Environmental Issues of Cloud Computing

Saravana Kumar, J
Assistant Professor, Department Of Computer Science, Periyar Govt. Arts College, Cuddalore.
Corresponding Author: Saravana Kumar, J


Keywords: Cloud Computing, Energy-Efficient Computing And Networking, Energy-Aware Data Centres, Energy Saving.

I. INTRODUCTION

Significant Savings In The Energy Budget Of A Data Centre, Without Sacrificing Service Level Agreements, Are An Excellent Economic Incentive For Data Centre Operators, And Would Also Make A Significant Contribution To Greater Environmental.

According To Amazon Company Estimates [1], At Its Data Centres (As Illustrated In Figure 1), Expenses Related To The Cost And Operation Of The Servers Account For 53% Of The Total Budget (Based On A 3-Year Amortization Schedule), While Energy-Related Costs Amount To 42% Of The Total, And Include Both Direct Power Consumption (~19%) And The Cooling Infrastructure (23%) Amortized Over A 15-Year Period.


The Colloquial Term ‘Smart’ In This Case Means ‘With Low Carbon Footprint’, Showing That The Adoption Of Such ‘Smart’ IT Solutions Will Enable A Potentially Large GHG Reduction, Including Information And Communication Technologies (ICT) Itself Which Is A Large Power Consumer (And Therefore A GHG Emitter), And IT Solutions That Have A Huge Potential Impact In Reducing GHG Emissions In Many Sectors (Fig. 1).
Environmental Issues Of Cloud Computing

Based On A Recent ‘Data Centre Energy Forecast Report’ [3], It Can Be Expected That Savings Of The Order Of 20% Can Be Achieved In Server And Network Energy Consumption With Respect To Current Levels And That These Savings May Induce An Additional 30% Saving In Cooling Needs As Detailed In A Study By HP And The Uptime Institute [4]. It Shows That ‘Most Of Data Centre Power Is Spent On Cooling ICT Equipment (Between 60 And 70%)’.

Thus There Are Very Significant Economic And Environmental Gains To Be Obtained From A Serious Research Thrust On Energy Efficiency In The General Area Of IT And Computer Networks. In Particular, Cloud Computing Is An Inherently Energy-Efficient Virtualization Technique [5], In Which Services Run Remotely In A Computing Cloud That Provides Scalable And Virtualized Resources. Thus Peak Loads Can Be Moved To Other Parts Of The Cloud And The Aggregation Of A Cloud’s Resources Can Provide Higher Hardware Utilization.

II. CURRENT STATE OF ENERGY EFFICIENCY IN ICT INFRASTRUCTURES

ICT Consumes An Increasing Amount Of Energy, But Is Also Instrumental In Increasing Productivity And Economic Prosperity And In Reducing Energy Expenditure From Other Sources Through E-Work, E-Commerce And E-Learning. Traditional Network Design Has Sought To Minimize Infrastructure Costs And Maximize Quality Of Service (Qos).

However, ICT Also Plays A Complex Role In Energy Consumption Via The ‘Communicate More And Travel Less’ Paradigm, As Well As Through The Use Of Smart Devices In Homes And Offices To Optimize Energy Management. Thus, ICT Can Reduce Energy Consumption And Carbon Emissions, But This Potential Reduction Is Partially Offset By The Power Used By Data Centres And Computer Networks [6] Which Runs Into Billions Of Dollars Or Euros.

Thus, A Fraction Of Energy Savings In ICT And Networks Could Lead To Significant Financial And Carbon Savings. In This Section, We Review Recent Research In Energy Efficiency For Standalone Hardware, And Then Review Work That Considers Energy Consumption As Part Of The Cost Functions To Be Used For Scheduling In Multiprocessor And Grid Architectures. Finally, Briefly Review Energy Consumption In Cluster Servers And Wired/Wireless Networks.

2.1. Energy Efficient Hardware


The Advanced Configuration And Power Interface (ACPI) Specification [15] Defines Four Different Power States That An ACPI-Compliant Computer System Can Be In. These States Range From G0-Working To
G3-Mechanical-Off. The States G1 And G2 Are Subdivided Into Further Sub States That Describe Which Components Are Switched Off In The Particular State. For Devices And The CPU, Separate Power States (D0–D3 For Devices And C0–C3 For Cpus) Are Defined, Which Are Similar To The Global Power States. Some Of The Mentioned Techniques Are Usually Applied To Mobile Devices But Can Be Used For Desktop Pcs As Well.

2.2. Energy Aware Scheduling In Multiprocessor And Grid Systems.

In Energy Aware Method To Schedule Multiple Real-Time Tasks In Multiprocessor Systems That Support Dynamic Voltage Scaling (DVS), The Key In Their Approach Is To Consider The Probabilistic Distribution Of The Tasks Execution Time In Order To Partition The Workload And Reduce Energy Consumption. Memory Energy Consumption [16] Can Also Be Reduced By Scheduling Techniques That Impact The Effectiveness Of Frequency Scaling By Combining The Effect Of Tasks Running On A Multicore System, Including Memory Contention And The Technical Constraint Of Chip-Wide Frequency And Voltage Settings.


Another Approach That Increases Energy Efficiency In Data Centres Is Based On Server Consolidation By Service Virtualization [20]. Virtualization Partitions Computational Resources And Allows The Sharing Of Hardware. Many Services Often Need Only A Small Fraction Of The Available Computational Resources Of A Data Centre Server. However, Even When Run At A Low Utilization, Servers Typically Need Up To 70% Of Their Maximum Power Consumption. Such Services Can Be Virtualized And Run Within A Virtual Machine (VM) Resulting In Significant Increases In Overall Energy Efficiency. Depending On Their Utilization, Many Vms Can Run On A Single Hardware Unit (Server Consolidation). Therefore, Less Hardware Is Needed Overall, Thus Reducing Energy Wasted For Cooling, While The Deployed Hardware Utilization Increases. This Consolidation Of Shared Hardware Fosters Energy Efficiency, Measured As Work Accomplished Per Unit Of Consumed Energy.


It Is Important To Point Out That Virtualization Comes At A Cost Which Needs To Be Managed Efficiently. When Resources Are Virtualized, Additional Management Of Vms Is Needed To Create, Terminate, Clone Or Move Vms From Host To Host. Migration Of Vms Can Be Done Off-Line (The Guest In The VM Is Powered Off) Or On-Line (Live Migration Of A Running VM To Another Host). The Management Solution Infrastructure 3 Of Vmware, For Instance, Supports Live Migration.

2.3. Power Minimization In Clusters Of Servers

Recent Research Has Considered Power Minimization In Server Clusters, With Guaranteed Throughput And Response Time. Energy Consumption Depends Primarily On CPU Utilization, But components, Such As Discs, Memory And Network Devices Also Use Energy So That A Server That Seems To Remain Idle May Still Use Up To 60% Of Its Peak Power. In Policies Are Developed That Use Economic Criteria And Energy As Criteria To Dispatch Jobs To A Small Set Of Active Servers, While Other Servers Are
Down To A Low-Power State. Similar Dynamic Provisioning Algorithms Are Studied For Long-Lived TCP Connections As In Instant Messaging And Gaming.

A Queuing Approach To Dynamic Provisioning Technique Has Also Been Studied To Obtain The Minimum Number Of Servers Required To Respect The Required QoS And Reactive Provisioning Can Be Used To Compensate For Sudden Surges In Load.

**2.4. Power Minimization In Wireless & Wired Networks**

According To Some Estimates, ‘The Internet’ May By Consuming More Than 860 Twh Annually, But Such Figures Can Only Be Considered As Educated Guesses Due To The Number Of Assumptions One Has To Make.

Traditionally Fixed Network Operators Have Not Considered Energy Consumption As A Major Cost Factor. Lately, However, As Sustainability Is Becoming A Key Business Objective, Fixed Network Operators Are Looking For Ways To Decrease Their Energy Footprint.


Surprisingly Enough, Energy Savings For Infrastructure Networks Have Not Received Much Attention Until Very Recently, While Energy-Saving Routing Protocols In Wireless Sensor Network Because Of The Specific Needs Of Battery Powered Networks And The Related Research Has Included The Use Of Topology Control That Modify The Network Graph To Optimize Properties Such As Network Capacity And Qos.


In A Wired Node, Power Consumption Depends And Influences Other Factors, Such As The Node’s Throughput; Furthermore, Up To 60% Of A Node’s Energy Consumption Can Originate With Peripheral Devices Such As Link Drivers. Turning Wired Network Nodes On And Off May Be Very Difficult In A Wired Context Because Of The High Volumes, Rate Of Traffic And The Stringent Qos Constraints.

Routing For Wireless Ad Hoc Networks With Battery-Powered Nodes has Also Been Considered In The Selection Paths So As To Satisfy Qos Constraints And Minimize Power. Generally Speaking, We Feel That The Research Community Has Begun To Seriously Consider Energy Consumption In Infrastructure Networks, And The IEEE Has Now Focused On Developing A Standard For Energy-Efficient Ethernet (IEEE 802.3 Az).


**III. TOWARDS ENERGY EFFICIENT CLOUD COMPUTING**

The Previous Discussion Highlights The Need To Develop A Comprehensive Approach For Energy Efficiency That Involves All System Layers And Aspects, Including Physical Nodes, Cooling Of Nodes, Networking Hardware, Communication Protocols And Finally The Servers And Services Themselves.

The Conceptual Framework Of Cloud Computing May Therefore Be A Way Forward To Analyze, Identify And Implement Overall Energy Savings In A System To Attain Truly ‘Green’ Computing Services.

In Contrast To Hardware-Oriented Optimization, Software Systems Can Potentially Be Optimized At Development Time By Specifying Their Energy Characteristics And By Adapting The Implementation.

However, This Requires Individual Adaptation Of Each Component And It Also Requires Understanding The Interaction Between Individual Components When They Operate As A System. A Major Challenge Is Therefore To Explore The Relations Among System Components And The Trade-Offs That Can Result In An Optimal Balance Between Performance, Qos And Energy Consumption And Include Self-Aware
Runtime Adaptation. Thus In This Section We Briefly Discuss Some Areas Of Energy-Efficiency Research Based On A Cloud Computing Perspective.


In A Business Environment Based On Cloud Computing, Workflows That Run Over Many Sites Will Tend To Be Popular. Thus, Developing Methods That Map The Workflow Onto Resources Under The Constraint Of Energy Optimization Becomes A Central Problem Of Great Value And Novelty. Furthermore, In Order To Comprehensively Raise The Energy Efficiency Of A System, All Of Its Layers Have To Be Considered, Including Application Layer Services. Services Have Different Needs Concerning The Environment They Are Running On Or Have Special Properties That Support The Energy Efficiency Of The Underlying System (E.G. Certain Usage Patterns). A Service, For Instance, Might Only Be Used Weekdays, Say, From 8 To 18 H Or Have Peak Usage At A Certain Time Of The Day. A User Also May Also Consider A Trade-Off Between A More Energy-Efficient Service And A More Reliable Or Faster Service, And Compose The Service In A Way That Fits Its Needs, Thus It Should Also Be Possible To Develop Accounting Mechanisms That Track And Depend On The Energy That Has Been Used By A Service.

3.1. Energy Aware Data Centres


However, The Degree Of Energy-Efficient Self-Management In Data Centres Is Still Limited Today. Services Should Not Only Be Virtualized And Managed Within A Data Centre Site But They Should Be Moved To Other Sites If Necessary. Not Only The Aspect Of Load Has To Be Considered, Also The ‘Heat’ Generated By A Service Has To Be Measured And Accounted For Before Migrating Operations. Every Operational Physical Node Produces Heat.

When A Particular Node Is Excessively Used Or Is Near Other High-Loaded Nodes, Hotspots Can Appear In A Given Data Centre. To Avoid Such Hotspots, Heat Can Be Distributed Across Sites. Furthermore, Services Can Be Moved From Sites With High Load Or High Temperature To Sites With Smaller Loads And Lower Temperatures. Generally, Services Should Be Moved To Those Locations, Where They Can Operate In The Most Energy-Efficient Way. This Kind Of Energy-Efficient Management Of Resources Has To Be Realized By An Autonomous Energy Management That Is As Transparent As Possible To The User Of A Service. Energy-Related Problems Have To Be Solved According To Defined Policies Without Needinghuman Interaction. Machine-Readable Descriptions Of The Needs And Features Of Services, Servers, Networks And Even Whole Sites Have To Be Available To Enable Energy Efficiency In The Highly Autonomous And Adaptive Systems Of The Future.

3.2. Energy Savings In Networks And Protocols

Some Hardware Already Offers Features That Create An Opportunity For Energy-Efficient Operation Such As Turning Off Network Interfaces And Throttling Of Processors. Network Protocols Could Also Be Optimized, Or Even Be Redeveloped In A Way That Enhances The Energy-Efficient Operation Of The Network Elements. Network Devices Could Be Enabled To Delegate Services To Other Devices So As To Transfer Services From Energy Inefficient To More Energy-Efficient Devices Or To Devices That Need To Be Always On, While Certain Other Devices Are Turned Off. The Delegating Device Can Then Become Dormant And Be Turned Off. Currently, Many Basic Network Services Have To Remain Active To Periodically Confirm Their Availability Even When No Communication Is Taking Place.

These ‘Soft States’ Make It Impossible To Turn Off Certain System Components. Therefore, New Protocols Need To Be Designed To Work Around Such Soft States So As To Increase The Energy Efficiency Of The Network. Signaling Can Also Be Revisited In This Context, Whereas Data And Signaling Traffic Vary Widely, The Same Technology And Mechanisms Are Used For Both (In So-Called In-Band Signaling). While Signaling Needs Only Low Bandwidth But Can Occur Anytime, Data Traffic Occurs After Signaling Has Taken Place, Usually Requires High Bandwidth And Traverses All Network Layers Up To The Application Layer, And Uses Processing Power Through Multiple Layers Of The Network.
Therefore the use of out-of-band signaling should also be evaluated to design and improve energy-aware communication protocols.

3.3. The Effect of Internet Applications

One large application area for the internet is in information dissemination. From digital cameras embedded in mobile phones to environmental sensors to the web 2.0, end users are generating and interconnecting unprecedented amounts of information and this trend is expected to continue unabated. However, the professional, expedited and reliable distribution of content requires increasing investments in infrastructure build-out and maintenance, and a matching electricity bill to run the underlying ICT.

Web, peer-to-peer and web-based video-on-demand services currently dominate internet traffic and, taken together, consistently comprise 85% or more of the internet traffic mix for several years. In practice, dissemination networks operate using methods and paradigms based on remote-access, replicating functionality in several parts of the protocol stack, and fail to benefit from recent advances in wired and wireless communications, storage technologies and Moore’s law. If cloud computing becomes a significant platform for producing and accessing information, the amount of data that will be transferred over the internet will increase significantly.

Content replication and dissemination algorithms will then need to consider energy as a key parameter of optimal operation, and therefore cloud computing calls for a thorough re-examination of the fundamentals of major computation/communication/storage and energy/performance trade-offs.

IV. CONCLUSION

This paper has reviewed the potential impact of energy-saving strategies for the management of integrated systems that include computer systems and networks. The propose that cloud computing with virtualization as a way forward to identify the main sources of energy consumption, and the significant trade-offs between performance, QoS and energy efficiency and offer insight into the manner in which energy savings can be achieved in large-scale computer services that integrate communication needs. Based on the approaches that identified and specific plug-ins and energy-control centres for networked large-scale hardware and software can be implemented and that they can have significant impact, including:

- Reducing the software and hardware related energy cost of single or federated data centres that execute “cloud” applications.
- Improving load balancing and hence QoS and performance of single and federated data centres.
- Reducing energy consumption due to communications.
- Saving GHG and CO2 emissions resulting from data centres and networks so as to offer computing power that is “environment protecting/conserving”.

Such improvements can have additional impact by reducing energy utilization for transportation and work by encouraging ‘green’ ICT-based smart solutions for E-work, E-learning and smart climate control for homes.

REFERENCES

[10]. AMD Powernow! Technology.
Environ
mental Issues Of Cloud Computing

[12]. AMD Cool’n’Quiet Technology.