Future proofing EGSE- is Virtualization the answer?

11th Int. WS on Simulation & EGSE facilities for Space Programmes
SESP 2010

28-30 September
at ESTEC, Noordwijk, The Netherlands
Topic – AIT Process

A.B. Armitage[(1)], K. Leadbeater[(1)], A. Polverini[(1)], R. Patrick[(2)]

[(1)]TermaB.V., Schuttersveld 9, 2316XG Leiden, The Netherlands
21)TermaA/S., Vasekær 12, DK-2730 Herlev, Denmark

ABSTRACT

This paper looks at how some newer developments in the information technology sector can be applied to EGSE/CCS systems. The aim is to improve the lifetime of systems, operations costs and improve the flexibility of Central Checkout Systems. This would apply particularly when addressing the needs of long term programs (Galileo), production Galileo) or when a CCS is considered part of an infrastructure to support many missions over a long period. The technologies in question include: - use of virtualization development and deployment of the systems - use of blade servers to provide a very high availability and efficiently packaged computing resource - use of thin clients management overhead of these systems These technologies were first investigated by Terma when addressing the needs of a large production line i.e. Galileo FOC AIT/feasibility studies and prototypes developed by Terma were so successful that they have also been deployed on systems to support other missions as well. This paper technologies identified above, describing the benefits of applying them in the CCS environment: Some examples include - Virtualization – allows old systems (OS et al) hardware, thus allowing users to benefit from new hardware without he need for software upgrades to achieve it. It also allows developers to rationalize dramatically use for developing and maintaining systems, thus keeping maintenance costs down. - Blade Servers – coupled with the correct disc drives and network topology, this increase of systems, with many hot redundant features that are available at very modest cost. They also can be deployed in very “dense” configurations – which is important in Thin clients – these replace traditional workstations as the MMI to the system. The system management overhead of these is far less, thus simplifying the system management From this, we have identified a number of different deployment alternatives that ensure that the CCS is “right-sized” for its designated use. The operational software (run unmodified on a great range of computer hardware and this feature is exploited here. Having been working with these technologies now for 2 years, including deploying operational, there were of course some problem areas, and it is worth to highlight these as well. We conclude that this approach is a major step forward for both developers We also identify other space applications – namely Mission Control Systems – where we believe the benefits are even greater.

INTRODUCTION

This paper outlines the benefits resulting from the use of four new technologies for the implementation of Central Checkout Systems which form the central part of an EGSE system. The technologies are:

- Blade Servers
- Virtualization
- Thin Clients
- Storage

The advantages of these technology is even greater when applied to similar systems with even higher availability requirements such as spacecraft control systems.

WHAT IS A BLADE?

A "blade" is a computer server normally mounted in a "blade enclosure". It is essentially a packaging concept developed for server farms allowing a very large number of servers to be concentrated into a small space. The packaging makes more efficient use of power and cooling than traditional rack mounted systems.

Server farms typically use virtualization software to give their customers the appearance of a completely private computer. The customer often uses "thin client" hardware or software to connect to the virtual machine running in the server farm. For a distributed or mobile, organization, the maintenance of central virtual applications is very cost efficient compared with applications installed on the desktop/laptop of every user.

The server virtualization industry has developed management tools that help to ensure continuity of service and easy maintenance.

We are now describing the advantages of this approach for spacecraft testing (Central Checkout Systems).

Blade HW Approach

A blade hardware architecture provides for optimal availability: all hardware within the enclosure is redundant (blades, network switches, power supplies, fans) however the shared infrastructure means that unnecessary duplication (e.g. fans and power supplied between servers) is avoided.

The blade "fabric" provides a backplane for all internal communication. This internal fabric is devoid of active electronic components or moving parts to ensure maximum reliability. All connection paths in the fabric are dual redundant.

The internal network infrastructure is set up in a cross-strapped dual hot redundant configuration so that while both network switches are operating, the full capacity of both redundant links is used, whereas if one network link or switch fails, all traffic can be routed via the remaining working switch. The blade management node allows for remote management and failure prediction.

The SAN internal or external to the blade enclosure provides a fully hot redundant self-rebuilding RAID array, with dual path so that system operation on all servers can continue after multiple successive failures of single disks.

This solution is environmentally efficient. Sharing of cooling fans and power supplies reduces both the materials used and power consumption compared with a traditional racked server system. The individual blade servers operate on surprisingly low power (e.g. 75 to 80W).

Benefits of Virtualization

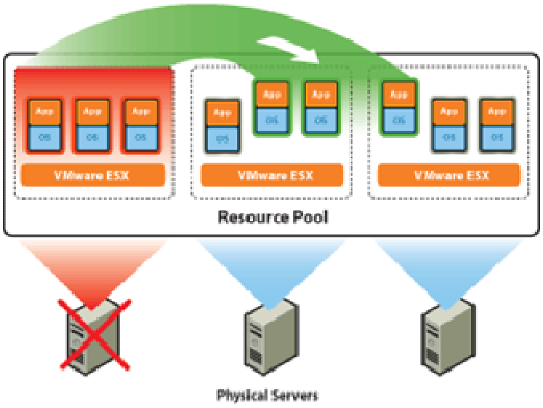
Virtualization allows one or more operating systems to be hosted by a physical machine. The physical and virtual machines are connected to a SAN (Storage Area Network) array of RAID protected hard disk drives. This SAN is partitioned in to private or shared disk areas that simulate real hard disk drives. The network layer is also virtualized.

Virtualization of the CCS servers provides several benefits, which are outlined below.

- A single physical server can host multiple virtual servers. As a virtual server becomes more busy, it can grow to take up more CPU resources on the physical platform on which it resides. The level of activity can be monitored, and if desired, limits can be placed on resource utilization, or virtual servers can be migrated between physical servers to optimize load balancing.
- In case of planned maintenance, a running server can be migrated seamlessly from one hardware platform to another, while still running. Only a few milliseconds interruption can be noticed in the network traffic during migration.
- After a migration, the original hardware platform can be powered down (from the server management application) and physically removed from the blade.

- In the same way, new physical server blades can be added to the enclosure, and the existing workload can be migrated to balance the load. Individual virtual machines can be migrated automatically when required for load balancing. Automatic migration can also be triggered by failure prediction.

Of course, during migration, no physical network disconnection or hardware disassembly is required.



High Availability

In case of an unplanned hardware shutdown due to CPU or memory failure, HA (high availability) can be configured so that an image of the failed server is automatically booted on another physical host within seconds of the hardware failure. This approach does not provide completely seamless uninterrupted migration in case of failure, but does allow the failed server to be back in action, automatically within a few minutes.

Full seamless high availability is expected to be a supported feature of VMWare during 2009, however it should not be assumed that this is necessarily the most appropriate solution: typically during long tests such as TBTV, a second CCS Lite is connected to the TMTC FE but no other SCOE, in order to ensure that a hardware failure does not allow loss of control or monitoring of the spacecraft. A second virtual CCS server connected to the TMTC FE would provide the equivalent high availability feature within the same server enclosure, allowing the primary CCS to be recovered while the spacecraft is still being monitored and able to be commanded. This approach to high availability, supported by a virtual server architecture could be more simple and effective than full literal server hot-standby.

Maintainability

Centralizing the storage of application software ensures that software upgrades can efficiently be deployed in a single central location without repetitive installations on many physical servers, or workstations.

New virtual servers can be created or “cloned” within minutes. If desired, the system can provide a physical platform for hosting other virtual servers running different operating systems from the CCS software. For example the customer may wish to install his own software applications such as an SRDB site, or specific user-preferred checkout applications. This is possible even if the operating system is not compatible with that of the CCS.

Server snapshots can be created to create a software image of the complete disk image of the virtual machine. This can be used as a means to create “instant backups”.

Server virtualization can help to avoid possible long term maintenance “deadlocks”. Consider the situation where application software and operating system can only run on a specific hardware platform which cannot be upgraded without major changes. In a non-virtual deployment the system can become stuck on an old hardware and operating system. For a long duration mission this can become a serious issue. In a virtual deployment, new version software can be run on one virtual server, while the previous version is still available on another. Alternatively a virtual server can continue to be supported on its original operating system after the hardware platform has been modernised.

Network Virtualization

The network approach (using VLAN’s) allows networks to be partitioned into virtual networks (such as EGSE, CCS LAN and Site LAN). Each of these LANS can be assigned dedicated IP address spaces. Via the physical blade fabric and partitioned virtual networks each virtual server can be connected to one or more VLAN without the need to perform any physical reconnection.

Network management running on the Virtual Center server gives a complete overview of the network links status and activity. The system manager can activate, deactivate or repartition LANS from a single graphical application.

Security

For missions where security or secure storage of spacecraft data is an issue, the blade approach allows central control of all data access.

Possibilities offered by the blade approach are:

- all network interaction with the thin client is securely encrypted
- possibilities to save data to USB stick on the thin client can be limited
- networks can be partitioned as isolated VLAN from the general CCS LAN
- in a SAN it would be difficult for a "blind thief" to find out which physical hard drive contained the actual sensitive data.
- storage of sensitive test results can be automatically encrypted by the hosted operating system

It would generally be recommended for secure missions that:

- the blade enclosure should be installed in a locked area with access only by authorised personnel
- the system management personnel should be security cleared to a level where they can be trusted to control the secure partition of data and networks within the CCS blade

A lockable enclosure ensures that the system is safe from unauthorised tampering or removal of any of its internal components.

THIN CLIENTS

Thin Clients are simple desktop based computers which allow remote connection to a server using one or other remote desktop protocol. The only software running locally on the thin client is that needed for the local user interface (in addition a web browser may be provided).

Thin clients are typically very compact, fitting on a desktop behind, next to, or even inside the user's display. They contain no moving parts (no hard drive, only flash memory) which means they are extremely robust, and can be dropped or knocked without damage.

A thin client uses very little power and is self-cooling which means that it is completely silent. This very low energy and materials use means that this approach is much more environmentally efficient than a traditional workstation approach.

Since there is no application software installed, Thin Clients are extremely flexible. The work required to shut down and move a thin client between two servers is zero. In the same way, a thin client can very easily be moved around inside the AIT room. During operating system or application software maintenance to the central CCS application, nothing needs to be done on the desktops.

Depending on the protocol used, a session can be stopped and resumed on the thin client without disturbance to the user session. This means that a thin client can be turned off, and when restarted, the user's session continues exactly where it left off (window positions, application state exactly as before). One such protocol is NX, and for this reason Terma prefer to use the NX protocol between thin client and CCS servers.

Access to SCOE MMI

The thin client need not be restricted to accessing the CCS, for example any Windows™ based SCOE, Instrument EGSE or front end equipment will support Remote Desktop Protocol (RDP), and most LINUX based EGSE will support NX, VNC or X11 connection.

Any of these can be used from a thin client. This allows several EGSE desktops to be accessed from one user interface without the operator moving from his seat.

A specific thin client may be set up so that whenever it is switched on, it automatically connects to the CCS with a specific login, or connects to a specific SCOE. This thin client can be considered "dedicated" to a specific setup. If the "dedicated" desktop is ever stopped, or disconnected then it can be configured to automatically restart. This could be appropriate for setting up a central test monitoring station.

Smart Cards

The thin clients allow the use of Smart Cards. This provides several benefits for system security, privilege management, and usability.

Each AIT user may be given his or her smart card that has been programmed with a set of sessions to start, user logins, and screen layout (e.g. CCS, front end, database, etc). This may be fixed to the user's allowed privileges (e.g. conductor, observer, administrator). The smart card may be programmed to require a password or PIN code as desired.

The user can then insert his Smart Card in a thin client in the AIT Room, and his own personalised desktop appears. Having performed some checks, he can then remove the Smart Card (immediately closing the thin client session, and freeing the session for other team members). If he then walks into the Clean Room and inserts his Smart Card into any free thin client inside the Clean Room, his user interface appears again exactly as it did (i.e. windows identical, cursor in the same place, etc). He can then continue immediately.

The same applies when turning off the thin client at the end of a day. After power on, a smart card records the sessions that shall be restarted, and re-establishes the user exactly where he finished work.

Alternatively Smart Cards can be used for certain roles, for example: "the test observer" role, or "the instrument expert" role.

The customer is free to choose how he wishes to organize these matters, and to what level of security they shall be managed. The hardware on each thin client allows cards to be programmed, however the privilege to program the cards can be globally allowed or can be protected by setting a "company key" so that only those knowing this password are able to program the smart cards.

A smart card need not be required, for example a thin client can the set up to continuously display a fixed set of sessions, and if desired restart them whenever they are terminated.

ENVIRONMENTAL CONSIDERATIONS

Robustness & Physical Factors

The CCS shall usually be operated in a clean room or AIT room environment in the vicinity of the SCOE, instrument EGSE and front-end equipment, some of which may be inside a clean room whereas the CCS may be outside. The CCS hardware shall need to operate almost continuously for several years.

It may be transported or moved around several times, and may be used in an environment where dust accumulates.

Physical shocks may come from movement over slightly uneven surfaces, accidental collision with other EGSE equipment while being moved, or impacts caused by cleaning staff. This environment is not extremely hostile, but nevertheless a robust and durable system is called for.

Note that some critical tests in the spacecraft AIT are at the end of the campaign. This is also the time when reliability problems with the CCS hardware may be more frequent due to increasing age. In this situation, a maintainable modular CCS becomes even more valuable.

The transportation and storage environments are more extreme but in these cases the CCS shall be stored in robust reusable transportation cases.

The blade approach ensures that equipment exposed to the end users on a day to day basis (thin clients) is extremely robust. More sensitive equipment can be centralised and secured. In general, equipment in the central blade enclosure is available in hot redundant configuration.

Energy Use

Given the long duration of some AIT campaigns, and the almost continual power use of the equipment, the energy efficiency of the CCS is a factor in its design.

The blade server approach avoids inefficiencies coming from duplicated redundant power supplies cooling fans and network equipment. This bring benefits to materials, bulk, noise and energy use

Audible Noise

The noise level generated by cooling fans of the CCS can be minimized to a level that does not disturb the concentration of the AIT engineers using the system. This depends on the specific enclosure system and the manufacturer. In some very large systems it may be necessary to install the blade server in a separate enclosed room for noise as well as for security reasons.

EU Standards Considerations

All CCS hardware considered is commercial off-the-shelf equipment, CE certified compliant by their suppliers to Low Voltage Directive (73/23/EEC, 2006/95/EC) and Electromagnetic Compatibility (2004/108/EC). Use of standard equipment, without performing any custom assembly ensures that the compliance with these standards at CCS level is automatic.

EU standards and proposals increasingly take into consideration the overall environmental impact (energy use) and the use of materials (e.g. hazardous chemicals) and disposal of the equipment at end of life (2002/96/EC).

Virtual CCS Deployment

For the future it could perhaps be considered to eliminate the CCS server hardware delivery altogether and to host the CCS software from a central site.

A “pay per user hour” cost model with service level agreement could be adopted. The service could be measured from the time of smart card insertion to the time at which it is removed. The advantage to the end user could be that costs could be avoided in case of mission delays, holidays, or quiet periods in the AIT. Any downtime would of course free of charge (or possibly involve penalties).

The fully virtual approach could avoid: travel for onsite commissioning and software updates, and equipment packing, unpacking and transportation

Another advantage would be to allow users to log in from various sites in parallel leading to a reduction in travel.

Such a configuration would require highly robust networks to be included in the service. For current missions this approach is probably too radical to be trusted. For example could this approach be trusted during a TBTV test, or for checkout on the launch pad in Kourou or Baikonur? On the other hand this kind of approach is already being used for business and safety critical applications.

CONCLUSION

For very large scale checkout applications, such as for large spacecraft fleet production lines, where there is a single production facility, there are immediate technical and cost benefits to use a centralised blade architecture. In these types of mission a fully coordinated approach with SCOE providers can also lead to optimizations of the overall EGSE infrastructure (e.g. Thin Clients used to access all SCOE).

For smaller, more distributed missions, the conclusion is less clear, because there is a procurement cost for multiple blade enclosures and virtualization licenses. In this situation the decision whether to use blade servers depends on the priority given to:

- *maintainability*
- *expandability*
- *robustness*
- *environmental issues*

In this type of mission, the total cost of ownership including maintenance, and the risk reduction provided by the inherent redundancy and robustness of the blade solution must also be evaluated. Smaller more compact and lower cost enclosures must be considered in these cases.

FURTHER READING

More information can be found at:

- Wikipedia Blade Servers
- IBM
- Dell
- HP
- Fujitsu
- Wikipedia Thin Clients
- IGEL