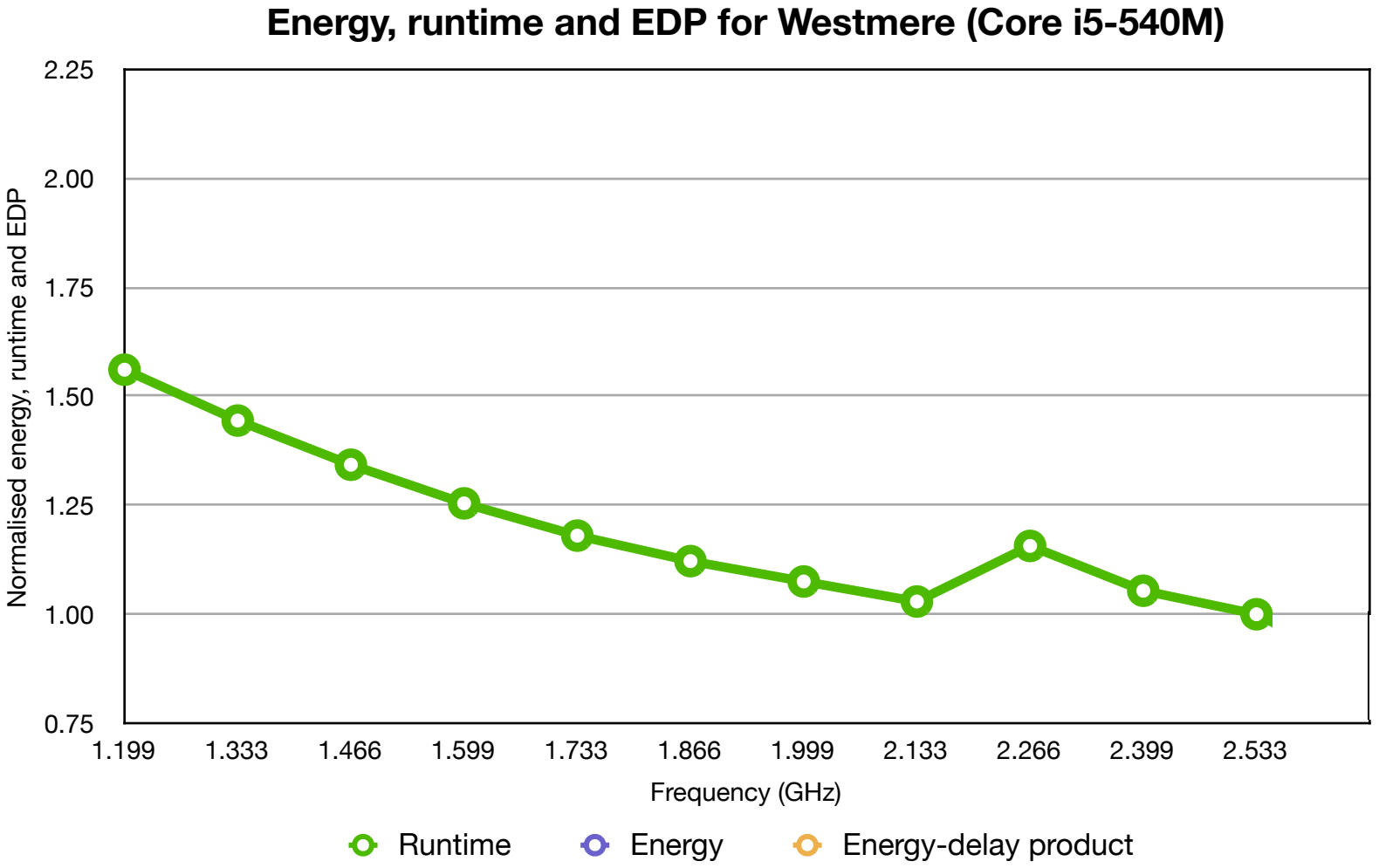
What about DVFS?

A glimpse

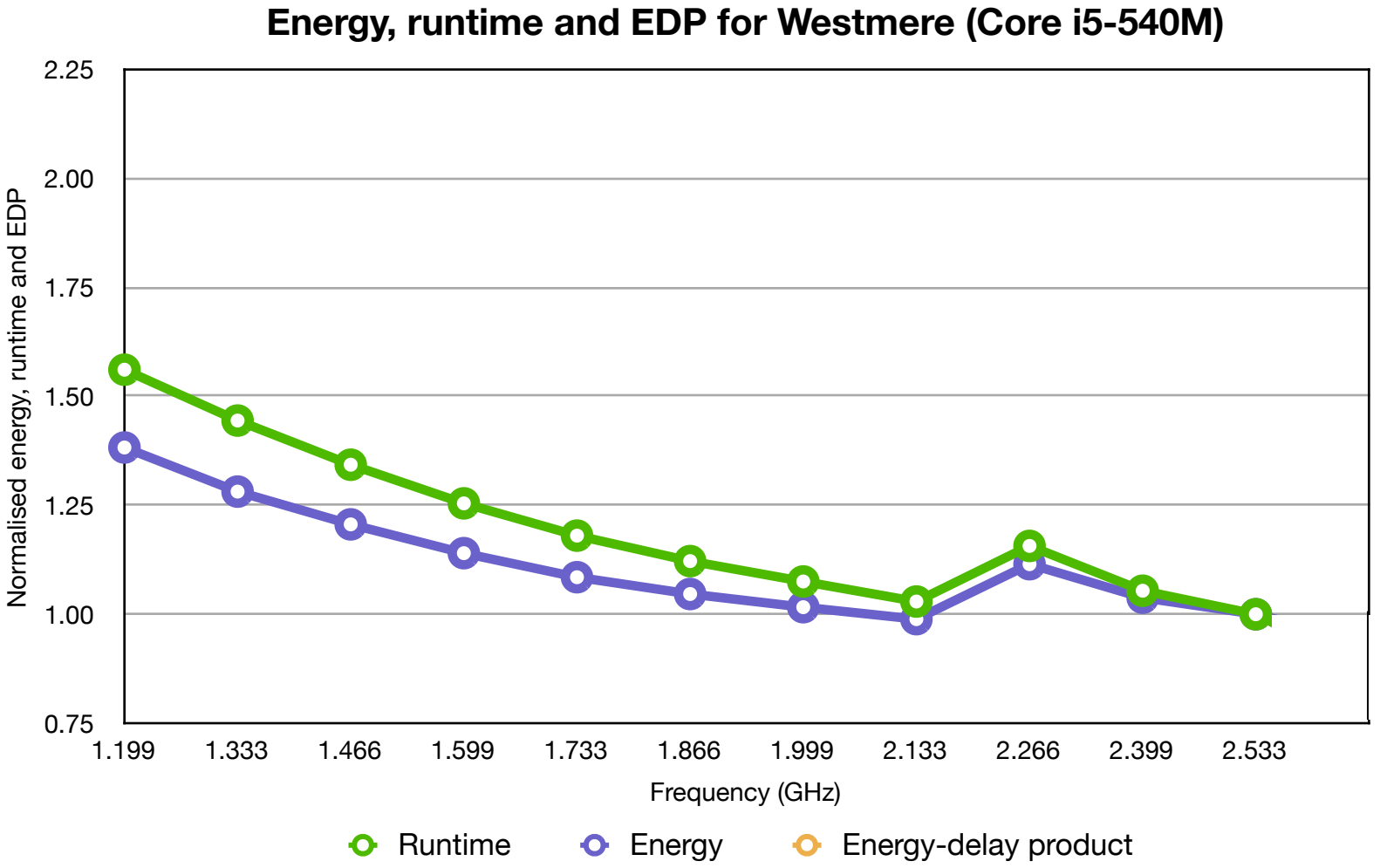
A glimpse



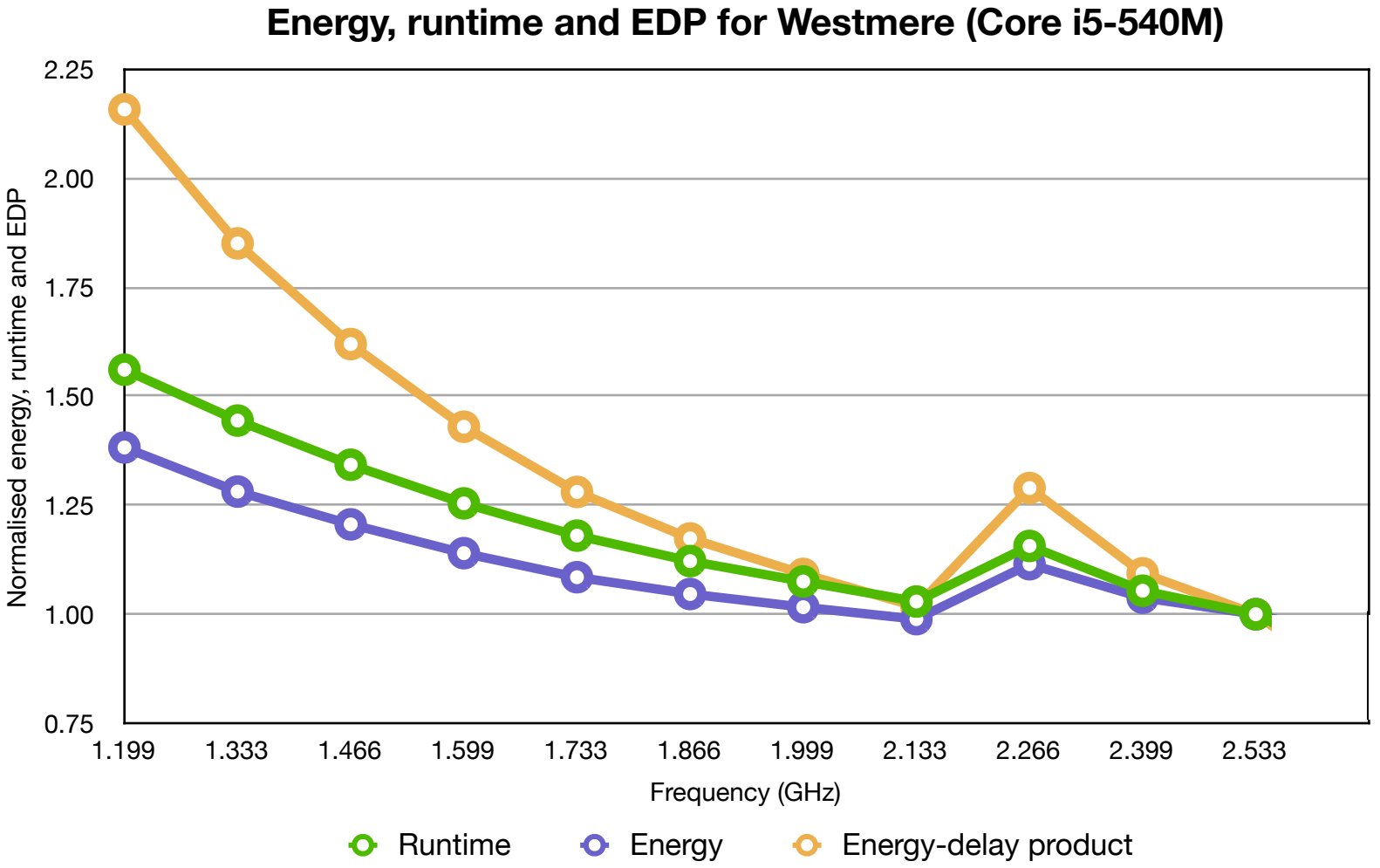
A glimpse



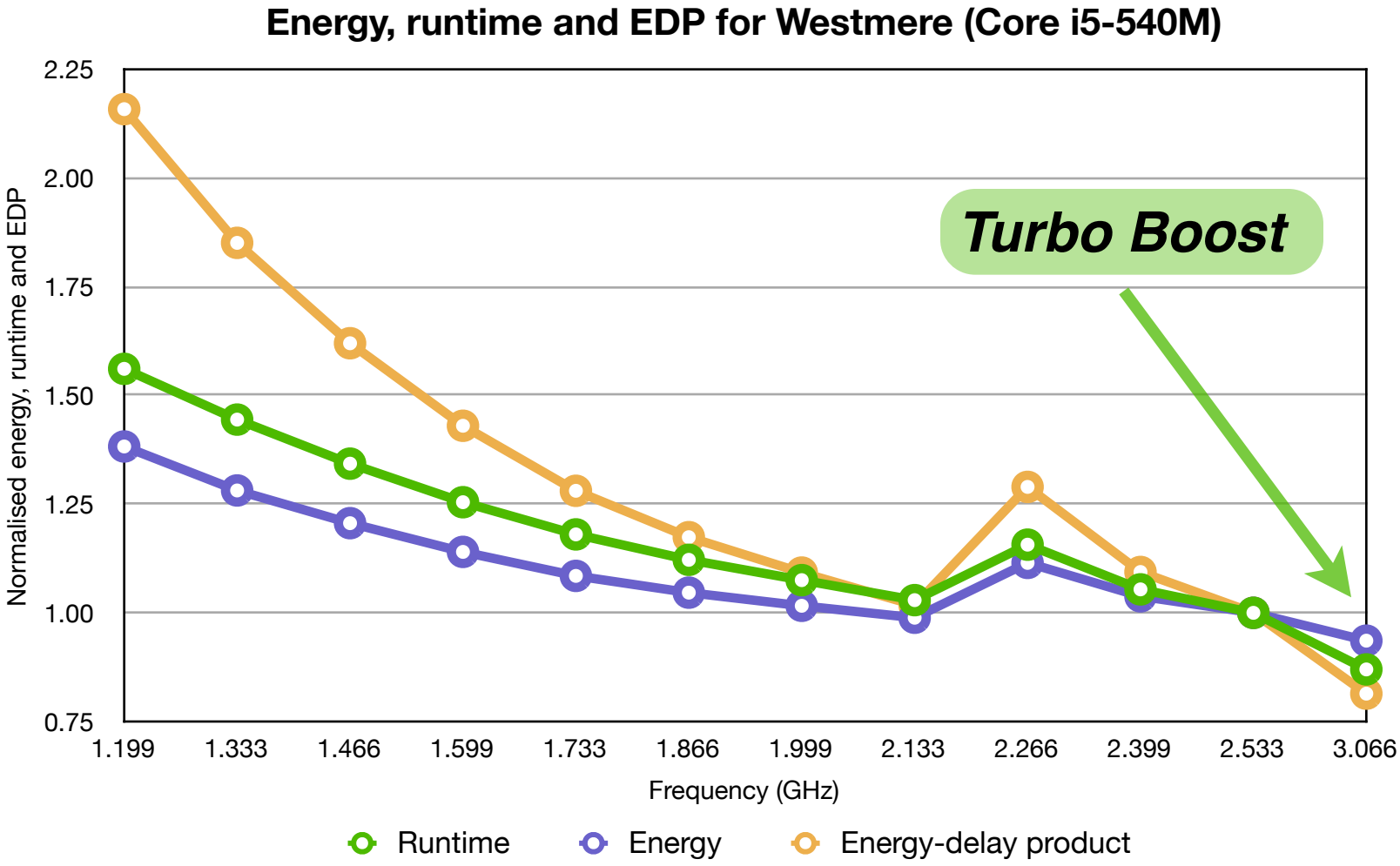
A glimpse



A glimpse



A glimpse



Concluding remarks

Concluding remarks

- Transistor scaling is causing higher proportions of static leakage power;

Concluding remarks

- Transistor scaling is causing higher proportions of static leakage power;
- smaller core voltages mean DVFS has less range to play with;

Concluding remarks

- Transistor scaling is causing higher proportions of static leakage power;
- smaller core voltages mean DVFS has less range to play with;
- improving memory performance means fewer opportunities to reduce CPU frequency without significant loss of performance;

Concluding remarks

NICTA

- Transistor scaling is causing higher proportions of static leakage power;
- smaller core voltages mean DVFS has less range to play with;
- improving memory performance means fewer opportunities to reduce CPU frequency without significant loss of performance;
- sleep/idle modes are becoming much more efficient;

Concluding remarks

- Transistor scaling is causing higher proportions of static leakage power;
- smaller core voltages mean DVFS has less range to play with;
- improving memory performance means fewer opportunities to reduce CPU frequency without significant loss of performance;
- sleep/idle modes are becoming much more efficient;
- DVFS implementations on multi-core processors are more complex and the cost-benefit is small;

Concluding remarks

NICTA

- Transistor scaling is causing higher proportions of static leakage power;
- smaller core voltages mean DVFS has less range to play with;
- improving memory performance means fewer opportunities to reduce CPU frequency without significant loss of performance;
- sleep/idle modes are becoming much more efficient;
- DVFS implementations on multi-core processors are more complex and the cost-benefit is small;
- Optimal energy-efficiency is achieved by running ***fast.***

Future direction

Future direction

- SIMD workloads

Future direction

- SIMD workloads
  - SSE/3Dnow! workloads require many 64bit operands, which could increase memory-boundedness

Future direction

- SIMD workloads
  - SSE/3Dnow! workloads require many 64bit operands, which could increase memory-boundedness
- Periodic workloads

Future direction

- SIMD workloads
  - SSE/3Dnow! workloads require many 64bit operands, which could increase memory-boundedness
- Periodic workloads
  - MPEG video/audio playback, analyse race-to-halt

Future direction

- SIMD workloads
  - SSE/3Dnow! workloads require many 64bit operands, which could increase memory-boundedness
- Periodic workloads
  - MPEG video/audio playback, analyse race-to-halt
  - usefulness of sleep modes

Future direction

NICTA

- SIMD workloads
  - SSE/3Dnow! workloads require many 64bit operands, which could increase memory-boundedness
- Periodic workloads
  - MPEG video/audio playback, analyse race-to-halt
  - usefulness of sleep modes
- Analyse and compare the trends for embedded devices

Future direction

- SIMD workloads
  - SSE/3Dnow! workloads require many 64bit operands, which could increase memory-boundedness
- Periodic workloads
  - MPEG video/audio playback, analyse race-to-halt
  - usefulness of sleep modes
- Analyse and compare the trends for embedded devices
  - Mobile phones and netbooks

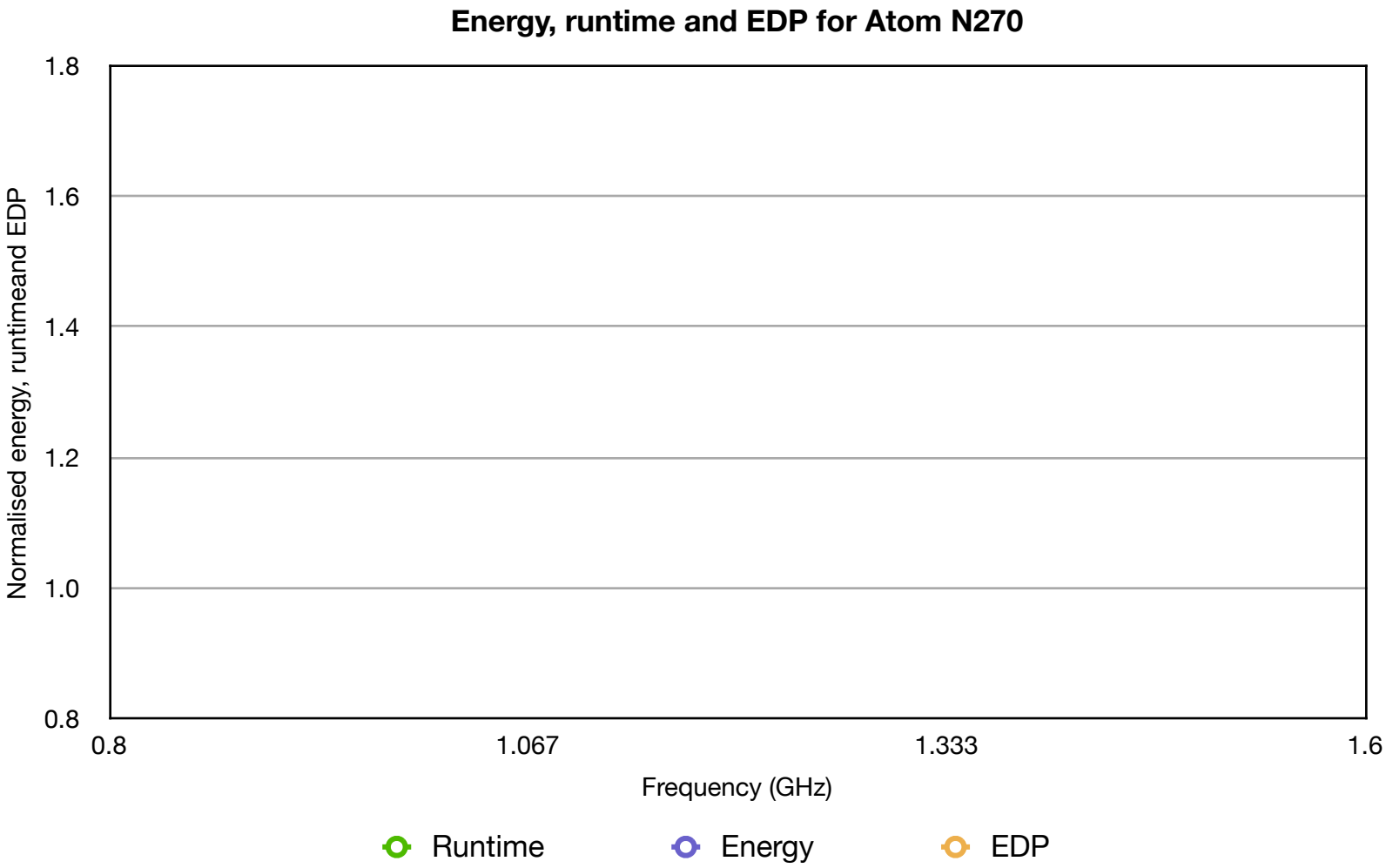
Questions?

Questions?

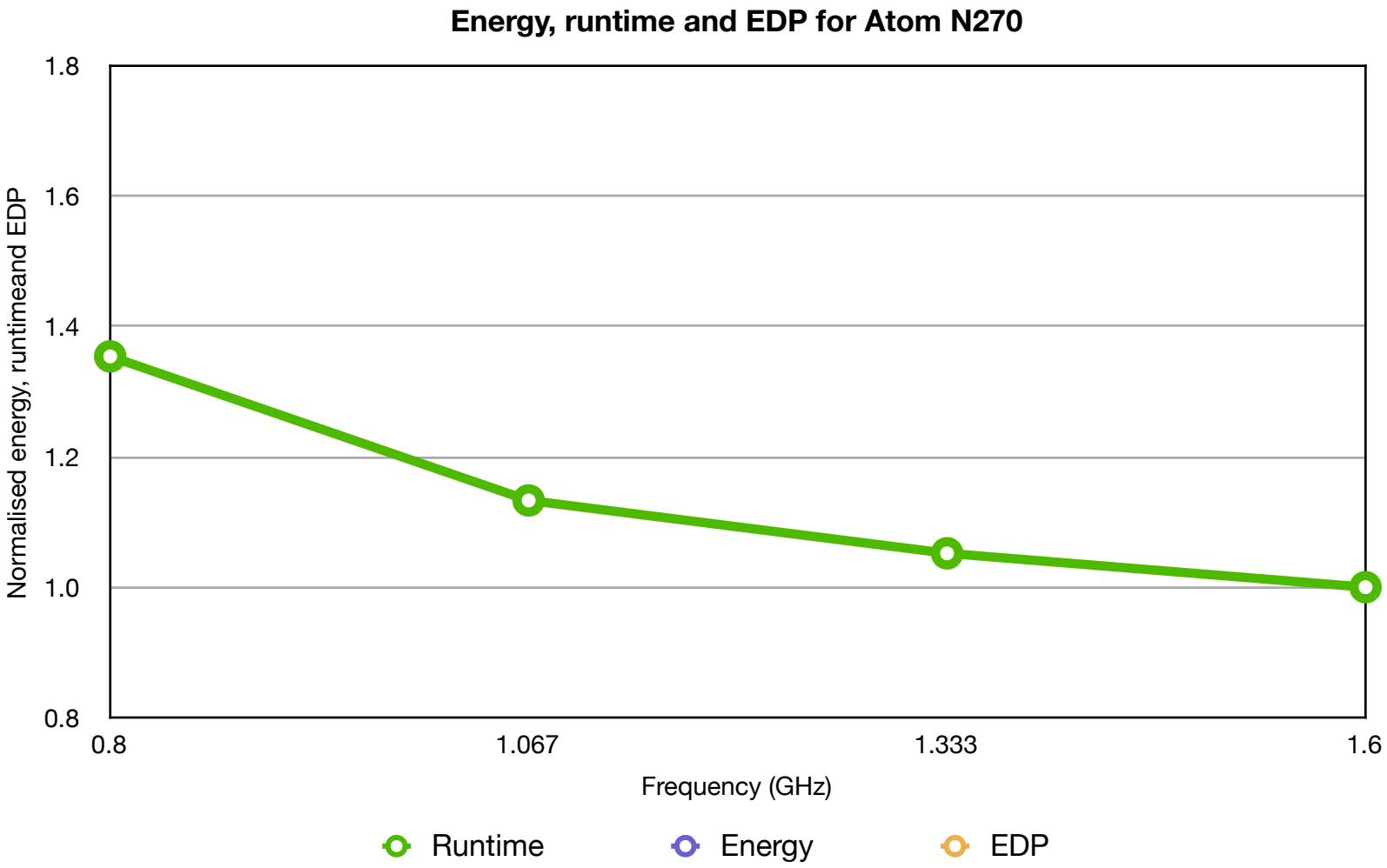
NICTA

Intel Atom N270

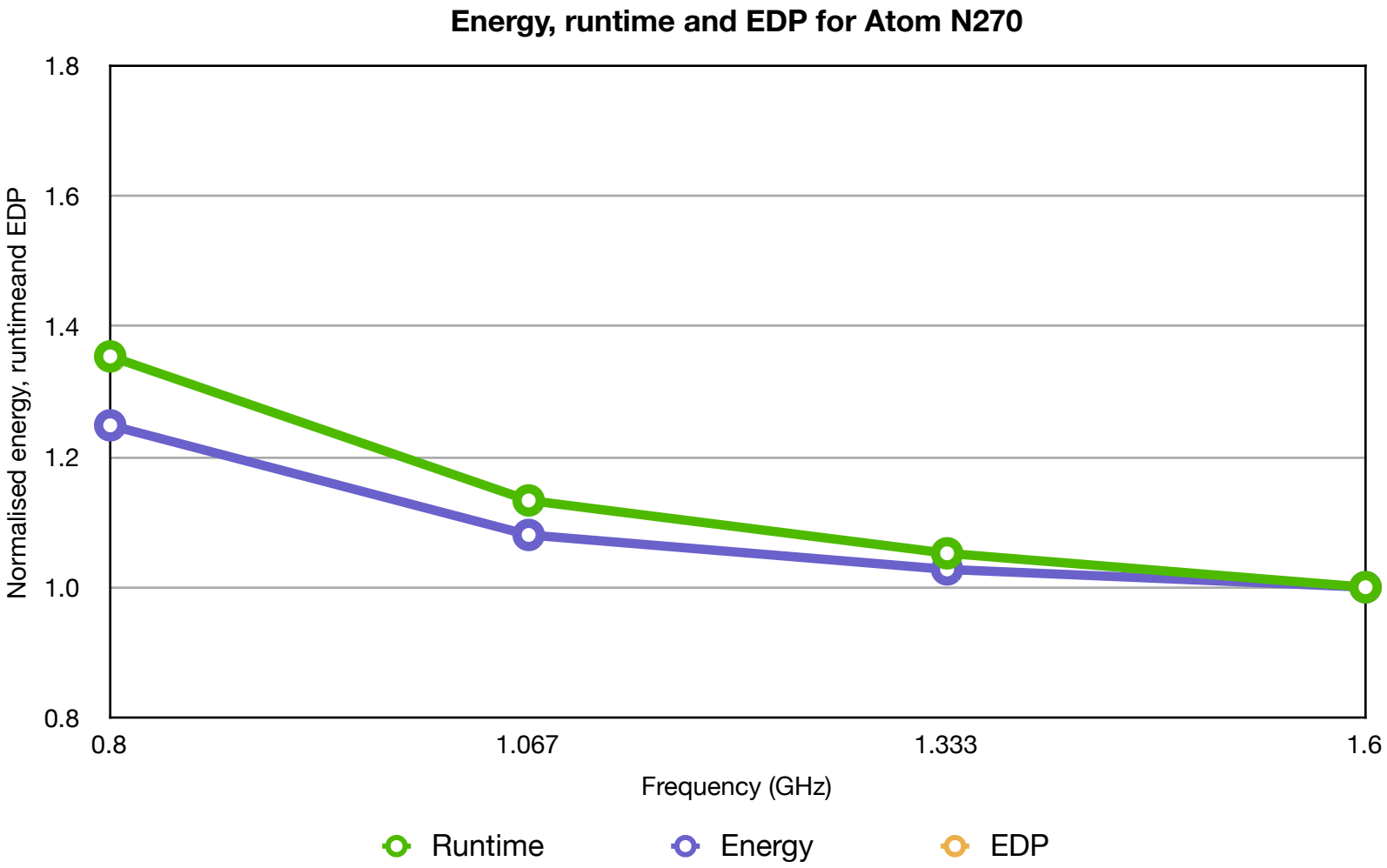
Intel Atom N270



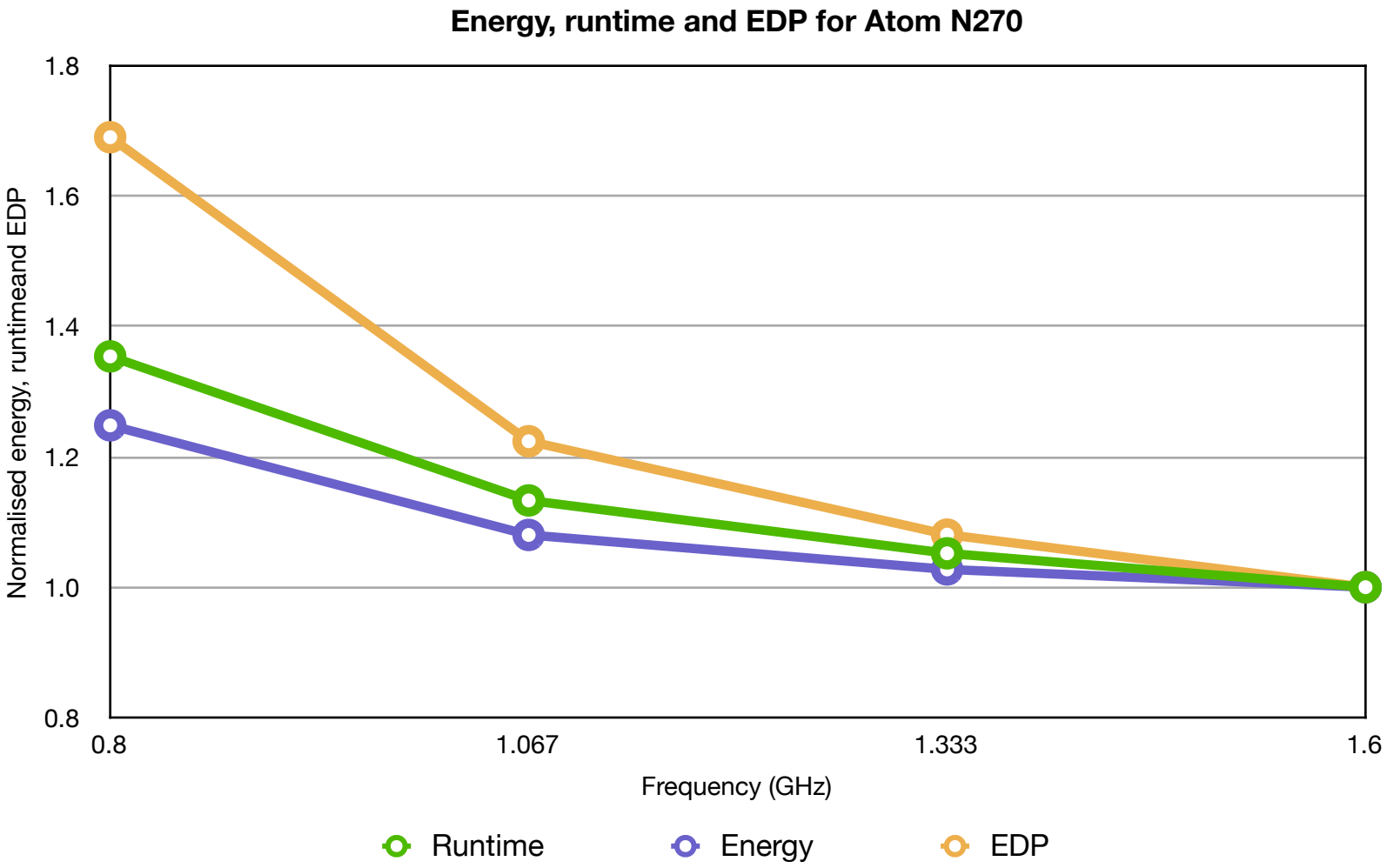
Intel Atom N270



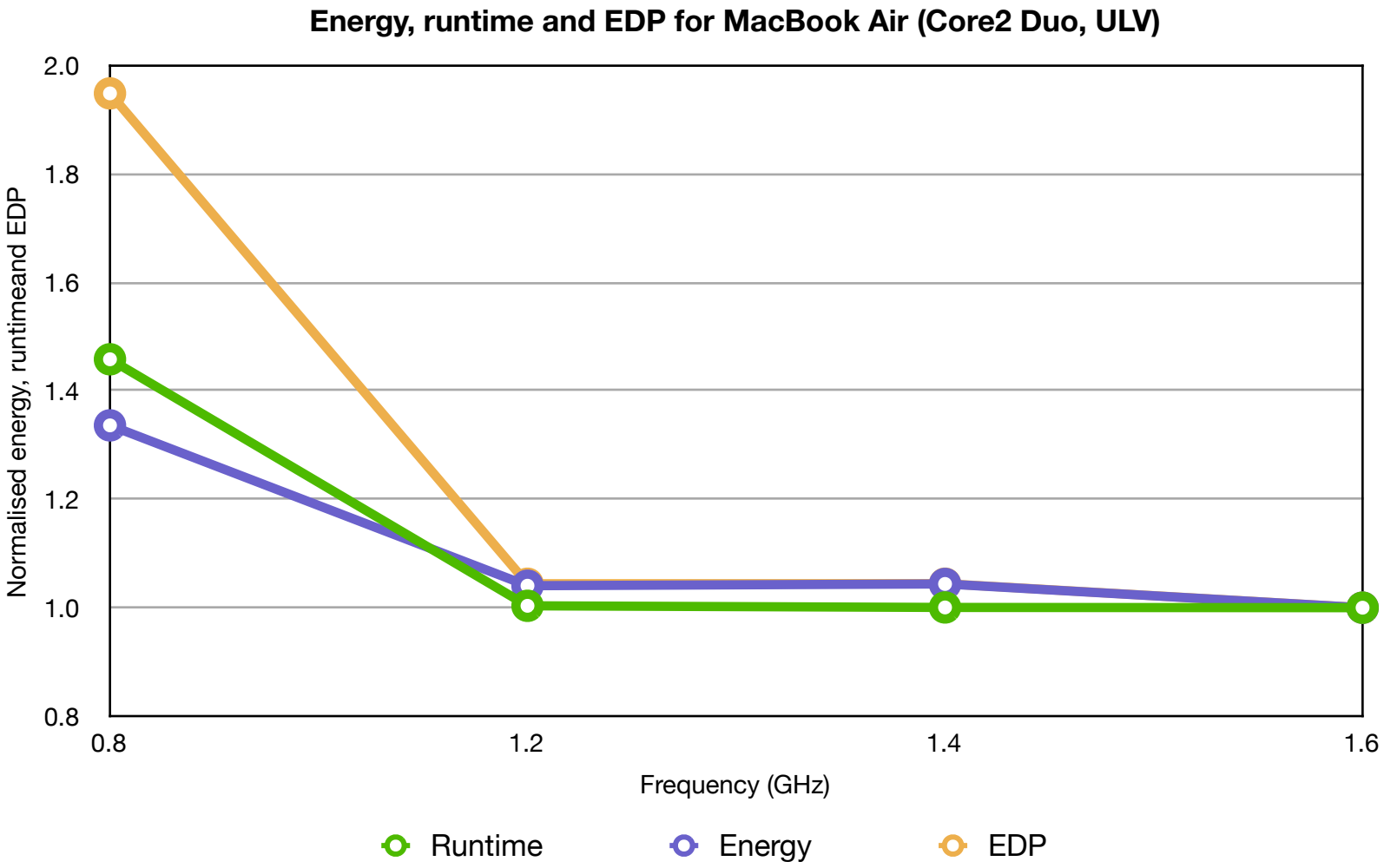
Intel Atom N270



Intel Atom N270



Intel Core 2 Duo (Ultra-low Voltage)




From imagination to impact

NICTA

From **imagination** to **impact**