Workload Characterization

Yoo Kiseong

[Index]

* Why Workload?
* definition of workload & components of workload
* workload model
* partitioning of workload parameters
* into the hot/cold identification algorithms

Why Workload

Why threashold for hot / cold data identification must be always “static” ?

There are variation between the periods in which the device becomes busy and those in which becomes idle.

Wouldn’t it be more efficient when we construct an algorithm works accordingly ?
- higher cache utilization
- higher throughput
- lower latency
- lower response time
- etc.

=> better performance of algorithm itself

Definition & Components of Workload

= Most of the source of stress on the system

Def. the set of all inputs that the system receives from its environment during any given period of time

A workload refers to a generic unit of work that arrives at the system from the **external source** .

→ transactions
→ interactive command
→ process
→ HTTP request, and
→ depends on the nature of the service

Workload Model

= a representation that mimics the workload of real world in study.

Workload model can be used for
- → the selection of system
- → performance tuning
- → capacity planning

1) Business Characterization : user-oriented description
2) Functional Characterization : describe program, command, requests that make up the workload
3) Resource-oriented characterization : describes the consumption of system resources by the workload (e.g. CPU time, I/O time, memory size)

Workload Parameters

Static Parameters : (analysis on precedence graph)

vs

Dynamic Parameters : (analysis on run-time events)

+++ some examples of workload parameters (cont.)

Workload Parameters

- Inter-Arrival Time : 시스템에 하나의 요청이 도달한 후 다음 요청이 도달하기까지 걸린 시간
- Service Demands : 개별 처리 과정에서 특정한 device 또는 resource 를 사용하는데 걸린 총 시간의 평균 (wait-time 은 포함하지 않는다)
- Number of (Interactive) Terminals
- Think Time
- Type of Request : RD-intesive or WR-intensive
  Sequential-Access or Random-Access
- Type of Resource Demanded
- Duration of the Request
- Quantity of Resources Demanded
- Mutablily of Data
- Network Speed
- more…
- more…
- more…

Partitioning of Workload

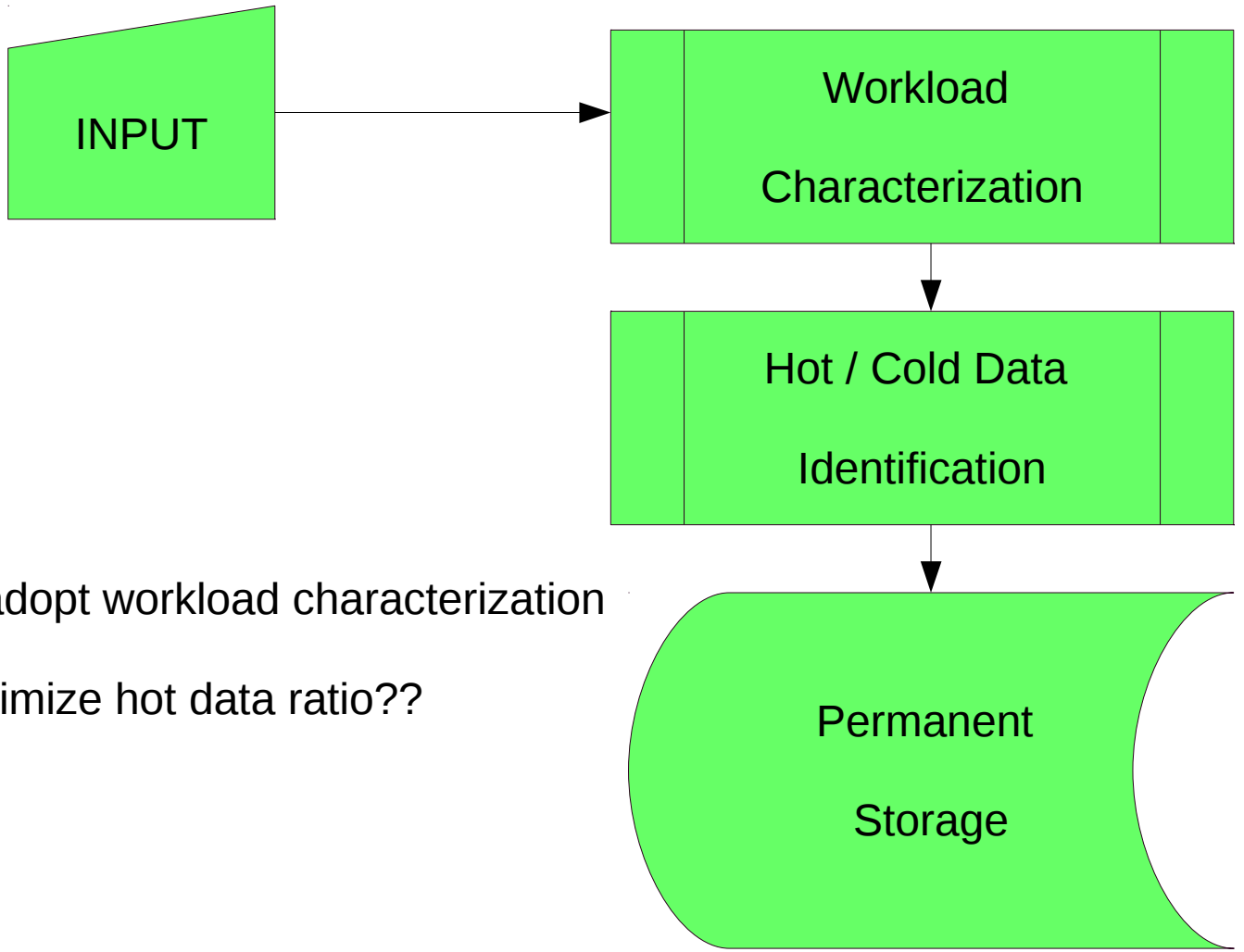
examples of baseline
- resource usage
- pupose of application
- geographical orientation
- functional
- organization unites
- mode

Averaging (with similar types of data)

vs

Clustering (with variant types of data)

Into Hot/Cold Identification



How can we adopt workload characterization

in order to optimize hot data ratio??

End of my presentation.

Thank you !