

DRPM: Dynamic Speed Control for Power Management in Server Class Disks

by: Gurumurthi et. al.
presented by: James Larkby-Lahet

Motivation

- 22.3 W IBM Ultrastar 36ZX server disk vs. 57.8 W 1.6 GHz Intel Xeon CPU
- 4-way Xeon SMP with 140 disks used by Dell for TPC benchmark uses 13.5 times more power for disks than CPUs
- current approaches
 - on - no perf compromise or power saved
 - spindown - savings, but serious penalty

DRPM vs TPM

- Spindle motor is 50-80% of idle power
- for short idle periods TPM may expend more energy than is saved
- DRPM doesn't have long spin up delay and can service requests in a lower power (speed) mode (with longer rotation and transfer delays)
- DRPM allows adaptive tradeoff between power and performance

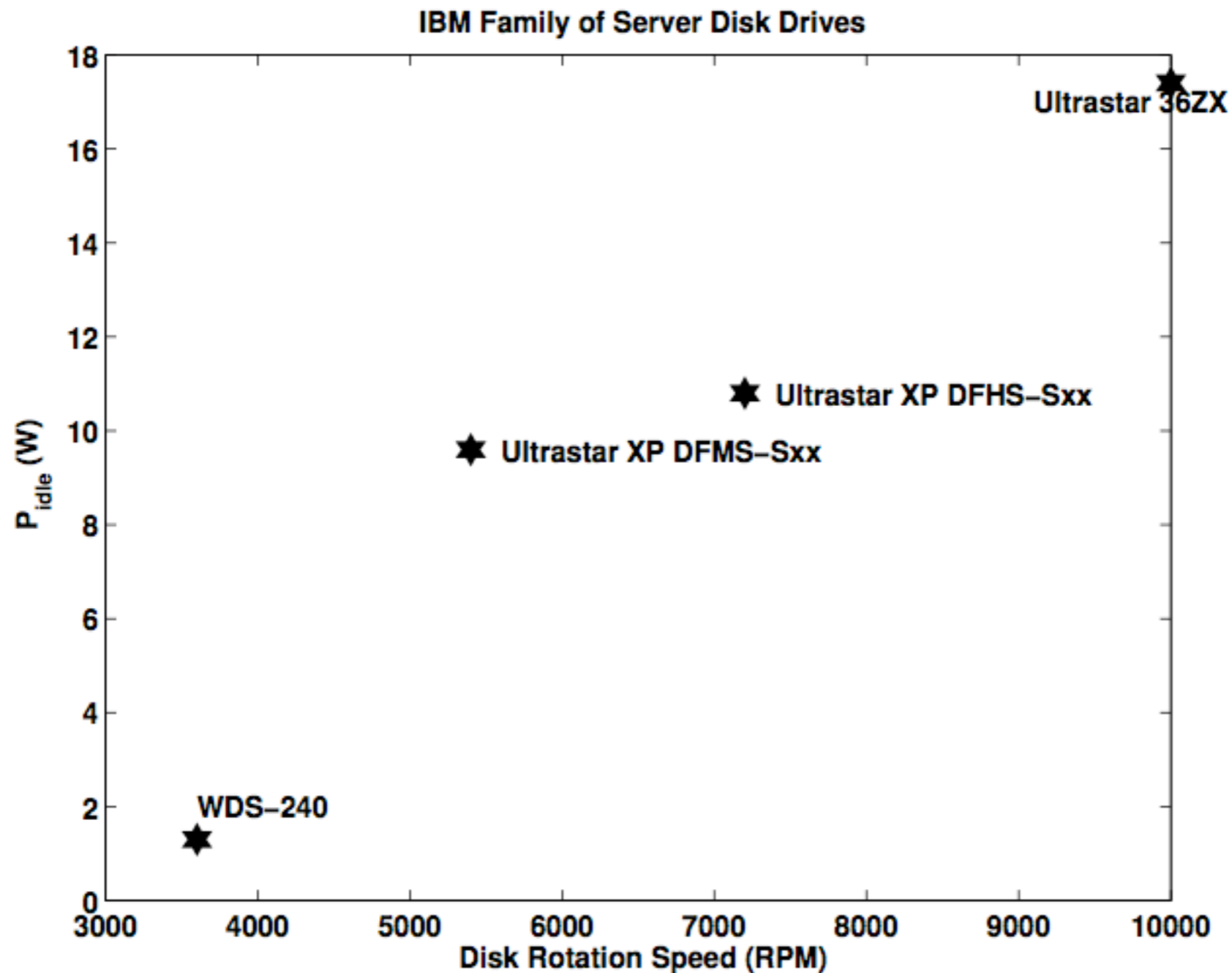
Spindle Motors

- permanent magnet DC brushless motor
- speed is achieved by 'pulsing' the motor power, it is not always on
- sensors provide speed feedback
- high torque needed to start platter spinning

Power Model

- 4 methods used to determine relationship between RPMs and power
- plot of multi-generational devices
- physics
- Sony Multimode Hard-disk
 - static configuration
 - not enterprise - 2 rather than 10 platters
- IBM paper on disk design

Multi-generation IBM Disks



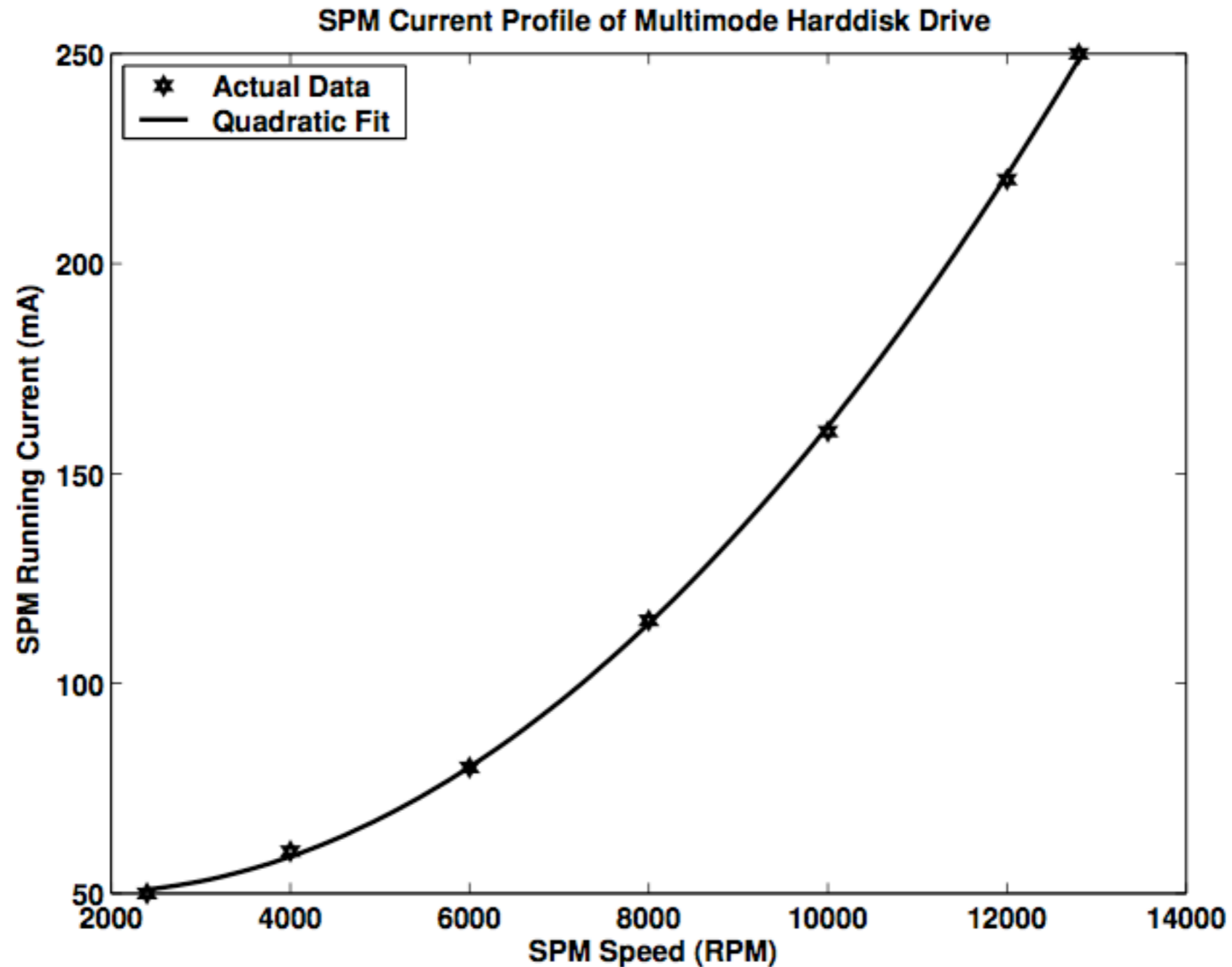
Physics

$$V = K_E \omega$$

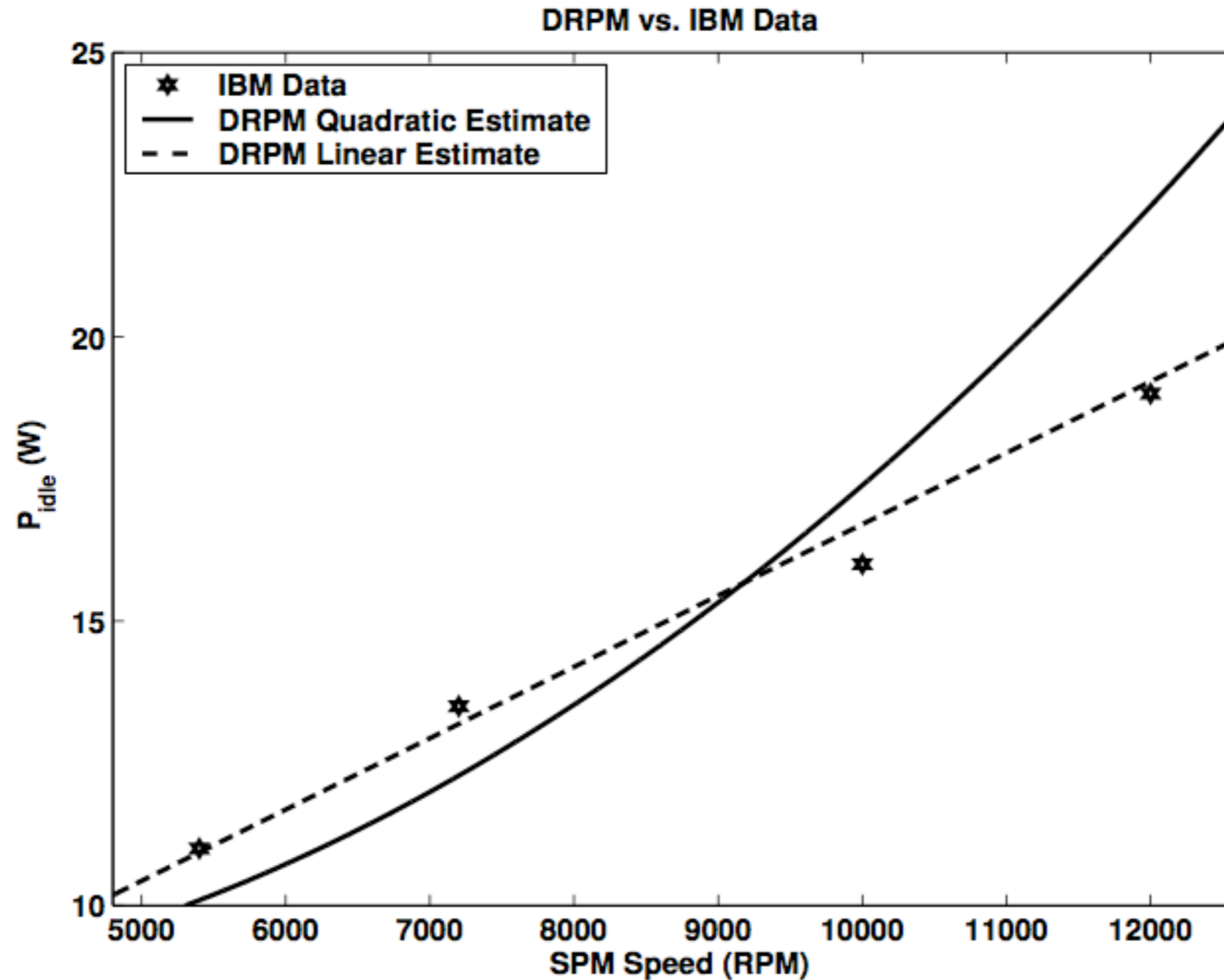
$$P = VI = \frac{V^2}{R}$$

$$P = \frac{K_E^2 \omega^2}{R}$$

Sony Multimode Disk



IBM Disk Design Study



Transition Times

Parameter	Value	Units
Max. Permissible Speed	15000	rpm
Rotor Inertia (J_0)	3.84	gcm^2
Torque Constant (K_T)	9.1	mNm/A
Max. Continuous Current at 12K rpm (I)	0.708	A

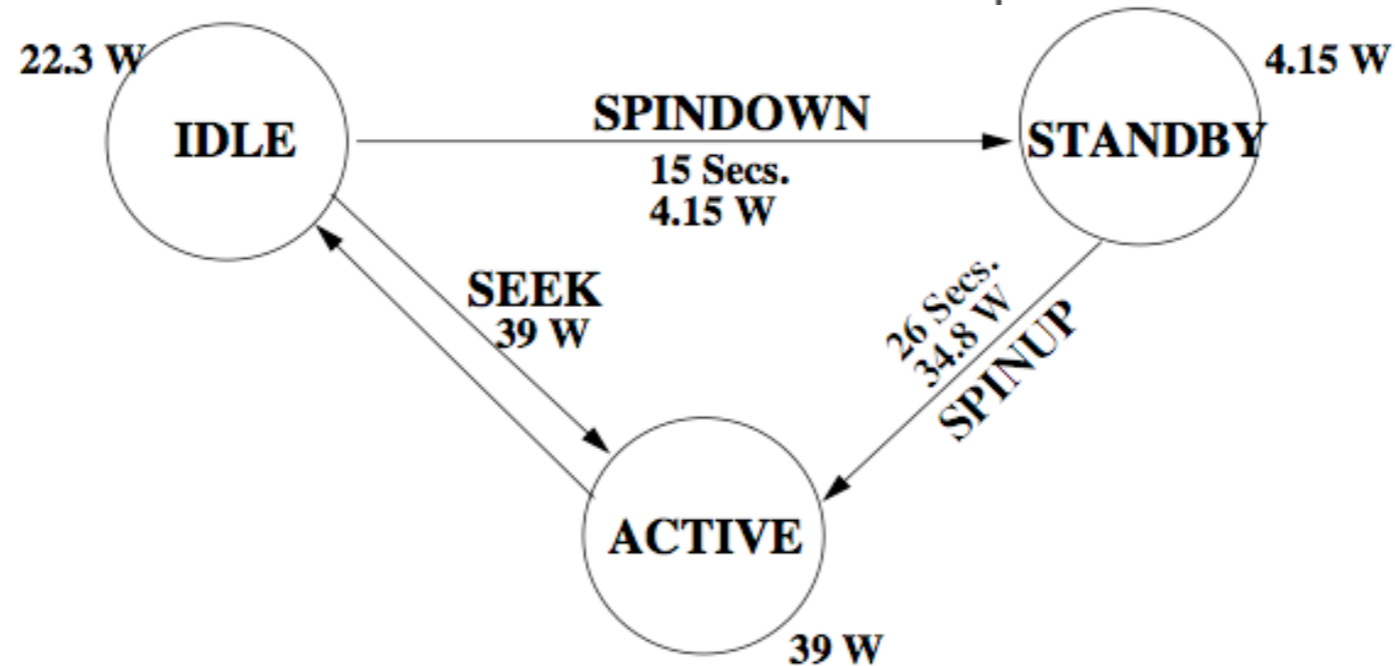
$$\Delta t = \frac{\pi}{300} \Delta n \frac{J_0 + J_L}{K_T I}$$

$$\Delta t = 2.693 \times 10^{-4} \Delta n$$

Parameter	Value
-----------	-------

Parameters Common to TPM and DRPM

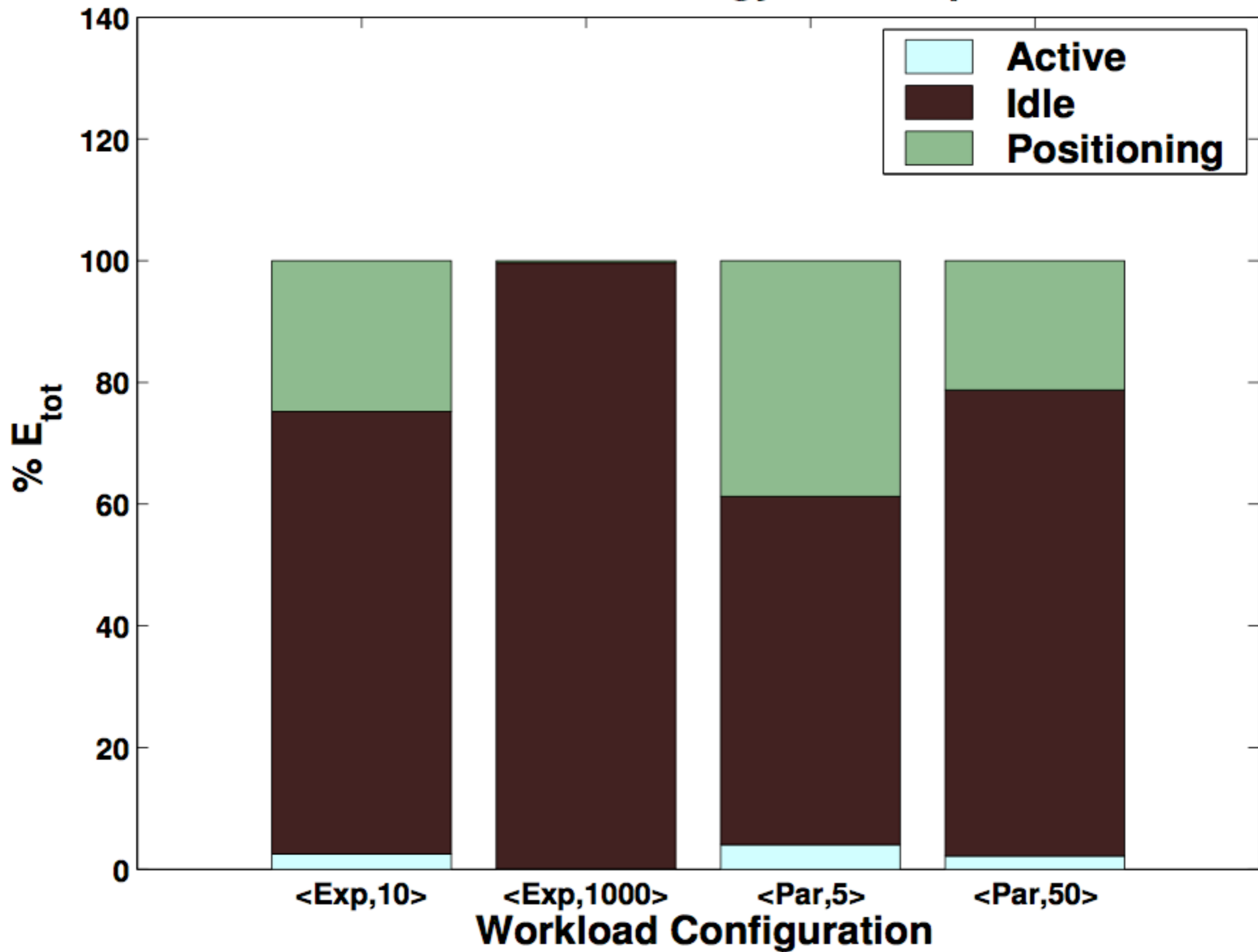
Number of Disks in the Array	<u>12,24</u>
Stripe Size	16 KB
RAID Level	<u>5,10</u>
Individual Disk Capacity	33.6 GB
Disk Cache Size	4 MB
Max. Disk Rotation Speed	12000 RPM



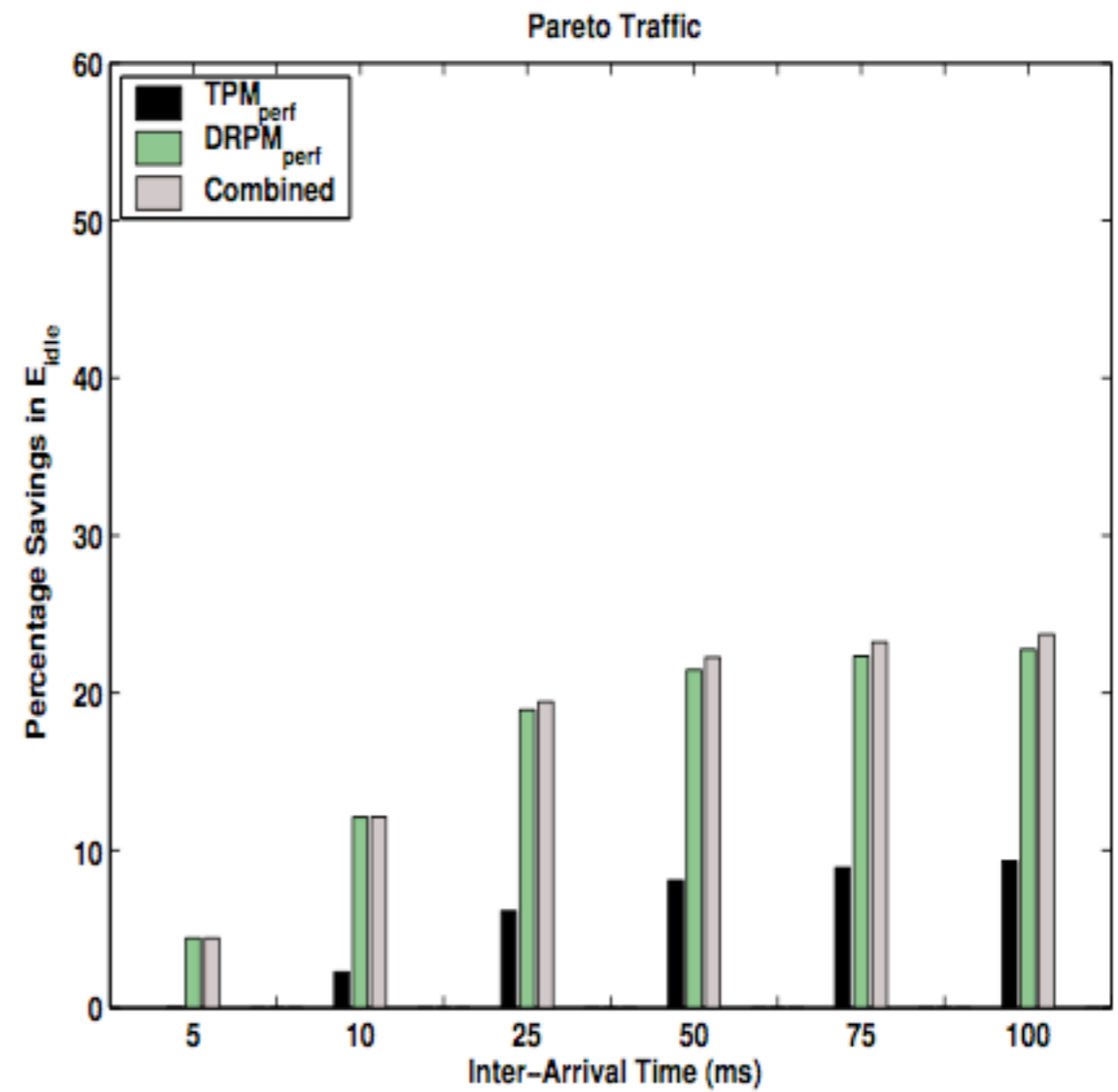
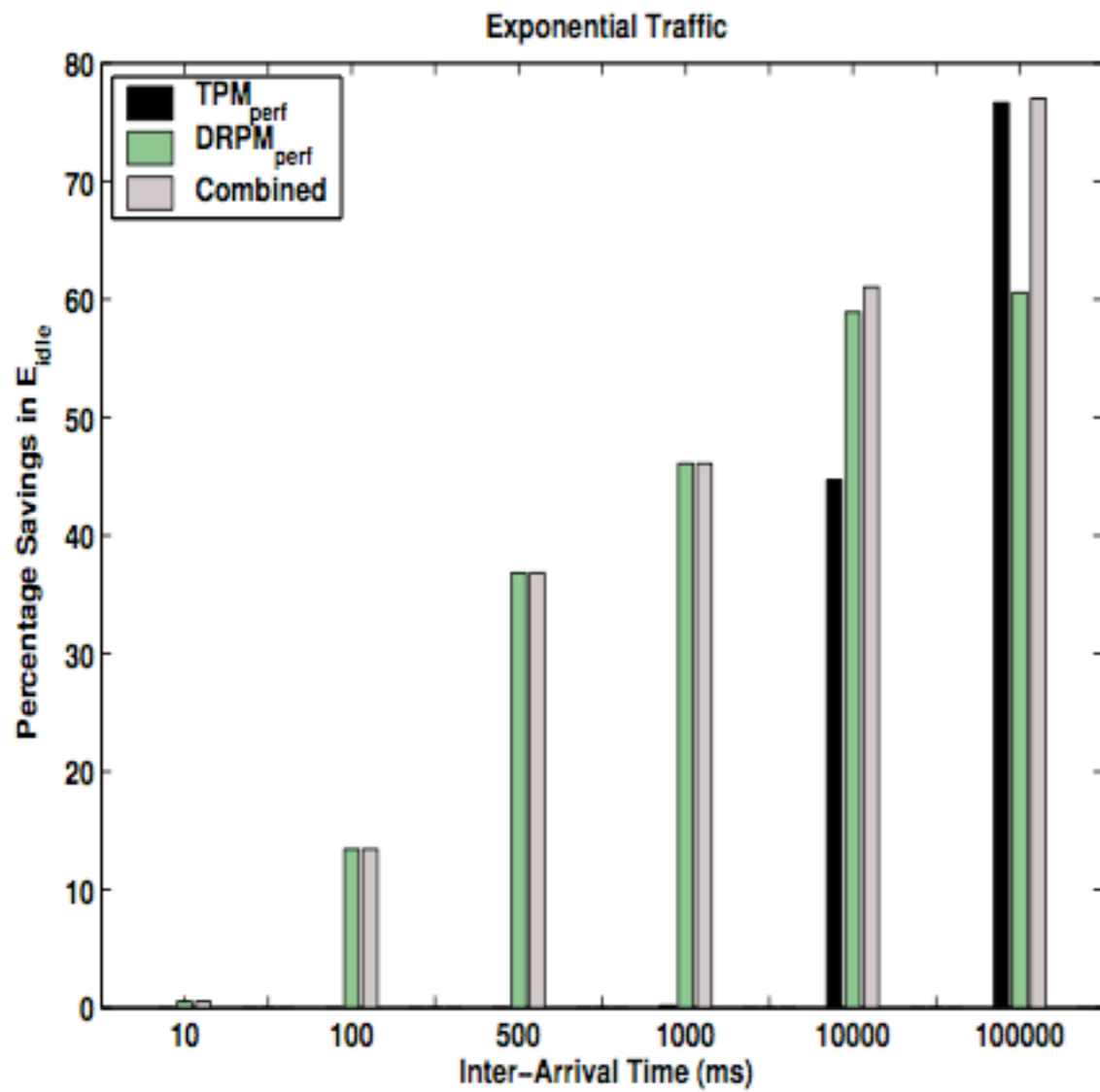
DRPM-Specific Parameters

Power Model Type	<u>Quadratic,Linear</u>
Minimum Disk Rotation Speed	3600 RPM
RPM Step-Size	<u>600,2100</u> RPM

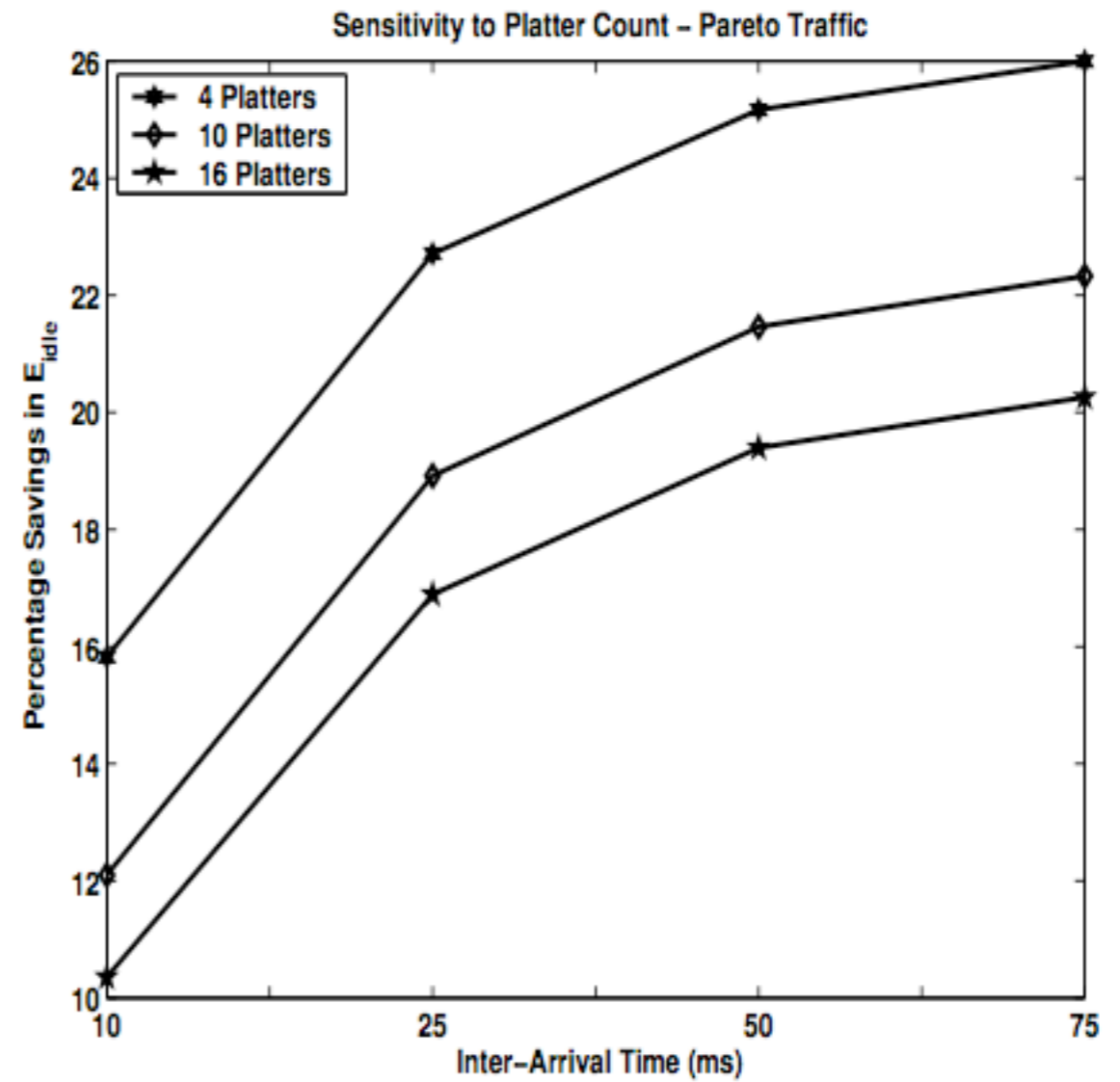
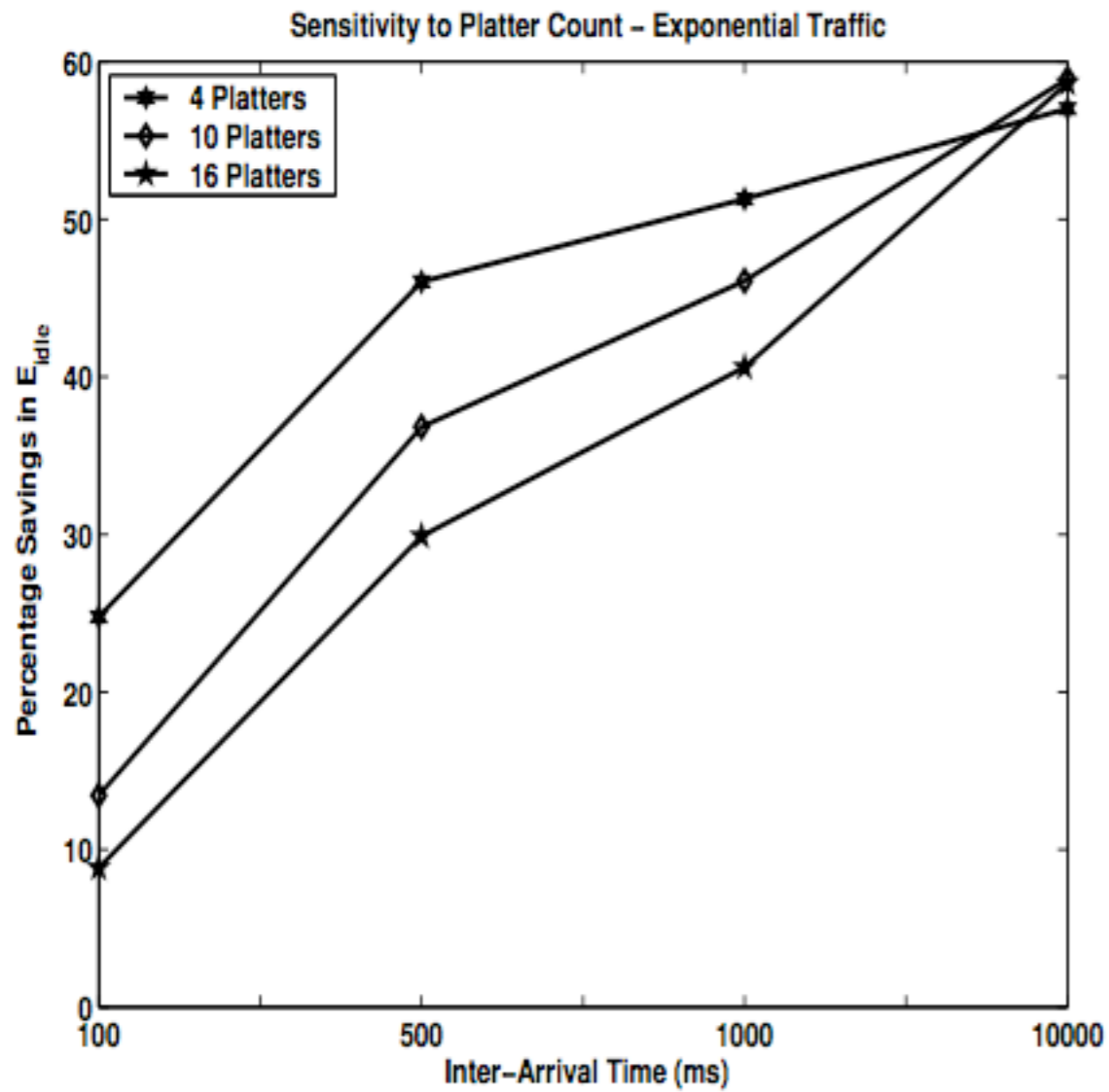
Breakdown of Energy Consumption



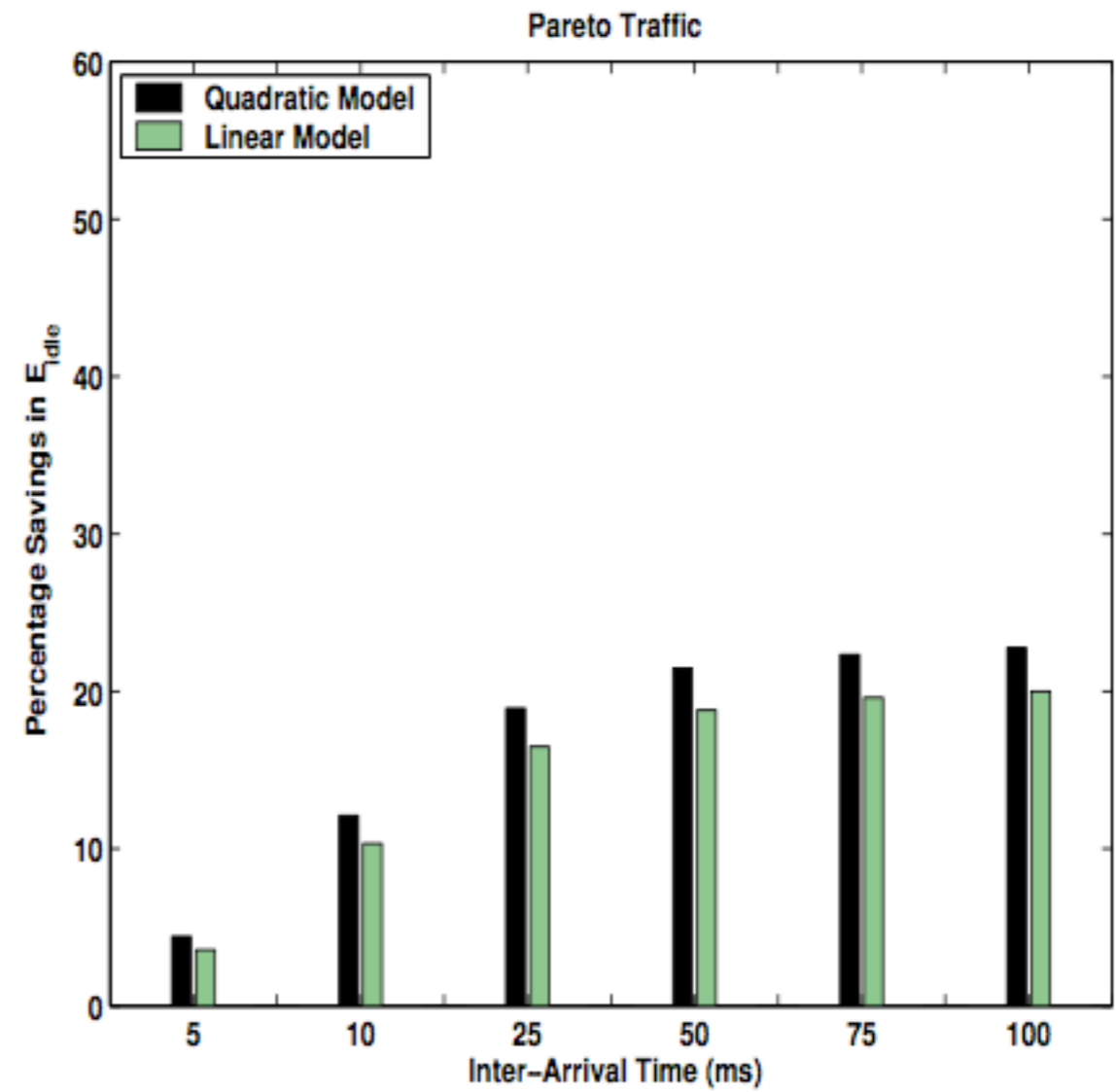
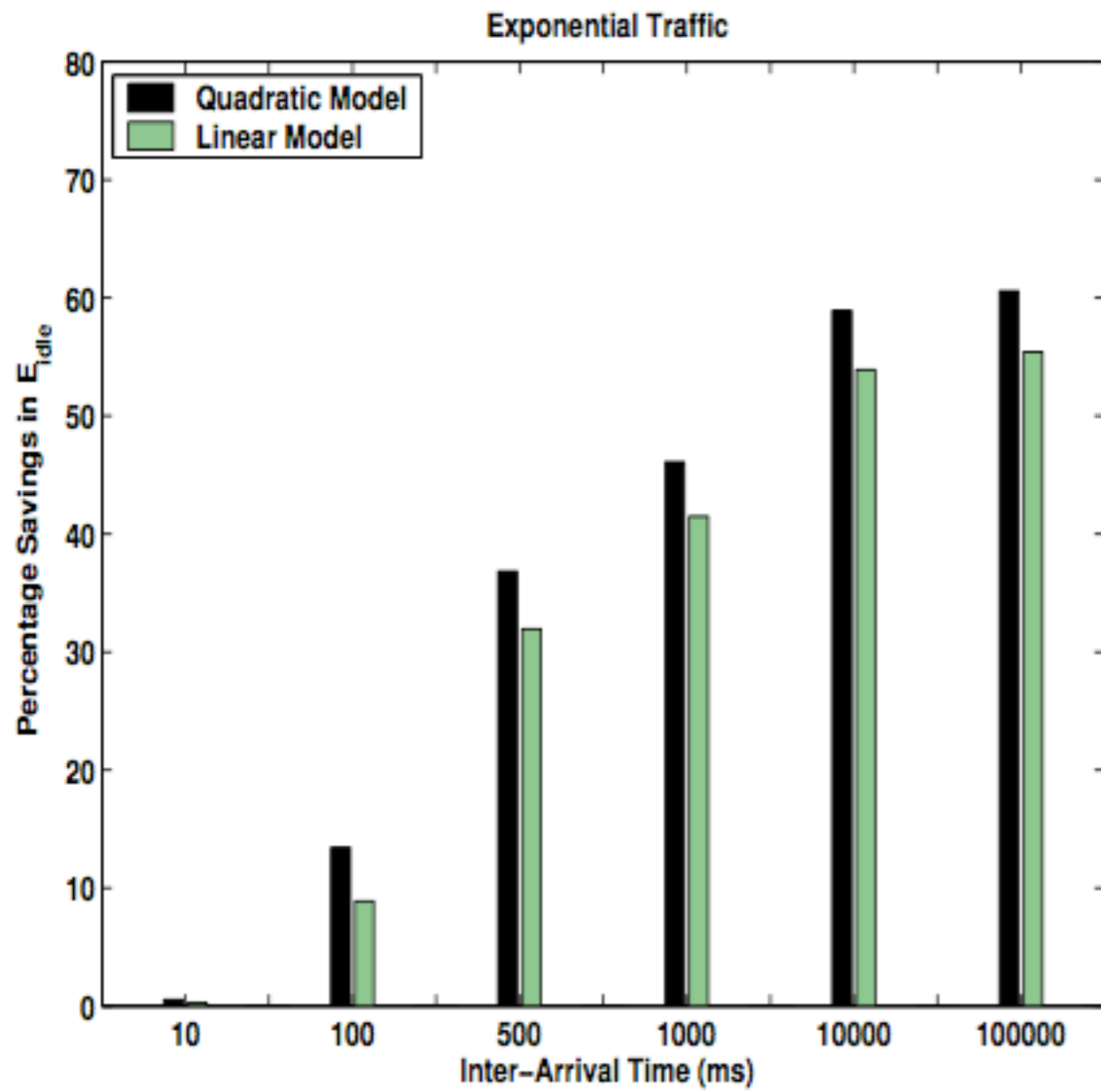
Oracle Comparison



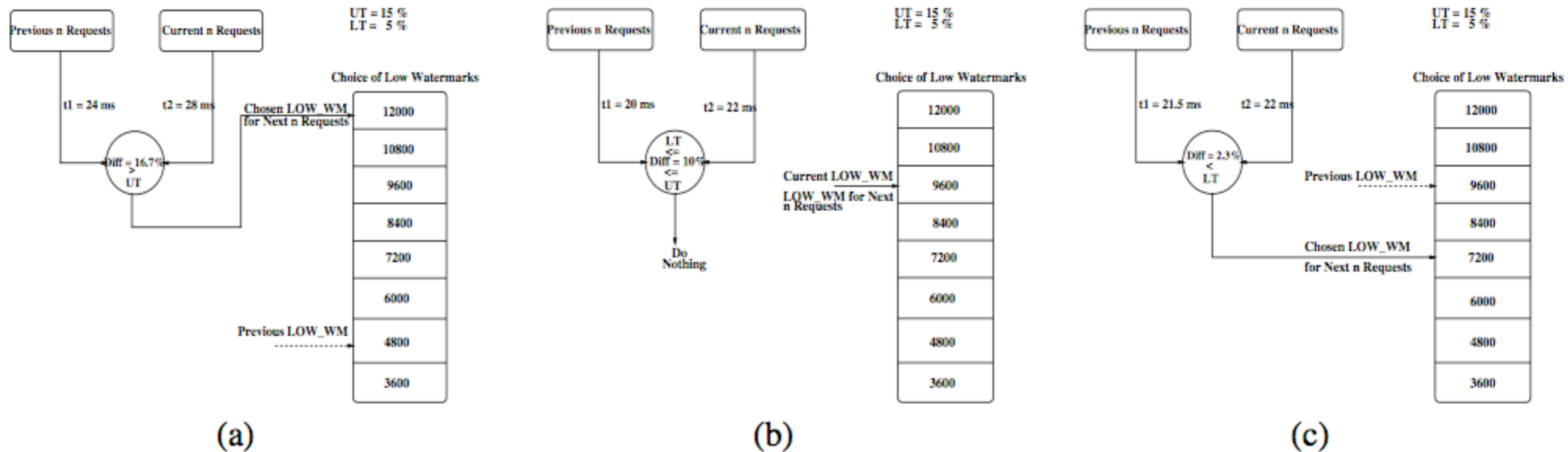
Platter Count



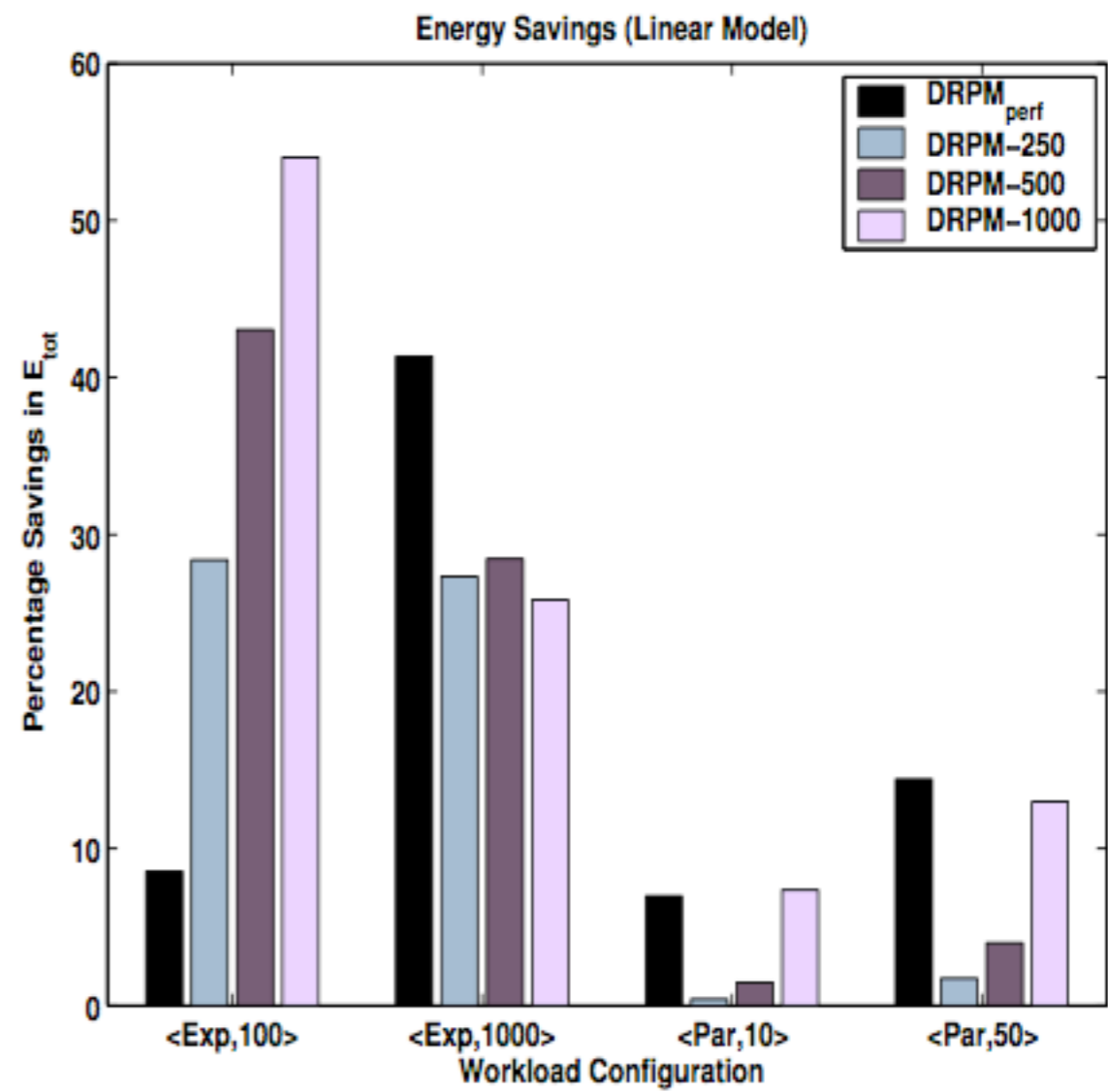
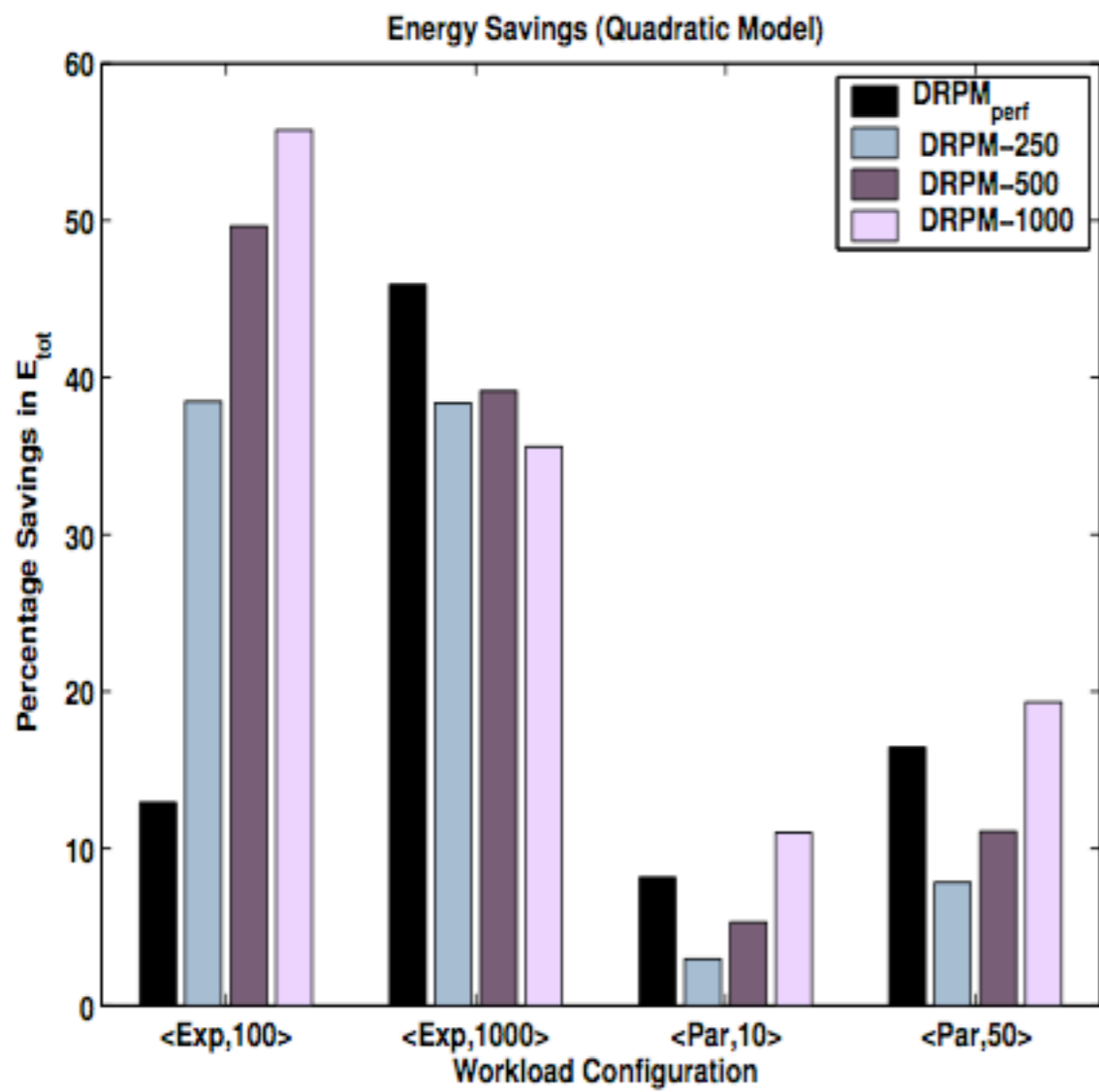
Power Model



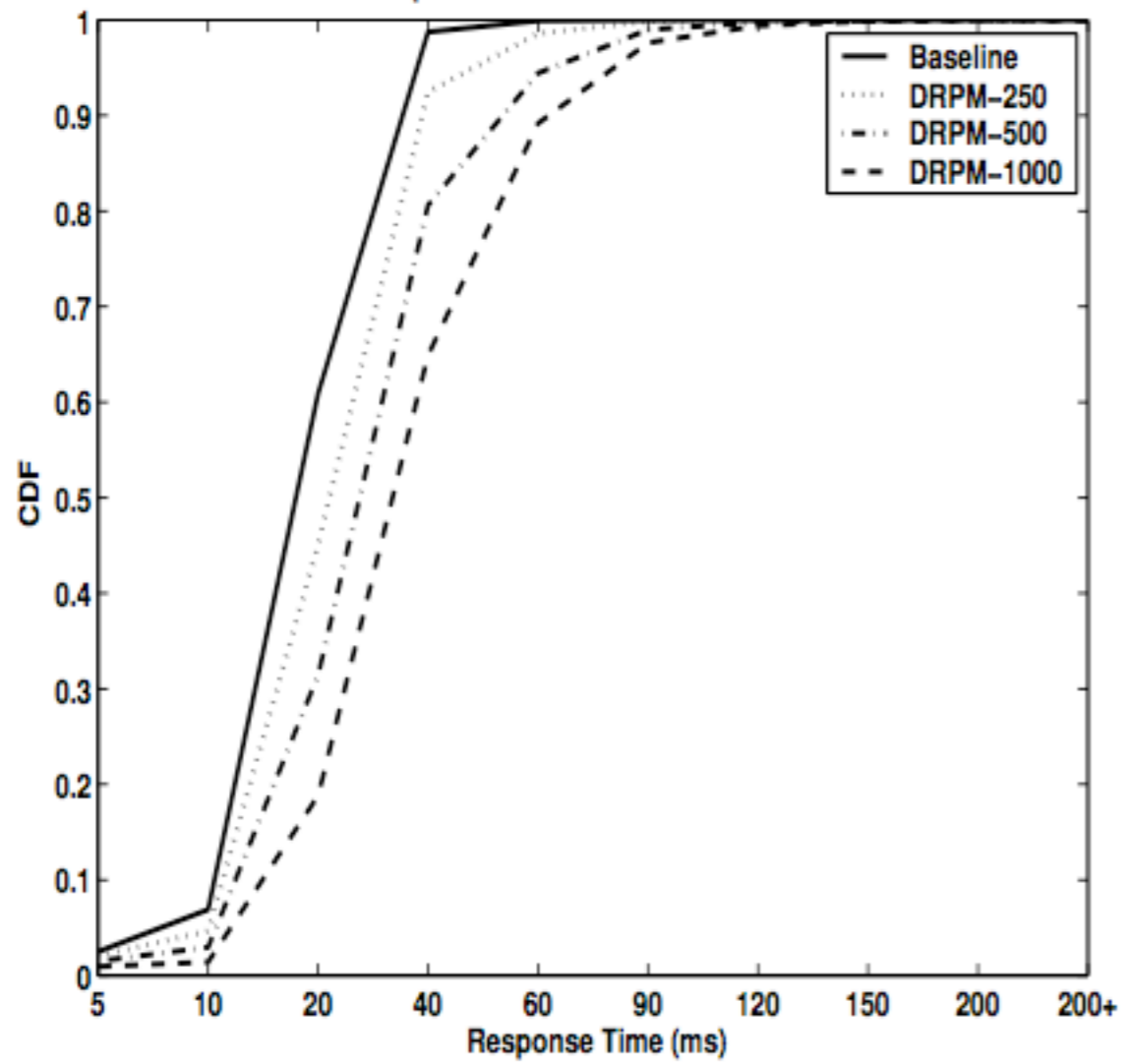
DRPM Heuristic



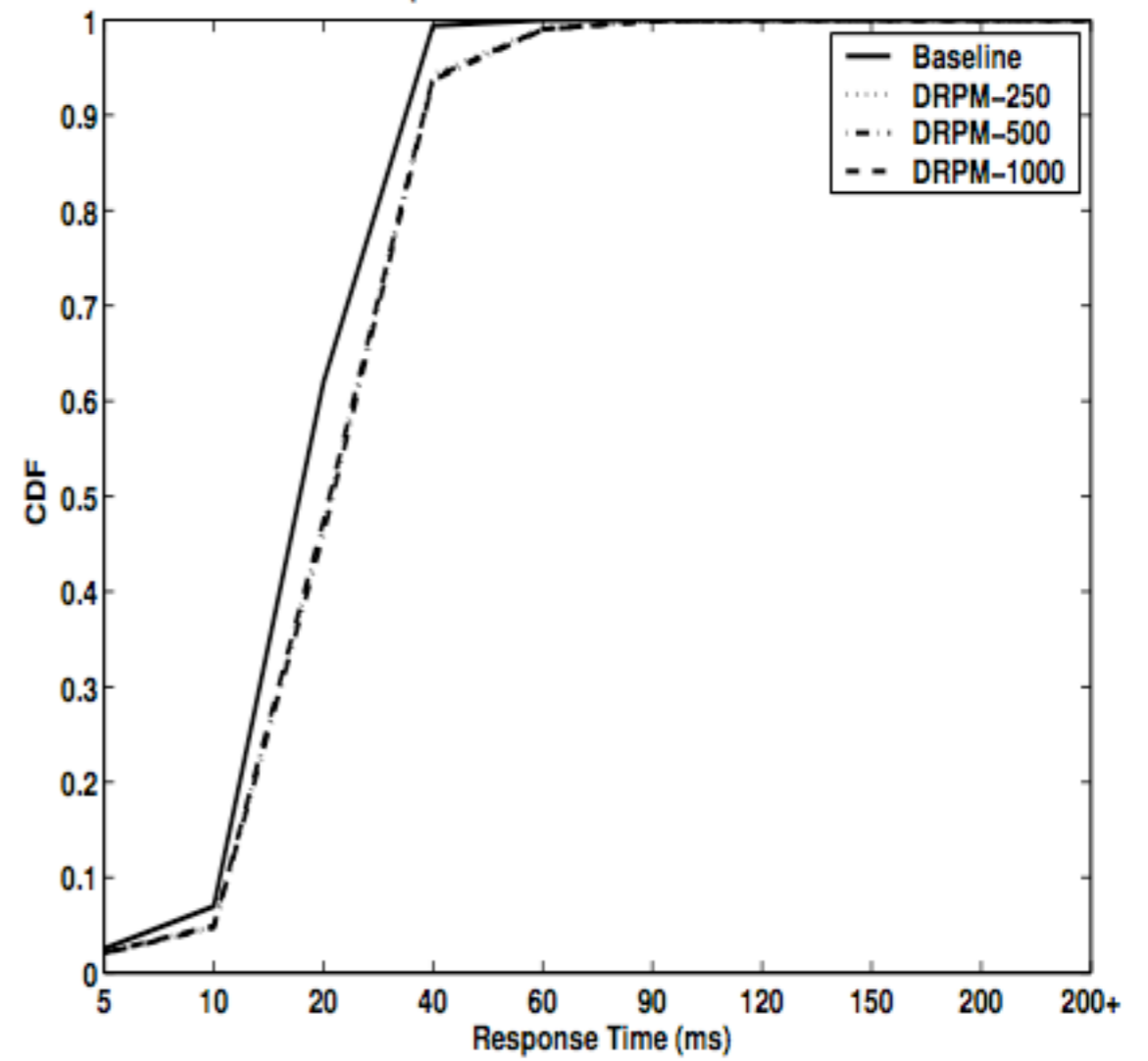
- watermark is set by array
- periodically, disk checks its queue length, if 0, moves to next lowest state unless at watermark



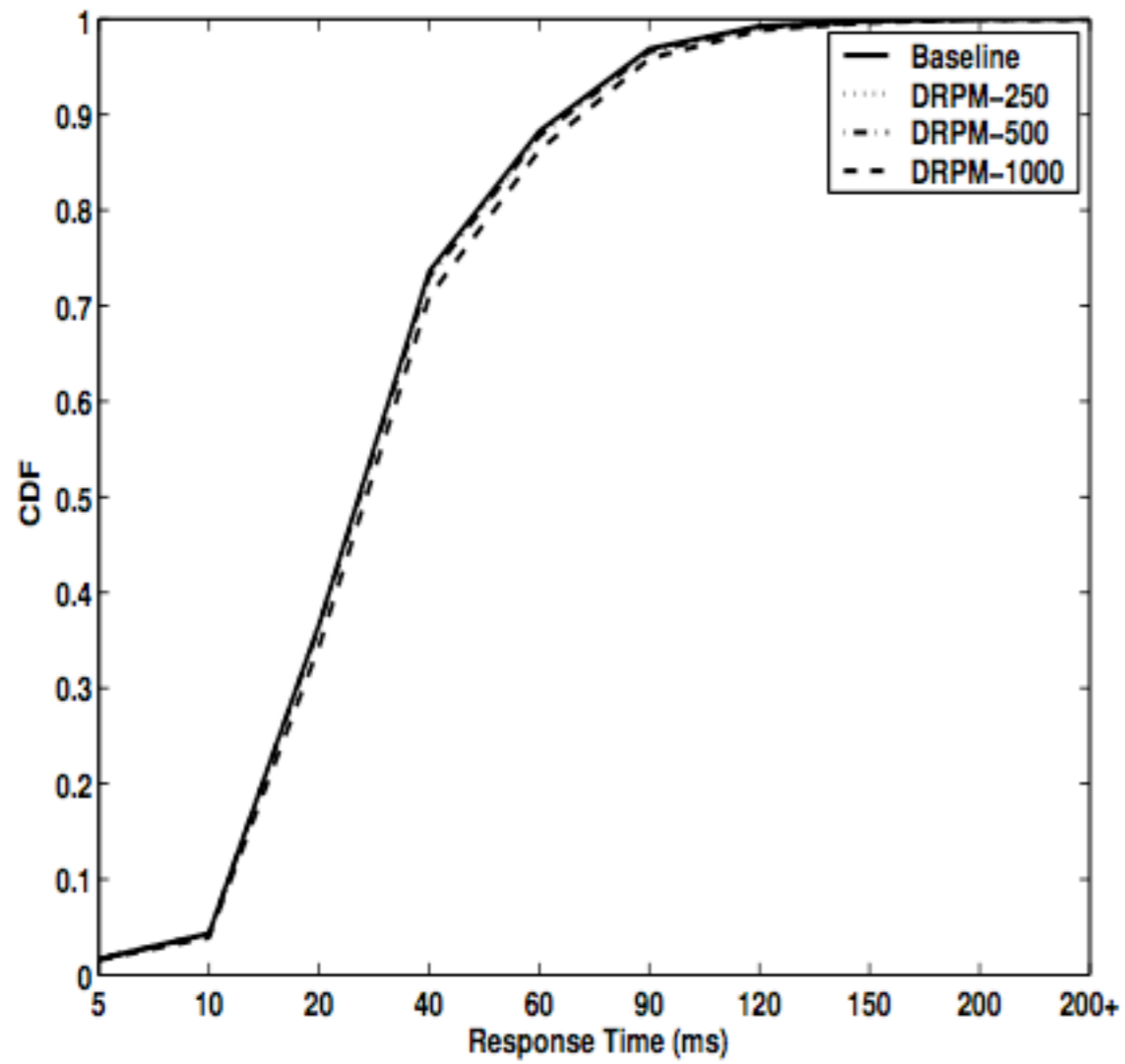
Exponential Traffic - 100 ms Mean



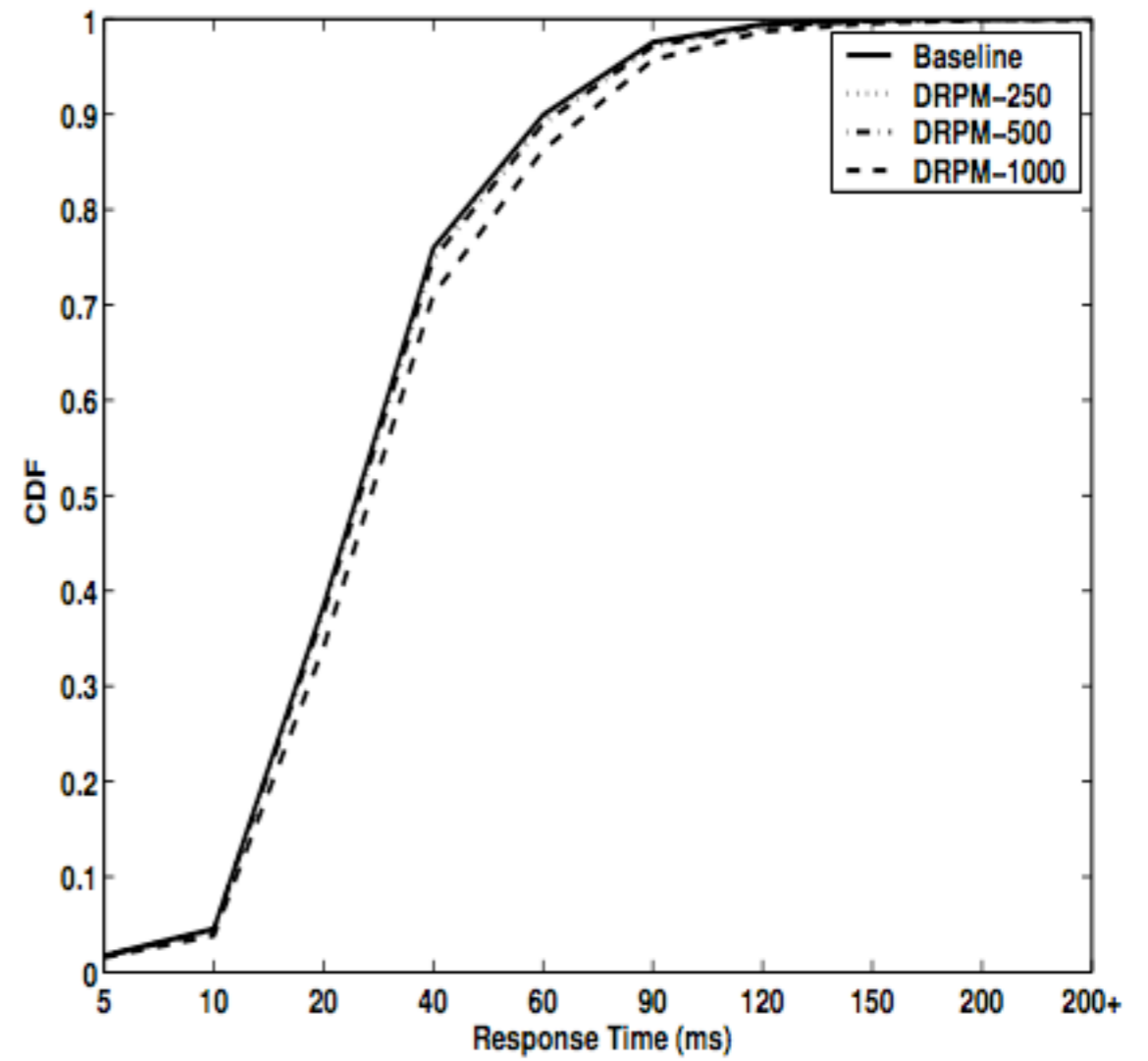
Exponential Traffic - 1000 ms Mean



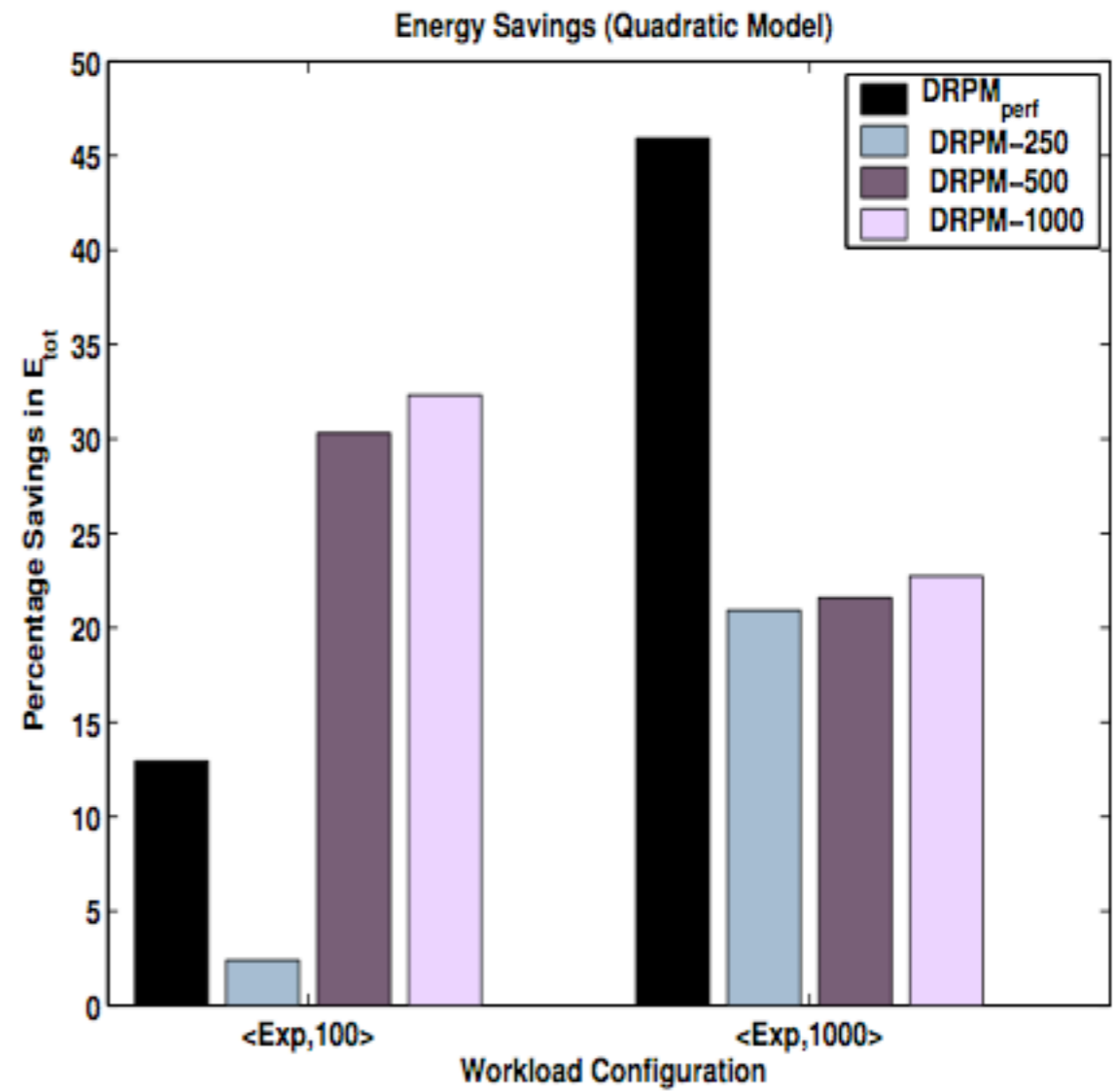
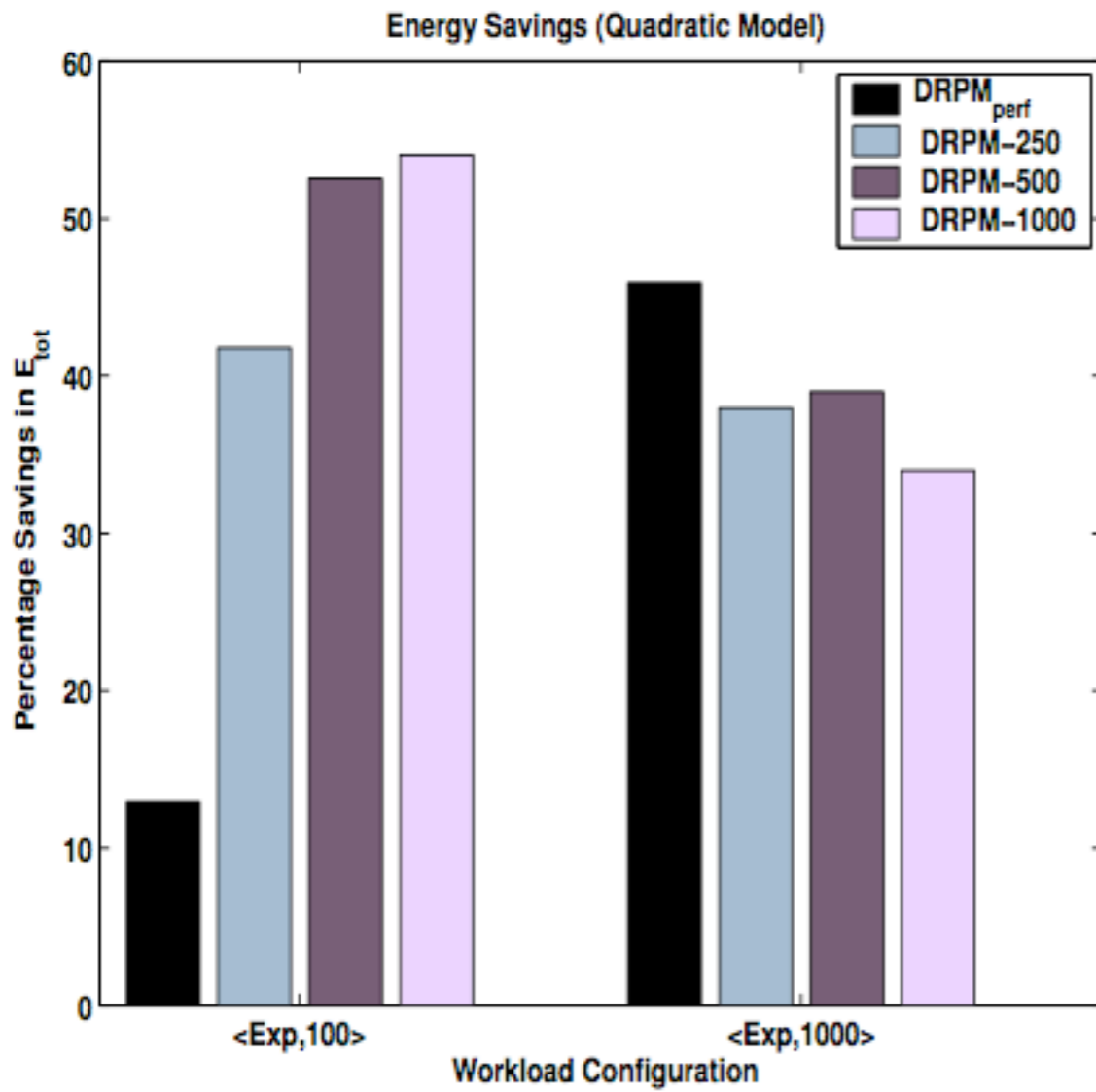
Pareto Traffic - 10 ms Mean



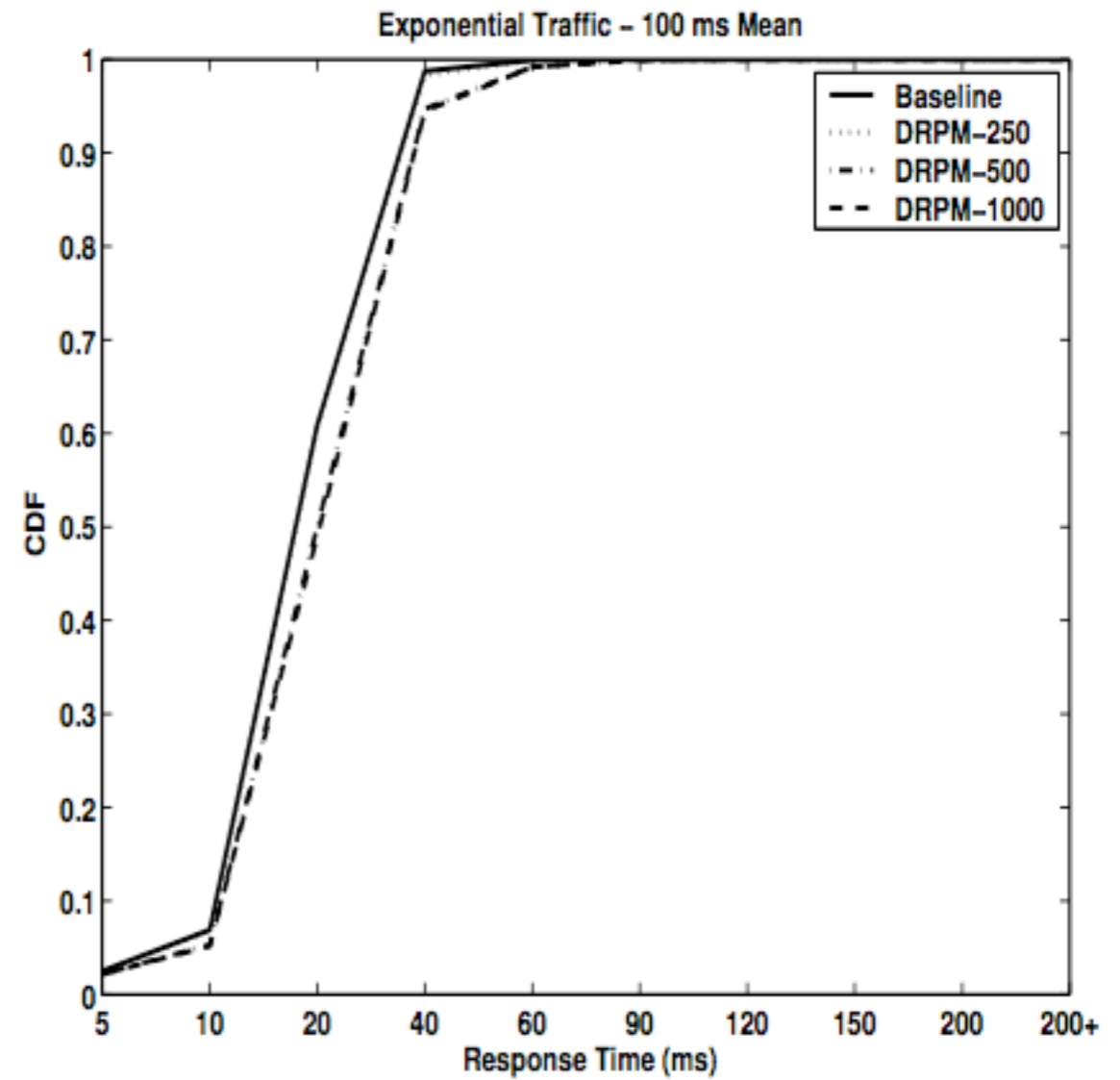
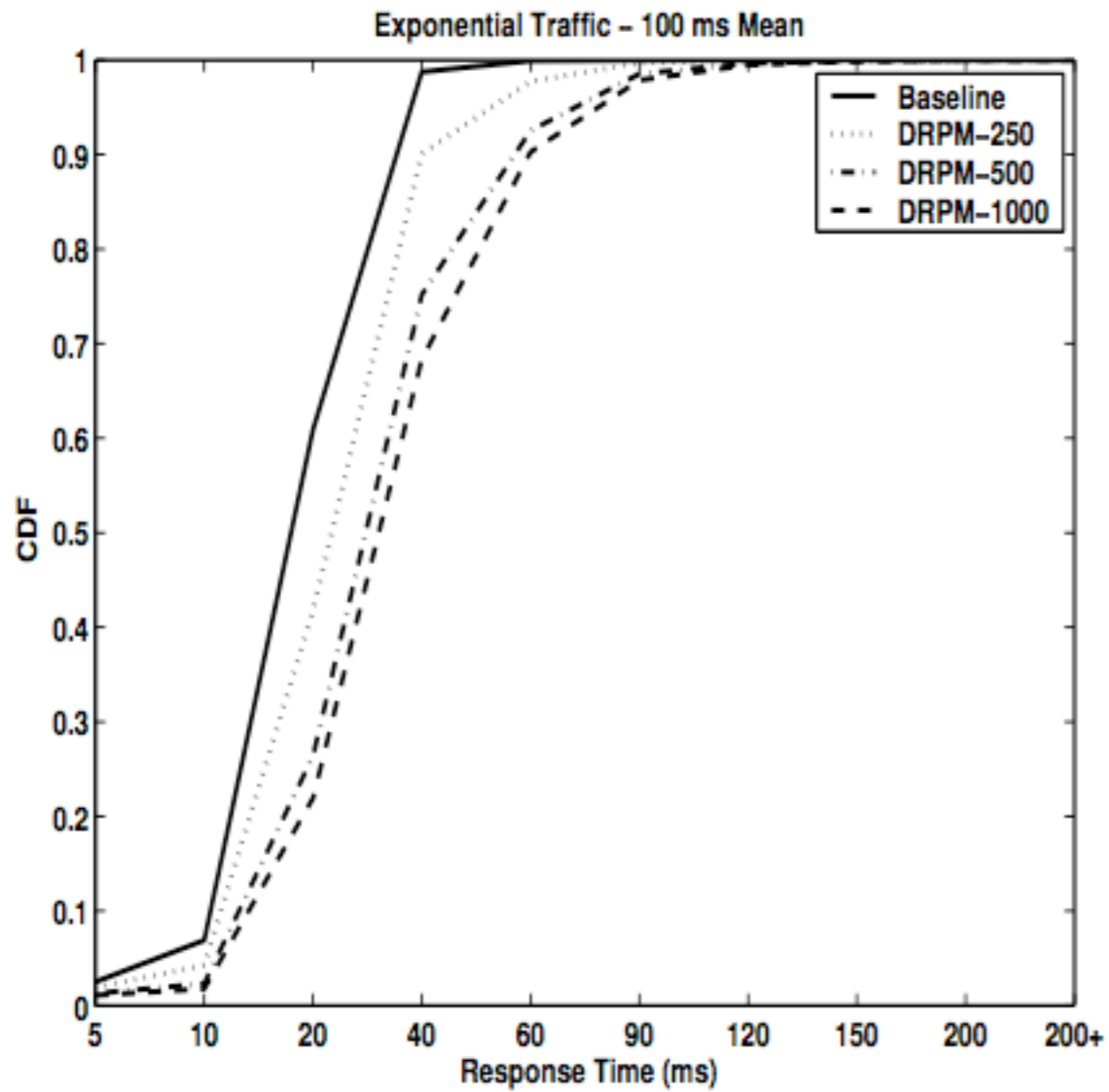
Pareto Traffic - 50 ms Mean



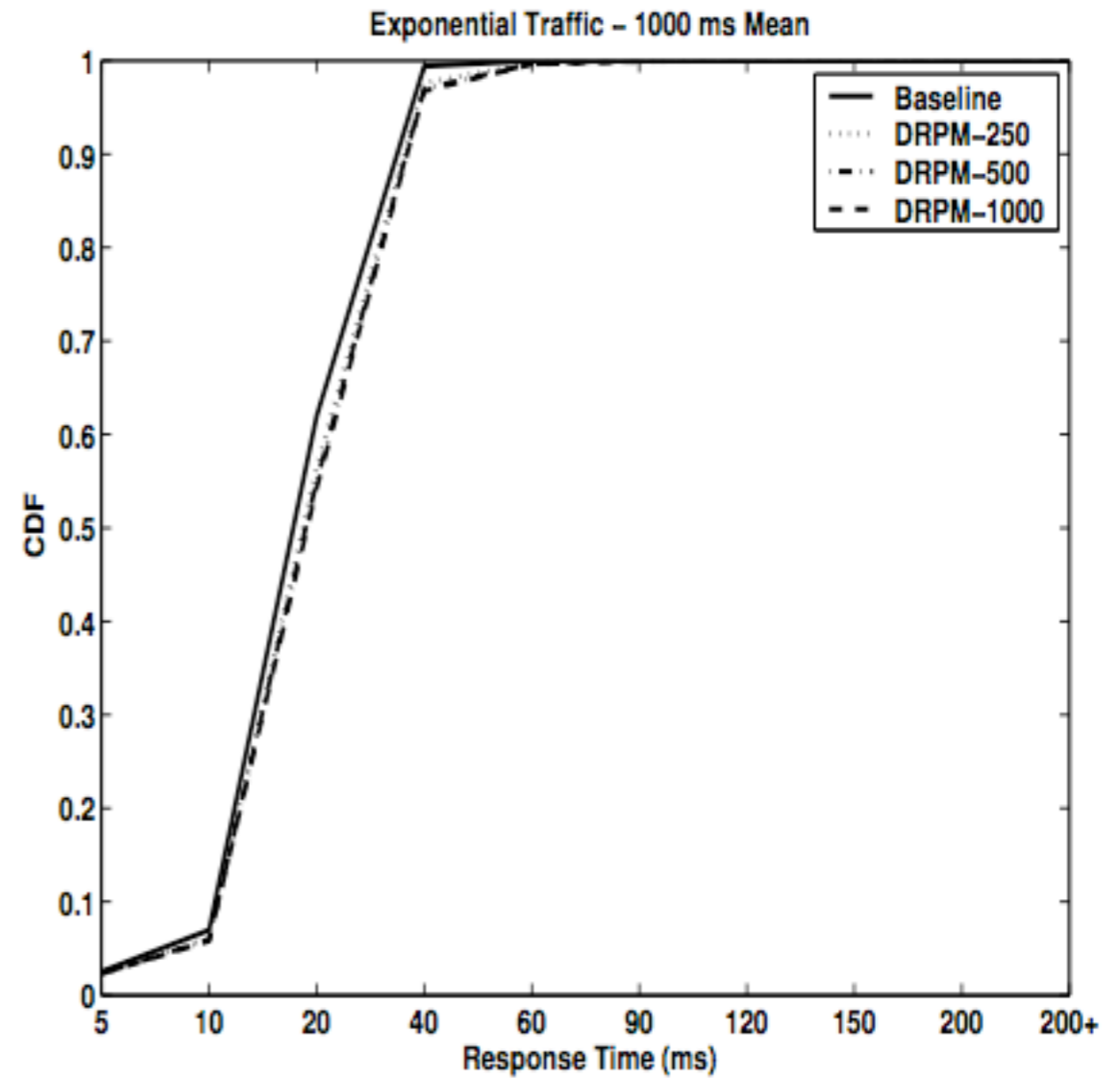
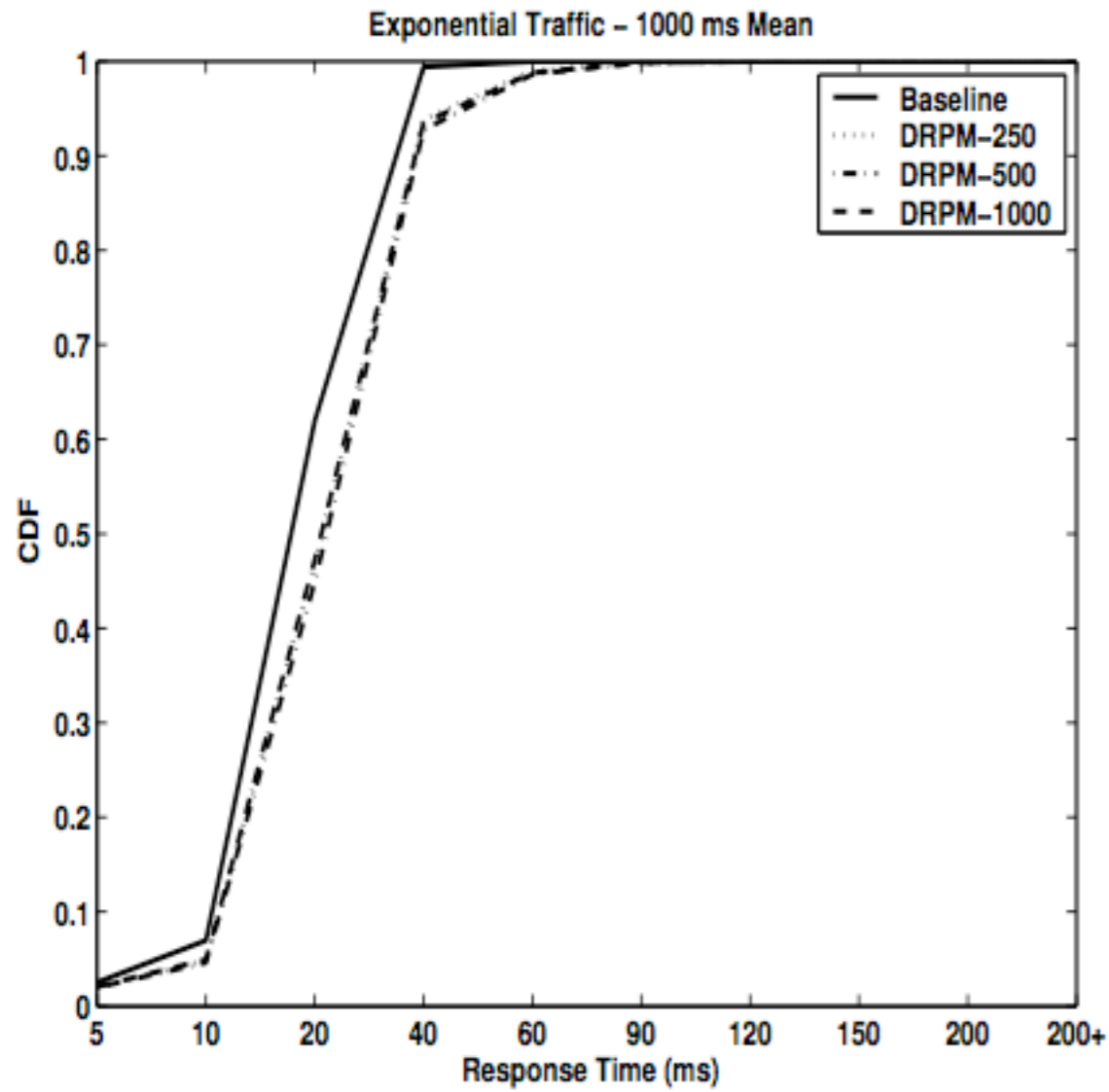
10-15% vs 5-8%



10-15% vs 5-8%



10-15% vs 5-8%



Issues

- Providing Speed Control - Pulse Width Modulation
- Fly-Height - Papillon slider is constant over DRPM range
- Head Positioning - servo sectors: may not see enough at low speed, but this is solved
- Idle-Time activities - DRPM good for QoS
- Reliability - DRPM may reduce? though avoids spin down duty cycling

Conclusion

- DRPM recovery is faster and takes less energy than spinup
- able to exploit short idle periods
- can provide a reduced service level
- orthogonal to spindown
- simple heuristic can save more power than oracle (DRPM or spindown)!