

# 計算機の構成とワークロードの違いによる 消費電力削減効果の現実

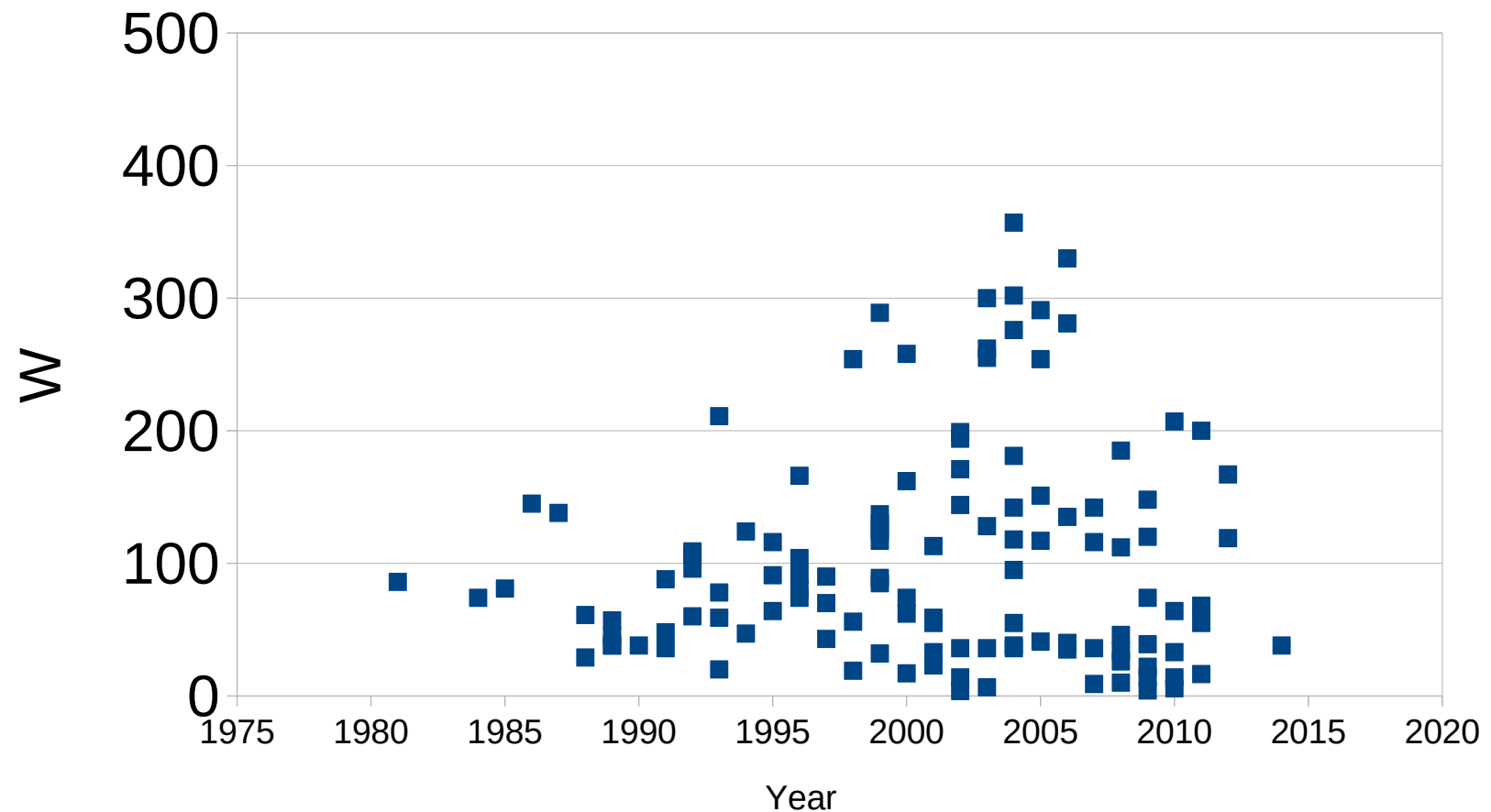
泊 久信, 平木 敬  
東京大学

# 消費電力

- 半導体製造技術の向上が十分な消費電力削減に繋がらない
  - 増えるリーク電流, 下がらない電圧
- 消費電力が性能向上の足枷に
- 性能向上を維持する<sup>↓</sup>ため、コンピュータ設計での消費電力を意識する重要性増す

# デスクトップ・サーバーでは


- 現在に至るまで、システム単位での消費電力はほぼ一定



# 高性能計算機の例

- システムあたりのソケット数は増え続けている
  - より多くのプロセッサを並べてる
- 電力増える
  - 電気代
  - 冷房代

# クラウド・コンピューティング

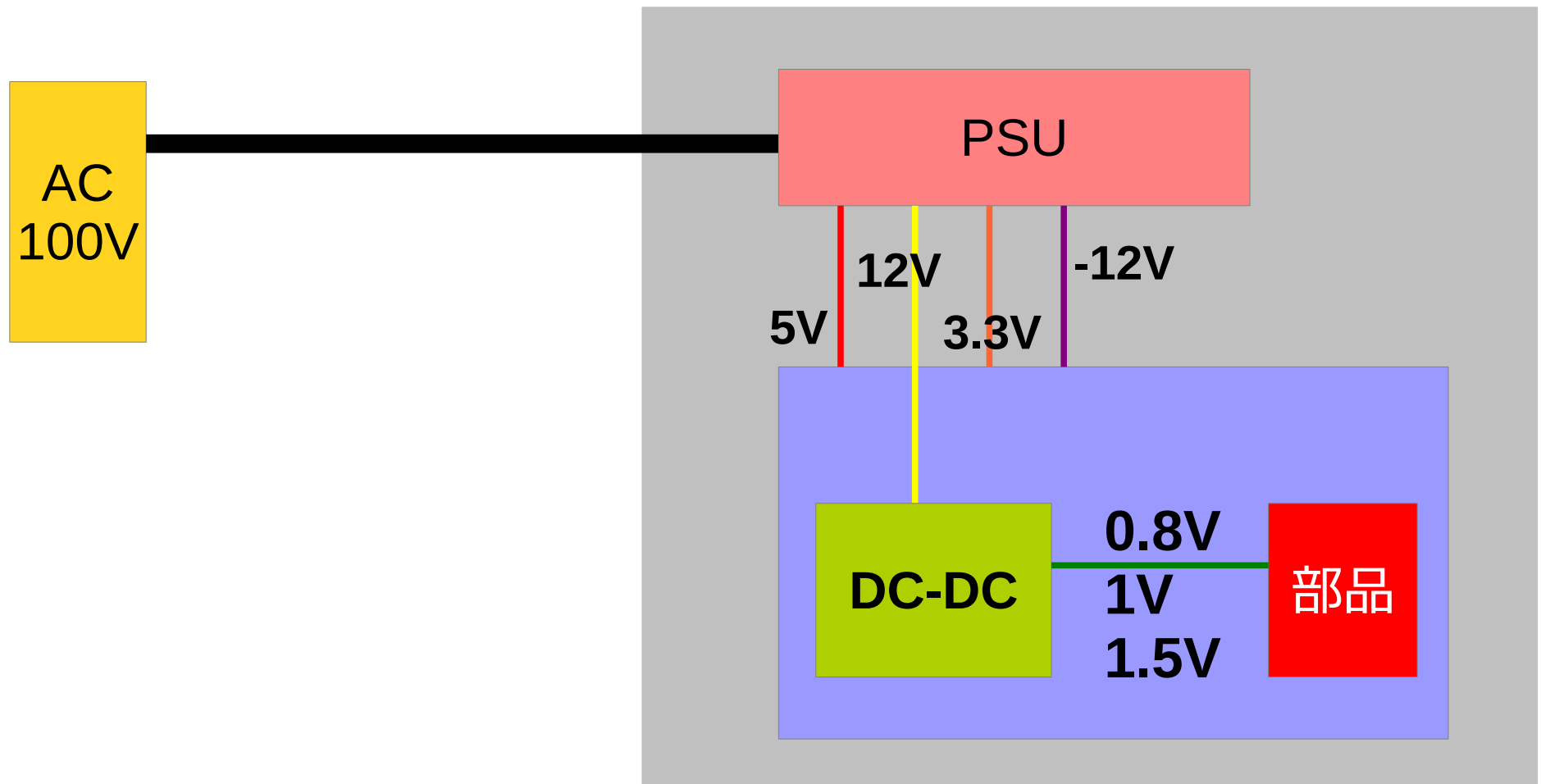
- サーバー仮想化・集約化
  - 仮想マシンを積極的にマイグレーションして消費電力を削減する、といった研究が存在
- 
- 負荷が変化することで、実際の計算機でどれだけの消費電力削減効果があるのか？

# 本研究

- コンピュータのどの部品がどれだけの電力を消費しているのか実測しよう
  - 「実測する」というのがポイント

# 測定方法

- 消費電力の実測にはいくつか方法がある



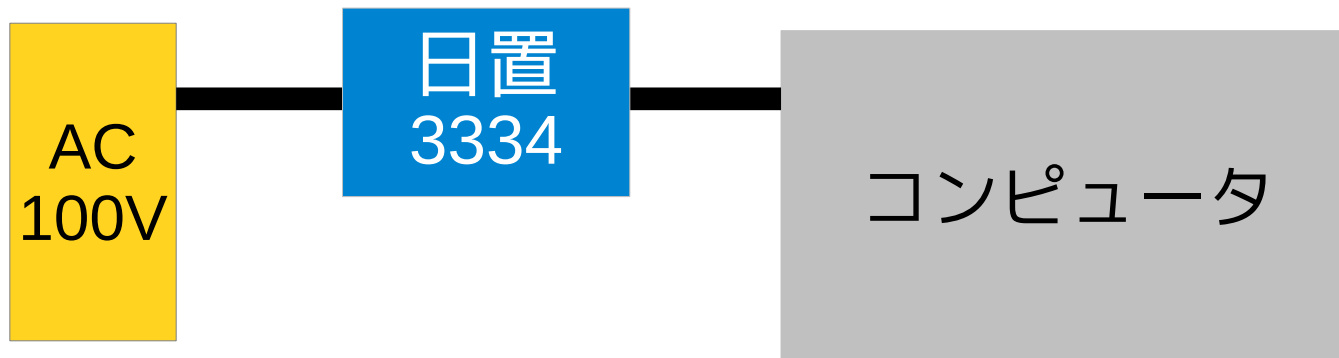
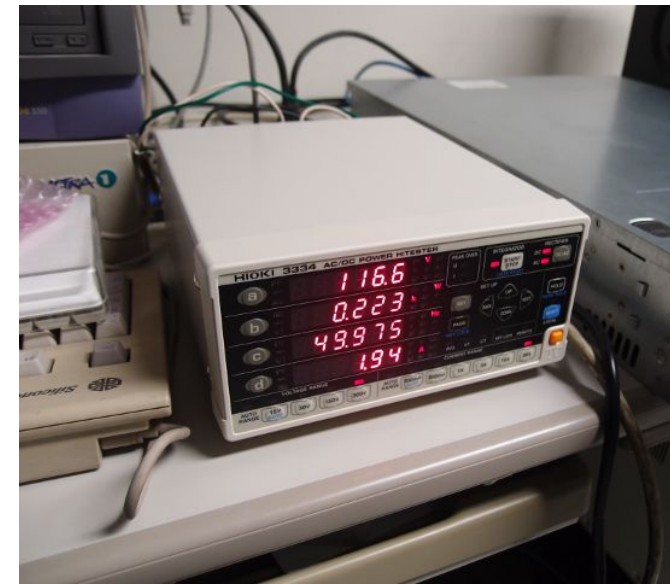
# どこで測る？

- 部品の電源ピン全て？
  - BGAの部品が多く、困難
  - 部品の消費電力を正確に測定可能
- PSUからの出力？
  - いくつかの部品で共通の場合がある(メモリ,CPU)
  - 直流なので簡単で正確に測定できる
- AC入力
  - 全ての部品の合計値+損失しか測定できない
  - 実際に課金される電力はここ



# 測定セットアップ

- 日置3334電力計で交流電源をタップ
  - 電流と電圧を測定
- 測定時, 部屋は26°Cに空調
- 電源は実測で117 V, 50.0 Hz



# 測定に利用した部品構成

- CPU: AMD Opteron 6282SE
  - OS6282YETGGGU
  - ACP 105W, TDP 140W
- 箱: Supermicro AS-2022G-URF+
- マザーボード: Supermicro H8DGU-LN4F+



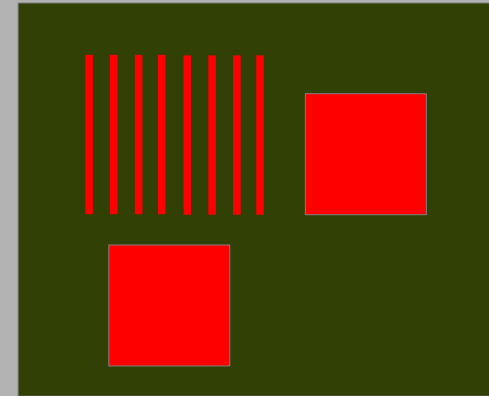
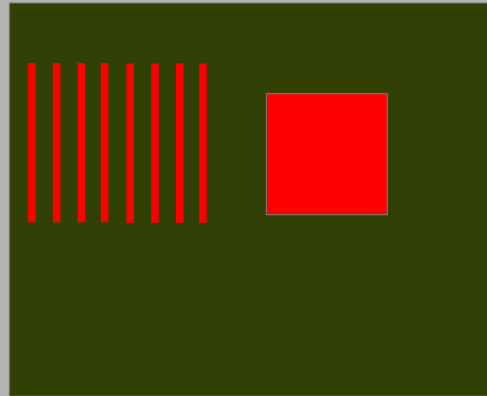
# 負荷

- Dhrystone
  - キャッシュからあふれない
  - CPU「だけ」に負荷を与えられる
- メモリ
  - 自作メモリ負荷プログラム、線形に読む
  - AVX, Non-temporal store, prefetchを利用

# 測定結果

- CPU
- メモリ
- NIC
- ファン
- HDD

# CPU消費電力の実際



待機時

121 W

147 W

1スレッド

148 W

195 W

全スレ  
ッド

238 W

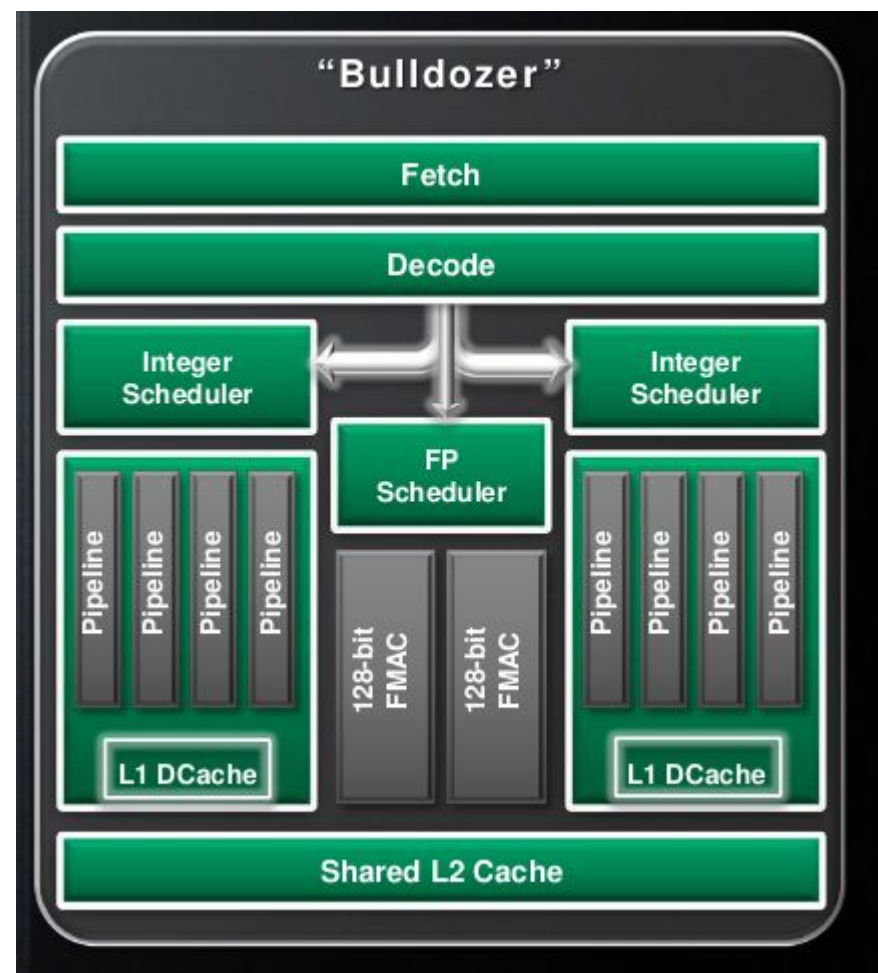
340 W

# CPU消費電力測定よりわかること

- プロセッサ単体のアイドル時電力は26 W
- Dhrystone時は102 W
  - FPU使っていないためTDP 140Wに比べ小さい
- 待機時にマザーボード＋メモリが90W近く消費
  - メモリの結果を合わせて議論する

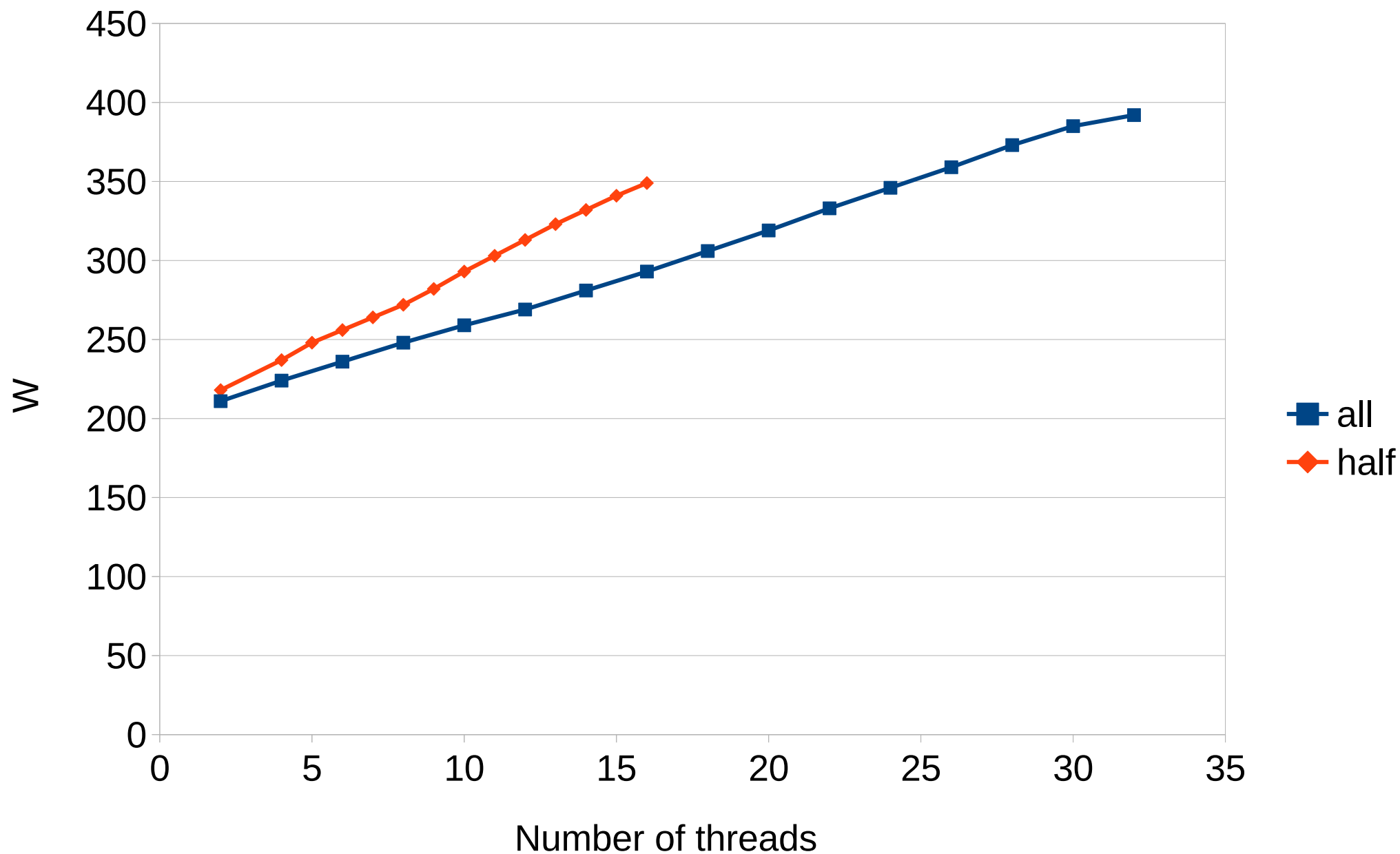
# CPU: スレッド数と消費電力

- Opteron 6282SEは、モジュールという単位で2スレッド動作する
  - デコーダ・L1I・FPUを共有
- モジュールごとに1スレッド割り当てた場合と全ての論理コアに1スレッド順に割り当てた場合
- numactl --physcpubind=XX



[Sean, Hotchips 23, 2011]

# CPU: スレッド数と消費電力





# CPU: スレッド数と消費電力

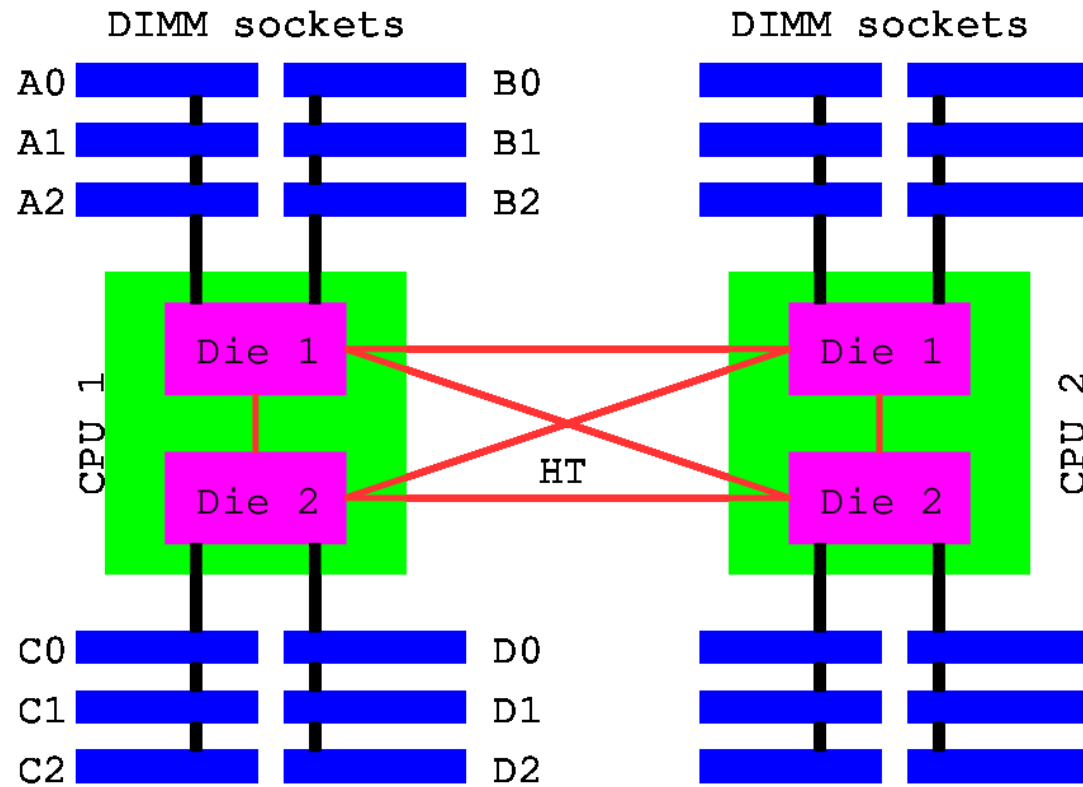
- ほぼ線形
  - モジュールあたり12.1 W
- 消費電力の差から
  - L1D, 整数実行パイプライン, 整数スケジューラ
    - 2.69 W
  - その他
    - 9.41 W

# CPUの性能/電力

- 1T/moduleで16スレッドを動かしたとき
- 2T/moduleで20スレッドを動かしたとき
  - これらの合計性能は概ね同じ
- 2T/moduleで20T動かしたときの電力
  - 1T/moduleで13T動かしたときと同じ
- 性能/電力を最大化するなら全スレッド使う

# メモリ

- 今回のOpteronはソケットあたり4ch DDR3
- メモリの配置で有効になるDDR3コントローラの数も変化させられる



# 使用したメモリモジュール

- DDR3-1333 RDIMM 36チップ, 16 GB
- DDR3-1600 RDIMM 36チップ, 16 GB



# 測定時の特記事項

- メモリのチャネル数が増加するとメモリバンド幅が増加
  - CPU使用率が上がりその分の消費電力も増加

# メモリ測定結果

待機時消費電力 (W)

負荷時消費電力 (W)

実測スループット (MBytes/s)

#channels	1 DIMM	2 DIMM	4 DIMM	8 DIMM	16 DIMM	8 DIMM (1600)
1 (1333)	142	142				
	291	297				
	8345	8104				
2 (1333)		142	142			
		300	306			
		12892	12757			
4 (1333)			144	147		
			315	331		
			24362	23865		
8				149	156	150
				350	402	352
				42179	40118	39420
4 (1333/Y)				147		
				340		
				24208		

# メモリ測定結果

待機時消費電力 (W)

負荷時消費電力 (W)

実測スループット (MBytes/s)

#channels	1 DIMM	2 DIMM	4 DIMM	8 DIMM	16 DIMM	8 DIMM (1600)
1 (1333)	142	142				
	291	297				
	8345	8104				
2 (1333)		142	142			
		300	306			
		12892	12757			
4 (1333)			144	147		
			315	331		
			24362	23865		
8				149	156	150
				350	402	352
				42179	40118	39420
4 (1333/Y)				147		
				340		
				24208		

# メモリ測定結果

待機時消費電力 (W)

負荷時消費電力 (W)

実測スループット (MBytes/s)

#channels	1 DIMM	2 DIMM	4 DIMM	8 DIMM	16 DIMM	8 DIMM (1600)
1 (1333)	142	142				
	291	297				
	8345	8104				
2 (1333)		142	142			
		300	306			
		12892	12757			
4 (1333)			144	147		
			315	331		
			24362	23865		
8				149	156	150
				350	402	352
				42179	40118	39420
4 (1333/Y)				147		
				340		
				24208		



# メモリ測定結果

待機時消費電力 (W)

負荷時消費電力 (W)

実測スループット (MBytes/s)

#channels	1 DIMM	2 DIMM	4 DIMM	8 DIMM	16 DIMM	8 DIMM (1600)
1 (1333)	142	142				
	291	297				
	8345	8104				
2 (1333)		142	142			
		300	306			
		12892	12757			
4 (1333)			144	147		
			315	331		
			24362	23865		
8				149	156	150
				350	402	352
				42179	40118	39420
4 (1333/Y)				147		
				340		
				24208		

# メモリ測定結果

待機時消費電力 (W)

負荷時消費電力 (W)

実測スループット (MBytes/s)

#channels	1 DIMM	2 DIMM	4 DIMM	8 DIMM	16 DIMM	8 DIMM (1600)
1 (1333)	142	142				
	291	297				
	8345	8104				
2 (1333)		142	142			
		300	306			
		12892	12757			
4 (1333)			144	147		
			315	331		
			24362	23865		
8				149	156	150
				350	402	352
				42179	40118	39420
4 (1333/Y)				147		
				340		
				24208		

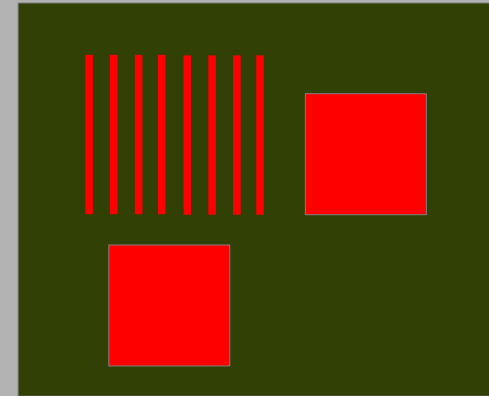
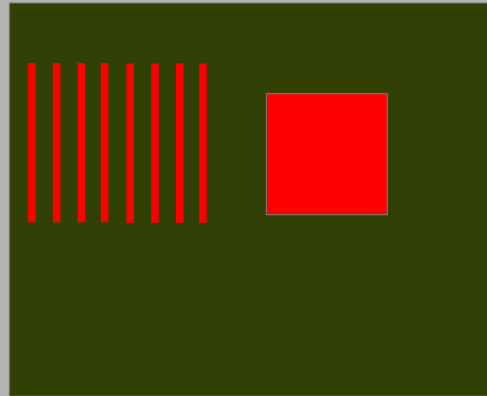
# メモリ結論

- PHYリンク電力 (CPU側) 0.5 W/チャネル
- DDR3-1333 SDRAM RDIMM
  - 待機時 0.75 W
  - 負化時 4 W, 5.13 W~6.5 W
- DDR3-1600にしても大きな差はなし
  - チップによる差の方が大きい

# ここまでの測定結果から

- CPU/メモリを除いたマザーボードの消費電力
  - 待機時の数値から計算すると
    - 89 W / 147 W
  - 負化時の結果から計算すると
    - 81 W / 340 W
- マザーボードの消費電力は大きい
  - 従来プロセッサばかりに目が行きがちだが、こちらでも低消費電力化の努力が必要

# CPU消費電力の実際



待機時

121 W

147 W

1スレッド

148 W

195 W

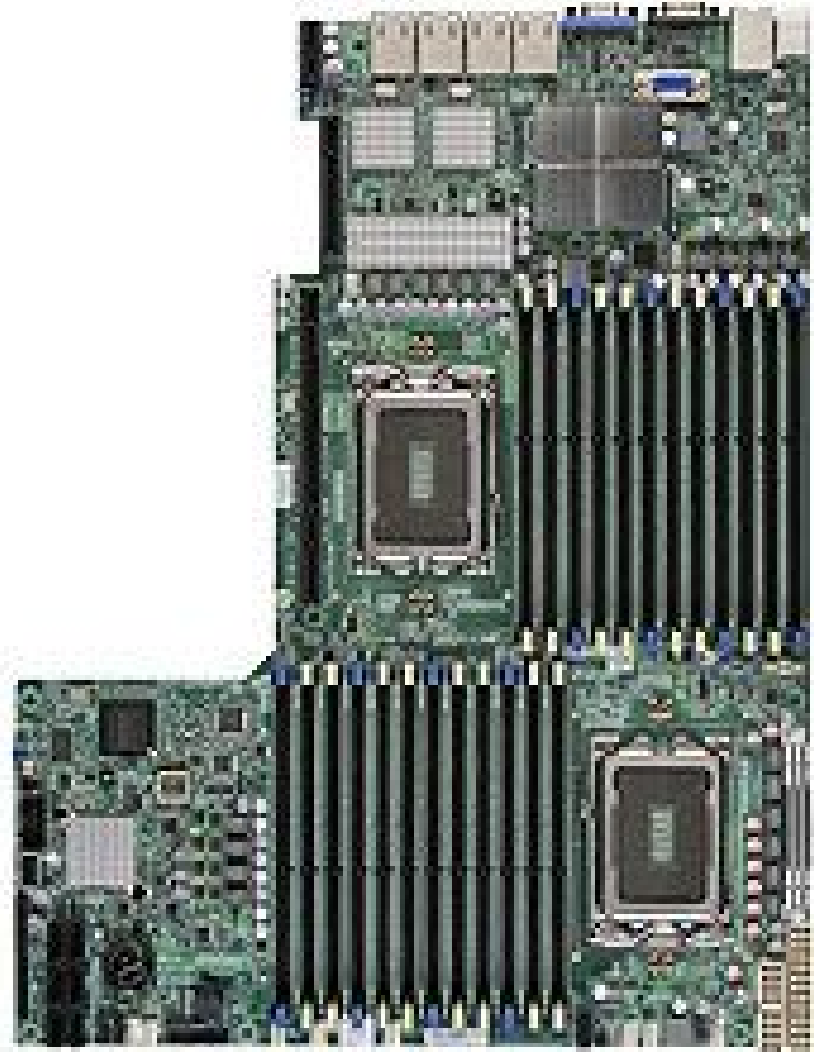
全スレ  
ッド

238 W

340 W

# マザーボード

- BMC
  - S5: 16W
- 4-port GbE
- チップセット
- ビデオ回路

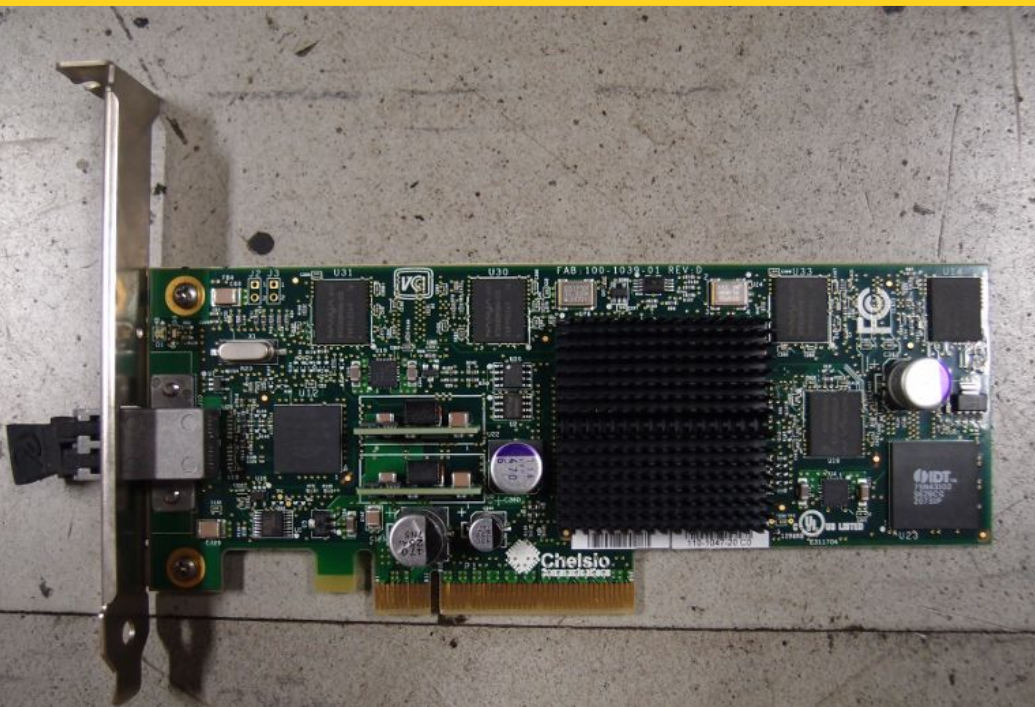


# NIC

- Chelsio T310 (10GBASE-SR)
  - リンクダウン時 9.5 W
  - リンクアップ時 14.5 W (リンク電力 5 W)
- MT25204 (IB 4X DDR)
  - リンクダウン時 4 W
  - リンクアップ時 5 W (リンク電力 1 W)
    - パッシブケーブル使用のため参考値

# NIC結果

- Chelsio T310はチップ数が多い
  - 世代として古い





# ファン

- 三洋電機Model 9A0812S4051
  - 12V, 0.18 A (2.16 W) [データシート]
  - 3400 rpm
  - 42.4 CFM
- 実測すると一台あたり1.33 W
  - データシート通りの値にはならない

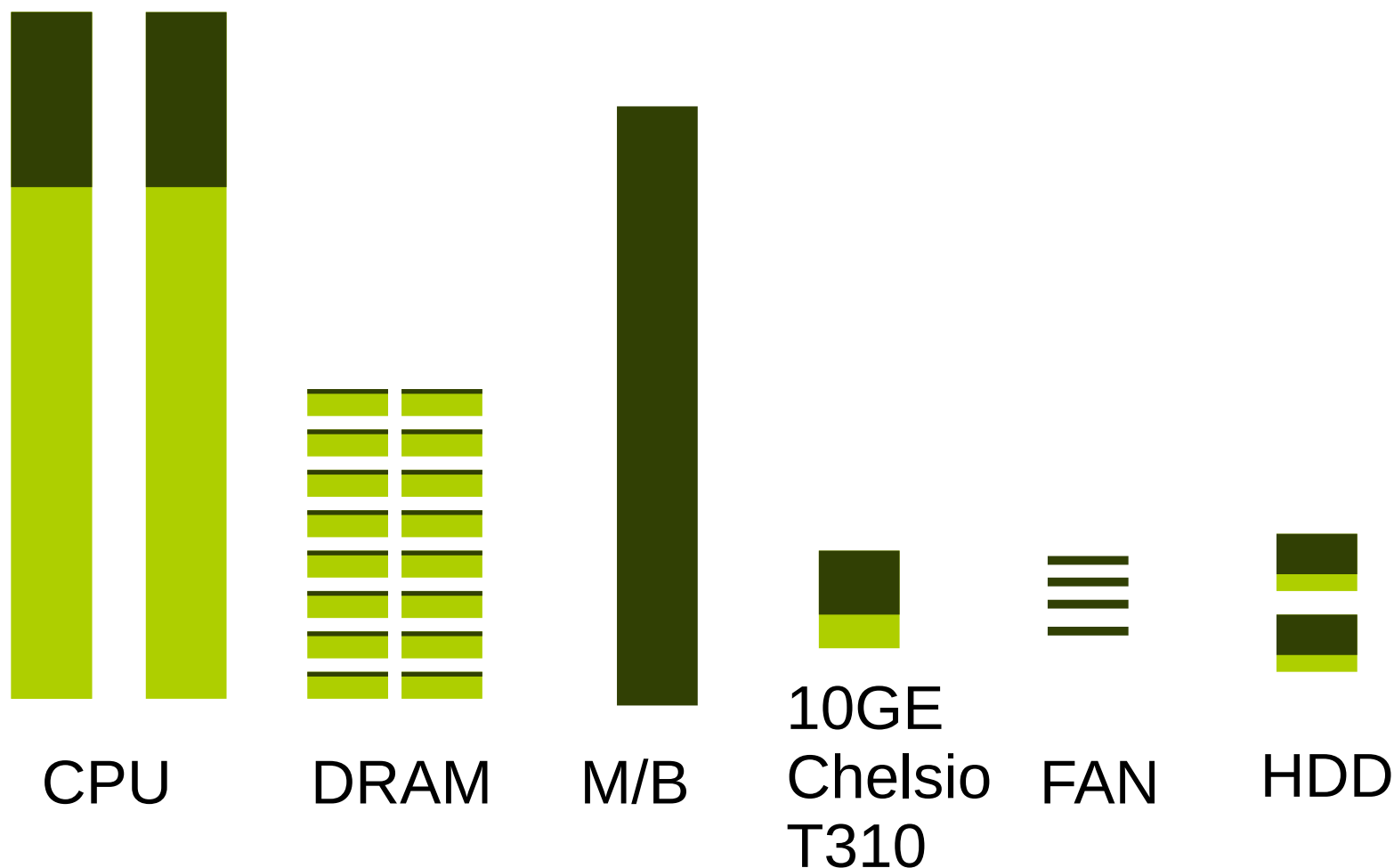
# HDD

- LinuxのSoftware RAIDの台数を変え、消費電力の差から算出
- WD30EZRX-00MMMB0
  - 無負化時 6 W
  - 書き込み時 8.5 W



# 消費電力の地図

- 濃い色が待機時、淡色が負化時の追加消費



# 結論

- コンピュータ実機の部品構成を変化させ、消費電力を測定
- 各部品の消費電力のブレークダウンを算出
- 必ずしもデータシート通りではない
- マザーボードの消費電力が大きい
  - 削減が必要



# Intel

CPU	nominal MHz	Motherboard	BIOS	PSU
Core i7-4770S	3100	ASUS Z87M-PLUS	1105	Antec EA-450 Platinum
Core i7-3770K	3500	ASRock Z77 Pro3	P2.10	Antec EA-450 Platinum
Core i7-2700K	3500	GIGABYTE GA-Z68M-D2H	F6	Antec EA-450 Platinum

