OPENSTACK AND INTEL TECHNOLOGY BEST PRACTICE

Contributors

Billy Cox
Software and Services Group,
Intel Corporation

Jiangang Duan
Software and Services Group,
Intel Corporation

Fleming Feng
Software and Services Group,
Intel Corporation

Das Kamhout
IT, Intel Corporation

Susie Li
Software and Services Group,
Intel Corporation

Tonny Liu
IT, Intel Corporation

As IT moves into the era of "new IT", cloud and related technologies are moving quickly. Many factors including cost constraints and consumerization of IT are driving these changes. Open source is leading the innovation tide and playing an indispensible role in shaping the future of new IT. OpenStack, as an emerging open source software stack for creating and managing clouds, has gained significant interest from both academia and industry leaders. In this article we illustrate Intel IT's OpenCloud project, which implemented an open source private cloud in a production environment based on OpenStack. Furthermore, we present how Intel® Node Manager and Intel® Trusted Execution Technology can help improve the power efficiency and security of a cloud environment.

Introduction

The "new" IT continues to emerge from the shadows of traditional IT. A few challenges are pushing IT to not only consider changes, but to do so quickly. The forces include reducing IT cost while significantly increasing agility, improving security and related controls, and the consumerization of IT.

The pressure to reduce the cost of IT is not new. But now, in addition to being more efficient, IT is being asked to be a key partner in driving the business forward. This role places IT squarely in the position to not just host content and applications, but to show the business how to use modern tools to advance the business.

Security remains a top concern for the use of clouds, especially public multitenant clouds. Security includes things like access control and logging, encryption, single sign-on, and trust.

Consumerization is a recent trend often associated with "bring your own." In reality, the trend is much more than just, "bring your own," because the new users of IT simply have different expectations. They look at tools like Dropbox* as compared to traditional enterprise sharing and expect the simple functionality of Dropbox. They expect to use a tool and see immediate results, as well as rapid updates. And as the industry innovation continues, there is an expectation that IT will keep up.

These are fundamental forces driving IT to change—not merely adjust.

While traditional IT would have delivered email, new IT is expected to also deliver full collaboration with large files and cross-group editing. Traditional IT delivers each user a laptop; new IT empowers the employee to use whatever

device they choose. Traditional IT would provide a structured database; new IT provides large scale means to analyze unstructured data.

The quest to develop solutions for the new IT have involved numerous players in many parts of the industry and have demanded innovations across the field. In many ways, open source is leading the way delivering key technologies. As an example, the well-known use of Xen in the Amazon Web Services public cloud would not have been possible without the hard work of a dedicated community of contributors.

OpenStack is a new and emerging open source software stack for creating and managing clouds. At its heart, it is about building an open, extensible, framework for managing the various resources in cloud environment (compute, network, storage, and so on). The project mission is "To produce the ubiquitous open source cloud computing platform that will meet the needs of public and private cloud providers regardless of size, by being simple to implement and massively scalable."[1]

OpenStack was originally launched by Rackspace in collaboration with NASA in July 2010. It gained significant interest in a short time from both academic and industry leaders. As of this writing, there are more than 170 companies[2] that have joined the project. It is the strong vibrant community of OpenStack contributors that is the key to the strength of OpenStack. All of the code of OpenStack is freely available under the Apache 2.0 license.

As shown in Figure1, OpenStack is composed of a set of interrelated projects, which make up the various components of a cloud computing platform. The latest stable release named Essex[3] includes five key projects:

- OpenStack Compute (codenamed Nova) provides a tool to provision and manage large number of virtual machine instances and networks. It is similar in scope to Amazon EC2* and Rackspace Cloud Servers*. OpenStack Compute is designed to be both hardware and hypervisor agnostic.
- OpenStack Object Storage (codenamed Swift) provides a distributed, eventually consistent virtual object store. Swift has built-in redundancy and failover management and allows scaling to multiples of petabytes, billions of objects distributed across nodes.
- OpenStack Image Service (codenamed Glance) provides discovery, registration, and retrieval services for virtual machine images.
- OpenStack identity (codenamed Keystone) provides a unified authentication across all OpenStack projects, including Token, Catalog, and Policy services.
- OpenStack Dashboard (codenamed Horizon) provides administrators and users the ability to manage their infrastructure using a simple web interface.

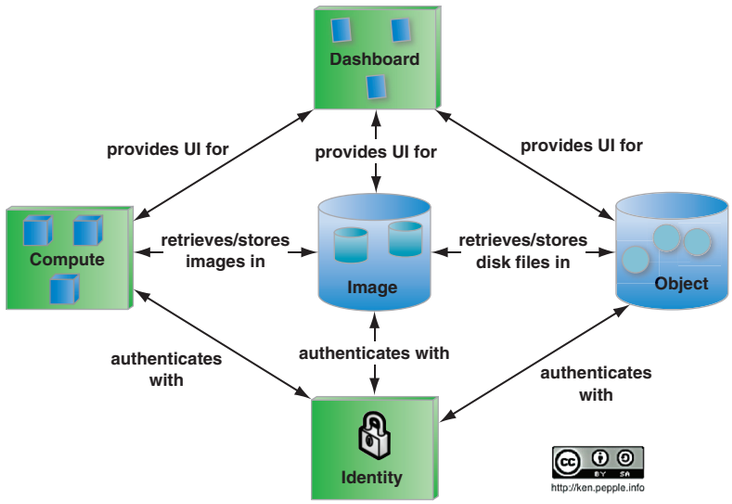

Figure 1: Key OpenStack projects (Source: www.openstack.org/Documentation/os-compute-adminguide.pdf, the document itself is under Apache 2.0 License)

In the balance of this article we examine in more detail a real-world implementation of OpenStack in Intel IT and extensions to OpenStack for support of trust and power management.

Intel IT OpenCloud

In late 2010, Intel IT implemented their first private cloud solution focused on delivering compute IaaS to its internal application developers and application owners. This was predominantly built through a combination of existing enterprise manageability tools and solutions coupled with integration software and databases. The goal and results were aimed to provide a cohesive solution that brought together compute, storage, and network resources to expose compute IaaS to Intel IT's end users. The success of the work was paramount in instituting rapid instantiation of new capacity for application developers (from 90 days to acquire a server to under three hour SLA with most happening in under 45 minutes), to establishing a federated capacity allowing increased sharing through multitenancy and resource pooling of assets, and to deliver numerous improvements in the automation and data transparency available to the operations team.

In late 2011, Intel IT decided to augment their cloud with an open source approach. This approach allows the team to get the latest technology the open source community has developed, focus only on the specific challenges they faced, and also provide feedback to the community with Intel's most advanced technologies. The new design and implementation based on open source is referred to as the Intel IT OpenCloud.

Foundation of the Intel IT OpenCloud

For Intel IT OpenCloud the team had an opportunity to revisit a number of design decisions and seek out areas allowing them to work in the open community for cloud acceleration.

The team analyzed numerous options based on performance, reliability, cost, and features, which led to an architecture consisting of Intel 5600 series blade servers for compute servers, utilizing 10GBe for network fabric, and Intel 5600 2U servers for the storage nodes. The modular nature of the underlying platform with large network pipes for connectivity is intended to allow for significant scale out, to achieve the necessary resiliency and performance required for a wide range of workloads with a primary focus on cloud-aware applications.

For the cloud operating environment, the team explored numerous solutions and closed in on OpenStack for its vibrant community and its ability to flexibly handle the operations for compute, network, and storage. The team implemented the 1.0 version of Intel IT OpenCloud based on OpenStack Diablo release. The environment is built to handle rolling upgrades to allow for no impact implementation of new infrastructure versions, therefore the OpenStack Essex version could be integrated no later than three months after release.

A key aspect of the foundation is the usage of Nova security groups to allow for automated logical segmentation amongst tenants in the environment as well as between VM roles inside of a single tenant. The utilization of security groups enables rapid configuration of IP tables at instantiation time, which significantly lowers the effort expended in ensuring reasonable segmentation in a multitenant resource pool.

Monitoring and Managing the OpenCloud

In the existing Intel IT private cloud, significant focus is placed on ensuring that we could monitor the entire environment, including compute, storage, and network resources, simultaneously in a single view. It is also critical that operators can manage all resources in an automated fashion for significant scale with minimal interaction. For the Intel IT OpenCloud, the team chose to implement a solution based on the open source monitoring tool Nagios by taking advantage of its extensive list of pre-built monitors, its support of multiple operating systems, and its ability to monitor many of the resources beyond OS (load balancers, firewalls, network switches, and so on).

Monitoring alone is not that useful in an automated environment. The team decided to make extensive use of a real-time configuration management system to control the environment as it scales. The team focused on the open source tool Puppet to handle most of configuration management actions. However to complete the monitoring and managing circle to enable full automation the team architected in a small but important component that handles business logic rules.

The basic aspects of the architecture and implementation are that the watcher (Nagios) sends alerts onto the message bus to which the event handler is subscribed, and the event handler makes decisions based on real-time configuration data of the infrastructure and application layout. As examples:

- If an event is sent onto the bus about a specific node having issues inside a given application scale unit (combination of server instances for scaling), the event handler will decide to destroy that node and instantiate a replacement.
- If a more catastrophic failure happens that affects the entire data center, the event handler can make a choice to disable a scale unit inside the data center, or even remove the data center completely from the global load balancer list of DNS end points.

Currently the team focused on three levels of automated remediation. The three levels are: destroy and create node; remove scale unit from load balancer (in some situations this means 20 servers removed from the load balancer); and remove data center from global load balancer pool. These three levels allow for cloud-aware applications to operate effectively in an active/inactive implementation dispersed across multiple data centers to ensure high reliability, which is a key goal of the Intel IT OpenCloud team.

Integration with Enterprise Systems

Intel IT has investments in numerous enterprise technologies from our service management tools to our authentication and entitlement tools. One of the goals of the team was to show how an open source infrastructure could integrate well with existing solutions that run an enterprise. The team picked a few areas to focus on first, and in this article we outline service management.

The integration into the service management system was paramount as the Intel IT team was in the midst of transforming into a complete Information Technology Information Library (ITIL) environment. The architecture and design goals were to have the systems themselves provide the necessary data. Therefore the utilization of the configuration management system coupled with the monitoring system and correlation engine allowed for provision time correlation of resources, which is fed onto the message bus and then imported into the service management tool. The monitor/watcher is also fed information at provision time to ensure that the resources are immediately monitored and so alerts on those resources can be easily ingested into the service management tool again through the use of the message bus. This allows for automated remediation to happen in a self-contained fashion and only exceptions requiring an operator to receive a ticket for problem management. By utilizing a message bus model with publish/subscribe methods, the design allows for a very flexible approach to what causes alerts, what causes auto-remediation, and what generates a ticket for operator analysis.

The team expects to continuously improve the environment, and the next areas of focus on are orchestration, block storage, auto-scaling policies, complex application deployment, as well as providing the foundation for the Intel IT PaaS solution.

Intel Technology Adding Value in Cloud

Intel IT OpenCloud provides a success story to use OpenStack to build up a viable private compute IaaS solution. Though this is a good foundation, there is still much work ahead to make cloud computing optimized. Security and energy efficiency are undisputedly among the top cloud computing challenges. In the following section, we will introduce Intel Node Manager and Intel Trusted Execution Technology, and what additional value they can bring to cloud and make the cloud environment more power-efficient and secure.

Server Power Management

Server power consumption is often an afterthought in data centers. For example, in many facilities the utility bill is bundled with the overall building charge which reduces the visibility of the data center cost.

Even though servers have become much more efficient, packaging densities and power have increased much faster. As a result, power and its associated thermal characteristics have become the dominant components of operational costs.[4]

Power and thermal challenges in data centers include:

- Increased total operational costs due to increased power and cooling demands.
- Physical limitations of cooling and power within individual servers, racks, and data center facilities.
- Lack of visibility into actual real-time power consumption of servers and racks.
- Complexity of management components and subsystems from multiple vendors with incompatible interfaces and management applications.

These challenges to manage data centers can be translated into the following requirements:

- Power monitoring and capping capabilities at all levels of the data center (system, rack identification, and data center). What can be done at an individual server level becomes much more compelling once physical or virtual servers are scaled up significantly.
- Aggregation of the power consumed at the rack level and management of power within a rack group to ensure that the total power does not exceed the power allocated to a rack.
- Higher level aggregation and control at the row or data center level to manage power budget within the average power and cooling resources available.
- Optimization of productivity per watt through management of power at the server, rack, row, and data center levels to optimize TCO.
- Application of standards-based power instrumentation solutions available in all servers to allow management for optimal data center efficiency. Extension of instrumentation to enable load balancing or load migration

based on power consumption, and close coupled cooling for the management of pooled power and cooling resources.

Intel® Node Manager

Intel Node Manager is a smart way to optimize and manage power and cooling resources in the data center. This server power management technology extends component instrumentation to the platform level and can be used to make the most of every watt consumed in the data center.[5]

Intel Node Manager is designed to address typical data center power requirements such as described above. It is implemented on Intel server chipsets starting with Intel® Xeon® processor 5500 series platforms. It provides power and thermal monitoring and policy based power management for an individual server and is exposed through a standards based IPMI interface on supported Baseboard Management Controllers (BMCs). Intel Node Manager requires an instrumented power supply conforming to the PMBus standard.

Intel Xeon processors regulate power consumption through voltage and clock frequency scaling. Reducing the clock frequency reduces power consumption, as does lowering voltage. The scale of reduction is accomplished through a series of discrete steps, each with a specific voltage and frequency. Voltage and frequency scaling also impacts overall system performance and therefore will constrain applications. The control range is limited to a few tens of watts per individual microprocessor. This may seem insignificant at the individual microprocessor level; however, when applied to thousands or tens of thousands of microprocessors typically found in a large data center, the potential power savings amount to hundreds of kilowatt hours per month.

Intel Node Manager is a chipset extension to the BMC for supporting in-band and out-of-band power monitoring and management at the node (server) level. Some of the key features include:

- Real-time power monitoring
- Platform (server) power capping
- Power threshold alerting

Figure 2 shows the Intel Node Manager server power management closed control loop.

The typical benefits brought by Intel Node Manager are illustrated in the following use cases.

Power Management Use Case 1: Real-Time Server Power Monitoring

Power monitoring is a critical capability that enables us to characterize workloads and identify opportunities and hotspots to increase data center energy efficiency. This use case needs to have Intel Node Manager–enabled hypervisor hosts, which can be a combination of open source Xen or KVM. Real-time power utilization of the server is shown in Figure 3.

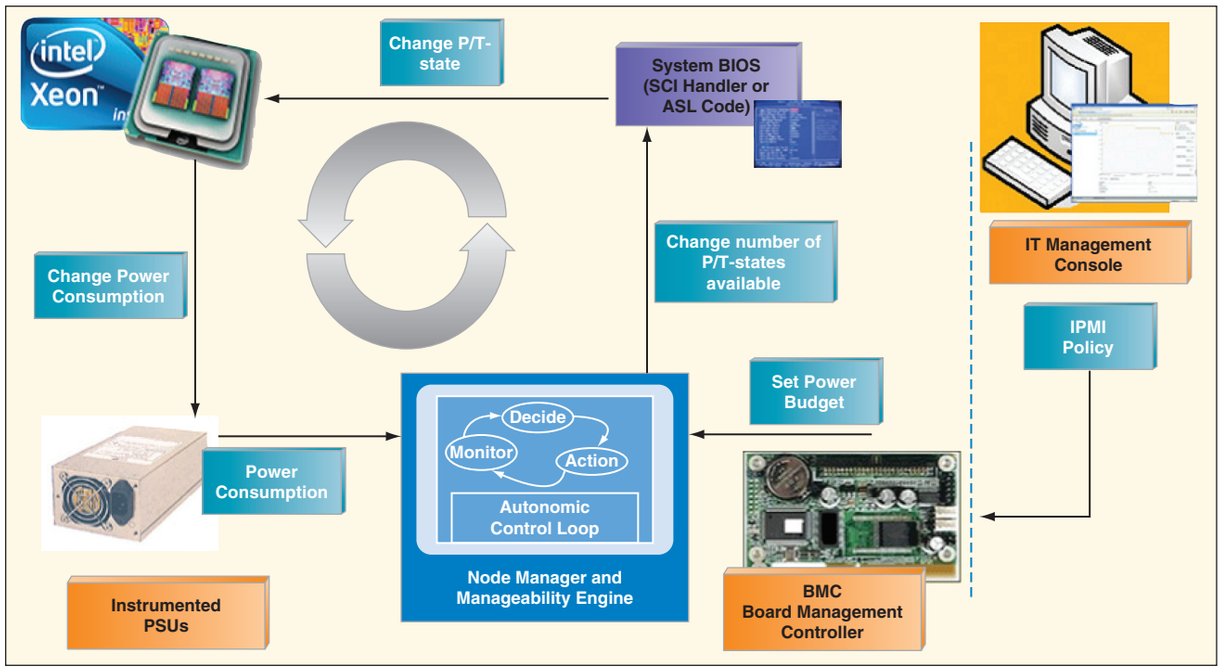

Figure 2: Intel® Node Manager closed control loop
(Source: Intel Corporation, 2012)

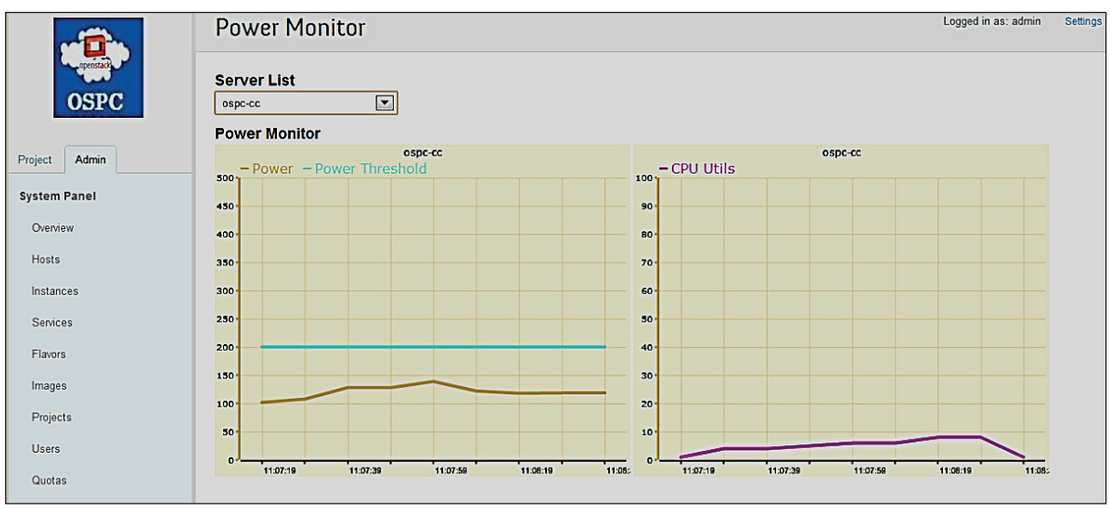

Figure 3: Real-time power consumption monitoring
(Source: OSPC)

Power Management Use Case 2: Policy-Based Resource Distribution Using VM Migrations

Real-time power consumption data allows us to perform power-aware resource distribution. Virtual machines that run the workload can be relocated

to optimize and rebalance power margins based on measurements. The virtual machines can be relocated from power constrained systems to unconstrained systems within the cluster or across different clusters for better system utilization and performance. Figures 4 and 5 show VM migration based on real-time power consumption data. Once the server power goes beyond the threshold, the VM has been migrated out of the server on to other unconstrained server within the cluster. Therefore, the power consumption of the original server has dropped down.

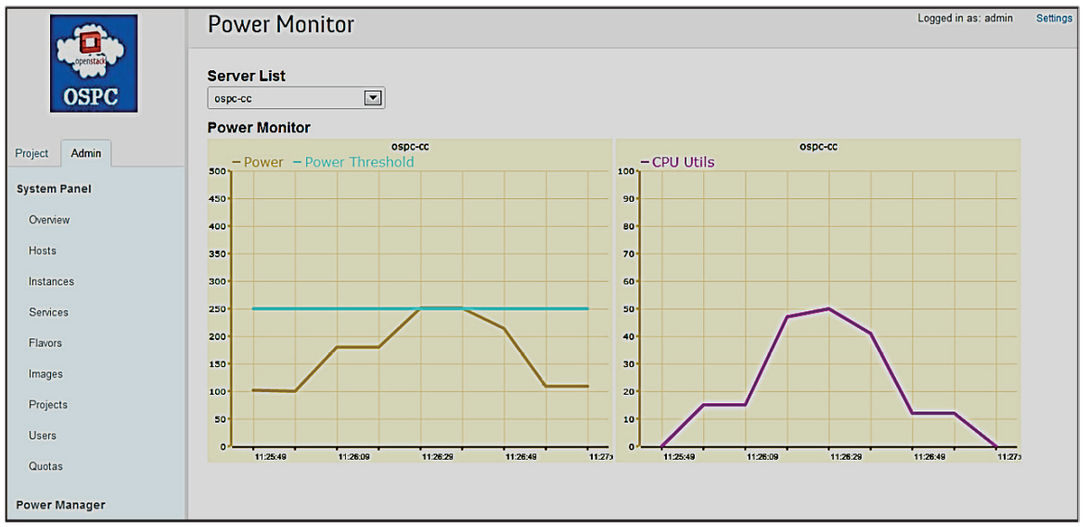

Figure 4: Policy-based VM migration—monitoring (Source: OSPC)

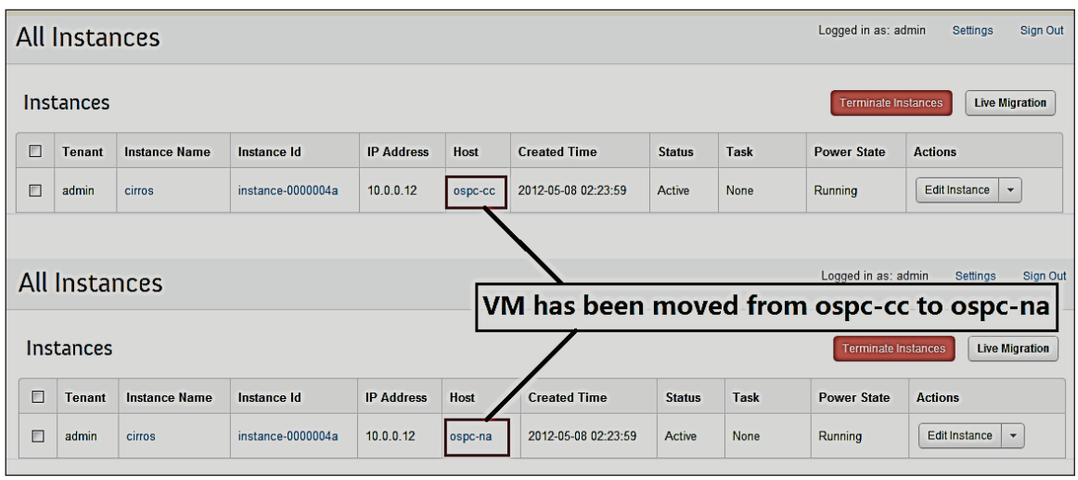

Figure 5: Policy-based VM migration—VM instances (Source: OSPC)

Power Management Use Case 3: Optimize Rack Density (Policy-Based Power Capping)

Traditionally, we can only estimate the power consumption data from the manufacturer's specifications. This requires the allowance of a hefty safety margin, and thus results in overprovisioned data center power, overcooling of IT equipment, and increased TCO.

The availability of power monitoring data allows management by numbers, which tightly matches servers by power quotas to available data center power. The use case is useful in older data centers underprovisioned for power and in host settings with power quotas in effect. Therefore, it can optimize the rack utilization and increase the data center density.

Figure 6 shows that the power utilization of the server stays within the threshold limit set from the server, which is 160 watts. In the earlier use case we observed that without any power capping, the power utilization of the server went up to about 200 watts

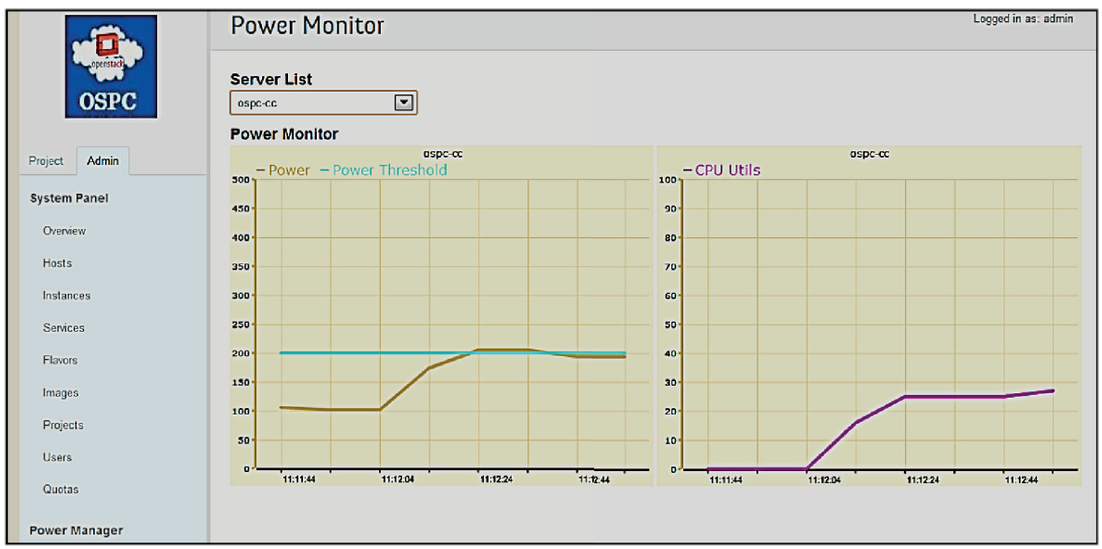

Figure 6: Policy-based power capping (Source: OSPC)

Security

Recent cloud computing customer surveys unanimously cite security, control, and IT compliance as primary issues that slow the adoption of cloud computing. Many customers have specific security requirements that must assure data location and integrity, and today they're using legacy solutions that rely on fixed hardware infrastructures. However, the means in legacy solution to verify a service's security compliance are labor-intensive, inconsistent, and nonscalable. For this reason, many businesses only deploy non-core applications in the public cloud and restrict sensitive applications to dedicated hardware.

Comprehensive security requires an uninterrupted chain of control from the application user's interfaces to the underlying hardware infrastructure.[6] Any gaps in this trust chain render them vulnerable to attacks. Intel® Trusted Execution Technology (Intel® TXT) provides hardware-based technologies to establish a root of trust that provides the necessary underpinnings for successful evaluation of the computing platform and its protection. It is specifically designed to harden platforms from the emerging threats of hypervisor attacks, BIOS, or other firmware attacks, malicious root kit installations, or other software-based attacks. Intel TXT gives IT and security organizations important enhancements to help ensure: more secure platforms; greater application, data, or virtual machine isolation; and improved security or compliance audit capabilities.

By providing controls to ensure only a trustable hypervisor is running on a platform, Intel TXT helps protect a server. While this basic protection and enhanced control is good on individual systems, it becomes even more powerful when one considers aggregated resources and dynamic environments such as cloud environment. Intel TXT can help create something known as *trusted computing pools*. In this usage model, a pool of trusted hosts can be created, each with Intel TXT enabled and by which the platform launch integrity has been verified. Data center administrators can set a policy that VMs with high requirements on host trustiness can be only scheduled on or migrated between the hosts in the trusted computing pool. This enables data center administrators to restrict confidential data or sensitive workloads to platforms that are better controlled and have had their configurations more thoroughly evaluated through the use of Intel TXT-enabled platforms.

We have enhanced OpenStack to use Intel TXT to implement the trusted computing pool usage model. See Figure 7 for the detail flow:

- In a cloud computing environment, a subscriber, who requires his VM to run on a trusted platform, can specify the trust level of that VM as Trusted.
- The request will pass along all the way to OpenStack Nova scheduler.
- The scheduler will invoke a web-based remote attestation (OpenAttestation) service to decide the platforms' trustworthiness.
- Based on the results, the scheduler will schedule the Trusted VM to one of the trusted platforms.

In Figure 7, the web-based OpenAttestation service is a standalone open source project in BSD license[7] (https://github.com/OpenAttestation/). It uses platform measurement credentials to complete the trust verification process and support compliance and audit activities. The implementation takes advantage of TCG (Trusted Computing Group) Infrastructure Work Group's Integrity Report Schema Specification. As a standalone SDK, OpenAttestation enables ISV software to remotely retrieve and verify target hosts' TPM PCRs and verify hosts' integrity, through exported Query API. This SDK is a critical open source building block upon which ISVs can build their Intel TXT–based security solution more easily.

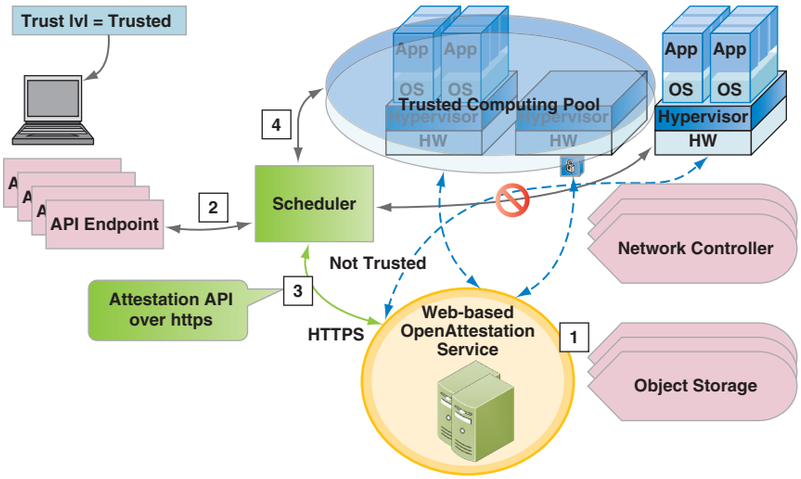

Figure 7: OpenStack trusted computing pool (Source: Intel Corporation, 2012)

Conclusion

As IT moves into the era of "new IT," cloud and related technologies are moving quickly. Many factors including cost constraints and consumerization of IT are driving these changes. Open source projects, including OpenStack, are leading the way to solutions for IT address these challenges. As one proof point, Intel IT has implemented an open source private cloud in a production environment based on OpenStack and other open source components. In addition to the challenge of integrating the various software elements into a solution, the Intel IT solution also shows the power of integration back into an existing enterprise IT organization.

Of the challenges faced by IT, power management and security are high on the list. Power is often ignored or misunderstood since the power bill is often buried in another part of the organization's budget. Through the use of Intel Node Manager, key use cases such as meeting operating cost constraints, physical limitations of power and thermal, and poor visibility into the actual power and thermal environment can be overcome. In the scope of security challenges, providing support for compliance and audit are important considerations. The Intel Trusted Execution Technology (Intel TXT) is a powerful tool to provide hardware based root of trust. Combined with remote attestation, Intel TXT provides a strong basis, rooted in hardware, for compliance and audit scenarios.

OpenStack has very strong industry momentum and enjoys an active and innovative community. Open source technologies, including OpenStack are well positioned to lead the market for solutions to key IT challenges utilizing Intel technology.

References

[1] *http://wiki.openstack.org/*

[2] *http://openstack.org/community/companies/*

[3] *http://www.openstack.org/projects/essex/press-release/*

[4] *http://www.intelcloudbuilders.com/index.php?option=com_productsearch&view=displaysearch&&lang=&filter_companyname=1322&filter_category=&filter_efficiency=&filter_region=&company_name=OpenStack&Itemid=181*

[5] *http://www.intel.com/content/www/us/en/data-center/data-center-management/node-manager-general.html*

[6] *http://www.intel.com/content/www/us/en/trusted-execution-technology/trusted-execution-technology-security-paper.html?wapkw=intel+trusted+execution+technology*

[7] *https://github.com/OpenAttestation/OpenAttestation*

Author Biographies

Billy Cox has been leading the cloud strategy efforts for Intel's Software and Services Group since 2007. He is also responsible for Intel's Cloud Builders program, which brings together cloud leaders to provide best practices and practical guidance on how to deploy, optimize, and support a cloud infrastructure. The program offers detailed reference architectures and best practices to build simplified, secure, and efficient cloud infrastructure. Previously, Billy was Director of Systems Engineering at HP for 14 years where he was responsible for the development of all infrastructure management tools used to manage the various server and storage platforms. With over 30 years of industry experience, Billy has led the design of compute, network, and storage solutions and actively participated in multiple standards efforts. His e-mail address is billy.cox at intel.com.

Jiangang Duan manages the Scalability Lab in Intel Asia-Pacific Research and Development Ltd. He has worked on enterprise solution tuning and optimization for more than ten years, including several generations of Intel processor performance evaluations and has worked jointly with local OEMs to complete industry benchmark publications. He now focuses on cloud computing and virtualization, and is responsible for development, deployment, and optimization of efficient cloud data centers with open source software (Xen/KVM/OpenStack). Before joining Intel, Jiangang got his bachelor's degree in 1999 and Master's degree in 2001 from the Electrical Engineering Department of Tsinghua University. His e-mail address is jiangang.duan at intel.com.

Fleming Feng is Chief Open Source Scientist in the Open Source Technology Center PRC team. He joined Intel in 1997 and in 2000 he started to work on open source software development from Carrie Grade Linux project; he then worked on various open source projects. His current role is mostly working as technical interface for PRC local engagement on government collaboration, the open source community, and business engagement based on open source solutions. His e-mail address is fleming.feng at intel.com.

Das Kamhout is a principal engineer in Intel IT responsible for the architecture, strategy, and execution of the Intel IT Cloud Journey. In his 15 years at Intel he has been responsible for everything from the clients that Intel IT uses, to the architecture that Intel's Design Grid runs upon. His e-mail address is das.kamhout at intel.com.

Susie Li is an engineering manager in the Open Source Technology Center in Shanghai, PRC. She joined Intel in 1999 and had been involved in IPMI, EFI, virtualization, and cloud projects. She currently manages the Cloud and Virtualization team in the Intel Open Source Technology Center. Her team is focusing on enabling Intel technologies in open source virtualization and cloud stack (Xen, KVM, and OpenStack). She received her Bachelor's and Master's degrees in Engineering from Shanghai Jiao Tong University, PRC. Her e-mail address is susie.li at intel.com.

Tonny Liu is an engineering manager from IT Flex China team. He and his team have been working on cloud computing, mobile development, and enterprise solutions for more than eight years. Tonny got his Bachelor's and Master's degrees from the Electrical Engineering department of Southwest Jiao Tong University in 1998 and 2002 respectively. He personally holds ICCP and PMP certifications. His e-mail address is tonny.h.liu at intel.com.