Intel
Intelligent
Power
Management
Intel®

How High Temperature Data Centers & Intel® Technologies save Energy, Money, Water and Greenhouse Gas Emissions

Power savings through the use of Intel®'s intelligent Power Management (Node Manager and Data Center Manager) enables a future High Temperature Data Center Environment.

Intel® Corporation

September 2011

DRAFT PUBLIC

Contents

Executive Summary

Intel®'s Power Management Technologies known as Node Manager (NM) and Data Center Manager (DCM) in combination with High Temperature Ambient Data Center operations was jointly tested over a 3 month Proof of Concept (POC)with KT at the existing Mok-dong Data Center in Seoul, South Korea. The objective was to maximize the number of servers compute nodes within space, power and cooling constraints of the data center (DC).

Intel® Intelligent Power Node Manager	High Temperature Ambient (HTA)
Intel® Intelligent Power Management is an Intel® platform-software-based tool featuring that provides policy-based power monitoring and management for individual servers, racks, and/or entire data centers.	Raising the operating temperature within the computer room in a data center decreases chiller energy costs and increases power utilization efficiency.

Table 1 – Intel® Intelligent Power Node Manager and HTA explained.

The POC proved the following:

- A Power Usage Effectiveness (PUE) [details in glossary]of 1.39 would result in approximately 27% energy savings at the Mok-dong Data Center, in Seoul. This could be achieved by using a 22°C chilled water loop.
- Node Manager and Data Center Manager made it possible to save 15% power without performance degradation using control policy.

DRAFT PUBLIC

- In the event of a power outage, the Seoul DC Uninterrupted Power Supply (UPS) uptime could be extended by up to 15% whilst still running its applications with little impact on the business Service Level Agreements.
- Potential additional annual energy cost savings greater than $2,000USD per rack by putting an under-utilized rack into a lower power state by implementing Intel® Intelligent Power Node Manager for power control.

This paper describes the procedures and results from testing Intel® Intelligent Power Management within a High Ambient Temperature operation Data Center at the Seoul Data Center.

Methodology

The High Temperature Ambient Data Center Operations Intel®'s Intelligent Power Manager POC's use Intel®'s Technical Project Engagement Methodology (TPEM). This is not meant to be a detailed look at each step of the mythology but a guideline to the approach.

#	Description	HTA	NM/DCM
1	Customer Goals and Requirements	√	√
2	Data Gathering	√	√
3	Design DC room constructed Server platform chosen	X	X
4	Instrumentation HTA- Covered under data collection	X	√
5	Design Test Cases	√	√
6	Run Test Case	√	√
7	Data Collection	√	√
8	Analysis	√	√
9	Data Modeling	√	√
10	Reporting and Recommendations	√	√

Table 2 - uses Intel®'s Technical Project Engagement Methodology (TPEM).

Software Tools

Three tools were used to measure and model KT environment:

1. Intel® Data Center Manager/Node Manager: (NM/DCM) was used to collect data on the workload, the inlet temperature and the actual power usage (both idle and under workload).

2. 3D Computer Room Modeling Tool was used to create a 3D virtual computer room that visually animates the airflow implications of power systems, cables, racks, pipes and Computer Room Air Conditioning (CRACs). This assisted in evaluating the thermal performance options of the different conceptual designs.

3. Data Center Architectural Design Tool additional software was used that enabled Intel® to accurately predict data center capacity, energy efficiency, total cost, and options including the impact of geographical location on the design for KT

Business Challenge

Increasing compute capabilities in DC's has resulted in corresponding increases in rack and room power densities. KT wanted to utilize the latest technology to gain the best performance and best energy efficiency across their Cloud Computing Business offering. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (seems to be hanging)

Table 3 - ASHRAE's Technical Committee (TC) 9.9, Mission Critical Facilities, Technology Spaces and Electronic Equipment

> ***Option 1*** *– use IT equipment optimized for a combination of attributes including energy efficiency and capital cost with the dominant attribute being RELIABILITY*
> ***Option 2*** *– use IT equipment optimized for a combination of attributes including some level of reliability with the dominant attribute being ENERGY and compressorless cooling (free-Cooling)*

KT chose to go with Option1 for the short term with the goal to move toward Option2.

Technical Challenge

KT's Mok-dong DC has been already been constructed with the floor dimensions construction was completed. (rephrase, 2 construct) Therefore the power & thermal optimization problem becomes primarily one of thermodynamics, i.e.: dissipating heat generated by the servers. This is done by transferring the heat to the air and then ducting hot exhaust air from the server. This hot air is then passed through one of the CRAC units where it is cooled by the chilled water (CW) and blown back into the computer room to support the desired ambient air temperature, or 'set point'. The CW is supplied in a separate loop by one or more chillers located outside the computer room.

The Seoul data center has a 5MW maximum power capacity. This must cover both server power and Data Center facilities cooling. This means that the more efficient the cooling infrastructure, the more power is available for the actual compute infrastructure (review PUE).

At KT there was 1,133.35 m^2 total fixed available space for IT equipment in the data center's computer room; space is another key metric for computer room layout.

Node Manager and Data Center Manger

Intel® Intelligent Power Node Manager provides users with a powerful tool for monitoring and optimization of DC energy usage, enhancing cooling efficiency, and identifying thermal hot spots in the DC. It can provide historical power consumption trending data at the server, rack and Data Center. For KT, Inlet thermal monitoring of node temperatures was a key capability to identify potential hot spots in the DC in real time.

Test Environments

The test environments were set up in two distinct configurations; Business and Lab environments – respectfully.

Devices	Descriptions
Server Platform	2* 5640 CPUs @2.66GHz, 48GB memory, Intel® NM v1.5
	Business environment for power monitoring: 2 racks (34 servers)
	Lab environment for power/thermal monitoring and power management 1: rack (18 servers)
Server OS	Business environment: Xen (v5.6.0) in
	Lab environment: Windows Server 2008
Tools	Power management tool: Intel®® Data Center Manager v2.0
	Workload simulation tool: SPECpower_ssj2008
	Infrared temperature meter; power meter

Table 4 - NM/DCM Test Environment

The uses cases used for the POC were as follows:

Power Monitoring and Power Guard

The 'Power Monitor and Power Guard' test scenario was designed to record real-time power monitoring and power policy management on server platforms. DCM can be used as a tool to balance useful workload across the DC. Workload and power management can be used to dynamically move workload and power in the DC where it is most needed.

Test Case #	Test Case	Descriptions
TC1.1	Power monitoring	Power monitoring at the server/rack/row level.
TC1.2	Performance-aware power optimization	Power optimization without workload performance degradation or with limited workload performance impact.
TC1.3	Priority based power optimization	DCM group power resolution algorithm lets high priority nodes receive greater power when power is limited.

Table 5 - NM/DCM Power Monitoring and Power Guard

Analysis of these uses cases demonstrated that KT could increase server utilization and improve power efficiency by increasing overall server usage thru the implementing of Virtual Machines (VM). Consolidating VMs would enable a power management policy that would automatically to decrease rack power consumption of individual nodes and servers when not in use. When these racks are in an idle and therefore low power consumption state they can still be quickly brought to full power by removing the power policy.

Additional results from Test Case 1.2 were a PUE of 1.39 and unit energy cost of US$0.07/KWh, the annual energy cost saving by changing the power setting of a rack would be US$2179.
(=2.369KW*1.39*24*365*$0.07)

Thermal Monitoring and Thermal Guard

The purpose of the 'Thermal Monitor and Thermal Guard' test scenario was to increase data center density under the constraint of cooling limitation

Test Case #	Test Case	Descriptions
TC2.1	Thermal monitoring	Inlet temperature monitoring at the node level
TC2.2	Thermal based power management	Power management under inlet temperature trigged event

Table 6 - Thermal Monitoring and Thermal Guard

As shown in Figure 9, the node's inlet temperature is kept at 19°C-20°'C with a boundary maximum value of 21°'C and a minimum value of 18°C that is close to the DC set temperature. This information can help the DC administrator manage cooling resources in the DC if there are abnormal thermal events occurred.

Business Continuity

The purpose of 'Business Continuity' test scenario was to prolong service availability at power outage.

Test Case #	Test Case	Descriptions
TC3.1	Business continuity at abnormal power event	Minimum power policy to prolong business continuity time under power outage

Table 7 - Business Continuity

Overall Node Manager and Data Center Manager Key Test Results:

Test Case Id	Benefits Description
1.2	Up to 15% power savings under a 50%-80% workload level without performance degradation through performance-aware power control policy;
1.3	Successful priority-based power control by allocating more power to workloads with higher priority under the power-limited scenario;
2.1&2.2	Alarms triggered when abnormal power/thermal events occur;
3.1	Prolonging the business continuity time up to 15% for 80% workload level when there is power outage;
2.1&2.2	Reducing the heat generation at a node when the inlet temperature exceeds the pre-defined thermal budget.

Computation Fluid Dynamics

The three configurations were reviewed and 208 Rack Computational Fluid Dynamics (CFD) was modeled. After the data was gathered, Intel® conducted detailed analyses of the three configurations. Intel® recommended configuration 3 based on the following:

CFD simulations modeled all three scenarios and found cooling capacity for 208 racks at 6.7kw could be supported. 208 racks will consume an IT load of 1.3964MW (6.7 KW X 208 racks) installed capacity (running load) and the nameplate or provisioned load (the worst case) of 1.664MW (208x8KW). There were 17 CRACs in an N+1 configuration. Cooling load is running at 73KW to 94kw out of a rating of 105KW. Supply temperature was 22°C and return temperature varies from 31°C to 36°C. Table 4 provides the detail system descriptions.

System Description	Analysis
Equipment Load	208 racks rated at 6.7kw per rack. Supply air is at 22C
CRAC Cooling	There are 17 CRACs in an N+1 configuration. Cooling load is running at 73kw to 94kw out of a rating of 105kw.Supply temperature is 22C
DC Ambient Temperature	The hot aisle containment area averages around 36C while the cold aisle and the rest of the DC averages 22°C.
Airflow	208 racks at 6.7Kw : The total airflow required for 208 racks is 188,241 CFMs which is within the capability of the 17 CRACs.

Table 8 - System Description

The CRACs were run at about 85% utilization and the conclusion was that 208 racks could be supported by the 17 CRAC units at 6.7kW/rack. To meet this cooling load; the CFM model delta T varied between 11°C to 14° and would not meet the CRAC unit Delta T of 10°C.

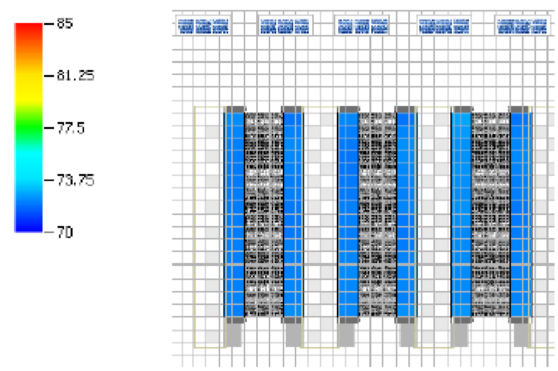

Figure 1 - view shows no issue with rack inlet temperatures

Intel®'s Intelligent Power Management
The CFD tool tested the 208 6.7kW/rack configuration. The sample clip above (extract from the larger CFD) clearly shows the Data Center environment at blue or approximately 21°C (70°F). This view shows no issue with rack inlet temperatures.

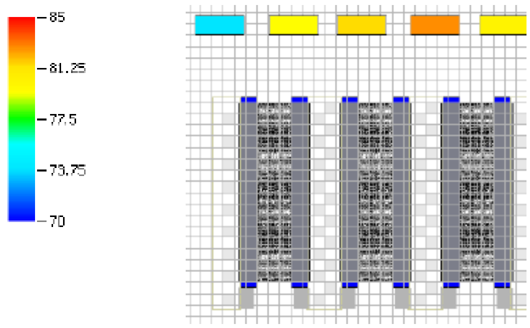

Figure 2 view shows the ACU cooling load (perimeter units only) ranging from 85-94kW (maximum of 105.6kW), so these units are operating at a range of 80-89% of capacity, with the aid of the in-row coolers.

The Air Conditioning Unit cooling load (perimeter units only) ranging from 85-94kW (maximum of 105.6kW), so these units are operating at a range of 80-89% of capacity.

High Temperature Ambient Data Center Simulations

With the CFD simulation showing a number of total racks supported, Intel® investigated additional ways to optimize the Data Center. After collecting data from key DC systems Intel® was able to focus an area that had the greatest opportunity for optimization and cost savings. The tool (see section heading - Data Center Architectural Design Tool) used to collect the data produces a visual report such as Figure3 below:

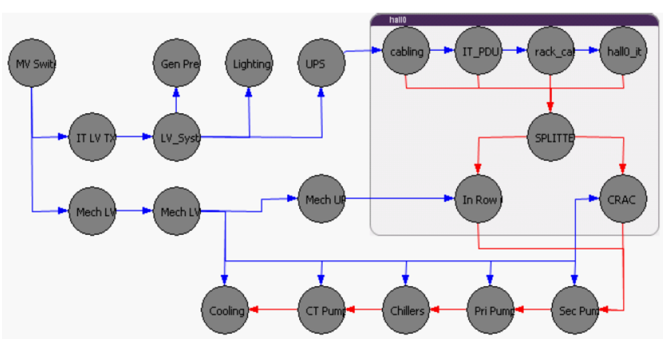

Figure 3 -Data Center Architectural Design Tool Graphic

The Data Center data gathering and investigation process identified the Chilled Water (CW) loop as the area for the greatest potential for improvement.
The results of Intel's analysis - there are significant chillers and CRAC improvements possible by utilizing a 22°C water loop (either with or without economizer). Using 22°C chilled water loop, the PUE is improved to 1.39. The improvements are mainly resulting from energy reduction by chillers and CRAC with 27% overhead energy reduction, as compared to the baseline (7°C). The best efficiency is achieved by 22°C chilled water loop with economizer with PUE=1.30 resulting from 43% overhead energy reduction. Even thought there is a better PUE to be gained by using economizers there were KT ruled it out due to:

1. Space constraints within the existing site, and;
2. Capital cost of an economizer at this late stage of the DC project.

Overall Results

NM/DCM, CFD modeling and HTA Data Center simulation POC has clearly shown the opportunity for optimized operations to maximum power and cooling efficiency savings.

NM/DCM should be activated on the existing servers and the use of Data Center Manager 2.0 is recommended to manage the data center operations based on node level inlet temperature and adjustments to computer room cooling based on actual cooling needs. As the racks are installed into the computer room, NM/DCM will allow KT to monitor the impacts on cooling as the additional servers are added to the existing data center.

When implemented in future DC, most notably when constructing the future outside of Seoul Data Center the data suggests the climate in Korea is highly conducive to optimized wet side and air side economizer's designs. These designs would be the cornerstone of the high ambient temperature setting and have been extensively explored in the modeling. The expectation is of PUE in the 1.05 to 1.1 range is achievable. This report will be a key reference architecture which will yield substantial power savings.

Glossary

Power usage effectiveness (PUE)

$$PUE \frac{\text{Total facility power}}{\text{IT equipment power}}$$

PUE is a measure of how efficiently a computer data center uses its power; specifically, how much of the power is actually used by the computing equipment (in contrast to cooling and other overhead – lower PUE is better).

BMC	Board Management Controller
CDC	Cloud Data Center
DC	Data Center
DCM	Intel® Data Center Manager
NM	Intel® Intelligent Power Node Manager
POC	Proof of Concept
VM	Virtual Machine
AHU	Air handling Unit, used inter-changeably with CRAC
ASHRAE	The American Society of Heating, Refrigerating and Air-Conditioning Engineers
CAPEX	Capital Expenditure
CFD	Computational Fluid Dynamics
CRAC	Computer Room Air Conditioner Unit, used inter-changeably with AHU
DC	Data Center
DCIE	Data Center Infrastructure Efficiency
Delta T	Delta Temperature typically refers to supply and return temperature of cooling systems or server temperature.
HTA	High Ambient Temperature
LV	Low Voltage
MIPS	Million Instructions Per Second
MV	Medium Voltage
ODM	Original Design Manufacturer
TX	Transfer