



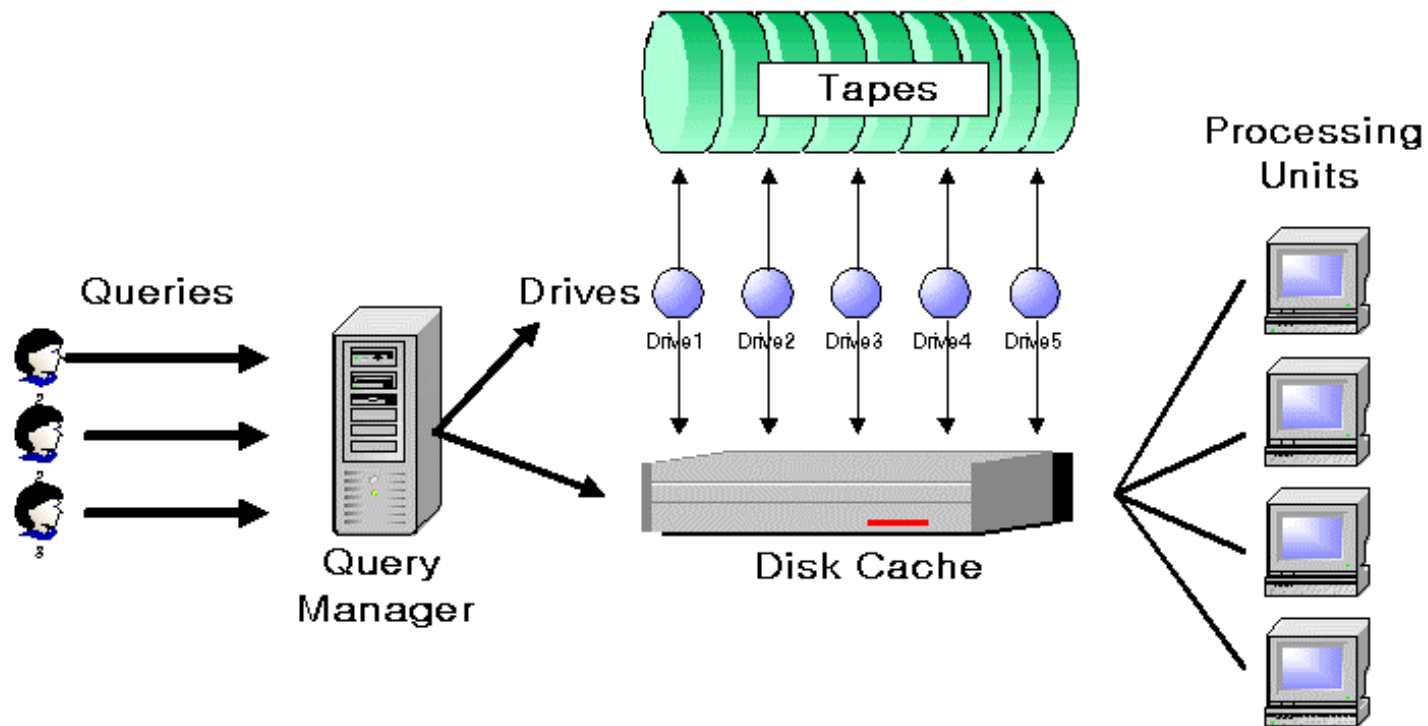
Simulation Analysis of the Optimal Storage Resource Allocation for Large HENP Databases

Jinbaek Kim and Arie Shoshani

**National Energy Research Scientific Computing
(NERSC) Division**

Lawrence Berkeley National Laboratory

The Simulation Architecture Model



Problem: How to balance disk and tape resources

Problem Description



- **Given**
 - Files are stored on a robotic tape system
 - selected subsets of the files are cached into disk caches for analysis or data mining based on user queries
- **Observation**
 - too little disk cache will force files to be removed from disk prematurely
 - too few tape drives will not make good use of a large disk cache, as the throughput from the tape system will form the bottleneck
- **Goal**
 - minimize the number of times that files are cached into disk (maximize file sharing)
 - Balance the tape and disk resources based on the patterns of the requests to the data

The Model Parameters



Physical parameters

- Cache Size: the size of the cache in GBs.
- Number of Drives: the number of tape drives in the tape robot.
- Read Speed: the average time to read a file from tape and to write it into cache.

Query and file parameters

- Number of Queries: the number of queries used in a simulation.
- Number of Files per Query: the average number of files needed by a query.
- File Size: the average file size
- Inter-arrival Time: the time between arrivals of two consecutive queries.

Access pattern parameters

- Overlap Factor: the number of files that overlap between any two queries.
- Clustering Factor: percent of data read from a file on the average by each query.
- Processing Rate: the number of seconds required to process 1 MB.
- Processing time: the time to process a file on the average, in seconds.

The Parameters Used in the Simulation



- The parameters we varied are:
 - 1) the number of drives, D
 - 2) the disk cache size, S_c
 - 3) the processing time per file, T_p
 - 4) the overlap factor, F_o
 - 5) the number of queries, Q , each with F files
- Note: processing time is calculated

$$T_p = \rho \times S_f \times F_c, \text{ where}$$

ρ = processing rate, S_f = file size, F_c = Clustering factor

e.g. $\rho = 2$ seconds/MB, $S_f = 1$ GB, $F_c = 0.01 \Rightarrow T_p = 20$ seconds

File Assignment to Queries According to an Overlap Factor



- **Definition:** the Overlap Factor (F_o) is the probability that any two queries have the same file
 - Given a pool of N files, and F files are selected at random per query, $F < N$
 - Then
 - F / N is the probability that a file is selected
 - $(F / N)^2$ is the probability that a file is selected by any two queries
 - Thus:
$$N = F / \sqrt{F_o}$$
 - e.g. $F=100$, $F_o = 0.1$, then $N = 100 / \sqrt{0.1} = 316$
 - Selecting 100 out 316 randomly per query will generate queries where 10 out of 100 are same in any two queries

The Performance Measure



- **Absolute measure: the average time to execute a collection of queries**
- **We chose to use a relative measure:**
 - **The absolute (actual measure) divided by the “ideal measure”**
 - **“ideal measure” is when all files are in the disk cache**
- **Reason: relative measure is more meaningful to the user**
 - **E.g. 500 sec improvement relative to ideal time of 1,000 has more value than 500 sec relative to 10,000.**

$$efficiency(i) = \frac{T_{actual}(i)}{T_{ideal}} = \frac{T_{actual}(i)}{F \times T_p} ; \quad efficiency = \frac{\sum_{i=1}^Q efficiency(i)}{Q}$$



(GB)	Number of Drives, D				
	$D=1$	$D=2$	$D=3$	$D=4$	$D=5$
10	8.4725	4.6087	3.3325	2.5982	2.0600
20	8.2350	4.4612	3.2564	2.5355	1.9831
30	8.0325	4.3887	3.1672	2.4807	1.9127
40	7.7425	4.2938	3.0643	2.3988	1.8259
50	7.5150	4.1850	2.9587	2.3079	1.7413
60	7.2200	4.0481	2.8513	2.1701	1.6662
70	7.0425	3.9588	2.7842	2.0669	1.5672
80	6.9575	3.7881	2.6512	2.0187	1.4779
90	6.9575	3.6562	2.5460	1.9028	1.4179
100	6.9575	3.5106	2.4460	1.8135	1.3625
110	6.9575	3.3789	2.3449	1.7427	1.3222
120	6.9575	3.3025	2.2575	1.6263	1.2863
130	6.9575	3.2431	2.1972	1.5678	1.2674
140	6.9575	3.1400	2.1084	1.5166	1.2443
150	6.9575	3.0800	2.0509	1.4615	1.2368
160	6.9575	3.0800	1.9507	1.4421	1.2299
170	6.9575	3.0800	1.9052	1.4258	1.2214
180	6.9575	3.0800	1.8829	1.4141	1.2174
190	6.9575	3.0800	1.8521	1.4103	1.2142
200	6.9575	3.0800	1.8291	1.4063	1.2080
210	6.9575	3.0800	1.8178	1.4009	1.2072
220	6.9575	3.0800	1.8157	1.3968	1.2025
230	6.9575	3.0800	1.8145	1.3944	1.1998
240	6.9575	3.0800	1.8145	1.3907	1.1975
250	6.9575	3.0800	1.8145	1.3871	1.1943
260	6.9575	3.0800	1.8145	1.3852	1.1896
270	6.9575	3.0800	1.8145	1.3842	1.1878
280	6.9575	3.0800	1.8145	1.3829	1.1861
290	6.9575	3.0800	1.8145	1.3810	1.1829
300	6.9575	3.0800	1.8145	1.3807	1.1816
310	6.9575	3.0800	1.8145	1.3807	1.1796
320	6.9575	3.0800	1.8145	1.3807	1.1796

The efficiency at the points-of-no-benefit (PNB)

Efficiency measures for:
 $F_0=0.1, T_p=100, Q=20$

PNB: Point of No-Benefit

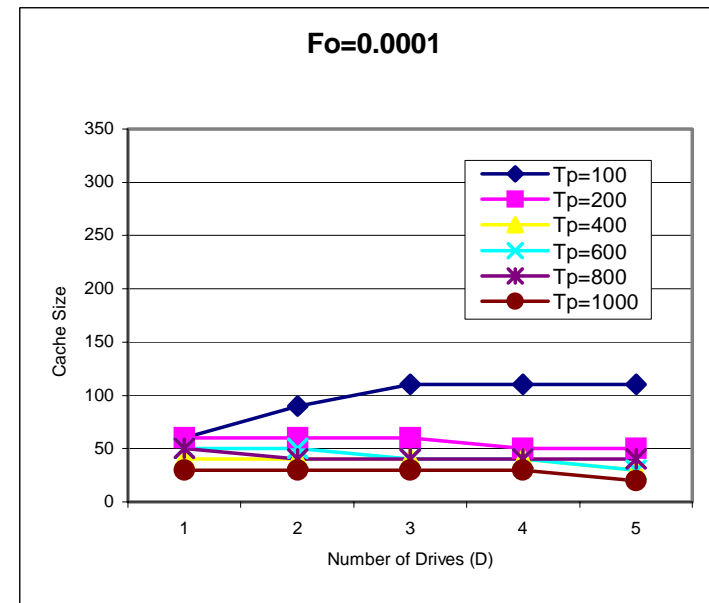
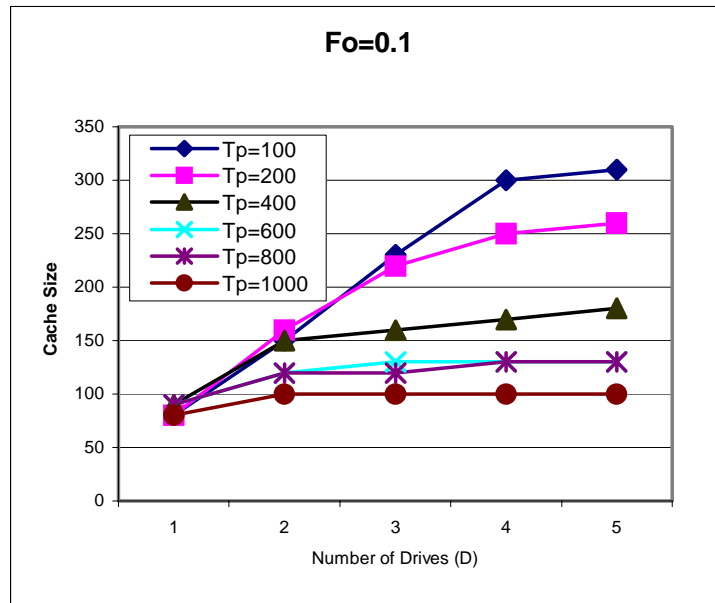


Definition:

Given a "threshold" chosen for the simulation,
and the number of drives D ,
the PNB(D) is defined as the pair (D, Sc) , such that:
 $(\text{efficiency}(D, Sc) - \text{efficiency}(D, Sc + 1)) < \text{threshold}$,
where $Sc + 1$ is the next increment of cache size after Sc

- Only PNB-pairs (D, Sc) are used to plot the results of the simulation

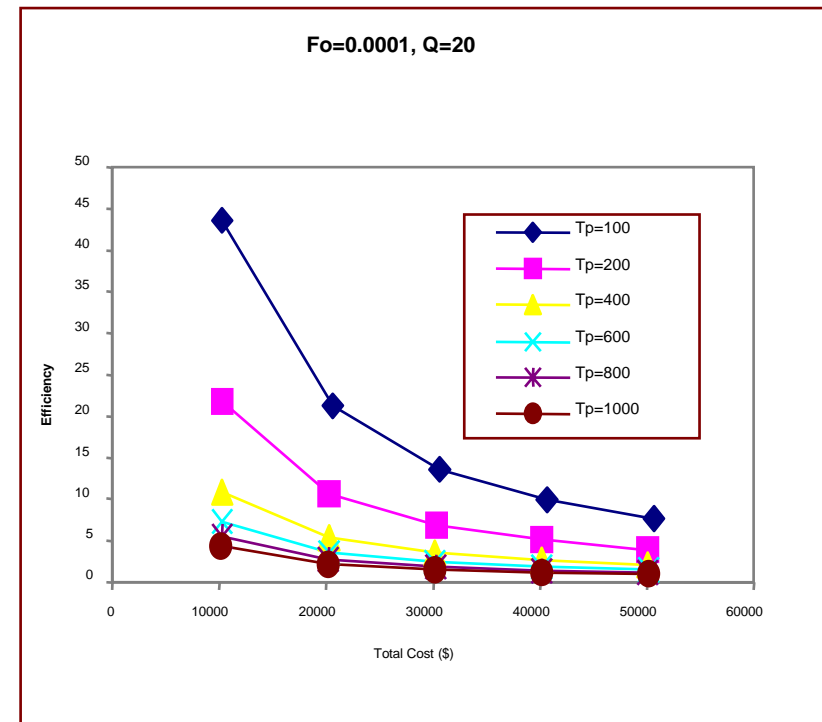
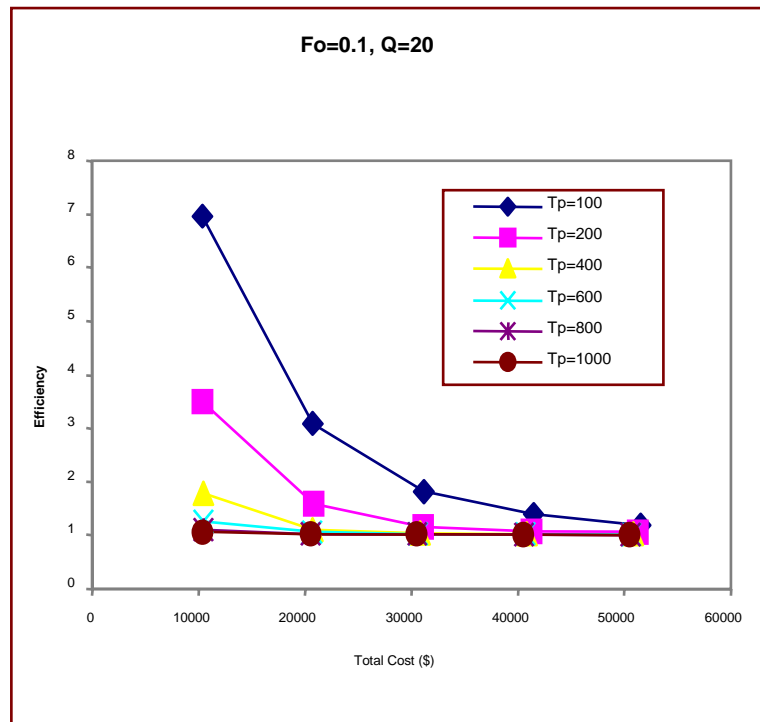
Does extra cache help?



The relationship between number of drives and cache size at PNBs

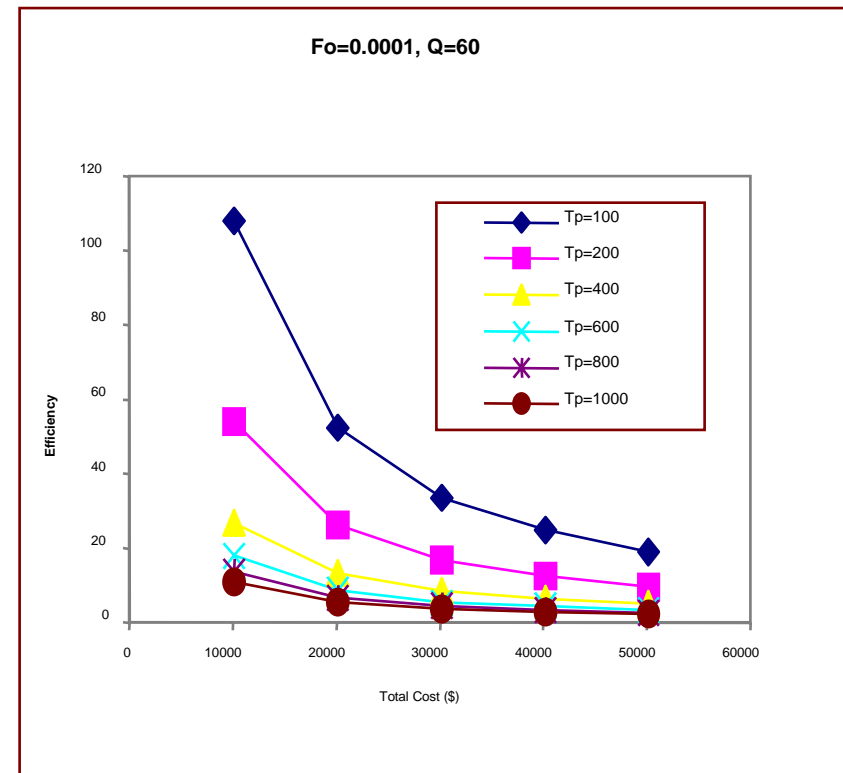
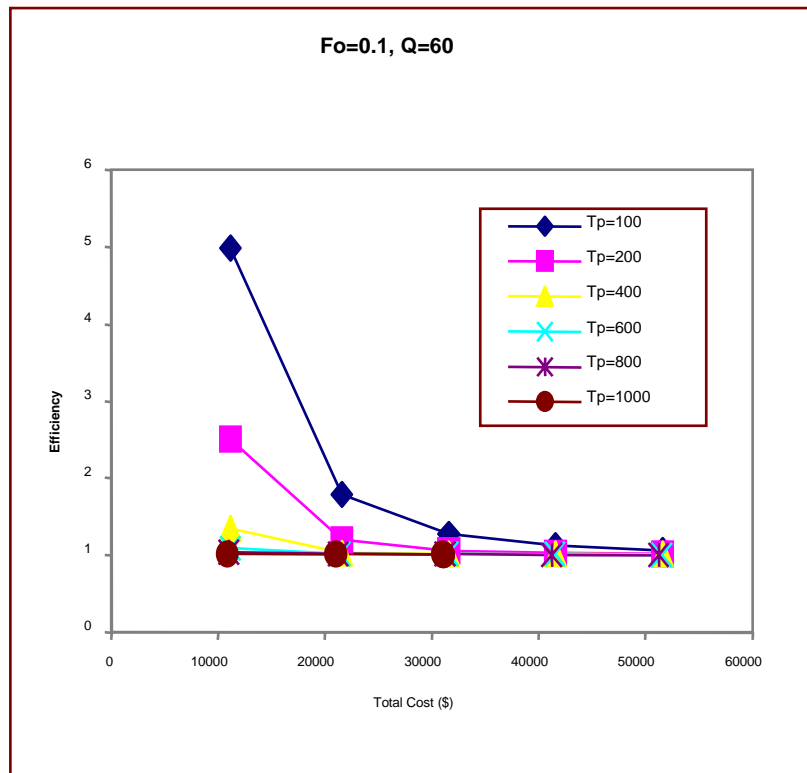
- **Observation 1**: more cache helps only if overlap factor is high
- **Observation 2**: more cache is especially helpful when query processing time is short

Efficiency (at PNBs) as a function of resources (\$ cost)



Observation: when overlap factor is low, efficiency figures are high, and more benefit is derived from additional resources

Efficiency vs. cost – larger number of queries



Observation: increasing the number of queries has little effect when overlap factor is high. Efficiency deteriorate proportionally with a small overlap factor.

Main Conclusions



- **Efficiency is very sensitive to Overlap Factor and Processing Time**
 - **The lower the Overlap Factor the more we benefit from additional resources**
 - **The shorter the Processing Time the more benefit from additional resources**
- **Using the PNB-pairs is an effective way of determining how to balance tape and disk resource**
 - **If budget is limited, what proportions to use**
 - **If budget is “unlimited”, where it is not worth spending further (see next slide)**

Is it worth spending more money?

