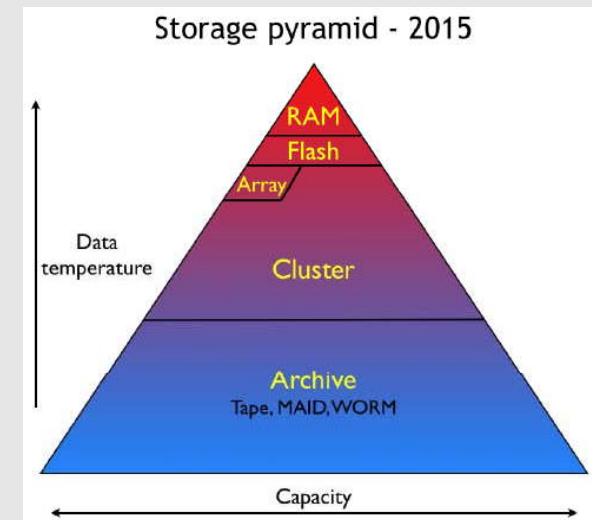
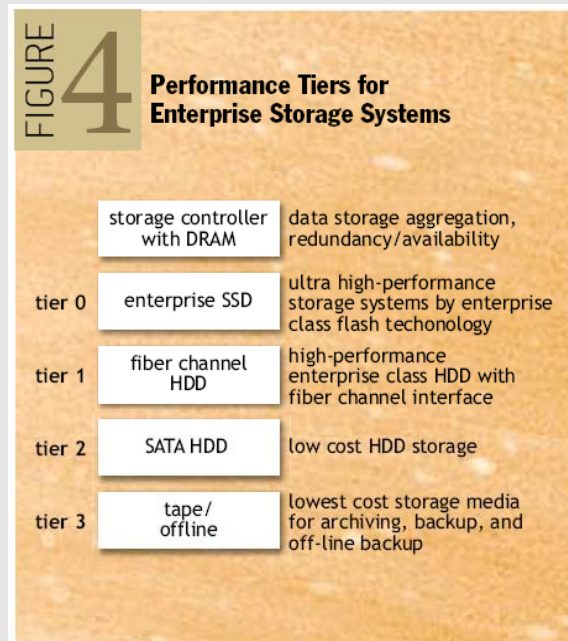
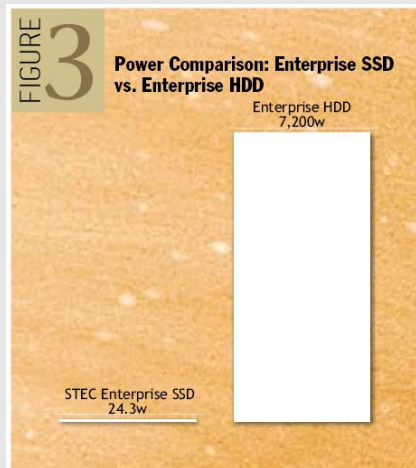
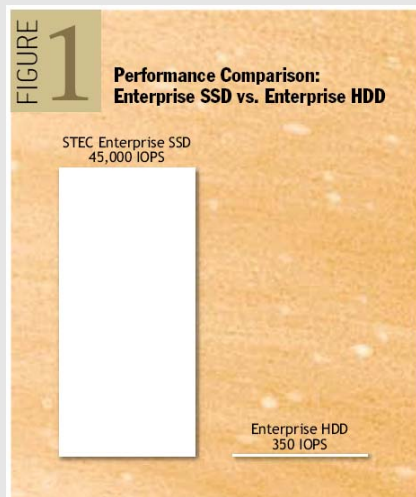


Let's make our own SSD Benchmark

2009. 4. 23.
Jaehyuk Cha, Hanyang Univ.

Advent of SSD

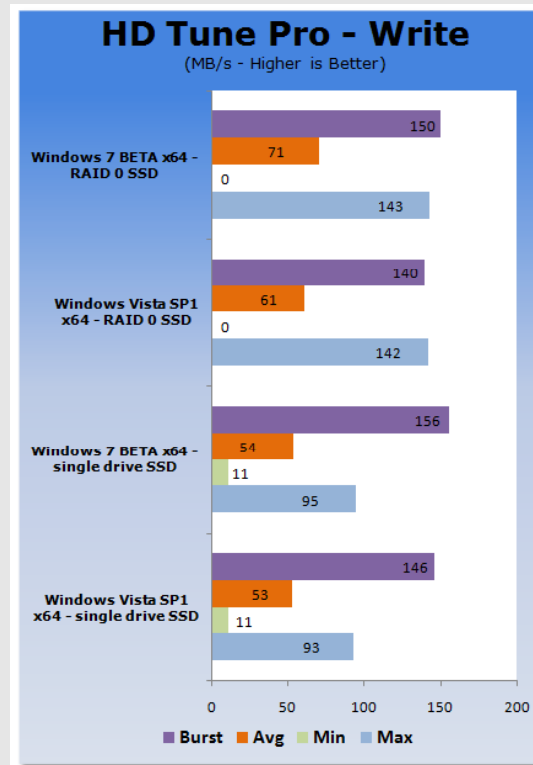
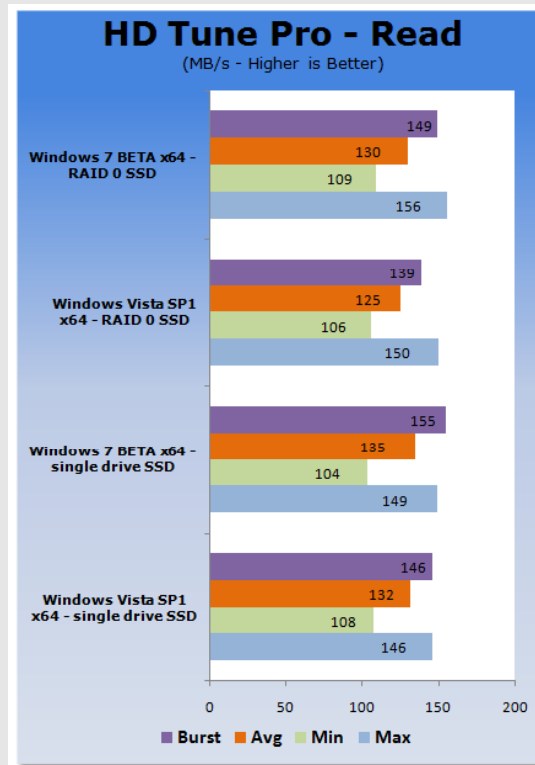
- Tape is dead, disk is tape, flash is disk, RAM locality is king (Jim Gray)
 - Is SSD ready for enterprise storage medium?
 - New storage pyramid?



- M Moshayedi (CTO of STEC), ACM QUEUE July/August 2008
- Robin Harris, <http://storagemojo.com>

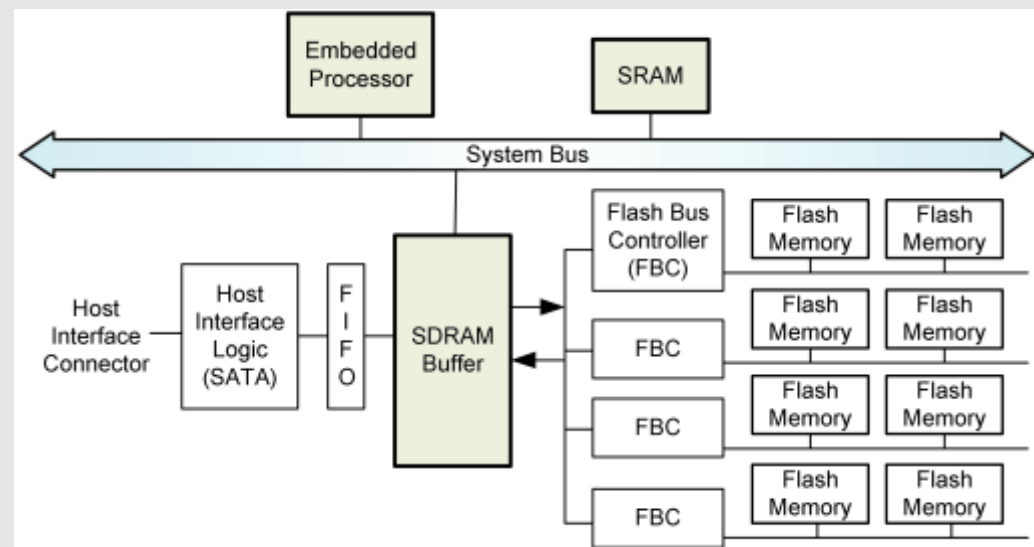
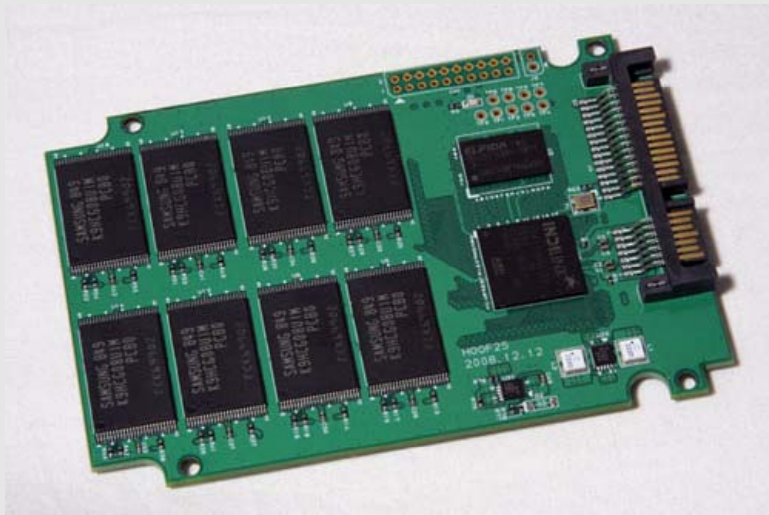
SSD is different!

- SSD is made of flash chips, but
 - is different from raw flash chips, and is similar to HDDs
 - Embedded controller hides the characteristics of raw flash chips
- SSD is a block device, but
 - is different from HDDs
 - Non-uniform performance of writes (and even reads!)



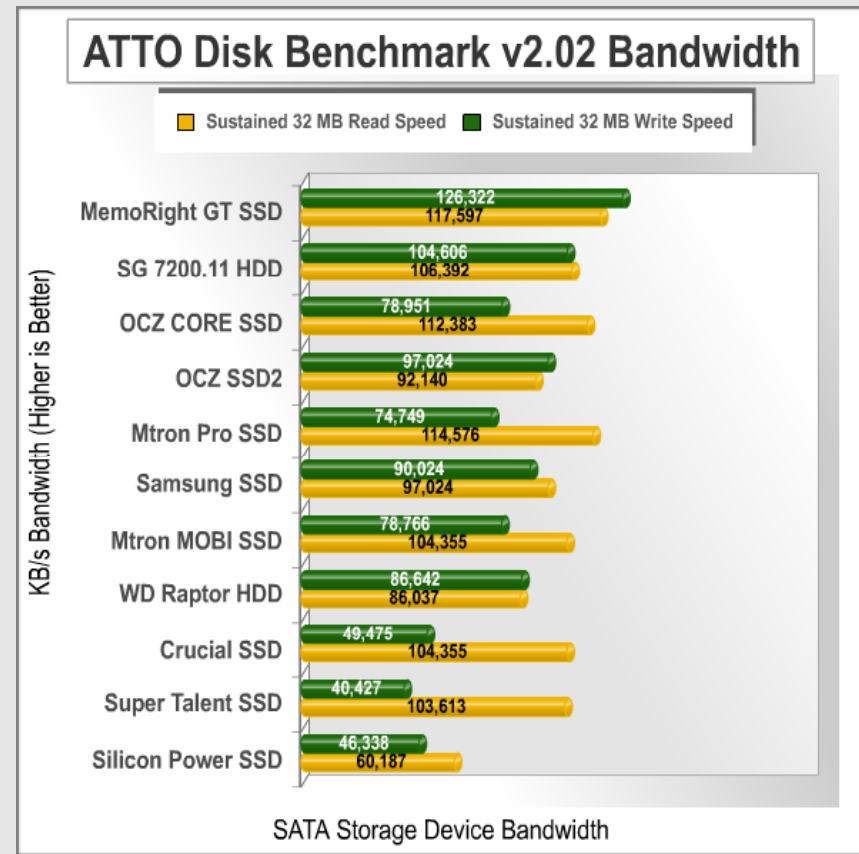
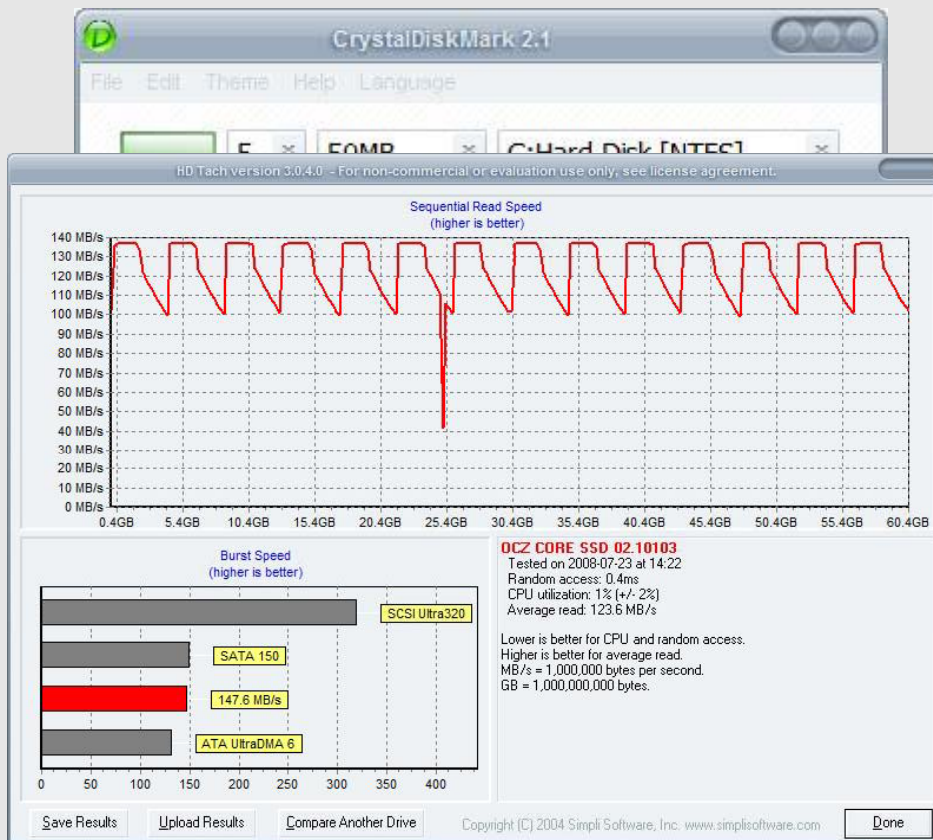
SSD internals

- Controller logics varies from vendor to vendor
 - Various FTL mapping algorithms
 - Various size and management policies of write buffer
 - Various degrees of parallelism
 - ...



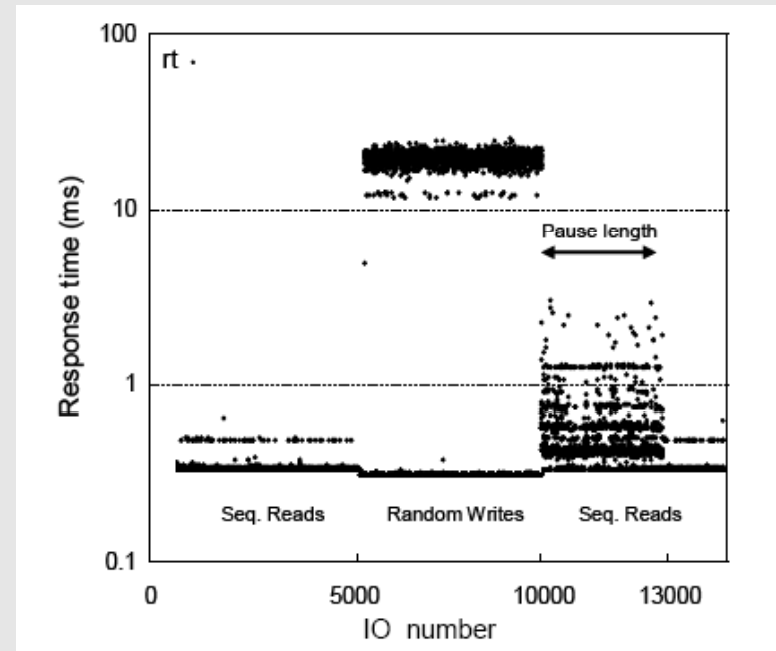
HDD benchmarks sufficient for SSDs? (1/2)

- HDD benchmarks
 - IO patterns: Sequential R/W, Random R/W, disk traces from real world
 - Performance metric: IOPS, bandwidth, CPU usage, access time, ...



HDD benchmarks sufficient for SSDs? (2/2)

- We need
 - Various I/O patterns that consider other than sequentiality
 - LBA interval
 - Physical-block alignment
 - Wear-out problem
 - High standard deviations on write throughput
 - Disk traces containing the TRIM operations
 - ...
 - Factors other than I/O patterns
 - Current status of the device
Ex) Intel's Slowing down problem
with extensive use
 - More performance metrics
 - Power consumption aspect
 - Endurance aspect



High-level Benchmarks sufficient for SSDs?

- Are file system benchmarks sufficient for SSDs?
 - File operations should include “file deletion”
 - What else performance metrics?
- Are other system benchmarks sufficient for SSDs?

Existing Benchmarks for SSD

- Previous efforts to open the blackbox
 - Benchmarks in database domain
 - SanDisk
 - uFLIP
 - Microsoft
 - Sun Microsystems

Previous Works – Database domain

- A Case Study for Flash Memory SSD in Enterprise Database Applications
 - SIGMOD'08, Lee et al.
 - Benchmarking SSD performance for typical DB access patterns
 - Measured only a single device
- On the Use of NAND Flash Memory in High-Performance Relational Databases
 - Master's thesis, MIT, 2008, Myers
 - Benchmarking SSD performance for DB workloads
 - Measured two devices

Previous Works - SanDisk

- vRPM (Virtual RPM)
 - “How fast would you have to spin a virtual HDD to achieve the level of performance seen by an SSD in a client PC?”
 - e.g.

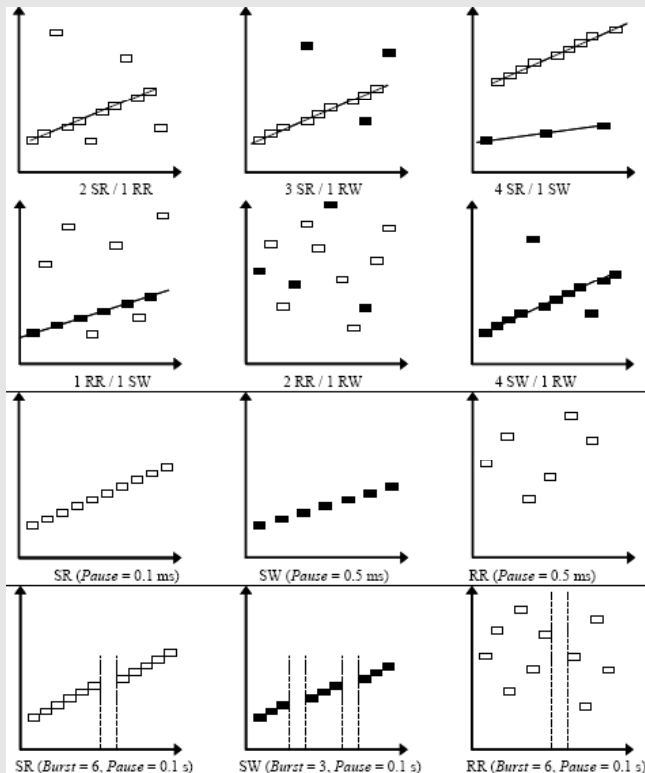
Table 4
Client PC SSD Performance in vRPM
(2009 rates are estimated)

SSD Generation	Read IOPS	Write IOPS	PC IOPS	vRPM
2006/7	5,000	10	20	1,000
2008	10,000	100	200	10,000
2009	20,000	400	785	40,000

- LDE (Longterm Data Endurance)
 - Endurance metric proposal for end customers
 - Methodology for testing and measuring LDE
 - e.g. Device with 40 TBW and the user writes 10 GB per day
 - LDE is not exhausted in the first 10 years

Previous Works - uFLIP

- uFLIP: Understanding Flash IO Patterns
 - CIDR'09, Bouganim et al.
 - Proposed a set of microbenchmarks and its methodology
 - Microbenchmark according to various IO patterns
 - Applied the benchmark to eleven flash drives
 - Opened everything related to their works



uFLIP, Understanding Flash IO patterns

[Home](#) | [Abstract](#) | [People](#) | [Benchmark](#) | [Results](#) | [Software](#) | [Publications](#) | [Links](#) | [Bibliography](#) | [Contact](#)

uFLIP Results

These results are complete for 9 devices.

We did not publish the results of the G-SKILL device since these were identical to the SAMSUNG's one.

Two results are missing. The Transcend SLC device (will be done very quickly) and the FusionIO Drive (in a longer term).

The excel files are very useful to analyze the results. The first folder are graphs (named Gxy where x is the micro benchmark ID and y a number from 1 to 6 corresponding to the experiment). These folders take their value in another folders named Axy.

We hope to publish quickly a set of macros to browse in an easier way all these results.

Thanks for your patience !

[Download results](#)

[View some photography of tested devices](#)

Previous Works – Sun Microsystems

- Sun Flash Analyzer
 - Collects, analyzes, and displays data from systems
 - Finds SSD candidate device

Swat SSD Configurator

File Help

Import C:\swat302\monitor_data\sample_data\sample_data.bin.gz

SSD Configuration: not saved

Clear Load Save Save as Set lun sizes Data subset Show SSDs only

Total gigabytes: 0.00

IOPS per terabyte: 0.00

Total IOPS: 0.00

Average read%: 0.00

Average transfer size: 0

Average response time: 0.00

SSD candidate search

Find candidates using filters below

IOPS/TB >= 100

and Read% >= 0

and Response time >= 10

and Queue depth >= 0

'Find SSD candidates' scans all devices that have been selected using the Devices tab on the left and will filter using the specified selection criteria.

Filter criteria are customer and/or workload dependent. The default settings are merely meant to be starting values.

Workload data available for January 26 2009 23:55:00 through January 27 2009 23:55:00.

End time	Lun	SSD	GB	Iops/tb	Iops	rd%	respon...	Qdepth	wait Q	wait tm	xfersize	MB/sec
01/27/09-23:55:00	c4t60060e8005b181000000b181000002efd0	<input type="checkbox"/>	n/a	n/a	125.03	12.51	1.94	0.21	0.03	0.27	29117	3.47
01/27/09-23:55:00	c4t60060e8005b181000000b181000002dcd0	<input type="checkbox"/>	n/a	n/a	40.48	48.46	45.12	0.33	1.50	36.98	36292	1.40
01/27/09-23:55:00	c4t60060e8005b181000000b181000002fbd0	<input type="checkbox"/>	n/a	n/a	39.85	46.73	44.48	0.32	1.45	36.35	36613	1.39
01/27/09-23:55:00	c4t60060e8005b181000000b181000002dad0	<input type="checkbox"/>	n/a	n/a	38.72	46.54	47.04	0.33	1.49	38.60	37134	1.37
01/27/09-23:55:00	c4t60060e8005b181000000b181000002fcd0	<input type="checkbox"/>	n/a	n/a	36.63	42.53	49.47	0.32	1.49	40.65	38296	1.34
01/27/09-23:55:00	c4t60060e8005b181000000b181000002d7d0	<input type="checkbox"/>	n/a	n/a	36.41	42.30	49.12	0.32	1.46	40.22	38342	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002fdd0	<input type="checkbox"/>	n/a	n/a	36.41	42.84	49.50	0.32	1.48	40.66	38445	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002ddd0	<input type="checkbox"/>	n/a	n/a	36.26	42.75	50.65	0.33	1.51	41.62	38534	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002f9d0	<input type="checkbox"/>	n/a	n/a	36.20	42.63	48.07	0.32	1.42	39.26	38658	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002e1d0	<input type="checkbox"/>	n/a	n/a	36.20	42.52	51.42	0.33	1.53	42.35	38633	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002f5d0	<input type="checkbox"/>	n/a	n/a	36.19	42.52	46.33	0.31	1.36	37.68	38704	1.34
01/27/09-23:55:00	c4t60060e8005b181000000b181000002f8d0	<input type="checkbox"/>	n/a	n/a	36.11	42.75	48.26	0.32	1.42	39.44	38621	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002fad0	<input type="checkbox"/>	n/a	n/a	36.08	42.80	48.94	0.32	1.45	40.07	38693	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002d8d0	<input type="checkbox"/>	n/a	n/a	36.07	42.86	49.44	0.32	1.46	40.46	38816	1.34
01/27/09-23:55:00	c4t60060e8005b181000000b181000002e2d0	<input type="checkbox"/>	n/a	n/a	36.04	42.71	51.23	0.33	1.52	42.13	38730	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002ded0	<input type="checkbox"/>	n/a	n/a	36.00	42.59	51.52	0.33	1.53	42.43	38706	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002f6d0	<input type="checkbox"/>	n/a	n/a	35.99	42.62	46.59	0.31	1.36	37.87	38764	1.33
01/27/09-23:55:00	c4t60060e8005b181000000b181000002e3d0	<input type="checkbox"/>	n/a	n/a	35.95	42.59	51.71	0.33	1.53	42.58	38716	1.33

Let's Discuss!