India Research Lab

Server Workload Analysis for Power Minimization using Consolidation

Akshat Verma, Gargi Dasgupta, Tapan Nayak,
Pradipta De, Ravi Kothari

IBM India Research Lab

Consolidation for Power Minimization

- Phases in a Typical Consolidation Scenario
  - Workload Profiling
  - Workload Sizing
  - Server Selection
  - Placement
- Important Considerations & Questions
  - A consolidation plan remains active for long durations.
  - Workload intensity may vary greatly with time.
    - How are workloads sized (Peak, Average, any other)??
    - Which workloads are co-located on the same server?
  - The servers may be heterogeneous with different power and performance limits.
    - Which servers are selected?

Overview

- Details of the workload studied
- Workload Characteristics
- Implications for the Design of Consolidation Algorithms
- Peak Clustering Based Placement
- Validation

Workload Details

- Production data center of Fortune Global 500 company
- Monitored using MDMS framework
  - 5 min granularity
  - 90 day period in 2007
- Selected 4 application suites with 10, 18, 13 and 16 servers
  - AppSuite1, 2 and 4 are 2-tiered applications
  - AppSuite3 is 3 tiered

Workload Characteristics

Single Workload Characteristic

- **Peak load is typically close to 100%** ➔ *If consolidation is performed by reserving the maximum utilization for each application, the application may require capacity equal to the size of its current entitlement.*
- **90 percentile is much less than Peak** ➔ *If we could size an application based on 90-percentile CPU utilization instead of maximum CPU utilization, it could lead to significant savings.*
- **The tail does not decay exponentially** ➔ *If a statistical measure that ignores the tail of the distribution is used for sizing an application, the consolidated server may observe a large number of SLA capacity violations.*

Workload Characteristics

Correlation between workloads

4. ***There are both positively correlated and uncorrelated applications in a typical server cluster*** ➔ *Hence, correlation needs to be considered during placement to avoid SLA capacity violations.*
5. ***Correlated Applications may not always peak together. Similarly, non-correlated applications may also peak together in some cases.***

Stability of statistical parameters

6. *Some servers exhibit periodic behavior and the future pattern can be reliably forecasted with a day or a week of data. For many non-periodic servers, the statistical properties are fairly stable over time. For highly variable servers, an adaptive prediction method like MovingAverage should be used to estimate the statistical properties.*
7. *The correlation between the CPU utilization of various servers is fairly stable across time* ➔ ***Consolidation based on statistical metrics is practical.***

Core Dump!!!!

Implications for Placement

- Peak based sizing wastes resources
- Sizing based on parameters like mean may have high performance risk
- Correlation between the applications may increase the risk further.

- Correlation-based Placement (CBP): Separate out positively correlated applications. Assumes that the peak of one workload can be handled by borrowing resources from other workloads.

Implications for Placement

- Problems with CBP
  - Correlation is a global metric and may not accurately capture correlated peaks.
  - Single Parameter based sizing can not capture both the tail and the body of the distribution

What next? Some complicated technique that nobody will ever use!!!!!!

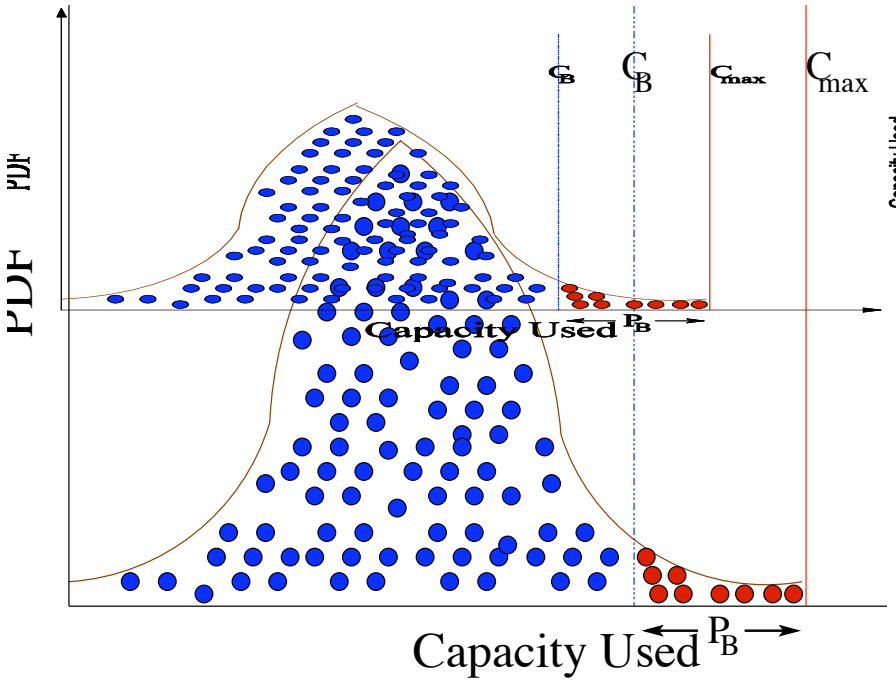
Well, it is really very simple!!!!!

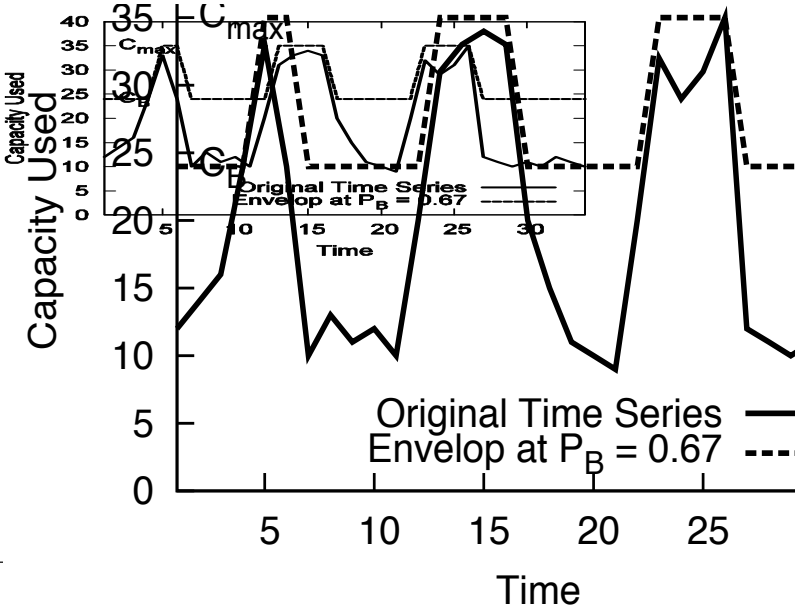
Use two parameters to size a workload

Implications for Placement

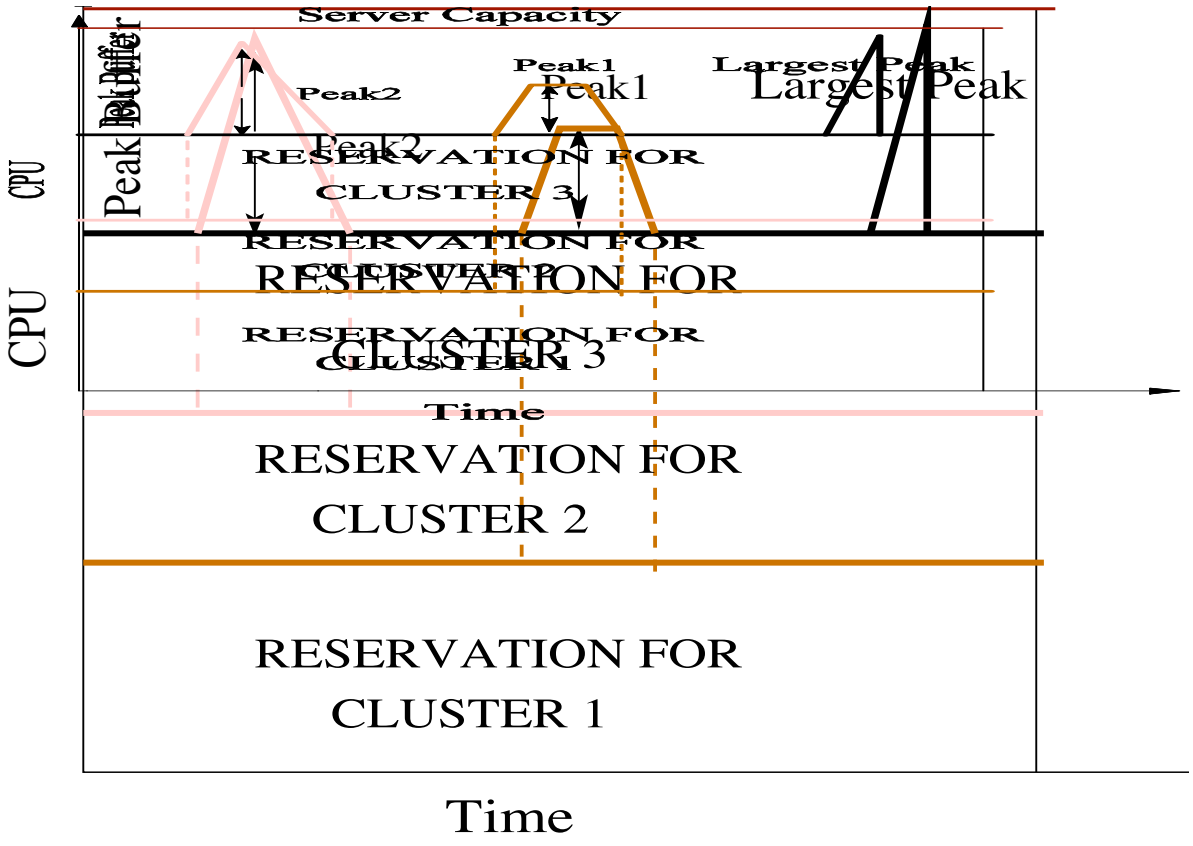
- Peak Clustering Based Placement
  - Use two metrics for sizing: (i) a body based metric when placing with workloads which do not peak together and (ii) a tail based metric when placing with workloads that peak together
  - Cluster workloads that peak together and ensure we proportionally allocate workloads to a server from each cluster.

Peak Clustering Based Placement



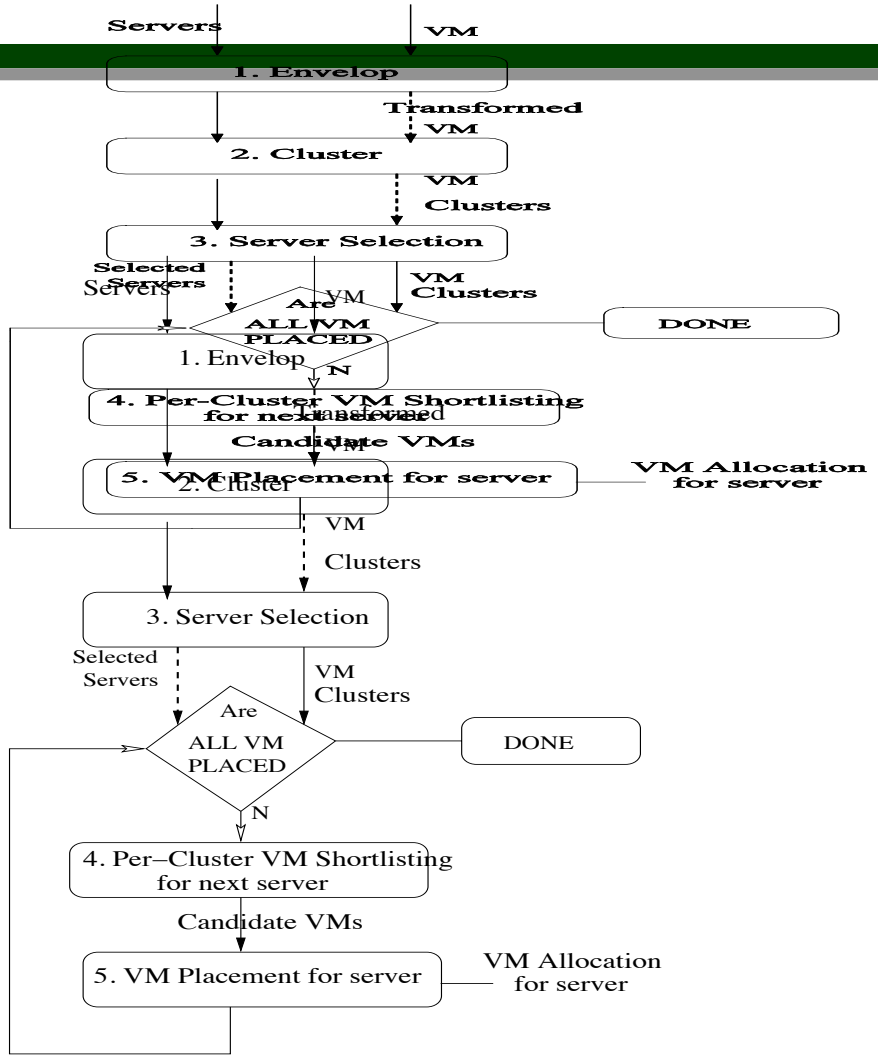


PCP (Contd..)


Server Capacity
Peak Buffer
CPU
Peak2
Peak1
Largest Peak
RESERVATION FOR CLUSTER 3
RESERVATION FOR CLUSTER 2
RESERVATION FOR CLUSTER 1
Time
Time

PCP (Cont ..)

- Create an envelop for each workload based on two parameters
- Cluster workloads based on their peaks
- Select the most power-efficient server
- Allocate workloads from each cluster to the server in a proportional manner.
- Keep a buffer for the peak.
- Use FFD to minimize fragmentation.

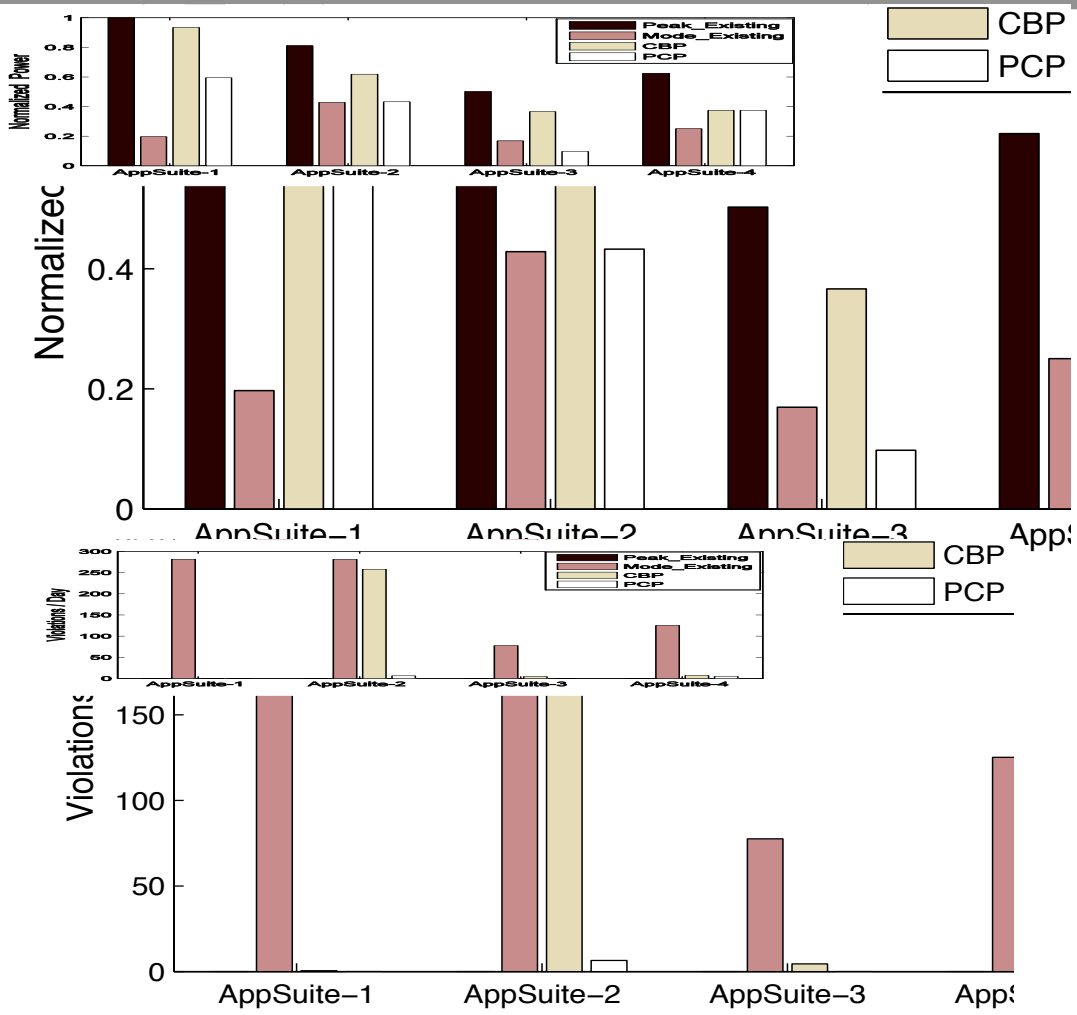


Experimental Setup

- Implemented in IBM Emerald 2.0 Consolidation Planning Tool
- Training Period of 5 days
- Evaluation Period of 1 day (the day following the training)
- Compared with best known methodology with Peak-based and mode-based sizing
- Metrics for Evaluation
  - Power Consumed
  - Capacity Violations
  - Workload Imbalance across servers (difference between the most highly loaded server and the average)

Results Summary

- Power
  - Consolidation based on peak sizes may not even save any power
  - PCP saves almost the same amount of power as mode.
- Violations
  - Mode has the most violations.
  - CBP may have violations depending on the correlation threshold cutoff
  - The capacity violations for Mode may be as large as the capacity of the server.



Conclusion

- CPU utilization of workloads exhibit a high degree of variance
- Most servers peak close to 100% utilization.
- The average utilization is usually much lower but consolidation based on average utilization has high consolidation risk.
- The statistical parameters are more stable than the workload itself.
- Placement based on two parameters for sizing may lead to aggressive consolidation with low risk of capacity violations.

Thanks

- Questions?????