Energy Efficient Virtual Machine Optimization

Vikram Yadav
IBM India Pvt. ltd
Noida, India

Pooja Malik
IBM India Pvt. ltd
Noida, India

G. Sahoo, Ph.D.
Birla institute of Technology
Mesra,
Ranchi, India

Ajay Singh Chauhan
IBM India Pvt. ltd
Noida, India

ABSTRACT
The optimization of energy consumption in the cloud computing environment is the question how to use various energy conservation strategies to efficiently allocate resources. The need of different resources in cloud environment is unpredictable. It is observed that load management in cloud is utmost needed in order to provide QOS. The jobs at over-loaded physical machine are shifted to under-loaded physical machine and turning the idle machine off in order to provide green cloud. For energy optimization, DVFS and Power-Nap are good strategies. As much of this energy is wasted in idle systems: in typical deployments, server utilization is below 30%, but idle servers still consume 60% of their peak power draw. In this paper, we have proposed an hybrid approach for energy optimization using Ant Colony optimization, Bee Colony optimization, PowerNap, DVFS and RAILS having the constraint QOS.

Keywords
Cloud Computing, Hadoop, BigData, DVFS, Power-Nap, Ant colony algorithm, bee colony algorithm etc.

1. INTRODUCTION
We ask that authors Google and Apache foundation launched GFS, Big-Table, MapReduce [1] [2] [3] and Hadoop [4] respectively. Hadoop is a free java based framework that support the processing of large data sets in a distributed computing environment. Hadoop use the concept of google's MapReduce. Amazon launched simple storage services(S3) [5] in 2006. It is a scalable, low cost, high speed web based service designed for online backup and achieving of data and application program. Datacenter are estimated to consumed 2% of the U.S. electrical grid capacity resulting $2 billion per monthly in utility bills.

A McKinery & Company study estimates carbon dioxide emission from datacenters will quadruples to exceed emission from the airline industry by 2020 owing to rapid growth in global demand for computing power [6]. Unfortunately maximum of the energy is wasted by systems that are idle. At idle, current servers still draw about 60% of the peak power [7-9]. In typical data-centers, average utilization is only 20-30% [7] [10]. With the variety of cloud computing model and their framework, some of the computing nodes and data centers form a cluster. Google has many datacenters, one of them in Oregon: consumes equal to a medium size city [11]. During the peak operation, a total of 29.4MW of electricity is consumed by Dublin Datacenter [12]. Idle servers or under - loaded servers impact double on the performance.

Green Computing is a concept of building computing centers and data centers that have an environmentally sustainable and friendly, low power consumption and which emits lesser CO_2. So the main role is being efficient as much as possible, with great concern for the environment. Sometimes, this limits the possibility of handling any workload due to limited resources. For achieving the energy efficient computing, we are using Hadoop as framework in datacenters.

A data center (sometimes spelled *datacenter*) is a centralized repository, either physical or virtual for the storage, management, and dissemination of data and information organized around a particular body of knowledge or pertaining to a particular business.

The National Climatic Data Center (NCDC), for example, is a public data center that maintains the world's largest archive of weather information. A private data center may exist within an organization's facilities or may be maintained as a specialized facility. According to Carrie Higbie, of Global Network Applications, every organization has a data center, although it might be referred to as a server room or even a computer closet.

Need of Datacenter:

1) To maintain high availability

2) To maintain Infrastructure at minimum cost

3) To maintain Connectivity – Data center connectivity is cheap. No other way to say it. When clients are situated in a carrier neutral facility, their cost of bandwidth is a fraction of what they would normally pay on the outside. Clients are putting their high bandwidth applications in data centers to lower to total costs. The savings realized can be spent on redundant connectivity to fully bullet-proof the services.

4) Remote Management – The days of having to "get on" a machine to make changes are over. System Admins are usually sitting at their work stations performing maintenance and updating applications on servers at some other location. Whether the server is down the hall: in a server room or across town in a data center is irrelevant. Proximity to the infrastructure is no longer a day to day consideration.

Apache™ Hadoop® is an open source software project that enables the distributed processing of large data sets across clusters of commodity servers. It is designed to scale up from a single server to thousands of machines, with a very high degree of fault tolerance. Rather than relying on high-end hardware, the resiliency of these clusters comes from the software's ability to detect and handle failures at the application layer.

Deployment Models:
Cloud services can be deployed in different ways, depending on the organizational structure and the provisioning location. Four deployment models are usually distinguished, namely public, private, community and hybrid cloud service usage.

Service Models:
Cloud computing providers offer their services according to several fundamental mode

Main subprojects
Apache Hadoop has two main subprojects: (1). MapReduce - The framework that understands and assigns work to the nodes in a cluster. (2). HDFS(Hadoop Distributed File System) - A file system that spans all the nodes in a Hadoop cluster for data storage. It links together the file systems on many local nodes to make them into one big file system. HDFS assumes nodes will fail, so it achieves reliability by replicating data across multiple nodes.

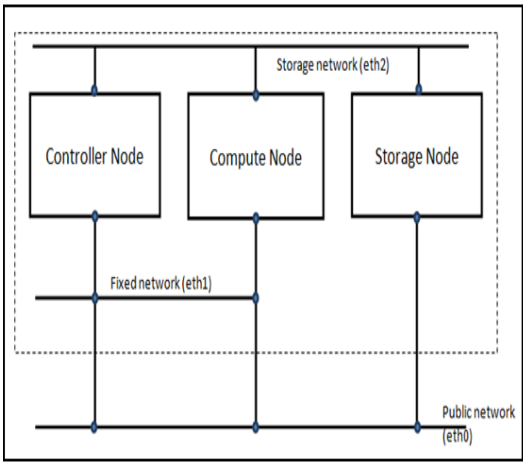
Hadoop is supplemented by an ecosystem of Apache projects, such as Pig, Hive and Zookeeper, that extend the value of Hadoop and improves its usability.

Server virtualization and consolidation can be used to transform data center into a flexible cloud infrastructure with the performance and reliability to run the most demanding application.

Cloud services can be deployed in different ways, depending on the organizational structure and the provisioning location. Four deployment models are usually distinguished, namely public, private, community and hybrid cloud service usage.

Organization of this paper is as follows: Related Work is discussed in section 2. Existing framework of Eco-Friendly Cloud Computing is discussed in section 3. Research methodology is discussed in section 4. Proposed algorithm is discussed in section 5. Section 6 gives details of the experimental results. Section 7 gives the conclusion with a direction of the future work.

2. RELATED WORK
David Meisner [13] presented a power nap strategy for eliminating server idle power. Liang Luo [14] presented a resource scheduling algorithm based on energy optimization methods. Eugen Feller [15] summarized the energy consumption and adjustment methods of standalone computer and cluster servers. In Chung-Hsing Hsu's work [16], They continuously observed the energy consumption and server utilization of server from 2007 to 2010, pointing out that server's defaults energy model has changed. In Charles lively & Xingtu Wu's work [17], they use precise experimental probes to collect energy consumption and performance characterstics in different parallel implementations. HoAydin et.al [18] have proposed the minimizing the energy consumption and subsequently the cost for the static system. K.Mukherjee & G. Sahoo [19] have proposed a framework for achieving better load balancing in Grid environment. Again, K. Mukherjee & G. Sahoo [20] proposed an ant colony optimization and Bee colony for service rescheduling. Once again, K. Mukherjee & G. Sahoo [21] proposed an algorithmic approach for Green Cloud.

3. EXISTING ARCHITECTURE


The overall energy consumption of a cloud computing system can be expressed as the following formula:

$$E_{cloud} = E_{controller\ node} + E_{compute\ node} + E_{storage\ node}$$

$E_{storagenode}$ represents the energy consumption by multichannel memory, $E_{computenode}$ represents the energy consumption by the computing nodes and $E_{Controllernode}$ represents the energy consumption by the controller nodes.

4. RESEARCH METHOLODOLOGY
In the proposed approach, we have used following existing algorithms for different purposes.

4.1 Introduction of new Module in Existing architecture of Data center.

4.2 Hybrid approach (Bee Colony optimization for Load-balancing and Ant Colony optimization for Traversing all nodes).

4.3 PowerNap and DVFS for energy optimization.

4.4 RAILS for distribution of power.

4.1 Introduction of New Module in Existing Architecture of Data Center
There will be a graphical user interfaces by which jobs can be assigned to the data center. Data center is a combination of computing nodes, data nodes and controller nodes. In typical data center all the nodes are in active mode either mode or in idle mode. From the last statics only 30% of the whole power is consumed for the required. 70 % of the power is wasted in idle nodes. We can save this energy wastage by applying prediction method on the incoming jobs. We will use Hadoop framework (A Apache foundation) for data analytics .

We will take last three years(approximation) data for analyzing the next job arrival times on the basis of this PCPM will decide how many nodes to be in inactive , powernap or active mode.

At the first, the consumer requests for service to Graphical user interface for Consumer(GUIC). From here the request for service is transferred to Service Acceptor(SA).SA decide on the basis of Tracker of Virtual machine(TVM) that service should be accepted or rejected. If service is accepted, check for Quality of service(QOS) checker. After that, services are submitted to execution manager which interacts with power consumption policy maker(PCPM). PCPM interacts with virtual machine turn on-off decider(VMTD). Then, it assigns services to VMs. A record tracker tracks the record. Active VMs execute the services on CPUs of IaaS. It turns the powers of the corresponding CPUs of IaaS ON and the rest are made OFF or idle. Then, it sends the services from active VMs to active CPUs of IaaS using mapping of service from virtual machine to IaaS. Thus, all CPUs of IaaS are not turned on, only those required are turned ON and the rest are turned OFF. In this way power consumption of energy is minimized along with minimization of operational cost and the emission of carbon dioxide gas to the environment. Thus cloud computing can be a very promising technology due to its eco-friendliness and minimized operational cost

New Module Power Consumption Tuner node(PCTN) will be in sync with the controller node. PCTN will be communicating to PCPM(Power Consumption policy maker) where PCPM is sub-module of PCTN.

We will take approximately last three years' data for analyzing the next job arrival times. On the basis of this, PCPM will decide how many nodes to be in inactive , powernap or active mode.

In the existing architecture, we need to add a new module called PCTN having 7main sub-units:

4.1.1.TVM (Tracker of Virtual Machine):

It will save the information about compute node and storage nodes as they are in cluster. Information data will be the metadata about the nodes -- whether they are

- overloaded or under loaded
- Node is active, inactive or powernap mode.

4.1.2. SEM (Service Execution Manager):

It will take inputs from TVM and assign the job to the available active node.

4.1.3. SA (Service Acceptor):

It will accept jobs from the Quality of service checker.

4.1.4. VMTD(Virtual Machine Turn on and off decider):

This module interacts will the PCPM module for turning off and on of any virtual

4.1.5. QoSC (Quality of Service Checker):

It will Manage and decide the response time, ensure availability, scalability for each application hosted on the cluster. It will take the input from predefined SLA for each domain/application hosted on the server. Accordingly, the response will be sent back to each requester application.

4.1.6. BS (Billing Service):

This will record the usage of controller node, compute node and storage node during the complete cycle of day having different per unit charge. For example, for storage node, charge would be different as compared to compute node during different timings for each node.

4.1.7. Power Consumption Policy Maker(PCPM):

This sub-module consists of two components:

4.1.7.1. Predictive Analysis Component:

This component will be deploying Hadoop for predication of new incoming job arrival time. It will decide the distribution of machines to be -- active, inactive and idle mode on the basis of results received from Hadoop framework. Hadoop will basically analyze the previous data and will correlate the results with the existing distribution models - exponential , poison, general, etc having some variance.

4.1.7.2. Data-Storage Component:

This Component consists of unstructured data . So, Big-data is the preferred one. For analyzing recent data, Apache Flume is used. It is a distributed system for collecting, aggregating and moving large amounts of data from multiple sources into HDFS or another central data store. Enterprises typically collect log files on application servers or other systems and archive the log files in order to comply with regulations. Being able to ingest and analyze that unstructured or semi-structured data in Hadoop. It can turn this passive resource into a valuable asset.

4.2 Hybrid approach (Bee Colony Optimization for Load-balancing and Ant Colony optimization for Traversing all Nodes)

4.2.1 Ant Colony Optimization

Ant Colony optimization technique was proposed by Marco Dorigo in the early of 90’s. In our problem, there are n nodes in a cluster. Some are under-loaded, over-loaded and some are idle. Ant colony optimization will help in traversing all nodes by shortest path(Traveling salesman problem). After traversing all the nodes, we will find the nodes that are idle and make them turn off. In our practical implementation, we have provided a color code (RGB code) to each node. We will set a threshold value like summation of RGB must be less than or equal to 300. The node will be turn off or on, on the basis of RGB scale. Ants (are blind) navigate from source to food with the help of Pheromone(secreted by ants). Initially, the ants navigate in all possible paths from source to food uniformly. On the long path, ants take more time to return as there will be more distance between the ants. So, the evaporation will be high as compared to small path, as there will be less distance between the ants and the evaporation will be less. In the next iteration, more ants will navigate in that path that has more pheromone means shortest path. So, shortest path is discovered via pheromone trails. Each ant move in random direction, Pheromone gets deposited on the path. Ants detect the lead ants path and inclined to follow more pheromone path. The probability of path being followed starting node are selected at random path based on the amount of trail present on possible paths from starting node. The path that has higher probability has more trails. Ant reached next node, select path and continues until reaches start node. Finished tour is a solution. A completed tour is analyzed for optimality. Trail amount adjusted favor for better solution. A better solution receive more trails. A worse solution receive less trail. Higher probability of ant selecting path: that is a part of better solution. New cycle is performed repeated until most ants select the same tour on every cycle.

4.2.2 Bee Colony optimization

Bee Colony optimization is used for service rescheduling. There are many nodes that are under-loaded and some are overloaded. We will uniformly distribute the load among all the node. Honey bee is a social insect. They work in a decentralized and well organized manner. There are two types of bees: One is forger bees, who collect the nectar and food-stores, who store that in hives. Forger bees moves out for searching the nectar(food). They move randomly in any direction. After finding the nectar, they come back on the hive and start dancing on the dance floor. The duration of this dance is closely correlated with the search time experienced by the dancing bees. There are two types of dance, waggle dance which implies poor quality of nectar and tremble dance (round dance) which implies good quality of nectar. If the dance is tremble dance, then the new born bee agents fly to collect the good quality nectar and store them in the hive. After this operation, the old bee agents die and the new born bee agents start to fly with the good quality nectar stored in the hive, and finally mix them with those sources which are holding poor quality nectar. This process of distribution goes on until there is a uniform quality of nectar in all the sources. Similarly, in cloud computing environment, we observe that some CPUs of IaaS are overloaded for processing consumers services, some are under loaded and some are totally idle. We can save the consumption of energy by turning this idle CPUs OFF and rescheduling services from overloaded CPUs to under loaded CPUs which are totally idle.

4.2.3 Hybrid Approach

For this, the combined Bee-Ants colony system can be used. Firstly, we can divide our jobs into two parts, the first part, which looks after the proper management of overloaded and under loaded CPUs (service rescheduling) and the second part, which helps to manage the idle CPUs. We propose bee colony algorithm for service rescheduling and ant colony algorithm for power consumption management. Thus, for service rescheduling, we can think the tracker as hive which consist of bee agents, from where agents start to forge for nectar. Here nectar implies the threshold value of CPU. Poor quality of nectar implies lower threshold value of CPU whereas good quality of nectar implies the higher threshold value of CPU. Good quality of nectar implies overloaded CPU and poor quality of nectar implies under loaded CPU. In addition, the dance floor represents the service scheduling table, where the information about the status of the CPUs that bee agents have visited are stored. Actually this process is implemented for rescheduling the services from overloaded CPU's to under loaded CPU's in order to provide good QoS to consumers. Again, if the dance of the bee agents is waggle dance then hive tracker indicates the service acceptor to accept new services from consumer. After this step, we apply ant colony based algorithm to find the idle CPUs and then turning them OFF in order to minimize the consumption of power and hence, lower the operational cost. The ant colony from VM turns ON-OFF decider start to find the idle CPUs. The algorithm involves artificial agents called ants who show a cooperative behavior to find the shortest path to the food source from their nest. Ant System is best suited in solving problems like travelling salesman problem which run in dynamic and varied environments without the help of any central control. This is why, it is very much applicable to distributed problem solving. As cloud computing is dynamic in nature, hence we propose to apply ant system in cloud computing in finding idle CPU's. The self-organized behavior of this social insects can be used to solve problems in cloud computing. Here the ants are analogous to cloud data centers and food forging to load allocation to reduce power consumption. We can make analogy with ant colony algorithm for the cloud environment considering its dynamic nature. The creation of controller i.e. Queen ant divides the VMs into two parts: the servers those are in ON state and the servers those are in OFF state. Worker ants are comparable with cloud servers, queen ant with randomly selected cloud server, its reproduction means switching more servers in ON state when needed and cleaner ant's dead body removal means switching OFF the unwanted or idle servers to reduce power consumption. The main challenges in this kind of environment is to manage the state of cloud servers ON and OFF by the controller. When the allocation is already done by using bee colony algorithm, the service scheduling table is updated containing the attributes like remaining power etc. which is maintained by the controller itself (queen ant) and is updated time to time which helps to make decisions that which servers should be ON or put into sleep mode as they are idle.

Proposed Approach:

Consumer request to GUIC for service and GUIC submit request to SA which checks virtual machine on TVM.

If (Virtual machine are available)

{

It will check PCPM

If(utilization power is not within the acquired range)

Required service cann't be fulfilled else (check for service violation and cost)

Required service can be fulfilled

If(virtual machine not available)

Request for service is rejected.

/* ant colony optimization*/

Initialize Trail

Do while (stopping criteria Not Satisfied)-cycle loop

Do until (Each Ant complete a tour)- Tour loop

Local trail update

End do

Analyze tours

Global trail update

End Do

4.3 POWERNAP

It is a state when only the system components that detects the arrival of new work is only powered except these units all the other high power consumption units like the CPU, DRAM and all the processing units are turned off and don't utilize power. It can be utilized to achieve enormous energy saving during idle states. It has two operating modes active mode and nap mode.

Mechanism:

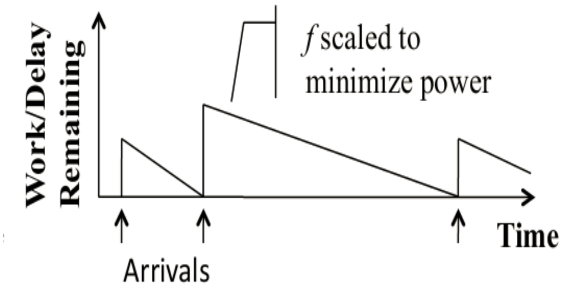
Wake delay: This is the time duration taken by the processing unit to turn on all it's services from the state of nap. Till the wake delay ends, the work remains fixed.

Nap delay: This is the time allotted between the period when the work/job become zero to that of the nap's starting period.

Explanation: Initially, the cluster is at nap. When the work arrives the services, the servers are powered on and the time take to turn it on from the nap state is referred as wake delay. During this wake delay, the work remains constant and after it all the servers starts functioning and executing the jobs. Again, when the jobs ends/finishes,then the cluster for jobs for a definite time, after which it sends the servers back to nap. This definite amount time wis referred as Nap delay. After which the nap remains, till any further work is notified.

4.4 DVFS (DYNAMIC VOLTAGE AND FREEQUENCY SCALING)

This model of job distribution eliminates any idle time arrival between two consecutive jobs. This is achieved by stretching time existing job, to fill any idle time until the next job arrives. To implement it, a prior knowledge of the upcoming jobs should be known.



Explanation: Initially, a certain amount of work is taken in the cluster and when it ends, it creates an idle time state for the servers, we have DVFS to scale the voltage and frequency such that the time taken by the servers can be modified and such intermediate idle states are avoided. This is possible only when the upcoming work's arrival time is predicted or known. So, for each work arrival the previous work is stretched till the new work is arrived.

4.5 Rails(Redundant Array of Inexpensive Load Sharing)

It is basically the replacement of the general(common) Power Supply units(PSU) with a combination(parallel) of multiple inexpensive PSU's. This is very feasible for PowerNap since in PowerNap state a very small energy is required and with need, each PSU is activated and power is taken into the consequently activated servers.

Key features of RAILS are:

1. Efficient dynamic power output switching.

2. Tolerance of PSU failure by using N+K model.

3. Minimal cost.

5. CONCLUSIONS WITH FUTURE DIRECTION OF WORK

In this paper, we have emphasized on developing an eco-friendly data centre architecture in Iaas of cloud by introducing a new module called PCTN (Power Consumption Tuning Node).We have proposed our finding using a flow diagram representation, that showcases the interaction between PCPM and Controller node, which in turn controls the power consumed in data centre using different modules such as TVM, VM, EM, QoSC, SA, VMTD,etc. PCPM contains mainly 2 components -- Analytics framework(Hadoop) and database(in our case it is big data). PCPM basically focuses on low power consumption and simultaneously using existing hardware with improved performance for analyzing large sets of big data using Mapreduce and HDFS functionality of Hadoop. This will help in achieving the scalability and fault-tolerance for a variety of applications by optimizing the execution engine once.

6. REFERENCES

[1] The Google File System, http://labs.google.com/papers/gfssosp2003.pdf.

[2] Bigtable:A Distributed Storage System for Structured Data ,http://labs.google.com/papers/bigtable-osdi06.pdf

[3] MapReduce:Simplifed Data Processing on LargeClusters,http://labs.google.com/papers/mapre uceosdi04.pdf

[4] Hadoop, http://lucene.apache.org/hadoop/

[5] Amazon Simple Storage Service, http://aws.amazon.com/s3/.

[6] The Datacenter Journal, http://www.datacenterjournal.com/facilities/the-green-data-center-opportunity/.

[7] L. Barroso and U. Holzle, "The case for energy-proportional computing,"IEEE Computer, Jan 2007.

[8] X. Fan, W.D. Weber, and L. A. Barroso, "Power provisioning for a warehouse-sized computer," in Proc. of the 34th Annual InternationalSymposium on Computer Architecture, 2007.

[9] C. Lefurgy, X. Wang, and M. Ware, "Server-level powercontrol,"in Proc. of the IEEE International Conference on Autonomic Computing, Jan 2007.

[10] P. Bohrer, E. Elnozahy, T. Keller, M. Kistler, C. Lefurgy, and R. Rajamony, "The case for power management in web servers,"Power Aware Computing, Jan 2002.

[11] http://www.netxt.com/power-103-megawatt-secret-google-container-data-center/

[12] Microsoft Dublin Data Center, http://www.datacenterknowledge.com/inside-microsofts-dublin-mega-data-center/dublin-data-center-generators/

[13] D.Meisner and B.Gold, "PowerNap: Eliminating Server Idle Power", ACM 978-1-60558-215-3/09/03.

[14] Liang Luo, Wenjun Wu "A Resource Scheduling Algorithm of Cloud Computing based on Energy Efficient Optimization Methods" IEEE-2010.

[15] E.Feller., D. Leprince and C. Morin. "State of the art of power saving in clusters results from the EDF", case study. 2010.

[16] C.H.Hsu & S. W. Poole. "Power Signature Analysis of the SPECpower_ssj2008 Benchmark[C]. Performance Analysis of Systems and Software (ISPASS)", 2011 IEEE International Symposium, 2011: 227-236.

[17] C.Lively , X. Wu, V. Taylor, S. Moore, H. Chang and K. Cameron. "Energy and performance characteristics of different parallel implementations of scientific applications on multicore systems", International Journal of High Performance Computing Applications, 2011, 25(3): 342- 350.

[18] H. Aydin, R.G Melhem, D Mosse, and P. Mejia-Alvarez."Power-Aware Scheduling for Periodic Ral-Time Task", IEEE Transactions on Computers, May 2004, , 53(5)

[19] K.Mukherjee and G.Sahoo "Mathematical Model of Cloud computing framework using Fuzzy Bee Colony optimization Technique", International Conference on Advances in Computing, Control and Telecommunication Technologies, IEEE Xplorer, 2009.

7. AUTHOR'S PROFILE

Vikram Yadav: Currently working with IBM. I had completed Master of Engineering in Software Engineering from BIT Mesra, Ranchi. I have already published two international paper on Cloud Computing on Cluster Energy optimization.

Pooja Malik: Dynamic professional with 12+ years of experience in SOA, Cloud, BPM, CEP, BI, Monitoring & Analytics concepts along with SDLC processes. Expertise in concepts of end-to-end project planning and implementation from scope management, to activity sequencing, effort & cost estimation, risk analysis to quality management in adherence to international guidelines and norms. Proficient in monitoring offshore delivery management for technical deliverables of concurrent projects as well as understanding client needs requirements, and expectations and devising solutions accordingly. Have worked closely with various Fortune 500 customers as an Architect Design analyst and Technical lead to understand their existing requirement and Integration, Problem statement and successfully provided them with architectural blueprints, migration roadmaps, Integration framework and life cycle templates

G. Sahoo :received his MSc in Mathematics from Utkal University in the year 1980 and PhD in the Area of Computational Mathematics from Indian Institute of Technology, Kharagpur in the year 1987. He has been associated with Birla Institute of Technology, Mesra, Ranchi, India since 1988, and currently, he is working as a Professor and Head in the Department of Information Technology. His research interest includes theoretical computer science, parallel and distributed computing, cloud computing, evolutionary computing, information security, image processing and pattern recognition.

Ajay Singh Chauhan : Having more then Twenty years of working experience out of which 15+ years (IT experience) in implementation, analysis, data modeling, architecture, design, development and testing of middleware / Enterprise Integration (EAI, BPM), SOA, Internet (B2B, B2C), Client-Server solutions.

Extensively used object oriented methodology with n-tier architecture on Java/J2EE framework along with Oracle/SQL Server and middleware/message brokers solutions like TIBCO and Websphere integration suit. Well versed with prevalent SDLC processes and CASE tools. Possess good communication skills, persistence, highly motivation and quick learning skills.

Experience includes successfully building and managing teams; guiding and mentoring offshore teams. Extensive experience on Pre Sales activities, providing solutions to Business/Sales team in proposal preparation. Have experience on Setting up of Integration competency center.