DL-TR-2005-004

CCLRC

The UK HPC Integration Market

Christine Kitchen, Martyn Guest, Miles Deegan, Igor Kozin *and* Richard Wain

October 2005

Council for the Central Laboratory of the Research Councils

The UK HPC Integration Market

October 2005 Version 1.0

Christine Kitchen, Martyn Guest, Miles Deegan, Igor Kozin and Richard Wain

Computational Science and Engineering Department, CCLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire, WA4 4AD, UK

Abstract

In overviewing the current HPC landscape, this paper considers the multitude of issues faced by an organisation when deciding how best to procure, maintain and maximise the usage of a given HPC resource. We concentrate on the potential role of HPC integrators in any partnership that looks to maximise this entire process, and whether such organisations in the UK have the ability to provide the necessary level of expertise required in all phases of the process, from procurement, through installation onto ongoing support of the resource throughout its life cycle.

We consider how current HPC technology roadmaps might impinge on the role of integrators in responding to the undoubted challenges that lie ahead. Crucial issues when considering potential integrator involvement include both size of the hardware solution i.e., number of nodes, and the ongoing robustness of open source software solutions that might be deployed on these platforms. We provide an in-depth analysis of the current status and capability of a number of the leading HPC Integrators within the UK. Our primary attention is given to the four major companies who supply the academic community and hence are well known to us – Streamline Computing, ClusterVision, OCF and Compusys. Five other integrators are also considered, albeit with less rigour.

It should be emphasised that all the information in Section 3 is provided directly by the companies and does not necessarily reflect the views of CCLRC. With the exception of Cambridge On-line systems, information in Section 4 has been obtained from the Cluster Integrator's official websites. The document concludes with a 10-point summary of important considerations when procuring HPC clusters (mid-to-high-end compute clusters).

If the reader has questions regarding issues raised in the document, or requires further feedback they are encouraged to contact us directly – details are available on our website and in Appendix C.

Contents

1. Introduction

This paper considers the current HPC landscape and the multitude of issues faced by an organisation when deciding how best to procure, maintain and maximise the usage of any associated HPC resource. Historically many organisations have responded to this challenge by forming partnerships with the associated vendor of the hardware in looking to maximise this entire process. Our specific interest here lies in the potential role of HPC integrators in such a partnership, and whether such organisations in the UK have the ability to provide the necessary level of expertise required in all phases of the process, from procurement, through installation onto ongoing support of the resource throughout its life cycle.

The paper is structured as follows. In section 2 we provide a somewhat extended overview of both current and future HPC landscapes, and how the associated roadmaps might impinge on the role of integrators in responding to the undoubted challenges that lie ahead. We will see from the outset that crucial issues involve both size of the hardware solution i.e., number of nodes, and the ongoing robustness of open source software solutions that might be deployed on these platforms. Sections 3 and 4 present the real detail of this paper - the current status and capability of a number of the leading HPC Integrators within the UK. In section 3 we provide an in-depth analysis of the four major integrators who supply the academic community and hence are well known to us - Compusys, ClusterVision, OCF Plc and Streamline Computing – this analysis has been conducted by a variety of mechanisms that are detailed at the appropriate point; section 4 summarises data from the smaller, less-known organisations.

Finally, we provide a 10-point summary of the findings of this paper in section 5.

2. The HPC Landscape

Today's supercomputing choices are the product of commodity market competition, technology evolution, historical hardware and software legacies, and leadership choices within industry. Computer vendors, driven by developments such as DOE's Accelerated Strategic Computing Initiative (ASCI [1]), have aggressively pushed the performance levels of parallel supercomputers higher and higher. Given the economies of scale, it is clear that the immediate future of supercomputing will be dominated by parallel supercomputers built with commodity compute servers, such as those used as web servers, tied together by a high-performance communications fabric. Although currently on this plateau in the evolution of parallel supercomputer architectures, this will not last long. New architectures are already on the drawing boards that will be capable of a quadrillion arithmetic operations per second

(PFlops). Such computers cannot be built using the same technology in today's TFlops computers – they would require too much space and consume too much power

Custom and Commodity Clusters: The evolution of cluster technologies has proceeded on two fronts: in the first development, beginning in the late nineties, IBM, Compaq and SGI - among others - began creating *proprietary clusters* using their shared-memory servers and custom-designed or semi-commodity networks. The IBM approach has been to use their own custom multi-level switch fabrics to interconnect shared-memory nodes based on their Power workstation processors. These nodes have 32 processors in the case of HPCx (Phase 2).

At the same time, true *commodity clusters* were being built and deployed based on uni-processor or dual-processor nodes utilizing Compaq Alpha or Intel X-86 processors. These clusters used mainly semi-commodity Myrinet [2] interconnects from Myricom; but smaller examples were sometimes based on gigabit (or slower) Ethernet switch fabrics. In all cases, the system software was built around the Linux open-source operating system. A number of Terascale commodity clusters have now been installed, including Intel x86-based clusters at Los Alamos (LANL) and Lawrence Livermore (LLNL) National Laboratories, the latter utilizing a Quadrics [3] fat-tree switch fabric rather than the Myrinet fabrics utilized at Sandia and LANL. The most notable systems here are:

- The 42TFlops MareNostrum JS20 Cluster at Barcelona Supercomputer Centre [8] featuring 4800 PowerPC processors
- The 22.9TFlop *Thunder* Cluster at LLNL[9][10] featuring 4096 Intel's Itanium2 processors
- The 11.2 TFlops *Lightening* Cluster at LANL [4] featuring 2,816 of AMD's 64-bit Opteron chip, a real competitor to Intel's costly Itanium product line[1].
- The 11.8 TFlops Cluster at PNNL [5], featuring 2,000 of Intel's Itanium 2 processors.

It is perhaps worth stressing that commodity clusters based on open source software (i.e., Beowulf clusters) are 2-10X more cost effective than clusters based on proprietary solutions.

None of these clusters - custom or commodity - have system balance between computation and communications that is competitive with that found on true MPPs such as the Cray T3E. Nevertheless, for many important classes of applications, they are capable of achieving high parallel efficiency on a thousand processors or more. They also, in general, lack full-system Reliability, Availability and Serviceability (RAS) features. For truly large systems, this has caused difficulties in running large jobs with long execution times. In addition, few of the large clusters deployed have truly scalable system software i.e. can provide service to dozens of simultaneous users and have fast, scalable system boot-up, and executable loading capabilities.

A number of new-generation cluster-based systems are now being developed. They include ASCI Purple, a next-generation shared-memory cluster from IBM with 64 Power-4 processors (each 8 GFlops) per node and a new Federation switch fabric deployed in 2005 at LLNL. Its peak speed will be at least 60 TFlops. HP is producing a new semi-commodity cluster based on Intel's 64-bit Itanium II processor. It utilizes 2-way shared-memory nodes, and it's interconnect is a Quadrics fat-tree. Similarly SGI is offering an Itanium 2-based cluster that is distinguished, however, by a very large non-uniform memory access shared memory node with 64 processors per node. The CSAR service at Manchester has been augmented with such a system, the SGI Altix 3000. Both the HP and SGI systems run Linux as their operating system.

Open Source Solutions: Open Source developments have yielded significant enhancements to the state-of-the-art in operating systems (c.f., Linux OS) and tools (c.f., Apache). Multiple accretions of open software mass have resulted in profitable enterprises that combine these tools into single offerings with support. In addition, there exist many efforts to build

[1] Note that Lightening was installed by the US HPC Integrator, Linux Networx (see section 4.2)

clustering tools (c.f., LLNL Chaos, LANL Clustermatic, Sandia CPLANT, NPACI Rocks, OSCAR and others) that extend the desktop environment to medium and high-performance computing. The potential deployment of open source solutions in satisfying requirements of High-end, Terascale Technical Computing is more an open question. What is clear is that current trends in developing and implementing ultra-scale computers fall well below the requirements capable of addressing many scientific challenges. Based on a number of factors, we believe that ultra-scale platforms will be developed using - among other things - commodity-based hardware components and open source software. It is clear, however, that a major investment is required in developing these software components, technologies, and enhancements to attain the 100X to 1,000X in computational capability and capacity that is vitally needed to address scientific requirements.

In a balanced computational environment, hardware and software improvements have contributed equally to the increase in applications performance. What matters most is the shortest time to solution, with code development efforts definitely on the critical path to success. The delivery of fully integrated heterogeneous, multi-vendor parallel computing environments pose considerable challenges from system administration to application development to resource management. If these challenges are beyond current Tier-1 vendors, assessing the potential of HPC integrators to respond to the same challenge raises a number of issues that need to be addressed.

2.1 *MOORE'S LAW AND HARDWARE DEVELOPMENT*

Chip performance is expected to continue along the Moore's law trend line for the next 5 to 10 years, with the transistor density at least continuing to grow during the foreseeable future. The processor core on IBM's Power4 processor occupies approximately 1/4 of the transistors (and area) of the processor using today's CMOS technology. This implies that in two processor generations from now, i.e., in about three years, the core will only occupy 1/16 of all available transistors. Development economics and high demands on time-to-market will drive the use of common processor cores across many applications using a common base instruction set and a common micro-architecture. It is therefore likely that the extra space enabled by the Moore's law development of transistor density will be used to obtain customised chips, i.e., chips for different purposes which are based on a common core. For example, a low end chip could be obtained by adding fast I/O and L2 cache using the extra space for transistors; a game chip by adding high bandwidth memory interface, high bandwidth graphics interface and special arithmetic processors; a network processor chip by adding external cache controller, packet processor, protocol accelerator, and Ethernet interface; or a dense server chip by placing several cores plus external cache controller and memory control on a single chip.

We would, however, note a major shift in the semiconductor industry strategy that has become readily apparent during the past 18 months, whereby increased processor throughput will be accomplished by providing more computing elements ("cores"), rather than by increasing the operating frequency. Indeed, the 5-year Semiconductor Industry Roadmap shows a factor of two increase in operating frequency from 3.2 GHz (2004) to 6.4 GHz (2009), whereas Intel, AMD, and others are migrating to multiple (2x, 4x, 8x ...) cores within a single "processor." Experience has already shown that balance on multi-core processors can be a significant issue; this change in processor design has important implications for Petascale computing (see 2.2).

Finally we note that an alternative trend aimed at improving performance / Watt and total cost of operation will be to use less powerful processors that use much less power, and constrained memory and I/O. Such a trend will mean that many more processors may be required to reach a desired aggregate performance, further increasing the burden on programmers and users of HPC systems.

2.2 *Operational Challenges with Large Systems*

When delivering and supporting the world's largest systems (e.g., the ASCI systems - Red, Blue Mountain, Blue-Pacific, White, and Q - and HPCx etc), the associated sites encounter many problems with the software environment. These "defects" are often difficult to isolate down to root-cause. In addition, *many of the vendor partners providing these platforms do not have systems even within 1/10 the size of these platforms for debugging and patch verification. If a Tier-1 vendor does not have these capabilities, there is little or no chance that an integrator could claim to be able address this shortcoming.* All these factors contribute to making the process of integrating large systems take inordinately long periods of time (over a year). Several lessons learned in these efforts suggest that high-performance computing goals might best be served by open source development efforts where the end-sites and their associated communities directly participate in the development and/or testing and/or integration of the technology on platforms at scale.

These conclusions are applicable to the broader high-end technical computing (HEC) community for the following reasons:

- HEC customers require the ability to have operating system and tools source code for root cause analysis and debugging. In many cases – at least in the USA – the HEC customer sites have the requisite in-house expertise not only to do root-cause fault isolation, but also to formulate and implement bug fixes (that may possibly include re-architecting portions of the solution). With Open Source community development efforts, these fixes can be fed back into the community for the benefit of all.
- Taken as a group, vendors of proprietary solutions indicate that they cannot make a profit providing solutions that span the entire HEC space. Thus HEC sites are left with two options: live with mediocre solutions that fulfil only part of their requirements, or implement major portions of the solution in-house on top of the vendor provided proprietary foundations. Neither of these two options is advantageous for the HEC or vendor community in the long run. With the Open Source community development model, HEC sites can contribute their solutions back to the community because it is based on a common software base and can be contributed back without fear that the development would benefit only one particular vendor offering (and thus inhibit competition in the future and lock the HEC site into one vendor proprietary foundation). With this wide HEC scope, the diversity of the debugging environment and developed solutions benefits the entire community.
- Change has been a dramatic feature of the HPC landscape for many years. However, recent introductions of many disruptive technologies (e.g., killer micros, IA-32 commodity hardware, open source software) have radically increased the rate of change in the industry. As a result, the time span for vendor support of provided hardware and software platforms has been reduced to shorter than the full lifecycle of implemented solutions at HEC sites. In addition, timescales under which support is withdrawn (from announcement to withdrawal of support) by vendors (measured in quarters) is typically much shorter than the governmental planning period for replacements (measured in years). This mismatch in timescales has led HEC sites to require Open Source based products as a hedge against "change of support" status.

Processors vs. Processing Elements. As noted above, a major shift in the semiconductor industry strategy will lead to increased processor throughput being accomplished through more computing elements ("cores"), thereby increasing the burden of parallelization. Future baseline architectures for 100 TFlops and 1 PFlop might thus appear as below:

- 100 TFlops:
  - 5 GFlops each "processing element"
  - 20,000 processing elements
- 1 PFlop

- 10 GFlops each processing element
- 100,000 processing elements

This increased reliance on parallelization and hence system size will merely act to emphasise the operational challenges associated with Large Systems, challenges that stretch the resources of proprietary vendors to the limit and are realistically beyond the reach of the HPC integrators central to this paper.

2.3 *Capacity Through Beowulf Clusters*

The success of ASCI and international HEC efforts to deliver complex modelling capability to the community has led to an overwhelming demand for "capacity at the capability level." Now that large scale predictive scientific simulation has been adopted as a critical component of scientific endeavour, there is great demand to run parallel applications for parameter studies, convergence studies and the like at small to medium scale of parallelism. These simulations run at a scale of parallelization that was a "capability run" for previous generations of ASCI platforms. However, rather than one or even a few capability runs, literally hundreds of delivered TFlops of computation are required for these production calculations.

There is now widespread debate on how best to meet these crushing capacity demands. Although the strategy - and indeed a detailed analysis of the requirements - are still ongoing, one aspect is certain: the solution will be dependent on commodity clusters that are based on open source software (i.e., Beowulf clusters). Again, this is being driven by the fact that Beowulf clusters are 2-10x more cost effective than clusters based on proprietary solutions. It would seem clear from the outset that *the role of system integrators in providing this scale of resource is far more credible than is the case for their involvement at the very high-end i.e. in the capability regime.*

2.4 *Proprietary Vendors and System Integrators*

Market dynamics will not support substantial unique developments for ultra large systems. A guiding principle for many of the proprietary vendors when building HPC systems is to leverage standard technology wherever possible, and to judiciously add custom technology only when this is required. Further, any new technology must ultimately be applicable to mainstream commercial applications and systems for it to be viable from a business perspective.

The vendor is becoming more and more of a system integrator. In this context it should be mentioned that commodity-based Beowulf systems are rapidly becoming the systems of choice, at least within the academic community. This growth is fostered in the UK by the emergence of a number of companies who do provide crucial added-value services around clusters that address existing shortcomings in the standard, open source based environments, while keeping costs to a reasonable level (see section 3). Procuring systems from such integrators may be more prudent than relying on the more traditional Tier-1 companies such as IBM or HP given the existing cost differential in the associated products, although caution should be applied and reference sites always sought for feedback regarding the previous installations (preferably of similar sized systems). While some way removed from providing credible high-end alternatives for e.g. HECToR, the emergence of enabling technology infrastructures provided by toolkits such as OSCAR [6] and ROCKS [7] has taken much of the hassle away from supporting such systems, at least up to 128 CPUs. The emergence of essential features such as check point restarting, concurrent file systems etc. needs to be closely monitored in judging the "fit-for-purpose" nature of commodity systems at the high-end.

2.5 HPC System Integrator Criteria

Continuing the theme of section 2.4, we need to consider an appropriate set of performance criteria that might enable us to differentiate between the various cluster integrators in the market. Based on an objective assessment of their standing against these criteria, it should be possible to identify potential HEC candidates - the more performance target-compliant a vendor is, the more likely they are to provide a viable alternative to the more traditional Tier-1 companies.

We would consider from the outset that the following issues need to be assessed when judging the viability of a given integrator:

1. *In-house technical expertise:*

   The 'added-value' that the vendor brings to the procurement through software stacks, finely tuned OS etc.

   Number of developers - the vendor's ability to develop, support and maintain the software so as to sustain 'cutting-edge' technology and assist with various code porting and optimization tasks.

2. *Size of the company:*

   Is the company sufficiently staffed to be able to support large scale compute clusters over multiple installation sites?

   Turn-over - is the company a practical long-term prospect with the potential to be 'self-sufficient' with respect to large clusters (> 1000 processors) within the next few years?

3. *Current install-base:*

   An important factor is the integrator's current success in the small- to mid-range computer cluster market and the actual install base (whether solely to the UK or if they have a presence overseas, particularly in Europe or the USA).

   This again provides information on the potential longevity of an integrator as well as feedback from the community regarding their overall performance.

4. *Support infrastructure:*

   The number of technical and software engineers in place to support the cluster throughout it's lifetime and importantly whether this is in-house or out-sourced. If out-sourced whether there are any plans to change this in the near future.

These points formed the basis of discussions with the key cluster integrators at the outset of this analysis exercise – a more detailed account of these points can be found at Appendix 1.

In addition to the above, the possibility of an integrator solution rather than one from the more traditional Tier-1 vendor is very much dependent on the nature of the HPC resource(s) under consideration. Some obvious examples would include:

- The *size of the system* in question – is this targeting less than 1000 CPUs, a domain in which most of the integrators have experience, or does the system in question exceed, say, 10+ TFlop. If the latter, it is worth mentioning that many national procurements in Europe have rejected the use of integrators at a fairly early stage in the proceedings. While US Integrators certainly have experience in the 1000+ CPU domain, this is not in general the case for their UK counterparts.
- The expected *usage pattern and environment* around the resource – is this being driven by *Capability* or *Capacity* requirements? We would expect, based on some of the considerations above, that integrators would be capable of providing the latter requirement far more effectively than the former.
- The level of *RAS features* expected of the HPC solution. Demanding levels of RAS (say 95+ %) around truly large systems are exceptionally difficult to sustain, particularly in a *Capability* regime when running large jobs with long execution

times. Few of the large clusters deployed have truly scalable system software i.e. can provide service to dozens of simultaneous users and have fast, scalable system boot-up, and executable loading capabilities. Assuming such features appear in any contract around the services to be provided, it is extremely unlikely that any integrator would be in the position to accept the risk involved in committing to high levels.

3. State of the UK HPC Integration Marketplace

Consideration is given in the remainder of this document to the current status and capability of a number of the leading HPC Integrators within the UK. We provide an in-depth analysis of the four major integrators who supply the academic community and hence are well known to us - Streamline Computing, ClusterVision, OCF and Compusys – plus a less rigorous analysis of some of the other players. Our analysis has been conducted by a variety of mechanisms; given that we know each well, we asked for a response to a set of discussion points following a phone conversation. These points are reproduced in Appendix 1. In this section we provide an overview of each of the four major UK HPC Integrators together with their response to each of the eight discussion points, while Section 4 summarises data from the smaller, less-known organisations.

As mentioned before, there is no doubt that the vendor is becoming more and more of a system integrator, driven by commodity-based Beowulf systems rapidly becoming the systems of choice, at least within the academic community. Section 5 provides a summary of the headline issues to be considered between procuring systems from the integrators central to this paper rather than from the traditional Tier-1 companies such as IBM or HP given the undoubted cost differential in the associated products. Important considerations such as where is the dividing line between simply rolling out technology infrastructures provided by toolkits such as OSCAR [6] and ROCKS [7], against providing the level of technical expertise to address the typical RAS requirements associated with high-end solutions for e.g. HECToR?

Note that much of the detailed data provided by each integrator makes up the company overviews below. To reiterate these are entirely the companies opinions and do not necessarily reflect CCLRC's views.

Additional company contact information is included in Appendix 2.

3.1 *ClusterVision*

As a Company, ClusterVision BV has been trading since 2002 and ClusterVision Ltd was introduced as a direct subsidiary responsible for the UK & Ireland HPC market in November 2004 as part of their European Expansion. Their growth and success since the start of trading has been impressive.

ClusterVision specialise in the design, implementation and support of large-scale compute, database and storage clusters. They are the only Euro-wide cluster company to focus solely on cluster technology and development. Their clustering technology provides an alternative to traditional proprietary-based supercomputing "by using a method of connecting multiple computers to form a unified and powerful computing system". Every ClusterVision cluster is delivered as a fully functional turnkey system with all the hardware and software integrated and configured for immediate deployment.

ClusterVision's sales and technical teams have designed, built and supported some of the "largest and complex computational and storage clusters in the UK, Benelux and Germany". Several of the key staff hold PhD's in Applied Science disciplines and have years of experience working with both traditional and clustered supercomputers for scientific research. These backgrounds in applied scientific research combined with practical experience of a wide range of supercomputing technologies "provides insight and understanding of

customer's requirements" enabling ClusterVision to provide tailor-made solutions to meet these requirements.

3.1.1 Install Base

The current division between Academic and Industrial installations is approximately 70/30% - ClusterVision does keep some private installations confidential for commercial reasons. Shipment figures in nodes:

2003	750 nodes
2004	1000 nodes
2005	2300 nodes

Specific reference sites include the four sites of the UK National Grid Service, at RAL and the Universities of Oxford, Leeds and Manchester. These four clusters comprise of 348 Intel Xeon 3.06GHz processors and a total of 46TB storage. Each cluster has a full Myrinet network; two clusters run Oracle 9i RAC. (ClusterVision are a World-Wide Oracle Partner).

The reference site at University of Lancaster was the largest Intel 64-bit Cluster in the UK totalling 418 Intel Xeon EMT64 3.06GHz processors and 84TB storage. ClusterVision work closely with Intel and AMD and were the first to deliver both an Opteron and a dual-core Opteron cluster in the UK, at the university of Manchester and RAL respectively. The University of Durham has a 137 Opteron cluster with Myrinet, and a 35TB storage cluster which is used for the International Millennium Simulation project.

In Belgium, CENAERO has an Intel Xeon cluster with Myrinet which featured in the Top 500. In Germany, the Universities of Dortmund and Frankfurt have recently ordered complex clusters with well over 400 Opteron processors, InifiniBand, Myrinet, large storage arrays and the Luster distributed file system.. In Switerzland, CERN will have a 750 processor Xeon cluster installed by ClusterVision this year.

Anyone of the above sites would "be happy to act as a reference" for the UK.

.

3.1.2 Company Details and Size

ClusterVision BV and ClusterVision Ltd have a combined head count of 21 permanent employees and 15 under part-time contract. This also includes ClusterVision Deutschland but does not include staff at their Reseller in France or Logistica subcontractors.

The 21 permanent members of staff are fall into the following categories:

- Management 3
- Sales 4
- After Sales Support/ Technical Support 12

  Of which 4 are hardware engineers and 8 are software/hardware engineers – also know as Cluster Engineers.
- Office Staff 3

The Head Office is in Amsterdam, with close proximity to Schiphol International Airport. It houses a dedicated assembly facility and is subject to "the highest standards of quality control and European guidelines on Health and Safety".

The main office for ClusterVision Ltd is in London.

3.1.3 International Presence

ClusterVision opened an office in Munich, Germany earlier this year, ClusterVision Deutschland and have a reseller, Oxyala, in France. They therefore have access to opportunities across Northern Europe, particularly the Benelux, France, Germany,

Switerzland and the UK, enabling close collaboration in Framework opportunities and a sharing of resources. The UK is currently their largest market due to the strong focus on Scientific Research and Technology advancement, followed increasingly closely by the German Market

The Company's international coverage extends to emerging HPC Markets including South Africa as a response to proactive approaches and also looking at partnerships in the US to support global companies who standardise procurements across Europe and the USA.

3.1.4 Company Expertise

ClusterVision have received a number of Awards for technical and entrepreneurial achievements, including: Vokso Trofee "Technical Breakthrough" (April 2003); Intermediair top 20 most innovative companies in the Netherlands (May 2003); Netherlands British Chamber of Commerce (NBCC) enterprise award for Dutch-British export (Nov 2004).

These awards are due in part to the development of the highly regarded ClusterVisionOS™, ClusterVision's own Linux-based cluster operating system and software environment, designed specifically to ease the use, management and support of large-scale compute, storage and database clusters.

Where benchmarking forms an integral part of the procurement decision, ClusterVision have a greater degree of success "due to the in-house expertise of their Technical Team". The strong scientific background and experience amongst their staff has resulted in a greater knowledge of scientific code in a number of areas: biophysical chemistry, material science, earth sciences, molecular dynamics, chemistry, hydrology, hydrogeology, stochastic modelling and nuclear physics. Tihs proves "useful in benchmarking and providing the optimised solution".

This scientific background is coupled with a solid commercial background demonstrated by the development of an HPC Division for another system provider before starting with ClusterVision[2] and the recruitment of key individuals from Blue Chip companies in component distribution, VAR, SI's and industrial applications such as drug candidate discovery.

3.1.5 Marketplace

ClusterVision from the outset has included grid computing in its product portfolio and one of their most important collaborations is with the University of Amsterdam's Professor Bob Hertzberger and his team working on Virtual Laboratory technology. The ClusterVisionOS, with its ability to include multiple images, has led to the provision of a grid training cluster to the National e-Science Centre to enable trials of grid middleware across the various nodes. The rise in grid enabled clusters is envisaged as a natural development from the continued rise in resource demand.

There is a move towards very low latency interconnects exhibited by the Cray XD1, InfiniPath, 10GE, InfiniBand Double Data Rate. There is also increasing attention towards cooling and energy saving.

Closer to home and welcome is the increasing demand for large, multiple complex benchmarks which allows the cluster specialists to demonstrate their expertise, and total cost of ownership ahead of aggressive Tier-1 price scales. The single selling price or 'buy now factor' cannot fund technology growth, plus technology refresh and 3, 4, or even 5 years support.

[2] Two of the founder members of ClusterVision originated from the HPC team at Compusys.

3.1.6 *Relationship to Tier-1 Organisations*

In all fields ClusterVision certainly "has the experience and expertise to build, deliver and support tera and petascale solutions. Our technical solutions are respected and our reputation reflects this". Therefore ClusterVision works independently of Tier 1 Companies but has strategic relationships in place for necessary collaboration, e.g. SGI in the UK and Official Cray reseller in the Benelux and France. It is possible for ClusterVision to finance projects up to 3M GBP without partners and therefore they generally choose to pursue project opportunities independently on this basis.

In trying to put into context with the USA SI, ClusterVision suggest that "UK integrators are competitive in price with the USA Integrators but to date no USA Integrator has the on-site engineering back up to support installations or shown a consistent or long term commitment to the UK market".

3.1.7 *Relationship to Other System Integrators*

ClusterVision has not seen these companies in the Academic sector.

3.1.8 *Additional Information*

ClusterVision provide tailor made turn key solutions and "can provide these across large and small scale requirements. We design, build, benchmark, install and support our own clusters with no reference to third party". In doing so, ClusterVision have achieved a number of technical firsts: First production ready InfiniBand cluster in Europe, first Opteron 64-bit single- and dual-core cluster in the UK, largest Intel EM64T processor cluster in the UK. They add that "this total solution is due to our own Linux-based cluster operating system and software environment, the ClusterVisionOS™".

3.2 COMPUSYS

Compusys is a privately owned, UK based Computer Solutions Provider and Systems Integrator, and was formed in 1987. Originally an Ollivetti PC Systems reseller, Compusys soon began building and assembling its own brand of PC systems, and by 1989, this had overtaken branded systems sales. Current platforms include Desktop PCs, Workstations, Servers and Storage, and laptops and portables.

The Compusys organisation has continued to grow, but "has retained it's strong focus on high levels of product quality and customer service". By the mid 1990's, Compusys began to realize that broadening its product and service portfolio was essential to the long term survival of the company.

In 1997, Compusys formed its Networking Division, to provide turn-key networking and integration projects for new and existing customers. This provided clients with "a one-stop shop for their entire desktop, networking and back-office requirements". Within two years, Compusys' Networking Division had enjoyed significant success with many turn key campuswide LAN and WAN deployments for further and higher education establishments, the Police, and commercial customers.

In 1999, Compusys formed its' HPC division, after recognising the future potential of this commodity-based solution. Compusys HPC has grown and evolved "to become one of the leading HPC Integrators in the UK", and is the UK's longest established HPC business unit. Compusys also has an e-Business Consulting Division, providing web, content management and e-business solutions to local government, emergency services and academic institutions, which has several significant and prestigious clients.

3.2.1 *Install Base*

Compusys HPC have been building, supplying, installing and supporting HPC Linux Clusters since 1999. In this time, the company has built up "experience and expertise in the design, building, deployment and support of Linux Computational Clusters".

Initial deployments were based on PA RISC and Alpha Processors, as these were the fastest processors available at that time. In their first two years of HPC, Compusys deployed Alpha Clusters to a significant number of Academic Sites, including the University of Liverpool, the University of Leeds, Cranfield University, and many other leading research sites.

During 2000 and 2001, Intel Architecture performance improved, with PIII Xeon systems catching and eventually overtaking the performance of the Alpha based platforms. Compusys again were "at the forefront of HPC Deployments in the UK, with significant deployments taking place at every leading research University in the UK". Over the last four years, Compusys have continued to provide cluster solutions to UK, European and Inter-continental customers.

The company's average run rate for the delivery and deployment of systems over the last four years has been around 50 per annum. The majority of these systems have been below 64 Nodes, with around 15-20% of these being above 128 Nodes. The majority of these installations have been to academic institutions, as the market for HPC has been far more mature than the commercial sector. However, Compusys HPC have continued to expand into commercial markets, with successes in the Automotive, Manufacturing, Bio-Informatics and Financial sectors. The commercial markets currently account for around 15% of Compusys HPC.

3.2.2 *Company Details and Size*

Compusys is arguably unique in that it provides all its own service, support and maintenance for all systems sold by the company. This includes all desktops, servers, laptops and HPC Cluster Solutions. The company employs its own field engineering, support, helpdesk and administration staff, and so does not rely on any third party in the provision of any of its service delivery. Compusys continues to provide "a broad range of solutions, to an even broader range of customers". Compusys now employs over 120 staff.

3.2.3 *International Presence*

Compusys HPC have been providing and supporting cluster solutions outside of the UK for over 5 Years. Their first significant international installation was a 1000 CPU Alpha Cluster for the Moscow Academy of Science, which was installed in 2001. Since then, Compusys have continued to provide systems and solutions to clients across mainland Europe, with installations in many major University and Research sites in Germany and Austria. Compusys HPC have also shipped a number of HPC Clusters to the United States, where their Clusters are incorporated into sophisticated manufacturing systems in Silicon Valley.

Compusys HPC continue to provide solutions across Europe, and are working on projects in Germany, Austria, France, Spain, Italy, Croatia, Switzerland, and other eastern European countries. The company's international coverage also extends to emerging HPC Markets, such as South Africa, Australia and India, and are exploring new opportunities and forging new partnerships.

3.2.4 *Company Expertise*

Building stable, manageable, high performance clusters from commodity components is a skilled job, and for this Compusys HPC have "both in-house expertise and a proven track record of over 250 successful cluster deployments. This experience continually feeds back into the cluster solutions that we sell, thereby enabling us to raise our standards even further".

Compusys stands apart from most other HPC specialists in the UK, as they are the only provider whose solutions are integrated using 100% in-house resources. No third parties are used in any part of the solution, as Compusys HPC is "a true cluster solutions provider". Their solutions are built on their own systems hardware, manufactured in their own assembly facilities, adjacently located to their HPC Labs. All of their HPC Clusters are fully built, configured, tested and signed off in the HPC Labs prior to shipment to site. This ensures the systems are fully operational on day one.

Even after installation, Compusys "continues to provide its own service and support to deal with any HPC issue, hardware or software, itself. Compusys HPC support staff are available to determine the source of the issue and to arrange and take the appropriate action, whether that be scheduling a Compusys field engineer to rectify a hardware fault, or remotely logging into the cluster to fix a software problem".

Compusys have direct relationships with all of the leading vendors of HPC cluster products; both hardware and software. For example, they "collaborate with motherboard and systems hardware manufacturers to help them to design suitable HPC platforms. This collaboration is reciprocated when it comes to support, allowing Compusys to efficiently resolve any operational issues that may occur through direct contact with the designers and engineers of the products sold. The direct relationships also support a preferential pricing and commercial support model – essential when fighting against stiff competition from Tier-1 Vendors".

The company's approach to the deployment of their solutions embodies transparency. "Compusys quotes are clear and explicit, detailing all of the components that are needed for a cluster. This approach allows customers to evaluate the offer effectively, and make like for like comparisons. This open approach is followed through to deployment, where a set of agreed sign-off parameters and formalised procedures are used to ensure that the customer get what they expect. Compusys' implementation specification document (ISD) is a key part of these procedures. It specifies how the cluster will be built, the exact software configuration, and the sign off tests that will be run for system acceptance. This reassures both the client and Compusys that the system will be completed to everyone's satisfaction, within an acceptable timeframe".

All Compusys HPC clusters are hot staged in their HPC Labs facility. After the cluster compute platforms leave their production line, they are built into a cluster in the company's Labs. This includes all Cabinet preparation, with the installation of all required communications cabling, mains wiring and power distribution, and the fitting of external 16 Amp mains connectors. All cables are fully labelled at both ends for easy service and system identification. Once fully built, the systems are passed into the hot staging area. It is here that the complete software environment is built and applied to the cluster, and where internal quality assurance sign off tests are performed. "The testing is designed to identify any components that could fail during the first few months of operation, and to ensure the software environment is configured to the agreed specification. The software environment provides all of the required tools to effectively use, schedule, monitor, manage and control the cluster, and incorporates several in-house developed modules and enhancements, only available with a Compusys Cluster".

The deployment of a cluster at the customer site is simple, as it is already fully operational, meaning that "it takes a matter of hours to deploy". All sites are surveyed prior to deployment, to make sure the systems will fit in the space allocated, and that they can be handled through the building to their final position.

3.2.5 Marketplace

Compusys see the HPC market continuing to grow, as the technology continues to mature, and new innovations drive performance ever higher. However, the issue of Full Economic Costing is now playing an increasing part in decision-making process, as the factors that affect FEC are becoming more visible, and sites are now reaching their capacity to provide the space, mains power and air conditioning required to run a super computer cluster.

Compusys are already seeing the cost of ownership benefits of Dual Core technology putting such solutions high on the shopping list of prospective buyers. Compusys own cost benefit analysis shows potential savings of up to 40% per annum in mains powers and air conditioning running costs with a Dual Core Processor cluster deployment. On a cluster with 500 CPU Cores, this can be a significant cost saving, resulting in a reduced Full Economic Cost.

Compusys are partners with AMD, and were launch partners for the Dual Core processors. This gave Compusys early access to the Dual Core parts, for testing and benchmarking ahead of the official launch. Compusys are already working on Dual Core deployments, with the first installation of this technology due to take place shortly. The company are also heavily involved with other leading HPC technology companies, and are looking forward to the launch of many other new technologies (most of which are currently under NDA), which will have a significant impact on the performance and ownership costs of commodity cluster solutions.

Compusys are also leading the drive to "take the mature products and technologies offered to the academic markets, to the commercial sector, with the drive for greater efficiency and value steering users away from large shared memory systems".

Finally, the company believes that "commodity clusters will be the utility resource that everyone dreams of. Developments are already underway to take standard commercial business applications, and deploy them on commodity clusters, to save cost, increase performance and enable true scale-out computing. In the future, a cluster will just be a resource for processing, with virtual machines running applications for an entire organization. As demands from individual instances increase, resources can be dynamically allocated to those VMs that need it. The days of deploying a new server each time you deploy a new application with be over. With Grid Tools linking each cluster into a pool, most organizations will have the flexibility to cope with all their computing demands, with less hardware, being more flexible, and saving money".

Compusys aim to play a leading part in bringing the above technologies to market, and is working closely with Global HPC vendors, as they prepare to launch their new products over the next 12 months.

3.2.6 Relationship to Tier-1 Organisations

Compusys has a neutral stance on the position of Tier-1 vendors in the HPC marketplace, especially in the UK and Europe. Although Compusys have competed head to head with every Tier-1 vendor at some point either currently, or in the past, they believe that their levels of experience and expertise, plus the fact that all their services are provided by in-house staff, ensures that they can match or exceed the levels of service provided by "the so-called big names".

Compusys is part of a group turning over in excess of $100 Million a year. Compusys are more than able to finance contracts of well in excess of £2 million and have done so several times in the past. Tier-1 support is only an issue, if a particular Tier-1 vendor has decided to offer products at prices below cost.

It is also a little known fact, that most Tier-1 vendors in the HPC space do not actually manufacture their own HPC products. For example, IBM use MSI to make their Dual AMD platforms, as Compusys have used and deployed this same chassis and motherboard, branded MSI, not IBM. SUN offers a 1U Dual AMD and a 3U Quad AMD server platform for its HPC solutions-these systems are made for SUN by Newisys, and are available from Compusys too.

The company believes that most of the time, Tier-1 vendors bring a low cost price for the computer hardware, and the comfort factor of a known name. However, when "you strip back the offerings from the Tier-1 vendors, and find that the hardware is not truly made from the vendor, and that the integration of the cluster is actually being carried out by a smaller

specialist, because the Tier-1 vendor doesn't have the knowledge or expertise to do it themselves, you have to ask where the value comes from".

3.2.7 *Relationship to Other System Integrators*

Compusys has and maintains contacts with the other HPC specialists in the UK, including Streamline and OCF. Compusys and the other vendors meet at various industry gatherings throughout the year, and obviously follow each others progress. With "the high degree of skill and understanding required to install and support Linux HPC Clusters", Compusys see themselves as specialist's contractors to the traditional Systems Integrators, as the steep ramp-up required for a new SI to enter the market would be hugely expensive, and without a track record, difficult to market. Compusys have already done work for Computacenter, and for HPC and Clustering projects.

3.3 *OCF LTD*

OCF is a specialist independent reseller of high performance technical compute (HPTC), high performance visualisation (HPV) solutions and Enterprise Computing Infrastructure for Storage and Server Technologies. The company "continues to evolve its skills in accordance with technological advances and breakthroughs in order to remain at the cutting edge of HPTC and HPV developments and Infrastructure provision". OCF have forged a strong relationship with IBM and the majority of their solutions are now based upon IBM hardware. These solutions range from individual workstations and Servers to large bespoke enterprise systems providing "organisations with maximum compute power and complete data management facilities by adopting a collaborative approach to solving customers IT challenges and working closely with a number of technology partners" in the following areas:

- Development and Deployment of Server Infrastructure
- Complete Data Management Solutions Providing automated ILM (Information Lifecycle Management) allowing organisations to attain the maximum return from all assets.
- Manufacturers of high performance visualisation workstations
- Manufacturers of high performance immersive group visualisation environments
- Manufacturers of the whole range of compute servers, from individual Linux servers, to clustered Linux based solutions, through to high end SMP servers
- Manufacturers of computer interconnect technology
- Independent software vendors

3.3.1 *Install Base*

The division between public/private installations is 70/30% - OCF does keep some private installations confidential for commercial reasons.
A number of customer sites would "be happy to act as a reference" for the UK and the associated market developments – please contact OCF for named contacts.

3.3.2 *Company Details and Size*

The overall size of the company is shown in the Table below:

Function	No. of FTEs
Managerial/Supervisory	1
Sales	5
Service	1
Operational/Administrative	2
Sub-contractors	
Consultants Design Experts	

Design Experts	3
Technical (Implementation, Roll out, Support &Maintain)	6
Others:	
Total:	**18**

OCF's Technical team comprises a team of eight under the leadership of Technical Director, Chris Bailey. This team has "expertise to configure, supply and install large storage solution infrastructure in addition to its well-known expertise in clustered and SMP HPC systems". In addition the company has recently invested in a software engineer to add to their capabilities, giving OCF the ability "to assist it's customers in such matters as Solution Software design, integration of Management software and Benchmarking Code for HPC Servers".

In-house warehousing, and engineering workshops at OCF Sheffield Head Quarters enable OCF's technicians to be able to test all equipment, evaluate and benchmark new technologies as they are released. The company also provide a customer help desk and support facility (available Monday – Friday 08:00 – 18:00 excluding public holidays). OCF build and support procedures are governed as part of their ISO9000 accreditation and IBM reseller partnership agreement. Strict internal procedures are in place to ensure that all IT equipment delivered to OCF plc is fully tested prior to shipment to their customers.

OCF provide maintenance for IBM equipment through IBM Global Services, with OCF acting as the first point of contact for complete projects. Maintenance provided can be customised from next day return to base, through to seven day, 24 hour on-site support. Options include service upgrades for in-warranty machines; extended maintenance for post-warranty machines; experienced technicians; and extensive parts distribution.

OCF have a dedicated Support Hotline to which all calls (technical queries, delivery information, returns, repairs etc.) are logged through the Hotline onto an electronic Call Management system. This in turn is set to escalate any calls that exceed the agreed time span. OCF can also run reports for specific problems, all customers with specific machines that have been notified with a problem etc. Problems relating to hardware are relayed to the manufacturer within 2 hours of receipt, and the Support Hotline monitors progress.

3.3.3 Geographical Outreach

OCF's primary market area is the United Kingdom and Southern Ireland. In the past OCF have performed contracts in Scandinavia and Holland and have acted as a European business partner for a Life Sciences ISV with business currently quoted in Italy, Switzerland and France. OCF have no overseas locations and all contracts are serviced with staff based in their Sheffield office.

3.3.4 Company Expertise

OCF has supplied computational clusters to a wide range of academic and corporate clients, and as a result have implemented a wide range of CPU architectures – Alpha, Intel IA32, Itanium 2, Opteron, and Power 5, utilising nodes from a wide range of Tier-1 vendors and white box manufacturers. OCF have also integrated clusters across a wide variety of interconnects Ethernet, Gigabit Ethernet, Dolphin Wulfkit, Infiniband, and Myrinet, and have expertise in the configuration of tools such as SCALI MPI & SCALI manage.

Recently OCF have undertaken research into the provision of the systems likely to be attached to the computational environment and have investigated the integration of high performance storage systems and the integration of visualisation environments. This work has been undertaken with IBM around GPFS & the DCV visualisation product.

Within the cluster environment OCF's expertise is centred on the infrastructure and the basic operating system framework. OCF do not pretend to be experts in applications and have limited abilities in optimising applications for optimum performance. The role OCF play is

supporting the application experts (usually ISV's) by allowing them to concentrate upon what they do best and taking care of the infrastructure challenges. The market is extremely fragmented, with a number of vendors operating at the server level; a whole host of different interconnect architectures; different flavours of Linux; different versions of each flavour of Linux causing driver compatibility issues etc etc. OCF acknowledge that they do not have all of the skills required in house but do believe that they are unsurpassed in their contact base.

As regards OCF's long term potential and viability in the market, it depends upon how you view the market! It is vital for a company such as OCF, operating as it does in between the manufacturer and the end user, to constantly reinvent itself to ensure that it does not die with today's technology. Such longevity is facilitated by strong relationships with its vendor, customer and technical partner community and OCF work very hard in all of these areas.

3.3.5 Market Place

Considering the impact of Full Economic Costings then it is necessary to take into account the full cost of ownership of a facility through the following factors:

Purchase cost
Installation
Configuration
Power
Cooling
Management personnel
Maintenance
Upgrade
Disposal
Real Estate costs

These make up different proportions of the total and have varying degrees of importance.

It is difficult to justify the use of valuable inner city real estate as a viable location. It is cost advantageous to position systems in lowest cost facilities taking advantage of regional development sites and even offshore hosting and management capability.

The location of systems where the cost of power is minimised is a sensible model. This view point becomes increasingly interesting if we take into account the development of GRID – e-Science initiatives – where a distributed facility could be constructed taking advantage of the regional development locations (Objective 1 regions) and then linked as a grid to fulfil the users requirements.

If a portion of the facility was set to one side to assist development within the host region and academic research support was made available to companies within the region alongside the computational facility then "we could be looking at creating a win/win scenario". Institutions such as Imperial College, UCL, Oxford Cambridge, Manchester etc could gain advantage from the objective 1 supported locations whilst the location can gain from available academic consultancy to assist business.

As regards OCF's own internal investments, they are continuing to increase the peoples skills in OCF's core competency areas i.e. infrastructure and operating systems. Taking the solution up the food chain from there requires the creation of strong mutually beneficial relationships – hence their continued focus upon a partner based approach and evolving relationships with such specialist organisations as Pathscale, Mellanox, Voltaire, IBM, Allinea, Arup, Fluent, Force 10, Level 5, etc etc.

3.3.6 *Relationship to Tier-1 Organisations*

In responding to the question – "Does your company have the strength and depth to build high-performance, scalable systems which can support Tera & Petascale solutions in the not-too-distant future – Are you able to provide this independently" – OCF believe that the only honest answer is "yes we can build them but not independently". They do not feel able to assess the relative abilities of UK and US integrators.

OCF's view is that "the only sensible way forward for commodity based Tera & Petascale computing is to leverage the financial strength of a Tier-1 vendor". OCF have backed that view by developing a very strong relationship with IBM. Risk, reliability and long-term support are vitally important in most areas of business, and they are critical in large scale computing environments. From an integration perspective, the major problem of fault identification and subsequent resolution means that pre-configured supported solutions are very much in vogue with the Tier-1's, with IBM's 1350 and 1600 clusters and HP's XC. The downside of such solutions is that they are significantly more expensive and customers face difficult decisions, especially those funded by the public purse!

In the near term OCF envisage performing a similar role for IBM that Streamline used to/still does perform for SUN, although transactions will be structured in a different way. Potential scenarios involve OCF taking on the credit exposure, getting IBM to sell the kit direct to the customer and OCF selling the services direct to the customer or IBM selling a total solution to the customer and OCF contracting with IBM to provide services.

In the medium term OCF are not convinced that today's delivery models will be appropriate. How users obtain their compute cycles may well change radically if there are advances in middleware, bandwidth etc. There has to come a time when Vendors question whether delivering kit is an economic model, or whether the delivery of cycles is more appropriate. IBM currently operate a very crude version of this (On Demand Computing) and the vision of them simply taking servers off the production line and into huge server farms serviced on demand may not be fanciful.

3.3.7 *Relationship to Other System Integrators*

OCF believe that the role of these organisations depends to a large degree on the business model being adopted by the provider of the compute cycles. Such organisations are good at billing and on going facilities management. Hence CSC's tie up with SGI at CESA. They tend to be exceptionally expensive having had a reasonably easy life looking after corporate networks and charging absolute fortunes.

3.4 *STREAMLINE COMPUTING*

Streamline Computing creates HPC solutions from commodity components, whether sourced internally through Streamline's own supply chain or in partnership with Tier 1 or Tier 2 vendors, both delivering and supporting these solutions. It has provided over 200 HPC Linux clusters to blue chip companies both nationally and internationally, as well as to over 35 top-ranking academic institutions since its formation in December 2000.

Since spinning out of the HPC facilities of Oxford and Warwick universities, eight of Streamline Computing's Linux supercomputing clusters have been included in the Top500 list at the time of their delivery. Six of these were delivered in collaboration with Sun Microsystems, with whom Streamline continues to work closely, most recently delivering a 1024 CPU Sun Opteron cluster to the University of Nottingham (Spring of 2005). On registration, this cluster was the most efficient gigabit Linux cluster in the Top500 list. Streamline computing is a trading division of Concurrent Thinking Ltd. Along side its software tools company Allinea.

Streamline has also worked closely with its clients to provide HP, IBM and Dell solutions as well.

In the majority of cases, Streamline provides its own software stack to extract maximum benefit from minimum investment in HPC hardware, irrespective of whether the requirement is for hardware from one of the major Tier-1 vendors, or white boxes from their own suppliers. The company has a proven track record in tuning operating systems plus specialized MPI and scheduling software layers, and has developed expertise in specific applications in the areas of engineering, simulation, life science and energy. Streamline delivers the solution stack through the configuration of the front-end nodes and a customised cluster management and control (CMA) solution which is currently going through a second stage development. This will extend the control, management and re-purposing of Linux clusters.

Streamline-Computing has experience in networking solutions (commodity Ethernet or low-latency alternative), storage and file-systems, visualization and integration of these within the operational environment and policies of the end-user.

Streamline partners: Streamline Computing partners with a number of organisations in the HPC sector to deliver solutions. These range across the complete stack from microprocessor through networking to application development and integration; the company works closely with AMD and Intel and their respective platform providers.

The majority of clusters delivered by Streamline offer Gigabit Ethernet as the networking solution which meet the demands of the majority of commercial ISV codes as well as meeting the primary requirements of academia in respect of maximising CPU count. In terms of high-performance networking however, Streamline is one of Myricom's premier resellers and integrators internationally, and is developing new partnerships with alternative commodity interconnect suppliers and Ethernet switch vendors where such networks can offer distinct advantages.

Streamline clusters are today shipped with the Streamline CMA where IPMI or proprietary lights-out management interfaces are available although in some legacy cases an 'out of network' advanced console server, providing access to hardware without having to rely on a live and functional network connection is necessary. The company is an authorised reseller for most of the software companies providing compiler technology to the HPC market, and if purchased as a component within a Streamline solution, these software products come preinstalled, licensed, ready for use and integrated with the Allinea tools for debugging. Their current list of compiler suppliers includes The Portland Group, the Absoft Corporation, Intel and PathScale

Streamline's software stack: Streamline Computing installs and configures a custom combination of open-source and licensed software on all systems as part of its package. Most systems ship with SuSE Professional or Enterprise Linux or RedHat Enterprise Server Linux. Other Linux distributions such as Scientific Linux, Rocks or older RedHat distributions, can be loaded on request with options available to support re-imaging of clusters for specific environments as well as the ability to support RAM disk in both disk and diskless clients. All distributions are fully installed and patched up to the prevailing levels, and each system can be configured for automatic operating system updates using Yum or Yast. For parallel applications using MPI, Streamline has installed most of the numerous choices available. Streamline also has unique access to an open source version of MPI called SCore, developed and widely used in Japan. Streamline is part of the SCore consortium which includes Fujitsu and Toshiba. The SCore parallel computing environment is installed as standard on systems shipped by Streamline. When installed and configured, SCore provides a parallel computing environment, including a custom MPI layer, that generally provides significant performance increases for parallel applications compiled to use SCore's drivers and subsystems. SCore provides multiple network support allowing an application to use different networks such as Myrinet and Gigabit Ethernet during execution without any user designation. SCore also

provides deadlock detection, fault tolerance with pre-emptive check-pointing, parallel process migration and flexible job distribution including gang and batch scheduling.

Streamline also has a wide range of expertise in Distributed Resource Management (DRM) and scheduling software and in setting up load distribution systems. Most Streamline clusters ship with SGE, but also deploy clusters using the PBS Pro and Open PBS as well as LSF.

3.4.1 Install Base

An idea of the relative number of "small" (32-64) systems compared to larger (128+) Machines

Whilst a clear majority of systems installed over the last few years fall in the small to medium category, more recently entries into the larger compute market are becoming more commonplace (64 to 128 nodes), with some deployments making the TOP 500 rakings. The largest machines that Streamline has installed include 1000+ processor systems for Nottingham University and a middle-eastern Oil company (both with Sun Microsystems).

Notable recent supercomputing clusters installed by Streamline Computing

In partnership with Sun Microsystems:

- 1024 CPU Opteron gigabit cluster with SCORE at the University of Nottingham – the most efficient gigabit cluster ever installed, reaching #109 on the current Top500 list (April 2005)
- 320 CPU Opteron mixed Myrinet/gigabit system at the University of Sheffield (January 2005)
- 256 CPU Opteron Myrinet cluster with SCORE at the University of Leicester (December 2004)
- 128 CPU Opteron Myrinet cluster at Schlumberger GeoQuest (December 2004)
- 2048 CPU Intel Xeon Myrinet cluster for major oil company in the Middle East (November 2003)
- 128 CPU UltraSparc III Myrinet cluster at University of Durham - the largest UltraSparc III cluster in Europe at the time (2002).

Streamline Computing systems:

- 256 CPU Opteron Myrinet cluster, racked 316 in the Top500 (June 2004) for Rutherford Appleton Laboratories
- 264 CPU Intel Xeon Myrinet cluster known as the White Rose Computational GRID, a computational infrastructure across three British universities
- Numerous Intel Xeon and Xeon EM64T clusters, representing more than 500 CPUs, installed for a major UK manufacturer.

What is the approximate split between HEI's and industrial installations.

The split in cluster terms is approx 60% HEI/research to 40% commercial. Notable commercial customers include the following (in alphabetical order):

- ADCO
- Astex Technology
- ChevronTexaco plc
- Corus Group pls
- Fujitsu Laboratories of Europe
- Kerr-McGee
- McLaren F1
- Offshore Hydrocarbon Mapping Plc
- Organon Laboratories Ltd
- Renault F1 team
- Rolls Royce plc
- Saudi Aramco
- Schlumberger GeoQuest
- Serco Assurance

3.4.2 *Company Details and Size*

Some of Streamline's company information is given in the table below; of a total of 28 staff within Concurrent Thinking Ltd, 18 reside in the Streamline Division. Details of the skills base within the organisation are captured above under the sub-heading "Streamline's Software Stack".

	Streamline Division	**Concurrent Thinking Limited**
Number of staff employed	18	28
Management/Supervision	3	5
Clerical/support	4	4
Operational/Support	7	15
Sales & Marketing	3	4

Streamline Computing is a young company currently working towards accreditation for BS EN ISO 9002.

3.4.3 *European Presence*

Streamline have installed clusters in France, Germany, (as well as the Middle East, USA and Canada), although its primary focus has been within the UK. Between Allinea and Streamline (the two trading divisions of Concurrent Thinking Ltd) they have direct sales in France and Austria (covering Germany, Eastern Europe and Russia). The historical approach has been to provide software and support skills to local cluster builders. In servicing Streamline's largest commercial customer, they have worked closely with Emplics in Europe and Aspen Systems in the USA as hardware build partners. Streamline also have provided the same service for Sun in Middle East and elsewhere.

In the Concurrent Thinking business and growth plans, coverage will be increased with offices in the US and Europe, although no further information is available at present.

3.4.4 *Company Expertise*

Streamline engineers have skills in a number of areas individually and "combine these skills to solve more difficult system level issues". The company's staff have a good grounding in Linux but bring skills in DRM and GRID, parallel computing, cluster management, system monitoring and high-end storage. Company staff members have PhDs in Parallel Computing, Computational Physics, Computational Biology, Numerical Linear Algebra, and Computational Fluid Dynamics.

Allinea's staff have specific skills relating to large scale parallelism. Allinea has already brought to market a parallel debugger (DDT) and a parallel optimiser (OPT), and is involved in other R&D projects in parallel graphics and fault tolerant parallel computing. On occasion, these skills are applied to benchmarks (although Streamline has its own ISV manager who works closely with ISVs on benchmarking). Allinea has very close ties with a number of companies developing compilers and libraries.

3.4.5 *Marketplace*

Streamline / Concurrent Thinking Strategy

Clearly it is only possible to discuss investment plans under formal NDA. However a few pointers as to how Streamline are thinking the market will develop and some of their comments on this may help.

As cluster sizes grow but commodity hardware becomes cheaper and cheaper, the "value" in the market will be the ability to make sure that clusters work efficiently for a broad range of applications and function in a more scalable manner and operate seamlessly across sub-systems. System management and monitoring capabilities will become more important, and will help the ability of Streamline and the end-user to support such a system through its lifetime. The ability to provide a high level of support, not only on a system level, but also on an application level, will help differentiate the 'box-shifters' from the serious HPC oriented companies.

Previous investment has been made in building this knowledge base within the company. In particular, Streamline have invested in staff and resources to develop Score and CMA (the Cluster Monitoring and Administration Tool) as clear differentiating technologies as well as clearly separating tools provision to Allinea to address the Linux cluster market as well as proprietary HPTC solutions together with embedded processing capabilities. The challenge now is to provide solutions and services which leverage this knowledge 1) by providing a commodity solution across a number of geographies in partnership, and 2) by positioning value-added services at the high end of the marketplace.

Trends in the Marketplace

A significant proportion of cluster sales have been, and continue to be, based on Gigabit technology, or a mix of Gigabit and Myrinet. For this reason, Streamline has invested significantly in R&D relating to SCore and MPICH-PM, which together outperform other MPIs over Gigabit and allow customers to run their MPI applications over heterogeneous networks (n.b. the efficiency of the system at the University of Nottingham). Streamline observe very few customers who run a large proportion of capability jobs on their clusters; the majority run mildly parallel applications on up to 32 processors. Benchmarks show that SCore outperforms MPICH and LAM by between 20% and 1800% on standard test cases using Computational Chemistry and Computational Fluid Dynamics codes using 32 processors and Gigabit. Streamline will soon announce new technology that it has been developing that is expected to widen this performance gap even further while providing highly advanced fault tolerant capabilities. This will help add even more 'value' to the Streamline offering.

At the higher end of the marketplace, most of Streamline's customers continue to adopt Myrinet 'D' or 'F' card technology (2Gbit/s full duplex), rather than the more expensive 'E' (4Gbit/s full duplex) and the new '10G' (10 Gbit/s full duplex) technology. This trend may be explained once again by a modest need to run capability jobs, and a wish to optimize price/performance and maximise cpu count. With the new wave of multi-core solutions, with 4 and 8-socket servers, Streamline is witnessing an increased requirement for clusters of modest SMPs connected by Myrinet 10G, Infiniband or Quadrics interconnects.

As all interconnect vendors reach the physical limits of their technologies (e.g. latencies of the order of 1 us), whatever technology is chosen to be embedded on the motherboard is likely to become the defacto standard. This is probably going to be either specialist 10GigE technology or Infiniband technology.

Challenges

It should be noted that Streamline has had exposure to, and has overcome, some very interesting technical and support challenges relating to large cluster systems (buffer overflows on switches and NICs; processor timing problems; interaction between job schedulers and Linux modules etc.).

A potential major challenge to SI is whether all public procurement decisions relating to clusters become a question of price, then companies like Streamline will be unable to sustain the skill-base needed to resolve such complex problems. Similarly, records demonstrate that Streamline currently provide a significant level of customer support (all of their support queries are logged in a company database) yet that the delivery of this level of service is

clearly not profitable for academic customers on a tight budget. As a growing company, Streamline see this as an investment for the future. However if University procurement procedures do not provide the mechanisms for companies like Streamline to demonstrate their value, then the provision of this level of service will become unviable.

3.4.6 Relationship to Tier-1 Organisations

Very large systems require the financial strength of a Tier-1 vendor to execute larger contracts but need the specialist skills of companies like Streamline to provide the knowledge and skills to build and support the systems. Relationships at all levels with Tier-1 vendors are thus critical as a success factor for future growth. Note that it is possible to include companies like Intel, AMD and Supermicro in this definition of Tier-1 companies, rather than just the traditional Tier-1s.

Streamline have built many systems with Sun Microsystems but have also assembled systems with Dell, HP and IBM hardware. They have also provided support for a number of other vendors, both in UK (for systems shipped from the USA) and in Europe, M. East and USA, where Streamline partners build and Streamline support. Experience so far demonstrates that Streamline's technical staff have more than equal skills to any other integrator globally, and can compete effectively if these skills can be channelled profitably.

3.4.7 Relationship to Other System Integrators

Streamline has developed relationships with large scale System Integrators where commercial and industrial end-users have out-sourced supply to these organisations. This arrangement meets the market demand for delivery of the commodity servers in the most efficient manner but enables Streamline to transfer its value in providing the solution. Streamline is continuing to develop further relationships in this area as part of a definitive strategy aimed primarily at the industrial and commercial HPTC sectors and where appropriate in the academic sector too.

3.4.8 Additional Information

"With a strong UK based team with deep technical skills" Streamline has much to offer UK and European Customers and as a company "can keep a UK technology flag flying in a global market". Whilst establishing their own Tier-1 relationships, Streamline feels that support from organisations such as CCLRC and other UK government organisations can help the company to succeed in this market and ultimately benefit UK PLC with exports and technology from a strong UK base. "Ultimately, hardware is the most commodity and in that sense the least important decision relating to the purchase of HPC clusters". As such, their strategy is focussed on the delivery of intellectual property and know-how in a manner that is independent of the actual hardware.

4. UK HPC Integrators II

Having considered the current status and capability of the four leading UK HPC Integrators – Streamline Computing, ClusterVision, OCF and Compusys – in section 3, we now provide a far briefer overview of some of the other, less recognised players. Note that much of the information here has been obtained from the organisations web pages and does not map naturally onto the discussion points raised with each of the integrators of section 3. Companies considered below include Cambridge Online Systems, Linux Networx, Western Scientific, SCC and Workstations UK. In only one case, Cambridge Online Systems, did we have the opportunity to raise and discuss the points noted at Appendix 1.

4.1 CAMBRIDGE ONLINE SYSTEMS LTD

Response from Cambridge Online Systems Ltd (www.cambridgeonline.net)

4.1.1 *Install Base*

The customer base is within the UK and made up of Research and Higher Education establishments (70%) and commercial organisations (30%).

The main architectures installed are HP (Digital) Alpha with Tru64 Unix, HP Proliant (with Intel Xeon and AMD Opteron processors) and HP Integrity (Itanium), both with Linux. The majority of systems are either clusters or compute farms and also include bladeservers. The split between 'small' (32-64) systems and larger (128) systems is roughly 75:25%

Over the past three years, noticeable trends have been:-

- shift from proprietary Unix to Linux (accelerating)
- increasing use of industry-standard, 'commodity' computing systems
- heterogeneous, mixed-vendor Linux environments.

4.1.2 *Company Details and Size*

Cambridge Online was established in 1978. Their Applied Technology Group specialises in the provision, development and support of IT infrastructure, covering computer systems, storage, networking and telecommunications. The company partners with industry leading technology vendors and hold ISO9001:2000 quality accreditation, and with a particular focus upon HPC systems, delivers "a highly technical and consultative approach to meeting solution requirements".

Total staffing of 60 employees is split by sales - 5, technical support/engineers - 15, software development - 30 and management and administration – 10. More than 250 customers are served.

Relevant customer names include the Wellcome Trust Sanger Institute, European Bioinformatics Institute, University of Cambridge, University of East Anglia, Cranfield University and the Medical Research Council. Reference details can be provided following customer requests

4.1.3 *Company Outreach and Presence*

Cambridge Online's business is largely UK-based with only a small, overseas presence (< 10 customers).

4.1.4 *Company Expertise*

Cambridge Online has "particular expertise and experience" in HPC. Their portfolio of products from industry leading vendors is complemented by value-added services which include technical consulting, system and network design, system build and configuration, systems integration, network infrastructure design and installation, technical support and system maintenance.

Linux and open-source software (especially Lustre file system technology) are key drivers.

The company's HPC Lab provides facilities for proof-of-concept demonstrations, benchmarking, training and support.

4.1.5 *Relationship to Tier-1 Organisations*

Cambridge Online claims to have the "strength and depth" to provide HPC, scalable systems supporting Tera / Petascale solutions and to have the necessary level of experience.

The company works in close partnership with leading vendors for full support and have accredited relationships with Hewlett-Packard, Intel Corporation, Silicon Graphics Inc, Platform Computing Corporation and RedHat amongst others.

They do feel that UK integrator partners are competitive with those in the USA, *provided always that strong Tier-1 relationships exist.*

4.2 *Linux Networx*

Linux Networx website (www.linuxnetworx.com)

4.2.1 *Install Base*

Linux Networxs has been responsible for building some of the most powerful supercomputers in the world, and as such have an install base of large systems that is far greater than any of their UK counterparts. The 10 fastest Linux Networx supercomputers are:

- JVN 13.9 TFlops, 1024 Dual Xeon Nodes (ARL)
- Lightning 11.26 TFlops, 1408 Dual AMD Opteron 2.0GHz Nodes (ACSI – Los Alamos, Lawrence Livermore & Sandia)
- MCR 11.2 TFlops, 1152 Dual Xeon 2.4GHz Nodes (production Oct 2003)
- Pink 10 TFlops, 1025 Dual Xeon 2.4GHz Nodes (Los Alamos National lab – LANL)
- AIST 3.1 TFlops, 278 Dual Xeon 3.06Ghz Nodes (Japan)
- Jazz 1.73 TFlops, 408 Intel Xeon Processors (Argonne National Lab – LCRC)
- Orange 1.63 TFlops, 256 AMD Opteron 1.6GHz (Los Alamos National Lab - LANL)
- Catalyst 1.5 TFlops, 128 Dual Xeon 3.06GHz (Sandia National Lab)
- Powell 1.5 TFlops, 128 Dual Xeon 3.06Ghz Nodes (Department of Defense)
- Brahms 1.4 TFlops, 128 Dual Xeon 3.06GHz Nodes (Boeing)

Note the information provided on the above URL is not representative of all Linux Networx installations, merely those from customer sites who have given permission for Linux Networx to publish details on their website.

Linux Networx client base covers a range of sectors including Manufacturing, Life sciences, Government & Research, Entertainment and Oil & Gas. From the website it is not possible to provide a breakdown of academic to commercial sales ratios.

4.2.2 *Company Details and Size*

Linux Networx is based in the USA but has offices/subsidiaries worldwide. The main objective of the company is to help customers improve product development and scientific research by delivering high productivity computing systems. Linux Networx strives to raise industry standards through new technologies, higher customer satisfaction by delivering proven computing systems to help customers overcome their most difficult computing challenges.

Linux networx's workforce consists of hardware and software engineers, installation & integration staff & sales representatives, although the distribution of employees in these fields is not provided.

4.2.3 *Company Outreach and Presence*

Linux Networx's outreach covers most of the world, with sales representatives in the United States, Europe, Middle East, Africa and Asia. The company also has a small number of 'approved resellers' to cover the UK, France and Egypt. Unlike many of the other integrators considered in this section, Linux Networx can point to a successful UK installation, having

recently installed an Evolocity II Linux cluster computing system at the European Centre for Medium-Range Weather Forecasts (ECMWF). The system will be used to evaluate the suitability of cluster technology for broader deployment within ECMWF's high performance production environment, primarily as a test bed for various aspects of ECMWF's operational workload. Linux Networx successfully met ECMWF's requirements for acceptance testing on June 18 2005. The cluster is fairly modest in size, consisting of 64 AMD Opteron 2.2 GHz processors, 128 GB of memory and uses InfiniBand high-speed interconnects from Mellanox. The cluster also includes the companies own cluster management tools, Clusterworx and Icebox, to provide total cluster management from one interface.

4.2.4 *Company Expertise*

Linux Networx clearly has a proven track record in the HPC market with a range of systems in the TOP100. The company has incorporated an Active cooling technology to the cluster designs, developed their own management software together with storage to provide a unified high performance computing solution.

The company has partnered with many top-tier application vendors to optimize and pre-integrate software onto the clusters. They also work directly with customers to understand their specific software application and integrate and optimize the performance on their Evolocity clusters.

4.2.5 *Marketplace*

Linux Networx covers a diverse market place with a range of compute solutions, from HPC to grid. Many installations appear to be focussed in the USA, however sales are global with one of the most recent being the ECMWF in the UK. The academic sector is not specifically mentioned on the website.

4.2.6 *Relationship to Tier-1 Organisations*

Linux networx aims to partner with key software and hardware vendors to jointly design and sell turnkey cluster solutions that are optimised for specific applications and markets. Partnerships have been made with both Intel and AMD in order to deliver some of the fastest supercomputers in the world.

Other important technology solution partnerships have been made with high performance interconnect firms (Mellanox, Myricom, Quadrics), compiler developers (Pathscale, Portland) as well as novel architecture manufacturers (Clearspeed).

4.2.7 *Additional Information*

In addition to providing hardware, Linux Networx have there own Total cluster management packages:

- Clusterworx A comprehensive Linux cluster management software package;
- Icebox. is a hardware management appliance that combines a serial terminal server and a remote controlled power distribution for simplified cluster management. Scalable to support thousands of nodes, each Icebox has a network connection that allows multiple boxes to create a highly scalable IP-based communication network.

Linux Networx also have partnerships with several leading independent software vendors (ISVs) to offer fully integrated systems. Optimising and integrating mission critical applications with the ISV enables delivery of high productivity compute clusters

4.3 *SCC*

Obtained from SCC website (www.scc.com)

4.3.1 *Install Base*

SCC international client base spans both public and private sectors with specific expertise in Banking, Financial and Professional Services, Manufacturing, Pharmaceuticals, Retail, Leisure, Telecommunications, Transport and Utilities, together with Defence and Intelligence, Education, Health and Local and Civil Government.

4.3.2 *Company Details and Size*

SSC has a 28 year history of successful growth, whose business has developed from an initial investment in the UK of €3,000 into a €3 billion turnover business with leading positions in seven key European markets and business partners in over 65 countries.

4.3.3 *Company Outreach and Presence*

SCC is a strong company within the European market place – no information is provided for activities outside the European region.

SCC has offices in the following European countries:

- United Kingdom
- Belgium
- France
- Germany
- Italy
- Netherlands
- Spain

4.3.4 *Company Expertise*

In order to develop and deliver core solution sets and leverage SCC has developed relationships with key technology vendors. The 'Enterprise Solutions' section of the company is formally organised into 6 key technology pillars.

Each pillar operates under its own management with specialist sales, consultancy and vendor relationship management resources. Each management team "expends considerable time and resource with vendors and service providers from both a leading and emerging marketplace position. This ensures a continual flow of new ideas and technology components for solutions architects to design, test and deliver best of breed solutions with rapid time to market. Distilling their experience with bringing solutions composed of leading technology components to a wide variety of customers allows SCC to develop thought leadership and trusted advisor status".

Enterprise Solutions does not operate independently of the traditional customer account manager or sales contact. Customer engagement is on a planned, integrated basis based on business needs which have been intelligently identified. Enterprise Solutions is a rich source of solutions expertise to be introduced and managed SCC account management teams assigned to customers.

At all stages of engagement SCC Enterprise Solutions works to a process that ensures that value goals are clearly defined at the outset, audited during the project lifecycle and measured at its conclusion.

- **Enterprise Computing – process and transact**
  Key vendor relationships are: HP, IBM, Sun, SGI
- **Enterprise Storage – store and retrieve**
  Key vendor relationships are: HP, IBM, Sun, Veritas, Network Appliance and EMC
- **Enterprise Communications – join and protect**

Key vendor relationships are: Cisco, Nortel, CheckPoint, Nokia, APC and QinetiQ

- **Enterprise Software – collaborate and analyse**
  Key vendor relationships are: IBM, Oracle, Microsoft, Citrix
- **Enterprise Print Solutions – copy and archive**
  Key vendor relationships are: HP and Xerox
- **Enterprise Solutions Architects – design and transform**
  Key vendor relationships are all the above, ISVs and emerging technology partners (e.g. VMWare)

4.3.5 Marketplace

As outlined above, SCC covers a variety of commercial clients as well as academic institutions. No data is provided on the website regarding the proportion of commercial to academic install bases.

4.3.6 Relationship to Tier-1 Organisations

SCC has a number of key partnerships with Tier-1 organisations including IBM, HP, Sun, SGI. See section 4.4.4 for further details.

4.4 Western Scientific

Obtained from Western Scientific website (www.wsm.com)

Unfortunately very little in the way of information on client base / technical assistance is provided at the above website.

4.4.1 Company Details and Size

Western Scientific is a global provider of high-performance computing and storage solutions. Founded 26 years ago, Western Scientific supply an extensive line of computing solutions including the latest Beowulf / HPC clusters, RAID and tape storage, high performance workstations & servers and networking solutions for the multi-users Linux, Unix & Windows marketplace.

4.4.2 Company Outreach and Presence

Western Scientific main Head quarters are located in the United States. The company also targets the European audience and has an office located in the United Kingdom. They did put in an appearance at the 2004 Machine Evaluation Workshop in December 2004, and while promising much, have been invisible since that event.

4.4.3 Company Expertise

Western Scientific has partnerships with key Tier-1 organisations including AMD, IBM and Intel. It also has key collaborations with major high performance interconnect providers (Mellanox & Cyclades) and main stream Linux OS (RedHat and SuSE)

4.5 Workstations UK

Obtained from Workstations UK (www.workstationsuk.co.uk)

Workstations UK Ltd is the European agent for Terrascale Technologies who provide the parallel storage platform TerraGrid, "the fastest, most scalable shared storage solution available".

4.5.1 Company Details and Size

Workstations UK is based in Amersham, Buckinghamshire in the United Kingdom. Inventors of the blade server, Workstations UK has experience of MPI, PVM, SCI, Infiniband interconnects , NAS & SAN storage.

4.5.2 Company Outreach and Presence

Workstations UK operates within EMEA space. The customer base includes:

- Conoco/Phillips
- Sandia National Laboratory
- EBI
- Raytheon
- NNSA
- Defense Intellegence Agency

Workstations UK is currently involved in HPC projects in Norway, Italy, Switerzland (CERN) and Brazil.

4.5.3 Company Expertise

Workstations UK is focussed on the TerraGrid parallel storage platform from TerraScale technologies, and acts as the European agent for TerraGrid.

TerraGrid is used in Geophysical processing, Biotechnology, Digital media, Mechanical and Electrical Engineers and High Performance Computing, where TerraGrid "makes a linux cluster behave like an SMP".

5. Summary and Conclusions

In overviewing the current HPC landscape, this paper has considered the multitude of issues faced by an organisation when deciding how best to procure, maintain and maximise the usage of any associated HPC resource. We have concentrated on the potential role of HPC integrators in any partnership that looks to maximise this entire process, and whether such organisations in the UK have the ability to provide the necessary level of expertise required in all phases of the process, from procurement, through installation onto ongoing support of the resource throughout its life cycle. Our primary conclusions are as follows;

1. Crucial issues when considering potential integrator involvement include both size of the proposed hardware solution i.e., number of nodes, and the ongoing robustness of open source software solutions that might be deployed on these platforms. Specifically;
    a. The *size of the system* in question – is this targeting less than 1000 CPUs, a domain in which most of the integrators have experience, or does the system in question exceed, say, 10+ TFlop. If the latter, it is worth mentioning that national procurements have rejected the use of integrators at an early stage. While US Integrators certainly have extensive experience in the 1000+ CPU domain, this is not in general the case for their UK counterparts. However in fairness to the UK integrator market, tenders of the 1000+ CPU systems are rare, which has repercussions restricting companies experience in this market place. It would certainly be of interest within Europe to give some of the integrators mentioned in

this document an opportunity to demonstrate there capabilities in the 10+TFlop arena.

b. The increased reliance on parallelization and hence system size to accomplish the highest levels of performance will merely act to emphasise the operational challenges associated with extremely large systems, challenges that stretch the resources of proprietary vendors to the limit and are realistically beyond the reach of most of the HPC integrators central to this paper

c. The expected *usage pattern and environment* around the resource – is this being driven by *Capability* or *Capacity* requirements? We would again suggest that integrators are capable of providing the latter requirement far more effectively than the former.

d. The level of *RAS features* expected of the HPC solution. Demanding levels of RAS (say 95+%) around truly large systems are exceptionally difficult to sustain, particularly in a *Capability* regime when running large jobs with long execution times. Assuming such features appear in any contract around the services to be provided, it is extremely unlikely that any integrator would be in the position to accept the risk involved in committing to high levels

2. Our considered view is that existing UK HPC Integrators certainly do have a valuable role to play in the on-going provision of capacity-based resources i.e. less than 1000 CPUs, but the majority are far less able to provide added value to high-end capability machines where the focus lies on stringent RAS requirements.

3. The HPC integration marketplace is growing rapidly. The install base of commodity-based clusters is accelerating at pace in the UK, funded through initiatives such as SRIF, and a large portion of that business is going to the integrators identified in this paper, and not to Tier-1 vendors such as IBM and HP. The reasons for this are easy to understand;

   a. The focus of Tier-1 activity remains on the larger, proprietary-based machines where the margins are greatest. Companies such as IBM remain focused on their proprietary CPU offerings – the power series – with much of their pre- and post-sales support targeting such solutions.

   b. The margins are less attractive for Tier-1 organisations when dealing with commodity solutions. We have certainly witnessed Tier-1 vendors discouraging commodity solutions in favour of their own proprietary-based solutions.

4. Most integrators see the HPC market continuing to grow, as the technology continues to mature, and new innovations drive performance ever higher. One possible caveat here however is the issue of Full Economic Costing. This is now playing an increasing part in the decision-making process as the factors that affect FEC are becoming more visible, and sites are now reaching their capacity to provide the space, mains power and air conditioning required to run a super computer cluster. The days of major injections of capital funding though Universities and SRIF may be drawing to a close, and with it a much needed funding stream for many of the UK integrators.

5. There is some confusion over the role that Tier-1 vendors actually play in the UK market, a point made by some of the integrators. 80-90% of all HPC Cluster developments in the UK have been carried out by systems integrators, even those 'sold' by a Tier-1 vendor. For example, IBM and HP Cluster solutions have been integrated and installed by OCF for a number of years. Dell works with Scali to position their PC offering into a "Dell Cluster". Streamline recently built most of the large HPC Clusters sold in the UK by SUN, while Compusys are the integration and support specialists for the Cray XD1 supercomputer.

6. This issue of technical competence is seen by all the integrators as the key differentiator, and key to the future of their organisations. As cluster sizes grow but commodity hardware becomes cheaper and cheaper, the "value" in the market will be the ability to make sure that clusters work efficiently for a broad range of applications and function in a

more scalable manner and operate seamlessly across sub-systems. System management and monitoring capabilities will become more important, and will help the ability of the integrator and the end-user to support such a system through its lifetime. The ability to provide a high level of support, not only on a system level, but also on an application level, will help differentiate the 'box-shifters' from the serious HPC oriented companies.

7. In the commodity-based solutions market, integrators provide cost-effective, technology compelling solutions rivalling those of alternative Tier-1 organisations.
8. One potential engagement model would be to form an Integrator Technology Partnership with those integrators who are deemed appropriate to the task in hand. These organisations typically do not have legacy turf wars that have made previous attempts to structure multi-Tier-1 vendor consortiums around high-end HPC solutions extremely difficult (e.g. in the UK's national HPC procurements - HPC'97 and HPCx), and are far more able to accept such a solution. This would have the obvious advantage of pooling highly competent, but thinly spread, technical expertise.
9. All integrators have existing relationships with Tier-1 vendors, although the nature of these interactions varies considerably. It is clear that deployment of very large systems requires the financial strength of a Tier-1 vendor to execute larger contracts and arguably needs the specialist expertise of key integrators to provide the knowledge and skills to build and support the systems. Productive relationships with Tier-1 vendors are seen as critical success factors for future growth of many of the integrators, and does provide an engagement model for other organisations, assuming that the Tier-1 vendor of choice does not inflict an ineffective integrator (or vice versa).
10. We do not feel that traditional SI houses have a role to play in this arena – they are invisible within the academic space, and would realistically have a steep learning curve to climb to be in a position to deal with the technology issues central to HPC provision. We would suggest that their value is clearly "only perceived at a corporate rather than operational level".

6. Bibliography

[1] http://www.sandia.gov/ASCI/, http://www.llnl.gov/asci/

[2] N. Boden, D. Cohen, R. E. Felderman, A. E. Kulawik, C. L. Seitz, J. N. Seizovic, and W. Su, "Myrinet: A Gigabit-Per-Second Local-Area Network," in *IEEE Micro*, **15**, 1995, pp. 29-36.

[3] F. Petrini, W.-C. Feng, A. Hoisie, S. Coll, and E. Frachtenberg, "The Quadrics Network: High-Performance Clustering Technology," in *IEEE Micro*, **22**, 2002, pp. 46-57.

[4] http://www.fcw.com/fcw/articles/2003/0825/tec-lightning-08-25-03.asp

[5] http://www.brightsurf.com/news/aug_03/PNNL_news_082703.php

[6] OSCAR: http://www.osl.iu.edu/publications/pubs/2003/oscar:ols03.pdf, http://oscar.openclustergroup.org/tiki-index.php

[7] ROCKS: http://www.rocksclusters.org/Rocks/

[8] http://www.bsc.org.es/

[9] http://www.llnl.gov/linux/thunder/

[10] http://news.com.com/2100-1008_3-5208220.html?tag=nefd.lede

7. APPENDIX 1: Integrator Questionnaire

Initial Integrator Discussion Points around HPC provision

To best inform the data gathering exercise associated with this document, a set of preliminary questions were devised and discussed with each of the integrators, typically by phone during the 2nd and 3rd weeks in June. These questions are sketched out below, with the responses of section 3 driven off the following eight points:

1. Understanding the current install base (both in the UK and abroad) and trends in the cluster marketplace. Information on current install base (ideally over the last 4-5 years, providing a picture of changing trends), including where possible the site, architecture, size, procurement date etc. (naturally no financial are expected). An idea of the relative number of "small" (32-64) systems compared to larger (128+) Machines. What is the approximate split between HEI's and industrial installations?

2. Details and Company size/status etc: Company overview - background, status, size etc. In providing this information it would be useful to have a breakdown of the relevant parts of the organisation - sales staff, after-sales support team, technical support (software and hardware if possible) etc., approximate turnover (machines not finances), customers (numbers - names are not necessary although they might prove useful). The details of, say, three customer reference sites would be helpful.

3. Company Outreach and Presence: An extension to the first two points - what presence do you have overseas - in particular in Europe and the USA - and what, if any, is the size of the current install base outside the UK

4. Company areas of Expertise - From the technical perspective, what level of technical expertise do you feel you bring to the system integration market that makes you competitive and a long term prospect in that market place. Please mention any other tie-in's you provide which you feel are relevant.

5. Company perspective of the Marketplace - How do you feel the cluster market place is changing (especially with FEC coming into effect). Any information you can provide as to changes/investments you are making to adapt to this changing climate, e.g. requirement to increase skills in middleware, compiler, database, files systems etc; need to forge strong links with software companies / interconnect solution; impact of GRID/e-science developments?

6. Relationship to Tier-1 Organisations: Do you feel your company has the strength and depth to build high-performance, scalable systems which can support Tera & Petascale solutions in the not-too-distant future. Are you able to provide this independently, or do you require Tier-1 support in dealing with areas such as risk and liability. Do you feel that the UK integrator providers are competitive with those in the USA?

7. Relationship to Other System Integrators - Do you perceive any role for the more traditional SIs e.g. EDS, CSC in this marketplace. These are pretty

invisible to us, but you may have a different perspective over what appears to be a more expensive alternative - at least in the non-academic space?

8. The above pointers are clearly not exhaustive, so if you feel that any other information is important in our trying to understand your position in the marketplace, please feel free to mention it. Can you provide a viable cost-effective alternative to the established blue chip/Tier-1 companies when it comes to procuring mid-range / high-end systems?

8. APPENDIX 2: Company Contact Details

8.1 ClusterVision

http://www.clustervision.com

Address: ClusterVision Ltd
17 Essington House
Lytton Grove
London SW15 2ET

Tel: +44 870 080 1990

8.2 Compusys

http://www.compusys.co.uk

Address: 1A Bessemer Cresent
Rabans Lane Industrial Area
Aylesbury
Buckinghamshire, HP19 8TF

Tel: 0870 745 7575

8.3 OCF ltd

http://www.ocf.co.uk

Address: Rotunda Business Centre
Thorncliffe Road
Thorncliffe Park
Sheffield S35 2PG

Tel: +44(0) 1142 572200

8.4 Streamline

http://www.streamline-computing.co.uk

Address: The Innovation Centre
Warrwick Technology Park
Gallows Hill
Warwick CV32 6UW

Tel:+44 (0)1926 623130

8.5 *Cambridge On-Line*

http://www.cosl.co.uk

Address: Cambridge Online Systems Limited
163 Cambridge Science Park
Milton Road
Cambridge CB4 0GP

Tel: +44(0) 1223 422600

8.6 *Linux Networxs*

Address: Linux Networx GmbH
Europaallee 10
67657 Kaiserslautern
Germany

Tel: +49 631 3031809

8.7 *SCC Plc*

Address: James House
Warwick Road
Birmingham
B11 2LB

Tel: +44(0) 121 766 7000

8.8 *Western Scientific*

Address: Studio 1
Waterside Park
Third Avenue Centrum 1000
Burton on Trent
Staffs
England
DE14 2WQ

Tel: +44(0) 1283 569989

8.9 *Workstations UK*

Address: Unit 1a Brazils Yard
Plantation Road
Amersham HP6 6HJ
United Kingdon.

Tel: 01494 724 498

9. APPENDIX C: Daresbury Contact Details

Should further information be required regarding any of the issues raised in this document, please contact us using any of the methods below:

http://www.cse.clrc.ac.uk/disco/contact.shtml

Address: DisCo Group A22
Christine Kitchen
CCLRC Daresbury Laboratory
Keckwick Lane
Daresbury
Warrington
Cheshire
England. WA4 4AD

Tel: +44 (0) 1925 603756

http://www.cse.clrc.ac.uk/disco/forums.shtml

Council for the Central Laboratory of the Research Councils
Chilton, Didcot, Oxfordshire OX11 0QX, UK
Tel: +44 (0)1235 445000 Fax: +44 (0)1235 445808

CCLRC Rutherford Appleton Laboratory
Chilton, Didcot,
Oxfordshire OX11 0QX
UK
Tel: +44 (0)1235 445000
Fax: +44 (0)1235 445808

CCLRC Daresbury Laboratory
Keckwick Lane
Daresbury, Warrington
Cheshire WA4 4AD
UK
Tel: +44 (0)1925 603000
Fax: +44 (0)1925 603100

CCLRC Chilbolton Observatory
Drove Road
Chilbolton, Stockbridge
Hampshire SO20 6BJ
UK
Tel: +44 (0)1264 860391
Fax: +44 (0)1264 860142